

## Reviewer One

We would like thank the reviewers for their helpful and insightful comments. Referee comments are in normal font in red and our reply in italics.

I was curious why it was decided to estimate a trend, instead of, for instance, monthly changes. A trend is a good model for some of the processes (i.e. GIA), but perhaps is a poor choice to represent processes such as surface mass balance since there can be significant inter-annual variability. Could you please comment on this?

*This statement is correct but raises several issues that represent computational and data challenges. In our framework, we are simultaneously inverting seven years' worth of ICESat, ENVISAT, GRACE and GPS data at relatively fine spatial scale. To make this computationally feasible we must employ dimensional reduction and this is partly achieved with the use of stochastic partial differential equations and a Gaussian Markov random field approach. This is not sufficient alone and we need, therefore, to reduce the data sampling in space and or time. In effect we do both. As explained in the m/s we are working on a time evolving solution but this will be with annual resolution. Monthly or seasonal solutions are challenging because of i) the computational cost and ii) that some of the data does not adequately resolve sub-annual signals. The ICESat data, in particular, which provide the lowest error  $dh/dt$  estimate, must be estimated over 3 year means to produce adequate spatial coverage.*

Why not explicitly use the posteriori correlations from the mascon solutions themselves? Additionally, in Luthcke's solution, a 2000 m elevation cutoff is dictated in the spatial correlation constraints. Is this explicitly considered in your analysis? I believe using the formal posteriori covariance matrix would be more favorable

*The reviewer suggests using the full covariance matrix for GRACE. Unfortunately, these data were not available for the release we are using here. Regarding the elevation cutoff, we have been provided with a version where the 2000m cut-off was not applied. This was not explicitly mentioned in the paper and will be added to the revised version.*

I believe a nice addition to the paper would be a Table which succinctly captures the processes and methodology. For instance, in this table it would be nice to list the following: 1. Observations (altimetry, GRACE, gps, etc) 2. State parameters (trend for GIA, ice dynamics, etc) 3. Weighting information on the observations (both diagonal and off-diagonal components) 4. Assumed apriori information on the state parameters (both diagonal and off-diagonal components). This would allow the reader to quickly assess exactly what is being done and what assumptions are being made.- Truly, there is some dependence on your solution with the choice of 3) and 4) in the above comment. Could you please remark on this, or provide some analysis on how sensitive the solution is to these choices? Additionally, there was no discussion of relative data weights. Do you weight any observations higher than others? What is the relative weight of the apriori information on the state relative to the observations? Some discussion of these matters would be appropriate.

*This point was also raised by referee 2. Including a sensitivity analysis is possible and we will add a table as suggested in an updated version of the paper, and include a more detailed section regarding the sensitivity of the results to any data errors. However, we would like to stress at this point that there is no weighting applied to the observations.*

The results that you presented left me wondering how well you are fitting the data. What are the RMS of the residuals? How does the misfit to each observation type look spatially? Is

your estimate fitting to one specific observation better than another? For instance, does your estimate agree better with altimetry than GRACE? If so, perhaps this was reflected in the initial choice of weights on the observations and choice of apriori information. This type of analysis would provide more credibility to the results presented.

*How closely the results fit to a particular data set is, of course, a function of the a-priori error estimates that are associated with each dataset. As a consequence of this, we agree that it is important to discuss how these errors were estimated for each data set and we will expand the discussion of this in a revised m/s. It is not so easy to show a “misfit” between the data and the solution because the data are not observing one of the solved-for fields directly but some mixture of these fields. Altimetry does not observe a mass change but a volume change, for example.*

**What physical processes could allow for a negative SMB in this region?**

**Could you please comment more on this, and why this result is believable? “**

*We should have stressed that our results are trends on the SMB anomalies, so a negative signal represents a negative trend in the anomalies to a long-term mean. This could mean that, e.g. there is less snowfall over the 2003-2009 period than over the former years. As a consequence a -ve anomaly is just as plausible as a +ve one. We will clarify this in the text.*