https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Numerical issues in modeling ice sheet instabilities such as binge-purge type cyclic ice stream surging

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

[1]Department of Physics and Physical Oceanograph, Memorial University of Newfoundland, St. John's, NL, A1B 3X7,Canada
[2]Institute for Marine and Antarctic Studies, University of Tasmania, 20 Castray Esplanade, Battery Point TAS 7004, Australia
[3]The Australian Centre for Excellence in Antarctic Science, University of Tasmania, Hobart, TAS, Australia
*khank@mun.ca

Correspondence: Kevin Hank (khank@mun.ca)

Abstract. As in any environmental system, modeling instabilities within the glacial system is a numerical challenge of potentially high real-world relevance. Differentiating between the impacts of physical system processes and numerical noise 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 numerical sensitivity of ice stream surge cycling in glaciological models. Through sensitivity tests, we identify some numerical requirements for a 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 higher horizontal grid resolutions. Discrepancies between high and coarse horizontal grid resolutions can be reduced by incorporating a resolution-dependent basal temperature ramp for basal sliding thermal activation. Inclusion of a diffusive bed thermal model reduces the surge cycling ice volume change by ~ 33 % as the additional heat storage dampens the change in basal temperature during surge events. The inclusion of basal hydrology, as well as a non-flat topography, leads to increased ice volume change during surge events (~ 20 and 17 %, respectively). Therefore, these latter three components are essential if one is endeavoring to maximize physical fidelity in ice stream surge cycle modeling. An abrupt transition between hard bedrock and soft sediment, as in the HEINO experiments, leads to ice stream propagation along this boundary but is not the cause of the main surge events.

1 Introduction

The use of Ice Sheet Models (ISMs) 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 components and implementations. Model validation is particularly important when modeling highly non-linear ice sheet insta-

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

bilities, for which it is hard to distinguish between numerical noise and physical phenomena. In addition, there are a number of numerical choices, such as for thermal activation of basal sliding, for which no model to date has documented sensitivities.

One archetypal ice sheet instability is the binge-purge ice stream cycling explanation of Heinrich Events (HEs) originally proposed by MacAyeal (1993). Within this hypothesis, the Laurentide Ice Sheet (LIS) in the Hudson Bay/Hudson Strait region gradually grows to a threshold thickness (binge phase). Once the ice sheet is thick enough to sufficiently isolate the ice sheet 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 sheet is no longer frozen to the bed and thus enables basal sliding. Localized warm-based ice streaming increases the ice sheet surface gradient (steeper slope) at the warm/cold-base transition point, leading to an increase in driving stress. The resultant increase of deformation work can warm the surrounding ice to the pressure melting point. The presence of water at the ice sheet/bed interface as well as in a deformable sediment layer can further increase ice velocities. Fowler and Schiavi (1998) call this sequence of events the propagation of an *"activation wave"*. Instead of the slow deformation flow (ice creep), the ice sheet now destabilizes rapidly (purge phase). As a consequence of the high ice velocities, the ice sheet thins and cold ice is advected from either upstream or the 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 *pinning point*, supporting some of the driving stress and decreasing the velocities and heat production in the adjacent ice. This *"deactivation wave"* (Fowler and Schiavi, 1998) ends the surge and the ice sheet enters the next binge phase.

Meaningful modeling of binge-purge type HEs and surges in general is challenging. Sensitivity in model response to different numerical choices are evident (Calov et al., 2010; Roberts et al., 2016; Ziemen et al., 2019) and small perturbations of the system can significantly vary the form, amplitude, and period of binge-purge oscillations (Souček and Martinec, 2011; Mantelli et al., 2016). The exact cause of the numerical sensitivities is often unclear. To date, uncertainties associated with the numerical aspects of a model have received limited attention in studies examining ice sheet surging (e.g., Payne, 1995; Marshall and Clarke, 1997; Calov et al., 2002; Papa et al., 2006; Steen-Larsen and Dahl-Jensen, 2008; Calov et al., 2010; Robel et al., 2013; Feldmann and Levermann, 2017). 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"*. Furthermore, the theory underpinning the understanding of the instability mechanisms is not fully developed, especially in the context of a spatially extended 3D system, thus precluding systematic benchmarking of numerical models.

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 higher grid resolution is necessary to resolve the Hudson Strait properly. However, studies examining the effect of different grid resolutions on surge behavior are sparse (Greve et al., 2006; Van Pelt and Oerlemans, 2012; Brinkerhoff and Johnson, 2015; Roberts et al., 2016) and an in-depth numerical analysis of Hudson Strait binge-purge type surges (to whatever idealized form) is missing entirely.

Greve and MacAyeal (1996) examined the impact of different time steps on binge-purge oscillations in a coupled dynamic/thermodynamic flowline model. They report a similar dynamic behavior across different time steps, but both test runs

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

crashed. Later studies using a three-dimensional version of the same but further developed model find shorter periods and a slight decrease in surge amplitude but otherwise reasonable convergence as the time step decreases. (Greve et al., 2006; Takahama, 2006). The same two studies also show only a minor effect of grid rotation on the general features of the oscillations. None of the more recent studies on binge-purge type surge events includes experiments with different time steps.

Accounting for discrepancies associated with modeling choices becomes even more important when investigating the effects of artificially induced noise. For example, Souček and Martinec (2011) show that low levels of surface temperature noise can lead to chaotic behavior in the time lag between subsequent ice stream oscillations with mean periods varying by ± 2 kyr (~ 20 % of characteristic period of the oscillations, Fig. 8 in Souček and Martinec (2011)). Similarly, Mantelli et al. (2016) find oscillations for an otherwise steadily streaming ice stream when surface temperature (within -22 to -32 °C) and accumulation rate (0.2 to 0.3 m yr^{-1}) are stochastically forced with time-correlated (red), low amplitude noise. Under such imposed climate variability, they find ice stream oscillation amplitudes larger than obtained for a constant climate forcing and exceeding the noise intensity. Additionally, 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 (weighted average of basal temperature at surrounding grid points vs. no temperature averaging). Note that their model is based on an Arakawa A grid (velocities and temperatures are calculated on the same node (Arakawa and Lamb, 1977)), enabling a thermal activation scheme without temperature averaging.

1.1 Research questions

Given the above context, in this paper we examine the effects of numerical choices on ice sheet surges as framed in the research questions detailed below. All research questions are investigated in Sec. 3 and a concise summary is provided in Sec. 4. We primarly use the 3D glacial systems model (GSM). To partly address potential non-linear dependencies of surge cycling on model parameters, we use a high variance subset of 5 base GSM parameter vectors (each comprising 8 model input parameters) for our numerical experiments.

To mitigate the possibility that these results are just due to specific numerical/modeling choices within the GSM, we repeat experiments that do not require source code changes 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, we do not intend to compare model results directly. Instead, we aim to increase confidence in model results by showing that the same conclusions can be drawn from two different models.

To provide a minimum estimate of the numerical noise in the models, we examine the differences in surge characteristics when applying a stricter numerical convergence in the GSM and adjusting the matrix solver used in PISM. These noise estimates set a minimum threshold for discerning physical significance of changes in surge characteristics due to physical model components (e.g., with and without subglacial hydrology): *Q1 - What is the threshold of numerical noise in the two models (Sec. 3.1.3)?*

An abrupt transition from hard bedrock to soft sediment 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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

surge characteristics by incorporating a smooth transition from $0\ \%$ sediment cover (hard bedrock) to $100\ \%$ sediment cover effectively changing the basal sliding coefficient $\boldsymbol{C}$ in Eq. (1b): *Q2 - Does the abrupt transition between a soft and hard bed significantly affect surge characteristics (Sec. 3.2.1)?* Given the topographic lateral bounds of Hudson Strait, we also examine: *Q3 - How does a non-flat topography affect the surge behavior (Sec. 3.2.1)?*

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 ($10\ \mathrm{K}$ colder minimum surface temperature, Calov et al. (2010))). We explore the role of the additional heat storage on surge activity by incorporating a $1\ \mathrm{km}$ deep bed thermal model in the GSM and PISM: *Q4 - Is the inclusion of a bed thermal model a controlling factor for surge activity (Sec. 3.2.2)?*

Ambiguity arises when determining the basal temperature at the grid cell interface. 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. We examine surge sensitivity to different interpolation schemes (see Sec. S3.2 for details on the three approaches): *Q5 - Do different approaches to determining the grid cell interface basal temperature significantly affect surge behavior, and if yes, which one should be implemented (Sec. 3.2.3)?*

At relatively coarse horizontal grid resolutions (e.g., $25\ \mathrm{km}$), the basal temperatures at the adjacent grid cell centers are of physical relevance as well. 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. (S5)). Therefore: *Q6 - 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. 3.2.4)?*

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 exists. Observational and experimental evidence for subtemperate 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., $25\mathrm{x}25\ \mathrm{km}$ in $1\ \mathrm{yr}$). 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. For higher horizontal grid resolutions, the average grid cell temperature is a better representation of the subgrid temperatures (e.g., Figs. 7, S27, and S28). Consequently, the thermal activation ramp should be sharper (smaller transition zone) for higher horizontal grid resolutions.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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 ($< 5°$C). A wide temperature ramp (e.g., $T_{ramp} = 1°$C, see Eq. (4)) enables an earlier sliding onset (for increasing basal temperature), spatially extended sliding, and a prolonged sliding duration (for decreasing basal temperature).

While on theoretical ground, the sensitivity of sliding speeds on temperature is expected to be strong near the pressure melting point, the appropriate functional form of the temperature ramp is not well constrained nor is the sensitivity to the functional form well documented. Herein, we use basal temperature gradients in high-resolution runs and approximations of the sub-grid warm-based connectivity between the faces of, e.g., a 25 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 apriori 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: *Q7 - How different are the model results for different basal temperature ramps and what ramp should be used (Sec. 3.2.5)?*

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 open review (Drew and Tarasov, 2022). Here we compare GSM surge statistics with and without a fully coupled 0D hydrology model: *Q8 - What is the effect of a simplified basal hydrology on surge characteristics (Sec. 3.2.6)?* PISM event characteristics are compared for local and mass-conserving horizontal transport hydrology models: *Q9 - How significant are the details of the basal hydrology model on surge characteristics in PISM (Sec. 3.2.6)?*

While both subglacial 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 subglacial hydrology is not an alternative for a basal temperature ramp. The ramp is still needed to prevent refreezing even when a description of subglacial hydrology is included (Mantelli et al., 2019). As such: *Q10 - 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. 3.2.7)?*

Incorporating the findings of the above experiments, we study numerical convergence with respect to horizontal grid resolution and time step for surge cycling. By convergence we mean decreasing differences between simulations when increasing the resolution or decreasing the time step. True model validation can only come from comparison with continuum model results, which are, however, only available to a very limited degree and do not encompass the process complexity considered here,

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

so we limit ourselves to a study of numerical convergence: *Q11 - Do model results converge (decreasing differences when increasing horizontal grid resolution and decreasing maximum time step, Sec. 3.3)?*

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 shallow ice physics (Tarasov et al., manuscript in preparation) is based on a slight variant of the ice dynamical core of Pollard and DeConto (2012). 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. 3.1.5). Ice dynamical time-stepping is subject to CFL (Courant–Friedrichs–Lewy) constraint (Courant et al., 1928) with further reductions upon ice dynamical solver convergence failure. A complete description of the GSM will soon be submitted for publication (Tarasov et al., manuscript in preparation).

