



Ensemble-Based Seismic and Production Data Assimilation Using Selection Kalman Model

Maxime Conjard¹  · Dario Grana²

Received: 14 December 2020 / Accepted: 20 March 2021 / Published online: 8 April 2021
© The Author(s) 2021

Abstract Data assimilation in reservoir modeling often involves model variables that are multimodal, such as porosity and permeability. Well established data assimilation methods such as ensemble Kalman filter and ensemble smoother approaches, are based on Gaussian assumptions that are not applicable to multimodal random variables. The selection ensemble smoother is introduced as an alternative to traditional ensemble methods. In the proposed method, the prior distribution of the model variables, for example the porosity field, is a selection-Gaussian distribution, which allows modeling of the multimodal behavior of the posterior ensemble. The proposed approach is applied for validation on a two-dimensional synthetic channelized reservoir. In the application, an unknown reservoir model of porosity and permeability is estimated from the measured data. Seismic and production data are assumed to be repeatedly measured in time and the reservoir model is updated every time new data are assimilated. The example shows that the selection ensemble Kalman model improves the characterisation of the bimodality of the model parameters compared to the results of the ensemble smoother.

Keywords Selection Kalman model · Seismic inversion · History matching · Ensemble smoother · Channelized reservoir · Multimodality

✉ Maxime Conjard
maxime.conjard@ntnu.no

¹ Department of Mathematical Sciences, NTNU, Trondheim, Norway

² Department of Geology and Geophysics, School of Energy Resources, University of Wyoming, Laramie, WY, USA

1 Introduction

The predictions of fluid flow reservoir models are generally affected by the uncertainty in the initial reservoir model of petrophysical properties, such as porosity and saturation. This model is usually built by integrating different types of data (core measurements, well logs, and surface seismic measurements) and by using geophysical data and statistical algorithms. The so-obtained static reservoir model is then used as an input for fluid flow simulators to predict the dynamic changes in fluid saturation during injection and production. However, there are several sources of uncertainties that might impact the accuracy of the static reservoir model, and as a consequence they affect the accuracy of the fluid flow predictions. Probability random fields can be used to represent the model uncertainty. When additional data become available, including production data or repeated seismic surveys, the static model is updated to reduce the initial uncertainty and improve the reservoir description (Oliver et al. 2008; Evensen 2009).

During injection and production, direct observations of the fluid saturation can be acquired at the well locations. However, this information is available only at sparse locations and cannot fully explain the local variations far away from the wells. Seismic data, repeatedly acquired at multiple times, are used to monitor the changes in seismic response during injection and production (Landrø 2001; Landrø et al. 2003; Dadashpour et al. 2010; Trani et al. 2012). The changes in fluid saturation affect the elastic response of the reservoir. Therefore, by measuring seismic and well data at different times, rock and fluid properties can be predicted by solving inverse problems. In reservoir modeling, this process is called history matching (Williams et al. 1998), which is a data assimilation problem where a set of unknown model variables are updated every time new data are available (Evensen 2009). Time-lapse seismic data have been previously used in history matching in reservoir applications (Huang et al. 1997; Landa et al. 1997; Aanonsen et al. 2003; Kretz et al. 2002; Dong et al. 2005). The integration of seismic data leads to several challenges associated with the resolution of the data and the dimensionality of the problem. In this work, the measured data include oil production and bottom hole pressure at the wells and surface seismic data, and the unknown model variables are the porosity and saturation fields, but the process can be extended to other rock and fluid properties.

Several methods have been presented to solve data assimilation problems, such as gradient-based techniques (e.g. Gauss–Newton), derivative-free optimization methods (e.g. Hooke–Jeeves direct search, genetic algorithm, and particle swarm optimization) and Monte Carlo methods (e.g. ensemble Kalman filter and particle filtering) as described in Oliver and Chen (2011). However, one of the main challenges in seismic and production data assimilation is the representation of the uncertainty. The Ensemble Kalman filter (EnKF) is a Bayesian updating method that relies on the Kalman filter predictor-corrector structure (Evensen 2009). This method has been successfully applied to several geoscience problems including subsurface monitoring (Nævdal et al. 2003; Oliver et al. 2008; Evensen 2009; Oliver and Chen 2011). Skjervheim et al. (2005) applied the EnKF approach to time-lapse seismic and production data in a three-dimensional field case.

Several variations of the ensemble approach have been presented to improve the accuracy of the predictions, including the ensemble smoother and the ensemble smoother with multiple data assimilation (Emerick and Reynolds 2013). Numerous attempts focusing on reproducing the uncertainty in the Kalman gain estimate have also been presented (Myrseth et al. 2010; Sætrom and Omre 2011). Mathematical formulations including localization, covariance inflation, and dimensionality reduction have also been tested for preventing overfitting, ensemble collapsing and spurious correlations and to make the approach applicable to practical applications, including non-linear models in large grids. However, ensemble based methods are generally based on Gaussian assumptions for the conditional distribution of the model, and therefore it is not theoretically applicable to models with multimodal random variables. A regression towards the mean occurs during the conditioning steps, thereby rendering posterior marginal distributions Gaussian. This is a challenging problem in reservoir simulation because subsurface properties, such as porosity and saturation, often show non-Gaussian distributions owing to the underlying geology. The focus on this study is on representing non-Gaussian features in the posterior ensemble of the variables of interest.

Statistical methods, such as Ensemble Randomized Likelihood (EnRML) (Chen and Oliver 2012), Gaussian anamorphosis (Bertino et al. 2003; Zhou et al. 2012), Gaussian mixture models (Dovera and Della Rossa 2011) and truncated pluri-Gaussian (Oliver and Chen 2018), have been developed to address this issue. EnRML and Gaussian anamorphosis attempt to modify the conditioning step to preserve the desirable features in the posterior ensemble, and they are compatible with any prior ensemble. Gaussian mixture and truncated pluri-Gaussian models assume specific prior distribution, and the ensemble algorithm is adapted to the prior model. Recent developments also include the indicator based data assimilation proposed in Kumar and Srinivasan (2018) and the hierarchical truncated pluri-Gaussian simulation (Silva and Deutsch 2019).

An alternative to Gaussian mixture models is the selection-Gaussian models (Arellano-Valle et al. 2006; Arellano-Valle and del Pino 2004; Omre and Rimstad 2018). These models have been applied to spatio-temporal inversion in Conjard and Omre (2020) and validated using synthetic examples. In this manuscript, the compatibility of the selection Kalman model framework detailed in Conjard and Omre (2020) is extended by integrating the selection Kalman model with the ensemble smoother. The resulting approach is named 'selection ensemble smoother'. The prior for the initial state vector is a selection-Gaussian distribution, but instead of sampling the initial ensemble directly from the selection-Gaussian distribution, the proposed approach takes advantage of the structure of the selection-Gaussian to sample from a Gaussian augmented state vector, thereby preventing unwanted regression towards the mean while keeping the computational cost to a minimum. A similar approach has been proposed for asymmetric priors using the closed skew Gaussian distribution (Naveau et al. 2005; Rezaie and Eidsvik 2014).

The proposed methodology is then tested on a two-dimensional channelized reservoir, with the goal of estimating the porosity and saturation fields based on seismic and production data. In Sect. 2, the problem is defined in a Bayesian formulation, the novel method is presented and the forward operator for the reservoir model is defined.

In Sect. 3, a synthetic two-dimensional case study featuring a channelized reservoir with two facies is presented. In Sect. 4, conclusions are presented.

2 Methodology

The Methodology section describes the proposed framework for data assimilation in reservoir modeling. The Bayesian formulation of the selection ensemble smoother is first presented. Then, the forward operators for the geophysics and fluid flow modeling are summarized.

In this paper, $f(\mathbf{y})$ denotes the probability density function (pdf) of a random variable \mathbf{y} , $\varphi_n(\mathbf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the pdf of the Gaussian n -vector \mathbf{y} with expectation n -vector $\boldsymbol{\mu}$ and covariance $(n \times n)$ -matrix $\boldsymbol{\Sigma}$. Furthermore, $\Phi_n(A; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the probability of the aforementioned Gaussian n -vector \mathbf{y} to be in $A \subset \mathbb{R}^n$. Further, \mathbf{i}_n is used to denote the all-ones n -vector.

2.1 Hidden Markov Model

From a Bayesian perspective, the data assimilation can be assimilated in a hidden Markov model formulation, where at a given time the current state is updated based on the previous state and the current measurements. The state of the chain represents the values of the variables of interest. If a variable of interest is \mathbf{r} and the observable measurement is \mathbf{d} , the goal of data assimilation is to predict the initial value of the variable of interest, i.e. the state of the chain $\mathbf{r}_0 = \mathbf{r}_{t=0}$ based on the data \mathbf{d}_t available at different times $t = 0, \dots, T$, assuming that the forward model and the likelihood function of the measurements are known.

Given the prior model of the initial state vector $\mathbf{r}_0 \in \mathbb{R}^n$, the distribution of the vector of the following states $[\mathbf{r}_{1:T+1} | \mathbf{r}_0] = [\mathbf{r}_1, \dots, \mathbf{r}_{T+1} | \mathbf{r}_0]$ is defined as

$$f(\mathbf{r}_{1:T+1} | \mathbf{r}_0) = \prod_{t=0}^T f(\mathbf{r}_{t+1} | \mathbf{r}_t), \quad (1)$$

with

$$[\mathbf{r}_{t+1} | \mathbf{r}_t] = \omega_t(\mathbf{r}_t) + \boldsymbol{\epsilon}_t^r \sim f(\mathbf{r}_{t+1} | \mathbf{r}_t), \quad (2)$$

where $\omega_t(\cdot) \in \mathbb{R}^n$ is the forward operator with random Gaussian n -vector $\boldsymbol{\epsilon}_t^r$, independent and identically distributed (iid) for each time t . The forward model defines a first-order Markov chain.

