We would like to thank Alex Robel for his thoughtful and helpful comments on our manuscript which have led to considerable improvements from our original submission. Alex's latest comments are given below in bold, with our response to the comments together with the changes made to the manuscript are given in blue.

Overall, this revised manuscript provides an excellent proof of concept for using Early Warning Indicators (EWI) as a tool for diagnosing the Marine Ice Sheet Instability in ice sheet model simulations. The revision now also serves as an excellent primer on the theory and practice of calculating EWIs that will be useful to those in the glaciology community who are new to this topic. My overall recommendation is that this paper is now close to being ready for publication, with a few more suggestions, which could be classified as "minor".

My more substantive suggestions:
1. In my original review I was fairly convinced that this study primary utility was as a proof of concept for showing how EWI could be identified observationally for MISI. However, in this version, it isn't until lines 352-354 that it became clear to me the exact utility of using EWI-type analysis on simulation output, beyond as a proof of concept. Here the authors state that this approach could be useful to lower the computational expense needed to verify the presence of bifurcations in a model simulations (i.e. without very long simulations). As a modeler, I find this to be a compelling argument, and should be front and center in the paper, particularly in the first section where you are trying to entice readers to continue learning more about EWIs.

We agree that this is an important potential application of the EWI and Alex makes an excellent point that this is not mentioned until very late in the manuscript. We have added several sentences covering this point in the second paragraph of the introduction.

2. Not to hammer too hard on the window length issue, but I think there needs to be a bit more discussion of two points.

(a) First off, on lines 280-282, you state 300 years is "the shortest window size for which the DFA indicator provides an accurate prediction for all tipping event". What does "accurate" mean in this context? From Figure 7, I would guess you mean where Kendall's Tau has a maximum, but why does that make this window the most accurate way to determine whether a tipping point will occur? In principal, wherever 0<tau<1 is indicative of increasing indicator, so one could argue that window lengths anywhere from 200 to 400/500 years seem to work in this regard.

This statement is admittedly rather strongly worded since it is true that a positive tau coefficient was obtained for smaller window sizes. Our intended meaning was simply that below ~300 years for all indicators and tipping point events (but closer to ~200 years) for some, the tau coefficient becomes less strongly positive, indicating a less clear increase in each indicator before the tipping point. This in practice would represent less certainty in early warning detection. We have reworded this and also the section that introduces the Kendall's tau coefficient so together these changes hopefully make our intended meaning clearer.

(b) There is also still not much investigation of window lengths relevant to observational time scales (i.e. decades). I wonder if increasing the forcing rate to levels that might be relevant to modern climate change (i.e. degrees/century) would make short window lengths more useful? This is perhaps beyond the scope of the paper, but if its computationally feasible to run a few simulations with higher forcing rate, and re-calculate figure 7 for those forcing rates, you could

begin to answer the questions of whether it would be possible to use modern observations to calculate EWIs.

We agree with Alex that this is something very interesting to look at in the future but think that it is beyond the scope of the current paper. This question of what determines a 'minimum' practical record length to apply EWIs to observations is a study in and of itself and would require a large number of computationally expensive simulations to address properly – not just by exploring how it is affected by window size but no doubt other factors such as the frequency of variability in the forcing are highly relevant. The choices that we have made in this paper represent carefully selected values and given the nature of this paper as a first step in applying EWIs to ice sheet modelling, and the significant delays to publication that these lengthy simulations would involve, we prefer to reserve this suggestion for future studies.

3. In my opinion, the point that is made throughout about separating tipping points by using slow forcing is a bit off the mark. Because you are not doing fully steady-state simulations, it is not clear that you have shown that these are actually three discrete tipping points (in the traditional sense of a bifurcation diagram which traces the stable manifold of the system). In the real system, these three tipping points likely involve a grounding line retreating over a region of reverse sloping bed with some prograde bumps. Perhaps if the forcing were even slower, there would be places where the grounding line stabilizes on these bumps, subdividing these tipping points even further. On the other hand, for realistically fast forcing rates, you would have a "tipping point cascade" that might look like one single rapid retreat. This would still be of interest, and could be valuable to identify using EWI since it is closer to what we are likely to see in reality. This is all to say that without a completely steady-state analysis it seems a bit premature to argue that you have found the three actual tipping points in the system, when there may be more in a mathematical sense, or when they might combine into one tipping event under real forcing.

