The Dissertation Committee for Hoonyoung Jeong Certifies that this is the approved version of the following dissertation:

FAST ASSESSMENT OF UNCERTAINTY IN BUOYANT FLUID DISPLACEMENT USING A CONNECTIVITY-BASED PROXY

Committee:

Kamy Sepehrnoori, Supervisor

Sanjay Srinivasan, Co-Supervisor

Mary Wheeler

Mojdeh Delshad

Mrinal Sen

FAST ASSESSMENT OF UNCERTAINTY IN BUOYANT FLUID DISPLACEMENT USING A CONNECTIVITY-BASED PROXY

by

Hoonyoung Jeong, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2016

Dedication

To my lovely wife for her patience and sacrifice

To my children for their love and patience

Acknowledgements

I would like to express my deepest gratitude to Dr. Sanjay Srinivasan for his support and guidance throughout my research. He is always kind and patient to discussing research problems with me. I could not have finished my dissertation without his enthusiasm and efforts. I am very grateful to my dissertation committee members, Dr. Kamy Sepehnoori, Dr. Mary Whleer, Dr. Mojdeh Delshad, and Dr. Mrinal Sen, for their time and guidance.

I want to thank Jin Lee for her guidance and support in administration issues; Frankie Hart for her guidance, and Roger Terzian and John Cassibry, for their technical support.

I also would like to thank all of my officemates, classmates, and friends I have had here at UT and in Austin (I do not want to leave anyone out accidently!). I would never forget all the fond memories I shared with my friends.

I acknowledge the cooperation of staff of BP and Statoil, and support from the US Department of Energy, grant #DE-DE-FE0004962. Financial support was provided by the sponsors of the Geologic CO_2 Storage Industrial Associates Project at The University of Texas at Austin: BP, Chevron, ExxonMobil, Foundation CMG, Halliburton/Landmark Graphics, Luminant, Shell, Statoil and USGS. Statoil and the Sleipner License are acknowledged for provision of the Sleipner 2010 Reference dataset. Any conclusions in this paper concerning the Sleipner field are my own opinions and do not necessarily represent the views of Statoil.

FAST ASSESSMENT OF UNCERTAINTY IN BUOYANT FLUID DISPLACEMENT USING A CONNECTIVITY-BASED PROXY

Hoonyoung Jeong, Ph.D.

The University of Texas at Austin, 2016

Supervisor: Kamy Sepehrnoori

Co-supervisor: Sanjay Srinivasan

It is crucial to estimate the uncertainty in flow characteristics of injected fluid. However, because a large suite of geological models is probable given sparse static data, it is impractical to conduct full physics flow simulations on the entire suite of models in order to quantify the uncertainty in fluid displacements. Thus a fast alternative to a full physics simulator is necessary to quickly predict the fluid displacements. Most of the proxies proposed thus far are inappropriate to approximate the buoyant flow of injected fluid for 3D heterogeneous rock during the injection period. In this dissertation, a new proxy will be proposed to quickly predict the buoyant flow of injected fluid during CO_2 sequestration.

The geological models are ranked based on the extent of the approximated CO_2 plumes. By selecting a representative group of models among the ranked models, the uncertainty in the spatial and temporal characteristics of the CO_2 plume migrations can be quickly quantified. About 90% of the computational cost of quantifying the uncertainty in the extent of CO_2 plumes was saved using the proposed connectivity based proxy.

In a geological carbon storage project, the spatial and temporal characteristics of CO_2 plume migrations can be monitored by 4D seismic surveys. The images of CO_2

plumes obtained from 4D seismic surveys are used as observed data to find subsurface models honoring the spatial and temporal characteristics of the observed CO_2 plumes. However, because manually comparing an observed CO_2 plume and prior CO_2 plumes in a large suite of subsurface models is inefficient, an automatic measure to calculate the dissimilarity between the CO_2 plumes is necessary.

The most intuitive way to calculate the dissimilarity is the Euclidean distance between vectors representing CO_2 plumes. However, this is inappropriate to measure the dissimilarity between CO_2 plumes because it does not consider spatial relation between the elements of the vectors. The shape dissimilarity between the CO_2 plumes that reflects the spatial relation can be calculated using the Hausdorff distance. The computational cost of calculating the shape dissimilarity between CO_2 plumes is significantly reduced by calculating the Hausdorff distance between the representations of the CO_2 plumes such as perimeter, surface, and skeleton instead of the original CO_2 plumes. An appropriate representation should be chosen according to the spatial characteristics of CO_2 plumes.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1. PROBLEM DESCRIPTION

During the operation of a geological carbon storage project, it is important to predict when and where injected CO_2 remains. The spatial and temporal flow characteristics of the injected CO_2 can be assessed using a full physics simulator in which the complex physics of flow behaviors are implemented. However, if a large suite of geologic models is possible given sparse static data, it takes too high computational cost to run the full physics simulator in the entire suite. A fast proxy is necessary to quickly assess the uncertainty in the spatial and temporal flow characteristics of the injected fluid. However, most of previous proxies are inappropriate to quickly approximate the buoyant flow of CO_2 plumes for 3D heterogeneous rocks. A new proxy for quickly approximating CO_2 plumes is necessary.

The spatial and temporal characteristics of observed CO_2 plumes obtained from 4D seismic surveys can be honored by finding the most probable models of which the CO_2 plumes are spatially and temporally similar to the observed CO_2 plumes. An automatic measure for the dissimilarity between the CO_2 plumes is required if a large suite of geologic models are possible. However, image dissimilarity measures in computer vision literature are inappropriate because they are rotation-, scale-, and location-invariant. Dissimilarity measures for CO_2 plumes must be sensitive to rotation, scale, and location of CO_2 plumes because they have different spatial characteristics. The most intuitive way is the Euclidean distance between vectors representing CO_2 plumes, but it does not take account into the spatial relation between the elements of the vectors.

---

* Reprinted with permission from “Fast assessment of CO_2 plume characteristics using a connectivity based proxy” by Jeong, H. and Srinivasan, S., 2016. International Journal of Greenhouse Gas Control.

Efficient shape dissimilarity measures for considering the spatial relation between the elements of the vectors representing CO_2 plumes are necessary.

1.2. RESEARCH OBJECTIVES

The main objective of this dissertation is to quickly quantify the uncertainty in fluid displacement and select the most probable models among a large suite of geologic models that honor the observed data. This objective is divided into the following:

[1] Development of a fast proxy for approximating the buoyant flow of injected fluid

[2] Incorporation of that proxy within a framework for model selection in order to identify the most probable models honoring an observed fluid displacement

[3] Development of a software module for the fast assessment of flow characteristics and the quick selection of probable models

1.3. DISSERTATION OUTLINE

This dissertation consists of eight chapters. Following an introductory chapter and a literature review chapter, Chapter 3 introduces a connectivity-based proxy for quickly approximating the buoyant flow of injected fluid. After the proxy is validated for several different conditions, the proxy is applied to two real field data. Chapter 4 describes how to quickly quantify the uncertainty in the extent of fluid displacements using the connectivity-based proxy. Chapter 5 introduces the Hausdorff distance and its application to measure the dissimilarity between 2D fluid displacements. It will show that a set of the most probable models honoring an observed 2D fluid displacement can be quickly selected using the Hausdorff distance between parameter vectors and compare it performance to the Euclidean distance. Chapter 6 introduce the Hausdorff distance between surfaces and skeletons to measure the dissimilarity between 3D fluid displacements. It will show that the computational cost of the model selection process can

be significantly saved using the Hausdorff distance between surfaces and skeletons compared to the Hausdorff distance. Chapter 7 describes the software development process, the implementation of the algorithms for the proxy and the model selection process, and the input parameters for the software module. The final chapter summarizes the key findings of this work and makes recommendations for future work.

Chapter 2. Literature Review

2.1. APPROXIMATING CO_2 PLUME MIGRATIONS

In a geological carbon storage project, one of the critical questions is whether injected CO_2 will remain confined within a permitted zone (Jeong and Srinivasan, 2016; Jeong et al., 2013). In order to answer this question, the spatial and temporal characteristics of the CO_2 plume should be estimated before CO_2 is injected. However, because sparse data would be available to characterize aquifers, a large suite of subsurface models has to be considered in order to account for the significant uncertainty associated with the geology of the storage formation. Senel and Chugunov (2013) assessed the uncertainty of injection performance and the extent of CO_2 plumes in a CO_2 storage site in Illinois Basin. They ran a full physics simulator for 200 cases described using different reservoir parameters. However, excessive computational cost is incurred in order to conduct full physics simulations over a large suite of subsurface models. Fast alternatives to full physics simulators have been researched to alleviate the computational cost of predicting CO_2 plume migrations.

The fastest way to approximate the buoyant flow of CO_2 plumes is to analytically or semi-analytically solve a simplified form of the governing equations describing the CO_2 migrations. In these approaches, the physics and the description of fluid and rock properties are simplified. The analytical solutions for a CO_2 plume migration are useful to screen potential CO_2 storage sites. Nordbotten et al. (2005) presented an analytical solution for the height of CO_2 column, and a semi-analytical solution for the pressure distribution as a function of time and radial distance from an injector. Mathias et al. (2009) developed an analytical solution to the temporal and spatial pressure distribution. Vilarrasa et al. (2013) presented a semi-analytical solution that considers the effect of a non-uniform injection rate along the entire aquifer interval. However, these existing

analytical and semi-analytical solutions are applicable to homogeneous rock. Thus they are inappropriate for approximating a CO_2 plume migration when the heterogeneity of the rock properties such as facies, permeability, and porosity significantly influences the CO_2 plume migration.

Vertical-equilibrium (VE) models save the computational cost of predicting a CO_2 plume migrationby simulating flow in the vertically averaged model instead of a fully 3-dimensional model. The key assumption of the VE models is that brine and CO_2 attain vertical equilibrium instantaneously due to strong buoyancy forces (Guo et al., 2014). Under such an assumption, the lateral migration of the plume is solved numerically over the vertically averaged 2D domain of the reservoir and subsequently, the vertical migration of the plume is solved analytically. The VE models thus provide 3-dimensional results for CO_2 saturations (Nilsen et al., 2011). Bandilla et al. (2014) showed that the CO_2 plumes for the Sleipner L9 model computed using their VE models are in good agreement with those computed using a fully 3-dimensional model. However, the vertically coarsened model still needs to be simulated. Court et al. (2012) reported that the VE models are inappropriate for formations that have low vertical permeability, large thickness, or brine and CO_2 that have a small contrast in density. In such cases, the vertical equilibrium between brine and CO_2 is not established along the thickness of the formation. In addition, the vertical equilibrium assumption does not work properly when a CO_2 plume migration is controlled by 3-dimensional heterogeneity. This will be shown in the verification section of the connectivity-based proxy.

Bhowmik et al. (2014) applied a random walk to diagnose flow of CO_2 plumes in heterogeneous aquifer models. The random walker approximates migration paths as well as flow responses by stochastically transitioning walkers from one location to the next. The random walker is a Markov process and it requires the current state of the walkers in

order to predict their next state. The transitions of a large number of walkers over a fairly large grid is not only required, but these transitions also have to be modeled sequentially adding to the computational time.

Invasion percolation theory can be used to simulate the invasion of slow moving immiscible fluids when viscous forces are negligible compared to capillary and gravity forces (Carruthers, 2003). In this theory a fluid invades or percolates into porous media if the fluid pressure is greater than the threshold capillary pressure. However, invasion percolation theory is not valid when CO_2 is injected at a high rate because viscous forces are dominant over capillary forces (Cavanagh, 2013; Singh et al., 2010).

Streamline simulation offers rapid modeling of fluid displacements in heterogeneous porous media. Kovscek and Wang (2005) showed that an analytic streamline-based proxy can provide a good approximation of the ratio of oil recovery to injected CO_2 volume in a large set of heterogeneous reservoir models during CO_2 enhanced oil recovery. However, the streamline proxy has two disadvantages in approximating a CO_2 plume migration. First, the streamline proxy cannot reproduce the buoyancy effect between two fluids with different densities (Thiele and Batycky, 2015). Second, a pressure equation must be solved at least once to calculate the streamlines using the potential field (Sharifi et al., 2014). Even a one-time calculation demands high computational cost in a geological model with millions of cells.

Sethian (1996) introduced the fast-marching method (FMM) to quickly compute the position of monotonically advancing fronts. FMM has been recently applied to approximate reservoir drainage volume (Sharifi et al., 2014; Xie et al., 2015). FMM is a computationally efficient method to approximate fluid movements in heterogeneous porous media. However, it cannot reproduce the buoyant flow of a CO_2 plume.

In this study, we propose a fast connectivity-based proxy to approximate a CO_2 plume migration in 3-dimensional reservoir models with heterogeneity in facies, permeability, and porosity during an injection period where viscous forces are dominant over capillary forces. CO_2 plume migration is governed by advective and buoyant forces during the injection period while dissolution, diffusion, and density-driven convection processes take over during the post-injection period (Darcis et al., 2011). In particular, density-driven convection can cause downward migration of the CO_2 plume in the form of fingers that have the opposite flow direction to the buoyant flow direction (Zhang, 2013). This study is aimed at approximating a CO_2 plume migration only during the injection period.

The connectivity-based proxy has advantages over the previous proxies in terms of the computational cost and accuracy because it considers the buoyant flow of a CO_2 plume without any complicate pressure calculations. CO_2 plume migrations in a synthetic aquifer model computed using the connectivity-based proxy are compared to those computed using a full physics simulator, CMG-GEM (CMG, 2012) for various combinations of injection periods, injection rates, a CO_2 density, a CO_2 viscosity, the degree of rock heterogeneity, permeability of cap rock, and topology of the aquifer. The full physics simulations are also compared to the CO_2 plume migrations computed using FMM and a VE model. The connectivity-based proxy is also verified for two field cases. Lastly, we will show that the uncertainty in the extent of CO_2 plumes in a large suite of geological models can be quickly quantified using the connectivity-based proxy.

2.2. SHAPE DISSIMILARITY MEASURE BETWEEN FLUID DISPLACEMENTS

This section introduces measures for shape dissimilarity between fluid displacements to find models honoring observed fluid displacement data. CO_2 plume

migrations are used as examples to measure the shape dissimilarity between fluid displacements.

4D seismic images that describe spatial and temporal CO_2 plume migrations are important for calibrating simulation models for CO_2 storage. Once simulated CO_2 plumes are matched to observed CO_2 plumes by adjusting the model parameters of simulation models, reliable prediction of future CO_2 plume migration can be achieved. This would be invaluable for risk assessment of a CO_2 storage project and its management.

Chadwick and Noy (2010) calibrated aquifer models to honor CO_2 plumes obtained using 4D seismic surveys for the Sleipner project. They manually matched simulated CO_2-water contacts (CWC) to the observed CWC by adjusting the aquifer model parameters such as CO_2 density, permeability, and the depth topography of the top surface. Singh et al. (2010) compared the observed CWC and that computed using a commercial simulator and relative permeability curves for the Sleipner benchmark model released by Statoil. Chadwick and Noy (2010) and Singh et al. (2010) compared the observed and simulated CWCs visually for a small number of models. However, a faster measure to compute the mismatch between the observed and simulated plumes is necessary to screen a large number of models.

Zhu et al. (2015) matched simulated and observed CWC inverted from 4D seismic data in the Sleipner benchmark model by adjusting the temperature of the aquifer, permeability anisotropy, and the composition of injected CO_2. They regarded the observed and simulated CO_2 plumes as images and minimized the mismatch by adjusting aquifer parameters. They used image analysis software to compute the percentage of area within the observed CO_2 plume that the simulated CO_2 plume failed to cover. However, the mismatch between the spatial characteristics of the simulated plumes for two different

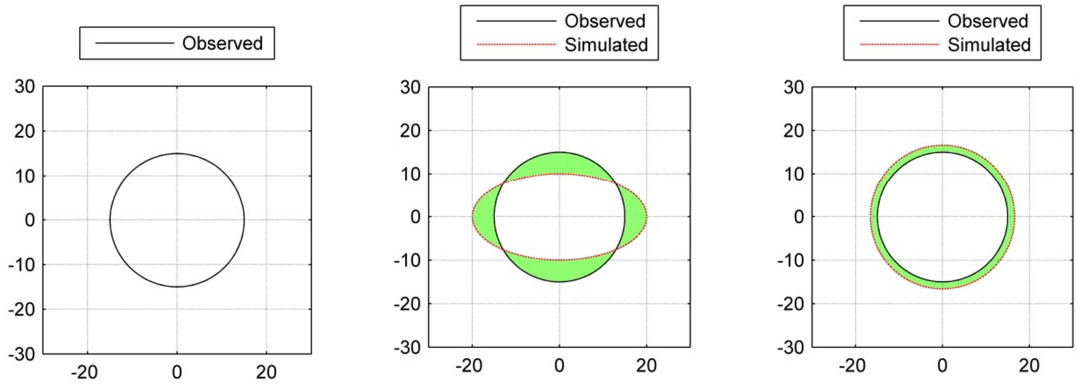
models can be different even though they have the same level of mismatch. An example of this will be presented later.

Similar to the application of image analysis software presented in Zhu et al. (2015), other techniques for measuring image dissimilarity in computer vision literature can be considered. However, most of those techniques are not applicable to measuring the dissimilarity between the shapes of CO_2 plumes during sequestration. In computer vision, rotation, scale, and location invariances are essential properties to consider when formulating image dissimilarity because images differ according to the perspective even though they include the same object (Nixon and Aguado, 2012). The objective for using such computer vision techniques is to detect and compare objects regardless of rotation, scale, and location of objects. However, in sequestration applications, differences in terms of rotation, scale, and location of CO_2 plumes should be discriminated. In this regard, techniques measuring image dissimilarity regardless of rotation, scale, and location are inappropriate for measuring the dissimilarity in shapes of CO_2 plumes.

The most common and intuitive way to measure image dissimilarity considering the rotation, scale, and location of CO_2 plumes is the Euclidean distance between vectors representing an observed and simulated CO_2 plumes. The Euclidean distance is computed as the square root of the squared difference between each element of the observed and simulated vector. Obidegwu (2015) measured the dissimilarity between binary images of gas distribution obtained from seismic data using the Hamming distance (Hamming, 1950). The Hamming distance between binary images is equivalent to the square of the Euclidean distance. However, both the Euclidean and Hamming distance do not indicate the difference between the vector elements at individual locations. In the Euclidean distance, the mismatch between locations that are spatially close or far are treated

equally. Even though simulated CO_2 plumes are spatially or visually different from an observed CO_2 plume, they can have the same Euclidean distance.

For example, Figure 2.1 shows an observed (left) and two simulated CWCs. In Figure 2.1(a), the observed CWC is a circle with a radius, 15 units. In Figure 2.1(b), the first simulated CWC is an ellipse with a horizontal radius, 20 units, and a vertical radius, 10 units. In Figure 2.1(c), the second simulated CWC is a circular plume with a radius, 16.60 units. In Figure 2.1(b) and (c), the green area represents the mismatch between the observed and simulated CWCs. The first and second simulated CWCs are spatially and visually different. However, the simulated CO_2 plumes have the same mismatch in terms of the mismatched area as well as the Euclidean distance to the observed CO_2 plume. In other words, they honor the observed data to the same extent even though the second simulated CWC is more similar to the observed CWCs. Thus the Euclidean distance is inappropriate to measure the dissimilarity between the shapes of CO_2 plumes because the Euclidean distance does not account for the spatial difference between the CO_2 plumes.



(a) Observed CWC

(b) First example of a simulated CWC

(c) Second example of a simulated CWC

Figure 2.1 Observed and two simulated CO_2-water contacts. The areas marked in green show the mismatches between the simulated and observed contacts.

Another possibility for comparing CO_2 plume migrations is to measure shape dissimilarity using the Hausdorff distance. The Hausdorff distance is used to match the shape of objects in binary images (Huttenlocher et al., 1993). The Hausdorff distance is not only sensitive to rotation, scale, and location (Zhang and Lu, 2004), but it also considers the mismatch between images spatially.

The Hausdorff distance has been applied to measure shape dissimilarity of geologic features and water fronts (Abadpour et al., 2013; Lee et al., 2013; Suzuki and Caers, 2008). However, the Hausdorff distance has not been used to measure the shape dissimilarity between CO_2 plumes in CO_2 sequestration. In this study, the Hausdorff distance is compared to the Euclidean distance for finding the most probable models honoring the spatial characteristics of an observed CO_2 plume and approaches for reducing the computational cost of the Hausdorff distance are introduced.

The computational cost of measuring shape dissimilarity between CO_2 plumes using the Hausdorff distance is subject to the number of grid blocks describing the CO_2 plumes. Specifically, the cost of computing the Hausdorff distance between point sets *A* and *B* is expressed as *O(mn)* where *m* and *n* are the numbers of points consisting *A* and *B*, respectively (Huttenlocher et al., 1993).

Many researchers have studied measuring shape dissimilarity of perimeters, surfaces, and skeletons using the Hausdorff distance (Aspert et al., 2002; Attali and Montanvert, 1997; Dubuisson and Jain, 1994; Enayatifar and Salam, 2013; Gang Pan et al., 2003; Huttenlocher et al., 1993; Lipikorn et al., 2004; Sim et al., 1999; TAKÁCS, 1998; Yue Lu et al., 2001; Zhao et al., 2005). We introduce shape representations of objects such as a perimeter (Chapter 5), a surface, and a skeleton (Chapter 6) to save the computational cost of calculating the shape dissimilarity between CO_2 plumes using the Hausdorff distance. Representations of an object are calculated using only the non-zero

pixels that are needed to represent the shape of the object. The number of non-zero pixels or voxels in the representations is smaller than that in the entire original image, but they are known as effective shape descriptors of an object (Chatbri et al., 2015; Nixon and Aguado, 2012; Zhang et al., 2007). The computational cost of calculating the Hausdorff distance that is proportional to the numbers of points can be reduced using the perimeter, surface, and skeleton objects. The computational cost of calculating the Hausdorff distance between the original CO_2 plume and between their representations is compared.

Chapter 3. Scaled Connectivity Analysis

3.1. SCALED CONNECTIVITY ANALYSIS

Connectivity analysis measures how connected a well is to every grid cell in a discrete rendition of the reservoir model. The migration of an injected fluid can be approximated using connectivity analysis by simply assuming that an injected fluid reaches a grid cell with better connectivity earlier. However, connectivity analysis based on only reservoir static properties such as porosity and permeability has to be modified to account for buoyant force. In this section, we investigate further why connectivity analysis cannot reproduce the buoyant flow behavior of a CO_2 plume and a remedy for the problem is presented.

3.1.1. Connectivity analysis

Hird and Dubrule (1998) presented a way to calculate the least resistive path from wells to grid cells and they defined the cost of the least resistive path as connectivity. The resistivity between two neighboring grid cells is calculated knowing their average effective permeability, the cross-sectional area to flow, and the distance between two adjacent grid cells. They obtained a good correlation between connectivity and water breakthrough times in 2-dimensional cross section models.

Hirsch and Schuette (1999) introduced principles of graph theory to quickly assess connectivity of reservoir models. They postulated that grid cells are connected by edges, and the edge weights are computed using reservoir parameters that impact connectivity such as porosity and permeability. They computed effective horizontal permeability, and determined connected components using graph theory concepts.

Pardo-Igúzquiza and Dowd (2003) developed a computer program for estimating connectivity in a 3-dimensional model. Once an indicator map that defines a permeable

phase or lithofacies is available, their program calculates some connectivity statistics such as minimum, mean, and maximum lengths of connected components.

De Lima et al. (2012) characterized well-to-reservoir and well-to-well connectivity in fractured reservoir models using the methods presented by Hirsch and Schuette (1999). They used the ratio of pore volume to transmissibility between neighboring grid cells as an edge weight, which is one of Hirsch and Schuette's suggestions. This edge weight can be also interpreted as the time needed to fill the given pore volume with a fluid of unit viscosity under a unit potential difference (De Lima et al., 2012).

$$Edge\ weight = \frac{Average\ pore\ volume}{flow\ rate\ from\ Darcy's\ law} = \frac{\sqrt{Vp_i Vp_j}}{T_{i,j} \cdot \frac{\Delta\Phi}{\mu}} = \frac{\sqrt{Vp_i Vp_j}}{T_{i,j}} \quad \text{Equation (3-1)}$$

The subscript i and j are the indices of grid cells, V_p is pore volume, $\Delta\Phi$ is the potential difference between grid cells i and j, μ is the fluid viscosity, $T_{i,j}$ is the transmissibility between grid cells i and j. ΔP and μ are dropped from the final expression because they have unit values. $T_{\mathrm{i,j}}$ of two regular cells connected along the x-direction is given by (Donnez, 2007):

$$T_{i,j} = \frac{T_i \cdot T_j}{T_i + T_j} \ where\ T_i = k_{x,i} \frac{\Delta y \cdot \Delta z}{\Delta x_i / 2} \ and\ T_j = k_{x,j} \frac{\Delta y \cdot \Delta z}{\Delta x_j / 2} \cdots\cdots\cdots\cdots\cdots \text{Equation (3-2)}$$

T_{i} and T_j are transmissibilities computed corresponding to half the cell. k_x is the absolute permeability in the x-direction. Δx, Δy, and Δz are the cell sizes in the x-, y-, z-directions, respectively. $T_{i,j}$ is the harmonic average of T_i and T_j. $T_{i,j}$ of non-regular cells is calculated using the equations that Donnez (2007) provides.

A smaller edge weight implies a shorter time to fill the grid cells or the shorter travel time of the fluid between the grid cells. The connectivity from a source to a grid cell is the sum of the edge weights along the shortest path, which is computed using

Dijkstra's algorithm (Fredman and Tarjan, 1987). In other words, the connectivity is the shortest time to travel from the source to the grid cell. The proposal is to use this approach to approximate the migration of injected fluid. The procedure of calculating fluid displacement using connectivity analysis (CA) is as follows:

i) Calculate edge weights between neighboring grid cells using Equation (3-1)

ii) Calculate connectivity using Dijkstra's algorithm

iii) Sort grid cells in ascending order of connectivity

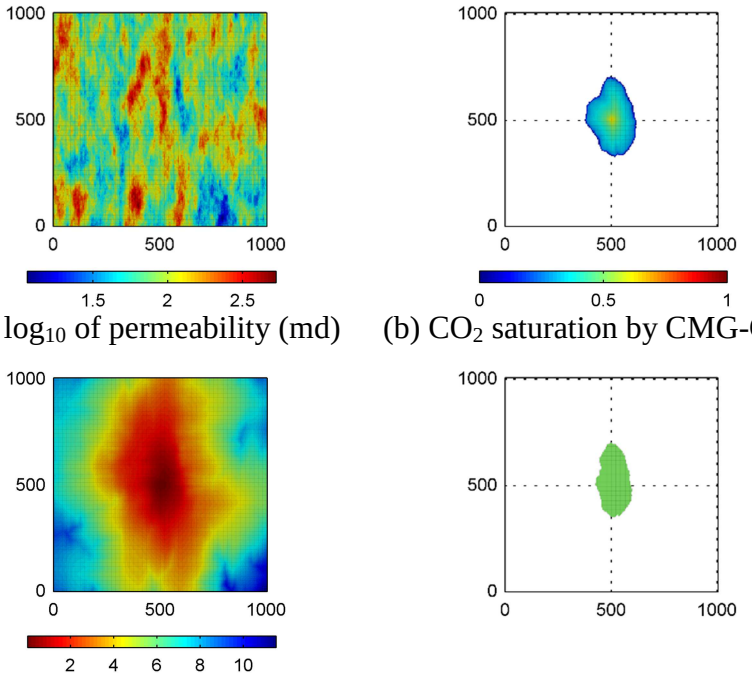
iv) Fill an unfilled grid cell with the smallest connectivity with as much fluid as the product of the cell volume, porosity, and average injected fluid saturation

v) Repeat the 4^{th} step until the filled total volume equals the injected volume

A method to compute the average injected fluid saturation will be discussed later. The key assumption is that the injected fluid migrates faster to a grid cell with better connectivity. The migrated region is calculated by truncating the flow path such that the saturated volume of the intersected grid cells equals the volume of the injected fluid. Each intersected grid cell is assigned a binary value: either saturated with the injected fluid or not.

CA is applied to a two-dimensional case as shown in Figure 3.1(a), where the dimensions of the grid system are 201 by 201 by 1 and the dimensions of the grid cell are 5m by 5m by 2m. The porosity is constant, 0.2. CO_2 is injected at the center of the aquifer at 500 m^3/day for a duration of 10 years. The reservoir is described using a single layer in the vertical direction. Figure 3.1(b) shows the CO_2 plume migration computed by a compositional numerical simulator (CMG-GEM). In Figure 3.1(c), redder means better connectivity and shorter travel time. In Figure 3.1(b) and Figure 3.1(c), CO_2 seems to migrate earlier to grid cells indicated in red. Because of the manner in which connectivity

is computed, the CO_2 plume migration is mainly governed by permeability heterogeneity. Compared to Figure 3.1(b), CA provides a good approximation for the CO_2 migration along the north-south permeable direction as shown in Figure 3.1(d).



(a) $\log_{10}$ of permeability (md) (b) CO_2 saturation by CMG-GEM

(c) Connectivity (1000 days) (d) CO_2 migration by CA

Figure 3.1 Log_{10} of permeability for a 2D case, corresponding CO_2 saturation computed using CMG-GEM, and corresponding results of CA. The unit of the axes is meters.

Now CA is applied to a three-dimensional case as shown in Figure 3.2 where the z-axis is exaggerated by a factor of 10. The dimensions of the grid system are 201 by 201 by 10 and the dimension of an individual grid cell is 5m by 5m by 2m. CO_2 is injected at the center of the aquifer and 15m below the aquifer top at 6,000m^3/day for a period of 2 years.

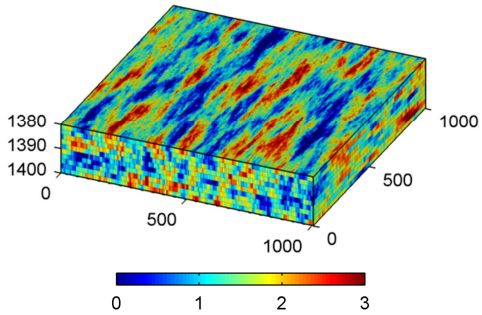
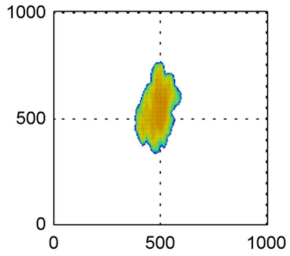
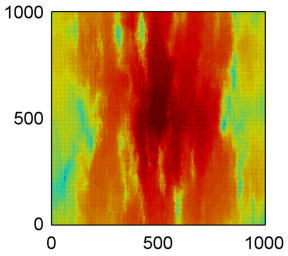
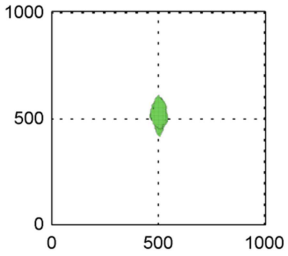
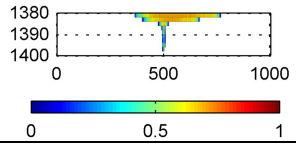
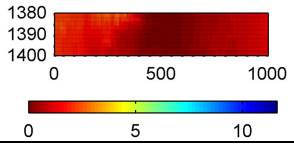
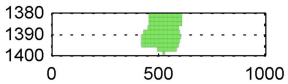


Figure 3.2 Log_{10} of permeability (md) of a 3D case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

Contrary to the 2D case, the CO_2 plume simulated using CMG-GEM in Table 3.1 exhibits buoyant characteristics. After the injected CO_2 migrates vertically in the vicinity of the injection well, it spreads horizontally. However, the CO_2 plume approximated using CA does not exhibit such buoyant behavior.

Table 3.1 Comparison of CO_2 migration by CMG-GEM and CA using Equation (3-1) in the geological model shown in Figure 3.2. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	CO_2 saturation by CMG-GEM	Connectivity (years)	CO_2 migration by CA
Areal view			
Slice at i=101			

In order to take the buoyancy effect into account, the potential difference caused by viscous and gravity forces should be considered in Equation (3-1). The relative permeability of CO_2 should be also considered in the transmissibility term because the

edge weight represent the travel time of CO_2 when two fluids are flowing concurrently. Likewise, the CO_2 volume should be used instead of the pore volume. Thus Equation (3-1) for the edge weight is extended to Equation (3-3):

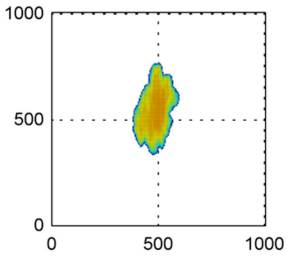
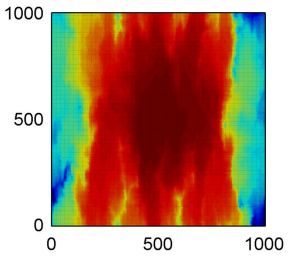
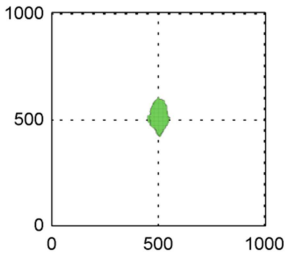
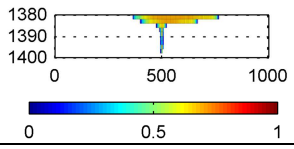
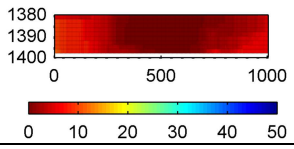
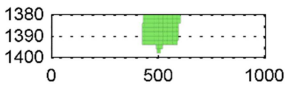
$$Modified\ Edge\ weight = \frac{Average\ CO_2\ volume}{CO_2\ flow\ rate\ from\ Darcy's\ law} = \overline{\mu_{CO_2}} \frac{\hat{S}_{CO_2}\sqrt{Vp_i \cdot Vp_j}}{\hat{k}_{r,CO_2} T_{ij}\left[\Delta P_{ij} + \left(\overline{\rho_{brine}} - \overline{\rho_{CO_2}}\right) g \Delta h_{ij}\right]} \quad \cdots\cdots\cdots \text{Equation (3-3)}$$

The i and j are cell indexes, $\overline{\mu_{CO_2}}$ average CO_2 viscosity in aquifer conditions, ΔP pressure difference, $\overline{\rho_{brine}}$ average brine density in aquifer conditions, $\overline{\rho_{CO_2}}$ average CO_2 density in aquifer conditions, g gravity constant, Δh depth difference between the i-th and j-th cells. $\hat{S}_{CO_2}$ and $\hat{k}_{r,CO_2}$ are the CO_2 saturation and the CO_2 relative permeability at the end point of the CO_2 relative permeability curve. ΔP is the pressure difference between the i-th and j-th grid cells. The procedure to calculate ΔP will be discussed later.

Now the connectivity is calculated for the 3D model shown in Figure 3.2 using Equation (3-3) instead of Equation (3-1). The results obtained using Equation (3-3), are shown in Table 3.2. In the results of the full physics flow simulation shown in Table 3.2, it is evident that the CO_2 plume exhibits strong buoyancy caused by the density difference between CO_2 and water. The CO_2 plume vertically migrates to upper layers and then horizontally spreads within the top layer. CA corresponding to Equation (3-3) considers buoyant force, but the spread of the CO_2 plume in all layers along the vertical migration path is more similar to a cylinder rather than a wedge as depicted by the full physics flow simulation result. However, in the full physics simulation, CO_2 mostly migrates in the vertical direction with high velocity in the vicinity of the injection well before reaching the top layer. The reason why CA cannot reproduce the buoyant behavior

is discussed in the next section and then, a modified connectivity analysis that reproduces the buoyant flow will be presented.

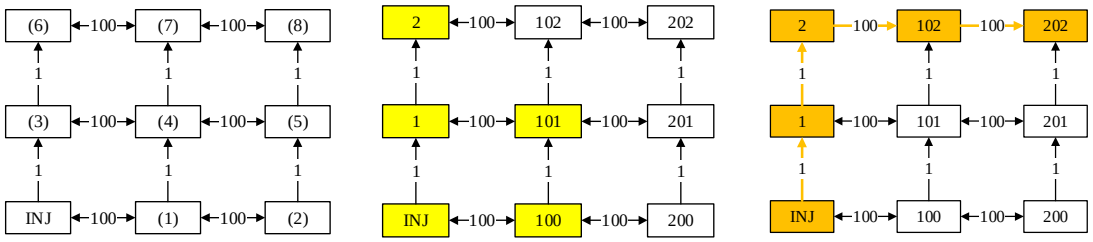
Table 3.2 Comparison of CO_2 migration by CMG and CA using Equation (3-3) in the geological model shown in Figure 3.2. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	CO_2 saturation by CMG-GEM	Connectivity (years)	CO_2 migration by CA
Areal view			
Slice at i=101			

3.1.2. Scaled Connectivity Analysis

Let us study a simple example to investigate why connectivity analysis cannot reproduce the flow pattern corresponding to buoyant flow. Figure 3.3 shows an example of tracing the migrated regions along a cross-section using the original connectivity analysis. In Figure 3.3(a), the bracketed numbers in the grid cells and the numbers on the arrows represent the indices of the grid cells and the edge weights, respectively. The vertical edge weight is 1 and the horizontal edge weight is 100. The edge weights in the vertically downward direction are infinite because the potential differences are negative in that direction. The total injected amount is 5 grid cells and the saturated blocks are colored in yellow. The connectivities of the grid cells are presented in Figure 3.3(b). For example, the shortest paths from the injector to the grid cell (7) are INJ-(3)-(6)-(7), INJ-

(3)-(4)-(7), and INJ-(1)-(4)-(7). The paths are different, but the costs or the connectivities to the grid cell (7) are the same, 1+1+100=102. In Figure 3.3(b), the connectivities are truncated with the injected volume in the order of the connectivity values. The grid cells with the top five connectivities (0, 1, 2, 100, 101) are marked as the grid cells contacted by the migrating CO_2 plume.



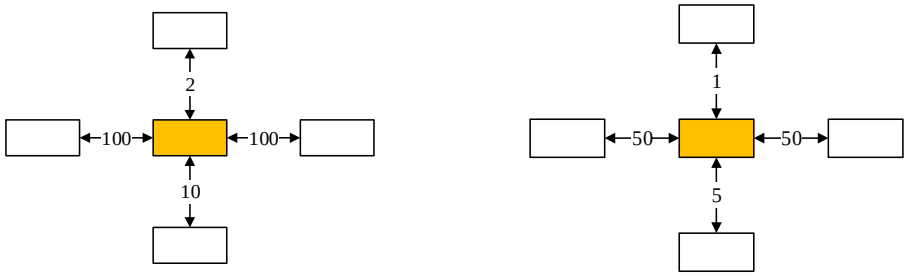
(a) A graph (b) Truncated with the injected volume, 5 (c) Expected CO_2 plume migration

Figure 3.3 An example of connectivity analysis

However, Figure 3.3(c) shows the expected plume migration when buoyancy is dominant. The original approach is inappropriate for representing buoyant flow. This is because a fluid does not migrate along the shortest path in terms of the sum of travel times between neighboring grid cells. In fact, the movement of the CO_2 plume is strongly affected by the local gradient at each location of the plume and the movement is in the direction of the maximum gradient. In Figure 3.3(c), most of the fluid migrates vertically rather than horizontally at first because there is big difference between the two edge weights, 1 and 100 in the vicinity of the injector. The injected fluid rapidly migrates to the top layer, and then the fluid migrates horizontally in the top layer along the local minimum edges, which is colored orange in Figure 3.3(c). For example, an edge weight of 1 is the minimum adjacent to the injector location, and an edge weight of 1 is also the minimum at the next location of the plume (middle left). Then an edge weight of 100 is the minimum at the top left because the edge weight in the downward direction is infinite.

