We thank reviewer 1 for his/her suggestions and comments. Original comments are given in bold below, followed by our responses.

I was unable to find the image mosaics and associated metadata on bprc.osu.edu/GDG/data.php, and so I cannot comment on their quality. This hinders the review process.

Due to a linking error on the project webpage, the image mosaics were not visible. We didn't realize the problem until we received this review.

Perhaps the main issue is that there is a 0.5 pixel mismatch between the edges of the tiles of masks versus the DEM and hillshade tiles i.e. the grids are misaligned. This sort of drift commonly occurs when reprojecting or interpolating – the grids should be snapped to exactly the same grid.

All image tiles are the same corner coordinates (and are thus on the same grid), however, the DEM has a 30 m resolution whereas the image mosaic and masks have a 15 m resolution. Therefore the center coordinates of each 15 m pixel will be offset ½ a pixel relative to the 30 m DEM. For example, the upper left coordinates of tile 1-2 are:
[-390700, -2005550] for all data sets. The upper left pixel for the DEM will then be centered at:
[(-390700 + 30) ,(-2005550 - 30)] = [-390655 -2005595] and for the 15 mask and image mosaics:
[-390662.50, -2005587.50].
To "snap" these to the same grid as suggested would require upsampling the DEM or downsampling the mask & mosaic.

In the DEM tiles, there is a 1-pixel edge artefact where the surface abruptly steps up in height e.g. the sea goes from 30m to~37m andbackdown to 30m e.g. in tile4-4 at 1171614 m north.

This has been corrected for the next version release.

Of lesser concern: there is a discretisation issue in open ice-sheet areas where the surface height steps up and down in 1 m increments in oddly geometric (i.e. non-physical) patterns (e.g. at 452778m, -1570164 m).

This is simply a result of the integer precision of the DEM areas of very low relief. Note the variations of +/- 1 m are well within the error of the DEM (which is one reason why they are distributed in integer precision). The "non-physical" patterns referred too are visible in the hillshade images over areas of low relief because of the very low grayscale stretch - i.e. such small variations would not be visible in areas where slopes exceeded 1 m/pixel.

There are isolated pit artefacts (e.g. at 466229, -826599 in tile 4-5) but these appear very rare.

Editing of the existing DEM is ongoing and this and other artifacts will be removed in the next release.

The hillshade product has speckle (that seems not always to reflect speckle in the DEM) and a 2-pixel blank gap between tiles but this is of little concern as this is a display product (rather than e.g. a modelling product).

As mentioned, the hillshade products are for display purposes only and are not a direct indication of DEM quality. The hillshade algorithm, which relies on slopes and aspects, results in the mentioned edge effect.

Section 3 – image mosaic
The USGS L1T images are (presumably) orthorectified with a different, older DEM (not the one presented here) which presumably means that there are substantial mis-matches between this new DEM and the image mosaic where the new DEM captures relief in higher resolution than the old. Presumably this is not captured in the 'image geo-registration error' in the metadata, as it would be unknown to USGS. The obvious solution is to orthorectify the images to the new DEM rather than the old. It may be that this is such a substantial task that it is beyond this project – is this so? What are the implications for the image mosaic?

Since the primary motivation for creating the image mosaic was for extracting the ice and ocean masks, high-precision re-orthorectification of the landsat imagery using the GIMP DEM was beyond the scope of this study. Further, since the DEM and the imagery are not coincident in time, thinning of the ice sheet, particularly near the margins, would result in some amount of orthorectification error, so it's not clear how much improvement this would make for our purposes. However, improved orthorectification of imagery is one of the primary motivations for the DEM and will be undertaken in future studies.
As explained in the text, error estimates, based on the residual in of the orthorectified imagery and predefined control points, are extracted from the USGS metadata and provided with the image mosaics.

Page 458, line 11 – what is the evidence for the assertion that the Radarsat images have smaller errors?

This has been amended to ". As with Landsat, the primary source of geolocation error in the RADARSAT imagery is error in the DEM used for terrain correction and are similar in magnitude to the Landsat mosaic (Moon and Joughin , 2008)."

Manual digitisation of the entire coast at high resolution is an impressive effort! The authors describe 3 error sources but for source 2 (image geo-registration), what about the error contribution from using orthorectified images that used the old DEM (assuming this was done)? Perhaps of similar order to the errors quoted?