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 temperature forcing is defined by a domain wide surface temperature (rTnorth, Tab. 1) and a specified vertical temperature gradient (atmospheric lapse rate (lapsr in Tab. 1)). The surface temperature forcing is asymmetric in time (Fig. S1), enabling 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 assymetric temperature forcing is shown in Fig. S2 for one parameter vector).

The GSM is initialized from ice-free conditions and the model runs last for $200\ \mathrm{kyr}$, of which the first $20\ \mathrm{kyr}$ are considered as spin-up interval, which is not included in the analysis. Note that this is a very conservative spin-up interval and that most runs reach their mean pseudo-Hudson Strait ice volume after $\sim 5\ \mathrm{kyr}$ (e.g., Fig. 4). 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 is $1\ \mathrm{yr}$ (automatically decreased as needed to meet CFL constraint or when convergence fails).

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 to date unfeasible for a glacial cycle contexts. Previous HE 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).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

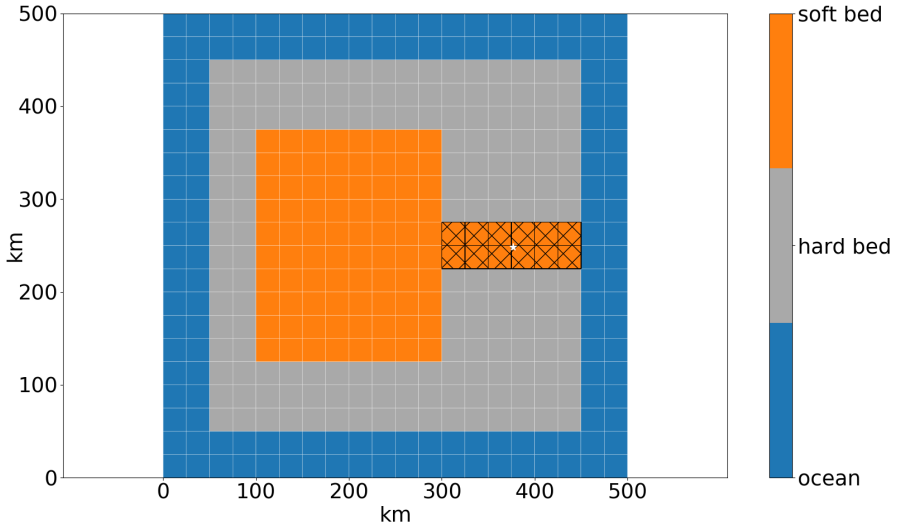

Figure 1. Modified ISMIP-HEINO geometry (Calov and Greve, 2006). The model domain is reduced to $500\text{x}500$ km to enable horizontal grid resolutions up to 3.125 km. The shown grid resolution is $25\text{x}25$ km. The basal topography is flat and the hatched area marks the soft-bedded pseudo-Hudson Strait. The white star indicates the location of the grid cell shown in Fig. 5+S22.

To improve on the drawbacks of the zeroth-order approximations, the GSM uses hybrid SIA/SSA ice dynamics. This heuristic combination links the two sets of equations by including a shear softening term in the calculation of the effective viscosity in the respective other set (SIA internal shear in the SSA viscosity calculation and SSA vertical-mean longitudinal stretching in the SIA viscosity calculation) (Pollard and DeConto, 2007, 2012). Additionally, horizontal shear and longitudinal stress gradient terms from the SSA equations reduce the driving stress in the SIA equations (Pollard and DeConto, 2007, 2012). A third coupling option adds the distinction between the depth-averaged internal-shear and basal velocity to the SSA basal stress term (Pollard and DeConto, 2007, 2012). This coupling term, however, tends to weaken numerical convergence without having much impact on ice sheet history and was therefore not used for the experiments in this paper. The hybrid SIA/SSA ice dynamics are activated for grid cells with a SIA velocity exceeding $30\ \text{m yr}^{-1}$ for soft beds and $200\ \text{m yr}^{-1}$ for hard beds. The full set of equations is described in Tarasov et al. (manuscript in preparation).

We set 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 warm-based area under an ice sheet and therefore sliding onset for coarse grid resolutions (Sec. 1.1). 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

$$\boldsymbol{u_b} = \boldsymbol{C_b}\,|\boldsymbol{\tau_b}|^{n_b-1}\,\boldsymbol{\tau_b} \tag{1a}$$

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

as a dependence of the basal sliding coefficient $\boldsymbol{C_b}$ on the estimated warm-based fraction of a grid cell (indirectly accounting for sub-temperate sliding) F_{warm} (Eq. (3))

$$\boldsymbol{C_b} = (1 - F_{warm})\, C_{froz} + F_{warm}\boldsymbol{C}, \tag{1b}$$

where $\boldsymbol{u_b}$ is the basal sliding velocity, $\boldsymbol{\tau_b}$ the basal stress, n_b the bed power strength (Tab. 1), and $\boldsymbol{C}$ the fully warm-based sliding coefficient (depends on the bed properties, see also Fig. S4). C_{froz} is the fully cold-based sliding coefficient for numerical regularization:

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

F_{warm} is calculated according to

$$F_{warm} = \max\left[0, \min\left(1, \frac{T_{bp,I} + T_{ramp}}{T_{ramp}}\right)\right]^{T_{exp}}, \tag{3}$$

where $T_{bp,I}$ is the grid cell interface basal temperature relative to the pressure melting point, negative T_{ramp} the temperature below which the entire grid cell is cold-based, and T_{exp} the exponent used for the ramp. The legacy values ($T_{ramp} = 1.0$ °C and $T_{exp} = 28$) were based on horizontal basal temperature gradients around the basal sliding activation zone and arguments of connectivity between grid cell interfaces (as basal sliding requires a connected subgrid warm-based path). Different values for T_{ramp} and T_{exp} are explored within this paper. T_{ramp} can be chosen as either constant or depending on the horizontal grid resolution (res, equal extent in x- and y-direction):

$$T_{ramp} = F_{T_{ramp}} \cdot \frac{\mathrm{res}}{50\ \mathrm{km}}\ ^{\circ}\mathrm{C} \tag{4}$$

This resolution dependence leads to a sharper temperature ramp for higher horizontal grid resolutions. The factor $F_{T_{ramp}}$ is used to conduct experiments with different temperature ramps at the same horizontal grid resolution (Sec. 3.2.5). The temperature ramps for all 4 horizontal grid resolutions and $F_{T_{ramp}} = 1$ (default value) are shown in Fig. 2. A temperature ramp similar to the one suggested by Fowler (1986) and later Mantelli et al. (2019)

$$F_{warm} = \exp\left(\frac{T_{bp,I}}{\delta}\right) \quad \text{for} \quad T_{bp,I} \leq 0 \tag{5}$$

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

The current idealized setup encompasses a maximum of 8 input parameters (Tab. 1). The 5 base 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 HE characteristics and soft bed sliding law exponent. Note that the soft and hard bed sliding law exponent in this study are equal (n_b in Tab. 1). Due to the significantly increased model run time, exponents larger than 3 are not considered

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

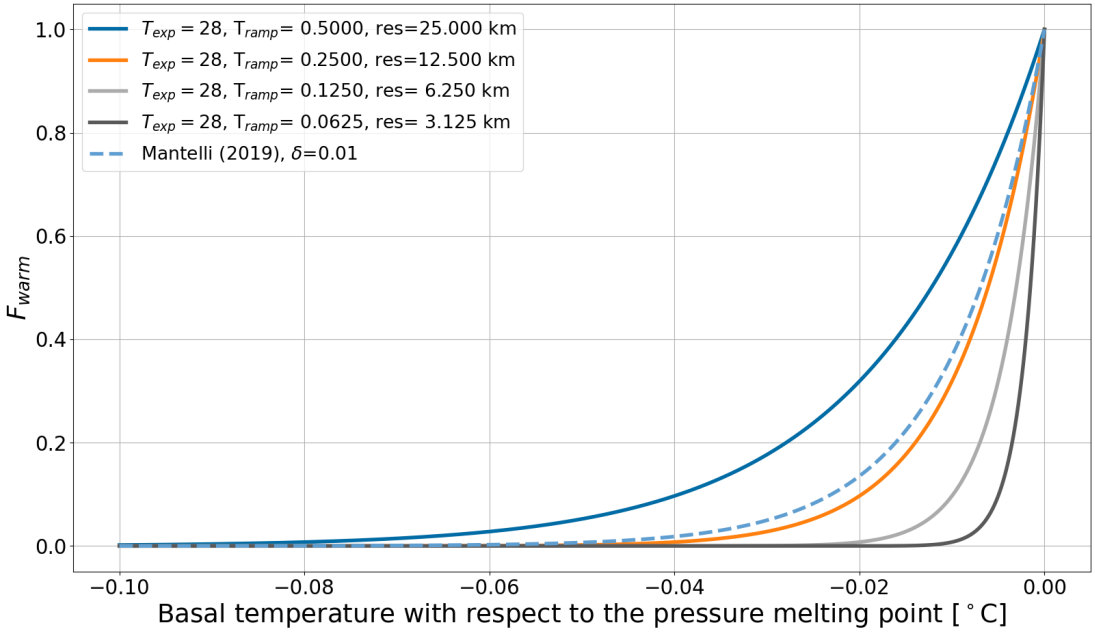

Figure 2. Temperature ramps for different values of T_{ramp} which depend on the horizontal grid resolution. A temperature ramp similar to the one suggested by Mantelli et al. (2019) (Eq. (5)) is shown for $\delta = 0.01$.

here. The GSM base setup used in this paper does not incorporate basal hydrology and glacial isostatic adjustment (GIA). Determining the effects of GIA is beyond the scope of this paper and will be addressed in future work. 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 a GSM setup with local basal hydrology (Eq. (S6) to (S8), Sec. 3.2.6) and examine resolution scaling (Sec. 3.3.2). However, experiments done with and without basal hydrology lead to qualitatively similar results (e.g., same conclusions from upscaling experiments in Sec. 3.2.5). We therefore omit subglacial hydrology coupling for the main analysis.

