Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-454-RC2, 2016

Hydrology and Earth System Sciences
Open Access
Discussions

HESSD

Interactive comment

Interactive comment on "Hybridizing sequential and variational data assimilation for robust high-resolution hydrologic forecasting" *by* Felipe Hernández and Xu Liang

Anonymous Referee #2

Received and published: 1 November 2016

HESS-Discussions

DOI: 10.5194/hess-2016-454

Title: Hybridizing sequential and variational data assimilation for robust high-resolution hydrologic forecasting

Authors: Felipe Hernandez and Xu Liang

Review

The paper describes a "hybrid" algorithm for streamflow data assimilation. The algorithm, termed OPTIMISTS, mixes and matches elements from ensemble-based and

Printer-friendly version

Discussion paper

variational concepts. A number of assimilation tests are conducted for two watersheds, the first with ~1,100 state variables and the second with ~33,000 state variables. These tests assimilate streamflow data for 2 weeks or 4 weeks and then conduct streamflow "forecasts" using the final time of the assimilation period as initial condition. The skill of the forecasts is compared to a control forecast for which the initial condition does not benefit from the streamflow assimilation.

The authors claim that their new "hybrid" algorithm is superior, at least conceptually, to other hybrid approaches. They acknowledge that forecasts based on streamflow assimilation using the new algorithm are in some cases worse than those without.

The topic of the paper is appropriate for HESS. However, my summary assessment is that the paper is unusually difficult to follow and lacking in specific, clear explanations and results. Most importantly, it is never quite clear what motivates the choices made in constructing the new algorithm and how (and by how much) they are better than the corresponding elements from existing algorithms that the authors reject. The tests only show the difference between the new assimilation algorithm and the default (calibrated) model. But since the authors are touting their new algorithm as an improvement over other hybrid algorithms, they need to show that their algorithm outperforms the others. But I am doubtful that this would be possible, given that, on balance, the new algorithm does not seem to outperform the model-only results.

Overall, I do not think that the paper is suitable for publication in HESS at this time.

Major comments:

1) By far my greatest concerns is that the tests in the results section only show the difference between the new assimilation algorithm and the default (calibrated) model. But since the authors are touting their new algorithm as an improvement over other hybrid algorithms, they need to show that their algorithm outperforms the others, or at least is capable of improving the model-only results (which, on balance, is not the case). The statement (p11, L26) that "a formal comparison with 4DEnVar (Buehner et al 2010) or

a similar hybrid method deserves an investigation on its own" is woefully inadequate. It is important to keep in mind that the authors put together the new algorithm by picking and choosing elements from existing concepts at will, using "features that [the authors] consider most valuable" (p2, L32). The proposed new algorithm as a whole has not been proven elsewhere (as is the case for standard algorithms if the application satisfies the usual conditions of linearity and error independence and whiteness). Put differently, is the new algorithm optimal under those assumptions?

2) It is extremely difficult to follow the methods section. The authors simply state what they do (and sometimes more, see next comment) without clearly explaining what it is that motivates a particular choice and why this choice is likely better than what the existing algorithms do. For example, step 2 in the 7-step list on p3 (re. random samples) replaces the resampling in traditional particle filter algorithms (p4, L4). But why should this new approach better "avoid sample impoverishment" (p4, L4) than the existing resampling strategies? Note that I am not saying that it does not. My comment is that it is not sufficiently motivated, justified, and explained.

3) Sections 2.1 and 2.2 include three different approaches for the likelihood (eq. 3, 7, and 8), each representing a consecutively more aggressive simplification. But it looks like only eq. 7 and 8 were used in the tests (p8, L31-32). Why is equation 5 needed? It seems that the discussion lacks focus.

4) Abstract: The text here is far too generic and not reflecting the results of the assimilation tests. p1, L17-18 talks about "the benefits of the coupled probabilistic/multi-objective approach", but quantitative statements are missing (how much improvement?), and the abstract does not include any hint re. the fact that, on balance, the tests do NOT show an improvement from the assimilation compared to the model-only results.

