Review of "On the Contribution of Grain Boundary Sliding to Firn Densification – an Assessment using an Optimisation Approach"
by Schultz et at.

Summary and general comments:

This paper presents work using the framework presented by Alley (1987) to investigate the common assumption that near-surface firn densification is driven by grain-boundary sliding. The authors use an optimization approach to find the best coefficients for four variants of their model at 159 different sites that have firn depth-density data in the SUMup dataset. The authors righty identify the importance of improving our ability to model firn for numerous problems in glaciology, and I applaud them for creatively leveraging the extensive SUMup database in a way that has not previously been done.

I believe this work holds promise to provide a worthy contribution to *The Cryosphere*, but I recommend major revisions prior to its publication. The paper is very heavy on its description of the model development and methods used, but the analysis of the results and their implication on the broader firn modeling/glaciological community are spare. I found a number of their analyses to be unconvincing and assumptions to be not fully explained and/or justified. (e.g. their assumption of linear relationships between their parameters and temperature and surface mass balance). Additionally, the manuscript needs significant editing to improve its structure and clarity – it has many grammatical errors and heavily uses multi-clause sentences that muddle the authors' points. I have noted some (but not all) of these in my specific comments below.

A few general comments:
- The manuscript purports to be based on the grain boundary sliding model presented by Alley (1987), but the authors wrap all of the grain structure parameters from Alley into a single parameter C, which makes their model quite similar to other macroscale firn models for stage 1 (albeit with a different dependence on the current density). The authors do point this out eventually, but I would like more discussion about this assumption. Is it still a grain boundary sliding model, and why?

- Related: The authors use RMSD to quantify the misfit between their model variants and firn-depth-density profiles – but are any of these variants significantly better than previously published firn densification models for stage 1? I think it would be a relatively simple exercise to run, e.g. the Herron and Langway (1980) or another firn model, for these sites to show that the authors' model provides an improvement.

- A pesky problem for the firn modeling community is layering in the firn. The core the authors chose to highlight shows significant layering. Like other models, their model does not capture the layering; but, I suggest that a physically based model should be better at simulating layering than the macroscale models. As the authors are claiming that their model formulation is more physically based, a bit of discussion about this model deficiency would be welcomed.

- I would like to see a more detailed uncertainty analysis as a part of the discussion (or as its own section). Not until the end of the conclusion do the authors mention that the results will be affected by biases in firn model forcing data that comes from a regional climate model; this is true, and it should come sooner. Additionally, how do other assumptions contribute to uncertainty, e.g. the assumption of surface grain size and Arrhenius activation energy? To what extent is C just accommodating other errors – e.g. the authors assume that the model calculates grain size r accurately using the Gow formula – but errors are introduced by the boundary condition, and this formula is just an empirical relationship itself.

The authors propose a linear relationship to predict C from SMB. If this relationship exists, they should include the equation for the line, and use the value of C predicted by it to run their model. How do those results compare to the outputs from the optimized model? Are they acceptable? Are they better than predicted by a previously published model?

Specific Comments

1. Introduction

Line 15: the first paragraph should be rewritten – it is filled with grammatical errors and does not contain citations. I would argue that cold content (Vandecrux et al., 2020) and ice lenses (MacFerrin et al. 2019) are just as important for governing melt water retention. I do not disagree that firn densification is important for meltwater retention, but I do not think this paragraph adequately makes a strong case that it is. The last sentence implies that current models are not adequate but does not provide evidence that they are not.

L22: I would not say that the other models neglect the overburden stress exactly – for models operating at small-ish (sub annual) timesteps, they use the mean accumulation rate over the lifetime of a firn parcel (Li and Zwally 2011, 2015), which is effectively the stress, rather than the mean site accumulation rate.

L29: Arnaud and Goujon are actually the same firn-densification model, Goujon just added heat conduction to the Arnaud densification physics. I think it is worth noting at this point that Anaud/Goujon used Alley (1987) equations for stage 1 densification because that is what your paper is about. Their equations for stage 2 densification are based on Maeno and Ebinuma.

