

Interactive comment on “Inferring the basal sliding coefficient field for the Stokes ice sheet model under rheological uncertainty” by Olalekan Babaniyi et al.

Douglas Brinkerhoff (Referee)

doug.brinkerhoff@mso.umt.edu

Received and published: 5 October 2020

In *Inferring the basal sliding coefficient field for the Stokes ice sheet model under rheological uncertainty*, Babaniyi and co-authors present the application of a Laplace-approximation based Bayesian inference procedure to the problem of determining the distribution of basal traction given observations. In particular, the authors make a significant contribution to inverse problems in glaciology by demonstrating what happens when we ignore the uncertainties in model parameters apart from the ones that we're primarily interested in. In particular they show that ignoring a reasonable amount of rheological uncertainty leads to a posterior distribution that is egregiously incorrect and

C1

far too sure of itself. They address this problem by introducing the Bayesian Approximation Error approach, in which random samples from the nuisance variable are used to generate an empirical covariance and bias in the model predictions, which are added to observational noise to produce a much more realistic pseudo-likelihood model and thus a much more reasonable posterior distribution.

This paper is an excellent and timely contribution. First, it builds upon a lineage of important papers (Bui-Tanh, 2013; Petra, 2014; Isaac, 2015) that have provided a framework for (Bayesian) uncertainty quantification in large-scale ice sheet model (something that is of critical importance, in my view), while also demonstrating the limitations of those methods that still remain. In some ways, this paper raises more questions than it answers (and I mean that in a good way): if even a relatively robust inversion framework like Bayesian inference can be led astray by unquantified aleatoric model uncertainty, how can we go to deal with unquantified epistemic uncertainty going forward? Second, numerical recipe that it introduces is general and viable across a large class of problems in ice sheet modelling, so long as there are a sufficient number of cores available for drawing Monte Carlo samples. The development of "embarrassingly parallel" methods for uncertainty quantification (as opposed to ones that require more intricate communication) make the implementation of these methods broadly feasible for other researchers. Finally, the results are compelling. The posterior distributions that the authors produce are indeed much more realistic than those produced in the absence of their method, and any work that makes the case for a *broader* posterior variance in ice sheet model inversions should be lauded.

Nonetheless, There are a handful of factors that may limit the paper's applicability. I describe these issues below, interspersed with a handful of technical corrections.

P6 L15 The choice of covariance function should be motivated. First, why bother with this inverse elliptic operation to begin with, when it doesn't seem any more physical in this case than using squared exponential or Matern kernel? If elliptic in-

C2

verse is justified, then why the exponent of 2? This differs from that used by Petra (2014), so why the change?

P7 L10 Fix reference formatting.

P8 L7 It would be better to use Isaac (2015a) as the reference here, since the "generalized eigenvalue problem" formulation does not appear in the previous works (although the solution to the problem stated in those works is equivalent).

P8 L12 What does *sufficiently* mean?

Eq. 14, 16 This 'right arrow' notation isn't very informative.

P9 L13 ϵ_* is defined (empirically) later, but it should at least be given a name or description here.

Eq. 21 This method seems to rely on the modelling error being well-approximated by a normal distribution. Is there any evidence to suggest that this is true for the problems being considered here? (The forward model is non-linear after all). What happens if it's not?

P10 L11–18 These 'rules of thumb' need some explanation. What are they telling us and why are they different? A qualitative description would help.

P11 L15–19 This is a very unrealistic prior relative to the typical situation in ice sheet modelling. In what case do we typically know the true value of a parameter's mean? That's a lot of information! I think a much more interesting case would be to provide a vague prior with a mean not equal to the true value. A frequent assumption for GRFs is mean zero and a large marginal variance, because there isn't any more information than that to go off of *a priori*. Does this method still work for a less informative prior, or does it ruin the estimates of ϵ_* and Γ_ϵ ?

C3

P13 L1 Why introduce rejection sampling when one could use a similar log-transform as for the enhancement factor? Why does constraining the parameter space to shear thinning matter in this synthetic case anyways?

P13 L13 Where does the weird constant 3 come from in the rate factor parameterization?

Table 1 These values on prior distributions seem arbitrary, and because they're parameters of an elliptic model rather than the more interpretable length scale and variance that we'd find in a more traditional GRF kernel, they are a bit opaque. Why were these values of γ and δ chosen?

Sections 5/6 I think that these sections should be reordered to put both setup and results for each example right next to one another. This would be more brief, and the reader wouldn't have to flip back and forth between sections.

Figure 5 and supporting text It's unusual for vertical velocity to be available in an inversion (we don't really get it from remotely sensed data). Would it change results much to not have access to it?

P17 L6 'Finally, we see that the off-diagonal blocks also display fairly complex behavior' is not a useful sentence. Complex in what way, and why?

P27 L6–7 'We omit...' This sentence is redundant.

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2020-229>, 2020.

C4