Review of Wang et al. (2025) ‘Quantifying Temperature-sliding Inconsistency in Thermomechanical Coupling: A Comparative Analysis of Geothermal Heat Flux Datasets at Totten Glacier’

Summary
This paper presents a new set of metrics to assess the inconsistencies between modelled basal temperatures and observed surface ice velocities in ice-sheet simulations, and applies them to evaluate the likely shortcomings of eight different geothermal-heat-flux datasets in the Totten basin in East Antarctica. The paper finds that its results in this regard agree with previous work that assessed the eight datasets using radar specularity observations to determine the presence of basal water and thus whether the ice was warm or cold, which suggests it is indeed performing well in identifying problem areas.

I think the new method is a useful way to quantify the mismatches between modelled ice temperatures and observed velocities, and the paper is well-executed. I do think it would benefit from a small amount of restructuring, as presenting the application of the method and the comparison of the geothermal datasets at the same time makes it harder for the reader to understand what’s going on. I also find it dissatisfying that the method cannot say anything about the source of the inconsistencies, however. Pointing out where the inconsistencies are is, of course, a useful first step, but without knowing where they’re coming from, there’s not necessarily much one can do with the information. That would obviously be too much for this paper though, and the authors suggest this might be the focus of future work – I would strongly encourage them to pursue it, as that would be really very useful!

Overall, I think the paper is suitable for publication subject to minor revisions to improve the clarity of the presentation.

Page and line numbers refer to those in the clean version of the submitted manuscript.

Major Comments

- Focus of the paper: This is a a bit of a minor gripe, but I feel the focus of the paper is not always very clear: is it presenting the new set of metrics or is it assessing the differences between the GTF datasets? I think it would be clearer to start with one simulation where the authors show how the method works in practice, and then move on to the eight simulations comparing the GTF data. As it stands, the authors are trying to do both things at the same time and the consequent seesawing makes it harder for the reader to understand what the metrics are showing in concrete terms. This need not be a new simulation or anything – just pick one of the eight and add a short section to the results showing the fields for the metrics and explaining the interpretation, before launching in to the full-bore comparison.
- Future steps: This paper is very much a first step – yes, knowing where the inconsistencies are is good, but knowing why they’re there is a lot more useful if one wants to improve one’s ice-sheet model. I know the authors hint at this being a possibility for future work, but maybe they could suggest in the discussion some ways in which modellers might use the method to investigate the causes of inconsistencies? This would make the paper a lot more directly useful to the ice-sheet-modelling community, one feels.

Minor Comments

- p.1, l.30-35: I can’t quite work out from the phrasing in the abstract for the first inconsistency if this means the GHF datasets are too cold or whether the model’s too cold. Similarly, with the second one, are you saying the model is overheating compared to the GHF datasets, or that the datasets themselves are too hot? Or both? Obviously, it’s explained later, but best to not confuse people with the first thing they’ll read! Some rephrasing to make it clearer what’s going on here would be beneficial.