Hydrol. Earth Syst. Sci. Discuss., 9, 13373–13414, 2012
www.hydrol-earth-syst-sci-discuss.net/9/13373/2012/
doi:10.5194/hessd-9-13373-2012

Hydrology and Earth System Sciences Discussions

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

H. Medina[1], **N. Romano**[2], **and G. B. Chirico**[2]

[1]Department of Basic Sciences, Agrarian University of Havana, Havana, Cuba
[2]Department of Agricultural Engineering, University of Naples Federico II, Naples, Italy

Received: 5 November 2012 – Accepted: 26 November 2012 – Published: 3 December 2012

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The purpose of this work is to evaluate the performance of a dual Kalman Filter procedure in retrieving states and parameters of a 1-D 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 measured data at 1 and 2 cm soil depths are in very good agreement with those obtained by exploiting the entire measured profiles. A reasonably good correspondence has been found between the parameters obtained from the proposed assimilation technique and those identified by applying a non sequential parameter estimation method. The dual Kalman Filter also performs very 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 chosen initial conditions and the involved uncertainty.

1 Introduction

Models applied for simulating hydrologic systems require parameters that often are not available or are not measureable in real-world applications. These parameters need then to be retrieved from the available hydrological data using numerical techniques which have to deal with the uncertainties involved in both the observations and the model simplifications (Vrugt et al., 2005). Given the relevance of the vadose zone processes in the large-scale hydrologic and hydro-meteorological model applications, a more accurate as possible identification of the soil hydraulic parameters is of paramount importance.

HESSD
9, 13373–13414, 2012

Several studies have pointed out the potential of using near-surface soil water content observations derived from remote sensing and ground-based networks, for retrieving states and parameters of soil hydrological models. The majority of the studies focussed on using near-surface soil moisture information to retrieve soil moisture profiles in the unsaturated zone, while assuming soil hydraulic properties to be known (e.g. Entekhabi et al., 1994; Walker et al., 2001; Walker and Houser, 2004; Dunne and Entekhabi, 2005). A few recent studies aimed at retrieving soil hydraulic properties by assimilating near-surface soil moisture. Ines and Mohanty (2008) used a genetic algorithm (GA) to identify soil water retention and hydraulic conductivity functions by inverting a soil-water-atmosphere-plant (SWAP) model using observed near-surface soil moisture (0–5 cm depth) as search criterion. To retrieve the effects of land surface subgrid heterogeneity and soil water phase changes, Tian et al. (2008) applied a dual Unscented Kalman Filter (UKF) for assimilating satellite observations of soil moisture in a land surface model, based on a vertical soil water balance coupled with the Darcy law. Lü et al. (2011) applied a direct insertion method for assimilating surface soil moisture within the Richards equation, coupled with a particle swarm optimisation algorithm, for identifying the optimal saturated hydraulic conductivity. Montzka et al. (2011) presented a synthetic study concerning the assimilation of the top soil moisture data in HYDRUS-1D simulations with the particle filter, estimating states and parameters by a perturbing strategy of the original parameter vector.

Apart from the work of Medina et al. (2012), we are not aware of any study that is explicitly concerned with the simultaneous retrieval of states and parameters in the Richards equation. The states and parameters of a dynamic system, described by state-space equations discretised in time domain, can be simultaneously retrieved from noise observations following to two different approaches. The first approach employs a single filter to predict the evolution of the joint probability distribution of states and parameters, which are combined in an augmented state vector. The second approach, referred to as the *dual filter* approach, uses two separate filters, one for the state space and the other for the parameter space.

HESSD
9, 13373–13414, 2012

As shown by Chirico et al. (2012), the optimal assimilation algorithm should be designed by considering three main factors: (i) the type of numerical scheme employed for the Richards equation, which corresponds to a state-space representation of the dynamic system, and can be either linear or nonlinear; (ii) the type of variable selected for describing the system state (e.g. soil water content or pressure head); (iii) the type of observed variable (e.g. soil water content or pressure head).

When retrieving states only, Chirico et al. (2012) demonstrated that the optimal algorithm consists of a standard Kalman Filter applied to a Crank-Nicolson numerical scheme. A linear, although implicit, numerical scheme provides an accurate solution to the recursive estimation problem with minor computational efforts as compared with a nonlinear numerical scheme. Actually, nonlinear numerical schemes, despite being more efficient from a strict numerical perspective, require the implementation of non-standard Kalman Filter schemes, such as the Extended Kalman Filter or the Unscented Kalman Filter. This implies larger computational efforts without granting improved retrieving performance. It has been also pointed out that it is quite effective to select the h-form or θ-form of the Richards equation consistently with the type of assimilated variable, thus ensuring a linear relationship between states and observations.

Consequently, Medina et al. (2012) proposed a dual Kalman Filter (Dual-KF) for real-time simultaneous prediction of soil water content or matric head profiles and soil hydraulic parameters, by assimilating near-surface information in a 1-D Richards' equation. The dual Kalman Filter is designed as follows: a standard Kalman Filter is implemented with a Crank-Nicolson numerical scheme to retrieve soil state profiles, while an Unscented Kalman Filter (UKF) 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. The UKF has been proved to be very efficient in retrieving soil hydraulic parameters, which are linked by highly nonlinear relationships and cover rather different ranges of variation, with limiting values defined by physical constrains. The advantage of the dual approach

HESSD
9, 13373–13414, 2012

is to limit the unscented transformation to the parameters space, which dimensionality is normally small as compared with the state space.

Medina et al. (2012) evaluated the dual Kalman Filter using a synthetic case study. The advantage of referring to synthetic values instead of 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 aspects of the system at hand, such as the soil heterogeneity and errors in the measurements. Actually, data assimilation studies are greatly demanded to develop methods that are feasible for real applications (Liu et al., 2012).

This paper aims at evaluating the dual Kalman Filter performance by exploiting data measured during evaporation experiments executed on soil cores in laboratory. The strength of the evaporation experiments is that the data are gathered during a transient flow, which is very close to natural processes occurring in real soils, thereby providing highly representative hydraulic behaviour of the porous medium 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 for the determination of unsaturated soil hydraulic properties. The method consists in 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 valuable experimental dataset to test the dual Kalman Filter approach developed by Medina et al. (2012).

The performance of the Dual-KF approach is herein deeply examined, accounting for the effects of the observation depth, the assimilation frequency and the parameter initialization, on both the state and parameter retrieving processes.

HESSD
9, 13373–13414, 2012

2 The dual Kalman filter formulation

Medina et al. (2012) provides a detailed description of the algorithm employed for the separate state-space representation used to retrieve states and parameters. At every time-step, the current estimate of the parameters is used in the standard linear state-filter, and the current estimate of the states is used in the unscented parameter-filter. In order to make this reading easier, the state and parameter filter equations employed in the assimilation algorithm is herein summarised.

The set of system equations can be written as:

$$\boldsymbol{x}_k = \mathbf{F}\left(\boldsymbol{x}_{k-1},\, \boldsymbol{u}_{k-1},\, \boldsymbol{v}_{k-1},\, \hat{\boldsymbol{w}}_{k-1}\right) \tag{1}$$

$$\boldsymbol{y}_k = \mathbf{H}\left(\boldsymbol{x}_k,\, \boldsymbol{n}_k,\, \hat{\boldsymbol{w}}_{k-1}\right) \tag{2}$$

for the state vector $\boldsymbol{x}$ at time k, and

$$\boldsymbol{w}_k = \boldsymbol{w}_{k-1} + \boldsymbol{r}_{k-1} \tag{3}$$

$$\boldsymbol{y}_k = \mathbf{H}\left(\mathbf{F}\left(\hat{\boldsymbol{x}}_{k-1},\, \boldsymbol{u}_{k-1},\, \boldsymbol{v}_{k-1},\, \boldsymbol{w}_{k-1}\right),\, \boldsymbol{n}_k,\, \boldsymbol{w}_k\right) \tag{4}$$

for the parameter vector $\boldsymbol{w}$. F is the state transition function, H is the observation function $\boldsymbol{v}_k$ is the process noise associated to the state equation, $\boldsymbol{n}_k$ is the observation noise, $\boldsymbol{u}_k$ represents and exogenous input to the system and $\boldsymbol{r}_k$ is the process noise linked to the parameter equation.

