~~Numerical issues in modeling~~ Modeling sensitivities of thermally and hydraulically driven ice stream surge cycling

Kevin Hank[1], Lev Tarasov[1], and Elisa Mantelli[2,3,4]

[1]Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL, A1B 3X7,Canada
[2]Department of Earth and Environmental Sciences , Ludwig-Maximillians-Universitaet Munich, Theresienstr. 41, 80333 Munich, Germany
[3]Alfred Wegener Institute for Polar and Marine Research, Am Alten Hafen 26, 27568 Bremerhaven, Germany
[4]Institute for Marine and Antarctic Studies, University of Tasmania, 20 Castray Esplanade, Battery Point TAS 7004, Australia
[*]khank@mun.ca

Correspondence: Kevin Hank (khank@mun.ca)

Abstract. Modeling ice sheet instabilities is a numerical challenge of potentially high real-world relevance. Yet, differentiating between the impacts of model physics, numerical implementation choices, and numerical errors is not straightforward. Here we use an idealized North American geometry and climate representation (similar to the HEINO experiments, Calov et al., 2010) to examine the process and numerical sensitivity of ice stream surge cycling in ice flow models. Through sensitivity tests, we identify some numerical requirements for a more robust model configuration for such contexts. To partly address model-specific dependencies, we use both the Glacial Systems Model (GSM) and Parallel Ice Sheet Model (PISM). We show that modeled surge characteristics are resolution-dependent though converging (decreasing differences between resolutions) at finer horizontal grid resolutions. Discrepancies between fine and coarse horizontal grid resolutions can be reduced by incorporating sliding at sub-freezing temperatures. The inclusion of basal hydrology increases the ice volume lost during surges, whereas the dampening of basal temperature changes due to a bed thermal model leads to a decrease.

1 Introduction

1.1 Motivation and background

The use of ice sheet models has grown at least an order of magnitude over the last two decades. The relevance of such modeling studies to the actual physical system can be unclear without careful consideration and testing of numerical aspects and implementations. This is especially true when modeling the highly non-linear ice sheet surge instability, which has significant implications not only for the ice sheet itself but also for the climate. In fact, it is often difficult to assess whether model results are physically significant (effects of physical system processes), a consequence of model-specific numerical choices,

or a combination of both. ~~This is especially important in the case of abrupt changes.~~ Whether ice sheet instabilities observed in numerical simulations are the result of physical instabilities of the underlying continuum models or spurious effects of the discretization and numerical implementation of said models has long been debated (e.g., Payne et al., 2000; Hindmarsh, 2009) and is a consequential matter. The present study is concerned with characterizing the impact of model physics, numerical choices, and numerical errors on ice stream surge cycling.

Binge-purge ice stream cycling was first introduced in the glaciological literature by MacAyeal (1993) as an explanation for Heinrich Events arising from the former Laurentide Ice Sheet (LIS) in the Hudson Bay/Hudson Strait region. The key idea is that the ice stream gradually grows to a threshold thickness (binge phase) driven by surface accumulation. Once the ice stream is thick enough to sufficiently isolate the ice stream base from the cold surface, heat from geothermal and deformation work sources can slowly bring the basal temperature to the pressure melting point. The bottom layer of the ice stream is no longer frozen to the bed and thus enables basal sliding. Localized warm-based ice streaming increases the ice stream surface gradient (steeper slope) at the warm/cold-base transition point, leading to an increase in driving stress. The resultant increase of heat from deformation work can warm the surrounding ice (close) to the pressure melting point, thus enabling (sub-temperate) basal sliding (Fowler, 1986). When the melting point is reached, the presence of water at the ice sheet/bed interface (Fowler and Schiavi, 1998) as well as in a deformable sediment layer (Bueler and Brown, 2009) can further increase sliding velocities. Instead of the slow deformation flow (ice creep), the ice stream now flows rapidly (purge phase). As a consequence of the high ice velocities, the ice stream thins and cold ice is advected from either upstream or the lateral boundaries of the ice stream. Cold ice advection in combination with changing heat source contributions (from both deformation work and basal sliding) and lowering of the pressure melting point as ice thins eventually leads to refreezing of the ice/bed interface. The first localized frozen patch of ice acts as a *sticky spot*, supporting some of the driving stress and decreasing the velocities and heat production in the adjacent ice. This marks the end of the surge, thus enabling the ice stream to enter the next binge phase. Whether hydraulically or thermally driven, these activation (purge) and stagnation (binge) phases can alternate in a quasi-periodic fashion (e.g., Souček and Martinec, 2011) - this is what we refer to as 'ice stream surge cycling' in the remainder of this paper.

As a result of the physics involved and the behaviors expected, modeling of ice stream surge cycling is challenging. The challenges entail, among others, rapid surge onset, high ice velocities, and non-linear (thermo-viscous, hydraulic, and thermo-frictional) feedbacks. In addition to the physical complexity, further challenges arise in the numerical modeling of ice stream surge cycling, whether in terms of model choices (e.g., choice of mechanical model, thermal modeling of the substrate, accounting for sub-glacial hydrology) and/or in terms of their numerical implementation (e.g., grid size, convergence under grid refinement, etc.).

Our focus here is on the challenges arising from numerical modeling, both those related to the physical system being modeled and those related to the numerical implementation. Previous studies examining numerical aspects of surge cycling suggest strong sensitivities in model response to implementation choices such as grid size (e.g., Calov et al., 2010; Roberts et al., 2016; Ziemen et al., 2019). The effects of different approximations of the Stokes equations have been previously addressed (e.g., Brinkerhoff and Johnson, 2015), and are therefore not discussed here.

The discretization and related numerical implementation choices (e.g., grid resolution and grid orientation) have been shown to affect numerical results. As far as the choice of grid is concerned, Ziemen et al. (2019), for example, find a constantly active ice stream at $40\ \mathrm{km}$ grid resolution and oscillatory behavior at $20\ \mathrm{km}$ grid resolution. They argue that this finer grid resolution is necessary to resolve the Hudson Strait properly. A few other studies examine the effect of different grid resolutions on surge behavior (e.g., Payne and Dongelmans, 1997; Greve et al., 2006; Van Pelt and Oerlemans, 2012; Brinkerhoff and Johnson, 2015; Roberts et al., 2016), but an in-depth numerical analysis of Hudson Strait ice stream surge cycling (to whatever idealized form) is entirely absent from the literature. In terms of grid rotation, Greve et al. (2006) and Takahama (2006) show only a minor effect of grid rotation on the general features of the oscillations.

An additional level of complexity in the modeling of ice sheet surge cycling arises from the fact that small perturbations of the initial or boundary conditions can significantly vary the surge characteristics (Souček and Martinec, 2011; Mantelli et al., 2016). For example, Souček and Martinec (2011) show that low levels of surface temperature noise can lead to chaotic behavior in the periodicity of ice stream oscillations, with mean periods varying by $\pm 2\ \mathrm{kyr}$ ($\sim 20\ \%$ of characteristic period of the oscillations, Fig. 8 in Souček and Martinec (2011)). Moreover, Souček and Martinec (2011) find differences in form, period, and amplitude of oscillations when using two different numerical implementations for calculating the basal temperature for thermal activation of basal sliding. However, whether this observed sensitivity arises from physical grounds (e.g., as in Mantelli et al., 2016) or is a spurious numerical effect, the numerical error remains unclear. Souček and Martinec (2011) thus rightfully conclude that '*... the implementation of surge-type physics in large-scale ice-sheet models is rather problematic since the information about the physical instability may be lost in the numerics*'.

1.2 Study overview

Herein, we disentangle the effects of numerical choices (e.g., grid size) and physical system processes (e.g., sub-temperate basal sliding) on ice sheet surges via numerical experiments.

In terms of ice flow models, we primarily use the 3D glacial systems model with hybrid shallow shelf/ice physics (GSM, Tarasov et al., 2023). However, to mitigate the possibility that our conclusions are biased by specific numerical/modeling choices within the GSM, we repeat experiments that do not require implementation of novel physics with the widely used Parallel Ice Sheet Model (PISM, Bueler and Brown, 2009; Winkelmann et al., 2011). As the two model setups and physics are somewhat different (see Table 2 for details), this permits more confident conclusions that are not model specific. ~~In order to~~ To partly address potential non-linear dependencies of surge cycling on model parameters, we run each ~~of our numerical experiments~~ numerical experiment with a high variance ensemble of 5 GSM and 9 PISM parameter vectors instead of just a single run.

In terms of different numerical choices, the impact on model results is usually determined by calculating the model error to the exact analytical solution. However, the theory behind the surge instability is not fully developed (no analytical solution exists) in the context of a spatially extended 3D system, thus precluding systematic benchmarking of numerical models.

To overcome this issue and provide at least a minimum estimate of the numerical model error, we first determine '**M**inimum **N**umerical **E**rror **E**stimates' (MNEEs). This is a ~~new metric that aims to minimally~~ minimal threshold to resolve whether a change in surge characteristics due to changes in the model configuration is significant (see Sec. 2.3 for details).

Equipped with these tools, we set out to tackle the research questions detailed in Sec. 1.3, which we denote with labels $Q_1 - Q_{11}$. The remainder of the paper is then structured as follows: we start by describing our models and experimental setups in Sec. 2. We then present detailed results that allow us to answer our research questions in Sec. 3, with a concise summary and discussion provided in Sec. 4. The results are organized into the following main themes: key surge characteristics of the reference setup (Sec. 3.1), MNEEs (Sec. 3.2), sensitivity experiments with and without a significant (with respect to the MNEEs) effect on the results (Sec. 3.3), and convergence study (Sec. 3.4).

1.3 Research questions

In this subsection, we detail the key research questions that we address through numerical experiments. Following the above-described structure in the description of the results, the research questions are divided into three sub-categories: minimum numerical error estimates (MNEEs), sensitivity experiments, and convergence study.

Minimum numerical error estimates

Q_1 *What is the threshold of MNEEs in the two models (Sec. 3.2)?*

Sensitivity experiments

We examine the significance of different model configurations on the surge characteristics. We are particularly interested in model configurations affecting the basal temperature and thus the surge behavior. Therefore, we first discuss the change in surge characteristics due to a bed thermal model (Q_2) and modeling choices affecting the basal temperature at the grid cell interface where the ice velocities are calculated (Q_3 and Q_4), including the basal sliding thermal activation criterion (Q_5). Previous studies examining the effects of ice stream behavior are often based on an idealized basal topography and sediment distribution and do not consider sub-glacial hydrology (e.g., Calov et al., 2010; Brinkerhoff and Johnson, 2015). Therefore, we determine the change in surge characteristics due to these aspects in Q_6, Q_7, Q_8, and Q_9, respectively. Since thermally and hydraulically driven ice stream surges are not exclusive, we also investigate the differences between the two mechanisms when used as the primary smoothing mechanism at the warm/cold-based transition zone (Q_{10}).

Q_2 *Is the inclusion of a bed thermal model a controlling factor for surge activity (Sec. 3.3.1)?*

Except for PISM, all models in the HEINO experiments did not include a bed thermal model (Calov et al., 2010). PISM is one of the few models that did not show oscillatory behavior in the HEINO experiments (except for experiment T1 (10 K colder minimum surface temperature, Calov et al. (2010))). We explore the role of the additional heat storage on surge activity by deactivating a $1\ \mathrm{km}$ deep bed thermal model in the GSM and PISM.

Q_3 *Do different approaches to determining the grid cell interface basal temperature significantly affect surge behavior, and if yes, which one should be implemented (Sec. 3.3.2)?*

On a staggered grid (commonly Arakawa C grid, Arakawa and Lamb, 1977), the velocities are calculated at the grid cell interfaces, whereas basal temperatures are situated in the grid cell center. Therefore, the basal temperature at the grid cell interface needed for the thermal activation of basal sliding needs to be determined as a function of the basal temperatures at the adjacent grid cell centers. Here we examine surge sensitivity to different interpolation schemes (see Sec. 3.3.2).

Q_4 *How much of the ice flow should be blocked by upstream or downstream cold-based ice, or equivalently, what weight should be given to the adjacent minimum basal temperature (Sec. S8.1)?*

At relatively coarse horizontal grid resolutions (e.g., $25\ \mathrm{km}$), the basal temperatures at the adjacent grid cell centers are of physical relevance. For example, a cold-based grid cell in the downstream direction should block at least part of the ice flow across a $25\ \mathrm{km}$ long warm-based interface (Eq. (S1)). Here we examine surge sensitivity to a change in the weight of the adjacent (grid cell center) minimum basal temperature when calculating the grid cell interface temperature.

Q_5 *How different are the model results for different basal temperature ramps and what ramp should be used (Sec. 3.3.3)?*

Another issue that is often ignored is the basal sliding thermal activation criterion. Based on the results of Souček and Martinec (2011), the basal temperature is a critical factor in the onset and termination of (surging) ice streams. Mantelli et al. (2019) show that an abrupt onset of sliding at the transition from a cold-based ice sheet to an ice sheet bed at the pressure melting point causes refreezing on the warm-based side and, therefore, cannot exist. Observational and experimental evidence for sub-temperate sliding further supports a smooth transition from cold-based no-sliding conditions to fully warm-based sliding, with sliding velocities increasing as the basal temperature approaches the pressure melting point (Barnes et al., 1971; Shreve, 1984; Echelmeyer and Zhongxiang, 1987; Cuffey et al., 1999; McCarthy et al., 2017).

An additional argument for sub-temperate sliding can be made on numerical grounds for coarse horizontal grid resolutions. It is unlikely that an entire grid cell reaches the pressure melting point within one time step (e.g., $25\text{x}25\ \mathrm{km}$ in $1\ \mathrm{yr}$). Furthermore, a sub-grid path at the pressure melting point would likely occur before the whole grid cell reaches the pressure melting point. As such, the activation of basal sliding should start at grid-cell basal temperatures below the pressure-melting point and ramp up as the pressure-melting point is approached. As the horizontal grid resolution becomes finer, the range of sub-grid temperatures in a grid cell decreases (e.g., Figs. 10, S27, and S28). Consequently, the thermal activation ramp should be sharper (smaller transition zone) for finer horizontal grid resolutions.

Experimental work (e.g., Barnes et al., 1971; McCarthy et al., 2017) supports the notion of sub-temperate sliding within a narrow range of temperatures below the pressure melting point ($< 5^\circ$C). A wide temperature ramp (e.g., ~~$T_{\mathrm{ramp}} = 1^\circ$~~$T_{\mathrm{ramp}} = 1^\circ$C, see Eq. (9)) enables an earlier sliding onset (for increasing basal temperature), spatially extended sliding, and a prolonged sliding duration (for decreasing basal temperature).

We use basal temperature gradients in fine-resolution runs and approximations of the sub-grid warm-based connectivity between the faces of, e.g., a $25\ \mathrm{km}$ grid cell (there should be no ice streaming across the grid cell if a frozen sub-grid

area disconnects warm-based patches) to constrain an a priori functional form of the basal temperature ramp. We then use upscaling and resolution-scaling experiments to constrain the dependency of the ramp on horizontal grid resolution.

Q_6 *Does the abrupt transition between a soft and hard bed significantly affect surge characteristics (Sec. 3.3.4)?*

An abrupt transition from hard bedrock to soft sediment (as, e.g., used in the HEINO experiments Calov et al., 2010) can lead to additional localized shear heating caused by the difference in basal resistance and therefore sliding velocities at that transition. We explore the impact of the bed-type transition on surge characteristics by incorporating a smooth transition from $0\ \%$ sediment cover (hard bedrock) to $100\ \%$ (soft) sediment cover effectively changing the basal sliding coefficient $\boldsymbol{C}$ in Eq. (6d).

Q_7 *How does a non-flat topography affect the surge behavior (Sec. 3.3.4)?*

Given the topographic lateral bounds of Hudson Strait, we examine the effects of a non-flat topography on the surge characteristics.

Q_8 *What is the effect of a simplified basal hydrology on surge characteristics in the GSM (Sec. 3.3.5)?*

The implementation of a fully-coupled basal hydrology model changes the basal drag and, therefore, has the potential to affect the surge characteristics. A basal hydrology model coupled to an effective-pressure dependent sliding law, or a Coulomb-plastic bed (as in PISM), introduces a positive feedback such that larger sliding speeds increase frictional heating, and thus meltwater availability which further weakens the bed and leads to even faster sliding. Different basal hydrology process representations have been proposed in the literature (e.g., a 0D (Gandy et al., 2019), poroelastic (Flowers et al., 2003), or linked cavity hydrology model (Werder et al., 2013)), and in-depth comparison is currently under review (Drew and Tarasov, 2022). Here we compare GSM surge statistics with and without a fully coupled 0D hydrology model.

Q_9 *How significant are the details of the basal hydrology model on surge characteristics in PISM (Sec. S8.2)?*

PISM surge characteristics are compared for local and mass-conserving horizontal transport hydrology models.

Q_{10} *What are the differences (if any) in surge characteristics between local basal hydrology and a basal temperature ramp as the primary smoothing mechanism at the warm/cold-based transition zone (Sec. S8.3)?*

While both sub-glacial hydrology and a basal temperature ramp provide a means for a smooth increase in sliding velocities, these processes operate in slightly different temperature regimes. The basal temperature ramp enables sub-temperate sliding and the maximum velocities occur once the pressure melting point is reached. In contrast, a local basal hydrology model increases sliding velocities once the basal temperature reaches the pressure melting point (basal melting), and basal ice velocities further ramp up with decreasing effective pressure (ice overburden pressure minus basal water pressure). Note that sub-glacial hydrology is not an alternative for a basal temperature ramp. The ramp is still needed to prevent refreezing even when a description of sub-glacial hydrology is included (Mantelli et al., 2019).

Convergence study

Q_{11} *Do model results converge (decreasing differences when increasing horizontal grid resolution, Sec. 3.4)?*

Incorporating the findings of the above experiments, we study numerical convergence with respect to horizontal grid resolution for surge cycling. By convergence, we mean decreasing differences between simulations when increasing the resolution.

2 Methods

2.1 GSM

2.1.1 GSM model description

The 3D thermo-mechanically coupled glacial systems model (GSM) has developed over many years (e.g., Tarasov and Peltier, 1997; Tarasov et al., 2012; Bahadory and Tarasov, 2018). It includes an energy-conserving finite volume ice and bed thermodynamics solver. The current hybrid shallow shelf/ice physics is based on a slight variant of the ice dynamical core of Pollard and DeConto (2012). As is standard for thermo-mechanically coupled glaciological ice sheet models, the GSM has a default explicit time step coupling between the thermodynamics and ice dynamics but also includes an optional implicit coupling scheme (c.f. Sec. 3.2.2). Ice dynamical time stepping is subject to CFL (Courant–Friedrichs–Lewy) constraint (Courant et al., 1928) with further automated reductions upon ice dynamical solver convergence failure. The source code of the model version used in this manuscript can be found in the supplementary material (Tarasov et al., 2023).

The GSM is run with an idealized down-scaled North American geometry (Fig. 1, modified after the ISMIP-HEINO setup (Calov and Greve, 2006)) and simplified climate representation. The surface temperature forcing in the GSM is given by

$$T_{\mathrm{surf}} = \mathrm{rTsurf} + \mathrm{lapsr} \cdot H + T_{\mathrm{asym}}, \tag{1}$$

where rTsurf and lapsr are input parameters for the domain-wide surface temperature constant and atmospheric lapse rate, respectively (Table 1), H the ice sheet thickness, and T_{asym} the asymmetric (in time) temperature forcing (maximum difference of 10°C, orange line in Fig. S1) calculated according to

$$T_{\mathrm{asym}} = \left| \left(\frac{t}{200\ \mathrm{kyr}} \cdot 3 + 2 \right) - 1 \right| \cdot 5^{\circ}\mathrm{C}, \tag{2}$$

where t is the model time ranging from -200 kyr to 0 kyr (instead of 0 kyr to 200 kyr). The asymmetric temperature forcing enables the analysis of the timing of cycling onset and termination under different physical and numerical conditions (a comparison of ice stream ice volume evolution under constant and asymmetric temperature forcing is shown in Fig. S2 for one parameter vector).

The surface mass balance forcing is then determined by

$$M_{\mathrm{tot}} = M_{\mathrm{acc}} - M_{\mathrm{melt}}, \tag{3}$$

where M_{acc} and M_{melt} are the surface accumulation and melt, respectively. The surface accumulation is defined by

$$M_{\mathrm{acc}} = \mathrm{precRef} \cdot \exp\left(\mathrm{hpre} \cdot T_{\mathrm{surf}}\right), \tag{4}$$

where precRef and hpre are the precipitation coefficient input parameters. Surface melt is calculated according to a Positive Degree Day (PDD) approach:

$$M_{\mathrm{melt}} = \mathrm{rPDDmelt} \cdot \max\left(0.0, \mathrm{POSdays} \cdot \left(T_{\mathrm{surf}} + 10.0^{\circ}\mathrm{C}\right)\right), \tag{5}$$

where rPDDmelt is the input parameter for melt per PDD and the PDD constant POSdays is set to $100\ \mathrm{days\ yr^{-1}}$. Note that we set $T_{\mathrm{surf}} = 0.1^{\circ}\mathrm{C}$ and $M_{\mathrm{tot}} = -100\ \mathrm{m\ yr^{-1}}$ for ocean grid cells, and $T_{\mathrm{surf}} = 0.1^{\circ}\mathrm{C}$ and $M_{\mathrm{tot}} = -200\ \mathrm{m\ yr^{-1}}$ at the boundaries of the model domain.

The GSM is initialized from ice-free conditions. The coarsest horizontal grid resolution is $25\mathrm{x}25\ \mathrm{km}$ and is progressively refined (halved) to $3.125\mathrm{x}3.125\ \mathrm{km}$. This gives a total of 4 different horizontal grid resolutions. The maximum time step size is $1\ \mathrm{yr}$ (automatically decreased as needed to meet CFL constraint or when convergence fails).

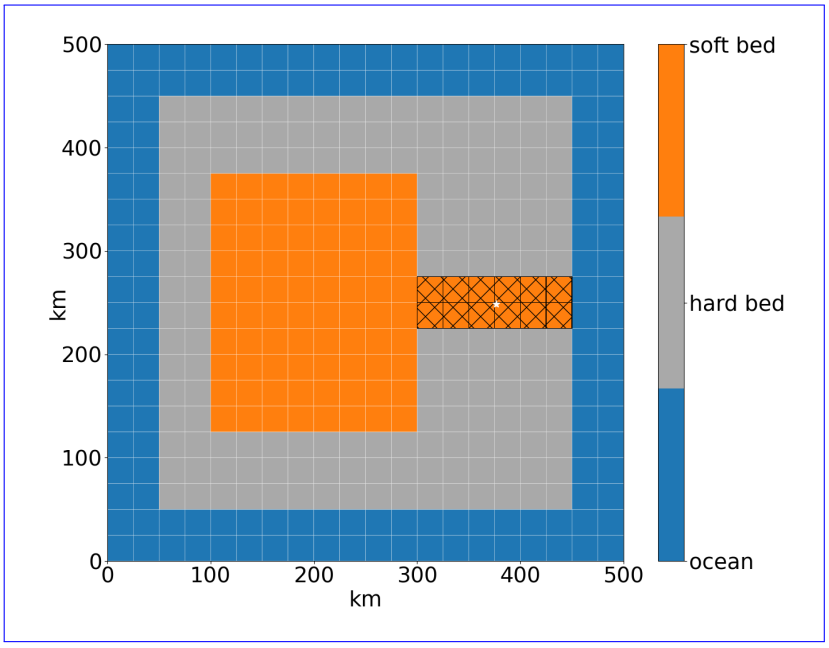


Figure 1. Modified ISMIP-HEINO geometry (Calov and Greve, 2006). The model domain is reduced to $500\mathrm{x}500\ \mathrm{km}$ to enable horizontal grid resolutions up to $3.125\ \mathrm{km}$. The shown grid resolution is $25\mathrm{x}25\ \mathrm{km}$. The basal topography is flat and the hatched area marks the soft-bedded pseudo-Hudson Strait. The white star indicates the location of the grid cell shown in Fig. 8+S21.

While Mantelli et al. (2019) conclude that Stokes mechanics are needed to arrive at a mathematically well-posed model, running numerical experiments with a thermo-mechanically coupled Stokes model is currently unfeasible over glacial cycle time scales. Previous ice stream surge modeling studies are often based on zeroth-order, thin-film approximations of the Stokes

problem, like the shallow-ice Approximation (SIA, e.g., 8 out of 9 models in the ISMIP HEINO experiments (Calov et al., 2010)). While resolving vertical shear, which is the dominant mode of motion in slow flowing regions, SIA-based models neglect longitudinal stress gradients and horizontal shear, which are known to be important for fast ice streams (Hindmarsh, 2009) and are instead captured by the zeroth-order shallow-shelf approximation (SSA).

To partially offset the limitations of the zeroth-order approximations, the GSM uses hybrid SIA/SSA ice dynamics (Pollard and DeConto, 2007, 2012). The hybrid SIA/SSA ice dynamics are activated for grid cells with a SIA velocity exceeding $30\,\mathrm{m\,yr^{-1}}$. Changing these activation velocities ($20\,\mathrm{m\,yr^{-1}}$ and $40\,\mathrm{m\,yr^{-1}}$) has no significant effect on the surge characteristics (Table S1). Activating the SSA everywhere leads to more, shorter, and weaker surges because no threshold velocity needs to be overcome to initiate basal sliding (Sec. S1.2). Note that we set an upper limit of $40\,\mathrm{km\,yr^{-1}}$ for the SSA velocity to ensure that sliding velocities stay within a physically reasonable range.

We configure the GSM with a 1 km deep (17 non-linearly-spaced levels) bed thermal model. A basal temperature ramp is used to ensure a smooth transition between cold-based regions of no sliding and temperate sliding, account for observational evidence of sub-temperate sliding, and more accurately represent the sub-grid warm-based ice fraction in a grid cell and therefore more accurately represent sliding onset for coarse grid resolutions (Q_5 in Sec. 1.3). However, the shape of such a basal temperature ramp is not well constrained. In the GSM, the basal temperature ramp is incorporated into a Weertman-type power law

$$\boldsymbol{u}_\mathbf{b} = \boldsymbol{C}_\mathbf{b} \left|\boldsymbol{\tau}_\mathbf{b}\right|^{n_\mathrm{b}-1} \boldsymbol{\tau}_\mathbf{b} \tag{6a}$$

(6b)

~~as a dependence of the basal sliding coefficient $\boldsymbol{C}_\mathbf{b}$ on the estimated warm-based fraction of a grid cell (indirectly accounting for sub-temperate sliding) F_warm (Eq. (8))~~

(6c)

as a dependence of the basal sliding coefficient $\boldsymbol{C}_\mathbf{b}$ on the estimated warm-based fraction of a grid cell (indirectly accounting for sub-temperate sliding) F_warm (Eq. (8))

$$\boldsymbol{C}_\mathbf{b} = \left(1 - F_\mathrm{warm}\right) C_\mathrm{froz} + F_\mathrm{warm} \boldsymbol{C}, \tag{6d}$$

where $\boldsymbol{u}_\mathbf{b}$ is the basal sliding velocity, $\boldsymbol{\tau}_\mathbf{b}$ the basal stress, n_b the bed power strength (Table 1), and $\boldsymbol{C}$ the fully warm-based sliding coefficient (depends on the bed properties, see also Fig. S4). ~~C_froz~~ C_froz is the fully cold-based sliding coefficient for numerical regularization:

$$C_\mathrm{froz} = 2 \cdot 10^{-3}\,\mathrm{m\,yr^{-1}} \left(5 \cdot 10^{-6}\,\mathrm{Pa^{-1}}\right)^{n_\mathrm{b}}. \tag{7}$$

~~F_warm~~ F_warm is calculated according to

$$F_\mathrm{warm} = \max\left[0, \min\left(1, \frac{T_\mathrm{bp,I} + T_\mathrm{ramp}}{T_\mathrm{ramp}} \frac{T_\mathrm{bp,I} + T_\mathrm{ramp}}{T_\mathrm{ramp}}\right)\right]^{T_\mathrm{exp}}, \tag{8}$$

where ~~$T_{\mathrm{bp,I}}$~~ $T_{\mathrm{bp,I}}$ is the grid cell interface basal temperature relative to the pressure melting point, negative ~~T_{ramp}~~ T_{ramp} the temperature below which the entire grid cell is cold-based, and ~~T_{exp}~~ T_{exp} the exponent used for the ramp. The values used in previous GSM modeling studies (~~$T_{\mathrm{ramp}} = 1.0\ ^{\circ}\mathrm{C}$ and $T_{\mathrm{exp}} = 28$~~ $T_{\mathrm{ramp}} = 1.0\ ^{\circ}\mathrm{C}$ and $T_{\mathrm{exp}} = 28$ (e.g., Bahadory and Tarasov, 2018)) were based on horizontal basal temperature gradients around the basal sliding activation zone with consideration of the sub-grid warm-based connectivity between grid cell interfaces (as basal sliding requires a connected sub-grid warm-based path). Different values for ~~T_{ramp} and T_{exp}~~ T_{ramp} and T_{exp} are explored within this paper. ~~T_{ramp}~~ T_{ramp} can be chosen as either constant or depending on the horizontal grid resolution (res, equal extent in x- and y-direction):

$$T_{\mathrm{ramp}} = P_{\mathrm{T_{ramp}}} \cdot \frac{\mathrm{res}}{50\ \mathrm{km}}\ ^{\circ}\mathrm{C} \tag{9}$$

This choice of resolution dependence leads to a sharper temperature ramp for finer horizontal grid resolutions. The parameter ~~$P_{\mathrm{T_{ramp}}}$~~ $P_{\mathrm{T_{ramp}}}$ is used to conduct experiments with different temperature ramps at the same horizontal grid resolution (Sec. 3.3.3). The temperature ramps for all 4 horizontal grid resolutions and ~~$P_{\mathrm{T_{ramp}}} = 1$~~ $P_{\mathrm{T_{ramp}}} = 1$ (default value) are shown in Fig. 2. For comparison, a temperature ramp similar to the one suggested by Fowler (1986) and later Mantelli et al. (2019)

$$F_{\mathrm{warm}} = \exp\left(\frac{T_{\mathrm{bp,I}}}{\delta}\frac{T_{\mathrm{bp,I}}}{\delta}\right) \quad \text{for} \quad T_{\mathrm{bp,I}} \leq 0 \tag{10}$$

is shown for $\delta = 0.01$, where δ is a parameter controlling the width of the transition zone. Based on experiments conducted by Barnes et al. (1971), Mantelli et al. (2019) expect δ to be small.

2.1.2 GSM ensemble input parameter vectors

Each GSM experiment is run with an ensemble based on 5 input parameter vectors. The current idealized setup encompasses a maximum of 8 input parameters (Table 1) per parameter vector. The 5 parameter vectors used in this study are hand-picked from an exploratory ensemble (Fig. S3). The criteria for these 5 parameter vectors was the highest subset variance in surge characteristics and soft bed sliding law exponent. Note that the soft and hard bed sliding law exponents in this study are equal (n_{b} in Table 1). Due to the significantly increased model run time, sliding law exponents larger than 3 are not considered here. To isolate interactions, the GSM reference setup used in this paper does not incorporate basal hydrology and glacial isostatic adjustment (GIA). Processes associated with basal hydrology, such as lubrication of the bed and decoupling of the ice sheet from the bed, are likely to have a major effect on surge patterns. To determine the impact of these effects, we run the GSM with local basal hydrology enabled (Eq. (19) to (21), Sec. 3.3.5) and examine resolution scaling (Sec. S9.2). However, experiments done with and without basal hydrology lead to qualitatively similar results (e.g., same conclusions from upscaling experiments in Sec. 3.3.3). We therefore omit sub-glacial hydrology coupling for the main analysis.

