

Petrophysical characterization of deep saline aquifers for CO₂ storage using ensemble smoother and deep convolutional autoencoder

Mingliang Liu*, Dario Grana

Department of Geology and Geophysics, School of Energy Resources, University of Wyoming, 1000 E. University Ave. Laramie, WY 82071, United States



ARTICLE INFO

Keywords:
 CO₂ storage
 CO₂ plume prediction
 Data assimilation
 Deep learning
 Deep saline aquifers

ABSTRACT

Carbon dioxide sequestration in deep saline aquifers requires accurate and precise methods to monitor carbon capture and storage and detect leakage risks. The assessment of the CO₂ plume location during injection and storage depends on the accuracy of the spatial distribution model of petrophysical properties, such as porosity and permeability. This work focuses on stochastic methods for petrophysical characterization and presents a method for the prediction of porosity and permeability using borehole observations and surface geophysical data. This study utilizes injection and monitoring measurements at the borehole locations and time-lapse seismic surveys. The proposed method is based on a stochastic approach to inverse problems with data assimilation, namely the ensemble smoother with multi-data assimilation. Ensemble-based methods are generally unfeasible when applied to large geophysical datasets, such as time-lapse seismic surveys. In the proposed approach, a machine learning method, namely the deep convolutional autoencoder, is applied to reduce the dimension of the seismic data. The ensemble smoother is then applied in a lower dimensional data space to predict the aquifer petrophysical properties. This method updated predictions of porosity and permeability every time new data, either seismic surveys or borehole data, are available, to reduce the uncertainty in the CO₂ plume prediction. The method is tested and validated on a synthetic geophysical dataset generated for the Johansen formation, a potential large-scale offshore site for CO₂ storage.

1. Introduction

CO₂ sequestration and storage in deep saline aquifers has been widely studied (Kopp et al., 2010; Goodman et al., 2013; Castelletto et al., 2013; Pool et al., 2013; Jeong et al., 2013; Ghorbanidehno et al., 2015; Li and Benson, 2015; Levine et al., 2016; Ma et al., 2019; Zhu et al., 2019). One of the main challenges for a successful implementation of carbon storage in deep saline formations is the ability to reliably monitor the CO₂ plume migration and the dynamic conditions during and after injection (Goodman et al., 2011 and 2017; Bachu, 2008 and 2015; Miller et al., 2014). Due to the complexity and heterogeneity of geological structures in the subsurface and the lack of spatially exhaustive datasets with direct measurements of rock and fluid properties, the assessment of the potential storage and forecasting of the dynamic conditions, such as CO₂ plume location, are uncertain. One of the main sources of uncertainty is the lack of knowledge of the spatial distributions of rock and fluid properties in the aquifer (Li et al., 2011).

If the aquifer spatial distributions of rock and fluid properties are known; then, numerical simulations of fluid flow through porous media allow predicting the fluid and pressure conditions during and after injection. However, the distribution of rock and fluid properties in the

reservoirs cannot be exactly determined due to the lack of direct measurements, noise in the indirect measurements (such as well logs and surface geophysical data), the approximations in the physical models, and the natural heterogeneity of the reservoir rocks.

Geostatistical realizations and fluid flow simulations have been previously combined to study the uncertainty in reservoir models and predictions in petroleum geoscience (Oliver et al., 2008; Mohamed et al., 2010; Caeiro et al., 2015) and hydrogeology (Capilla et al., 1997; Gómez-Hernández et al., 1997; Kitanidis, 1997; Zimmerman et al., 1998). The application of these methods to carbon storage studies is not completely new (Szulczeński and Juanes, 2009; Bellenfant et al., 2009; Bergmo et al., 2011; Gasda et al., 2012; Szulczeński et al., 2012; Deng et al., 2012; Bhownik et al., 2013; Castelletto et al., 2013; Jeong et al., 2013; Li and Benson, 2015; Levine et al., 2016). Risk assessment in storage capacity has been investigated in several works (Wilson et al., 2003; Zhou et al., 2008; Sifuentes et al., 2009; Burruss et al., 2009; Brennan et al., 2010; Nordbotten et al., 2012; Sun et al., 2013; Ellett et al., 2013; Peck et al., 2014; Gorecki et al., 2015; Jung et al., 2018). However, accurate analysis of all the sources of uncertainty in the pre-injection (static) model and complete studies

* Corresponding author.

E-mail address: mliu4@uwyo.edu (M. Liu).

on the propagation of the uncertainty in the model predictions are still missing in literature.

Several geophysical measurements have been utilized for static reservoir characterization and reservoir monitoring, including 3D prestack seismic data, electromagnetic data, gravity data, and time-lapse seismic data (Chadwick et al., 2005; Chadwick et al., 2006; Chadwick et al., 2009; Alnes et al., 2011; Grude et al., 2013; Ivandic et al., 2015; Grana et al., 2017; Roach and White, 2018; Wang et al., 2018; Glubokovskikh et al., 2019). The above-mentioned studies focus on the location of the CO₂ plume; several works focusing on the pressure front extent are also available (Nordbotten and Celia, 2010; Oruganti et al., 2011; Senel and Chugunov, 2013; Mathias et al., 2013). Finally, some of the published works also include the integration of studies of enhanced oil recovery (Liu et al., 2013; Dai et al., 2014), water production to optimize the aquifer storage capacity (Bergmo et al., 2011), and history matching for reservoir updating (Chadwick and Noy, 2010). Most of the presented methods have been tested in depleted reservoirs and some of the most successful and documented applications are located in the North Sea, in particular the Sleipner field and the Johansen field (Arts et al., 2008; Eigestad et al., 2009; Wei and Saaf, 2009; Bergmo et al., 2009a; Bergmo et al., 2009b; Chadwick and Noy, 2010; Gasda et al., 2012; Sundal et al., 2013).

Different methods, including stochastic optimization and sampling algorithms, can be adopted to update the static reservoir model of rock and fluid properties and make the dynamic model more predictive for future forecasting (Bellenfant et al., 2009; Tavakoli et al., 2013; Sun et al., 2013; Grana et al., 2017). In particular, data assimilation algorithms are commonly used for history matching in petroleum engineering. Two of the main common algorithms for history matching of reservoir data are the Ensemble Kalman filter and the Ensemble Smoother (Evensen, 2009; Emerick and Reynolds, 2013). These methods have been applied to production and geophysical data (Chen and Oliver 2012 and 2017; Luo et al., 2016, 2018; Liu and Grana 2018; Lorentzen et al., 2019; Canchumuni et al., 2019a, 2019b; Ma et al., 2019). Machine learning methods have also been used to solve the data assimilation problem (Tahmasebi et al., 2018; Laloy et al., 2018; Chen et al., 2018; Liu and Grana, 2019; Liu et al., 2019; Sun and Durlofsky, 2019; Luo, 2019; Puzyrev, 2019; Etienam, 2019).

The goal of this work is to present a mathematical method to improve the aquifer characterization in terms of petrophysical properties and their spatial distribution. We propose to integrate injection data and well observations with time-lapse seismic surveys to update the static model of porosity and permeability every time new data (seismic or borehole data) are measured and improve the accuracy and precision of the predictive fluid flow model. The advantage of the proposed approach is that it provides more reliable estimates of the CO₂ plume migration thanks to an improved understanding and characterization of the spatial petrophysical model. The mathematical method is based on a stochastic approach to data assimilation, namely the Ensemble Smoother, and it combines geostatistical realizations of spatial property models, fluid flow simulations, geophysical modeling of measured surface seismic data, and machine learning algorithms. In particular, we generate multiple static (pre-injection) models of reservoir properties to obtain a set of potential reservoir models that account for all the possible geological scenarios, i.e. all the possible combinations of properties that cannot be excluded by geological prior knowledge. This ensemble of models represents the uncertainty in the reservoir model before data acquisition. We then update the reservoir model using time-lapse seismic data and monitoring wells. The result is a set of equiprobable petrophysical models that allows the prediction and uncertainty quantification of aquifer conditions at every time of the CO₂ storage process, before, during, and after injection. In particular, this assessment allows predicting the uncertainty in the storage capacity as well as in the CO₂ plume location. The methodology accounts for several sources of information. The accurate quantification of these properties for carbon storage studies allows successful risk analysis and decision making for CO₂ storage studies.

The dataset in this study is the Johansen formation, located offshore Norway, (Bergmo et al., 2009a, 2009b and 2011). The static and dynamic fluid flow simulation models have been presented in Bergmo et al., 2011. In this work, a synthetic time-lapse seismic dataset was created to demonstrate the methodology and prove the validity of stochastic data assimilation methods for improving the petrophysical characterization of deep saline aquifers and the value of information of time-lapse seismic data for reservoir monitoring of the CO₂ plume migration.

2. Methodology

We aim to predict porosity and permeability from borehole injection and monitoring data and time-lapse seismic data. The modeling can be formulated as an inverse problem

$$\mathbf{d} = \mathbf{H}(\mathbf{m}) + \mathbf{e} \quad (1)$$

where \mathbf{d} includes injection and monitoring borehole data and time-lapse seismic surveys, \mathbf{m} is the spatial petrophysical model of porosity and permeability, and \mathbf{H} is the forward operator including rock physics, seismic wave propagation, and fluid flow simulation (Tarantola, 2005; Oliver et al., 2008).

The goal is to study the most likely value and the uncertainty in the predictions of model variables \mathbf{m} based on the measured data \mathbf{d} . Therefore, we aim to compute the posterior distribution of the model properties \mathbf{m} given the observed data \mathbf{d} . The probability $p(\mathbf{m}|\mathbf{d})$ indicates the posterior density function of the model properties \mathbf{m} given the data \mathbf{d} and it is computed as

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \quad (2)$$

where $p(\mathbf{m})$ is the prior distribution of the model properties \mathbf{m} , $p(\mathbf{d}|\mathbf{m})$ is the likelihood function that predicts the probability of the data \mathbf{d} of being observed as a response of the model properties \mathbf{m} , and $p(\mathbf{d})$ is a normalizing constant, equal to the integral $\int p(\mathbf{d}|\mathbf{m})p(\mathbf{m})d\mathbf{m}$.

2.1. Forward operator

In this section, we describe the physical models linking the model parameters (porosity and permeability) to the measured borehole and geophysics data.

The geophysics operator includes a rock physics model to predict elastic properties such as P- and S-wave velocity and density from petrophysical properties (porosity, mineralogy, and saturations) and the seismic wave propagation model to predict travel time and seismic wave amplitudes from the elastic properties. Aquifer rocks might have different mineral fractions, porosity and fluid saturation, and consequently generate different elastic responses. The elastic variables of interest \mathbf{v} are generally expressed as a function of rock and fluid properties \mathbf{p}

$$\mathbf{v} = \mathbf{g}(\mathbf{p}), \quad (3)$$

where \mathbf{g} is the rock-physics model. There is an extensive literature on rock-physics modeling for different rock types, such as shale, poorly and highly consolidated sandstone, and carbonate. Given that the target formation of this CO₂ storage study consists of sandstones, the adopted rock-physics model in this study is the soft sandstone model (Mavko et al., 2009). In this model, at each location in the model, the matrix bulk and shear moduli are computed using the Voigt-Reuss-Hill average whereas the bulk and shear moduli of the dry (unsaturated) rock at the critical porosity φ_0 are obtained using the Hertz-Mindlin equations (Appendix A). The moduli of the dry rock with variable porosity φ in the range $[0, \varphi_0]$ are estimated by interpolating the two end members (matrix and dry rock at critical porosity) using the modified Hashin-Shtrikman lower bounds (Appendix A). Finally, by applying Gassmann's fluid model (Gassmann, 1951), we can obtain the bulk and shear moduli of the saturated-rock (K_{sr} and μ_{sr}) and compute the P- and S-wave

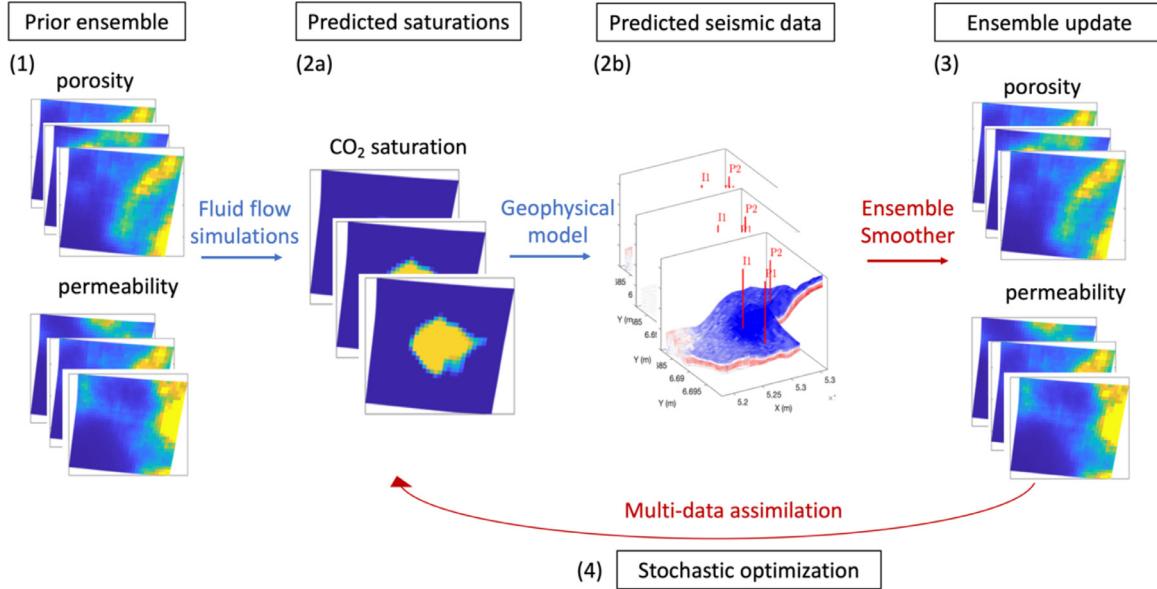


Fig. 1. Flowchart of ES-MDA algorithm for porosity and permeability prediction.

velocity by definition:

$$V_p = \sqrt{\frac{K_{sr} + \frac{4}{3}\mu_{sr}}{\rho_{sr}}} \quad (4)$$

$$V_s = \sqrt{\frac{\mu_{sr}}{\rho_{sr}}} \quad (5)$$

where ρ_{sr} is the density of the saturated rock and it is computed as the weighted linear average of the matrix density ρ_m and fluid density ρ_f :

$$\rho_{sr} = (1 - \phi)\rho_m(f_m) + \phi\rho_f(s_f) \quad (6)$$

where ϕ is the porosity of the rock and the matrix density ρ_m and fluid density ρ_f depend on the mineral fractions f_m and fluid saturations s_f , respectively. We can then predict the seismic response through a simplified seismic wave propagation model. In this work, we assume that the geophysical data are time-lapse post-stack seismic $s(t)$, measured in time domain, generated by the convolutional model

$$s(t) = \mathbf{w}(t) * \mathbf{r}(t, \mathbf{v}(t)) + \mathbf{e} \quad (7)$$

where $\mathbf{w}(t)$ is the seismic wavelet (assumed to be known), \mathbf{r} represents the reflection coefficients calculated from the elastic properties $\mathbf{v}(t)$ and \mathbf{e} represents the measurement errors. Therefore, assuming that the mineral fractions are homogeneous in the aquifer, the seismic data are modeled as the response of the spatial rock and fluid property model

$$s(x, y, t) = \mathbf{G}(\phi(x, y, z), s_f(x, y, z)) + \mathbf{e} \quad (8)$$

where x, y, z are the spatial locations of the model and \mathbf{G} is the geophysical operator.

