Review: Physics-based modeling of Antarctic snow and firn density

Eric Keenan, Nander Wever, Marissa Dattler, Jan T. M. Lenaerts, Brooke Medley, Peter Kuipers Munneke, and Carleen Reijmer

Summary
The authors apply the physically-based SNOWPACK snow model over the Antarctic ice sheet in order to simulate snow and firn densification. The SNOWPACK simulation is compared with in situ measurements and two other firn densification models. The authors find that biases in SNOWPACK are generally lower than in the two other empirically-based models, especially for locations where observations have not been used to calibrate the semi-empirical models. They suggest that in future projections of Antarctic firn densification, SNOWPACK would produce more reliable results, because of more detailed representation of physical processes, compared with simpler semi-empirical models.

General Comments
In general, the manuscript is well written and well organized. The topic is an important and relevant one, especially given the recent launch of the ICESat-2 altimetry satellite. The scheme introduced is more sophisticated than that of other models applied over the Antarctic ice sheet, and the paper therefore represents an important advance on other recent studies. The paper clearly confirms that the SNOWPACK model is capable of realistically simulating near-surface density over the Antarctic ice sheet, and has advantages over other models in that it is more detailed in its representation of physical processes governing snow evolution. I do have some concerns regarding the interpretation of results, however, as noted below. I feel that overall these revisions do not require major changes to the paper, but should be addressed before the paper can be published.

(1) Given that the uncertainty ranges of the density simulated by the different models overlap to some degree, it is not completely clear whether there is a statistically significant difference between them, or between the models and the observations at different levels. The authors should test whether this is the case.

(2) The authors should be careful to note some of the limitations of the current implementation (e.g. the validation is over the top 10 m, not the entire firn column; and the SNOWPACK bias is larger below 6 m depth) particularly in the abstract and conclusions sections.

(3) The available evidence doesn't seem to necessarily support the argument that biases are substantially larger in the semi-empirical models at locations that were not used to calibrate those models. The authors should clarify whether this is indeed the case and revise the text accordingly. It would be interesting to include both the GSFC-FDM and IMAU-FDM in this comparison if possible. Further specific comments are provided below.

Specific Comments

1. **Title:** I would suggest adding SNOWPACK to the title, and mentioning the near-surface e.g. "Physics-based modeling of near-surface Antarctic snow and firn density with the SNOWPACK model". I would argue that the other models utilized here are also physically-based, they just employ simpler parameterizations for the process of firn densification.
2. **Lines 1-11:** In general, some quantitative evidence should be provided here. Some of the limitations of SNOWPACK applied over Antarctica should be discussed, for example the larger bias for higher accumulation areas and the larger biases deeper in the snowpack, as well as the fact that this approach focuses on the near-surface, not the full firn column.
3. **Lines 7-8:** It would be best to quantify the magnitude of the biases here.
4. **Line 9:** It isn't entirely clear from this sentence that this is one of the findings of the study; it would be best to provide some quantitative results here. Also I believe the performance does degrade somewhat at these sites, just not as much as for the semi-empirical models?
5. **Line 17:** It would be informative to mention other methods of estimating mass balance (e.g. gravity measurements, e.g. Velicogna et al., 2020; or the input output method, e.g. Rignot et al., 2019).

   Velicogna, I., Mohajerani, Y., A, G., Landerer, F., Mouginot, J., Noël, B., Rignot, E., Sutterley, T., van den Broeke, M. R., van Wessem, M., and Wiese, D. (2020) Continuity of ice sheet mass loss in Greenland and Antarctica from the GRACE and GRACE Follow-On missions. *Geophysical Research Letters* 47, e2020GL087291.

   Rignot, E., Mouginot, J., Scheuchl, B., van den Broeke, M., van Wessem, M. J., and Morlighem, M. (2019) Four decades of Antarctic Ice Sheet mass balance from 1979-2017. *Proceedings of the National Academy of Sciences*, 116, 4, 1095-1103.

6. **Line 38:** What is meant by "all local and temporal density variability"? No model can capture "all" variability. Please clarify.
7. **Line 41:** These models do employ "physical principles"; they are not entirely empirical. Suggest simply removing the phrase "rather than physical principles".
8. **Lines 50-53:** Describe how the model is forced, briefly.
9. **Line 50:** Instead of "we apply", do you mean "we compare results from">
10. **Line 61:** SNOWPACK also seems to include parameterizations that are empirically based. Perhaps mention explicitly how SNOWPACK is different from the other models mentioned in earlier sections.
11. **Line 75:** Perhaps change "new drifting snow compaction routine" to "new snow compaction routine", as drifting snow is just a component of the routine.
12. **Line 80:** Can the authors briefly note how this parameterization is derived?

