We appreciate the Referees' valuable remarks and recommendations and carefully addressed them in the new version of the manuscript. Please find final response attached as supplement.

On behalf of all authors,

Elena Zakharova

Referee 1

General comment

The authors have done substantial revisions and improved the quality of presentation (both text and figures). However, I still have a lot of concerns and suggestions. As for the majority of my questions the authors just removed a lot of text and figures from the manuscript, I feel that we should have another detailed look into what is left.

I provide here the structure of the paper for an overview:

1. Introduction
2. Regional setup and data
2.1. Study Region
2.2. Data
    2.2.1. In situ data
    2.2.2. Altimetry data
3. Temporal variability of radar altimetry signal over frozen rivers
3.1. Backscatter variability
3.2. Waveform changes

4. Methods
4.1. Ice onset and break up algorithm
4.2. Ice thickness algorithm

5. Results
5.1. Ice phenology algorithm verification
5.2. Ice thickness retrievals
5.3. Ice thickness estimation for the entire studied river reach
5.4. Winter ice bridge roads operation forecast

6. Discussion
6.1. Factors affecting ice thickness retrievals from altimetry
6.2. Potential improvement of algorithms

7. Conclusions

In the **Introduction** I still miss a thorough description of the knowledge gap which authors fill with your study. It is clear that river ice needs monitoring, that the remote sensing is the great tool for that, and that authors saw and implemented a good potential of the altimetry backscatter. I would like to see more details on the altimetry principles as well as some review on the existing studies using altimetry for the fresh ice monitoring. From this authors can draw the knowledge gap and the objectives. Also, for example, authors mention some SAR studies on the ice thickness (there are also many on the ice phenology which I think you should mention too) but do not provide any drawbacks of them, i.e. why do we need to use the altimetry at all, if we have SAR?
(See also my comments for the first review).

Reply. The details on the altimetry principles are added and review of the existing studies dedicated to the altimetry application for freshwater ice is extended. Very few studies used radar altimetry for lake ice. We sited 2 studies applied altimetry for monitoring lake ice phenology (no other studies exists). Unfortunately, **ONLY ONE** study dedicated to the lake ice thickness exists (Beckers et al., 2017). This study used another approach based on height estimation from the altimetric waveform. This method is more difficult to implement for the **narrow** rivers. We cited this study in the Introduction and explained the approach used in the section 4.2. We also mentioned that their method can have a potential issue when applied to narrow rivers as the intermediate peak on the waveform exists, but additional studies are needed to well understand from which surface it comes (floodplain surface, ice/air interface ?) . Several more studies used SAR instruments for river ice phenology dates and ice thickness were added. (Sobiech et al., 2013. Sun and Trevor 2018, Zhang et al., 2019).

Regional setup and data: I would split the section into two sections, to avoid numbering of the third order (2.2.1 etc).

Reply: The section was split.

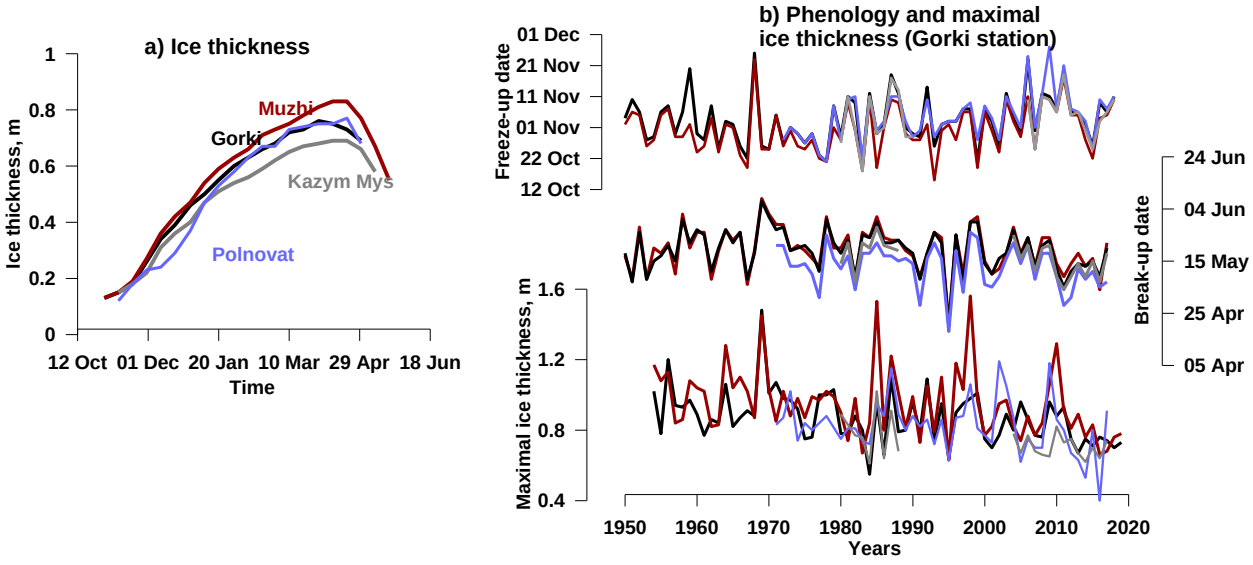
Figure 2b: I actually see some trend for the maximum ice thickness.

Reply: Thank you. The text was modified.

Consider adding similar graphs for four other stations in a supplementary figure.