2.1.3 GSM model setups

The reference setup (Table 2) has a 3.125 km horizontal grid resolution and 1 year maximum time step size. The bed topography is flat (at sea level) and an asymmetric temperature forcing is used (Fig. S1). For the sake of generality, we chose a flat

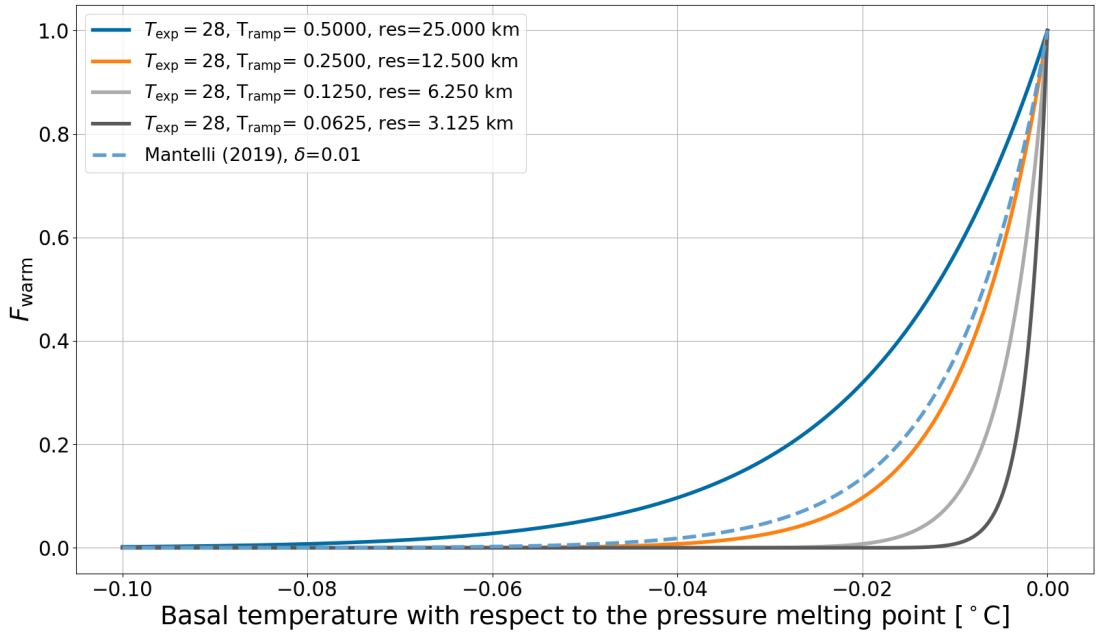


Figure 2. Temperature ramps for different values of ~~T_{ramp}~~ T_{ramp} which depend on the horizontal grid resolution. A temperature ramp similar to the one suggested by Mantelli et al. (2019) (Eq. (10)) is shown for $\delta = 0.01$.

topography for the reference setup, while the effect of a basal trough is investigated at a later stage (Sec. 3.3.4). Branching off this reference set-up, we carry out one-factor-at-a-time sensitivity experiments to isolate numerical and process impacts. These experiments, in turn, examine the response to: 3 numerical aspects related to the MNEEs, 4 model aspects affecting the thermal onset of basal sliding, a change in sediment cover, a non-flat topography, the addition of local basal hydrology, and
285 different horizontal grid resolutions [25 km, 12.5 km, 6.25 km]. The 3 numerical aspects are stricter numerical convergence criteria, the addition of surface temperature noise (± 0.1°C and ± 0.5°C), and an approximate implicit time step coupling between the thermodynamics and ice dynamics. The 4 thermal model aspects are switching to thin (20 m) bed thermal model, different approaches to determining the basal temperature at the grid cell interface, different weights of the adjacent minimum basal temperature for the basal sliding temperature ramp (~~$W_{Tb,min}$~~ $W_{Tb,min}$), and different basal temperature ramps (~~T_{ramp}~~
290 ~~and T_{exp}~~ T_{ramp} and T_{exp}) for thermal activation of basal sliding. See Table 1 for details on parameter ranges.

Category	Parameter	Description	Range	Unit
Ensemble parameter - ISM	C_{rmu}	soft bed sliding coefficient	0.3 - 1	
	C_{fslid}	hard bed sliding coefficient	0 - 3	
	lapsr	atmospheric lapse rate	-5 - -10	$^{\circ}\mathrm{C\ km}^{-1}$
	PDDmelt	melt per Positive Degree Day (PDD)	0.005 - 0.012	$\mathrm{m\ PDD}^{-1}\ (^{\circ}\mathrm{C})^{-1}$
	hpre	precipitation coefficient	0.02 - 0.2	$(^{\circ}\mathrm{C})^{-1}$
	PrecRef	precipitation coefficient	1 - 3	$\mathrm{m\ yr}^{-1}$
	rTsurf	domain wide surface temperature constant	-9 - -15	$^{\circ}\mathrm{C}$
	n_{b}	soft and hard bed sliding law exponent, bed power strength	1 - 3	
Hydrology parameters	~~$hwb,Crit$~~ $h_{\mathrm{wb,Crit}}$	effective bed roughness scale (Eq. (19))	0.01 - 1	m
	rBedDrainRate	constant bed drainage rate	0.001 - 0.01	$\mathrm{m\ yr}^{-1}$
	$N_{\mathrm{eff,Fact}}$	effective pressure factor (Eq. (21))	$2 \cdot 10^4$ - $2 \cdot 10^5$	Pa
Additional parameters	~~P_{Tramp}~~ P_{Tramp}	basal temperature ramp scaling factor (Eq. (9))	0.125 - 16 (**1**)	
	~~T_{ramp}~~ T_{ramp}	basal temperature (with respect to the pressure melting point) at which sub-temperate sliding becomes important (Eq. (8), (9))	0.03125 - 1 (**0.0625**)	$^{\circ}\mathrm{C}$
	~~T_{exp}~~ T_{exp}	basal temperature ramp exponent (Eq. (8))	5 - 56 (**28**)	
	~~$W_{\mathrm{Tb,min}}$~~ $W_{\mathrm{Tb,min}}$	weight of adjacent minimum basal temperature for basal sliding temperature ramp (Eq. (S1))	0.0 - 1.0 (**0.5**)	

Table 1. Model parameters are listed with respect to their purpose/category. Ice Sheet Model - ISM. Hydrology parameters used when running the GSM with local basal hydrology. Additional (non-regular) input parameters that are usually set to a fixed value. The default values of the 3.125 km horizontal grid resolution reference setup are shown in the brackets for the additional parameters.

2.2 PISM

2.2.1 PISM model description

In contrast to the GSM, the Parallel Ice Sheet Model (PISM) is not specifically developed for glacial cycle ensemble modeling. Therefore, the two models use distinct sets of numerical optimizations for computational speed. To minimize the model dependency of our analysis, experiments are also carried out with v2.0.2 of PISM.

Similar to the GSM, PISM is a 3D thermodynamically-coupled ice sheet model and the SSA is used as a 'sliding law' once the sliding velocity exceeds $100\ \mathrm{m\ yr^{-1}}$. For further details on the model itself, refer to Bueler and Brown (2009) and Winkelmann et al. (2011). The details on the default PISM setup, together with the default GSM values, are listed in Table 2. Given the higher computational cost of PISM experiments, the relatively high sensitivity of PISM to the number of parallelized cores for these experiments (Table 6), and run time limitations of the computational cluster, the reference setup is run at $25\ \mathrm{km}$ horizontal grid resolution.

For stability reasons, the PISM adaptive time stepping ratio (used in the explicit scheme for the mass balance equation) was reduced to 0.01 when using small till friction angles (Constantine Khrulev, personal communication).

The default sliding law in PISM is a purely-plastic (Coulomb) model where

$$|\boldsymbol{\tau_b}| \leq \tau_c \quad \text{and} \quad \boldsymbol{\tau_b} = -\tau_c \frac{\boldsymbol{u}}{|\boldsymbol{u}|} \quad \text{if} \quad |\boldsymbol{u}| > 0. \tag{11}$$

Therefore, the basal shear stress $\boldsymbol{\tau_b}$ can never exceed the yield stress τ_c, and basal sliding only occurs when $\boldsymbol{\tau_b}$ reaches τ_c.

2.2.2 PISM ensemble input parameter vectors

The PISM configuration encompasses 6 model input parameters (Table 3). These parameters define the input fields for surface temperature, surface accumulation, and till friction angle. As for the GSM, PISM is initialized from ice-free conditions. Similar to Calov and Greve (2006), the surface temperature at every grid cell is calculated as

$$T_{\mathrm{surf}} = T_{\mathrm{min}} + S_t \cdot d^3, \tag{12}$$

where S_t represents the horizontal surface temperature gradient, d the distance from the domain center (~~x_{center}, y_{center}~~ x_{center}, y_{center}) in km, defined as:

$$d = \sqrt{(x - x_{\mathrm{center}})^2 + (y - y_{\mathrm{center}})^2}\sqrt{(x - x_{\mathrm{center}})^2 + (y - y_{\mathrm{center}})^2} < R, \tag{13}$$

and R denotes the radius and sets an upper limit for d. A comparable equation is used to calculate the surface mass balance (accumulation/ablation) rate input field.

$$B_{\mathrm{surf}} = B_{\mathrm{max}} - S_b \cdot d^5, \tag{14}$$

where S_b is the horizontal surface mass balance gradient. The input field for the till friction angle is defined by simple grid assignment and a somewhat smoothed transition between the soft and hard bed region. Input fields for one parameter vector are shown for surface temperature, surface accumulation, and till friction angle in Fig. S6, S7, and S8, respectively.

Setup component	GSM	PISM
horizontal grid resolution	3.125 km x 3.125 km	25 km x 25 km
number of grid cells	160 x 160	120 x 120
model domain	500 km x 500 km	3000 km x 3000 km
vertical layers	65	60
run time	200 kyr	200 kyr
maximum time step size	1 yr	1 yr
number of cores/processes	1	8
ice dynamics	hybrid SIA/SSA	hybrid SIA/SSA (maximum SIA diffusivity of 1000 $\mathrm{m}^2\ \mathrm{s}^{-1}$)
sliding law	Weertman-type power law (Eq. (6a))	Coulomb friction law (Eq. (11))
bed topography	flat (at sea level)	flat (at sea level)
bed thermal model	1 km deep (17 non-linearly-spaced levels)	1 km deep (20 equally-spaced levels)
basal hydrology	not included	local basal hydrology model based on an undrained plastic bed model (Tulaczyk et al., 2000a)

Table 2. Comparison between the GSM and PISM reference setup.

Category	Parameter	Description	Range	Unit
Ensemble parameters	$soft$	soft bed till friction angle	0.5 - 12.0	°
	$hard$	hard bed till friction angle	15.0 - 30.0	°
	~~B_{max}~~ B_{max}	maximum surface mass balance (accumulation/ablation) rate	50 - 450	$\mathrm{kg\ m^{-2}\ yr^{-1}}$
	S_{b}	horizontal surface mass balance gradient	$(0.15 - 1.00) \cdot 10^{-11}$	$\mathrm{kg\ m^{-2}\ yr^{-1}\ km^{-5}}$
	~~T_{min}~~ T_{min}	minimum surface temperature	220 - 245	K
	S_{t}	horizontal surface temperature gradient	$(0.10 - 1.0) \cdot 10^{-8}$	$\mathrm{K\ km^{-3}}$
Constant parameters	~~x_{center}~~ x_{center}	location of the domain center in x-direction	1500	km
	~~y_{center}~~ y_{center}	location of the domain center in y-direction	1500	km
	R	maximum radius of the domain	1500	km

Table 3. Parameters used to generate the PISM input fields.

The 6 model ensemble parameters (Table 3) were selected via Latin Hypercube sampling. After sieving an ensemble of 100 runs for those that show oscillatory behavior, a 9-member high-variance (with respect to the surge characteristics) subset was

extracted by visual identification (Fig. S10). Each PISM experiment is run with an ensemble based on these 9 input parameter vectors.

2.2.3 PISM bed properties

A PISM ensemble parameter restriction arose as experiments carried out with PISM only show oscillatory behavior for small yield stresses τ_c. This can be achieved by either a small till friction angle Φ or low effective pressure on the till (N_{till}, Eq. (S2)) (Bueler and Van Pelt, 2015):

$$\tau_c = c_0 + \tan(\Phi)\, N_{\text{till}}, \tag{15}$$

where $c_0 = 0$ Pa is the till cohesion (Tulaczyk et al., 2000b). For convenience, we decide to vary only the till friction angle between 0.5 and 1°, for which PISM shows oscillatory behavior, and otherwise use PISM default values (see Sec. S2.3 for details).

The resulting very slippery beds enabled occasional maximum sliding velocities of up to $\sim 600\ \text{km yr}^{-1}$ in the simulations (Fig. S11, Sec. S2.4). For comparison, observed outlet glacier velocities at Jakobshavn Isbræ(Greenland) approach $20\ \text{km yr}^{-1}$ (Joughin et al., 2012, 2014). Similar to the GSM, we, therefore, set an upper limit of $40\ \text{km yr}^{-1}$ for the SSA velocity.

2.2.4 PISM model setups

As for the GSM, we carry out one-factor-at-a-time sensitivity experiments branching off the PISM reference set-up (Table 2) for all 9 parameter vectors. These experiments, in turn, examine the response to: 2 numerical aspects related to the MNEEs, removing the bed thermal model, an abrupt sediment transition zone, a non-flat topography (Fig. S9), a mass-conserving horizontal transport model for basal hydrology (Bueler and Van Pelt, 2015), and different horizontal grid resolutions (50 km,12.5 km). The 2 numerical aspects are different number of cores/processes ($n = 2, 4, 16, 32$), and addition of surface temperature noise ($\pm 0.1^\circ$C and $\pm 0.5^\circ$C).

2.3 Run analysis approach

Surge characteristics

The quantities being analyzed are: the number of surges, the surge duration, the ice volume change during a surge, and the period between surges (Fig. 3). The surge time is defined as the time of minimum (pseudo-Hudson Strait) ice volume, and the duration of a surge includes the surge itself as well as the time it takes the ice sheet to recover approximately half the ice volume lost during the surge (Sec. S3). The calculated ice volume change is the difference between the pre-surge and minimum (pseudo-Hudson Strait) ice volume in that particular surge (Sec. S3). The period between surges is the time span between two subsequent occurrences of minimum (pseudo-Hudson Strait) ice volume (not defined for the very last event). The spin-up interval (first 20 kyr of every run) is not incorporated in the analysis, and only surges with a (pseudo-Hudson Strait) ice volume change of more than $500\ \text{km}^3$ and $4 \cdot 10^4\ \text{km}^3$ are considered in the GSM and PISM analyses, respecitvely ($\sim 5\ \%$

of mean ice volume across all runs). Note that this is a very conservative spin-up interval. For example, most GSM runs reach their mean pseudo-Hudson Strait ice volume after ~ 5 kyr (e.g., Fig. 11).

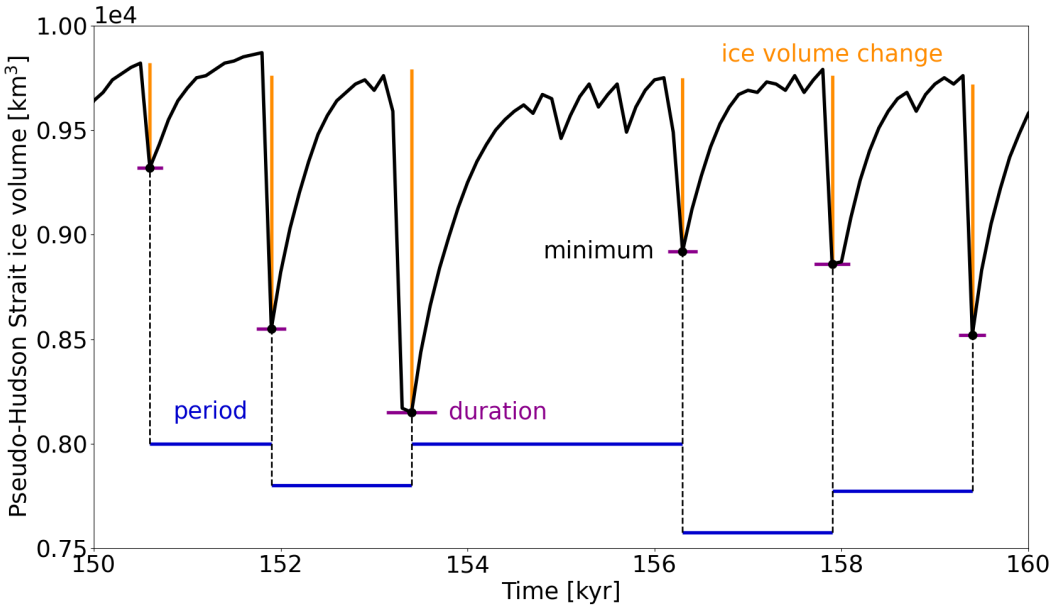


Figure 3. Pseudo-Hudson Strait ice volume of a GSM model run with visual illustration of the surge characteristics used to compare different model setups. The horizontal grid resolution is 3.125 km.

In addition to the surge characteristics, the Root Mean Square Error (RMSE) and mean bias are calculated as a percentage deviation from the reference (pseudo-Hudson Strait) ice volume time series for all setups (each parameter individually) and then averaged over the 5 parameter vectors (Eq. (S3) and (S4)). The full run time is considered (no spin-up interval).

Percentage differences

We compare different model setups by calculating the percentage difference between the reference setup and all other setups for every parameter vector individually and then average this difference over all parameter vectors. Crashed runs are not considered and runs with less than 2 surges require special treatment (see Sec. S5 for further details on the analysis).

Surge area

In the GSM, the whole pseudo-Hudson Strait (Fig. 1) is ice-covered and at maximum ice volume at the beginning of a surge. Surges in the GSM, therefore, consistently appear as ice volume minima, which allows us to directly use the pseudo-Hudson Strait ice volume for the GSM results.

For PISM, a large fraction of the pseudo-Hudson Strait area is only ice-covered when a surge occurs (e.g., Fig. 5), leading to an inconsistency in the surge detection. This issue is addressed by including the ice volume over the eastern half of the pseudo-Hudson Bay, the area most affected by the surge drained through the pseudo-Hudson Strait. See Sec. S2.5 for further details and a comparison between the two approaches.

Minimum numerical error estimates

We compute the new '**M**inimum **N**umerical **E**rror **E**stimates' (MNEEs) ~~metric~~ threshold by examining the model response to changes in the model configuration that are not part of the physical system. The MNEEs are defined as the percentage differences in surge characteristics when applying a stricter (than default) numerical convergence in the GSM and ~~adjusting the matrix solver~~ changing the number of processor cores used in PISM~~(changing the number~~. The differences between PISM runs with different numbers of processor cores ~~used). They~~ can be caused by, for example, a different order of floating point arithmetic operations and the processor-number-dependent preconditioner used in PISM (PISM 2.0.6 documentation, 2023). The MNEEs are then used as a threshold to determine if model sensitivities to changes in the model configuration that affect the physical system (e.g., the inclusion of a bed thermal model or sliding depence on effective pressure from basal hydrology) are above the level of background noise induced by iterative numerical solvers in the model. We refrain from drawing conclusions about the effects of a change in model configuration with physical relevance when the model sensitivities in question are smaller than the MNEEs. In these cases, the actual physical response of the model might be hidden within the numerics.

While the MNEEs are useful to our purpose, we wish to emphasize that they can not replace proper model verification and validation and are missing uncertainties due to, e.g., different approximations of the Stokes equations and other physical processes not included in the models. Nonetheless, they provide a minimum estimate of the numerical model error, which is still a significant improvement over ignoring this issue entirely.

3 Results

3.1 Key surge characteristics of the reference setup

Before analyzing ensemble characteristics, it is crucial to understand how surges initiate, propagate and terminate. Surges in the GSM originate at the pseudo-Hudson Strait mouth ($x = 450$ km, $y = 225$ to 275 km) and propagate towards the center of the pseudo-Hudson Bay ($x = 200$ km, $y = 250$ km, Fig. 1 and 4). The surging onset is a complex interplay between heating at the ice sheet bed, basal temperature, and ice sheet velocity. The beginning of a surge is shown in ~~an online video (~~video 01 ~~of Hank (2023))~~ (Hank, 2023) and Fig. 4. Just before the start of the surge, the entire South-North extent of pseudo-Hudson Strait grid cells close to the ocean is warm-based. At $t = 6.69$ kyr, the SIA velocities exceed 30 m yr^{-1} and the SSA is activated (Sec. 2.1.1). The longitudinal stress gradient and horizontal shear terms provide additional heating~~(heating due to shelfy stream dynamics in video 01 of Hank (2023))~~. This leads to several small ice streams with relatively strong heating due to basal sliding ($\sim 10^7$ J m^{-2} yr^{-1}) at $t = 6.70$ kyr in the video. This is an order of magnitude larger than heat production from deformation

work. The additional heat fosters higher ice velocities, leading to even more heating, the extension of the warm-based area to the West, and therefore the upstream propagation of the small ice streams ($t = 6.71$ kyr). The narrow ice streams draw in warm-based ice from the surrounding grid cells, increasing the velocities and heat production in the area between the ice streams. This leads to a merger of the ice streams with now high velocities occurring over the full South-North extend of the pseudo-Hudson Strait ($t = 6.72$ kyr). The warm-based area rapidly extends towards the West due to the strong heating and high ice velocities, causing a pseudo-Hudson Strait surge.

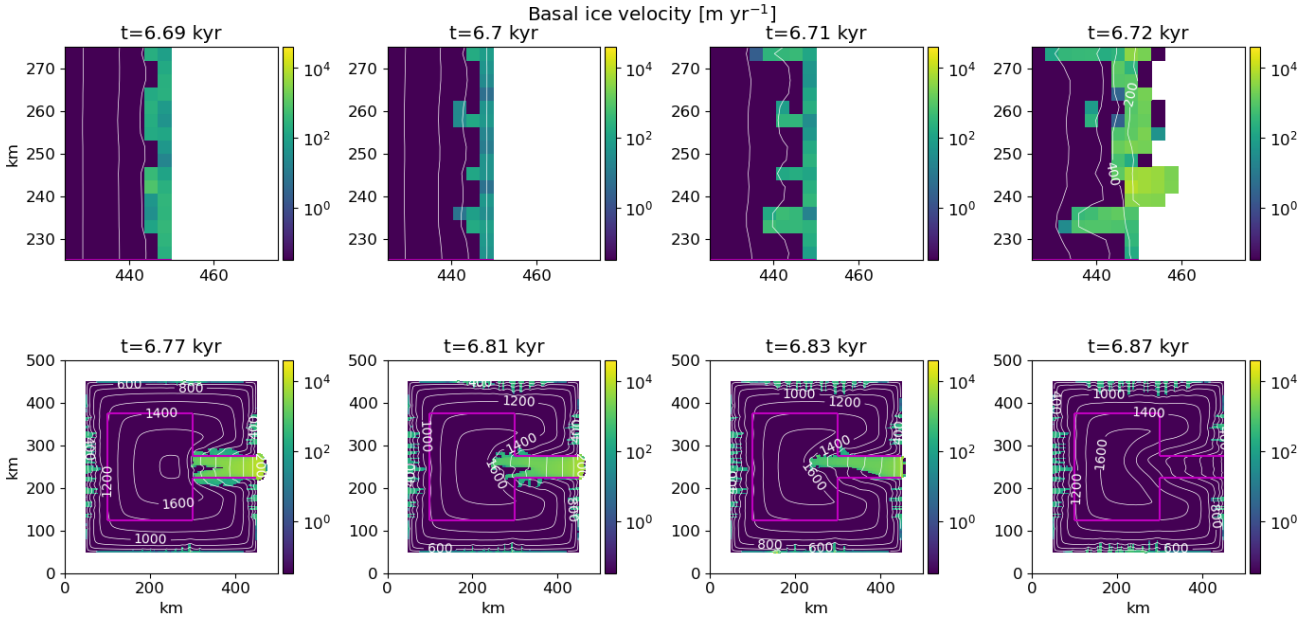


Figure 4. Basal ice velocity for parameter vector 1 at different time steps using the GSM. The horizontal grid resolution is 3.125 km and the maximum model time step is 1 yr. The contour lines show the ice sheet surface elevation in m. The magenta line outlines the soft-bedded pseudo-Hudson Bay and Hudson Strait. Note that the top and bottom rows show different areas of the domain, with the top zooming in on the surge onset area.

The surge propagates nearly symmetrically until the pseudo-Hudson Bay area is reached ($t = 6.77$ kyr in Fig. 4 and video 02 of Hank (2023)). After this point, the northern branch of the ice stream propagates more rapidly and extends further to the West than the southern branch. While the smaller southern branch starts to shrink at $t = 6.81$ kyr, the northern part propagates until $t = 6.83$ kyr. At this time, the southern branch has vanished almost completely due to a thinner ice sheet (than at the start of the surge) and the advection of cold ice into the surge area. After $t = 6.83$ kyr, the available heating is no longer sufficient to keep the ice sheet bed at the pressure melting point, and the northern part collapses as well. The surge ends after 150 yr (at $t = 6.87$ kyr).

Since the GSM setup and climate forcing are symmetric about the horizontal axis in the middle of the pseudo-Hudson Strait ($y = 250$ km in Fig. 1), we interpret the induced asymmetry as '*spontaneous symmetry breaking*' similar to the results described in Sayag and Tziperman (2011). We define the asymmetry as positive when the surge is stronger Northward (Fig. 4 and video 02 of Hank (2023)) or shifted Northward. The asymmetry sign varies across the first surges (i.e., the surge least biased by previous asymmetries) of the 5 reference runs, ruling out any persistent numerical bias.

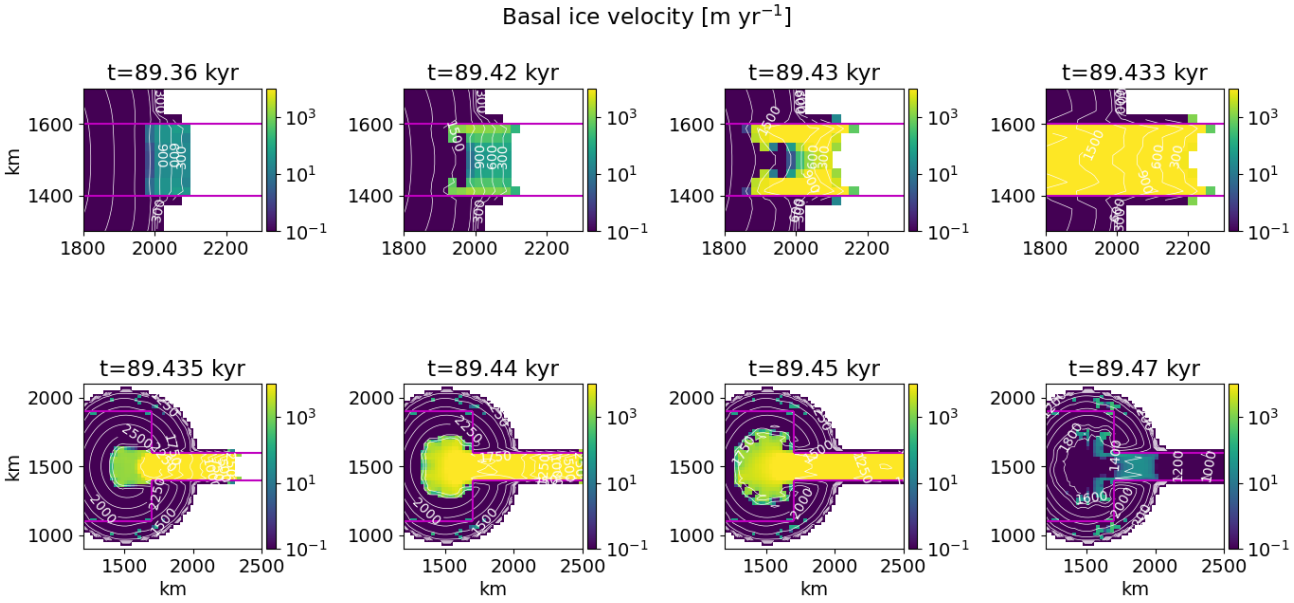


Figure 5. Basal ice velocity for parameter vector 8 at different time steps using PISM. The horizontal grid resolution is 25 km and the maximum model time step size is 1 yr. Otherwise as in Fig. 4.

Surges in PISM originate at the ice sheet margin in the soft-bedded pseudo-Hudson Strait (exact position varies between runs) and propagate towards the center of the pseudo-Hudson Bay ($x = 1300$ km, $y = 1500$ km, Fig. S8 and 5). The ice near the margin is already flowing downstream before the start of the surge ($t = 89.36$ kyr). However, the basal temperature is below the pressure melting point, and the ice velocities are low (< 100 m yr^{-1}). As the ice sheet upstream of the margin thickens, the warm-based area extends further downstream, particularly along the 100 % soft-bedded contour line (magenta line in Fig. 5).

Once the warm-based area connects with the margin ($t = 89.42$ kyr), the ice velocities increase beyond 100 m yr^{-1}, activating the SSA (Sec. 2.2.1). Similar to the surges in the GSM, the sliding velocities then increase rapidly, quickly extending the warm-based area ($t = 89.43$ kyr and $t = 89.433$ kyr). The surge propagates upstream into the pseudo-Hudson Bay and the ice is transported along the pseudo-Hudson Strait into regions with increasingly negative surface mass balance rates ($t = 89.435$ kyr to $t = 89.45$ kyr, Fig. S7). The ice sheet thins, the basal temperature at the margin falls below the pressure melting point, blocking parts of the upstream ice stream, and the surge ceases at $t = 89.47$ kyr (~ 100 yr surge duration). The ice volume in

the surge-affected area continues to decrease for, on average, another 2.5 kyr due to the large amounts of ice in the negative surface mass balance regions. In contrast to the GSM, PISM results remain symmetrical about $y = 1500$ km throughout the surge.

Due to the differences in model setup, physics, and numerics (Table 2), the GSM and PISM reference setup yield different surge characteristics (Table 4). While resembling the inferred ice-rafted debris (IRD) interval duration as closely as possible is not a goal of this study, the modeled values are in agreement with the literature (200 to 2280 yr (Hemming, 2004)). The mean modeled GSM period is shorter than the observed period of, on average, 7 kyr (K.M. Cuffey and W.S.B. Paterson., 2010). However, exploratory GSM runs with a dimensionally accurate (not downscaled) model domain (but otherwise identical experimental setup) yielded periods within the range of geological inferences. The mean modeled PISM period is within limits set by the literature. The mean (pseudo-Hudson Strait) ice volume change in the GSM corresponds to 15 % of a 1.5 km thick ice sheet covering the downscaled pseudo-Hudson Strait area (150x50 km). In PISM, the mean ice volume change is 7.1 % of the mean (across reference setup runs) maximum ice volume in the eastern half of the pseudo-Hudson Bay and pseudo-Hudson Strait.

Metric	GSM reference setup	PISM reference setup
number of surges	180 ± 100	35 ± 25
mean period	1.1 ± 0.5 kyr	10 ± 10 kyr
mean duration	0.3 ± 0.1 kyr	3 ± 2 kyr
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	$1.1 \pm 0.3 \cdot 10^5$ km^3

Table 4. Surge characteristics of the GSM (~~$T_{\text{ramp}} = 0.0625°$~~$T_{\text{ramp}} = 0.0625°$C, ~~$T_{\text{exp}} = 28$~~ $T_{\text{exp}} = 28$ (black line in Fig. 2), ~~$W_{\text{Tb,min}} = 0.5$~~$W_{\text{Tb,min}} = 0.5$, TpmTrans for the interface calculation, sharp transition between hard and soft bed) and PISM reference setup (Table 2). No runs crashed and all runs had more than 1 surge event. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

3.2 Minimum numerical error estimates

Differences in surge characteristics (compared to the reference setup) are considered significant when they exceed the MNEEs given in Table 5 and 6 for the GSM and PISM, respectively. However, this does not necessarily mean that smaller changes have no physical relevance but rather that their interpretation is difficult (if not impossible) because the physical response is hidden within the numerical sensitivities. Likely sources of the MNEEs are the iterative SSA solutions and floating point accuracy.

To determine a minimum significant threshold in the GSM, we re-run a set of GSM runs with 3.125 km horizontal grid resolution, imposing a stricter numerical convergence (decreasing final iteration thresholds). In a second experiment, we additionally increase the maximum iterations from 2 to 3 for the outer Picard loop ~~(ice dynamics)~~ solving for the ice thickness and from 2 to 4 when solving the non-linear elliptic SSA equation for horizontal ice velocities.

The largest differences between simulations occur for the mean period (7 %, Table 5) when using stricter convergence thresholds (no change in the maximum number of iterations). The standard deviations are on the same order of magnitude as the values themselves, indicating different responses across the 5 parameter vectors. Determining the MNEEs at 12.5 km instead of 3.125 km horizontal grid resolution yields similar results, except for the mean pseudo-Hudson Strait ice volume change (21 %, Table S2).

Metric	reference setup	stricter numerical convergence [% difference]	stricter numerical convergence with increased maximum iterations [% difference]
number of surges	180 ± 100	$\mathbf{-4.1 \pm 4.9}$	-0.9 ± 3.6
mean period	1.1 ± 0.5 kyr	$\mathbf{7.0 \pm 10.6}$	4.7 ± 10.6
mean duration	0.3 ± 0.1 kyr	2.5 ± 3.2	$\mathbf{3.9 \pm 4.8}$
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	-1.1 ± 3.1	$\mathbf{4.6 \pm 4.6}$

Table 5. Percentage differences of surge characteristics between GSM runs with regular and stricter numerical convergence and increased maximum iterations for the ice dynamics loops at 3.125 km horizontal grid resolution. The values represent the average of 5 parameter vectors. No runs crashed and all runs had more than 1 surge event. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. The bold numbers mark the largest MNEE for each surge characteristic.