Again an edge weight of 100 is the minimum at the next locations (top middle and top right). The actual migration follows the path of local minimum edges. Thus CA is modified so that the shortest paths contain the local minimum edges as many as possible.

Scaled connectivity analysis (SCA) scales edge weights with the local minimum edge weight so that fluid migrates along the edge with the local minimum edge weight at each grid cell. In SCA, the minimum edge weight among the neighboring edges is found and then the edge weights are divided by the minimum edge weight. In Figure 3.4, the minimum edge weight is 2 and then the edge weights are scaled by dividing all the edge weights to that block by 2.



(a) original edge weights (b) scaled edge weights

Figure 3.4 An example of locally scaled edge weights

Figure 3.5 shows the result of scaled connectivity analysis for the example in Figure 3.3. Table 3.3 shows the sequential steps of calculating the shortest path from INJ to the grid cell (8) in the scaled graph in Figure 3.5. As shown in Figure 3.5, the truncated connectivity in the scaled graph reproduces the local minimum path. As is evident, by scaling edge weights with their local minimum, the buoyant flow behavior is reproduced.

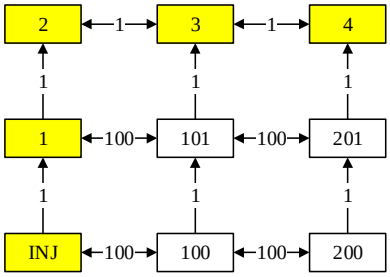


Figure 3.5 Result of scaled connectivity analyses for the example shown in Figure 3.3

Table 3.3 Sequential steps of calculating the shortest path from INJ to the grid cell (8) in shown Figure 3.5

Current grid cell	Connected grid cells	Connected edge weights	Local minimum edge weight	Scaled edge weights	Next grid cell
INJ	(1), (3)	100, 1	1	100, 1	(3)
(3)	(4), (6)	100, 1	1	100, 1	(6)
(6)	(7)	100	100	1	(7)
(7)	(6), (8)	100, 100	100	1, 1	(8)

There are three differences between connectivity analysis and scaled connectivity analysis. The first difference is to calculate connectivity in a scaled graph instead of the original graph. The second difference is to use a modified Dijkstra's algorithm that calculates the cumulative sum of original edge weights along the shortest path computed in a scaled graph. The last difference is that the modified Dijkstra's algorithm rejects the path along which the travel time is greater than a critical time computed based on the injection period. The modified Dijkstra's algorithm can save computation time by ignoring paths that are infeasible given the specified time.

3.1.3. Procedure of SCA

The procedure of approximating CO_2 plume migration using SCA is as follows:

i) Calculate a radial pressure gradient at the centroid of each cell using the following equation:

$$\frac{dP}{dr} = \frac{Q_{CO_2}\overline{\mu_{brine}}}{2\pi r \bar{k}\bar{h}} \quad \text{Equation (3-4)}$$

ii) Calculate edge weights between neighboring grid cells using Equation (3-3)

iii) Change negative edge weights to infinity (implying that flow cannot occur against gravity)

iv) Scale edge weights with their local minimums

v) Calculate scaled connectivity using the modified Dijkstra's algorithm

vi) Fill grid cells with an injected fluid in the order of the scaled connectivity such that the filled total volume equals the injected volume

In Equation (3-3) and Equation (3-4), $\overline{\mu_{CO_2}}$, $\overline{\rho_{CO_2}}$, $\overline{\mu_{brine}}$, and $\overline{\rho_{brine}}$ can be obtained from a full physics simulation result or thermo-dynamically calculated under bottom hole conditions.

Equation (3-4) is the analytical solution for the pressure gradient in the radial flow in homogeneous rock under steady state flow conditions (Ahmed, 2006). In Equation (3-4), P is pressure and r is the distance from an injector. Q_{CO_2} is the flow rate in aquifer conditions, $\overline{\mu_{brine}}$ average viscosity of brine in aquifer conditions, $\bar{k}$ average permeability, $\bar{h}$ average thickness of the aquifer.

In SCA, $\bar{k}$ is used to approximate the pressure gradient in Equation (3-4). Because $\bar{k}$ is likely to be related to the permeability in the CO_2 swept zone, $\bar{k}$ might be changed according to operating conditions such as injection time and injection rate. An arithmetic average can be attempted first for $\bar{k}$ in Equation (3-4). However, $\bar{k}$ should be manually adjusted until SCA gives a good approximation of a full physics simulated CO_2 plume migration. Our final goal is not to obtain a good approximation of a CO_2 plume migration for only one geological model, but to acquire proxy responses that represent CO_2 plume migrations in a large suite of geological models. The adjustment of $\bar{k}$ is conducted for only one geological model, and then the adjusted $\bar{k}$ is also used for the other models that have the same geological setting for permeability and porosity such as mean and variogram.

The steady state pressure solution corresponding to flow in a cylindrical homogenenous medium is given by (Ahmed, 2006):

$$P = P_{wf} + \frac{Q\mu}{2\pi kh} log \frac{r}{r_w} \quad \text{Equation (3-5)}$$

where P_{wf} is the well bottom hole pressure, r is a radial distance from the well, r_w is a well radius. In this solution to the pressure equation in radial coordinates, two grid cells that are located at a considerable distance, d_{ij}, from each other might have small pressure drop across them if they are separated by a small radial distance. This implies that the edge weights along the radial direction will be small while in other directions, the edge weights will be large. This in turn causes the shortest paths to be always in the radial direction and the final approximation of plume migration to have a radial shape. Consequently, we propose a modified expression for the pressure drop as:

$$\Delta P_{ij} = \frac{\frac{dP}{dr}\Big|_i + \frac{dP}{dr}\Big|_j}{2} d_{ij} \quad \text{Equation (3-6)}$$

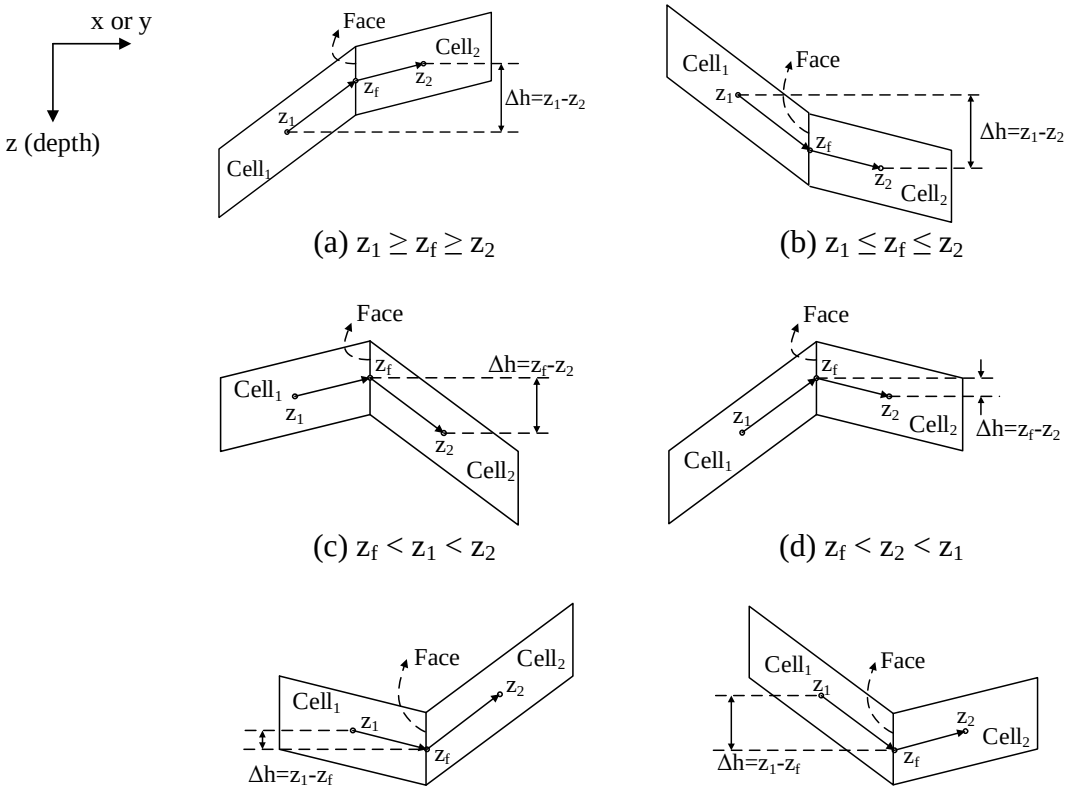
where d_{ij} is the distance between the grid cells i and j. In Equation (3-6), the distance between two neighboring grid cells is used instead of the radial distance to consider the effect of heterogeneity in directions other than the radial direction.

Brine viscosity is used in Equation (3-4) instead of CO_2 viscosity even though CO_2 is injected. The velocity of CO_2 plume is controlled by pre-existing brine even if CO_2 is much less viscous than brine because the aquifer is filled with brine. Thus brine viscosity is more appropriate than the CO_2 viscosity to approximate the pressure gradient.

The edge weight in Equation (3-3) is the time needed to fill the given pore volume. The maximum pore volume filled with CO_2 is the pore volume times the CO_2 saturation at the end point of the CO_2 relative permeability curve. Thus the end point of the CO_2 relative permeability curve is used for $\hat{S}_{CO_2}$ and $\hat{k}_{r,CO_2}$.

The gravity term, $\Delta(\rho g h)_{ij}$, in Equation (3-3), assumes that $\Delta\rho g$ is constant and Δh depends on the vertical locations of two adjacent grid cells. If the face in contact with

the two adjacent cells is between the center of the two adjacent cells, as shown in Figure 3.6(a) and (b), then Δh is calculated as the depth (z_1) of the starting cell minus the depth (z_2) of the ending cell. In Figure 3.6(b), the negative gravity term means the direction from the starting cell to the ending cells is opposite to the buoyant flow direction. If the sum of a pressure difference and a negative gravity term is still negative, then the edge weight becomes infinite, which means CO_2 cannot move from the starting cell to the ending cell. If the face is located either above, or below the depth to the center of the two adjacent cells, as shown in Figure 3.6(c) to (f), then the flow paths may be in a direction opposite to buoyant flow. The flow path in the direction opposite to buoyant flow should have a disadvantage because it is hard to migrate against gravity. Thus in these cases, Δh is the negative depth difference between the face and the cell center instead of the depth difference between the starting and ending cells.



(e) $z_f > z_1 > z_2$ (f) $z_f > z_2 > z_1$

Figure 3.6 Examples of calculating Δh in a 2D cross section

Flow in the up-dip direction is preferred more than in the down-dip direction in a CO_2 migration because of strong buoyancy effect. However, a small down-dip angle such as 1° does not make significant changes in CO_2 plume migrations in most of aquifers that have a much larger horizontal extent than the vertical extent. Although a CO_2 plume can migrate to small down-dip angles, those can make the potential term, $\Delta\Phi = \Delta P_{ij} + \Delta(\rho g h)_{ij}$, in Equation (3-3) negative. The negative potential term is adjusted using Equation (3-7) and Equation (3-8) in order to allow the down-dip flow of CO_2 plume. In Equation (3-7) and Equation (3-8), a dip angle of two adjacent grid cells is denoted as α. α_{min} is the minimum down-dip angle ignoring the buoyant force and α_{max} is the maximum down-dip angle factoring the buoyant force. If a down-dip angle is less than α_{max}, then the effect of the negative buoyant force is reduced as shown in Equation (3-7) and Equation (3-8). ω is the power in the power model of Equation (3-7) to control the shape of the interpolated curve. The effect of the negative buoyant force varies according to the structure in vicinity of a cell, but it is simply interpolated as formulated in Equation (3-7). Figure 3.7 shows the change in the down dip factor as a function of ω. ω = 1 represents the linear regression between α_{min} and α_{max}. α_{min}, α_{max}, and ω are aquifer-structure-dependent parameters. They should be adjusted when SCA does not reproduce the down dip migration in a full physics simulation. However, good results are obtained in the examples of this study with α_{min} = 1°, α_{max} = 5°, and ω = 0.5.

$$down\ dip\ factor = \begin{cases} 0, & 0 < \alpha \leq \alpha_{min} \\ \left(\dfrac{\alpha - \alpha_{min}}{\alpha_{max} - \alpha_{min}}\right)^{\omega}, & \alpha_{min} < \alpha < \alpha_{max} \\ 1, & otherwise \end{cases} \quad \cdots\cdots\cdots\cdots \text{Equation (3-7)}$$

$$\Delta\Phi = \Delta P_{ij} + down\ dip\ factor \cdot \Delta(\rho g h)_{ij} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \text{Equation (3-8)}$$

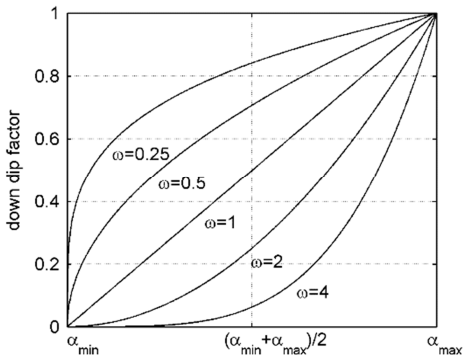


Figure 3.7 Down dip factor as a function of ω

As mentioned in the previous section, the modified Dijkstra's algorithm rejects paths that have connectivity greater than a certain time. This time is different from an injection period because the edge weights are calculated using approximate parameters such as viscosity, density, and pressure gradients. This cutoff time is obtained by adding 50% of the injection duration in order to account for errors that the approximations results in. The modified Dijkstra's algorithm therefore ignores paths that have the connectivity larger than one and half times of an injection period.

In the last step, grid cells are sorted in the order of their scaled connectivity. Each grid cell can have CO_2 as much as the product of its volume, porosity, and average CO_2 saturation ($\overline{S_{CO_2}}$). The grid cells are filled with CO_2 until the cumulative CO_2 volume is equal to the total volume of injected CO_2 in aquifer conditions. $\overline{S_{CO_2}}$ should be adjusted to better CO_2 plume migration computed by a full physics simulator in one of many probable geological models. A small $\overline{S_{CO_2}}$ results in a large CO_2 plume in SCA while a large $\overline{S_{CO_2}}$ results in a small CO_2 plume. Several $\overline{S_{CO_2}}$ are attempted until a SCA

approximation matches an observed CO_2 plume. Good matches are obtained in most of the examples of this study with the CO_2 saturation at the end point in a CO_2 relative permeability curve for $\overline{S_{CO_2}}$.

Before SCA is applied to a large suite of geological models, $\bar{k}$, $\overline{S_{CO_2}}$, $\alpha_{\min}$, $\alpha_{\max}$, and ω are adjusted for a few small and synthetic models. $\bar{k}$ that is influenced by the degree of heterogeneity is manually adjusted unless an SCA approximation computed using an arithmetic average of permeability is in good agreement with a full physics simulated CO_2 plume migration. $\overline{S_{CO_2}}$ should be increased or decreased when the size of the CO_2 plume in a SCA approximation is large or small compared to its full physics simulation, respectively. The aquifer-structure-dependent parameters should be adjusted when SCA does not reproduce the down dip migration in a full physics simulation.

3.2. VERIFICATION OF SCA

The objective of SCA is to quickly approximate CO_2 plume migrations rather than to delineate it as accurately as possible in heterogeneous porous media. In this section, the application of SCA for approximating CO_2 plume migrations is verified under different operating conditions, fluid properties, aquifer structures, and rock heterogeneity. SCA is also compared against a full physics simulator (CMG-GEM) and other proxies for CO_2 plume migrations.

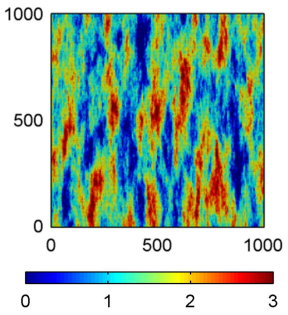
3.2.1. Base case

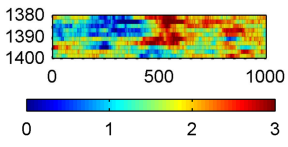
The aquifer shown in Figure 3.2 is used as a base case. The input data of SCA shown in Table 3.4 are used for the base case. The initial average pressure and temperature of the aquifer is 18,145 kPa and 75 °C, respectively. The fluid properties are obtained from the CMG-GEM results at the injector or they are calculated corresponding to the conditions at the injector. The permeability shown in Figure 3.8 follows a log

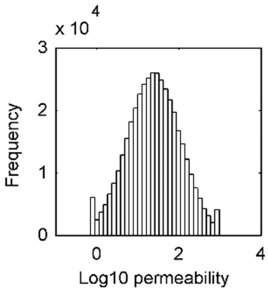
normal distribution with mean 1.3880 and standard deviation 0.6376 of $\log_{10}$ permeability. As shown in Figure 3.8, the major correlation direction is north-south. The major, minor and vertical correlation lengths are 250 m, 50 m and 15 m, respectively. The permeability is generated using the sequential Gaussian simulation in SGeMS (Remy et al., 2011). The porosity is assumed to be constant at 0.2. Capillary pressure is ignored and Figure 3.9 shows the relative permeability curves of brine and CO_2. The boundary is connected to an infinite aquifer.

Table 3.4 Input data of SCA for the base case

Aquifer structure	Size	1,005 m * 1,005 m * 20 m
	Dimensions of a grid cell	5 m * 5 m * 2 m
Operating conditions	Injection period	2 years
	Injector location	(i, j, k) = (101, 101, 3)
	Injection rate	11214.5 kg/day
Rock and fluid properties in aquifer conditions	Average permeability	67 md
	Average porosity	0.2
	CO_2 saturation at the end point	0.7
	CO_2 relative permeability at the end point	0.8
	Average CO_2 viscosity	0.0332 cp
	Average CO_2 density	411.3 kg/m^3
	Average brine viscosity	0.3786 cp
	Average brine density	1007 kg/m^3
	Average CO_2 saturation	0.7
Others	Minimum downdip	1°
	Maximum downdip	5°
	Power of downdip factor	0.5







(a) Areal view (b) Slice at i=101 (c) Histogram of $\log_{10}$ permeability

Figure 3.8 Different views and histogram of $\log_{10}$ permeability (md) field shown in Figure 3.2. The axis unit in (a) and (b) is in meters. The z-axis is exaggerated by a factor of 10.

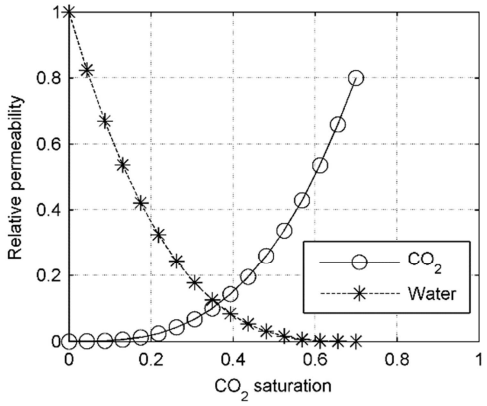


Figure 3.9 Relative permeability of CO_2 and water in the base case

Table 3.5 shows scaled connectivity and connectivity computed by SCA for the base case. Table 3.6 shows CO_2 plume migration computed using SCA and CMG-GEM in the base case. In the CMG-GEM results, CO_2 migrates longer in the north-south direction than the east-west direction along the major correlated direction of permeability. This implies the spatial distribution of permeability is the main factor that determines CO_2 plume migration. SCA presents an approximation of CO_2 plume horizontal migration similar to the full physics simulation results. The wedge-shaped migration of CO_2 plume in the slice at i = 101 of Table 3.6 is reproduced in the results of SCA. SCA does not capture the small horizontal migration in the second layer, but the SCA proxy does perform adequately to quantify the uncertainty in dominant migration paths.

Table 3.5 Scaled connectivity and connectivity in the base case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	Scaled connectivity	Connectivity, years

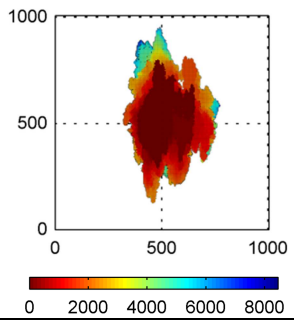
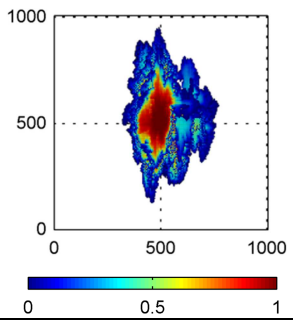
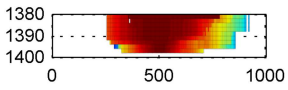
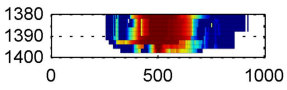
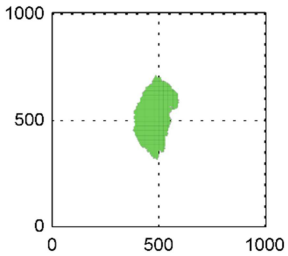
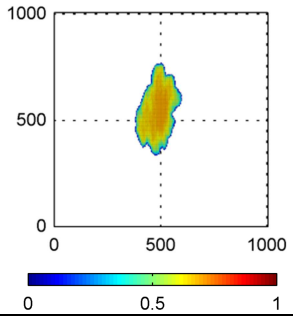
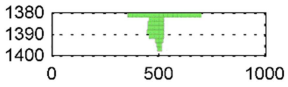
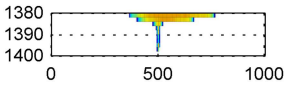
Areal view		
Slice at i = 101		

Table 3.6 Comparison of CO_2 migration by SCA and CMG-GEM in the base case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG (S_{CO_2})
Areal view		
Slice at i = 101		

3.2.2. Injection period

Table 3.8 shows the CO_2 plume migration after CO_2 is injected during 1 year while the injection period is 2 years in the base case. SCA reproduces the CO_2 plume migration that is horizontally shorter than in the base case.

Table 3.7 Comparison of CO_2 migration by SCA and CMG-GEM for 1 year. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})

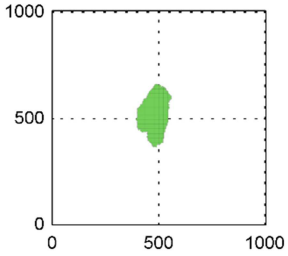
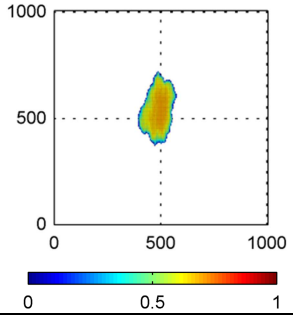
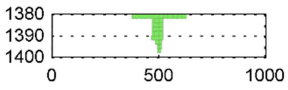
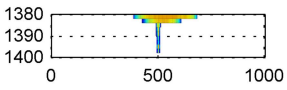
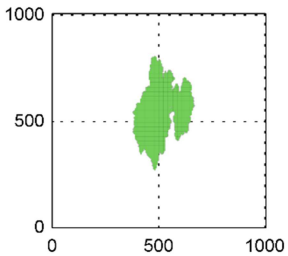
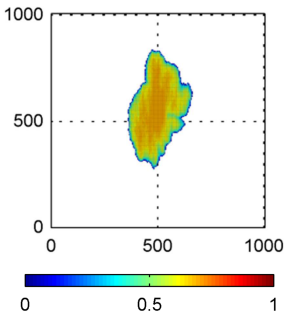
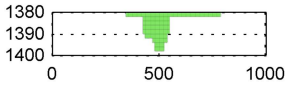
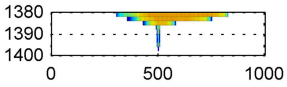
Areal view		
Slice at i = 101		

Table 3.8 shows the CO_2 plume migration after CO_2 is injected for 4 years. 80 md is used in this case for $\bar{k}$ in the approximate pressure gradient equation (Equation (3-4)) instead of 67 md, which is the arithmetic average. Because Equation (3-4) is an approximation of the spatially and temporally different pressure gradients, $\bar{k}$ should be adjusted if operating conditions such as an injection period are changed.

In Table 3.8, the east part of the SCA approximation does not have the CO_2 plume migration in the south direction. This is because the impermeable zone in the east part shown in Figure 3.8 impedes the CO_2 flow. However, the overall shape of the SCA approximation and the CMG-GEM result is still similar.

Table 3.8 Comparison of CO_2 migration by SCA and CMG-GEM for 4 years. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})
Areal view		

Slice at i = 101		

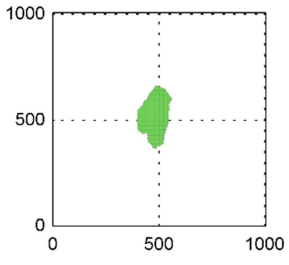
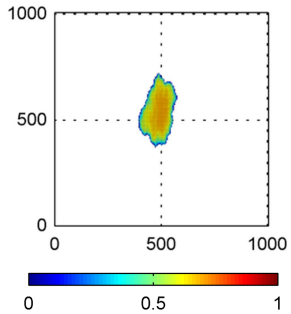
3.2.3. Injection rate

Table 3.9 and Table 3.10 show the results of SCA and CMG-GEM corresponding to half and double of the CO_2 injection rate in the base case, respectively. In the double injection rate case, 80 md is used for $\bar{k}$ in Equation (3-4) instead of the arithmetic average of permeability.

The horizontal migrations approximated by SCA in Table 3.9 and Table 3.10 are slightly different from those of the CMG-GEM results, but SCA captures the essence of the horizontal migrations of CO_2 plumes at the half and double injection rates.

In Table 3.9, the wedge-shaped vertical migration in the CMG-GEM result is reproduced in the SCA approximation. As shown for the slice i=101 in Table 3.6 and Table 3.10, the double injection rate case has longer horizontal migration in each layer than the base case. This is because the higher injection rate leads to higher pressure gradient. In the SCA result for the double injection rate case, CO_2 also spreads longer horizontally in each layer.

Table 3.9 Comparison of CO_2 migration by SCA and CMG-GEM for half injection rate case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})
Areal view		

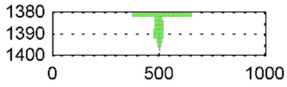
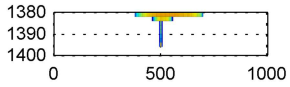
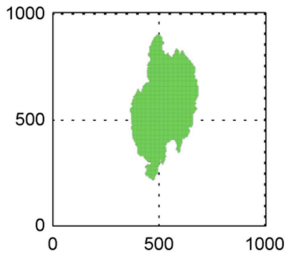
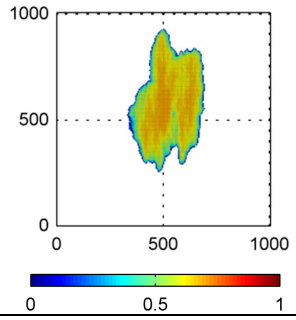
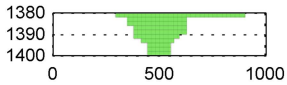
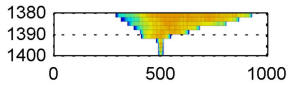
Slice at i = 101		

Table 3.10 Comparison of CO_2 migration by SCA and CMG-GEM for double injection rate case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

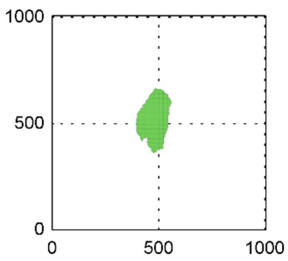
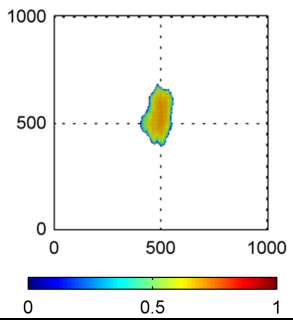
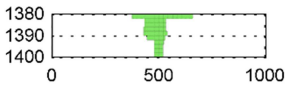
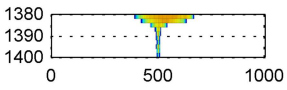
	SCA	CMG-GEM (S_{CO_2})
Areal view		
Slice at i = 101		

3.2.4. Density and viscosity

In this case, SCA and CMG-GEM results are compared when the aquifer has the higher initial pressure (28145 kPa) than the base case (18145 kPa). If the initial pressure increases, then the density and viscosity of CO_2 increase. The density and viscosity of CO_2 for the SCA input data increase from 411.3 kg/m^3 and 0.0332 cp used in the base case to 735.7 kg/m^3 and 0.0622 cp. The size of the CO_2 plume in this case is smaller than that in the base case because the total volume and the mobility of the injected CO_2 decrease. SCA reproduces such the spatial characteristics of the small CO_2 plume for this case as shown in Table 3.11.

Table 3.11 Comparison of CO_2 migration by SCA and CMG-GEM for higher CO_2 density and viscosity. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

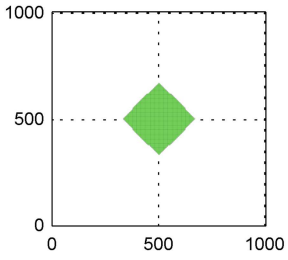
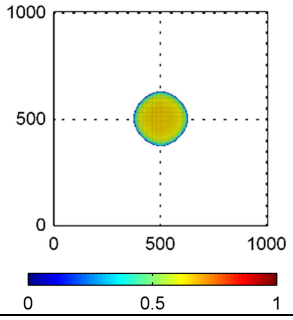
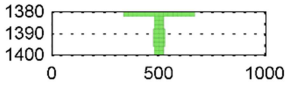
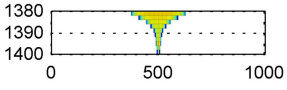
	SCA	CMG-GEM (S_{CO_2})

Areal view		
Slice at i = 101		

3.2.5. Rock heterogeneity

Table 3.12 shows CO_2 migration in homogeneous rock that has the mean of $\log_{10}$ permeability as the base case. SCA reproduces the buoyant flow in the cross-sectional view. However, SCA fails to reproduce the radial migration of the CO_2 plume because it does not consider diagonal connections such as the north-east/south-west directions. It is not appropriate to apply SCA to analyze displacement in homogeneous rock, but the goal of this study is to quantify the uncertainty in CO_2 displacement in heterogeneous rock.

Table 3.12 Comparison of CO_2 migration by SCA and CMG-GEM for homogeneous rock. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})
Areal view		
Slice at i = 101		

However, permeability in most of reservoirs is not homogeneous. The most common measure of permeability heterogeneity is V_{DP}, the Dykstra-Parsons coefficient (Jensen, 2000):

$$V_{DP} = \frac{k_{0.50} - k_{0.16}}{k_{0.50}} \quad \text{Equation (3-9)}$$

where $k_{0.50}$ is the median of permeability and $k_{0.16}$ is the 0.16 quantile of permeability. V_{DP} in most reservoirs ranges from 0.5 to 0.9 where $V_{DP} = 0$ means homogeneous (Willhite, 1986).

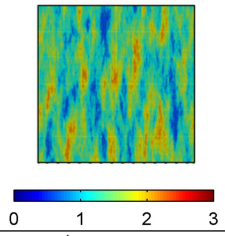
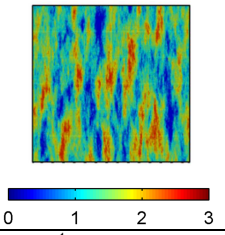
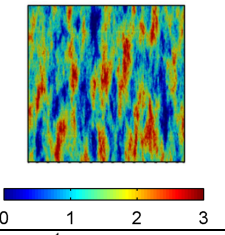
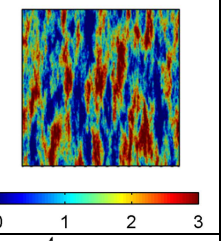
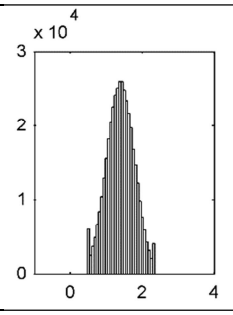
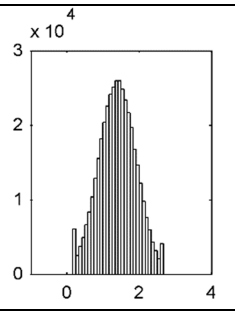
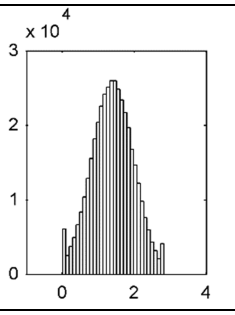
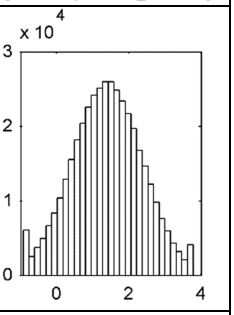
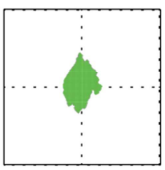
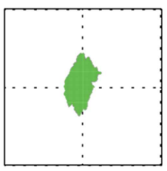
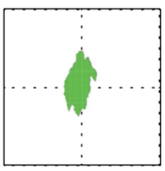
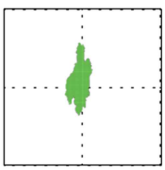
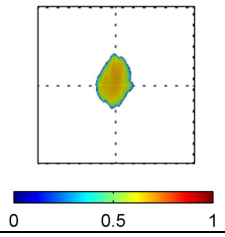
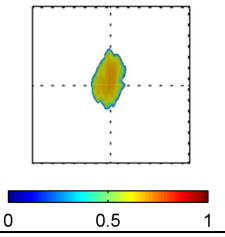
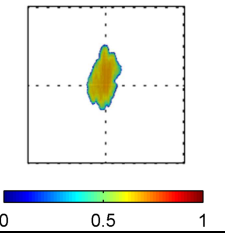
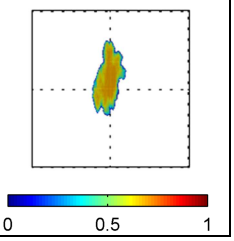
Now SCA is tested in four heterogeneous permeability fields. The four heterogeneous permeability fields are generated by normalizing $\log_{10}$ permeability of the base case and back-transforming them with the same mean and the different standard deviations, 0.5, 0.7, 0.9, and 1.5 times of the standard deviation of the base case. The 2nd and 3rd rows in Table 3.13 show the areal view and histogram of $\log_{10}$ permeability in the four permeability fields. V_{DP} of the four permeability fields are 0.52, 0.65, 0.73, and 0.89 and V_{DP} in the base case is 0.77.

The 4th and 5th rows in Table 3.13 are the arithmetic averages and adjusted $\bar{k}$ of the four permeability fields. As discussed before, $\bar{k}$ in the approximate pressure gradient equation (Equation (3-4)) should be adjusted if reservoir model parameters are changed. The arithmetic average of an entire permeability field is a good reference for $\bar{k}$. However, the arithmetic average is not always a good representation of heterogeneous rock because the behaviors of produced or injected fluid are not only mostly affected by permeability near a well, but the spatial distribution of permeability is also complicated.

The CMG-GEM results (the 7th row) in Table 3.13 show the CO_2 plume migrations according to V_{DP}. As V_{DP} becomes higher, the CO_2 plume spreads longer in the north-south direction and it is more affected by permeability heterogeneity. As shown

the areal views of SCA and CMG-GEM in Table 3.13, SCA reproduces the overall trend in the shape of the CO_2 plume simulated by CMG-GEM. In this regard, SCA fails to reproduce the radial migration of the CO_2 plume in the homogeneous model, but SCA is successful in capturing the spatial characteristics of the CO_2 plumes in the heterogeneous models whose V_{DP} lie in the normal range.

Table 3.13 Comparison of CO_2 migration by SCA and CMG-GEM for four heterogeneous permeability fields

	V_{DP}=0.52	V_{DP}=0.65	V_{DP}=0.73	V_{DP}=0.89
Log_{10} k (md)				
Histogram of log_{10} k (md)				
Arithmetic k	32 md	41 md	56 md	207 md
Adjusted $\bar{k}$	50 md	50 md	65 md	Not adjusted
Areal view of SCA				
Areal view of CMG-GEM (S_{CO_2})				

3.2.6. Cap rock

As shown in Figure 3.10, the 3th layer from the top is modeled as a shale barrier in this variation of the base case. It is investigated if SCA reproduces a CO_2 leakage through the shale barriers according to the permeability of the shale barriers.

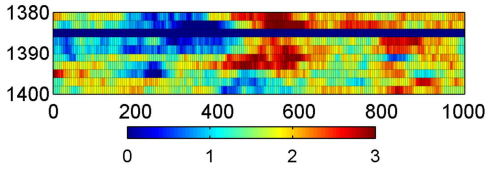
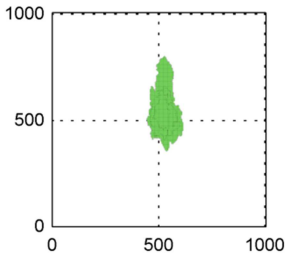
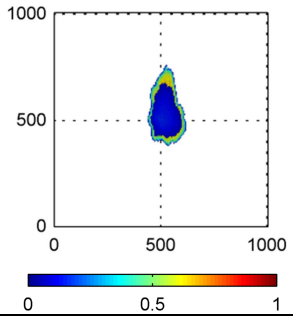
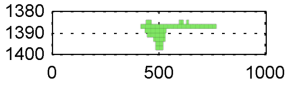
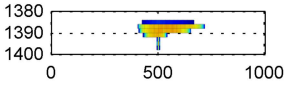
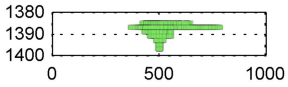
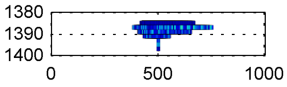


Figure 3.10 Cross-sectional view (i=101) of $\log_{10}$ permeability (md) in the base case with shale barriers. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

Table 3.14 shows the CO_2 migration when the shale barrier has a permeability of 0.01 md. Even though the permeability of the shale barrier is small, CO_2 migrates through the shale barrier. The CO_2 saturation above the shale barriers is small, but considering the short injection period (2 years), CO_2 cannot be confined safely below the shale barrier. SCA captures the horizontal extent of the plume migration in the top layer as well as the vertical migration of the CO_2 plume.

