Fast and robust scheme for uncertainty quantification in naturally fractured reservoirs

de Hoop, Stephan; Voskov, Denis

DOI
10.2118/203968-MS

Publication date
2021

Document Version
Final published version

Published in
Society of Petroleum Engineers - SPE Reservoir Simulation Conference 2021, RSC 2021

SPE-203968-MS

Fast and Robust Scheme for Uncertainty Quantification in Naturally Fractured Reservoirs

Stephan de Hoop, Delft University of Technology; Denis Voskov, Delft University of Technology, Stanford University

This paper was prepared for presentation at the SPE Reservoir Simulation Conference, available on-demand, 26 October 2021 – 25 January 2022. The official proceedings were published online 19 October 2021.

Abstract

The main objective of this study is to perform Uncertainty Quantification (UQ) using a detailed representation of fractured reservoirs. This is achieved by creating a simplified representation of the fracture network while preserving the main characteristics of the high-fidelity model. We include information at different scales in the UQ workflow which allows for a large reduction in the computational time while converging to the high-fidelity full ensemble statistics.

Ultimately, it allows us to make accurate predictions on geothermal energy production in highly heterogeneous fractured porous media. The numerical reservoir simulation tool we use in this work is the Delft Advanced Research Terra Simulator (DARTS). It is based on Finite Volume approximation in space, fully coupled explicit approximation in time, and uses the novel linearization technique called Operator-Based Linearization (OBL) for the system of discretized nonlinear governing equations. We use a fracture network generation algorithm that uses distributions for length, angles, size of fracture sets, and connectivity as its main input. This allows us to generate a large number of complex fracture networks at the reservoir scale. We developed a pre-processing algorithm to simplify the fracture network and use graph theory to analyze the connectivity before and after pre-processing. Furthermore, we use metric space modeling methods for statistical analysis.

A robust coarsening method for the Discrete Fracture Matrix model (DFM) is developed which allows for great control over the degree of simplification of the network's topology and connectivity. We apply the framework to modeling of geothermal energy extraction. The pre-processing algorithm allows for a hierarchical representation of the fracture network, which in turn is utilized in the reduced UQ methodology. The reduced UQ workflow uses the coarser representation of the fracture networks to partition/rank the high-fidelity parameter space. Then a small subset of high-fidelity models is chosen to represent the full ensemble statistics. Hereby, the computational time of the UQ is reduced by two/three orders of magnitude, while converging to similar statistics as the high-fidelity full ensemble statistics.

The methods developed in this study are part of a larger project on a prediction of energy production from carboniferous carbonates. The final goal is to perform UQ in pre-salt reservoirs which are characterized by complex reservoir architecture (i.e., large karstification and fracture networks). The UQ of fractured reservoirs is typically done in the petroleum industry using effective media models. We present a

methodology that can efficiently handle a large ensemble of DFM models, which represent complex fracture networks and allow for making decisions under uncertainty using more detailed high-resolution numerical models.

Introduction

Society heavily depends on the availability of energy (e.g., industry and agriculture, logistics, and households) and raw materials (e.g., plastic, medicine, tires, and solvents). Carbonate reservoirs host a major part of the world's hydrocarbon reserves, approximately 60% of the oil and 40% of the gas reserves (Akbar et al., 2000). This percentage is likely to increase due to several large discoveries made in the past decade, such as pre-salt carbonates in off-shore Brazil (Mello et al., 2011; Boyd et al., 2015) and the Tarim basin in China (Huang et al., 2017). The ongoing energy transition has also resulted in an increase in geothermal discoveries, for example, the geothermal potential in Dutch onshore Carboniferous carbonates (Reijmer et al., 2017).

Developments of any project related to the subsurface are often associated with large uncertainty and risks (Caers, 2011). The main reason for this is the lack of information on some of the modeling parameters, such as, in the case of natural fracture networks, the spatial distribution of matrix permeability and fracture aperture. These parameters largely control the resulting fluid-flow and geomechanical behavior of the porous media, hence greatly impact any decision making related to these subsurface activities.

A popular remedy for the lack of information is to assume, based on the available data, some (prior) distribution of the modeling parameters and generate a large set of models. This set of models, generally denoted as an ensemble, is subjected to the modeling effort and subsequently attempts to explain the uncertainty in the flow- or geomechanical response. This class of uncertainty quantification (UQ) methods belongs to the Monte Carlo (MC) methods (Hammersley, 2013). MC methods often need a large amount of random samples to converge and, in the case of large fractured reservoirs, this can lead to an insurmountable computational effort.

In response, the ensemble of fractured reservoir models is regularly upscaled to effective media models (e.g., dual-porosity or Embedded Discrete Fracture Models) to reduce the computational time and still attempt to make meaningful predictions. However, in the case of realistically complicated fracture orientation and connectivity, these models might not actually reflect reality. Therefore, the use of a Discrete Fracture Matrix (DFM) model, which explicitly represents the fracture geometry, is preferred (Moinfar et al., 2011).

Since the mesh is conformal to the fractures in DFM models, the main problem, especially in the case of complex realistic fracture networks, is the gridding step. This is particularly problematic whenever complex fracture intersections are present. The resulting mesh will often contain several artifacts: non-orthogonal control volume connections, control volumes with a large aspect ratio (i.e., flat triangles), and a massive difference in the size of the control volumes. This will eventually diminish the quality of the resulting numerical solution and also lead to poor nonlinear convergence of the simulation process. Therefore, we have developed a pre-processing algorithm, which borrows concepts from early work done is this area (Koudina et al., 1998; Maryvska et al., 2005) as well as more recent approaches (Mustapha and Mustapha, 2007; Mallison et al., 2010; Karimi-Fard and Durlofsky, 2016). The main idea is to simplify the fracture network such that the meshing artifacts are reduced while keeping the main characteristics of the fracture network intact. This allows for a reliable and fast way of constructing DFM models on any desired scale.

Finally, we present a methodology that can efficiently handle a large ensemble of DFM models, which represent complex fracture networks and allow for making decisions under uncertainty using more detailed high-resolution numerical models. This approach is similar to those in Scheidt and Caers (2009); Scheidt et al. (2011); de Hoop et al. (2018), however, here we utilize information of coarse-scale models, obtained via our efficient and robust pre-processing procedure, to rank and partition the high-fidelity parameter space.

Subsequently, a small subset of high-fidelity models is chosen to represent the full ensemble statistics. Hereby, the computational time of the UQ is reduced by several orders of magnitude, while the reduced approach is converging to similar statistics as the full high-fidelity ensemble of models.

Method and theory

In this section all the relevant theory is briefly described.

Generation fracture networks

In this section, we present a generic framework which allows us to generate various types of fracture networks with different properties (e.g., connectivity, orientation, length, etc.). Fractures, as observed in outcrops, typically occur in sets with characteristic angles of intersection, connectivity, and spacing (Bour and Davy, 1997; Manzocchi, 2002; Bakay et al., 2016). Therefore, we propose to generate the fractures as sets which are parameterized by the number of fractures in the set, lengths, and the angles at which they intersect. Connectivity of the final network, expressed as the total length of the largest connected sub-network divided by the total length of all the fractures in the domain, is calculated and adjusted (i.e., fracture sets are connected or disconnected) based on the target connectivity.

