Covariance resampling for particle filter – state and parameter estimation for soil hydrology

Daniel Berg[1,2], Hannes H. Bauser[1,2], and Kurt Roth[1,3]

[1]Institute of Environmental Physics (IUP), Heidelberg University
[2]HGS MathComp, Heidelberg University
[3]Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University

Correspondence to: Daniel Berg (daniel.berg@iup.uni-heidelberg.de)

Abstract. Particle filters are becoming increasingly popular for state and parameter estimation in hydrology. One of their crucial parts is the resampling after the assimilation step. We introduce a resampling method that uses the full weighted covariance information calculated from the ensemble to generate new particles and effectively avoids filter degeneracy. The ensemble covariance contains information between observed and unobserved dimensions and is used to fill the gaps between them. The covariance resampling approximately conserves the first two statistical moments and partly maintains information of higher order moments in the retained ensemble. The effectiveness of this method is demonstrated with a synthetic case – an unsaturated soil consisting of two homogeneous layers – by assimilating time domain reflectometry (TDR)-like measurements. Using this approach we can estimate state and parameters for a rough initial guess with just 100 particles. The estimated states and parameters are tested with a free run after the assimilation, which is found to be in good agreement with the synthetic truth.

1 Introduction

Mathematical models represent hydrological and other geophysical systems. They aim to describe the dynamics and the future development of system states. These models need the current state and certain system parameters (e.g. material properties, forcing) for state prediction. Both, state and system parameters, are typically ill-known and have to be estimated.

Data assimilation methods, originally used for state estimation only, are adapted to also estimate parameters and other model components like the boundary condition. The ensemble Kalman filter (EnKF) (Evensen, 1994; Burgers et al., 1998) is a popular data assimilation method in hydrology. It has the advantage of using the ensemble covariance to correlate dimensions with observations to unobserved dimensions. The EnKF with parameter estimation is applied to several hydrological systems. Moradkhani et al. (2005b) used the EnKF for a rainfall-runoff model and Chen and Zhang (2006) for saturated flow modeling. The EnKF based on Richards equation is mostly applied in synthetic studies (e.g. Wu and Margulis, 2011; Song et al., 2014; Erdal et al., 2015; Shi et al., 2015; Man et al., 2016). However, some applications to real data exist (e.g. Li and Ren, 2011; Bauser et al., 2016).

As the EnKF is based on Bayes' theorem, all uncertainties have to be represented correctly, otherwise the method has a poorer performance (Liu et al., 2012; Zhang et al., 2015). Nonlinear systems (e.g. systems described by Richards equation) violate the

EnKF assumption of Gaussian probability density functions (Harlim and Majda, 2010; DeChant and Moradkhani, 2012). The dynamics of Richards equation is generally dissipative and the Gaussian assumption is appropriate. However, jumps at layer boundaries, soliton-like fronts during strong infiltration and diverging potentials for strong evaporation deform the probability density function and lead to non-Gaussianity. In this case the probability density function requires higher statistical moments

5 to be described correctly. Particle filter can accomplish this task.

The particle filter has already been used for state and parameter estimation for various hydrological systems. Since parameters do not have their own model dynamics like the state, the resampling step is crucial. Moradkhani et al. (2005a) suggested to perturb the parameters using Gaussian noise with zero mean after the resampling step. They used an unweighted variance of the ensemble modified with a damping factor such that the ensemble spread does not become too large. This method or similar

10 has been used for instance for land surface models (Qin et al., 2009; Plaza et al., 2012), rainfall-runoff models (Weerts and El Serafy, 2006) and soil hydrology (Montzka et al., 2011; Manoli et al., 2015). A common challenge is that with only a rough initial guess, perturbing only the parameters does not guarantee a sufficient ensemble spread and the filter can diverge.

Further development of the resampling for parameter estimation was done by Moradkhani et al. (2012) and Vrugt et al. (2013). They used a Markov chain Monte Carlo (MCMC) method to generate new particles. This method was further used

15 by e.g. Yan et al. (2015) and Zhang et al. (2017). The latter compared the performance of this method with an EnKF and the particle filter presented by Moradkhani et al. (2005a) and found that the performance of the filters were similar with slight advantages for the EnKF. While the MCMC is accurate, it is also expensive, as it requires additional model runs. To increase the efficiency, Abbaszadeh et al. (2018) additionally combined it with a generic algorithm.

In this paper we introduce the covariance resampling, a resampling method that generates new particles using the ensemble

20 covariance. This method conserves the first two statistical moments in the limit of large numbers while partly maintaining higher order moments in the retained ensemble. With the covariance, the unobserved parameters of the new particles are correlated to the observed state dimensions. The particle filter with covariance resampling is able to estimate state and parameters in case of a difficult initial condition without additional model evaluations.

2 Particle Filter

25 The particle filter is an ensemble-based sequential data assimilation method that consists of a forecast and an analysis step. The ensemble members are called particles. It is used to combine information from observation and model to a posterior estimate. For a detailed review consider e.g. van Leeuwen (2009)

If new information in the form of observations becomes available, the system is propagated forward to the time the observation is taken (forecast). This results in a prior probability density function. The prior is updated with the information of the

30 observation to get the posterior. This is accomplished using Bayes' theorem,

$$P(\boldsymbol{u}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{u})P(\boldsymbol{u})}{P(\boldsymbol{d})} \, , \tag{1}$$

which describes the probability of an event $\boldsymbol{u}$ under the condition of another event $\boldsymbol{d}$. In data assimilation this is used to combine the information of the prior $P(\boldsymbol{u})$ of the state $\boldsymbol{u}$ with the observation $\boldsymbol{d}$. The probability $P(\boldsymbol{d})$ is a normalization constant

$$P(\boldsymbol{d}) = \int d\boldsymbol{u}\, P(\boldsymbol{d}|\boldsymbol{u})P(\boldsymbol{u})\,. \tag{2}$$

Both, EnKF and particle filter are Monte Carlo approaches based on Bayes' theorem with the difference that the EnKF assumes a Gaussian prior and posterior. The particle filter instead directly approximates the prior by an ensemble of realizations (particles). This direct sampling of the prior distribution allows the particle filter to have non-Gaussian probability density functions. Sampling the prior with N particles, represented by Dirac delta functions, $P(\boldsymbol{u})$ is approximated as

$$P(\boldsymbol{u}) \simeq \frac{1}{N}\sum_{i=1}^{N}\delta_D(\boldsymbol{u}-\boldsymbol{u}_i)\,. \tag{3}$$

Inserting this equation into Bayes' theorem results in an approximated posterior distribution

$$P(\boldsymbol{u}|\boldsymbol{d}) \simeq \sum_{i=1}^{N} w_i\, \delta_D(\boldsymbol{u}-\boldsymbol{u}_i)\,, \tag{4}$$

where the weights w_i are given by

$$w_i = \frac{P(\boldsymbol{d}|\boldsymbol{u}_i)}{P(\boldsymbol{d})}\,. \tag{5}$$

Since the weights are normalised to 1, the normalisation constant $P(\boldsymbol{d})$ can be obtained by:

$$\sum_{i=1}^{N} w_i = 1 \;\Rightarrow\; P(\boldsymbol{d}) = \sum_{i=1}^{N} P(\boldsymbol{d}|\boldsymbol{u}_i)\,. \tag{6}$$

In general, $P(\boldsymbol{d}|\boldsymbol{u}_i)$ is an arbitrary distribution that represents the observation error. We assume Gaussian distributed observation errors which results in:

$$P(\boldsymbol{d}|\boldsymbol{u}_i) \propto \exp\left[(\boldsymbol{d}-\mathbf{H}(\boldsymbol{u}_i))^{\intercal}\mathbf{R}^{-1}\,(\boldsymbol{d}-\mathbf{H}(\boldsymbol{u}_i))\right]\,, \tag{7}$$

where $\mathbf{R}^{-1}$ is the inverse of the observation error covariance and $\mathbf{H}$ is the observation operator that projects the state $\boldsymbol{u}$ from state-space to observation-space.

