Integrating geomechanical proxy models with data assimilation for energy transition applications

Saifullin, Ilshat; Seabra, Gabriel Serrão; Pluymakers, Anne; Vossepoel, Femke C.; Voskov, Denis

DOI
10.1016/j.gete.2024.100618

Publication date
2024

Document Version
Final published version

Published in
Geomechanics for Energy and the Environment

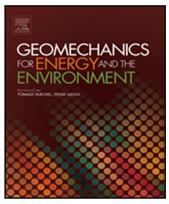
Citation (APA)
Saifullin, I., Seabra, G. S., Pluymakers, A., Vossepoel, F. C., & Voskov, D. (2024). Integrating geomechanical proxy models with data assimilation for energy transition applications. *Geomechanics for Energy and the Environment*, *40*, Article 100618. https://doi.org/10.1016/j.gete.2024.100618

Geomechanics for Energy and the Environment 40 (2024) 100618

Contents lists available at ScienceDirect

Geomechanics for Energy and the Environment

journal homepage: www.elsevier.com/locate/gete



Integrating geomechanical proxy models with data assimilation for energy transition applications

Ilshat Saifullin [a,*], Gabriel Serrão Seabra [a,b,d], Anne Pluymakers [a], Femke C. Vossepoel [a], Denis Voskov [a,c]

[a] Faculty of Civil Engineering and Geosciences, TU Delft, Stevinweg 1, 2628 CN Delft, Netherlands
[b] Petroleo Brasileiro S.A. (Petrobras), Rio de Janeiro, Brazil
[c] Department of Energy Science & Engineering, Stanford University, CA, USA

ARTICLE INFO

Editors-in-Chief:
Professor Lyesse Laloui and Professor Tomasz Hueckel

Dataset link: https://gitlab.com/open-darts/subsidence-esmda

Keywords:
Energy transition
Geomechanics
Subsidence
Uncertainty quantification
Data assimilation

ABSTRACT

This study presents a method to address the significant uncertainties in subsurface modeling that impact the efficiency of energy transition applications such as geothermal energy extraction and CO_2 geological sequestration. The approach combines a physics-based geomechanical proxy model with an ensemble smoother with multiple data assimilation (ES-MDA), aimed at enhancing uncertainty quantification through the integration of vertical displacement measurements from fluid production and injection. The data from wells is limited in spatial coverage, while these measurements offer extensive spatial information, improving the understanding of subsurface behavior by reflecting changes in reservoir pressure and temperature. The open-DARTS simulator for fluid flow and a geomechanical proxy are used to perform data assimilation with ES-MDA. By generating an ensemble of model realizations with varied permeability, calculating vertical displacements at the surface, and applying ES-MDA, we effectively identify the probability distribution of the vertical displacement of the model conditioned to observed subsidence data. Entropy is used as a statistical measure to quantify the reduction of uncertainty of subsurface models based on observations. Our approach was tested on a 2D conceptual and 3D realistic datasets, demonstrating its capability to provide data assimilation. This workflow represents an advancement in subsurface modeling, supporting informed decision-making in geothermal energy production and CO_2 sequestration by offering an improved alternative for data assimilation and enhancing tools for uncertainty quantification.

1. Introduction

The transition toward renewable sources of energy and the need to manage carbon emissions call for efficient and safe utilization of subsurface resources. The strategies needed to achieve such complexity are outlined with the scenarios by international organizations such as the International Energy Agency (IEA), pointing to the needed marked difference in sources of energy generation to achieve the set levels of greenhouse gas reductions. This is clear in the "World Energy Outlook 2022"[1] report, which states that sustainable energy solutions rest on the rapid development of technologies and energy sources playing a unique and complementary role. While Geological Carbon Sequestration (GCS) has been singled out as one of the core enabling technologies with the substantive potential for CO_2 emissions reduction in a fossil fuel-constrained scenario, geothermal energy is another important sustainable source of energy, providing a low-emission and reliable source to act as an alternative for the energy transition.

Exploiting subsurface resources involves significant uncertainties in properties such as permeability and porosity, which impact project feasibility, operational efficiency, and environmental safety. These uncertainties necessitate advanced methods for risk management.[2,3] The accuracy requirements for subsurface characterization vary by project, based on local economic, regulatory, and geological factors. In geothermal energy, operators must prevent temperature drops below critical thresholds (e.g., 10 °C) at production wells to avoid premature thermal breakthrough, with thermal plume estimation being even more critical in multiple well cases. For CO_2 sequestration projects, monitoring CO_2 concentrations near sensitive areas like faults or legacy wells

* Corresponding author.
E-mail addresses: I.S.Saifullin@tudelft.nl (I. Saifullin), g.serraoseabra@tudelft.nl (G.S. Seabra), Anne.Pluymakers@tudelft.nl (A. Pluymakers), f.c.vossepoel@tudelft.nl (F.C. Vossepoel), D.V.Voskov@tudelft.nl (D. Voskov).
[d] Contributed equally to this work as the first author.

https://doi.org/10.1016/j.gete.2024.100618
Received 30 May 2024; Received in revised form 26 October 2024; Accepted 21 November 2024
Available online 2 December 2024

is required to meet regulatory standards and assess leakage risks. In both cases, pressure and temperature variations near existing faults must be estimated for induced seismicity risk assessment. Additionally, energy transition operations often lack comprehensive monitoring systems, and integrating new measurement techniques can supplement production data, improving both risk mitigation and performance. As we discuss in this work, surface displacement data can provide insights into reservoir behavior and help mitigate project risks.[4,5] Modern instruments like tiltmeters enable high-precision measurements of these displacements at sub-millimeter scale.[6,7] However, incorporating such data requires advanced geomechanical modeling, which poses challenges due to the complexity of rock-fluid interactions, limited calibration data, and the computational demands of fully-coupled simulations.

To address these challenges, researchers have developed surrogate geomechanical models, including neural networks[8] and physics-based proxies combining analytical and numerical methods. Ref. 9 and Ref. 10 introduced methods using Green's function and the nucleus of strain concept to model stress responses and predict compaction and subsidence. These approaches, validated against finite element models, have been applied to real-world challenges like assessing induced seismicity in the Groningen gas field.[11] We have enhanced these models by incorporating thermoelastic responses, developing a framework for efficient displacement evaluation, and adapting to non-rectangular grids (e.g., Corner Point Geomertry). By integrating our model with the open-DARTS simulator,[12–14] we aim to refine physics-based geomechanical proxy models for geothermal energy extraction and CO_2 sequestration applications. These computationally efficient proxies enable advanced data assimilation techniques that would otherwise be prohibitively expensive, addressing the challenges of subsurface modeling and uncertainty quantification in energy transition applications.

Data-assimilation methods, including ensemble-based strategies like EnKF and ES-MDA,[15] and gradient-based methods such as RML[16,17] and 4DVar,[18–20] integrate observational data with model realizations to refine numerical model accuracy. ES-MDA was successfully applied in various subsurface applications.[21–23] This method's iterative nature, coupled with the use of an inflated covariance matrix, allows for more accurate and robust updates to the model, even with the sparse and diverse data characteristic of subsurface modeling challenges.[24] We incorporate ES-MDA to reduce uncertainty in reservoir permeability, enhancing understanding of key variables in geothermal and GCS projects. While vertical displacements may be subject to uncertainties in rock geomechanical properties, the comparative behavior of models remains important for estimating thermoporoelastic responses. Our framework is tested on a 2D conceptual dataset with Gaussian distributed permeability, followed by application to the more complex Brugge Dataset.[25] While there is no standardized metric for uncertainty reduction in geothermal and GCS applications, our approach demonstrates significant improvements in subsurface characterization. Using entropy as a measure of uncertainty, we achieved reductions of up to 53.1% for CO_2 sequestration and 17.3% for geothermal applications. These improvements substantially enhance decision-making processes and risk assessment in subsurface resource management and energy transition initiatives. This sequence evaluates the model's performance from simplified contexts to real-world simulations, demonstrating its applicability for decision-making in energy transition projects.

In summary, the main contributions of this work include the development and application of a physics-based geomechanical proxy model, which is integrated with ES-MDA. Additionally, this study enhances the monitoring and management of CO_2 injection and geothermal energy extraction through the data assimilation of surface displacement measurements into subsurface models. This method has been effectively applied to both a conceptual 2D dataset and a realistic scenario inspired by the Brugge dataset.

This paper covers the methodology, including the geomechanical proxy model development and ES-MDA framework, followed by test cases for geothermal and CO_2 injection models. Results are presented for both conceptual and Brugge dataset grids, with a final discussion synthesizing the findings and their implications for energy transition applications.

2. Methodology

2.1. Hydrodynamic model

For the hydrodynamic simulation part of our methodology, we used open Delft Advanced Research Terra Simulator (open-DARTS). open-DARTS is an open-source reservoir simulation software with a focus on energy transition applications. Within this framework, we simulated the operation of a geothermal field using a single-component physics model, and modeled CO_2 injection scenario employing compositional thermal physics, described in Section 3. In our study we leverage the Operator-Based Linearization (OBL) interpolation caching feature of open-DARTS simulator. An OBL approach exploited in open-DARTS replaces the computation of fluid and rock properties at a particular state (pressure, temperature and composition) with the construction of parameter space and approximation of PDE operators in it using multi-linear interpolation[26] which has been shown its applicability to both geothermal[27] and compositional physics relevant to GCS.[28] The advantage of the OBL approach is that once the operator's space has been constructed and dynamically filled with the computed operators, it can be further reused, thereby diminishing the time required for Jacobian evaluations. Moreover, this feature facilitates the recycling of computation points from previous simulations when conducting subsequent model simulations within the same ensemble, and across all iterations of the ES-MDA. This approach reduces the computational cost associated with running extensive model simulations.

The source/sink is implemented in open-DARTS as an additional term in the PDEs, which is non-zero for the reservoir blocks with well perforations. The inflow/outflow is proportional to so called well index, computed by the commonly used Peaceman's formula and the pressure difference between the perforated reservoir block and the bottom hole pressure defined at the well. This approach does not require mesh refinement, which would increase computational time. Instead, we add a source/sink term in the PDEs for the reservoir blocks containing well perforations. The well index, which determines the pressure change in the cell with a well perforation for a rate-controlled well, is computed using Peaceman's formula. This index depends on the well radius, grid cell size, and permeability. Notice that for the transient analysis, the refined mesh should be implemented around the well.[29]

One of the recently implemented features of open-DARTS is a fully-coupled geomechanical modeling using Finite Volume discretization Method with a multi-point flux and stresses approximations.[30,31] Although the coupled solver allows us to neglect some assumptions we have made in this study and resolves the geomechanical state precisely, in order to preserve performance, we used a proxy geomechanical model described below. The mathematical formulation and the solution approach are described in the Appendices A and B correspondingly.

2.2. Geomechanical proxy model

In this Section, we describe a developed numerical methodology of thermoporoelastic rock response approximation based on the nucleus of strain approach. This model, referred to as a geomechanical proxy model, calculates the poroelastic and thermoelastic displacements at the surface for the requested points at the surface using pressure and temperature changes from a hydrodynamic simulator. One of the advantages of this approach is that evaluation points do not have to be at cell centers or nodes of the original grid as in fully-coupled geomechanical simulators. The nucleus of strain concept allows the computation of displacements at any location, even outside the reservoir grid. Thus, the model can provide the values of the vertical displacements at the

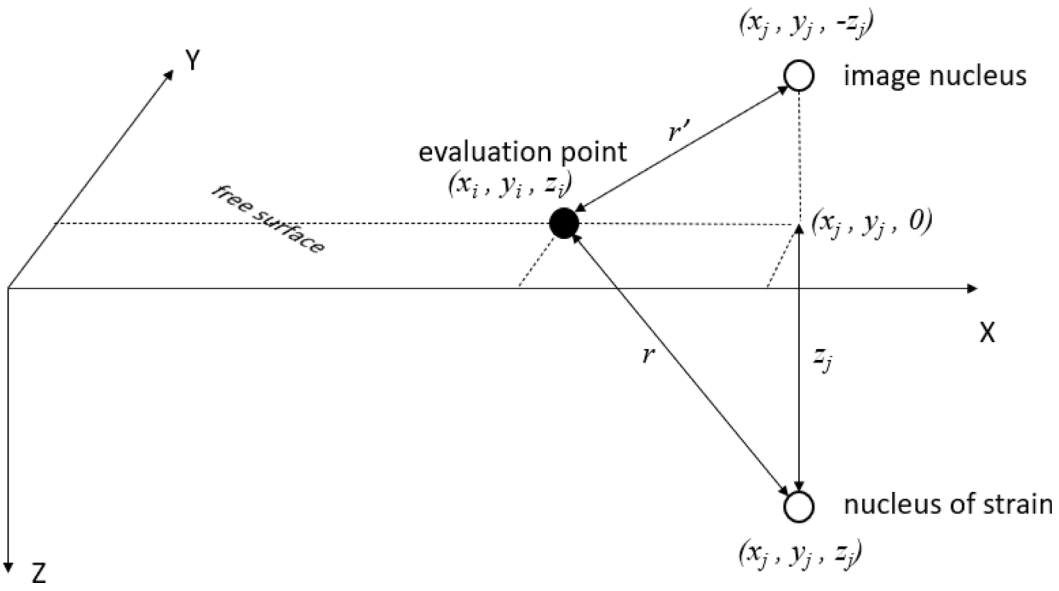


Fig. 1. The image nucleus illustration for the semi-infinite reservoir.

surface without constructing a computational grid up to the surface. The inputs of the proxy model are: cells geometry, pressure and temperature changes at each cell and geomechanical properties (Poisson ratio, Young modulus, and thermal expansion coefficient). Note that this approach has the limitation that geomechanical rock properties should be homogeneous (also in the surroundings of the reservoir) as it has been derived from an analytical solution.

The poroelastic displacement potential at evaluation point formulated by Ref. 32 is

$$\Phi = -\frac{C_m V \Delta p}{4\pi R}, \tag{1}$$

where R is the distance from the evaluation point to the nucleus, V is the volume of the nucleus, and Δp is the pressure change and C_m - uniaxial compaction coefficient [1/MPa] defined as

$$C_m = \frac{(1+\nu)(1-2\nu)}{\mathrm{E}(1-\nu)}. \tag{2}$$

The gradient potential gives the solution for the displacements[33]:

$$\begin{pmatrix} u_x^p \\ u_y^p \\ u_z^p \end{pmatrix} = \begin{pmatrix} \frac{\partial \Phi}{\partial x} \\ \frac{\partial \Phi}{\partial y} \\ \frac{\partial \Phi}{\partial z} \end{pmatrix} \tag{3}$$

Assuming there are N grid cells (nucleus of strain) with a pressure change $\Delta p_j = p_{initial} - p_{current}$ [MPa], and evaluation points with index i having coordinates (x_i, y_i, z_i), the analytic formulas for the poroelastic displacements for an infinite domain are derived using Ref. 32. This nucleus-of-strain concept is extended to a half-infinite space using so called image nucleus[34] leading to additional terms which will be indicated further. The image nucleus is schematically illustrated in Fig. 1 and has been derived using a condition that the stresses acting through the free surface (the horizontal plane $z = 0$) are nullified.[33]

The poroelastic displacements at point i are evaluated as a superposition of displacement terms that represent the influence of all nuclei of strain, which can be represented by the grid cells:

$$\begin{pmatrix} u_x^p \\ u_y^p \\ u_z^p \end{pmatrix}_i = -\frac{C_m}{4\pi} \sum_{j=1}^{N} \Delta p_j \begin{pmatrix} G_x \\ G_y \\ G_z \end{pmatrix}_{ij}, \tag{4}$$

with the geometrical part

$$\begin{pmatrix} G_x \\ G_y \\ G_z \end{pmatrix}_{ij} = \int_{V_j} \begin{pmatrix} \frac{\partial}{\partial x} 1/r \\ \frac{\partial}{\partial y} 1/r \\ \frac{\partial}{\partial z} 1/r \end{pmatrix}_{ij} dv_j + C_v \int_{V_j} \begin{pmatrix} \frac{\partial}{\partial x} 1/r' \\ \frac{\partial}{\partial y} 1/r' \\ -\frac{\partial}{\partial z} 1/r' \end{pmatrix}_{ij} dv_j + 2z_j \int_{V_j} \begin{pmatrix} \frac{\partial^2}{\partial x \partial z} 1/r' \\ \frac{\partial^2}{\partial y \partial z} 1/r' \\ \frac{\partial^2}{\partial z^2} 1/r' \end{pmatrix}_{ij} dv_j, \tag{5}$$

where the first term represents the infinite-space solution using (1) and (3), whereas the next two terms are the semi-infinite space corrections,[35] r - the distance vector between the evaluation point i and jth nucleus of strain, i.e. the jth prism center, z_j - the depth of jth prism, dv_j - infinitesimal element of volume located in jth nucleus of strain, r' - the distance vector between the evaluation point i and jth image nucleus of strain, $C_v = 3 - 4\nu$.

Similarly to Eq. (4), the thermoelastic displacements at point i can be computed:

$$\begin{pmatrix} u_x^T \\ u_y^T \\ u_z^T \end{pmatrix}_i = -\frac{\alpha_T}{4\pi} \sum_{j=1}^{N} \Delta T_j \begin{pmatrix} G_x \\ G_y \\ G_z \end{pmatrix}_{ij}, \tag{6}$$

where ΔT_j [°C] is the temperature change and α_T [1/°C] is a linear thermal expansion coefficient. Finally, thermoporoelastic displacements are the sums of poroelastic and thermoelastic displacements:

$$u = u^p + u^T. \tag{7}$$

Ref. 35 and Ref. 10 present validations of the described geomechanical proxy model method for the poroelasticity problem, comparing the solution in elastic homogeneous cylindrical reservoir with Geertsma's method for the uniform pressure depletion. Ref. 10 also validates the approach against Finite Element Method solution for the non-uniform pressure depletion and the arbitrary shaped reservoir.

In this formulation, for the poroelastic and thermoelastic displacements have a common geometric term, which has been utilized in the implemented code.

It is worth emphasizing that this model has intrinsic mechanical boundary conditions, such that the displacements at boundaries are expected to be non-zero and the resulting vertical displacements are expected to be asymmetric on the top and bottom boundaries of the reservoir grid.

