The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2018-120-AC1, 2018

Interactive comment on "Sentinel-3 Delay-Doppler Altimetry over Antarctica" *by* Malcolm McMillan et al.

Malcolm McMillan et al.

m.mcmillan@leeds.ac.uk

Received and published: 7 December 2018

Response to Reviewers

Thank you for the reviews of our manuscript. We have now completed revisions in order to address the comments raised by both reviewers. In summary, the main changes are:

- Adding a full continent-wide cross-over analysis to assess measurement precision.
- Adding a full continent-wide evaluation of the Sentinel-3 measurements against IceBridge.
- Assessing the relationship between Sentinel-3 – IceBridge elevation differences

and the magnitude of surface slope.
- Adding validation statistics for all three retrackers.
- Including further glaciological analysis of the data, specifically demonstrating the capability of Sentinel-3 to identify possible subglacial lake activity.
- Comprehensive redesign of the figures, to make the graphics clearer and to ensure better use of space.

Full details of all of our responses to the individual comments made by the reviewers can be found below. Reviewers' comments are in italics and our responses are in bold type. We believe that the changes have substantially improved the manuscript and we are grateful to both the reviewers for their comments.

Reviewer 1

This paper presents some of the first results from Sentinel 3's doppler focused radar altimeter over Antarctica. It describes the instrument and processing strategies, and it gives evidence that the data have decimeter precision over flat targets, and up to a few meters precision over sloping surfaces.

The paper is well written and needed very little editing. Its introduction to Sentinel-3 data should be valuable to the research community, and it makes the case well that the data are useful. The paper is somewhat weak in the analysis (rather than illustration) of the data.

I would have liked to see the elevation error magnitudes explained by mechanisms, rather than just quantified by region: it appears that even in the DML and Wilkes Land areas, the errors are multimodal, with a narrow distribution in flat areas and a wide distribution in areas with high relief. Were I using the data, I would want to assign error estimates to the data based on, for example, the local surface slope or the footprint-scale relief. Calculating errors based on histograms for areas hundreds of km on a *side does not do much to help me do this. Likewise, it would be useful if the authors could explore a quantitative connection between the roughness differences at Dome C and Vostok and the biases, and perhaps include the other sites in this analysis. Without understanding the roughness at the DML and Wilkes sites, I don't know if the roughness difference between Vostok and Dome C is significant.*

We agree with the reviewer's point and, as requested, we have added a new investigation of the impact of surface slope on our error estimates. In order to do this, we have expanded our previous site-specific evaluation to cover the whole ice sheet. We have then compared the error magnitude to surface slope in order to assess the relationship between the two, and to evaluate the impact of increasing slope upon the accuracy of the Delay-Doppler altimeter measurements. We have also undertaken a continental-scale cross-over analysis, which we use to assess the precision of the measurements as a function of surface slope. As requested, we have quantified the roughness – based upon our IceBridge profiles – at both Dome C and Vostok, and show that this is 4 times greater at Dome C. We have not repeated this type of analysis at the coastal sites because this analysis is specifically designed to investigate whether the relative bias at *low slope* sites can be explained by differences in the small-scale surface roughness. This is not relevant to the other sites because (1) there is no significant bias, and (2) they are not flat (at long wavelengths), and so at these more complex sites the radar response is a convolution of a broad range of topographic wavelengths and we believe it does not make sense to consider the small-scale topography in isolation.

My other concerns about the figures. Fonts are consistently too small, and lines are consistently too fine. The authors should print out their paper and try to read the figures!

We have redrawn most of the figures in the revised manuscript and have made sure that we increase the font and line size so that the figures are now much

clearer. Further details are provided below in the response to Reviewer 2's comment 3, related to the figure presentation.

Figure 2: Panels d and e show the same thing as b and c. I recommend just showing just one of these profiles. The green bars do not seem to show anything different from one part of the pass to another. I would recommend showing a histogram of the data contained in the bars (i.e. a histogram of the 400-m standard deviations).

As suggested, we have condensed this figure by replacing the profiles of the standard deviation – the green bar panels – with histograms of the data instead. We do, however, prefer to still keep both ground tracks because we want to show that the analysis holds for more than one track, and in particular is independent of a specific viewing geometry.

