



Iterative multilevel assimilation of inverted seismic data

Mohammad Nezhadali^{1,2} · Tuhin Bhakta² · Kristian Fossum² · Trond Mannseth²

Received: 28 June 2021 / Accepted: 12 December 2021 / Published online: 16 February 2022
© The Author(s) 2022

Abstract

In ensemble-based data assimilation (DA), the ensemble size is usually limited to around one hundred. Straightforward application of ensemble-based DA can therefore result in significant Monte Carlo errors, often manifesting themselves as severe underestimation of parameter uncertainties. Localization is the conventional remedy for this problem. Assimilation of large amounts of simultaneous data enhances the negative effects of Monte Carlo errors. Use of lower-fidelity models reduces the computational cost per ensemble member and therefore renders the possibility to reduce Monte Carlo errors by increasing the ensemble size, but it also adds to the modeling error. Multilevel data assimilation (MLDA) uses a selection of models forming hierarchies of both computational cost and computational accuracy, and tries to balance between Monte Carlo errors and modeling errors. In this work, we assess a recently developed MLDA algorithm, the Multilevel Hybrid Ensemble Smoother (MLHES), and introduce and assess an iterative version of this algorithm, the Iterative Multilevel Hybrid Ensemble Smoother (IMLHES). In our assessments, we compare these algorithms with conventional single-level DA algorithms with localization. To this end, a typical example of large amount of spatially distributed data, i.e. inverted seismic data, is considered and three data sets of this kind are assimilated in three different petroleum reservoir models. Qualitatively evaluating the DA outcomes, it is found that multilevel algorithms outperform their conventional single-level counterparts in obtaining the posterior statistics of both uncertain parameters and model forecasts. Additionally, it is observed that IMLHES performs better than MLHES in the same regard, and also successfully converges to the proximity of solution in a case where the considered iterative single-level algorithm did not converge to the global optimum.

Keywords Data assimilation · Multilevel methods · Iterative smoother · Ensemble-based history-matching

1 Introduction

Sound decision making in petroleum reservoir management depends on reliable production forecasts from reservoir models, including accurate estimates of uncertainty in the forecasts. The reliability is increased by utilization of available data for calibration of the models.

Ensemble-based Data Assimilation (DA) methods, using statistically correct formulations, have accordingly become popular for automated reservoir history-matching [7, 11, 25, 27, 30, 36, 37].

Monte Carlo approximations play a crucial role in ensemble-based DA. Due to computational-cost limitations,

the ensemble size is limited to roughly one hundred. Using straightforward ensemble-based DA, the degrees of freedom of the problem would equal the ensemble size, and such an approach would result in significant Monte Carlo errors.

The negative effects of Monte Carlo errors are enlarged in the presence of large amounts of simultaneous data, e.g. inverted seismic data, resulting in underestimation of variance of the unknown parameters, and in more severe cases ensemble collapse. There have been several efforts on balancing the degrees of freedom of the problem and the amount of data [29, 31].

The most widely used treatment for Monte Carlo errors is distance-based localization [24]. The basic assumption underlying distance-based localization is that true correlations between a parameter and a datum decrease when the distance between their respective locations increase, and disappears if the distance exceeds a critical distance. This assumption may not always hold for subsurface problems. Different correlation functions and their utilization in DA can be found in [6, 9, 17]. A proper choice of correlation function, and the critical distance in

✉ Mohammad Nezhadali
mone@norceresearch.no

¹ University of Bergen, Nygårdsgaten 112, 5008
Bergen, Norway

² NORCE, Bergen, Norway

particular, depends on parameter and data types as well as on other problem settings. This reduces the robustness of distance-based localization, also for problems where its basic assumption does hold. Papers using ensemble-based methods for assimilation of seismic data [1, 10, 26], typically use localization methodologies developed originally for production data.

Simply increasing the ensemble size will of course reduce Monte-Carlo errors, but it will also increase computational cost. Utilization of a lower-cost and lower-fidelity model renders the possibility of increasing the ensemble size without increasing the total computational cost. Use of a lower-fidelity reservoir model will, however, introduce modeling errors in addition to those already present in conventional-fidelity simulation results. The underlying assumption when applying lower-fidelity models in DA is therefore that the gain in reducing Monte Carlo errors is larger than the loss in numerical simulation accuracy. DA using various types of lower fidelity models has been applied to several inverse problems, e.g., within petroleum reservoir modeling [13, 21, 40] and atmospheric science [20]. Note that since lower-fidelity simulations are applied to the forecast step and localization is applied to the analysis step, the two techniques can be combined, if desired.

Multilevel simulations utilize a selection of models for the same entity that constitute hierarchies in both fidelities and computational costs (multilevel models). The idea is to decrease Monte Carlo errors without increasing numerical errors too much. There are a number of ways to realize multilevel models. We choose to construct them by spatial coarsening of the conventional simulation grid to several levels of coarseness, and correspondingly upscale the associated grid-based parameter functions. Multilevel Data Assimilation (MLDA) [8, 14, 15, 22, 23, 32, 35] utilizes multilevel models in the forecast step. Since inverted seismic data are given on the conventional grid (denoted the fine grid from now on), MLDA with such data must be able to handle differences in grid levels between data and model forecasts.

As utilization of iterative methods helps to improve the quality of history matching in standard Ensemble Smoothers (ES) [7], a similar advancement in the domain of MLDA is possible. An MLDA smoother for assimilation of spatially distributed data, the Multilevel Hybrid Ensemble Smoother (MLHES), was developed in [32] and was assessed for assimilation of inverted seismic data. In this work we further investigate this algorithm and introduce an iterative version of MLHES, the Iterative Multilevel Hybrid Ensemble Smoother (IMLHES). We will also evaluate the performance of these algorithms in comparison with conventional DA algorithms, i.e. fine-level DA with

localization, for assimilation of inverted seismic data in petroleum reservoir problems.

The rest of this paper is organized as follows. Section 2 is devoted to introducing some standard DA algorithms to establish a base for comparison. Section 3 introduces the MLDA algorithms. Section 4 explains the test models used for comparison of the performance of DA schemes. In Section 5 we describe the numerical investigations, which are followed by their results in Section 6. Finally, in Section 7 we summarize and conclude the paper.

2 Standard data assimilation schemes

Ensemble-based DA is a robust method for solving the parameter and state estimation problems using Bayesian methodology. We explain two widely used DA algorithms; Ensemble Smoother (ES), introduced in [41], and Iterative Ensemble Smoother (IES), introduced in [7]; both with localization [16], as conventional methods for DA.

2.1 Ensemble Smoother

Consider the prior ensemble $\{z_j^{pri}\}_{j=1}^{N_e}$, containing N_e realizations from the prior parameters random vector Z^{pri} . Their corresponding forecasts, $\{y_j\}_{j=1}^{N_e}$ (realizations from the random vector Y) are then obtained by running the forward model, \mathcal{M} , on each of the ensemble members as

$$y_j = \mathcal{M}(z_j^{pri}). \quad (1)$$

Consequently, the empirical estimation of the mean and the covariance of the forecast random vector, $E(Y)$ and $C(Y)$, respectively, can be calculated as

$$E(Y) = \frac{1}{N_e} \sum_{i=1}^{N_e} y_j, \quad (2)$$

$$C(Y) = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} (y_j - E(Y)) (y_j - E(Y))^T \quad (3)$$

Let d_j denote a realization of data to be assimilated, drawn from $D \sim \mathcal{N}(\mu_D, \mathbb{C}(D))$, where μ_D is the mean and $\mathbb{C}(D)$ is the data-error covariance. The linear-Gaussian assumption enables the possibility of formulating a closed form for the analysis step in the ES. Accordingly the analysis step for an arbitrary ensemble member z_j can be written as

$$z_j^a = z_j^{pri} + K(d_j - y_j), \quad (4)$$

where K is the Kalman gain defined by

$$K = C(Z, Y) (C(Y) + \mathbb{C}(D))^{-1}, \quad (5)$$

and $C(Z, Y)$ is the empirical estimate of cross-covariance between parameters and model forecasts.

2.2 Iterative ensemble smoother

Iterative versions of ES are developed for improved performance of DA in nonlinear problems. Here we present the algorithm introduced in [7] with confined step length [34], and denote it Iterative Ensemble Smoother (IES) in the rest of this paper.

Writing the posterior logarithm of likelihood using Bayesian update equation and ignoring the constant terms, the objective function is given as [39]

$$J(Z) = \frac{1}{2} \|Z - E(Z^{pri})\|_{C(Z^{pri})}^2 + \frac{1}{2} \|\mathcal{M}(Z) - \mu_D\|_{C(D)}^2 \tag{6}$$

The corresponding objective function for z_j is [7]

$$J_j(z_j) = \frac{1}{2} \|z_j - z_j^{pri}\|_{C(Z^{pri})}^2 + \frac{1}{2} \|\mathcal{M}(z_j) - d_j\|_{C(D)}^2 \tag{7}$$

In the linear-Gaussian case J_j is minimized by the update in Eq. 4. However, assuming nonlinearity in \mathcal{M} , this will not be the case. Accordingly, the Gauss-Newton scheme with confined step length is employed to minimize J_j for all realizations.

The Gauss-Newton scheme requires the gradient and the Hessian of J_j for each of the ensemble members. In this approach, these are computed using approximations based on the ensemble. Accordingly, the gradient, ∇J_j^i , and the Hessian, H^i , for the parameter vector realizations at iteration i , z_j^i , are given as

$$\nabla J_j^i \approx C(Z^{pri})^{-1} (z_j^i - z_j^{pri}) + M^{iT} C(D)^{-1} (\mathcal{M}(z_j^i) - d_{i,j}) \tag{8}$$

$$H^i \approx C(Z^{pri})^{-1} + M^{iT} C(D)^{-1} M^i \tag{9}$$

respectively. In these formulae, M^i denotes the approximation to Jacobian of $\mathcal{M}(Z)$, and H^i is approximated by neglecting the derivative of M^i . M^{iT} denotes the transpose (or more generally, the adjoint) of M^i . In IES [7, 19], the

approximation,

$$M^i \approx C(Y^i, Z^i) C(Z^i)^+ \tag{10}$$

is used for calculation of M^i , where superscript $+$ denotes Moore-Penrose pseudo inverse. A mathematical justification for this can be found in [36, Theorem 1].

The realization z_j^i is then updated as

$$z_j^{i+1} = z_j^i - \beta (H^{i-1} \nabla J_j^i) \tag{11}$$

where β is the step length which is defined based on “restricted-step algorithm” in [34]. The update equation can be written as

$$z_j^{i+1} = z_j^i + \beta (\Delta_j^{i,pri} + \Delta_j^{i,lik}) \tag{12}$$

where $\Delta_j^{i,pri}$ and $\Delta_j^{i,lik}$ are given by

$$\Delta_j^{i,pri} = (I_Z - K^i M^i) [z_j^{pri} - z_j^i] \tag{13}$$

$$\Delta_j^{i,lik} = K^i [d_j - \mathcal{M}(z_j^i)] \tag{14}$$

In these Equations, I_Z is the identity matrix of the parameter vector dimension, and the Kalman gain K^i is given by

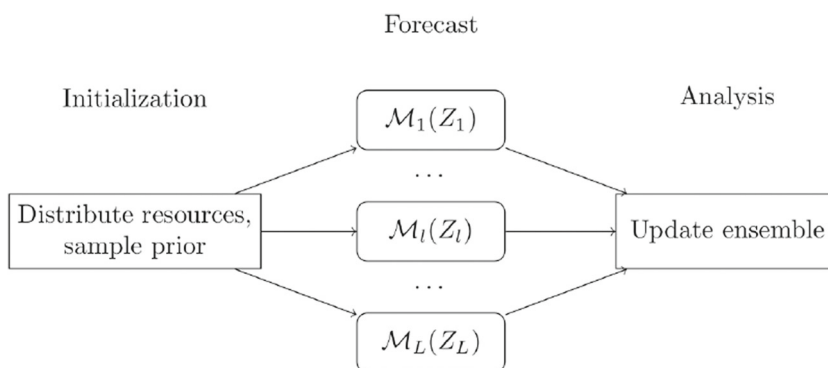
$$K^i = C(Z^{pri}) M^{iT} (M^i C(Z^{pri}) M^{iT} + C(D))^{-1} \tag{15}$$

The iterations are continued until convergence is obtained.

3 Multilevel data assimilation

In MLDA the forecast step is performed using a set of models which have different costs and fidelities. Here, we define $\mathcal{M}_L := \mathcal{M}$, and $\{\mathcal{M}_l\}_{l=1}^{L-1}$ are approximations to \mathcal{M}_L with increasing accuracy and computational cost as l increases. We will denote $\{\mathcal{M}_l\}_{l=1}^L$ a multilevel model. After sampling from the prior distribution, the ensemble of prior state vectors is divided into L sub-ensembles. Each of the sub-ensembles are modeled using its corresponding forward model, as seen in Fig. 1, where subscript l denotes the sub-ensemble index.

