Review of
Characteristics of the contemporary Antarctic firn layer simulated with IMAU-FDM v1.2A (1979-2020)
by Veldhuisen et al.

Reviewer: Vincent Verjans

This study presents a modeling approach to simulate the evolution of the Antarctic firn layer. It presents the re-calibration of a firn model (IMAU-FDM), with forcing from an updated regional climate model (RACMO2.3p2). The calibration mostly follows the same approach as in the initial model release but with a more extensive dataset of firn cores and surface snow density measurements. Based on a 1979-2020 model simulation, the study presents an analysis of spatio-temporal features in firn thickness, firn air content (FAC) and surface elevation change over the entire Antarctic ice sheet. I welcome this contribution to firn model improvement. The modeling capabilities and the full coupling between climate and firn models are two remarkable aspects of this study. The model pair IMAU-FDM – RACMO is often used for evaluating the firn height change component in altimetry studies, and it is thus paramount to accurate estimates of ice sheet mass balance change by the glaciological community. But this study often gives the feeling to be a simple update of the work of Ligtenberg et al. (2011). For this reason, I find this study slightly too superficial, and several points can be evaluated and analyzed in more details. Nevertheless, I believe that, by building upon the work already achieved for this first version of the manuscript, a future revised and updated version will be a good contribution to *The Cryosphere*, and to firn model development in general.

This review is separated in Major, Minor, and Specific comments. My Major comments highlight current weaknesses of the studies that require a change in the methods or more in-depth work. My Minor comments require more clarity, small modifications, and/or strong justifications from the authors. The Specific comments are remarks concerning specific statements in the manuscript, and mostly relate to the structure of the text.

Major comments

1) Uncertainty quantification
As mentioned in the introduction, IMAU-FDM is commonly used in ice sheet mass balance assessments. In addition to this, the firn component of elevation changes is often cited as a major component of uncertainty in altimetry-based mass balance assessments. Combining these two aspects together, it is obvious that our community needs better uncertainty estimates from firn models themselves. The uncertainty quantification in this study is, in my view, not sufficient. The authors present many components of uncertainty in Section 2.6, as well as ways to estimate their contribution to IMAU-FDM output uncertainty. However, it is not clear how these points highlighted for a sensitivity analysis are subsequently used for uncertainty quantification in the results.

1a) Results of the sensitivity analysis
Many of the possible experiments presented in Section 2.6 are not even mentioned in the results section about uncertainty (Section 6). In Section 2.6, the authors mention various scenarios of accumulation and temperature reduction for the spin-up period (three for each variable). However, they only discuss a single test per variable in Section 6, and it is not specified which one it is. This obviously requires more clarity.
Similarly, the sensitivity tests using "*the 95 % confidence intervals of the MO fits*", and the "*uncertainty of the fresh snow density*" are not discussed. I strongly recommend to clarify the links between Sections 2.6 and 6. For example, each sensitivity test should be given a name. And there should be a Table that specifies the sensitivity tests, their corresponding variable adjustments, and results.
The Table 4 needs some further adjustments. I suppose that the FAC values in Table 4 are calculated only over the dataset of firn cores, and not over the entire ice sheet, which should be specified in the caption. Table 4 should also include the results from the spin-up perturbation experiments: 3 per variable (see Section 2.6), and the experiments combining temperature and accumulation perturbations (not explained in Section 2.6).
Finally, the sensitivity experiments use ice sheet wide averages of precipitation spread (+/- 8 %) and of temperature uncertainty (+/- 1.5 K). However, local uncertainties can differ strongly from the ice sheet averages. Where possible, I recommend using spatially variable uncertainties in the sensitivity analysis.