The fluid flow operator includes equations for fluid flow in porous media to predict the saturation and pressure at each time step given the initial petrophysical model and injection parameters, which is commonly refer to as reservoir simulation. In particular, the CO₂ injection and migration in a deep saline aquifer can be modeled as a two-phase miscible flow of the brine-CO₂ system that can be numerically solved using the black-oil framework (Eigestad et al., 2009; Lie, 2016). To investigate the long-term CO₂ migration after injection (for instance, hundreds of years), the standard three-dimensional simulation is computationally prohibitive. For this reason, alternative simplified approaches have been developed with the assumption that the flow system is in vertical equilibrium (VE; Nilsen et al., 2011) so that the analytical expressions

of the vertical distribution of fluid phases are available. In such simplified models, the three-dimensional fluid flow simulation is reduced to two-dimensional problems while correctly approximating the three-dimensional behavior. Consequently, the so-called VE models largely speeds up the computational time (Nilsen et al., 2011). In this work, we adopt the MRST-co2lab (Lie, 2016) in which a wide class of vertical-equilibrium models have been implemented, to simulate the process of CO₂ injection and migration. In general, given a spatial model of rock properties (porosity $\phi(x, y, z)$ and permeability $\mathbf{k}(x, y, z)$) and initial fluid parameters at time T_0 in the aquifer and assuming that rock properties do not change over time T , the simulation returns the spatial model of saturation and pressure conditions at any given time T

$$[s_f(x, y, z, T), p_f(x, y, z, T)] = \mathbf{F}(\phi(x, y, z), \mathbf{k}(x, y, z), s_f(x, y, z, T_0), p_f(x, y, z, T_0)). \quad (9)$$

The fluid flow model \mathbf{F} in Eq. 9 is used to predict the saturation conditions at the monitoring well locations, during and after injection.

In our approach, we assume that the initial fluid parameters at time T_0 in the aquifer are constant and known (fully brine-saturated aquifer with constant fluid pressure), whereas rock properties (porosity $\phi(x, y, z)$ and permeability $\mathbf{k}(x, y, z)$) are unknown.

2.2. Inverse problem

In this section, we propose a mathematical approach for the solution of the inverse problem in Eq. 1 where the unknown model variables \mathbf{m} are porosity and permeability, the observed data \mathbf{d} are the surface seismic data and the injection and monitoring data at the well locations, and the forward operator \mathbf{H} includes the rock physics, seismic, and fluid flow models in Eqs. 4-9.

The method we adopt is the Ensemble Smoother with Multi-Data Assimilation (ES-MDA, Emerick and Reynolds, 2013). In particular, we generate an initial set of petrophysical models \mathbf{m} (porosity and permeability) and we update the ensemble using the ES-MDA to match the measured data \mathbf{d} (borehole and time-lapse seismic data). The ensemble of updated models provides an approximation of the posterior distribution of the model variables. The large dimension of the geophysical dataset would make the application of the ES-MDA unfeasible, because it would require a large amount of petrophysical models in the initial ensemble to avoid the ensemble collapse. For this reason, we propose

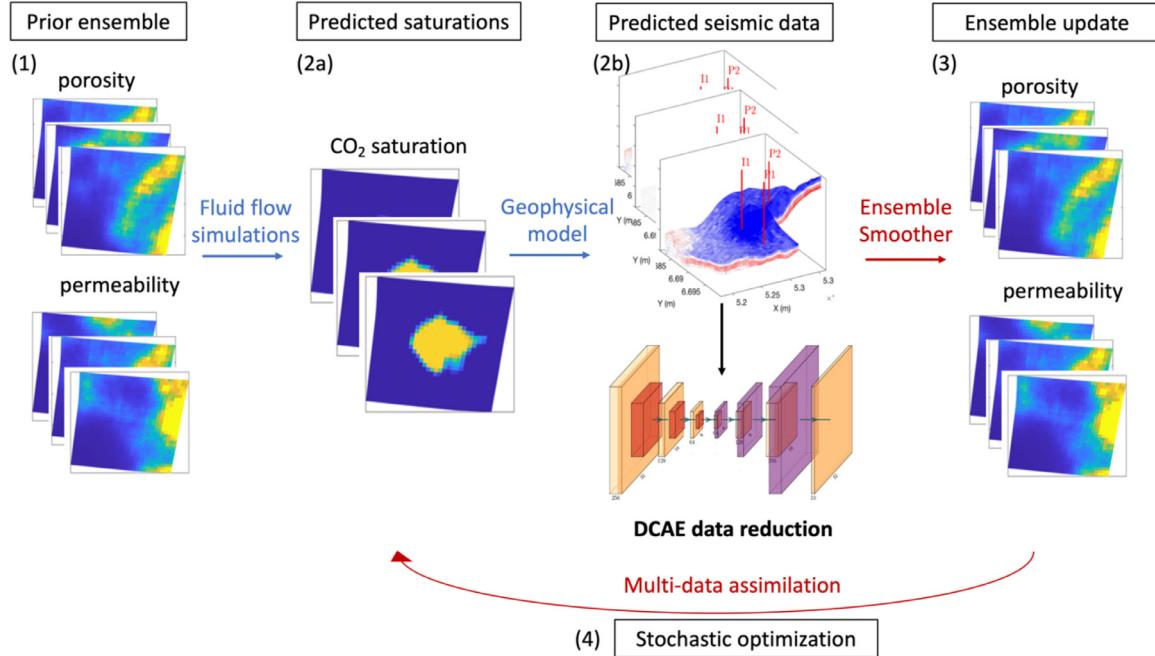


Fig. 2. Flowchart of ES-MDA algorithm with data reduction using the deep convolutional autoencoder for porosity and permeability prediction.

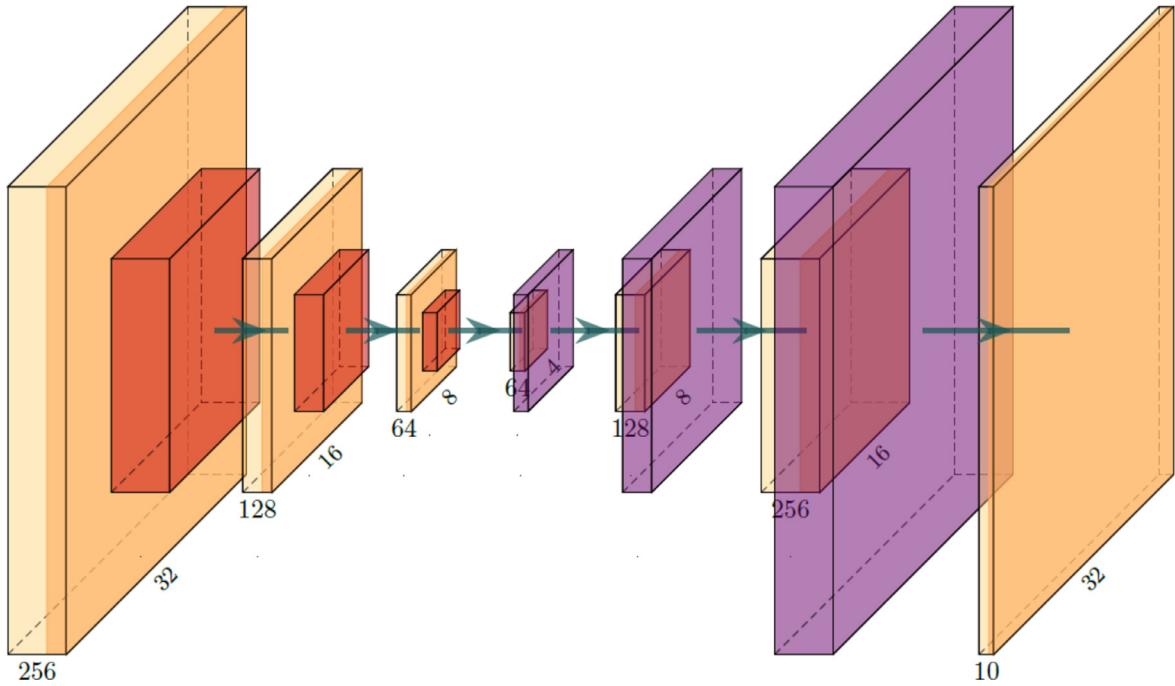


Fig. 3. Architecture of the DCAE: Orange blocks represent the convolutional layers; red blocks represent the max-pooling layers; and purple blocks represent the up-pooling layers. The numbers of thickness and width of the layers are the number of channels and data dimensions, respectively.

to apply the deep convolutional autoencoder (DCAE; Goodfellow et al., 2016) to reduce the dimension of the data space.

In ES-MDA, the model variables are assumed to be Gaussian defined in the domain $(-\infty, +\infty)$. When model variables are defined in a bounded domain, a logit transformation can be applied to map them from the domain $[m_{min}, m_{max}]$ to the domain $(-\infty, +\infty)$, apply the ES-MDA, and back transform the results in the original model space after updating (Liu and Grana, 2018). Because the inverse problem is not linear, the mean and covariance matrices cannot be analytically computed. In the ES-MDA, the covariance matrices of the model and of the

data and their cross covariance are approximated using the sample covariance matrix of the model and prediction ensemble.

The algorithm includes the following steps:

1. We define the number N_e of models in the ensemble and a sequence of inflation coefficients $\{\alpha_i\}_{i=1, \dots, N}$ with $\sum_{i=1}^N \frac{1}{\alpha_i} = 1$, where N is the number of data assimilations (Emerick and Reynolds, 2013). We generate an ensemble of N_e prior realizations $\{\mathbf{m}_j^i\}_{j=1, \dots, N_e}$ of the model properties using geostatistical algorithms, where the superscript i indicates the iteration number, that is equal to 1 for the initial iteration (Fig. 1, step 1).

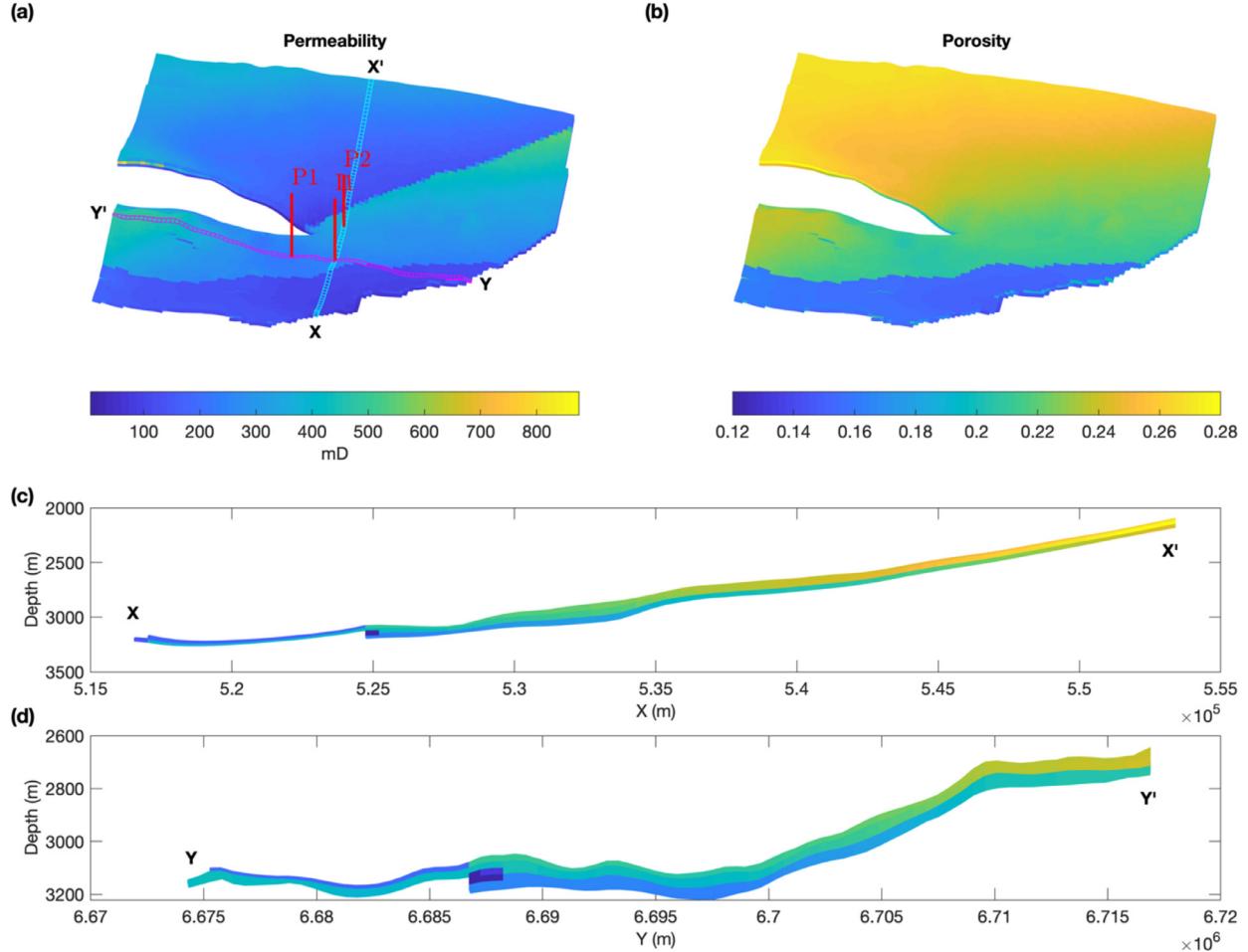


Fig. 4. Johansen model: (a) permeability; (b) porosity; (c) vertical section of the porosity model crossing I1 and P2; (d) vertical section of the porosity model crossing I1 and P1.