2.1.3 GSM model setups

The base setup (Tab. 2) has a $3.125\ \mathrm{km}$ horizontal grid resolution and $1\ \mathrm{year}$ maximum time step. 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 topography for the base setup, while the effect of a basal trough is investigated at a later stage (Sec. 3.2.1). Branching off this base setup, we carry out one at a time sensitivity experiments to isolate numerical and process impacts. These experiments in turn examine: sediment cover and non-flat topography (Sec. S3.1), a thin ($20\ \mathrm{m}$) bed thermal model, approach to determining the

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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	K km^{-1}
	PDDmelt	melt per Positive Degree Day (PDD)	0.005 - 0.012	m PDD^{-1}
	hpre	precipitation coefficient	0.02 - 0.2	($^{\circ}$C)$^{-1}$
	PrecRef	precipitation coefficient	1 - 3	m yr^{-1}
	rTnorth	northern surface temperature constant	-9 - -15	$^{\circ}$C
	n_b	soft and hard bed sliding law exponent, bed power strength	1 - 3	
Hydrology parameters	$\mathrm{h}_{\mathrm{wb,Crit}}$	effective bed roughness scale (Eq. (S6))	0.01 - 1	m
	rBedDrainRate	constant bed drainage rate	0.001 - 0.01	m yr^{-1}
	$N_{\mathrm{eff,Fact}}$	effective pressure factor (Eq. (S8))	$2 \cdot 10^4$ - $2 \cdot 10^5$	Pa
Additional parameters	$F_{T_{ramp}}$	basal temperature ramp scaling factor (Eq. (4))	0.125 - 16 (**1**)	
	T_{ramp}	basal temperature (with respect to the pressure melting point) at which sub-temperate sliding becomes important (Eq. (3), (4))	0.03125 - 1 (**0.0625**)	$^{\circ}$C
	T_{exp}	basal temperature ramp exponent (Eq. (3))	5 - 56 (**28**)	
	$W_{Tb,\mathrm{min}}$	weight of adjacent minimum basal temperature for basal sliding temperature ramp (Eq. (S5))	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 base setup are shown in the brackets for the additional parameters.

basal temperature at the grid cell interface (Sec. S3.2), weight of the adjacent minimum basal temperature for the basal sliding temperature ramp ($W_{Tb,\mathrm{min}}$, Sec. S3.3), basal temperature ramp (T_{ramp} and T_{exp}), horizontal grid resolution [25 km, 12.5 km, 6.25 km], maximum time step [0.5 year, 0.25 year], and local basal hydrology (Sec. S3.4). See Tab. 1 for details in parameter ranges.

2.2 PISM

2.2.1 PISM model description

The GSM is an ice sheet model developed specifically for glacial cycle ensemble modelling. The GSM is therefore numerically optimized for computational speed. To increase confidence in the GSM model results and reduce the model dependency of our analysis, experiments are also carried out with v2.0.2 of the Parallel Ice Sheet Model (PISM). Note that these experiments are not intended to be a direct comparison of the two models, but rather to show that the same conclusions can be achieved with different models (despite their differences in model setups, physics and numerics).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Similar to the GSM, PISM is a 3D thermodynamically-coupled ice sheet model. For further details on the model itself, refer to Bueler and Brown (2009); Winkelmann et al. (2011). The details on the default PISM setup, together with the default GSM values, are listed in Tab. 2. Given the higher computational cost of the 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 base setup is run at 25 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

$$|\boldsymbol{\tau_b}| \leq \tau_c \quad \text{and} \quad \boldsymbol{\tau_b} = -\tau_c \frac{\boldsymbol{u}}{|\boldsymbol{u}|} \quad \text{if} \quad |\boldsymbol{u}| > 0. \tag{6}$$

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.

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	1 yr	1 yr
number of cores/processes	1	8
ice dynamics	hybrid SIA/SSA	hybrid SIA/SSA (maximum SIA diffusivity of 1000 $m^2\ s^{-1}$)
sliding law	Weertman-type power law (Eq. (1a))	Coulomb friction law (Eq. (6))
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 base setup.

2.2.2 PISM ensemble input parameter vectors

The PISM configuration encompasses 6 model input parameters (Tab. 3). These parameters define the input fields for surface temperature, surface accumulation, and till friction angle. Similar to Calov and Greve (2006), the surface temperature at every grid cell is calculated as

$$T_{surf} = T_{min} + S_t \cdot d^3, \tag{7}$$

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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

$$d = \sqrt{(x - x_{center})^2 + (y - y_{center})^2} < R, \tag{8}$$

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.

$$B_{surf} = B_{max} - S_b \cdot d^5, \tag{9}$$

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.

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}	surface mass balance (accumulation/ablation) rate	50 - 450	kg m^{-2} yr^{-1}
	S_b	horizontal surface mass balance gradient	(0.15 - 1.00) $\cdot 10^{-11}$	kg m^{-2} yr^{-1} km^{-5}
	T_{min}	minimum surface temperature	220 - 245	K
	S_t	horizontal surface temperature gradient	(0.10 - 1.0) $\cdot 10^{-8}$	K km^{-3}
Constant parameters	x_{center}	location of the domain center in x-direction	1500	km
	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 (Tab. 3) were sampled via a Latin-Hypercube for specifying a 100 run ensemble. After filtering for runs that show oscillatory behavior, a 10-member high-variance subset was extracted by visual identification (Fig. S9).

2.2.3 PISM bed properties

Note that 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}) (Bueler and Van Pelt, 2015):

$$\tau_c = c_0 + \tan(\Phi)\, N_{\text{till}}, \tag{10}$$

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

where $c_0 = 0$ Pa is the till cohesion (Tulaczyk et al., 2000b). N_{till} is given by

$$N_{\text{till}} = N_0 \left(\frac{\delta_e P_0}{N_0} \right)^s 10^{\left(\frac{e_0}{C_c}\right)(1-s)}, \tag{11}$$

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 $\frac{W_{\text{till}}}{W_{\text{till}}^{\text{max}}}$ (Tulaczyk et al., 2000b; Bueler and Van Pelt, 2015). W_{till} and $W_{\text{till}}^{\text{max}} = 2$ 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., 2000b) 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., 2000b) 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, oscillations do not occur unless $\Phi < 1^\circ$ (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. S10 and S11). These values are all well within the ranges set by laboratory measurements. 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 the PISM default values.

However, these experiments incorporate very slippery beds enabling maximum sliding velocities of up to ~ 600 km yr^{-1} for single grid cells and time steps (Fig. S10). For comparison, observed surge velocities range from tens of meters per year for several years to hundreds of meters per day for short periods (K.M. Cuffey and W.S.B. Paterson., 2010).

Excluding runs that show maximum sliding velocities > 50 km yr^{-1} from the analysis yields similar results to the full 10-member ensemble (Sec. S6 and Fig. S12), indicating a stable solution of the numerical matrix solver even for runs with very high velocities. In addition, the 50 km yr^{-1} is exceeded no more than 7 times per 200 kyr run (100 yr output) and the maximum sliding velocities are generally within the observed range (Fig. S10).

2.2.4 PISM model setups

The default PISM setup (Tab. 2) is rerun for all 10 parameter vectors with different number of cores/processes ($n = 2, 4, 16, 32$), without a bed thermal model, different horizontal grid resolutions (50 km,12.5 km), a different maximum time step (0.5 and 0.25 year), and with a mass-conserving horizontal transport model for basal hydrology (Bueler and Van Pelt, 2015).

2.3 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 event characteristics. 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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

requirements and temporal resolution (e.g., Fig. S14 and S15). The Python analysis scripts are provided as supplementary material.

The quantities being analyzed are: the number of surges, the surge duration, the ice volume change during a surge. and the period between surges. The event time is defined as the time of minimum (pseudo-Hudson Strait) ice volume, and the duration of an event 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 (as per *scipy.signal.peak_widths*). The calculated ice volume change is the difference between the pre-event and minimum (pseudo-Hudson Strait) ice volume in that particular event (as per *scipy.signal.peak_prominences*). The period between events 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 events with a (pseudo-Hudson Strait) ice volume change of more than 500 km^3 and 40^4 km^3 are considered in the GSM and PISM analyses, respecitvely (~ 5 % of mean ice volume across all runs). A graphical illustration of the different event characteristics can be found in Fig. S18.

Additionally, the Root Mean Square Error (RMSE) and mean bias are calculated as a percentage deviation from the base (pseudo-Hudson Strait) ice volume time series for all setups (each parameter individually) and then averaged over the 5 parameter vectors (Eq. (S9) and (S10)). The full run time is considered (no spin-up interval).

We compare different model setups by calculating the difference between the base setup and all other setups for every parameter vector individually and then average this difference over the 5 parameter vectors. Crashed runs are not considered and runs with less than 2 events require special treatment (see Sec. S12 for further details on analysis).

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.

Due to the setup of the PISM model domain (Sec. 2.2.1 and S4), a large fraction of the pseudo-Hudson Strait area is ice-free when no surge occurs, leading to an inconsistency in the surge detection. This issue can be avoided 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. S9 for further details and a comparison between the two approaches.

3 Results

3.1 Key features of the base setup

3.1.1 Surge onset, propagation, and termination

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 3). 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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

01 of Hank (2023)) and Fig. 3. 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 SSA is activated (Sec. 2.1.1), and 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 the basal shear stress ($\sim 10^7$ J m^{-2} yr^{-1}) at $t = 6.70$ kyr. 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.

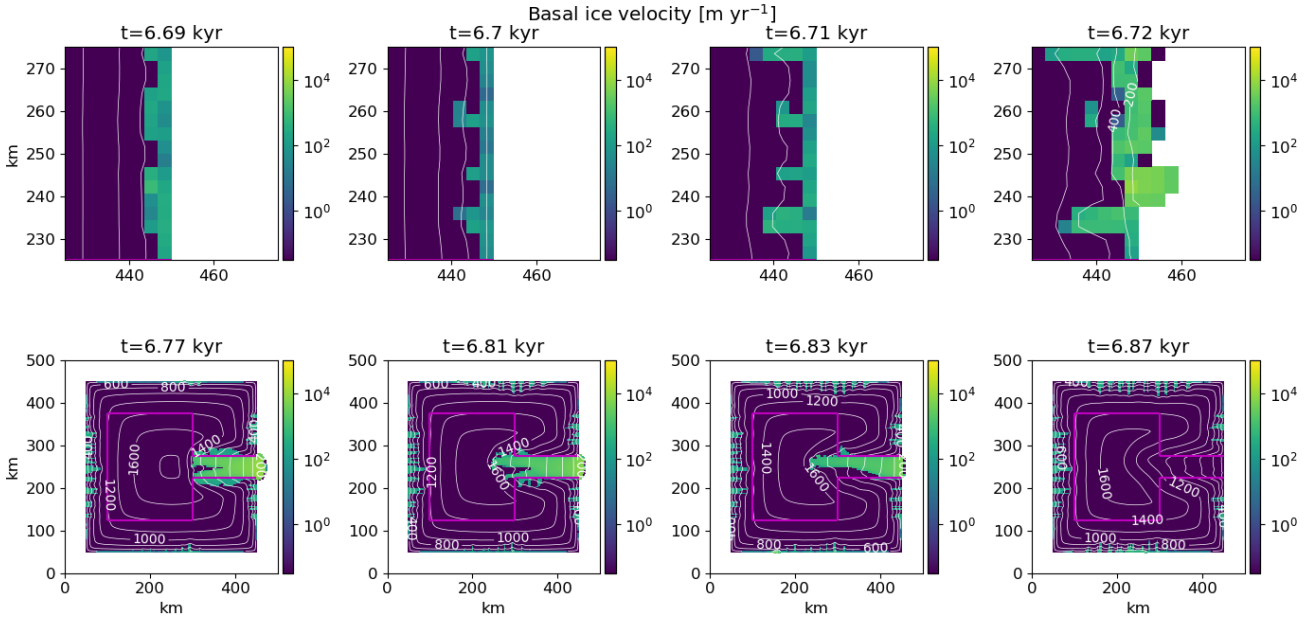

Figure 3. Basal ice velocity for parameter vector 1 at different time steps using the GSM. The horizontal grid resolution is 3.125 km and the maximum model time step is 1 yr. The contour lines show the ice sheet surface elevation in m. The magenta line outlines the soft-bedded pseudo-Hudson Bay and Hudson Strait. Note that the top and bottom rows show different areas of the domain, with the top being zoomed in.