To sequentially assimilate observations, it is necessary to propagate the ensemble forward in time to the next observation. Consider the following generic model equation:

$$\boldsymbol{u}^k = \boldsymbol{f}(\boldsymbol{u}^{k-1}) + \boldsymbol{\beta}^k\,, \tag{8}$$

where $\boldsymbol{f}(\cdot)$ is the deterministic part of the model and $\boldsymbol{\beta}^k$ is a stochastic model error. The index k denotes the time. To estimate state and parameters simultaneously we use an augmented state. In our case the augmented state $\boldsymbol{u}$ consists of the state $\boldsymbol{\theta}$ (water content) and a set of parameters $\boldsymbol{p}$

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{p} \end{bmatrix}\,. \tag{9}$$
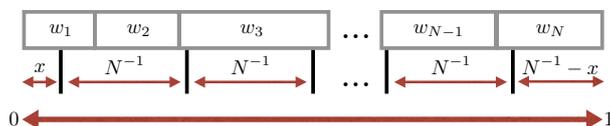
Figure 1. Illustration of the universal resampling process. A random number x is drawn from a uniform distribution in the interval $[0, N^{-1}]$. The endpoint of this number indicates the first particle. Then N^{-1} is added $(N-1)$-times to this random number, where every endpoint is a particle of the new ensemble. In the illustration, particle one is chosen once, particle two not once and particle three twice. This way some particles are replicated and other particles are dropped. If the model does not have a stochastic model error, it is necessary to perturb the new particles, otherwise they would be identical and the filter would degenerate.

3 Resampling

Particle filters tend to filter degeneracy, which is also referred to as filter impoverishment. After several analysis steps, one particle gets all statistical information as its weight becomes increasingly large, whereas the remaining particles only get a small weight such that the ensemble effectively collapses to this one particle. In this case, the filter does not react on new
5   observations and follows the particle with the large weight.

Gordon et al. (1993) introduced resampling to particle filters, a technique that reduces the variance in the weights and has the potential to prevent filter degeneracy. The idea of resampling is that after the analysis, particles with large weights are replicated and particles with small weights are dropped. This helps that the regions with high weighted particles are represented better by the ensemble, which alleviates the degeneracy of the filter. There are many different resampling algorithms (see van Leeuwen
10   (2009) for a summary). One of these methods is the stochastic universal resampling.

3.1 Stochastic Universal Resampling

The stochastic universal resampling (Kitagawa, 1996) can be summarized as follows (see also Fig. 1): All weights are aligned after each other on an interval $[0, 1]$. A random number in the interval $[0, N^{-1}]$ is drawn from a uniform distribution. This number points to the first particle of the new ensemble, selected by the corresponding weight. Then N^{-1} is added $(N-1)$-times
15   to x. Each of the endpoints selects the corresponding particle for the new ensemble. This way some particles get duplicated and some particles are dropped. With this approach, particles with a weight smaller than N^{-1} can be chosen maximally once, whereas a weight larger than N^{-1} guarantees that the particle is at least chosen once. If all particles have equal weights, no particle is dropped. The result is a new set of N particles. After the resampling step, all weights are set to N^{-1}. The stochastic universal resampling has a low sampling noise compared to other resampling methods (van Leeuwen, 2009).

20   ### 3.2 Covariance Resampling

If the model does not have a stochastic model error, like we consider in this study, it is necessary to perturb the particles, otherwise they would be identical and the filter would still degenerate. Even in the presence of a model error it can be useful to

perturb the particles after the resampling step. For example if the model error is ill-known or structurally incorrect, it can help to guarantee a sufficient ensemble spread and diversity.

There are different suggestions how to do the perturbation. For example, Moradkhani et al. (2005a) used the ensemble variance to perturb the parameters with a Gaussian with zero mean. Pham (2001) proposed to sample new particles by perturbing the identical particles using a Gaussian with the (damped) ensemble covariance matrix as covariance. Xiong et al. (2006) sampled the whole ensemble from a Gaussian using the first two moments specified by the ensemble (full covariance information), which neglects the particle filter ability to use non-Gaussian distribution.

We neither perturb the duplicated states nor draw a complete new ensemble. The covariance resampling we propose uses the universal resampling – other resampling methods can be equally used – to choose the ensemble members that are kept. Instead of duplicating the particles and setting the weights to N^{-1}, the weight of the particles is changed to

$$w_i = \frac{z}{N} \quad \text{with} \quad i \in \{1, 2, .., N'\} ,$$ (10)

where the particle i is chosen z-times and N' is the number of kept particles. This ensures that the estimated statistical moments are conserved.

The total ensemble reduces to N'. To have N ensemble members again, $N - N'$ new particles have to be generated. These particles are sampled from a Gaussian $\mathcal{N}\left(\overline{\boldsymbol{u}}, \mathbf{P}^{\mathrm{f}}\right)$ with the weighted mean

$$\overline{\boldsymbol{u}} = \sum_{i=1}^{N} w_i \, \boldsymbol{u}_i .$$ (11)

and the weighted covariance

$$\mathbf{P}^{\mathrm{f}} = \frac{1}{1 - \sum_{i=1}^{N} w_i^2} \sum_{i=1}^{N} [\boldsymbol{u}_i - \overline{\boldsymbol{u}}] [\boldsymbol{u}_i - \overline{\boldsymbol{u}}]^{\mathsf{T}} ,$$ (12)

where the factor $\frac{1}{1 - \sum_{i=1}^{N} w_i^2}$ is Bessel's correction for a weighted covariance. Mean and covariance are calculated using the weights before resampling (Eq. (5)).

Sampling only the dropped particles from Gaussian conserves the first two statistical moments in the limit of large numbers with the advantage that information of higher moments is partly conserved in the retained ensemble. A weight of N^{-1} is assigned to each of the new particles, which results in a sum of all weights larger than one. Therefore, it is necessary to normalize the weights again. The whole resampling process is illustrated in Fig. 2.

New particles are generated with a Cholesky decomposition of the covariance matrix. The calculation of the covariance from the ensemble can result in small numerical errors that have to be regularised, otherwise the decomposition would fail. For details about the generation of new particles and regularisation of the covariance matrix see Appendix A.