2. We then run the fluid flow simulation (Eq. 9) to obtain the saturation distributions for each ensemble model at each time (Fig. 1, step 2a) and we compute the corresponding the geophysical response (Eq. 8) to predict the seismic data $\{d_j^i\}_{j=1,...,N_e}$ (Fig. 1, step 2b) and apply a perturbation $d_{p_j}^i = d_j^i + \sqrt{\alpha_i} \Sigma_d^{1/2} z_{p_j}^i$ to each model prediction, where Σ_d is the covariance matrix of the error measurements and $z_{p_j}^i \sim N(\mathbf{0}, \mathbf{I}_n)$ for $j = 1, \dots, N_e$, i.e. $z_{p_j}^i$ is distributed according to a Gaussian distribution with zero mean covariance matrix \mathbf{I}_n with \mathbf{I}_n being identity matrix of size $n \times n$ and n being the number of data measurements.
3. We update the model ensemble (Fig. 1, step 3) according to the following expression

$$\mathbf{m}_j^{i+1} = \mathbf{m}_j^i + \mathbf{K}^i (\mathbf{d}_{p_j}^i - \mathbf{d}_j^i) \quad (10)$$

for $j = 1, \dots, N_e$ (where the superscript i indicates the iteration). In Eq. 10, \mathbf{K}^i is computed as:

$$\mathbf{K}^i = \Sigma_{m,d}^i (\Sigma_{d,d}^i + \alpha_i \Sigma_d)^{-1} \quad (11)$$

where $\Sigma_{m,d}^i$ is the $m \times n$ cross-covariance matrix of models \mathbf{m}^i and predicted data \mathbf{d}^i (where m is the number of model variables) and $\Sigma_{d,d}^i$ is the $n \times n$ covariance matrix of the predicted data \mathbf{d}^i :

$$\Sigma_{m,d}^i = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} (\mathbf{m}_j^i - \bar{\mathbf{m}}^i)(\mathbf{d}_j^i - \bar{\mathbf{d}}^i) \quad (12)$$

$$\Sigma_{d,d}^i = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} (\mathbf{d}_j^i - \bar{\mathbf{d}}^i)(\mathbf{d}_j^i - \bar{\mathbf{d}}^i) \quad (13)$$

with $\bar{\mathbf{m}}^i$ and $\bar{\mathbf{d}}^i$ being the means of the models and data predictions of the ensemble.

4. We repeat steps 2 and 3 for N iterations (Fig. 1, step 4).

For the geostatistical simulations in step 1, we use the Fast Fourier Transform – Moving Average method (FFT-MA); however other algorithms such as sequential Gaussian simulations could be used. For bounded properties, logit transformation or normal score transformations could be applied to map the properties in the set of real numbers. For highly bimodal properties, sequential Gaussian mixture simulations or multi-point statistics algorithms (Mariethoz and Caers, 2014) can be adopted. The forward model in step 2 combines the fluid flow simulations (step 2a) and the seismic forward model (step 2b). The saturation model predictions of the initial porosity and permeability models are obtained using fluid flow simulations based on the vertical-equilibrium as in Nilsen et al., (2011); however, other approaches such as percolation-based simulators could be used. The elastic and seismic response of the saturation models are computed using a rock physics model and seismic convolutional operators; poroelastic and wave equations could also be applied.

In our application, the number of model variables m is the number of unknown porosity and permeability values and the number of data measurements n is the number of borehole observations plus the number of seismic data measures. For example, for an aquifer model with 100,000

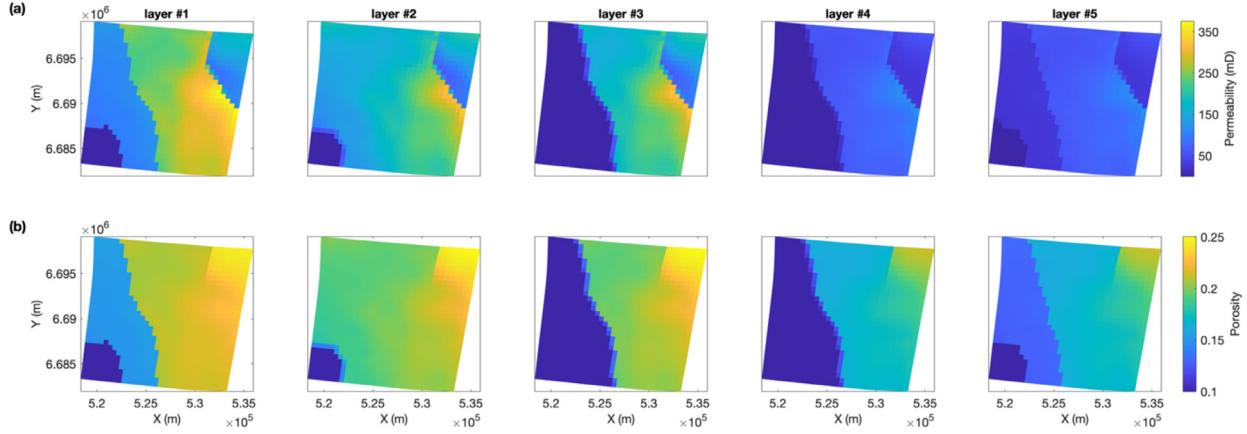


Fig. 5. Reduced model of interest: (a) permeability; (b) porosity of each of the 5 layers.

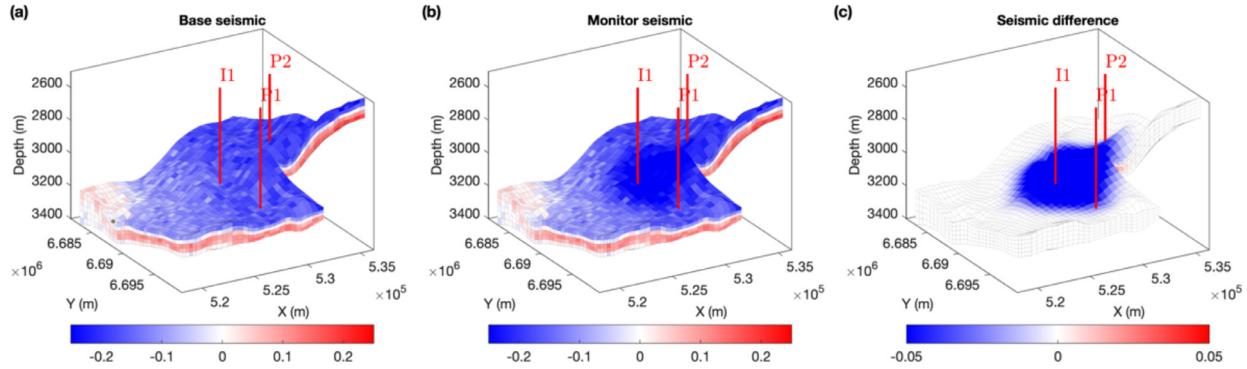


Fig. 6. Time-lapse seismic surveys: (a) base seismic survey acquired before injection; (b) monitor seismic survey acquired after 10 years; (c) the difference between the base and monitor seismic.

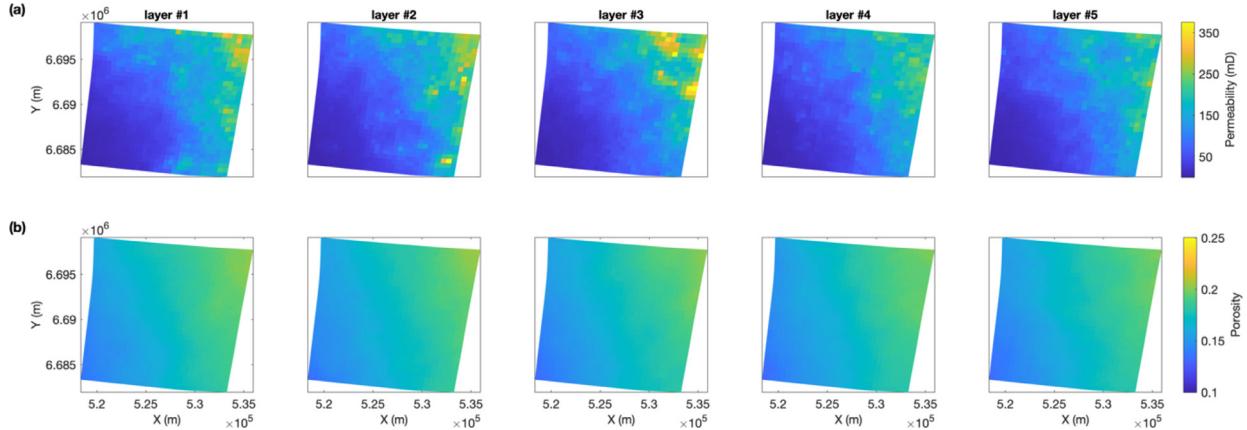


Fig. 7. Mean of the prior models: (a) permeability; (b) porosity.

cells the number of model variables m is 200,000 (100,000 unknown values of porosity and 100,000 unknown values of permeability). For a dataset with borehole measurements every month, in 1 injection and 1 monitoring well, and seismic data every 5 years for 20 years of injection, the number of data measurements n is $n = 480 + 500,000 = 500,480$, assuming that the seismic acquisition grid coincide with the aquifer model grid. The dimension of the dataset makes the application of the ES-MDA algorithm unfeasible. Indeed, in the ES-MDA, the ensemble of models must be large enough to avoid the collapse of the ensemble (i.e., model variances converging to 0). Because of the large dimension of the data, the required number of models in the ensemble is extremely large and would require a long computational time for the calculation of the

data predictions using the fluid flow and geophysical forward models. Instead of increasing the size of the ensemble, we propose to reduce the dimension of the data and apply the ES-MDA in the reduced data space (Fig. 2). In particular, we apply the deep convolutional autoencoder (DCAE) to reduce the data as in Liu and Grana (2019).

Autoencoders are an unsupervised learning technique for the sparse representation learning using neural networks (Goodfellow et al., 2016). Unlike the multilayer perceptron (MLP), the architecture of the autoencoder is generally designed by imposing a bottleneck in the network to compress the original input data (Fig. 3). Such bottleneck divides the autoencoder network into two parts: the encoder component in which the network learns how to compress the input data x into a sparse la-

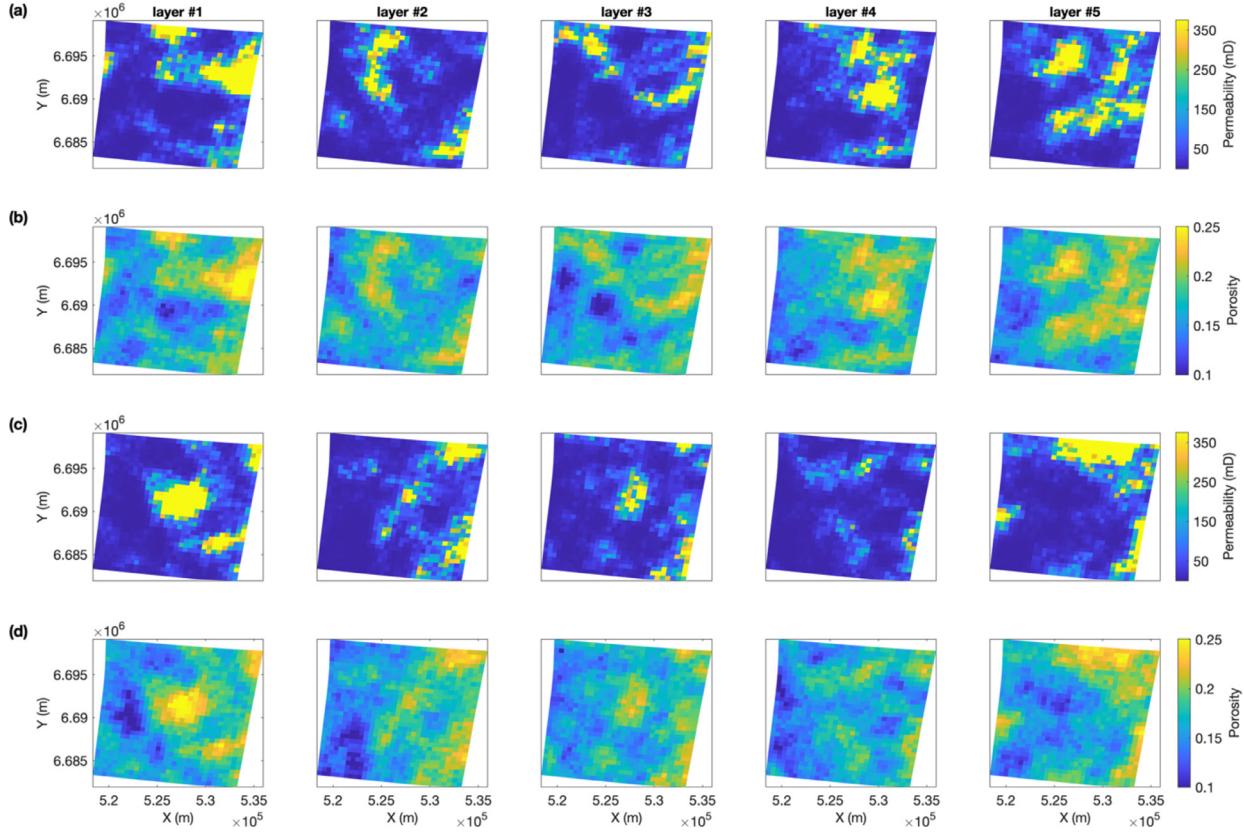


Fig. 8. Two prior realizations: (a) permeability realization #100; (b) porosity realization #100; (c) permeability realization #500; (b) porosity realization #500.