1b) Ice-sheet wide method for uncertainty quantification

The authors have investigated the results of the sensitivity experiments only at the firn core locations. However, this is of little interest to the community. A method to compute uncertainty estimates over the entire ice sheet in the different components of Eq. (6), and thus ultimately on dh/dt, should be developed. A straightforward approach could be to sample perturbations in all uncertainty sources mentioned in Section 2.6, compute simulations at key locations, and regress uncertainty estimates against climatic variables. However, I leave the choice to the authors of how to best estimate uncertainty from their model, and I note that they have already worked on similar issues (e.g., Kuipers Munneke et al., 2015). The outcome of a more thorough uncertainty analysis should be:

- uncertainty bands in time series of Figure 7 and Figure 9
- maps of total uncertainty in modeled surface elevation change (and possibly of the uncertainty components)
- a quantification and discussion of the different components to uncertainty in dh/dt across Antarctica

2) Statistical procedure

In statistical calibration of parameters such as the MOs, it is well-known that validation should be, to some degree, independent of calibration. Here, the authors use the same dataset to calibrate their new parameters (MOs and surface density) as to evaluate their calibration. As such, the validation is not meaningful. I realize that firn data is sparse, and it can be argued that excluding a part of the dataset from calibration might be detrimental to model fitting. But that does not exclude a form of k-fold cross validation to better evaluate the fits. Furthermore, this approach would provide more robust uncertainty estimates on the MO and surface density values.

For the evaluation, the authors use only 10 additional measurements in the evaluation dataset, in addition to the ones used in the calibration dataset. This modest introduction of independent data in the evaluation results in (line 243): "*a slightly deteriorated correlation for the* MO_{550}." when compared to the FDMv1.1p1 parameterization. Furthermore, in Section 6, the authors demonstrate that the calibration of the MOs is sensitive to realistic uncertainty in climate forcing, and that this sensitivity strongly impacts results in FAC, z_{550}, and z_{830} (Table 4). These two aspects thus show evidence that the statistical fits performed in this study are probably sensitive to noisy features inherent to the observations used in the calibration.

In Table 3, I recommend providing the RMSE values in all the rows, to give the bias values in addition to the RMSEs, and to compute these fit statistics also with respect to the FAC values of the dataset.

Finally, if I understand the process correctly, the comparison of fits between FDMv1.1p1 and FDM v1.2A for z550, and z830 is unfair. FDMv1.1p1 was calibrated with climate output of a previous RACMO version and another surface density parameterization. As such, it is obvious that it performs worse than FDM v1.2A when it is used with RACMO2.3p2 forcing and another equation for surface density. Furthermore, it is unclear how much of the improved fit to observations is due to the re-calibration of IMAU-FDM versus the updated climatic forcing and the updated surface density parameterization. All these aspects should be addressed in the manuscript, and identified as caveats in the comparison. The same holds when comparing FDMv1.1p1 and FDM v1.2A to the altimetry product (Figure 8): it remains unknown what part of the improvement is due to changes in the densification equation, changes in the climatic forcing, and changes in the surface density parameterization. As pointed out by the authors, Figure 9 suggests that most of the improvements are due to the update of RACMO, which raises questions concerning the improved performance of FDM v1.2A in simulating firn processes. This should be discussed more in depth.

3) Neglect of melt areas

The MO parameterization is constrained only in dry firn areas, but the model is used in wet firn areas also. One can reasonably expect errors to be much larger in modeling wet firn densification. While this topic is not the focus of the study, any ice-sheet wide study of firn evolution should at least discuss this limitation. Ideally, I would encourage the authors to evaluate IMAU-FDM in the wet firn areas also, by comparing modeled FAC to observed FAC from firn cores. They could also provide uncertainty estimates that are valid for melt areas.

Minor comments

1) There is a general lack of quantification. I encourage the authors to identify all the uses of words such as “reasonably”, “somewhat”, “roughly”, “improved”, and “substantially”. These should be complemented by quantitative values.

2) In the densification equation (Eq. (3)), it seems to me that the mean long-term accumulation rate is used. An alternative approach is to use the mean accumulation rate over the lifetime of each specific firn layer, which is more representative of the effect of overburden stress (Li and Zwally, 2011). This aspect could be important given that decadal snowfall variability can be large on the Antarctic ice sheet. And there is no physical reason for the densification of a firn layer to be a function of past accumulation rates. Why did the authors choose to use the mean long-term accumulation rate?

