Review of 'Inferring on-glacier meteorology from physical modeling and remote sensing' by Ren et al.

Doug Brinkerhoff

July 31, 2026

This article describes the application of a Bayesian scheme for inferring temperature and precipitation biases in downscaled ERA5-land such that when used in conjunction with a surface energy balance model, the resulting predicted SMB and albedo are consistent with some independent satellite derived products. This method is applied to 4 glaciers (2 in the Alps and 2 in high-mountain Asia), and the incorporation of these additional constraints is shown to improve model fidelity to a held-out set of stake-based SMB measurements. In general, I find this manuscript to be a step in the right direction via its more-careful-than-most treatment of uncertainties, and its incorporation of novel observational datasets. However, it also suffers a little bit from confusing descriptions in a few places, and also from what is in my view an insufficient acknowledgment of some of the obstacles that stand in the way of the broad generalization of the methods presented here.

1 Line-by-line comments

L95 In the broader context of this paper, I am not sure that I can co-sign the assertion that mass/energy balance models of the class described here are more easily transferable – the paper itself discusses the need to do a surprising amount of hand-tuning and subjective modification to inversion procedures, even after incorporating additional constraints.

P3.2.2 This continuity approach (ASTER SMB) to surface balance inference should be viewed as exceptionally error prone. First, the actual uncertainties in each of the remotely sensed constituents to the calculation (surface elevation change, surface velocity, ice thickness) are all not well constrained. In particular, they all have spatial autocorrelation that does not appear in reported errors (with the exception of Hugonnet's dH/dt). More challengingly, using them here requires the computation of numerical derivatives of the computed flux, which amplifies the error in the underlying products. While the authors do a credible job of showing that this product does a reasonable job of reproducing observations for some of their glaciers, I think that this product should be presented with much more skepticism and a fairer accounting of its limitations.

L230 The asterisk operator should probably be a $\times$ instead, here and elsewhere.

L221–239 This section would be more helpful if the downscaling methods that are currently relegated to a reference had a very short (like one sentence) description accompanying them.

L231 If the goal is to develop a method that is spatially transferable, then using different methods of downscaling for different glaciers is less than ideal. I think it would be better to either apply a precipitation lapse rate in all cases or in no cases.

Eq. 2 This equation is incomplete without either (1) an explicit (mathematical) statement that θ lives within bounds, or (2, preferred) retention of the prior term.

L276 It's not clear that MCMC is actually intractable here - especially if one were to start with a good initial guess, they would probably yield a significantly higher sample density in the vicinity of the MAP point. Brute force methods are good when they can be applied (like here), but it's worth mentioning that this will not be feasible in the event that the size of the parameter space were to increase (or if hierarchical priors were adopted to enforce spatial correlation between locations on a glacier).

Eq. 7 It's worth being a bit more rigorous in the presentation of this equation, and to describe in what ways $\theta_{inferred}$ might end up being a poor approximation to the true posterior mean.

Eq. 8 It's also worth noting that very small ESS (i.e. size approximately 1) is the characteristic of the circumstance in which *none* of the sampled parameters provide a good fit to observations, and so the one that is the least bad absorbs all of the probability mass. Perhaps comment on whether that is occurring here?

L306 I am totally lost when it comes to this pre-inference prior choosing business. Where does the modelled SMB used in making this distinction come from? Is this some sort of peculiar form of empirical Bayes?

Eqs. 11–15 Don't use the asterisk.

L355–359 I don't follow this spatial extrapolation reasoning. Please try to describe more clearly.

L378 and elsewhere I am not sure that I understand the value that the cloud frequency comparison provides, since this is not strongly indicative of either of the inferred values. I suppose if there were never any clouds, you could confidently say that the precipitation was zero, but I don't think that's the case in any of your domains.

Fig. 4, row 2, col 1 The concentration of samples at the edge of the plot like this is a classic indication of prior mis-specification, and a very dangerous side-effect of using uniform-in-a-box priors – the parameter wants to be outside this bound, but the prior doesn't allow it. Is a temperature correction of 1.99 actually infinitely more probably than a correction of 2.01? For the other plots - these exceptionally concentrated posteriors might imply that the data uniquely constrains a single model - or it might imply that all models fit badly.

Fig. 5 I don't understand why the gray dots are different between panels. Wouldn't it be better to use the same gray dots and just change the methods?

Fig. 7 This plot is pretty incomprehensible to me. There are a large number of axes, colors, and unlabeled numbers. It would be better to figure out how to present this information in a more targeted (and larger) format.

L427 This line about spatial variability in precipitation correction factor is sort of unsatisfying. First, the associated plot is relegated to the supplement. Second (and more importantly), there's no discussion of what these precipitation patterns might mean. In the supplemental figure, we see theses factors varying wildly across pretty short spatial scales, while there also doesn't appear to be any temporal correlation between the same locations. Why? If the climate model/downscaling method is just missing a piece of physics, then shouldn't this be represented in these recovered biases? The patterns here seem to me much more likely to be a result of mis-specification of either the physics or the likelihood. I recognize that hierarchical or spatially tied modelling is beyond the scope of the present work, but the inferred correction factors would be much more defensible if they had some modicum of agreement with one another.

L4684 It would be nice if these were reported alongside a notion of relative error. Does a 1m RMSE correspond to a 10% or 100% error relative to the stake measurement (recognizing that this gets tricky when mass balance is near zero).

L524–527 I am having trouble parsing what this line means – how do these observed products make a prior inappropriate? Is the prior wrong, or is the likelihood model associated with this product wrong?

L572 It's worth mentioning that adjusting your prior in response to a perceived inadequacy in the data product is a little bit of a Bayesian crime and that having to make these sorts of judgement calls limits global applicability of these methods.