Computational Geosciences (2023) 27:265–287
https://doi.org/10.1007/s10596-023-10191-9

ORIGINAL PAPER

# Sequential multilevel assimilation of inverted seismic data

**Mohammad Nezhadali**[1,2] · **Tuhin Bhakta**[1] · **Kristian Fossum**[1] · **Trond Mannseth**[1]

Received: 9 June 2022 / Accepted: 30 December 2022 / Published online: 18 February 2023


**Abstract**

We consider estimation of absolute permeability from inverted seismic data. Large amounts of simultaneous data, such as inverted seismic data, enhance the negative effects of Monte Carlo errors in ensemble-based Data Assimilation (DA). Multilevel (ML) models consist of a selection of models with different fidelities. Multilevel Data Assimilation (MLDA) attempts to obtain a better statistical accuracy with a small sacrifice of the numerical accuracy. Spatial grid coarsening is one way of generating an ML model. It has been shown that coarsening the spatial grid results in a problem with weaker nonlinearity, and hence, in a less challenging problem than the problem on the original fine grid. Accordingly, formulating a sequential MLDA algorithm which uses the coarser models in the first steps of the DA, followed by the finer models, helps to find an approximation to the solution of the inverse problem at the first steps and gradually converge to the solution. We present two variants of a sequential MLDA algorithm and compare their performance with both conventional DA algorithms and a simultaneous (i.e., using all the models on the different grids simultaneously) MLDA algorithm using numerical experiments. Both posterior parameters and posterior model forecasts are compared qualitatively and quantitatively. The results from numerical experiments suggest that all MLDA algorithms generally perform better than the conventional DA algorithms. In estimation of the posterior parameter fields, the simultaneous MLDA algorithm and one of the variants of sequential MLDA (SMLES-H) perform similarly and slightly better than the other variant (SMLES-S). While in estimation of the posterior model forecasts, SMLES-S clearly performs better than both the simultaneous MLDA algorithm and SMLES-H.

**Keywords** Sequential data assimilation · Multilevel methods · Ensemble-based history-matching · Seismic data

## 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–through the process known as history-matching. Ensemble-based Data Assimilation (DA) methods, using statistically correct formulations, have become popular for automated reservoir history-matching [1–7]. We consider estimation of absolute permeability from inverted seismic data.

Ensemble-based methods have limited degree of freedom, which in conjunction with massive amounts of data, e.g. time-lapse seismic data, can result in over-fitting. There have been several efforts to balance the degrees of freedom of the problem and the information content in the data, including use of localization [8], reduction of data using machine learning techniques [9, 10], reduction in data size using the correlation between the data and wells' cumulative production [11], sparse representation of data using a wavelet transform [12], assimilation of only the saturation front or transformation of the data into position of fluid fronts [13–15], combination of coarsening the data and coarse model simulations [16], and projection of data into ensemble subspace in combination with local analysis [17].

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

✉ Mohammad Nezhadali
monezhadali@gmail.com

[1] NORCE Norwegian Research Center, Nygårdsgaten 112, 5008, Bergen, Norway

[2] University of Bergen, Natural Science Building, Allégaten 41, 5007, Bergen, Norway

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, such as inverted seismic data, resulting in underestimation of variance, and in more severe cases ensemble collapse.

The most widely used treatment for Monte Carlo errors is distance-based localization [18]. 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 increases, and disappears if the distance exceeds a critical distance. This assumption may not always hold for subsurface problems. Different localization functions and their utilization in DA can be found in [19–21]. A proper choice of localization function, and the critical distance in 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 [22–24], typically use localization methodologies developed originally for meteorological science which were later adapted to petroleum problems with production data.

Simply increasing the ensemble size will reduce Monte-Carlo errors, but it will also increase the 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 [25–27] and atmospheric science [28]. 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.

Lower-fidelity models can have high numerical errors. Additionally, choosing an optimal level of fidelity for these models is not straightforward. Multilevel (ML) simulations utilize a selection of models, which form hierarchies of both computational accuracy and computational costs (ML model). Multilevel Data Assimilation (MLDA) [29–37] utilizes an ML model in the forecast step of the DA. The levels of the multilevel model have different numerical accuracies. The MLDA, allocating some of the computational power to the models with lower fidelity, tries to achieve a low total error by keeping a balance between the numerical errors and statistical errors.

Some conventional DA methods (i.e., single-level ensemble-based DA methods) , like the ensemble smoother (ES) [38], the ES with multiple DA (ESMDA) [2], and the iterative ES (IES) [1], assimilate data simultaneously, i.e. assimilate all data over a certain period at once, while other conventional methods, like the ensemble Kalman filter (EnKF) [39] and the EnKF with multiple DA (EnKFMDA) [40], assimilate data sequentially. In the MLDA domain, a Simultaneous MLDA (SiMLDA) algorithm was developed for assimilation of inverted seismic data [33]. Numerical experiments show that this algorithm, the Multilevel Hybrid Ensemble Smoother (MLHES) whose formulation uses a hybrid Kalman gain based on several levels, outperforms its conventional DA counterparts with which it was compared [33, 37, 41]. However, strong nonlinearity can affect its performance negatively. Hence, another SiMLDA algorithm, the iterative version of MLHES (IMLHES) [37], was designed to handle this problem. Numerical experiments show that use of iterations improves performance of MLHES [37], but IMLHES is not without limitations. It is not obvious how to optimally formulate convergence criteria for the different levels of IMLHES. This could cause failed iterations and additional computational cost.

There are indications that utilization of sequential MLDA (SeMLDA) algorithms can benefit from certain properties of the problem of estimation of permeability from inverted seismic data. Firstly, analytical [42] and numerical [43] results show that sequential DA is expected to outperform simultaneous DA for weakly nonlinear problems. Secondly, 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 degree of nonlinearity of the mapping from parameter field to model forecast, see, e.g., [44, 45]. 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. Taking advantage of this, we consider several resolutions of inverted seismic data, and develop a SeMLDA scheme which first assimilates the coarsest resolution of the data corresponding to the coarsest forward model, followed by the data in higher resolution corresponding to more nonlinear models. This construction corresponds to the optimal ordering of data as suggested by [42, 43]. Note that, here, the term ‘sequential’ pertains to a sequence of data with different resolutions, not data at different times.

In this work, we will introduce two variants of a SeMLDA algorithm and assess their performance in comparison with a conventional sequential DA algorithm, a conventional simultaneous DA algorithm, and a SiMLDA algorithm. This will be done with the help of numerical experiments.

The rest of this paper is organized as follows. Since the SeMLDA methods developed in this paper are partially inspired by the ESMDA, the ESMDA will be briefly presented in Section 2. Section 3 describes MLDA in general, and introduces two variants of a novel SeMLDA algorithm. Section 4 explains the test models used for numerical investigation. 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 Ensemble smoother with multiple data assimilation

The forward models used in the parameter estimation process are often nonlinear. In the case of assimilation of inverted seismic data, the nonlinearity comes from both fluid flow equations and rock physics modeling. This nonlinearity has motivated the development of several DA algorithms, including ESMDA [2]. This algorithm assimilates the same data $N_a$ times while inflating the data-error covariance matrix. By doing so, several small assimilation steps are taken instead of one big assimilation step. This helps to better account for the nonlinearity of the problem.

At step $i$ of ESMDA, an ensemble of $N_e$ realizations $\{z_{i,j}^{\text{pri}}\}_{j=1}^{N_e}$ from the prior parameter vector $Z_i^{\text{pri}}$ is considered. The ESMDA update consists of three steps.

Firstly, running the forward simulator for every realization, $z_{i,j}^{\text{pri}}$, $1 \leq j \leq N_e$, we have

$$y_{i,j} = \mathcal{M}\left(z_{i,j}^{\text{pri}}\right) , \tag{1}$$

where $y_{i,j}$ is a realization from the model forecasts random vector at assimilation step $i$, $Y_i$.

Secondly, the original observation data-error model, $D \sim \mathcal{N}(\mathbb{E}(D), \mathbb{C}(D))$, is slightly manipulated such that the random observation data vector at step $i$ is given as $\widehat{D}_i \sim \mathcal{N}(\mathbb{E}(D), \alpha_i \mathbb{C}(D))$. The data-error covariance matrix is inflated using the scalar value $\alpha_i$ so that multiple assimilations of the same data does not violate Bayes' rule. Hence, the updated parameter vector of an arbitrary ensemble member is given by

$$z_{i,j}^{\text{pos}} = z_{i,j}^{\text{pri}} + K_i \left(\widehat{d}_{i,j} - y_{i,j}\right) , \tag{2}$$

where $\widehat{d}_{i,j}$ is a realization of $\widehat{D}_i$, and the Kalman gain, $K_i$, is given as

$$K_i = \mathrm{C}\left(Z_i^{\text{pri}}, Y_i\right)\left(\mathrm{C}(Y_i) + \alpha_i \mathbb{C}(D)\right)^{-1} . \tag{3}$$

The terms $\mathrm{C}(Y_i)$ and $\mathrm{C}\left(Z_i^{\text{pri}}, Y_i\right)$ denote the the covariance of $Y_i$ and cross-covariance between $Z_i^{\text{pri}}$ and $Y_i$, respectively.

Finally, while $i < N_a$, the prior ensemble at step $i+1$ is set equal to the posterior ensemble at step $i$;

$$z_{i+1,j}^{\text{pri}} := z_{i,j}^{\text{pos}} , \tag{4}$$

for $1 \leq j \leq N_e$.

In general, the likelihood can be written as a product of inflated likelihoods. This process, known as tempering, fulfills Bayes theorem exactly. ESMDA is a special case of tempering. If the forward model, $\mathcal{M}$, is a linear model and the distributions of the prior parameters and the data errors are Gaussian, there exists a condition (denoted the MDA condition) for $\alpha_i$ which ensures that the ESMDA samples correctly from the posterior distribution. This condition is given as

$$\sum_{i=1}^{N_a} \frac{1}{\alpha_i} = 1 . \tag{5}$$

It is, however, unclear how important (5) is for problems where normality or linearity does not hold.

## 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 $\{\mathcal{M}_l\}_{l=1}^{L-1}$ where $\mathcal{M}_l$ is an approximation to $\mathcal{M}_L$ with increasing accuracy and computational cost as $l$ increases. We will denote $\{\mathcal{M}_l\}_{l=1}^{L}$ an ML model.

### 3.1 Multilevel model

Members of an ML model 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, increasing the length of the forward model time steps, 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 estimate are kept in the fine grid, meaning that upscaling the parameters is considered as part of the ML forward models).

Fossum and Mannseth [32] proposed a robust technique for grid coarsening, which was also used in [33, 37, 41]. In each coarsening 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 $8 \times 8$ example grid can be found in Fig. 1. The figure illustrates that coarsening has occurred in a uniform manner across both directions, except for the

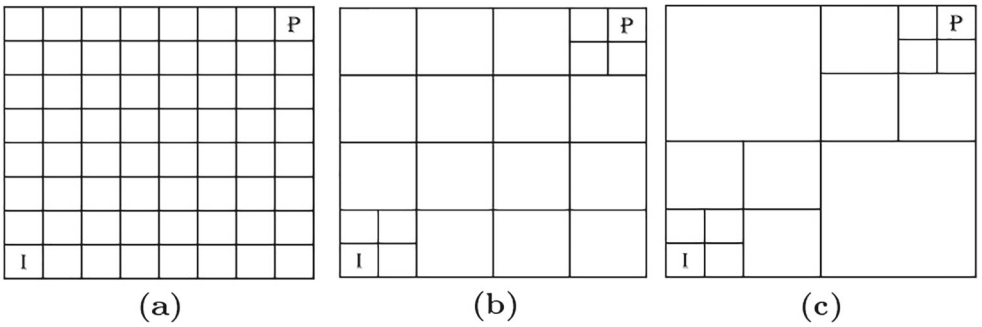



**Fig. 1** Grid coarsening proposed by [32] performed on an $8 \times 8$ grid (a) Finest level (b,c) Coarser levels

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 should not change the physics of the problem too much.

### 3.2 Transformation of model forecasts

The discrepancy in coarseness of the ML 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 ML sample statistics of model forecasts, 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. A standard volume-weighted arithmetic averaging technique is used for upscaling. Since the grids of the ML model used here have a nested structure, 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 $\zeta_f$ and $\zeta_c$ 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 level $c$.