MNEEs in PISM are determined by comparing runs with different numbers of cores. Although most parameter vectors show similar results at the beginning of the runs, minor differences can slowly accumulate and lead to significant discrepancies in surge activity by the end of the run (Fig. S18). ~~These minor differences can be caused by, for example, a different order of floating point arithmetic operations and the processor-number-dependent preconditioner used in PISM (PISM 2.0.6 documentation, 2023) .~~The largest differences occur for the number of surges (16 %) and mean ice volume change (16 %) for nCores= 32, but the standard deviations are large due to a more than ~ 200 % increase in both surge characteristics for parameter vector 6.

The differences in surge characteristics between different numbers of cores can be minimized (but not removed entirely) by decreasing the relative Picard tolerance in the calculation of the vertically-averaged effective viscosity (10^{-4} to 10^{-7}) and the relative tolerance for the Krylov linear solver used at each Picard iteration (10^{-7} to 10^{-12}, Table S5 and Fig. S19). However, this leads to an unreasonable increase in model run time (~ 300 %) that is not feasible for an ensemble-based approach (more than 50 % of all runs did not finish within the time limit of the computational cluster). Intermediate decreases in the relative tolerances still lead to significant differences in surge characteristics while increasing the model run time and are, therefore, not used in the PISM reference setup. Considering that small differences prevail for all tested relative tolerances, comparing model configurations with different numbers of cores for, e.g., finer horizontal grid resolution experiments is not straightforward.

Setup	number of surges	mean period	mean duration	mean ice volume change	nS1
25 km reference setup	35 ± 25	10 ± 10 kyr	3 ± 2 kyr	$1.1 \pm 0.3 \cdot 10^5$ km^3	0
nCores= 2	-7.1 ± 19.5	6.8 ± 36.2	-0.4 ± 9.5	1.5 ± 10.3	0
nCores= 4	-8.2 ± 22.9	-3.8 ± 6.6	2.8 ± 18.3	0.6 ± 4.8	1
nCores= 16	-10.9 ± 26.0	$\mathbf{-8.2 \pm 14.7}$	7.6 ± 21.2	-0.7 ± 13.3	1
nCores= 32	$\mathbf{16.0 \pm 56.2}$	6.9 ± 48.5	$\mathbf{-8.0 \pm 17.4}$	$\mathbf{16.3 \pm 35.1}$	0

Table 6. Percentage differences of surge characteristics (except first row) between the PISM reference setup and setups with different numbers of cores at 25 km horizontal grid resolution. The values represent the average of 9 parameter vectors. No runs crashed and all runs showed at least 1 event. Runs with just one surge (nS1) are ignored when calculating the change in mean period. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. The bold numbers mark the largest MNEE for each surge characteristic.

3.2.1 Adding surface temperature noise

Low levels of surface temperature noise have previously been shown to cause chaotic behavior in the mean periods of oscillations (Souček and Martinec, 2011). Adding low levels of uniformly distributed surface temperature noise (maximum amplitude of ± 0.1°C and ± 0.5°C) to the climate forcing (updated every 100 yr) does not significantly affect the surge characteristics for the GSM (Table S3). For example, the effect of adding ± 0.5°C surface temperature noise on the mean period is only 4 % (compared to the ~ 20 % for ± 0.01°C reported by Souček and Martinec (2011)). Adding the same levels of uniformly distributed surface temperature noise to PISM increases the mean duration by 12 % (for ± 0.1°C), but has no significant effect on the other surge characteristics (Table S6).

3.2.2 Implicit thermodynamics/ice dynamics coupling

In contrast to the commonly used explicit time step coupling between the thermodynamics and ice dynamics in glaciological ice sheet models, we test the impact of approximate implicit time step coupling via an iteration between the two calculations for each time step. The implicit coupling decreases the mean duration and pseudo-Hudson Strait ice volume change (-13 % and -25 %, respectively). The number of surges and mean period show no significant change (Table S4). While the changes in mean duration and pseudo-Hudson Strait ice volume change are larger than the MNEEs, they do not justify an increase in run time of ~ 265 % and the implicit coupling is therefore omitted for the GSM reference setup.

3.3 Sensitivity experiments

Here we discuss differences in surge characteristics due to changes in the model setup. An overview of the results can be found in Fig. 6 and 7 for the GSM and PISM, respectively. The exact values of the percentage differences and information on crashed runs or runs without oscillations are provided in the supplement. We first examine the model aspects affecting the thermal activation of basal sliding (Sec. 3.3.1 to 3.3.3), followed by the analysis of a smooth sediment transition zone,

non-flat topography, and local basal hydrology (Sec 3.3.4 and 3.3.5). Experiments without significant differences in the surge characteristics are only briefly mentioned here (Sec. 3.3.6). A more in-depth discussion of these latter experiments is available in the supplement.

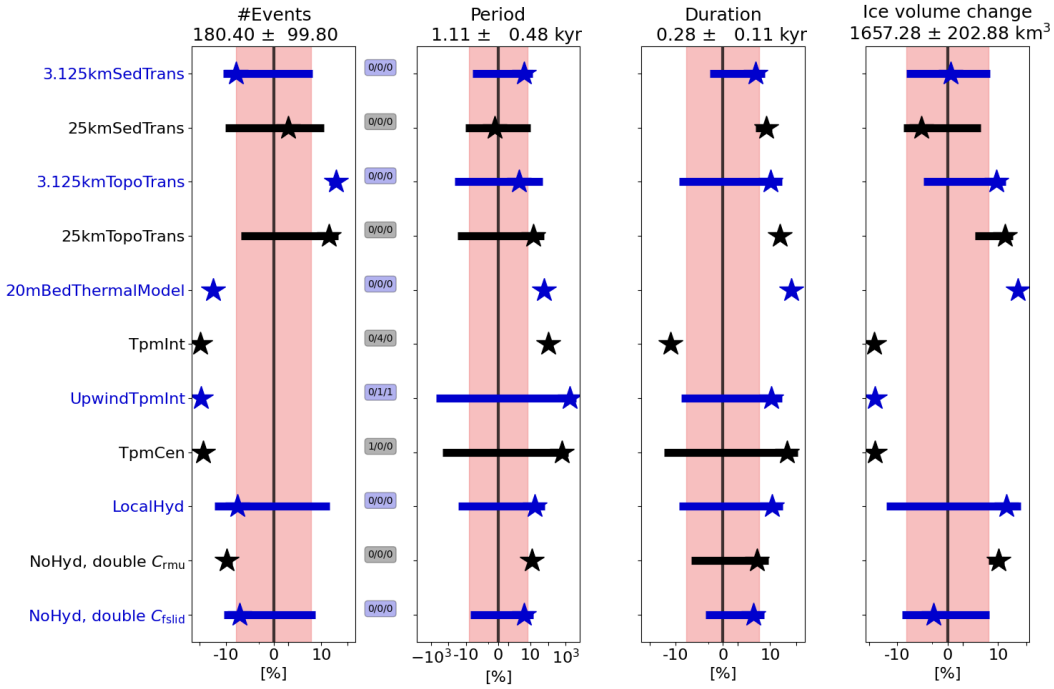


Figure 6. Percentage differences in surge characteristics compared to the GSM reference setup for model setups discussed in Sec. 3.3 (average of the 5 parameter vectors). The horizontal grid resolution is 3.125 km. The different colors were added for visual alignment of the individual model setups, the stars are the ensemble mean percentage differences, and the horizontal bars represent the ensemble standard deviations. The shaded pink regions mark the MNEEs (Table 5) and the black numbers in the title of each subplot represent the mean values of the reference setup. The 3 small numbers between the first two columns represent the number of crashed runs (nC), the number of runs without a surge (nS0), and the number of runs with only one surge (nS1), respectively. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. The x-axis is logarithmic. Further details of each individual experiments are provided in the subsequent sections and the supplement. The model setups, from top to bottom, are: 3.125 km wide sediment transition zone (instead of an abrupt transition in the reference setup), 25 km wide sediment transition zone, 3.125 km wide sediment transition zone with pseudo-Hudson Bay/Hudson Strait topography (instead of a flat topography in the reference setup), 25 km sediment transition zone with pseudo-Hudson Bay/Hudson Strait topography, 20 m deep (1 layer) bed thermal model (instead of a 1 km deep bed thermal model (17 non-linearly-spaced layers) in the reference setup), 3 different approaches to calculate basal grid cell interface temperature (TpmInt, upwind TpmInt, TpmCen), local hydrology (instead of no hydrology), and doubling the values of the soft and hard bed sliding coefficients (as an attempt to represent basal hydrology without actually adding it).

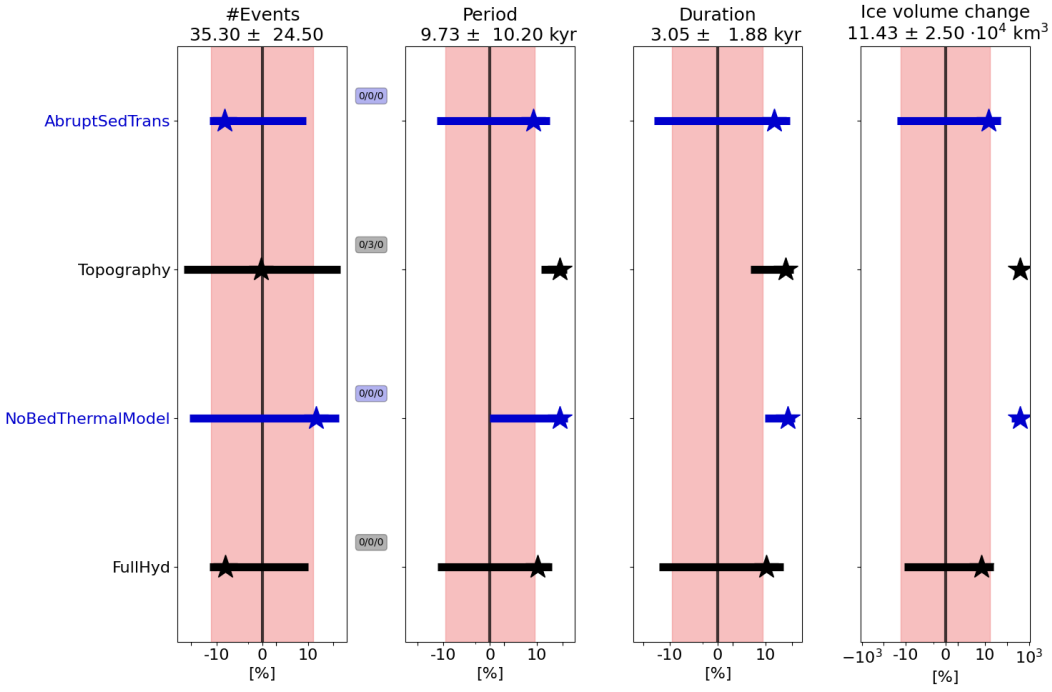


Figure 7. Percentage differences in surge characteristics compared to the PISM reference setup for model setups discussed in Sec. 3.3 (average of the 9 parameter vectors). The horizontal grid resolution is 25 km. Otherwise same as Fig. 6. The model setups, from top to bottom, are: abrupt sediment transition (instead of the transition shown in, e.g., Fig. S8), pseudo-Hudson Bay/Hudson Strait topography (instead of a flat topography in the reference setup, Fig. S9), no bed thermal model (instead of a 1 km deep bed thermal model (20 equally-spaced layers) in the reference setup), and a mass-conserving horizontal transport model for basal hydrology (instead of a local hydrology).

3.3.1 Bed thermal model

First, we examine the effects of a 1 km deep bed thermal model on the basal temperature and the surge characteristics in the GSM as well as PISM. Both models show significant differences when limiting the bed thermal model to one layer (GSM) or removing it entirely (PISM).

Advection of cold ice near the end of a surge rapidly decreases the basal ice temperature and, therefore, increases the temperature gradient between the basal ice and the bed. In GSM runs with the 1 km deep (17 non-linearly-spaced levels) bed thermal model (reference setup), this stronger gradient increases the heat flux from the bed into the ice and dampens the actual change in basal ice temperature. Similarly, a rapid increase in basal ice temperature due to higher basal ice velocities at the beginning of a surge reverses the existing temperature gradient at the base of the ice sheet, leading to a heat flux from the ice into the bed. Consequently, less heat is available to warm the surrounding cold-based ice, counteracting the surge propagation (Fig. 8).

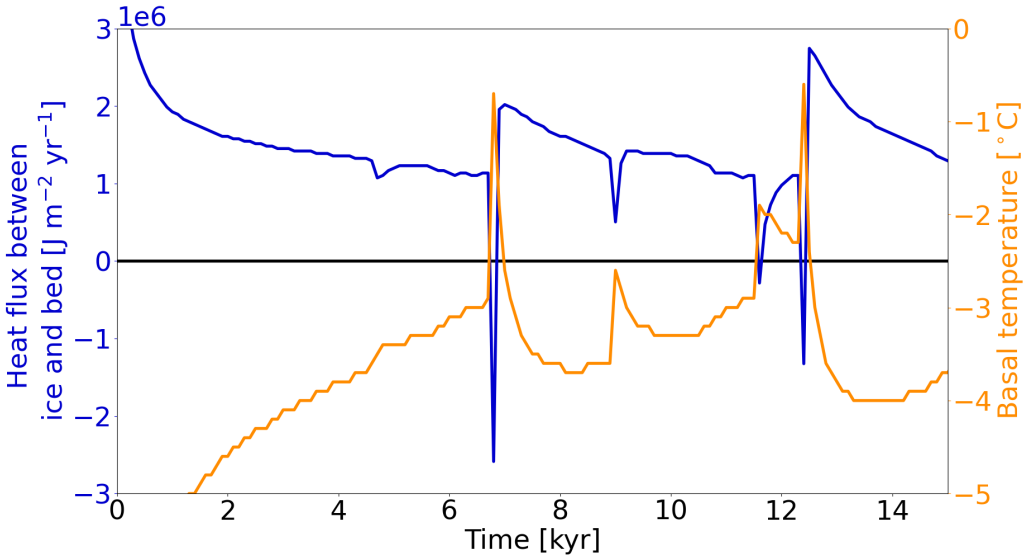


Figure 8. Heat flux at the base of the ice sheet (positive from bed into ice) and basal ice temperature for a grid cell in the center of the pseudo-Hudson Strait (grid cell center at $x = 376.5625$ km and $y = 248.4375$ km, white star in Fig. 1) and parameter vector 1 with the 1 km deep bed thermal model (17 non-linearly-spaced levels) using the GSM. The horizontal grid resolution is 3.125 km.

With only one bed thermal layer (20 m deep, removing most of the heat storage), the variance of the average basal temperature with respect to the pressure melting point in the pseudo-Hudson Strait increases (Fig. S20) and more heat is available to warm the surrounding ice (no or smaller heat flux into the bed, Fig. S21). The additional heat increases the mean pseudo-Hudson Strait ice volume change and duration (50 % and 65 %, respectively, Fig. 6). Due to the larger changes in pseudo-Hudson Strait ice volume and average basal temperature with respect to the pressure melting point, the ice sheet requires more time to reach the pre-surge state when only one bed thermal layer is used. Therefore, the period increases (60 %) while the number of surges drops. These differences in surge characteristics exceed the MNEEs (Table 5). The stronger surges (larger pseudo-Hudson Strait ice volume change) lead to overall less ice volume in the pseudo-Hudson Strait (Table S7).

Running PISM without the 1 km deep (20 equally-spaced levels) bed thermal model yields similar behavior as the GSM, further underlining the impact of a bed thermal model. The mean period, mean duration, and mean ice volume change all increase (80 %, 70 %, and 396 %, respectively; Fig. 7). In contrast to the GSM characteristics, the number of surges increases for runs without a bed thermal model. However, the standard deviation is large and the change in the number of surges is somewhat misleading. The number of surges decreases for 6 out of 9 runs. Parameter vectors showing an increase in the number of surges without a bed thermal model show very few surges (e.g., Fig. S22) or transition to a constantly active ice stream when the bed thermal model is included. As for the GSM, the stronger surges lead to an overall smaller ice sheet in the surge affected-area (Table S8).

3.3.2 Basal temperature at the grid cell interface

Another modeling choice that affects the thermal activation of basal sliding is the approach to determining the basal temperature at the grid cell interface. The most straightforward approach to determining the basal temperature with respect to the pressure melting point at the grid cell interface ($T_{\mathrm{bp,I}}$) is to use the mean of the two adjacent basal **T**emperatures with respect to the **p**ressure **m**elting point at the grid cell **Cen**ters (TpmCen).

$$T_{\mathrm{bp,I}} = 0.5 \cdot \left(T_{\mathrm{bp,L}} + T_{\mathrm{bp,R}} \right), \tag{16}$$

where $T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$ $T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$ are the grid cell center basal temperatures with respect to the pressure melting point to the left and right of the interface, respectively. Similarly for upper and lower grid cells adjacent to a horizontally aligned interface. However, this approach does not explicitly account for ice thickness changes at the grid cell interface.

TpmInt, on the other hand, calculates the basal temperature at the **Int**erface (T_{I}) by averaging the adjacent grid cell center basal temperatures (T_{L} and T_{R}, Eq. (17a)). $T_{\mathrm{bp,I}}$ $T_{\mathrm{bp,I}}$ is then determined by using the interface ice sheet thickness (average of adjacent grid cell center ice thicknesses H_{L} and H_{R}, Eq. (17b)).

$$T_{\mathrm{I}} = 0.5 \cdot (T_{\mathrm{L}} + T_{\mathrm{R}}) \tag{17a}$$

$$T_{\mathrm{bp,I}} = T_{\mathrm{I}} + \beta_{\mathrm{P}} \frac{H_{\mathrm{L}} + H_{\mathrm{R}}}{2}, \tag{17b}$$

where $\beta_{\mathrm{P}} = 8.7 \cdot 10^{-4}\,^{\circ}\mathrm{C\,m^{-1}}$ is the standard basal melting point depression coefficient. When TpmInt is used with the upwind scheme and the basal ice velocity exceeds 20 $\mathrm{m\,yr^{-1}}$, Eq. 17a is replaced by $T_{\mathrm{I}} = T_{\mathrm{up}}$, where T_{up} $T_{\mathrm{I}} = T_{\mathrm{up}}$, where T_{up} is the upstream adjacent grid cell center basal temperature.

The last approach (TpmTrans) attempts to represent heat transfer from sub-glacial hydrology and ice advection by accounting for extra warming above the pressure melting point given by

$$T_{\mathrm{add}} = M_{\mathrm{b}} \cdot \frac{L_{\mathrm{H}}}{c_{\mathrm{H}}} \cdot \frac{1}{H_{\mathrm{b}}} \cdot \Delta t, \tag{18a}$$

where M_{b} is the basal mass balance in $\mathrm{m\,yr^{-1}}$ (positive for melt), $L_{\mathrm{H}} = 3.35 \cdot 10^{5}$ $\mathrm{J\,kg^{-1}}$ the specific latent heat of fusion of water/ice, $c_{\mathrm{H}} = 2097$ $\mathrm{J\,kg^{-1}\,K^{-1}}$ the heat capacity of ice at 273.03 K, H_{b} the basal ice layer thickness in m, and Δt the current model time step in yr. In an intermediate calculation step, the temporary basal temperature at the grid cell center $T_{\mathrm{Im,C}}$ is calculated by accounting for the additional heating T_{add}

$$T_{\mathrm{Im,C}} = T_{\mathrm{C}} + T_{\mathrm{add}}, \tag{18b}$$

where T_{C} is the basal temperature at the grid cell center. The basal temperature with respect to the pressure melting point at each adjacent grid cell center $T_{\mathrm{bp,Im,C}}$ is then calculated using the interface ice thickness.

$$T_{\mathrm{bp,Im,C}} = T_{\mathrm{Im,C}} + \beta_{\mathrm{P}} \frac{H_{\mathrm{L}} + H_{\mathrm{R}}}{2} \tag{18c}$$

In the intermediate steps to calculate the interface temperature (Eq. (18b) and (18c)), $T_{\mathrm{Im,C}}$ and $T_{\mathrm{bp,Im,C}}$ are allowed to exceed the pressure melting point. This temporary higher basal temperature is an attempt to account for heat transported to the interface by ice advection and basal water.

$$\text{IF } T_{\mathrm{bp,Im,C}} > 0^{\circ}\mathrm{C}: \quad T_{\mathrm{bp,Im,C}} = \min\left(0.5^{\circ}\mathrm{C}, 0.5 \cdot T_{\mathrm{bp,Im,C}}\right) \tag{18d}$$

(18e)

~~Averaging the adjacent basal temperatures with respect to the pressure melting point at the grid cell center ($T_{\mathrm{bp,Im,L}}$ and $T_{\mathrm{bp,Im,R}}$) yields the final basal temperature with respect to the pressure melting point at the interface ($T_{\mathrm{bp,I}}$).~~

(18f)

Averaging the adjacent basal temperatures with respect to the pressure melting point at the grid cell center ($T_{\mathrm{bp,Im,L}}$ and $T_{\mathrm{bp,Im,R}}$) yields the final basal temperature with respect to the pressure melting point at the interface ($T_{\mathrm{bp,I}}$).

$$T_{\mathrm{bp,I}} = 0.5 \cdot \left(T_{\mathrm{bp,Im,L}} + T_{\mathrm{bp,Im,R}}\right) \tag{18g}$$

Note that neither the grid cell center nor the interface basal temperature may exceed the pressure melting point (only the basal temperature in the intermediate calculation steps).

The GSM reference setup (no hydrology) uses TpmTrans. The additional heat embodied in T_{add} warms up the grid cell interface. Without the extra warming (TpmInt), 4 out 5 parameter vectors do not show any surges. For the only run that still has cyclic behavior (parameter vector 1), the number of surges decreases by $84\ \%$ (note that runs without surges are considered for the number of surges in Fig. 6). Using TpmInt with an upwind scheme leads to slightly more surges (difference of $7\ \%$ and, therefore, on the same order of magnitude as the MNEE ($4\ \%$, Table 5)). Sporadic surges now occur in all but one run, leading to a large increase in the mean period ($1645\ \%$, Fig. 6).

The most straightforward approach, TpmCen, leads to $75\ \%$ fewer surges, and an increase in mean period and mean duration ($609\ \%$ and $43\ \%$, respectively). The mean pseudo-Hudson Strait ice volume change decreases ($-61\ \%$). Note that the TpmInt, TpmInt uwpind, and TpmCen surge characteristics are difficult to compare due to the different number of runs considered (except for the number of surges, decrease of $97\ \%$ vs. $90\ \%$ vs. $75\ \%$, respectively). Due to significantly fewer surges, the mean pseudo-Hudson Strait ice volume increases for runs with TpmInt, TpmInt uwpind, and TpmCen (Table S9).

3.3.3 Basal temperature ramps at different resolutions

Here we examine the effect of different basal temperature ramps (thermal activation criteria for basal sliding) at $3.125\ \mathrm{km}$ horizontal grid resolution and determine ramps for the coarse resolution runs that best match the $3.125\ \mathrm{km}$ model results (later used in Sec. 3.4.1). For coarse resolutions, changing the basal temperature ramp can lead to a shift from oscillatory to non-oscillatory behavior (compare $25\ \mathrm{km}$ runs in Fig. S23 and S24).

When running the GSM at $3.125\ \mathrm{km}$ horizontal grid resolution, surges are apparent for all tested basal temperature ramps. Due to an earlier sliding onset and easier surge propagation, increasing the width of the temperature ramp generally increases

the mean pseudo-Hudson Strait ice volume change and duration (Fig. 9). The ice sheet takes longer to recover from the surge (longer regrowth phase), increasing the mean period and decreasing the average number of surges. Running the GSM without a basal temperature ramp leads to small but significant (according to the MNEEs) differences in the mean duration (-7 %).

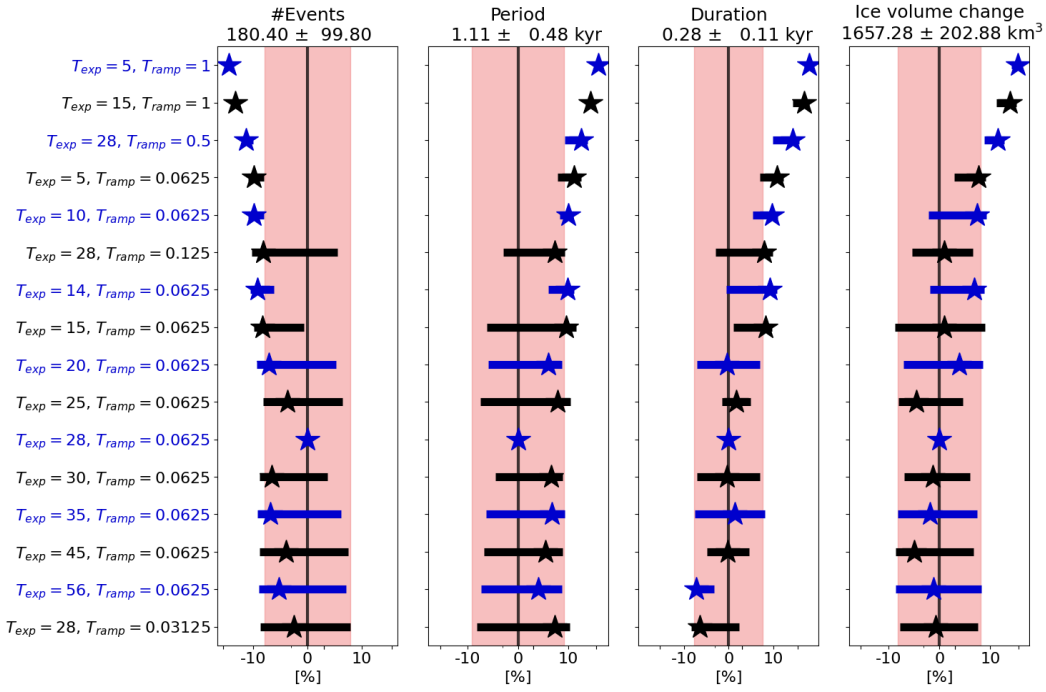


Figure 9. Percentage differences in surge characteristics compared to the GSM reference setup (~~$T_{\text{ramp}} = 0.0625$~~$T_{\text{ramp}} = 0.0625$, ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$) for different basal temperature ramps at 3.125 km horizontal grid resolution (average of the 5 parameter vectors). The ramps are sorted from widest (first row) to sharpest (last row, see Fig. S25 for a visualization of all ramps). Otherwise same as Fig. 6. No runs crashed and all runs had more than 1 surge event. The exact values are given in Table S10.

Except for the three widest ramps, the mean ice volume bias is less than one percent. The RMSE, on the other hand, is roughly 8 %, indicating that the average pseudo-Hudson Strait ice volume is similar, but the timing of surges varies even for small differences in the width of the ramp (Table S10).

We compare the different temperature ramps at 25 km, 12.5 km and 6.25 km horizontal grid resolution by calculating a single score for the mean and standard deviation of all surge characteristics (Sec. S7.3). The ramps yielding the smallest differences compared to the 3.125 km reference setup are listed in Table S11 and shown in Fig. S26. These results may be different for a different reference setup (see Table S22 for a comparison of different reference setups with local basal hydrology).

At 25 km horizontal grid resolution, only 3 out of 12 basal temperature ramps remain after removing the ramps for which the sum of scores (score-mean + score-std, last column in Table S11) differs by more than 50 % from the minimum sum of scores (bold numbers in last column in Table S11). The minimum scores for the mean and standard deviation occur for the same ramp

($T_{\mathrm{exp}} = 5$, $T_{\mathrm{ramp}} = 0.5$ $T_{\mathrm{exp}} = 5$, $T_{\mathrm{ramp}} = 0.5$), clearly identifying it as the ramp that best resembles the 3.125 km horizontal grid resolution reference runs. For the two finer horizontal grid resolutions, the minimum mean and standard deviation scores arise for different temperature ramps, preventing the determination of a single best ramp.

A more physically-based approach to determining an appropriate scale-compensating temperature ramp stems from our motivation for research question Q5 above. We bundle all 3.125x3.125 km grid cells of our reference runs into patches of, e.g., 64 grid cells. Each patch represents a coarser, e.g., 25x25 km grid cell. We then determine the warm-based fraction (basal temperature at the pressure melting point) and the mean basal temperature with respect to the pressure melting point of each patch. We can then estimate the parameters T_{ramp} and T_{exp} T_{ramp} and T_{exp} of the basal temperature ramp (Eq. (8)) by plotting the warm-based fraction against the mean basal temperature for all patches (e.g., Fig. 10) and fitting a basal temperature ramp with the preliminary assumption that a corresponding coarse grid cell should have an ice streaming fraction proportional to the sub-grid warm-based area.

However, this upscaling analysis does not account for the connectivity between the faces of, e.g., a 25 km grid cell. Without a continuous warm-based channel from one grid cell interface to another, there should be effectively no basal sliding across the grid cell, even when the average basal temperature is close to the pressure melting point. Consequently, the best estimate for the two parameters of the basal temperature ramp should be a lower bound to the points in the scatter plot.

Furthermore, the upscaling results depend on the bed properties (soft sediment vs. hard bedrock) and the specific scenario (surge vs. quiescent phase). Therefore, we only consider patches within the pseudo-Hudson Strait area during surges. Due to the limited storage capacity for the 10 yr output fields, only the first 10 kyr after the first surge are used for the upscaling experiments.

The upscaling results agree well with the score analysis at 25 km horizontal grid resolution. Both indicate that at this resolution, the ramp $T_{\mathrm{exp}} = 5$, $T_{\mathrm{ramp}} = 0.5$ $T_{\mathrm{exp}} = 5$, $T_{\mathrm{ramp}} = 0.5$ (first row in Table S11, Fig. 10) gives results that best match those of the 3.125 km reference run. The two approaches yield a similar range of temperature ramps at 12.5 and 6.25 km horizontal grid resolution, but the upscaling experiments generally favor wider temperature ramps (Table S11 and Fig. S27 and S28). This is likely a consequence of the above-mentioned role of sub-grid warm-based connectivity not accounted for in the upscaling analysis. When using the resolution-dependent ramp of Eq. (9), the upscaling experiments, therefore, provide a lower bound of $T_{\mathrm{exp}} = 5$ $T_{\mathrm{exp}} = 5$. Upscaling experiments with local basal hydrology lead to similar results.

3.3.4 Smooth sediment transition zone and non-flat topography

The effects of a smooth sediment transition zone (instead of an abrupt transition from hard bedrock (0 % sediment cover) to 100 % (soft) sediment cover) and a non-flat topography on surge characteristics are examined here.

The abrupt transition from hard bedrock to soft sediment (pseudo-Hudson Bay and Hudson Strait) in the GSM reference setup and the corresponding difference in basal sliding coefficient provide an additional heating source due to shearing between slow and fast-moving ice. This additional heat appears to foster the propagation of small surges along the transition zone (e.g., 6 to 6.3 kyr in the upper row of video 03 of Hank (2023)). Incorporating a smooth transition zone (3.125 km or 25 km wide) affects the location of the small-scale surges (not considered in surge characteristics) but shows only minor differences for the

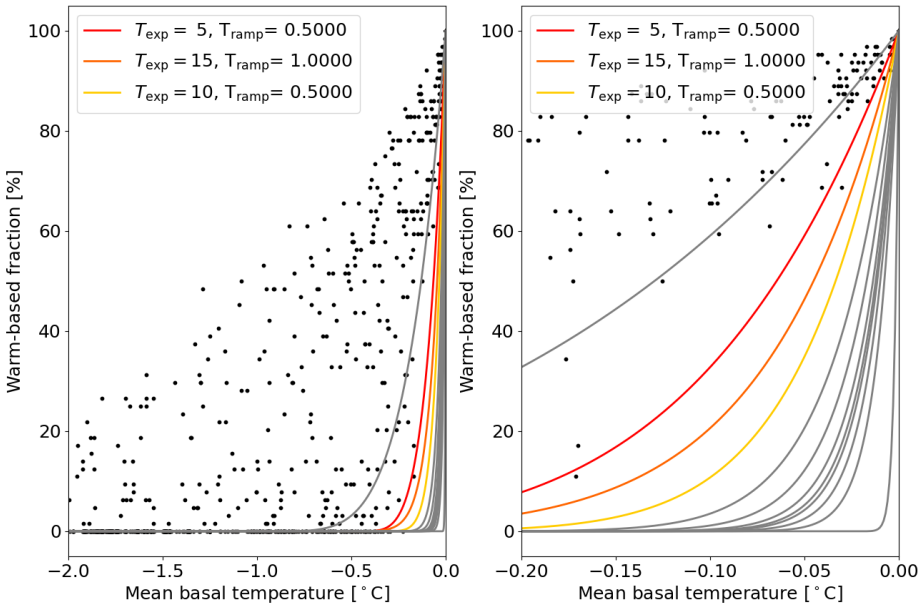


Figure 10. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 25 km horizontal grid resolution including all 5 parameter vectors using the GSM. For example, an upscaled 25 km patch (containing 64 3.125 km grid cells) with 32 3.125 km grid cells at the pressure melting point and 32 3.125 km grid cells at −1°C with respect to the pressure melting point has a warm-based fraction of 50 % and a mean basal temperature of −0.5°C. Only grid cells within the pseudo-Hudson Strait and time steps within the surges of the 10 kyr after the first surge are considered. The restriction to the 10 kyr after the first surge for these experiments is set by storage limitations due to the high temporal resolution of the model output fields (10 yr). The colored ramps correspond to the 25 km horizontal grid resolution basal temperature ramps in Table S11 and the gray lines show all other ramps that were tested at this resolution.