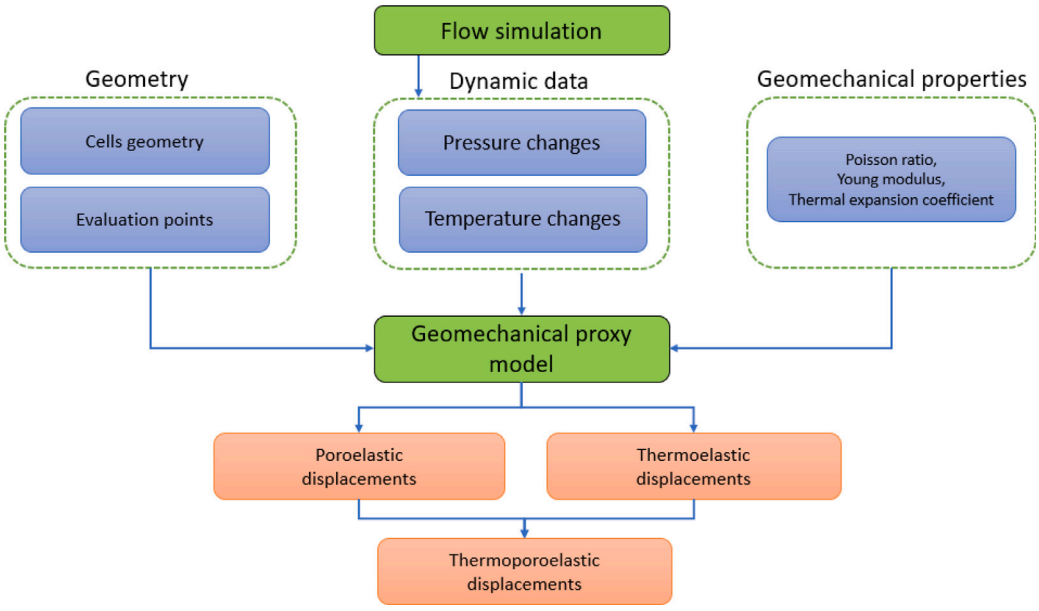


Fig. 2. Scheme of the geomechanical proxy model of displacements evaluation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The flow chart of the proxy model is shown in Fig. 2, where the inputs are in blue boxes, and the outputs are in red boxes. The displacement evaluation is implemented in C++ programming language and parallelized using OpenMP. Parallelization is straightforward, as the evaluation of the displacement at different locations is independent. A Python interface is created using the pybind11 open-source library, and the geomechanical proxy model is integrated into the open-DARTS hydrodynamic model and ES-MDA using Python. In our implementation, the proxy model is linked with the hydrodynamic simulator at runtime with displacement computation at desired time steps. However, it can also be applied as a post-processing tool. An analytic expression for the volumetric integrals in Eq. (5) has been derived by Ref. 36 for the grid aligned to XYZ-axes and consisting of cells of a rectangular prism shape.

In real-world applications, grids usually consist of complex cell geometries and they are specifically oriented. For such cases, we implemented and tested two approaches of the proxy model application. In the first approach, we construct an auxiliary structured grid, which fulfills the requirements and interpolates the pressure changes Δp and the temperature changes ΔT to it. The auxiliary grid is then used in the geomechanical proxy model. This is illustrated schematically in Fig. 3a, where the original grid is shown with black lines and the geomechanical grid with white lines. The disadvantage of the interpolation approach is that a structured grid of rectangular prisms should be refined enough to approximate the original mesh.

Another way of different cell shape and orientation treatment is the construction of a rectangular prism with averaged cell sizes of the original cell and oriented to the XYZ-axes. The idea for a 2D case is shown in Fig. 3b which extension to 3D is straightforward. The cells with dashed black borders are cells of the original grid and the blue cells are the rectangular prisms with geometrically averaged sizes which can be used in the geomechanical proxy model. The case on the left side represents a rectangular prism, which violates the axes' orientation condition. The case on the right side shows a cell of the original grid which violates both shape and orientation conditions.

The presented proxy method assumes the homogeneous and isotropic rock geomechanical properties and linear poroelastic medium. Since the method explicitly evaluates the displacements from the pressure and temperature changes, it ignores the geomechanical effects such as porothermoelastic diffusion. Thus, it does not capture the time-dependent geomechanical response, e.g. the consolidation effect. We also did not consider the stress-dependent behavior of the rock geomechanical properties in this study. Nevertheless, the proxy method can be a good alternative for preliminary evaluations in complex problems, including heterogeneous reservoirs using averaged homogeneous rock geomechanical properties.

2.3. Prior distribution of reservoir properties for data assimilation

We generate prior distributions for the logarithm of reservoir permeability, a common practice in uncertainty quantification studies.[8,15,37] The spatial field for the logarithm of permeability, $\log p(x, y)$, is represented as a Gaussian distribution:

$$\log p(x, y) \sim \mathcal{N}(\mu, \Sigma),$$

where μ represents the mean of the logarithm of permeability, and Σ is the covariance matrix derived from the specified spatial correlation function. The correlation between the values at two points (x_0, y_0) and (x_1, y_1) on the grid is modeled using a modified Gaspari–Cohn function, providing a smooth correlation[38]:

$$C(x_0, y_0, x_1, y_1) = \begin{cases} -\frac{h^5}{4} + \frac{h^4}{2} + \frac{5h^3}{8} - \frac{5h^2}{3} + 1 & \text{for } h < 1, \\ \frac{h^5}{12} - \frac{h^4}{2} + \frac{5h^3}{8} + \frac{5h^2}{3} - 5h + 4 - \frac{2}{3h} & \text{for } 1 \le h < 2, \\ 0 & \text{for } h \ge 2. \end{cases}$$

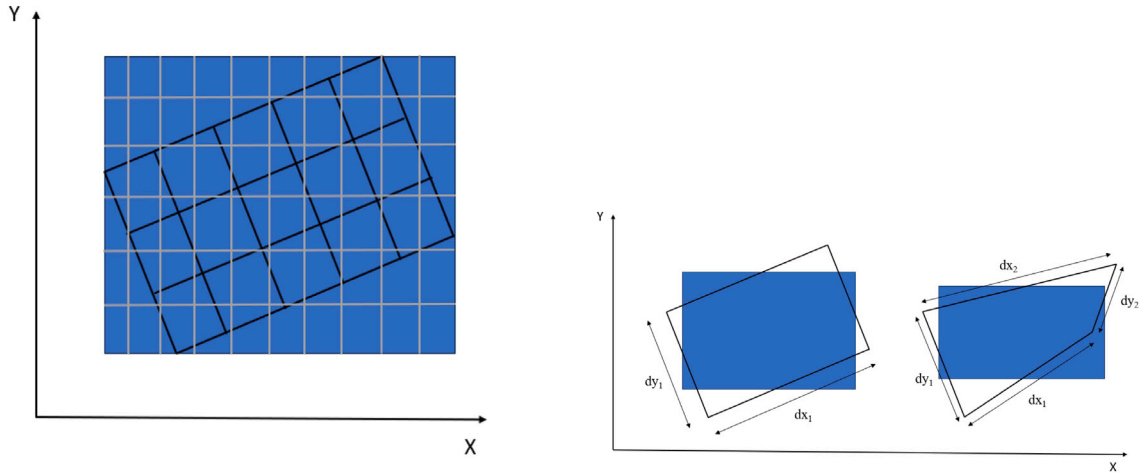
In this equation, h is the normalized distance scaled by correlation lengths L_x and L_y, and rotated by angle θ to align with principal axes of anisotropy:

$$h = \sqrt{\left(\frac{\Delta x'}{L_x}\right)^2 + \left(\frac{\Delta y'}{L_y}\right)^2},$$

$$\Delta x' = \cos(\theta)\Delta x - \sin(\theta)\Delta y, \quad \Delta y' = \sin(\theta)\Delta x + \cos(\theta)\Delta y.$$

This covariance structure was discretized over the grid to form the matrix Σ, and the Cholesky decomposition of Σ was used to simulate spatially correlated random fields that form the ensemble of log-permeability fields:

$$m_i = \mu + L_{\text{chol}} z_i,$$



(a) Original grid for flow simulation and auxiliary grid for the geomechanical proxy model (interpolation)

(b) Original grid cell and averaged rectangular prism for geomechanical proxy method

Fig. 3. Two approaches of a proxy model application to CPG grid: (a) interpolation, (b) averaging.

where L_{chol} is the Cholesky factor of Σ, and z_i is a vector of standard normal random variables.

The implementation involves generating a grid-based model domain, computing the covariance matrix using the Gaspari–Cohn correlation function, and using Cholesky decomposition to generate multiple realizations of the log-permeability field. This approach ensures that the simulated fields respect specified statistical properties, aligning with geostatistical practices as illustrated in SGEMS and GSLIB user guides.[37,39]

2.4. Data assimilation framework with ES-MDA

The ES-MDA technique is applied to assimilate observations into open-DARTS reservoir model realizations, using the physics-based geomechanical proxy model to account for vertical displacements computation. This method is an excellent choice for this problem, as it offers flexibility by requiring minimal alterations to the existing forward model code and avoiding the computation of adjoints gradients.[18] ES-MDA performs data assimilation through iterative assimilation of the same observations, adjusting the influence of the observations at every assimilation step to enhance model updates. This adaptation allows taking into account non-linearities in the forward model through its step-wise approach,[15,24] mainly vertical displacements measured at the surface.

We can describe the analysis for each ensemble member's model parameter $\boldsymbol{z}_j^a$, in our case, the reservoir model permeabilities vector $\boldsymbol{k}_j^a$ as follows:

$$\boldsymbol{z}_j^a = \boldsymbol{z}_j^f + C_{ZD}^f \left(C_{DD}^f + \alpha_i C_D \right)^{-1} \left(\boldsymbol{d}_j - G(\boldsymbol{z}_j)^f \right), \quad \text{for } j = 1, 2, \dots, N_e. \tag{8}$$

Here N_e is the total number of ensemble members, $\boldsymbol{z}_j^a$ and $\boldsymbol{z}_j^f$ represent the analyzed and forecasted permeabilities of the jth ensemble member, respectively, C_{ZD}^f and C_{DD}^f are the cross- and auto-covariance matrices of the model permeabilities and vertical displacements observations in the forecast step, α is a scaling factor, and $\boldsymbol{d}_j$ and $G(\boldsymbol{z}_j)^f$ are the perturbed and forecasted observations for the jth ensemble member. Besides, $G()$ represents the forward model. For more details, see Eq. (9):

$$\boldsymbol{d}_j = \boldsymbol{d}_{\text{obs}} + \sqrt{\alpha} C_D^{1/2} r. \tag{9}$$

In this equation, $\boldsymbol{d}_{\text{obs}}$ is the vertical displacements observed data, C_D is the measurement error covariance, and r is sampled from a standard normal distribution with zero mean and an identity matrix as the covariance. The factor $\sqrt{\alpha}$ scales the perturbations. In this study, we employed 4 ES-MDA iterations, with $\alpha = 0.25$.

The ES-MDA workflow, summarized below, iteratively employs these equations to update each ensemble member.

1. **Initialization**: Initialize the ensemble with randomized reservoir permeability parameters, collect measured vertical displacement data $\boldsymbol{d}_{\text{obs}}$, and define the covariance matrix for measurement errors C_D.
2. **Iteration Parameter Determination**: Establish the number of assimilation cycles N_a and select an inflation factor α for the iterations, spanning $i = 1, \dots, N_a$.
3. **Iterative Update Cycle**: Execute the following steps in each iteration for every ensemble member:

   (a) **Prediction Phase**: Calculate the expected vertical displacements $G(\boldsymbol{z}_j)^f$ based on the current estimates of reservoir permeabilities.

   - Compute pressure and temperature fields using open-DARTS, incorporating the updated permeability fields.
   - Calculate vertical displacements employing the geomechanical proxy, which integrates the pressure and temperature effects from open-DARTS to predict deformation.

   (b) **Observation Perturbation**: Formulate $\boldsymbol{d}_j$ by implementing the perturbed observation technique to consider observational uncertainty.

   (c) **Parameter Update**: Adjust the reservoir permeability parameters $\boldsymbol{z}_j^a$ via the update equation to enhance alignment between the predicted and observed vertical displacements.

4. **Posterior Assessment**: Upon completing the iterative process, assess the adjusted ensemble of permeabilities $\boldsymbol{z}_j$ using open-DARTS for the final posterior computation of the pressure and temperature field and the geomechanical proxy for the vertical displacements field.

To quantitatively assess the reduction in uncertainty achieved through ES-MDA, we employ Shannon entropy.[40] The Shannon entropy, H, for a discrete probability distribution $p(x)$ over variable X is defined as:

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log(p(x_i)), \tag{10}$$

where $p(x_i)$ is the probability of the ith outcome among n possible outcomes, and log denotes the natural logarithm. In the context of ES-MDA, we compute the entropy of probability maps reflecting critical reservoir characteristics (e.g., CO_2 molar fraction above a safety threshold or critical reservoir temperatures) before and after data assimilation. The reduction in entropy is calculated as:

$$\Delta H = H(\text{prior}) - H(\text{posterior}). \tag{11}$$

Here, ΔH quantifies the uncertainty reduction, with a higher value indicating greater reduction. To provide a normalized measure of improvement, we compute the relative entropy reduction:

$$\text{Relative Entropy Reduction (\%)} = \left(\frac{H(\text{prior}) - H(\text{posterior})}{H(\text{prior})}\right) \times 100. \tag{12}$$

This approach provides a quantitative and interpretable measurement of uncertainty reduction, allowing for an objective assessment of the ES-MDA's effectiveness in improving model certainty. The entropy calculations are implemented using probability maps generated from ensemble simulations, estimating the likelihood of exceeding specific critical thresholds for variables of interest at each grid point. The entropy is measured in *nats*, as it is calculated using the natural logarithm.

3. Test cases

We used two grids: a conceptual model with a structured grid and a Brugge grid upscaled to one layer. In both cases, the model setups are designed to test the ES-MDA and geomechanical proxy integrated framework. The conceptual model allows for controlling the known variables in a more predictable environment, whereas the Brugge dataset model introduces complexities that challenge the method's capability to handle real-world unpredictability and heterogeneity.

The initial pressure and temperature are defined by a gradient of 100 bar/km and a thermal gradient of 30 °C/km from 1 bar and 20 °C at the surface, respectively.

The flow boundary conditions are defined by assigning a large effective volume to the lateral boundary cells, thereby simulating an infinite reservoir, while enforcing no-flow restrictions at the top and bottom boundaries. The impact of overburden and underburden layers in the hydrodynamic model is out of scope for this study. This leads to the assumption that the pressure and the temperature remain constant above and below the reservoir. The assumption of no-flow boundary conditions for the reservoir top and bottom is reasonable for most of the applications. For example, in CO_2 injection its leakage from the reservoir is commonly avoided and reservoir with sufficient cap rock integrity is preferable to be chosen. Although the temperature change is mostly driven by the convection, additional layers should be added in order to account for the vertical heat conduction. Since the geomechanical proxy model uses only the pressure changes in the reservoir part, imposing drained boundary conditions at the reservoir top and bottom can impact the geomechanical response, the reservoir grid should be extended vertically up to non-permeable barriers to properly evaluate the displacements. The presence of the pressure changes in the lateral reservoir surrounding is neglected in our study since one can extend the reservoir laterally as well until the impact to the displacements in the interest area will be negligible due to the large distance and low pressure changes there.

The simulation time is 30 years. In the Data assimilation Section 5, the simulation runs only for the first 10 years to imitate the field development-in-progress study. However, application to the whole time-series might improve the data assimilation results. The well rates for each scenario are listed in Table 1.

3.1. Geothermal model

The geothermal model is based on a single-component two-phase formulation with dynamic water properties evaluated using IAPWS97 incorporated into the open-DARTS framework.[27,41] The temperature ranges in the cases we used correspond to one phase regime. Operational controls for wells are established as rate control listed in Table 1 with a bottom hole pressure (BHP) constraint maintaining an 80 bar differential relative to initial reservoir conditions. We align the injection well's control with the preceding time step's production rate for the geothermal case, ensuring a continuous re-injection process with water density adjustment to ensure the mass balance due to the temperature difference between the produced and injected liquid. This is particularly important in scenarios where low permeability near the production well prompts a shift to BHP control, thus potentially restricting output to the predefined rate.

3.2. Model for CO_2 injection

The GCS model operates within the open-DARTS compositional formulation, which utilizes equilibrium flash calculations to establish phase distributions.[42] Comprising CO_2 and H_2O components, the system is initially at water saturation with CO_2 as the injection component in the liquid phase. The simulation time consists of 10 years of injection proceeded with 20 years of monitoring.

3.3. Conceptual reservoir

A conceptual reservoir model with 2D structured grid utilized in this study is a simplified representation to emulate the geothermal and GCS scenarios for our experiments (Fig. 4). The reservoir comprises a $51 \times 51 \times 1$ structured grid with right prism-shaped cells and uniform properties. In the geothermal case, the model features a production well and an injection well positioned along the X-axis in the middle of the grid. For the GCS case, the injection well is situated in the center of the grid to explore the radial flow and pressure propagation characteristics typical of single-well injection systems. The rock properties, maintained uniform across the model, are specified within the accompanying Table 1. We will plot hydrodynamic model variables along the monitoring points line at reservoir depth and geomechanical variables (vertical displacements) at the surface (depth = 0).

3.4. Brugge dataset

The Brugge dataset offers a more challenging and realistic scenario for our data assimilation and subsurface modeling approaches. Unlike the conceptual model, the Brugge dataset model includes complex geological features and heterogeneity that are common in actual reservoirs. This model is constructed using a Corner Point Geometry (CPG) grid, which is commonly used to capture the geological layers and faults within the reservoirs. The geometry of the Brugge dataset has been upscaled to one layer and permeability from the first layer of the original model is used. The rock properties, well rate and geometry summary are specified in Table 1. The grid contains 6672 cells with 4950 active cells. The grid exaggerated 5 times by Z-axis, the permeability field used for the single run and monitoring points distributed along the wells locations as displayed in Fig. 5. Well locations for the geothermal case are I-index=37 and J-index=31 for the production

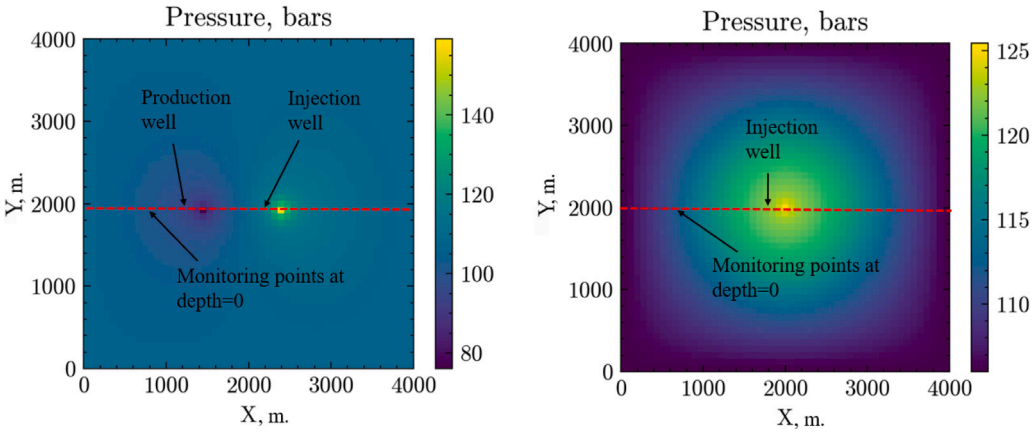
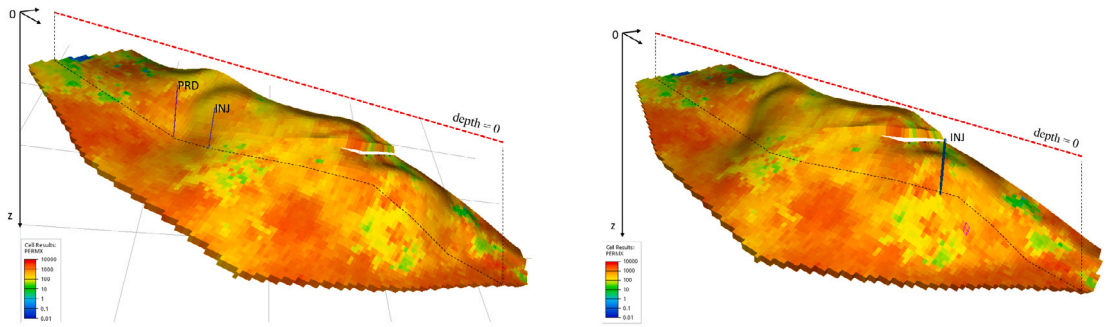


Fig. 4. 2D Conceptual Model setup showing the grid layout (top view), well positions, and monitoring points for the geothermal extraction case (left) and GCS (right).