Fig. 1 Representation of a single forecast step of MLDA algorithms



3.1 Multilevel models

Multilevel models will form hierarchies of both accuracy and computational cost.

One can think of several schemes to devise the hierarchy, including but not limited to coarsening the spatial grid of the forward model, coarsening the temporal grid of the forward model, and relaxing the convergence criteria in the iterative linear solvers. All of these methods reduce the computational cost of the models and increase their numerical error. Coarsening the spatial grid and performing simulations on such grids is chosen for the current work (Note, however, that the parameters that we invert for are kept in the fine grid, meaning that upscaling the parameters is considered as part of the multilevel forward models). The techniques presented in this work are, however, robust enough so that with minor manipulations, they can be used for other lower fidelity models.

As for coarsening the grid of the forward models, [14] proposed a robust technique, which was also used in [32]. This technique occurs in a sequence of steps. In each step, neighboring cells of the grid at the previous step are merged into a coarser cell unless they are to be kept fine deliberately. A representation of the grid coarsening process for an 8×8 sample grid can be found in Fig. 2. As it can be seen in the figure, coarsening has occurred in a uniform manner across both directions, except for the vicinity of two opposite corners, where the grid cells are kept in fine scale to boost the local numerical accuracy around the two wells, producer (P) and injector (I). The aim is that the grid coarsening does not change the physics of the problem too much.

3.2 Transformation of model forecasts

The discrepancy in coarseness of the multilevel grids results in the spatially distributed model forecasts to be in different resolutions for different levels. Therefore, in order to be able to compute the multilevel sample statistics of model forecast, a robust transformation scheme should be devised

for converting a model forecast from the resolution at one level to another.

In the problem at hand, transformation of the model forecast requires either upscaling or downscaling. Standard volume-weighted arithmetic averaging technique is used for upscaling.

Downscaled model forecasts are simply put equal to the corresponding coarse grid values. Accordingly, both upscaling and downscaling are linear transformations of model forecasts. Hence, we define a set of linear transformations, $\{U_f^c : \mathbb{R}^{\zeta_f} \mapsto \mathbb{R}^{\zeta_c} | 1 \leq c, f \leq L\}$, where ζ_f and ζ_c denote the dimension of model forecast vector at arbitrary levels f and c , respectively, and U_f^c transforms the model forecast vector from level f to be compatible with level c .

Figure 3 gives two examples of transformation of spatially distributed model forecast, one from a coarser grid to a finer grid, and one vice versa. Each model forecast component is represented in its corresponding spatial grid cell. As can be seen in Fig. 3, in the upscaling procedure, the arbitrarily named model forecast components $\{a_i\}_{i=1}^4$ in the northwest zone from the finer grid (level f), are averaged to form their corresponding model forecast component, \bar{a} , in the coarsened grid. Similar procedure has been performed for the rest of model forecast components, shown by *. In the downscaling procedure, on the other hand, the model forecast components in the coarse grid are simply copied to their corresponding components of the finer grid.

3.3 Upscaling of observation data

As part of the DA process, the mismatch between the model forecasts and observation data needs to be calculated. Here, it is assumed that inverted seismic data is given in the resolution of the finest simulation grid. Accordingly, for each of the levels, either the observation data should be upscaled to the resolution of model forecast or the model forecast should be downscaled to the resolution of the observation data. In this paper, we take the former approach.

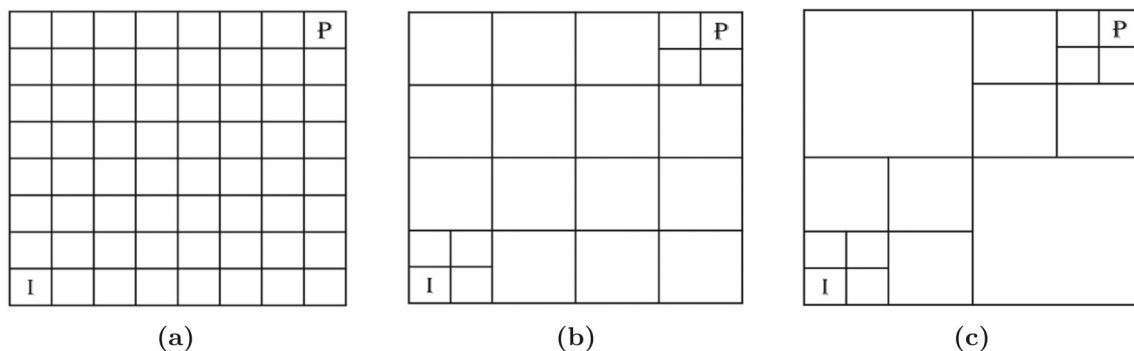
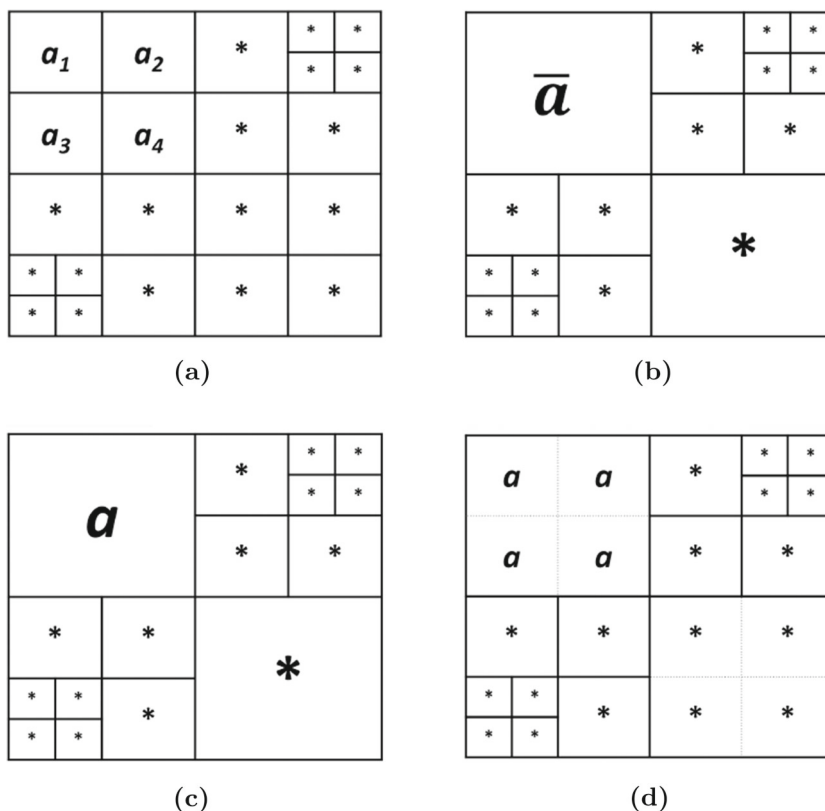


Fig. 2 Grid coarsening proposed by [14] performed on an 8×8 grid (a) Finest level (b, c) Coarser levels

Fig. 3 Transformation of model forecast between two levels f (finer) and c (coarser) (a) model forecast in resolution of level f (b) transformation of model forecast from resolution of level f to the resolution of level c (c) model forecast in resolution of level c (d) transformation of model forecast from resolution of level c to the resolution of level f



Since the observation data is in the resolution of the finest model, using the same transformation functions as those designed for model forecasts on the fine observation data will result in upscaling of observation data into the preferred resolution. Accordingly, the transformed random vector of observation data at level l is given as

$$D_l = U_L^l D . \tag{16}$$

3.4 Multilevel statistics

Assuming we have approximations of the model forecasts, Y , being a function of the unknown parameter vector, Z , on several levels, a statistically correct scheme for approximation of multilevel statistics for Y is required. As for MLDA, the mean and the covariance of model forecast are of foremost interest. Accordingly, formulations for these multilevel statistics are proposed.

Assuming the model with the highest fidelity, \mathcal{M}_L , to be exact, [22] proposed an unbiased formulation for approximation of multilevel statistics for DA under certain conditions. Under these set of conditions, the proposed method outperformed its alternatives [15]. For reservoir problems, however, these conditions typically do not hold, and another formulation inspired by Bayesian Model Averaging (BMA) was proposed [15]. In this formulation the statistics are computed based on reliability weights w_l

for each of the levels l . This formulation is, by definition, a biased scheme for computation of multilevel moments; however, it will be a useful technique for problems in which variance error dominates bias, which is often the case for petroleum reservoir problems [15]. Using this formulation and transformations of the forecast, [32] proposed a formulation of multilevel statistics for spatially distributed model forecasts which will be used in the current work. According to this scheme, the multilevel mean of the model forecast at level l is given as

$$E_{ML,l}(Y) = \sum_{k=1}^L w_k U_k^l E(Y_k) , \tag{17}$$

$$\sum_{k=1}^L w_k = 1 , \tag{18}$$

where $E(Y_k)$ denotes sample mean of the model forecast at level k . Using the law of total variance, the multilevel approximation of covariance of the model forecast at level l is formulated as

$$C_{ML,l}(Y) = \sum_{k=1}^L w_k \{ C(U_k^l Y_k) + (E(U_k^l Y_k) - E_{ML,l}(Y)) (E(U_k^l Y_k) - E_{ML,l}(Y))^T \} . \tag{19}$$

In addition, the parameter-forecast cross-covariance can be written as

$$C_{ML,l}(Z, Y) = \sum_{k=1}^L w_k \{ C(Z_k, U_k^l Y_k) + (E(Z_k) - E_{ML}(Z)) (E(U_k^l Y_k) - E_{ML,l}(Y))^T \}, \quad (20)$$

where $E_{ML}(Z)$ is the multilevel formulation of the parameter-vector mean. This statistic is formulated using the same weights as in forecasts multilevel statistics, but since the parameters are in the same resolution for all levels, no transformation is needed for formulating it,

$$E_{ML}(Z) = \sum_{k=1}^L w_k E(Z_k). \quad (21)$$

Similarly, the multilevel covariance of the parameter vector is defined as

$$C_{ML}(Z) = \sum_{k=1}^L w_k \{ C(Z_k) + (E(Z_k) - E_{ML}(Z)) (E(Z_k) - E_{ML}(Z))^T \}. \quad (22)$$

If all the Z_k , $1 \leq k \leq L$, share the same probability distribution function, which is the case in the first iteration of DA algorithms to be discussed in Section 3.5, Equations 21 and 22 will reduce to $E(Z)$ and $C(Z)$, respectively.

3.5 Multilevel data assimilation algorithms

Multilevel Hybrid Ensemble Smoother (MLHES) was proposed in [32] and tested on a petroleum reservoir model. Here, we firstly explain MLHES algorithm. Following that, we develop an iterative MLDA algorithm based on MLHES; i.e. IMLHES. The coarse model forecasts in both of these algorithms entail Multilevel Modeling Error (MLME), defined as the discrepancy between the upscaled fine model forecasts of a certain realization and the coarse model forecast of that realization. Several methods for correction of this error were assessed in [33]. Here, we utilize mean bias correction for addressing the MLME, which was also used in [32].

3.5.1 Multilevel hybrid ensemble smoother

Based on the decision on resource allocation, N_l ensemble members are simulated using \mathcal{M}_l . Running the forward simulator for every prior realization $z_{l,j}^{pri}$, where $1 \leq l \leq L$

and $1 \leq j \leq N_l$, we have

$$\hat{y}_{l,j} = \mathcal{M}_l(z_{l,j}^{pri}), \quad (23)$$

where $\hat{y}_{l,j}$ is the model forecast pertaining to realization $z_{l,j}^{pri}$. It should be mentioned that $z_{l,j}^{pri}$ vectors are essentially i.i.d samples from the prior distribution regardless of l and have the same dimension. The subscript l is used to denote that the forecasts pertaining to this realization are modeled using \mathcal{M}_l . The models forecasts, $\hat{y}_{l,j}$, however, are in different resolutions for different l , and accordingly have different dimensions.

Next, the multilevel simulations are corrected for their mean bias by

$$y_{l,j} = \hat{y}_{l,j} + \left(U_L^l E(\hat{Y}_L) - E(\hat{Y}_l) \right). \quad (24)$$

This correction will help to account for part of the modeling error associated with coarser models, and results in the forecast mean to be unbiased.

In MLHES, multilevel approximation of the mean and covariance of the model forecasts are utilized for computation of the Kalman gain, and since the model forecasts are in different resolutions for different levels, this is done separately for each level. Accordingly, the updated parameters vector of an arbitrary ensemble member at level l is given by

$$z_{l,j}^a = z_{l,j}^{pri} + K_l(d_{l,j} - y_{l,j}), \quad (25)$$

where the observation data sample, $d_{l,j}$, is a random pick from $\mathcal{N}(U_L^l \mu_D, U_L^l C(D) U_L^l T)$, and the level-specific Kalman gain, K_l , is given as

$$K_l = C_{ML,l}(Z^{pri}, Y) \left(C_{ML,l}(Y) + U_L^l C(D) (U_L^l)^T \right)^{-1}. \quad (26)$$

Here, the multilevel forecast covariance and the parameter-forecast cross-covariance are calculated by Eqs. 19 and 20, respectively.

