Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-454-AC1, 2016

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

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

Felipe Hernández and Xu Liang

feh17@pitt.edu

Received and published: 14 December 2016

First we would like to thank Referee #1 and Referee #2 for their thorough and careful review of our manuscript, and for their very helpful comments. We will try our best to incorporate their suggestions, address their concern, and modify our manuscript accordingly when we are given the green light. We believe implementing their valuable comments/suggestions will improve the presentation, clearness, and accuracy of our work.

We agree that a comparison between OPTIMISTS and an established method would be highly desirable. We actually developed a Particle Filter (PF) with "regularized" re-sampling using a Gaussian kernel for this purpose, but were unable to run VIC reliably

for a single time step (we only managed to allow saving and recovering the routing state variables for simulations of at least a couple of days). This is due to a problem in our coupled version between VIC and the routing scheme, which will require additional efforts to be fixed. On the other hand, comparisons using the DHSVM model would not work because traditional PFs are not suited for high-dimensional models (this is precisely one of the problems we are addressing with OPTIMISTS) and, similarly, the DHSVM is not designed for modeling with coarse-resolutions.

Moreover, we had an additional reason why we were not totally convinced about including a comparison in the first version of the manuscript. OPTIMISTS is as much a new data assimilation (DA) method as it is a hybrid of existing methods. For example, it can be configured to behave as an evolutionary 3D-Var algorithm (select a sequential time step, evaluate many solutions with a single cost-function objective to be minimized, and do not use any root samples). Enlarging the assimilation time step transforms it into a type of 4D-Var. By using only root and random samples, a sequential time step, and an objective that evaluates the likelihood of candidate solutions given the observations, one can get something similar to a PF. Therefore, to some extent, the comparisons we performed between multiple configurations of OPTIMISTS could represent comparisons between those methods. We are making modifications to the manuscript to better convey this idea.

This flexibility in OPTIMISTS can be very advantageous, as it allows one to find configurations that better match specific conditions. However, as can be seen in the experiments, it comes at the cost of not allowing to consistently get adequate results if it is not well parameterized. Part of the objective of the experiments was to determine which components from the methods that OPTIMISTS borrows from are the most beneficial. It is very clear, for example, that sequential DA performs poorly in our tests, suggesting that traditional EnKF, PF, 3D-Var, etc. would most likely underperform compared to 4D-Var or some other method with time-extended evaluations. Similarly, one should configure OPTIMISTS to take advantage of the extended-period evaluations.

From Figures 6 and 7 it is clear that selecting the adequate time step will almost guarantee that the method will yield improved forecasts. On the other hand, choosing the sequential route will result in very poor performance. We will modify the manuscript to better convey the idea that, after selecting adequate parameters (extended assimilation time frame, multiple objectives), OPTIMISTS should, on balance, provide good results. In other words, it is not the entire range of possible configurations of OPTIMISTS that results in better forecasts, but a subset of them which the user should adapt based on their specific watersheds. The information shown in our results would help identify a good starting point for this adaptation.

Nonetheless, we used this opportunity to implement a 4D-Var algorithm and we performed comparisons on the VIC model. This implementation of 4D-Var uses a single-objective version of the ensemble optimization algorithm in MAESTRO and, given its evolutionary nature, is able to solve the non-linear optimization problem. A traditional two-term cost function (normalized squared deviations from observations and from the background) was used as the objective to be minimized. It must also be noted that, to speed convergence, we seeded the initial population with the root/base states (as it is done for OPTIMISTS).

Before discussing the comparison results we want to clarify the nomenclature of our experiments, which we will also improve in the manuscript during the revision process. The "scenarios" correspond to four-week time frames in which we evaluate the models, with a two-week assimilation period and a two-week forecasting period. We have three scenarios for the Blue River VIC model started in the following dates: 11/16/96, 4/7/97, and 2/24/98. Similarly, we have two scenarios for the Indiantown Run DHSVM model starting in 7/28/09 and 8/26/09.

The attached figure shows the comparison between one of the best-performing configurations of 4D-Var (using 2,000 iterations) and OPTIMISTS (1-week period, 500 members, 2 objectives, 40% samples, 95% roots, K-class kernels, and a greed value of 0.75). The numbers in the figure correspond to the three error metrics (NSEl2, NSEl1,

MARE) during the two forecast weeks. This comparison shows that, even though 4D-Var outperforms OPTIMISTS in Scenario 1, its forecasting accuracy is still inferior to the default model. This gives validity to the explanation in the manuscript, according to which the special circumstances of the scenario are unfavorable (no significant precipitation occurs during that period after a large storm) for DA in general. This result shows that it is not necessarily a sign of a weakness particular to OPTIMISTS. In addition, 4D-Var also outperforms OPTIMISTS in Scenario 2, but OPTIMISTS gets the upper hand on Scenario 3.

While the results with 4D-Var are in general slightly better than those of OPTIMISTS, our approach does have important redeeming qualities. First, it requires a considerable smaller number of evaluations to reach results that are comparable or even better than those of 4D-Var. Second, it provides a probabilistic estimate. And third, it leverages the inclusion of the Bayesian approach to mitigate the exponential growth of the solution space for cases of higher dimensionality. We will include this comparison in the manuscript to provide the readers with a more complete picture on what performance one can expect from OPTIMISTS in contrast to state of the art methods. Of course, as we mentioned, more systematic experiments should be carried out to better establish the benefits and weaknesses of the algorithm.

