The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2019-54-RC1, 2019

Interactive comment on "Assimilation of surface observations in a transient marine ice sheetmodel using an ensemble Kalman filter" *by* Fabien Gillet-Chaulet

Dan Goldberg (Referee)

dan.goldberg@ed.ac.uk

Received and published: 28 June 2019

This study is one of the first to apply Ensemble Kalman Filtering methods to an ice-sheet model with a nontrivial stress balance (attempts have been made with Shallow-ice models). Such methods, rather than using deterministic means to optimise a cost/misfit function in order to infer hidden properties from error-prone observations, essentially generate an ensemble meant to encompass a probability distribution, and repeatedly apply model dynamics and bayesian inference on this ensemble in order to refine the statistical properties of an unknown state and parameter set. The paper implements a variant of the EnKF known as the Ensemble Subspace Transform

Kalman Filter, which is simply a particular choice and one which is meant to avoid costly computation of an ensemble covariance (which i do question – please see specific comments).

The methodology presented is really just one step in a long long road toward operational ice-sheet forecasting (compare with over a half-century of development in numerical weather prediction) but an important one – especially when considering there is at least one person with a question about filtering methods every time a talk is presented in state and parameter estimation for ice sheets (as the authors have pointed out, there are already others working toward the ice-sheet version of the other main tool of weather prediction, 4Dvar). Thus i feel it is a worthwhile study which should be worthy of publication as a methodological investigation (and the authors frame it as such, at least in the conclusions section. However I do think the manuscript needs work before this can happen.

On the basis of the extent of the comments below I choose "major revisions" – but there is no formal definition of what this means, and the editor may choose to ignore this classification. I am not suggesting modification of the algorithm and/or results, simply clarity of text.

GENERAL COMMENTS

For one thing upon reading I had significant detail understanding what was done. There were a number of points on which i felt clarity was needed, and most of these are addressed below in line-by-line comments so I will not list them here. Note that these specific questions compose the bulk of my review – and this is because without having a better idea of what was actually implemented, it is difficult to critique the results further!

However something I will state in the general comments is that despite similarities, ice-sheet models are distinct from e.g. atmospheric models in that the unknown parameters most sought cannot generally be observed directly, in contrast to models in which the initial conditions in a forecast/analysis cycle represent the parameters of greatest interest. This is exemplified by the fact that the “state” vector contains non-dynamic variables (friction and bed elevation) and the fact that the observation operator, rather than being a simple averaging or restriction, encapsulates a fully nonlinear solve of an elliptic partial differential equation. I think this is something that should be made very clear to members of the climate and meteorological community who read this work.

One overall comment is regarding the distribution of the ensemble. As I understand it, even if the initial ensemble is evenly distributed given the prior, it is difficult to know a priori whether the projection of the ensemble will represent a favorable distribution of the projected space. That is, what if the forecast “clusters”, underrepresenting important regions of state space? As I ask below, it is unclear whether there is a “reinitialisation” of the ensemble at every step. Clearly this topic has already been considered in the NWP literature, for example Song et al (2013) makes use of a time-dependent adjoint (a tool the authors state the work here is meant to circumvent the need for) in order to generate a more representative ensemble.

I also question whether the “toy problem” proposed by the authors truly tests all of the difficulties a filtering approach might encounter. I bring this up in more detail below but some aspects of the approach seem to hang on the “locality” of the problem (a technique called “localisation” is employed to ignore long-distance correlations of the state). I wonder if this only works because the problem is one-dimensional with no buttressing involved, so essentially to a strong degree (though not completely) the velocities depend locally on basal friction and geometry? Would this still be a good approach in a 2D domain with an expansive embayed ice shelf (such as the Ross or FRIS), or very weak basal traction over a large part of the domain (such as Pine Island)?

The methodology essentially uses a whole “family” of geometries and velocities to infer hidden parameters of the system. This somewhat bears similarity to a differnet paper led by the author, “Assimilation of surface velocities acquired between 1996 and 2010

to constrain the form of the basal friction law under Pine Island Glacier” – aside from the statistical formality, and the introduction of consistency between these geometries by way of the continuity equation (which is actually not so consistent if the analysis updates do not conserve mass!!!) – I wonder if the author would consider comparing and contrasting these approaches.

Finally, i point out that, despite the divide between filtering and adjoint-based methods, there is a growing sentiment in NWP to take what is “best” from the various approaches and form more hybrid schemes (for instance the Song paper referenced above, see also Kalnay 2010). Therefore I urge the author to reflect on such innovations and how they might be useful in further developments for filtering of ice-sheet models.

Line by Line comments:

p2.l3: would be good to state this is a point when only resolving 1 horizontal dimension

p3.10 variational

p3.13 "use of linearised or adjoint models" this assumes a trivial mapping from model vars to observations – see my comments below.

p3.15 rewrited?

p3.35 – would be good to explain as soon as possible i.e. here what you mean by a twin experiment, or give a reference, as this is jargony

p 5 thru eq (10): this is a well written explanation of the EnKF. However I have a few questions which might be due to my lack of familiarity with filtering methods, but I think this might be true of many readers of this paper. This is also important as, though it is not the algorithm used, the one used is far more complex so this is a chance to explain your methods to the reader.