A pseudo-code of MLHES is presented in Appendix A.

3.5.2 Iterative multilevel hybrid ensemble smoother

In order to be able to handle more nonlinear cases, an iterative version of MLHES is developed. In this algorithm, like with the MLHES, after the forecast step, we correct for the multilevel mean bias. To ease the notations, we consider the bias-correction as part of the new-defined forward models $\{\tilde{\mathcal{M}}_l\}_{l=1}^L$ given as

$$\tilde{\mathcal{M}}_l(Z) = \mathcal{M}_l(Z) + \left(U_L^l E(\hat{Y}_L) - E(\hat{Y}_l) \right). \quad (27)$$

Similarly to how (6) was obtained in Section 2.2, we can write the objective function for level l as

$$J_l(Z) = \frac{1}{2} \|Z - E(Z^{pri})\|_{C(Z^{pri})}^2 + \frac{1}{2} \|\tilde{\mathcal{M}}_l(Z) - U_L^l \mu_D\|_{(U_L^l C(D) U_L^l)^T}^{-1} \cdot \quad (28)$$

Randomizing this objective function using the ensemble of realizations pertaining to level l , the objective function for realization j in sub-ensemble l can be written as

$$J_{l,j}(z_{l,j}) = \frac{1}{2} \|z_{l,j} - z_{l,j}^{pri}\|_{C(Z^{pri})}^2 + \frac{1}{2} \|\tilde{\mathcal{M}}_l(z_{l,j}) - d_{l,j}\|_{(U_L^l C(D) U_L^l)^T}^{-1} \cdot \quad (29)$$

In order for minimization of $J_{l,j}$ for all the realizations, the Gauss-Newton scheme with confined step length is employed. Accordingly, the gradient, $\nabla J_{l,j}^i$, and the Hessian, $H_{l,j}^i$, for the parameter vector realizations at iteration i , $z_{l,j}^i$, are defined by

$$\nabla J_{l,j}^i = C(Z^{pri})^{-1} (z_{l,j}^i - z_{l,j}^{pri}) + M_l^{iT} (U_L^l C(D) U_L^l)^{-1} (\tilde{\mathcal{M}}_l(z_{l,j}^i) - d_{l,j}), \quad (30)$$

$$H_{l,j}^i \approx C(Z^{pri})^{-1} + M_l^{iT} (U_L^l C(D) U_L^l)^{-1} M_l^i, \quad (31)$$

respectively. Here, M_l^i denotes the ensemble approximation to Jacobian of $\tilde{\mathcal{M}}_l$ at iteration i , and $H_{l,j}^i$ is approximated by neglecting the derivative of M_l^i .

If the same procedure that was used in Section 2.2 to obtain (10) was followed for \mathcal{M}_l , the approximation,

$$M_l^i \approx C(Y_l^i, Z_l^i) C(Z_l^i)^+, \quad (32)$$

would be obtained. However, in order to be able to use the information also from other levels for approximation of M_l^i , the multilevel formulation of the parameter-forecast cross-covariance is utilized instead of its single-level formulation,

$$M_l^i \approx C_{ML,l}(Y^i, Z^i) C_{ML}(Z^i)^+ \quad (33)$$

The update formula for an arbitrary realization j of sub-ensemble l at iteration i , $z_{l,j}^i$, can be written as

$$z_{l,j}^{i+1} = z_{l,j}^i - \beta_l (H_{l,j}^{i-1} \nabla J_{l,j}^i), \quad (34)$$

where β_l is the step length at level l which is updated as in the restricted step algorithm, [34], at every iteration. The update equation can be written as

$$z_{l,j}^{i+1} = z_{l,j}^i + \beta_l (\Delta_{l,j}^{i,pri} + \Delta_{l,j}^{i,lik}), \quad (35)$$

where by using Woodbury matrix lemma, $\Delta_{l,j}^{i,pri}$ and $\Delta_{l,j}^{i,lik}$ are given by

$$\Delta_{l,j}^{i,pri} = (I_Z - K_l^i M_l^i) [z_{l,j}^{pri} - z_{l,j}^i], \quad (36)$$

$$\Delta_{l,j}^{i,lik} = K_l^i [d_{l,j} - \tilde{\mathcal{M}}(z_{l,j}^i)], \quad (37)$$

and the level-specific Kalman gain, K_l^i , is obtained by

$$K_l^i = C(Z^{pri}) M_l^{iT} (M_l^i C(Z^{pri}) M_l^{iT} + U_L^l C(D) (U_L^l)^T)^{-1} \quad (38)$$

The iterations are then separately performed for each of the levels until convergence is obtained for all of them.

A pseudo-code of the IMLHES algorithm is presented in Appendix B.

4 Test models

Three different reservoir models are set up to assess the algorithms performances. These reservoir models have some shared properties. They are two-dimensional with Cartesian grids. For all of them, compressible two-phase flow (oil and water), no-flow boundary conditions, and standard equations for capillary pressure and relative permeability are considered. A description of the other shared general properties of the reservoir models is given in Table 1. Unique features of the reservoir models are explained separately in Sections 4.1–4.3.

The forward models used for forecasting each consist of two segments. A reservoir flow model is used to predict the state variables in time, and a petro-elastic model is utilized for computing the elastic rock properties from parameters and predicted state variables.

The flow segment of the forward model is derived by substitution of the Darcy’s law into the mass conservation equation for each of the phases, resulting in [3]

$$\nabla \cdot \left[\frac{k_{ro}}{v_o B_o} k(\nabla p_o - \rho_o g \nabla z) \right] = \frac{\partial}{\partial t} \left(\frac{\phi S_o}{B_o} \right) + q_o, \quad (39)$$

$$\nabla \cdot \left[\frac{k_{rw}}{v_w B_w} k(\nabla p_w - \rho_w g \nabla z) \right] = \frac{\partial}{\partial t} \left(\frac{\phi S_w}{B_w} \right) + q_w, \quad (40)$$

where

$$S_o + S_w = 1, \quad (41)$$

$$p_{cow} = p_o - p_w. \quad (42)$$

Table 1 Shared properties of the reservoir models

Fine cell dimensions:	30 × 30 × 30 (m ³)	Porosity:	0.2
Initial Oil saturation:	0.85	Initial Pressure:	200 bar

In these Equation, k denotes absolute permeability, and k_{r*} denotes the relative permeability of the corresponding phase. k_{r*} is a function of saturation of that phase, S_* . The pressure of a phase is denoted by p_* , and the capillary pressure, p_{cow} , is a function of S_w . Furthermore, g denotes the gravitational constant; ν_* , B_* , and ρ_* are the viscosity, the formation volume factor, and density of their corresponding phases; and q_* denotes the sink or source term of its corresponding continuity equation.

The flow segment of the forward models is performed using Eclipse-100 [38]. Coarsening the grid is done by using the Eclipse keyword COARSEN, which merges groups of pre-defined neighboring cells to form a coarser grid. The upscaling of permeabilities is performed using pore-volume weighted arithmetic averaging, and transmissibilities between two neighboring coarse cells in each direction are calculated based on harmonic averaging in that direction and summing it in other directions [38].

As for the petro-elastic segment of the forward model, an in-house model based on standard rock-physics [4], [12, Report 1] was used.

4.1 Reservoir model I

This model has a 40×40 grid, and two wells, one producer (P) at southwest corner and one injector (I) at northeast corner. Both of the wells are pressure-controlled, the injector at 275 bar and the producer at 100 bar.

4.2 Reservoir model II

This model has a 64×64 grid, and two wells, one producer (P) at southwest corner, and one injector at northeast corner (I). Additionally, an oblique fault stretching from 8 grid cells distant from the northwest corner to 8 grid cells distant from southeast corner is added to the general reservoir model structure. As can be seen in Fig. 5, the coarsening scheme in the presence of such a fault, which will be discussed in Section 5.2, results in some permeability values that are located on one side of the fault in the fine grid to contribute to an upscaled permeability value located on the other side of the fault in the coarsened grid.

4.3 Reservoir model III

This model has a 70×70 grid and three wells, one injection well (I) in southeast corner and two producers (P1 and P2) in southwest and northwest corners. All the wells are pressure-controlled, I at 300 bar and P1 and P2 at 110 bar. This model has three zones, each of which contains one of the wells and has its own variogram for the permeability field, and there exist a smooth transition from one zone to another.

5 Numerical investigation

In order to compare the quality of the multilevel algorithms presented in this work with the standard DA methods, three experiments are conducted. Each experiment is performed on one of the three reservoir models discussed in Section 4.

The unknown parameter fields in all the experiments are logarithmic permeability fields, which have different distributions in each of the experiments.

The observation data are two sets of time-lapse bulk-impedance data taken based on a baseline (day zero of production) and two vintages, which are different for each experiment and will be described separately. These observation data are generated using the results of simulation of a random draw from the prior parameter distribution. As inverted seismic data typically are spatially correlated, we use a correlated covariance matrix for the data error. In doing so, a variogram with the specifications given in Table 2 is considered. The marginal standard deviation of each observation value is given as

$$\sigma = r \max\{|\delta|, \eta\}, \quad (43)$$

where $r = 0.1$, δ is the value of observation data at a certain location, and η is a threshold put to avoid too much certainty in the observation data whose absolute values are very small. This threshold is defined as the 1st smallest percentile of the absolute value of the observation data.

For each numerical experiment we will compare plots of results obtained with multilevel algorithms, MLHES and IMLHES, with the results obtained from their standard DA counterparts, ES with localization (ES-LOC) and IES with localization (IES-LOC). The localization scheme used in ES-LOC and IES-LOC, is a distance-based localization based on the covariance structure given in [16] and spherical model for variogram.

The gold standards (reference solutions) for the comparison will be results obtained using ES with an exceedingly large ensemble (ES-REF) for smoothers, and results obtained using IES with an exceedingly large ensemble (IES-REF) for iterative algorithms. By utilizing such unrealistically large ensembles we obtain results that are visually indistinguishable from the best results that can be achieved using ES and IES. This was assured by running these algorithms with perturbations in their ensemble sizes.

Table 2 Variogram used for observation data error, the unit for range is grid cells

variogram type	range	mean	anisotropy ratio
spherical	5	0	1

Furthermore, we will show plots of the log permeability realizations used when generating the synthetic data (“Truth”).

In each experiment, a fixed computational power is considered for each iteration of all algorithm runs (except for reference solutions). As the dominant cost of the DA process is pertaining to simulations of forward models, where iterative linear solvers dominate the computational costs for large problems, the computational cost pertaining to each ensemble member to be simulated using forward model, \mathcal{M}_l , is assumed to be proportional to G_l^γ , where G_l is the number of the active grid cells in the forward model at level l , and $\gamma \in [1.25, 1.5]$, [2]. Here, we take $\gamma = 1.35$. Additionally, as usual for large-scale cases, the ensemble size for standard single-level DA algorithms is set to be 100. Using this basis for calculations, the computational power allocated for all the runs will be equal if the following equation holds for all of them,

$$100G_L^{1.35} = \sum_{l=1}^L N_l G_l^{1.35}. \tag{44}$$

Considering this equation, we set N_l for different levels of the MLHES. There exists $L - 1$ degrees of freedom for specification of the $\{N_l\}_{l=1}^L$. No optimization was performed for this specification, the only aim pursued was to keep the size of sub-ensembles ascending with decrease in model cost. Several other similar settings that were tried resulted in similar DA outcomes.

For all experiments, the convergence criterion for the iterative algorithms was that improvements in the relative data mismatch should be smaller than 0.0001. The number of iterations required for convergence was approximately the same for IES-LOC and IMLHES. Accordingly, no adjustments were performed for equalizing the total computational cost used by these two algorithms.

For the MLHES and IMLHES, there is a possibility to improve the results by tuning the weights in Eqs. 17–22 for specific cases, but here we use the simplest choice—weights being all equal to $1/L$.

5.1 Experiment I

This experiment is conducted on Reservoir Model I. The multilevel algorithms have four levels, corresponding to 85, 154, 436, and 1600 grid cells, respectively. A summary of the resource allocation for the different runs carried out in this experiment can be found in Table 3. The observation data for this experiment are generated based on seismic vintages at 2000 and 4000 days after production starts.

The unknown logarithmic permeability field is based on a spherical variogram with mean and variance constant at 5 and 1, respectively, anisotropy angle and anisotropy ratio

Table 3 A summary of resource allocation for different runs in experiment I

	level 1	level 2	level 3	level 4
	$G_1 = 85$	$G_2 = 154$	$G_3 = 436$	$G_4 = 1600$
	N_1	N_2	N_3	N_4
ES-LOC	-	-	-	100
MLHES	525	471	231	30
ES-REF	-	-	-	8000
IES-LOC	-	-	-	100
IMLHES	525	471	231	30
IES-REF	-	-	-	8000

of -20 degrees and 0.33, and range 20 grid cells. Randomly selected realizations from this logarithmic permeability field can be found in Fig. 4.