2.1 The standard Kalman Filter formulation for linear state retrieving

The linear algorithm for the states retrieving is summarised in the following three phases.

I. Initialization:

$$\hat{\boldsymbol{x}}_0 = E\left[\boldsymbol{x}_0\right] \tag{5}$$

HESSD
9, 13373–13414, 2012

$$\mathbf{P}_{x_0} = E\left[\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right)\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right)^T\right] \quad (6)$$

$$\mathbf{R}_{v_0} = E\left[\left(\boldsymbol{v}_0 - \overline{\boldsymbol{v}}_0\right)\left(\boldsymbol{v}_0 - \overline{\boldsymbol{v}}_0\right)^T\right] \quad (7)$$

$$\mathbf{R}_{n_0} = E\left[\left(\boldsymbol{n}_0 - \overline{\boldsymbol{n}}_0\right)\left(\boldsymbol{n}_0 - \overline{\boldsymbol{n}}_0\right)^T\right] \quad (8)$$

where $\mathbf{P}_x$, $\mathbf{R}_v$ and $\mathbf{R}_n$ are the system auto covariance matrix, the auto covariance matrix of the process noise, and the auto covariance matrix of the observation covariance, respectively. Subscript "0" indicates initial values.

II. Prediction phase, by computing the state mean and covariances:

$$\hat{\boldsymbol{x}}_k^- = \mathbf{F}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{v}_{k-1}, \hat{\boldsymbol{w}}_{k-1}\right) \quad (9)$$

$$\mathbf{P}_{x_k}^- = \mathbf{F}\mathbf{P}_{x_{k-1}}\mathbf{F}^T + \mathbf{F}\mathbf{R}_v\mathbf{F}^T \quad (10)$$

where subscript k indicates the time step, $\hat{\boldsymbol{x}}_k^-$ and $\mathbf{P}_{x_k}^-$ represents the optimal prediction (prior mean at time t_k) of $\boldsymbol{x}_k$ and $\mathbf{P}_{x_k}$.

III. Correction phase, for updating estimates with the last observation:

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

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

$$\mathbf{P}_{x_k} = \mathbf{P}_{x_k}^- - \mathbf{K}_k^x \mathbf{P}_{\hat{y}_k}^x \left(\mathbf{K}_k^x\right)^T \quad (13)$$

where $\mathbf{K}_k^x$ is the Kalman gain, expressing the ratio of the expected cross-covariance matrix of the process prediction error and the observation prediction error, $\mathbf{P}_{x_k y_k} = \mathbf{P}_{x_k}^- \mathbf{H}_{x_k}^T$, and the expected auto-covariance matrix of the observation prediction error, $\mathbf{P}_{\tilde{y}_k}^x = \mathbf{H}_{x_k}\mathbf{P}_{x_k}^- \mathbf{H}_{x_k}^T + \mathbf{H}_{x_k}\mathbf{R}_n\mathbf{H}_{x_k}^T$. $\hat{\boldsymbol{x}}_k$ and $\mathbf{P}_{x_k}$, represent the states posterior density mean and covariance.

HESSD
9, 13373–13414, 2012

2.2 The Unscented Kalman filter (UKF) formulation for parameter estimation

From an optimization perspective, the Kalman filter parameter estimation looks for minimizing the following prediction-error cost:

$$J = \sum_{t=1}^{k} \left(\boldsymbol{y}_t - \hat{\boldsymbol{y}}_t^-\right)^T \left(\mathbf{R}_e\right)^{-1} \left(\boldsymbol{y}_t - \hat{\boldsymbol{y}}_t^-\right) \tag{14}$$

where $\mathbf{R}_e$ is an artificial noise parameter covariance, assumed as a constant diagonal matrix.

In the UKF the distribution of the parameters is represented by a Gaussian random variable, being specified using a minimal set of L_w sample points, so called sigma points, which are deterministically chosen to completely capture the true mean and covariance of the variable 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).

The algorithm for the dynamic retrieving of the unknown parameters can be summarised in the following four phases.

I. Initialization:

$$\hat{\boldsymbol{w}}_0 = E\left[\boldsymbol{w}\right] \tag{15}$$

$$\mathbf{P}_{w_0} = E\left[\left(\boldsymbol{w} - \hat{\boldsymbol{w}}_0\right)\left(\boldsymbol{w} - \hat{\boldsymbol{w}}_0\right)^T\right]. \tag{16}$$

II. Time update phase:

$$\hat{\boldsymbol{w}}_k^- = \hat{\boldsymbol{w}}_{k-1} \tag{17}$$

$$\mathbf{P}_{w_k}^- = \mathbf{P}_{w_{k-1}} + \mathbf{R}_{r_{k-1}} \tag{18}$$

where $\mathbf{R}_{r_k} = \left(\lambda_{\mathrm{RLS}}^{-1} - 1\right)\mathbf{P}_{w_k}$, being $\lambda_{\mathrm{RLS}} \in (0,\ 1]$ a forgetting factor as defined in the recursive least-squares (RLS) algorithm. This relationship for $\mathbf{R}_r$ is chosen on the basis

HESSD
9, 13373–13414, 2012

of the results illustrated in Medina et al. (this issue), where alternative expressions have been compared.

III. Computation of the sigma points for the measuring update:

$$\boldsymbol{\mathcal{W}}_{k-1} = \left[\hat{\boldsymbol{w}}_{k-1} \quad \hat{\boldsymbol{w}}_{k-1} + \sqrt{\gamma \mathbf{P}_{k-1}} \quad \hat{\boldsymbol{w}}_{k-1} - \sqrt{\gamma \mathbf{P}_{k-1}}\right] \tag{19}$$

where γ is a coefficient scaling the sampled parameter distribution around the mean parameter vector.

IV. Measuring update equations:

$$\boldsymbol{\mathcal{Y}}_{k|k-1} = \mathbf{H}\left(\mathbf{F}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{v}_{k-1}, \boldsymbol{\mathcal{W}}_{k-1}\right)\right). \tag{20}$$

The output function is obtained as:

$$\hat{\boldsymbol{y}}^{-}_{w_k} = \sum_{i=0}^{2L_w} \mu_i^{(m)} \mathcal{Y}^{z}_{i,k|k-1} \tag{21}$$

$$\mathbf{P}_{w_k \boldsymbol{y}_k} = \sum_{i=0}^{2L} \mu_i^{(c)} \left(\boldsymbol{\mathcal{W}}_{i,k|k-1} - \hat{\boldsymbol{w}}^{-}_{k}\right)\left(\boldsymbol{\mathcal{Y}}_{i,k|k-1} - \hat{\boldsymbol{y}}^{-}_{k}\right)^{T} \tag{22}$$

$$\mathbf{P}_{\tilde{y}_k} = \sum_{i=0}^{2L} \mu_i^{(c)} \left(\boldsymbol{\mathcal{Y}}_{i,k|k-1} - \hat{\boldsymbol{y}}^{-}_{k}\right)\left(\boldsymbol{\mathcal{Y}}_{i,k|k-1} - \hat{\boldsymbol{y}}^{-}_{k}\right)^{T} + \mathbf{R}_{e_k} \tag{23}$$

$$\mathbf{K}^{w}_{k} = \mathbf{P}_{w_k \boldsymbol{y}_k} \mathbf{P}^{-1}_{\tilde{y}_k} \tag{24}$$

$$\hat{\boldsymbol{w}}_{k} = \hat{\boldsymbol{w}}^{-}_{k-1} + \mathbf{K}^{w}_{k}\left(\boldsymbol{y}_k - \hat{\boldsymbol{y}}^{-}_{k}\right) \tag{25}$$

$$\mathbf{P}_{w_k} = \mathbf{P}^{-}_{w_k} - \mathbf{K}^{w}_{k} \mathbf{P}^{w}_{\tilde{y}_k} \left(\mathbf{K}^{w}_{k}\right)^{T} \tag{26}$$

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

HESSD
9, 13373–13414, 2012

m and c, respectively. More details about the UKF implementation can be found in Medina et al. (2012).

