Alexander Robel (Referee)

This manuscript explores the stability of Pine Island Glacier under forcing from ocean melt, using a high order model of marine ice sheet dynamics. Additionally, this is the first study to bring the concept of Early Warning Signals (EWS) to the stability transition known as the "marine ice sheet instability" (MISI). Though EWS have been explored in a limited degree in some other areas of glaciology, this is an interesting and useful application to the MISI problem, which has recently been a focus of intense study in the glaciology community.

The central concept and design of this study are sufficiently novel and important that it eventually should be published in The Cryosphere, though I think it requires some revision first. In particular, since this is the first application of EWS to the MISI problem, it needs to be clear why exactly EWS are a useful tool for studying marine ice sheet stability. Additionally, the methodological details of EWS, while established in the dynamical systems literature, are not well known in glaciology. If the authors wish other glaciologists to follow their lead in using this approach (which I think should be the case), then they need to do a better job explaining the methods they use and the assumptions inherent to these methods. I lay out these critiques in more detail below through major and minor suggestions:

Many thanks, Alex, for the positive evaluation of our study and the helpful comments. We replied to all comments and hope that our replies and changes made to the manuscript address all issues raised. Our replies below are indicated in red and references can be found at the end of the document.

Major:

1. Here is the question that you should answer in this manuscript: Why/When are EWS a useful tool for understanding MISI at a particular glacier? At the moment, my interpretation (perhaps erroneous) of the purpose of EWS laid out in this manuscript is to show that there is a bifurcation (in fact several) in the model. However, you don't need EWS to show that this is the case when you have a model available, since you have the quasi-steady simulations which show the bifurcation structure of PIG. Rather, the point of EWS is to detect a bifurcation before it occurs. You do so in the context of your model, however, solely within the context of a predictive model, EWS are not strictly necessary, because the model can be run forward to determine whether a bifurcation will occur with continued forcing along some trajectory (this is the point of physical models!). However, what you could argue here is that your study is a proof-of-concept to indicate the circumstances under which we would expect to detect EWS in observations, which would be immensely useful for the community. This is what I find currently lacking in the study - any discussion of the implications of your study for observations. For example, you touch on this issue later in the paper about the fact that in the real world, ice sheets do not stay on the stable manifold because forcing is much faster than the response time, but then don't really explore whether this makes EWS useless in practice (I think no, but that isn't my take away from the current way its written). Another issue (which you don't remark on) is the fact that a 300 year averaging window for the EWS indicators is not super useful when the entirety of the observational record is 40 years long (perhaps a bit longer if we include some lower quality historical obs). This is all to say that showing EWS exist in a model is not very insightful in of itself if it doesn't provide some indication for what we should be looking for in observations (since again, we already know that there are bifurcations associated with MISI in models).

Both reviewers state that our paper needs to do a better job of explaining the background and methodology of EWS and why our results are useful in a glaciological context. We agree that both of these things could be greatly improved on and will endeavour to do so in a revision of the paper. The aim of this paper is a 'proof of concept', demonstrating that EWS can be found for the marine ice sheet instability in realistic geometries. The theoretically-proven hysteresis of MISI (Schoof 2007) is similar to a fold (aka saddle node) bifurcation. For this class of dynamic systems the existence of EWS can be theoretically proven. However, it is by no means clear that EWS can be still detected in more complex, realistic systems, as for example discussed in Darkos et al., (2015). We show for the first time for a realistic geometry (including for example ice shelf buttressing that is not considered in the theoretical hysteresis for MISI) that indeed EWS can be detected. We will add more discussion regarding the potential for observations of ice sheets to be used in the context of EWS since we agree that this is arguably the most important potential outcome of our findings. Our study is a first step in understanding how this methodology could be used. Certainly, as Alex points out, a 300 year averaging window is not necessarily useful in terms of available datasets. An important caveat that we will make much clearer in a revised paper is that the 300 year averaging window that we found was optimal does not necessarily represent a lower bound and this is something that certainly requires and warrants further investigation.

2. You haven't necessarily explained why EWS show up in certain types of systems. To me, this is key to then explaining why you are calculating these things (ACF, variance, etc). In a canonical saddle-node bifurcation, we expect the stable eigenvalue of the linearized system state to smoothly decrease towards zero as you approach the bifurcation, which causes weaker damping of noisy forcing back towards the stable manifold. So, do EWS only occur where there is a saddle-node bifurcation? How do we know that MISI at PIG has such a smooth approach to the bifurcation? i.e. if the eigenvalue associated with the stable mode is controlled at first order by bed topography (which it is in the canonical formulation of MISI, see Schoof 2012 and others), then does the topography in the vicinity of the bifurcation need to vary smoothly towards a bedrock peak to produce EWS? Maybe these are questions for another study, but there needs to be some indication that you have grappled with the question of why you expect EWS to occur for PIG.