615 major surges (< 7.5 %) for all surge characteristics, Fig. 6). The mean bias for both widths is < 1 %, indicating only minor differences in ice volume between an abrupt and smooth transition. However, the timing of surges varies for different transition zones (RMSE ≤ 8 %, Fig. 11). A wider transition zone (more sediment surrounding the pseudo-Hudson Strait and Hudson Bay) generally favors an earlier sliding onset (e.g., Fig. 11), but the details depend on the parameter vector in question.

Similar to the GSM results, the PISM percentage differences between a smooth (reference setup) and abrupt sediment
620 transition show no significant effect, except a 22 % increase in surge duration (Fig. 7).

Adding a 200 m deep pseudo-Hudson Strait and Hudson Bay with a smooth transition zone and 500 m deep ocean to the GSM setup displaces the origin of surges slightly further inland. Due to both the resultant warmer basal temperature and depressed pressure melting point, the surges propagate faster, last longer, and evacuate more ice volume (Fig. 6). The

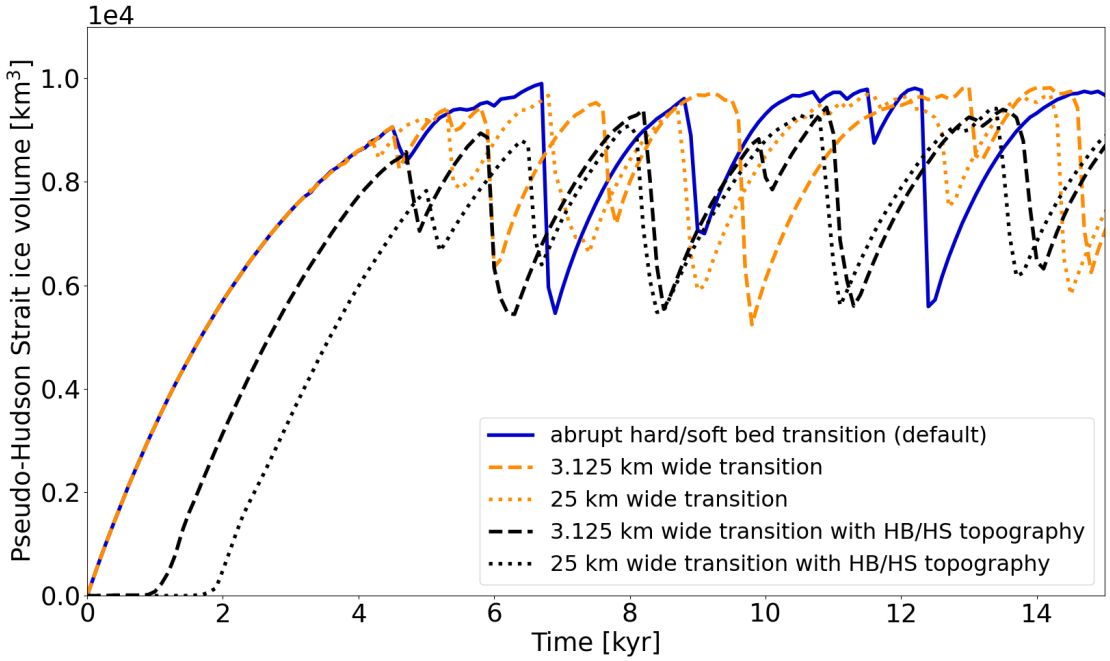


Figure 11. Pseudo-Hudson Strait ice volume for GSM parameter vector 1 and three different bed configurations. The horizontal grid resolution is 3.125 km. Note that the width of the topographical transition zone matches the width of the soft to hard bed transition zone. In experiments with a pseudo-Hudson Bay/Hudson Strait (HB/HS) topography, the pseudo-Hudson Strait topography is below sea level, increasing the time required for glaciation. A wider transition zone (larger area below sea level) leads to a later glaciation.

topography slopes down towards the pseudo-Hudson Strait, increasing the ice inflow from the surroundings. The ice sheet recovers faster from the previous surge, decreasing the mean period. Note that Fig. 6 shows an increase in the mean period, but this is somewhat misleading due to the now early surges for parameter vector 0 and the subsequent large increase in the mean period (~ 100 %, no surges in the middle part of the run due to cold surface temperatures (Fig. S29)). All other parameter vectors show a decrease in the mean period for both widths of the transition zone. The mean bias indicates a decrease in ice volume of ~ 6.5 % for runs with a non-flat topography caused by the larger surges. The pseudo-Hudson Strait topography also suppresses the small surges otherwise observed in the vicinity of the pseudo-Hudson Strait. A detailed comparison of an individual run is presented in Sec. S7.4.

Comparing the results for two different widths of the topographic transition zone (-200 m to sea level) indicates fewer but stronger surges (increase of mean pseudo-Hudson Strait ice volume change by 9 %, Fig. 6) for a wider transition zone. The gentler slope increases the width of the ice stream and, thereby, the ice flux out of the pseudo-Hudson Strait. (video 04 of Hank (2023)). The increased flux leads to a decreased pseudo-Hudson Strait ice volume at the end of the surge. The stronger surges

for a wider transition zone increase the recovery time, leading to a smaller increase in the number of surges than for the narrow transition zone (difference of 16 %, Fig. 6).

While imposing a non-flat topography fosters surges in both models, the increase in mean ice volume change is much larger in PISM (390 %) than in the GSM (maximum $\sim$ 17 %), leading to a longer regrowth-phase (79 % increase in mean period) and overall less ice volume (mean bias -30 %, Table S13). The longer recovery times in PISM outweigh the effect of earlier sliding onsets leading to more surges described above for the GSM. Therefore, the number of surges decreases in PISM (while increasing in the GSM) when using a non-flat topography (Fig. 7).

Since the topography will vary from ice stream to ice stream, we stick to a flat topography for the remaining experiments.

3.3.5 Basal hydrology

The effects of adding a simple local basal hydrology model to the GSM are examined here. The local basal hydrology sets the basal water thickness by calculating the difference between the basal melt rate and a constant basal drainage rate (rBedDrainRate in Table 1). This sub-glacial hydrology provides a simple and computationally efficient way to capture changes in basal sliding velocities due to effective pressure variations (Drew and Tarasov, 2022, under review). However, it does not account for basal ice accumulation, englacial or supraglacial water input, or horizontal water transport.

The basal water thickness (~~hwb~~ h_{wb}) and an estimated effective bed roughness scale (~~$hwb,Crit$~~ $h_{\mathrm{wb,Crit}}$ in Table 1) determine the effective pressure coefficient

$$N_{\mathrm{C,eff}} = 1 - \min\left(\frac{\mathrm{h}wb}{\mathrm{h}wb,Crit}\frac{h_{\mathrm{wb}}}{h_{\mathrm{wb,Crit}}}, 1.0\right)^{3.5} \tag{19}$$

The basal water thickness is limited to ~~$hwb,Crit = 10$ m~~ $h_{\mathrm{wb,Crit}} = 10$ m and is set to ~~$hwb = 0$ m~~ $h_{\mathrm{wb}} = 0$ m where the ice thickness is less than 10 m and where the temperature with respect to the pressure melting point is below -0.1°C. Experiments with ~~$hwb,Crit = 5$ m~~ $h_{\mathrm{wb,Crit}} = 5$ m yield the same results, and removing all the water for $H < 1$ m, $H < 50$ m, and $T_{\mathrm{bp}} < -0.5$°C does not significantly (according to Sec. 3.2) affect the model results. The effective pressure at the grid cell interface is then

$$N_{\mathrm{eff}} = g\rho_{\mathrm{ice}} \cdot 0.5\,(H_{\mathrm{L}} N_{\mathrm{C,eff,L}} + H_{\mathrm{R}} N_{\mathrm{C,eff,R}})\,, \tag{20}$$

where $g = 9.81$ m s^{-2} is the acceleration due to gravity, $\rho_{\mathrm{ice}} = 910$ kg m^{-3} the ice density, H the ice thickness and the subscripts L and R denote the adjacent grid cells to the left and right of the interface, respectively (similarly for upper and lower grid cells adjacent to a horizontally aligned interface). We enforce that N_{eff} never falls below 10 kPa (denominator in Eq. (21), similar results for $N_{\mathrm{eff,min}} = 5$ kPa). Finally, the effective pressure of each grid cell alters the basal sliding coefficient in the sliding law (Eq. (6a)) according to

$$C_{\mathrm{b}} = C_{\mathrm{b}} \cdot \min\left(10, \max\left(0.5, \frac{N_{\mathrm{eff,Fact}}}{N_{\mathrm{eff}} + 10^4\ \mathrm{Pa}}\right)\right), \tag{21}$$

where $N_{\mathrm{eff,Fact}}$ is the effective pressure factor (Table 1). The change of the basal sliding coefficient C_{b} is, therefore, limited to $C_{\mathrm{b}} \cdot 0.2$ to $C_{\mathrm{b}} \cdot 10$. Allowing a larger change of $C_{\mathrm{b}} \cdot 0.1$ to $C_{\mathrm{b}} \cdot 20$ does not significantly (according to Sec. 3.2) change the model results.

When running the GSM with the local sub-glacial hydrology model, intermediate values are used for all 3 parameters (the effective bed roughness scale ~~$h_{wb,Crit} = 0.1$ m~~ $h_{\mathrm{wb,Crit}} = 0.1$ m (Eq. (19)), the constant bed drainage rate rBedDrainRate $\simeq 0.003$ m yr^{-1}, and the effective pressure factor $N_{\mathrm{eff,Fact}} \simeq 63246$ Pa (Eq. (21))) for all 5 parameter vectors. However, different values were tested for all 3 parameters (not shown). In general, a larger $N_{\mathrm{eff,Fact}}$ increases the basal sliding coefficient (Eq. (21)) and, therefore, leads to fewer but stronger surges. The results for ~~$h_{wb,Crit}$~~ $h_{\mathrm{wb,Crit}}$ and rBedDrainRate are not as straightforward to interpret. The model response varies for the 2 tested parameter vectors, and the changes are generally smaller than the MNEEs of Table 5.

Adding the local basal hydrology model to the GSM increases the mean ice volume change and duration by 20 % and 12 %, respectively (Fig. 6, exceeding the MNEEs (Table 5)). The stronger surges are due to the reduction of effective pressure and, thus, increased sliding (Eq. (21) and (6a)). The mean period increases (17 %) while the number of surges decreases (−4 %), but the standard deviations are large.

Since the local hydrology model effectively increases the basal sliding coefficient, we test if this impact can be replicated simply by increasing the sliding coefficients (Table 1) in a GSM configuration without basal hydrology. Doubling the soft bed sliding coefficient leads to similar or larger maximum basal sliding velocities and, consequently, maximum ice fluxes but a smaller increase in the mean period (12 % vs. 17 %) and mean pseudo-Hudson Strait ice volume change (11 % vs. 20 %) than that of the local hydrology model. Increasing the hard bed sliding coefficient has no significant effect on the surge characteristics (pseudo-Hudson Bay and Hudson Strait are soft-bedded, Fig. 6). Intermediate increases in the basal sliding coefficients (not shown) also do not capture the surge characteristics of the basal hydrology model. Therefore, simply changing the basal sliding coefficients cannot replace the basal hydrology model.

3.3.6 Sensitivity experiments without a significant effect

The effect of an experiment is considered insignificant when the change in surge characteristics is smaller than the MNEEs (Sec. 3.2). This is the case for different weights of the adjacent minimum basal temperature when calculating the basal interface temperature (Q_6), for different implementations of the basal hydrology (Q_9), and when using basal hydrology instead of the basal temperature ramp as the primary smoothing mechanism (Q_{10}). The details of these experiments are presented in Sec. S8.1, S8.2, and S8.3, respectively. We want to emphasize that experiments without a significant effect can still have physical relevance, but it is currently hidden within the numerical sensitivities.

3.4 Convergence study

In this section, we examine the horizontal grid resolution dependence of the GSM and PISM model results. Model results are considered as converging when the differences in surge characteristics decrease with increasing horizontal grid resolutions.

3.4.1 GSM convergence study

Significant differences in surge characteristics occur when changing the horizontal grid resolution. These differences can be as large as a highly oscillatory behavior at 3.125 km and no oscillations at 25 km horizontal grid resolution (Fig. S23). Changing the basal temperature ramp can somewhat counteract this discrepancy by enabling basal sliding at lower basal temperatures for coarser grid resolutions (Fig. S24 and video 05 of Hank (2023)). Further details on discrepancies between horizontal grid resolutions for individual parameter vectors are discussed in Sec. S9.1.

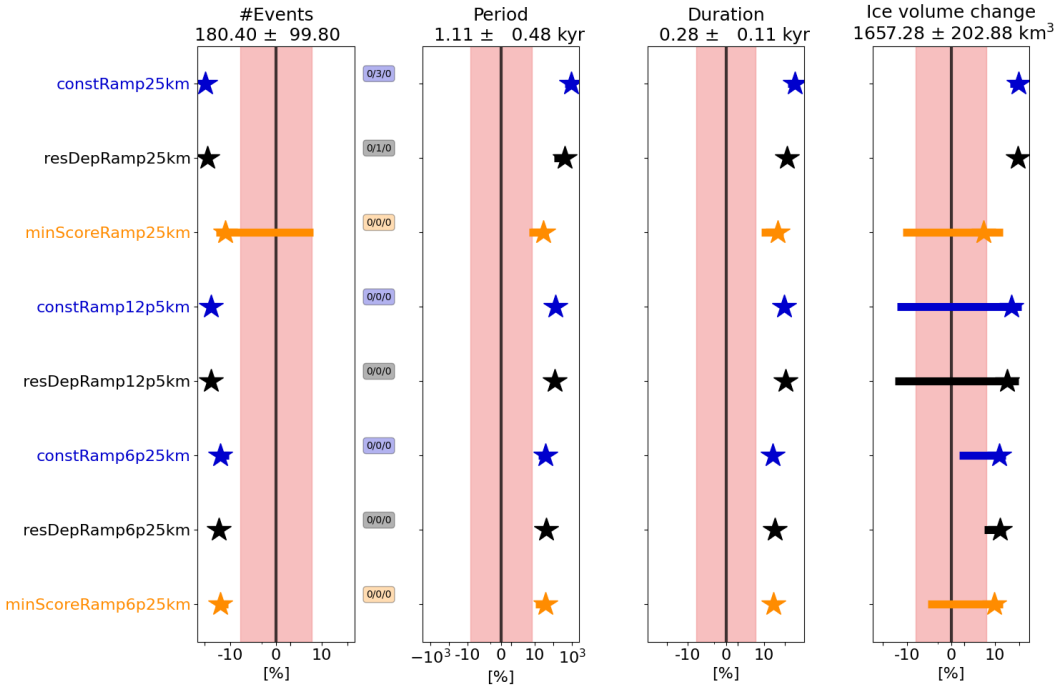


Figure 12. Percentage differences in surge characteristics compared to the GSM reference setup (3.125 km horizontal grid resolution) for model setups with coarser (25 km, 12.5 km, and 6.25 km) horizontal grid resolutions (average of the 5 parameter vectors). The different colors were added for visual alignment of the individual model setups and mark model setups with constant (blue), resolution-dependent (black), and minimum score (orange, Sec. S7.3) basal temperature ramps. The resolution-dependent ramps (~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$) and constant ramp (black line, ~~$T_{\text{ramp}} = 0.0625$~~$T_{\text{ramp}} = 0.0625$, ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$) are shown in Fig. 2. The minimum score basal temperature ramps are ~~$T_{\text{ramp}} = 0.5$~~$T_{\text{ramp}} = 0.5$, ~~$T_{\text{exp}} = 5$~~$T_{\text{exp}} = 5$ at 25 km and ~~$T_{\text{ramp}} = 0.125$~~$T_{\text{ramp}} = 0.125$, ~~$T_{\text{exp}} = 45$~~$T_{\text{exp}} = 45$ at 6.25 km horizontal grid resolution. At 12.5 km, the minimum score ramp is the same as the resolution-dependent ramp. Otherwise same as Fig. 6. Further details of each individual experiments are provided in Sec. 3.4.1 and the supplement.

We compare the differences in surge characteristics for different basal temperature ramps at each resolution (Fig. 12). We examine: a constant ramp (~~$T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 28$~~$T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 28$), a resolution-dependent temperature ramp (~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$, Fig. 2), and the ramp with the smallest differences in surge characteristics (bold mean score in Ta-

ble S11). Note that the large differences in mean period at 25 km resolution are caused by long time intervals without any oscillations in the coarse resolution runs (Table S18). 25 km, 12.5 km, and 6.25 km runs show progressively smaller differences for the constant and resolution-dependent ramp, indicating model convergence. Convergence of the GSM results with increasing grid resolutions is further supported by successively smaller pseudo-Hudson Strait ice volume RMSE and mean bias values (Table S19). RMSE and mean bias are smaller across all resolutions when using a resolution-dependent instead of a constant temperature ramp (except for the RMSE at 12.5 km horizontal grid resolution).

All three basal temperature ramps lead to similar differences in surge characteristics at 6.25 km and 12.5 km horizontal grid resolution (Table S18). At 25 km resolution, the ramp with the minimum differences in surge characteristics significantly improves the agreement with the 3.125 km runs, with differences smaller than for any other ramp or resolution. This could either be a coincidence or indicate that despite thorough testing, the best ramp has not been found at 6.25 km and 12.5 km horizontal grid resolution. Since other ramps at 25 km horizontal grid resolution show only slightly larger differences in surge characteristics (e.g., difference of 0.23 in the mean score, Table S11), it is unlikely that it is just a coincidence. However, the sensitivity of the surge characteristics to grid refinement remains, no matter the choice of the temperature ramp, with differences significantly exceeding the MNEEs from Sec. 3.2.

Since including a sub-glacial hydrology model significantly affects the surge characteristics, we also examine the horizontal grid resolution scaling with a local basal hydrology model (Sec. S9.2). The results show overall smaller differences (relative to the 3.125 km reference simulations) in surge characteristics than without (Table S22 vs. S18). The analysis of the convergence study (with and without basal hydrology) and the upscaling experiments in Sec. 3.3.3, therefore, suggest a resolution-dependent temperature ramp with ~~T_{exp}~~ T_{exp} between 5 and 10.

3.4.2 PISM convergence study

Similar to the results presented for the GSM, the ice volume RMSE and mean bias show convergence under systematic grid refinement (Table S26). However, for the three resolutions examined here, the PISM surge characteristics show convergence for the mean duration and ice volume change but not the number of surges and mean period (Table S25). Note that 4 out of 9 ~~12.5 km runs~~ runs at 12.5 km horizontal grid resolution did not finish within the time limit of the computational cluster and are considered as crashed runs (potentially skewing the statistics). Additionally, 1 run at 12.5 km resolution did not show any surges and was also excluded from the analysis. The differences in surge characteristics for different grid resolutions are, in general, larger than the MNEEs, but can be smaller (mean ice volume change of the 25 km runs).

4 Results Summary and Discussion

This section summarizes our modeling results in the context of the research questions outlined in Sec. 1.3 and previous modeling studies.

Minimum numerical error estimates

Q_1 *What is the threshold of MNEEs in the two models?*

~~Since the model error with respect to the exact analytical solution can not be determined in the context of this study, we use MNEEs to determine a minimum threshold for the significance of a change in the model configuration (Sec. 3.2).~~

The MNEEs can be as large as $16\ \%$ (Table 5 and 6).

Given the nonlinearities in the SSA (or higher approximation) ice sheet system, there is no a priori reason to confidently assume other ice sheet models will have ignorable MNEEs for unstable contexts such as surge cycling and grounding line response. Therefore, it is crucial to determine MNEEs (or a comparable ~~metric~~threshold) to minimize the possibility of interpreting numerical errors as a physical response to a change in model setup.

In contrast to the findings of Souček and Martinec (2011), adding low levels of surface temperature noise does not significantly affect the GSM and PISM results (Table S3 and S6). Potential reasons for the different model responses are the use of an Arakawa A grid (velocities and temperatures are calculated on the same node, Arakawa and Lamb, 1977) and the JOSH (JOint Shallow-ice/Higher-order model) ice sheet dynamics in Souček and Martinec (2011).

Sensitivity experiments with a significant effect

Q_2 *Is the inclusion of a bed thermal model a controlling factor for surge activity?*

Including a $1\ \mathrm{km}$ deep bed thermal model significantly (according to the MNEEs) affects the surge characteristics in the GSM and PISM. The additional heat stored in the bed changes the thermal conditions at the ice-bed boundary, dampening the ice volume change during a surge (Fig. 6 and 7). Models with similar setups but without a bed thermal model likely overestimate the ice volume change during a surge (e.g., Calov et al., 2010; Brinkerhoff and Johnson, 2015). Therefore, the inclusion of a bed thermal model is a key aspect of modeling ice stream surge cycling.

Q_3 *Do different approaches for determining the grid cell interface basal temperature significantly affect surge behavior, and if yes, which one should be implemented?*

The choice of approach for determining the basal temperature at the grid cell interface significantly changes the surge characteristics. Without considering additional heat transfer to the grid cell interface (as an attempt to represent heat contributions from sub-glacial hydrology and sub-grid ice advection), the number of surges decreases by at least $75\ \%$. The additional heat is, therefore, an essential component for modeling surges in the GSM.

This additional heat transfer to the grid cell interface is comparable to spreading $50\ \%$ of the basal heating effect from sliding in a grid cell to the surrounding grid cells used in mPISM (latest version based on PISM v0.7.3) (e.g., Ziemen et al., 2014, 2019; Schannwell et al., 2023). This spreading of basal heating warms the grid cells adjacent to an ice stream and was necessary to model Heinrich Event-like surges (Florian Ziemen, personal communication). While no additional heat transfer was added to PISM v2.0.2 used within this study, the till friction angles had to be reduced to model surges.

Q_5 *How different are the model results for different basal temperature ramps and what ramp should be used?*

Similar to Souček and Martinec (2011), we find significant differences in the period and amplitude of surges at all tested resolutions when using different implementations for thermal activation of basal sliding (the basal temperature ramp). In the GSM, a wider temperature ramp enables sliding onset at lower temperatures, fostering more extensive surge propagation and leading to stronger surges. However, the choice of the most appropriate temperature ramp at the finest resolution tested (3.125 km, Fig. 9) is unclear and identifying a single best ramp (fit of coarse resolution runs to 3.125 km runs) is challenging (Table S11). In general, a resolution-dependent ramp with ~~$T_{\exp}$~~ $T_{\exp}$ between 5 and 10 (Eq. (8) and (9)) yields the smallest differences between fine and low resolution simulations. However, given potential dependencies on the particular ice sheet model, we recommend resolution testing to determine the optimal basal temperature ramp.

~~To account for observational and experimental evidence of sub-temperate sliding (Barnes et al., 1971; Shreve, 1984; Echelmeyer and , avoid an abrupt onset of sliding at the warm/cold-based transition that causes refreezing on the warm-based side (Mantelli et al., 2019), and minimize resolution dependencies~~Nevertheless, a basal temperature ramp (or similar mechanism) should be implemented in all ice sheet models for contexts where surge onset/termination are important.

Q_6 *Does the abrupt transition between a soft and hard bed significantly affect surge characteristics?*

Incorporating a smooth transition zone with two different widths (3.125 km and 25 km) in the GSM does affect the location of proximal small-scale ice streams (video 03 of Hank (2023)). However, the abrupt transition is not the cause of the major surges in the GSM (Fig. 6) and PISM experiments (Fig. 7). Since the sediment cover can change within a few kilometers (e.g., Andrews and MacLean, 2003), we conclude that, despite the minor differences, an abrupt transition between soft and hard beds is a reasonable simplification, especially considering horizontal grid cell dimensions of 25 km or larger.

Q_7 *How does a non-flat topography affect the surge behavior?*

Imposing a non-flat topography leads to significantly longer and stronger surges (Fig. 6 and 7). As such, and in agreement with previous modeling studies (e.g., Winsborrow et al., 2010, and references within), ice streaming is sensitive to the basal topography.

Q_8 *What is the effect of a simplified basal hydrology on surge characteristics?*

Activating the local basal hydrology model (including the addition of effective pressure dependence into the sliding law) in the GSM significantly increases the surge duration and amplitude (Fig. 6). Somewhat stronger surges are expected due to the reduction in effective pressure introduced by the sub-glacial water. Model runs without sub-glacial hydrology will therefore tend to underestimate the strength of surges. In general, this also holds for subglacial hydrology models with higher complexity (Drew and Tarasov, 2022, under review). The importance of sub-glacial hydrology has also been shown in several other studies examining the effects of ice sheet surges and ice streaming within a continuum model approach (e.g., Fowler and Johnson, 1995; Fowler and Schiavi, 1998; Benn et al., 2019).

Sensitivity experiments without a significant effect

Q_4 *How much of the ice flow should be blocked by upstream or downstream cold-based ice, or equivalently, what weight should be given to the adjacent minimum basal temperature?*

Changing the weight of the adjacent minimum basal temperature for the basal sliding temperature ramp in the GSM yields a maximum difference of $15\ \%$ (Table S15). These somewhat small effects on surge characteristics are likely due to the fact that most surges propagate upstream (from the ocean to the pseudo-Hudson Bay) and the adjacent minimum basal temperatures (almost exclusively located upstream) have little potential to affect (e.g., partly block) the ice flow.

Q_9 *How significant are the details of the basal hydrology model on surge characteristics in PISM?*

Incorporating a mass-conserving horizontal transport hydrology model does not significantly change the surge characteristics in PISM (Fig. 7), indicating that the computationally much cheaper local hydrology model is a reasonable simplification for this context. More nuanced results, depending on the surge characteristics examined, are observed for the GSM (Drew and Tarasov, 2022, under review).

Q_{10} *What are the differences (if any) in surge characteristics between local basal hydrology and a basal temperature ramp as the primary smoothing mechanism at the warm/cold-based transition zone?*

Once included, the local basal hydrology is the primary smoothing mechanism. However, since the two smoothing mechanisms operate in different temperature regimes, a basal temperature ramp (representing sub-temperate sliding) cannot be replaced by a basal hydrology scheme (as in, e.g., Robel et al., 2013; Kyrke-Smith et al., 2014; Brinkerhoff and Johnson, 2015). The differences in surge characteristics are smaller than the MNEEs preventing further analysis.

Convergence study

Q_{11} *Do model results converge (decreasing differences when increasing horizontal grid resolution)?*

In general, both models exhibit convergence under systematic horizontal grid refinement for the overall ice volume (mean bias, Table S19+S23 and S26), but the solution is not fully converged at the finest resolutions tested. However, while all surge characteristics converge for the GSM (Table S18), PISM results do not show convergence for the number of surges and mean period (Table S25). This clearly illustrates that mean ice volume and, consequently, mean ice thickness, as presented, e.g., in Van Pelt and Oerlemans (2012), are insufficient metrics to determine whether cyclic model results exhibit a resolution dependency.

While other studies examining thermally induced ice streaming do not find a strong resolution dependence (Hindmarsh, 2009; Brinkerhoff and Johnson, 2015), these studies are not directly comparable. The different results are likely due to differences in the experimental design. For example, neither Hindmarsh (2009) nor Brinkerhoff and Johnson (2015) consider a bed thermal model. While Hindmarsh (2009) considers sub-temperate sliding, his model allows sliding far below the pressure melting point (order of $\delta = 1$ compared to $\delta = 0.01$ within this study, Eq. (10)) and focuses on steady

ice streams, not ice stream surge cycling. Both of the above studies analyze just one parameter vector, and there are some parameter vectors for which, e.g., the GSM exhibits only a minor resolution dependence.

Even though the studies are not directly comparable, the results of Brinkerhoff and Johnson (2015) offer some insight relevant to this study. For example, they suggest membrane stresses are necessary for convergence under horizontal grid refinement. The hybrid SIA/SSA ice dynamics used in the GSM and PISM might be insufficiently 'higher-order' and lead to a stronger resolution dependence than the schemes used in Hindmarsh (2009) and Brinkerhoff and Johnson (2015). However, GSM experiments with the SSA active everywhere show a resolution dependence comparable to the velocity-dependent SSA activation criteria (Table S24 and S18, respectively), indicating that the hybrid SIA/SSA ice dynamics are not the sole reason for the strong resolution dependence.

5 Conclusions

Within the limitations of hybrid SIA/SSA ice dynamics, we investigate the effect of ice sheet model numerics and discretization choices on surge characteristics often neglected in ice sheet modeling studies. We show how to reduce numerical and discretization sensitivities given finite computational resources and then how to determine the significance of model results given residual computationally unavoidable numerical sensitivities for surge cycling contexts. In particular, our analyses offer guidance in minimizing the resolution dependency by implementing a resolution-dependent basal temperature ramp for basal sliding thermal activation and increasing confidence in model results by determining Minimum Numerical Error Estimates (MNEEs). Based on these MNEEs, our results indicate that surge characteristics are significantly affected by the inclusion of a basal hydrology model. Not including the dampening effect of a bed thermal model on basal temperature changes, as has been the tendency in idealized process studies, overestimates the surge amplitude. The key takeaways of this study are the physical modeling choices and numerical sensitivities that must be considered when numerically modeling ice stream surge oscillations.

Code availability. TEXT

Data availability. TEXT

Code and data availability. The GSM source code (v01.31.2023) and run instructions are available at https://doi.org/10.5281/zenodo.7668472 (Tarasov et al., 2023). Instructions on how to install and run PISM and the PISM source code (v2.0.2) can be acquired from the repository at https://zenodo.org/record/6001196. Further information on how to recreate this work's results, input files, parameter vectors, and the analysis scripts used to determine the surge characteristcs can be found at https://doi.org/10.5281/zenodo.7905404 (Hank, 2023).

Sample availability. TEXT

Author contributions. TEXT

K.H. and L.T. conceptualized the ideas behind this study. All authors were involved in designing the experimental setup of the GSM. K.H. designed the experimental setup for PISM and performed the modeling analysis for both models under the supervision of L.T. All authors contributed to the results, interpretation, and writing of the manuscript.

Competing interests. The authors have no competing interests.

Disclaimer. TEXT

Acknowledgements. The authors thank *Andy Aschwanden*, *Ed Bueler* and *Constantine Khrulev* for support with the Parallel Ice Sheet Model (PISM). We thank *Ed Bueler* and *Daniel F Martin* for fruitful discussions about the bed thermal model and the numerical tolerances, respectively. We also thank *Florian Ziemen* and *Clemens Schannwell* for insightful discussions on modeling Heinrich Event-like surges. Finally, we thank two anonymous reviewers and the handling topic editor Ludovic Räss for their constructive comments. This research has been supported by an NSERC Discovery Grant (number RGPIN-2018-06658), the Canadian Foundation for Innovation, and the German Federal Ministry of Education and Research (BMBF) as a Research for Sustainability initiative (FONA) through the PalMod project. EM was supported by the European Union (ERC-2022-STG, grant number 101076793). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

Andrews, J. T. and MacLean, B.: Hudson Strait ice streams: A review of stratigraphy, chronology and links with North Atlantic Heinrich events, Boreas, 32, 4–17, https://doi.org/10.1080/03009480310001010, 2003.

Arakawa, A. and Lamb, V. R.: Computational Design of the Basic Dynamical Processes of the UCLA General Circulation Model, in: General Circulation Models of the Atmosphere, edited by CHANG, J., vol. 17 of *Methods in Computational Physics: Advances in Research and Applications*, pp. 173–265, Elsevier, https://doi.org/https://doi.org/10.1016/B978-0-12-460817-7.50009-4, 1977.

Bahadory, T. and Tarasov, L.: LCice 1.0-a generalized Ice Sheet System Model coupler for LOVECLIM version 1.3: Description, sensitivities, and validation with the Glacial Systems Model (GSM version D2017.aug17), Geoscientific Model Development, 11, 3883–3902, https://doi.org/10.5194/gmd-11-3883-2018, 2018.

Barnes, P., Tabor, D., and Walker, J. C. F.: The Friction and Creep of Polycrystalline Ice, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 324, 127–155, http://www.jstor.org/stable/77933, 1971.

Benn, D. I., Fowler, A. C., Hewitt, I., and Sevestre, H.: A general theory of glacier surges, Journal of Glaciology, 65, 701–716, https://doi.org/10.1017/jog.2019.62, 2019.

