Review of Thompson-Munson et al.: *Observed and modeled Greenland firn properties (1980–2020)*
by Vincent Verjans.

This study applies two state-of-the-art firn models at the scale of the Greenland Ice Sheet (GrIS). The Community Firn Model (CFM) is used with the semi-empirical firn densification formulation NASA GSFC-FDMv1 of Medley et al. (2022). SNOWPACK is a more physically-detailed snow compaction model. The goal of this study is to compare results from these two different approaches to firn modeling over the GrIS. The authors also perform a comparison of model output against in-situ firn core observations.

I believe that this study demonstrates a comprehensive modeling effort, and the results are a valuable large-scale comparison of firn model behavior. The modeling experiments are rigorous, well-explained, and are undoubtedly a great contribution to the firn modeling community. The use of a same atmospheric forcing allows to identify differences only related to firn model and to parameterization choices. The figures are of good quality. And the authors perform a thorough evaluation by using an extensive dataset of 766 firn cores. However, I believe that there is a problem in the methodology of the evaluation, as I explain in this review. I appreciate the modeling effort and the results of this study, but I have some reservations concerning the interpretation and the lack of in-depth investigation of the structural differences between the two models. In other words, the results are excellent, but what can we conclude from this study? What is the main message for the firn modeling community? I believe that with a little more analysis, this study can be much more than simply providing model output from two models at the GrIS scale. This work is in the scope of The Cryosphere, and I welcome its publication pending some revisions.
I have separated my review in Major comments that require a re-evaluation of some steps of the study, Minor comments that require more clarity in the manuscript and/or small changes, and Specific comments, which focus on specific aspects, and are mostly of technical nature. Despite my numerous comments, I strongly encourage the authors to re-submit the manuscript after the revisions have been made.

<u>Major comments</u>
1) Problems of the evaluation
The use of 766 firn cores in the evaluation process is noteworthy. However, as pointed out by the authors (l.207-208):
"*Since most observations are from shallow cores (median depth = 2.0 m; Fig. 2a) the observed FAC values are relatively low (median FAC = 1.3 m; Fig. 2b) and do not represent the FAC of the full firn column.*"
This skewed distribution of the observations make the evaluation very biased and difficult to interpret. In any modeled firn profile, FAC in the upper two meters is essentially dictated by the surface density (ρ0). In the CFM, ρ0 is fixed to 350 kg m^{-3}. This is purely a choice of the authors, as any other constant value or parameterization of ρ0 would be equally valid (e.g., Kuipers Munneke et al., 2015; Fausto et al., 2018; Medley et al., 2022). Thus, the statistics of the evaluation (NSE, relative bias) essentially reflect how well the choice of ρ0 fixed to 350 kg m^{-3} fits the SUMup surface densities, rather than showing the performance of the CFM GSFC-FDMv1 densification scheme. Similarly for SNOWPACK, the evaluation reflects the performance of the surface density scheme compared to the SUMup surface densities, and not its performance in densification physics.
Another issue with using FAC as an evaluation metric is that the best predictor of FAC is the core depth. However, this does not bear any information about model performance. For example, a pair of modeled and observed FAC values over 2m depth will (almost) always be close to each other, and a pair of modeled and observed FAC values over 20m depth will also (almost) always be relatively close to each other. The good correlation between modeled and observed FAC values is due to cores being compared over a same depth. This problem arises because the authors have decided to use all the SUMup cores in the evaluation, and not to restrict their analysis to cores with a minimum depth threshold.
In order to alleviate these two problems, I encourage the authors to make an evaluation by binning cores based on their depth. A separate evaluation for each depth bin should be performed. For example, all the cores can be separated in groups of depth<2m, 2m<depth<5m, 5m<depth<15m, depth>15m or something similar. The binning should be made appropriately in order to have sufficient cores in each bin, but also meaningful evaluation

statistics at the same time. Furthermore, I would like the authors to highlight more clearly that the CFM results at low FAC values are mostly determined by the ρ0 choice and not by the GSFC-FDMv1 densification scheme. In the current version of the manuscript, this is not clear for readers less familiar with firn modeling. Finally, the authors point out (l197-200): "*SNOWPACK simulates more variability between layers compared to the CFM. This partly results from the fixed surface density of 350 kg m^{-3} set for the CFM, while the surface density in SNOWPACK varies based on atmospheric conditions, and partially because the CFM outputs are interpolated onto a grid.*" This is important and needs to be quantified. How much of this low-variability error is due to the ρ0 assumption? And how much is due to the interpolation? I strongly recommend to run the CFM at some firn core locations with ρ0 set to the SNOWPACK ρ0 time series, and without the interpolation scheme. This would bring better insights into the impact of these aspects.

