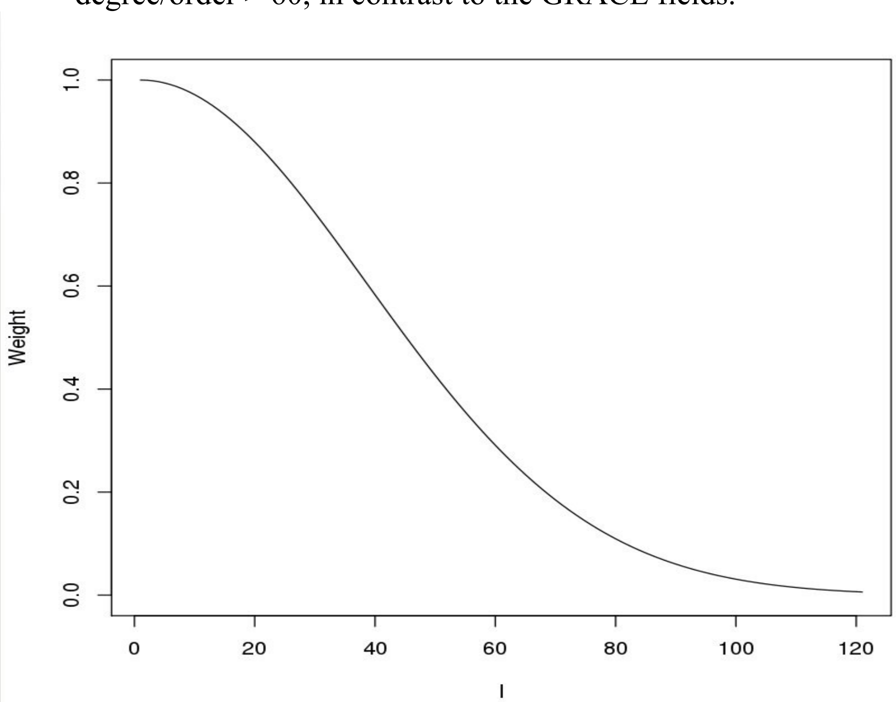
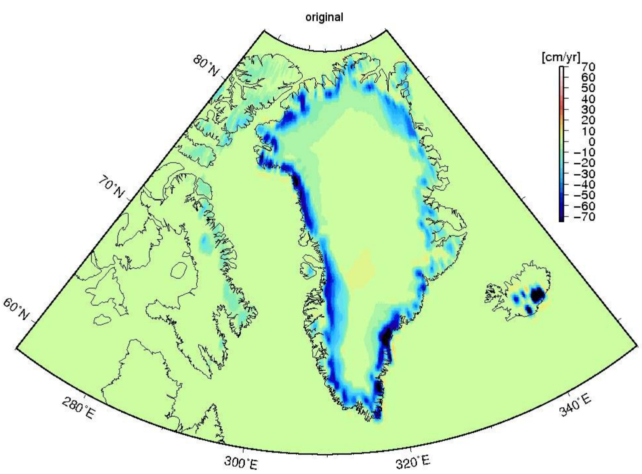
This manuscript builds on previous work of the lead author. In Colgan et al., 2013, an inversion method was presented that attempts to increase the spatial resolution of the GRACE data. Due to inherent characteristics of these data, a strong spatial correlation between adjacent grid points remained, resulting in mass balance fields lacking the spatial details seen in other remote sensing data. This paper uses additional altimetry data to overcome these limitations. Compared to other studies combining these data (e.g., Zammit-Mangion et al, 2013), the method is rather straightforward, but to my knowledge, this is the first time such a combined inversion has been applied to the Greenland ice sheet (potentially because of the challenges of the firn densification problem, which is ignored in this work).