Fracture networks themselves are only a means in any modeling workflow and (almost) never constitute the final objective. Usually the objective is to make predictions on a subsurface system, which in turn can contain a fracture network. Our proposed algorithm considers the discretization accuracy, denoted as l_c, at which the final high-fidelity subsurface model will be constructed from the start of the generation. Fractures that belong to the same fracture set will not intersect at a scale or have a spacing that is smaller than l_c. This results in a uniform discretization of each fracture set and minimizes the numerical artifacts which are often observed in Discrete Fracture Network (DFM) models (Berre et al., 2018). Interaction of different fracture sets can still result in a complex fracture network where intersections occur below the discretization accuracy, hence the need for a robust pre-processing method as explained below in section Pre-processing fracture network.

The algorithm sequentially populates the domain with fractures sets, where the characteristics of the fracture sets are sampled from a specified distribution. For example, the length of the fractures is repeatedly described by a power-law distribution, but you could use any distribution that models your input data (outcrop analogue, seismic, or borehole data). The length of each fracture in the domain is some (random) multiple of the discretization accuracy l_c. Before adding a fracture to the domain, it is split into n_i^s segments of length l_c, where n_i^s is the number of smaller fracture segments of fracture i. Additional fractures of the same set as fracture i always intersect at the nodes of the smaller segments, hence a uniform distribution of control volumes is expected after meshing the fracture network.

The following parameters are input to the generation framework:

- Connectivity, C;
- Number of fractures pairs, N_f;
- Maximum number of fracture pairs per set, $N_{f,s}$;
- Set of length distributions, L_f ;
- Set of angle distributions, Θ_f.

Since the number of fractures and maximum number of fractures per set are highly correlated with the total connectivity of the network, you can choose to express those parameters as a function of connectivity. For example, $N_f = N_f^0\sqrt{C}$, where N_f^0 is the some base number of fractures (e.g., at maximum connectivity).

Also, as could be inferred from the parameters above, the fractures are always added in pairs within each fracture set. This means that there are two length and angle distributions per fracture set. We designed a detailed algorithm for sequential generation of fracture network following the approach suggested in Sartori Suarez (2018).

After the algorithm has finished and N_f fracture pairs are added to the domain, the connectivity is calculated using graph theory. First construct the unweighted undirected graph $G = (E,V)$, where V is the set of vertices of the graph and E is the set of paired vertices called edges. Based on these two sets we can construct the following matrices:

- Incidence matrix: $B(G)$, which is a $n \times m$ matrix, where n is the number of vertices and m the number of edges in the graph. $B_{ij} = 1$ if vertex i is on edge j otherwise $B_{ij} = 0$;
- Degree matrix: $D(G)$, which is a $n \times n$ matrix describing the number edges attached to each vertex. The degree matrix can be obtained using the follow equation $D = \text{diag}\,(B\mathbf{1})$, where diag($\mathbf{v}$) is a function that constructs a square matrix with vector $\mathbf{v}$ on its diagonal and $\mathbf{1}$ is a vector of ones with size $m \times 1$;
- Discrete Laplacian matrix: $L(G)$ which can be found via the following equation $L = D - BB^T$.

The Laplacian of the graph can be used for finding spanning trees of a given graph (i.e., connected fracture sets in the fracture network). Particularly, each element of the rational basis of the null-space of the Laplacian describes a connected component of the graph (Spielman, 2010). With this basis we can find the number of connected fracture sets in our network but also the connectivity of the network (as defined above). If this connectivity is larger than the specified connectivity, fracture sets are disconnected until the actual connectivity reaches the desired connectivity (within some threshold). The reverse (i.e., connecting fracture sets) is done when the specified connectivity is lower than the desired connectivity. Components of the graph are sequentially merged or disconnected based on this smallest distance between each vertex in the component of the graph and each vertex in the fracture sets.

Pre-processing fracture network

All the fluid flow evaluations are done with the Delft Advanced Research Terra Simulator (DARTS), recently developed at the Delft University of Technology (Khait and Voskov, 2017, 2018a,b). DARTS is finite volume based simulator capable of solving thermal flow and transport problem in complex and realistic fractured reservoirs by using the DFM model adapted from Karimi-Fard et al. (2004). Current DFM implementation in DARTS is based on a Two-Point Flux Approximation (TPFA) scheme where fractures are represented explicitly in the numerical model. However, explicit representation of the fractures, especially in fracture networks generated from outcrops, often contain complex fracture intersection which in turn might pose meshing difficulties. The resulting mesh, therefore, can contain abundant artefacts which negatively impact the performance of the reservoir simulator.

Meshing difficulties related to fracture networks are thoroughly explained in Koudina et al. (1998); Mustapha and Mustapha (2007); Mallison et al. (2010); Karimi-Fard and Durlofsky (2016); Berre et al. (2018). All of the meshing challenges are related to scale. If we choose a discretization scale which is sufficiently small, any complex fracture interaction can be accurately represented. However, this is often not feasible especially when simulating a large ensemble of models required for uncertainty quantification.

Most of the existing fracture network processing methods attempt to achieve the same, particularly to increase the orthogonality of control volume connections and reduce the number of extremely small control volumes. Subsequently this improves the accuracy of the numerical solution. All processing should be achieved while minimizing the changes of the configuration and topology of the actual fracture network input data.

As mentioned in the introduction, this study is greatly inspired by previous efforts in this area (e.g. Karimi-Fard and Durlofsky, 2016). Besides the merging of fracture nodes that violate some algebraic constrain based on l_c, we also identify the angles at which fractures intersect and correct any intersection that falls below a certain threshold value (θ_{min}). Particularly, we perform the following steps to obtain a suitable fracture network for DFM simulations:

1. Find all intersections of all the fractures in the domain;
2. Sequentially discretize each fracture (from largest to smallest fracture) and check algebraic constraint (related to l_c);
3. Straighten fractures elements of node degree 2-2 or 2-1 if their angle is below a tolerance angle (θ_{min});
4. Solve remaining and/or created low angle intersections (i.e., intersection that occur at an angle below θ_{min}.

Governing equations of the physical systems

This section briefly covers the governing equations of the physical system that was used in the uncertainty quantification workflow, particularly the equations describing geothermal energy production.

Continuous form. The conservation of mass, in general form, is written as

$$\frac{\partial}{\partial t}\left(\phi\sum_{p=1}^{np} x_{cp}\rho_p s_p\right)+\nabla\cdot\sum_{p=1}^{np} x_{cp}\rho_p \mathbf{v}_p+\sum_{p=1}^{np} x_{cp}\rho_p q_p=0,\quad c=1,\ \ldots,\ n_c \tag{1}$$

where φ represents the porosity, x_{cp} is the molar mass fraction of component c in phase p, ρ_p is the density, s_p is the saturation, and q_p is the source term of the p-th phase respectively, and $\mathbf{v}_p$ is the velocity of the p-th phase.

The Darcy velocity of the p-th phase is given by

$$\mathbf{v}_p=-\frac{k_{r,p}}{\mu_p}\mathbf{K}\nabla\left(p_p-\rho_p\mathbf{g}\right),\quad p\in\{o,w\} \tag{2}$$

where $k_{r,p}$ is the relative permeability, μ_p is the viscosity and p_p is the pressure of the p-th phase respectively, $\mathbf{K}$ is the permeability tensor and $\mathbf{g}$ is the directional gravitational acceleration defined as $g\#z$. The conservation of energy, required for the geothermal simulations, is described by the following equation

$$\frac{\partial}{\partial t}\left(\phi\sum_{p=1}^{np}\rho_p s_p U_p+\left(1-\phi\right)U_r\right)+\mathrm{div}\sum_{p=1}^{np} h_p\rho_p\mathbf{v}_p+\mathrm{div}\left(\kappa\nabla T\right)+\sum_{p=1}^{np} h_p\rho_p q_p=0, \tag{3}$$

where U_p is the internal energy of fluid phase p, U_r is the rock internal energy, h_p is the enthalpy of phase p, κ is the thermal conduction, and T is the temperature.