5.2 Experiment II

This experiment is conducted on Reservoir Model II. In this experiment, the presence of the oblique fault in the field interferes with coarsening the model. One way to handle this issue would be to avoid coarsening the grid around the fault area; however, this would reduce the computational efficiency of the multilevel scheme. In order to keep the grid coarsening as it is, the fault is approximated with bigger “zigzags” as depicted in Fig. 5 for one realization of the logarithmic permeability field at different levels of coarseness.

The multilevel algorithms have four levels, corresponding to 124, 310, 1060, and 4096 grid cells, respectively. A summary of the resource allocation for the different runs carried out in this experiment can be found in Table 4. The observation data for this experiment are generated based on seismic vintages at 5000 and 10000 days after production starts.

The unknown logarithmic permeability field is based on a spherical variogram with mean and variance constant at 5 and 1, respectively, anisotropy angle and anisotropy ratio of -30 degrees and 0.7, and range 25 grid cells. Randomly selected realizations from this logarithmic permeability field can be found in Fig. 6.

5.3 Experiment III

This experiment is conducted on Reservoir Model III.

The multilevel algorithms have four levels, corresponding to 163, 412, 1279, and 4900 grid cells, respectively. The observation data for this experiment are generated based on seismic vintages at 3000 and 6000 days after production starts.

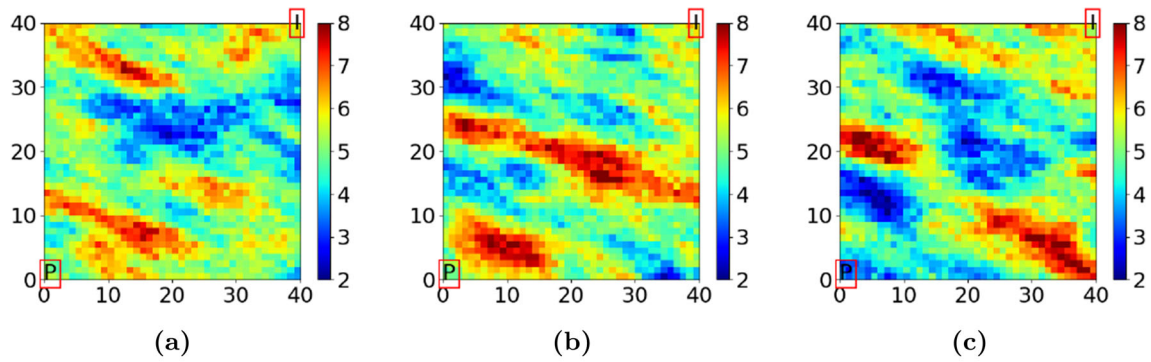


Fig. 4 Randomly selected realizations from prior distribution of the logarithmic permeability field of Reservoir model I

Fig. 5 Approximation of the fault for simulations, a: original fault, b-d: approximations at coarser levels

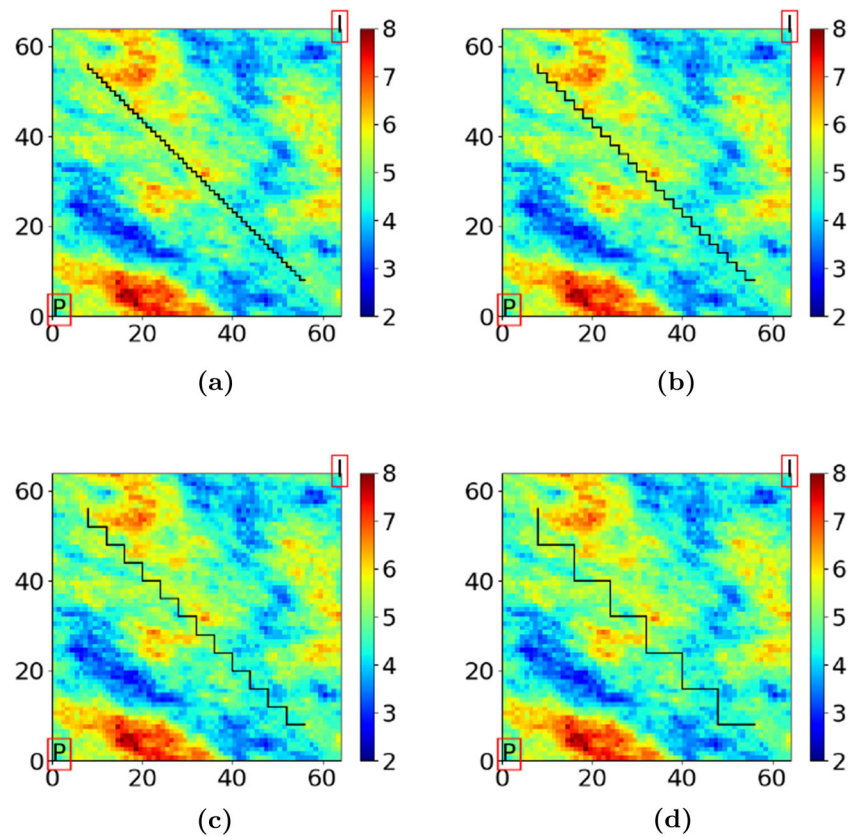


Table 4 A summary of resource allocation for different runs in experiment II

	level 1	level 2	level 3	level 4
	$G_1 = 124$	$G_2 = 310$	$G_3 = 1060$	$G_4 = 4096$
	N_1	N_2	N_3	N_4
ES-LOC	-	-	-	100
MLHES	1404	652	170	40
ES-REF	-	-	-	10000
IES-LOC	-	-	-	100
IMLHES	1404	652	170	40
IES-REF	-	-	-	10000

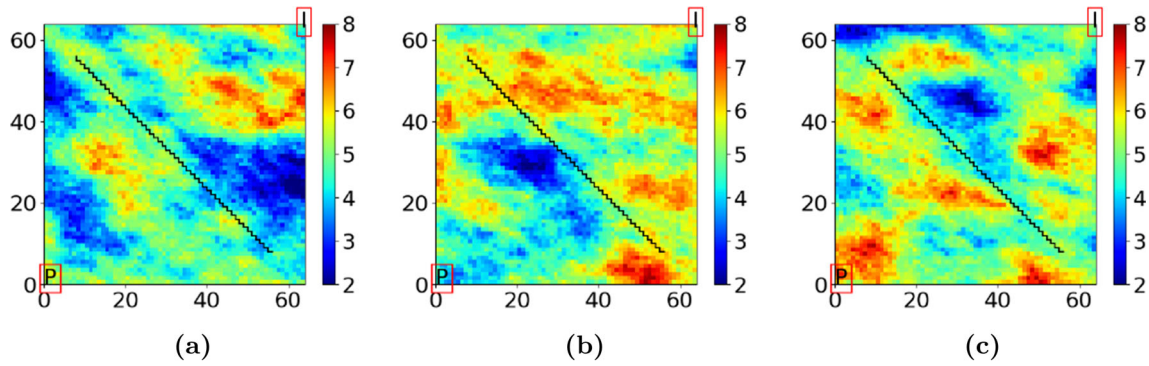


Fig. 6 Randomly selected realizations from prior distribution of the logarithmic permeability field of Reservoir model II

Table 5 A summary of resource allocation for different runs in Experiment III

	level 1	level 2	level 3	level 4
	$G_1 = 163$	$G_2 = 412$	$G_3 = 1279$	$G_4 = 4900$
	N_1	N_2	N_3	N_4
ES-LOC	-	-	-	100
MLHES	791	707	226	30
ES-REF	-	-	-	10000
IES-LOC	-	-	-	100
IMLHES	791	707	226	30
IES-REF	-	-	-	10000

Fig. 7 Reservoir Model III. (a) three zones of the model (b-d) randomly selected realization from prior logarithmic permeability field

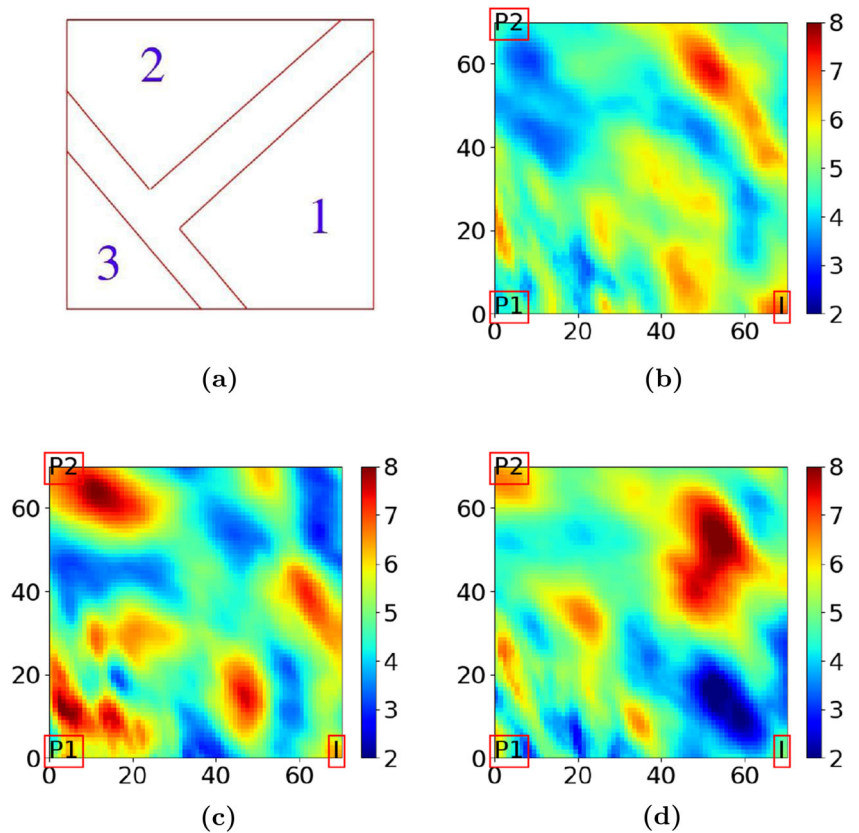


Table 6 The three variograms of Reservoir Model III

Zone	variance	mean	range	ratio	angle	type
1	1	5	20	0.4	-70	cubic
2	1	5	40	0.7	-30	cubic
3	1	5	30	0.6	-60	cubic

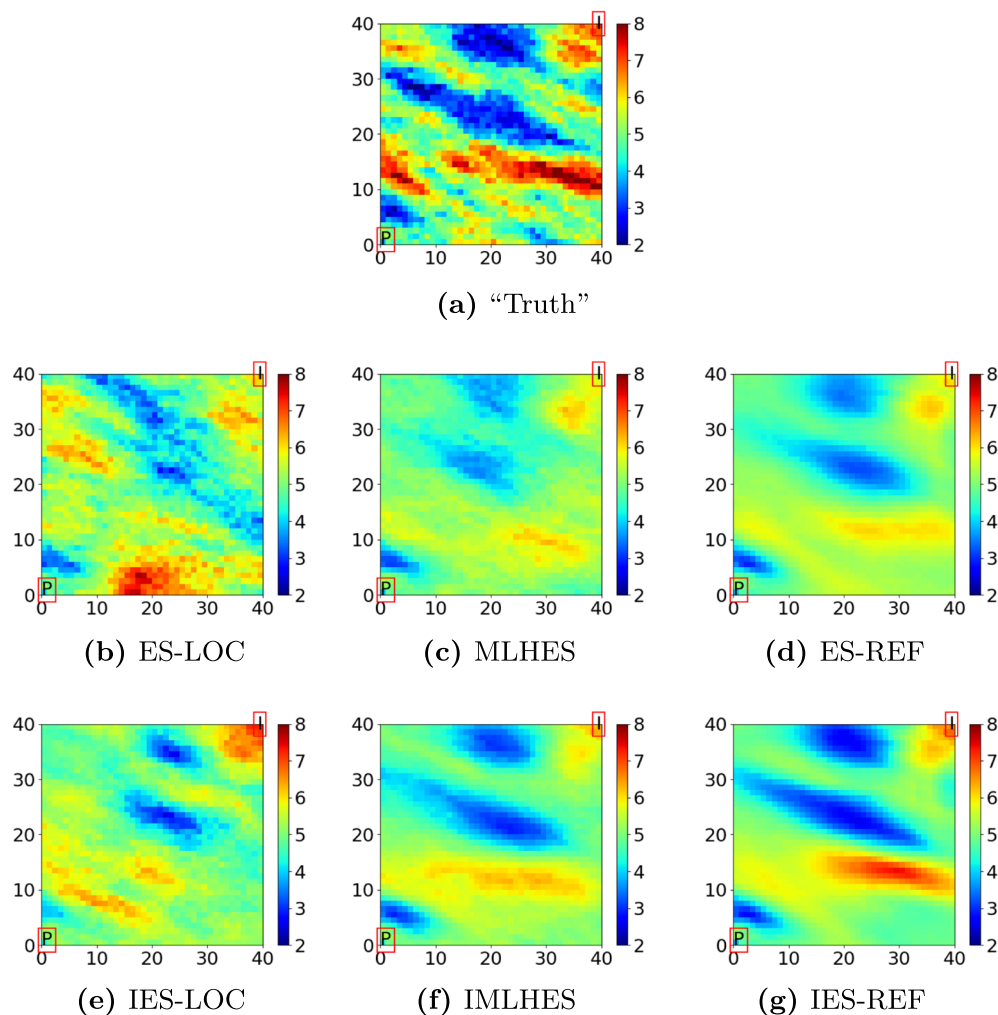
A summary of the resource allocation for different tests carried out in this experiment can be found in Table 5.

