Response to reviewers

"A facet based numerical model to retrieve ice sheet topography from Sentinel-3 altimetry"

Jérémie Aublanc, François Boy, Franck Borde, and Pierre Féménias

---

Note to editor and reviewers

First and foremost, we would like to warmly thank the reviewers for taking the time to examine the paper, and for their positive and valuable feedbacks. We did our best to consider all the comments and complete this work and the manuscript accordingly. Whilst also trying to keep a reasonnable size for the manuscript length, which is necessary for the clarity of the messages deliver to the readers.

The main changes made to the manuscript are summarised below:

- In order to demonstrate one of the points stated in section 5.2 (lines 512-518), and reply to the questions of Reviewers n°2 and n°3, a new supplementary section S8 was written. This section addresses the impact of a vertical bias in the input DEM on the surface topography retrieved with AMPLI. This new supplementary section is available at the end of this document, pages 14-15.
- The "areas of improvements" section was substantially rewritten to address comments of Reviewers n°2 and n°3, and also to deliver clearer messages to the readers. In particular, the effect of surface roughness and the perspectives offered by neural networks are discussed. The rewritten sub-section 5.3 "areas of improvements" is also available at the end of this document, page 16.
- Most of the figures in the main text were redesigned for improving legibility, as observed by all reviewers. It must be noted that in the first manuscript version the image quality was unfortunately poor in the PDF, which did not make the image easy to read. Apologies for that. The updated figures are available at the end of the document, pages 17-19.
- The only error found after submission in the data analysis is related to the "*relocation distance*" quality control. In fact, for AMPLI we realised that the data discarded with this check are always associated to measurements taking "*NaN*" elevation and coordinates values (the test was not correctly applied). Thus, the data population involved was found to be the same as for the "*relocation failures*" test. To conclude, this quality control is just not relevant and removed from Table S3 (in the supplementary materials S3). This modification has no impact on the results presented sections 4 and 5 (for both AMPLI and ESA L2 data).

Response to Reviewer 1

Please find the author's responses in blue below the reviewer's comments.

General comments

The paper presents a new methodology (AMPLI) that uses the REMA DSM to simulate Sentinel-3 waveforms to improve altimeter retracking over the Antarctic Ice sheet. Results are compared against existing ESA Level-2 products, and bias calculations are performed against ICESat-2 to evaluate the performance of AMPLI over the Antarctic Ice Sheet.

The manuscript is well written, well-structured and clear, with the key references cited throughout. The methodology is well described, and the quality of the supplement is high and adds detail where needed. The results section supports the interpretations and conclusions in the manuscript. The figures are of high quality and support the manuscript text. I particularly liked the areas of improvement section, which tries to ascertain the impact of snow volume scattering.

Specific comments

I think when referencing Brown, 1977, you should add a few sentences talking a bit more about Brown's 5 assumptions in the context of your simulations, as he mentions that some of the assumptions may be violated over a land surface. Specifically, the distinction that your surface heights are spatially correlated; whereas, brown assumes a non-correlated (random) surface (i.e. The scattering surface may be considered to comprise a sufficiently large number of random independent scattering elements).

This is a fair and relevant point. To avoid overloading the text, I would not comment Brown's validity directly in the main text. You will find below a discussion for each of the five assumptions stated in his paper. If the reviewer and editor find the information relevant, this text could be added among the supplementary materials (I would suggest in the supplementary section S1).

It is also worth noting that this is not the first time that the Brown model is a starting point for the radar signal modelling over ice sheets. There are many references available in the literature. Among those cited in this paper: Lacroix et al. (2008), Larue et al. (2021), Helm et al. (2024).

The five assumptions that Brown (1977) stated in the introduction of his paper are discussed below. These assumptions must be met to ensure the model validity, as extracted from his paper: "*For land scatter, the situation is somewhat different, and some of the above assumptions may be violated*".

"1 - The scattering surface may be considered to comprise a sufficiently large number of random independent scattering element"

=> Over ice sheets the radar wave reflection is considered "diffuse". Therefore, as for ocean, it is possible to take the assumption that the altimetry waveform is built with a large number of independent scatterers (as sampled within the radar footprint).

"2 - The surface height statistics are assumed to be constant over the total area illuminated by the radar during construction of the mean return"

=> This assumption is violated over ice sheets simply because the surface is not flat. Thus, Brown's analytical formulation cannot be directly used to accurately reproduce waveforms acquired over ice sheets (except specific cases, like over the lake Vostok area).

Thus, the advantage of a facet-based modelling is to account for the effect of terrain variation in the radar waveform shape, using an external DEM (down to metre scale roughness with REMA). As mentioned in the text, line 103: "*In this work, we use the same FSIR formulation* [as Brown] *and discretized it at the DEM grid points. Hence, through this so-called "facet-based simulation", the effect of terrain roughness is integrated in the radar signal modelling, down to decametre scale variations*".

"3 - The scattering is a scalar process with no polarization effects and is frequency independent."

=> To our understanding, the antenna polarisation does not act on the radar waveform shape over ocean. However, polarisation effects have been reported for LRM altimetry over ice sheets in Remy et al. (2006) and Armitage et al. (2014). To the best of our knowledge, such effect is not yet investigated in SAR altimetry. Nonetheless, it is anticipated that the impact would be relatively minor in the waveform leading edge shape in SAR mode, because it is mainly generated with energy backscattered from surface, and not sub-surface (Aublanc et al., 2018).

"4 - The variation of the scattering process with angle of incidence (relative to the normal to the mean surface) is only dependent upon the backscattering cross section per unit scattering area, go, and the antenna pattern."

=> The assumption remains overall valid over the Antarctica and Greenland ice sheets, considering the relative surface homogeneity at the footprint scale, in terms of backscattering properties. Nevertheless, this assumption is likely violated in case of ice sheet melting (i.e. meltwater on the surface of the ice sheet), as it can occur in the Greenland ice sheet margins. Specific analyses should be planned in the future on this subject.

"5 - The total Doppler frequency spread (4 V_r/λ due to a radial velocity V, between the radar and any scattering element on the illuminated surface is small relative to the frequency spread of the envelope of the transmitted pulse (2/T, where Tis the width of the transmitted pulse)."

=> The nature of the surface has no impact on this assumption.

References

Armitage, T. W. K., Wingham, D. J., and Ridout, A. L.: Meteorological Origin of the Static Crossover Pattern Present in Low-Resolution-Mode CryoSat-2 Data Over Central Antarctica, IEEE Geoscience and Remote Sensing Letters, 11, 1295–1299, https://doi.org/10.1109/LGRS.2013.2292821, 2014.

Aublanc, J., Moreau, T., Thibaut, P., Boy, F., Rémy, F., and Picot, N.: Evaluation of SAR altimetry over the Antarctic ice sheet from CryoSat-2 acquisitions, Advances in Space Research, 62, 1307–1323, https://doi.org/10.1016/j.asr.2018.06.043, 2018.

Remy, F., Legresy, B., and Benveniste, J.: On the Azimuthally Anisotropy Effects of Polarization for Altimetric Measurements, IEEE Transactions on Geoscience and Remote Sensing, 44, 3289–3296, https://doi.org/10.1109/TGRS.2006.878444, 2006.

Technical corrections

The figures are of high quality and compliment the manuscript, the only (minor) comment that I would have is to increase the font size in the figures to improve legibility. Also in Figure 1, the cyan lines are difficult to see, perhaps try another color i.e magenta.

As also replied to other reviewers the font size was increased for most of the figures. Fig. 1 was also redesigned. The updated Fig. 1 is available page 17 of this document.

Response to Reviewer 2 (Veit Helm)

Please find the author's responses in blue below the reviewer's comments.

The paper presents a new method to determine the point of closest approach (POCA) for SAR altimetry by using a facet-based numerical model. The approach is called AMPLI and makes use of the Antarctic reference elevation model (REMA) in the simulation of the 20 Hz SAR waveforms. In the simulation step a cross track backscatter distribution (CTBD) map is generated which is used in a second step to determine the POCA.

The results are evaluated over the Antarctic ice sheet against ESA Level-2 products, by performing a cross point error analysis with ICESat-2 elevation measurements, which are taken as reference.

