Review egusphere 225-2242:

MinSIA v1: a lightweight and efficient implementation of the shallow ice approximation

General Impression

The submitted manuscript describes the implementation of an ice-flow model based on the Shallow Ice Approximation (SIA) formulated using the vertically integrated mass balance and implemented in MATLAB. The manuscript further discusses measures in spatial smoothing and time-discretization that had to be taken to stabilize the method in order to produce simulations using synthetic forcings of glaciations over selected regions of the Alps and the Black Forest, the latter used as the benchmark case presented in this study.

Whereas the description of the model seems to be almost complete (I ask for a few more details) and implemented along the guidelines of the target journal (GMD), to which general purpose model descriptions fit, I have to express major concerns on the applicability of the model in view of its approximation and the applied slope correction and the evaluation of its performance and reproducibility by the reader, which I try to elaborate in the sections below.

Critical issues

Looking at the abstract text, I already was surprised to read about 25 m spatial resolution (assuming it is the one in horizontal direction) in connection with the applied Shallow Ice Approximation (SIA). This seems to have been confirmed within the text. There is a long series of studies on the validity and accuracy of the SIA by several authors (Jóhannesson, 1992; Gudmundsson, 2003; Hindmarsh, 2004 - to list a few) that later were accompanied by numerical investigations comparing full-stress Stokes (e.g., Le Meur et al. 2004, Seddik et al. 2017) and/or higher order approximation models (Pattyn et al., 2008) to SIA. The quintessence of all these studies is that **SIA is inaccurate** with respect to three aspects that apply to the studies presented in this manuscript, namely,

- fast sliding (though this is difficult to estimate speeds from the information given)
- steep slopes and increased accumulation
- terrain resolutions significantly below a few multiples of the local ice thickness.

The last point, I would even say, is a validity criterion, as shallowness (which also applies to longitudinal gradients to be resolved) is the fundamental parameter behind the expansions of the Stokes equation leading to the zeroth-order equations, the SIA (Hindmarsh, 2004). In other words, based on the above mentioned literature, to me a resolution below about a few times the local ice-thickness is in violation with the principle assumption that go into the derivation of the approximation applied in this study. I am not surprised by the instabilities arising from resolving steep glaciers at 25 m with SIA. I would also like to see a mathematical argument to whether the ad-hoc introduced slope correction in equation (3) is consistent with the assumption of the SIA (I think it is not). The author leaves the reader with no means of quantifying the inaccuracies introduced by the choice of the method in combination with the resolution and steepness of the terrain, since the reference run is provided with the method itself. It either would take comparison with observations (not possible for synthetic ELA forcing) or a higher order (if not Stokes) solution to the problem. To me, these issues seem not to be fixable by improved numerical methods, as they are inherent to the approximation applied. Thus, I would see the necessity to demonstrate the mathematical consistency of the presented model with the zeroth order approximation (SIA). Further, I would expect a clear proof of the advantage of running SIA on – how I conclude - 2 orders of magnitude above the with theory consistent mesh resolutions and quantify the errors introduced by steep slopes/high accumulation/fast sliding to comparison of either observed results or results obtained with to the task suitable approximations to or the Stokes equations. This not necessarily has to be done on domains as big as whole parts of the Alps.

The second concern is linked to the main motivation given by the author to provide a computationally fast method to the community in order to compete with machine learning (ML) algorithms. My deviating opinion on this competition aside (see in specific comments), I see the following **issues** concerning the **description of model performance and reproducibility**, namely

- the proprietary package needed to run the implementation (MATLAB) prohibiting reproduction of the results for readers without a valid license
- the lack of detailed description on applied parallel computation paradigms (if any), like OpenMP (threading and SIMD) or even distributed memory, like MPI. Can the setup perhaps even utilize accelerators?
- the detailed justification on the choice of the methods (one direct and one Krylov) used to solve the sparse matrix problem

I will in detail refer to these items in the specific comments section.

