Response to reviewer 1

Responses to the comments are shown in black and proposed changes to the manuscript are shown in blue.

The authors are performing data assimilation (DA), for a dynamical hydrological model with time-varying parameters. I fully agree that, in hydrology, time-varying parameters very often lead to a more realistic description of reality than constant ones, with the caveat that a good stochastic model for their dynamics is used. That being said, I have serious doubts about the validity of the method used in this paper.

We thank the reviewer for their time. Please see below our responses to the comments.

First of all, there is no clear separation between the hydrological model assumptions and the numerical method that is used for DA. The model assumptions should not only comprise the deterministic hydrological equations and the observational error assumptions, but here also a precise definition of the assumed stochastic dynamics of the model parameters. These assumptions have nothing to do with DA, but are part of our prior knowledge about the system. Together with prior probability distributions, e.g. for initial conditions, and measured data they completely define the posterior as well as predictive distributions of states and parameters.

We respectfully note that the assumed stochastic dynamics of the model parameters were detailed in Step 1 of the algorithm in Section 3.2. The assumed parameter dynamics are a Gaussian random walk with time varying mean and variance. This is a reasonable assumption for cases where no prior knowledge of the parameter dynamics (or changes to catchment properties) is available. The prior mean is assumed to evolve locally linearly in time, such that it is estimated by the drift rate from the previous two time steps (as indicated in Line 204). This is shown in Figure 1. The prior variance is assumed to be also time dependent, and based on the variance at the previous time (as indicated in Lines 205-207). The statistics of the prior parameter ensemble are also updated sequentially by the DA algorithm based on assimilated observations.

We note that these details are not the main focus of the paper, as they have already been investigated and discussed in detail in Pathiraja et al. (2016). We have provided references to these publications in the manuscript (see Lines 184, 191-192). However, we will provide additional discussion in the revised manuscript, as detailed below.

We propose to add the following to Step 1 of Section 3.2 so that the assumed stochastic dynamics are clearer: "The parameters are assumed to evolve through a Gaussian random walk with time varying mean and variance. The prior mean is assumed to evolve locally linearly in time, such that it is estimated by the drift rate from the previous two time steps. The prior variance is based on the variance of the update from the previous time." This discussion would be inserted in Line 202 before "The prior (or background)…"

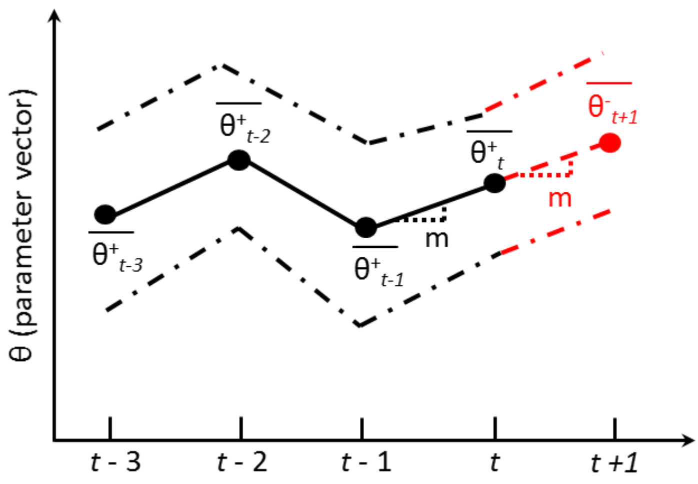

Figure 1. Schematic showing how the mean of the prior parameter ensemble $\overline{\theta}^{-}_{t+1}$ is proposed at a given time *t+1*. [*taken from* Pathiraja et al. (2016)]

All DA methods must lead to the same distributions and the choice is only a matter of computational efficiency. If this separation is not clearly made, there is a risk that inference (here DA) and prediction is done under different model assumptions, which would be inconsistent and lead to a loss of interpretability of the results. I've been trying to do this separation and unveil the assumptions behind the dynamics of the parameters from step 1 of the numerical DA algorithm, but it is not obvious to me what these assumptions are. Without constraining data (i.e. in a predictive mode), what would be the dynamics of the parameter distribution?