Pham (2001) introduced a tuning parameter to modify the covariance matrix and Moradkhani et al. (2005a) for the variance, respectively. They used the tuning factor to reduce the amplitude of the perturbation. For the covariance resampling we also introduce a tuning parameter. If the model dynamics does not support a sufficient spread for the ensemble, the perturbation of the covariance resampling has to be large enough to prevent the ensemble from degeneracy. One example for such a case are
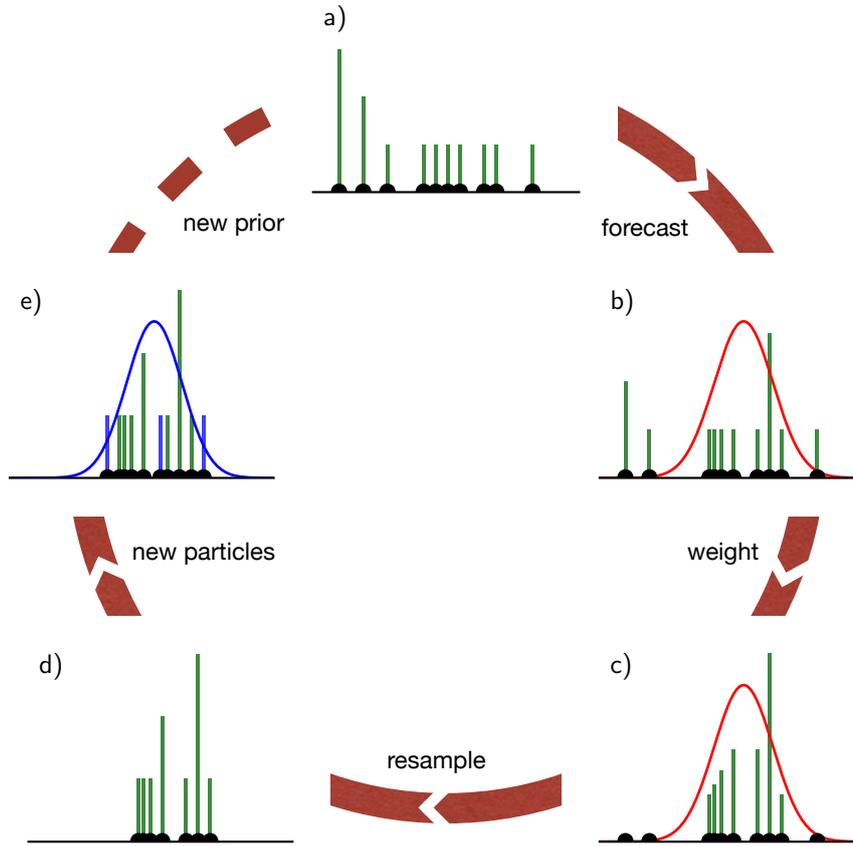
Figure 2. Illustration of the particle filter with covariance resampling. The green bars show the weight of each ensemble member (ten in this example) and the black dots the position of it. (a) The prior represented through the ensemble. (b): The ensemble is propagated to the next observation (depicted as Gaussian, red curve). (c): The particles are weighted according to the observation. At this point, some particles have already negligible weight. (d): The universal resampling drops partilces with low weight (three in this example). Instead of adding new particles at the same position, only the weights of the kept particles are changed. If a particle is resampled k-times, it will get the weight $k\,N^{-1}$. The ensemble size is reduced and new particles have to be added to conserve the ensemble size and avoid filter degeneration. (e): The new particles are drawn from the full covariance of the ensemble (Eq. (12)) and their weight is set to N^{-1}. Since new particles with weights are added to the ensemble, it is necessary to normalize the weights to one again. This results in the posterior or the next prior respectively.

parameters. The covariance matrix is modified by a multiplicative factor γ

$$\mathbf{P'}^{\mathrm{f}} = (\gamma\,\gamma^{\mathsf{T}}) \circ \mathbf{P}^{\mathrm{f}},\tag{13}$$

where $\circ$ is the entrywise product (Hadamard product). In the case of parameters the factor is chosen larger than one for the parameter space to provide a sufficient ensemble spread.

4   Case study

The algorithm is tested using a synthetic case study of a one-dimensional unsaturated porous medium with two homogeneous layers. The system has a vertical extent of $1\,\mathrm{m}$ with the layer boundary in the middle at $50\,\mathrm{cm}$. The representation of the considered system is described following the structure of Bauser et al. (2016). The general representation of a system has four components: *dynamics*, *forcing*, *subscale physics* and *state*. The dynamics propagates the state forward in time, conditioned on the subscale physics and forcing.

The dynamics in an unsaturated porous medium can be described by the Richards' equation

$$\partial_t \theta - \nabla \left[K(\theta) \left[\nabla h_m - 1 \right] \right] = 0 \,, \tag{14}$$

where $h_m (\mathrm{L})$ is the matric head, $K (\mathrm{LT^{-1}})$ the isotropic hydraulic conductivity and $\theta \, (-)$ the volumetric water content. We use the finite-element solver MuPhi (Ippisch et al., 2006) to solve Richards' equation numerically. The resolution is set to $1\,\mathrm{cm}$ which results in a 100-dimensional water content state.

The macroscopic material properties represent the averaged subscale physics with the functions $K(\theta)$ and $h_m(\theta)$ and a set of parameters. In this study, the Mualem van Genuchten parametrisation is used (Mualem (1976), Van Genuchten (1980)):

$$K(\Theta) = K_w \, \Theta^\tau \left[1 - \left[1 - \Theta^{n/[n-1]} \right]^{1-1/n} \right]^2 \,, \tag{15}$$

$$h_m(\Theta) = \frac{1}{\alpha} \left[\Theta^{-n/[n-1]} - 1 \right]^{1-1/n} \,, \tag{16}$$

with the saturation $\Theta \, (-)$

$$\Theta := \frac{\theta - \theta_r}{\theta_s - \theta_r} \,. \tag{17}$$

With these equations the subscale physics is described by six parameters for each layer. The parameter $\theta_s \, (-)$ is the saturated water content and $\theta_r \, (-)$ the residual water content. The matric head h_m is scaled with the parameter $\alpha \, (\mathrm{L^{-1}})$ that can be related to the air entry value. The parameter $K_w \, (\mathrm{LT^{-1}})$ is the saturated hydraulic conductivity, $\tau \, (-)$ a tortuosity factor and $n \, (-)$ is a shape parameter. In this study the parameters α, n and K_w will be estimated for each layer. Combining Eq. (16) and Eq. (15) results in a conductivity function

$$K(h_m) = K_w \left[1 + (\alpha h_m)^n \right]^{-\tau(1-1/n)} \left[1 - (\alpha h_m)^{n-1} (1 + (\alpha h_m)^n)^{-1+1/n} \right]^2 \tag{18}$$

that incorporates all estimated parameters.