L26 – 35: Since your paper is about modeling grain boundary sliding in stage 1, I think it would be prudent for you to focus on how the models handle densification in stage 1. That is, you talk about sintering before you talk about grain boundary sliding (paragraph starting on line 32) – this is just distracting.

L36: change to ‘first applied grain boundary sliding theory …’

L36: be specific of what you mean with ‘low density’ – less than the critical density, I think.

L39: numerous times: cite numerous studies

L40: ‘tried to’ – colloquial, vague language. Hypothesized? Showed?

L48: ‘in the sense of Alley’ → ‘Following Alley,’

Section 3.1

My general feeling with Section 3.1 (and much of section 3) is that it is distracting from the what the authors state is the goal of their paper (at the end of the introduction), which is to evaluate the efficacy of a grain-boundary sliding densification model. I think that these details of the numerical treatment help the paper achieve that goal. My suggestion would be to move this section to an appendix at the end of the paper. Perhaps I feel this way because I got lost in some of the details, which I attribute to several instances of imprecise language (See below comments). I am admittedly not familiar with the updated Lagrangian formulation, but I don’t think that understanding the numerics of that numerical scheme are

important for understanding the results (hence my suggestion that this section be moved to an appendix). The meat of the paper begins with section 4, so reduce section 3 to only what is necessary.

Equation 1: you say you let *v_b = v*, but when you do that, the left term implies that the time derivative of the density (i.e. the densification rate) is zero – is that what you mean? I am not sure what you are getting at with this.

L74: 'build' → built

Line 71: I am confused a bit with *v* and *v_b* – can you clarify the language? You use multiple terms ('flow velocity', 'boundary velocity' and 'material velocity', and later 'grid point velocity'), but only have the two variables. Please choose one term for each variable and stick with that. It is also not clear which velocity is which – e.g. with 'boundary velocity', are you talking about the boundary of one grid point/control volume, or the boundary of the entire model domain (i.e. downward advection due to underlying ice dynamics?).

Equation 2: please use a different variable for stress – you already use *t* for time (eq. 1). This gets especially confusing because Alley (1987), equation 5, matches your equation 5; but your *t* has a different meaning than the *t* in the Alley paper. I don't think it is sufficient to differentiate between *t* and *t_zz*. Alley used *P* for pressure; Morris and Wingham (2014) and Arthern et al. (2010) used σ.

Equation 4: With this formulation, it seems that in the accumulation zone the model domain will continue to thicken indefinitely?

Section 3.2

- write Equation 5 as a part of a sentence (as you have done with other equations)

- Stress term in Eq. 5 – Alley (1987) makes a steady-accumulation assumption. In this paper it appears that you do not – see comment regarding this at Section 4 comments.

- Line 107: 'seeming arbitrary at first' is a subjective comment (is it more arbitrary than the 5/3?); remove it.

- Do not use 'term' and 'factor' interchangeably. (e.g. you use 'term' to describe $8D_{BD}\Omega/k_bTh^2$ but this is a factor, not a term.)

-Line 117: 'kind of fade out behavior' – please use specific language and avoid colloquialisms – 'fade out' is used several more times, but I think you can be more precise, e.g. how does it fade out? e.g. "the influence of grain boundary sliding on the strain rate decreases asymptotically as the density approached the critical density".

L117: 'inverse relative density of rho_c = 550 kg m^{-3}' – that value is the actual density, not a relative density. I think you mean something else here; please clarify.

-Line 120: previously you said *T* 'resembles the temperature' (what does that mean?), and here you explicitly say that *T* is the temperature.

-Line 120: you previously said *r* is grain radius, and here you redefine it slightly differently ('grain size' is a broader term than 'grain radius'). Which is it? Please be more specific of what you mean by grain

radius, because the grains are not perfect spheres and different people mean different things when they talk about grain radius.

3.4
L128: Again, I think you have just not explained your method clearly – what is a 'test volume'? This is the only time you use that phrase. Upon first glance, this sentence says that the densification rate is zero; the next sentence clarifies this a bit – but not until the third sentence do I figure out that your control volumes are changing. The upshot is that I don't doubt that your numerical scheme is correct; I think it is not explained clearly.

