Selecting appropriate model complexity: An example of tracer inversion for thermal prediction in enhanced geothermal systems

Hui Wu[1], Zhijun Jin[2], Su Jiang[3], Hewei Tang[4], Joseph P. Morris[5], Jinjiang Zhang[1], and Bo Zhang[6]

[1]Peking University
[2]Institute of Energy, Peking University
[3]Stanford University
[4]Lawrence Livermore National Laboratory
[5]Lawrence Livermore National Laboratory (DOE)
[6]The Key Laboratory of Orogenic Belts and Crustal Evolution, School of Earth and Space Sciences, Peking University

December 7, 2022

Abstract

A major challenge in the inversion of subsurface parameters is the ill-posedness issue caused by the inherent subsurface complexities and the generally spatially sparse data. Appropriate simplifications of inversion models are thus necessary to make the inversion process tractable and meanwhile preserve the predictive ability of the inversion results. In the present study, we investigate the effect of model complexity on the inversion of fracture aperture distribution as well as the prediction of long-term thermal performance in a field-scale single-fracture EGS model. Principal component analysis (PCA) was used to map the original cell-based aperture field to a low-dimensional latent space. The complexity of the inversion model was quantitatively represented by the percentage of total variance in the original aperture fields preserved by the latent space. Tracer, pressure and flow rate data were used to invert for fracture aperture through an ensemble-based inversion method, and the inferred aperture field was then used to predict thermal performance. We found that an over-simplified aperture model could not reproduce the inversion data and the predicted thermal response was biased. A complex aperture model could reproduce the data but the thermal prediction showed significant uncertainty. A model with moderate complexity, although not resolving many fine features in the "true" aperture field, successfully matched the data and predicted the long-term thermal behavior. The results provide important insights into the selection of model complexity for effective subsurface reservoir inversion and prediction.

Selecting appropriate model complexity: An example of tracer inversion for thermal prediction in enhanced geothermal systems

Hui Wu[1,2], Zhijun Jin[1,2,3], Su Jiang[4], Hewei Tang[5], Joseph P. Morris[5], Jinjiang Zhang[1], Bo Zhang[1]

[1] School of Earth and Space Sciences, Peking University, Beijing, China.

[2] Institute of Energy, Peking University, Beijing, China.

[3] Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 100083, China

[4] Department of Energy Resources Engineering, Stanford University, Stanford, CA, USA.

[5] Lawrence Livermore National Laboratory, Livermore, CA, USA.

Corresponding author: Hui Wu (hui.wu@pku.edu.cn)

Key Points:

- Combined dimensionality reduction and data assimilation to infer fracture aperture distribution from tracer recovery data.
- Quantitatively investigated the effect of model complexities on the aperture inversion and thermal prediction of a field-scale EGS.
- A moderate model complexity is sufficient to reproduce tracer recovery data and provide accurate thermal predictions.

Abstract

A major challenge in the inversion of subsurface parameters is the ill-posedness issue caused by the inherent subsurface complexities and the generally spatially sparse data. Appropriate simplifications of inversion models are thus necessary to make the inversion process tractable and meanwhile preserve the predictive ability of the inversion results. In the present study, we investigate the effect of model complexity on the inversion of fracture aperture distribution as well as the prediction of long-term thermal performance in a field-scale single-fracture EGS model. Principal component analysis (PCA) was used to map the original cell-based aperture field to a low-dimensional latent space. The complexity of the inversion model was quantitatively represented by the percentage of total variance in the original aperture fields preserved by the latent space. Tracer, pressure and flow rate data were used to invert for fracture aperture through an ensemble-based inversion method, and the inferred aperture field was then used to predict thermal performance. We found that an over-simplified aperture model could not reproduce the inversion data and the predicted thermal response was biased. A complex aperture model could reproduce the data but the thermal prediction showed significant uncertainty. A model with moderate complexity, although not resolving many fine features in the "true" aperture field, successfully matched the data and predicted the long-term thermal behavior. The results provide important insights into the selection of model complexity for effective subsurface reservoir inversion and prediction.

1 Introduction

Flow and transport processes in geothermal reservoirs highly depend on spatially heterogeneous reservoir properties, such as permeability distribution in a hydrothermal system (Cox et al., 2001; Dobson et al., 2003; Shi et al., 2018) and fracture aperture distribution in an enhanced geothermal system (EGS) (Chen & Zhao, 2020; Guo, Fu, Hao, Peters, & Carrigan, 2016; Okoroafor et al., 2022; Wu, Fu, Morris, et al., 2021). Characterizing permeability/aperture fields is important for the modeling, prediction, optimization and long-term risk management of geothermal reservoirs. However, due to the high cost and technical difficulties in directly measuring subsurface fields, available permeability/aperture data are generally spatially sparse. A comprehensive characterization is often performed through the inversion of indirect hydraulic or geophysical data, such as hydraulic and tracer testing data (Berkowitz, 2002; Chen et al., 2013; Somogyvári et al., 2017; Vogt et al., 2012; Wu, Fu, Hawkins, et al., 2021), electrical resistivity (Johnson et al., 2021; Wu et al., 2019), seismic (Emerick, 2018; Liu & Grana, 2020), and so on. A key component of hydraulic/geophysical inversion is a reliable model that can properly simulate the underlying physical processes and output model responses for given model parameters. As analytical models are only applicable to idealized scenarios with over simplified fields, numerical models are required for the inversion of representative fields in real-world applications. The infinite-dimensional space of a heterogeneously distributed field is projected to a finite-dimensional parameter space by discretizing the model on a finite element mesh.

A major challenge in permeability and aperture inversion is the ill-posedness issue caused by the high dimensionality of model parameter space and the scarcity of hydraulic/geophysical data. A reliable numerical approximation of model responses requires a relatively fine discretization, which inevitably leads to a high-dimensional parameter space. Practically available hydraulic/geophysical data are usually insufficient to constrain such a high-dimensional parameter space. To tackle this challenge, dimensionality reduction methods have been used to map the discretization-dependent, cell-based high-dimensional parameter space to a low-dimensional latent space (Jiang et al., 2021; Laloy et al., 2013; Marzouk & Najm, 2009; Tang et al., 2021; Xiao et al., 2022; Zhu & Zabaras, 2018). Principal component analysis (PCA) is a conventional dimensionality reduction method, which learns spatial similarities in training samples (prior permeability or aperture models) and compress the most salient features into a latent space defined by orthogonal principal components (Hawkins et al., 2020; Sarma et al., 2008; Wu, Fu, Hawkins, et al., 2021; Zhang et al., 2020). As a linear transform method, PCA is applicable to Gaussian and log-normal fields that can be fully characterized by two-point statistics. For fields that follow non-Gaussian distributions, nonlinear transform methods are required for effective dimensionality reduction, such as deep learning-based methods (e.g., generative adversarial network and variational autoencoder) that have been widely explored in the recent literature (Canchumuni et al., 2020; Jiang & Jafarpour, 2021; Laloy et al., 2018; Mo et al., 2020).

The reduction of model dimensionality essentially leads to the reduction of model complexity. In a cell-based parameter space, each cell value is tuned independently during inversion, and the model has the maximum variance. Through dimensionality reduction, the prior knowledge in training samples, such as spatial auto-correlation and statistical features, are learned and used to reduce the model degree of freedom (model complexity). The learned prior knowledge serves to constrain the heterogeneous distribution of permeability/aperture fields and regularize subsequent inversion. Inversion on a low-dimensional latent space not only mitigates the ill-posedness issue and make the inversion computationally tractable, but also better honors the spatial auto-correlation nature of permeability/aperture fields than inversion on a cell-based parameter space does. The tuning of a latent parameter changes the overall spatial distribution of a field rather than its value at a single cell, which is a highly desired feature for inversion in a data-scarce environment.

Latent space dimensionality, as a quantitative measure of model complexity, is a hyperparameter that needs to be carefully determined prior to dimensionality reduction. On the one hand, model complexity should be deliberately compromised to accommodate the limited information in hydraulic/geophysical data. On the other hand, the model needs to capture adequate variations in the unknown field to appropriately simulate the underlying physical processes. An extremely complex model is prone to overfitting and may undermine the predictive ability of the inferred permeability/aperture fields, while an over-simplified model may fall into the opposite error of underfitting and be unable to reproduce hydraulic/geophysical data. Unfortunately, due to the many inherent complexities of subsurface reservoirs, it is difficult, if not impossible, to predetermine an ideal model complexity for permeability/aperture inversion from hydraulic/geophysical data. In most previous studies, the dimensionality of latent space is subjectively determined (Yang et al., 2021). Some studies used relatively large latent spaces to preserve at least 90% of the total variance in original cell-based parameter spaces when using PCA for dimensionality reduction (Hawkins et al., 2020; Laloy et al., 2013; Tang et al., 2021; Zhao & Luo, 2020). Some other studies, on the other hand, preserved 50% ~ 60% of the total variance through relatively small latent spaces (Romary, 2009; Fernández-Martínez et al., 2012; Emerick, 2017). The dimensionality of the resultant latent spaces in these studies varies from 30 to 1,000.

Although dimensionality reduction methods have been widely used in permeability/aperture inversion, how to select an appropriate model complexity/latent space to circumvent the dilemma of overfitting and underfitting remains unclear. The effects of model complexity on permeability/aperture inversion and subsequent reservoir performance prediction require further investigation. Several studies examined the effect of latent space dimensionality on forward simulation accuracy by first generating permeability/facies fields from latent spaces with different dimensionalities, and then performing forward simulations on these generated fields (Romary, 2009; Fernández-Martínez et al., 2012). The results indicated that a small latent space (preserving 50% - 60% of the total variance in the original cell-based parameter spaces) was sufficient to accurately simulate the underlying physical processes. Li & Cirpka (2006) investigated the effect of latent space dimensionality on the inversion of a 2D hydraulic conductivity field. With the increase of latent space dimensionality, the consumed computational resources increased, while the inversion error (defined as the root mean square error between true and inferred fields) gradually decreased and converged to a stable value. These studies focus on the effect of model complexity on forward and inversion modeling, but lack analysis of the predictive ability of the inversion results. Results from these studies provide insights into the lower limit of model complexity to prevent underfitting. However, the upper limit of model complexity to avoid overfitting, which manifests as good data match but poor predictive ability, remains unexplored.

The main goal of the present study is to investigate the effect of model complexity on the inversion of fracture aperture distribution as well as the prediction of long-term thermal recovery in an EGS. PCA is used to map the original cell-based aperture distribution to a latent space. The corresponding model complexity is quantitatively represented by the percentage of total variance in the original aperture fields preserved by the latent space. An ensemble-based inversion method, ensemble smoother with multiple data assimilation (ES-MDA), is used for aperture inversion from practically available tracer, pressure and flow rate data. Through this investigation, we aim to analyze not only the minimum model complexity required to appropriately reproduce tracer/pressure/flow rate data, but also the impact of overfitting on thermal performance prediction due to excessive model complexity. The paper is organized as follows. In Section 2, we introduce PCA for the dimensionality reduction of spatially auto-correlated aperture fields. Aperture fields generated from latent spaces with different dimensionalities are compared to demonstrate the effect of latent space dimensionality on model complexity. Section 3 describes a field-scale synthetic EGS model with a predominant horizontal fracture, followed by the introduction of forward simulation methods (flow, tracer and thermal) as well as a data assimilation framework using ES-MDA. In Section 4, synthetic tracer, pressure and flow rate data are provided to ES-MDA to invert for the latent space obtained from PCA. The inverted latent space is then mapped back to a cell-based aperture field to predict the thermal performance of the EGS model. A series of latent spaces with different dimensionalities are analyzed to investigate the effect of model complexity. Section 5 provides discussions regarding the implication of the results.

2 Principal component analysis for dimensionality reduction

Principle component analysis (PCA), also known as Karhunen-Loève (KL) expansion, is a well-established dimensionality reduction method. It has been broadly used in many subsurface inversion problems to map Gaussian or log-normal fields (e.g., permeability and aperture) to low-dimensional latent spaces that follow the standard normal distribution (Hawkins et al., 2020; Wu, Fu, Hawkins, et al., 2021). To perform PCA on the field of interest, we first generate an ensemble of training fields based on our prior knowledge of the field obtained from geological/geophysical investigations such as core logs, wellbore images and outcrop analysis. The reduction of dimensionality is then achieved by first computing orthogonal principal components from the training fields, and then retaining the most significant principal components as the basis functions to generate new fields through linear combination. The principal components can be calculated through either computing the eigenvectors and eigenvalues of the covariance matrix of the training fields, or directly performing singular value decomposition (SVD) on the training fields. The significance of a principal component is represented by the percentage of the total variance in the training fields preserved by the principal component. For a spatially auto-correlated field, most of the variance in the training fields can be effectively preserved by a small number of principal components. New fields generated from the linear combination of the retained principal components have the same dimensionality as the training fields, and are fully controlled by the weights of the retained principal components. These weights form the latent space to be inferred in subsequent inversion. The detailed procedure of PCA has been widely reported in the literature (Liu & Durlofsky, 2020; Wu, Fu, Hawkins, et al., 2021) and therefore not repeated here.