We already tried to plot all stations, but decided that the look of figure with 4 stations is not nice (see figure below). Moreover, as one can see, the figure with 4 stations does not provide any additional information comparing to the general description of the Ob R. ice regime that is already given in the text. In the text we mentioned main differences between the stations and provided all important dates and values in order to avoid the presenting of the messy plot. Also, at the southern stations the observations began in 1980ies and have an important gap in 1990ies (see Table1).

Nevertheless, we updated the figure in the manuscript adding 2 more station with long observations without the gaps.



The whole section **Temporal variability of radar altimetry signal over frozen rivers** is a mix of own results from this study and some discussion of the previous studies. I understand that authors first investigate your data and then build an algorithm based on own findings, previous studies, and known facts. But I think it should be possible to find a way for rigorous presentation of own results and their discussion with respect to other studies. In general, there are a lot of speculations in this section which are given without any references or proofs.

Reply: The section was re-written. Many references were added and many general descriptions of low importance related to the ice formation and regime, that can be found in the cited literature were

deleted. Unfortunately, to our knowledge there are no studies dedicated to analysis of altimetric backscatter behavior over frozen rivers. Several studies, done by the authors investigated the variability of the altimetric backscatter over lake ice (Kouraev et al., 2007, 2015). We cited these works. Many statements has been drawn from authors' own research conducted during the period between the cited publications and present work. We hope that the new version of the section is now better structured. We also think that although several statements looks speculative, they could be seen as hypothesis or assumptions, which are proved during validation of proposed algorithms.

Here I also would repeat my suggestion to use a typical (or atypical) backscatter cycle over one year as an example, and to illustrate all ice events and corresponding backscatter changes within one year.
Reply: The plot was added
Authors added color in Figure 3 for the ice cover period, this is already helpful but still it is difficult to follow references to this figure in the text when authors describe the seasonal changes. This also includes Methods. Also, please add the cycle of TB on such graph.
Reply: The TB lines were added

Results

I could follow the results until the attempt to validate the 2D product. Do I understand correctly that:
1) authors retrieve the ice thickness at all VS using the relationship between backscatter and in situ ice thickness
2) then interpolate the ice thickness from the VS to the entire river
3) then extract again the ice thickness at some VS (why those?) from the interpolated product
4) and then compare it with the in situ ice thickness?
Sounds like a lot of data juggling here, especially considering interpolating and smoothing. I understand the intention to validate the 2D product but I am not sure that it is possible to achieve here. Why not to extract the ice thickness exactly at the location of the gauging stations?

Reply: Yes, that is right. The Hice extracted from the 2D product was used for case study, namely, for prediction of Salekhard ice road opening/closure dates. And in the paragraph preceding the case (4) we explain why we need to assess the quality of interpolated product exactly for selected parameters : Hice at December 1 (for ice roads) and Hice_max (for climate change monitoring). Of cause, we could stop the processing of satellite-derived data at the level of Hice time series at 48 VSs. However, the idea behind elaboration of the 2D product was a potential application of satellite retrievals for areas between VSs and for other ice roads existing in the study region. Many of these ice roads are not maintained by local authorities and used by local population at one's own risk. The objective was not the validation, but the evaluation of goodness for the specified task.
We modified the text to better present this idea. " The elaborated maps can be used for evaluation of ice thickness and ice phenology dates in areas between virtual stations. For instance, two useful parameters could be extracted from the 2D product: the maximum ice thickness and ice thickness observed on 1 December. From a practical standpoint, knowledge of the maximum river ice thickness is relevant for hydro-climate change monitoring, while the ice thickness determined on 1 December is crucial for local and regional socio-economic stakeholders, as this is the average date for the opening of the ice bridge road to the north of the study area at Salekhard. To assess the goodness of the 2D product for practical use, we compare the interannual dynamics of the mentioned parameters derived from 2D product and observed on gauging stations."
The other Referee suggested to move the sub-sections dedicated to algorithms' verification to the Methods. We did it. We also moved, for consistency, the part of the current sub-section dedicated to the elaboration of the 2D product to the Methods, as this part took some volume after adding details on window selection. So, we tried to combine and address all Referees' requests for restructuring of the manuscript.

Also, there are some discussion and speculations in this part of the results which should be moved to the Discussion section. In the Discussion section I still miss some discussions on the place of this study in the context of the other river ice studies, be that in situ observations or other remote sensing

techniques, other Arctic rivers, or maybe even some connection to the lake ice studies. The Discussion in the form as it is now would fit to a purely methodological paper but this one is a combination of methodology and scientific results. And that is of course then hard to fit into a common paper structure too. I think that after my questions and suggestions, authors would need to make some amendments to the Discussion and Conclusions parts, as well as to the Abstract. Therefore, I leave it for the next round.
Reply: The paragraph with discussion of the source of errors was moved to the Discussion. The discussion on altimetry drawbacks (required by other Referee) was added. A paragraph about potential combination of altimetry with other RS techniques for river ice studies was extended. We hope that we now clearly expressed our opinion that we see the place of altimetry as a one of the element within multi-instrument approaches of river ice monitoring. We can't say more, as the number of studies dedicated to the river ice is really low comparing to other cryospheric thematic. It is too early for the "Road maps" and it is out of scope of this study.
Unfortunately, we could not understand which amendments we should to make to the Conclusion and especially to the Abstract. In these sections we presented only the results of study, which did not change since last review.

