Supplement to "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)

S1 GSM - Climate forcing

The asymmetric temperature forcing in the GSM (orange line in Fig. S1) is calculated according to

$$T_{asym} = \left| \left(\frac{t}{200\ \mathrm{kyr}} \cdot 3 + 2 \right) - 1 \right| \cdot 5 \tag{S1}$$

, where t is the model time ranging from -200 kyr to 0 kyr (instead of 0 kyr to 200 kyr).

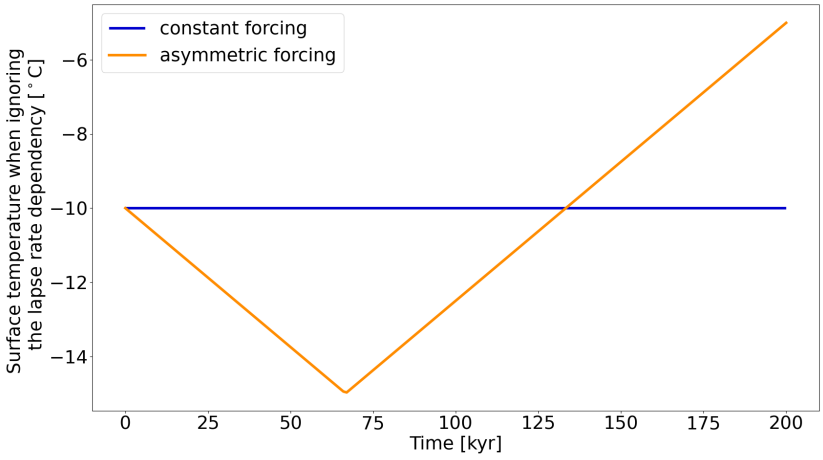


Figure S1. Constant and asymmetric temperature forcing in the GSM. The coldest temperature is reached at 66.7 kyr. For the case shown here, the surface temperature is set to rTnorth$= -10°$C (Tab. 1). All model runs within this paper use the asymmetric forcing.

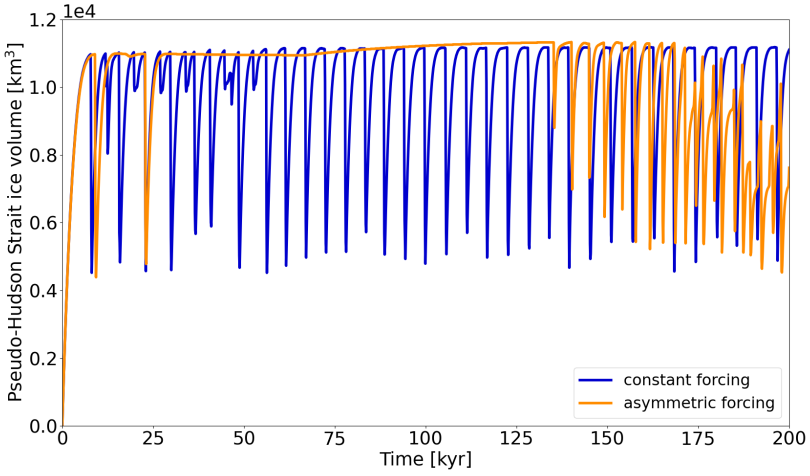


Figure S2. Pseudo-Hudson Strait ice volume for a constant and asymmetric temperature forcing in the GSM. This plot shows parameter vector 1 with a horizontal grid resolution of 25 km.

S2 GSM - Parameter vectors

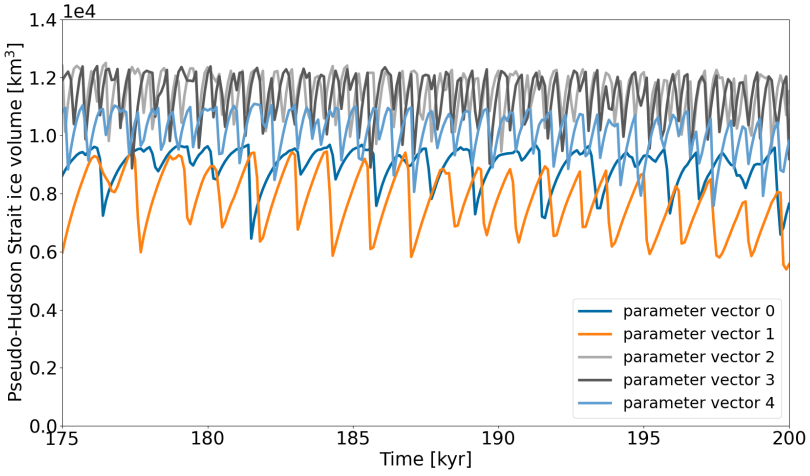


Figure S3. Pseudo-Hudson Strait ice volume for the last 25 kyr of all 5 GSM parameter vectors when using the base setup. Note that only the last 25 kyr are shown for better visibility of the individual oscillation pattern.

S3 GSM - Details of different model setups

S3.1 Bed properties

The effects of an abrupt transition from hard bedrock (0 % sediment cover) to soft sediment (100 % sediment cover) are examined by adding a smooth transition zone (Fig. S5 a)). Two widths of this transition zone (25 km and 3.125 km) are investigated. The basal velocity (or more precisely the sliding coefficient $\boldsymbol{C}$ in Eq. (1b)) then depends on the sediment cover within a grid cell (Fig. S4). In the experiments with a non-flat topography, the bed of the pseudo-Hudson Bay and Hudson Strait is placed 200 m and the surrounding ocean 500 m below the sea level (Fig. S5 b). The topographic transition zones (25 km and 3.125 km wide) align with the sediment transition zones.

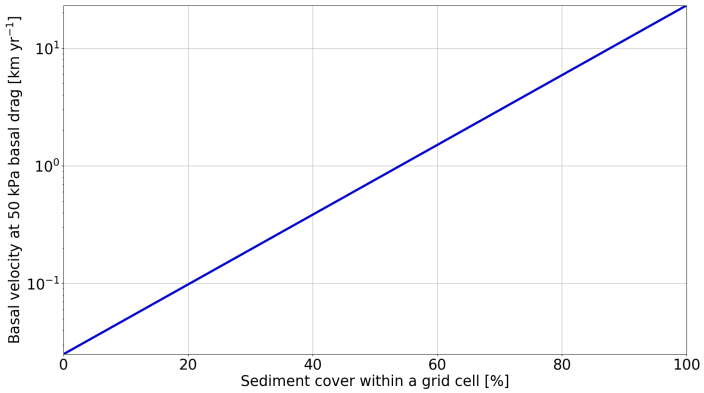


Figure S4. Basal velocity at 50 kPa basal drag for variable sediment cover and a power-law exponent of 3 (n_b in Tab. 1).

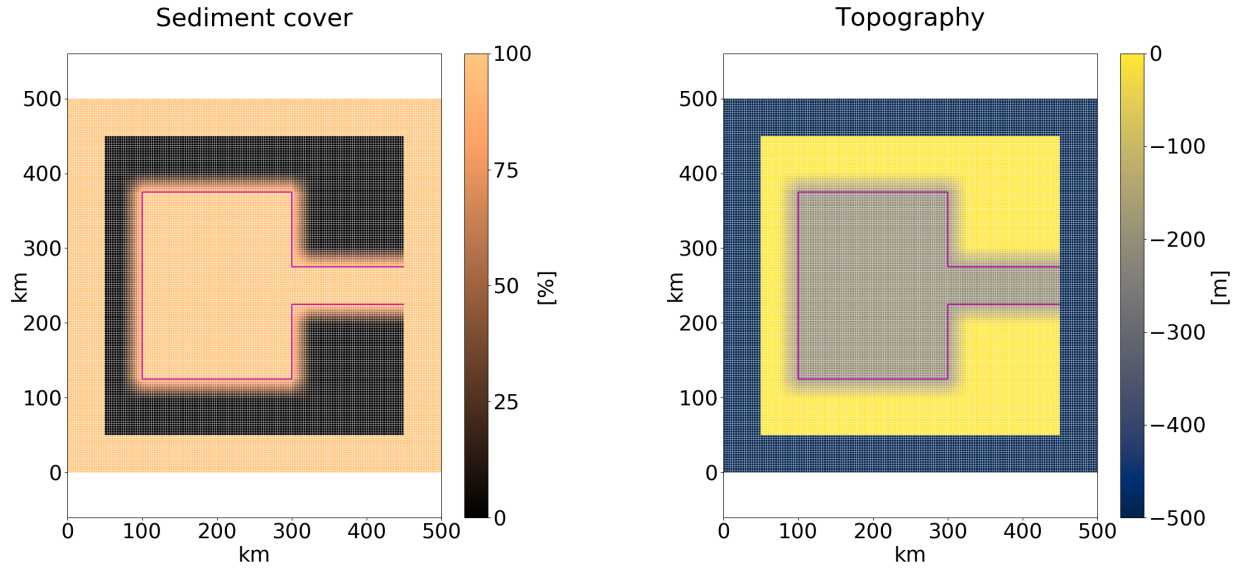
S3.2 Approches to determine the basal temperature at the grid cell interface

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

$$T_{bp,I} = 0.5 \cdot (T_{bp,L} + T_{bp,R}) \tag{S2}$$

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

TpmInt, on the other hand, calculates the basal temperature at the **Int**erface (T_I) by averaging the adjacent grid cell center basal temperatures (T_L and T_R, Eq. (S3a)). $T_{bp,I}$ is then determined by using the interface ice sheet thickness (average of



(a) (b)

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

adjacent grid cell center ice thicknesses H_L and H_R, Eq. (S3b)).

$$T_I = 0.5 \cdot (T_L + T_R) \tag{S3a}$$

$$T_{bp,I} = T_I + \beta_P \frac{H_L + H_R}{2} \tag{S3b}$$

where $\beta_P = 8.7 \cdot 10^{-4}\ ^\circ\mathrm{C\ m^{-1}}$ is the standard basal melting point depression coefficient.

The last approach (TpmTrans) accounts for extra heat available beyond that needed to reach the local pressure melting point (T_{add}).

$$T_{Im,C} = T_C + T_{add} \tag{S4a}$$

, where T_C and $T_{Im,C}$ are the basal temperature at the grid cell center and the basal temperature in the intermediate calculation step, respectively. The basal temperature with respect to the pressure melting point at each adjacent grid cell center $T_{bp,Im,C}$ is then calculated using the interface ice thickness.

$$T_{bp,Im,C} = T_{Im,C} + \beta_P \frac{H_L + H_R}{2} \tag{S4b}$$

In the intermediate steps to calculate the interface temperature (Eq. (S4a) and (S4b)), $T_{Im,C}$ and $T_{bp,Im,C}$ are allowed to exceed the pressure melting point. This temporary higher basal temperature is an attempt to account for heat transported to the interface by ice advection and basal water.

$$\text{IF } T_{bp,Im,C} > 0°\text{C}: \quad T_{bp,Im,C} = \min\left(0.5°\text{C}, 0.5 \cdot T_{bp,Im,C}\right) \tag{S4c}$$

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

$$T_{bp,I} = 0.5 \cdot \left(T_{bp,Im,L} + T_{bp,Im,R}\right) \tag{S4d}$$

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

S3.3 Weighting function of the adjacent minimum basal temperature

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

$$T_{bp,I} = W_{Tb,\min} \cdot \min\left[T_{bp,L}, T_{bp,R}\right] + T_{bp,I} \cdot \left(1 - W_{Tb,\min}\right) \tag{S5}$$

, where $T_{bp,I}$ is the basal temperature with respect to the pressure melting point at the grid cell interface, and $T_{bp,L}$ and $T_{bp,R}$ are the basal temperatures with respect to the pressure melting point at the adjacent grid cell centers. Note that $T_{bp,Im,L}$ and $T_{bp,Im,R}$ instead of $T_{bp,L}$ and $T_{bp,R}$ are used when calculating $T_{bp,I}$ according to TpmTrans (Eq. (S4)). In this way, the additional heat T_{add} is still considered even when $W_{Tb,\min} = 1$.