The surge propagates nearly symmetrically until the pseudo-Hudson Bay area is reached ($t = 6.77$ kyr in Fig. 3 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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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 the induced asymmetry as a numerical noise induced bifurcation. We define the asymmetry as positive when the surge is stronger Northward (Fig. 3 and video 02 of Hank (2023)) or shifted Northward. The asymmetry sign varies across the first surge events (i.e., the event least biased by previous asymmetries) across the 5 base runs, ruling out any persistent numerical bias.

3.1.2 Event characteristics of the GSM and PISM base setup

Due to the differences in model setup, physics, and numerics (Tab. 2), the GSM and PISM base setup yield different event characteristics (Tab. 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)). Due to the downscaled GSM domain, 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). The mean modeled PISM period is within limits set by the literature (Hemming, 2004). 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 7.4 % of the mean (across base setup runs) maximum ice volume in the eastern half of the pseudo-Hudson Bay and pseudo-Hudson Strait.

Metric	GSM base setup	PISM base setup
number of events	180 ± 100	28 ± 17
mean period	1.1 ± 0.5 kyr	10 ± 12 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	$1.2 \pm 0.3 \cdot 10^5$ km^3

Table 4. Event characteristics of the GSM ($T_{ramp} = 0.0625°$C, $T_{exp} = 28$, $W_{Tb,\min} = 0.5$, TpmTrans for the interface calculation, sharp transition between hard and soft bed) and PISM base setup (Tab. 4). 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 basal temperature ramp in GSM uses $T_{ramp} = 0.0625$ and $T_{exp} = 28$ (black line in Fig. 2.)

3.1.3 Numerical noise estimation

Differences in event characteristics are considered significant when they exceed the numerical noise estimates given in Tab. 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 noise. Likely sources of the numerical noise are the iterative SSA solutions and floating point accuracy.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

To estimate the numerical noise, we re-run a set of GSM runs with 3.125 km horizontal grid resolution, imposing a stricter numerical convergence (decreasing final iteration thresholds, see Tarasov et al. (manuscript in preparation) for further details). The largest differences between simulations with base and stricter convergence thresholds occur for the mean period ($\sim 7 \pm 11$ %, Tab. 5). The standard deviations are on the same order of magnitude as the values themselves, indicating different responses across the 5 parameter vectors. Note that estimating the numerical noise 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 ($\sim 21 \pm 53$ %, Tab. S2).

Metric	base setup	stricter numerical convergence [% difference]
number of events	180 ± 100	-4.1 ± 4.9
mean period	1.1 ± 0.5 kyr	7.0 ± 10.6
mean duration	0.3 ± 0.1 kyr	2.5 ± 3.2
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	-1.1 ± 3.1

Table 5. Percentage differences of event characteristics between GSM runs with regular and stricter numerical convergence at 3.125 km. 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.

Numerical noise in PISM is estimated 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. S19). The largest differences occur for the mean period (~ 14 % for nCores$= 16$) and duration (~ 9 % for nCores$= 32$).

Setup	number of events	mean period	mean duration	mean ice volume change	nE1
25 km base setup	28 ± 17	10 ± 12 kyr	3 ± 2 kyr	$1.2 \pm 0.3 \cdot 10^5$ km^3	0
nCores$= 2$	0.8 ± 25.0	8.3 ± 45.0	-2.3 ± 19.8	1.5 ± 6.2	0
nCores$= 4$	1.9 ± 34.1	-7.0 ± 15.0	-8.8 ± 15.0	0.8 ± 4.4	1
nCores$= 16$	-1.2 ± 16.5	$\mathbf{14.3 \pm 46.1}$	6.2 ± 42.1	$\mathbf{-2.5 \pm 7.5}$	0
nCores$= 32$	$\mathbf{2.8 \pm 12.9}$	-6.3 ± 15.3	$\mathbf{-9.4 \pm 16.0}$	0.3 ± 12.7	0

Table 6. Change in event characteristics compared to the 25 km PISM base setup in percent (except first row). The values represent the average of 10 parameter vectors. No runs crashed and all runs showed at least 1 event. Runs with just one event (nE1) are ignored when caluclating 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 noise estimate for each event characteristic.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

3.1.4 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 surface temperature noise (± 0.1°C and ± 0.5°C) to the climate forcing does not significantly affect the event characteristics for the GSM (Tab. 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)).

3.1.5 Implicit thermodynamics/ice dynamics coupling

Here we test the impact of implicit coupling (via an iterative implementation) between the thermodynamics and ice dynamics in the GSM (see Tarasov et al. (manuscript in preparation) for details). The implicit coupling decreases the mean duration and pseudo-Hudson Strait ice volume change (~ -13 % and ~ -25 %, respectively). The number of events and mean period show no significant change (Tab. S3). While the changes in mean duration and pseudo-Hudson Strait ice volume change are larger than the numerical noise estimates, they do not justify an increase in run time of ~ 265 % and the implicit coupling is therefore omitted for the GSM base setup.

3.2 Sensitivity experiments

3.2.1 Smooth sediment transition zone and non-flat topography

The effects of a smooth sediment transition zone (instead of an abrupt transition from hard bedrock to soft sediment) and a non-flat topography on surge characteristics are examined here. The smooth transition zone does not significantly affect the surge characteristics. A non-flat topography leads to more, longer, and stronger (larger mean pseudo-Hudson Strait ice volume change) surges.

The abrupt transition from hard bedrock to soft sediment (pseudo-Hudson Bay and Hudson Strait) in the GSM base 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 surge events 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 event characteristics) but shows only minor differences for the major surge events (< 7.5 % for all event characteristics, Tab. 7). 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 events varies for different transition zones (RMSE ≤ 8 %, Fig. 4). A wider transition zone (more sediment surrounding the pseudo-Hudson Strait and Hudson Bay) generally favors an earlier sliding onset, but the details depend on the parameter vector in question (e.g., Fig. 4).

Imposing non-flat basal topography has a more significant effect than the sediment transition zone. In general, the number of events, mean duration, and mean pseudo-Hudson Strait ice volume change all increase compared to a flat topography (Tab. 7). Note that Tab. 7 also shows an increase in the mean period, but this is somewhat misleading due to the now early surges for

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

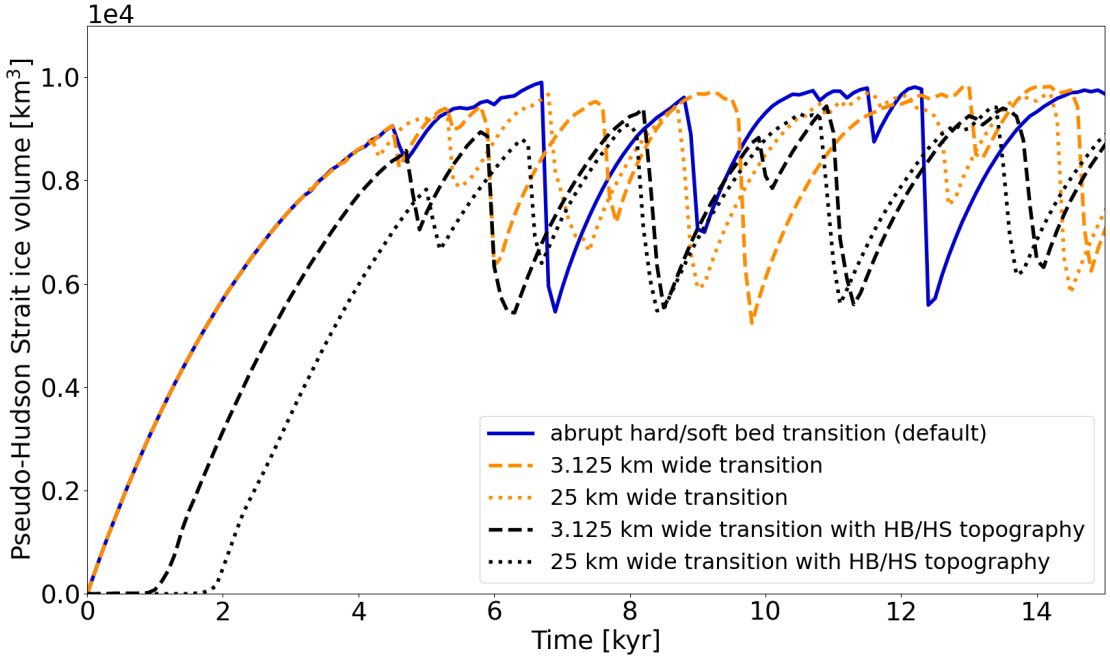

Figure 4. Pseudo-Hudson Strait ice volume for parameter vector 1 and three different bed configurations. The horizontal grid resolution is 3.1125 km. Note that the width of the topographical transition zone matches the width of the soft to hard bed transition zone. In experiments with a pseudo-Hudson 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.

parameter vector 0 and the subsequent large increase in the mean period (~ 200 %, no surges in the middle part of the run due to cold surface temperatures (Fig. S20)). 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 surge events. A wider transition zone (smaller slope) leads to fewer (difference of ~ 16 %) but stronger events (difference in mean pseudo-Hudson Strait ice volume change and mean duration of ~ 9 % and ~ 14 %, respectively, Tab. 7). A detailed comparison of an individual run is presented in Sec. S15.

The width of the transition zone (-200 m to sea level) affects the position and width of the major surge events (tendency towards wider surges for a wider transition zone (video 04 of Hank (2023))). The pseudo-Hudson Strait topography also suppresses the small surge events otherwise observed in the vicinity of the pseudo-Hudson Strait itself.

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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Metric	base 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 events	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 7. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base 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. 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 for the event characteristics (except for the RMSE and mean bias).