Discretization and Operator-Based Linearization. We apply a finite-volume discretization on a general unstructured grid (using a TPFA for the fluxes across interfaces with upstream weighting) and a backward (implicit) Euler time discretization strategy to both the conservation equations and obtain the following system of equations (assuming no gravity and capillarity)

$$\begin{aligned} V\left[\left(\phi\sum_{p=1}^{np} x_{cp}\rho_p s_p\right)^{n+1}-\left(\phi\sum_{p=1}^{np} x_{cp}\rho_p s_p\right)^{n}\right]-\Delta t\sum_{l}\left(\sum_{p=1}^{np} x_{cp}^{l}\rho_p^{l}\Gamma_p^{l}\Delta p^{l}\right) \\ +\Delta t\sum_{p=1}^{np}\rho_p x_{cp} q_p=0,\quad c=1,\ \ldots,\ n_c \end{aligned}, \tag{4}$$

and

$$V\left[\left(\phi\sum_{p=1}^{np}\rho_p s_p U_p+\left(1-\phi\right)U_r\right)^{n+1}-\left(\phi\sum_{p=1}^{np}\rho_p s_p U_p+\left(1-\phi\right)U_r\right)^{n}\right] - \Delta t\sum_{l}\left(\sum_{p=1}^{np}h_p^l\rho_p^l\Gamma_p^l\Delta p^l+\Gamma_c^l\Delta T^l\right)+\Delta t\sum_{p=1}^{np}h_p\rho_p q_p=0, \tag{5}$$

where Γ_p^l is the convective and Γ_c^l is the thermal transmissibility of interface l and phase p respectively.

For the linearization of the above system of nonlinear equations we use the Operator-Based Linearization (OBL). In this approach, the discrete form of the mathematical equations are grouped into the state dependent operators. These operators are subsequently calculated, exactly, in a set of supporting points. The solution, or rather the value of the operator, for a particular state, is then multi-linearly interpolated inside a particular hypercube in the parameter-space. The multi-linear interpolation means simple, exact, and above all flexible (partial) derivatives for the nonlinear solution procedure. For details on the OBL framework the reader is referred to Voskov (2017); Khait and Voskov (2017, 2018a).

Uncertainty quantification framework

The uncertainty quantification workflow utilizes a distance-based method and ranking procedure similar to those proposed in Scheidt and Caers (2009); Scheidt et al. (2011); de Hoop et al. (2018). Distance-based methods attempt to reduce the dimensionality and complexity of the modeling task. They can be extremely effective for visualizing the variability of an ensemble of models, which is often high dimensional. This dimensionality reduction is achieved by projecting onto a (much) lower-dimensional space. This is especially efficient when a distance, or dissimilarity measure, is used that has a high correlation with the modeling objective (e.g., predicting the distribution of total energy produced of the ensemble) (Caers, 2011). Subsequently, the global and local structure of the response uncertainty can be visually and quantitatively analyzed in this reduced space.

A commonly used approach for dimensionality reduction is the Multidimensional Scaling (MDS) algorithm (Borg and Groenen, 2005). In MDS, the square symmetric distance matrix **D**, obtained by taking the distances of a particular property between each ensemble member, is projected onto a Cartesian hyperplane (after double centering it around a origin) of dimension $N_{\text{reduced}} = N_{\text{m}} << N_{\text{M}}$, where N_{M} is the size of the ensemble **M**. For the details of the MDS algorithm we refer the reader to Borg and Groenen (2005) chapter 12 and Caers (2011).

The dimension N_{m} of the hyperplane required to accurately replicate the original **D**, depends on the magnitude of the eigenvalues of the decomposition of the distance matrix. Typically N_{m} attains one or two when production data from wells are used in the computation of **D**. Generally in geoscience problems N_{m} can range from one utill six dimension(s) (Caers, 2011). The following simple example shows both the effectiveness of reducing the size as well as detecting possible redundancy in the output space.

Fluid-flow simulations generally have high dimensional input data in the form of complex high-fidelity models. However, decision are usually made on simple statistics, such as the amount of energy produced from a (relatively) small set of wells. This often leads to redundancies in the output space of the simulations. Sadly, due to the nonlinearity of the simulations, it is not necessarily easy to (a priori) predict which models behave similarly. Therefore, using a fast way of detecting these possible redundancies in the output space is the basis of the reduced UQ methodology. If we can predict which subset of models explains the same variability of the full ensemble (i.e., first two stochastic moments), we can greatly reduce the computational time of the UQ process.

In this work we utilize the coarse information, obtained from our pre-processing algorithm, to construct the squared dissimilarity matrix, $\mathbf{D}^{(2)}$, of the full ensemble. Subsequently, we apply the MDS procedure to this matrix and perform clustering based on the projected variability of the coarse representation of the

full ensemble. The medoids of the clusters are the representative subset of models which are simulated on the high-fidelity scale. Finally, three quantiles, particularly the P10, P50, and P90, are used to quantify the mismatch between the full ensemble statistics and the approximate statistics from the subset of models. Precisely, we calculate the mismatch with the following equation (Scheidt et al., 2009)

$$\varepsilon_{N_K} = \frac{1}{3N_t}\sum_{i}^{N_t} \left(\left|P_{10}^{\text{full}}(t_i) - P_{10}^{\text{sub}}(t_i)\right| + \left|P_{50}^{\text{full}}(t_i) - P_{50}^{\text{sub}}(t_i)\right| + \left|P_{90}^{\text{full}}(t_i) - P_{90}^{\text{sub}}(t_i)\right|\right), \tag{6}$$

where N_K represents the number of clusters (i.e., size of the subset of HF models), N_t is the number of time-steps, and t_i represents the i-th time step.

Results

In this section the results of the fracture generation and cleaning are presented followed by the results of the reduced uncertainty quantification workflow.

Fracture generation results

Figure 2 shows three examples of fracture networks generated with the above described algorithm. In these examples the number of fractures and size of the fracture sets is a function of connectivity. The angles for each fracture set are sampled from a random distribution. Characteristic length l_c is chosen substantially small in order to achieve a complex fracture network.