The unknown logarithmic permeability field is based on three different variograms in three zones of the field. The Zones, i.e. Zone 1, Zone 2, and Zone 3 can be seen in Fig. 7. The area without any assigned zone is a continuous transition from one zone to others. The details about the variograms based on which the distribution of the unknown parameters are defined can be found in Table 6. Randomly selected realizations from this logarithmic permeability field can be found in Fig. 7.

6 Numerical results

The results from the numerical experiments are assessed qualitatively, using the posterior parameters and forecasts.

Firstly, mean and variance of the posterior parameter fields obtained by different algorithms and reference cases are compared with each other. Additionally, since both the model forecasts and observation data are in different resolutions for different levels of multilevel algorithms, comparison of posterior forecasts as such is not a possibility. Instead, we run fine-scale simulations of the posterior

**Fig. 8** Experiment I—Mean posterior logarithmic permeability field

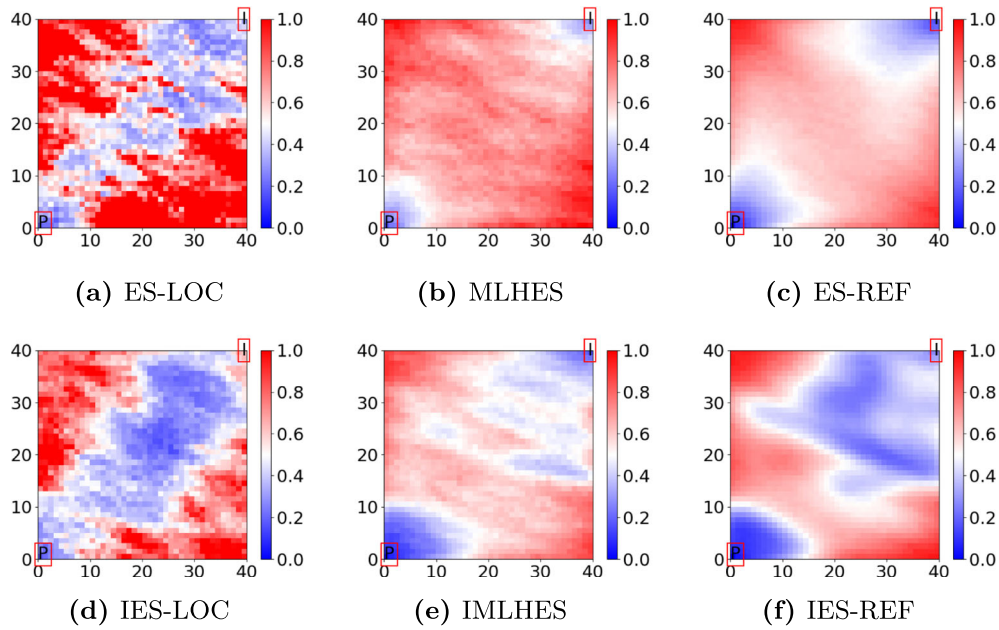


Fig. 9 Experiment I–Variance of posterior logarithmic permeability field

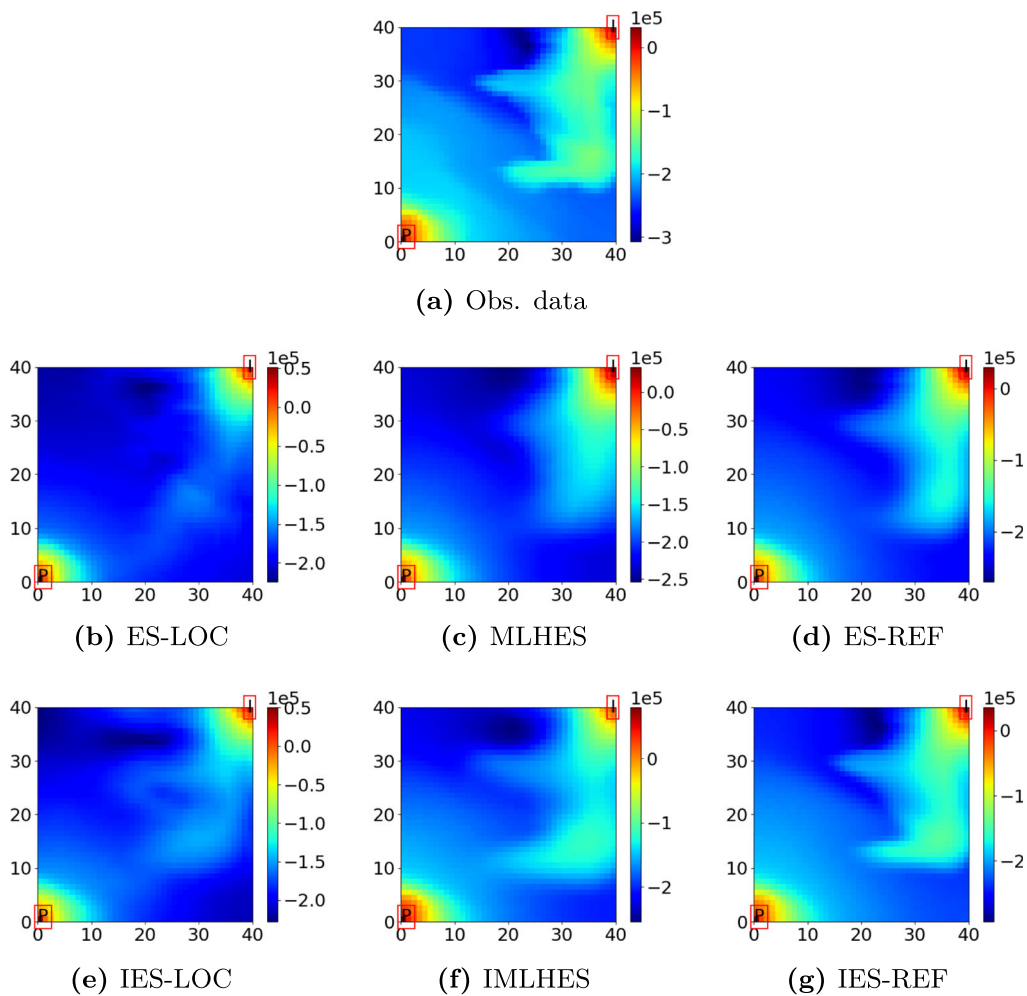


Fig. 10 Experiment I–Posterior time-lapse bulk impedance field ($\frac{m}{s} \frac{kg}{m^3}$) in comparison with observation data in the second vintage

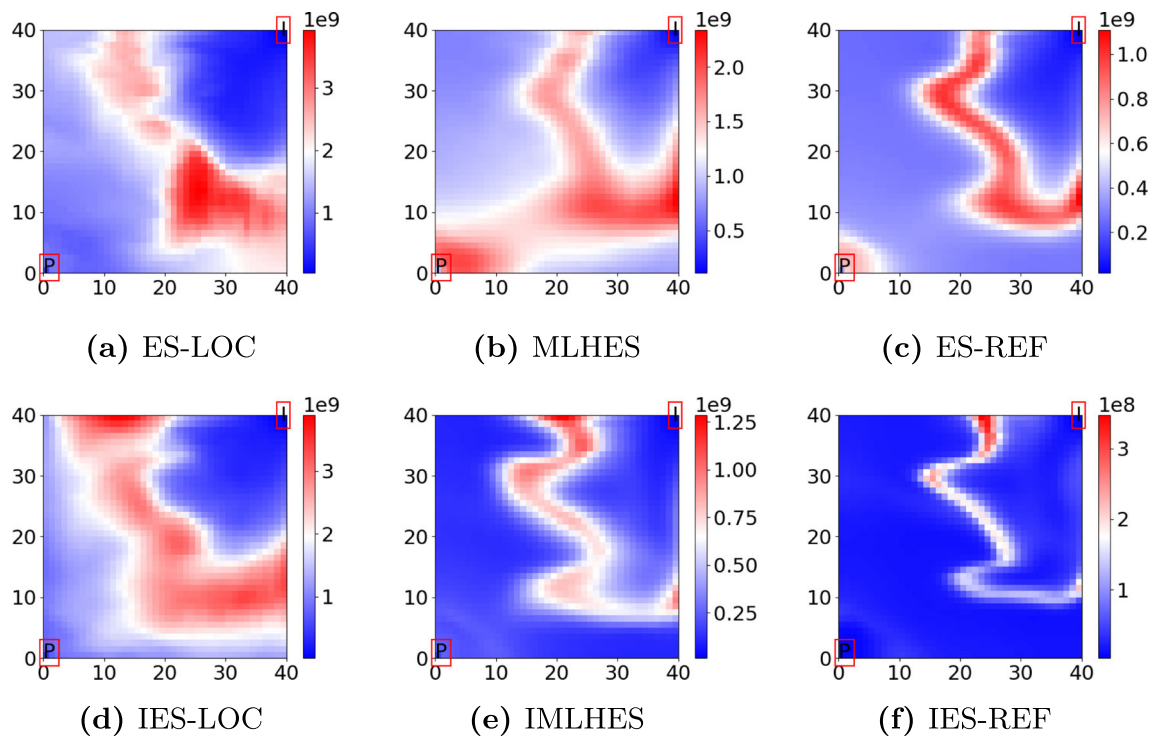


Fig. 11 Experiment I—Variance of posterior time-lapse bulk impedance field ($(\frac{m}{s} \frac{kg}{m^3})^2$), in the second vintage

ensemble for all algorithms and then plot the mean and variance of the model forecasts for all the algorithms and compare them together. The model forecasts of the second vintage are presented for all the experiments.

The simplest formulation is chosen for computation of posterior statistics of MLHES and IMLHES, i.e. re-uniting all the sub-ensembles in these algorithms and treating them as one ensemble for computation of posterior mean and variance fields.

ES-LOC was tested with several ranges for localization (critical distances), and the best results are presented for each of the experiments.

6.1 Results of Experiment I

Visual analysis of the mean permeability fields in Fig. 8 shows that MLHES and IMLHES results are more similar to ES-REF and IES-REF results, respectively, than ES-LOC and IES-LOC results are. This is confirmed by comparison of the variance fields in Fig. 9. In Fig. 8, improvements are seen in approximation of the “Truth” by utilization of IMLHES compared with MLHES.

From Figs. 10 and 11, it is seen that the statistics of posterior forecasts obtained by use of multilevel algorithms are more similar to those of reference cases than the results obtained by use of conventional algorithms. In Fig. 10, the

mean of model forecasts obtained from IMLHES is more similar to the observation data than that of MLHES.

The ES-LOC and IES-LOC results presented here are based on the localization range of 30 grid cells.

6.2 Results of Experiment II

Visual analysis of the mean permeability fields in Fig. 12 shows that MLHES results are more similar to ES-REF results than ES-LOC results are. This can be further confirmed by comparison of the variance fields given in Fig. 13 and the statistics of model forecasts in Figs. 14 and 15.

In this experiment, IES-LOC does not converge to the proximity of global optimum. Since the same holds for IES-REF, there exists no reference for comparison of the iterative algorithms. Nonetheless, as can be seen in Fig. 12, IMLHES results are more similar to “True” permeability field than those of MLHES. Also, in Fig. 14, slight improvements in approximation of the observation data is seen in IMLHES results compared with MLHES results.