(a) Geothermal case (b) GCS case

Fig. 5. Brugge dataset model setup illustrating CPG grid, well locations, and monitoring points along J-index=31.

well and I-index=46 and J-index=31 for the injection well. GCS model setup has a different injection well location situated at I-index=100 and J-index=31.

According to the application to the generic grid discussed in 2.2, we used the interpolation in order to apply the geomechanical proxy method for Brugge dataset geometry. Two different resolutions are used for the construction of a structured grid: with 5 and 10 refinement factors of the original grid.

The results showed that the basic method of a cell approximation with a right prism and averaged sizes provides a solution for the subsidence/uplift similar to the interpolation with a refinement factor of 10. Meanwhile, the implementation with interpolation significantly increases the number of prisms which affects the performance of geomechanical evaluation. Thus, we used a simple averaging approach described in Section 2.2, since data assimilation requires a fast displacement evaluation.

Table 1
Model setup.

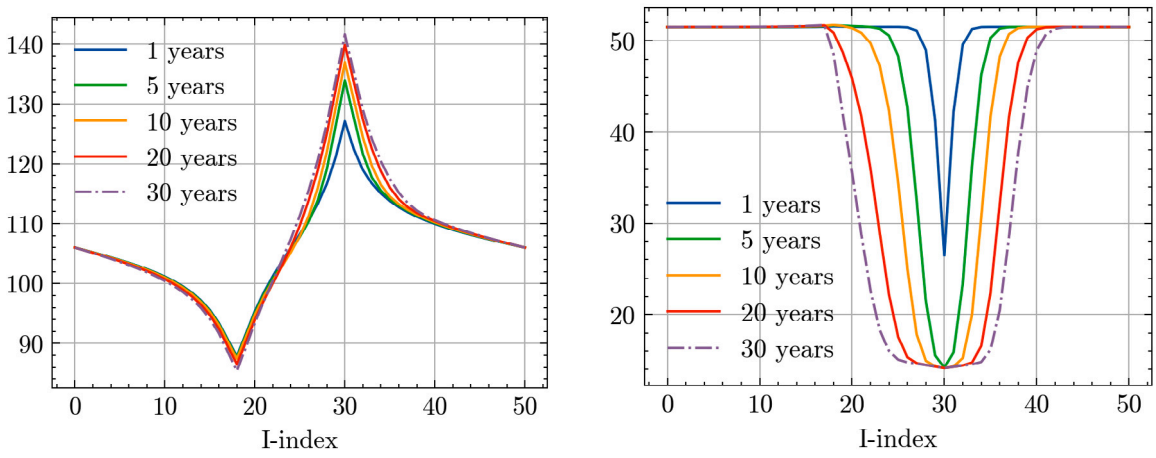
Property	Conceptual case	Brugge case
Grid size (XY), km	4 × 4	17 × 9
Thickness, m	100	58.2-224.7
Grid dimensions	51 × 51 × 1	139 × 48 × 1
Reservoir top depth, km	1000	1.545-1.945
Well spacing (geothermal), km	1	1.25
Production rate (geothermal), m^3/day	7500	7500
Injection rate (GCS), Mt/year	2.5	10
Injection temperature, °C	11.5	36.85
Porosity	0.2	0.18 (mean)
Permeability, mD	100	540 (mean)
Heat capacity, kJ/m^3/K	2200	2200
Heat conductivity, kJ/m/day/K	181.44	181.44
Poisson ratio	0.25	0.25
Young modulus, MPa	3300	3300
Linear thermal expansion coefficient, 1/°C	1.3e−5	1.3e−5

4. Geomechanical proxy results

This section presents an integrated analysis of results from hydrodynamic simulations combined with the application of a geomechanical proxy model, focusing on monitoring points at various times evaluated at the surface for the test cases described earlier. It is important to note that as shown in Fig. 5, for this study, the Z-axis points downwards, so negative values of the vertical displacements at the surface correspond to uplift, and positive values correspond to subsidence.

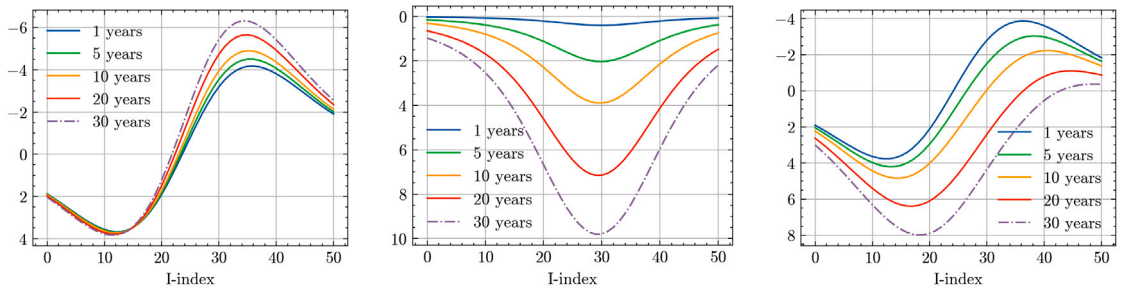
4.1. Geothermal application in conceptual model

In the geothermal scenario, the pressure reaches the steady-state since the model has constant-pressure boundary conditions, except a slight increase in the pressure which is observed around the injection well. This phenomenon is attributed to the reduction in fluid viscosity associated with a drop in the temperature due to the injection of cold fluid, as illustrated in Fig. 6. Due to the approach of well modeling described in the Appendix B and the relatively large size of the reservoir blocks, it takes time for the pressure and temperature in the grid block to approach the injection values in Fig. 6b, since they represent the averaged values in the grid block. The geomechanical proxy model outputs, depicted in Fig. 7a, show variations in vertical poroelastic displacements indicating both uplift and subsidence across the model domain. The maximum uplift observed around the injection well peaks at approximately 6 mm after 30 years of injection, which gradually increases due to the pressure change increase. The propagation of the



(a) Pressure, bars. (b) Temperature, °C.

Fig. 6. Geothermal model results for the conceptual case. 1D plots at the monitoring points at different times.



(a) Poroelastic displacements, mm. (b) Thermoelastic displacements, mm. (c) Thermoporoelastic displacements,

Fig. 7. Geothermal model results for the conceptual case. 1D plots of the displacements at the monitoring points at different times.

thermal front, leads to thermal contraction of the rock which results in the subsidence shown in Fig. 7b. The Fig. 7c depicts the combined porothermoelastic effects. Although the poroelastic effect dominates at the beginning of the production, after 30 years, mostly subsidence is observed as a result of the thermal plume growth in the reservoir.

4.2. Geothermal application in Brugge dataset

The hydrodynamic simulation results for geothermal energy production using the Brugge dataset are shown in Fig. 8. The pressure increase near the injection well is higher than its decrease near the production well due to heterogeneity. Thus, the poroelastic vertical displacement at the surface demonstrate preferably an uplift shown in Fig. 9a increasing at injection area with time due to the water viscosity decrease. The thermoelastic displacement gradually increases as shown in Fig. 9b with the thermal plum propagation shown in Fig. 8b and reaches a similar peak as poroelastic displacements but with opposite sign a bit earlier than 20 years. This determines the shape of thermoporoelastic displacement. At 30 years, there is the subsidence around both wells, and the uplift in the area between the injection well and a boundary shown in Fig. 8c.

In the Brugge grid geothermal case, the temperature at the reservoir grid block containing the injection well drops faster Fig. 8b than in the conceptual grid case since the well index of the injection well in the Brugge case is few times larger compared to the conceptual case.

In fact, the vertical displacements at the surface for the Brugge dataset are lower than the conceptual case due to the Brugge dataset deeper placement, so the elastic impact of the pressure and the temperature changes at the surface is lower.

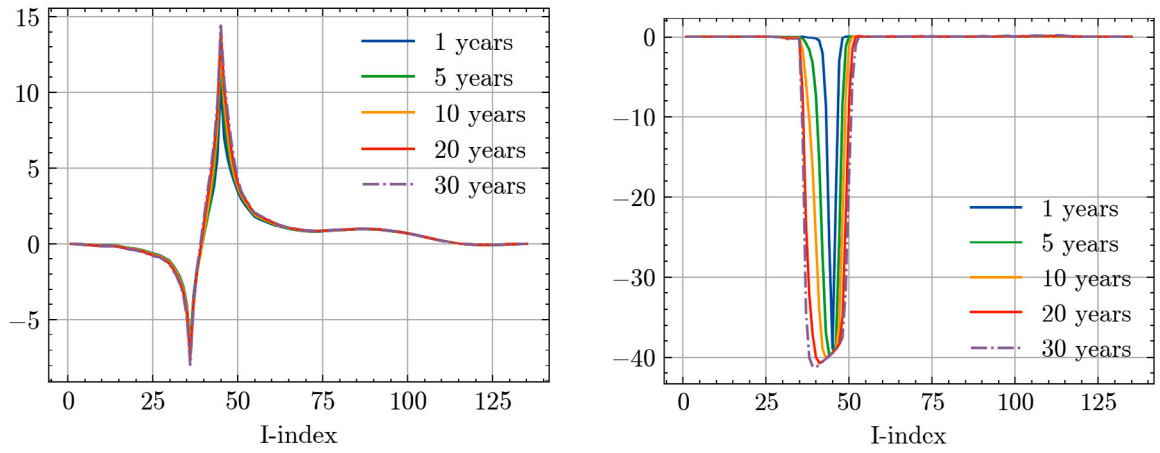
4.3. GCS application in conceptual model

Hydrodynamic model for the GCS scenario results are illustrated in Fig. 10, where the pressure increase in the reservoir leads to the uplift. The conceptual case grid is 2D, so there is no gravity or gas migration after injection stops at 10 years. The pressure change decreases over time as the CO_2 plume propagates, as CO_2 has lower viscosity than water. These results in a decrease of uplift from a peak of 20 mm. (at 1 year) to 15 mm. (at 10 years) are shown in Fig. 11. In the meantime, thermoelastic subsidence is observed due to the rock thermal contraction as the injection gas has a lower temperature with a maximum value of 4 mm. (at 10 years). Thus, thermoporoelastic displacement has a cap-cutted parabola shape. After stopping the injection at 10 years, the pressure change vanishes, which leads to the dissipation of the poroelastic effect. However, the thermoelastic effect stays leading to the subsidence of 1.8 mm. in maximum above the injection well.

4.4. GCS application in Brugge dataset

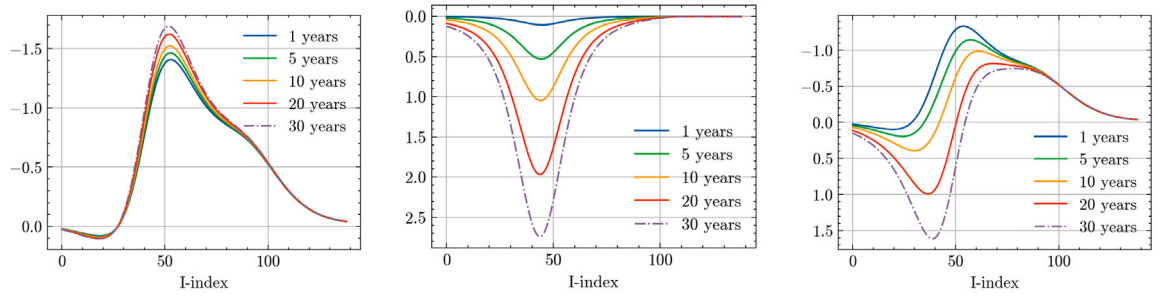
These are the hydrodynamic model simulation results for the GCS model using the Brugge dataset. Fig. 12 shows the pressure and temperature distributions and the CO_2 saturation at different times. The corresponding elastic displacements at the surface are illustrated in Fig. 13. The overall behavior is similar to the conceptual case results. However, the results slightly differ due to the heterogeneity and gravity effects. Similarly to the geothermal model results, the vertical displacement amplitude at the surface for the Brugge dataset is lower than the conceptual case due to the deeper location of the Brugge dataset grid.

Apart from the conceptual case, the Brugge grid cells have different depths, and gas migration can be observed due to the gravity effect



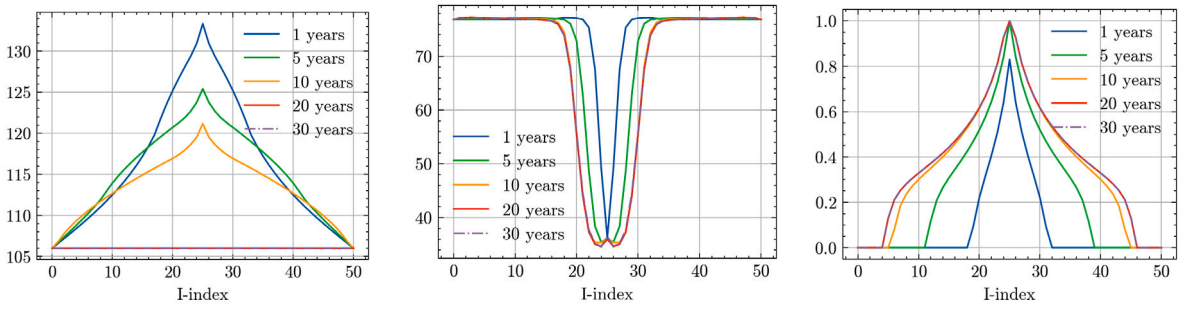
(a) Pressure change, bars. (b) Temperature change, °C.

Fig. 8. Geothermal model results for the Brugge dataset. 1D plots at the monitoring points.



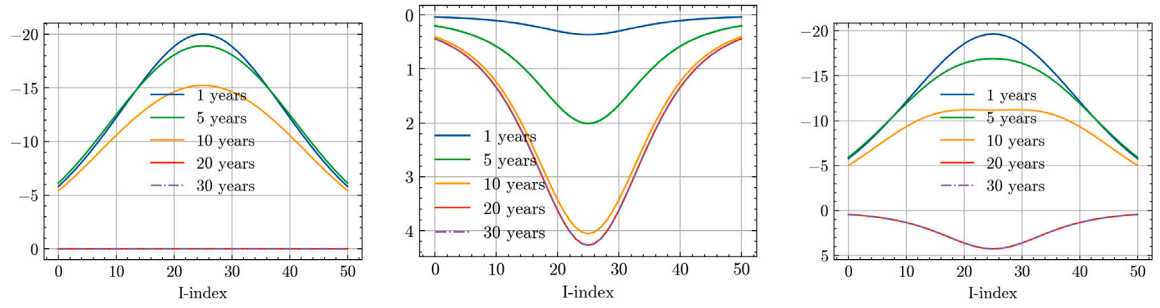
(a) Poroelastic displacements, mm. (b) Thermoelastic displacements, mm. (c) Thermoporoelastic displacements,

Fig. 9. Geothermal model results for the Brugge dataset. 1D plots of the displacements at the monitoring points.



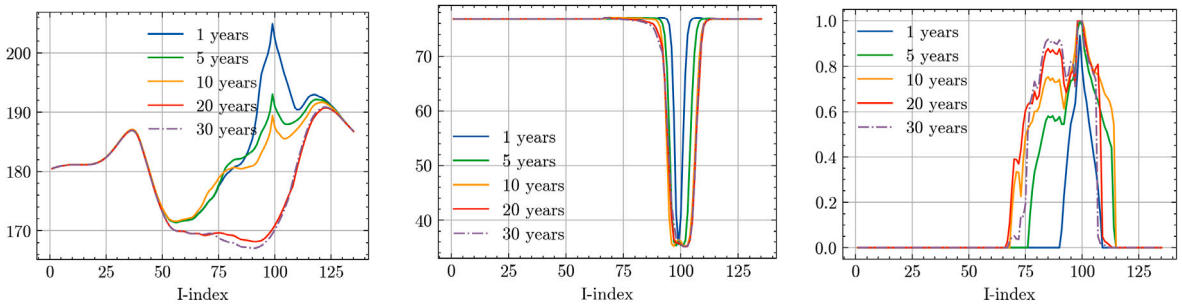
(a) Pressure, bars. (b) Temperature, °C. (c) CO_2 saturation

Fig. 10. GCS model results for the conceptual case. 1D plots at the monitoring points at different times.



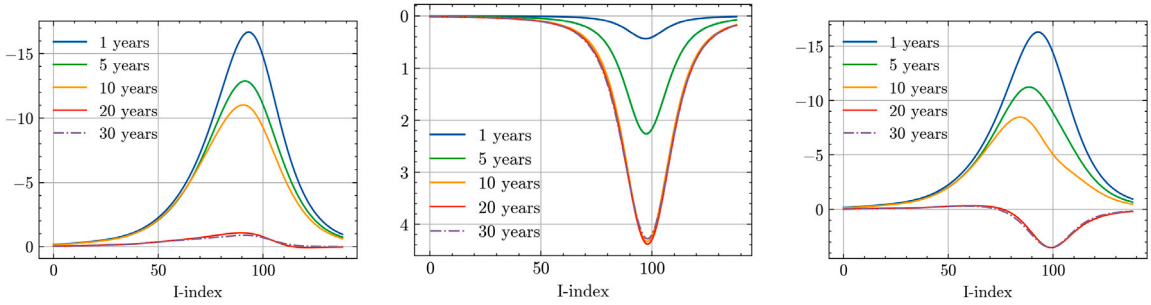
(a) Poroelastic displacements, mm. (b) Thermoelastic displacements, mm. (c) Thermoporoelastic displacements, mm.

Fig. 11. GCS model results for the conceptual case. 1D plots of the displacements at the monitoring points at different times.