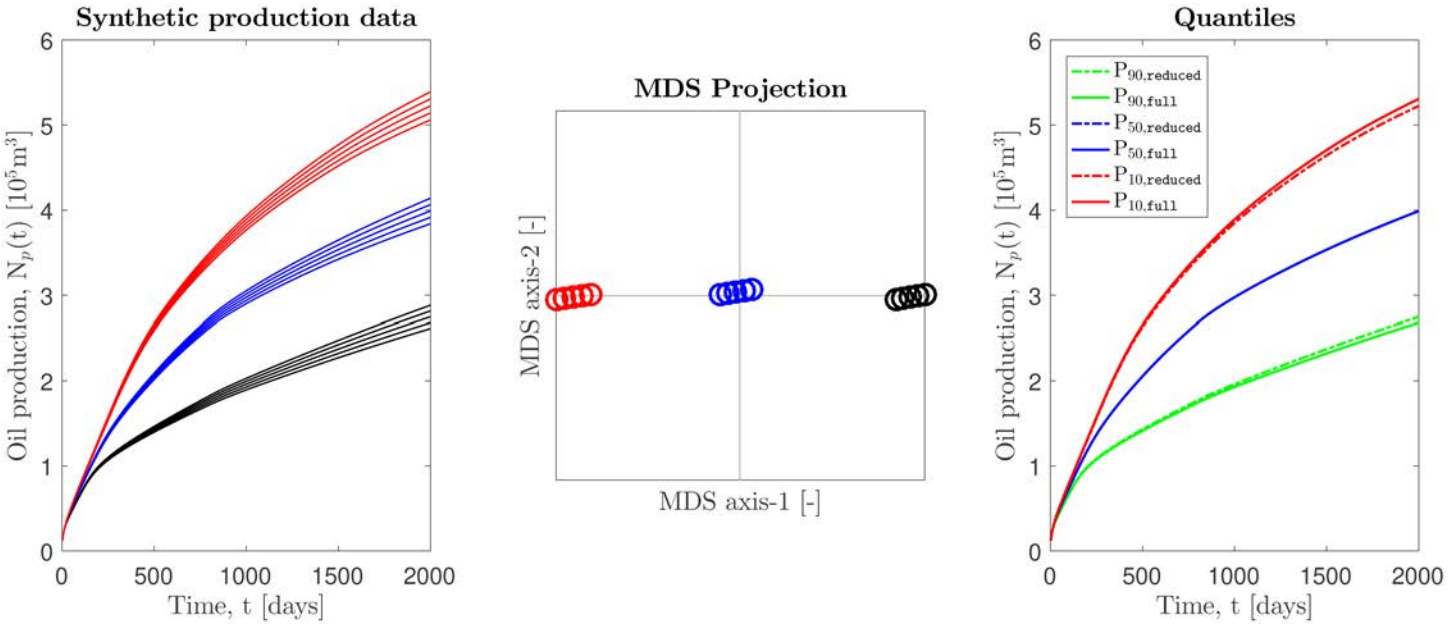

Figure 1—Distances and similarities in the high-dimensional flow-response output space are preserved in the reduced dimensional space (left and middle graph). Clustering will identify similar responses and this subset of models converges to the full ensemble uncertainty characteristics (right graph). Taken from de Hoop et al. (2018).

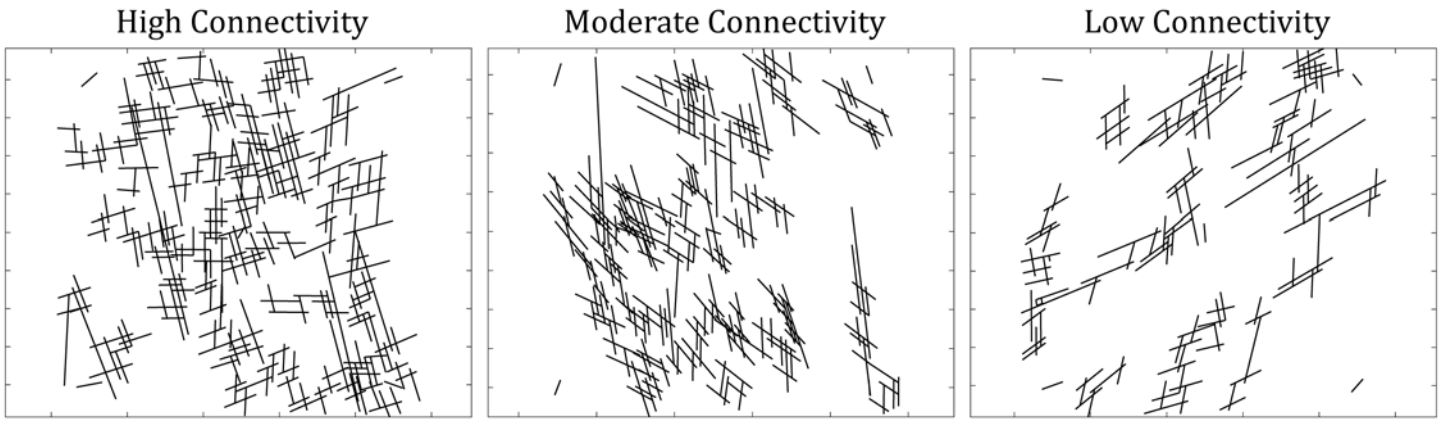

Figure 2—Three examples of the fracture networks with different connectivity, generated using the algorithm in section Generation fracture networks.

An ensemble of fracture networks was generated for the analysis of the pre-processing strategy and subsequent uncertainty quantification. A total number of 600 fracture networks were generated with a varying degree of connectivity (200 low, 200 moderate, and 200 high connectivity fracture networks respectively), number of fractures, and orientation of the fracture sets. The fluid-flow results presented in the next section were performed on the 184 models with the largest connectivity. Moderate to low connectivity results will be included in the future work.

Fracture pre-processing results

The pre-processing algorithm and its performance is first illustrated using a realistic fracture network, generated from outcrop data (Bisdom et al., 2016). The meshes in Figure 3 are obtained using the free software Gmsh (Geuzaine and Remacle, 2009). The figure attempts to illustrate the comparison between the meshing results obtained from the unprocessed fracture input versus the pre-processed fracture network. A large reduction in the degrees of freedom is observed in part (A) while a more uniform volume distribution is observed in part (B).

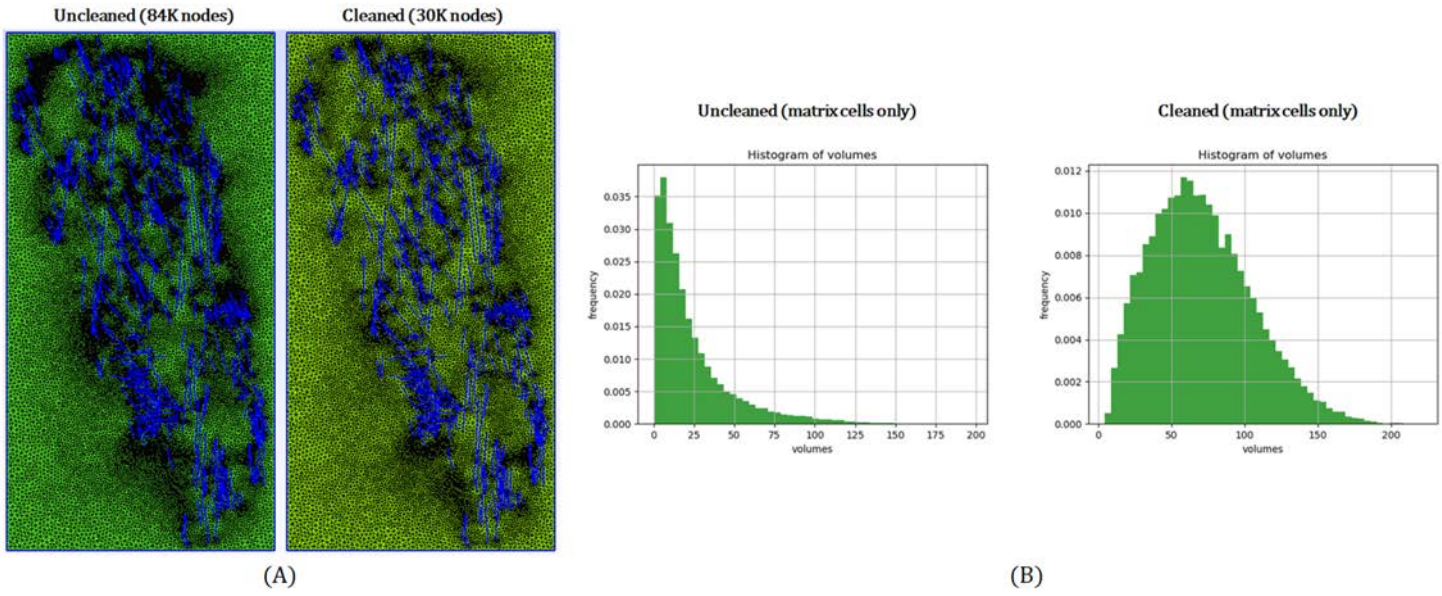

Figure 3—(A) Comparison between meshing results of the uncleaned and cleaned network. (B) Comparison of the control volume size distribution of the uncleaned and cleaned fracture network. Modified from De Hoop et al. (2020)

Figure 4 illustrates the differences between each level of coarsening for a geothermal simulation. The numbers below the temperature fields correspond to the degrees of freedom in fracture and matrix cells respectively. Even though large variations can be observed in the temperature fields, due to the large reduction in spatial resolution, a similar response in terms of temperature at the production well can be observed.

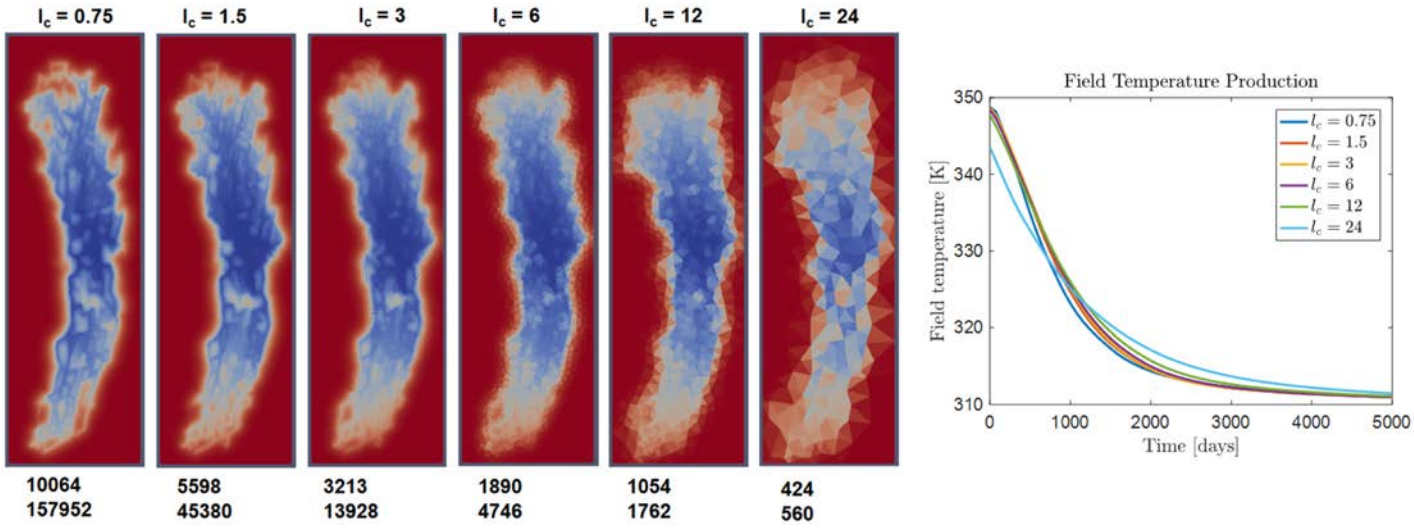

Figure 4—Temperature response of the fracture network at different scales (field on the left and rates on the right). The two numbers below the temperature fields at different scales indicate the number of fracture and matrix cells respectively. Modified from De Hoop et al. (2020)

Besides studying the effect of the pre-processing on a single realizations, it is important to see the effect on stochastic properties. Usually, decisions are made based on an ensemble of models, for example, the ensemble of models generated in section Fracture generation results. It is important to see how the quantiles are changing for each level of coarsening. In order to investigate this, numerical simulations were performed.

For the geothermal simulations, cold water is injected in the bottom left of the domain and water is produced from the top right of the domain. Each well (i.e., doublet system) penetrates a fractures and the distance between each well is roughly 800 meters. The simulation models have no flow boundary conditions on each side of the domain. The injection well is controlled by rate (500 m^3/day) and an injection temperature of 30 degrees Celsius, while the production well is set to bhp control (450 bar). The initial reservoir temperature is 150 degrees Celsius and the initial pressure is 400 bar (i.e., high-enthalpy geothermal system). The matrix permeability is set to 1 mD and the fracture aperture is 1 mm, resulting in a contrast in permeability of roughly 10^8 (K_{frac}/K_{matrix}), resembling an average naturally fractured carbonate system. Figure 5 shows the energy production rate for the 184 high connectivity models. Figure 6 shows how the quantiles are changing for each of the coarsening level.

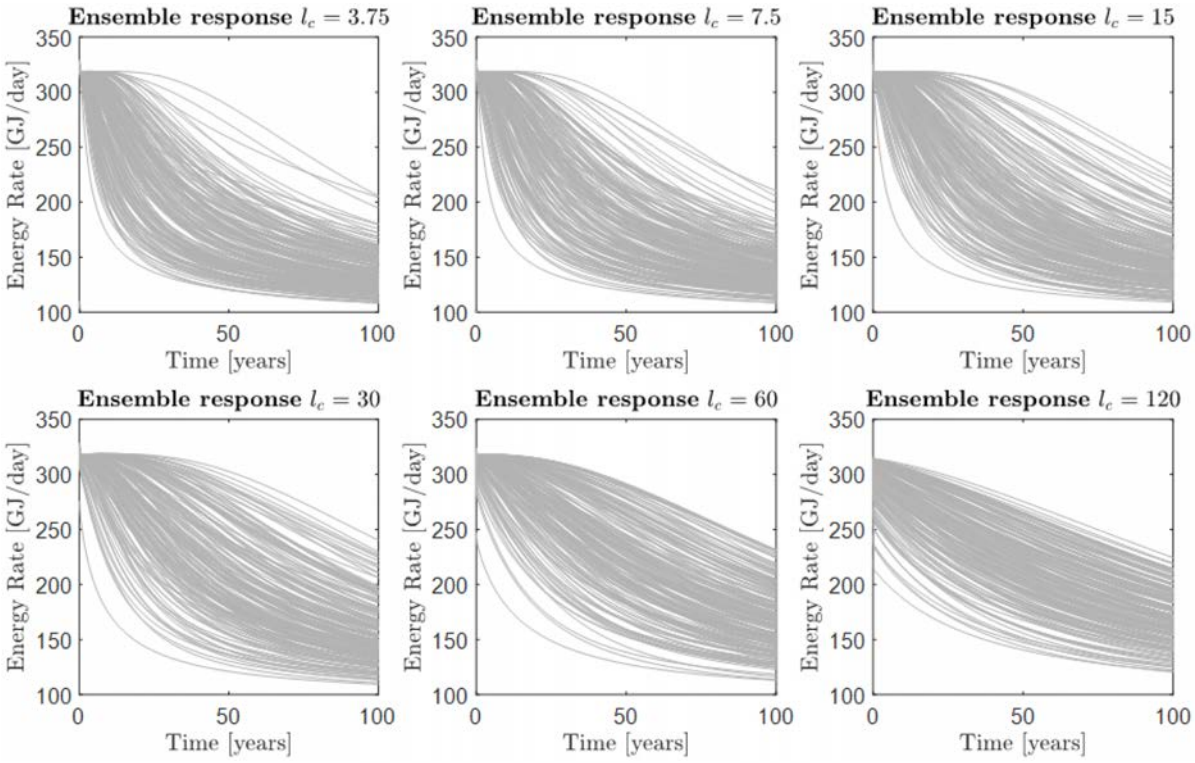

Figure 5—Flow response of the ensemble of fracture networks for each coarsening level.