3.2.2 Bed thermal model

We examine the effects of a 1 km deep bed thermal model on 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 (base 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. 5).

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. S21) and more heat is available to warm the surrounding ice (no or smaller heat flux into the bed, Fig.S22). The additional heat increases the mean pseudo-Hudson Strait ice volume change and duration (~ 50 % and ~ 65 %, respectively, Tab. 8). 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 events drops ($\sim$ 32 %). These differences in event characteristics exceed the numerical noise estimates (Tab. 5). The stronger events (larger pseudo-Hudson Strait ice volume change) also lead to a mean ice volume bias of -2.3 % (Tab. 8).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

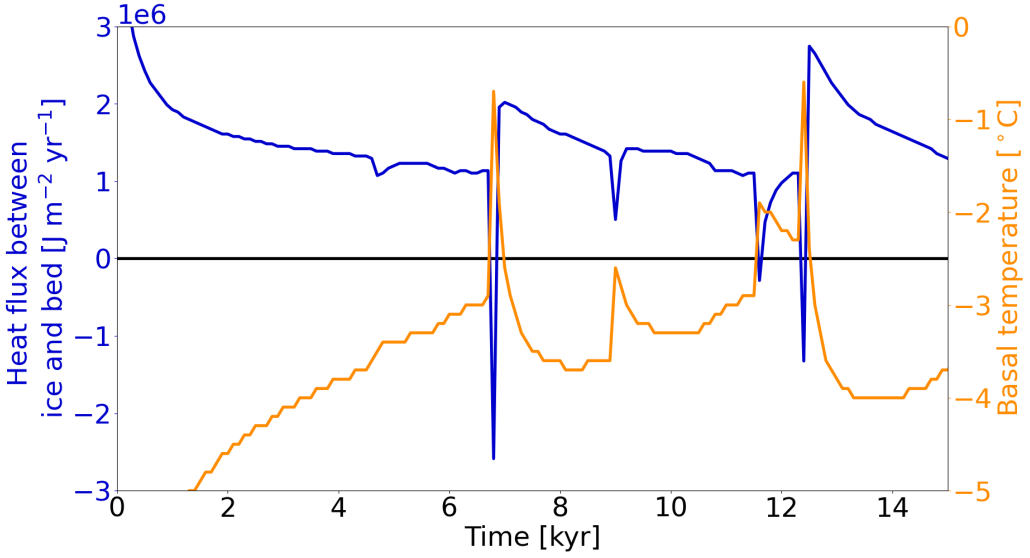

Figure 5. Heat flux at the base of the ice sheet (positive from bed into ice) and basal ice temperature for a grid cell in the center of the pseudo-Hudson Strait (grid cell center at $x = 376.5625$ km and $y = 248.4375$ km, white star in Fig. 1) and parameter vector 1 with the 1 km deep bed thermal model (17 non-linearly-spaced levels) using the GSM. The horizontal grid resolution is 3.125 km.

Metric	base setup	20 m deep (1 layer) bed thermal model
number of events	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 8. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup for runs with only one bed thermal layer (20 m deep). 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 for the event characteristics (not the RMSE and mean bias).

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 (~ 75 %, ~ 65 %, and ~ 375 %, respectively; Tab. 9). In contrast to the GSM characteristics, the number of events increases (~ 28 %) for runs without a bed thermal model. However, the standard deviation is large (~ 113 %) and the change in the number of events is somewhat misleading. The number of events decreases for 6 out of 10 runs. Parameter vectors

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

showing an increase in the number of events without a bed thermal model show very few surges (e.g., Fig. S23) or transition to a constantly active ice stream when the bed thermal model is included. As for the GSM, the stronger events lead to an overall smaller ice sheet in the surge affected-area (mean ice volume bias of ~ -27 %, Tab. 9).

Metric	base setup	no bed thermal model
number of events	28 ± 17	27.9 ± 113.3
mean period	10 ± 12 kyr	74.8 ± 83.6
mean duration	3 ± 2 kyr	63.2 ± 69.9
mean ice volume change	$1.2 \pm 0.3 \cdot 10^3$ km^3	374.5 ± 176.2
RMSE	-	36.5 ± 4.9
Mean Bias	-	-26.6 ± 4.9

Table 9. Percentage differences of event characteristics, ice volume RMSE and mean bias compared to the PISM base setup for runs without a bed thermal model. 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 for the event characteristics (not the RMSE and mean bias).

3.2.3 Basal temperature at the grid cell interface

Here we compare the effects of different approaches to determining the basal temperature at the grid cell interface. The GSM base setup (no hydrology) attempts to represent heat transfer from subglacial hydrology and ice advection by accounting for extra warming above the pressure melting point (TpmTrans, Sec. S3.2). This additional heat warms up adjacent colder ice. Without the extra warming, 4 out 5 parameter vectors do not show any surges when using TpmInt (Sec. S3.2). For the only run that still has cyclic behavior (parameter vector 1), the number of events decreases by ~ 84 % (note that runs without surges are considered for the number of events in Tab. 10). Using TpmInt with an upwind instead of a downwind scheme leads to slightly more events (difference of ~ 7 % and, therefore, on the same order of magnitude as the numerical noise estimate ($\sim 4 \pm 5$ %, Tab. 5)). Sporadic surges now occur in all but one run, leading to a large increase in the mean period (~ 1650 %, Tab. 10).

The most straightforward approach (TpmCen, Sec. S3.2) leads to ~ 75 % fewer events, and an increase in mean period and mean duration (~ 610 % and ~ 43 %, respectively). The mean pseudo-Hudson Strait ice volume change decreases (~ -61 %). Note that the TpmInt, TpmInt uwpind, and TpmCen event characteristics are difficult to compare due to the different number of runs considered (except for the number of events, decrease of ~ 97 % vs. ~ 90 % vs. ~ 75 %, respectively).

Due to significantly fewer events, the mean bias in pseudo-Hudson Strait ice volume increases for runs with TpmInt, TpmInt uwpind, and TpmCen (4 %, ~ 7 % and ~ 2 %, respectively). RMSE values are ~ 7 % for TpmInt and TpmCen, and ~ 9 % for TpmInt upwind (Tab. 10).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Metric	base setup (TpmTrans)	TpmInt	TpmInt, upwind	TpmCen
nC	0	0	0	1
nE0	0	4	1	0
nE1	0	0	1	0
number of events	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 10. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup for different approaches to calculate the basal temperature at the grid cell interface in percent (except first column). Crashed runs (nC) are not considered and runs without events (nE0) only contribute to the change in event number. Runs with only 1 event (nE1) 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 event characteristics (not the RMSE and mean bias).

3.2.4 Weight of adjacent minimum basal temperature

Here we compare the event characteristics for three different weights when calculating the basal interface temperature in the GSM (Eq. (S5)): no consideration of adjacent minimum basal temperature ($W_{Tb,\min} = 0.0$), basal temperature at the interface depends to 50 % on the adjacent minimum basal temperature at the grid cell center (base setup, $W_{Tb,\min} = 0.5$), and basal temperature at the interface is equal to the adjacent minimum basal temperature at the grid cell center ($W_{Tb,\min} = 1.0$).

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. 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, therefore, affects the surge propagation rather than blocking parts of the ice flow.

The surge cycling response to changes in $W_{Tb,\min}$ is not coherent (Tab. S4). For instance, the mean surge period increases for both $W_{Tb,\min} = 0.$ and $W_{Tb,\min} = 1.0$ compared to the base $W_{Tb,\min} = 0.5$. However, standard deviations are large, indicating a different model response for different parameter vectors.

3.2.5 Basal temperature ramps at different resolutions

Here we examine the effect of different basal temperature ramps at 3.125 km horizontal grid resolution and determine ramps for the coarse resolution runs that best match the 3.125 km model results (later used in Sec. 3.3.1). For coarse resolutions,

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

changing the basal temperature ramp can lead to a shift from oscillatory to non-oscillatory behavior (compare 25 km runs in Fig. S24 and 8).

When running the GSM at 3.125 km horizontal grid resolution, surge events 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. 6). The ice sheet takes longer to recover from the surge (binge-phase), increasing the mean period and decreasing the average number of events. The largest differences (compared to the base setup) in event characteristics occur for the widest ramp. Running the GSM without a basal temperature ramp leads to small but significant (according to Sec. 3.1.3) differences (compared to the base setup) in the mean duration (-7 ± 3 %) and mean pseudo-Hudson Strait ice volume change (-4 ± 6 %).

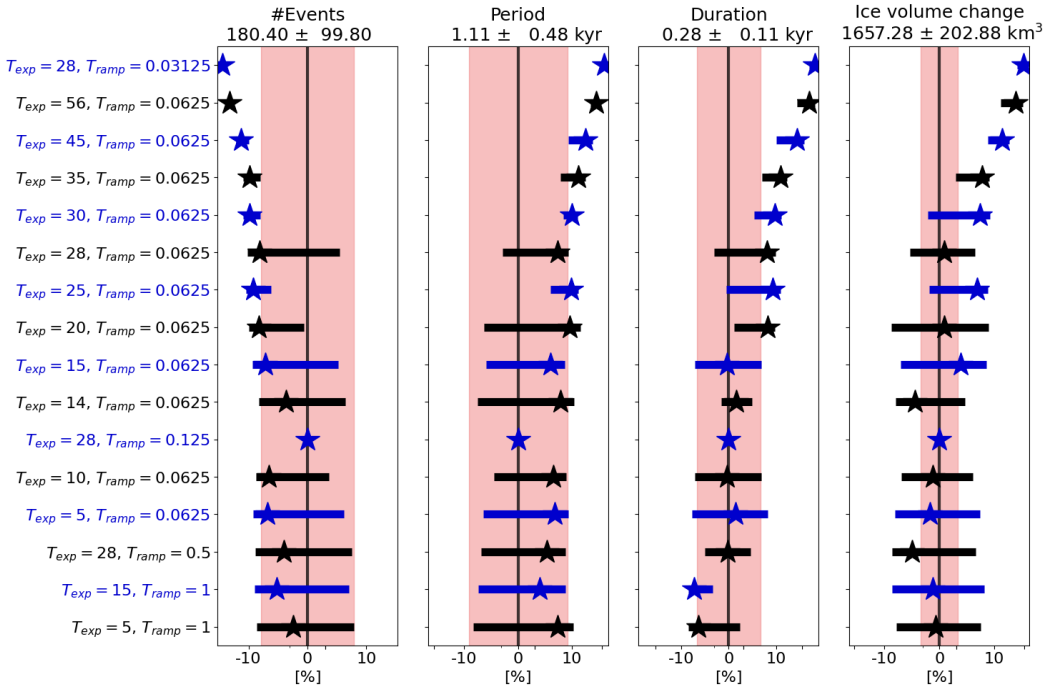

Figure 6. Percentage differences in event characteristics compared to the GSM base setup ($T_{ramp} = 0.0625$, $T_{exp} = 28$) for different basal temperature ramps at 3.125 km horizontal grid resolution. The ramps are sorted from widest (first row) to sharpest (last row, see Fig. S25 for a visualization of all ramps). The shaded pink regions mark the numerical noise estimates (Tab. 5) and the black numbers in the title of each subplot represent the mean values of the base setup. 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 x-axis is logarithmic. The exact values are given in Tab. S5.