Brinkerhoff, D. J. and Johnson, J. V.: Dynamics of thermally induced ice streams simulated with a higher-order flow model, Journal of Geophysical Research F: Earth Surface, 120, 1743–1770, https://doi.org/10.1002/2015JF003499, 2015.

Bueler, E. and Brown, J.: Shallow shelf approximation as a "sliding law" in a thermodynamically coupled ice sheet model, J. Geophys. Res., 114, https://doi.org/10.1029/2008JF001179, 2009.

Bueler, E. and Van Pelt, W.: Mass-conserving subglacial hydrology in the Parallel Ice Sheet Model version 0.6, Geoscientific Model Development, 8, 1613–1635, https://doi.org/10.5194/gmd-8-1613-2015, 2015.

Calov, R. and Greve, R.: ISMIP HEINO. Ice Sheet Model Intercomparison Project - Heinrich Event INtercOmparison, pp. 1–15, http://www.pik-potsdam.de/~calov/heino/he_setup_2006_11_02.pdf, 2006.

Calov, R., Greve, R., Abe-Ouchi, A., Bueler, E., Huybrechts, P., Johnson, J. V., Pattyn, F., Pollard, D., Ritz, C., Saito, F., and Tarasov, L.: Results from the Ice-Sheet Model Intercomparison Project-Heinrich Event INtercOmparison (ISMIP HEINO), Journal of Glaciology, 56, 371–383, https://doi.org/10.3189/002214310792447789, 2010.

Courant, R., Friedrichs, K., and Lewy, H.: Über die partiellen Differenzengleichungen der mathematischen Physik, Mathematische Annalen, 100, https://doi.org/10.1007/BF01448839, 1928.

Cuffey, K. M., Conway, H., Hallet, B., Gades, A. M., and Raymond, C. F.: Interfacial water in polar glaciers and glacier sliding at -17 °C, Geophysical Research Letters, 26, 751–754, https://doi.org/10.1029/1999GL900096, 1999.

Drew, M. and Tarasov, L.: Surging of a Hudson Strait Scale Ice Stream: Subglacial hydrology matters but the process details don't, The Cryosphere Discussions, 2022, 1–41, https://doi.org/10.5194/tc-2022-226, 2022.

Echelmeyer, K. and Zhongxiang, W.: Direct Observation of Basal Sliding and Deformation of Basal Drift at Sub-Freezing Temperatures, Journal of Glaciology, 33, 83–98, https://doi.org/10.3189/s0022143000005396, 1987.

Flowers, G. E., Björnsson, H., and Pálsson, F.: New insights into the subglacial and periglacial hydrology of Vatnajökull, Iceland, from a distributed physical model, Journal of Glaciology, 49, 257–270, https://doi.org/10.3189/172756503781830827, 2003.

Fowler, A. C.: Sub-Temperate Basal Sliding, Journal of Glaciology, 32, 3–5, https://doi.org/10.3189/S0022143000006808, 1986.

Fowler, A. C. and Johnson, C.: Hydraulic run-away: a mechanism for thermally regulated surges of ice sheets, Journal of Glaciology, 41, 554–561, https://doi.org/10.3189/S002214300003478X, 1995.

Fowler, A. C. and Schiavi, E.: A theory of ice-sheet surges, Journal of Glaciology, 44, 104–118, https://doi.org/10.3189/s0022143000002409, 1998.

Gandy, N., Gregoire, L. J., Ely, J. C., Cornford, S. L., Clark, C. D., and Hodgson, D. M.: Exploring the ingredients required to successfully model the placement, generation, and evolution of ice streams in the British-Irish Ice Sheet, Quaternary Science Reviews, 223, 105 915, https://doi.org/10.1016/j.quascirev.2019.105915, 2019.

Greve, R., Takahama, R., and Calov, R.: Simulation of large-scale ice-sheet surges: The ISMIP HEINO experiments, Polar Meteorology and Glaciology, pp. 1–15, http://hdl.handle.net/2115/30205, 2006.

Hank, K.: Supplementary material for "Numerical issues in modeling thermally and hydraulically driven ice stream surge cycling", https://doi.org/10.5281/zenodo.7905404, 2023.

Hemming, S.: Heinrich events: Massive late Pleistocene detritus layers of the North Atlantic and their global climate imprint, Reviews of Geophysics - REV GEOPHYS, 42, https://doi.org/10.1029/2003RG000128, 2004.

Hindmarsh, R. C.: Consistent generation of ice-streams via thermo-viscous instabilities modulated by membrane stresses, Geophysical Research Letters, 36, 1–6, https://doi.org/10.1029/2008GL036877, 2009.

Joughin, I., Smith, B. E., Howat, I. M., Floricioiu, D., Alley, R. B., Truffer, M., and Fahnestock, M.: Seasonal to decadal scale variations in the surface velocity of Jakobshavn Isbrae, Greenland: Observation and model-based analysis, Journal of Geophysical Research: Earth Surface, 117, 1–20, https://doi.org/10.1029/2011JF002110, 2012.

Joughin, I., Smith, B. E., Shean, D. E., and Floricioiu, D.: Brief communication: Further summer speedup of Jakobshavn Isbræ, Cryosphere, 8, 209–214, https://doi.org/10.5194/tc-8-209-2014, 2014.

K.M. Cuffey and W.S.B. Paterson.: The Physics of Glaciers, Butterworth-Heinemann/Elsevier, Burlington, MA, 4th edn., 2010.

Kyrke-Smith, T. M., Katz, R. F., and Fowler, A. C.: Subglacial hydrology and the formation of ice streams, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 470, https://doi.org/10.1098/rspa.2013.0494, 2014.

MacAyeal, D. R.: Binge/purge oscillations of the Laurentide Ice Sheet as a cause of the North Atlantic's Heinrich events, Paleoceanography, 8, 775–784, https://doi.org/10.1029/93PA02200, 1993.

Mantelli, E., Bertagni, M. B., and Ridolfi, L.: Stochastic ice stream dynamics, Proceedings of the National Academy of Sciences, 113, E4594–E4600, https://doi.org/10.1073/pnas.1600362113, 2016.

Mantelli, E., Haseloff, M., and Schoof, C.: Ice sheet flow with thermally activated sliding. Part 1: the role of advection, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475, https://doi.org/10.1098/rspa.2019.0410, 2019.

McCarthy, C., Savage, H., and Nettles, M.: Temperature dependence of ice-on-rock friction at realistic glacier conditions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 375, https://doi.org/10.1098/rsta.2015.0348, 2017.

Payne, A. J. and Dongelmans, P. W.: Self-organization in the thermomechanical flow of ice sheets, Journal of Geophysical Research: Solid Earth, 102, 12 219–12 233, https://doi.org/10.1029/97jb00513, 1997.

Payne, A. J., Huybrechts, P., Abe-Ouchi, A., Calov, R., Fastook, J. L., Greve, R., Marshall, S. J., Marsiat, I., Ritz, C., Tarasov, L., and Thomassen, M. P.: Results from the EISMINT model intercomparison: The effects of thermomechanical coupling, Journal of Glaciology, 46, 227–238, https://doi.org/10.3189/172756500781832891, 2000.

PISM 2.0.6 documentation: PETSc options for PISM users, https://www.pism.io/docs/manual/practical-usage/petsc-options.html, 2023.

Pollard, D. and DeConto, R. M.: A Coupled Ice-Sheet/Ice-Shelf/Sediment Model Applied to a Marine-Margin Flowline: Forced and Unforced Variations, Glacial Sedimentary Processes and Products, pp. 37–52, https://doi.org/10.1002/9781444304435.ch4, 2007.

Pollard, D. and DeConto, R. M.: Description of a hybrid ice sheet-shelf model, and application to Antarctica, Geoscientific Model Development, 5, 1273–1295, https://doi.org/10.5194/gmd-5-1273-2012, 2012.

Robel, A. A., Degiuli, E., Schoof, C., and Tziperman, E.: Dynamics of ice stream temporal variability: Modes, scales, and hysteresis, Journal of Geophysical Research: Earth Surface, 118, 925–936, https://doi.org/10.1002/jgrf.20072, 2013.

Roberts, W. H., Payne, A. J., and Valdes, P. J.: The role of basal hydrology in the surging of the Laurentide Ice Sheet, Climate of the Past, 12, 1601–1617, https://doi.org/10.5194/cp-12-1601-2016, 2016.

Sayag, R. and Tziperman, E.: Interaction and variability of ice streams under a triple-valued sliding law and non-Newtonian rheology, Journal of Geophysical Research: Earth Surface, 116, https://doi.org/10.1029/2010JF001839, 2011.

Schannwell, C., Mikolajewicz, U., Ziemen, F., and Kapsch, M.-L.: Sensitivity of Heinrich-type ice-sheet surge characteristics to boundary forcing perturbations, Climate of the Past, 19, 179–198, https://doi.org/10.5194/cp-19-179-2023, 2023.

Shreve, R. L.: Glacier sliding at subfreezing temperatures., Journal of Glaciology, 30, 341–347, https://doi.org/10.1017/S0022143000006195, 1984.

Souček, O. and Martinec, Z.: ISMIP-HEINO experiment revisited: Effect of higher-order approximation and sensitivity study, Journal of Glaciology, 57, 1158–1170, https://doi.org/10.3189/002214311798843278, 2011.

Takahama, R.: Heinrich Event Intercomparison with the ice-sheet model SICOPOLIS, Master's thesis, http://hdl.handle.net/2115/28749, 2006.

Tarasov, L. and Peltier, W. R.: A high-resolution model of the 100 ka ice-age cycle, Annals of Glaciology, 25, 0–7, https://doi.org/10.3189/s026030550001380x, 1997.

Tarasov, L., Dyke, A. S., Neal, R. M., and Peltier, W. R.: A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling, Earth and Planetary Science Letters, 315-316, 30–40, https://doi.org/10.1016/j.epsl.2011.09.010, 2012.

Tarasov, L., Hank, K., and Lecavalier, B. S.: GSMv01.31.2023 code archive for LISsq experiments, https://doi.org/10.5281/zenodo.7668472, 2023.

Tulaczyk, S., Kamb, W. B., and Engelhardt, H. F.: Basal mechanics of Ice Stream B, west Antarctica: 2. Undrained plastic bed model, Journal of Geophysical Research: Solid Earth, 105, 483–494, https://doi.org/https://doi.org/10.1029/1999JB900328, 2000a.

Tulaczyk, S., Kamb, W. B., and Engelhardt, H. F.: Basal mechanics of Ice Stream B, West Antarctica 1. Till mechanics, Journal of Geophysical Research: Solid Earth, 105, 463–481, https://doi.org/10.1029/1999jb900329, 2000b.

Van Pelt, W. J. and Oerlemans, J.: Numerical simulations of cyclic behaviour in the Parallel Ice Sheet Model (PISM, Journal of Glaciology, 58, 347–360, https://doi.org/10.3189/2012JoG11J217, 2012.

Werder, M. A., Hewitt, I. J., Schoof, C. G., and Flowers, G. E.: Modeling channelized and distributed subglacial drainage in two dimensions, Journal of Geophysical Research: Earth Surface, 118, 2140–2158, https://doi.org/10.1002/jgrf.20146, 2013.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, http://www.the-cryosphere.net/5/715/2011/tc-5-715-2011.pdf, 2011.

Winsborrow, M. C., Clark, C. D., and Stokes, C. R.: What controls the location of ice streams?, Earth-Science Reviews, 103, 45–59, https://doi.org/10.1016/j.earscirev.2010.07.003, 2010.

Ziemen, F. A., Rodehacke, C. B., and Mikolajewicz, U.: Coupled ice sheet-climate modeling under glacial and pre-industrial boundary conditions, Climate of the Past, 10, 1817–1836, https://doi.org/10.5194/cp-10-1817-2014, 2014.

985 Ziemen, F. A., Kapsch, M. L., Klockmann, M., and Mikolajewicz, U.: Heinrich events show two-stage climate response in transient glacial simulations, Climate of the Past, 15, 153–168, https://doi.org/10.5194/cp-15-153-2019, 2019.

Supplement to "~~Numerical issues in modeling~~ Modeling sensitivities of thermally and hydraulically driven ice stream surge cycling"

Kevin Hank[1], Lev Tarasov[1], and Elisa Mantelli[2,3,4]

[1]Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL, A1B 3X7,Canada
[2]Department of Earth and Environmental Sciences , Ludwig-Maximillians-Universitaet Munich, Theresienstr. 41, 80333 Munich, Germany
[3]Alfred Wegener Institute for Polar and Marine Research, Am Alten Hafen 26, 27568 Bremerhaven, Germany
[4]Institute for Marine and Antarctic Studies, University of Tasmania, 20 Castray Esplanade, Battery Point TAS 7004, Australia
[*]khank@mun.ca

Correspondence: Kevin Hank (khank@mun.ca)

S1 GSM - Details of different model aspects

S1.1 Climate forcing

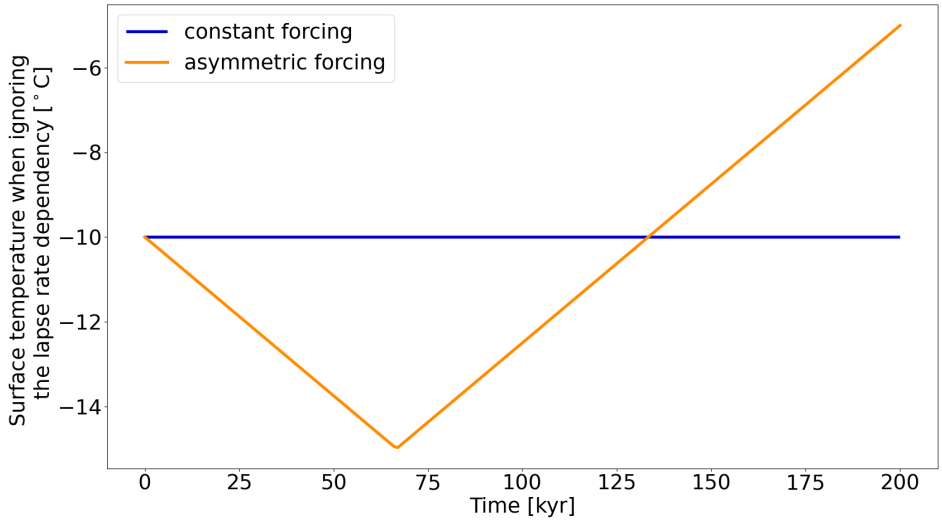


Figure S1. Constant and asymmetric temperature forcing in the GSM. The coldest temperature is reached at 66.7 kyr. For the case shown here, the surface temperature constant is set to rTsurf$= -10°$C (Table 1). All model runs within this paper use the asymmetric forcing.

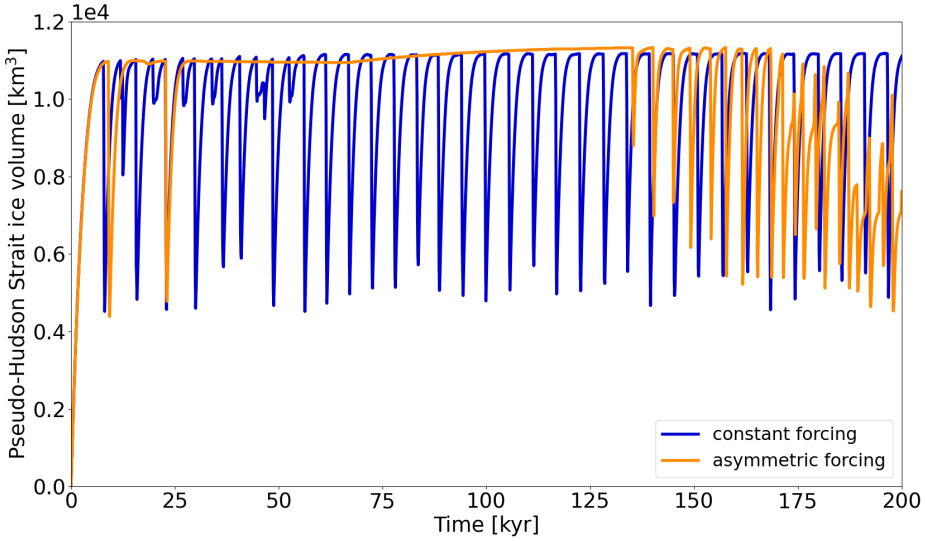


Figure S2. Pseudo-Hudson Strait ice volume for a constant and asymmetric temperature forcing in the GSM (Fig. S1). This plot shows parameter vector 1 with a horizontal grid resolution of 25 km.

S1.2 SSA activation velocities

Setup	number of surges	mean duration	mean period	mean pseudo-Hudson Strait ice volume change
reference setup	180 ± 100	1.1 ± 0.5 kyr	0.3 ± 0.1 kyr	$1.7 \pm 0.2 \cdot 10^3$ km^3
$v_{\mathrm{SIA,crit}} = 20$ m yr^{-1}	-3.7 ± 7.0	3.2 ± 6.4	1.5 ± 2.1	3.2 ± 2.4
$v_{\mathrm{SIA,crit}} = 40$ m yr^{-1}	-5.5 ± 5.4	6.1 ± 6.8	2.4 ± 5.7	3.5 ± 9.0
SSA everywhere	7.3 ± 24.8	1.7 ± 27.6	-9.3 ± 14.1	-17.7 ± 29.7

Table S1. Percentage differences of surge characteristics between the GSM reference setup (first row) and runs with different SSA activation velocities at 3.125 km. By default, the SSA is activated once the SIA velocity exceeds $v_{\mathrm{SIA,crit}} = 30$ m yr^{-1}. No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

S1.3 Parameter vectors

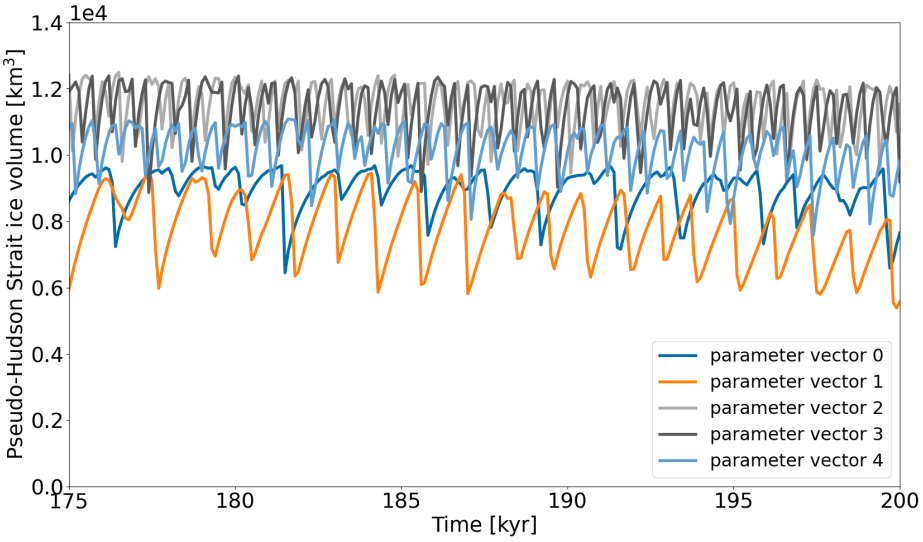


Figure S3. Pseudo-Hudson Strait ice volume for the last 25 kyr of all 5 GSM parameter vectors when using the reference setup. Note that only the last 25 kyr are shown for better visibility of the individual oscillation pattern.

S1.4 Bed properties

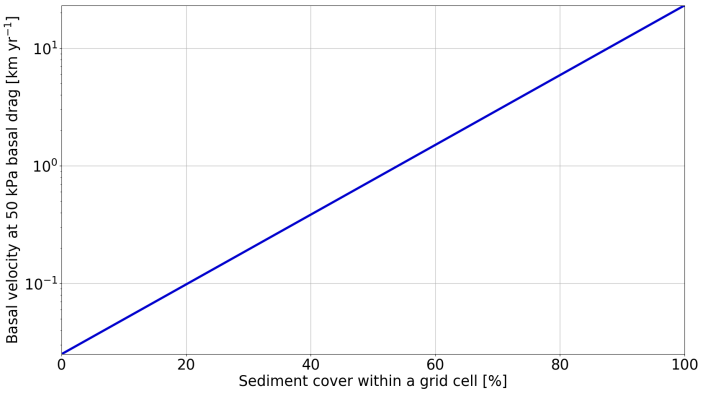
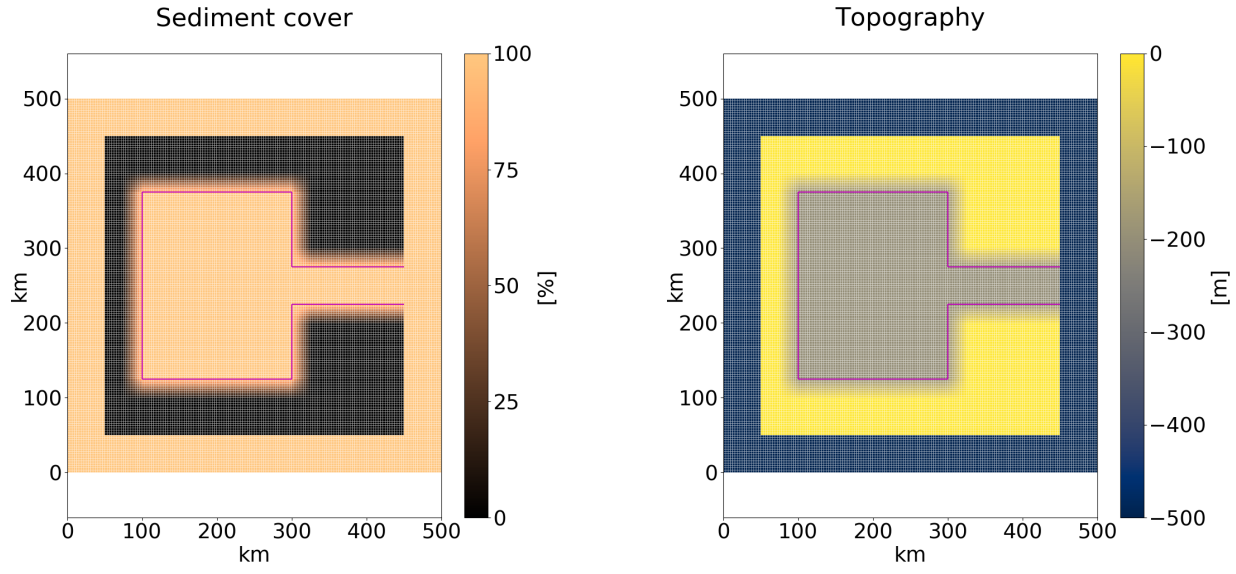


Figure S4. Basal velocity at 50 kPa basal drag for variable sediment cover and a power-law exponent of 3 (n_b in Table 1).



(a) (b)

Figure S5. Sediment cover and topography map for a 25 km wide transition zone at 3.125 km horizontal grid resolution. The transition zones for topography and sediment cover are at the same locations. The magenta line outlines the 100 % soft-bedded pseudo-Hudson Bay and Hudson Strait.

The effects of an abrupt transition from hard bedrock (0 % sediment cover) to soft sediment (100 % sediment cover) are examined by adding a smooth transition zone (Fig. S5 a)). Two widths of this transition zone (25 km and 3.125 km) are investigated. The basal velocity (or more precisely the sliding coefficient $\boldsymbol{C}$ in Eq. (6b)) then depends on the sediment cover within a grid cell (Fig. S4). In the experiments with a non-flat topography, the bed of the pseudo-Hudson Bay and Hudson Strait is placed 200 m and the surrounding ocean 500 m below the sea level (Fig. S5 b). The topographic transition zones (25 km and 3.125 km wide) align with the sediment transition zones.

S1.5 Weighting function of the adjacent minimum basal temperature

A weighting function takes into account the adjacent minimum basal temperature for the basal sliding temperature ramp.

$$T_{\mathrm{bp,I}} = W_{\mathrm{Tb,min}} \cdot \min\left[T_{\mathrm{bp,L}}, T_{\mathrm{bp,R}}\right] + T_{\mathrm{bp,I}} \cdot \left(1 - W_{\mathrm{Tb,min}}\right), \tag{S1}$$

where ~~$T_{\mathrm{bp,I}}$~~ $T_{\mathrm{bp,I}}$ is the basal temperature with respect to the pressure melting point at the grid cell interface, and ~~$T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$~~ $T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$ are the basal temperatures with respect to the pressure melting point at the adjacent grid cell centers. Note that $T_{\mathrm{bp,Im,L}}$ and $T_{\mathrm{bp,Im,R}}$ instead of ~~$T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$~~ $T_{\mathrm{bp,L}}$ and $T_{\mathrm{bp,R}}$ are used when calculating ~~$T_{\mathrm{bp,I}}$~~ $T_{\mathrm{bp,I}}$ according to TpmTrans (Eq. (18)). In this way, the additional heat T_{add} is still considered even when ~~$W_{\mathrm{Tb,min}} = 1$~~ $W_{\mathrm{Tb,min}} = 1$.

S2 PISM - Details of different model aspects

S2.1 Input fields

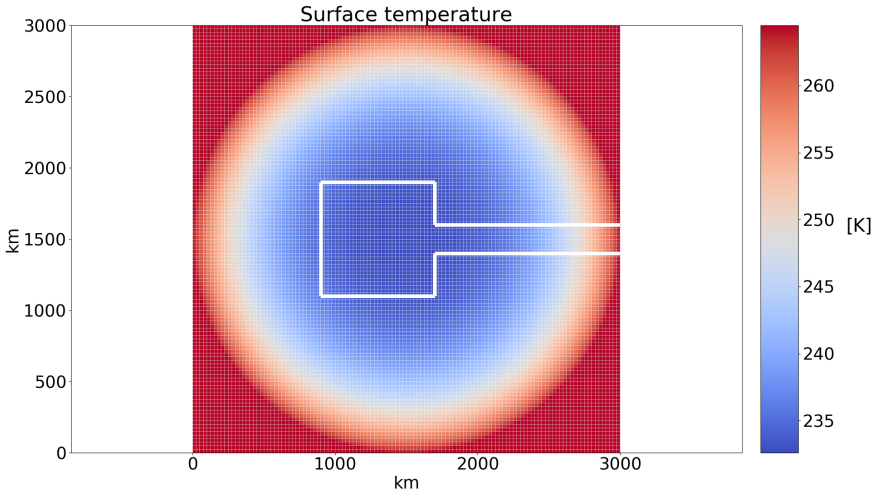


Figure S6. PISM surface temperature input field for parameter vector 1. The corresponding parameter values of ~~T_{min}~~ T_{min} and S_{t} are 232.60 K and $9.45 \cdot 10^{-9}$ K km^{-3}, respectively. Thick white lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution is 25x25 km.

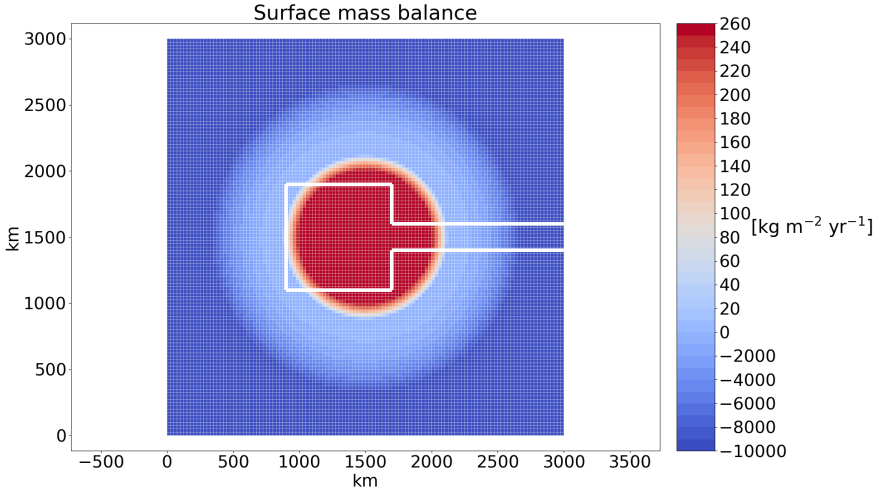


Figure S7. PISM surface mass balance input field for parameter vector 1. The corresponding parameter values of ~~B_{max}~~ B_{max} and S_{b} are 408.81 kg m^{-2} yr^{-1} and $4.55 \cdot 10^{-12}$ kg m^{-2} yr^{-1} km^{-5}, respectively. Thick white lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution is 25x25 km.

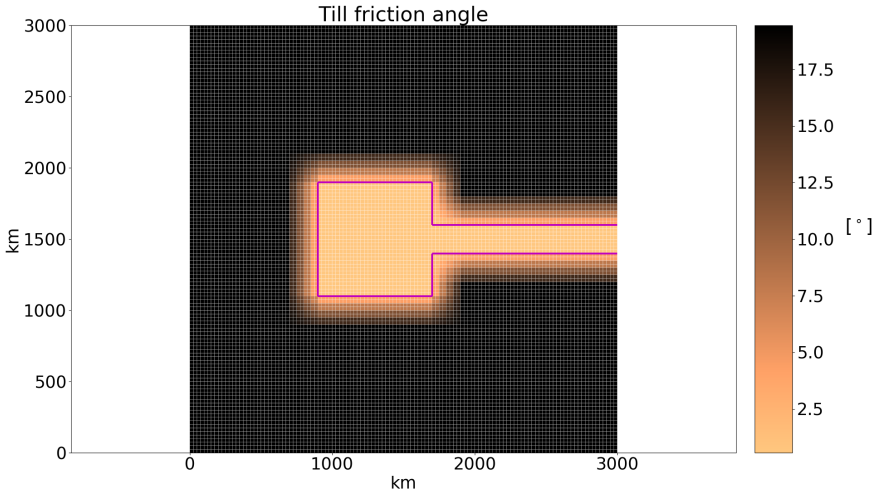


Figure S8. PISM till friction angle input field for parameter vector 1. The corresponding parameter values of $soft$ and $hard$ are $0.56°$C and $19.44°$C, respectively. Magenta lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution is 25x25 km.

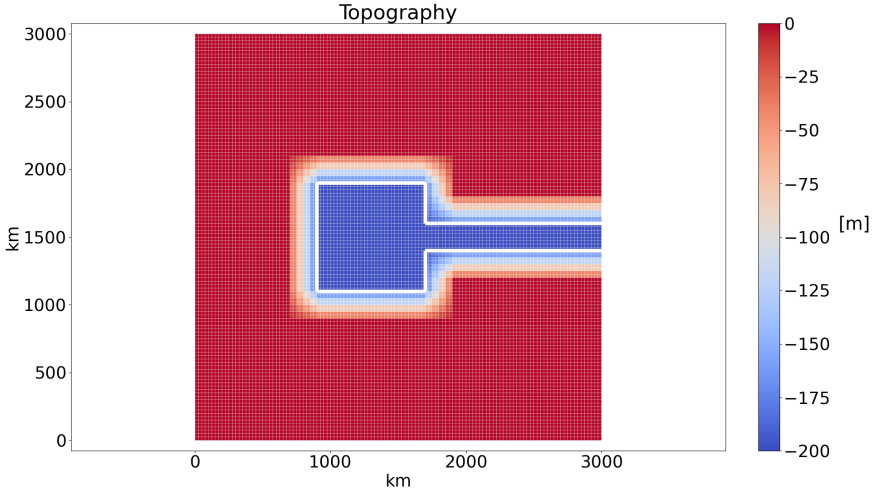


Figure S9. PISM topography input field (same for all parameter vectors). The white lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution is 25x25 km.

S2.2 Parameter vectors

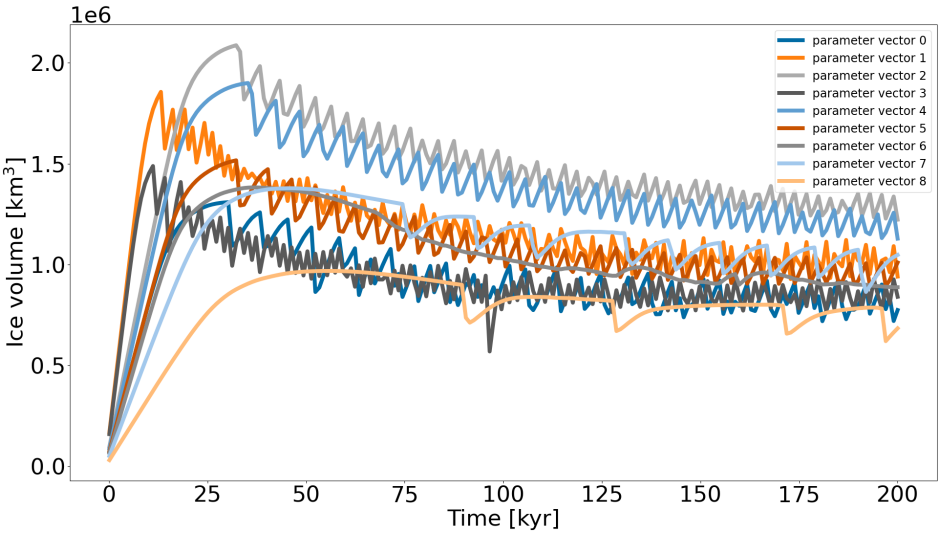


Figure S10. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for all 9 PISM parameter vectors when using the reference setup.

