Response to Anonymous Referee #1

We summarized the response to Referee #1 to two main points: 1) the Ensemble Kalman filter applied to forward modelling and 2) the assumptions behind the choice of boundary conditions.

1- We would like to start by clearly stating that way of discussing is not the intention of this discussion forum. Point 3 in the "General obligations for referees" for HESS states "A referee of a manuscript should judge objectively the quality of the manuscript and respect the intellectual independence of the authors. In no case is personal criticism appropriate." Remarks such as “This also suggests that the authors may not have understood the EnKF" in the first review is clearly personal criticism. Had no personal remarks been made, then the statement in the second review “Don't take criticism of your work as an insult" would not have been necessary. Personal comments get taken personally. Furthermore, the code of conduct of EGU states “Reviewers are expected to provide a brief, clearly written, constructive and unbiased feedback in a timely manner." This is clearly not the case here. Had the reviewer provided a structured, objectively and concisely written list of dot-points instead of a rant, as good reviews do, the discussion would have been much more effective. Responding to such an excessively long review takes far too much time.

We will briefly recall here how nonlinear operation systems work. We will not specify the meaning of each variable here, assuming the reviewer knows what they are, and report the key equations required to explain what we did. We start from the state update equation

$$x_k^{i,a} = x_k^{i,f} + K_k \left[y_k - y_{k(i,f)} + v_k^i\right], \qquad (1)$$

where the gain is calculated as:

$$\mathbf{K}_k = \frac{\mathbf{PH^T}}{\mathbf{HPH^T} + \mathbf{R}_k}. \qquad (2)$$

The two matrices products in the Kalman gain are:

$$\begin{cases} \mathbf{PH}^T = \frac{1}{M-1} X_k^f {Y_k^f}^T \\ \mathbf{HPH}^T = \frac{1}{M-1} Y_k^f {Y_k^f}^T \end{cases} \qquad (3)$$

The state and observation-simulation deviation matrices are written as:

$$\begin{cases} \mathbf{X}_k^f = \left[x_k^{1,f} - \bar{x}_k^f \quad x_k^{2,f} - \bar{x}_k^f \quad \dots \quad x_k^{M,f} - \bar{x}_k^f \;\right] \\ \mathbf{Y}_k^f = \left[y_k^{1,f} - \bar{y}_k^f \quad y_k^{2,f} - \bar{y}_k^f \quad \dots \quad y_k^{M,f} - \bar{y}_k^f \;\right] \end{cases} \qquad (4)$$

Pauwels and DeLannoy (2009) provide an in-depth analysis of what this means, which we will summarize here. Assume one state variable (x, for example, catchment wetness) and one observation (y, streamflow). Entering Equations 3 and 4 into the expression for the Kalman gain leads to:

$$\mathbf{K}_k = \frac{\sigma_{xy}}{\sigma_y^2 + v_k^i}. \qquad (5)$$

Assume the observation error is zero. The Kalman gain then becomes:

$$\mathbf{K}_k = \frac{\sigma_{xy}}{\sigma_y^2}. \qquad (6)$$

This is very simply a linear regression, across all ensemble members, between the observation and the state. Thus, if the model predicts different streamflow than is observed, the state update becomes:

$$x_k^{i,a} = x_k^{i,f} + \frac{\sigma_{xy}}{\sigma_y^2} \left[y_k - y_{k(i,f)} \right]. \qquad (7)$$

In other words, the gain maps the difference in observation space to state space. If the model underestimates the streamflow, Equation 7 will increase the modeled catchment wetness (if the covariance between catchment wetness and streamflow is positive, which it usually is). If the model overestimates the streamflow, Equation 7 will reduce the modeled wetness. So, yes, streamflow is a proxy for the wetness, because you update the catchment wetness without observing it. If the observation error is nonzero, the update will be reduced. For multiple observations and state variables, similar reasoning can be made.

