# Application of ensemble transform data assimilation methods for parameter estimation in reservoir modelling

Sangeetika Ruchi[1] and Svetlana Dubinkina[1]

[1]Centrum Wiskunde & Informatica, P.O. Box 94079, 1098 XG Amsterdam, The Netherlands

**Correspondence:** Sangeetika Ruchi (s.ruchi@cwi.nl)

**Abstract.** Over the years data assimilation methods have been developed to obtain estimations of uncertain model parameters by taking into account a few observations of a model state. The most reliable methods of MCMC are computationally expensive and sequential ensemble methods such as ensemble Kalman filers and particle filters provide with a favourable alternative. However, Ensemble Kalman Filter has an assumption of Gaussianity. Ensemble Transform Particle Filter does not have this assumption and has proven to be highly beneficial for an initial condition estimation and a small number of parameter estimation in chaotic dynamical systems with non-Gaussian distributions. In this paper we employ Ensemble Transform Particle Filter (ETPF) and Ensemble Transform Kalman Filter (ETKF) for parameter estimation in nonlinear problems with 1, 5, and 2500 uncertain parameters and compare them to importance sampling (IS). The large number of uncertain parameters is of a particular interest for subsurface reservoir modelling as it allows to parametrise permeability on the grid. We prove that the updated parameters obtained by ETPF lie within the range of an initial ensemble, which is not the case for ETKF. We examine the performance of ETPF and ETKF in a twin experiment setup, where observations of pressure are synthetically created based on the know values of parameters. For small number of uncertain parameters (1 and 5) ETPF performs comparably to ETKF in terms of the mean estimation. For large number of uncertain parameters (2500) ETKF is robust with respect to the initial ensemble while ETPF is sensitive due to sampling error. An issue of an increase in the root mean square error after data assimilation is performed in ETPF for a high-dimensional test problem is resolved by applying distance-based localization, which however deteriorated the posterior estimation of the leading mode by largely increasing the variance, which is due to a combination of less varying localized weights, not keeping the imposed bounds on the modes via the Karhunen-Loeve expansion and the main variability explained by the leading mode. A possible remedy is instead of applying localization to use only leading modes that are well estimated by ETPF, which demands a knowledge at which mode to truncate.

## 1 Introduction

An accurate estimation of subsurface geological properties like permeability, porosity etc. is essential for many fields specially where such predictions can have large economic or environmental impact, for instance prediction of oil or gas reservoir locations. Knowing the geological parameters a so-called forward model is solved for the model state and a prediction can be made. The subsurface reservoirs, however, are buried thousands of feet below the earth surface and exhibit a highly heterogeneous structure, which makes it difficult to obtain their geological parameters. Usually a prior information about the parameters is

given, which still needs to be corrected by observations of pressure and production rates. These observations are, however, known only at well locations that are often hundreds of meter apart and corrupted by errors. This gives instead of a well-posed forward problem an ill-posed inverse problem of estimating uncertain parameters, since many possible combinations of parameters can result in equally good matches to the observations.

5    Different inverse problem approaches for groundwater and petroleum reservoir modelling, generally termed as history matching, have been developed over the past years, e.g. Oliver et al. (1997) implemented Markov chain Monte Carlo methods with different perturbations and tested it on a 2-D reservoir model; Reynolds et al. (1996) obtained reservoir parameters estimations using Gauss-Newton method; Vefring et al. (2006) used Levenberg–Marquardt method to characterize reservoir pore pressure and permeability. A review of history matching developments is written by Oliver and Chen (2011).

10    For reservoir models the term data assimilation and history matching are used interchangeably, as the goal of data assimilation is the same as that of history matching, where observations are used to improve a solution of a model. Ensemble data assimilation methods such as Ensemble Kalman filters (Evensen, 2009) have been originally developed in meteorology and oceanography for the state estimation. Now it is one of the frequently employed approaches for parameter estimation in subsurface flow models as well (e.g. Oliver et al., 2008). A detailed review of ensemble Kalman filter developments in reservoir

15    engineering is written by Aanonsen et al. (2009). An ensemble Kalman filter efficiently approximates a true posterior distribution if the distribution is not far from Gaussian, as it corrects only the mean and the variance. For nonlinear models with multimodal distributions, however, an ensemble Kalman filter fails to correctly estimate the posterior, as shown by Dovera and Della Rossa (2011).

Importance Sampling (IS) is quite promising for such models as it does not have any assumptions of Gaussianity. It is

20    also an ensemble based method in which the probability density function is represented by a number of samples. One sample corresponds to one configuration of uncertain model parameters. The forward model is solved for each sample and predicted data is computed. The weight is assigned to samples based on the observations of the true physical system and the predicted data. The drawback of IS is that it does not update the uncertain parameters but only their weight, thus a computationally unaffordable ensemble is required. In order to decrease this cost a family of particle filters (Doucet et al., 2001) has been

25    developed where IS is with resampling and a sample is called particle. A significant work for parameter estimation using particle filtering has been done in hydrology. Moradkhani et al. (2005) used it to estimate model parameters and state posterior distributions for a rainfall-runoff model. Weerts and El Serafy (2006) compared an ensemble Kalman filter and a particle filter with different resampling strategies for a rainfall-runoff forecast and obtained that as the number of particles increases the particle filter outperforms the ensemble Kalman filter. Guingla et al. (2012) employed particle filtering to correct the soil

30    moisture and to estimate hydraulic parameters.

The resampling in particle filtering is, however, stochastic. Ensemble Transform Particle Filter (ETPF) developed by Reich and Cotter (2015) is a particle filtering method that *deterministically* resamples the particles based on their weights and covariance maximization among the particles. ETPF has been used for initial condition estimations and for parameter estimations in chaotic dynamical systems with a small number of uncertain parameters (Lorenz 63 model). It has not been applied,

35    however, in subsurface reservoir modelling for estimating a large number of uncertain parameters. In this paper we employ

it for estimating uncertain parameters in subsurface reservoir modelling. ETPF provides the equations that are solved in the space defined by the ensemble members. Therefore for comparison we employ Ensemble Transform Kalman Filter (ETKF) developed by Bishop et al. (2001) that also transforms the state from the model space to the ensemble space, minimises the uncertainty in the ensemble space and transforms the estimation back to the model space.

5    In this paper we investigate the performance of ETPF and ETKF for parameter estimation in nonlinear problems and compare them to IS with a large ensemble. This paper is organized as follows: in the section 2 we describe IS, ETPF, and ETKF for parameter estimation. We apply these methods in Sect. 3 to a one parameter nonlinear test case, where the posterior can be computed analytically, and in Sect. 4 to a single-phase Darcy flow, where the number of parameters is 5 and 2500. In Sect. 5 we draw the conclusions.

## 2    Data assimilation methods

We implement an ensemble transform Kalman filter and an ensemble transform particle filter for estimating parameters of subsurface flow. Both of these methods are based on Bayesian framework. Assume we have an ensemble of $M$ model parameters $\{\boldsymbol{u}_m\}_{m=1}^M$, then according to this framework, the posterior distribution, which is the probability distribution $\pi(\boldsymbol{u}_m|\boldsymbol{y}_{\mathrm{obs}})$ of the model parameters $\boldsymbol{u}_m$ given a set of observations $\boldsymbol{y}_{\mathrm{obs}}$, can be estimated by the pointwise multiplication of the prior probability distribution $\pi(\boldsymbol{u}_m)$ of the model parameters $\boldsymbol{u}_m$ and the conditional probability distribution $\pi(\boldsymbol{y}_{\mathrm{obs}}|\boldsymbol{u}_m)$ of the observations given the model parameters, which is also referred as the likelihood function,

$$\pi(\boldsymbol{u}_m|\boldsymbol{y}_{\mathrm{obs}}) = \frac{\pi(\boldsymbol{y}_{\mathrm{obs}}|\boldsymbol{u}_m)\pi(\boldsymbol{u}_m)}{\pi(\boldsymbol{y}_{\mathrm{obs}})}.$$

The denominator $\pi(\boldsymbol{y}_{\mathrm{obs}})$ represents the marginal of observations and can be expressed as:

$$\pi(\boldsymbol{y}_{\mathrm{obs}}) = \sum_{m=1}^M \pi(\boldsymbol{y}_{\mathrm{obs}}, \boldsymbol{u}_m) = \sum_{m=1}^M \pi(\boldsymbol{y}_{\mathrm{obs}}|\boldsymbol{u}_m)\pi(\boldsymbol{u}_m),$$

20   which shows that $\pi(\boldsymbol{y}_{\mathrm{obs}})$ is just a normalisation factor.

### 2.1    Ensemble Transform Kalman Filter

Assume we have an ensemble of $M$ initial model parameters $\{\boldsymbol{u}_m^b\}_{m=1}^M$, where $b$ refers to a background (prior) ensemble, which are sampled from a chosen prior probability density function, then the ensemble Kalman estimate (or analysis) $\{\boldsymbol{u}_m^a\}_{m=1}^M$ is given by:

25   $$\boldsymbol{u}_m^a = \sum_{l=1}^M \mathrm{diag}\left(s_{lm} + q_l - \frac{1}{M}\right)\boldsymbol{u}_l^b, \quad m = 1,\ldots,M,$$

where diag is a diagonal matrix, $s_{lm}$ is the $(l,m)$ entry of a matrix $\mathbf{S}$

$$\mathbf{S} = \left[\mathbf{I} + \frac{1}{M-1}(\mathbf{A}^b)^T\mathbf{R}^{-1}\mathbf{A}^b\right]^{-1/2}, \tag{1}$$

and $q_l$ is the $l$-th entry of a column $\boldsymbol{q}$

$$\boldsymbol{q} = \frac{1}{M-1}\mathbf{1}_M - \mathbf{S}^2(\mathbf{A}^b)^T\mathbf{R}^{-1}(\bar{\boldsymbol{y}}^b - \boldsymbol{y}_{\text{obs}}).$$

Here $\mathbf{I}$ is an identity matrix of size $M \times M$, $\mathbf{1}_M$ is a vector of size $M$ with all ones, $\bar{\boldsymbol{y}}^b$ is the mean of the predicted data defined by

$$5 \quad \bar{\boldsymbol{y}}^b = \frac{1}{M}\sum_{m=1}^{M}\boldsymbol{y}_m^b,$$

$\mathbf{A}^b$ is the background ensemble anomalies of the predicted data defined as

$$\mathbf{A}^b = \left[ (\boldsymbol{y}_1^b - \bar{\boldsymbol{y}}^b) \quad (\boldsymbol{y}_2^b - \bar{\boldsymbol{y}}^b) \quad \ldots \quad (\boldsymbol{y}_M^b - \bar{\boldsymbol{y}}^b) \right],$$

and $\mathbf{R}$ is the measurement error covariance. To ensure that the anomalies of analysis remain zero centered we check whether $\mathbf{A}^a\mathbf{1}_M = \mathbf{A}^b\mathbf{S}\mathbf{1}_M = \mathbf{0}$, given $\mathbf{S}\mathbf{1}_M = \mathbf{1}_M$ and $\mathbf{A}^b\mathbf{1}_M = \mathbf{0}$. The model parameters $\boldsymbol{u}_m^b$ and the predicted data $\boldsymbol{y}_m^b$ are related by

10 $\quad \boldsymbol{y}_m^b = h(\boldsymbol{u}_m^b)$, where $h$ is a nonlinear function and here we assume that the function $h$ is known.

