



# Kalman filters for assimilating near-surface observations into the Richards equation – Part 3: Retrieving states and parameters from laboratory evaporation experiments

H. Medina<sup>1</sup>, N. Romano<sup>2</sup>, and G. B. Chirico<sup>2</sup>

<sup>1</sup>Department of Basic Sciences, Agrarian University of Havana, Havana, Cuba

<sup>2</sup>Department of Agricultural Engineering, University of Naples Federico II, Naples, Italy

Correspondence to: G. B. Chirico (gchirico@unina.it)

Received: 5 November 2012 – Published in Hydrol. Earth Syst. Sci. Discuss.: 3 December 2012

Revised: 12 November 2013 – Accepted: 19 November 2013 – Published: 4 July 2014

**Abstract.** The purpose of this work is to evaluate the performance of a dual Kalman filter procedure in retrieving states and parameters of a one-dimensional soil water budget model based on the Richards equation, by assimilating near-surface soil water content values during evaporation experiments carried out under laboratory conditions. The experimental data set consists of simultaneously measured evaporation rates, soil water content and matric potential profiles. The parameters identified by assimilating the data measured at 1 and 2 cm soil depths are in very good agreement with those obtained by exploiting the observations carried out in the entire soil profiles. A reasonably good correspondence has been found between the parameter values obtained from the proposed assimilation technique and those identified by applying a non-sequential parameter estimation method. The dual Kalman filter also performs well in retrieving the water state in the porous system. Bias and accuracy of the predicted state profiles are affected by observation depth changes, particularly for the experiments involving low state vertical gradients. The assimilation procedure proved flexible and very stable in both experimental cases, independently from the selected initial conditions and the involved uncertainty.

adequately addressing the different sources of uncertainties embedded in the simulation process (Hoeben and Troch, 2000; Vrugt et al., 2005; McLaughlin, 2002; Vereecken et al., 2008), in particular those arising from the parameterization of the soil hydraulic properties, i.e. the soil water retention and hydraulic conductivity functions, which are fundamental for reliably modelling soil water dynamics in the vadose zone.

Several studies have pointed out the potential of data assimilation (DA) techniques, in particular those based on sequential algorithms as the Kalman filter (Kalman, 1960), to improve the forecast offered by a hydrological model by using near-surface soil water content observations derived from remote sensing or ground-based networks. As defined by Liu and Gupta (2007), DA aims at providing consistent estimates of the dynamical behaviour of a system by merging the information available in imperfect models and uncertain data in an optimal way. Nonetheless, the majority of soil hydrology studies deal with the use of near-surface soil moisture information to retrieve soil moisture profiles in the unsaturated zone while assuming the soil hydraulic properties as known attributes (e.g. Entekhabi et al., 1994; Walker et al., 2001, 2004; Dunne and Entekhabi, 2005). Poorly identified parameters usually generate error and bias, which should be faced using bias-correction strategies (De Lannoy et al., 2007a, b) or, more generally, a model calibration.

Classical algorithms for model calibration, either the traditional approaches accounting for deterministic methods, or those based on the standard Bayes' law, calibrate model parameters by implementing an optimization procedure that

## 1 Introduction

With the unprecedented availability of soil moisture information, new opportunities emerge to improve the accuracy in hydrologic predictions. Nonetheless, the effective use of this information entails the implementation of techniques

minimizes long-term prediction errors for a given historical data set. This type of algorithms assumes time-invariant parameters and thus does not make any attempt to include information from new observations (Moradkhani et al., 2005a; Liu and Gupta, 2007; Vrugt et al., 2013), thus hindering the identification of time-varying errors associated with the several components of the uncertainty. In addition, the applicability of such algorithms is limited in those regions where the lack of historical data makes the optimization procedure less practicable (Thiemann et al., 2001).

Therefore an increasing number of hydrological DA applications has been aimed at exploring the capabilities of DA methods to improve the accuracy of modelled physical phenomena and processes by recursively retrieving both states and parameters. An interesting discussion about pros and cons of the common approaches dealing with simultaneous state and parameter estimation in hydrological applications was provided by Liu and Gupta (2007). Todini et al. (1978) provided significant insights about the simultaneous state–parameter estimation in pioneering hydrological applications. Other important studies have been carried out by Vrugt et al. (2005), who combined the ensemble Kalman filter (EnKF) to update model states recursively with an optimization technique for parameter estimation. Moradkhani et al. (2005a, b) provided a framework for dual state–parameter estimation using the EnKF and particle filtering (PF).

One controvertible issue is which of the two approaches, namely the dual or the joint Kalman filtering, performs better for the simultaneous retrieval of states and parameters. In the dual filter approach, two separate filters – one for the state space and the other for the parameter space – are implemented. Instead, the joint filter methods predict the evolution of the joint probability distribution of states and parameters, which are combined in an augmented state vector. Theoretically, the joint approach is more robust because it accounts for the cross-covariance between the states and parameter estimates (van der Merwe, 2004). However, several studies highlighted the instability and intractability of this approach as a result of the complex interactions between states and parameters in a nonlinear dynamic system (Moradkhani et al., 2005; Liu and Gupta, 2007). Medina et al. (2014) have shown that the decoupled state-space representation in the dual KF provides higher flexibility since different Kalman filters can be selected for the states and parameters, respectively, according to the specific characteristics of both the observation and system equations.

A few studies have explored the possibility of simultaneously retrieving soil moisture profiles and soil hydraulic parameters, by assimilating surface soil moisture observations (e.g. Qin et al., 2009; Yang et al., 2009). Lü et al. (2011) applied a direct insertion method for assimilating surface soil moisture within the Richards equation, coupled with a particle swarm optimization algorithm, for identifying the optimal saturated hydraulic conductivity. Monzka et al. (2011) and Plaza et al. (2012) performed recursive state and parameter

retrieval for the soil moisture estimation, but using a particle filter algorithm, which is another sequential method widely applied in the most recent studies (Vrugt et al., 2013).

Medina et al. (2014) performed a synthetic numerical study to evaluate a dual Kalman filter (named DSUKF) for real-time simultaneous prediction of soil water content or matric pressure head profiles and soil hydraulic parameters, by assimilating near-surface information in a one-dimensional Richards equation. In this approach, a standard Kalman filter is implemented with a Crank–Nicolson numerical scheme to retrieve soil state profiles, while an unscented Kalman filter (UKF) (Julier and Uhlman, 1997, 2004; van der Merwe 2004) is implemented for retrieving soil hydraulic parameters. The UKF is based on a statistical linearization of the nonlinear operators (unscented transformation), without the need of performing any analytic differentiation.

One distinguishing issue of this approach concerns the adopted linearized forward state model. The noniterative integration arising from the linearized Crank–Nicolson formulation represents a simplification of the solution of the Richards equation. However, the time stepping adopted within the numerical scheme can be adjusted to reduce the numerical error to negligible values (e.g. Haverkamp et al., 1977). Paniconi et al. (1991) show that the linearization of the Richards equation, when applied to second-order time stepping formulations as the Crank–Nicolson, reduces the accuracy to first order. Nevertheless, further improvements can be easily incorporated, as for instance reported by Kavetski et al. (2002), who demonstrated that the reliability and the efficiency of these noniterative linearized schemes can be increased by adaptive truncation error control, and can be even more efficient than analogous time stepping schemes with iterative solvers.

Medina et al. (2014) provided valuable insights about the identifiability of the parameters featuring in the van Genuchten–Mualem soil hydraulic analytical relations (van Genuchten, 1980) by means of dual Kalman filters. This study also discussed the impact exerted by parameter initialization, observation depth, and assimilation frequency on the overall DSUKF retrieval performance, as well as the influence of the formulation of the Richards differential equation (either  $h$ -based or  $\theta$ -based form). The performance of the adopted approach is compared with that using the dual ensemble KF with ensemble sizes equal to 12 (DEnKF(12)) and 50 (DEnKF(50)), respectively. Reichle and Koster (2003) found that 12 ensembles consistently reduced the uncertainty of the soil moisture estimates during a one-dimensional EnKF state retrieval application. Camporese et al. (2009) suggest a minimum of about 50 realizations for ensuring a suitable level of accuracy in analogous applications. The results show that the root mean square error (RMSE) using DEnKF(50) is about 1.5 times lower, but at cost of seven times more computational central processing unit (CPU) time. However, DEnKF(12) is outperformed in

terms of both uncertainty reduction and computational efficiency.

The advantage of referring to synthetic values instead of using measured values of the variables is that the former are known a priori, and thus the assessment of the algorithm performance is facilitated. On the other hand, synthetic studies greatly simplify the inherent complexity of real-world applications, where uncertainties arise from several complexities of the system at hand, such as the soil heterogeneity and measuring errors. Nevertheless, one important question concerning the validity of the approach is whether it is able to cope with real data suitably.

This paper aims at evaluating the DSUKF performance by exploiting data measured during evaporation experiments carried out on undisturbed soil cores under laboratory conditions. The strength of the evaporation experiment is that the data are gathered during a transient flow that is very close to natural processes occurring in real soils, thereby providing a highly representative hydraulic response of the soil under study. The evaporation tests selected for this study are two of those employed by Romano and Santini (1999) for evaluating a parameter optimization method developed to determine the unsaturated hydraulic properties of different soil types. The method consists of a non-sequential inversion procedure, also implemented with a Crank–Nicolson numerical scheme, but using matric pressure head data instead of soil water content data. The evaporation experiments carried out by Romano and Santini (1999) have been selected mainly because soil water content values were also measured with a relatively high vertical resolution using the gamma ray attenuation method, thus providing a valuable experimental data set to test the DSUKF approach developed by Medina et al. (2014).

The performance of the DSUKF approach is herein examined, accounting for the effects of the observation depth, the assimilation frequency, and the parameter initialization, on both state and parameter retrieval processes.

## 2 The dual Kalman filter formulation

A detailed description of the algorithm employed for the separate state-space representation used to retrieve states and parameters is in the paper by Medina et al. (2014). At each time step, the current estimate of the parameters is used in the state filter, and the current estimate of the states is used in the parameter filter. For the sake of effectiveness, the state and parameter filter equations employed in the assimilation algorithm are herein summarized. In the most general case, the set of system equations can be written as

$$\mathbf{x}_k = F_{k-1,k}(\mathbf{x}_{k-1}, \mathbf{u}_k, \hat{\mathbf{w}}_{k-1}) + \mathbf{v}_{k-1} \quad (1)$$

$$y_k = H_k(\mathbf{x}_k, \hat{\mathbf{w}}_{k-1}) + \eta_k \quad (2)$$

for the state vector  $\mathbf{x}$  at instant  $k$ , and

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \xi_{k-1} \quad (3)$$

$$y_k = H_k(F_{k-1,k}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathbf{w}_k), \mathbf{w}_k) + \zeta_k \quad (4)$$

for the parameter vector  $\mathbf{w}$ . In the equation above,  $F$  is the state transition function,  $H$  the observation function,  $\mathbf{v}_{k-1}$  the zero mean Gaussian process noise with covariance  $\mathbf{Q}_{k-1}$ ,  $\eta_k$  the zero mean observation or measurement noise with covariance  $\mathbf{R}_k$ , and  $\mathbf{u}_k$  represents an exogenous input to the system. The parameter update is artificially set up by means of a stationary process with an identity state transition matrix.  $\xi_{k-1} \in N(0, \mathbf{Q}_{w,k-1})$  is the noise driving the parameter updating, and  $\zeta_k \in N(0, \mathbf{R}_{w,k})$  is the noise corrupting the observation equation relative to the parameters, both artificially settled. The caret symbol (“”) denotes the density mean of the variable.