Figures 3 and 4: It is hard to tell if the DEM spans the axes, or if there is extra white space around the DEM surface. The elevation axis is illegible, as are the labels on the inset panels.

We have revised these figures to address the points raised here, and to make them generally easier to interpret. We have replaced the 3-d DEM plots with 2-d versions overlaid upon MOA. We have also increased the font size. See also our response to the related comments by Reviewer 2 below.

Figure 3: The span of the colour scale for the coloured points is too large, and should be reduced to the span of the residuals.

As requested, we have reduced the span of the colour scale to +/- 1 m.

Figure 5: I could not find a description of the differences between panels A-C, and between panels E-G, either in the text or the caption. Are there multiple profiles in d and h? The lines and points in D and H are too small, and the axis label font sizes are unreadable.

This was an omission on our part, thanks for spotting it. We have added details

in the caption to explain that the profiles relate to the three different sub-tracks resolved within the ATM swath. We have also redrawn the figure with larger points, lines, and axis labels.

Reviewer 2

This manuscript assesses the quality of Sentinel-3 delay/Doppler radar altimetry over Antarctica through analysis of: precision at Vostok Subglacial Lake and Dome C; accuracy at Vostok, Dome C, Dronning Maud Land, and Wilkes Land; and elevation change across the continent. The authors find that Sentinel-3 altimetry achieves a precision of <0.1 m over flat topography, has a mean bias <0.2 m in the flat interior and <1.5 m in complex coastal terrain, and can reproduce known patterns in elevation-change rate at the continent scale. The paper represents the first use of Sentinel-3 altimetry data over an ice sheet, is generally well-written, and will be of significant value to the community as an introduction to the capabilities of this Sentinel mission, but there are three major weaknesses that should be addressed before publication:

1. Limited Analysis: While the overall analysis presented in the manuscript is sound, it is also much more limited than I would expect from a rigorous treatment of altimetry uncertainty. The authors only assess precision in two flat, interior locations, yielding a "best-case" estimate, while opting not to perform a continent-wide crossover analysis, which would provide the readers with much better insight how and why precision varies. Crossover analysis is a standard technique in satellite altimetry error analysis and should be included in any rigorous assessment.

We realise that our motivation for starting our analysis at these "best case" sites was not clear, and so we have added further text to clarify this point. The analysis at Lake Vostok was not supposed to quantify measurement precision across the entire continent. Rather it was meant as an initial analysis of the instrument precision, in order to characterise the influence on the SAR measurements of factors such as radar speckle and retracker imprecision.

This type of analysis – to quantify the shot-to-shot noise – we believe is best done in the absence of additional complicating factors such as complex topography, and – as the reviewer points out – is not indicative of measurement precision in more challenging regions. It is, however, a useful benchmark as it can be compared to similar analyses of other satellite missions that have previously been conducted at Lake Vostok.

We do agree with the reviewer's comment that a single-cycle cross-over analysis was an obvious omission from the original manuscript, and would be a useful addition. We have therefore undertaken a continent-wide cross-over analysis for two full cycles of data and used this to investigate the relationship between the magnitude of the cross-over differences and the ice sheet surface slope. As well adding new text, we have also added a new figure and a table to report these cross-over results.

In a similar vein, there are only four regions included in the accuracy analysis, when Operation IceBridge has surveyed across the continent. It is not clear why this assessment was so limited in scope, when including the rest of the Operation IceBridge airborne laser altimetry dataset would help illuminate the causes of error in the Sentinel-3 altimetry and provide the community with much more direction as to where and how to use this new dataset.

We had initially focussed on specific sites because we wanted to provide a detailed analysis over different categories of topographic terrain. We do, however, agree with the reviewer's point here, and so we have now (1) added an evaluation that utilises the continent-wide IceBridge record, and (2) analysed the relationship between the magnitude of the elevation differences and the ice sheet surface slope. We have added a new figure and supporting text to discuss these results.