## 2.2 Ensemble Transform Particle Filter

In particle filtering we represent the probability distribution function using ensemble members (also called particles) as in ensemble Kalman filter. We start by assigning prior (background) weights $\{w_m^b\}_{m=1}^M$ to $M$ particles and then compute new (analysis) weights $\{w_m^a\}_{m=1}^M$ using the Bayes' formula and observations $\boldsymbol{y}_{\text{obs}}$

$$15 \quad w_m^a = \frac{\pi(\boldsymbol{y}_{\text{obs}}|\boldsymbol{u}_m^b)w_m^b}{\pi(\boldsymbol{y}_{\text{obs}})}. \tag{2}$$

We assume that initially all particles have equal weight, thus $w_m^b = 1/M$ for $m = 1, \ldots, M$, and that the likelihood is Gaussian with error covariance matrix $\mathbf{R}$, then from Eq. (2) $w_m^a$ is given by

$$w_m^a = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})^T\mathbf{R}^{-1}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})\right]}{\sum_{j=1}^{M}\exp\left[-\frac{1}{2}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})^T\mathbf{R}^{-1}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})\right]}, \quad m = 1, \ldots, M. \tag{3}$$

In Importance Sampling (IS), which will be used in this paper as a "ground" truth, these weights define the posterior pdf. The

20 mean parameter for IS is then

$$\bar{\boldsymbol{u}}^a = \sum_{m=1}^{M}\boldsymbol{u}_m^b w_m^a.$$

It is important to note that IS *does not* change the parameters $\boldsymbol{u}$, it only modifies the weight of the particles (samples). Therefore a resampling needs to be implemented for parameter estimation, which is usually stochastic. Instead particle filtering has been modified using a deterministic coupling methodology which resulted in an ensemble transform particle filter of Reich and

25 Cotter (2015). ETPF looks for a coupling between two discrete random variables $B_1$ and $B_2$ so as to convert the ensemble

members belonging to the random variable $B_2$ with probability distribution $\pi(B_2 = \boldsymbol{u}_m^b) = w_m^a$ to the random variable $B_1$ with uniform probability distribution $\pi(B_1 = \boldsymbol{u}_m^b) = 1/M$. The coupling between these two random variables is an $M \times M$ matrix $\mathbf{T}$ whose entries should satisfy

$$t_{mj} \geq 0, \quad m, j = 1, \ldots, M, \tag{4}$$

$$\sum_{m=1}^{M} t_{mj} = \frac{1}{M}, \quad j = 1, \ldots, M, \tag{5}$$

$$\sum_{j=1}^{M} t_{mj} = w_m^a, \quad m = 1, \ldots, M. \tag{6}$$

An optimal coupling matrix $\mathbf{T}^*$ with elements $t_{mj}^*$ minimizes the squared Euclidean distance

$$J(t_{mj}) = \sum_{m,j=1}^{M} t_{mj} \|\boldsymbol{u}_m^b - \boldsymbol{u}_j^b\|^2 \tag{7}$$

and the analysis model parameters are obtained by the linear transformation

$$\boldsymbol{u}_j^a = M \sum_{m=1}^{M} t_{mj}^* \boldsymbol{u}_m^b, \quad j = 1, \ldots, M. \tag{8}$$

Then the mean parameter for ETPF is

$$\bar{\boldsymbol{u}}^a = \sum_{m=1}^{M} \boldsymbol{u}_m^a \frac{1}{M}.$$

We use $FastEMD$ algorithm of Pele and Werman (2009) to solve the linear transport problem and get the optimal transport matrix.

**Remark:** An important property of ETPF is preservation of imposed interval bounds on ensemble members. Consider an ensemble of parameters $\{\boldsymbol{u}_m^b\}_{m=1}^{M}$ given by

$$\boldsymbol{u}_m^b = (a_m^b \; b_m^b \; c_m^b)^T, \quad m = 1, \ldots, M,$$

where we assume all the parameters $\{a_m^b\}_{m=1}^{M}$, $\{b_m^b\}_{m=1}^{M}$ and $\{c_m^b\}_{m=1}^{M}$ are bounded between 0 and 1. Therefore, the following inequalities hold:

$$0 < a_{\min} \leq a_m^b \leq a_{\max} < 1, \quad m = 1, \ldots, M,$$
$$0 < b_{\min} \leq b_m^b \leq b_{\max} < 1, \quad m = 1, \ldots, M,$$
$$0 < c_{\min} \leq c_m^b \leq c_{\max} < 1, \quad m = 1, \ldots, M.$$

Now we assume two discrete random variables $B_1$ and $B_2$ have probability distributions given by

$$\pi(B_1 = \boldsymbol{u}_m^b) = 1/M, \quad \pi(B_2 = \boldsymbol{u}_m^b) = w_m^a,$$

with $w_m^a \geq 0$, $m = 1, \ldots, M$ and $\sum_{m=1}^{M} w_m^a = 1$. As ETPF looks for a matrix $\mathbf{T}^*$ which defines coupling between these two probability distributions, each entry of this coupling matrix satisfies the conditions given by Eq. (4)–(6). These conditions assure that each entry of the coupling matrix will be non-negative and less than 1. Since the analysis given by Eq. (8) is

$$
\boldsymbol{u}_m^a = \begin{bmatrix} a_1^b(Mt_{1m}^*) + a_2^b(Mt_{2m}^*) + \cdots + a_M^b(Mt_{Mm}^*) \\ b_1^b(Mt_{1m}^*) + b_2^b(Mt_{2m}^*) + \cdots + b_M^b(Mt_{Mm}^*) \\ c_1^b(Mt_{1m}^*) + c_2^b(Mt_{2m}^*) + \cdots + c_M^b(Mt_{Mm}^*) \end{bmatrix}, \quad m = 1, \ldots, M,
$$

5    these conditions lead to

$$
0 < a_{\min} \leq a_m^a \leq a_{\max} < 1, \quad m = 1, \ldots, M,
$$

$$
0 < b_{\min} \leq b_m^a \leq b_{\max} < 1, \quad m = 1, \ldots, M,
$$

$$
0 < c_{\min} \leq c_m^a \leq c_{\max} < 1, \quad m = 1, \ldots, M.
$$

Thus the coupling matrix bounds the analysis ensemble members to be in the desired range. This is not observed in ETKF as
10    the matrix $\mathbf{S}$ given by Eq. (1) does not impose any of the non-equality and equality constraints, so it results in values outside the bound.

### 2.3 Localization

All variations of ensemble Kalman filter and particle filter are limited by the ensemble size. Since, even if the dimension of the problem is just up to a few thousands, a large ensemble size will make each run of the model computationally very expensive.
15    This limit of a small ensemble size introduces sampling errors. To deal with this issue localized ETKF (LETKF) was introduced by Hunt et al. (2007) and for localized ETPF (LETPF) by Reich and Cotter (2015). More recent approaches to particle filter localization include Penny and Miyoshi (2016) and Poterjoy (2016)

For the local update of a model parameter $\boldsymbol{u}_m(X_i)$ at a grid point $X_i$, we introduce a diagonal matrix $\hat{\mathbf{C}}_i \in R^{N_y \times N_y}$ in the observation space with an element

20    $$(\hat{\mathbf{C}}_i)_{ll} = \rho\left(\frac{||X_i - r_l||}{r_{\text{loc}}}\right), \tag{9}$$

where $i = 1, \ldots, n^2$, $l = 1, \ldots, N_y$, $n^2$ is the number of model parameters, $N_y$ is the dimension of the observation space, $r_l$ denotes the location of the observation, $r_{\text{loc}}$ is a localisation radius and $\rho(\cdot)$ is a taper function, such as Gaspari-Cohn function Gaspari and Cohn (1999)

$$
\rho(r) = \begin{cases} 1 - \frac{5}{3}r^2 + \frac{5}{8}r^3 + \frac{1}{2}r^4 - \frac{1}{4}r^5, & 0 \leq r \leq 1, \\ -\frac{2}{3}r^{-1} + 4 - 5r + \frac{5}{3}r^2 + \frac{5}{8}r^3 - \frac{1}{2}r^4 + \frac{1}{12}r^5, & 1 \leq r \leq 2, \\ 0, & 2 \leq r. \end{cases}
$$

25    Then the estimated model parameter at the location $X_i$ is

$$
\boldsymbol{u}_m^a(X_i) = \sum_{l=1}^{M} \text{diag}\left(s_{lm}(X_i) + q_l(X_i) - \frac{1}{M}\right)\boldsymbol{u}_l^b(X_i), \quad m = 1, \ldots, M,
$$

where diag is a diagonal matrix, $s_{lm}(X_i)$ is the $(l,m)$ entry of the localized transformation matrix $\mathbf{S}(X_i)$

$$\mathbf{S}(X_i) = \left[\mathbf{I} + \frac{1}{M-1}(\mathbf{A}^b)^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})\mathbf{A}^b\right]^{-1/2}$$

and $q_l(X_i)$ is the $l$-th entry of the localized column $\boldsymbol{q}(X_i)$

$$\boldsymbol{q}(X_i) = \frac{1}{M-1}\mathbf{1}_M - \mathbf{S}(X_i)^2(\mathbf{A}^b)^T\mathbf{R}^{-1}(\bar{\boldsymbol{y}}^b - \boldsymbol{y}_{\text{obs}}).$$

5     LETPF modifies the likelihood and thus the weights given by Eq. (3) are computed locally at each grid $X_i$

$$w_m^a(X_i) = \frac{\exp\left[-\frac{1}{2}(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})(\boldsymbol{y}_m^b - \boldsymbol{y}_{\text{obs}})\right]}{\sum_{j=1}^M \exp\left[-\frac{1}{2}(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})^T(\hat{\mathbf{C}}_i\mathbf{R}^{-1})(\boldsymbol{y}_j^b - \boldsymbol{y}_{\text{obs}})\right]}, \quad m = 1,\dots,M, \tag{10}$$

where $\hat{\mathbf{C}}_i$ is the diagonal matrix given by Eq. (9). Then the estimated model parameter $\boldsymbol{u}_j^a(X_i)$ at the grid $X_i$ is given by

$$\boldsymbol{u}_j^a(X_i) = M\sum_{m=1}^M t_{mj}^* \boldsymbol{u}(X_i)_m^b, \quad j = 1,\dots,M,$$

where $t_{mj}^*$ is an element of an optimal coupling matrix $\mathbf{T}^*$ which minimizes the squared Euclidean distance at the grid point
10   $X_i$

$$J(t_{mj}) = \sum_{m,j=1}^M t_{mj}[u_m^b(X_i) - u_j^b(X_i)]^2, \tag{11}$$

which reduces the LETPF to a univariate transport problem. It should be noted that localization can be applied only for grid-dependent parameters.