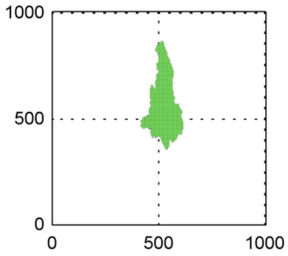
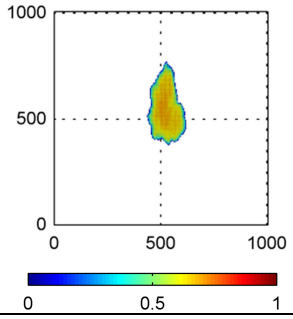
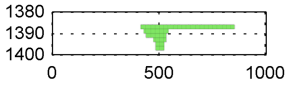
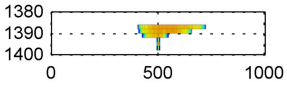
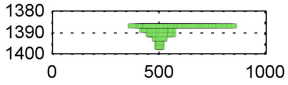
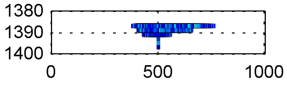
Table 3.14 Comparison of CO_2 migration by SCA and CMG-GEM for shale barriers with 0.01 md. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})
Areal view		
Slice at i = 101		

Side view from east		

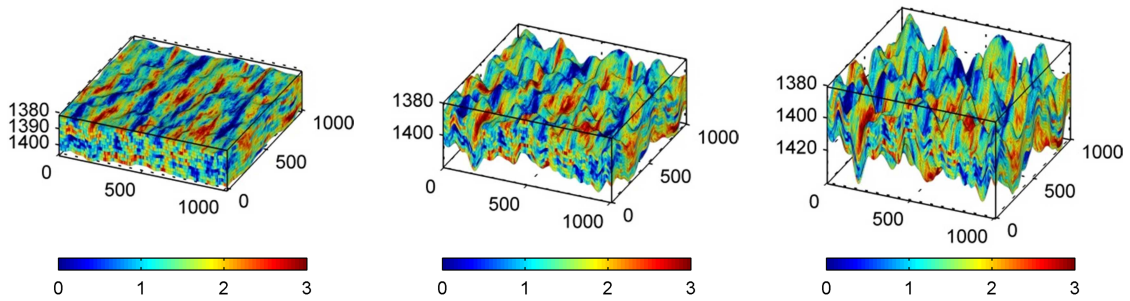
Reducing the permeability of the shale barrier to 0.001 md prevents a CO_2 leakage as shown in Table 3.15. SCA represents the vertical migration as well as the horizontal migration of the CO_2 plume.

Table 3.15 Comparison of CO_2 migration by SCA and CMG-GEM for shale barriers with 0.001 md. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})
Areal view		
Slice at i = 101		
Side view from east		

3.2.7. Topology of the aquifer

SCA is applied to three folded aquifer layers where amplitudes in the vertical direction are added to the structure in the base case. Figure 3.11 shows the $\log_{10}$ permeability distribution in the three structures. The spatial distribution of permeability is identical to that in the base case. The medium and strong amplitudes in Figure 3.11(b) and Figure 3.11(c) are four times and eight times larger than the weak amplitudes in Figure 3.11(a), respectively.

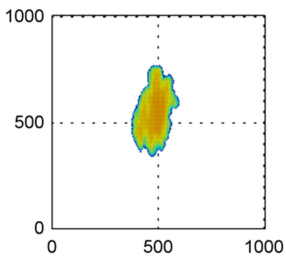
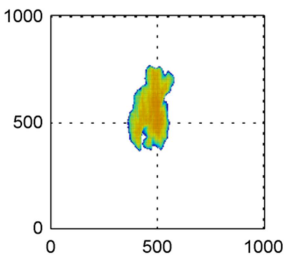
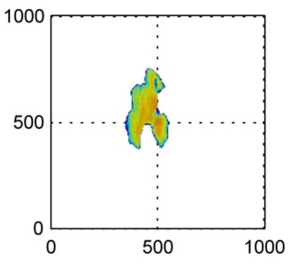


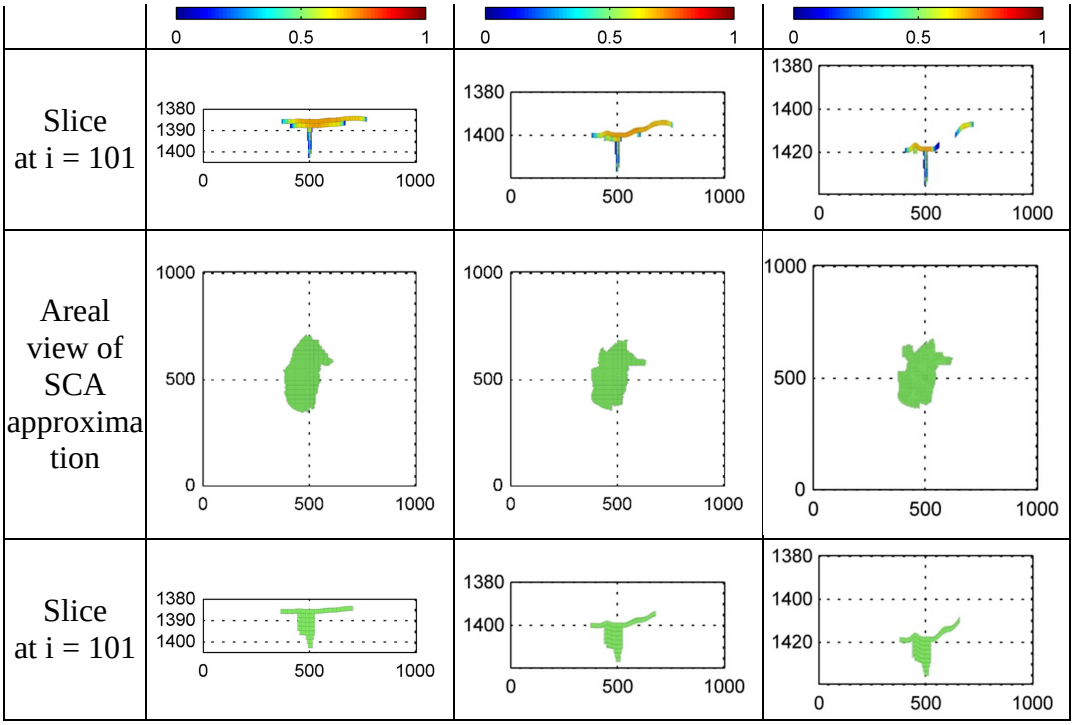
(a) weak amplitudes (b) medium amplitudes (c) strong amplitudes

Figure 3.11 Log_{10} permeability (md) in three aquifer models with variable topology. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

Compared to the base case, the weak amplitudes in the vertical direction shown in Figure 3.11(a) do not make significant variation in the CO_2 plume migration as shown in the CMG-GEM results for the weak amplitudes of Table 3.16. The CO_2 plumes for the base case and the case with medium amplitudes have slightly different shape in the south and east regions, but they are still almost the same despite stronger amplitudes. This implies that the structural patterns shown in Figure 3.11(a) and Figure 3.11(b) do not have strong influence on the CO_2 plume migrations. For the weak and medium amplitudes, SCA does not reproduce the exact frontal shapes of the CO_2 plumes, but the SCA approximations are similar to the CMG-GEM results in the vertical and horizontal migrations.

Table 3.16 Comparison of CO_2 migration by SCA and CMG-GEM for the three aquifer structures shown in Figure 3.11. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	weak amplitudes	medium amplitudes	strong amplitudes
Areal view of S_{CO_2} by CMG-GEM			



Compared to the weak and medium amplitudes, the CO_2 plume migration computed using CMG-GEM for the strong amplitudes is quite different from that for the base case as shown in Table 3.16. This means the structural pattern in the strong amplitudes has stronger influence on the CO_2 plume migration than those in the weak and medium amplitudes. SCA reproduces the vertical migration and the longer horizontal migration along the north-south direction, but it does not reproduce the frontal shapes exactly compared to the weak and medium amplitudes. In this regards, SCA might not work properly when structural heterogeneity has significant influence on a CO_2 plume migration. Unfortunately, we cannot know if SCA works for a complex structure before applying because there is no direct way to measure how much influence structural heterogeneity has on a CO_2 plume migration. However, the final goal of SCA is to provide approximate CO_2 plumes for a large suite of geological models. We can judge if

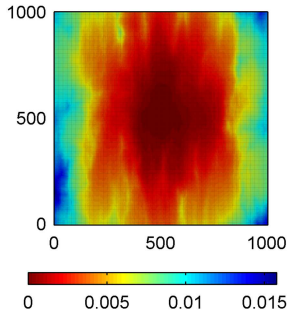
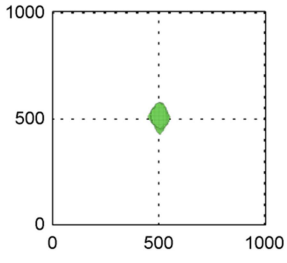
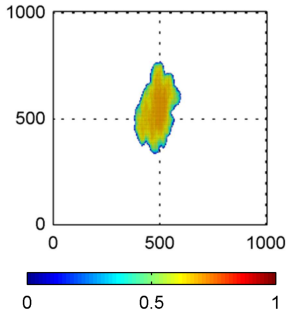
SCA works for the complex structure by comparing the results of SCA and a full physics simulator in one model of the suite.

3.2.8. Other proxies

In this case, two proxies, the fast marching method (FMM) and a vertical equilibrium (VE) model, are applied to the base case and the case with the shale barrier, and then their results are compared with those using CMG-GEM. For the VE model, the co2lab module in the Matlab Reservoir Simulation Toolbox (MRST) is used (Lie et al., 2011).

In the left column in Table 3.17, the diffusive time of flight to every grid cell is calculated using FMM. The diffusive time of flight implies the arrival time of a pressure wave (Kulkarni et al., 2001). The diffusive time of flight does not imply the arrival time of injected CO_2, but here we investigate if FMM can consider the buoyant effect. The grid cells are filled with CO_2 in the order of the diffusive time of flight as much as the volume of injected CO_2. As shown in Table 3.17, FMM does not reproduce the buoyant flow of CO_2 plume because it is not designed to consider the buoyancy effect between two fluids.

Table 3.17 CO_2 migration by FMM in the base case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	Diffusive time of flight years	Approximation of CO_2 plume	CMG-GEM (S_{CO_2})
Areal view			

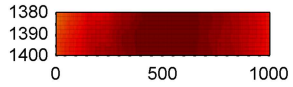
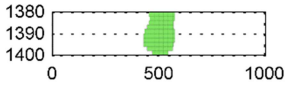
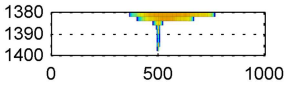
Slice at i = 101			

Table 3.18 and Table 3.19 show the migration of CO_2 plume calculated using VE and CMG-GEM for the base case and the shale barrier case, respectively. In Table 3.18, the result of VE for the base case is acceptable because the VE model is able to correctly represent the buoyant migration of the CO_2 plume, and the horizontal migration due to permeability heterogeneity. However, VE overestimates CO_2 saturations in the east-southern region of the reservoir compared to the result of CMG-GEM. The vertically averaged permeability shown in Table 3.20 is higher in the east-southern zone than the permeability of the top layer in the base case shown in Figure 3.8(a) and this causes the CO_2 plume computed using the VE model to spread more in the east-southern zone.

VE is applied to the cap rock case where the whole 3rd layer from the top has a shale barrier of permeability equal to 0.001 md as shown in Figure 3.10. The shale barrier has constant permeability and consequently, the vertically averaged permeability is similarly impacted at all locations in the reservoir. This evident in the map of the vertically averaged permeability shown in Table 3.20 where the observed permeability features look virtually the same. For this reason, the CO_2 migration computed by VE in the cap rock case is almost identical to that in the base case. VE is therefore unable to capture the influence of the shale barrier on the vertical migration of the CO_2 plume. In this regard, VE is inappropriate to represent the CO_2 plume migration affected by vertical heterogeneity.

Table 3.18 Comparison of CO_2 migration by VE and CMG-GEM for the base case. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	VE (S_{CO_2})	CMG-GEM (S_{CO_2})

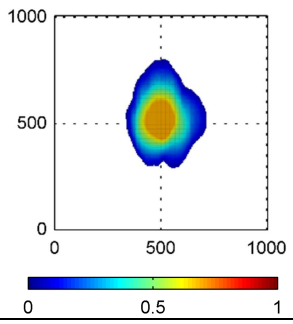
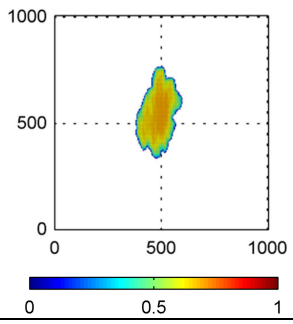
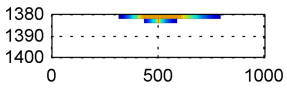
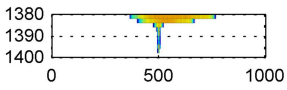
Areal view		
Slice at i = 101		

Table 3.19 Comparison of CO_2 migration calculated by VE and CMG-GEM for the cap rock case where there is a shale barrier of permeability 0.001md. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

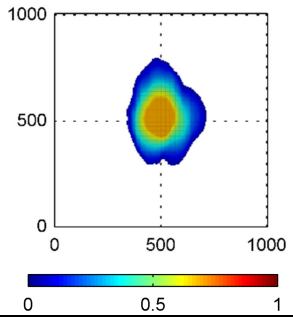
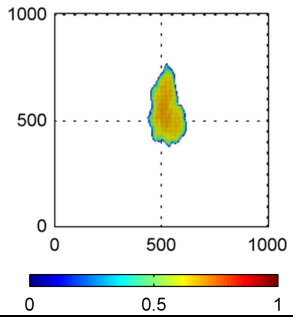
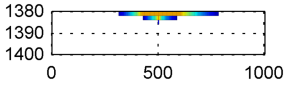
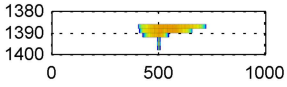
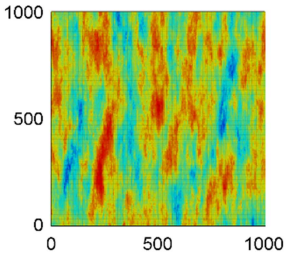
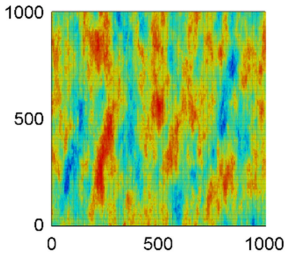
	VE (S_{CO_2})	CMG-GEM (S_{CO_2})
Areal view		
Slice at i = 101		

Table 3.20: Vertically averaged $\log_{10}$ permeability for the cases in Table 3.18 and Table 3.19. The unit of the axes is meters.

	Vertically averaged $\log_{10}$ peambility in the base case	Vertically averaged $\log_{10}$ peambility in the cap rock case
Areal view		

3.3. APPLICATION TO FIELD CASES

3.3.1. The Johansen model

The Johansen Formation is a candidate site for geological CO_2 storage located in the south-western coast of Norway (Eigestad et al., 2009). The data set for the Johansen Formation is available online. SCA is applied to approximate a CO_2 migration in the Johansen Formation and it is compared to the results of a full physics simulator, CMG-GEM.

The grid system of the Johansen model is 100 by 100 by 11. The typical cell size is 500 m by 500 m, and its thickness varies from 16 to 24 m. Figure 3.12 shows the structure and faults of the Johansen model where the red means a fault and the yellow means an active cell. A major fault is located in the middle of the Johansen model along the north-south direction. The boundaries which do not face a boundary fault are connected to an aquifer allowing brine displacement. The initial average pressure and temperature of the aquifer are 31,310 kPa and 94 °C, respectively.

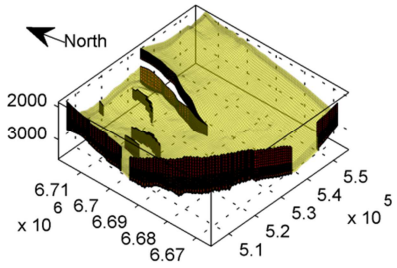
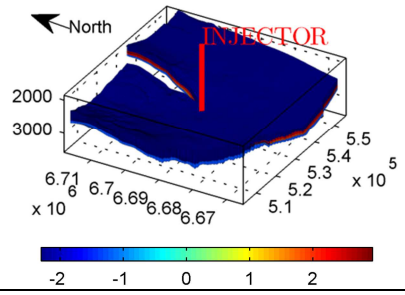
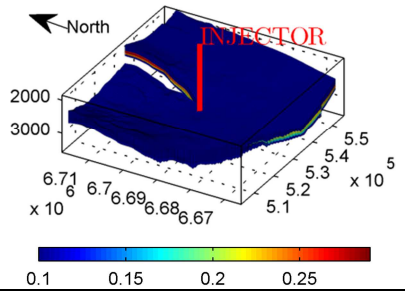
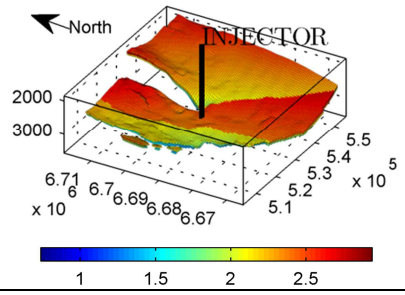
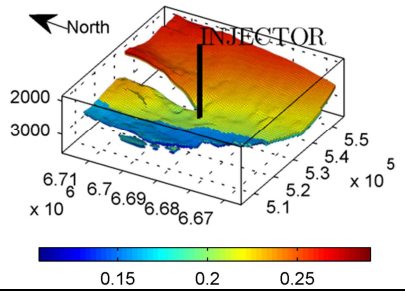


Figure 3.12 Structure and faults of the Johansen model. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

The Johansen model has shale layers at the top and bottom, and sandstone layers in the middle, which is the Johansen Formation. The first row in Table 3.21 shows $\log_{10}$ horizontal permeability and porosity of the entire 3D model including the shale layers at

the bottom and top. The ratio of vertical to horizontal permeability is 0.1. The horizontal permeability of the shale layers are 0.01 md and 0.1 md at the top and bottom, respectively. The porosity of the shale layers is constant, 0.1. The second row in Table 3.21 shows $\log_{10}$ horizontal permeability and porosity of sandstone in the Johansen Formation. The horizontal permeability and porosity of the Johansen Formation range from 64 to 1,660 md and from 0.10 to 0.29, respectively. CO_2 is injected at (48,48) over the Johansen Formation during 500 years.

Table 3.21 Log_{10} horizontal permeability and porosity of the Johansen model. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	Log_{10} horizontal permeability	Porosity
Entire		
Sandstone		

In the original Johansen model, the horizontal dimension of the grid cell containing the injector is the same to that of the other grid cells, 500 m by 500 m. This big grid cell size might cause inaccuracy of the CMG-GEM simulation results. Thus the grid cell containing the injector is refined to 31 by 31, and the refined dimension is 16.13 m by 16.13 m. Figure 3.13 shows the local refinement of the grid cell containing the

injector on the 6th layer. The grid cells containing the injector on the perforated layers (6th to 10th) are refined as shown in Figure 3.13.

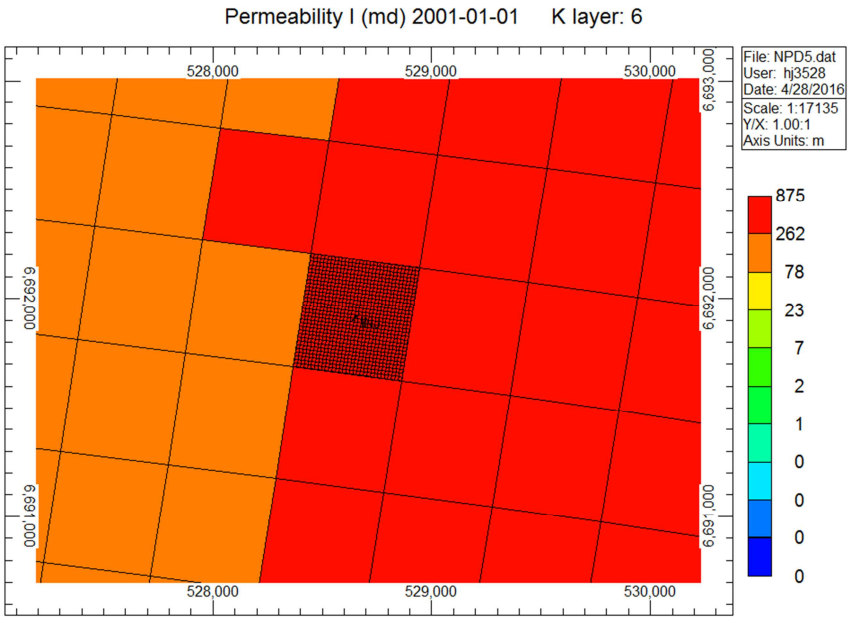


Figure 3.13 Local refinement of the grid cell containing the injector on the 6th layer in the Johansen model. The unit of the axes is meters.

Figure 3.14 shows the relative permeability curves of CO_2 and water in the Johansen model. The end point values in the relative permeability curve of CO_2 are used for $\hat{S}_{CO_2}$ and $\hat{k}_{r,CO_2}$ in Equation (3-3). Table 3.22 contains the input parameters for SCA. The fluid properties are obtained from the CMG-GEM results at the injector.

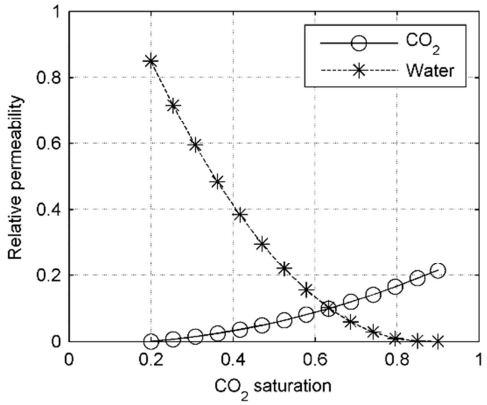


Figure 3.14 Relative permeability in the Johansen model

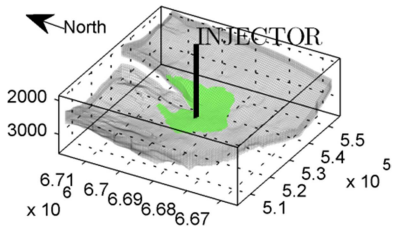
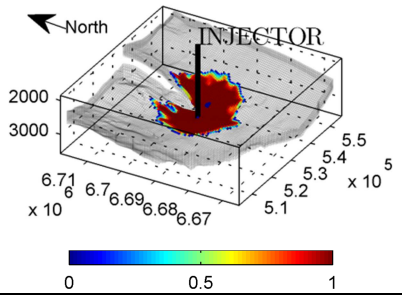
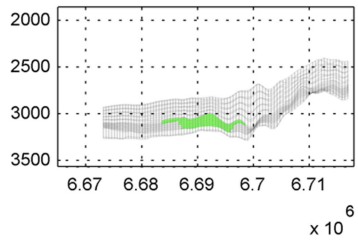
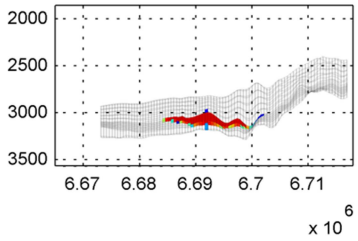
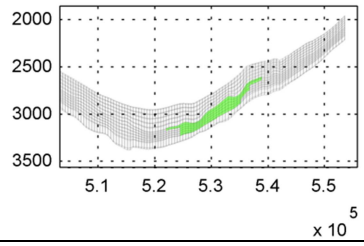
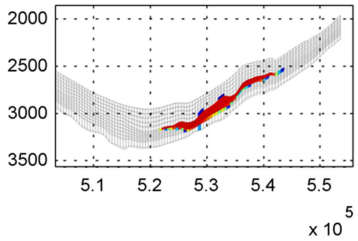
Table 3.22 Input data of SCA for the Johansen model

Aquifer structure	Average thickness	103.5194 m
Operating conditions	Injection period	500 years
	Injector location	(i, j, k) = (48, 48, 6) ~ (48, 48, 10)
	Injection rate	14,000 m^3/day in aquifer conditions
Fluid properties in aquifer conditions	CO_2 saturation at the end point	0.9
	CO_2 relative permeability at the end point	0.2142
	Average CO_2 viscosity	0.058 cp
	Average CO_2 density	700 kg/m^3
	Average brine viscosity	0.3027 cp
	Average brine density	1014.5 kg/m^3
	Average CO_2 saturation	0.9
Others	Minimum downdip	1°
	Maximum downdip	5°
	Power of downdip factor	0.5

The right column in Table 3.23 shows the 3D, top, and cross-sectional views of CO_2 migration after 500 years computed by CMG-GEM. SCA is able to reflect the migration of the CO_2 plume to both the eastern and western regions adjacent to the main fault. The CO_2 plume spreads to the longest extent in the east direction because the eastern regions are updip. The CO_2 plume in the updip direction will continue to move to the north and east directions because of buoyancy even though CO_2 is no longer injected. As shown in the slices at i = 48 and j = 48, SCA captures the spatial features of the CO_2 plume that migrates farthest away in the east direction.

Table 3.23 Comparison of CO_2 migration by SCA and CMG-GEM in the Johansen model. The unit of the axes is meters. The z-axis is exaggerated by a factor of 10.

	SCA	CMG-GEM (S_{CO_2})

3D view		
Slice at i = 48		
Slice at j = 48		

3.3.2. The Sleipner L9 model

The Sleipner L9 model is a benchmark model for the topmost layer of the Utsira formation in the Sleipner gas field where CO_2 has injected since 1999 (Singh et al., 2010). The data set for the Sleipner L9 model was released by Statoil in 2011. The approximation of the CO_2 migration in the Sleipner L9 model computed by SCA is compared with the full physics simulation result presented by Singh et al (2010).

The grid system of the Sleipner L9 model is 65 by 119 by 43. The lateral cell size is 50 by 50 m, and its thickness ranges from 0.0745 to 1 m. Table 3.24 shows the $\log_{10}$ horizontal permeability and porosity of the entire model and the sandstone layers. The sandstone layers are between the top and bottom shale layers. The shale layers act like no flow boundaries and the lateral boundaries have constant pressure (Bandilla et al., 2014).

The horizontal and vertical permeabilities range from 1692 to 1977 md and from 482 to 564 md, respectively. As shown in the sand permeability of Table 3.24, the shallower zone has higher porosity and permeability than the deeper zone.

Table 3.24 Log_{10} horizontal permeability and porosity of the Sleipner L9 model. The unit of the axes is meters. The z-axis is exaggerated by a factor of 20.

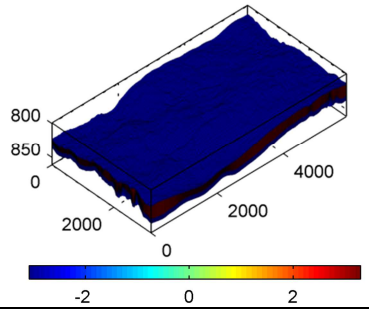
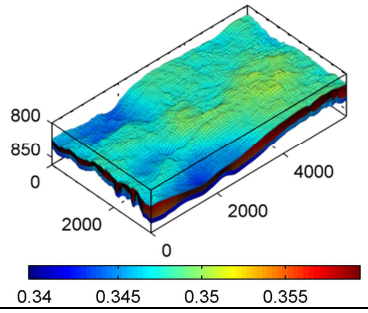
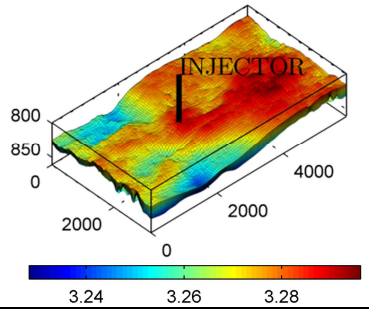
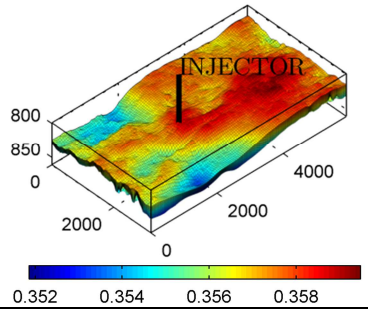
	Log_{10} horizontal permeability	Porosity
Entire		
Sandstone		

Table 3.25 includes the input data for SCA for the Sleipner L9 model. The initial aquifer pressure (105 bar) and the bottom hole temperature (48°C) that Alnes et al. (2011) presented are used to calculate the properties of the CO_2 plume within SCA. Especially, we used the bottom hole temperature instead of the initial aquifer temperature (35.5°C) for SCA because it takes some time for the injected CO_2 to cool down to the initial aquifer temperature. The CO_2 density gradually changes from 461.5 kg/m^3 at the bottom-hole temperature to 810.3 kg/m^3 at the initial aquifer temperature as CO_2 moves far from the injector. However, the CO_2 density (461.5 kg/m^3) at the bottom-hole temperature is used to reproduce the fast movement to the top layer of the buoyant CO_2

plume at the injector in SCA. Thus the viscosity and density of CO_2 and brine are thermodynamically calculated at the initial aquifer pressure and the bottom-hole temperature.

Table 3.25 Input data of SCA for the Sleipner L9 model for 2008

Aquifer structure	Average thickness	11.3 m
Operating conditions	Injection period	10 years
	Injector location	(i, j, k) = (36,75,36)
	Injection rate at the initial aquifer conditions	290 kg/day
	Total injected volume at the initial aquifer conditions	2453608 m^3
Fluid properties in aquifer conditions	CO_2 saturation at the end point	0.89
	CO_2 relative permeability at the end point	0.75
	Average CO_2 viscosity	0.035 cp
	Average CO_2 density	461.5 kg/m^3
	Average brine viscosity	0.567 cp
	Average brine density	993 kg/m^3
	Average CO_2 saturation	0.89
Others	Minimum downdip	1°
	Maximum downdip	5°
	Power of downdip factor	0.5

The solid line in Figure 3.15 shows the cumulative mass of the injected CO_2 that Singh et al. (2010) used in the Sleipner L9 model. The dashed line with asterisks in Figure 3.15 represents the average mass rates from 1999 to 2008, which are the cumulative masses divided by the injection periods. CO_2 plume migrations should be approximated using SCA for multiple times (1999, 2001, 2002, 2004, 2006, 2008), but the average mass rate is not constant and it continues to increase as the dashed line with asterisks in Figure 3.15 shows. However, pressure gradients are calculated at a constant flow rate because steady state is assumed in SCA. Thus one that represents the varying flow rate should be chosen for each year. When SCA was conducted with the average

mass rates corresponding to the multiple times, the SCA approximations had the longer and larger extent of the plumes. This is because the average mass rate causes the pressure gradients to be overestimated because the average mass rate increases sharply. Therefore the median of the average mass rates is used to calculate the pressure gradients instead of the average mass rate. The dashed line with squares in Figure 3.15 represents the median mass rates from 1999 to 2008 and they are the median values in the average mass rates from 1999 to their years. For example, the average mass rates are 114.16, 171.23, 190.26, and 228.31 for 1999, 2001, 2002, and 2004, respectively. The median mass rate for 2004 is the median of 114.16, 171.23, 190.26, and 228.31. The median mass rate for each year is used to calculate the pressure gradients in each year. The end point of the CO_2 relative permeability curve that Singh et al. (2010) provides is used for $\hat{S}_{CO_2}$ and $\hat{k}_{r,CO_2}$ in Equation (3-3).

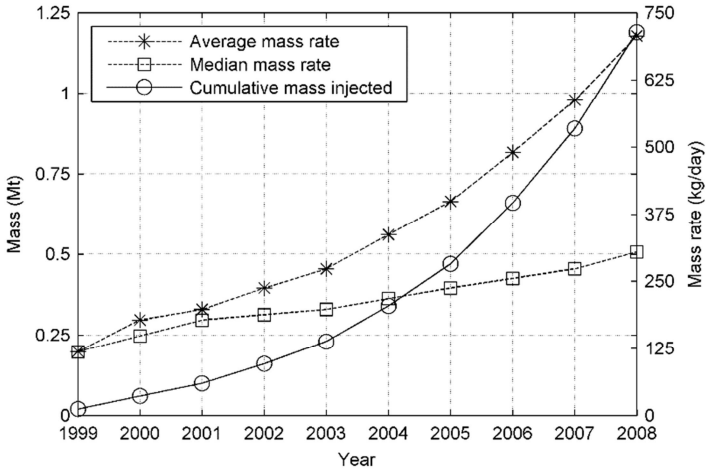
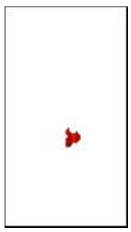
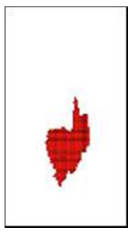
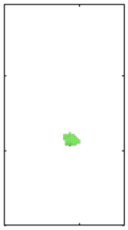
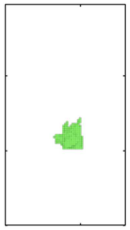
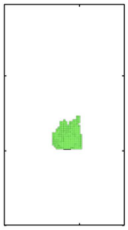
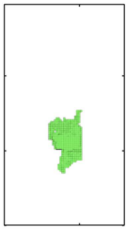
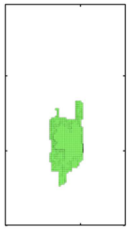
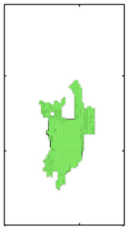


Figure 3.15 Cumulative masses, average mass rates, and median mass rates of injected CO_2 in the Sleipner L9 model from 1999 to 2008 (Singh et al., 2010)

Singh et al. (2010) provided the results of ECLIPSE 300 (©Schlumberger) simulation in the Sleipner L9 model from 1999 to 2008. Table 3.26 shows the results obtained using SCA for the Sleipner L9 model for the same duration. SCA reproduces the

migration of the CO_2 plume along the north-south direction observed in ECLIPSE 300 results presented by Singh et al. (2010). SCA captures the spatial features of the CO_2 plume that migrates to a longer extent in the north-east region than in the north-west region. In SCA, the migration to the north-west region is overestimated compared to the ECLIPSE 300 results. However, SCA presents similar approximations in terms of the shape and size of the CO_2 plume during all the periods compared to the ECLIPSE 300 results.

Table 3.26 CO_2 plume migrations computed by ECLIPSE 300 and SCA for the Sleipner L9 model from 1999 to 2008

ECLIPSE 300 results presented by Singh et al. (2010)					
					
SCA approximations					
					
1999	2001	2002	2004	2006	2008

3.4. CONCLUSIONS

The final goal of this study is to quickly quantify the uncertainty in the spatial and the temporal features of CO_2 plume migrations. However, it is impractical to run a full physics simulator on a large suite of geological models. We propose a fast alternative

method named SCA for approximating the CO_2 plume migrations, and to save the computational cost of quantifying the uncertainty in the CO_2 plume migrations.

In SCA, adjacent grid cells are assumed to be connected with edges and the edge weight are defined as a travel time. SCA scales the edge weights with their local minimum in order to reproduce a CO_2 plume flowing along the locally minimum edges because of strong buoyancy. The scaled connectivity that is the cost of the shortest path from an injector in the scaled graph is used to determine the sequence in which grid cells are filled with CO_2.

SCA was validated by changing several conditions in the base case. Results that were close to the full physics responses were obtained corresponding to variations in an injection period, an injection rate, a CO_2 density, a CO_2 viscosity, the degree of rock heterogeneity, permeability of a cap rock. SCA presented better approximations than CA, the fast marching method, and the vertical equilibrium model in the heterogeneous cases. However, SCA does not work properly when the rock heterogeneity is weak or the aquifer structure has an overwhelming influence on a CO_2 plume migration.

SCA was applied to two real field cases, the Johansen model and the Sleipner L9 model. In both of the cases, SCA captured the spatial and temporal characteristics of the CO_2 plumes.

Chapter 4. Quantification of Uncertainty in CO_2 Migration Using SCA

The final goal of this study is to quickly quantify the uncertainty in the spatial features of CO_2 plumes over a large suite of heterogeneous models. However, it takes a tremendous computation cost to run a full physics simulator in the large suite of the models. The uncertainty can be quickly quantified through a ranking process using a proxy (Ballin et al., 1992; Deutsch and Srinivasan, 1996; Scheidt and Caers, 2009). A fast proxy, SCA, is used to compute CO_2 plume migrations in the large suite instead of the full physics simulator. The models are ranked based on the extent of the approximated CO_2 plumes and then they are sampled with an interval. The sampled models are considered as the representatives of the entire suite. The uncertainty in the spatial features of CO_2 plumes are quickly quantified by conducting full physics simulations in the representative models instead of the entire suite.

The procedure of the ranking process using SCA is as follows:

i) Approximate the CO_2 plume migration in a suite of models using SCA

ii) Calculate the extent of the CO_2 plumes in the north, south, east, and west directions in the proxy results

iii) Rank and sort the models based on the proxy extent of the CO_2 plumes

iv) Sample representative models with an interval

v) Run a full physics simulator in the representative models

vi) Calculate the P10, P50, P90 quantiles of the extent of the CO_2 plumes in the north, south, east, and west directions based on the full physics simulation results in the representative models

The ranking process is applied to two cases: a variogram based modeling case and an object modeling case. Geological models for the two cases are generated based on a

variogram based modeling technique and an object modeling technique, respectively. SCA is applied to quantify the uncertainty in the extent of CO_2 plumes in geological models generated using two different geostatistical techniques.

4.1. THE JOHANSEN MODEL

In the Johansen model, only geologic models for the sandstone layers are developed while keeping the shale properties fixed. Figure 4.1 shows the scatter plot in $\log_{10}$ scale between porosity and horizontal permeability on the data for the Johansen model. There are 6 linear relationships between the logarithm values of porosity and horizontal permeability. However, the lithology for each linear relationship is not given, so one of the six trends shown in Figure 4.1 is used to describe the relationship between porosity and permeability. The irreducible water saturation of the selected trend is assumed to be constant and the relationship is described as $\mathrm{k} = 1.2589\mathrm{E}6\phi^{5.5}$.