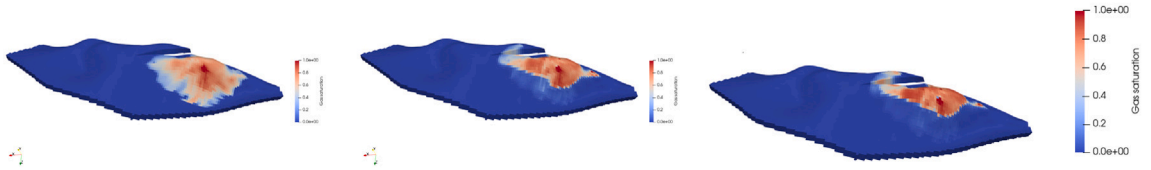
(a) Pressure, bars. (b) Temperature, °C. (c) CO_2 saturation

Fig. 12. GCS model results for the Brugge dataset. 1D plots at the monitoring points at different times.



(a) Poroelastic displacements, mm. (b) Thermoelastic displacements, mm. (c) Thermoporoelastic displacements, mm.

Fig. 13. GCS model results for the Brugge dataset. 1D plots of the displacements at the monitoring points at different times.



(a) CO_2 saturation at 10 years (b) CO_2 saturation at 20 years (c) CO_2 saturation at 30 years

Fig. 14. GCS model results for the Brugge dataset. CO_2 saturation at different times.

after stopping the injection at 10 years, which is shown in Fig. 14. A part of the fault is impermeable due to juxtaposition, which adjusts the gas movement direction. Gas migration leads to the shape change of the CO_2 plume, as illustrated in Fig. 14. The displacements at the surface indicate an uplift at the injection stage, as the poroelastic response is higher than the thermoelastic response working in the opposite direction.

Since we did not aim for detailed CO_2 modeling, we used grids with one layer only. However, CO_2 injection modeling often requires a high-fidelity multilayered grid to properly represent the gas migration, given the fingering effect and gas dissolution in the water.[42,43]

One must notice that the combined thermoelastic and poroelastic effects are highly impacted by the parameters applied in this study, and the final results can change if different parameters are used.

5. Data assimilation

This section outlines results for the application of ES-MDA in a geomechanical proxy model for the conceptual model and Brugge dataset cases. These evaluations reveal the method's efficacy in both synthetic environments and actual field applications when applied to geothermal energy extraction and CO_2 geological sequestration. We generated a prior ensemble for each case as described in Section 2.3, designating one realization as the reference 'true' model. Synthetic observations of vertical displacements were sampled from this reference at the final time step along a transect intersecting the well trajectories (Fig. 4). These data, despite inherent noise, are used to adjust the model's permeability distribution, enhancing productivity of the models and reducing uncertainty in temperature and CO_2 plume distributions for geothermal and GCS cases, respectively.

5.1. Assimilating the 2D conceptual dataset

The prior ensemble consists of 100 permeability fields, with an additional realization serving as the reference model. The reference case has a log-permeability of log(6) mD, with variability between log(3) mD and log(9) mD. Fig. 15 shows the reference model and three prior realizations, illustrating the permeability variability.

5.1.1. Geothermal energy extraction

We apply ES-MDA using vertical displacement measurements captured at the last time step of the production period along a line crossing the well paths on the surface projection. These displacements serve as observations of subsurface changes caused by geothermal operations.

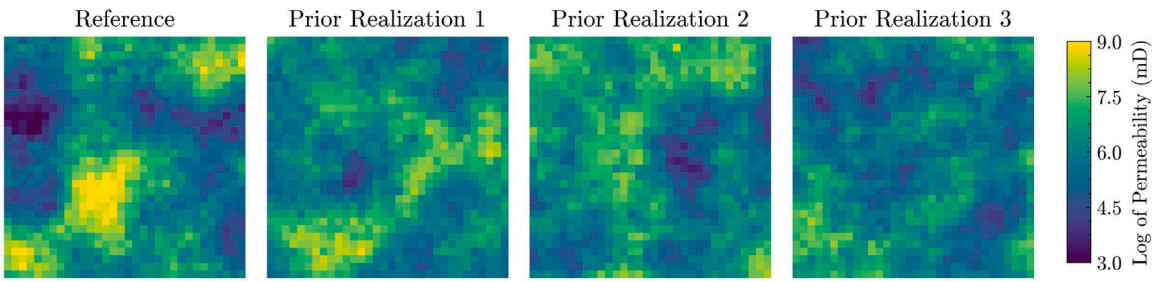


Fig. 15. Reference model and three realizations of the prior permeability distribution for the 2D synthetic dataset.

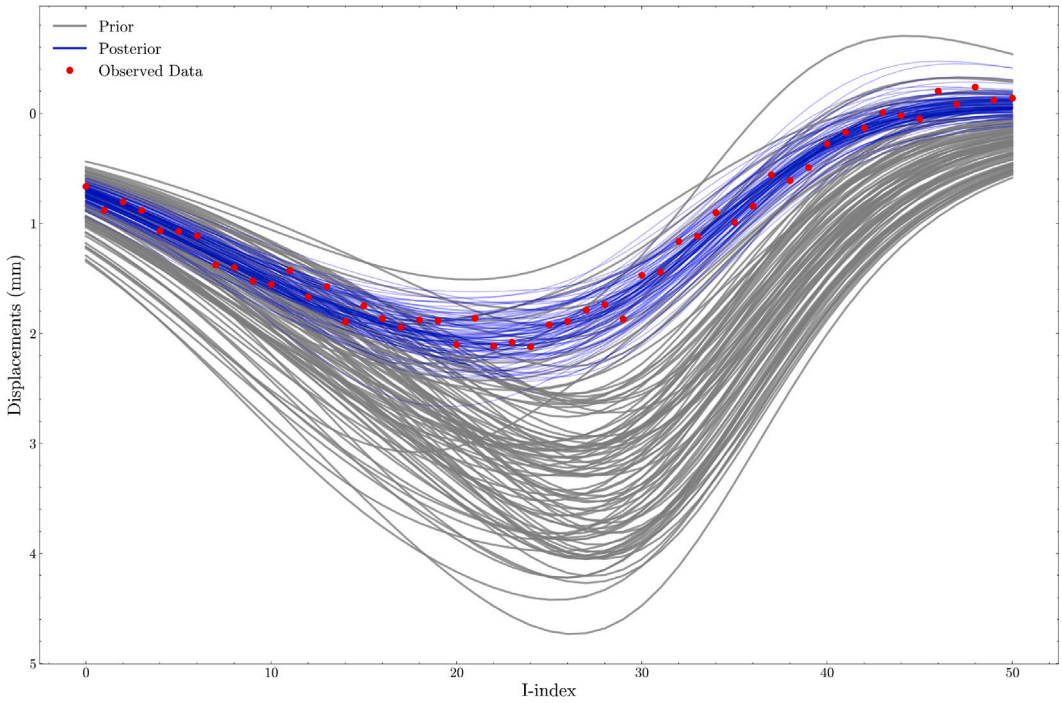


Fig. 16. Prior and posterior ensemble realizations for the conceptual geothermal case, illustrating the adjustments in the model realizations post-ES-MDA. The prior realizations (gray lines) show a wide spread, reflecting initial uncertainties. In contrast, the posterior realizations (blue lines) are tightly clustered around the observed data (red dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 16 illustrates ES-MDA's impact on model realizations. The posterior realizations (blue lines) show significantly improved agreement with observed data (red dots) compared to the prior realizations (gray lines). This indicates a reduction in uncertainty and better integration of observational data. The displacement curves' shape reflects the reservoir's physical response to geothermal operations, showing uplift near the injector and subsidence near the producer.

The difference plots in Fig. 17 show the magnitude and location of permeability adjustments. Areas of pronounced changes indicate where initial model permeability caused significant deviations in reservoir pressure and temperature, impacting surface displacements. While multiple permeability realizations could produce similar subsidence signals, making the recovered field not necessarily the true permeability map, these results are valuable for reducing uncertainty in the posterior states of the reservoir. Constraining the temperature distribution for geothermal applications is often more critical for decision-making than the exact permeability distribution. This showcases the value of integrating surface displacement data into subsurface modeling for geothermal applications, even when the permeability field cannot be uniquely determined. Additional data types could further constrain the permeability field if needed for specific geothermal project requirements.

The histograms and CDF curves in Fig. 18 further quantify the permeability changes caused by the data assimilation. Initially, the variability in permeability is large, as seen in the spread of the prior histograms. After assimilation, the posterior histograms shift, with Realizations 1 and 3 slightly shifting to lower permeability values and Realization 2 shifting to higher permeability values, respectively. The cumulative distribution functions (CDFs) corroborate these shifts. The shapes of the histograms also present small changes, showing that the permeability spread of the models in the posterior ensemble is a simple redistribution of values within the model.

The transition from a broad and dispersed prior understanding to a more concentrated and confident posterior model indicates the capability of ES-MDA in constraining model realizations with observations. It demonstrates a successful data-assimilation process, where the final model not only agrees with the observed data but also does so by acknowledging and accounting for the noise inherent in real-world measurements.To quantitatively evaluate the impact of data assimilation on the probability distribution of temperature within the reservoir, we constructed probability maps considering a threshold of 50 °C. These maps, depicted in Fig. 19, reveal a shift in the probability distribution post-data assimilation.

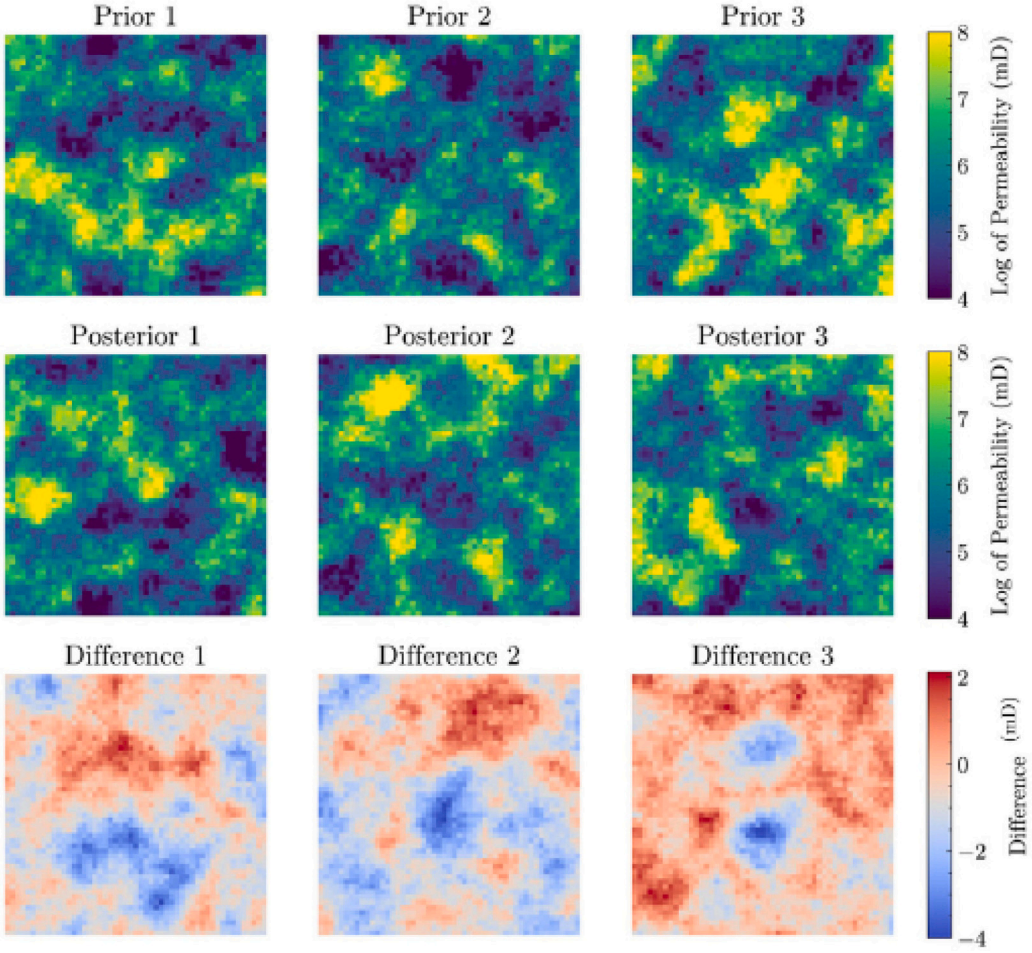


Fig. 17. Prior and posterior permeability estimates, and the difference between these two for the conceptual case. The color scale in the difference plots highlights areas of increased and decreased permeability, corresponding to the permeability adjustments made by ES-MDA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Quantification of the uncertainty reduction in terms of entropy for the temperature distribution before and after data assimilation.

	Entropy (*nats*)	Uncertainty reduction (*nats*)	Relative uncertainty reduction
Prior	78.2	–	–
Posterior	64.7	13.5	17.3%

As described in Section 2.4, we compute the entropy of temperature distribution probabilities to quantify model uncertainty. Table 2 shows the entropy reduction from 78.2 *nats* to 64.7 *nats*, representing a 17.3% reduction in relative uncertainty after data assimilation.

The reduction in entropy demonstrates the data-assimilation process's ability to leverage observed vertical displacement measurements, enhancing model realizations and reducing uncertainty in subsurface modeling for geothermal applications. However, it is crucial to acknowledge the inherent non-uniqueness of permeability fields in geothermal data assimilation. Various permeability distributions can yield similar results, indicating a range of plausible models rather than a single 'true' model. Consequently, directly comparing posterior permeability maps with the reference case is not conclusive. Instead, we focus on the constrained posterior states, particularly the temperature distribution, which is often more critical for geothermal applications than the exact permeability field. Performance indicators include changes in entropy, which quantify the reduction of uncertainty in the key variables of interest, and the variability within the ensemble of model realizations—encompassing both parameter distributions and state responses, such as vertical displacements. A reduction in entropy reflects a more unified understanding of reservoir properties, while the reduction in the variability of vertical displacements demonstrates how effectively the data assimilation process constrains the model's predictions. However, a narrow range of variability in the posterior ensemble should be interpreted cautiously, as it could indicate undesirable ensemble collapse rather than successful data assimilation. Evaluating both entropy reduction and the behavior of variability within the ensemble is crucial for accurately assessing the data-assimilation process's effectiveness in constraining the temperature distribution, which is key for decision-making in geothermal energy projects.

5.1.2. CO_2 geological sequestration scenario

Adapting our model to the CO_2 geological storage scenario necessitated adjustments due to the distinct fluid mobility involved in CO_2 injection. Unlike the water in geothermal systems, in this case study, CO_2 is injected in a supercritical state, exhibiting much lower viscosity and leading to a faster pressure propagation within the reservoir. In this conceptual case, the uniform pressure buildup observed at the end of the simulation across all models results in a nearly uniform uplift. This scenario illustrates a fundamental challenge in demonstrating the value of data-assimilation methods, which rely on the presence of differences between prior realizations and observations. When these differences are minimal, as seen with the uniform uplift in this case, it becomes challenging to showcase the effectiveness of data assimilation. This

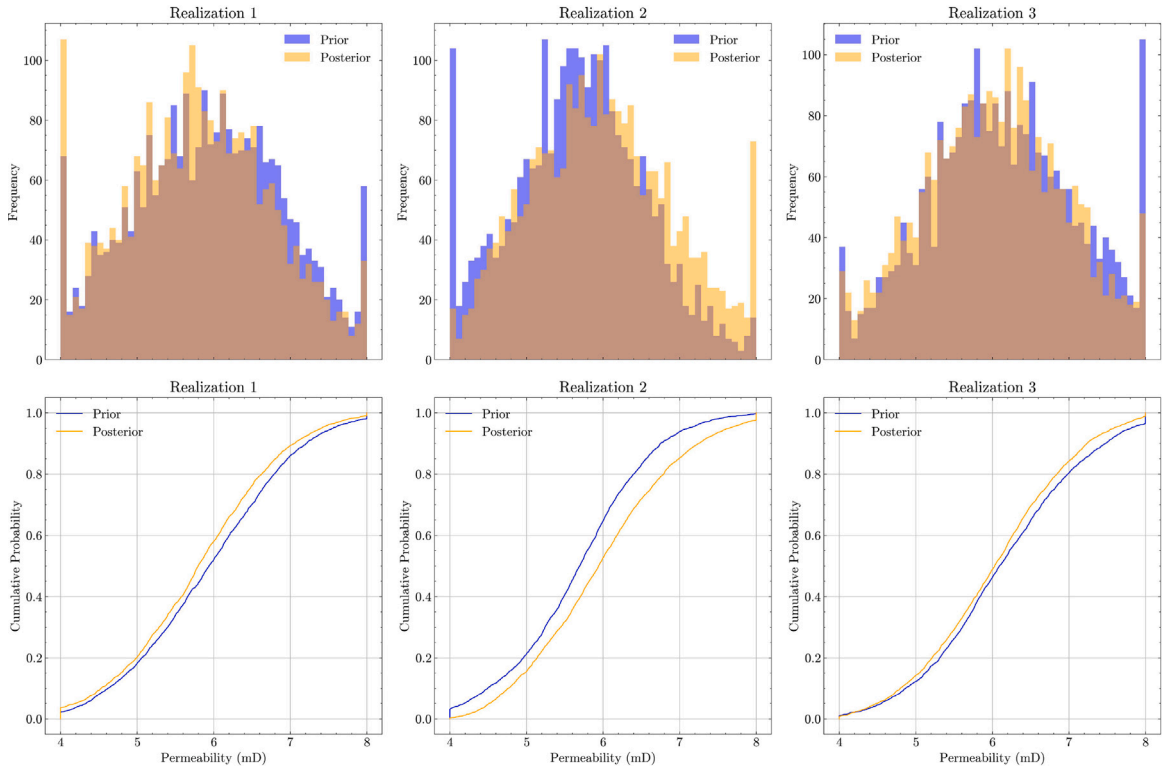


Fig. 18. Histograms and CDFs for the permeability values before (blue) and after (yellow) ES-MDA. The brown area is the overlap region of the distributions.