In the current study, the field of interest is the aperture distribution of a 2D fracture. To demonstrate the relationship between latent space dimensionality (i.e., the number of retained principal components) and model complexity, we perform PCA on training aperture fields and then compare the aperture fields reconstructed/generated with different numbers of principal components. We generate 5,000 aperture fields on an 800 m × 800 m domain discretized into a 160 × 160 regular grid. We use sequential gaussian simulations and assume a spherical variogram with a mean of 0.6 mm, a standard deviation of 0.45 mm and a correlation length of 75 m. The generated aperture fields follow a log-normal distribution, and are provided to PCA as training fields. After PCA, 5,000 principal components are obtained and ranked in a descending order according to their significance, i.e., percentage of preserved variance. We then use the first l principal components to reconstruct training aperture fields as well as generate new aperture fields.

We first analyze the effect of l on training aperture field reconstruction. A random aperture field is selected from the training ensemble and reconstructed with l = 2, 10, 50, 200 and 800 as shown in Fig. 1. The preserved percentage of total variance is 1%, 5%, 21%, 56% and 84% for l = 2, 10, 50, 200 and 800 respectively. With a small l, the reconstructed aperture field is almost uniform and misses most of the variance in the training aperture field. With the increase of l, the complexity of the reconstructed aperture field increases, manifesting as the capture of fine features in the training aperture field.

We then analyze the effect of l on new aperture field generation. A random l-dimensional latent parameter vector is first sampled from the standard normal distribution, and then used as the weights of the retained l principal components to generate a new aperture field. For each l, we generated 30 aperture fields and analyze their mean, standard deviation and correlation length (Fig. 2). The new aperture fields also follow log-normal distributions, but the mean, standard deviation and correlation length are different from that of the training fields. Compared with the training aperture fields, aperture fields generated from latent spaces exhibit smaller mean and standard deviation, and larger correlation length. For an extremely small latent space (l = 2), the mean of the generated aperture fields is slightly smaller than that of the training aperture fields. The standard deviation is significantly smaller than that of the training aperture fields, and the correlation length shows the opposite trend. A smaller standard deviation and a larger correlation length lead to a smoother aperture distribution, and therefore a less complex aperture model. With the increase of l, the mean, standard deviation and correlation length gradually approaches to their corresponding values in the training aperture fields.

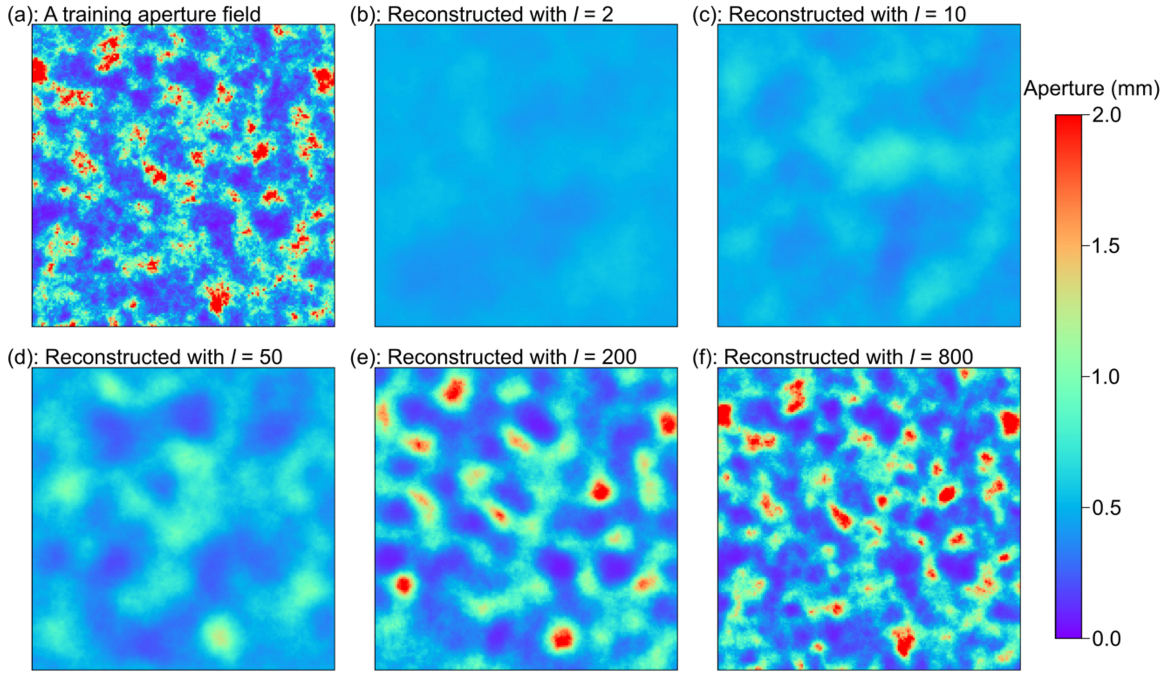


Fig. 1 Reconstruction of a training aperture field. (a) A randomly selected aperture field from the training ensemble. (b) ~ (f) Reconstructed aperture fields using different numbers of principal components.

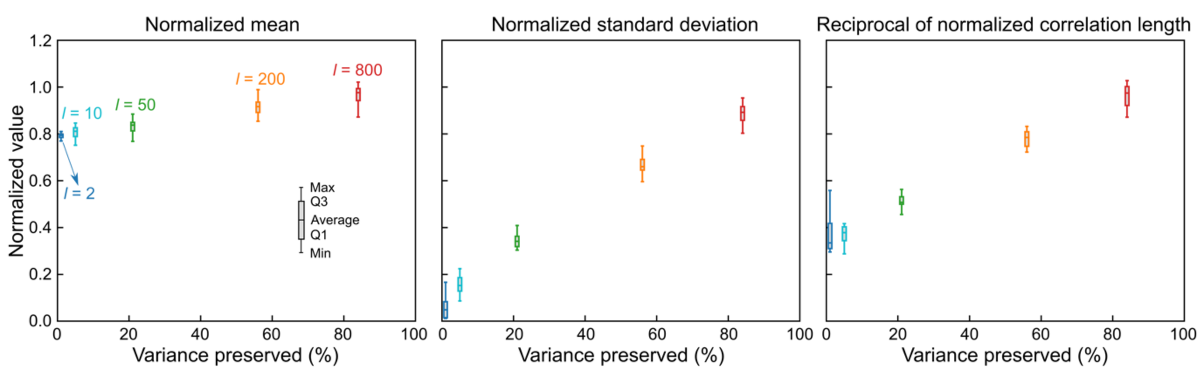


Fig. 2 Box plots of mean, standard deviation and correlation length of aperture fields generated with different numbers of principal components. The box plots show the minimum, maximum, average, as well as the 25% (Q1) and 75% (Q3) percentiles. The mean, standard deviation and correlation length are normalized by their corresponding values used to generate the training aperture fields, i.e., 0.6 mm, 0.45 mm and 75 m respectively. For correlation length, we show the reciprocal of the normalized correlation length.

3 An EGS model and forward simulation/inversion methods

In this section, we develop a field-scale single-fracture EGS model to demonstrate the effect of model complexity on aperture inversion and thermal prediction (Fig. 3). Data used for aperture inversion include practically available tracer, pressure and flow rate data. In what follows, we first describe the model details and then the numerical simulation of flow, tracer transport and thermal extraction in the EGS model. Finally, we briefly introduce a data assimilation framework developed in our previous work (Wu, Fu, Hawkins, et al., 2021), which has proven an effective method for aperture inversion and thermal prediction.

3.1 A field-scale single-fracture EGS model

The developed EGS model is 3000 × 3000 × 3000 m^3 in dimension with a horizontal circular fracture 800 m in diameter, located at the center of the model. An injection well and two production wells are connected by the fracture (Fig. 3(a)). A vertical temperature gradient of 40 °C/km is assumed in the model with an initial temperature of 200 °C at the fracture depth.

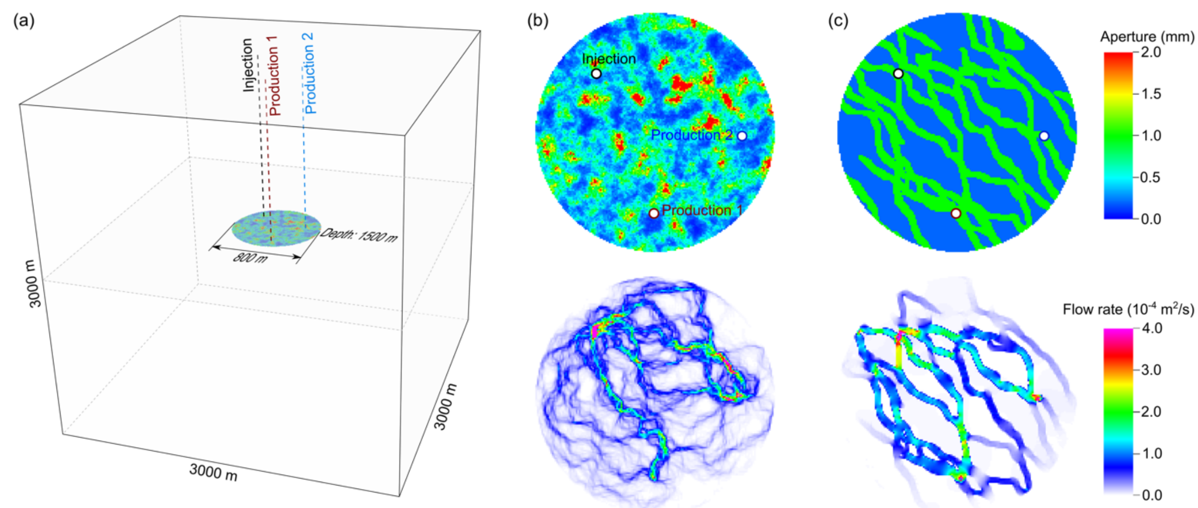


Fig. 3 (a) A field-scale EGS model with a horizontal circular fracture located at 1,500 m depth. (b) A Gaussian aperture field and the corresponding flow field under 20 L/s injection rate and constant pressure at the two production wells. (c) A non-Gaussian, two facies aperture field and the corresponding flow field under 20 L/s injection rate and constant pressure at the two production wells.

We investigate two “true” aperture fields, one is a spatially auto-correlated log-normal field (Fig. 3(b)) and the other is a two facies field (Fig. 3(c)). The log-normal aperture field is randomly generated from sequential gaussian simulation assuming a spherical variogram with a mean of 0.6 mm, a standard deviation of 0.45 mm and a correlation length of 75 m. The two facies aperture field is generated using the ‘snesim’ geostatistical algorithm (Strebelle, 2002) from a geostatistical tool box SGeMS (Remy et al., 2009). The background aperture is 0.2 mm and the aperture of flow channels is 1 mm. Note that the three wells are connected by flow channels. Although the two aperture fields follow different statistical distributions, we use the same aperture model in subsequent inversion for them, i.e., log-normal aperture model. We use the two facies aperture example to demonstrate the scenario where the statistical distribution of the ground truth field fundamentally differs from that of the assumed aperture distribution in the inversion process. This is commonly encountered in real-world problems as the ground truth field is complex and we do not have sufficient data to correctly characterize its statistical distribution.

3.2 Flow, tracer and thermal simulation

Flow and tracer simulation is performed to generate synthetic data for the “true” aperture fields, including tracer breakthrough curves (BTCs) and flow rates at the two production wells, as well as the pressure difference between the injection and production wells. The data are then provided to a data assimilation framework (Section 3.3) for aperture inversion, during which tracer simulation is used as forward model to simulate tracer, pressure and flow rate responses under various aperture scenarios. After inversion, thermal simulation is performed for both the “true” and inferred aperture fields to examine the predictive ability of the inferred aperture fields.

The discretization of the model is as follows. The fracture plane is represented by a thin layer 4 mm in thickness, and the in-plane mesh resolution is 5×5 m^2 within the circular fracture and gradually increases to 150×150 m^2 in the far field. For the surrounding rock formations, the mesh resolution is $5 \times 5 \times 2.5$ m^3 near the fracture plane and becomes progressively coarser in the far field. The resulting computational domain consists of approximately 2,800,000 elements. A massively parallel multi-physics simulation platform developed at the Lawrence Livermore National Laboratory (Settgast et al., 2017), GEOS, is used for flow, tracer and thermal simulation. GEOS provides a thermal-hydro-mechanical-chemical framework to simulate various physical processes in subsurface reservoirs, such as fluid flow, mass and heat transport, and hydraulic fracturing (Fu et al., 2013; Fu et al., 2016; Vogler et al., 2018; Wu, Fu, Morris, et al., 2021). The implementation of flow, tracer and thermal modules relevant to the present study has been described in the literature (Guo, Fu, Hao, & Carrigan, 2016; Guo, Fu, Hao, Peters, & Carrigan, 2016), and therefore not repeated here.

We first simulate the flow field and then solve the advection-dispersion-sorption equation based on the obtained flow field to simulate tracer transport processes. Note that we do not consider mechanical simulation, indicating that the fracture aperture distribution does not evolve during flow and tracer transport processes. Table 1 lists the parameters for flow and tracer modeling. As the fracture plane is represented by a thin layer, we calculate the equivalent porosity and permeability of the fracture through $\phi = w/H$ and $k = w^3/12H$ respectively (Guo, Fu, Hao, & Carrigan, 2016), where w is the aperture and H is the thickness of the fracture layer. Due to the relatively low rock formation permeability and the minor effect of matrix diffusion on tracer transport (Wu, Fu, Hawkins, et al., 2021), we assume that tracer transport is confined within the circular fracture and only consider the fracture for tracer modeling. Fracture boundaries are assumed impermeable. A hydrostatic initial pressure is assumed in the model with a pressure of 30 MPa at the fracture depth. The flow field in the fracture is simulated with an injection rate of 20 L/s and a constant downhole pressure of 30 MPa at the two production wells (Fig. 3(b) and (c)). According to the simulation results, the pressure difference between the injection and production wells and flow rates at production well 1 and 2 are 824 kPa, 5.9 L/s, and 14.1 L/s respectively for the log-normal aperture field (Fig. 3(b)), and 178 kPa, 12.2 L/s, and 7.8 L/s respectively for the two facies aperture field (Fig. 3(c)).