Regarding the style, I noticed the following issues:

- choice and mixing of tenses. Past or present, be consistent.
- missing words
- order of the words in sentences
- wrong or missing prepositions
- typos
- spaces
- lower case instead of indices
- dates format is inconsistent
- writing out versus spelling out numbers
- in situ or in-situ, italic or not? Corrected

I can see that the professional proofreading was not accomplished and would like to see that done for the next round of the revisions.

Specific comments:

41: icy conditions – do you mean “ice conditions”?
Reply: corrected for "ice conditions"
42: “…for people who are required…” – seems to me redundant in the sentence
Reply: changed for ".. for people who perform..."
51: please add that clouds are limiting factors for the optical sensors, not for the river monitoring in general
Reply: corrected for "... are limiting factors for monitoring river ice at high latitudes using optical sensors."
58: please check the order of the words in the sentence
Reply: corrected for " Passive microwave and thermal satellite instruments have demonstrated capability for the retrieval of ice thickness for large lakes (Kang et al., 2014; Duguay et al., 2002, 2015; Gunn et al., 2015; Kheyrollah Pour et al., 2017)."
81: extends approximately… a preposition is missing?
Reply: corrected for " The lower reach of the Ob River extends for approximately 800 km..."
84: reference to the Figure 1 is odd here – move it to a more general description of the study area.
Reply: the reference to the Fig.1 was moved to one of the previous sentences.
Figure 1: when I proposed to add color to the Figure 1, I mainly meant the overview part from the first version. I think the black and white zoom-in figure looks better. Just add colors to the gauging station symbols, the main cities, and the new overview map to the previous version. Sorry for the confusion.
Reply: We modified the figure and added colors as it was asked
101: something is missing between “stations” and “water”
Reply: modified for "... all gauging stations **monitoring** water level."
113: order of words
Reply: modified for " According to in-situ observations at the gauging stations, ice formation begins between 23and 27 October. For the last 20 years, the earliest and latest records were 1 October and 18 November, respectively."
114: “installation of ice cover” sounds odd to me
Reply: modified for " …the full freezing can take up to 10 days."
123: ice onset and melt **date**
Reply: the word "date" was introduced.
145-147: you mention studies for the ice thickness retrievals but you use AMR for the ice phenology. Please clarify.
Reply: the sentence was modified: " Brightness temperature measurements acquired with other passive microwave radiometers, such as SSM/I and AMSR-E, have demonstrated good performance for the retrieval of **ice phenology dates** and ice thickness on large lakes **in Russia** and Canada (**Kouraev et al.,2007**, Kang et al., 2014)."
150: I think you can remove “Jason-2 and -3”, as you do not introduce Jason-3 at this point.
Reply: phrase was modified
154: difference (bias) – why do you need the word “bias” in parenthesis here? Again, is it the difference or the bias?
Reply: the word " bias" was removed
157: can you provide a more recent access date?
Reply: the link was verified 2021/06/30.
159: what is ICE1 algorithm? Any references? Please provide a short explanation how is the backscatter coefficient defined and retrieved?
Reply: the citations were provided and the description of how the radar altimeter backscatter is estimated is added.
163: it is not the Python code which overlaps Jason measurements?
Reply: modified for "... Python code allowing the overlapping along-track Jason measurements and Landsat images."
166: “…the stations names were extended” – “the names of the stations located on the secondary branch were extended…”
Reply: modified as suggested.
178: “installation of ice cover” – ice does not install, please use another word.

Reply: modified for "The freezing in river channels starts from the banks..."
180: what do you mean with the word “intercepts”?
Reply: modified for "traps"
193: please introduce Sig0, see also my comment to the line 159
Reply: The Sig0 is introduced in the line 159 (altimetry data description) and explained in the lines 165-170 presenting now the ICE 1 retracking algorithm.
193: Δt is the time period between two consecutive observations? Please mention.
Reply: explanation for Δt was added.
Figure 3: the vertical lines indicating the new year are not visible. Open water line in the legend is also barely visible, especially when printed.
Reply: the figure was modified, the vertical lines were highlighted. All figures are now in 400dpi. Some degradation of quality die to figures re-sizing by the Word can be expected.
You decided to show the backscatter time series for these 2 stations based on their location (north and south), correct? Please mention it in the caption.
Reply: correct, two virtual stations are located in north and in south of study region. Necessary details are added in the caption.
How did you decide where to start and finish the red line, i.e. the ice cover period? Is it based on the in situ observations? Or these are the results of your algorithm implementation? Please mention.
Reply: details were added :" **Data for period of ice cover retrieved from altimetric measurements are shown as thick dark red line."**
226: I do not see what do you call an intermediate peak in Figure 4. Please indicate it on the figure.
Reply: arrows are added.
248: “Freezing on the floodplain and banks” – do you mean freezing of the land surface, i.e. soil or sand? Please reformulate.
Reply: the sentence was modified " Freezing of small oxbow lakes on the floodplain and soils and bogs on the banks usually occurs earlier than in the big channels of the Ob river."
252: I think it is the other way around – the backscatter increase marks the ice decay.
Reply: we slightly modified the sentence: as this is the ice decay which is responsible for the backscatter increase."The beginning of the ice cover decay (thermal melting) leads to the spring backscatter increase. "
268: “we used a relative backscatter decrease…” – add for what exactly
Reply: modified for "... we used a relative backscatter decrease for ice thickness estimation instead of the absolute backscatter values..."
278-280: and then what dates did you use for the ice thickness estimation period? The best of automated vs manual?
Reply: the phrase was modified for "Starting from the first date of freezing (defined using manual algorithm),..."
284: power function, not equation. In the equation, you have Hice_alti in the left part – should it be H in situ? Please use subscript instead of lowercase.
Reply: sorry, here is in situ ice thickness. Thank you pointing this out.. The equation was modified. As the same equation was used for estimation of satellite-derived ice thickness, the general form of Hice was used.
You decided to exclude the scatterplot but I think it is important to show it. Please show all gauging stations (and corresponding virtual stations) with different symbols or colors as well as the power fits. Alternatively, one set of all stations together, with one fit. We can decide later whether to include it or not.
Reply: additional figure showing all VS-GS pairs from the training set with their power fits was added. Please, do not exclude/reshape this figure from the manuscript during the next revision. We deleted the short version of the figure (showed only 2 VS-GS pairs) from the first version of the manuscript to reduce the total length of the text (as it was recommended). But from the new referee comments we see now, that the length is no more a drawback for the article and we can provide all necessary information regardless total manuscript volume. We considered as well that the plotting of points from all VS-GS pairs together is not a good solution. Moreover, the coefficients of the specific fitting curves are presented in the table 2, so it is not necessary to plot/compare these curves on the same subplot.

