

Rebuttal to the review by an anonymous reviewer

We thank the reviewer for their insightful and constructive comments on our manuscript. We'd hereby like to address their concerns, and propose revisions to our manuscript to alleviate them. Reviewer's comments are displayed in boldface, our replies in regular type.

Please note that all comments are answered in the original order, with one exception; since the reviewer's comments about the sub-grid melt scheme, the sliding law, and the ABUMIP results are strongly related, we answer those together.

Major comments

One goal of this paper is to convince the reader that the model can be used for scientific studies using a resolution of 16km. While many studies have shown this is too coarse to accurately represent ice sheet dynamic accurately, it can be a good trade off between computing cost and model accuracy especially for paleo time scale simulations. Keeping this in mind, I would have liked this paper focusing more on convincing the reader it could do so and quantify the errors with higher resolution results in the experiments presented here. In many of the experiments, either 16km is the highest resolution, or higher resolutions is used without comparing with results at 16km. For this reason, I find this study incomplete, and I strongly encourage the authors to add simulations to their suite of runs. Also, please, be a bit more quantitative in your analysis of each experiment.

The main goal of our paper is to convince the reader that the model can be used for palaeoglaciological applications. We typically use a resolution of 40 km for this kind of work, because of the very long timescales involved; while 16 km might be regarded as "coarse" within the context of future projections, for palaeo-applications this is generally already too high to be feasible. By showing that, in the different benchmark experiments, our model results at 40 km do not appreciably deviate from those at higher resolutions, we demonstrate that this relatively coarse resolution still produces reliable results. We will clarify the text in the introduction to more clearly reflect the fact that 40 km should be viewed as the default, and that the higher-resolution experiments serve only to illustrate that 40 km does not introduce prohibitively large truncation errors. Having said that, the suite of experiments also show that the model can in principle be used at a higher resolution.

I am a bit concerned about your need to scale your basal stress by the square of the grounded area fraction as opposed to its real value. It reduces the amount of basal stress applied in the cell containing the grounding line, and in some cases, by a lot (by half when the grounding line is at the middle of that cell). I am not sure why and it might be due to the implementation of your numerics, or else. For area fraction less than 0.5, it will behave closely to a model not using any grounding line parameterization which will lead to reduced model accuracy. Have you tried using the same logic used for the basal stress parameterization as the one used for your basal melt rate parameterization (i.e., considering the cell grounded if the cell center is grounded and floating otherwise)? Also, you applied this scaling using numerical justification, do you have any physical justification you could add to it?

We based our approach to the sub-grid friction scaling largely on that used in PISM, as described by Feldmann et al., 2014. In that paper, they describe not only a sub-grid scaling scheme, but also how the driving stress is calculated using two-sided differencing everywhere except for near the grounding line, where one-sided differencing is used. In practice, this introduces the same quadratic dependence on grounded area that we use. The difference between the two approaches has to do with the choice of staggered grids. PISM solves the SSA on the regular grid, so that the driving stress can be calculated using either two-sided or one-sided differencing. In IMAU-ICE, the SSA and the DIVA are both solved on the staggered grid, where one-sided differencing is a bit more problematic (because it would involve information from non-nearest-neighbour grid cells). We believe that scaling the basal friction with the square of the sub-grid grounded fraction is therefore simply a slightly different way to achieve the same result. Our preliminary experiments, where we did not include the square, showed similar results to what Feldmann et al. describe when they do not include the one-sided differencing scheme, which further strengthens this belief.

We will clarify the text of the manuscript to reflect these thoughts.

About your MISIMIP experiments: the advantage of this experimental setup is that it has an analytical solution. It would be good to see how your results compare to it. Pattyn (2017) uses a flux condition at the grounding line that is derived from the boundary layer solution from Schoof (2007), hence his spun up steady state is really accurate despite the use of a coarse resolution. It looks like your results are still about 100km away from the analytical solution at the end of the original spun up state with your 16km (highest) resolution. Also, at the end of the re-advance experiment, the difference in grounding line location between the 32km and the 16km resolutions results is much bigger than the difference between the 64km and the 32km resolutions results, and these differences seem to be about the same at the end of the retreat experiment. It would be sensible for you to show results at higher resolutions (8km and 4km for example) for this experiment. This will help quantify the error with 16km.

The experiment we performed follows that of Pattyn (2017), where the original MISIMIP experiment is modified from a 1-D flowline to a 2-D plan-view setting more appropriate for continental ice-sheet models. This means that the analytical solution by Schoof (2007), which was derived for a flowline case, is no longer applicable. We will clarify this difference in the manuscript to avoid confusion.

We will also perform a few additional simulations so that both the (modified) MISIMIP and the ABUMIP experiments are presented at the same 40, 32, 20, 16, and 10 km resolutions. We will furthermore add a panel to the MISIMIP figure showing the relation between grid resolution and GL hysteresis, which clearly shows how the amount of hysteresis decreases with resolution, and is always smaller than the grid resolution (which we believe is the most important result of this experiment).