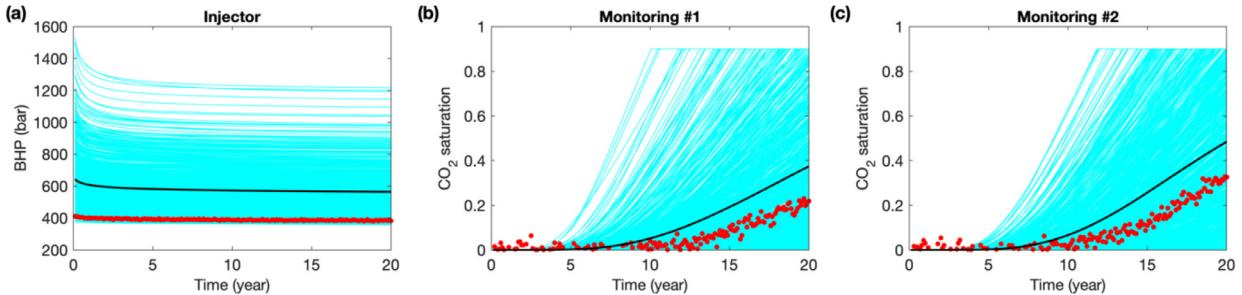


Fig. 9. Predicted injection and monitoring data from prior models: (a) BHP of the injection well I1; (b) CO₂ saturation at the monitoring well P1; (c) CO₂ saturation at the monitoring well P2. Red dots represent the true data; cyan lines represent predicted data from prior models; and black lines represent the mean of the predicted data.

tent feature space \mathbf{z} , and the decoder component in which the network learns how to recover the data $\hat{\mathbf{x}}$ from the sparse features \mathbf{z} with minimum reconstruction error $L = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2$.

An autoencoder consisting of fully connected layers with linear activation function is theoretically equivalent to the principle component analysis (Goodfellow et al., 2016); therefore, it can only capture the linear patterns and correlations in the data space. To avoid the limitation, we combine the autoencoder with convolutional neural network (CNN), namely the deep convolutional autoencoder DCAE. The building blocks of DCAE are convolutional layers with non-linear activation function and max-pooling layers rather than fully connected layers, which provides the network the ability of exploiting spatial and non-linear relationships in the data space. Fig. 3 depicts the network architecture of the DCAE that we design to reduce the seismic data in this study. The detailed parameters of each layer are summarized in Table 1. The network is almost symmetric. The encoder consists of three convolutional layers with decreasing number of filters and three max-pooling layers

to extract the most relevant features in the input space. By contrast, the decoder consists of three convolutional layers with increasing number of filters and up-pooling to recover the input data from the sparse features. In this study, the model is implemented using Keras with TensorFlow as backend engine and the NVIDIA GeForce GTX 1080 GPU is used to speed up the training.

3. Application

In this section, we apply the ES-MDA with DCAE to the petrophysical characterization of deep saline aquifers in terms of porosity and permeability for CO₂ storage and plume prediction. The dataset in this study is built based on the Johansen formation, a potential candidate for large scale CO₂ storage offshore the south-western coast of Norway (Bergmo et al., 2011). The burial depth of the Johansen formation varies between 2200 m and 3100 m below sea level and the thickness ranges from 80 m to 120 m. Such depth and thickness allow storing considerable amount of CO₂ in the form of a dense phase.

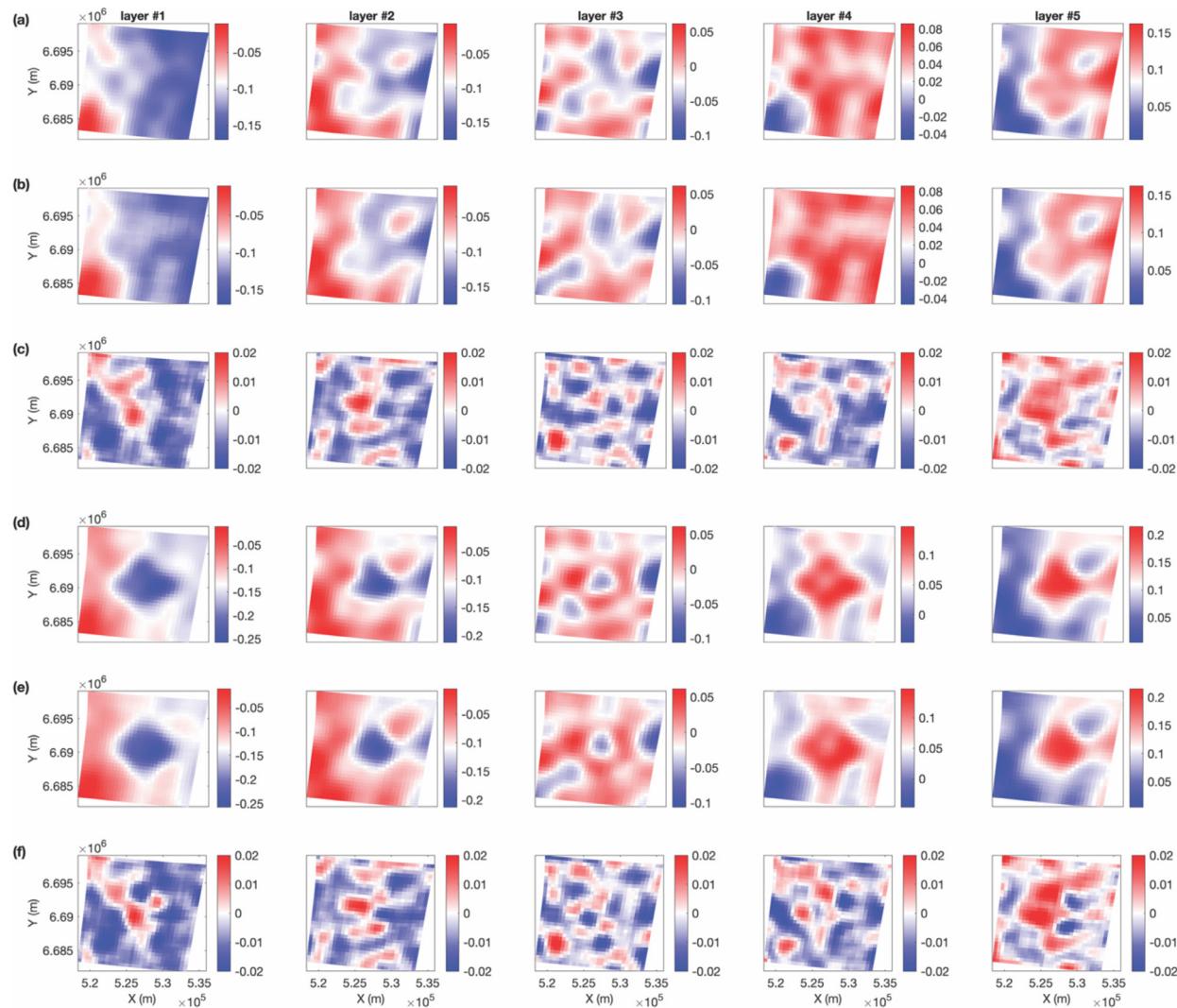


Fig. 10. Reconstruction of the time-lapse seismic data for one realization by the DCAE: (a) original base seismic; (b) reconstructed base seismic; (c) residual between the original and reconstructed base seismic; (d) original monitoring seismic; (e) reconstructed monitoring seismic; (f) residual between the original and reconstructed monitoring seismic.

Table 1
Parameters of each layer in the DCAE.

Layer	Parameters	Layer	Parameters		
1. Input	Input shape	(32, 32, 10)	9. UpSampling	Pooling size	(2, 2)
2. Conv	Filters	256	10. Conv	Filters	128
	Filter size	(3, 3)		Filter size	(3, 3)
	Activation	ReLU		Activation	ReLU
3. MaxPooling	Pooling size	(2, 2)	11. UpSampling	Pooling size	(2, 2)
4. Conv	Filters	128	12. Conv	Filters	256
	Filter size	(3, 3)		Filter size	(3, 3)
	Activation	ReLU		Activation	ReLU
5. MaxPooling	Pooling size	(2, 2)	13. UpSampling	Pooling size	(2, 2)
6. Conv	Filters	64	14. Conv	Filters	10
	Filter size	(3, 3)		Filter size	(3, 3)
	Activation	Sigmoid		Activation	Sigmoid
7. MaxPooling	Pooling size	(2, 2)			
8. Conv	Filters	64			
	Filter size	(3, 3)			
	Activation	ReLU			

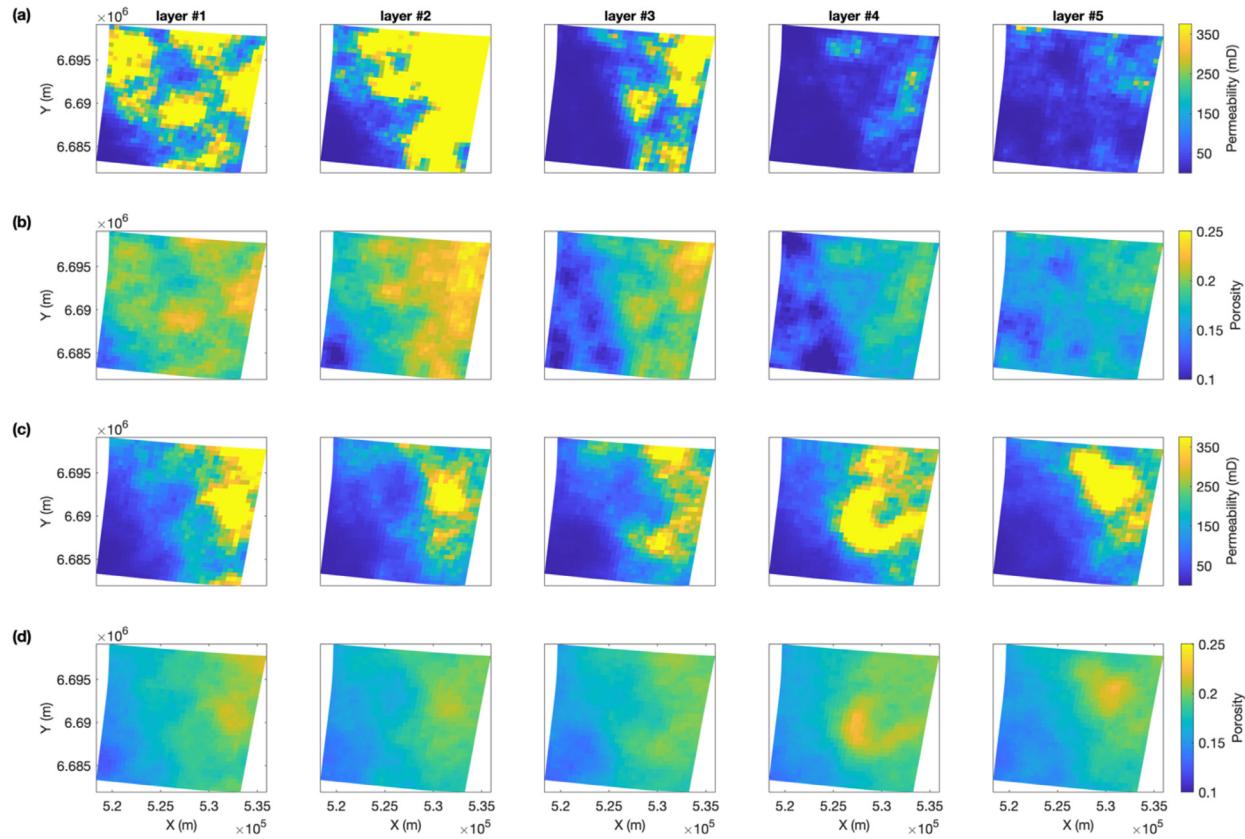


Fig. 11. Posterior mean of the updated models: (a) and (b) represent posterior permeability and porosity with seismic data, respectively; (c) and (d) represent posterior permeability and porosity without seismic data, respectively.