Finally, the analysis of elevation-change rate is entirely qualitative; while I certainly *understand that the processing chain is still in development, there should be some quantitative analysis of the result if it is to be included in the manuscript, such as a comparison to CryoSat-2 derived elevation-change rates over the same time window. At the very least, a map of "known" dh/dt from CryoSat-2 or ICESat should be shown in conjuncture with the Sentinel-3 map to show the patterns are generally similar.*

To address this point, we have expanded the quantitative description and scope of this section by (1) reporting rates of elevation change close the grounding lines of Amundsen Sea glaciers, and (2) adding an additional quantitative analysis of localised ice sheet lowering in East Antarctica, which we believe is likely to be the signature of subglacial lake drainage event. Together these provide a quantitative, if preliminary, demonstration of the capabilities of Sentinel-3 for mapping glaciological signals of change. We would, however, prefer not to give a detailed quantitative inter-comparison between datasets, because we feel that the Sentinel-3 record is still too short for this type of analysis. Also, given the comments about efficient use of figure space, and the fact that numerous published maps of dz/dt are readily available within the literature, we believe that it is better to refer the reader to this published material, rather than to use more space reproducing them here. We have added more text to point the reader to published baseline datasets and figures that can be used for this purpose.

2. Data and Method Detail: There are quite a few pieces of information missing with respect to the data sources used and the methods applied. There is a long list of corrections applied to the altimetry data, but no citations that provide the details of the models. There is no data citation related to the ATM laser altimetry (though there is for the Riegl data), which prevents the reader from knowing what level of data processing was used or other important details for understanding cm-scale accuracy. It appears to be L2 data; why was L2 chosen over L1b (full-swath)? Are the same corrections with the same models applied to the airborne altimetry as to the satellite altimetry? What are the dates of the Operation IceBridge flights (which is needed understanding the

potential impact of the dh/dt correction)? Last, the method for correcting space and time elevation differences for the satellite-airborne analysis requires some more detail: how is a 1 km DEM being used to correct for spatial separations of <100 m? What is the error associated with the dh/dt correction? What is the magnitude of these corrections (mm? cm? m?)? Also, at a basic level, are the results presented as hsat − hair or hair − hsat? The former is typical, but the manuscript implies the latter ("IceBridge-Sentinel-3 measurement pair"); resolving this issue is fundamental to the analysis of why the bias exists (p.7, lines 7–9).

As requested, we have added considerable extra detail to the description of our data are methods, namely:

1. **We have added references to describe where details of the methods and models used to generate the instrument and geophysical corrections can be found, together with a reference for the ATM laser altimetry. We choose to use the corrections supplied with, or applied within, each of the satellite and airborne products because the corrections will be instrument dependent – for example radar vs laser – and we believe that the teams that operate and distribute the data are best placed to select the most appropriate models to use.**

2. **We have provided additional clarification that we use the Level 2 ATM data, because our understanding is that this is effectively a smoothed version of the L1b, which both reduces the impact of uncorrelated shot-to-shot noise, and also brings the measurement cell closer to the resolution of the SAR footprint. We select this because we want the validation data to be as accurate, and as similar in scale to the data being validated, as possible.**

3. **We have added details of the dates of the individual IceBridge campaigns at each site, which are now included in Table 3.**

4. **We have provided further detail relating to the correction of spatial and temporal elevation differences, both in terms of the method used for the spatial correction (a bicubic interpolation based upon the surrounding 4x4 pixels of the 1 km DEM), and the magnitude – and implications – of the temporal correction (< 1cm/yr at Vostok, Dome C and Dronning Maud Land, but larger – 8 cm/yr – at Wilkes land).**

5. **We have clarified that the results are indeed the typical hsat – hair, and made sure this is clear and consistent within the manuscript.**

3. Figure Presentation: The figures could be greatly improved to make them easier to understand and to use journal space better. Cyan and yellow for groundtracks on Figures 1, 3, and 5 are hard to see, printed or on a screen. I am not sure why the maps on Figures 2, 3, and 4 are plotted as 3D perspective plots; this style makes it harder to understand the spatial context, particularly how the tracks align with easting and northing (which is important since this is how the transects are plotted). The axis and legend labels on all figures are unreadably small and the line widths are too small. I am not convinced that Figures 2c and 2e add to the discussion over a single histogram of uncertainty in 400 m along-track bins across the entire elevation range. The histograms and cumulative distribution functions in Figures 3b/3c and 3e/3f can likely be combined into a single panel. The authors should rethink how the each of the figures is presented to make the best use of space and to ensure they are fully legible at production size.