We also fixed a small error in the plotting script that creates this figure, so that the GL position curves look smoother now (before, it calculated the position from the discrete flotation mask, this has now been changed to utilise the sub-grid GL position).

As you mention in your introduction, only a few existing ice sheet models are using DIVA. Thus, the best comparison possible to other models would be to employ similar experimental setup they have used too. You cite the work from Leguy et al. 2021 which has the advantage to present results with several sliding laws (including coulomb friction) and the same melt parameterization you chose for your model (on top of using DIVA as well). IT would have been sensible to implement some of their experimental setup and present results from your model. Especially since you are using their study to justify your choice of melt parameterization. I will get back to this below.

About your ABUMIP experiments: you are using ABUMIP to test your melt parameterization. While this experiment is an application of using your basal melt rate parameterization, it is not designed to demonstrate how well it works in your model. In this experiment, the melt rate is large and Leguy et al. (2021) showed that the melt rate parameterization is very resolution sensitive under these circumstances. Your model barely shows sensitivity with resolution. Also, IMAU-ice is the only model using a coulomb friction law preventing you to make a sensible comparison with other models in the study. Right now, you are comparing 2 different version of your own model and your results are even more extreme compared to the previous version. What do you make of this? Which results would you believe is the most sensible? (I am not saying it is wrong but comparing one outlayer with another just takes you so far.) A more meaningful comparison would have been to run with a Weertman sliding law and compare your results with CISM which is the only model using DIVA in Sun et al. (2020) (listed as L1L2). Physically speaking, you should expect ABUK to produce more sea level rise than ABUM. Eventually both experiments should plateau towards the same results but at the least, ABUM should be lagging behind ABUK. I am concerned that all your runs for a given resolution are similar for both ABUM and ABUK, and your results at 16km are the one leading to the most sea level rise. Also, you can see that your run at 40km returns more sea level rise with ABUM than with ABUK. You should discuss this more in your text and discard the results at that resolution for this reason. (A few models in Sun et al. (2020) exhibit this behavior and I don't think it is right either.) Again, showing performance with the setup from Leguy et al. (2021) or Seroussi et al. (2018) would help explain a lot of these behaviors. I do think adding this experiment is adequate for this paper and does not need to wait for another publication especially since you are claiming good results with your newly implemented melt parameterization.

I am concerned about your first paragraph in your future research section which I highlights here in 3 points:

- 1) The study from Leguy et al. (2021) clearly highlights that the choice of basal melt parameterization is model dependent. In fact, using a similar melt parameterization as in Seroussi et al. (2018), they found opposite results. This means that you should test your model first and convince yourself and reviewers which conclusion you model reaches.
- 2) You mention a factor of 2 of sea level rise differences depending on the choice of melt parameterization; was it twice as much or half of the presented results? If the former, that is indeed scary. If the later, why would you think it is not the right choice to make? Leguy et al. (2021) found that using PMP or FCMP lead to similar results.

- 3) You state that the NMP and PMP schemes were resolution dependent, inferring that the FCMP scheme isn't, which would explain your ABUK and ABUM results. Leguy et al. (2021) shows that all melt parameterizations are grid resolution dependent for high melt experiments, even more so when using a coulomb friction law. (This was true as well in Seroussi et al. (2018) in some cases and their coarsest resolution was 2km.) I would expect this to be true in your model as well.**

We agree that the interplay between the sliding law, the sub-grid melt scheme, and the modelled grounding-line retreat deserve to be investigated more thoroughly, and that the schematic MISMIP+-type of experiments performed by Leguy et al. (2021) and the more realistic ABUMIP-type of experiments provide a solid basis for this. In fact, we have already performed a large number of these experiments over the past few months, as part of the basal sliding / basal melt upgrade mentioned in the Discussion section of our manuscript. Regarding the relative “performance” of the sub-grid melt schemes, our findings match those reported by Leguy et al., meaning that it depends on the experimental set-up (i.e. schematic MISMIP+ v.s. realistic Antarctica) and on the model resolution. Likewise, we find different results in the MISMIP+ vs the Antarctica set-ups regarding the effect of sliding laws; in MISMIP+, the effect of choosing a different sliding law is larger than that of choosing a different melt scheme, whereas in Antarctica we find the opposite. Regarding the comparison to other, similar models: we have also performed all of the MISMIP+ experiments (again with different resolutions, sliding laws, sub-grid schemes, and also different stress balance approximations), and find that all of our simulations produce grounding-line retreat curves that lie well within the range of other models (as reported by Cornford et al., 2020).