2) Interpretation of the results
This study is a model intercomparison. I believe that this warrants more discussion of why the models diverge, and what conditions make them more prone to diverge/agree.
This firstly necessitates a better description of the model physics. The governing equations (densification, heat conduction, etc.) should be provided in the manuscript or in an Appendix. Based on these equations and on their results, the authors should provide some explanation on the different sensitivities of CFM-GSFC-FDM and SNOWPACK to temperature, accumulation, melt, wind forcing, etc. For example, I found Figure 9 very interesting. But the analysis does not tell why GSFC-FDM and SNOWPACK agree well in the Northeast and Southeast, but show strong discrepancies in the Central West and Norhwest. As another example, from Figure 5, why does SNOWPACK simulate much larger FAC at low summer temperatures than CFM-GSFC-FDM? Throughout the manuscript, I have been somewhat frustrated by the dichotomy between impressive results but lack of in-depth explanations.
Finally, the authors have related the sensitivity of both models to climatic conditions (LTSR and summer temperature) in the steady-state climate configuration. It would be interesting to expand such an analysis to the transient climate configuration. This would involve quantifying the sensitivity of FAC loss/gain to changes in atmospheric forcing.
As a final note concerning this Major comment 2, I should emphasize that addressing the sort of questions that I raise is not an absolute necessity for publication. The study is already a thorough modeling effort, with a good quantitative evaluation of the results. I simply believe that a thorough analysis of model behavior would bring this study to the next level.

<u>Minor comments</u>
1) References
I find that this study does not sufficiently recognize previous work from the firn science community. I provide here some examples, but I also encourage the authors to proceed to a more in-depth literature review, and to cite other previous relevant studies in their manuscript.
- l42-43 "*Changes in the amount of air-filled pore space within the firn, known as the firn air content (FAC), have been investigated in both observations (e.g., Vandecrux et al., 2019) and models (e.g., Medley et al., 2022).*": please cite Benson (1962); Braithwaite et al. (1994); Sorensen et al. (2011); Kuipers Munneke et al. (2015); etc.
- l48-49 "*Modeling firn has become important for estimating mass balance (MB) from satellite altimetry, since this method relies on firn models to interpret the causes of surface height changes (e.g., Li and Zwally, 2011).*": please cite Arthern and Wingham (1998); Morris and Wingham (2014).
- l52-53 "*Additionally, understanding the limits and deficiencies in firn models is essential for quantifying uncertainties in altimetry-based MB estimates.*": please cite Morris and Wingham (2015); Verjans et al. (2021).
- l57-59 "T*hese models use empirical relationships between densification, accumulation, and temperature, and they are often tuned to observations (e.g., Ligtenberg et al., 2011; Medley et al., 2022; Li and Zwally, 2011).*": please cite Herron and Langway (1980); Arthern et al. (2010); Simonsen et al. (2013); Verjans et al. (2020)
- l68-69 "*Still, both semi-empirical as well as physics-based firn models have been successfully used in Greenland (e.g., Vandecrux et al., 2020b; Dunmire et al., 2020; Medley et al., 2022).*": please cite Sorensen et al. (2011); Kuipers Munneke et al. (2015).

l71-72 "*At an ice-sheet scale, few comparisons of semi-empirical and physics-based models exist*": please recognize the work of Steger et al. (2017) here.
-l356-357 "*Neither model captures the high densities resulting from the firn aquifer because the use of bucket scheme in the models prevents full saturation in the firn.*": when discussing this aspect, please note that firn aquifer formation has been modeled by Verjans et al., (2019) and that conditions for aquifer development have been investigated by Kuipers Munneke et al. (2014).
- l361-362 "*Our results agree with these findings that model differences are highest where liquid water is present, indicating that poor representation of meltwater percolation processes is still a substantial limiting factor in firn model performance.*": please cite Verjans et al. (2019).
- l370-372 "*This demonstrates the limitations of a coarsely-gridded forcing, especially in steeply sloped areas where climate is likely to be highly variable within a single grid cell*.": please mention the downscaling work of Noël et al. (2016).
- l419-420 "*This could be related to the fact that SNOWPACK was developed using data from seasonal alpine snow which may not be representative of the physical processes driving deep firn densifcation.*": please cite Maeno and Einuma (1983); Arnaud et al. (2000).