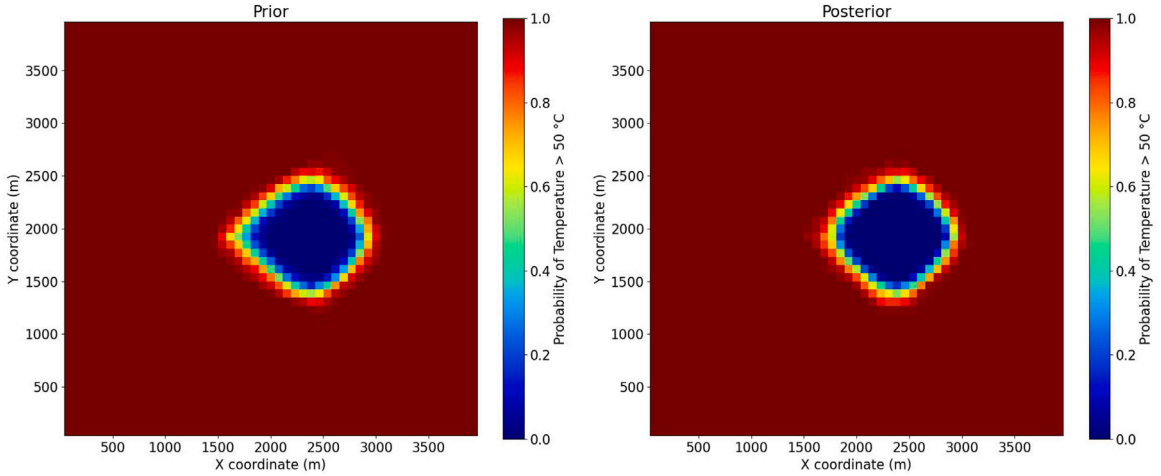


Fig. 19. Probability maps of temperature exceeding 50 °C for the prior (left) and posterior (right) models. The post-data-assimilation map shows a distinct change in the probability distribution.

uniform response does not reflect a limitation of the data-assimilation method itself but highlights the limitations of using this particular case to demonstrate the method's capacity to enhance model accuracy and reduce uncertainties under more variable conditions.

To better illustrate the method's capability, we adjusted the ensemble's permeability distribution, setting the mean log-permeability to 3 mD, with bounds from 1 mD to 8 mD (Fig. 20). This adjustment resulted in more varied pressure propagation and uplift patterns, allowing for a clearer demonstration of the data-assimilation process in CO_2 sequestration. Future studies could consider assimilating vertical displacement measurements at multiple time steps, potentially providing insights into the dynamic interplay between pressure changes and displacement processes during CO_2 injection.

The posterior realizations exhibit uplift curves characteristic of CO_2 injection, with pronounced uplift centered around the injection site (Fig. 21). This pattern reflects the pressure diffusion from the injection well and aligns well with the observed reservoir response. A detailed analysis of permeability changes near the injection site and their implications for model performance is provided in Appendix C, along with histograms and cumulative distribution functions for the model realizations.

To quantitatively evaluate the impact of data assimilation on the probability distribution of CO_2 molar fraction within the reservoir, we constructed probability maps considering a threshold of CO_2 molar fraction equal to 0.1. Fig. 22 shows the change in the probability distribution post-data assimilation, indicating how the observations inform on CO_2 molar fraction.

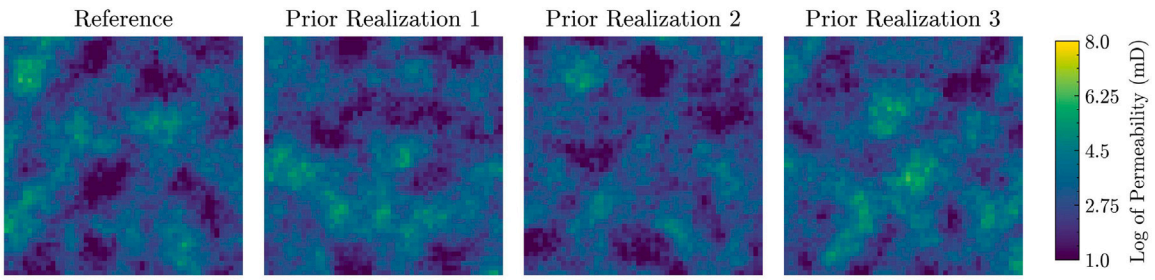


Fig. 20. Reference model realization and three realizations of the prior permeability distribution for the conceptual dataset for the GCS case.

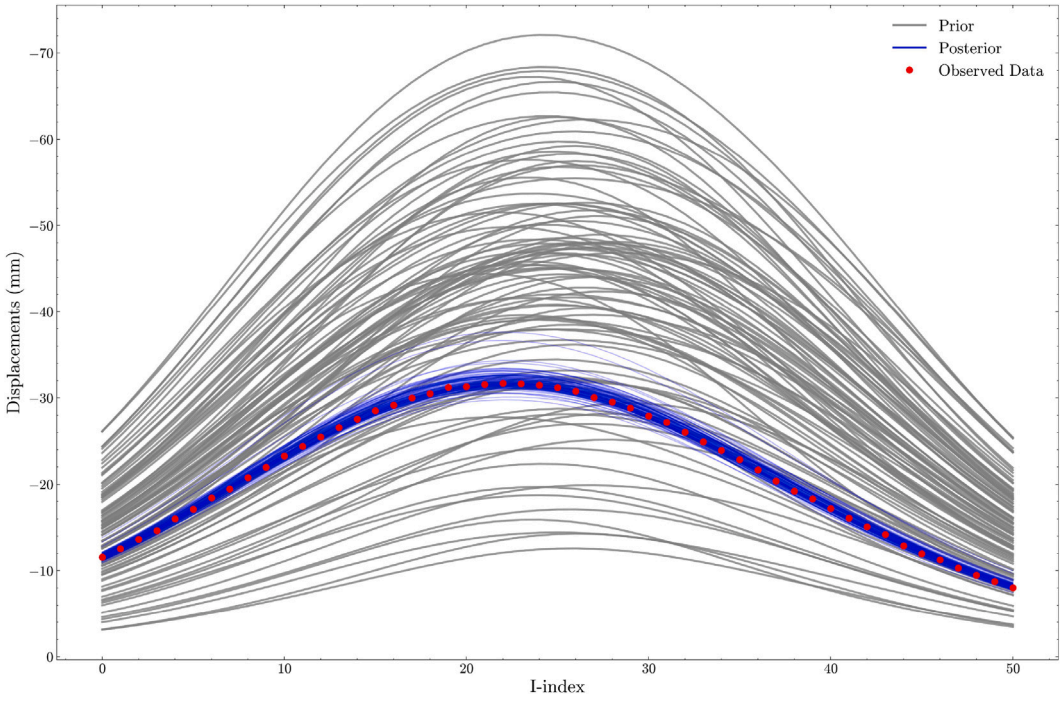


Fig. 21. ES-MDA results for the CO_2 sequestration scenario, illustrating the vertical displacement realizations. The observed data are shown as red dots, with prior realizations in gray and posterior realizations in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

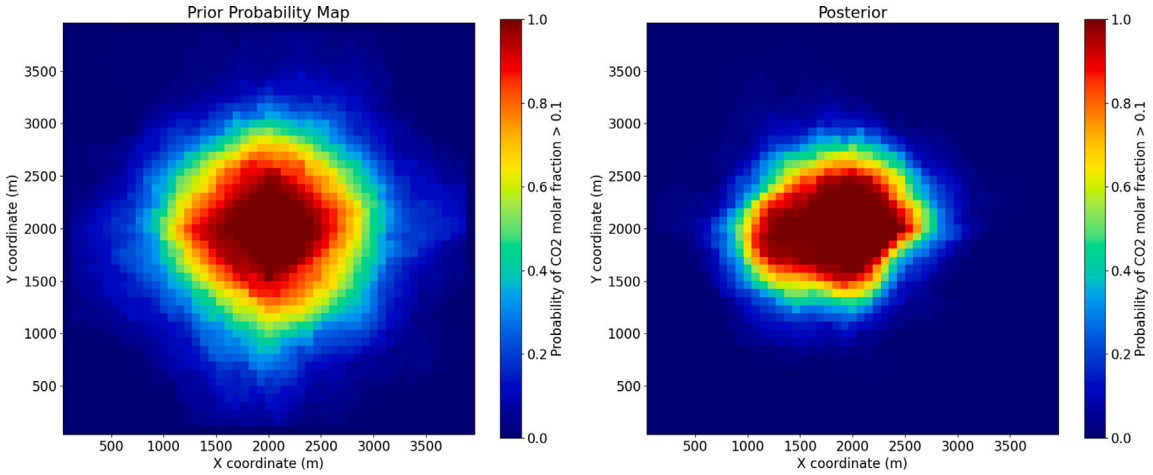


Fig. 22. Probability maps of CO_2 molar fraction exceeding 0.1 for the prior (left) and posterior (right) models. The post-data-assimilation map shows a distinct change in the probability distribution.

Following the approach described in Section 2.4, we quantified the model's uncertainty using the entropy of CO_2 molar fraction distribution probabilities. Table 3 shows the entropy reduction from 725.2 *nats* to 340.1 *nats*, representing a 53.1% decrease in relative uncertainty after data assimilation. This substantial reduction is particularly significant because it demonstrates the ability to constrain the CO_2

Table 3
Quantification of the uncertainty reduction in terms of entropy for the CO_2 molar fraction distribution before and after data assimilation.

	Entropy (*nats*)	Uncertainty reduction (*nats*)	Relative uncertainty reduction
Prior	725.2	–	–
Posterior	340.1	385.1	53.1%

plume prediction, even in the face of non-unique permeability distributions. This improvement in predictive capability is important for risk assessment and decision-making in CO_2 sequestration projects, where accurate plume tracking is more critical than determining the exact permeability field.

In the GCS conceptual experiment, we can observe larger displacements across the ensemble compared to those in the geothermal case, as there is no production from the reservoir. However, the evaluations across both scenarios evaluated in this study, geothermal energy extraction and CO_2 geological sequestration, highlight the efficacy of applying the combination of flow simulation with a physics-based proxy model and ES-MDA for a better estimation of reservoir permeability and dynamic behavior. In particular, using surface vertical displacements has been shown to have potential for the evaluation of the permeability distribution in both cases.

6. Data assimilation for the Brugge dataset

The Brugge dataset presents a more complex challenge than the conceptual dataset due to its geological features.[25] In this section, we apply the established data assimilation framework, consisting of ES-MDA, open-DARTS and the physics-based proxy model, to understand the reservoir's permeability and dynamic responses within this more realistic context when assimilating observations of vertical displacement.

We apply the same basic approach used in the previous cases to prepare the ensemble of permeability fields. However, we adjusted the features of our Gaussian distribution of permeability to match the specific data collected from past studies of the Brugge dataset. This means setting the average permeability at 540 mD and allowing for variation within the observed range in the field, as can be seen in Fig. 23. The prior distribution created for the Brugge dataset is designed to be a realistic starting point for our simulations. Similarly to the conceptual case, we create an ensemble of 101 permeability fields. Out of this ensemble, 100 realizations are designated as the prior ensemble for data assimilation. The remaining realization serves as the reference realization.

After setting up the prior ensemble of permeability fields, we proceed with the application of the ES-MDA framework for both geothermal energy extraction and CO_2 geological sequestration scenarios. We utilize the vertical displacement measurements along a transect intersecting the well trajectories (Fig. 5) as the observational data to be integrated into the data-assimilation process.

The results of the ES-MDA application for both scenarios are presented in Fig. 24. For the geothermal energy extraction case, we observe a complex response with a combination of uplift and subsidence across the surface, reflecting the intricate geology of the Brugge dataset. In the CO_2 sequestration scenario, we observe a pronounced uplift primarily concentrated around the injection site, consistent with the behavior observed in the 2D conceptual case.

To quantitatively evaluate the impact of data assimilation, we constructed probability maps for both scenarios. For the geothermal case, we considered temperatures exceeding 50 °C, while for the CO_2 sequestration case, we used a threshold of 0.1 CO_2 molar fraction. The entropy of these probability distributions was computed to quantify the model's uncertainty before and after data assimilation. The results are summarized in Table 4.

Table 4
Entropy reduction for the Brugge dataset scenarios, indicating enhanced model reliability and predictive accuracy.

Scenario		Entropy (*nats*)	Uncertainty reduction (*nats*)	Relative uncertainty reduction
Geothermal	Prior	29.3	–	–
	Posterior	27.1	2.2	7.5%
CO_2 sequestration	Prior	354.0	–	–
	Posterior	288.5	65.1	18.5%

For the geothermal case, the entropy reduction from 29.3 nats to 27.1 nats indicates a slight improvement in the model's reliability, corresponding to a relative uncertainty reduction of 7.5%. This modest improvement highlights the challenge of reducing uncertainties in complex fields like Brugge, especially for geothermal applications where the interplay between pressure and thermal effects results in smaller displacements. In contrast, the CO_2 sequestration scenario showed a more substantial improvement. The entropy reduction from 354.0 nats to 288.5 nats signifies an 18.5% reduction in relative uncertainty post-data assimilation. This larger reduction can be attributed to the more pronounced uplift patterns and larger vertical displacements associated with CO_2 injection, providing more informative data for the assimilation process.

The application of the ES-MDA framework to the more complex Brugge dataset highlights the challenges associated with modeling real-world reservoirs. While the fundamental approach remains similar to the 2D conceptual case, the increased complexity of the Brugge dataset impacts the performance of the data-assimilation method. The CO_2 sequestration scenario benefited more from the data assimilation process compared to the geothermal case, likely due to the larger and more distinct displacement patterns associated with CO_2 injection.

These results demonstrate the potential of integrating surface displacement data through the use of a physics-based proxy model and the ES-MDA technique to enhance our understanding of subsurface properties and dynamic behavior. The method shows promise in reducing uncertainties and improving model predictions, particularly for CO_2 sequestration applications. However, the more modest improvements in the geothermal case suggest that additional data types or more sophisticated assimilation techniques may be necessary to fully constrain complex geothermal systems. Nonetheless, these findings support the value of this approach in improving decision-making processes for energy transition applications, even in challenging, realistic reservoir scenarios.

A detailed analysis of the changes in permeability fields, including histograms and cumulative distribution functions for both the geothermal and CO_2 sequestration scenarios, is presented in Appendix C.2. This comprehensive examination provides further insights into the effects of data assimilation on the Brugge dataset models.

Furthermore, to validate the data assimilation results, we examined prior as posterior displacements at locations outside the original assimilation area and compared them to the reference model. These results, presented in Appendix C.3, confirm the improved model accuracy even in non-assimilated regions, as posterior predictions better align with independent reference data points. The detailed validation plots for each case are shown in Fig. C.35.

7. Discussion

In this study, we integrate vertical displacement observations with open-DARTS reservoir simulations and a physics-based geomechanical proxy model to enhance subsurface model predictions. While different permeability fields may generate similar vertical displacement responses, leading to non-unique permeability estimates, our ES-MDA

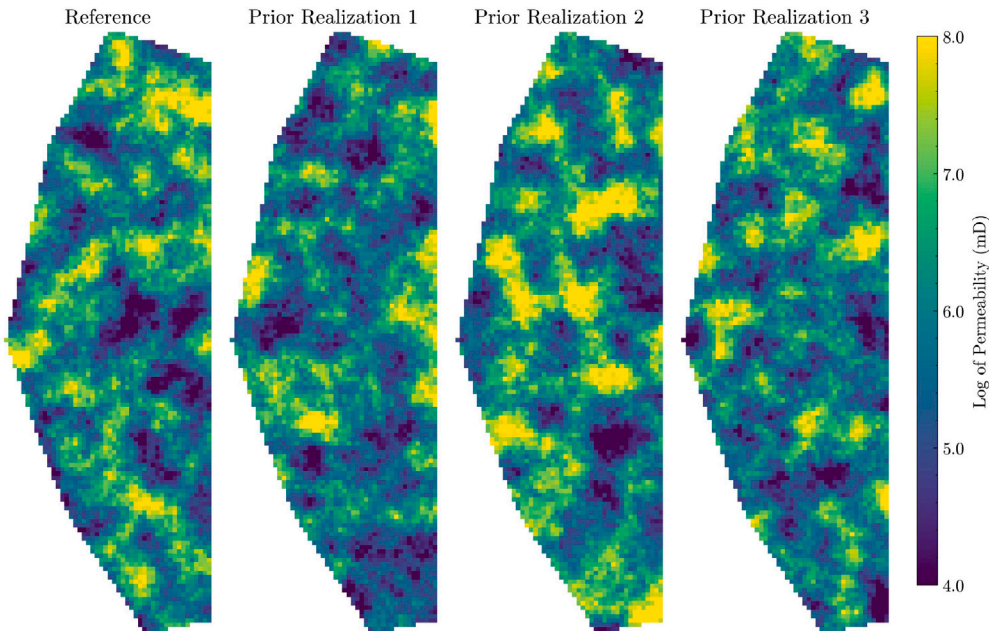
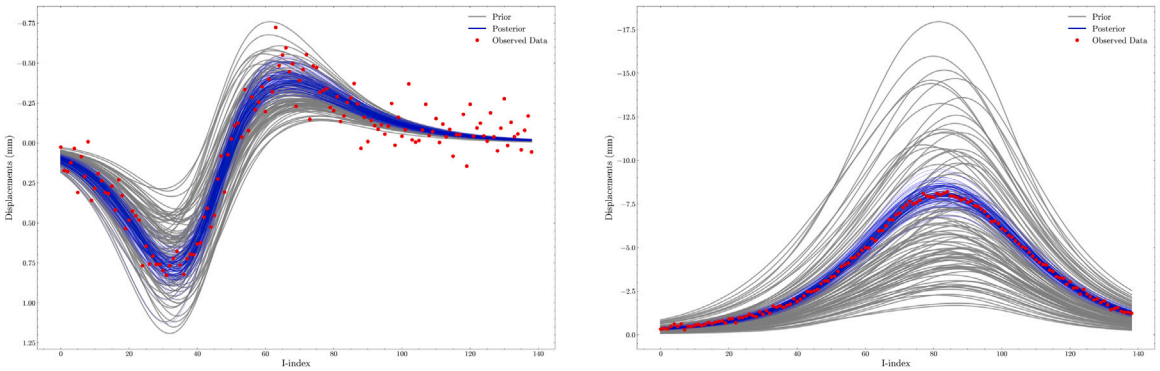


Fig. 23. Reference realization and three realizations of the prior permeability distribution for the Brugge dataset.



(a) Geothermal energy extraction (b) CO_2 geological storage

Fig. 24. Post-ES-MDA comparison of realizations and observations for the Brugge dataset scenarios. The observed data are shown as red dots, with prior realizations in gray and posterior realizations in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method effectively estimates likely permeability fields and their probability distributions. More importantly, this approach successfully constrains the posterior states of the reservoir, reducing uncertainty in critical variables such as temperature distribution and CO_2 plume extent, even when the exact permeability field cannot be uniquely determined.

The integration of our geomechanical proxy model with ES-MDA enhances subsurface modeling for geothermal energy extraction and CO_2 sequestration by leveraging often underutilized surface displacement measurements. This approach improves interpretation of complex subsurface behaviors in response to operational activities. However, we acknowledge that the assumption of homogeneous geomechanical properties may not fully reflect real-world formation heterogeneity, potentially affecting the accuracy of displacement and stress distributions.