3 Materials and methods

3.1 Governing equation

The water movement along a vertical soil column, modelled as a homogeneous, variably saturated porous medium under isothermal conditions, is simulated using the θ-based form of the Richards equation (Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[\left(D(\theta)\frac{\partial \theta}{\partial z} - K(\theta)\right)\right] \tag{27}$$

where t is time, z is soil depth taken positive downward, with $z = 0$ at the top of the profile, θ is the soil water content [$L^3\,L^{-3}$], $D(\theta) = K(\theta)\frac{\partial h}{\partial \theta}$ [$L^2\,T^{-1}$] is the unsaturated diffusivity, with K [$L\,T^{-1}$] being the unsaturated hydraulic conductivity and h the soil water pressure head [L].

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

$$\theta(h) = \theta_r + \theta_s - \theta_r[1 + |\alpha h|^n]^{-m} \tag{28}$$

$$K(\theta) = K_S S_e^{\lambda}\left[1 - \left(1 - S_e^{1/m}\right)^m\right]^2 \tag{29}$$

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

HESSD
9, 13373–13414, 2012

3.2 Crank-Nicolson finite difference scheme

The numerical solution of the θ-based form of the Richards equation (Eq. 27) 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:

$$\left(-\frac{D^k_{i-1/2}}{2\,\Delta z_i\,\Delta z_u};\ \frac{1}{\Delta t^k}+\frac{\frac{D^k_{i-1/2}}{\Delta z_u}+\frac{D^k_{i+1/2}}{\Delta z_l}}{2\,\Delta z_i};\ -\frac{D^k_{i+1/2}}{2\,\Delta z_i\,\Delta z_l}\right)\begin{pmatrix}\theta^{k+1}_{i-1}\\ \theta^{k+1}_{i}\\ \theta^{k+1}_{i+1}\end{pmatrix}$$
$$=\left(\frac{D^k_{i-1/2}}{2\,\Delta z_i\,\Delta z_u};\ \frac{1}{\Delta t^j}-\frac{\frac{D^k_{i-1/2}}{\Delta z_u}+\frac{D^k_{i+1/2}}{\Delta z_l}}{2\,\Delta z_i};\ \frac{D^k_{i+1/2}}{2\,\Delta z_i\,\Delta z_l}\right)\begin{pmatrix}\theta^{k}_{i-1}\\ \theta^{k}_{i}\\ \theta^{k}_{i+1}\end{pmatrix}+\frac{K^k_{i-1}-K^k_{i+1}}{2\,\Delta z_i} \quad (30)$$

where subscript i is the node number (increasing downward), superscript k is 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 discretised, with $\Delta z_u = z_i - z_{i-1}$, $\Delta z_l = z_{i+1} - z_i$ and Δz_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:

$$\left(\frac{1}{\Delta t^j}+\frac{D^k_{1+1/2}}{2\,\Delta z_1\,\Delta z_l};\ -\frac{D^k_{1+1/2}}{2\,\Delta z_1\,\Delta z_l}\right)\begin{pmatrix}\theta^{k+1}_{1}\\ \theta^{k+1}_{2}\end{pmatrix}$$
$$=\left(\frac{1}{\Delta t^k}-\frac{D^k_{1+1/2}}{2\,\Delta z_1\,\Delta z_l};\ \frac{D^k_{1+1/2}}{2\,\Delta z_1\,\Delta z_l}\right)\begin{pmatrix}\theta^{k}_{1}\\ \theta^{k}_{2}\end{pmatrix}+\frac{Q_{\text{top}}-K^k_{1+1/2}}{\Delta z_1} \quad (31)$$

HESSD
9, 13373–13414, 2012

$$\left(-\frac{D^k_{n-1/2}}{2\Delta z_n \Delta z_u}; \frac{1}{\Delta t^k} + \frac{D^k_{n-1/2}}{2\Delta z_n \Delta z_u}\right)\begin{pmatrix}\theta^{k+1}_{n-1}\\ \theta^{k+1}_{n}\end{pmatrix}$$
$$= \left(\frac{D^k_{n-1/2}}{2\Delta z_n \Delta z_u}; \frac{1}{\Delta t^k} - \frac{D^k_{n-1/2}}{2\Delta z_n \Delta z_u}\right)\begin{pmatrix}\theta^{k}_{n-1}\\ \theta^{k}_{n}\end{pmatrix} + \frac{K^k_{n-1/2} - Q_{\text{bot}}}{\Delta z_n} \quad (32)$$

being Q_{top} and Q_{bot} the fluxes at the top and the bottom of the soil profile, respectively.

3.3 Experimental data set

We used the data of two evaporation tests reported in the paper by Romano and Santini (1999) and performed on the two undisturbed soil core named as GA3 and GB1. Each of these soil core 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 Dual-KF 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, after being completely sealed at the bottom, is then 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 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.

HESSD
9, 13373–13414, 2012

Table 1 lists the basic physical properties of the two soil samples together with the VGM model parameters α, n and Ks 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 time series, with a time update of 600 s, have been built by polynomial interpolation of the gamma-ray measurements at all depths and are taken as true state values.

In the assimilation algorithm, the observation values are sampled from the interpolated soil water content time series. 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 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 corresponds to 7 and 5 1/2 days, respectively.

4 Dual-KF implementation

Similarly to the synthetic study of Medina et al. (2012), the saturated (θ_s) and the residual (θ_r) soil water contents are assumed to be known, as these parameters can be easily determined by direct or indirect methods (e.g. Chirico et al., 2007, 2010; Pringle et al., 2007). As reported in the paper by Romano and Santini (1999), the values of parameter θ_s are fixed to 0.31 $cm^3\,cm^{-3}$ for GA3 and 0.35 $cm^3\,cm^{-3}$ for GB1, respectively. The values of parameter θ_r are instead defined according to the values suggested by Carsel and Parrish (1988) for soils of the same textural class: $\theta_r = 0.067\,cm^3\,cm^{-3}$ for a silt-loam soil as GA3, and $\theta_r = 0.08\,cm^3\,cm^{-3}$ 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).

HESSD
9, 13373–13414, 2012

The Unscented Kalman filter is thus implemented to retrieve the remaining parameters K_S, α, and n. As shown in Medina et al. (2012), a variable transformation is applied to constrain the retrieved parameter values to a certain physically meaningful range and thus to guaranty 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:

$$w_i = w_{i_{\min}} + \left(w_{i_{\max}} - w_{i_{\min}}\right) \boldsymbol{g}(\delta w_i) \tag{33}$$

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

$$\boldsymbol{g}(\delta w_i) = \frac{\delta w_i}{2\left(1 + |\delta w_i|\right)} + 0.5. \tag{34}$$

Table 2 summarises 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 $\theta_0 = 0.28\,\mathrm{cm^3\,cm^{-3}}$ for soil core GA3 and $\theta_0 = 0.31\,\mathrm{cm^3\,cm^{-3}}$ 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.

The adopted experiment involved three 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 is the same as assimilating the entire profile, has been included as benchmark performance for OD = 1 cm and OD = 2 cm in the parameter identification.

HESSD
9, 13373–13414, 2012

Two assimilation frequencies have been used: a finer frequency with assimilations every 2 h ($AF = 1/2\,h^{-1}$), and a coarser one with assimilations every 12 h ($AF = 1/12\,h^{-1}$).

The initial covariance matrices are all diagonals and are defined analogously to Medina et al. (2012): the initial state covariance matrix is set to a value of one thousand percent 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 auto covariance matrix is updated using the 2 % of the observed state vector, while the system noise covariance is updated assuming the 5 % of the profile state vector.

The limiting and initialization values of parameters *Ks*, α and n are also identical to those employed by Medina et al. (2012). Table 3 summarizes the minimum value (w_{min}) and the prescribed range ($w_{max} - w_{min}$) 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 retrieving algorithm.

