Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2017-431-RC2, 2017

Hydrology and Earth System Sciences
Open Access
Discussions

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

Anonymous Referee #2

Received and published: 21 December 2017

Before beginning review of this manuscript, although not mentioned in the text and reference, this should be considered a re-submission of the previous HESSD manuscript entitled "Hybridizing sequential and variational data assimilation for robust high-resolution hydrologic forecasting (https://doi.org/10.5194/hess-2016-454)" by the same authors, which was rejected in 2016. I suggest the editorial board compare the final revision of the previous HESSD manuscript with the current one if the track record was not screened yet. I cannot examine whether the authors submitted the final revision in the previous submission in 2016 or not. However, if so, improvement and uniqueness of the current manuscript over the rejected final manuscript should be carefully evaluated. In addition, since HESSD is independent publication, the previous manuscript in

2016 should be cited and discussed in this manuscript.

This manuscript proposed a hybrid DA method, OPTIMISTS, combining sequential and variational methods, and compared performance of developed methodology over PF and VAR using distributed hydrologic models. The topic is of interest to a wide range of hydrologic modelling community. The strategy of the proposed methodology to leverage different DA approaches, sequential and variational DA, is one of the important trends in recent studies. However, there are major gaps in experimental setup and evaluation, and incomplete reasoning in new methodology which require significant changes before publication. I hope the followings would be helpful to improve the quality of manuscript.

1) Evaluation period and methods In this manuscript, the total evaluation period is 10 weeks (5 cases with a 2-weeks period each): 3 scenarios for 2-weeks forecasts in the Blue River and 2 scenarios for 2-weeks forecasts in the Indiantown Run, not including assimilation period.

The evaluation period for hydrologic modelling and data assimilation is usually longer than at least 6-8 months and up to multiple decades. The total 10-weeks forecasts (2-weeks piecewise each) and associated metrics cannot be accepted as a rigorous evaluation.

Given that the selected events in the Blue River in the 2016 manuscript are different from those in the current one, there seems to a potential to further increase evaluation period. In Table 3, the authors also mentioned calibration periods are 85 and 20 months, respectively.

Considering the availability of observation data, what is the maximum evaluation period for two catchments? Why don't you use the whole or most calibration period for DA evaluation? Was there any reason to use the limited period for evaluation?

For the larger domain, the Blue River catchment, is there just one streamflow observation gage over 3,000 square kilometer area? Why don't you assimilate observations in multiple locations to reduce equifinality and overfitting?

In this study, evaluation metrics were estimated for the whole 2-weeks forecast period. However, it is more common to evaluate metrics for varying forecast lead times because the impact of updating varies and disappears over time.

I highly suggest the evaluation period and method should be reconsidered to qualify a kind of general standard shown in many forecast and DA-related papers: simulating more than several months for each catchment and evaluating metrics for varying forecast lead times.

2) Probabilistic evaluation Although the proposed method is a stochastic approach, probabilistic metrics were not measured and analyzed. At least, basic metrics such as reliability, CRPS, predictive QQ plot and Brier score should be compared over the conventional method such as PF. Without such evaluation, improvements and features of the hybrid ensemble method cannot be understood in terms of stochastic perspectives.

In addition, Figs 5 and 8 (streamflow hydrographs) should include traces or spreads of ensemble for visual inspection.

3) uncertainty specification on hydrologic DA In order to apply DA for hydrologic modelling, uncertainties for states and observations should be carefully taken care of. Sometimes, not surprisingly, noise configuration or specification may significantly affect DA performance. However, there is no description on how uncertainties of different state variables and observations such as interception, snow, soil moisture and streamflows were formulated and implemented for hydrologic ensemble modelling, which should impact DA process to generate ensemble, optimize state variables and estimate likelihood or weight. A detailed description is required for reproducibility of this study.

Regarding this issue, for example, how different particles of distributed hydrologic mod-

els are generated in "the sampling step" of this DA algorithm? More specifically, how high-dimensional model states are being perturbed to avoid sample impoverishment in this step?

4) Under-simulation or filter degeneracy in assimilation step In the analysis or assimilation step which corresponds the first 2-weeks in Figs 5 and 8, under-simulation or filter degeneracy (scenario 3 in Fig. 5 and scenario 2 in Fig. 8) is found. Usually, whatever filter is used, traces of simulated states (here streamflow) overlap observations in the assimilation step since uncertainty of observation is set smaller than that of state variables. It is common that NSE values of the assimilation step or the first forecast step are higher than 0.9 – 0.95. However, a large gap between simulation and observation exists even in the assimilation step, which should be clearly diagnosed and discussed.

5) Comparison of posteriors of state variables What potential readers want to see in the result section may be not only comparison of NSE at the outlet location. The authors need to address why and how their DA method can improve over the conventional ones in hydrologic forecasting from perspectives of distributed modelling. A comparison of posterior distributions of state variables updated by the new and conventional methods may be useful to show how and why the new DA works for high dimensional applications.

Especially, given that the authors urged OPTIMISTS employed essential features from but outperformed particle filters, a comparison of posteriors between two methods is also required to demonstrate whether non-Gaussian and multi-modal distributions are preserved or not.

6) Evaluation and optimization steps for hydrologic modelling It is not clear how the cost function is formulated for distributed hydrologic models. The authors need to show explicitly how multiple spatially-distributed state variables and associated uncertainties are taken into account to formulate the cost function in evaluation and optimization steps.

7) Tuning hyper-parameters There are numerous hyper-parameters such as time step, objectives, no. particles, optimization, Wroot and Kf-class, Psamp and g, related to this DA method which may increase uncertainty and subjectivity of forecasting. However, analysis methods and results on hyper-parameters shown in Figs. 3, 6 and7 are still confusing and do not provide well-organized understandings. A summary or guideline is required for proper range or values of hyper-parameters.

8) Terms and units In Table 3, use of two different units for one variable (m3/s and l/s) is not recommended.

Throughout the manuscript, the term 'time step' is used to represent 'assimilation time step' or 'assimilation window'. Since the time step usually stands for a temporal increment for numerical schemes, 'assimilation window' or 'analysis window' may be more appropriate to avoid possible confusion.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2017-431, 2017.