3) Lack of clarity about the firn core dataset
Depsite re-reading several time Section 2.5, it is still unclear to me how the authors selected their dataset.
- “*We used 125 density profiles from firn cores and 8 density profiles from neutron density probe measurements*”: but the wet firn cores were discarded for the calibration, so how were they used in this work?
- “*For the MO fits, 104 dry firn cores could be used (…) To evaluate the firn density profiles from the simulation using the derived MO fits, 122 firn cores could be used.*”: what explains this difference of 18 cores?
- in Section 2.6 “*105 observational locations shown in Figure 1*”: 105 is not even mentioned in Section 2.5.
- in caption of Figure 1 “*The grey circles indicate ten additional locations that were included in the sensitivity analysis*”: give more details about these 10 additional firn cores in Section 2.5.
- Concerning access to the firn core data, please see the *Data policy* section of *The Cryosphere*: “*Authors are required to provide a statement on how their underlying research data can be accessed. This must be placed as the section "Data availability" at the end of the manuscript.*”

4) Interpretation of FAC change versus mass change
I believe that there is a confusion when interpreting the role of FAC change for conversion of elevation changes to mass changes. For example, on line 328: “*63 to 68 % of the seasonal surface elevations fluctuations are caused by a change in air content rather than actual mass change*”. However, FAC changes are principally caused by changes in snowfall, which implies a corresponding mass change. Thus, there is an underlying mass-related component to fluctuations in elevation caused by FAC changes.

5) I find that there is a general lack of explanation about some model results. I give some examples here below.
- In Figure 7, FAC variability is smaller than snowfall variability. What explains the dampened response of FAC?
- Figure 6d shows that there is a lot of spatial variability in the phase of firn height. This is an interesting result, and some explanation should be provided for why such patterns appear.
- The seasonal amplitude comparison with altimetry (Section 5.1) is averaged across the ice sheet. Are there no interesting patterns that appear at finer scales? Furthermore, the authors state that (line 374): “*The performance of the new model thus appears to represent an improvement*” when comparing FDMv1.1p1 and FDM v1.2A. But because they only analyze the amplitude averaged over the entire ice sheet, the better performance could be due to error compensations in different areas. A finer-scaled analysis is therefore required.
- The evaluation against the surface altimetry data is not sufficient. There should be more focus on regional patterns and on relations between discrepancies and climate among other things. And especially, the comparison should be much more quantified (see Minor comment 1).
- On line 323: “*This agrees with the modelled firn thickness amplitude of 3.5 cm by Medley et al. (2020).*”. This is despite the lower accumulation rates than in the study of Medley et al. (2020). Thus, how is that compensated? Is IMAU-FDM more sensitive to seasonal temperature variability?

Specific comments

-l4
"*improved*": specify that this refers to a previous model version.
-l5
"*observations*": change to "firn core observations".
-l15
The reference cited shows that firn thickness can exceed 100 m.
-l17 – 19
"*Firstly, firn depth and density estimates are required to convert altimetry observed volume-to-mass changes, which remains a major source of uncertainty in mass balance studies*": please also reference the work of Morris and Wingham (2015).
-l27 – 29
Please reference the work of Arthern and Wingham (1998).
-l27
"*are used as measures of its dynamics and mass balance*": clarify and rephrase.
-l30
Change "*in a mass and density component*" to "in a mass- and density-change component".
-l32 – l37
Is this paragraph only about ice shelves. If yes, this should be specified in the first sentence. If not, the reference to ice shelf hydrofracturing is confusing.
-l33
I do not believe that there is evidence of "*reduced accumulation*" for a warmer future climate in Antarctica.
-l33
Use "potentially lead to".
-l40
Please reference the work of Medley et al. (2015).
-l44
"*Firn models can roughly be divided in two classes: physically based and semi-empirical models.*": please reference the work of Morris and Wingham (2011).
-l46
It is not clear to me what "*which*" refers to.
-l47
Change "*less poorly known parameters*" to "a smaller number of poorly constrained parameters".
-l49
Please reference the work of
-l56
Start the sentence as: "This study shows that, at the basin scale, (…)".
-l58
Add comma: "climatic conditions, firn densification".
-l63
Specify the sort of "*field measurements*".
-Table 1
In the column "*Other*", please refer to Equation numbers.
-l75
"*FDM V1.2G*" has not been defined yet. Maybe, refer earlier to Table 1.
-l77
"*comparable*": with respect to what metric?
-Section 2.1.1 and in the remainder of the manuscript
Please do not use the same symbols for different parameters. If the authors want to preserve a connection between related parameters, I recommend the use of subscripts.
-l83