We agree with many of these points and they are largely covered in the manuscript already. We state at the start of our results that 'We focus our results on these three major changes in the glacier configuration and ignore any possible smaller tipping points that do not result in significant grounding line retreat or changes in ice volume'. We do not claim anywhere that there are only three tipping points but state that three distinct tipping points can be identified from our model runs. While it is true that we first identify 'potential' tipping points using transient calculations, but we then go on to do fully steady-state simulations to verify that these are tipping points in Section 4.2. It is also true that a faster forcing might cause all three to be crossed at once and we cover this point in the manuscript. Finally, we agree that it is possible that a slower forcing, or alternatively using a smaller interval in the control parameter for our steady-state simulations, might identify additional potential tipping points. We have altered our wording to better reflect that we have not necessarily identified every single tipping point in a mathematical sense and added a sentence to the discussion on this point. In general, however, we chose to continue to refer to three tipping points because we are very confident that our methodology detected the three largest and most societally relevant tipping points in our model simulations.

Minor issues:
Line 12: what does "this" refer to in this sentence? Replaced 'this' with 'ongoing and future changes'

Line 18: indicators in model simulations robustly detect Done

Line 23: delete "a major component of the earth system" to make sentence clearer This part of the sentence was originally added in response to a previous request to clarify what is meant by a 'tipping element' but we have changed the sentence to hopefully make it read better.

Line 27: If grounding line retreat causes grounding line flux to increase Done

Line 31: a small perturbation results in the system Done

Line 43: delete "externally forced" since there needs to be an external forcing trigger for MISI to occur in the first place Done

Line 47: I'm not sure you need to mention the lower stable branch since technically it does not "participate" in the bifurcation (and there doesn't need to be a lower stable branch at all to have a saddle node bifurcation) The lower stable branch is shown in figure 1 and as is commonly done with diagrams of this type and so we prefer to keep this description.

Line 46-55: It would be useful to indicate what are the assumptions under which it is the case that MISI is a saddle-node bifurcation? i.e. that bed slope is negligible and changes very slowly in space (i.e. Schoof 2007/2012) These are the assumptions used by Schoof but do not necessarily represent necessary assumptions in general for the MISI to be a saddle-note bifurcation. We show that the MISI tipping point behaves as a saddle-node bifurcation but these assumptions are not made in any of our simulations that use a realistic geometry.

Line 51: parameter range Done

Line 58 and elsewhere: I always thought this was called "critical slowing down". A cursory search in the literature indicates that this is the most common usage and should perhaps be used that way here if you want readers to relate this to the broader EWS literature. Replaced all instances with critical slowing down

Line 76: are you calling these EWI or EWS? You use both in the same sentence We believe this usage is correct: the early warning signals manifest in the data or model output e.g. critical slowing down but this are analysed with the use of early warning indicators e.g. lag-1 autocorrelation.

Section 2.1: So, I am typically loathe to reference my own papers, but I think it bears noting that Robel et al. 2018 shows analytically that the response time (calculated directly from the eigenvalues of the system) increases towards the MISI bifurcation (see Fig. 3 in that paper). This paper is certainly directly relevant and was missed in a literature search so we are grateful to Alex for pointing it out and a sentence has been added to section 2.1 on this point with the reference added.

Line 123-124: This sentence could be written a bit more clearly since its unclear what you are saying about variance here This explanation has been expanded

Line 154: typo at exponent Done

Line 191: Also, by using realistic noise you can assess EWI detectability that would be expected in observations Yes indeed, added this point

Line 199: How do you know what is not related to system recovery time? Sentence is a bit vague Reworded this to make it clearer what we mean

Line 220-224: can you be clearer about the different interpretations of tau=1 and 0<tau<1? We have rewritten this section slightly to hopefully make this clearer

Line 247: MISI event begins Done

Line 314: of ice flow are Done

Line 337: and also decreasing window length? True but this is mentioned elsewhere, the point here is specifically that increasing distance from the tipping point eventually reduces the predictive power of EWIs

Line 346: related to issue #3 above, but it isn't clear why this cascade is a problem from the POV of EWI detectability We have considered this further and realise some confusion may arise from our using of the term 'tipping cascade' which has been used with various meanings in the past. Perhaps the stricter meaning is one tipping point causing a second tipping point to be crossed and in that sense this is not precisely what we are referring to here. A tipping point can be crossed either through changes in the system state or the control parameter. However, these processes are not instantaneous and if the control parameter is changing sufficiently quickly then it could trigger a second tipping point to happen soon after the first in a way that might not have happened if it were held constant or changed very slowly after the first tipping point was crossed. We called this a tipping cascade in the sense that one tipping point 'cascades' into the next but it is a different mechanism than the definition above. That being said, cascading tipping points are also an issue since they would complicate any interpretation further and the methodology we use whereby a control parameter changes continuously with time could not distinguish between one tipping point or two cascading tipping points. We have changed the wording here to make our meaning clearer.

Line 349: tipping points is one that Done

Figure 4: in panel (b) you have open markers where you have done steady-state calculations, but not in panel (a), so it is unclear where you have actually done simulations to determine steady-

states. This is done just to improve the readability of panel a and avoid clutter and the last two sentences of the figure caption explains where steady-state calculations were done but we have expanded on this explanation to make it clearer.