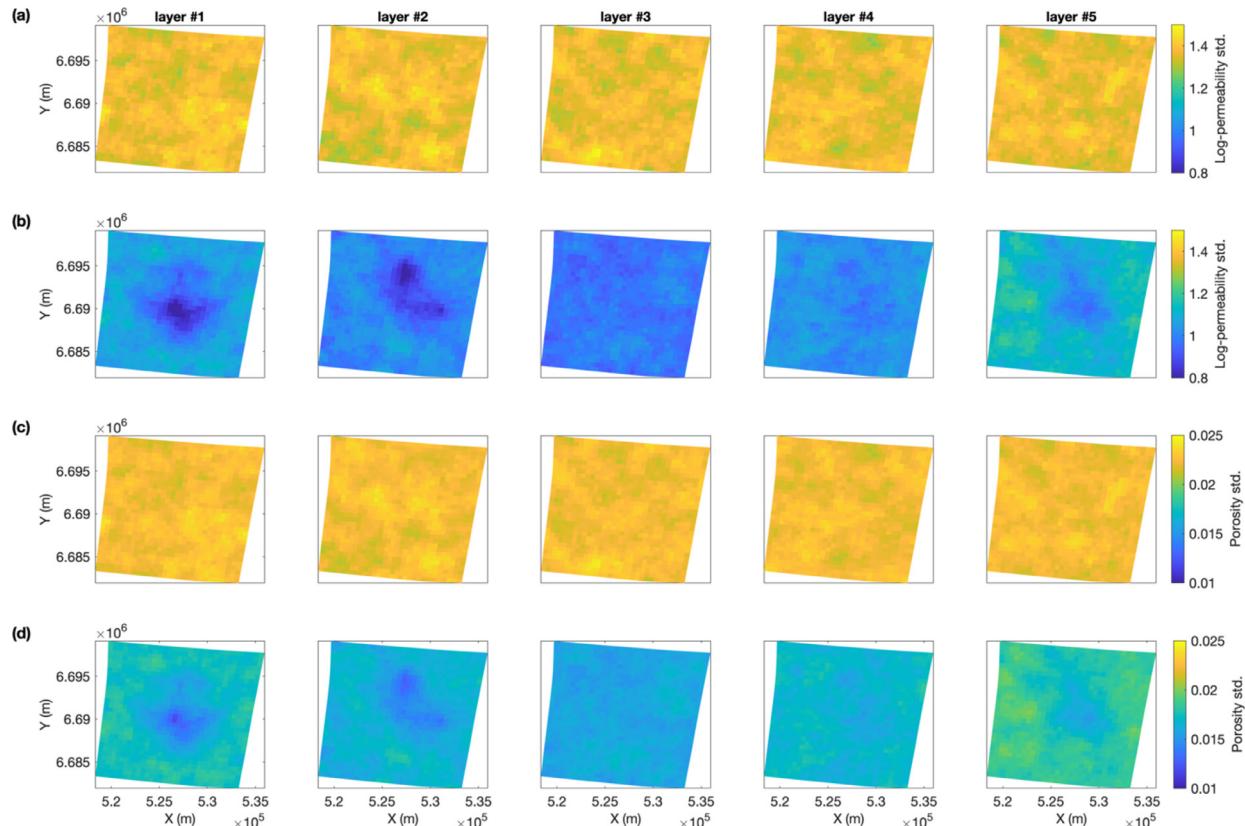


Fig. 12. Prior and posterior standard deviation of the updated models: (a) prior standard deviation of permeability; (b) posterior standard deviation of permeability; (c) prior standard deviation of porosity; (d) posterior standard deviation of porosity.

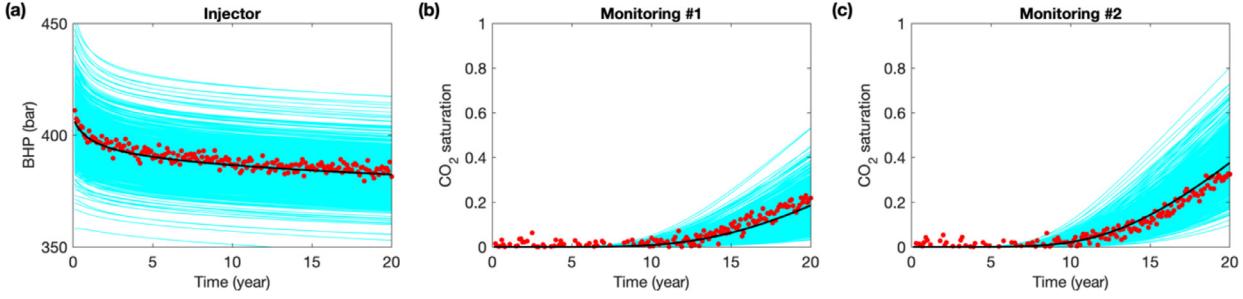


Fig. 13. Predicted injection and monitoring data from posterior models: (a) BHP of the injection well I1; (b) CO₂ saturation at the monitoring well P1; (c) CO₂ saturation at the monitoring well P2. Red dots represent true data; cyan lines represent predicted data from prior models; and black lines represent mean of the predicted data.

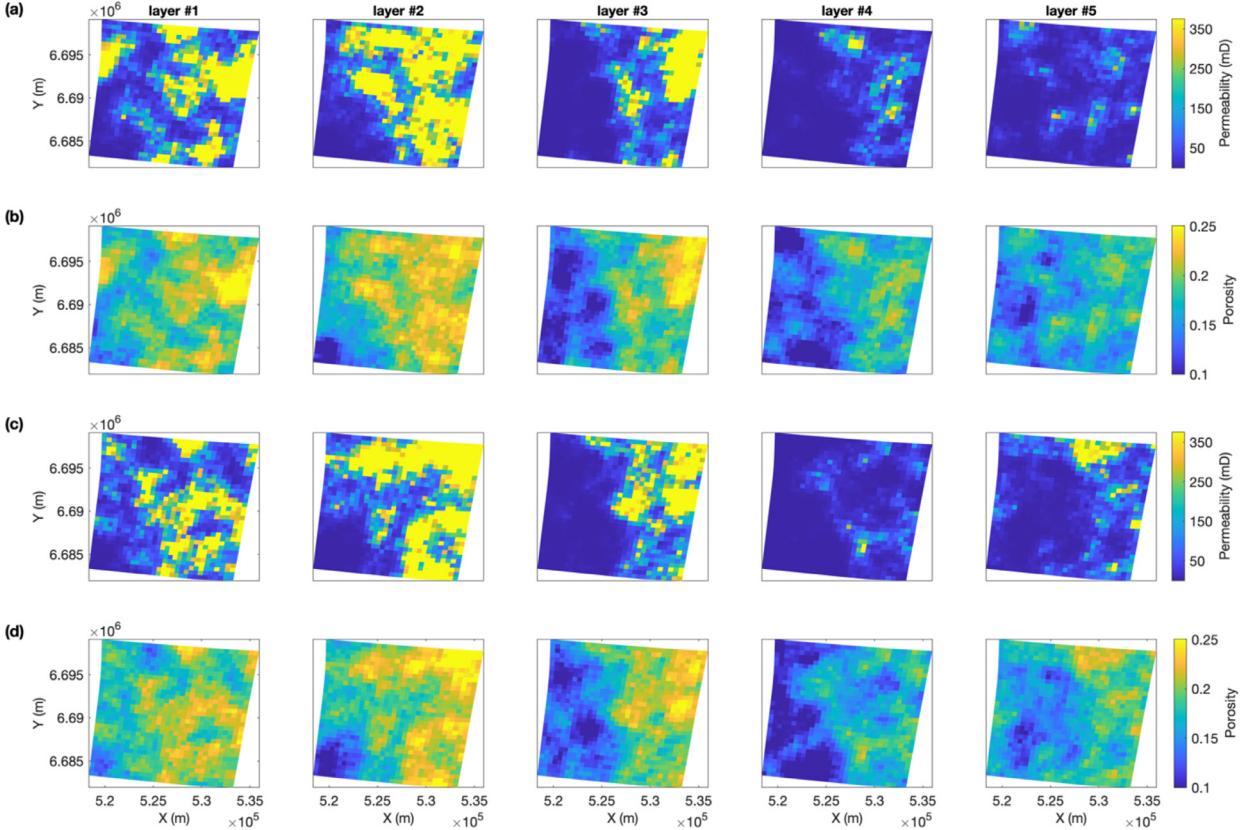


Fig. 14. Two posterior realizations (corresponding to the prior realizations in Fig. 8): (a) permeability realization #100; (b) porosity realization #100; (c) permeability realization #500; (d) porosity realization #500.

3.1. Aquifer model and observations

The simulation grid in the Johansen data set is $100 \times 100 \times 11$ grid cells, each of size $500 \text{ m} \times 500 \text{ m}$ laterally and variable heights from 16 to 24 m (Fig. 4). The model consists of three geological zones and is divided into two sectors by a major fault. The uppermost five layers correspond to the Dunlin shale above the Johansen formation and the lowermost layer corresponds to the Amundsen shale. The permeabilities of the Dunlin and Amundsen shale are assumed to be constant, 0.01 and 0.1 milli-Darcy (mD), respectively. Both shale formations are set to be inactive grid cells in the simulation. The permeability and porosity of the Johansen formation are presented in Figs. 4a and 4b, respectively. The injection well (well I1 in Fig. 4a) is positioned near the main fault with a constant CO₂ injection rate of $60,000 \text{ m}^3/\text{day}$. The injection period is 20 years. When injection stops, the induced pressure gradient gradually dissipates. As a result, the fluid phases are driven by the gravity and

capillary forces to rearrange themselves in the aquifer, as part of the migration process. To monitor the CO₂ sequestration, two monitoring wells (wells P1 and P2 in Fig. 4a) are configured ten grid cells east and north of the injector well respectively, corresponding to a distance of approximately 5 km. Figs. 4c and d show two sections of porosity crossing the injector and monitoring wells. To reduce the computational cost, we reduce the study area of interest to a reduced $32 \times 32 \times 5$ with the injector at the center. The fluid flow simulation shows that the CO₂ plume remains in this region in the period under study. The permeability and porosity of each of the 5 layers of the Johansen formation are presented in Fig. 5.

In this work, the bottom hole pressure (BHP) of the injection well and CO₂ saturation at the monitoring wells are simulated using MRST-co2lab and available every month during the entire injection period (i.e., 20 years). Time-lapse post-stack seismic data are computed using the convolution seismic and rock physics models: the generated geophys-

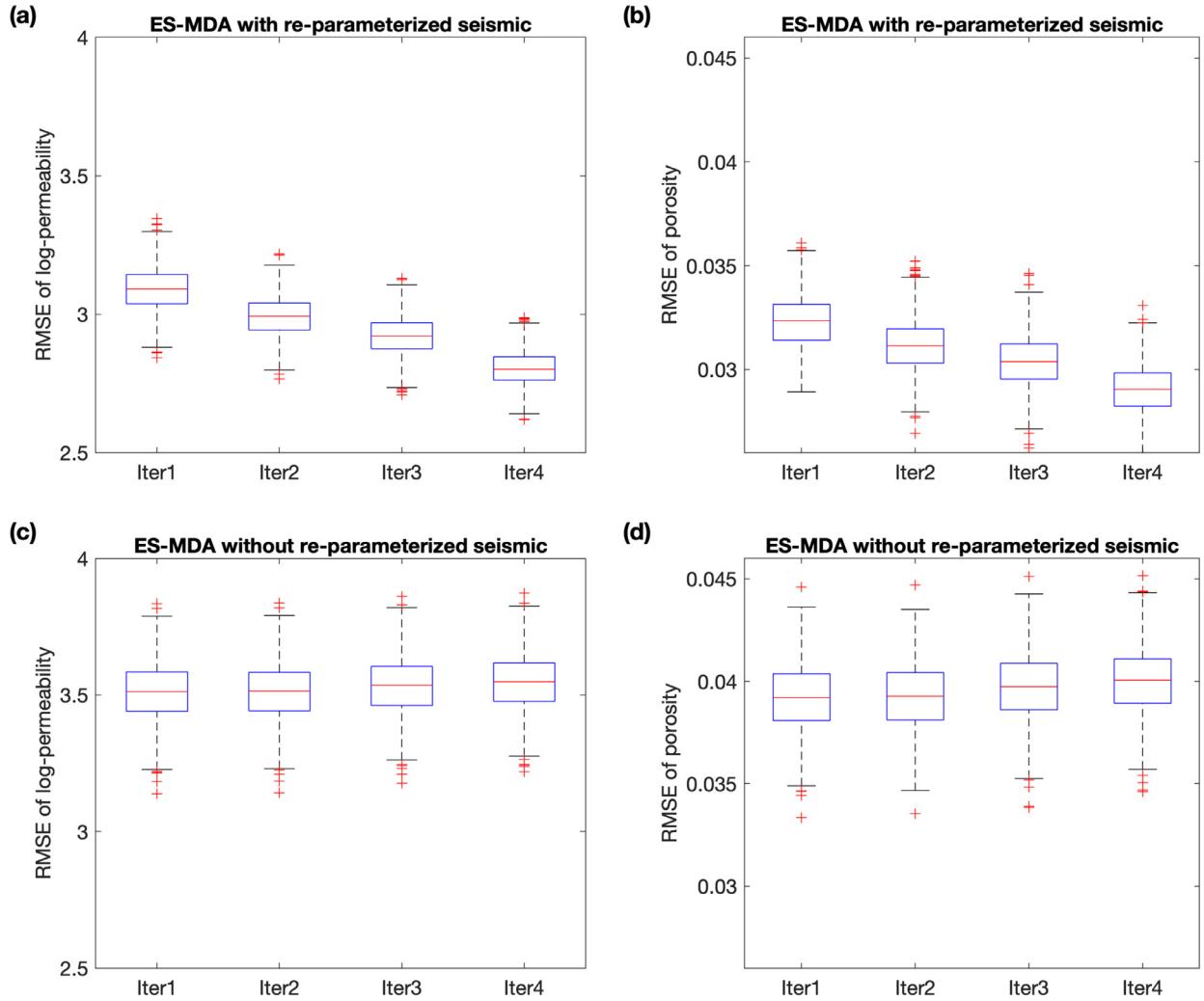


Fig. 15. RMSE of the updated models as function of the ES-MDA iterations: (a) and (b) represent the RMSE of updated permeability and porosity after assimilation of borehole and seismic data, respectively; (c) and (d) represent the RMSE of permeability and porosity after assimilation of borehole data only (i.e. without seismic data), respectively.

ical dataset includes a base survey at the beginning of injection and a monitoring survey after 10 years. The time-lapse seismic data are shown in Fig. 6.

3.2. Prior models and predicted data

In this study, we generate 1000 prior petrophysical models through the FFT-MA and probability field simulation algorithm (Srivastava, 1992; Froidevaux, 1993). Such realizations constitute the initial ensemble for the application of ES-MDA. In the geostatistical simulation, the local mean is set to be the low-frequency trend of the true petrophysical model. Figs. 7a and 7b show the mean permeability and porosity of the prior models, respectively. Two prior realizations (realization #100 and #500) are also shown in Fig. 8. The corresponding injection and monitoring data are simulated using MRST-co2lab with the same well configurations and simulation parameters as the true model (Fig. 9).

