Reponse to Reviewer 3

In this paper, McNabb et al. investigate the effect of missing data (called voids in the article) on the glacier volume change that can be obtained from digital elevation models (DEMs) differencing. The methodology is rather straightforward, they differentiate two DEMs acquired over Southeast Alaska that (almost) cover the entire glacierized area. These data are used as reference data. Then, they artificially generate voids in the data and evaluate the impact of different void-filling/interpolation methods on the on the regional glacier volume change estimate, but also for each individual glacier. They investigate 11 different void-filling/interpolation methods that are often used in the literature, providing a unique and comprehensive assessment. They conclude that most interpolation methods introduce very little bias (<1%) on the regional glacier volume change. However, individual glacier volume change estimates can be severely affected by the choice of the interpolation strategy.

This paper is rather narrow-focused, but its scope fits very well within The Cryosphere, where it will certainly reach an adequate audience. The topic is timely and very relevant, as the geodetic method is more and more widespread in glaciology. To my opinion, this paper has the potential to become a classic paper in the field of geodetic mass balance. However, and while I appreciate the concision of the paper, I have the feeling that the authors could discuss some aspects more in depth. Moreover, I sometimes had a hard time following the paper and found that it lacks clarity in its current form. These are my two major comments.

We would like to thank the referee for their careful and constructive comments that have helped to improve both the clarity and focus of the manuscript, and we hope our revisions will be satisfactory. Our responses to the comments below are in blue, with the original comments in black.

1 Major comments

1. **Volume change vs. geodetic mass balance**
   The title of the paper mention the sensitivity of geodetic glacier mass balance, but actually discuss only glacier volume changes. This decision is somehow understandable, because it is the quantity that is directly affected by the void-filling strategy. However, the impact of void-filling strategies on the individual glacier volume change expressed in km^3 is not very intuitive, and not as informative as it could be. First of all, the IfSAR DEM was acquired over two years and it would be better to present the annual mean instead of the totals, in order to get rid of this temporal inconsistency. Second, the results are largely dependent on the glacier area considered, larger glaciers being more sensitive to the interpolation (P8L12-13), mostly because for a similar elevation change they have larger volume change, due to their larger area. For example, for figures 7 and 8 (and 5?), I suggest to present the results in kg m^{-2} a^{-1} or in m a^{-1} (if the authors do not want to make any density assumption).

   In accordance with this comment, as well as the comments from another reviewer, we have rewritten the analysis using units of m a^{-1}, that is, the area-averaged rate of volume change.

   If the authors want their study to be reproduced and the conclusions of this article to be applied elsewhere, they need to analyze the influence of the void-filling strategy more in depth. I feel like the paper misses some basic, yet interesting analysis. For instance, what is the influence of the percentage of voids for individual glaciers? Of the glacier-wide mass balance/mean rate of elevation change? Of the glacier area?

The authors probably analyzed these influences already and found that they were limited/not interesting, but I think it is probably worth mentioning them, in order to apply their conclusions to a different setting.

We have significantly expanded the discussion on the impact of voids for individual glaciers and methods in section 4.5, in particular adding Figures 10, 11, and 12, and focusing on the percentage of voids where the 'best' methods start to give unreliable results. We hope that this helps to increase the impact and depth of the analysis.

2. **Some clarifications needed**
The objective of the study is quite straightforward, but a number of confusions and unclear statements prevent from an easy understanding of the paper. I had to go back and forth a number of time reading the paper, and I have the feeling that the clarity of the paper could be much improved if the authors address the three comments below. First, the author mimic the voids of a standard DEM difference, based on ASTER correlation map patterns. Consequently, I expected that they would investigate the influence of the void-filling strategy for this purpose. However, they also investigate such methods as the global ones, which are generally used for regionalization of Lidar surveys. They should make a clear distinction between these two applications when relevant. In other words, I do not think that the global methods are relevant for DEM differences void filling at the scale of individual glaciers. Correct me if I'm wrong, but I do not know any paper which studied individual glacier mass balances obtained with such global methods to fill in the holes of a DEM difference.

   You are correct in saying that the global methods have not really been used to estimate individual glacier mass balances in prior studies. Our goal with including these methods was to investigate whether useful results could be obtained for an individual glacier using global methods, as well as to compare the regional estimates obtained using a variety of methods. In the updated text, we have tried to make this distinction more clear, in particular at P6L20-24.

   Second, the different methods described are relatively basic, however their description should be clearer. For instance, adding equations to the description of each method would be beneficial. Alternatively, you could share the code you wrote, which would also support your conclusion in which you encourage others to test different methods when dealing with voided data.

   We have added a statement in both the Methods section, and the Code Availability about where the scripts used to generate and interpolate the voids will be available.