Incorporating 2D surface displacement measurements and using displacement data from multiple time steps could significantly enhance the accuracy of subsurface models by providing more comprehensive temporal and spatial constraints. While our current study focused on 1D displacement data at the final time step to demonstrate the effectiveness of our methodology, future research could explore these extensions as part of a broader sensitivity analysis. Such analyses – including the impact of data sparsity, topographical density, and the timing of data collection – are valuable but beyond the scope of the present work. Addressing these aspects in future studies would help optimize data collection strategies and further validate the robustness of our approach in various geological settings.

Our study established a Data Assimilation framework, which has several opportunities for future research. These include comparing the effectiveness of various data types, quantifying uncertainty in geomechanical parameters, conducting more extensive validation of the proxy model, and exploring other data assimilation techniques beyond ES-MDA. Future studies should also focus on incorporating heterogeneous geomechanical properties, extending the framework to different geological settings, and integrating diverse data types such as seismic data to further validate its robustness and expand its applicability.

For geothermal projects, the interplay of pressure and thermal effects results in smaller displacements compared to GCS projects, posing a challenge when using vertical displacements as observations. Future research should address the integration of heterogeneous geomechanical properties and optimize data assimilation locations and timing to enhance subsurface modeling accuracy. This is crucial for advancing

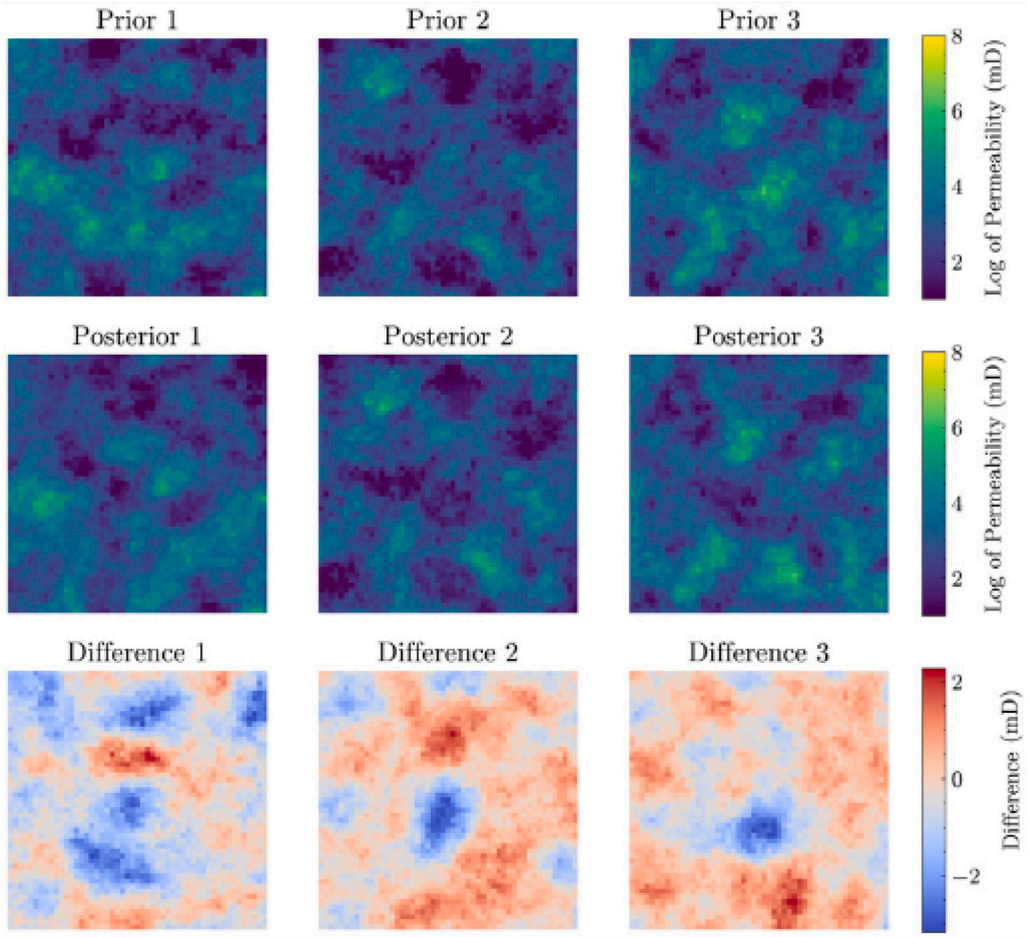


Fig. C.25. Comparison of prior, posterior, and the difference in permeability fields for the CO_2 sequestration case. The difference plots emphasize the localization and magnitude of permeability adjustments around the injection site made by the ES-MDA.

such models in energy transition initiatives, ensuring more accurate representation and management of both geological and operational complexities.

It is important to note that the use of vertical displacement data for assimilation in geothermal and CO_2 sequestration applications is not common practice. Traditional approaches often rely on well data or seismic surveys, overlooking the potential of surface deformation measurements. Our study demonstrates the value of incorporating these often-underutilized observations, offering a novel approach to constrain subsurface models and reduce uncertainties in key parameters such as permeability distributions and plume extents. This innovative use of vertical displacement data opens new avenues for improving subsurface characterization in energy transition projects.

8. Conclusion

This study presents the development and application of a physics-based geomechanical proxy model for approximate thermoporoelasticity solutions, designed to integrate vertical displacement observations within a data assimilation framework. Our approach employs ES-MDA to quantify uncertainties in thermal front and CO_2 plume distribution for geothermal and Geological Carbon Sequestration (GCS) projects, respectively. The model's flexibility allows for runtime application or post-processing in hydrodynamic simulations, with displacement evaluations at desired time steps and compatibility with various simulators. Its adaptability to non-rectangular grids enhances its suitability for realistic geological settings. The integration with open-DARTS' OBL caching feature significantly improved computational efficiency, demonstrating the potential for more complex and extensive simulations in geothermal and CO_2 sequestration applications.

Our ES-MDA implementation demonstrated efficacy in uncertainty quantification across four case studies. In both conceptual models (geothermal and GCS), the method yielded satisfactory posterior ensembles with reduced uncertainty. The geothermal case showed a 17.3% decrease in temperature plume distribution entropy, while the GCS scenario achieved a 53.1% reduction in CO_2 plume uncertainty, benefiting from larger vertical displacements and ensemble variance. The application to the Brugge dataset further validated the method's effectiveness in complex geological settings. For geothermal extraction, entropy reduction was 7.5%, while the GCS case showed an 18.5% reduction in CO_2 molar fraction uncertainty. These results highlight the method's robustness in handling realistic reservoir complexities.

The comparative analysis between geothermal and GCS applications underscores the versatility of our approach. While the degree of improvement varied based on fluid properties and injection dynamics, all cases benefited from ES-MDA application. Notably, our study pioneers the use of vertical displacement data for assimilation in these applications, a practice not commonly employed in the field. This innovative approach demonstrates the potential of integrating geomechanical proxy models with ES-MDA to enhance subsurface modeling accuracy and reliability, leveraging often overlooked surface deformation data. By doing so, we offer a valuable tool for advancing energy transition projects through improved reservoir characterization and uncertainty quantification, potentially changing how the industry approaches data assimilation in geothermal and CO_2 sequestration projects.

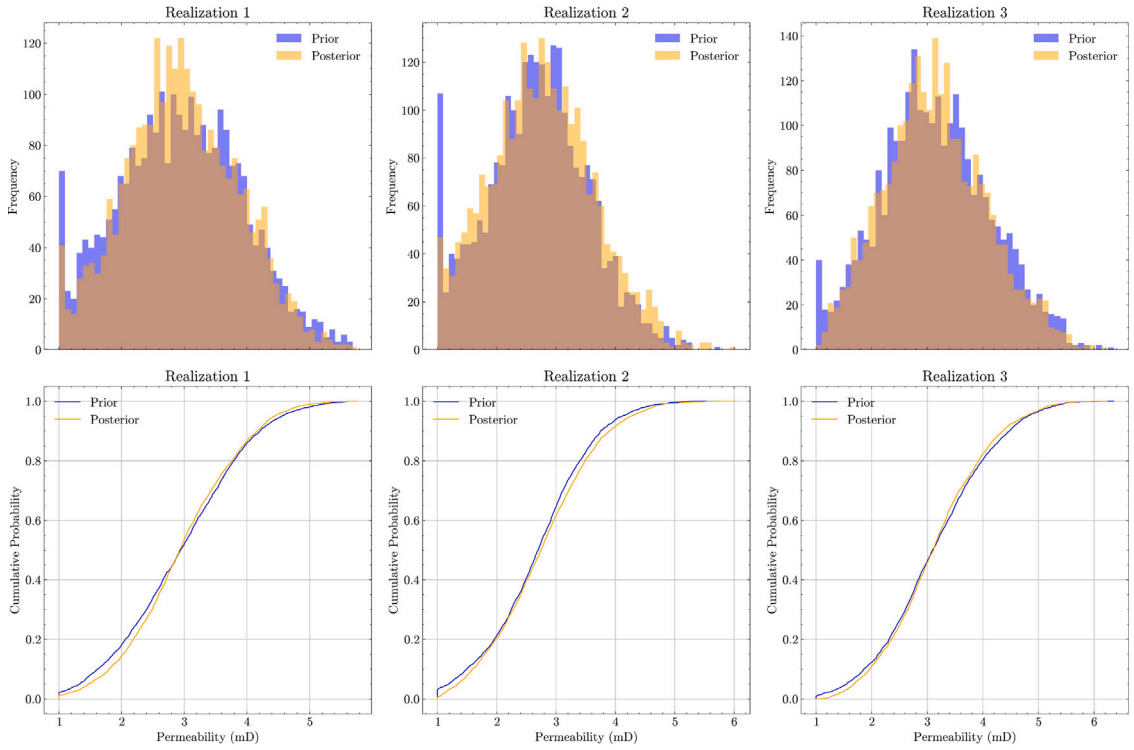


Fig. C.26. Histograms and cumulative distribution functions displaying the permeability values before (blue) and after (yellow) the use of ES-MDA for the GCS case. The brown area is the overlap region of the distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Ilshat Saifullin: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Gabriel Serrão Seabra:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Anne Pluymakers:** Writing – review & editing, Supervision, Project administration, Funding acquisition. **Femke C. Vossepoel:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Methodology, Conceptualization. **Denis Voskov:** Writing – review & editing, Supervision, Software, Methodology, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

In the course of preparing this manuscript, the authors utilized ChatGPT from OpenAI to improve readability. Additionally, Grammarly was employed to guarantee grammatical accuracy and to improve overall readability. Subsequent to employing these tools, the authors meticulously reviewed and amended the content as necessary. The authors bear full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Gabriel Serrao Seabra reports financial support was provided by Petrobras. Gabriel Serrao Seabra reports a relationship with Petrobras that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank NWO, Netherlands for sponsoring the doctoral research of Ilshat Saifullin (grant number 18816) and Petróleo Brasileiro S.A. (Petrobras), Brazil for sponsoring the doctoral research of Gabriel Serrão Seabra.

We would like to thank Aleks Novikov for the help in geomechanical part of this study.

Appendix A. Governing equations for the hydrodynamic model

Our hydrodynamic model is based on the open-DARTS formulation[27,43] which includes the mass and energy conservation equations for the system with n_c components and n_p phases:

$$\frac{\partial}{\partial t}\left(\phi \sum_{j=1}^{n_p} x_{cj}\rho_j s_j\right) + div \sum_{j=1}^{n_p} x_{cj}\rho_j \mathbf{u_j} + \sum_{j=1}^{n_p} x_{cj} q_j = 0, \quad c = 1, \ldots, n_c, \quad \text{(A.1)}$$

where ϕ is the porosity, x_{cj} is the molar fraction of component c in phase j, s_j is the saturation of phase j, and ρ_j is phase density [kmol/m^3].

$$\frac{\partial}{\partial t}\left(\phi \sum_{j=1}^{n_p} s_j\rho_j U_j + (1-\phi)U_r\right) + div \sum_{j=1}^{n_p} h_j\rho_j \mathbf{u_j} + div(\kappa \nabla T) + \sum_{j=1}^{n_p} h_j\rho_j q_j = 0 \quad \text{(A.2)}$$

where U is the specific fluid internal energy [kJ/kmol], U_r is the rock internal energy in [kJ/m^3], T is the temperature [K], h_j is the phase enthalpy [kJ/kmol], and κ_j is total thermal conductivity for the solid rock and the phase j [kJ/m/day/K]. The phase j velocity $\boldsymbol{u_j}$ evaluation is based on the Darcy's law:

$$\boldsymbol{\nu_j} = -\mathbf{K}\frac{k_{rj}}{\mu_j}(\nabla p_j - \gamma_j \nabla z), \quad \text{(A.3)}$$

where $\mathbf{K}$ is the permeability tensor [mD], k_{rj} is the relative permeability of phase j, μ_j is the viscosity of phase j [mPa · s], p_j is the pressure of

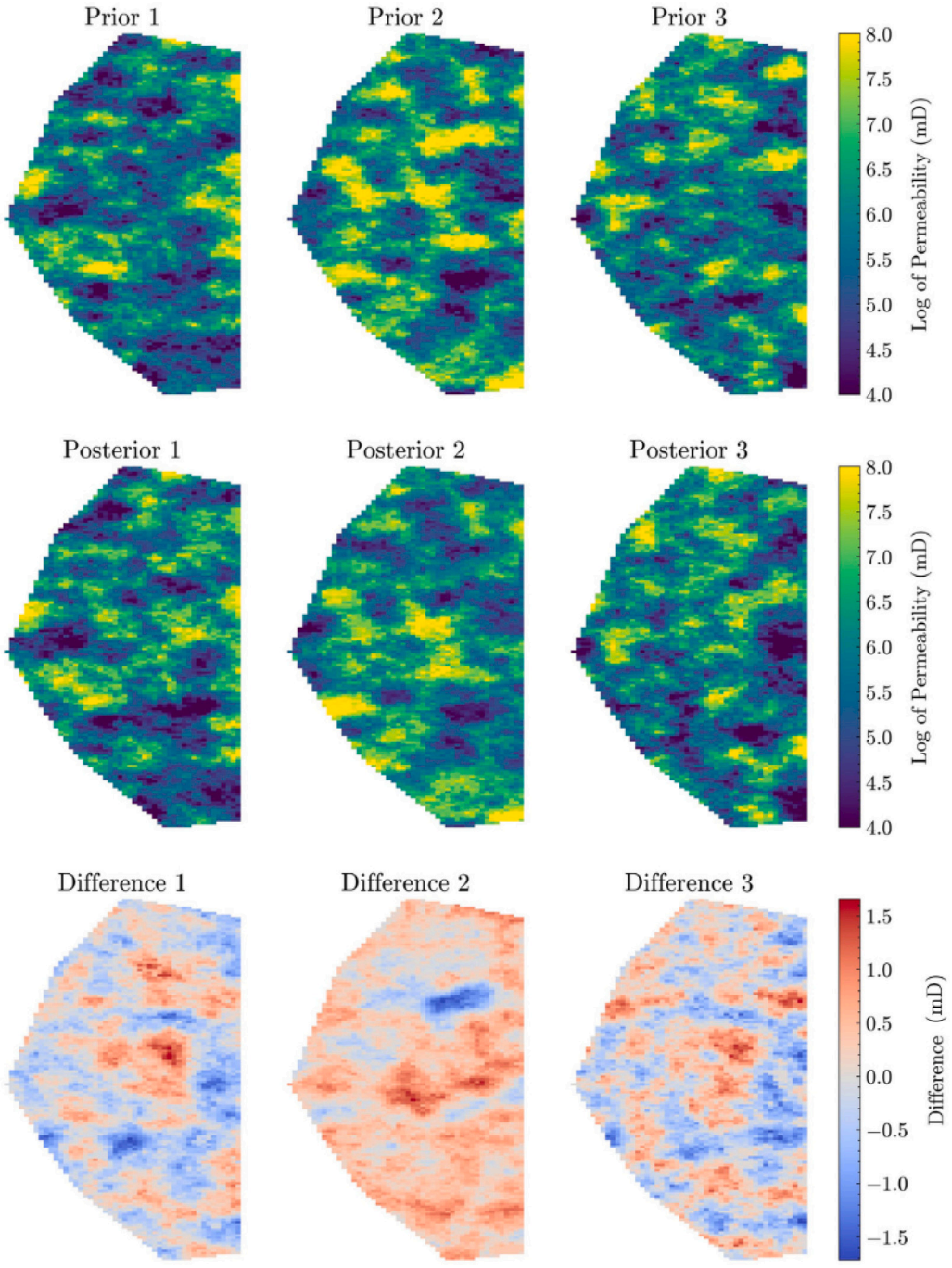


Fig. C.27. Comparison of prior, posterior, and the difference in permeability fields for the Brugge dataset geothermal energy extraction case. The difference plots illustrate the spatially heterogeneous adjustments made to the permeability field by the ES-MDA.

phase j [bars], $\gamma_j = \rho_j g$ is the specific weight of phase j [N/m^3] and z is depth [m], g is the gravitational acceleration [m s^{-2}].

The porosity is updated using

$$\phi = \phi_0 + c_r(p - p_0) \tag{A.4}$$

where ϕ_0 is the porosity at the reference pressure p_0 [bars] and c_r is the rock compressibility [1/bars].