2) The Reference Climate Interval (RCI)
The use of an RCI is necessary for the spin-up of firn models. However, this implies assumptions which must be properly understood, explicitly stated, and discussed. Here, the authors state (l126-127): "*We make the assumption that this period is representative of the longer-term Greenland climate.*" And further, they state (l225-227): "*The RCI used for model spin-up spans 1980 through 1995, and since we assume that this period represents a relatively steady-state, long-term Greenland climate (Fig. A2)*". I believe that the message conveyed to the readers about the RCI is misleading. The RCI is used to develop the initial model firn column, from which transient experiments over the period of interest (1980-2020) start. As such, ideally, the initialization should be computed with the true climate forcing of the decades and centuries preceding 1980. This is true regardless of whether the long-term conditions were in steady-state (i.e., without trends) or not. In other words, the "perfect" RCI should not represent steady-state conditions if the true conditions were not in steady-state prior to 1980. However, we have only incomplete knowledge of the true climate, especially prior to 1980. In light of this incomplete knowledge, using steady-state conditions over the RCI is a reasonable simplification, but not a necessary condition for a valid firn model initialization procedure. In the manuscript, it should be clearer that steady-state conditions over the RCI are used in order to isolate effects of climatic deviations from the RCI on firn column changes. But such steady-state conditions are not representative of true conditions, and the true changes in firn thickness are influenced by the unsteady nature of past climate conditions. I would like the authors to mention these points in the Discussion, as well as other studies that have shown that the pre-1980 climate of GrIS was not in steady-state (e.g., Hanna et al., 2011).

3) Evaluation of sensitivity to climatic changes
The evaluation of model performances is performed with firn core data. Because firn cores only provide a snapshot of firn density in time (as pointed out by the authors), they cannot serve to evaluate the sensitivity of firn models to changes in climatic forcing. For example, good performance of a model when evaluated against firn cores does not imply that the model would accurately capture FAC changes under a +1ºC change in mean surface temperature. This is particularly important to keep in mind when firn models are used to compute FAC change estimates in time, as done in this study. Evaluation with firn core data is legitimate given the scarcity of transient in-situ firn data, however I think that this limitation of the evaluation process deserves a paragraph in the Discussion section.

4) Clarification about the "CFM"
The authors repeatedly use the name "CFM" for one of the two firn models used. However, as far as I am aware of, the CFM allows to choose among various different firn densification formulations. Thus, I would find it more appropriate to call the model CFM- GSFC-FDM (or something similar). This is important because much of the FAC patterns are due to the use of the GSFC-FDM densification scheme, and not to the CFM itself, which is

simply a numerical tool. Furthermore, the authors state that (l144-145) "*The densification rate is determined with the NASA GSFC-FDMv1 firn densification equation (Medley et al., 2022)*". If this is the case, I suggest that they specify GSFC-FDMv1.2.1 to avoid any possible confusion with other versions of GSFC-FDMv1. Also, the CFM allows for a range of thermal conductivity parameterizations, the choice of which likely influences the results. I ask the authors to specify the thermal conductivity parameterization used. Finally, it should be clearer in the manuscript that the CFM itself is a numerical tool, and that the CFM output thus largely depends on the parameterization of the CFM (densification scheme, ρ0, thermal conductivity), and not on the CFM itself. This is important for readers less familiar with the CFM.