We inject tracers into the fracture for one hour and simulate tracer transport for 20 hours to obtain tracer BTCs at the two production wells (Fig. 4). We consider both conservative and sorptive tracers. Note that for the simulation of a sorptive tracer, we assume an equilibrium sorption process with a typical partition coefficient of 1 mm. Compared with the conservative tracer BTCs, the sorptive tracer BTCs exhibit delayed peaks and reduced peak concentrations due to sorption effects. The tracer BTCs from the two facies aperture field show earlier arrival and larger peak concentration magnitude than that from the log-normal aperture field do, especially for the BTCs at production well 1.

For thermal simulation, we circulate water among the injection and production wells for 50 years, with an injection rate of 20 L/s, an injection temperature of 50 °C, and a constant downhole pressure of 30 MPa at the two production wells. The upper, lower and lateral model boundaries are assumed impermeable to both fluid and heat. Parameters for thermal simulation are also listed in Table 1. For the log-normal aperture field, the production temperature at production well 1 decreases slower than that at production well 2 does (Fig. 4(a)), while for the two facies aperture field, the production temperature at production well 1 decreases faster.

Table 1 Parameters for flow, tracer and thermal simulations of the EGS model.

Parameter	Value
Porosity of rock matrix	0.01
Permeability of rock matrix (m^2)	1×10^{-16}
Density of rock matrix (kg/m^3)	2500
Specific heat capacity of rock matrix (J/kg/K)	790
Thermal conductivity of rock matrix (W/m/K)	2.5
Density of water (kg/m^3)	887.2
Dynamic viscosity of water (Pa·s)	1.42×10^{-4}
Compressibility of water (Pa^{-1})	5×10^{-10}
Specific heat capacity of water (J/kg/K)	4460
Longitudinal dispersivity (m)	0.2
Transverse dispersivity (m)	0.02
Partition coefficient (mm)	1

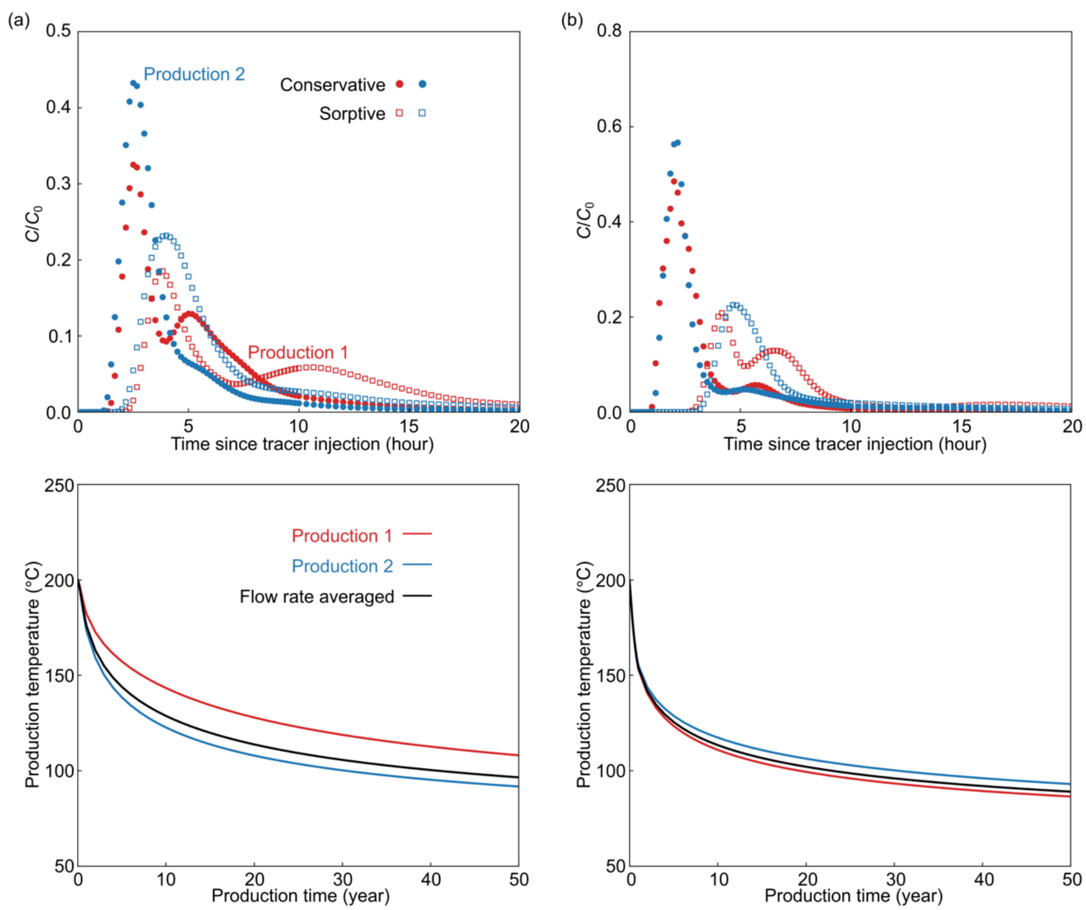


Fig. 4 Tracer (upper) and thermal (lower) breakthrough curves at the two production wells. (a) Results for the log-normal aperture field. (b) Results for the two facies aperture field. The simulated tracer concentration is normalized by injection concentration C_0.

3.3 A data assimilation framework for aperture inversion and thermal prediction

The data assimilation framework developed in Wu, Fu, Hawkins, et al. (2021) is used in the current study for aperture inversion and thermal prediction. The framework includes three major components, i.e., parameterization, inversion and prediction. Here we briefly introduce the key procedures of applying the framework to the aperture inversion and thermal prediction in this study. We refer to Wu, Fu, Hawkins, et al. (2021) for more details of the framework.

3.3.1 Parameterization

The latent spaces generated in Section 2 are used as parameter spaces for aperture inversion. Five latent spaces with dimensionalities of 2, 10, 50, 200 and 800 (Fig. 1) are considered. Note that the aperture field generated from the latent spaces has a square shape (800 m × 800 m), while the aperture field in the EGS model has a circular shape (800 m in diameter). Therefore, only the field within the inscribed circle of the generated 800 m × 800 m aperture field is used for flow, tracer and thermal simulation.

3.3.2 Aperture inversion using ES-MDA

The synthetic tracer, pressure and flow rate data from the “true” aperture fields (Section 3.2) are used for latent space inversion through ES-MDA. An advantage of ES-MDA over deterministic inversion methods is that a posterior ensemble of viable realizations (instead of a single optimal realization) can be obtained to quantify the uncertainties associated with the aperture field. To perform ES-MDA, we first generate 720 l-dimensional latent parameter sets as the prior ensemble by randomly sampling from the standard normal distribution. For each latent parameter set in the ensemble, we use the retained principal components from PCA to map the latent parameter set to an aperture field, and run flow and tracer simulation based upon the aperture field. ES-MDA is then used to update the latent parameter sets according to the simulated and “true” tracer, pressure and flow rate data. The two-step procedure (flow/tracer simulation and latent parameter update) is repeated for 12 iterations to get the posterior ensembles of latent parameter sets and aperture distributions. The major steps and update equation of ES-MDA have been widely described in the literature (Emerick and Reynolds, 2013; Wu, Fu, Hawkins, et al., 2021) and are not repeated here. The key parameters for ES-MDA, such as data standard deviation and inflation coefficient, are the same as that in Wu, Fu, Hawkins, et al. (2021). Note that 3% random Gaussian noise is added in the synthetic data before inversion with ES-MDA.

3.3.3 Thermal prediction based on posterior aperture fields

After ES-MDA, the obtained posterior aperture fields are incorporated into the EGS model to perform thermal simulation and predict temperature responses at the two production wells. Predictions from different posterior aperture fields are used to analyze the associated uncertainties.

4 Aperture inversion and thermal prediction with different model complexities

4.1 Comparison between “true” and simulated tracer, pressure and flow rate data

In each ES-MDA iteration, we record the simulation results (tracer, pressure and flow rate) to compare with the “true” data (Figs. 5 and 6). The latent space dimensionality (l) shows similar effects on the fit of tracer/pressure/flow rate data for the log-normal and two facies aperture fields, as summarized below.

- For $l = 2$, the simulated tracer BTCs, pressure and flow rates show little variation among the 720 prior realizations (before ES-MDA) (Figs. 5(a) and 6(a)). This is because the aperture fields generated from such a low-dimensional latent space have little variance and are relatively smooth (i.e., low model complexity), as shown in the first column of Figs. 7 and 8. After ES-MDA, the 720 parameter sets collapse to the same posterior parameter set (Fig. S1 in the Supporting Information), similar to the small ensemble size-induced ensemble collapse phenomenon reported in many previous studies (Nejadi et al., 2017; Xiao & Tian, 2020). However, the simulation results from this posterior parameter set cannot correctly fit the “true” data (Figs. 5(a) and 6(a)), which is an indicator of underfitting. The second peak of the tracer BTC at production well 1 is not resolved as the underlying aperture model is unable to capture the complexities in the “true” aperture fields.

- When l increases to ten, the complexity of the aperture model increases and the variation among the prior simulation results also increases (Figs. 5(b) and 6(b)). Nevertheless, the 720 parameter sets still collapse to the same posterior parameter set after ES-MDA (Fig. S1). The fit of the "true" data, especially the pressure and flow rate data, is better than that for $l = 2$. The second peak of the tracer BTC at production well 1 is successfully resolved by the posterior realizations, although there still exist some discrepancies between the "true" and simulated tracer BTCs.
- When l further increases to 50, the collapse of parameter sets is greatly alleviated and the posterior latent parameters show considerable uncertainties (Fig. S1). Correspondingly, the uncertainty of the simulation results from the posterior ensemble also increases, especially for the tracer BTCs (Figs. 5(c) and 6(c)). The 90% credible intervals of the simulated tracer BTCs properly match the "true" tracer BTCs, and both the arrival time and magnitude of the second peak of the tracer BTC at production well 1 are correctly reproduced. An aperture model with moderate complexity is able to capture the necessary variations in the "true" aperture field to reproduce the tracer, pressure and flow rate data, even if the aperture model and the true" aperture field follow fundamentally different statistical distributions.
- When l increases to 200 and 800, the latent parameter uncertainties in the obtained posterior ensemble further increase (Fig. S1), and the simulation results from the posterior ensemble show even larger uncertainties compared with that for $l = 50$ (Figs. 5(d), 5(e), 6(d) and 6(e)). With a relatively large latent space dimensionality, the tracer BTCs, pressure and flow rate data can be matched but the associated uncertainties are significant. Further analysis of the predictive ability of the obtained posterior realizations is necessary to examine possible overfitting of the inversion results.