3.3. Data reparameterization with DCAE

To avoid the ensemble collapse due to the large dimension of seismic data, we apply a data reduction to reduce the time-lapse seismic data to a much lower dimensional data space by applying the DCAE. With the DCAE depicted in Fig. 3, the $32 \times 32 \times 10$ input data that are extracted

from time-lapse data at the interface of model layers are sparsely represented by a 1024-dimensional vector. In this application, we used all the predicted seismic data from prior realizations and the true observation to train the DCAE. With one NVIDIA GeForce GTX 1080 GPU, the training takes approximately 5 minutes to converge. Fig. 10 shows the original and reconstructed seismic of one realization randomly selected from the training set. We can see that the reconstruction error is relatively small and the main space structures are well captured, which indicates that the extracted sparse features contain most of the information of the input data. Therefore, it is applicable and efficient to use the trained DCAE to reduce the order of seismic data.

3.4. Data assimilation results

In this study, the iteration number of ES-MDA is set to be 4 with the inflation coefficients of 9.333, 7.0, 4.0 and 2.0, respectively as suggested for history matching studies in Emerick and Reynolds (2013). We assume that the covariance matrix Σ_d is a diagonal matrix where the elements on the diagonal represent the noise level in the data. Since the deep neural network cannot be expressed explicitly, we use the noise level in the original seismic data to estimate the noise standard deviation in the latent space. The standard deviations of noise of CO₂ saturation, BHP and the original seismic amplitudes are 0.01, 1, and 0.02 respectively. After encoding using the DCAE, the seismic data are scaled

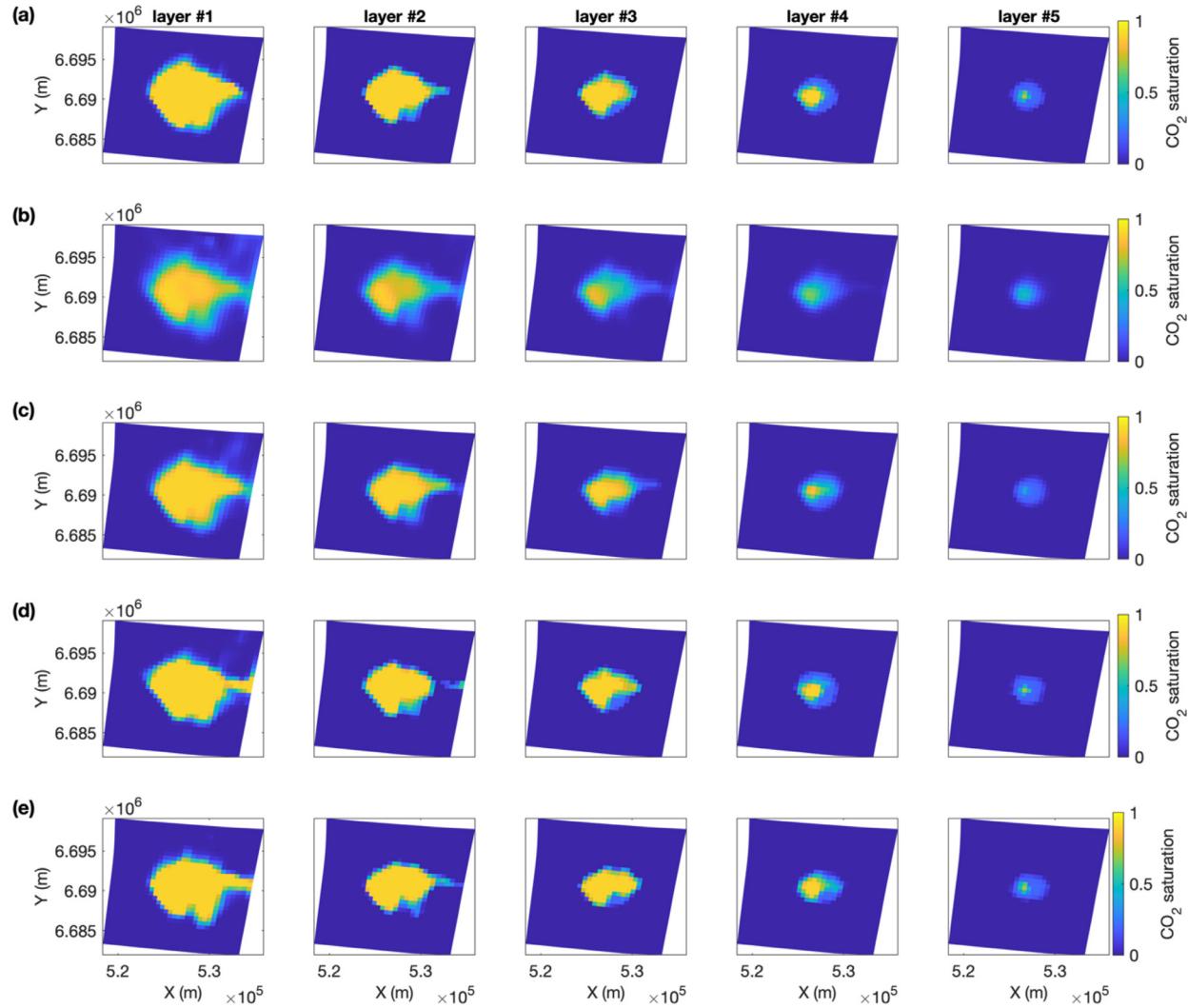


Fig. 16. Predicted CO₂ plume from true, prior and posterior models after 100 years: (a) true CO₂ plume; (b) prior mean of the predicted CO₂ plume; (c) posterior mean of the predicted CO₂ plume; (d) predicted CO₂ plume from the updated realization #100; (e) predicted CO₂ plume from the updated realization #500.

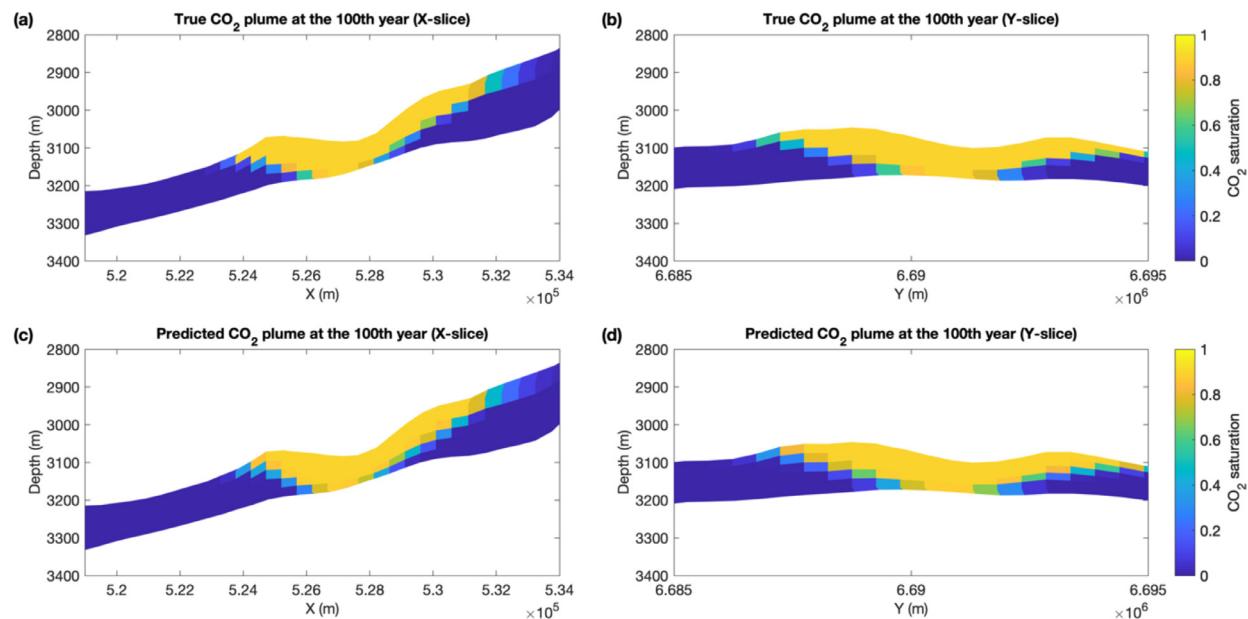


Fig. 17. Vertical sections of the true and predicted CO₂ plume after 100 years, crossing the injection and monitoring wells: (a) true CO₂ plume crossing I1 and P2; (b) true CO₂ plume crossing I1 and P1; (c) predicted CO₂ plume crossing I1 and P2; (d) predicted CO₂ plume crossing I1 and P1.

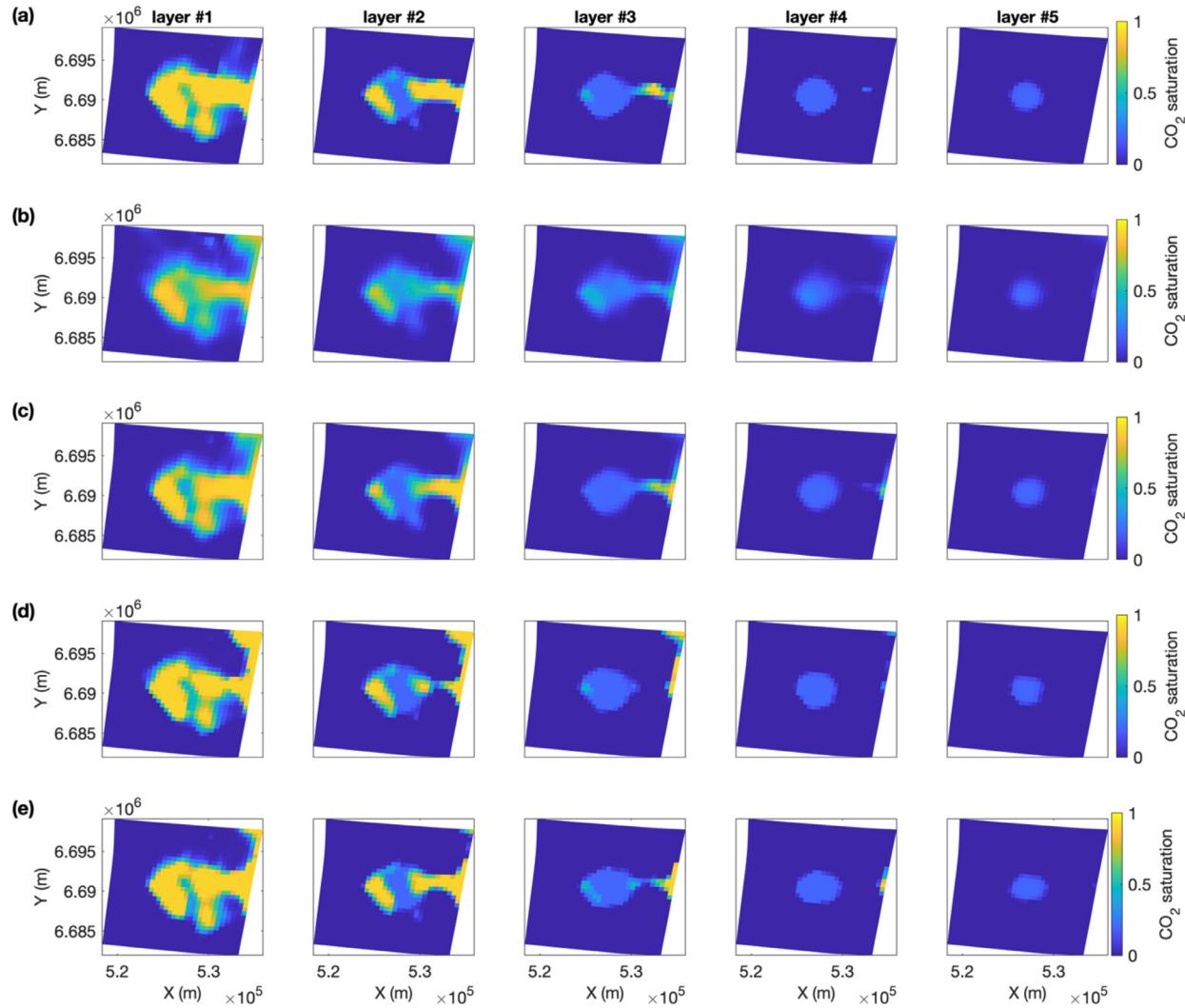


Fig. 18. Predicted CO_2 plume from true, prior and posterior models after 400 years: (a) true CO_2 plume; (b) prior mean of the predicted CO_2 plume; (c) posterior mean of the predicted CO_2 plume; (d) predicted CO_2 plume from the updated realization #100; (e) predicted CO_2 plume from the updated realization #500.