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

$$ME_j = \frac{1}{N}\sum_{i=1}^{N}\left(x_{i,j}^{guess} - x_{i,j}^{true}\right) \quad (35)$$

$$RMSE_j = \left[\frac{1}{N-1}\sum_{i=1}^{N}\left(x_{i,j}^{guess} - x_{i,j}^{true}\right)^2\right]^{1/2} \quad (36)$$

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

HESSD
9, 13373–13414, 2012

5 Results

5.1 Parameter retrieving

Figures 2 and 3 show the temporal pattern of the retrieved VGM parameters for GA3 and GB1, respectively, by assimilating the observed soil water content every two hours 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 parameters values retrieved using OD = 2 cm (Fig. 2b) and OD = 12 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., 2012). The final retrieved K_S value is approximately 1×10^{-4} cm s^{-1}, higher than both the value estimated by Romano and Santini (1999) for the VGM hydraulic model (0.222×10^{-4} cm s^{-1}), and the value directly estimated with the falling head method (0.345×10^{-4} 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 using OD = 1 cm, parameter n converges to two main values: one approximately 2.5 and the other one approximately 1.7 (Fig. 2a). For OD = 2 cm and OD = 12 cm, the convergence patterns are less dependent on the initial set of parameter values. For OD = 2 cm, n converges to a value approximately of 2.5, while for OD = 12 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 retrieving process provides values of K_S and n that converge toward those identified by Romano and Santini (1999) ($K_S = 1.565 \times 10^{-4}$ cm s^{-1} and $n = 2.7$, respectively). Using OD = 2 cm, the means of Ks and n are found practically equal to those reported by Romano and Santini (1999). The parameters retrieved by assimilating the entire profile (OD = 12 cm, Fig. 3c) follow two distinct patterns. The parameter values retrieved with initial sets S2, S3, and S6 follow patterns fairly close to the values found by Romano and Santini (1999), with mean K_S, α, and n of approximately 1.5×10^{-4} cm s^{-1}, 0.028 cm^{-1} and 2.8, respectively. The parameter values retrieved with the initial sets S1, S4, and S5 follow

HESSD
9, 13373–13414, 2012

patterns with average values $K_S = 2.4 \times 10^{-4}$ cm s^{-1}, $\alpha = 0.043$ cm^{-1} and $n = 1.8$, which are relatively close to the values reported by Carsel and Parrish (1988) for loam soils: $Ks = 5.0 \times 10^{-4}$ cm s^{-1}, $\alpha = 0.036$ cm^{-1} and $n = 1.56$. The convergence patterns obtained with OD = 1 cm and 2 cm (Fig. 3a and b, respectively) reflect rather well these two alternative parameter space solutions.

Compared with K_S and n, parameter α 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 α, as a result of a relatively higher amount of information of αretrievable for soil water states close to the air entry value (Vrugt et al., 2001, 2002; Medina et al., 2012). In both experiments GA3 and GB1, parameter α converges predominantly toward a value of approximately 0.04 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 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 the soil hydraulic parameters with a non-sequential inverse method. Indeed, parameter α acts as a scaling factor of the pressure head values with respect to the soil moistures in the VGM model, and its identifiability with inverse methods is highly affected by the type of information employed (e.g. Simunek and van Genuchten, 1996; Ritter et al., 2004; Wöhling and Vrugt, 2011). As discussed later, differences between the estimated parameter values can be also attributed to the non uniqueness of the solution.

In Fig. 4 the soil water retention and hydraulic conductivity functions pertaining to soil core GA3 using the 18 retrieved parameter vectors, are compared with those using the parameter values optimized by 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 cm and 12 cm, have a slope similar to the

HESSD
9, 13373–13414, 2012

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 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 estimated for soil core GB1 and those obtained by Romano and Santini (1999). The estimated soil water retention curves are also shifted with respect to the reference curve, except for the dry range due to the marked difference between the residual soil moisture of the cited study ($0.12\,cm^3\,cm^{-3}$) and the value assumed in this study ($0.078\,cm^3\,cm^{-3}$). 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 different residual soil moisture values.

The non uniqueness of the solution is related to the multivariate correlation structure induced in the parametric distribution, as shown in 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. 32 and 33) assuming OD = 12 cm. This is roughly the pattern observed for both experiments, independently from the adopted observation depths. The temporal reduction of $\delta(Ks)$ variance (Fig. 6a) is small compared with that of $\delta(\alpha)$ and $\delta(n)$ (Fig. 6b and c). Moreover, $\delta(Ks)$ 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 sign 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, it is interesting a closer inspection of the performance of the assimilation algorithm with relatively low assimilation frequencies. Figure 7 illustrates the temporal patterns of the retrieved parameters using AF = $1/12\,h^{-1}$ and OD = 1 cm, thus involving

HESSD
9, 13373–13414, 2012

only 15 assimilation events for GA3 and 11 for GB1. There is a very good agreement between these patterns and the analogous patterns using AF = 1/2 h^{-1} (Figs. 2a and 3a). The identifiability of *Ks* is partially affected, while that of α and n is almost unchanged. Also the covariance and correlation structure follow the general trends previously described and the effect of the negative correlation between α and n can be visually perceived. For example, notice 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 α.

In summary, all these outcomes demonstrate the good performance of the proposed approach in terms of parameter retrieving, despite the limited amount of explored observations, the error embedded as part of the experimental data, and the wide spectrum of conditions considered. In general, by using both OD = 1 cm and OD = 2 cm, it is possible to efficiently identify sets of parameters similar to those obtained by assimilating the entire soil moisture profile.

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 retrieving algorithm can in principle sample parametric solutions involving a "wrong" saturation, i.e. giving place to state values being higher than θ_s, or "wrong" dry conditions, i.e. giving place to state values being smaller than θ_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 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. 18), 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

HESSD
9, 13373–13414, 2012

"temporarily adaptive" θ_s (or in principle θ_r), i.e. making θ_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 θ_s as an additional unknown parameter to be retrieved would not either avoid this issue, but would rather make it more evident.

5.2 State retrieving

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 two hours, within the three observation depths examined (1, 2 and 12 cm). The results show that dual filter algorithm is generally able to retrieve the true state profiles with a relatively low dependence from the identified parametric array. For OD = 1 cm using GA3, 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, 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 retrieving process is forced to deal with significant soil vertical heterogeneity, as testified by the variation of the soil water content along the soil sample (Fig. 9). This vertical variability makes the retrieving 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 at the fifth day is very small.

The comparison between the retrieving performances achieved with the observation depths OD = 1 and OD = 2 cm, provides contrasting results for the two experiments. For GA3, the differences between the soil water content profiles retrieved with OD = 1 and OD = 2 are still marked at the fourth and seventh day, whilst for GB1 the profiles are similar both at the third and the fifth day. This is more clearly appreciated in Fig. 10,

HESSD
9, 13373–13414, 2012

illustrating the temporal evolution of the ME and RMSE for both the 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 in general, the overall performance of the retrieving algorithm, is scarcely affected by the initial soil water content.

Figure 10a shows that relatively high errors occur for GA3 with OD = 1 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 cm is that water fluxes during the last stage of this experiment are very small (see Fig. 1), 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 this physical constrains impact on the dynamic of the states retrieving 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. Normally 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. 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 four 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 the information provided by the second node, while the top node is barely contributing to the retrieved value. The

HESSD
9, 13373–13414, 2012

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 retrieving value of the bottom node, almost till the end of the experiment. 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 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 retrieving 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 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 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 retrieving 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.

HESSD
9, 13373–13414, 2012

Table 4 summarises the ME and RMSE values computed for OD = 1 cm and 2 cm and AF = 1/2 h^{-1} and 1/12 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 h^{-1} produces higher errors than AF = 1/2 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 capability of the dual Kalman Filter approach in retrieving both parameters and states simultaneously. The performances of the proposed approach have been evaluated with reference to 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 θ-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 depths of 1 and 2 cm, the approach permits to properly identify 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

HESSD
9, 13373–13414, 2012

retrieved parameter α 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 provides a good performance in terms of states retrieving and proved to be able to deal even with the heterogeneity of the GB1 soil sample. 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 to 1/12 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 efficiently cope with two or three dimensional problems of soil water flux, 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 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.

References

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.: Uncertainty in predicting soil hydraulic properties at the hillslope scale with indirect methods, J. Hydrol., 334, 405–422, 2007.

