Kalman filters for assimilating near-surface observations in the Richards equation – Part 1: Retrieving state profiles with linear and nonlinear numerical schemes
G. B. Chirico, H. Medina, and N. Romano
Hydrol. Earth Syst. Sci. Discuss., 9, 13291-13327, 2012

Reply to Referee#2

We are grateful to Referee#2 for his/her thoughtful review, which certainly helps us in improving our paper. If the Editor agrees, the assimilation techniques will be described more in-depth in the revised manuscript. We admit that we have been too synthetic when describing the Kalman algorithms and a bit imprecise in the usage of the notations for some equations.
We hope that the following replies clarify all points raised by Referee#2.

Ref#2
Summary: the paper repeats a synthetic assimilation study by Walker et al. 2001 with (i) different numerical methods to solve the Richards equations and (ii) different Kalman-filter-based assimilation schemes.
Reply
We would like to point out that the synthetic experiment (evaporation from a uniform soil column) is also similar to the one presented by Enthekabi et al. (1994). Analogously to Walker et al. (2001), we selected this synthetic experiment to take advantage of the outcomes of these previous works while exploring the feasibility of different assimilation algorithms applied to different numerical schemes of the Richards equation.

Ref#2
1) System description
It would help if the system equations for each of the numerical schemes were written out upfront. That is, similar to Eq. 31 for the CN-scheme, please add the equations for the EX and NL schemes. Also, more explanation of the expected relative importance of the numerical scheme versus the chosen assimilation technique would be helpful.
E.g.: if a forward solution converges faster with one numerical approach than another, then this should be indicated separately from the impact of assimilating.
Reply
We agree with Referee#2. In sections 2 and 3 we will add a more comprehensive description of the system equations (similarly to what we provided for the CN scheme) and we will also illustrate the performance of the numerical schemes as well as the relative importance of these schemes versus the chosen assimilation technique.

Ref#2
P.13295 Eq. 2 and L20: Please correct the equation. What are the state system parameters w doing in the observation system? The parameters in the observation system should solely reflect the parameters of the observation operator. The model state x is itself already a function of state system parameters w.
Reply
We employed this expression for the observation function H to emphasize that the parameters of the state system might also directly affect the observation system. For instance, this is the case when we observe the soil water content (θ) and we would retrieve the matric pressure head (h) profiles (see section 4.4). The observed soil water contents are directly affected by both the state $\boldsymbol{x}$ and the system parameters $\boldsymbol{w}$, i.e

the parameters of the soil water retention function, which in our study is modelled by the van Genuchten relationship (van Genuchten, 1980):

$$\theta(h) = \theta_r + (\theta_s - \theta_r)\left(1 + |\alpha h|^n\right)^{-m} \quad (27)$$

However, to avoid any misunderstanding, in the revised manuscript we will not introduce the model parameters $\boldsymbol{w}$ in the state space equations, given that in this paper we are dealing with time-invariant parameters. Eqs. (1) and (2) will be written as follows:

$$\boldsymbol{x}_k = F(\boldsymbol{x}_{k-1}, \boldsymbol{u}_k, \boldsymbol{v}_k) \quad (1)$$

$$\boldsymbol{y}_k = H(\boldsymbol{x}_k, \boldsymbol{n}_k) \quad (2)$$

We think that reporting the system parameter vector $\boldsymbol{w}$ in H is instead relevant and necessary when we also retrieve model parameters, as shown in Parts II and III of this study (as also done by van der Merwe (2004), Moradkhani et al. (2005), Liu and Gupta (2007) and others).

Ref#2

*P.13295 L17: the forecasted state is a result of *both* the internal dynamics in F and the exogenous input u.*

Reply

We will improve the sentence accordingly. Our statement, at least as we conceived it, is in agreement with this explanation. When we say "*in response to the current exogenous input vector*" we interpret it as "*conditioned to*" and consequently also dependent on the exogenous input vector.

Ref#2

*P.13295 L24: Q_k reflects the model error variance, *not* the full uncertainty in the model predictions (the latter is P^-)*

Reply

We certainly agree on this. We are sorry for this imprecise sentence and we will change it accordingly.

Ref#2

P.13295 L7: p(u_k): this u_k should probably read n_k for obs error.

Reply

You are correct, thank you.

