Reviewer Two

The reviewer is particularly concerned with the sensitivity of the approach to prior assumptions. Below we discuss this in some detail, and provide an outline of what will be included in the revised m/s.

Very little discussion on error analysis is provided. I think this should be addressed, particularly since the approach is statistically driven, so I would think error estimates for all of components would be available. The reader doesn't get a feeling for the errors for most of the input data sets and, other than the final ice mass loss values, none of the estimated components. Sigmas for the altimetry trends are given in Fig 2, but what are the uncertainties on the input GRACE mascons, the final SMB estimates, the final GIA rates? What is their spatial variation? Without this, it makes interpretation of the results and assessment of the comparisons more difficult (e.g., p3008 and p3009).

This point was also raised by the first reviewer and a subsection will be devoted to it in the revised m/s. This issue is exacerbated by the fact that we employ several data sets and provide solutions for several fields; a detailed analysis of all uncertainties will prove length and tedious to the reader. We will thus report the following:

a) We will plot a spatial map of the uncertainty associated with GRACE.

b) We will provide uncertainties over the summed contribution of SMB estimates, and the summed dm/dt due to GIA over the domain under study. These will be included in all of the comparisons (e.g. p3008 and p3009).

c) With regards to the spatial variation, we have stippled the mean plots where the field is deemed to be significant. This, together with what we report in b), we feel is sufficient. Plotting spatial uncertainties will increase the figure count by four and will considerably lengthen the manuscript.

Related to the first item, I would have liked to have seen more discussion on the influence of the various constraints that are employed (dh/dt error cutoff, static surface density, length scales, ice velocity constraint on elevation rates, etc.). I suspect that these have a significant influence on where the mass change is allocated within the framework, particularly the ice velocities constraint outlined on p.3004. What if a different constraint is used? What if no constraint it used?

The reviewer is correct in saying that our prior assumptions considerably affect the results. Several parameters are assumed known (e.g. surface density), and, moreover, we impose soft maxima for all processes which in turn are dependent on some prior belief on how these relate to the processes (for example the ice velocity constraint). On this point, putting no constraint will give a solution where the height change due to ice dynamics follows what is being observed in altimetry, since the spatial correlation length is relatively small. The ice velocity constraint helps to reduce this over-fitting and allocate height change in areas where ice velocity is low to other processes (in this case SMB and firn compaction). With regards to a difference constraint, the reviewer is right in saying that a different constraint will give different results. However, the aim of the framework is not to provide solutions which are independent of prior beliefs, rather to provide a solution under realistic prior assumptions (which could be verified using independent tests). Unfortunately, due to the under-determined nature of the system, estimating many of the parameters in our prior assumptions (for example the functional relationship of the constraint with respect to ice velocity) is not really tractable.

The resulting GIA uplift rates seem very smoothed...much more than the 100km smoothing constraint mentioned on p2004 ln17 would suggest. Please comment.

We think the reviewer is referring to p3005 line 17. Here we report the mesh density of the process to be on the order of 100 km (i.e. the side of the triangles in the triangulation are roughly 100km). This is not the length scale of the process, indeed the length scale needs to be much larger than this to be reconstructed from a 100 km mesh density. In this work we use 3000 km (p3004, ln13). This parameter was extracted from the Ivins and James (2005) GIA model.

A great deal of detail has been skipped regarding the methodology. The reader is not really left with a sense of how the whole system works. I realize you can't reproduce everything from earlier Zammit-Mangion et al paper, but I believe more can be done to describe the methodology. For example, the parameter layer isn't explained. And it's unclear how you go from the three layers to the FE mesh of the different processes to a final "statistically sound" result. How are you able to effectively separate the four different processes discussed in the Results section. Please consider adding some more explanation, figures, etc. in this section. At the moment, the methodology is very much a black box.

The reviewer is right in remarking that the methodology is detailed in the Zammit-Mangion paper. We have tried not to clutter the paper with too much mathematical detail but unfortunately this has made the analysis less clear in some areas. In the revised version we will expand Section 3, and include an overview diagram (similar to that in the paper of Zammit-Mangion et al.) to facilitate the understanding of this section.

Comparisons with ice core data is presented in support of the SMB results derived. Given the variability seen in the SEAT cores, it's difficult to accept any conclusions from the MEDLEY result, which represents just a few cores. What if the MEDLEY trend was an anomaly like SEAT 10-5? The Ligtenberg et al 2011 paper, which discusses the FDM derived using RACMO, made comparisons with 48 ice cores and looks to show good agreement with these cores. Many of these cores were in the WAIS, so there looks to be many other ice cores in the region that could be used to validate your model.