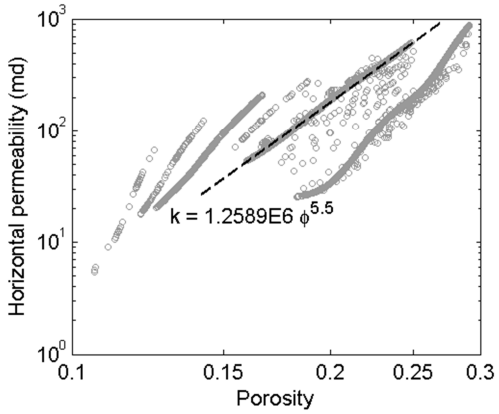
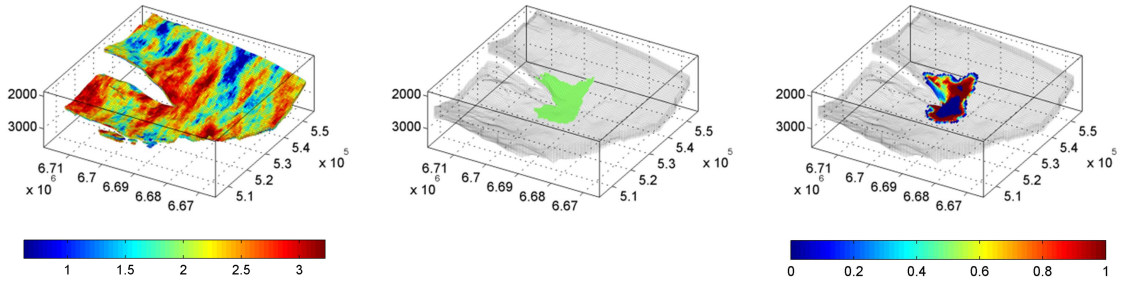


Figure 4.1 Scatter plot and linear regression of porosity and horizontal permeability in the given Johansen model data

200 realizations for the porosity are generated using the sequential Gaussian simulation in SGeMS and then the permeability is calculated using the selected transform. The porosity follows a normal distribution with mean 0.2 and standard deviation 0.04. The minimum and maximum values of porosity are 0.1 and 0.3,

respectively. The porosity is correlated in the east-west direction with anisotropy ratio of 2.5:1 and a correlation length of 10 grid blocks.

Figure 4.2 shows the results of SCA and CMG-GEM for one of the 200 models. In Figure 4.2(c), most of the injected CO_2 plume migrates to the north in the region to the right. This is because the strong updip in the region causes the buoyant CO_2 flow. SCA reproduces this spatial feature of the CO_2 plume.



(a) Horizontal permeability in the sandstone layers

(b) SCA result

(c) CO_2 saturation computed by CMG-GEM

Figure 4.2 Horizontal permeability, SCA result, and CO_2 saturation computed by CMG-GEM for one of the 200 models

The extent of the CO_2 plume in the north, south, east, and west directions are measured in order to rank the 200 models. The 200 models Figure 4.3 shows the travel distance of the CO_2 plume from the injector to the north, south, east, and west directions calculated by CMG-GEM and SCA. The correlations between these results calculated over the ensemble of the 200 models are 0.73, 0.70, 0.83, and 0.30, respectively. In the CMG-GEM results, the extent of the CO_2 plume in each direction is calculated in the zone that has CO_2 saturations greater than 0.01. As the results indicate, there is some disparity between the CMG-GEM and SCA results regarding the plume migration to the west. Even though the correlation between the CMG-GEM and SCA results is low, the absolute difference in the predictions made by CMG-GEM and SCA regarding the

displacement of CO_2 in the west direction is 1 km as shown in Figure 4.3(d) and this is below the resolution of what can be detected using SCA.

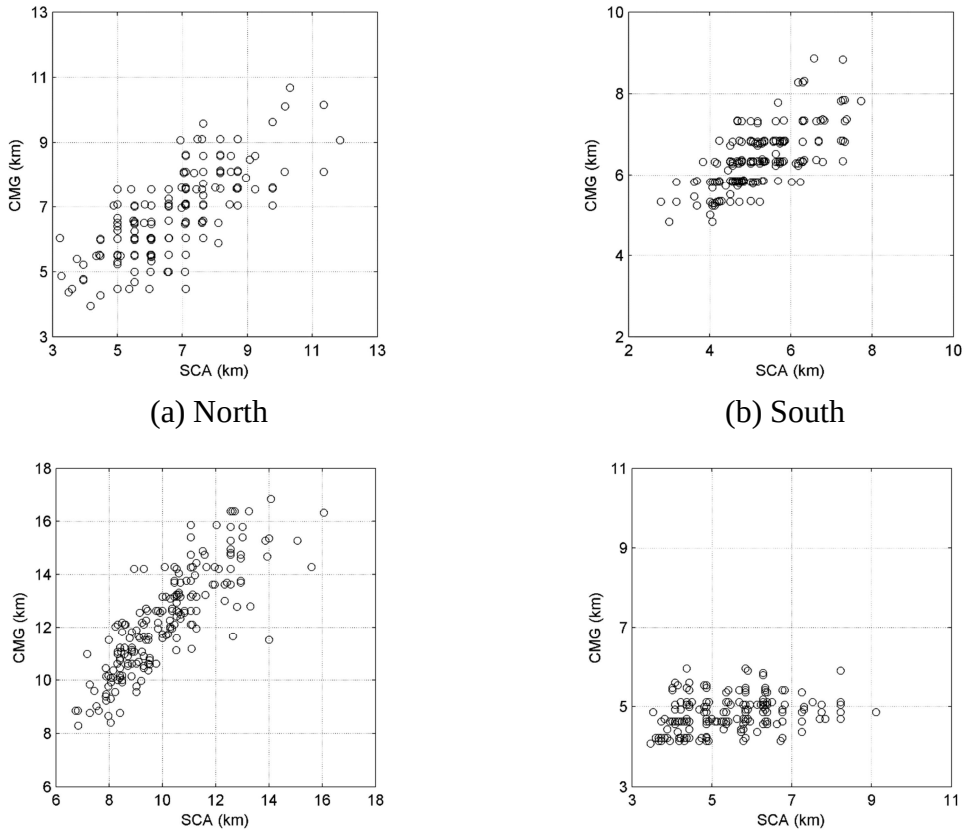
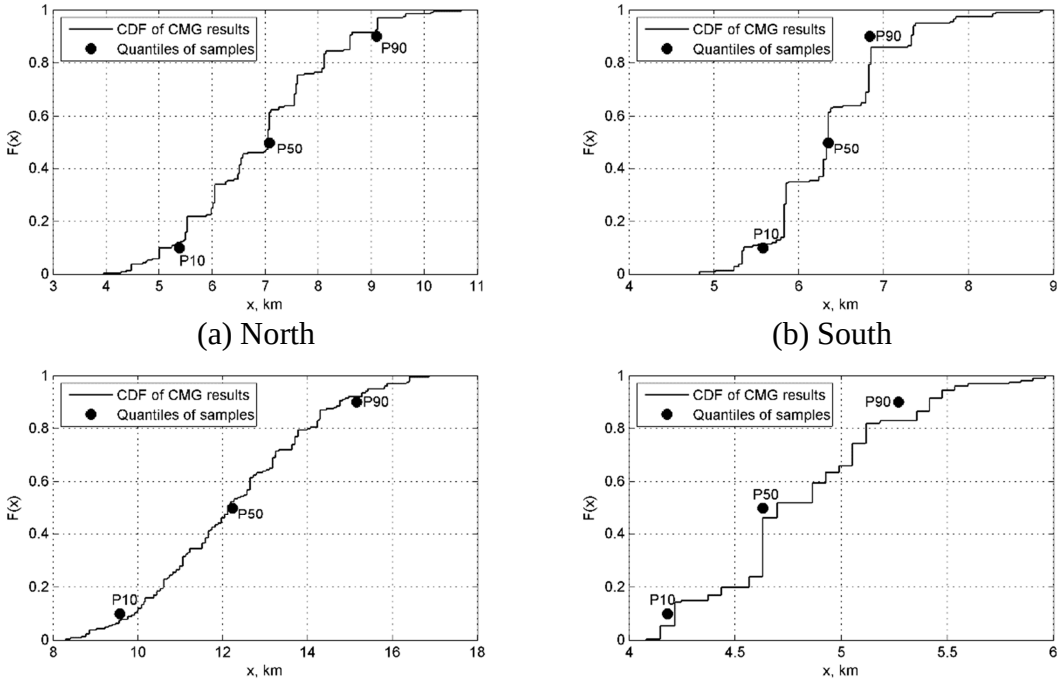


Figure 4.3 Scatter plots of the lengths of the CO_2 plume from the injector in the north, south, east, and west directions computed by CMG-GEM and SCA results

The 200 models are sorted in the order of the lengths of the CO_2 plume computed by SCA. Models are sampled corresponding the equally space quantiles of the sorted distribution. The full physics simulator, CMG-GEM, is then performed on the sampled 20 models The P10, P50, and P90 of the plume travel lengths are calculated based on the 20 full physics simulation results. In Figure 4.4, the solid line is the cumulative distribution function (CDF) of lengths computed based on the CMG-GEM results. The

three points are the P10, P50, and P90 of the lengths based on the SCA results. It is evident from Figure 1.22 that the P10, P50, and P90 quantiles from the SCA results are close to those obtained on the basis of the CMG-GEM results. In Figure 4.4(d), the P10, P50, P90 of the extent travelled by the CO_2 plume in the west direction are close to the true values despite the poor correlation.



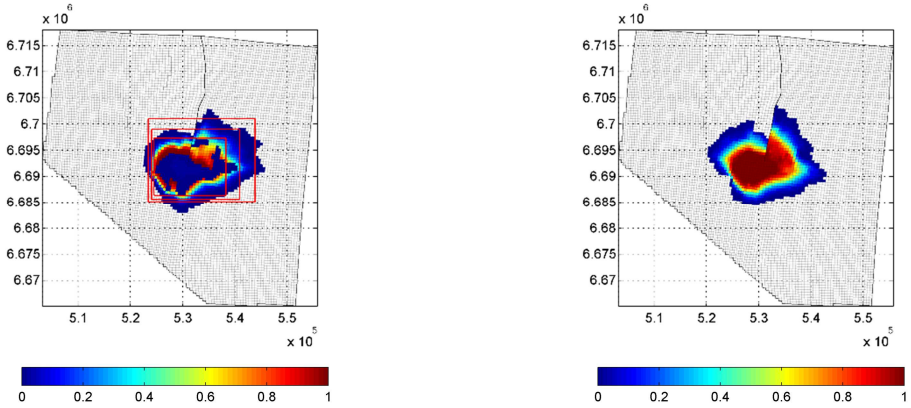
(a) North (b) South

(c) East (d) West

Figure 4.4 Cumulative distribution functions and quantiles of the CO_2 plume lengths in the CMG-GEM results and the samples selected based on the SCA results

It takes about 500 seconds to run a full physics CMG-GEM simulation per model while SCA takes only 0.7 seconds per model. 6 processors were used for CMG-GEM, but single processor was used for SCA. The total computational cost is 200 * 500 / 3600 = 27.78 hours for the ranking process without SCA and (200 * 0.7 + 20 * 500) / 3600 = 2.82 hours for the ranking process with SCA. Thus in this case, about 90% of the computational cost is saved through the ranking process with SCA.

As shown in Figure 4.5, the means profiles of CO_2 saturations based on the CMG-GEM simulations and the SCA approximations have similar shape. In Figure 4.5(a), the red rectangles connect the P10, P50, P90 quantiles in the north, south, east, west directions estimated on the basis of the 20 models sampled from the SCA results. The uncertainty regarding the extent of the CO_2 plume can be visualized in the red rectangles. Figure 1.23 indicates that there is more uncertainty regarding the plume migration in the east direction.



(a) Mean of CMG-GEM CO_2 saturations (b) Mean of SCA approximations

Figure 4.5 Means of CMG-GEM CO_2 saturations and SCA approximations

4.2. VARIOGRAM BASED MODELING CASE

In this case, 400 heterogeneous geological models are generated using the sequential Gaussian simulation and the sequential Gaussian co-simulation, which are variogram based techniques. Table 4.1 shows the details of the three different distributions of porosity and horizontal permeability in the 400 models. In the case names of Table 4.1, short and long mean short and long correlation lengths, and NS and EW mean the north-south and east-west correlation directions. The horizontal-to-vertical permeability is 0.1.

Table 4.1 Three kinds of geological models in the variogram based modeling case

Number of models	200 models	100 models	100 models

Dataset name	Short NS	Long NS	Long EW
Mean and standard deviation of ϕ	(0.2, 0.05)	(0.2, 0.04)	(0.2, 0.04)
Mininum and maximum of ϕ	(0.05, 0.35)	(0.1, 0.3)	(0.1, 0.3)
Mean and standard deviation of $\log_{10}$ k	(2.5, 0.35)	(2, 0.53)	(2, 0.53)
Mininum and maximum of $\log_{10}$ k	(1.6, 3.4)	(0.8, 3.2)	(0.8, 3.2)
Major correlated direction of ϕ and k	North-south	North-south	East-west
Correlation lengths of ϕ and k in terms of cells	(20, 15, 5)	(40, 10, 3)	(40, 10, 3)
Correlation of ϕ and k	0.8	0.7	0.7

Figure 4.6 shows the aquifer structure. The grid system is 151 by 151 by 10 and the cell size is 100 m by 100 m by 5 m. A CO_2 injector is located at the center of the aquifer. The boundary is connected to an infinite aquifer allowing brine displacement. The relative permeability curves shown in Figure 3.9 are used.

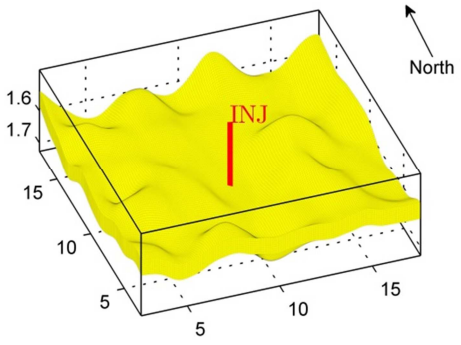


Figure 4.6 Aquifer structure in the variogram based modeling case. The unit of the axes is km. The z-axis is exaggerated by a factor of 20.

Table 4.2 shows the input data of SCA. The fluid properties are obtained from the CMG-GEM results. Different average permeabilities are used for the three datasets when the pressure gradients are calculated using Equation (3-4). 1500 md is used for the short NS and 500 md is used for the long NS and the long EW. This is because as described in

Table 4.1, the long NS and long EW have the same cumulative distribution of permeability, but their distribution is different from that in the short NS.

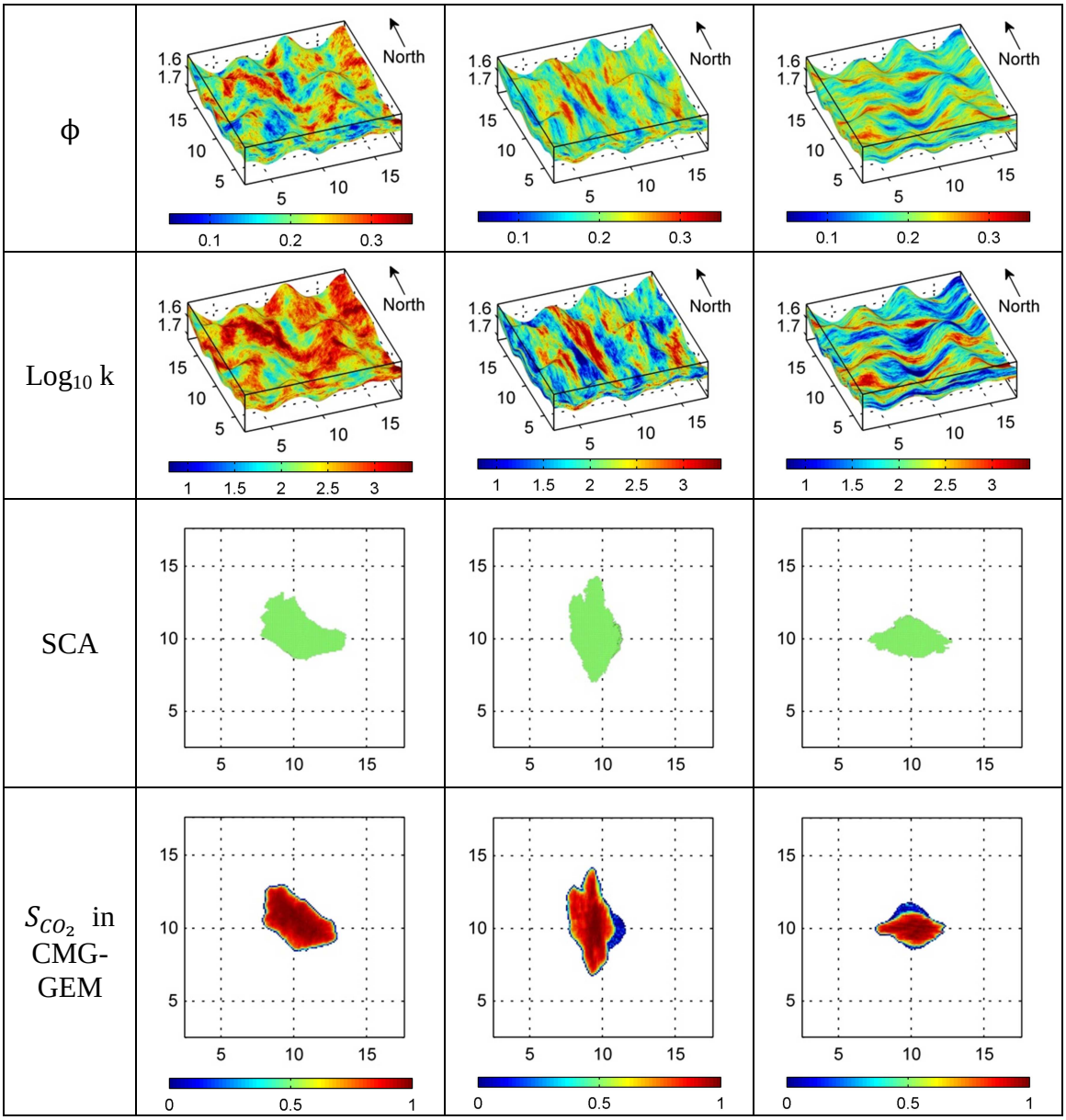
Table 4.2 Input data of SCA for the variogram based modeling case

Aquifer structure	Size	15.1 km * 15.1 km * 50 m
	Dimensions of a grid cell	100 m * 100 m * 5 m
Operating conditions	Injection period	3 years
	Injector location	(i, j, k) = (101, 101, 3) ~ (101, 101, 5)
	Injection rate	20,000 m^3/day in bottom hole conditions
Fluid properties in aquifer conditions	CO_2 saturation at the end point	0.7
	CO_2 relative permeability at the end point	0.8
	Average CO_2 viscosity	0.0445 cp
	Average CO_2 density	545 kg/m^3
	Average brine viscosity	0.3538 cp
	Average brine density	1008 kg/m^3
	Average CO_2 saturation	0.4
Others	Minimum downdip	1°
	Maximum downdip	5°
	Power of downdip factor	0.5

In Table 4.3, the CO_2 plumes in the CMG-GEM results migrate along the 45° counter-clockwise, the north-south, and the east-west permeable features near the injector in the three models. The CO_2 plume in the long EW is smaller than the short NS and the long NS because the permeable zones near the injector in the top layer of the long EW are smaller than the others. SCA reproduces such spatial characteristics of the CO_2 plumes in the CMG-GEM results.

Table 4.3 Porosity, $\log_{10}$ horizontal permeability, SCA result, CMG-GEM result in each dataset of the variogram based modeling case. The unit of the axes is km. The z-axis is exaggerated by a factor of 20.

Dateset name	Short NS	Long NS	Long EW



The extent of the CO_2 plume in the north, south, east, and west directions are measured in order to rank the 400 models. Figure 4.7 shows the extent of the CO_2 plume in the north, south, east, and west directions calculated by CMG-GEM and SCA in the 400 models. The correlations between these results in the 400 models are 0.91, 0.84, 0.88, and 0.87 in the north, south, east, and west directions, respectively. In the CMG-

GEM results, the extent of the CO_2 plume from the injector in each direction is calculated in the zone that has CO_2 saturation greater than 0.01.

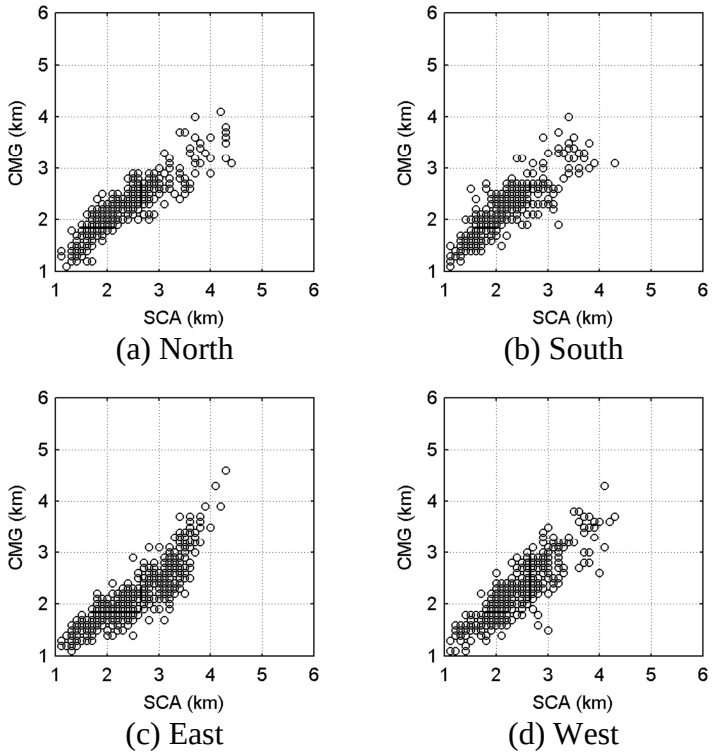
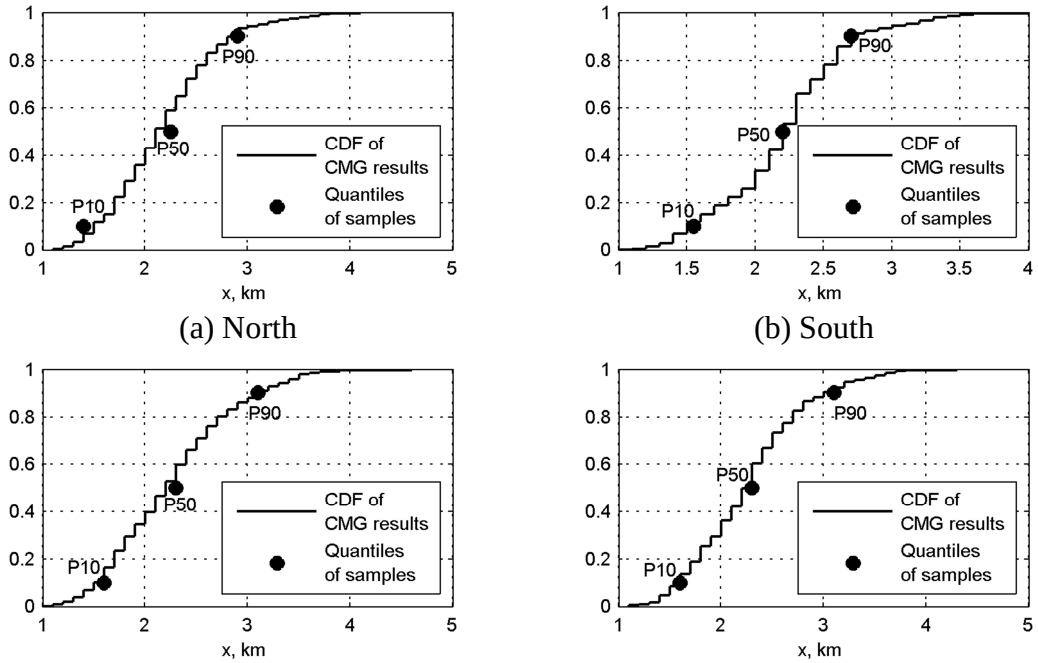


Figure 4.7 Scatter plots of the extent of the CO_2 plume from the injector in the north, south, east, and west directions in the CMG-GEM and SCA results

The 10% of the 400 models are sampled based on the extent of the CO_2 plume computed using SCA in each direction. The P10, P50, and P90 of the extent in each direction are calculated in the full physics simulation results of the sampled 40 models. Figure 4.8 shows the cumulative distribution function (CDF) of the plume extent in the four directions based on the CMG-GEM results and the P10, P50, and P90 of the plume extent computed in the sampled 40 full physics simulation results. As shown in Figure 4.8, the P10, P50, and P90 values in the sampled models are close to the full physics results.



(a) North (b) South

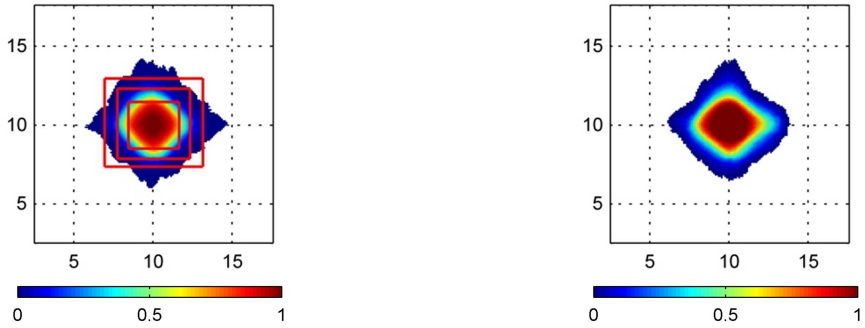
(c) East (d) West

Figure 4.8 Cumulative distribution functions and quantiles of the CO_2 plume lengths in the CMG-GEM results and the samples selected based on the SCA results

It takes about 300 seconds and 2.2 seconds to run CMG-GEM and SCA in a model, respectively. The parallel version of the CMG-GEM simulator was run on 6 processors, but a single processor was used for SCA. The total computational cost is 400 * 300 / 3600 = 33.33 hours for the ranking process using CMG-GEM and (400 * 2.2 + 40 * 300) / 3600 = 3.58 hours for the ranking process using SCA. About 90% of the computational cost is saved in computing P10, P50, and P90 by ranking the 400 models using SCA.

As shown in Figure 4.9, SCA captures the overall spatial characteristics of the mean profile of the CO_2 saturations computed using CMG-GEM. In Figure 4.9 (a), the red rectangles connect the P10, P50, and P90 quantiles of the extent of the CO_2 plumes in the north, south, east, and west directions estimated on the basis of the 40 models sampled from the SCA results. The small, medium, and large rectangles represent P10,

P50, and P90, respectively. The red rectangles cover the extent of the CO_2 plume in the CMG-GEM results.



(a) Mean (>0.01) of CO_2 saturations computed using CMG-GEM in 400 models

(b) Mean (>0.01) of approximations computed using SCA in 400 models

Figure 4.9 Means of CMG-GEM CO_2 saturations and SCA approximations in 400 models. The unit of the axes is km.

4.3. OBJECT BASED MODELING CASE

Issautier et al. (2014) pointed out that variogram-based modeling techniques (or two-point statistics techniques) fail to reproduce CO_2 storage capacity for fluvial reservoirs with high sinuosity because they ignore the connectivity characteristics of fluvial sand channels. In this case, we see if SCA works for fluvial models generated using an object modeling technique (Bridge and Leeder, 1979; Deutsch and Wang, 1996; Keogh et al., 2007).

For this object modeling case, we use the model parameters of the base case in the section 3.2.1. such as the aquifer structure shown in Figure 3.11(a), the operation conditions, the fluid properties, and the relative permeability curves shown in Figure 3.9. Thus the input data for SCA are the same to those in Table 3.4. Only different geological models are used. The model is assumed to have a north-south oriented fluvial depositional environment shown in Figure 4.10. In Figure 4.10, yellow is sand and the background is shale. The horizontal permeability and porosity of sand and shale are 300

md, 0.2 and 1 md, 0.1, respectively. The ratio of horizontal and vertical permeabilities is 0.1. 200 fluvial models are generated using the TiGenerator in SGeMS (Boucher et al., 2010). The thickness is 5 vertical grid cells and the width is 15 lateral grid cells. CO_2 is injected at (i, j, k) = (101, 101, 3) at 18,690.8 kg/day during 1 year. All of the 200 models have sand facies at the injector location. The boundary is connected to an infinite acting aquifer where fluid displacement is allowed.

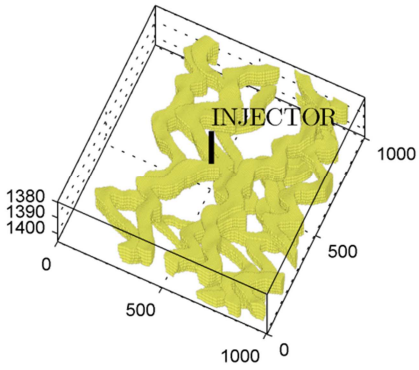
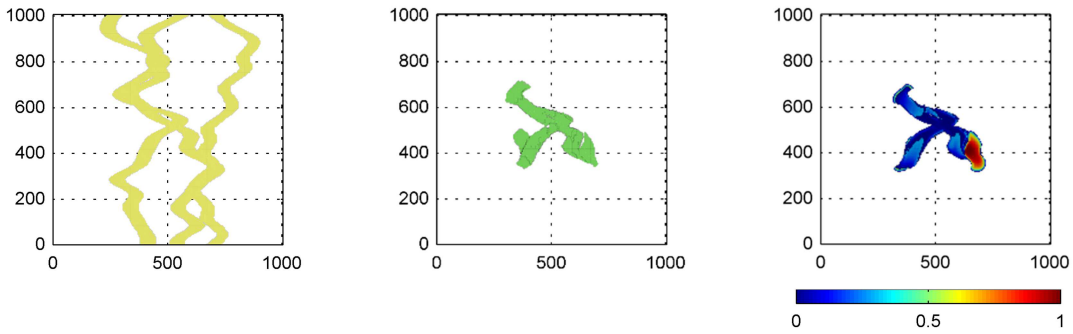


Figure 4.10 A fluvial geological model. The unit of the axes is meters. The z-axis is exaggerated by a factor of 20.

Figure 4.11 shows the results of SCA and CMG-GEM in the geological model shown in Figure 4.10. In this model, there is no sand facies near the injector at k = 9 and 10 where the k index at the top is 10. The 8th layer is the top most layer where the injector is connected to the sand facies. Thus the CO_2 plume vertically migrates up to the 8th layer and then starts to spread horizontally in the 8th layer. As shown Figure 4.11(b) and (c), SCA reproduce the spatial features of the CO_2 plume that migrates along the permeable zone in the 8th layer.

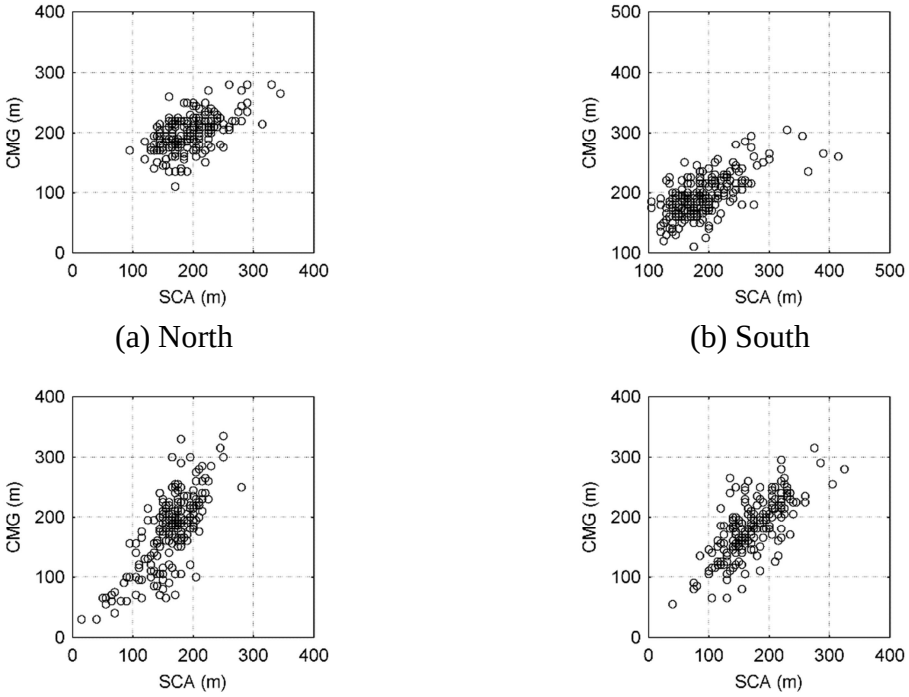


(a) Sand facies at k = 8 (b) SCA result (c) CO_2 saturation computed by CMG-GEM

Figure 4.11 Sand facies in the k = 8 slice through the 3D reservoir model, SCA result, and CO_2 saturation computed by CMG-GEM in that slice obtained after performing the simulation on the 3D model shown in Figure 4.10. The unit of the axes is meters. The z-axis is exaggerated by a factor of 20.

The extent of the CO_2 plume from the injector in the north, south, east, and west directions are calculated using the results obtained by SCA and CMG-GEM. In the CMG-GEM results, the extent of the CO_2 plume is the maximum distance from the injector in each direction in the zone that has CO_2 saturation greater than 0.01. Figure 4.12 shows the extent of the CO_2 plumes in the north, south, east, and west directions in the CMG-GEM and SCA results. In Figure 4.12, the x-axis and the y-axis represent the results in SCA and CMG-GEM, respectively. The correlations computed using the 200 fluvial models are 0.52, 0.65, 0.74, and 0.73, respectively. The correlations in the north and south directions are smaller than those in the east and west directions. This is because the variations in the extent of the plume in the north and south directions are smaller than those in the other two directions as shown Figure 4.12. The main orientation of the channels is the north-south direction, and most of the models have the CO_2 plume extending in those directions. In contrast, the plume migration in the east and west directions depends on the sinuosity of the channel bodies connected to the injector. For this reason, the permeable zones have larger variations in the east and west directions

than in the north and south directions. These variations are greater than the resolution of the SCA model and consequently the correlations in the east and west directions are better.



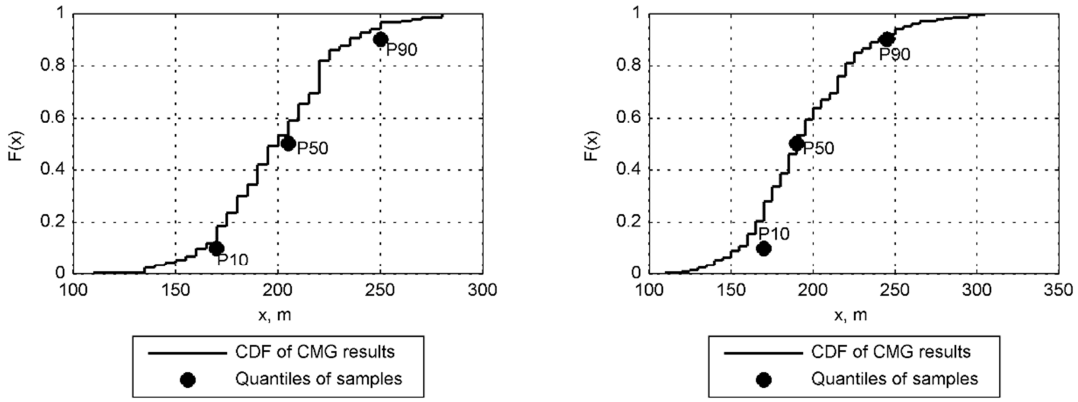
(a) North (b) South

(c) East (d) West

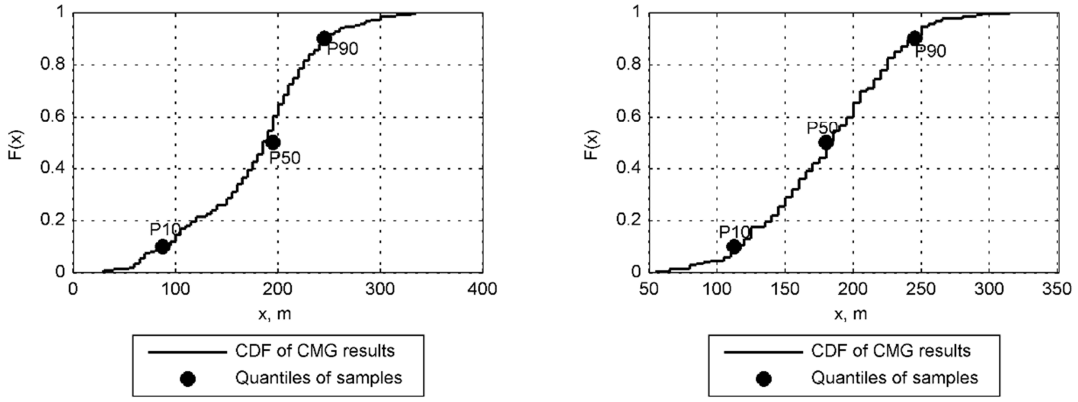
Figure 4.12 Scatter plots of the extent of the CO_2 plume from the injector in the north, south, east, and west directions in the CMG-GEM and SCA results

The 200 models are ranked in the order of the extent of the CO_2 plume computed using SCA and then 10% of the 200 models are sampled (corresponding to equally spaced quantiles of the cdf of the plume front distance from the injector). The P10, P50, and P90 of the plume extent are computed by running the full physics simulator, CMG-GEM, on the 20 sampled models. Figure 4.13 shows the cumulative distribution function (CDF) of the plume extent in different directions based on the CMG-GEM results and the P10, P50, and P90 of the plume extent computed in the sampled 20 full physics

simulation results. As shown in Figure 4.13, the P10, P50, and P90 values for the plume extent in the sampled models are close to the full physics results.



(a) North (b) South

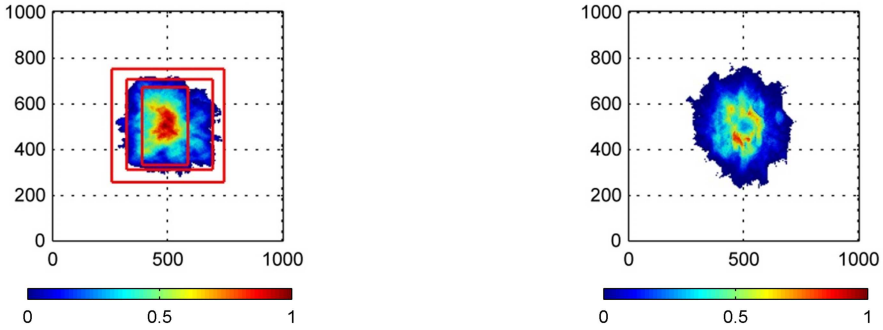


(c) East (d) West

Figure 4.13 Cumulative distribution functions and quantiles of the CO_2 plume lengths based on the CMG-GEM results and the samples selected based on the SCA results

It takes about 1100 seconds and 4 seconds to run CMG-GEM and SCA using a geological model, respectively. The parallel version of the CMG-GEM simulator was run on 6 processors, but a single processor was used for SCA. The total computational cost for the CMG-GEM simulation is 200 * 1100 / 3600 = 61.11 hours and (200 * 4 + 20 * 1100) / 3600 = 6.33 hours for SCA. About 90% of the computational cost is saved in estimating P10, P50, and P90 using SCA.

In Figure 4.14(a) and (b), the migration of the CO_2 plume in the north and south directions are slightly overestimated in the mean profile computed using the SCA approximations compared to that by the CMG-GEM. In Figure 4.14(a), the red rectangles connect the P10, P50, and P90 estimates for the extent of the plume in the north, south, east, and west directions estimated using the sampled 20 models. The red rectangles cover most of the CO_2 saturated zone in Figure 4.14(a).