Ref#2

P.13295 L22: y^- = E[], remove the E[], the observation predictions are straight deterministic forward simulations

Reply

You are correct, thank you.

Ref#2

P.13295

Eq 4: No idea how this equation was obtained, but it cannot be right – the dimensions are wrong and the covariances do not make any sense. If you prefer a fancy equation different from a regular PH/[HPH+R], please do provide a few lines to allow the reader to follow.

Eq. 5: This equation is also wrong in my eyes and I cannot trace back where it comes from: should it not be Pˆ- - KHPˆ-; please prove me wrong by giving a derivation. Besides, the corresponding Eq. for the EKF (Eq. 10) is right. . .

Reply

We already replied to Referee#1 on these two points. First, we notice that there is an editing error in the first term on the right hand side of Eq. (4), as printed in the published manuscript. The correct form is the following:

$$\mathbf{K}_k = E\left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^-\right)\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)^T\right] E\left[\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)^T\right]^{-1} = \mathbf{P}_{\boldsymbol{xy},k}^- \left[\mathbf{P}_{\boldsymbol{y},k}^-\right]^{-1} \quad \text{(Eq. 4)}$$

Ref#2 refers to the more common expression of the Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}_k}^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_{\mathbf{x}_k}^- \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \quad \text{(c.\#1\_1)}$$

The expression of Eq. (c.1#1) is equivalent to Eq. (4), as shown below.

Reminding that

$$\boldsymbol{y}_k = \mathbf{H}\boldsymbol{x}_k + \boldsymbol{n}_k \quad \text{(c.\#1\_2)}$$

the covariance matrices $\mathbf{P}_{\boldsymbol{xy},k}^-$ and $\mathbf{P}_{\boldsymbol{y},k}^-$ can be expressed as follows

$$\mathbf{P}_{\boldsymbol{xy},k}^- = E\left[\left(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k^-\right)\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)^T\right] = \mathbf{P}_{\mathbf{x}_k}^- \mathbf{H} \quad \text{(c.\#1\_3)}$$

$$\mathbf{P}_{\boldsymbol{y},k}^- = E\left[\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}_k^-\right)^T\right] = \left(\mathbf{H}_k \mathbf{P}_{\mathbf{x}_k}^- \mathbf{H}_k^T + \mathbf{R}_k\right) \quad \text{(c.\#1\_4)}$$

Eq. (c.#1_1) can be then obtained by combining (c.#1_3) and (c.#1_4).

With respect to Eq. 5, Ref#2 refers to the common expression for the calculus of $\mathbf{P}_{\boldsymbol{x},k}$:

$$\mathbf{P}_{\boldsymbol{x},k} = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right)\mathbf{P}_{\boldsymbol{x},k}^- = \mathbf{P}_{\boldsymbol{x},k}^- - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^- \quad \text{(c.\#1\_5)}$$

This expression is equivalent to Eq. 5, which we repeat here for the sake of clarity:

$$\mathbf{P}_{\boldsymbol{x},k} = \mathbf{P}_{\boldsymbol{x},k}^- - \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \mathbf{K}_k^T \quad \text{(Eq. 5)}$$

With the following equations we show that $\mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \mathbf{K}_k^T = \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^-$:

$$\mathbf{K}_k \mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^- = \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \left(\mathbf{P}_{\boldsymbol{y},k}^-\right)^{-1} \mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^- = \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \left\{\left(\mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^-\right)^T \left[\left(\mathbf{P}_{\boldsymbol{y},k}^-\right)^{-1}\right]^T\right\}^T = \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \left\{\left(\mathbf{P}_{\boldsymbol{x},k}^-\right)^T \left(\mathbf{H}_k\right)^T \left[\left(\mathbf{P}_{\boldsymbol{y},k}^-\right)^{-1}\right]^T\right\}^T \quad \text{(c.\#1\_6)}$$

Since $\mathbf{P}_{\boldsymbol{x},k}^-$ and $\left(\mathbf{P}_{\boldsymbol{y},k}^-\right)^{-1}$ are symmetric matrices, they do not suffer any change with the transpose operator and hence:

$$\mathbf{K}_k \mathbf{H}_k \mathbf{P}_{\boldsymbol{x},k}^- = \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \left[\left(\mathbf{P}_{\boldsymbol{x},k}^-\right)\left(\mathbf{H}_k\right)^T \left(\mathbf{P}_{\boldsymbol{y},k}^-\right)^{-1}\right]^T = \mathbf{K}_k \mathbf{P}_{\boldsymbol{y},k}^- \left(\mathbf{K}_k\right)^T \quad \text{(c.\#1\_7)}$$

