MODELING, MONITORING, AND CONTROLLING TOOLS FOR RISK MANAGEMENT OF SHALLOW AQUIFER CONTAMINATION BY FAR-FIELD BRINE LEAKAGE FROM CO₂ STORAGE FORMATIONS: INTERMEDIATE-SCALE LABORATORY TESTING

by

Ahmad Hamed Askar
A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Civil and Environmental Engineering).

Golden, Colorado
Date ____________________

Signed: ________________________
Ahmad Hamed Askar

Signed: ________________________
Dr. Tissa H. Illangasekare
Thesis Advisor

Golden, Colorado
Date ____________________

Signed: ________________________
Dr. Junko Munakata Marr
Professor and Head
Department of Civil and Environmental Engineering


ii
ABSTRACT

Carbon-dioxide geological sequestration (CGS) into deep saline formations is accepted as a promising technology to reduce CO₂ emissions in the atmosphere. Injected super-critical CO₂ pressurizes and displaces native brine to significant lateral distances within the storage formation. Far-field caprock discontinuities (e.g., faults and fractures) can act as conductive pathways posing a contamination risk of brine leakage into drinking water aquifers. In response to such potential risk and other adverse impacts by CGS operations, the EPA requires a multidimensional risk management plan to be submitted prior to CO₂ injection. The key requirements of this plan include the development of a risk assessment study, a monitoring system design, and an emergency remedial response strategy. Developing and validating tools for enhancing the management of CGS activities are essential for secure and permanent geologic storage of CO₂ and safe CGS operations.

Due to technical and cost constraints on field testing, this research incorporated intermediate-scale laboratory experimentation with numerical modeling to (1) provide insights about brine leakage simulations necessary for improving risk assessment and (2) validate developed approaches to design leakage monitoring and control systems. In this study, an ~8m long two-dimensional soil tank was constructed to simulate brine leakage from a storage zone to a shallow aquifer under a multidimensional flow field and complex geological settings mimicking field conditions. A series of experiments were conducted in this tank to generate high-resolution data on brine leakage propagation under various scenarios to support different investigations. For example, data from the first three of five experiments were collected to validate a FEFLOW-based transport model that was used as a testbed to evaluate the impacts of source condition uncertainties.
on brine leakage predictions. The fourth experiment was performed to test a developed framework for designing cost-efficient leakage monitoring systems. In this analysis, predictions made by the calibrated model using monitoring data were compared with experimental measurements to validate the framework’s applicability. The final experiment was performed to test the efficacy of using deep extraction wells to control brine leakage concentrations into a shallow aquifer. As handling extracted brine can be extremely costly, an optimization algorithm was employed to place and design the wells’ pumping rates. The brine extraction technique was also numerically tested on a hypothetical leakage problem at San Joaquin Basin, California, to assess its practicality under field-scale settings.

The results of the integrated experimental and numerical analyses suggest that the storage zone's heterogeneity and pressure field are the most influential factors on leakage pathway predictions. Optimally integrating less-expensive shallow observation wells with deep sensors can minimize the monitoring cost while allowing for early leakage detection and providing useful data to improve model predictions. The analysis showed also that brine leakage can be controlled by extracting a brine volume less than half that of the injected CO$_2$ by volume. Further, utilizing the dilution capacity of the overlaying formations can significantly reduce the required extraction rate to control brine leakage through a fracture. This research provides designers and operators with tested and validated tools to enhance the management of CGS operations, meet the EPA requirements, and improve site-selection. The findings of this study could be extended to shed light on managing waste fluid leakage from other deep injection applications such as hydraulic fracturing and deep waste disposal.
# TABLE OF CONTENTS

LIST OF FIGURES.................................................................................................................. viii

LIST OF TABLES ....................................................................................................................... xiii

ACKNOWLEDGMENTS .............................................................................................................. xiv

CHAPTER 1 INTRODUCTION ................................................................................................... 1

1.1 Research Motivation ........................................................................................................... 6

1.2 Dissertation Outline .......................................................................................................... 7

CHAPTER 2 LITERATURE REVIEW AND OBJECTIVES ....................................................... 9

2.1 CO₂ Geological Sequestration ............................................................................................ 9

2.1.1 Injection-Pressure Consequences ............................................................................... 9

2.1.2 Caprock Fracture Settings .......................................................................................... 10

2.1.3 Leakage Conditions and Controlling Factors ............................................................... 16

2.2 Pressure Build-up Management ....................................................................................... 18

2.3 CGS Leakage Contamination Risk .................................................................................... 20

2.3.1 Brine Leakage Potential and Prediction ....................................................................... 22

2.3.2 Impact of Leakage Intrusion into Shallow Aquifers .................................................... 24

2.4 Hydraulic Fracturing and Waste Disposal ......................................................................... 27

2.5 CGS Monitoring ................................................................................................................. 30

2.5.1 Source Identification .................................................................................................... 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.9</td>
<td>Water Head and Flow Rate Observation</td>
<td>81</td>
</tr>
<tr>
<td>3.1.10</td>
<td>Performed Experiments</td>
<td>82</td>
</tr>
<tr>
<td>3.2</td>
<td>Numerical Methods</td>
<td>84</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Forward Simulation</td>
<td>85</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Model Calibration</td>
<td>86</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>90</td>
</tr>
<tr>
<td>4.2</td>
<td>Experimental Approach</td>
<td>93</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Testing System Setup</td>
<td>93</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Experimental Procedures</td>
<td>94</td>
</tr>
<tr>
<td>4.3</td>
<td>Numerical Modeling</td>
<td>96</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Model Setup</td>
<td>96</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Model Calibration and Validation</td>
<td>99</td>
</tr>
<tr>
<td>4.4</td>
<td>Results and Discussion</td>
<td>102</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Experimental Results and Model Performance</td>
<td>102</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Numerical Experiments</td>
<td>106</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Relative Error Analysis</td>
<td>111</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary</td>
<td>117</td>
</tr>
</tbody>
</table>
CHAPTER 5  MONITORING BRINE LEAKAGE FROM DEEP GEOLOGIC FORMATIONS STORING CARBON DIOXIDE: DESIGN FRAMEWORK VALIDATION USING INTERMEDIATE-SCALE EXPERIMENT ........................................ 119

5.1 Introduction ................................................................................................................. 119

5.2 Global Optimal Design Framework (OD Framework) ............................................. 124
  5.2.1 Linear Predictive Uncertainty Analysis ............................................................... 124
  5.2.2 Data Worth ............................................................................................................ 125
  5.2.3 DW to Multiple Predictions .................................................................................... 127
  5.2.4 Sequential Evaluation ......................................................................................... 128
  5.2.5 Global Optimization ......................................................................................... 129

5.3 Monitoring Scheme Design ........................................................................................ 131
  5.3.1 Testing System Design ....................................................................................... 131
  5.3.2 Base-Model Rationale and Assumptions ............................................................ 133
  5.3.3 Model Setup and Parametrization ...................................................................... 136
  5.3.4 Model Predictions ............................................................................................... 139
  5.3.5 OD Framework Application ............................................................................... 140
  5.3.6 Experimental Procedures ................................................................................... 142

5.4 Results and Discussion .............................................................................................. 143
  5.4.1 Data Worth and Value Index Calculations ......................................................... 144
  5.4.2 Monitoring System Design ............................................................................... 147
  5.4.3 Leakage Source Conditions Analysis ................................................................. 150
### 5.4.4 OD Experimental Validation

5.5 Summary

#### CHAPTER 6

CONTROLLING POTENTIAL BRINE LEAKAGE FROM CO₂ STORAGE FORMATIONS USING DEEP EXTRACTION WELLS: NUMERICAL AND EXPERIMENTAL TESTING

6.1 Introduction

6.2 Optimization Method

6.3 Experimental Validation

6.3.1 Experimental Setup

6.3.2 Modeling the Simulated Leakage in the Tank

6.3.3 Optimization Problem Setup for the Experimental Problem

6.3.4 Sensitivity Analysis of the Optimization Results

6.3.5 Optimality Check for the Tank Extraction System

6.3.6 Experimental Procedures for Testing the Optimal Extraction System

6.3.7 Results of Experimental Validation

6.4 Numerical Testing

6.4.1 Modelling Simplifications and Assumptions

6.4.2 San Joaquin Basin Model Setup

6.4.3 Optimization Problem Description

6.4.4 Optimal Extraction System

6.5 Discussion
6.6 Summary ........................................................................................................................................... 190

CHAPTER 7 SUMMARY, CONCLUSIONS AND FUTURE WORK .................................................. 192

7.1 Summary of Hypotheses Testing Outcomes ................................................................. 192

7.2 Specific Conclusions from Chapter 4 (Paper I) ...................................................... 195

7.3 Specific Conclusions from Chapter 5 (Paper II) ..................................................... 197

7.4 Specific Conclusions from Chapter 6 (Paper III) .................................................. 199

7.5 Recommendations for Further Research ............................................................. 201

REFERENCES ............................................................................................................................................. 203

APPENDIX A MATERIAL REPRODUCTION PERMISSIONS .................................................. 250
LIST OF FIGURES

Figure 1.1  The Keeling Curve of CO$_2$ Concentration at Mauna Loa Observatory, Hawai (SIO, 2021).................................................................2

Figure 1.2  Schematic depicting potential deep brine leakage pathways as consequence of CO$_2$ sequestration. Numbers correspond to (1) natural fractures, (2) fault damage zone, (3) abandoned deep wells, (4) cement failure in the injection well, (5) brine diffusion, (6) brine leakage plume, (7) injection well (CO$_2$/liquid waste/HF fluids), (8) injected CO$_2$/liquid waste plume, (9) displaced brine (Nelson et al., 2005; Li & Liu, 2016; Pan et al., 2013). Dashed red line represents the domain of interest to this study........................................................................................................................................5

Figure 2.1  Different fracturing patterns based on rock matrix material (a) a limestone outcrop at the south margin of the Bristol Channel Basin, UK (Belayneh & Cosgrove, 2004), (b) sandstone exposures in the Dounreay area, Scotland (Zhang & Sanderson, 1996), and (c) fault zone structures in the Valley of Fire State Park of southern Nevada, USA (Jourde et al., 2002). The figure was copied from (Lei et al., 2017)........................................................................................................11

Figure 2.2  “Fluid pathways in fracture networks under heterogeneous stresses. (a) With the increase of the boundary stress ratio, fluid flow becomes more concentrated in only part of the fractures in the network due to the shear dilation effect (Min et al., 2004). (b) The vertical fluid flow through a jointed layer exhibits a highly localized pattern when the fractured rock is deformed under a critical stress state (Sanderson & Zhang, 1999). (c) The vertical flow in a 3D persistent fracture network shifts from a homogeneous to localized pattern with the increase of the horizontal stress ratio (Lei et al., 2015).” (Lei et al., 2017)..................................................................................................................13

Figure 2.3  Fracture dissolution simulations carried out over a broad range of Da and Pe, where bo is the initial aperture width of fracture, Pe is Peclet number (relates advection transport to diffusion transport) and Da, is Damköhler number (relates reaction rate to advection transport) (Detwiler & Morris, 2019) .................................................................................................................................14

Figure 2.4  Schematic depicting potential deep brine leakage pathways as consequence of the hydraulic fracturing process..................................................................................................................28
Figure 2.5  Ln(K) distribution of two well-characterized research sites of (left) Borden and (right) Cape-Cod (Sudicky, 1986; Sudicky et al., 2010)…………………………..45

Figure 2.6 Transverse concentration profile in the plume; Solid line represents the simulated normalized concentrations, crosses represent the discrete concentration measurements, and the arrows represent the margin of error in the predicted plume location (Schwede & Cirpka, 2009)……………………………………..58

Figure 3.1 Diagram describing (a) the developed conceptual mode and (b) the 90° transformation of the natural upwelling leakage scenario in the horizontal soil tank………………………………………………………………………………………………..65

Figure 3.2 Depiction of (a) the soil tank, (b) instrumentation layout and packing configurations. All dimensions shown in centimeters. The numbering scheme for the injection ports, instrumentation, and sampling ports increases sequentially from top to bottom and right to left. Note, flow occurs from right to left. All dimensions are in centimeters………………………………………………………69

Figure 3.3 (a, b) Target lnK distributions (dashed lines) overlaying histograms of the five sands used for packing the selected realizations for Zones 1 and 3 (lnK); (c, d) Semi-variogram plots showing the selected experimental (black solid line) and theoretical variograms (red solid line) compiled with seven other realizations that also exhibited smooth variograms (gray dashed lines). It should be noted that Real stands for realization…………………………………………71

Figure 3.4 The process of tank packing ……………………………………………………………………………………..74

Figure 3.5 Projected view of the final packed heterogeneity in the three zones of the soil tank……………………………………………………………………………………………………..74

Figure 3.6 (a) Schematic of the developed testing hypothetical problem showing the locations where the concentration profiles were drawn; (b) concentration profiles of the 10,000 ppm NaBr tracer under the vertical (solid lines) and horizontal (markers) gravity directions; (c) concentration profiles of the 50 ppm NaBr tracer (solid lines and solid markers) and 100ppm NaBr tracer (dashed lines and hollow markers); In this graph, all solid lines are for gravity in v-direction, markers are for gravity in Hz-direction and colors reflect the type of sand…………………………………………………………………………………..76

Figure 3.7 (a) Aqueous Sampling from the soil tank through the self-sealing septa ports and (b) the used IC for the chemical analysis………………………………………………………………………………..78
Figure 3.8  IC response for the 75 ppm standard solution………………………………………………79

Figure 3.9  IC calibration curve to calculate NaBr concentrations of unknown samples using area of the Br- peak……………………………………………………………………80

Figure 3.10  Example of one of the developed relationships to correct the error in concentration data due to vial evaporation………………………………………………………81

Figure 3.11  (a) Three sets of manometers connected to the pressure ports in the tank and (b) the comparison between water levels in the manometers and the water column …………………………………………………………………………………82

Figure 4.1  Projected view of NaBr distribution in the three experiments. Contour lines denote the extents of the plume boundaries at specific times. Note that the tank elbow (i.e., turn) is not included in this figure ………………………………………96

Figure 4.2  The hydraulic conductivities, boundary conditions, injection rates and tracer concentrations that were assigned in the model to generate the presented plume distributions………………………………………………………………98

Figure 4.3  (a) calculated percent absolute error (PAE) for each calibrated flow variable with the median (50th percentile) value shown and (b, c, d) observed versus calibrated concentration data for each experiment. Table includes Normalized Root Mean Square (NRMS), sample size, mean and standard deviation (STD) of the residuals………………………………………………103

Figure 4.4  Visual comparison between observed and simulated plume distributions which overlay spatial maps of the final calibrated parameters (hydraulic conductivity (K), Porosity (ϕ), and transverse dispersion coefficients (α_T)) in (a) Experiment 1, (b) Experiment 2, and (c) Experiment 3…………………………105

Figure 4.5  One example of the conducted visual comparison between the model and experimental flow fields at Zone 1: (a) the flow pathways and their trajectories (red dashed lines) of the bromide-free florescence tracer in Zone 1 during the flow field delineation experiments, (b) the uncalibrated model streamlines (yellow dashed lines) in Zone 1 and (c) the calibrated model streamlines (yellow dashed lines) in Zone 1……………………………………106
(a) The simulated steady state plume of the reference scenario (“reference plume”) including assigned BCs and (b) Zone 1 heterogeneity used to develop the new reference scenarios. Note the direction of flow is from bottom to top and all outer boundaries were assigned no-flow BCs unless otherwise stated.

The selected realizations of lnK for simulating the possible uncertainty in storage formation heterogeneity architecture: (a) realizations with $\lambda_V / \lambda_{Hz} = 35/11$, $\sigma^2_{lnK}=1.0$ for scenarios A4-A9 and A1`, A1`` (b) realizations with $\lambda_V / \lambda_{Hz} = 35/11$, $\sigma^2_{lnK}=3.5$ for scenarios A10-A15, (c) realizations with $\lambda_V / \lambda_{Hz} = 11/11$, $\sigma^2_{lnK}=1.0$ for scenarios A16-A21, and (d) realizations with $\lambda_V / \lambda_{Hz} = 80/11$, $\sigma^2_{lnK}=1.0$ for scenarios A22-A27 and B4`, B4``.

(a, b) Relative error (RE) in the location of the plume center and (c, d) plume width relative to the reference plume under the studied 5 cases. The graph includes a cartoon illustrating what is meant under each type of RE and the coordinates adopted in the estimation of the RE (Figure 4.6). Note, y-axis changes dramatically from a, b to c, d.