## 3   One parameter nonlinear problem

15   First we consider a one parameter nonlinear problem from Chen and Oliver (2013). The prior distribution is Gaussian distribution with mean 4 and variance 1. The nonlinear forward model is

$$h(u) = \frac{7}{12}u^3 - \frac{7}{2}u^2 + 8u.$$

The true parameter $u^{\text{true}}$ gives $h(u^{\text{true}}) = 48$ and the observation error is drawn from a Gaussian distribution with zero mean and variance 16. In Fig. 1 we plot the posterior probability density functions estimated by ETPF (top), ETKF (bottom) with
20   ensemble sizes $10^2$ (left), $10^3$ (center), and $10^4$ (right). The prior distribution is shown in red and the posterior estimated by IS with ensemble size $10^5$ is shown in black. We can see that ETPF provides better approximation of the true probability density function, while ETKF gives a skewed posterior. It should be noted that ETKF is able to give a non-Gaussian (though wrong) posterior due to the nonlinearity of the map between the uncertain parameters and observations.
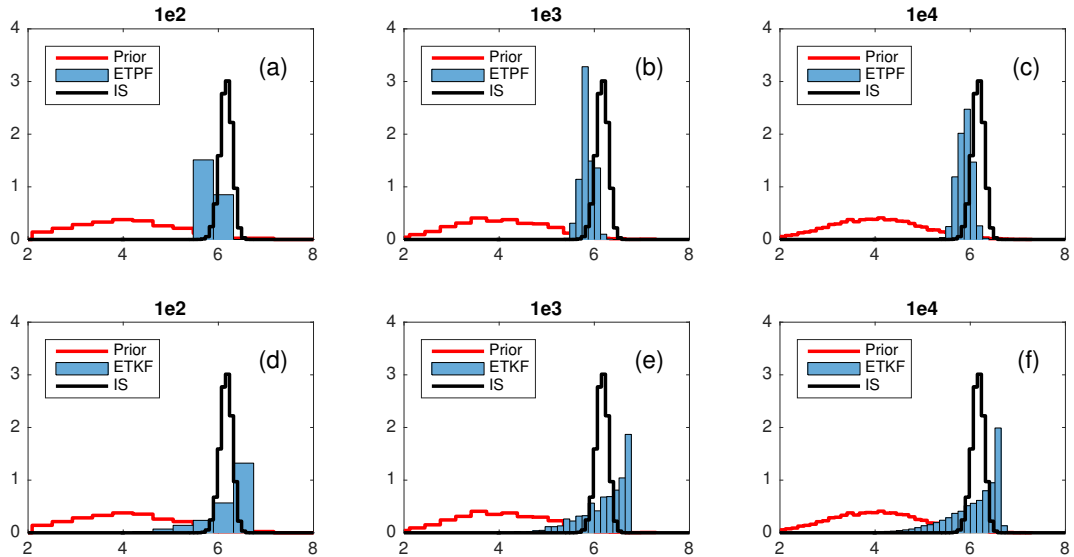
**Figure 1.** Probability density functions for the one parameter nonlinear problem. Top: ETPF, bottom: ETKF. Left: ensemble size $10^2$, center: ensemble size $10^3$, right: ensemble size $10^4$. Prior is in red. True pdf obtained by IS with ensemble size $10^5$ is in black.

## 4 Single-phase Darcy flow

We consider a steady-state single-phase Darcy flow model defined over an aquifer of two-dimensional physical domain $D = [0,1] \times [0,1]$, which is given by,

$$-\nabla \cdot (k(x,y)\nabla P(x,y)) = f(x,y), \quad (x,y) \in D$$

$$P(x,y) = 0, \quad (x,y) \in \partial D$$

where $\nabla = (\partial/\partial x \; \partial/\partial y)^T$, $\cdot$ denotes the dot product, $P(x,y)$ the pressure, $k(x,y)$ the permeability, $f(x,y)$ the source term, which we assume to be $2\pi^2 \cos(\pi x)\cos(\pi y)$, and $\partial D$ the boundary of domain $D$. The forward problem of this second order elliptical equation is to find the solution of pressure $P(x,y)$ for given $f(x,y)$ and $k(x,y)$. We, however, are interested in finding permeability given noisy observations of pressure at a few locations.

We perform numerical experiments with synthetic observations, where instead of a measuring device a model is used to obtain observations. We implement a cell-centered finite difference method to discretize the domain $D$ into $n \times n$ grid cells $X_i$ of size $\Delta x^2$ and solve the forward model with the true parameters. Then the synthetic observations are obtained by

$$\boldsymbol{y}_{\text{obs}} = \mathbf{L}(\mathbf{P}) + \eta,$$

with an element of $\mathbf{L}(\mathbf{P})$ being a linear functional of pressure, namely

$$L_l(\mathbf{P}) = \frac{1}{2\pi\sigma^2} \sum_{i=1}^{n^2} \exp\left(-\frac{||X_i - r_l||^2}{2\sigma^2}\right) P_i \Delta x^2, \quad l \in 1, \ldots, N_y$$

where $n = 50$, $\sigma = 0.01$, $r_l$ denotes the location of the observation and $N_y = 16$, which is the number of observations. The observation locations are spread uniformly across the domain $D$ and $\eta$ denotes the observation noise drawn from a normal distribution with zero mean and standard deviation of $0.09$. This form of the observation functional and parametrization of the uncertain parameters given below guaranty the continuity of the forward map from the uncertain parameters to the observations and thus the existence of the posterior distribution as shown by Iglesias et al. (2014).

## 4.1 Five parameter nonlinear problem

For our first numerical experiment with Darcy flow, we consider a low-dimensional problem where the permeability field is defined by mere 5 parameters similarly to Iglesias et al. (2014). We assume that the entire domain $D = [0,1] \times [0,1]$ is divided into two subdomains $D_1$ and $D_2$ as shown in Fig. 2. Each subdomain of $D$ represents a layer and is assumed to have a permeability function $k(\mathbf{X})$, where an element of $\mathbf{X}$ is defined by $X_i$ for $i = 1, \ldots, n^2$. The thickness of a layer on both sides $a$ and $b$, correspondingly, defines the slope of the interface and a parameter $c$ defines a vertical fault. The layer moves up or down depending on $c < 0$ or $c > 0$, respectively, and its location is assumed to be fixed at $x = 0.5$.

Further, for this test case we assume piecewise constant permeability within each of the subdomains, hence $k(\mathbf{X})$ is given by

$$k(\mathbf{X}) = k_1 \delta_{D_1}(\mathbf{X}) + k_2 \delta_{D_2}(\mathbf{X}),$$

where $k_1$ and $k_2$ represent permeability of the subdomain $D_1$ and $D_2$, respectively, and $\delta$ is Dirac function. Then the parameters defining the permeability field for this configuration are

$$\boldsymbol{u} = (a \; b \; c \; \log(k_1) \; \log(k_2))^T.$$

We assume that the true parameters are $a^{\text{true}} = 0.6$, $b^{\text{true}} = 0.3$, $c^{\text{true}} = -0.15$, $k_1^{\text{true}} = 12$ and $k_2^{\text{true}} = 5$. These parameters are used to create synthetic observations. Figure 2 shows the true permeability with dots representing the observation locations. Next, we assume that the five uncertain parameters are drawn from a uniform distribution over a specified interval, namely $a, b \sim \mathcal{U}[0,1]$, $c \sim \mathcal{U}[-0.5, 0.5]$, $k_1 \sim \mathcal{U}[10, 15]$ and $k_2 \sim \mathcal{U}[4, 7]$.

As it was pointed out in Sect. 2.2, ETPF updates the parameters within the original range of an initial ensemble, while ETKF does not. Therefore a change of variables has to be performed for ETKF so that the updated parameters are physically viable. In order to be consistent we perform the change of variables for ETPF as well. As the domain $D$ is $[0,1] \times [0,1]$, the parameters $a$ and $b$ should lie within the interval $[0,1]$. To enforce this constraint we substitute $a$ according to

$$a' = \log\left(\frac{a}{1-a}\right), \quad a' \in R$$

and similarly $b$ is substituted by $b'$. Thus the uncertain parameters are now $\boldsymbol{u}' = (a' \; b' \; c \; \log(k_1) \; \log(k_2))^T$.
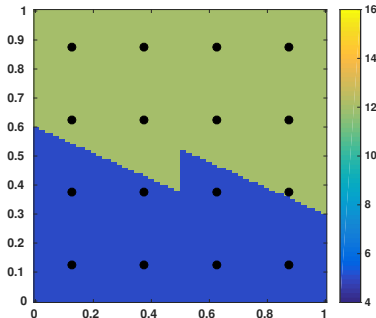
**Figure 2.** True permeability of the 5 parameter nonlinear problem with dots representing the observation locations.

In Fig. 3 we plot probability density functions for parameters $a$ (a)–(d), $c$ (e)–(h) and $\log(k_2)$ (i)–(l), as the parameters $b$ and $\log(k_1)$ show similar results. The posterior obtained by IS with ensemble size $10^6$ is plotted as a black line and the true value of parameters is plotted as a black line with crosses. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. ETPF and ETKF used $10^3$ (odd columns) and $10^4$ (even columns) ensemble members. In order to perform an objective comparison between the probabilities we compute the Kullback-Leibler divergence of a posterior $\pi$ obtained by either ETPF or ETKF and the posterior $\pi^{\text{IS}}$ obtained by IS

$$\mathrm{D}_{\text{KL}}(\pi^{\text{IS}} \parallel \pi) = \sum_{i=1}^{N_{\text{b}}} \pi^{\text{IS}}(u_i) \log \frac{\pi^{\text{IS}}(u_i)}{\pi(u_i)} (u_i - u_{i-1}), \tag{12}$$

where $N_{\text{b}} = 20$ is the number of bins. The Kullback-Leibler divergence for parameters $a$, $c$ and $\log(k_2)$ is displayed in the titles of Fig. 3, where we observe that ETKF outperforms ETPF.