Consider, then, a set of measurements $\mathbf{d} = \{\mathbf{d}_1, \dots, \mathbf{d}_T\}$ where $\mathbf{d}_t \in \mathbb{R}^m, \forall t = 1, \dots, T$. It is assumed that the measured data depend on the variables of interest according to a likelihood model. The likelihood model for $[\mathbf{d} | \mathbf{r}]$ is conditionally independent and defined as

$$f(\mathbf{d}|\mathbf{r}) = \prod_{t=0}^T f(\mathbf{d}_t|\mathbf{r}_t), \tag{3}$$

with

$$[\mathbf{d}_t|\mathbf{r}_t] = \psi_t(\mathbf{r}_t) + \epsilon_t^d \sim f(\mathbf{d}_t|\mathbf{r}_t), \tag{4}$$

where $\psi_t(\cdot, \cdot) \in \mathbb{R}^m$ is the likelihood function with random Gaussian m -vector ϵ_t^d , iid for each t .

Let $\Phi_t = \omega_t \circ \omega_{t-1} \circ \dots \circ \omega_1$ such that

$$[\mathbf{r}_{t+1}|\mathbf{r}_0] = \Phi_t(\mathbf{r}_0) + \epsilon_t^r, \tag{5}$$

then, it is possible to write $[\mathbf{d}_t|\mathbf{r}_0]$ as

$$[\mathbf{d}_t|\mathbf{r}_0] = \psi_t \circ \Phi_{t-1}(\mathbf{r}_0) + \epsilon_t^d, \tag{6}$$

and consequently $[\mathbf{d}|\mathbf{r}_0]$ as

$$[\mathbf{d}|\mathbf{r}_0] = \begin{bmatrix} \psi_1 \circ \Phi_0(\mathbf{r}_0) \\ \dots \\ \psi_T \circ \Phi_{T-1}(\mathbf{r}_0) \end{bmatrix} + \begin{bmatrix} \epsilon_1^d \\ \dots \\ \epsilon_T^d \end{bmatrix}. \tag{7}$$

Equations 1 and 4 provide the framework for the EnKF, whereas Eq. 7 gives the framework for the ensemble smoother using the same hidden Markov model formulation.

2.2 Ensemble Smoother

The ensemble smoother is an inverse method that assimilates the measured data \mathbf{d} in a unique step by evaluating the likelihood $f(\mathbf{r}_0|\mathbf{d})$. The likelihood model described in Eq. 7 is represented as

$$\mathbf{d} = g(\mathbf{r}_0) + \delta, \tag{8}$$

where $g(\mathbf{r}_0) = \begin{bmatrix} \psi_1 \circ \Phi_0(\mathbf{r}_0) \\ \dots \\ \psi_T \circ \Phi_{T-1}(\mathbf{r}_0) \end{bmatrix}$ and $\delta = \begin{bmatrix} \epsilon_1^d \\ \dots \\ \epsilon_T^d \end{bmatrix}$.

In this data assimilation approach, an ensemble of n_e members $\mathbf{r}^{p(i)}$, $i = 1, \dots, n_e$ is generated from the prior distribution of the initial state $f(\mathbf{r}_0)$ and the ensemble members are updated as

$$\mathbf{r}^{u(i)} = \mathbf{r}^{p(i)} + \hat{\Gamma}_{r,d} \hat{\Sigma}_d^{-1} (\mathbf{d} - \mathbf{d}^i), i = 1, \dots, n_e, \tag{9}$$

where \mathbf{d}^i represents the pseudo-data (i.e. the data predicted from the model according to the forward operator) generated for the ensemble member i as $\mathbf{d}^i = g(\mathbf{r}^{p(i)}) + \boldsymbol{\delta}$, and where the crosscovariance and covariance matrices, $\hat{\boldsymbol{\Gamma}}_{r,d}$ and $\hat{\boldsymbol{\Sigma}}_d$, are computed as

$$\hat{\boldsymbol{\Gamma}}_{r,d} = \frac{1}{n_e} \sum_{i=1}^{n_e} (\mathbf{r}^{p(i)} - \bar{\mathbf{r}}^p)(\mathbf{d}^i - \bar{\mathbf{d}})^T, \quad (10)$$

$$\hat{\boldsymbol{\Sigma}}_d = \frac{1}{n_e} \sum_{i=1}^{n_e} (\mathbf{d}^i - \bar{\mathbf{d}})(\mathbf{d}^i - \bar{\mathbf{d}})^T, \quad (11)$$

with $\bar{\mathbf{r}}^p$ being the empirical mean from the initial ensemble and $\bar{\mathbf{d}}$ the empirical mean of the pseudo-data. Localization has been proposed in data assimilation problems to avoid spurious correlations and ensemble collapse (Houtekamer and Mitchell 1998; Chen and Oliver 2011). The method is based on distance-dependent functions to constrain the updates of the model variables and their variability to a specific area based on the observed data in that region. Localization should be applied on a scale larger than the correlation length of the spatial variograms of the model variables to avoid introducing biases on the spatial correlation and resolution of the model. In the case study, localization is used on both $\hat{\boldsymbol{\Gamma}}_{r,d}$ and $\hat{\boldsymbol{\Sigma}}_d$ because, as is often the case, the size of the ensemble is small compared to the dimension of the state vector.

One of the most popular inverse methods in reservoir modeling simulation is the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds 2013). In the MDA version of the ensemble smoother, the data are assimilated multiple times to improve convergence. This approach requires a limited number of runs of the forward operator and handles the non-linearity in the forward and likelihood models. The mathematical formulation of the MDA approach is detailed in Emerick and Reynolds (2013).

In the ensemble smoother, the prior distribution for the initial state can be chosen freely, however it is difficult to preserve non-Gaussian features in the posterior distribution owing to the linearized update (Eq. 9). The selection ensemble smoother attempts to represent non-Gaussian features in the posterior distribution.

2.3 Selection Ensemble Smoother

The selection-Gaussian pdf can represent multimodality, skewness, and heavy-tailed distributions; it is an extension of the skew-Gaussian distribution (Omre and Rimstad 2018). Consider a Gaussian $(n + n)$ -vector $[\tilde{\mathbf{r}}, \boldsymbol{\kappa}]$

$$\begin{bmatrix} \tilde{\mathbf{r}} \\ \boldsymbol{\kappa} \end{bmatrix} \sim \varphi_{2n} \left(\begin{bmatrix} \tilde{\mathbf{r}} \\ \boldsymbol{\kappa} \end{bmatrix}; \begin{bmatrix} \boldsymbol{\mu}_{\tilde{\mathbf{r}}} \\ \boldsymbol{\mu}_{\boldsymbol{\kappa}} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{\tilde{\mathbf{r}}} & \boldsymbol{\Sigma}_{\tilde{\mathbf{r}}} \boldsymbol{\Gamma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}}^T \\ \boldsymbol{\Gamma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}} \boldsymbol{\Sigma}_{\tilde{\mathbf{r}}} & \boldsymbol{\Sigma}_{\boldsymbol{\kappa}} \end{bmatrix} \right), \quad (12)$$

where $\boldsymbol{\mu}_{\tilde{\mathbf{r}}}$ and $\boldsymbol{\mu}_{\boldsymbol{\kappa}}$ are n -vectors, $\boldsymbol{\Gamma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}}$ is a $(n \times n)$ -matrix, and $\boldsymbol{\Sigma}_{\tilde{\mathbf{r}}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\kappa}}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}}$ are three covariance $(n \times n)$ -matrices with $\boldsymbol{\Sigma}_{\boldsymbol{\kappa}} = \boldsymbol{\Gamma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}} \boldsymbol{\Sigma}_{\tilde{\mathbf{r}}} \boldsymbol{\Gamma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}}^T + \boldsymbol{\Sigma}_{\boldsymbol{\kappa}|\tilde{\mathbf{r}}}$.

If \mathbf{r}_0 is defined as $\mathbf{r}_0 = [\tilde{\mathbf{r}}|\boldsymbol{\kappa} \in A]$, where $A \subset \mathbb{R}^n$ is a selection set of dimension n , then \mathbf{r}_0 is distributed according to a selection-Gaussian pdf given by

$$f(\mathbf{r}_0) = [\Phi_n(A; \boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa)]^{-1} \Phi_n(A; \boldsymbol{\mu}_\kappa + \boldsymbol{\Gamma}_{\kappa|\tilde{\mathbf{r}}}(\mathbf{r}_0 - \boldsymbol{\mu}_{\tilde{\mathbf{r}}}), \boldsymbol{\Sigma}_{\kappa|\tilde{\mathbf{r}}}) \times \varphi_n(\mathbf{r}_0; \boldsymbol{\mu}_{\tilde{\mathbf{r}}}, \boldsymbol{\Sigma}_{\tilde{\mathbf{r}}}). \tag{13}$$

The structure of the selection-Gaussian model provides a suitable model for non-Gaussian features of the posterior distribution of the inverse problem.

A selection-Gaussian random variable $\mathbf{r}_A \in \mathbb{R}^n$ can be written as $[\mathbf{r}|\boldsymbol{\kappa} \in A]$ with $\mathbf{r} \in \mathbb{R}^n, \boldsymbol{\kappa} \in \mathbb{R}^n, A \subset \mathbb{R}^n$, and where $[\mathbf{r}, \boldsymbol{\kappa}]$ is jointly Gaussian. When conditioning on data, it is always possible to consider $[\mathbf{r}, \boldsymbol{\kappa}]$, condition first on the data \mathbf{d} and then on $\boldsymbol{\kappa} \in A$. Similarly when transforming $\mathbf{r}_A \rightarrow g(\mathbf{r}_A)$, one can show that the distribution of $g(\mathbf{r}_A)$ is equal to that of $[g(\mathbf{r})|\boldsymbol{\kappa} \in A]$. In any data assimilation process where a selection-Gaussian distribution is chosen as a prior, it is therefore possible to work on the augmented Gaussian vector $[\mathbf{r}, \boldsymbol{\kappa}]$ throughout the assimilation step(s), and condition on the latent variable $\boldsymbol{\kappa}$ being in A afterwards. The selection Ensemble Kalman filter detailed in (Conjard and Omre 2020) adopts this approach.