We consider as well that the fitting of the points by a unique (for all VS) curve (while we hopefully have a chance to obtain a specific coefficients) is not proven. The unique curve will, obviously, lead to deterioration of the results of validation. Probably the "unique curve approach" can be used for extrapolation of the method for other river reaches or for other rivers of the region (north of the Western Siberia) in absence (or limited number) of in situ observation. However, this question takes more investigations.

287: “gauging VSs”: “-“is missing?

Reply: Thank you for the correction.

289: why do you use “mean” and “average” interchangeably?

Reply: the sentence was modified for "mean" in both cases.

291: “thickness” is missing

Reply: Thank you for the correction.

321: please explain in the text of the paper why automated algorithm was better for the melt end date retrieval compared to the manual approach.

Reply: the phrase was modified for "**The method was designed for detection of the melt start.** Manual estimation of dates associated with break-up allows for better control of the complex variability of the backscatter during the spring than automated estimation. It is likely that the automatic approach passes over complex cases and detects in these cases the melt end or even provides unrealistic early/late melt dates estimates."

Figure 6: if you show the histogram, should the y-axis be called pdf? I am not sure, please check and correct if needed. “Freeze up” lost “e” in the title, and also the legend. You can explain M and A in the legend instead of the caption. In general, the quality of this figure is not good enough, please work on it (font size, visibility, etc).

Reply: The figure was re-plotted with several corrections. The quality of the figure is 400 dpi. We called the y-axis "Norm pdf" as it presents the normalised on the total number of observations values of pdf. We mentioned this in the figure caption as well. Other reviewer asked us to plot all plots in one line. We did not manage to plot the long words in the legend within the plot limits in readable font size, so in the legend; we kept only the letters.

329: If I understood correctly, for the ice phenology you do not train anything - the 10 VS are just for the validation of the retrieved dates.

Reply: Yes this is true.

336-337 and Figure 7: 1) as you show and describe the results of the manual retrieval, why do you refer to the algorithm here? 2) Please also add why you do not provide the graph for the melt end (poor results?). 3) Do I understand correctly that you show that the manual approach works, in general, better, and for the ice thickness retrievals you use the manually retrieved dates? 4) Please make it clear in the text.

Reply. 1) The reference to algorithm in the figure caption was removed. 2) The new phrase related to the question 321 " **The method was designed for detection of the melt start**" now answers to the current question about the absence of results for melt end detection. We did not aim the development of "open water" or melt end date algorithm, as considered that ice weakening associated with thermal degradation (melt start) is more important information for people safety. However, it is interesting task that we will explore in our future work. 3) Yes, we demonstrated that altimetry-based approach works and our manual retrievals are good. Unfortunately, our automated retrievals are not as good as manual ones and further development (coding) is necessary for their amelioration (when the funding will be available). We consider that now we better explained in the text why. 4) We added the following phrase " As the manual algorithm of ice dates detection demonstrated better accuracy than automated algorithm, it was selected for further analysis of results and for use with the ice thickness retrieval algorithm."

339-344: I think you do not need to explain the legend in the caption, it should be clear enough. The information on what stations are in (4 gauging, 20 virtual) is important. The max-min red lines on b) are not visible.

Reply: Number of stations used is provided in the figure caption. The line width was increased. The legend was removed.

345-350: I think, we, in general, see the decrease of the accuracy from the north to the south? Interesting! But it may also be a result of the proximity of the VS to the gauging station: Pitlar station is the closest to its VS. Include the distance to the gauging station for each of the VS in the table. Please include that point into the discussion.

Reply: The distance to gauging station is now provided in the Table 2. We do really observe some tendency of degradation of validation scores with the distance. However, the correlation of R and RMSE with the distance is low -0.45 and 0.39 respectively and the p-values for these correlations are very high and equal correspondingly to 0.20 and 0.26. Sorry, we do not know how to correctly present these results and even not sure that they have any importance. The main reasons for lower accuracy of Hice retrievals on southern stations, as we think, are given in the manuscript in the current section (we suppose this is the effect of the polynia for VS12) and in the Discussion in the section 6.1 (ice hummocking/ridging in area of VS109).