The thermo-compositional model used in the CO_2 injection case, involves n_c components, liquid and vapor phases. In a molar formulation, the nonlinear unknowns are the pressure, the temperature and compositions: $p, T, x_{c_1}, x_{c_2}, \ldots, x_{c_{n_c}}$. The geothermal model involves a water component, liquid and vapor phases, and the pressure and the phase enthalpy as unknowns: p, h, with the state-dependent water properties evaluated using the IAPWS97 formulation.[44]

Appendix B. Solution method for the hydrodynamic model

In order to numerically solve the system of PDEs (A.1) and (A.2) we apply the Finite Volume Method for the discretization in space and backward Euler discretization in time, leading to the Fully Implicit Method. The discretized PDEs result to a system of non-linear equation, solved using the conventional Newton–Raphson method. Our linearization approach involves the Jacobian assembly with analytic evaluation of fluid and rock properties and their derivatives with respect to the unknowns. The linear system with a Jacobian matrix is solved using an

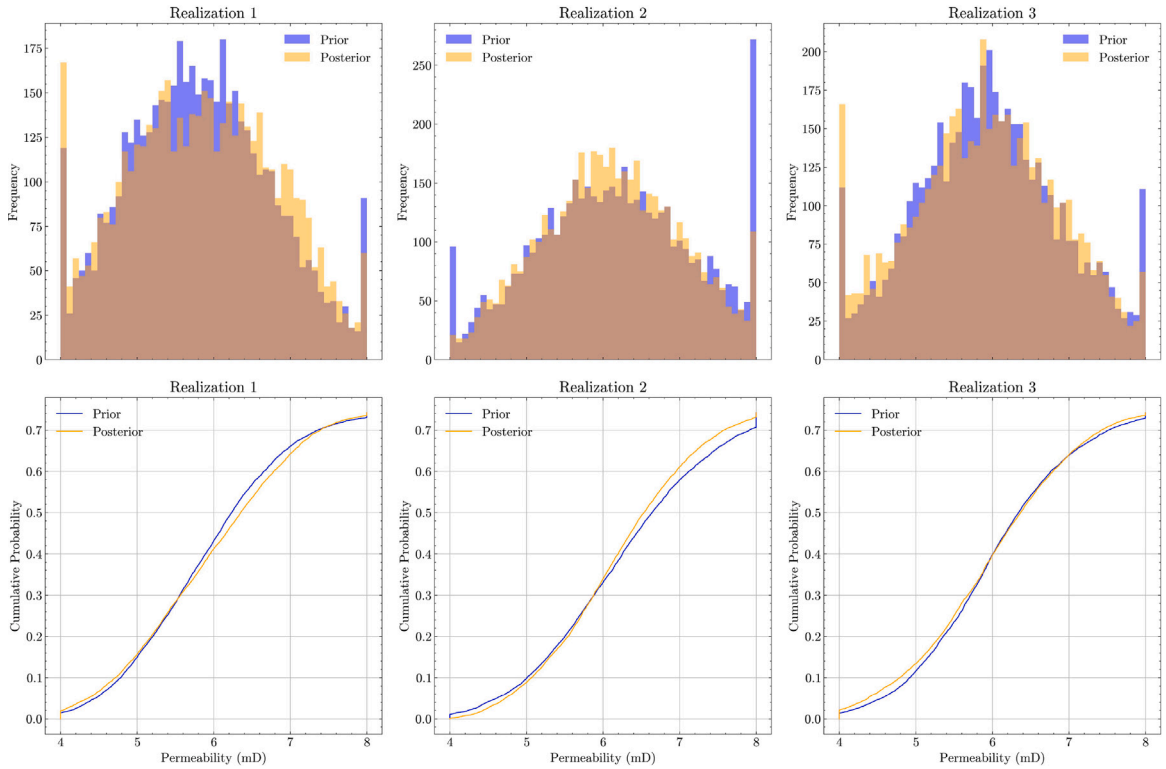


Fig. C.28. Histograms and CDFs displaying the permeability values before (blue) and after (yellow) the ES-MDA for the Brugge dataset geothermal energy extraction case.The brown area is the overlap region of the distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

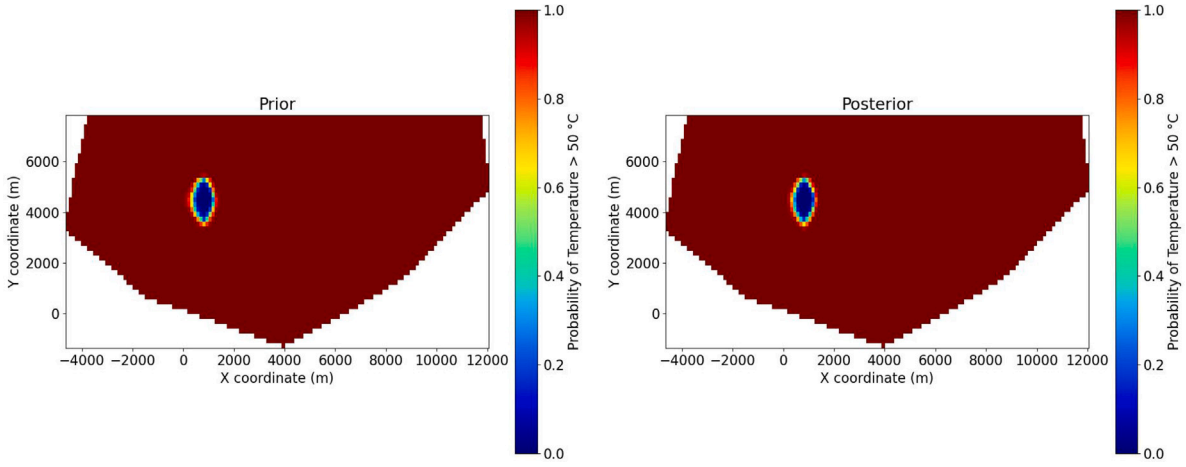


Fig. C.29. Probability maps of temperature exceeding 50 °C for the Brugge dataset prior (left) and posterior (right) realizations for geothermal energy extraction scenario.

iterative method GMRES with CPR preconditioner at each iteration of the Newton–Raphson method. A part of the solution for multicomponent model (CO_2 injection in our case) is the multiphase flash which resolves mole fractions for each component and phase fractions for the given composition. The detailed description of discretization and linearization can be found in Refs. 26, 27, 45. The convergence parameters for the geothermal model used are: tolerance for the Newton method = 1e−3, tolerance for the linear solver = 1e−3, maximum allowed time step size = 30 days. The convergence parameters for the CO_2 injection model used are: tolerance for the Newton method = 1e−3, tolerance for the linear solver = 1e−5, maximum allowed time step size = 5 days.

Appendix C. Complementary analyses of data assimilation results

C.1. Detailed analyses of the 2D conceptual model

This appendix provides complementary analyses of the main discussions presented in the paper. The purpose of these analyses is to present a view of how data assimilation influences the reservoir model's dynamics and to substantiate the core findings with deeper statistical analyses.

For the 2D conceptual model of CO_2 injection, there are impacts of data assimilation on areas experiencing substantial pore pressure

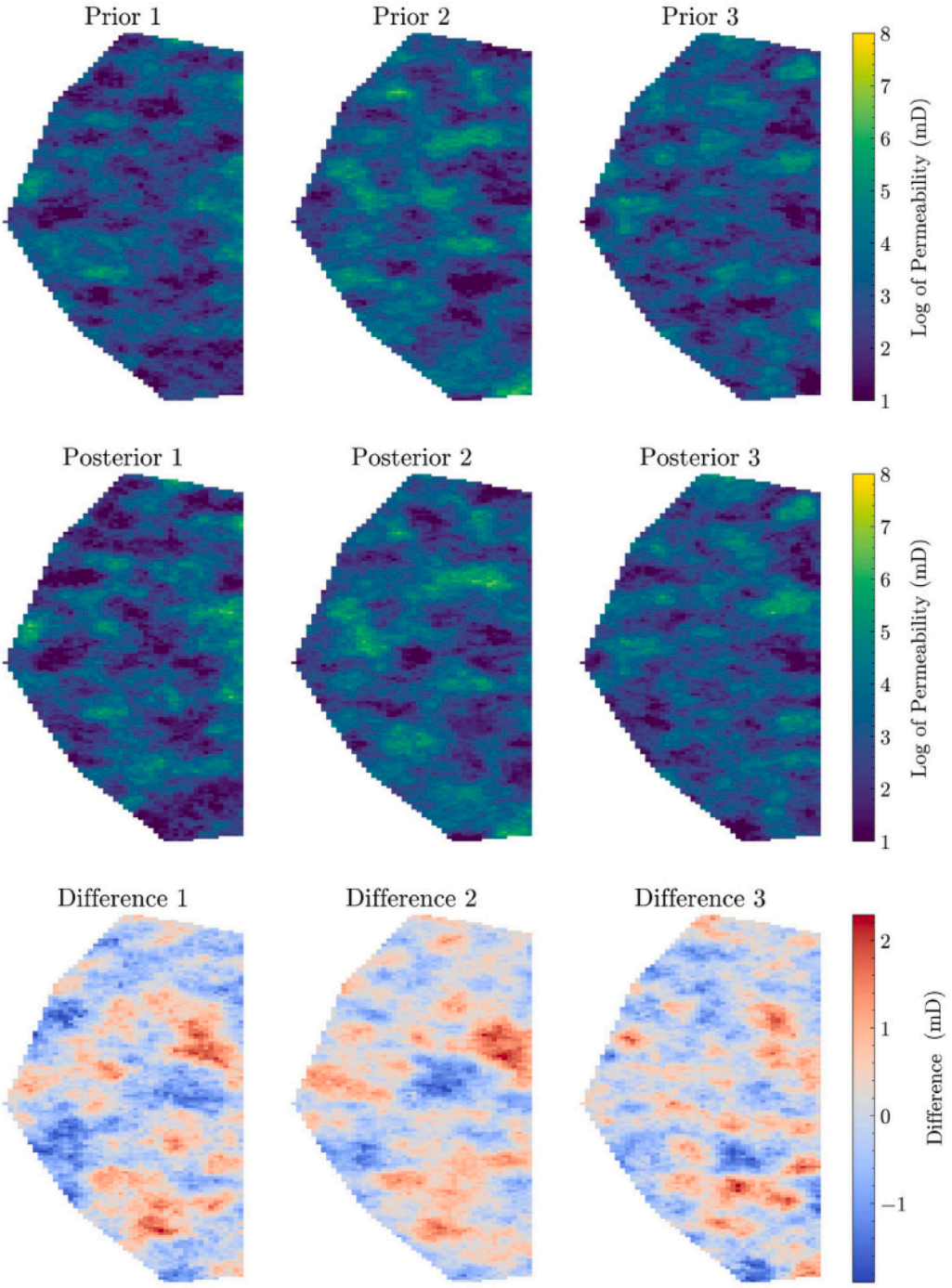


Fig. C.30. Comparison of prior, posterior, and the difference in permeability fields for the Brugge dataset CO_2 geological sequestration case. The difference plots showcase the localized adjustments made to the permeability field around the injection site.

changes. The following difference plots and statistical distributions provide an examination of how data assimilation modifies the permeability fields around the CO_2 injection site, leading to observable surface displacements and alterations in the reservoir's dynamic behavior. The difference plots show the shifts in permeability close to the CO_2 injection site (Fig. C.25), indicating that data assimilation strongly impacts areas experiencing substantial pore pressure changes, which can be due to the more pronounced pressure variations in this area, leading also larger displacements at the surface in the vertical projection of the injector well. Reviewing the histograms and CDFs across the three realizations, the posterior distributions show a striking similarity to the prior distributions.Fig. C.26 illustrates the impact on the overall permeability distribution of three realizations.

C.2. Detailed analyses of Brugge dataset results

Following the detailed exploration of the 2D conceptual model's response to data assimilation, we now show some complementary results for the Brugge dataset analyses.

C.2.1. Geothermal energy extraction

For Geothermal Energy Extraction in the Brugge dataset, the difference plots (Fig. C.27) reveal complex and spatially heterogeneous

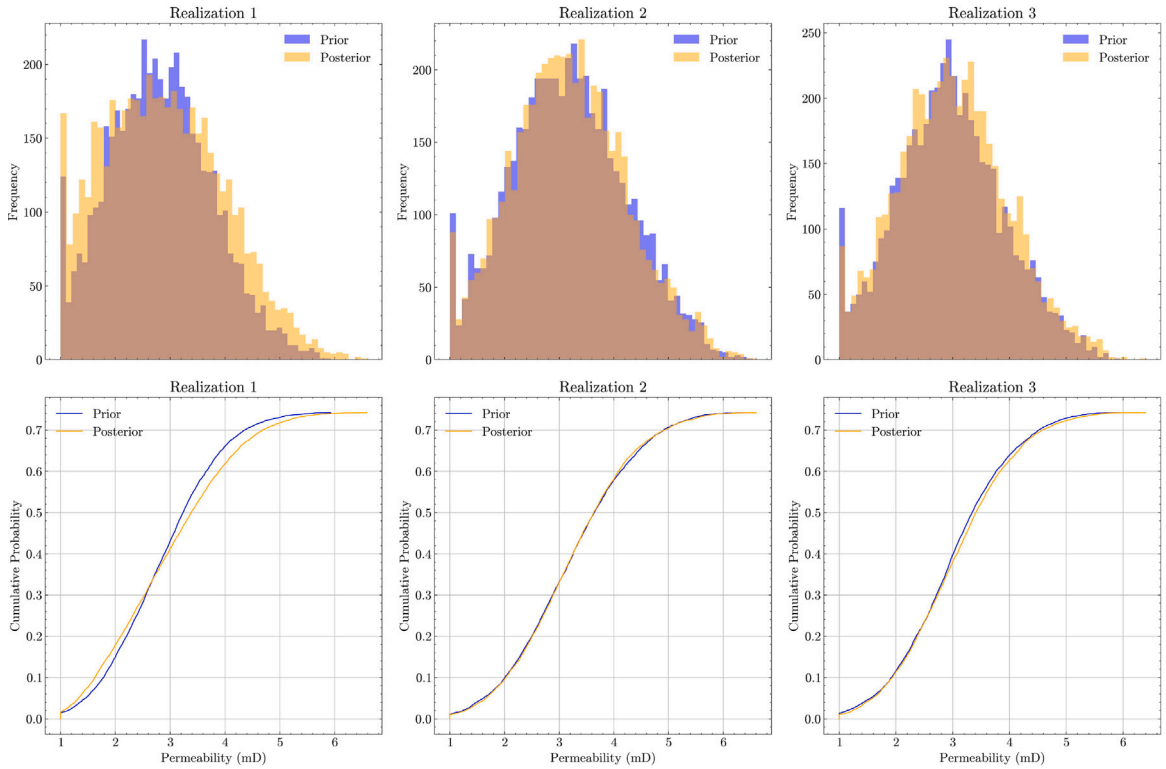


Fig. C.31. Histograms and cumulative distribution functions displaying the permeability values before (blue) and after (yellow) the ES-MDA for the Brugge dataset CO_2 geological sequestration case. The brown area is the overlap region of the distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

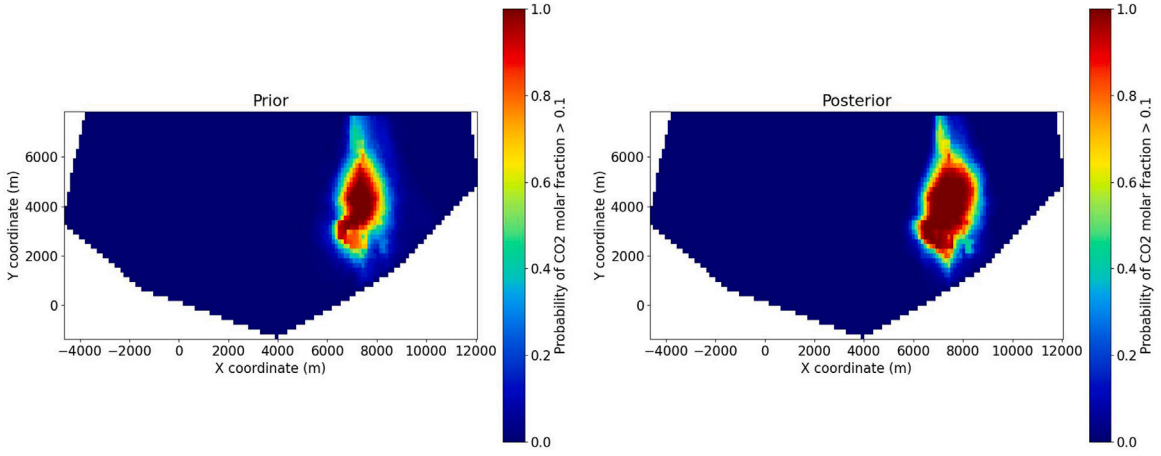


Fig. C.32. Probability maps of CO_2 molar fraction exceeding 0.1 for the prior (left) and posterior (right) models. The post-data-assimilation map shows a distinct change in the probability distribution.

adjustments to the permeability field. The changes are not limited to specific regions but are distributed across the model domain, reflecting the intricate effect of geology on the extraction process. This highlights the importance of using a detailed, realistic geological model to capture the nuances of subsurface behavior, especially when dealing with complex reservoirs.

The histograms and CDFs (Fig. C.28) further demonstrate the impact of ES-MDA on the permeability distribution. In contrast to the conceptual case, the changes are more subtle, with the posterior distributions exhibiting a slightly narrower spread compared to the prior. This suggests that while the ES-MDA refines the permeability model to match better the observed vertical displacements, the inherent complexity of the Brugge dataset limits the degree of adjustment possible with this implementation of ES-MDA.

The probability maps for temperatures exceeding 50 °C in the Brugge dataset (Fig. C.29) illustrate the impact of data assimilation on the temperature distribution. Due to the significantly larger scale of the field compared to the conceptual dataset, the cold front propagation is less pronounced, resulting in minimal observable differences in the probability maps pre- and post-assimilation.