The expressions employed for $\mathbf{K}_k$ (Eq. 4) and $\mathbf{P}_{\boldsymbol{x},k}$ (Eq. 5) are equal to those employed by van der Merwe (2004). Alternative expressions can be found in the literature, as reported for instance by Grewal and Andrews (2008). We can change these equations and write them according the common usage if this might be an issue. We believe it is interesting to provide a different interpretation of the Kalman gain and the covariance P.

Ref#2

P.13298 Eq. 7-8 and 11-12: please correct: all derivatives are calculated at x_{k-1} (all4 equations) and for the error term (either v or n) set to 0 in the first of each pair of equations.

Reply

We agree that the error term in Eq. 11 should be set to 0. However, the derivates in Eq. 11-12 should be calculated using the prior estimate of the state value $\hat{\boldsymbol{x}}_k^-$. In P.13298, L8, the text in parenthesis should be corrected as follows: "*(computed at the a priori estimate* $\hat{\boldsymbol{x}}_k^-$*)*".

Ref#2

P.13299, L17, Eq.15: xˆa is an unfortunate choice as symbol for the augmented state. This xˆa is used as a symbol for the updated state ('analysis') in the DA community. Maybe choose another symbol?

Reply

We used the same symbols employed by van der Merwe (2004). If this is an issue, we can change it.

Ref#2

P.13300: Eq. 19: second diagonal term should be Q_k, not R_v, for consistency

Reply

We suggest changing the manuscript by using R (with different subscripts) for indicating the covariances of the different types of noise sources, while using "*q*" (lower case) for indicating the flux at the boundary conditions (see also the reply below).

Ref#2

P13304, Eq. 13: why here beta?, P. 13310, Eq. 33: why here Q? Please reserve Q for model error covariance.

P13310, Eq. 33 and 35: why is the time in superscripts, rather than subscripts? I thought that iterations are indicated in superscripts in this manuscript. What is f() in this equation 33?

Reply

We will correct these inconsistencies in the usage of the notations. The symbol β in equation (31) refers to the generic boundary conditions, while Q in equation (33) refers to the flux boundary condition. We suggest using the lower case q instead of Q in the revised manuscript.

The symbol f in equation (33) refers to the terms of the discrete system equation which are independent from the states.

Please, also read our reply below, where we provide clear descriptions of the SKF-CN and SKF_v-CN algorithms.

Ref#2

P13306, L17 and Eq.13-14: the SKF is really designed with additive noise terms in both the state and observation system. Consequently, naming the SKF with a subscript SKF_v is irrelevant: an SKF should not be applied with propagating errors through the dynamical state system (also, the _v is not added everywhere, why? E.g. section 4.5). I suppose the real problem is that the SKF is not described per se in this paper. How about rewriting the section 2 on 'Kalman filtering' to describe the exact basic SKF instead of giving a general description?

Reply

We admit that the manuscript might not devote enough space to explaining the difference between SKF and SKF_v. In the revised manuscript, we will provide a detailed description of these algorithms as suggested by the Referee.

Below we anticipate some key differences between SKF-CN and SKF_v-CN

To ensure physical consistency in the time-dependency one has to account (either implicitly or explicitly) for the continuous-time formulation of the dynamic model, as done, for example, by Katul et al. (1993), Entekhabi et al. (1994) and Reichle et al. (2002). When passing from the continuous formulation to the equivalent discrete-time system model, the zero-mean white Gaussian noise [$v(t)$ in our case] normally appears multiplied by a term, necessarily proportional to the computational time-step. Otherwise, it is necessary to resort to an artificial normalization of the error (see Walker (1999), Chapter 6, pp. 6.26). Maybeck (1979) provides many details about these formal aspects.