For the true trajectories and the observations, parameters by Carsel and Parrish (1988) for sandy loam are used for the upper layer (layer 1) and loamy sand for the lower layer (layer 2). Table 1 gives the true values for the estimated parameters and Table 2 the values for the fixed parameters, respectively. In the following the parameters will have an index representing their corresponding layer.
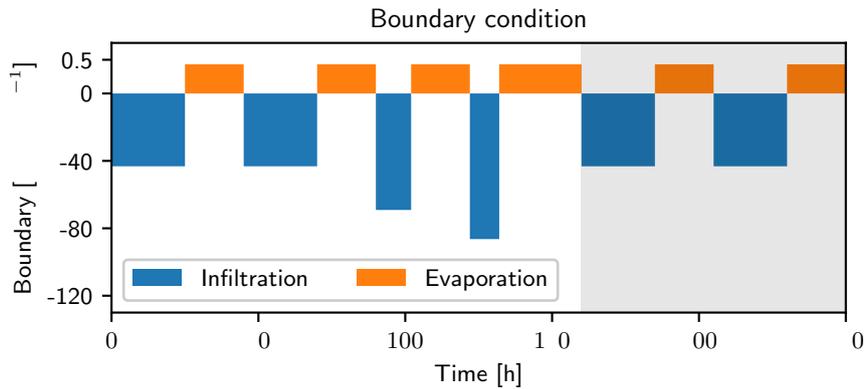
Figure 3. Upper boundary condition for the data assimilation case. Four rain events (blue) followed by a dry period (orange) are used for the test of the data assimilation run. After this run, two additional rain events and dry periods are used in a free run to test the assimilation results (grey background). The intensity and duration of these events is set equal to the first events of the data assimilation run. Note the different axes for infiltration and evaporation.

Since state and parameters are estimated, we separate the model equation Eq. (8) into

$$\mathbf{u}^n = \begin{bmatrix} \boldsymbol{\theta}^k \\ \boldsymbol{p}^k \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}(\boldsymbol{\theta}^{k-1}, \boldsymbol{p}^{k-1}) \\ \boldsymbol{p}^{k-1} \end{bmatrix}, \tag{19}$$

with a constant model for the parameters $\boldsymbol{p}$ and Richards' equation as $f(\cdot)$. Note that the model error of equation Eq. (8) is set to zero. In hydrology the model error is typically ill-known and can vary both in space and time.

5   The forcing is reflected in the boundary condition of the simulation. For the lower boundary, a Dirichlet condition with zero potential (groundwater table) is used. The upper boundary condition is chosen as a flux boundary (Neumann), representing four rain events with increasing intensity and dry periods in between (see Fig. 3).

Using infiltrations with an increasing intensity has the advantage that the system has more time to adjust the parameters. This way less observations are necessary to resolve the infiltration front and the information of the observations can be incorporated
10   in the state and parameters. The stronger infiltration front in the end is used to test the robustness of the estimated state and paramters.

The last component of the system is the state. Initially, the system is in equilibrium and will be forced by the boundary condition. The initial state is depicted in Fig. 4. Six TDR-like observations are taken equidistantly in each layer at the positions $(0.1, 0.25, 0.3)$ m for layer 1 and $(0.6, 0.75, 0.9)$ m for layer 2. The measurement error is chosen to be $\sigma_{\mathrm{Obs}} = 0.007$ (e.g.
15   Jaumann and Roth, 2017). Observations are taken hourly for the duration of $160\,\mathrm{h}$.

For the initial state of the data assimilation, the observations at time zero are used. The measured water content is interpolated linearly between the measurements and kept constant towards the boundary. Additionally, the saturated water content for loamy
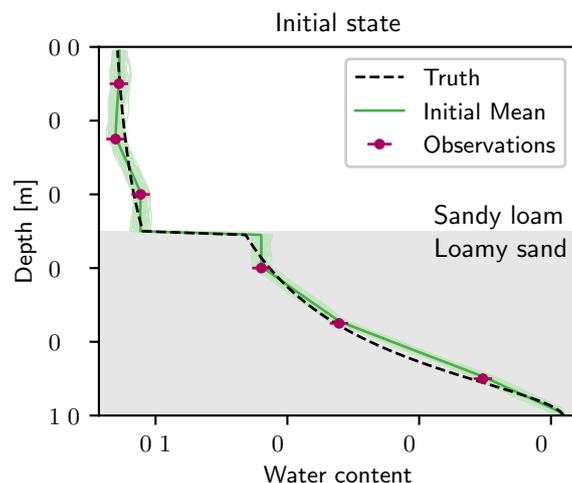
Figure 4. Initial state for the data assimilation run. Observations at time zero are connected linearly and set constant towards the layer boundary. The ensemble (light green) with 100 ensemble members is generated by perturbing this state with a Gaussian with a specified covariance. The initial truth that is used for the observations is shown in black.

sand, which is 0.41 is taken as the lower boundary. The approximated state is used as the ensemble mean for the particle filter. This procedure is a viable option for real data although it represents a rather crude approximation of the real initial condition.

To generate the ensemble, the approximated state is perturbed by a correlated multivariate Gaussian. The standard deviation of this perturbation is chosen to be 0.003. The value is chosen such that the ensemble represents the uncertainty of the water content in most parts (see Fig. 4). The square of it is used for the whole covariance matrix but is modified using the following two steps:

1. All covariances between the two layers are set to zero to ensure no correlations across the layer boundary.

2. The matrix is multiplied with the fifth-order polynomial Gaspari and Cohn (Gaspari and Cohn, 1999) function using a length scale $c = 10\,\mathrm{cm}$. This function has a compact support and approximates a Gaussian with half-width c. This way, the water content is only correlated in the range of $20\,\mathrm{cm}$.

The use of the covariance increases the diversity of the ensemble and also helps to avoid degeneration. Using uncorrelated Gaussian random numbers with zero mean would result in a fast degeneration of the particle filter caused by the dissipative nature of the system. The perturbation would simply dissipate and the ensemble collapses.

The initial parameters for the ensemble are uniformly distributed. The ranges of the uniform distributions are given in Tab. 1. Note that the logarithm of the saturated conductivity K_w is used for the estimation, so K_w spans three order of magnitudes. The ensemble size is 100. The multiplicative factor Eq. (13) is set to:

$$\gamma = \begin{bmatrix} 1_{100} \\ 1.2_6 \end{bmatrix}, \tag{20}$$

Table 1. True Mualem van Genuchten parameters and range of the uniformly distributed initial guess.

Parameter	Truth	Lower	Upper
$\log(K_{w,1})$	-4.40	-7	-4
$\log(K_{w,2})$	-4.91	-7.5	-4
n_1	2.28	2.2	3.5
n_2	1.89	1.8	3.2
α_1	-12.4	-14	-12
α_2	-7.5	-10.5	-6.5

Table 2. Fixed Mualem van Genuchten parameters.

Parameter	Layer 1	Layer 2
θ_s	0.41	0.41
θ_r	0.057	0.065
τ	0.5	0.5

where the index denotes the dimension. The covariance in the 100-dimensional state space is unchanged. For the parameter space a factor of 1.2 is used to compensate the missing dynamics.

After the assimilation, the final mean of the augmented state is used to run a forecast without data assimilation. In this run two additional infiltration events and evaporation periods (see Fig. 3) are used to test the forecasting ability of the estimated states and parameters.

5 Results

The development of the parameters is depicted in Fig. 5. The saturated conductivity $K_{w,1}$ (Fig. 5a) can be estimated quickly because the infiltration front induces dynamics in the first layer which is strongly dependent on K_w. Instead $K_{w,2}$ (Fig. 5b) is not sensitive to the dynamics in the first hours, as the infiltration front did not reach the second layer yet. At around $46\,\mathrm{h}$, the infiltration front reaches the first observation position in the second layer and the estimation for $K_{w,2}$ improves quickly.