### 2.1 The standard Kalman filter formulation for linear state retrieval

The linear algorithm for the state retrieval is summarized in the following three phases.

I. Initialization:

$$\hat{\mathbf{x}}_0 = E[\mathbf{x}_0] \quad (5)$$

$$\mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \quad (6)$$

$$\mathbf{Q}_0 = E[(\mathbf{v}_0 - \bar{\mathbf{v}}_0)(\mathbf{v}_0 - \bar{\mathbf{v}}_0)^T] \quad (7)$$

$$\mathbf{R}_0 = E[(\eta_0 - \bar{\eta}_0)(\eta_0 - \bar{\eta}_0)^T] \quad (8)$$

Subscript “0” indicates initial values.

II. Prediction phase is carried out by computing the state mean and covariances respectively as follows:

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_{k-1,k}(\hat{\mathbf{w}}_{k-1})\hat{\mathbf{x}}_{k-1} + \mathbf{g}_{k-1,k}(\mathbf{u}_k, \hat{\mathbf{w}}_{k-1}) \quad (9)$$

$$\mathbf{P}_k^- = \mathbf{F}_{k-1,k} \mathbf{P}_{k-1}, \mathbf{F}_{k-1,k}^T + \mathbf{Q}_k, \quad (10)$$

where subscript  $k$  indicates the time step;  $\hat{\mathbf{x}}_k^-$  and  $\mathbf{P}_k^-$  represent the a priori predictions of the state mean and covariance, respectively;  $\mathbf{F}_{k-1,k}$  is the linear state transition matrix of operator  $F$  in Eq. (1), which depends nonlinearly on the parameters  $\mathbf{w}_{k-1}$ ;  $\mathbf{g}_{k-1,k}$  is a vector accounting for the effect of the exogenous input  $\mathbf{u}_k$ , and it is also nonlinearly dependent on the parameters  $\mathbf{w}_{k-1}$ .

III. Correction phase is carried out for updating estimates with the last observation:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \left( \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1}, \quad (11)$$

$$\hat{\mathbf{x}}_k = \left( \hat{\mathbf{x}}_k^- \right) + \mathbf{K}_k \left( \mathbf{y}_k - \mathbf{H}_k \left( \hat{\mathbf{x}}_k^- \right) \right), \quad (12)$$

$$\mathbf{P}_k = \mathbf{P}_k^- \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^-, \quad (13)$$

where  $\mathbf{K}_k$  is the Kalman gain, expressing the ratio of the expected cross-covariance matrix of the process prediction error and the observation prediction error,  $\mathbf{y}_k$  represents the actual measurements, and  $\hat{\mathbf{x}}_k$  and  $\mathbf{P}_k$  represent the states posterior density mean and covariance, respectively.  $\mathbf{H}_k$  stands for the linear observation operator (Eq. 2) in matrix form.

## 2.2 The unscented Kalman filter (UKF) formulation for parameter estimation

In the UKF formulation, the distribution of the parameters is represented by a Gaussian random variable, being specified using a minimal set of sample points, so-called sigma points, which are deterministically selected to capture the true mean and covariance of the variable completely and, when propagated through the true nonlinear system, to capture the posterior mean and covariance accurately up to the second order for any nonlinearity (van Der Merwe, 2004; Chirico et al., 2014; Medina et al., 2014).

The algorithm for the dynamic retrieval of the unknown parameters can be summarized in the following four phases.

I. Initialization of the parameter vector,  $\hat{\mathbf{w}}_0$ , and covariance,  $\mathbf{P}_{w,0}$ :

$$\hat{\mathbf{w}}_0 = E[\mathbf{w}] \quad (14)$$

$$\mathbf{P}_{w,0} = E \left[ (\mathbf{w} - \hat{\mathbf{w}}_0) (\mathbf{w} - \hat{\mathbf{w}}_0)^T \right] \quad (15)$$

II. Time update phase, in order to get the corresponding a priori predictions:

$$\hat{\mathbf{w}}_k^- = \hat{\mathbf{w}}_{k-1} \quad (16)$$

$$\mathbf{P}_{w,k}^- = \mathbf{P}_{w,k-1} + \mathbf{Q}_{w,k-1}, \quad (17)$$

where  $\mathbf{Q}_{w,k} = \left( \lambda_{\text{RLS}}^{-1} - 1 \right) \mathbf{P}_{w,k}$ , being  $\lambda_{\text{RLS}} \in (0, 1]$  a forgetting factor as defined in the recursive least-squares (RLS) algorithm. This relationship for  $\mathbf{Q}_{w,k}$  is chosen on the basis of the results illustrated in Medina et al. (2014), where alternative expressions have been compared.

III. Computation of the sigma points  $\mathcal{W}_{k,i}$  for the measuring update:

$$\mathcal{W}_{k,i} = \left[ \hat{\mathbf{w}}_k^-, i=0 \quad \hat{\mathbf{w}}_k^- + \left( \sqrt{\gamma \mathbf{P}_{w,k}^-} \right)_i, i=1, \dots, N_{\text{par}} \quad \hat{\mathbf{w}}_k^- - \left( \sqrt{\gamma \mathbf{P}_{w,k}^-} \right)_i, i=N_{\text{par}}+1, \dots, 2N_{\text{par}} \right], \quad (18)$$

where  $N_{\text{par}}$  is equal to the number of parameters to be retrieved, while  $\gamma$  is a coefficient scaling the sampled parameter distribution around the mean parameter vector.

IV. Measuring update equations:

$$\mathcal{Y}_k = H_k \left( F_{k-1,k} \left( \hat{\mathbf{x}}_{k-1}, \mathbf{u}_k, \mathcal{W}_k \right) \right) \quad (19)$$

Then mean,  $\hat{\mathbf{y}}_{w,k}^-$ , cross-covariance,  $\mathbf{P}_{wy,k}$ , and covariance,  $\mathbf{P}_{yy,k}^w$ , of the transformed sigma points  $\mathcal{Y}_k$  are respectively calculated as follows:

$$\hat{\mathbf{y}}_{w,k}^- = \sum_{i=0}^{2N_{\text{par}}} \mu_i^{(m)} \mathcal{Y}_{k,i}, \quad (20)$$

$$\mathbf{P}_{wy,k} = \sum_{i=0}^{2N_{\text{par}}} \mu_i^{(c)} \left( \mathcal{W}_{k,i} - \hat{\mathbf{w}}_k^- \right) \left( \mathcal{Y}_{k,i} - \hat{\mathbf{y}}_{w,k}^- \right)^T, \quad (21)$$

$$\mathbf{P}_{yy,k}^w = \sum_{i=0}^{2N_{\text{par}}} \mu_i^{(c)} \left( \mathcal{Y}_{k,i} - \hat{\mathbf{y}}_{w,k}^- \right) \left( \mathcal{Y}_{k,i} - \hat{\mathbf{y}}_{w,k}^- \right)^T, \quad (22)$$

where  $\mu_i$  are the weights related to the sigma point  $i$ , conditioned to  $\sum_{i=0}^{2N_{\text{par}}} \mu_i = 1$ . Weight values for calculating the mean and the covariance are indicated by the superscripts  $m$  and  $c$ , respectively.

The Kalman gain,  $\mathbf{K}_{w,k}$ , and the posterior parameter mean,  $\hat{\mathbf{w}}_k$ , and covariance,  $\mathbf{P}_{w,k}$ , are calculated respectively as follows:

$$\mathbf{K}_{w,k} = \mathbf{P}_{wy,k} \left( \mathbf{P}_{yy,k}^w + \mathbf{R}_{w,k} \right)^{-1} = \mathbf{P}_{wy,k} \left( \mathbf{P}_{\nu\nu,k}^w \right)^{-1}, \quad (23)$$

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_k^- + \mathbf{K}_{w,k} \left( \mathbf{y}_k - \hat{\mathbf{y}}_{w,k}^- \right), \quad (24)$$

$$\mathbf{P}_{w,k} = \mathbf{P}_{w,k}^- - \mathbf{K}_{w,k} \mathbf{P}_{\nu\nu,k}^w \left( \mathbf{K}_{w,k} \right)^T, \quad (25)$$

where  $\mathbf{P}_{\nu\nu,k}^w$  represents the covariance of  $\mathbf{y}_k - \hat{\mathbf{y}}_{w,k}^-$ . More details about the UKF implementation can be found in Medina et al. (2014).

## 3 Materials and methods

### 3.1 Soil water flow governing equation

Water movement in a vertical soil column, modelled as a homogeneous, variably saturated porous medium under isothermal conditions, is simulated using the  $\theta$ -based form of the Richards equation:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[ \left( D(\theta) \frac{\partial \theta}{\partial z} - K(\theta) \right) \right], \quad (26)$$

where  $t$  is time,  $z$  soil depth taken positive downward, with  $z = 0$  at the top of the profile,  $\theta$  the soil water content [ $L^3 L^{-3}$ ],  $D(\theta) = K(\theta) \frac{dh}{d\theta}$  [ $L^2/T$ ] is the unsaturated diffusivity, with  $K$  [ $L/T$ ] being the unsaturated hydraulic conductivity and  $h$  the soil water matric pressure head [ $L$ ].

The constitutive relationships characterizing the soil hydraulic properties are the van Genuchten–Mualem (VGM) parametric relations (van Genuchten, 1980):

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

$$K(\theta) = K_s S_e^\lambda \left[ 1 - \left( 1 - S_e^{1/m} \right)^m \right]^2, \quad (28)$$

where  $\theta_s$  is the saturated soil water content,  $\theta_r$  the residual soil water content,  $S_e = (\theta - \theta_r) / (\theta_s - \theta_r)$  the effective saturation, and  $K_s$  the saturated hydraulic conductivity, whereas  $\alpha$  [ $L^{-1}$ ],  $n$  (-),  $m$  (-) and  $\lambda$  (-) are empirical parameters. A common assumption, also adopted in this work, is to consider  $\lambda = 0.5$  and  $m = 1 - 1/n$ .