Secondly, the rates of change in surface elevation for the period (2019 to 2022) are calculated by comparing two gridded POCA-corrected elevation maps derived from three months of processed data for each year. SEC is evaluated against SEC derived from ICESat-2 and ESA Level2 corrected POCA, respectively.

The results of the new method are showing a great increase in the quality compared to the ESA Level_2 product for both, POCA correction and derived SEC.

The manuscript is well written, well-structured and clear, with most of the references cited. The quality of scientific content is high. The analysis applied are of high standard and well selected to show the improved performance of the new approach. Methods and supplements are described in detail in order to follow the developed method. The section on results supports the conclusions drawn and interpretation.

The figures are of high quality and support the text.

In my opinion the manuscript is worth to publish and fits well in the scope of The Cryosphere.

I do have minor correction/suggestions, which I would like to see included in the final accepted paper:

My main point is the following:

Although your approach shows great progress, I think in your comparison or evaluation one data set is missing. This is a relocation based on adapted Roemer/LEPTA for SAR. You can easily implement this. You can search the POCA in the SAR footprint of the REMA DEM centered around the 20 Hz records and derive the across track lock angle as given in eq. 5. I would like to see the results of adapted Roemer/LEPTA in comparison to AMPLI and ESA L2 here as well. Because AMPLI needs lots of computing resources, which can be argued, if gained improvement is significant – as the comparison to ESA Level 2 shows. However, if the adapted Roemer / LEPTA for SAR shows similar results as AMPLI, it would be easier to implement or propose it as a new standard POCA correction for SAR, as the time consuming simulation step is no longer necessary. Therefore, I think that this comparison is needed to be included in the manuscript (including comparison to ICESat2 cross overs as well as SEC).

I can identify three alternative relocation methods to which we could compare AMPLI (this short list includes the two you also reported above). This list is not exhaustive, and there are likely other ones.

- (1) **Roemer** et al. (2007) https://doi.org/10.1016/j.rse.2006.02.026
- (2) **LEPTA**: Li et al. (2022) https://doi.org/10.5194/tc-16-2225-2022
- (3) **MPI**: Huang et al. (2024) https://doi.org/10.1016/j.rse.2024.114020

The comparison to other alternative relocation methods was initially envisaged when we began writing this manuscript, but was finally discarded for several reasons, that I will try to explain below:

- This manuscript is dedicated to the presentation of the AMPLI software, which is relatively complex regarding the innovative facet-based modelling approach developed in UnFocused-SAR mode. The manuscript also includes a detailed evaluation made against ICESat-2 in (1) the elevation retrieved along the track (2) the Surface Elevation Change. The addition of a benchmark including other relocation methods requires a substantial amount of work, which was considered out of scope of this manuscript.
- Corollary as first bullet, the text is already quite long for a research article (~10 000 words, from abstract to conclusion). It also includes 8 supplementary sections. Even by comparing AMPLI to a single one of these methods, it implies: adding a description of the numerical implementation + additional results/figures in the sections 3.4 and 4.3 + additional text to discuss the differences observed… The manuscript will be more complicated to read, and therefore less accessible.
- To the best of my knowledge, there is no Sentinel-3 data set produced with the three relocation methods mentioned before. In their respective publications, Roemer and LEPTA were initially developed for LRM, not SAR altimetry. This means that the adaption and implementation of these methods to Sentinel-3 have to be done, and properly documented in the manuscript. This also implies the risk of "sub-optimal" implementation on our side, which can occur involuntarily, or which may be criticised by the authors of these methods.
- The manuscript still includes a comparison to the ESA Sentinel-3 Land Ice Thematic Products, which are the "official" data from the ground segment, openly available for the users.

It is important keeping conciseness as much as possible, for the sake of the article readability. Moreover, the benchmark/comparison of different relocation methods should be probably carried out by another group, to guarantee the fairness of results and interpretations.

Finally, it is highlty likely that AMPLI requires much more computing resources compared to the three relocation methods mentioned above. However, when run in a computational cluster environment, it does not prevent to perform mission reprocessing. The CPU time of AMPLI is therefore not a "blocking issue" for climate studies applications. For instance, we almost finished reprocessing the complete Sentinel-3A and Sentinel-3B missions with AMPLI (~7 years of data over the Antarctic and Greenland ice sheets), by subsampling REMA from 10 m to 20 m.

Figures:

- labelling of all figures (axis and inserts) is too small. Please enlarge all.
- Figure 3,5: I hardly can't read the red and green on top of the dark blue. It's too small and of low contrast.
- Figure 4: contrast in not sufficient. Please adapt the selected range or consider to show the normalized backscatter in dB.

Most of the figures have been redesigned. In Fig. 4, the energy range is now displayed between "0" and "0.75" (it was between "0" and "1" in the first manuscript version) to increase contrast (if the radargram would be displayed in dB scale, it would not be possible to see the low energy peaks). The updated Figures are available pages 17-19 of this document.

Introduction:

- A facet based model has been introduced before, but was so far only applied to LRM waveforms. Please add the reference: https://doi.org/10.5194/tc-18-3933-2024

I agree it is worth highlighting this paper, considering the common approach taken in the waveform modelling. It will be introduced in section 2.2, lines 108-109, in the text below. The reference is also added and mentioned section 5.3, when discussing the perspectives with neural networks.

« It is worth noticing that similar facet-based approaches were developed in the recent past for sea ice (Landy et al., 2019) and ice sheet (Helm et al., 2024) radar altimetry. »

Section 2 Modelling:

- Is there a need to consider beam steering in your approach, as this is the case for delay doppler processing. Usually when you form the DDMs at consecutive 20 Hz records, the individual doppler beams from different looks, which are collected in a stack before range migration, usually do not perfectly overlap. If no beam steering is applied the stacked waveform will be slightly blurred. Maybe it's not needed as you use the simuated waveform in a correlation and later only use the CTBD for the POCA estimate, but at least you should mention it here or in the supplements.

As mentioned in the text, lines 128-133, in this facet-based modelling the delay-Doppler Maps are directly simulated in amplitude (I^2+Q^2), at 20 Hz rate. Therefore, there is no need to perform the beam steering operation with this approach. The beam steering operation is performed on-ground to a burst of 64 complex radar pulses (I & Q) acquired by the satellite, indeed, to align the beam sampling at the 20 Hz locations defined in the ground segment processing.

This will be clarified in the supplementary materials S1, as follow:

*"**Note**: In this facet-based modelling, the DDMs are directly simulated in amplitude (I^2+Q^2), at 20 Hz rate. Therefore, there is no need to perform the beam-forming (i.e. along track FFT) and beam-steering operations, that are part of the UF-SAR processing as applied to a burst of real complex pulses."*

- What is the consequence to only use 45 instead of 180 looks? Can you exemplarily compare waveforms generated using 45 and 180 looks?

The principle of the multi-looking approach in UF-SAR is to collect single look waveforms that sample the same surface, but from different viewing angles. This reduces the speckle noise inherently present in the single look waveforms. As already mentioned in the main text, lines 128-133, the simulation is directly performed in amplitude. Therefore, there is no speckle noise present in the single look waveforms simulated. Consequently, there is no need to further increase the number of looks in our modelling.

However, it is useful keeping a 20 Hz rate, and a same range of viewing angles than the ground segment, for simulating the single look waveforms. Because the surface sampled within the altimeter range window changes from one single look to another, due to (1) the viewing angle variation (2) the on-board tracking command variation, which is updated at 20 Hz rate on-board. In the second case, large tracker range variations can occur, especially over the ice sheet margins (because the Sentinel-3 closed-loop tracker can have difficulty following the topography variations). This is why it was chosen to keep the same range of viewing angles as the ground segment for simulating the delay-Doppler stack.

In a nutshell, we just sub-sample the stack in azimuth by a factor of 4, from 80 Hz rate (in real data) to 20 Hz rate (in our simulation). Or stated another way, from 180 looks (in real data) to 45 looks (in our simulation).

- Please, give a number of the processing cost. How long does it take to simulate 10s of real data on a CPU. How long did it take you to process the full 3 months of data? This is important, if your new correction should become a new standard, then also processing costs need to be considered.