Change “*average*” to “averages of”.
-l86
Cite a reference for “*Snow crystal size and therefore fresh snow density indeed increase with increasing temperature.*”.
-l94
Specify the frequency of “*instantaneous*”: hourly/daily/…
-Equation 2
I believe that there should be no overbar on T_s and V_{10} in this equation.
-l101
“*defined as the top 0.5 m*”: am I correct that the authors calibrate the density of the modeled upper 0.5 m to the density of the observed upper 0.5 m, but that the calibrated surface density is then used only for the top layer of IMAU-FDM? If so, please clarify the approach as well as the slight discrepancy between calibration and usage of the surface density parameterization.
-l102
Specify the model time step here or elsewhere.
-Section 2.1.2
All equations should be specified as two different cases for rho<550 and rho>550 kg m^{-3}.
-l110
I suggest replacing “*processes*” by “mechanisms”.
-l111 - 112
A citation is needed for this sentence.
-l114
Change “*turn out to depend on the accumulation rate*” to “are chosen as functions of the long-term mean accumulation rate”.
-l120
Why was the power-law function not tested for MO_{550}?
-l124
Provide formulation of the thermal conductivity.
-l128
Provide formulation of the irreducible water content.
-l129
The description of the refreezing algorithm is unclear and confusing. It suggests that no meltwater refreezes if a layer cannot accommodate all the incoming meltwater.
-l131
Typo: remove “*the*”.
-l131
Specify the amount of melt.
-l131
Do not use “*significant*” because it does not refer to statistical significance here.
-l134 – 135
What is the depth of the model domain? And what are the boundary conditions at the lower boundary?
-l137
I believe that “*total thickness*” should be replaced by “total mass”.
-l142
Please quantify the “*minor trend*”.
-l142
Change “*ice*” to “material with ρ>830 kg m^{-3}”.
-Equation 6
Format of the variables in the equation does not correspond to format of the variables in the main text.
-l152
SMB units are wrong.
-l160
Change “;” to “and”.

-l161
Explain briefly the notion of "*upper-air relaxation*".
-l168
Split this sentence in two: "This results in an improved forcing. For example, (…)".
-l169
Remove "*e.g.*".
-l178
Typo: "describes".
-Figure 1
As I understand it, all the firn cores are used for the sensitivity analysis. For this reason, I recommend changing the label for the grey dots in the legend to "Sensitivity analysis only".
-Figure 1 caption
Replace "*on top of*" by "in addition to".
-Section 2.6
Provide a name for each sensitivity experiment (see Major comment 1).
-l197
"*To improve the representation*": I do not see the causal link between improvement and the rest of the sentence.
-l203 – 204
Change "+-" to +/-".
-l204
Does +/- correspond to the RACMO2.3p2 RMSE? If so, please specify this.
-l206
Specify that 30 kg m^{-3} corresponds to the RMSE in surface density.
-l206
"*Section 3.2*" should be "Section 3.1".
-Table 2
There is no "*D*" parameter in Eq. (2) for FDM FS-L and FDM v1.2A.
-l218 – 219
Refer to Equation numbers.
-l231
Change "*reduced with*" to "reduced by".
-l240
"*a simulation of FDM v1.2A without MO fits*": does that mean a simulation with the FDMv1.1 MO values?
-l241
"less steep": specify that this is with respect to accumulation (bdot).
-l253
Replace "*approaches zero*" by "decreases asymptotically towards zero".
-l255
Provide references for the accumulation rate values.
-Table 3
In the column "*Version*", I believe that MO_{550} and MO_{830} should be replaced by z_{550} and z_{830*}. In the column "*Fit*", please refer to Equation numbers. Add a Bias column (see Major comment 2). Provide RMSE values for all models (see Major comment 2). Add FAC rows (see Major comment 2).
-l258
Specify "27 km horizontal resolution".
-l259
Typo: "in Figure 4".
-l261
What does "*calm*" mean?
-l262
Change "*On top of this*" to "In addition to this".
-l264 – 265