### 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, level $L$. Therefore, for each of the levels, either the observation data should be upscaled to the resolution of the model forecasts or the model forecasts should be downscaled to the resolution of the observation data. We take the former approach. Since the observation data is in the resolution of the finest model, using the same transformation functions as those

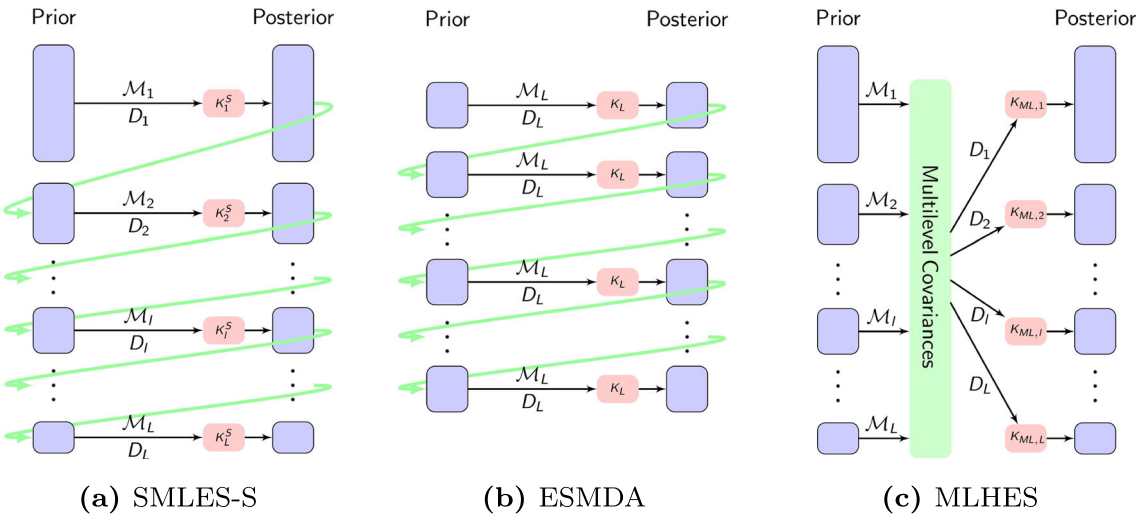



**Fig. 2** The schematic of two MLDA algorithms compared with ESMDA

designed for the model forecasts on the fine observation data will result in upscaling of observation data into the required resolution. Hence, the transformed random vector of observation data at level $l$ is given as

$$D_l = U_L^l D \,. \tag{6}$$

### 3.4 Sequential multilevel data assimilation

The idea of sequential assimilation of spatially distributed data resonates well with the nature of the ML model used. Since coarsening the model is expected to result in weaker nonlinearity [44, 45], utilization of coarse models in the first steps of the DA followed by more non-linear fine models in the next steps can help to gradually zoom in on the solution of the inverse problem. Therefore, in SeMLDA the observation data are upscaled into several levels corresponding to the levels of the ML model. Afterwards, the data are assimilated sequentially starting from the coarsest resolution followed by the finer ones.

#### 3.4.1 Sequential multilevel ensemble smoother

In this section we discuss two versions of the Sequential Multilevel Ensemble Smoother (SMLES). This algorithm draws on the ESMDA algorithm [2], MLHES algorithm [33], and constraints associated with assimilation of linearly dependent data [46].

Initially, based on the available computational resources, the number of simulations performed on each level should be decided. Since the simulations are cheaper on the coarser levels, a higher number of simulations will be performed on those levels. Considering this, a decision is made on the resource allocation. Based on the resource allocation, a sample of $N_1$ ensemble members is generated based on the prior information. Afterwards, at step $l$ the model pertaining to level $l$ is used to assimilate the data transformed to match the resolution of the model forecasts at that level. Running the forward simulator for every realization, $z_{l,j}^{\text{pri}}$, $1 \leq j \leq N_l$, from the prior parameters random vector at level $l$, $Z_l^{\text{pri}}$, we have

$$y_{l,j} = \mathcal{M}_l \left( z_{l,j}^{\text{pri}} \right) \,, \tag{7}$$

where $y_{l,j}$ is a realization from the model forecasts random vector at step $l$, $Y_l$.

In the analysis step, realizations from the random observation data vector at level $l$, $\widetilde{D}_l \sim \mathcal{N}(U_L^l \mathbb{E}(D), \mathrm{C}(\widetilde{D}_l))$ are generated as $\{\widetilde{d}_{l,j}\}_{j=1}^{N_l}$. The form of $\mathrm{C}(\widetilde{D}_l)$ will be discussed in Section 3.4.2. The updated parameter vector of an arbitrary ensemble member is, then, given by

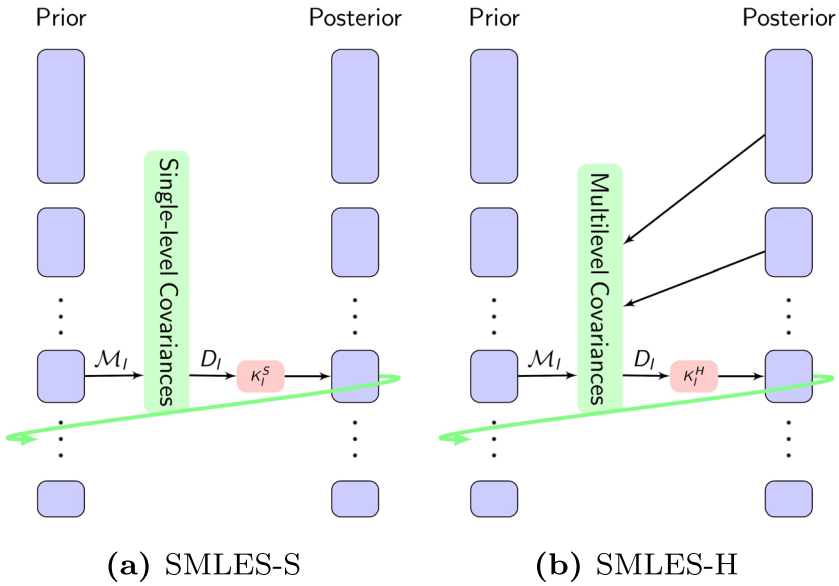
$$z_{l,j}^{\text{pos}} = z_{l,j}^{\text{pri}} + K_l^* \left( \widetilde{d}_{l,j} - y_{l,j} \right) \,. \tag{8}$$

Here, $K_l^*$ is defined generically with * being a wildcard notation, and we introduce two flavors of the algorithm.

The straightforward flavor (SMLES-S) utilizes the data error and model at level $l$ for formulation of the Kalman gain at step $l$. Accordingly, the Kalman gain is given as

$$K_l^S = \mathrm{C}\left( Z_l^{\text{pri}}, Y_l \right) \left( \mathrm{C}(Y_l) + \mathrm{C}(\widetilde{D}_l) \right)^{-1} \,, \tag{9}$$

where $\mathrm{C}(Y_l)$ and $\mathrm{C}\left( Z_l^{\text{pri}}, Y_l \right)$ denote the covariance of $Y_l$ and cross-covariance between $Z_l^{\text{pri}}$ and $Y_l$, respectively.




**(a)** SMLES-S **(b)** SMLES-H

**Fig. 3** Step $l$ of SMLES-S and SMLES-H

Finally, while $l < L$, the prior for step $l+1$ is obtained from the posterior at level $l$ as

$$z_{l+1,j}^{\text{pri}} \coloneqq z_{l,j}^{\text{pos}} , \tag{10}$$

for $1 \leq j \leq N_{l+1}$.

In SMLES-S only the realizations which have been simulated by all the models are considered for the final statistics. With a subtle manipulation of the data-error covariance in Eq. 9 for the realizations at level $l$ which are not simulated by $\mathcal{M}_{l+1}$, the resulting posterior realizations in the linear-Gaussian case would sample correctly from the posterior distribution up to the data at level $l$. The condition that should hold at any level $l$ is explained in Section 3.4.2.

Figure 2 represents the schematic of SMLES-S in comparison with ESMDA and MLHES algorithms. As can be seen from the figure, the ensemble size of SMLES-S shrinks as the level increases due to limited computational resources. Accordingly, at level $L$ the ensemble size becomes small and it may end up in the situation that MLDA wants to avoid at the first place. A possible treatment would be utilization of localization in the finer levels, but that would reduce the robustness of the algorithm. Another alternative is to allow for transfer of information between different levels by utilization of ML statistics. By doing so, a hybrid version of the SMLES algorithm is formulated similar to MLHES and IMLHES. The hybrid Kalman gain can be formulated as

$$K_l^H = \mathrm{C}_l^H(Z,Y)\left(\mathrm{C}_l^H(Y_l) + \mathrm{C}(\widetilde{D}_l)\right)^{-1} , \tag{11}$$

where $\mathrm{C}_l^H(Y_l)$ and $\mathrm{C}_l^H(Z,Y)$ denote the ML covariance of $Y_l$ and the ML covariance between $Z$ and $Y$, respectively. The ML statistics will be discussed in Section 3.4.3.

The SMLES-H algorithm takes advantage of both conditioning all the realizations to the data up to the last level on which they are simulated and the ML statistics. A depiction of the difference between SMLES-S and SMLES-H is presented in Fig. 3. Pseudo-codes of SMLES-S and SMLES-H are presented in Appendices A and B, respectively.

#### 3.4.2 Partially multiple data assimilation condition

For the convenience of the reader, we present the main result from [46] (omitting the derivation). Similar to formulating a condition for assimilating a set of data multiple times such that the updated ensemble will sample correctly from the posterior in the linear-Gaussian case, there exists a condition which should hold if the data are linearly dependent, known as the Partially Multiple Data Assimilation (PMDA) condition [46].

In the problem at hand, the data at any coarser level, $D_c$, is a linear transformation of the data at any finer level, $D_f$, such that for any $\{c, l, f\}$, $1 \leq c < l < f \leq L$, we have

$$U_f^c = U_l^c U_f^l . \tag{12}$$

Hence, we can define the PMDA condition for the data up to any arbitrary level $\mathcal{L}$, $1 \leq \mathcal{L} \leq L$.

Consider the data-error covariance matrix at level $l < \mathcal{L}$ to be given as

$$\mathrm{C}(\widetilde{D}_l) = A_l U_L^l \mathbb{C}(D) {U_L^l}^T {A_l}^T , \tag{13}$$

where $A_l$ is an inflation matrix. Assuming the prior distribution is Gaussian, the forward model $\mathcal{M}_{\mathcal{L}}$ is linear, and for any coarser level $c$ we have

$$\mathcal{M}_c = U_{\mathcal{L}}^c \mathcal{M}_{\mathcal{L}} , \tag{14}$$

the straightforward formulation of the SMLES algorithm, given in Section 3.4.1, samples correctly from the posterior distribution of the parameter random vector up to the data at level $\mathcal{L}$ if the following condition holds,

$$\sum_{l=1}^{\mathcal{L}} {U_{\mathcal{L}}^l}^T \mathrm{C}(\widetilde{D}_l)^{-1} U_{\mathcal{L}}^l = \left(U_L^{\mathcal{L}} \mathbb{C}(D) {U_L^{\mathcal{L}}}^T\right)^{-1} . \tag{15}$$

#### 3.4.3 Multilevel statistics

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

Assuming the model with the highest fidelity, $\mathcal{M}_L$, to be exact, and $\{\mathcal{M}_l\}_{l=1}^{L-1}$ to be approximations to $\mathcal{M}_L$ with a decreasing numerical error, [29] proposed a formulation for ML statistics. In this formulation, inspired by Bayesian Model Averaging, 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 ML 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 [29]. Using this formulation and transformations of the forecast, [33] proposed a formulation of ML statistics for spatially distributed

**Table 1** Shared properties of the reservoir models

| | | | |
|---|---|---|---|
| Fine cell dimensions: | $30 \times 30 \times 30$ ($m^3$) | Porosity: | 0.2 |
| Initial Oil saturation: | 0.85 | Initial Pressure: | 200 bar |

model forecasts. According to this scheme, the ML mean of the model forecast at level $l$ is given as

$$\mathrm{E}_l^H(Y) = \sum_{k=1}^{L} w_k U_k^l \,\mathrm{E}(Y_k)\,, \tag{16}$$

$$\sum_{k=1}^{L} w_k = 1\,, \tag{17}$$

where $\mathrm{E}(Y_k)$ denotes sample mean of the model forecast at level $k$. At step $l$ the model forecasts pertaining to levels finer than $l$ are not available, since the proposed algorithm is sequential. Hence, with a small change in the above-mentioned formulation, the formulation used in this work will be based on levels $1 \le k \le l$,

$$\mathrm{E}_l^H(Y) = \sum_{k=1}^{l} w_k U_k^l \,\mathrm{E}(Y_k)\,, \tag{18}$$

$$\sum_{k=1}^{l} w_k = 1\,. \tag{19}$$

Using the law of total variance, the ML approximation of covariance of the model forecast at level $l$ is formulated as

$$\mathrm{C}_l^H(Y) = \sum_{k=1}^{l} w_k \left\{ \mathrm{C}\left(U_k^l Y_k\right) + \left(\mathrm{E}\left(U_k^l Y_k\right) - \mathrm{E}_l^H(Y)\right) \left(\mathrm{E}\left(U_k^l Y_k\right) - \mathrm{E}_l^H(Y)\right)^T \right\}\,. \tag{20}$$