This is the way assimilation of brightness temperatures or backscatter values or streamflow into hydrologic models works. The observation system

$$y_k = h\,(x_k) + v_k \quad (8)$$

in this case is a radiative transfer model (for brightness temperature assimilation), a backscatter model (for backscatter data assimilation), or the hydrologic model (for streamflow assimilation, as just explained). Thus, the model does not have to directly predict the variable it assimilates. They can be calculated through the observation system (Equation 8).

What one CANNOT do, as is unfortunately done frequently in streamflow assimilation papers (and also suggested by the reviewer), is to enter these values in the state vector, simply because they are not state variables. One of the two prerequisites of a correct system description is that the system must be controllable. This means that an external input needs to be able to move the internal state of a system from any initial state to any other final state in a finite time interval. If discharge is in the state vector, at the beginning of the time step one can assign any value to it; regardless of the forcing, it will have no impact on the soil moisture content and streamflow at the end of the time step. We know no hydrologic models for which streamflow is an initial condition. Another prerequisite is that the system must be observable, meaning that all state variables must be able to be inferred from the observations. The two prerequisites are very closely linked to each other.

Because flood forecasting models (especially the ones that are used for streamflow assimilation) usually have a very limited amount of state variables, when streamflow is entered into the state vector, problems with these two prerequisites usually do not occur, even though they are not met. But in a more complicated system such as ours, working with a system that is not correctly described can lead to significant problems, most likely excessive state updates.

In the case of assimilating backscatter values, which the reviewer correctly points out no hydrologic models directly compute, how would this approach work, using an observation matrix with ones and zeros? How can a model enter a variable it doesn't compute in the state vector? Same question for assimilating brightness temperatures. The reason for the confusion when assimilating discharge is that, in this case, the model does directly compute the variable that is assimilated. And the poor system description does not lead to problems because of the limited amount of state variables.

Reading the second review, we believe that the reviewer assumes that we are updating the parameters (which we do not) or that we disturb them at every time step (which we also do not). This is not clear. The pdf we infer is simply $\mathrm{P}\,(x_k|x_{k-1}, y_k)$. Disturbing initial parameter values is common practice in data assimilation with the EnKF (most of the papers listed below and most others). There is no parameter space in the Kalman filter equations, only state and observation space. This makes the last page in the second review very confusing.

2- In order to justify the assumptions behind the choice of the fixed boundary conditions (BC), we compared the water table (WT) dynamics of the bore used by the Referee to another, more appropriate, bore. To minimize the BC influence on the simple domain conceptualization, we have purposely chosen a location and an aquifer where the time variability of BC was low. Figure 1 compares the WT dynamics (in red) of a bore located in the center of a forestry block, more than two kilometers from any groundwater extraction, and the bore selected by the Referee (in blue), which is monitoring the levels of an incorrect aquifer. Constrained by the availability of the observations, the comparative analysis of the WT dynamics is only possible during the first part of the 70s, but a similar trend has been reported for other bores with more recent observations.

