

## ***Interactive comment on “What is the Surface Mass Balance of Antarctica? An Intercomparison of Regional Climate Model Estimates” by Ruth Mottram et al.***

### **Anonymous Referee #3**

Received and published: 5 March 2020

Review of Mottram et al., What is the Surface Mass Balance of Antarctica? An Intercomparison of Regional Climate Model Estimates

#### Summary

The authors present an intercomparison exercise of five different regional climate model surface mass balance estimates, as well as the near surface climate, over Antarctica. The authors find a large spread in total SMB (1961 to 2519 Gt year<sup>-1</sup>), which largely stems from differences in West Antarctica and the Antarctic Peninsula. Variability is quite consistent between models, which is unsurprisingly since they are all forced by ERA-Interim, but the trends differ in sign and magnitude and are quite sen-

C1

sitive to the time period selected. Also, not surprisingly, the nudged models simulate the near-surface climate better as they are not allowed to deviate as substantially from ERA-Interim as the un-nudged models. Finally, the authors discover that the biases are typically consistent between models. The paper presents a significant amount of work but still requires improvements. First, the manuscript has numerous mistakes throughout and needs refinement of the language in several places (see Minor Comments). More importantly, there are several major issues with the analysis that need to be addressed to improve the scientific rigor of the paper.

#### Major Comments

1. Throughout the manuscript, it is not clear what time periods are being used. There is the common model interval, climatological mean, etc. The authors need to be very clear throughout the paper because it often seems that different intervals are being confused in nomenclature. It's not clear to me why the common reference interval is not the common period between all models: 1987 – 2015. Throughout the paper sometimes its 1980-2010, 1987-2015, and 1987-2018. I recommend using the same interval through to avoid confusion. If the authors have a reason to use different intervals, then please make it clear what interval is being used. It is additionally unclear why 1980-2010 is representative of the climatological period, please explain.

2. Similarly, there is no discussion of significance for the statistics presented. There are claims within the text that certain models perform better than, but without significance levels, these claims lack strength and are more speculative. Trends are discussed at both long (1987-2015) and short (decadal) time intervals, but the significance is never discussed. I would caution the authors descriptions of trends, especially at short time scales, since it is very hard to observe a significant trend in SMB since its highly variable year to year. Furthermore, because this is an intercomparison paper, its important for the authors to be very clear concerning the metrics of how they conclude one model outperforms the other. Is it RMSE? R<sup>2</sup>? Bias? And what is the threshold? Is an RMSE of 93 better than 97? What if one model performs differently at different el-

C2

evaluation bands? I did not find the argument compelling that the models tuned to specific Antarctic conditions outperformed the others because there was not a clear framework for comparison. The authors need to make clear the evaluation metrics and how they evaluate model performance, which will require more detailed statistical analysis throughout. Model means are compared, but its not clear if the paper considers even a simple statistic of the standard error of the mean. The Student's t-test can be used to evaluate whether the means are different. Please be transparent with the limitations of the analysis and provide meaningful significance tests on all of the comparisons, otherwise the conclusions are speculative rather than significant.

3. All of the RCMs presented are forced by the ERA-Interim reanalysis product. I find it concerning that there is no discussion of the role of using a single reanalysis to force all of the RCMs. Thus, this is not a definitive evaluation of the full range of possibilities in SMB, but rather a range due to RCM differences alone. There should be more discussion about how there would be additional spread due to varying choice of forcing; specifically, what is the impact of comparing models that are all forced by the same reanalysis. I think the paper needs to tone down the claims about the work providing the "likely range of SMB" in the first sentence of section 4.1, as it is more the likely range of RCMs forced by ERA-Interim. Basically, this explores the range in RCM space, but not reanalysis forcing space.

4. It's obviously quite a challenge to compare these models, which have differing levels of complexity. But it seems that the comparison would be better suited by comparing all the variables consistent between models (Precip-Evap-Subl). Otherwise, an inter-comparison doesn't shed much light on direct model to model differences. In fact, it appears that the authors could investigate whether these Antarctic specific physics actually provide improvement, which would be of great interest to the community. Therefore, the paper should do an ideal comparison of all 5 models with common variables (P – E – S) and evaluate performance. Then evaluate the models with extra physics (RACMO/MAR) to see if and how much model performance improves. Otherwise, its

C3

difficult to untangle whether those additional processes provide any more information.