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

$$\mathrm{C}_l^H(Z,Y) = \sum_{k=1}^{l} w_k \left\{ \mathrm{C}\left(Z_k, U_k^l Y_k\right) + \left(\mathrm{E}(Z_k) - \mathrm{E}^H(Z)\right) \left(\mathrm{E}\left(U_k^l Y_k\right) - \mathrm{E}_l^H(Y)\right)^T \right\}, \tag{21}$$

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

$$\mathrm{E}^H(Z) = \sum_{k=1}^{l} w_k \,\mathrm{E}(Z_k)\,. \tag{22}$$

## 4 Test models

Three different reservoir models are set up for numerical investigations. 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. All the wells in these reservoir models are controlled by their bottom-hole pressure; the injectors at 300 bar, and the producers at 110 bar. 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 consist of two segments. A reservoir flow model is used to predict the state variables, i.e. the pressure and saturation of the reservoir fluids, 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 Darcy's law into the mass conservation equation for each of the fluid phases, resulting in [47]

$$\nabla.\left[\frac{k_{ro}}{\nu_o B_o} k(\nabla p_o - \rho_o g \nabla z)\right] = \frac{\partial}{\partial t}\left(\frac{\phi S_o}{B_o}\right) + q_o\,, \tag{23}$$

$$\nabla.\left[\frac{k_{rw}}{\nu_w B_w} k(\nabla p_w - \rho_w g \nabla z)\right] = \frac{\partial}{\partial t}\left(\frac{\phi S_w}{B_w}\right) + q_w\,, \tag{24}$$

where

$$S_o + S_w = 1\,, \tag{25}$$

$$p_{cow} = p_o - p_w\,. \tag{26}$$

In these equations, $k$ denotes absolute permeability, and $k_{r*}$ denotes the relative permeability of the corresponding phase, while $*$ is a wildcard notation. $k_{r*}$ 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; $\nu_*$, $B_*$, and $\rho_*$ 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 [48]. 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 [48].

As for the petro-elastic segment of the forward model, an in-house model based on standard rock-physics equations [49], [50, Report 1] was used.

### 4.1 Reservoir model I

This model has a 50 × 50 grid, and two wells, one producer (P) at center east and one injector (I) at center

**Table 2** A summary of resource allocation for different runs in experiment I

| | Level 1 $G_1 = 157$ | Level 2 $G_2 = 259$ | Level 3 $G_3 = 685$ | Level 4 $G_4 = 2500$ |
|---|---|---|---|---|
| | $N_1$ | $N_2$ | $N_3$ | $N_4$ |
| ESMDA-REF | – | – | – | 10000 |
| SMLES-S | 951 | 880 | 710 | 412 |
| SMLES-H | 1512 | 1134 | 756 | 378 |
| ESMDA-LOC | – | – | – | 100 |
| IES-REF | – | – | – | 10000 |
| IMLHES | 459 | 438 | 216 | 30 |
| IES-LOC | – | – | – | 100 |

west. This model is designed to evaluate the performances of the different algorithms in parameter estimation of an oil reservoir with relatively long-range correlation in permeability field.

### 4.2 Reservoir model II

This model has a $64 \times 64$ grid and five-spot well pattern, four injectors at the corners and a producer at the center of the field. This model is designed to assess the performances of the different algorithms in history-matching of a field with relatively short-range correlation length.

### 4.3 Reservoir model III

This model has a $70 \times 70$ grid and two wells, an injection well in southwest corner and a production well in the northeast corner. This model has two permeability zones, one with a long-range correlation length and one with a short-range correlation length, and there exist a smooth transition from one zone to another.

## 5 Numerical investigation

Three numerical experiments are conducted, one for each of the three reservoir models discussed in Section 4.

The unknown parameter field in all the experiments is the logarithmic permeability field, which has a different distribution in each of the experiments.

The observation data are two sets of time-lapse bulk-impedance data 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 non-diagonal covariance matrix for the data error, based on isotropic spherical variograms with mean 0. The ranges of the variograms are different for different experiments so that the robustness of the algorithms towards variogram range is assessed. The marginal standard deviation of each observation value is given as

$$\sigma = r \ \max\{|\delta|, \eta\}\,, \tag{27}$$

where $r = 0.1$, $\delta$ is the value of observation data at a certain location, and $\eta$ is a threshold introduced to avoid too much certainty in observation data whose absolute values are very small. This threshold is defined as the $1^{st}$ smallest percentile of the absolute value of the observation data.

In order for comparison of SMLES with standard ESMDA, three algorithms are run; SMLES-S, SMLES-H, and ESMDA with localization (ESMDA-LOC). The gold standard for this comparison is the DA results obtained

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

| | Level 1 $G_1 = 265$ | Level 2 $G_2 = 433$ | Level 3 $G_3 = 1147$ | Level 4 $G_4 = 4096$ |
|---|---|---|---|---|
| | $N_1$ | $N_2$ | $N_3$ | $N_4$ |
| ESMDA-REF | – | – | – | 10000 |
| SMLES-S | 1673 | 1380 | 1343 | 415 |
| SMLES-H | 1496 | 1122 | 748 | 374 |
| ESMDA-LOC | – | – | – | 100 |
| IES-REF | – | – | – | 10000 |
| IMLHES | 644 | 498 | 111 | 40 |
| IES-LOC | – | – | – | 100 |

from vanilla ESMDA with an exceedingly large ensemble (ESMDA-REF). (Obviously, such an ensemble size would be computationally infeasible for a real application.)

In addition, in order for comparison of performance of SMLES with the iterative DA algorithms, the following algorithms are run: Iterative Ensemble Smoother (IES) [1] with localization (IES-LOC) and IMLHES. The gold standard for this comparison is the DA results obtained from vanilla IES with an exceedingly large ensemble (IES-REF).

What we want to estimate is the posterior distribution of the parameters and the model forecasts. The correct estimates would be given by Markov Chain Monte Carlo with an exceedingly large chain length. However, the focus of this work is comparing the novel MLDA algorithms with algorithms of the same class which are widely used in reservoir history matching, i.e. ensemble-based DA algorithms. Hence, ESMDA and IES with exceedingly large ensemble sizes are selected to remove the Monte Carlo errors and serve as the gold standards for comparison.

In order for these comparisons to be fair, there exists the "equal computational power" constraint. 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 relating to simulation of each ensemble member, using forward model $\mathcal{M}_l$, is assumed to be proportional to $G_l^\gamma$, where $G_l$ is the number of the active grid cells in the forward model at level $l$, with $\gamma \in [1.25, 1.5]$, [51]. Here, we take $\gamma = 1.35$. Accordingly, the computational power associated with each algorithm run can be written as

$$\Omega = \sum_{l=1}^{L} n_l G_l^{1.35} \,. \tag{28}$$

where $n_l$ is the total number of simulations using $\mathcal{M}_l$. Using Eq. 28, the computational cost of the algorithm runs are set to be equal.

The iterative algorithms do not always have a fixed computational cost and many iterations are performed to satisfy the convergence criteria without considerable improvements in the objective function. Therefore more computational power (approximately twice the computational power of the other algorithms) was observed for these algorithms (IMLHES and IES-LOC).

Setting a fixed computational cost, there exists $L-1$ degrees of freedom for specification of the $\{N_l\}_{l=1}^{L}$ in the ML algorithms. No optimization was performed for this specification, the only aim pursued was to keep the size of sub-ensembles ascending with decreasing model cost. Several other similar settings that were tried resulted in similar DA outcomes.

The convergence criterion for the iterative algorithms was that improvements in the relative data mismatch should be smaller than 0.0001.

The localization scheme in ESMDA-LOC was based on covariance structure given in [52], spherical variogram, and the DA was performed using subspace inversion method proposed by [53]. As for IES-LOC, the localization scheme was based on covariance structure given in [52] and spherical variogram.

For the SMLES-H and IMLHES, there is a possibility to improve the results by tuning the weights in calculation of ML statistics for specific cases, but here we use the simplest choice–$\{w_l = \frac{1}{\mathcal{L}} \mid 1 \le l \le \mathcal{L}\}$.

Both ESMDA-LOC and ESMDA-REF are run with 6 steps, with $\alpha_i = 6$, $1 \le i \le 6$, in all the steps. Runs with higher number of steps were conducted but no major improvement in the DA results was observed.

As for SMLES-S, the inflation matrices are set as $A_l = L * I_{\zeta_l}$ for $1 \le l < L$, where $I_{\zeta_l}$ is the identity matrix of size $\zeta_l$. For $\mathcal{L} = L$, Eq. 15 is solved to calculate $A_L$. Unlike satisfying the MDA condition, which is computationally trivial, satisfying the PMDA condition is computationally very expensive for real field problems and becomes practically unfeasible. This is due to the presence of the costly inversions in Eq. 15. However, since the PMDA condition only secures correct sampling for linear-Gaussian problems, approximately satisfying this condition may suffice for the real field cases.

Regarding SMLES-H, the ensemble at level $l$ is divided into two sub-ensembles; (I) the realizations which are simulated using $\mathcal{M}_{l+1}$ after the analysis step, $\{z_{l,j}^{\text{pri}} \mid 1 \le j \le N_{l+1}\}$, and (II) those that are not, $\{z_{l,j}^{\text{pri}} \mid N_{l+1} < j \le N_l\}$. For Sub-ensemble (I) the inflation matrix is set as $A_l = L * I_{\zeta_l}$. For Sub-ensemble (II), Eq. 15 is solved for $\mathcal{L} = l$, so that $A_{\mathcal{L}}$ is calculated. This is done to assure

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

| | Level 1 $G_1 = 390$ | Level 2 $G_2 = 1261$ | Level 3 $G_3 = 4900$ |
|---|---|---|---|
| | $N_1$ | $N_2$ | $N_3$ |
| ESMDA-REF | – | – | 10000 |
| SMLES-S | 949 | 925 | 420 |
| SMLES-H | 1266 | 844 | 422 |
| ESMDA-LOC | – | – | 100 |
| IES-REF | – | – | 10000 |
| IMLHES | 768 | 246 | 32 |
| IES-LOC | – | – | 100 |

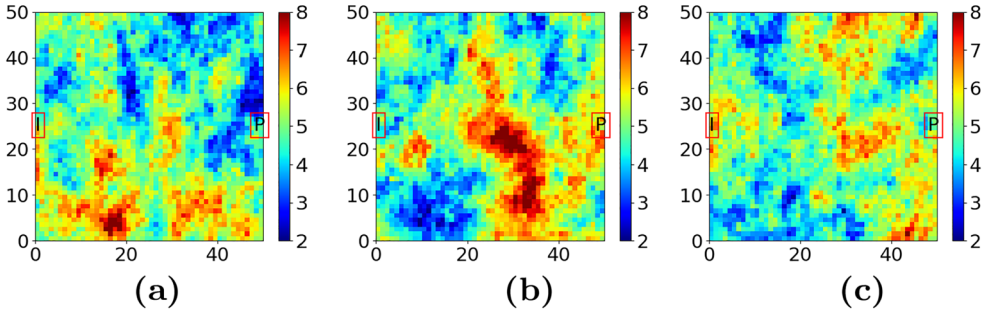



**Fig. 4** Randomly selected realizations from prior distribution of the logarithmic permeability field of Reservoir model I

that the PMDA condition is satisfied up to the last level that each realization is simulated. In order to avoid additional computations, the ML statistics are calculated based on the posterior parameters and their corresponding linear update for the model forecasts.

### 5.1 Experiment I

This experiment is conducted on Reservoir Model I. The ML algorithms have four levels, corresponding to 157, 259, 685, and 2500 grid cells, respectively. A summary of the resource allocation for the different runs carried out in this experiment can be found in Table 2. All the numbers in the table are per assimilation step or iteration, except for SMLES-S and SMLES-H. For these two algorithms the total number of simulations performed at each of the levels are reported. This also holds for Tables 3 and 4 pertaining to Experiment II and Experiment III, respectively.

The observation data for this experiment are generated based on seismic vintages at 2500 and 5000 days after production starts. The range of the variogram used for data-error covariance in this experiment is 15 grid cells.

The prior unknown logarithmic permeability field is based on an exponential variogram with mean and variance constant at 5 and 1, anisotropy angle and anisotropy ratio of 80 degrees and 0.7, 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. The ML algorithms have four levels, corresponding to 265, 433, 1147, 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 3.

The observation data for this experiment are generated based on seismic vintages at 4000 and 8000 days after production starts. The range of the variogram used for data-error covariance in this experiment is 10 grid cells.

The prior unknown logarithmic permeability field is based on a spherical variogram with mean and variance constant at 5 and 1, the anisotropy angle and anisotropy ratio of $-30$ degrees and 0.7, and range 10 grid cells. Randomly selected realizations from this logarithmic permeability field can be found in Fig. 5.

