Comment on tc-2021-176
Anonymous Referee #2

The authors present a novel method of correcting slope-induced biases in radar satellite altimetry. A major challenge in assessing elevation and elevation changes of ice sheets. First, I have to applaud the authors for revisiting this challenge, which has been a considerable error source in radar altimetry since the early work by (Brenner et al., 1983). Novel strategies for dealing with this issue are of interest to the radar community, but I will let it be up to the editor to decide if the topic is within the scope of “The Cryosphere”.

I share many of the same general concerns as the first reviewer, and the following review will mainly supply additional comments. However, first I would like to highlight a couple of common issues also raised by the first reviewer. 1) Impact of radar penetration. Operating in the LRM area of Greenland, one would expect a considerable difference between the raw elevation measurements derived by leading-edge retracking at >10-20% and a validation dataset of real surface elevation observations. Hence, before venture into assessing the biases, this needs to be addressed. I cannot see any mentioning of surface penetration in the paper. 2) Performance of LEPTA relative to other approaches. The limited description of the implementation of both the LEPTA and reference methods leaves the readers with whether the observed differences are due to the method or implementation. 3) The figures are not of publication quality:

- The figure is hard to follow from the caption. Besides the equations, the main text only offers the figure to be “methods are briefly illustrated”. What is “briefly illustrated” I suggest adding an extensive description of the model flow, both in the main text and caption.
- This figure could be one image, I guess h_ice should be h_ice2, what is d_min. Text inside the figure should be avoided as much as possible. What is the geographical location of the plot?
- This is a fine figure, however, see the following comments about grid sizes.
- Same as above
- Could have been one figure, with a double y-axis. The curves show steps, which
suggest the tested delta r values to be too coarsely spaced.
- Work is needed to better resolve the signals in this illustration; the flat lines do not offer much information.

This leaves me with the following suggestions for improvements.

- When assessing the performance of retrackers an informative measure is “slope vs. elevation bias”. A better-performing retracker will have a flat response to an increase in the surface slope. Assessing this response would be beneficial for the paper.
- The differences between the different methods should be judged in terms of statistical significance.
- How does the gridding of 50x50 km tiles influence the results? Why is the point-based method the only method missing data in the trunk of Jakobshavn isbrae? Why is there no data for all the methods east of the line from 79fjord to Helheim of figure 3? The data coverage seems different in figure 4.
- Table 1, for the discussion it would be informative to also have the arcticDEM vs. ICESat statistics.
- Table 2, The surface penetration biases may relate to the retracking threshold chosen. How is the statistics changing between valid choices? (10%-90%)

Minor comments:

L2: What is “assessing snow/ice anomalies”

L14: Is the difference between the 1cm and 0cm bias between LEPTA and ESA significant and therefore needs to be differentiated from the performances of LEPTA. I think that it is the standard deviation that is the important measure.

L16: Reformulated: “we recommend the LEPTA method for obtaining... “

L17: What is complex topography? The work is done in the LRM area.
Concerning elevation change, you could add a reference to (Hurkmans, Bamber, and Griggs, 2012).

Suggestion to move this to the last part of the introduction.

Why not use the official releases of the downsampling ArcticDEM?

Is it ATL03 or ATL06 being used? From the link it seems to be ALT06, is there a bias of using the downsampling product?

This should be moved to the acknowledgment.

This sentence needs to be elaborated.

8x8 km seems small. When looking at the SARIn retracted data from ESA, relocation distances of up to 12 km from the nadir point can easily be found.

The bias can be evaluated in monthly intervals, but at some of 50x50 km tiles closer to the coast a seasonal difference in the bias is expected. How is this seen in your data?

Why use both the nearest and natural neighbor interpolations? You give some reasoning. However, would the two algorithms not converge in your case, and thereby there is no need for adding a user-defined threshold?

I guess the eastside is a result of topography? Could you give some insights into the differences on the east and west-side which will be the reasoning for this reported difference.

Having a setup at 50km tiles it would be rather easy to take the time-tagged ArcticDEM tiles into the analyses. This might be a large job to undertake this effort, but one or two tiles would be very informative for the analysis.

The observed change in bias is an important observation, please elaborate on this.
L244: Please clarify the statement: “relative sensitive”. Relative is a difficult word as it might be different for you and me.

L247: “although not directly visible” please improve the figure.

L284: Any insights into why ESA outperforms the other methods?

L290: Please elaborate on this last statement.

Reference:
