In this article, Hank and Tarasov, perform several simulations of the Laurentide ice sheet of the last glacial period in order to elucidate the most likely candidates to explain the origin of Heinrich events. The subject and the way it is addressed in the introduction and discussion are of very high relevance. The paper is potentially very well suited for publication in Climate of the Past. I see, however, two major deficiencies of the current manuscript: A questionable and irreproducible experimental setup and the lack of adequate conclusions. I expand these aspects in the following.

General comments regarding the experimental setup and the exploration of the parameter space:

Your experimental setup is firstly very hard to understand and secondly not reproducible at all. It is hardly understandable because your entire exploration of the parameter space relies on perturbing the values of the parameters of your table S1. However the majority of these parameters are not explained at all in the manuscript.
The very pertinent questions you are addressing in this article are conveniently described at the end of the Introduction paragraph. They summarize the different mechanisms triggering Hudson Strait ice surges in the context of HEs. However, it is difficult to understand how exploring, for example, the values of the “northwestern desert-elevation cutoff” is relevant in this context. The equations on which these not-defined parameters apply are not present in the manuscript, so the reader can not even guess what influence they have on the scientific questions you are addressing. This is the case because the exploration of these parameters only helps to see whether a given realization of the model has fulfilled or not your sieves, but does not give any mechanistic picture of their influence on ice stream behavior. Nor these experiments can be reproduced with other models of similar physics. An article should be as shelf-contained as possible, so if you analyze an ensemble of simulations that have fulfilled a given sieve, the reader can understand why that is the case and the implications for your conclusions without having to open the experimental setup of another manuscript or guess that it is going to be more deeply explained in another article in preparation.

In my opinion, both these two problems would be solved by the following strategy: You could fix several of the parameters that have already allowed you to fulfill the sieves and that are mechanistically irrelevant for the ice stream behavior to a single documented value. The equations on which these parameters appear could be summarized in an appendix. Then, you could systematically explore the parameter space of the processes that are suspected to be very influential on the ice-stream dynamics (e.g. sliding laws, effective pressure, dependence on bed type and base thermal state…; see below). And finally you could describe how these different values affect the nature of the simulated Hudson Strait ice stream variability.

Specific comments regarding the experimental setup and the exploration of the parameter space:

What are the values of C_{warm} and C_{froz}? (absent in table S1)

What is the value of n_b for hard bedrock? (I could guess it is a 3 from a dimensional analysis based on the units of table S1, but it is not present in that table)

Have you explored different values of $n_{b, hard}$? I assume not (from what is shown in table S1). If not, why not simply use the same value as for $n_{b, soft}$? That will reduce one degree of freedom.

According to equations (2) and (3) your sliding coefficient, C_b is a function of the bedrock type (soft sediments vs hard), the thermal character of the base and the effective pressure. Why do you need additionally to change the exponent of the sliding law over soft and hard bedrocks?

Please justify the 3.5 exponent in equation (4)

I deduce from table S1 that your maximum allowed value for $h_{wb,Crit}$ is 1.0. Why limit it to that relatively low value? What happens if you significantly increase it?

Additionally, what does the model do if h_{wb} reaches $h_{wb,Crit}$?
Neff could not be reduced further (it would be purely 0!! From equation (4)), but what happens with the excess of heat? Can you create additional basal water? If not, could you diagnose the "free" values of h_{wb}?
How sensitive are the surging cycles to this basal water limit?

You introduce a Coulomb dragging law in equations (5a) and (5b).
How is C_c variable? In space or in time? According to what? (not deducible from the manuscript or table S1)

The necessity of equation (6) is justified in the following manner: " To account for possible Weertman-type sliding when Coulomb drag is high …"
A high Coulomb drag will potentially stabilize the flow and prevent the appearance of a limit cycle. Why convoluting a Weertman-type sliding law and a Coulomb one?
What is the effect of using the "pure" Coulomb law (without limiting τ_b ad hoc) on the surges?

GHF reconstructions:

Blackwell and Richards, 2004 is just a map without any reference to a published peer-reviewed work. Shapiro and Ritzwoller, 2004 show values around 55 mW/m² with a standard deviation around 20 mW/m². The minimum value for the whole region in Pollack 1993 is higher than 45 mW/m² and the Hudson Bay and Strait regions would show a mean value closer to 60 mW/m². Goutorbe et al., 2011 showed a Hudson Bay around 40-45 mW/m² and higher values in Hudson Strait for their first method, and a regional mean around 50 mW/m² for the same region with some hot spots of more than 70 mW/m². Lucazeau, 2019 reconstructed values in Hudson Bay are around 50 mW/m² in their first method and around 40 mW/m² in the other two methods, while they show significantly higher values (around 60 mW/m²) for the three methods in the Hudson Strait Area. Finally, as far as I could see, Cuesta-Valero et al., 2021 do not show any reconstruction of geothermal heat flow nor seems the intention of their paper.

Line 345 reads: "The exact transition point [to the binge-purge mode] depends on the parameter vector in question but generally requires a Hudson Strait/Hudson Bay GHF ave ≤ 37 mW/m². And line 357 states: "Therefore both types of Hudson Strait ice stream surge cycling are consistent within available GHF constraints."

In light of the reconstructed values described in your referenced studies, this last sentence seems highly inaccurate or simply wrong. It should rather say something like: "The binge-purge mode is, under our experimental setup, only accessible if GHF<37 mW/m² which represents the lower bound of available constraints".

Conclusions of the manuscript in light of the experimental setup

In the following I will enumerate a number of questions and concerns that, in my opinion, make some of the current conclusions highly questionable.

What is the influence of h_wb,Crit on the transition to binge-purge?
Would this transition be possible if basal water is conserved?

Your basal friction laws are designed in a way that effective pressure can be reduced several orders of magnitude (even to purely 0!) when there's enough basal water.
Also, Cb will accordingly reach extremely high values (equation (3)) when effective pressure drops. Do you think this is glaciologically realistic (as opposed to Leguy's parameterisation, for example, or even the PISM treatment?

The regularized Coulomb law is thought to perform well, even capturing the heterogeneous character of the bed without needing to guess different friction coefficients in soft and hard beds (Joughin et al., 2019). In section 3.3 it is stated that the majority of the Coulomb runs crashed. And that it significantly increased the needed computational time for those who survived.
Why is this? Could it be because of equation (6)? Or because of the additional complexity of the parameterisations over the friction coefficient and effective pressure introduced in equations (2), (3) and 4? Again, what would be the effect of letting the Coulomb law do its job, without limiting tau_b?

Runs that are subjected to severe numeric problems can often illustrate that the physical problem is not well-posed. Can you discard this is what is happening here?
How many of your reference ensemble runs crashed?

Looking at figure 5, the reader might notice that the upper limit of the basal velocities bar is 40,000 m/yr. It is not known whether the simulated velocities ever approach that value, because the manuscript does not contain the time series of velocities over a whole ice stream cycle, but I imagine that they are not far from the upper limit of the depicted scale. From figure 5 it is clear that they reach up to 20,000 m/yr for several hundred of kilometers upstream of the grounding line. Current observed surface velocities in Antarctica do not go beyond ~1,500 m/yr downstream of the grounding line. Do you think your simulated velocities are realistic?

Illustrating the dependence of your results to different GHF values seems completely adequate. However, as described above, spontaneous cycling ice stream behavior seems to appear only (given aside other questionable choices of your basal sliding laws) for the lower limit of available constraints.

All in all, unless the authors show otherwise, their experimental setup is constructed in a way that the relevant-for-the-problem parameter space (and thus the phase space of the associated physical problem) is very narrow and situated in a very specific region.

In the absence of a rigorous illustration of the influence of the particular assumptions of the experimental setup (basal sliding laws, basal water, effective pressure, dependence of the friction coefficient on the nature and temperature of the bed, numerical issues…) on the mechanisms favoring spontaneous ice stream cycling, the conclusions reached here are based on the analysis of a very particular experimental setup. (Which in my opinion is, by construction, prompt to oscillate in a questionable physical manner).

And this leads me to my next point:
What happens if you run the Antarctic ice sheet under such a configuration?

To answer this question, you would need to make some assumptions concerning the bed type. One reasonable choice would be to assume the presence of soft sediments in every Antarctic marine sector. Do you expect that such a simulation would give reasonable results?

In my modeling experience, this is unlikely.

GRISLI and Yelmo can also show spontaneous ice stream oscillations under extreme conditions (e.g, limiting basal water and having very distinct spatial basal frictions). But when you apply such a physics to Antarctica, it becomes very difficult to approach observed velocity values, and furthermore some ice streams become very noisy and others dramatically oscillate in a manner that so far is not observed in Antarctica.

Therefore, I believe the current experimental setup extremely conditions the current conclusions of the paper.

The introduction section nicely ends with the presentation of very relevant questions regarding the different mechanisms exciting the Hudson Strait ice stream variability during HEs. The discussion section also nicely re-addresses those questions (see some minor comments below) in the context of the new results shown in the manuscript. The conclusions, however, do not fairly summarize the findings and are biased towards the spontaneous ice-stream cycling mechanism.

After everything that is shown in the paper, and given all the limitations of the experimental setup pointed out in this review, concluding that “Based on our results, Hudson Strait ice stream surge cycling is the most likely Heinrich Event mechanism…” seems unjustified.

Under a questionable and very particular experimental setup, your simulations show that you have to go to the very low bound of Geothermal heat forcing for the ice stream surge cycling to emerge. But, even assuming that a binge-purge like oscillation is a good candidate to explain HEs, there is a remaining puzzling question that you completely ignored. Why are HEs happening at the middle of the cold NH phases, or stadials?
Some synchronization mechanisms have been explored in the literature to potentially answer this question (.e.g. the work of Calov and Ganopolski and Shanwell et al., 2024) but you do not inform whether these synchronization mechanisms are captured in your simulations. So, under your experimental setup, the question of HEs occurrence during the surface NH cold phases is still of concern for your conclusions (see also Barker et al., 2015; Nature).
Therefore, the authors seem to be evaluating certain hypotheses with a level of criticism and rigor not present in the case of their announced more likely mechanism.

You discard the ice-shelf breakup related hypothesis (Q2) because your simulated ice shelves are not big enough to significantly buttress the Hudson strait ice stream and because the paleoreconstructions of the Labrador Sea conditions during MIS3 do not seem compatible wit the existence of a big ice shelf in the Area.
This is a fair criticism, but I would like to point out here that in GRISLI and Yelmo (both codes are available; the first upon request to the former developer, Catherine Ritz or myself, and the second here: https://github.com/palma-ice/yelmo) the emergence of a big Labrador Sea ice shelf is a pretty natural characteristic provided oceanic temperatures are low enough. Even with a relatively warm ocean, Yelmo simulates the existence of very developed fringing ice shelves around the mouth of Hudson Strait (Moreno_parada et al., 2023; The Cryosphere).

The questions Q2, Q3 and Q4 of your discussion are compatible with each other. When answering Q3 you state that your maximum simulated ice shelf (leaving aside the one with inhibited calving) is close to the minimum required by Hulbe 2004 to explain the IRD signal. So, can you discard that a combination of Q2, Q3 and Q4 produces both the icebergs from breaking the ice shelves and the increase in the flux at the grounding line necessary for explaining HEs?

I believe that referencing our paper (Alvarez-Solas et al., 2013) in lines 95 and 548 in a technical and discussing context respectively, without having it cited in the introduction, when the different hypotheses are explained, is academically incorrect. Note, we also explored the conceptual idea of triggering HEs through changes in the oceanic temperature (Alvarez-Solas et al., 2010; NatGeo), and the effects of an ice-shelf breakup during H1 (Alvarez-Solas et al, 2011, ClimPast), coetaneous with Marcott et al., 2011

Jorge Álvarez-Solas