Referee comment on "A fine-scale digital elevation model of Antarctica derived from ICESat-2" by Xiaoyi Shen et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2021-204-RC2, 2021

Review:
A fine-scale digital elevation model of Antarctica derived from ICESat-2
Shen et al. 2021

The study presents a new elevation model of Antarctica based on one-year of ICESat-2 observations. The authors provide a specific time-stamped DEM with a final pixel size of 250m, following the same approach as presented by Slater et al. (2017). The new DEM is validated against OIB data and compared to existing Antarctic DEMs. Results show an improved accuracy compared to DEMs based on Radar altimetry but with less accuracy than DEMs based on Radar interferometry or Stereo-Photogrammetry.

In general, it is an interesting project and worth to be published as ICESat2 provides precise point information with high accuracy and good coverage. This large database should be used to generate a gridded data product of high quality which is easily accessible and to be used in different applications. The authors did this approach; however, I do have some concerns and questions related to the method, the validation and comparison with existing DEMs.

General comments:
The paper is well written, however in some instances the statements are not accurate (see below). Figures are mostly ok with room for improvements (see below).

Structure is fine and easy to follow.

I have some concerns about the selected 250m posting and the fitting method used for 1 year of data.

Major question marks arise when looking at figure 3b illustrating the difference between the three-postings used to generate the final DEM.

I have some major concerns about the method used to validate the new DEM and the way
to compare to existing DEMs.

Specific comments and questions

Section 2.1

Please state which processing version was used. Do you apply any additional filtering prior fitting the data than the at06_quality_summary?

The spatial resolution of 20m is true for the along track sampling. However, the track spacing is latitude dependent and I doubt that a 250m spatial resolution across track is reached for lower latitudes. Please be correct here, and add a figure showing the latitude dependency of the track spacing. This is important as one of your arguments is the dense spatial coverage and I would like to know if 250m is a reliable grid size. As seen later in the manuscript most of the grid cells used are of coarser resolution.

Section 2.2

L84 ATM L2 Icessn?

OIB Slope and roughness is mentioned two times and plotted in Fig 1. I don’t see where this information is used in the paper and how the histograms add information. I suggest to remove b,c,d in the Figure. In the text it is also not clear what filtering is exactly applied (Line 85).

Labels in Fig 1 b,c,d are too small.

Section 2.3

It is not correct that the Altimetric based DEMs have no specific time stamp. Slater DEM corrects for elevation change and though the DEM effective time is 1\textsuperscript{st} July 2013. The Helm DEM is July 1\textsuperscript{st} 2012. One cycle (369 days) was used to generate this DEM. IceSat/ERS1 DEM has also a clear timestamp as elevation change was taken into account. This means one of your arguments doesn’t hold and you need to check the whole paper where this misleading information is given (it is stated already in the abstract).

Additional in Table1: REMA and TDX DEM: Instead of marking time stamp unclear you should give the acquisition time of the sensors. If I remember correctly TDX was in Antarctic acquisition mode for two seasons (2013/2014 and to fill gaps 2016/2017). REMA used data mostly from the 2015 and 2016 austral summer seasons. Please carefully check the relevant publications of Wessel (2021) and Howat (2019).

Section 2.4

L128 elevation gasps

I don’t understand the labeling of Figure 2 (The numbers of ICESat-2 measurement points in grid cells at 250 m and 500 m are both resampled to a resolution of 1 km). Is this figure showing the 1km coverage or not? The final DEM is 250m so why do you show the 1km coverage?

Why don’t you show the coverage for all three resolutions (e.g. panels a), b), c)) and include in the figure label how many of the grid cells have data coverage (e.g 1km 74 %, 500m 70 %, 250m 73 %).
Doing so it would be much easier to evaluate if a 250m DEM posting makes sense. In your case you only have 26% coverage, so most of the 250m DEM is based on resampled coarser DEMs and interpolation. Therefore I think, starting from 500m, refill with 1km and do the kriging would make more sense. Later one can resample to whatever posting is needed. Doing so, your fitting should be more robust as more points are used per pixel. As comparison, Slater stated a 60% coverage for his finest 1km grid, which seems to me more reasonable.