While we agree that these are certainly valuable experiments, and we definitely want to publish them, we feel that they are of a different nature than the work presented in the current manuscript. Whereas the current work consists of more basic numerical verification experiments, these new experiments move to a next level, studying the effects of different physical parameterisations on the rate of change of the Antarctic ice sheet. A discussion of this breadth and depth would require more text than could be feasibly added to the current manuscript, which is already almost 30 pages long. We therefore decided to move them to a separate publication, where we can leave out a lot of model description and benchmark experiments (as they will already be in the current manuscript), and spend the available text on a more thorough investigation of these model choices. We will mention these ongoing developments more extensively in the “future research” part of the Discussion section.

Regarding the results for the ABUM/ABUK experiments: the extremely high melt rate of 400 m/yr that is prescribed in the ABUM experiment should lead to a lag of 3 or 4 years at the most with respect to the ABUK results; in the 500-year figures used here and by Sun et al. (2020), this difference would not be visible. In the extreme case of a grounding-line thickness and velocity of 2,000 m and 1,000 m/yr, respectively, this would lead to a shelf of only 2.5 km long, which we doubt would generate enough buttressing to cause any significant differences between ABUM and ABUK. We will reflect these thoughts in the manuscript.

Minor comments

P2, I32: your reference to Bueler and Brown (2009) is not listed in your references. Please add it.

We will do so.

P4, eq2: Please define your integration bounds “b” and “s”, and your exponent “n”. I note that you define “b” later on page 5 and it is always nicer for the reader to know about the different notations as they see them.

Following a suggestion by reviewer #1, we will include a table listing all model symbols and parameters.

P5, I9: please explicitly write down your Neuman boundary condition.

We will do so.

P5, I17: please define z_{sl} .

This will be included in the symbol table.

P6, I1: You mention that λ_w is limited between 0 and 1. As written on P5, I18 it is not bounded to do so. I would suggest to write it as: $\lambda_w = \min(z_{sl} - b, 1000)$, when $b < z_{sl}$, or something similar. If your code does not do that, please modify your text accordingly. Similar remark for w_b and equation 10.

We will add the 0 – 1 limits to the equations.

P3, Sec2: what kind of horizontal grid discretization are you using? On P6, I18 you mention a staggered grid without talking about what kind of staggering you are using. Also, refer to the appendix for more details about it.

A reference to Appendix A was indeed missing from Section 2; we will include it.

P7, I29-30: please write down the equation you are using for basal mass balance for clarity since you are combining 2 methods. Plus, it reduces the amount of paper the reader has to go through.

We will do so.

P9, I3: replace Sec3.4 by Sec 3.2. At least in the printable version, there are no sections 3.2 and 3.3.

We will do so.

P9, Sec.3.4: why did you choose your coarsest resolution to be 5km here while you test much coarser resolution in all the other experiments? See remark above. (You could simply show results at 0.5km, 1km, 2km, 4km, 8km, 16km, 32km.)

We initially chose the same set-up as Bueler and Brown (2009), which describes a narrow (~80 km) ice stream. We agree that it is more meaningful to choose a set-up where the same resolutions as for the other experiments (40, 32, 20, 16, 10 km) can be used. We have redone the experiments with a new set-up, describing a wider (~300 km) ice stream, so that a 40 km resolution can still resolve the low-friction channel. We will update the manuscript to reflect these changes.

P12, Sec.4.3: Please be more quantitative about the grounding line position at the different resolution and at the different stage of the experiment. You could list them in a table. Also, I am not sure I understand the reference to Leguy et al. (2021) in this section as they use the MISMIP3d (Pattyn et al. 2013) and MISMIP+ experiments (Asay-Davis et al. 2016) in their paper, not MISMIP as opposed to the Leguy et al. (2014) paper. Also showing results at higher resolution would be sensible.

The second panel we will add to Fig. 7, showing the magnitude of the grounding-line hysteresis as a function of the model resolution, will help clear this up.

The reference to Leguy et al. (2021) should have been to Leguy et al. (2014), which is the study looking at MISMIP.

The new experiments we have done now go as high as 10 km.

P13, Sec.4.4: Please show a sea level time series of your control experiment.

We will replace Figs. 8 and 9 with a single figure showing ABUC (the control experiment), ABUM, and ABUK. We will also add a few lines to the text to discuss the results of this experiment.

P13, l12: I suggest removing “Before starting the 500-year simulation” and begin your sentence directly with “We initialize”. It could be confusing especially since you are using a relaxation time of 500 year in your spinup procedure which is not part of the simulation.

We will do so.

Fig2: please indicate what $O(R^{1.1})$ represents and what R is

R is the grid resolution; $O(R^{1.1})$ means that the error in the modelled margin position scales with the grid resolution to the 1.1'th power (i.e. approximately linearly). We will clarify this in the figure caption.

Fig 8. And 9.: the results from Sun et al. (2020) show the difference w.r.t to the control experiment. Your figure caption does not state this. Can you clarify whether you subtracted it or not?

Sun et al. (2020) show the sea-level change with respect to the initial (i.e. spun-up) state, not to the sea-level change in the control run. We do the same.