Review of Noel et al

The authors present the results of one dynamically downscaled Earth System Model (ESM) simulation over the Greenland Ice Sheet (GrIS) and present the resulting historical surface mass balance (SMB) output from their regional climate model RACMO. After dynamical downscaling of the ESM input, the SMB is furthermore statistically downscaled to a nominal horizontal resolution of 1km.

Overall Impression (Text and Figures):

In general, the authors are doing a very good job in keeping their sentence and paragraph structure easy to follow and all their figures are well presented. Therefore, the manuscript is good to read.

Scientific assessment:

Overall, it's hard to make a case for how the study in its present form will benefit the wider cryospheric and climate community. The point of the authors here is to create a scientific foundation for additional papers that they want to write on the future contribution of the GrIS to sea level rise via (surface) mass loss. Overall, 21st century simulations of the GrIS climate and SMB would be very beneficial for the community, however, the presented analysis currently lacks the needed depth to be considered a valuable contribution to the field. Therefore, I would encourage the authors to consider the following points.

(1) The authors present only one RCM simulation forced with one GCM/ESM run to create a foundation for a future paper on 21st century GrIS climate projections. However, in its current form, the paper lacks (a) a consideration of the inter-model spread between all of the different GCMs in the CMIP5/6 model domain (b) a consideration of how the authors made their specific selection for the one run they choose out of their CESM2 ensemble. Fettweis et al (2013) for example analyse all the CMIP5 models over the current climate, selectively find the most suitable boundary forcings and create a downscaled RCM ensemble for multiple emission scenarios and models. This point is unfortunately omitted in this study.

(2) The authors focus their analysis only on the GrIS surface mass balance. If this study should become a standalone piece of work without the promised future projections, then the authors should be highly encouraged to consider at least a subset of other parameters to validate their single-simulation analysis to exclude the likelihood of compensating biases leading to a "correct" SMB due to "false" physical reasons - (a) Surface energy budget vs. observations (b) Albedo vs. observations (c) Temperature and/or cloud properties vs. observations.

(3) If the reader assesses the novelty based on what the authors highlight "…for the first time an ESM (CESM2) can be used to reconstruct historical SMB…" then the science of the paper would need to be judged either on the on the claim that is "the first time" or that the "historical SMB" is more accurate than from other model setups.

However, (a) e.g. Fettweis et al. (2013) as a benchmark already show that GCMs/ESMs can be used to force RCMs over the historical period and roughly get the magnitude of the SMB components right.
(b) The most accurate "historical SMB" does not come from this model setup, but rather from regional climate models that downscale observation-based reanalysis data (e.g.RCM with ERA-I or ERA-5). The presented results (Figure 3) unsurprisingly show that CESM2-RACMO does not capture the interannual SMB variability and extremes (e.g. melt in 2012) which is expected with GCM boundary forcings. However, it means that the accuracy of historical SMB representation is also not an advancement of the scientific knowledge.

Recommendations:

The reviewer would like to encourage the authors to either add significant extra analysis to their current model and study setup to create a solid foundation for their promised future attribution studies, or potentially add the presented analysis to their upcoming future projections altogether.

The authors could potentially consider some of the following points/questions when considering the next steps for their analysis post-review.

(a) Given the limited amount of future GrIS mass loss studies with RCMs and GCM forcing, the scientific interest of the presented approach lies in the actual future projections, not necessarily on the historical SMB reconstructions due to obvious limitations when using GCM/ESM boundary conditions.
(b) How representative is this one CESM2 run compared to the spread in CMIP5/6 simulations? Other recent studies have found great uncertainties in future GrIS projections using RCMs to downscale GCMs/ESMs which is/are not really discussed yet in the manuscript. What if the authors would force RACMO with other GCMs? How well does the current setup represent the surface energy budget, temperature, albedo, cloud properties?
(c) If forcing RACMO with other GCMs is technically not feasible, then one approach would be to force RACMO with additional ensemble members presented in Figure 4. The robustness of the SMB and potential underlying compensating errors can hardly be assessed by only one simulation.

Minor comments:

P1.L9: "without assimilating observations" is this correct? The methods of the paper claim that RACMO uses satellite albedo to constrain the surface albedo. Please clarify.

P3.L19: "bare ice albedo is prescribed from … MODIS.." – please see first minor comment and clarify.

P3.L28 Also in the statistical downscaling technique the authors use observed MODIS albedo. Please see the first comment on how this fits with the claim that this study doesn't use assimilated observations.

P3.L32-33: Does it only *change* the runoff and SMB or also *improve* the statistical comparison?

P4.L24: "due to the high quality of the CESM2 climate" but also e.g. **P1.L5** "good comparison" and **P5.L6** "shows excellent agreement" and at other points in the manuscript - these are quite colloquial expressions with little scientific meaning. What does a "high quality" climate in a GCM mean? The manuscript doesn't even currently evaluate the CESM2 climate for example.

P4.L25ff: But what about other parameters such as the surface energy budget, temperature and clouds? How does it compare to recent circulation and cloud anomalies over Greenland which have been shown to be important for future projections?

P5.L6-8: The acceleration (i.e. dSMB/dt) is likely not discussed here but rather a "total mass loss".

P5.L30-32ff: ad HadGEM; "did not accurately reproduce SMB". Throughout this study the reader is often left in the dark as to "Why?" certain numbers or results are mentioned, and why certain processes behave the way they do. At the moment, the paper is an ensemble of nice figures and easy-to-follow text, but the study and the reader would highly benefit if the authors would more often dig into the question of "Why?" some processes and numbers are reported here and apparently deemed important for the reader. This would also be a good point to address the matter why HadGEM and CESM2 produce such different SMB/ME/RU results (+-50%)? Is it due to differences in the lateral forcings/ the internal RACMO physics/ circulation / cloud physics? Hofer et al. (2019) for example show the large spread in GrIS SMB that can result from different GCM forcing.

P7.L8-9 "can reliably reproduce … variability of historical SMB" – When looking at Figure 3 the GCM-forced SMB reconstruction clearly lacks the ability to reproduce the interannual SMB variability and extremes shown in Figure 3A when RACMO is forced by reanalysis. Just as an example, the extreme melt summer of 2012 accurately captured in Figure 3A is not present in Figure 3B, therefore the the reader considers this to be a doubtful assumption.

P7.L3-4: unclear phrasing "is for 60%"

P7.L7-10: What are the uncertainties coming from the lack of a multi model forcing (e.g. Fettweis et al. (2013).

Figure 1: How does it compare during melt season? How does the SEB compare to the observational networks of PROMICE, DMI and/or GCNET?

Figure 2: Please clarify the choice of HadGEM and not other GCMs? If it is feasible to force RACMO with other GCMs then please consider analysing the intermodel spread of the GrIS climate when RACMO is forced by other GCMs

References

Fettweis, Xavier, et al. "Estimating Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR." Cryosphere discussions 6 (2012): 3101-3147.

Hofer, Stefan, et al. "Cloud microphysics and circulation anomalies control differences in future Greenland melt." Nature Climate Change 9.7 (2019): 523.