Specific Comments

I indicate the line number as given in the egusphere-document. I display quotes from the manuscript in a greenish colour. It will be pointed out if a comment relates to a major point of critics (previous section)

- Line 2: This paper aims at keeping classical numerics competitive in this field. This is merely a comment conveying my divergent opinion: I do not think that the competition (if there is, as I rather see them as complementary) between machine learning (ML) algorithms and process based models (*classical* in terms used here) is about speed. In my opinion, it is about accuracy and physical completeness. ML reproduces what it has been trained with, and it does this very fast and efficient. The particular reason why these techniques (or at least in my opinion should be) utilized is exactly to speed things up. On the other hand, ML is restricted to a given setup (fixed set of geometry and/or processes) defined by the training set. These surrogate models need new training (which tends to be computationally intensive, so no free lunch also here) to introduce new physics, often even if moving to another topography. That – in my opinion – is where the strength of process models lies – they easier can introduce improved physics and they are more generally applicable. I, though, think it is futile to try to compete in terms of speed with a ML implementation using the same setup in the training set as in a process-based model (of which complexity it ever may be), in particular, if the training phase of the ML is left out of the comparison.
- Line 5: … a lightweight implementation of the new scheme in MATLAB. This relates to the second major point before. In my opinion, it would have been better to wait for the Python implementation to be ready before publishing. At this point in time, the reader needs a licensed software (*MATLAB*) to reproduce results and is left without any insight on the performance of the Python implementation to come. To palliate the first problem, perhaps the author can test with the open source software *Octave* and report to what extend this can be used in replacement to *MATLAB* to run your current version of the code?
- Line 16: Three-dimensional simulations of the Stokes equations with a free surface (as implemented, e.g., in the model Elmer/Ice) and the shallow ice approximation (SIA) are end-members in the hierarchy of ice-flow models. May I in this connection point out that there are existing Stokes simulations of the Western part of the Alps (in fact, including Black Forest, which

was not part of the investigation but the domain) spanning several thousand years around the LGM (Cohen et al, 2018 and 2023).

- Line 17: The SIA considers vertically averaged velocities, assumes hydrostatic pressure, and neglects all stresses arising from horizontal shearing. To my knowledge, SIA defines horizontal velocities as a function of the vertical coordinate (Greve and Blatter, 2009), but only its force-balance considers vertically integrated stresses (expressed in terms of integrals of the vertically varying horizontal velocity field along columns). So, the fluxes, but not the velocities are vertically integrated variables. Could it be that in the presented application the fact that one would have to perform a quadrature of the force balance in each column is circumvented by reducing the solution to the mass-balance and the introduction of the pre-factors f_d and f_s therein? If so, I would ask to explain what assumptions the replacement of the vertical integrals by – what occurs to me – fixed pre-factors are implicitly introduced by this procedure (e.g. on vertical distributions of enhancement factors, temperatures).
- Line 18: In between, there are two-dimensional approaches that account for these stresses, … . Even Stokes can be solved in two dimensions (flow line). I would suggest to write depth-integrated instead.
- Line 26: The still limited performance of two-dimensional models arises from a combination of assuming hydrostatic pressure and the explicit time-stepping scheme implemented in almost all contemporary models (Bueler, 2023). Like before, I would suggest using depth-averaged or integrated models, instead. Further I would ask to explain what performance is referred to? Pure computational performance or beyond that also in terms of accuracy and stability? In view of the latter (which links to my primary point of criticism), I would drop the "still", as to me it implies that the instabilities/inaccuracies described in this manuscript are something that could be overcome. In my opinion, they are built-in, as models based on hydrostatic approximation are not capable to resolve horizontal gradients in stresses that act on scales smaller than a few multiples of the ice thickness and assume small slopes (Hindmarsh, 2004) .
- Line 62, eq. (3) and line 68 and eq. (5): The factor involving the cosine of the slope angle β of the ice surface with … is a simple correction for steep slopes. I will come back to this later also in the appendix, but to me it appears that introducing this slope correction violates a principal assumption of the SIA that the surface normal in lowest order points exactly into the vertical direction which by shallowness also would demand that $\cos\beta \approx 1 \;\rightarrow\; \beta \approx 0$. In the other extreme, for $\beta \rightarrow \pi/2$ the 5th and 8th power of $\cos\beta$ in (9) will result in a vastly vanishing pseudo diffusivity and (8) basically reduces to the geodetic mass balance. In other words, mass transport shuts down on steep slopes. That in my view is artificial and covers the fact that the hydrostatic pressure approximation is not valid on slopes significantly deviating from zero (Hindmarsh, 2004) - see point one in major critics.
- Line 67: The factor f_d mainly depends on temperature in reality. This factor should also contain some contributions from the vertical integral of the force balance, including the temperature dependence and enhancement factors. As demanded before, I think for completeness, the exact composition of f_d as well as f_s leading to the constant values reported in the text should be written out and explained – if not in the main part, at least in the Appendix. Further, as this seems to be the only mentioning of temperature, which, by the Arrhenius factor (Greve and Blatter, 2009) has a strong influence on the ice viscosity and hence the flow: I would kindly ask how the model – if at all - treats temperature variations in the ice, in particular as I would think that during a glaciation of the Alps the thermal ice conditions significantly vary, both, spatially and in time? Would the presented model be in principle capable of including advective temperature transport?