This is a relevant point, and I therefore made a computational time benchmark, with different spatial resolutions for the input DEM, as you also suggest below. This benchmark is available hereunder. This text will be added in the supplementary materials, section S1.

Computational time benchmark

The Central Processing Unit (CPU) time was evaluated on one of the CPU nodes available in the high-performance cluster of the Centre National d'Etudes Spatiales (CNES), in Toulouse (France). We used a single core of the node, with a processor clocked at ~3 GHz.

This benchmark has been performed with a Sentinel-3A track portion located in the East Antarctic Ice Sheet interior. The CPU time is not supposed to be significantly sensitive to the area sampled by the satellite. The CPU time is evaluated in the processing steps including "radar equation calculation" and "Delay-Doppler Map generation", steps (c) and (d) as listed above. In fact, these two processing steps are by far the most demanding ones in terms of computational time.

The computational time is evaluated with the AMPLI software version used to produce the results presented in this paper. As described section 2.1, with this version the facet-based simulation is performed by means of the 10 m resolution REMA DEM. The CPU time was also evaluated with REMA sub-sampled at 20 m and 40 m, as they could represent relevant alternatives to reduce the computational time, but at the expense of the modelling accuracy. It was nonetheless out of the scope of this study making a further sensitivity analysis of the AMPLI software performance, related to the spatial resolution of the input DEM.

The benchmark was performed using 10 seconds of Sentinel-3 acquisitions, which represents a segment length of about 66 km on-ground.

- 10 m REMA (version used in this study): ~**480** seconds of data processing (CPU time) => *real time factor of ~48*
- 20 m REMA: ~**70** seconds (CPU time) => *real time factor of ~7*
- 40 m REMA: ~**25** seconds (CPU time) => *real time factor of ~2.5*

- Include processing costs, for the suggested adapted Roemer/LEPTA for SAR.

Roemer/LEPTA discussed above

Section 3 Results:

- Can you please put the outlier detection from supplements S3 to results. I think it is important to see that AMPLI has also the potential to flag bad waveforms.

Most of the different controls can be made without the developed facet-based modelling. This is for example the case for the "*SNR*" and "*WF peak detection*" tests, that together flag 5.58% of the 20 Hz measurements, as reported in Supplementary S3. However, AMPLI has the ability to detect "ambiguities" in the first radar return. We think this is a clear added value, and this was already emphasised in the discussions, section 5.2.

For the sake of conciseness and readability, a choice was made to describe the different quality controls in the supplementary sections. The main text still includes a brief description of the major quality controls applied, and the total ratio of discarded measurements for the two data set.

- For a fair comparison against ICESat-2, I think you need to use exactly the same amount of data for both methods. Therefore, please consider only values were both ESA level2 and AMPLI are identified as valid.

The objective was to apply the same quality controls to both data set, as much as possible. More precisely, the quality controls that can be achieved without the facet-based modelling are applied to the ESA L2 data (for example, the "*SNR*" and "*WF peak detection*" tests, as already mentioned). However, some quality tests are intrinsically parts of their respective processing. This is especially the case of the "surface ambiguity" control performed with AMPLI (as also already mentioned), which cannot be made with the ESA L2 processing alone. In the case of the ESA L2 Processing, relocation errors can also occur, that are not present in the AMPLI data set. Thus, I think it is important to keep the statistics/results as they are currently provided in the manuscript, that consider (1) the new opportunities of data quality analysis bring by AMPLI (2) the different robustness of the two methods.

However, to cover this point, and check the impact of the quality controls, you will find below in the Table 1 the statistics obtained by keeping the same measurements for both AMPLI and ESA L2 processing (when both ESA L2 and AMPLI measurements are identified "valid"). The new numbers are reported in green colour. Overall, where surface slope is lower than 1°, the population of data remains relatively similar, thus the impact on the statistics is limited (negligible?). The only notable changes occur for surface slope > 1°, because the data population is lower, as it represents areas where the data quality decline. In particular, there is a decrease in the MAD, for both AMPLI and ESA L2. Nevertheless, for this slope range, AMPLI still provides a comparable reduction of the MAD compared to ESA L2 (factor of ~13 against factor of ~12.7, without and with same quality controls applied, respectively). The maps in Fig. 6 have also been remade, but the changes cannot be seen visually (which is not surprising considering the low impact in the statistics). So, they are not added here.

To briefly conclude, this new data selection does not change the results and outcomes. But these results can be added in another supplementary section, if required by the editor and reviewer.

Surface Slope		**< 0.1°**	**0.1° - 0.5°**	**0.5° - 1°**	**> 1°**
Population ($x10^3$ count)	ESA L2	393 $x10^3$ / 390 $x10^3$	568 $x10^3$ / 550 $x10^3$	65 $x10^3$ / 59 $x10^3$	11 $x10^3$ / 10 $x10^3$
	AMPLI	390 $x10^3$ / 390 $x10^3$	551 $x10^3$ / 550 $x10^3$	63 $x10^3$ / 59 $x10^3$	15 $x10^3$ / 10 $x10^3$
Median bias (m)	ESA L2	+0.04 / +0.04	+0.29 / +0.28	+1.86 / +1.84	+2.95 / +2.87
	AMPLI	+0.09 / +0.09	+0.14 / +0.14	+0.31 / +0.31	+0.42 / +0.38
Mean bias (m)	ESA L2	+0.05 / +0.05	+0.64 / +0.62	+3.95 / +3.93	+6.00 / +5.97
	AMPLI	+0.09 / +0.09	+0.15 / +0.15	+0.37 / +0.35	+0.50 / +0.45
MAD (m)	ESA L2	0.14 / 0.14	0.42 / 0.41	2.87 / 2.71	5.6 / 5.07
	AMPLI	0.10 / 0.10	0.16 / 0.16	0.30 / 0.29	0.43 / 0.40
STD (m)	ESA L2	0.15 / 0.15	0.93 / 0.90	5.21 / 5.09	8.87 / 8.05
	AMPLI	0.10 / 0.10	0.17 / 0.17	0.36 / 0.35	0.51 / 0.47

- Once again, please include a comparison to adapted Roemer/LEPTA for SAR.

Roemer/LEPTA discussed above

Section 5.1

- Please include in your discussion as well reference https://doi.org/10.5194/tc-18-3933-2024 as this paper addresses the penetration affect for Ku-Band and its worth to note that you observe with KU-SAR a penetration effect, which is smaller than for LRM but still present.

The reference from your work was added in the updated section 5.3 (available page 16 of this document), when discussing potential improvements that can be made with neural networks. Indeed, the Ku-band radar-wave signal penetrates into the snow medium, generating a "volume scattering signal". This effect and the impact on the estimated elevation with AMPLI, are (already) discussed in this section 5.1.

Section 5.2

- Line 517: I disagree with the statement, that the relocation method (Roemer, LEPTA) is affected by temporal elevation changes more than AMPLI. In both cases this will only be the case if the topography is changing within the footprint, as one takes the spatial closest position (x,y) coordinates of the DEM but use z from the radar range measurement itself.

Li et al. (2022) mentioned that the accuracy of the surface elevation estimated with LEPTA is impacted by temporal elevation changes. Cited from the paper, section 5: "*LEPTA shows sensitivity to a bias in the DEM […] Changes in the elevation over time will affect the applied correction as well as the location of the impact point*". This is documented by the Figures available in Appendix C of their article, showing the impact of a constant elevation bias applied to the input DEM (thus without topography change in the footprint). This is also confirmed by the study from Huang et al. (2024), for both LEPTA and MPI (Fig. 12 of the article).

Regarding AMPLI, because this point was also questioned by Reviewer #3, an analysis was added in supplementary S8, demonstrating that AMPLI is (almost) not affected by a constant vertical elevation bias in the input DEM. Thus, as a corollary, by temporal elevation changes when using a static DEM (if changes are homogeneous over the footprint, as mentioned section 5.2). At least for the bias values evaluated here, between ± 5 m, which corresponds to the same range of values tested by Huang et al. (2024). This new supplementary section S8 is available pages 14-15 of this document.