### 3.2 Crank–Nicolson finite difference scheme

The numerical solution of the  $\theta$ -based form of the Richards equation (Eq. 26) is implemented according to the Crank–Nicolson finite difference scheme, with an explicit linearization of both the soil hydraulic conductivity  $K$  and the diffusivity  $D$ , which takes on the following form for the intermediate nodes of the soil column:

$$\begin{aligned} & \left( -\frac{D_k^{i-1/2}}{2\Delta z^i \Delta z^u}; \frac{1}{\Delta t_k} + \frac{\frac{D_k^{i-1/2}}{\Delta z^u} + \frac{D_k^{i+1/2}}{\Delta z^l}}{2\Delta z^i}; -\frac{D_k^{i+1/2}}{2\Delta z^i \Delta z^l} \right) \begin{pmatrix} \theta_k^{i-1} \\ \theta_k^{i+1} \\ \theta_{k+1}^{i+1} \end{pmatrix} \\ & = \left( \frac{D_k^{i-1/2}}{2\Delta z^i \Delta z^u}; \frac{1}{\Delta t_k} - \frac{\frac{D_k^{i-1/2}}{\Delta z^u} + \frac{D_k^{i+1/2}}{\Delta z^l}}{2\Delta z^i}; \frac{D_k^{i+1/2}}{2\Delta z^i \Delta z^l} \right) \\ & \quad \begin{pmatrix} \theta_k^{i-1} \\ \theta_k^i \\ \theta_{k+1}^i \end{pmatrix} + \frac{K_k^{i-1} - K_k^{i+1}}{2\Delta z^i}, \end{aligned} \quad (29)$$

where subscript  $i$  is the node number (increasing downward), superscript  $k$  the time step, and  $\Delta t^k = t^{k+1} - t^k$ . All the nodes, including the top and bottom nodes, are located in the centre of each soil compartment in which the soil column is discretized, with  $\Delta z^u = z^i - z^{i-1}$ ,  $\Delta z^l = z^{i+1} - z^i$  and  $\Delta z^i$  the compartment thickness (cm). The spatial averages of  $K$  are calculated as arithmetic means.

Flux conditions imposed at the upper and lower boundaries are expressed, respectively, by the following equations:

$$\begin{aligned} & \left( \frac{1}{\Delta t_k} + \frac{D_k^{1+1/2}}{2\Delta z^1 \Delta z^l}; -\frac{D_k^{1+1/2}}{2\Delta z^1 \Delta z^l} \right) \begin{pmatrix} \theta_{k+1}^1 \\ \theta_{k+1}^2 \end{pmatrix} \\ & = \left( \frac{1}{\Delta t_k} - \frac{D_{1+1/2}^k}{2\Delta z^1 \Delta z^l}; \frac{D_{1+1/2}^k}{2\Delta z^1 \Delta z^l} \right) \begin{pmatrix} \theta_k^1 \\ \theta_k^2 \end{pmatrix} + \frac{q_{\text{top}} - K_k^{1+1/2}}{\Delta z^1} \end{aligned} \quad (30)$$

$$\begin{aligned} & \left( -\frac{D_k^{n-1/2}}{2\Delta z^n \Delta z^u}; \frac{1}{\Delta t_k} + \frac{D_k^{n-1/2}}{2\Delta z^n \Delta z^u} \right) \begin{pmatrix} \theta_{k+1}^{n-1} \\ \theta_{k+1}^n \end{pmatrix} \\ & = \left( \frac{D_k^{n-1/2}}{2\Delta z^n \Delta z^u}; \frac{1}{\Delta t_k} - \frac{D_k^{n-1/2}}{2\Delta z^n \Delta z^u} \right) \begin{pmatrix} \theta_k^{n-1} \\ \theta_k^n \end{pmatrix} + \frac{K_k^{n-1/2} - q_{\text{bot}}}{\Delta z^n} \end{aligned} \quad (31)$$

being  $q_{\text{top}}$  and  $q_{\text{bot}}$  the fluxes at the top and the bottom of the soil profile, respectively.

In the numerical scheme, an explicit linearization of  $K$  and  $D$  is implemented by taking their values at the previous time step  $k-1$ . Then, a linear state-space representation of the dynamic system can be easily derived by combining the set of Eqs. (29)–(31) written for each node and accounting for the boundary conditions:

$$\mathbf{B}_{k-1} \theta_k = \mathbf{A}_{k-1} \theta_{k-1} + \mathbf{f}_{k-1}, \quad (32)$$

where  $\mathbf{A}_{k-1}$  and  $\mathbf{B}_{k-1}$  are tri-diagonal matrices obtained by assembling the terms in the first pair of parentheses on the right- and left-hand side of Eqs. (29)–(31), respectively. The term  $\mathbf{f}_{k-1}$  is a vector obtained by assembling the terms on the right-hand side of the state variable at time step  $k-1$ . More explicitly, Eq. (32) becomes

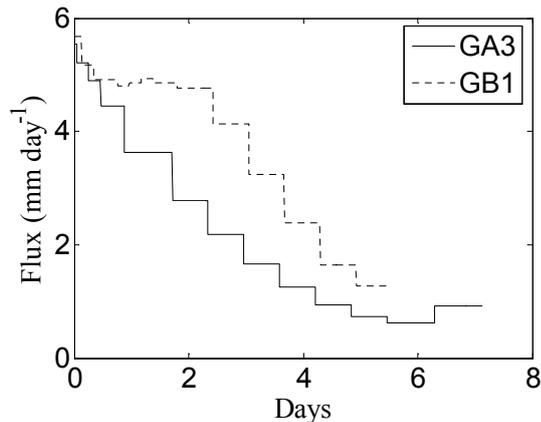
$$\theta_k = \mathbf{F}_{k-1} \theta_{k-1} + \mathbf{g}_{k-1} \quad (33)$$

by making  $\mathbf{F} \mathbf{B}^{-1} \mathbf{A}$  and  $\mathbf{g} = \mathbf{B}^{-1} \mathbf{f}$ , which correspond to the analogous terms in Eqs. (9)–(10).

### 3.3 Experimental data set

We used data published on an earlier study since it was judged to be helpful to explore advanced methods about a common problem already mentioned in the literature. We used the data reported in the paper by Romano and Santini (1999) and collected during evaporation tests on the two undisturbed soil core named as GA3 and GB1. Each of these two soil cores had an inner diameter of 8.0 cm and a length of 12.0 cm. In order to provide a clear view of the soil water dynamic process, against which the performance of DSUKF has been evaluated, a brief description of the evaporation tests is herein provided, while further details can be found in Romano and Santini (1999).

The undisturbed soil core, after being completely saturated from the bottom, is induced to a state of hydrostatic equilibrium with the matric pressure head value at the bottom end almost at zero. The sample cylinder is then completely sealed at the bottom and positioned on a plate, supported by a strain-gauge load cell measuring the soil sample weight, while a small fan is positioned near the top. Tensiometers connected to pressure transducers are inserted at various depths to monitor the soil water pressure head. The evaporation experiment is carried out until the formation of air bubbles causes the breakdown of the hydraulic connection between the last working tensiometer and the corresponding pressure transducer. Tensiometers were inserted at the following three soil



**Figure 1.** Evaporation flux at the soil surface calculated from the water balance of each soil sample between consecutive measurements of the soil profiles.

depths: 3, 6, and 9 cm. Additionally, soil water content profiles during the experiment were measured with a gamma ray attenuation device, with a vertical resolution of 1.0 cm.

Table 1 lists the basic physical properties of the two soil samples together with the VGM model parameters  $\alpha$ ,  $n$  and  $K_s$ , which have been estimated by Romano and Santini (1999) by applying a non-sequential parameter estimation method and are employed as reference values in this study.

Soil water content profiles, with a time update of 600 s, have been built by polynomial interpolation of the gamma ray measurements at all soil depths and taken as “true” state values.

The assimilation algorithm samples the observation values from the interpolated soil water content profiles. The evaporation rate at the soil surface is estimated by applying a water balance equation between two consecutive measurements of the soil water content profiles, under the assumption of a constant evaporation flux during the measurement interval. This approach provides an approximate temporal pattern of the upper boundary fluxes, with step changes, as depicted in Fig. 1.

The duration of the assimilated evaporation process is 170 h for GA3 and 131 h for GA1, which approximately correspond to 7 and 5.5 days, respectively.

#### 4 DSUKF implementation

Similarly to the synthetic study of Medina et al. (2014), the saturated ( $\theta_s$ ) and the residual ( $\theta_r$ ) soil water contents are assumed to be known, as these parameters can be easily determined by direct or indirect methods (e.g. Pringle et al., 2007; Chirico et al., 2010). As reported in the paper by Romano and Santini (1999), the values of parameter  $\theta_s$  are fixed to  $0.31 \text{ cm}^3 \text{ cm}^{-3}$  for GA3 and  $0.35 \text{ cm}^3 \text{ cm}^{-3}$  for GB1, re-

spectively. The values of parameter  $\theta_r$  are instead defined according to the values suggested by Carsel and Parrish (1988) for soils of the same textural class:  $\theta_r = 0.067 \text{ cm}^3 \text{ cm}^{-3}$  for a silt loam soil as GA3, and  $\theta_r = 0.08 \text{ cm}^3 \text{ cm}^{-3}$  for a loam soil as GB1, which are both slightly smaller than the air-dried values assumed by Romano and Santini (1999) (see Table 1).

The unscented Kalman filter is thus implemented to retrieve the remaining parameters  $K_s$ ,  $\alpha$  and  $n$ . As shown in Medina et al. (2014), a variable transformation is applied to constrain the retrieved parameter values to a certain physically meaningful range and thus to guarantee operational stability. Considering that  $w_i$  is the true value of the  $i$ th parameter, the parameter estimation procedure makes use of the following variable transformation:

$$w_i = w_{i_{\min}} + (w_{i_{\max}} - w_{i_{\min}})s(\delta w_i), \quad (34)$$

where  $w_{\min}$  and  $w_{\max}$  represent user-defined nominal values constraining the minimum and maximum values of the parameter, respectively; the correction terms  $\delta w_i$  are the actual variables under estimation and are expressed as independent terms of a nonlinear sigmoidal function  $s(\delta w)$ . The sigmoidal function, designed to limit the absolute magnitude of the estimated adjustment, is defined as follows:

$$s(\delta w_i) = \frac{\delta w_i}{2(1 + |\delta w_i|)} + 0.5. \quad (35)$$

Table 2 summarizes the initial conditions employed for the state variables and the covariance matrices, as well as the examined observation depths and assimilation frequencies.