In order to check the sensitivity of the results to the initial parameter ensemble we perform 10 simulations based on a random draw of an initial ensemble from the same prior distributions. We conduct the numerical experiments for ensemble sizes varying from 10 to $10^3$ with an increment of 50. In Fig. 4 we plot the true parameters, the mean estimated by IS, the mean $\bar{\bar{u}}^a$ and the spread $\bar{\bar{u}}^a \pm \bar{u}^a_{\text{std}}$ of estimated parameters averaged over 10 simulations

$$\bar{\bar{u}}^a_i = \frac{1}{10} \sum_{r=1}^{10} \bar{u}^{a,r}_i, \quad \bar{u}^a_{\text{std}} = \frac{1}{10} \sum_{r=1}^{10} \sqrt{\frac{1}{M-1} \sum_{m=1}^{M} (u^{a,r}_{i,m} - \bar{u}^{a,r}_i)^2}, \text{ where } \bar{u}^{a,r}_i = \frac{1}{M} \sum_{m=1}^{M} u^{a,r}_{i,m}, \ r = 1, \ldots, 10,$$

$M$ is ensemble size, $i = 1, \ldots, 5$ is parameter index, and the superscript $a$ is for the analysis. We observe that all methods including IS have bias in the estimations of geometrical parameters, which is due to the small number of observations. ETPF and ETKF perform comparably in terms of mean estimation, though some are better estimated by ETKF and other are better estimated by ETPF. Comparing the error in pressure of the mean parameters we observe that methods are equivalent (not shown), which is a manifestation of the ill-posedness of the problem. In Fig. 4 we see that the spread from ETPF is smaller than from ETKF for each parameter. Even though both methods are slightly underdispersive, as the spread to error ratio is
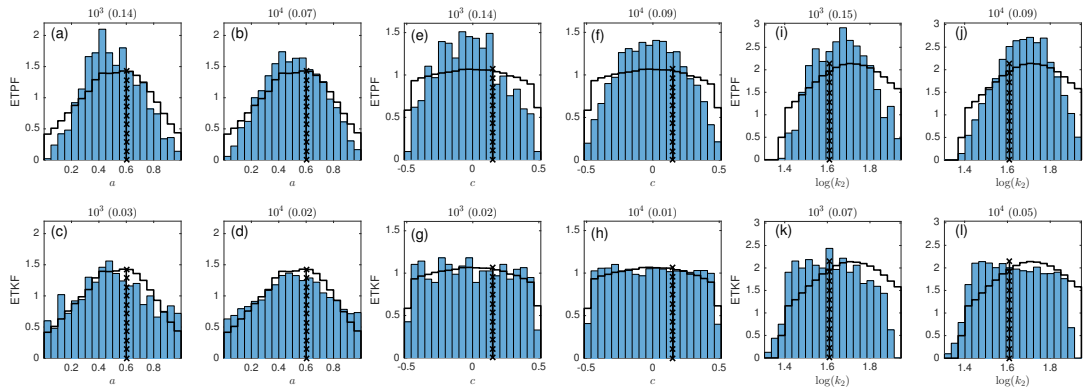
**Figure 3.** Probability density functions for the parameters $a$ (a)–(d), $c$ (e)–(h), and $\log(k_2)$ (i)–(l). The posterior obtained by IS with ensemble size $10^6$ is plotted as a black line and the true values of parameters are plotted as black crosses. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. ETPF and ETKF used $10^3$ (odd columns) and $10^4$ (even columns) ensemble members.

below 1, ETKF gives better ratio for all parameters but $\log(k_2)$, namely for the ensemble size $10^3$ and each parameter it is () for ETKF and () for ETPF.

We compute an average of the relative error over all parameters

$$\text{RE}^{a,r} = \frac{1}{5}\sum_{i=1}^{5}\frac{|\bar{u}_i^{a,r} - u_i^{\text{true}}|}{|u_i^{\text{true}}|},\ r = 1,\ldots,10,$$

5    and the data misfit

$$\text{misfit}^{a,r} = (\bar{\boldsymbol{y}}^{a,r} - \boldsymbol{y}_{\text{obs}})^T R^{-1}(\bar{\boldsymbol{y}}^{a,r} - \boldsymbol{y}_{\text{obs}}),\ r = 1,\ldots,10 \tag{13}$$

after data assimilation. The same metrics are computed before data assimilation and denoted by a superscript $b$. In Fig. 5(a)–(b) we plot $(\text{misfit}^{a,r} - \text{misfit}^{b,r})$ and $(\text{RE}^{a,r} - \text{RE}^{b,r})$, respectively, for each simulation $r$ as a function of ensemble size. ETPF is shown in blue and ETKF in red. Black line is at zero level. Positive values of the differences mean an increase of either data

10    mismatch or relative error after data assimilation. We observe a data misfit decrease for both ETPF and ETKF except at an ensemble size 10. RE does not always decrease for ETPF: for some simulations ETPF is at zero level or slightly above it, while for ETKF the sole exception is at an ensemble size 10.

## 4.2   High-dimensional nonlinear problem

Next, we consider a high-dimensional problem where the dimension of the uncertain parameter is $n^2 = 2500$. The domain $D$ is

15    now not divided into subdomains. However, unlike in the previous test case here we implement a spatially varying permeability field. We assume the log permeability is generated by a random draw from a Gaussian distribution $\mathcal{N}(\log(\mathbf{5}), \mathbf{C})$. Here $\mathbf{5}$ is an $n^2$ vector with all 5. $\mathbf{C}$ is assumed to be an exponential correlation with an element of $\mathbf{C}$ being

$$C_{i,j} = \exp(-3(|h_{i,j}|/v)),\ i,j = 1,\ldots,n^2.$$
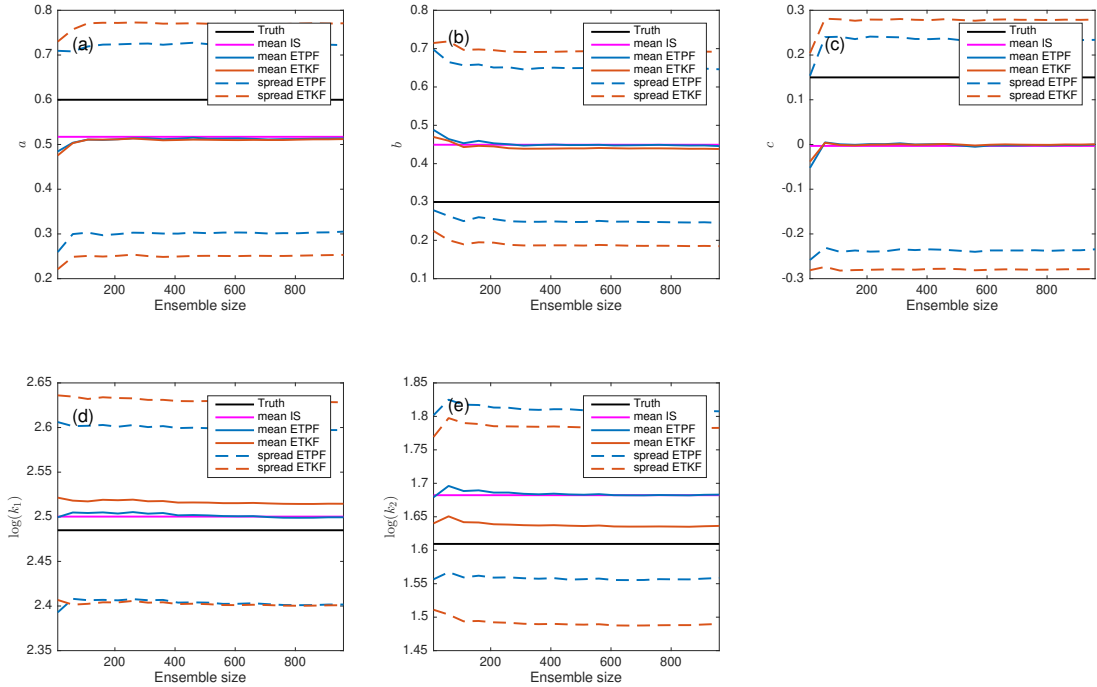
**Figure 4.** $\bar{\bar{u}}^a$ and $\bar{\bar{u}}^a \pm \bar{u}^a_{\text{std}}$ w.r.t ensemble size: (a) for the parameter $a$, (b) for $b$, (c) for $c$, (d) for $\log(k_1)$, (e) for $\log(k_2)$. ETPF is shown in blue, ETKF in red, the true parameters are in black and the mean of IS in magenta.

Here $h_{i,j}$ is the distance between two spatial locations and $v$ is the correlation range which is taken to be $0.5$. For the log permeability we use Karhunen-Loeve expansions of the form

$$\log(k_j) = \log(5) + \sum_{i=1}^{n^2} \sqrt{\lambda_i} \nu_{i,j} \mathcal{Z}_i, \quad \text{for} \quad j = 1, \ldots, n^2 \tag{14}$$

where $\lambda$ and $\nu$ are eigenvalues and eigenfunctions of $\mathbf{C}$, respectively, and the vector $\mathcal{Z}$ is of dimension $n^2$ iid from a Gaussian distribution with zero mean and variance one. Making sure that the eigenvalues are sorted in descending order $\mathcal{Z}_i \sim \mathcal{N}(0,1)$ produces $\log(\mathbf{k}) \sim \mathcal{N}(\log(\mathbf{5}), \mathbf{C})$. The uncertain parameter is thus $\boldsymbol{u} = \mathcal{Z}$ with the dimension $n^2 = 2500$.

We perform 10 different simulations based on a random draw of an initial ensemble from the prior distribution. We conduct the numerical experiments for ensemble sizes varying from 10 to $10^3$ with an increment of 50. We compute the root mean square error (RMSE) of the log permeability field

$$\text{RMSE}^{r,a} = \sqrt{\left(\log(\bar{\mathbf{k}}^{a,r}) - \log(\mathbf{k}^{\text{true}})\right)^T \left(\log(\bar{\mathbf{k}}^{a,r}) - \log(\mathbf{k}^{\text{true}})\right)}, \quad r = 1, \ldots, 10,$$
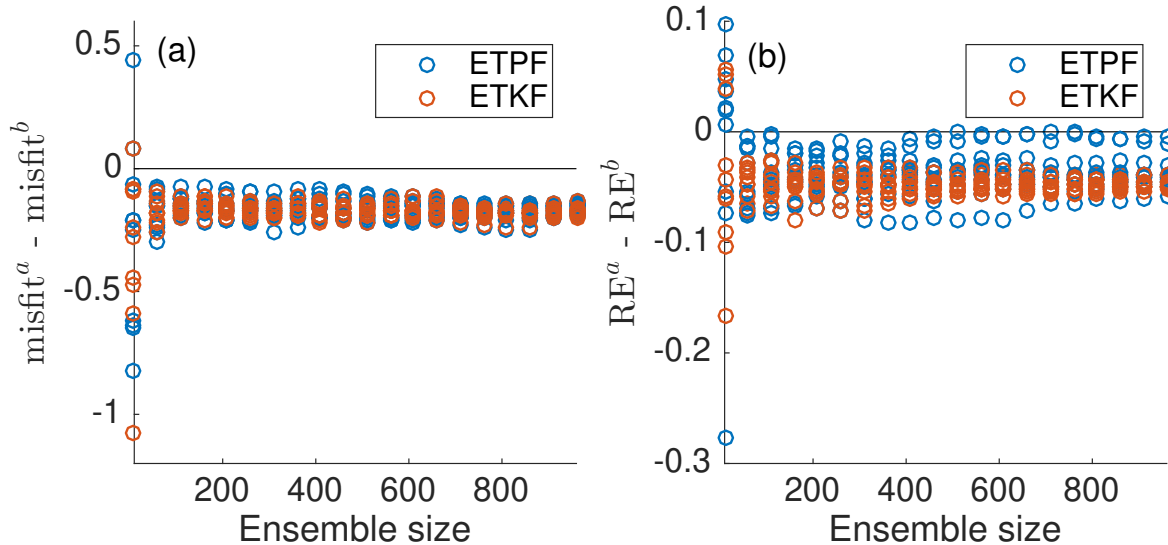
**Figure 5.** misfit$^{a,r}$ − misfit$^{b,r}$ (a) and RE$^{a,r}$ − RE$^{b,r}$ (b) w.r.t ensemble size. ETPF is shown in blue, ETKF in red and the zero level in black. A circle is for one simulation.