(c) Mean (>0.01) of CO_2 saturations computed using CMG-GEM in 200 models

(d) Mean (>0.01) of approximations computed using SCA in 200 models

Figure 4.14 Means of CMG-GEM CO_2 saturations and SCA approximations in 200 models. The unit of the axes is meters. The z-axis is exaggerated by a factor of 20.

4.4. CONCLUSIONS

SCA was used to quickly quantify the uncertainty in the extent of CO_2 plumes in models that are described using two-point statistics and fluvial models with strong sinuosity. In the cases, SCA provided the accurate proxy responses that reflect the plume migration in the different directions. The initial suite of geological models was ranked and sampled based on the proxy responses. The P10, P50, and P90 of the full physics simulation in the sampled models were calculated. About 10% of the total computational cost was required to quantify the P10, P50, and P90 responses bracketing the uncertainty in the extent of the CO_2 plumes in each direction.

Chapter 5. A Shape Dissimilarity Measure between 2D Fluid Displacements Using Perimeters

5.1. HAUSDORFF DISTANCE

The Euclidean distance between two binary images is given by:

$$Euclidean(A,B) = \|A - B\| \quad \text{Equation (5-1)}$$

where A and B represent the vectors describing two spatial objects such as a CO_2 plume and $\|\cdot\|$ is a norm operator. Each element of the vector corresponds to a binary value at a location e.g. either saturated with CO_2 or not computed by applying a threshold. The norm operator is applied over all elements of the vectors regardless of their spatial location and so in that sense, the Euclidean distance measures the dissimilarity between images regardless of the spatial relation between the elements of the vectors representing the images.

Contrary to the Euclidean distance, the Hausdorff distance considers the spatial relation between two point sets. The Hausdorff distance is given by (Huttenlocher et al., 1993):

$$H(A,B) = max\big(h(A,B), h(B,A)\big) \quad \text{Equation (5-2)}$$

where

$$h(A,B) = \max_{a\in A} \min_{b\in B} \|a - b\| \quad \text{Equation (5-3)}$$

and a and b represent elements of A and B, respectively. Figure 5.1 shows the four distances from a_3 to the four elements of B. The distance from a_3 to b_1 is the smallest and the other minimum distances from a to b are also obtained. The maximum distance among the minimum distances from a to b is h(A, B) in Equation (5-3). Likewise, h(B, A) is computed and then the maximum of h(A, B) and h(B, A) is the Hausdorff distance between and A and B.

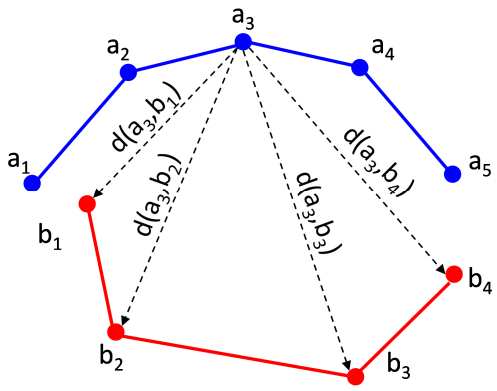


Figure 5.1 An example of calculating the Hausdorff distance

The computational cost of the Hausdorff distance between A and B is proportional to the product of the numbers of elements making up A and B (Huttenlocher et al., 1993). The computational cost of the Hausdorff distance can be reduced by decreasing the numbers of points describing the spatial objects using representations such as perimeter, surface, and skeleton.

5.2. PERIMETER

A CO_2 plume obtained 4D seismic surveys may be given as a 2D object. The common representation of a 2D object is a perimeter. In a 2D binary image, a perimeter pixel has at least one neighbor pixel with a zero attribute value where a pixel has four neighbor pixels (north, south, east, and west). 3D simulated CO_2 plumes are vertically averaged to 2D objects in order to compare them to a given 2D observed CO_2 plume. The vertically averaged 2D objects are converted to binary images by thresholding them. The converted 2D images do not represent the vertical characteristics of the CO_2 plume. Thus the Hausdorff distance between perimeters can be meaningfully computed only for CO_2 plumes that do not exhibit strong vertical characteristics compared to horizontal migrations.

The procedure for measuring the dissimilarity between the perimeters of 2D binary images of CO_2 plumes using the Hausdorff distance is as follows:

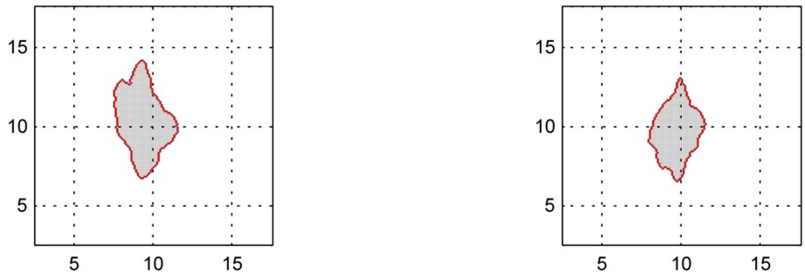
i) Fill holes in the 2D binary images

ii) Find perimeter pixels

iii) Calculate the Hausdorff distance between two perimeters

In the first step, holes in an image are the zero pixels which are not connected to the perimeter (Soille, 2004). The holes are filled because the perimeter of a CO_2 plume is assumed to be a closed curve.

Figure 5.2(a) and Figure 5.2(b) show the 2D binary images and the perimeters of two CO_2 plumes where the perimeter pixels are colored red. Each perimeter has about 10% of the total number of the non-zero value pixels in each 2D CO_2 plume image. Because the time for computing the Hausdorff distance is proportional to the product of the numbers of two point sets, using only the perimeter pixels considerably reduces the cost.



(a) Areal view of 1st 2D CO_2 plume and its perimeter

(b) Areal view of 2nd 2D CO_2 plume and its perime

Figure 5.2 2D binary images and perimeters of two CO_2 plumes. The meshes and red-colored grid blocks represent 2D CO_2 plumes and perimeter pixels. The unit in x- and y-axis are in kilometers.

5.3. MODEL SELECTION ALGORITHM

The final goal of this study is to select the most probable models honoring an observed CO_2 plume over a large suite of geologic models. The observed CO_2 plume has to be compared to the CO_2 plumes that are estimated over the large suite of prior geologic models. The geologic models that exhibit CO_2 plume migrations similar to the observed data are chosen as the probable models. A model selection algorithm (Bhowmik et al., 2010) is used to select the most probable models honoring an observed CO_2 plume. The procedure is as follows:

1) Run SCA in prior models
2) Convert the 3D SCA approximations to 2D projections by vertically averaging them
3) Convert the 2D projections to 2D binary images by thresholding them
4) Measure the dissimilarity between all pairs of the 2D SCA approximations
5) Project the models onto a N-dimensional space using multi-dimensional scaling (MDS)
6) Group the models in the N-dimensional space using the k-mean clustering algorithm
7) Conduct fully physics flow simulations in the representative models of the groups
8) Convert the full physics simulated CO_2 saturation images of the representative models to 2D projections by vertically averaging them
9) Convert the 2D projections to 2D binary images by thresholding them
10) Measure the dissimilarity between the observed 2D CO_2 plume and the 2D binary images of the representative models

11) Select the representative model that has the most similar 2D CO_2 plume to the observed 2D CO_2 plume

12) Select the group as the most probable models that contains the selected representative model

In the 4th and 10th steps, the dissimilarity between 2D objects is measured using the following methods:

i) The Euclidean distance

ii) The Hausdorff distance

iii) The Hausdorff distance between the perimeters

In the 5th step, multivariate analysis of the dissimilarity matrix obtained in the 4th step is performed using MDS (Kruskal, 1964) in order to project the models in a metric space. MDS finds an optimal projection for a given dimension to minimize the "stress" function (Kruskal, 1964).

$$MDS\ Stress\ function = \sqrt{\frac{\sum_i \sum_j \left(d_{ij}^* - d_{ij}\right)^2}{\sum_i \sum_j {d_{ij}}^2}} \quad \cdots\cdots\cdots\cdots\cdots\cdots \text{Equation (5-4)}$$

The subscripts i and j are an index of reservoir models, d is the distance between two models, and the superscript * implies the distance in the space projected by MDS. The acceptable level of the stress function is not mathematically proven, but the rule of thumb is that it should be under 0.1. As the dimension of the projected space increases, the stress function decreases. N is set to the minimum dimension in which the stress function is less than 0.1.

In the 6th step, the k-means clustering algorithm (Bishop, 2007) selects M locations randomly in a space such that the sum of distances between the models within each cluster is minimized and the distance to models in other clusters or groups are

maximized. The optimal M locations are called centroids of the clusters, and the closest models to the centroids are called medoids, which are considered as representative models. The representative models are be processed through a full physics flow simulator.

For the k-means clustering algorithm, the elbow method (Thorndike, 1953) is used to determine the optimal number of groups. The elbow method is to choose the number of clusters at which the sum of the squared distance (SSD) between the cluster members and their centroid within each cluster decreases suddenly and reaches to a plateau. SSD is given by:

$$SSD = \sum_{i=1}^{K} \sum_{j=1, x_{i,j} \in c_i}^{N_i} \left\| x_{i,j} - c_i \right\|^2 \quad \text{Equation (5.5)}$$

The subscripts i and j are the index of a cluster and the index of a cluster member in the cluster. K and N_i are the number of clusters and the number of cluster members of the i-th cluster. c and x represent the centroid of a cluster and a cluster member of the cluster. $\|\cdot\|$ is a norm operator.In the 7th step, the full physics simulations in the representative models are conducted using CMG-GEM (CMG, 2012).

5.4. APPLICATION

The model selection scheme is applied to select the most probable models honoring CO_2 plumes observed in an aquifer shown in Figure 5.3. This example will serve to emphasize the efficacy of the Hausdorff distance to capture the differences between models. The grid system consists of 151 × 151 × 10 grids with $\Delta x = \Delta y = 100$ and $\Delta z = 5$m. A CO_2 injector is located in the middle of the aquifer. The initial average pressure and temperature of the aquifer is 18,145 kPa and 75 °C, respectively. The

boundary is connected to an infinite aquifer allowing outflow of CO_2. The relative permeability curves are shown in Figure 3.9 and capillary pressure is ignored.

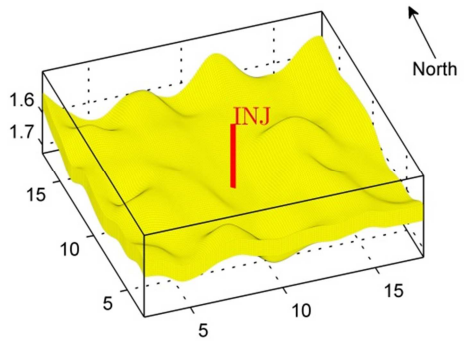


Figure 5.3 Structure of the aquifer and the location of the CO_2 injector in the application example. The unit of the horizontal axes are in kilometers. The z-axis is exaggerated by factor of 10.

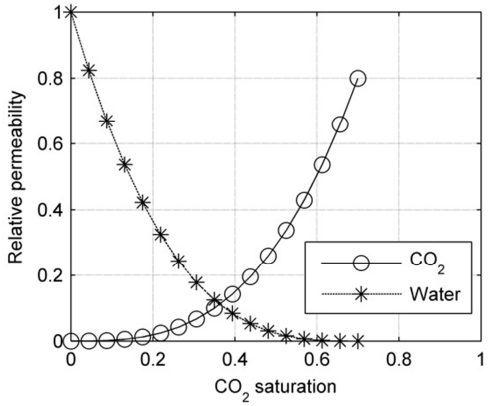


Figure 5.4 Relative permeability curves of CO_2 and water

Table 4.1 shows the geostatistical parameters for three datasets of 400 geologic models. Porosity of the 400 geologic models is generated using the sequential Gaussian simulation of SGeMS (Remy et al., 2011). Permeability is assumed to be correlated with porosity, and it is generated using the sequential Gaussian co-simulation of SGeMS (Remy et al., 2011). The dataset names in Table 4.1 refer to short and long correlation lengths, and north-south (NS) and east-west (EW) correlation directions. The ratio of vertical to horizontal permeability is 0.1.

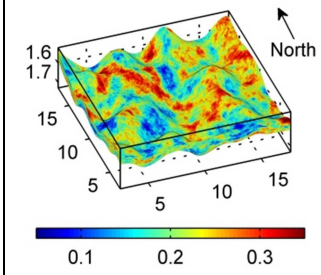
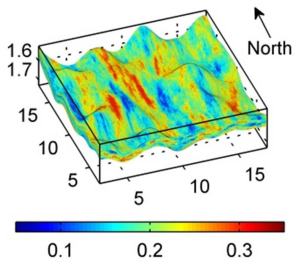
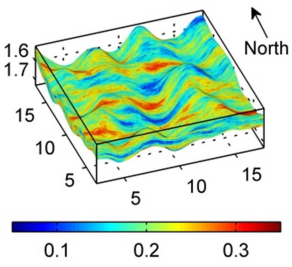
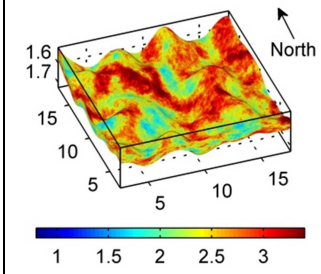
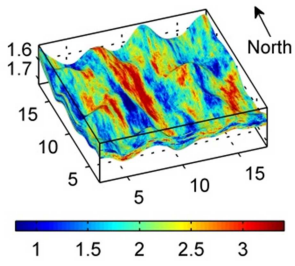
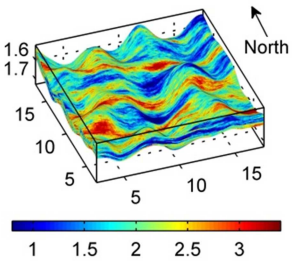
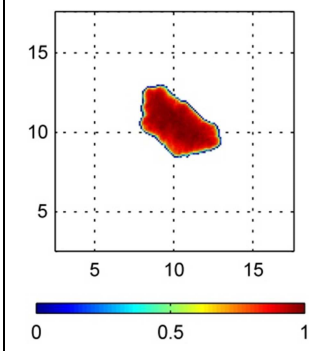
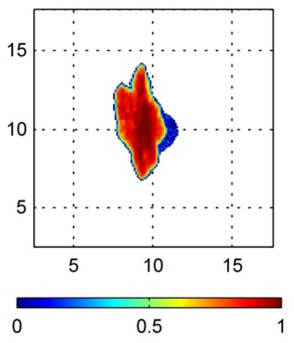
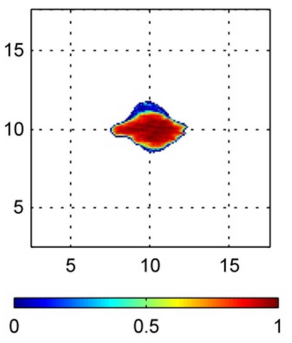
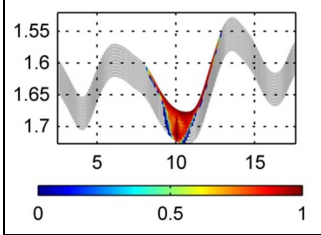
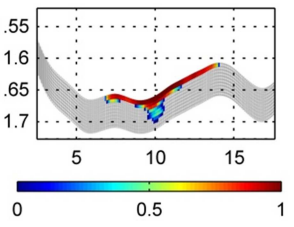
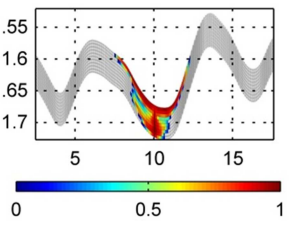
Table 5.1 Geostatistical parameters used for generating the three sets of geologic models

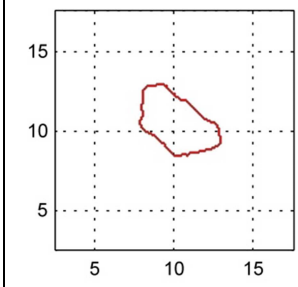
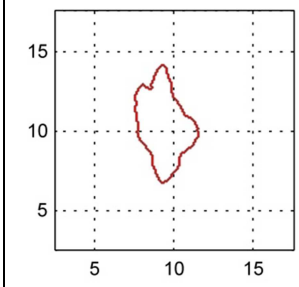
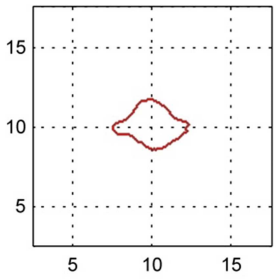
Number of models	200 models	100 models	100 models
Dataset name	short NS	long NS	long EW
Mean and standard deviation of ϕ	(0.2, 0.05)	(0.2, 0.04)	(0.2, 0.04)
Mininum and maximum of ϕ	(0.05, 0.35)	(0.1, 0.3)	(0.1, 0.3)
Mean and standard deviation of $\log_{10}$ k	(2.5, 0.35)	(2, 0.53)	(2, 0.53)
Mininum and maximum of $\log_{10}$ k	(1.6, 3.4)	(0.8, 3.2)	(0.8, 3.2)
Major correlated direction of ϕ and k	north-south	north-south	east-west
Major, minor, and vertical correlation lengths of ϕ and k in terms of cells	(20, 15, 5)	(40, 10, 3)	(40, 10, 3)
Correlation of ϕ and k	0.8	0.7	0.7

Table 4.3 shows porosity and horizontal permeability of some of the models, and the top and cross-sectional views of the corresponding CO_2 plumes simulated using CMG-GEM for one model in each of the three sets. The perimeters of the simulated CO_2 plumes for the same models are also shown.

As evident by the simulated CO_2 saturations in Table 4.3, the injected CO_2 migrates along permeable features. In the cross-sectional view of CO_2 saturation in Table 4.3, CO_2 migrates first to the top of the structure and then spreads horizontally. The wedge shape of the CO_2 plumes is common in CO_2 plume migrations because of strong buoyancy. This implies that the simulated vertical characteristics of the CO_2 plumes are not likely to be too different for the 400 models. The last row of Table 4.3 shows the perimeters of the projected 2D CO_2 plumes.

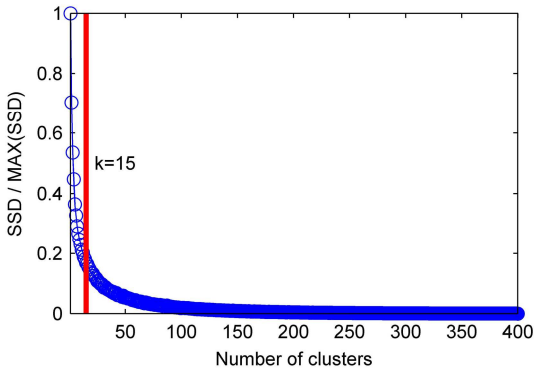
Table 5.2 Porosity, $\log_{10} k_h$, CO_2 saturations computed using CMG-GEM for the three datasets. The units of the horizontal axes are in kilometers. The z-axes in the 3D views and the cross-sectional views are exaggerated by factors of 10 and 40, respectively

Dateset name	Short NS	Long NS	Long EW
ϕ			
$Log_{10}\ k_h$			
Areal view of simulated CO_2 saturation.			
Cross-sectional view of CO_2 saturation computed using CMG-GEM			
	J = 101	I = 95	J = 101

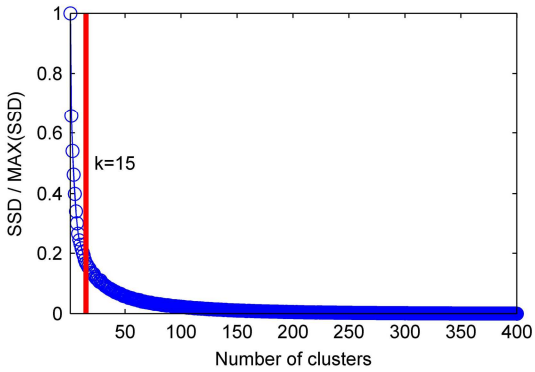
Perimeter of 2D CO_2 plume			

As described in the procedure of the model selection process in the previous section, the 400 models are grouped based on the shape dissimilarity between the 400 SCA approximations measured using three measures: the Euclidean distance, the Hausdorff distance, the Hausdorff distance between perimeters. The representative models of the groups will then be processed through a full physics flow simulator. The full physics simulated CO_2 plumes of the representative models are compared to the 2D observed CO_2 plume and a conclusion regarding the efficacy of the proposed Hausdorff distance using perimeters is made. The group containing the best-matched representative model is selected as the most probable models honoring the observed CO_2 plume characteristics.

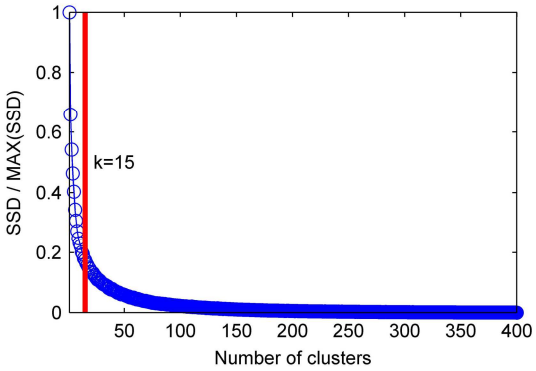
Figure 5.5 shows the optimal numbers of clusters determined using the elbow method. The optimal numbers of clusters are 15 for all the three approaches. The 400 geologic models are classified into 15 groups using the k-means clustering algorithm, and then the CO_2 plumes simulated in the representative models of the 15 groups are compared to an observed CO_2 plume.



(a) Euclidean distance



(b) Hausdorff distance



(c) Hausdorff distance between perimeters

Figure 5.5 Relative SSD (=SSD over the maximum value of SSD) of the clusters grouped based on the distance matrix measured using the three approaches and the optimal number of clusters

5.4.1. 2D Case #1

Figure 5.6 shows an observed CO_2 plume for the 2D case #1 and the observed CO_2 plume is given as a 2D image. The extent of the observed CO_2 plume in the east-west direction is longer than that in the north-south direction. The shape of the observed CO_2 plume can be described as an ellipse in the center with a long and narrow protrusion toward the east.

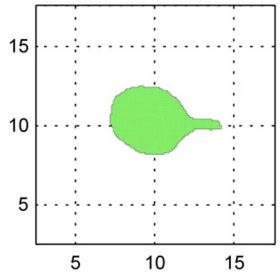


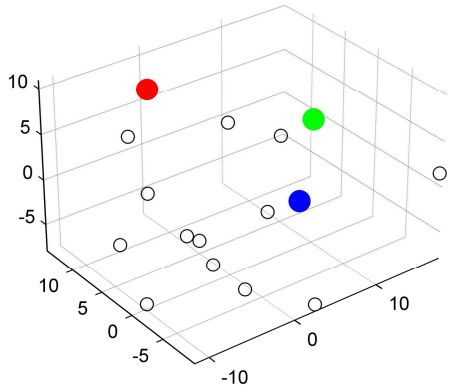
Figure 5.6 Observed CO_2 plume for the 2D case #1. The units of the axes are in kilometers.

The SCA approximations for the 400 prior models are computed, and then the dissimilarity between the 400 SCA approximations is measured using the Euclidean distance, the Hausdorff distance, and the Hausdorff distance between perimeters. The 400 models are projected onto a 10-dimensional metric space using MDS based on the distance matrix. The 10-dimensional metric space was used to make the MDS stress function shown in Equation (5-4) less than 0.1.

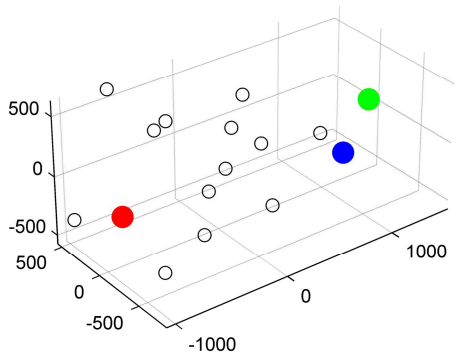
The 400 models are clustered into 15 groups based on the distances between the SCA approximations measured using the three approaches. The 15 representative models (medoids) are full physics simulated using CMG-GEM, and then the dissimilarity between the 15 full physics simulated CO_2 plumes and the observed CO_2 plume shown in Figure 5.6.

Figure 5.7 shows the relationship between the observation and the 15 medoids in 3D metric spaces. The 3D metric spaces are derived using MDS based on the distance matrix computed using each measure. The dimension (=10) of the space used to project the 400 models is too high to visualize the locations of the 400 models in the space. For this reason, the 3D metric space of the observation and the 15 medoids shown in Figure 5.7 is plotted only for visualization. In the 3D space, the distance represents the shape dissimilarity between the 2D CO_2 plumes corresponding to the observation and the 15

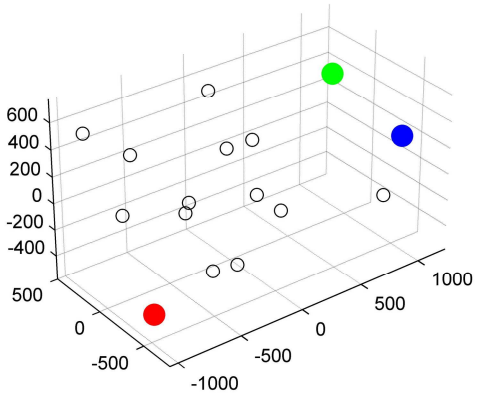
medoids. The green, blue, and red dots correspond to the observation, the selected medoid, and the farthest medoid from the observation. The group members of the medoid corresponding to the blue dot are selected as the most probable models. The red dot corresponds to the medoid of the cluster of models that exhibit most dissimilarity to the observed CO_2 plume.



(a) The Euclidean distance



(b) The Hausdorff distance

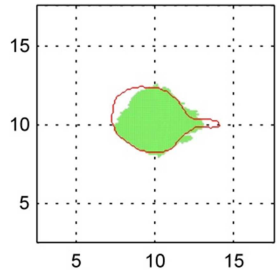
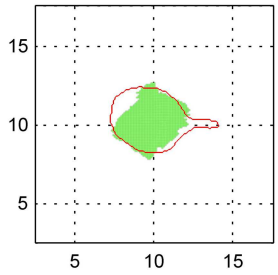
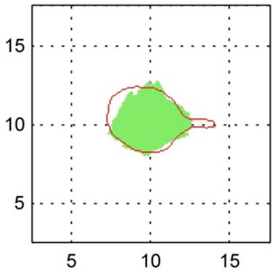
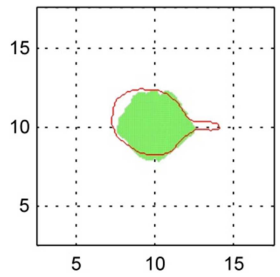
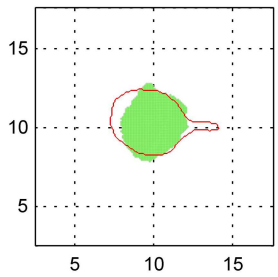
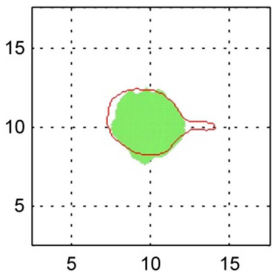


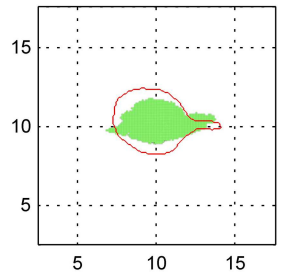
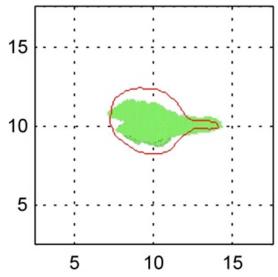
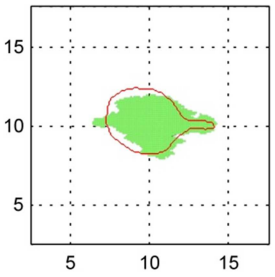
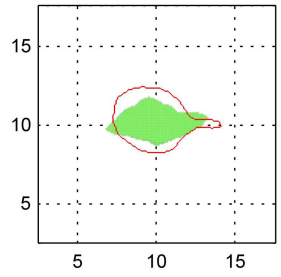
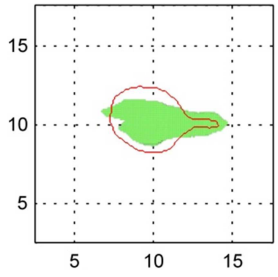
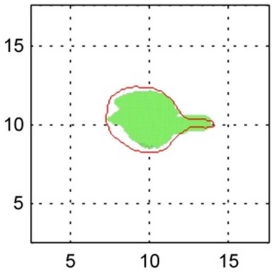
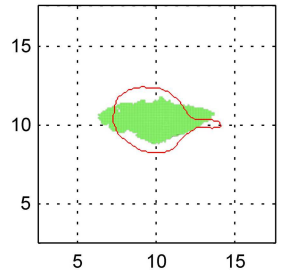
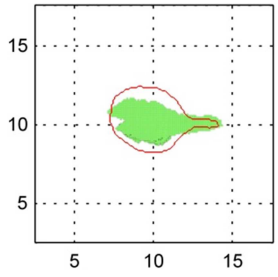
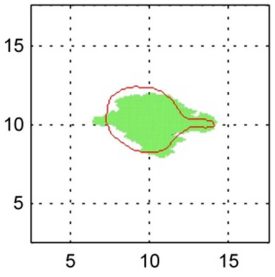
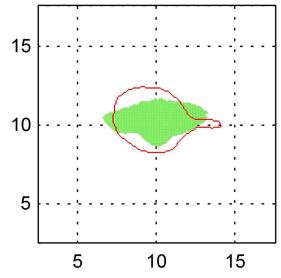
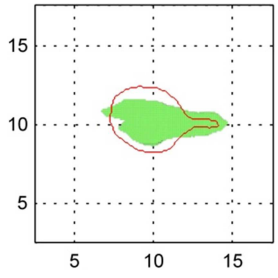
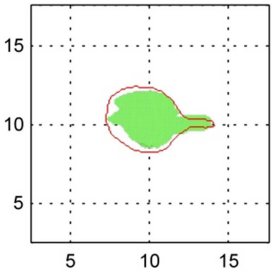
(c) The Hausdorff distance between perimeters

Figure 5.7 3D metric space obtained by MDS depicting the positions of the observation and the medoids obtained by cluster analysis. The green, blue, and red dots represent the observation, the selected medoid, and the medoid farthest from the observation.

Table 5.3 shows the CO_2 plumes of the models selected using the model selection process and using the three distance measures for the observed CO_2 plume shown in Figure 5.6. Each medoid selected using each approach corresponds to the blue dot in the 3D space shown in Figure 5.7.

Table 5.3 CO_2 plumes computed using SCA and CMG-GEM for the medoid and two group members of the group selected using the model selection process. The results are shown corresponding to the three distance measures. The units of the axes are in kilometers. The red curve is the perimeter of the observed CO_2 plume shown Figure 5.6.

	Selected Medoid	Group member #1	Group member #2
Euclidean distance	Proxy CO_2 plumes computed using SCA		
			
	Corresponding CO_2 plumes computed using CMG-GEM		
			
Hausdorff	Proxy CO_2 plumes computed using SCA		

distance			
	Corresponding CO_2 plumes computed using CMG-GEM		
			
Hausdorff distance between perimeters	Proxy CO_2 plumes computed using SCA		
			
	Corresponding CO_2 plumes computed using CMG-GEM		
			

In Table 5.3, the CO_2 plumes computed using SCA and using CMG-GEM are similar enough justifying the conclusion that the SCA approximations is sufficient for being used within the model selection process for discriminating the models.

The CO_2 plumes corresponding to the models obtained using the Euclidean distance do not exhibit the plume extension toward the east. This is because the Euclidean distance measures the overall dissimilarity between models regardless of the spatial distribution of the CO_2 plumes.

On the contrary, the CO_2 plumes corresponding to the models obtained using the Hausdorff distance and the Hausdorff distance between perimeters have an ellipse in the center and a long and narrow protrusion toward the east. In Table 5.3, the second model in the selected set obtained using the Hausdorff distance and the Hausdorff distance between perimeters exhibit CO_2 plume characteristics that is more similar to the observed CO_2 plume than the others. These results are because the Hausdorff distance and the Hausdorff distance between perimeters consider the details of the spatial distribution CO_2 plumes when computing the dissimilarity between models.

Table 5.4 shows the CO_2 plume of the farthest medoid from the observation and the CO_2 plumes of the group members of that cluster. The medoid based on each distance measure corresponds to the green dot in the 3D space shown in Figure 5.7. The major orientation of the CO2 plume shown in

Table 5.4 is the north-south direction while the major orientation of the observed CO2 plume is the east-west direction. In this case, the three approaches do not have significant difference in finding the most dissimilar CO2 plumes to the observation.

Table 5.4 CO_2 plumes computed using CMG-GEM for the medoid and the two members of the group farthest from the observation. The units of the axes are in kilometers. The red curve is the perimeter of the observed CO_2 plume shown Figure 5.6.

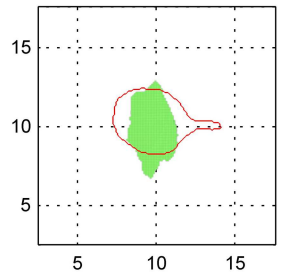
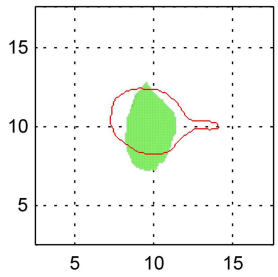
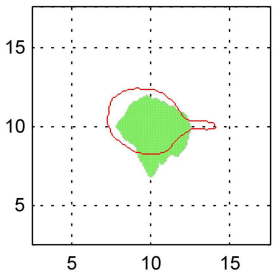
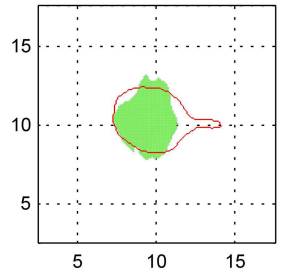
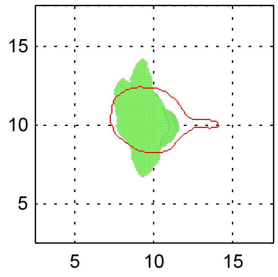
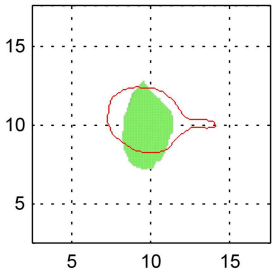
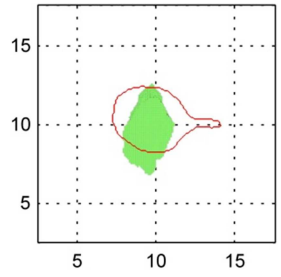
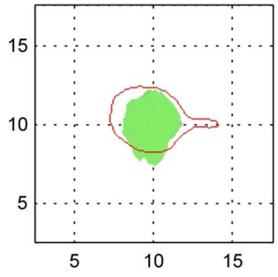
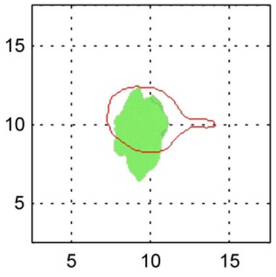
	Farthest medoid	Group member #1	Group member #2
Euclidean distance			
Hausdorff distance			
Hausdorff distance between perimeters			

Figure 6.10 shows the computational costs associated with the three distance measures. It took less than 2 seconds to calculate the perimeters of the 400 proxy CO_2 plumes computed using SCA. The cost of calculate the perimeters is not considered because it is negligible compared to the cost of measuring the dissimilarity. The Hausdorff distance between CO_2 plume perimeters saved about 98% of the computational cost associated with the computation of the Hausdorff distance even though they both yielded the similar clustering result.

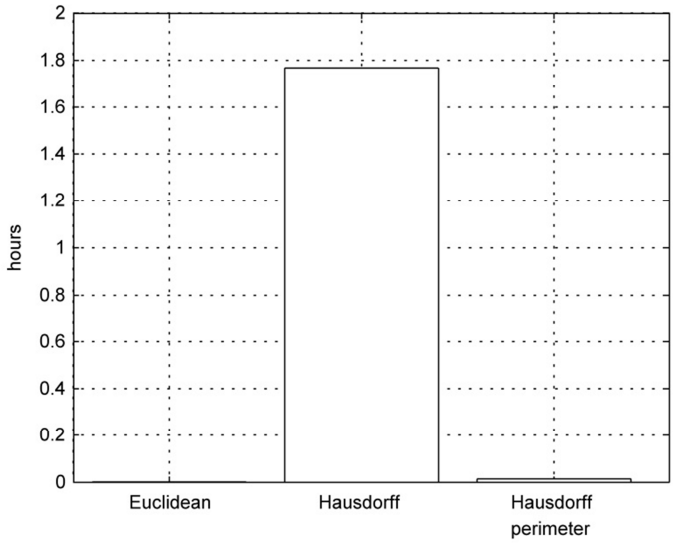


Figure 5.8 Computational costs of the three approaches for the 2D case #1

It took about 300 seconds and 2.2 seconds to run CMG-GEM and SCA for a geological model as mentioned in the section 3.2. The total computational cost of the model selection process using the Hausdorff distance is (400 * 2.2 + 15 * 300 + 6364) / 3600 = 3.26 hours. The total computational costs of the model selection process using the Hausdorff distance between perimeters is (400 * 2.2 + 15 * 300 + 55) / 3600 = 1.51 hours. The Hausdorff distance between perimeters saved about 54% of the total computational cost of the model selection process using the Hausdorff distance. It is to be emphasized that at the end of this computation, the residual uncertainty associated with the reservoir models after the history matching process can be quantified as opposed to a single deterministic history matched model obtained using the traditional history matching approaches. More savings in computational cost (as compared to traditional workflows) would be evident if more temporal CO_2 plume observations are matched or more grid blocks are used to resolve the CO_2 plume.

5.4.2. 2D Case #2

Figure 5.9 shows an observed CO_2 plume for the 2D case #2. The extent of the observed CO_2 plume in the east-west direction is longer than that in the north-south direction. The shape of the observed CO_2 plume is described as an ellipse in the center, however the extent of the observed CO_2 plume is assumed to be uncertain. This mimics the real case where a cutoff value on a seismic survey may determine the extent of the plume and there may be some ambiguity in the specification of that threshold. The objective is to demonstrate that proposed measure for dissimilarity of CO_2 plumes discriminates between models regardless of the uncertainty associated with the extent of the CO_2 plume. The 400 prior models do not have CO_2 plumes that extend as far as the observed CO_2 plume shown in Figure 5.9.

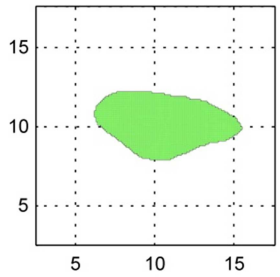
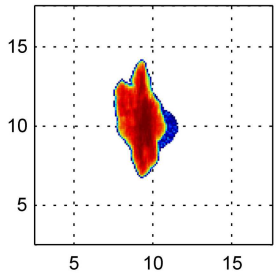


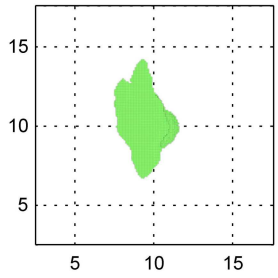
Figure 5.9 Observed CO_2 plume for the 2D case #2. The units of the axes are in kilometers.