The proposed selection ensemble smoother for data assimilation problems is defined by a prior distribution $f(\mathbf{r}_0)$ that is selection-Gaussian such that $\mathbf{r}_0 = [\tilde{\mathbf{r}}_0|\boldsymbol{\kappa} \in A], A \subset \mathbb{R}^n$, and where the prior state vector is represented by the augmented vector $[\tilde{\mathbf{r}}_0, \boldsymbol{\kappa}]$. When assessing $[\mathbf{r}_0|\mathbf{d}_{1:T}]$, the selection ensemble smoother first evaluates $[\tilde{\mathbf{r}}_0, \boldsymbol{\kappa}|\mathbf{d}_{1:T}]$ and then adopts a Markov chain Monte-Carlo (MCMC) algorithm (Omre and Rimstad 2018) to calculate $[\mathbf{r}_0|\mathbf{d}_{1:T}] = [\tilde{\mathbf{r}}_0|\boldsymbol{\kappa} \in A, \mathbf{d}_{1:T}]$. There are two advantages in using the augmented state vector. First, in ensemble methods, the update step is optimal if the prior is Gaussian and the prior and the data are jointly Gaussian. Therefore, by enforcing a Gaussian initial state vector, the optimal update conditions are ensured. Then, since the update is optimal only if the prior is Gaussian and the forward and likelihood models linear and Gaussian, the traditional update tends to make the posterior ensemble more and more Gaussian; however, conditioning on the latent variable after the assimilation is performed allows representing non-Gaussian features in the posterior distribution.

The selection ensemble smoother proceeds as follows. First, n_e ensemble members are generated

$$[\tilde{\mathbf{r}}^{p(i)}, \boldsymbol{\kappa}^{p(i)}], i = 1, \dots, n_e, \tag{14}$$

from the augmented Gaussian prior distribution of the initial state $f(\mathbf{r}_0, \boldsymbol{\kappa})$, and they are updated as

$$\begin{bmatrix} \tilde{\mathbf{r}}^{u(i)} \\ \boldsymbol{\kappa}^{u(i)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{r}}^{p(i)} \\ \boldsymbol{\kappa}^{p(i)} \end{bmatrix} + \hat{\boldsymbol{\Gamma}}_{\tilde{\mathbf{r}}, \mathbf{d}} \hat{\boldsymbol{\Sigma}}_d^{-1} (\mathbf{d} - \mathbf{d}^i), i = 1, \dots, n_e, \tag{15}$$

where \mathbf{d}^i represents the pseudo-data generated for the ensemble member i as $\mathbf{d}^{(i)} = g(\tilde{\mathbf{r}}^{p(i)}) + \delta$, and where $\hat{\mathbf{\Gamma}}_{\tilde{\mathbf{r}}\kappa,d}$ and $\hat{\mathbf{\Sigma}}_d$ are calculated as follows

$$\hat{\mathbf{\Gamma}}_{\tilde{\mathbf{r}}\kappa,d} = \frac{1}{n_e} \sum_{i=1}^{n_e} \left(\begin{bmatrix} \tilde{\mathbf{r}}^{p(i)} \\ \kappa^{p(i)} \end{bmatrix} - \begin{bmatrix} \bar{\tilde{\mathbf{r}}^p} \\ \bar{\kappa}^p \end{bmatrix} \right) (\mathbf{d}^i - \bar{\mathbf{d}})^T, \tag{16}$$

$$\hat{\mathbf{\Sigma}}_d = \frac{1}{n_e} \sum_{i=1}^{n_e} (\mathbf{d}^i - \bar{\mathbf{d}})(\mathbf{d}^i - \bar{\mathbf{d}})^T, \tag{17}$$

with $\bar{\tilde{\mathbf{r}}^p}$ and $\bar{\kappa}^p$ being the empirical means from each component of the initial ensemble and $\bar{\mathbf{d}}$ being the empirical mean of the pseudo-data. In the case study, localization is used on both $\hat{\mathbf{\Gamma}}_{\tilde{\mathbf{r}}\kappa,d}$ and $\hat{\mathbf{\Sigma}}_d$ because the size of the ensemble is small compared to the dimension of the state vector.

After conditioning on the data, the ensemble represents $[\mathbf{r}_0, \kappa|\mathbf{d}]$ which is assumed to be Gaussian. The target distribution $f(\mathbf{r}_0|\kappa \in A, \mathbf{d})$ can be written as

$$f(\mathbf{r}_0|\kappa \in A, \mathbf{d}) = \frac{1}{f(\kappa \in A|\mathbf{d})} \int_A f(\mathbf{r}_0|\kappa, \mathbf{d})f(\kappa|\mathbf{d})d\kappa, \tag{18}$$

$$= \int_{\mathbb{R}^n} f(\mathbf{r}_0|\kappa, \mathbf{d}) \frac{f(\kappa, \mathbf{d})\mathbf{I}_A(\kappa)}{f(\kappa \in A, \mathbf{d})} d\kappa, \tag{19}$$

$$= \int_{\mathbb{R}^n} f(\mathbf{r}_0|\kappa, \mathbf{d})f(\kappa|\kappa \in A, \mathbf{d})d\kappa. \tag{20}$$

Equation 20 shows how sampling from the posterior distribution is performed: first from the truncated Gaussian $\kappa^* \sim [\kappa|\kappa \in A, \mathbf{d}]$ and then from the Gaussian vector $\mathbf{r}^* \sim [\mathbf{r}_0|\kappa^*, \mathbf{d}]$. Hence \mathbf{r}^* are distributed according to the target distribution $f(\mathbf{r}_0|\kappa \in A, \mathbf{d})$.

The challenging step is the sampling from the truncated Gaussian, which is performed using an MCMC sampling algorithm (Omre and Rimstad 2018). The MCMC algorithm, denoted by S , generates κ^*

$$\kappa^* = S(\hat{\boldsymbol{\mu}}_{\kappa|\mathbf{d}}, \hat{\mathbf{\Sigma}}_{\kappa|\mathbf{d}}, A), \tag{21}$$

where $\hat{\boldsymbol{\mu}}_{\kappa|\mathbf{d}}$ and $\hat{\mathbf{\Sigma}}_{\kappa|\mathbf{d}}$ are the empirical mean and covariance matrix of the posterior ensemble for κ . Note that the selection set A is considered constant throughout the algorithm. The empirical covariance matrix estimate $\hat{\mathbf{\Sigma}}_{\kappa|\mathbf{d}}$ is known to be subject to sampling errors, which are addressed by using localization on $\hat{\mathbf{\Sigma}}_{\kappa|\mathbf{d}}$. The algorithm itself has to be tailored to the specific application as rank deficiency is to be expected in high-dimensional inverse problems. The original algorithm expected a positive definite covariance matrix, and even though localization does increase the rank of the localised matrix, it is unlikely that it becomes full rank after the transformation. Therefore, singular value decomposition is used to evaluate the pseudo inverse instead of outright inverting the matrix in the algorithm. The singular value decomposition of

$\hat{\Sigma}_{\kappa|d}$ gives,

$$\hat{\Sigma}_{\kappa|d} = U \Sigma V^T, \tag{22}$$

where U and V are $(n \times n)$ orthogonal matrices, and Σ is an $(n \times n)$ diagonal matrix whose diagonal entries are non-negative. The pseudo inverse $\hat{\Sigma}_{\kappa|d}^+$ of $\hat{\Sigma}_{\kappa|d}$ is then given by

$$\hat{\Sigma}_{\kappa|d}^+ = V \Sigma^+ U^T, \tag{23}$$

where Σ^+ is the pseudoinverse of Σ , which is formed by replacing every non-zero diagonal entry by its inverse and transposing the resulting matrix.

2.4 Forward Operator

The data assimilation problem discussed in this work requires two operators: the fluid flow simulation operator used as a forward model and the geophysical operator used for the likelihood model of the measured data.

The forecast of the dynamic behavior of fluid displacement in the subsurface requires a mathematical-physical model generally called fluid flow simulation. For hydrocarbon-water two-phase flow in subsurface reservoirs, the model is based on Darcy’s equation and mass balance law and includes two partial differential equations (PDE) in two unknowns, namely water saturation $s_w(t)$ and water pressure $p_w(t)$. This model assumes that there are only two immiscible fluids, the flow is isothermal, and the capillary pressure is a function of saturation. For a system of oil and water, the PDE system can be written as follows

$$\nabla \cdot \left[\frac{kk_w}{B_w \mu_w} (\nabla p_w(t) - \gamma_w \nabla d) \right] + q_w = \frac{\partial}{\partial t} \left(\frac{\phi s_w(t)}{B_t} \right), \tag{24}$$

$$\nabla \cdot \left[\frac{kk_o}{B_o \mu_o} (\nabla (p_w(t) - p_c(t)) - \gamma_o \nabla d) \right] + q_o = \frac{\partial}{\partial t} \left(\frac{\phi (1 - s_w(t))}{B_o} \right), \tag{25}$$

where k is the absolute permeability, ϕ is the porosity, d is the depth, p_c is the capillary pressure, q_w is the water production rate, q_o is the oil production rate and the constants k_w , B_w , μ_w , γ_w , γ_o , k_o , B_o , μ_o , and γ_o are parameters associated with the fluid properties of oil and water. If porosity and permeability are known, then the system can be solved using finite difference methods to compute the water and saturation and pressure, $s_w(t)$ and $p_w(t)$, at any time t . The oil saturation and pressure can then be computed from water saturation and water pressure.

The permeability field is assumed to be a point-wise function of the porosity field given by Kozeny–Carman’s equation

$$k = a \times \frac{\phi^3}{(1 - \phi)^2}, \tag{26}$$

where a is a geometrical factor estimated by fitting the true permeability field. This approximation allows removing k from the state vector because it is then a known function of the porosity ϕ .