and variance

$$\text{variance}^{r,a} = \frac{1}{M-1} \sum_{m=1}^{M} \left( (\log(\mathbf{k}_m^{a,r}) - \log(\overline{\mathbf{k}}^{a,r}) \right)^T \left( \log(\mathbf{k}_m^{a,r}) - \log(\overline{\mathbf{k}}^{a,r}) \right), \quad r = 1, \ldots, 10.$$

We also compute the data misfit for each simulation after data assimilation by Eq. (13). In Fig. 6 we plot mean, minimum and maximum over 10 simulations after data assimilation for the data misfit (left), RMSE (center), and variance (right). ETPF is shown in blue and ETKF in red. We observe that ETPF is underdispersive compared to ETKF as particle filters are highly degenerative compared to Kalman filters. Misfit given by ETPF is smaller than the one given by ETKF for almost all simulations at ensemble sizes greater than 150. The RMSE on the contrary is larger. In Fig. 7(a)–(b) we plot (misfit$^{a,r}$ − misfit$^{b,r}$) and (RMSE$^{a,r}$ − RMSE$^{b,r}$), respectively, as a function of ensemble size for a simulation $r = 1, \ldots, 10$. The superscript $b$ is for
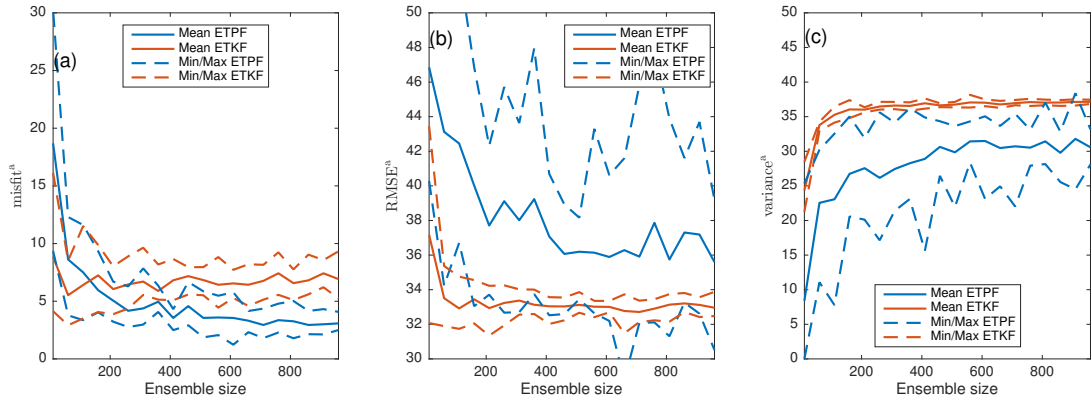
**Figure 6.** Mean, minimum and maximum over 10 simulations after data assimilation for the data misfit (a), RMSE (b), and variance (c). ETPF is shown in blue and ETKF in red.
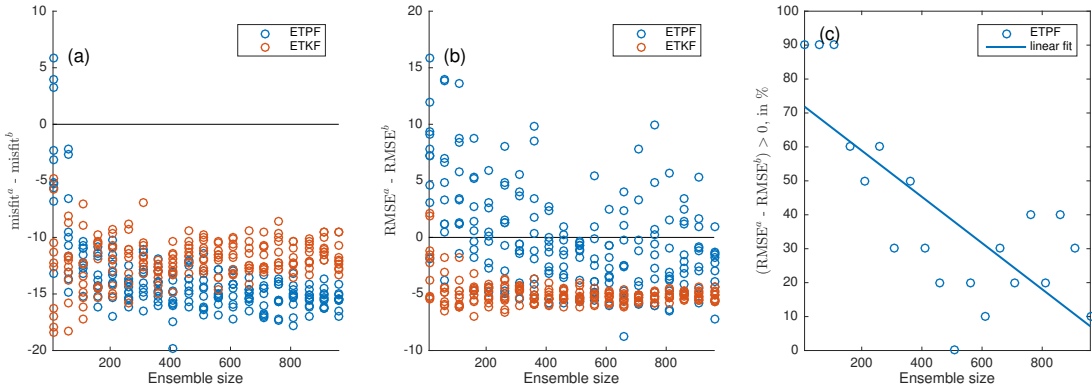
**Figure 7.** misfit$^{a,r}$ − misfit$^{b,r}$ (a) and RMSE$^{a,r}$ − RMSE$^{b,r}$ (b) w.r.t ensemble size. ETPF is shown in blue, ETKF in red and zero level in black. One circle is for one simulation. For ETPF % of simulations that result in (RMSE$^a$ − RMSE$^b$) > 0 and a linear fit as a function of ensemble size are shown in (c).

the metrics before data assimilation and the superscript $a$ is for the metrics after data assimilation. ETKF always provides a decrease in both the data misfit and RMSE except at ensemble size 10. ETPF gives a decrease in the data misfit though an increase in RMSE, which indicates that ETPF overfittes the data. However, as the ensemble size increases this happens less often as can be seen in Fig. 7(c), where we plot for ETPF a percentage of simulations that result in (RMSE$^a$ − RMSE$^b$) > 0 and a linear fit as a function of ensemble size.

In Fig. 8 we plot log permeability fields. In Fig. 8(a) the true permeability is shown with dots representing the observation locations, and in Fig. 8(d) the mean permeability field obtained by IS with ensemble size $10^5$. The RMSE provided by IS is 32.62. In Fig. 8(b–e) and Fig. 8(c–f) we display mean permeability fields obtained with ensemble size $10^3$ by ETPF and ETKF,

respectively. In Fig. 8(b–c) we plot the mean log permeabilities for the smallest RMSE over simulations, which is 30.51 for ETPF and 32.48 for ETKF. In Fig. 8(d–e) we plot the mean log permeabilities for the largest RMSE over simulations, which is 39.2 for ETPF and 33.87 for ETKF. We observe that ETKF as well as IS provide smooth mean permeability fields that have smaller absolute values than the true permeability. ETPF gives higher variations of the mean permeability field and is

5   in an excellent agreement with the true permeability for a good initial ensemble shown in Fig. 8(b). This means that ETPF is sensitive to the initial sample due to the sampling error and that the spatial variability of ETPF is a result of the sampling error. In Fig. 9 we plot the variance of the permeability fields obtained with ensemble size $10^5$ by IS (d), with ensemble size $10^3$ by ETPF (b–e) and ETKF (c–f) . Fig. 9(b–c) is for the smallest RMSE and Fig. 9(e–f) is for the largest RMSE. ETKF provides smoother variance than ETPF due to smaller sampling errors.

10      In Fig. 10 we show squared error $(\overline{\mathcal{Z}}^a - \mathcal{Z}^{\text{true}})^2$ in blue for ETPF and in red for ETKF for three leading modes $\mathcal{Z}_1$ (a), $\mathcal{Z}_2$ (b), and $\mathcal{Z}_3$ (c), where solid line is for median and shaded area is for 25 and 75 percentile over 10 simulations. We observe that in terms of the estimation of the three leading modes ETPF outperforms ETKF. In Fig. 11 we plot the posterior of $\mathcal{Z}_1$ (left), $\mathcal{Z}_2$ (center), and $\mathcal{Z}_3$ (right) obtained by IS with ensemble size $10^6$ and by ETPF (top) and ETKF (bottom) with ensemble size $10^4$. The posterior of these modes is roughly approximated by ETPF as shown in Fig. 11 (a)–(c). ETKF provides a skewed posterior

15   of the modes shown in Fig. 11 (d)–(f), which was also observed in the one parameter nonlinear problem, see Fig. 1(f). In order to perform an objective comparison between the probabilities we compute the Kullback-Leibler divergence of a posterior $\pi$ obtained by either ETPF or ETKF and the posterior $\pi^{\text{IS}}$ obtained by IS according to Eq. (12) ETPF gives the Kullback-Leibler divergence 0.21, 0.42, and 0.6, while ETKF 0.16, 0.07, and 0.5 for the modes $\mathcal{Z}_1$, $\mathcal{Z}_2$, and $\mathcal{Z}_3$, respectively. Thus ETKF gives a better approximation of the true pdf. Since first modes are well estimated by ETPF and last modes not (not shown), we use

20   only three leading modes in the Karhunen-Loeve expansion given by Eq. (14) when computing the estimated log permeability keeping the number of uncertain parameters the same, namely 2500. In Fig. 12(a) we observe that ETPF outperforms ETKF for large ensemble sizes independent of an initial sample. Moreover, ETPF is not overfitting the data anymore since RMSE always decreases after data assimilation except at small ensemble sizes shown in Fig. 12(b). In Fig. 13 we show the mean fields for the best and worst initial samples of $10^4$ size. ETPF gives RMSE at the best sample 31.1 and the worst sample 32.98.

25   By comparing it to 30.51 and 39.2 obtained using the full Karhunen-Loeve expansions, we observe that the maximum RMSE over simulations decreased substantially, while the minimum RMSE only slightly increased. ETKF gives RMSE at the best sample 32.27 and the worst sample 33.23. (Compare to 32.48 and 33.9 using the full Karhunen-Loeve expansions). Thus ETKF slightly decreases both maximum and minimum RMSE over simulations. ETPF is more affected by sampling noise at small scales, so using a truncated representation of the fields significantly improves the results for ETPS. ETKF is filtering out the

30   small scales that are not observed and thus is less affected by the truncation.

