Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-296-RC2, 2016

Hydrology and Earth System Sciences Discussions
Open Access

Interactive comment on "EnKF with closed-eye period – towards a consistent aggregation of information in soil hydrology" *by* H. H. Bauser et al.

Anonymous Referee #2

Received and published: 15 August 2016

In the paper by Bauser et al., the authors perform data assimilation experiments for a 1D real-world case (Grenzhaus site, Germany) for a time period of two months. An iterative EnKF is used to correct states (soil moisture), parameters (layer-based hydraulic conductivity, Miller scaling factors, porosity, van Genuchten parameters) and forcings (upper boundary condition) with measured soil moisture profiles. The authors introduce a so called 'closed-eye' EnKF scheme, in which parameters are not estimated during an intensive rainfall event. This is done in order to prevent the incorporation of model structural errors into the estimated parameters.

General comments:

1) The introduction would benefit from a more detailed description of previous applica-

Printer-friendly version

Discussion paper

tions of the EnKF in soil hydrology. The authors should highlight the novelties of their approach compared to previous studies. Additionally, one of the aspects of this paper is a proper description of uncertainties for the utilized test case, so relevant literature on uncertainty description in hydrologic data assimilation should also be cited (e.g. Jafarpour & Tarrahi (2011), Zhang et al. (2015)).

2) More details on the iterative EnKF are needed because this method is not commonly used in hydrologic data assimilation: Are the initial conditions the same for each EnKF iteration? What is the criterion for choosing 50 or 10 EnKF iterations? How do the updated parameters evolve during the different EnKF iterations? In order to judge the EnKF performance better, it would be beneficial if the authors could also provide simulation results with the initial guess of parameters (no data assimilation) as a benchmark.

3) The assimilation time period is rather short and already split into different parts for the estimation of initial conditions/ Miller scaling factors (time periods A/B) and the estimation of the full state-parameter vector (time period C). In between time periods A/B and C there is an inconsistency in terms of ensemble size (40/40 versus 100), the number of EnKF iterations (50/1 versus 10) and an additional perturbation of parameters. I suggest the authors to be more consistent in the data assimilation set-up by using the same meta-parameters (ensemble size/ EnKF iterations) for all time periods. Especially for the change of the ensemble size it is unclear how the states and parameters from the 40 ensemble members are re-sampled to 100 ensemble members (see specific comments below). This data assimilation scheme also does not allow for an equilibration of model states towards the perturbed/ updated parameters. Therefore, I would suggest to add a certain spin-up phase for the ensemble at the beginning of the assimilation phase and between the EnKF iterations. In my opinion, it would also strengthen the paper, if the comparison between standard and 'closed-eye' EnKF could be made for more than one single rainfall event in order to provide more evidence for the effect of model structural errors during rainfall events. This could be achieved,

e.g. by applying parameter estimation with EnKF during both rain events (time period B+C) or by extending the simulation period for additional rain events. It would also be advantageous to use an independent verification period (including one or more rain events, not only the recession phase in time period D) to compare the performance of the updated parameter values from the two assimilation schemes.

4) It is quite unclear from the manuscript how the upper boundary condition is updated. As a filtering method, the EnKF only adapts the state-parameter vector for the current model time step. How is the updated upper boundary condition then incorporated in the subsequent model integration? Additionally, the results for the update of the upper boundary condition are not really discussed in detail. So it is quite unclear, how this update affects the model simulations. The state vector includes almost all components of the water balance for the 1D soil profile. These components are all adapted with the same measurements and it is unclear whether the soil moisture data contain enough information to update all water balance components at once. As the main focus of the paper is on parameter estimation, I suggest the authors to leave out the estimation of the upper boundary condition and only use perturbed forcing data instead.

5) A discussion is missing on how a 'closed-eye' EnKF period can be defined in other practical applications. In the paper, a quite heuristic criterion (maximal parameter change within a defined time period) is used which relies on a prior application of the standard EnKF assimilation scheme. Can such a criterion be generalized? The EnKF is usually applied as a pure sequential method without iterations (often in the context of real-time simulations), so the question is whether the 'closed-eye' period can also be defined properly for such cases, or if the iterative EnKF is a prerequisite for the 'closed-eye' EnKF. An additional question is how such a 'closed-eye' period would be defined for a 3D model where precipitation is spatially distributed and the computational demand is much larger than for a 1D soil column.

6) To my knowledge, the covariance inflation method proposed in Anderson (2009) has not been applied in hydrologic data assimilation so far. It would be interesting to see,

how this inflation scheme affects model results, e.g. by providing some information on the calculated inflation factors or by comparison with a standard EnKF assimilation scheme.