All ramps (wider and sharper than base setup) show fewer events and a longer mean period than the base temperature ramp setup ($T_{ramp} = 0.0625$, $T_{exp} = 28$). However, changes in both event characteristics are relatively small (with standard deviations on the same order of magnitude) for ramps with a similar width to the base setup (Fig. S25). The mean duration and

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

mean pseudo-Hudson Strait ice volume change show a consistent response (increase/decrease for both event characteristics for a wider/sharper ramp) except for the four basal temperature ramps with the smallest difference to the base setup.

Except for the three widest ramps, the mean 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 (Tab. S5).

We compare the different temperature ramps at $25\ \mathrm{km}$, $12.5\ \mathrm{km}$ and $6.25\ \mathrm{km}$ horizontal grid resolution by calculating a single score for the mean and standard deviation of all event characteristics (Sec. S19). The ramps yielding the smallest differences compared to the $3.125\ \mathrm{km}$ base setup are listed in Tab. S6 and shown in Fig. S26. These results may be different for a different base setup (see Tab. 14 for a comparison of different base setups with local basal hydrology).

At $25\ \mathrm{km}$ 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 S6) differs by more than $50\ \%$ from the minimum sum of scores (bold numbers in last column in Table S6). The minimum scores for the mean and standard deviation occur for the same ramp ($T_{exp} = 5$, $T_{ramp} = 0.5$), clearly identifying it as the ramp that best resembles the $3.125\ \mathrm{km}$ horizontal grid resolution base runs. For the two higher horizontal grid resolutions, the minimum mean and standard deviation scores arise for different temperature ramps, preventing the determination of a single best ramp.

We complement the above analysis by upscaling the $3.125\ \mathrm{km}$ base runs. For example, a $25\mathrm{x}25\ \mathrm{km}$ grid cell contains a patch of 64 $3.125\mathrm{x}3.125\ \mathrm{km}$ grid cells. The scatter plot of the warm-based fraction (basal temperature with respect to the pressure melting point at $0\ ^\circ\mathrm{C}$) and the mean basal temperature with respect to the pressure melting point of the patch can be used to estimate the parameters T_{ramp} and T_{exp} of the basal temperature ramp (Eq. (3)). However, this does not account for the connectivity between the faces of, e.g., a $25\ \mathrm{km}$ 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, this estimate for 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). As such, the upscaling statistics only consider grid cells within the pseudo-Hudson Strait area during surges. Due to the limited storage capacity for the $10\ \mathrm{yr}$ output fields, only the first $10\ \mathrm{kyr}$ after the first surge are used for the upscaling experiments.

The upscaling results agree well with the score analysis at $25\ \mathrm{km}$ horizontal grid resolution. Both indicate that at this resolution, the ramp $T_{exp} = 5$, $T_{ramp} = 0.5$ (first row in Table S6, Fig. 7) gives results that best match those of the $3.125\ \mathrm{km}$ base run. The two approaches yield a similar range of temperature ramps at 12.5 and $6.25\ \mathrm{km}$ horizontal grid resolution, but the upscaling experiments generally favor wider temperature ramps (Table S6 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. (4), the upscaling experiments, therefore, provide a lower bound of $T_{exp} = 5$. The upscaling experiments with local basal hydrology lead to similar results.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

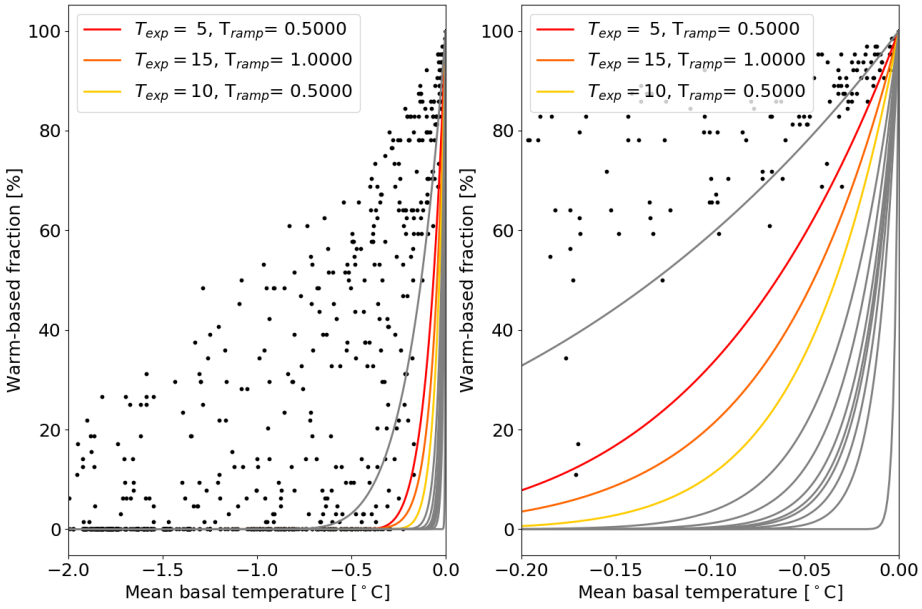

Figure 7. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 25 km horizontal grid resolution including all 5 parameter vectors using the GSM. Only grid cells 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 S6 and the gray lines show all other ramps that were tested at this resolution.

3.2.6 Basal hydrology

The effects of a simple local basal hydrology model in the GSM (Sec. S3.4) and the comparison of a local and horizontal transport model in PISM are examined here. When running the GSM with the local subglacial hydrology model, intermediate values are used for all 3 parameters (the effective bed roughness scale $\mathrm{h_{wb,Crit}} = 0.1$ m (Eq. (S6)), 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. (S8))) 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. (S8)) and, therefore, leads to fewer but stronger events. The results for $\mathrm{h_{wb,Crit}}$ 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 numerical noise estimates of Table 5.

Adding a local basal hydrology model to the GSM increases the mean ice volume change and duration by ~ 20 % and ~ 12 %, respectively (Table 11, exceeding the numerical noise estimates (Table 5)). The stronger surge events are due to the

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

reduction of effective pressure and, thus, increased sliding (Eq. (S8) and (1a)). The mean period increase (~ 17 %) while the number of events decreases (~ 4 %), but the standard deviations are large.

Since the local hydrology model effectively increases the basal sliding coefficient, we also compare setups with higher soft and hard bed sliding coefficients (C_{rmu} and C_{fslid} in Table 1, respectively) and no hydrology. Doubling the soft bed sliding coefficient leads to a similar model response but with a smaller increase in mean pseuod-Hudson Strait ice volume change (~ 11 % vs. ~ 20 %) and a stronger effect on the number of events and mean period (~ -10 % vs. ~ 4 %) than the local hydrology model. Increasing the hard bed sliding coefficient has no significant effect on the event characteristics (pseudo-Hudson Bay and Hudson Strait are soft-bedded, Table 11). Simultaneously increasing the soft and hard bed sliding coefficient yields similar results to increasing the soft bed sliding coefficient alone (not shown).

Metric	no hydrology	local hydrology	no hydrology, double C_{rmu}	no hydrology, double C_{fslid}
number of events	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 11. Percentage differences of event characteristics, ice volume RMSE and mean bias of GSM runs with a local basal hydrology model compared to runs without subglacial hydrology in percent (except first column). Additionally shown are the changes in event 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 event. The first 20 kyr of each run are treated as a spin-up interval for the event characteristics (not the RMSE and mean bias).

PISM experiments with a mass-conserving horizontal transport hydrology model yield similar results to the local hydrology model (Table 12). The mean duration, period, and ice volume change increase (0.5 %, 8.9 %, and 1.7 %, respectively), while the number of events decreases (3.5 %). These differences are generally within the numerical noise estimates (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 (+3.2 % and −0.3 %, respectively).

3.2.7 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 ($T_{ramp} = 0.001$, $T_{exp} = 28$), minimizing the smoothing effect of the basal temperature ramp. The change in event characteristics between runs with local basal hydrology and the sharp temperature ramp and the GSM base setup is similar (maximum difference of 3 %; compare Table 11 and S8) to the runs with local basal hydrology and the base basal temperature ramp ($T_{ramp} = 0.0625$, $T_{exp} = 28$), indicating that the local basal hydrology is the

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Metric	local hydrology	horizontal transport
number of events	28 ± 17	-3.5 ± 18.5
mean period	10 ± 12 kyr	8.9 ± 40.0
mean duration	3 ± 2 kyr	0.5 ± 23.4
mean ice volume change	$1.2 \pm 0.3 \cdot 10^3$ km^3	1.7 ± 13.6
RMSE	-	3.2 ± 3.1
Mean Bias	-	-0.3 ± 1.9

Table 12. Percentage differences of event 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 in percent (except first column). No runs crashed and all runs had at least 1 surge event. One run with the horizontal transport model showed just one event and was ignored when caluclating the change in mean period. The first 20 kyr of each run are treated as a spin-up interval for the event characteristics (not the RMSE and mean bias).

primary smoothing mechanism in both cases. The differences in the change of event characteristics between the base and the steeper ramp are smaller than the numerical noise estimates, preventing further analysis.

3.3 Resolution and time step dependence

In this section, we examine the horizontal grid resoltion and time step dependence of the GSM and PISM model results. Model results are considered as converging when the differences in event characteristics decrease with increasing horizontal grid resolutions and decreasing time steps. In general, both models show convergence, but the discrepancies between different horizontal grid resolutions are significant.

3.3.1 GSM convergence study without basal hydrology

In the GSM, significant surge pattern differences occur when computationally more feasible (coarser) horizontal grid resolutions are used ($\sim$ 200-fold increase in run time when increasing the horizontal grid resolution from 25 km to 3.125 km). 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. S24). Changing the basal temperature ramp can somewhat counteract this discrepancy by enabling basal sliding at lower basal temperatures for coarser grid resolutions (Fig. 8 and video 05 of Hank (2023)). Further details on discrepancies between horizontal grid resolutions for individual parameter vectors are discussed in Sec. S20.