If dynamic is induced in the system, the ensemble spread in the water content space increases because of different parameter sets. This makes the particles and their corresponding parameter sets distinguishable and parameter estimation possible. The parameters n_1 and n_2 (Fig. 5c and d) as well as α_2 (Fig. 5f) can be estimated well. One exception is α_1 (Fig. 5e). This parameter is insensitive to the observations. The effect of α on the trajectory of the ensemble members is limited to a small region next to the layer boundary. Further away, wrong values can be compensated by n. The parameter α_1 jitters randomly around a value slightly larger than the truth. In a synthetic case, the estimation of α_1 can be improved easily by having observations directly next to the boundary. This way a different value for α_1 would have a direct influence on the trajectory and α_1 would become
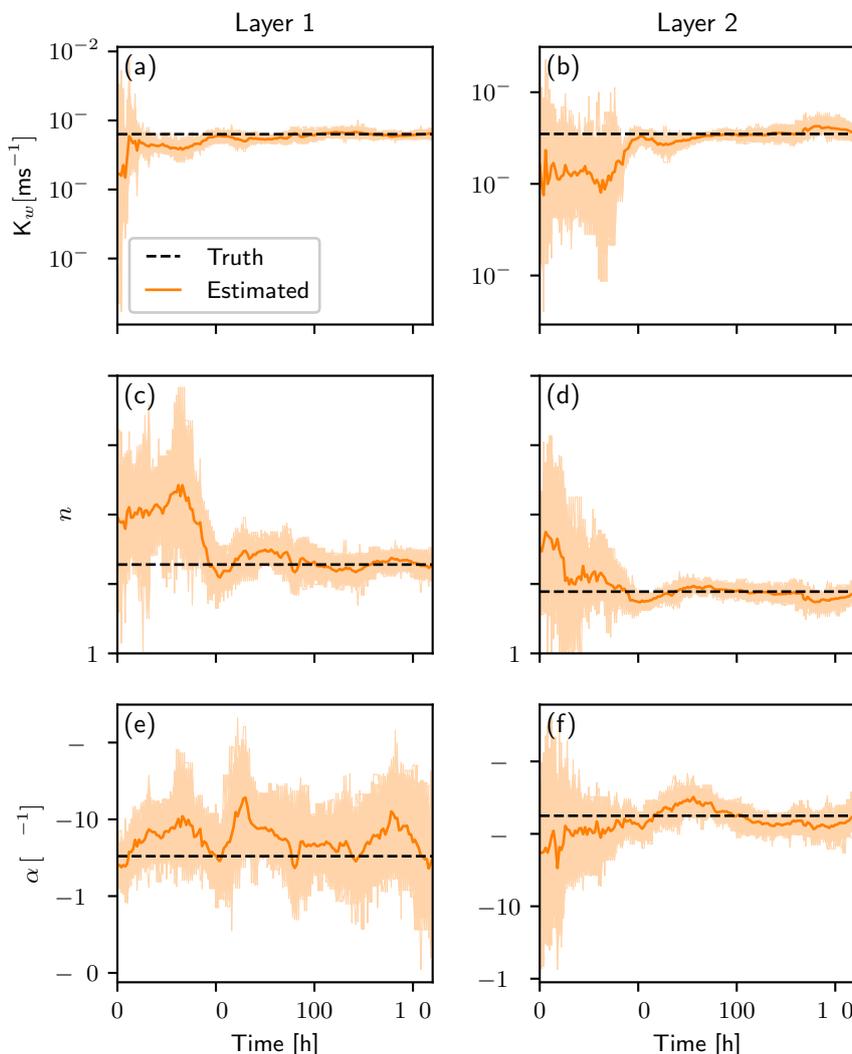
Figure 5. Estimation of all six parameters ((a): $K_{w,1}$, (b): $K_{w,2}$, (c): n_1, (d): n_2, (e): α_1, (f): α_2) over time. The ensemble mean is shown in orange and the ensemble in light orange. The truth is a dashed black line.

sensitive to the observations. However, in reality it is not feasible to change the measurement position or measure directly next to the layer interface. We decided to retain these positions to illuminate an often encountered practical difficulty.

To see the effect of the parameters on the forward propagation, it is necessary to have a closer look at the conductivity function Eq. (18). This function is used for the model forward propagation and many parameter sets can effectively describe 5 the same situation in a limited regime of the hydraulic head. The function is shown in Fig. 6 for both layers and reveals that the difference between the truth and the estimated parameters is small.
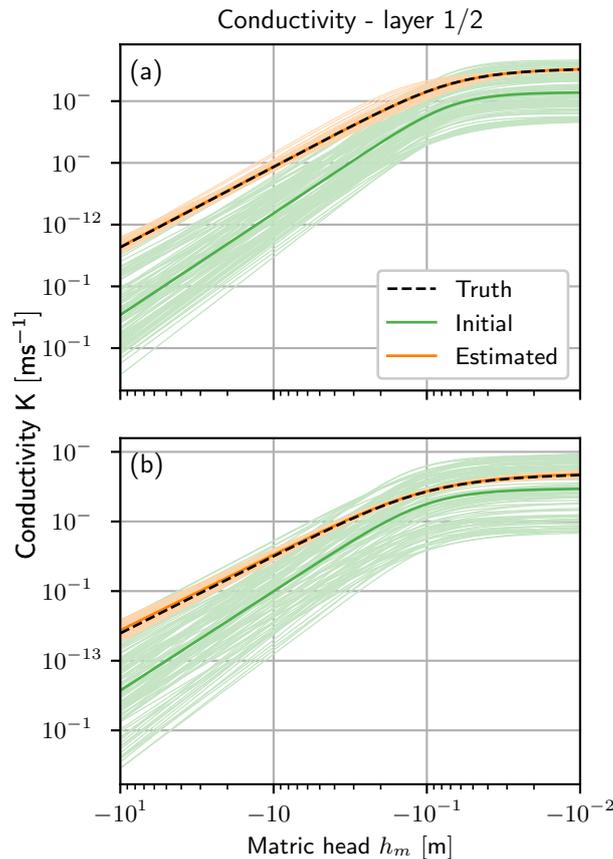
Figure 6. Conductivity function $K(h_m)$ (Eq. (18)) for (a): layer 1 and (b): layer 2. In this function all estimated parameters are represented. The initial ensemble (light green) with the mean (green) are shown. The truth (dashed black) is almost congruent with the estimated mean (orange), such that only the final ensemble (light orange) is visible.

The final water content state agrees with the synthetic truth in a narrow band (see Fig. 7). The estimated state is slightly biased to higher water contents. We checked that the direction of the bias is a random effect and is different for different seeds. The observation of a bias instead is caused by long range correlations of the model. In this case, the system has started to relax after the last infiltration and a higher water content in one part results in a higher water content in the rest of the layer. The ensemble spread next to the layer boundary is caused by the large spread of α_1.