3.5
Line 142: This is an incomplete sentence. Discretized by what? Or maybe it is just written oddly?

Line 145: please cite page numbers with textbooks.

Equation 11: Is your equation for conductivity your formulation? I do not see it in Paterson 1994; please either cite source material or state that you came up with this parametrization (and how).

3.6
- Please provide what you use for a surface boundary condition for the grain size.

- "This is suitable as the grain radius is used in the constitutive equation by Alley (1987)" – please justify this assumption. Also, second use of word 'is' appears to be a typo.

Section 4.
Alley (1987) did not include heat diffusion (he assumed isothermal conditions), and he assumed steady accumulation. How does your inclusion of it affect your results vis-à-vis those of Alley? Please provide justification that it is appropriate to "test the concept of the material model developed by Alley.

Eq. 14 and 15: You are adding an Arrhenius term, which was not included in Equation 5 (but is of course included in many other firn models. Given that your paper is about optimizing the grain boundary sliding approach, can you justify this addition, or add text describing why you felt it was appropriate?

L172: How does lumping these parameters together affect the results? I would expect that they will vary by site/climate, and so what does it mean for your grain boundary sliding model that you remove the specific parameters that describe grain geometry?

L187: what does 'its' refer to? i.e. to test what's influence?

L191: It is not clear to me why a possible dependency (of the strain rate?) will reflect in the optimal values of Cv3 and Cv4? There is still a factor of 1/T in the equations.

L193: → 'The aim of the optimization…'

L195: Where is the Wilhelms core taken from? Please mark in on the map (figure 3) in a different color and provide coordinates. I suggest coming up with an easier-to-read name (i.e. state the full name once here when introducing and citing it, and give it an easy-to-recognize abbreviation to use for the remainder of the paper). I suggest this because I get concerned when I am reading if there are multiple cores with similar names that I need to be differentiating between.

L200: this implies that Lundin and Verjans did not use objective measures for their studies, which is not the case. Verjans did in fact use RMSE. I will also suggest that RMS deviation is not wholly objective in this scenario (which is not to say it is not useful) – imagine that you had a scenario where you were able to model the density perfectly at all depths except a single depth where it is off by 20 kg m^{-3}. If this misfit point is near the surface where the density is low, it is a much higher percent misfit than at greater depths where the density is greater.

L209: It occurs to me at this point that you have not described how you are modeling densification in stage 2 – ostensibly your model goes beyond stage 1 densification because this is important for getting the temperature profile/heat transfer correct. With the Arnaud/Goujon/Breant models, there is a factor introduced to allow the densification rates to match at the critical density – are you doing something similar? I do not think you need to provide a lot of detail but you should add a sentence describing what you do for density greater than 550.

L214: Please be more specific about what tests you did.

L215: A general note: Thank you for consistently including the units with your variables and equations.

L223: It would be helpful to explicitly say that there are 21 densities you tested, e.g. 'we tested 21 different values of the surface density…"

Section 5

L232: Be specific of how many sources, or leave that out. Probably you could just remove your short description of what SUMup is (the fact that is is available online and has a lot of measurements is not germane for the point of your paper)

Figure 2: Since you include the SMB and temperature in a plot, it might be useful to also include the mean of those as predicted by RACMO.

L241: Positive SMB and melt are not exclusive - quite a few sites experience significant melt in Greenland yet still have positive mass balance. Quite a few of your sites, as far as I can tell on your map, are in wet firn locations. How did you handle these sites with your simulation? Your model handles densification due to compaction but not densification due to meltwater percolation and refreezing. How does this assumption affect your results? Do you see different optimal values of C for sites where melt > 0? Or higher optimal surface densities?

L255: Except it appears that the cold, low-accumulation center of the East Antarctic ice sheet is not represented, and this is one place of high interest for firn studies because of its relevance to ice core studies. (The Arnaud model specifically used Vostok as a test site).

Section 5: I suggest moving section 5 to before section 4. At this point my feeling as a reader is that I have made it to the 12th page and am finally getting to the part of the paper where you tell me more specifics of the science you did (this is also related to my suggestion that you move a majority of section 3 to an appendix.)