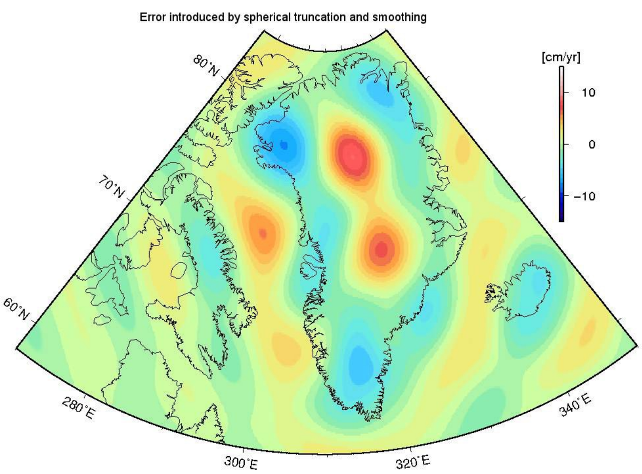
General Comments:

- On p. 551, l. 10-15, the pocket of anomalous positive values on Baffin Island is attributed to the inability of the inversion to satisfy high oceanic mass gain within the parameter space employed. My hypothesis is that this is caused by a more fundamental problem: the incompatibility of the spatial characteristics of the input GRACE data (M^G in eq. 3 in Colgan et al. 2013) and the modelled fields (M^k in eq. 1 in Colgan et al., 2013). If I interpret the method in Colgan et al., 2013 correctly, the GRACE mascons are converted into spherical harmonics, which are cut off at degree and order 60 (and I believe this is a good choice, since in the mascon processing a maximum degree/order 60 is used as well, see Luthcke et al., 2013). No smoothing is applied to the GRACE fields. The model fields consist of 26 km cells, which are smoothed with a 200 km Gaussian filter. This is done in the spatial domain (eq. 2 in Colgan et al. 2013). Below, I have plotted the representation of such a 200 km Gaussian filter in the spherical harmonic domain, as a function of degree *l*. Clearly, the modelled fields contain data of degree/order > 60, in contrast to the GRACE fields.



Cutting of the GRACE harmonics will result in the well-known Gibbs phenomenon (or, 'ringing' effect), since no information is available past degree/order 60, see for example, Swenson et al, 2002. This effect is illustrated in the figure below: the figure on the left shows trends in cumulative surface mass balance of the Greenland ice sheet from a regional climate model, for 2002-2010 (note that this does not include the contribution of glacier discharge and is not meant to represent the actual trends, it's only meant to illustrate my point). I then created two other fields. For the first field, I simply smoothed the SMB field with a 200 km Gaussian filter (=modelled field in Colgan et al., 2013). For the second field, I transformed the map to spherical harmonics and cut them off at degree/order 60 (= GRACE input in the Colgan 2013 method). These spherical harmonics were then transformed back to a spatial field. The figure on the right shows the difference between these two fields. The 'ringing' is clearly visible along the edges of Greenland, but, importantly, also in the interior of the ice sheet. This figure compares well to the bottom right of figure 6 in Colgan et al., 2013. What happens is the following: the modelled fields are smoothed and, in the spherical domain, contain information up to degree/order ~100-120. Therefore, they will not suffer from the ringing effect. Yet, the method will try to minimize the difference between the GRACE field and the model fields and therefore place a mass gain in the cells in the interior of the ice sheet to accommodate the (unphysical) ringing signal in the GRACE field. Furthermore, note that the figure below shows a generally negative signal along the ice sheet periphery. The ringing effect may therefore also bias the estimates of the peripheral glaciers mass balance. The southern tip of Baffin Island is located in a positive 'ring', there, the method will again put a positive mass balance in the cells to fit the GRACE data. Since mass changes are only fitted in cells containing ice, the model cannot reproduce the ringing signal in the ocean regions and only to some extent in the ice-free land regions (due to the constraints), explaining the differences seen between model and observations in fig. 6 of Colgan et al., 2013.





I can think of two solutions for this fundamental problem:

1. smooth the GRACE data with a Gaussian filter, so that the harmonics are not cut off abruptly at degree/order 60. This will require a Gaussian smoother with a radius of ~400-500 km, this way the harmonics with degrees > 60 are given a nearly zero weight by the Gaussian filter. This will probably require to lower the resolution of the 26 km grid cells in the model; or:
2. Convert the model fields to spherical harmonics and cut them off at degree/order 60. Some smoothing may still be required. The same smoothing should be applied to the model fields and the GRACE fields, for consistency.