350-352: "for many years and many locations…" – I do not really see that by looking at the figure, except for the station VS12. Please provide then some kind of a quantification of your statement.

Reply: We meant relative errors. The phrase was modified.

361: I do not understand why do you need to refer to the south or north here. You simply do it for all 5 gauging stations, correct?

Reply: the reference on "south/north" was deleted

Table 2. please explain better the content of the table in the caption. I understand now what do you show there but it took me 2 rounds of revisions and very careful repetitive reading. For example, it is really confusing that you provide coefficients for the power fit and R and RMSE for the validation regression – both next to each other.

Reply: for clarity, the Table 2 was split into 2 tables. Table 2 now presents only the results of validation and the Table 3 presents the scores from cross-validation experiment. The details were also added to the header of the Table 3. In many scientific publications of AGU and Elsevier publishers the scores and different parameters of equations are often provided altogether in one table for compaction or facilitating the results overview. Following this practice, we decided to put the fitting coefficients and the validation scores together in one table.

Figure 8: please be consistent with the used terminology: backscatter measurements or altimetric measurements.

Reply : backscatter measurements was changed for altimetric measurements

I think it would be nice to arrange the figure vertically from north to south.

Reply: the figure was re-arranged vertically with VS pairs from north to south

386-391: please include a short explanation how did you choose the window size

Reply: The details were added: "The size of applied window allowed for preserving the magnitudes and spatial heterogeneity of ice thickness in spatial domain, as well as for reducing the residual noise in temporal domain, which is left after smoothing of backscatter time series with Loess filter". We did not enter into technical details how we adapt our window as it can be found in different corresponding manuals and explanation could result in one more paragraph of important size. Instead, we provided the criteria, which were important for selection of the window.

Thank you for the details. I think that some more details would not hurt here. What are the corresponding manuals, can you cite them? For example, why 40 km window size preserve the spatial heterogeneity of ice thickness? What is changing on the scale of 40 km? Why is it important at all to smooth temporally? Is there a possibility of oversmoothing and affecting results too much by it?

Reply: The manuals and an article providing theoretical background we based on, when selected the smoothing window are below .

- de Smith M J (2015) STATSREF: Statistical Analysis Handbook -A comprehensive handbook of statistical concepts, techniques and software tools . The Winchelsea Press, Winchelsea, UK
- Lotov et al., INTERACTIVE DECISION MAPS, Approximation and Visualization of Pareto Frontier. Applied Optimization Series, Volume 89. SPRINGER SCIENCE+BUSINESS MEDIA, LLC. Ed.: Pardalos P. ISBN 978-1-4613-4690-6, 2004.
- Kamenev G., A Multicriteria Method for Identification and Forecasting. Mathematical Models and Computer Simulations, 2018, Vol. 10, No. 2, pp. 154–163.

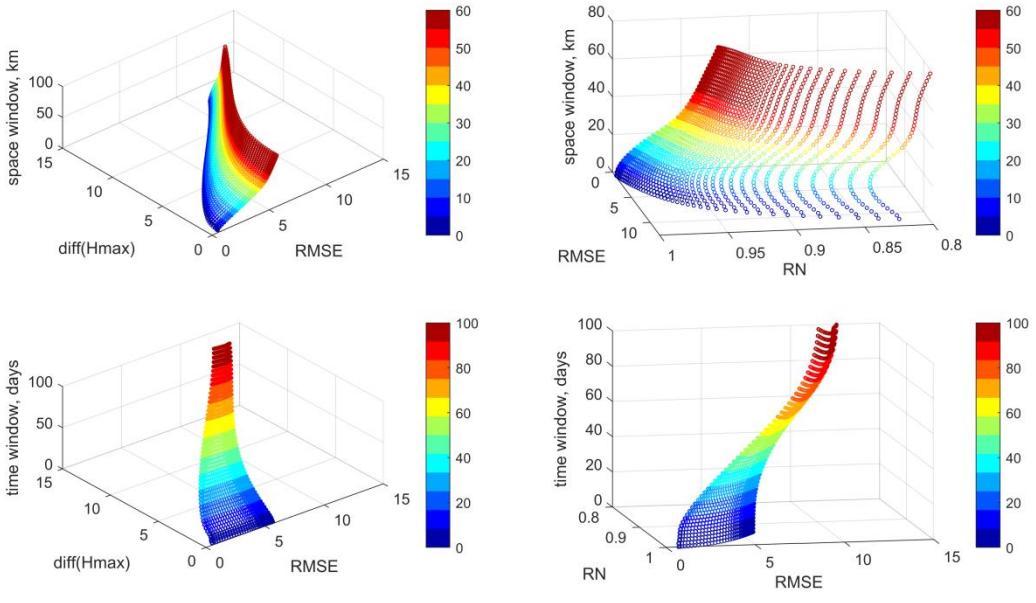
The smoothing was necessary because of irregular temporal satellite sampling (the satellite over-passes different VSs at different dates), gaps in satellite data product and Hice_alti time series, eventual outliers in Hice retrievals etc. This is a general procedure applied for many spatial altimetric high level (L3 and L4) data products (see for example https://sextant.ifremer.fr/record/bd5a176b-350e-4d5f-8683-da457637bdcb/). When selecting the window, we compared the Hice extracted from smoothed product at location of VSs (20 VSs for the Big Ob) with the "unsmoothed" Hice retrieved from satellite. Three criteria were employed for window selection: correlation coefficient, RMSE and difference between maximal Hice retrieved for each year (see Figures A and B bellow).