S3.4 Local basal hydrology

The local basal hydrology sets the basal water thickness by calculating the difference between the basal melt rate and a constant basal drainage rate (rBedDrainRate in Tab. 1). This subglacial hydrology provides a simple and computationally efficient way to capture changes in basal sliding velocities due to effective pressure variations (Drew and Tarasov, 2022, under open review). However, it does not account for basal ice accumulation, englacial or supraglacial water input, or horizontal water transport.

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

$$N_{\mathrm{C,eff}} = 1 - \min\left(\frac{\mathrm{h_{wb}}}{\mathrm{h_{wb,Crit}}}, 1.0\right)^{3.5} \tag{S6}$$

The basal water thickness is limited to $\mathrm{h_{wb,Crit}} = 10\ \mathrm{m}$ and is set to $\mathrm{h_{wb}} = 0\ \mathrm{m}$ where the ice thickness is less than $10\ \mathrm{m}$ and where the temperature with respect to the pressure melting point is below $-0.1°\mathrm{C}$. Experiments with $\mathrm{h_{wb,Crit}} = 5\ \mathrm{m}$ yield the same results, and removing all the water for $H < 1\ \mathrm{m}$, $H < 50\ \mathrm{m}$, and $T_{bp} < -0.5°\mathrm{C}$ does not significantly (according to Sec. 3.1.3) affect the model results. The effective pressure at the grid cell interface is then

$$N_{\mathrm{eff}} = g\rho_{\mathrm{ice}} \cdot 0.5\left(H_L N_{\mathrm{C,eff,L}} + H_R N_{\mathrm{C,eff,R}}\right) \tag{S7}$$

, where $g = 9.81\ \mathrm{m\ s^{-2}}$ is the acceleration due to gravity, $\rho_{\mathrm{ice}} = 910\ \mathrm{kg\ m^{-3}}$ the ice density, H the ice thickness and the subscripts L and R denote the adjacent grid cells to the left and right of the interface, respectively (similarly for upper and lower grid cells adjacent to a horizontally aligned interface). We enforce that N_{eff} never falls below $10\ \mathrm{kPa}$ (denominator in Eq. (S8), similar results for $N_{\mathrm{eff,min}} = 5\ \mathrm{kPa}$). Finally, the effective pressure of each grid cell alters the basal sliding coefficient in the sliding law (Eq. (1a)) according to

$$C_b = C_b \cdot \min\left(10, \max\left(0.5, \frac{N_{\mathrm{eff,Fact}}}{N_{\mathrm{eff}} + 10^4\ \mathrm{Pa}}\right)\right) \tag{S8}$$

, where $N_{\mathrm{eff,Fact}}$ is the effective pressure factor (Tab. 1). The change of the basal sliding coefficient C_b is, therefore, limited to $C_b \cdot 0.2$ to $C_b \cdot 10$. Allowing a larger change of $C_b \cdot 0.1$ to $C_b \cdot 20$ does not significantly (according to Sec. 3.1.3) change the model results.

S4 PISM - Input fields

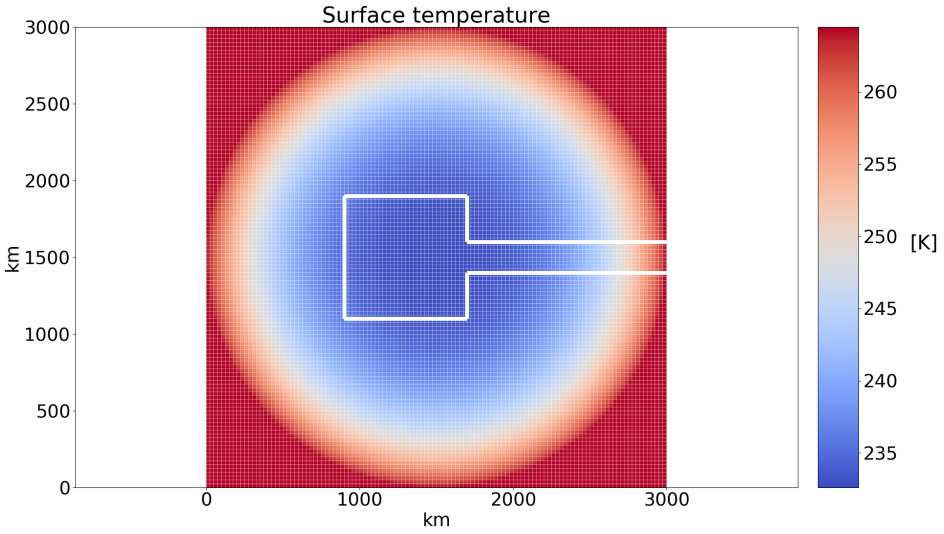


Figure S6. PISM surface temperature input field for parameter vector 1. The corresponding parameter values of T_{min} and S_t are 232.60 K and $9.45 \cdot 10^{-9}$ K km^{-3}, respectively. Thick white lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution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

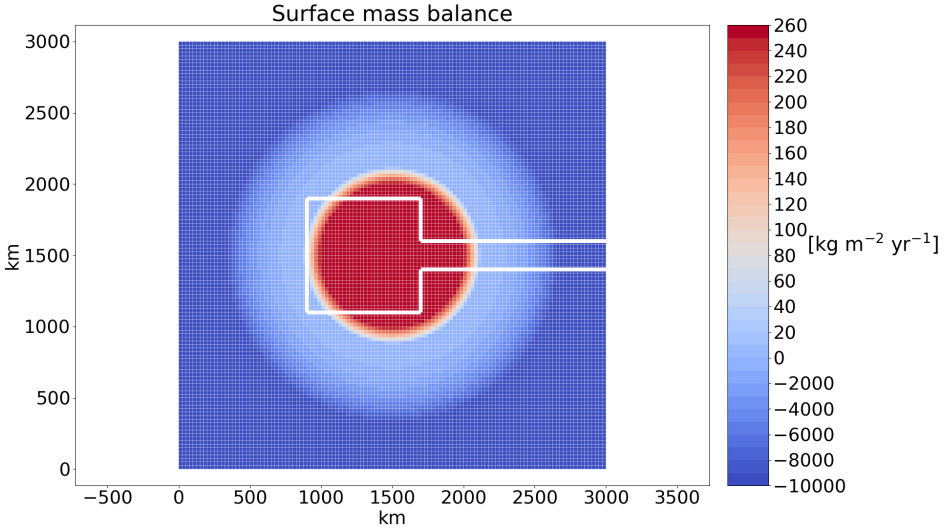


Figure S7. PISM surface mass balance input field for parameter vector 1. The corresponding parameter values of B_{max} and S_b are 408.81 kg m^{-2} yr^{-1} and $4.55 \cdot 10^{-12}$ kg m^{-2} yr^{-1} km^{-5}, respectively. Thick white lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson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

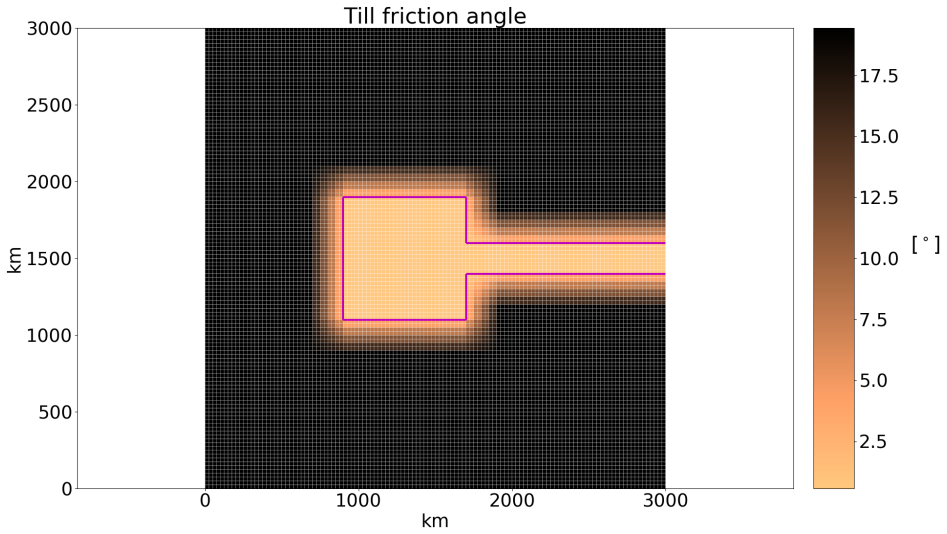


Figure S8. PISM till friction angle input field for parameter vector 1. The corresponding parameter values of $soft$ and $hard$ are $0.56°$C and $19.44°$C, respectively. Magenta lines outline the simplified soft-bedded pseudo-Hudson Bay/Hudson Strait area. The horizontal grid resolution is 25x25 km.

S5 PISM - Parameter vectors

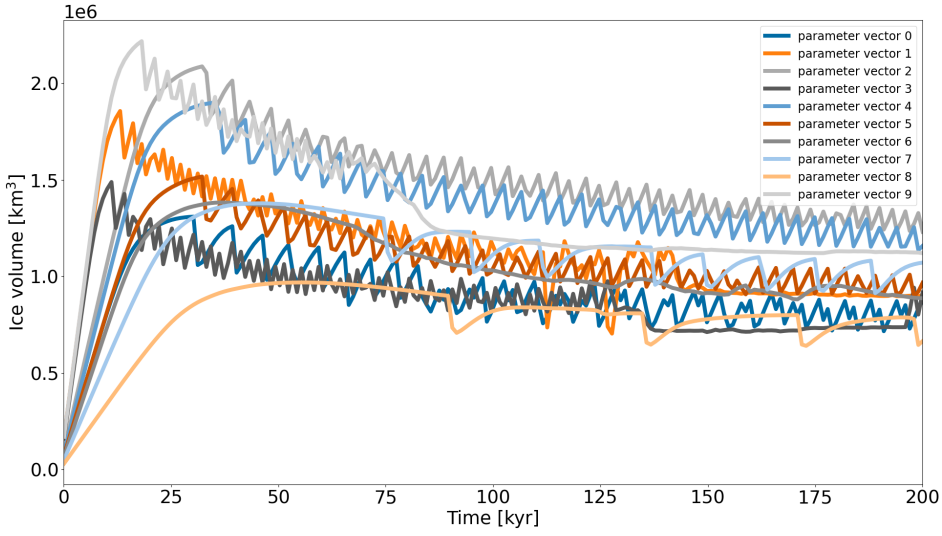


Figure S9. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for all 10 PISM parameter vectors when using the base setup.

S6 PISM - Maximum magnitude of basal ice velocity

Small till friction angles (0.5 to $1.0°$) lead to slippery beds and high maximum basal sliding velocities (up to $\sim 600\ \mathrm{km\ yr^{-1}}$) for a small number of time steps in some runs. A maximum of 7 out of 2000 time steps exceeds $50\ \mathrm{km\ yr^{-1}}$ (parameter vector 1 in Fig. S10). While observed velocities can reach several hundreds of meters per day for short periods (K.M. Cuffey and W.S.B. Paterson. (2010), e.g., $300\ \mathrm{m\ d^{-1}} = 109.5\ \mathrm{km\ yr^{-1}}$), high modeled velocities might lead to instabilities in the numerical matrix solver.

To test if the model response to different setups is affected by such numerical instabilities, we create a subensemble of the 10 base parameter vectors, including only runs with a maximum basal sliding velocity $v_{max} < 50\ \mathrm{km\ yr^{-1}}$ (parameter vectors 6, 7, and 8 in Fig. S10). We repeat the analysis in Sec. 2.3 for the subensemble and compare the event characteristics to the full ensemble.

Setup	#Events	mean period	mean duration	mean ice volume change	nE1
full ensemble base setup (nCores= 8)	28 ± 17	10 ± 12 kyr	3 ± 2 kyr	$1.2 \pm 0.3 \cdot 10^5\ \mathrm{km^3}$	0
subensemble base setup (nCores= 8)	5 ± 2	26 ± 11 kyr	7 ± 2 kyr	$1.2 \pm 0.5 \cdot 10^5\ \mathrm{km^3}$	0
nCores= 2, full ensemble	0.8 ± 25.0	-8.3 ± 45.0	-2.3 ± 19.8	1.5 ± 6.2	0
nCores= 2, subensemble	-22.2 ± 15.7	$\mathbf{38.7 \pm 70.2}$	11.5 ± 15.3	7.5 ± 7.2	0
nCores= 4, full ensemble	1.9 ± 34.1	-7.0 ± 15.0	-8.8 ± 15.0	0.8 ± 4.4	1
nCores= 4, subensemble	$\mathbf{-33.3 \pm 27.2}$	-15.8 ± 17.6	$\mathbf{-12.2 \pm 6.0}$	1.1 ± 5.2	1
nCores= 16, full ensemble	-1.2 ± 16.5	$\mathbf{14.3 \pm 46.1}$	6.2 ± 42.1	$\mathbf{-2.5 \pm 7.48}$	0
nCores= 16, subensemble	-11.1 ± 15.7	37.5 ± 68.6	-1.4 ± 18.5	-6.1 ± 4.5	0
nCores= 32, full ensemble	$\mathbf{2.8 \pm 12.9}$	-6.3 ± 15.3	$\mathbf{-9.4 \pm 16.0}$	0.3 ± 12.7	0
nCores= 32, subensemble	-4.2 ± 5.9	-1.2 ± 17.7	-9.8 ± 5.3	$\mathbf{8.0 \pm 19.1}$	0
N_{diff}	35.2	51.8	21.0	8.9	-

Table S1. Noise estimates in percent (except first two rows) for the full ensemble (all 10 parameter vectors) and the subensemble including only the 3 runs with a maximum basal sliding velocity $v_{max} < 50\ \mathrm{km\ yr^{-1}}$ (Fig. S10). 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 maximum changes across all 4 experiments are marked as bold numbers for the the full- and subensemble. The last row shows the maximum noise estimates for the difference in event characteristics between the full- and subensemble described in the text and used in Fig. S12.

The maximum noise estimates for each event characteristic are on the same order of magnitude (mean duration and ice volume change) or larger for the low-velocity subensemble (number of events and mean period, bold numbers in Tab. S1). Note that the larger noise estimates might be partly due to the lower number of runs considered in the subensemble. Since the 4 noise experiments show positive and negative changes for the same event characteristic (e.g., for the mean duration of

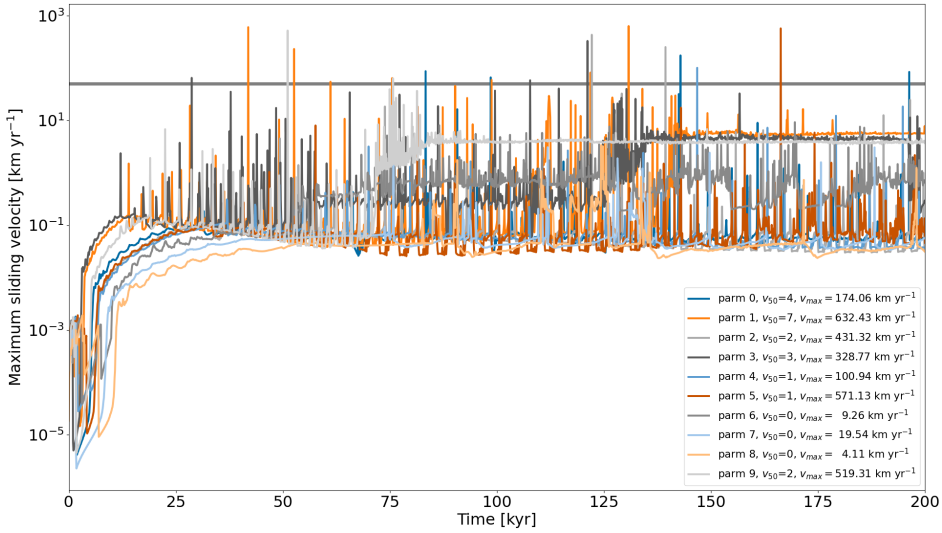


Figure S10. Maximum sliding velocity (max(max(abs($\boldsymbol{u}$)), max(abs($\boldsymbol{v}$)))) at each time step (100 yr interval) within the whole model domain for all 10 parameter vectors using PISM. The black horizontal line marks 50 km yr^{-1} and v_{50} indicates the number of time steps exceeding this velocity. v_{max} is the highest maximum sliding velocity in a run.

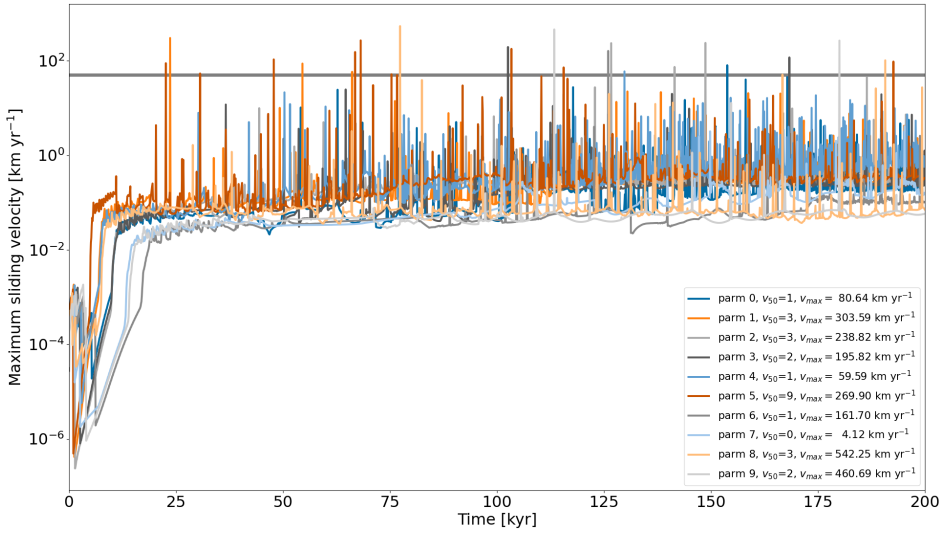


Figure S11. Maximum sliding velocity (max(max(abs($\boldsymbol{u}$)), max(abs($\boldsymbol{v}$)))) at each time step (100 yr interval) within the whole model domain for 10 parameter vectors with till friction angles between 5 and 10° and values of $C_c = 0.2$, $e_0 = 0.6$, and $\delta = 0.01$ using PISM. The black horizontal line marks 50 km yr^{-1} and v_{50} indicates the number of time steps exceeding this velocity. v_{max} is the highest maximum sliding velocity in a run.

the subensemble), we assume that the maximum noise estimates also represent changes in both directions (e.g., mean duration $\pm 12.2\ \%$). When calculating the difference in model response between the full ensemble and subensemble, the total maximum noise estimate N_{diff} is, therefore, the sum of the absolute values of the noise estimates of the full- and subensemble (e.g., for the number of events $N_{\text{diff}} = \pm(2.8\ \% + 33.3\ \%) = \pm 35.2\ \%$, last row in Tab. S1 and red area in Fig. S12).

The differences in event characteristics between the full- and subensemble are generally within the N_{diff} noise estimates indicating no significant change in model response when excluding runs with $v_{max} \geq 50\ \text{km yr}^{-1}$ (Fig. S12). Exceptions occur for increasing the horizontal grid resolutions to $50\ \text{km}$ and removing the bed thermal model. The large differences for the $50\ \text{km}$ runs are likely due to the large number of crashed runs (3 out of 7 and 1 out of 3 runs for the full- and subensemble, respectively). Therefore, the only significant differences in model response between the full- and subensemble occur when the bed thermal model is removed.

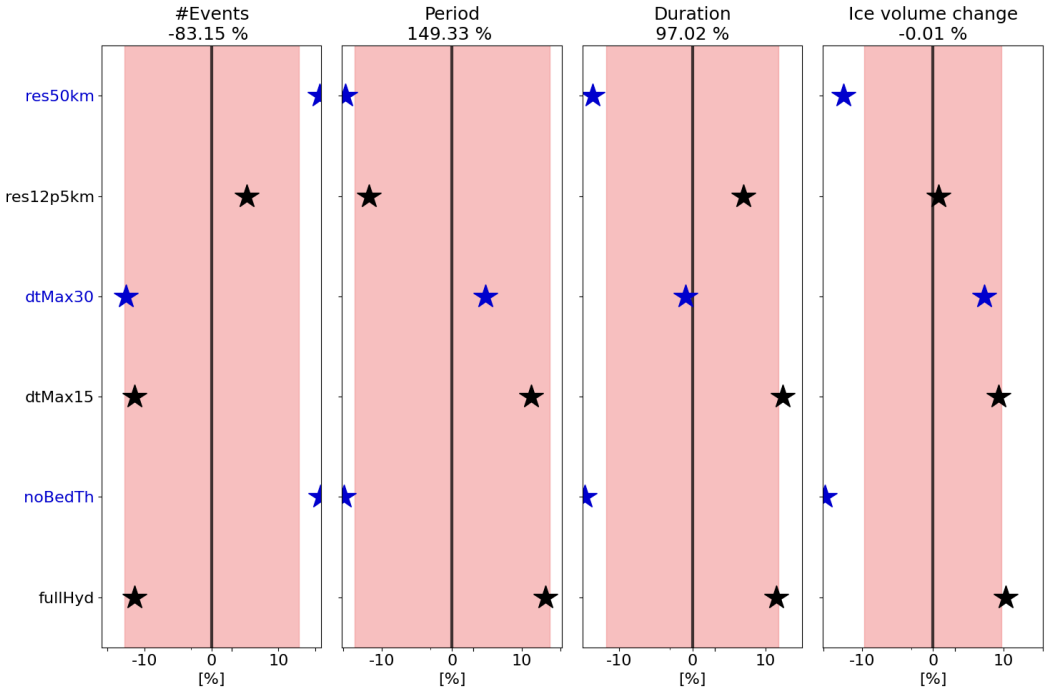


Figure S12. Differences in the percentage change of the mean event characteristics (between comparison and base setup) for the full 10 run PISM ensemble and a subensemble including only the 3 runs with a maximum basal sliding velocity $v_{max} < 50\ \text{km yr}^{-1}$. A positive difference indicates a larger change for the subensemble. The percentages in the titles of each subplot represent the differences in the event characteristics of the base runs. For example, the mean number of events in the subensemble is $\sim 83\ \%$ smaller than in the full ensemble. The shaded pink regions mark the combined noise estimates described in the text and listed in the last row of Tab. S1. The different colors resemble different model setups. The x-axis is logarithmic.

Without a bed thermal model, both ensembles show an increase in the number of events (28 vs. 163 $\%$ for the full- and subensemble, respectively). The large difference between the two ensembles is due to a low number of events across all 3 subensemble runs (maximum of 8 events per run for the subensemble (base setup) compared to 46 for the full ensemble). In fact, 6 out of 7 runs with $v_{max} \geq 50$ km yr^{-1} show a decreasing number of events when removing the bed thermal model. Therefore, the larger model response in the number of events for the subensemble resembles the differences in the base model runs (black numbers in subtitles of Fig. S12) rather than a numerical solver instability. Even larger differences in the event characteristics of the base runs occur for the mean period and duration.

The difference in mean ice volume change between the base runs of the two ensembles is only -0.01 $\%$. Both ensembles show an increase in mean ice volume change when removing the bed thermal model, but the effect is stronger when including the high-velocity runs (283 vs. 374 $\%$).

Based on the above analysis, the model response is qualitatively similar for the full- and subensemble, but differences can occur when the mean values of the base runs are far apart, e.g., for the mean period when removing the bed thermal model (75 vs. -24 $\%$ for the full- and subensemble, respectively). However, this represents a physically different response rather than a numerical instability. Since the maximum sliding velocity of 50 km yr^{-1} is only exceeded for maximum 7 time steps per run (Fig. S10), all runs are included in the analysis.

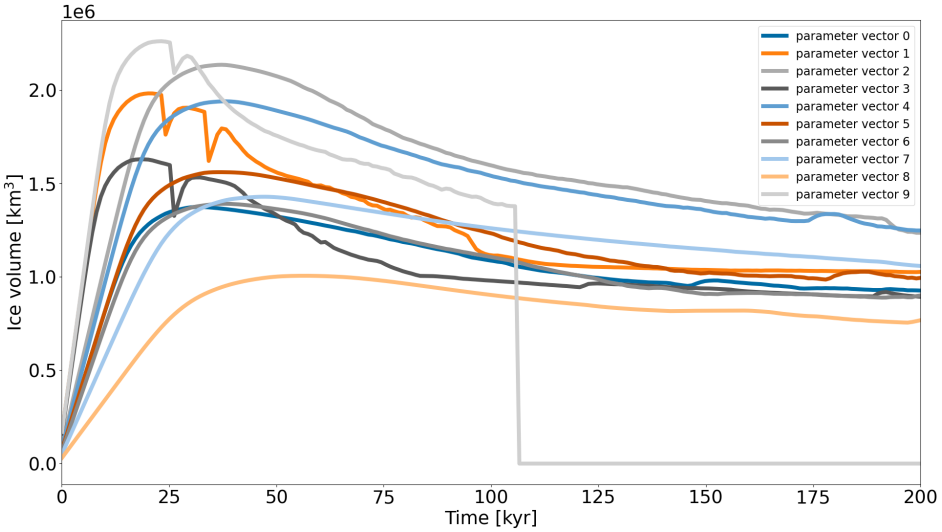


Figure S13. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for all 10 PISM parameter vectors when using the base setup but a soft bed till friction angle of 1°. Note that the run for parameter vector 9 crashed (sudden drop to 0 km^3 ice volume).

S7 GSM - temporal resolution of output time series

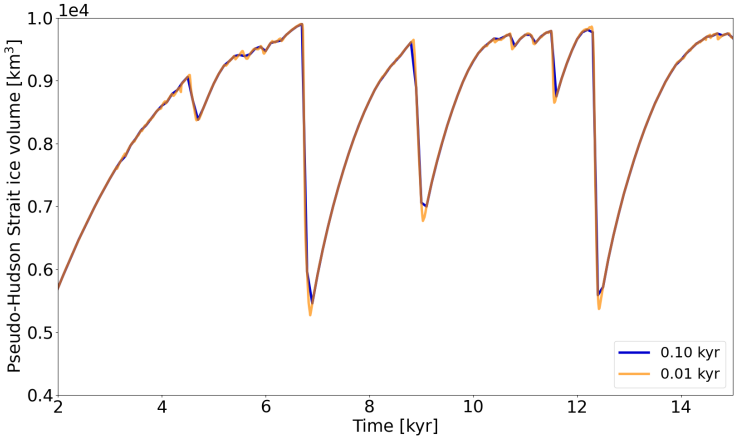


Figure S14. Pseudo-Hudson Strait ice volume of a GSM model run with different output time steps. The horizontal grid resolution is 3.125 km.

S8 PISM - temporal resolution of output time series

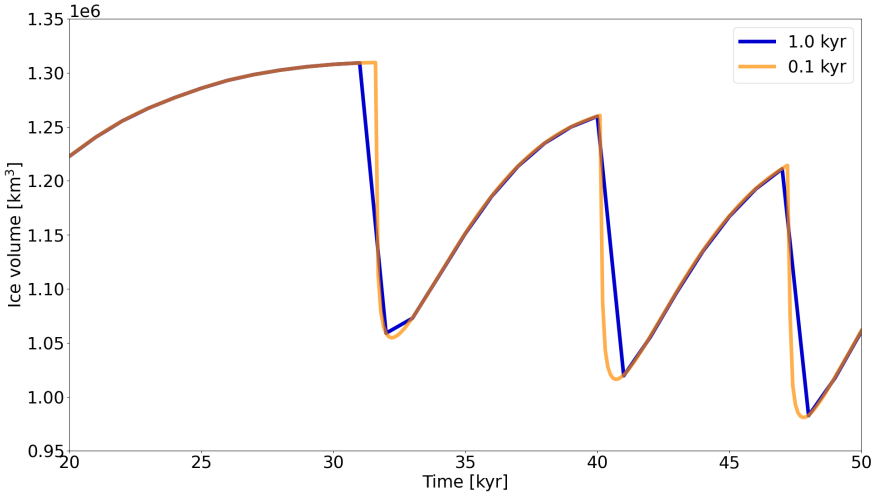


Figure S15. Ice volume in the surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) of a PISM model run with different output time steps. The horizontal grid resolution is 25 km.

S9 PISM - Ice volume - pseudo-Hudson Strait vs. surge-affected area

During a surge, ice from the pseudo-Hudson Bay and areas surrounding the pseudo-Hudson Strait is rapidly transported into the mostly ice-free pseudo-Hudson Strait. Consequently, the ice sheet extends further to the East (increasingly stronger melting), covering almost the entire pseudo-Hudson Strait area. Due to the complex interaction between ice transport and melting area, times of minimum ice volume over the area most affected by the pseudo-Hudson Strait surge (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait, e.g., video 06 of Hank (2023)) correspond to maxima in the pseudo-Hudson Strait ice volume for most surges (grey lines in Fig. S16). However, some ice volume minima do not align with a maximum of the pseudo-Hudson Strait ice volume (red lines in Fig. S16). This inconsistency hampers the detection of surges when using the pseudo-Hudson Strait ice volume and can lead to flawed statistics. To avoid this issue, we use the ice volume in the surge-affected area, for which surges appear as minima, for all PISM results. A comparison between PISM results based on the pseudo-Hudson Strait and the surge-affected area is shown in Fig. S17.

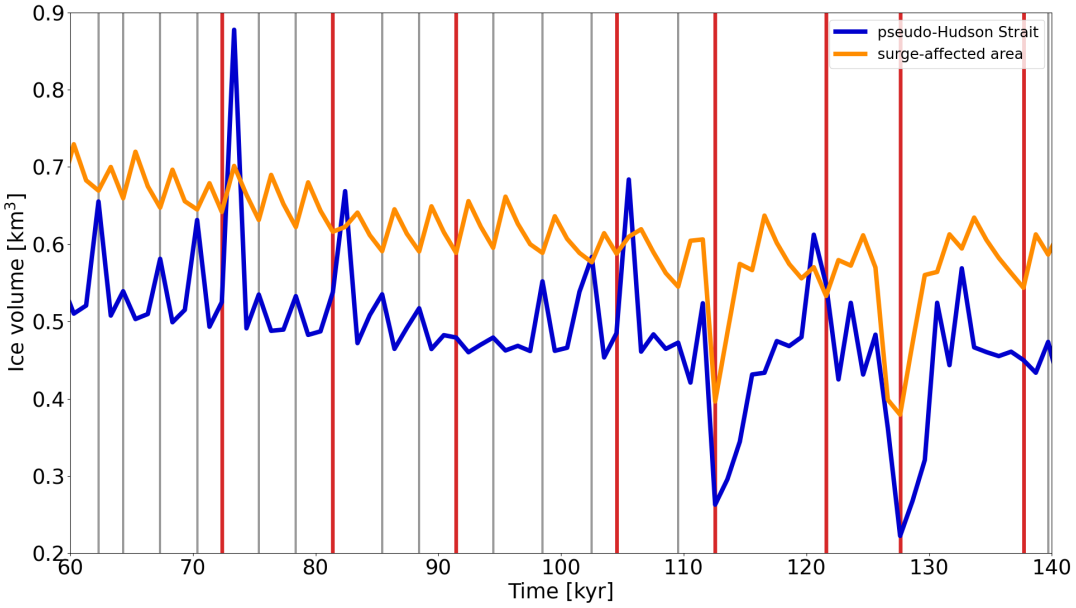


Figure S16. Normalized pseudo-Hudson Strait and surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) ice volume for parameter vector 1 using PISM. For most surges events, the minimum ice volume over the surge-affected area aligns with a maxima in the pseudo-Hudson Strait ice volume (grey lines). This is, however, not true for all surge events (thick red lines) and can lead to flawed statistics. See also video 06 of Hank (2023).

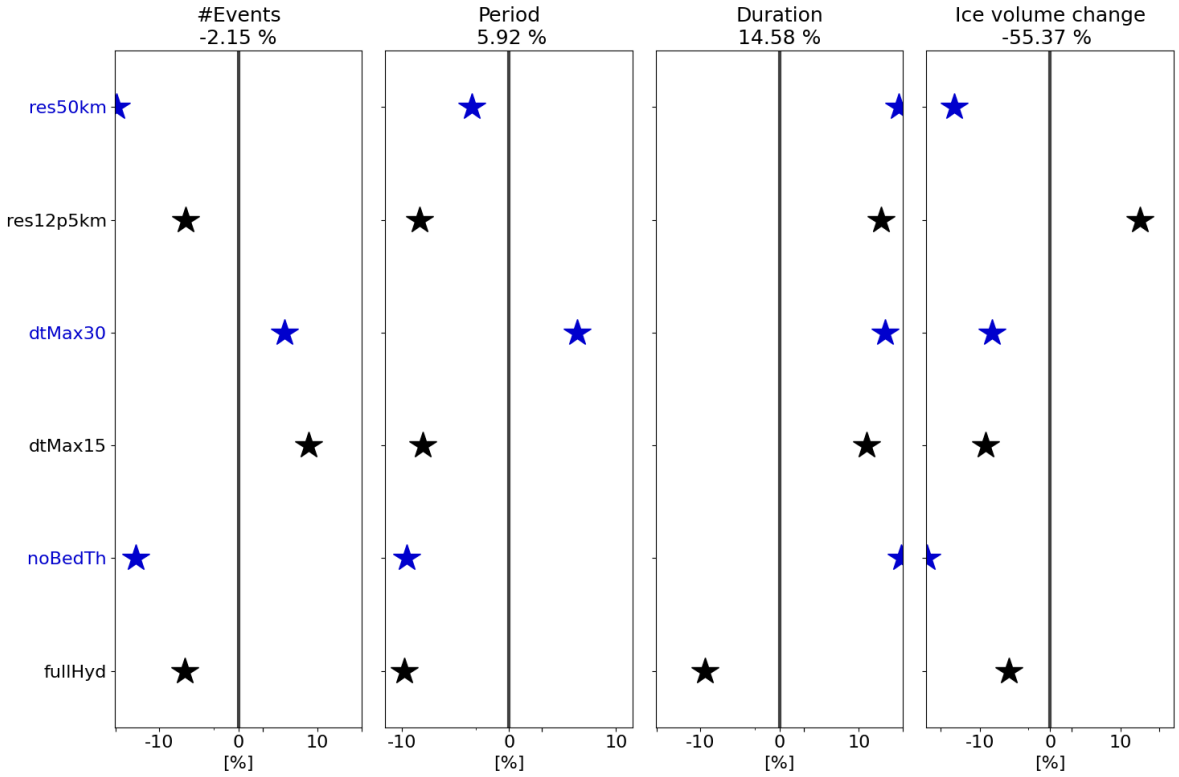


Figure S17. Differences in the percentage change of the mean event characteristics (between comparison and base setup) when using the ice volume of the surge-affected area (eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait) compared to only the pseudo-Hudson Strait ice volume. A positive difference indicates a larger change for the analysis based on the pseudo-Hudson Strait ice volume. The percentages in the titles of each subplot represent the differences in the event characteristics of the base runs. For example, the mean number of events based on the pseudo-Hudson Strait ice volume is ~ 2.0 % smaller than for the ice volume of the surge-affected area. The different colors resemble different model setups. Note that the surge threshold is 40^4 km^3 when using the surge-affected area ice volume and $0.5 \cdot 10^4$ km^3 for the pseudo-Hudson Strait ice volume (~ 5 % of mean ice volume across all runs). The x-axis is logarithmic.

S10 Event characteristics

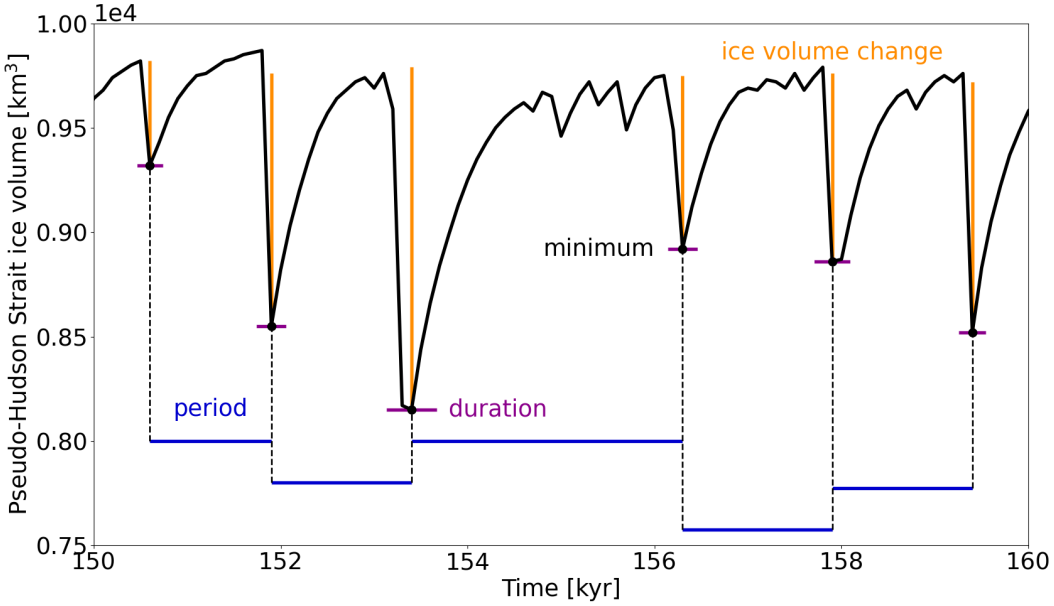


Figure S18. Pseudo-Hudson Strait ice volume of a GSM model run with visual illustration of the event characteristics used to compare different model setups. The horizontal grid resolution is 3.125 km.

S11 RMSE and mean bias

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

$$rmse = \sqrt{\frac{\sum_{t=0}^{t_{max}} (x_t - b_t)^2}{N}} \cdot \frac{100}{b_m} \tag{S9}$$

$$mean\ bias = \frac{\sum_{t=0}^{t_{max}} (x_t - b_t)}{N} \cdot \frac{100}{b_m} \tag{S10}$$

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

S12 Comparison between different model setups

The analysis to compare the different model setups follows

1. run 1 parameter vector with the base setup (Tab. 2)
2. calculate the event characteristics for this base run
3. re-run the same parameter vector for one of the comparison setups (Sec. 2.1.3 and 2.2.4)
4. calculate the event characteristics for the comparison run
5. calculate the differences in event characteristics between the base run and comparison run expressed as percentage deviations from the base run (positive for increase compared to the value of the base run)
6. repeat steps 1) to 5) for all parameter vectors (5 for the GSM, 10 for PISM)
7. take the average of all percentage deviations for each event characteristic

Percentage deviations for crashed comparison runs are not considered for the final average and runs with less than 2 events require special treatment. In these cases, the period is set to a NaN value, leading to a NaN difference between that particular run and the corresponding base run. We use Numpys *numpy.nanmean()* and *numpy.nanstd()* to ignore these NaN values when averaging over all parameter vectors. Similarly, all event characteristics except for the number of events are set to NaN values for runs with no events at all.

S13 GSM - 12.5 km noise estimate

Metric	original $12.5\ \mathrm{km}$ runs	stricter numerical convergence [%]
#Events	81 ± 42	2.3 ± 8.5
mean period	$2.3 \pm 0.8\ \mathrm{kyr}$	-2.1 ± 7.5
mean duration	$0.6 \pm 0.2\ \mathrm{kyr}$	-1.4 ± 9.4
mean pseudo-Hudson Strait ice volume change	$2.2 \pm 1.1 \cdot 10^3\ \mathrm{km}^3$	20.9 ± 53.0

Table S2. Percentage differences of event characteristics between GSM runs with regular and stricter numerical convergence at $12.5\ \mathrm{km}$. No runs crashed and all runs had more than 1 surge event. The first $20\ \mathrm{kyr}$ of each run are treated as a spin-up interval and are not considered in the above.

Metric	base setup	± 0.1°C noise	± 0.5°C noise	implicit coupling
#Events	180 ± 100	-4.0 ± 4.3	-4.1 ± 7.0	1.1 ± 4.9
mean period	1.1 ± 0.5 kyr	4.8 ± 5.3	3.8 ± 6.8	-0.3 ± 5.3
mean duration	0.3 ± 0.1 kyr	1.3 ± 4.4	0.9 ± 4.3	-12.7 ± 9.5
mean pseudo-Hudson Strait ice volume change	$1.7 \pm 0.2 \cdot 10^3$ km^3	0.9 ± 4.1	2.1 ± 5.5	-25.1 ± 18.7
RMSE	-	8.0 ± 2.5	7.8 ± 2.1	7.3 ± 2.5
Mean Bias	-	-0.1 ± 0.2	0.1 ± 0.0	1.8 ± 1.5

Table S3. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup for two different amplitudes of surface temperature noise and implicit coupling between the thermodynamics and ice dynamics (except first column). 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).

S14 PISM - Noise estimate

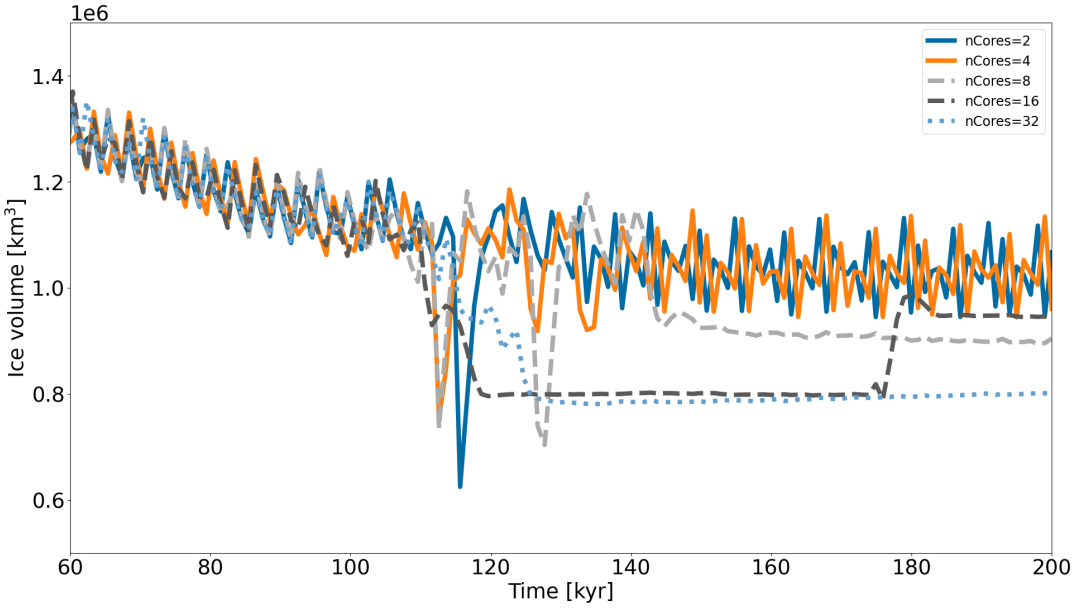


Figure S19. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 1 and different number of cores/processes using the Parallel Ice Sheet Model (PISM).

S15 GSM - Smooth sediment transition zone and non-flat topography

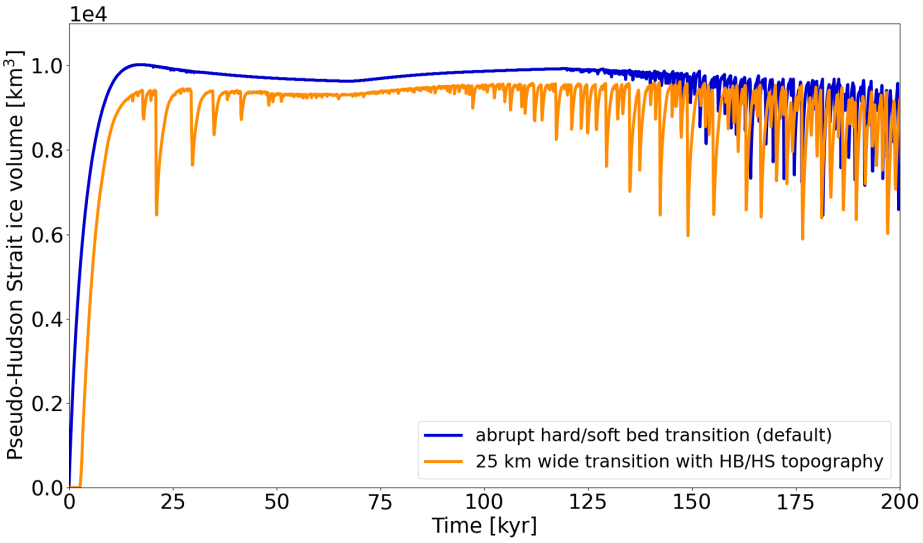


Figure S20. Pseudo-Hudson Strait ice volume for parameter vector 0 with and without a 200 m deep topography in the pseudo-Hudson Bay and Hudson Strait region using the GSM. In runs with a non-flat topography, the initial glaciation is delayed because the pseudo-Hudson Strait topography is below sea level. The horizontal grid resolution is 3.125 km.

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

S16 GSM - Bed thermal model

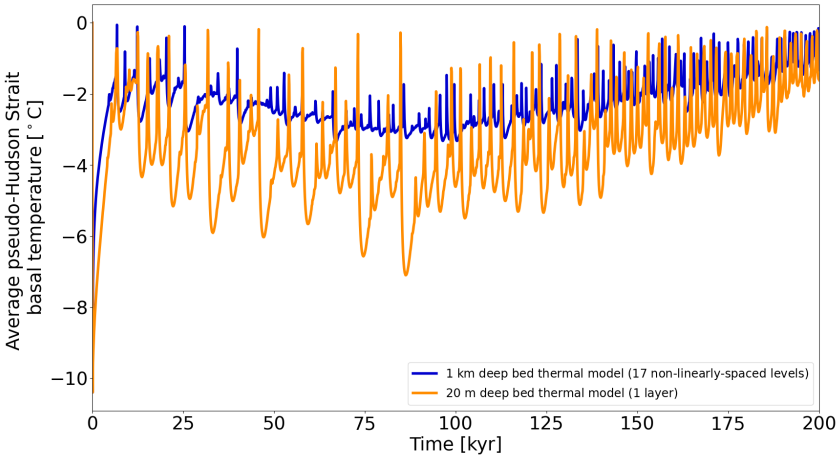


Figure S21. Average pseudo-Hudson Strait basal ice temperature with respect to the pressure melting point for parameter vector 1 with a 20 m and 1 km deep bed thermal model (17 non-linearly-spaced levels) using the GSM. The horizontal grid resolution is 3.125 km.

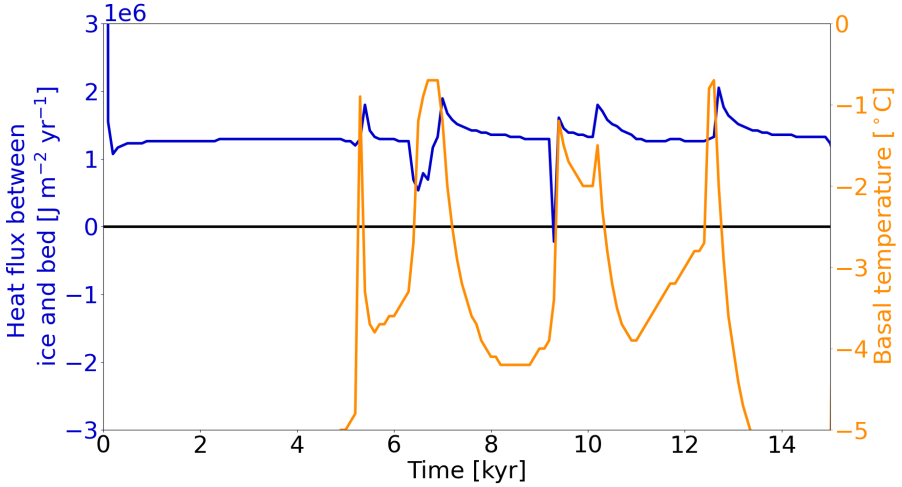


Figure S22. Heat flux at the base of the ice sheet (positive from bed into ice) and basal ice temperature for a grid cell in the center of the pseudo-Hudson Strait (grid cell center at $x = 376.5625$ km and $y = 248.4375$ km, white star in Fig. 1) and parameter vector 1 with only one bed thermal layer (20 m deep) using the GSM. The horizontal grid resolution is 3.125 km.

S17 PISM - Bed thermal model

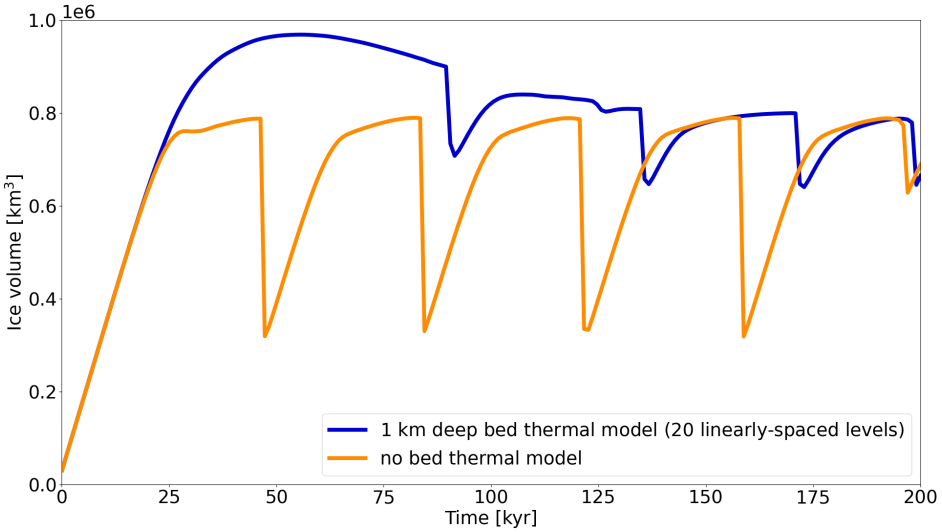


Figure S23. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 8 with and without the 1 km deep (20 linearly-spaced levels) bed thermal model using the Parallel Ice Sheet Model (PISM). The horizontal grid resolution is 25 km.

S18 GSM - Weight of adjacent minimum basal temperature

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

Table S4. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup ($W_{Tb,\min} = 0.5$) for different weights of the adjacent minimum basal temperature for the basal sliding temperature ramp (except first column). No runs crashed and all runs had more than 1 surge 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).

S19 GSM - Basal temperature ramps at different resolutions

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

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

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

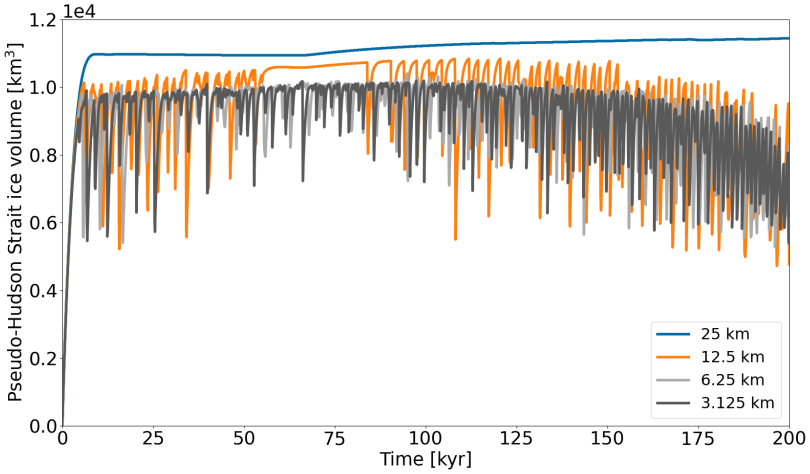


Figure S24. Pseudo-Hudson Strait ice volume for parameter vector 1 and different horizontal grid resolutions using the GSM. A constant temperature ramp with $T_{ramp} = 0.0625$°C and $T_{exp} = 28$ is used for all horizontal grid resolutions (magenta line in Fig. 2).

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

Table S5. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup ($T_{ramp} = 0.0625$, $T_{exp} = 28$) for different basal temperature ramps (except for base setup). The ramps are sorted from widest (first row) to sharpest (last row, see Fig. S25). The bold base values in the middle of the table seperate the ramps that are wider (above) and sharper (below) than 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 for the event characteristics (not the RMSE and mean bias).

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

Table S6. Single value scores for the mean and standard deviation of the basal temperature ramps. The temperature ramps are shown in Fig. S26. A total of 12, 13, and 13 ramps were tested at 25 km, 12.5 km, and 6.25 km horizontal grid resolution, respectively. Note that ramps whose sum (score-mean + score-std) differ by more than 50 % from the minimum sum at the corresponding resolution are not listed here. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. No runs crashed and all runs had more than 1 event. Note that the sum of scores can be slightly off due to rounding (± 0.01).

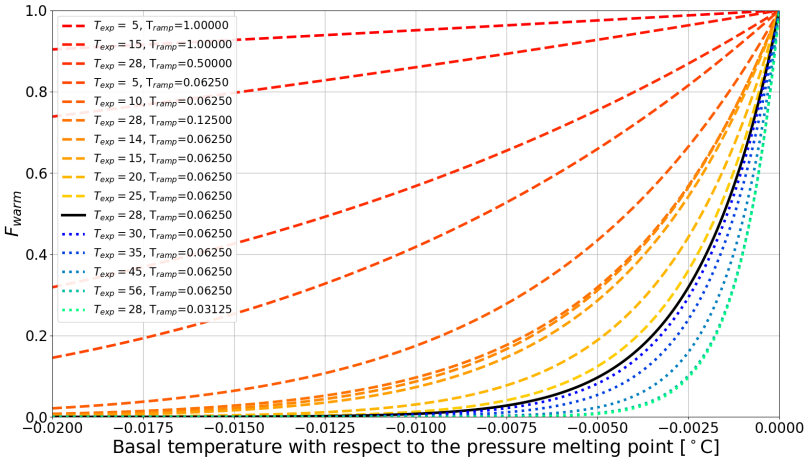


Figure S25. Temperature ramps for different values of T_{ramp} and T_{exp}. The black solid line shows the ramp used for the 3.125 km horizontal grid resolution base setup ($T_{ramp} = 0.0625$, $T_{exp} = 28$). The solid and dotted lines show ramps that are wider and sharper than the base setup, respectively. The depicted temperature ramps are the same as the ones listed in Fig. 6 and Tab. S5.

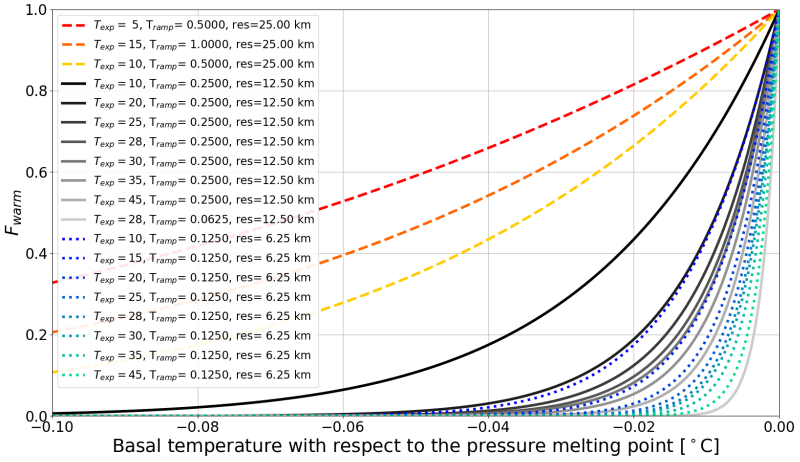


Figure S26. Shown are the temperature ramps listed in Tab. S6 at 25 km (solid lines), 12.5 km (dashed lines), and 6.25 km horizontal grid resolution (dotted lines).

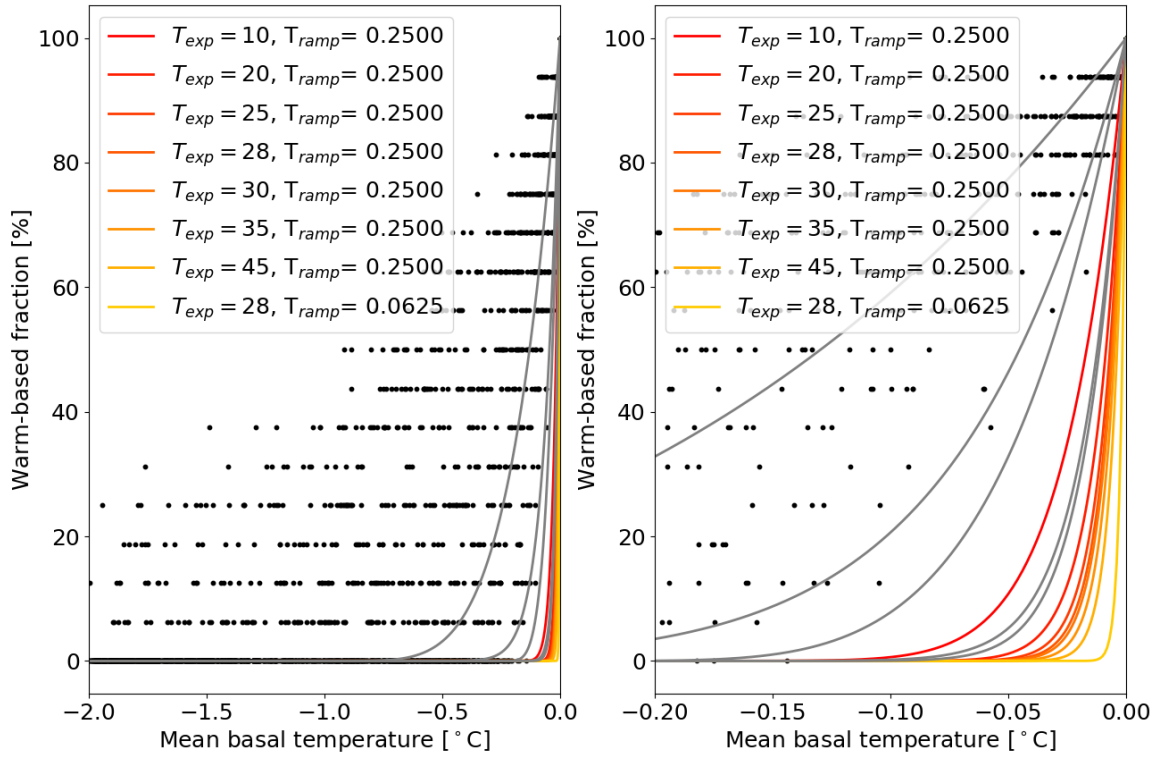


Figure S27. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 12.5 km horizontal grid resolution including all 5 parameter vectors using the GSM. Only grid cells within the pseudo-Hudson Strait and time steps within the surges of the 10 kyr after the first surge are considered. The restriction to the 10 kyr after the first surge for these experiments is set by storage limitations due to the high temporal resolution of the model output fields (10 yr). The colored ramps correspond to the 12.5 km horizontal grid resolution basal temperature ramps in Tab. S6 and the gray lines show all other ramps that were tested at this resolution.

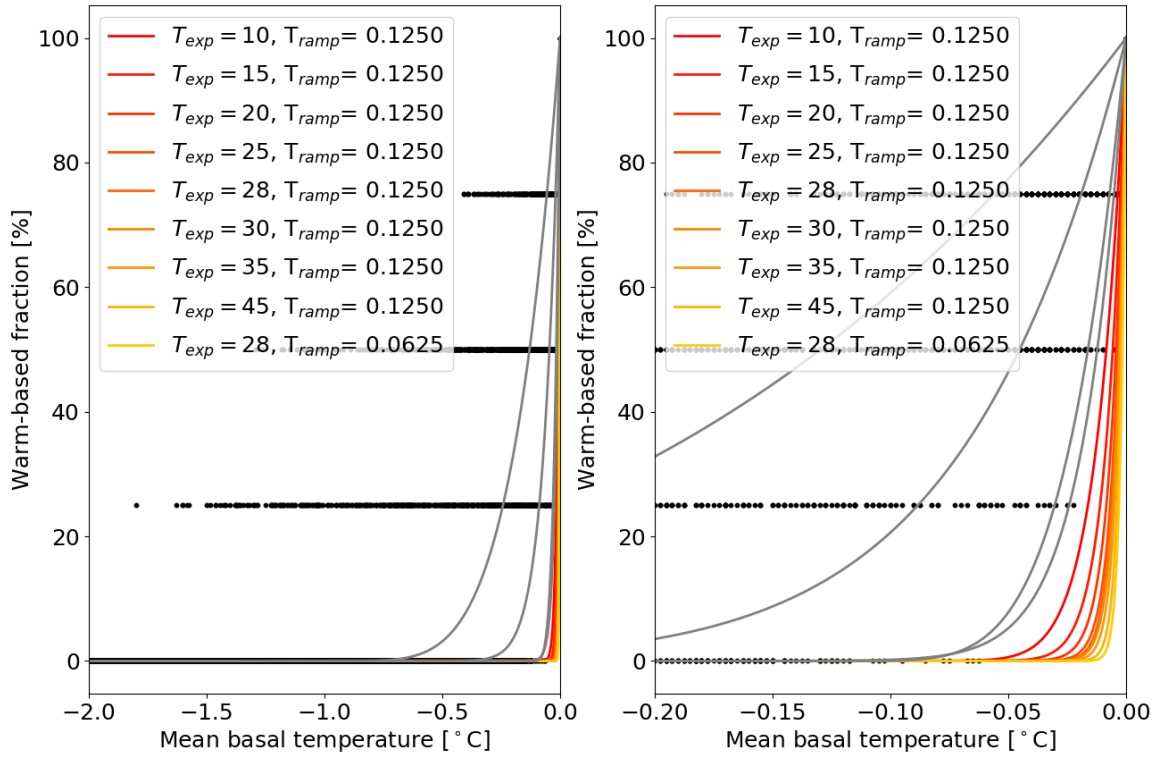


Figure S28. Warm-based fraction (basal temperature with respect to the pressure melting point at 0 °C) vs. mean basal temperature with respect to the pressure melting point when upscaling a 3.125 km run to 6.25 km horizontal grid resolution including all 5 parameter vectors using the GSM. The colored ramps correspond to the 6.25 km horizontal grid resolution basal temperature ramps in Tab. S6 and the gray lines show all other ramps that were tested at this resolution. Otherwise same as Fig. S27.

S20 GSM - Resolution and time step dependence

Analyzing individual parameter vectors in detail shows that some discrepancies prevail even when using a resolution-dependent temperature ramp. In the case of parameter vector 1, for example, surges do still not occur for the coldest temperatures (Fig. S29). Note the asymmetry in termination and onset of surge cyclicity ($\Delta t_1 < \Delta t_2$). For increasing temperatures after the minimum surface temperature $T_{\min} = -15°\mathrm{C}$ at $t_{\min} = 66.7$ kyr, the first surge occurs at a surface temperature slightly higher than the initial temperature T_{init}, for which oscillations occur. The difference between Δt_1 and Δt_2 is ~ 25 kyr and closely resembles the lag of the average pseudo-Hudson Strait basal temperature with respect to the pressure melting point behind the surface temperature changes. For example, the minimum average pseudo-Hudson Strait basal temperature with respect to the pressure melting point ($T_{\mathrm{bpm}} = -3.2°\mathrm{C}$) occurs 23 kyr after the minimum surface temperature (not shown). The period without oscillations in the 25 km run corresponds to a period of somewhat smaller and less frequent oscillations in the higher resolution runs (Fig. 8).

