

Manuscript Number: JCOMP-D-14-00215R1

Title: An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment

Article Type: Regular Article

Keywords: Particle filter; Data Assimilation; Coupled Hydro-Geophysical Inversion; Electrical Resistivity Tomography

Corresponding Author: Mr. Gabriele Manoli,

Corresponding Author's Institution: University of Padova

First Author: Gabriele Manoli

Order of Authors: Gabriele Manoli; Matteo Rossi; Damiano Pasetto; Rita Deiana; Stefano Ferraris; Giorgio Cassiani; Mario Putti

Abstract: The modeling of unsaturated groundwater flow is affected by a high degree of uncertainty related to both measurement and model errors. Geophysical methods such as Electrical Resistivity Tomography (ERT) can provide useful indirect information on the hydrological processes occurring in the vadose zone. In this paper, we propose and test an iterataed particle filter method to solve the coupled hydrogeophysical inverse problem. We focus on an infiltration test monitored by time-lapse ERT and modeled using Richards equation. The goal is to identify hydrological model parameters from ERT electrical potential measurements. Traditional uncoupled inversion relies on the solution of two sequential inverse problems, the first one applied to the ERT measurements, the second one to Richards equation. This approach does not ensure an accurate quantitative description of the physical state, typically violating mass balance. To avoid one of these two inversions and incorporate in the process more physical simulation constraints, we cast the problem within the framework of a SIR (Sequential Importance Resampling) data assimilation approach that uses a Richards equation solver to model the hydrological dynamics and a forward ERT simulator combined with Archie'slaw to serve as measurement model. ERT observations are then used to update the state of the system as well as to estimate the model parameters and their posterior distribution. The limitations of the traditional sequentialBayesian approach are investigated and an innovative iterative approach is proposed to estimate the model parameters with high accuracy. The numerical properties of the developed algorithm are verified on both homogeneous and heterogeneous synthetic test cases based on a real-world field experiment.



NICHOLAS SCHOOL OF THE ENVIRONMENT
DUKE UNIVERSITY

DIVISION OF EARTH & OCEAN SCIENCES

November 18, 2014

Dear Editor,

thank you for reconsidering the manuscript "*An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment*" by G. Manoli, M. Rossi, D. Pasetto, R. Deiana, S. Ferraris, G. Cassiani, and M. Putti (JCOMP-D-13-00723) for publication after minor revisions.

We have addressed all the reviewer's minor comments and, in the attached documents, we report the revised version of the paper together with a summary of the revisions made.

Thank you for your time and consideration,

A handwritten signature in blue ink that reads "Gabriele Manoli".

Gabriele Manoli



Revision notes

"An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment"

(JCOMP-D-14-00215)

Reviewer #2:

The manuscript presents an iterative Bayesian approach to data assimilation in the context of hydro-geophysical inversion. I did not review the first version of this manuscript; based on the first round of reviews and the authors' reply I feel the revised version (the current submission) is a significant improvement that warrants the manuscript's eventual publication in JCP. The following are a few (relatively minor) comments that should be addressed before the manuscript becomes publishable.

1. The introduction misses a discussion of the state-of-the-art in "hydro-geophysical inversion of...infiltration experiment[s]". In the absence of such a discussion, it is hard to judge the novelty of the current contribution to that particular application area. A cursory Google search with these keywords reveals a number of papers, e.g., Tartakovsky and others, Hydro-geophysical approach for identification of layered structures of the vadose zone from electrical resistivity data, Vadose Zone Journal, 2008.

REPLY: We agree with the reviewer's comment and we have added a discussion on state-of-the-art applications of hydro-geophysical inversion for the characterization of vadose zone processes (p. 3 lines 17-20):

"ERT has been widely used to monitor vadose zone processes (e.g., Daily et al. 1992, LaBrecque et al. 2004, Tartakovsky et al. 2008) but it is well known that the inversion procedure can produce mass balance errors (Singha and Gorelick, 2005) especially when surface ERT is used to monitor water infiltration into soil (Michot et al. 2003, Cassiani et al. 2012, Travelletti et al. 2012) due to a rapid decrease of ERT resolution with depth."

2. It would be helpful to provide a motivation for the computational example presented in Figure 6(a). What physical setting does this example represent? Would it not be more natural to consider a layered soil?

REPLY: The reviewer is right, the choice of the physical setting presented in Fig. 6a has been clarified. Just to add to the discussion with the reviewer, we would like to point out that a layered soil with vertical-only infiltration would not test completely the ability of the method to identify spatially heterogeneous patterns. In some sense, the proposed test case actually contains the layered case, but adds also a horizontal component to the front movement. We have added the following sentence (p. 27 lines 424-427):

"The physical setting in Fig. 6(a) is not intended to represent typical field conditions but aims to provide a simple setup generating both vertical and lateral infiltration patterns to test the proposed approach in a truly multidimensional heterogeneous setting."

An iterative particle filter approach for coupled hydro-geophysical inversion of a controlled infiltration experiment

Gabriele Manoli^{a,b}, Matteo Rossi^c, Damiano Pasetto^a, Rita Deiana^d,
Stefano Ferraris^e, Giorgio Cassiani^c, Mario Putti^a

^a*Department of Mathematics, University of Padova, Via Trieste 63, 35121 Padova, Italy*

^b*Nicholas School of the Environment, Duke University, Durham, North Carolina 27708, USA.*

^c*Department of Geosciences, University of Padova, Via Gradenigo 6, 35131 Padova, Italy*

^d*Dipartimento dei Beni Culturali, University of Padova, Piazza Capitaniato 7, 35139 Padova, Italy*

^e*Interuniversity Department of Regional and Urban Studies and Planning, Politecnico and University of Torino, Viale Mattioli 39, 10125 Torino, Italy*

Abstract

The modeling of unsaturated groundwater flow is affected by a high degree of uncertainty related to both measurement and model errors. Geophysical methods such as Electrical Resistivity Tomography (ERT) can provide useful indirect information on the hydrological processes occurring in the vadose zone. In this paper, we propose and test an iterataed particle filter method to solve the coupled hydrogeophysical inverse problem. We focus on an infiltration test monitored by time-lapse ERT and modeled using Richards equation. The goal is to identify hydrological model parameters from ERT electrical potential measurements. Traditional uncoupled inversion relies on the solution of two sequential inverse problems, the first one applied to the

Email address: manoli@dmsa.unipd.it (Gabriele Manoli)

ERT measurements, the second one to Richards equation. This approach does not ensure an accurate quantitative description of the physical state, typically violating mass balance. To avoid one of these two inversions and incorporate in the process more physical simulation constraints, we cast the problem within the framework of a SIR (Sequential Importance Resampling) data assimilation approach that uses a Richards equation solver to model the hydrological dynamics and a forward ERT simulator combined with Archie's law to serve as measurement model. ERT observations are then used to update the state of the system as well as to estimate the model parameters and their posterior distribution. The limitations of the traditional sequential Bayesian approach are investigated and an innovative iterative approach is proposed to estimate the model parameters with high accuracy. The numerical properties of the developed algorithm are verified on both homogeneous and heterogeneous synthetic test cases based on a real-world field experiment.

Keywords: Particle filter, Data Assimilation, Coupled Hydro-Geophysical Inversion, Electrical Resistivity Tomography

1. Introduction

Electrical Resistivity Tomography (ERT) is a practical, cost-effective, indirect tool for collecting soil and moisture content data in subsurface environments [1–5]. When applied to the simulation of the dynamics of the vadose zone, ERT relies on the inversion of the direct current (DC) flow equation providing an image of the electrical resistivity [4], with the soil moisture pattern reconstructed from petrophysical relations, such as, e.g., Archie's Law [6]. A second inverse problem is finally used to estimate hydrological model param-

9 eters. It is well known that inverse modeling of a parabolic diffusion equation
10 is generally an ill-posed problem and regularization techniques are often em-
11 ployed to achieve well-posedness [2, 7–9]. Traditional geophysical inversion
12 is at the same time an over- and under- constrained problem, in the sense
13 that the problem character can change in space, and benefits from the use of
14 prior information embedded in the regularization procedure [10]. However,
15 imposing smoothness via regularization may introduce inaccuracies or even
16 unphysical constraints into the estimates of the hydrological properties [11].
17 ERT has been widely used to monitor vadose zone processes [e.g. 1, 12, 13]
18 but it is well known that the inversion procedure can produce mass balance
19 errors [14] especially when surface ERT is used to monitor water infiltration
20 into soil [15, 5, 16] due a rapid decrease of ERT resolution with depth. To
21 cope with this limitation coupled hydro-geophysical approaches seem highly
22 promising [17]. By these procedures, the spatial distribution and the tem-
23 poral dynamics of the geophysical properties are enforced by a physically
24 based hydrologic model combined with petrophysical relations, and explicit
25 assumptions for spatial and temporal regularization are no longer needed.

26 Even though the coupled approach avoids an independent geophysical
27 inversion, estimation of the hydrologic properties (e.g. soil hydraulic pa-
28 rameters) is still a highly non-linear, mixed-determined inversion problem.
29 For these reasons, although parameter estimation can be made theoretically
30 well-posed, the physical interpretation of the estimated parameters is still
31 not well understood [18]. The presence of structural model errors (model
32 approximations, uncertain initial conditions, etc.), as well as measurement
33 uncertainties, suggests that a deterministic search for the best parameters is

34 not likely to converge to a single set of “true” values. A stochastic approach
35 based on ensemble forecasting seems therefore the most appropriate solution
36 procedure [18, 19].