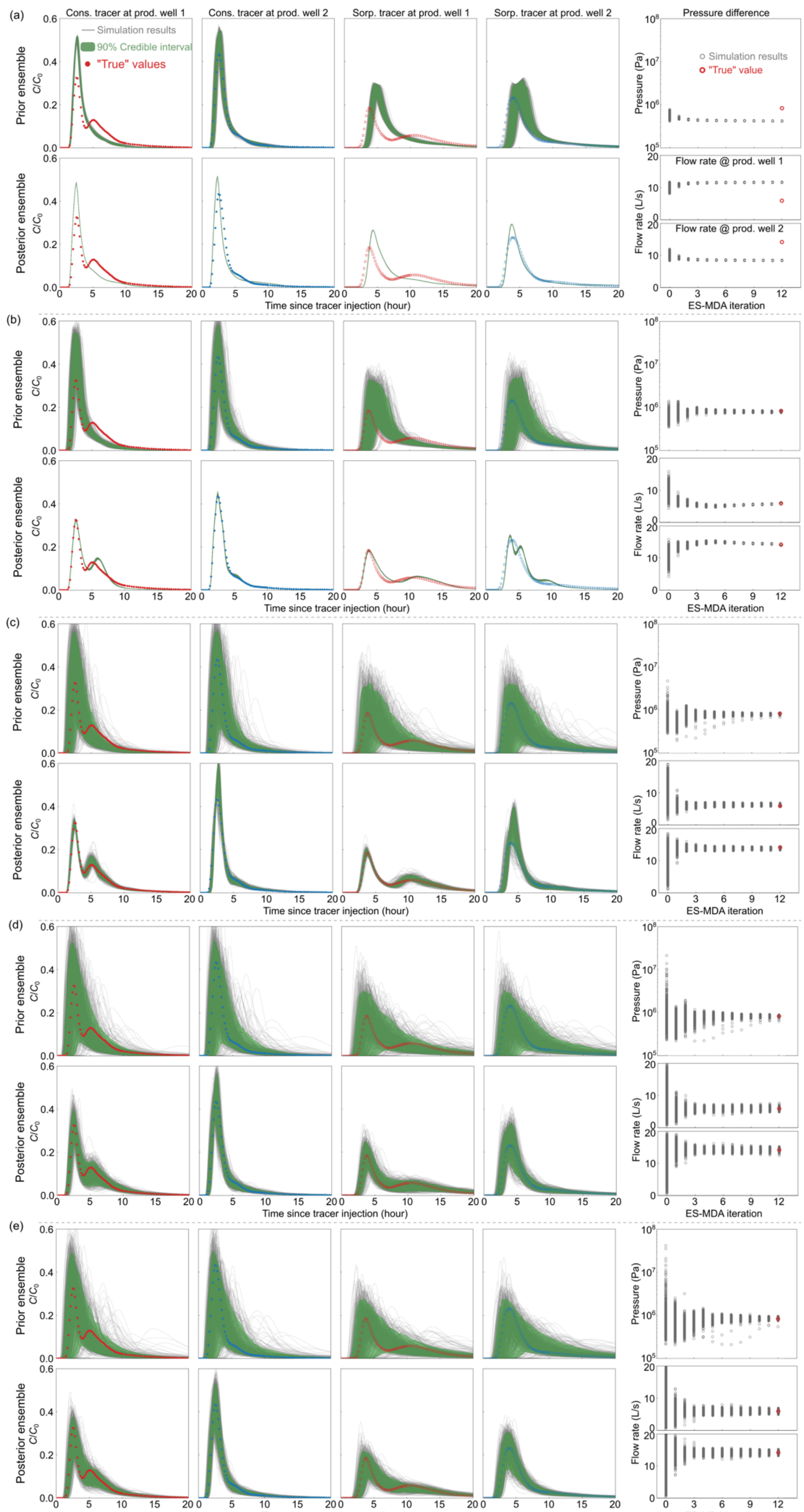
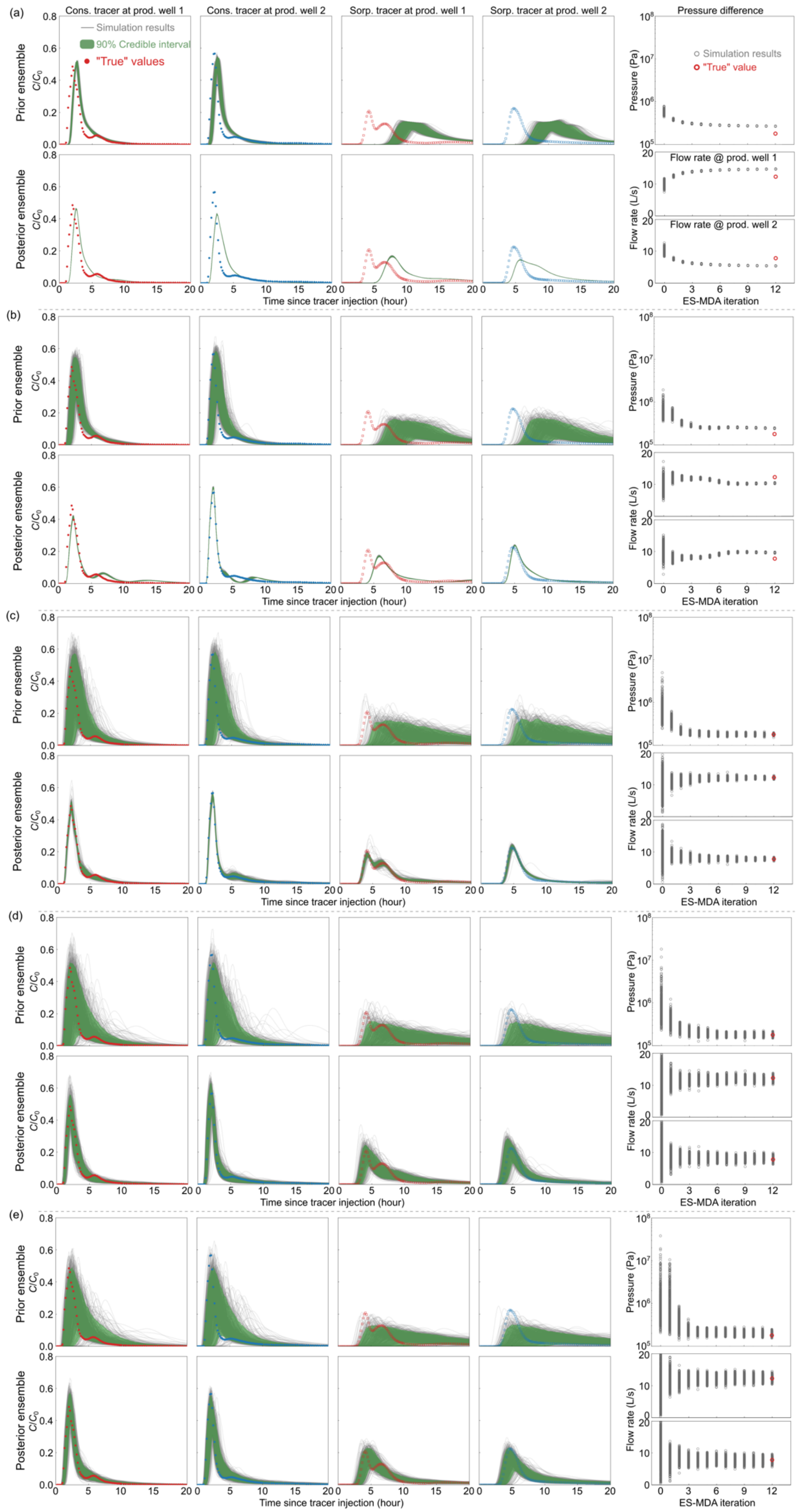

(a)
(b)
(c)
(d)
(e)
Cons. tracer at prod. well 1
Cons. tracer at prod. well 2
Sorp. tracer at prod. well 1
Sorp. tracer at prod. well 2
Pressure difference
Simulation results
90% Credible interval
"True" values
Simulation results
"True" value
Prior ensemble
Posterior ensemble
C/C0
Pressure (Pa)
Flow rate @ prod. well 1
Flow rate @ prod. well 2
Flow rate (L/s)
Time since tracer injection (hour)
ES-MDA iteration

Fig. 5 Comparison of tracer BTCs, pressure and flow rates between the "true" data and the simulation results for the log-normal aperture scenario. (a) Inversion with a latent space dimensionality of $l = 2$. (b) $l = 10$. (c) $l = 50$. (d) $l = 200$. (e) $l = 800$. For tracer BTCs (the first to fourth columns), the upper row shows the results from prior realizations, and the lower row shows the results from posterior realizations. Note that we show the tracer BTCs from all the 720 realizations (grey curves), as well as the corresponding 90% credible intervals (green shadings). For pressure difference and flow rate (the fifth column), we show the evolution of simulation results (grey circles) with respect to ES-MDA iterations. The "true" values are annotated by red circles situated along the final iteration.


(a)
Cons. tracer at prod. well 1
Cons. tracer at prod. well 2
Sorp. tracer at prod. well 1
Sorp. tracer at prod. well 2
Pressure difference
Simulation results
90% Credible interval
"True" values
Simulation results
"True" value
Prior ensemble
Posterior ensemble
C/C0
Pressure (Pa)
Flow rate @ prod. well 1
Flow rate @ prod. well 2
Flow rate (L/s)
Time since tracer injection (hour)
ES-MDA iteration
(b)
(c)
(d)
(e)

394

Fig. 6 Comparison of tracer BTCs, pressure and flow rates between the “true” data and the simulation results for the two facies aperture scenario. (a) Inversion with a latent space dimensionality of $l = 2$. (b) $l = 10$. (c) $l = 50$. (d) $l = 200$. (e) $l = 800$.

4.2 Aperture distribution and flow field in the fracture

We now analyze the aperture distribution and flow field in the fracture (Figs. 7 and 8). We generate aperture distributions from the prior and posterior ensembles, and then perform flow simulations to obtain the corresponding flow fields. For both the log-normal and two facies aperture scenarios, we observe some narrow flow channels connecting the injection and production wells in the “true” flow fields (Fig. 3). When the latent space dimensionality is low ($l = 2$), the aperture field generated from PCA is relatively smooth, and the posterior aperture distributions cannot resolve these narrow flow channels (first column of Figs. 7 and 8). A larger latent space dimensionality leads to a more heterogeneous aperture distribution and therefore a more channelized flow field. When the latent space dimensionality is high, both the prior and posterior realizations exhibit some narrow channels between the injection and production wells (Figs. 7 and 8). Of course, as the prior realization is not conditioned on the tracer, pressure and flow rate data, the corresponding flow field is significantly different from the “true” flow field. For example, the flow field from a prior realization with $l = 800$ (second row, fifth column in Fig. 7) shows two major channels connecting the injection well and production well 2, but misses the channel between the injection well and production well 1. After ES-MDA, the obtained posterior realization shows a flow field that resembles the “true” flow field better than the prior realization does, especially for the high latent space dimensionality cases (fourth row in Figs. 7 and 8).

However, not every flow channel in the “true” flow field is resolved by the posterior realizations. For the log-normal aperture scenario, the “true” flow field shows four major and several minor flow channels between the injection and production wells (Fig. 3(b)), while the posterior realizations only resolve three major flow channels (sixth row in Fig. 7). For the two facies aperture scenario, there are seven flow channels in the “true” flow field (Fig. 3(c)), but only two major flow channels in the flow fields from the posterior realizations (sixth row in Fig. 8). Compared with the “true” flow fields, the flow fields from posterior realizations have fewer flow channels but larger channel width. The relatively large channel width is necessary for the posterior realizations to maintain comparable effective fracture areas as that in the “true” flow fields, so that the tracer, pressure and flow rate data can be matched (especially the sorptive tracer BTC which highly depends on the interaction area between fracture fluid and surrounding rocks). The overall effect of the many narrow flow channels in the “true” flow field is represented by the two or three relatively wide flow channels in the flow fields from posterior realizations.

The latent space dimensionality shows significant effect on the posterior aperture distributions. For a small latent space dimensionality ($l = 2$ or 10), the posterior aperture distributions (flow fields) are almost identical, which is a direct consequence of latent parameter collapse. Besides the randomly selected posterior realization in Figs. 7 and 8 (third and fourth rows in Fig. 7 for the log-normal aperture scenario, and in Fig. 8 for the two facies aperture scenario), we provide two additional, randomly selected posterior realizations for comparison in the Supporting Information (Fig. S2). All the three aperture distributions (flow fields) are almost the same as the average aperture distribution (flow field) displayed in the fifth and sixth rows in

Figs. 7 and 8. The standard deviation of the posterior realizations is negligible (Fig. S3 in the Supporting Information). With the increase of latent space dimensionality, the variations among the posterior aperture distributions (flow fields) increase (Fig. S3). The uncertainty in the latent parameters propagates to the aperture distribution and flow field. For a large latent space dimensionality (l = 200 or 800), we observe considerable variations in aperture distribution among the posterior realizations, but the major flow channels from these posterior realizations are similar, as a result of conditioning on tracer, pressure and flow rate data.

In the next section, we will analyze how the uncertainty in the latent parameters further propagates to thermal predictions and examine possible overfitting of the posterior realizations obtained with a relatively large latent space dimensionality.

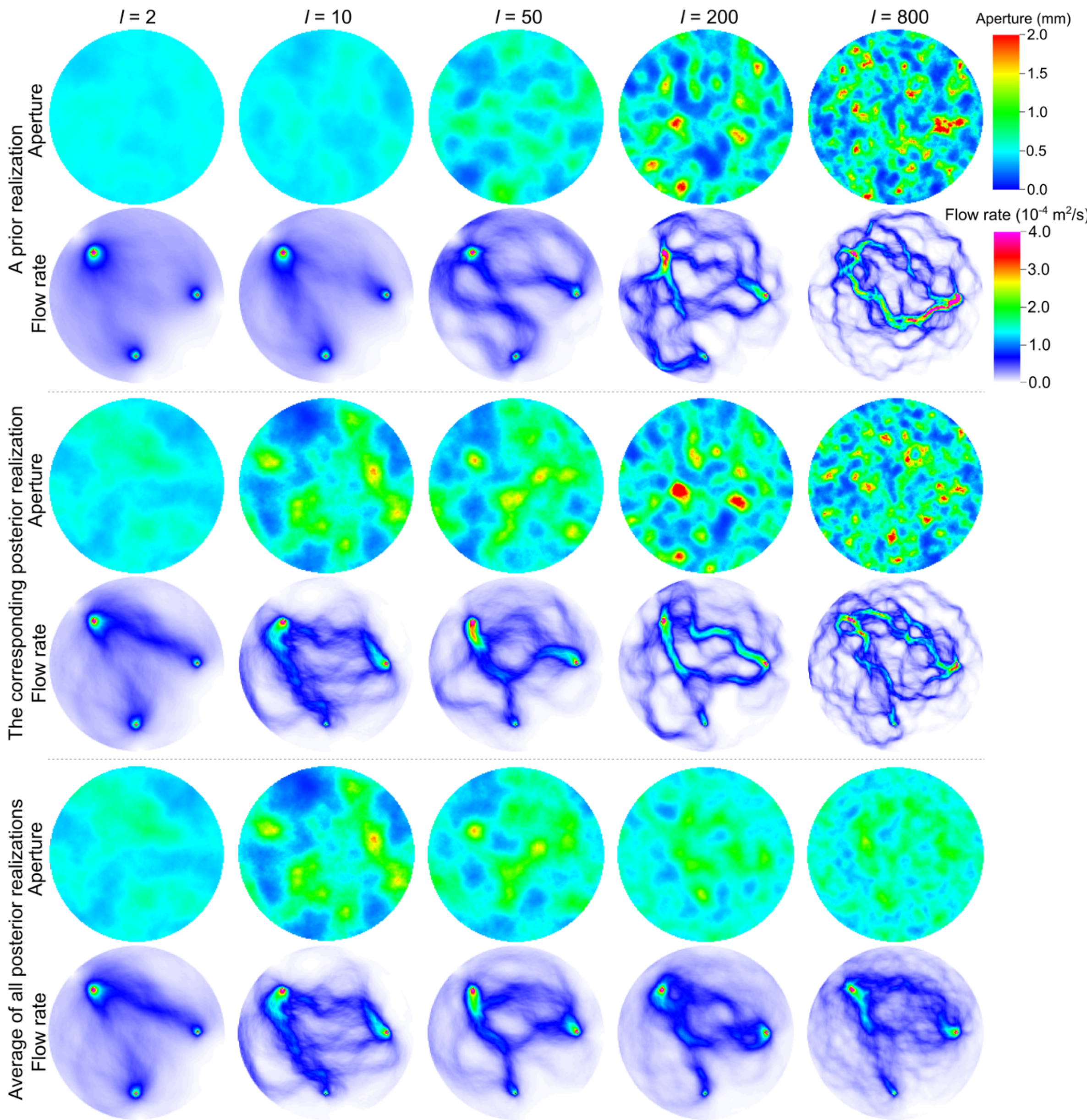


Fig. 7 Aperture distribution and flow field in the fracture for the log-normal aperture scenario. The first and second rows are results from a randomly selected realization in the prior ensemble (before ES-MDA), and the third and fourth rows are results from the corresponding posterior realization (after 12 ES-MDA iterations). The fifth and sixth rows are average results of all the realizations in the posterior ensemble.