13. **Lines 87-88:** How much do these parameters change the comparison with observed profiles. Provide some additional details either in the main manuscript or a supplemental section.
14. **Line 90**: Briefly explain the physical meaning of the "threshold friction velocity".
15. **Line 124**: Is this a bias over the entire Antarctic ice sheet? Are there spatial variations in the bias?
16. **Lines 130-131:** Is there a reference for these statements?
17. **Line 133:** Why use 19.4% and not 15.1 W m^{-2} ?
18. **Line 134:** Why is there still a bias after the bias is removed?
19. **Lines 154-170:** It would be helpful here to describe these two models in a bit more detail, in particular to highlight how they differ from SNOWPACK in terms of key physical processes (e.g. compaction), as the model differences are important to the conclusions of the study.
20. **Line 166:** Explain the meaning of "replay".
21. **Line 194:** Suggest changing "reduction in both RMSE…" to "statistically significant reduction in both RMSE…"
22. **Line 196:** This section could potentially be moved to later in the manuscript. It might logically follow the section on comparison with observations.
23. **Line 200:** Clarify why these two stations were chosen.
24. **Line 221:** It is a bit unclear what is meant by "we tested for explanatory variables". Please clarify.
25. **Lines 245-247:** It might be useful to have a table here for the bias and RMSE for different models above and below 400 kg m^{-3}.
26. **Line 253:** This sentence is confusing. Suggest revising to read something like: "Additionally, we cannot rule out the possibility of larger errors in the observational data for densities above 400 kg m^{-3}."
27. **Line 256:** The SUMup dataset does include information on measurement methods. It might be interesting to see if dividing by measurement method changes these biases in any way.
28. **Lines 259-271, Fig. 6:** Can the authors note whether the differences are statistically significant? It might also be useful to provide an uncertainty range on the biases. Also, at first glance at it appears that all the model simulation uncertainty ranges overlap in Fig. 6, but this is not the case. Perhaps the figure can be modified slightly to make this clearer, e.g. changing the transparency for different models or changing the colors. (Not sure how easy this would be).
29. **Lines 293-294:** It seems this would not be difficult to find out? It would also be interesting to see the IMAU-FDM results.
30. **Lines 296-297:** From Fig. 8, it actually looks like there is a larger change in the SNOWPACK density bias (at least at different levels). The numbers here do not seem to match with the figure. Please clarify.
31. **Lines 335-336:** This portion is interesting but seems disconnected from the rest of the manuscript. Perhaps these temporal variations could be placed in the context of temporal variations from in situ data. Are there any locations

where a timeseries of measurements is available that could be compared with the SNOWPACK runs?

32. **Line 342:** Without validation of the temporal variability of the in situ measurements, I'm not sure the model results would qualify as "evidence". Please revise.

33. **Lines 360-364:** I'm not sure these statement is completely supported by the results. For example, SNOWPACK seems to show a larger bias at higher accumulation locations, and the SNOWPACK and the GSFC-FDM both seem to show a positive bias in locations that were not used to constrain GSFC-FDM between 0 and 6 m in depth. In general, however, I would agree that including a more physically realistic simulation of snowpack processes should produce a better projection of future conditions. Perhaps revise this statement to note that this is likely the case, but not entirely certain.

Technical Corrections

1. **Line 14:** Change "with an increasing…" to "at an increasing…"
2. **Line 91:** Change "from MERRA-2" to "from the MERRA-2".
3. **Fig. 2 caption:** Change "SNOWPACK simulations" to "SNOWPACK simulation locations" for clarity. Note that the borehole depths are 10 m below the surface for clarity.
4. **Line 172:** Change "as depths" to "the average density between depths of"
5. **Line 185:** Change "of average" to "of the average".
6. **Line 189:** Suggest changing "almost perfect" to "an excellent".
7. **Line 198:** Change "is bias-corrected MERRA-2" to "in bias-corrected MERRA-2".
8. **Line 215:** Change "windspeed represents" to "density variations due to uncertainty in windspeed represent".
9. **Line 219:** Change "observations" to "observed density values".
10. **Lines 259-260:** This sentence is quite wordy. Suggest revising, e.g. "In a comparison at 122 observed density profiles, SNOWPACK exhibits a lower bias compared to IMAU-FDM for the entire near-surface, and a lower bias compared to GSFC FDM between from 0 to 7 m depth (Fig. 6)".
11. **Lines 273-274:** Suggest changing to read "low SMB categories by applying a threshold of 200 kg m^{-2} yr^{-1} to MERRA-2 mean annual SMB (Fig. 7)."
12. **Line 276:** Change "reduced" to "lower" or "smaller"
13. **Line 277:** Again change "reduced" to "lower" or "smaller".
14. **Caption, Fig.** 7: Change "GSFM-FDM" to "GSFC-FDM". Change "a MERRA-2 1980-2017 mean annual SMB threshold of 200 kg m^{-2} yr^{-1}" to "a 200 kg m^{-2} yr^{-1} threshold applied to MERRA-2 1980-2017 mean annual SMB."
15. **Lines 303-304:** Change "as well as their different level of process representation complexity" to "as well as their different level of complexity in representing physical processes."
16. **Line 325:** Change "compared" to "compared".
17. **Line 373:** The heading for Appendix A is out of place.