The governing soil water flow equation in its stochastic form is:

$$C(h)\frac{\partial h}{\partial t} = \frac{\partial\left[K(h)\left(\frac{\partial h}{\partial z}+1\right)\right]}{\partial z} + v(t) \qquad \text{(c\#2.1)}$$

The corresponding discrete form according to the CN differential scheme for node i at time-step k+1 is for i=2,…,N-1:

$$\left(\frac{-K_k^{i-1/2}}{2\Delta z^i \Delta z^u};\frac{C_k^i}{\Delta t_k}+\frac{\frac{K_k^{i-1/2}}{\Delta z^u}+\frac{K_k^{i+1/2}}{\Delta z^l}}{2\Delta z^i};\frac{-K_k^{i+1/2}}{2\Delta z^i \Delta z^l}\right)\begin{pmatrix}h_{k+1}^{i-1}\\ h_{k+1}^{i}\\ h_{k+1}^{i+1}\end{pmatrix} = \left(\frac{K_k^{i-1/2}}{2\Delta z^i \Delta z^u};\frac{C_k^i}{\Delta t_k}-\frac{\frac{K_k^{i-1/2}}{\Delta z^u}+\frac{K_k^{i+1/2}}{\Delta z^l}}{2\Delta z^i};\frac{K_k^{i+1/2}}{2\Delta z^i \Delta z^l}\right)\begin{pmatrix}h_k^{i-1}\\ h_k^{i}\\ h_k^{i+1}\end{pmatrix} + \frac{K_k^{i-1}-K_k^{i+1}}{2\Delta z^i} + v_k^i \qquad \text{(c\#2.2)}$$

For nodes i=1 and i=N the discrete forms at time-step k+1 are:

$$\left(\frac{C_k^1}{\Delta t_k}+\frac{K_k^{1+1/2}}{2\Delta z^i \Delta z^l};\frac{-K_k^{1+1/2}}{\Delta z^1 \Delta z^l}\right)\begin{pmatrix}h_{k+1}^{1}\\ h_{k+1}^{1+1}\end{pmatrix} = \left(\frac{C_k^1}{\Delta t_k}-\frac{K_k^{1+1/2}}{2\Delta z^1 \Delta z^l};\frac{K_k^{1+1/2}}{2\Delta z^1 \Delta z^l}\right)\begin{pmatrix}h_k^{1}\\ h_k^{1+1}\end{pmatrix} + \frac{q^1-K_k^{1+1/2}}{\Delta z^1} + v_k^1 \qquad \text{(c\#2.3)}$$

$$\left(\frac{-K_k^{N-1/2}}{\Delta z^N \Delta z^l};\frac{C_k^N}{\Delta t_k}+\frac{K_k^{N-1/2}}{2\Delta z^N \Delta z^l}\right)\begin{pmatrix}h_{k+1}^{N}\\ h_{k+1}^{N+1}\end{pmatrix} = \left(\frac{K_k^{N-1/2}}{2\Delta z^N \Delta z^l};\frac{C_k^N}{\Delta t_k}-\frac{K_k^{N-1/2}}{2\Delta z^N \Delta z^l}\right)\begin{pmatrix}h_k^{N}\\ h_k^{N+1}\end{pmatrix} + \frac{K_k^{N-1/2}-q^N}{\Delta z^N} + v_k^N \qquad \text{(c\#2.4)}$$

The forecasting equation of the system state $\boldsymbol{x}_k$, coinciding with the matric pressure head ($x_k^i = h_k^i$), can be obtained by combining the discrete equations written for all N nodes in the following linear state-space form:

$$\hat{\boldsymbol{x}}_k^- = \mathbf{A}_k^{-1}\mathbf{B}_k\hat{\boldsymbol{x}}_{k-1} + \mathbf{A}_k^{-1}\boldsymbol{g}_k + \mathbf{A}_k^{-1}\boldsymbol{v}_k \qquad \text{(c\#2.5)}$$

where $\mathbf{A}_k$ is the tri-diagonal matrix obtained by assembling the terms in the first parenthesis on the right hand-side of Eqs. c#2.2-2.4, $\mathbf{B}_k$ is the tri-diagonal matrix obtained by assembling the terms in the first parenthesis on the left hand-side of Eqs. c#2.2-2.4. The term $\boldsymbol{g}_k$ is a vector obtained by assembling the source terms on the right hand-side of Eqs. c#2.2-2.4 which are independent from the process noise.