Specific comments

Title

This study does not bring any new observations about Greenland firn properties. For this reason, I find the use of the word "*Observed*" in the title inadequate.

l8

Specify: "isolate firn model differences".

l9

"*Both models perform well*": this needs to be quantified.

l10

"*is hindered by meltwater percolation*": this not really evaluated in this study. The authors only analyze in details the performance at firn aquifer sites, but do not compare performance in dry areas versus percolation areas more generally.

l10

Change "*the full ice-sheet*" to "ice-sheet-wide".

l11

Please move "*(i.e., air volume in the firn)*" to the line where FAC is used for the first time.

l13

Change "*the models' treatment of atmospheric input*" to "the sensitivity of the models to atmospheric forcing".

l15-16

Specify "spatially-integrated FAC decrease of".

l22

Change "*in a thick*" to "by a thick".

l23

Change "*density of firn varies across the ice sheet*" to "density of firn varies in depth and across the ice sheet".

l24

Remove "*in time and*".

l25

Specify: "can buffer the contribution of increased melt rates to sea-level rise".

l26-27

I find this sentence confusing, as it is not specifically about firn. Please consider rephrasing.

l31

To my knowledge, we are not sure whether ice slabs make deeper pore space completely inaccessible. I recommend using "potentially inaccessible".

l39

Change "*firn has lost its capacity to store meltwater*" to "the meltwater storage capacity of firn has abruptly decreased".

l46

Change "*the memory effect of changes to the firn from previous years*" to "firn changes evolving on multi-year timescales".

l52

Change "*measured in satellite altimetry*" to "measured from satellite altimetry".

l59

"*Semi-empirical models are beneficial because they do not rely on the physics of firn densification*": this statement is too crude and needs more nuance (see Arthern et al., 2010 for example).

l63-64

What is "*the constitutive relationship*"? As far as I know, even formulations linking stresses to firn strain rates rely on some form of parameterization, and there is no single universal constitutive relationship.

l65

Change "*observations from tuning*" to "observations for tuning".

l67-68

"*since snow physics have been more-thoroughly studied*": this requires one more line of explanation, and also remove the hyphen.

l68

"*have been successfully used*": what do the authors mean by "*successfully*"?

l70-71

"*have seen significant development for polar regions in recent years*": can the authors please list some of these developments?

l78

Remove "*completely*".

l94-95

"*Regional climate models are not always widely available or regularly updated, and no single reanalysis clearly outperforms others over the GrIS (Zhang et al., 2021).*": either provide more details, or simply say that the method could equally well be applied with a regional climate model or with another reanalysis product.

l95-97

I think that these two sentences should be rephrased as they do not read very well.

l99

Replace "*full ice-sheet*" by "ice-sheet-wide".

l99

Typo: "*gird*".

l105

What do the authors mean by "*successfully*"?

l109-110

"*the constitutive relationship*": see comment above.

l113-114

"*SNOWPACK uses the MeteoIO library (Bavay and Egger, 2014) for preparing the meteorological forcing data for the simulations.*": please explain.

l115-116

I don't think that storing output every 7 days conserves "*computational expenses*", but only reduces storage size. Please correct this statement, or provide explanations if I am wrong.

l117

"*impacted by the layer-merging scheme*": please explain.

l117

Typo: "*use*" should be "used".

l119-120

"*We set the surface roughness to 0.002 m for solving the energy balance with the Michlmayr et al. (2008) stability correction when a stable boundary layer is diagnosed*.": please explain more.

l121

Specify: "varies across the GrIS".

l134-135

"*The CFM uses a layer-merging scheme at 5- and 10-m depth to reduce computational demands.*": please explain more.

l149

"*was chosen to be near the depth at which the firn reaches the ice density*": please be more specific.