- Line 70: The correction is explained in the derivation of Eq. (3) in Appendix A. It should, however, be mentioned that it is less elaborate than the correction for rapid mass movements developed by Savage and Hutter (1989) and only exact if the ice surface is parallel to the bed. In particular, it does not capture the effect of steep walls in a valley. I discuss this in detail later where the method is explained in the Appendix. As mentioned before, I think the cosine in (3) is in violation of the basic assumptions behind SIA. I also do not see the parallel slab problem of a creeping shear thinning fluid in a direct connection to the Savage-Hutter theory, which is an expansion of the depth-averaged Navier-Stokes equations (i.e., including acceleration terms) for yielding granular flows, neglecting terms of order $O\left(\epsilon^{3/2}\right)$ of the aspect ratio ϵ, and in contrary to the SIA, usually already defined in a locally rotated coordinate system (e.g., Zwinger et al. , 2003).
- Line 80: It should, however, be emphasized that modeling sliding along the bed is still one of the major challenges in this field. It is even questionable whether this process can be described well by the SIA or whether lateral stress components have to be taken into account. Please, refer to literature where these topics are addressed. For the latter: It is established (see literature cited in the major points section) that SIA is unsuited for fast sliding conditions.
- Line 89: Formally, it looks like a nonlinear diffusion equation, although it is mathematically not a diffusion equation due to the occurrence of ∇s in D. Can you please elaborate: If not a diffusion equation, what is it then?
- Line 120: The upstream scheme ensures D = 0 in this situation and thus avoids a permanent systematic increase in ice volume. What about the considerations on the second term of the r.h.s. of (8) at the contact line in order to avoid negative heights? There is often negative net mass balance, r, at the front of the glacier. Is it masked out? Generally speaking, to stay consistent in mass balance one should solve the mass balance or in higher order models the kinematic boundary condition together with the inequality for positive or zero flow-heights (see, e.g., Gagliardini et al. 2013, section 6.5) , which in my opinion should be applicable to the scheme applied in this manuscript.
- Line 128: Therefore, a dynamic smoothing of the slope of the ice surface in the diffusivity $D(h,|\nabla s|)$ is the central novel idea in MinSIA. The idea was originally motivated by Airy's linear theory of ocean waves (gravity waves in an inviscid fluid with a free surface). According to this theory, oscillations in pressure and particle velocity decrease exponentially with depth below the surface. The respective depth of penetration is proportional to the wavelength. Based on this result, MinSIA uses a smoothed version of $|\nabla s|$ with the length scale of smoothing being proportional to the ice thickness h. I am not able to see a connection from a theory based on gravity waves in Newtonian fluids to the mechanical behaviour of a (strongly non-linear and highly viscous) shear thinning free-surface thin-film Stokes problem. My theory: this smoothing algorithm works because it re-introduces the basic principle of thin-film approximations to resolve horizontal gradients with respect to a large aspect ratio (i.e., over several ice-thicknesses), even if run on mesh resolutions that apparently violate this assumption. To probe that theory, I suggest to run the simulation on a to the SIA appropriate mesh resolution (>1000 m range) without smoothing and report on instabilities of such a run. I would also ask to include a clear motivation to opt for such small horizontal resolutions of 25 m – 100 m in connection with SIA. Thinking of the main motivation presented in the text, a fast solution, I would even say that increasing the resolution should work exactly in this direction as the problem size gets smaller. This links to the first in the main point of critics.
- Line 163: The preconditioned conjugate gradients (PCG) method is used for solving the linear equation system. An incomplete Cholesky factorization is used as a preconditioner, whereby the version that compensates dropped nondiagonal elements at the respective diagonal elements