To analyse the ensemble, we take a closer look at the effective sample size and the number of particles that are resampled. The effective sample size is defined as (Doucet, 1998):

$$N_{\text{eff}} = \frac{1}{\sum_{i=1}^{N} w_i^2} \tag{21}$$
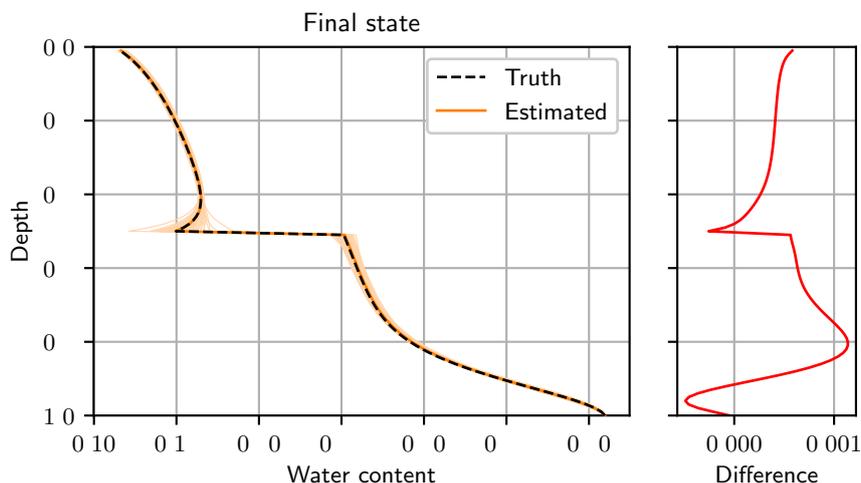
Figure 7. Final water content state after the assimilation run. The truth (dashed black) is almost congruent with the estimated mean (orange), such that only the final ensemble (light orange) is visible. The estimated mean is used to start a free forward run afterwards. The difference of the estimated water content and the synthetic truth lies in a narrow band, with a small bias to larger water contents.

and gives an estimate how many particles are significant. For example, if one particle accumulates all the weight $N_{\text{eff}} = 1$ and the ensemble is effectively described only by this particle.

Fig. 8 shows the effective sample size and the number of new particles over time. The effective sample size drops every time new information is available and the number of resampled particles increases. For times $t < 15\text{h}$, the effective sample

5   size drops to small values. The infiltration front propagates through the first layer, which leads to a large ensemble spread caused by unknown parameters and only a few particles have a significant weight. The filter assimilates the information from the observations and resamples particles with low weight. This improves the state and parameters and leads to an increasing effective sample size until the infiltration front reaches the second layer ($t \approx 46\text{h}$). The effective sample size decreases rapidly because the parameters in the second layer are still unknown and lead to a large ensemble spread again. This variation of the

10 effective sample size occurs every time the filter gets new information that leads to a discrepancy between states of the particles and the observations.

The effective sample size is a crucial parameter for the covariance resampling. If it drops to low values the filter can degenerate because effectively too few particles contribute to the weighted covariance (Eq. (12)) and the covariance information becomes insignificant.

15 For further analysis, the RMSE is calculated based on the difference of the true water content and the weighted mean at every observation time. This includes also the unobserved dimensions. The RMSE shows a similar behaviour as the parameters and the effective sample size (see Fig. 9). During the first infiltration, the state and the parameters are improved, the RMSE becomes smaller. Then the infiltration front reaches the boundary interface. The parameters of the second layer and α_1 are still
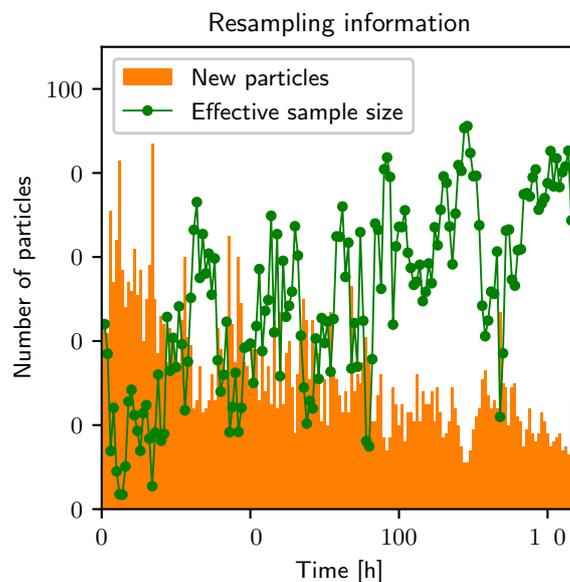
Figure 8. Amount of particles that are resampled (orange) and the effective ensemble size (green dots). The lines connecting the dots are for better visibility of the time dependent behaviour. The effective sample size increases while the number of resampled particles decreases. Every infiltration reduces the effective sample size and leads to more resampled particles.

wrong and diverse. This leads to a spread of the ensemble by the system dynamics and also a shift of the mean away from the truth. The parameters in the second layer are estimated and the state is corrected, which leads to a decrease in the RMSE. The increase for the next infiltration events becomes smaller since state and parameters are already improved.

The forward run without data assimilation shows that the RMSE oscillates in a narrow range. These oscillations are part of

5 the unobserved space next to the boundary and are mainly caused by the wrong value of α for the first layer. In the beginning, the RMSE stays small, but when the infiltration front reaches the boundary of the two layers, the mean state and the truth begin to deviate from each other (limited to the boundary interface). After the infiltration front passed, the state starts to equilibrate and is increasingly defined by the whole parameter set, which has a certain distance to the true equilibrium.

6 Summary and Conclusions

10 We introduced a resampling method for particle filters that uses the covariance information of the ensemble to generate new particles and effectively avoids filter degeneracy. The method was tested in a synthetic one-dimensional unsaturated porous medium with two homogeneous layers. Even with just a rough initial guess, a broad parameter range and only 100 ensemble members, the estimation shows excellent results. After the assimilation, the results are verified in a free run with the final results.
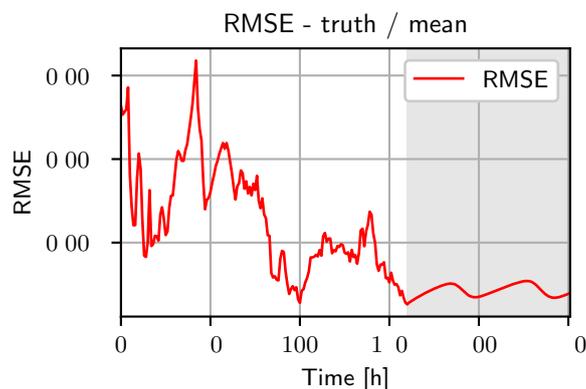
Figure 9. The RMSE (red) of the water content calculated between the truth and the estimated mean including all dimensions. After 160 hours the free run starts (grey background). During this forward run, the RMSE is about 10^{-3}. For the assimilation and the free run the RMSE increases with each infiltration.

The covariance connects information between observed and unobserved dimensions. This has some similarity to the ensemble Kalman filter but in our approach information of higher order moments is partly maintained in the retained ensemble. Even tough the RMSE of the water content includes the unobserved state dimensions, it stays in a narrow range (RMSE is about 10^{-3}) during the forecast. With every infiltration, the RMSE shows excursions caused by a wrong value of parameter α in the first layer that results in a wrong state near the layer boundary during the infiltration.

