Comment on gmd-2021-352
Anonymous Referee #2

Referee comment on "Benchmarking the vertically integrated ice-sheet model IMAU-ICE (version 2.0)" by Constantijn J. Berends et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-352-RC2, 2022

This manuscript presents the recent development made with the model IMAU-ICE version 2.0. It describes the new numerical options available, and the improvement made since v1.0 of the model. Several experiments are presented with the goal to show the performance of the model with respect to some benchmark experiments well known in the ice sheet modeling community.

The manuscript is well written, easy to follow and written in a logical fashion. I particularly appreciated the author’s effort on presenting high quality figures that were easy to read and understand.

Unfortunately, at this point, I cannot recommend the publication of this manuscript without major revisions. I will highlight my major concerns first followed by minor comments throughout the text.

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.

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 faction 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?

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 MISMIP 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.

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 oultier 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
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.

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 latter, 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.

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.

**Minor comments**

P2, l32: your reference to Bueler and Brown (2009) is not listed in your references. Please add it.

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.

P5, l9: please explicitly write down your Neuman boundary condition.

P5, l17: please define \(z_{sl}\).

P6, l1: You mention that \(\lambda_w\) is limited between 0 and 1. As written on P5, l18 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.

P3, Sec2: what kind of horizontal grid discretization are you using? On P6, l18 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.

P7, l29-30: 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 as to go through.

P9, l3: replace Sec3.4 by Sec 3.2. At least in the printable version, there are no sections 3.2 and 3.3.

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.)
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.

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

Figures

Fig2: please indicate what O(R^1.1) represents and what R is.

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?

References:

