Impact of sensor failure on the observability of flow dynamics at the Biosphere 2 LEO hillslopes

D. Pasetto[a,*], G.-Y. Niu[b,c], L. Pangle[b], C. Paniconi[d], M. Putti[a], P. A. Troch[b,c]

[a]*Department of Mathematics, University of Padua, Padua, Italy*
[b]*Biosphere 2 - University of Arizona, Tucson, USA.*
[c]*Department of Hydrology and Water Resources, University of Arizona, Tucson, USA.*
[d]*Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ETE), Quebec City, Canada.*

Abstract

The Biosphere 2 Landscape Evolution Observatory (LEO) has been developed to investigate hydrological, chemical, biological, and geological processes in a large-scale, controlled infrastructure. The experimental hillslopes at LEO are instrumented with a large number of different sensors that allow detailed monitoring of local and global dynamics and changes in the hydrological state and structure of the landscapes. Sensor failure, i.e., a progressive reduction in the number of active or working sensors, in such an evolving system can have a dramatic impact on observability of flow dynamics and estimation of the model parameters that characterize the soil properties. In this study we assess the retrieval of the spatial distributions of soil water content and saturated hydraulic conductivity under different scenarios of heterogeneity (different values of correlation length of the random field describing the hydraulic conductivity) and a variable number of active sensors. To avoid the influence of model structural errors and measurement bias, the analysis is based on a synthetic representation of the first hydrological experiment at LEO simulated with the physically-based hydrological model CATHY. We assume that the true hydraulic conductivity is a particular random realization of a stochastic field with lognormal distribution and exponential correlation length. During the true run, we collect volumetric water content measurements at an hourly interval. Perturbed observations are then used to estimate the total water storage via linear interpolation and to retrieve the conductivity field via the ensemble Kalman filter technique. The results show that when less than 100 out of 496 total sensors are active, the reconstruction of volumetric water content may introduce large errors in the estimation of total water storage. In contrast, retrieval of the saturated hydraulic conductivity distribution allows the CATHY model to reproduce the integrated hydrological response of LEO for all sensor configurations investigated.

Keywords: Sensor failure analysis; Data assimilation; Landscape Evolutionary Observatory; Soil moisture;

Hydraulic conductivity; Richards' equation

1. Introduction

Determination of the number and location of sensors needed to monitor a real-world hydrological process is a classical problem in experimental design, where the best compromise between maximum amount of information and minimum number of sensors is sought (e.g., [1, 2]). In this framework, an aspect that is rarely taken into consideration is that sensors may fail during long-term experiments, thereby putting at risk the observability of the system since it may not always be possible to replace broken sensors. The lifetime of sensors is thus a crucial unknown in experiments of long duration, and it becomes important to be able to predict how the information obtained from the active sensors changes over time as the sensor network deteriorates.

This is the premise for the present study, which is based on the setup of the Landscape Evolution Observatory (LEO) of the Biosphere 2 facility near Tucson, Arizona. The three synthetic, controlled hillslopes at LEO were constructed with the aim of improving our predictive understanding of the coupled physical, chemical, biological, and geological processes at Earth's surface in changing climates [3]. Each hillslope is 30 m long and 11.15 m wide and has an average slope of 10 degrees. The 1 m deep soil consists of basaltic tephra, ground to homogeneous loamy sand texture. For the first years of LEO operation, vegetation is not present and the research is focused on the characterization of the hydrological response of the hillslopes in terms of water transit times, generation of seepage and overland flow, internal dynamics of soil moisture, and evaporation. The second part of the experiment envisages the presence of plants growing on the hillslopes and aims to monitor the oxygen and carbon cycles inside LEO, as well as the impact of vegetation on the spatial distribution of soil water content and on changes in the soil hydraulic properties [4, 5].

To monitor these processes, each hillslope is equipped with a dense network of soil sensors (496 locations) that measures volumetric water content (496 sensors), soil water potential, and soil temperature. These local observations of the internal state of the soil are combined with measurements of the global system response, such as the total

*Tel: +39 049 827 1319; Fax: +39 049 827 1499
Email address: damiano.pasetto@unipd.it (D. Pasetto)

weight of the infrastructure (and thus the water storage), the rate of irrigation/evaporation, and the water outflow at the seepage face. Finally, geochemical analysis of irrigation water, soil water, and seepage outflow are available to monitor solute transport processes along the hillslopes.

As sensors fail, the number of active sensors, m, will decrease in time. For example, assuming that the time of failure of a sensor, t_f, follows a Gamma distribution with shape parameter α and rate parameter β, the probability that a sensor is working at time t is

$$P_w(t) = P(t_f > t) = 1 - \int_0^t g(\tau;\alpha,\beta)\, d\tau \tag{1}$$

where $g(\,\cdot\,;\alpha,\beta)$ is the probability density function (pdf) of the Gamma distribution. With the further assumption that the times of failure of the sensors are independent and identically distributed random variables, the number of active sensors at time t has a binomial distribution with parameters $p = P_w(t)$ and $n = 496$, and the expected value of active sensors at time t is $E[m] = np$. Figure 1 shows the probability of the lifetime of a sensor, $P_w(t)$, for two possible combinations of parameters α and β (the expected value of the failure time in this example is $E[t_f] = \alpha/\beta = 10$ years).

In this study we assess the impact of the number of active sensors on the observability of the LEO hillslopes. The physically-based hydrological model CATHY [6] is employed to numerically simulate the water dynamics on the LEO landscapes. CATHY couples a finite element solver of the Richards equation for subsurface flow developed by Paniconi and Putti [7] with a surface routing scheme developed by Orlandini and Rosso [8]. Surface flow occurs along a conceptual channel network derived from the digital elevation model (DEM) of the landscape [9], and the coupling between the surface and subsurface modules is resolved via a boundary condition-based partitioning of the atmospheric inputs into soil infiltration and land surface ponding. To account for heterogeneities in the LEO soil [10], we represent the saturated hydraulic conductivity as a three-dimensional random field with a lognormal probability distribution and an anisotropic exponential covariance function.

We use two different approaches to quantitatively assess the information associated with the network of active sensors of volumetric water content. In the first approach we are interested in knowing if LEO's sensor network allows us to accurately retrieve the spatial and temporal distribution of the water content in the entire landscape. To assess the accuracy of the retrieval, we compare the integral of the computed water content over the entire domain with the measured variation of water storage in the landscape. In the second approach we assess the sensor network's ability

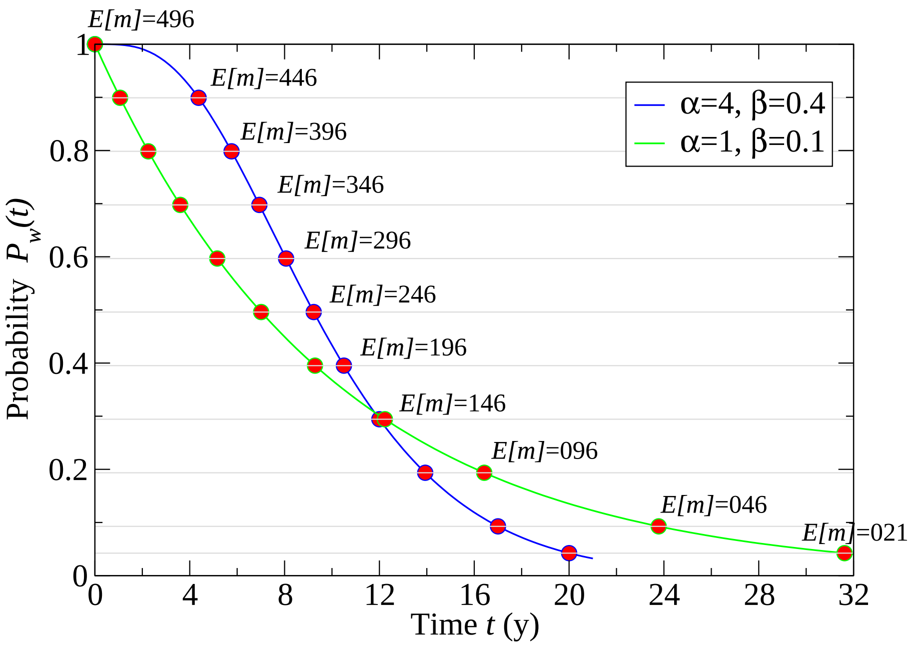

Figure 1: The probability of a sensor being active at time t for two distributions of failure time. The two red circles along each horizontal line give, for each distribution, the time at which the number of active sensors is expected to have dropped to the indicated value of $E[m]$. For instance, the expected value of active sensors is 46 ($E[m] = P_w n = 0.093 \times 496$) after 17 y for the blue distribution and after 23.8 y for the green distribution.

to allow retrieval of the saturated hydraulic conductivity of the soil, a critical parameter for numerical modeling of the future hydrological experiments at LEO. To account for parameter and measurement uncertainties, the ensemble Kalman filter (EnKF) [11–14] is used to compute the posterior probability distribution of the saturated hydraulic conductivity. EnKF performs a Gaussian approximation of sequential Bayesian inversion, thereby extending the Kalman filter to nonlinear models. The evolution in time of the state pdf is simulated using a Monte Carlo (MC) technique. The ensemble of model solutions is associated with random realizations of the unknown parameters. These MC realizations are then used in the update step to compute the covariance matrices required in the Kalman filter. Due to its straightforward implementation and its computational efficiency [15], EnKF is largely employed in engineering applications for measurement assimilation in real time. Moreover, since EnKF seeks a probability distribution of the parameters, this approach reduces the issues associated with non-uniqueness of the solution that typically occurs in inverse problems (e.g, [16]).

One of the major drawbacks of the EnKF technique is the so-called ensemble inbreeding (i.e., the strong reduction of the ensemble variance after few updates). For this reason, Drecourt et al. [17] and De Lannoy et al. [18] suggest that it is important to ensure that the ensemble spread is large enough at the assimilation time. Recent enhancements to the EnKF technique for estimation of two-dimensional stochastic parameters include introduction of a damping parameter [19] to reduce ensemble inbreeding, and covariance localization to clean the ensemble covariance matrices of spurious terms [20, 21]. Sun et al. [22, 23] combine EnKF with grid-based localization and Gaussian mixture model clustering techniques to estimate a multimodal parameter distribution. Panzeri et al. [24] couple EnKF with the ensemble moment equation of the transient groundwater flow equation to circumvent the MC simulation. Alzraiee et al. [25] compare centralized and decentralized fusion to invert the measurements generated with different pumping tests. Amongst applications of EnKF for estimating the spatial distribution of parameters in three-dimensional hydrological models, Chen and Zhang [26] showed that EnKF provides a satisfactory estimation of the three-dimensional hydraulic conductivity field assimilating measurements of pressure head in a synthetic example of saturated flow.

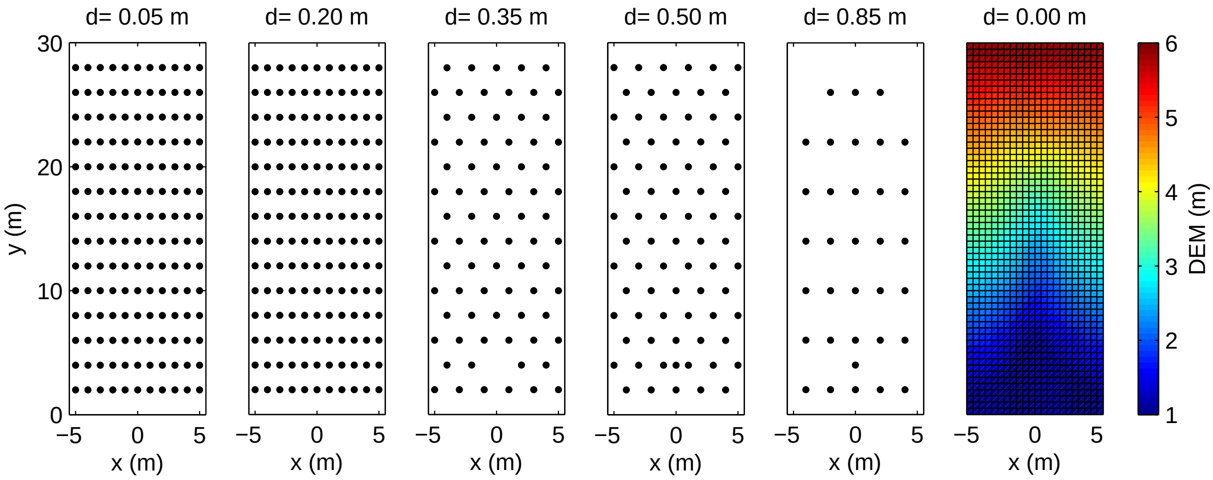

Figure 2: Digital elevation model of the surface of the hillslope domain Ω and the locations $\vec{x}_1, \ldots, \vec{x}_m$ of the m=496 sensors of water content, installed at five different depths d.

2. Problem representation

We represent the hillslope (the three LEO hillslopes are identical) as a three-dimensional domain Ω with the DEM depicted in Figure 2 and a 1 m deep soil. The bottom of the hillslope, the two side boundaries (the edges along the y axis in Figure 2), and the upslope boundary are impermeable, while the downslope boundary (at y=0 m, hereafter denoted by Γ) is the outflow face, and is modeled as a seepage face boundary condition. Let $\theta(t, \vec{x})$ be the soil water content $[-]$ at a time t $[T]$ at a point $\vec{x} = (x, y, z) \in \Omega$. Given a spatial distribution of θ at a reference time t_0=0 (initial condition), rainfall and evaporation boundary conditions are imposed at the surface, and θ responds according to this forcing term and to the soil hydraulic properties.

The dense sensor network allows the system to be monitored every 15 minutes from the reference time t_0 (times t_i with $t_i - t_{i-1}$=15 min). To test the procedure with a known distribution of water content, in this study we consider the following synthetic measurements generated with the numerical model CATHY:

1. The outgoing water volume V_Γ at the seepage face Γ $[L^3]$,

$$V_\Gamma(t_i) = \int_{t_{i-1}}^{t_i} \int_\Gamma K_S(\vec{x}) \cdot \nabla\psi(t, \vec{x}) \cdot \vec{n}_\Gamma(\vec{x})\, d\vec{x}\, dt \tag{2}$$

   where ψ is the pressure head, K_s is the saturated hydraulic conductivity tensor $[L/T]$, and $\vec{n}_\Gamma$ is the outward normal vector at Γ.

2. The water storage V_Ω $[L^3]$, computed as the time variation of the total weight P_Ω of the LEO infrastructure measured by 10 load cells:

$$V_\Omega(t_i) = \int_\Omega \theta(t_i, \vec{x})\, d\vec{x} = V_\Omega(t_0) + \gamma(P_\Omega(t_i) - P_\Omega(t_0)) \tag{3}$$

where γ is the specific weight of water.

3. The measurements of the soil water content $\mathbf{y}_i^o = \left\{\theta^o(t_i, \vec{x}_j^o)\right\}_{j=1}^{496}$ at 496 sensors (5TM Decagon probes). The sensor locations, $\vec{x}_1^o, \ldots, \vec{x}_{496}^o$, are distributed at 5 depths within the domain (0.05, 0.20, 0.35, 0.50, and 0.85 m) as shown in Figure 2.

These measurements are subject to error. Measurement error on V_Γ and V_Ω is modeled as a multiplicative random noise ϵ with a lognormal distribution, $\xi \sim LogN$, unitary expected value, $E[\xi] = 1$, and coefficients of variation cv_{V_Γ} and cv_{V_Ω} for V_Γ and V_Ω, respectively. For example:

$$V_\Omega^o(t_i) = V_\Omega(t_i)\xi(t_i), \tag{4}$$

where V_Ω^o is the observed storage, V_Ω is the real storage and

$$\log(\xi(t_i)) \sim N(-0.5\log(1 + cv_{V_\Omega}^2), \log(1 + cv_{V_\Omega}^2)). \tag{5}$$

This approach, suggested by Camporese et al. [27], guarantees the positivity of the perturbed measurements.

The volumetric water content measurement error is modeled as an additive Gaussian process with mean 0 and variance σ_θ^2, according to the calibration of the sensors:

$$\theta^o(t_i, \vec{x}_j^o) = \theta(t_i, \vec{x}_j^o) + \epsilon_j(t_i), \tag{6}$$

with

$$\epsilon_j(t_i) \sim N(0, \sigma_\theta^2)). \tag{7}$$

Note that the perturbed measurements lower than the residual moisture content, θ_r $[L^3/L^3]$, or higher than the saturated moisture content, θ_s $[L^3/L^3]$, are corrected to the range limits.

Considering only the measurements of volumetric water content, the internal observability of the system depends on the number of active sensors m and on their spatial distribution σ_m, where σ_m is a possible selection of m sensors among the initial 496 sensors. The number of possible combinations σ_m is the binomial coefficient of Newton $\begin{pmatrix} 496 \\ m \end{pmatrix}$. This is a huge number (e.g., $\begin{pmatrix} 496 \\ 446 \end{pmatrix} \approx 10^{69}$, $\begin{pmatrix} 496 \\ 21 \end{pmatrix} \approx 5 \cdot 10^{36}$). To take into account the influence of the spatial distribution of the m sensors in light of the impossibility of considering all possible combinations, we conduct the sensor failure analysis over a fixed number r of random distributions $\sigma_{m,1}, \ldots, \sigma_{m,r}$.

We propose two methodologies to assess the critical number of active sensors, m^*, that might compromise the observability of the LEO system. In the first approach we employ the measurements of volumetric water content to evaluate the spatial distribution of θ over the entire domain, thus providing an estimate of the total storage of the system. In the second approach, the measurements are employed for the calibration of the saturated hydraulic conductivity in a numerical model of LEO.

3. Estimation of volumetric water content

Local measurements of water content allow us to infer the water content $\tilde{\theta}(t_i, \vec{x})$ at any point $\vec{x} \in \Omega$ by interpolation. To ensure monotonicity between the measured values, here we compute $\tilde{\theta}(t_i, \vec{x})$ with linear interpolation

$$\tilde{\theta}(t_i, \vec{x}) = \sum_{j=1}^{m} \theta^o(t_i, \vec{x}_{\sigma_{m,r}(j)}) \phi_{\sigma_{m,r}(j)}(\vec{x}) \tag{8}$$

where $\phi_{\sigma_{m,r}(j)}(\vec{x})$ are piecewise linear interpolation functions such that $\phi_{\sigma_{m,r}(j)}(\vec{x}_l) = \delta_{\sigma_{m,r}(j),l}$. The functions $\phi_{\sigma_{m,r}(j)}$ are defined on the Delaunay triangularization associated with the location of the active sensors. The water content outside the convex hull delimited by the sensors is approximated with linear extrapolation and eventually corrected to the physical limits $[\theta_r, \theta_s]$. The subscript i in (8) refers to the measurement time, and $\sigma_{m,r}(j)$ indicates the j-th sensor among the m active sensors in the combination $\sigma_{m,r}$.

A simple control to check the reliability of the estimated water distribution at a time t_i consists in comparing the estimated water volume, $\tilde{V}_\Omega(t_i) = \int_\Omega \tilde{\theta}(t_i, \vec{x})\, d\vec{x}$, with that measured by the load cells $V_\Omega(t_i)$. If the estimated water volume $\tilde{V}_\Omega(t_i)$ falls outside the 90% confidence interval of the measure, then the distribution of sensors $\sigma_{m,r}$ is not trustworthy for the estimation of water content in Ω.

3.1. Numerical model

The CATHY (CATchment HYdrology) model [6] has been developed for simulation of water dynamics in catchments and hillslopes. The model solves the three-dimensional Richards equation describing the dynamics of pressure head ψ in variably saturated porous media:

$$S_w(\psi) S_s \frac{\partial \psi}{\partial t} + \theta_s \frac{\partial S_w(\psi)}{\partial t} = \nabla \cdot \left[K_s K_r(S_w(\psi)) \left(\nabla \psi + \eta_z \right) \right] + q \tag{9}$$

where S_s is the aquifer specific storage coefficient $[L^{-1}]$, $S_w = \theta/\theta_s$ is the saturation $[L^3/L^3]$, K_r is the relative hydraulic conductivity [-], ∇ is the gradient $[1/L]$, and $\eta_z = (0, 0, 1)^T$, with z the vertical coordinate directed upward $[L]$. The term q represents source/sink fluxes $[L^3/L^3 T]$ internal to the domain Ω or forcing terms that control the fluxes along the domain boundary $\partial\Omega$.

Retention curves establish a one-to-one relationship between the pressure head ψ and the water content θ (hysteresis is not considered in this work). The characteristics of the porous media, such as the pore size distribution n_p [-] and the pore entry suction $1/\alpha$ $[L]$, determine the local shape of these curves. Here the van Genuchten curves [28] are employed to model the water content and the relative hydraulic conductivity K_r as a function of ψ:

$$\theta(\psi) = \begin{cases} \theta_r + \dfrac{\theta_s - \theta_r}{(1 + (\alpha|\psi|)^{n_p})^{m_p}} & \text{if } \psi < 0, \\ 1 & \text{if } \psi \geq 0; \end{cases} \tag{10}$$

$$K_r(\psi) = \begin{cases} \sqrt{1 + (\alpha|\psi|)^{n_p}} (1 - (1 - \dfrac{\theta(\psi) - \theta_r}{\theta_s - \theta_r})^{m_p})^2 & \text{if } \psi < 0, \\ 1 & \text{if } \psi \geq 0; \end{cases} \tag{11}$$

where $m_p = 1 - 1/n_p$. If saturation or infiltration excess runoff occurs, CATHY couples the Richards equation solver with a one-dimensional diffusion wave approximation of the Saint-Venant equation for overland flow routing. The mass balance at the surface/subsurface interface is enforced by a boundary condition switching algorithm described in [6].

CATHY can simulate the main outputs measured at LEO, such as the seepage face outflow V_Γ, the water storage V_Ω, and, in case of overland flow, the hydrograph at the outlet of the landscape.

3.2. State space model

After discretizing the domain Ω with n nodes and e tetrahedral elements, the numerical solution of equation (9) is obtained by way of the finite element method (FEM) with piecewise linear basis functions [6]. The time-integration is performed with the backward Euler method combined with Picard or Newton iterations. To simplify the notation, in the following the CATHY state-space model is described considering only the subsurface processes. A more general notation should include in the state vector also the surface discharge and the ponding water volumes, as described by Pasetto et al. [29]. We indicate with

$$\mathbf{x}_i = \{\psi_k(t_i)\}_{k=1}^{n} \tag{12}$$

the numerical solution, i.e., the vector of pressure heads ψ_k at the location of the grid nodes at time t_i. The state vector of the model, $\mathbf{x}_i$, can be formally expressed as a nonlinear function, $\mathcal{F}$, of the pressure head $\mathbf{x}_{i-1}$ at the previous observation time, of the forcing term q_{i-1}, and of the vector $\boldsymbol{\gamma}$ representing the soil properties (in our case the saturated hydraulic conductivity K_S):

$$\mathbf{x}_i = \mathcal{F}\left(\mathbf{x}_{i-1}, q_{i-1}, \boldsymbol{\gamma}\right). \tag{13}$$

The vector

$$\mathbf{y}_i = \left\{\theta(t_i, \vec{x}^{o}_{\sigma_m(j)})\right\}_{j=1}^{m} \tag{14}$$

collects the numerical observations of the water content. This vector is a nonlinear function $\mathcal{H}$ of the state vector $\mathbf{x}_i$, described through the van Genuchten relation (10):

$$\mathbf{y}_i = \mathcal{H}\left(\mathbf{x}_i, \boldsymbol{\gamma}\right). \tag{15}$$

Equations (13) and (15) define the state-space model for the numerical simulations with CATHY.

4. Estimation of saturated hydraulic conductivity

The numerical simulation of LEO using CATHY with adequate initial and boundary conditions results in the computation of θ on the entire domain Ω, and thus represents an alternative to the estimation of water content via interpolation of the measurements. A key issue in the use of distributed models such as CATHY is the identification of the model parameters that allow us to correctly describe the system and retrieve the measurements. In this context the sensor network is fundamental for calibration of the numerical model and improving the forecast of the system state variables.

4.1. Data assimilation for inverse problems

The large number of sensors in LEO is useful in an inverse modeling framework, i.e., to infer the spatial distribution of the soil hydraulic properties to be adopted in equations (9), (10), and (11). The inverse problem can be stated as follows: find the parameter vector, $\boldsymbol{\gamma}_p$, that minimizes the objective function given by the squared error between the observed and the simulated measurements:

$$\boldsymbol{\gamma}^p = \arg\min_{\gamma \in \mathcal{K}} \sum_{i=1}^{i_F} \|\mathbf{y}_i - \mathbf{y}_i^o\|^2 \tag{16}$$

where i_F is the total number of measurement times in the inversion experiment and $\mathcal{K}$ is the search space where we look for the optimal solution. The numerical minimization of such an objective function is impractical for large state-space models and mainly depends on the dimension of the search space $\mathcal{K}$.

As an alternative, we consider sequential data assimilation methods (filtering) for parameter estimation [12, 30]. Starting from a prior pdf of the parameters, $p_0(\boldsymbol{\gamma})$, the joint pdf of the state variables and the parameters is sequentially updated each time that an observation becomes available (assimilation time). Using the Bayes formula we obtain

$$p^a(\mathbf{x}_i, \boldsymbol{\gamma}_i \,|\, \left\{\mathbf{y}_j\right\}_{j=1}^{i}) = C\mathcal{L}(\mathbf{y}_i \,|\, \mathbf{x}_i, \boldsymbol{\gamma}_i) p^f(\mathbf{x}_i, \boldsymbol{\gamma}_i \,|\, \left\{\mathbf{y}_j\right\}_{j=1}^{i-1}) \tag{17}$$

where p^a and p^f are the analysis and forecast pdfs, $\mathcal{L}$ is the likelihood function, and C is a normalization constant. The forecast pdf, $p^f(\mathbf{x}_i, \boldsymbol{\gamma}_i \,|\, \left\{\mathbf{y}_j\right\}_{j=1}^{i-1})$, represents the evolution in time of the previous analysis, $p^a(\mathbf{x}_{i-1}, \boldsymbol{\gamma}_{i-1} \,|\, \left\{\mathbf{y}_j\right\}_{j=1}^{i-1})$, and is computed from the model equation (13) (see, e.g., [31]). The successive computation of the forecast and analysis pdfs represents the filtering problem. In the case of a linear state-space model with additive and Gaussian noise, the analysis and forecast pdfs are Gaussian and the Kalman filter [32] directly computes the expected values and the covariances of these distributions by minimising the variance of the analysis pdf.

For a nonlinear state-space model with update of both state vector and parameters, such as the one defined in (13) and (15), the filtering solution requires the use of random realizations to discretize the pdfs of interest and approximate their evolution in time.

4.2. *Ensemble Kalman filter*

In the EnKF method, the Kalman gain is still used to evaluate the analysis step and the relative expected values and covariance matrices. The optimality of the original Kalman filter (minimum variance) is lost and the Gaussian hypothesis describes the analysis pdf. The filter is initialized with an ensemble of N random samples of the prior distribution of the state, $\{\mathbf{x}_0^{a,j}\}_{j=1}^N \sim p(\mathbf{x}_0)$ and parameters $\{\boldsymbol{\gamma}_0^{a,j}\}_{j=1}^N \sim p(\gamma_0)$. The empirical distribution of the numerical solutions $\{\mathbf{x}_i^{f,j}\}_{j=1}^N$,

$$\mathbf{x}_i^{f,j} = \mathcal{F}\left(\mathbf{x}_{i-1}^{a,j}, q_{i-1}, \boldsymbol{\gamma}_{i-1}^{a,j}\right), \tag{18}$$

approximates the forecast step, while the parameters are assumed constant in the time interval $[t_{i-1}, t_i]$, i.e., $\boldsymbol{\gamma}_i^{f,j} = \boldsymbol{\gamma}_{i-1}^{a,j}$. The variable q_{i-1} represents the forcing term during the forecast step. In the analysis step of EnKF, both the state vector and the parameters are updated with the augmented state technique:

$$\begin{pmatrix} \mathbf{x}_i^{a,j} \\ \\ \boldsymbol{\gamma}_i^{a,j} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_i^{f,j} \\ \\ \boldsymbol{\gamma}_i^{f,j} \end{pmatrix} + \mathbf{K}_i^f \left(\mathbf{y}_i^{o,j} - \mathbf{y}_i^j\right) \tag{19}$$

where $\mathbf{y}_i^j = \mathcal{H}(\mathbf{x}_i^{f,j})$. The vector $\mathbf{y}_i^{o,j}$ represents random perturbations of the observed measurements $\mathbf{y}_i^o$ according to (7). The use of the perturbed observations $\mathbf{y}_i^{o,j}$ theoretically guarantees that, in the Kalman filter hypothesis, the correct variance of the updated variables will be computed (see, e.g., [33]). K_i^f is an ensemble approximation of the Kalman filter taking into account the cross correlations between the augmented state and the observations (for more details see, e.g., [27] and [29]). The empirical distribution of the parameters obtained at the last assimilation time, $\left\{\gamma_{i_F}^{a,j}\right\}$, represents the posterior distribution of the parameters, which depends on the number of active sensors m and their location in the domain Ω.

The inverse methodology presented in this section employs the measurements at the active sensors to estimate the posterior distribution of the hydraulic conductivity field at LEO. Decreasing the number of active sensors might result in high uncertainty associated with the parameters or in unrealistic parameter estimates. In a synthetic scenario we can directly compare the true and the posterior distributions of the conductivity field. Since this approach is not possible in real applications, to understand if the number of active sensors is sufficient to obtain an accurate posterior distribution of the parameters, we re-run the simulation (without assimilation) employing the posterior distribution of the parameters. If the numerical outputs are inside the confidence interval of the measurements, then we can conclude that the estimated parameters are admissible. In this case we can say that the number of active sensors is sufficient for the calibration of the model and for the observability of the system.

5. Numerical experiments

For the estimation of water content and the retrieval of the saturated conductivity we consider only synthetic measurements generated with the numerical model CATHY. This allows evaluation of the accuracy and robustness of the proposed identification approaches. We describe below the setup of the numerical experiment used to obtain the measurement set.

The LEO hillslope domain Ω is discretized with 22×60=1320 square cells of dimension 0.5×0.5 m at the surface and 20 vertical layers of thickness of 0.05 m. The obtained parallelepipeds are further subdivided into six tetrahedra for the FEM solution of Eq. (9) with piecewise linear basis functions. The numerical simulations reproduce the first LEO experiment, conducted on February 18, 2013. A rainfall rate of 12 mm/h is imposed for a duration of 22 h followed by an evaporative forcing of 1 mm/d until the end of the experiment. Gevaert et al. [34] presents a detailed description of the experimental results. This experiment is a suitable platform for testing the reliability of the sensor network since the hillslope undergoes a broad range of dynamics and system states from fully unsaturated conditions to significant surface and subsurface outflows. Niu et al. [10] were able to reproduce the overland flow, seepage face flow, and total water storage responses with the CATHY model and a heterogeneous configuration of soil parameters.

Generation of the synthetic truth. We assume that the true saturated hydraulic conductivity field K_S is a realization of a stationary three-dimensional random field with a lognormal distribution,

$$Y = Log(K_S), \quad Y \sim N(\mu_Y, C_Y), \tag{20}$$

and exponential covariance function, C_Y, with vertical anisotropy:

$$C_Y(l_x, l_y, l_z) = \sigma_Y \exp\left(-\sqrt{\left(\frac{l_x}{\lambda_x}\right)^2 + \left(\frac{l_y}{\lambda_y}\right)^2 + \left(\frac{l_z}{\lambda_z}\right)^2}\right) \tag{21}$$

where l_x, l_y, and l_z are the lag distances and λ_x, λ_y, and λ_z are the integral scales in direction x, y, and z, respectively. μ_Y and σ_Y are the expected value and variance of Y. The other parameters describing the soil hydraulic properties

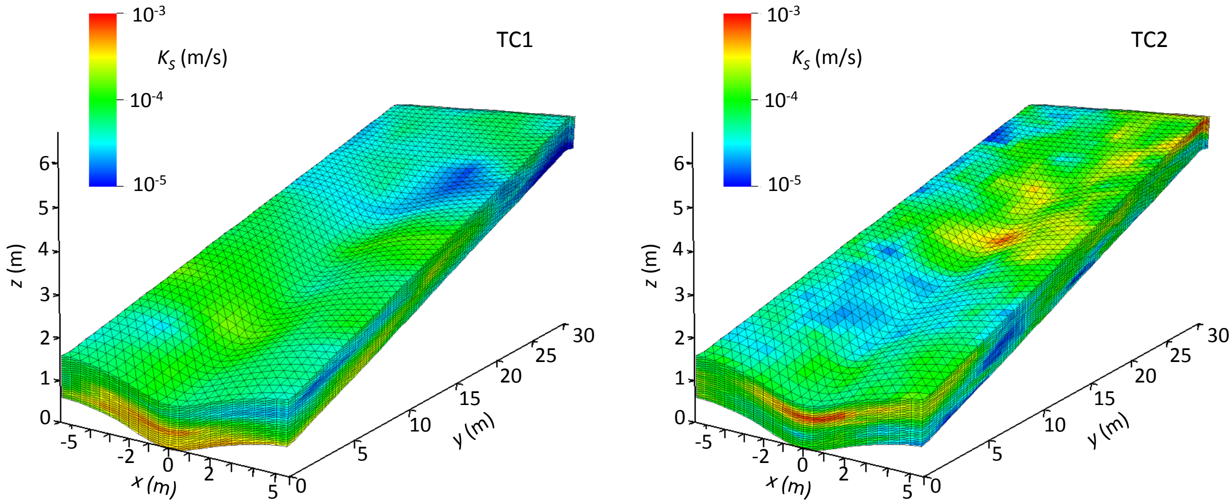

Figure 3: Spatial distribution of the hydraulic conductivity K_S in the domain Ω for the true runs of TC1 and TC2.

are considered homogeneous, and Table 1 summarizes their values. The parameters μ_Y and σ_Y are calculated using an expected value of K_S equal to $1.0{\times}10^{-4}$ m/s, in agreement with the calibration proposed in [10], and a coefficient of variation of K_S equal to 100%, corresponding to variations of K_S of two orders of magnitude. The resulting parameters of the Gaussian field Y are μ_Y=-9.56 log(m/s) and σ_Y=0.83 log(m/s). To assess the sensitivity of the results with respect to variations of the Y integral scale, which is unknown for the real LEO soil, we consider two scenarios: in test case 1 (TC1) we set $\lambda_x = \lambda_y$=8 m and λ_z=0.5 m whereas in test case 2 (TC2) the integral scales are halved, $\lambda_x = \lambda_y$=4 m and λ_z=0.25 m. A 3-dimensional adaptation of the random generator of stochastic fields HYDRO_GEN [35, 36] is employed to sample the realizations of the log-conductivity Y.

Figure 3 shows the true hydraulic conductivity fields that characterize test cases TC1 and TC2. The hydrological response associated with these two configurations of hydraulic conductivity is represented by the black straight lines shown in Figure 4 denoted as “true” state. Starting from unsaturated initial conditions ($V_\Omega(t_0) \approx$35 m^3), the rain water infiltrates in the soil and accumulates at the base of Ω, increasing the water storage. At around t = 10 h the formation of base flow generates an outgoing flux at the seepage face. Overland flow appears at the outlet after 20 h. From a qualitative point of view, these results are in accordance with the observations and simulations presented in [10]. The intensity and timing of the seepage face and overland flows are different in TC1 and TC2, due to the different soil hydraulic properties.

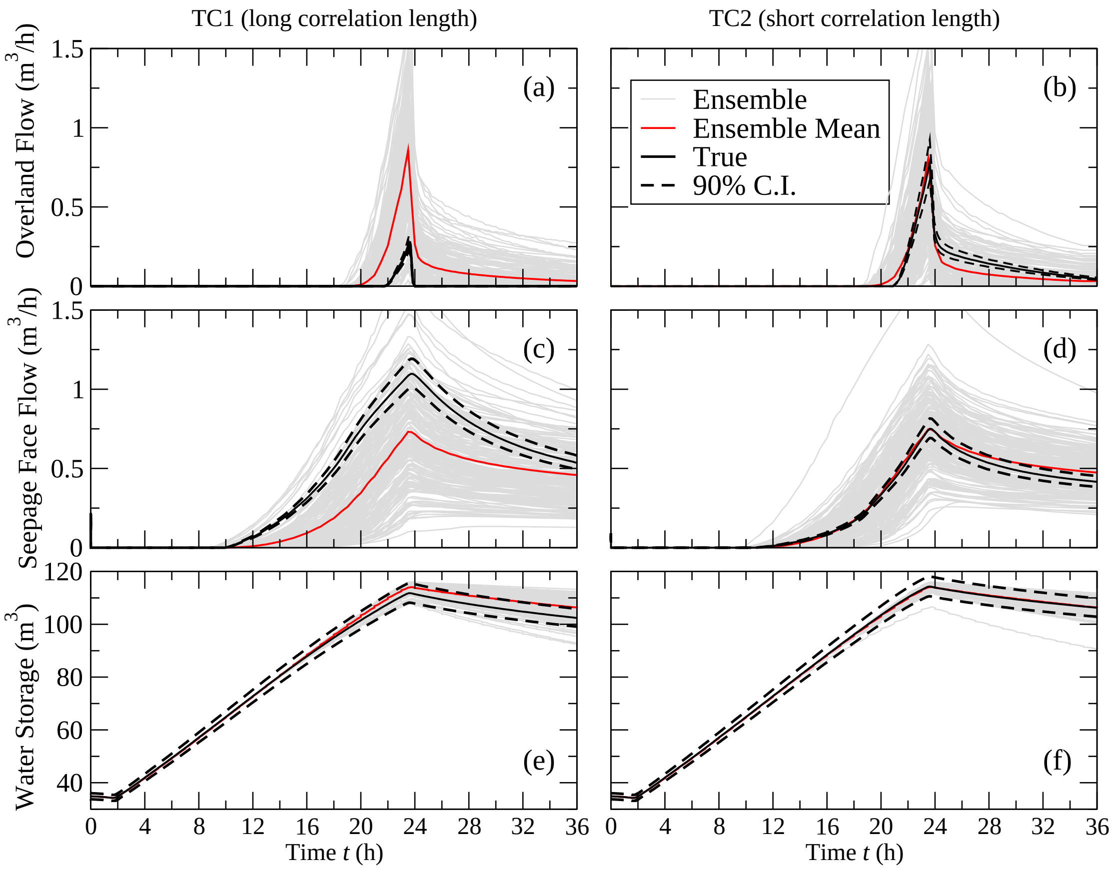

Figure 4: Hydrological response of the CATHY model for test cases TC1 and TC2. The grey lines represent the outputs obtained with 200 random realizations of the prior pdf of the saturated hydraulic conductivity and the red line is their mean. The solid black line is the response associated with the true run, and the dashed lines delimit the 90% confidence interval of the true measurments.

Grid information	
DEM cell dimensions	0.5×0.5 m
Number of cells in the surface grid	$22 \times 60 = 1320$
Soil depth (uniform)	1 m
Vertical discretization (# of layers)	20
Soil layer thickness (uniform)	0.05 m
# of nodes in the 3D grid	$23\times 61\times21=29463$
# of tetrahedral elements in the 3D grid	158400
Parameters	Heterogeneous
Saturated hydraulic conductivity	lognormally distributed
	$E[K_S] = 1.0 \times 10^{-4}$ m/s, $CV_{K_S} = 100\%$
	$Y = \log(K_S)$ has exponential covariance function
	TC1: $\lambda_x = \lambda_y = 8$ m; λ_z=0.50 m
	TC2: $\lambda_x = \lambda_y = 4$ m; λ_z=0.25 m
Parameters	Homogeneous
Aquifer specific storage	$S_s = 5 \times 10^{-4}$ m^{-1}
Porosity	$\theta_s = 0.37$
Van Genuchten curve fitting parameters	$\alpha = 0.6$ m^{-1}, $n = 2.26$
Residual moisture content	$\theta_r = 0.002$
Simulation period	129600 s (36 h)
Initial conditions	Linear interpolation of LEO water content sensors at 8:00 a.m. 02/18/2013
Atmospheric forcing	Spatially distributed rainfall (12 mm h^{-1}) and evapotranspiration (0.04 mm h^{-1})
Measurements	
Measures of volumetric water content θ	normally distributed with $\sigma_\theta \approx 0.012$
Measures of storage V_Ω	lognormally distributed with $cv_{V_\Omega} = 2\%$
Measures of seepage face volumes V_Γ	lognormally distributed with $cv_{V_\Gamma} = 2\%$
Ensemble size N	200

Table 1: Domain discretization and parameter values for the LEO sensor failure experiment. *cv* indicates the coefficient of variation.

Generation of the open loop. The open loop (OL) represents the sensitivity of the model response to variations of the parameters. We evaluate OL from a set of model solutions constructed from 200 independent Y realizations sampled from the prior pdf described above. The OL results are shown in Figure 4. Note that the response of the true realization may differ drastically from the response of the OL ensemble mean (red line), as occurring in TC1. In the TC2 case the opposite behaviour is observed, but this is just a random result. In both scenarios, the spread of the OL response reveals a large uncertainty in the model forecast, suggesting that the ensemble mean is not a statistically accurate estimator of the true response.

Generation of synthetic observations. The water content measurements employed in the interpolation and inverse problems are selected from the results of the true runs in TC1 and TC2. To simulate the measurement error, the values of water content are perturbed according to equation (7).

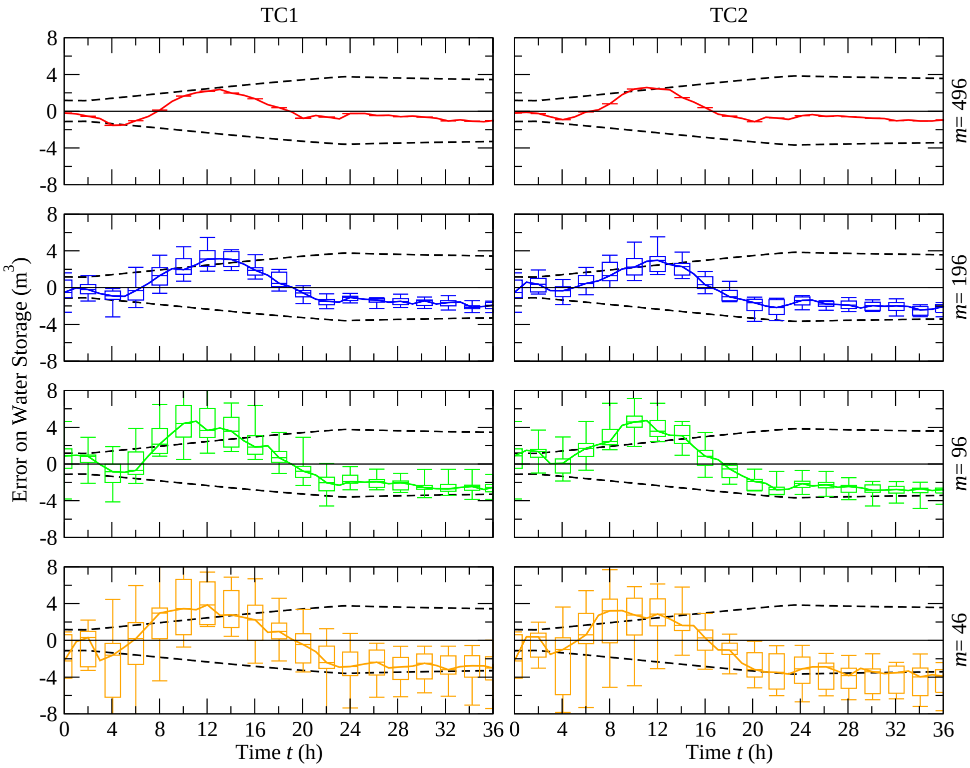

Figure 5: Error between the measured water storage in Ω and the one estimated with $\tilde{\theta}$ (equation (8)). The dashed lines represent the 90% confidence intervals, while the box plots indicate the variability of the results with respect to the spatial distribution of the sensors.

6. Results and discussion

6.1. Reconstruction of volumetric water content

In this section the perturbed measurements of volumetric water content are employed to estimate the total water volume in the entire domain Ω, as described in equation (8). The sensor failure analysis is performed considering a decreasing number of active sensors, m=[496, 446, 396, 346, 296, 246, 196, 146, 96, 46, 21], and r=10 random distributions of the active sensors in space.

Figure 5 presents the time behaviour of the error between the measured volumes (shown as the solid black line in Figure 4, panels (e) and (f)) and the estimated volumes with m sensors, m= 496, 196, 96, and 46. The box plots for the m=496, 196, 96, and 46 represent the 2.5, 25, 50, 75, and 97.5 percentiles of the errors obtained with r=10 different spatial configurations of the active sensors. The estimation of water storage using the entire sensor network (m=496) produces errors bounded by the 90% confidence interval (dashed lines), showing the accuracy of this estimate. The largest errors are recorded during the infiltration period, 2 h < t < 12 h, i.e., when the soil is still mostly unsaturated and the linear interpolation is not optimal in describing the nonlinear distribution of the infiltration front. A better estimation of V_Ω is obtained for t >20 h, i.e., when the domain is almost completely saturated. In fact, in this scenario the distribution of water content is more homogeneous, especially at the base of LEO (where sensors are sparser) resulting in more accuracy of the linear interpolation/extrapolation. Similar water volumes are estimated interpolating the measurements from less than the half of the sensors, m=196. The location of the active sensors slightly influences the estimation of water storage, especially during the infiltration period, i.e., when a large number of measurements is required to capture the water distribution in Ω. In such a period, some spatial configurations of the sensors generate an error that exceeds the limits of the 90% confidence interval, although all the placements are able to reproduce the correct volumes with sufficient accuracy during the rest of the simulation. When the number of sensors drops below 100 (m=96, m=46), the estimation of water storage becomes less reliable and highly dependent on the distribution of the active sensors. The results of the smaller sensor sets show unreliable estimates during most of the simulation. These results are not sensitive to the integral scale of the hydraulic conductivity.