The a priori estimate of the covariance matrix is calculated as follows:

$$\mathbf{P}_{\boldsymbol{x},k}^{-} = \mathbf{A}_k^{-1}\mathbf{B}_k\mathbf{P}_{\boldsymbol{x},k-1}\left(\mathbf{A}_k^{-1}\mathbf{B}_k\right)^T + \mathbf{A}_k^{-1}\mathbf{R}_{v,k}\left(\mathbf{A}_k^{-1}\right)^T \quad \text{(c\#2.6)}$$

where $\mathbf{R}_{v,k}$ is the covariance matrix of the zero-mean Gaussian noise $\boldsymbol{v}_k$. Notice that in Eq. c#2.6, the effect of the system noise variance is transformed by a matrix $\mathbf{A}_k$ which includes terms proportional to the time-step Δt_k.

In what we call SKFv, as implemented by Walker et al. (2001), the process noise $v(t)$ is added after the discretisation. In this case, the forecasting equations of the system state $\boldsymbol{x}_k$ and the corresponding covariance matrix are:

$$\hat{\boldsymbol{x}}_k^{-} = \mathbf{A}_k^{-1}\mathbf{B}_k\hat{\boldsymbol{x}}_{k-1} + \mathbf{A}_k^{-1}\boldsymbol{g}_k + \boldsymbol{v}_k \quad \text{(c\#2.7)}$$

$$\mathbf{P}_{\boldsymbol{x},k}^{-} = \mathbf{A}_k^{-1}\mathbf{B}_k\mathbf{P}_{\boldsymbol{x},k-1}\left(\mathbf{A}_k^{-1}\mathbf{B}_k\right)^T + \mathbf{R}_{v,k} \quad \text{(c\#2.8)}$$

From our point of view, this form of assuming the process noise explains the extremely large and unrealistic covariances reported by Walker et al. (2001) while adding a 5% of the state values, as conceived by Entekhabi et al. (1994).

This approach demands an artificial normalization of the process noise, in order to make it the independent from the time step. This "unnatural" normalization should also modify the noise covariance matrix for the calculus of the state covariance.

We implemented the SKF_v just to show that with SKF_v-CN we could get exactly the same results obtained by Walker et al. (2001) with SKF_v-EX, i.e. SKF_v implemented with an explicit finite difference scheme, as illustrated in Section 4.2. In all other cases we implemented the SKF algorithm.

Ref#2

Overall: unless I am mistaken, only scalar (one-dimensional) observations are assimilated in this paper. The use of boldface vectors for obs, and matrices for obs error covariances is thus not relevant. It may be an idea to simplify the notation to reflect the scalar nature of the obs.

Reply

The observations refer to more than one node. Only when the observation depth is equal to 0.5 cm the observation is a single scalar value.

Ref#2

*P.13293 L21: the Kalman filter is *not* a technique "to describe dynamic systems", but a technique to filter observations or to merge observations with dynamic systems.*

*L23: the Kalman filter does *not* provide a prediction of the state system, but instead it provides an *analysis* (or posterior estimate or update). The system itself provides the prediction or forecast (or prior estimate).*

Reply

The Kalman filter, as a DA technique, effectively merges information from different sources. We will change the sentence accordingly, to avoid any possible misunderstanding.

We agree, "prediction" is not a precise term. Sorry for this, we will change it accordingly.

Ref#2

P.13294 L2: EKF: 'but still widely used': where for example? Either insert a reference or delete. EK is still used in e.g. ECMWF-operations and Meteo-France, but in reality, all institutes have moved or are moving to EnKF. Also, P. 13297: I question if EKF is "undoubtedly the most widely used approach for dealing with nonlinearity".

Reply

As already replied to Referee#1 for a similar comment, we agree on this point and suggest changing lines 16-17 of page 13297 as follows:

"Within the general framework of the Kalman Filter, the Extended Kalman Filter (EKF) has been the first approach suggested for dealing with nonlinearity."

At Page 3294 Line 2, we remove "*but still widely used*" from the sentence.

Ref#2

L19: if it is important to think about DA techniques for operational settings, then it would be good to explain why the UKF is preferred in this study over the Ensemble KF (EnKF) or the particle filter (PF, admitted, the last one is no KF and may not fit in this paper): these are the most commonly used techniques in hydrologic DA. The UKF is not a commonly used KF-technique in hydrologic DA to deal with nonlinearities. What is the exact reasoning for trying the computationally more intensive UKF?