C.2.2. CO_2 geological sequestration

For CO_2 Geological Sequestration in the Brugge dataset, the difference plots (Fig. C.30) highlight the most significant adjustments to permeability occurring in the vicinity of the injection site, similar to

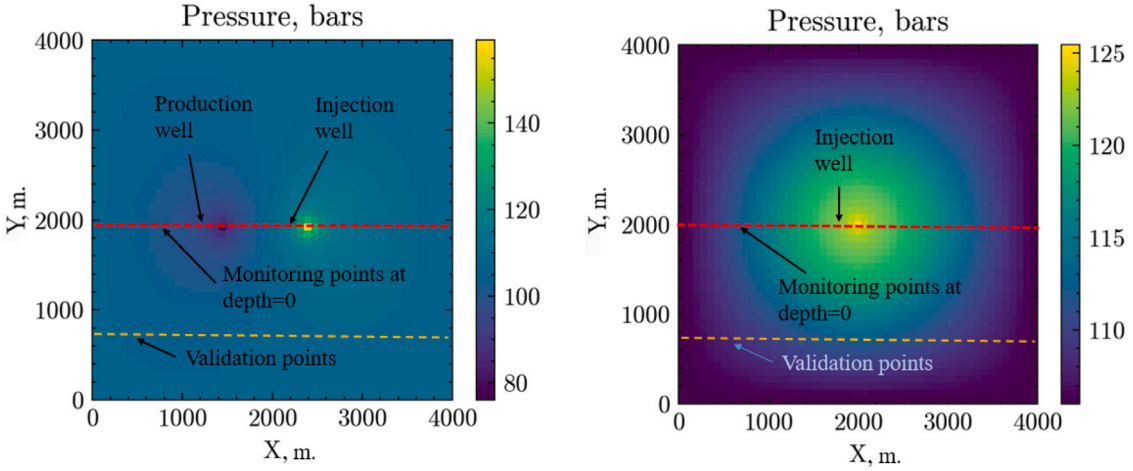
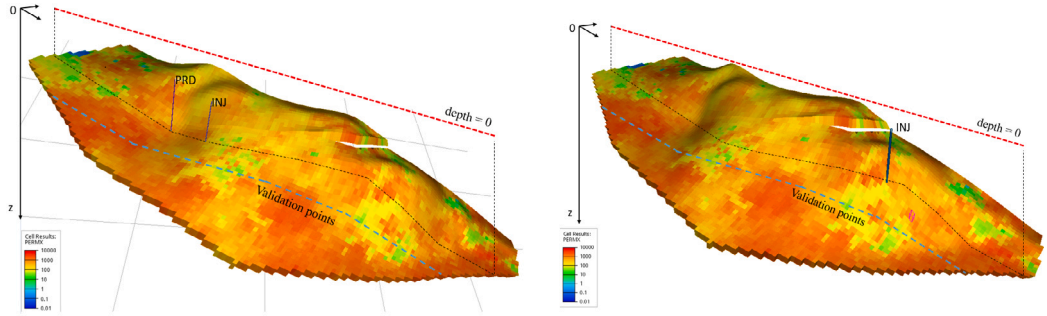
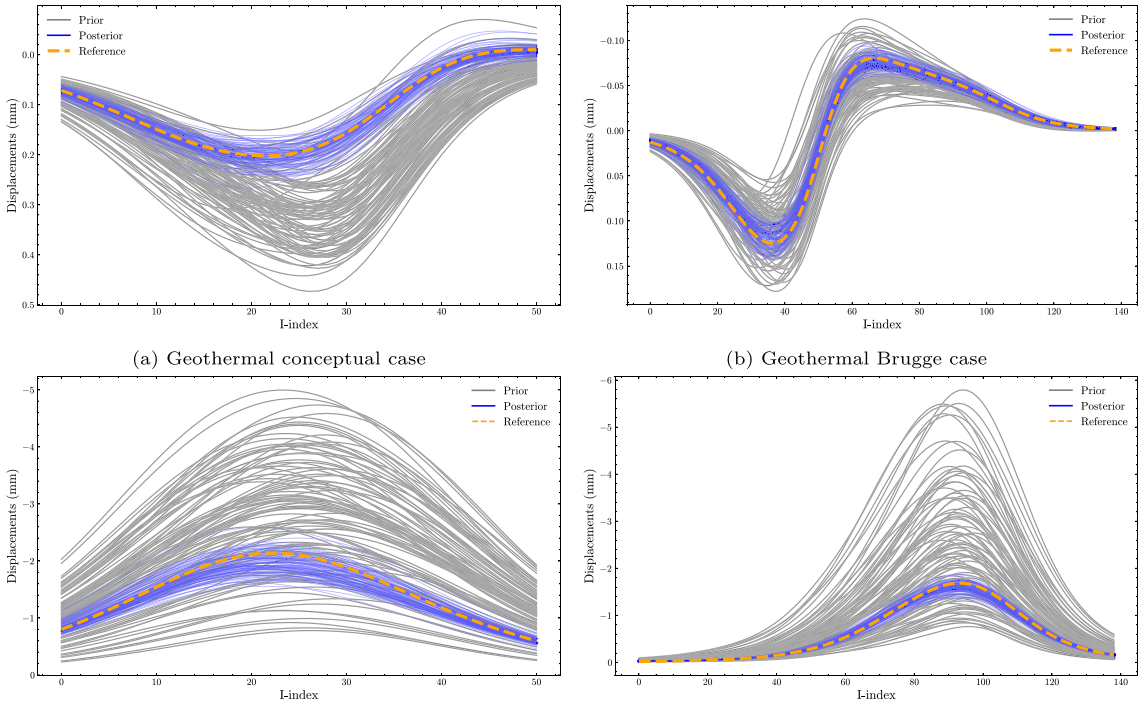


Fig. C.33. 2D Conceptual Model setup showing the grid layout (top view), well positions, and monitoring points for the geothermal extraction case (left) and GCS (right).



(a) Geothermal case (b) GCS case

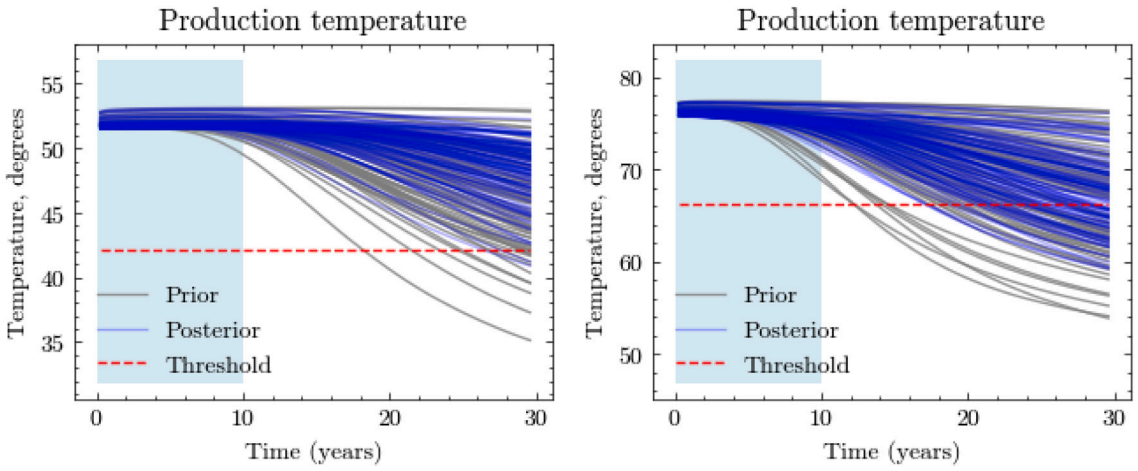
Fig. C.34. Brugge dataset model setup illustrating CPG grid, well locations, and monitoring points along J-index=31.



(a) Geothermal conceptual case (b) Geothermal Brugge case

(c) CO_2 sequestration conceptual case (d) CO_2 sequestration Brugge case

Fig. C.35. Validation results: Vertical displacement along lines outside data assimilation areas. Posterior predictions (blue) show improved agreement with reference data (orange) compared to prior predictions (gray). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) Geothermal conceptual case (b) Geothermal Brugge case

Fig. C.36. Forecasting showcase of ES-MDA on the well production temperature.

the 2D case. However, the spatial distribution of these changes is more complex, reflecting the heterogeneity and structural complexity of the Brugge dataset. This emphasizes the need for detailed, high-resolution models to capture the nuances of CO_2 injection and its impact on the subsurface in a realistic case.

The histograms and CDFs (Fig. C.31) exhibit a similar trend to the geothermal case, with a slight narrowing of the posterior permeability distribution compared to the prior. This suggests that while the ES-MDA was able to refine the permeability realizations to better match the simulated vertical displacements with the observed ones, the inherent complexity of the Brugge dataset limits the degree of adjustment possible.

The probability maps for CO_2 molar fraction exceeding 0.1 (Fig. C.32) show a more pronounced change in the probability distribution post-data assimilation compared to the geothermal case. This indicates a more significant impact of the data assimilation process on the CO_2 plume prediction.

The detailed analyses of both the geothermal and CO_2 sequestration cases in the Brugge dataset reveal several important insights. Both cases show complex spatial patterns in the permeability adjustments, reflecting the geological intricacy of the Brugge dataset and underscoring the importance of using realistic, detailed geological models in reservoir simulations and data assimilation processes. The histograms and CDFs for both cases show relatively modest changes in the overall permeability distributions, suggesting that while ES-MDA is able to refine the model to better match observations, the complex nature of the Brugge dataset constrains the extent of these adjustments. The CO_2 sequestration case shows more localized changes around the injection site, while the geothermal case exhibits more distributed adjustments, likely stemming from the distinct physical processes and operational setups of these two applications. These results highlight the challenges of applying data assimilation techniques to complex, realistic reservoir models, with improvements being more subtle than in simpler, conceptual models. The modest adjustments in permeability distributions align with the entropy reduction results presented in the main text, suggesting that while vertical displacement data can improve model predictions, additional data types or more sophisticated assimilation techniques may be necessary to achieve more substantial uncertainty reductions in complex reservoirs. These detailed analyses support the main conclusions drawn in the paper, providing a more nuanced understanding of how ES-MDA performs when applied to a complex, realistic reservoir model like the Brugge dataset, and underscore both the potential and the limitations of using surface displacement data for subsurface characterization in energy transition applications.

C.3. Validation analyses

To validate our data assimilation approach, we compared model predictions before and after assimilation at locations outside the original assimilation area. In both CO_2 sequestration and geothermal *conceptual cases*, data assimilation was performed using displacement measurements along line $j = 25$, with validation along line $j = 10$ as indicated with an orange line in Fig. C.33. For the *Brugge cases*, assimilation occurred along line $j = 31$ and validation along line $j = 20$ which is shown by the green line in Fig. C.34. Fig. C.35 shows that posterior model predictions align more closely with reference data at these independent locations, validating our approach beyond the assimilated data points.

To demonstrate the forecasting ability we simulated the Geothermal models for a longer period of 50 years using the permeability distributions from the prior and the posterior, generated by data assimilation using the vertical displacements at 10 years. The temperature at the production well for prior and posterior ensembles is shown in Fig. C.36 for the conceptual and the Brugge cases indicating the reduction in the uncertainty. The temperature drop by 10 °C has been set as a threshold. The period, used in the data assimilation is indicated by the light blue rectangle.

Data availability

The developed code is open source and available at the gitlab repository https://gitlab.com/open-darts/subsidence-esmda.

References

1. IEA. *Carbon Capture, Utilisation and Storage - Fuels & Technologies*. Technical Report; International Energy Agency; 2022 URL https://www.iea.org/reports/carbon-capture-utilisation-and-storage-2.
2. Witter JB, Trainor-Guitton WJ, Siler DL. Uncertainty and risk evaluation during the exploration stage of geothermal development: A review. *Geothermics*. 2019;78:233–242. http://dx.doi.org/10.1016/j.geothermics.2018.12.011.
3. March R, Doster F, Geiger S. Assessment of CO2 storage potential in naturally fractured reservoirs with dual-porosity models. *Water Resour Res*. 2018;54(3):1650–1668. http://dx.doi.org/10.1002/2017WR022159.
4. Ringrose P, Mathieson A, Wright I, et al The in salah CO2 storage project: Lessons learned and knowledge transfer. 2013:6226–6236. 37, http://dx.doi.org/10.1016/j.egypro.2013.06.551,
5. Shi J-Q, Durucan S, Korre A, Ringrose P, Mathieson A. History matching and pressure analysis with stress-dependent permeability using the In Salah CO2 storage case study. *Int J Greenh Gas Control*. 2019;91. http://dx.doi.org/10.1016/j.ijggc.2019.102844.

6. Thomas F, Livio FA, Ferrario F, Pizza M, Chalaturnyk R. A review of subsidence monitoring techniques in offshore environments. *Sensors*. 2024;24(13). http://dx.doi.org/10.3390/s24134164.
7. Davis E, Wright C, Demetrius S, Choi J, Craley G. Precise tiltmeter subsidence monitoring enhances reservoir management. In: *SPE/AAPG Western Regional Meetings*. 2000:443–448. http://dx.doi.org/10.2118/62577-ms.
8. Tang M, Ju X, Durlofsky LJ. Deep-learning-based coupled flow-geomechanics surrogate model for CO2 sequestration. *Int J Greenh Gas Control*. 2022;118. http://dx.doi.org/10.1016/j.ijggc.2022.103692.
9. Van Wees JD, Pluymaekers M, Osinga S, et al 3-D mechanical analysis of complex reservoirs: A novel mesh-free approach. *Geophys J Int*. 2019;219(2):1118–1130. http://dx.doi.org/10.1093/gji/ggz352.
10. Paullo Muñoz LF, Roehl D. An analytical solution for displacements due to reservoir compaction under arbitrary pressure changes. *Appl Math Model*. 2017;52:145–159. http://dx.doi.org/10.1016/j.apm.2017.06.023.
11. Candela T, Pluymaekers M, Ampuero J-P, et al Controls on the spatio-temporal patterns of induced seismicity in Groningen constrained by physics-based modelling with Ensemble-Smoother data assimilation. *Geophys J Int*. 2022;229(2):1282–1308. http://dx.doi.org/10.1093/gji/ggab497.
12. Voskov D, Saifullin I, Wapperom M, et al *open Delft Advanced Research Terra Simulator (open-DARTS)*. Zenodo; 2023. http://dx.doi.org/10.5281/zenodo.8046982.
13. Voskov D, Saifullin I, Novikov A, et al open Delft advanced research terra simulator (open-DARTS). *J Open Source Softw*. 2024;9(99):6737. http://dx.doi.org/10.21105/joss.06737.
14. Saifullin I, Voskov D, Novikov A, et al *DARTS Open-Source Reservoir Simulation Framework*. (1):European Association of Geoscientists & Engineers; 2024:1–19. vol. 2024, http://dx.doi.org/10.3997/2214-4609.202437086,
15. Emerick AA, Reynolds AC. Ensemble smoother with multiple data assimilation. *Comput Geosci*. 2013;55:3–15. http://dx.doi.org/10.1016/j.cageo.2012.03.011.
16. Tian X, Volkov O, Voskov D. An advanced inverse modeling framework for efficient and flexible adjoint-based history matching of geothermal fields. *Geothermics*. 2024;116:102849.
17. Tian X, Wapperom M, Gunning J, et al A history matching study for the fluidflower benchmark project. *Transp Porous Media*. 2024:1–27.
18. Evensen G, Vossepoel FC, van Leeuwen PJ. Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem. Cham: Springer International Publishing; 2022. In: Springer Textbooks in Earth Sciences, Geography and Environment; http://dx.doi.org/10.1007/978-3-030-96709-3.
19. Tarantola A. Inverse problem theory and methods for model parameter estimation. Society for Industrial and Applied Mathematics; 2005. In: Other Titles in Applied Mathematics; http://dx.doi.org/10.1137/1.9780898717921,
20. Oliver DS, Reynolds AC, Liu N. *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press; 2008. http://dx.doi.org/10.1017/cbo9780511535642.
21. Tang H, Fu P, Jo H, et al Deep learning-accelerated 3D carbon storage reservoir pressure forecasting based on data assimilation using surface displacement from InSAR. *Int J Greenh Gas Control*. 2022;120. http://dx.doi.org/10.1016/j.ijggc.2022.103765.
22. Gazzola L, Ferronato M, Frigo M, et al A novel methodological approach for land subsidence prediction through data assimilation techniques. *Comput Geosci*. 2021;25(5):1731–1750. http://dx.doi.org/10.1007/s10596-021-10062-1.
23. Candela T, Chitu A, Peters E, et al Subsidence induced by gas extraction: A data assimilation framework to constrain the driving rock compaction process at depth. *Front Earth Sci*. 2022;10. http://dx.doi.org/10.3389/feart.2022.713273.
24. Emerick AA, Reynolds AC. Investigation of the sampling performance of ensemble-based methods with a simple reservoir model. *Comput Geosci*. 2013;17(2):325–350. http://dx.doi.org/10.1007/s10596-012-9333-z.
25. Peters E, Arts RJ, Brouwer GK, et al Results of the brugge benchmark study for flooding optimization and history matching. *SPE Reserv Eval Eng*. 2010;13(03):391–405. http://dx.doi.org/10.2118/119094-PA,
26. Voskov DV. Operator-based linearization approach for modeling of multiphase multi-component flow in porous media. *J Comput Phys*. 2017;337:275–288. http://dx.doi.org/10.1016/j.jcp.2017.02.041.
27. Khait M, Voskov D. Operator-based linearization for efficient modeling of geothermal processes. *Geothermics*. 2018;74:7–18. http://dx.doi.org/10.1016/j.geothermics.2018.01.012.
28. Lyu X, Khait M, Voskov D. Operator-based linearization approach for modeling of multiphase flow with buoyancy and capillarity. *SPE J.*. 2021;26(4):1858–1878. http://dx.doi.org/10.2118/205378-PA.
29. Li L, Voskov D, Yao J, Li Y. Multiphase transient analysis for monitoring of CO2 flooding. *J Pet Sci Eng*. 2018;160:537–554. http://dx.doi.org/10.1016/j.petrol.2017.10.075.
30. Novikov A, Voskov D, Khait M, Hajibeygi H, Jansen JD. A scalable collocated finite volume scheme for simulation of induced fault slip. *J Comput Phys*. 2022;469. http://dx.doi.org/10.1016/j.jcp.2022.111598.
31. Novikov A. *A Finite Volume Framework for Accurate Modeling of Fault Reactivation in Poroelastic Rocks* [Ph.D. thesis]. Delft University of Technology; 2024. http://dx.doi.org/10.4233/uuid:1f4731b3-cbdd-4317-98f7-8520b7446684,
32. Geertsma J. A basic theory of subsidence due to reservoir compaction; the homogeneous case. 1973.
33. Tempone P, Fjær E, Landrø M. Improved solution of displacements due to a compacting reservoir over a rigid basement. *Appl Math Model*. 2010;34(11):3352–3362. http://dx.doi.org/10.1016/j.apm.2010.02.025.
34. Segall P. Induced stresses due to fluid extraction from axisymmetric reservoirs. *Pure Appl Geophys PAGEOPH*. 1992;139(3–4):535–560. http://dx.doi.org/10.1007/BF00879950.
35. Barbosa VCF, Oliveira VC, Arelaro AD, Borges FAS. 3D displacement and stress fields of compacting reservoir: alternative solutions. *Rev Brasileira de Geofisica*. 2022;40(spl):7–19. http://dx.doi.org/10.22564/brjg.v40i5.2139.
36. Nagy D, Papp G, Benedek J. The gravitational potential and its derivatives for the prism. *J Geod*. 2000;74(7–8):552–560. http://dx.doi.org/10.1007/s001900000116.
37. Remy N, Boucher A, Wu J. *Applied Geostatistics with Sgems: A User's Guide*. Cambridge University Press; 2009.
38. Gaspari G, Cohn SE. Construction of correlation functions in two and three dimensions. *Q J R Meteorol Soc*. 1999;125(554):723–757.
39. Swan A. DEUTSCH, CV & journel, AG 1998. GSLIB. Geostatistical software library and user's guide, x+ 369 pp.+ CD-ROM. Oxford, New york: Oxford university press. Price£ 45.00 (hard covers). ISBN 0 19 510015 8.. *Geol Mag*. 1999;136(1):83–108.
40. Shannon CE. A mathematical theory of communication. *Bell Syst Techn J*. 1948;27(3):379–423.
41. Wang Y, Voskov D, Khait M, Bruhn D. An efficient numerical simulator for geothermal simulation: A benchmark study. *Appl Energy*. 2020;264. http://dx.doi.org/10.1016/j.apenergy.2020.114693.
42. Wapperom M, Tian X, Novikov A, Voskov D. FluidFlower Benchmark: Lessons Learned from the Perspective of Subsurface Simulation. *Transp. Porous Media*. 2024;151(5):1033–1052. http://dx.doi.org/10.1007/s11242-023-01984-8.
43. Lyu X, Voskov D. Advanced modeling of enhanced CO2 dissolution trapping in saline aquifers. *Int J Greenh Gas Control*. 2023;127. http://dx.doi.org/10.1016/j.ijggc.2023.103907.
44. Kretzschmar HJ, Wagner W. International steam tables: Properties of water and steam based on the industrial formulation IAPWS-IF97. 2019. http://dx.doi.org/10.1007/978-3-662-53219-5.
45. Khait M, Voskov DV. Operator-based linearization for general purpose reservoir simulation. *J Pet Sci Eng*. 2017;157:990–998. http://dx.doi.org/10.1016/j.petrol.2017.08.009.