37 Sequential Data Assimilation (S-DA) methods (typically called filters)
38 have been successfully applied to improve model predictions by incorporat-
39 ing real system observations onto the dynamical model and have been already
40 employed to correct the hydrological states of groundwater infiltration mod-
41 els [20]. Their ability to include structural and parametric error distributions
42 make them particularly attractive for application to the problem of dynamic
43 parameter estimation [18]. Because of the high nonlinearity of porous media
44 infiltration models, the typical filtering method used in hydrological applica-
45 tions is the Ensemble Kalman filter (EnKF) [21]. Notwithstanding the linear
46 optimality properties of the Kalman Gain [22], the main limitation of EnKF
47 is that it is based on the Gaussian approximation of the filtering probability
48 distribution, possibly leading to inaccurate results or even divergence of the
49 posterior pdfs in presence of a strongly nonlinear relation between observa-
50 tions and state variables [23–25]. To cope with arbitrary non-Gaussian prior
51 distributions, the family of particle filters is a highly attractive alternative, as
52 it is directly based on the Bayesian filtering rule [26, 27]. Particle filters have
53 been recently introduced into hydrology [28–31, 25] and used also for estima-
54 tion of hydrological model parameters [32–34]. All these latter studies focus
55 on the assimilation of direct hydrological information (e.g. discharge [25] or
56 soil moisture data [35–37]). A coupled hydro-geophysical parameter estima-
57 tion procedure by S-DA has been presented by [38], but its ability to provide
58 accurate estimates of unknown model parameters remains to be proven, as

59 shown by the consistent underestimation of saturated hydraulic conductivity
60 in the results of [38]. As a matter of fact, the structural uncertainties of both
61 the hydrologic evolution and geophysical observation models strongly affect
62 the estimated parameters. Sequential filters correct both model parameters
63 and state variables at each assimilation time, yielding identified parameter
64 values that vary in time [18]. Compared to smoothers or other more sophis-
65 ticated inversion methods (e.g., Markov Chain Monte Carlo methods [39, 40])
66 the filtering approach is computationally more efficient when dealing with a
67 detailed and spatially resolved simulation model such as the coupled Richards
68 equation-ERT solver here employed.

69 In this paper we propose an iterative procedure to overcome the prob-
70 lem of the sensitivity to the initial guess and provide accurate identifica-
71 tion of unknown model parameters from indirect state information. The
72 method is grounded on a Sequential Importance Resampling (SIR) particle
73 filter, already tested in similar hydrological applications [25, 38], whereby an
74 ERT forward simulation model is embedded into the observation equation
75 and both parameter and state distributions are updated at each assimilation
76 step. Iteration is introduced by sequentially repeating until convergence the
77 same simulation period, using as initial guess the state values and parameter
78 pdfs evaluated from the results of the previous iteration. Compared to more
79 sophisticated statistical updates, the use of iterations allows the inclusion of
80 a less accurate but computationally more efficient inversion scheme able to
81 cope with large dimensional problems.

82 We validate the methodology on synthetic test cases and apply the meth-
83 ods to a field experiment comparing the results of our procedure with tra-

ditional uncoupled inversion of ERT data. We focus on both homogeneous and heterogeneous systems with parameters distributed by zones. The proposed procedure displays convergence of the posterior distribution towards the correct value of the hydraulic conductivity in both the homogenous and heterogeneous scenarios independently from the initial guess. The numerical results obtained from the synthetic test cases show that the iterative approach yields faster convergence with respect to standard DA methods, using consistently smaller ensemble sizes and a drastic reduction of the number of forward model runs, in particular for the heterogeneous test case. The results obtained in the application to the real world problem are consistent with the desired physical constraints at relatively low computational costs, thus improving significantly on existing coupled flow-ERT procedures.

2. Parameter estimation by sequential data assimilation

The state space model describing the S-DA problem can be written as:

$$x_t = \mathcal{F}(x_{t-1}, \lambda, w_t), \quad (1)$$

$$y_t = \mathcal{H}(x_t, \lambda, v_t), \quad (2)$$

where x_t is the state vector at assimilation time t , \mathcal{F} is the evolution operator, λ is the time-independent parameter vector, w_t is the stochastic model error, y_t is the observation vector, \mathcal{H} is the observation model, and v_t is the stochastic error term in the observations. Model uncertainty is connected, e.g., to structural model errors, parameter errors, initial solution errors, etc. Casted in a stochastic framework, the objective of S-DA is to estimate the posterior probability density function (pdf) of the state vector at time t conditioned to

105 the observations y_t^{obs} that become available at time t . Because of model non-
106 linearity, Monte Carlo-based approaches are used to discretize the state and
107 observation pdfs in equations (1) and (2). To relax the Gaussian hypothesis
108 inherent to Kalman-filter based algorithms we estimate the state and pa-
109 rameter pdfs employing a SIR (Sequential Importance Resampling) particle
110 filter, which has been successfully tested in hydrological applications [25] in
111 standard S-DA mode.

112 *2.1. Sequential Importance Resampling for parameter estimation*

113 Let the state vector x_t be characterized by a probability density function
114 denoted by $p(x_t)$ and let $p(\lambda)$ be the prior distribution of the parameters
115 λ . The sequence of random variables $\{x_0, x_1, \dots\}$ defines a Markov chain
116 where (1) and $p(w_t)$ uniquely identify the transition probability density func-
117 tion $p(x_t|x_{t-1}, \lambda)$. The variance associated to $p(x_t)$ typically increases with
118 time during the numerical simulation, leading to highly uncertain forecasts.
119 Our goal is to obtain the posterior distribution of the parameters λ and of the
120 state variables x_t , conditioned to the field observations $y_{1:t}^{obs}$, i.e., the filtering
121 pdf $p(x_t, \lambda|y_{1:t}^{obs})$. Sequential data assimilation allows to compute a posterior
122 distribution as soon as a field observation y_t^{obs} becomes available. For this
123 reason in the following we will assume that the parameters are time depen-
124 dent, λ_t , in the sense that they may change when their posterior distribution
125 changes.

126 The S-DA technique consists of two basic steps that are repeated sequen-
127 tially. In the forecast step the state pdf is propagated in time to obtain the
128 forecast pdf, $p(x_t, \lambda_t|y_{1:t-1}^{obs})$. This is expressed by the Chapman-Kolmogorov

129 equation as:

$$p(x_t, \lambda_t | y_{1:t-1}^{obs}) = \int p(x_t, \lambda_t | x_{t-1}, \lambda_{t-1}) p(x_{t-1}, \lambda_{t-1} | y_{1:t-1}^{obs}) dx_{t-1} d\lambda_{t-1}. \quad (3)$$

130 Note that in this step we have the effective propagation from time $t - 1$ to
 131 time t of the system state by formal application of (1) using constant values
 132 of the parameters. The second step is called analysis or update and consists
 133 in correcting the forecast pdf using the new field observation y_t^{obs} . Bayes'
 134 theorem allows the factorization of the filtering pdf as:

$$p(x_t, \lambda_t | y_{1:t}^{obs}) = C p(y_t^{obs} | x_t, \lambda_t) p(x_t, \lambda_t | y_{1:t-1}^{obs}),$$

135 where C is a normalization constant and the other two factors are the like-
 136 likelihood function, to which we assign a known distribution, and the forecast
 137 pdf, computed in (3), respectively. The analysis step essentially consists in a
 138 reinitialization of the system state variables and of the parameters given the
 139 forecast and the observations.

140 In the SIR algorithm the forecast and filtering pdfs are approximated
 141 using an ensemble of N random samples (also called particles), $\{x_t^{(i)}, \lambda_t^{(i)}\}$,
 142 $i = 1, \dots, N$, with associated weights $\{\omega_t^{(i)}\}$, $i = 1 \dots, N$:

$$p(x_t, \lambda_t | y_{1:t-1}^{obs}) \approx \sum_{i=1}^N \omega_t^{(i-)} \delta(x_t - x_t^{(i-)}) \delta(\lambda_t - \lambda_t^{(i-)}), \quad (4)$$

$$p(x_t, \lambda_t | y_{1:t}^{obs}) \approx \sum_{i=1}^N \omega_t^{(i+)} \delta(x_t - x_t^{(i+)}) \delta(\lambda_t - \lambda_t^{(i+)}), \quad (5)$$

143 where $\delta(\cdot)$ is the Dirac delta function, and superscripts '-' and '+' denote
 144 the realizations before and after the update, respectively. The SIR algorithm
 145 starts by assigning uniform weights to the N realizations of the ensemble.

146 The Monte Carlo discretization reduces the forecast step to the propagation
 147 in time of the ensemble members using the system dynamics and, in the up-
 148 date step, new weights are calculated recursively, by means of the likelihood
 149 function, as:

$$\omega_t^{(i)} = C\omega_{t-1}^{(i)}p(y_{1:t}^{obs}|x_t^{(i-)}, \lambda_t), \quad (6)$$

150 where C is a normalization constant. To avoid the ensemble deterioration
 151 phenomenon [41], resampling is performed when $N_{eff} < 0.5N$, where N_{eff}
 152 is the effective ensemble size, evaluated as:

$$N_{eff} = \left[\sum_{i=1}^N (\omega_t^{(i)})^2 \right]^{-1},$$

153 and is representative of the number of realizations that have non-negligible
 154 weights. We adopt the systematic resampling method [42], to duplicate sam-
 155 ples with large weight and discard samples with negligible weight. The re-
 156 sampling procedure maintains the ensemble size equal to N by generating
 157 new members using parameters drawn from the posterior distribution and
 158 assigning to them uniform weights. The duplicated realizations will then dif-
 159 ferentiate in the following forecast step. If the resampling step does not occur,
 160 i.e., all the particles have sizable weights, then $x_t^{(i+)} = x_t^{(i-)}, \lambda_t^{(i+)} = \lambda_t^{(i-)}$ and
 161 only the weights are changed according to (6), yielding an effective weighted
 162 distribution given by (4) and (5).