The extents of an observed and simulated CO_2 plumes can be different according to the threshold employed on a 4D seismic survey or a flow simulation. As shown in Figure 5.10, the extent of the CO_2 plumes expressed in terms of binary values are different according to the threshold values. A dissimilarity measure of CO_2 plumes should find simulated CO_2 plumes similar to an observed one regardless of the uncertainty associated with thresholds. However, because the Euclidean distance does not spatially consider the difference between CO_2 plumes as shown in Figure 2.1, the

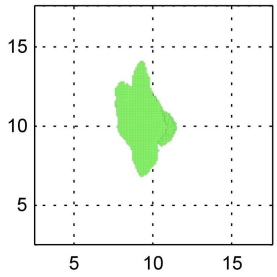
Euclidean distance might measure the dissimilarity between CO_2 plumes inaccurately according to thresholds.



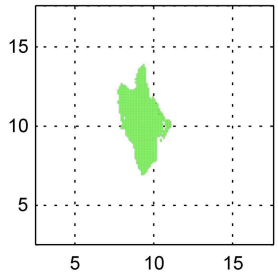
(a) CO_2 saturation computed using CMG-GEM



(b) Threshold = 0.01



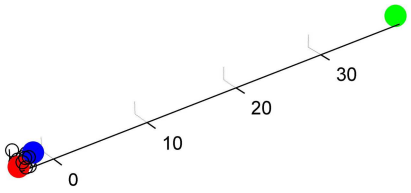
(c) Threshold = 0.25



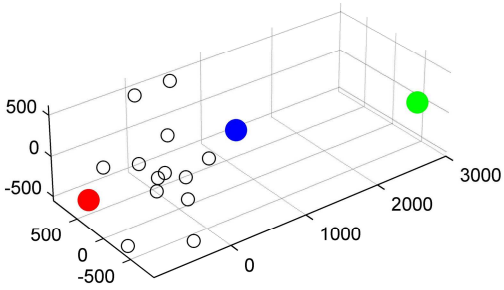
(d) Threshold = 0.50

Figure 5.10 Extent of CO_2 plumes for different threshold values

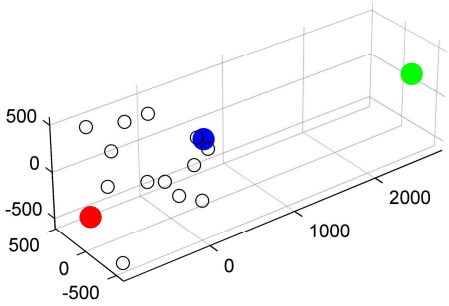
As described in the previous section, the 400 prior models are grouped using cluster analysis, and then the medoid closest to the observation is selected. Figure 5.11 shows the selected medoid, the observation, and the medoid of the cluster farthest from the observation in the MDS calculated 3D space. In Figure 5.11(a), all of the medoids obtained using the Euclidean distance have almost the same distance to the observation and they are not significantly distinguishable in terms of the dissimilarity to the observation.



(a) The Euclidean distance



(b) The Hausdorff distance

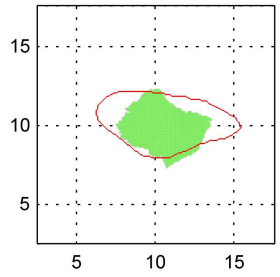
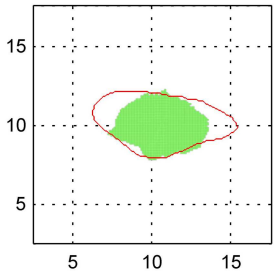
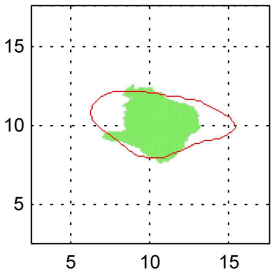
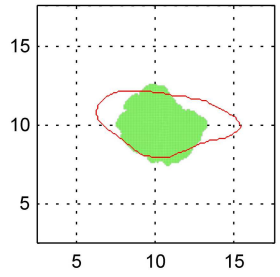
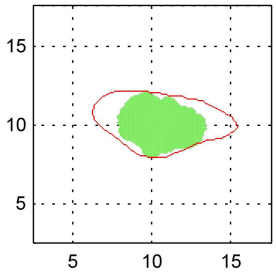
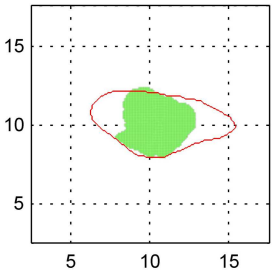
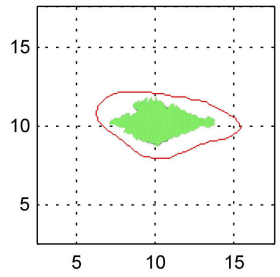
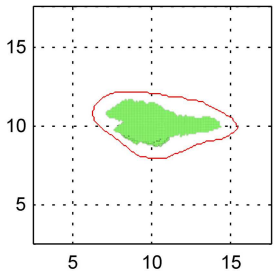
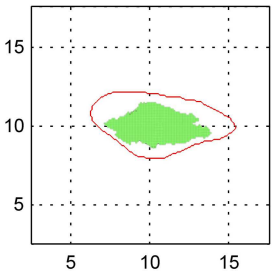


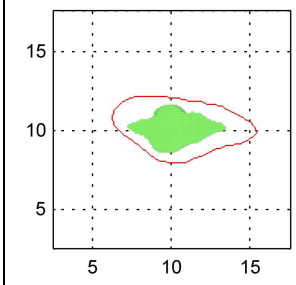
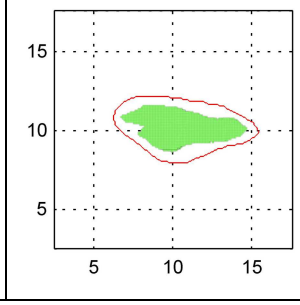
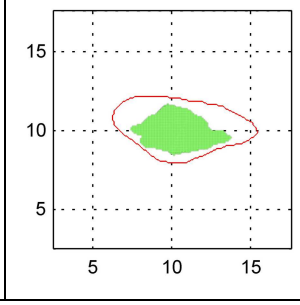
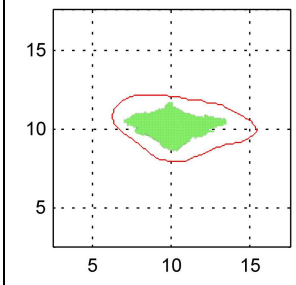
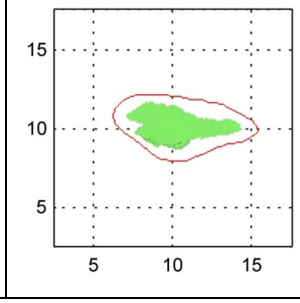
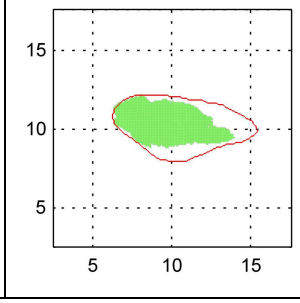
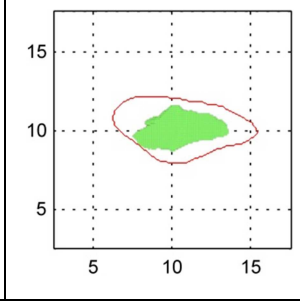
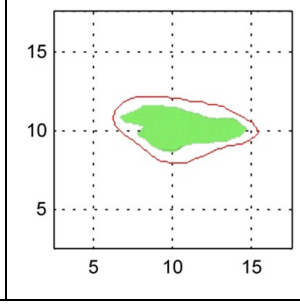
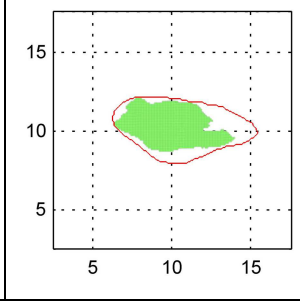
(c) The Hausdorff distance between perimeters

Figure 5.11 3D metric space obtained by MDS depicting the positions of the observation and the medoids obtained by cluster analysis. The green, blue, and red dots represent the observation, the selected medoid, and the medoid farthest from the observation.

Table 5.5 shows the CO_2 plumes of the selected medoid and its group members. The medoid closest to the observed plume computed using each distance measure corresponds to the green dot in the 3D space shown in Figure 5.11.

Table 5.5 CO_2 plumes computed using SCA and CMG-GEM for the medoid and the two group members of the group selected using the model selection process and the three approaches for the 2D case #2. The units of the axes are in kilometers. The red curve is the perimeter of the observed CO_2 plume shown in Figure 5.9.

	Selected Medoid	Group member #1	Group member #2
Euclidean distance	Proxy CO_2 plumes computed using SCA		
			
	Corresponding CO_2 plumes computed using CMG-GEM		
			
Hausdorff distance	Proxy CO_2 plumes computed using SCA		
			
	Corresponding CO_2 plumes computed using CMG-GEM		

			
Hausdorff distance between perimeters	Proxy CO_2 plumes computed using SCA		
			
	Corresponding CO_2 plumes computed using CMG-GEM		
			

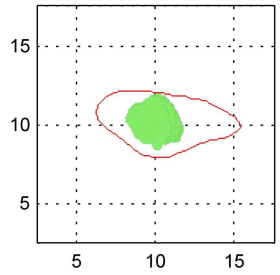
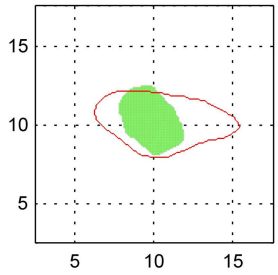
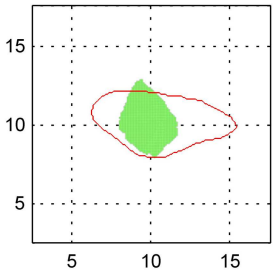
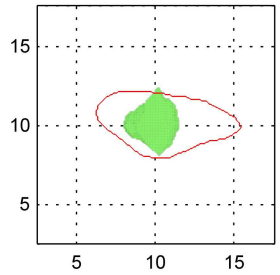
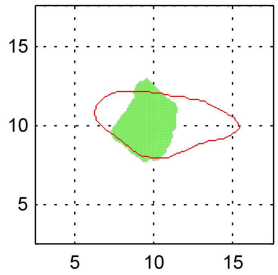
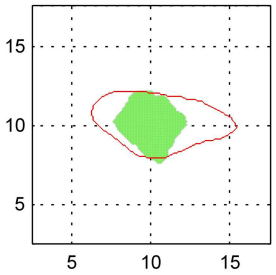
The major orientation of the CO_2 plumes selected using the Euclidean distance is almost isotropic. This is because the dissimilarity measured using the Euclidean distance is mostly affected by how many grid blocks are saturated within the observed CO_2 plume rather than the overall shape of the simulated CO_2 plumes.

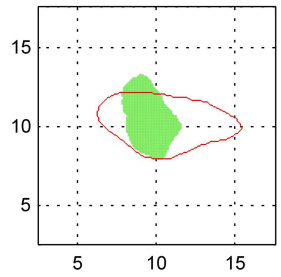
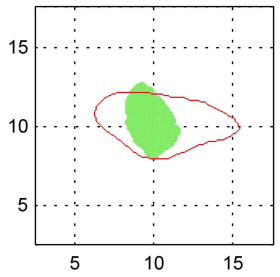
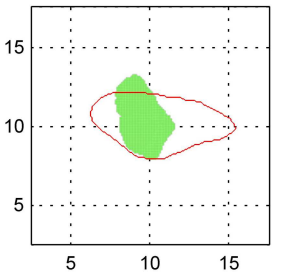
The CO_2 plume selected using the Hausdorff distance and the Hausdorff distance between perimeters is slightly different from the observed CO_2 plume in terms of size, but their perimeters have similar shape. Even though CO_2 plumes have slightly different

extent, the Hausdorff distance and the Hausdorff distance between perimeters are able to yield selected models that exhibit similar extension of the CO_2 plume.

Table 5.6 shows the CO_2 plume of the medoid farthest from the observation and response of some group members. The medoid in each approach corresponds to the green dot in the 3D space shown in Figure 5.7. The major orientation of the CO_2 plume shown in Table 5.6 is the north-south direction while the major orientation of the observed CO_2 plume is the east-west direction. All three measures are equally effective in discriminating the cluster exhibiting the maximum dissimilarity to the observed plume.

Table 5.6 CO_2 plumes computed using CMG-GEM for the medoid and the two group members of the group farthest from the observation for the 2D case #2. The unit of the axes is km. The red curve is the perimeter of the observed CO_2 plume shown Figure 5.6.

	Farthest medoid	Group member #1	Group member #2
Euclidean distance			
Hausdorff distance			

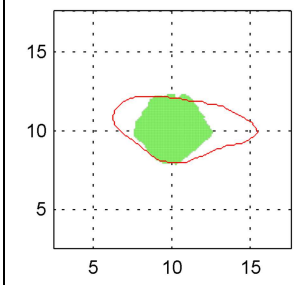
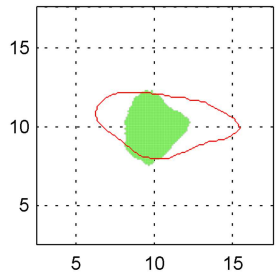
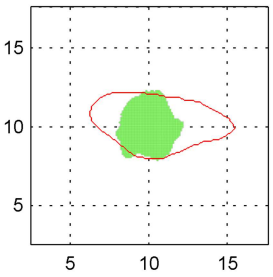
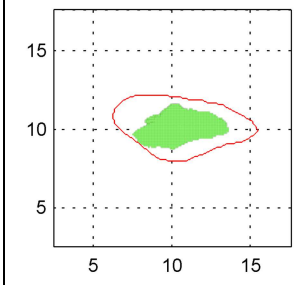
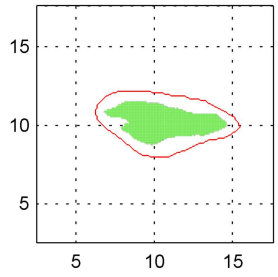
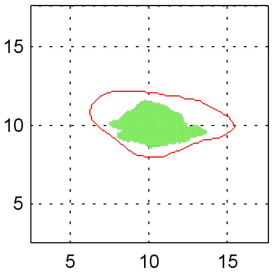
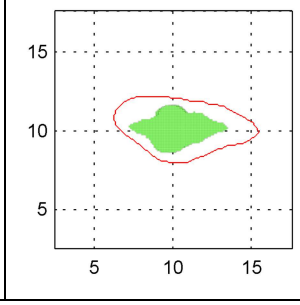
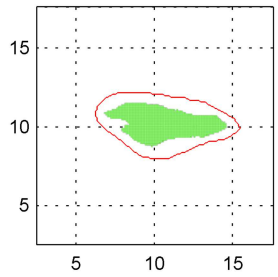
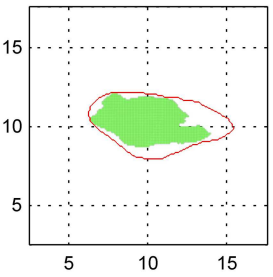
Hausdorff distance between perimeters			

The computational costs of the three approaches for this case were almost the same to those for the 2D case #1. The Hausdorff distance between perimeters spent about 2% of the computational cost of the Hausdorff distance measuring the dissimilarity between the 2D CO_2 plumes. In terms of the total computational cost of the model selection process, the Hausdorff distance between perimeters saved about 54% compared to the Hausdorff distance.

As shown in Figure 5.10, the images of CO_2 plumes can be different according to threshold values where the threshold values are used to threshold CO_2 saturation maps computed using CMG-GEM. Table 5.7 shows the results for a CO_2 saturation threshold of 0.2 while Table 5.6 shows the results for a CO_2 saturation threshold of 0.1. The Hausdorff distance and the Hausdorff distance between perimeters present the similar clustering results regardless of the threshold values the Euclidean distance presents the different clustering results. The Hausdorff distance based measures is advantageous over the Euclidean distance because the Hausdorff distance based measures is less sensitive to threshold values than the Euclidean distance.

Table 5.7 CO_2 plumes computed using CMG-GEM for the medoid and the two group members of the group farthest from the observation for the 2D case #1 using threshold = 0.2. The unit of the axes is km. The red curve is the perimeter of the observed CO_2 plume shown Figure 5.6.

	Farthest medoid	Group member #1	Group member #2

Euclidean distance			
Hausdorff distance			
Hausdorff distance between perimeters			

5.5. CONCLUSIONS

The objective of model selection is to find a set of most probable models that honor the visual and spatial characteristics of an observed 2D CO_2 plume. In order to select models similar to an observed CO_2 plume, a measure that computes the shape dissimilarity between the CO_2 plumes computed over suite of models is necessary. A good distance measure should result in accurate grouping of models based on the plume characteristics and also be able to result in a robust selection of the group of models that exhibit the closest characteristics to the observation. However, previous measures for

shape dissimilarity of contours in computer vision are not applicable to CO_2 plumes because they are invariant to rotation, scale, and location of an object.

The shape dissimilarity between 2D CO_2 plumes can be measured using the Hausdorff distance that exhibits sensitivity to rotation, scale, and location. However, the cost for computing the Hausdorff distance can become prohibitive if the number of grid blocks used for describing the CO_2 plume increases. The number of grid blocks describing the CO_2 plumes can be significantly reduced if the shape of the CO_2 plumes can be described using a perimeter. The perimeters of the CO_2 plumes are used to measure the shape dissimilarity between the CO_2 plumes instead of the CO_2 plumes themselves.

In the examples presented in this chapter, the Euclidean distance failed to find the models that have similar spatial characteristics to the observed 2D CO_2 plumes. The Hausdorff distance, and the Hausdorff distance between perimeters were successful in selecting the group of the most probable models honoring the observed 2D CO_2 plumes. However, the Hausdorff distance between perimeters used about 2% of the computational cost of the Hausdorff distance for measuring the dissimilarity between the 2D CO_2 plumes. Overall, the model selection process using the Hausdorff distance between perimeters saved about 54% of the computational cost of the model selection process compared to that using the Hausdorff distance.

Chapter 6. A Shape Dissimilarity Measure between 3D Fluid Displacements Using Surfaces and Skeletons

In the previous chapter, given a 2D observed CO_2 plume, measures for the dissimilarity between 2D CO_2 plumes were presented. In this chapter, measures for the dissimilarity between 3D CO_2 plumes are introduced. We explore the calculation of Hausdorff distance based on the surfaces and the skeletons of fluid displacements.

6.1. SURFACE

A surface voxel has at least one neighbor voxel with a zero attribute value in a 3D binary image where a voxel has six neighbor voxels (up, down, north, south, east, and west). Contrary to a perimeter, the surface of a CO_2 plume is influenced by the vertical characteristics of the CO_2 plume. However, because a surface has more voxels with non-zero attributes than a perimeter, it takes more computational cost to calculate the Hausdorff distance between surfaces than perimeters.

The procedure for measuring the dissimilarity between the surfaces of two CO_2 plumes using the Hausdorff distance is as follows:

i) Fill holes in 3D binary images

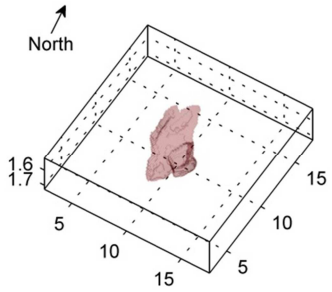
ii) Find surface voxels

iii) Calculate the Hausdorff distance between two surfaces

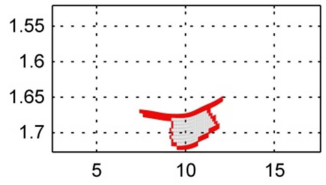
In the first step, holes in a 3D binary image are the zero voxels that are not connected to the surface. The holes are filled because the surface of a CO_2 plume is assumed to be closed.

Figure 6.1(a) through Figure 6.1(d) show the surfaces of two CO_2 plumes that are used to calculate the perimeters in Figure 5.2 and cross-sectional slices through the

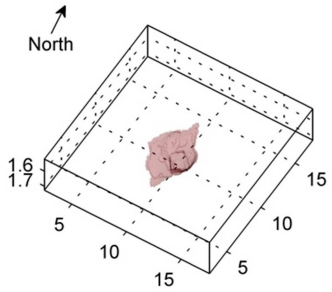
plumes where the surface voxels are colored in red. Each surface has about 60% of the total number of the non-zero value voxels in each 3D CO_2 plume image.



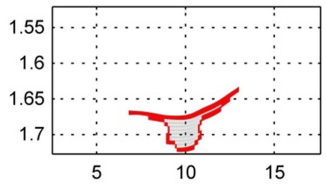
(a) Surface of 1st CO_2 plume



(b) Slice through the 1st CO_2 plume in the north-south direction at the center. The surface of the plume is shown in red.



(c) Surface of 2nd CO_2 plume



(d) Slice through the 2nd CO_2 plume in the north-south direction at the center. The surface is shown in red.

Figure 6.1 Surfaces of two CO_2 plumes and slices through the plume at the center. The surface pixels are shown in red.

6.2. SKELETON

A skeleton is calculated by thinning or removing voxels in a 3D image until the image does not change while retaining topological equivalence. A skeleton is also called a medial axis. Similar to a surface, a skeleton can also represent the vertical characteristics of a CO_2 plume, but it has less non-zero voxels than a surface.

We compute a skeleton using Homann's program (Homann, 2007), which is an implementation of a thinning algorithm presented by Lee et al. (1994). In the thinning algorithm, a voxel is removed from a skeleton if the following conditions are satisfied:

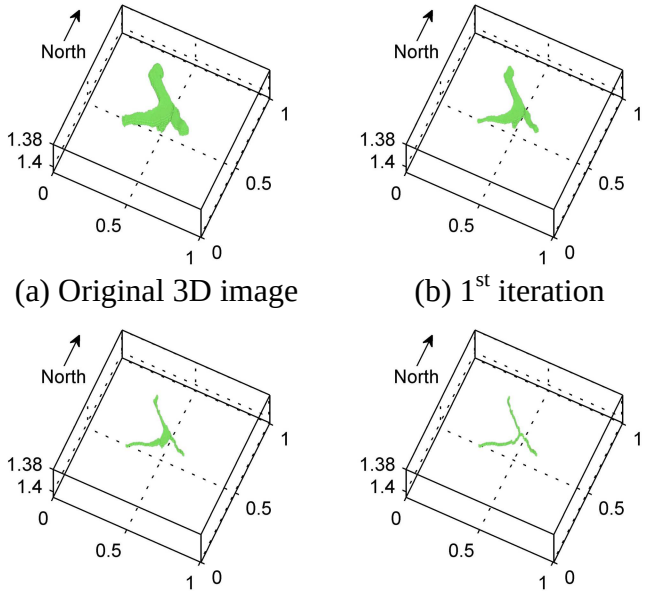
i) if the voxel is a surface voxel

ii) if the voxel is not the end of a line

iii) if the voxel does not change the Euler characteristic

iv) if the voxel does not change the number of connected objects

These steps are repeated until the 3D binary image does not change any longer. In the 3rd step, the Euler characteristic is a number representing a topological space's shape (Ivancevic and Ivancevic, 2008). The Euler characteristic is checked to keep the topological shape of the original object. Figure 6.2 shows the thinned 3D images after the 1st, 2nd, and final iterations computed using the above thing algorithm.



(a) Original 3D image (b) 1st iteration

(c) 2nd iteration (d) Final

Figure 6.2 An example of the thinning process

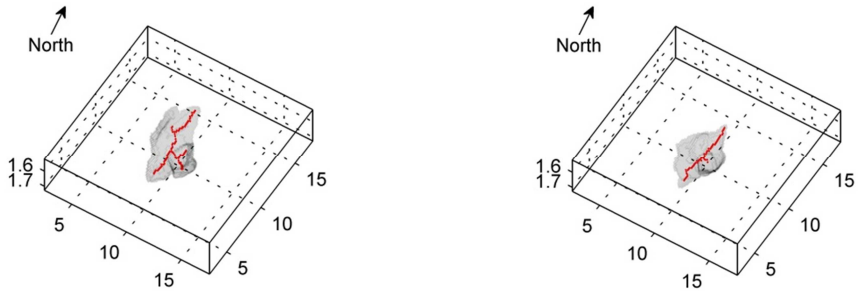
The procedure of measuring the dissimilarity between the skeletons of two CO_2 plumes using the Hausdorff distance is as follows:

i) Calculate skeletons

ii) Find skeleton voxels

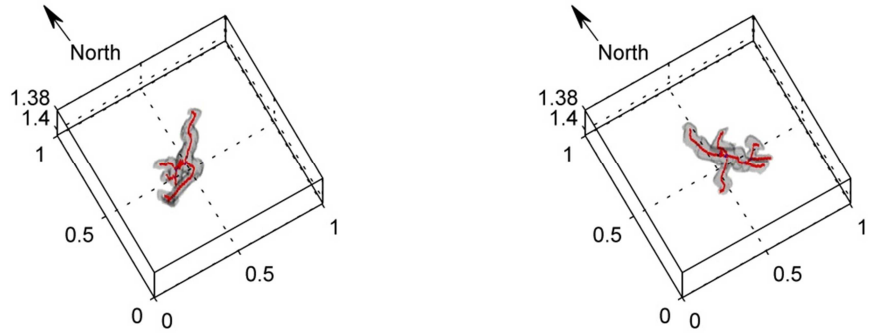
iii) Calculate the Hausdorff distance between two skeletons

Figure 6.3(a) and Figure 6.3(b) show the skeletons of the two CO_2 plumes that are used to calculate perimeters in Figure 5.2 where the skeleton voxels are colored in red. Each skeleton has about 5% of the total number of the non-zero voxels in the corresponding 3D CO_2 plume image. In Figure 6.3(a) and Figure 6.3(b), the 2nd skeleton (left) has less branches than the 1st skeleton (right). Although the western region in the 2nd CO_2 plume is saturated with CO_2, the 2nd skeleton does not have non-zero valued grid blocks in the western region. The 2nd skeleton does not adequately represent the spatial characteristics of the 2nd CO_2 plume when compared to the 1st CO_2 plume. Skeletons are more appropriate to represent channel like CO_2 plumes shown in Figure 6.4.



(a) Skeleton of 1st CO_2 plume (b) Skeleton of 2nd CO_2 plume

Figure 6.3 Skeletons of two CO_2 plumes



(a) Skeleton of 1st CO_2 plume (b) Skeleton of 2nd CO_2 plume

Figure 6.4 Skeletons of two channel like CO_2 plumes

6.3. MODEL SELECTION ALGORITHM

The basic process of selecting the most probable models is the same to the process presented in section 4.4, but the methods to assess the dissimilarity between CO_2 plumes are different. The procedure is as follows:

1) Run SCA in prior models
2) Measure the dissimilarity between all pairs of the proxy CO_2 plumes
3) Project the models onto a N-dimensional space using MDS
4) Group the models in the N-dimensional space using the k-mean clustering algorithm
5) Conduct fully physics flow simulations in the representative model for each group
6) Convert the simulated CO_2 saturation images to 3D binary images by thresholding the full physics simulated CO_2 saturation images of the representative models
7) Measure the dissimilarity between the observed CO_2 plume and the 3D binary images of the representative models

8) Select the representative model that has the most similar CO_2 plume to the observed CO_2 plume

9) Use the most probable models to assess the uncertainty in CO_2 migration

In the 2nd and 7th steps, the dissimilarity between 3D CO_2 plumes is measured using the following methods:

i) The Euclidean distance between 3D CO_2 plumes

ii) The Hausdorff distance between 3D CO_2 plumes

iii) The Hausdorff distance between the surfaces of 3D CO_2 plumes

iv) The Hausdorff distance between the skeletons of 3D CO_2 plumes

6.4. APPLICATION

The aquifer structure shown in Figure 6.5 has 201 × 201 × 10 grid cells with Δx = 5m, Δy = 5m, and Δz = 2m. CO_2 is injected at the middle of the aquifer model at 18,690.8 kg/day for 1 year. The boundary is connected to an infinite-acting aquifer. The relative permeability curves are shown in Figure 3.9 and capillary pressure is ignored.

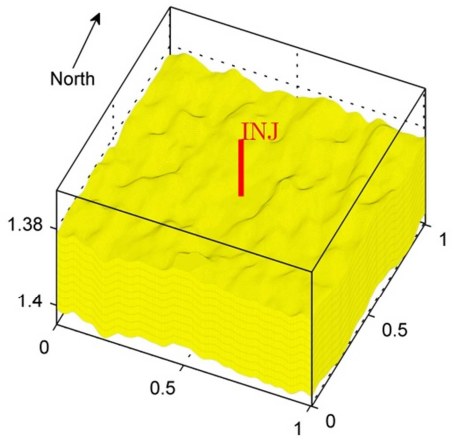
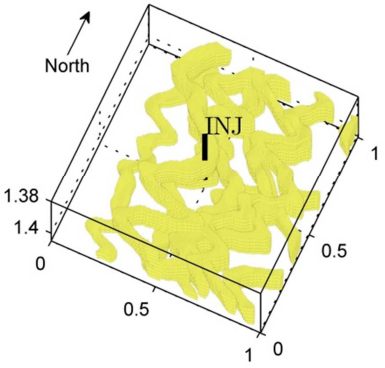
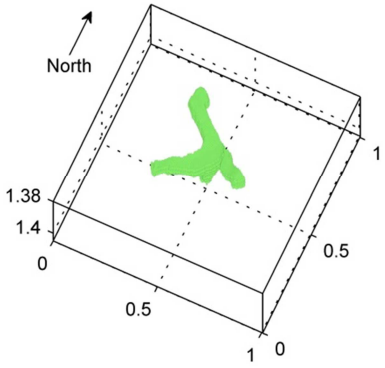


Figure 6.5 Structure of the aquifer for the 3D application example and the location of the CO_2 injector. The units of the axes are kilometers. The z-axis is exaggerated by a factor of 20.

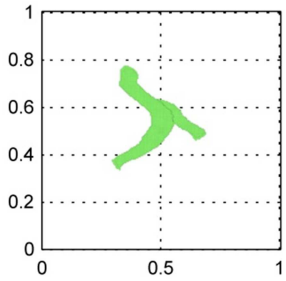
Figure 6.6 shows facies and the observed CO_2 plume for the reference model. In Figure 6.6(a), yellow and background are sandstone and mudstone, respectively. The major orientation of fluvial channels is north-south, and their thickness and width are 5*Δz and 15*Δx, respectively. The fluvial model is built using the TiGenerator in SGeMS (Boucher et al., 2010). k_h of sandstone and mudstone are 300 md and 1 md respectively, and the porosities are 0.2 and 0.1, respectively. The ratio of k_v to k_h is assumed to be 0.1. The observed CO_2 plume shown in Figure 6.6(b) is obtained by simulating flow in the reference model using CMG-GEM and then applying a CO_2 saturation threshold of 0.1.



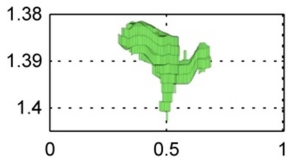
(a) Facies



(b) 3D view of observed CO_2 plume



(c) Areal view of observed CO_2 plume

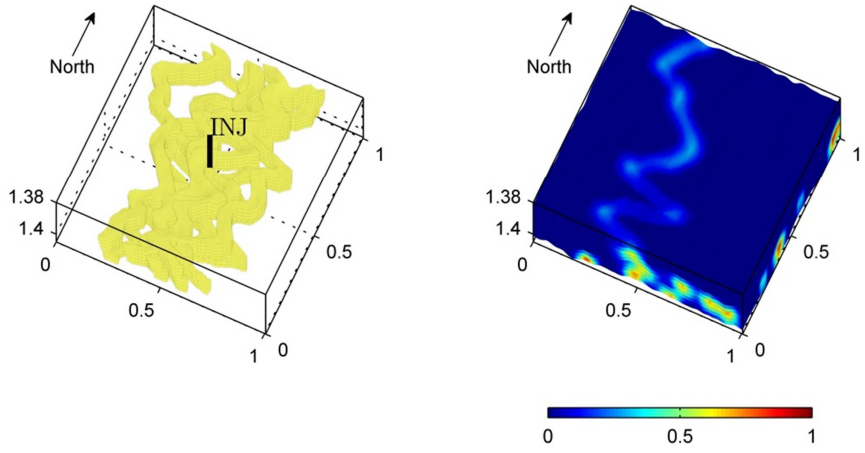


(d) Side view of observed CO_2 plume from south

Figure 6.6 Facies and CO_2 plume simulated in the reference model for the 3D case. The units for the axes are in kilometers. The z-axis in (d) is exaggerated by a factor of 20.

The CO_2 plume shown in Figure 6.6(b) through Figure 6.6(d) is used as observed data. In Figure 6.6(c), it can be observed that the injected CO_2 trifurcates in three directions from the injector located at the middle of the aquifer model to the north-west, south-west, and south-east directions. The flow in the south-eastern direction is over a shorter length than in the other directions. In Figure 6.6(d), the injected CO_2 migrates up and then migrates horizontally along the top of the aquifer.

Figure 6.7 shows data used for simulating the prior models. The training image shown in Figure 6.7 (a) is generated using the TiGenerator with the parameters that are used to build the true model. However, because of the stochastic nature of the TiGenerator, the training image for the prior models exhibits different characteristics than the reference. The probability of sandstone shown in Figure 6.7(b) is obtained by filtering the true facies shown in Figure 6.6(a) with a moving average filter with window dimensions 15*Δx, 15*Δy, and 5*Δz. The probability of mudstone is calculated by subtracting the probability of sandstone from one. The true facies at the injector (101,101,1) ~ (101,101,10) is assumed known. k_h, k_v, and ϕ of sandstone and mudstone are also assumed known.

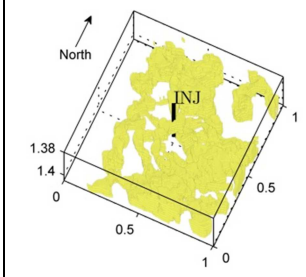
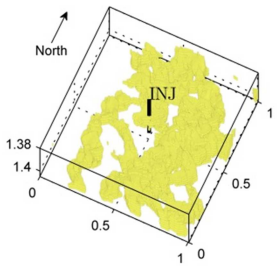
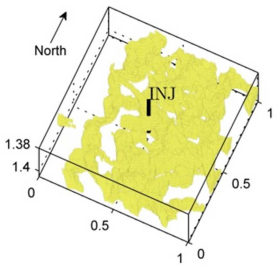
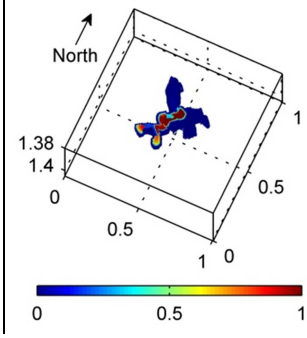
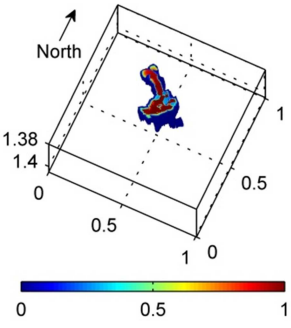
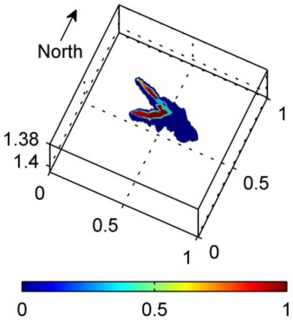


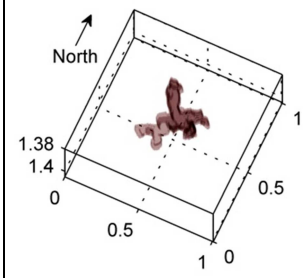
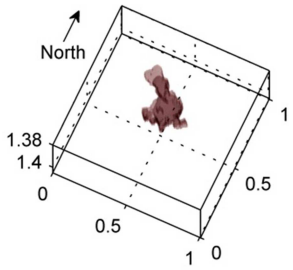
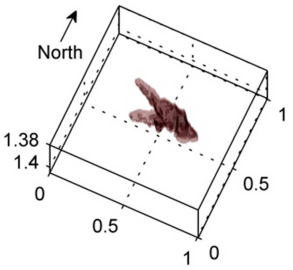
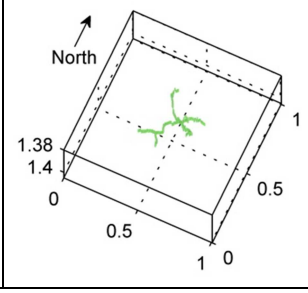
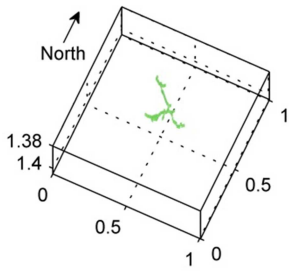
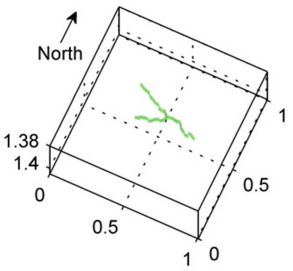
(a) Training image (b) Probability of sandstone

Figure 6.7 Training image and probability of sandstone for simulating prior models using SNESIM in the 3D case. The unit of the axes is km.

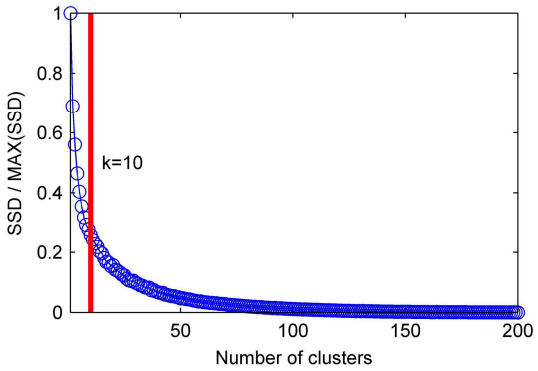
200 models honoring the prior data are generated using the single normal equation simulation (SNESIM) in SGeMS. The first row of Table 6.1 shows facies in three models out of that suite. The middle row of Table 6.1 shows the regions where the CO_2 saturation is greater than 0.01. That threshold value for CO_2 saturation is used in order to determine if a grid block is saturated with CO_2. The simulated 3D CO_2 saturation maps are converted to binary images by applying the same threshold. The last two rows of Table 6.1 show the surfaces and skeletons of the 3D binary images.