Reply

There is plenty of literature about EnKF, but only few studies deal with alternative non-linear approaches, such as the UKF technique.

In the revised manuscript, we will acknowledge that EnKF is a feasible approach for practical large-scale problems, as shown by several studies.

In the UKF, similarly to the EnKF, the update mean and covariance of the system states are approximated by the sample mean and covariance of the update ensemble. However, while the ensemble size required in the EnKF is heuristic, in the case of the Unscented Kalman Filter the ensemble is defined by the sigma points which are deterministically chosen and the number of points required is of the same order as the dimension of the system.

The sensitivity analysis carried out by Camporese et al. (2009) showed that an ensemble size greater than 50 did not add accuracy to the data assimilation scheme when considering a synthetic soil column experiment similar to that used by Walker et al. (2001). However, many of the practical drawbacks found during the implementation of the retrieving algorithms with the Richards equation using the UKF might affect the EnKF. Luo and Moroz (2009) showed that in the estimation scheme of the ordinary EnKF, the random samples generally introduce spurious modes in the transformed distribution even if the set of sample points has the correct mean and covariance.

Our decision to evaluate the UKF should be also seen in the context of our overall study, which includes retrieving model parameters. We consider that many of the generalizations emerged from state retrieving applications cannot be straightforwardly applied to parameter retrieving. In several cases parameter retrieving is normally associated with a lesser dimensionality, but a higher nonlinearity and more complex physical constraints.

However, the EnKF-UKF comparison will be a focus of subsequent studies.

We would also be interested in studying the particle filtering, following the recent discussions on this topic (Morakhadni et al. 2013; Vrugt et al., 2013).

Ref#2

4)– Numerical experiment: Unless I missed it, please indicate in the text (not just in the table) how the synthetic observations are generated, both for the pressure heads and the soil moistures. I understand that the R (obs error variance) is defined in the table, but only for the pressure heads, not for soil moisture. Did the authors also perturb the observations themselves to generate them based on the truth?

Or was the exact true value assimilated? It looks like almost perfect observations are assimilated, which is not too realistic.

Reply

Section 4.1 describes how the synthetic observations are generated.

We agree that the explanation of the numerical experiments should be also more informative.

The observations were also perturbed with respect to the "true" values, adding a random error with mean zero and a standard deviation of five percent of these values. We noted that only for deviation larger than 20% of the true values, the responses start to be sensibly distorted.

Please also note that Part III of this study is focused on real experimental data.

Ref#2

*P.13308, L7: '. . .propagated from the initial conditions' –> '. . .propagated from the initial *uniform* conditions': were the profiles each propagated with their respective numerical scheme (CN, EX, NL)? See above comment: I would like to find out the relative effect of the numerical scheme versus that of the DA, i.e. the effect on convergence speeds.*

Reply

As we mentioned above, the numerical solutions obtained with the different numerical schemes are almost identical. We could not notice any difference in terms of accuracy of the numerical solution obtained with the different numerical schemes. This was achieved by choosing an appropriate time-step for each numerical solution.

We will also specify this in the manuscript, including some statistics about the relative effect of the numerical scheme versus that of the data assimilation algorithm.

Ref#2

Section 4.2, fig 1 and the subsequent figures: assimilation at different depths is shown in fig 1: please discuss in the text what you see and relate to Walker et al. 2001 in this paragraph. The next figures do not need the results for the different assimilation depths after having discussed them once in fig 1. Instead, the subsequent figures should merge information in the different panels for 1 assimilation depth, to focus on the differences caused by the different DA aspects or numerical schemes.

Reply

We will provide more comments about the effect of the assimilation depths. However, we would like to keep the figures showing the results for all assimilations depths. Although the assimilation depth is not relevant for hourly assimilated pressure head values, it becomes more relevant when pressure heads are assimilated every two days (Figure 4) and when soil water content is the assimilated variable (Figure 5).

Ref#2

Section 4.3: could part of the success of the UKF be solely due to the iterative nature of the numerical scheme in the NL (as opposed to no iterations in the other schemes), rather than to the intrinsics of the UKF?