In addition, I have some concerns if one year of data is enough to estimate a reliable elevation change which is internally used to reference your DEM to a specific time stamp. For a six-year time series as in Slater et.al. this makes sense but for one year I don’t see the point. Could you please provide the elevation change product (a5 parameter of your fit) to see if the method makes sense?

Otherwise, one could remove a5 from equation 1 and make the fit more robust.

Could you please explain why you chose the fitting method? To my opinion the fitting method forces a quadratic surface in each grid cell (but mostly the real surface is not quadratic – sastrugis, small scale undulations etc.). The fitting method is minimizing the advantage of ICESat2. As the accuracy of each single IceSat-2 measurement is very high and the footprint small the quality of the input data is very high (compared to Radar altimetry). So why not make use of all valid measurements by taking all data and run the kriging interpolation (or whatever you prefer) – similar to the Helm or Bamber DEM approach?

Could you please spent more explanation or any equation of how your uncertainty map was derived. What is the 95% confidence level for elevation estimation and how exactly is this derived?

What kriging method did you apply (model, nugget, sill, radius). How is the variance error calculated. Which software was used for kriging)

Fig3: I don’t understand the large differences between the DEMs of different resolution. How can a 100-300m offset be explained? Two options: Your method isn’t working or the evaluation as shown in Fig3b makes no sense. It would be better to show Antarctic wide difference plots of (DEM_250m – resampled DEM_500m and DEM_250 – resampled DEM_1km) and the corresponding histogram and statistics.

Section 2.4.3

Why don’t you resample the DEM to the OIB data locations and calculate the difference and its statistics? OIB is your reference elevation and you shouldn’t replace is by a median. By calculating a median for each grid cell, you assume the surface in the grid cell is flat. In the interpolation you assumed a quadratic surface.

Table3: Why are ice shelves less accurate? Do you have any explanation? These are very flat areas and your argument that the DEM accuracy is better in flat areas doesn’t hold anymore. Did you apply a tide correction for the OIB data?

Section 3

Figure 6b and 6d. It seems that the DEM has artefacts, especially over the ice shelf. What is the reason for those artefacts? (Fitting routine or kriging artefacts or IceSat2 data problems?)
Do you observe similar artefacts in other areas as well? Could you please zoom into 6b and compare this region to hill shades of the other DEMs?

What is the reason for plotting the grounding line? To me it is an unnecessary information.

Figure 7c: too small and low resolution. The color of 500m and 1km grid cells cannot be distinguished (I’m red green blind). Colors of 7a and 7c should be the same.

Figure 8: Please chose a different color scale (e.g., red-white-blue like in Figure 9)

Section 4

Comparison to other DEMs and DEM comparison to OIB data is difficult to evaluate. The problem is the different time stamp of the DEMs. As you use for each DEM different OIB data which are not in the same area one can’t compare the results. The table shows clearly different numbers of grid cells, so the results can’t be compared.

This means that your arguments that the new DEM shows a better accuracy cannot derived from the applied analysis (I don’t doubt that this is not the case but the analysis is inadequate to show this.)

Furthermore, your suggested approach to calculate a median OIB elevation for each grid cell will certainly influence the results as the DEMs have different pixel spacing (again resample the DEM to the OIB location).

A valid approach would be to choose OIB data in areas of low elevation change. This would enable you to take the whole OIB data set and compare the chosen data to all DEMs.

A similar comparison you applied in Table 6. However, the number of grid cells are not the same. Again, you can’t compare the results.

In addition, for comparison one could use published GPS transects like:


It would be also worth to show an OIB elevation profile and the differences of the different DEMs to this OIB profile (similar to Fig 9 in Slater et.al.) This could nicely show the performance of the new ICESAT-2 DEM.

Resolution of e.g. Fig 9 is too low

Section 5

I miss a clear statement, why this DEM is needed. REMA and TDX seem to outperform the new ICESat2-DEM by a factor of 2. Furthermore your uncertainty map shows values of < 1m, however the standard deviation of the differences to OIB shows 8m indicating that the uncertainty map is not representing this. Please discuss this.