On another topic, we especially appreciate Referee #2 for pointing out the problems related to the broad classification of DA methods we used (i.e., sequential and variational). Even though we have seen this very distinction being used in the literature before, it is a very good point that these two categories are not sufficiently differentiated. However, despite the misleading labels, we believe that the distinction we were illustrating was between Bayesian (KF, EnKF, PF) and variational (1D-4D) methods (a fair one, given that the key difference is the actual method for exploring the solution space). We will both adopt these more accurate labels in the revised manuscript (including the title) and acknowledge that even these modified categories do not allow for a perfect separator for the wealth of methods that exist–they are more a practical framework from which to build a productive discussion. After all, one of our objectives with OPTIMISTS is to build bridges between these approaches.

We similarly appreciate the referees for identifying inaccurate uses of technical terms. We will fix these problems in the revision process and attempt a better integration of the terminologies typically used in the multiple fields that make up the intersection where our work stands, including Kalman filtering, particle filtering, variational DA, and evolutionary computation.

Response to other comments:

- We will gladly revise the introduction and the method sections to make them serve the audience of the journal better. In the original manuscript, we intentionally made the introduction and the methods sections relatively domain-agnostic, with the hope that our approach could be applied beyond our own domain of Hydrology. We understood that this came at the cost of denying hydrologists a better degree of familiarity throughout the explanation, but we had hoped it would open the possibility of furthering the dialogue between additional disciplines that use DA and it may be worth the cost. But as the reviewer pointed out, the sacrifice of the clarity and concise presentation needs to be considered. In the revision, we'll try to keep a good balance and make the clarity the first priority.

- We decided to include the full array of state variables that both models (VIC and DHSVM) possess for our tests, even those which would indeed just represent noise in the large scale of things (such as the interception). This is because we want to test the robustness of OPTIMISTS. A robust DA algorithm should be able to manage the inclusion of such type of variables (especially under the threat of equifinality), while omitting these variables would help simplifying the problem. Although it would make practical sense to focus on the most important state variables, their inclusion in our tests effectively represent an additional handicap on the method, increasing the significance of any positive results.

- We use an ensemble optimization algorithm because of its proved advantages from the evolutionary optimization literature (in essence, by trying to mitigate the problems associated with the no-free-lunch theorems). Yes, one could use a simpler algorithm (like the standard NSGA-2).

- While we did not use Eqs. 3, 4, and 5 in our tests, these would provide the most accuracy for models with low dimensionality. We have them included for the sake of completeness and for the applicability of our method to be easily extended to problems with a low dimensionality. In the revised manuscript we will make this point clear.

We will also make improvements in response to other comments, including:

- Dividing the explanation of the method into sub-sub-sections to allow for an easier understanding

- Adding reasons for design decisions of the algorithm and their contrasts with existing methods where they are currently missing

- Better explaining the difference between using random samples and using traditional resampling

- Further discussing the differences in the availability of meteorological information in our tests and in a hypothetical operational case

- Reformatting of Table 4 to improve readability (Values in brackets are not ranges but rather a set: these were all the levels assigned to the factors/parameters in the factorial experiment.)

- Adding a table with a summary of the performance metrics of the multiple configurations We would like to thank the referees again for their valuable comments and suggestions. With the approval of the editor, we will work on the aforementioned modifications to revise the manuscript and will be glad to implement additional changes to meet the journal's and the referees' standards.

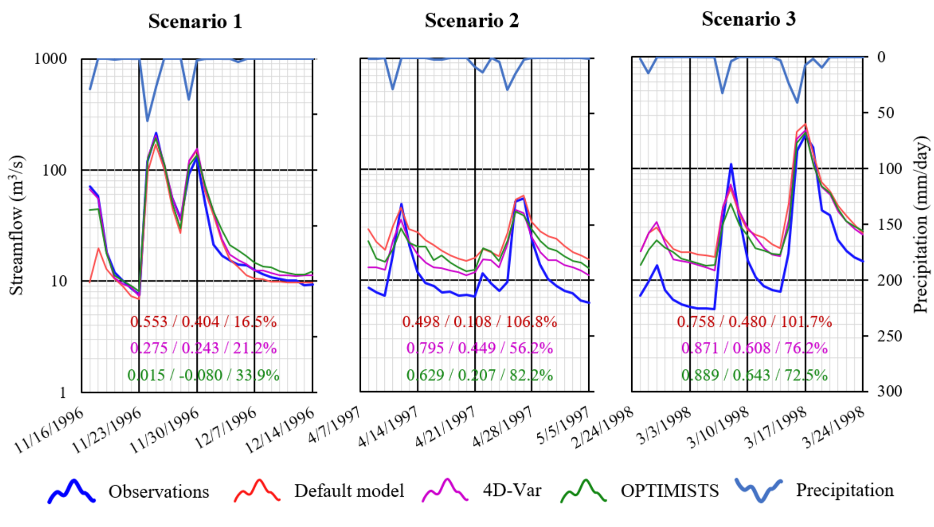

Fig. 1. Comparison between OPTIMISTS and 4D-Var on the Blue River model