Author's response to RC2

Text in black: Reviewer's comments
Text in blue: Author's response

This manuscript introduces a 3D dynamics model for mountain glaciers based on the Community Ice Sheet Model, and its application to study the evolution of mountain glaciers using protocols from the third phase of the Glacier Model Intercomparison Project (GlacierMIP3). While the paper is overall well written, I have several minor concerns that should be addressed before the manuscript is considered suitable for publication.

We thank the reviewer for their constructive feedback. We have incorporated their suggestions and provide detailed responses to their comments below.

I would like the authors to better motivate, in the Introduction, the need to use high-fidelity models for mountain glaciers, particularly considering their increased sensitivity to data errors. The authors touch on this in the Discussion, but this should be featured in the Introduction as well.

We have included a motivation to include glacier modelling within an ESM framework in the introduction:

"Integrating glacier modeling within an Earth System Model (ESM) framework offers several emerging advantages for studying glaciated regions. It will enable dynamic coupling with the climate system, enhancing the representation of feedback mechanisms with the land-atmosphere-hydrology components. This integration will allow more comprehensive assessments of climate, ecological, and hydrological impacts across glaciated regions worldwide."

The initialization of the model is done with ad hoc tuning methods (e.g., Pollard and DeConto, 2012). However, I think the reader would benefit from a discussion of more advanced initialization techniques based on (transient) PDE-constrained optimization methods, which have become increasingly standard in the literature and offer a more rigorous approach to model initialization.
Pollard, D. and DeConto, R. M.: A simple inverse method for the distribution of basal sliding coefficients under ice sheets, applied to Antarctica, The Cryosphere, 6, 953–971, https://doi.org/10.5194/tc-6-953-2012, 2012.

We agree with the reviewer on the use of more sophisticated approaches in literature for initialization. We used the approach described in the manuscript for its simplicity, computational efficiency, and ease of implementation, and because we have used a similar method for ice sheet studies.

To address the reviewer's concern, we have included a new Section 7.2 on "*Glacier initialization*" under Section on "Model limitations and future work".

"CISM uses a simple, computationally efficient inverse method to estimate the spatial distribution of the basal friction coefficient Cp, adjusting the values to minimize the mismatch between modeled and observed ice thickness (Sect. 4.2). It is similar to the method developed by Pollard and DeConto (2012) to derive basal sliding coefficients for Antarctica. Other

studies have used more sophisticated approaches such as adjoint-based optimization methods, which compute gradients of a cost function (typically the mismatch between modeled and observed velocities) with respect to control parameters using the adjoint of the governing equations, or transient (time-evolving) inversion that assimilates time series of observations (Morlighem et al., 2013; Goldberg and Heimbach, 2013; Perego et al., 2014). At present, CISM does not have these capabilities for ice-flow modeling."

Detailed comments:

Equation (1): For clarity and completeness, please include the y-component of the DIVA model alongside the x-component.

Done - the equation for the y-direction is also added:

$$
\begin{aligned}
&\frac{1}{H}\frac{\partial}{\partial x}\left[2\bar{\eta}H\left(2\frac{\partial\bar{u}}{\partial x}+\frac{\partial\bar{v}}{\partial y}\right)\right]+\frac{1}{H}\frac{\partial}{\partial y}\left[\bar{\eta}H\left(\frac{\partial\bar{u}}{\partial y}+\frac{\partial\bar{v}}{\partial x}\right)\right]+\frac{\partial}{\partial z}\left(\eta\frac{\partial u}{\partial z}\right)=\rho_i g\frac{\partial s}{\partial x},\\
&\frac{1}{H}\frac{\partial}{\partial x}\left[\bar{\eta}H\left(\frac{\partial\bar{u}}{\partial y}+\frac{\partial\bar{v}}{\partial x}\right)\right]+\frac{1}{H}\frac{\partial}{\partial y}\left[2\bar{\eta}H\left(\frac{\partial\bar{u}}{\partial x}+2\frac{\partial\bar{v}}{\partial y}\right)\right]+\frac{\partial}{\partial z}\left(\eta\frac{\partial v}{\partial z}\right)=\rho_i g\frac{\partial s}{\partial y},
\end{aligned}
\tag{1}
$$

Equation (2): Please clarify whether the cap applies to the absolute values of the separate x and y derivatives of the surface elevation, or to the magnitude of the surface elevation gradient. Either way, this seems a bit of a crude fix. How frequently is this correction active in simulations performed in this study?

The cap applies to the absolute values of the separate x and y derivatives at each cell edge. We have added the clarification in the revised text (Section 2.1):

"When solving Eq. (1), CISM has an option to cap the magnitude of the surface slope ($\|\partial s/\partial x\|$ at east and west cell edges and $\|\partial s/\partial y\|$ at north and south edges) at a value of m_{max} to maintain model stability in regions of steep topography. For this study we set $m_{max} = 1.0$ (a 45° slope). The cap is applied to about 1% of ice-covered cell edges in the Alps domain described in (Sect. 4)."

We agree that slope-limiting is a crude fix. We have used this approach in Greenland and Antarctic ice-sheet simulations where relatively few cells exceed the limit, but in the steep terrain of the Alps, the limit was being applied to a substantial fraction of the ice-covered cells.

For the revised simulation, we carried out a new set of runs with $m_{\max} = 1.0$, i.e. an angle of 45° relative to the horizontal. With this change, the slope is limited at only about 1% of cell edges. When we first made this change, flow speeds increased over steep terrain and the spun-up ice was too thin. However, we found that increasing the values of the C_p^{min}, C_p^{init}, and C_p^{max} parameters from 3K/30K/100K to 5K/50K/200K reduced the maximum speeds (so that we did not have to reduce the 1-month timestep) and enabled a good match to the target thicknesses. The rms error between the simulated and target thickness is now 12.5 m, compared to 14.8 m in the previous submission. This suggests that minimizing the use of the slope cap has improved the representation of the flow.

We highlight that the new results, based on these changes, are consistent with the original simulations, and our analysis remains unchanged.

Equation (3): The scalar representation of the sliding law is potentially confusing, as basal shear stress and basal velocity are vector fields.
Please rewrite this equation in vector form, for example: tau_b = C_p |u_b|^{1/m-1} u_b.

Done - we have modified the equation as suggested:

$$\tau_b = C_p |\mathbf{u}_b|^{\frac{1}{m}-1} \mathbf{u}_b,$$

Section 5.1: Data Blocks and Repartitioning: The term "data block" needs to be defined. Are data blocks the fundamental units used for distributing computational workload across processor cores? How is the repartitioning of the data blocks performed, especially considering they are now unstructured?

Yes, the blocks are the fundamental units for distributing the workload. Each square block of data is assigned to one processor core. The resulting blocks are not unstructured. Rather, we can think of them as squares on a checkerboard in which some squares are labeled as active, while the rest are labeled as inactive. The active blocks are initialized with information that tells them which adjacent blocks are also active. During the run, each active block exchanges information only with neighboring blocks that are also active.

We have added additional information in Section 6.1 (*Spatial resolution and computational efficiency*) to clarify how this works:

"On initialization, CISM partitions the global domain into square blocks of data. For the 100-m grid, the initial domain consists of about 9300 blocks, each with 75 or 76~cells on a side. Of these blocks, only about 900 contain one or more of the grid cells included in the glacier mask. CISM labels these blocks as active, assigns one block to each processor core, and discards the remaining blocks. We modified CISM's parallel routines (halo updates, gather/scatters, global sums, and broadcasts) to operate only on active blocks, exchanging data with adjacent blocks that are also active. This allows a tenfold reduction in cost compared to a simulation with inactive blocks included."