

Reviewers' comments are *in italics*, while our replies are in roman text. Serif font indicates our action.

Reply to Reviewer 2

We thank reviewer 2 for providing a detailed review.

There are several major concerns with the study itself. They are listed here in no particular order of importance.

We address the reviewer's major concerns as part of our general reply, where we outline a new introduction.

First, the hindcasting period of ~20 years is extremely short by glaciological standards. Apart from the fast flowing outlet glaciers, whose dynamics is barely resolved in this model due to a number of reasons (e.g. resolution, the absence of appropriate boundary conditions/forcing at fjord terminated glaciers etc.), that can move some appreciable distance (~10-60 km), the rest of the ice sheet hardly moves at all, perhaps, remaining within the original model grid cell. Hence, all possible changes that simulated during this hindcast period are due to the surface mass balance, which is identical in all simulations. Therefore, the authors could have easily omitted this hindcasting stage, and compared the ice-sheet configurations at 1989. If this assessment is incorrect, the authors need to demonstrate that significant dynamic changes happened in these 20 years of simulations. In other words, the authors need to demonstrate that hindcasting for such a short period of time is indeed a reasonable approach to test sensitivities of ice-sheet models to their initial states.

The hindcasting period of 20 years is indeed a short period for an ice sheet modeling study, however, recent observations clearly show changes at short time-scales and we therefore, for the first time, assess an ice sheet model at this short time scale.

The reviewer's assertion that the hindcasting stage could have been omitted is incorrect. A major aspect of our paper is that rates of change are better metrics for model evaluation. Hindcasting is needed to obtain simulated rates of change for a time period where observations are available. We now make this clear in the Discussion section:

“A comparison of observed and simulated rates of change requires a reference period covering both observations and simulations. Hindcasting provides simulated rates of change for this reference period. In other words, it adds a temporal dimension to validation efforts.”.

As explained in the suggested introduction there are also practical limitations that control the choice of the hindcasting period, as validation data is only available for short periods. The reviewer's assessment that all possible changes during the hindcast period must be

due to the climatic mass balance is incorrect. This would be only true if the initial state represents a steady-state in the mathematical sense. This, however, is not the case for “paleo-climate” and “flux-corrected”, and may not even be the case for the “constant-climate” initial state, as pointed out on p. 5078, l. 10–12: “Even sufficiently long constant-climate initializations are not expected to be free of transients; for example, basal hydrology may prevent a steady-state configuration in the mathematical sense (e.g. Kamb et al., 1985).”

Second, the climate forcing used during the hindcasting stage come from an atmospheric model (HIRHAM5). Undoubtedly, as any model, it has errors and biases. Since the changes during the hindcasting period are dominated by the evolution of the surface mass-balance, one could argue that the effects of errors in climate forcing on the final states is substantial. There is no analysis or discussion of such errors or performance of this model. How different would the results be if outputs from other atmospheric models would be used as climate forcings? Since the changes during the hindcasting period are dominated by the evolution of the surface mass-balance, one could argue that the effects of errors in climate forcing on the final states is substantial.

We agree that errors in climatic mass balance have a large effect on the final state. As explained in our general reply, we examine how hindcasts with our particular choice of an ice sheet system model compare to observations. The performance of the regional climate model of our choice, HIRHAM5, has been assessed in a separate publication (Rae et al., 2012) and we refer to that discussion in our paper, but a formal error estimate is not available. Undoubtedly there will be a differences in results when using different climate forcing. This is in fact explored in a companion paper by Aðalgeirsdóttir and others (manuscript in prep.).

Third, the choice of ways to obtain initial states is not obvious. Though, it is the authors’ choice, and they are free to use any approaches, the manuscript does not provide any discussions or justifications for these approaches.

Regarding our choice of initialization procedure, we added a sentence that we chose three initialization techniques that have been used in published literature. These three initializations merely serve as examples used to explain the hindcasting method.