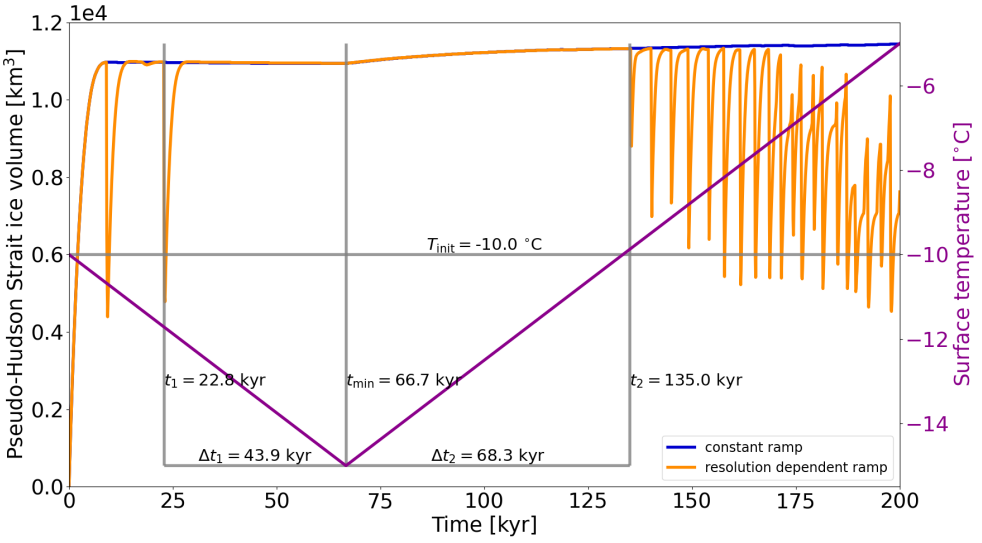


Figure S29. Pseudo-Hudson Strait ice volume for parameter vector 1 and different basal temperature ramps using the GSM (constant ramp: $T_{ramp} = 0.0625°\mathrm{C}$ and $T_{exp} = 28$; resolution-dependent ramp: $T_{ramp} = 0.5°\mathrm{C}$ and $T_{exp} = 28$, see Fig. 2). The right axis shows the surface temperature when ignoring the lapse rate dependency. $t_{\min}$, t_1, and t_2 mark the time of the minimum surface temperature, the start of the last surge before $t_{\min}$, and the start of the first surge after $t_{\min}$, respectively. Δt_1 and Δt_2 represent the time difference between $t_{\min}$ and t_1 and t_2, respectively. T_{init} indicates the surface temperature at the beginning of the run.

Setup	mean RMSE	mean Bias
25 km, constant ramp	17.1 ± 4.7	14.6 ± 4.4
12.5 km, constant ramp	10.3 ± 2.5	4.0 ± 1.1
6.25 km, constant ramp	8.7 ± 2.2	0.4 ± 0.5
25 km, resolution-dependent ramp	15.3 ± 3.0	10.3 ± 2.5
12.5 km, resolution-dependent ramp	10.3 ± 2.8	3.0 ± 2.3
6.25 km, resolution-dependent ramp	8.5 ± 2.3	0.2 ± 0.6
25 km, $T_{ramp} = 0.5$, $T_{exp} = 5$	12.8 ± 2.4	6.4 ± 1.7
12.5 km, $T_{ramp} = 0.25$, $T_{exp} = 28$, same as resolution-dependent ramp	10.3 ± 2.8	3.0 ± 2.3
6.25 km, $T_{ramp} = 0.125$, $T_{exp} = 45$	8.5 ± 2.0	0.6 ± 0.5
0.5 year maximum time step	7.7 ± 2.4	0.0 ± 0.1
0.25 year maximum time step	8.0 ± 2.5	0.0 ± 0.1

Table S7. Pseudo-Hudson Strait ice volume RMSE and mean bias compared to the 3.125 km GSM base setup in percent. The values represent the average of 5 parameter vectors. No runs crashed and the entire 200 kyr run time is used (no spin-up interval).

S21 GSM - Basal hydrology instead of basal temperature ramp as the primary smoothing mechanism

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

Table S8. Percentage differences of event characteristics, pseudo-Hudson Strait ice volume RMSE and mean bias compared to the GSM base setup with local basal hydrology instead of a basal temperature ramp as the primary smoothing mechanism (except first column). No runs crashed and all runs had more than 1 surge 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).

S22 GSM - Resolution scaling with basal hydrology

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

Table S9. Resolution scaling of pseudo-Hudson Strait ice volume RMSE and mean bias with local basal hydrology in percent. The three upper ramps are compared to the 3.125 km GSM setup with $T_{exp} = 28$, the lower three to $T_{exp} = 5$. The values represent the average of 5 parameter vectors. No runs crashed and the entire 200 kyr run time is used (no spin-up interval).

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

Table S10. Single value scores for the mean and standard deviation of the basal temperature ramps and the number of DWINN failures (maximum 4) for a resolution-dependent base temperature ramp with $T_{exp} = 28$. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. At $= 25$ km, 1 run crashed for $T_{exp} = 10$ and 1 run showed no events for $T_{exp} = [15, 20, 28]$. Note that the sum of scores can be slightly off due to rounding (± 0.01).

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

Table S11. Single value scores for the mean and standard deviation of the basal temperature ramps and the number of DWINN failures for a resolution-dependent base temperature ramp with $T_{exp} = 5$. The minimum scores for the mean, standard deviation, and sum at each resolution are marked as bold numbers. At $= 25$ km, 1 run crashed for $T_{exp} = 10$ and 1 run showed no events for $T_{exp} = [15, 20, 28]$. Note that the sum of scores can be slightly off due to rounding (± 0.01).

S23 PISM - Resolution and time step dependence

Similar to the results presented for the GSM (Sec. S20), analyzing individual parameter vectors for PISM shows significant differences in surge behavior for different horizontal grid resolutions. Parameter vector 1 at 25 km horizontal grid resolution, for example, first transitions from high-frequency small-scale oscillations to low-frequency oscillations with larger amplitudes (110 kyr) before it remains at a constant ice volume (starting at 150 kyr, Fig. S30 and video 09 of Hank (2023)). In contrast, high-frequency oscillations are present until the end of the run at 12.5 km horizontal grid resolution. The 50 km run also shows oscillations for the entire run period, but the ice volume at 200 kyr is roughly $0.3 \cdot 10^6$ km^3 (27.5 %) larger than the corresponding value at 12.5 km.

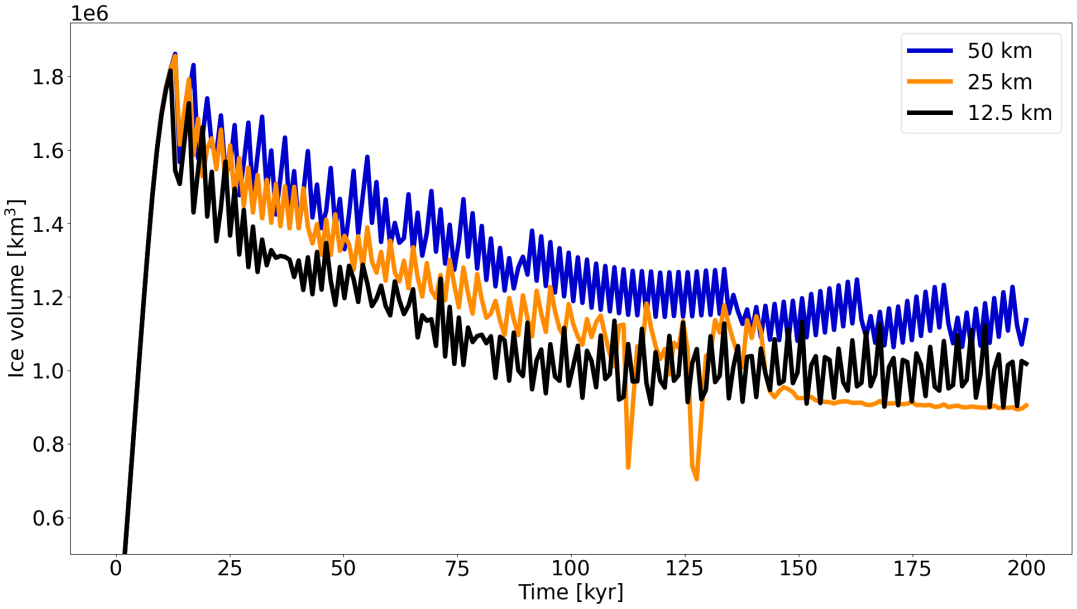


Figure S30. Ice Volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 1 and different horizontal grid resolutions using the Parallel Ice Sheet Model (PISM). See also video 09 of Hank (2023).

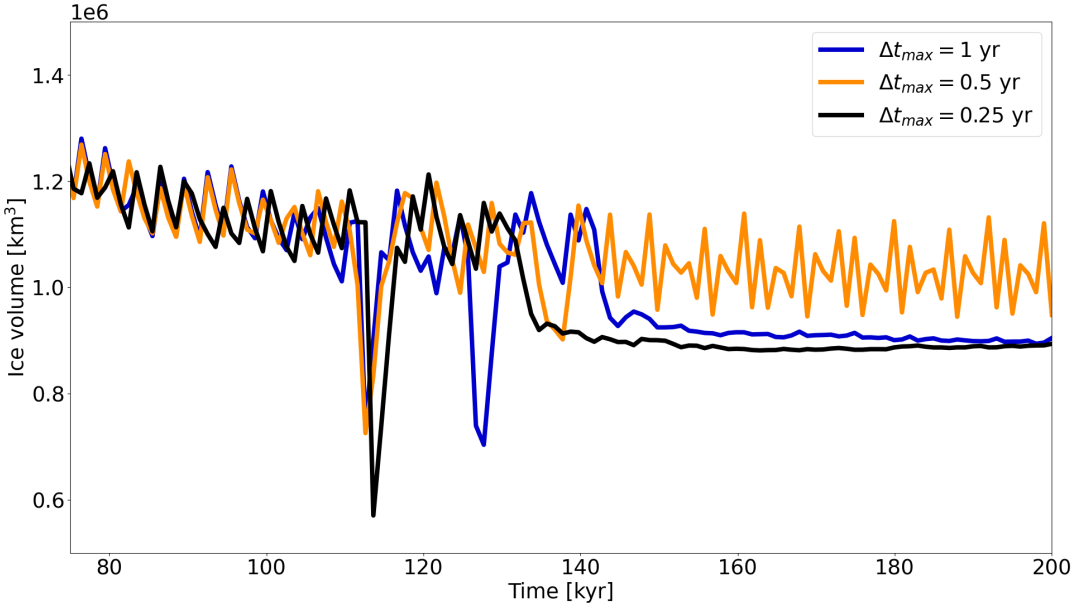


Figure S31. Ice volume in the eastern half of the pseudo-Hudson Bay and the pseudo-Hudson Strait for parameter vector 1 and different maximum time steps using the Parallel Ice Sheet Model (PISM). The horizontal grid resolution is 25 km.

Setup	nC	mean RMSE	mean Bias
$50\ \mathrm{km}$	4	12.9 ± 4.1	9.0 ± 5.5
$25\ \mathrm{km}$	0	8.5 ± 2.0	3.3 ± 1.4
$0.5\ \mathrm{year}$ maximum time step	0	4.7 ± 3.0	-0.1 ± 2.4
$0.25\ \mathrm{year}$ maximum time step	0	3.8 ± 3.1	-0.7 ± 2.3

Table S12. Ice volume RMSE and mean bias (in percent) due to different horizontal grid resolutions and maximum time steps. Note that the $12.5\ \mathrm{km}$ (highest resolution tested, one crashed run) is used as base for the grid resolution convergence study, whereas the $25\ \mathrm{km}$ setup is used for the maximum time step experiments. The values represent the average of 10 parameter vectors. Crashed runs (nC) are not considered. The entire $200\ \mathrm{kyr}$ run time is used (no spin-up interval).

References

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.

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.

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