For a class of problems (including the problem considered here) where the model forecast can be seen as a spatially integrated response to a spatially varying parameter field, there exists a correlation between small-scale oscillations in the parameter domain and the nonlinearity strength

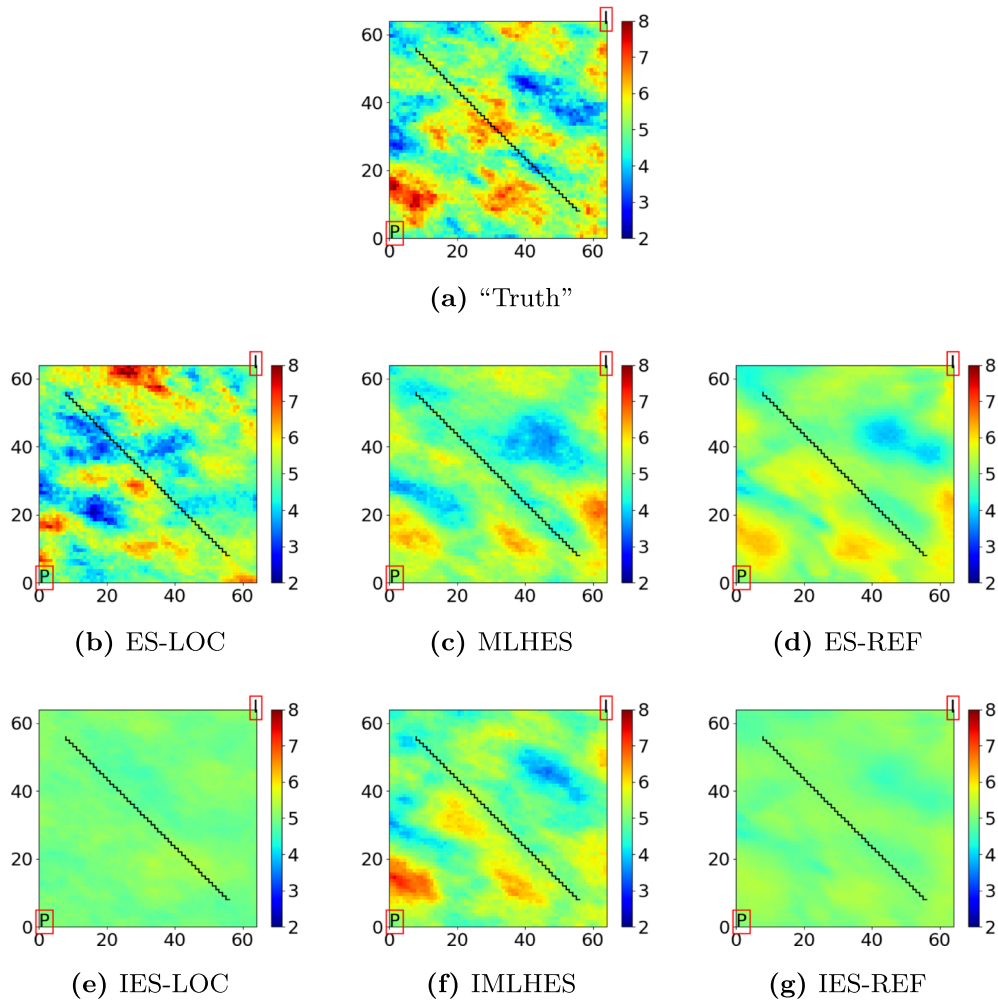


Fig. 12 Experiment II—Mean posterior logarithmic permeability field

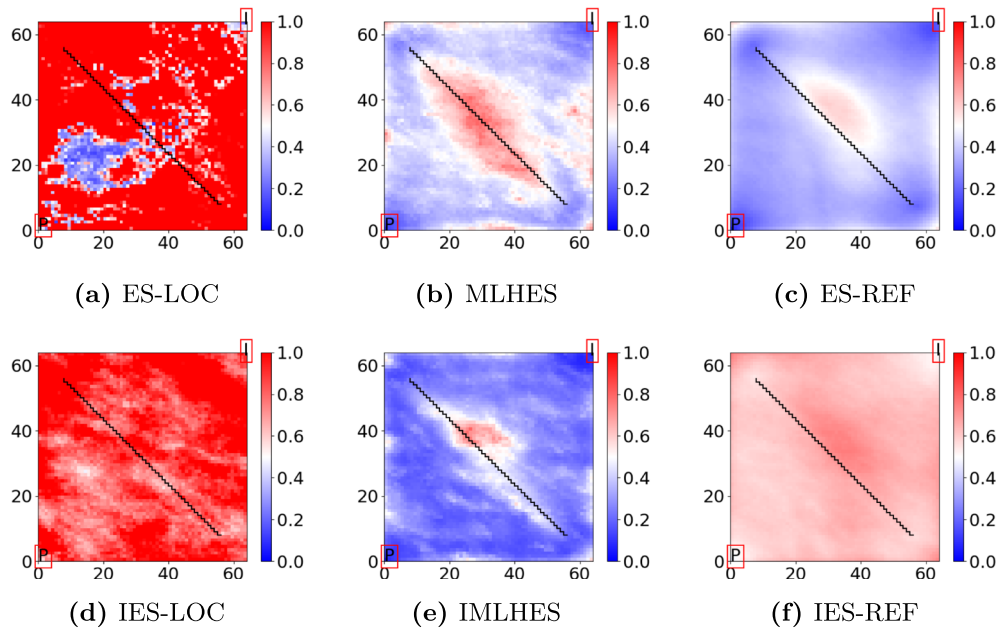


Fig. 13 Experiment II—Variance of posterior logarithmic permeability field

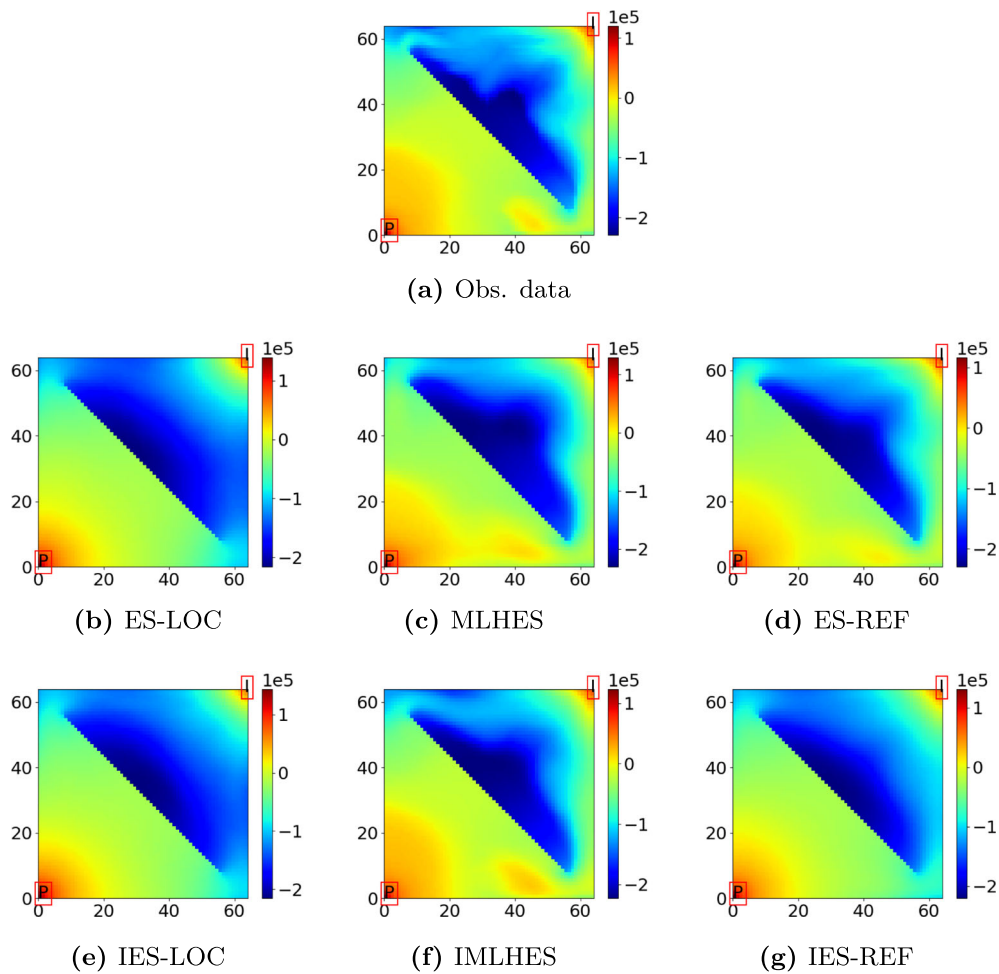


Fig. 14 Experiment II—Posterior time-lapse bulk impedance field ($\frac{m}{s} \frac{kg}{m^3}$) in comparison with observation data in the second vintage

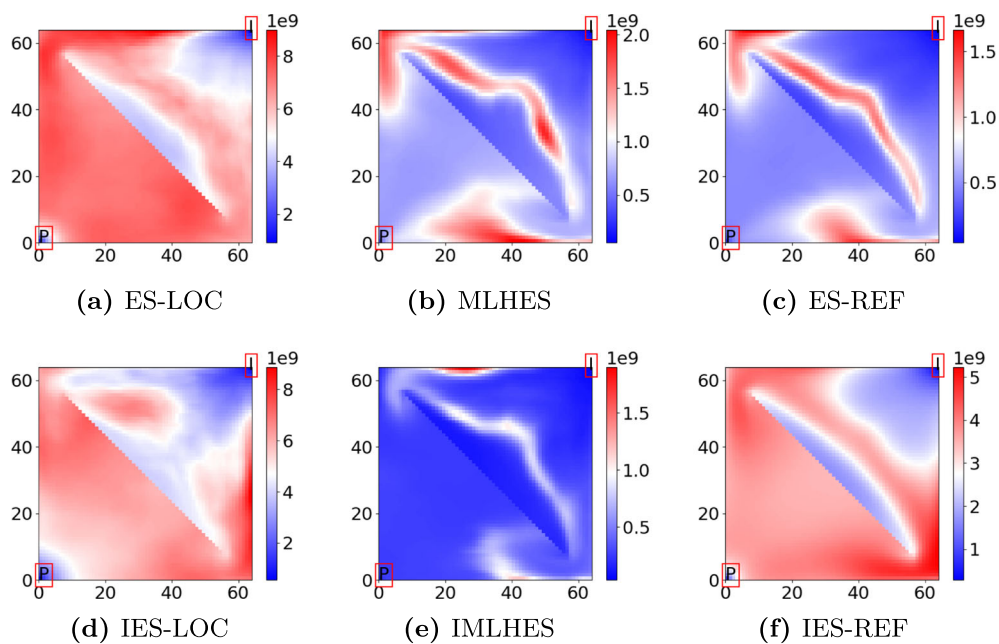


Fig. 15 Experiment II—Variance of posterior time-lapse bulk impedance field ($(\frac{m}{s} \frac{kg}{m^3})^2$), in the second vintage

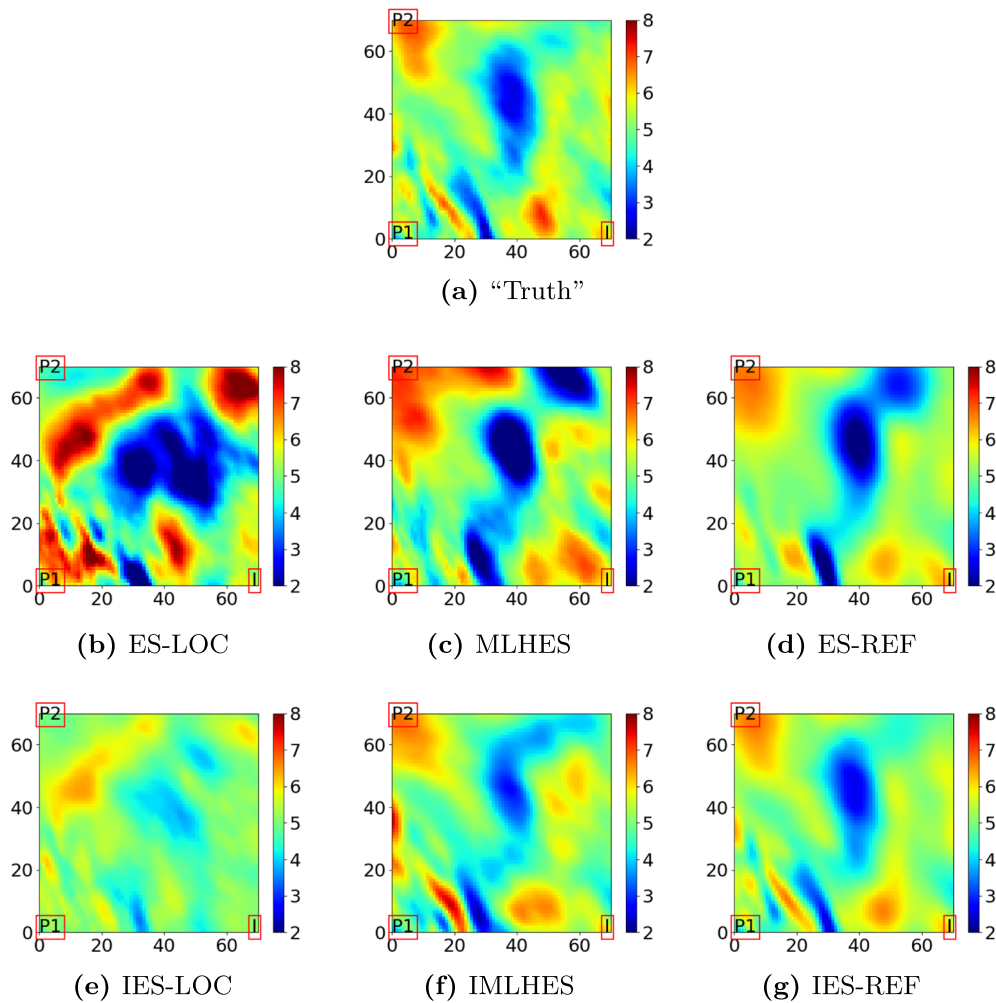


Fig. 16 Experiment III—Mean posterior logarithmic permeability field

of the mapping from parameter field to model forecast, see, e.g., [5, 18]. This correlation is such that coarsening the simulation grid and upscaling the associated parameters will generally result in weaker nonlinearity in the coarser forward models compared to the finer ones. Accordingly, use of IMLHES instead of IES could improve convergence, as was the case in experiment II.