l151
Remove "*For example, if the firn needs 1000 years to spin up, the RCI would repeat 63 times.*".
l166
Remove "*(also known as "depth-integrated porosity (DIP)")*".
l168
Specify "where z = 0 m represents the surface, and is increasing downwards".
Section 3.1
Please re-order this section. The authors start with specific results at two individual cores and then provide the general results at the GrIS-scale. I recommend starting from the general results, and then focusing on specific results.
l193
"*collected in southwest Greenland*": please specify the core date also in the main text.
l197
Change "*in the surface*" to "near the surface".
l200-201
"*For polar regions in particular, temporal variations in wind and the presence of drifting snow translate into vertical density variations with increasing accumulation.*": I do not understand "*with increasing accumulation.*". Also, this statement requires some references.
l211
Specify: "are overestimating FAC on average".
Section 3.2
In general, when comparing SNOWPACK and CFM results, I strongly recommend using the Root Mean Squared Deviation metric. This would provide more quantitative information in the analysis.
l232-233
Remove this sentence as this is repetitive information.
Equation (6)
Does "*snow*" account for sublimation and blowing snow? And are these fluxes identical for SNOWPACK and CFM?
Figure 1b
Notice that SNOWPACK does not underestimate density at high depth here. This is interesting to me, and warrants more analysis about the model physics (see Major comment 2).
l242
Here and everywhere in the manuscript, change "*an FAC*" to "a FAC".
l244
Remove "*the*" in "*the FAC values*".
l246
Does "*response*" refer to FAC or to something else? Please avoid any confusion.
l248
"*generally similar spatial patterns in FAC across the ice sheet*": please quantify.
Figure 2
Please use log color scale in Fig 2c and Fig 2d. With the current color scale, almost all the data points simply appear white.
Figure 3
Because the equations of the linear regressions are not discussed in the main text, I recommend removing the black solid lines and the regression equations from Figure 3.
l253
Specify: "is on average greater than".
l253-254
"*difference between the models increases with depth*": please quantify.
l257
Refer to an appropriate Figure when introducing the basins.

l262
“fall in the middle”: please rephrase.
l271
Change “*between the two years*” to “over the RCI”.
l271-272
Change “*which verifies the steady-state assumption of the RCI and the design of the spin-up (Table 2).*” to “which is a consequence of our choice of RCI and of the design of the spin-up (Table 2).”
l274-277
Much of this information is repetitive information with respect to the paragraph above. Furthermore, please be more careful about the impacts of the RCI assumptions when discussing these results (see Minor comment 2).
l276
Change “*Greenland’s sea level rise buffering capacity*” to “the sea-level rise buffering capacity of the Greenland firn layer”.
l280
“*somewhat*”: please quantify.
l282
Throughout their manuscript, the authors use the term “*seasonal breathing*” to designate the seasonal fluctuations in FAC. I know that this term has been used in some previous studies. However, I personally dislike this term. In my view, it is scientifically incorrect: firn does not breathe. I would appreciate if the authors replace this term by another one. For example: the FAC seasonal amplitude. I thank the authors for their understanding.
l285-286
“*the seasonality was undetectable by the chosen methods*”: please explain in Figure A3 why the seasonality was undetectable in some cases.
Caption of Figure 5
Add comma: “air temperature, all calculated”.
Figure 6
Why are is there missing data from one or both models in some areas? Please explain in the main text.
Caption of Figure 6
Please rephrase the first sentence of the caption to make it more intelligible grammatically.
Caption of Figure 7
Specify: “The bottom row also includes the percent difference averaged over the GrIS.”
Caption of Table 1
Specify: “and the average percent difference between”.
Table 1
Provide root mean squared deviation between CFM-GSFC-FDM and SNOWPACK in each basin. And please discuss these values in the main text.
Caption of Table 2
Specify: “The average percent change between”.
Table 2
Firstly, I am surprised by the magnitude of changes between 1980 and 1995 (up to 0.5% in magnitude). How is this possible given that the 1980-1995 simulation is simply a repetition of the climatic RCI loop imposed during the spin-up. As far as I understand the spin-up process, repeating the RCI once more should cause only negligibly small changes. Can the authors please explain this?
Secondly, I think that for both periods (1980-1995 and 2005-2020), it would be more relevant to analyze the trend over the period instead of the change between two individual years. Analyzing only two individual years means that conclusions can be influenced by inter-annual variability. This may explain my first point of this comment.
l292
Change “*statistically*” to “significantly”.
l297-298