Schematic of the main components of global OD framework.

Schematic depicting (a) migration process of a leaking brine from a CO2 storage formation under natural field conditions, (b) experimental setup, packing heterogeneity, and instrumentation. Dominant flow direction occurs from the right to the left side of the tank.

Developed base-models including the generated hypothetical plumes overlying soil properties; model boundary and initial conditions are presented.

The overall worth of potential monitoring locations to the four combined categories of predictions calculated using the sequential selection technique. Indicative contour lines were included to delineate the highest DW zones.

The increase in uncertainty of all predictions, as a result of sequentially disregarding monitoring location data in the three base-models. Conc, Inj, and Frac respectively refer to the five concentration prediction points in the shallow aquifer, the injection flux at the three considered leakage sources, and the six fracture permeabilities.
Figure 5.6  (a) Value Index of the selected set of monitoring locations relative to the full dataset, and (b) the final implemented design in the experiment. GA and IG refer to the Genetic Algorithm and Initial Guess, respectively ..........148

Figure 5.7  Normalized Percent Error (NPE) in model predictions under the three data availability scenarios .................................................................152

Figure 6.1  Schematic of the GA adapted procedures for the optimization problem ..........163

Figure 6.2  Schematic depicting (a) brine Leakage migration from a CO\(_2\) storage formation under natural settings, (b) experimental setup including packing heterogeneity, tank instrumentation and the potential locations of extraction wells. Note the flow occurs from right to left........................................165

Figure 6.3  The three forward models, used by the GA in the optimization problem, including the boundary conditions, potential extraction wells, calibrated K-field and the location of the concentration constraining points ..................167

Figure 6.4  The concentration profiles of the plume at the two sets of the constraining points locates at X= 5.76 (Left) and X=6.83m (Right) under the three considered leakage scenarios .........................................................173

Figure 6.5  (a) Time-lapse photos of the controlled plume and (b) measured concentrations of the plume at the last day of the experiment ........................................176

Figure 6.6  (a) Simplified cross-section of San Joaquin Basin stratigraphy and (b) transient simulation of the leakage plume propagation after fault activation by the equivalent injection of 5 Mt CO\(_2\)/year for 50 years ......................181

Figure 6.7  (a) The sensitivity of the extraction ratios to the change in the target concentrations. (b-e) the concentration profiles of the plume along the interface between Kern River and Etchegoin formations (marked by a red line in Figure 6.6a) for different target concentration scenario under the leakage mechanisms of (b, c) continuous or (d, e) buried pathway ..........188
LIST OF TABLES

Table 1.1  CO₂ Storage projects across USA using saline aquifers, extracted from Ajayi et al., (2019) .................................................................3

Table 3.1  Physical Properties of sands used in soil tank. Data extracted from Agartan (2015), Barth et al., (2001), Peng & Brusseau (2005) and Sakaki & Illangasekare (2007) .................................................................72

Table 3.2  Retention times, peak heights, areas under the peak curves for the used Br standard solutions .................................................................79

Table 3.3  Summary of the Performed Experiments .................................................................83

Table 4.1  Summary of Experimental boundary condition and Collected Data statistics ......95

Table 4.2  Summary of the assigned initial values and bounds for parameters subjected to optimization process in PEST for each experiment .................................................................99

Table 4.3  Summary of the 46 Tested Uncertainty Scenarios .................................................110

Table 5.1  Summary of parameter zones and assigned bounds .............................................138

Table 5.2  Summary of boundary conditions and collected data ..........................................143

Table 6.1  Tank Experiment Optimization Results .................................................................170

Table 6.2  Numerical Experiment Optimization Results ........................................................186
I would like to express my gratitude to my advisor, Dr. Tissa Illangasekare for his guidance and trust. I am also thankful to my committee members Dr. Paul Sava, Dr. Kamini Singha, and Dr. Ahmadreza Hedayat for their insightful comments. Special thanks are extended to Dr. Andrew Trautz and Dr. Ye Zhang for their contributions and technical assistance in developing the foundation of this study and designing the experimental system. I would like also to acknowledge the incredible help from Sarah Awara in proofreading my scientific papers and dissertation. I am indebted to Dr. David Lipson and Dr. John Spear for their limitless support and willingness to listen, advice, contribute, and help during my PhD studies.

A long list of people have contributed to make my PhD journey at Colorado School of Mines less stressful. I have to thank John Stults, Allan Foster, and William Taylor for being real good friends. My lab fellows including post-docs, collaborators, grad- and undergrad-students helped to make the experimental work easier by assisting me in building and packing the tank and sometimes collecting and analyzing data. For such tasks, many thanks goes to Ana Ilie, Jakub Solovský, Jordan Skipwith, Forrest Pilone, Joon Moon, William Konishi, Anouk Uragoda, Cody Goebe and Gianna Pittman. In addition, Stefanie Shea, Juliane Brown, Andrew Maizel, Youn, Choi, and Anastasia Nickerson, provided fruitful discussions and feedbacks. Analise Butler deserves a special mention because of her unconditional welling to help in the lab work.

Throughout my research journey at CESEP, I have received a great deal of support from my lovely wife Shimaa, who was always patient and wise with all of the financial, social and health hardships we faced during my studies, my beautiful daughter Mariah, who brought so much joy and happiness to my life during the last two years of this research, my sister Manal, who kept me
focused and determined through our constructive discussions and my friends Abdo and Omar, who never hesitated to rescue me every time I needed help in the lab or in my daily life. I should also acknowledge the emotional containment and academic guidance that were offered by my parents, Dr. Hamed Askar and Dr. Magda Shehata. Moreover, big “thank you” goes to my other amazing sisters, Rehab and Aya Askar, for their continuous encouragement and support. All my family members and friends overseas deserve a recognition for their extreme kindness during the time of my research abroad. Last but not least, I would like to thank all my brothers in law, Ahmed Ali Sedeek, Moahmed Al-Azali, Jason Piwarski, and Taha Hammad, my cousin Hany Shehata and my dear friend Ahmad Farahat, for taking care of my family, which relieved me a lot.
CHAPTER 1
INTRODUCTION

Since the industrial revolution in 1760, the energy demand around the world has increased significantly causing a large spike in carbon dioxide (CO\textsubscript{2}) concentrations in the atmosphere due to the intensive combustion process of fossil fuels. The average annual global rise of CO\textsubscript{2} concentrations in the atmosphere has reached 2.46 ppm in 2019 (Olivier et al., 2017). CO\textsubscript{2} concentrations continued to increase through 2021 crossing a new climate-change threshold of +419 ppm of CO\textsubscript{2} on the air (Figure 1.1) (SIO, 2021). Recently, Friedlingstein et al., (2019) estimated that 440 ± 20 gigatons of carbon (10\textsuperscript{15} grams) were emitted to the atmosphere as CO\textsubscript{2} from 1850 to 2018 due to fossil fuel burning, which concerned the environmental authorities and scientific communities around the world.

The Intergovernmental Panel on Climate Change (IPCC), in compliance with the call and efforts of the United Nations Framework Convention on Climate Change (UNFCC), has proposed several methods for mitigating CO\textsubscript{2} loads in the atmosphere: (1) regulating the energy use all over the world, (2) finding alternatives to fossil fuels (e.g., renewable resources) and, (3) sequestering captured CO\textsubscript{2} into deep geologic formations (IPCC, 2005; Lackner, 2003). Carbon geologic sequestration (CGS) was the most promising technique among the proposed options for two reasons: (1) the short implementation time and (2) the sustainable use of fossil fuels that currently supply 84% of the world’s energy (Looney, 2020). The International Energy Agency (IEA) highlighted that at least 14% of the planned reduction of CO\textsubscript{2} emissions should be handled by CGS
technology to halve the global CO₂ emissions by 2050, as agreed in Paris during the UNFCCC meeting in 2015 (IEA, 2017).

In 1972, CGS technology was employed to enhance oil recovery from low oil production wells (EoR) (King et al., 2013), but later the technology was applied to geologically sequestrate CO₂ at a commercial scale. The Sleipner project in Norway was the first launched large-scale CGS facility for commercial purposes. The plant marked 20 years of operation in 2016 with a total injected CO₂ of 17 Mt into a deep offshore aquifer (Global CCS Institute, 2017). In-Salah facility, located in Algeria, is an onshore CGS facility with an approximate injection rate of 1 Mt of CO₂ per year (Ringrose et al., 2013). Lately, Abu Dhabi established its first CGS facility in 2016 for serving the booming iron and steel industry in the country by capturing and storing 0.8 Mt of CO₂ per year (Ustadi et al., 2017).

![Figure 1.1 The Keeling Curve of CO₂ Concentration at Mauna Loa Observatory, Hawai (SIO, 2021)](image)

Three types of geological storages can be potentially used for CGS applications, such as unmineable coal seams, deep saline formations, and depleted oil and gas reservoirs (Bachu, 2003;
IPCC, 2005). Several studies reported the efficiency of using deep saline formations as repositories of CO₂ because of their large storage capacity (i.e., $21,633 \times 10^3$ Mt of CO₂) and redundancy across the world and the U.S.A (IPCC, 2005; Lackner, 2003; Bachu, 2003; U.S. DOE, 2015). However, the sparse information and limited resolution data on the deep saline formations necessitates intensive caprock-integrity investigations to ensure the safe operation of CGS facilities, which can be financially challenging (Bentham & Kirby, 2005). From the first commercial implementation of CGS technology in 1996 until 2008, deep saline formations were able to host approximately 20 Mt of CO₂ alone (Michael et al., 2010). Table 1-1 presents the planned and currently operated CGS projects in the USA using deep saline aquifers.

Table 1.1 CO₂ Storage projects across USA using saline aquifers, extracted from Ajayi et al., (2019).

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Storage</th>
<th>CO₂ fate</th>
<th>Status</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cranfield</td>
<td>1–1.5 Mt/y</td>
<td>Saline reservoir, Tuscaloosa Sandstone Formation, down dip of the mature Cranfield Oil Field</td>
<td>In operation since 2009</td>
<td>(Lu et al., 2012)</td>
</tr>
<tr>
<td>Citronelle</td>
<td>0.25 Mt/y</td>
<td>The southern flank of the Citronelle dome</td>
<td>In operation since 2011</td>
<td>(Haghighat et al., 2013)</td>
</tr>
<tr>
<td>Decatur</td>
<td>1 Mt/y</td>
<td>Mount Simon sandstone</td>
<td>In operation since 2011</td>
<td>(Zhou et al., 2010)</td>
</tr>
<tr>
<td>Frio</td>
<td>177 t/d</td>
<td>Frio Formation</td>
<td>In operation since 2004</td>
<td>(Hovorka et al., 2006)</td>
</tr>
<tr>
<td>Teapot Dome</td>
<td>170 t/d</td>
<td>Tensleep and Red Peak Formations</td>
<td>In operation since 2006</td>
<td>(S.J. &amp; V.W., 2006)</td>
</tr>
<tr>
<td>Kevin Dome</td>
<td>0.125 Mt/y</td>
<td>The Duperow Formation (3900ft)</td>
<td>Planning</td>
<td>(Riding &amp; Rochelle, 2005)</td>
</tr>
<tr>
<td>Wasatch Plateau</td>
<td>1 Mt/y</td>
<td>The Jurassic Entrada Formation and Navajo sandstone</td>
<td>Planning</td>
<td>(Parry et al., 2007)</td>
</tr>
</tbody>
</table>

In contrast to deep saline formations, available data on depleted oil and gas reservoirs and explored coal seams minimizes the required site characterization effort in the project and reduces
the likelihood of undetected natural leakage pathways to exist in the system (e.g., faults and fractures) (IPCC, 2005). However, the poorly sealed boreholes and improperly cemented abandoned wells, drilled during the exploration phase, can subject the underground sources of drinking water (USDWs) to contamination risk by CO$_2$ leakage or the upwelling formation fluids (Hannis et al., 2017). Therefore, using deep saline formations to store CO$_2$ was found to be less risky and more efficient than the other alternatives.

In CGS operations, supercritical CO$_2$ (ScCO$_2$) is injected into a deep saline aquifer under very high pressure to overcome the formation entry pressure. Naturally existing or pressure-induced discontinuities in the caprock (e.g., activated faults, new fractures, etc.), have the potential to act as leakage pathways (Figure 1.2) extending from the storage zone to the overburden layers including the USDWs (Wang et al., 2017; Lee et al., 2012). CO$_2$ injection pressure, transmitted to the storage formation fluids, propagates faster and farther than the CO$_2$ plume (Celia et al., 2011; Birkholzer et al., 2011a), which makes susceptible areas for brine leakage larger than of those suspected to host CO$_2$ leakage events (Wunsch et al., 2013).

Birkholzer et al., (2009) demonstrated that the injection pressure response can travel laterally several kilometers away from the injection well (Cavanagh & Wildgust, 2011; Birkholzer et al., 2011a), which pressurizes and displaces native brine to remote areas posing the risk of far-field brine leakage (Figure 1.2). It is worth mentioning that there is no existing solid evidence of brine leakage from the currently operated CO$_2$ injection facilities yet. However, brine leakage events are already reported from CGS analog systems, such as produced-water disposal operations (Thamke & Craig, 1997; Thamke & Midtlyng, 2003; Jacobs, 2009), natural CO$_2$ gas reservoirs (Keating et al., 2010; 2013; 2013; 2014) and leaky deep saline aquifers (Llewellyn, 2014).
Figure 1.2 Schematic depicting potential deep brine leakage pathways as consequence of CO$_2$ sequestration. Numbers correspond to (1) natural fractures, (2) fault damage zone, (3) abandoned deep wells, 4) cement failure in the injection well, (5) brine diffusion, (6) brine leakage plume, (7) injection well (CO$_2$/liquid waste/HF fluids), (8) injected CO$_2$/liquid waste plume, (9) displaced brine (Nelson et al., 2005; Li & Liu, 2016; Pan et al., 2013). Dashed red line represents the domain of interest to this study.

The carbonic acid produced in the storage formation due to the reaction between the injected CO$_2$ and native brine can deteriorate the integrity and sealing capacity of the caprock by the rock mineral dissolution process. Such chemical impact on the caprock enhances the geomechanical stresses imposed by CO$_2$ injection by facilitating the induction of new leaky
pathways in the system or widening the existing minor cracks in the caprock (Rathnaweera et al., 2016; Andreani et al., 2008; Ellis et al., 2011). Although the quality of native brine is a site-specific property, the total dissolved solids (TDS) of deep brine water typically exceed 10,000 ppm, while in some basins can even reach 100,000 ppm (Wunsch et al., 2013). Moreover, high levels of chloride, iron, manganese, nitrates can be found in the deep brine with toxic trace metals such as arsenic, lead, cadmium, and poisonous levels of boron (Wunsch et al., 2013; Qafoku et al., 2017). Hazardous contaminants can exist naturally in the native brine or introduced via the dissolution process of rock minerals in the storage formation or along the migration pathway of brine leakage (Qafoku et al., 2017). Several studies demonstrated the possibility of contaminating the USDWs by brine leakage using high-fidelity multiphase reactive-transport simulations (Carroll et al., 2014b) and geochemical mixing models (Wunsch et al., 2013). Contamination of the USDWs by brine leakage can occur through two mechanisms; (1) poor-quality native brine leaks upward through conductive pathways and mixes with potable groundwater or (2) leaking storage-formation fluids trigger the migration of the overlying saline water affecting the shallow aquifer quality (Tillner et al., 2013; 2016; Delfs et al., 2016; Kissinger et al., 2017).

1.1 Research Motivation

In response to the potential contamination risk from CGS operations on USDWs, the EPA established the underground injection control (UIC) program in 1980 to provide a regulatory framework for shallow and deep injection activities (U.S. EPA, 2018). Under this program, a set of federal requirements regarding the planning, construction, and operation of CGS projects must be fulfilled to ensure both efficient long-term trapping of the injected CO₂ and management of the potential leakages from CO₂ storage formations. These requirements include developing,
implementing, and periodically reviewing the risk assessment analysis, monitoring system design, and emergency-remedial response strategy of CGS facilities (U.S. EPA, 2010).

Exploring the most influential leakage source parameters on the model predictions of the plume pathway and distribution can help in better planning site characterization. Such source parameter sensitivity analysis will also improve the accuracy of leakage simulations for risk assessment and thus the identification of the likely areas/activities to be impacted in the shallow aquifer. Additionally, developing and validating tools for leakage monitoring and control systems’ designs will contribute to safer operation of CGS projects. Addressing such knowledge gaps in modeling, monitoring, and controlling brine leakage within a research framework will help to enhance the overall management of CGS applications, which was the main motive for this study. It is worth mentioning that the problems addressed in this study are of strong relevance to other injection activities under the UIC program, such as deep waste disposal and high-volume hydraulic fracturing. These operations involve the injection of pressurized non-potable fluids or chemically amended water into deep geologic formations as well.