Uniform profiles have been assumed as initial state condition, with values  $\theta_0 = 0.28 \text{ cm}^3 \text{ cm}^{-3}$  for soil core GA3 and  $\theta_0 = 0.31 \text{ cm}^3 \text{ cm}^{-3}$  for soil core GB1. These are approximately average values between saturation and the soil water content measured at the soil surface at the beginning of the experiments.

Our evaluations considered three different observation depths (OD): 1, 2, and 12 cm. Escorihuela et al. (2010) found that an OD of 2 cm is the most effective soil moisture sampling depth for L-band radiometry, even in comparison with smaller depths. The observation depth of OD = 12 cm, which means the assimilation of the entire soil core, has been included as benchmark performance for OD = 1 cm and OD = 2 cm in the parameter identification. Two assimilation frequencies have been used: a finer frequency with assimilations every 2 h ( $AF = 1/2 \text{ h}^{-1}$ ), and a coarser one with assimilations every 12 h ( $AF = 1/12 \text{ h}^{-1}$ ).

The initial covariance matrices are all diagonals and are defined similarly to Medina et al. (2014): the initial state covariance matrix is set to a value of 1000 % of the mean state profile on the diagonal elements; the initial matrix of the normalized correction terms associated with the soil hydraulic parameters is assigned to a value of 0.01 on the diagonal elements; the diagonal of the observation noise autocovariance matrix is updated using the 2 % of the observed state vector,

**Table 1.** Physical and soil hydraulic properties of the two soil samples employed for the evaporation experiments.

Soil Sample	Texture	$\rho_b$ , g cm <sup>-3</sup>	$\theta_s$	$\theta_r^*$	$\alpha^{**}$ (10 <sup>-2</sup> cm <sup>-1</sup> )	$n^{**}$	$K_s^{**}$ (10 <sup>-4</sup> cm s <sup>-1</sup> )
GA3	Silty loam	1.592	0.310	0.080	1.75	2.27	0.222
GB1	Loam	1.572	0.348	0.120	1.67	2.70	1.56 5

\* Air-dried value. \*\* Estimated by inversion method (Romano and Santini, 1999).

**Table 2.** Values adopted for the initialization and implementation of the retrieval algorithm.

Input variable	Soil sample	
	GA3	GB1
Initial state variable	0.28 cm <sup>3</sup> cm <sup>-3</sup>	0.31 cm <sup>3</sup> cm <sup>-3</sup>
Observation depths (ODs)	1, 2 and 12 cm	
Assimilation frequency (AF)	2 and 12 h	
<i>State covariance matrices</i>		
Initial state covariance matrix $P_0^{i,i}; i = 1 \dots N_{\text{nod}}$	0.8 cm <sup>6</sup> cm <sup>-6</sup>	
Process noise updating $Q_0^{i,i}; i = 1 \dots N_{\text{nod}}$	0.05 $x_i$ cm <sup>6</sup> cm <sup>-6</sup>	
Observation noise updating $R_0^{i,i}; i = 1 \dots N_{\text{obs}}$	0.02 $y_i$ cm <sup>6</sup> cm <sup>-6</sup>	
<i>Parameter covariance matrices</i>		
Initial normalized correction terms matrix $P_{w,0}^{i,i}; i = 1 \dots N_{\text{par}}$	0.01	
Forgetting factor $\lambda_{\text{RLS}}$ (linked to $Q_w$ , Eq. 17)	0.9999	
Artificial noise covariance $R_w^{i,i}; i = 1 \dots N_{\text{obs}}$	1.0 × 10 <sup>-5</sup>	

$N_{\text{nod}}$  is the number of nodes (states);  $N_{\text{par}}$  is the number of parameters under scrutiny;  $N_{\text{obs}}$  is the number of observations;  $x$  and  $y$  represent the state and the observation vectors, respectively.

while the system noise covariance is updated assuming the 5 % of the profile state vector.

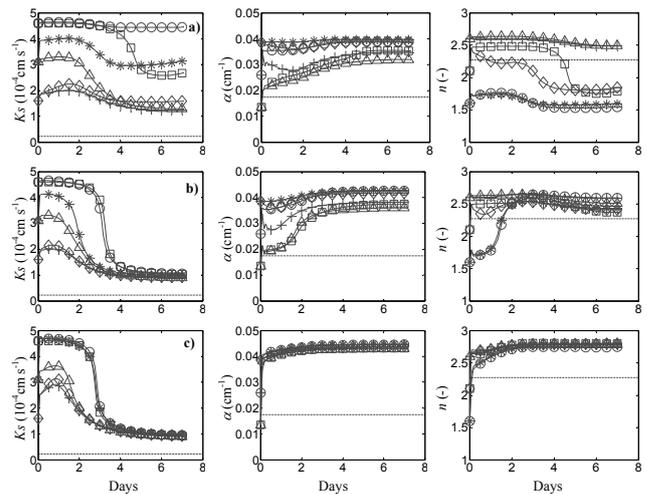
The limiting and initialization values of parameters  $K_s$ ,  $\alpha$  and  $n$  are also identical to those employed by Medina et al. (2014). Table 3 summarizes the minimum value ( $w_{\text{min}}$ ) and the prescribed range ( $w_{\text{max}} - w_{\text{min}}$ ) of each parameter, as well the resulting values of the sets of initial parameters. The six sets of initial parameters are employed for evaluating the influence of the initial condition on the performance of the parameter retrieval algorithm.

For quantitatively evaluating the performance of the involved schemes, the mean error (ME) and the root mean square error (RMSE) between retrieved and true state profiles are computed as follows:

$$ME_j = \frac{1}{N} \sum_{i=1}^N (x_{i,j}^{\text{guess}} - x_{i,j}^{\text{true}}), \quad (36)$$

$$RMSE_j = \left[ \frac{1}{N-1} \sum_{i=1}^N (x_{i,j}^{\text{guess}} - x_{i,j}^{\text{true}})^2 \right]^{1/2}, \quad (37)$$

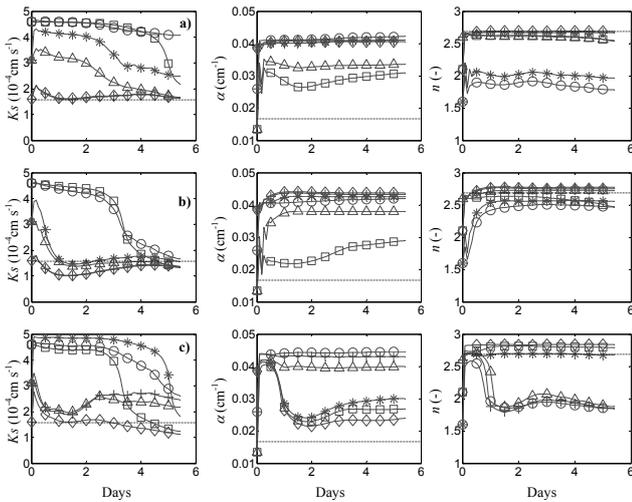
where  $x_{i,j}^{\text{guess}}$  and  $x_{i,j}^{\text{true}}$  represent the guess and true state value at node  $i$  and time  $j$ , respectively.



**Figure 2.** Retrieved VGM parameters  $K_s$ ,  $\alpha$  and  $n$  for GA3 soil sample with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$ , and observation depth (a) OD = 1 cm, (b) OD = 2 cm and (c) OD = 12 cm. Comparisons account for the six pondered sets of initial parameters S1(O), S2(S), S3(\*), S4(Delta), S5(+), and S6(Diamond). The dotted line indicates the value of the parameter found by Romano and Santini (1999).

**Table 3.** Values of  $w_{\min}$  and  $w_{\max}-w_{\min}$  used to constrain the distribution of the parameters  $K_s$ ,  $\alpha$  and  $n$ , and the resulting six sets of parameter values employed for initializing the assimilation algorithm.

Parameter	$w_{\min}$	$w_{\max}-w_{\min}$	S1	S2	S3	S4	S5	S6
$K_s$ ( $\text{cm s}^{-1}$ )	$1 \times 10^{-5}$	$6 \times 10^{-4}$	$4.6 \times 10^{-4}$	$4.6 \times 10^{-4}$	$3.1 \times 10^{-4}$	$3.1 \times 10^{-4}$	$1.6 \times 10^{-4}$	$1.6 \times 10^{-4}$
$\alpha$ ( $\text{cm}^{-1}$ )	$1 \times 10^{-3}$	$5 \times 10^{-2}$	$2.6 \times 10^{-2}$	$1.35 \times 10^{-2}$	$3.85 \times 10^{-2}$	$1.35 \times 10^{-2}$	$2.6 \times 10^{-2}$	$3.85 \times 10^{-2}$
$n$ (-)	1.1	2.0	1.6	2.1	1.6	2.6	2.6	2.1



**Figure 3.** Retrieved VGM parameters  $K_s$ ,  $\alpha$  and  $n$  for GB1 soil sample with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$ , and observation depth (a)  $OD = 1 \text{ cm}$ , (b)  $OD = 2 \text{ cm}$  and (c)  $OD = 12 \text{ cm}$ . Comparisons account for the six pondered sets of initial parameters S1( $\circ$ ), S2( $\square$ ), S3( $*$ ), S4( $\Delta$ ), S5( $+$ ) and S6( $\diamond$ ). The dotted line indicates the value of the parameter found by Romano and Santini (1999).

## 5 Results

### 5.1 Parameter retrieval

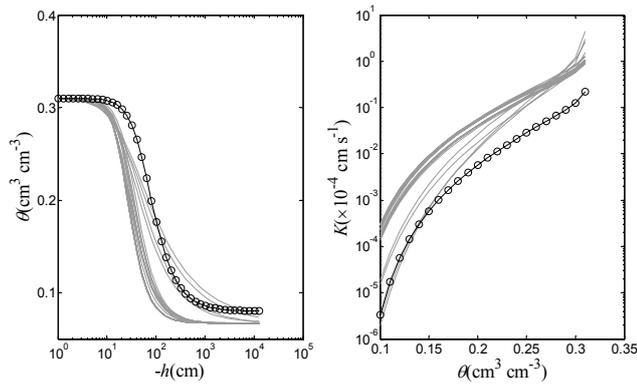
Figures 2 and 3 show the temporal patterns of the retrieved VGM parameters for GA3 and GB1, respectively, obtained by assimilating the observed soil water content every 2 h within the three different observation depths (ODs): 1, 2, and 12 cm. For soil core GA3 (Fig. 2), there is a very good agreement between parameter values retrieved using  $OD = 2 \text{ cm}$  (Fig. 2b) and  $OD = 12 \text{ cm}$  (Fig. 2c). It is particularly interesting to note the consistency in the convergence of  $K_s$ , which in general is the less identifiable parameter (Medina et al., 2014). The final retrieved  $K_s$  value is approximately  $1 \times 10^{-4} \text{ cm s}^{-1}$ , higher than both the value estimated by Romano and Santini (1999) for the VGM hydraulic model ( $0.222 \times 10^{-4} \text{ cm s}^{-1}$ ) and the value directly estimated with the falling head method ( $0.345 \times 10^{-4} \text{ cm s}^{-1}$ ). However, the final retrieved  $K_s$  is in the range of values obtained by Romano and Santini (1999) with analytical models of the soil hydraulic properties other than the VGM. When us-

ing  $OD = 1 \text{ cm}$ , parameter  $n$  converges to two main values: one approximately 2.5 and the other one approximately 1.7 (Fig. 2a). For  $OD = 2 \text{ cm}$  and  $OD = 12 \text{ cm}$ , the convergence patterns are less dependent on the initial set of parameter values. For  $OD = 2 \text{ cm}$ ,  $n$  converges to a value approximately of 2.5, while for  $OD = 12 \text{ cm}$  it converges to 2.7, thus in both cases higher than the reference value of 2.27.