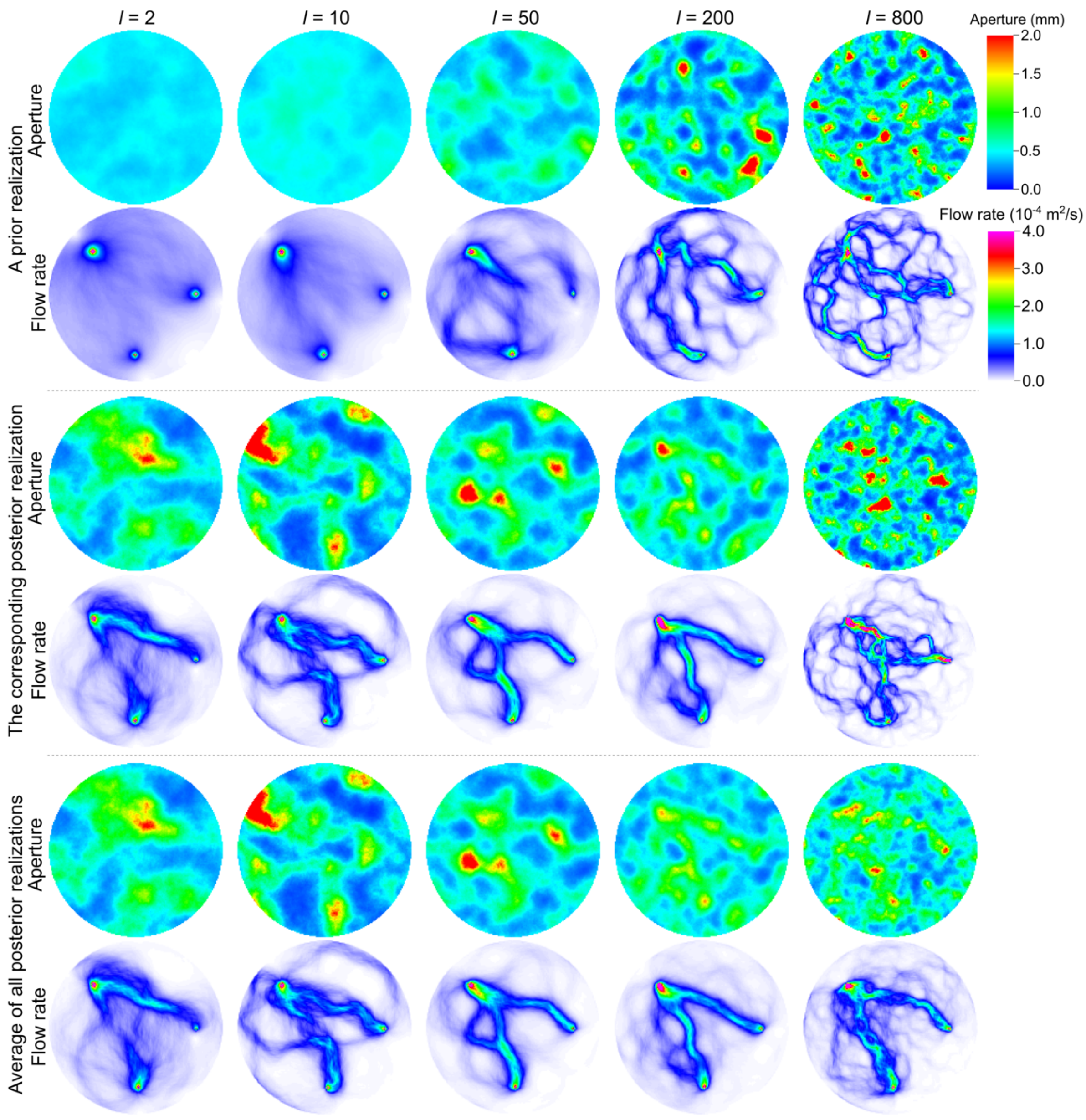


Fig. 8 Aperture distribution and flow field in the fracture for the two facies aperture scenario. The first and second rows are results from a randomly selected realization in the prior ensemble (before ES-MDA), and the third and fourth rows are results from the corresponding realization in the posterior ensemble (after 12 ES-MDA iterations). The fifth and sixth rows are average results of all the realizations in the posterior ensemble.

4.3 Thermal performance prediction

In this sub-section, we perform thermal simulations with both prior and posterior realizations to analyze their abilities in predicting the long-term thermal performance of the EGS model. For each latent space dimensionality, we randomly select ten prior realizations and their corresponding posterior realizations to perform thermal simulation, and then compare the simulated temperature responses with the "true" temperature responses (Figs. 9 and 10).

We first analyze the prior predictions (first and third rows in Figs. 9 and 10). With the increase of latent space dimensionality, the variations among the prior predictions of temperature responses increase (first row in Figs. 9 and 10). For relatively large latent space dimensionalities, the prior predictions vary in broad ranges and many predictions significantly underestimate the temperature reductions. Interestingly, the variation among the prior predictions of the flow rate-averaged temperature response (third row in Figs. 9 and 10), although considerable for large latent space dimensionality cases, is substantially smaller than that of the temperature responses at individual production wells. To understand the reduced variation of the flow rate-averaged temperature response, we select a prior realization for further analysis (Fig. 11). For the prior realization, the predicted flow rate at production well 1 is 2.7 L/s, much smaller than the "true" flow rate (5.9 L/s). As a result, the predicted flow rate at production well 2 is much larger than the corresponding "true" flow rate. The temperature decrease is highly related to the flow rate. The underestimated flow rate at production well 1 results in a slow temperature decrease at production well 1, and the overestimated flow at production well 2 leads to a fast temperature decrease at production well 2. Since the flow rates at the two production wells are not independent, the underestimation of flow rate at one production well means the overestimation of flow rate at the other production well. Therefore, when the temperature decrease at one production well is significantly underestimated (e.g.: production well 1 in Fig. 11(a)), the temperature decrease at the other production well is likely to be overestimated (production well 2 in Fig. 11(a)). As a result, the variation of the temperature prediction at one production well counteracts the variation of the temperature prediction at the other production well, causing a reduced variation in the flow-rated averaged temperature prediction.

Compared with prior realizations, the posterior realizations provide more accurate predictions for both the individual and flow-rate averaged temperature responses (second and fourth rows in Figs. 9 and 10). When latent space dimensionality is low ($l = 2$ or 10), temperature predictions from the ten posterior realizations are almost identical due to the collapse of latent parameters, and cannot match the "true" temperature responses. With the increase of latent space dimensionality, the posterior predictions match the "true" temperature responses better but also show larger uncertainties. When the latent space dimensionality increases to 800, the posterior predictions at individual production wells exhibit significant uncertainties, indicating the overfitting of the obtained posterior realizations. However, the posterior prediction of the flow-rated averaged temperature response still successfully reproduces the "true" response with relatively small uncertainty (fourth row, fifth column in Figs. 9 and 10). Once again this is mainly caused by the counteraction between the underestimation of temperature response at one production well and the overestimation of temperature response at the other production well, as shown by the example in Fig. 11(b).

According to Figs. 9 and 10, to correctly predict the thermal responses at the two production wells, the latent space dimensionality should be in the range of 10 ~ 200 for the log-normal aperture scenario, and 50 ~ 200 for the two facies aperture scenario. While for the flow-rated averaged temperature response, the latent space dimensionality should not be smaller than ten for the log-normal aperture scenario, and a larger latent space dimensionality not smaller than 50 appears to be necessary for the two-facies aperture scenario.

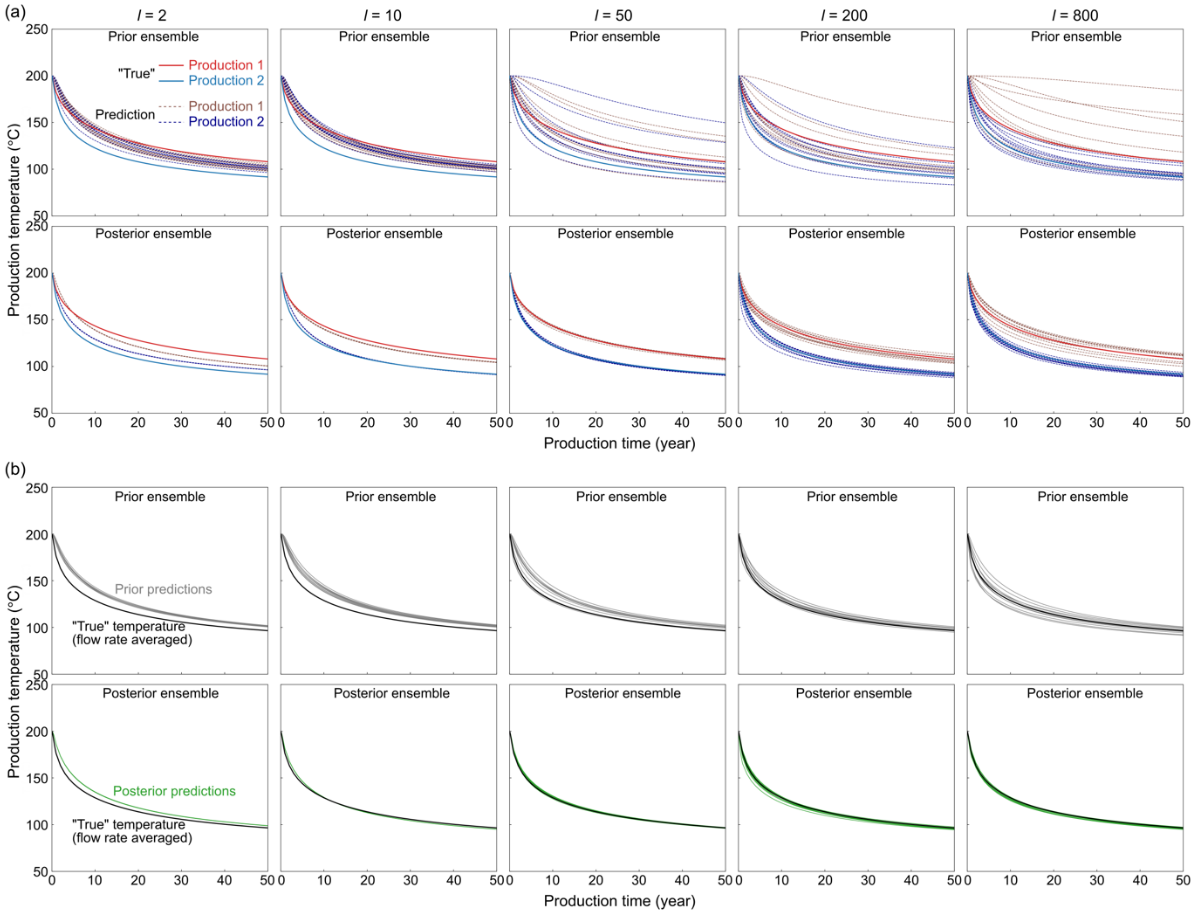


Fig. 9 Prediction of thermal responses from prior and posterior realizations for the log-normal aperture scenario. (a) Temperature responses at the two production wells. The solid lines are “true” temperature responses, and the dash lines are predictions. The upper row shows the predictions from prior realizations, and the lower row shows the predictions from posterior realizations. (b) Flow rate averaged temperature response. The black line is the “true” temperature response. The gray and green lines are predictions from prior (upper row) and posterior (low row) realizations respectively.

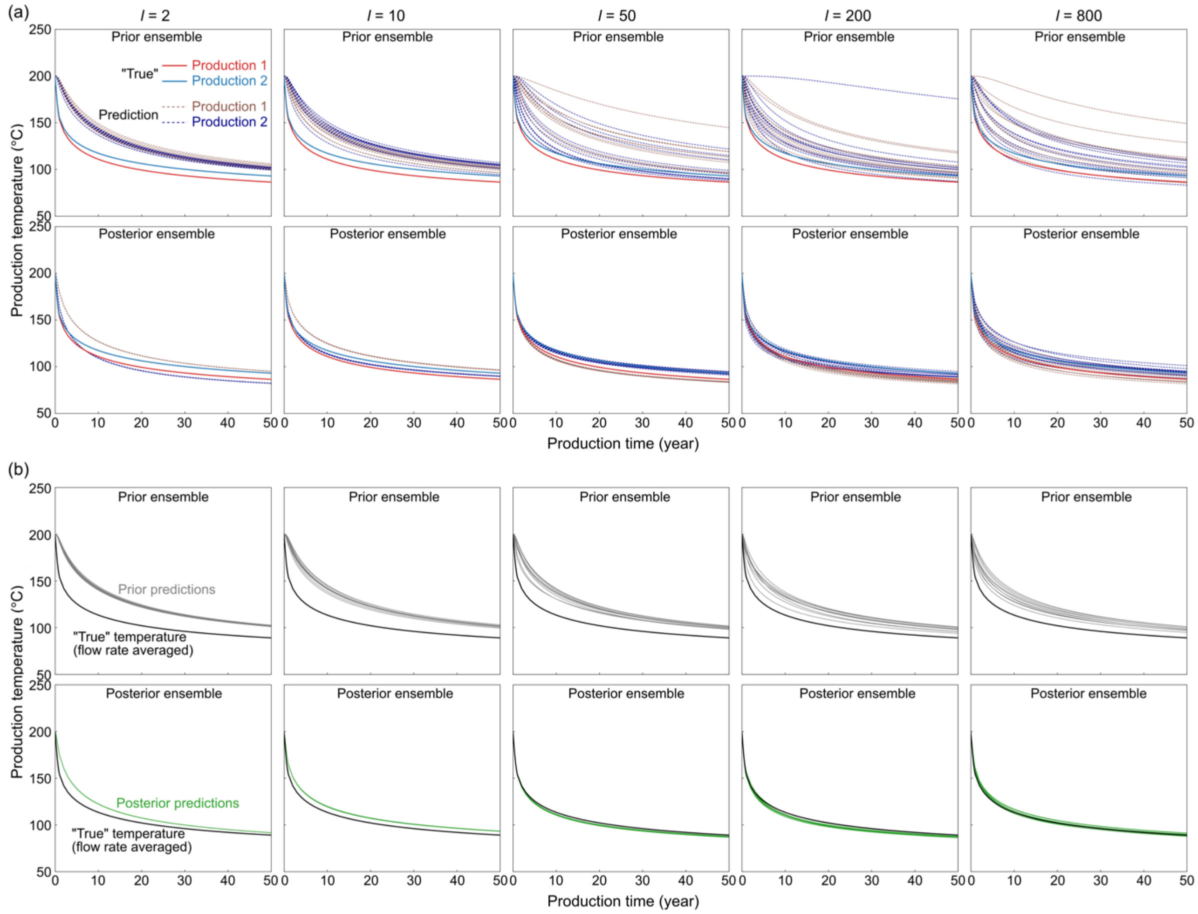


Fig. 10 Prediction of thermal responses from prior and posterior realizations for the two facies aperture scenario. (a) Temperature responses at the two production wells. (b) Flow rate averaged temperature response.

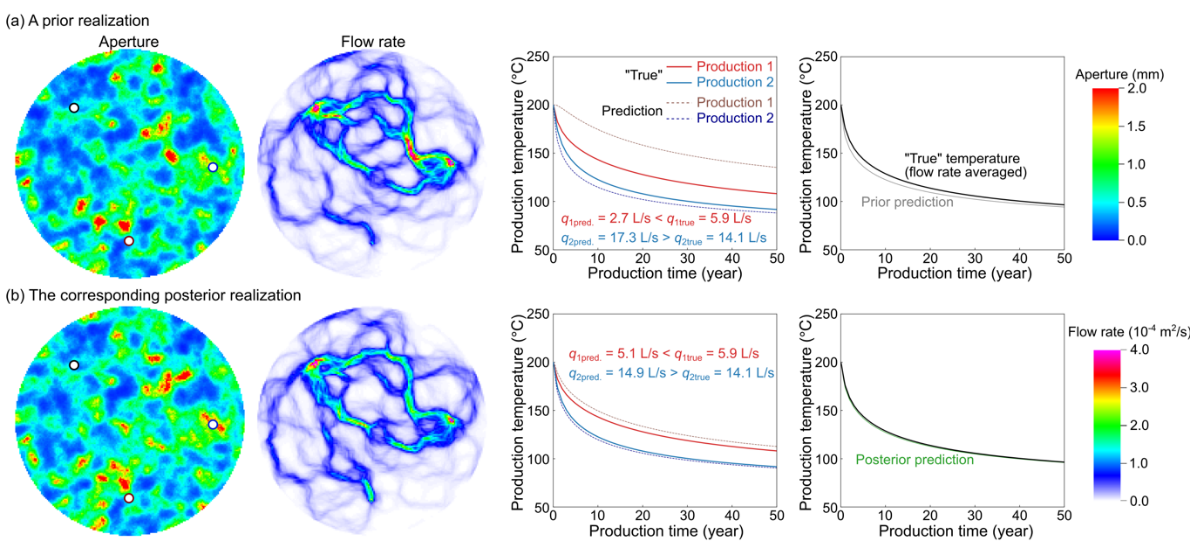


Fig. 11 Comparison of aperture distribution, flow field and temperature responses between a prior realization and the corresponding posterior realization. The prior and posterior realizations are from the ES-MDA case with a latent dimensionality of 800 for the log-normal aperture scenario.

Predicted ($q_{1\text{pred.}}$ and $q_{2\text{pred.}}$) and true ($q_{1\text{true}}$ and $q_{2\text{true}}$) flow rates at the two production wells are annotated.

5 Discussion

5.1 Selecting appropriate model complexity

Model reduction has been considered essential to tackle the challenges associated with complex subsurface conditions and data scarcity in many subsurface inversion and characterization problems (Jiang & Ou, 2017; Marzouk & Najm, 2009; Zhu & Zabaras, 2018). The present study attempts to investigate the effect of model complexity on the inversion and prediction of subsurface reservoirs, and more importantly, to provide some insights into the selection of model complexity to avoid underfitting and overfitting. Through a field-scale EGS model, we demonstrate both underfitting behavior under low model complexity (poor data match) and overfitting behavior under high model complexity (good data match but poor prediction). For the log-normal aperture scenario considered in the present study, an inversion model that preserves 21% of the total variance in the "true" aperture field (corresponding to a latent space dimensionality of $l = 50$) is sufficient to correctly reproduce tracer/pressure/flow rate data and accurately predict long-term thermal performance. Increasing the model complexity to preserve 56% of the total variance ($l = 200$) can also produce satisfactory data fit and thermal prediction results but the associated uncertainties increase. Further increasing the model complexity to preserve 84% of the total variance ($l = 800$) leads to significant uncertainties and the thermal performance cannot be accurately predicted. Therefore, a model corresponds to a latent space dimensionality between 50 and 200, i.e., with 21% to 56% of the total variance preserved, is appropriate for tracer/pressure/flow rate data inversion and thermal prediction in the present study. This is also true for the two facies aperture scenario where the "true" aperture field and the aperture model used for inversion follow different statistical distributions.

The selection of model complexity actually depends on the purpose of inversion. For the presented EGS model, we note that although the posterior aperture fields obtained from the latent space with $l = 50$ (Figs. 7, 8 and S2) can reproduce the "true" data and make accurate predictions, they look different from the "true" aperture fields in Fig. 3. Many fine features in the "true" aperture field could not be resolved due to the lack of necessary complexities in the aperture model generated from such a low-dimensional latent space. Hence, if the primary goal is to infer fracture aperture, a relatively high model complexity is required, but if the primary goal is to predict thermal performance, then a moderate model complexity is sufficient. For many subsurface characterization problems, people are mainly concerned with the predictive ability rather than the realism of the inversion results, and therefore a moderate model complexity could be employed.

The selection of model complexity also needs to consider the amount of information contained in the data for inversion. In general, the more information the data contain, the more complex the model should be to avoid underfitting. In the present study, the inversion data include tracer BTCs/flow rates at two production wells and pressure difference between injection and production wells. The information in these data is spatially limited and far from sufficient to characterize the aperture distribution in the 2D fracture plane. As a result, a relatively simple model is able to retrieve the information and reproduce the data. Fortunately, since both tracer transport and heat extraction processes are tightly related to the flow field among the injection and production wells, the information retrieved from tracer BTCs, although limited, still provide accurate thermal predictions. If more data, for example, tracer BTCs at other locations are available, the model complexity needs to be increased to accommodate the increased amount of information in the inversion data.

The result that a model preserving only 21% of the total variance in the "true" aperture field is able to reproduce tracer/pressure/flow rate data and predict thermal responses is surprising as many studies preserved at least 50% ~ 60% total variance when using PCA for model reduction (Fernández-Martínez et al., 2012; Hawkins et al., 2020; Laloy et al., 2013). An important implication from the current study is that we should use a relatively simple model for inversion/data assimilation in subsurface reservoirs, especially when the available measurements are scarce and prediction is the primary goal. A low complexity model can not only mitigate the overfitting pitfall but also alleviate the computational burden in many subsurface inversion problems. Of course, the model should not be too simple otherwise it may fail to reproduce inversion data.

5.2 Geologic facies model

Geologic facies models have been used to describe highly channelized subsurface reservoirs such as the two facies aperture model in Fig. 3(c). The characterization of such facies models has been widely investigated in recent years (Chang et al., 2010; Jafarpour & McLaughlin, 2009; Jiang & Jafarpour, 2021). To preserve the geologic realism of facies models during inversion/data assimilation, a model reduction method that can directly generate facies models from low-dimensional latent space is required. PCA is inappropriate as facies models do not follow a Gaussian or log-normal distribution. Many methods have been proposed for the reduction of facies models, such as optimization-based PCA (Vo & Durlofsky, 2014), discrete cosine transform (Jafarpour & McLaughlin, 2007), and deep learning algorithms such as variational autoencoder (VAE) (Canchumuni et al., 2019; Laloy et al., 2017; Mo et al., 2020) and generative adversarial network (GAN) (Canchumuni et al., 2020; Laloy et al., 2018).

The current study provides an alternative strategy for the characterization of geologic facies models. Instead of developing advanced model reduction methods for the two facies aperture model in Fig. 3(c), we directly use a log-normal aperture model generated from PCA latent space for data assimilation. Although the "true" and inversion aperture models follow fundamentally different statistical distributions, the obtained posterior aperture models are able to reproduce tracer/pressure/flow rate data and predict long-term thermal performance after data assimilation. A log-normal aperture model with appropriate correlation length is capable of inducing relevant channelized flow structures analogous to those of a facies-based model. The posterior aperture distributions fail to preserve the geologic realism in the two facies aperture model (Fig. 8). However, as Murray (2007) concluded, if prediction instead of explanation is the primary goal, the realism of model parameters should not be considered an essential model-evaluation criterion.

5.3 Prior realization

An interesting observation is that a prior aperture realization and the corresponding posterior aperture realization have many common features, especially when the aperture model is relatively complex (fourth and fifth columns in Figs. 7 and 8, Fig. 11). In another words, ES-MDA tends to perturb a prior realization as slightly as possible to match the data being assimilated. The obtained posterior realization largely depends on the prior realization provided to ES-MDA. If the prior realization is not well constrained or even physically unrealistic, then the corresponding posterior realization may also show unrealistic features. Therefore, it is of great importance to constrain prior realizations with available prior knowledge. Fortunately, for subsurface reservoirs, prior knowledge of the field of interest (aperture or permeability) can be obtained from geological/geophysical measurements, such as core logs, wellbore images and outcrop analysis. In the current study, the prior knowledge used to constrain prior aperture realizations includes the spatially autocorrelated nature as well as the mean, standard deviation and correlation length of the aperture field.

Interpretation results from other geophysical investigations, such as seismic and ground penetrating radar (GPR), can also be used to constrain prior realizations. For example, Wu, Fu, Hawkins, et al. (2021) used the results of GPR survey to constrain prior aperture realizations in a horizontal fracture at a meso-scale field test site (the Altona Field Laboratory located in northern New York State, USA). The GPR survey of the field test site indicated a narrow flow channel between injection and production wells, which meant that the underlying aperture field was anisotropic with larger correlation length in the direction from injection to production wells (west to east) than that in the south to north direction. During subsequent tracer data assimilation, such an anisotropic feature was used as prior knowledge to constrain prior aperture realizations.

6 Conclusion

We investigated the effect of model complexity on the inversion of fracture aperture distribution as well as the prediction of long-term thermal recovery in a field-scale EGS model. Inversion models with different complexities were used to invert for fracture aperture distribution through the assimilation of tracer/pressure/flow rate data using an ensemble-based method (ES-MDA). Thermal simulations were then performed to examine the predictive ability of the inferred aperture distributions. With a low model complexity, ensemble collapse occurred. The inferred aperture distributions failed to reproduce tracer/pressure/flow rate data, and the predicted long-term thermal response was biased. With a high complexity model, the data could be properly matched, but the inferred aperture distribution and predicted thermal response exhibit significant uncertainties. A moderate model complexity is sufficient to retrieve the information contained in tracer/pressure/flow rate data and provide accurate thermal predictions.