“*The spatial pattern of the depths of the critical density levels z550 and z830 are shown in Figures 4b and c, and are roughly the inversed pattern of the surface snow density*”: not in melt areas, please discuss (see Major comment 3).
-l266 – 267
“*The patterns vary spatially across climatic regions with temperature as a primary driver and accumulation as a secondary driver.*”: please discuss the impact of surface melt (see Major comment 3).
-l270 – 271
Please rephrase this sentence with more formal language.
-Figure 3
I believe that there is a mismatch between the numbers of points shown in the subplots, and the number of cores mentioned in Section 2.5. For example, it seems to me that there are less than 100 data points shown in Figure 3a. Also, please show scatter plots of the match between modeled and observed z_{830*}, as well as between modeled and observed FAC.
-l281 – 282
Change “*can be used*” to “must be used”.
-Figure 4 caption
Change “*firn age of the crtical density level*” to “ firn age at the crtical density level”.
-Figure 5
Define “*peak-to-peak*” in the caption. Please also include a map of average seasonal amplitude.
-l286
Change “*most*” to “more”.
-l286
Typo: “parts of the coast”.
-l292
If statistical significance has not been tested for, please do not use “*significant*”.
-l294
Replace “*closer to the mean*” by “less spatially variable”.
-l303
Please remind the reader about the study period.
-l303 – 304
“*Large values indicate that seasonal and interannual climate variability cause large temporal variations in FAC.*”: this statement is not supported by the map of peak-to-peak variability, which depends only on two single values in the entire time series. In my view, peak-to-peak does not characterizes temporal variability well.
-l313
V_{tot} does not appear in Equation 6. Maybe simply replace dh/dt by V_{tot} in Equation 6.
-l320
Specify that dh/dt in summer months has contributions from sublimation and melt.
-Figure 6 caption
Change “*indicate the standard deviations*” to “indicate the inter-annual standard deviations”
-l329
Change “*biases*” to “errors”.
-Section 4.4
In general, this section requires much more quantitative assessments (see Minor comment 1).
-l340
“cumulative surface temperature anomaly”: is this the cumulative anomaly in surface temperatures from the long-term mean? Please clarify.
-l341
“*the seasonal firn thickness and FAC variability is driven by*”: please discuss why there is no one-to-one correspondence between firn thickness and FAC variability.
-l341
Typo: “*is*” should be “are”.
-l343 – 344

“*firn densification, despite the long time scale, reduces these snowfall-induced fluctuations by about 15 %*”: please clarify how this is calculated.
-l356
Specify: “captures the strong spatial variation in firn thickness and density observed in our firn core dataset.”.
-l363 – 364
Specify: “shown in grey in Figs. 3d and 3e”.
-Section 5.1
Please provide maps of seasonal amplitude and of discrepancy in seasonal amplitude.
-l368 – 369
Provide references to support that lower seasonal variability in IMAU-FDM can be explained by altimetry errors.
-Figure 7
I find the color codes in this figure confusing. I suggest to show FAC in a color other than blue.
-Figure 7 caption
Change the caption to: “Time series from FDM v1.2A of FAC, of the cumulative anomalies of surface temperature, of the vertical firn surface velocity, and of the separate components of the vertical velocity from Eq. (6). Time series are shown for (a) the entire ice sheet, (b) the part of the ice sheet situated above 2,000 m a.s.l. and (c) the part of the ice sheet situated below 2,000 m a.s.l.”.
-Figure 8
Remove ice shelves from the maps, or show them in a separate color.
-Figure 8 caption
I suggest changing “*Maps of average surface elevation change*” to “Maps of trends in surface elevation”.
-l384
“*which yields the improvement of FDM v1.2A compared to FDM v1.1p1*”: I do not understand why the authors call this an “*improvement*”. As I understand it, the residual of FDM v1.2A is calculated as altimetry minus FDM v1.2A. The residual of FDM v1.1p1 is calculated as altimetry minus FDM v1.1p1. Thus, subtracting the residuals results in FDMv1.2A minus FDMv1.1p1. In other words, it is simply the difference between both models because the altimetry term cancels out in this operation. If I misunderstand something, please clarify. Otherwise, please revise the use of “*improvement*” when referring to the difference between the residuals throughout the manuscript.
-l389
Typo: “*trends*” should be “trend”.
-l394 – 395
Please provide a quantitative justification for why the 11 glaciers chose are “*representative locations*”.
-l401
“*(+9 % sd)*”: is that compared to post-2003 altimetry?
-l401
“*likely related to the measurement imprecision*”: please provide a reference.
-l402
“*(+13 % sd)*”: is that compared to FDM v1.2A?
-l402
“*the altimetry variability remains higher than the simulated variability (+13 % sd)*”: please discuss possible reasons.
-Figure 9
Why do the authors use a 6-months running average? This masks out all the seasonality. I recommend using a shorter averaging window, 3 months for example.
-Figure 9 caption
Change “*altimetry observed*” to “altimetry observations”.
-Section 6
This entire section should be thoroughly reworked (see Major comment 1).
-Table 4
This table should be thoroughly reworked (see Major comment 1).
-l430