S2.3 Bed properties

In the PISM, oscillatory behavior only occurs for small yield stresses τ_c (Eq. (15)). This can be achieved by either a small till friction angle Φ or low effective pressure on the till (N_{till}) (Bueler and Van Pelt, 2015). N_{till} is given by

$$N_{\text{till}} = N_0 \left(\frac{\delta_e P_0}{N_0} \right)^s 10^{\left(\frac{e_0}{C_c}\right)(1-s)}, \tag{S2}$$

where $N_0 = 1$ kPa is the reference effective pressure, $e_0 = 0.69$ the void ratio at N_0, $C_c = 0.12$ the dimensionless coefficient of compressibility, δ_e the effective fraction of the overburden pressure, P_0 the ice overburden pressure, and s the ratio $\frac{W_{\text{till}}}{W_{\text{till}}^{\max}}$ (Tulaczyk et al., 2000; Bueler and Van Pelt, 2015). W_{till} and $W_{\text{till}}^{\max} = 2$ m are the effective and maximum thickness of water in the till, respectively. The values listed here are the PISM defaults. C_c is on the lower end of measured values (Tulaczyk et al., 2000) with significantly larger (up to 17) values reported (Sauer et al., 1993; Mitchell and Soga, 2005). e_0 can vary between 0.45 (Tulaczyk et al., 2000) and approximately 4 (Fig. 10.2 in Mitchell and Soga, 2005). The default value of δ_e is based on Greenland and Antarctic model runs, but δ_e is generally considered as a tuning parameter to match observed surface velocities, which are not available in a paleo context (Andy Aschwanden, personal communication).

When only changing the till friction angle (Eq. (15)), oscillations do not occur unless $\Phi < 1°$ (Fig. S13). This is well below the measured values of about 10 to 40° (K.M. Cuffey and W.S.B. Paterson., 2010). However, similar oscillatory results are obtained for till friction angles between 5 and 10° when slightly adjusting the values of $C_c = 0.2$, $e_0 = 0.6$, and $\delta_e = 0.01$ to favor sliding (compare Fig. S11 and S12). These values are all well within the ranges set by laboratory measurements.

S2.4 Maximum magnitude of basal ice velocity

Small till friction angles (0.5 to $1.0°$) lead to slippery beds and high maximum basal sliding velocities (up to ~ 600 km yr^{-1}) for a small number of time steps in some runs. A maximum of 7 out of 2000 time steps exceeds 50 km yr^{-1} (parameter vector 1 in Fig. S11). While observed velocities can reach several hundreds of meters per day for short periods (K.M. Cuffey and W.S.B. Paterson. (2010), e.g., 300 m d$^{-1} = 109.5$ km yr^{-1}), high modeled velocities might lead to instabilities in the numerical matrix solver. Therefore, we set an upper limit of 40 km yr^{-1} for the SSA velocity.

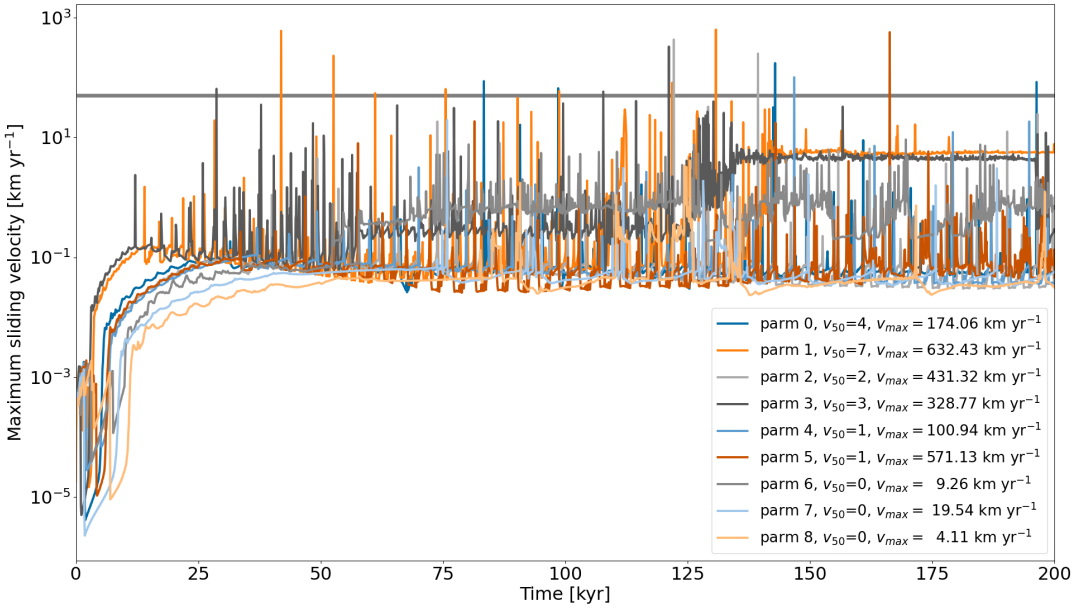


Figure S11. Maximum sliding velocity (max(max(abs($\boldsymbol{u}$)), max(abs($\boldsymbol{v}$)))) at each time step (100 yr interval) within the whole model domain for all 9 parameter vectors using PISM without an upper limit for the SSA velocity. The black horizontal line marks 50 km yr^{-1} and v_{50} indicates the number of time steps exceeding this velocity. $v_{\max}$ is the highest maximum sliding velocity in a run.

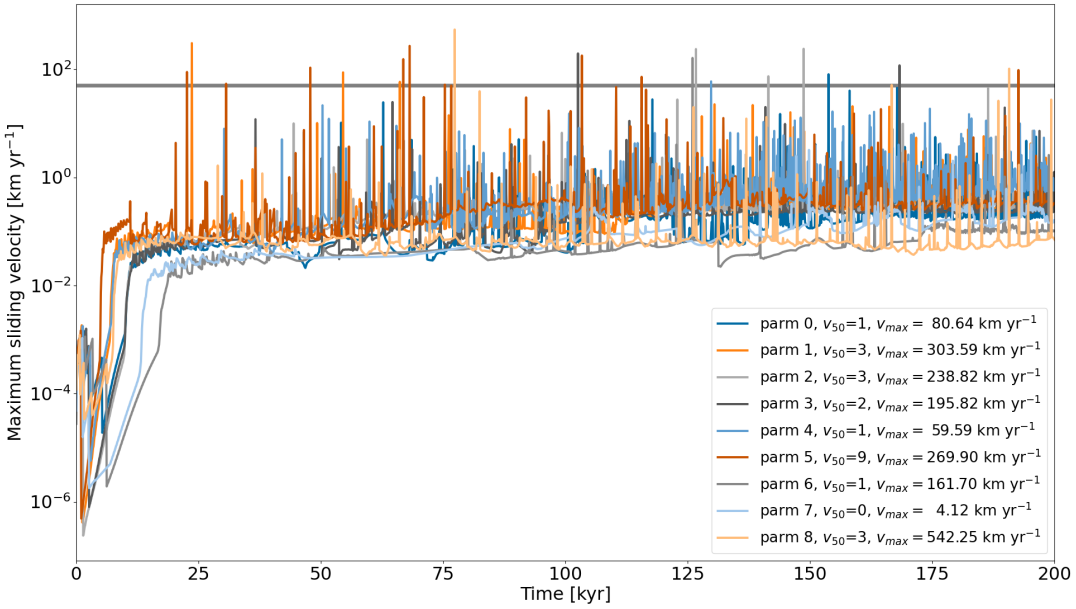


Figure S12. Maximum sliding velocity (max(max(abs($\boldsymbol{u}$)), max(abs($\boldsymbol{v}$)))) at each time step (100 yr interval) within the whole model domain for 9 parameter vectors with till friction angles between 5 and 10° and values of $C_c = 0.2$, $e_0 = 0.6$, and $\delta_e = 0.01$ using PISM without an upper limit for the SSA velocity. The black horizontal line marks 50 km yr^{-1} and v_{50} indicates the number of time steps exceeding this velocity. v_{max} is the highest maximum sliding velocity in a run.

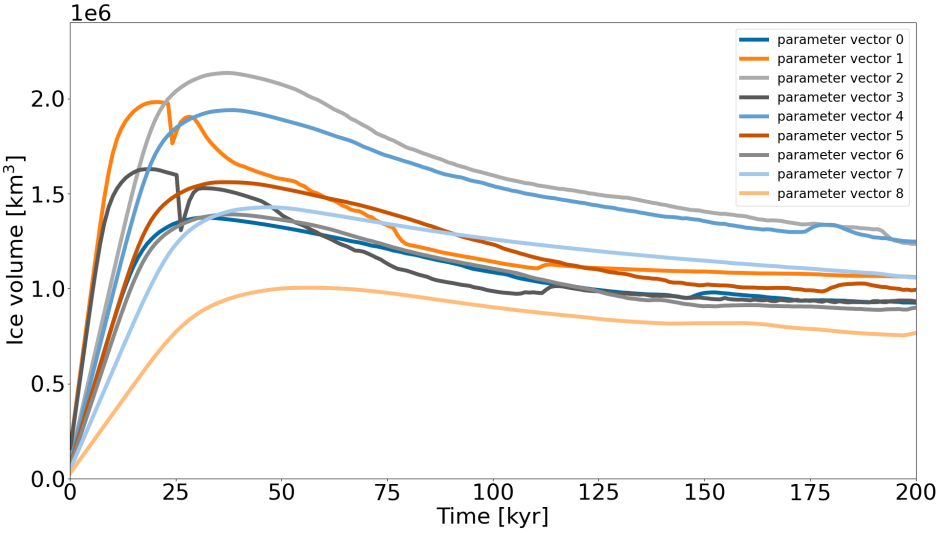


Figure S13. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for all 9 PISM parameter vectors when using the reference setup but a soft bed till friction angle of 1°.

S2.5 Ice volume - pseudo-Hudson Strait vs. surge-affected area

During a surge, ice from the pseudo-Hudson Bay and areas surrounding the pseudo-Hudson Strait is rapidly transported into the mostly ice-free pseudo-Hudson Strait. Consequently, the ice sheet extends further to the East (increasingly stronger melting), covering almost the entire pseudo-Hudson Strait area. Due to the complex interaction between ice transport and melting area, times of minimum ice volume over the area most affected by the pseudo-Hudson Strait surge (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait, e.g., Fig. 5 and video 06 of Hank (2023)) correspond to maxima in the pseudo-Hudson Strait ice volume for most surges (grey lines in Fig. S14). However, some ice volume minima do not align with a maximum of the pseudo-Hudson Strait ice volume (red lines in Fig. S14). This inconsistency hampers the detection of surges when using the pseudo-Hudson Strait ice volume and can lead to flawed statistics. To avoid this issue, we use the ice volume in the surge-affected area, for which surges appear as minima, for all PISM results. A comparison between PISM results based on the pseudo-Hudson Strait and the surge-affected area is shown in Fig. S15. Note that we only consider the eastern half of the pseudo-Hudson Bay because some runs also show surges on the Western side of the ice sheet (e.g., $50\ \mathrm{km}$ run in video 09 of Hank (2023)).

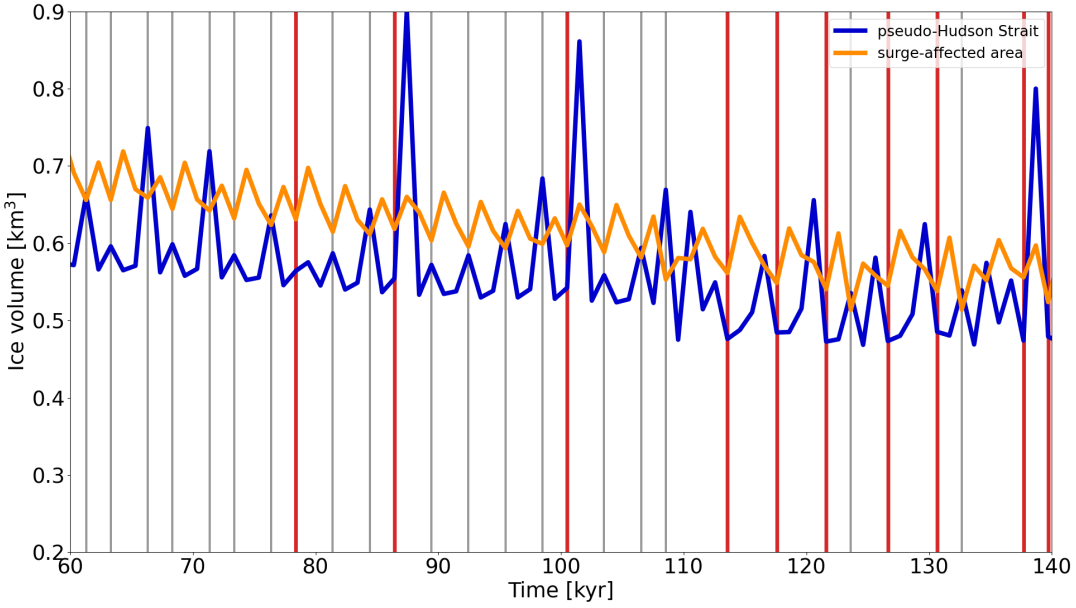


Figure S14. Normalized pseudo-Hudson Strait and surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) ice volume for parameter vector 1 using PISM. For most surges, the minimum ice volume over the surge-affected area aligns with a maxima in the pseudo-Hudson Strait ice volume (grey lines). This is, however, not true for all surges (thick red lines) and can lead to flawed statistics. See also video 06 of Hank (2023).

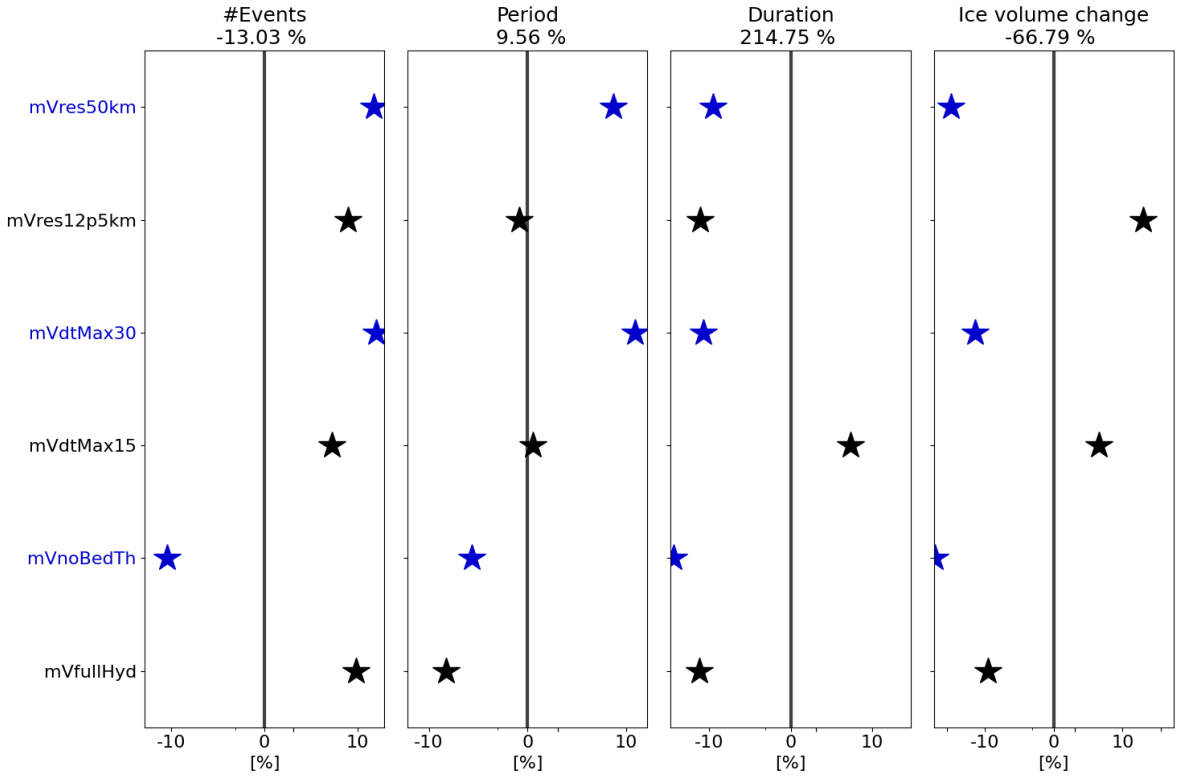


Figure S15. Differences in the percentage change of the mean surge characteristics (between comparison and reference setup) when using the ice volume of the surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) compared to only the pseudo-Hudson Strait ice volume. A positive difference indicates a larger change for the analysis based on the pseudo-Hudson Strait ice volume. The different colors were added for visual alignment of the individual model setups, the stars are the ensemble mean differences, and the horizontal bars represent the ensemble standard deviations. The percentages in the titles of each subplot represent the differences in the surge characteristics of the reference runs. For example, the mean number of surges based on the pseudo-Hudson Strait ice volume is ~ 13 % smaller than for the ice volume of the surge-affected area. The different colors resemble different model setups. Note that the surge threshold is $4 \cdot 10^4$ km^3 when using the surge-affected area ice volume and $0.5 \cdot 10^4$ km^3 for the pseudo-Hudson Strait ice volume (~ 5 % of mean ice volume across all runs). The x-axis is logarithmic.

S3 Run analysis approach

For both models, we use the Python module *scipy* (version 1.5.2 on GSM cluster and 1.7.0 on PISM cluster, different versions due to the availability on computational clusters) and its built-in function *scipy.signal.find_peaks* on the ice volume output to determine the surge characteristics. The surge duration and ice volume change during a surge are determined by the functions *scipy.signal.peak_widths* and *scipy.signal.peak_prominences*, respectively. The Python analysis scripts are provided as supplementary material.

Temporal resolution of output time series

The standard output time steps in the GSM and PISM are 0.1 and 1 kyr, respectively. Note that these time steps might not exactly capture the minimum ice volume but are generally a good compromise between storage requirements and temporal resolution (e.g., Fig. S16 and S17).

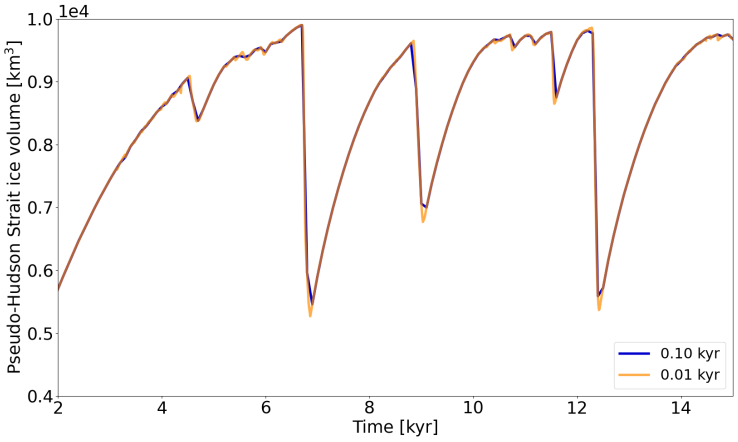


Figure S16. Pseudo-Hudson Strait ice volume of a GSM model run with different output time steps. The horizontal grid resolution is 3.125 km.

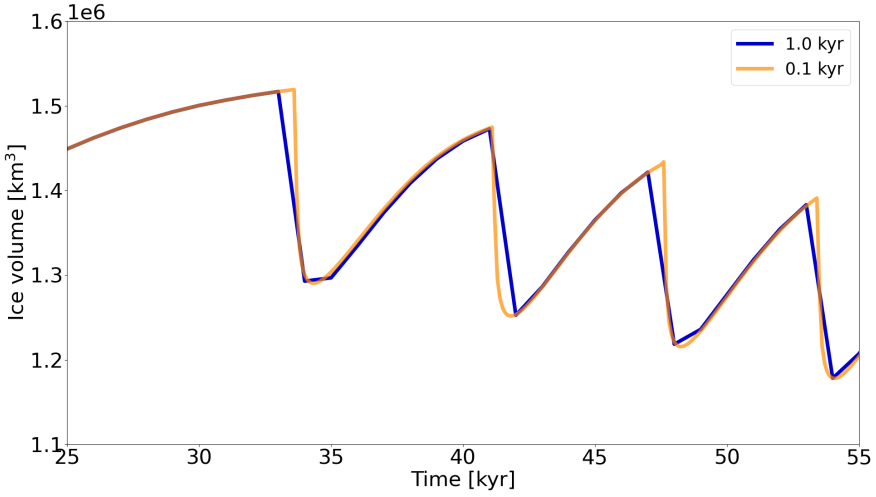


Figure S17. Ice volume in the surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) of a PISM model run (parameter vector 5) with different output time steps. The horizontal grid resolution is 25 km.

S4 RMSE and mean bias

The RMSE and mean bias values presented throughout the paper are calculated according to the following equations

$$rmse = \sqrt{\frac{\sum_{\mathrm{t}=0}^{t_{\max}} (x_{\mathrm{t}} - b_{\mathrm{t}})^2}{N}} \cdot \frac{100}{b_{\mathrm{m}}}, \text{ and} \tag{S3}$$

$$mean\ bias = \frac{\sum_{\mathrm{t}=0}^{t_{\max}} (x_{\mathrm{t}} - b_{\mathrm{t}})}{N} \cdot \frac{100}{b_{\mathrm{m}}}, \tag{S4}$$

where x_{t} and b_{t} are the (pseudo-Hudson Strait) ice volume values at time t of the comparison setup and reference setup, respectively. $t_{\max}$ is the maximum time, N the number of time steps, and b_{m} the mean of the reference setup time series. These values are then averaged over all 5 parameter vectors. Crashed runs are excluded from the averaging process.

S5 Comparison between different model setups

The analysis to compare the different model setups follows

1. run 1 parameter vector with the reference setup (Table 2)
2. calculate the surge characteristics for this reference run (s_{ref})
3. re-run the same parameter vector for one of the comparison setups (Sec. 2.1.3 and 2.2.4)
4. calculate the surge characteristics for the comparison run (s_{comp})
5. calculate the differences in surge characteristics between the reference run and comparison run expressed as percentage differences from the reference run (positive for increase compared to the value of the reference run): $p = \frac{s_{\mathrm{comp}} - s_{\mathrm{ref}}}{s_{\mathrm{ref}}} \cdot 100$
6. repeat steps 1) to 5) for all parameter vectors (5 for the GSM, 9 for PISM)
7. the values shown in the tables and figures are the mean and standard deviation of all percentage differences for each surge characteristic

Percentage differences for crashed comparison runs are not considered for the final average and runs with less than 2 surges require special treatment. In these cases, the period is set to a NaN value, leading to a NaN difference between that particular run and the corresponding reference run. We use Numpys *numpy.nanmean()* and *numpy.nanstd()* to ignore these NaN values when averaging over all parameter vectors. Similarly, all surge characteristics except for the number of surges are set to NaN values for runs with no surges at all. Note that the values for the reference setup stated in the tables are the mean and standard deviation of the actual surge characteristics of all reference runs, not percentage differences.

S6 Minimum numerical error estimates

S6.1 GSM

S6.1.1 Minimum numerical error estimates at 12.5 km

Metric	original 12.5 km runs	stricter numerical convergence [% difference]	stricter numerical convergence with increased maximum iterations [% difference]
#Surges	81 ± 42	2.3 ± 8.5	$\mathbf{2.8 \pm 9.2}$
mean period	2.3 ± 0.8 kyr	$\mathbf{-2.1 \pm 7.5}$	-1.5 ± 9.3
mean duration	0.6 ± 0.2 kyr	-1.4 ± 9.4	$\mathbf{2.6 \pm 14.2}$
mean pseudo-Hudson Strait ice volume change	$2.2 \pm 1.1 \cdot 10^3$ km^3	$\mathbf{20.9 \pm 53.0}$	-5.1 ± 12.2

Table S2. Percentage differences of surge characteristics between GSM runs with regular and stricter numerical convergence and increased maximum iterations for the ice dynamics loops at 12.5 km (except first column). The values represent the average of 5 parameter vectors. No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. The bold numbers mark the largest MNEE for each surge characteristic.

S6.1.2 Adding surface temperature noise

Metric	reference setup	± 0.1°C noise	± 0.5°C noise
#Surges	180 ± 100	-4.0 ± 4.3	-4.1 ± 7.0
mean period	1.1 ± 0.5 kyr	4.8 ± 5.3	3.8 ± 6.8
mean duration	0.3 ± 0.1 kyr	1.3 ± 4.4	0.9 ± 4.3
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	0.9 ± 4.1	2.1 ± 5.5
RMSE	-	8.0 ± 2.5	7.8 ± 2.1
Mean Bias	-	-0.1 ± 0.2	0.1 ± 0.0

Table S3. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup for two different amplitudes of surface temperature noise (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S6.1.3 Implicit thermodynamics/ice dynamics coupling

Metric	reference setup	implicit coupling
#Surges	180 ± 100	1.1 ± 4.9
mean period	1.1 ± 0.5 kyr	-0.3 ± 5.3
mean duration	0.3 ± 0.1 kyr	-12.7 ± 9.5
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	-25.1 ± 18.7
RMSE	-	7.3 ± 2.5
Mean Bias	-	1.8 ± 1.5

Table S4. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup for implicit coupling between the thermodynamics and ice dynamics in the GSM (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S6.2 PISM

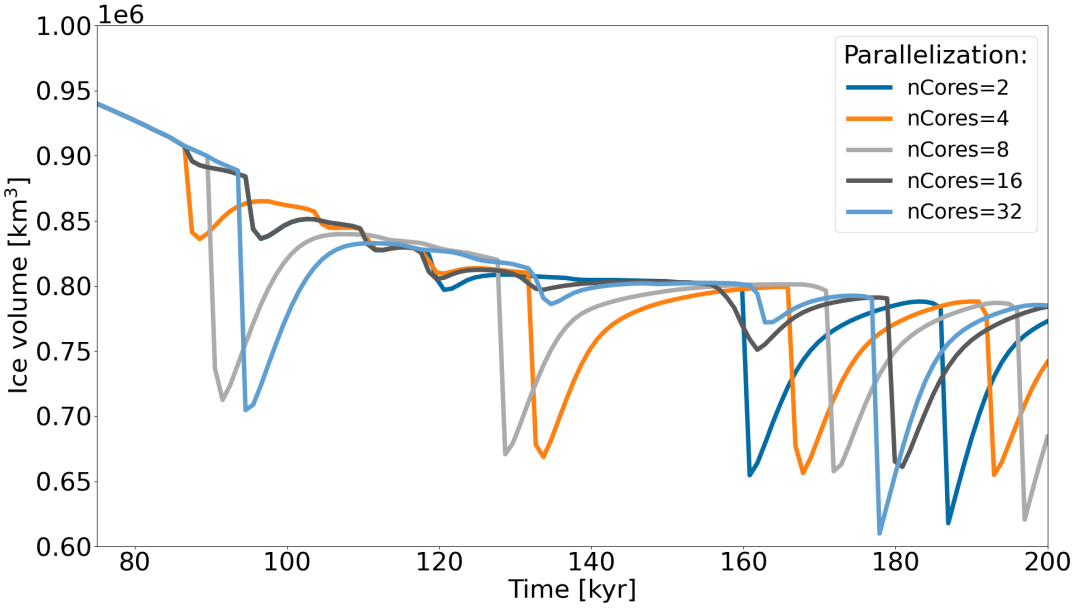


Figure S18. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 8 and different numbers of cores/processes using the PISM.

S6.2.1 Relative tolerances

Setup	number of surges	mean period	mean duration	mean ice volume change	nC	nS1
25 km reference setup	35 ± 25	10 ± 10 kyr	3 ± 2 kyr	$1.1 \pm 0.3 \cdot 10^5$ km^3	0	0
nCores$= 8$, KSP$= 10^{-10}$	-10.7 ± 21.6	-1.8 ± 5.6	11.2 ± 32.8	5.3 ± 11.2	0	1
nCores$= 16$, KSP$= 10^{-10}$	-4.7 ± 22.4	-2.7 ± 4.8	0.1 ± 9.3	-0.4 ± 5.1	0	1
nCores$= 32$, KSP$= 10^{-10}$	-10.5 ± 28.7	-3.7 ± 6.2	-1.2 ± 12.4	1.2 ± 7.2	4	1
nCores$= 8$, KSP$= 10^{-10}$, PIC$= 10^{-7}$	8.3 ± 11.8	-28.6 ± 17.4	-17.1 ± 35.2	9.0 ± 21.5	6	0
nCores$= 16$, KSP$= 10^{-10}$, PIC$= 10^{-7}$	98.1 ± 135.9	-22.5 ± 10.1	-14.4 ± 24.6	15.7 ± 54.4	5	0
nCores$= 32$, KSP$= 10^{-10}$, PIC$= 10^{-7}$	83.8 ± 125.1	-18.2 ± 11.6	-13.9 ± 25.7	23.2 ± 56.4	5	0

Table S5. Percentage differences of surge characteristics compared to the PISM reference setup with different numbers of cores and adjusted relative tolerances for the Picard iteration in the calculation of the vertically-averaged effective viscosity (PIC, default is 10^{-4}) and the Krylov linear solver used at each Picard iteration (KSP, default is 10^{-7} (except first row). The values represent the average of 9 parameter vectors. Crashed runs (nC) are not considered and runs with just one surge (nS1) are ignored when calculating the change in mean period. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. Note that more than 50 % of all runs with KSP$= 10^{-10}$ and PIC$= 10^{-7}$ did not finish within the time limit set by the computational cluster and are considered as crashed runs (nC). A direct comparison of runs with these tolerances can be found in Fig. S19. Note that all test runs without preconditioning (removes processor-number-dependence of results) crashed during the spin-up phase and long before the first surge occurs.

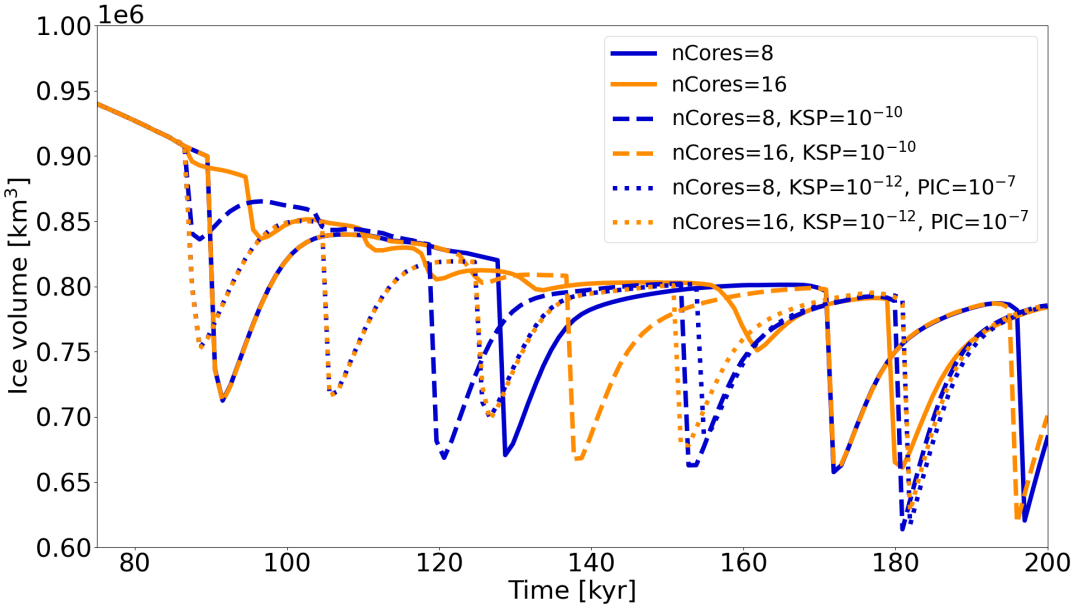


Figure S19. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 8 and different number of cores/processes using the PISM with different relative tolerances for the Picard iteration in the calculation of the vertically-averaged effective viscosity (PIC, default is 10^{-4}) and the Krylov linear solver used at each Picard iteration (KSP, default is 10^{-7}.