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 in the Richards equation – Part 1: Retrieving state profiles with linear and nonlinear

HESSD

9, 13373–13414, 2012

numerical schemes, Hydrol. Earth Syst. Sci. Discuss., 9, 13291–13327, doi:10.5194/hessd-9-13291-2012, 2012.
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 multifrecuency remote sensed observations, IEEE T. Geosci. Remote, 32, 438–448, 1994.
Ines, A. V. M. and Mohanty, B. P.: Near-surface soil moisture assimilation for quantifying effective soil hydraulic properties using genetic algorithm: 1. Conceptual modeling, Water Resour. Res., 44, W06422, doi:10.1029/2007WR005990, 2008.
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, Hydrol. Earth Syst. Sci., 16, 3863–3887, doi:10.5194/hess-16-3863-2012, 2012.
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.
Medina, H., Romano, N. and Chirico, G. B.: Kalman filters for assimilating near-surface observations in the Richards equation – Part 2: A dual filter approach for simultaneous retrieving of states and parameters, Hydrol. Earth Syst. Sci. Discuss., 9, 13329–13372, doi:10.5194/hessd-9-13329-2012, 2012.
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.
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.
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.

HESSD
9, 13373–13414, 2012

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.

Tian, X., Xie, Z., and Dai, A.: A land surface soil moisture data assimilation system based on the dual-UKF method and the Community Land Model, J. Geophys. Res., 113, D14127, doi:10.1029/2007JD009650, 2008.

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

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

Vrugt, J. A., 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.

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.

HESSD
9, 13373–13414, 2012

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

Soil sample	Texture	ρb, g cm^{-3}	θ_s	θ_r^*	α^{**} (10^{-2} cm^{-1})	n^{**}	Ks^{**} (10^{-4} cm s^{-1})
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).

HESSD
9, 13373–13414, 2012

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

Input variable	Soil sample	
	GA3	GB1
Initial state variable	0.28 $cm^3\ cm^{-3}$	0.31 $cm^3\ cm^{-3}$
Observation depths (OD)	1, 2 and 12 cm	
Assimilation frequency (AF)	2 and 12 h	
Initial state covariance matrix $\mathbf{P}^x_{i,i}$; $i = 1 \ldots N_{nod}$	0.8 $cm^6\ cm^{-6}$	
Initial normalized correction terms matrix $\mathbf{P}^w_{i,i}$; $i = 1 \ldots N_{par}$	0.01	
Process-noise updating $\mathbf{R}_{v_{i,i}}$; $i = 1 \ldots N_{nod}$	0.05 x_i $cm^6\ cm^{-6}$	
Observation noise updating $\mathbf{R}_{n_{i,i}}$; $i = 1 \ldots N_{obs}$	0.02 y_i $cm^6\ cm^{-6}$	

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

HESSD
9, 13373–13414, 2012

Table 3. Values of w_{min} and $w_{max} - w_{min}$ used to constrain the distribution of the parameters *Ks*, α 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
Ks ($cm\,s^{-1}$)	1×10^{-5}	6×10^{-4}	4.6×10^{-4}	4.6×10^{-4}	3.1×10^{-4}	3.1×10^{-4}	1.6×10^{-4}	1.6×10^{-4}
α (cm^{-1})	1×10^{-3}	5×10^{-2}	2.6×10^{-2}	1.35×10^{-2}	3.85×10^{-2}	1.35×10^{-2}	2.6×10^{-2}	3.85×10^{-2}
n (–)	1.1	2.0	1.6	2.1	1.6	2.6	2.6	2.1

HESSD
9, 13373–13414, 2012

Table 4. Mean error (ME) and RMSE (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 h^{-1} and AF = 1/12 h^{-1}).

		GA3				GB1			
AF	Initial	OD = 1 cm		OD = 2 cm		OD = 1 cm		OD = 2cm	
	set	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE
1/2 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 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

HESSD
9, 13373–13414, 2012

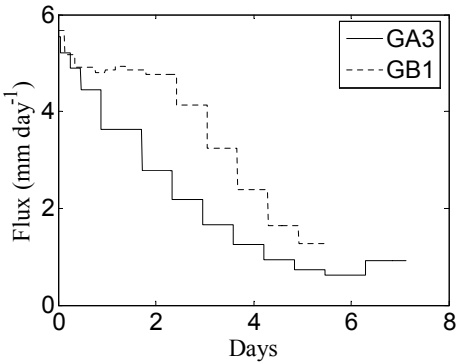

Fig. 1. Evaporation flux at the soil surface calculated from the water balance of each soil sample between consecutive measurements of the soil profiles.

HESSD
9, 13373–13414, 2012

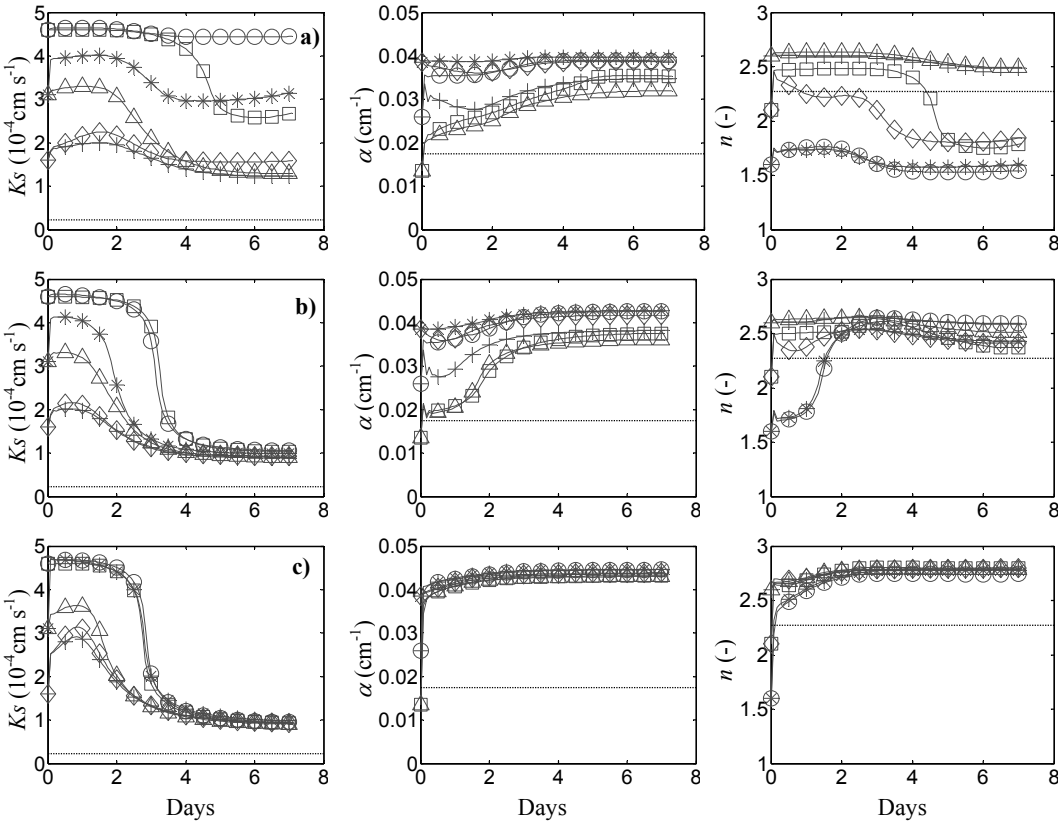

Fig. 2. Retrieved VGM parameters *Ks*, α and *n* for GA3 soil sample with assimilation frequency AF = 1/2 h^{-1}, and observation depth **(a)** OD = 1 cm, **(b)** OD = 2 cm and **(c)** OD = 12 cm. Comparisons account for the six pondered sets of initial parameters S1(◯), S2(□), S3(∗), S4(Δ), S5(+) and S6(◊). The dotted line indicates the value of the parameter found by Romano and Santini (1999).