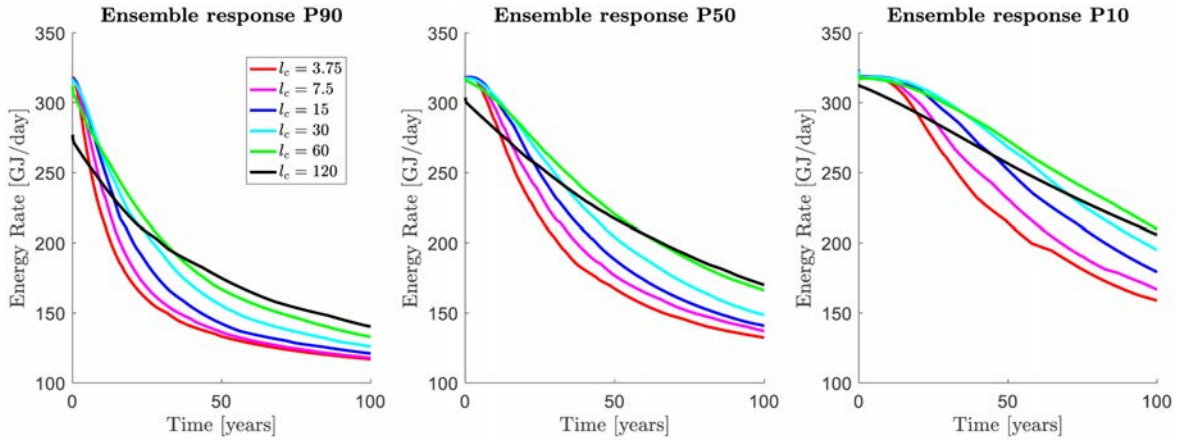

Figure 6—Quantiles of energy production for each ensemble scale. Monotonous deviation of each coarse-scale ensemble from the fine-scale can be observed, as well as an increase in the energy production for each quantile. This is most likely caused by the increase in connectivity when coarsened is increased.

One artifact of the pre-processing algorithm is that disconnected networks become connected when increasing the characteristic length scale l_c (i.e., discretization accuracy). See Figure 7 as an example. This is also observed in the flow-response as an increase in energy production due to a larger area of the reservoir being depleted. Please note that matrix permeability and conductivity are not upscaled in these coarse-scale realizations which also contributes to deviation from fine-scale response. This seems to imply that the low to moderate connectivity models will deviate more from the fine-scale than the higher connectivity models. For the UQ workflow, this would mean that the distance used for the MDS procedure would be less representative of the distance between fine-scale ensemble members. To remedy this, we could keep track to which fracture set each individual fracture belongs in the pre-processing algorithm and ultimately disconnect fractures that were not connected in the unprocessed input data.

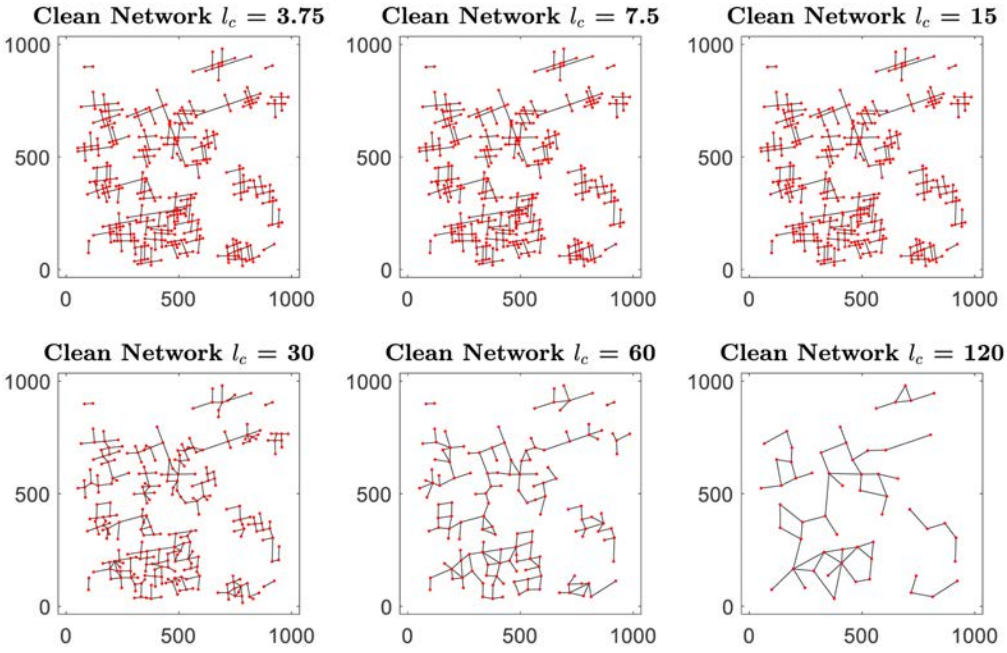

Figure 7—Example of alterations to the fracture network based on desired discretization accuracy l_c. The proposed remedy for ensuring similar flow-behavior in the coarse scale models is to dicsonnect fractures that have become connected due to coarsening.

In Figure 7, the $l_c = 60$ (i.e., $16 - l_c^{\texttt{finest}}$) can be seen as a scale at which the coarsening fails to accurately represent the fine-scale fracture network (in terms of topology, orientation of fractures, and connectivity). The question remains if this scale can still be used for ranking and selecting representative models for the reduced UQ workflow, which will be answered for highly connected networks in the following section. Finally, to further illustrate the performance of the cleaning, Figure 8 shows the connectivity of the pre-processed versus connectivity of the input data for each level of coarsening.