Since the measurements are generated from a numerical simulation, we can compute the spatial behaviour of the error between the estimated and the true distribution of volumetric water content in Ω. Panel (a) in Figure 6 shows the spatial distribution of volumetric water content at time t=12 h for TC2. As expected, at this time the uppermost layers of the landscape are partially saturated, while base flow starts forming at the bottom of the domain. Panels (b) and (c) show the errors for the m=496 and m=46 cases of test TC2, respectively. It is evident that the error is smaller where the density of observations is higher.

To summarize the results for all the numerical simulations, we compute the time average of the L_2-norm of the error between the estimated and the true water content,

$$e_\theta = \frac{1}{i_F} \sum_{i=1}^{i_F} \left[\int_\Omega \left(\tilde{\theta}(t_i) - \theta(t_i) \right)^2 \, d\Omega \right]^{\frac{1}{2}} \tag{22}$$

Figure 7 shows the values of e_θ for all the combinations and number of active sensors considered. As expected, the water content estimation with m=496 sensors has the minimum error. The error increases and is more sensitive to the location of the sensors as the number of active sensors decreases. In particular the interpolation errors obtained with m <100 active sensors are consistently higher than the errors obtained with m >300, meaning that there is a strong

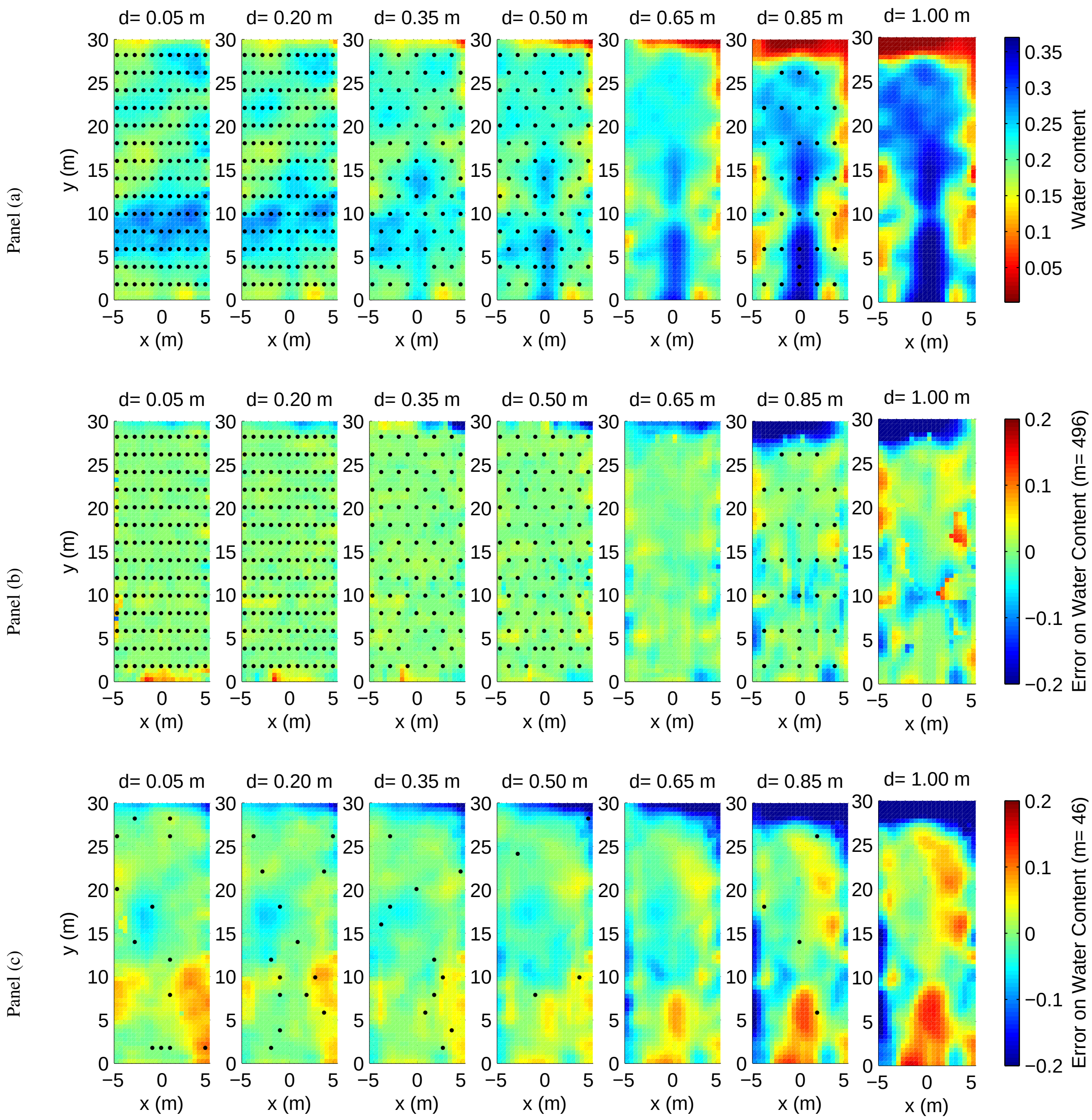

Figure 6: True volumetric water content for test TC2 after 12 h at seven depths (a) and the errors associated with the estimated volumetric water content interpolating the measurements from m=496 (b) and m=46 (c) active sensors. The black dots show the locations of the active sensors.

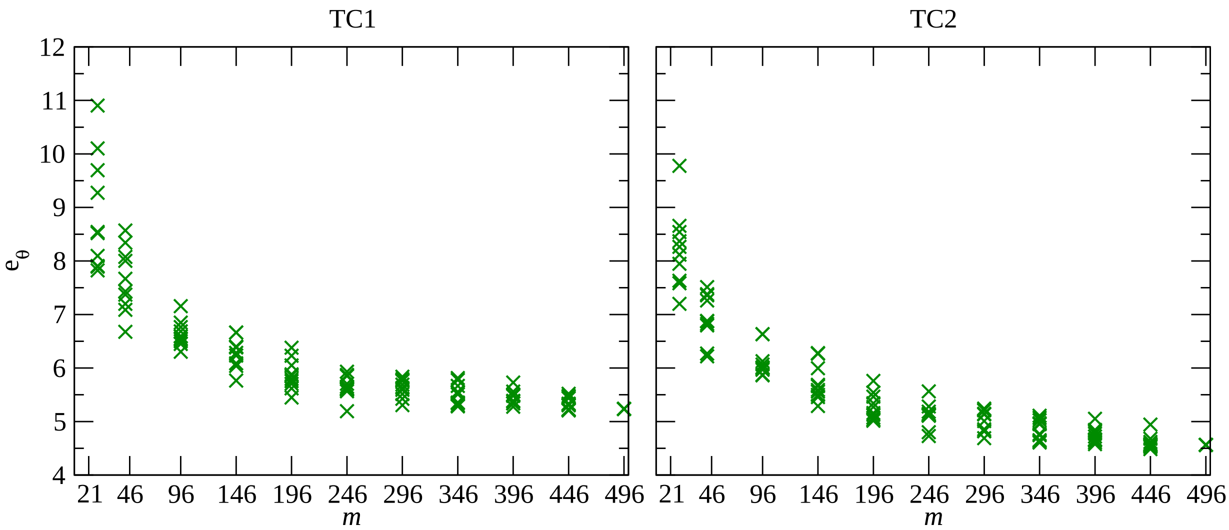

Figure 7: Time-averaged L^2-norm of the error between the true and estimated water contents (equation (22)) for the different number and configurations of active sensors.

deterioration in the observability of the system.

6.2. Reconstruction of hydraulic conductivity

In Figure 4 we show that the response of the hydrological model CATHY is highly sensitive to the different samples of the hydraulic conductivity field. It is evident that the prior distribution of K_S entails high uncertainty in the model results, implying that the model, with an incorrect parameterization, is not suitable for the prediction of the water dynamics in the landscape. Here we assimilate the measurements of volumetric water content using EnKF to obtain a posterior distribution of the hydraulic conductivity field that is closer to the realization adopted in the true runs. The sensor failure analysis is performed considering a decreasing number of active sensors, m=[496, 196, 146, 96, 46, 21], and one random distribution of the active sensors in space (r=1) .

At every assimilation time EnKF updates the model state variables (i.e., pressure head and, if present, surface discharge) and the hydraulic conductivity field associated with each MC realization. Thus, we expect that the discrepancy between the estimated log-conductivity field, $\tilde{Y}$, and the true field, Y, decreases during the filtering process. We consider the root mean squared error between the ensemble of log-conductivity realizations and the true log-conductivity,

$$e_Y(t) = E\left[\frac{1}{\Omega}\int_\Omega \left(\tilde{Y}(t) - Y\right)\, d\Omega\right] \tag{23}$$

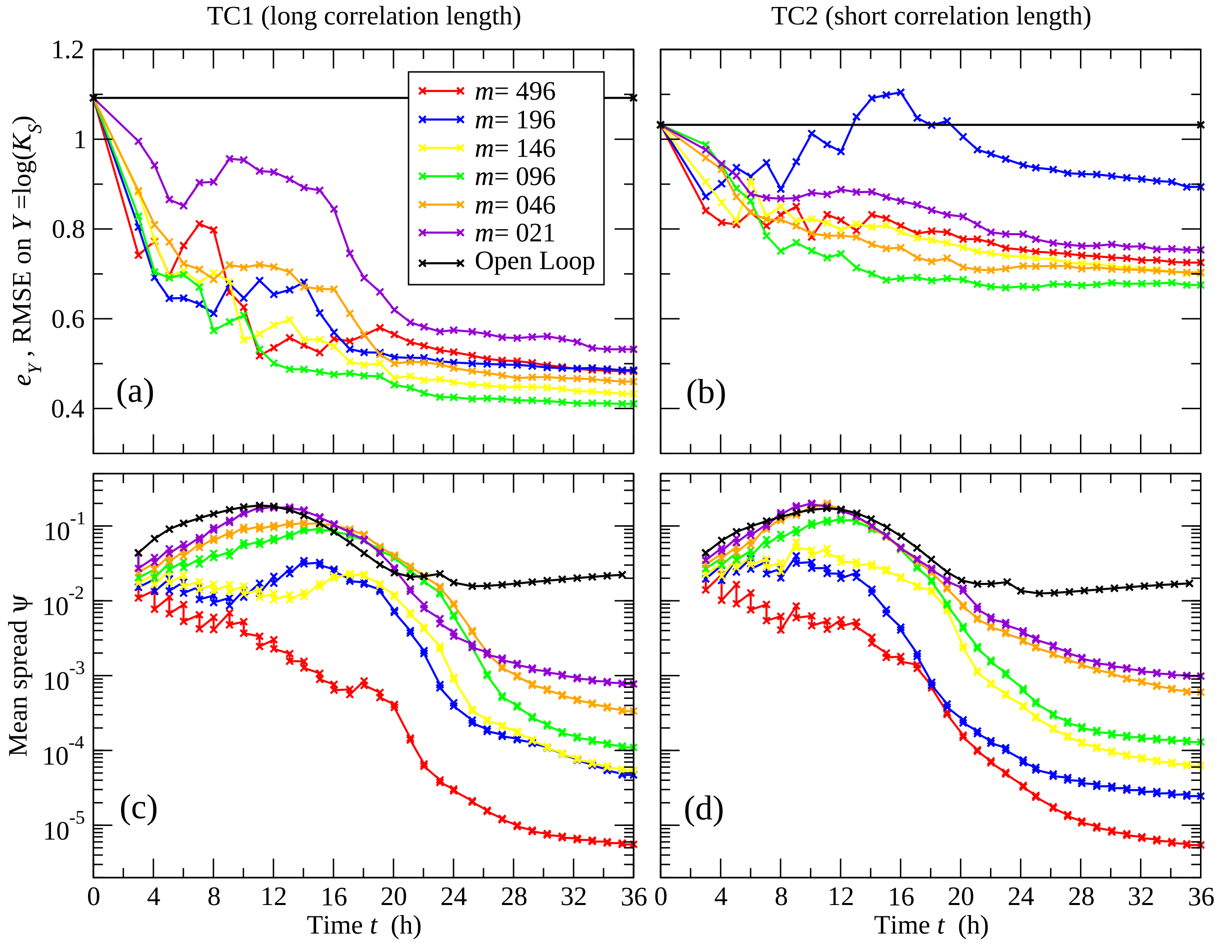

Figure 8: Root mean squared error (equation (23)) between the true and estimated distributions of K_S (panels (a) and (b)) and mean ensemble spread (equation (24)) of the pressure heads (panels (c) and (d)) at each assimilation of EnKF for different numbers of active sensors. The first update is at t=3 h.

where $\tilde{Y}(t)$ is the log-conductivity random field described by the MC realizations at time t.

