The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2017-275-RC2, 2018

Interactive comment on "A Bayesian Hierarchical Model for Glacial Dynamics Based on the Shallow Ice Approximation and its Evaluation Using Analytical Solutions" *by* Giri Gopalan et al.

Anonymous Referee #2

Received and published: 6 June 2018

Overall Review:

I believe this is an interesting, useful contribution and publishable with some revisions. Essentially, your computations assess the errors in using the numerical approximations for "f" using analytical solutions as a base line. That is, you generate "Y's" with analytical solutions but then forget about that and use numerical approximations in the BHM. "error" is then viewed as differences between Bayesian results and the analytical "truth". This is valuable work, though as you make clear, it doesn't make any assurances when the analytical model is "replaced by nature" in producing data. You also considered several cases, but I do think that your paper would be strengthened if you

also studied the impact sampling plans and sample sizes (ie. What if "every other observation (in time) was removed? This is also critical in judging the impacts of your approximations used in computations (see the next paragraph).

My first concern is correctness of all contributions. Errors can occur when manipulating equations rather than probability distributions. I think yours turned out right, but all conditioning assumptions are not clear. Consider Appendix B1 beginning at the bottom of p. 20. The "overall model" as written at the top of p. 21 is quite brief and does not include probability assumptions. I sense that you understand the key issues based on the sentence in lines 16-17, p. 21. Namely, equations like Y = m(variables) + error are code for "the conditional distribution of Y given "variables" and the mean of "error" = 0 and some variance of "error" has conditional mean m and conditional variance equal to the variance of "error". The assertion that all "errors" in you models have mean zero seems to be missing, but more importantly, when you do the manipulation leading to line 14, you must have assumed both models for Yck and Y(c-1)k are conditioned on the same quantities so you can simply subtract their conditional means, etc. Further, simply taking differences of Yck and Y(c-1)k is based on their joint distribution, so cavalierly moving Y(c-1)k to the left hand side and claiming you're now looking as the distribution of Yck given Y(c-1)k and the other variables. That requires a probability computation (moving from joint to a conditional distribution) in general. Fortunately, it is common that the algebraic versions can actually be proven to be correct probabilistically for "linear manipulations", but in complicated settings, this needs to be checked (based on my quick check, I think you're OK but think you should check as well). This all relates to my suggestion that your model isn't simply lines 2-4, p. 21. What are the conditional distributions assumption (the Z's are independent etc.)? One more related issue involves discussion of inference for X's. In a sense, you should be careful in posterior inferences about both S and X simultaneously, given q. (they are simply linear functions of each other). Again, I think you're OK but it merits your attention.

I think the approximations you used on p. 21 are reasonable, but a bit more defense would be good. Further, I'm not comfortable with the way you needed all the approximations so that you could use grid sampling to claim genuine posterior inference. I think that you could ski p the approximations and did a full MCMC approach, it wouldn't be as easy as what you did but it's not that much harder. I think you should at least try some MCMC to confirm your computations and approximations. Further, what is the dependence of the value of your approximations on f. Surely you need to answer this if you plan to suggest operational use of you programs as you suggest you will do in the future.

Other Notes:

(1) The model for X is an explosive autoregression and hence you have built-in a limitation. A non-explosive model could be Xj = r Xj-1 + error where 0 < r < 1. If you make r a parameter and let the data tell you about r, you may be able to predict further in the future if the data suggests r can be much smaller than 1.

(2) I think you missed emphasizing a crucial (and related) contribution of Berliner et al (2008). Namely they also treat model error through their "corrector process" and this should be mentioned.

(3) As a minor point, you should include at least one reference to

Berliner, L.M. 1996. Hierarchical Bayesian time series models. In Hanson, K. and R. Silver, eds. Maximum entropy and Bayesian methods. Dordrecht, etc., Kluwer Academic Publishers, 15–22.

The references by Wikle and Cressie both reference it but you should too since it urges the "data model, process model, parameter model" view. Also, since that paradigm is so key in your paper, I think you should break out the formula in line 22, p. 2 as a separate line for emphasis.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-275, 2018.