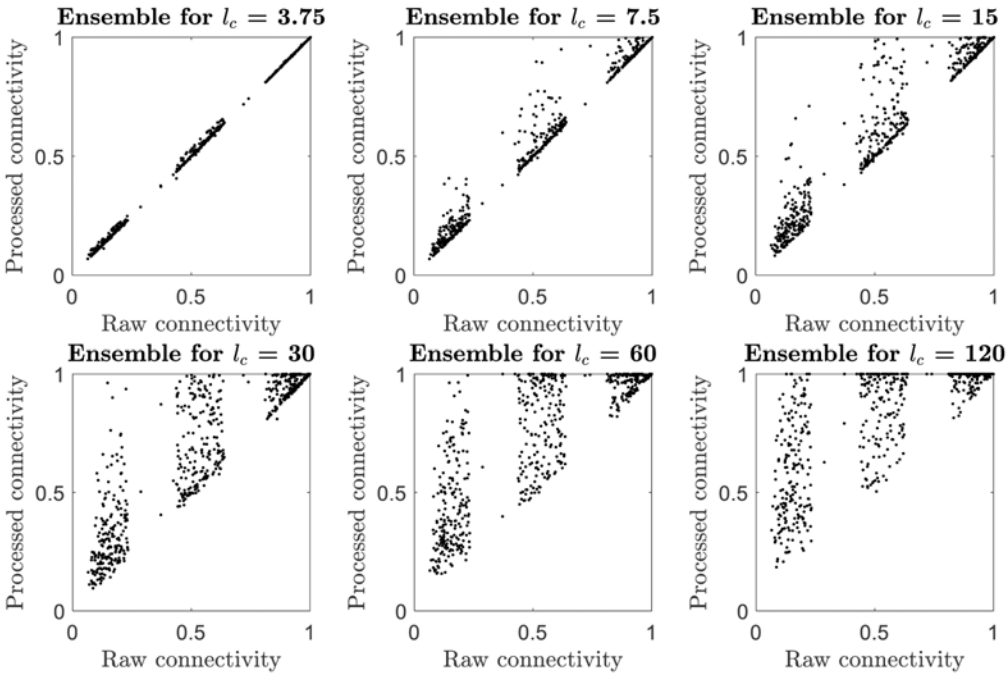

Figure 8—Displaying the behavior of the connectivity of the processed vs. raw input data as a function of coarsening levels.

Uncertainty Quantification results

As mentioned in Uncertainty quantification framework, performing high-fidelity DFM simulations on a large ensemble is often unfeasible and impractical. Instead of using dual-porosity models to speed up the computation, we propose to utilize the more realistic DFM models at a coarser scale to identify representative models. Ultimately, simulating this subset of high-fidelity models should convergence to the same statistics as the full ensemble.

For the reduced UQ workflow in this section, we present the analysis on the 184 highest connectivity models, as mentioned in the previous section. The coarse scale response that we utilize in the example is based on the l_c = 60 model scale. The reason for this is a trade-off between the quality of the distance (higher for lower l_c) and the computational gain of using a larger l_c. Using a lower quality distance for the MDS procedure will likely increase the number of representatives you need to select, since the correlation between the coarse-scale and the fine-scale response is lower for larger l_c. This will increase the overall computational time since more representative models are required (which are simulated on the finest scale). The total simulation time of the reduced UQ, as a function of l_c, is given by

$$t_{\texttt{red.UQ}}(l_c) = N_M \times t_{\texttt{processing}}(l_c) + N_M \times t_{\texttt{sim.coarse}}(l_c) + N_K(l_c) \times t_{\texttt{sim.fine}} \,. \tag{7}$$

The $t_{\text{red.UQ}}$ is expect to be much smaller than $t_{\text{fullUQ}} = N_M \times t_{\text{sim.fine}}$, since $t_{\text{sim.fine}} >> t_{\text{sim.coarse}}$ and $N_M >> N_K$. Computational time for the MDS and clustering procedure is negligible compared to the total time. Processing time is dependent on the l_c as well, but much smaller than $t_{\text{sim.fine}}$. Therefore, the objective is to minimize $t_{\text{red.UQ}}$ while also minimizing ε_{N_K}.

The results of the reduced UQ workflow are depicted in Figure 9. The coarse-scale response is translated to z-scores (subtract mean and normalize by the standard deviation of the energy rate at each time-step). Then the distances are projected into lower dimensional space using MDS and clustering is performed using k-medoids algorithm. The centroids of each cluster is chosen as a representative and comprises the subset of models. These models are then simulated on the fine-scale and for the computation of the cumulative distribution function (CDF) the responses are weighted by the size of the cluster which they present (similar to Scheidt et al. (2009)).

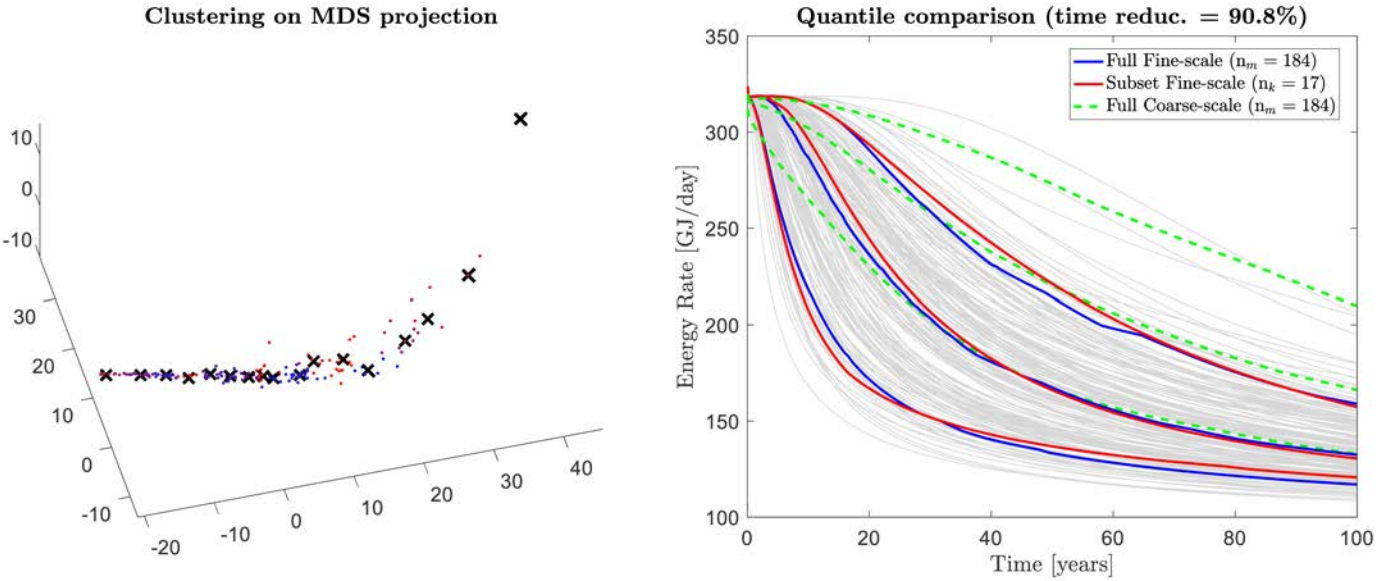

Figure 9—The general workflow of reconstructing the stochastic response. First, use coarse scale information to construct distances between ensemble members. Project into lower dimensional space using MDS and perform clustering. Reconstruct the quantiles based on subset and compare with full ensemble statistics.

Conclusions

In this work we have presented three main aspects which contribute to a robust and UQ quantification workflow for naturally fractured reservoirs. Instead of the using a simplified effective media models (i.e., dual-porosity), we show that it's possible to use high-fidelity DFM models to accurately represent the complex geometry of the fracture networks. We achieve this by utilizing coarse-scale information and selecting a subset of high-fidelity models which describe the same statistics as the full-ensemble of models.

The fracture network generation algorithm allows us to quickly generate a large ensemble of models, depending on some input parameters (e.g., connectivity, number of fractures, orientation, etc.). Connectivity of the final model is ensured by disconnected or connected realizations until target connectivity is achieved. Creating a graph and computing its Laplacian allows for an easy computation of connected fracture sets and therefore speed up this process.