163 2.2. Iterative parameter estimation

164 Since the resampling step is a reinitialization of the system state variables
 165 at an observation time, it is convenient to use this step to sample new realiza-
 166 tions from the posterior pdf of the parameters. Let $\{\hat{\lambda}_t^{(i)}\}, i = 1, \dots, N$ be the

167 parameter values of the realizations after the resample. Most of these param-
168 eters are equal, the number of different values corresponding to the number
169 of realizations that have non-negligible weights. Maintaining these values
170 for the parameter update, i.e. $\lambda_t^{(i+)} = \hat{\lambda}_t^{(i)}$, may yield an impoverishment
171 of the ensemble with the consequence that the posterior distribution is not
172 adequately explored and erroneous parameter estimations may be identified.
173 This can be exemplified in the case that only one realization is duplicated
174 after the resample. In this case the posterior distribution collapses in one
175 single value that cannot change in the subsequent updates. To guarantee a
176 good performance of the filter it is then necessary to perturb the duplicated
177 parameters to effectively explore the relevant pdf. Moradkhani et al. [28]
178 propose a perturbation of the parameters with independent additive Gaus-
179 sian variates, $\lambda_t^{(i+)} = \hat{\lambda}_t^{(i)} + \xi_t^{(i)}$, $\xi_t^{(i)} \sim N(0, Var(\lambda_t^{(i-)}))$, while [43, 44] use
180 a Markov-Chain sampling of the parameters with the computation of the
181 Metropolis ratio to accept or eventually reject the sampled values. While the
182 first approach requires a large number of realizations, the second strategy
183 incurs in increased computational effort due to the repetition of the fore-
184 cast step necessary for the computation of the Metropolis ratio. Here we
185 propose to sample the updated parameters from a probability distribution
186 that maintains the initial structure, but employing the moments updated
187 with the ensemble statistics. For example, assuming an initial distribution
188 defined only by the first and second moments (e.g., uniform, normal, log-
189 normal distributions), the proposed scheme updates the expected value μ_{λ_t}
190 and the coefficient of variation cv_{λ_t} on the basis of the prior $\{\lambda_t^{(i-)}\}$ and the
191 resampled $\{\hat{\lambda}_t^{(i)}\}$ parameters. To this aim, we impose that the expected value

192 of the new distribution be given by the mean of the resampled parameters:

$$\mu_{\lambda_t} = E[\hat{\lambda}_t^{(i)}], \quad (7)$$

193 and the coefficient of variation be given by the maximum between the coef-
 194 ficient of variations of the forecasted and the updated parameters,

$$cv_{\lambda_t} = s \cdot \max \left(cv_{\lambda_t^{(-)}}, cv_{\hat{\lambda}_t} \right), \quad (8)$$

195 where s is a tuning coefficient used to force a gradual reduction of the variance
 196 of the distribution (typically $s=0.9$) and the use of the maximum value avoids
 197 the fast collapse of the filter when only a few realizations are resampled. The
 198 sequence of posterior parameter distributions obtained with this procedure
 199 needs several updates to converge and hence we iterate the filtering procedure
 200 by cyclic repetition of the assimilation interval until the resampling step is
 201 no longer performed at any update of the period. This stopping criterion
 202 ensures that no further progresses are obtained by continuing the iterations.
 203 A more computationally savvy approach would be to stop on the basis of
 204 average residual or parameter update metrics. At each restart of the filtering
 205 process (external iteration) the mean and variance of the prior distribution
 206 of the parameters is updated by:

$$\begin{aligned} \mu_{\lambda_0}^{k+1} &= \frac{1}{n_t} \sum_{t=1}^{n_t} \mu_{\lambda_t}^k, \\ cv_{\lambda_0}^{k+1} &= \frac{1}{n_t} \sum_{t=1}^{n_t} cv_{\lambda_t}^k, \end{aligned}$$

207 where n_t is the number of updates in each S-DA cycle (k -th external iter-
 208 ation). Instead of restarting the S-DA procedure with the posterior distri-
 209 bution at the previous S-DA cycle, we use a “mean posterior disitribution”

210 to reduce the effect of the initial bias on the parameter estimation. The
 211 procedure is illustrated schematically in Figure 1.

212 **3. Evolution and Observation models of water infiltration and ERT**

213 In this study we are interested in applying the S-DA method to a coupled
 214 hydro-geophysical model. The evolution model (1) describes the soil mois-
 215 ture dynamics in the vadose zone and ERT observations are used to update
 216 system state and parameters by means of a geophysical electrical current flow
 217 observation model (2).

218 *3.1. Evolution model*

219 We use Richards' equation to describe the infiltration process in a variably-
 220 saturated isotropic porous medium:

$$S_s S_w(\psi) \frac{\partial \psi}{\partial t} + \phi \frac{\partial S_w(\psi)}{\partial t} = \vec{\nabla} \cdot \left[\mathbf{K}_s K_r(\psi) \left(\vec{\nabla} \psi + \eta_z \right) \right] + q, \quad (9)$$

221 where S_s is the elastic storage term, S_w is water saturation, ψ is water pres-
 222 sure, t is time, ϕ is the porosity, \mathbf{K}_s is the saturated hydraulic conductivity
 223 tensor, K_r is the relative hydraulic conductivity, $\eta_z = (0, 0, 1)^T$ with z the ver-
 224 tical coordinate directed upward and q is a source/sink term. The saturated
 225 hydraulic conductivity is modeled as a diagonal matrix and its components
 226 K_x , K_y and K_z are the saturated hydraulic conductivities along the coordi-
 227 nate directions x , y and z , respectively. Equation (9) is highly nonlinear due
 228 to the pressure head dependencies of saturation and relative hydraulic con-
 229 ductivity. These constitutive functions are modeled using the characteristic

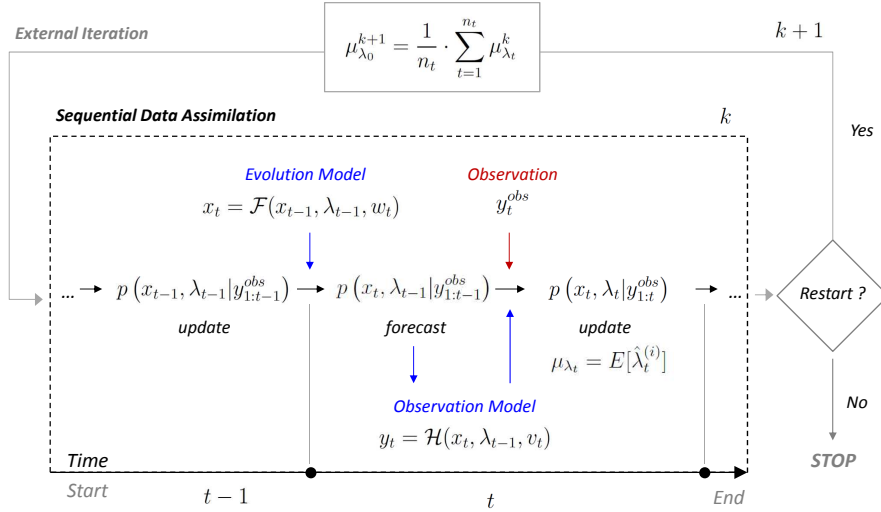


Figure 1: Scheme of the iterative particle filter method (modified from [45]). The data assimilation cycle starts with a distribution of the system state at time $t - 1$ which is used by the evolution model to provide a forecast at time t . The forecast state is converted by the observation model into a forecasted observation which is combined with the field observation y_t to produce the update at time t . When all the available data are assimilated, the data assimilation cycle is restarted (k -th external iteration) until convergence of the model parameter λ_t (see main text for details).

230 relations proposed by [46]:

$$S_w(\psi) = \begin{cases} (1 - S_{wr})(1 + \beta_\psi)^{-m} + S_{wr} & \psi < 0, \\ 1 & \psi \geq 0, \end{cases} \quad (10)$$

$$K_r(\psi) = \begin{cases} (1 + \beta_\psi)^{-m/2} [(1 + \beta_\psi)^m - \beta_\psi^m]^2 & \psi < 0, \\ 1 & \psi \geq 0, \end{cases} \quad (11)$$

231 where S_{wr} is the residual water saturation, $\beta_\psi = (\psi/\psi_s)^\alpha$, ψ_s is the capillary
 232 or air entry pressure, α is a constant and $m = 1 - 1/\alpha$, with $1.25 < \alpha <$
 233 6 . Equation (9) is numerically solved using the subsurface module of the
 234 CATHY model (CATchment HYdrology [47]), a linear tetrahedral finite ele-
 235 ment method with backward Euler scheme with adaptive time stepping and
 236 Newton-like iterations for the solution of nonlinear system [48]. The system
 237 state vector x_t of (1) collects the nodal pressure head ψ at simulation time t .
 238 The nonlinear function \mathcal{F} is a formal representation of the numerical solver
 239 and comprises a number of time steps to advance within the assimilation
 240 interval $[t - 1, t]$. The stochastic noise w_t , kept constant during the forecast
 241 step, represents model uncertainty and is generally specified by a normal or
 242 lognormal distribution of the parameters.

243 3.2. Observation model

244 We monitor the infiltration process with ERT measurements. ERT emits
 245 direct current (DC) from evenly spaced electrodes installed at the soil sur-
 246 face and monitors the electrical potential differences at other locations. The
 247 DC injection pairs are moved sequentially to generate a number of electri-
 248 cal potential fields. Using moisture content-resistivity relationships (e.g.,

249 Archie’s Law [49, 50]) and assuming that changes in conductivity correspond
 250 to changes in moisture content, the water flow in the vadose zone can be mon-
 251 itored [17, 38, 51]. The intensity of the electrical potential field Φ induced in
 252 the soil by the input current can be modeled as [51]:

$$-\vec{\nabla} \cdot \left[\kappa(S_w) \vec{\nabla} \Phi \right] = I [\delta(\vec{r} - \vec{r}_{S+}) - \delta(\vec{r} - \vec{r}_{S-})], \quad (12)$$

253 where κ is the scalar electrical conductivity of the bulk (porous medium
 254 plus contained fluid), I is the applied current, δ is the Dirac delta function,
 255 and \vec{r}_{S+} and \vec{r}_{S-} are the current source and sink electrode position vectors,
 256 respectively. The soil electrical conductivity is related to saturation accord-
 257 ing to the following petrophysical relationship that is derived from Archie’s
 258 law [6]:

$$\kappa(S_w) = \kappa(t_0) \left(\frac{S_w(t)}{S_w(t_0)} \right)^n, \quad (13)$$

259 where $S_w(t_0)$ and $\kappa(t_0)$ are the initial water saturation and the corresponding
 260 initial electrical conductivity of the soil, respectively, and n is a dimensionless
 261 parameter generally calibrated in the lab using soil samples. Since water sat-
 262 uration varies during the infiltration process, the induced electric field is time
 263 dependent. Let y_t^{obs} be the vector collecting the electrical potential differences
 264 that are observed at the measurement electrodes at time t . Equations (10)-
 265 (11), (12) and (13) imply that there exists a nonlinear relation between the
 266 water pressure in the soil and the electrical potential differences at all elec-
 267 trodes. In fact, van Genuchten relations (10)-(11) and Archie’s law (13) allow
 268 us to calculate the soil electrical conductivity field from the water pressure.
 269 Equation (12) is solved numerically using a three-dimensional linear finite
 270 element solver. In order to avoid boundary effects on the simulated electrical

271 potential, the model domain used to simulate the infiltration experiment for
 272 both the hydrological and DC current models is enlarged in the three spatial
 273 directions to accommodate the geophysical simulations. The solution of (12)
 274 gives the electrical potential differences $y_{t,i}$, $i = 1, \dots, N_{obs}$, at the N_{obs} elec-
 275 trode positions to be compared to the corresponding field measurements y_t^{obs} .
 276 The general observation model of equation (2) becomes $y_t = \mathcal{H}(\psi_t)$, where \mathcal{H}
 277 embeds the nonlinear relation between the soil moisture and the electric po-
 278 tential. The observation y_t^{obs} can then be related to the measurement model
 279 using the measurement uncertainties as:

$$y_t^{obs} = y_t (1 + v_t),$$

280 where v_t is the observation error, modeled as an unknown realization of a
 281 normal random variable with zero mean and standard deviation equal to σ_y .
 282 The term v_t incorporates both measurement errors and observation model
 283 uncertainties. From the previous equation and the probability distribution
 284 of v_t we can now explicitly derive the expression for the likelihood function
 285 $p(y_t^{obs}|x_t)$, which in the case of a normal distribution becomes:

$$p(y_t^{obs}|x_t) = C \cdot \exp \left[-\frac{1}{2} \sum_{j=1}^{N_{obs}} \left(\frac{y_{t,j}^{obs} - y_{t,j}}{\sigma_y y_{t,j}} \right)^2 \right],$$

286 where C is a normalization constant. This pdf is estimated from the MC
 287 ensemble, hence completing the overall inversion algorithm.