Similarly, the prediction of the geophysical response of subsurface models requires a mathematical-physical model that includes rock physics relations and seismic wave propagation models. Rock physics models allow computing P-wave and S-wave velocities and densities given the petrophysical properties, porosity, and fluid saturation. The seismic wave propagation model allows calculating the seismic response, in terms of amplitude and travel time, based on the velocities and densities. For a partially-saturated porous rock of porosity ϕ and water saturation s_w , the density $\rho(\phi, s_w)$ can be computed as

$$\rho(\phi, s_w) = (1 - \phi)\rho_m + \phi(\rho_w s_w + \rho_o(1 - s_w)), \quad (27)$$

and the P-wave and S-wave velocities, $v_p(\phi, s_w)$ and $v_s(\phi, s_w)$, are given by Raymer–Dvorkin’s relations (Dvorkin et al. 2014)

$$v_p(\phi, s_w) = (1 - \phi)^2 \sqrt{\frac{K_m + \frac{4}{3}G_m}{\rho_m}} + \phi \sqrt{\frac{\left(\frac{s_w}{K_w} + \frac{1-s_w}{K_o}\right)^{-1}}{\rho_w s_w + \rho_o(1 - s_w)}}, \quad (28)$$

$$v_s(\phi, s_w) = (1 - \phi)^2 \sqrt{\frac{G_m}{\rho_m}} \sqrt{\frac{(1 - \phi)\rho_m}{\rho(\phi, s_w)}}, \quad (29)$$

where ρ_m is the density of the solid phase, ρ_w is the density of water, ρ_o is the density of oil, K_m is the bulk modulus of the solid phase, G_m is the shear modulus of the solid phase, K_w is the bulk modulus of water, and K_o is the bulk modulus of oil. All these parameters are assumed to be constant and known. The reflection coefficients $r_{pp}(t, \theta)$ associated with seismic wave propagation are a function of the seismic travel time and incident angle θ and depend on the relative change in P-wave and S-wave velocity and density. For small incident angles, the reflection coefficients can be approximated using Aki–Richards equations (Aki and Richards 1980) as

$$r_{pp}(t, \theta) = r_p(\theta) \frac{\partial}{\partial t} v_p(t) + r_s(\theta) \frac{\partial}{\partial t} v_s(t) + r_\rho(\theta) \frac{\partial}{\partial t} \rho(t), \quad (30)$$

where

$$r_p(\theta) = \frac{1}{2}(1 + \tan^2(\theta)), \quad (31)$$

$$r_s(\theta) = -4c \sin^2(\theta), \quad (32)$$

$$r_\rho = \frac{1}{2}(1 - 4c \sin^2(\theta)), \quad (33)$$

with c being a constant value equal to the square of the average v_s/v_p ratio.

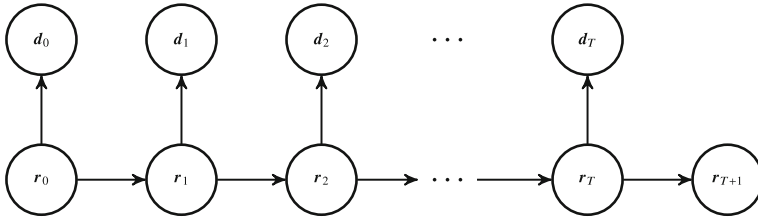


Fig. 1 Graph of the hidden Markov model

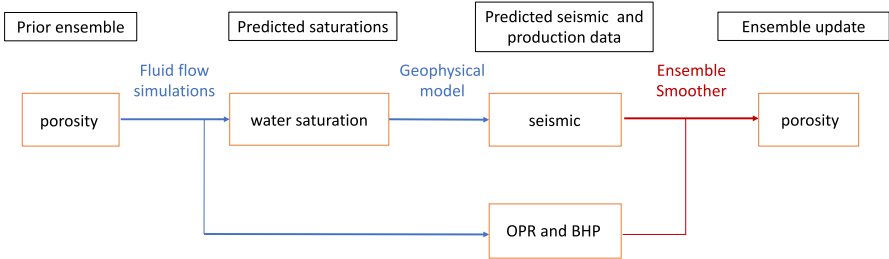


Fig. 2 Data assimilation flowchart

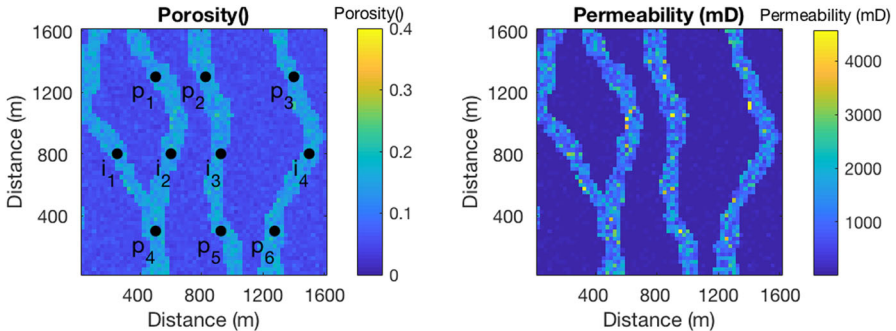


Fig. 3 Reference porosity and permeability fields, with the location of the injection and producer wells

The amplitudes $s(t, \theta)$ of the seismogram are then approximated as a convolution of a wavelet $w(t, \theta)$ and the reflection coefficients $r_{pp}(t, \theta)$ as

$$s(t, \theta) = \int w(u, \theta)r_{pp}(t - u, \theta)du, \tag{34}$$

where the wavelet $w(t, \theta)$ is generally assumed to be known. As shown in Buland and Omre (2003), the convolution can be discretized by decomposing the matrix associated with the forward operator as $G = WAD$, where W is the wavelet matrix, A is the reflection coefficient matrix, and D is the first order differential matrix.

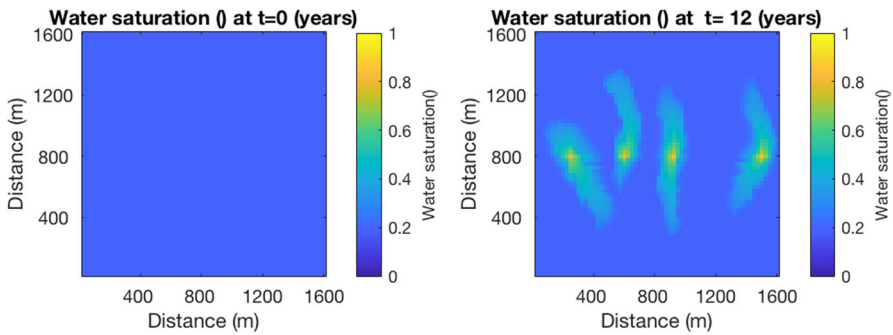


Fig. 4 Water saturation at initial and final times

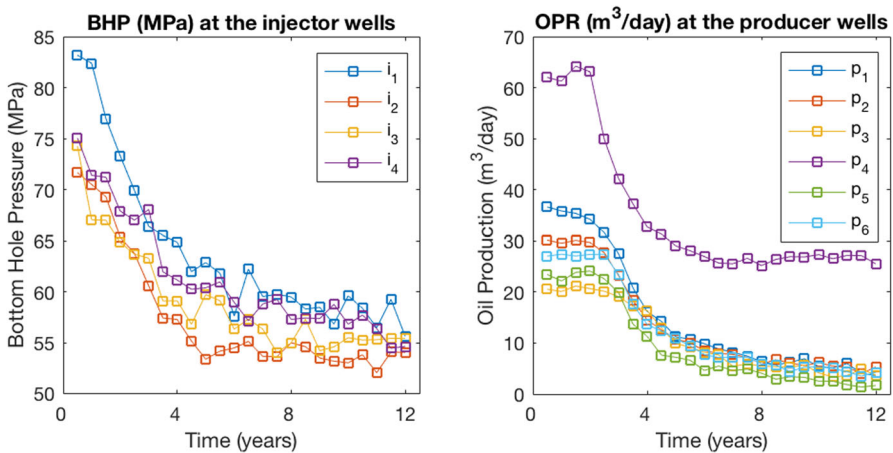


Fig. 5 Well production data: bottom hole pressure and oil production rate

In the current implementation, we adopted a forward model based on the convolution of the wavelet with the linearized approximation of Zoeppritz equations (Shuey 1985; Aki and Richards 1980) because it is mathematically tractable, and the computational cost is limited. The proposed approach can be extended to other forward operations including convolutional models based on full Zoeppritz equations, Kennett's invariant imbedding method and the Born weak scattering approximation (Zoeppritz 1919; Kennett 1984; Russell 1988; Sheriff and Geldart 1995; Yilmaz 2001; Aki and Richards 1980). In theory, the same formulation can also be applied to the full waveform model for wave propagation, however, the computational cost of the approach is still prohibitive for field-scale applications. Applications of ensemble-based methods to full waveform inversion problems have been proposed Thurin et al. (2019) and Gineste et al. (2020), according to Gaussian assumptions of the error and model variables.

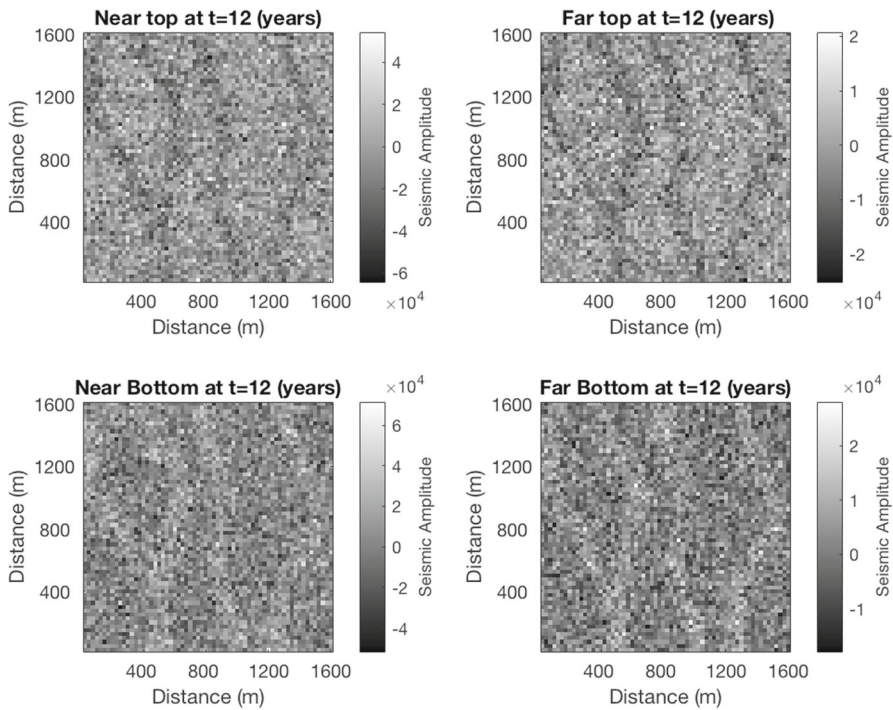


Fig. 6 Seismic data (near and far angle) at the top and bottom of the reservoir after 12 years of injection and production

Table 1 Parameter values

$\mu_r^\phi()$	$\mu_k^\phi()$	$\sigma_r^\phi()$	$\gamma^\phi()$	A()	$d_x(m)$	$d_y(m)$
-2.5	0	0.2	0.99	$(]-\infty, -0.2] \cap [0.3, 2])^n$	100	200

3 Application

The proposed method is applied to a two-dimensional synthetic case to demonstrate the validity of the selection ensemble smoother (S-ES) and the advantages compared to the traditional ensemble smoother (ES).