     Next we apply LETPF and LETKF. The optimal localization radius between 0.2 and 1.2 was obtained in terms of the smallest RMSE and shown in Table 1. It should be noted that smaller localization radius for LETPF than for LETKF was also observed by Cheng and Reich (2015) and it is probably related to more noisy approximation of the posterior by LETPF than by LETKF. In Fig. 14 we plot misfit, RMSE and variance. At small ensemble sizes both LETKF and LETPF give smaller misfit, smaller

35   RMSE but larger variance than ETKF and ETPF. For large ensembles LETKF performs worse than ETKF, which is due to
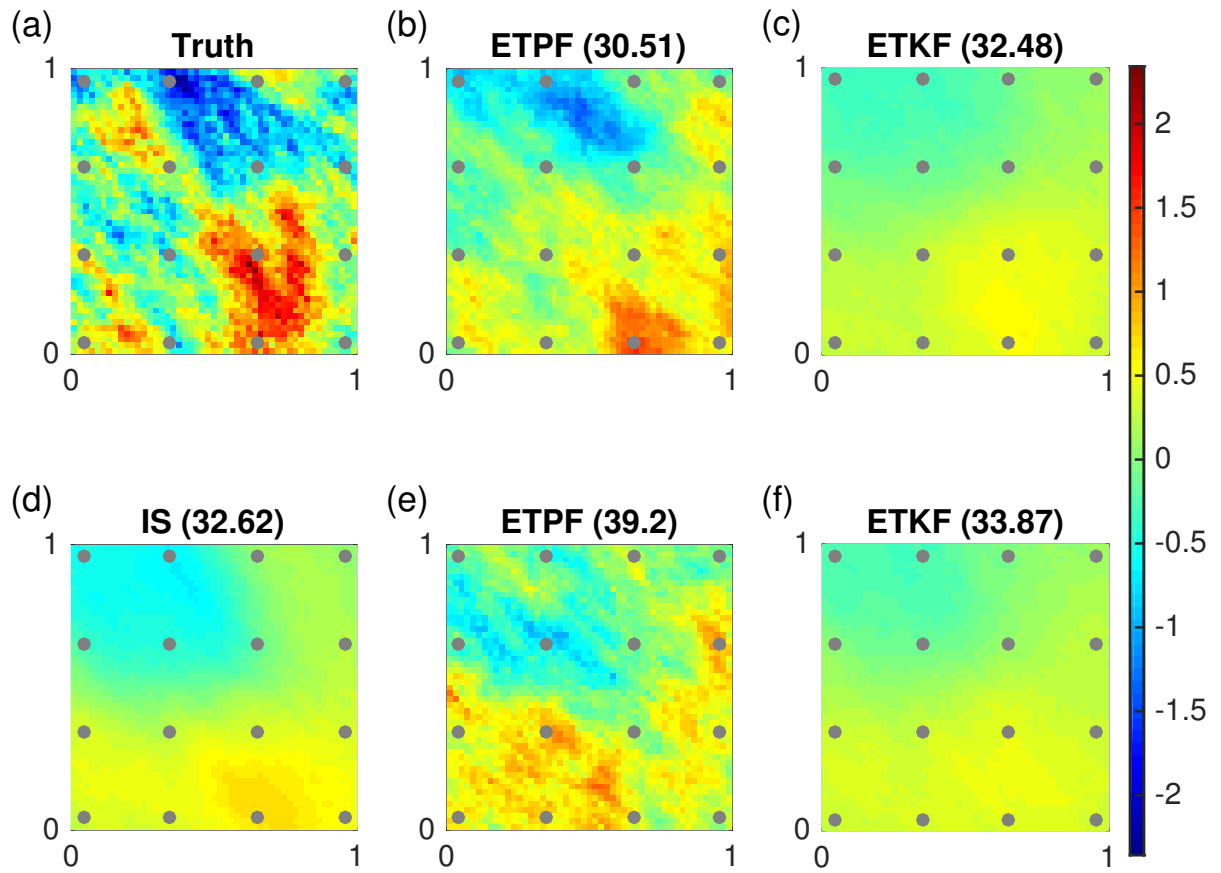
**Figure 8.** Log permeability field with dots representing the observation locations. Truth is shown in (a) and mean obtained by IS with ensemble size $10^5$ in (d). Mean obtained with ensemble size $10^3$ by ETPF shown in (b–e) and by ETKF in (c–f), where (b–c) are at the smallest RMSE and (e–f) are at the largest RMSE over simulations. The corresponding RMSE is given in brackets.

**Table 1.** Optimal localization radius for LETPF and LETKF at different ensemble sizes M.

| M | 10 | 110 | 210 | … | 910 |
|---|---|---|---|---|---|
| LETPF | 0.2 | 0.6 | 0.6 | … | 0.6 |
| LETKF | 0.2 | 1.2 | 1.2 | … | 1.2 |

the imposed range on localization radius, meaning that 1.2 is not optimal. Comparing the performance of LETPF to (L)ETKF we observe that at small ensemble sizes LETKF still outperforms ETPF but at large ensemble sizes LETPF performs now
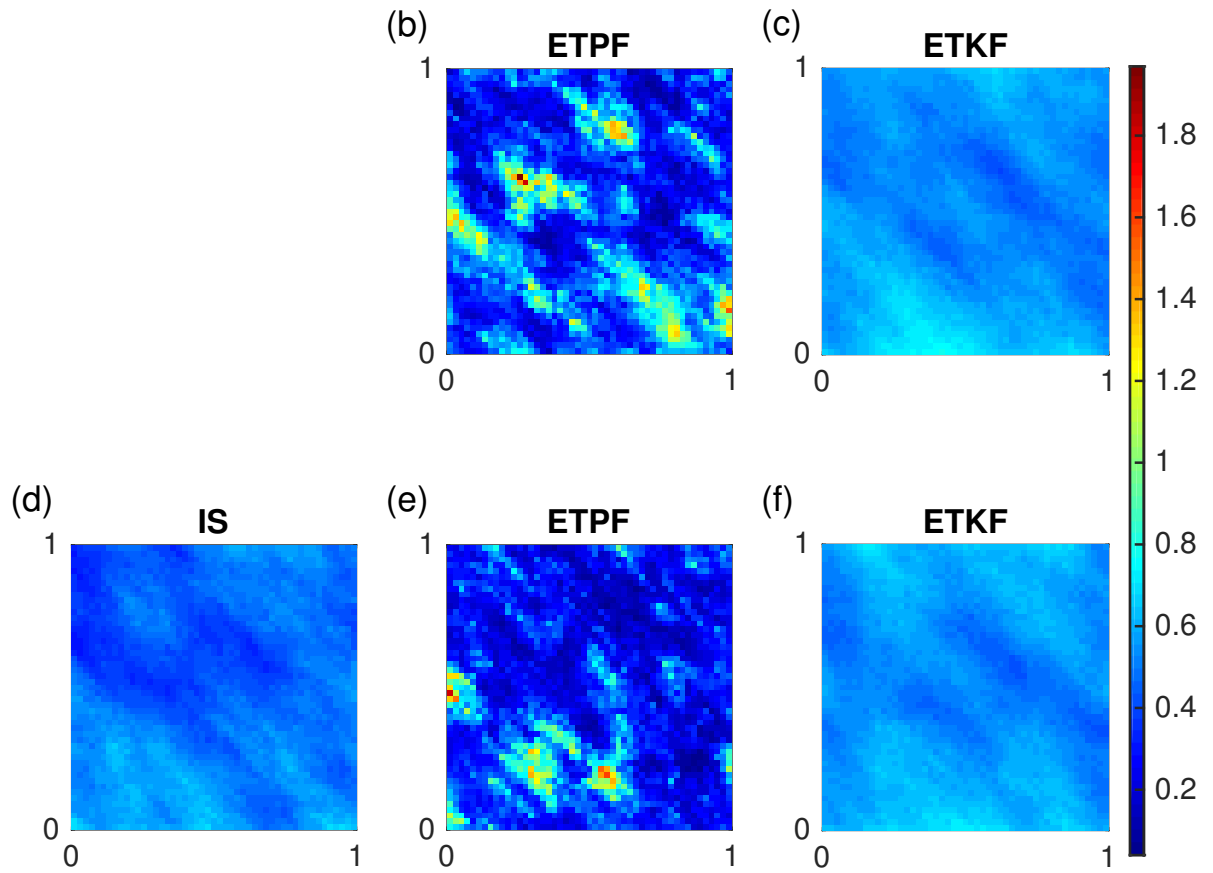
**Figure 9.** Variance of log permeability fields: obtained with ensemble size $10^5$ by IS (d), with ensemble size $10^3$ by ETPF (b–e), and ETKF (c–f). Variance at the smallest RMSE (b–c) and at the largest RMSE (e–f) over simulations.

comparably to ETKF. Moreover, LETPF overfits the data less often than ETPF: $40\%$ against $90\%$ for ensemble size 10 and $0\%$ against non-zero$\%$ for ensemble sizes greater than 150 (not shown).

In Fig. 15–16 we plot mean and variance of the log permeability field at ensemble size $10^3$ for ETPF (b)–(e) and ETKF (c)–(f) with localization at the smallest RMSE (b)–(c) and largest RMSE (e)–(f) over simulations, which are 32.29 and 34.08 for ETPF and 32.92 and 34.09 for ETKF, respectively. We observe that localization decreases the sampling noise and the spatial variability of the mean field obtained by ETPF at ensemble size $10^3$ resembles IS at ensemble size $10^5$. The variance obtained by ETPF with localization shown in Fig. 16(b–e) has also improved.
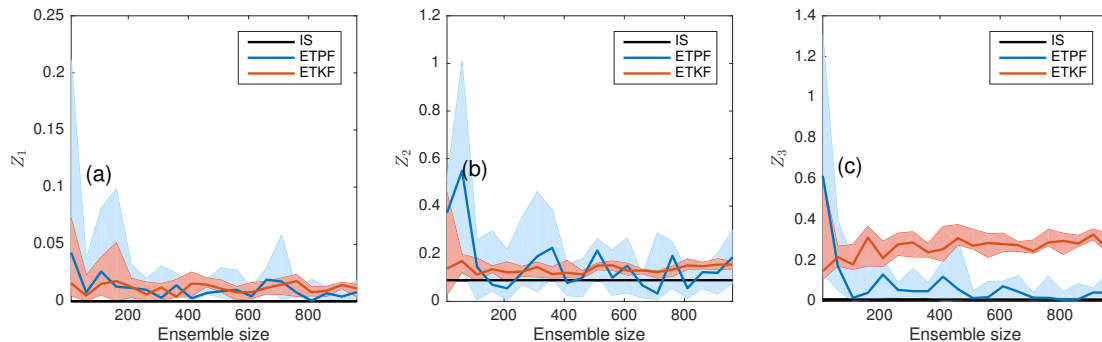
**Figure 10.** Squared error between the true and the mean estimated modes for $\mathcal{Z}_1$ (a), $\mathcal{Z}_2$ (b), and $\mathcal{Z}_3$ (c) w.r.t ensemble size. ETPF is shown in blue and ETKF in red with solid lines for median and shaded area for 25 and 75 percentile over 10 simulations. IS with ensemble size $10^5$ is in black.

The posterior estimation of the leading mode $\mathcal{Z}_1$, however, degraded, while of $\mathcal{Z}_2$ and $\mathcal{Z}_3$ improved. The Kullback-Leibler divergence for the leading mode is 0.73 (compare to 0.21 without localization), and for second and third is 0.2 and 0.18, correspondingly (compare to 0.42 and 0.6 without localization). Variance of the posteriors is larger when localization is applied for both methods. The localized weights given by Eq. (10) vary less than the non-localized weights given by Eq. (3). Therefore the localized pdf is less noisy than the non-localized. However, localization applied in the form of the Karhunen-Loeve expansion given by Eq. (14) does not retain the imposed bounds on the modes $\mathcal{Z}$ as we need to invert a matrix product of eigenvalue and eigenvector matrices to obtain the modes. Moreover unlike ETKF, LETPF does not converge to ETPF as the localization radius goes to infinity due to the transport problem being univariate for LETPF and multivariate for ETPF.

