Review of 'Antarctic surface climate and surface mass balance in the Community Earth System Model 2 (1850-2100) by Dunmire et al. 2022'

Dunmire et al. exhaustively evaluate the results of CESM2 over the Antarctic Ice Sheet against AWS observations and several often-used products such as RACMO, MAR, ERA5 or the reconstruction from Medley and Thomas (2019). The comparison is honest by highlighting both remaining biases and improvements against the previous version of the model. Furthermore, they discuss the future evolution of SMB using three different scenarios as it is a key variable for the ice sheet dynamics. One could argue than the topics is not particularly new, but this is an important study that would deserve a publication in the Cryosphere. CESM2 (one of the few polar-oriented ESMs) is often used either directly or downscalled to study the Antarctic climate and force ice sheet models. Knowing its biases is then important and I think that updates comparing to CESM1 (new emission scenarios, improved physics) are sufficient to justify a new study to present SMB projections. Furthermore, I found the paper well-structured, with clear figures and conclusions that are supported by results.

I have only minor comments listed hereafter that I hope will help the authors to improve their manuscript before its potential publication.

Minor comments

P4 Section 2.5: I recommend to specify the forcing of the RCM s(I think ERA-Interim for RACMO, ERA5 for MAR). Furthermore, you used ERA5 as a reanalysis but the reconstruction based on MERRA2. Yet this choice is explained and fully justified by the better performance on the reconstruction with this specific reanalysis but I wonder if the different forcings (MERRA2, ERA-Interim, ERA5) could results in different trends and change your comparison with CESM2. Since you also used a collection of products that could all give you the same information, is there any reason why you selected ERA5? For instance, precipitation or temperature could be also compared with RACMO or MAR. Try to justify why you selected ERA5 for these comparisons.

P11 L205: According to Figure 6a, I guess the values are for the grounded ice sheet. What is the SMB of CESM2 over the ice shelves? Are you well using a common mask? if yes specify it to help the readers interpretation, if not I strongly recommend to do it to compare something similar. (Also true for CMIP5 and CMIP6 values in Fig6.). There seems to have no evaluation/comparison over the ice shelves while other products than "the reconstruction" could be used there. Since ice shelves are particularly important for the Antarctic mass balance, this should be corrected.

Trends: Why do you use normalized trends? I understand that it better highlights the importance of small changes over areas with low values (eg., SMB over the high plateau) but at the same time it masks the real changes. Importance of small changes that are significant can still be highlighted by dots or crosses as you did. For Figures 3b and 7c indicate in the caption what crosses represent.

Most CESM2 trends are compared against other products (which is really interesting), but not the melt trend. I'd suggest to perform a similar comparison or at least cite a study (eg. Kuipers Munneke et al., 2012) that presents melt trends

P14 L246 and after: SSP scenarios are only mentioned using their first category (SSP5 instead of SSP5-8.5). Since there are several under scenarios in each category, keep mentioning the full name to remain clear.

Specific comments and stylistic suggestions

P1 L4 : maybe « climate models » in general is enough than only ESM.

P1 L14 : I suggest to replace « a coupled Antarctic Ice Sheet » by an « coupled ice sheet model » as it's not the real AIS that will be integreaed into CESM3. This is only a suggestion which the authors can obviously accept or refuse.

P1 L24 : I agree about the stronger regional warming over these regions but the references are not adequate. The mass losses in West Antarctica are mainly due to ocean warming and not to the atmosphere that the references refer to. Increasing air-temperatures are more likely to contribute to hydrofracturing over the AP and subsequent glacier speed-up, but this is still a small contribution against the total mass loss over these two regions. Please reformulate/change your references.

P2 L30 : Consider to remove « Studies have shown that »

P2 L32 : Barthel et al. 2020 do not discuss the SMB uncertainty.

P2 L54 : Add a reference (Gorthe et al., 2020?)

P3 and P4 (Section 2.1 and 2.2): Do you use a specific member for the comparison or also the average of the 11 members ?

P4 L106-108 : I suggest to specify that the SMB of the RCMs (and CESM2?) also includes the runof.

P4 L111-112 : « The » reconstruction is perhaps a little over-emphasized given that the other products (MAR and RACMO) also give reconstructions. (Again a suggestion, feel free to take into account or not). I'd suggest to refer to something like "the SMB reconstruction of Medley and Thomas (2019)" (or any abbreviation like MT2019 reconstruction).

P6 L132 and 136 : Consider to replace « affect » by « effect ».

P7 L139-149 : Are the temperature trends in ERA5 reliable ? If I'm not mistaken, most evaluations (eg., Gossart et al., 2019) only assessed the mean climate and not the trends. I would like more discussion on the potential reasons for these differences. Perhaps just mentioning that CESM2 is not constrained would be enough. Do you have a simulation where CESM2 is constrained that you could also compare to ERA5 (or AWS if ERA5 is not reliable) ? (see also the minor comment about trends above)

P11 L205-204: It's confusing that CESM2 SMB is significantly greater than RACMO SMB (1997 Gt/yr) but not significantly greater than "the reconstruction" (1953 Gt/yr). I guess this come from the large variability in the reconstruction. Do you know why this variability is so large? Is the variability computed on the same period (as all the other products have almost the same variability)?

P13 L236-240: Consider to divide the sentence in several ones to make it clearer.

P14 L246: Specify if you're presenting temperatures over the (grounded or full) ice sheet or over the regions.

P14 L247: "the first ten years of the future scenario (2015-2025) to the final ten years of the scenario (1990-2100)" Is there a mistake for the second period? (Shouldn't be 2090-2100?). Note that changes are more often compared to a selected period over the historical period than over a "future" period (I mean by "future", after 2014 where the scenario is no more the "historical" concentrations). The choice of the period should be consistent with P14 L268. Furthermore, are 10 years representative of the climates of both the "historical/start of the future period" and the end of the century?

P14 L250: Could you explain these differences? Are they due to the inertia of the system?

P14 L270: This is a really interesting analysis. The negative SMB in summer for all the scenario suggests high runoff values and in general strong melt and melt ponds. Since runoff indicates remaining liquid water at the surface (sometimes considered to be a proxy of potential hydro-fracturing – Donat-Magnin et al., 2021; Gilbert and Kittel, 2021), this might suggest that even for the low-emission scenario, surface melt could lead to severe damages over the ice shelves and strongly contribute to their disintegration with large consequences for the ice sheet stability. Maybe you could discuss/mention this in your manuscript.

Appendix : Change the order of the figures to match their order of appearance in the manuscript.

Figures (clear and adapted. I particularly appreciated Fig6.) For Figures 3b and 7c indicate in the caption what crosses represent.

Reference in this review:

Donat-Magnin, M., Jourdain, N. C., Kittel, C., Agosta, C., Amory, C., Gallée, H., ... & Chekki, M. (2021). Future surface mass balance and surface melt in the Amundsen sector of the West Antarctic Ice Sheet. The Cryosphere, 15(2), 571-593.

Gilbert, E., & Kittel, C. (2021). Surface melt and runoff on Antarctic ice shelves at 1.5 C, 2 C, and 4 C of future warming. Geophysical Research Letters, 48(8), e2020GL091733.

Kuipers Munneke, P., Picard, G., Van Den Broeke, M. R., Lenaerts, J. T. M., & Van Meijgaard, E. (2012). Insignificant change in Antarctic snowmelt volume since 1979. Geophysical Research Letters, 39(1).