Although EWS have largely been used to detect saddle node bifurcations these methods can be applied to other types of bifurcation (Scheffer, 2009), for example they have successfully been used in the context of Hopf bifurcations (Chisholm & Filotas, 2009). We will add this point to the revised paper. The second comment, regarding how bed roughness might affect the performance of EWS, is something we looked at during our modelling work but ended up not including in the paper submission. In fact, this was originally a major motivation of the PIG experiments because there was no guarantee that observing critical slowing in the idealised MISMIP experiment would work for a realistic topography. Before we did the PIG experiment we superimposed varying types and amounts of Perlin noise onto the smooth MISMIP bed, such that the overall retrograde bed slope became more obscured by smaller scale bed features. Repeating the same experiments on these new beds and extracting the relaxation time to predict the tipping point became successively less accurate as more noise was added. However even for a very bumpy bed there was still a trend towards longer relaxation times as the tipping point was approached and the accuracy of the predicted tipping point was still good enough that we had confidence this approach would work for a real glacier. The resolution of both our PIG model and the bed topography are such that we resolve as much of this 'bumpiness' as possible in these simulations. Discussion around both of these points will be added into a revised version of the manuscript.

3. Related to the issue above, more of the detail about how the EWS indicators are calculated would be helpful to bring into the main text, since this is a topic most TCD Interactive comment Printer-friendly version Discussion paper TC readers are not familiar with. How do you ensure statistical significance? Could we see EWS away from a bifurcation by chance? Why not?

The issue of so-called 'false alarms', whereby an EWS is detected and incorrectly anticipates a bifurcation, is something that has been written about extensively in the literature and indeed it is entirely possible using this methodology. In our paper we do various statistics on the sensitivity and the significance of the indicators that we have calculated. These are currently in an appendix and would probably warrant being moved into the main text which we will do, along with reference to the possibility of false alarms.

4. You say that you force your model with variability from a "surrogate model" based on ocean variability in the Amundsen Sea region, but don't provide further details. First, more detail is needed on the surrogate model. Second, presumably this surrogate model produces ocean variability with significant power in the decadal range, as many studies (e.g. from Jenkins and others) have found that such variability is important in this region. However, the typical formulation of EWS (e.g. Lenton et al. 2008, and previous studies) assumes a martingale process for the noise forcing (i.e. white noise) which is not the case here. Can you explain why this doesn't affect your interpretations of EWS indicators? Third, it is unclear whether the steady-state and quasi-steady-state simulations used to make the bifurcation diagrams in Figure 3 include noise forcing. If not, then this is concerning, because it is well known that marine ice sheets have a different steady-state with and without noise in the forcing (e.g. Robel et al. 2018, Hoffman et al. 2019, Mikkelsen et al. 2018 (but for non-marine ice sheets)). This could be quite important in your simulations, since the location of the bifurcation is important to know for calculating EWS.

The surrogate model was an AR based surrogate and does indeed contain a lot of decadal variability and this is important to capture as Alex and other authors have shown. We will provide more details on how this surrogate was made in a revised version of the paper. The second important benefit of adding this variability to our melt rate forcing is that, by perturbing the model, it helps to extract information on the system response time to identify critical slowing. It is true that the paper by Lenton 2008 shows results from earlier work by Held and Kleinen 2004 and in this case they used white noise with the same aim. The variability we add to our forcing is not white noise but there is no reason why this should affect our ability to extract information on the system response time. Indeed, other authors have found critical slowing in a wide variety of paleo data (e.g. Dakos et al. 2008) and the natural variability that drove these systems is presumably also not white noise. The key to be able to detect EWS is to add some perturbation with a measurable impact on the system state and our surrogate time series succeeds in that with the added benefit of replicating natural variability as closely as possible. Indeed, using natural variability to test for EWS arguable has more relevance to this system than using white noise. With regards to our steady-state simulations (the squares and circles in figure 3b) no natural variability was added to these simulations. With the addition of strong variability in the forcing it is possible that some of these steady states might be nudged far enough that they end up in a different basin of attraction. However, we disagree with the reviewer that this is concerning. The location of the bifurcation for the purposes of searching for EWS was not defined based on these points but on the 'quasi steady state' points and so have no bearing on our results and merely serve to demonstrate that several hystereses exist. Having said that, it is more consistent *to add the same variability in these runs. We have re-run these steady state simulations with the same natural variability that was added to the main PIG simulations and this has no significant impact on the location of these points but we will use these results, rather than the ones shown in this version of the paper, for a revised submission.*

All minor comments will be addressed in a revised version of the manuscript and where appropriate we have added a response to specific questions. If not specified otherwise, we will correct/expand on all issues raised below in a revised version of the manuscript.