Section 5.2: I suggest rewriting the first paragraph of this section for clarity.

L259 – 268: I am confused here. You said earlier that your model time resolution is 48 steps per year (roughly weekly), but you say here that you are using annual data from RACMO. Why do you not use the daily RACMO data downsampled to weekly resolution? Are you using the same value for the temperature

and accumulation for each of the time steps during a given year? On Line 260 you say you are interpolating, but what are you interpolating (spatially? temporally?), and how?

L269: In the cases where you have cores that have measurements from the near surface, why don't you just use the measured surface density from the cores? Are there cases where the surface density from the optimization scheme is significantly different than the observations from SUMup? It seems possible that your scheme could allow a surface density that is different than the observation in order to get a better overall RMSD fit. Do you use the same density at all time steps for a given site/simulation?

L270: What is the basis of your choice of 0.5mm for initial grain size? Is it a realistic assumption to assume it is the same everywhere? (I would think not – southern Greenland is quite different than e.g. the South Pole). How sensitive is your model to this parameter choice?

L274: But aren't the surface densities in good agreement because the range you chose for them is based on what the literature says in the first place? This seems circular.

L277: I think there is literature that you could find to cite regarding Antarctic vs. Greenland surface densities. Or you could query SUMup and find all measurements shallower than 50cm or 1m for Greenland and Antarctica.

L279 – 280: This seems out of place – should it be in the temperature section in section 3? Also you need to specify what your model domain depth is since you are using a Neumann condition.

Section 6:

L290: 'even better agreement' – this implies that that core has good agreement in the first place – but does it? You do not provide context for whether ~23 kg m^{-3} is actually a good agreement. I suggest that RMSD is a good metric to intercompare the models' performance – v2 is doing better in general than v3. But it does get murkier when actually deciding if this is a "good" fit to the data – so this would be better if you could normalize the RMSD somehow, and provide a quantification of what you consider to be a good fit (any why).

Figure 5: please label which variant is which in a more obvious manner (the v1 is small and hard to spot).

L291: Isn't this more a function of the SMB (from RACMO) being accurate than it is the strain rate being correct? If you were to optimize your model to fit the depth-age profile best (rather than depth-density), would the same optimal parameters be the same?

L296: This is confusing. Alley (1987) as far as I can tell does not include an Arrhenius term. (Although he did derive an activation energy for grain boundary sliding). This makes me wonder: the activation energy you use is notably smaller than recent studies have implied it should be (Arthern et al. 2010; Morris and Wingham 2014) – how does your choice of activation energy affect your results?

Note on figures in general (especially Fig. 6): the colors you have chosen can be challenging to see (especially the yellow) when text is written in those colors.

L305/Figure 7: To me these data do not appear to be linear, and the lines do not appear to be good fits (especially at low and high temperatures). A Pearson correlation coefficient is based on the assumption that the data are linear, which I am not convinced they are, so I think it is deceiving to include that metric. To me the most obvious conclusion is that there is not a clear linear relationship between any of the c parameters and temperature. If this is not what I should take away from figure 7, you need to be more thorough convincing me.

Figure 7: what does a negative value of c indicate? (it is hard to tell if there are any; but figure 7's y axis extends to -2, and it looks like there could be? Make the axis the same as figure 8.)

310: Again, 'even higher' is implying that you had a good fit between temperature and c, but the a high Pearson is meaningless because the data in figure 7 are not linear.

L312-315/Figure 8: These do appear to be more linear than the C vs temperature plot, but I think it would be appropriate to do a statistical test to actually show that they data are linear (e.g. a statistical model with higher order terms that have coefficients not significantly different than zero, or some other test). I am concerned with the values at the high accumulation sites for variants 1 and 2 – there are clusters that are half of what would be predicted by your linear model. And, what is going around 0.4 m w.eq./a? (And 0.6?) These vertical series of dots would indicate to me that there is no correlation between SMB and C – i.e. you can have a number of sites with SMB = 0.4 and C can vary by a factor of 10.