1.2 Dissertation Outline

The present chapter included an overview on the basic problems facing brine leakage management that were addressed in this research. Chapter 2 describes more details about the technical and scientific background of brine leakage causes, impacts and the proposed monitoring and remediation design tools in the literature. Chapter 3 presents the experimental procedures, laboratory facilities, instruments, and numerical methods used in this study. Chapters 4, 5 and 6, reproduce three journal articles addressing the problems described above based on the research conducted under this dissertation. The article in Chapter 4 is already published in the Water Resources Research journal and the second article (Chapter 5) is submitted to the same journal but
is still under review. While the third article (Chapter 6) is ready to be submitted to a peer-reviewed journal. Finally, a summary of the research findings, conclusions, and suggested future research is described in Chapter 7.

The articles’ respective titles and authors are as follows:


2.1 **CO₂ Geological Sequestration**

CO₂ geological sequestration into abandoned deep saline reservoirs have received a broad attention from the scientific community as it can adversely impact the natural water resources and the intact conditions of deep aquifers. In this section, the potential consequences of injection pressure, caprock fracturing causes, and leakage controlling factors are reviewed.

2.1.1 **Injection-Pressure Consequences**

Injection induced pressure associated with CO₂ sequestration in deep saline formation can displace native brine to significant lateral distances from the injection point (Zhou et al., 2010), which enlarges the injection footprint and increases the potential leakage risk from remote areas in the storage reservoir. According to Birkholzer et al., (2009), dissolved particles in native brine can travel hundreds of meters due to injection pressure. Additionally, the boundaries of the pressure bulb inside a CO₂ storage formation can reach 100 Km after 100 years of operation. Based on a numerical study conducted by Zhou et al. (2010) on Mt. Simon aquifer at Illinois Basin, a pressure rise of 1.0 and 0.1 bar can occur at distances of 150 and 300 Km from the injection site, respectively. In the same study, after 50 years of injection, the predicted build-up pressure near the injection location ranged from 25 to 35 bars, which can affect the caprock integrity (Pan et al., 2013; Zhou et al., 2010). In another stochastic study, it was found that the injection pressure footprint covered about 98% of a 100 km²-cross-sectional domain, while the CO₂ plume covered only 3.5% of the same domain after 50 years of brine leakage (Nogues et al., 2011). Such rapid
and extensive propagation of the injection pressure throughout CO$_2$ storage formations indicates the possibility of brine leakage to occur at nearby or remote areas from the injection site before CO$_2$ leakage. By disturbing the in-situ stresses in the caprock, the elevated pressure in the storage zone can activate existing faults, closed fractures, or induce new leaky pathways (Lee et al., 2012). Storage formations bounded by a damaged caprock are highly subjected to brine leakage risk (Nordbotten et al., 2004).

2.1.2 Caprock Fracture Settings

Fractures and damage zones present in the rock matrices (e.g., caprock) as discrete discontinued features due to the complex self-organized mechanical dynamics deriving them. This makes the rock fragmentation and cracks to occur in different scales and directions in the matrix based on the physical properties of the host rock (Lei & Wang, 2016) as shown in Figure 2.1. Wang et al., (2017) explored the geomechanical factors controlling the propagation of newly induced fractures in the caprock by CO$_2$ injection. They concluded that the heterogeneity of the tri-axial stresses in the caprock, developed by injection process, governs the direction of the new fractures. The evolution of a fracture network or a damage zone in the caprock is controlled by the settings of the formation natural and/or induced stresses (Schultz, 2000; Pollard & Aydin, 1988). Generally, heterogeneous stress fields result in local nonuniform normal and shear stresses in the rock matrix (e.g., caprock). Such differential stresses lead to various complex geomechanical responses from the existing cracks in the rock, such as sliding, closing, opening, branching, abutting, cross-cutting and propagating mechanisms. These responses directly affect the permeability, dominant fluid pathways, and mass transport of the fracture network (Lei et al., 2017; Detwiler & Morris, 2019) as shown in Figure 2.2. Zhang et al., (1996) reported that the dominant fluid pathways of a fracture network can be altered significantly by the stress configuration in the
formation. In addition, Zhang & Sanderson (1996) highlighted that the randomly oriented small fractures in a non-symmetrical fracture network can increase the connectivity of the damage zone. Min et al., (2004), Baghbanan & Jing (2008) and Figueiredo et al., (2015) showed that under high differential stresses, in the rock matrix (e.g., caprock), local small fractures can be dilated forming large connected flow channels, as shown in Figure 2.2a.

![Figure 2.1](image-url)

Figure 2.1 Different fracturing patterns based on rock matrix material (a) a limestone outcrop at the south margin of the Bristol Channel Basin, UK (Belayneh & Cosgrove, 2004), (b) sandstone exposures in the Dounreay area, Scotland (Zhang & Sanderson, 1996), and (c) fault zone structures in the Valley of Fire State Park of southern Nevada, USA (Jourde et al., 2002). The figure was copied from (Lei et al., 2017).

Sanderson & Zhang (1999; 2004) found that the vertical flow through a fracture network can be intercepted by encountering large apertures, high pore pressure or heterogeneous stress field (Figures 2.2b and 2.2c). Under differential formation stresses, the permeability of a damage zone can be initially reduced due to fracture closure (Zhang & Sanderson, 1996; Min et al., 2004). However, such permeability rebounds quickly as a result of the fracture shear dilation, crack propagation and brittle failure in the rock matrix (e.g., caprock) (Baghbanan & Jing, 2008; Latham et al., 2013; Paluszny & Matthäi, 2009; Renshaw, 1996). In addition, the mass transport dynamics
through a fracture network can be indirectly impacted by the geomechanical stresses via altering the parameters controlling the advection and dispersion properties (Lei et al., 2015). Zhao et al., (2010) found that the compressive stresses in a fractured rock (e.g., caprock) can attenuate dispersion coefficients due to fracture closure. While the solute spreading will be enhanced eventually by shear dilation effects. Zhao et al., (2011) noted that with relatively low differential stress ratio, the average residence time of a solute in a fracture network can be increased due to permeability reduction. However, when this ratio increases over a specific threshold, the solute residence time can be dropped significantly due to the newly connected channels in the system. Above discussion shows how dynamic are the transport properties of a critically stressed fractured system, which poses a great challenge to characterize leakage source conditions in a CGS site.

Characterizing the permeability of a fractured system is further complicated when the mixture of native formation fluids (i.e., brine) and injected ScCO₂ is not at equilibrium state with the host rock minerals (i.e., caprock). This non-equilibrium state leads to a chain of geochemical reactions between the CO₂-laden brine and rock minerals (i.e., dissolution or precipitation process), which can vary the fracture hydraulic settings. Rathnaweera et al., (2016) conducted a long-term experiment (1.5 years) to examine the induced mineralogical changes by CO₂-laden brine in sandstone rock samples. They observed a significant dissolution of rock minerals including quartz, Fe²⁺, Ca²⁺, Al³⁺, and barite minerals, which considerably enhanced rock permeability. Ross et al., (1981) reported an increased permeability of a flooded calcite- and dolomite-cemented sandstone by CO₂. Moreover, Andreani et al., (2008) found a linear relationship between the increase of fractures’ apertures in claystone (i.e., CO₂ confining layer) and their exposure time to CO₂-brine flow. In case of carbonate fractured caprock, CO₂-acidified
brine can rapidly dissolve and deteriorate the naturally existing small fractures/damages causing enlargement in their width by a factor of \(~ 2.7\) (Ellis et al., 2011).

![Diagram of fluid pathways in fracture networks under heterogeneous stresses.](image)

Figure 2.2 “Fluid pathways in fracture networks under heterogeneous stresses. (a) With the increase of the boundary stress ratio, fluid flow becomes more concentrated in only part of the fractures in the network due to the shear dilation effect (Min et al., 2004). (b) The vertical fluid flow through a jointed layer exhibits a highly localized pattern when the fractured rock is deformed under a critical stress state (Sanderson & Zhang, 1999). (c) The vertical flow in a 3D persistent fracture network shifts from a homogeneous to localized pattern with the increase of the horizontal stress ratio (Lei et al., 2015).” (Lei et al., 2017).

The reactive-transport simulations conducted by Groves & Howard (1994) and Dreybrodt (1996) showed that the breakthrough time needed for a small fracture to grow significantly and start transferring large volumes of fluids depends on the kinetics of rock dissolution-reactions, which is closely related to the chemical composition of the leaky fluid and rock minerals. Moreover, the instable fronts of the dissolution-reactions play an essential role in reducing such breakthrough time and accelerating the dissolution process (Hanna & Rajaram, 1998). This finding
was later confirmed by a set of dissolution experiments conducted by Durham et al., (2001) and Detwiler et al. (2003). The dissolution rate of the internal surfaces of a fracture is controlled by both fluid transport properties (i.e., advection versus dissolution ($P_e$)) and reaction constant rate (Detwiler & Morris, 2019) as illustrated in Figure 2.3.

Figure 2.3 Fracture dissolution simulations carried out over a broad range of Da and Pe, where $b_0$ is the initial aperture width of fracture, Pe is Peclet number (relates advection transport to diffusion transport) and Da, is Damköhler number (relates reaction rate to advection transport) (Detwiler & Morris, 2019).

On the other hand, several studies documented a noticeable reduction in the fracture permeability due to mineral precipitation that can be caused by the ionic solubility decrease of
rock minerals (Shiraki & Dunn, 2000; Muller et al., 2009; Zhang et al., 2010; Chaudhuri et al., 2008). Ionic solubility can be dropped due to (1) mixing the leaking fluid with other over-saturated solutions, (2) temperature changes, and (3) heterogeneity of the rock matrix mineralogical structure (Chaudhuri et al., 2008; Tartakovskv et al., 2008; Li et al., 2007; Jones & Detwiler, 2016). The geochemical reactions can also be reversed due to the effects of geomechanical stresses, system mineralogical structure, and fracture rock minerals (Detwiler & Morris, 2019), which can alter the dominant process in the storage formation from mineral dissolution to precipitation, or vice versa.

Most of the developed studies exploring the impact of CO₂/brine leakage on overlying aquifer and formations, represented fractured caprock by a single leaky pathway (Vermeul et al., 2016; Liu et al., 2017; Delfs et al., 2016). This can be attributed to the large scale of modeled area compared to fracture sizes, coarse discretization of the problem domain, and the assumption that small fractures can hardly contribute to leakage flow. Although this simplification reduces the computational burden of the model, it ignores and obfuscates the micro-contribution of the small-discrete fractures to the local flow field at the source vicinity. The role of small fractures in forming connected flow channels was highlighted multiple of times in the literature (Zhang & Sanderson, 1996; Berkowitz et al., 2000; Darcel et al., 2003; Min et al., 2004; Baghbanan & Jing, 2008; Figueiredo et al., 2015). Using an equivalent single leaky pathway may be insignificant to the overall water-budget balance but could greatly affect the predicted leakage transport behavior (Vialle et al., 2016; Lei et al., 2017). Pan et al., (2013) emphasized that small cracks and minor fissures in the caprock can facilitate diffusing the CO₂ injection pressure throughout the caprock creating new leaky pathways. Poorly designed injection pressure of CO₂ can lead to fault activation/permeabilization, fracturing and induced seismicity in the CO₂ bearing formation or in
the overlying zones (Rutqvist et al., 2016). Based on the reviewed literature in this section, the prediction and characterization of the hydraulic and structural settings of caprock discontinuities can be very challenging, which rises concerns about the reliability of the risk assessment studies submitted to EPA for permission purposes.

2.1.3 Leakage Conditions and Controlling Factors

CO₂/brine leakage may occur early during the injection phase as a result of poorly designed ScCO₂ injection rates or arise after years from ceasing the injection process due to the physiochemical interactions between the injected ScCO₂ and the confining caprock layer (Pan et al., 2013; Wang et al., 2017; Ellis et al., 2011). Different studies explored the potential causes, transport dynamics, and controlling parameters (i.e., fracture/fault permeability, aperture, brine density, and etc.) of the CO₂/brine leakage in attempts to understand and predict the leakage fate in the overburden formations. For example, Birkholzer et al., (2011a) studied the upward brine leakage through vertical abandoned wells from a saline sedimentary basin after CO₂ injection. They highlighted that the formation brine can only leak through abandoned wells when the induced pressure in the storage formation is high enough to overcome the fluid column weight in these wells. Moreover, they emphasized that the leakage rate is controlled by the wellbore effective permeability and the magnitude of injection pressure. Recently, Wang et al., (2018) found that the brine leakage rate is directly proportional to the excess in pore pressure (i.e., caused by injected CO₂) and inversely to the degree of brine salinity. The numerical studies conducted by Nicot (2008), Birkholzer & Zhou (2009), Yamamoto et al., (2009), Person et al., (2010) and Lions et al., (2014) showed that brine leakage rate is mainly controlled by both pressure increase and extent in the storage formation.
Keating et al., (2013) studied the CO₂/brine leakage through natural faults using 2D and 3D numerical models that were calibrated on collected observations from two CGS natural analog sites [Chimayo, New Mexico site and Springerville, Arizona site]. They found that narrow faults and damage zones can promote faster upward brine migration. While Liu & Zhang (2011) and Lu et al., (2012) reported the high sensitivity of CO₂/brine leakage rate and volume to the permeability of the leakage pathways. In addition, Bricker (2012), Zhao et al., (2012), Wainwright (2013) and Xie et al., (2015a; 2015b) highlighted the important role of the permeability field of the storage zone and fractured caprock in controlling the rate of brine leakage. Jeanne et al. (2013) and Vialle et al., (2016) showed how the permeability distribution of caprock fractures can impact the prediction of CO₂/brine leakage pathways. While Hyman et al., (2020) demonstrated a high sensitivity of the scCO₂ leakage to the fractured caprock permeability, connectivity, and structure (i.e., fracture density).

Slow diffusive leakage of formation brine can also occur from the CO₂ storage zone to the overlying formations across the caprock-mass under the induced injection pressure. Dejam & Hassanzadeh (2018) found that this diffusive leakage rate is governed by the confining layer thickness, storage unit depth and the contrast between the permeabilities and porosities of the storage zone and caprock. Delfs et al., (2016) conducted an interesting study to investigate the brine migration through a leakage pathway in a layered system under the effects of density and temperature variations. They concluded that leaking brine will not likely travel the whole distance to the shallow aquifer, but it would rather consequentially displace the overlying formations’ fluids until the USDWs are intruded.

Brine leakage development and migration are site-specific problems that vary with the hydraulic and geological conditions of every CO₂ storage site. Although, the expected high
variability in the CO₂/brine leakage dynamics from one site to another, Nelson et al., (2005) emphasized that leakage through existing or induced caprock fractures can be the most probable leakage mechanism in CGS operations. Therefore, leakage through caprock fractures and damage zones was given a particular focus in this study.

2.2 Pressure Build-up Management

Due to the potential geomechanical risks on the caprock that can be posed by the pressure increase in the storage formation, different pressure management strategies were proposed in the literature. Buscheck et al. (2011) introduced the concept of “Active CO₂ Reservoir Management (ACRM)”, where the native brine is extracted from the storage zone, desalinated, and then reinjected back with the CO₂. They showed that the ACRM technique decreases the pressure build-up, alleviates the leakage risk, enlarges the CO₂ storage capacity, and enhances the CO₂ dissolution entrapment in the host formation by mixing it with fresh brine. The ACRM also reduces the danger of caprock degradation, fault activation, induced seismicity and shrinks the project area of review (AoR) (Buscheck et al., 2011). Notably, the AoR is defined by “the region surrounding the geologic sequestration project where USDWs may be endangered by the injection activity.” (U.S. EPA, 2010). Moreover, the ACEM technique can be employed to decouple neighboring CGS operations by eliminating the hydraulic interconnection between them, which eases the permitting process from authorities. Based on numerical experiments, Court et al., (2012) confirmed that ACEM technique can control the storage zone pressure, reduce the AoR extent and mitigate the leakage risk of CO₂ and brine.

Extracted brine for pressure management can be reinjected in different over- or underling saline formations for disposal purposes or can be used as a feedstock for cooling/heating facilities, or desalination plants to secure the water demands in arid regions (Buscheck et al., 2011; Court et
al., 2012). Harto & Veil (2011) and Sullivan et al. (2013) presented cost analyses showing how expensive the handling process of the extracted brine can be. Therefore, Birkholzer et al., (2012) introduced the concept of “impact-driven pressure management”, where the build-up pressure in the storage zone is managed locally at areas where high geomechanical and leakage risks are anticipated (e.g., faults and caprock damage zones). This local-scale management method minimizes the volume of extracted brine, which makes the capital cost of CGS operations financially feasible. To design the brine extraction scheme, Birkholzer et al., (2012) used the derivative-based optimizer in iTOUGH2 (Finsterle, 2007) with an analytical solution of the pressure field in multi-layered systems.

Buscheck et al., (2014) introduced the idea of dual-mode wells, where brine is pre-extracted from the storage formation, until a significant decrease in the intact pressure (i.e., pre-injection pressure) is reached, and then the CO\textsubscript{2} is injected using the same extraction wells. Under this approach, the pressure in the vicinity of the dual-mode wells takes reasonable time to recover, which gives the system an additional pressure buffer capacity during the injection phase. Moreover, the dual-mode idea reduces the number of required wells for pressure management and thus minimizes the capital cost of the CGS project. Cihan et al., (2015) extended the method of Birkholzer et al., (2012) by incorporating the global optimizer “Differential Evolution” to find the best locations and pumping strategies of multiple extraction wells. The optimal well locations were found sensitive to the storage zone heterogeneity, while the pumping rates were sensitive to both storage zone heterogeneity and formation slope. Owing to the inherent uncertainty in the initial site characterization data, González-Nicolás et al., (2019) proposed the approach of “adaptive pressure management” that entails the steps of (1) data collection, (2) model calibration and (3) periodic brine extraction rate optimization. Under this approach, the extraction rates are cyclically
adjusted based on the collected pressure data, which optimizes the required amount of extracted brine based on site conditions. It is worth mentioning that no studies were found in the literature addressing the implementation or design of leakage control systems, when one of the EPA requirements is to submit an emergency-remedial response plan.

2.3 CGS Leakage Contamination Risk

Once ScCO₂ is injected into the brine-bearing formation, a series of physiochemical reactions (i.e., CO₂ phase-partitioning and buoyancy effects) occurs between the injected CO₂, native brine and the confining caprock, which translates into various trapping mechanisms: (a) structural, (b) residual, (c) solubility and (d) mineral trapping (IPCC, 2005). These reactions are controlled by different factors including storage formation pressure, temperature, salinity, mineralogy, residual saturation of CO₂, and the presence of intra-formation shales (IPCC, 2005; Johnson et al., 2004; Wunsch, 2013). Under structural trapping, ScCO₂ moves upward to the top of the storage reservoir by bouncy forces to be physically trapped by the caprock unit. This mechanism relies on the fact that CO₂ plume will be initially less dense than native brine. Meanwhile, the ScCO₂ will be partitioning between gaseous and dissolved phases to eventually transform to different dissolved carbonate species through a series of chemical reactions (Espinoza et al., 2011). Residual amounts of ScCO₂ will be trapped within the formation matrix by capillary forces reflecting the residual trapping mechanism (Pruess, 2008; Gaus, 2010). Under solubility trapping, the non-aqueous CO₂ phase will be slowly dissolved into native brine producing bicarbonates and carbonic acids (i.e., HCO₃⁻ and CO₃²⁻) at the plume fringes (i.e., CO₂/brine mixing zones). It is worth mentioning that this dissolution process leads to a local convective mixing of the host formation fluids that accelerates the dissolution process (Ennis-King & Paterson, 2005). Finally, the mineral trapping mechanism occurs when the dissolved rock minerals
(i.e., Ca, Mg and Fe carbonates) by CO₂-acidified brine precipitate under the conditions of solution oversaturation, high temperature, and low pressure in the storage formation (Lin et al., 2008; Oelkers, 2005).

The produced carbonic acid due to CO₂ dissolution into native brine lowers the system pH and dissolves the surrounding rock minerals of the storage formation and caprock, which potentially elevates the trace metal concentrations in the formation fluids and/or activates leakage pathways in the caprock (Rathnaweera et al., 2016; Andreani et al., 2008; Ellis et al., 2011). The leakage chemical composition is determined based on the dominant phase of CO₂ in the vicinity of the source. Three leakage forms can be expected at CGS sites: (1) free-CO₂ gas leakage that may occur at early times of CGS operation from proximate locations to the injection well; such leakage can be intercepted by capillary pressure at pore spaces and fracture apertures in the caprock (Wunsch, 2013; Siirila et al., 2012; Navarre-Sitchler et al., 2013; Carroll et al., 2014b), (2) CO₂-laden brine leakage, which may arise after CO₂ dissolution into native brine at the plume fringes (Horner et al., 2015; Andreani et al., 2008; Jung et al., 2015; 2014; Mason et al., 2013) and (3) far-field brine leakage, which can take place at significant distances from the injection site as a result of the induced injection pressure in the storage formation (Zhou et al., 2010; Jones et al., 2015; Birkholzer et al., 2011a; Celia et al., 2011; Wunsch et al., 2013).

In the absence of pH buffer reactions (i.e., presence of dissolvable calcite minerals), the upward leakages of free-CO₂ and CO₂-laden brine can considerably increase the acidity levels in the shallow aquifers. This encourages the in-situ mobilization of the major/minor elements (e.g., Ca, Mg, K, Na, Fe, Al, Ba, etc.) and hazardous trace metals (e.g., As, Pb, Cd, etc.) between the system solid and aqueous phases (Zheng et al., 2014). Moreover, CO₂-laden brine leakage can be already enriched by such toxic metals from the dissolution process of the rock minerals in the
storage formation or the overburden layers (Qafoku et al., 2017); rock mineral dissolution process is controlled by the carbonic-acid concentration in the leaking brine, which varies according to solution temperature, pressure, CO₂ solubility, and brine salinity (Horner et al., 2015). USDWs can also be contaminated by the direct mixing with the upwelling far-field brine leakage through conductive pathways (Zheng et al., 2014). Wunsch et al. (2013) conducted a comprehensive statistical analysis on the chemistry database of deep-brine in USA (67,134 data points). The results of this analysis highlighted the potential of deep saline water to contain toxic trace metals and high TDS levels (i.e., $\sim 84,400$ ppm) compared to the USDWs (TDS <10,000 ppm) (U.S. EPA, 2020; U.S. EPA, 2018). In addition, brine leakage can also indirectly contaminate the USDWs by triggering the upward migration of overlying saline groundwater bodies (Tillner et al., 2013; 2016; Delfs et al., 2016; Kissinger et al., 2017).

2.3.1 Brine Leakage Potential and Prediction

Studies based on natural analog systems to CGS application were used to understand the potential leakage mechanisms of storage formation fluids and the impacts of CO₂-enriched brine on shallow aquifers (Keating et al., 2010; 2013; Viswanathan et al., 2012). Keating et al., (2010) discovered that the elevated concentrations of U, As, and Pb in the shallow aquifer of Chimayó, New Mexico were caused solely by the upward brine leakage from deep CO₂ natural reservoirs and not by CO₂-laden brine. Kirk et al. (2009) demonstrated that the redox mobilization of Fe and Mn to Rio-Grande floodplain, New Mexico was caused by mixing the shallow groundwater with deep leaking brine. Keating et al., (2014) discussed multiple evidence to show that the fresh water at Springerville, Arizona was intruded by an upwelling brine leakage coming from deep saline formations. In a numerical analysis, Zhou et al. (2010) found that 9.5% and 62% from the displaced
brine at Mt. Simon aquifer due to injection can leak upward after 50 and 200 years of operation, respectively.

Risk assessment analysis of a CGS operation should include a relatively accurate prediction of the leakage chemical composition upon arrival at the shallow receptors to evaluate the contamination risk on the USDWs. Keating et al., (2010) and Qafoku et al., (2017) highlighted that such prediction require developing a sophisticated coupled geochemical and reactive-transport model based on our knowledge about: (1) the system flow and transport parameters (i.e., advection, mixing and dilution related parameters), (2) the pH buffering zones and scavenging capacities (e.g., silicate and calcite zones), (3) the geological and mineralogical structure of the overburden and storage formations (i.e., rock minerals and trace metal content), (4) the expected scavenging secondary reactions of hazardous trace metals (i.e., metal resorption to soil minerals) and (5) baseline measurements of the groundwater major ions. It is worth noting that sorption, ion-exchange, and precipitation processes of organic compounds and trace metals represent the “scavenging” capacity of the receptive formations. This is an important property that describe the formation ability to attenuate contaminants’ migration and mobilization. However, estimating the scavenging capacity of a formation is very difficult under field conditions, which affects the accuracy of leakage modeling (Viswanathan et al., 2012; Jones et al., 2015; Keating et al., 2013; 2010; Marcon & Kaszuba, 2015). Due to the high complexity of geochemical reactions associated with CO₂/brine leakage and the lack of above needed information to calibrate and validate models, there is an inevitable predictive uncertainty in the chemical signature of CO₂/brine leakage on the quality of shallow aquifers (i.e., USDWs).

Although, all CO₂ injection sites are monitored since injection commencement, no solid evidence of CO₂/brine leakage was found yet and most of the leakage studies are numerically
based (Emberley et al., 2005; Shevalier et al., 2013; de Caritat et al., 2013). However, leakage from deeply disposed produced-water (Thamke & Midtlyng, 2003; Jacobs, 2009, January), natural CO$_2$ reservoirs (Keating et al., 2010; 2013a; 2013b; 2014) and leaky deep saline formations (Llewellyn, 2014) has been reported. Given the long time needed for CO$_2$/brine leakage to be detectable at the shallow zones, the non-existence of reported leakage events until now can be attributed to the fact that CGS technology is relatively new compared to deep waste-fluid disposal applications (since the 1970th) (King et al., 2013). Moreover, deep brine leakage incidents may have already happened at the currently operated CGS sites but were never been detected due to the difficulty of identifying brine leakage intrusion after the intensive mixing process that occurs with the surrounding fluids during its migration phase (Choi et al., 2012).

2.3.2 Impact of Leakage Intrusion into Shallow Aquifers

In the context of CO$_2$/brine leakage impacts on the USDWs’ pH and TDS levels, Wang et al., (2016) showed that by exposing the Edwards unconfined rocks in Texas to CO$_2$-rich groundwater, the pH value dropped to only 6 due to the presence of high calcite content. Similarly, Wunsch et al., (2014; 2013) found a limited drop in the pH value (from 9.5 to ~6) after exposing limestone and dolomite rocks to CO$_2$-rich water. Zheng et al., (2016) developed a reactive-transport model, based on field data from Kharaka et al., (2010b) and Trautz et al., (2013), to interpret the changes occurred in the pH during a series of CO$_2$-exposure column experiments on rock samples from the High Plains aquifer in Texas. They found that the calcite and magnesite dissolutions (i.e., pH buffering reactions) are the main processes responsible for eliminating the decrease of pH. Bacon et al., (2016) and Carroll et al., (2014b) used reactive-transport models to predict the impact of CO$_2$ and CO$_2$-rich brine leakages on the Edwards carbonate aquifer. They found that this aquifer are likely to be subject to pH reduction and TDS increase by any of these
leakage mechanisms. Horner et al., (2015) investigated the effect of storage formation salinity on the dissolution of CO₂ to form CO₂-laden brine. They found that the low-salinity brine-bearing reservoirs are more likely to produce mineral-rich leakages than high-salinity reservoirs. Thus, CO₂-laden brine that leaks from a relatively low-salinity formation will most probably contain high concentrations of dissolved rock minerals that may include hazardous trace metals, compared to the leaking fluids from high-salinity reservoirs.

The chemistry of groundwater in sandstone and gravel aquifers can be also changed due to the interaction with CO₂-rich brine leakage. Lawter et al., (2016) showed a significant increase in the major/minor elements (i.e., Si, Ca, Mg, Na, K, Al, Fe, Mn, etc.) beside trace metals (i.e., Sb, Sr, Ba, Cr, Se, As, Pb, Cs, Cu, and Zn) in the column effluent of CO₂-exposure experiments conducted on unconsolidated aquifer sediments. The reactive transport model developed by Zheng et al., (2016) predicted a dissolution-release of Sb, Sr, Ba, Cr, Se, As, Pb, Cs, Cu, and Zn and an adsorption/desorption-release of As and Pb in the groundwater of the High Plains aquifer in Texas after the intrusion of CO₂-laden brine leakage. Numerous other experimental and modeling studies predicted an increase in the elements of As and Pb due to the interaction between the shallow aquifer sediments and CO₂/brine leakage (Apps et al., 2010; Carroll et al., 2009; Wang & Jaffe, 2004; Zheng et al., 2009; Humez et al., 2013; Kirsch et al., 2014; Little & Jackson, 2010; Lu et al., 2010; Smyth et al., 2009; Varadharajan et al., 2013).

For carbonate aquifers, Bacon et al., (2016) showed an increase in the concentration of all trace metals, except Cr, after exposing the Edwards aquifer sediments to elevated CO₂ conditions. The laboratory experiments performed by Wunsch et al., (2014; 2013) proved the responsibility of interbedded carbonates in the limestone and dolomite rocks for releasing different trace metals after being in contact with CO₂-rich groundwater. Wigley et al., (2013) looked at the characteristics
of aquifer sediments that have been flooded once by CO$_2$-laden brines. They found that the hematite grain coating was bleached out during the intrusion process, which released Cu, Pb, and Zn metals. Qafoku et al. (2017) highlighted that hazardous trace metals (e.g., As and Pb) located at deep storage saline zones or overlying formations can migrate upward with CO$_2$/brine leakage and adversely impact the USDWs quality.

Organic contaminants can also migrate with CO$_2$/brine leakage and affect the water quality of the USDWs. Kharaka et al., (2006) demonstrated that CO$_2$ injection can elevate the concentration of dissolved organic carbon in native brine. While different researchers highlighted the potential risk of increasing the concentrations of dissolved phenols, PAHs, BTEX and other organic compounds in the leaking fluids due to CO$_2$ injection (Scherf et al., 2011; Kharaka et al., 2009; 2010a; 2010b). Several laboratory experiments showed the potential mobilization of organic compounds from organic-rich storage formations to native brine due to the injection of ScCO$_2$ (Zhong et al., 2014; Kolak et al., 2015; Jarboe et al., 2015). Phenols, PAHs, and BTEX were identified as the most likely organic compounds to leach as a consequence of CO$_2$ injection (Zheng et al., 2013; Cantrell & Brown, 2014). However, field evidence on the existence of organic compounds in native brine at deep storage formations were rarely reported in the literature (Qafoku et al., 2017). Carroll et al., (2014a) developed a reactive-transport model to simulate the benzene migration with CO$_2$ leakage from a storage formation. They found that the plume size of the benzene leakage is constantly smaller and more concentrated close to the source compared to CO$_2$ leakage plume due to the plume continuous dissipation by biodegradation and adsorption of the benzene plume.
2.4 Hydraulic Fracturing and Waste Disposal

Under the UIC program, there are other deep injection applications that can potentially lead to deep brine leakage as well and thus can benefit from the findings of this study. In this section, the potential brine leakage resulted from the operations of hydraulic fracturing and deep waste disposal are briefly discussed. Hydraulic fracturing (HF) is a well-established method for stimulating oil and gas productions from unconventional reservoirs (e.g., low permeability black shale, tight sandstone, or coal formations) by developing small permeable fractures via injecting highly pressurized fracturing fluids (Barati & Liang, 2014; U.S. EPA, 2016), as shown in Figure 2.4. Several studies demonstrated the potential of vertical fluid migration from the target formations as a consequence of HF operations using numerical approaches (Warner et al., 2012; Rozell & Reaven, 2012; Myers, 2012; Gassiat et al., 2013; Kissinger et al., 2013; Brownlow et al., 2016) and field observations (Llewellyn et al., 2015; Llewellyn, 2014). However, other studies opposed these suggestions either by field observations or by counter probabilistic and numerical model findings that exhibit the low likelihood for such leakage incidents (Shanafield et al., 2019; Cohen et al., 2013; Saiers & Barth, 2012; Flewelling & Sharma, 2014; 2015; Engelder et al., 2014).

Jackson et al., (2013) summarized the possible occasions of formation fluid to leak due to fracking operations in the following cases: when the new fracture (1) extends beyond the target formation limits, (2) connects with natural geological discontinuities (e.g., fracture or faults) or (3) intersects with existing abandoned wells. They also argued that out of all these potential leakage mechanisms, the interaction between the newly induced fractures and existing faults, fractures or abandoned wells in the system can be the most expected leakage mechanism in the HF process. A comprehensive review about the potential risks on shallow water resources promoted by HF operations can be found in Vengosh et al., (2014) and McIntosh et al (2018). In the same context,
McIntosh & Ferguson (2019) urged the need to monitor enhanced oil recovery (EoR) operations and deep waste disposal activities for their potential to cause a higher risk of formation fluid migration than HF.

Figure 2.4 Schematic depicting potential deep brine leakage pathways as consequence of the hydraulic fracturing process.

Deep waste disposal (DWD) technique was basically introduced as a mean to get rid of the oil and gas byproducts (i.e., frack chemicals, saline produced water, radioactive material, and heavy metals) by injecting them into deep saline formations (Clark et al., 2005; Harlow, 1939). Other industrial waste fluids, municipal wastewater, and hazardous liquids have been also disposed
into subsurface formations via deep injection wells (Saripalli et al., 2000; Aust, 1985). Reused depleted oil reservoirs as waste disposal repositories can contain a considerable volume of residual brine. This is because the depleted oil reservoirs can be recharged by brine through several ways (Fakhru’l-Razi et al., 2009); (1) vertical brine leaching, (2) lateral brine inflow and (3) injected brine during EoR process. Thus, injecting liquid wastes under high pressure into depleted oil reservoirs for disposal purposes can lead to brine leakage through natural pathways (e.g., faults and fractures), or via the induced rock matrix cracks during oil production or HF (Gurevich & Chilingarian, 1993).

The high content of suspended solids and oil droplets in the liquid wastes can partially or completely clog the recipient formation pore-space and shrink the available storage capacity (Melcer & Gerrish, 1996; Ochi & Vernoux, 1996). This loss in the storage capacity elevates the pressure levels in the injection formation causing unfavorable rock fracturing or fault activation process in the overlying caprock (Saripalli et al., 2000). This increases the probability of formation brine leakage. Lately, Marsac (2019) explored the impacts of disposing calcium-rich produced water into deep saline formations. She concluded that at high levels of calcium content and formation salinity, the injected produced water can clog the porous media located immediately around the injection well, which may lead to well cementing failure and fluid leakage. Wesson & Nicholson (1987) emphasized that formation clogging can exponentially accelerate the pressure build-up in the disposal formation promoting high leakage risk. Moreover, Wesson & Nicholson (1987) and Folger & Tiemann (2016) highlighted that the induced pressure in the waste host formation propagates horizontally away from the injection point by several Km and could last for long time, in order of years, after ceasing injection, which make a rise to the potential of far-field brine leakage.
2.5 CGS Monitoring

A comprehensive CGS management strategy that involves a monitoring, measurement, and verification (MMV) plan is essential to be developed prior to CO₂ injection to ensure safe and efficient long-term confinement of the injected ScCO₂ (Romanak et al., 2012; Pawar et al., 2013; 2014; 2015). This MMV plan should describe the techniques that will be used to detect the leakage at the overlaying formations during and after CO₂ injection (U.S. EPA, 2010; U.S. EPA, 2013).

Different monitoring techniques were developed to detect and track CO₂/brine leakage, such as: (1) monitoring micro seismicity (Arts et al., 2008), (2) monitoring surface deformation using satellite-borne synthetic-aperture radar (Onuma & Ohkawa, 2009), (3) monitoring vegetation stresses by remote sensing (Rouse et al., 2010), (4) collecting and analyzing carbon isotopes from the shallow aquifers and sediments (Krevor et al., 2010), (5) monitoring the vertical turbulent fluxes at the ground surface using the eddy-covariance towers (Lewicki et al., 2009), (6) monitoring the gaseous phase of CO₂ using flux accumulation chambers (Yang et al., 2012; Ren et al., 2016; Schroder et al., 2016), (7) collecting and analyzing deep and near-surface soil, gas and groundwater samples (Benson, 2006; Dai et al., 2014; Yang et al., 2015), (8) monitoring pressure anomalies (Nogues et al., 2011; Keating et al., 2014; Wang & Small, 2014), (9) observing the bulk electric-conductivity field (Jewell et al., 2015) and (10) using partitioning tracers (Myers et al., 2012; 2015). These methods vary in the detection time (e.g., early to late detection), data resolution/accuracy, monitoring extent/depth (e.g., shallow, or deep systems) and target parameters (e.g., leakage rate, occurrence, and concentration). In addition, many of these techniques can also be integrated for data cross-validation to ensure accurate leakage identification, extending the monitoring area, and applying deep monitoring techniques for early detection of the leakage.
2.5.1 Source Identification

The ability of a CGS monitoring system to provide useful data to identify the leakage source is one of the goals that need to be considered during the design of such systems to ensure efficient planning, designing and implementation of corrective actions and remediation measures (Jung et al., 2013). Various numerical and analytical approaches were developed to address different aspects related to contaminant source identification. For example, number of approaches were developed to reconstruct the contaminant release history (Skaggs & Kabala, 1995; Atmadja & Bagtzoglou, 2001a; Wagner, 1992), identify the source location and leakage age (Dimov et al., 1996; Neupauer & Wilson, 1999; 2005) or regenerate the contaminant distribution history (Michalak & Kitanidis, 2004). Atmadja & Bagtzoglou (2001b) classified these approaches with respect to their mathematical foundations as follows; (1) optimization-based approaches (Jury & Roth, 1990; Mahar & Datta, 2000; Skaggs & Kabala, 1994; Liu & Ball, 1999), (2) analytical methods (Alapati & Kabala, 2000; Macdonald, 1995), (3) direct methods (Woodbury & Ulrych, 1996; Neupauer et al., 2000), (4) probabilistic/geostatistical approaches (Neupauer & Wilson, 1999; Wilson & Liu, 1994; Woodbury & Ulrych, 1996; Woodbury et al., 1998) and (5) inversion methods (Clark & Oppenheimer, 1994; Skaggs & Kabala, 1995).

Generally, source identification problems have a relatively high ill-posedness nature due to the lack of knowledge about hydrodynamic dispersion histories, model structure errors and data uncertainty (Milnes & Perrochet, 2007; Neupauer & Lin, 2006). Dealing with complex heterogeneous systems requires detailed information about the velocity field and transport parameters with accurate boundary conditions (BCs). Problems with multidimensional heterogenous environments were not commonly addressed in the literature because of its complexity and the need of several assumptions to be given (Milnes & Perrochet, 2007). However,

In this section, methods related to identifying the source location and initial leakage concentration are reviewed. The first attempt to locate a contaminant source using an optimization method was made by Gorelick et al., (1983), who adapted a linear least-square regression technique. Neupauer & Wilson (1999; 2001; 2002) and Neupauer & Lin (2006) used downstream concentration data to estimate a backward probability density function (PDF) for the contaminant source location by applying an adjoint inverse method. Mahar & Datta (2000) adapted a non-linear optimization approach to estimate the released contaminant concentrations from several hypothetical sources using discrete downstream breakthrough curves (BTCs), which enabled the identification of the source location.

Mahar & Datta (2001) developed a non-linear optimization-based method for simultaneously estimating model parameters and identifying pollution source using scattered observed concentrations over the domain. Aral et al. (2001) utilized observed concentration data from four wells to reconstruct the contaminant release history and estimate the source location by adapting an optimization scheme. Mahinthakumar & Sayeed (2005) used the Genetic Algorithm to estimate the location of an instantaneous non-point contamination source with uniform initial concentration. Milnes & Perrochet (2007) proposed a deterministic approach to identify the contaminant source location by solving the advection-dispersion equation for a reversed flow field. Lately, artificial neural networks (ANNs) also played an important role in the field of groundwater-pollutant source characterization (Singh et al., 2004; Singh & Datta, 2007; Srivastava & Singh, 2014).
All above methods rely on the ability of the forward model underlying the search process to capture the hydro-physical processes controlling the contaminant transport. Otherwise, a biased estimate of the contaminant source location will be resulted. To enhance the performance of source identification methods, the computationally intensive forward models were commonly replaced by less computational surrogate models (Leichombam & Bhattacharjya, 2016; Khalil et al., 2005). Despite the fact that surrogate models (e.g., reduced order models) can be of a great benefit to reduce the computational burden of the source identification problem, they can also add extra error to the results due to the high-level of approximation to the real system complexity applied in these models.

2.5.2 Monitoring System Design

Our conceptual understanding of groundwater systems depends mainly on the available information about the hydrogeological regime in the area. Due to the high cost of site investigation, having incomplete data on the geological and hydraulic settings of subsurface formations is a typical case. Therefore, the question of "where are the most critical locations in the system that can provide the most informative data to a groundwater model?" is frequently raised among groundwater practitioners. Most of the developed strategies to answer this question can be found in the literature under different topics, such as data worth analysis, value of information, monitoring system optimization, and investigation plan design (Kikuchi, 2017), but eventually they are all addressing the same issue.

One essential driver for optimizing the design of groundwater monitoring systems is to reduce the predictive uncertainty of groundwater models (Herrera, 2005; Leube et al., 2012; Hill et al., 2013; Wöhling et al., 2016). Three main factors determine what type of technique can be used to design a monitoring system (Kikuchi, 2017): (1) monitoring purpose, (2) implementation
phase (e.g., pre-, or post-data collection) and (3) data dimensionality (i.e., temporal, and spatial data density). Since these factors are case dependent and range widely within the groundwater management environmental issues, various optimization methods were developed that rely on different quantitative ranking metrics of the proposed monitoring schemes.

Optimization techniques can be categorized into four types based on the minimum-search engine employed in the optimization code: “(1) gradient-based algorithms that rely on derivative methods; (2) heuristic algorithms that follow empirical guidelines and incorporate elements of structured randomness; (3) direct-search algorithms that utilize deterministic stencil-based procedures; and (4) hybrid methods that combine different aspects from the three previous categories.” (Yonkofski et al., 2016). Examples of these optimization techniques include: the Levenberg–Marquardt code (i.e., gradient-based approach) that was used by Ruprecht (2014) to find the minimum injection cost of CO₂ that guarantees a maximum CO₂ secondary trapping under different injection rates, the GATOUGH optimizer (i.e., heuristic optimization algorithm) that was used by Zhang & Agarwal (2013) to improve the storage capacity and safety of CO₂ in saline aquifers, and the Hooke–Jeeves code (i.e., direct search algorithm) that was used by Cameron & Durlofsky (2012) to find the optimum CO₂ injection well location, pumping rate and brine extraction rate. One example of using hybrid methods in CGS-related optimization problems can be found in the work of Jahangiri & Zhang (2012); their objective was to co-optimize the sequestration state of CO₂ and the corresponding oil recovery.

CGS-related optimization problems are typically formulated to minimize the leakage detection time, the overall operation cost, the number of monitoring points, or maximize the storage efficiency (Yonkofski et al., 2016). Yang et al. (2011) developed a tool that estimates the detection probability of CO₂-flux by near-surface sensors, which was then used as a metric to
evaluate the worth of different configurations of monitoring sensors. Yang et al., (2017) improved the sensors’ worth evaluation of this code through defining a threshold value of the CO₂-flux in the shallow zone that indicates the occurrence of real CO₂ leakage event. Seto & McRae (2011a) proposed a design framework for CGS monitoring systems that selects the best observation points based on a trade-off analysis between operation cost and environmental safety. In the same year, Seto & McRae (2011b) developed a selection operator of the best monitoring locations of CO₂ leakage using a Bayesian approach to handle different leakage pathway realizations. Sun & Nicot (2012) and Sun et al. (2013a) introduced two optimization methods that rely on evaluating the detectability of a set of proposed monitoring systems to leakage-driven pressure anomaly.

Fang et al., (2014) used the parameter estimator “PEST” to evaluate the efficacy of proposed monitoring systems based on the pressure and saturation data informativity to CO₂ storage properties. Dai et al., (2015) developed a probabilistic optimization tool for CO₂ leakage surveillance operation based on data worth analysis using a collocation-based Kalman Filter. Dai et al. (2016) expanded this approach by incorporating different datasets from multiple potential monitoring schemes. Then, they evaluated the uncertainty reduction in the prediction of the contaminant migration by each dataset to select the best monitoring scheme. Yonkofski et al., (2016) proposed a simulation-based optimizer to find the best sensor locations that provide the minimum detection time of CO₂ leakage using an annealing search code. A stochastic optimization approach was suggested by Chen et al., (2017) to evaluate the predictive uncertainty reduction under different monitoring data realizations using a Markov Chain Monte Carlo method.

Chen et al., (2018) developed a filtering-based data assimilation algorithm to design a monitoring system of CGS operations. The algorithm finds the optimum locations of monitoring wells and determine the most informative data type under the uncertainty conditions of the
information about the CO$_2$ storage system. Jeong et al., (2019) adapted a probabilistic approach to evaluate the CO$_2$ leakage detectability by five potential near-surface monitoring systems. They tested the five systems under a set of 200 heterogeneity realizations of the overburden formations. They concluded that the most efficient configuration of the near-surface sensors is to be evenly spaced, particularly when available data is limited. However, if there were enough data available about the system, placing the sensors between the low and high permeability zones turned to be the most efficient scheme (Jeong et al., 2019).

Incooperating predictive uncertainty analysis for determining the worth of data to be collected from the proposed monitoring locations was applied by several researchers. Dausman et al., (2010) and Fienen et al., (2010) have used the “PREDUNC” utility in “PEST” (Doherty et al., 2010; 2015) to conduct a predictive uncertainty analysis and design the groundwater monitoring system. While Wöhling et al., (2016) incorporated the “PREDUNC” utility with the Generic Algorithm (GA) (Goldberg, 1989), to simultaneously evaluate the worth of groups of new sensors instead of sequentially evaluating the worth of single sensors. Vilhelmsen & Ferré (2018) extended the capacity of Wöhling et al.’s (2016) approach to select the best monitoring design that provides informative data to a set of different predictions of interest. Safi et al (2019) used Vilhelmsen & Ferré’s (2018) approach to select the best locations and depths of observation wells, while constraining the selection algorithm on predefined limits of the project capital cost. Zell et al., (2018) used both linear and non-linear uncertainty analysis methods to evaluate the importance of specific measurements (i.e., head, discharge, and tracer information) to characterize the age of base-flow discharge and accordingly design the monitoring system.
2.5.3 Prediction Uncertainty Analysis

Several non-linear techniques, with different advantages and limitations, were developed and used to quantify the posterior model parameter uncertainties and propagate them to model predictions. The most commonly applied non-linear techniques in the literature are the Generalized Likelihood Uncertainty Estimator (GLUE) and the Markov Chain Monte Carlo (MCMC) based methods. In GLUE method (Beven & Binley, 1992; Beven & Freer, 2001), the forward model has to be run many times to conduct Monte Carlo simulations based on parameter realizations sampled from predefined prior probability density functions (PDFs). The fit between the simulated and observed data is then evaluated using an appropriate likelihood function to rank the used parameter realizations. Finally, the posterior PDFs of model parameters and predictions (i.e., posterior PDF variance is the predictive uncertainty) are calculated using only the highly ranked parameter realizations (Wu, 2013). GLUE is a simply structured technique based on informal/random sampling operator, which gives it a wide applicability range under different environmental models (Freer et al., 1996; Aronica et al., 1998; Beven & Freer, 2001; Brazier et al., 2001; Mertens et al., 2004; Hassan et al., 2008; Jin et al., 2010).

GLUE method was considered inefficient in exploring the predictive uncertainty of complex models because of the large number of required forward runs to reach a proper convergence of the Monte Carlo simulations, which boosts the computational burden of the method (Blasone et al., 2008). Moreover, Mantovan and Todini (2006; 2007) raised concerns about the potential inconsistency and incoherence in GLUE evaluation of model predictive uncertainty due to the informal/random sampling of model parameters. The dynamic sampling-based method MCMC was then employed to provide a formal search of model parameters via applying the evolving Markov chain procedures. MCMC relies on using proposal and acceptance functions for
updating model parameters, which gives it a formal sampling nature. MCMC is usually applied within a Bayesian context to enhance inferring the posterior parameter and prediction uncertainties (Kuczera & Parent, 1998; Gallagher & Doherty, 2007a).

The procedure of parameter sampling is the key difference between the MCMC and GLUE methods. Various MCMC-based algorithms were emerged by altering the parameter sampling procedure from single-chain sampling methods (Metropolis et al., 1953; Hastings, 1970; Haario et al., 1999) to multi-chain sampling methods (Vrugt et al., 2003; 2009; Ter Braak, 2006). MCMC has a superior performance in dealing with complex parameter spaces, featured by strong nonlinearity, high dimensionality, and multimodal probability distributions, over the traditional GLUE method (Blasone et al., 2008; Hassan et al., 2009; Wu, 2013).

Zell et al., (2018) highlighted that the burdensome computational cost of 3D regional groundwater models (e.g., multi-million node models) hinders using non-linear methods to estimate prediction uncertainty. This is because of the large number of forward runs required for obtaining a stationary posterior distribution of model parameters. In recognition to such numerical obstacle, researchers started to explore different techniques to minimize the computational burden of non-linear uncertainty analysis. For example, Hassan et al., (2009) replaced the forward model by a relationship between model parameters (i.e., the ratio between the recharge and conductivity) and the likelihood function parameters (i.e., the sum of squared errors and error variance) to reduce the computational cost of the MCMC method. Tonkin & Doherty (2009) and Herckenrath et al., (2011) reduced the number of required forward runs by the MCMC method by decreasing the dimensions of the parameter space using the sensitivity matrices developed during parameter estimation. This approach is called calibration-constraint MC or Null-Space MC. Chen & Oliver (2013) used the iterative ensemble smoother (iES) technique (Van Leeuwen & Evensen, 1996) to
minimize the number of model runs needed for constructing the necessary sensitivity matrices of highly parameterized models for uncertainty analysis. Lately, White (2018) incorporated the iES technique in PEST (i.e., pestpp-ies) to lessen the computational load posed by high-dimensional parameter spaces during predictive uncertainty analysis. iES has been successfully implemented to enhance data assimilation for different applications (Crestani et al., 2013; Emerick & Reynolds, 2013; Bocquet & Sakov, 2014).

Despite the effort devoted in minimizing the computational burden of non-linear uncertainty analysis tools, high fidelity transport models still present a significant obstacle to conduct such type of analysis under limited computational resources. Therefore, the linear first-order second-moment (FOSM) technique became particularly important for analyzing the predictive uncertainty of the computationally intensive models. In the FOSM, the model is assumed to be linear (i.e., the relationship between model outputs and inputs can be represented by a matrix) and the noise of model parameters, observations and predictions vary according to a Gaussian distribution (Dausman et al., 2010). Under these assumptions, the uncertainty of measurements and parameters can be propagated to predictions using the parameters-to-data and predictions-to-parameters sensitivity matrices. To calculate these matrices, the forward model has to be run only once, which reduces the computational cost associated with using the FOSM technique significantly.

2.6 Numerical Modeling

The fundamentals, assumptions, and limitations of flow and solute transport governing equations are briefly discussed in this section. For more details, a comprehensive review can be found in Bear and Cheng (2010).
2.6.1 Deterministic Modeling

Modeling groundwater flow in porous media can be traced back to the empirical work of Darcy (1856), when he found that the fluid flux through porous media in a cylindric soil column is proportional to the hydraulic head difference between upstream and downstream sides. His work was extended by Bear (1972) to cover the three dimensions of groundwater systems using the fluid mechanics continuum theorem (Darcy, 1856; Bear, 1972). The empirically founded Darcy’s law presented in Equation 1 can also be derived by averaging the Navier-Stokes equations for porous media flow and discounting all local and convective acceleration terms (Bear, 1972; Barnhart, 2010). Thus, applying Darcy’s law to porous media results in an average velocity at every point in the system (Barnhart, 2010), which masks the microscopic velocity variation occurring at the pore-scale. By substituting Darcy’s law in the mass conservative equation, a partial differential equation describing the hydraulic head distribution across the system domain can be derived, as presented in Equation 2, where \( q \) is Darcy’s flux, \( \phi \) is porosity, \( \mathbf{v} \) is the seepage velocity tensor, \( \mathbf{K} \) is the hydraulic conductivity tensor, \( \mathbf{h} \) is the hydraulic head vector, \( S_s \) is specific storage, \( t \) is time, and \( q_s \) are the sources and/or sinks.

\[
q = \phi \mathbf{v} = -\mathbf{K} \nabla \mathbf{h} \tag{1}
\]

\[
\nabla \cdot (\mathbf{K} \nabla \mathbf{h}) = S_s \frac{\partial \mathbf{h}}{\partial t} + q_s \tag{2}
\]

Equation 2 is solved in the framework of solute transport modeling to obtain the hydraulic head field from which the average velocity field (i.e., Darcy’s velocity) is calculated using Equation 1. The spreading rate of the solute particles in porous media is not only driven by the flow in the pore space (i.e., velocity field) but also is affected by other factors such as solid phase adsorption (i.e., soil particles), chemical reactions, and molecular diffusion. Thus, a different
equation was necessary to be developed to track the actual transport of the solute particles across the model domain. Advection-dispersion equation (ADE) was classically used for this task, as it contains components that describe the solute mass flux due to mechanical dispersion, advection, and molecular diffusion and can be written as expressed in Equation 3. Although, the ADE ignores the changes in the mass transfer due to chemical reactions between the solid and liquid phases (Barnhart, 2010), it still can be used for simulating the transport of conservative solutes.

\[
\frac{\partial \bar{C}}{\partial t} = -q \nabla \bar{C} + \nabla \cdot (\phi D \nabla \bar{C}) \quad (3)
\]

\[
D = \nu \alpha L + D^* \quad (4)
\]

In Equations 3 and 4, \((\bar{C})\) is the solute concentration vector, \((D)\) is the hydrodynamic dispersion tensor (i.e., the sum of the molecular diffusion \((D^*)\) and the mechanical dispersion \((\nu \alpha L)\), and \((\alpha L)\) is the longitudinal dispersivity. The credit for developing the ADE may be given to the early work of Bird et al. (1960), while the concept of separating the dispersion term from the advection and diffusion terms dates back to Taylor (1953). Wide variety of codes are available nowadays for solving these equations using different numerical approaches (e.g., finite element and finite difference-based solvers – FEM and FDM) (Wang & Anderson, 1995; Anderson et al., 2015).

The first term of Equation 3 describes the mass flux due to the advection transport of the solute particles, which is strongly related to the velocity field in the system. While the second term describes the mass flux due to molecular diffusion and local mechanical dispersion merged under one term called “hydrodynamic dispersion”. Theoretically, if we know the “true” micro-velocity field of a system, transport can be described by the advection and diffusion processes only (Molz, 2015). It is practically prohibited to map the micro-velocity field over the pore space structure of
a large-scale problem to simulate the micro flow through the media erratic flow paths (i.e., tortuous micro channels). Therefore, the mechanical dispersion was tied to the velocity field generated by applying Darcy’s law that basically disregard the pore-scale velocities giving an average velocity field across the system. This is called the macrodispersion theory, where the local dispersion processes at the pore space is upcaled to the representative elementary volume (REV) of the transport domain.

The mechanical dispersion term in the ADE is assumed to follow Fick’s first law by analogy to molecular diffusion (Taylor, 1953; Taylor., 1922), which might be true only in ideal homogeneous isotropy media (Bear, 1961; Deng et al., 2004). This assumption was later critiqued by field studies and was considered as a potential source of error in field scale transport models (Harvey & Gorelick, 2000). Corey & Auvermann (2003) discussed some empirical evidence against this assumption using a controlled laboratory experiment. Thus, the ADE is only valid for macroscale domains of porous media, where the observation scale (e.g., prediction scale or plume extent) is larger than the fickian limits of dispersivity or larger than the heterogeneity correlation length at which asymptotic conditions can be reached. In some cases, the dispersivity fickian limits cannot even be reached (Howington et al., 1997), as in Borden site where a data point at 1,040 days was disregard to report a longitudinal field dispersivity of 0.49m (Molz, 2015).

In 1992, Gelhar provided a critical review for 59 field scale dispersivity studies showing three main facts: (1) dispersivity values change with the study-scale, (2) dispersivity may reach asymptotic conditions after extremely large spatial scales, and (3) there is a great variation in the dispersivity values of a given scale of investigation (Gelhar et al., 1992). Several studies in the literature support the scale dependency of dispersivity by showing that field dispersion coefficients vary by orders of magnitude than the experimentally founded ones (Anderson & Cherry, 1979;
Gelhar, 1986; Neuman, 1990a; Sudicky & Cherry, 1979). For this reason, most of the modelers started to treat dispersivity as a fitting parameter that depends on model scale. This practice was later censured by Gelhar et al., (1992) and Molz (2015) because, by that, dispersivity loses its physical meaning and is adjusted to compensate our lack of knowledge about the system. Despite these concerns, modelers still consider dispersivity as a fitting parameter besides hydraulic conductivity (K). While some studies concluded its low sensitivity to the transport model compared to K (Ma et al., 2012; Sarris et al., 2018).

The dispersion term in the ADE is positively correlated with the solute velocity, which is controlled by the K-field of the system (Barnhart, 2010; Ma et al., 2012). The more detailed K data incorporated into the flow model (i.e., micropore scale data), the smaller value of the mechanical dispersion is needed to simulate the solute transport, as it will be more controlled by molecular diffusion (Ma et al., 2012; Zheng et al., 2011). Thus, the inherent uncertainty in the K-field description, due to limited spatial resolution of available geological data, can impact the accuracy of transport modeling predictions (Zheng & Jiao, 1998; Yoon & McKenna, 2012). Zheng & Jiao (1998) emphasized on the need of new innovative monitoring systems that can help in providing information about the aquifer properties (e.g., K-field) at a scale useful for developing a detailed transport modeling.

2.6.2 Stochastic Modeling

Stochastic methods were founded to count for the inherent uncertainty in groundwater modelling parameters (e.g., K). The early analysis of uncertainty in groundwater flow, conducted by Freeze (1975), helped establishing the baseline principles of stochastic modeling in hydrogeology. Later, large number of researchers worked on developing and enhancing stochastic techniques after realizing the high complexity of modeling subsurface environments
deterministically (Freyberg, 1986; Sudicky, 1986; Garabedian et al., 1991; Farrell et al., 1994b; Mackay et al., 1994; Dagan, 1989; Gelhar, 1993; Kitanidis, 1997; Zhang, 2002). In stochastic modeling, parameters are described in the groundwater flow and transport equations as random space functions that represent their variability in space under natural heterogeneous conditions (e.g., the \( K \) and \( D \) are defined using PDFs). Given that \( K \) is a constitutive parameter in the groundwater flow equation, the results of head, flux and thus solute concentrations will be displayed as probabilistic distributions as well. Such results can be described by a set of realizations (i.e., results ensemble) that represent all possible solutions under the uncertainty of input parameters (e.g., \( K \) and \( D \)). A comprehensive derivation of the stochastic forms of the governing Equations 2 and 4 can be found in Barth (1999).

In order to develop a representative space random function for the \( K \)-field, extensive knowledge about the probability of each possible spatial distribution (i.e., realization) of the \( K \) in the space is required, which can be very challenging to characterize in the field (El-Kadi, 1995; Yeh et al., 2015). By that, stochastic modeling can be more complicated and data demanding than conventional deterministic numerical modeling as argued by Phillips et al., (1989). Thus, different researchers started to explore field data from well-characterized sites (e.g., Borden and Cape-Cod) in attempts to find an adequate statistical description (i.e., PDF function) of the \( K \) probability distribution in the space (Freeze, 1975; Sudicky, 1986; Gómez-Hernández & G., 1989; Sudicky et al., 2010). It was found that the log-normal PDF can reasonably represent the distribution of \( K \) data in the field as shown in Figure 2.5. The advantage of the log-normal PDF is that it can be fully characterized by the first and second statistical moments only (i.e., mean (\( \mu \)) and variance (\( \sigma^2 \))). The PDF mean and variance do not describe the spatial distribution of \( K \) values in the space but represent the overall statistics of the \( K \)-field. However, the spatial distribution of the \( K \) values can
be described using the covariance matrix, which measures the linear variability of $K$ values between any two points over the study domain. The covariance of a system can be displayed graphically by the variogram (Barnes, 1991).

High order moments are not used to describe the $K$-PDF in groundwater models because they are more vulnerable to field data noise and need a lot of time and effort to be characterized (Barth, 1999). To enable the statistical representation of the $K$-field, simplifying assumptions like stationarity and ergodicity were essential to linearize the spatial correlation functions. Stationarity suggests that the first moment (i.e., mean) remains constant over the entire system domain and covariance varies solely with the separation distance between any two points (Kelkar et al., 2002). Therefore, stationarity assumption can mask some important geological features that occupies large spatial stretches, such as fossil channel deposits. This makes the stationarity assumption always successful in cases where only few data are available on the system (Chilès & Desassis, 2018). However, ergodicity implies that the characterized statistical properties based on field data (i.e., one realization of $K$) can be used as an average of the realization ensemble representing the entire study area (i.e., population distribution of $K$).

Figure 2.5 $\ln(K)$ distribution of two well-characterized research sites of (left) Borden and (right) Cape-Cod (Sudicky, 1986; Sudicky et al., 2010).
Several techniques were developed to generate multiple realizations of the $K$-field, such as geostatistical methods (De Marsily et al., 2005), stochastic simulations (Deutsch & Journel, 1998; Journel & Zhang, 2006), geological-process models (Koltermann & Gorelick, 1996), and multi-point geostatistical approaches (Hu & Chugunova, 2008). In geostatistical methods, a random field generator is combined with kriging interpolation and Monte Carlo sampling engines to create the $K$-field realizations based on predefined statistical properties ($\mu$, $\sigma^2$ and variogram). Relatively similar approach is applied by the stochastic kriging-based simulators, such as Sequential Gaussian Simulator (SGSIM) and Sequential Indicator Simulator (SISIM), to produce multiple equiprobable realizations for the $K$-field. However, geological-models are used to predict all possible heterogeneities of the $K$-field based on the expected depositioning mechanisms of the sediments. Multi-point geostatistics were mainly founded to overcome the limitations of two-point based simulators in reproducing complex heterogeneities and non-smooth features. The required multiple data points for this method can be retrieved from, what so called, training images that are developed based on prior information about the system geological settings (Tahmasebi, 2018).

Stochastic simulation is generally computationally expensive, particularly for multimillion-node models because of the number of required realizations to reach reasonable convergence of the flow and transport stochastic equations. Convergence means that the overall statistics of the ensemble results (e.g., mean and variance) do not show variation with increasing the number of realizations anymore. The complexity of aquifer heterogeneity contributes also to the computational cost of stochastic modeling as it increases the number of required realizations for convergence. It should be noted that the number of required realizations to obtain convergence of second order moment (i.e., variance) in transport models is much larger than of those needed to converge the first statistical moment (i.e., mean) (Hassan et al., 1998). Bellin et al., (1992) showed
that 1500 Monte Carlo realizations were needed to reach an acceptable convergence of the mean and variance of the migrating particles trajectory in a mild heterogeneity transport model. While Hassan et al., (1998) used around 1200 to 2400 realizations to reach an appropriate convergence of the same statistical moments under almost the same \( \text{ln}K \) variance of Bellin et al.’s (1992) study area. Several ensemble reduction methods were introduced in the literature because of the high computational cost of stochastic modeling, including Latin hypercube sampling LHS (Gwo et al., 1996; Lahkim et al., 1999), realizations ranking (Kupfersberger & Deutsch, 1999), and percolation theory-based realization ranking (Deutsch, 1998). Awad et al., (2018) showed that only 10% of the ensemble realizations can be used to approximate the full ensemble results in a 3D flow and transport model.

2.7 Inverse Modeling

The deterministic transport modeling discussed in Section 2.6.1, employs the input parameters of \( K \) and \( D \) with the initial and BCs (i.e., constant flux or heads) in a numerical solver (e.g., FEM or FDM) to make the predictions of interest (e.g., solute concentration). Such process is called a single forward model run and results in a unique solution (e.g., one concentration distribution) that corresponds to the used input values and BCs. However, if we reversed such process by treating model parameters (e.g., \( K \) and \( D \)) as dependent variables and field observation data (e.g., concentration distribution) as input data under initial and BCs (i.e., independent variables) of the system, this operation will be termed as direct inverse modeling, direct parameter estimation or direct model calibration (Neuman, 1973). The credit of developing the earliest fundamentals of the direct inversion methods can be given to Stallman (1956).