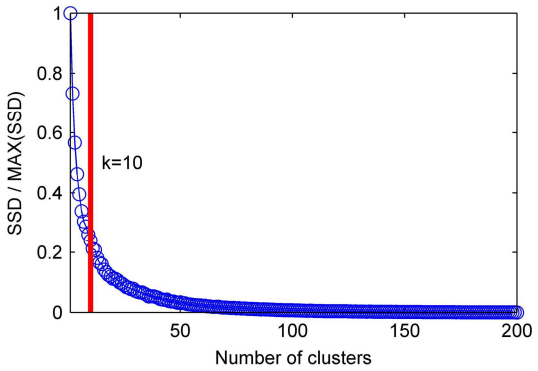
Table 6.1 Three facies models simulated using SNESIM, CO_2 saturations simulated using CMG-GEM, surfaces and skeletons of the CO_2 plumes. The unit of the axes are in kilometers.

	1st model	2nd model	3rd model
Facies			
$S_{CO_2} > 0.1$			

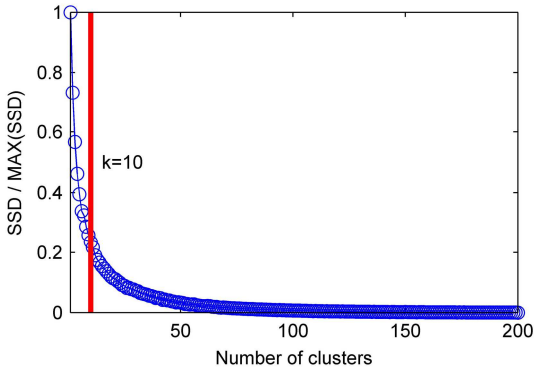
Surface of CO_2 plume			
Skeleton of CO_2 plume			

The CO_2 plume migrations are approximated on the 200 prior models using SCA. The distance measures between all pairs of prior models are computed and subsequently subjected to MDS and cluster analysis. As shown in Figure 6.8, the optimal numbers of clusters determined using the elbow method are 10 for all the four approaches. The 200 geologic models are clustered to 10 groups using the k-means clustering algorithm, and then only the medoids of the groups are processed through the full physics simulator. The full physics simulated CO_2 plumes of the medoids are compared to the observed CO_2 plume shown in Figure 6.6.

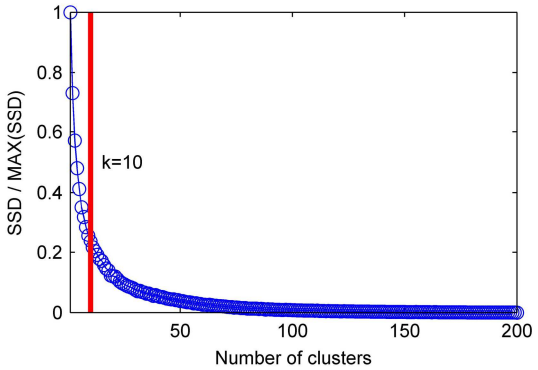




(a) Euclidean distance



(b) Hausdorff distance

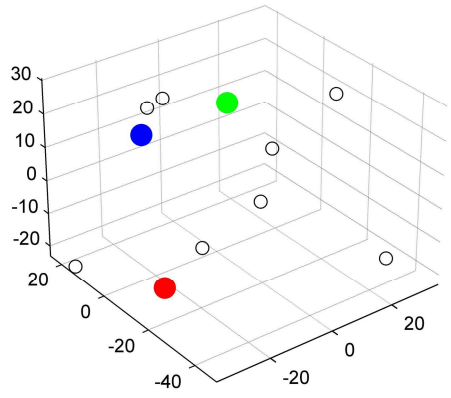


(c) Hausdorff between surfaces

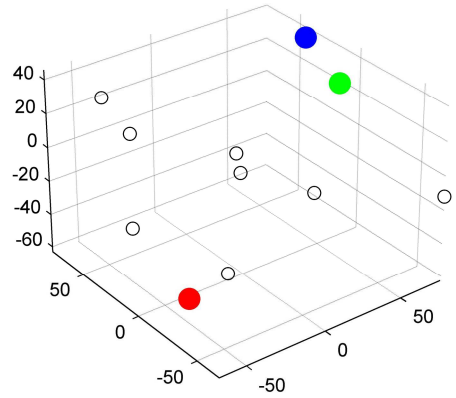
(d) Hausdorff between skeletons

Figure 6.8 Relative SSD (=SSD over the maximum value of SSD) of the clusters grouped based on the distance matrix measured using the four approaches and the optimal number of clusters

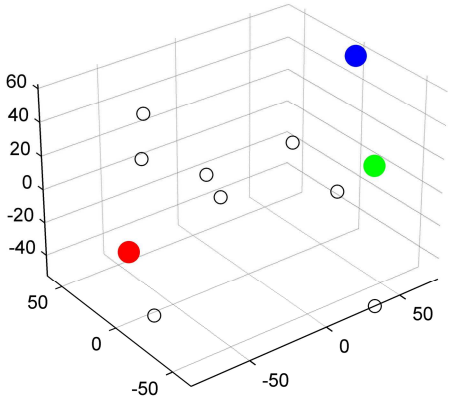
Figure 6.9 shows the dissimilarity between the observed and representative CO_2 plumes in the 3D spaces that are projected using MDS based on the distance matrix computed using each distance measure. Figure 6.9 showing the 3D metric space of the observations and the 10 medoids is plotted only for visualization because the dimension (=10) of the space used to project the 200 models using MDS is too high to visualize the locations of the 200 models in the space. The 3D space describes the shape dissimilarity between the 3D CO_2 plumes corresponding to the observation and the 15 medoids computed using each distance measure. The green dot corresponds to the observation, and the blue and red dots correspond to the closest medoid and the farthest medoid to the observation. The group members of the medoid corresponding to the blue dot are selected as the most probable models.



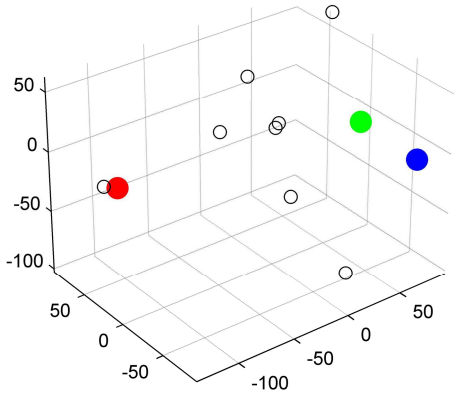
(a) Euclidean distance



(b) Hausdorff distance



(c) Hausdorff distance between surfaces

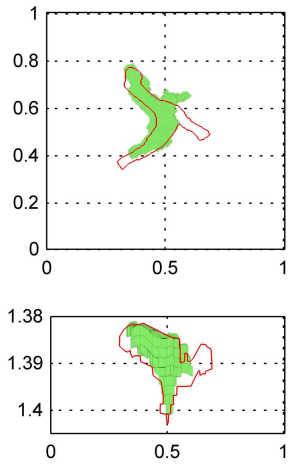
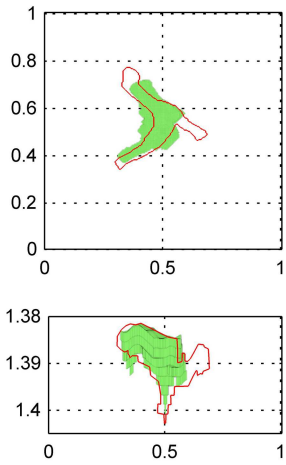
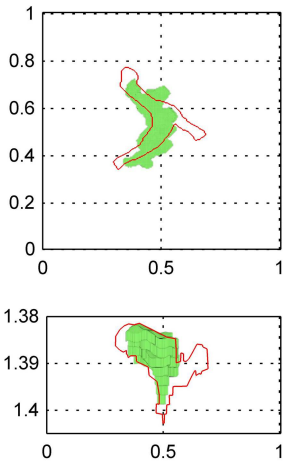
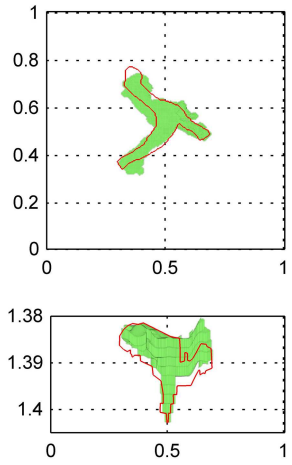
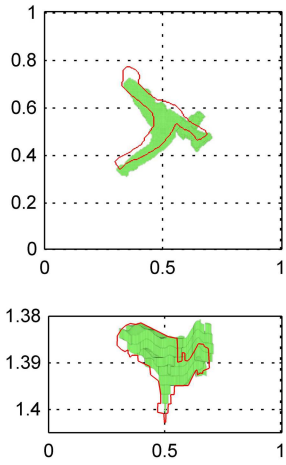
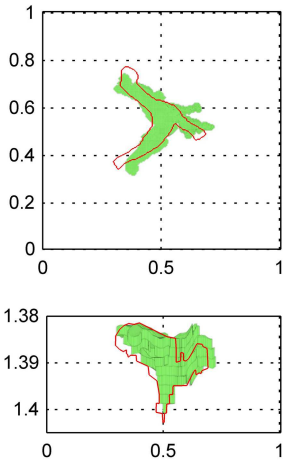
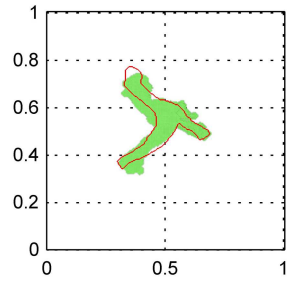
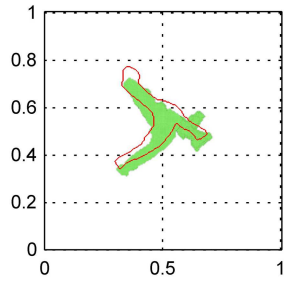
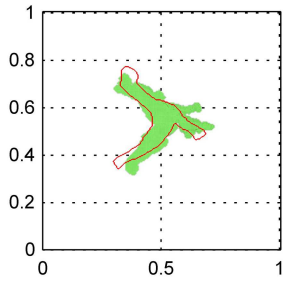


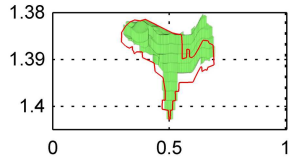
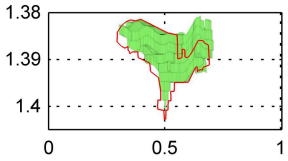
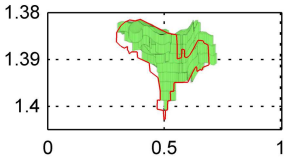
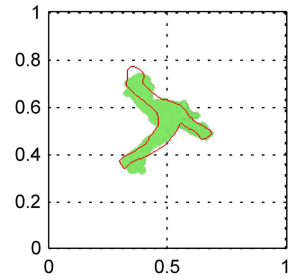
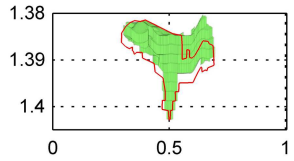
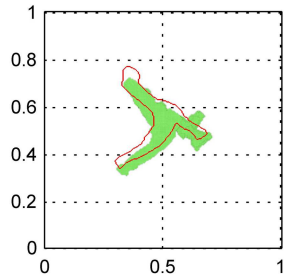
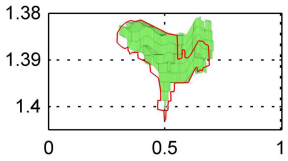
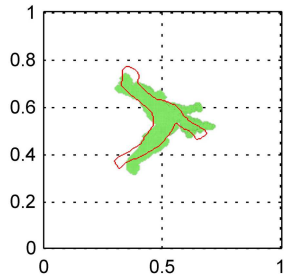
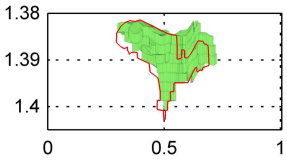
(d) Hausdorff distance between skeletons

Figure 6.9 3D metric space obtained by MDS depicting the positions of the observation and the medoids obtained by cluster analysis. The green, blue, and red dots represent the observation, the selected medoid, and the medoid farthest from the observation.

Table 6.2 shows the top and side views the CO_2 plumes computed using CMG-GEM for the medoid and two group member of the selected group corresponding to the blue dot in the 3D space shown in Figure 6.9.

Table 6.2 Top and side (from south) views of the CO_2 plumes computed using CMG-GEM for the medoid and two group members of the selected group. The results corresponding to the four distance measures are shown. The units for the axes are in kilometers. The red curve represents the observed CO_2 plume. The z-axis in the side view is exaggerated by a factor of 20.

	Selected Medoid	Group memebr #1	Group memebr #2
Euclidean distance			
Hausdorff distance			
Hausdorff distance between surfaces			

			
Hausdorff distance between skeletons	 	 	 

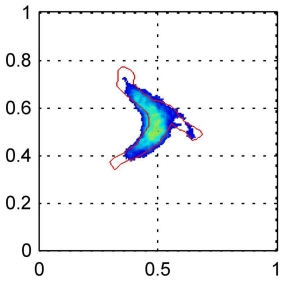
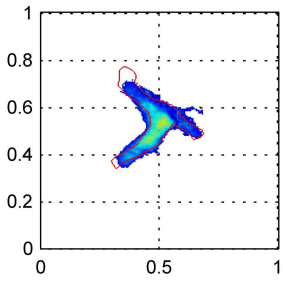
In Table 6.2, the CO_2 plumes selected based on the Euclidean distance are well confined within the boundary of the observed CO_2 plume. However, the most probable models selected using the Euclidean distance do not reflect the south-eastern flow. The difference in the vertical extent looks large due to the 20 fold exaggeration of the vertical axis. In reality, the difference between the vertical extent of the observed and simulated CO_2 plumes is small compared to the match in the horizontal extent. However, the Euclidean distance does not discriminate between the horizontal and vertical extent of the CO_2 plume and consequently models that exhibit excellent match in the horizontal extent of the plume may not be selected because they differ in the vertical migration characteristics. In this regard, unless prior models have CO_2 plumes that are almost identical to the observed CO_2 plume, the Euclidean distance fails to find models that yield plumes similar to the observed CO_2 plume.

Contrary to the Euclidean distance, the CO_2 plumes selected using the Hausdorff distance based approaches have the similar horizontal flow characteristics to the observed

CO_2 plume. The models selected based on the Hausdorff distance do not seem to exhibit the vertical migration characteristics of the observed CO_2 plume. However, this is understandable because the vertical thickness of the aquifer is much smaller than the horizontal extent of the aquifer. The Hausdorff distance takes into consideration, this anisotropy in the geometry of the reservoir and consequently, the models are selected more based on the similarity of the horizontal extent of the plume.

In Table 6.2, two of the group members selected using the Hausdorff distance, the Hausdorff distance between surfaces, and the Hausdorff distance between skeletons are exactly the same. Because the groups selected using the Hausdorff distance, the Hausdorff distance between surfaces, and the Hausdorff distance between skeletons have the similar mean profiles of CO_2 saturation as shown in Table 6.3, they exhibit the similar CO_2 flow characteristics. On the other hand, the group selected using the Euclidean distance has small average values of CO_2 saturation in the east-south region. This implies that few group members selected using the Euclidean distance exhibit CO_2 migration in the east-south region.

Table 6.3 Mean profiles of the CO_2 saturation (>0.1) computed using CMG-GEM for the selected group members. The results corresponding to the four distance measures are shown. The units for the axes are in kilometers. The red curve represents the observed CO_2 plume. The z-axis in the side view is exaggerated by a factor of 20.

	Euclidean distance	Hausdorff distance
Top view		

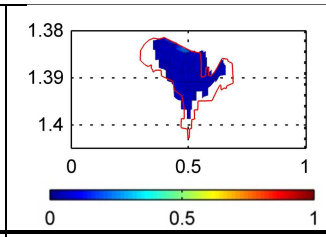
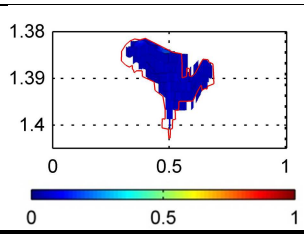
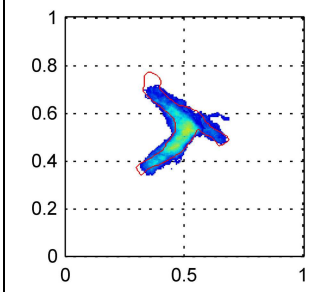
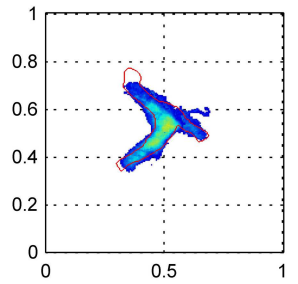
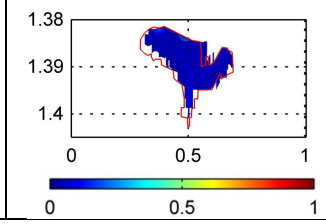
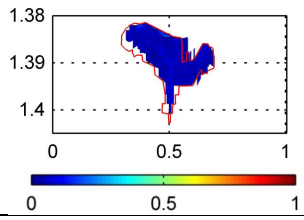
Side view from south		
	Hausdorff distance between surfaces	Hausdorff distance between skeletons
Top view		
Side view from south		

Table 6.4 shows the CO_2 plume of the medoid and two group members of the group corresponding to the green dot in the 3D space shown in Figure 6.9. The CO_2 plumes shown in Table 6.4 do not cover the horizontal extent of the observed CO_2 plume.

Table 6.4 CO_2 plumes computed using CMG-GEM for the medoid and two members of the group farthest from the observation . The units of the axes are in kilometers. The red curve represents the observed CO_2 plume shown Figure 6.6.

	Farthest medoid	Group member #1	Group member #2

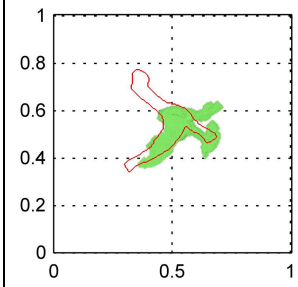
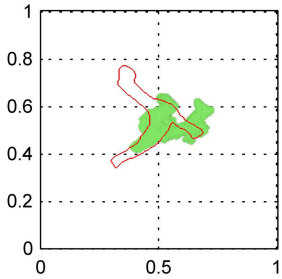
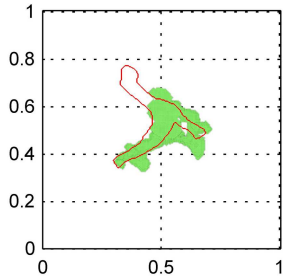
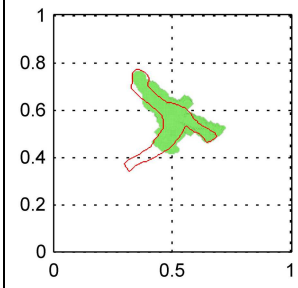
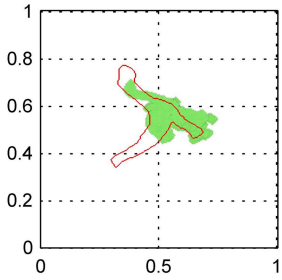
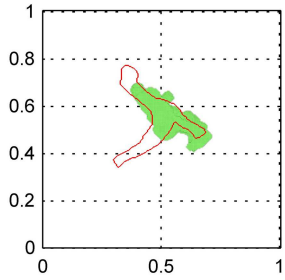
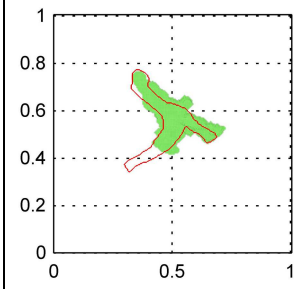
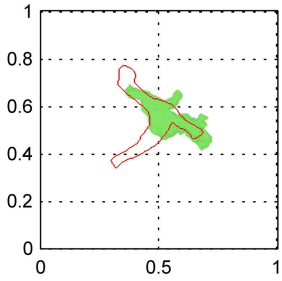
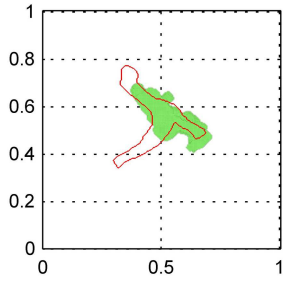
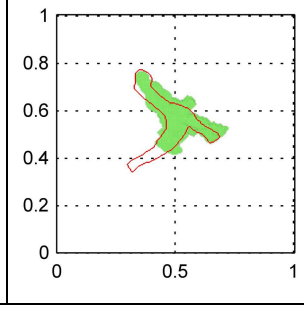
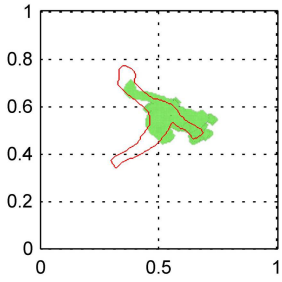
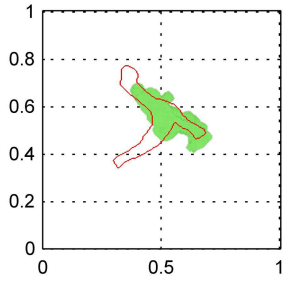
Euclidean distance			
Hausdorff distance			
Hausdorff distance between surfaces			
Hausdorff distance between skeletons			

Figure 6.10 shows the computational costs associated with the four methods for computing the distance between the models. The process of calculating surfaces and skeletons of the 200 proxy CO_2 plumes took less than 30 seconds. The costs of calculating surfaces and skeletons are negligible compared to the cost of calculating the

dissimilarity measures and so they are not considered in the computational costs. The Hausdorff distance between surfaces and the Hausdorff distance between skeletons saved about 64% and 99.76% of the computational cost associated with computing the Hausdorff distance s although they yielded the similar results.

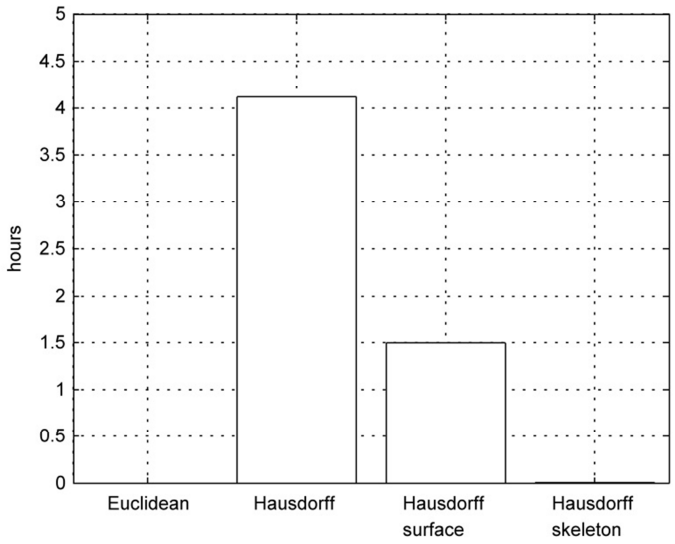


Figure 6.10 Computational costs of the four approaches for the 3D case

It took about 2500 seconds to run CMG-GEM simulations and about 4 seconds for calculating the plume using SCA for a geological model. The total computational cost of the model selection process using the Hausdorff distance is (200 * 4 + 10 * 1100 + 14821) / 3600 = 7.39 hours. The total computational cost of the model selection process using the Hausdorff distance between surfaces and skeletons is (200 * 4 + 10 * 1100 + 5388) / 3600 = 4.77 hours and (200 * 4 + 10 * 1100 + 36) / 3600 = 3.29 hours, respectively. The Hausdorff distance between surfaces and skeletons saved 35% and 56% of the total computational cost of the model selection process. More computational cost is expected to be saved if more grid cells are needed to resolve CO_2 plumes.

6.5. CONCLUSIONS

The objective of this chapter was to quickly select a set of the most probable models based on the 3D CO_2 plume characteristics. The Euclidean distance and the Hausdorff distance can also be applied to measure the dissimilarity between 3D CO_2 plumes.

As discussed in the previous chapter, the cost associated with the calculation of the Hausdorff distance is proportional to the number of grid cells describing CO_2 plumes. The computational cost of calculating the Hausdorff distance can be reduced by removing nonzero grid cells that are not necessary to represent the shape of CO_2 plumes. Surface and skeleton based approaches are used to obtain simplified representations of the 3D CO_2 plume.

The Euclidean distance, the Hausdorff distance, the Hausdorff distance between surfaces, and the Hausdorff distance between skeletons were applied to measure the dissimilarity between the CO_2 plumes migrating in a fluvial depositional environment. The Euclidean distance failed to find the models honoring the spatial characteristics of the observed 3D CO_2 plume.

The Hausdorff distance based approaches were successful in find the most probable models honoring the observed CO_2 plume migrating along the fluvial channel. The Hausdorff distance between surfaces and the Hausdorff distance between skeletons selected acceptable models at a fraction of the cost associated with the computation of the Hausdorff distance. The Hausdorff distance between surfaces and the Hausdorff distance between skeletons costed 36% and 0.24% of the computational cost of the Hausdorff distance. In terms of the total computational cost of the model selection process, the Hausdorff distance between surfaces and skeletons saved 35% and 56% of the cost, respectively. The Hausdorff distance between surfaces can be applied more widely than

the Hausdorff distance between skeletons because a skeleton might not provide an adequate shape representation of a more isotropic migration pattern.

Chapter 7. Development of a Software Module for SCA and Model Selection Process

In the previous chapters, a fast proxy (SCA) for approximating CO_2 plume migrations is introduced and then several approaches for measuring the dissimilarity between the CO_2 plumes are presented to select the most probable models honoring an observed CO_2 plume. In this chapter, a software module for the SCA and the model selection process will be detailed.

7.1. Stanford Geostatistical Modeling Software

The Stanford Geostatistical Modeling Software (SGeMS) is a free and powerful geostatistics software including most of geostatistics algorithms such as kriging, co-kriging, the sequential Gaussian simulation, and the single normal equation simulation. The software module for model selection using SCA was developed as a plugin in SGeMS. The main goal of SCA and the model selection is to quickly quantify the uncertainty in CO_2 plume migrations using a large suite of geologic models. The suite of geologic models is generated using the geostatistical algorithms of SGeMS. The generated geologic models are used in the SGeMS plugin for SCA and the model selection without any importing/exporting process. The development time for visualization of results is also saved by developing the SGeMS plugin.

The SGeMS interface has three primary panels, as shown in Figure 7.1. The left panel, called the Algorithms panel, contains all the geostatistical software under two main headings: estimation and simulation. There is a third heading called Utilities, which contains various smaller programs for post-processing and manipulation of models within the program. The center panel, called the Objects panel, lists the models that are part of

the current project. The visualization panel displays any selected realization within an object, and also allows actions like panning, zooming etc. of the image.

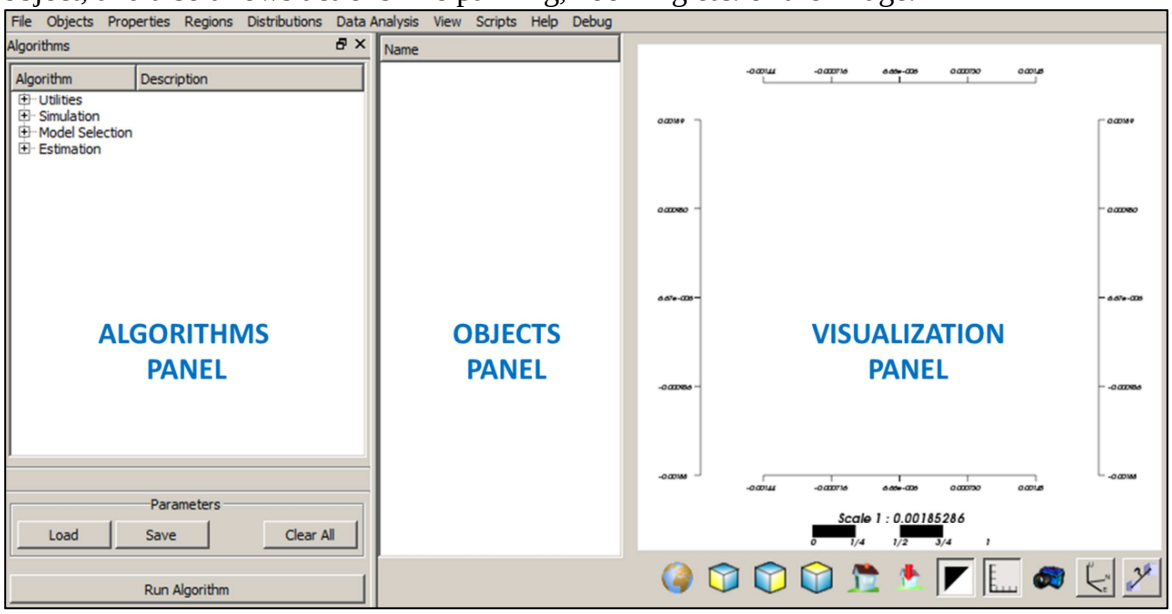


Figure 7.1 User interface for SGeMS

7.2. SGeMS PLUGIN FOR SCA AND MODEL SELECTION

7.2.1. Installation

Two filee, setup.exe and UTGS.msi, are needed to install the SGeMS containing the plugin for SCA and the model selection process. The installation files only work on the 64-bit version of windows. The instructions for the plugin installation are given in Appendix A.

7.2.2. Main input panel

The model selection algorithm is implemented as a separate heading in the algorithms panel shown in Figure 7.2. Once this option is selected, it brings up the input window for the model selection process, where the user needs to provide all the parameters and data used to the process. The main tab in the model selection input panel

contains drop-down menus to select the particular permeability and porosity objects to be used, together with details about the connectivity proxy and reservoir flow simulator to be used. The various input options are listed below, and referenced in Figure 7.2:

[1] Proxy selection: radio button menu to select the particular proxy to be used

[2] Permeability object: contains all the models in the initial model set

[3] Porosity object: similar to permeability object

[4] Unit system to be used for data: the code allows for using either field units or SI units.

[5] Type of K direction: two options (DOWN and UP). DOWN means K increase from top to bottom. UP means K increase from bottom to top.

[6] Depth of grid blocks: an object containing a single realization of depths to each grid block. It is assumed that the structure of the grid is same for all initial models. If the depth of grid blocks is given, the type of K direction is automatically set to DOWN. The depth of grid blocks can be obtained by exporting grid top or grid bottom data in CMG. If the random walker algorithm is selected, the depth of grid blocks must be given.

[7] Corner point file: This is optional. You can input the corner point data including the CMG keywords for the corner points such as COORD, ZCORN, and NULL. If the corner point file is given, the type of K direction is automatically set to DOWN. This parameter is available only in SCA.

[8] Observation file: Observed data. This can be chosen by clicking the right button "…" and browsing.

[9] Simulator input file: the input file of CMG-GEM simulator. This can be chosen by clicking the right button "…" and browsing.

[10] Perm include file name: the name of the permeability include file used in the input.

[11] Poro include file name: the name of the porosity include file used in the input.

[12] Simulator file: the path of the reservoir flow simulator to be used. Our current implementation uses CMG-GEM simulator. This can be chosen by clicking the right button "…" and browsing.

[13] CMG simulator command: the command to run a simulator. This must include [#simulator_file] and [#simulator_input_file]. Our plugin will replace [#simulator_file] with the actual path of the simulator you put in [9], [#simulator_input_file] with the actual path of the input file.

[14] Report input file: the input file of the RESULTS GRAPH© program in CMG. This can be chosen by clicking the right button "…" and browsing.

[15] # of header lines in report output: the number of header lines skipped in the report output file.

[16] Report file: the path of the RESULTS GRAPH© program in CMG. This can be chosen by clicking the right button "…" and browsing.

[17] CMG report command: the command to obtain the simulation results. This must include [#report_file], [#report_input_file], and [#report_output_file]. Our plugin will replace [#report_file] with the actual path of the RESULTS GRAPH© program you put in [13], [#report_input_file] with the actual path of the input file, [#report_output_file] with the actual path of the output file.

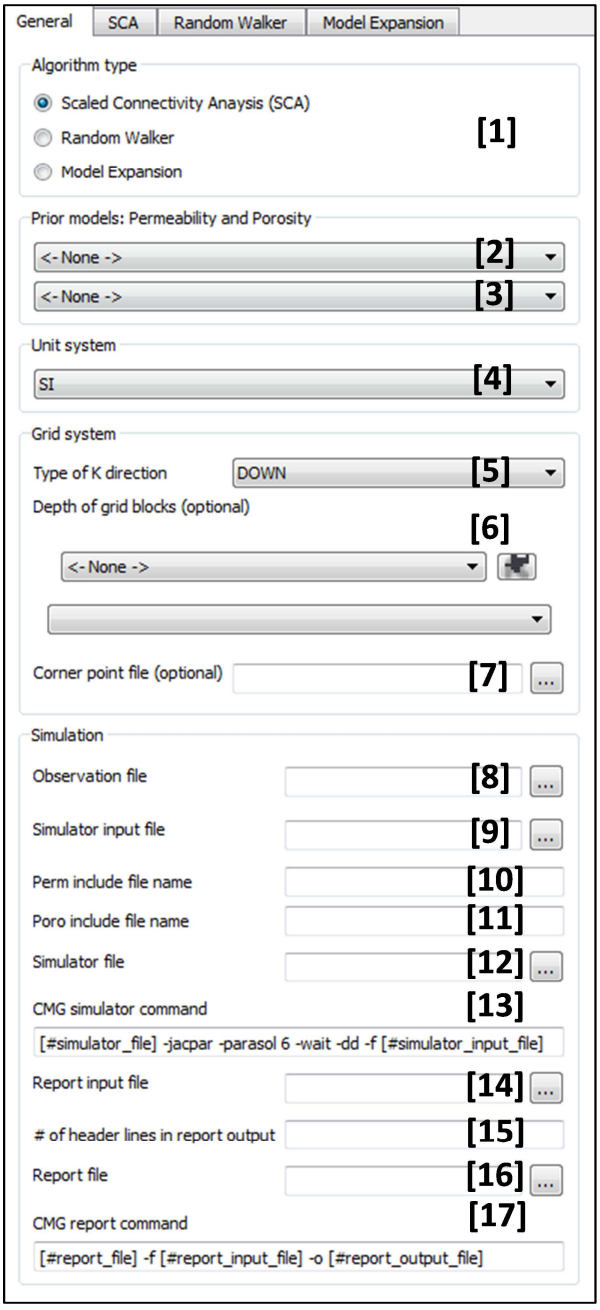


Figure 7.2 Main input panel for the model selection algorithm

7.2.3. Input panel for SCA and model selection

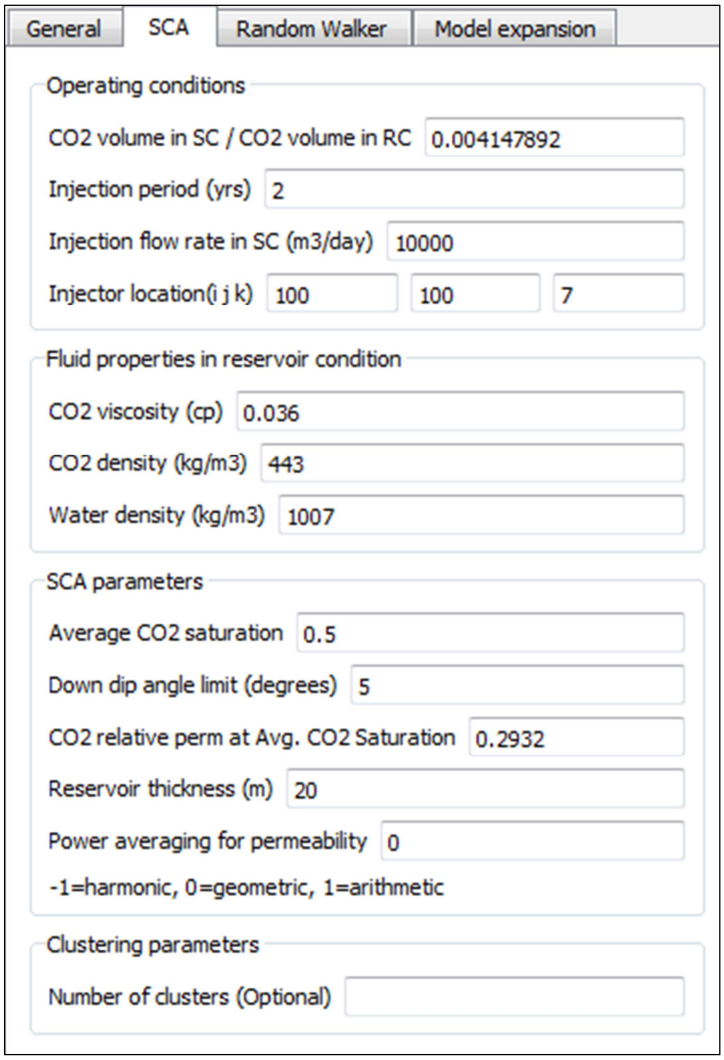


Figure 7.3 SCA tab for the model selection plugin

The SCA proxy is used when the algorithm type chosen in the general tab shown in Figure 7.2 is SCA. The interface is shown in Figure 7.3, and described below.

1) Operating conditions
    a) CO_2 volume factor (related to compressibility): the ratio of CO_2 volume at standard conditions (P=1atm and T=20°C) to the CO_2 volume at reservoir conditions.

b) Injection period: CO_2 injection period. SCA calculates the shortest paths from an injector to every grid block. In SCA, the path along which the travel time is greater than the injection period is ignored in order to save the computation time.

c) Injection flow rate in SC: CO_2 injection flow rate in standard conditions. The CO_2 injection flow rate is used to calculate the pressure gradient in the grid blocks.

d) Injector location: IJK coordinates. The first index starts from 1.

2) Fluid properties in reservoir conditions: CO_2 viscosity, CO_2 density, water density.

3) SCA parameters

a) Average CO_2 saturation

b) Down dip angle limit: maximum angle allowing down dip CO_2 flow. In SCA, if the down dip causes the negative potential ($\Delta p + \Delta \rho g h$), the flow does not take place in the direction. In reality, if the down dip is small or local, CO_2 plume is pushed in the down dip direction. Thus the potential is assumed to be Δp if the down dip angle is less than the down dip angle limit that cause a negative potential.

c) CO_2 relative perm at avg. CO_2 saturation: CO_2 relative permeability at the average CO_2 saturation.

d) Reservoir thickness: The reservoir thickness is used to calculate the pressure gradient at the grid blocks.

e) Power averaging for permeability: The average permeability is used to calculate the pressure gradient in the grid blocks. The geometric average is generally used for average permeability, but the power averaging for permeability can be varied according to the characteristics of reservoir heterogeneity. The power exponents for harmonic, geometric, arithmetic averages are -1, 0, 1, respectively.

4) Results Processing

a) Number of clusters: Optional. The number of clusters or groups.

7.3. CONCLUSIONS

In this chapter, the installation and input data of the SGeMS plugin for SCA and the model selection process were outlined. Because SGeMS is an open source software, any new plugin can be implemented in SGeMS. By developing the computer program for SCA and the model selection process as a SGeMS plugin, existing geostatistics algorithms and visualization tools can be used without redeveloping them. Geologic models can be also used in the SGeMS plugin without any importing process.