S6.2.2 Adding surface temperature noise

Metric	reference setup	± 0.1°C noise	± 0.5°C noise
#Surges	35 ± 25	-12.4 ± 24.2	-12.0 ± 25.2
mean period	10 ± 10 kyr	-5.6 ± 8.0	-4.0 ± 8.8
mean duration	3 ± 2 kyr	11.5 ± 37.9	2.0 ± 12.9
mean pseudo-Hudson Strait ice volume change	$1.1 \pm 0.3 \cdot 10^5$ km^3	1.9 ± 15.9	2.5 ± 8.1
RMSE	-	4.1 ± 3.5	4.3 ± 2.6
Mean Bias	-	-0.8 ± 1.9	0.1 ± 0.3
nS1	0	1	1

Table S6. Percentage differences of surge characteristics, ice volume RMSE and mean bias compared to the PISM reference setup for two different amplitudes of surface temperature noise (except first column). No runs crashed and all runs showed at least 1 surge. Runs with just one surge (nS1) are ignored when caluclating the change in mean period. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S7 Sensitivity experiments with a significant effect

S7.1 Bed thermal model

Metric	reference setup	20 m deep (1 layer) bed thermal model
number of surges	180 ± 100	-31.6 ± 5.6
mean period	1.1 ± 0.5 kyr	60.2 ± 22.4
mean duration	0.3 ± 0.1 kyr	65.2 ± 24.5
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	49.6 ± 14.6
RMSE	-	10.4 ± 2.2
Mean Bias	-	-2.3 ± 1.7

Table S7. Percentage differences (except first column) of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup for runs with only one bed thermal layer (20 m deep). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

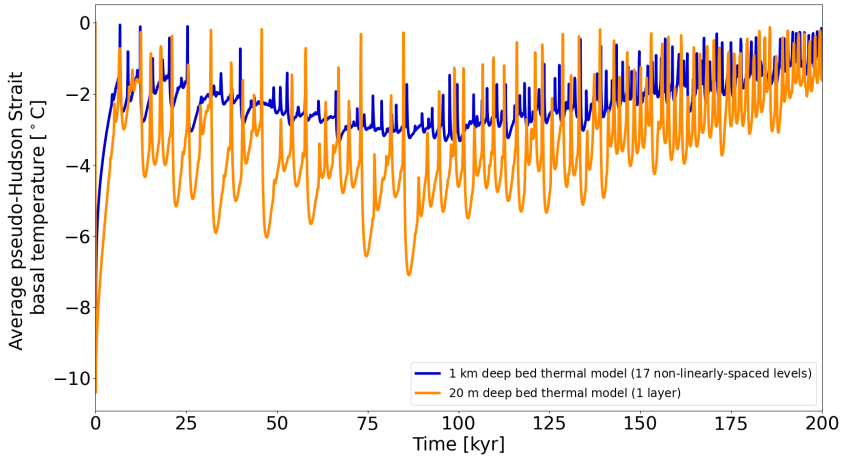


Figure S20. Average pseudo-Hudson Strait basal ice temperature with respect to the pressure melting point for parameter vector 1 with a 20 m and 1 km deep bed thermal model (17 non-linearly-spaced levels) using the GSM. The horizontal grid resolution is 3.125 km.

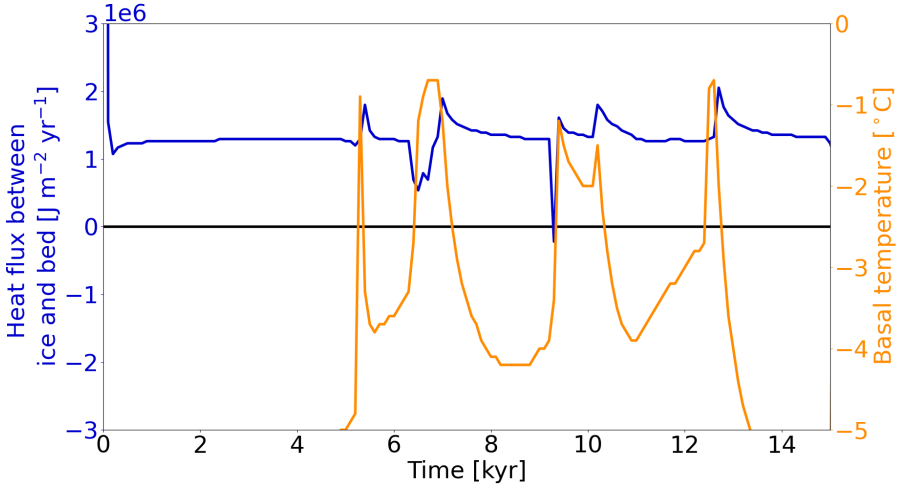


Figure S21. Heat flux at the base of the ice sheet (positive from bed into ice) and basal ice temperature for a grid cell in the center of the pseudo-Hudson Strait (grid cell center at $x = 376.5625$ km and $y = 248.4375$ km, white star in Fig. 1) and parameter vector 1 with only one bed thermal layer (20 m deep) using the GSM. The horizontal grid resolution is 3.125 km.

Metric	reference setup	no bed thermal model
number of surges	35 ± 25	20.7 ± 140.5
mean period	10 ± 10 kyr	79.9 ± 80.0
mean duration	3 ± 2 kyr	69.8 ± 60.4
mean ice volume change	$1.1 \pm 0.3 \cdot 10^5$ km^3	395.8 ± 240.5
RMSE	-	36.0 ± 5.3
Mean Bias	-	-27.1 ± 5.6

Table S8. Percentage differences of surge characteristics, ice volume RMSE and mean bias compared to the PISM reference setup for runs without a bed thermal model (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

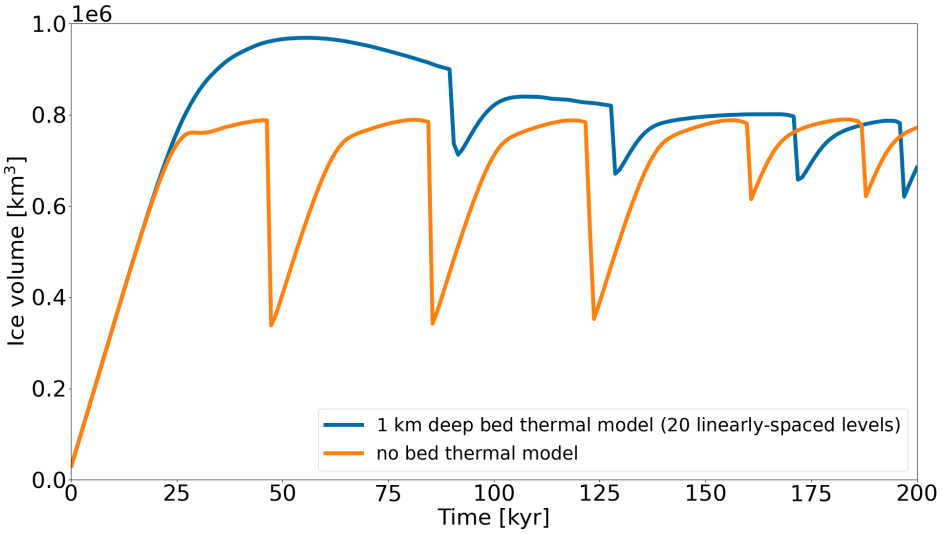


Figure S22. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 8 with and without the 1 km deep (20 linearly-spaced levels) bed thermal model using the PISM. The horizontal grid resolution is 25 km.

S7.2 Basal temperature at the grid cell interface

Metric	reference setup (TpmTrans)	TpmInt	TpmInt, upwind	TpmCen
nC	0	0	0	1
nS0	0	4	1	0
nS1	0	0	1	0
number of surges	180 ± 100	-96.9 ± 6.3	-90.2 ± 15.4	-74.6 ± 13.9
mean period	1.1 ± 0.5 kyr	106.2 ± 0.0	1645.4 ± 2136.8	609.4 ± 832.22
mean duration	0.3 ± 0.1 kyr	-15.9 ± 0.0	11.1 ± 17.4	43.3 ± 71.1
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	-66.2 ± 0.0	-60.4 ± 6.5	-61.3 ± 5.6
RMSE	-	7.4 ± 2.4	9.4 ± 2.6	6.9 ± 2.5
Mean Bias	-	4.0 ± 1.6	6.7 ± 2.4	2.1 ± 2.1

Table S9. Percentage differences (except first column) of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup for different approaches to calculate the basal temperature at the grid cell interface (Sec. 3.3.2). Crashed runs (nC) are not considered and runs without surges (nS0) only contribute to the change in surge number. Runs with only 1 surge (nS1) are excluded from the calculation of the mean period. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S7.3 Basal temperature ramps at different resolutions

To simplify the comparison of different temperature ramps, we calculate a single value score based on all surge characteristics. The calculation steps are as follows.

1. calculate the absolute values for all surge characteristic means
2. calculate the average across all ramps for all characteristics (means and standard deviations separately, total of 4 means and 4 standard deviations)
3. for each ramp, divide all surge characteristics by their corresponding average
4. sum the values for all surge characteristics (separately for mean and std)

The above calculation combines the 4 surge characteristics to a single value for the mean and standard deviation of each ramp. We keep separate values for the mean and standard deviation since the two metrics contain different information. Smaller values indicate a better agreement with the 3.125 km reference setup.

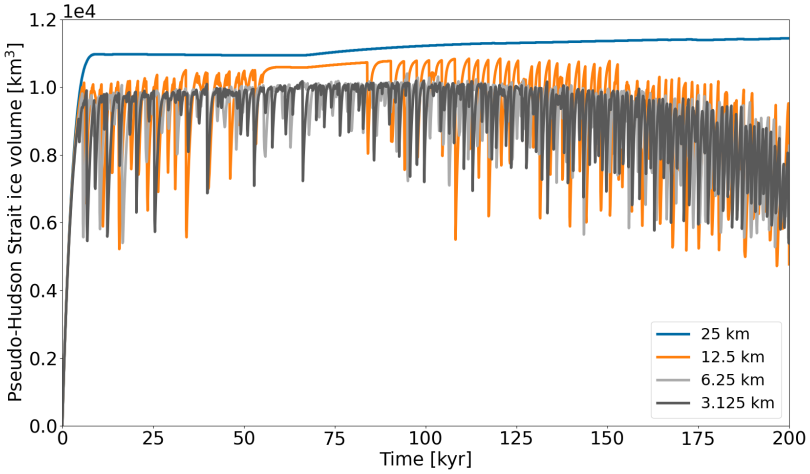


Figure S23. Pseudo-Hudson Strait ice volume for parameter vector 1 and different horizontal grid resolutions using the GSM. A constant temperature ramp with ~~$T_{\text{ramp}} = 0.0625°$~~ $T_{\text{ramp}} = 0.0625°$C and ~~$T_{\text{exp}} = 28$~~ $T_{\text{exp}} = 28$ is used for all horizontal grid resolutions (magenta line in Fig. 2).

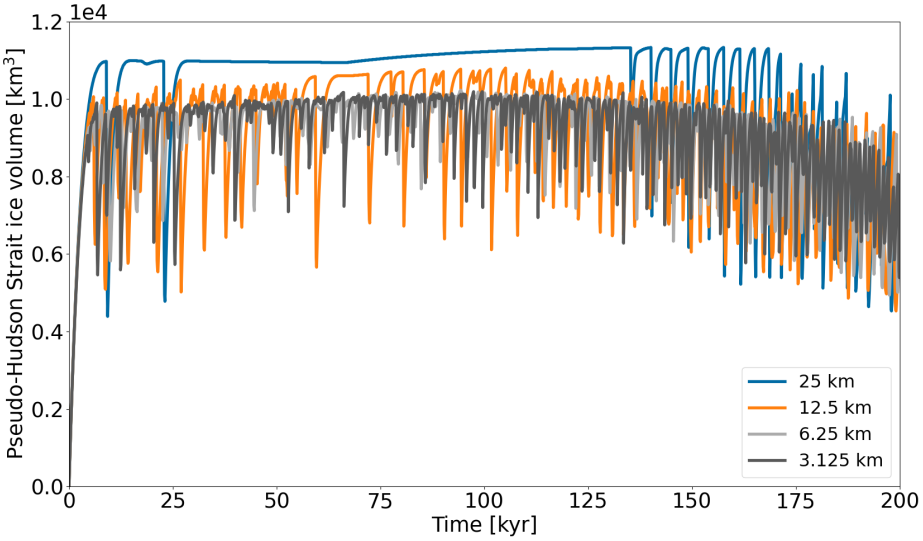


Figure S24. Pseudo-Hudson Strait ice volume for parameter vector 1 and different horizontal grid resolutions using the GSM. A resolution-dependent temperature (Eq. (9)) with ~~$P_{\mathrm{T_{ramp}}} = 1$~~ $P_{\mathrm{T_{ramp}}} = 1$ and ~~$T_{\mathrm{exp}} = 28$~~ $T_{\mathrm{exp}} = 28$ is used for all horizontal grid resolutions (matching colors in Fig. 2).

Metric	#Surges	mean period	mean duration	mean pseudo-Hudson Strait ice volume change	RMSE	Mean Bias
~~$T_{exp} = 5, T_{ramp} = 1$~~ $T_{exp} = 5, T_{ramp} = 1$	-63.5 ± 17.1	121.7 ± 29.5	300.0 ± 116.6	95.5 ± 39.5	21.8 ± 4.8	-17.8 ± 5.6
~~$T_{exp} = 15, T_{ramp} = 1$~~ $T_{exp} = 15, T_{ramp} = 1$	-39.0 ± 10.2	64.6 ± 22.8	179.2 ± 117.3	51.5 ± 35.3	17.4 ± 3.7	-11.2 ± 4.6
~~$T_{exp} = 28, T_{ramp} = 0.5$~~ $T_{exp} = 28, T_{ramp} = 0.5$	-17.1 ± 7.1	28.6 ± 21.0	64.0 ± 54.1	18.7 ± 12.4	10.0 ± 3.3	-3.5 ± 3.2
~~$T_{exp} = 5, T_{ramp} = 0.0625$~~ $T_{exp} = 5, T_{ramp} = 0.0625$	-9.5 ± 5.1	16.5 ± 12.4	14.9 ± 12.0	3.9 ± 2.9	8.1 ± 2.4	-0.8 ± 0.6
~~$T_{exp} = 10, T_{ramp} = 0.0625$~~ $T_{exp} = 10, T_{ramp} = 0.0625$	-9.3 ± 5.0	10.1 ± 5.0	8.8 ± 7.0	3.4 ± 4.0	8.0 ± 2.4	-0.4 ± 0.3
~~$T_{exp} = 28, T_{ramp} = 0.125$~~ $T_{exp} = 28, T_{ramp} = 0.125$	-4.6 ± 6.4	3.2 ± 4.2	4.4 ± 5.4	0.3 ± 2.1	7.9 ± 2.2	-0.3 ± 0.2
~~$T_{exp} = 14, T_{ramp} = 0.0625$~~ $T_{exp} = 14, T_{ramp} = 0.0625$	-7.1 ± 5.1	9.3 ± 7.4	7.3 ± 7.5	2.8 ± 3.4	7.8 ± 1.9	-0.2 ± 0.1
~~$T_{exp} = 15, T_{ramp} = 0.0625$~~ $T_{exp} = 15, T_{ramp} = 0.0625$	-4.9 ± 4.7	8.4 ± 10.5	4.8 ± 4.4	0.3 ± 6.3	7.8 ± 2.0	-0.2 ± 0.1
~~$T_{exp} = 20, T_{ramp} = 0.0625$~~ $T_{exp} = 20, T_{ramp} = 0.0625$	-3.0 ± 4.7	2.0 ± 3.9	-0.1 ± 2.9	1.3 ± 4.2	7.9 ± 2.4	-0.1 ± 0.1
~~$T_{exp} = 25, T_{ramp} = 0.0625$~~ $T_{exp} = 25, T_{ramp} = 0.0625$	-1.2 ± 3.5	4.1 ± 7.7	0.5 ± 1.1	-1.5 ± 3.0	7.8 ± 2.4	-0.0 ± 0.1
3.125 km reference setup	**180 ± 100**	**1.1 ± 0.5 kyr**	**0.3 ± 0.1 kyr**	**$1.7 \pm 0.2 \cdot 10^3$ km^3**		-
~~$T_{exp} = 30, T_{ramp} = 0.0625$~~ $T_{exp} = 30, T_{ramp} = 0.0625$	-2.4 ± 3.6	2.4 ± 3.9	-0.1 ± 2.9	-0.4 ± 2.4	7.9 ± 2.2	0.0 ± 0.1
~~$T_{exp} = 35, T_{ramp} = 0.0625$~~ $T_{exp} = 35, T_{ramp} = 0.0625$	-2.6 ± 4.7	2.6 ± 4.8	0.5 ± 4.3	-0.6 ± 4.1	7.9 ± 2.3	0.1 ± 0.2
~~$T_{exp} = 45, T_{ramp} = 0.0625$~~ $T_{exp} = 45, T_{ramp} = 0.0625$	-1.3 ± 4.8	1.8 ± 4.4	-0.1 ± 1.6	-1.6 ± 4.1	7.8 ± 2.2	0.1 ± 0.1
~~$T_{exp} = 56, T_{ramp} = 0.0625$~~ $T_{exp} = 56, T_{ramp} = 0.0625$	-1.7 ± 4.7	1.3 ± 4.7	-3.2 ± 2.1	-0.4 ± 5.2	7.7 ± 2.1	0.2 ± 0.0
~~$T_{exp} = 28, T_{ramp} = 0.03125$~~ $T_{exp} = 28, T_{ramp} = 0.03125$	-0.8 ± 4.9	3.2 ± 8.0	-2.3 ± 3.1	-0.2 ± 3.8	7.8 ± 2.3	0.2 ± 0.1

Table S10. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM

ramp	score-mean	score-std	sum of scores
res= 25 km, ~~$T_{\text{exp}} = 5, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 5, T_{\text{ramp}} = 0.5$	**0.85**	**2.92**	**3.77**
res= 25 km, ~~$T_{\text{exp}} = 15, T_{\text{ramp}} = 1$~~ $T_{\text{exp}} = 15, T_{\text{ramp}} = 1$	1.08	3.53	4.61
res= 25 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.5$	1.58	3.88	5.46
res= 12.5 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.25$	3.87	4.11	7.98
res= 12.5 km, ~~$T_{\text{exp}} = 20, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 20, T_{\text{ramp}} = 0.25$	3.04	2.55	**5.59**
res= 12.5 km, ~~$T_{\text{exp}} = 25, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 25, T_{\text{ramp}} = 0.25$	3.43	3.31	6.74
res= 12.5 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.25$	**2.93**	2.72	5.65
res= 12.5 km, ~~$T_{\text{exp}} = 30, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 30, T_{\text{ramp}} = 0.25$	3.54	**2.45**	5.99
res= 12.5 km, ~~$T_{\text{exp}} = 35, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 35, T_{\text{ramp}} = 0.25$	3.30	2.98	6.28
res= 12.5 km, ~~$T_{\text{exp}} = 45, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 45, T_{\text{ramp}} = 0.25$	3.36	3.17	6.54
res= 12.5 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.0625$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.0625$	3.40	2.87	6.27
res= 6.25 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.125$	2.73	1.69	4.42
res= 6.25 km, ~~$T_{\text{exp}} = 15, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 15, T_{\text{ramp}} = 0.125$	2.13	**1.17**	**3.30**
res= 6.25 km, ~~$T_{\text{exp}} = 20, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 20, T_{\text{ramp}} = 0.125$	1.92	2.98	4.90
res= 6.25 km, ~~$T_{\text{exp}} = 25, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 25, T_{\text{ramp}} = 0.125$	2.09	2.35	4.44
res= 6.25 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.125$	2.05	2.10	4.15
res= 6.25 km, ~~$T_{\text{exp}} = 30, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 30, T_{\text{ramp}} = 0.125$	1.95	1.63	3.58
res= 6.25 km, ~~$T_{\text{exp}} = 35, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 35, T_{\text{ramp}} = 0.125$	1.94	1.66	3.60
res= 6.25 km, ~~$T_{\text{exp}} = 45, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 45, T_{\text{ramp}} = 0.125$	**1.70**	2.67	4.37
res= 6.25 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.0625$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.0625$	1.80	2.56	4.36

Table S11. Single value scores for the mean and standard deviation of the basal temperature ramps. The temperature ramps are shown in Fig. S26. A total of 12, 13, and 13 ramps were tested at 25 km, 12.5 km, and 6.25 km horizontal grid resolution, respectively. Note that ramps whose sum (score-mean + score-std) differ by more than 50 % from the minimum sum at the corresponding resolution are not listed here. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. No runs crashed and all runs had more than 1 surge. Note that the sum of scores can be slightly off due to rounding (± 0.01).

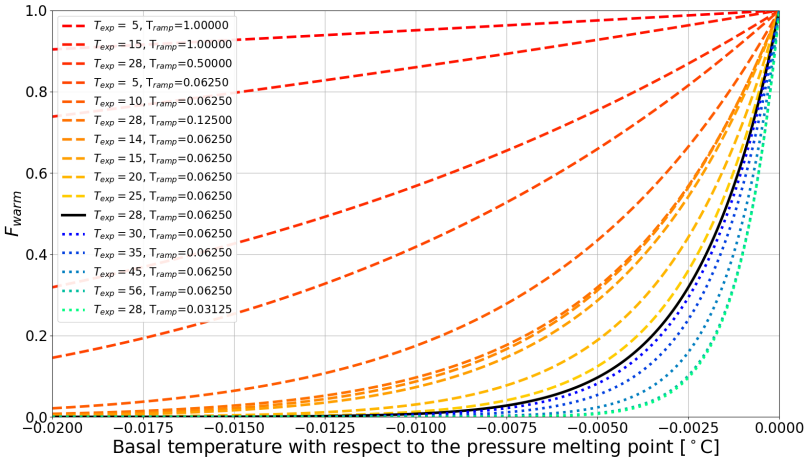


Figure S25. Temperature ramps for different values of ~~T_{ramp}~~ T_{ramp} and ~~T_{exp}~~ T_{exp}. The black solid line shows the ramp used for the 3.125 km horizontal grid resolution reference setup (~~$T_{\text{ramp}} = 0.0625$~~$T_{\text{ramp}} = 0.0625$, ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$). The solid and dotted lines show ramps that are wider and sharper than the reference setup, respectively. The depicted temperature ramps are the same as the ones listed in Fig. 9 and Table S10.

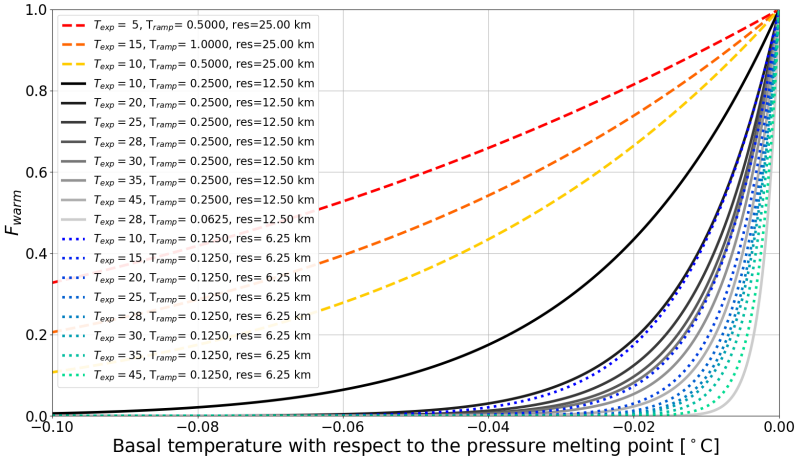


Figure S26. Shown are the temperature ramps listed in Table S11 at 25 km (solid lines), 12.5 km (dashed lines), and 6.25 km horizontal grid resolution (dotted lines).

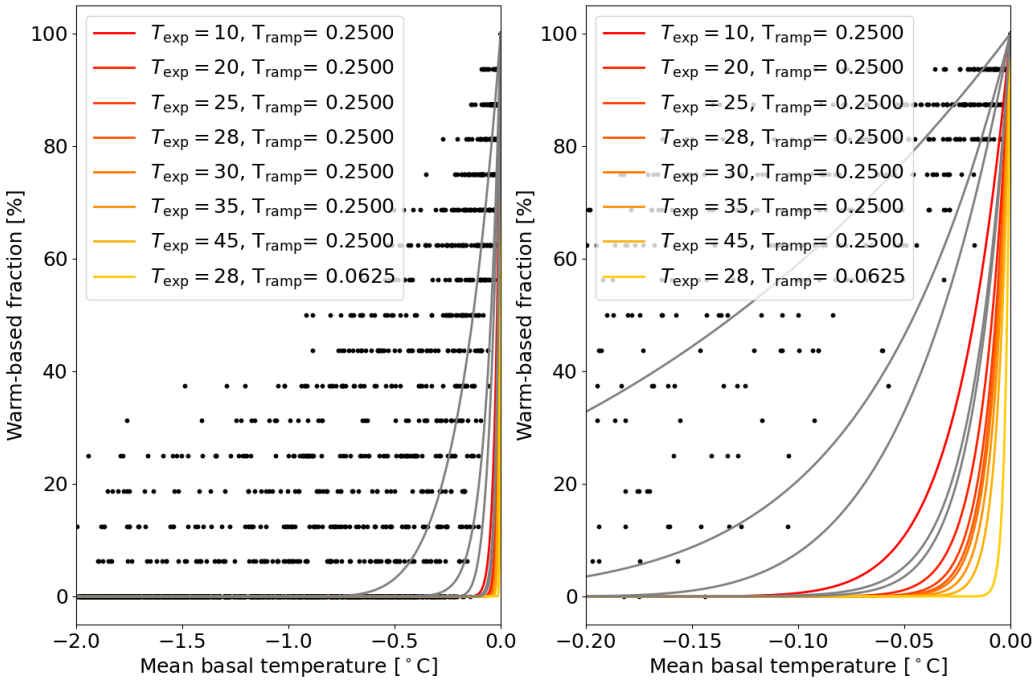


Figure S27. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 12.5 km horizontal grid resolution including all 5 parameter vectors using the GSM. Only grid cells within the pseudo-Hudson Strait and time steps within the surges of the 10 kyr after the first surge are considered. The restriction to the 10 kyr after the first surge for these experiments is set by storage limitations due to the high temporal resolution of the model output fields (10 yr). The colored ramps correspond to the 12.5 km horizontal grid resolution basal temperature ramps in Table S11 and the gray lines show all other ramps that were tested at this resolution.

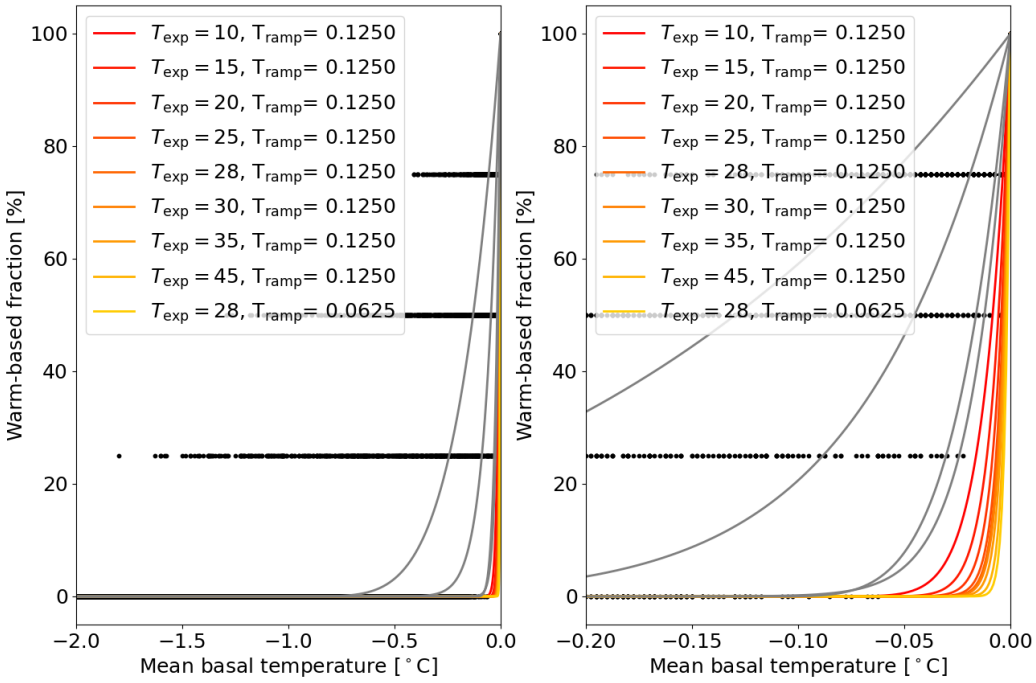


Figure S28. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 6.25 km horizontal grid resolution including all 5 parameter vectors using the GSM. The colored ramps correspond to the 6.25 km horizontal grid resolution basal temperature ramps in Table S11 and the gray lines show all other ramps that were tested at this resolution. Otherwise same as Fig. S27.

S7.4 Smooth sediment transition zone and non-flat topography

Metric	reference setup (abrupt transition) reference values	3.125 km wide transition	25 km wide transition	3.125 km wide transition with HB/HS topography	25 km wide transition with HB/HS topography
number of surges	180 ± 100	-4.2 ± 8.9	1.0 ± 11.4	36.3 ± 17.3	19.9 ± 22.6
mean period	1.1 ± 0.5 kyr	4.6 ± 9.2	-0.4 ± 10.4	2.2 ± 48.0	14.5 ± 45.1
mean duration	0.3 ± 0.1 kyr	2.7 ± 3.6	7.2 ± 4.4	10.2 ± 17.5	24.3 ± 9.0
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	0.2 ± 4.8	-1.7 ± 4.1	8.4 ± 10.0	17.3 ± 15.6
RMSE	-	7.9 ± 2.3	8.0 ± 2.2	11.2 ± 1.8	12.2 ± 2.0
Mean Bias	-	0.0 ± 0.2	-0.6 ± 0.5	-6.2 ± 1.9	-6.6 ± 2.2

Table S12. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup for runs with a smooth transition between hard bedrock and soft sediment, and runs with a pseudo-Hudson Bay/Hudson Strait (HB/HS) topography (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (except for the RMSE and mean bias).

Comparing the runs for parameter vector 1 more closely shows that when using a non-flat topography, the surges now start and propagate at the southernmost and northernmost end of the pseudo-Hudson Strait, where the topography is deepest and begins to slope upwards. Additionally, the surges tend to propagate faster and extend further to the West and in North-South direction than without topography (e.g., 8.0 to 8.3 kyr in the bottom row of video 07 of Hank (2023)). This is mainly due to warmer basal conditions in the transition zone and Hudson Bay region before the start of the surge (200 m bed depression increases the heat generation at the bed (video 08 of Hank (2023)) which, in turn, increases the average basal temperature with respect to the pressure melting point). An interesting effect of the 200 m deep Hudson Strait and 500 m deep ocean is that the pressure melting point is first reached further inland and not at the eastern end of the pseudo-Hudson Strait, as is the case for a flat topography (e.g., 7.8 to 8.1 kyr in the bottom row of video 07 of Hank (2023)).

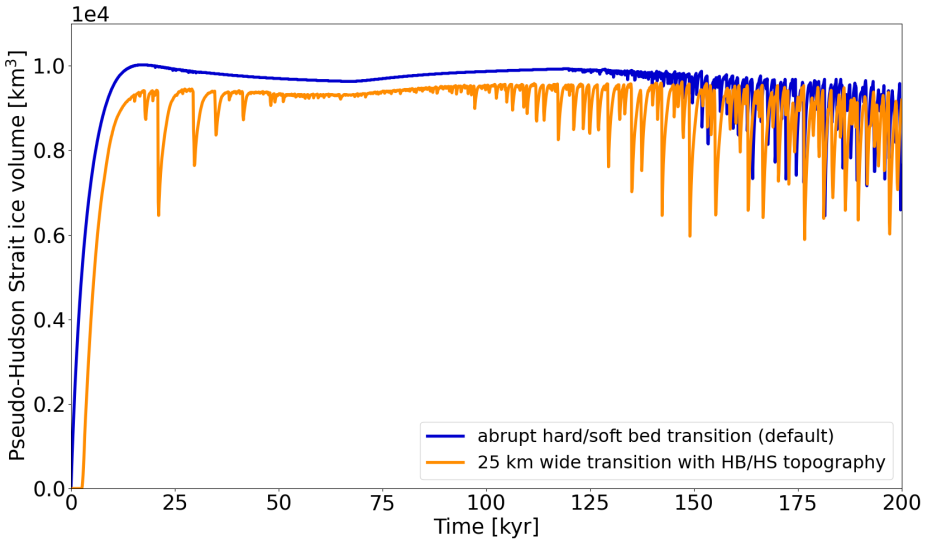


Figure S29. Pseudo-Hudson Strait ice volume for parameter vector 0 with and without a 200 m deep topography in the pseudo-Hudson Bay and Hudson Strait region using the GSM. In runs with a non-flat topography, the initial glaciation is delayed because the pseudo-Hudson Strait topography is below sea level. The horizontal grid resolution is 3.125 km.