Minor:

Line 23: please define what you mean by "tipping element" for the uninitiated

Line 29: grounding line flux

Line 47: the 2012 Schoof JFM paper makes more sense to cite here

Line 48: what do you mean by catastrophic?

We use the word catastrophic in the sense that it is sometimes in literature on tipping points, that is a large change in the system state. We have made an effort to keep all terminology clear and consistent and since this is the only instance of that word in the paper we will rephrase it.

Line 89: So is accumulation held constant in time or does it have a seasonal cycle? Be more specific here, because its important to know if there is variability in more than just ocean melt.

Accumulation is held constant in time, specifically to avoid adding variability in more than one forcing which could potentially muddy our results.

Line 119: Again, to be clear, it isn't necessarily the case that EWS exist for all tipping points (if we define any bifurcation as a tipping point).

We addressed this comment above and will add a clarification on this point.

Line 126: it would help here to explain exactly how you force the simulations to produce the grey dashed and black lines in Figure 3 (also in caption). How fast is the forcing? How large are the step increments to determine the steady-states? How do you determine when there is a tipping point (just a large enough jump in the grounding line?)? What if you have a tipping point that causes the grounding line to retreat only a small amount?

We will expand on this in a revised version of the manuscript.

Line 119-129: This whole paragraph is confusing

Line 136: What do you mean by short tem "weather noise"? Isn't this the thing that is detected by EWS?

This is a term that has been used in EWS literature to describe rapid but low amplitude fluctuations in the system state that do not represent

Line 184: What do you mean by "equivalent to a random walk"?

A random walk describes an evolving system that consists of a succession of random steps but this is probably unnecessary jargon and will be removed.

Line 197-198: is this hysteresis related to the domain extent? If you aren't simulating flow from outside the domain, then when the whole domain collapses you won't be able to regrow the glacier for any parameter (because theres no catchment).

The catchment is not completely gone and the grounding line remains within the basin but presumably you are correct that by extending the domain it might make it easier to regrow the glacier. That does not detract from our point in this case but we will include this as a possible explanation.

Line 203-205: what about much shorter windows? (related to point 1 above related to observations). How short of a window would actually be calculable from observations (a decade? does this start to run into the AC time scale of forcing?)

This is addressed above and will be discussed in more detail in the revised paper.

Line 214-215: this is a confusing sentence which leads me to think that you are saying the EWS are not actually "early". To what extent does this depend on the speed of the trend in forcing? Can you test it for different trend rates?

We feel that the next sentence addresses this question to some extent but we can expand on that further. The EWS are identifying tipping in the transient simulation for which they are calculated, and this would be equivalent to the 'steady-state' simulations if the forcing was increased sufficiently slowly, but since the transient simulation evolves quicker than the glacier can adjust then these are not the same. To what extent the forcing rate affects various things is a potentially large and complex research question that we believe is beyond the scope of our study.

Figure 3a: If the black line doesn't fall on top of the gray dashed line, then the black line simulation isn't really quasi-steady. Why not call it something else? Also, it is unclear how the grey dashed line is determined, and which parts are stable and unstable?

We will think of a word other than 'quasi steady' as this is admittedly contradictory. The dashed grey line is just a line through the individual model results represented by symbols, all symbols are steady states but it is possible that these are not true stable steady states.

Figure 3b: please explain the different between "tipping point" and "instability onset". I can guess that the latter has to do with the region in which an unstable manifold exists (i.e. there is hysteresis), but I'm not sure readers will necessarily pick up on this without you explaining it explicitly.

Figure 4: Looking at this (and Fig. B2), it seems clear to me that the length of time ahead of the tipping point needed to detect EWS is directly correlated to the speed of the forcing. Yet, that isn't really made clear here or in the text.

The rate at which the forcing is changed is constant throughout the simulation (ignoring the added variability)

Figure B2: can you extend this to shorter windows?

Appendix A: Related to some of the issues I raised above, it may be valuable to bring the flowline simulations into the main text to demonstrate, in a very simple system where the exact location of the tipping points are known, how EWS work

References

Dakos, V. et al. Resilience indicators: prospects and limitations for early warnings of regime shifts. Phil. Trans. R. Soc. B 370, 20130263 (2015)

Chisholm, R. A. & Filotas, E. Critical slowing down as an indicator of transitions in two-species models. J. Theor. Biol. 257, 142–149 (2009).

Scheffer, M. et al. Early-warning signals for critical transitions. Nature reviews. 461. 53-59 (2009)

Dakos, V. et al. Slowing down as an early warning signal for abrupt climate change. PNAS. 14308-14312 (2008).

Held, H. and Kleinen, T. Detection of climate system bifurcations by defenerate fingerprinting. GRL. 31. L23207 (2004).

Brock, W. A. and Carpenter, S. R. Interacting regime shifts in ecosystems: implication for early warnings. Ecological Monographs 80, 353-367 (2010)