The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2019-321-RC2, 2020

The Cryosphere
Discussions

Interactive comment on "GrSMBMIP: Intercomparison of the modelled 1980–2012 surface mass balance over the Greenland Ice sheet" *by* Xavier Fettweis et al.

Anonymous Referee #2

Received and published: 19 March 2020

The manuscript presents an experiment in which the surface mass balance (SMB) output of five regional climate models, four surface energy balance models, and two positive degree day (PDD) schemes for the Greenland Ice Sheet (GrIS) are each forced with ECMWF-Interim atmospheric reanalyses over the period 1980-2012. They are compared with each other, with available in situ observations, with MODIS-derived bare ice extent, and with a derived gravimetric data set in which an observed terminal glacier discharge has been incorporated. The output from two global general circulation models is also considered. The main results presented are that the models simulate a statistically significant decrease in SMB over the period, that the largest differences between models occur on the ice sheet margins, and that regional climate

models generally perform well in comparison to the validation data.

The manuscript is around 7500 words with 6 figures and 4 tables, which is a reasonable length for the topic. It represents a considerable community effort in organizing and executing the experiment. The author list comprises most major modeling efforts for contemporary GrIS SMB with the outstanding exception of atmospheric reanalyses (e.g., the Arctic System Reanalysis; MERRA-2). The initial reaction is that this is a significant update on earlier efforts of Vernon et al. (2013) and perhaps Rae et al. (2012) in model assessment. While those studies were mostly focused on regional climate models, this manuscript aspires beyond that with the inclusion of a large number of surface energy balance and PDD models. I have a few points for the authors to consider below, and so would suggest some revision of the manuscript.

1. The experiment necessarily relies on the common use of one forcing product, ECM-Interim. By itself, this study is then not a complete characterization of SMB and its uncertainty from model sources, as the uncertainty of the forcing would also need to be considered. The use of different forcing products is beyond the scope of this study, but it would seem that the forcing selection plays a significant role in determining trends. Consider that if one wished to comprehensively evaluate a forcing product for the GrIS, a possible approach would be this experiment: a comparison of many forced models with observations may be seen as an elaborate validation of the forcing product. Is that not so? The purpose of this study is an appraisal of the different models, and for this purpose the key results are in how models compare with each other, and the systematic differences between them. These would seem to be the results that should be emphasized. Comparisons with observations are of interest (e.g., Fig. 1 is very interesting) but would not seem to be the principal outcome to be emphasized. The manuscript presents a considerable amount of information on the intercomparison in the form of figures and tables. Beginning in section 4, the text focuses primarily on the direct comparison with observations. The intercomparison is largely covered in the second paragraph of section 5. It is suggested that the results be re-ordered with the

intercomparison presented first. Some additional aspects of the intercomparison may be highlighted, as suggested below.

2. For tables and Figs. 2, 4, 5, and 6, the plots should be sorted by model type rather than alphabetically, and perhaps labelled accordingly. It may also be useful to plot the spread for each model type.

3. It is useful to continually compare this experiment with the previous efforts cited in the introduction. Vernon et al. found SMB estimates were within 34% of the multi-model mean of 4 models. Table 4 suggests this value is now something like 22% for the 5 RCMs but close to 100% when all of the models are considered. Does this suggest an increasing proficiency within the RCMs. Also, it is noticeable that the manuscript does not indicate surface temperature sensitivity. It is difficult to include and assess the PDD and EBM models without that consideration. For example, this was a focus of Bougamont et al., who found that PDD models were more sensitive than EBM models. Given the same forcing and the trends shown in Table 2, it does not appear that a similar conclusion holds here, is that correct?

4. At 445 words, the abstract is too long by half. A large part of the abstract is devoted to motivation, which should instead be mostly left to the introduction. As suggested in the previous point, may consider adding more text regarding the resulting differences between the models.

5. A concern is the very lengthy description of the models contained in section 2. The descriptions include sub-model components, the forcing time scale, vertical resolution, ancillary forcing data, etc. It is of course useful for close examination of individual model results, but this is generally available elsewhere from the cited literature, and it is not clear that all of it is directly pertinent to the aggregate results presented for understanding cryospheric modeling. It is suggested that this material may be incorporated into supplementary text and/or condensed with a table that includes model type, references, and major points. Otherwise it could be argued that this type of material is

more appropriate for a publication such as Geoscientific Model Development.

6. Lines 455 and following. As indicated, it is apparent from Fig. 6 that the snowfall from the EBM models, which is directly imported from the forcing, is low (mostly blue) compared to the mean over the interior regions of the ice sheet. Would it be correct in saying that models that compute snowfall generally show higher amounts than the forcing? This appears to be true for most of the HIRHAM, NHM-SMAP, and RACMO RCMs and to some extent for the BOX13. Is that an expected systematic response?

7. A list of acronyms in the appendix would be useful.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-321, 2020.