"*A consistent decreasing trend is modeled from ~2002 and through ~2011.*": please keep the 2005-2020 as a baseline for analysis. Switching between different periods of analysis makes the messages more confusing.
l305
Change "*associated water percolation processes*" to "associated water percolation and refreezing processes".
l312
"*SNOWPACK simulates greater negative changes in FAC compared to the CFM once the models diverge.*": I find this sentence unclear. Please consider rephrasing.
l314
Please specify: "followed by an increase in FAC only in the northeast and southeast".
l317-329
The differences between CFM-GSFC-FDM and SNOWPACK that are described in these two paragraphs are important. This deserves more detailed investigation into the causes of these differences (see Major comment 2).
l335-336
"*the ice slab nearest to the surface, which in some cases could be bare ice at the surface since there is no condition that the ice slab must be beneath a layer of snow or firn*": I apologize, but I will make another pedantic comment. The term "ice slab" has been used a lot over the last 6 years, and it is now often used inappropriately. Ice slabs are thick layers of ice that develop within a layer of porous firn. They should not be confused with the expansion of the ablation area. In this study, the algorithm of the authors makes no difference between development of ice slabs and ablation area extension. For this reason, I ask the authors to replace their use of "*ice slabs*" by "ice slabs or ablation area extension".
l340
Change "*depth to those slabs*" to "depth of those slabs".
l343-344
Please re-evaluate this following Major comment 1.
l350-351
"*The signature of model biases differs across the ice sheet as climate, topography, and the impact of firn hydrology vary*.": where is this assessed specifically?
Caption of Figure 9
Please add "note the different y-axis scales."
l353
"*Some of the highest model biases in SNOWPACK and the CFM occur in southeast Greenland*": please quantify.
l354
Please clarify: "First, some of the observed density profiles are from cores that were drilled directly into a perennial firn aquifer (Miller et al., 2018)."
l361
"*Our results agree with these findings that model differences are highest where liquid water is present*": please quantify the performances of the models in conditions of high LTSR.
l361
Please specify: "where liquid water is abundant".
l366
"*Within these five cores*": please rephrase.
Figure 10
I am puzzled by the big differences in the northeast ice caps. I wonder how much this influences the results provided in this study. If this influence is significant, the authors should reconsider the presentation of their results. The focus of this study is "*the Greenland ice sheet*" and not "the Greenland ice sheet and its surrounding ice caps".
l368
Change "*observed*" to "local".
l374-375
"*gives confidence in the models' abilities to simulate firn properties across the full ice sheet.*": please nuance this statement as the performances of the models in the percolation zone are limited.

l379
Specify: “in the firn models themselves.”
l382-383
“*the LTSR is a stronger predictor of FAC in the CFM compared to SNOWPACK*”: please quantify using the coefficient of determination.
l383-384
“*The large range of possible SNOWPACK-simulated FAC values at low LTSR values is likely due to the model’s sophisticated new-snow density scheme (…)*”: thus, is the difference in LTSR sensitivity between CFM-GSFC-FDM and SNOWPACK mostly caused by the ρ0 assumption for CFM-GSFC-FDM? (see Major comment 1)
l386
Change “*This indicates*” to “Our results show”.
l393
Typo: “*has*” should be “have”.
l396
“*which requires more detailed output from a dedicated firn model*”: please mention that this is only a particularity of MERRA2 because it does not provide melt as an output.
l400
“*where the spread in FAC is less than in SNOWPACK*”: please quantify.
l402
Change “*is easier to predict using LTSR*” to “shows less variability for a given LTSR”.
l409-410
The total Greenland FAC values provided here do not agree with the values given in Table 2. Please correct this.
l411-412
“*is close to a regional climate model’s (HIRHAM5_MOD) estimate for this period (Vandecrux et al., 2019).*”: please provide the value.
l416
“*uses the constitutive relationship between stress and strain*”: see comment above.
l422-423
Remove “*(and by proxy, whether the physics-based or empirical approach is recommended)*”.
l427
Change “*corresponds with*” to “corresponds to”.
l429-432
“*The stronger seasonality in the CFM is indicative of the model’s more simple treatment of forcing data like accumulation and temperature, which have strong seasonal patterns. SNOWPACK’s same sophisticated new-snow density scheme that leads to a complex relationship between LTSR and FAC also results in this smaller seasonal signal.*”: it is unclear to me how the authors reach these conclusions from their results. Please explain more.
l434
“*the 2012 extreme melt season can be seen as an abrupt drop in FAC in most basins*”: please be more nuanced. There is an abrupt drop only in the southwest, southeast, and to a lesser degree northeast basins.
l441
Please be more nuanced: “Ice slabs, which may render deep pore space inaccessible”.
l443-444
Please rephrase this sentence.
l445-447
“*Pore space depletion can also be a sign of firn densification, which has been found to increase cold content in the firn and amplify meltwater freezing and ice slab formation in the near-surface (Vandecrux et al., 2020a).*”: this statement is a very crude simplification of a complex process with many interactions and feedbacks. I suggest rephrasing: “Pore space depletion can also be caused by firn densification, which in turn modifies the meltwater refreezing and retention capacities of the firn in a complex manner.”
l459