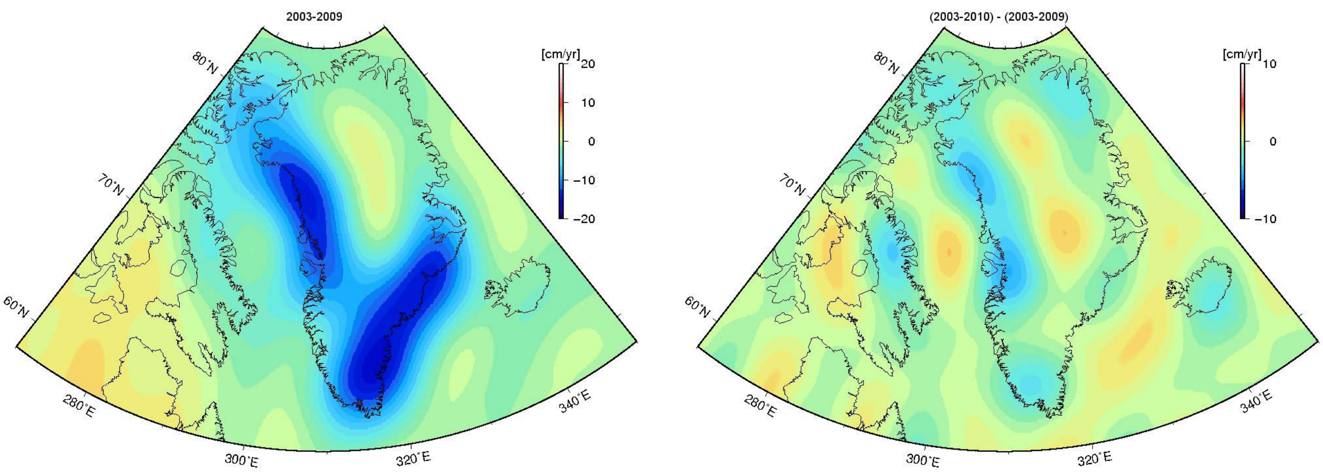
- According to Luthcke et al., 2013, each GRACE mascon covers an area of ~12,390 km^2, or about 111x111 km near the equator. This means that, when using 26 km grid cells, the GRACE data is being oversampled by a factor of ~16 (in fact, the oversampling is probably even worse, as Luthcke et al. 2013 mention that the signal at a particular mascon is affected by mass changes in mascons at distances up to 600 km). In Colgan et al., 2013, this resulted in a strong spatial correlation between adjacent grid points, evident from the 'smoothness' of the mass rate fields (e.g, fig. 11 in Colgan et al., 2013). Here, the authors attempt to overcome this spatial correlation by including altimetry data with a higher spatial resolution. However, as they point out, altimetry provides surface elevation changes. Converting these to mass changes would require a density correction. The authors have opted not to apply such a correction, but I have difficulties following the motivation for this choice, which is only very briefly motivated. On 15-16 on p. 543, the altimetry observations are classified as 'relative trends in cryospheric mass changes'. This implicitly assumes that elevation changes at adjacent nodes are taking place at the same density. Take the hypothetical case where two neighbouring cells are experiencing the same negative elevation rate, and the firn is compacting in one cell, but not in the other. So the effective mass loss in the two cells will be different, but because of its low resolution, GRACE cannot distinguish this difference. Since the elevation rates are identical in the two cells, the algorithm will incorrectly assign an identical mass balance to the cells. Although it is true that there is some spatial correlation in firn fields, the typical length scales are smaller than the GRACE resolution (compare fig. 4a and fig. 8 in Zwally et al., 2011) and this likely will affect the results. Several groups have developed firn correction fields (e.g., Zwally, 2011, JoG; Khan, 2014, Nat. Comm.;...), so such a correction should be relatively easy to implement. Otherwise, the consequences of omitting this correction should be discussed rigorously and accounted for in the error bars.

- A few Operation Icebridge campaigns have been flown over the Canadian Arctic. Given that you do not incorporate these airborne altimetry data in the Canadian Arctic, they offer a great possibility to verify the HIGA mass rates in this region. You will need to make some assumptions concerning density to convert from volume to mass, and the time intervals of the airborne observations will not exactly match, but I recommend looking into this.