We compare the differences in event characteristics for different basal temperature ramps at each resolution. We examine: a constant ramp ($T_{ramp} = 0.0625$, $T_{exp} = 28$), a resolution-dependent temperature ramp ($T_{exp} = 28$, Fig. 2), and the ramp with the smallest mean score (bold number in Table S6). 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. 25 km, 12.5 km, and 6.25 km runs show progressively smaller differences for the constant and resolution-dependent ramp, indicating model convergence (Table 13).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

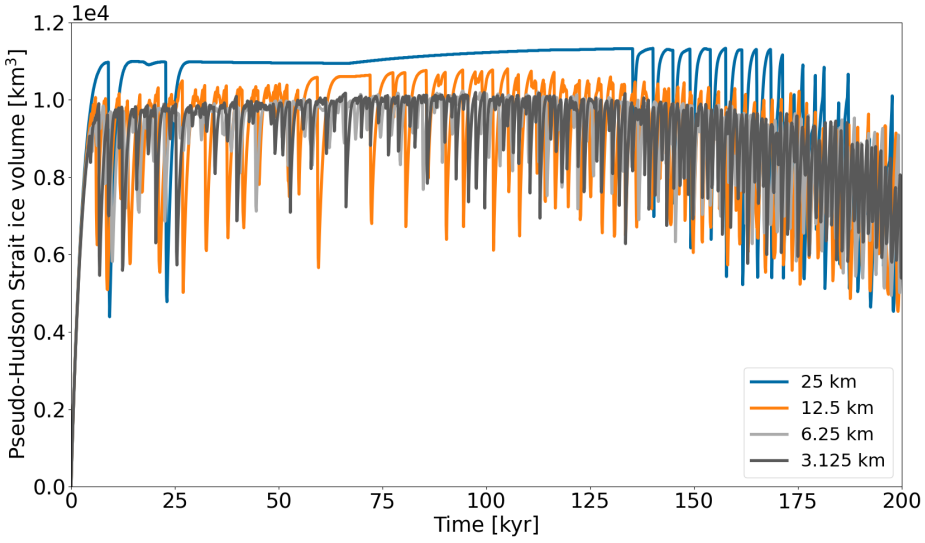

Figure 8. Pseudo-Hudson Strait ice volume for parameter vector 1 and different horizontal grid resolutions using the GSM. A resolution-dependent temperature (Eq. (4)) with $F_{T_{ramp}} = 1$ and $T_{exp} = 28$ is used for all horizontal grid resolutions (matching colors in Fig. 2).

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 S7). 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 event characteristics at 6.25 km and 12.5 km horizontal grid resolution (Table 13). At 25 km resolution, the ramp with the minimum mean score 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 scores (e.g., difference of 0.23, Table S6), it is unlikely that it is just a coincidence. However, the sensitivity of the event characteristics to grid refinement remains, no matter the choice of the temperature ramp, with differences significantly exceeding the numerical noise estimates from Sec. 3.1.3.

Experiments with different maximum time steps show only minor (less than 6 %) differences compared to the base runs for all event characteristics (Table 13). The pseudo-Hudson Strait ice volume RMSE values for smaller maximum time steps are on the same order of magnitude as the ones for 6.25 km grid resolution (~ 8 %), while the mean bias shows no difference for the given precision (Table S7). Overall, changes due to different maximum time steps are considerably smaller than for different horizontal grid resolutions and within the range of numerical noise (Table 5). Therefore, the implemented CFL condition is adequate for determining the ice dynamical time step, even though the condition is only sufficient for the solution of linear partial differential equations.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Setup	number of events	mean period	mean duration	mean pseudo-Hudson Strait ice volume change	nE0
3.125 km base 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, $T_{ramp} = 0.5$, $T_{exp} = 5$	-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, $T_{ramp} = 0.125$, $T_{exp} = 45$	-25.3 ± 13.6	37.9 ± 26.7	28.2 ± 7.0	9.8 ± 11.6	0
0.5 year maximum time step	-4.4 ± 4.5	5.4 ± 4.8	2.0 ± 2.4	-0.5 ± 5.6	0
0.25 year maximum time step	-1.8 ± 3.2	2.6 ± 4.4	-0.1 ± 3.4	-0.1 ± 3.4	0

Table 13. Change in event characteristics compared to the 3.125 km GSM base setup in percent (except first row). The values represent the average of 5 parameter vectors. No runs crashed and runs without events (nE0) only contribute to the change in event 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 ($T_{exp} = 28$) and constant ramp (black line, $T_{ramp} = 0.0625$, $T_{exp} = 28$) are shown in Fig. 2. The third ramp listed for each resolution is the ramp with the smallest mean score (Table S6).

3.3.2 GSM convergence study with basal hydrology

Since including a subglacial hydrology model significantly affects the event characteristics, we further examine the horizontal grid resolution scaling with a local basal hydrology model (Sec. S3.4). Based on the results without basal hydrology (Sec. 3.3.1), 5 basal temperature ramps ($T_{exp} = [5, 10, 15, 20, 28]$) with a resolution-dependent T_{ramp} (Eq. (4)) are tested for all resolutions. As it is unclear which basal temperature ramp should be used at the highest horizontal grid resolution (3.125 km), we test two different ramps ($T_{exp} = [5, 28]$). The experiments that yield the smallest differences in event characteristics (smallest mean score in Table S10 and S11) compared to the corresponding 3.125 km base runs (bold rows) are presented in Table 14.

Similar to the results without a basal hydrology model, the smallest differences in event characteristics (except the mean pseudo-Hudson Strait ice volume change) occur for the coarsest horizontal grid resolution (25 km, Table 14). This likely indicates that the optimal ramps at 12.5 and 6.25 km horizontal grid resolution have not been found.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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

Table 14. Change in event characteristics compared to the 3.125 km GSM setups with local basal hydrology (bold rows, $T_{exp} = [5, 28]$) for the ramps with the smallest mean score (analysis steps described in Sec. S19) in percent. 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.

In general, the resolution-dependent ramp with $T_{exp} = 5$ leads to the smallest differences between coarse and high-resolution runs. The differences in event 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 14 vs. 13), further underlining the importance of the basal hydrology.

Except for 12.5 km horizontal grid resolution, the resolution-dependent ramp with $T_{exp} = 5$ yields a self-consistent response across all resolutions. At 12.5 km, the next closest exponent ($T_{exp} = 10$) 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 numerical noise (DWINN). 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 event characteristics (Table S10 and S11). We rule out ramps for which the differences exceed the maximum numerical noise estimates (maximum of Table 5 and S2) for more than one event characteristic (DWINN failures).

Under these criteria and when using $T_{exp} = 5$ at 3.125 km horizontal grid resolution, the resolution-dependent ramp with $T_{exp} = 10$ remains within the DWINN ensemble for all resolutions (Table S11). The results for $T_{exp} = 28$ at 3.125 km horizontal grid resolution do not yield a single ramp that remains within the DWINN ensemble at all resolutions (Table S10). However, except for 6.25 km, for which the differences between the tested basal temperature ramps are the smallest, $T_{exp} = 5$ yields the minimum mean-score. The above analysis and the upscaling experiments in Sec. 3.2.5, therefore, suggest a resolution-dependent temperature ramp with T_{exp} between 5 and 10.

The pseudo-Hudson Strait ice volume RMSE and mean bias show convergence (smaller differences) for both 3.125 km horizontal grid resolution setups (Table S9).

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

3.3.3 PISM convergence study

Similar to the results presented for the GSM (Table 13 and 14), running PISM with different resolutions can lead to significant differences in surge behavior. In contrast to the GSM resolution scaling results, the PISM event characteristics do not show convergence for the three resolutions examined here (Table 15). Note that 4 out of 10 50 km runs crashed, indicating that this resolutions is too coarse to model binge-purge-type surges and skewing the statistics. The differences in event characteristics between the 12.5 km and 25 km PISM runs show a similar response than, e.g., the differences between the 3.125 km and 6.25 km runs in the GSM (Table 15 vs. Table 13).

The ice volume RMSE and mean bias converge (but not the event characteristics, see Table S12 and Fig. S30). The differences in event characteristics for different grid resolutions are, in general, larger than the numerical noise estimates (different numbers of cores), but can be smaller (mean period of the 50 km runs).

Setup	number of events	mean period	mean duration	mean ice volume change	nC	nE0
12.5 km base setup	$\mathbf{38 \pm 23}$	$\mathbf{6 \pm 5}$ **kyr**	$\mathbf{2 \pm 1}$ **kyr**	$\mathbf{1.2 \pm 0.2 \cdot 10^5}$ **km^3**	**1**	**1**
50 km	10.0 ± 47.3	9.6 ± 48.5	10.6 ± 40.2	29.6 ± 45.7	4	0
25 km	-27.5 ± 16.8	38.5 ± 40.7	21.7 ± 30.0	9.2 ± 14.4	0	0
25 km base setup	$\mathbf{28 \pm 17}$	$\mathbf{10 \pm 12}$ **kyr**	$\mathbf{3 \pm 2}$ **kyr**	$\mathbf{1.2 \pm 0.3 \cdot 10^5}$ **km^3**	**0**	**0**
0.5 year maximum time step	-2.0 ± 37.0	-6.8 ± 13.0	-8.6 ± 16.1	-0.3 ± 3.2	0	1
0.25 year maximum time step	-3.4 ± 16.5	-0.8 ± 25.8	1.5 ± 31.9	0.9 ± 6.6	0	0

Table 15. Change in PISM event characteristics due to different horizontal grid resolutions and maximum time steps. Note that the 12.5 km (highest resolution tested) is used as base for the grid resolution convergence study, whereas the 25 km setup is used for the maximum time step experiments. Except for the two rows containing the base setups (bold), all values are in percent. The values represent the average of 10 parameter vectors (8 for the resolution convergence study because one of the 12.5 km runs crashed and one did not show any oscillations (infinite difference)). Crashed runs (nC) are not considered and runs without events (nE0) only contribute to the change in event numbers. The first 20 kyr of each run are treated as a spin-up interval and are not considered in the above.

Similar to the results for different numbers of cores, small differences slowly accumulate for different maximum time steps, leading to a different pattern at the end of the run (e.g., Fig. S31). As for the GSM, different horizontal grid resolutions have a larger effect on surge behavior than the maximum time step for all event characteristics (Table 15).

4 Results Summary and Discussion

This section summarizes our modelling results in the context of the research questions outlined in Sec. 1.1.

Q1 - What is the threshold of numerical noise in the two models? When modeling binge-purge type surge events, numerical noise is apparent in both the GSM and PISM. The differences in event characteristics when applying a stricter numerical convergence criteria in the GSM can be as large as $\sim 7 \pm 11$ % (Table 5). Adjusting the matrix solver used in PISM (different

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

number of cores) leads to differences in event characteristics of up to $\sim 14 \pm 46$ % (Table 6). Consequently, the effects of physical model components cannot be determined if the differences in event characteristics are smaller than these numerical noise estimates and are considered insignificant.

In contrast to the findings of Souček and Martinec (2011), adding low levels of surface temperature noise does not significantly affect the GSM results (Table S3). Potential reasons for the different model responses are the use of an Arakawa A grid in Souček and Martinec (2011), the different implementations of the ice sheet dynamics, and the asymmetric in time temperature forcing in this paper (compared to a constant forcing in Souček and Martinec (2011)).