The reviewer addresses a valid point regarding the scarcity of ice core data for independent comparisons. Unfortunately, very few ice cores cover the observation period. In Ligtenberg's paper, the effects of firn compaction are addressed. This requires a larger observation period but it also allows for temporally cutting out a period of interest from the available data. Our comparison of a 7 year SMB anomaly trend is more dependent on perfect temporal agreement, and we did not have any other cores available for the 2003-2009 period. We will mention this in the revised version of the paper.

The abstract suggests it would be easily scalable for the whole of Antarctica. If so, then why was only the WAIS explored? All of the GIA/SMB/ice-mass change comparison studies (e.g., King et al, 2012, Shepherd et al, 2012, etc.) cover the AIS, so the same comparisons could be made. It would have made for a more complete comparison.

We agree with the reviewer that covering the whole of the AIS would have made it easier to better assess the performance of the framework. Unfortunately, as we hope the reviewer appreciates, the framework is computationally intensive (particularly in terms of memory) and at the time of writing we did not have the algorithms, nor the computational resources in place to consider the AIS as a whole. The study on the WAIS allowed us to explore computational tools to tackle this problem with ease, and establish a way forward. We are now implementing a similar approach for the entire AIS and that is why we know of the scaleability. We will alter the abstract to reflect this.

p2997, ln14: the Velicogna & Wahr is a bit dated, and their later papers show a lower proportion of GIA error. Consider updating reference.

We agree with the reviewer and will use more recent references (e.g. King et al, 2013, Sasgen et al 2014).

P2998, ln5: What if the SMB models have more than just systematic biases in them? For example, if the SMB variations themselves are modeled incorrectly (over/under estimated), then this would necessarily impact the spatial relationships used in the combination. This gets back to the earlier comment regarding the error analysis.

An advantage of the proposed approach is that any biases in the models, whether systematic or not, are not propagated to the framework. We use the models to extract typical length scales and orders of magnitudes; these are widely accepted to be correctly represented in the models. Biases would have affected our results if, for example, we had set the model output as the prior expectation of our process, which we do not. The amplitudes of inter-annual variability have been corroborated using independent estimates and we have confidence, therefore, that they are reproducing the correct of order of magnitude variability.

p3000, ln16: I assume these are formal errors on the trend. These tend to be optimistic, so I would recommend in the future applying some sort of error adjustment (bootstrapping, scaling, etc.) to make them more realistic.

This is an important point but it has not been tackled at this stage of the project. In fact, we have dealt with this problem in a recent paper by taking into account small-scale variations, and can include a sentence detailing this in an updated version of the paper.

p.3001, ln2: Considering you are using a RL04 GRACE mascon solution, which is a now dated release, it would have been very insightful to see how the results were affected when only the GRACE component was changed to, say, the CSR RL05 fields. In addition, how do the mascons relate to the FE mesh? The mascon discs won't be aligned with the mesh triangle boundaries, so how is this treated (if indeed it's even a problem)?

Our collaborators have only very recently, and thereby later than the rest of the community, pre-released a new version of their mascon solution. Given that this paper is only a proof-of-concept project, we will detail effect of the new release on the results in a forthcoming paper. To be clear, however, the Release versions referred to above are not relevant to the Goddard mascon solutions which are determined directly from the K band range-rate data. They have no relationship to CSR, JPL, GFZ or other spherical harmonic solutions. The most recent version of this mascon solution is version 2.

p3001, ln 20: The concern here is that the correlations would be more accurate if the mass loss was only due to surface mass changes, but a considerable amount of the observed mass change is related to GIA, which may have a different spatial signal. Plus, you're correlating mass variations using volume/height estimates. Most areas will have some correlation, but the degree of correlation will certainly vary, and introduce error. If this correlation between mascons is important, which I assume it is, it would be useful to see a more in-depth treatment of the error from the altimetry-based correlation, and its potential impact on the solution.

This relates to a query from referee 1. It is a non-trivial and important point which we will tackle in a forthcoming paper where we resolve the time varying component of the signals.

p.3001, ln23: What do you mean here by "averaging strength"?