- Using the one sigma standard deviation of all elevation rates within the grid cell (p. 543) basically gives an indication of the spatial spread of the observations. Using this method, you are ignoring systematic errors in the ICESat data and the derived elevation changes, which may be spatially correlated. For example, across track slope and dH/dt cannot always be separated accurately from each other, if the sampling of the altimetry tracks is unfavorable . This will lead to a large standard error in the dH/dt estimates, which may be non-random in the 26 km boxes.

  Furthermore, the proposed method (and error bars) doesn't address the fact that ICESat is undersampling outlet glaciers . Hurkmans et al, 2012, JGR have shown that volume loss may be underestimated by up to 20% by ICESat in the Jakobshavn Isbræ area and the same likely applies to other large marine terminating glaciers. I believe this will affect the HIGA mass rates: in such regions ICESat underestimates the surface lowering, but GRACE still senses the associated mass loss, although at a much lower resolution. As a result, the altimetry doesn't 'guide' (cft. p. 543, l. 21) the mass loss to the correct locations . The mass balance in the 26x26 km cells near the glacier outlet will be biased high (i.e, not negative enough), whereas in the regions the method will result in a low bias (too negative). Please comment.

- According to p. 543, l. 8., the ICESat period spans Sep. 2003 to October 2009. This is combined with GRACE data for Dec 2003.-Dec. 2010. You assume that the ICESat spatial distribution is representative of the longer GRACE period, but this assumption is never substantiated. It's a well know fact that the trend patterns observed by GRACE are non-stationary: in the course of the mission, mass loss has expanded to the northwest, whereas the mass loss in the southeast has decelerated around 2008, and started to accelerate again around 2010. Furthermore, according to Luthcke et al, 2013, 2010 was a year with a record mass loss. Therefore, I'm a bit skeptical about this assumption. To illustrate my point, the figure below shows the GRACE trends for Sep. 2003 to 2009 on the left, in cm/yr of equivalent water height. The figure on the right shows the difference in trend between Sep. 2003—2009 and Dec. 2003—Dec. 2010. Indeed, mass loss has increased along the west coast. Note that a 200 km Gaussian smoother was applied to the GRACE data to reduce noise: locally, as they would be observed by ICESat, mass loss rate differences will be significantly larger than in the plot. There's no pressing reason to use GRACE data for Dec. 2003—Dec 2010, except maybe to facilitate the comparison to the IMBIE paper. For consistency, I recommend using identical time periods for both the GRACE and ICESat data, which shouldn't require too much additional work since the GRACE observations are available at monthly intervals. Alternatively, if you want to stick to the present time periods, you should provide solid evidence that the effect of using different time periods is negligible.



- According to Gardner et al., 2011, the mass loss in the Canadian Arctic is driven by surface mass balance. High resolution SMB fields are available from regional climate models (e.g., Gardner et al., 2011; Lenaerts et al., 2013). Given the paucity of in-situ validation data, it seems worthwhile to compare the HIGA results to these SMB fields.
- Recently, a method has been published which combines GRACE gravity fields and ICESat elevation changes with *a priori* information of the spatial characteristics of SMB and GIA in a Bayesian framework, at a similarly high resolution as this work (Zammit-Mangion et al, Resolving the Antarctic contribitution to sea-level rise: a hierarchical modelling framework, Environmetrics, 2013). A short comparison of the two methods, or at least a reference to this work, seems appropriate.

Specific Comments:

Abstract: lines: 15-18 The part describing the continuity equation is rather general and should give a brief description of the results of the method and its validation.

p. 539: line 17: It seems more correct to state that altimetry characterizes the spatial variabiltiy of VOLUME changes..."

lines 23-24: I don't think you can say tat this stage that you have overcome 'the dependence on modelling complex firn processes': you have chosen to ignore these processes rather than actually overcome them, see General Comments.