"*place our results in a context of uncertainty*": this study does not perform any uncertainty analysis.
l460-461
Change "*lets us isolate the differences in the models themselves and examine how they respond to the same forcing*" to "lets us examine how the models respond differently to the same forcing due to structural model differences and parameterization choice".
l461
Change "*metric*" to "factor".
l464
Change "*While FAC in the CFM more gradually decreases as*" to "While FAC more gradually decreases in the CFM as".
l465
"*reaches near-zero FAC values that the CFM does not capture*": please consider rephrasing. This sentence suggests that SNOWPACK is right and that CFM is wrong.
l466
Change "*even more*" to "on larger areas".
l470
The FAC values given here do not agree with values in Table 2. Please correct this.
l471
"*reasonable estimations when compared to other studies*": please quantify.
l477
Change "*the pore space depletion is more extreme*" to "the pore space is more depleted".
l480
Specify: "the firn layer in the CFM loses only an equivalent of 1 mm."
Caption of Table A1
Does the standard deviation refer to (a) the within-basin standard deviations of the RCI mean values or (b) the annual standard deviation of the basin mean values? If it is (a) specify "mean plus/minus basin standard deviation", and if it is (b) specify "mean plus/minus inter-annual standard deviation".
Caption of Table A2
Same comment as for the caption of Table A1.
Caption of Figure A1
Move "*(green vertical line)*" next to "*350 kg m^{-3}*".
Caption of Figure A2
Change "*full ice sheet*" to "Greenland ice sheet".
Section "*Code and data availability*"
I thank the authors for providing the entire model output from both models as an open-source dataset. However, I ask the authors to include a README file in the dataset that explains in details how one can read and extract information from the files.

References used in this review

Arnaud, L., Gay, M., Barnola, J.-M., and Duval, P.: Physical modeling of the densification of snow/firn and ice in the upper part of polar ice sheets, in: Physics of Ice Core Records, edited by: Hondoh, T., Hokkaido University Press, Sapporo, Japan, 285–305, 2000.

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.

Arthern, R. J., Vaughan, D. G., Rankin, A. M., Mulvaney, R., and Thomas, E. R.: In situ measurements of Antarctic snow compaction compared with predictions of models, J. Geophys. Res.-Earth Surf., 115, 1–12, https://doi.org/10.1029/2009JF001306, 2010.

Benson, C. S.: Stratigraphic Studies in the Snow and Firn of the Greenland Ice Sheet, U.S. Army Snow, Ice and Permafrost Research Establishment (SIPRE–CRREL), Research Report 70, reprinted with revisions by CRREL, 1996, 1962.

Braithwaite, R., Laternser, M., and Pfeffer, W. T.: Variation of near-surface firn density in the lower accumulation area of the Greenland ice sheet, Pâkitsoq, West Greenland, J. Glaciol., 40, 477–485, https://doi.org/10.3189/S002214300001234X, 1994.

Fausto, R. S., Box, J. E., Vandecrux, B., van As, D., Steffen, K., MacFerrin, M., Machguth H., Colgan W., Koenig L. S., McGrath D., Charalampidis, C., and Braithwaite, R. J.: A Snow Density Dataset for Improving Surface Boundary Conditions in Greenland Ice Sheet Firn Modeling, Front. Earth Sci., 6, 51 pp., https://doi.org/10.3389/feart.2018.00051, 2018

Hanna, E., Huybrechts, P., Cappelen, J., Steffen, K., Bales, R. C., Burgess, E., McConnell, J. R., Peder Steffensen, J., Van den Broeke, M., Wake, L., Bigg, G., Griffiths, M., and Savas, D.: Greenland Ice Sheet surface mass balance 1870 to 2010 based on Twentieth Century Reanalysis, and links with global climate forcing, J. Geophys. Res.-Atmos., 116, D24121, https://doi.org/10.1029/2011JD016387, 2011.