5. The manuscript needs to justify the use of SMB observations starting in 1950. There are regions of strong trends in snow accumulation that might end up biasing the comparison. If the issue relates mainly to reducing the number, the authors could present a comparison of only coincident SMB observations with the data, but then also provide the more liberal comparison as it currently exists in the text.

6. Finally, the paper needs to discuss the impacts of its findings. With the present day mass loss from Antarctica on the order of 100 Gt per year, this is quite concerning finding the differences in SMB from RCM choice alone are several hundred Gt per year. Please contextualize the findings in regard to how we can measure the mass balance of the ice sheets.

#### Minor Comments

Several model names and versions are discussed before they are described, which makes it quite hard to follow. Please reorder the sections to ease. For instance, section 2.1 and the end of Section 1 mention several models and different version, but there is no description, so its hard for the reader to follow. It would also be appropriate to cite the papers that refer to these model versions.

P1, Line 7: Is this for grounded ice only? Does it include islands and ice shelves?

P1, Line 7-8: What do the values after the  $\pm$  represent? The standard deviation of all the models?

P1, Line 10-11: Why is 1980-2010 chosen as the climatological period? Later in section 2.3.3, it appears that 1987-2015 is the common modeling period that is used to "compute a climatological mean" (P10, Line 12). Please rectify.

P2, Line 1: change "compar" to "compare"

P2, Line 11: remove either "potentially" or "potential" since its repetitive

C4

P2, Line 13: add "and" after "2002,"

P3, Line 16: remove the comma after "published"

P3, Line 23: remove "drive"

P4, Line 1: please describe what a "reinitialized hindcast" is

P4, Lines 5-7: While this is true, it might have a limit. See Lenaerts et al., 2018, which shows that often the snow is not dumped in the proper place when moving from 27 km to 5.5 km. Please add a sentence clarifying this.

P4, Line 8: add "is used" after "ensemble mean"

P5, Line 6: change "developed in" to "developed for"

P5, Line 18: the end of this sentence needs to be reworded

P6, Line 4: Do you mean "processes" not "process"?

P6, Line 22: change "includes no" to "does not include"

Table 1: What does SMB scheme mean?

P7, Line 10: change "schemes" to "scheme"

P10, Line 10: add "that are" after "2015 and 2018"

P10, Lines 14-20 need clarification

P12, Lines 4-7: This sentence is very long and needs to be split in two.

P12, Line 13-14: Please reword the sentence as its confusing.

P13, Line 6: Remove "In"

Figure 3. Please add the statistics to these plots (RMSE, etc.). Also, its very difficult to distinguish the colors here. Maybe large dots would help.

C5

Figure 4. This figure should be much bigger. It's very hard to see the colors. Also, in the caption there are "a", "b", etc., but they do not exist on the plots.

P18, Lines 9-11: are these values consistent with what is listed in the abstract?

Figure 5. This needs to be in landscape orientation. The numbers are much too small to read.

P20, Lines 5-6: What does "much clearer mean," please clarify

Figure 6. Again, these plots are too small, and the numbers are nearly impossible to read. P22, Line 2: remove "below"

There should be significance values associated with the trends. It looks like none would be statistically significant and thus are effectively no different than zero.

P24, Line 6: remove "very"

P24, Line 9: change "bring" to "brings"

P25, Line 7: Should the interval be 1987-2015?

P27, Line 29-30: Do your results actually support "Models that have not undergone specific adjustments for Antarctica clearly represent the SMB in Antarctica more poorly". Look at the RMSE value in Table 2, it looks liked sometimes they perform better. Please clarify.

P 28, Line 11-12: please give values in Gt of these processes to show that they are effectively negligible

P 28, Line 20: add "fig." before "7"

P28, Lines 20-21: the sentence needs to be improved.

P29, Line 2: replace "mod-latitudes" with "mid-latitudes"

---

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2019-333>, 2020.

C6