### 5.3 Experiment III

This experiment is conducted on Reservoir Model III. The ML algorithms have three levels, corresponding to 390, 1261, and 4900 grid cells, respectively. A summary of the

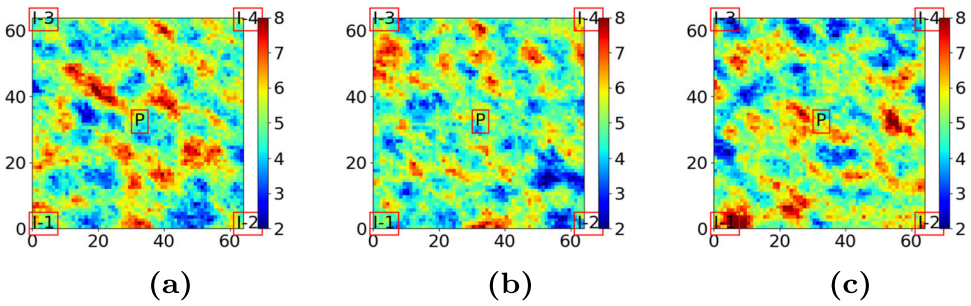



**Fig. 5** Randomly selected realizations from prior distribution of the logarithmic permeability field of Reservoir model II

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 4000 and 8000 days after production starts. The range of the variogram used for data-error covariance in this experiment is 5 grid cells.

The unknown logarithmic permeability field is based on two different variograms in two zones of the field, one encompassing the northeastern part of the field and one encompassing the southwestern part of the field with a smooth transition between them. The details about the variograms based on which the distribution of the unknown parameters are defined can be found in Table 5. Randomly selected realizations from this logarithmic permeability field can be found in Fig. 6.

## 6 Numerical results

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

As for qualitative analysis, firstly, mean and variance of the posterior parameter fields obtained by different algorithms are compared to the reference cases. Additionally, since the model forecasts are in different resolutions for different levels of the ML algorithms, comparison of the posterior model forecasts as such is not a possibility. Instead, we run fine-scale simulations of the posterior ensemble for all algorithms and then plot the mean and variance of the model forecasts for all the algorithms and compare them. The model forecasts of the second vintage are presented for all the experiments.

As for SMLES-H and IMLHES, the simplest formulation is chosen for computation of posterior statistics ($\mathrm{E}(Z^{\mathrm{pos}})$, $\mathrm{Var}(Z^{\mathrm{pos}})$, $\mathrm{E}(Y^{\mathrm{pos}})$, and $\mathrm{Var}(Y^{\mathrm{pos}})$). All posterior ensemble members at different levels are united into one ensemble and the empirical mean and variance are calculated.

The conventional DA algorithms (ESMDA-LOC and IES-LOC) were tested with several ranges for localization (critical distances), and the best results are presented for each of the experiments.

The quantitative analysis is performed using a measure suggested by [54] for comparison of DA results obtained by different schemes with a reference case. Consider $\nu$ to be a vector of interest for quantitative analysis, e.g. parameter estimate vector or model forecast vector. The following two metrics are computed for each of the algorithm runs [54],

$$\epsilon_{\mathrm{Mean}} = \frac{||\left(\mathrm{Mean}(\nu_*) - \mathrm{Mean}(\nu^{pri})\right) - \left(\mathrm{Mean}(\nu^{pos}) - \mathrm{Mean}(\nu^{pri})\right)||_2}{||\left(\mathrm{Mean}(\nu^{pos}) - \mathrm{Mean}(\nu^{pri})\right)||_2}, \tag{29}$$

$$\epsilon_{\mathrm{Var}} = \frac{||\mathrm{Var}(\nu_*) - \mathrm{Var}(\nu^{pos})||_2}{||\mathrm{Var}(\nu^{pos})||_2}, \tag{30}$$

where Mean and Var represent the empirical mean and variance for different cases; superscripts pri and pos denote the prior and the reference posterior, respectively; the subscript $*$ is a wildcard notation for the algorithm of interest; and the distance is measured in 2-norm. Here, the reference posterior is calculated based on the results obtained by ESMDA-REF for all the experiments.

When estimation of $\nu$ is concerned, the $\epsilon_{\mathrm{Mean}}$ metric has the property that it will be close to 0 for the algorithms that perform similar to the reference and will be close to 1 for the algorithms whose estimate of $\nu$ are similar to the prior estimate. Similarly, the $\epsilon_{\mathrm{Var}}$ metric calculates the difference between the variance obtained by the algorithm of interest and the reference variance normalized by the norm of the reference variance, meaning that smaller values are preferred.

We will compute these two metrics for both the posterior parameter estimates and the second vintage of the posterior model forecasts.

**Table 5** The two variograms of Reservoir Model III

| | Variance | Mean | Range | Ratio | Angle | Type |
|---|---|---|---|---|---|---|
| Variogram 1 | 1 | 5 | 30 | 0.7 | −30 | cubic |
| Variogram 2 | 1 | 5 | 10 | 0.4 | −70 | cubic |

### 6.1 Results of experiment I

Visual analysis of the mean permeability fields in Fig. 7 shows that the SMLES-S and SMLES-H results are more similar to the ESMDA-REF result than the ESMDA-LOC result is. It also shows that IMLHES performs more similar to IES-REF than IES-LOC does. There does not seem to be any considerable advantage in performance of a specific ML algorithm when the mean permeability fields obtained by the ML algorithms are compared with the IES-REF result. However, due to convergence issues mentioned above, the SiMLDA algorithm used more computational

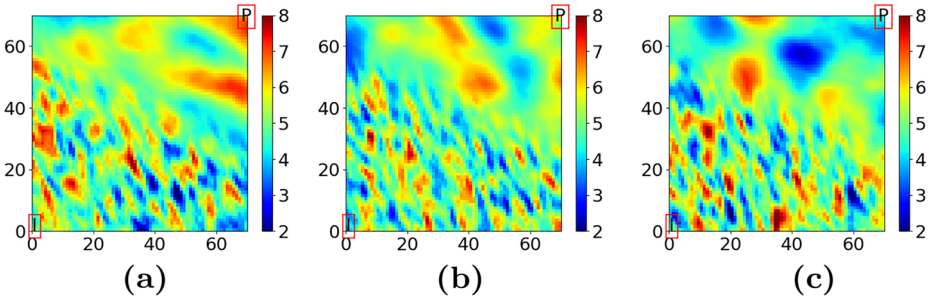



**Fig. 6** Randomly selected realizations from prior distribution of the logarithmic permeability field of Reservoir model III

resources than SeMLDA algorithms. This holds for all the experiments.

Checking the variance fields in Fig. 8 confirms that both SMLES-S and SMLES-H perform better than ESMDA-LOC. It further confirms that IMLHES performs more similar to IES-REF than IES-LOC does. However, in this figure it is evident that SMLES-S underestimates the uncertainty in the posterior parameters while SMLES-H and IMLHES overestimate it. No indication of superiority of either SMLES-H or IMLHES over each other is noticeable in the variance fields, and both perform slightly better than SMLES-S. It is worth noting that, changing the color scale of the plot of the variance field for SMLES-S (denoted by SMLES-S* in the figure), shows that, in spite of underestimation of uncertainty, the structure of variance field is predicted accurately by this algorithm.

Superiority of performance of ML algorithms over the performance of the conventional DA algorithms is further confirmed by checking the statistics of the model forecasts in Figs. 9 and 10. There is no clear indication of advantage of either SMLES-H or IMLHES over each other. However, SMLES-S performs better than both of them particularly in estimating the variance field of the model forecasts.

Quantitative measures given in Tables 6 and 7 also confirm that the ML algorithms generally perform more similar to ESMDA-REF than the conventional DA algorithms. Among the ML algorithms, SMLES-H performs slightly better in estimation of both the mean posterior parameters and the mean model forecasts, while IMLHES performs slightly better in estimation of the variance of the posterior parameter field, and SMLES-S performs significantly better in estimation of the variance of the posterior model forecasts.

The ESMDA-LOC and IES-LOC results presented here are based on the localization range of 40 grid cells (1200 meters).

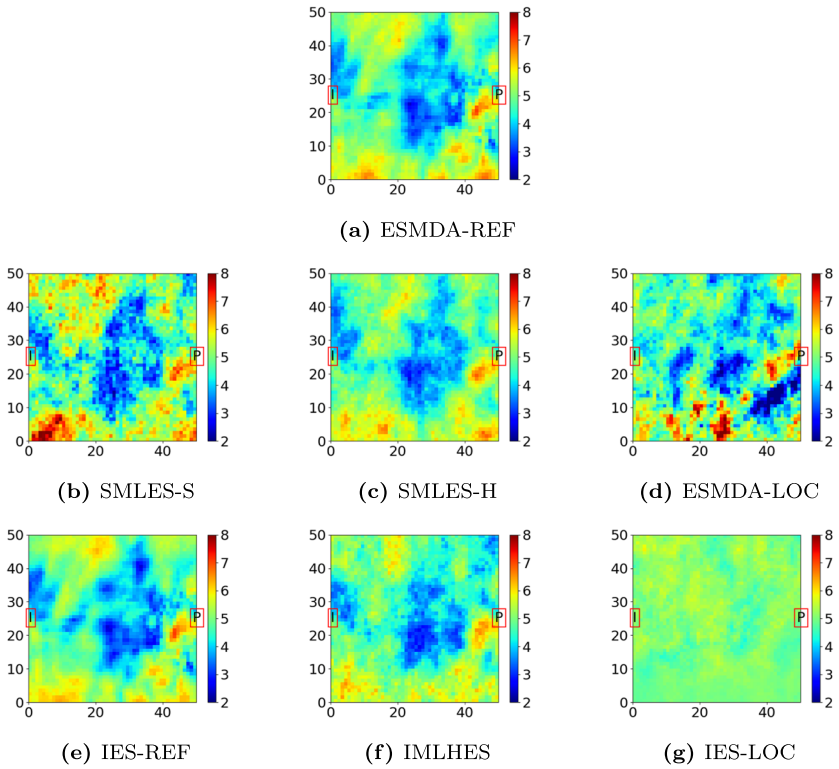



**Fig. 7** Experiment I–Mean posterior logarithmic permeability field

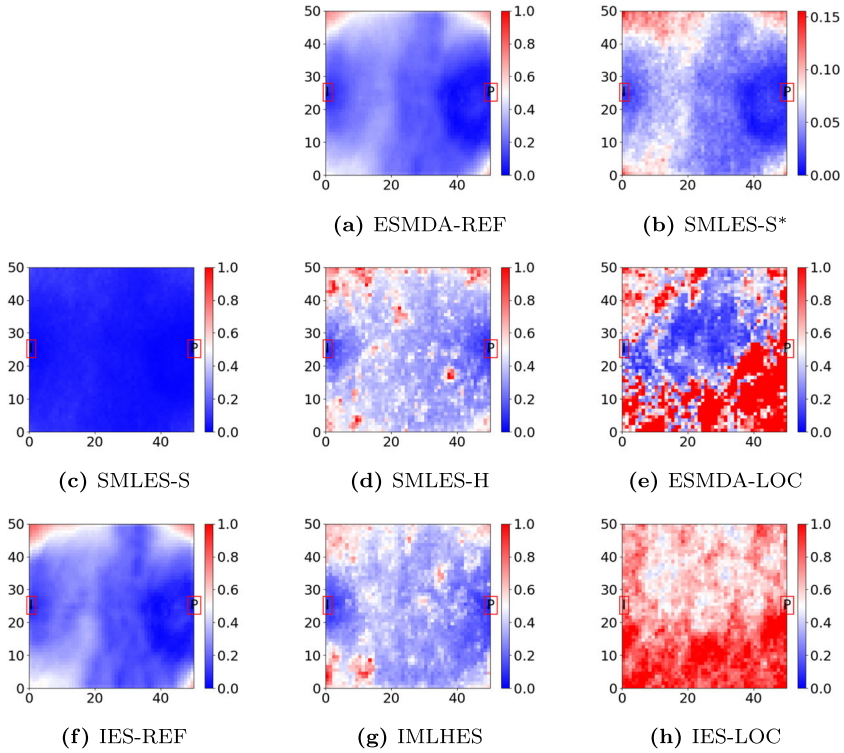



**Fig. 8** Experiment I–Variance of posterior logarithmic permeability field. The only difference between SMLES-S and SMLES-S* is in their scales

## 6.2 Results of experiment II

Visual analysis of the mean permeability fields in Fig. 11 shows that the SMLES-S and SMLES-H results are more similar to the ESMDA-REF result than the ESMDA-LOC result is. It also shows that IMLHES performs more similar to IES-REF than IES-LOC does. There does not seem to be any considerable advantage in performance of a specific ML algorithm when the mean permeability fields obtained by ML algorithms are compared with the IES-REF result.