- Your discussion should include discussion of why the optimal C values vary by site and how that influences our understanding of grain boundary sliding.

Section 7:
L323: Again, I suggest that getting depth-age correct is more dependent on getting the accumulation rate correct – and, since the accumulation rate is often determined by counting annual layers, the science can become a bit circular.

L326: I agree entirely with what you say in this paragraph – your model formulation is not too different than the other models that have been published in recent years, but up until this point you have been claiming that you are optimizing a grain boundary sliding model (including in the title). I think it would be appropriate for you to rework the text a little bit to be more forthcoming with this, rather than waiting until the very end to point it out.

L338: Morris (2018) did in fact create a 'transition model' to move from stage 1 to stage 2.

L345: I am confused here – are you suggesting that because you do not have a separate Arrhenius factor in equations 16 and 17 there is implicitly an Arrhenius factor wrapped into C_v3 and C_v4? If this is the case, say so clearly in Section 4. It seems that you have a baseline assumption in your study that firn strain has an Arrhenius temperature dependence. But, if Variants 3 and 4 give good answers, couldn't this just as easily indicate that firn does not have an Arrhenius dependence? Part of my point here is that you have coefficients, C_v3 and C_v4, that have all sorts of 'physics' wrapped into them, and I don't think you can just decide what those physics are.

L349: The first two sentences of this paragraph need to be rewritten to improve clarity. After multiple readings it is still not clear to me what you are trying to convey.

L349, 352: 'resulting factors' – do you mean the 4 values of C for your variants? C and the surface density? Please use more specific language, e.g. 'indicate a clear dependency of C_v3 and C_v4 on the mean surface mass balance'.

L350: This is not obvious to me – are you saying that the C factors are functions of temperature, and since temperature and SMB are correlated, then C must also be a function of SMB?

L353: What do you mean by "Dependency of the mismatch"?

L354: You can at least do a scale estimate of the effect that horizontal stresses would contribute – see Horlings et al., 2020. Your paper is about stage 1 densification/grain boundary sliding – on the timescales of firn densification on ice sheets, is horizontal stress going to make a significant difference? I am also skeptical that an 3D model incorporating horizontal stresses would significantly reduce your misfit – at least in the example core you highlighted, the layering of the density profile likely contributes to a significant portion of the misfit.

Section 8:

L367: not represented well in the model – are you effectively saying here that the stress should have an exponent other than 1?
L371: I agree with this paragraph; I think it should be moved into the discussion section along with a more complete uncertainty analysis.

References:

Arthern, Robert J., et al. "In situ measurements of Antarctic snow compaction compared with predictions of models." *Journal of Geophysical Research: Earth Surface* 115.F3 (2010).

Herron, Michael M., and Chester C. Langway. "Firn densification: an empirical model." *Journal of Glaciology* 25.93 (1980): 373-385.

Horlings, Annika N., et al. "Effect of horizontal divergence on estimates of firn-air content." *Journal of Glaciology* 67.262 (2021): 287-296.

Li, Jun, and H. Jay Zwally. "Modeling of firn compaction for estimating ice-sheet mass change from observed ice-sheet elevation change." *Annals of Glaciology* 52.59 (2011): 1-7.

Li, Jun, and H. Jay Zwally. "Response times of ice-sheet surface heights to changes in the rate of Antarctic firn compaction caused by accumulation and temperature variations." *Journal of Glaciology* 61.230 (2015): 1037-1047.

MacFerrin, Michael, et al. "Rapid expansion of Greenland's low-permeability ice slabs." *Nature* 573.7774 (2019): 403-407.

Morris, Elizabeth. "Modeling dry-snow densification without abrupt transition." *Geosciences* 8.12 (2018): 464.

Morris, E. M., and D. J. Wingham. "Densification of polar snow: Measurements, modeling, and implications for altimetry." *Journal of Geophysical Research: Earth Surface* 119.2 (2014): 349-365.

Vandecrux, Baptiste, et al. "Firn cold content evolution at nine sites on the Greenland ice sheet between 1998 and 2017." *Journal of Glaciology* 66.258 (2020): 591-602.