   Third, I found the example about Taku Glacier extremely confusing. If I understand correctly figure 2, the tongue of Taku Glacier is mostly free of voids. However, the global methods (panels e to g) totally change the pattern of areas where data are available! Consequently, the methods should be described as Interpolation and not Void-filling (for instance the title of section 3.2 should be changed), because they also apply to areas without voids (and it is technically a non-exact interpolation method). If I did not understand correctly figure 2, you can ignore this comment, but you should consider changing figure 2.

   You are correct that the tongue of Taku Glacier is mostly free of voids; you are also correct that these methods would be better described as 'void interpolation' rather than 'void filling', given that we aren't actually 'filling' the voids, per se. We have updated the text to reflect this.

2 Specific comments

- Fig. 2: confusion between the glacier and voids outlines. You should draw the glacier outlines in a different color/line thickness, such as the panel 3 is easier to understand.

  We have changed the color of the voids in Fig. 2 in order to avoid this confusion.

- Fig. 5: add a scale/grid on the inset.

  The updated figure no longer has an inset.

- Fig. 6: this figure is extremely confusing to me. First of all, I'm missing the location of the voids on panel a, and I had to go back on the previous figures to understand where the voids where. Then the figure shows the strong impact of the void-filling procedure on the tongue of Taku Glacier, but from what I understood of figure 2, there were no voids on the tongue. This comment is in line with my major comment 2.

  See response to your major comment 2 above. We have shown the void locations in the updated Figure, panel a, to help make this figure more clear.

- Fig. 7: I think here the reader loses the information about the difference to truth relative to the total glacier volume change. I guess the larger the volume, the larger the error.

  The updated figure shows the differences averaged by glacier area, rather than the volume change in km^3, in accordance with other comments.

- P1L3 and P1L17: I do not fully agree with your definition of the glaciological method, which does not really monitor changes in surface height (that would actually be the geodetic method). The glaciological method directly measures the surface accumulation and melt.

  Changed to read, "Traditional estimates of glacier mass balance have involved *in-situ* seasonal or annual measurement of accumulation and ablation at select locations, and extrapolation of these sparse measurements to the entire glacier..."

- P1L8-9: add a word about the artificially generated voids

  Done.

- P1L11: define ASTER

  Done.

- P2L6: Digital Elevation Models-¿ DEMs

  Done.

- P2L33-P3L6: this paragraph is not really well structured in my opinion. You should describe more clearly the philosophy of your study. I suggest to move completely your warning statement about the radar DEM difference (P3L1-2) to the other place where it is mentioned (P4L19-24). You need to add something about the artificially generated voids and to better justify the choice of ASTER, among a large choice of (optical) sensors.

  We have moved the warning statement to the end of the paragraph, and added something about artificially generating the voids used. We feel that the warning is important to have in the introduction, as well as later on in the paper. We don't feel that the choice of ASTER needs justification here, as it is a widely-used sensor in glacier studies with a long (nearly 20-year) record.

- P3L9-14: what is the glacierized area?

  $\sim$5900 km^2, added to text.

- P3L19-20: provide references about studies that used SRTM to estimate geodetic glacier mass balance

  Done.

- P3L29: what proportion of the glacierized are if affected by these small voids?

  Fewer than 2.5% of the glacierized area; we have added this information here.

- P4L2-6: more references and details are needed in this paragraph. What is the precision of the IfSAR DEM? What is the proportion of voids? Has it been used in other glaciological studies?

  We have added more information here. The metadata report an accuracy of ∼1 m, while the USGS-reported accuracy is 3 m. Voids in the original acquisitions are small (<1% according to the metadata), and are filled in post-processing using proprietary algorithms.

- P4L5-6: is this sentence useful?

  Probably not, so we removed it.

- P4L12: can you justify the exclusion of glaciers smaller than 1 km^2 from your analysis?

  We omit these smaller glaciers because errors/inaccuracies in glacier outlines are much larger for smaller glaciers. As our goal is to investigate the effects of void interpolation methods on estimated volume changes, it is best to have a larger sample of on-glacier pixels to work with; voids over small glaciers result in more limited data from which to extrapolate. Also, since our objectives are methods oriented, the question about small glaciers being important is not so relevant. For further comparison of results over small glaciers, we would suggest that higher spatial resolution DEMs are required as opposed to the medium resolution DEMs used here. We now clarifiy this in the text.

- Method section: at some point you need to explain how you calculate the regional estimate. Is it the sum of the individual glaciers, or do you consider all the glaciers as a single body of ice? Do you include the glaciers smaller than 1 km^2 in this estimate? If not you might bias it.

  It is the sum of the individual glaciers. We have included this to help clarify this point.

