The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2020-323-RC2, 2021

The Cryosphere
Discussions
Open Access

Interactive comment on "High-resolution topography of the Antarctic Peninsula combining TanDEM-X DEM and REMA mosaic" *by* Yuting Dong et al.

Romain Hugonnet (Referee)

romain.hugonnet@legos.obs-mip.fr

Received and published: 10 January 2021

1 General comments

The paper is well-prepared and the authors make great effort to present their study rigorously. The text is generally well-written and the quality of Figures is good. The introduction accurately paints the context and limitations for Digital Elevation Models and related studies of the Antarctic Peninsula. The methodology presented for height correction is, to my knowledge, novel and its implementation is sound. The validation effort of the results with high-precision data is valuable. Finally, the resulting impacts

Printer-friendly version

Discussion paper

are significant and clearly highlighted in the conclusions.

However, I have several major concerns. In decreasing order of importance, those are:

1. the general logic when combining the TDM and REMA DEMs,
2. the relation to existing methodologies,
3. the limits of the validation with IceBridge and ICESat-2,
4. the statistical estimators used in the study.

Those are detailed below.

2 Specific comments

2.1 Combining TDM and REMA DEMs

With 2 DEMs covering approximately all the AP, one can ask himself: what is the best reference DEM to use? What I draw from the authors presentation is that they decided to correct TDM DEM with REMA DEM due to:
A) the short timespans of acquisitons of TanDEM-X (2013-2014) while REMA is based on WorldView acquisitions spanning a longer period of 2009-2017. This shorter timespan of TDM leads to less issues with glacier elevation change over time on the entire AP.
B) less data gaps in TanDEM-X compared to REMA.
These factors that motivated their choice are somewhat “implicit” along the flow of paper, except in the abstract. I think the authors should clearly state all their arguments at once at some point of the main text to make the entire reasoning behind their choice understandable (at the end of introduction? Or the start of methods?)

Additionally, I think that this choice is still subject to some discussion. And this for several reasons:

- firstly, I do not find the argument “less data gap in TDM DEM” completely valid. As the authors show, TDM DEM contains many height errors, or artefacts, that require correction. Those could be interpreted as data gaps as well. It is not clear how much surface area is affected, but the authors should have quantified this after application of their methodology. Many of REMA data gaps likely originate from photogrammetric blunders, very much like TDM height errors originate from interferometric ones. This would partly invalidate argument B) for selecting TDM.
- secondly, the REMA DEM is a mosaic based on WorldView 2 m and 8 m strips, freely available through the Polar Geospatial Center. I have not checked, but it is possible that most of the AP is covered by WorldView acquisition within a 2/3-year timespan of each other instead of suffering from the full 2009-2017 deviation. This would partly invalidate argument A) for selecting TDM.
- finally, the validation effort shows REMA DEM to consistenly have higher vertical precision than the TDM DEM (Table 1, 2, 3: value of 90

I understand that this choice is complex and that the methodology developed by the authors is already dependent on the type of blunders present in TDM DEM which might not be reapplicable to REMA with its own blunders and larger-scale data gaps. Therefore, although this choice is directionally important for the study, it certainly remains fully with the authors. A correction based on TDM DEM is undeniably valuable. In any case, the authors should provide:

- a Figure (maybe in the Supplementary) showing the surface area coverage over time of available REMA DEM strips for the AP. If many strips are closely available in time with good coverage of the AP (>80%), this really poses the question of correcting TDM blunders instead of correcting/gap-filling REMA.

- a Table with coverage statistics + mean vertical correction for identified blunders at each multi-scale step of the methodology somewhere in the main manuscript.
- a discussion on the influence of the 8% data gaps in REMA on the correction of TDM DEM should be provided. I acknowledge that this effect is unavoidable, yet adding a paragraph quantifying the possibly omitted TDM blunders in these 8% REMA data gaps would be useful (maybe based on the average surface area of blunders found on the rest of the AP?). Providing a map of the areas possibly affected (where there is no REMA coverage, and possibly high slopes/larger errors in the TDM Height Error Map?) would also be valuable for future users.

Optionally, the authors should consider using individual strips instead of the REMA mosaic product for height error correction. This additional effort could significantly reduce temporal biases related to glacier elevation change by selecting strips closest to 2013-2014 when available (e.g. biases shown on Figure 7f). The authors would however need to co-register those strips individually to TanDEM-X before using them for corrections.

2.2 Relation to existing methodologies

The paper surprisingly lacks references to previous DEM correction/fusion methodologies. This is true both for the introduction (supposed to explain the context of existing methods and why a new one is necessary), for the methods (supposed to reference/compare to existing methods, if applicable) and discussion (based on the results, what are the benefits of using this specific method compared to others? qualitatively at least. quantitatively would be even better e.g. by comparing with other methods locally).

Many studies have looked at merging DEMs, removing data gaps and improving general DEM quality, for example: Reuter et al. (2007), Papasaika et al. (2009)

(full thesis here: https://ethz.ch/content/dam/ethz/special-interest/baug/igp/igp-dam/documents/PhD_Theses/109.pdf), Yamazaki et al. (2017), etc.

The authors should:

- provide a scientific context referencing existing methods and justifying a new methodology,
- identify and cite possible similar existing methodologies, if applicable.

2.3 Validation with ICESat-2 and IceBridge

Seasonal and temporal biases of validation exist but are omitted in the validation methods and its discussion. Those should be quantified and discussed.