Transferring the information to the unobserved dimensions helps the filter to not degenerate when only a rough initial guess is available. The states and parameters are both altered actively. For the used initial condition, perturbing the parameters only (Moradkhani et al., 2005a), can lead to filter degeneracy because the state is only changed by the dynamics of the system. Compared to the particle filter with MCMC resampling (Moradkhani et al., 2012; Vrugt et al., 2013), the covariance resampling presented in this study has the advantage that it does not need additional model runs to generate new particles. However, the covariance resampling has to calculate the covariance matrix and perform a Cholesky decomposition every assimilation step. Similar to localization for the ensemble Kalman filter (Houtekamer and Mitchell, 2001; Hamill et al., 2001), it is possible to localize the covariance in the resampling to increase the efficiency.

The effective sample size (Eq. (21)) is a crucial parameter for this method. The covariance resampling needs a sufficient effective sample size, otherwise the calculation of the covariance matrix (Eq. (12)) becomes inaccurate and the filter may degenerate. In such a situation, the filter can be improved by resetting the weights to N^{-1} or increasing the ensemble size. In our example this was not necessary because the effective sample size was critical only for single assimilation steps.

Different parameter sets can approximately describe the same conductivity function (Eq. (18)) in a certain matric head regime. Model dynamics is necessary to differentiate between those sets. If the infiltration covers only a small regime, the conductivity function is only significant in the observed range and can differ from the truth otherwise. This is also reflected in

the chosen boundary condition. Starting with infiltrations with low intensity but longer duration helps the filter to explore the water content range slowly and the observations can resolve the infiltration front.

The covariance resampling connects observed with unobserved dimensions to effectively estimate parameters and prevent filter degeneracy. It conserves the first two statistical moments in the limit of large numbers, while partly maintaining higher order moments in the retained ensemble. The method is able to estimate state and parameters in case of a difficult initial condition without additional model evaluations and using a rather small ensemble size.

Appendix A: Generation of correlated random numbers

A1 Cholesky Decomposition

Correlated random numbers are generated using the Cholesky decomposition. We use the LDLT decomposition which is part of the *Eigen3* library (Guennebaud, Jacob et al., 2010). Decomposing the covariance matrix $\mathbf{Q}$ leads to

$$\mathbf{Q} = \mathbf{LDL}^{\mathsf{T}}, \tag{A1}$$

where $\mathbf{D}$ is a diagonal matrix and $\mathbf{L}$ is a lower unit triangular matrix. The LDLT form of the decomposition is related to the LLT-form by

$$\mathbf{Q} = \mathbf{L}'\mathbf{L}'^{\mathsf{T}} \quad \text{with} \quad \mathbf{L}' := \mathbf{LD}^{\frac{1}{2}}. \tag{A2}$$

To draw a random vector $\mathbf{y}$ from a Gaussian distribution $\mathcal{N}\left(\boldsymbol{\mu}, \mathbf{Q}\right)$ with mean $\boldsymbol{\mu}$, we first generate a normal distributed ($\mathcal{N}\left(\mathbf{0}, \mathbf{I}\right)$) random vector $\mathbf{x}$. This vector is multiplied with $\mathbf{L}'$ and the mean $\boldsymbol{\mu}$ is added:

$$\mathbf{y} = \mathbf{L}'\mathbf{x} + \boldsymbol{\mu} \tag{A3}$$

To verify that this gives the correct result the covariance matrix of $\mathbf{y}$ is calculated:

$$\overline{(\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^{\mathsf{T}}} = \overline{\mathbf{L}'\mathbf{x}(\mathbf{L}'\mathbf{x})^{\mathsf{T}}} = \mathbf{L}'\overline{\mathbf{x}\mathbf{x}^{\mathsf{T}}}\mathbf{L}'^{\mathsf{T}} = \mathbf{L}'\mathbf{I}\mathbf{L}'^{\mathsf{T}} = \mathbf{Q} \tag{A4}$$

yields $\mathbf{Q}$ as required.

A2 Regularisation of the ensemble covariance matrix

The calculation of the Cholesky decomposition (LDLT-version) is only possible if the matrix is not indefinite. Mathematically, a covariance matrix has to be positive semidefinite:

$$\boldsymbol{v}^{\mathsf{T}}\mathbf{Q}\boldsymbol{v} = \boldsymbol{v}^{\mathsf{T}}\sum(\boldsymbol{y}_i - \boldsymbol{\mu})(\boldsymbol{y}_i - \boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{v} \tag{A5}$$

$$= \sum \boldsymbol{v}^{\mathsf{T}}(\boldsymbol{y}_i - \boldsymbol{\mu})(\boldsymbol{y}_i - \boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{v} \tag{A6}$$

$$= \sum (\boldsymbol{v}^{\mathsf{T}}(\boldsymbol{y}_i - \boldsymbol{\mu}))^2 \geq 0 \quad \text{with} \quad \boldsymbol{v} \in \mathrm{I\!R}^{\mathrm{d}}, \tag{A7}$$

but the covariance matrix calculated from our ensemble is occasionally indefinite. The reason for the covariance matrix being indefinite is a numerical error in the calculation of this matrix. In fact, the calculation of the eigenvalues λ results in negative values in the order of $\mathcal{O}(10^{-20})$.

For this purpose, the identity matrix $\mathbf{I}$, which is multiplied by a scalar λ_{max}, is added to the covariance matrix. The value of
5 λ_{max} is in the order of magnitude of the largest negative eigenvalue of $\mathbf{Q}$. Thus, the regularised covariance matrix reads

$$\mathbf{Q}_{\mathrm{Reg.}} = \mathbf{Q} + \lambda_{\mathrm{max}}\mathbf{I}. \tag{A8}$$

In our experiments, the smallest variance on the main diagonal of the covariance matrix is still 16 orders of magnitude larger than λ_{max} such that the influence of this correction is negligible and does not change the results.

Competing interests. All authors declare that they have no competing interests.

10 *Acknowledgements.* This research is funded by Deutsche Forschungsgemeinschaft (DFG) through project RO 1080/12-1. Hannes H. Bauser and Daniel Berg were supported in part by the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), funded by DFG grant GSC 220 in the German Universities Excellence Initiative.

References

Abbaszadeh, P., Moradkhani, H., and Yan, H.: Enhancing hydrologic data assimilation by evolutionary Particle Filter and Markov Chain Monte Carlo, Advances in Water Resources, 111, 192 – 204, https://doi.org/10.1016/j.advwatres.2017.11.011, 2018.

Bauser, H. H., Jaumann, S., Berg, D., and Roth, K.: EnKF with closed-eye period – towards a consistent aggregation of information in soil hydrology, Hydrology and Earth System Sciences, 20, 4999–5014, https://doi.org/10.5194/hess-20-4999-2016, 2016.

Burgers, G., van Leeuwen, P. J., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2, 1998.

Carsel, R. F. and Parrish, R. S.: Developing joint probability distributions of soil water retention characteristics, Water Resources Research, 24, 755–769, https://doi.org/10.1029/WR024i005p00755, 1988.

Chen, Y. and Zhang, D.: Data assimilation for transient flow in geologic formations via ensemble Kalman filter, Advances in Water Resources, 29, 1107 – 1122, https://doi.org/10.1016/j.advwatres.2005.09.007, 2006.