The ES-LOC and IES-LOC results presented here are based on the localization range of 50 grid cells.

6.3 Results of Experiment III

Visual analysis of the mean permeability fields in Fig. 16 shows that MLHES and IMLHES results are more similar to ES-REF and IES-REF results, respectively, than ES-LOC and IES-LOC results are. This is confirmed by comparison of the variance fields in Fig. 17. In Fig. 16, improvements are seen in approximation of the “Truth” by utilization of IMLHES compared with MLHES.

From Figs. 18 and 19, it is seen that the statistics of posterior forecasts obtained by use of multilevel algorithms are more similar to those of reference cases than the results obtained by use of conventional algorithms. In Fig. 18, the mean of model forecasts obtained from IMLHES is more similar to the observation data than that of MLHES.

The ES-LOC and IES-LOC results presented here are based on the localization range of 60 grid cells.

7 Summary and Conclusions

In this work, a recently devised MLDA algorithm (MLHES) for assimilation of spatially distributed data was discussed, and an iterative version of it (IMLHES) was introduced. Both of these methods utilize generalizations of multilevel statistics introduced in [15] for Monte Carlo approximations of mean and covariance of model forecasts. In addition, performance of these algorithms were evaluated in comparison

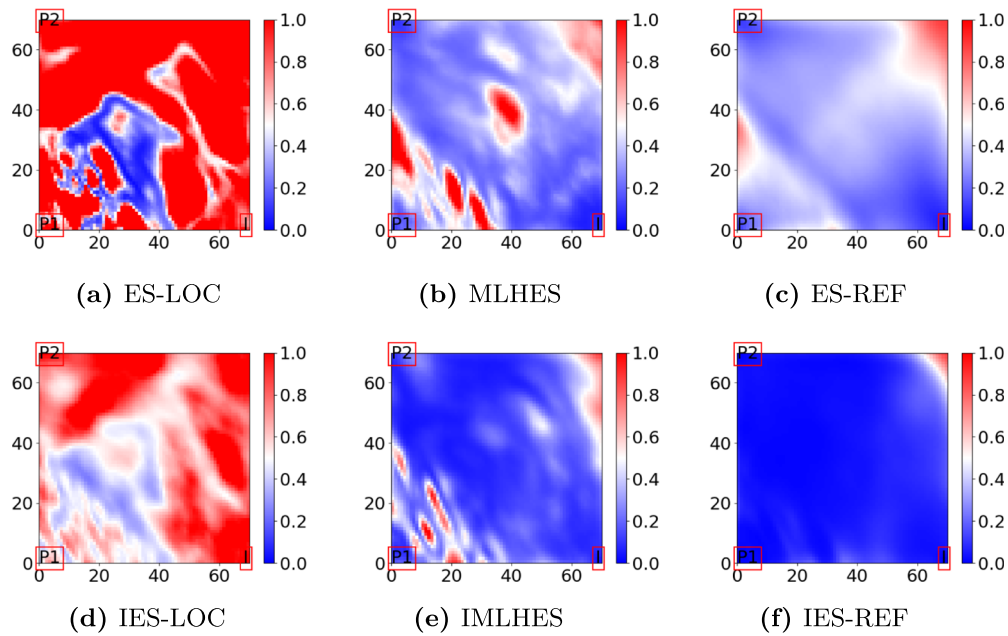


Fig. 17 Experiment III–Variance of posterior logarithmic permeability field

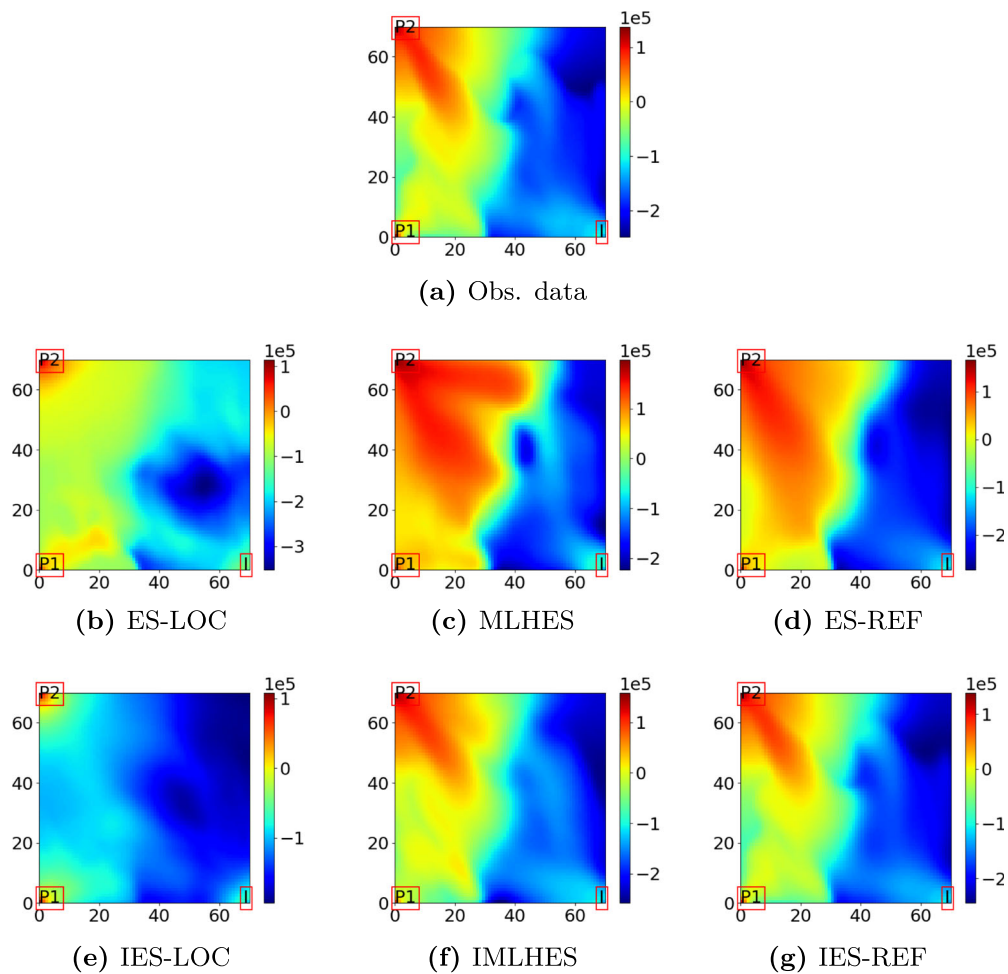


Fig. 18 Experiment III–Posterior time-lapse bulk impedance field ($\frac{m}{s} \frac{kg}{m^3}$) in comparison with observation data in the second vintage

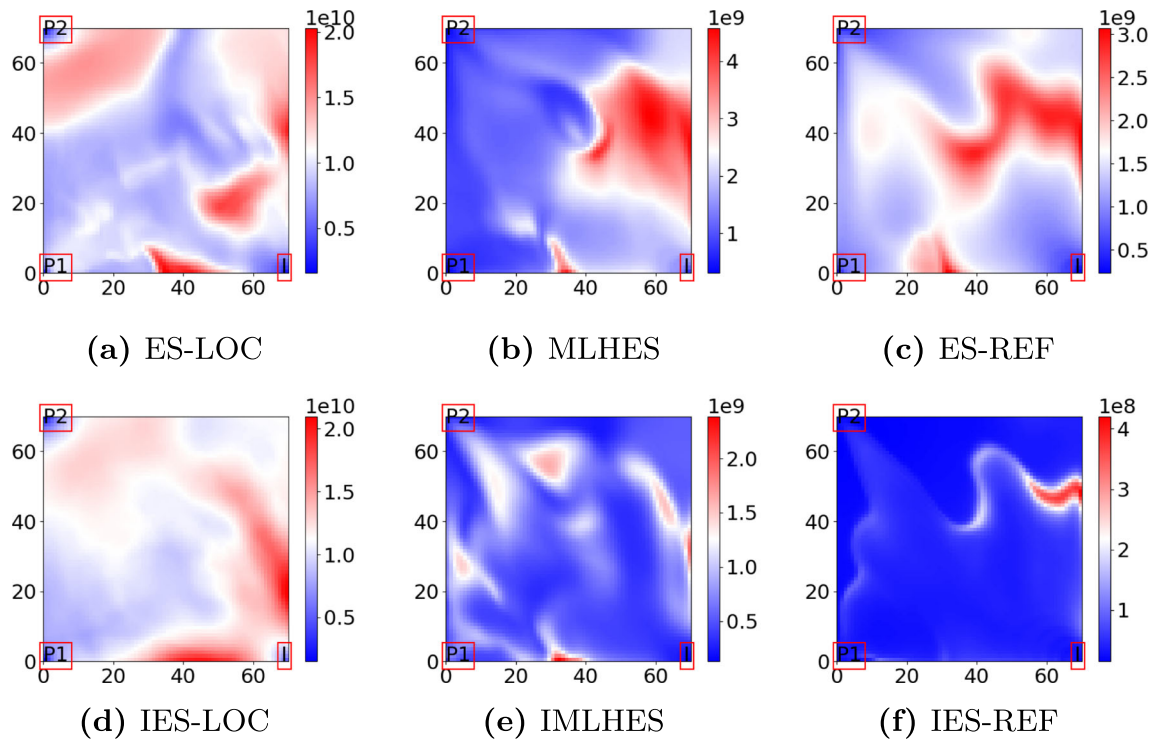


Fig. 19 Experiment III–Variance of posterior time-lapse bulk impedance field ($(\frac{m}{s} \frac{kg}{m^3})^2$), in the second vintage

with standard DA techniques. In doing so, three experiments were conducted. Each experiment was performed on a reservoir model and consisted of six algorithm runs: conventional Ensemble Smoother with localization (ES-LOC), MLHES, ES with an exceedingly large ensemble size (ES-REF), conventional Iterative Ensemble Smoother with localization (IES-LOC), IMLHES, and IES with an exceedingly large ensemble (IES-REF).

In order to assess the numerical results, firstly, the mean and variance fields of posterior unknown parameters (log permeability) were generated and assessed visually. The assessments suggest that in all experiments the results from MLHES were more similar to those of ES-REF than the results from ES-LOC were. Except for one experiment where both IES-LOC and IES-REF did not converge to the proximity of global optimum, the same conclusion was true about the iterative algorithms. The exception suggests an additional advantage of IMLHES over IES. It was also observed that iterations resulted in all the mean posterior fields obtained by IMLHES to be closer to the permeability field from which the observation data were generated than the mean posterior field obtained by MLHES.

Secondly, fine-scale simulations were conducted for all the posterior ensembles of all algorithms. Plots of simulated time-lapse bulk impedance means and variances

were compared to plots of observed time-lapse bulk impedance. Visual analysis of these plots showed that in all the applicable cases, the multilevel algorithms performed more similar to the reference cases than the conventional algorithms did. Additionally, the means of model forecasts obtained from IMLHES were closer to the observation data than the means of model forecasts obtained from MLHES were.

In addition to the presented conventional DA algorithms, which utilize distance-based localization, ES and IES were used in conjunction with correlation-based localization [28] for assimilation of the same data, but this did not improve the conventional DA results. Furthermore, results from investigations that were not presented here suggest that MLDA algorithms show consistency in quality of history matching with respect to variation in observation data error.

There are several issues about the presented MLDA algorithms that can be investigated further. Firstly, the optimal extent of coarsening the grid was not discussed. The rule of thumb was to coarsen the grid until further coarsening results in marginal reduction in grid size, due to restrictions with coarsening around the wells. Additionally, the number of levels and allocation of resources between them to obtain the optimal result can also be further investigated. In this research, the weights for different

levels in the multilevel statistics was set to be all equal; however, there is no such constraint. Accordingly, tuning these weights can be a matter of investigation. Furthermore, the posterior statistics that was presented in this paper was based on a pool comprised of all the updated realizations of different levels put together. Since the accuracy and the computational power allocated per realization differs from level to level, the optimal choice of weights for presenting the posterior statistics needs further research. Moreover, the multilevel modeling error, which was partially accounted for by mean bias correction, can be studied in more detail. For this we refer to [33]. Finally, as realistic reservoir cases are more complex than the fields tested in this paper, the increase in the dimensionality of the parameters may call for a combination of localization and the proposed MLDA algorithms.