HESSD
9, 13373–13414, 2012

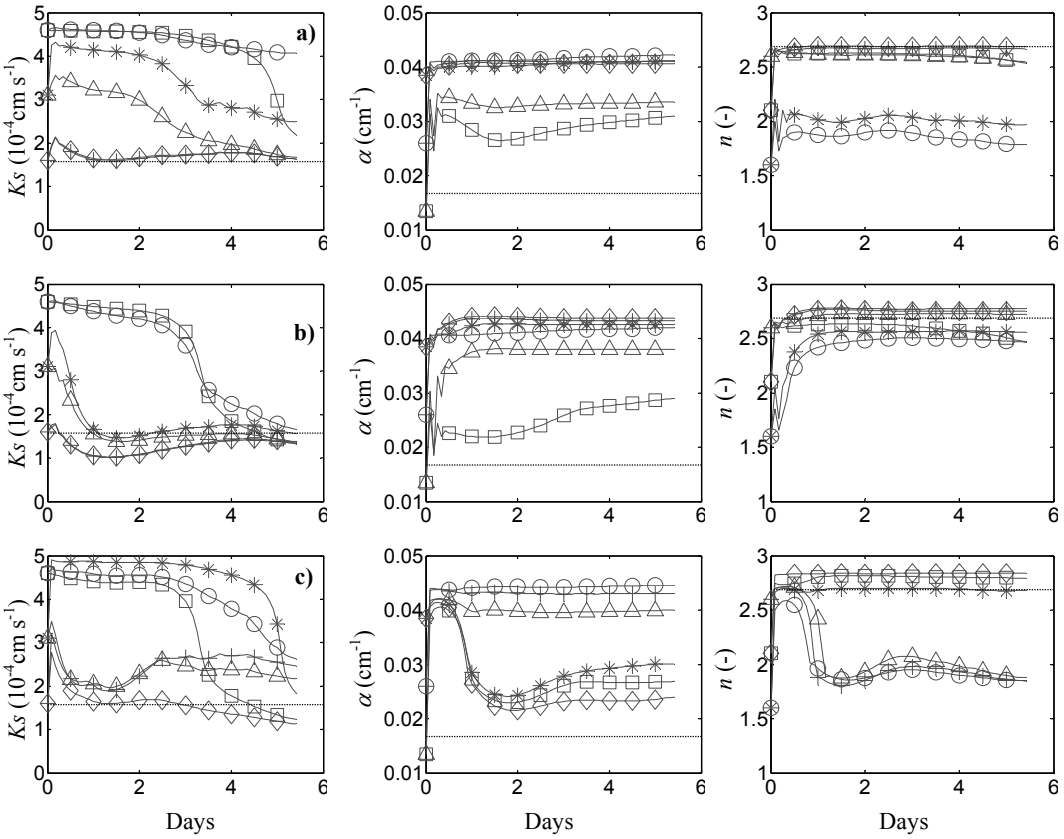

Fig. 3. Retrieved VGM parameters *Ks*, α and *n* for GB1 soil sample with assimilation frequency AF = 1/2 h^{-1}, and observation depth **(a)** OD = 1 cm, **(b)** OD = 2 cm and **(c)** OD = 12 cm. Comparisons account for the six pondered sets of initial parameters S1(◯), S2(□), S3(∗), S4(Δ), S5(+) and S6(◊). The dotted line indicates the value of the parameter found by Romano and Santini (1999).

HESSD
9, 13373–13414, 2012

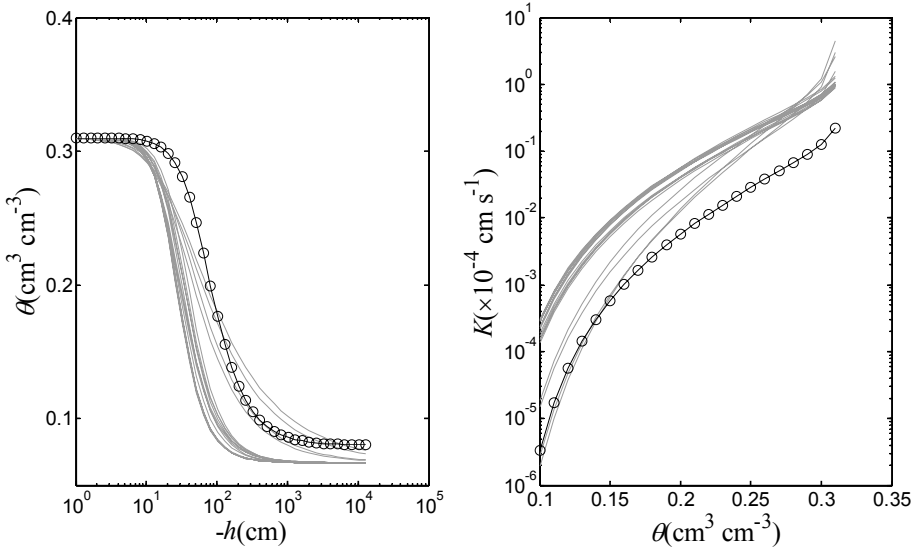

Fig. 4. Comparison of the 18 (corresponding to 3 observation depths times 6 initial parameter sets) soil water retention curves $\theta(h)$, and hydraulic conductivity functions, $K(\theta)$, using the converging parameters for GA3 (gray solid lines). The solid lines with markers indicate the corresponding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

HESSD
9, 13373–13414, 2012

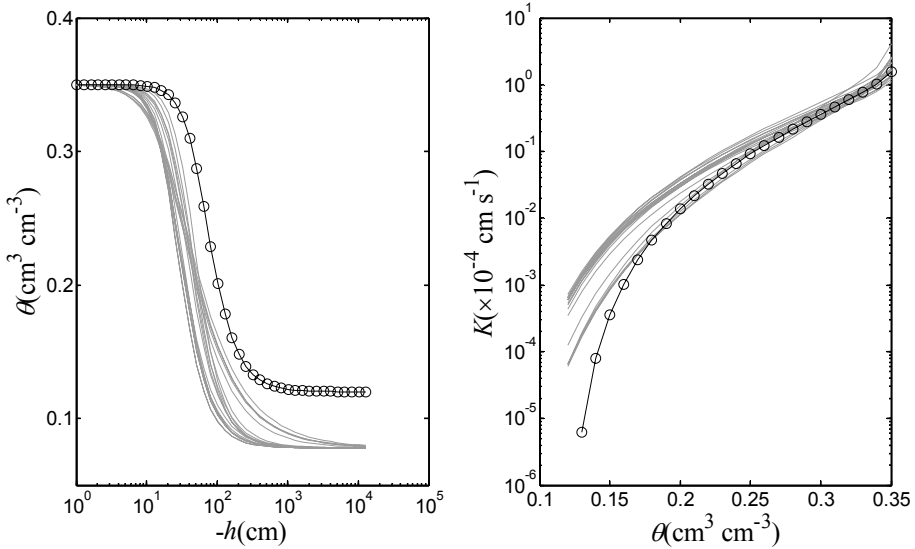

Fig. 5. Comparison of the 18 (corresponding to 3 observation depths times 6 initial parameter sets) soil water retention curves $\theta(h)$, and hydraulic conductivity functions, $K(\theta)$, using the converging parameters for GB1 (gray solid lines). The solid lines with markers indicate the corresponding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