The values of $e_Y(t)$ at the assimilation times are shown in panels (a) and (b) of Figure 8 for TC1 and TC2, respectively. We can see that even for a low number of active sensors the assimilation process computes a posterior distribution of $\tilde{Y}$ that has a lower error than the prior. The first assimilation step produces the most significant correction on the ensemble of $\tilde{Y}$ and, at this assimilation time, there is a clear correspondence between the simulations with the smallest error e_Y and the simulations with the largest number of active sensors m. However, the errors do not show a monotone decrease in time. This undesired result, that would be expected when using a low number of sensors (e.g., m=21), is particularly affecting the simulations with a large number of measurements (e.g., m=196, m=496), resulting in a final distribution of the errors that is not consistent with the number of active sensors. We associate this behaviour to two main drawbacks of EnKF. On the one hand, EnKF performs a global update of the parameters at each assimilation time. The ensemble approximation of the cross covariances between the state variables and the parameters may result in spurious correlations and a (wrong) update of parameters that are not directly involved in the measured process. On the other hand, EnKF tends to underestimate the variance of the updated state variables, thus reducing the efficiency of the Kalman gain in correcting the erroneous updates. This results in the reduction of the ensemble spread, especially for the scenarios with a large number of sensors. To show this phenomenon, we evaluate the mean variance associated with the nodal pressure head, which is the main state variable of the CATHY model and is directly related to the volumetric water content, as a measure of the ensemble spread:

$$\bar{\sigma}_{\psi}^2(t) = \frac{1}{\Omega}\int_{\Omega} E\left[(\psi(t) - E[\psi(t)])^2\right] d\Omega \tag{24}$$

where E indicates the mean over the ensemble. Panels (c) and (d) of Figure 8 show the temporal value of $\bar{\sigma}_{\psi}^2(t)$ for TC1 and TC2, respectively. We can see that, in all the scenarios explored, the variance decreases with respect to the OL run, indicative of a lower uncertainty on the model response. It is evident that the simulations with a high number of sensors have a faster decrease in the ensemble spread.

In our experiment, the infiltration process at the first assimilation time (t = 3 h) involves only the superficial layers of the domain. Thus the first assimilation steps are able to correct the Y field at the surface but have less accuracy at the bottom layers of the domain. As a consequence, the following update steps are necessary to improve the parameter estimation for these layers. Common techniques adopted to improve the EnKF update are covariance localization [21], to remove spurious correlations, and covariance inflation [19], to increase some components of the covariance and improve the ensemble spread. Here we adopt a different approach. The idea is to increase the forecast time between assimilations, in such a way as to have a better correlation between the state realizations and the associated parameters. Firstly, we focus on the time of the first assimilation, which entails the maximum decrease in the error and in the ensemble spread. It is important to perform the first update when the infiltrating water has also reached the sensors at the deepest layer, so that the process dynamics influences the response of the sensors throughout the entire domain. In our simulations, this occurs at t=12 h. As shown in panels (c) and (d) of Figure 8, this time

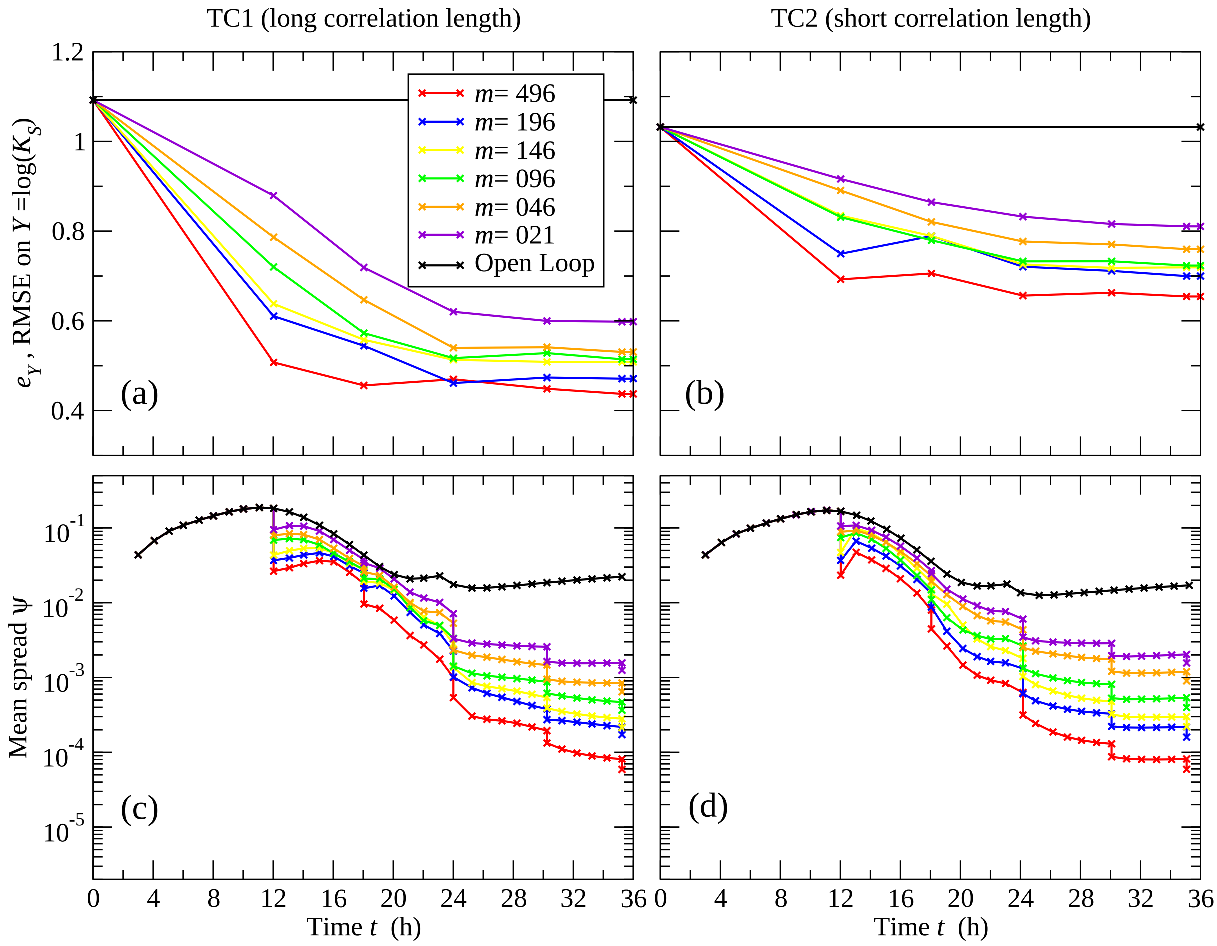

Figure 9: Root mean squared error (equation (23)) between the true and estimated distributions of K_S (panels (a) and (b)) and mean ensemble spread (equation (24)) of the pressure heads (panels (c) and (d)) at each assimilation of EnKF for different numbers of active sensors. The first update is at t=12 h and the assimilation frequency is six hours.

corresponds to the maximum variance on the OL pressure heads, and thus to the maximum differentiation between the ensemble realizations obtained with the prior parameter distribution. Secondly, as Shi et al. [37] have pointed out, the accuracy of an EnKF analysis may degrade when using short assimilation intervals, thus we increased the update time from hourly to every 6 h.

Figure 9 shows the convergence profiles of the errors e_Y obtained with the first assimilation timed at t=12 h and an assimilation frequency of six hours. It can be seen that the errors at the first update are lower than in the previous scenario (Figure 8) for all sensor failure configurations explored. Moreover, the error now increases only at a small number of assimilation times, very slightly, and without compromising the final correlation between the errors and the number of active sensors. The greatly reduced number of updates owing to the postponed first update and larger assimilation interval may explain the higher final errors obtained with fewer active sensors. However, the higher values of the ensemble spread should allow improved performance for longer simulations, thus restoring (or improving) the previously computed errors on the parameters.

Figure 9 shows that, for both TC1 and TC2, the error at the end of the simulation increases consistently as the number of active sensors drops. This suggests that the hydrological dynamics of LEO is well captured by the sensor network even when a large fraction of sensors has failed.

Figures 10 and 11 compare the true conductivity field and the a posteriori ensemble mean of the conductivity field estimated with m=496 and m=21 active sensors at six different depths d for TC1 and TC2, respectively. For both test cases, the estimated conductivity with m=496 captures most of the features of the true field, with small differences at the bottom layer where there are no sensors. The final K_S estimated with m=21, although having larger errors (see Figure 9), manages to capture several zones of high and low conductivity of the true field, especially for the scenario with long correlation length (TC1). This means that, also with only 21 active sensors, EnKF is able to improve the distribution representing the true soil hydraulic conductivity with respect to the a priori pdf.

This is even more evident in Figure 12, where the final ensemble means of the conductivities are employed in a forward run of CATHY to compare the hydrological response of the system for the different distributions of K_S. The true K_S and the K_S estimated with m=496, 46, and 21 active sensors produce very similar integrated responses of the system in terms of overland flow (panels (a) and (b)), seepage face flow (panels (c) and (d)), and total water storage (panels (e) and (f)) for both TC1 and TC2. The differences in the soil conductivity fields have a noticeable impact only when we examine the water content measured by the sensor network. Panels (g) and (h) report the root mean squared error on the measured volumetric water content, and we see that, for both TC1 and TC2, the results with m=496 active sensors are more accurate than those using m=21 or m=46 sensors.