Checking the variance fields in Fig. 12 confirms that both SMLES-S and SMLES-H perform better than ESMDA-LOC. It further confirms that IMLHES performs more similar to IES-REF than IES-LOC does. However, in this figure it is evident that SMLES-S underestimates the uncertainty in the posterior parameters while SMLES-H and IMLHES overestimate it. No indication of superiority of either SMLES-H or IMLHES over each other is noticeable in the variance fields, and both perform slightly better than SMLES-S. As for Experiment I, changing the color scale of the plot of the variance field for SMLES-S (denoted by SMLES-S* in the figure), shows that, in spite of underestimation of uncertainty, the structure of variance field is predicted accurately by this algorithm.

Superiority of performance of the ML algorithms over the performance of the conventional DA algorithms is further confirmed by checking the statistics of the model forecasts in Figs. 13 and 14. There is no clear indication of advantage of either SMLES-H or IMLHES over each other. However, SMLES-S performs better than both of them particularly in estimating the variance field of the model forecasts.

Quantitative measures given in Tables 6 and 7 also confirm that the ML algorithms generally perform more similar to ESMDA-REF than the conventional DA algorithms. All the MLDA algorithms perform reasonably similar in estimation of the mean of posterior parameters, while IMLHES and SMLES-H perform slightly better in estimation of the variance of the posterior parameter field. As for estimation of the statistics of the posterior model forecasts, SMLES-S performs best, and its superiority is most evident in estimation of the variance of the posterior model forecasts.

The ESMDA-LOC and IES-LOC results presented here are based on the localization range of 60 grid cells (1800 meters).

## 6.3 Results of experiment III

Visual analysis of the mean permeability fields in Fig. 15 shows that the SMLES-S and SMLES-H results are more similar to the ESMDA-REF result than the ESMDA-LOC result is. It also shows that IMLHES performs more similar to IES-REF than IES-LOC does. There does not seem to be any considerable advantage in performance of a specific ML algorithm when the mean permeability fields obtained by ML algorithms are compared with the IES-REF result.

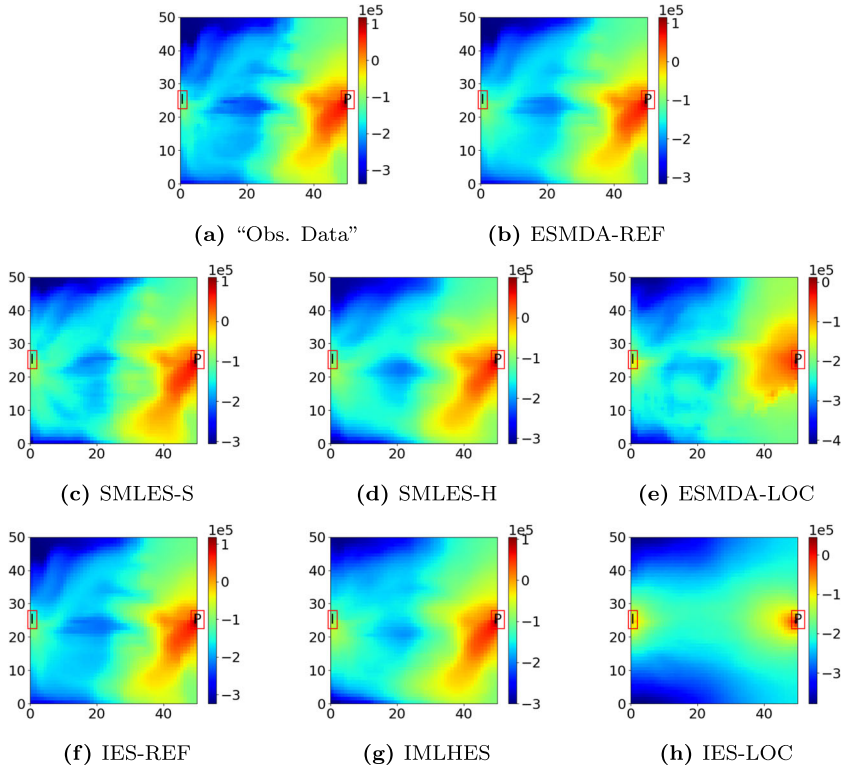



**Fig. 9** Experiment I–Mean of posterior time-lapse bulk impedance field ($\frac{m}{s}\frac{kg}{m^3}$) in comparison with observation data in the second vintage

Checking the variance fields in Fig. 16 confirms that both SMLES-S and SMLES-H perform better than ESMDA-LOC. It further confirms that IMLHES performs more similar to IES-REF than IES-LOC does. However, in this figure it is evident that SMLES-S underestimates the uncertainty in the posterior parameters while SMLES-H and IMLHES overestimate it. No indication of superiority of either SMLES-H or IMLHES over each other is noticeable in the variance fields, and both perform slightly better than SMLES-S. As for Experiments I and II, changing the color scale of the plot of the variance field for SMLES-S (denoted by SMLES-S* in the figure), shows that, in spite of underestimation of uncertainty, the structure of variance field is predicted accurately by this algorithm.

Superiority of performance of the ML algorithms over the performance of the conventional DA algorithms is further confirmed by checking the statistics of the model forecasts in Figs. 17 and 18. There is no clear indication of advantage of either SMLES-H or IMLHES over each other. However, SMLES-S performs better than both of them particularly in estimating the variance field of the model forecasts.

Quantitative measures given in Tables 6 and 7 also confirm that the ML algorithms generally perform more similar to ESMDA-REF than the conventional DA algorithms. All the MLDA algorithms perform reasonably similar in estimation of the mean of posterior parameters, while IMLHES and SMLES-H perform slightly better in estimation of the variance of the posterior parameter field. As for estimation of the statistics of the posterior model forecasts, SMLES-S performs best, and its superiority is most evident in estimation of the variance of the posterior model forecasts. SMLES-H performs slightly better than IMLHES in this regard.

The ESMDA-LOC and IES-LOC results presented here are based on the localization range of 60 grid cells (1800 meters).

## 7 Summary and conclusions

In this work, two variants of a novel sequential MLDA algorithm, SMLES-S and SMLES-H, were introduced. In addition, performances of these algorithms were assessed in comparison with two conventional DA algorithms and a simultaneous MLDA algorithm. In doing so, three experiments were performed on three reservoir models. The three experiments were designed such that the performance of the algorithms were evaluated in different settings for the prior parameter fields (different variogram types; different anisotropies; and various correlation lengths including long-range correlation, short-range correlation, and mixture of long-range and short-range correlations) and different ranges for the variograms used for the data-error covariance. Each of the experiments consisted

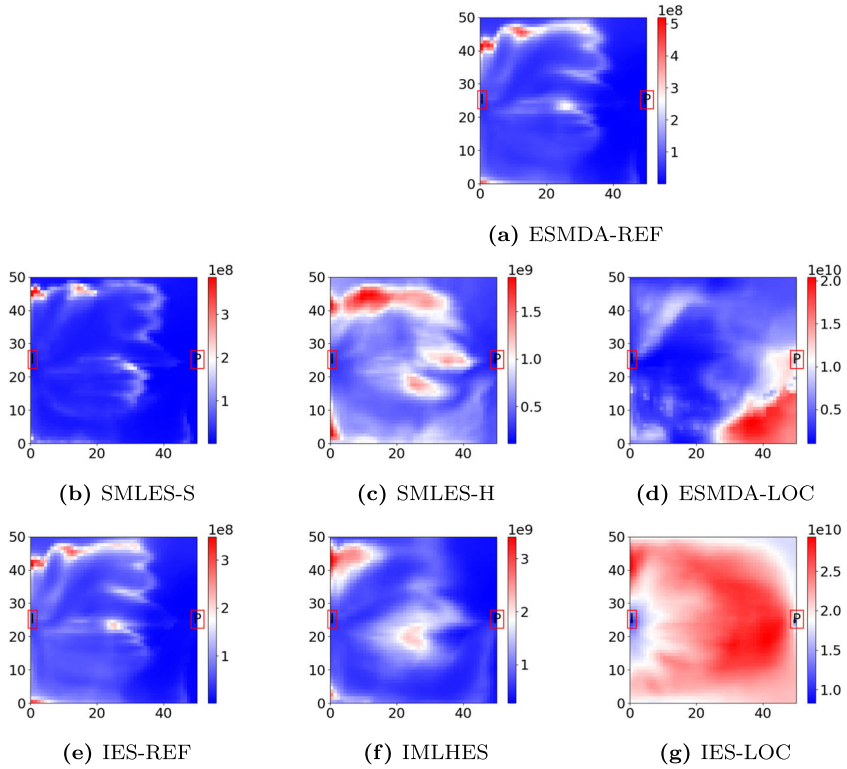



**Fig. 10** Experiment I–Variance of posterior time-lapse bulk impedance field $\left(\left(\frac{m}{s}\frac{kg}{m^3}\right)^2\right)$, in the second vintage. Note that the color bars have very different scales

of seven algorithm runs: SMLES-S, SMLES-H, ESMDA with localization (ESMDA-LOC), vanilla ESMDA with an exceedingly large ensemble (ESMDA-REF), an iterative MLDA algorithm (IMLHES), iterative ensemble smoother with localization (IES-LOC), and vanilla iterative ensemble smoother with an exceedingly large ensemble (IES-REF). Results of the experiments were assessed both qualitatively and quantitatively.

In order for qualitative evaluation of the numerical results, firstly, the mean and the variance of posterior parameter fields were generated and assessed visually. The relative performances of the different methods were similar for all three experiments. The assessments showed that both SMLES-S and SMLES-H performed more similar to ESMDA-REF than ESMDA-LOC did in estimation of the posterior parameter mean field. Regarding estimation of the variance fields, SMLES-H overestimated the variance while SMLES-S underestimated it. The superiority of performance of both SeMLDA algorithms over ESMDA-LOC was evident, also for the variance fields. Among the iterative algorithms, IMLHES performed more similar to IES-REF than IES-LOC did. There was no indication of superior performance of either SMLES-H or IMLHES over each other in any of the experiment when their performances were compared to IES-REF results. However, IMLHES used more computational resources than either of the SeMLDA algorithms. Both IMLHES and SMLES-H performed slightly better than SMLES-S.

Additionally, fine-scale simulations were run for all the posterior ensembles obtained by the different algorithms in

**Table 6** Summary of quantitative analysis of the experiments for posterior parameter estimates

| | Experiment I | | Experiment II | | Experiment III | |
|---|---|---|---|---|---|---|
| | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ |
| ESMDA-REF | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SMLES-S | 0.76 | 0.79 | 0.76 | 0.70 | 0.55 | 0.51 |
| SMLES-H | 0.51 | 0.67 | 0.87 | 0.39 | 0.54 | 0.21 |
| ESMDA-LOC | 1.34 | 9.09 | 1.21 | 0.65 | 1.07 | 0.95 |
| IES-REF | 0.26 | 0.09 | 0.34 | 0.11 | 0.22 | 0.08 |
| IMLHES | 0.60 | 0.58 | 0.80 | 0.36 | 0.61 | 0.20 |
| IES-LOC | 1.05 | 2.00 | 0.97 | 0.63 | 1.08 | 0.64 |

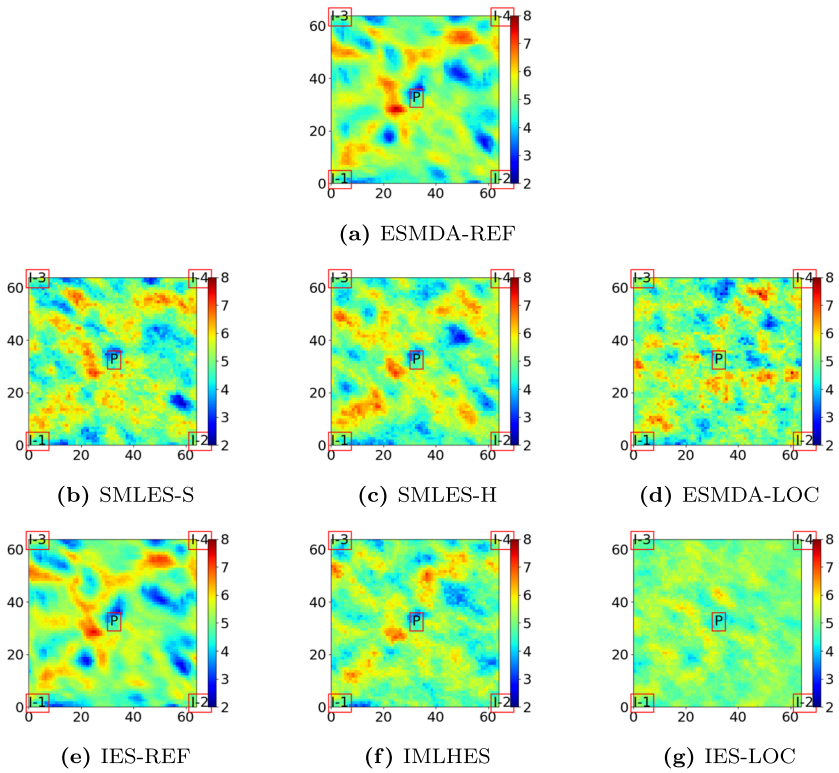
**Table 7** Summary of quantitative analysis of the experiments for posterior model forecasts

| | Experiment I | | Experiment II | | Experiment III | |
|---|---|---|---|---|---|---|
| | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ | $\epsilon_{\text{Mean}}$ | $\epsilon_{\text{Var}}$ |
| ESMDA-REF | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| SMLES-S | 0.16 | 0.58 | 0.076 | 0.46 | 0.037 | 0.30 |
| SMLES-H | 0.12 | 8.23 | 0.13 | 4.28 | 0.07 | 14.23 |
| ESMDA-LOC | 1.26 | 84.48 | 0.16 | 4.11 | 0.075 | 4.73 |
| IES-REF | 0.09 | 0.21 | 0.029 | 0.47 | 0.011 | 0.39 |
| IMLHES | 0.16 | 11.28 | 0.12 | 5.80 | 0.19 | 19.88 |
| IES-LOC | 1.27 | 316.33 | 0.21 | 13.12 | 0.31 | 97.79 |

all the experiments. Plots of the mean and the variance of model forecasts from the different algorithms were compared to each other. Visual analysis of these plots showed that in all the experiments the ML algorithms performed better than their conventional DA counterparts. Among the ML algorithms, SMLES-S consistently performed better than both SMLES-H and IMLHES in estimation of the variance of the posterior model forecasts. Either of the other two MLDA algorithms did not have a clear advantage over each other.