- P5L9: see my major comment 2 -¿ you might want to rename this paragraph Interpolation instead of Void Filling

  Changed to "Void Interpolation"

- Void-filling section: how do you deal with the temporal inconsistency between your two IfSAR DEMs? I guess for most method you interpolate the rate of elevation change and not the elevation change? This should be written clearly. However, this is not possible for the method based on the interpolation of elevation.

  As stated in section 2.2.3, we remove any glacier outlines that fall 10% by area or more in both collection years, to avoid this problem, adjusting the area of the remaining glacier outlines that fall within both DEM acquisition years accordingly.

- P6L4-6: here the reader wonders why using ASTER voids for regionalization applications (i.e., Lidar based studies)? In order to test the influence of regionalization, one could extract elevation changes in your DEM difference along Lidar flight lines... but this would be another paper!

  Indeed, and it has been done by other papers, for example Berthier and others (2010). Here, we are attempting to show that while these methods may not be a useful way of estimating the mass balance of an individual glacier, it can still be useful to estimate the mass balance of a region using data from the glaciers where you do have measurements.

- P7L6: define normalized glacier elevation

  Added "(i.e., the elevation divided by the elevation range)"

- P7L14: consider switching the order between the sections 4.2 and 4.3.

  We prefer the section order as-is.

- P7L15-23: this paragraph is rather disconnected from the rest of the analysis.

  We disagree. This paragraph discusses the variability of elevation changes in the region, which helps to explain some of the differences seen between the methods.

- P7L16: the pattern of elevation change is negative -¿ the phrasing is not clear to me

  Changed to read "elevation changes in the region are negative, especially at lower elevations."

- P7L26-27: the average geometric volume change has probably little influence, contrary to the mean elevation difference.

  Per your and another reviewer's suggestions, we have changed the units from km^3 to $m\,a^{-1}$.

- P8L3-10: can you say a word about the constant methods? And about the 1km neighborhood method?

  Added in the updated text in the following paragraphs at lines .

- P8L4: define RMS. This sentence is not completely clear to me.

  We have removed this paragraph. RMS values are now defined in the text at P8L16.

- P8L13-14: the larger glaciers are more sensitive than the others, because they have larger volume change for a similar elevation change, due to their larger area. This is one on the limitations of your analysis, because you look at volume changes only (see my major comment 1). You should consider extending your analysis to glacier mass balance, or rate of elevation change.

  Please see the updated text, which compares the area-averaged volume changes as suggested.

- P9L6-8: this paragraph is very short, while Fig. 8 is probably a key figure! Could you elaborate a bit? I found an unbalance with the previous paragraph (P8L30-P9L5), which is less important and much longer than this one.

  We have expanded the discussion of this figure, including adding analysis for smaller glaciers as well.

- P9L20: the 1km neighborhood is never mentioned earlier in the section 4.2.

  We now mention it in the paragraph discussing the updated Fig. 8.

- Section 4.3: consider adding a column to Table 1, which summarizes the regional totals for each method. It might also be interesting to discuss the difference between the regional estimates obtained by summing the individual glaciers versus the regional totals obtained by applying each method to the glacierized area considered as a single body of ice.

  Given that the goal is to show how the regional total changes using each method, we feel that the column showing the change is sufficient here. I'm not sure I understand the second point here - the regional estimates are achieved by summing the volume change estimates for each glacier, not by applying the methods as though the total area were a single body of ice.

- Section 4.4: why don't you use the percentage of voids of individual glaciers to study the influence of the percentage of voids on the distance to truth?

  See updated text (now section 4.5), as well as the newly-added figures which discuss this in much greater detail.

- P10L12: give the total percentage of voids for each threshold.

  See the newly-added Figure 10, which shows this.

- P10L28: again I got confused because you look at 91 individual glaciers and then you mention the global mean hypsometric as one of the best performing method... Please clarify.

  The global method uses elevation differences from all of the glaciers in a sampled area and is applied to each glacier based on that glacier's hypsometry. Hence, the global method here is only using the data from these 91 glaciers.

- P10L31: you can mention in the text that outliers are more often located near the voids, which increases their influence in a linear interpolation.

  A good point, which we have included in the text.

- P11L5-7: the order (regional volume change then individual glacier volume) is the opposite from section 4.2 and 4.3.

  Re-written to change the order.

- P11L25-26: you actually did not demonstrate this in your analysis...

  Hopefully, you agree with us that the updated analysis serves to demonstrate this.

References

Berthier, E., E. Schiefer, G. K. C. Clarke, B. Menounos and F. Rémy, 2010. Contribution of Alaskan glaciers to sea-level rise derived from satellite imagery, *Nature Geoscience*, **3**(2), 92–95.