<u>The paragraph where these results are discussed will be updated (lines 512-518 in the first manuscript version). Updates are reported in red colour.</u>

The cross-correlation performed between the Sentinel-3 UF-SAR mode waveform and the numerical model (illustrated Fig. 3b) adjusts for potential vertical ~~offset~~ bias in the input DEM. In practice, this was mainly implemented to correct for temporal height variations between the altimetry measurement and the static DEM. To examine the veracity of this statement, positive and negative vertical biases were introduced in REMA (fixed offset), using the same range of values as Huang et al. (2024, section 4.5), up to ± 5 m. This analysis is available in supplementary material, section S8. As a result, even with the largest bias values (+5 m and -5 m), no performance degradation could be noticed in the surface topography retrieved with the AMPLI software. We consider this is a major advantage in comparison to LEPTA and MPI, as both methods are affected in case of vertical bias in the input DEM (Huang et al., 2024), and as a corollary, by temporal surface height variations. This aspect of the processing can be still enhanced, as the echo relocation is however not executed if the vertical bias on the DEM is found higher than ~14 m (as estimated via the cross-correlation delay between the simulated and the measured waveforms). This evolution is stated among the areas of improvements reported in the next section 5.3.

- I think it is also worth to discuss if and how the simulation process could be speed up – how important is a 10m facet resolution? Will 25m or 50m will result in similar simulated waveforms and CTBD's. If yes this could significantly speed up the processing. Maybe put some numbers in. How long does it take to simulate 10s of real data for 10m, 20m, 50m resolution?

I agree this is relevant communicating on the computational time required for this new facet-based approach, because the method indeed requires extensive computations. This "Computational time benchmark" is available at the page 7 of this document, and will be added in supplementary S1. As also mentioned before, the computational time does not prevent for a full mission reprocessing of the two Sentinel-3 satellites. It was also out of the scope of this study making a further sensitivity analysis of the AMPLI software performance, related to the spatial resolution of the input DEM. But this will be covered in other studies.

Response to Reviewer 3 (Melody Sandells)

Please find the author's responses in blue below the reviewer's comments.

This paper concerns a new method of retrieving ice sheet elevation by estimating the point of radar first return with a combination of modelling altimeter waveform and *a priori* information from the Reference Elevation Model of Antarctica (REMA). This is then applied to the entire Antarctic ice sheet and an improved performance was obtained in comparison to existing Sentinel-3 products, particularly in topographically challenging areas. The improvements demonstrated in Figure 8a are particularly impressive, and this new method offers a promising development that can enhance understanding of the Antarctic Ice Sheet behaviour in coastal areas, where changes are most prominent. The paper is generally well written and methodology explained.

General comments:

1. What is the impact of the REMA accuracy on the AMPLI accuracy? If a <+/- 1m random noise is added to REMA, how would Figure 7 differ (if at all) (1m error from https://tc.copernicus.org/articles/15/4421/2021/)

If we are referring to "accuracy", then this should be related to potential bias in the input DEM (rather than a "noise" in the model). And, indeed, the 1 m error mentioned in Dong et al. (2021) is related to vertical biases/offsets in the DEM.

This is anyway a good point, and the influence of the DEM accuracy was also assessed in Li et al. (2022) and Huang et al. (2024) for the LEPTA and MPI algorithms. Therefore, using same range of values as Huang et al. (2024), positive and negative biases were applied to the input DEM, up to ± 5 m. Because it would be too time-consuming to reprocess all data needed to generate Fig. 7, one single orbit cycle of Sentinel-3A was processed to check the impact in the surface elevation estimated. The results will be reported in a new supplementary section S8 (available pages 14-15 in this document). This study is also used to demonstrate one of the assumptions stated in section 5.2 (third bullet point), as discussed with Reviewer #2.

2. How are data gaps in REMA handled? It's probably worth highlighting the issues with REMA (e.g. it's static) to emphasize the benefits of AMPLI.

This point is covered in Supplementary section S3: "*the measurement is discarded if the DEM is not 100% complete in the cross-track direction to the nadir point, up to ± 8 km*". This corresponds to ~0.73% of the measurements over the Antarctic ice sheet. It might be noted that in this study we used REMA version n°2 (released in 2022). While in Dong et al. (2021), they used REMA version n°1.1. If I am not mistaken, the DEM coverage was improved in the version n°2, compared to n°1.1. But I unfortunately cannot find a source to confirm this assumption.

I am not sure how to cover the second point, as by nature a DEM is static, while radar altimetry can provide regular temporal elevation changes thanks to the periodic revisit of the satellite. So, I do not think we should oppose the two, they generally just have different applications. On top of this, the achievements reached with AMPLI partly rely on REMA performance, to simulate accurate model waveforms.

3. It appears from Table S1 that the surface backscattering coefficient in equation 1 is constant. This was also covered in the discussion as a possible source of error. How was this determined and what is the impact on surface elevation retrieval for a realistic range of values?

I think there are two aspects to be discussed related to the constant value of the backscattering coefficient (Sigma-0), and I am not sure which one is questioned here. Therefore, I reply on both:

1) A 6 dB value was chosen for all the simulations, wherever the location of the satellite, because this is approximatively the average value found over Antarctica. But, since the simulated waveform is normalised

by its maximum value during the relocation processing, the absolute value of the Sigma-0 has no impact on the relocation processing, and therefore on the topography retrieved (this information will be reported in supplementary S1, for the sake of clarity). I also realised that this normalisation was not clearly indicated in the initial manuscript. It will be stated line 199, by adding: "*Both WF and SWF signals are normalised by their respective maximum value, as displayed in Fig. 3.*"

2) When solving Brown's equation, we ignored the variation of backscattering coefficient with angle of incidence (i.e. leading to a constant Sigma-0 in equation n°1). This point will be mentioned in the text, line 126, and will be further discuss in supplementary S1, as follow:

Added sentence in the main text *(line 126)*

"As Brown (1977), we ignored the variation of backscattering coefficient with angle of incidence. This simplification was made to facilitate and speed up the numerical calculations. and is further discussed in supplementary material, section S1."

Updates in Supplementary S1 *(added text is underlined)*

"σ_0 *is taken as a constant (= 6 dB). We neglect the* ***σ_0*** *variation with angle of incidence, by making the assumption that the impact remains relatively minor compared to the antenna aperture. We assume this assumption can be taken, based on the relative homogeneity of the ice sheet surface (at the footprint scale), in terms of backscattering properties. Brown (1977) also took this assumption for the ocean surface. <u>This simplification was confirmed to be valid, given the relatively good agreement obtained in the simulated waveforms, when they were compared to Sentinel-3 ones (as presented section 2.4, and in supplementary material S2). Nevertheless, it would be still worthwhile considering the</u>* ***<u>σ_0</u>*** *<u>variation with angle of incidence, for further refining the physical modelling.</u>"*

4. Equation 5 – I found this confusing – should delta H and delta R be written as vectors rather than scalars for dot/cross products? Although simple, a geometry image may help. Are x_r, y_r and z_r derived from REMA?

(x_r, y_r, z_r) are derived from the location of point of first radar return, determined via *COG_{CTBD}*, defined in equation (4). To clarify, The text + equation (5) will be modified as follow:

Added text is underlined (line 221):

<u>COG_{CTBD} provides the cross-track distance (in meter unit) between nadir and point of first radar return. This information is employed to geolocate the point of first radar return on a polar stereographic plane (EPSG:3031).</u> The look angle between nadir direction and point of first radar return is computed as:

$$Look\ angle = arccos\left(\frac{\overrightarrow{\Delta H} \cdot \overrightarrow{\Delta R}}{\Delta H\ \ \Delta R}\right) \qquad (5)$$

5. Section 4.1. Please could you clarify why it is more computationally efficient to calculate the elevation DEM residuals against REMA rather than differencing the elevations derived in the two time periods? Does this magnify or reduce the effect of REMA uncertainties? Although presumably possible to use a small test area to process the full time-series of S3, including this would likely complicate this paper unnecessarily – perhaps you could comment in response whether you plan to do this in future work.