Chapter 8. Summary, Conclusions, and Recommendations

8.1. SUMMARY

The objective of this dissertation was to quickly quantify the uncertainty in the buoyant flow of injected fluid and select the most probable models honoring an observed fluid displacement. In order to achieve this objective, a fast alternative to full physics simulators is proposed.

The proposed proxy is validated for CO_2 plume migrations, which are the typical case for the buoyant flow of injected fluid. The CO_2 plume approximated using CA and SCA were compared to that computed using CMG-GEM. The reason why CA fails to reproduce the buoyant flow of the CO_2 plume was studied with a simple example.

SCA was tested for various combinations of injection periods, injection rates, CO_2 density, CO_2 viscosity, the degree of rock heterogeneity, permeability of cap rock, and topology of the aquifer. The results of SCA were compared to those obtained using the Fast Marching algorithm (FMM) and the Vertical Equilibrium (VE) model. SCA was also verified for two real field cases: the Johansen Formation and the Sleipner L9 model.

SCA was applied to quickly assess the uncertainty in the extent of CO_2 plumes in models that were generated using two-point and multi-point statistics. The P10, P50, and P90 of the CO_2 plume extent obtained using SCA and the ranking process are compared to those of the full physics simulation results. The results indicate that SCA can provide reliable ranking of reservoir models that can subsequently be used to bracket the uncertainty in performance prediction.

The Hausdorff distance and the Hausdorff distance based on reduced parametrization of the plume geometry are introduced to measure the shape dissimilarity between fluid displacements. They are applied and compared for CO_2 plume migrations. The Hausdorff distance based on reduced parameter approaches take less computational

time than the approach based on exhaustive computation of Hausdorff distance by removing unnecessary grid cells to describe the representative shape of CO_2 plumes. The Hausdorff distance-based approaches use perimeters, surfaces, skeletons for the representations of CO_2 plumes.

The Hausdorff distance between perimeters of the CO_2 plume were applied to find the reservoir models most similar to the observation. These models were then compared to the models identified on the basis of the Euclidean distance and the Hausdorff distance. The computation of the perimeter of CO2 plume shapes is based on the 2D projection of the 3D plume geometry. This assumes that there is not much variation in the vertical extent of the plume. The results indicate that the distance based on the perimeter is reliable when the vertical heterogeneity in the reservoir is not significant. The Hausdorff distances between surfaces and between skeletons extracted from the 3D plume geometry were also applied in order to retrieve the most probable reservoir models. These models were then compared to those obtained using the Euclidean distance and the Hausdorff distance for the 3D observed CO_2 plume. The skeleton-based Hausdorff distance can result in a quick assessment of the differences in dynamic characteristics of models especially in situations where reservoir heterogeneity has a dominating influence on the flow such as in channel systems. In such cases, the skeleton-based Hausdorff distance measure can be used to quickly find the most probable models honoring the observed CO_2 plume data through the model selection process.

The SGeMS plugin for the model selection process using SCA and the Hausdorff based approaches was developed. The dissertation provides a summary of the key inputs and the installation procedure for that software.

8.2. CONCLUSIONS

This section lists the significant conclusions drawn from the chapters of this dissertation.

8.2.1. Scaled Connectivity Analysis

[1] Previously developed proxies for reservoir flow are inappropriate to quickly approximate the buoyant flow of a CO_2 plume in 3D heterogeneous rock during an injection period.

[2] The buoyant flow of a CO_2 plume is represented in connectivity analysis by scaling edge weights with their local minimum.

[3] The approximate CO_2 plumes computed using SCA were close to to the full physics responses for the various combinations of injection periods, injection rates, a CO_2 density, a CO_2 viscosity, the degree of rock heterogeneity, permeability of cap rock, and topology of the aquifer.

[4] SCA successfully reproduced the buoyant flow of the CO_2 plume in the test cases while CA and the fast marching method failed to reproduce the buoyant flow.

[5] SCA captured the influence of the intermediate shale layer on the CO_2 plume migration in the 3D heterogeneous case, but the Vertical Equilibrium (VE) model did not.

[6] SCA did not work properly when the rock heterogeneity is weak or the aquifer has a significant influence on the CO_2 plume migration.

[7] SCA reproduced the essential characteristics of the CO_2 plumes in the Johansen Formation model and the Sleipner L9 model.

8.2.2. Quantification of Uncertainty in CO_2 Migration Using SCA

[1] SCA provided the proxy responses correlated to the extent of the CO_2 plumes computed using the full physics simulator in the geologic models using two-point and multi-point statistics.

[2] The P10, P50, and P90 of the extent of the CO_2 plumes in the representative models sampled using SCA and the ranking process were close to those of the full physics simulation results.

[3] About 90% of the total computational cost was saved in estimation of P10, P50, and P90 of the extent of the CO_2 plumes by sampling the representative models based on the proxy CO_2 plumes computed using SCA.

8.2.3. A Shape Dissimilarity Measure between 2D Fluid Displacements Using Perimeters

[1] The Hausdorff distance and the Hausdorff distance between perimeters were more accurate than the Euclidean distance in measuring the shape dissimilarity between the 2D CO_2 plumes.

[2] The most probable models selected using the Hausdorff distance and the Hausdorff distance between perimeters were almost the same.

[3] The Hausdorff distance between perimeters costed about 2% of the computational cost of the Hausdorff distance.

[4] The Hausdorff distance between perimeters results in the reduction of about 54% of the computational cost of the model selection process compared to that using the Hausdorff distance.

8.2.4. A Shape Dissimilarity Measure between 3D Fluid Displacements Using Surfaces and Skeletons

[1] The CO_2 plumes selected using the Hausdorff distance between surfaces and skeletons are more similar to the 3D observed CO_2 plume than those selected using the Euclidean distance.

[2] The Hausdorff distance between surfaces was more accurate than the Hausdorff distance between skeletons in measuring the shape dissimilarity between the 3D CO_2 plumes.

[3] The Hausdorff distance between surfaces and Hausdorff distance between skeletons saved about 64% and 99% of the computational cost associated with the computation of the exhaustive Hausdorff distance measuring the dissimilarity between the CO_2 plumes.

[4] The Hausdorff distance between surfaces and skeletons saved about 35% and 56% of the total computational cost of the model selection process using the Hausdorff distance.

8.3. RECOMMENDATIONS

The following is a list of recommendations for future research:

[1] SCA does not reproduce the radial migration of a CO_2 plume in homogeneous rock. If SCA considers the diagonal connections between neighboring grid cells, it could work properly when the rock heterogeneity is weak.

[2] The equation for an edge weight given in Equation (3-3) can be extended to consider more physics such as capillary pressure and chemical reaction between CO_2 and rock.

[3] SCA is designed to approximate CO_2 plume migrations when viscous forces are dominant during an injection period. However, in the post-injection period,

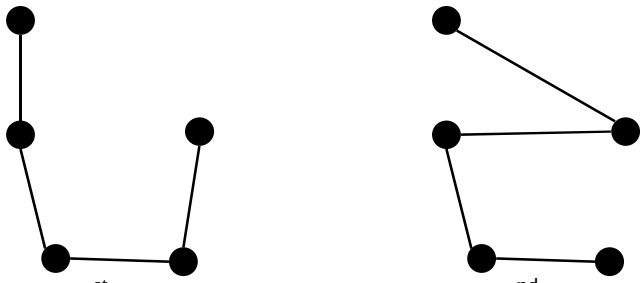
capillary and buoyant forces are dominant over viscous forces. The CO_2 plume migration in the post-injection period can be also approximated by applying conditions for CO_2 trapping or migration to the SCA approximation based on capillary and buoyant forces. For example, Ren et al. (2015) calculated the entry capillary pressure field based on the permeability at each grid block using the Leverett J-function. They estimated the locally trapped CO_2 in the SCA approximation by comparing the entry capillary pressure to the gravity potential.

[4] The model selection process was not applied for multiple times and CO_2 injectors. Time-lapse data for CO_2 plume migration can be obtained at multiple times as shown in Table 3.26. The model selection process was conducted for a CO_2 plume observed at a single CO_2 injector and at a single time. The model selection process could be applied for multiple times and CO_2 injectors by combining multiple dissimilarities between CO_2 plumes.

[5] CO_2 saturation information can be obtained from seismic surveys. If the CO_2 saturation information is available, a dissimilarity measure that considers the shapes of CO_2 plumes and the CO_2 saturations within the plumes at the same time is necessary.

[6] The Hausdorff distance does not take account into the direction or course of curves (Alt et al., 2004; Eiter and Mannila, 1994). The two sets of points shown in Figure 8.1 are the same, but they have different paths connecting them. Because the Hausdorff distance does not distinguish the order of the point set, the two sets of points shown in Figure 8.1 are considered as the same in the Hausdorff distance. Thus the Hausdorff distance between the two sets of points shown in Figure 8.1 is 0. The Hausdorff distance is not accurate in measuring the

dissimilarity between two curves if the curves are different according to their paths as shown in Figure 8.1.



(a) 1st chain (b) 2nd chain

Figure 8.1 Two different chains of the same set of 5 points

However, the Fréchet distance considers the direction or course of curves while the Hausdorff distance does not (Alt et al., 2004; Eiter and Mannila, 1994). The Fréchet distance is intuitively defined by supposing a man walking with his dog on a leash. He and his dog are walking on two different curves with different speeds, but they are not allowed to backtrack. The Fréchet distance is the minimal length of the a leash required to connect him and his dog walking along their respective curves (Chambers et al., 2010). The Fréchet distance is a measure of dissimilarity between curves accounting for the location and ordering of the points along the curves (Eiter and Mannila, 1994). The mathematical definition of the Fréchet distance is as follows:

$$\delta_F(f,g) = min \quad [\ max_{\ t\in[0,1]} \{\ \ d(\ \ f(\alpha(t)), g(\beta(t))\ \ )\ \ \}\ ] \cdots\cdots\cdots\cdots \text{Equation (8-1)}$$

Where $f: [a, b]$ and $g: [a', b']$ are curves, $d(\cdot)$ the Euclidean distance operator, α and β are arbitrary non-decreasing functions from $[0, 1]$ onto $[a, b]$ and $[a', b']$, respectively. Eiter and Mannila (1994) presented a discrete variation of

the Fréchet distance. The discrete Fréchet distance can be used to measure the dissimilarity between skeleton paths.

However, the Hausdorff distance and the Fréchet distance had no difference in measuring the dissimilarity between skeleton paths in the examples of this dissertation. The Hausdorff distance and the Fréchet distance could give different dissimilarity between skeleton paths of CO_2 plumes migrating along fluvial sand channels with extremely high sinuosity.

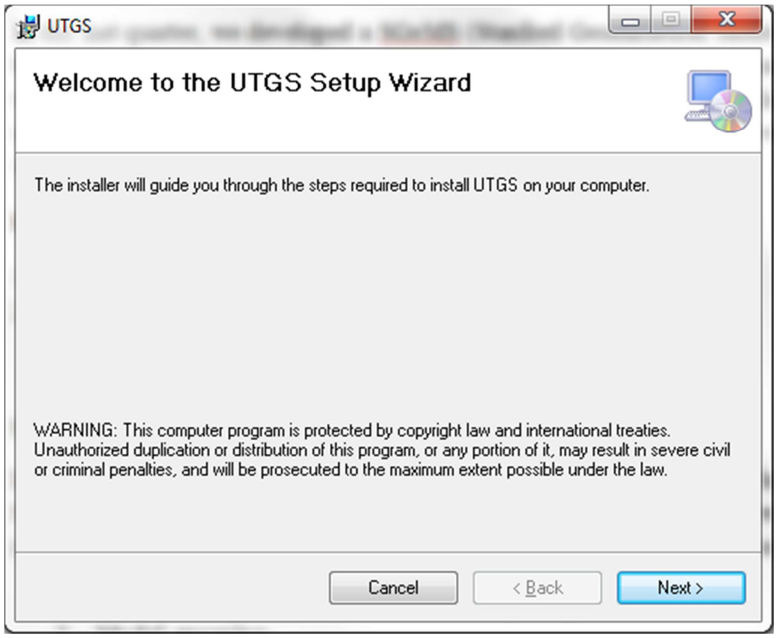
[7] Obidegwu et al. (2015) improved the history matching results by matching production data as well as the images of the gas distribution obtained from time lapse seismic data. Trani et al. (2012) considerably reduced the uncertainty associated with forecasts of production by history matching production data and water front positions obtained from seismic data. Abadpour et al. (2013) presented an approach that measures the misfit between an observed and predicted waterfronts using the Hausdorff distance. As discussed in these papers, the uncertainty in CO_2 injection performance can be improved by incorporating CO_2 injection data and the shape dissimilarity between CO_2 plumes measured using the Hausdorff distance based approaches.

Appendix A: Installation of the SCA plugin

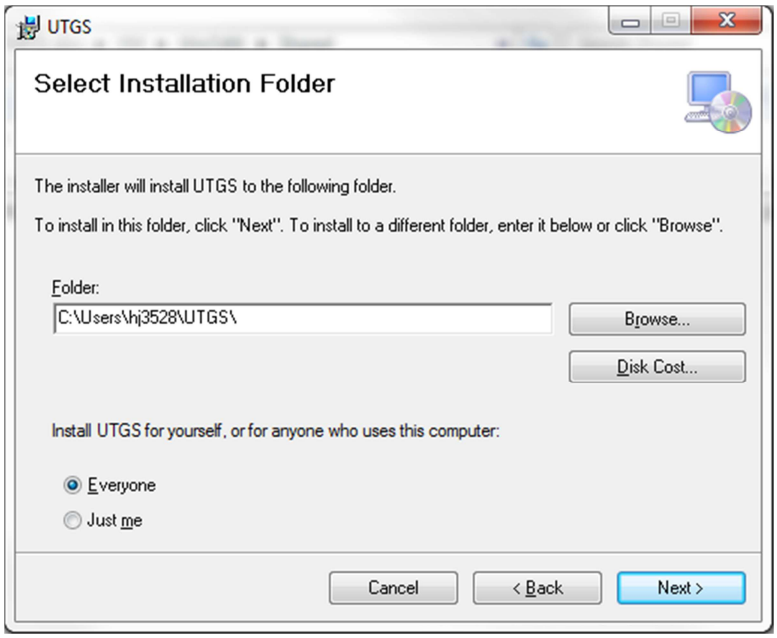
The installation of plugins proceeds by:

1) Run either setup.exe or UTGS.msi

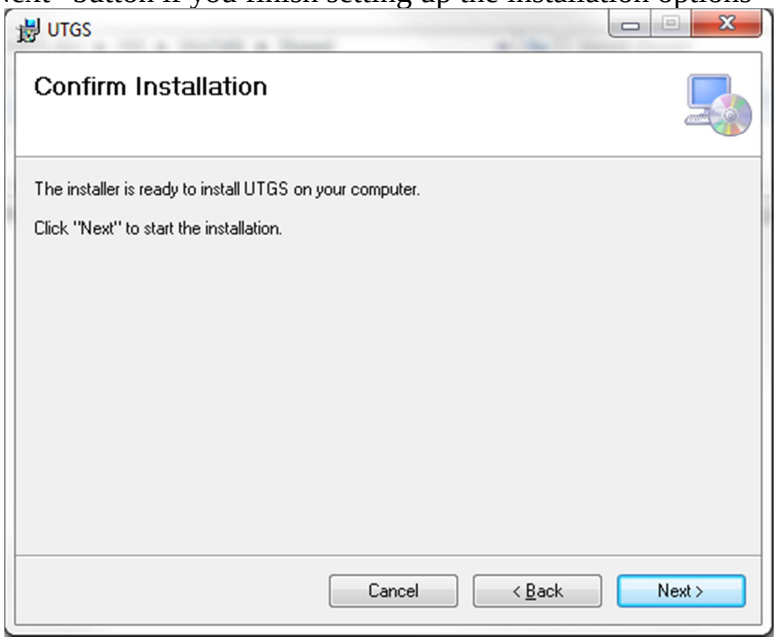
2) Click the “Next” button



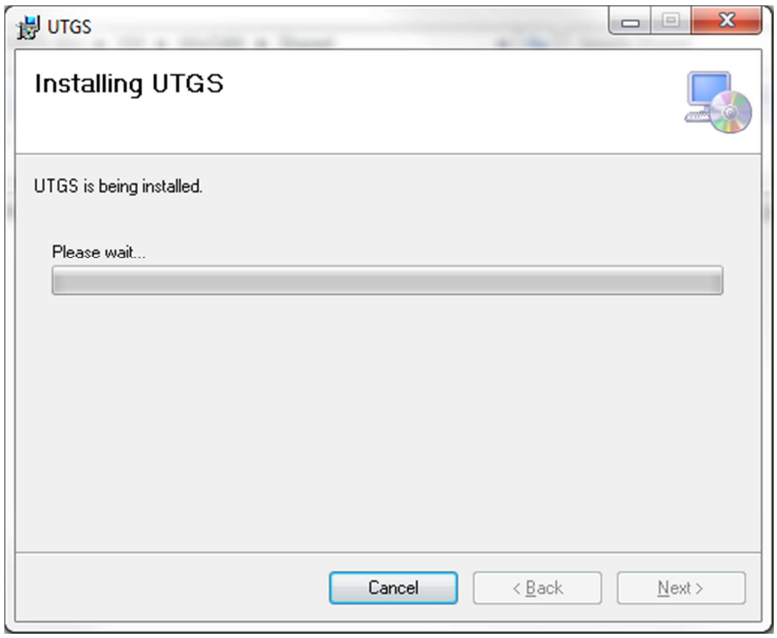
3) Select the installation folder by clicking the “Browse” button and choose if the SGeMS is available for anyone or only you on your computer. After selecting them, click the “Next” button



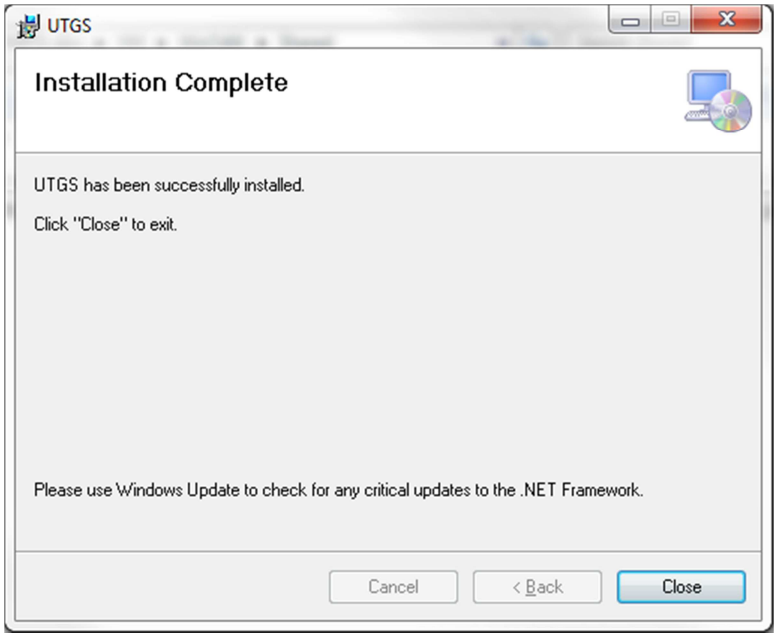
4) Click the “Next” button if you finish setting up the installation options



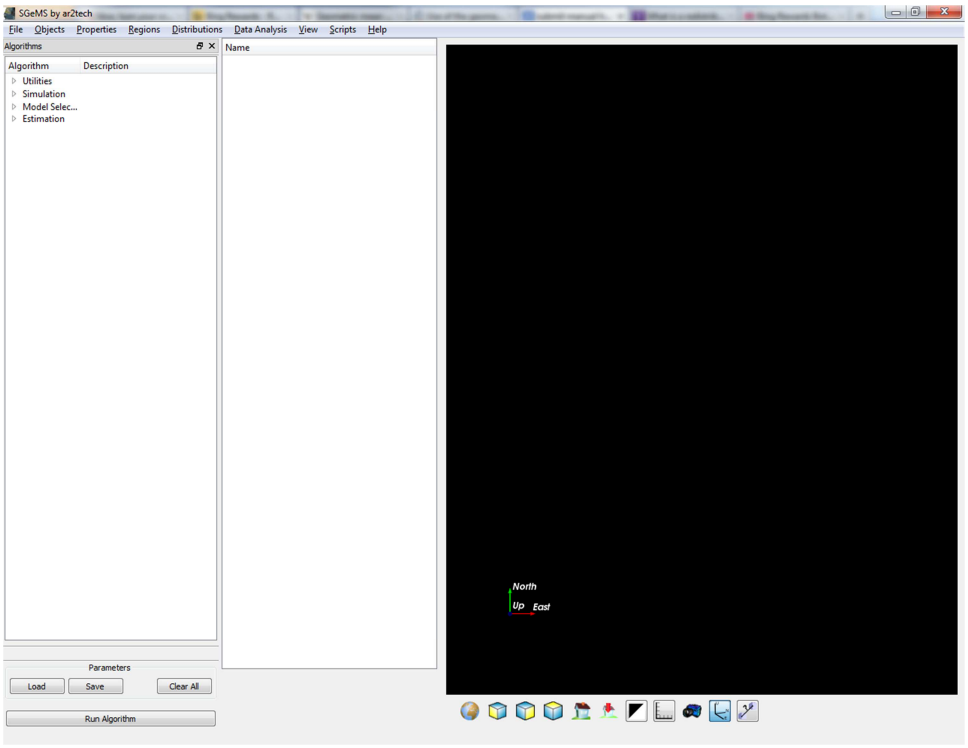
5) Contact a system manager if the screen asking an administration right might show up while installing the SGeMS



6) Click the “Close” button



7) Double-click a shortcut named UTGS for the SGeMS on the desktop or “Windows button-All programs-UTGS”. The below figures show the first screen of the SGeMS shown and the installed files and directories after installing the SGeMS.




```
├──UTGS                    : Main folder. Python scripts, DLL files for
│   │                        SGeMS modules, VTK  and QT
│   ├──encodings           : Python scripts for encodings
│   ├──examples            : Examples for the model selection algorithm
module
│   └──plugins
│       ├──colorscales     : Files defining color scales for data
visualization
│       ├──designer : DLL files for QT designer
│       ├──Geostat         : User interface files for SGeMS modules
│       ├──iconengines     : DLL files for QT icon engines
│       └──imageformats    : DLL files to read image files in QT
```

References

Abadpour, A., Bergey, P., Piasecki, R., 2013. 4D Seismic History Matching With Ensemble Kalman Filter-Assimilation on Hausdorff Distance to Saturation Front, in: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers. doi:10.2118/163635-MS

Ahmed, T., 2006. Reservoir Engineering Handbook. Elsevier Inc., pp. 354–357.

Alnes, H., Eiken, O., Nooner, S., Sasagawa, G., Stenvold, T., Zumberge, M., 2011. Results from Sleipner gravity monitoring: Updated density and temperature distribution of the CO_2 plume. Energy Procedia, 10th International Conference on Greenhouse Gas Control Technologies 4, 5504–5511. doi:10.1016/j.egypro.2011.02.536

Alt, H., Knauer, C., Wenk, C., 2004. Comparison of Distance Measures for Planar Curves. Algorithmica 38, 45–58. doi:10.1007/s00453-003-1042-5

Aspert, N., Santa-Cruz, D., Ebrahimi, T., 2002. MESH: measuring errors between surfaces using the Hausdorff distance, in: Proceedings. IEEE International Conference on Multimedia and Expo. IEEE, pp. 705–708. doi:10.1109/ICME.2002.1035879

Attali, D., Montanvert, A., 1997. Computing and Simplifying 2D and 3D Continuous Skeletons. Comput. Vis. Image Underst. 67, 261–273. doi:10.1006/cviu.1997.0536

Ballin, P.R., Journel, A.G., Aziz, K., 1992. Prediction Of Uncertainty In Reservoir Performance Forecast. J. Can. Pet. Technol. 31. doi:10.2118/92-04-05

Bandilla, K.W., Celia, M.A., Leister, E., 2014. Impact of Model Complexity on CO2 plume modeling at Sleipner. Energy Procedia 63, 3405–3415. doi:10.1016/j.egypro.2014.11.369

Bhowmik, S., Srinivasan, S., Bryant, S., 2014. Improved Particle Tracking Proxy for Assessing Plume Migration during Geologic Sequestration. Energy Procedia, 12th International Conference on Greenhouse Gas Control Technologies, GHGT-12 63, 3853–3863. doi:10.1016/j.egypro.2014.11.415

Bhowmik, S., Srinivasan, S., Bryant, S., 2010. Predicting the Migration of CO2 Plume Using Injection Data and a Distance-Metric Approach to Reservoir-Model Selection. Society of Petroleum Engineers. doi:10.2118/139709-MS

Bishop, C.M., 2007. Pattern Recognition and Machine Learning, 1st ed. 20. ed. Springer.

Boucher, A., Gupta, R., Caers, J., Satija, A., 2010. Tetris: a training image generator for SGeMS. Stanford Cent. Reserv. Forecast.

Bridge, J.S., Leeder, M.R., 1979. A simulation model of alluvial stratigraphy. Sedimentology 26, 617–644. doi:10.1111/j.1365-3091.1979.tb00935.x

Carruthers, D.J., 2003. Modeling of Secondary Petroleum Migration Using Invasion

Percolation Techniques. Multidimens. Basin Model. AAPG/Datap, 21–37.

Cavanagh, A., 2013. Benchmark Calibration and Prediction of the Sleipner CO_2 Plume from 2006 to 2012. Energy Procedia, GHGT-11 37, 3529–3545. doi:10.1016/j.egypro.2013.06.246

Chadwick, R.A., Noy, D.J., 2010. History-matching flow simulations and time-lapse seismic data from the Sleipner CO2 plume. Geol. Soc. London, Pet. Geol. Conf. Ser. 7, 1171–1182. doi:10.1144/0071171

Chambers, E.W., Colin de Verdière, É., Erickson, J., Lazard, S., Lazarus, F., Thite, S., 2010. Homotopic Fréchet distance between curves or, walking your dog in the woods in polynomial time. Comput. Geom. 43, 295–311. doi:10.1016/j.comgeo.2009.02.008

Chatbri, H., Kameyama, K., Kwan, P., 2015. A comparative study using contours and skeletons as shape representations for binary image matching. Pattern Recognit. Lett. doi:10.1016/j.patrec.2015.04.007

CMG, 2012. User's Guide GEM. Computer Modelling Group Ltd., Calgary, Canada.

Court, B., Bandilla, K.W., Celia, M.A., Janzen, A., Dobossy, M., Nordbotten, J.M., 2012. Applicability of vertical-equilibrium and sharp-interface assumptions in CO_2 sequestration modeling. Int. J. Greenh. Gas Control 10, 134–147. doi:10.1016/j.ijggc.2012.04.015

Darcis, M., Class, H., Flemisch, B., Helmig, R., 2011. Sequential Model Coupling for Feasibility Studies of CO 2 Storage in Deep Saline Aquifers. Oil Gas Sci. Technol. – Rev. d'IFP Energies Nouv. 66, 93–103. doi:10.2516/ogst/2010037

De Lima, A., Lange, A., Schiozer, D., 2012. Assisted History-Matching for the Characterization and Recovery Optimization of Fractured Reservoirs Using Connectivity Analysis. Society of Petroleum Engineers. doi:10.2118/154392-MS

Deutsch, C. V., Wang, L., 1996. Hierarchical object-based stochastic modeling of fluvial reservoirs. Math. Geol. 28, 857–880. doi:10.1007/BF02066005

Deutsch, C., Srinivasan, S., 1996. Improved Reservoir Management Through Ranking Stochastic Reservoir Models. Society of Petroleum Engineers. doi:10.2118/35411-MS

Donnez, P., 2007. Essentials of Reservoir Engineering. Editions Technip, Paris, pp. 285–296.

Dubuisson, M.-P., Jain, A.K., 1994. A modified Hausdorff distance for object matching, in: Proceedings of 12th International Conference on Pattern Recognition. IEEE Comput. Soc. Press, pp. 566–568. doi:10.1109/ICPR.1994.576361

Eigestad, G.T., Dahle, H.K., Hellevang, B., Riis, F., Johansen, W.T., Øian, E., 2009. Geological modeling and simulation of CO_2 injection in the Johansen formation. Comput. Geosci. 13, 435–450. doi:10.1007/s10596-009-9153-y

Eiter, T., Mannila, H., 1994. Computing discrete Fréchet distance.

Enayatifar, R., Salam, R.A., 2013. Similarity Measure Using Hausdorff Distance in 2D Shape Recognition System. Proc. 2nd Int. Symp. Comput. Commun. Control Autom. 194–197. doi:10.2991/3ca-13.2013.49

Fredman, M.L., Tarjan, R.E., 1987. Fibonacci heaps and their uses in improved network optimization algorithms. J. ACM 34, 596–615. doi:10.1145/28869.28874

Gang Pan, Yijun Wu, Wu, Z., Wenyao Liu, 2003. 3D face recognition by profile and surface matching, in: Proceedings of the International Joint Conference on Neural Networks, 2003. IEEE, pp. 2169–2174. doi:10.1109/IJCNN.2003.1223744

Guo, B., Bandilla, K.W., Doster, F., Keilegavlen, E., Celia, M.A., 2014. A vertically integrated model with vertical dynamics for CO_2 storage. Water Resour. Res. 50, 6269–6284. doi:10.1002/2013WR015215

Hamming, R.W., 1950. Error Detecting and Error Correcting Codes. Bell Syst. Tech. J. 29, 147–160. doi:10.1002/j.1538-7305.1950.tb00463.x

Hird, K.B., Dubrule, O., 1998. Quantification of Reservoir Connectivity for Reservoir Description Applications. SPE Reserv. Eval. Eng. 1. doi:10.2118/30571-PA

Hirsch, L.M., Schuette, J.F., 1999. Graph theory applications to continuity and ranking in geologic models. Comput. Geosci. 25, 127–139. doi:10.1016/S0098-3004(98)00116-2

Homann, H., 2007. Implementation of a 3D thinning algorithm. Insight J.

Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 15, 850–863. doi:10.1109/34.232073

Issautier, B., Viseur, S., Audigane, P., le Nindre, Y.-M., 2014. Impacts of fluvial reservoir heterogeneity on connectivity: Implications in estimating geological storage capacity for CO2. Int. J. Greenh. Gas Control 20, 333–349. doi:10.1016/j.ijggc.2013.11.009

Ivancevic, V.G., Ivancevic, T.T., 2008. Complex Nonlinearity: Chaos, Phase Transitions, Topology Change and Path Integrals. Springer Science & Business Media, pp. 368–369.

Jensen, J.L., 2000. Statistics for Petroleum Engineers and Geoscientists. Gulf Professional Publishing, pp. 136–137.

Jeong, H., Srinivasan, S., 2016. Fast assessment of CO2 plume characteristics using a connectivity based proxy. Int. J. Greenh. Gas Control 49, 387–412. doi:10.1016/j.ijggc.2016.03.001

Jeong, H., Srinivasan, S., Bryant, S., 2013. Uncertainty Quantification of CO2 Plume Migration Using Static Connectivity of Geologic Features. Energy Procedia 37, 3771–3779. doi:10.1016/j.egypro.2013.06.273

Keogh, K.J., Martinius, A.W., Osland, R., 2007. The development of fluvial stochastic modelling in the Norwegian oil industry: A historical review, subsurface implementation and future directions. Sediment. Geol. 202, 249–268. doi:10.1016/j.sedgeo.2007.05.009

Kovscek, A.R., Wang, Y., 2005. Geologic storage of carbon dioxide and enhanced oil recovery. I. Uncertainty quantification employing a streamline based proxy for reservoir flow simulation. Energy Convers. Manag. 46, 1920–1940. doi:10.1016/j.enconman.2004.09.008

Kruskal, J., 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika 29, 1–27. doi:10.1007/BF02289565

Kulkarni, K.N., Datta-Gupta, A., Vasco, D.W., 2001. A Streamline Approach for Integrating Transient Pressure Data Into High-Resolution Reservoir Models. SPE J. 6, 273–282. doi:10.2118/74135-PA

Lee, K., Jeong, H., Jung, S., Choe, J., 2013. Characterization of channelized reservoir using ensemble Kalman Filter with clustered covariance. Energy, Explor. Exploit. 31, 17–30. doi:10.1260/0144-5987.31.1.17

Lee, T.C., Kashyap, R.L., Chu, C.N., 1994. Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms. CVGIP Graph. Model. Image Process. 56, 462–478. doi:10.1006/cgip.1994.1042

Lie, K.-A., Krogstad, S., Ligaarden, I.S., Natvig, J.R., Nilsen, H.M., Skaflestad, B., 2011. Open-source MATLAB implementation of consistent discretisations on complex grids. Comput. Geosci. 16, 297–322. doi:10.1007/s10596-011-9244-4

Lipikorn, R., Shimizu, A., Kobatake, H., 2004. Three-Dimensional Object Recognition Using a Modified Exoskeleton and Extended Hausdorff Distance Matching Algorithm, in: Campilho, A., Kamel, M. (Eds.), Image Analysis and Recognition, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 697–704. doi:10.1007/b100437

Mathias, S.A., Hardisty, P.E., Trudell, M.R., Zimmerman, R.W., 2009. Approximate Solutions for Pressure Buildup During CO_2 Injection in Brine Aquifers. Transp. Porous Media 79, 265–284. doi:10.1007/s11242-008-9316-7

Nilsen, H.M., Herrera, P.A., Ashraf, M., Ligaarden, I., Iding, M., Hermanrud, C., Lie, K.-A., Nordbotten, J.M., Dahle, H.K., Keilegavlen, E., 2011. Field-case simulation of CO2 plume migration using vertical-equilibrium models. Energy Procedia 4, 3801–3808. doi:10.1016/j.egypro.2011.02.315

Nixon, M.S., Aguado, A.S., 2012. Feature Extraction & Image Processing for Computer Vision. Academic Press, pp. 181–182.

Nordbotten, J.M., Celia, M.A., Bachu, S., 2005. Injection and Storage of CO_2 in Deep Saline Aquifers: Analytical Solution for CO_2 Plume Evolution During Injection. Transp. Porous Media 58, 339–360. doi:10.1007/s11242-004-0670-9

Obidegwu, D., Chassagne, R., MacBeth, C., 2015. Seismic Assisted History Matching Using Binary Image Matching, in: EUROPEC 2015. Society of Petroleum Engineers. doi:10.2118/174310-MS

Pardo-Igúzquiza, E., Dowd, P.A., 2003. CONNEC3D: a computer program for connectivity analysis of 3D random set models. Comput. Geosci. 29, 775–785. doi:10.1016/S0098-3004(03)00028-1

Remy, N., Boucher, A., Wu, J., 2011. Applied Geostatistics with SGeMS: A User's Guide. Cambridge University Press.

Ren, B., Bryant, S.L., Lake, L.W., 2015. Fast Modeling of Local Capillary Trapping during CO2 Injection into a Saline Aquifer, in: Carbon Management Technology Conference. Carbon Management Technology Conference. doi:10.7122/439486-MS

Scheidt, C., Caers, J., 2009. Representing Spatial Uncertainty Using Distances and Kernels. Math. Geosci. 41, 397–419. doi:10.1007/s11004-008-9186-0

Senel, O., Chugunov, N., 2013. CO2 Injection in a Saline Formation: Pre-Injection Reservoir Modeling and Uncertainty Analysis for Illinois Basin – Decatur Project. Energy Procedia 37, 4598–4611. doi:10.1016/j.egypro.2013.06.368

Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts. Proc. Natl. Acad. Sci. 93, 1591–1595.

Sharifi, M., Kelkar, M., Bahar, A., Slettebo, T., others, 2014. Dynamic Ranking of Multiple Realizations By Use of the Fast-Marching Method. SPE J.

Sim, D.-G., Kwon, O.-K., Park, R.-H., 1999. Object matching algorithms using robust Hausdorff distance measures. IEEE Trans. Image Process. 8, 425–429. doi:10.1109/83.748897

Singh, V., Cavanagh, A., Hansen, H., Nazarian, B., Iding, M., Ringrose, P., 2010. Reservoir Modeling of CO_2 Plume Behavior Calibrated Against Monitoring Data From Sleipner, Norway. Society of Petroleum Engineers. doi:10.2118/134891-MS

Soille, P., 2004. Morphological Image Analysis: Principles and Applications. Springer, Berlin ; New York, pp. 208–209.

Suzuki, S., Caers, J., 2008. A Distance-based Prior Model Parameterization for Constraining Solutions of Spatial Inverse Problems. Math. Geosci. 40, 445–469. doi:10.1007/s11004-008-9154-8

TAKÁCS, B., 1998. COMPARING FACE IMAGES USING THE MODIFIED HAUSDORFF DISTANCE. Pattern Recognit. 31, 1873–1881. doi:10.1016/S0031-3203(98)00076-4

Thiele, M.R., Batycky, R.P., 2015. Discussion of Streamline Simulation: A Technology Update. J. Pet. Technol. 53, 26–27. doi:10.2118/0501-0026-JPT

Thorndike, R.L., 1953. Who belongs in the family? Psychometrika 18, 267–276. doi:10.1007/BF02289263

Trani, M., Arts, R., Leeuwenburgh, O., 2012. Seismic History Matching of Fluid Fronts Using the Ensemble Kalman Filter. SPE J. 18, 159–171. doi:10.2118/163043-PA

Vilarrasa, V., Carrera, J., Bolster, D., Dentz, M., 2013. Semianalytical Solution for CO_2 Plume Shape and Pressure Evolution During CO_2 Injection in Deep Saline Formations. Transp. Porous Media 97, 43–65. doi:10.1007/s11242-012-0109-7

Willhite, G.P., 1986. Waterflooding. Society of Petroleum Engineers.

Xie, J., Yang, C., Gupta, N., King, M., Datta-Gupta, A., 2015. Depth of Investigation and Depletion in Unconventional Reservoirs With Fast-Marching Methods. SPE J. doi:10.2118/154532-PA

Yue Lu, Chew LIm Tan, Weihua Huang, Liying Fan, 2001. An approach to word image matching based on weighted Hausdorff distance, in: Proceedings of Sixth International Conference on Document Analysis and Recognition. IEEE Comput. Soc, pp. 921–925. doi:10.1109/ICDAR.2001.953920

Zhang, D., Lu, G., 2004. Review of shape representation and description techniques. Pattern Recognit. 37, 1–19. doi:10.1016/j.patcog.2003.07.008

Zhang, L., Fonseca, M.J., Ferreira, A., 2007. Survey on 3D shape descriptors.

Zhang, W., 2013. Density-driven enhanced dissolution of injected CO2 during long-term CO2 geological storage. J. Earth Syst. Sci. 122, 1387–1397. doi:10.1007/s12040-013-0342-7

Zhao, C., Shi, W., Deng, Y., 2005. A new Hausdorff distance for image matching. Pattern Recognit. Lett. 26, 581–586. doi:10.1016/j.patrec.2004.09.022

Zhu, C., Zhang, G., Lu, P., Meng, L., Ji, X., 2015. Benchmark modeling of the Sleipner CO_2 plume: Calibration to seismic data for the uppermost layer and model sensitivity analysis. Int. J. Greenh. Gas Control In press. doi:10.1016/j.ijggc.2014.12.016