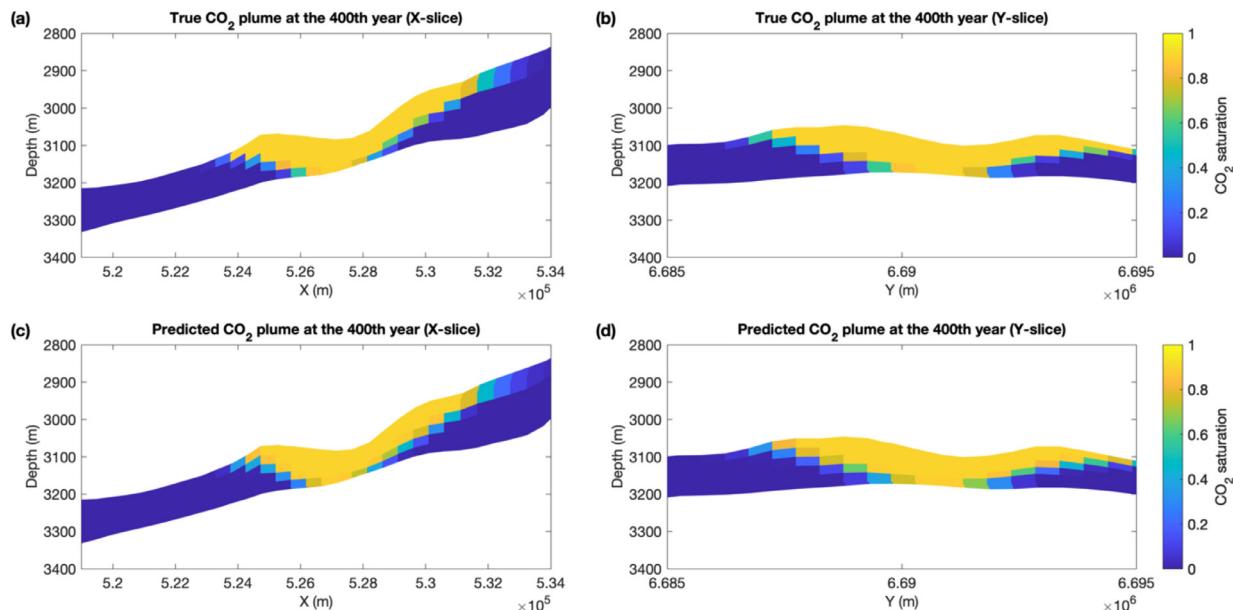


Fig. 19. Vertical sections of the true and predicted CO_2 plume after 400 years, crossing the injection and monitoring wells: (a) true CO_2 plume crossing I1 and P2; (b) true CO_2 plume crossing I1 and P1; (c) predicted CO_2 plume crossing I1 and P2; (d) predicted CO_2 plume crossing I1 and P1.

from [-0.3, 0.2] to [0, 1] and the standard deviation of the noise of the re-parameterized seismic data is 0.08. Figs. 11a and 11b show the posterior mean of the petrophysical models and illustrate a good match between the predicted models and the true models. Fig. 12 shows that the uncertainty in the posterior models is largely reduced compared to the prior models after the assimilation of both the borehole data and re-parameterized seismic. Furthermore, Fig. 13 shows that the predicted injection and monitoring data from the obtained posterior models match the true observations. Not only the mean, but also the individual realizations converge to the true solution. Fig. 14 shows the posterior realizations obtained from the prior models in Fig. 8 and illustrates that both realizations converge to the true model after updating.

To determine the value of seismic data in the prediction of the petrophysical properties, we also perform the inversion without time-lapse seismic data, i.e. with borehole data only. The results are shown in Figs. 11c and 11d. In general, the results with seismic data show a much better agreement with the true models compares with the results with borehole data only, especially in the regions away well locations. From the root mean squared errors (RMSE) in Fig. 15, we can also conclude that seismic data provide important information to reduce the model mismatch.

3.5. Prediction of CO₂ plume

In the CO₂ storage studies, it is of great significance to accurately predict the CO₂ plume location in the aquifer in the long term. Figs. 16 and 17 compare the true and predicted CO₂ plume after 100 years of migration from the prior and posterior models. The predictions are conditioned to borehole data and re-parameterized seismic data and are illustrated layer by layer (Fig. 16) and along the vertical sections crossing the injection and monitoring wells (Fig. 17). Figs. 18 and 19 show the same results after 400 years of migration. In both cases, the posterior predicted CO₂ plume has a much better agreement with the CO₂ plume from the true model in each layer, compared to the prior models.

4. Conclusion

Traditionally, geophysical methods have been used for static modeling of pre-injection aquifer conditions; however, due to several sources of uncertainty in the data and in the physical models, pre-injection models of rock and fluid properties are often inaccurate. In this work, we proposed the integration of geophysical data in history matching, in particular the simultaneous assimilation of borehole data and time-lapse seismic survey. To overcome the limitations due to the large dimension of geophysical data, we combined the stochastic inversion method with the deep convolutional autoencoder to reduce the dimension of the observed data and perform the data assimilation in a lower dimensional space. This approach allows improving the aquifer characterization in terms of petrophysical properties such as porosity and permeability and obtaining more precise predictions of the dynamic behavior of the CO₂ plume during and after injection. The method was illustrated and validated on a synthetic deep saline aquifer model and shows the ability of the proposed methodology to improve the petrophysical characterization as well as the model forecasting during migration.

CRediT authorship contribution statement

Mingliang Liu: Conceptualization, Methodology, Software, Writing - original draft. **Dario Grana:** Conceptualization, Supervision, Writing - review & editing.

Acknowledgements

Authors acknowledge the School of Energy Resources, and the Department of Geology and Geophysics of University of Wyoming for the support. The authors also would like to thank SINTEF for providing the Matlab Reservoir Simulation Toolbox.

APPENDIX A. Soft sand model

The soft sand model (Mavko et al., 2009) allows computing the dry-rock elastic moduli of a rock with known porosity, based Hertz-Mindlin contact theory and Hashin-Shtrikman elastic bounds. First, the bulk and shear moduli of the dry-rock at the critical porosity value, K_{HM} and G_{HM} , are computed using Hertz-Mindlin equations:

$$K_{HM} = \sqrt[3]{\frac{P_e[n(1-\phi_0)G_{mat}]^2}{18[\pi(1-\nu)]^2}} \quad (A1)$$

$$G_{HM} = \frac{5-4\nu}{5(2-\nu)} \sqrt[3]{\frac{3P_e[n(1-\phi_0)G_{mat}]^2}{2[\pi(1-\nu)]^2}} \quad (A2)$$

where ϕ_0 is the critical porosity, ν is the grain Poisson's ratio, P_e is the effective pressure, and n is average number of contacts per grain (Mindlin, 1949). Then, for porosity values within the range [0, ϕ_0], the bulk and shear moduli of the dry-rock K_{dry} and G_{dry} are estimated by interpolating the matrix elastic moduli at zero porosity and the Hertz-Mindlin elastic moduli at the critical porosity, using the modified Hashin-Shtrikman lower bounds (Hashin and Shtrikman, 1963):

$$K_{dry} = \left(\frac{\frac{\phi}{\phi_0}}{K_{HM} + \frac{4}{3}\mu_{HM}} + \frac{1 - \frac{\phi}{\phi_0}}{K_{mat} + \frac{4}{3}\mu_{HM}} \right)^{-1} - \frac{4}{3}\mu_{HM} \quad (A3)$$

$$\mu_{dry} = \left(\frac{\frac{\phi}{\phi_0}}{\mu_{HM} + \frac{1}{6}\xi\mu_{HM}} + \frac{1 - \frac{\phi}{\phi_0}}{\mu_{mat} + \frac{1}{6}\xi\mu_{HM}} \right)^{-1} - \frac{1}{6}\xi\mu_{HM} \quad (A4)$$

where

$$\xi = \frac{9K_{HM} + 8\mu_{HM}}{K_{HM} + 2\mu_{HM}} \quad (A5)$$

References

- Alnes, H., Eiken, O., Nooner, S., Sasagawa, G., Stenvold, T., Zumberge, M., 2011. Results from Sleipner gravity monitoring: updated density and temperature distribution of the CO₂ plume. *Energy Procedia* 4, 5504–5511.
- Arts, R.J., Chadwick, A., Eiken, O., Thibeau, S., Nooner, S., 2008. Ten years' experience of monitoring CO₂ injection in the Utsira Sand at Sleipner, offshore Norway. *First break* 26 (1).
- Bachu, S., 2008. CO₂ storage in geological media: role, means, status and barriers to deployment. *Progr. Energy Combust. Sci.* 34, 254–273.
- Bachu, S., 2015. Review of CO₂ storage efficiency in deep saline aquifers. *Int. J. Greenhouse Gas Control* 40, 188–202.
- Bellenfant, G., Guyonnet, D., Dubois, D., Bouc, O., 2009. Uncertainty theories applied to the analysis of CO₂ plume extension during geological storage. *Energy Procedia* 1 (1), 2447–2454.
- Bergmo, P.S., Grimstad, A.A., Lindeberg, E., Riis, F., Johansen, W.T., 2009a. Exploring geological storage sites for CO₂ from Norwegian gas power plants: Utsira South. *Energy Procedia* 1 (1), 2953–2959.
- Bergmo, P.S., Lindeberg, E., Riis, F., Johansen, W.T., 2009b. Exploring geological storage sites for CO₂ from Norwegian gas power plants: Johansen formation. *Energy Procedia* 1 (1), 2945–2952.
- Bergmo, P.S., Grimstad, A.A., Lindeberg, E., 2011. Simultaneous CO₂ injection and water production to optimise aquifer storage capacity. *Int. J. Greenhouse Gas Control* 5 (3), 555–564.
- Bhowmik, S., Srinivasan, S., Bryant, S., 2013. Prediction of plume migration using injection data and a model selection approach. *Energy Procedia* 37, 3672–3679.
- Brennan, S.T., Burruss, R.C., Merrill, M.D., Freeman, P.A., Ruppert, L.F., 2010. A Probabilistic Assessment Methodology for the Evaluation of Geologic Carbon Dioxide Storage. U.S. Geological Survey, pp. 1–31.
- Burruss, R.C., Brennan, S.T., Freeman, P.A., Merrill, M.D., Ruppert, L.F., Becker, M.F., Herkelrath, W.N., Kharaka, Y.K., Neuizil, C.E., Swanson, S.M., Cook, T.A., Klett, T.R., Nelson, P.H., Schenk, C.J., 2009. Development of a Probabilistic Assessment Methodology for Evaluation of Carbon Dioxide Storage. U.S. Geological Survey, pp. 1–81.
- Caeiro, M.H., Demyanov, V., Soares, A., 2015. Optimized history matching with direct sequential image transforming for non-stationary reservoirs. *Mathematical Geosci.* 47 (8), 975–994.
- Canchumuni, S.W., Emerick, A.A., Pacheco, M.A.C., 2019a. Towards a robust parameterization for conditioning facies models using deep variational autoencoders and ensemble smoother. *Comp. & Geosci.* 128, 87–102.
- Canchumuni, S.W., Emerick, A.A., Pacheco, M.A.C., 2019b. History matching geological facies models based on ensemble smoother and deep generative models. *J. Petroleum Sci. and Eng.* 177, 941–958.