The developed pre-processing method is essential in the reduced UQ workflow. Particularly, it allows us to generate a uniformly meshed DFM model of complex fracture networks for a desired scale l_c, which has a large mesh quality and therefore smooth numerical convergence. This allows for a quick way and reliable way of ranking and partitioning the ensemble space in order to identify representative ensemble members. The preprocessing method seems to be working bests for highly connected networks. Moderate to low connectivity fracture networks seem to be more affected by the preprocessing strategy. Hence it is expected that the coarser realizations result in a lower quality distance for the uncertainty quantification procedure. A remedy for this will be included in future work and consists of a connectivity modification, where necessary, to preserve the main flow characteristics.

Acknowledgments

The project was carried out with a subsidy (reference TKI2017-07-UG) from the Ministry of Economic Affairs, National schemes EZ subsidies, Top sector Energy, carried out by the Netherlands Enterprise Agency. We also want to thank Mohammed Karimi-Fard for his insightful comments and suggestions regarding the pre-processing strategy. Finally we would like to thank Andrea Sartori for the initial development of the fracture generation algorithm.

References

Akbar, M., Vissapragada, B., Alghamdi, A. H., Allen, D., Herron, M., Carnegie, A., Dutta, D., Olesen, J.-R., Chourasiya, R., Logan, D., et al. (2000). A snapshot of carbonate reservoir evaluation. *Oilfield Review*, **12**(4):20–21.

Bakay, A., Demyanov, V., and Arnold, D. (2016). Uncertainty quantification in fractured reservoirs based on outcrop modelling from northeast brazil. In 7th EAGE Saint Petersburg International Conference and Exhibition, pages cp–480. *European Association of Geoscientists & Engineers*.

Berre, I., Doster, F., and Keilegavlen, E. (2018). Flow in fractured porous media: A review of conceptual models and discretization approaches. *Transport in Porous Media*, pages 1–22.

Bisdom, K., Bertotti, G., and Nick, H. M. (2016). The impact of different aperture distribution models and critical stress criteria on equivalent permeability in fractured rocks. *Journal of Geophysical Research: Solid Earth*, **121**(5):4045–4063.

Borg, I. and Groenen, P. J. (2005). Modern multidimensional scaling: Theory and applications. Springer Science & Business Media.

Bour, O. and Davy, P. (1997). Connectivity of random fault networks following a power law fault length distribution. *Water Resources Research*, **33**(7):1567–1583.

Boyd, A., Souza, A., Carneiro, G., Machado, V., Trevizan, W., Santos, B., Netto, P., Bagueira, R., Polinski, R., Bertolini, A., et al. (2015). Presalt carbonate evaluation for santos basin, offshore brazil. *Petrophysics*, **56**(06):577–591.

Caers, J. (2011). Modeling uncertainty in the earth sciences. John Wiley & Sons.

De Hoop, S., Voskov, D., and Bertotti, G. (2020). Studying the effects of heterogeneity on dissolution processes using operator based linearization and high-resolution lidar data. In *ECMOR XVII*, volume **2020**, pages 1–13. European Association of Geoscientists & Engineers.

de Hoop, S., Voskov, D., Vossepoel, F., and Jung, A. (2018). Quantification of coarsening effect on response uncertainty in reservoir simulation. In ECMOR XVI-16th European Conference on the Mathematics of Oil Recovery, volume **2018**, pages 1–16. *European Association of Geoscientists & Engineers*.

Geuzaine, C. and Remacle, J.-F. (2009). Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International journal for numerical methods in engineering*, **79**(11):1309–1331.

Hammersley, J. (2013). Monte carlo methods. Springer Science & Business Media.

Huang, S., Zhang, Y., Zheng, X., Zhu, Q., Shao, G., Cao, Y., Chen, X., Yang, Z., and Bai, X. (2017). Types and characteristics of carbonate reservoirs and their implication on hydrocarbon exploration: A case study from the eastern tarim basin, nw china. *Journal of Natural Gas Geoscience*, **2**(1):73–79.

Karimi-Fard, M., Durlofsky, L., and Aziz, K. (2004). An efficient discrete-fracture model applicable for general-purpose reservoir simulators. *SPE Journal*, **9**(2):227–236.

Karimi-Fard, M. and Durlofsky, L. J. (2016). A general gridding, discretization, and coarsening methodology for modeling flow in porous formations with discrete geological features. *Advances in water resources*, **96**:354–372.

Khait, M. and Voskov, D. (2018a). Adaptive parameterization for solving of thermal/compositional nonlinear flow and transport with buoyancy. *SPE Journal*, **23**:522–534.

Khait, M. and Voskov, D. (2018b). Operator-based linearization for efficient modeling of geothermal processes. *Geothermics*, **74**:7–18.

Khait, M. and Voskov, D. V. (2017). Operator-based linearization for general purpose reservoir simulation. *Journal of Petroleum Science and Engineering*, **157**:990–998.

Koudina, N., Garcia, R. G., Thovert, J.-F., and Adler, P. (1998). Permeability of three-dimensional fracture networks. *Physical Review E*, **57**(4):4466.

Mallison, B. T., Hui, M.-H., and Narr, W. (2010). Practical gridding algorithms for discrete fracture modeling workflows. In ECMOR XII-12th European Conference on the Mathematics of Oil Recovery, pages cp–163. *European Association of Geoscientists & Engineers*.

Manzocchi, T. (2002). The connectivity of two-dimensional networks of spatially correlated fractures. *Water Resources Research*, **38**(9):1–1.

Maryvska, J., Sever`yn, O., and Vohralík, M. (2005). Numerical simulation of fracture flow with a mixed-hybrid fem stochastic discrete fracture network model. *Computational Geosciences*, **8**(3):217–234.

Mello, M. R., Bender, A. A., De Mio, E., et al. (2011). Giant sub-salt hydrocarbon province of the greater campos basin, brazil. In OTC Brasil. Offshore Technology Conference.

Moinfar, A., Narr, W., Hui, M.-H., Mallison, B. T., Lee, S. H., et al. (2011). Comparison of discrete-fracture and dual-permeability models for multiphase flow in naturally fractured reservoirs. In SPE reservoir simulation symposium. Society of Petroleum Engineers.

Mustapha, H. and Mustapha, K. (2007). A new approach to simulating flow in discrete fracture networks with an optimized mesh. *SIAM Journal on Scientific Computing*, **29**(4):1439–1459.

Reijmer, J. J., Johan, H., Jaarsma, B., and Boots, R. (2017). Seismic stratigraphy of dinantian carbonates in the southern netherlands and northern belgium. *Netherlands Journal of Geosciences*, **96**(4):353–379.

Sartori Suarez, A. (2018). Uncertainty quantification based on hierarchical representation of fractured reservoirs. Master's thesis, Delft University of Technology, the Netherlands.

Scheidt, C. and Caers, J. (2009). Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, **41**(4):397.

Scheidt, C., Caers, J., Chen, Y., and Durlofsky, L. J. (2011). A multi-resolution workflow to generate high-resolution models constrained to dynamic data. *Computational Geosciences*, **15**(3):545–563.

Scheidt, C., Caers, J., et al. (2009). Uncertainty quantification in reservoir performance using distances and kernel methods–application to a west africa deepwater turbidite reservoir. *SPE Journal*, **14**(04):680–692.

Spielman, D. A. (2010). Algorithms, graph theory, and linear equations in laplacian matrices. In Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures, pages 2698–2722. *World Scientific*.

Voskov, D. (2017). Operator-based linearization approach for modeling of multiphase multi-component flow in porous media. *Journal of Computational Physics*, **337**:275–288.