Herron, M. and Langway, C.: Firn densification: an empirical model, J. Glaciol., 25, 373–385, https://doi.org/10.3189/S0022143000015239, 1980.

Kuipers Munneke, P., Ligtenberg, S. R. M., van den Broeke, M. R., van Angelen, J. H., and Forster, R. R.: Explaining the presence of perennial liquid water bodies in the firn of the Greenland Ice Sheet, Geophys. Res. Lett., 41, 476–483, https://doi.org/10.1002/2013GL058389, 2014.

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.

Maeno, N. and Ebinuma, T.: Pressure sintering of ice and its implication to the densification of snow at polar glaciers and ice sheets, J. Phys. Chem., 87, 4103–4110, https://doi.org/10.1021/j100244a023, 1983.

Medley, B., Neumann, T. A., Zwally, H. J., Smith, B. E., and Stevens, C. M.: Simulations of firn processes over the Greenland and Antarctic ice sheets: 1980–2021, The Cryosphere, 16, 3971–4011, https://doi.org/10.5194/tc-16-3971-2022, 2022.

Morris, E. M. and Wingham, D. J.: Densification of polar snow: Measurements, modeling, and implications for altimetry, J. Geophys. Res.-Earth Surf., 119, 349–365, https://doi.org/10.1002/2013JF002898, 2014.

Morris, E. M. and Wingham, D. J. : Uncertainty in mass balance trends derived from altimetry; a case study along the EGIG line, Central Greenland. Journal of Glaciology, 61(226), 345–356, https://doi.org/10.3189/2015JoG14J123, 2015.

Noël, B., van de Berg, W. J., Machguth, H., Lhermitte, S., Howat, I., Fettweis, X., and van den Broeke, M. R.: A daily, 1 km resolution data set of downscaled Greenland ice sheet surface mass balance (1958–2015), The Cryosphere, 10, 2361–2377, https://doi.org/10.5194/tc-10-2361-2016, 2016.

Simonsen, S. B., Stenseng, L., Adalgeirsdóttir, G., Fausto, R. S., Hvidberg, C. S., and Lucas-Picher, P.: Assessing a multilayered dynamic firn-compaction model for Greenland with ASIRAS radar measurements, J. Glaciol., 59, 545–558, https://doi.org/10.3189/2013JoG12J158, 2013.

Sørensen, L. S., Simonsen, S. B., Nielsen, K., Lucas-Picher, P., Spada, G., Adalgeirsdottir, G., Forsberg, R., and Hvidberg, C. S.: Mass balance of the Greenland ice sheet (2003–2008) from ICESat data – the impact of interpolation, sampling and firn density, The Cryosphere, 5, 173–186, https://doi.org/10.5194/tc-5-173-2011, 2011.

Steger, C. R., Reijmer, C. H., van den Broeke, M. R., Wever, N., Forster, R. R., Koenig, L. S., Kuipers Munneke, P., Lehning, M., Lhermitte, S., Ligtenberg, S. R. M., Miège, C., and Noël, B. P. Y.: Firn Meltwater Retention on the Greenland Ice Sheet: A Model Comparison, Front. Earth Sci., 5, 3, https://doi.org/10.3389/feart.2017.00003, 2017

Verjans, V., Leeson, A. A., Stevens, C. M., MacFerrin, M., Noël, B., and van den Broeke, M. R.: Development of physically based liquid water schemes for Greenland firn-densification models, The Cryosphere, 13, 1819–1842, https://doi.org/10.5194/tc-13-1819-2019, 2019.

Verjans, V., Leeson, A. A., Nemeth, C., Stevens, C. M., Kuipers Munneke, P., Noël, B., and van Wessem, J. M.: Bayesian calibration of firn densification models, The Cryosphere, 14, 3017–3032, https://doi.org/10.5194/tc-14-3017-2020, 2020.

Verjans, V., Leeson, A., McMillan, M., Stevens, C., van Wessem, J. M., van de Berg, W. J., van den Broeke, M., Kittel, C., Amory, C., Fettweis, X., Hansen, N., Boberg, F., and Mottram, R.: Uncertainty in East Antarctic firn thickness constrained using a model ensemble approach, Geophys. Res. Lett., 48, e2020GL092060, https://doi.org/10.1029/2020GL092060, 2021.