The synthetic example mimics the production of hydrocarbon in a reservoir by water injection as shown in Liu and Grana (2018, 2020). The geology of the reservoir includes four channels in the north-south direction with an average porosity of 0.2. The channels are surrounded by shale with effective porosity close to 0. It is assumed that the entire reservoir is filled by oil with an irreducible water saturation of 0.2. Water is injected in four wells and oil is produced from six wells. All the wells are located within the channels. Fluid flow occurs predominantly within the channels due to the low porosity and permeability of the surrounding shale, and the injected water displaces the oil in place towards the producing wells.

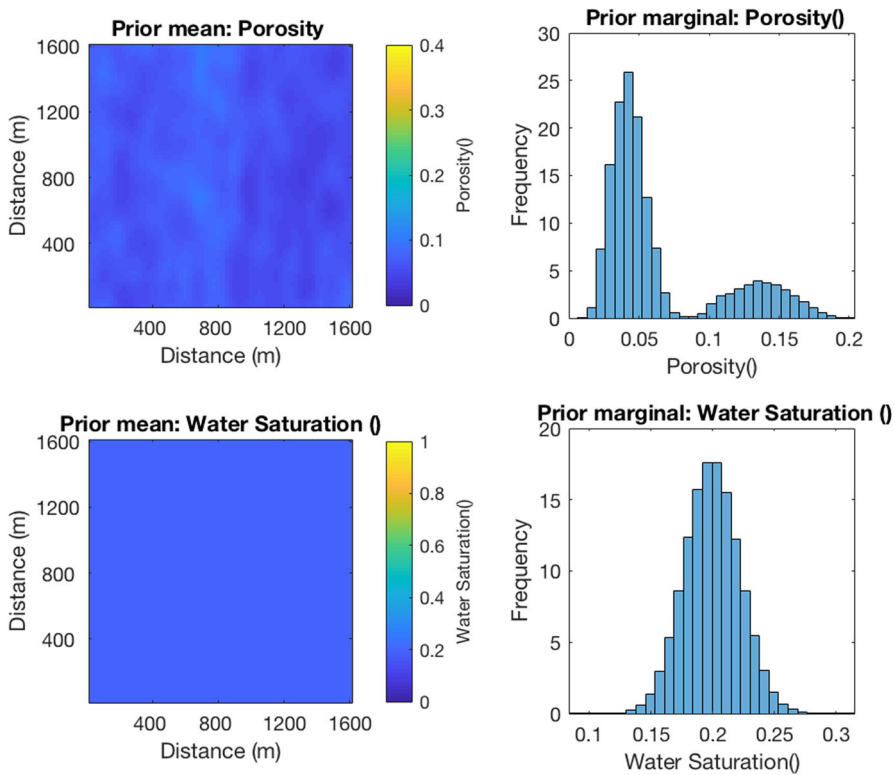


Fig. 7 Prior ensemble of models: mean maps and marginal histograms

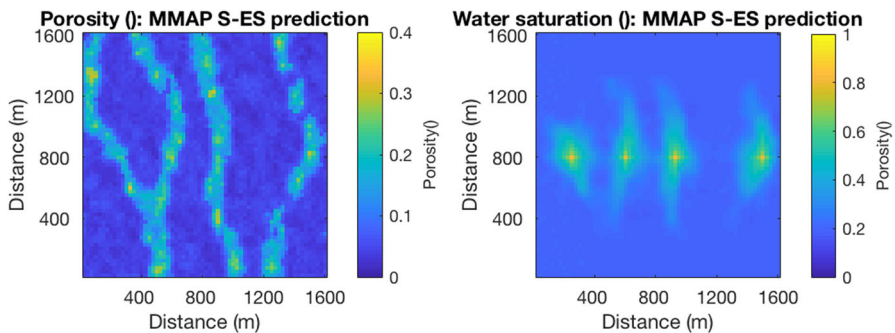


Fig. 8 Posterior ensemble prediction from the selection ensemble smoother: marginal maximum a posteriori (MMAP) predictions

The reservoir model mimics the two-phase (oil and water) flow due to water injection and oil production in a geological system represented in a two-dimensional uniform grid of $n = 64 \times 64$ cells of dimension $40 \times 40 \times 25$ m. Figure 3 shows the true porosity ϕ_r and permeability k_r fields. The reference porosity model is generated using the following geostatistical modeling approach: First, a facies model for a channelized

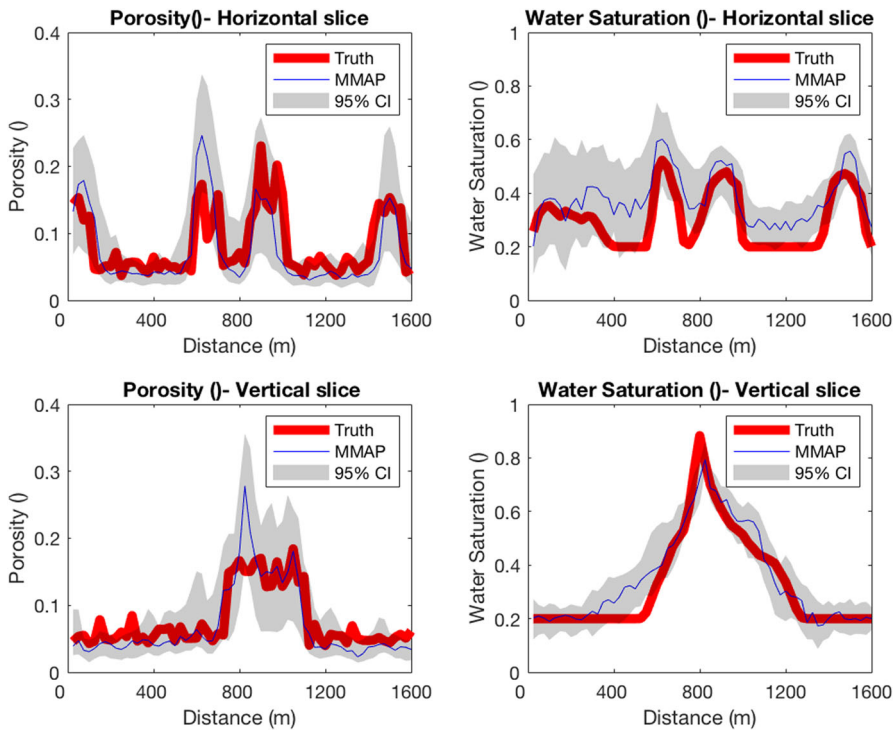


Fig. 9 Posterior ensemble prediction from the selection ensemble smoother: predictions along horizontal and vertical lines

system is generated using object modeling with randomly sampled parameters of the channel geometry (length, width, and tortuosity). Then a porosity model is generated using sequential Gaussian simulation within each facies using facies-dependent prior distributions to mimic medium-high porosity within the channels and low porosity in the background shale. Finally, permeability is simulated in the logarithmic domain, conditioned on the porosity field using sequential Gaussian co-simulation. Figure 3 shows the geostatistical realization used as the reference (true) model for the application. Figure 3 also shows the locations of the four water injection wells (i_1, i_2, i_3, i_4) and the six producer wells (p_1, \dots, p_6).

Fluid flow is then simulated to create the reference (true) saturation maps for a time period of 12 years using the Matlab Reservoir Simulation Toolbox (MRST) (Lie 2019). Figure 4 shows the evolution of the water saturation in the field for the reference porosity and permeability field specified in Fig. 3. The initial water saturation s_{w0} is set constant to 0.2, then water propagates predominantly within the channels until it reaches, after 12 years, one of the producers. Figure 5 shows the bottom hole pressure (BHP) and the oil production rate (OPR) every six months at the injector and producer wells, respectively.