As requested we have undertaken a thorough revision of the figures. We have (1) removed the 3-d perspective plots in Figures 3 and 4, and replaced them with 2-d plots overlaid upon MOA, which hopefully makes them more easily interpretable, (2) increased the size of axis labels and line widths, (3) removed the along-track profiles in Figures 2c and 2e, and replaced them with single histograms, (4) combined the histograms and cumulative distribution functions into a single panel in all relevant figures, and (5) condensed Figures 3 and 4 into a single figure. In

a few cases, we have kept the ground track lines intentionally thin, because we tested using a thicker line and the figures became too cluttered and difficult to interpret. Our intention is that the priority for these figures is to clearly see the elevation differences, and that the track lines are purely in the background for reference.

Specific comments

line 13: "the first step towards a new era" is a bit over the top. The Antarctic coast is already blanketed in better altimetry (that includes along-track delay/Doppler processing). Please tone down these statements throughout (especially the use of "novel").

Our intention here was to refer to the fact that Sentinel-3 represents a new class of *operational* altimetry satellites (which is important for the purpose of moving to more long-term, systematic monitoring), rather than the novelty of the delay-doppler system (which as the reviewer highlights, has already been accomplished by CryoSat-2). However, we appreciate that our wording was not clear and so we have revised the text accordingly. As requested, we have also toned down similar statements throughout the manuscript.

p.1, line 20: "accuracy decreases slightly" is an understatement, given most of the measurements are up-to an order of magnitude less accurate.

We agree this wording was too strong and have revised the text accordingly.

p.2, line 16: the rest of the details of what controls pulse-limited footprint should go in this parenthetical statement (i.e., satellite altitude, pulse length).

We have added these details as requested.

p.3, lines 5–7: Figure 1 should have the SARIn/LRM mode mask on it to show where Sentinel-3 altimetry will make the biggest improvement compared to existing datasets (i.e., where Sentinel-3 covers LRM areas). It also should show the latitudinal limit of Sentinel-3.

We have added the CryoSat-2 mode mask boundary to Figure 7 as we felt it was more useful to show the location of it here, relative to the signals of elevation change. The latitudinal limit is also clear in this plot and we have stated this within the figure caption.

p.4, lines 12–14: The rest of the study only uses the TCOG retracker; I think adding some of the statistics from the other retrackers to Table 2 would be a fantastic addition and helpful for the community.

As requested, we have added statistics from the other retrackers into Table 3. We have also added a more complete description of the method in Section 3.

p.4, lines 23–24: I would like to see the % of data points removed by this filtering step.

As requested, we have added statistics relating to the percentage of points removed, which is 3% and 3.7% for the DEM and slope correction filtering, respectively.

p.4, line 27: This date range is insufficient given the dh/dt correction. Please provide exact flight dates, either here or in Table 2.

As requested, we have added the dates of the airborne campaigns for each site into Table 3.

p.4, line 28: The along-track sampling of ATM L2 data is 0.25 s, which results in variable along-track distance spacing due to changes in aircraft velocity.

Thanks for the clarification, we have revised the text to make this point clear.

p.5, line 3: Why was the single closest measurement chosen, instead of some sort of average of all the points within the approximate SAR footprint? Given that surface roughness is later suggested as the cause of bias and that lidar can easily resolve surface roughness, it seems appropriate to try a less simplistic approach that can account for surface roughness.

