Hydrol. Earth Syst. Sci. Discuss., 8, C4212–C4215, 2011
www.hydrol-earth-syst-sci-discuss.net/8/C4212/2011/

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.

Anonymous Referee #2

Received and published: 26 September 2011

This study presents a procedure for combining Bayesian parameter estimation with data assimilation to improve operational streamflow forecasts. The Integrated unCertainty and Ensemble-based data Assimilation (ICEA) method combines the previously presented ISURF (combination of GSA and DREAM to estimate parameters) and data assimilation with the EnKF. The authors had developed this method to be applicable to operational forecasting to increase the skill of predictions. This is applied to the SAC-SMA/SNOW-17 models to compare performance with operational forecasts. The method utilizes SNOTEL and CDEC observations of SWE to improve streamflow estimates. Overall the paper shows improvement over current operational forecasts but several issues need to be addressed (clarified and corrected) before the paper is ready for publication. Therefore, I recommend major revision of the manuscript given the comments below:

Comments:

1.) Page 7719, Line 17 states that assimilation is only performed once a week. Why is the assimilation only performed once a week when the SNOTEL observations are typically daily?

2.) Page 7720, line 3 states that the second scenario used "ISURF-derived optimal model parameters" but ISURF estimates parameter distributions not the optimal parameter set. How is optimal defined here (e.g. mode, mean)?

3.) Equation 5 appears to have an error. If the difference between the SNOTEL and modeled SWE is minimized, this will not produce an areal estimate of the SWE, as suggested in this equation. It seems from this method that the sum of the SNOTEL SWE multiplied by the weights is the areal SWE estimate but this needs to be clarified in the manuscript.

4.) Section 3.3 describes the method for estimating the areal SWE for the upper elevation band. This method combines SNOW17 model estimates (with prior RFC parameters?) and SNOTEL observations to estimate the spatially averaged SWE. The Ck values are then estimated based on the model estimates and the in-situ observations. This makes the SWE values highly dependent on the SNOW-17 model. The model dependent SWE observations are then used for calibration and assimilated into the model. In my opinion this is very problematic because the model is calibrated, in part, to the prior model runs and not solely on observations. In addition, it cannot be ensured that the estimate of SWE is representative of the spatial average and thus an accurate calibration and assimilation will not necessarily lead to more accurate streamflow forecast. A further clarification of the technique with justification or a method for estimating areal SWE independently of the model is necessary.

5.) In paragraph 2 of section 3.4, the method of using the SCF and PXTEMP parameters to account for forcing error is described. This technique adds noise to these two parameters to account for errors in forcing data. From my understanding in He et al. 2011a, this is not actually noise added to the parameters but an estimated posterior distribution from DREAM. If this is the case here, the forcing uncertainty will likely be underestimated. Since the DREAM algorithm works in a batch framework, the uncertainty in these parameters (and thus the forcing error) will be estimated over the entire length of the batch instead of daily, as is more commonly performed. Estimating the SCF across the entire batch length will find the uncertainty relating to the average error of the precipitation measurements as opposed to the daily measurement errors. It is likely that the daily precipitation measurement error has a much greater variance than its long term average uncertainty, which will likely lead to degraded prior distributions. In my opinion, the added ease of application through assuming the SCF and PXTEMP parameters handle forcing error is not worth risking the potential problems associated with this approach. Further, I would suggest examining this method as compared to the traditional methods of adding forcing error to ensure that this method is not significantly altering the results.

6.) Page 7724 line 7 states that the UR95 has a perfect score of 0%. This statement is not entirely correct because an uncertainty ratio of 0% indicates no uncertainty is estimated. In any practical scenario, there would be some uncertainty, due to forcing, model, parameter and observation error, and therefore an UR95 of 0 % will be an overconfident prediction. 7.) Page 7729 line 24 states "from day 230 to day 252, the EnKF ensemble is much wider than ICEA ensemble" but it is not mentioned that during this time the EnKF ensemble encompasses the observation while the ICEA ensemble does not. This means that the EnKF actually performed better than ICEA during this period. This is followed by the statement "from day 265 to day 289, the ICEA ensemble reasonably captures the recession pattern while the EnKF ensemble follows the variation of RFC predictions which deviate from the observed streamflow". Though the ICEA is closer to the observation during this period, it appears that the observation is outside the ensemble of the ICEA for the majority of this period. For this reason, I disagree that the "ICEA ensemble reasonably captures the recession pattern".

8.) Page 7730 lines 14-17 explains that after day 265 the EnKF SWE melts more rapidly than the ICEA SWE leading to poorer performance from the EnKF than the ICEA in terms of streamflow. However, the EnKF appears to match the observed SWE more closely than the ICEA throughout the melt season. Given that the ICEA performs worse in matching the observed SWE but better in matching the streamflow, it is likely that the SWE used for assimilation is not representative of the true basin SWE (this relates back to problems suggested in comment 3). This would explain the poor performance of the EnKF in terms of streamflow despite a relatively accurate assimilation of SWE. Once again I find it necessary to show that the method for spatially averaged SWE generation is representative of the true basin SWE.

9.) Page 7731 lines 7-10 states "the whole EnKF predicted streamflow ensemble is wider than the ICEA ensemble at several lead times (day 2, day 6, and day 7), while the ensemble is narrower at other lead times (figure 9f)". Figure 9f shows the NRR which is a measure of the accuracy of the ensemble spread, not a direct measure of the width of the ensemble. The UR95 is a more accurate measure for comparing which has a wider ensemble spread. This statement would also be more accurate if phrased "the whole EnKF predicted streamflow ensemble is less overconfident than the ICEA ensemble at several lead times (day 2, day 6, and day 7), while the ensemble is more overconfident at other lead times (figure 9f)".

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