Since the number of estimated parameters in a hydrological model, through an inverse problem, will always be greater than the available data, such problems will always be ill-posed
Moreover, some data can be insensitive to target model parameters (e.g., $K$ and $D$), which leads to an ill-posed problem as well (Zhou et al., 2014). Ill-posedness may cause three mathematical problems: non-existence, non-uniqueness, and non-steadiness of the solution (Anderson et al., 2015; Zhou et al., 2014). During the last few decades, a considerable research effort was advocated to alleviate the ill-posedness in groundwater inverse problems. Examples of the proposed techniques in the literature to resolve this issue include (Zhou et al., 2014; Carrera et al., 2005; McLaughlin & Townley, 1996): (1) balancing between the number of target parameters and observations either by reducing the unknown parameters or increasing observed data, (2) limiting the target parameter values within a narrow prior knowledge bounds (Neuman, 1973), (3) considering a regularization term in the objective function (i.e., a function that expresses the misfit between observed data and model results), (4) collecting sensitive data to model parameters, and (5) linearizing the governing equations of the forward model underlying the inverse problem (Carrera & Neuman, 1986).

To obtain a reasonable unique solution of the target model parameters using direct inverse modeling, error-free data over the entire model domain (at each node/cell) with accurate initial and BCs are required (Zhou et al., 2014). Due to the impossibility of obtaining such dataset, a new class of inverse modeling was developed, which is called indirect inverse modeling. This approach is based on minimizing the variance between the simulated and observed data through reducing the error of an objective function. Thus, model parameters are iteratively adjusted until an output error criterion is met (Nelson, 1960; Nelson, 1961; Neuman, 1973). This process requires running the forward model multiple of times that in turn leads to a great computational burden to solve the inverse problem. This gives a rise to the automatic calibration/inversion procedures over the manual trial-and-error techniques (Carrera et al., 2005; Poeter & Hill, 1997).
The inverse modeling computational burden can be alleviated using one of the following techniques: parallel computing (Tavakoli et al., 2013), realizations filtering (Scheidt & Caers, 2009; Ginsbourger et al., 2013), or/and replacing the forward model by less computationally expensive surrogate models (e.g., Reduced Order Models (RoMs)) (He et al., 2013). Unfortunately, the less computationally expensive models (i.e., RoMs, AAN, analytical solutions, linearized models) are not always adequate for simulating highly non-linear groundwater problems (i.e., contaminant transport in heterogeneous media). Thus, a non-linear indirect inverse modeling approach had to be developed to estimate and calibrate model parameters of such highly non-linear problems.

Different techniques and theories were developed to indirectly solve linear and non-linear inverse problems (Carrera et al., 2005; Carrera, 1988; McLaughlin & Townley, 1996; Franssen et al., 2009; Zhou et al., 2014). Examples of these techniques include: maximum likelihood method (MLM) (Carrera & Neuman, 1986), Bayesian-based methods (Woodbury & Ulrych, 1996; Jiang et al., 2004), gradual deformation technique (Hu, 2000), ensemble filtering (Nowak, 2009; Bailey & Baù, 2011), GLUE (Beven & Binley, 1992), and self-calibrating sequential methods (SC) (Gómez-Hernández et al., 2003). Interestingly, Franssen et al., (2009) explored seven different inverse techniques and concluded that the parametrization methods may differ among the tested techniques but their overall performance is relatively close (i.e., the predictive accuracy).

Of a practical interest to this study, the mathematical foundation of the MLM method is discussed. MLM is one of the most commonly employed techniques under inverse modeling codes (i.e., PEST) for solving non-linear inverse problems. The method is capable of simultaneously optimizing several parameters through maximizing the likelihood function expressed in Equation 5 or minimizing its exponent term as described in Equation 6 (i.e., objective function). $L(x|y^{obs})$
is the likelihood of the estimated parameter $x$ (e.g., $K$ or $D$) conditioned by observed data $y^{obs}$ (e.g., $\tilde{C}$ and $\tilde{h}$), $y_i$ is a vector of model outputs at the locations of observed data ($y^{obs}$), $(y_i - y^{obs})$ is a vector of residuals between estimated and observed data, $N_m$ is the number of observations, and $C_{y,i}$ is the covariance matrix between measurement uncertainties and model-to-data misfit errors.

The likelihood function in Equation 5 is derived based on the assumption that all datasets and parameters can jointly form multi-Gaussian spaces. An iterative search algorithm, such as gradient descent, Gauss-Newton, or Marquardt Levenberg methods (Marquardt, 1963; Madsen et al., 2004), is then needed to find the minimum of Equation 6 and meet a predefined convergence criterion. The objective function, represented by Equation 6, can be more constrained by adding a \textit{regularization term} as expressed in Equation 7. This term constraints the minimization process of the objective function on specific prior information about model parameters, which helps in well-posing the inverse problem with providing stability and uniqueness to the inverse problem solution (Zhou et al., 2014).

\begin{equation}
L(x|y^{obs}) \propto \exp\left\{-\frac{1}{2} \sum_{i=1}^{N_m} (y_i - y^{obs})^T C_{y,i}^{-1} (y_i - y^{obs})\right\}
\end{equation}

\begin{equation}
J = \sum_{i=1}^{N_m} (y_i - y^{obs})^T C_{y,i}^{-1} (y_i - y^{obs})
\end{equation}

\begin{equation}
J = \sum_{i=1}^{N_m} (y_i - y^{obs})^T C_{y,i}^{-1} (y_i - y^{obs}) + (x - x^{pri})^T C_x^{-1} (x - x^{pri})
\end{equation}

In Equation 7, $x^{pri}$ is the priori guess of the estimated parameter $x$, $C_x$ is the covariance matrix between the expected uncertainties in the prior model parameters. Equations 5-7 can be derived using Bayesian theorem in the framework of probabilistic inversion (Aster et al., 2019).
The multi-Gaussian assumption for the model parameter space in the MLM method imposes a continual smooth description of the system properties (e.g., \( K \)) that may not be valid for cases with extreme variability in the geological settings and structures. For instance, geological features like fluvial deposits or fossil buried channels, where a high contrast in the \( K_s \) can be found between the involved aquifers will be smoothed out (Journel & Deutsch, 1993; Gómez-Hernández & Wen, 1998). Such critical geological features to model predictions will be unreasonably adjusted to compensate model structure errors and reproduce field observations (e.g., \( \bar{C} \) and \( \bar{h} \)). Thus, such multi-Gaussian assumption can be a source of frequent error in the inverse problem results (Carrera et al., 2005). To resolve this issue, some studies proposed and discussed non-Gaussian-based models to represent target parameters in the inverse problem to enhance mapping the \( K \)-field (Kerrou et al., 2008; Rubin & Journel, 1991; Zinn & Harvey, 2003). However, Rubin et al., (2010) pointed out that constraining target parameters on any sort of predefined geostatistical structure, limits the degree of data assimilation between observations and parameters.

The appropriate representation of model parameters by either a single property map or a set of likely possible realizations is a controversial issue between scientists and decision-makers since decades. The unique representation of system properties will result in a single prediction by the forward model, which may or may not be close to the “true” future behavior of model outputs (i.e., contaminant fate). The inherent uncertainty in observation data and the existence of non-inferable parameters in the system entail large number of possible solutions of model outputs/predictions that cannot be explored using a unique representation of system properties. Such unexplored possible model outputs/predictions can include critical and risky scenarios of the future system behavior. Thus, it is more conservative to display all possible model predictions as
a probabilistic expression, so that model predictive uncertainty can be considered while taking management-related decisions.

Predictive uncertainties are not assuredly reduced by model calibration or history matching, because of the insensitivity of model predictions to target parameters (Doherty, 2015). From a philosophical perspective, calibrated models “cannot predict what will happen in the future. However, following proper uncertainty analysis, it may be able to predict what will NOT happen in the future” (Doherty, 2015). Such statement endorses the concept of model falsification introduced by Tarantola (2006). The falsification concept suggests that observed data should not be used by inverse modeling to find a unique solution for model parameters. Instead, it should be used to falsify unacceptable solutions from a large set of possible parameter distributions developed using prior information and expert knowledge about the system. Thus, we can ensure that the ensemble of model predictions, generated based on the residual acceptable parameter distributions, include all likely possible results necessary for taking risk-related decisions.

2.7.1 Model Parametrization

One option to obtain a well-posed and stable inverse problem is to reduce the number of unknown model parameters (i.e., \(x\) in Equation 5) compared to available data density as discussed before. This can be done through describing the real-world complexity and heterogeneity using limited number of parameters that are adjustable through history matching or calibration process. Number of studies have reported a direct relationship between the degree of model parametrization (i.e., degree of model complexity) in the inverse problem, and the biasness in the estimated parameters (e.g., \(K\) or \(D\)), and model predictions (e.g., \(\bar{C}\) and \(\bar{h}\)) (Cooley, 2004; Cooley & Christensen, 2006; Gallagher & Doherty, 2007b; Yoon & McKenna, 2012; Kowalsky et al., 2012; Knowling et al., 2019). The level of biasness in model predictions control the correctness of the
decisions taken based on these predictions (Gosses & Wöhling, 2019). Studying the adequate degree of complexity in environmental models to represent the real-world heterogeneity have received a great attention from the scientific community (Castilla-Rho, 2017; Simmons & Hunt, 2012). Groundwater modeling practitioners are required to select an appropriate level of parametrization - called defensible model complexity (Guthke, 2017) - so that they balance between incorporating all necessary system details to adequately reproduce field conditions and the computational burden resulted from this level of complexity (Knowling et al., 2019). Some guidelines can be found in Doherty & Simmons (2013) and Ferré (2017).

The parameterization degree can be classified as coarse, moderate, and high with respect to the number of considered parameters in the inverse problem (i.e., ≤10, 10-100K, and ≥100K, respectively) (Knowling et al., 2019). Hunt (2007) argued that highly parametrized models are more expressive in representing the uncertainty of model inputs (i.e., $K$ or $D$) and provide an enhanced assimilation between model parameters and observed data, which improves the accuracy of model predictions. Moreover, Knowling et al. (2019) recommended applying as fine-parametrization as possible to ensure numerical stability. They added that an in-depth understanding of the inverse problem theory and powerful computational resources are required to efficiently implement the high-dimensional parameterization techniques. The best practices for using this technique were discussed by Doherty & Hunt (2010) and Doherty et al. (2010a; 2010b).

In counter to Hunt’s argument, Voss (2011a; Voss, 2011b) claimed that low-dimensional parametrization allows for capturing the most important processes of the modeled system, which underpins the aim of gaining a quick overview about the most influential parameters in the system. Such knowledge is extremely important for site-characterization and model development, particularly under the fact that it can be impossible to reproduce the complexity and heterogeneity
of the natural subsurface environments in a numerical model with high accuracy. On the other hand, using simplified models (i.e., applying zonal parametrization) to represent complex subsurface environments can introduce high level of biasness in model outputs (e.g., $\bar{C}$ and $\bar{h}$). Such biasness can even be redoubled through history matching, as the optimization code will compensate the forward model deficits (i.e., unrepresented features/heterogeneity) by estimating biased parameters to fit observed data (Clark & Vrugt, 2006; White et al., 2014). Unknowable magnitude of biasness in the estimated parameters leads to inestimable uncertainty of model predictions, which will not allow to quantify the risk associated with the taken decisions based on these predictions (Knowling et al., 2019). Many examples are available in the literature for using the reduced parametrization techniques (Ala-aho et al., 2017; Sun et al., 2011), while other researchers had the propensity to adapt the highly parametrization methods (Cui et al., 2018; Knowling et al., 2015).

Parameterization Techniques:

One of the early developed techniques based on the reduced parameterization concept is the zonation method. Zonal parameterization was classically adapted by Carrera & Neuman (1986) to reduce the number of unknowns and control the ill-posedness of the inverse problem. For applying such method, the model domain is divided into discrete number of regions, where spatially uniform properties are assumed (Neuman, 1973; Hill, 2006; Sun & Yeh, 1985; Eppstein & Dougherty, 1996) or predefined geostatistical structure of aquifer properties is imposed (Clifton & Neuman, 1982). Zonal information, such as shape and location, can be retrieved from the data of geophysical surveys, borehole logs, or hydrogeological tests (Kowalsky et al., 2012).

Zonal parameterization can lead to biased estimated parameters because of one or more of the following reasons: (1) calibrated parameter field (i.e., $\mathbf{K}$ and $\mathbf{D}$) is whether a single smooth
map or discontinued interfaces, which may not be geologically justified (Zhou et al., 2014), (2) small errors in the zone shape limits the model capability to reproduce observed data (Kowalsky et al., 2012), (3) estimated parameters can be highly sensitive to errors in model structure and BCs, and (4) ignoring the micro-processes occurring at fine-scale heterogeneity can impact the accuracy of transport modelling (Knowling et al., 2019; Klise et al., 2008).

Yoon et al. (2012) and Barth et al. (2001b) reported that despite the explicit representation of synthetic aquifers’ heterogeneities in models, observed experimental data were not perfectly reproduced after calibration due to the uncaptured micro-processes at fine-scale heterogeneity (Yoon & McKenna, 2012; Barth et al., 2001b). Lately, Knowling et al. (2019) provided some guidelines to mitigate the ill-effects resulted from applying coarse parametrization in environmental models; these guidelines included: (1) wide range of parameter bounds should be assigned during history matching, (2) predictions should not be sensitive to micro-scale processes, and (3) model outputs should be in the form of quantity changes (e.g., ΔC) rather than absolute values (e.g., C).

In recognition to the limitations of zonation method, De Marsily et al., (1984) developed the Pilot Points Method (PPM), which is a high-dimensional parameterization technique. In this method, the available data (e.g., $\mathbf{K}$ or $\mathbf{D}$) are interpolated using kriging to produce a property base-map, where fictitious points (pilot point) are introduced at free-data locations. The property value at these fictitious points is estimated by minimizing the objective function of the inverse problem. A new map that incorporates the optimized property values (i.e., at the pilot points) is then developed. Such map represent a better assimilation between observed data and model parameters. These steps are repeated in an iterative framework until an acceptable fit between model outputs...
and observed data is accomplished (De Marsily et al., 1984). PPM considers the spatial correlation structure of the property values during the kriging interpolation process (Kowalsky et al., 2012).

The advantages of the PPM could be summarized in; (1) providing more flexibility in space discretization, (2) resulting in property maps (i.e., equally likely realizations) that honor the same geostatistical structure, (3) optimizing property values within a predefined margin of variability (Zhou et al., 2014; Cooley, 2000; Cooley & Hill, 2000), and (4) entailing accuracy tolerance in the assumed spatial correlation structure of the property data (Kowalsky et al., 2012). However, PPM method can also result in biased estimates of model parameters due to the smooth map of the property field generated by kriging interpolation, which may not be geologically justified. Different attempts of improvement can be found in the literature to resolve the PPM ill-effects (GÃomez-HernÃandez et al., 1997; Doherty, 2003; Kowalsky et al., 2004; 2005; Alcolea et al., 2006; Finsterle & Kowalsky, 2008; Rubin et al., 2010).

2.7.2 Impact of Sparse Data on Inversion Results

Although, inverse modeling started to be a recognized tool for aquifer characterization at regions where accurate geological data in not available (Tian-chyi et al., 2015; Poeter & Hill, 1997), estimated properties using inversion techniques can be highly biased as discussed before. Carrera & Neuman (1986) stated that “There is a general consensus among groundwater modelers that the inverse problem may, at times, result in meaningless solutions. However, the reasons for this “misbehavior” of the inverse solution are not always well understood: some hydrologists attribute them to non-uniqueness, some to non-identifiability, and others to instability”. Tian-chyi et al. (2015) set two conditions to classify the inverse problems addressing heterogeneous systems well-defined: (1) the model should be constrained on specified flux BCs or sources and sinks and (2) transient head data should be available everywhere across the domain. Since such conditions
are difficult to meet given the typical sparse field-data, sensitive models to subsurface heterogeneity (e.g., contaminant transport models) will always lead to ill-defined inverse problems. This gives a rise to the necessity of designing the monitoring systems carefully to collect the most informative data to model predictions (Zhou et al., 2014; McLaughlin & Townley, 1996; Knopman & Voss, 1988; Tian-chyi et al., 2015).

Data sensitivity to model parameters plays an essential role in determining the plausibility of estimating system properties via solving inverse problems. Carrera et al., (2005) and McLaughlin & Townley (1996) reported that the common case of having unsensitive hydraulic data to different areas in the K-field can be the main challenge in solving the hydrogeological inverse problems. Therefore, McLaughlin & Townley (1996), Franssen et al., (2003), Schwede & Cirpka (2009), and Yoon & McKenna (2012) recommended integrating different data types in the inverse problem (i.e., concentration data with hydraulic heads) instead of relying on a single high-resolution dataset (e.g., head or concentration data) to improve the overall data sensitivity and reduce solution space dimensionality. The challenge of using concentration measurements (i.e., brine leakage) as a standalone data-source, is the extremely low sensitivity and identifiability of parameters located outside the plume cloud to concentration data. Consequently, only a very narrow fringe from the target property field (e.g., K or D) can be reasonably estimated using concentration data (Schwede & Cirpka, 2009).