For temporal biases, the authors could use low-resolution, large-scale elevation change maps (Smith et al. (2020)) to partition their validation data over the AP. Binning the validation points by category of expected elevation change during the period (e.g., near stable; <0.2 m yr-1, small elevation change >0.2 and <0.5, strong elevation change >0.5) could provide improved statistics to evaluate the results through the validation effort. The impact of seasonal biases in elevation changes between the two validation datasets should also be discussed using known estimates of seasonal cycle in the AP.

2.4 Statistical estimators

Along the study, the authors provide Tables of the same format to quantify the improvement brought out by their methodology (Table 1, 2, 3 S2, S3). They use the mean of elevation difference to the validation data, the root mean square of differences and the 90% quantile of the distribution of elevation differences.

Currently, the mean does not bring much information and is hard to interpret due to temporal biases (preceding comments), but also because this statistical estimator is not very robust to outliers. I suggest using the median, as well as binning by category of expected elevation change during the period (preceding comments). Possibly, a Table showing median residuals normalized by the time difference between the validation dataset and the TanDEM-X date (1-2 years for LVIS 2015, 5-6 years for ATL06 2019) would allow for a better comparison between the biases identified in the two datasets.

The RMSE is generally a good metric, it is however overly sensitive to outliers which is exactly what the TDM DEM is here suffering from. Using this estimator might "oversell" the improvement in the results, especially to the reader unfamiliar with these effects. Consider:

- splitting your statistics by category of initial height differences, or showing the statistics independently for the corrected regions (before and after) and the untouched ones (once). I feel the second choice would be preferable.
- using the Mean Average Error (MAE), less sensitive to outliers.

For the 90% quantile: I see this is currently the raw quantile of the distribution (Table 1, line 3, the value is negative). I imagine that the authors want to show a measure of distribution spread (estimate of elevation precision). For this, consider either:

- taking the 90% quantile of the absolute value of the elevation difference.
- calculating the half-width between the 5% and 95% quantiles of the distribution.

3 Technical corrections

General on the text:
Unless the authors can justify a specific reason, I advocate for the two following changes along the manuscript:

1. Pick “elevation” or “height”, don’t use both interchangeably. Example in the caption of Table 2: Height differences calculated as DEM elevation minus laser height. This is confusing. I suggest using “elevation” everywhere, as it is the term most commonly used (Digital Elevation Model, Reference Elevation Model of Antarctica, etc).

2. The use of “height error” seems questionable to me. “Error” by itself is not precise enough as it can refer to random errors, (i.e. uncertainties) or refer to systematic errors (i.e. biases). Here, the artefacts in the TDM DEM and the related correction methods developed by the authors fit clearly in the box of systematic errors. Thus, it seems to me that it would be clearer to use “elevation bias” instead of “height error” along the text.

Text line by line:
75-96: Mention % value of data gaps in TDX
81: high-precision
88: stereophotogrammetry
89: such as the Advanced Spaceborne...
98-112: Please specify the REMA release used (is it r1 or r1.1?)
114-123: Mention exact acquisition date for the LVIS data (season)
136: “the” buffer zone: This is the first mention of such a zone. Change to “a buffer zone” and refer to the related section.
140-178: All this information is not really specific to the paper Methods: shorten +

optionally, move to “Data”?
204: To my knowledge, “path propagation algorithm” is not a nomenclature commonly used for this kind of method. This type of “flood-fill” method (https://en.wikipedia.org/wiki/Flood_fill) for region extraction is generally called “region extraction”, “blob extraction” or most generally connected-component labeling (CCL) (https://en.wikipedia.org/wiki/Connected-component_labeling). Please adapt the nomenclature, and cite a reference for the algorithm used if applicable, and also possibly its relation to an existing computing package/parallel implementation.
252: increases
253: increases
254: high-precision
255: multi-scale (for consistency)
258: “Their spatial extent increased from. . . to. . .”
264: Figure 6b
271: Why use the “geographically closest point” in the DEM instead of a bilinear interpolation to the center of the LVIS/ATL06 point? With the TDM DEM at a posting of 12 m, the potential 6 m horizontal bias using a “nearest neighbour” approach from the center of the point can lead to a 3 m vertical error on a 25° slope (50% slope), and higher for larger slopes. This procedure might be deteriorating the quality of the validation effort, consider switching to bilinear interpolation of the raster data.
321: ATL06
333: elimination of the residual height errors
341: Refer to Figure 10 before mentioning profiles L3/L4
371: The vertical scales should be specified sooner, in the Methods section maybe?

On the Figures and Tables:
Fig. 2:
Axis labels of histogram are too small in I. Text of schematics is squeezed vertically in III.

Fig. 5:
Specify if this is a schematic was created for demonstration purposes, or from real AP data at a specific transect.
Fig. 6, 7:
Add glacier outlines.

4 References for the review

Yamazaki, D. et al. A high-accuracy map of global terrain elevations: Accurate Global Terrain Elevation map. Geophys. Res. Lett. 44, 5844–5853 (2017)

Papasaika, H., Poli, D. Baltsavias, E. Fusion of Digital Elevation Models from Various Data Sources. in 2009 International Conference on Advanced Geographic Information Systems Web Services 117–122 (2009).

Reuter, H. I., Nelson, A. Jarvis, A. An evaluation of voidâĂŘfilling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci. 21, 983–1008 (2007)

Smith, B. et al. Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes. Science 368, 1239–1242 (2020)

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2020-323, 2020.