As mentioned in the text; lines 356-357: "*In general, ice sheet elevation rates are derived from radar altimetry measurements using repeat-track observations of consecutive cycles (Moholdt et al., 2010). Such approaches require continuous time series to be applied*". In fact, one of the problematics in the repeat-track method, is to account for (1) the ± 1 km deviation of the satellite ground track (2) the sparse surface sampling, with one elevation estimates available every ~300 m along the track (~20 Hz rate). For these two reasons, the surface sampling is never the same from one orbit cycle to another.

To sum-up, the elevation change calculation through a direct difference of the along-track elevations derived in the two periods is complicated, because the radar altimetry measurements are not co-located in space. Thus, the choice of using elevation residuals against DEM was not made for “computational efficiency” reasons, but rather because we considered it was the most accurate method with the data set available. As mentioned in the text, this is not a new method. As the method is sensitive to DEM precision (not accuracy), we made a validation by comparing the SEC obtained with this approach to an alternative one. This is already presented in the supplementary materials, section S5.

We also initially envisaged processing the full time series over a small area, in order to develop a repeat-track method for the SEC calculation. However, this would further complicate the paper, require additional work, and we would not be able anyway to provide a global assessment over whole Antarctica. And I think it is highly important ensuring a complete validation of the AMPLI software over the whole Antarctica, rather than arbitrarily choosing a specific region.

As it will be mentioned in an update of the “data availability section”, a full mission reprocessing is on-going with the AMPLI software, and the data will be available in open access. We will likely further assess Sentinel-3 SEC with AMPLI, but there is nothing planned at present. We also hope this will be endorsed by other groups, given the data will be openly available.

6. Figure 7. Please could you comment on why the Antarctic Peninsula thinning shown in IceSat is not captured by S3. Excluding central Antarctica centre is justified for statistical calculations, but should be included in this figure.

The coverage of the SEC maps calculated with Sentinel-3 measurements is discussed at the beginning of section 4.3, lines 396-405, and the missing grid points in the SEC maps are explained in the text. As mentioned, the ice sheet margins are not entirely covered due to (1) on-board tracking failures (2) bad quality measurements (because of the irregular and/or steep topography over these regions). There is a reference to the “Sentinel-3 2023 annual quality report”, written by the Sentinel-3 Mission Performance Cluster, which contains more information on this subject.

Moreover, the SEC maps calculated with Sentinel-3 measurements do not cover beyond 81.5 °S, because this is the limit of the Sentinel-3 orbit, as stated in introduction. Thus, central Antarctica cannot be sampled by Sentinel-3.

Specific comments:

- Figure 1a. It’s difficult to see the ‘nadir’ because it’s green on cyan on green. Please could you change this font colour. Also state yellow line is F_i=0).

Done. I enlarged ‘nadir’, so it is more distinguishable. There were unsuccessful attempts with other colours. Available page 17 of this document.

- Figure 2. Include Point of First Radar Return on this (presumably instead of POCA)

Done. Available page 17 of this document.

- Figure 3 Please increase font size

Done. Available page 18 of this document.

- Line 215. Define u_begin and u_end. State here what % measurements were ambiguous and therefore discarded (or refer to section S3).

This definition will be added for u_{begin} *and* u_{end}*:*

"The cross-track surface finally considered sampled by the waveform leading edge is delimited by the u_{begin} *and* u_{end} *indexes (green dotted lines in Fig. 3d)."*

Because at this point of the text we are still describing the method and discussing case study, it is a bit early to provide global statistics in Antarctica. Thus, these numbers will be added in the main text, where the Sentinel-3 data set and the quality controls are introduced, lines 286-288.

- Line 310-311. Separate S3 and IceSat 3 numbers – it is unclear whether there are 1,787 or 1,787 million S3 observations. Line 272 suggests 27 million.

There is a misunderstanding in the sentence interpretation, probably because it is ambiguously written. Thus, it will be simplified as follow:

"In total, approximately 1,787 and 1,750 million ~~Sentinel-3 and ICESat-2~~ co-located ~~elevations~~ measurements were identified over the Antarctic ice sheet, with the ESA L2 and AMPLI data set, respectively"

(note: "elevations" was also changed by "measurements")

- Line 321. 'As the population…' is an incomplete sentence.

Thanks for spotting this. The start of the sentence was changed ("in fact" instead of "as"). As additional change, "*with the decreased latitude*" is substitute by "*in the southernmost latitudes*"

« In fact, the population of co-located measurements significantly increases ~~with the decreased latitude,~~ in the southernmost areas due to narrower track spacing. »

- Line 460. Please could you comment about the effects of surface roughness e.g. sastrugi for both radar and laser observations.

In the rewritten section 5.3, available pages 16-17 of this document, the manuscript has been updated to elaborate more on the effect of surface roughness in the radar signal.

- Figure title for Figure S6c should be (b)-(a) not (b)-(c)

Thanks for spotting this.

New supplementary section S8

Section S8: Impact of a bias in the input DEM used for the facet-based simulation

In this analysis, artificial vertical elevation biases (fixed offset) were introduced in REMA to assess the impact on the surface topography retrieved with the AMPLI software. These bias values are ranging from −5 m to +5 m, in increments of 2.5 m, same intervals as taken by Huang et al. (2024, section 4.5). The processing chain was run with these configurations over the orbit cycle n°45 of Sentinel-3A (acquisitions from 17 May to 13 June, 2019). As reference data set, the AMPLI software was also run in its “nominal” configuration (i.e. no bias in the input DEM). The surface elevation from Sentinel-3A and ICESat-2 ATL06 is compared at nearly co-located points, following the methodology described in section 3. The spatial search radius was enlarged, from 25 m to 50 m, to increase the population of co-located points.

Table S4 displays the median bias and Median Absolute Deviation (MAD) between Sentinel-3A and ICESat-2 ATL06 for several slope ranges (same ones as in Table 1). The median bias between Sentinel-3A and ICESat-2 ATL06 was found to be identical in the few centimetres range, between the five configurations tested, and for all slope bins. The population of co-located measurements slightly decreases over the ice sheet margins when the vertical bias added to the DEM increases, as more measurements are rejected when controlling the vertical bias between DEM and altimetry data. In fact, when checking the agreement between the simulated and the measured waveform, the absolute range bias between both must remain below 30 gates (~14 m), as stated in supplementary section S3.

In Fig. S9 the elevation bias between Sentinel-3A and ICESat-2 ATL06 is mapped using a 100 km stereographic grid (EPSG:3031), for the “nominal” configuration (Fig. S9a) and those with + 5 m (Fig. S9b) and -5 m (Fig. S9c) elevation biases added in REMA. The differences between the map grid points are plotted in Fig. S9d (“*+5m*” - “*nominal*”) and Fig. S9e (“*-5m*” - “*nominal*”). In these maps no major change in performance is detected between the three configurations. The absolute differences between the configurations remain below 5 cm for 92% and 86% of the map grid points, in Fig. S9d and Fig. S9e, respectively. Whereas the effect of a bias in the input DEM can generate several metres of elevation errors with LEPTA and MPI algorithms, as reported in Huang et al. (2024, section 4.5).

This result is crucial, attesting that AMPLI can monitor the vast majority of the polar ice sheets over a period of several years, without noticeable errors introduced in case of temporal elevation changes (if homogeneous over the radar footprint), as discussed section 5.2. This outcome is also corroborated by the high agreement in the SEC estimated by Sentinel-3 AMPLI and ICESat-2 ATL15, as presented in section 4.