In addition to the saturation and production data, a synthetic geophysical dataset is generated for the data assimilation problem. To generate seismic data, two additional

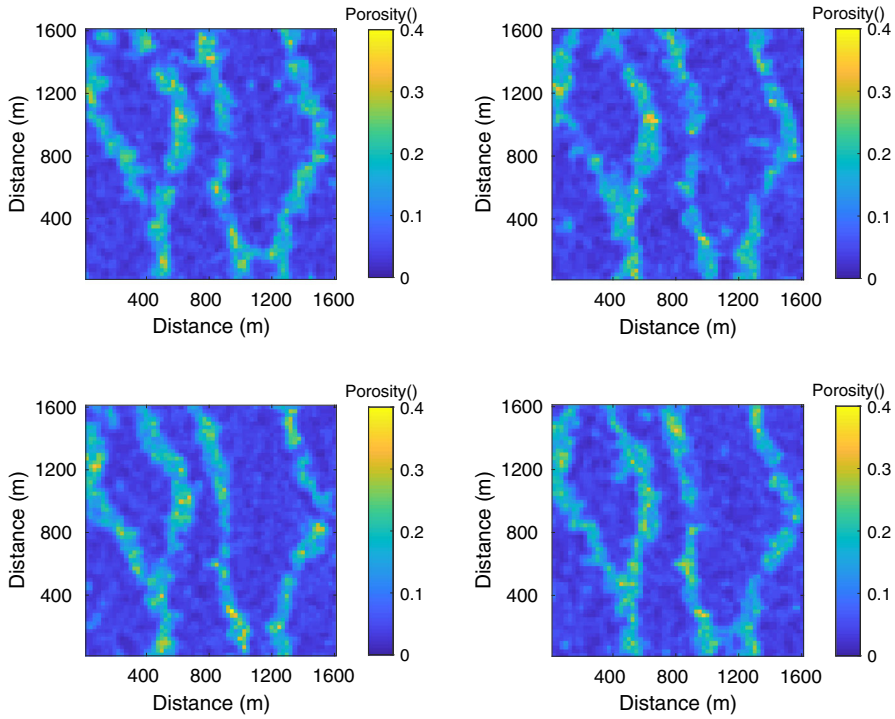


Fig. 10 Posterior realizations from the posterior distribution of the porosity field obtained from the selection ensemble smoother

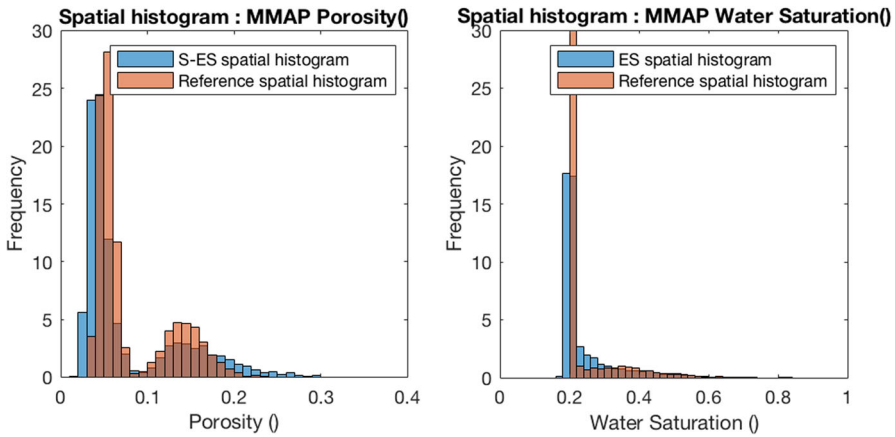


Fig. 11 Posterior statistics of the selection ensemble smoother results compared to the true model

geological layers representing non-permeable shale formations are added above and below the reservoir model. The model then includes three layers (top shale, reservoir, and bottom shale) and two interfaces (top reservoir and bottom reservoir). Seismic data are then generated using the rock physics model with constant mineral properties

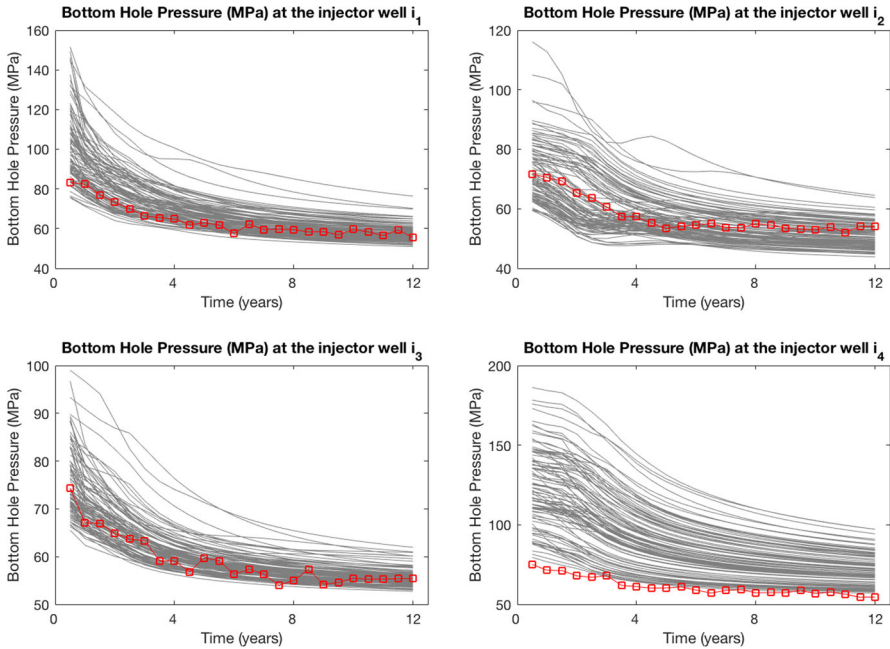


Fig. 12 Predicted bottom hole pressure (BHP) at the 4 injection wells obtained from the selection ensemble smoother; the ensemble members are in black, the measured BHP in red

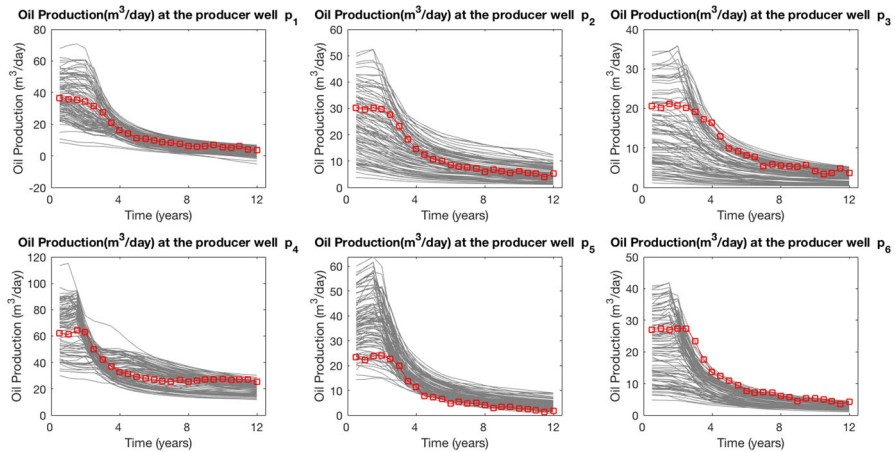


Fig. 13 Predicted oil production rate (OPR) at the 6 producer wells obtained from the selection ensemble smoother; the ensemble members are in grey, the measured OPR in red

in sand and shale and the seismic convolutional operator. The seismic data are calculated for two incident angles of 8° (near) and 24° (far). Figure 6 shows the seismic data computed 12 years after injection and production started. Because the reservoir properties are assumed to be vertically homogeneous within the model, two main

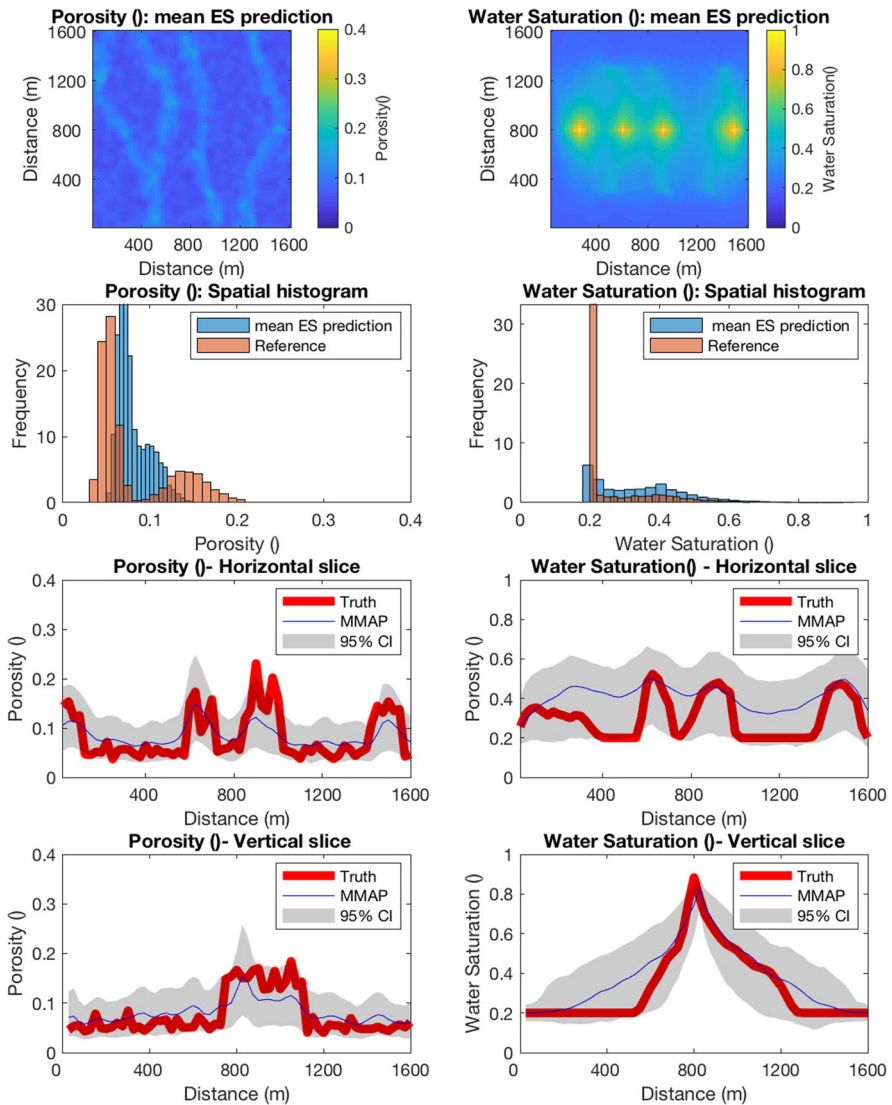


Fig. 14 Ensemble smoother predictions: mean maps, marginal histograms, and predictions along horizontal and vertical lines

reflections are generated in the seismic response, at the top and at the bottom of the reservoir. For this reason, for each trace, the dimension of the data is 4×1 (two angles and two measurements per angle), the dimension of the vector of elastic variables is 9×1 (three properties and three locations per property), and the dimension of the vector of porosity and initial saturation is 6×1 (two properties and three locations per property). The seismic operator is represented by a matrix $G = WAD$ of dimensions