HESSD
9, 13373–13414, 2012

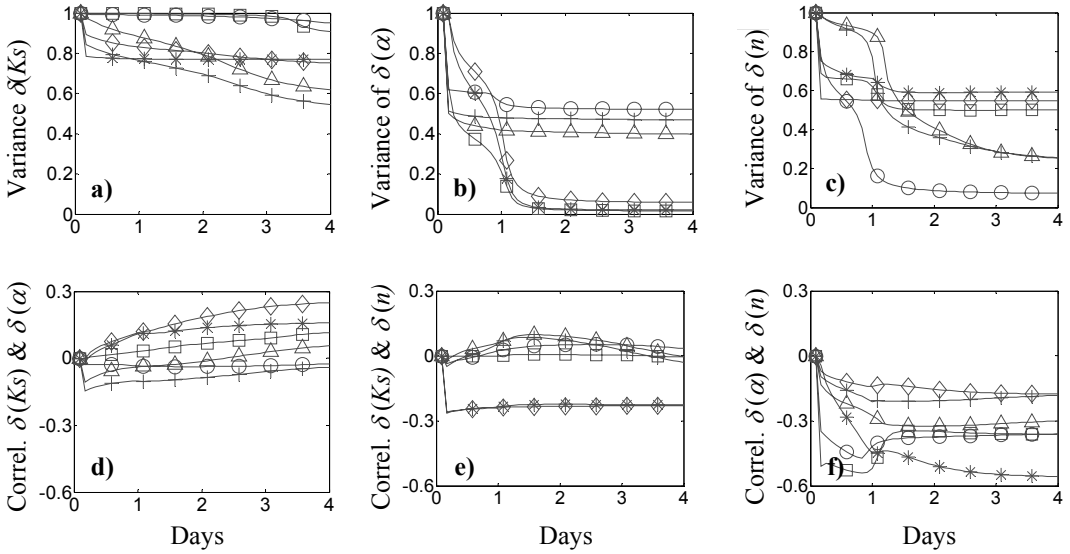

Fig. 6. Evolving variances of the correction terms **(a)** δ (Ks), **(b)** $\delta(\alpha)$ and **(c)** $\delta(n)$ associated with the VGM parameters and correlations **(d)**–**(e)** between these terms during the first four days using the GA3 data series, with assimilation frequency AF = 1/2 h^{-1}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

HESSD
9, 13373–13414, 2012

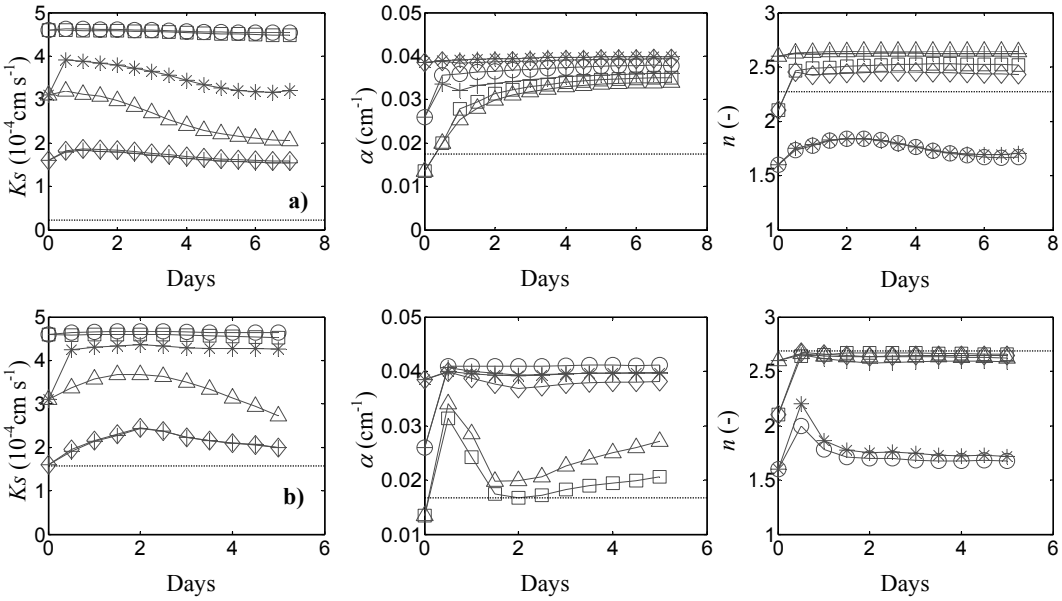

Fig. 7. Retrieved VGM parameters *Ks*, α and *n* using **(a)** GA3 and **(b)** GB1 experimental data, with assimilation frequency AF = 1/12 h^{-1} and observation depth OD = 1 cm. Comparisons account for the six pondered sets of initial parameters S1(○), S2(□), S3(∗), S4(Δ), S5(+) and S6(◊). The dotted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

HESSD
9, 13373–13414, 2012

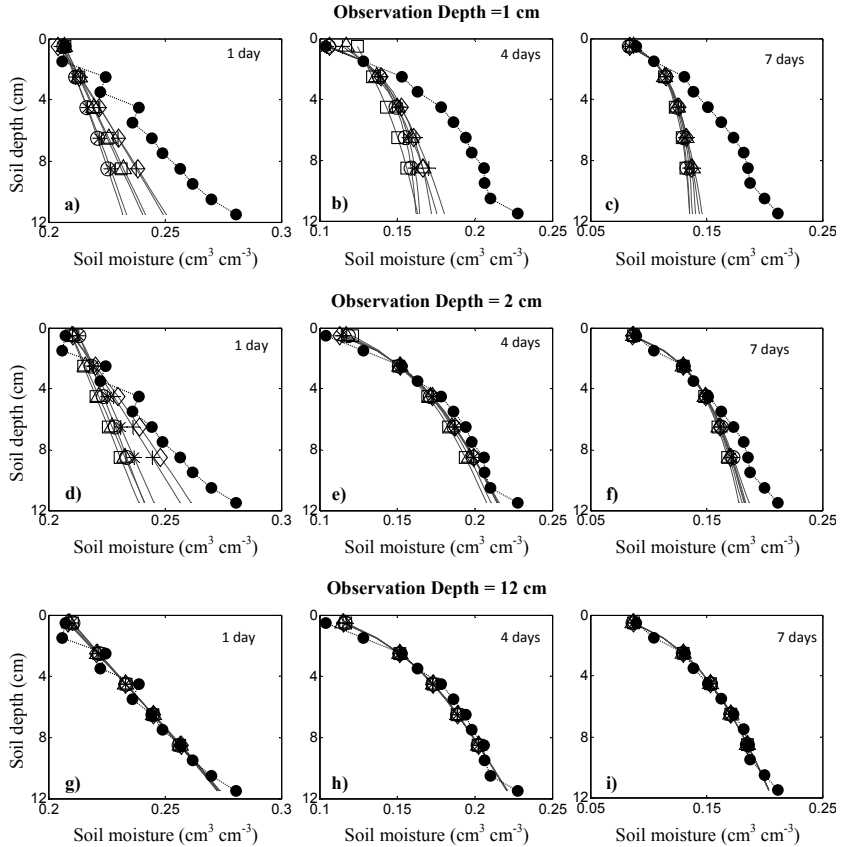

Fig. 8. Retrieved soil moisture profiles (solid lines) using the GA3 data series after 1 day **(a, d, g)**; 4 days **(b, e, h)** and 7 days **(c, f, i)**, with assimilation frequency AF = 1/2 h^{-1} and observation depths: **(a)**–**(c)** OD = 1 cm; **(d)**–**(f)** OD = 2 cm; **(g)**–**(i)** OD = 12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