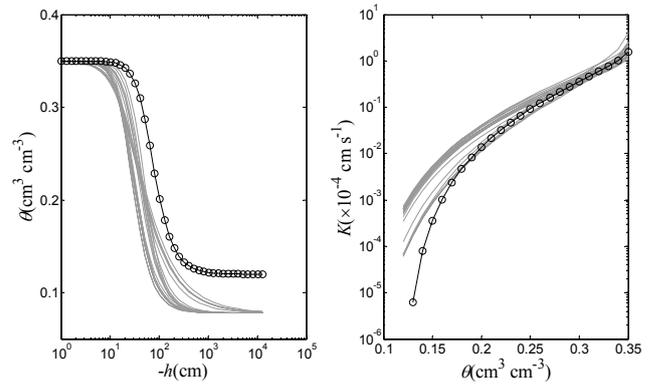
For soil core GB1 (Fig. 3), the retrieval process provides values of  $K_s$  and  $n$  that converge toward those identified by Romano and Santini (1999) ( $K_s = 1.565 \times 10^{-4} \text{ cm s}^{-1}$  and  $n = 2.7$ , respectively). Using  $OD = 2 \text{ cm}$ , the means of  $K_s$  and  $n$  are found practically equal to those reported by Romano and Santini (1999). The parameters retrieved by assimilating the entire profile ( $OD = 12 \text{ cm}$ , Fig. 3c) follow two distinct patterns. The parameter values retrieved with the initial sets S2, S3, and S6 follow patterns fairly close to the values found by Romano and Santini (1999), with mean  $K_s$ ,  $\alpha$ , and  $n$  of approximately  $1.5 \times 10^{-4} \text{ cm s}^{-1}$ ,  $0.028 \text{ cm}^{-1}$  and 2.8, respectively. The parameter values retrieved with the initial sets S1, S4, and S5 follow patterns with average values  $K_s = 2.4 \times 10^{-4} \text{ cm s}^{-1}$ ,  $\alpha = 0.043 \text{ cm}^{-1}$  and  $n = 1.8$ , which are relatively close to the values reported by Carsel and Parrish (1988) for loam soils:  $K_s = 5.0 \times 10^{-4} \text{ cm s}^{-1}$ ,  $\alpha = 0.036 \text{ cm}^{-1}$  and  $n = 1.56$ . The convergence patterns obtained with  $OD = 1 \text{ cm}$  and  $2 \text{ cm}$  (Fig. 3a and b, respectively) reflect rather well these two alternative parameter space solutions.

Compared with  $K_s$  and  $n$ , parameter  $\alpha$  is much less affected by the observation depth and the initial parameterization. Preliminary sensitivity analyses (not presented here for the sake of brevity) have highlighted that higher initial soil water content values favour the parameter identifiability and the increase of the convergence rate of  $\alpha$ , as a result of a relatively higher amount of information of  $\alpha$  retrievable for soil water states close to the air entry value (Vrugt et al., 2001, 2002; Medina et al., 2014). In both experiments GA3 and GB1, parameter  $\alpha$  converges predominantly toward a value of approximately  $0.04 \text{ cm}^{-1}$ , which is notably higher than those estimated by Romano and Santini (1999).

This relative inconsistency can be justified considering that the assimilation algorithm is implemented by exploiting the soil water content as an observation variable, whilst Romano and Santini (1999) employed pressure head values measured at three depths (3.0, 6.0, and 9.0 cm) to estimate



**Figure 4.** Comparison of the 18 (corresponding to 3 observation depths times 6 initial parameter sets) soil water retention curves  $\theta(h)$ , and hydraulic conductivity functions,  $K(\theta)$ , using the converging parameters for GA3 (grey solid lines). The solid lines with markers indicate the corresponding functions defined with the parameters found by Romano and Santini (1999). Note that in this study  $\theta_r = 0.067 \text{ cm}^3 \text{ cm}^{-3}$ , while Romano and Santini (1999) assumed  $\theta_r = 0.08 \text{ cm}^3 \text{ cm}^{-3}$ .

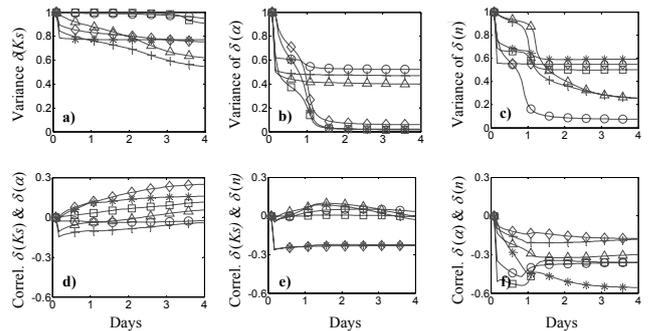


**Figure 5.** Comparison of the 18 (corresponding to 3 observation depths times 6 initial parameter sets) soil water retention curves  $\theta(h)$ , and hydraulic conductivity functions,  $K(\theta)$ , using the converging parameters for GB1 (grey solid lines). The solid lines with markers indicate the corresponding functions defined with the parameters found by Romano and Santini (1999). Note that in this study  $\theta_s = 0.35 \text{ cm}^3 \text{ cm}^{-3}$  and  $\theta_r = 0.08 \text{ cm}^3 \text{ cm}^{-3}$ , while Romano and Santini (1999) assumed  $\theta_s = 0.348 \text{ cm}^3 \text{ cm}^{-3}$  and  $\theta_r = 0.12 \text{ cm}^3 \text{ cm}^{-3}$ .

the soil hydraulic parameters with a non-sequential inverse method. Indeed, parameter  $\alpha$  acts as a scaling factor of the pressure head values with respect to the soil moistures in the VGM model, and its identifiability through an inverse method can be highly affected by the type of information employed (e.g. Šimůnek and van Genuchten, 1996; Ritter et al., 2004; Wöhling and Vrugt, 2011). As discussed later, the observed discrepancies can also be partly attributed to some inaccuracy in the identification of the final solution due to the high correlation between the van Genuchten parameters. The narrow range covered by the state variables in the considered experiment can have a negative impact on the final results. Several authors have emphasized some difficulties in the identification of the VGM parameters as a consequence of the narrow variability of naturally occurring boundary conditions (Scharnagl et al., 2011; Vrugt et al., 2001, 2002).

With reference to soil core GA3, Fig. 4 shows the comparisons between the soil water retention and hydraulic conductivity functions obtained using the 18 retrieved parameter vectors, and those functions described by the optimized parameter values of Romano and Santini (1999). These vectors are obtained through 18 combinations of 3 different observation depths and 6 initial parameter sets. The denser groups of water retention and hydraulic conductivity curves, corresponding to the solutions using  $OD = 2 \text{ cm}$  and  $12 \text{ cm}$ , have a slope similar to the corresponding reference curves of Romano and Santini (1999), but are shifted toward higher values of  $h$  and  $K$ . The groups of curves obtained using  $OD = 1 \text{ cm}$  have a different slope, and they match fairly well the reference curves only in the dry range.

Similarly to Fig. 4, Fig. 5 compares the soil hydraulic functions obtained in this study for soil core GB1 and those



**Figure 6.** Evolving variances of the correction terms (a)  $\delta(K_s)$ , (b)  $\delta(\alpha)$  and (c)  $\delta(n)$  associated with the VGM parameters and correlations (d–e) between these terms during the first 4 days using the GA3 data series, with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$  and observation depth  $OD = 12 \text{ cm}$ .

optimized by Romano and Santini (1999). Again, our estimated soil water retention curves are shifted with respect to the reference curve, except in the dry range of the graph because of the difference between the residual soil moisture of the cited study ( $0.12 \text{ cm}^3 \text{ cm}^{-3}$ ) and the value assumed in this study ( $0.078 \text{ cm}^3 \text{ cm}^{-3}$ ). The estimated hydraulic conductivities match very well the reference curve in the wet range, but depart in the dry one, also as result of the difference in the residual soil water contents.

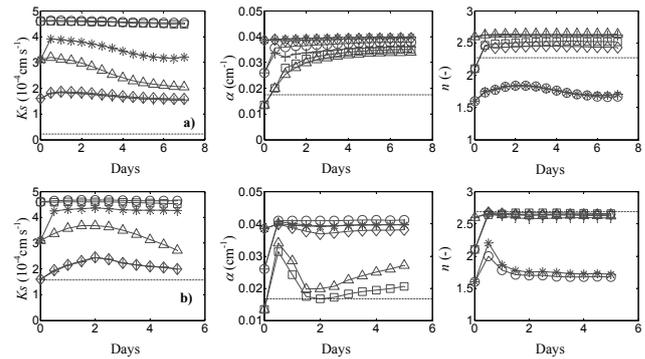
Some difficulty in correctly identifying the solution is definitely related to the multivariate correlation structure induced in the parametric distribution. This question is evident from Fig. 6, which illustrates the behaviour of the evolving variance and correlation of the correction terms associated with the estimated parameters for GA3 (see also Eqs. 34 and

35) assuming  $OD = 12$  cm. This is roughly the pattern observed for both experiments, independently from the adopted observation depths. The temporal reduction of  $\delta(K_S)$  variance (Fig. 6a) is small compared with that of  $\delta(\alpha)$  and  $\delta(n)$  (Fig. 6b and c). Moreover,  $\delta(K_S)$  predominantly exhibits a positive correlation with  $\delta(\alpha)$  and a negative correlation with  $\delta(n)$  (Fig. 6d and e), while the correlation is always negative between  $\delta(\alpha)$  and  $\delta(n)$ . The signs of these correlation values are consistent with those found by Romano and Santini (1999), although they examined the actual parameter values rather than a nonlinear transformation of them, as done in this study. It is also important to point out that the correlation structure influences the retrieved parameters in different ways, depending on the initial conditions and on the type of retrieval algorithm employed, which can be either sequential, such as in the present study, or non-sequential such as that employed by Romano and Santini (1999).