"*that our results are robust*": what do the authors mean here? In contrast, I understand from the results that the MO fits are not robust to realistic uncertainties in climate forcing, and that this sensitivity induces strong discrepancies (Table 4) in FAC estimates (see Major comment 2).
-l430 – 432
Repetition: "*A difficulty of the data to model comparison is that (…) but also hampers the comparison.*".
-l434
Typo: "*dependent*" should be "to depend".
-l440 – 441
Repetition: "*in these regions (…) in these regions*".
-l438 – 440
Please reference the work of Medley and Thomas (2019).
-Section Conclusions
Make sure to be consistent in using past or present tense.
-l446 – 448
This statement of improvement is misleading, because FDMv1.1p is forced in this study with climatic forcing that was different than the climatic forcing used for its calibration (see Major comment 2).
-l450
Typo: "*it has*" should be "they have".
-l451 – 452
"*the firn thickness and density patterns vary spatially across climatic regions with with temperature as a primary and accumulation as a secondary driver*": Here, I believe that more nuance is needed. Spatial variability in firn thickness is primarily dictated by accumulation rate patterns. Also, the sensitivity tests have shown that the trend in surface elevation change is more sensitive to accumulation uncertainty than temperature uncertainty.
-l455
"*As variations in firn air content and firn thickness align*": the meaning of this statement is not clear to me, please clarify.
-l463 – 464
"*our model in general is robust*": again, what do the authors mean here?

References used in this review

Arthern, R. J. and Wingham, D. J.: The Natural Fluctuations of Firn Densification and Their Effect on the Geodetic Determination of Ice Sheet Mass Balance, Clim. Change, 40, 605–624, https://doi.org/10.1023/A:1005320713306, 1998.

Kuipers Munneke, P., Ligtenberg, S. R. M., Noël, B. P. Y., Howat, I. M., Box, J. E., Mosley-Thompson, E., McConnell, J. R., Steffen, K., Harper, J. T., Das, S. B., and van den Broeke, M. R.: Elevation change of the Greenland Ice Sheet due to surface mass balance and firn processes, 1960–2014, The Cryosphere, 9, 2009–2025, https://doi.org/10.5194/tc-9-2009-2015, 2015.

Li, J. and Zwally, H. J.: Modeling of firn compaction for estimating ice-sheet mass change from observed ice-sheet elevation change, Ann. Glaciol., 52, 1–7, https://doi.org/10.3189/172756411799096321, 2011.

Ligtenberg, S. R. M., Helsen, M. M., and van den Broeke, M. R.: An improved semi-empirical model for the densification of Antarctic firn, The Cryosphere, 5, 809–819, https://doi.org/10.5194/tc-5-809-2011, 2011.

Medley, B., Ligtenberg, S. R. M., Joughin, I., Van Den Broeke, M. R., Gogineni, S., and Nowicki, S.: Antarctic firn compaction rates from repeat-track airborne radar data: I. Methods, Ann. Glaciol., 56, 155–166, https://doi.org/10.3189/2015AoG70A203, 2015.

Medley, B. and Thomas, E. R.: Increased snowfall over the Antarctic Ice Sheet mitigated twentieth-century sea-level rise, Nat. Clim. Change, 9, 34–39, https://doi.org/10.1038/s41558-018-0356-x, 2019.

Morris, E. M. and Wingham, D. J.: The effect of fluctuation in surface density, accumulation and compaction on elevation rates along the EGIG line, central Greenland, J. Glaciol., 57, 416–430, 2011.

Morris, E. M. and Wingham, D. J.: Uncertainty in mass-balance trends derived from altimetry: a case study along the EGIG line, central Greenland, J. Glaciol., 61, 345–356, 2015.