Forth, the comparison of the final states to observations, although not surprising, in my view, is not very informative. There are many other issues (e.g. unknown and/or unresolved physics, unknown/unresolved boundary conditions at the ice-sheet margins and bed) that can result in the simulated present-day states that are very far from the observed one. As a side comment, comparison to the GRACE observations, most likely, cannot be treated as an appropriate metric, or at least cannot have the same weight as comparison to other observations (e.g. surface elevation and its changes, surface velocities, ice sheet extent, etc.). This is due to the GRACE coarse resolution and a number of issues in processing (e.g. postglacial rebound, etc).

As outlined in our general reply, we discuss the current limitations of our approach in the revised discussion.

Of course, it is the authors' choice how to conduct their investigation, however, if I may offer a suggestion, it would be to use a synthetic approach to illustrate the usefulness of hindcasting. One could create an artificial state of an ice sheet (either using a realistic bed topography or idealized) by forcing an ice sheet with prescribed climate conditions. This state would be the true "present-day" ice-sheet configuration. Then, one could repeat a procedure somewhat similar to what is done in this study, i.e. create different initial states for the hindcasting period and compare the different ice-sheet configurations obtained at the end of hindcasting to the true state. It would be interesting to see what the optimal hindcasting period might be. My hunch is that 20 years is too short. By doing so, the authors would be able to truly isolate the effects of initial states and investigate/demonstrate advantages of the hindcasting approach.

This is an excellent suggestion for further research on how to define a "validation" time scale.

p. 5070:, l. 4–5: *what is "the quality of projections" and how can one measure it?*

As outlined in the general reply, we rephrased the introduction. The new version does not include this sentence anymore.

p. 5070:, l. 6: *"... initial states" of what?*

We mean initial states for ice sheet model simulations. The new version does not include this sentence anymore.

p. 5070:, l. 11: *what is "dynamic state"*

The dynamic state is the distribution of momentum within a system. Changed to "thermodynamic state (i.e. distribution of momentum and energy)" for clarity.

p. 5070:, l. 21: *"Ice sheet models integrate such physical process understanding." an awkward sentence*

As outlined in the general reply, we rephrased the introduction.

p. 5070:, l. 25: *what are "spatially-rich" observations?*

"Spatially-rich" means having lots of measurements in the spatial domain sufficient to reflect spatial variability of the measured quantity.

p. 5071; l. 9–23: *this paragraph is out of place and unnecessary, either remove it or re-write it to be stylistically similar to the previous and the next ones.*

As outlined in the general reply, we rephrased the introduction. The new version does not include this paragraph anymore.

p. 5073; l. 1–4: *this paragraph is unnecessary, the paper is fairly short.*

Removed.

p. 5073; l. 6: *why is 1989 used as a datum?*

1989 is the start of the ERA-Interim reanalysis product that has been used as lateral boundary conditions for the regional climate model HIRHAM5 and therefore the start of the climate forcing time-series (1989–2011) used in this study.

p. 5073; l. 6–7: *flux-correction is not used in climate models anymore.*

We agree, see Discussion, p. 5079, lines 1–4: “Improved physics, higher resolutions, and more physically-consistent coupling have rendered flux corrections mostly unnecessary in AOGCMs. Thus flux correction methods may be seen as a temporary remedy, until better coupling to the atmosphere is established.” Please note that we chose three published initialization methods to explain hindcasting. In particular we make no a priori statement whether we consider flux correction methods as viable initialization techniques or not.

p. 5073; l. 9: *what does “overall dynamic state” means?*

It means we expect the general flow pattern to agree between observations and simulations, while accepting disagreement at a local scale. Informally speaking, we’re trying to get the “bigger picture right”. The meaning of ‘dynamic’ is explained in reply to the comment on p. 5070, l. 11. Not changed.

p. 5073; l. 16–17: *what does that mean “reduce model complexity”? What are boundary conditions there?*