turned out to be particularly suitable. A direct solver is also implemented for testing. As mentioned in the second point of major critics. I would ask for a more detailed explanation of this particular choice. What method was used for the direct solver? Were these methods utilizing specific features of the hardware (shared-memory or distributed memory parallelism, vector units, accelerators, if present)? What can the reader expect if this model is run on other, more modern platforms than the CPU that was reported in the text? Are there special license conditions to utilize the mentioned methods in MATLAB?

- Line 187: As a first test, the accuracy of the solutions for different values of the smoothing factor *f* with respect to the reference solution (*f* = 0.25, δt = 64 yr) is measured. In relation to the major point of critics, I do not think that comparison to a result obtained with the method is able to show the accuracy of the method itself. To me, it merely shows the sensitivity of the method to a change in numerical parameters and resolutions.
- Line 209: Since even explicit schemes are stable for sufficiently small time increments, the occurrence of the staircase oscillation must be related to both the smoothing factor *f* and the time increment δt. I would like to see a proof on the statement of stability for small timesteps. Either by citing literature or running the model (at least for some time) with no smoothing on the 25 m resolution. What I think, is that the oscillations are a reaction of the SIA to significant undulations with a length-scale way below the ice-thickness. Consequently, I would conclude that the oscillations are natural to the SIA (if run on resolutions that violate the shallowness assumption) and damped out by the smoothing, similar to as they would by running the SIA without smoothing on coarser meshes (which I already asked you to test) that resolve horizontal scales above the typical ice-thicknesses.
- Line 267: The simulations over the entire time span with δt = 0.25 yr took 35400 s for δx = 25 m, 6740 s for δx = 50 m, and 1290 s for δx = 100 m on an Intel Core i5-7600 CPU (3.50 GHz) from 2018. This, to me, seems to be the only information on how some sort of performance evaluation has been done. In my view, this is not enough such that the reader can get a clear picture what to expect if running the model themselves. I would ask to specify the utilization of the CPU (single core or multiple cores, as there seem to be 4 included in the reported model). If multiple cores have been utilized, please report on what parallel paradigm was deployed (shared or distributed memory). How much memory was available in the test platform and how was the saturation of the available memory at the runtime? This links to the second point of my major critics.
- Line 288: Figure 8 illustrates the weakness of the scheme at steep slopes for δx = 25 m and δt = 0.25 yr. I would rather say it reveals the inapplicability of the approximation than a weakness in the numerical scheme (see major point of critics).
- Line 305: At present, a MATLAB implementation of MinSIA is available under the GNU General Public License. A Python implementation is under development. As mentioned under the major points and before in this section – despite the setup being shared with an open license – the proprietary nature of the software needed to run it imposes a big hurdle for people to reproduce these findings. This all would have been no problem if the publication would have been postponed to the point in time when the announced Python version would have been available.
- Line 334: A 2-element array corresponds to $r_{\max} = \infty$ and a scalar value additionally to $g_- = g_+$. I am not able to understand the meaning behind this sentence. Can you please elaborate? Do you allow for infinitely large accumulation rates?
- Line 390: Let us, for simplicity, assume that the ice surface is parallel to the bed with a slope angle β and that the z-coordinate is perpendicular to the surface with z = 0 at the bed. I think this is the culprit of the discrepancy about the bed-slope correction I raised in the main issues. The parallel slab example (see Greve and Blatter, 2009), which seems to be adapted here, in my view, cannot