## 5  Conclusions

MCMC methods remain the most reliable methods for estimating the posterior distributions of uncertain model parameters and states. They, however, also remain computationally expensive. Ensemble Kalman filters provide computationally affordable approximations but rely on the assumptions of Gaussian probabilities. For nonlinear models even if the prior is Gaussian the posterior is not Gaussian anymore. Particle filtering on the other hand does not have such an assumption but requires a resam-
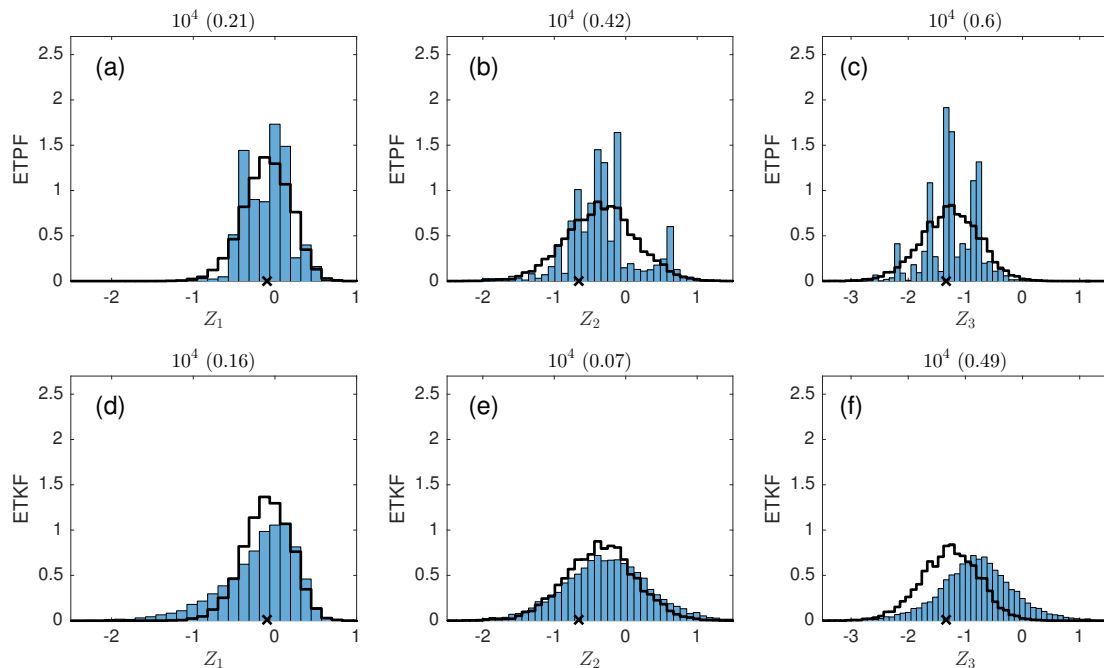
**Figure 11.** The posterior probability density function of parameters $\mathcal{Z}_1$ (left), $\mathcal{Z}_2$ (center), and $\mathcal{Z}_3$ (right). The posterior obtained by IS with ensemble size $10^6$ is plotted as a black line and the true parameter as a black cross. The posterior of ETPF is shown at the top and the posterior of ETKF at the bottom. Both ETPF and ETKF used $10^4$ ensemble members. The Kullback-Leibler divergence is in brackets.

pling step, which is usually stochastic. Ensemble transform particle filter is a particle filtering method that deterministically resamples the particles based on their importance weights and covariance maximization among the particles.

ETPF certainly outperforms ETKF for a one parameter nonlinear test case by giving a better posterior estimation. For the five parameter test case, the mean estimations obtained by ETPF are not consistently better than the ones obtained by ETKF and the spread is smaller. The Kullback-Leibler divergence from ETKF is smaller than from ETPF for all parameters. When the number of uncertain parameters is large (2500) a decrease of degrees of freedom is essential. This is performed by localization. At large ensemble sizes ETPF performs as well as ETKF, while at small ensemble sizes ETKF still outperforms ETPF. Even though localized ETPF overfits the data less often than non-localized, localization destroys the property of ETPF to retain the imposed bounds. This results in deterioration of the leading mode posterior approximation. Another approach to improve ETPF performance is instead of applying localization to use only leading modes in the approximation of log permeability as they are better estimated by the method. However, one needs to know at which mode to make a truncation and this is highly dependent on the covariance matrix of the log permeability.
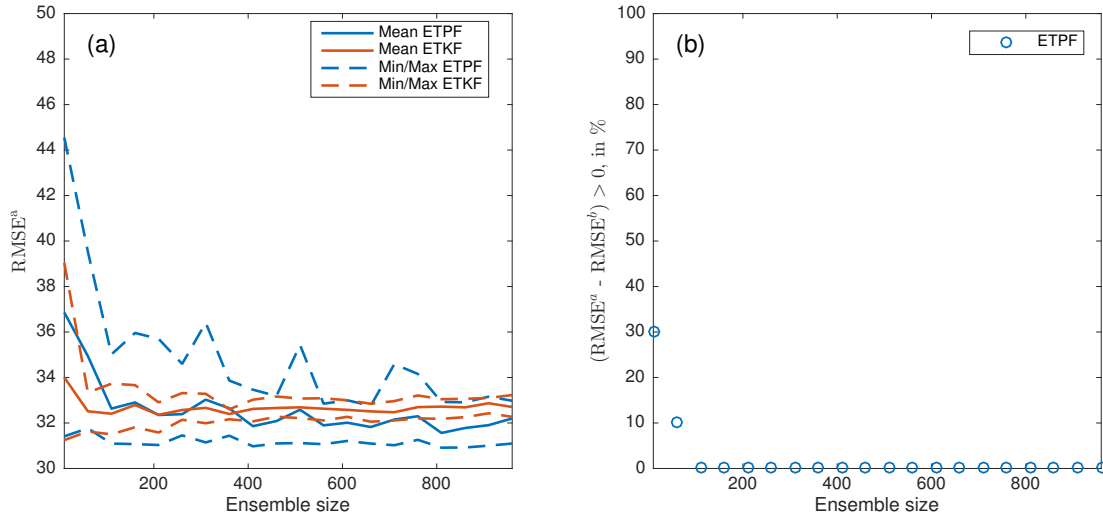
**Figure 12.** Using only three leading modes in the KL expansion. Panel (a): RMSE after data assimilation w.r.t ensemble size with mean, minimum and maximum over 10 simulations for ETPF shown in blue and ETKF in red. Panel (b): % of simulations that result in $(\text{RMSE}^a - \text{RMSE}^b) > 0$ for ETPF.

We believe that ETPF is promising for the inverse modelling. However, more theoretical studies have to be performed for ETPF before it is considered for realistic applications. Plausible issues related to realistic application are numerous accurate observations, time-dependency of an underlined model, and flow being multiphase, for example.

*Competing interests.* The authors declare that they have no conflict of interest.

5   *Acknowledgements.* This work is part of the research programme Shell-NWO/FOM Computational Sciences for Energy Research (CSER) with project number 14CSER007 which is partly financed by the Netherlands Organization for Scientific Research (NWO).
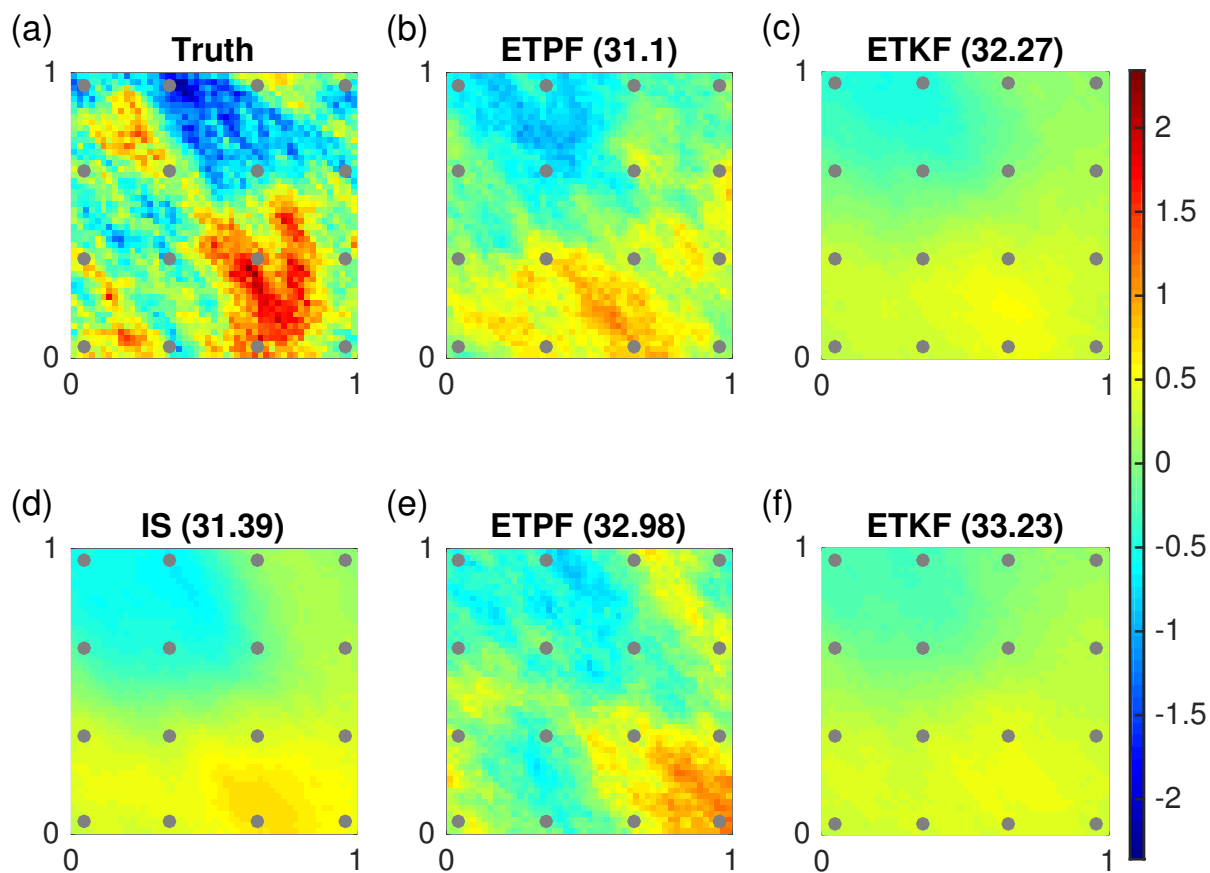
**Figure 13.** Same as figure 8, but using only three leading modes in the KL expansion.

## References

Aanonsen, S. I., Nævdal, G., Oliver, D. S., Reynolds, A. C., Vallès, B., et al.: The ensemble Kalman filter in reservoir engineering–a review, SPE Journal, 14, 393–412, 2009.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J.: Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical
5   Aspects, Monthly Weather Review, 129, 420–436, 2001.

Chen, Y. and Oliver, D. S.: Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification, Computational Geosciences, 17, 689–703, 2013.

Cheng, Y. and Reich, S.: Data assimilation: a dynamical system perspective, Frontiers in Applied Dynamical Systems: Reviews and Tutorials, 2, 75–118, 2015.