(a) is P^f = P_k?

(b) you do not say how the individual ensemble members (x_i^{k,a}) arise/are updated, only the state vector (which looks like the mean of the analysed ensemble)?

(c) is the posterior/analysis covariance used at all in subsequent time/filtering steps, as from eq 7 the covariance is always formed from the present ensemble – so i am struggling to grasp what is done in the algorithm in a multi-time step (k>1) framework.

(d) For each new forecast/analysis cycle, is the ensemble generated anew from the analysis-generate ensemble statistics?

(e) the formula given assumes normality of the ensemble does it not?

P5 eq(8): M_k is trivially the identity on the time-invariant components of x, i.e. b and C, correct?

p6.4. I am struggling to see why P^f need be formed, as it is a tensor product of X with itself (subject to (a) above). For instance, the last term in 9(a) is written

X (X^T H^T) ((H X) (X^T H^T) + R)^-1 dimension

so the largest matrix that need be formed is HX, and no matrix of (Nx x Nx) need be formed. Perhaps I do not understand where and how P^a is actually used however.

P6.5 I don't feel the concept of "error subspace" is ever suitably explained as i read the paper still wondering about this. \Omega as defined in eq 11 simply seems to be a "mixing" matrix that slightly changes the ensemble members – how is this an "error subspace"? (Assumiing that X \in R^{Nx x Ne} – i take it this is the case for eq 11 to make sense...)

Eqs 11-16: in contrast to the discussion of the EnKF this is very nonintuitive. You state (P6 line 7) that you approximate the covariance matrix by a low-rank matrix, which seems intuitive, but where is the equation describing this low-rank approximation and how is it done? (for what it is worth, low-rank approximations of covariance generally involve eigenvalue decompositions to retain the leading order covariance structure, but i do not see this here...

P6.23-25: can you give a more intuitive description of inflation? Why do you need it and what does it achieve? As it is I am not even sure if inflation corresponds to lower or higher rho.

P7, first paragraph. Im sorry but I am struggling to follow this paragraph. For instance, how does the non-linear observation operator applied to x_i lead to the product HX^f? i imagine they are related, as H is a linearisation of H (which, by the way, clashes with the symbol for ice thickness) but this is not explained.

P7.8. This assumption, i imagine, is valid in many NWP settings given the hyperbolicity of the equations. Are you confident it is a good approach for marine ice-sheet modelling?

P8.26. I was surprised by your suggestion that annual DEMs would be available over a multiple decadal period, as i do not know of such products for antarctica. The best i have seen is decadal or semidecadal with MUCH lower spatial resolution (e.g. Konrad et al 2017, GRL). Having skimmed the ArcticDEM website i do see mention of the spatial resolution, but not the temporal. Unless you can argue that such spatiotemporal resolution is reasonable and available, i suggest caveating this discussion by saying it is an idealised experiment and this is the type of spatiotemporal resolution to which the community should aspire.

P9.6: prognostic ice sheet models generally step forward the ice thickness, not surface elevation (as shown in your eq 4). In the analysis step you are updating z_s. Is there a simple mapping from your model state X to thickness?

P9.6: as mentioned in the previous comment you are updating z_s in each analysis step, which i am inferring then maps on to an update in thickness (tell me if i am wrong). Is this update at all volume conserving? If not should this be a concern?

P9.21-28: Lots of jargony language in this paragraph, likely not to be understood by the target audience. What is a sill and a nugget? You talk of the prediction obtained by kriging – is this something you have calculated? Is there any way to evaluate whether the ensemble does converge to it? Is this a way of evaluating whether the ensemble is large enough?

Section 4.1: Upon reading this, I realised that (a) i am unsure what time step you used, and (b) more importantly, whether each time step is a forecast/analysis step, as M_k in eq 8 could easily encompass multiple time steps – is this the case?

P10.15: again, these factors seem very important, and as mentioned above not overly well explained.

Section 4.2 – section headings should be capitalised

P12.25: Two comments about this paragraph: (a) Code that is continuously being updated and new algorithms developed might be an issue for *analytically* derived adjoint models, but not as much for automatic differentiation, which is specifically designed to generate new adjoint code when the "primal" code is changed; and (b) I return to the my confusion over the first paragraph on P7. Your observation matrix H contains, at the very least, a linearisation of the stress balance equation mapping geometry and basal friction onto velocity. It is not clear how you are finding this operator if not through some sort of forward model linearisation.

References:

Song et al, 2013: An adjointbased adaptive ensemble Kalman filter. Mon. Wea. Rev., 141, 3343–3359, https://doi.org/10.1175/MWR-D-12-00244.1.

Kalnay, E. (2010), Ensemble Kalman Filter: Current status and potential, in Data Assimilation: Making Sense of Observations, edited by W. Lahoz, B. Khattatov, and R. Menard, 24 pp., Springer, New York

---

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-54, 2019.