- Capilla, J.E., Gomez-Hernández, J.J., Sahuquillo, A., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. *J. Hydrol.* 203, 175–188.
- Castelletto, N., Teatini, P., Gambolati, G., Bossie-Codreanu, D., Vincké, O., Daniel, J.M., Battistelli, A., Marcolini, M., Donda, F., Volpi, V., 2013. Multiphysics modeling of CO₂ sequestration in a faulted saline formation in Italy. *Advances in Water Resources* 62, 570–587.
- Chadwick, R.A., Arts, R., Eiken, O., 2005. 4D seismic quantification of a growing CO₂ plume at Sleipner, North Sea. In: Geological Society, London, Petroleum Geology Conference series, 6. Geological Society of London, pp. 1385–1399.
- Chadwick, A., Arts, R., Eiken, O., Williamson, P., Williams, G., 2006. Geophysical monitoring of the CO₂ plume at Sleipner, North Sea. In: Advances in the Geological Storage of Carbon Dioxide. Springer, Dordrecht, pp. 303–314.
- Chadwick, R.A., Noy, D., Arts, R., Eiken, O., 2009. Latest time-lapse seismic data from Sleipner yield new insights into CO₂ plume development. *Energy Procedia* 1 (1), 2103–2110.
- Chadwick, R.A., Noy, D.J., 2010. History-matching flow simulations and time-lapse seismic data from the Sleipner CO₂ plume. In: Geological Society, London, Petroleum Geology Conference series, 7. Geological Society of London, pp. 1171–1182.
- Chen, B., Harp, D.R., Lin, Y., Keating, E.H., Pawar, R.J., 2018. Geologic CO₂ sequestration monitoring design: a machine learning and uncertainty quantification based approach. *Applied energy* 225, 332–345.
- Chen, Y., Oliver, D.S., 2012. Ensemble randomized maximum likelihood method as an iterative ensemble smoother. *Mathematical Geosci.* 44 (1), 1–26.
- Chen, Y., Oliver, D.S., 2017. Localization and regularization for iterative ensemble smoothers. *Computational Geosci.* 21 (1), 13–30.
- Dai, Z., Viswanathan, H., Fessenden-Rahn, J., Middleton, R., Pan, F., Jia, W., Lee, S.Y., McPherson, B., Ampomah, W. and Grigg, R., 2014. Uncertainty quantification for CO₂ sequestration and enhanced oil recovery. *arXiv preprint arXiv:1411.4900*.
- Deng, H., Stauffer, P.H., Dai, Z., Jiao, Z., Surdam, R.C., 2012. Simulation of industrial-scale CO₂ storage: Multi-scale heterogeneity and its impacts on storage capacity, injectivity and leakage. *Int. J. Greenhouse Gas Control* 10, 397–418.
- Eigestad, G.T., Dahle, H.K., Hellevang, B., Riis, F., Johansen, W.T., Øian, E., 2009. Geological modeling and simulation of CO₂ injection in the Johansen formation. *Computational Geosci.* 13 (4), 435.
- Ellett, K.Z., Medina, C., Rupp, J., Wang, G.C., Carr, T., 2013. Uncertainty in regional-scale evaluation of CO₂ geologic storage resources—comparison of the Illinois Basin (USA) and the Ordos Basin (China). *Energy Procedia* 37, 5151–5159.
- Emerick, A.A., Reynolds, A.C., 2013. Ensemble smoother with multiple data assimilation. *Computers & Geosci.* 55, 3–15.
- Etienam, C., 2019. 4D Seismic History Matching Incorporating Unsupervised Learning. *arXiv preprint arXiv:1905.07469*.
- Evensen, G., 2009. Data Assimilation: the Ensemble Kalman Filter. Springer Science & Business Media.
- Froidevaux, R., 1993. Probability field simulation. *Geostat. Troia*. 92, 73–83.
- Gasda, S.E., Nordbotten, J.M., Celia, M.A., 2012. Application of simplified models to CO₂ migration and immobilization in large-scale geological systems. *Int. J. Greenhouse Gas Control* 9, 72–84.
- Gassmann, F., 1951. Elastic waves through a packing of spheres. *Geophysics* 16 (4), 673–685.
- Glubokovskikh, S., Pevzner, R., Gunning, J., Dance, T., Shulakova, V., Popik, D., Popik, S., Bagheri, M., Gurevich, B., 2019. How well can time-lapse seismic characterize a small CO₂ leakage into a saline aquifer: CO2CRC Otway 2C experiment (Victoria, Australia). *Int. J. Greenhouse Gas Control*, 102854.
- Ghorbanieh, H., Kokkinaki, A., Li, J.Y., Darve, E., Kitanidis, P.K., 2015. Real-time data assimilation for large-scale systems: the spectral Kalman filter. *Advances in water resources* 86, 260–272.
- Gómez-Hernández, J.J., Sahuquillo, A., Capilla, J., 1997. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—I. Theory. *J. Hydrol.* 203 (1–4), 162–174.
- Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep learning. MIT Press.
- Goodman, A., Hakala, A., Bromhal, G., Deel, D., Rodosta, T., Frailey, S., Small, M., Allen, D., Romanov, V., Fazio, J., Huerta, N., McIntyre, D., Kutchko, B., Guthrie, G., 2011. U.S. DOE methodology for the development of geologic storage potential for carbon dioxide at the national and regional scale. *International Journal of Greenhouse Gas Control. Int. J. Greenhouse Gas Control* 5, 952–965.
- Goodman, A., Bromhal, G., Strazisar, B., Rodosta, T., Guthrie, W.F., Allen, D., Guthrie, G., 2013. Comparison of methods for geologic storage of carbon dioxide in saline formations. *Int. J. Greenhouse Gas Control* 118, 329–342.
- Goodman, A., Sanguinito, S., and Levine, J.S., 2017. Prospective CO₂ saline resource estimation methodology: refinement of existing US-DOE-NETL methods based on data availability. *Int. J. Greenhouse Gas Control*
- Gorecki, C.D., Ayash, S.C., Liu, G., 2015. A comparison of volumetric and dynamic CO₂ storage resource and efficiency in deep saline formations. *Int. J. Greenhouse Gas Control* 42, 213–225.
- Grana, D., Verma, S., Pafeng, J., Lang, X., Sharma, H., Wu, W., McLaughlin, F., Campbell, E., Ng, K., Alvarado, V., Mallik, S., Kaszuba, J., 2017. A rock physics and seismic reservoir characterization study of the Rock Springs Uplift, a CO₂ sequestration site in Southwestern Wyoming. *Int. J. Greenhouse Gas Control* 63, 296–309.
- Grude, S., Landrø, M., Osdal, B., 2013. Time-lapse pressure-saturation discrimination for CO₂ storage at the Snøhvit field. *Int. J. Greenhouse Gas Control* 19, 369–378.
- Hashin, Z., Shtrikman, S., 1963. A variational approach to the theory of the elastic behaviour of multiphase materials. *Journal of the Mechanics and Physics of Solids* 11 (2), 127–140.
- Ivandic, M., Juhlin, C., Lueth, S., Bergmann, P., Kashubin, A., Sopher, D., Ivanova, A., Baumann, G., Henninges, J., 2015. Geophysical monitoring at the Ketzin pilot site for CO₂ storage: New insights into the plume evolution. *Int. J. Greenhouse Gas Control* 32, 90–105.
- Jeong, H., Srinivasan, S., Bryant, S., 2013. Uncertainty quantification of CO₂ plume migration using static connectivity of geologic features. *Energy Procedia* 37, 3771–3779.
- Jung, H., Singh, G., Espinoza, D.N., Wheeler, M.F., 2018. Quantification of a maximum injection volume of CO₂ to avert geomechanical perturbations using a compositional fluid flow reservoir simulator. *Advances in water resources* 112, 160–169.
- Kitanidis, P.K., 1997. Introduction to geostatistics. Applications in hydrogeology..
- Kopp, A., Binning, P.J., Johannsen, K., Helmig, R., Class, H., 2010. A contribution to risk analysis for leakage through abandoned wells in geological CO₂ storage. *Advances in Water Resources* 33 (8), 867–879.
- Laloy, E., Hérault, R., Jacques, D., Linde, N., 2018. Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research* 54 (1), 381–406.
- Levine, J.S., Fukai, I., Soeder, D.J., Bromhal, G., Dilmore, R.M., Guthrie, G.D., Rodosta, T., Sanguinito, S., Frailey, S., Gorecki, D., Peck, W., Goodman, A.L., 2016. U.S. DOE NETL Methodology for estimating the prospective CO₂ storage resource of shales at the national and regional scale. *Int. J. Greenhouse Gas Control* 96, 103006.
- Li, B., Benson, S.M., 2015. Influence of small-scale heterogeneity on upward CO₂ plume migration in storage aquifers. *Advances in water resources* 83, 389–404.
- Li, S., Zhang, Y., Zhang, X., 2011. A study of conceptual model uncertainty in large-scale CO₂ storage simulation. *Water Resources Research* 47 (5).
- Lie, K.A., 2016. An Introduction to Reservoir Simulation Using Matlab: User guide for the Matlab Reservoir Simulation Toolbox (MRST). SINTEFICT.
- Liu, M., Grana, D., 2019. Time-lapse seismic history matching with iterative ensemble smoother and deep convolutional autoencoder. *Geophysics* 85 (1), 1–63.
- Liu, M., Grana, D., 2018. Stochastic nonlinear inversion of seismic data for the estimation of petroelastic properties using the ensemble smoother and data reparameterization. *Geophysics* 83 (3), M25–M39.
- Liu, Y., Sun, W., Durlofsky, L.J., 2019. A deep-learning-based geological parameterization for history matching complex models. *Mathematical Geosci.* 51 (6), 725–766.
- Liu, F., Ellett, K., Xiao, Y., Rupp, J., 2013. Assessing the feasibility of CO₂ storage in the New Albany Shale (Devonian-Mississippian) with potential for enhanced gas recovery using reservoir simulation. *Int. J. Greenhouse Gas Control* 17, 111–126.
- Lorentzen, R.J., Luo, X., Bhakta, T., Valestrand, R., 2019. History matching the full norne field model using seismic and production data. *SPE Journal*.
- Luo, X., 2019. Ensemble-based kernel learning for a class of data assimilation problems with imperfect forward simulators. *arXiv preprint arXiv:1901.10758*.
- Luo, X., Bhakta, T., Nævdal, G., 2018. Correlation-based adaptive localization with applications to ensemble-based 4D-seismic history matching. *SPE Journal* 23 (02), 396–427.
- Luo, X., Bhakta, T., Jakobsen, M. and Nævdal, G., 2016. An ensemble 4D seismic history matching framework with sparse representation based on wavelet multiresolution analysis. *arXiv preprint arXiv:1603.04577*.
- Ma, W., Jafarpour, B., Qin, J., 2019. Dynamic characterization of geologic CO₂ storage aquifers from monitoring data with ensemble Kalman filter. *Int. J. Greenhouse Gas Control* 81, 199–215.
- Mariethoz, G., Caers, J., 2014. Multiple-Point Geostatistics: Stochastic Modeling with Training Images. John Wiley & Sons.
- Mathias, S.A., Gluyas, J.G., de Miguel, G.J.G.M., Bryant, S.L., Wilson, D., 2013. On relative permeability data uncertainty and CO₂ injectivity estimation for brine aquifers. *Int. J. Greenhouse Gas Control* 12, 200–212.
- Mavko, G., Mukerji, T., Dvorkin, J., 2009. The Rock Physics Handbook. Cambridge University Press.
- Miller, D.C., Syamlal, M., Mebane, D.S., Storlie, C., Bhattacharyya, D., Sahinidis, N.V., Agarwal, D., Tong, C., Zitney, S.E., Sarkar, A., Sun, X., 2014. Carbon capture simulation initiative: a case study in multiscale modeling and new challenges. *Ann. Rev. Chem. Biomol. Eng.* 5, 301–323.
- Mindlin, R.D., 1949. Compliance of elastic bodies in contact. *J. Appl. Mech., ASME* 16, 259–268.
- Mohamed, L., Christie, M.A., Demyanov, V., 2010. Comparison of stochastic sampling algorithms for uncertainty quantification. *SPE Journal* 15 (01), 31–38.
- Nilsen, H.M., Herrera, P.A., Ashraf, M., Ligaarden, I., Iding, M., Hermanrud, C., Lie, K.A., Nordbotten, J.M., Dahle, H.K., Keilegavlen, E., 2011. Field-case simulation of CO₂-plume migration using vertical-equilibrium models. *Energy Procedia* 4, 3801–3808.
- Nordbotten, J.M. and Celia, M.A., 2010. Analysis of plume extent using analytical solutions for CO₂ storage.
- Nordbotten, J.M., Flemisch, B., Gasda, S.E., Nilsen, H.M., Fan, Y., Pickup, G.E., Wiese, B., Celia, M.A., Dahle, H.K., Eigestad, G.T., Pruess, K., 2012. Uncertainties in practical simulation of CO₂ storage. *Int. J. Greenhouse Gas Control* 9, 234–242.
- Oliver, D.S., Reynolds, A.C., Liu, N., 2008. Inverse Theory for Petroleum Reservoir Characterization and History Matching. Cambridge University Press.
- Oruganti, Y., Gupta, A.K., Bryant, S.L., 2011. Analytical estimation of risk due to pressure buildup during CO₂ injection in deep saline aquifers. *Energy Procedia* 4, 4140–4147.
- Peck, W., Glazewski, K., Klenner, R., Gorecki, C., Steadman, E., Harju, J., 2014. A workflow to determine CO₂ storage potential in deep saline formations. *Energy Procedia* 63, 5231–5238.
- Pool, M., Carrera, J., Vilarrasa, V., Silva, O., Ayora, C., 2013. Dynamics and design of systems for geological storage of dissolved CO₂. *Advances in water resources* 62, 533–542.
- Puzyrev, V., 2019. Deep learning electromagnetic inversion with convolutional neural networks. *Geophysical Journal International* 218 (2), 817–832.
- Roach, L.A., White, D.J., 2018. Evolution of a deep CO₂ plume from time-lapse seismic imaging at the Aquistore storage site, Saskatchewan, Canada. *Int. J. Greenhouse Gas Control* 74, 79–86.

- Sifuentes, W.F., Giddins, M.A., Blunt, M.J., 2009. Modeling CO₂ storage in aquifers: Assessing the key contributors to uncertainty. Offshore Europe. Society of Petroleum Engineers.
- Senel, O., Chugunov, N., 2013. CO₂ injection in a saline formation: pre-injection reservoir modeling and uncertainty analysis for Illinois basin-decatur project. Energy Procedia 37, 4598–4611.
- Srivastava, R.M., 1992. Reservoir characterization with probability field simulation. In: SPE Annual Technical Conference and Exhibition, pp. 4–7.
- Sun, W., Durlofsky, L.J., 2019. Data-space approaches for uncertainty quantification of CO₂ plume location in geological carbon storage. Advances in water resources 123, 234–255.
- Sun, Y., Tong, C., Trainor-Guitton, W.J., Lu, C., Mansoor, K., Carroll, S.A., 2013. Global sampling for integrating physics-specific subsystems and quantifying uncertainties of CO₂ geological sequestration. Int. J. Greenhouse Gas Control 12, 108–123.
- Sundal, A., Nystrøm, J.P., Dypvik, H., Miri, R., Aagaard, P., 2013. Effects of geological heterogeneity on CO₂ distribution and migration-A case study from the Johansen Formation, Norway. Energy Procedia 37, 5046–5054.
- Szulczewski, M., Juanes, R., 2009. A simple but rigorous model for calculating CO₂ storage capacity in deep saline aquifers at the basin scale. Energy Procedia 1, 3307–3314.
- Szulczewski, M.L., MacMinn, C.W., Herzog, H.J., Juanes, R., 2012. Lifetime of carbon capture and storage as a climate-change mitigation technology. Proc. Natl. Acad. Sci. U. S. A. 109, 5185–5189.
- Tahmasebi, P., Sahimi, M., Shirangi, M.G., 2018. Rapid learning-based and geologically consistent history matching. Transport in Porous Media 122 (2), 279–304.
- Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. SIAM.
- Tavakoli, R., Yoon, H., Delshad, M., ElSheikh, A.H., Wheeler, M.F., Arnold, B.W., 2013. Comparison of ensemble filtering algorithms and null-space Monte Carlo for parameter estimation and uncertainty quantification using CO₂ sequestration data. Water Resources Research 49 (12), 8108–8127.
- Wang, Z., Harbert, W.P., Dilmore, R.M., Huang, L., 2018. Modeling of time-lapse seismic monitoring using CO₂ leakage simulations for a model CO₂ storage site with realistic geology: Application in assessment of early leak-detection capabilities. International Journal of Greenhouse Gas Control 76, 39–52.
- Wei, L., Saaf, F., 2009. Estimate CO₂ storage capacity of the Johansen formation: numerical investigations beyond the benchmarking exercise. Computational Geosciences 13 (4), 451.
- Wilson, E.J., Johnson, T.L. and Keith, D.W., 2003. Regulating the ultimate sink: managing the risks of geologic CO₂ storage.
- Zimmerman, D.A., De Marsily, G., Gotway, C.A., Marietta, M.G., Axness, C.L., Beauheim, R.L., Bras, R.L., Carrera, J., Dagan, G., Davies, P.B., Gallegos, D.P., 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow. Water Resources Research 34 (6), 1373–1413.
- Zhou, Q., Birkholzer, J.T., Tsang, C.-F., Rutqvist, J., 2008. A method for quick assessment of CO₂ storage capacity in closed and semi-closed saline formation. Int. J. Greenhouse Gas Control 2, 626–639.
- Zhu, T., Ajo-Franklin, J., Daley, T.M., Marone, C., 2019. Dynamics of geologic CO₂ storage and plume motion revealed by seismic coda waves. Proceedings of the National Academy of Sciences 116 (7), 2464–2469.