Hydrol. Earth Syst. Sci. Discuss., 8, C4317–C4321, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C4317/2011/

Hydrology and Earth System Sciences Discussions

Interactive comment on "An integrated uncertainty and ensemble-based data assimilation approach for improved operational streamflow predictions" *by* M. He et al.

N. Voisin (Referee)

nathalie.voisin@pnl.gov

Received and published: 5 October 2011

The paper presents a multi step approach to generate ensemble streamflow forecasts where multiple sources of uncertainties are taken into account and merged into one ensemble: forcing errors, optimized model parameters and errors (ISURF), and initial conditions errors via SWE assimilation, all merged with an EnKF. In particular, each uncertainty type (mostly model parameters and initial conditions here) is individually characterized and merged into one "big ensemble" via an EnKF. The approach is evaluated with respect to an EnKF without the individual characterization of the uncertain-

ties (blend of undifferentiated sources of errors) and standard operational performance (removing the human forecaster's adjustment skill). The approach is validated over the American River Basin, CA.

The paper is very well written with a great and extensive literature review and justification of choices. The approach is a great contribution to the scientific community effort; characterize individual uncertainties of different sources and merge them in a single but meaningful ensemble in an effort to assess modeling uncertainties in a more accurate way. The analysis currently assesses if the simulations are more accurate which allows a comparison with the deterministic NWS forecasts. But it does not assess if the uncertainty is better assessed. I would recommend accepting the paper with major revisions. The paper would benefit from a couple of additional explanation of the performance and limitations of the approach, and most important a schematic diagram for clarity. Also, it is presently difficult to isolate the performance of the data assimilation with the overall water-year long approach performance. The results analysis would benefit from using another measure than the dispersion only – it needs to be the dispersion with respect to the observation (is the ensemble representative?) in order to assess if the uncertainty is better assessed than a blended uncertainty (ICEA vs EnKF analysis only). See specific comments below.

Specific comments

1/ add a schematic diagram that explains the chain of models and processes, and the variables being transferred (single value, or ensemble), the time scale of the analysis etc. For example: Observed precipitation, temperature, PET -> SNOW17; i)ISURF – optimized parameters and uncertainty), ii) EnKF for SWE assimilation and merging parameters uncertainties ->precip and snowmelt ensemble, PET (single value?)->SAC-MA -> ensemble streamflow forecast to be verified with respect to observed flow, for several days after the assimilation. The assimilation is performed every 7 days. Etc.

2/The forcing uncertainty is accessed via a proxy by SNOW17 parameters. It seems to me that if SNOW17 was driven by an ensemble weather forecast, the SCF is adjusting the forcing in a way that the uncertainty information is decreased, unless the ensemble weather forecast was calibrated with respect to a specific observed meteorological dataset knowing the SCF to be used. As such, I am not sure how much of the forcing uncertainty is really taken into account. Similarly, it means that in a full probabilistic approach where ensemble weather forecasts were to be used, the approach does not allows yet to merge forcing uncertainty with model parameterization and initial conditions uncertainties. This being said, this approach allows generating an ensemble merging initial conditions and parameters uncertainties with an estimate of forcing uncertainty. As far as I know, this is a first on how forcing uncertainty can be merged in an ensemble with parameters uncertainties and initial conditions errors.

3/ metrics used for the evaluation: the annual bias, the correlation, the RMSE, NRR and UR95.When presenting the results about the dispersion, I would suggest making sure that not only the dispersion is being discussed as what matters is if the dispersion represents the observed variability. The authors mention in the discussion that other metrics could be used to assess the reliability. I would suggest using some of them here, like rank histogram for example. They would help assessing if the dispersion is fast enough for short lead times, and if the information in the ensemble is right, because too large of an ensemble has no value, too narrow either. It needs to bracket the observation in a representative way (uniform histogram). I believe that this is another important component of your approach – you need to evaluate if the uncertainty is better assessed when individually characterizing the uncertainties, or if blending them drive to the same result.

4/ The results are presented for the entire WY. It is difficult to isolate the performance due to the SWE assimilation on top of the approach and the one due to the approach when the SWE assimilation is not in used. What is the performance of the system for the snow period only - overall. What is the performance of the snow-free period (hear glacial melt instead of perennial snow melt if applicable)? It is common to look at

the 95th percentile for looking at figure 6, it is not obvious if low flow/average flow has improved. Please comment as to here add again about the overall performance of the approach and its best application / limitations.

5/ the ISURF approach allows defining the optimized parameter sets, with their uncertainty structure. This is still more or less equivalent to a calibration prior to the data assimilation approach, which can affect the water balance. How does it theoretically affect the parameter uncertainty structure? And operationally, it might be okay (is it?) to not meet the annual water balance as long as we are getting the next few days peak flows right. It would contribute to the description of the approach (performances and limits of applicability discussion) to comment on it.

6/ Why did you choose one week for the frequency of the analysis? Others have use 3 days for example (Clark and Slater 2006), agreeing that the prior distribution does not necessarily change that fast. I would think though that during snowmelt period, when snow depletes faster, the frequency of the assimilation should not be any longer than the time of concentration of the basin in order to avoid any incoherency in the flow and in the ensemble flow forecast characteristics, i.e. about 3-4 days over the American? For example, it would help looking at the ensemble/dispersion over a continuous period of time; i.e. day 7 might display a large dispersion and then assimilation is performed and the ensemble dispersion narrows again? Please explain or show a time series of the ensemble dispersion over a period of time.

Technical revisions:

Why 6 years of training and 6 years of validation when 23 years are available?

P7716, line9: what is the model providing the meteorological forcing to SNOW17/SAC-MA?

p7716 line 11: is the observed daily flow regulated?

How do the uncertainty in the forcing compares to actual short or medium range ensemble forecasts?

P7721, line17: How are the station weights computed? Or give a reference.

P7732, line1: I would suggest substituting “current” with “automatic” as the current prediction depends heavily on the human forecasters who make a difference, as seen on the statistics used for this analysis and those seen on the RFC website.

P7743, table 1; It seems there a typo for year 1994 peak flow, should be higher.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 7709, 2011.