5) The text includes a lot of optimization and particle filter jargon, but I am not sure that it is always used correctly. I am not an expert in particle filtering per se, but I am

convinced that breaking data assimilation algorithms down into "sequential" and "variational" is not a good way to approach the topic (p1, L7; p2, L4-5). There are variational algorithms (3DVAR, PSAS) that are sequential. Also, the Kalman filter update equation can be derived from the same objective function that is usually considered the starting point for the derivation of variational algorithms such as 3DVAR (that is, under certain assumptions 3DVAR and the Kalman filter are just different ways of solving the same problem). Finally, there are non-sequential (ensemble) Kalman smoother algorithms that are based on Bayesian principles and cover an assimilation window of non-zero length. How would the authors classify these Kalman smoothers? The point here is that the sloppy use by the authors of the most basic terminology raises suspicions that the more arcane details of the particle filter and optimization language are equally inaccurate, which casts doubts on the validity of the entire development.

6) Another example of an odd statement is on p1, L11: "...which promotes the reduction of observational errors..." I am guessing by this statement the authors refer to the reduction of the errors in the *analysis* or forecast estimates resulting from the data assimilation. The "observational errors" themselves cannot be reduced by data assimilation. Moreover, all assimilation algorithms promote the reduction of the error in the *analysis* estimates. This is not just the case for the algorithm proposed by the authors. Besides being sloppy, the statement therefore also provides literally no useful information.

7) Another sloppy statement is on p1, L21: "...geophysical models are as underdetermined as ever...". (also p1, L7-8) I am guessing that the authors mean that model *estimates* of parameters or states based on the *assimilation* of (relatively few) observations are underdetermined. The geophysical models themselves are not underdetermined. The optimization (or assimilation) problem is.

8) Section 1 (Introduction): The text is missing focus on hydrological data assimilation. Only the title reveals that the introduction is really about estimating streamflow, while references are a mix of hydrological assimilation and atmospheric or ocean as-

similation. While some features of the assimilation or optimization are shared across disciplines, an algorithm designed for an NWP system cannot just be applied to hydrological models without consideration of the fundamental differences in the two kinds of models. E.g., processes in GCMs are chaotic, but hydrological models are about damped physics. The dimensionality of hydrological models is far smaller than that of GCMs used on modern NWP systems. The point is that the Introduction could do a much better job of introducing the research and results.

9) p8,L17: The state variables include interception? Is there any chance that the amount of water intercepted by the canopy can be adequately estimated from the assimilation of streamflow observations? I would venture to guess that the estimated interception reservoirs are simply noise.

10) p8, L21-22 mentions three forecasting "Scenarios" but I could not figure out what exactly those are. There are two watersheds and 3 Experiments, each with a larger number of "factorial" experiments. But what do the authors mean by "Scenario"? Lack of this information (or its prominent exposition) makes it very difficult to understand the results.

11) p9, L25-26: "For Experiment 1, the correlation between the improvement of the NSEl2 during the assimilation period and the improvement during the forecast period was of -0.344, and of -0.669 for the NSEl1." I did not understand this sentence at all.

12) p10, L4: *Why* could it be that the performance is indifferent to n (number of particles or ensemble members?)

13) p11, L27-30: So why does the "conceptually superior" new algorithm (p11, L27) give such poor results of the test runs (p11, L30)? The authors do not provide an adequate explanation. Stating how the algorithm could be improved (as the they do in the subsequent lines) is not sufficient. At best, the new algorithm can still be improved, which needs to be implemented and shown. Thereafter it could be considered as a step forward. As things stand, the reader is left wondering whether there is a problem

Printer-friendly version

Discussion paper

with the new algorithm or whether the problem is ill-posed and standard assimilation algorithms would also fail.

Minor comments:

a) I am not sure whether this is a journal requirement, but it is odd to have the equations placed at the end of each paragraph. Typically, an equation is referenced but then shown only a few lines further down.

b) p9, L1-3: The "forecasts" use NLDAS-2 forcing data, which is equivalent to using perfect meteorological forecasts. Using the term "forecasts" throughout the manuscript is therefore somewhat misleading. The experiments are really more like "simulations" than "forecasts".

c) p9, L8: replace "valleys" with "periods between runoff peaks" or something similar? It is not obvious what is meant by "valleys".

d) p11, L1: Caption of Fig 7 says that the results are for Experiment 2, whereas the text here says they are for Experiment 3. Which is it?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-454, 2016.