Surface Slope		< 0.1°	0.1° - 0.5°	0.5° - 1°	> 1°
Population ($x10^3$ count)	-5 m	94 $x10^3$	133 $x10^3$	16 $x10^3$	3.9 $x10^3$
	-2.5 m	94 $x10^3$	133 $x10^3$	16 $x10^3$	3.8 $x10^3$
	Nominal	94 $x10^3$	133 $x10^3$	15 $x10^3$	3.7 $x10^3$
	+2.5 m	94 $x10^3$	133 $x10^3$	15 $x10^3$	3.5 $x10^3$
	+5m	94 $x10^3$	131 $x10^3$	14 $x10^3$	3.3 $x10^3$
Median bias (m)	-5 m	+0.11	+0.16	+0.35	+0.46
	-2.5 m	+0.11	+0.16	+0.34	+0.45
	Nominal	+0.10	+0.14	+0.31	+0.42
	+2.5 m	+0.11	+0.15	+0.32	+0.45
	+5m	+0.11	+0.14	+0.32	+0.45
MAD (m)	-5 m	0.10	0.16	0.34	0.49
	-2.5 m	0.10	0.16	0.34	0.50
	Nominal	0.10	0.15	0.33	0.50
	+2.5 m	0.10	0.16	0.33	0.50
	+5m	0.10	0.16	0.33	0.49

Table S4: Statistics of the elevation difference between Sentinel-3 AMPLI and ICESat-2 ATL06 nearly co-located measurements (calculated as Sentinel-3 – ICESat-2) over the Antarctic ice sheet, for different slope intervals. Five AMPLI configurations are assessed: the one presented in section 2 ("Nominal") using an unbiased DEM, and four others with vertical biases introduced in REMA, ranging from -5 m to 5 m. The measurements acquired further south than 80°S are not considered in this analysis, in order to mitigate statistical over-representation of southern observations, where the population of co-located measurements significantly increases.

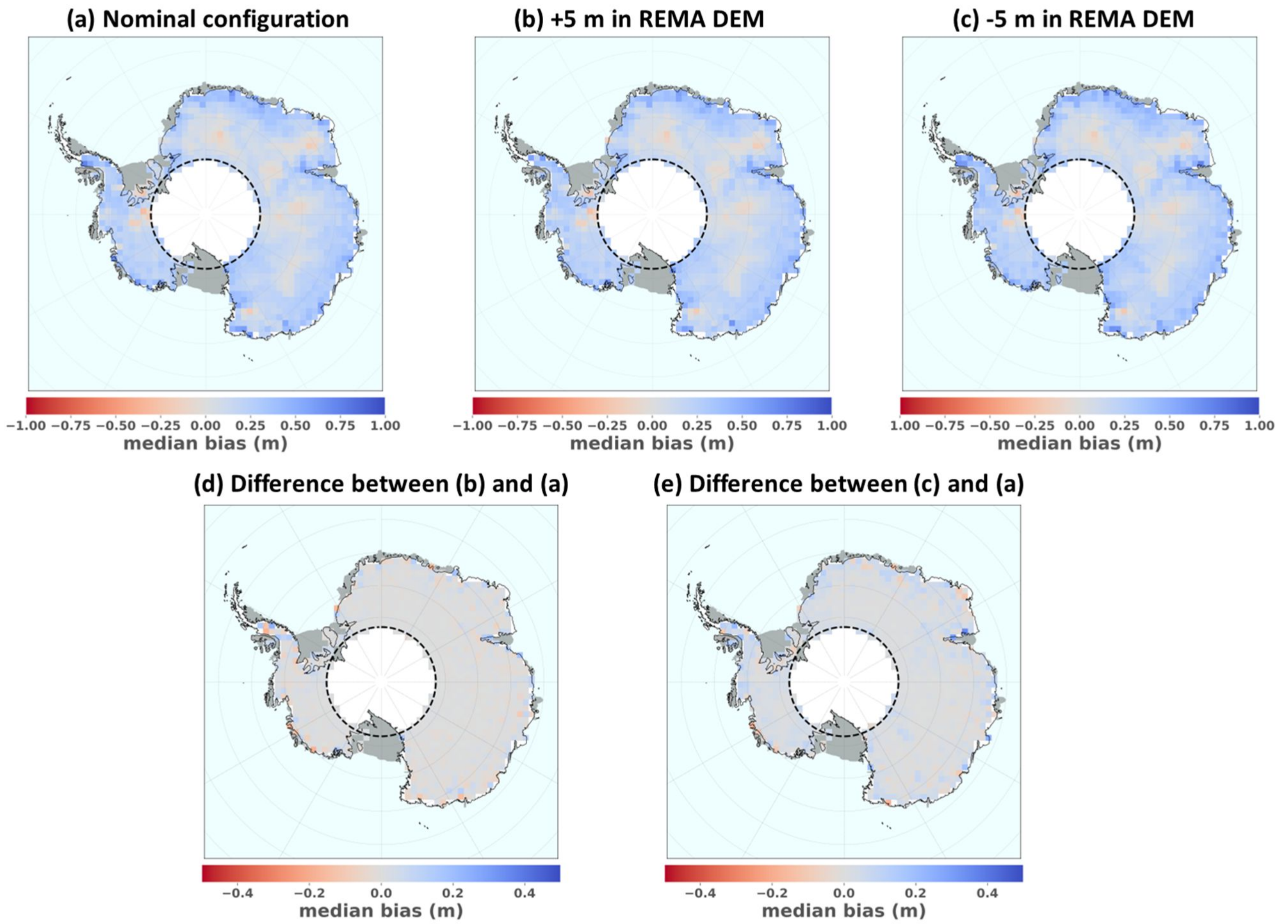


Figure S9: (top panels) Median elevation bias between Sentinel-3 AMPLI and ICESat-2 ATL06 co-located elevations, mapped using a 100 km grid (a) nominal AMPLI configuration, as presented section 2 (b) AMPLI configuration with a +5 m bias introduced in REMA (c) AMPLI configuration with a -5 m bias introduced in REMA. (bottom panels) Map differences between the nominal configuration and ones with +5 m (d) and -5 m (e) biases introduced in REMA.

Update of section 5.3 (in blue colour)

5.3 Areas of improvements

Among the approximations taken in the numerical modelling described in section 2.2, the main geophysical effect missing is likely the snow volume scattering (Ridley and Partington, 1988). It was outside the scope of this work exploring an accurate modelling for this effect, applicable and validated all-around Antarctica. However, previous studies have shown that snow volume scattering has a limited impact on the waveform leading edge in Ku-band SAR mode, and rather affects the waveform trailing edge (Aublanc et al., 2018). In supplementary material, section S7, we provide an example illustrating the difference between the measured and simulated Sentinel-3 UF-SAR waveforms over the EAIS interior. Since the volume scattering signal is not yet modelled in AMPLI, the relative energy level of the simulated waveform is lower on the trailing edge part, in comparison to the measured waveform. However, in the relocation procedure developed, the position of the first radar return is estimated within the area sampled by the waveform leading edge part only. This likely explains why neglecting snow volume scattering seems to have a limited impact on the Sentinel-3 AMPLI performances, as reported sections 3 and 4. Nevertheless, integrating the snow volume scattering effect to the numerical model remains a priority. It will further improve the agreement between the simulated and measured waveforms. In the current processing, this will be beneficial for better detecting simulation errors. In addition, such modelling enhancement would be ~~valuable~~ used to remove the artificial topography variations induced by temporal changes in the snowpack parameters, as mentioned in section 5.1.

Alongside the numerical modelling aspects, we identify that there are two main different geophysical effects not yet corrected in the developed processing. Firstly, it is critical compensating for the impact of the snow volume scattering in the altimeter range, which generates errors in the SEC retrieval, as stressed in section 5.1. This can be done by modelling the effect and adding it into the facet-based simulation with the aim of developing a retracking algorithm. Such modelling work has been initiated for LRM altimetry (Lacroix, 2008; Larue et al., 2021), but it has never been extensively tested over the entire ice sheet. Alternatively, a post-correction could be added to the elevation retrievals, as commonly done in LRM altimetry using waveform shape parameters (Flament and Rémy, 2012; Simonsen and Sørensen, 2017).

Secondly, we anticipate that the surface roughness, from metre (e.g. sastrugi) to kilometre (e.g. megadunes) scales, would act on the waveform leading edge width. In principle, this would generate errors in the altimeter range, because this effect cannot be compensated by an empirical threshold retracker, as currently implemented in the AMPLI or ESA L2 processing chains. While the associated elevation errors are not yet precisely quantified, we estimate they do not exceed few decimetres with AMPLI, given the results presented in section 3.4. Furthermore, since such topography features remain relatively temporally static, we also expect these errors to be constant over the time for a given location, and therefore not significantly influence the SEC estimation. Nonetheless, this hypothesis remains to be confirmed. Furthermore, the relationship between the surface roughness (metre-to-kilometre scales) and the radar waveform shape shall be further assessed, for improving or redefining the retracking algorithm.