From the text: "All error sources are expected to vary randomly in space, although there is likely a systematic component of error source (2) over distances equivalent to the size of a single image (e.g. 185 km for Landsat 7) due to errors in the registration model used to orthorectify the image, which typically is on the order of ±5 m, or 1/3 of a pixel for L1T-processed imagery."

Page 459, line 23: "Due to the failure of stereo-photogrammetric methods for Digital Elevation Model (DEM) extraction on featureless ice and snow surfaces, the difficult logistics involved in aerial LiDAR surveying, and the latitudinal limit of the Shuttle Radar Topography Mission, the coverage and accuracy of elevation data for polar regions are poor, especially over the interiors of ice sheets." - This simply isn't true. What about ERS-1, ERS-2, Envisat, ICESat and Cryosat2???

Amended to: "....and the latitudinal limit of the Shuttle Radar Topography Mission, DEM coverage and spatial resolution for polar regions are poor, especially over the interiors of ice sheets."

Page 460, line 3: the radar altimeter used by Bamber et al. would not have been a SAR altimeter.

Corrected to: "....created from a combination of satellite radar altimeter and …."

Which ICESat product - GLA12?

We now specify the GLA12 product in the text.

Is Shepherd et al., 2012 the original reference for the filtering technique? (unlikely)

hepherd et al., 2012 is the most detailed reference for this filtering technique and we now specify that the description on the technique is in the supporting online material for that paper.

Page 460, line 26: the variations are not month-to- month, they are intercampaign.

We now describe the biases as campaign-to-campaign .

The Borsa reference focuses on one aspect of the intercampaign bias. From this paragraph, it is not clear that the authors have corrected the intercampaign biases or applied other corrections in an appropriate way. More and better detail is needed here.

The Borsa reference explores the gaussian-centroid bias in the release 633 data, the only aspect of the inter-campaign bias that is well understood, and quantifies the remaining bias from other, unknown sources. It also gives an explanation for why correcting for sea-surface bias removes almost all of the gain-centroid bias. We hope that in the long run, the Borsa et al. paper is not the final word on biases in ICESat data, but for now it is the best available reference. Detailed, original exploration of this topic is beyond the scope of our paper.

This section estimates error based on the mismatch of the DEM to ICESat reference heights. This is useful but does not constitute a thorough error budget. It doesn't, for example, include the errors in the ICESat heights themselves. Also, the mismatch statistics are given as RMS but this masks any potential systematic bias in the DEM vs ICESat heights. It may be that the process of fitting the DEM to the ICESat point cloud results in a zero-mean difference in heights but this isn't shown. The question is: are the validation errors normally distributed around zero (no bias) or not? I suggest calculating the mean differences as well as the RMS, and also plotting the distribution of the differences to demonstrate normal distribution.

We have added the following to section 5.1 (regarding ICESat uncertainty):
"Elevations calculated in this way should be accurate to better than 0.1 m, or two orders of magnitude smaller than the expected DEM uncertainty."

We have added the following text to section 5.2:
"The re-gridded data were then co-registered to the ICESat GLAS point cloud using an iterative, 3-D conformal transformation (Noh and Howat, 2013). This procedure results in residuals between the DEM surface and ICESat point cloud with a normal distribution and a mean of zero."

Also, the authors assert that the validation error is dominated by slope – I suggest calculating the validation error by slope class (e.g. 0-1 degree, 1-2 or similar). Plot the distributions for these classes to look for bias and non-normal distribution. This error-by-slope-class would allow users to calculate the error anywhere in the data simply from the slope.

We have completed this analysis and include a new plot (new Fig. 7) showing slope versus error and bias, and these results are discussed in section 5.7

Furthermore, I suggest producing a map from the actual validation errors so that the spatial distribution of both the ICESat validation data and the error magnitude are clearly shown. This would have the advantage of making clear where temporal changes in height have led to large apparent validation errors.

This map is now included as Fig. 7 and discussed in the text in section 5.7.

Section 6 Conclusions

These aren't really conclusions, they're pointers to the datasets and an indication of future work.

We have reorganized the conclusions and added a sentence summarizing the objective of the paper.

Figure 2 – this figure is too small and dark to make the point that the authors want it to make.

We have made the figure brighter and increased the rendering resolution.

In various places: some typos and grammatical errors that should be obvious in proof reading.

We have completed another careful proof reading.