Specific comments:

page 1 line 6: I think the definition of the 'closed-eye EnKF' should already been given here or at least in the introduction.

page 2 line 15-16: This not fully representative for the study. You should also add your experiments with the standard/ 'closed-eye' EnKF here.

page 3 line 22-25: Please clarify what you define as 'non-equilibrium conditions' in this context. Do you mean preferential flow?

page 4 line 20-27: The definition of Miller scaling requires more details. How exactly are equations 5+6 incorporated into your model simulations? I guess you define a scaling factor for each model grid cell and use the parameters of the respective soil layer as the "reference material property". Additionally, why does a scaling factor of 1 introduce large uncertainties? This is not obvious from equations 5+6.

page 5 line 5-9: The TDR measurements for 111 and 136 cm in Figure 4b suggest that there is some dynamics in the lower part of the soil profile. On the contrary, the simulated soil moisture contents are almost constant. This is an indication that there might be some problems with the assigned lower boundary condition. The constant simulated soil moisture values are probably an effect of the sand layer that was introduced as layer 5 which acts as a hydraulic barrier to the upper soil profile. You could try to use the soil hydraulic parameters from layer 4 also for layer 5 to see, if the simulated soil moisture dynamics can be improved in the lower part of the soil profile. Additionally, it is not clear why you have chosen a water table depth of 4m. Is this value based on water level data from surrounding wells?

page 7 line 25: If the soil moisture measurements exceed the calibrated saturated water, why don't you use the data from the volumetric soil samples (page 5 line 24) as an initial guess?

page 8 line 9-11: Does this mean that you only use daily averages of precipitation? In the data assimilation experiments rather short time periods are investigated and the simulations are compared to hourly soil moisture measurements. So it would make much more sense to also use atmospheric forcing data with a similar time resolution which should be available from the weather station (page 3 line 3-4).

page 8 line 30-31: Which value for the cut-off radius did you use in the 5th order polynomial for soil water content?

page 9 line 2-4: Why is the localization different for hydraulic conductivity and Miller scaling factors? Miller scaling factors are strongly related to hydraulic conductivities and should therefore also have a similar localization.

page 9 line 9-11: Did you also use localization in the derivation of the inflation factors? This is necessary because otherwise uncertainties are increased in model areas that are not updated by the EnKF.

page 10 line 20-23: How are initial Miller scaling factors perturbed?

page 11 line 7-11: It is unclear how the initial conditions for time period C were exactly created. How were the final (updated) states from time period B re-sampled from 40 to 100 ensemble members? Did you perturb the ensemble mean of Miller scaling factors from time period A or each of the 40 ensemble members individually? Was this perturbation constant in space or was the perturbation on the grid cell/ soil horizon level? Please also take into account that this perturbation creates inconsistencies between states and parameters which could be alleviated by a spin-up period.

page 11 line 32 - page 12 line 4: It is not obvious from Figure 5 whether hydraulic conductivity reaches its initial value by the end of time period C. Please be more quantitative here: How large is the change within the 'closed-eye' period and how large is the difference between initial and final values? It would also be good to show how the update of hydraulic conductivity evolves over the different iterations of the EnKF. Is there a continuous increase of this parameter during the iterations? Additionally, the increase of hydraulic conductivity during the closed-eye period could also be related to the adaptive covariance inflation. The calculation of the inflation factors is influenced by the mismatch between simulations and measurements. As this mismatch increases during the infiltration event, this could also lead to higher inflation factors which increase the Kalman gain and thus the parameter update for this period. Did you check how the inflation factors evolve over this time period? Maybe it's also worthwhile to repeat these simulations without covariance inflation in order to exclude possible artefacts from the inflation scheme.

page 12 line 14-17: The 'closed-eye' EnKF experiments are added on top of the standard EnKF experiments with an additional parameter perturbation. This makes it quite difficult to compare these two experiments. I suggest to use exactly the same model set-up for both experiments, i.e. to repeat the 'closed-eye' EnKF experiment with the same initial conditions as the experiment with standard EnKF. Otherwise, a direct comparison of both methods is not possible. An additional question is why the 'closed-eye' EnKF and the parameter perturbation is only performed for the first soil layer.

page 12 line 23-24: What are the 'believed true material properties'? Given the fact that this is a real-world experiments, where the true material properties are unknown, this is quite speculative.

page 13 line 13-22: Please be more quantitative here by providing performance measures such as root mean square error or Nash-Sutcliffe efficiency.

page 13 line 26-30: In Figure 10b the simulated ensemble mean for each layer is always above the measured soil moisture values at the end of time period C. Why does this offset not appear in the beginning of time period D in Figure 10d? Shouldn't period D start with the simulation results from period C?

page 14 line 28-31: You should also discuss how such a 'closed-eye' period can be defined in practical applications.

Figure 2: Why is there no evaporation flux in time period A and at the beginning of time period B? If is is a dry period, there should also be evaporation.

Technical corrections:

page 13 line 3: Change 'andit' to 'and it'.

Table 2: Saturated soil moisture content for layers 3+4 is above 1. Please correct.

References:

D. Zhang, H. Madsen, M. E. Ridler, J. C. Refsgaard, K. H. Jensen (2015), Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model, Advances in Water Resources, 86, 400-413, doi:10.1016/j.advwatres.2015.07.018.

B. Jafarpour, M. Tarrahi (2011), Assessing the performance of the ensemble Kalman filter for subsurface flow data integration under variogram uncertainty, Water Resources Research, 47, W05537, doi:10.1029/2010WR009090.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-296, 2016.