Moreover, neural network solutions can be also envisaged to enhance the performance of the AMPLI software. As a first application, these techniques could be employed to better detect modelling errors. In fact, a disagreement between the simulated and measured waveform can for example occur in case of discrepancies in the input DEM, or heterogeneous elevation change variations over the radar footprint, as discussed section 5.2. A more advanced use would be to develop a neural network retracker, as recently done by Helm et al. (2024). Their retracking solution, named AWI-ICENet1, was applied to CryoSat-2 LRM waveforms acquired over the interiors of the Antarctica and Greenland ice sheets. The results show a reduction of the elevation errors associated to snow volume scattering, compared to other conventional retracking solutions. In parallel to the development of innovative neural network solutions, we also underline the paramount importance to continue improving the radar signal modelling. This is crucial for an efficient and accurate training of the neural networks.

Additionally, in the current processing, we decided to tolerate a maximum of ~14 m elevation difference between the altimetry measurement and the static DEM (as estimated in the cross-correlation, illustrated in Fig. 3b). A future processing evolution ~~shall~~ could include a dynamic vertical readjustment of the DEM, for successfully monitoring regions undergoing rapid elevation changes. Finally, a multiple peak retracker would also be valuable to increase the data sampling over the ice sheet margins, as shown in Huang et al. (2024).