Lines 29-32: This statement seems to be a bit premature, given that the results can't really be validated due to the paucity of in-situ data and the work of Zammit-Mangion (General comments). Delete this statement

p. 541: line 18: "Little Ice Age"

lines 11-28: Mention the spatial resolution of the GRACE mascons

line 20: Luthcke et al. 2013 use the GLDAS/Noah model to remove hydrology signals, but global hydrology models are know to perform poorly in Arctic regions. Luthcke et al., 2013 attempt to minimize leakage from hydrology signal by imposing constraints between glacier and non-glacier mascons, but due to the low resolution of the mascons, snow load within glacier mascons will still affect the mass balance estimates. E.g., in the Baffin Island region, GRACE picks up a large seasonal snow signal which is not captured by GLDAS (see Lenaerts et al. 2013, GRL). Please discuss how interannual snow variability affects the mass balance estimates for the Canadian Arctic.

Line 20: In Gardner et al., 2011, the ICESat data are processed by fitting planes of 700 m in length and a few hundred meters in width to the ICESat repeat tracks. How did you obtain the 2 km resolution from these planes?

Line 25: What is your motivation for not supplementing the Canadian Arctic ICESat altimetry by airborne observations?

p. 543: l. 6-9: "The IMBE ICESat period starts nine months after, and ends thirteen months before,. The IMBIE GRACE comparison period. Confusing, on p. 540, l. 12 you define the GRACE IMBIE comparison period as Dec. 2003-Dec. 2010. The ICESat period is defined here as Sep. 2003 to Oct. 2009, i.e., it starts a few months before the GRACE period.

p. 545: l.14: Zwally et al., 2011 first correct for temperature and accumulation-driven elevation changes before resolving mass loss. This should be mentioned here.

p. 549: l. 4: Longuevergne et al, 2010 discusses GRACE hydrological estimates on the High Plains Aquifer. Is this the most appropriate reference for 'cryosphere-attributed spherical harmonic solutions'?

p. 550: l. 15.: The estimates of Schrama and Wouters (2011) are likely biased due to the coarse ice mask used for Baffin Island, see Bonin et al, 2012, GJI. A better reference would be Jacob et al. 2012, who found -34+/-6 Gt/yr for basin 9 and -33+/-5 for basin 10. Also, I suggest to compare the mass loss of basins 9 and 10 separately to the cited studies.

l. 25: How do the estimates for the Greenland peripheral glaciers compare to other estimates? For example, Sector 3, Gardner et al. Give -18+/-3 Gt/a, for sector 2, -2.8+/-1 Gt/a.

p. 551.: l. 1-6: Briefly discuss how these findings compare to other GRACE studies, e.g. Schrama and Wouters, 2011 and Sasgen et al., 2012.

l. 20: There reason why GRACE products are never compared to point measurements is that they are representative for much larger areas. Even at the 26 km resolution suggested here, this is problematic, as is mentioned further on in the text. I suggest to delete this sentence.

p. 554: l. 14-16: It would be good to mention that two of the site are located in the ice sheet interior where the mass balance fields are smooth and the expected uncertainty is low (according to fig. 5, the ICESat data don't provide much additional information here and the GRACE measurements may be representative for the local mass balance) and the other two in the marginal zone, where the gradients in mass balance are large. Also, discuss the results separately (i.e., in the interior both points agree, in the marginal zone, only one of two).

l. 23-24: “consistent with the notion that mass balance generally has decreased”: this will depend on the locations where the in-situ measurements were made. Most of the points appear to lay in regions where your method suggests an increase in mass balance. Furthermore, it may also suggest a systematic bias in the method. This possibility should be mentioned as well.

l. 27-28: Mention that almost all sites are located in the interior and that in the more challenging marginal zone the method remains to be validated.

p. 555: l. 21: shouldn't dot_m be dot_h here?

p. 556: l. 24: Is this what you expect to see (divergence in marine-terminating regions, convergence in land terminating regions). Briefly discuss this here, rather than at the end in the conclusions.

p. 557: l. 14: overestimating the dot_m may be a result of the ringing problem discussed under 'General Comments'

l. 22: given the large uncertainty in the ice flux divergence, are the values in the interior significantly different from zero?

Fig. 8: It's hard to tell the points in the upper right corner apart. It would be helpful to include a blow-up of this part of the plot.