In space domain (upper panel of Figure A) the correlation coefficient (RN) and RMSE statistics deteriorate when applying the windows in the range 15-40 km. The degradation slows down at windows higher than 40 km. We did not analyse the reason, probably, it is related to the fact that the maximal distance between the VS in the study region is 42 km. The difference between maximuml Hice increases proportionally to the size of spatial window throughout whole tested range of windows (5-60 km) and served for control of reducing of Hice seasonal magnitude due to the smoothing.

In temporal domain (low panel of the Figure A), the window size has highest effect on RMSE statistic. In windows lower than 15 days (for 0 km spatial window) and 40 days (for 60 km spatial window) the changes in RMSE are low (Fig. A b, low panel). For 40 km spatial window, the selection of 30 days window (== smoothing on monthly scale) looks adequate. The decrease in correlation coefficient and in Hmax difference for 30 days window is low (0.99-0.94 and 0.02-0.05 m for Rcor and RMSE respectively).

Figure B represents 3D view of variation of the statistics regarding the window size. In theory (Lotov 2004, Kamenev, 2018), the optimal solutions (can be multiple) lies within an cross-over of surfaces. For this first version of the Hice spatial product we did not solve the problem analytically as it is described in these publications, as at the present time in the current manuscript, the spatial product served only for demonstration of potential use of satellite Hice and phenology dates retrievals, for example, for ice road operation dates forecast. We hope to do this in the future, when corresponding funding for product amelioration and development will be available.

Of cause, there was a risk of oversmoothing. However, we hope that we avoided this situation as the selected window allowed for keeping the average Hice_max difference and RMSE within uncertainties estimated during validation of the product (4 and 3 cm respectively). The correlation between smoothed and "unsmoothed" Hice was 0.99.



a b

Figure A. Effect of spatial and temporal window size on correlation coefficient, RMSE, and difference in Hice_max between Hice time series extracted from smoothed product and Hice time series retrieved at 20 VSs along the main channel of the Ob (used for gridding and smoothing).

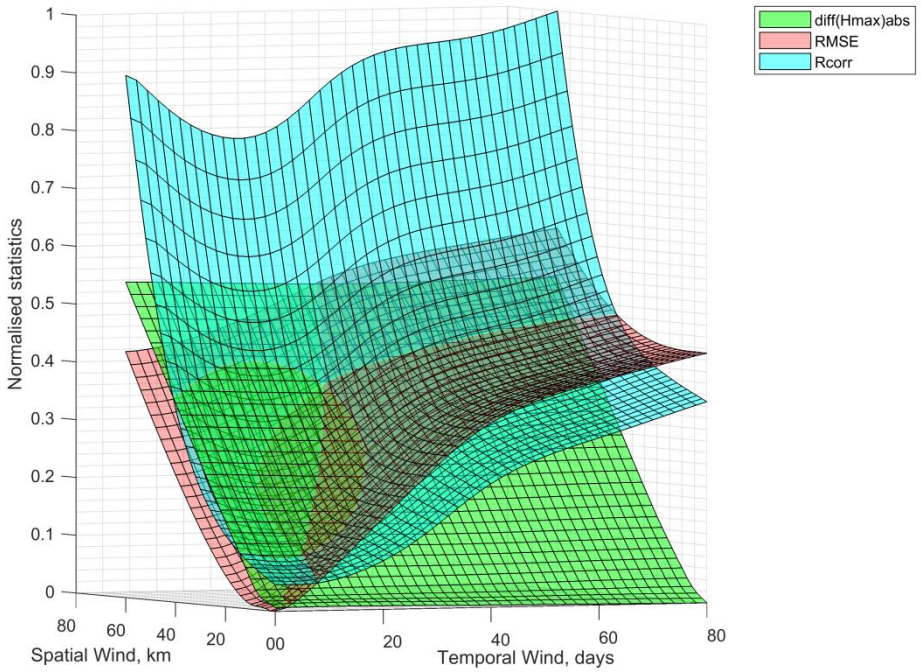


Figure B. Correlation coefficient, RMSE and difference between maximal Hice Normalised on their magnitude(max-min) values. Normalisation allowed for better graphical representation.

394: starting form this point I am again lost. Which in situ observations are you referring to when speaking about interstation areas?

Reply: We meant that there are no any additional in situ observation (published by other researchers or done by us) in areas between gauging stations. The phrase was modified " In the absence of validation data for reaches located far from gauging stations,..."

399-400: how do you derive this information from the Figure 9? What VS are there?

Reply: The color of the figure (varying from blue to yellow as shows the colorbar) corresponds to the ice thickness. The maximum ice thickness is observed in the end of ice season. The yellow color on the plots (highest values) degrades from 2009 to 2012 to the greenish color and becomes again more yellow starting from 2013 . We think that this is well seen on the figure. We hope that modification introduced makes the phrase more clear. "The interannual variability in maximum ice thickness retrieved from altimetric measurements indicates a clear decrease from 2008 to 2012 for 90% of the area of the studied river reach (Figure 9)".

442: please give some reference to the data

Reply: The reference was added." The information on dates of operation of the ice road was kindly provided by the State Traffic Service of Yamalo-Nenetzky Autonomous District (Russia)."