DA methods in general do not lead to the same posterior distributions, particularly when applied to complex non-linear/non-Gaussian problems. This is even when one is specifying the prior, initial conditions and observation error in exactly the same way. One only has to look at the plethora of existing DA algorithms to appreciate this (Ensemble Kalman Filter, Ensemble Transform Kalman Filter, Ensemble Adjustment Kalman Filter, Particle Filter-SIR, Particle Filter-MCMC, Auxiliary Particle Filter, just to name a few). They do not differ solely in terms of computational efficiency (e.g. Arulampalam et al., 2002; Liu et al., 2007).

The assumed parameter stochastic dynamics were discussed in the response to the previous comment; these apply in predictive mode also. In this case, the proposed local linear extrapolation is only valid close to the current time point. This is acceptable given that the Locally Linear Dual EnKF is designed for short term predictive modelling. This will be emphasised in the revised manuscript.

We propose to insert the following sentence at Line 64: "It can be used to either retrospectively estimate time variations in model parameters or for short-term predictive modelling."

Putting aside my concerns about these model assumptions, I have even more serious concerns about the chosen DA method, which seems to violate the assumptions behind Kalman filters in at least two ways:

(i) Kalman filters assume normality of the distribution of the augmented state (incl. parameters). Is there any reason to believe that the non-linearity of the chosen hydrological model is weak enough for this assumption to be approximately valid? Given the dimensionality of the model and the data, respectively, I'm almost certain that it will be grossly violated.

We note that the method relies on an Ensemble Kalman Filter, not a standard Kalman Filter. Ensemble Kalman Filters were specifically designed to handle the non-linear/non-Gaussian case, although they are sub-optimal in that only the mean and covariance are considered in the update. Nevertheless, they are a practical alternative to Sequential Monte Carlo methods (which approximate the full Bayesian posterior), as these methods suffer from several practical implementation issues. We also note that Ensemble Kalman Filters have been used with considerable success in a wide range of complex non-linear/non-Gaussian problems (see for Reichle et al., 2002; Gu et al., 2005; Komma et al., 2008; Sun et al., 2009; Xu et al., 2016).

(ii) Updating the states, based on prior predictions that have been made with parameters that have already been updated seems to use the data twice. This again seems to violate model assumptions, or in other words, I have no idea what the model assumptions are, for which the proposed method is a valid DA.

In regards to the second issue, in Pathiraja et al. (2016), we demonstrated that the dual update filter accounts for correlations between observation and process noise associated with the use of a dual update (see Step 6 in Section 3.2). Full details of the dual update procedure were provided in Pathiraja et al. (2016). The purpose of this manuscript is not to revisit details of this method, references to the relevant publications were provided in Lines 184 and 191-192 of this manuscript.

References

Arulampalam, M. S., S. Maskell, N. Gordon, and T. Clapp (2002), A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, *IEEE Trans. Signal Process.*, *50*(2), 174–188, doi:10.1109/78.978374.

Gu, Y., and D. S. Oliver (2005), History matching of the PUNQ-S3 reservoir model using the ensemble Kalman filter, *SPE J.*, *10*(2), 217–224, doi:10.2118/89942-PA.

Komma, J., G. Blöschl, and C. Reszler (2008), Soil moisture updating by Ensemble Kalman Filtering in real-time flood forecasting, *J. Hydrol.*, *357*(3–4), 228–242, doi:10.1016/j.jhydrol.2008.05.020.

Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*(7), doi:10.1029/2006WR005756.

Pathiraja, S., L. Marshall, A. Sharma, and H. Moradkhani (2016), Hydrologic modeling in dynamic catchments: A data assimilation approach, *Water Resour. Res.*, *52*, 3350–3372, doi:10.1002/2015WR017192.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi (2002), Hydrologic Data Assimilation with the Ensemble Kalman Filter, *Am. Meteorol. Soc. - Mon. Weather Rev.*, *130*(1), 103–114, doi:10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2.

Sun, A. Y., A. Morris, and S. Mohanty (2009), Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data, *Adv. Water Resour.*, *32*(2), 280–292, doi:10.1016/j.advwatres.2008.11.006.

Xu, T., and J. . Gomez-Hernandez (2016), Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble kalman filtering, *Water Resour. Res.*, 600–612, doi:10.1002/2015WR018249.