Finally, a closer inspection of the performance of the assimilation algorithm with relatively low assimilation frequencies is obtained from Fig. 7, which depicts the temporal patterns of the retrieved parameters using  $AF = 1/12 \text{ h}^{-1}$  and  $OD = 1$  cm, thus involving only 15 assimilation events for GA3 and 11 for GB1. Nevertheless, there is a good agreement between these patterns and the analogous patterns using  $AF = 1/2 \text{ h}^{-1}$  (Figs. 2a and 3a). Also the covariance and correlation structure follow the general trends previously described and the effect of the negative correlation between  $\alpha$  and  $n$  can be visually perceived. For example, it is worth noting that the lowest convergent value of  $n$  in Fig. 7b, obtained with the parameter initializations S1 and S3, corresponds to the highest convergent value of  $\alpha$ .

In summary, all these outcomes illustrate the performance of the proposed approach in terms of parameter retrieval, in a real scenario, characterized by a limited amount of assimilated observations and by observation errors embedded as part of the experimental data. In general, by using both  $OD = 1$  cm and  $OD = 2$  cm, it is possible to identify efficiently the sets of parameters similar to those obtained by assimilating the entire soil moisture profile. Although, the performance using  $OD = 2$  cm was found markedly better. The retrieval for GB1 was clearly affected by the initial parameterization; instead the response of GB3 was practically insensitive to this factor, except for  $OD = 1$  cm.

It is important to note that the physically constrained nature of soil water content, in this study chosen as state variable, precludes the simulation of saturated conditions and entails mathematical shortcomings for the UKF sampling strategy. In fact, the retrieval algorithm can in principle sample parametric solutions involving a “wrong” saturation (i.e. giving place to state values being higher than  $\theta_s$ ) or “wrong” dry conditions (i.e. giving place to state values being smaller than  $\theta_r$ ). Notice that these limitations are also attributable to the ensemble Kalman filter, which also involves parameter sampling around a mean state vector. When these meaningless values are sampled, the algorithm simply changes them



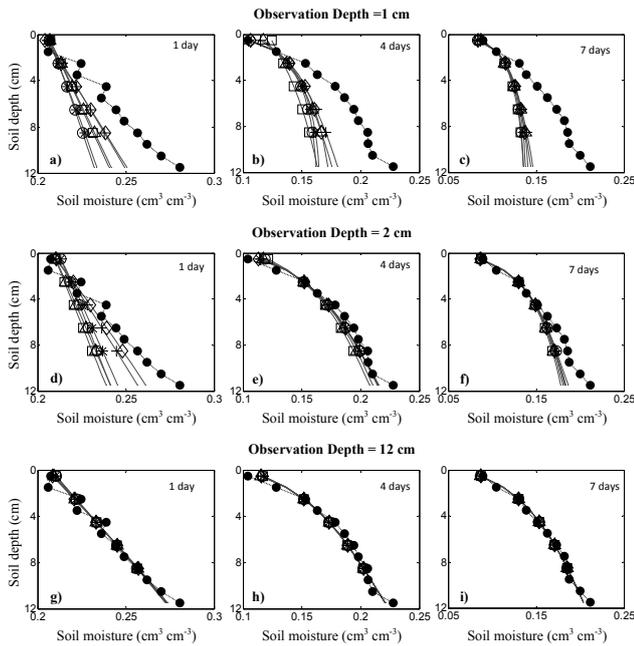
**Figure 7.** Retrieved VGM parameters  $K_S$ ,  $\alpha$  and  $n$  using (a) GA3 and (b) GB1 experimental data, with assimilation frequency  $AF = 1/12 \text{ h}^{-1}$  and observation depth  $OD = 1$  cm. Comparisons account for the six pondered sets of initial parameters S1(○), S2(□), S3(\*), S4(△), S5(+) and S6(◇). The dotted line indicates the value of the parameter found by Romano and Santini (1999).

to keep the state vector solutions within the valid range. Aimed to provide a general solution circumventing this issue, several alternative strategies have been unsuccessfully pondered. One of them was to use adaptive coefficients, scaling the sigma point distribution in the unscented approach (see Eq. 17), in order to shrink the deterministic sampling of the parameter around the mean. However, this demands a high computational cost, and provides temporarily biased retrieved parameters, affecting also tracking and convergence. It was also pondered the use of a “temporarily adaptive”  $\theta_s$  (or in principle  $\theta_r$ ), i.e. making  $\theta_s$  as the maximum soil water content value whenever at least one state value exceeds the adopted actual value. However, this gives place to an irreversible state biasing. Even including  $\theta_s$  as an additional unknown parameter to be retrieved would not avoid this issue, but would rather make it more evident.

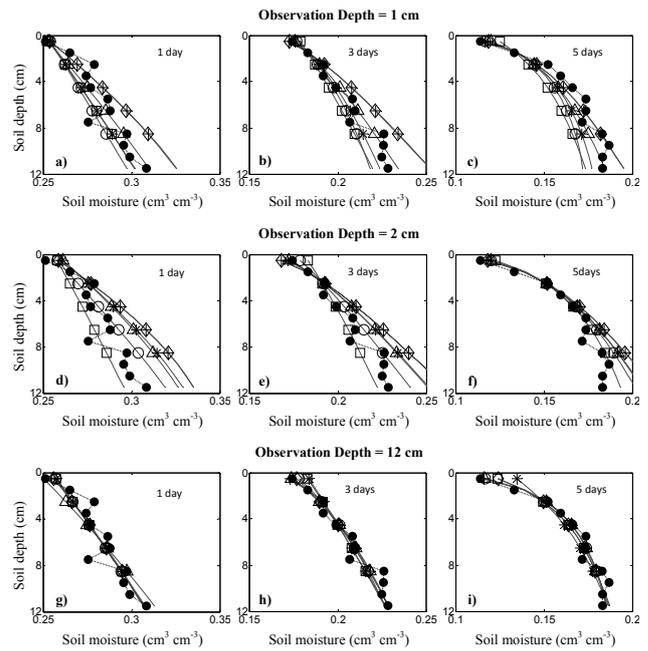
## 5.2 State retrieval

Figures 8 and 9 depict the retrieved states after 1, 4 and 7 days for GA3 and 1, 3 and 5 days for GB1, by assimilating observations every 2 h, within the three observation depths examined (1, 2 and 12 cm). The results show that the dual filter algorithm is generally able to retrieve the true state profiles with a relatively low dependence from the identified parametric array. For soil core GA3 and using  $OD = 1$  cm, the differences between retrieved and measured soil moistures are still large after 7 days (Fig. 8c). Nevertheless, the variability between the six involved simulations is barely noticeable, indicating that the initial parameterization has a low weight at this stage of the assimilation process. Using  $OD = 2$  cm, a very good match between retrieved and measured profiles is found already at the fourth day (Fig. 8e).

In the case of the GB1 experiment, the retrieval process is forced to deal with a relatively larger soil heterogeneity,



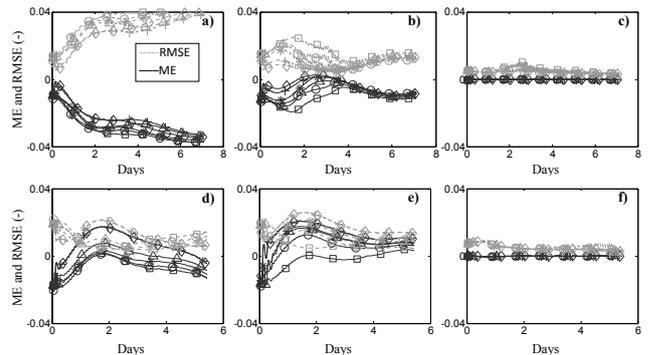
**Figure 8.** Retrieved soil moisture profiles (solid lines) using the GA3 data series after 1 day (a, d, g), 4 days (b, e, h) and 7 days (c, f, i), with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$  and observation depths: (a–c)  $OD = 1 \text{ cm}$ , (d–f)  $OD = 2 \text{ cm}$ , and (g–i)  $OD = 12 \text{ cm}$ . Comparisons account for the six pondered sets of initial parameters  $S1(\circ)$ ,  $S2(\square)$ ,  $S3(*)$ ,  $S4(\triangle)$ ,  $S5(+)$  and  $S6(\diamond)$ . The dotted line with solid circles represents the measured profile.



**Figure 9.** Retrieved soil moisture profiles (solid lines) using the GB1 data series after 1 day (a, d, g), 4 days (b, e, h) and 7 days (c, f, i), with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$  and observation depths: (a–c)  $OD = 1 \text{ cm}$ , (d–f)  $OD = 2 \text{ cm}$ , and (g–i)  $OD = 12 \text{ cm}$ . Comparisons account for the six pondered sets of initial parameters  $S1(\circ)$ ,  $S2(\square)$ ,  $S3(*)$ ,  $S4(\triangle)$ ,  $S5(+)$  and  $S6(\diamond)$ . The dotted line with solid circles represents the measured profile.

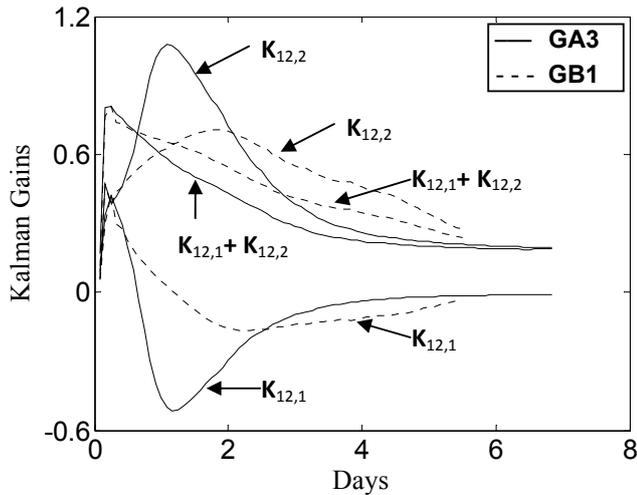
as shown by the irregular scatter of the soil water content data points along the vertical direction (Fig. 9). This vertical variability makes the retrieval process more sensitive to the parameter initialization, allowing for a wider spectrum of probable soil moisture profiles. However, this spectrum of probable soil moisture profiles well represents the soil moisture “anomalies” and the differences between the retrieved soil moisture profiles on the fifth day is very small.