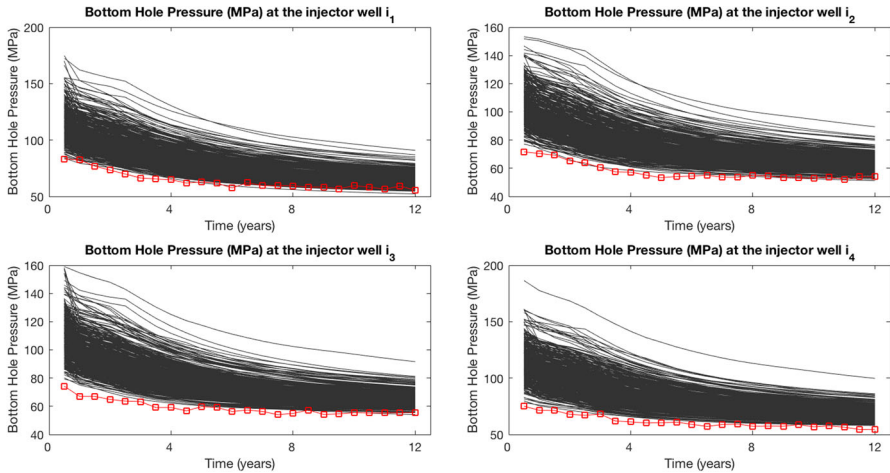


Fig. 15 Predicted bottom hole pressure (BHP) at the 4 injection wells obtained from the ensemble smoother; the ensemble members are in black, the measured BHP in red

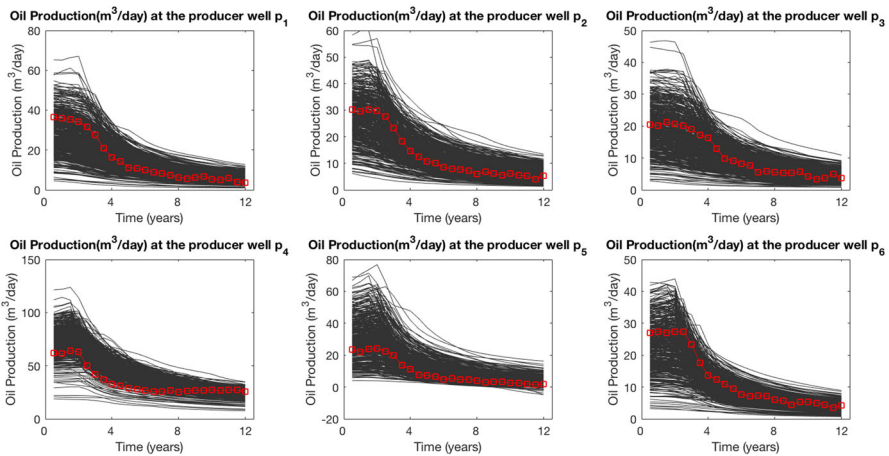


Fig. 16 Predicted oil production rate (OPR) at the 6 producer wells obtained from the ensemble smoother; the ensemble members are in grey, the measured OPR in red

4×9 (\mathbf{W} is 4×4 , \mathbf{A} is 4×6 , \mathbf{D} is 6×9) and the rock physics model is a system of 3 equations (one for each elastic property) with two unknowns (porosity and saturation).

The focus is on the prediction of the porosity field assumed to be spatially variable but constant in time, and the water saturation field before and after the experiment. Following the formalism in Eq. 8, all the measured data, production and seismic are saved in a vector \mathbf{d} that contains OPR and BHP measurements every six months for twelve years and the seismic amplitudes measured after twelve years. Porosity and initial water saturation are stored in \mathbf{r}_0 . The forward model f generates the model predictions according to the discretization of Eqs. 24, 25, 30 and 34. The measured data are perturbed by adding white noise represented by the Gaussian error term δ with

standard deviation 1, $\sqrt{2}/2$ for BHP and OPR respectively. The seismic amplitudes have a signal to noise ratio of approximately 2.5.

In this case study, data assimilation is performed using two methods: S-ES and ES. Figure 2 displays a flowchart summarizing the data assimilation procedure. For the S-ES, the prior distribution of the porosity field ϕ is defined on log-scale as $f(\log(\phi))$ and is selection-Gaussian with parameters

$$\Theta_\phi^{SG} = (\mu_r^\phi, \mu_k^\phi, \sigma_r^\phi, \Sigma_r^\rho, \gamma^\phi, A). \tag{35}$$

The parameters (γ^ϕ, A) are chosen so that the prior marginal is bimodal, with modes approximately matching the average porosity values of the two facies in the reference field. The spatial correlation $(n \times n)$ -matrix Σ_r^ρ is defined by the second order exponential spatial correlation function $\rho(\tau; d_x, d_y) = \exp[-(\frac{\tau_x^2}{d_x^2} + \frac{\tau_y^2}{d_y^2})]$. The parameters values are listed in Table 1. The chosen values for d_x and d_y assume longer correlation along the y -axis than along the x -axis. The prior distribution for the initial water saturation field is Gaussian centered at $\mu_{sw} = 0.2 \times \mathbf{i}_n$ with covariance matrix $\Sigma_{sw} = 0.0005 * \Sigma_r^\rho$, where Σ_r^ρ is the covariance matrix of the porosity field.

An initial ensemble $\mathbf{e}_0 = \{\mathbf{r}_0^{p(i)}, i = 1, \dots, n_e\}$ of size $n_e = 500$ is generated. Covariance localization (Gaspari and Cohn 1999) is used for the conditioning. The localization was chosen restrictive enough to preserve an acceptable rank in the estimated covariance matrix, thereby preventing ensemble collapse. The mean of the realizations and the marginal distributions are shown in Fig. 7. The spatial stationarity is shown in the maps, whereas the marginal distribution is displayed by the histograms, which show the bimodality of the selection-Gaussian prior distribution for porosity.

The posterior distributions of porosity and water saturation are then computed using the S-ES method. Figure 8 shows the posterior mean for the porosity field and the water saturation field after 12 years obtained by the S-ES. Comparing the posterior mean of the porosity field to the reference porosity field, it can be observed that channels are correctly identified with good accuracy. The predicted water saturation field is smoother than the reference water saturation, but the overall pattern is aptly reproduced. Figure 9 shows the prediction and 95% confidence interval for porosity and water saturation along a vertical and horizontal line of the two-dimensional field. Even though the uncertainty is slightly underestimated, the S-ES captures the characteristics of the variations of the porosity and water saturation fields. Inflation was not adopted to show the uncertainty quantification of the methods, but the use of inflation factors might mitigate the underestimation in the uncertainty quantification. Figure 10 shows realizations from the posterior distribution of the porosity field obtained with the S-ES. The channels are correctly predicted with some uncertainties on the boundaries. Figure 11 shows the spatial histograms of the predicted porosity field and water saturation field obtained with the S-ES and compares them the reference values. The marginal maximum a posteriori (MMAP) is used as a predictor due to the marginal bimodality of the posterior ensemble. The spatial histogram of the mean porosity field shows an accurate match with the reference porosity field, whereas the water saturation spatial histogram is slightly smoother than the references one. Both findings are consistent

Table 2 Statistics comparing the S-ES prediction to the ES prediction

	S-ES	ES
MAE $_{\phi}$	1.96×10^{-2}	2.6×10^{-2}
MAE $_{sw}$	6.6×10^{-2}	8.5×10^{-2}
corr $_{\phi}$	0.85	0.82
corr $_{sw}$	0.84	0.75
80% coverage	0.85	0.71

with the analysis of Fig. 8. Figures 12 and 13 show the predicted OPR and BHP for the S-ES respectively, showing a good agreement with the true data.

For the traditional ES approach, the prior ensemble is generated assuming that the prior distributions for the log-porosity and initial water saturation fields are Gaussian. The prior distribution for the porosity field is defined as $f(\log(\phi)) \sim \varphi_n(\log(\phi), \boldsymbol{\mu}_r, \sigma_r^2 \times \boldsymbol{\Sigma}_r^{\rho})$, where $\boldsymbol{\mu}_r = -2.5\mathbf{i}_n$ and $\sigma_r^2 = 0.2$ are chosen such that the marginal prior distribution approximates the spatial histogram of the reference porosity field. The covariance matrix $\boldsymbol{\Sigma}_r^{\rho}$ and the prior distribution for the initial water saturation field are the same as in the S-ES case. Figure 14 shows the results for the posterior ensemble for the porosity field and water saturation field obtained using the ES. The posterior mean of the porosity field seems to identify the channels but the porosity values are underestimated within the channels. The predicted water saturation is much smoother than the reference water saturation field and the prediction from the S-ES. The spatial histogram of the mean porosity field shows that the ES fails to capture the bimodal nature of the porosity field, while the spatial histogram of the water saturation field severely underestimates the number of locations where saturation remains equal to the initial value of 0.2 and overpredicts saturation in several locations. By analyzing the predictions of porosity and water saturation along the horizontal and vertical lines, the ES prediction fails to capture the sharp changes, even though the coverage of the confidence intervals is satisfactory. Figures 15 and 16 display the predicted OPR and BHP for the ES, showing a better match for OPR and a worse match for BHP compared to the S-ES results.

The two methods are compared using different statistics: the mean absolute error (MAE), the correlation, and the 80% coverage probability are considered. Let MAE $_{\phi}$ and corr $_{\phi}$ define, respectively, the MAE and the correlation between the predicted porosity field and the reference porosity field. Further, let MAE $_{sw}$ and corr $_{sw}$ define, respectively, the MAE and the correlation between the predicted water saturation field and the reference water saturation at the final time step. The 80% coverage probability is calculated for the predicted porosity field. The results are detailed in Table 2. The MAE measures for the ES are about 30% higher than those of the S-ES. The correlations for the porosity are comparable. The correlation for the water saturation for the S-ES is about 10% higher than that of the ES. The 80% coverage probability seems to indicate the ES underestimates the uncertainty more than the S-ES overestimates it. Overall these statistics are favorable to the S-ES. The sensitivity to the data does not seem to cause over-fitting. The use of multiple data assimilation and iterative methods could improve the data match for both the ensemble smoother

and selection ensemble smoother. The S-ES has a higher computational demand than the ES because the posterior distribution needs to be sampled. On a 64×64 grid, it takes a few minutes to generate 200 samples, which is considered acceptable in view of the results. When the dimension of the data is very large, such as for field case reservoir models defined on 3D grids with millions of nodes, the matrix $\hat{\Sigma}_d$ in Eqs. 9 and 15 is too large to be stored, let alone inverted. In that case, one could propose a parametric approach where the covariance is defined by a few model parameters and maximum likelihood estimation is used to estimate those parameters (Skauvold and Eidsvik 2019). Alternatively, a methodological development along the lines of the parametric Kalman filter could be considered (Pannekoucke et al. 2016).