Agreed that this is a bit unclear, we will modify the sentence to: "the extent of diffusivity is characterised by a parameter akin to the thermal coefficient in the heat equation, which is also estimated during inference (refer to Section 4.1 in Zammit-Mangion et al. for more details)."

p.3001, ln26: Should read "Thomas et al (2011)" instead of "Thomas and King (2011)". This occurs in other places as well.

Agreed, this will be corrected.

p.3006, ln9: It's not completely without prior information because the mass loss due to dynamics is constrained by the ice velocities described on p 3004, ln 16. This is equivalent to applying a type of forward modeling approach where all of the mass loss in essentially forced to go to regions of high velocity.

Yes and no! There is no one-to-one equivalence to a forward modeling approach; rather, it is only the probability for ice loss that is higher in areas with high ice velocity. We should stress that the framework does not "prevent" or disallow a dynamic signal in slow flow areas, it gives it a lower probability than in fast flow areas. Thus, for example, if GRACE detects a mass anomaly in a slow flow area with a length scale not characteristic of SMB (from altimetry) then it is possible to assign this to dynamics. In addition, in the time-evolving version of the framework, separation of SMB and dynamics will be improved because the former has, in general, high temporal variability, while the latter varies smoothly in time.

p.3007, ln25: If elastic effects are removed from the GPS displacements, wouldn't this impact your firn/elastic estimates?

In principle, the reviewer is correct. In this version of the framework however, the GPS stations are modeled to only measure GIA, therefore the elastic signal is removed in advance.

p3008, ln4: Why wasn't ICE-5G or the new ICE-6G included in the comparison analysis?

The key reason for this is that these are global solutions while the ones we compare to have been developed specifically for Antarctica. In addition, we wanted to include example of solutions derived from both forward modelling (W12a, IJ05-R2) and from data inversion (AGE-1, Gunter14).

p3008, ln 8: Isn't agreement with the GPS data nearly guaranteed since it is one of the input data sets?

This is an interesting point and it is worth mentioning. It is indeed not guaranteed, as the GPS data set is only one of several inputs (see comment above about data/solution misfits), and we are looking at a combination of these data. For example, in p2008 L10 we mention that there is poor agreement with the W06A station.

p3008, ln21: Wouldn't this agreement be mostly attributed to the smoothed nature of the RATES and AGE-1 solutions? Neither solution predicts GIA rates above 4mm/yr. Comparing a smoother solution to one with higher resolution and signal variation (W12a, IJ05-R2, and Gunter14) is an apples-to-oranges comparison, since they have different spectral content. Also, discussions of agreement should be done with uncertainties involved. Are the differences statistically significant? What are the uncertainties of the various components?

It is correct that we prescribe a certain spatial smoothness for the GIA solution. It is entirely possible that the real GIA signal is spatially less smooth than what we have assumed here. The degree of smoothness strongly depends on the "smoothness parameter" κ_{GIA} mentioned above and the reviewer's comment highlights the fact that we have failed to state this clearly enough in the paper. However, the mean basin uplift rates should not, in our view, depend strongly on a higher resolution of the GIA field, so we cannot say that we agree with the conclusion that the agreement between AGE-1 and RATES is due only to the fact that they are both smooth solutions. Also, we believe that a basin comparison of uplift rates is a useful comparison for GIA models.

p3012, ln5-25: It should be noted here that proper uncertainties of the input data sets is key towards generating reliable results. If, for example, the SMB estimates had 2-3cm uncertainties, then this would be reflected in the final estimate, i.e., the GIA rates would have large error bars. The same goes for the other data sets (altimetry, gravimetry, GPS). It's only a problem if the errors in the input data are too optimistic, i.e., lower than they are in truth.

*Yes, this is correct. However, we **do not** include any prior uncertainty estimates for SMB, only for the input data fields. Nonetheless, the referee is correct in stating that the solution is dependent on using the "proper uncertainties" and this is something that is surely a good thing! We would like to think that any reliable estimation would be dependent on the input errors. As stated earlier, we will detail our error estimation approach more fully in a revised m/s.*

p3012, ln26: The agreement with AGE-1 has been stated a couple of times, but when I visually compare the RATES and AGE-1 results, I don't see that much similarity, mainly because the RATES results have smoothed out most of the features. You might consider having a discrete color scale to better visualize the variations in Fig 7.

We do mention that the spatial pattern in RATES is different from that in AGE-1 (p3008 L8), but an updated figure can be included in the updated version of the paper to improve visualization.

Perhaps the station names can be added in one of the figures (e.g., Fig 8?).

We will include an updated figure.