The comparison between the retrieval performances achieved with the observation depths  $OD = 1 \text{ cm}$  and  $OD = 2 \text{ cm}$  provides contrasting results for the two experiments. For soil core GA3, the differences between the soil water content profiles retrieved with  $OD = 1 \text{ cm}$  and  $OD = 2 \text{ cm}$  are still significant on the fourth and seventh day, whereas for soil core GB1 the profiles are similar at both the third and the fifth day. This feature can be more clearly appreciated by visual inspection of Fig. 10 showing the temporal evolution of the ME and RMSE indices for both these experiments. As the initial soil water content is close to the profile mean value, the errors at the beginning of the simulation are relatively small. Nonetheless, it is important to point out that the evolving pattern of these statistics, as the overall performance of the retrieval algorithm, is scarcely affected by the initial soil water content.



**Figure 10.** Evolving mean errors (MEs) and root mean square errors (RMSEs) between predicted and measured profiles using GA3 (a–c) and GB1 (d–f) experimental data series, with assimilation frequency  $AF = 1/2 \text{ h}^{-1}$ , and observation depths: (a, d)  $OD = 1 \text{ cm}$ , (b, e)  $OD = 2 \text{ cm}$ , and (c, f)  $OD = 12 \text{ cm}$ . The analysis accounts for the six pondered sets of initial parameters  $S1(\circ)$ ,  $S2(\square)$ ,  $S3(*)$ ,  $S4(\triangle)$ ,  $S5(+)$  and  $S6(\diamond)$ .

Figure 10a shows that relatively high errors occur for GA3 with  $OD = 1 \text{ cm}$ , with increasing RMSE and ME (in absolute terms) values up to the second day and after the fourth day. A main cause of this relatively poor performance of the approach using GA3 with  $OD = 1 \text{ cm}$  is that water fluxes during the last stage of this experiment are very small (see Fig. 1),



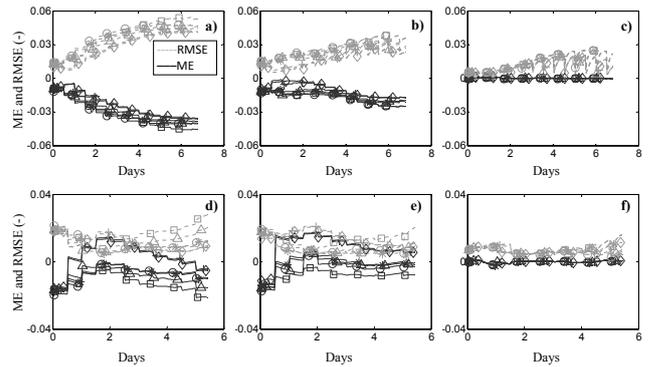
**Figure 11.** Evolving Kalman gain coefficients  $K_{12,1}$ ,  $K_{12,2}$  and the sum  $K_{12,1} + K_{12,2}$  for GA3 and GB1 using observation depths of 1 cm and the parameter initialization S6.  $K_{12,1}$  and  $K_{12,2}$  describe the influence of the first and second observation nodes, respectively, on node 12 (the bottom one).

which implies that the soil water content gradients are very small at the surface and, in turn, poor information is provided by the very top observation nodes to the lower nodes, about the ongoing process. For the same reasons, a small increase of RMSE and ME is also noticeable after the fourth day with  $OD = 2$  cm.

To see how these physical constraints impact on the dynamics of the states retrieval process, Fig. 11 shows the evolving Kalman gain coefficients  $K_{12,1}$ ,  $K_{12,2}$  and  $K_{12,1} + K_{12,2}$  for both data series using  $OD = 1$  cm and parameter initialization S6 (the parameter initialization has a limited impact on this aspect). Provided that the Kalman gain for  $OD = 1$  cm is a matrix of 12 rows (equal to the nodes of the soil profile) and 2 columns (equal to the observed nodes),  $K_{12,1}$  and  $K_{12,2}$  describe how the value retrieved at the twelfth node is influenced by the observations assimilated from the first and second nodes, respectively.

Usually, a value of  $K_{i,j}$  close to one indicates a high effect of the assimilated node  $j$  on the retrieved node  $i$ , whilst a value close to zero indicates a neglecting effect. Given the correspondence between states and observations, then involving an observation operator  $\mathbf{H}$  with ones in the diagonal elements and zero in the off-diagonal elements, a negative value of  $K_{i,j}$  manifests the negative cross-covariance between the states  $i$  and  $j$ . The sum  $K_{12,1} + K_{12,2}$  gives an idea of the combined effect of the two observation nodes, provided that the differences between model predictions and observations are similar.

$K_{12,1}$  for GA3 is almost zero after approximately 4 days. The sum  $K_{12,1} + K_{12,2}$  after this moment is constant and essentially equal to  $K_{12,2}$ . In other words, the assimilation algorithm is still acting on the bottom node, but just using



**Figure 12.** Evolving mean errors (MEs) and root mean square errors (RMSEs) between predicted and measured profiles using GA3 (a–c) and GB1 (d–f) experimental data series, with assimilation frequency  $AF = 1/12 \text{ h}^{-1}$ , and observation depths: (a, d)  $OD = 1$  cm, (b, e)  $OD = 2$  cm, and (c, f)  $OD = 12$  cm. The analysis accounts for the six pondered sets of initial parameters S1 ( $\circ$ ), S2 ( $\square$ ), S3 ( $*$ ), S4 ( $\triangle$ ), S5 ( $+$ ) and S6 ( $\diamond$ ).

the information provided by the second node, while the top node is barely contributing to the retrieved value. The term  $K_{12,1} + K_{12,2}$  obtained for GB1 is higher than that obtained for GA1 practically during the entire assimilation period. In the case of GB1, the assimilation of the top node, by means of the  $K_{12,1}$  coefficient, slightly influences the retrieval value of the bottom node, almost till the end of the experiment.

The negative values of  $K_{12,1}$  seem to be favoured by the adopted experimental settings. At the beginning of the experiment, the guess states transits from a constant matric potential profile to a distribution similar to a constant head profile; hence, the bottom node moves to a wetter range, while the evaporative process drives an opposite trend on the top node. The zero flux condition at the bottom profile favours the relatively large duration of this phase of the process. During these stages the correlation between the first and the bottom nodes use to be markedly negative. According to Fig. 11 this phase lasts approximately 1 and 2 days for GA3 and GB1, respectively. When this trend changes, with the bottom nodes also moving to a drier range (although much more slowly than the top node), the estimated  $K_{12,1}$  increases (i.e. it gets less negative values).

Notice that GB1 Kalman gain coefficients tend to converge toward those of the GA3 experiment after the fourth day. In fact, as it occurs for GA3, the ME and RMSE values for GB1 using  $OD = 1$  cm also slightly increase at the end of the experiment (Fig. 10d). More precisely, the GB1 flux at its last stage is about  $1 \text{ mm day}^{-1}$ , and it is similar to that found for GA3 from about the fourth day (see Fig. 1).

Thus, the dual filter approach performs well in all cases, except for GA3 using  $OD = 1$  cm, where the state retrieval process is relatively slow mainly due to the characteristics of the experiment, beside other factors such as the diminishing state covariance as the experiment advances, the narrow

**Table 4.** Mean error (ME) and root mean square error (RMSE) between predicted and measured soil moisture profiles for GA3 and GB1 data series at the end of the evaporation tests. Values account for the six sets of initial parameters (S1–S6), two observation depths, (OD = 1 and OD = 2 cm) and two assimilation frequencies (AF =  $1/2 \text{ h}^{-1}$  and AF =  $1/12 \text{ h}^{-1}$ ).

		GA3				GB1			
		OD = 1 cm		OD = 2 cm		OD = 1 cm		OD = 2 cm	
AF	Initial set	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE
$1/2 \text{ h}^{-1}$	S1	−0.0334	0.0393	−0.0137	0.0193	−0.0129	0.0145	0.0061	0.0080
	S2	−0.0344	0.0405	−0.0151	0.0209	−0.0126	0.0144	0.0032	0.0046
	S3	−0.0310	0.0366	−0.0123	0.0176	−0.0106	0.0118	0.0074	0.0096
	S4	−0.0319	0.0376	−0.0132	0.0188	−0.0083	0.0096	0.0076	0.0099
	S5	−0.0307	0.0362	−0.0107	0.0157	−0.0041	0.0063	0.0101	0.0132
	S6	−0.0311	0.0367	−0.0118	0.0170	−0.0042	0.0064	0.0100	0.0131
$1/12 \text{ h}^{-1}$	S1	−0.0427	0.0491	−0.0213	0.0275	−0.0096	0.0118	−0.0028	0.0073
	S2	−0.0402	0.0461	−0.0219	0.0281	−0.0211	0.0278	−0.0079	0.0201
	S3	−0.0334	0.0382	−0.0158	0.0210	−0.0101	0.0124	−0.0009	0.0056
	S4	−0.0504	0.0588	−0.0250	0.0342	−0.0155	0.0201	−0.0017	0.0112
	S5	−0.0501	0.0589	−0.0259	0.0360	−0.0046	0.0088	0.0056	0.0090
	S6	−0.0283	0.0325	−0.0148	0.0200	−0.0053	0.0095	0.0053	0.0089

range of the soil moisture content values covered by the top node, which also implies small differences between model predictions and observations. As a demonstration of the relative efficiency of the proposed approach, ME and MAE values decrease when the analysis of the errors obtained for GA3 using OD = 1 cm is limited to the top five nodes.

Analogously to Fig. 10, Fig. 12 depicts the temporal evolution of the ME and RMSE values, but using AF =  $1/12 \text{ h}^{-1}$ . The reduced assimilation frequency gives mainly place to an increase of the ME and RMSE for GA3, while a higher dependence of these statistics from the initial parameter sets for GB1. The error temporal patterns evidence that the profile retrieval process follows trends similar to those observed for the higher resolution, but with a slower convergence rate. It is interesting to observe that when assimilating the entire profile for GA3, the algorithm predicts the correct average soil water content (as the ME is almost null), but the retrieved profiles present deviations from the observed values along the soil column, as testified by the increasing RMSE. Unfortunately, the analysis is limited in time by the short duration of the experimental data series.

Table 4 summarizes the ME and RMSE values computed for OD = 1 cm and 2 cm and AF =  $1/2 \text{ h}^{-1}$  and  $1/12 \text{ h}^{-1}$  at the end of the each simulation. The approach is able to provide good results within the limited time conceded by the duration of the experiments. The average ME with OD = 1 cm is larger than that obtained with OD = 2 cm, by 2.25 times for GA3 and by 1.3 times for GB1. Similar results occur for the average RMSE values. Instead, AF =  $1/12 \text{ h}^{-1}$  produces higher errors than AF =  $1/2 \text{ h}^{-1}$  by 1.45 times for GA3, while by 1.4 times for GB1 with OD = 1 cm and 0.34 times for GB1 with OD = 2 cm. The results are generally more sensible to

the observation depths than to the assimilation frequency, due to the small state vertical gradients, particularly for GA3.