4 Conclusions

Seismic and production data assimilation is performed to predict the spatial distribution of porosity and water saturation using a novel method named 'selection ensemble smoother', a Monte Carlo ensemble method that extends the ensemble smoother to selection-Gaussian models. By using a selection-Gaussian prior and an augmented initial state vector, the posterior distribution can represent multimodal posterior distributions. The main advantage of the selection ensemble smoother is the use of the augmented state vector which makes possible conditioning on the latent variable after data assimilation. The posterior distribution is then sampled assuming that the model is selection-Gaussian, thereby allowing multimodal features in the posterior distribution. The design of the experiment is optimal for the selection ensemble smoother approach since the channelized system leads to a bimodal porosity distribution. The results from the selection ensemble smoother are promising since the predictions are accurate and the data predictions match the measurements. The limited band-width of the seismic measurements, which limits the resolution of the data, prevents the inversion from clearly identifying the non-stationarity in the porosity field but still allows the selection ensemble smoother method to select the right mode, whereas the ensemble smoother results regress towards the mean.

Acknowledgements Authors acknowledge the Uncertainty in Reservoir Evaluation (URE) activity at the Norwegian University of Science and Technology (NTNU), the School of Energy Resources and the Department of Geology and Geophysics of the University of Wyoming for their support. Authors acknowledge Mingliang Liu (University of Wyoming) for developing the case study, Henning Omre (NTNU) for the helpful comments and suggestions, and SINTEF for providing the Matlab Reservoir Simulation Toolbox MRST.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital).

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

- Aanonsen SI, Aavatsmark I, Barkve T, Cominelli A, Gonard R, Gosselin O, Kolasinski M, Reme H et al (2003) Effect of scale dependent data correlations in an integrated history matching loop combining production data and 4D seismic data. In: Society of Petroleum Engineers, SPE reservoir simulation symposium
- Aki K, Richards P (1980) Quantitative seismology: theory and methods. W. H. Freeman & Co, San Francisco
- Arellano-Valle R, del Pino G (2004) From symmetric to asymmetric distributions: a unified approach. In: Genton M (ed) Skew-elliptical distributions and their applications: a journey beyond normality. Chapman & Hall / CRC, New York, pp 113–133
- Arellano-Valle RB, Branco MD, Genton MG (2006) A unified view on skewed distributions arising from selections. *Can J Stat* 34(4):581–601
- Bertino L, Evensen G, Wackernagel H (2003) Sequential data assimilation techniques in oceanography. *Int Stat Rev* 71(2):223–241
- Buland A, Omre H (2003) Bayesian linearized AVO inversion. *Geophysics* 68(1):185–198
- Chen Y, Oliver DS (2011) Localization of ensemble-based control-setting updates for production optimization. *SPE J* 17(01):122–136
- Chen Y, Oliver DS (2012) Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Math Geosci* 44(1):1–26
- Conjard M, Omre H (2020) Data assimilation in spatio-temporal models with non-Gaussian initial states-the selection ensemble Kalman model. *Appl Sci* 10(17):5742
- Dadashpour M, Echeverria Ciaurri D, Mukerji T, Kleppe J, Landrø M (2010) A derivative-free approach for the estimation of porosity and permeability using time-lapse seismic and production data. *J Geophys Eng* 7(4):351–368
- Dong Y, Oliver DS et al (2005) Quantitative use of 4D seismic data for reservoir description. *SPE J* 10(01):91–99
- Dovera L, Della Rossa E (2011) Multimodal ensemble Kalman filtering using Gaussian mixture models. *Comput Geosci* 15(2):307–323
- Dvorkin J, Gutierrez MA, Grana D (2014) Seismic reflections of rock properties. Cambridge University Press, Cambridge
- Emerick AA, Reynolds AC (2013) Ensemble smoother with multiple data assimilation. *Comput Geosci* 55:3–15
- Evensen G (2009) Data assimilation: the ensemble Kalman filter. Springer, Berlin
- Gaspari G, Cohn S (1999) Construction of correlation functions in two and three dimensions. *QJR Meteorol Soc* 125(554):723–757
- Gineste M, Eidsvik J, Zheng Y (2020) Ensemble-based seismic inversion for a stratified medium. *Geophysics* 85(1):29–39
- Houtekamer PL, Mitchell HL (1998) Data assimilation using an ensemble Kalman filter technique. *Mon Weather Rev* 126(3):796–811
- Huang X, Meister L, Workman R et al (1997) Reservoir characterization by integration of time-lapse seismic and production data. In: SPE annual technical conference and exhibition, 38695, Society of Petroleum Engineers
- Kennett BLN (1984) Guided wave propagation in laterally varying media, I. Theoretical development. *Geophys J Int* 79(1):235–255
- Kretz V, Ravalec-Dupin L, Roggero F, et al (2002) An integrated reservoir characterization study matching production data and 4D seismic. In: SPE annual technical conference and exhibition, 77516, Society of Petroleum Engineers
- Kumar D, Srinivasan S (2018) Ensemble-based assimilation of nonlinearly related dynamic data in reservoir models exhibiting non-Gaussian characteristics. *Math Geosci* 51:75–107
- Landa J L, Horne R N, et al. (1997) A procedure to integrate well test data, reservoir performance history and 4-D seismic information into a reservoir description. In: SPE annual technical conference and exhibition, 38653, Society of Petroleum Engineers

- Landrø M (2001) Discrimination between pressure and fluid saturation changes from time-lapse seismic data. *Geophysics* 66(3):836–844
- Landrø M, Veire HH, Duffaut K, Najjar N (2003) Discrimination between pressure and fluid saturation changes from marine multicomponent time-lapse seismic data. *Geophysics* 68(5):1592–1599
- Lie KA (2019) An introduction to reservoir simulation using MATLAB/GNU Octave: User guide for the MATLAB reservoir simulation toolbox (MRST). Cambridge University Press, Cambridge
- Liu M, Grana D (2018) Ensemble-based seismic history matching with data reparameterization using convolutional autoencoder. In: SEG technical program expanded abstracts 2018, Society of Exploration Geophysicists, pp 3156–3160
- Liu M, Grana D (2020) Time-lapse seismic history matching with an iterative ensemble smoother and deep convolutional autoencoder. *Geophysics* 85(1):15–31
- Myrseth I, Omre H et al (2010) Hierarchical ensemble Kalman filter. *SPE J* 15(02):569–580
- Nævdal G, Johnsen L M, Aanonsen S I, Vefring E H, et al. (2003) Reservoir monitoring and continuous model updating using ensemble Kalman filter. In: SPE annual technical conference and exhibition, 84372, Society of Petroleum Engineers
- Naveau P, Genton MG, Shen X (2005) A skewed Kalman filter. *J Multivar Anal* 94(2):382–400
- Oliver D, Chen Y (2018) Data assimilation in truncated Plurigaussian models: impact of the truncation map. *Math Geosci* 50:867–893
- Oliver DS, Chen Y (2011) Recent progress on reservoir history matching: a review. *Comput Geosci* 15(1):185–221
- Oliver DS, Reynolds AC, Liu N (2008) Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, Cambridge
- Omre H, Rimstad K (2018) Bayesian spatial inversion and conjugate selection Gaussian prior models. arXiv preprint [arXiv:1812.01882](https://arxiv.org/abs/1812.01882)
- Pannekoucke O, Ricci S, Barthelemy S, Ménard R, Thual O, Tellus A (2016) Parametric Kalman filter for chemical transport models. *Dyn Meteorol Oceanogr* 68(1):31547
- Rezaie J, Eidsvik J (2014) Kalman filter variants in the closed skew normal setting. *Comput Stat Data Anal* 75:1–14
- Russell B H (1988) Introduction to seismic inversion methods. Society of Exploration Geophysicists
- Sætrom J, Omre H (2011) Ensemble Kalman filtering with shrinkage regression techniques. *Comput Geosci* 15(2):271–292
- Sheriff RE, Geldart LP (1995) Exploration seismology, 2nd edn. Cambridge University Press, Cambridge
- Shuey RT (1985) A simplification of the Zoeppritz equations. *Geophysics* 50(4):609–614
- Silva D, Deutsch C (2019) Multivariate categorical modeling with hierarchical truncated pluri-Gaussian simulation. *Math Geosci* 51:527–552
- Skauvold J, Eidsvik J (2019) Parametric spatial covariance models in the ensemble Kalman filter. *Spatial Stat* 29:226–242
- Skjervheim J A, Evensen G, Aanonsen S I, Ruud B O, Johansen T A, et al. (2005) Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter. In: SPE annual technical conference and exhibition, 95789, Society of Petroleum Engineers
- Thurin J, Brossier R, Métivier L (2019) Ensemble-based uncertainty estimation in full waveform inversion. *Geophys J Int* 219(3):1613–1635
- Trani M, Arts R, Leeuwenburgh O et al (2012) Seismic history matching of fluid fronts using the ensemble Kalman filter. *SPE J* 18(01):159–171
- Williams M, Keating J, Barghouty M et al (1998) The stratigraphic method: a structured approach to history matching complex simulation models. *SPE Reserv Eval Eng* 1(02):169–176
- Yilmaz Ö (2001) Seismic data analysis: processing, inversion, and interpretation of seismic data. Society of Exploration Geophysicists
- Zhou H, Li L, Franssen HJ, Gomez-Hernandez J (2012) Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman filter. *Math Geosci* 44:1–17
- Zoeppritz K (1919) On the reflection and propagation of seismic waves. *Gottinger Nachrichten* 1(5):66–84