Updated Figures

Figure 1

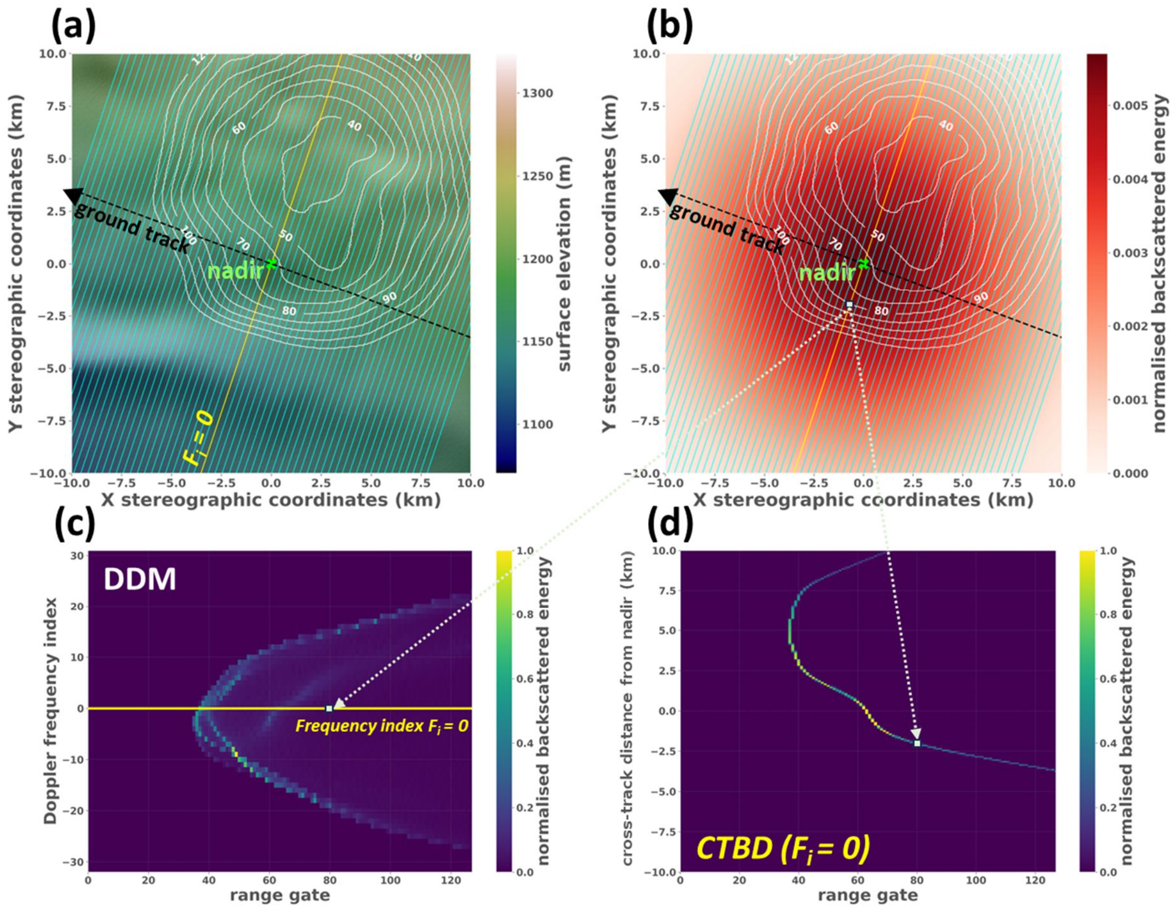


Figure 2

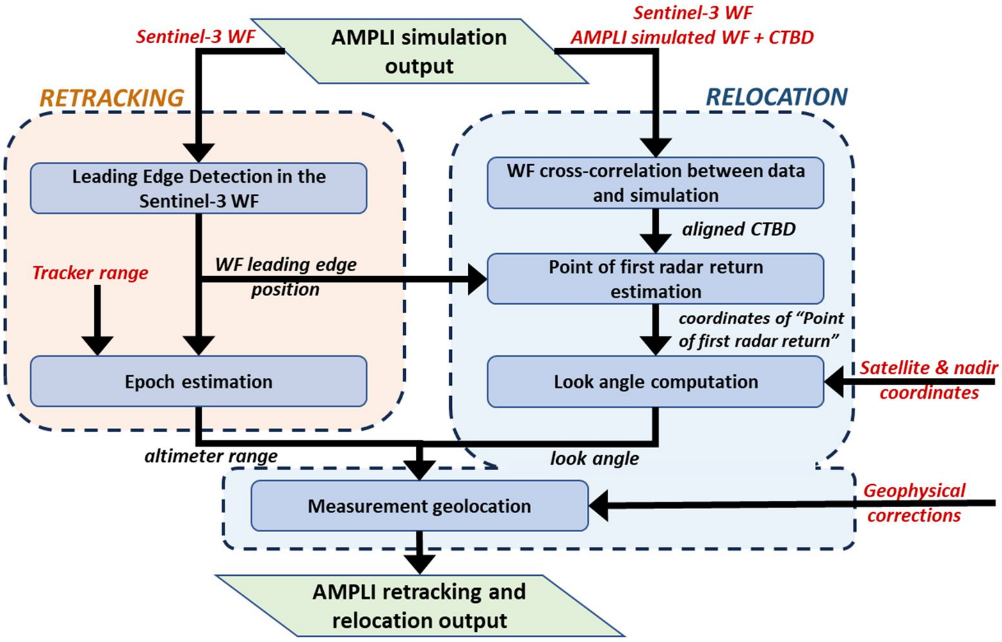


Figure 3

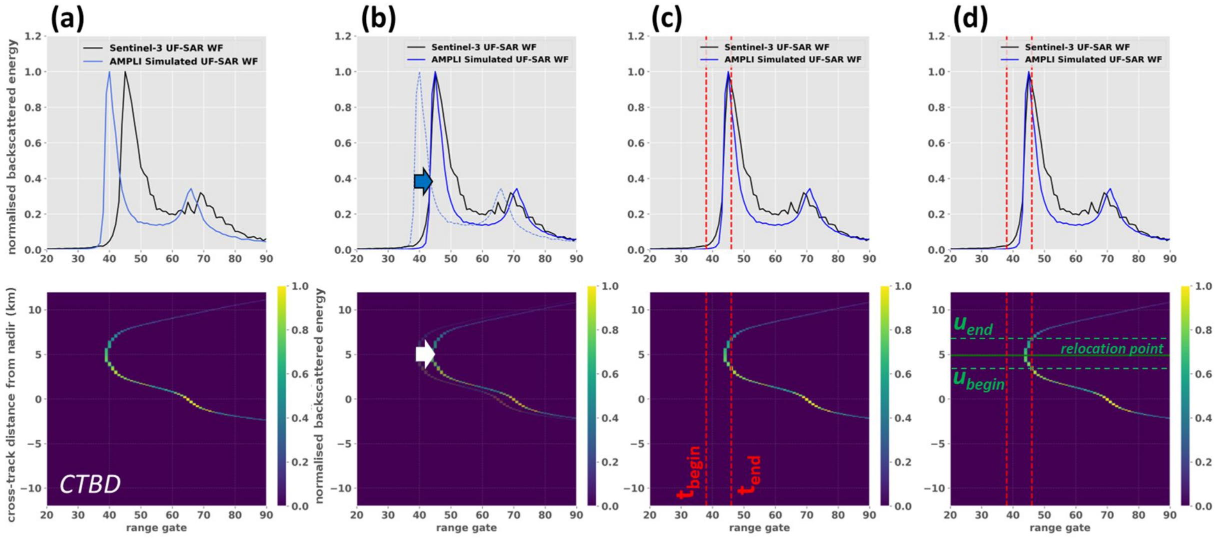


Figure 4

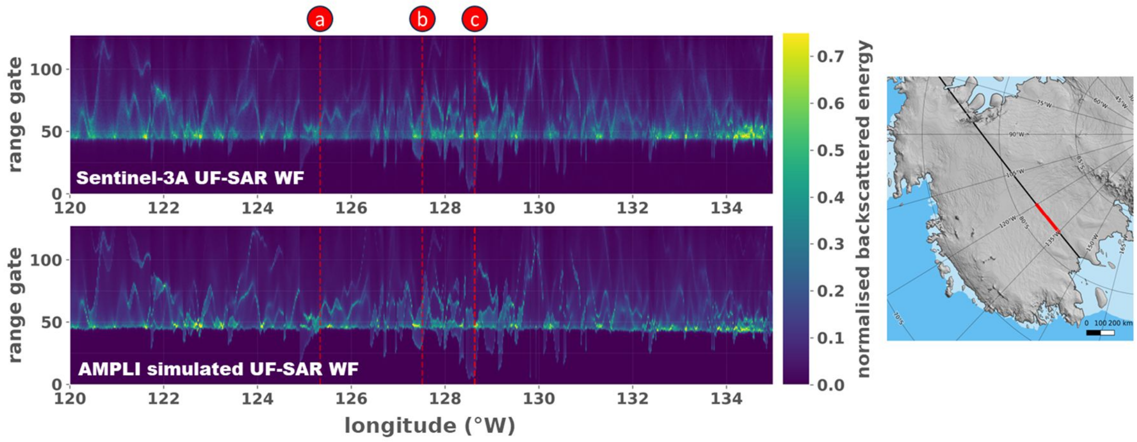


Figure 5

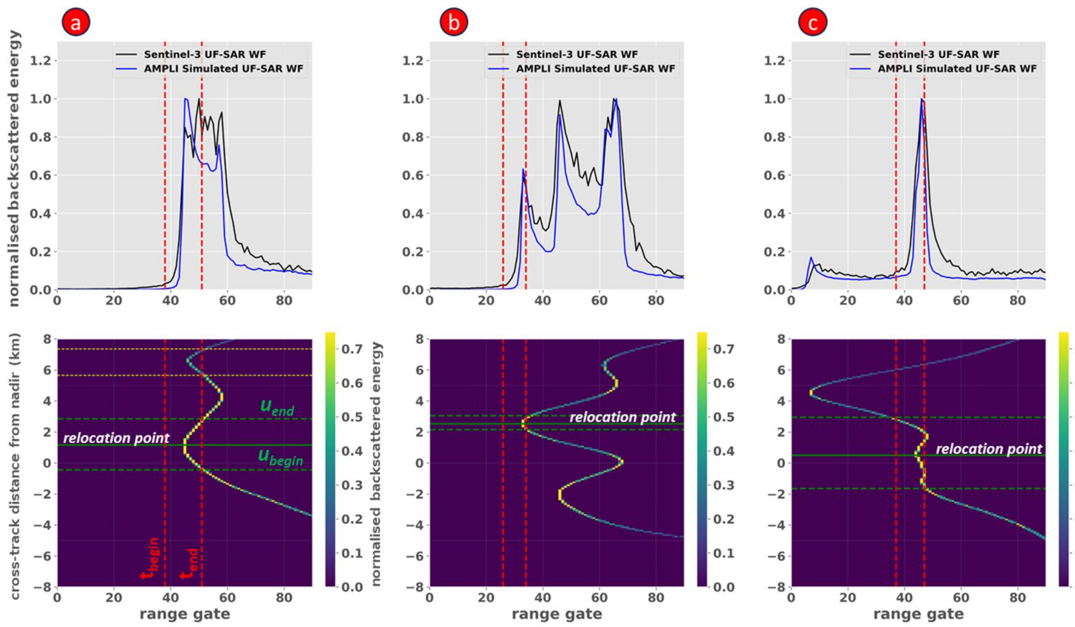
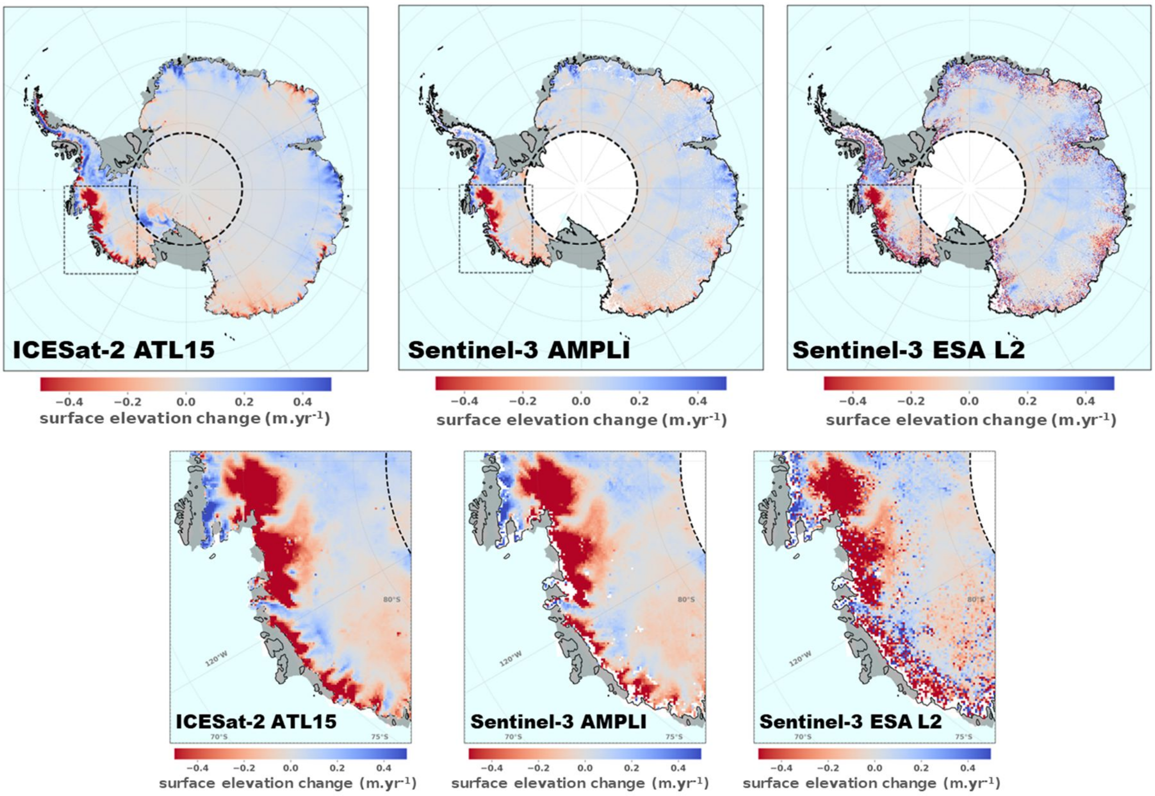


Figure 7


ICESat-2 ATL15
Sentinel-3 AMPLI
Sentinel-3 ESA L2
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)
ICESat-2 ATL15
Sentinel-3 AMPLI
Sentinel-3 ESA L2
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)
−0.4 −0.2 0.0 0.2 0.4
surface elevation change (m.yr-1)