Q2 - Does the abrupt transition between a soft and hard bed significantly affect surge characteristics? The abrupt transition between the soft and hard bed sliding law is not the cause of the major surge events (Table 7). However, incorporating a smooth transition zone with two different widths (3.125 km and 25 km) does affect the location of proximal small-scale ice streams (video 03 of Hank (2023)).

Q3 - How does a non-flat topography affect the surge behavior? Adding a 200 m deep pseudo-Hudson Strait and Hudson Bay with a smooth transition zone and 500 m deep ocean displaces the origin of surges slightly further inland. Due to both the resultant warmer basal temperature and depressed presure melting point, the surges propagate faster, last longer, and evacuate more ice volume (Table 7). The 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.

Comparing the results for two different widths of the transition zone indicates fewer but larger events for a wider transition zone. Due to the gentle slope, the topography affects a larger area, increasing the width of the ice stream. More ice is available for evacuation, prolonging the surge and decreasing the 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 events than for the narrow transition zone (Table 7).

Q4 - Is the inclusion of a bed thermal model a controlling factor for surge activity? Including a 1 km deep bed thermal model significantly (according to Sec. 3.1.3) 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 event (Table 8 and 9). 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).

Q5 - 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 event characteristics. Without considering additional heat transfer to the grid cell interface (as an attempt to represent heat contributions from subglacial hydrology and ice advection), 4 out of 5 runs do not show any events, and the number of events for the remaining run decreases by ~ 97 % when using TpmInt (Table 10). The additional heat is, therefore, an essential component for modeling surge events in the GSM. Using an upwind instead of downwind scheme for TpmInt has no significant effect (Table 10).

Q6 - 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

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

for the basal sliding temperature ramp in the GSM yields a maximum difference of ~ 15 % (Table S4). These somewhat small effects on event 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 have little potential to affect (e.g., partly block) the ice flow.

Q7 - 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 differences in the period and amplitude of surges when using different implementations for thermal activation of basal sliding (the basal temperature ramp). A wider temperature ramp enables sliding onset at lower temperatures, fostering surge propagation and leading to stronger surge events. However, the choice of the most appropriate 3.125 km temperature ramp (Table S5) is unclear and identifying a single best ramp (fit of coarse resultions runs to 3.125 km runs) is challenging (Table S6). In general, a resolution-dependent ramp with T_{exp} between 5 and 10 (Eq. (3) and (4)) yields the smallest differences between high and low resolution simulations. Even at the highest tested horizontal grid resolution, running the GSM without a basal temperature ramp leads to significant (according to Sec. 3.1.3) differences in the mean duration and mean pseudo-Hudson Strait ice volume change, underlining the importance of the basal temperature ramp across all resolutions.

Q8 - What is the effect of a simplified basal hydrology on surge characteristics? The local basal hydrology model in the GSM increases the mean ice volume change, mean period, and mean duration while the number of events slightly decreases (Table 11). Somewhat stronger events are expected due to the reduction in effective pressure introduced by the subglacial water. Model runs without subglacial hydrology will therefore tend to underestimate the amplitude of surges (mean ice volume change and duration).

Increasing the soft bed sliding coefficient in model runs without basal hydrology has a smaller increase in the mean duration, mean period and pseudo-Hudson Strait ice volume change than including the local basal hydrology model, but a stronger effect on the number of events (Table 11). Therefore, simply changing the basal sliding coefficient cannot replace the basal hydrology model. The importance of subglacial hydrology has also been shown in several other studies examing the effects of ice sheet surges and ice streaming within a continuum model approach (Fowler and Johnson, 1995; Fowler and Schiavi, 1998; Benn et al., 2019, e.g.,).

Q9 - 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 (Table 12), indicating that the computationally more efficient 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 open review).

Q10 - 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? Event characteristics in runs with an active local basal hydrology and a sharp temperature ramp ($T_{ramp} = 0.001$, $T_{exp} = 28$, minimizing the smoothing effect of the basal temperature ramp, Table S8) show only minor differences compared to the GSM setup with a local hydrology and the base temperature ramp ($T_{ramp} = 0.0625$, $T_{exp} = 28$, Table 11). Once included, the local basal hydrology is the primary smoothing

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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 numerical noise prevents further analysis.

Q11 - Do model results converge (decreasing differences when increasing horizontal grid resolution and decreasing maximum time step)? Systematic grid refinement shows a converging overall ice volume (mean bias) in both models (Table S7+S9 and S12). Event characteristics, on the other hand, converge for a constant and resolution-dependent ramp in the GSM (Table 13), but not PISM (Table 15). 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. The highest horizontal grid resolution used for PISM is 4 times coarser than the highest resolution in the GSM (12.5 km vs. 3.125 km), which might explain why PISM results do not converge. In the GSM, the agreement between coarse and high-resolution runs can be significantly improved when applying a resolution-dependent temperature ramp (Table 13 and Sec. S19). GSM resolution scaling experiments with activated local basal hydrology lead to overall smaller differences (relative to the 3.125 km reference simulations) in event characteristics than without (Table 14 vs. 13).

Event characteristics in both the GSM and PISM show a strong resolution dependence for all sensitivity tests (Table 13+14 and 15). This is in contrast to the findings of other studies examining thermally induced ice streaming Hindmarsh (2009); Brinkerhoff and Johnson (2015). However, both of these studies analyze just one parameter vector, and it is relatively easy to find a parameter vector for which, e.g., the GSM exhibits only a minor resolution dependence. 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. (5)) and focuses on steady ice streams, not binge-purge-type surges. Over 200 kyr, even minor differences at the beginning of a run can slowly accumulate and yield overall different surge characteristics (e.g., Fig. S31). Furthermore, Brinkerhoff and Johnson (2015) examine ice stream statistics over the whole domain and not a specific soft-bedded region. Neither Hindmarsh (2009) nor Brinkerhoff and Johnson (2015) consider a bed thermal model.

Decreasing the maximum time step leads to only minor (< 9 %, Table 13 and 15) changes in event characteristics for both PISM and the GSM. This is in agreement with the findings of earlier studies (Greve and MacAyeal, 1996; Greve et al., 2006; Takahama, 2006). However, individual parameter vectors might still show a different surge pattern (e.g., Fig. S31).

5 Conclusions

We investigate the effect of ice sheet model numerics on surge characteristics and determine key model components for simulating binge-purge type surge events. Numerical noise estimates (differences in surge characteristics of up to ~ 15 %) are used to discern the physical significance of the process in question. For some experiments (e.g., the weight of the adjacent minimum basal temperature), the numerical noise estimates are on the same order of magnitude or larger than the modeled differences in surge characteristics, hindering the analysis of the underlying physical process.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Experiments showing only minor changes in surge characteristics include: a mass-conserving horizontal transport hydrology model (instead of a local hydrology model), a smoothed transition between regions of soft sediment and hard bedrock (instead of an abrupt transition), and smaller (than $1\ \mathrm{yr}$) maximum time steps in the CFL condition.

On the other hand, surge characteristics are sensitive to the basal sliding activation function and show a strong resolution dependency. Since both the GSM and PISM show a resolution dependency, it is likely that it also exists in other ice sheet models with similar approximations. Incorporating a resolution-dependent basal temperature ramp for basal sliding thermal activation reduces the resolution dependency in the GSM. Based on our results, we suggest that those interested in modeling ice stream cycling at horizontal grid resolutions $> 3\ \mathrm{km}$ should use a resolution-dependent ramp with $T_{exp} = 10$ as a base test configuration. However, we strongly recommend resolution testing to determine the configuration with the smallest resolution dependency. Additionally, our results indicate that modelling of binge-purge surge instabilities that aims to reflect the physical behaviour of actual ice streams should include a non-flat topography, a bed thermal model, and a basal hydrology model.

The key takeaway of this study is the numerical sensitivities that must be considered when numerically modeling binge-purge-type oscillations of ice streams. Our analyses offer guidance in minimizing these sensitivities for research contexts that limit horizontal grid cell resolution to larger than about $3\ \mathrm{km}$. Significant (albeit smaller) numerical sensitivities to the choice of thermal activation ramp remain at our highest tested horizontal grid resolution ($3.125\ \mathrm{km}$). Analytical examination (where possible) and/or higher-resolution numerical modeling with higher-order glaciological models is needed to verify that modeling approaches represent the actual physical system for this context.

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.7668490 (Hank, 2023).

Author contributions.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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.

785

Acknowledgements. The authors thank *Andy Aschwanden*, *Ed Bueler* and *Constantine Khrulev* for support with the Parallel Ice Sheet Model (PISM). We thank *Ed Bueler* for a fruitful discussion about the bed thermal model. This research has been supported by the Bundesministerium für Bildung und Forschung (Research for Sustainability initiative, FONA, through the PalMod project) and an NSERC Discovery Grant held by LT.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

References

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., Ganopolski, A., Petoukhov, V., Claussen, M., and Greve, R.: Large-scale instabilities of the Laurentide ice sheet simulated in a fully coupled climate-system model, Geophysical Research Letters, 29, 1–4, https://doi.org/10.1029/2002GL016078, 2002.

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.

Feldmann, J. and Levermann, A.: From cyclic ice streaming to Heinrich-like events : the grow-and-surge instability in the Parallel Ice Sheet Model, The Cryosphere, 11, 1913–1932, https://doi.org/10.5194/tc-11-1913-2017, 2017.

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.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

Fowler, A. C. and Johnson, C.: Hydraulic run-away: a rn.echanisrn. for ther:rnally regulated surges of ice sheets, 1, 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. and MacAyeal, D. R.: Dynamic/thermodynamic simulations of Laurentide ice-sheet instability, Annals of Glaciology, 23, 328–335, https://doi.org/10.3189/S0260305500013604, 1996.

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 ice sheet instabilities such as binge- purge type cyclic ice stream surging", https://doi.org/10.5281/zenodo.7668490, 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.

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.

Marshall, S. J. and Clarke, G. K. C.: A continuum mixture model of ice stream thermomechanics in the Laurentide Ice Sheet 2. Application to the Hudson Strait Ice Stream, Journal of Geophysical Research: Solid Earth, 102, 20 615–20 637, https://doi.org/10.1029/97jb01189, 1997.

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.

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

Papa, B. D., Mysak, L. A., and Wang, Z.: Intermittent ice sheet discharge events in northeastern North America during the last glacial period, Climate Dynamics, 26, 201–216, https://doi.org/10.1007/s00382-005-0078-4, 2006.

Payne, A. J.: Limit cycles in the basal thermal regime of ice sheets, 1995.

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.

https://doi.org/10.5194/egusphere-2023-81
Preprint. Discussion started: 24 February 2023

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.

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.

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.

Steen-Larsen, H. C. and Dahl-Jensen, D.: Modelling binge-purge oscillations of the Laurentide ice sheet using a plastic ice sheet, Annals of Glaciology, 48, 177–182, https://doi.org/10.3189/172756408784700635, 2008.

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.

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.