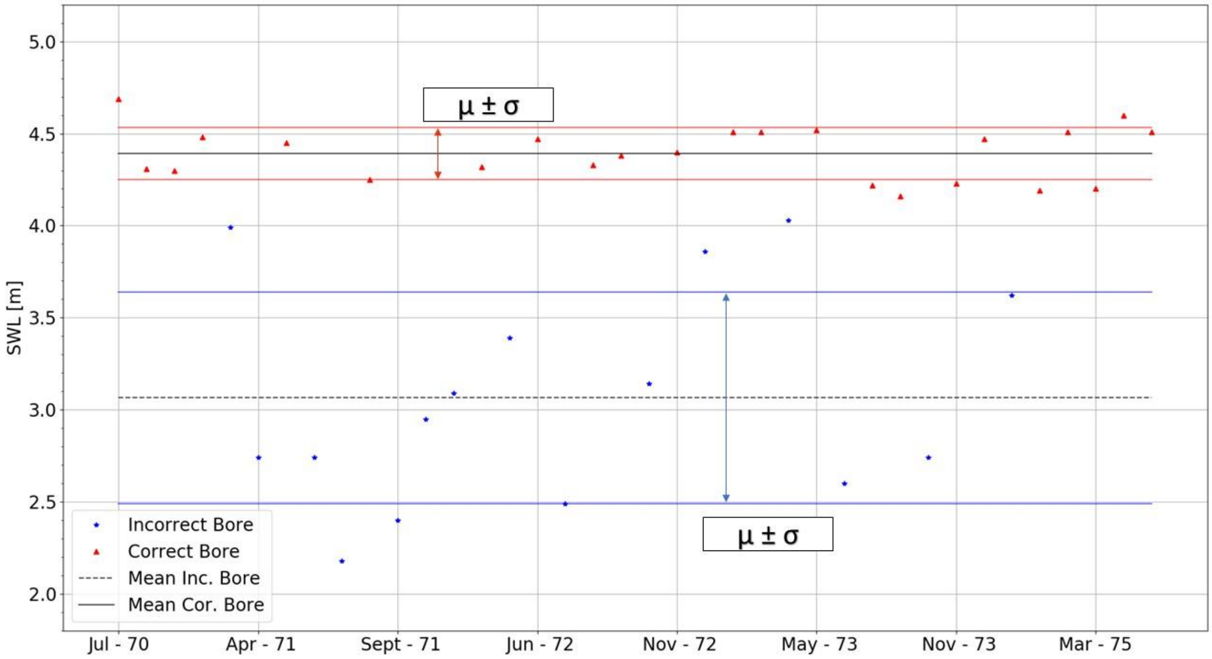

Figure 1 - Water level dynamics for two bores. SWL is distance from the surface.

As the WT dynamics at the bore selected by the referee is driven by conditions that are not representative of our simulation, a large "seasonal" fluctuation of the WT levels (in the order of 2 meters) is observable. Additionally, the aquifer monitored by this bore is classified as "Qpcb", part of the Pleistocene Bridgewater Formation. This aquifer is described as "formed by Aeolianitic calcarenite, a partially calcretised ancient dunal system". Further description of the aquifer characteristics reports that it contains minor quantities of groundwater, unconfined to semiconfined aquifers, often containing small fresh lenses that provide small stock/domestic supplies. This bore is also close to anthropogenic activity. In Figure1, the WT dynamics shown in red refers to a bore screened in correspondence to the aquifer reported in our study, which is classified as "Thgg", consisting of grey marl with coarse bioclastic presenting frequent chert band. In the figure, the substantial differences in the WT fluctuations of the two bores are also indicated by the solid lines that mark the interval within one standard deviation (σ) from the means. For a large part of the observed dataset, the seasonal fluctuation of this bore is in the order of 10-30 cm, making it hard to distinguish regional patterns from the effects of localized net-recharge. Therefore, we decided to maintain the BC head elevation constant over the simulated period.

References:

De Lannoy, G.J.M., and R.H. Reichle, Assimilation of SMOS bright-
ness temperatures or soil moisture retrievals into a land surface model, Hydrology and Earth System Sciences, 20, 4895-4911, 2016.

Lievens, H., et al. (17 co-authors), Assimilation of SMOS soil moisture and brightness temperature products into a land surface model, Remote Sensing of Environment, 180, 292-304, 2016.

Margulis, S.A., D. McLaughlin, D. Entekhabi, and S. Dunne, Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, Water Resources Research, 38(12), 1299, DOI:10.1029/2001WR001114, 2002.

Pauwels, V.R.N., and G.J.M. De Lannoy, Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space, Water Resources Research, 45, W08428, DOI:10.1029/2008WR007590, 2009.

Reichle, R.H., D.B. McLaughlin, and D. Entekhabi, Hydrologic Data Assimilation with the Ensemble Kalman Filter, Monthly Weather Review, 130, 103-114, 2002.