288 4. Experimental Results

289 The performance of the proposed approach was tested on a controlled
 290 infiltration field experiment. First, using the geometry of the real case study,

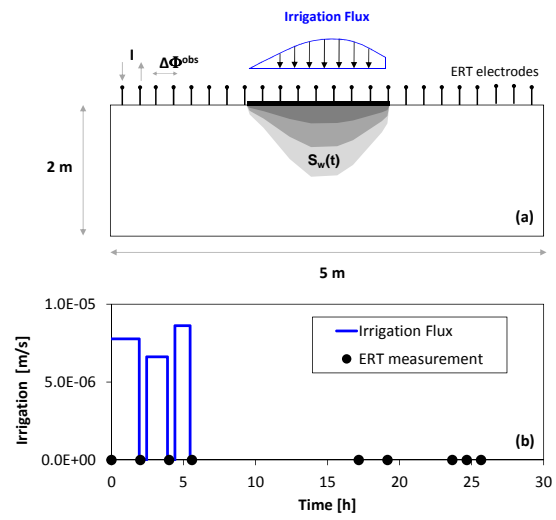


Figure 2: Schematic representation of the system geometry (a) and time-behavior of the infiltration flux rates imposed at the surface boundary (b). Black dots indicate the time of ERT measurements.

Table 1: Time invariant model parameters

Parameter	Description	Unit	Value	Reference
Evolution model				
ϕ	Soil porosity	-	0.33	[52]
S_s	Elastic storage	m^{-1}	5.0E-04	Assumed
S_{wr}	Residual saturation	-	0.003	[53]
ψ_r	Capillary pressure	m	-0.185	[53]
α	VG model parameter	-	2.0	[53]
Observation model				
n	Archie's law parameter	-	1.27	[52]
$S_w(t_0)$	Initial value of S_w	-	0.21	Field data
$\kappa(t_0)$	Initial value of κ	S m^{-1}	7.69E-04	Field data

291 a synthetic problem is designed in order to assess the convergence properties
292 of the developed scheme, then, the real field experiment is simulated.

293 The controlled infiltration experiment is described in [54] and is similar to
294 a previous experiment discussed by [55]. The experimental site is located in
295 Grugliasco (Turin, Italy), nearby the campus of the Agricultural Faculty of
296 the University of Turin. It is characterized by a regular stratigraphic sequence
297 of sandy soil composed mainly of eolic sands with low organic content [52,
298 56]. In the unsaturated zone, sand grains are relatively homogeneous with
299 a median diameter (d_{50}) of 200 μm and porosity of $\phi = 0.33$ forming a
300 homogeneous and isotropic soil in the horizon interested by the infiltration
301 process [52]. The water table is located approximately 20 m below the surface
302 and the vadose zone is not influenced by the underlying aquifer. A line of
303 sprayers was used to wet an area of about 3 m \times 20 m for 6 hours using
304 variable in time irrigation rates (shown in Figure 2(b)).

305 The infiltration front was monitored by means of both ERT and GPR
306 WARR surveys [54] along a cross section of the irrigated area. ERT was
307 performed in time-lapse mode using a dipole-dipole configuration, using 24
308 electrodes placed on the soil surface with a regular spacing of 0.2 m. ERT
309 data were acquired before irrigation (background ERT), during short in-
310 tervals within the irrigation period, and after the end of irrigation for the
311 following 24 hours. The exact timings of the ERT acquisitions used in the
312 data assimilation procedure (i.e. during and after irrigation) are shown as
313 bullets in Figure 2(b).

314 Soil samples at different depths were collected and used to obtain labora-
315 tory estimates of the hydrological parameters S_s , ϕ , α , ψ_s , and S_{wr} , as well

316 as Archie’s law constant n . Initial volumetric water content was estimated
317 from GPR measurements at $0.07 \text{ m}^3 \text{ m}^{-3}$, corresponding to an initial water
318 saturation $S_w(t_0) = 0.21$, while background ERT measurements were used to
319 determine the initial soil electrical conductivity $\kappa(t_0) = 7.69 \times 10^{-4} \text{ S m}^{-1}$,
320 corresponding to a resistivity of $1300 \text{ } \Omega \text{ m}$. This value is in accordance with
321 Archie’s law parameter calibrated during the laboratory experiments [52].
322 The values of these parameters are reported in Table 1.

323 Inverted resistivity data, obtained from the uncoupled approach devel-
324 oped by [4], revealed that irrigation was not uniformly distributed in the
325 direction orthogonal to the sprinkler line, probably due to the presence of
326 wind [54]. This was taken into account in order to properly define the top
327 boundary conditions and the irrigation flux was thus modeled with a Gaus-
328 sian distribution centered at 2.5 m (top boundary), with variance equal to 0.6
329 m, both values calculated such that the total flux equals the real irrigation
330 rate.

331 The model of the field experiment is developed using a vertical cross-
332 section orthogonal to the irrigation line, whose schematic representation is
333 illustrated in Figure 2(a). For the hydrologic simulation, no-flow boundary
334 conditions (BCs) were set all over the model domain, except for the top
335 boundary where the irrigation rate was imposed as a Neumann flux. Spa-
336 tially varying input infiltration is considered as a potential rate, and actual
337 infiltration is evaluated based on system state condition allowing the switch-
338 ing between Neumann and Dirichlet BCs in the case of ponding [47].

339 The finite element grid of the hydrologic model consists of 9792 nodes
340 and 49500 elements while the stationary geophysical model was solved on an

341 enlarged mesh characterized by 21240 nodes and 112404 elements.

342 4.1. Synthetic case

343 In the synthetic cases, a forward simulation of both the hydrological and
 344 the ERT models with pre-imposed parameters was used to generate the true
 345 state and the ERT measurements. We are interested in identifying saturated
 346 homogeneous or spatially heterogeneous hydraulic conductivity, simulated
 347 with a lognormal distribution to ensure positivity of the parameters val-
 348 ues [e.g. 57, 58]. All other model parameters are based on the values used
 349 in the field case study as listed in Table 1. The synthetic dataset of ERT
 350 observations was generated by the coupled hydro-geophysical forward model
 351 assuming the same dipole-dipole configuration of the field experiment. It was
 352 then used to constrain the particle filter simulations assuming different levels
 353 of measurement errors ($\sigma_y = 5 - 20\%$).

354 The convergence of the proposed coupled inversion method is tested by
 355 looking at the behavior of a number of error statistics. The discrepancy be-
 356 tween measured and simulated observations (electrical potential at the elec-
 357 trodes) is evaluated in terms of ensemble mean relative error (ϵ_y), maximum
 358 relative error ($\epsilon_{y,max}$) and root mean square error ($RMSE_y$):

$$\begin{aligned} \epsilon_y &= \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{N_{obs}} \sum_{j=1}^{N_{obs}} \frac{|y_{t,j}^{i,\Phi} - y_{t,j}^{obs}|}{|y_{t,j}^{obs}|} \right] \\ \epsilon_{y,max} &= \max_i \left\{ \max_j \left\{ \frac{|y_{t,j}^{i,\Phi} - y_{t,j}^{obs}|}{|y_{t,j}^{obs}|} \right\} \right\} \\ RMSE_y &= \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{N_{obs}} \sqrt{\sum_{j=1}^{N_{obs}} \frac{|y_{t,j}^{i,\Phi} - y_{t,j}^{obs}|^2}{|y_{t,j}^{obs}|^2}} \right]_i \end{aligned}$$

359 We also look at the L_2 -norm of the error ϵ_{ψ} between the true and the sim-
 360 ulated system state values, soil water pressure head, named the pressure
 361 error:

$$\epsilon_{\psi} = \|\bar{\psi}_t - \psi_t^{true}\|_2$$

362 where $\bar{\psi}_t$ is the ensemble mean pressure field at time t . For all the simulations
 363 we require a fixed number (8) of iterations chosen so that convergence is
 364 reached within a reasonable computational time and a reliable assessment of
 365 error statistics is obtained. The use of one of the stopping criteria proposed
 366 in section 2.2 would yield faster convergence in all test cases.

367 4.1.1. Homogeneous test case

368 In this test case, an isotropic and homogeneous soil with hydraulic con-
 369 ductivity equal to $\mathbf{K}_s = 10^{-5}$ [m s⁻¹] was employed. The saturated hydraulic
 370 conductivity tensor is thus the only unknown parameter $\lambda_t = \{\mathbf{K}_s\}$ with
 371 $K_x = K_y = K_z$ (homogeneous and isotropic soil).

372 A preliminary sensitivity analysis on the ensemble size carried out with
 373 $N = 20, 50, 100$ suggests that 20 particles are enough for this case study to
 374 obtain reliable estimates. Hence, a value $N = 20$ particles is used to test the
 375 performance of the method with the different measurement errors. Figure 3
 376 reports the convergence results in terms of both parameter values (left panel)
 377 and errors (right panel). To better illustrate the behaviour of the pdf of the
 378 hydraulic conductivity during the iterative procedure, the simulation results
 379 obtained with 100 particles are also shown (Figure 4).

380 The hydraulic conductivity estimated by the iterative particle filter method
 381 is shown to converge to the true value \mathbf{K}_s^{true} as the number of updates is suf-
 382 ficiently large (Figure 3). The number of updates necessary for convergence

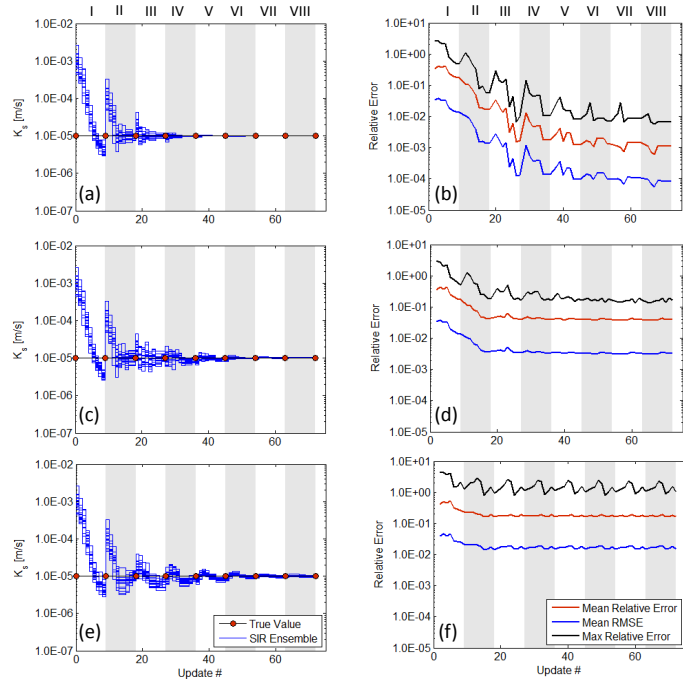


Figure 3: Synthetic test case results: convergence of the hydraulic conductivity (a,c,e) and relative errors between true and simulated observations (b,d,f). Mean relative error (ϵ_y), Mean $RMSE_y$ and Maximum Relative error ($\epsilon_{y,max}$) are shown. The performance of the method for different measurements error is illustrated: (a,b) $\sigma_{\Phi} = 5\%$ with measurements not randomly perturbed, (c,d) $\sigma_{\Phi} = 5\%$, and (e,f) $\sigma_{\Phi} = 20\%$ with randomly perturbed measurements. Red dots indicate the true value of \mathbf{K}_s . The roman numerals indicate the external iteration step. Each external iteration consists of 8 SIR updates

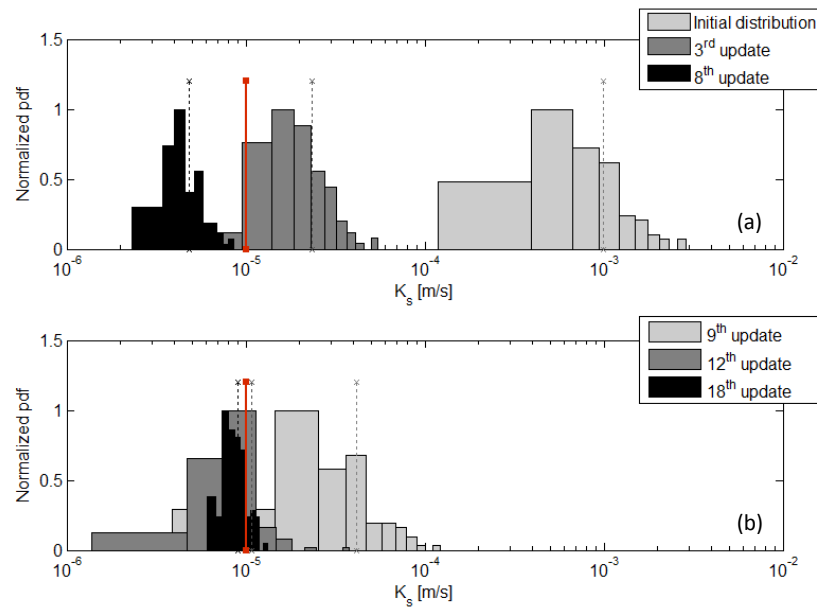


Figure 4: Synthetic test case results: pdf of the hydraulic conductivity normalized on the maximum value of the pdf. Panel (a) and (b) refer to the first and second external iteration of the SIR method, respectively. The simulation was run with an ensemble size $N=100$. Dotted lines indicate the ensemble mean and the red line indicates the true value of K_s .

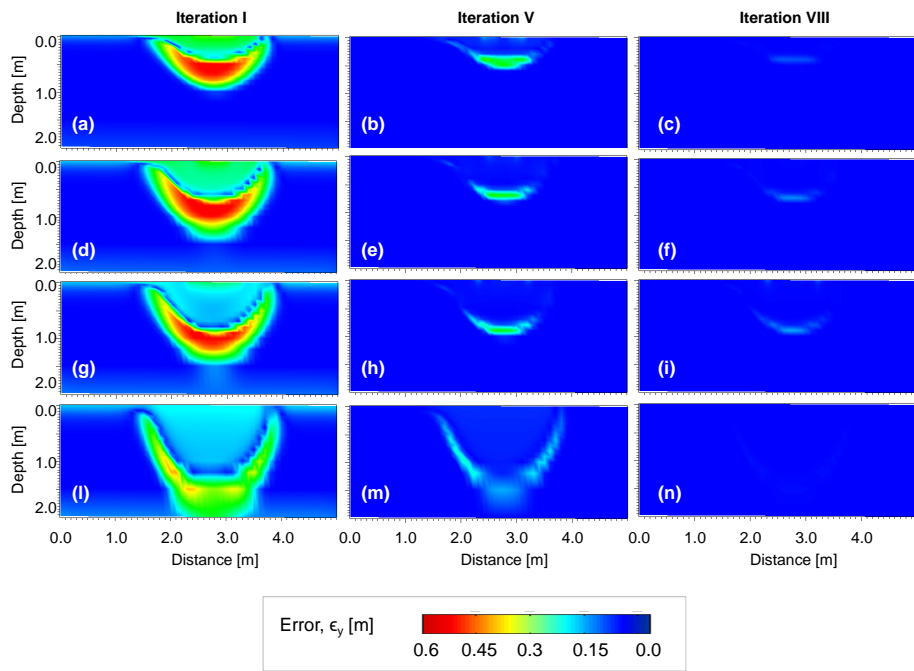


Figure 5: Spatial distribution of the error ϵ_ψ , representing the discrepancy between simulated (ensemble mean) and true system state (pressure field).

383 depends on the measurement error: when the true observations are assimilated, i.e. when the observations are not randomly perturbed, the method
 384 approaches \mathbf{K}_s^{true} after four iterations (Figure 3a) but for increasing noise, more iterations are needed to achieve convergence. As a matter of fact, for
 385 $\sigma_\phi = 5\%$ and 20% the estimated value $\mu_{\lambda_{t,k}}$ keeps oscillating until the 6th
 386 and 7th external iteration, respectively (Figure 3(c) and Figure 3(e)). The convergence speed depends on σ_ϕ , observing slower convergence for higher
 387 noises. The results demonstrate that the traditional particle filter (i.e. the non-iterative approach) may provide a biased estimate of the model parameter
 388 unless larger ensemble sizes are used. This is highlighted in Figure 4 where the pdf of the hydraulic conductivity at different updates of the first
 389 and second iterations are shown. If the initial guess of the model parameter is overestimated, the predicted value at the end of the first iteration
 390 (8th update in Figure 4(a)) is underestimated. This is due to the fact that the particle filter has to correct the model parameter more than necessary
 391 to balance the bias on the predicted state during the initial updates. For example, a higher initial estimate of \mathbf{K}_s corresponds to a higher infiltration
 392 capacity and thus causes an over-estimated total infiltrated water, with a corresponding over-estimation of the front speed. Hence, at later times, the
 393 inversion procedure must identify an under-estimated \mathbf{K}_s to accommodate the slower observed saturation front depth. As a result, the pdf of the parameter
 394 is shifted further than necessary on the opposite direction of the initial guess. The iterative approach allows the filter to “forget” the initial bias and
 395 converge more efficiently to the true parameter (Figure 4(b)). The results in Figure 5 show that the error ϵ_ψ develops at the edge of the infiltration front
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407

408 where sensitivities are highest. The iterative procedure successfully reduces
409 the discrepancy between simulated and true system state and the restart is
410 shown to be fundamental to achieve negligible errors. The traditional SIR
411 method corrects also the system state after each update but errors up to
412 0.6 m (in term of predicted pressure head) are still observed at the end of
413 the first iteration of the sequential procedure. The iterated approach allows
414 instead a reduction of the error ϵ_y down to negligible values ($\epsilon_y < 10^{-3}$ m).
415 The synthetic simulations confirmed that the particle filter is an efficient
416 method to update the system state and the iterative procedure is shown to
417 be essential to provide precise estimates of the model parameters at lower
418 computational effort.

419 *4.1.2. Heterogeneous test case*

420 The ability of the proposed methodology to estimate multiple model pa-
421 rameters is investigated. We consider the same infiltration experiment, now
422 characterized by an isotropic heterogenous soil (Figure 6(a-b)). The model
423 domain is divided into four zones with different hydraulic conductivities (thus
424 providing four unknown model parameters). The physical setting in Fig. 6(a)
425 is not intended to represent typical field conditions but aims to provide a sim-
426 ple setup generating both vertical and lateral infiltration patterns to test the
427 proposed approach in a truly multidimensional heterogeneous setting. The
428 results of the iterative SIR scheme, shown in Figure 6(c-d), demonstrate that
429 this approach successfully estimates multiple model parameters. To assess
430 the sensitivity to the initial condition, we simulated the same test problem
431 with different values of the initial guess. Figure 7 reports an example of
432 the identification results in the case of underestimated initial solutions. We

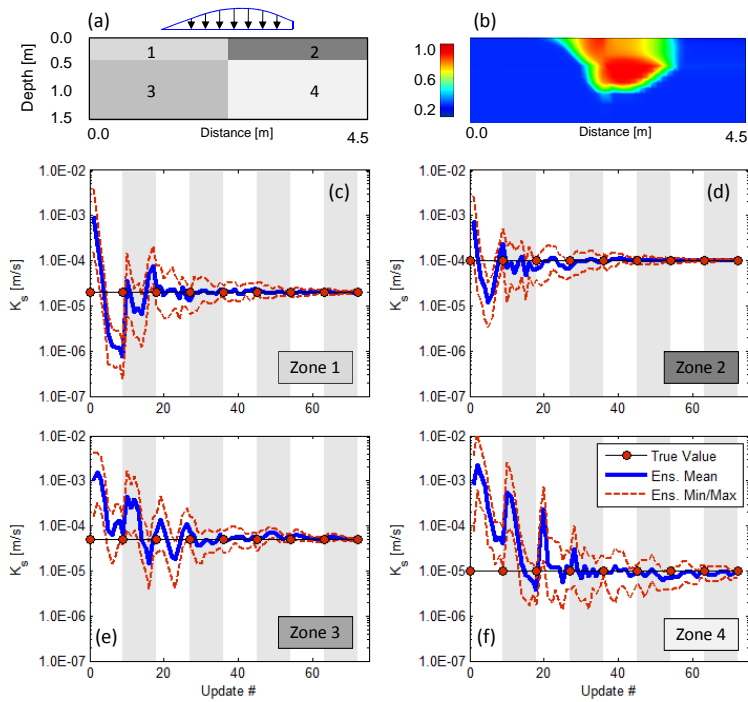


Figure 6: Heterogeneous test case results: (a) conceptual model of the model domain (divided into 4 zones with different soil properties) and (b) the simulated soil saturation at $t = 5.5h$. Convergence of the hydraulic conductivities of the four zones is shown in panels c-f. The results are relative to $\sigma_\phi = 5\%$ with randomly perturbed measurements.

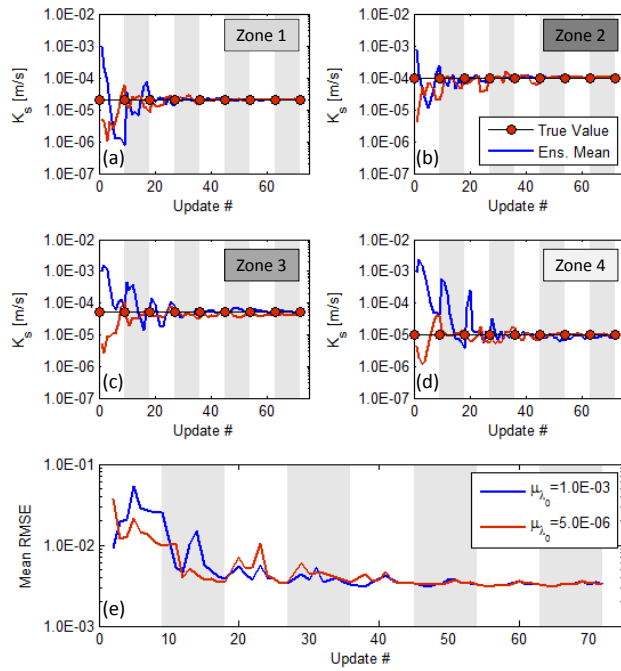


Figure 7: Heterogeneous test case results: convergence of the hydraulic conductivities (ensemble mean values) of the four zones (a-d) and mean RMSE (e) for different initial values μ_{λ_0} . The results are relative to $\sigma_{\Phi} = 5\%$ with randomly perturbed measurements.

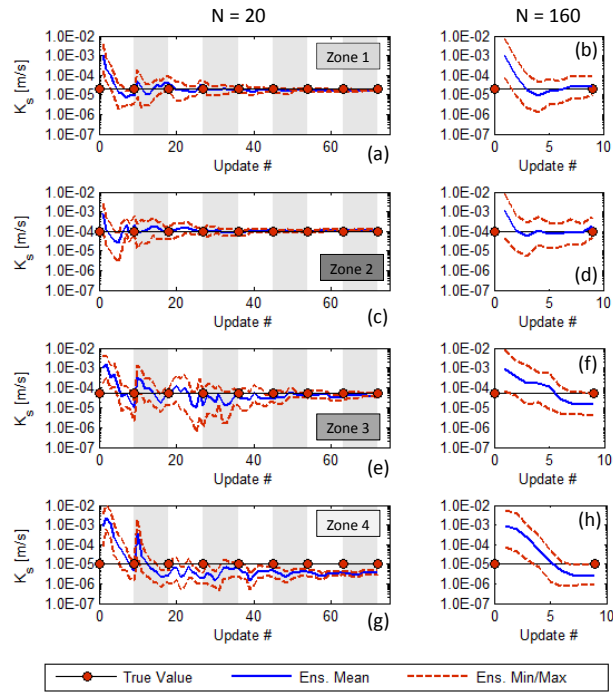


Figure 8: Heterogeneous test case results: comparison of the iterative approach ($N = 20$) with a non-iterative simulation with ensemble size $N = 160$. The convergence of the four hydraulic conductivities for the iterative (panels a,c,e,g) and non-iterative (panels b,d,f,h) cases is illustrated (runs with $\sigma_\phi = 20\%$ and randomly perturbed measurements).

433 notice that the behavior of the iterative SIR method is qualitatively similar
434 independently on the initial solution, thus confirming the reliability of the
435 proposed approach.

436 We note that the identification is practically achieved after four iterations,
437 for a total of 80 forward model runs. At later iterations the identified values of
438 zones 3 and 4 display small oscillations whose amplitude seem to decrease as
439 the scheme progresses (Figure 6(e-f)). This is likely due to the fact that both
440 zones 3 and 4 receive information from the infiltration experiment at later
441 times. At the first 4 observation times the true infiltration front is shallower
442 than the material interface, and only the last 4 measurements contribute
443 information towards the identification of hydraulic conductivity of zones 3
444 and 4.

445 To test the improvements obtained by our proposed iterative method
446 with respect to standard (non iterative) DA methods, we solve the same
447 problem with a one-iteration SIR approach but with an ensemble size $N =$
448 160. This value corresponds the same number of forward model runs used in
449 the previous simulations using (pre-fixed) eight iterations. We perform this
450 comparison for the case of $\sigma_\phi = 20\%$ and randomly perturbed measurements.

451 The convergence results of the iterative and non-iterative procedures for
452 this case are compared in Figure 8. The iterated simulation converges to the
453 correct hydraulic conductivities of zones 1, 2 and 3, and only a small dis-
454 crepancy persists in the estimation of \mathbf{K}_s in zone 4. The value of this bias is
455 consistent with the 20% measurement uncertainty, implying that the inverse
456 procedure has arrived at the correct solution. On the contrary, the results for
457 the non-iterative SIR show a bias in the identification of the parameters of

458 zones 3 and 4 that is larger than the variability dictated by the measurement
459 error. The corresponding ensemble means underestimate the true values,
460 reflecting the earlier observation that starting from a large \mathbf{K}_s leads to an
461 underestimation of the parameter value. The final posterior distributions
462 of the parameters have a higher ensemble variance than the corresponding
463 iterative-results, yielding an uncertain characterization of the soil structure.
464 The non-iterative SIR procedure shows a parameter distribution with strong
465 variations during the assimilation, corresponding to a large variance of the
466 posterior distribution.

467 *4.2. Field experiment*

468 The results of the field data inversion are shown in Figure 9. The as-
469 simulation of ERT measurements provides similar results to the synthetic
470 test case, thus confirming the reliability of the method. The iterative par-
471 ticle filter is shown to converge to a value of hydraulic conductivity \mathbf{K}^*
472 which is independent to the initial guess μ_{λ_0} . As a matter of fact, start-
473 ing from $\mu_{\lambda_0} = 10^{-3} \text{ ms}^{-1}$ the method provides a final estimate $\mathbf{K}^* =$
474 $8.9 \times 10^{-6} \pm 3.6 \times 10^{-7} \text{ ms}^{-1}$ and starting from $\mu_{\lambda_0} = 10^{-7} \text{ ms}^{-1}$, the fi-
475 nal estimate is $\mathbf{K}^* = 9.8 \times 10^{-6} \pm 2.9 \times 10^{-7} \text{ ms}^{-1}$. Note that in both cases
476 the initial guess is two orders of magnitude away from the final estimate
477 and the two final intervals for the identified parameter value are overlapping.
478 It must be emphasized that the method does not provide just an estimate
479 of hydraulic conductivity but a full probability distribution of the estimate.
480 As shown by the synthetic test, in the case of large measurement noise, the
481 relative errors slightly decrease during the first updates and quickly stabi-
482 lize. The residual errors are larger than observed in the synthetic test case

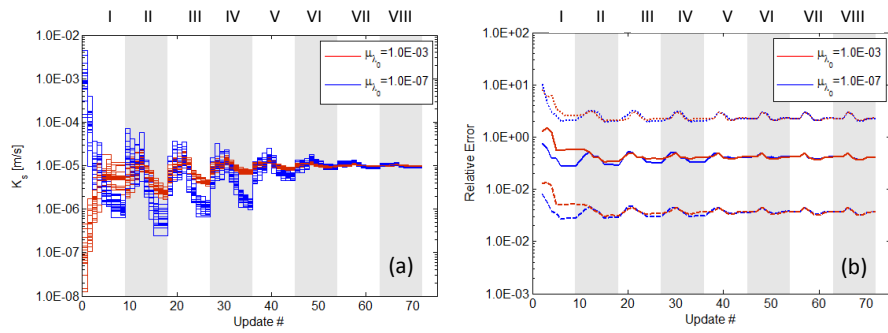


Figure 9: Field experiment results: convergence of the hydraulic conductivity (a) and relative errors (b) for different initial values of hydraulic conductivity μ_{λ_0} . The roman numerals at the top of the panels indicate the external iteration count. Mean relative error (ϵ_y), Mean $RMSE_y$ and Maximum Relative error ($\epsilon_{y,max}$) are shown.