HESSD
9, 13373–13414, 2012

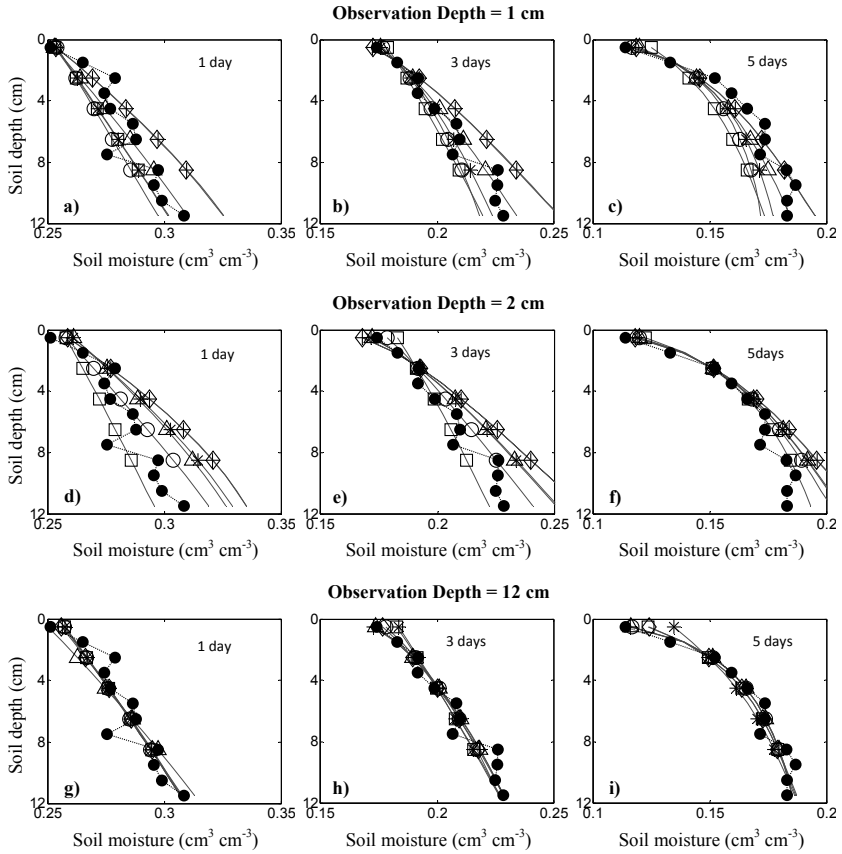

Fig. 9. Retrieved soil moisture profiles (solid lines) using the GB1 data series after 1 day **(a, d, g)**; 4 days **(b, e, h)** and 7 days **(c, f, i)**, with assimilation frequency AF = 1/2 h^{-1} and observation depths: **(a)–(c)** OD = 1 cm; **(d)–(f)** OD = 2 cm; **(g)–(i)** OD = 12 cm. Comparisons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

HESSD
9, 13373–13414, 2012

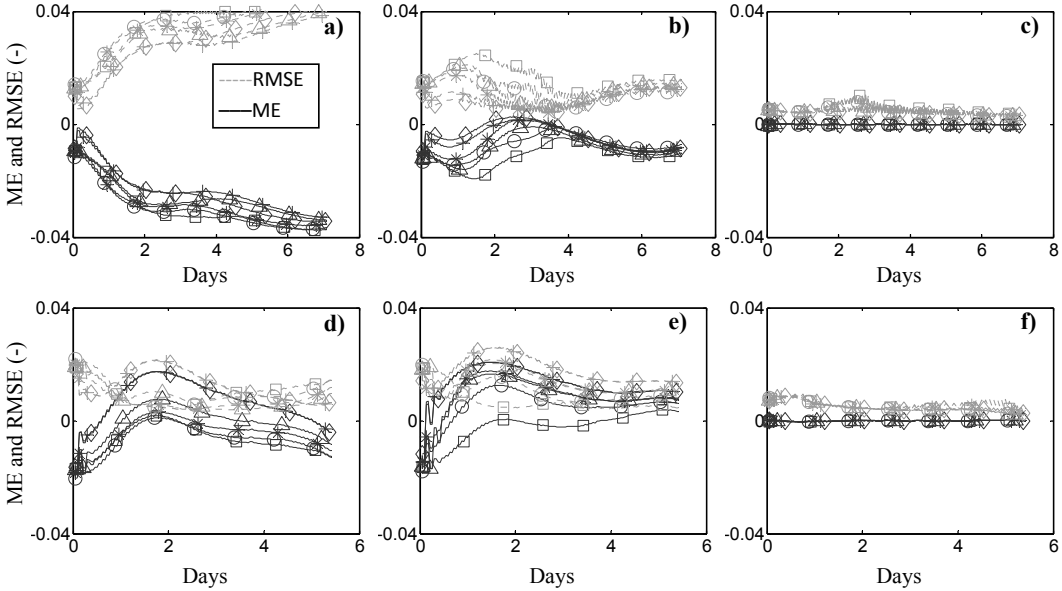

Fig. 10. Evolving mean errors (ME) and root mean square errors (RMSE) between predicted and measured profiles using GA3 **(a)–(c)** and GB1 **(d)–(f)** experimental data series, with assimilation frequency AF = 1/2 h^{-1}, and observation depths: **(a, d)** OD = 1 cm; **(b, e)** OD = 2 cm; **(c, f)** OD = 12 cm. The analysis accounts for the six pondered sets of initial parameters S1(◯), S2(□), S3(∗), S4(Δ), S5(+) and S6(◊).

HESSD
9, 13373–13414, 2012

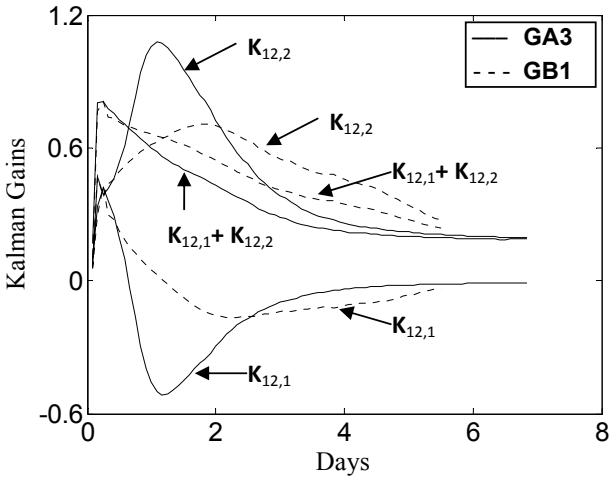

Fig. 11. Evolving Kalman gain coefficients $K_{12,1}$, $K_{12,2}$ and the sum $K_{12,1} + K_{12,2}$ for GA3 and GB1 using observation depths of 1 cm and the parameter initialization S6. $K_{12,1}$ and $K_{12,2}$ describes the influence of the first and second observation nodes, respectively, on node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

HESSD
9, 13373–13414, 2012

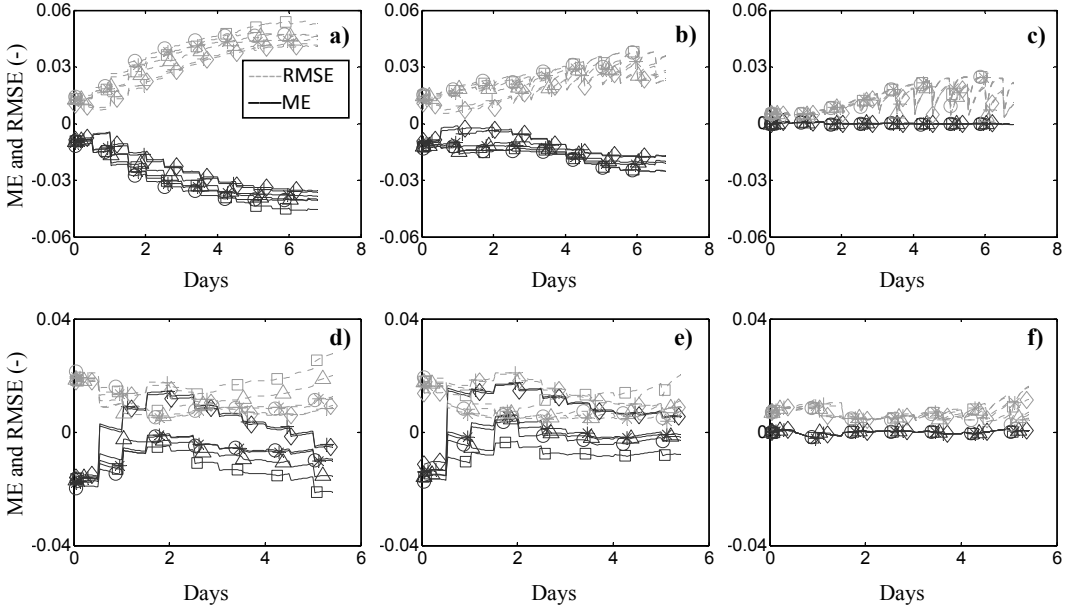

Fig. 12. Evolving mean errors (ME) and root mean square errors (RMSE) between predicted and measured profiles using GA3 **(a)–(c)** and GB1 **(d)–(f)** experimental data series, with assimilation frequency AF = 1/12 h^{-1}, and observation depths: **(a, d)** OD = 1 cm; **(b, e)** OD = 2 cm; **(c, f)** OD = 12 cm. The analysis accounts for the six pondered sets of initial parameters S1(◯), S2(□), S3(∗), S4(Δ), S5(+) and S6(◊).

HESSD
9, 13373–13414, 2012