Two metrics were adopted for quantitative comparison of the DA results obtained by different algorithms for estimation of both mean and variance of the posterior parameters and model forecasts. The metrics indicated that the ML algorithms generally performed better than the conventional DA algorithms in estimation of both mean and variance of the posterior parameters. They also indicated that SMLES-H and IMLHES performed slightly better than SMLES-S in estimation of the variance of the posterior parameters, and that all the MLDA algorithms performed better than IES-LOC in estimation of mean and variance of the posterior model forecasts. SMLES-S also performed consistently superior to ESMDA-LOC in estimation of mean and variance of the posterior model forecasts, while this was not observed for IMLHES and SMLES-H. Among the ML algorithms, SMLES-S clearly performed best when it came to estimation of the variance of the model forecasts. The other two algorithms did not consistently perform better than one another.

There were significant differences between the results from the MLDA algorithms and the results from the conventional DA algorithms in all the experiments. Simultaneous




(a) ESMDA-REF

(b) SMLES-S (c) SMLES-H (d) ESMDA-LOC

(e) IES-REF (f) IMLHES (g) IES-LOC

**Fig. 11** Experiment II–Mean posterior logarithmic permeability field

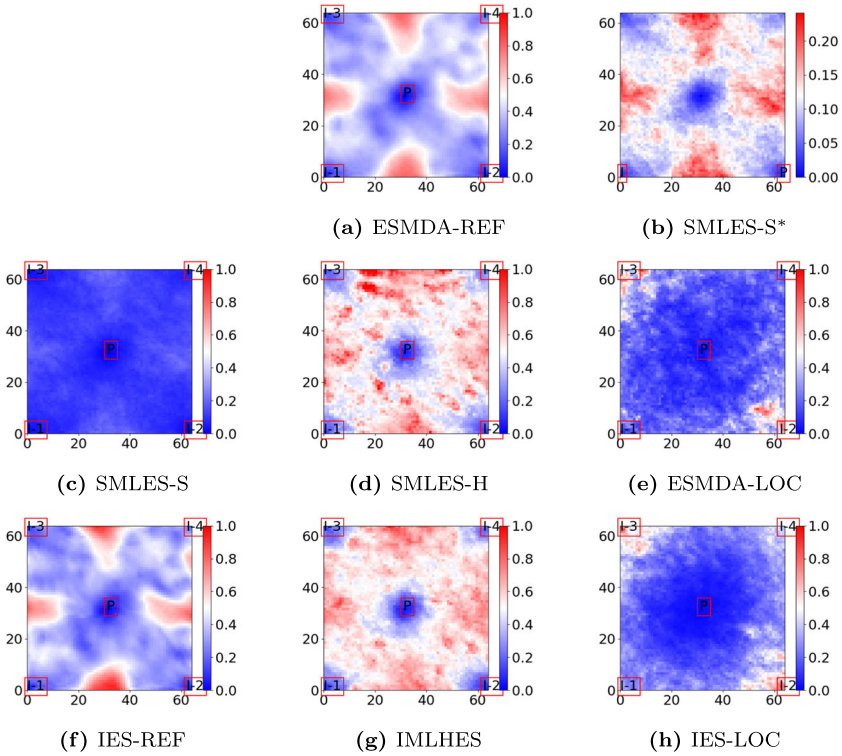



**Fig. 12** Experiment II–Variance of posterior logarithmic permeability field. The only difference between SMLES-S and SMLES-S* is in their scales

assimilation of large amounts of data into ensembles of small size partly explains the under-performance of the conventional algorithms. In the case of inverted seismic data we noticed true long-range correlations between the data and parameters. Regularization of the Kalman gain using distance-based localization, which disregards these true correlations and distorts the update formula, is another cause for the significant difference between the results.

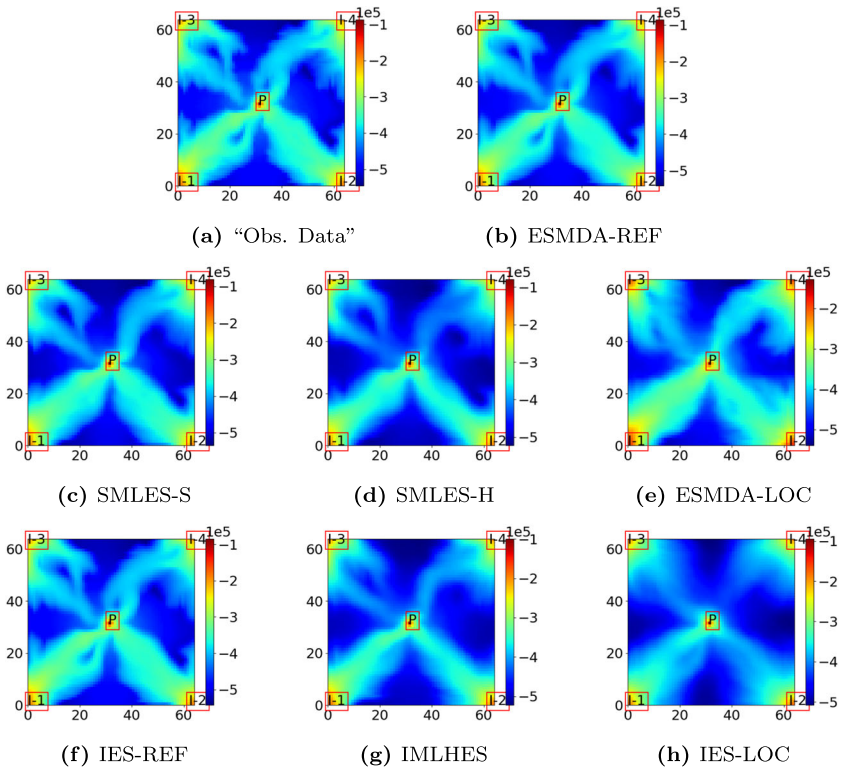
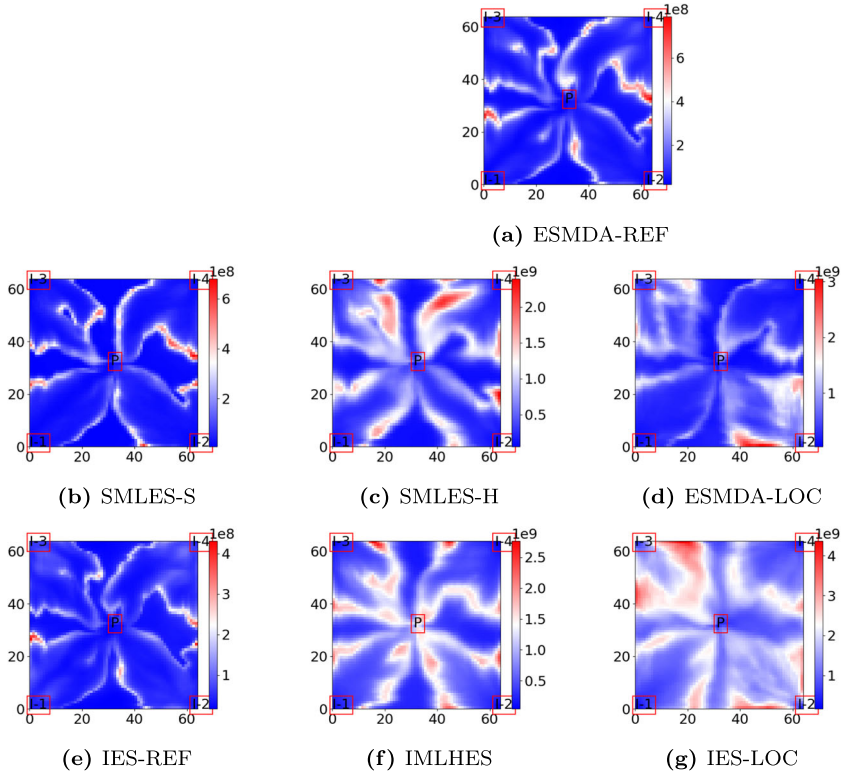



**Fig. 13** Experiment II–Mean of posterior time-lapse bulk impedance field ($\frac{m}{s}\frac{kg}{m^3}$) in comparison with observation data in the second vintage

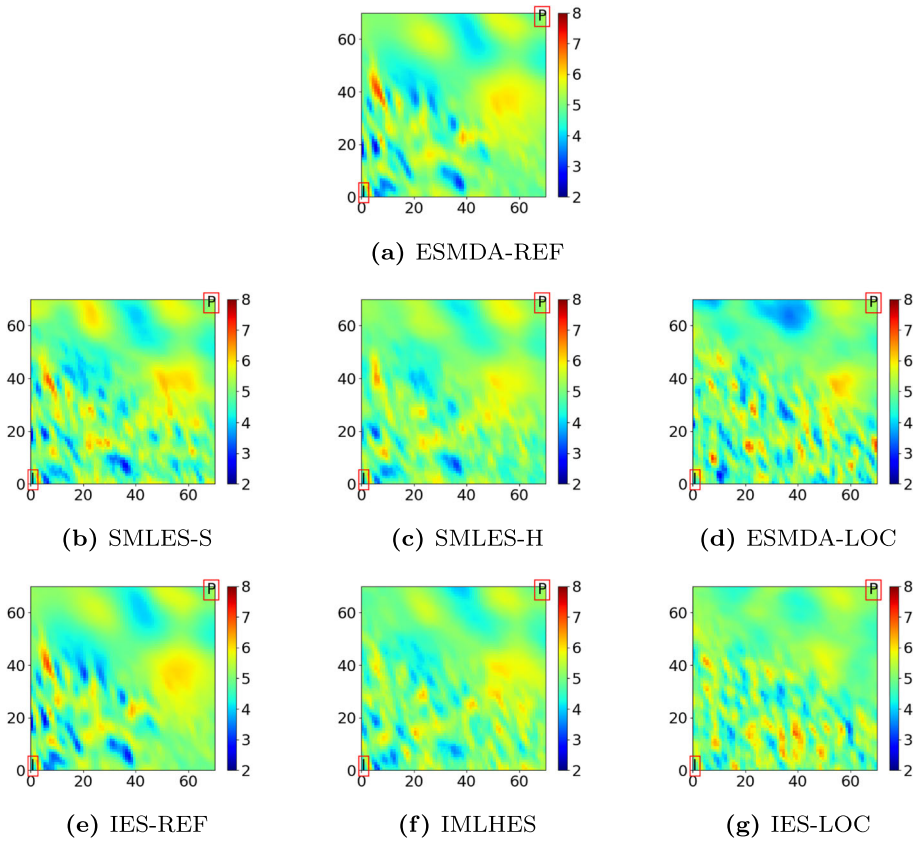


(a) ESMDA-REF

(b) SMLES-S (c) SMLES-H (d) ESMDA-LOC

(e) IES-REF (f) IMLHES (g) IES-LOC

**Fig. 14** Experiment II–Variance of posterior time-lapse bulk impedance field $\left(\left(\frac{m}{s}\frac{kg}{m^3}\right)^2\right)$, in the second vintage. Note that the color bars have very different scales