## 6 Conclusions

This study has shown the potential of the dual Kalman filter approach in retrieving both parameters and states simultaneously. The performances of the proposed approach have been evaluated with data obtained from evaporation experiments carried out in the laboratory on two different soil cores.

The dual Kalman filter approach is based on a standard Kalman filter for retrieving state values and an unscented Kalman filter for retrieving the parameters of the soil hydraulic property functions. The approach adopts a linearized numerical scheme of the  $\theta$ -based form of the Richards equation, based on the Crank–Nicolson finite differences, granting the linearity of both the state and the observation equations and thus enabling a direct optimal retrieval of the first and second moment of the states with a standard Kalman filter.

By assimilating soil moisture observations up to soil depths of 1 and 2 cm, the approach allows properly identifying a set of parameters that is in a very good agreement with that one obtained by assimilating the entire observed profiles. The retrieved parameters are also in a reasonably good agreement with the parameters found by Romano and Santini (1999), particularly for  $K_s$  and  $n$  in the case of the GB1 experiment. The retrieved parameter  $\alpha$  is larger than that estimated by Romano and Santini (1999), as result of the different type of information employed for estimating the parameters.

The method also shows a good performance in terms of state retrieval and proved to be able to deal even with the somewhat larger heterogeneities of soil core GB1. The prediction performance proved to be more sensitive to the observation depths than to the assimilation frequency, when changing the observation depths from 1 to 2 cm, while the assimilation frequency from  $1/2 \text{ h}^{-1}$  to  $1/12 \text{ h}^{-1}$ .

It is important to note the marked flexibility and stability of the approach, independently from the errors associated with the initial states and parameter sets. Further work is needed to investigate whether this approach is also able to cope with two- or three-dimensional problems of soil water flux efficiently, by undertaking more complex assimilation processes.

*Acknowledgements.* The laboratory evaporation tests have been conducted in the Soil Hydrology Laboratory of the Department of Agricultural Engineering – University of Napoli Federico II, under the supervision of prof. Alessandro Santini. This study has been supported by P. O. N. project AQUATEC – New technologies of control, treatment, and maintenance for the solution of water emergency. Hanoi Medina has been also supported by the Abdus Salam International Centre for Theoretical Physics (ICTP), where he has been appointed as Junior Associate.

Edited by: N. Verhoest

## References

- Camporese, M., Paniconi, C., Putti, M., and Salandin, P.: 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, 2009.
- Carsel, R. F. and Parrish, R. S.: Developing joint probability distributions of soil water retention characteristics, *Water Resour. Res.*, 24, 755–769, doi:10.1029/WR024i005p00755, 1988.
- Chirico, G. B., Medina, H., and Romano, N.: Functional evaluation of PTF prediction uncertainty: An application at hillslope scale, *Geoderma*, 155, 193–202, 2010.
- Chirico, G. B., Medina, H., and Romano, N.: Kalman filters for assimilating near-surface observations into the Richards equation – Part 1: Retrieving state profiles with linear and nonlinear numerical schemes, *Hydrol. Earth Syst. Sci.*, 18, 2503–2520, doi:10.5194/hess-18-2503-2014, 2014.
- De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency, *Water Resour. Res.*, 43, W06401, doi:10.1029/2006WR005100, 2007a.
- De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, *Water Resour. Res.*, 43, W09410, doi:10.1029/2006WR005449, 2007b.
- Dunne, S. and Entekhabi, D.: An ensemble-based reanalysis approach to land data assimilation, *Water Resour. Res.*, 41, W02013, doi:10.1029/2004WR003449, 2005.
- Entekhabi, D., Nakamura, H., and Njoku, E. G.: Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remote sensed observations, *IEEE Trans. Geosci. Remote Sens.*, 32, 438–448, 1994.
- Escorihuela, M. J., Chanzy, A., Wigneron, J. P., and Kerr, Y. H.: Effective soil moisture sampling depth of L-band radiometry: A case study, *Remote Sense Environ.*, 114, 995–1001, 2010.
- Haverkamp, R. M., Vauclin, M., Tourna, J., Wierenga, P. J., and Vachaud, G.: A comparison of numerical simulation models for one-dimensional infiltration, *Soil Sci. Soc. Am. J.*, 41, 285–294, 1977.
- Hoeben, R. and Troch P. A.: Assimilation of active microwave observation data for soil moisture profile estimation, *Water Resour. Res.*, 36, 2805–2819, doi:10.1029/2000WR900100, 2000.
- Julier, S. J. and Uhlmann, J. K. A New Extension of the Kalman Filter to Nonlinear Systems, in: *Proc. SPIE – Int. Soc. Opt. Eng. (USA)*, 3068, 182–193, 1997.
- Julier, S. J. and Uhlmann, J. K.: Unscented Filtering and Nonlinear Estimation, *Pro. IEEE*, 92, 401–422, 2004.
- Jury, W. A., Gardner, W. R., and Gardner, W. H.: *Soil Physics*, 5th Edn., John Wiley, New York, 1991.
- Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, *ASME J. Basic Eng.*, 82D, 35–45, 1960.
- Kavetski, D., Binning, P., and Sloan, S. W.: Noniterative time stepping schemes with adaptive truncation error control for the solution of Richards equation, *Water Resour. Res.*, 38, 1211, doi:10.1029/2001WR000720, 2002.
- 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.
- Lü, H., Yu, Z., Horton, R., Zhu, Y., Wang, Z., Hao, Z. and Xiang, L.: Multi-scale assimilation of root zone soil water predictions, *Hydrol. Process.*, 25, 3158–3172, doi:10.1002/hyp.8034, 2011.
- McLaughlin, D. B.: An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering, *Adv. Water Resour.*, 25, 1275–1286, 2002.
- Medina, H., Romano, N., and Chirico, G. B.: Kalman filters for assimilating near-surface observations into the Richards equation – Part 2: A dual filter approach for simultaneous retrieval of states and parameters, *Hydrol. Earth Syst. Sci.*, 18, 2521–2541, doi:10.5194/hess-18-2521-2014, 2014.
- Montzka, C., Moradkhani, H., Weihermuller, L., Hendricks-Franssen, H. J., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, *J. Hydrol.*, 399, 410–421, 2011.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P.: Dual state–parameter estimation of hydrologic models using ensemble Kalman filter, *Adv. Water Resour.*, 28, 135–147, 2005a.
- Moradkhani, H., Hsu, K. L., Gupta, H. V., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604, 2005b.
- Paniconi, C., Aldama, A. A., and Wood, E. F.: Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards equation, *Water Resour. Res.*, 27, 1147–1163, 1991.

- 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, *Hydrol. Earth Syst. Sci.*, 16, 375–390, doi:10.5194/hess-16-375-2012, 2012.
- Pringle, M. J., Romano, N., Minasny, B., Chirico, G. B., and Lark, R. M.: Spatial evaluation of pedotransfer functions using wavelet analysis, *J. Hydrol.*, 333, 182–198, 2007.
- Qin, J., Liang, S. L., Yang, K., Kaihotsu, I., Liu, R. G., and Koike, T.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, *J. Geophys. Res. Atmos.*, 114, D15103, doi:10.1029/2008JD011358, 2009.
- Reichle, R. H. and Koster, R. D.: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation, *J. Hydrometeorol.*, 4, 1229–1242, 2003.
- Ritter, A., Muñoz-Carpena, R., Regalado, C. M., Vanclooster, M., and Lambot, S.: Analysis of alternative measurement strategies for the inverse optimization of the hydraulic properties of a volcanic soil, *J. Hydrol.*, 295, 124–139, 2004.
- Romano, N. and Santini, A.: Determining soil hydraulic functions from evaporation experiments by a parameter estimation approach: Experimental verifications and numerical studies, *Water Resour. Res.*, 35, 3343–3359, 1999.
- Scharnagl, B., Vrugt, J. A., Vereecken, H., and Herbst, M.: Inverse modelling of in situ soil water dynamics: investigating the effect of different prior distributions of the soil hydraulic parameters, *Hydrol. Earth Syst. Sci.*, 15, 3043–3059, doi:10.5194/hess-15-3043-2011, 2011.
- Šimůnek, J. and van Genuchten, M. T.: Estimating Unsaturated Soil Hydraulic Properties from Tension Disc Infiltrometer Data by Numerical Inversion, *Water Resour. Res.*, 32, 2683–2696, doi:10.1029/96WR01525, 1996.
- Thiemann, M., Trosset, M., Gupta, H. V., and Sorooshian, S.: Bayesian recursive parameter estimation for hydrologic models, *Water Resour. Res.*, 37, 2521–35, 2001.
- Todini, E.: Mutually interactive state/parameter estimation (MISP), in *Application of Kalman Filter to Hydrology, Hydraulics and Water Resources*, edited by: Chiu, C.-L., University of Pittsburgh, Pittsburgh, Pa., 1978.
- 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.
- Vereecken, H., Huisman, J. A., Bogaen, H., Vanderborght, J., Vrugt, J. A., and Hopmans, J. W.: On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resour. Res.*, 44, W00D06, doi:10.1029/2008WR006829, 2008.
- Vrugt, J. A., Bouten, W., and Weerts, A. H.: Information content of data for identifying soil hydraulic parameters from outflow experiments, *Soil Sci. Soc. Am. J.*, 65, 19–27, 2001.
- Vrugt, J. A., Bouten, W., Gupta, H. V., and Sorooshian, S.: Toward improved identifiability of hydrologic model parameters: the information content of experimental data, *Water Resour. Res.*, 38, 1312, doi:10.1029/2001WR001118, 2002. 2044, 2045, 2002.
- Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, 41, W01017, doi:10.1029/2004WR003059, 2005.
- Vrugt, J. A., ter Braak, C. J. F., Diks, C. G. H., and Schoups, G.: Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications, *Adv. Water Resour.*, 51, 457–478, doi:10.1016/j.advwatres.2012.04.002, 2013.
- 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.
- Walker, J. P., Houser, P. R., and Willgoose, G. R.: Active microwave remote sensing for soil moisture measurement: a field evaluation using ERS-2, *Hydrol. Process.*, 18, 1975–1997, 2004.
- Wöhling, Th. and Vrugt, J. A.: Multi-response multi-layer vadose zone model calibration using Markov chain Monte Carlo simulation and field water retention data, *Water Resour. Res.*, 47, W04510, doi:10.1029/2010WR009265, 2011.
- Yang, K., Koike, T., Kaihotsu, I., and Qin, J.: Validation of a dual-pass microwave land data assimilation system for estimating surface soil moisture in semiarid regions, *J. Hydrometeorol.*, 10, 780–793, 2009.