First of all we would like to thank the referee for a very encouraging and thoughtful review. Many good points are addressed and we feel this discussion will help to improve the paper. Note that there is a figure at the end of this letter in response to one of the comments below. We feel that its inclusion in the main text is not necessary as its "bulk" behaviour is given in Fig 5 of the paper and other aspects are explained in the text, and it might add "clutter" with 4 additional images -- but this is at the discretion of the Editor.

"It is possible that our snapshot calibration is equifinal" : that seems likely – if you have the same number of beta values as velocity values, and a normal stress too, then even in 1D there is a null-space. I think it is a vector made up from a perturbation to beta-squared one cell upstream from the face and a perturbation to the normal stress. But even if that were eliminated, there are a number of vectors associated with small singular values – such as oscillations in beta-squared some way upstream – which might end up being determined by the initial guess/choice of iterative method/regularization rather than data

I agree completely with the above statements. In this passage, however, I was referring to the possibility of compensation between heterogeneous control variables.

Still, the issue you bring up is a real one, and is wrapped up in the issue of subjectivity of priors – this is highlighted by Figs 4(b,c), which you mention below. As is commonly done, a smoothing term is applied to \beta in the cost function, the weighting of which for the snapshot calibration was chosen on the basis of bounding large variation over short length scales, and was applied to both snapshot and transient calibrations, there being no rationale to use different values. It is possible, however, that this prior information manifested differently in the different calibrations, leading to the "rib" features being more prevalent in one than the other. This highlights the subjectivity of such weightings and the need for minimizing subjectivity in prior specification, though we do not attempt to address this issue here. Mention of the above has been made now in section 5.2.

The additional information provided by the transient observations is sufficient to generate a better ice-stream state estimate" is a big claim in that case (not saying it is not true), but how does it come about? It seems to me that the transient calibration might work out better just because it matches velocity and surface in time. Put another way, the snapshot might be weaker largely because it mismatches, so that it insists on acceleration extending further upstream from the grounding line than it ought, which would look like a lighter pull (more buttressing) on a weaker bed.

This last point is a very good one, and one that had not occurred to us. This sentence from the text is now expounded on a bit further.

Stronger bed: Not uniformly stronger, though ? There is also an interesting ribbed structure in fig 4c (with a rib of strong bed close to where the GL seems to slow in the prediction).

Yes, as mentioned above (and now in the paper) this rib structure does arise, and it is difficult to know whether this is a result of different observations/different equations, or effectively less-strong smoothing. We do not make any speculation as to the physical underpinnings of the rib structures.

Negative buttressing: I like the idea that an Hσ that is larger than the non-ice shelf value might

imply that the DEM h is too low. Might there be another explanation, too? That some parts of the grounding line are being pulled by faster flowing parts via the ice shelf. In that case you might expect the negative buttressing to line up with shear margins, which looks like it might be the case in fig 4a

This is a very good point and one which was considered but ultimately left out of the paper in favour of the other hypothesis. We now make mention of it.

Abstract: 'inverse methods'. This seems a bit slang to me.

We agree that this might be an overused or misleading term – particularly in instances where it has not been established that there is a unique minimum of the cost function in the control space. We have changed the mention in the abstract to the more descriptive "control method".

P4465, line 14-: The text doesn't actually say which method (AD, correct?) is used to compute the gradient of Jtrans. Is there space for a one or two sentence summary of the particular AD method?

This is explained in the reference (Goldberg and Heimbach, The Cryosphere, 2013) but we now say which AD tool is used.

P4467, line 27 : not so much the thickness, but the vertically integrated effective viscosity including crevassing etc.

Well, there is a distinction here between "estimate" and "control variable", I think you are saying I should limit this to the latter. But one could argue that both the control variables and the derived model state can be considered "estimates" which is why the sentence is worded as it is. In any event, the issue being addressed by this passage is now discussed in more detail in the Discussion.

p4470, line 1 ; 'high accuracy': maybe give numbers

done.

p4470, line 11, 'very weak bed' : perhaps give a number

done.

P4474 : line 8 : 'decreasing beta anywhere increases ice loss, lowering the bed only increases ice loss upstream of the projected 2041 grounding line.' is that quite correct? For the most part, there seems to be no sensitivity to beta downstream of the 2014 GL. The region where it seems to matter most and the bed does not looks to correspond to a grounded promontory in 2014. Is that bit lightly grounded?

If I understand the feature to which you are referring, it is lightly grounded, you can see this in Fig 6.

As for the region downstream of the 2041 (not 2014?) grounding line, we were bringing attention to the fact they are positive, albeit not as large as the negative sensitivities upstream. Certainly the sensitivities are small here – but a point made in section 6.1 is that on the whole bed elevation

influence is marginal, or at least not overwhelmingly large.

Fig 6 : I'd like to see the same figure for the snapshot calibration. I'm guessing it has more even thinning?

Please see the figure on the following page. We have elected not to include this figure in the paper as we feel its point is made by Fig 5 – there is very little grounding line retreat. Your intuition is correct – thinning is less skewed toward the grounding line than in the transiently calibrated simulation, and is generally even aside from strong thinning in Kohler trough and, toward the end, strong thinning at the upstream boundary. This latter feature is likely seen because the input fluxes that were inferred from the transient calibration (and, we believe, minimize anomalous thinning) were not used in this simulation, and those that **were** used were too small.

We point out, though, that thinning and ungrounding does eventually occur downstream, and the slight thickening signal apparent at first does reverse. This could indicate that the anomalous thickening (relative and absolute) may indeed be a transient resulting from the snapshot initialisation, and the long term tendency is, indeed, retreat.

Fig 7: the legend has 'linear friction parameter', which I confused with a linear sliding law until I read the text properly. Maybe 'time dependent friction parameter'

Agreed, thanks for pointing this out.

Fig 9 (a): Do the upper schematics (the views of the front/join) add much? The planview could be larget without them

If you think this improves the understandability we are happy to make this modification, thank you.

Figure 1. The equivalent of Fig 6, corresponding to the snapshot calibration (calibrated to MEaSUREs velocities with 2002 surface DEM).