(a) ESMDA-REF

(b) SMLES-S (c) SMLES-H (d) ESMDA-LOC

(e) IES-REF (f) IMLHES (g) IES-LOC

**Fig. 15** Experiment III–Mean posterior logarithmic permeability field

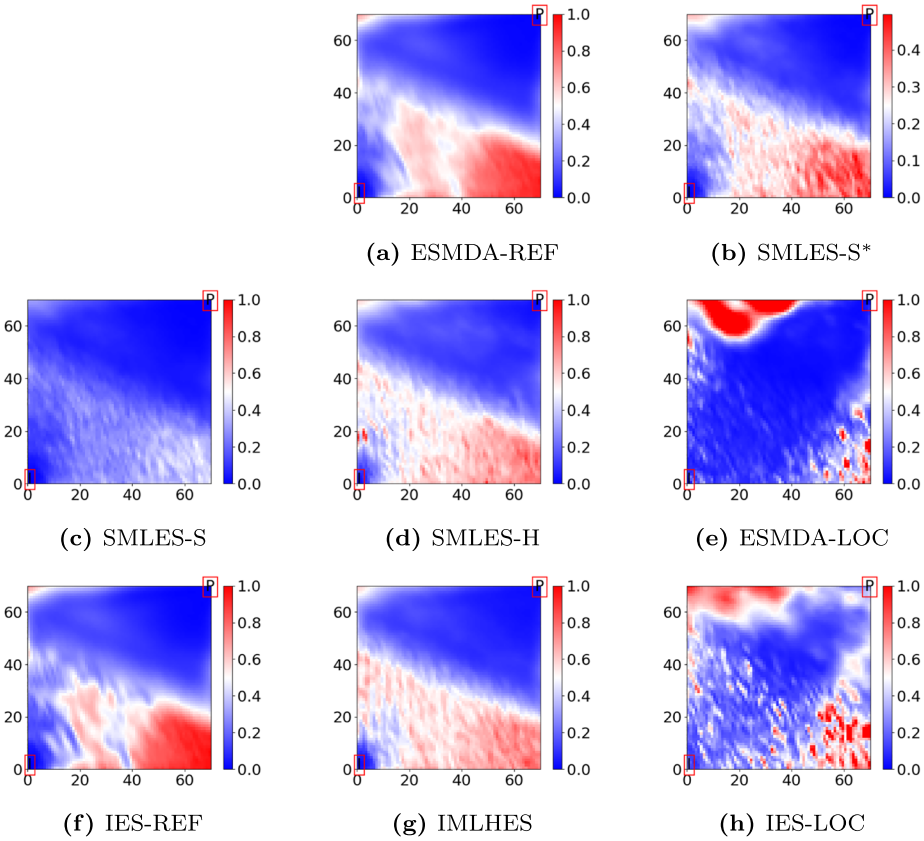


**Fig. 16** Experiment III–Variance of posterior logarithmic permeability field. The only difference between SMLES-S and SMLES-S* is in their scales

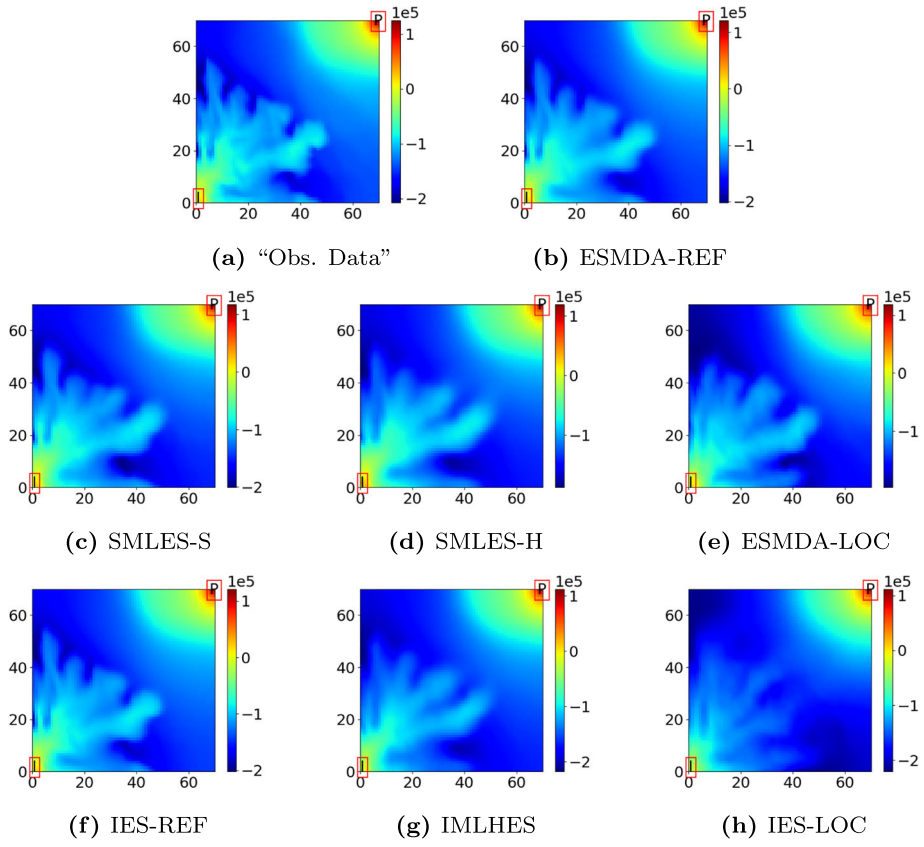


**Fig. 17** Experiment III–Mean of posterior time-lapse bulk impedance field $\left(\frac{m}{s}\frac{kg}{m^3}\right)$ in comparison with observation data in the second vintage

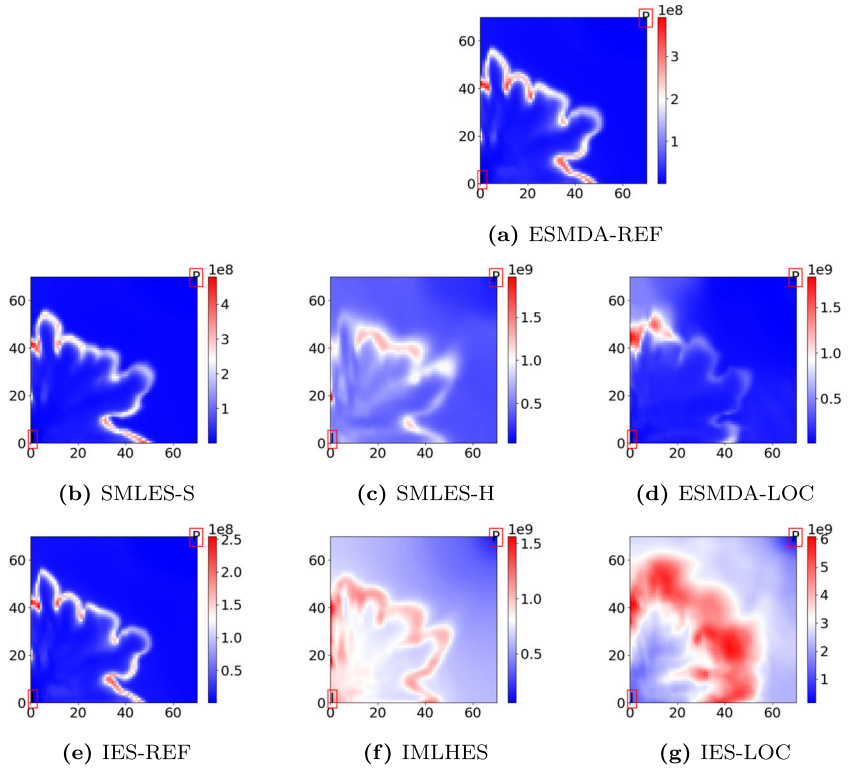



**Fig. 18** Experiment III–Variance of posterior time-lapse bulk impedance field $\left(\left(\frac{m}{s}\frac{kg}{m^3}\right)^2\right)$, in the second vintage. Note that the color bars have very different scales

Even though the results obtained from the experiments did not show any clear indication of superiority of SeMLDA over SiMLDA, they suggest that similar quality of DA results can be obtained by SeMLDA using a fixed and smaller computational power compared to SiMLDA.

There were several issues in this work that can be further investigated. As satisfying the PMDA condition is not a possibility in real field cases, devising robust techniques for approximately satisfying this condition can be studied. In order for optimization of both SeMLDA algorithms several of their characteristics can be further researched, e.g. the optimal extent of coarsening the grid, the number of levels and allocation of resources between them, the weights for different levels in the ML statistics, and the formulation of posterior statistics, to name a few. Additionally, generalizations of SMLES algorithms can also be studied, e.g. by assimilating the data more than once in some of the levels using more inflated data-error covariance matrices. Finally, as realistic reservoir cases are more complex than the fields investigated in this work, extensive coarsening of the grid may result in large numerical error and model bias. This limitation, and increase in the dimensionality of the parameters, may call for combination of localization and the proposed MLDA algorithms.

## Appendix A: Algorithm. Sequential multilevel ensemble smoother–straightforward formulation

Define $\{\mathcal{M}_l\}_{l=1}^{L}$
Define $\{A_l\}_{l=1}^{L}$
**for** $j = 1, ..., N_1$ **do**
  $z_{1,j}^{\text{pri}} \leftarrow \mathcal{N}\left(\mathrm{E}(Z^{\text{pri}}), \mathrm{C}(Z^{\text{pri}})\right)$
**end for**
**for** $l = 1, ..., L$ **do**
  **for** $j = 1, ..., N_l$ **do**
    $y_{l,j} \leftarrow \mathcal{M}_l\left(z_{l,j}^{\text{pri}}\right)$
  **end for**
  $\mathrm{C}(\widetilde{D}_l) \leftarrow A_l U_L^l \mathbb{C}(D) {U_L^l}^T {A_l}^T$
  $K_l^S \leftarrow \mathrm{C}\left(Z_l^{\text{pri}}, Y_l\right)\left(\mathrm{C}(Y_l) + \mathrm{C}(\widetilde{D}_l)\right)^{-1}$
  **for** $j = 1, ..., N_l$ **do**
    $\widetilde{d}_{l,j} \leftarrow \mathcal{N}\left(U_L^l \mathbb{E}(D), \mathrm{C}(\widetilde{D}_l)\right)$
    $z_{l,j}^{\text{pos}} \leftarrow z_{l,j}^{\text{pri}} + K_l^S\left(\widetilde{d}_{l,j} - y_{l,j}\right)$
  **end for**
  **for** $j = 1, ..., N_{l+1}$ **do**
    $z_{l+1,j}^{\text{pri}} \leftarrow z_{l,j}^{\text{pos}}$
  **end for**
**end for**