445: 1) what means 4 days ahead – simply the difference between predicted and observed dates? Or that one can predict the date 4 days in advance? 2) What would that mean?

Reply: . We meant 4 days in advance. The prediction is good four days in advance. Similar to the weather forecast, the longer the prediction interval, the worse the forecast.

2) It is difficult to say what it does mean exactly. The "predictor area" is located in 65-75 km upstream. Probably, the ice formation starts at both reaches at the same time and the Traffic Service considers that the ferry exploitation for next 4 days is still safe. It could be also that in the southern reach the ice appears earlier than for Salekhard city reach due to difference in morphology (island, sand bank, shallower water etc). It could be other reasons. Unfortunately, the area of the Salekhard ice road is not covered by Jason satellites and we can't say what happens between these two reaches without an additional study with use of other satellites (other altimeters, SARs or optical missions). An investigation

of capacity of other altimeters takes an additional efforts as the missions covering polar regions (> 66°N) have 27-35 days period. Such a period is not adapted for the operational monitoring. However, during our recent studies dedicated to the Ob R. discharge retrievals from the altimetric measurements (Zakharova et al., 2020), a multi-satellite method of water level time series construction was developed. Probably, similar approach could be applied for construction of multi-satellite backscatter time series.

447: how do you come up with 4 days of accuracy?
Reply: We added explication. " The dates when the Hice reaches 30 cm in location of four northermost VSs were extracted from the spatio-temporal smoothed product. These dates were compared with the dated provided by the State Traffic Service. The mean difference between these dates was four days. " Corrections were introduced.

455: mention that the dates are for Salekhard ice road and also add more details in the caption.
Reply: Details were added.

457-466: I also have troubles following this. You use the relationship between altimetric melt onset and real ice road closure dates to correct the altimetric date. But then you cannot compare it with the real closure date again, because the two datasets are not independent anymore, can you?

Reply: The main information about predictive capacity of the approach is contained in the Fig 12a. From the Fig12 a we see that there are both systematic and random differences between observed dates and AMO2. Moreover, the AMO2 parameter is a predictor, not the forecast. We supposed that we can correct the "predictor dates" on systematic (modelled or known) difference (what we did) and, then, we can evaluate only the residual difference left due to random or unknown errors. The latter gave us 3 days of RMSE. We modified the phrase for: " The residual difference of the forecasted and observed dates evaluated as RMSE is 3 days". Unfortunately, the length of available data was not enough to do full calibration/validation procedure for the forecast method (e.g. develop correction equation from one period and validate it over other period). It can be done in the future with longer time series. Moreover, the aim of this section was the demonstration of capacity and potential application of our results. We do not pretend to develop a robust operational system for the road opening/closure prediction. Development and comparison of different predictive algorithms is another very interesting task for future work.

Please, note as well, that in the forecast two parameters are important : the forecast precision and the forecast leading time (for AMO2 predictor the leading time is given in Fig12b) . We probably could elaborate a predictor that is independent on ground observations (for example, the certain value of ice thickness in specific area/period), however with the AMO2 predictor we could demonstrate one more approach how the product can be used for medium-term/ short-term forecast.

Referee2

Summary:

In this study, the authors have developed an algorithm to derive river ice phenology and thickness from satellite altimetry more specifically using the backscatter coefficient. The paper intends to provide complementary methods to estimate ice phenology and thickness as these parameters are important for ice road safety and climate studies. The results were compared to in situ data. Though the algorithm to predict the date of ice onset and melt is premature, the overall result has potential and adds value. Before the paper potentially can be accepted I recommend some changes. Please, see the comments below

General comments:

1) The methodology is only vaguely described, this needs to be improved. In section 4.1 a flow diagram could help to show the algorithm and then adding a more mathematical description would also help. Please, also provide more detail regarding the manual approached which seems to work better.

Reply: The flow diagram for ice phenology and more mathematical description were added.

Due to variety of factors affecting the temporal variability of the backscatter in winter and multi-peaky character of time series (notably in area of known development of winter polynya) we decided to run a manual retrievals to understand, how well we coded the algorithm. The criteria in both routines (manual and automates) were the same. We modified the text to clarify this question.

"...Their complex impact on the backscatter variability during freezing and melting makes it difficult to address all variations in an automated manner. Because of this, we decided to retrieve the ice phenology dates manually using visual analysis of backscatter (and TB if necessary) time series for each VS and to compare the performance of the automatic freeze/melt detection routine with its manual implementation. Both, manual and automated routine used the same criteria. "

In 4.2 It is not clear how it was decided that expression (1) was the best choice. If possible please add a figure that demonstrates how the relationship between the accumulated backscatter and the in situ ice thickness was established and what it looks like.

Reply: The scatterplot of $\Sigma(\Delta Sig0/\Delta t)$ vs Hice insitu is added. Details about fitting procedure were added. " Among tested fitting functions (linear, polynomial and power), the power equation (1) produced the best fit between the cumulative backscatter difference and in-situ ice thickness measurements. The selection of the function was based on maximisation of correlation coefficient between $\Sigma(\Delta Sig0/\Delta t)$ and in situ Hice and minimisations of root mean square error (RMSE) calculated between retrieved and observed Hice. "

2) The method sections should be expanded with a section describing how the results are validated and all the additional analysis performed, which currently is described in the result section.