10   Doucet, A., de Freitas, N., and Gordon, N.: Sequential Monte-Carlo Methods in Practice, Springer-Verlag, New York, 2001.
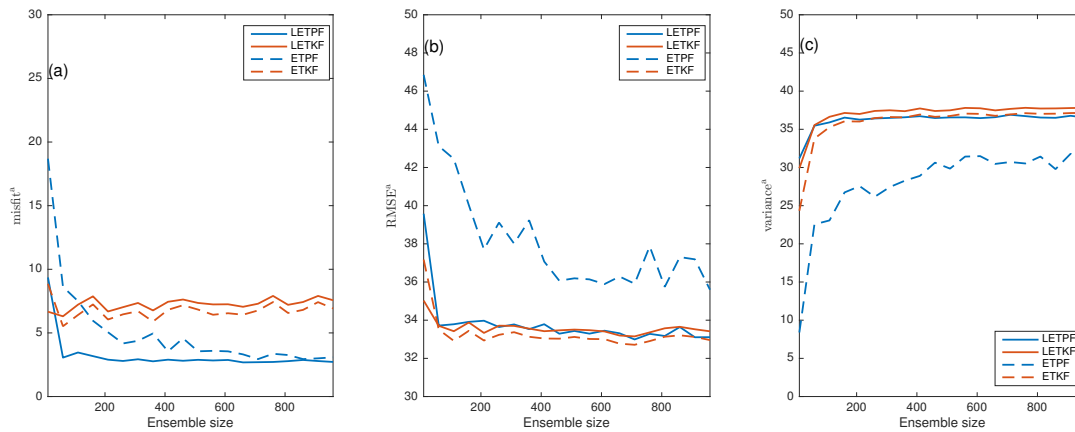
**Figure 14.** Mean over 10 simulations after data assimilation for the data misfit (a), RMSE (b), and variance (c). LETPF is shown in solid blue and LETKF in solid red. ETPF is shown in dashed blue and ETKF in dashed red.

Dovera, L. and Della Rossa, E.: Multimodal ensemble Kalman filtering using Gaussian mixture models, Computational Geosciences, 15, 307–323, 2011.

Evensen, G.: Data assimilation: the ensemble Kalman filter, Springer Science & Business Media, 2009.

Gaspari, G. and Cohn, S. E.: Construction of correlation functions in two and three dimensions, Quarterly Journal of the Royal Meteorological

5    Society, 125, 723–757, 1999.

Guingla, P., Antonio, D., De Keyser, R., De Lannoy, G., Giustarini, L., Matgen, P., and Pauwels, V.: The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter, Hydrology and Earth System Sciences, 16, 375–390, 2012.

Hunt, B., Kostelich, E., and I., S.: Efficient data assimilation for spatialtemporal chaos: A local ensemble transform Kalman filter., Physica

10   D, 230, 112–137, 2007.

Iglesias, M. A., Lin, K., and Stuart, A. M.: Well-posed Bayesian geometric inverse problems arising in subsurface flow, Inverse problems, 30, 114 001, 2014.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state–parameter estimation of hydrological models using ensemble Kalman filter, Advances in water resources, 28, 135–147, 2005.

15   Oliver, D. S. and Chen, Y.: Recent progress on reservoir history matching: a review, Computational Geosciences, 15, 185–221, 2011.

Oliver, D. S., Cunha, L. B., and Reynolds, A. C.: Markov chain Monte Carlo methods for conditioning a permeability field to pressure data, Mathematical Geology, 29, 61–91, 1997.

Oliver, D. S., Reynolds, A. C., and Liu, N.: Inverse theory for petroleum reservoir characterization and history matching, Cambridge University Press, 2008.

20   Pele, O. and Werman, M.: Fast and robust earth mover's distances, in: Computer vision, 2009 IEEE 12th international conference on, pp. 460–467, IEEE, 2009.
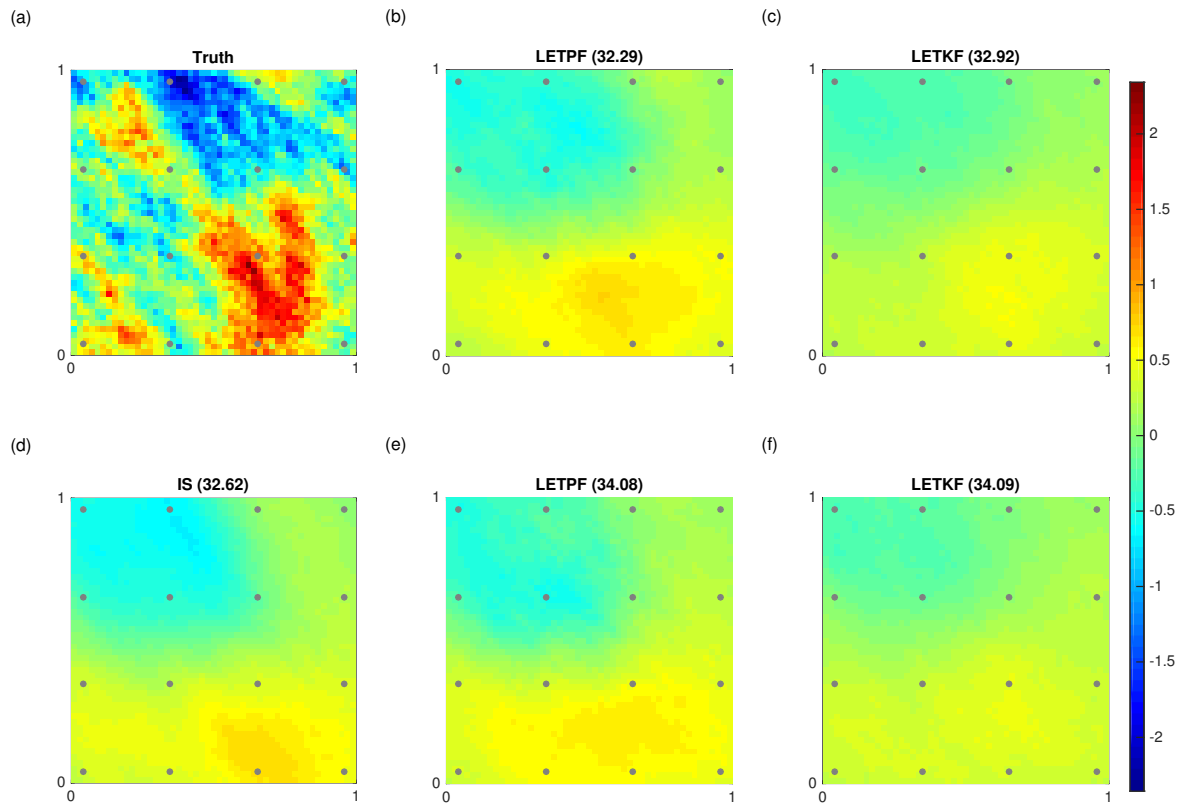
**Figure 15.** Same as figure 8, but with localization.

Penny, S. G. and Miyoshi, T.: A local particle filter for high-dimensional geophysical systems., Nonlin. Processes Geophys., 23, 391–405, 2016.

Poterjoy, J.: A localized particle filter for high-dimensional nonlinear systems., Monthly Weather Review, 144, 59–76, 2016.

Reich, S. and Cotter, C.: Probabilistic forecasting and Bayesian data assimilation, Cambridge University Press, 2015.

5 Reynolds, A. C., He, N., Chu, L., Oliver, D. S., et al.: Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data, SPE Journal, 1, 413–426, 1996.

Vefring, E. H., Nygaard, G. H., Lorentzen, R. J., Naevdal, G., Fjelde, K. K., et al.: Reservoir characterization during underbalanced drilling (ubd): methodology and active tests, SPE Journal, 11, 181–192, 2006.

Weerts, A. H. and El Serafy, G. Y.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-
10    runoff models, Water Resources Research, 42, 2006.
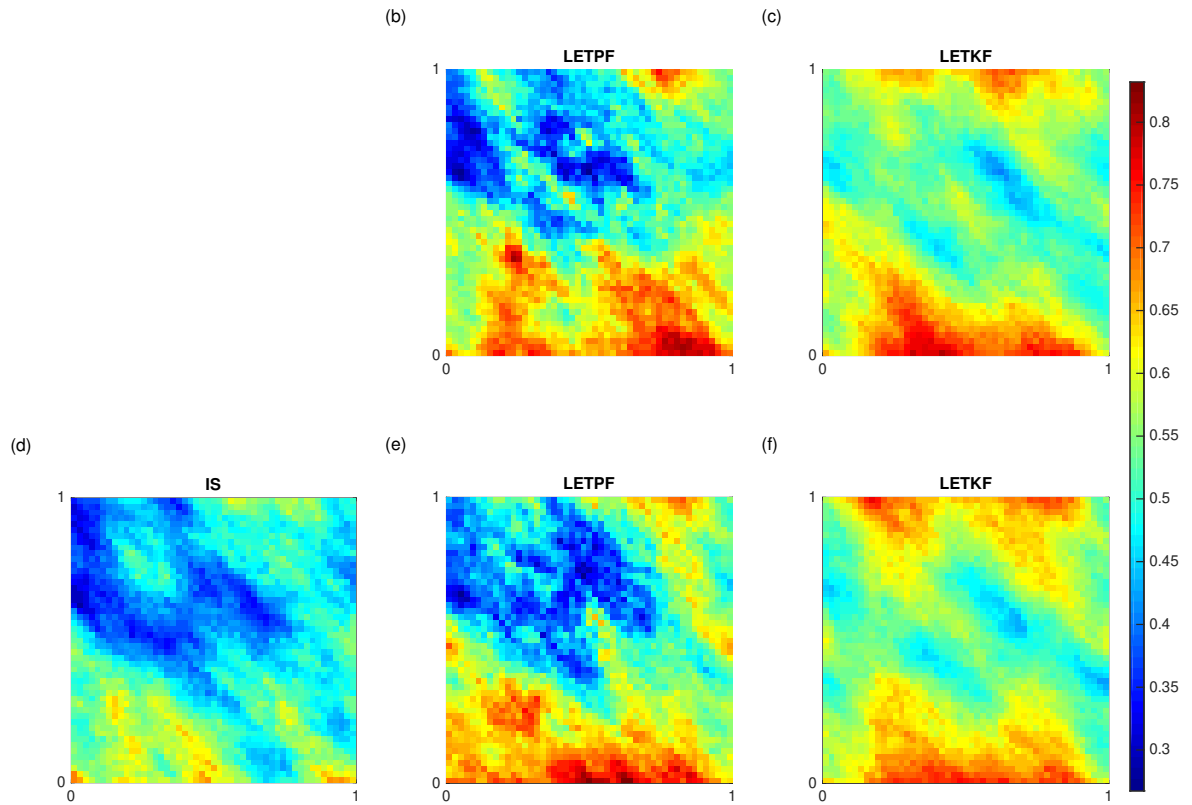
**Figure 16.** Same as figure 9, but with localization.