Reply

As stated above we verified that for the considered time steps identical responses of the numerical schemes were obtained.

Ref#2

Section 4.3 L21: LKF?? SKF instead??

Reply

Yes, thank you for evidencing this typo error.

Ref#2

Section 4.4: please indicate what is new here compared to what is already found in Walker et al. 2001?

Reply

We provide valuable insights about the practical implementation of the UKF, a scarcely pondered approach, for soil moisture retrieving using the Richard equation, some of which are extensible to the EnKF.

We provide insights about the links between the numerical scheme for formulating the Richards equation and the KF approach.

We deal with the problem of the nonlinearity in the observation equation, in particular with that associated to the observation of soil moisture while retrieving pressure heads.

New insights are also provided concerning the formulation of the process noise (also after Referee#2's suggestions) and the evolving state covariances.

Ref#2

Conclusions: that 'general guideline' is not generally useful as indicated later on: in general, we start with a model that has its own numerical scheme. Maybe think of a conclusion more in terms of suggestions for model development, rather than in terms of evaluating DA?

Reply

Thank you for the suggestion, we will do that.

As highlighted also in the conclusion, it is true that generally the assimilation algorithm is implemented with closed on-hand model software, however it is also important to point out that this is not the most desirable practice. It would be much more desirable that models are developed simultaneously with the assimilation algorithms, in order to achieve the optimal combination.

In this paper we show that, for the specific case study, a CN numerical scheme, although less numerically efficient than NL (which is implemented in the most popular soil water dynamics simulation software), when combined with SKF provides assimilation performances similar to those obtained with a UKF-NL algorithm, with much less practical drawbacks. However, we acknowledge that an explicit formulation of the covariance propagation using the standard KF constitutes a limitation of the SKF-CN in large scale applications. The UKF-based approaches represent a very suitable alternative for circumvent this issue.

References

Camporese, M., C. Paniconi, M. Putti, and P. Salandin (2009), Ensemble Kalman filter data assimilation for a process-based catchment scale model of surface and subsurface flow, Water Resour. Res., 45, W10421, doi:10.1029/2008WR007031.

Entekhabi, D., Nakamura, H., and Njoku, E. G.: Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrecuency remote sensed observations, IEEE Trans. Geosci. Remote Sensing, 32(2), 438-448, 1994.

Grewal, M.S, Andrews, A.P.: Kalman filtering: theory and practice using MATLAB— 3rd ed. John Wiley & Sons, Inc., Hoboken, New Jersey, 2008.

Katul, G.G., Wendroth, O.B., M.B. Parlange, Puente, C.E. and D.R. Nielsen: Estimation of in situ hydraulic conductivity function from nonlinear filtering theory, Water Resources Research, 29, 1063-1070, 1993.

Liu, Y. and Gupta H. V.: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007.

Luo, X. and Moroz I.M.: Ensemble Kalman filter with the unscented transform, Physica D 238, 549-562, 2009.

Maybeck, P.: Stochastic models, estimation, and control, vol. 1, Academic Press, 1979.

McLaughlin, D. B.: An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, Adv. Water Resour., 25, 1275–1286, 2002.

Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28(2), 135–147, 2005.

Moradkhani, H., DeChant, C.M., and Sorooshian S.: Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method, Water Resour. Res., 48, W12520, doi:10.1029/2012WR012144, 2013.

Reichle, R.H., McLaughlin, D.B., and Entekhabi D.: Hydrologic data assimilation with the Ensemble Kalman filter, Monthly Weather Review, 130, 103-114, 2002.

van der Merwe, R.: Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models, PhD dissertation. University of Washington, USA, 2004.

van Genuchten, M. Th.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J., 44, 892–898, 1980.

Vrugt, J.A., ter Braak, C.J. F., Diks, C.G.H., and G. Schoups: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, Adv. Water Resour., 51, 457-478, 2013

Walker, J. P.: Estimating soil moisture profile dynamics from near-surface soil moisture measurements and standard meteorological data. Ph.D. thesis, The University of Newcastle, Callaghan, New South Wales, Australia, 766 pp, 1999.

Walker, J. P., Willgoose G. R., and Kalma, J. D.: One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms, Adv. Water Resour., 24, 631-650, 2001.