During the development stage, we did test several different approaches to deriving the IceBridge validation measurement. After testing, we selected our chosen approach because (1) we believe that it is less affected by sampling bias than the method suggested above, because we require that the IceBridge validation measurement is close to the centre of the satellite footprint, whereas the method above can give poor agreement between the satellite and airborne measurements simply because the IceBridge tracks only cross one extreme of the 300 x 2000 m SAR footprint, and (2) it provides a larger validation dataset – and therefore more robust statistics – than another alternative that we tested, which was based on bilinear interpolation of the IceBridge measurements that bracketed each altimeter record. We have added text to discuss this point and justify our approach.

p.5, lines 8–10: A few words about how the distributions deviate from normal would be appreciated here. I am also somewhat worried about the use of median absolute deviation given that it is >20x smaller than the standard deviation in coastal areas, demonstrating that outliers are a significant problem for Sentinel-3 altimetry. One path forward would be to suggest a filtering method for Sentinel-3 users could apply that would bring the two statistics closer to one-another.

We agree that the presence of outliers has a large effect on increasing the standard deviation relative to the MAD. This is precisely why we choose to primarily report the MAD, because for non-normal distributions this statistic provides a more representative measure of the mid-point of the dispersion, which we believe is a more useful metric for the reader. Similarly, we also report the cumulative distribution because we believe it is more helpful to know that a certain percentage of data are within a particular elevation threshold of the validation data. Nonetheless, we emphasise that we do provide both the MAD and the standard deviation within Table 2, so that the reader has full visibility of both statistics. We have also now added further text to clarify why we choose the MAD metric,

and why it is appropriate for non-normal distributions. Furthermore, we have added a supplementary analysis to suggest how filtering based on the deviation from a pre-existing DEM can improve the Sentinel-3 minus IceBridge elevation differences, although we reiterate that developing these strategies is not a priority for this work, and that the most appropriate approach will depend upon each individual application.

p.5, lines 3–4: What % of IceBridge records were removed by this filtering step?

This step removed 7.3 % of the total Antarctic IceBridge dataset. We now report this within the manuscript.

p.5, line 22: It should be noted that you are using 1σ as the definition of precision.

We have added text to clarify this point, as requested.

p.6, lines 4–5: While it may be true, I do not think that the conclusion that the SAR waveform leading edge is insensitive to penetration is defensible given the analysis presented in the manuscript. Perhaps it is the Sentinel-3 sampling geometry that limits observed variability, or, even more simply, there was not much surface variability during the one-year period of observations. This statement should be removed, or alternatively there needs to be a thorough comparison between Sentinel-3 SAR and CryoSat-2 LRM mode using a similar retracker over the overlapping period to demonstrate that SAR processing (specifically) reduces sensitivity to penetration and subsurface scattering.

We agree with the reviewer that the conclusion about the leading edge was not supported by the analysis that we had presented here. We had confused work based on CryoSat-2 SAR data from a different study, which is not part of the current analysis, and so we have removed this sentence from the manuscript.

p.6, lines 8–9: A "slope" correction, in its original meaning, is not applied in this processing chain; it is a "relocation" correction.

We agree this was not the most appropriate term, and have reworded the text to correct this point.

p.6, lines 28–29: Given the importances of these outliers on the statistics, I think filtering strategies should be touched on in this manuscript. All the individual pieces are there to assert a solid filtering pipeline and demonstrate how it improves data statistics.

As suggested, we have added a supplementary analysis of how filtering based on the deviation from a pre-existing DEM can improve the Sentinel-3 minus IceBridge elevation differences, although we do reiterate that developing these strategies is not a priority for this work.

p.6, line 34: Why is the bias related to the airborne campaign rather than perhaps a bad dh/dt correction?

The dh/dt correction at Dome C is $<$ 1 cm/yr, and so we think it is unlikely that it would be in error by this magnitude. We have clarified in the text that the magnitude of the correction is only small. However, we do also acknowledge that there may be some small contribution from the dh/dt correction and so we have also added text to cover this point.

p.6, lines 31–32: The method presented does not derive a surface slope (but rather a linear slope of elevation residuals), so where is this coming from?

Yes, thanks, we agree that this was mistaken wording here and have corrected it in the revised manuscript.

p.19: It should be clarified that the footprint diameter is the beam-limited footprint. I would like to see when the rest of the satellites are planned for launch in this table as well.

As requested, we have clarified that this refers to the beam-limited footprint diameter, and we have also added the planned launch dates for Sentinel-3C and Sentinel-3D.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2018-120, 2018.