DeChant, C. M. and Moradkhani, H.: Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting, Water Resources Research, 48, n/a–n/a, https://doi.org/10.1029/2011WR011011, w04518, 2012.

Doucet, A.: On sequential simulation-based methods for bayesian filtering, Tech. rep., 1998.

Erdal, D., Rahman, M., and Neuweiler, I.: The importance of state transformations when using the ensemble Kalman filter for unsaturated flow modeling: Dealing with strong nonlinearities, Advances in Water Resources, 86, 354 – 365, https://doi.org/10.1016/j.advwatres.2015.09.008, data assimilation for improved predictions of integrated terrestrial systems, 2015.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99, 10 143–10 162, https://doi.org/10.1029/94JC00572, 1994.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological Society, 125, 723–757, https://doi.org/10.1002/qj.49712555417, 1999.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, IEE Proceedings F - Radar and Signal Processing, 140, 107–113, https://doi.org/10.1049/ip-f-2.1993.0015, 1993.

Guennebaud, G., Jacob, B., et al.: Eigen v3, http://eigen.tuxfamily.org, 2010.

Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter, Monthly Weather Review, 129, 2776–2790, https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2, 2001.

Harlim, J. and Majda, A. J.: Catastrophic filter divergence in filtering nonlinear dissipative systems, Commun. Math. Sci., 8, 27–43, https://projecteuclid.org:443/euclid.cms/1266935012, 2010.

Houtekamer, P. L. and Mitchell, H. L.: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 129, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2, 2001.

Ippisch, O., Vogel, H.-J., and Bastian, P.: Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation, Advances in Water Resources, 29, 1780 – 1789, https://doi.org/10.1016/j.advwatres.2005.12.011, 2006.

Jaumann, S. and Roth, K.: Effect of unrepresented model errors on estimated soil hydraulic material properties, Hydrology and Earth System Sciences, 21, 4301–4322, https://doi.org/10.5194/hess-21-4301-2017, 2017.

Kitagawa, G.: Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models, Journal of Computational and Graphical Statistics, 5, 1–25, http://www.jstor.org/stable/1390750, 1996.

Li, C. and Ren, L.: Estimation of unsaturated soil hydraulic parameters using the ensemble Kalman filter, Vadose Zone Journal, 10, 1205–1227, 2011.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A. I. J. M., van Velzen, N., He, M., Lee, H., Noh, S. J., Rakovec, O., and Restrepo, P.: Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, Hydrology and Earth System Sciences, 16, 3863–3887, https://doi.org/10.5194/hess-16-3863-2012, https://www.hydrol-earth-syst-sci.net/16/3863/2012/, 2012.

Man, J., Li, W., Zeng, L., and Wu, L.: Data assimilation for unsaturated flow models with restart adaptive probabilistic collocation based Kalman filter, Advances in Water Resources, 92, 258 – 270, https://doi.org/https://doi.org/10.1016/j.advwatres.2016.03.016, http://www.sciencedirect.com/science/article/pii/S0309170816300793, 2016.

Manoli, G., Rossi, M., Pasetto, D., Deiana, R., Ferraris, S., Cassiani, G., and Putti, M.: An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment, Journal of Computational Physics, 283, 37 – 51, https://doi.org/10.1016/j.jcp.2014.11.035, 2015.

Montzka, C., Moradkhani, H., Weihermüller, L., Franssen, H.-J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, Journal of Hydrology, 399, 410 – 421, https://doi.org/10.1016/j.jhydrol.2011.01.020, 2011.

Moradkhani, H., Hsu, K.-L., Gupta, H., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resources Research, 41, 2005a.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state–parameter estimation of hydrological models using ensemble Kalman filter, Advances in Water Resources, 28, 135 – 147, https://doi.org/10.1016/j.advwatres.2004.09.002, 2005b.

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 Resources Research, 48, n/a–n/a, https://doi.org/10.1029/2012WR012144, w12520, 2012.

Mualem, Y.: A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resources Research, 12, 513–522, https://doi.org/10.1029/WR012i003p00513, 1976.

Pham, D. T.: Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems, Monthly Weather Review, 129, 1194–1207, https://doi.org/10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2, 2001.

Plaza, D. A., De Keyser, R., De Lannoy, G. J. M., Giustarini, L., Matgen, P., and Pauwels, V. R. N.: The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter, Hydrology and Earth System Sciences, 16, 375–390, https://doi.org/10.5194/hess-16-375-2012, 2012.

Qin, J., Liang, S., Yang, K., Kaihotsu, I., Liu, R., and Koike, T.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, Journal of Geophysical Research: Atmospheres, 114, n/a–n/a, https://doi.org/10.1029/2008JD011358, d15103, 2009.

Shi, L., Song, X., Tong, J., Zhu, Y., and Zhang, Q.: Impacts of different types of measurements on estimating unsaturated flow parameters, Journal of Hydrology, 524, 549 – 561, https://doi.org/10.1016/j.jhydrol.2015.01.078, 2015.

Song, X., Shi, L., Ye, M., Yang, J., and Navon, I. M.: Numerical comparison of iterative ensemble Kalman filters for unsaturated flow inverse modeling, Vadose Zone Journal, 13, 2014.

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil science society of America journal, 44, 892–898, 1980.

van Leeuwen, P. J.: Particle Filtering in Geophysical Systems, Monthly Weather Review, 137, 4089, https://doi.org/10.1175/2009MWR2835.1, 2009.

Vrugt, J. A., ter Braak, C. J., Diks, C. G., and Schoups, G.: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, Advances in Water Resources, 51, 457 – 478, https://doi.org/10.1016/j.advwatres.2012.04.002, 35th Year Anniversary Issue, 2013.

Weerts, A. H. and El Serafy, G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resources Research, 42, n/a–n/a, https://doi.org/10.1029/2005WR004093, w09403, 2006.

Wu, C.-C. and Margulis, S. A.: Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach, Journal of Hydrology, 399, 313 – 325, https://doi.org/https://doi.org/10.1016/j.jhydrol.2011.01.011, http://www.sciencedirect.com/science/article/pii/S0022169411000369, 2011.

Xiong, X., Navon, I. M., and Uzunoglu, B.: A Note on the Particle Filter with Posterior Gaussian Resampling, Tellus A, 58, 456–460, https://doi.org/10.1111/j.1600-0870.2006.00185.x, 2006.

Yan, H., DeChant, C. M., and Moradkhani, H.: Improving soil moisture profile prediction with the particle filter-Markov chain Monte Carlo method, IEEE Transactions on Geoscience and Remote Sensing, 53, 6134–6147, 2015.

Zhang, D., Madsen, H., Ridler, M. E., Refsgaard, J. C., and Jensen, K. H.: Impact of uncertainty description on assimilating hydraulic head in the MIKE SHE distributed hydrological model, Advances in Water Resources, 86, 400 – 413, https://doi.org/10.1016/j.advwatres.2015.07.018, data assimilation for improved predictions of integrated terrestrial systems, 2015.

Zhang, H., Hendricks Franssen, H.-J., Han, X., Vrugt, J. A., and Vereecken, H.: State and parameter estimation of two land surface models using the ensemble Kalman filter and the particle filter, Hydrology and Earth System Sciences, 21, 4927–4958, https://doi.org/10.5194/hess-21-4927-2017, 2017.