I am not an official reviewer, but I have a comment that I'd like to share, hoping to improve the manuscript.

I'm wondering if the authors could provide a justification for the initialization ("spin-up") procedure. A fixed-surface initialization leads to an initialized ice sheet which is not in equilibrium with its climate and contains unphysical transients. As soon as the surface is allowed to evolve, the modeled ice sheet will quickly adjust towards a state that is in equilibrium with the forcing. The timescale of this adjustment is at least on the order of the length of the transient simulations made here (50 years). Consequentially, the response will be a mix of this adjustment and the applied reanalysis climate. Thus any interpretation of model results will be biased. This bias can be large, possibly dominating the signal over the modeled 50 years. Applying a surface relaxation may help to remove unphysical transients (e.g. Seddik et al., 2012; Gillet-Chaulet et al., 2012) or a different initialization procedure may be more suitable for this type of sensitivity study. For illustration, Figure 1 compares time series of mass change for two hindcasts, both forced with climatic mass balance and 2-m air temperature from RACMO for 1958–2011. One was obtained with a fixed surface (as in this manuscript) and the other with a free surface. While the differences are striking, this may be not used to make a case for one or the other initialization method. To detect whether a simulation is biased by unphysical transients, flux divergence or surface elevation changes are probably better metrics than the total mass change I've used in my illustration. In any case, validation with independent metrics is needed (c.f. Aschwanden et al., 2013). I thus recommend that the authors provide a strong case that their simulations are not strongly affected by unphysical transients.

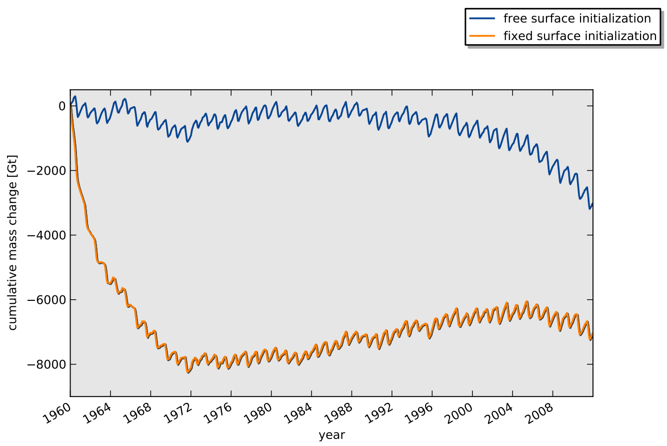

Figure 1: Modeled cumulative mass changes starting from the two initial states.

Minor comment The manuscript states: "...in order to validate a number of existing SMB parameterizations and our new approach against the results of the high-resolution model RACMO2/GR and recent satellite observations." I think this is an unlucky choice of language, and what I believe the authors meant to say is: "... in order to validate a number of existing SMB parameterizations and our new approach against recent satellite observations and compare to the results of the high-resolution model RACMO2/GR." Validation, even if interpreted somewhat loosely (as often the case in glaciology), means comparison to observations. Therefore, validation against another model is not permissible. One may weigh in that some modeling is required to obtain time-series of mass change from the L1 GRACE signal. However the major difference is that GRACE measures mass changes directly, and modeling is only needed remove contaminations in the signal (e.g. GIA).

Kind regards,

Andy Aschwanden

References

Aschwanden, A., G. Aðalgeirsdóttir, and C. Khroulev (2013). Hindcasting to measure ice sheet model sensitivity to initial states. *The Cryopshere*, **7**, 1083–1093. doi:10.5194/tc-7-1083-2013.

Gillet-Chaulet, F., O. Gagliardini, H. Seddik, M. Nodet, G. Durand, C. Ritz, T. Zwinger, R. Greve, and D. G. Vaughan (2012). Greenland ice sheet contribution to sea-level rise from a new-generation ice-sheet model. *The Cryosphere*, **6**(6), 1561–1576. doi:10.5194/tc-6-1561-2012. URL http://www.the-cryosphere.net/6/1561/2012/.

Seddik, H., R. Greve, T. Zwinger, F. Gillet-Chaulet, and O. Gagliardini (2012). Simulations of the Greenland ice sheet 100 years into the future with the full Stokes model Elmer/Ice. *J. Glaciol.*, **58**(209), 427–440. doi:10.3189/2012JoG11J177. URL http://www.igsoc.org/journal/58/209/t11J177.html.