Reply: We moved the 2 sub-sections dedicated to verification of retrievals from the Results to the Methods as recommended.

3) Currently, the result section is a mixture of results, methods, and an interpretation of the results. Though, a matter of style, this in my opinion makes the paper difficult to read and understand. In this case, it is more difficult to separate the actual result obtained from data, methods, and statistics from the author's interpretation. I, therefore, recommend rearranging this section, so the result section only objectively presents the results.

Reply: We hope that we guessed correctly the difference between results obtained from data and statistics from our interpretation. The 2 verification sub-sections were moved to the Method section. One paragraph with discussion of sources of uncertainties was moved to the Discussion. We were asked by other Referee to provide more methodology about elaboration of 2D product. Now, the volume of corresponding text allowed us to create a full subsection in the Methods dedicated to this part : "**5.3 Elaboration of 2D spatio-temporal ice thickness product".** We hope that the new structure will facilitate the reading.

4) Section 3.2, I do not understand why this section is in the data section, since you do not use the waveform info in the algorithm, or am I missing something. I would not expect that you will find two peaks (related to the ice/air and ice/water surfaces) in a Jason waveform due to the bin distance of 46 cm and an expected ice thickness of approximately 1 m, at least not often. The two peaks are not clearly seen in figure 4. Maybe this section should be moved to a discussion section regarding improving the algorithm.

Reply: The sections 3.1 and 3.2 were re-written as requested by other Referee and present now the background for the methods. The section name (now sec. 4) was changed. The waveform figure was kept here to demonstrate that the use of backscatter has a physical sense. Namely, the statement given in Beckers et al., that the main peak is from ice bottom, is important and the figure illustrates the decrease in amplitude of this peak in winter. This provides the proof that the backscatter decrease is related to the main peak decrease and, consequently, to the ice growth. Concerning the bin distance, we verified the waveform shape evolution over freshwater ice of different thickness on many lakes, and everywhere, we found this intermediate peak in the presence of snow-on-ice layer. Even for Hice < 40 cm the intermediate peak exists in Jason2 measurements ...in waveforms of other altimetric missions... and even in C-band... Actually, several ongoing studies investigate this phenomena and I hope upcoming publications will help to clarify this question.

5) In the discussion section I lack some comments on the strengths and weaknesses of the method and its use. e.g. One limitation is that you need in situ data to establish the thickness relationship, but if the relationship can be applied for another river it will add value.

Reply: Corresponding paragraph was added

6) The paper contains many language errors and should be proofread by a native English speaker or a proofreading service.

Reply: We paid the Elsevier publisher Service for English grammar correction for the 2nd manuscript version....It always worked well for our previous publications. .. Not this time... In new version we corrected many typos, missing words, rearranged the sentences. I hope we detected all catchy errors; Few occasional errors (articles and prepositions) could hopefully be corrected by specialists from journal technical team (what was the case with our last article published by EGU publisher).

Specific comments

Title: consider changing it to "River ice phenology and thickness from satellite altimetry. Potential for climate studies and ice bridge road operation

Reply: Thank you for the suggestion. The title was changed

L 61-62: How does this work, maybe add a sentence.

Reply: the phrase was added." ..., mainly via establishing a statistical relation between backscatter and in situ ice thickness"

L 63: "High resolution" please define.

Reply: values were provided

L 114 "installation of .." -> "formation of "

Reply: modified

L 122-123: Do not write "significant" unless you have performed a test

Reply: Yes we performed the test for all in situ stations. The results are not presented in full form as the manuscript is too big. We added several details in the sentence.

L 171. "In our previous studies (Kouraev et al., 2005; Zakharova et al., 2019, 2020), we noted that over" -> Previous studies (Kouraev et al., 2005; Zakharova et al., 2019, 2020), showed ..."

Reply: Thank you for the suggestion. The text was modified.

L 322 What do you mean by an "acceptable accuracy"?

Reply: we explained why we selected ±10 days as acceptable accuracy in the second sentence of the subsection. " Considering the 10-day repeat overpass of the satellite and the distance between the gauging stations and VS, we considered a 10-day time-step difference (e.g. ±10 days) as an acceptable accuracy for altimetry-derived ice phenology dates." The Jason repeat cycle is 10 days. The algorithm can be perfect with 0 days difference or deviate for some value multiples of 10. We think that 1 step difference from perfect match (==1 cycle==10 days) is acceptable accuracy when using Jason. It is evident, that we can't select criteria between 0 and 10 days, and 20 days bias is unpractical (==bad results) for river ice regime (from geophysical and from practical points of view). Anyway, we provided the percentage of 0 days cases and 10 days cases as for freeze up as for breakup. Thus, one can decide which criterion/accuracy is suitable for a specific task.

L 335 "A significant difference" rephrase if you have not performed a test

Reply: changed for "considerable"

L 350: "Error" is an imprecise formulation please clarify; RMSE, sd, ...

Reply: changed for "RMSE"

+ several languages errors not specified

Figures:

Figure 1: Please add lines to indicates the branches of the Ob River. Color code the VS used as training and test.

Reply: The Ob R. channels were highlighted, the training VS stations were colored in yellow.

Figure 3: Are the red lines based on in situ data? please explain in the figure text.

Reply: plotted Ice period was taken from the satellite retrievals done using manual routine. Details are added to the figure caption.

Figure 6: put the three figures on one line

Reply: Done