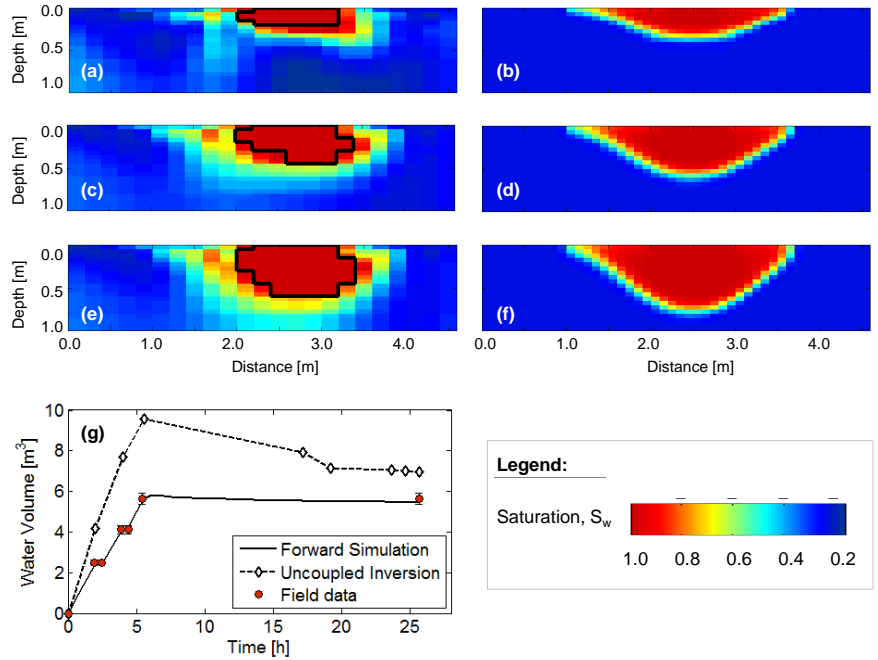


Figure 10: Time-lapse soil saturation estimated by uncoupled inversion (a,c,e) and by the forward simulation with the hydraulic conductivity estimated by the coupled approach (b,d,f). The results are shown at (a,b) $t = 2h$, (c,d) $t = 4h$, and (e,f) $t = 5.5h$. The black contour indicates the area where uncoupled inversion provides unphysical saturation estimates ($S_w > 1$). Mass balance (g): the forward simulation (black line) matches the volume of water injected at the site (red circles with a 5% error bar) while the estimate from uncoupled inversion of ERT data overestimates the mass of water in the system.

483 thus indicating a bias due to external factors not accounted for in the model
484 setup. The hydraulic conductivity estimated by the iterative particle filter
485 method is shown to converge to the \mathbf{K}^* value. However, the reliability of the
486 estimate has to be proven. For this purpose, a forward hydrologic simulation
487 is run with $\mathbf{K}_s = \mathbf{K}^*$ and the results are compared with field observations
488 (Figure 10). The robustness of the estimated parameter is confirmed by the
489 spatial agreement of simulated soil moisture fields obtained by the coupled
490 and uncoupled inversion procedures (Figure 10(a-f)) and by the excellent
491 agreement between the amount of injected water and the predicted mass
492 balance (Figure 10(g)). Further comparison between the forward simulation
493 and field data are presented in [54] where the simulated infiltration is shown
494 to match the front depth estimated by the GPR survey. The discrepancy
495 between the simulation and the time-lapse saturation estimated by uncou-
496 pled inversion increases for increasing front depth. As a matter of fact the
497 resolution of traditional ERT inversion decreases with depth and, given the
498 electrode configuration used in this study, the inverted resistivity is not re-
499 liable for depth higher than 1 m. In addition, the conversion of inverted
500 resistivity to soil saturation by Archie’s law (calibrated in the lab) provided
501 regions of $S_w > 1$ (black contour in Figure 10). Even though these regions
502 can be corrected empirically to ensure a consistent saturation field (accord-
503 ing to common practice in geophysical applications anyway) the uncoupled
504 approach over-estimates the total water present in the system at any time
505 (Figure 10). Therefore, while the forward simulation provides a full conser-
506 vation of mass, the traditional inversion approach provides a good qualitative
507 description of the physical process but does not ensure a correct mass balance

508 (Figure 10g).

509 **5. Discussion**

510 The results presented in this paper demonstrate the accuracy and ro-
511 bustness of the proposed iterative methodology and highlight the weaknesses
512 of both, uncoupled ERT inversion and traditional particle filter applications
513 with ERT data. As shown in Figures 3, 4, and 8 for the synthetic test cases
514 and in Figure 9 for the field simulations, a single iteration of the particle filter
515 method does not provide a reliable estimate of the soil hydraulic conductiv-
516 ity. To verify this hypothesis, we use as initial guess the parameter value
517 $\mu_{\lambda_0} = 10^{-3} \text{ m s}^{-1}$ and then employ the identified parameter μ_{λ_8} estimated
518 at the end of the first iteration to run a forward simulation of the infiltration
519 experiment. In this case, the irrigation intensity is found to be higher than
520 the infiltration capacity, thus leading to surface ponding not observed at the
521 site during the experiment. Therefore, if the particle filter is used to esti-
522 mate the model parameters without enough updates to ensure convergence,
523 the method may lead to wrong predictions of the system dynamics. The re-
524 sults of our simulations further show that a non-iterative SIR approach with
525 a large ensemble is not fully capable of performing a correct identification,
526 suggesting that the iterative approach is computationally more efficient for
527 solving the problem of interest.

528 The proposed coupled hydro-geophysical modeling framework presents
529 the following advantages compared to more traditional approaches: (1) a
530 forward geophysical model is used and the inversion of the geophysical data
531 is avoided thus guaranteeing physical consistency with the hydrologic quan-

532 tities; (2) the sequential approach provides a dynamic correction of the sim-
533 ulated system state, thus correcting intrinsic model errors (i.e. unknown
534 initial condition), with relatively small computational requirements; (3) the
535 data assimilation approach is particularly interesting for field applications
536 where the geophysical measurements can be affected by external factors (e.g.
537 soil evaporation, a rainfall event during the geophysical survey, etc.) that
538 can be easily included in the hydro-geophysical modeling framework; (4) the
539 filtering approach describes quantitatively both model and observation er-
540 rors, and provides the probability density functions of both system state and
541 model parameters.

542 **6. Conclusions**

543 A sequential Bayesian approach for coupled hydro-geophysical assimi-
544 lation of ERT measurements in a variably saturated flow model is presented.
545 An innovative iterative approach is proposed to achieve accurate identifica-
546 tion of the model parameters. The robustness of the methodology is tested
547 on spatially homogeneous and heterogeneous synthetic test cases and vali-
548 dated on a field infiltration experiment. We show that the new approach has
549 several advantages compared to uncoupled inversion and traditional sequen-
550 tial data assimilation techniques. In particular the iterative particle filter
551 provides accurate parameter estimation as opposed to traditional SIR that
552 may lead to biased results. Further work will focus on testing the method-
553 ology for the estimation of multiple and spatially varying parameters (e.g.
554 Archie's law, retention curves, heterogeneous soil, etc.).

555 **7. Acknowledgments**

556 This study was funded by the University of Padova, Italy, within the
557 Research Programme “GEO-RISKS: Geological, morphological and hydro-
558 logical processes: monitoring, modeling and impact in the north-eastern
559 Italy”, WP4. The authors would like to acknowledge partial funding from
560 the EU FP7 project CLIMB (“Climate Induced Changes on the Hydrology
561 of Mediterranean Basins - Reducing Uncertainty and Quantifying Risk”) and
562 the Italian Ministry of Education, Universities and Research, PRIN 2010-11
563 (“Innovative methods for water resources management under hydro-climatic
564 uncertainty scenarios”).

565 **References**

- 566 [1] W. Daily, A. Ramirez, D. J. La Brecque, J. Nitao, Electrical resistivity
567 tomography of vadose zone water movement, *Water Resour. Res.* 28
568 (1992) 1429–1442.
- 569 [2] T. C. J. Yeh, J. Šimůnek, Stochastic fusion of information for char-
570 acterizing and monitoring the vadose zone, *Vadose Zone J.* 1 (2002)
571 207–221.
- 572 [3] Q. Y. Zhou, J. Shimada, A. Sato, Temporal variations of the three-
573 dimensional rainfall infiltration process in heterogeneous soil, *Water*
574 *Resour. Res.* 38 (2002) 1–1–1–15.
- 575 [4] A. Binley, A. Kemma, Dc resistivity and induced polarization meth-
576 ods, in: Y. Rubin, S. S. Hubbard (Eds.), *Hydrogeophysics*, volume 50,
577 Springer, 2005, pp. 129–156.

- 578 [5] G. Cassiani, N. Ursino, R. Deiana, G. Vignoli, J. Boaga, M. Rossi, M. T.
579 Perri, M. Blaschek, R. Duttmann, S. Meyer, R. Ludwig, A. Soddu, P. Di-
580 etrich, U. Werban, Non-invasive monitoring of soil static characteristics
581 and dynamic states: a case study highlighting vegetation effects, *Vadose*
582 *Zone J.* 11 (2012) vzt2011.0195.
- 583 [6] G. E. Archie, The electrical resistivity log as an aid in determining some
584 reservoir characteristics, *Trans. Am. Inst. Min. Metall. Eng.* 146 (1942)
585 54–61.
- 586 [7] T. Ha, S. Pyun, C. Shin, Efficient electric resistivity inversion using
587 adjoint state of mixed finite-element method for Poissons equation, *J.*
588 *Comp. Phys.* 214 (2006) 171–186.
- 589 [8] E. Chung, T. Chan, X. Tai, Electrical impedance tomography using
590 level set representations and total variation regularization, *J. Comp.*
591 *Phys.* 205 (2005) 357–372.
- 592 [9] K. van den Doel, U. M. Ascher, On level set regularization for highly
593 ill-posed distributed parameter estimation problems, *J. Comp. Phys.*
594 216 (2006) 707–723.
- 595 [10] W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory*, Else-
596 vier, New York, 1984.
- 597 [11] J. Rings, C. Hauck, Reliability of resistivity quantification for shallow
598 subsurface water processes, *J. Appl. Geophys.* 68 (2009) 404–416.
- 599 [12] D. J. La Brecque, G. Heath, R. Sharpe, R. Versteeg, Autonomous

- 600 monitoring of fluid movement using 3-d electrical resistivity tomography,
601 J. Environ. Eng. Geoph. 9 (2004) 167–176.
- 602 [13] A. M. Tartakovsky, D. Bolster, D. M. Tartakovsky, Hydrogeophysical
603 approach for identification of layered structures of the vadose zone from
604 electrical resistivity data, Vadose Zone J. 7 (2008) 1–8.
- 605 [14] K. Singha, S. M. Gorelick, Saline tracer visualized with three-
606 dimensional electrical resistivity tomography: Field-scale spatial mo-
607 ment analysis, Water Resour. Res. 41 (2005) W05023.
- 608 [15] D. Michot, Y. Benderitter, A. Dorigny, B. Nicoullaud, D. King, A. Tab-
609 bagh, Spatial and temporal monitoring of soil water content with an
610 irrigated corn crop cover using surface electrical resistivity tomography,
611 Water Resour. Res. 39 (2003) 1138.
- 612 [16] J. Travelletti, P. Sailhac, J. P. Malet, G. Grandjean, J. Ponton, Hy-
613 drological response of weathered clay-shale slopes: water infiltration
614 monitoring with time-lapse electrical resistivity tomography, Hydrol.
615 Process. 26 (2012) 2106–2119.
- 616 [17] A. C. Hinnell, T. P. A. Ferré, J. A. Vrugt, J. A. Huisman, S. Moysey,
617 J. Rings, M. Kowalsky, Improved extraction of hydrologic information
618 from geophysical data through coupled hydrogeophysical inversion, Wa-
619 ter Resour. Res. 46 (2010) W00D40.
- 620 [18] G. A. Hansen, C. Penland, On stochastic parameter estimation using
621 data assimilation, Physica D 230 (2007) 88–98.