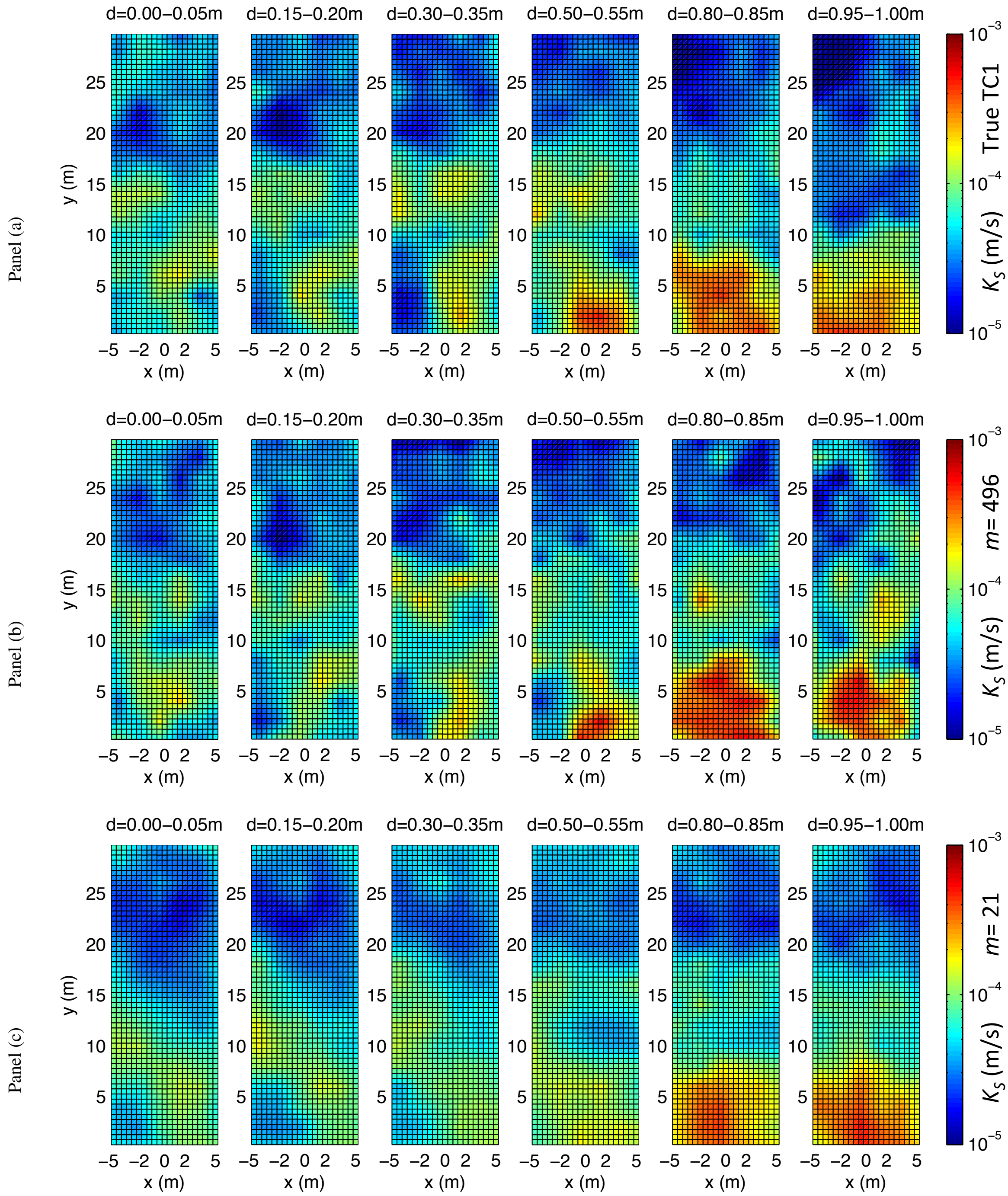

Figure 10: Test case TC1. Comparison between the true distribution of K_S (a) and the ensemble mean of K_S estimated with the EnKF procedure using m=496 (b) and m=21 (c) active sensors.

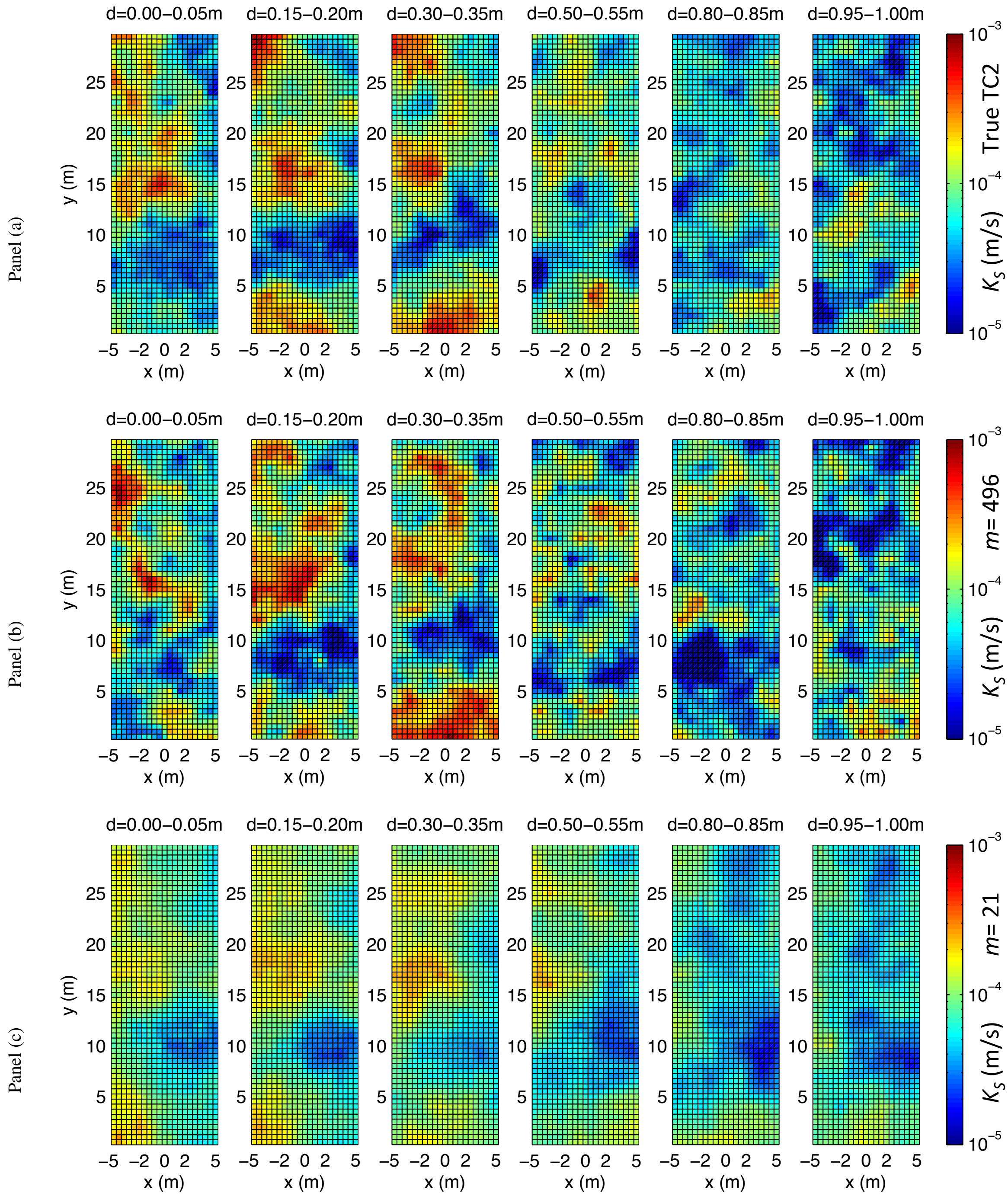

Figure 11: Test case TC2. Comparison between the true distribution of K_S (a) and the ensemble mean of K_S estimated with the EnKF procedure using m=496 (b) and m=21 (c) active sensors.

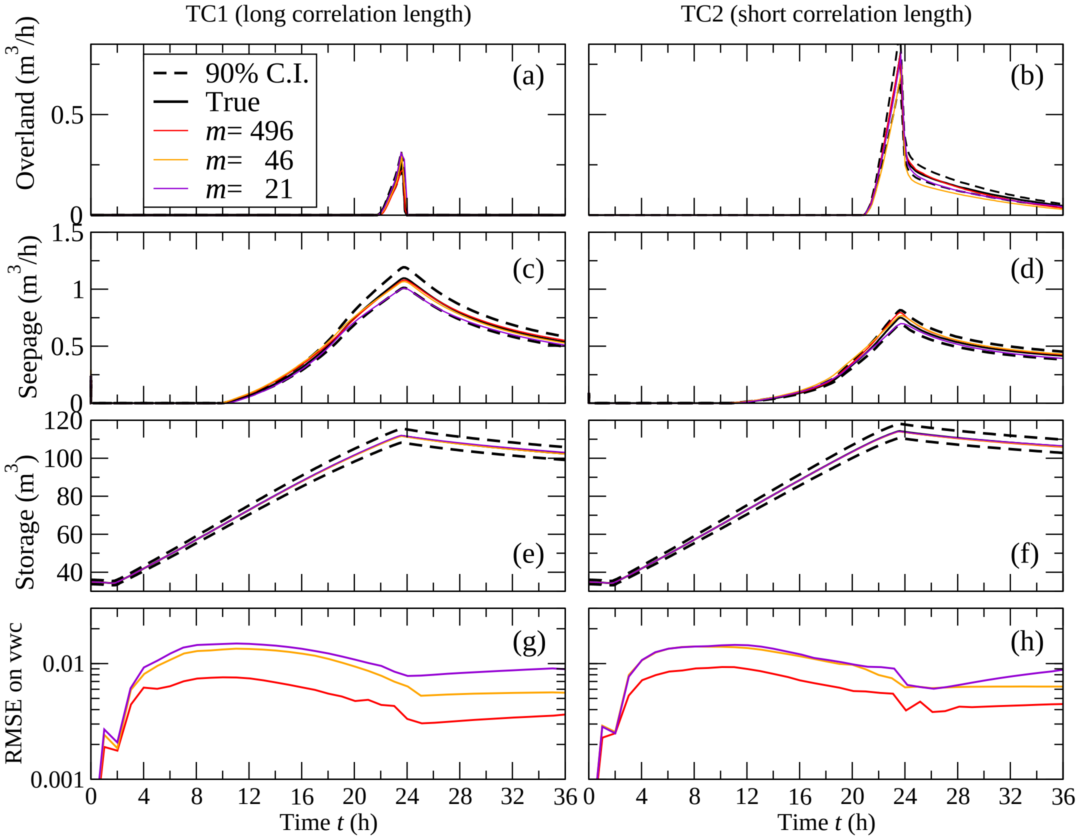

Figure 12: Comparison between the hydrological responses of CATHY using the true K_S field and the estimated distribution of K_S for m= 496, 46, and 21 active sensors, for both the TC1 (left graphs) and TC2 (right graphs) test cases. The hydrological responses shown are, from top graphs to bottom, overland flow, seepage face flow, total water storage, and root mean squared error on the volumetric water content. C.I. is confidence interval.

7. Conclusions

We have presented a failure analysis of the sensor network deployed in the experimental hillslopes of the Landscape Evolution Observatory at Biosphere 2. The main objective of this study was the determination of the minimum number of active sensors that are necessary for a reliable observation of the water dynamics in the system. To reach this goal we elected to work with two general quantities of interest in hydrological modeling. The first quantity is the total water volume inside the hillslope at a given time. The second quantity is the total discharge at the outlet, as formed by the sum of the seepage and overland fluxes. Synthetic test cases were developed for 11 different sensors sets having numbers of active sensors varying between 496 and 21. The soil was modeled as a three-dimensional second order stationary random field with exponential covariance function. A simulation with a given K_s distribution was used to obtain a "true" hillslope behavior, as simulated by the CATHY model, from which appropriate synthetic measurements were taken. For the first quantity, the assessment of the sensor failure was carried out by comparing the temporal variation of the total water storage as measured by virtual load cells and the total water volume obtained by linear interpolation of the observed values at the spatially distributed water content sensors. In the case of outlet discharge, the sensor network reliability was tested by using the measured water content inside the domain to identify via EnKF the spatially heterogeneous saturated hydraulic conductivity. The accuracy of the identification was verified by simulating the event with the hydrological model CATHY and comparing simulated and observed seepage and overland flow.