Metric	reference setup (smooth transition, e.g., Fig. S8) reference values	abrupt transition	smooth transition with HB/HS topography
number of surges	35 ± 25	-4.8 ± 13.4	-0.1 ± 190
mean period	10 ± 10 kyr	7.4 ± 24.5	79.0 ± 64.1
mean duration	3 ± 2 kyr	21.5 ± 59.4	56.2 ± 53.4
mean pseudo-Hudson Strait ice volume change	$1.1 \pm 0.3 \cdot 10^5$ km^3	14.1 ± 37.8	389.5 ± 169.8
RMSE	-	4.4 ± 2.5	37.7 ± 2.5
Mean Bias	-	-0.2 ± 0.3	-30.4 ± 1.4
nS0	0	0	3

Table S13. Percentage differences of surge characteristics, ice volume RMSE and mean bias compared to the PISM reference setup for runs with an abrupt transition between hard bedrock and soft sediment, and runs with a pseudo-Hudson Bay/Hudson Strait (HB/HS) topography (except first column). No runs crashed and runs without surges (nS0) only contribute to the change in surge number. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (except for the RMSE and mean bias).

S7.5 Basal hydrology

Metric	no hydrology	local hydrology	no hydrology, double C_{rmu}	no hydrology, double C_{fslid}
number of surges	180 ± 100	-3.8 ± 23.8	-9.5 ± 3.9	-3.0 ± 8.8
mean period	1.1 ± 0.5 kyr	17.4 ± 44.9	12.4 ± 4.1	4.5 ± 10.3
mean duration	0.3 ± 0.1 kyr	11.6 ± 19.1	3.1 ± 5.6	2.3 ± 3.5
mean ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	20.2 ± 44.7	10.5 ± 5.9	-0.9 ± 5.8
RMSE	-	8.7 ± 2.6	8.5 ± 2.7	7.8 ± 2.2
Mean Bias	-	-0.9 ± 0.8	-0.4 ± 0.4	-0.1 ± 0.1

Table S14. Percentage differences of surge characteristics, ice volume RMSE and mean bias of GSM runs with a local basal hydrology model compared to runs without sub-glacial hydrology in percent (except first column). Additionally shown are the changes in surge characteristics when doubling the values of the soft and hard bed sliding coefficient (C_{rmu} and C_{fslid} in Table 1, respectively). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S8 Sensitivity experiments without a significant effect

S8.1 Weight of adjacent minimum basal temperature

Depending on the location of the adjacent minimum grid cell center basal temperature, either the ice flow (when the adjacent minimum basal temperature is downstream) or upstream propagation of the surge should be affected (decreasing basal interface temperature with increasing weight). For the large-scale surges, the adjacent minimum basal temperature is almost exclusively located upstream (e.g., video 02 of Hank (2023)). Changing the weight of the adjacent minimum basal temperature, therefore, affects the surge propagation rather than blocking parts of the ice flow.

Here we compare the effect of three different weights on the GSM surge characteristics (Eq. (S1)): no consideration of adjacent minimum basal temperature (~~$W_{\mathrm{Tb,min}} = 0.0$~~$W_{\mathrm{Tb,min}} = 0.0$), basal temperature at the interface depends to 50 % on the adjacent minimum basal temperature at the grid cell center (reference setup, ~~$W_{\mathrm{Tb,min}} = 0.5$~~$W_{\mathrm{Tb,min}} = 0.5$), and basal temperature at the interface is equal to the adjacent minimum basal temperature at the grid cell center (~~$W_{\mathrm{Tb,min}} = 1.0$~~$W_{\mathrm{Tb,min}} = 1.0$).

The surge cycling response to changes in ~~$W_{\mathrm{Tb,min}}$~~ $W_{\mathrm{Tb,min}}$ is not coherent (Table S15). For instance, the mean surge period increases for both ~~$W_{\mathrm{Tb,min}} = 0.$ and $W_{\mathrm{Tb,min}} = 1.0$~~ $W_{\mathrm{Tb,min}} = 0.$ and $W_{\mathrm{Tb,min}} = 1.0$ compared to the reference ~~$W_{\mathrm{Tb,min}} = 0.5$~~$W_{\mathrm{Tb,min}} = 0.5$. However, standard deviations are large, indicating a different model response for different parameter vectors.

Metric	reference setup	~~$W_{\mathrm{Tb,min}} = 0.0$~~ $W_{\mathrm{Tb,min}} = 0.0$	~~$W_{\mathrm{Tb,min}} = 1.0$~~ $W_{\mathrm{Tb,min}} = 1.0$
#Surges	180 ± 100	-9.6 ± 6.5	-3.7 ± 7.8
mean period	1.1 ± 0.5 kyr	14.7 ± 13.5	3.0 ± 0.8
mean duration	0.3 ± 0.1 kyr	5.1 ± 4.9	-2.6 ± 3.3
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	-1.9 ± 4.0	4.0 ± 6.7
RMSE	-	7.8 ± 2.5	8.0 ± 2.5
Mean Bias	-	-0.1 ± 0.1	0.3 ± 0.1

Table S15. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup (~~$W_{\mathrm{Tb,min}} = 0.5$~~$W_{\mathrm{Tb,min}} = 0.5$) for different weights of the adjacent minimum basal temperature for the basal sliding temperature ramp (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S8.2 Different approaches to basal hydrology

Here we compare the effects on surge characteristics when using a horizontal transport model in PISM instead of a simple local basal hydrology. In general, PISM experiments with a mass-conserving horizontal transport hydrology model yield similar results to the local hydrology model (Fig. 7 and Table S16). The mean duration, period, and ice volume change increase (11 %, 10 %, and 7 %, respectively), while the number of surges decreases (5 %). These differences are on the same level as the MNEEs (Table 6) and show large standard deviations, indicating a different model response for different parameter vectors. The ice volume RMSE and mean bias are also small (+3.9 % and −0.1 %, respectively).

Metric	local hydrology	horizontal transport
number of surges	35 ± 25	-4.6 ± 14.5
mean period	10 ± 10 kyr	10.8 ± 27.8
mean duration	3 ± 2 kyr	10.5 ± 35.4
mean ice volume change	$1.1 \pm 0.3 \cdot 10^5$ km^3	6.8 ± 17.9
RMSE	-	3.9 ± 2.5
Mean Bias	-	-0.1 ± 0.3

Table S16. Percentage differences of surge characteristics, ice volume (eastern half of pseudo-Hudson Bay and the pseudo-Hudson Strait) RMSE and mean bias of PISM runs with a mass-conserving horizontal transport hydrology model compared to the local hydrology model (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S8.3 Basal hydrology instead of basal temperature ramp as the primary smoothing mechanism

We examine the effects of a local basal hydrology as main smoothing mechanism for basal sliding (compared to a basal temperature ramp) by using a very sharp ramp (~~$T_{\text{ramp}} = 0.001$, $T_{\text{exp}} = 28$~~$T_{\text{ramp}} = 0.001$, $T_{\text{exp}} = 28$), minimizing the smoothing effect of the basal temperature ramp. The change in surge characteristics between runs with local basal hydrology and the sharp temperature ramp and the GSM reference setup is similar (maximum difference of 3 %; compare Table S14 and S17) to the runs with local basal hydrology and the reference basal temperature ramp (~~$T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 28$~~$T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 28$), indicating that the local basal hydrology is the primary smoothing mechanism in both cases. The differences in the change of surge characteristics between the reference and the steeper ramp are smaller than the MNEEs, preventing further analysis.

Metric	reference setup	steeper ramp (~~$T_{\text{ramp}} = 0.001$, $T_{\text{exp}} = 28$~~$T_{\text{ramp}} = 0.001$, $T_{\text{exp}} = 28$), local hy
#Surges	180 ± 100	-3.8 ± 24.6
mean period	1.1 ± 0.5 kyr	16.0 ± 42.0
mean duration	0.3 ± 0.1 kyr	8.7 ± 17.0
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	21.5 ± 43.3
RMSE	-	8.9 ± 3.2
Mean Bias	-	-0.6 ± 0.9

Table S17. Percentage differences of surge characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM reference setup with local basal hydrology instead of a basal temperature ramp as the primary smoothing mechanism (except first column). No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval for the surge characteristics (not the RMSE and mean bias).

S9 Convergence study

S9.1 GSM convergence study without basal hydrology

Analyzing individual GSM parameter vectors in detail shows that some discrepancies prevail even when using a resolution-dependent temperature ramp. In the case of parameter vector 1, for example, surges do still not occur for the coldest temperatures (Fig. S30). Note the asymmetry in termination and onset of surge cyclicity ($\Delta t_1 < \Delta t_2$). For increasing temperatures after the minimum surface temperature $T_{\text{min}} = -15°\text{C}$ at $t_{\text{min}} = 66.7$ kyr, the first surge occurs at a surface temperature slightly higher than the initial temperature T_{init}, for which oscillations occur. The difference between Δt_1 and Δt_2 is ~ 25 kyr and closely resembles the lag of the average pseudo-Hudson Strait basal temperature with respect to the pressure melting point behind the surface temperature changes. For example, the minimum average pseudo-Hudson Strait basal temperature with respect to the pressure melting point ($T_{\text{bpm}} = -3.2°\text{C}$) occurs 23 kyr after the minimum surface temperature (not shown). The

period without oscillations in the 25 km run corresponds to a period of somewhat smaller and less frequent oscillations in the finer resolution runs (Fig. S24).

Setup	number of surges	mean period	mean duration	mean pseudo-Hudson Strait ice volume change	nS0
3.125 km reference setup	180 ± 100	1.1 ± 0.5 kyr	0.3 ± 0.1 kyr	$1.7 \pm 0.2 \cdot 10^3$ km^3	0
25 km, constant ramp	-95.1 ± 7.4	942.3 ± 517.70	300.0 ± 172.22	95.9 ± 52.6	3
25 km, resolution-dependent ramp	-78.1 ± 18.2	414.5 ± 309.0	119.5 ± 17.6	91.9 ± 23.6	1
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	-15.9 ± 20.4	29.7 ± 24.6	43.8 ± 36.6	3.5 ± 18.7	0
12.5 km, constant ramp	-59.2 ± 16.5	129.0 ± 41.8	90.3 ± 17.9	50.3 ± 76.5	0
12.5 km, resolution-dependent ramp, also minimum mean score	-56.5 ± 15.1	115.7 ± 46.8	101.1 ± 20.5	33.0 ± 66.3	0
6.25 km, constant ramp	-24.2 ± 13.1	36.4 ± 20.9	24.8 ± 8.5	14.9 ± 14.2	0
6.25 km, resolution-dependent ramp	-27.9 ± 9.9	42.2 ± 18.9	32.1 ± 6.3	15.9 ± 12.3	0
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 45$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 45$	-25.3 ± 13.6	37.9 ± 26.7	28.2 ± 7.0	9.8 ± 11.6	0

Table S18. Percentage differences of surge characteristics compared to the 3.125 km GSM reference setup (except first row). The values represent the average of 5 parameter vectors. No runs crashed and runs without surges (nS0) only contribute to the change in surge numbers. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above. The resolution-dependent ramps (~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$) and constant ramp (black line, ~~$T_{\text{ramp}} = 0.0625$~~$T_{\text{ramp}} = 0.0625$, ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$) are shown in Fig. 2. The third ramp listed for each resolution is the ramp with the smallest mean score (Table S11).

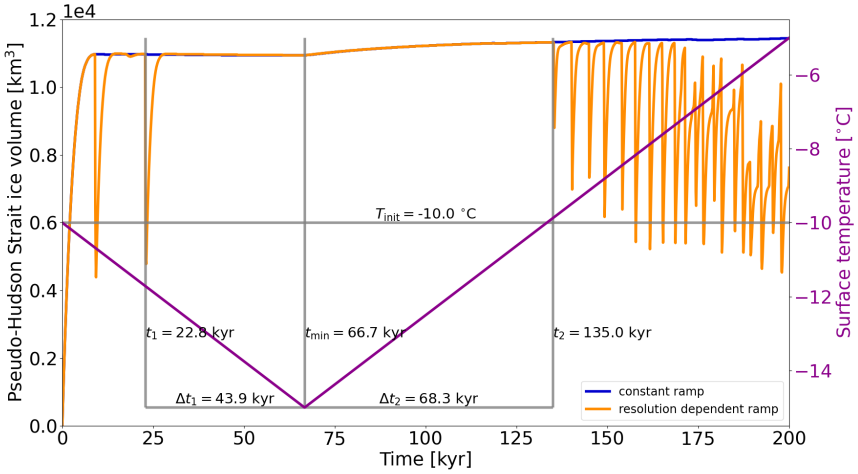


Figure S30. Pseudo-Hudson Strait ice volume for parameter vector 1 and different basal temperature ramps using the GSM (constant ramp: ~~$T_{\text{ramp}} = 0.0625°$~~$T_{\text{ramp}} = 0.0625°$C and ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$; resolution-dependent ramp: ~~$T_{\text{ramp}} = 0.5°$~~$T_{\text{ramp}} = 0.5°$C and ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$, see Fig. 2). The right axis shows the surface temperature when ignoring the lapse rate dependency. t_{min}, t_1, and t_2 mark the time of the minimum surface temperature, the start of the last surge before t_{min}, and the start of the first surge after t_{min}, respectively. Δt_1 and Δt_2 represent the time difference between t_{min} and t_1 and t_2, respectively. T_{init} indicates the surface temperature at the beginning of the run.

Setup	mean RMSE	mean Bias
25 km, constant ramp	17.1 ± 4.7	14.6 ± 4.4
12.5 km, constant ramp	10.3 ± 2.5	4.0 ± 1.1
6.25 km, constant ramp	8.7 ± 2.2	0.4 ± 0.5
25 km, resolution-dependent ramp	15.3 ± 3.0	10.3 ± 2.5
12.5 km, resolution-dependent ramp	10.3 ± 2.8	3.0 ± 2.3
6.25 km, resolution-dependent ramp	8.5 ± 2.3	0.2 ± 0.6
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	12.8 ± 2.4	6.4 ± 1.7
12.5 km, ~~$T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 28$~~ $T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 28$, same as resolution-dependent ramp	10.3 ± 2.8	3.0 ± 2.3
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 45$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 45$	8.5 ± 2.0	0.6 ± 0.5

Table S19. Pseudo-Hudson Strait ice volume RMSE and mean bias compared to the 3.125 km GSM reference setup in percent. The values represent the average of 5 parameter vectors. No runs crashed and the entire 200 kyr run time is used (no spin-up interval).

S9.2 GSM convergence study with basal hydrology

Based on the results without basal hydrology (Sec. 3.4.1), 5 basal temperature ramps (~~$T_{\mathrm{exp}} = [5, 10, 15, 20, 28]$~~ $T_{\mathrm{exp}} = [5, 10, 15, 20, 28]$) with a resolution-dependent ~~T_{ramp}~~ T_{ramp} (Eq. (9)) are tested for all resolutions. As it is unclear which basal temperature ramp should be used at the finest horizontal grid resolution (3.125 km), we test two different ramps (~~$T_{\mathrm{exp}} = [5, 28]$~~ $T_{\mathrm{exp}} = [5, 28]$). The experiments that yield the smallest differences in surge characteristics (smallest mean score in Table S20 and S21) compared to the corresponding 3.125 km reference runs (bold rows) are presented in Table S22.

Similar to the results without a basal hydrology model, the smallest differences in surge characteristics (except the mean pseudo-Hudson Strait ice volume change) occur for the coarsest horizontal grid resolution (25 km, Table S22). This likely indicates that the optimal ramps at 12.5 and 6.25 km horizontal grid resolution have not been found.

In general, the resolution-dependent ramp with ~~$T_{\mathrm{exp}} = 5$~~ $T_{\mathrm{exp}} = 5$ leads to the smallest differences between coarse and fine-resolution runs. The differences in surge characteristics are significantly smaller than for a resolution-dependent temperature ramp without local basal hydrology (except for the mean pseudo-Hudson Strait ice volume change, Table S18 vs. S22), further underlining the importance of the basal hydrology.

Except for 12.5 km horizontal grid resolution, the resolution-dependent ramp with ~~$T_{\mathrm{exp}} = 5$~~ $T_{\mathrm{exp}} = 5$ yields a self-consistent response across all resolutions. At 12.5 km, the next closest exponent (~~$T_{\mathrm{exp}} = 10$~~ $T_{\mathrm{exp}} = 10$) has the minimum mean score. However, given that there is no single best ramp across all resolutions, we assess different ramps as to whether differences are within inferred MNEEs (DWINS). To this end, we calculate the differences between the ramp with the minimum mean score and all other ramps at each resolution and for all surge characteristics (Table S20 and S21). We rule out ramps for which the differences exceed the maximum MNEEs (maximum of Table 5 and S2) for more than one surge characteristic (DWINS failures).

Under these criteria and when using ~~$T_{\mathrm{exp}} = 5$~~ $T_{\mathrm{exp}} = 5$ at 3.125 km horizontal grid resolution, the resolution-dependent ramp with ~~$T_{\mathrm{exp}} = 10$~~ $T_{\mathrm{exp}} = 10$ remains within the DWINS ensemble for all resolutions (Table S21). The results for ~~$T_{\mathrm{exp}} = 28$~~ $T_{\mathrm{exp}} = 28$ at 3.125 km horizontal grid resolution do not yield a single ramp that remains within the DWINS ensemble at all resolutions (Table S20). However, except for 6.25 km, for which the differences between the tested basal temperature ramps are the smallest, ~~$T_{\mathrm{exp}} = 5$~~ $T_{\mathrm{exp}} = 5$ yields the minimum mean-score.

The pseudo-Hudson Strait ice volume RMSE and mean bias show convergence (smaller differences) for both 3.125 km horizontal grid resolution setups (Table S23).

ramp	score-mean	score-std	sum of scores	DWINS failures
res= 25 km, ~~$T_{\text{exp}} = 5, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 5, T_{\text{ramp}} = 0.5$	**0.99**	4.31	**5.31**	0
res= 25 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.5$	1.44	5.29	6.74	2
res= 25 km, ~~$T_{\text{exp}} = 15, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 15, T_{\text{ramp}} = 0.5$	4.80	**3.05**	7.85	4
res= 25 km, ~~$T_{\text{exp}} = 20, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 20, T_{\text{ramp}} = 0.5$	5.65	3.47	9.11	4
res= 25 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.5$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.5$	7.11	3.88	11.00	4
res= 12.5 km, ~~$T_{\text{exp}} = 5, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 5, T_{\text{ramp}} = 0.25$	**3.69**	4.60	8.29	0
res= 12.5 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.25$	3.81	5.07	8.88	2
res= 12.5 km, ~~$T_{\text{exp}} = 15, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 15, T_{\text{ramp}} = 0.25$	3.82	4.11	7.93	2
res= 12.5 km, ~~$T_{\text{exp}} = 20, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 20, T_{\text{ramp}} = 0.25$	4.21	3.42	7.63	3
res= 12.5 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.25$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.25$	4.47	**2.81**	**7.28**	4
res= 6.25 km, ~~$T_{\text{exp}} = 5, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 5, T_{\text{ramp}} = 0.125$	4.03	4.29	8.33	3
res= 6.25 km, ~~$T_{\text{exp}} = 10, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 10, T_{\text{ramp}} = 0.125$	3.94	**3.76**	7.70	3
res= 6.25 km, ~~$T_{\text{exp}} = 15, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 15, T_{\text{ramp}} = 0.125$	4.65	3.90	8.55	1
res= 6.25 km, ~~$T_{\text{exp}} = 20, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 20, T_{\text{ramp}} = 0.125$	3.79	3.82	**7.60**	1
res= 6.25 km, ~~$T_{\text{exp}} = 28, T_{\text{ramp}} = 0.125$~~ $T_{\text{exp}} = 28, T_{\text{ramp}} = 0.125$	**3.59**	4.23	7.82	0

Table S20. Single value scores for the mean and standard deviation of the basal temperature ramps and the number of DWINS failures (maximum 4) for a resolution-dependent reference temperature ramp with ~~$T_{\text{exp}} = 28$~~ $T_{\text{exp}} = 28$ in the GSM. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. At $= 25$ km, 1 run crashed for ~~$T_{\text{exp}} = 10$~~ $T_{\text{exp}} = 10$ and 1 run showed no surges for ~~$T_{\text{exp}} = [15, 20, 28]$~~ $T_{\text{exp}} = [15, 20, 28]$. Note that the sum of scores can be slightly off due to rounding (± 0.01).

ramp	score-mean	score-std	sum of scores	DWINS failures
res= 25 km, ~~$T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.5$~~ $T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.5$	**0.84**	3.91	**4.75**	0
res= 25 km, ~~$T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.5$~~ $T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.5$	1.21	5.04	6.25	1
res= 25 km, ~~$T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.5$~~ $T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.5$	4.89	**3.29**	8.18	4
res= 25 km, ~~$T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.5$~~ $T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.5$	5.76	3.63	9.40	4
res= 25 km, ~~$T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.5$~~ $T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.5$	7.30	4.13	11.43	4
res= 12.5 km, ~~$T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.25$~~ $T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.25$	3.97	4.49	8.45	2
res= 12.5 km, ~~$T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.25$~~ $T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.25$	**3.77**	4.60	8.37	0
res= 12.5 km, ~~$T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.25$~~ $T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.25$	3.79	4.13	7.93	1
res= 12.5 km, ~~$T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.25$~~ $T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.25$	4.10	3.50	**7.59**	1
res= 12.5 km, ~~$T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.25$~~ $T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.25$	4.37	**3.28**	7.65	2
res= 6.25 km, ~~$T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.125$~~ $T_{\mathrm{exp}} = 5, T_{\mathrm{ramp}} = 0.125$	**3.53**	4.44	7.97	0
res= 6.25 km, ~~$T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.125$~~ $T_{\mathrm{exp}} = 10, T_{\mathrm{ramp}} = 0.125$	4.27	3.77	8.04	0
res= 6.25 km, ~~$T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.125$~~ $T_{\mathrm{exp}} = 15, T_{\mathrm{ramp}} = 0.125$	4.59	3.82	8.42	1
res= 6.25 km, ~~$T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.125$~~ $T_{\mathrm{exp}} = 20, T_{\mathrm{ramp}} = 0.125$	3.91	**3.64**	**7.55**	1
res= 6.25 km, ~~$T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.125$~~ $T_{\mathrm{exp}} = 28, T_{\mathrm{ramp}} = 0.125$	3.70	4.33	8.03	3

Table S21. Single value scores for the mean and standard deviation of the basal temperature ramps and the number of DWINS failures (maximum 4) for a resolution-dependent reference temperature ramp with ~~$T_{\mathrm{exp}} = 5$~~ $T_{\mathrm{exp}} = 5$ in the GSM. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. At $= 25$ km, 1 run crashed for ~~$T_{\mathrm{exp}} = 10$~~ $T_{\mathrm{exp}} = 10$ and 1 run showed no surges for ~~$T_{\mathrm{exp}} = [15, 20, 28]$~~ $T_{\mathrm{exp}} = [15, 20, 28]$. Note that the sum of scores can be slightly off due to rounding (± 0.01).

Setup	number of surges	mean period	mean duration	mean pseudo-Hudson Strait ice volume change
3.125 km, $T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 28$	**197 ± 131**	**1.5 ± 1.1 kyr**	**0.3 ± 0.2 kyr**	**$2.0 \pm 0.7 \cdot 10^3$ km^3**
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	9.7 ± 59.9	15.5 ± 42.3	24.3 ± 36.1	13.6 ± 46.7
12.5 km, ~~$T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 5$	-36.1 ± 17.6	68.0 ± 49.8	97.1 ± 60.3	3.0 ± 26.4
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 28$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 28$	-13.2 ± 31.1	27.0 ± 40.6	25.7 ± 25.2	5.6 ± 27.5
3.125 km, $T_{\text{ramp}} = 0.0625$, $T_{\text{exp}} = 5$	**190 ± 118**	**1.3 ± 0.7 kyr**	**0.3 ± 0.2 kyr**	**$1.8 \pm 0.4 \cdot 10^3$ km^3**
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	-2.4 ± 35.8	16.1 ± 31.4	20.7 ± 30.3	14.3 ± 35.8
12.5 km, ~~$T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 10$~~ $T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 10$	-37.7 ± 12.1	61.7 ± 44.1	63.4 ± 34.8	20.5 ± 39.0
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 5$	-25.6 ± 13.9	37.8 ± 23.8	41.1 ± 21.3	0.3 ± 19.8

Table S22. Percentage differences (except bold rows) of surge characteristics compared to the 3.125 km GSM setups with local basal hydrology (bold rows, ~~$T_{\text{exp}} = [5, 28]$~~$T_{\text{exp}} = [5, 28]$) for the ramps with the smallest mean score (analysis steps described in Sec. S7.3). The values represent the average of 5 parameter vectors. No runs crashed and all runs had more than 1 surge. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

Setup	mean RMSE	mean Bias
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	14.3 ± 3.2	6.0 ± 0.9
12.5 km, ~~$T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 5$	11.2 ± 4.2	0.6 ± 2.4
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 28$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 28$	10.0 ± 3.0	0.5 ± 0.6
25 km, ~~$T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.5$, $T_{\text{exp}} = 5$	14.5 ± 3.2	6.8 ± 0.4
12.5 km, ~~$T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 10$~~ $T_{\text{ramp}} = 0.25$, $T_{\text{exp}} = 10$	11.7 ± 4.2	1.6 ± 2.5
6.25 km, ~~$T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 5$~~ $T_{\text{ramp}} = 0.125$, $T_{\text{exp}} = 5$	10.1 ± 1.8	0.6 ± 0.8

Table S23. Resolution scaling of pseudo-Hudson Strait ice volume RMSE and mean bias with local basal hydrology in percent. The three upper ramps are compared to the 3.125 km GSM setup with ~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$, the lower three to ~~$T_{\text{exp}} = 5$~~$T_{\text{exp}} = 5$. The values represent the average of 5 parameter vectors. No runs crashed and the entire 200 kyr run time is used (no spin-up interval).

S9.3 GSM convergence study with active SSA everywhere (no basal hydrology)

Setup	number of surges	mean period	mean duration	mean pseudo-Hudson Strait ice volume change	nS0
3.125 km reference setup	216 ± 146	1.3 ± 0.8 kyr	0.3 ± 0.1 kyr	$1.4 \pm 0.6 \cdot 10^3$ km^3	0
25 km	-76.9 ± 17.9	432.2 ± 384.3	151.7 ± 49.0	163.8 ± 65.9	1
12.5 km	-61.3 ± 23.8	179.8 ± 127.1	154.3 ± 79.3	40.1 ± 54.4	0
6.25 km	-46.2 ± 11.6	66.7 ± 12.9	59.8 ± 19.5	75.1 ± 42.9	0

Table S24. Percentage differences of surge characteristics compared to the 3.125 km GSM reference setup (except first row) with a resolution-dependent basal temperature ramp (~~$T_{\text{exp}} = 28$~~$T_{\text{exp}} = 28$, Fig. 2) and active SSA everywhere. The values represent the average of 5 parameter vectors. No runs crashed and runs without surges (nS0) only contribute to the change in surge numbers. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

S9.4 PISM convergence study

Similar to the results presented for the GSM (Sec. S9.1 and S9.2), analyzing individual parameter vectors for PISM shows significant differences in surge behavior for different horizontal grid resolutions. Parameter vector 8 at 25 km horizontal grid resolution, for example, only shows 4 oscillations (Fig. S31). In contrast, more oscillations occur for both the 12.5 km and 50 km horizontal grid resolution run. Additionally, most of the 50 km surges transport ice toward the West, whereas the 25 and 12.5 km runs almost exclusively surge through the pseudo-Hudson Strait (video 09 of Hank (2023)).

Setup	number of surges	mean period	mean duration	mean ice volume change	nC	nS0	nS1
12.5 km reference setup	$\mathbf{22 \pm 19}$	$\mathbf{9 \pm 6}$ **kyr**	$\mathbf{3 \pm 1}$ **kyr**	$\mathbf{1.3 \pm 0.2 \cdot 10^5}$ **km^3**	**4**	**1**	**0**
50 km	4.1 ± 46.0	15.3 ± 47.4	11.9 ± 33.7	30.6 ± 39.6	0	0	1
25 km	-28.3 ± 12.0	46.5 ± 31.9	6.3 ± 13.8	4.0 ± 20.3	0	0	0

Table S25. Percentage differences (except bold rows) of PISM surge characteristics due to different horizontal grid resolutions. Note that the 12.5 km (finest resolution tested) is used as a reference for the grid resolution convergence study. 4 of the 12.5 km runs crashed after $\sim$ 50 kyr because they hit the run-time limit on the computational cluster (7 days) and 1 12.5 km run does not show a surge (nS0). Crashed runs (nC) are not considered and runs without surges in the comparison setup only contribute to the change in surge numbers. Runs without surges in the reference setup are not considered. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

Setup	nC	mean RMSE	mean Bias
50 km	0	11.1 ± 2.6	6.5 ± 4.1
25 km	0	7.4 ± 1.4	3.7 ± 0.8

Table S26. Ice volume RMSE and mean bias (in percent) due to different horizontal grid resolutions. Note that the 12.5 km (finest resolution tested) is used as a reference for the grid resolution convergence study. 4 of the 12.5 km runs crashed after $\sim$ 50 kyr because they hit the run-time limit on the computational cluster (7 days) and 1 12.5 km run does not show a surge (nS0). Crashed runs (nC) are not considered. The entire 200 kyr run time is used (no spin-up interval).

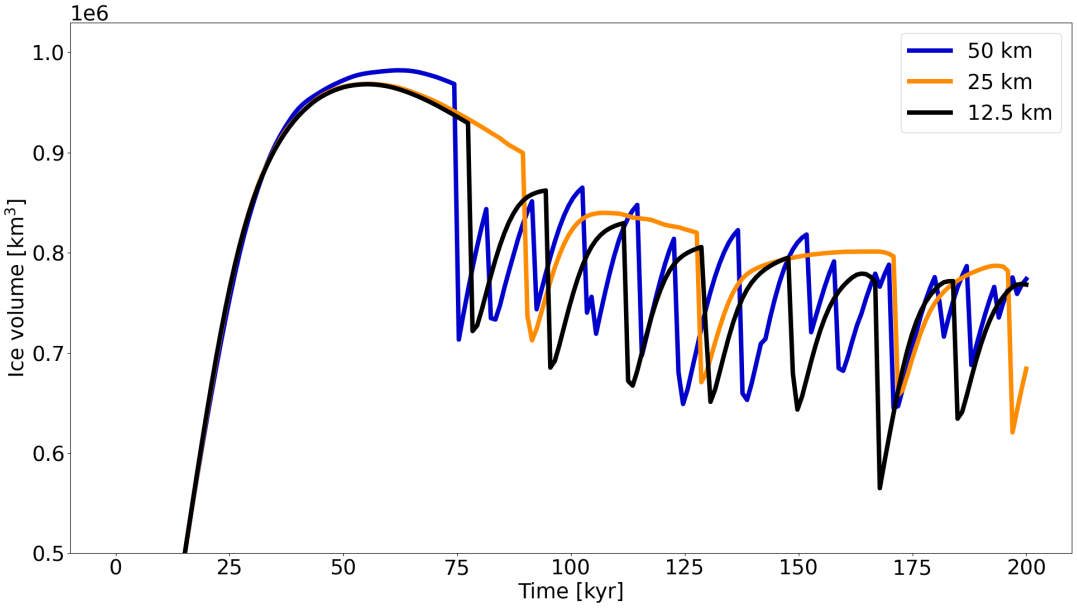


Figure S31. Ice Volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 8 and different horizontal grid resolutions using the PISM. See also video 09 of Hank (2023).

References

Bueler, E. and Van Pelt, W.: Mass-conserving subglacial hydrology in the Parallel Ice Sheet Model version 0.6, Geoscientific Model Development, 8, 1613–1635, https://doi.org/10.5194/gmd-8-1613-2015, 2015.

Hank, K.: Supplementary material for "Numerical issues in modeling thermally and hydraulically driven ice stream surge cycling", https://doi.org/10.5281/zenodo.7905404, 2023.

K.M. Cuffey and W.S.B. Paterson.: The Physics of Glaciers, Butterworth-Heinemann/Elsevier, Burlington, MA, 4th edn., 2010.

Mitchell, J. and Soga, K.: Fundamentals of Soil Behavior, John Wiley & Sons, Inc., 3ed edn., 2005.

Sauer, E. K., Egeland, A. K., and Christiansen, E. A.: Preconsolidation of tills and intertill clays by glacial loading in southern Saskatchewan, Canada, Canadian Journal of Earth Sciences, 30, 420–433, https://doi.org/10.1139/e93-031, 1993.

Tulaczyk, S., Kamb, W. B., and Engelhardt, H. F.: Basal mechanics of Ice Stream B, West Antarctica 1. Till mechanics, Journal of Geophysical Research: Solid Earth, 105, 463–481, https://doi.org/10.1029/1999jb900329, 2000.