be intermixed with the derivation of the SIA. The basic assumption behind the SIA is a directly into the vertical direction pointing surface normal (Greve and Blatter, 2009). The resulting full alignment with the negative direction of the gravity is a necessary condition for the zero Cauchy stress vector at the surface leading to identically vanishing stress-components and pressure, which constitutes another necessary condition for the vertical integration of the shear stresses (Greve and Blatter, 2009) in SIA. If the integration in the vertical direction is done from bedrock, b, to the free surface *h,* one obtains the shear stress at the bedrock as $\tau = \rho g\ h\ \nabla s$, where $h = s - b$, i.e., the total ice-depth in vertical direction. Thus, in SIA the shear stress is imposed by the hydrostatic pressure gradient induced by the horizontal change in surface elevation (with no influence of the bedrock slope at all), whereas in the parallel slab, i.e., equation (A1), the shear stress is the result of the downslope weight component of a in the rotated reference frame bed-parallel surface with in this rotated reference frame identically vanishing gradients of the free surface. As I see it, the author is setting (A1) identical to the bed-shear stress as defined in SIA, which I claim violates the basic assumptions of the SIA. If the author disagrees, I would ask to include a detailed derivation of the SIA (starting from the Stokes equations) in a rotated reference frame and show that the derived equations still comply with the lowest order approximation, even for large bedrock angles, as they occur in the studied topographies of the Alps and even the Black Forest.

Typos and technical corrections

Line 245: PGC -> PCG

Line 281: … which ~~is~~ also is still stronger than …

Line 396: $\tau(z)^n$ suggestion to change to $\tau^n(z)$

Line 409: suggestion: first term -> first multiplier on the right-hand-side

References

Cohen, D., F. Gillet-Chaulet, W. Haeberli, H. Machguth, and U.H. Fischer, 2018. *Numerical reconstructions of the flow and basal conditions of the Rhine glacier, European Central Alps, at the Last Glacial Maximum*, The Cryosphere, **12**, 2515-2544, doi: 10.5194/tc-12-2515-2018

Cohen D., G. Jouvet, T. Zwinger, A. Landgraf and U. H. Fischer, 2023. *Subglacial hydrology from high-resolution ice-flow simulations of the Rhine Glacier during the Last Glacial Maximum: a proxy for glacial erosion*,. E&G Quaternary Sci. J., **72**, 189–201, doi:10.5194/egqsj-72-189-2023

Gudmundsson, G. H., 2003, *Transmission of basal variability to a glacier surface*, J. Geophys. Res., 108(B5), 2253, doi:10.1029/2002JB002107.

Greve and Blatter, 2009. Dynamics of Ice Sheets and Glaciers, Springer

Hindmarsh, R.C.A. , 2004, *A numerical comparison of approximations to the Stokes equations used in ice sheet and glacier modelling*, J. Geophys. Res.: Earth Surface, 109, F01012, doi:10.1029/2003JF000065

Johannesson, T., 1992, *The landscape of temperate ice caps*, Ph.D. thesis, University of Washington, Seattle

Le Meur E., O. Gagliardini, T. Zwinger, J. Ruokolainen, 2004, *Glacier flow modelling: a comparison of the Shallow Ice Approximation and the full-Stokes equation*, C. R. Physique **5**, 709-722, doi:10.1016/j.crhy.2004.10.001

Pattyn, F., Perichon, L., Aschwanden, A., Breuer, B., de Smedt, B., Gagliardini, O., Gudmundsson, G. H., Hindmarsh, R. C. A., Hubbard, A., Johnson, J. V., Kleiner, T., Konovalov, Y., Martin, C., Payne, A. J., Pollard, D., Price, S., Rückamp, M., Saito, F., Souček, O., Sugiyama, S., and Zwinger, T., 2008, *Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–HOM)*, The Cryosphere, 2, 95–108, doi:10.5194/tc-2-95-2008

Seddik, H., R. Greve, T. Zwinger, and S. Sugiyama, 2017. *Regional modeling of the Shirase drainage basin, East Antarctica: full Stokes vs. shallow ice dynamics*, The Cryosphere, **11**, 2213-2229, doi:10.5194/tc-11-2213-2017.

Zwinger, T., Kluwick, A., and Sampl, P., 2003, *Numerical Simulation of Dry-Snow Avalanche Flow over Natural Terrain*. In: Hutter, K., Kirchner, N. (eds) *Dynamic Response of Granular and Porous Materials under Large and Catastrophic Deformations*. Lecture Notes in Applied and Computational Mechanics, vol 11. Springer, Berlin, Heidelberg, doi:10.1007/978-3-540-36565-5_5