**Algorithm 1** SMLES–straighforward formulation.

## Appendix B: Algorithm. Sequential multilevel ensemble smoother–hybrid formulation

Define $\{\mathcal{M}_l\}_{l=1}^{L}$
Define $\{A_l\}_{l=1}^{L}$
Define $\{w_l\}_{l=1}^{L}$
**for** $j = 1, ..., N_1$ **do**
  $z_{1,j}^{\text{pri}} \leftarrow \mathcal{N}\left(\mathrm{E}\left(Z^{\text{pri}}\right), \mathrm{C}\left(Z^{\text{pri}}\right)\right)$
**end for**
**for** $\mathcal{L} = 1, ..., L$ **do**
  **for** $j = 1, ..., N_{\mathcal{L}}$ **do**
    $y_{\mathcal{L},j} \leftarrow \mathcal{M}_{\mathcal{L}}\left(z_{\mathcal{L},j}^{\text{pri}}\right)$
  **end for**
  $\mathrm{E}(Z_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-1}\sum_{j=1}^{N_{\mathcal{L}}} z_{\mathcal{L},j}^{\text{pri}}$
  $\mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(\mathrm{Z}) \leftarrow \left(\sum_{l=1}^{\mathcal{L}} w_l \mathrm{E}(Z_l)\right) / \sum_{l=1}^{\mathcal{L}} w_l$
  $\mathrm{E}(Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-1}\sum_{j=1}^{N_{\mathcal{L}}} y_{\mathcal{L},j}$
  $\mathrm{C}(Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-1}\sum_{j=1}^{N_{\mathcal{L}}} (y_{\mathcal{L},j} - \mathrm{E}(Y_{\mathcal{L}}))(y_{\mathcal{L},j} - \mathrm{E}(Y_{\mathcal{L}}))^T$
  $\mathrm{C}(Z, Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-1}\sum_{j=1}^{N_{\mathcal{L}}} (z_{\mathcal{L},j}^{\text{pri}} - \mathrm{E}(Z_{\mathcal{L}}))(y_{\mathcal{L},j} - \mathrm{E}(Y_{\mathcal{L}}))^T$
  $\mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(Y) \leftarrow \left(\sum_{l=1}^{\mathcal{L}} w_l U_l^{\mathcal{L}} \mathrm{E}(Y_l)\right) / \sum_{l=1}^{\mathcal{L}} w_l$
  $\mathrm{C}_{\mathcal{L}}^{\mathrm{H}}(Z, Y) \leftarrow \sum_{l=1}^{\mathcal{L}} w_l \Big\{ \mathrm{C}(Z, Y_l) U_l^{\mathcal{L}^T} + \left(\mathrm{E}(Z_l) - \mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(\mathrm{Z})\right) \left(U_l^{\mathcal{L}} \mathrm{E}(Y_l) - \mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(Y)\right)^T \Big\} / \sum_{l=1}^{\mathcal{L}} w_l$
  $\mathrm{C}_{\mathcal{L}}^{\mathrm{H}}(Y) \leftarrow \sum_{l=1}^{\mathcal{L}} w_l \Big\{ U_l^{\mathcal{L}} \mathrm{C}(Y_l) U_l^{\mathcal{L}^T} + \left(U_l^{\mathcal{L}} \mathrm{E}(Y_l) - \mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(Y)\right) \left(U_l^{\mathcal{L}} \mathrm{E}(Y_l) - \mathrm{E}_{\mathcal{L}}^{\mathrm{H}}(Y)\right)^T \Big\} / \sum_{l=1}^{\mathcal{L}} w_l$
  **for** $j = N_1, ..., N_{\mathcal{L}}$ **do**
    **if** $j \leq N_{\mathcal{L}+1}$ **then**
      $\mathrm{C}(\widetilde{D}_{\mathcal{L}}) \leftarrow A_{\mathcal{L}} U_L^{\mathcal{L}} \mathbb{C}(D) U_L^{\mathcal{L}^T} A_{\mathcal{L}}^T$
    **else**
      $\mathrm{C}(\widetilde{D}_{\mathcal{L}}) \leftarrow \left(\left(U_L^{\mathcal{L}} \mathbb{C}(D) U_L^{\mathcal{L}^T}\right)^{-1} - \sum_{l=1}^{\mathcal{L}-1} U_{\mathcal{L}}^{l\,T} \mathrm{C}(\widetilde{D}_l)^{-1} U_{\mathcal{L}}^{l}\right)^{-1}$
    **end if**
    $K_{\mathcal{L}}^{H} \leftarrow \left[\mathrm{C}_{\mathcal{L}}^{\mathrm{H}}(Z, Y)^T, \mathrm{C}_{\mathcal{L}}^{\mathrm{H}}(Y)\right]^T \left(\mathrm{C}_{\mathcal{L}}^{\mathrm{H}}(Y) + \mathrm{C}(\widetilde{D}_{\mathcal{L}})\right)^{-1}$
    $\widetilde{d}_{\mathcal{L},j} \leftarrow \mathcal{N}\left(\mathrm{E}(\widetilde{D}_{\mathcal{L}}), \mathrm{C}(\widetilde{D}_{\mathcal{L}})\right)$
    $\left[z_{\mathcal{L},j}^{\text{pos}\,T}, y_{\mathcal{L},j}^{\text{pos}\,T}\right]^T \leftarrow \left[z_{\mathcal{L},j}^{\text{pri}\,T}, y_{\mathcal{L},j}{}^T\right]^T + K_{\mathcal{L}}^{H}\left(\widetilde{d}_{l,j} - y_{l,j}\right)$
  **end for**
  **for** $j = 1, ..., N_{l+1}$ **do**
    $z_{l+1,j}^{\text{pri}} \leftarrow z_{l,j}^{\text{pos}}$
  **end for**
  $\mathrm{E}(Z_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-N_{\mathcal{L}+1}-1}\sum_{j=N_{\mathcal{L}+1}+1}^{N_{\mathcal{L}}} z_{\mathcal{L},j}^{\text{pos}}$
  $\mathrm{E}(Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-N_{\mathcal{L}+1}-1}\sum_{j=N_{\mathcal{L}+1}+1}^{N_{\mathcal{L}}} y_{\mathcal{L},j}^{\text{pos}}$
  $\mathrm{C}(Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-N_{\mathcal{L}+1}-1}\sum_{j=N_{\mathcal{L}+1}+1}^{N_{\mathcal{L}}} \left(y_{\mathcal{L},j}^{\text{pos}} - \mathrm{E}(Y_{\mathcal{L}})\right)\left(y_{\mathcal{L},j}^{\text{pos}} - \mathrm{E}(Y_{\mathcal{L}})\right)^T$
  $\mathrm{C}(Z, Y_{\mathcal{L}}) \leftarrow \frac{1}{N_{\mathcal{L}}-N_{\mathcal{L}+1}-1}\sum_{j=N_{\mathcal{L}+1}+1}^{N_{\mathcal{L}}} \left(z_{\mathcal{L},j}^{\text{pos}} - \mathrm{E}(Z_{\mathcal{L}})\right)\left(y_{\mathcal{L},j}^{\text{pos}} - \mathrm{E}(Y_{\mathcal{L}})\right)^T$
**end for**

**Algorithm 2** SMLES–hybrid formulation.

**Funding** Open Access funding provided by NORCE Norwegian Research Centre AS. 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., TotalEnergies EP 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.

## Declarations

**Financial interests** There are no financial interests to disclose.

**Non-financial interests** The fourth author is an editorial board member in Computational Geosciences. Otherwise, there are no non-financial interests.



## References

1. Chen, Y., Oliver, D.S.: Ensemble randomized maximum likelihood method as an iterative ensemble smoother. Math. Geosci. **44**(1), 1–26 (2012)
2. Emerick, A.A., Reynolds, A.C.: Ensemble smoother with multiple data assimilation. Comput. Geosci. **55**, 3–15 (2013)
3. Luo, X., Bhakta, T., Nævdal, T., et al.: Correlation-based adaptive localization with applications to ensemble-based 4D-seismic history matching. SPE J. **23**(02), 396–427 (2018)
4. 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)
5. Iglesias, M.A.: Iterative regularization for ensemble data assimilation in reservoir models. Comput. Geosci. **19**(1), 177–212 (2015)
6. Raanes, P.N., Stordal, A.S., Evensen, G.: Revising the stochastic iterative ensemble smoother. Nonlinear Process. Geophys. **26**(3), 325–338 (2019)
7. 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)
8. Trani, M., Arts, R., Leeuwenburgh, O., Brouwer, J., Douma, S.: The importance of localization in the assimilation of 4D seismic data in the data assimilation process using the EnKF. In: SEG Technical Program Expanded Abstracts 2009, pp. 3835–3839. Society of Exploration Geophysicists (2009)
9. Liu, M., Grana, D.: Time-lapse seismic history matching with an iterative ensemble smoother and deep convolutional autoencoder. Geophysics **85**(1), M15–M31 (2020)
10. Soares, R.V., Luo, X., Evensen, G., Bhakta, T.: 4D seismic history matching: Assessing the use of a dictionary learning based sparse representation method. J. Pet. Sci. Eng. **195**, 107763 (2020)
11. Yin, Z., Feng, T., MacBeth, C.: Fast assimilation of frequently acquired 4D seismic data for reservoir history matching. Comput. Geosci. **128**, 30–40 (2019)
12. 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)
13. Przybysz-Jarnut, J.K., Jansen, J.D., Gisolf, A.: Joint assimilation of production data and time-lapse density changes from seismics using the representer method. In: ECMOR XI-11th European Conference on the Mathematics of Oil Recovery, pp. cp–62. European Association of Geoscientists & Engineers (2008)
14. Trani, M., Arts, R., Leeuwenburgh, O.: Seismic history matching of fluid fronts using the ensemble Kalman filter. SPE J. **18**(01), 159–171 (2013)
15. Zhang, Y., Leeuwenburgh, O.: Image-oriented distance parameterization for ensemble-based seismic history matching. Comput. Geosci. **21**(4), 713–731 (2017)
16. Mannseth, T., Fossum, K.: Assimilating spatially dense data for subsurface applications–balancing information and degrees of freedom. Comput. Geosci. **22**(5), 1323–1349 (2018)
17. Neto, G.M.S., Soares, R.V., Evensen, G., Davolio, A., Schiozer, D.J.: Subspace ensemble randomized maximum likelihood with local analysis for time-lapse-seismic-data assimilation. SPE J. **26**(02), 1011–1031 (2021)
18. Houtekamer, P.L., Mitchell, H.L.: A sequential ensemble Kalman filter for atmospheric data assimilation. Mon. Weather. Rev. **129**(1), 123–137 (2001)
19. Gaspari, G., Cohn, S.E.: Construction of correlation functions in two and three dimensions. Q. J. Roy. Meteorol. Soc. **125**(554), 723–757 (1999)
20. Chen, Y., Oliver, D.S.: Cross-covariances and localization for EnKF in multiphase flow data assimilation. Comput. Geosci. **14**(4), 579–601 (2010)
21. 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)
22. Leeuwenburgh, O., Brouwer, J., Trani, M.: Ensemble-based conditioning of reservoir models to seismic data. Comput. Geosci. **15**(2), 359–378 (2011)
23. Emerick, A.A.: Analysis of the performance of ensemble-based assimilation of production and seismic data. J. Pet. Sci. Eng. **139**, 219–239 (2016)
24. Alfonzo, M., Oliver, D.S.: Seismic data assimilation with an imperfect model. Comput. Geosci. **24**(2), 889–905 (2020)
25. 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)
26. 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)
27. Fossum, K., Mannseth, T.: Coarse-scale data assimilation as a generic alternative to localization. Comput. Geosci. **21**(1), 167–186 (2017)

28. 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)
29. 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)
30. Hoel, H., Law, K.J.H., Tempone. R.: Multilevel ensemble Kalman filtering. SIAM J. Numer. Anal. **54**(3), 1813–1839 (2016)
31. de Moraes, R.J., Hajibeygi, H., Jansen, J.D.: A multiscale method for data assimilation. Comput. Geosci. **24**(2), 425–442 (2020)
32. 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)
33. 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)
34. 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)
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. Moldovan, G., Lehnasch, G., Cordier, L., Meldi, M.: A multigrid/ensemble Kalman filter strategy for assimilation of unsteady flows. J. Comput. Phys., 110481 (2021)
37. Nezhadali, M., Bhakta, T., Fossum, K., Mannseth, T.: Iterative multilevel assimilation of inverted seismic data. Comput. Geosci. **26**(2), 241–262 (2022)
38. Leeuwen, P.J.V., Evensen, G.: Data assimilation and inverse methods in terms of a probabilistic formulation. Mon. Weather. Rev. **124**(12), 2898–2913 (1996)
39. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. Oceans **99**(C5), 10143–10162 (1994)
40. Emerick, A.A., Reynolds, A.C.: History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. Comput. Geosci. **16**(3), 639–659 (2012)
41. 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)
42. Fossum, K., Mannseth, T.: Parameter sampling capabilities of sequential and simultaneous data assimilation: I. Analytical comparison. Inverse Probl. **30**(11), 114002 (2014a)
43. Fossum, K., Mannseth, T.: Parameter sampling capabilities of sequential and simultaneous data assimilation: II. statistical analysis of numerical results. Inverse Probl. **30**(11), 114003 (2014b)
44. Chavent, G., Liu, J.: Multiscale parametrization for the estimation of a diffusion coefficient in elliptic and parabolic problems. IFAC Proc. Vol. **22**(4), 193–202 (1989)
45. Grimstad, A.A., Mannseth, T.: Nonlinearity, scale, and sensitivity for parameter estimation problems. SIAM J. Sci. Comput. **21**(6), 2096–2113 (2000)
46. Mannseth, T.: Assimilation of multiple linearly dependent data vectors. Comput. Geosci. **24**(1), 349–354 (2020)
47. Aziz, K., Settari, A.: Petroleum Reservoir Simulation. Applied Science Publ. Ltd, London, UK (1979)
48. Schlumberger Ltd: Eclipse reservoir simulation software v2016. Technical Description Manual (2016)
49. Batzle, M., Wang, Z.: Seismic properties of pore fluids. Geophysics **57**(11), 1396–1408 (1992)
50. Fahimuddin, A.: 4D seismic history matching using the ensemble Kalman filter (EnKF): possibilities and challenges. PhD thesis, The University of Bergen (2010)
51. Axelsson, O.: Iterative solution methods. Cambridge University Press, Cambridge (1996)
52. 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)
53. Evensen, G.: Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn. **54**(6), 539–560 (2004)
54. Iglesias, M.A., Law, K.J.H., Stuart, A.M.: Evaluation of Gaussian approximations for data assimilation in reservoir models. Comput. Geosci. **17**(5), 851–885 (2013)