The results showed that the reconstruction of volumetric water content via linear interpolation of the sensor network values is an accurate procedure when the level of failure is low. When only few sensors are active (i.e., m <100) this methodology may introduce sizeable errors in the estimation of total water storage, especially during water infiltration periods. This corresponds to the fact that when the sensor network is only partially active there are large portions of the hillslope where information is lost and interpolation identifies incorrect water content values, especially near the soil surface.

Retrieval of the saturated hydraulic conductivity distribution, on the other hand, seems to be more robust. In fact, calibration of CATHY via EnKF is able to improve the estimation of the true conductivity fields yielding accurate predictions of the global hydrological response of the LEO hillslope also at the lowest tested numbers of active sensors (m=46 and m=21). This result is achieved thanks to the role played by the hydrological model in correlating in space and time the water content measurements with the hillslope response, and by the assumption of no model bias that is inherent in our synthetic experiment. To thoroughly verify this result, we conducted two test cases where the "true" saturated conductivity distribution was characterized by different correlation lengths. For both test cases we were able to assert a high reliability of the sensor network. The CATHY simulations were able to accurately simulate the behavior of the system also for the lowest number of active sensors. We remark that this was true for the overall global response of the system, i.e., by looking only at total seepage and overland flow at the outlet. The recovery of the distribution of the internal states was much more inaccurate especially when few sensors are active. In this case, the simulations suffer from the problem that the EnKF identification of K_s yields a field that is statistically equivalent (up to small errors) to the true realization. Hence, local details of the system dynamics, in terms of spatially distributed water content, are lost, leading to inaccurate simulated water content values at the sensor positions, with errors that decrease drastically when the entire sensor network is considered active. The outlet discharge, on the other hand, does not suffer from this problem, a sign that the statistical properties (mean and variance) of the K_s field were correctly identified.

The results from the two test cases showed that the identification process is more accurate for the larger correlation length, corresponding intuitively to a lower number of parameters to be identified. Moreover, to improve the parameter identification and limit the effects of erratic covariance estimations in the Kalman gain in areas of the domain where the measurements contribute negligible information, we showed that decreasing the assimilation frequency and delaying the first update helps incorporate responses to the infiltration signal from all the sensors. Consequently the sample spread increases, yielding reduced inconsistencies between filter behavior and number of active sensors. In this case, however, achieving small estimation errors requires larger simulation times, corresponding to a larger number of update steps. The test case results also showed differences in reconstruction according to the degree of sensor failure. Nonetheless we are able to reproduce the integrated hydrological response of LEO also with the lowest numbers of active sensors.

Acknowledgments

The Biosphere 2 facility and the capital required to conceive and construct LEO were provided through a charitable donation from the Philecology Foundation, and its founder, Mr. Edward Bass. We gratefully acknowledge this charitable donation.

References

[1] N-S. Hsu and W. W-G. Yeh. Optimum experimental design for parameter identification in groundwater hydrology. *Water Resour. Res.*, 25: 1025–1040, 1989. doi: 10.1029/WR025i005p01025.

[2] T. T. Ushijima and W. W-G. Yeh. Experimental design for estimating unknown groundwater pumping using genetic algorithm and reduced order model. *Water Resour. Res.*, 49(10):6688–6699, 2013. doi: 10.1002/wrcr.20513.

[3] L. Hopp, C. Harman, S. L. E. Desilets, C. B. Graham, J. J. McDonnell, and P. A. Troch. Hillslope hydrology under glass: confronting fundamental questions of soil-water-biota coevolution at Biosphere 2. *Hydrol. Earth Syst. Sci.*, 13:2105–2118, 2009. doi: 10.5194/hess-13-2105-2009.

[4] K. Dontsova, C. I. Steefel, S. Desilets, A. Thompson, and J. Chorover. Solid phase evolution in the Biosphere 2 hillslope experiment as predicted by modeling of hydrologic and geochemical fluxes. *Hydrol. Earth Syst. Sci.*, 13:2273–2286, 2009. doi: 10.5194/hess-13-2273-2009.

[5] V. Y. Ivanov, S. Fatichi, G. D. Jenerette, J. F. Espeleta, P. A. Troch, and T. E. Huxman. Hysteresis of soil moisture spatial heterogeneity and the “homogenizing” effect of vegetation. *Water Resour. Res.*, 46:W09521, 2010. doi: 10.1029/2009WR008611.

[6] M. Camporese, C. Paniconi, M. Putti, and S. Orlandini. Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resour. Res.*, 46:W02512, 2010. doi: 10.1029/2008WR007536.

[7] C. Paniconi and M. Putti. A comparison of Picard and Newton iteration in the numerical-solution of multidimensional variably saturated flow problems. *Water Resour. Res.*, 30(12):3357–3374, 1994. doi: 10.1029/94WR02046.

[8] S. Orlandini and R. Rosso. Diffusion wave modeling of distributed catchment dynamics. *J. Hydraul. Eng.-ASCE*, 1(3):103–113, 1996. doi: 10.1061/(ASCE)1084-0699(1996)1:3(103).

[9] S. Orlandini, G. Moretti, M. Franchini, B. Aldighieri, and B. Testa. Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models. *Water Resour. Res.*, 39(6):1144, 2003. doi: 10.1029/2002WR001639.

[10] G.-Y. Niu, D. Pasetto, C. Scudeler, C. Paniconi, M. Putti, P. A. Troch, S. B. DeLong, K. Dontsova, L. Pangle, D. D. Breshears, J. Chorover, T. E. Huxman, J. Pelletier, S. R. Saleska, and X. Zeng. Incipient subsurface heterogeneity and its effect on overland flow generation - Insight from a modeling study of the first experiment at the Biosphere 2 Landscape Evolution Observatory. *Hydrol. Earth Syst. Sci.*, 18:1873–1883, 2014. doi: 10.5194/hess-18-1873-2014.

[11] G. Evensen. *Data Assimilation: The Ensemble Kalman Filter*. Springer, New York, 2007.

[12] G. Evensen. The ensemble Kalman filter for combined state and parameter estimation. *IEEE CONTR. Syst. Mag.*, 29(3):83–104, 2009. doi: 10.1109/MCS.2009.932223.

[13] M. Camporese, G. Cassiani, R. Deiana, and P. Salandin. Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. *Water Resour. Res.*, 47(12):W12508, 2011. doi: 10.1029/2011WR010528.

[14] E. Crestani, M. Camporese, D. Baú, and P. Salandin. Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation. *Hydrol. Earth Syst. Sci.*, 17(4):1517–1531, 2013. doi: 10.5194/hess-17-1517-2013.

[15] H.-J. Hendricks Franssen and W. Kinzelbach. Ensemble Kalman filtering versus sequential self-calibration for inverse modeling of dynamic groundwater flow systems. *J. Hydrol.*, 365:261–274, 2009. doi: 10.1016/j:jhydrol.2008.11.033.

[16] J. Carrera and S. P. Neumann. Estimation of aquifer paremeters under transient and steady state conditions: 2. Uniqueness, stability and solution algorithms. *Water Resour. Res.*, 22(2):211–227, 1986.

[17] J. P. Drecourt, H. Madsen, and D. Rosbjerg. Calibration framework for a Kalman filter applied to a groundwater model. *Adv. Water Resources*, 29(5):719–734, 2006. doi: 10.1016/j.advwatres.2005.07.007.

[18] G. J. M. De Lannoy, R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resour. Res.*, 43(9):W09410, 2007.

[19] H.-J. Hendricks Franssen and W. Kinzelbach. Real-time groundwater flow modeling with the ensemble Kalman filter: Joint estimation of states and parameters and the filter inbreeding problem. *Water Resour. Res.*, 44:W09408, 2008. doi: 10.1029/2007WR006505.

[20] Y. Chen and D. S. Oliver. Cross-covariances and localization for EnKF in muliphase flow data assimilation. *Comput. Geosci.*, 14:579–601, 2010. doi: 10.1007/s10596-009-9174-6.

[21] J. L. Anderson. Localization and sampling error correction in ensemble Kalman filter data assimilation. *Mon. Weather Rev.*, 140(7):2359–2371, 2012. doi: 10.1175/MWR-D-11-00013.1.

[22] A. Y. Sun, A. Morris, and S. Mohanty. Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data. *Adv. Water Resources*, 32(2):280 – 292, 2009. doi: 10.1016/j.advwatres.2008.11.006.

[23] A. Y. Sun, A. Morris, and S. Mohanty. Sequential updating of multimodal hydrogeologic parameter fields using localization and clustering techniques. *Water Resour. Res.*, 45:W07424, 2009. doi: 10.1029/2008WR007443.

[24] M. Panzeri, M. Riva, A. Guadagnini, and S. P. Neuman. Data assimilation and parameter estimation via ensemble Kalman filter coupled with stochastic moment equations of transient groundwater flow. *Water Resour. Res.*, 49(3):1334–1344, 2013. doi: 10.1002/wrcr.20113.

[25] A. H. Alzraiee, D. Baú, and Elhaddad A. Estimation of heterogeneous aquifer parameters using centralized and decentralized fusion of hydraulic tomography data. *Hydrol. Earth Syst. Sci.*, 18:3207–3223, 2014. doi: 10.5194/hess-18-3207-2014.

[26] Y. Chen and D. Zhang. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Adv. Water Resources*, 29(8): 1107–1122, 2006. doi: 10.1016/j.advwatres.2005.09.007.

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

[28] M. T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Amer. J.*, 44: 892–898, 1980.

[29] D. Pasetto, M. Camporese, and M. Putti. Ensemble Kalman filter versus particle filter for a physically-based coupled surface-subsurface model. *Adv. Water Resources*, 47(1):1–13, 2012. doi: 10.1016/j.advwatres.2012.06.009.

[30] H. Moradkhani, K.-L. Hsu, H. Gupta, and S. Sorooshian. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.*, 41(5):W05012, 2005. doi: 10.1029/2004WR003604.

[31] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.

[32] R. E. Kalman. A new approach to linear filtering and prediction problems. *J. Basic Eng.-T. ASME*, 82(Series D):35–45, 1960.

[33] G. Evensen. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dyn.*, 53(4):343–367, 2003. doi: 10.1007/s10236-003-0036-9.

[34] A. I. Gevaert, A. J. Teuling, R. Uijlenhoet, S. B. DeLong, T. E. Huxman, L. Pangle, D. D. Breshears, J. Chorover, J. D. Pelletier, S. R. Saleska, X. Zeng, and P. A. Troch. Hillslope-scale experiment demonstrates role of convergence during two-step saturation. *Hydrol. Earth Syst. Sci.*, 18:3681–3692, 2014. doi: 10.5194/hess-18-3681-2014.

[35] A. Bellin and Y. Rubin. HYDRO_GEN: A spatially distributed random field generator for correlated properties. *Stoch. Hydrol. Hydraul.*, 10 (4):253–278, 1996. doi: 10.1007/BF01581869.

[36] M. A. Cushy, A. Bellin, and Y. Rubin. Generation of three-dimensional flow fields for statistically anisotropic heterogeneous porous-media. *Stoch. Hydrol. Hydraul.*, 9(1):89–104, 1995. doi: 10.1007/BF01581760.

[37] Y. Shi, K. J. Davis, F. Zhang, C. J. Duffy, and X. Yu. Parameter estimation of a physically based land surface hydrologic model using the ensemble Kalman filter : A synthetic experiment. *Water Resour. Res.*, 50:706–724, 2014. doi: 10.1002/2013WR014070.