An appropriate model complexity is essential to the inversion and prediction of subsurface reservoirs, and deserves careful deliberation based on the primary purpose of the inversion as well as the type and amount of the inversion data. In a real-world application, it is difficult to predetermine model complexity and one might need to manually adjust model complexity in a trial-and-error manner. According to the results in the current study, we recommend starting with a relatively simple model rather than an extremely complex model, and the quality of the fit to tracer BTCs appears to be a reasonable indicator of an appropriate model complexity.

Data Availability Statement

The synthetic flow, pressure, tracer and thermal data used in this study is obtained from numerical simulations with GEOS. The data assimilation framework is available in Wu, Fu, Hawkins, et al. (2021).

Acknowledgments, Samples, and Data

We thank Dr. Pengcheng Fu for providing critical and constructive comments on various aspects of this study. This research was performed in support of the EGS Collab project and the EGS Collab team is gratefully acknowledged. This work was supported by U.S. Department of Energy, Geothermal Technologies Office, and performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. This document is LLNL report LLNL-JRNL-838540. This work is also supported by the National Key Research and Development Program of China (No. 2021YFA0716000), and the China National Petroleum Corporation-Peking University Strategic Cooperation Project of Fundamental Research.

References

Berkowitz, B. (2002). Characterizing flow and transport in fractured geological media: A review. *Advances in Water Resources*, *25*, 861-884. https://doi.org/10.1016/S0309-1708(02)00042-8

Canchumuni, S. W. A., Castro, J. D. B., Potratz, J., Emerick, A. A., & Pacheco, M. A. C. (2020). Recent developments combining ensemble smoother and deep generative networks for facies history matching. *Computational Geosciences*, *25*, 433-466. https://doi.org/10.1007/s10596-020-10015-0

Canchumuni, S. W. A., Emerick, A. A., & Pacheco, M. A. C. (2019). Towards a robust parameterization for conditioning facies models using deep variational autoencoders and ensemble smoother. *Computers & Geosciences*, 128, 87-102. https://doi.org/10.1016/j.cageo.2019.04.006

Chang, H., Zhang, D., & Lu, Z. (2010). History matching of facies distribution with the EnKF and level set parameterization. *Journal of Computational Physics*, 229, 8011-8030. https://doi.org/10.1016/j.jcp.2010.07.005

Chen, X., Hammond, G. E., Murray, C. J., Rockhold, M. L., Vermeul, V. R., & Zachara, J. M. (2013). Application of ensemble-based data assimilation techniques for aquifer characterization using tracer data at Hanford 300 area. *Water Resources Research*, *49*(10), 7064-7076. https:// doi.org/10.1002/2012WR013285

Chen, Y., & Zhao, Z. (2020). Heat transfer in a 3D rough rock fracture with heterogeneous apertures. *International Journal of Rock Mechanics and Mining Sciences*, 134, 104445. https://doi.org/10.1016/j.ijrmms.2020.104445

Cox, S. F., Knackstedt, M. A., & Braun, J. (2001). Principles of structural control on permeability and fluid flow in hydrothermal systems. *Reviews in Economic Geology*, 14, 1-24. https://doi.org/10.5382/Rev.14.01

Dobson, P. F., Kneafsey, T. J., Hulen, J., & Simmons, A. (2003). Porosity, permeability, and fluid flow in the Yellowstone geothermal system, Wyoming. *Journal of Volcanology and Geothermal Research*, 123, 313-324. https://doi.org/10.1016/S0377-0273(03)00039-8

Emerick, A. A. (2017). Investigation on principal component analysis parameterizations for history matching channelized facies models with ensemble-based data assimilation. *Mathematical Geosciences*, 49(1), 85-120. https://doi.org/10.1007/s11004-016-9659-5

Emerick, A. A. (2018). Deterministic ensemble smoother with multiple data assimilation as an alternative for history-matching seismic data. *Computational Geosciences*, *22*(5), 1175-1186. https://doi.org/10.1007/s10596-018-9745-5

Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55, 3-15. https://doi.org/10.1016/j.cageo.2012.03.011

Fernández-Martínez, J. L., Mukerji, T., García-Gonzalo, E., & Suman, A. (2012). Reservoir characterization and inversion uncertainty via a family of particle swarm optimizers. Geophysics, 77(1), M1-M16. https://doi.org/10.1190/geo2011-0041.1

Fu, P., Hao, Y., Walsh, S. D. C., & Carrigan, C. R. (2016). Thermal drawdown-induced flow channeling in fractured geothermal reservoirs. *Rock Mechanics and Rock Engineering*, 49(3), 1001–1024. https://doi.org/10.1007/s00603-015-0776-0

Fu, P., Johnson, S. M., & Carrigan, C. R. (2013). An explicitly coupled hydro-geomechanical model for simulating hydraulic fracturing in arbitrary discrete fracture networks. *International Journal for Numerical and Analytical Methods in Geomechanics*, 37, 2278-2300. https://doi.org/10.1002/nag.2135

Guo, B., Fu, P., Hao, Y., & Carrigan, C. R. (2016). Investigating the possibility of using tracer tests for early identification of EGS reservoirs prone to flow channeling. 41st workshop on geothermal reservoir engineering.

Guo, B., Fu, P., Hao, Y., Peters, C. A., & Carrigan, C. R. (2016). Thermal drawdown-induced flow channeling in a single fracture in EGS. *Geothermics*, 61, 46-62. https://doi.org/10.1016/j.geothermics.2016.01.004

Hawkins, A. J., Fox, D. B., Koch, D. L., Becker, M. W., & Tester, J. W. (2020). Predictive inverse model for advective heat transfer in a short-circuited fracture: Dimensional analysis, machine learning, and field demonstration. *Water Resources Research*, 56, e2020WR027065. https:// doi.org/10.1029/2020WR027065

Jafarpour, B., & McLaughlin, D. B. (2007). Efficient permeability parameterization with the discrete cosine transform. *SPE reservoir simulation symposium*.

Jafarpour, B., & McLaughlin, D. B. (2009), Estimating channelized-reservoir permeabilities with the ensemble Kalman filter: The importance of ensemble design. *SPE Journal*, 14(2), 374-388. https://doi.org/10.2118/108941-PA

Jiang, A., & Jafarpour, B. (2021). Deep convolutional autoencoders for robust flow model calibration under uncertainty in geologic continuity. *Water Resources Research*, 57(11), e2021WR029754. https://doi.org/10.1029/2021WR029754

Jiang, L., & Ou, N. (2017). Multiscale model reduction method for Bayesian inverse problems of subsurface flow. *Journal of Computational and Applied Mathematics*, 319, 188-209. https://doi.org/10.1016/j.cam.2017.01.007

Jiang, Z., Zhang, S., Turnadge, C., & Xu, T. (2021). Combining autoencoder neural network and Bayesian inversion to estimate heterogeneous permeability distributions in enhanced geothermal reservoir: Model development and verification. *Geothermics*, 97, 102262. https://doi.org/10.1016/j.geothermics.2021.102262

Johnson, T. C., Burghardt, J., Strickland, C., Knox, H., Vermeul, V., White, M., et al. (2021). 4D proxy imaging of fracture dilation and stress shadowing using electrical resistivity tomography during high pressure injections into a dense rock formation. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB022298. https://doi.org/10.1029/2021JB022298

Laloy, E., Hérault, R., Jacques, D., & Linde, N. (2018). Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, 54, 381-406. https://doi.org/10.1002/ 2017WR022148
Laloy, E., Hérault, R., Lee, J., Jacques, D., & Linde, N. (2017). Inversion using a new low-dimensional representation of complex binary geo- logical media based on a deep neural network. *Advances in Water Resources*, 110, 387-405. https://doi.org/10.1016/j.advwatres.2017.09.029
Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., & Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. *Water Resources Research*, 49(5), 2664-2682. https://doi.org/10.1002/wrcr.20226
Li, W., & Cirpka, O. A. (2006). Efficient geostatistical inverse methods for structured and unstructured grids. *Water Resources Research*, 42, W06402. https://doi.org/10.1029/2005WR004668
Liu, M., & Grana, D. (2020). Petrophysical characterization of deep saline aquifers for CO_2 storage using ensemble smoother and deep convolutional autoencoder. *Advances in Water Resources*, *142*, 103634. https://doi.org/10.1016/j.advwatres.2020.103634
Liu, Y., & Durlofsky, L. J. (2020). 3D CNN-PCA: A deep-learning-based parameterization for complex geomodels. *Computers & Geosciences*, 148, 104676. https://doi.org/10.1016/j.cageo.2020.104676
Marzouk, Y. M., & Najm, H. N. (2009). Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics*, 228(6), 1862-1902. https://doi.org/10.1016/j.jcp.2008.11.024
Mo, S., Zabaras, N. J., Shi, X., & Wu, J. (2020). Integration of adversarial autoencoders with residual dense convolutional networks for estimation of non-Gaussian hydraulic conductivities. *Water Resources Research*, 56, e2019WR026082. https://doi.org/10. 1029/2019WR026082
Murray, A. B. (2007). Reducing model complexity for explanation and prediction. *Geomorphology*, 90, 178-191. https://doi.org/10.1016/j.geomorph.2006.10.020
Nejadi, S., Trivedi, J. J., & Leung, J. (2017). History matching and uncertainty quantification of discrete fracture network models in fractured reservoirs. *Journal of Petroleum Science and Engineering*, 152, 21-32. https://doi.org/10.1016/j.petrol.2017.01.048
Okoroafor, E. R., Co, C., & Horne, R. N. (2022). Numerical investigation of the impact of fracture aperture anisotropy on EGS thermal performance. *Geothermics*, 100, 102354. https://doi.org/10.1016/j.geothermics.2022.102354
Remy, N., Boucher, A., & Wu, J. (2009). Applied geostatistics with SGeMS: A user's guide. Cambridge University Press, Cambridge, U. K.
Romary, T. (2009). Integrating production data under uncertainty by parallel interacting Markov chains on a reduced dimensional space. *Computational Geosciences*, 13, 103-122. https://doi.org/10.1007/s10596-008-9108-8
Sarma, P., Durlofsky, L. J., & Aziz, K. (2008). Kernel principal component analysis for efficient differentiable parameterization of multipoint geostatistics. *Mathematical Geosciences*, 40, 3–32. https://doi.org/10.1007/s11004-007-9131-7
Settgast, R. R., Fu, P., Walsh, S. D. C., White, J. A., Annavarapu, C., & Ryerson, F. J. (2017). A fully coupled method for massively parallel simulation of hydraulically driven fractures

in 3-dimensions. *International Journal for Numerical and Analytical Methods in Geomechanics*, 41, 627–653. https://doi.org/10.1002/nag.2557

Shi, Z., Zhang, S., Yan, R., & Wang, G. (2018). Fault zone permeability decrease following large earthquakes in a hydrothermal system. *Geophysical Research Letters*, 45, 1387-1394. https://doi.org/10.1002/2017GL075821

Somogyvári, Y., Jalali, M., Jimenez Parras, S., Bayer, P. (2017). Synthetic fracture network characterization with transdimensional inversion. *Water Resources Research*, *53*(*6*), 5104-5123. https://doi.org/10.1002/2016WR020293

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical geology*, 34(1), 1-21. https://doi.org/10.1023/A:1014009426274

Tang, H., Fu, P., Sherman, C. S., Zhang, J., Ju, X., Hamon, F., et al. (2021). A deep learning-accelerated data assimilation and forecasting workflow for commercial-scale geologic carbon storage. *International Journal of Greenhouse Gas Control*, 112, 103488. https://doi.org/10.1016/j.ijggc.2021.103488

Vo, H. X., & Durlofsky, L. J. (2014). A new differentiable parameterization based on principal component analysis for the low-dimensional rep- resentation of complex geological models. *Mathematical Geosciences*, 46(7), 775-813. https://doi.org/10.1007/s11004-014-9541-2

Vogler, D., Settgast, R. R., Annavarapu, C., Madonna, C., Bayer, P., & Amann, F. (2018). Experiments and simulations of fully hydro-mechanically coupled response of rough fractures exposed to high-pressure fluid injection. *Journal of Geophysical Research: Solid Earth*, 123, 1186-1200. https://doi.org/10.1002/2017JB015057

Vogt, C., Marquart, G., Kosack, C., Wolf, A., & Clauser, C. (2012). Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter. *Water Resources Research*, 48, W08517. https://doi. org/10.1029/2011WR011673

Wu, H., Fu, P., Morris, J. P., Mattson, E. D., Neupane, G., Smith, M. M., et al., & EGS. (2021). Characterization of flow and transport in a fracture network at the EGS Collab field experiment through stochastic modeling of tracer recovery. *Journal of Hydrology*, 593, 125888. https://doi. org/10.1016/j.jhydrol.2020.125888

Wu, H., Fu, P., Hawkins, A. J., Tang, H., & Morris, J. P. (2021). Predicting thermal performance of an enhanced geothermal system from tracer tests in a data assimilation framework. *Water Resources Research*, 57, e2021WR030987. https://doi.org/10.1029/2021WR030987

Wu, H., Fu, P., Yang, X., Morris, J. P., Johnson, T. C., Settgast, R. R., & Ryerson, F. J. (2019). Accurate imaging of hydraulic fractures using templated electrical resistivity tomography. *Geothermics*, 81, 74-87. https://doi.org/10.1016/j.geothermics.2019.04.004