- 622 [19] P. J. Smith, S. L. Dance, M. J. Baines, N. K. Nichols, T. R. Scott,
623 Variational data assimilation for parameter estimation: application to a
624 simple morphodynamic model, *Ocean Dyn.* 56 (2009) 697–708.
- 625 [20] M. Camporese, C. Paniconi, M. Putti, P. Salandin, Ensemble Kalman
626 filter data assimilation for a process-based catchment scale model of
627 surface and subsurface flow, *Water Resour. Res.* 45 (2009) W10421.
- 628 [21] G. Evensen, The ensemble Kalman filter: theoretical formulation and
629 practical implementation, *Ocean Dyn.* 53 (2003) 343–367.
- 630 [22] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic
631 Press, New York, 1970.
- 632 [23] P. Gauthier, P. Courtier R, P. Moll, Assimilation of simulated wind lidar
633 data with a Kalman filter, *Mon. Weather Rev.* 121 (1993) 1803–1820.
- 634 [24] M. S. Arulampalam, B. Ristic, Comparison of the particle filter with
635 range-parametrized and modified polar EKF's for angle-only tracking,
636 in: *Proc. SPIE*, volume 4048, pp. 288–299.
- 637 [25] D. Pasetto, M. Camporese, M. Putti, Ensemble Kalman filter versus
638 particle filter for a physically-based coupled surface-subsurface model,
639 *Adv. Water Resources* 47 (2012) 1–13.
- 640 [26] N. J. Gordon, D. J. Salmond, A. F. M. Smith, Novel approach to
641 nonlinear/non-Gaussian Bayesian state estimation, *IEE Proc.-F* 140
642 (1993) 107–113.

- 643 [27] A. Doucet, S. Godsill, C. Andrieu, On sequential Monte Carlo sampling
644 methods for Bayesian filtering, *Stat. Comput.* 10 (2000) 197–208.
- 645 [28] H. Moradkhani, K.-L. Hsu, H. Gupta, S. Sorooshian, Uncertainty assess-
646 ment of hydrologic model states and parameters: Sequential data assim-
647 ilation using the particle filter, *Water Resour. Res.* 41 (2005) W05012.
- 648 [29] A. H. Weerts, G. Y. H. El Serafy, Particle filtering and ensemble Kalman
649 filtering for state updating with hydrological conceptual rainfall-runoff
650 models, *Water Resour. Res.* 42 (2006) W09403.
- 651 [30] Y. Zhou, D. McLaughlin, D. Entekhabi, Assessing the performance of
652 the ensemble Kalman filter for land surface data assimilation, *Mon.*
653 *Weather Rev.* 134 (2006) 2128–2142.
- 654 [31] K. L. Hsu, H. Moradkhani, S. Sorooshian, A sequential Bayesian ap-
655 proach for hydrologic model selection and prediction, *Water Resour.*
656 *Res.* 45 (2009) W00B12.
- 657 [32] J. A. Vrugt, C. G. H. Diks, H. V. Gupta, W. Bouten, J. M. Verstraten,
658 Improved treatment of uncertainty in hydrologic modeling: Combining
659 the strengths of global optimization and data assimilation, *Water Re-*
660 *sour. Res.* 41 (2005) W01017.
- 661 [33] H.-J. Hendricks Franssen, W. Kinzelbach, Real-time groundwater flow
662 modeling with the ensemble Kalman filter: Joint estimation of states
663 and parameters and the filter inbreeding problem, *Water Resour. Res.*
664 44 (2008) W09408.

- 665 [34] P. Salamon, L. Feyen, Assessing parameter, precipitation, and predictive
666 uncertainty in a distributed hydrological model using sequential data
667 assimilation with the particle filter, *J. Hydrol.* 376 (2009) 428–442.
- 668 [35] M. Camporese, C. Paniconi, M. Putti, P. Salandin, Comparison of data
669 assimilation techniques for a coupled model of surface and subsurface
670 flow, *Vadose Zone J.* 8 (2009) 837–845.
- 671 [36] D. A. Plaza, R. De Keyser, G. J. M. De Lannoy, L. Giustarini, P. Mat-
672 gen, V. R. N. Pauwels, The importance of parameter resampling for
673 soil moisture data assimilation into hydrologic models using the particle
674 filter, *Hydrol. Earth Syst. Sci.* 16 (2012) 375–390.
- 675 [37] C. Montzka, H. Moradkhani, L. Weihermüller, H.-J. Hendricks Franssen,
676 M. Canty, H. Vereecken, Hydraulic parameter estimation by remotely-
677 sensed top soil moisture observations with the particle filter, *J. Hydrol.*
678 399 (2011) 410–421.
- 679 [38] J. Rings, J. A. Huisman, H. Vereecken, Coupled hydrogeophysical pa-
680 rameter estimation using a sequential Bayesian approach, *Hydrol. Earth*
681 *Syst. Sci.* 14 (2010) 545–556.
- 682 [39] H. Haario, E. Saksman, J. Tamminen, An adaptive Metropolis algo-
683 rithm, *Bernoulli* 7 (2001) 223–242.
- 684 [40] J. A. Vrugt, C. J. F. ter Braak, C. G. H. Diks, D. Higdon, B. A. Robin-
685 son, J. M. Hyman, Accelerating Markov chain Monte Carlo simulation
686 by differential evolution with self-adaptive randomized subspace sam-
687 pling, *Int. J. Nonlinear Sci. Numer. Simul.* 10 (2009) 273–290.

- 688 [41] J. S. Liu, R. Chen, Sequential Monte Carlo methods for dynamical
689 systems, *J. Am. Stat. Assoc.* 93 (1998) 1032–1044.
- 690 [42] G. Kitagawa, Monte Carlo filter and smoother for non-Gaussian non-
691 linear state space models, *J. Comput. Graph. Stat.* 5 (1996) 1–25.
- 692 [43] H. Moradkhani, C. M. DeChant, S. Sorooshian, Evolution of ensem-
693 ble data assimilation for uncertainty quantification using the particle
694 filter-Markov chain Monte Carlo method, *Water Resour. Res.* 48 (2012)
695 W12520.
- 696 [44] J. A. Vrugt, C. J. F. ter Braak, C. G. H. Diks, G. Schoups, Hydrologic
697 data assimilation using particle markov chain Monte Carlo simulation:
698 Theory, concepts and applications, *Adv. Water Resources* 51 (2012)
699 457–478.
- 700 [45] M. Dowd, Bayesian statistical data assimilation for ecosystem models
701 using Markov Chain Monte Carlo, *J. Marine Syst.* 68 (2007) 439–456.
- 702 [46] M. T. van Genuchten, D. R. Nielsen, On describing and predicting
703 the hydraulic properties of unsaturated soils, *Ann. Geophys.* 3 (1985)
704 615–628.
- 705 [47] M. Camporese, C. Paniconi, M. Putti, S. Orlandini, Surface-subsurface
706 flow modeling with path-based runoff routing, boundary condition-based
707 coupling, and assimilation of multisource observation data, *Water Re-
708 sour. Res.* 46 (2010) W02512.
- 709 [48] C. Paniconi, M. Putti, A comparison of Picard and Newton iteration

- 710 in the numerical-solution of multidimensional variably saturated flow
711 problems, *Water Resour. Res.* 30 (1994) 3357–3374.
- 712 [49] A. Brovelli, G. Cassiani, E. Dalla, F. Bergamini, D. Pitea, A. M. Binley,
713 Electrical properties of partially saturated sandstones: a novel compu-
714 tational approach with hydro-geophysical applications, *Water Resour.*
715 *Res.* 41 (2005) W08411.
- 716 [50] A. Brovelli, G. Cassiani, Combined estimation of effective electrical
717 conductivity and permittivity for soil monitoring, *Water Resour. Res.*
718 47 (2011) W08510.
- 719 [51] V. Nenna, A. Pidlisecky, R. Knight, Application of an extended Kalman
720 filter approach to inversion of time-lapse electrical resistivity imaging
721 data for monitoring recharge, *Water Resour. Res.* 47 (2011) W10525.
- 722 [52] G. Cassiani, A. Kemna, A. Villa, E. Zimmermann, Spectral induced
723 polarization for the characterization of free-phase hydrocarbon contam-
724 ination of sediments with low clay content, *Near Surf. Geophys.* 7 (2009)
725 547562.
- 726 [53] D. Canone, S. Ferraris, G. Sander, R. Haverkamp, Interpretation of
727 water retention field measurements in relation to hysteresis phenomena,
728 *Water Resour. Res.* 44 (2008) W00D12.
- 729 [54] M. Rossi, G. Manoli, D. Pasetto, R. Deiana, S. Ferraris, M. Putti,
730 G. Cassiani, Quantitative hydro-geophysical monitoring and coupled
731 modeling of a controlled infiltration experiment, Submitted (2013).

- 732 [55] G. Cassiani, M. Giustianini, S. Ferraris, R. Deiana, C. Strobbia, Time-
733 lapse surface-to-surface gpr measurements to monitor a controlled in-
734 filtration experiment, *Bollettino di Geofisica Teorica ed Applicata* 50
735 (2009) 209–226.
- 736 [56] M. Baudena, I. Bevilacqua, D. Canone, S. Ferraris, M. Previati,
737 A. Provenzale, Soil water dynamics at a midlatitude test site: Field
738 measurements and box modeling approaches, *J. Hydrol.* 414-415 (2012)
739 329–340.
- 740 [57] R. J. Hoeksema, P. K. Kitanidis, Analysis of the spatial structure of
741 properties of selected aquifers, *Water Resour. Res.* 21 (1985) 563–572.
- 742 [58] G. Dagan, *Flow and Transport in Porous Formations*, Springer, New
743 York, 1989.