Rephrased to: “Therefore no forcing at the ocean boundary is applied. ” Regarding boundary conditions, in the Supplement we state: “At the ocean boundary, ice is calved-off at the initial calving position, which is held fixed throughout the simulations.”

p. 5073; l. 25: *join with the previous paragraph.*

Changed as suggested.

p. 5074:, l. 6–7: *what are resolutions of spinup simulations?*

We explain our grid-refinement strategy and resolutions in the Supplement.

p. 5074:, l. 23–: *maybe say something why these data sets are used. This might be a good place to say something about metrics.*

We added a sentence clarifying our choice of data sets.

p. 5075 :, l. 16–: *suggest to use “constant climate”, “paleo-climate”, “flux correction” in quotation marks*

Changed as suggested.

p. 5075:, l. 24–25: *what does it mean “normalized to the beginning of the GRACE period”?*

Changed to: “Figure 4 shows the time-series of mass change since the beginning of the GRACE period (January 2004)”

p. 5076:, l. 24–26: *“... the three initial states” do not “respond differently” they adjust to a shock (or a jump) in the surface forcing.*

“Respond differently” is more general than “adjust to a shock”. We use “respond differently” at the beginning of the discussion section because, at this point, the nature of the differences in the response has not yet been established. We address unphysical transients (i.e. a shocks) on p. 5078, l. 8ff. Not changed.

p. 5077:, l. 3–6: *the last two sentences of this paragraph are unclear.*

We replaced

“Trend under-estimation is expected because of the absence of ocean forcing that could lead to an increase in ice discharge.”

with

“As mentioned earlier, observations show a rapid increase in ice discharge since the late 1990s, which was attributed to changes at the ocean boundary. Our model does not include ocean forcing that could lead to an increase in simulated ice discharge. Therefore, under-estimation of the simulated mass loss trend is expected.”

p. 5077:, l. 11: *why should the split between ice dynamic and surface processes be the*

same?

We believe our suggestion above clarifies this question.

p. 5078:, l. 1–7: *for how long has the “interim forcing” been applied? What is “ERA-forcing”? What are its errors?*

ERA-interim is a reanalysis product of ECMWF providing the best available lateral boundary conditions for regional climate models. This is available for the period 1989–2011 (at the time of our study, it is continuously updated). We mention on p. 5071, l. 6. that the climate forcing was applied 1989–2011. A comment on p. 5071, l. 6 is added stating that this corresponds to ERA-interim period and we added a reference to Dee et al. (2011). The performance of HIRHAM5 has been assessed in Rae et al. (2012) and we refer to that discussion in our paper, but a formal error estimate is not available.

p. 5078:, l. 23–29: *This paragraph is too cryptic, either clarify or remove it.*

Reviewer 3 considers this paragraph as an interesting result. **Not changed.**

p. 5079:, l. 5: *what does that mean “Surface elevation changes corrected for model drift”?*

We explain this on p. 5078, l. 12–14: “An approach to mitigate the effect of model drift involves calculating model drift and subtracting it from the experiment.” Changed to: “Model drift is removed by subtracting the surface elevation change time-series of the drift experiment from the hindcast. Drift-corrected surface elevation changes are shown in Fig. 7.”

Section 5 Conclusions needs to be rewritten. The presented conclusions have a very loose connection to the material described in the manuscript.

As outlined in our generally reply, we rephrased the conclusions.

References

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.

Rae, J. G. L., Aðalgeirsdóttir, G., Edwards, T. L., Fettweis, X., Gregory, J. M., Hewitt, H. T., Lowe, J. a., Lucas-Picher, b., Mottram, R. H., Payne, a. J., Ridley, J. K., Shannon, S. R., van de Berg, W. J., van de Wal, R. S. W., and van den Broeke, M. R.: Greenland ice sheet surface mass balance: evaluating simulations and making projections with regional climate models, *The Cryosphere*, 6, 1275–1294, doi:10.5194/tc-6-1275-2012, URL <http://www.the-cryosphere.net/6/1275/2012/>, 2012.