Xiao, C., & Tian, L. (2020). Surrogate-based joint estimation of subsurface geological and relative permeability parameters for high-dimensional inverse problem by use of smooth local parameterization. *Water Resources Research*, 56, e2019WR025366. https://doi.org/ 10.1029/2019WR025366

Xiao, C., Zhang, S., Ma, X., Jin, J., & Zhou, T. (2022). Model-reduced adjoint-based inversion using deep- learning: Example of geological carbon sequestration modeling. *Water Resources Research*, 58, e2021WR031041. https://doi.org/10.1029/2021WR031041

Yang, H., Lin, Y., Wohlberg, B., & Tartakovsky, D. M. (2021). Consensus equilibrium for subsurface delineation. *Water Resources Research*, 57(10), e2021WR030151. https://doi.org/10.1029/2021WR030151

Zhang, J., Zheng, Q., Chen, D., Wu, L., & Zeng, L. (2020). Surrogate-based bayesian inverse modeling of the hydrological system: An adaptive approach considering surrogate approximation error. *Water Resources Research*, 56, e2019WR025721. https://doi.org/10.1029/2019WR025721

Zhao, Y., & Luo, J. (2020). Reformulation of Bayesian geostatistical approach on principal components. *Water Resources Research*, 55, e2019WR026732. https://doi.org/10.1029/2019WR026732

Zhu, Y., & Zabaras, N. (2018). Bayesian deep convolutional encoder-decoder networks for surrogate modeling and uncertainty quantification. *Journal of Computational Physics*, 366, 415-447. https://doi.org/10.1016/j.jcp.2018.04.018

AGU PUBLICATIONS

Water Resources Research

Supporting Information for

Selecting appropriate model complexity: An example of tracer inversion for thermal prediction in enhanced geothermal systems

Hui Wu[1, 2], Zhijun Jin[1, 2, 3], Su Jiang[4], Hewei Tang[5], Joseph P. Morris[5], Jinjiang Zhang[1], Bo Zhang[1]

[1] School of Earth and Space Sciences, Peking University, Beijing, China.

[2] Institute of Energy, Peking University, Beijing, China

[3] Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 100083, China

[4] Department of Energy Resources Engineering, Stanford University, Stanford, CA, USA.

[5] Lawrence Livermore National Laboratory, Livermore, CA, USA.

Contents of this file

Figures S1 to S3

Introduction

This supporting information provides the box plots of prior and posterior latent parameters in Fig. S1. Fig. S2 shows two randomly selected posterior realizations in terms of aperture distribution and flow field. Fig. S3 shows the standard deviation of aperture distribution/flow field in posterior ensemble.

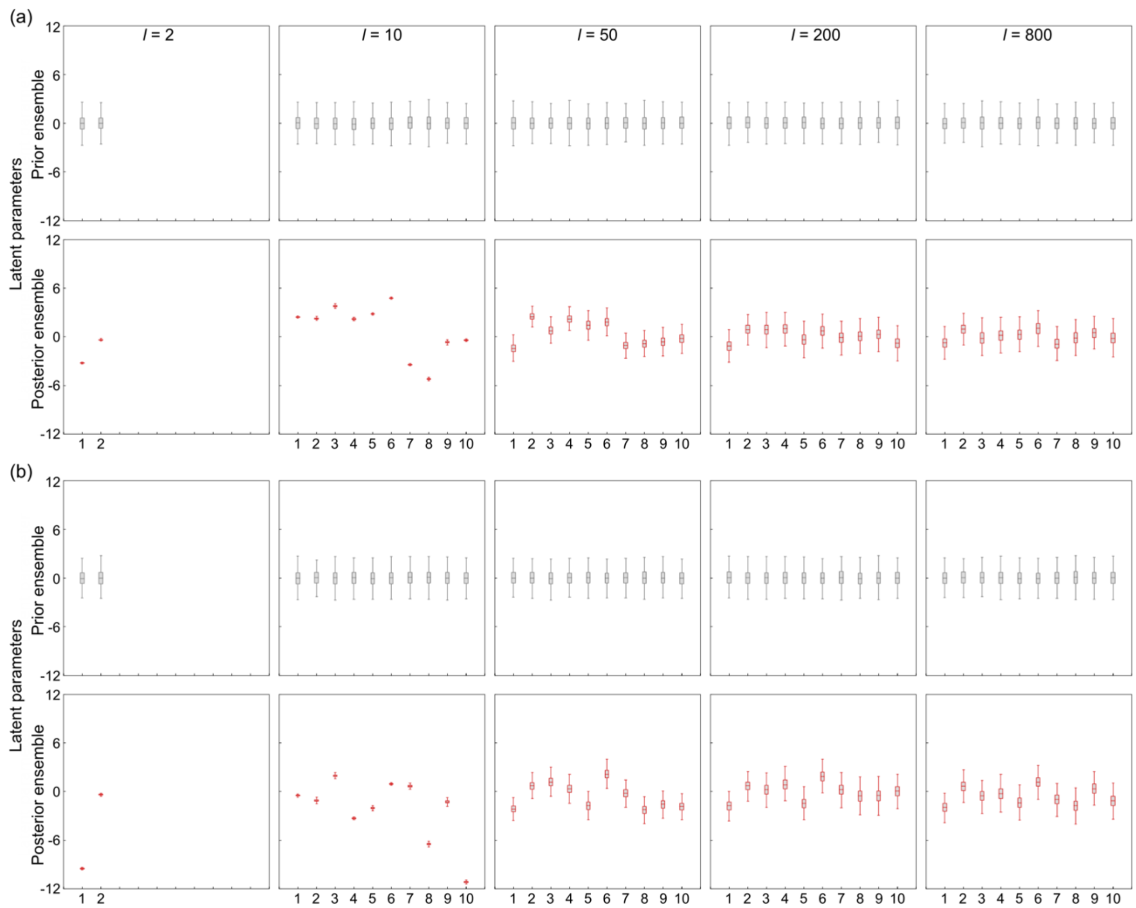


Figure S1. Box plots of ten latent parameters from prior (upper row) and posterior (lower row) ensembles. The ten latent parameters correspond to the ten most significant principal components after PCA. (a) Log-normal aperture field scenario. (b) Two facies aperture field scenario. For the latent space with a dimensionality (*l*) of two, only two latent parameters are shown.

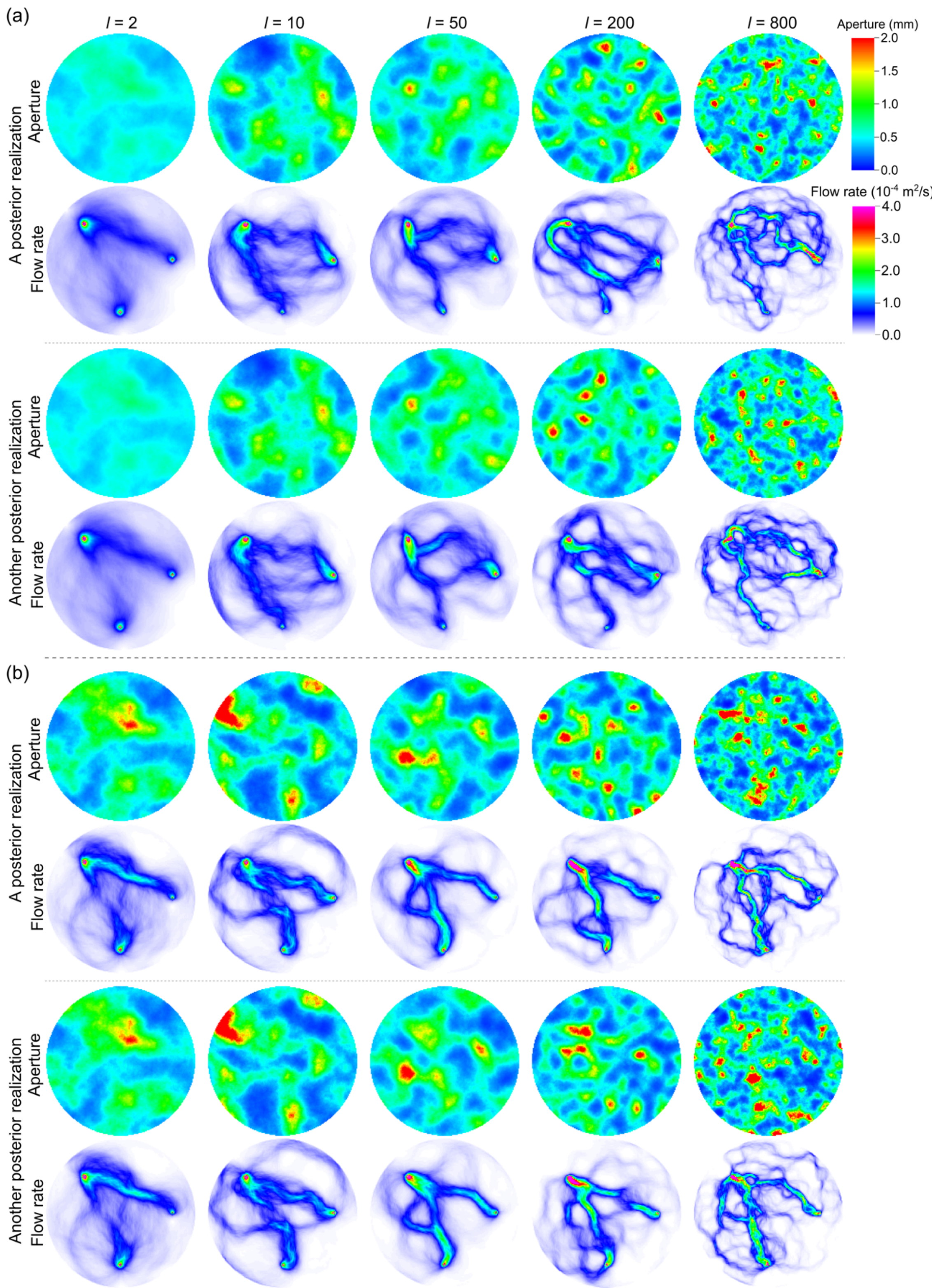


Figure S2. Aperture distribution and flow field from two randomly selected posterior realizations. (a) Log-normal aperture field scenario. (b) Two facies aperture field scenario.

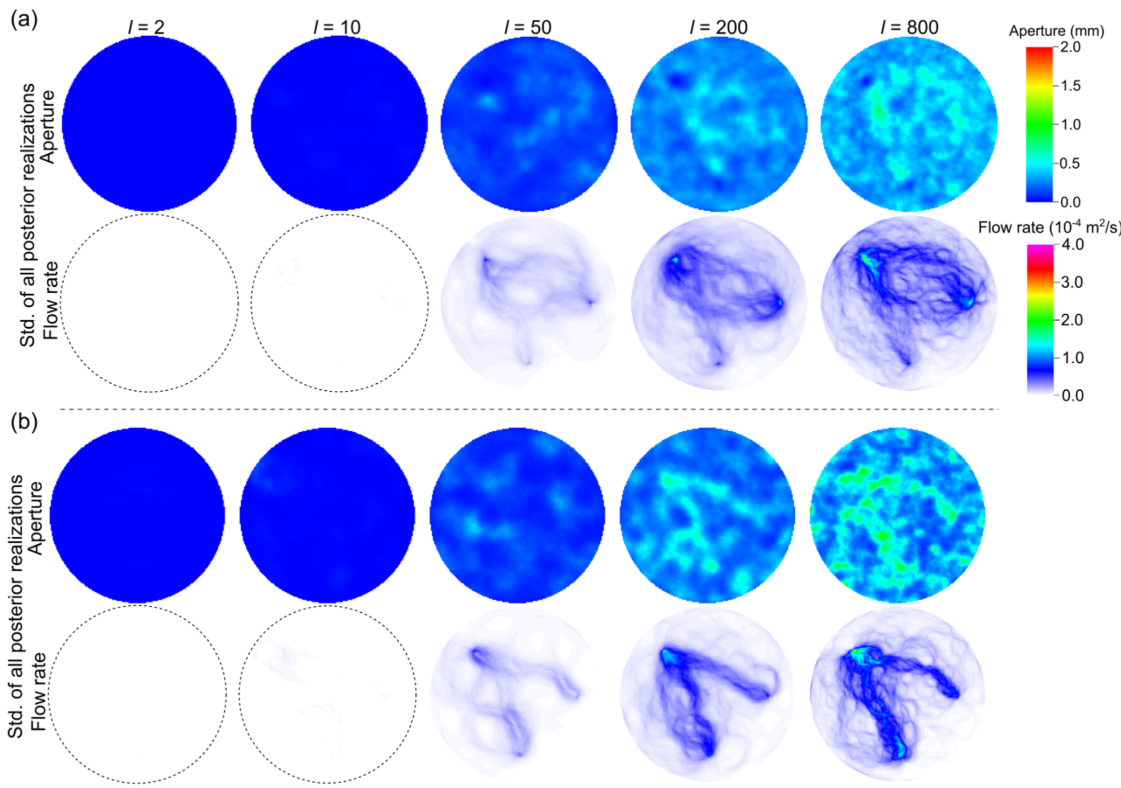


Figure S3. Standard deviation (Std.) of aperture distribution and flow field in the posterior ensemble. (a) Log-normal aperture field scenario. (b) Two facies aperture field scenario.