Appendix

A Multilevel hybrid ensemble smoother algorithm

Algorithm 1 MLHES algorithm.

```

Define  $\{\mathcal{M}_l\}_{l=1}^L$ 
Define  $\{w_l\}_{l=1}^L$ 
for  $l = 1, \dots, L$  do
  for  $j = 1, \dots, N_l$  do
     $z_{l,j} \leftarrow \mathcal{N}(E(Z^{pri}), C(Z^{pri}))$ 
     $\hat{y}_{l,j} \leftarrow \mathcal{M}_l(z_{l,j})$ 
  end for
end for
for  $l = 1, \dots, L$  do
   $y_{l,j} \leftarrow \hat{y}_{l,j} + \{U_L^l E(\hat{Y}_L) - E(\hat{Y}_l)\}$ 
end for
for  $l = 1, \dots, L$  do
   $E_{ML,l}(Y) \leftarrow \sum_{k=1}^L w_k U_k^l E(Y_k)$ 
   $C_{ML,l}(Y) \leftarrow \sum_{k=1}^L w_k \{C(U_k^l Y_k) + (E(U_k^l Y_k) - E_{ML,l}(Y)) (E(U_k^l Y_k) - E_{ML,l}(Y))^T\}$ 
   $C_{ML,l}(Z, Y) \leftarrow \sum_{k=1}^L w_k \{C(Z_k, U_k^l Y_k) + (E(Z_k) - E_{ML}(Z)) (E(U_k^l Y_k) - E_{ML,l}(Y))^T\}$ 
  for  $j = 1, \dots, N_l$  do
     $d_{l,j} \leftarrow \mathcal{N}(U_L^l \mu_D, U_L^l C(D) U_L^{lT})$ 
     $z_{l,j}^a \leftarrow z_{l,j}^{pri} + C_{ML,l}(Z, Y) (C_{ML,l}(Y) + U_L^l C(D) U_L^{lT})^{-1} (d_{l,j} - y_{l,j})$ 
  end for
end for

```

B Iterative Multilevel Hybrid Ensemble Smoother Algorithm

Algorithm 2 IMLHES algorithm.

```

Define  $\{\mathcal{M}_l\}_{l=1}^L$ 
Define  $\{w_l\}_{l=1}^L$ 
for  $l = 1, \dots, L$  do
  for  $j = 1, \dots, N_l$  do
     $d_{l,j} \leftarrow \mathcal{N}(U_L^l \mu_D, U_L^l C(D) U_L^{lT})$ 
     $z_{l,j}^{pri} \leftarrow \mathcal{N}(E(Z^{pri}), C(Z^{pri}))$ 
     $z_{l,j}^0 \leftarrow z_{l,j}^{pri}$ 
  end for
end for
 $i \leftarrow 0$ 
while not Converged do
  for  $l = 1, \dots, L$  do
    for  $j = 1, \dots, N_l$  do
       $\hat{y}_{l,j}^i \leftarrow \mathcal{M}_l(z_{l,j}^i)$ 
    end for
end for
for  $l = 1, \dots, L$  do
  for  $j = 1, \dots, N_l$  do
     $y_{l,j}^i = \hat{y}_{l,j}^i + (U_L^l E(\hat{Y}_L) - E(\hat{Y}_l))$ 
  end for
end for
for  $l = 1, \dots, L$  do
  if Level has not converged then
     $E_{ML,l}(Y^i) \leftarrow \sum_{k=1}^L w_k U_k^l E(Y_k^i)$ 
     $E_{ML}(Z^i) \leftarrow \sum_{k=1}^L w_k E(Z_k^i)$ 
     $C_{ML,l}(Z^i, Y^i) \leftarrow \sum_{k=1}^L w_k \{C(Z_k^i, U_k^l Y_k^i) + (E(Z_k^i) - E_{ML}(Z^i)) (E(U_k^l Y_k^i) - E_{ML,l}(Y^i))^T\}$ 
     $C_{ML}(Z^i) \leftarrow \sum_{k=1}^L w_k \{C(Z_k^i) + (E(Z_k^i) - E_{ML}(Z^i)) (E(Z_k^i) - E_{ML}(Z^i))^T\}$ 
     $M_l^i \leftarrow C_{ML,l}(Y^i, Z^i) C_{ML}(Z^i)^{-1}$ 
     $K_l^i \leftarrow C(Z^{pri}) M_l^{iT}$ 
     $(M_l^i C(Z^{pri}) M_l^{iT} + U_L^l C(D) (U_L^l)^T)^{-1}$ 
  for  $j = 1, \dots, N_l$  do
     $\Delta_{l,j}^{i,pri} \leftarrow (I - K_l^i M_l^i) [z_{l,j}^{pri} - z_{l,j}^i]$ 
     $\Delta_{l,j}^{i,lik} \leftarrow K_l^i [d_{l,j} - y_{l,j}^i]$ 
     $z_{l,j}^{i+1} \leftarrow z_{l,j}^i + \beta_l (\Delta_{l,j}^{i,pri} + \Delta_{l,j}^{i,lik})$ 
  end for
  Update  $\beta_l$ 
  end if
end for
   $i \leftarrow i + 1$ 
end while

```

Acknowledgments The authors acknowledge financial support from the NORCE research project “Assimilating 4D Seismic Data: Big Data Into Big Models” which is funded by industry partners, Equinor Energy AS, Lundin Energy Norway AS, Repsol Norge AS, Shell Global Solutions International B.V., Total E&P Norge AS, and Wintershall Dea Norge AS, as well as the Research Council of Norway (PETROMAKS2). We also thank Schlumberger for providing academic software licenses to ECLIPSE.

Funding Open Access funding provided by NORCE Norwegian Research Centre AS.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alfonzo, M., Oliver, D.S.: Seismic data assimilation with an imperfect model. *Comput. Geosci.* **24**(2), 889–905 (2020)
- Axelsson, O.: *Iterative solution methods*. Cambridge University Press (1996)
- Aziz, K., Settari, A.: *Petroleum Reservoir Simulation*. Applied Science Publ Ltd., London (1979)
- Batzle, M., Wang, Z.: Seismic properties of pore fluids. *Geophysics* **57**(11), 1396–1408 (1992)
- Chavent, C., Liu, J.: Multiscale parametrization for the estimation of a diffusion coefficient in elliptic and parabolic problems. *IFAC Proc. Vol.* **22**(4), 193–202 (1989)
- Chen, Y., Oliver, D.S.: Cross-covariances and localization for EnKF in multiphase flow data assimilation. *Comput. Geosci.* **14**(4), 579–601 (2010)
- Chen, Y., Oliver, D.S.: Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Math. Geosci.* **44**(1), 1–26 (2012)
- de Moraes, R.J., Hajibeygi, H., Jansen, J.D.: A multiscale method for data assimilation. *Comput. Geosci.* **24**(2), 425–442 (2020)
- Emerick, A., Reynolds, A.: Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. *Comput. Geosci.* **15**(2), 251–269 (2011)
- Emerick, A.A.: Analysis of the performance of ensemble-based assimilation of production and seismic data. *J. Pet. Sci. Eng.* **139**, 219–239 (2016)
- Emerick, A.A., Reynolds, A.C.: Ensemble smoother with multiple data assimilation. *Comput. Geosci.* **55**, 3–15 (2013)
- Fahimuddin, A.: *4D Seismic History Matching Using the Ensemble Kalman Filter (EnKF): Possibilities and Challenges*, PhD thesis. The University of Bergen (2010)
- Fossum, K., Mannseth, T.: Coarse-scale data assimilation as a generic alternative to localization. *Comput. Geosci.* **21**(1), 167–186 (2017)
- Fossum, K., Mannseth, T.: A novel multilevel method for assimilating spatially dense data. In: *ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery*, vol. 2018, pp. 1–12. European Association of Geoscientists & Engineers (2018)
- Fossum, K., Mannseth, T., Stordal, A.S.: Assessment of multilevel ensemble-based data assimilation for reservoir history matching. *Comput. Geosci.* **24**(1), 217–239 (2020)
- Furrer, R., Bengtsson, T.: Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *J. Multivar. Anal.* **98**(2), 227–255 (2007)
- Gaspari, G., Cohn, S.E.: Construction of correlation functions in two and three dimensions. *Q. J. Roy. Meteorol. Soc.* **125**(554), 723–757 (1999)
- Grimstad, A.A., Mannseth, T.: Nonlinearity, scale, and sensitivity for parameter estimation problems. *SIAM J. Sci. Comput.* **21**(6), 2096–2113 (2000)
- Gu, Y., Oliver, D.S.: An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *Spe J.* **12**(04), 438–446 (2007)
- Hatfield, S., Subramanian, A., Palmer, T., Düben, P.: Improving weather forecast skill through reduced-precision data assimilation. *Mon. Weather Rev.* **146**(1), 49–62 (2018)
- He, J., Sarma, P., Durlofsky, L.J.: Reduced-order flow modeling and geological parameterization for ensemble-based data assimilation. *Comput Geosci.* **55**, 54–69 (2013)
- Hoel, H., Law, K.J., Tempone, R.: Multilevel ensemble Kalman filtering. *SIAM J. Numer. Anal.* **54**(3), 1813–1839 (2016)
- Hoel, H., Shaimerdenova, G., Tempone, R.: Multilevel ensemble Kalman filtering based on a sample average of independent EnKF estimators. *Found. Data Sci.* **2**(4), 351 (2020)
- Houtekamer, P.L., Mitchell, H.L.: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weather Rev.* **129**(1), 123–137 (2001)
- Iglesias, M.A.: Iterative regularization for ensemble data assimilation in reservoir models. *Comput. Geosci.* **19**(1), 177–212 (2015)
- Leeuwenburgh, O., Brouwer, J., Trani, M.: Ensemble-based conditioning of reservoir models to seismic data. *Comput. Geosci.* **15**(2), 359–378 (2011)
- Lorentzen, R.J., Bhakta, T., Grana, D., Luo, X., Valestrand, R., Nævdal, G.: Simultaneous assimilation of production and seismic data: application to the Norne field. *Comput. Geosci.* **24**(2), 907–920 (2020)
- Luo, X., Bhakta, T.: Towards automatic and adaptive localization for ensemble-based history matching. In: *ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery*, vol. 2018, pp. 1–26. European Association of Geoscientists & Engineers (2018)
- Luo, X., Bhakta, T., Jakobsen, M., Nævdal, G.: An ensemble 4D-seismic history-matching framework with sparse representation based on wavelet multiresolution analysis. *SPE J.* **22**(03), 985–1010 (2017)
- Luo, X., Bhakta, T., Nævdal, G., et al.: Correlation-based adaptive localization with applications to ensemble-based 4D-seismic history matching. *SPE J.* **23**(02), 396–427 (2018)
- Mannseth, T., Fossum, K.: Assimilating spatially dense data for subsurface applications—balancing information and degrees of freedom. *Comput. Geosci.* **22**(5), 1323–1349 (2018)
- Nezhadali, M., Bhakta, T., Fossum, K., Mannseth, T.: A novel approach to multilevel data assimilation. In: *ECMOR XVII*, vol. 2020, pp. 1–13. European Association of Geoscientists & Engineers (2020)
- Nezhadali, M., Bhakta, T., Fossum, K., Mannseth, T.: Multilevel assimilation of inverted seismic data with correction

- for multilevel modeling error. *Front. Appl. Math. Stat.*, 7. <https://doi.org/10.3389/fams.2021.673077> (2021)
34. Oliver, D.S., Reynolds, A.C., Liu, N.s.: *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press (2008)
 35. Popov, A.A., Mou, C., Sandu, A., Iliescu, T.: A multifidelity ensemble Kalman filter with reduced order control variates. *SIAM J. Sci. Comput.* **43**(2), A1134–A1162 (2021)
 36. Raanes, P.N., Stordal, A.S., Evensen, G.: Revising the stochastic iterative ensemble smoother. *Nonlinear Process. Geophys.* **26**(3), 325–338 (2019)
 37. Sakov, P., Haussaire, J.-M., Bocquet, M.: An iterative ensemble Kalman filter in the presence of additive model error. *Q. J. Roy. Meteorol. Soc.* **144**(713), 1297–1309 (2018)
 38. Schlumberger, L.: *Eclipse reservoir simulation software v2016. Technical Description Manual* (2016)
 39. Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics (2005)
 40. Tarrahi, M., Elahi, S.H., Jafarpour, B.: Fast linearized forecasts for subsurface flow data assimilation with ensemble Kalman filter. *Comput. Geosci.* **20**(5), 929–952 (2016)
 41. Van Leeuwen, P.J., Evensen, G.: Data assimilation and inverse methods in terms of a probabilistic formulation. *Mon. Weather. Rev.* **124**(12), 2898–2913 (1996)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.