Kowalsky et al., (2012) highlighted that, errors in the transport model structure can lead to extremely biased estimated parameters, which will result in inaccurate predictions, even if the model fit was improved. They also added that the inverse problem domain should be extended beyond the area of interest, as the heterogeneity outside this area plays a critical role in controlling the plume migration within it. Schwede & Cirpka (2009) showed how low spatial-density
concentration data (i.e., discontinues spatial BTCs) can mislead the optimization algorithm resulting in poor predictions; with low-density data the calibration process can terminate before a reasonable prediction of the entire plume is obtained because a well-fit to the sparse/discrete data points is accomplished early in the optimization process without capturing the actual shape and location of the plume, as shown in Figure 2.6.

Figure 2.6 Transverse concentration profile in the plume; Solid line represents the simulated normalized concentrations, crosses represent the discrete concentration measurements, and the arrows represent the margin of error in the predicted plume location (Schwede & Cirpka, 2009).

2.8 Research Hypotheses

Based on the reviewed literature above and the EPA requirements to mitigate the potential adverse impacts of CGS operations, a set of testable hypotheses were formulated as follows to derive this research:

I. The inability to accurately capture the early behavior of brine plume development in the vicinity of the leakage location, due to the uncertainties associated with source conditions,
reduces the predication reliability of brine plume migration pathways and shallow aquifer contamination risks,

II. Monitoring systems for brine leakage detection and contamination management can be developed by integrating less-expensive shallow observation wells with deep sensors in an optimal manner to reduce the overall costs associated with the placement, operation, and maintenance of the observation points,

III. Locations and pumping rates of the extraction wells, designed to control brine leakage, can be determined using optimization approaches to satisfy the regulatory constraints on the water quality in the shallow aquifer by capitalizing on the attenuation capacity of the overburden formations through natural dilution.

2.9 Study Scope & Specific Objectives

The goal of this study is to contribute towards the overall management of the potential brine leakage from CO2 storage formations through satisfying a set of safety-related requirements from the EPA to protect shallow aquifers. This study focuses mainly on far-field brine leakage (see Figure 1.2) because distant areas from the injection site pose higher leakage risk compared to those located in proximate ranges, where the system is well-characterized to avoid operational failures. High-resolution investigation of caprock integrity may not be available at the remote zones from the injection point due to the inaccurate delineation of the AoR using uncertain field information. Moreover, given the extent and speed of native brine migration through overlying formations, brine leakage can act as an early warning of CO2 leakage (Sun et al., 2013a; Réveillère & Rohmer, 2011) and as an indicator of the leakage source and migration pathway. The study focuses on brine leakage rather than the leakage of gaseous or aqueous phases of CO2.
Due to safety concerns and cost constraints, it was not practical to test the above hypotheses in the field. Thus, this study was based on intermediate-scale experiments in an ~8m long soil tank where brine leakage from a storage zone to a shallow aquifer was simulated. Numerical modeling was also found to be essential in this study to explore multiple uncertainty scenarios of source conditions expected in field settings (Hypothesis 1) and to design both brine leakage monitoring and control systems (Hypotheses 2 and 3) that were both tested using the soil tank. For simulating far-field brine leakage, Cihan et al., (2013) proved the suitability of using a single-phase flow model. Therefore, the experimental and numerical components of this study do not include multi-phase simulations of the CO₂ plume dynamics.

A list of specific objectives was set to enable the laboratory experimentation, model validation, optimization methods’ development and hypotheses testing presented in this dissertation. These objectives are:

a) Constructing an intermediate-scale, porous-media, test tank, where brine migration from a storage formation to a shallow aquifer through multiple intermediate layers is simulated, to collect high-resolution data on brine leakage plume propagation,

b) Validating the capability of a numerical model, based on the transport code FEFLOW (Diersch, 2014), to capture the brine leakage plume migration in a multidimensional flow field of a complex geological settings mimicking field conditions,

c) Investigating the impact of uncertainty in source conditions on the accuracy of predicting brine leakage pathways and plume distribution,

d) Evaluating the use of the Genetic Algorithm (GA) and linear uncertainty analysis tools to design the monitoring system of CGS operations with the least number of deep sensors to minimize the monitoring cost,
e) Testing the applicability of a design approach that optimizes the placement and extraction rates of deep wells to be used as an emergency control system for brine leakage to protect the shallow aquifers against contamination.
CHAPTER 3

METHODS AND MATERIALS

This chapter describes in detail the experimental and numerical approaches that were used in this study\(^1\).

3.1 Experimental Approach

Exploring the impacts of uncertainty in source conditions on the accuracy of brine leakage predictions (Paper I) needs a validated transport model for simulating brine leakage migration under different uncertainty scenarios of source conditions. High-resolution data on the system hydrogeological settings, source conditions and the spatiotemporal development of the brine plume is necessary for this validation. Moreover, testing the performance of brine leakage monitoring system and validate its design approach (Paper II), requires high-resolution data to be collected during a leakage event. This data is essential to evaluate the monitoring system leakage detectability and its informativity to the predictions of interest. Finally, to assess the efficacy of a brine leakage control system and the applicability of its design approach (Paper III), data should be gathered from the overlaying formations of the storage zone during a leakage event. Therefore, the research of this dissertation requires high-resolution data to be collected on the leakage

migration from the CO$_2$ storage zone to the shallow aquifer under various scenarios to support different investigations.

As discussed in Section 2.9, it is technically, practically, and financially infeasible to initiate a brine leakage event at a real CGS site to collect the required data for this study. Keating et al. (2010) suggested three potential research strategies other than field testing to develop CO$_2$/brine leakage related studies: (1) using data from engineered and well-characterized CGS pilot sites, such as Frio or Ketzin sites (Doughty et al., 2008; Kharaka et al., 2006; 2009; 2013; Nowak et al., 2013; Zeidouni et al., 2014), (2) generating data under controlled laboratory experiments (Luyun Jr et al., 2011; Trevisan et al., 2017; Agartan 2015; Plampin et al., 2017; Solovský et al., 2019) and (3) utilizing the data from analog systems to CGS operations, such as leakage of re-injected produced water (Thamke & Craig, 1997; Thamke & Midtlyng, 2003; Jacobs, 2009) or natural brine leakage from deep formations (Keating et al., 2010; 2013a; 2013b; 2014; Llewellyn, 2014).

Given that the objective of engineered pilot sites is to prevent leakage to examine CO$_2$ trapping under field conditions and the availability of high-resolution spatiotemporal data from CGS analog sites is scarce, this study uses a laboratory-based approach. Lenhard et al., (1995) and Oostrom (2007) pointed out that intermediate-scale experiments can reproduce field-scale groundwater flow processes under laboratory settings while keeping initial and boundary conditions under full control. Several researchers have adopted this approach to explore various aspects of CGS (Luyun Jr et al., 2011; Agartan, 2015; Trevisan et al., 2017; Plampin et al., 2017; Solovský et al., 2019). Therefore, an intermediate-scale physical model was developed for the research of this study.
Since brine leakage development and migration are site-specific problems that vary with the hydraulic and geological conditions of every CO₂ storage site, the developed physical model was not intended to reproduce the complexity of a specific CGS site. However, the goal of the experimental work was to capture some of the important processes in the field to provide a much more reasonable way to validate a model for conducting theoretical analyses and test the developed tools for designing the leakage monitoring and controlling systems.

3.1.1 Conceptual Model Development

To design an intermediate-scale physical model that simulates brine leakage under field-like conditions, a generic conceptual model for far-field brine leakage was first developed based on the important factors deriving and controlling the leakage process. This conceptual model was basically synthesized from a number of field studies of brine leakage in CGS analog systems [e.g., Keating et al., (2010; 2013a; 2013b; 2014), Llewellyn (2014), Jacobs (2009), Thamke & Craig (1997), and Thamke & Midtlyng, (2003)]. The developed conceptual model, presented in Figure 1.2, reflects the typical settings of a brine leakage problem; the model included multiple, hydraulically connected aquifers with complex interacting multidimensional flow fields that affects the plume development and travel times. These flow fields are controlled simultaneously by different boundary conditions, such as regional (shallow and deep) hydraulic gradients, local fracture flows, and the overall pressure perturbations associated with the active phase of CO₂ injection.

In this conceptual model, a far-field brine plume is leaking from a CO₂ storage formation (denoted as Zone 1) through fractured caprock, that in turn migrating via advection and dispersion across multiple intermediate overburden layers (denoted as Zone 2) and eventually propagating in the shallow aquifer (denoted as Zone 3), as shown in Figure 3.1a. The advective flux is driven
primarily by the high pressure transmitted from the storage formation during the leakage event, which creates a vertical hydraulic gradient causing the plume upward migration with minimal density effects in the vicinity of the source. The further migration of the plume through Zone 2 will be slightly retarded because of density effects (Wunsch et al., 2013), which will become less significant with distance from the leakage source as the plume dilutes. In this scenario, the leakage plume is transmitted through a steady flow field created by the head gradient and superposed over the existing regional groundwater flow. Close to the source, the velocity filed is dominated by the fracture flow. While the flow field in the upper layers, further away from the source, is a result from both fracture and regional flows.

Figure 3.1 Diagram describing (a) the developed conceptual mode and (b) the 90° transformation of the natural upwelling leakage scenario in the horizontal soil tank.
Natural systems display depth-dependent temperature variations, which from a modeling perspective, requires that variable-density flow be coupled with both heat flow and brine transport (Zhang et al., 2005). Following Birkholzer et al. (2011a), this study assumes isothermal conditions given the opposing effects that temperature and pressure exert on the brine density at varying depths. Moreover, as the multiphase flow dynamics induced by CO₂ injection have a negligible impact on the pressure response of far-field native brine (Cihan et al., 2013; Bandilla et al., 2015), single-phase flow conditions were assumed in this study. In natural systems, flow within Zones 1 and 3 is relatively horizontal and controlled by local and regional gradients. Flow within Zone 2 in the vicinity of a leak is predominantly vertical as it is controlled by the overall head difference between Zones 1 and 3 as discussed above. For simplification, numerical studies related to CGS applications and other deep injection activities assumed negligible regional flow in the upper layers to simulate the fluid migration from deep formations to shallow aquifers (e.g., Oldenburg and Unger, 2003; Pan et al., 2013; Gassiat et al., 2013). The same assumption was made in this study due to practical constraints on assigning regional flow to the Zone 2, representing the overlaying formations, in the soil tank.

3.1.2 Experiment System Design

Designing an experiment for understanding fundamental processes about brine leakage and generate high-resolution data on the leakage development is challenging because it requires; (1) the experiment to be able to capture the broad range of time and length scales of the CGS related processes as brine leakage, (2) detailed knowledge about the subsurface geological settings, source conditions, and flow fields, and (3) accurate data with high spatiotemporal resolutions on the development and migration of brine leakage. In this study, many of these challenges were addressed through its reliance on the concept of intermediate-scale testing.
Intermediate-scale testing is performed with spatial length scales that are intermediary to the traditional column-scale and field scale, allowing scale-dependent natural phenomena to be reproduced in a laboratory setting. In this study the strata in natural geologic basins and formations, 1-2 km in depth, was downscaled significantly to the order of 10 meters to enable laboratory experimentation. Moreover, the heterogeneity and layering associated with this problem were explicitly defined using geostatistical methods. Meanwhile the multidimensional flow fields expected in this problem were prescribed and maintained via multiple boundary conditions. By this, the natural system complexities were relatively reproduced under the laboratory settings with efficient imitation of the multiscale interactive processes.

Despite downscaling this CGS problem to the laboratory setting, it still offers a significant space requirement challenge. Instead of creating a large vertical soil tank to conduct this work, a novel approach was adopted in which a long horizontal tank was used to simulate the brine leakage problem, essentially rotating the overall leakage scenario 90° (Figure 3.1b). The soil tank was packed in a series of vertical layers analogous to the horizontal layering of the natural sedimentary formations. An impermeable “caprock” containing fractures was located between Zones 1 and 2. A vertical flow gradient was established at both ends of the soil tank (i.e., horizontal flow in natural system) to represent the regional/local gradients controlling the horizontal groundwater flow. Horizontal flow (i.e., vertical flow in natural system) across Zone 2 was induced in the experiments by adjusting the head difference between Zone 1 and Zone 3 (Figure 3.2). The native brine was assumed to be uniform throughout the storage formation and therefore no transport resulting from concentration gradients is possible in Zone 1.

A dyed bromide tracer was used as a surrogate for the brine found in natural formations to track the plume propagation through the transparent side of the tank. The brine surrogate was
injected in the outflow of the fractures to simulate brine leakage at sufficiently low concentrations so as to prevent creating a large density contrast between the NaBr tracer and background water that can lead to plume sinking – an important consideration given that the direction of gravity is different in the experimental setup than in the field setting.

### 3.1.3 Soil Tank Settings

Experiments were performed in the soil tank shown schematically in Figure 3.2. The total internal dimensions of the tank were 800 cm × 123 cm × 6.5-8.0 cm (length × height × width). Tank sections with materials presented in Figure 3.2 were connected via a quarter cylindrical aluminum tank, creating the L-shape depicted in Figure 3.2; this was a necessary design feature given the space constraints in the laboratory. The low flow velocities (Reynold’s number < 1.0) ensured negligible head loss along the cylindrical portion where the flow was turned 90°. The top of the soil tank was sealed with a flat wooden plate so as to create a no-flow boundary.

The use of glass walls in the soil tank allowed the dyed tracer to be visually tracked using digital photography during the experiments. This facilitated determining the general boundaries of the plume as it spreads and helped informing sampling strategies. A screened section of pea gravel, 10-cm thick, was placed at the top and bottoms of either ends of the soil tank to control the head gradients (Figure 3.2); in the experiments, either constant-head reservoirs or peristaltic pumps (Cole-Parmer Masterflex L/S: 7526-60) were connected to these sections to control the head level or fluid flux at the tank boundaries, respectively.

One wall of the acrylic section of the tank, containing Zone 3, had a 16 × 30 rectangular grid with 7.0 cm wide squares. A self-sealing septa port was installed at each intersecting point of the grid for high-resolution (480 individual points) aqueous sampling with syringes (Figure 3.7a).
Initially, there were 175 septa ports installed at a much coarser resolution in Zones 1 and 2. However, later, extra 40 septa ports were added to Zone 2 to give more flexibility in selecting the best monitoring locations at this zone, as will be discussed in Chapter 5.

Figure 3.2 Depiction of (a) the soil tank, (b) instrumentation layout and packing configurations. All dimensions shown in centimeters. The numbering scheme for the injection ports, instrumentation, and sampling ports increases sequentially from top to bottom and right to left. Note, flow occurs from right to left. All dimensions are in centimeters.
Initially, a total of 30 manometers were connected to hollow-brass-tubes installed across the tank with the highest spatial distribution near the brine injection point so as to detect pressure changes due to tracer injection. Then, 17 hollow-brass-tubes were placed in Zone 1 to increase the number of potential deep extraction wells/tubes for the selection process of the most optimal extraction location to control the leakage plume in the tank, as will be discussed in Chapter 6.

A total of 21 electrical conductivity sensors (5TE-Decagon Devices) were installed in Zones 1 and 2 to automatically track plume propagation. A 2 cm acrylic thick sheet with a width equal to that of the internal width of the soil tank was installed at an angle of 76° to represent an impermeable caprock. Prior to its installation, six 4 cm holes were drilled in the sheet (Figure 3.2) to simulate fractures within the damaged zone of a caprock. The inclined confining sheet was modeled after the upward dipping caprock (Freeman Shale) of the potential CO₂ storage (Vedder formation) in Southern San Joaquin Basin, California (Cihan et al., 2015). This inclination imposed different fluid fluxes at the inflow and outflow boundaries of Zone 1, which in-turn created a complex flow field at this zone and in the source vicinity. Six external injection lines were similarly installed at a distance of 5 cm from the fractures/holes to inject the bromide tracer at distinct rates and locations during the experiments.

3.1.4 Porous Media Configuration

The soil tank was packed with six well-characterized manufactured silica sands, with the properties presented in Table 3.1, to create the layered system shown in Figure 3.2. The packing configurations applied in Zones 1 and 3 were designed to resemble natural sedimentary strata with correlated lognormal hydraulic conductivity (lnK) fields. The approach used to develop the discrete heterogeneity fields was designed to ensure that predefined geostatistical structures are honored (Barth et al., 2001b; Trevisan et al., 2017). To design the discrete heterogeneous lnK-
fields of Zones 1 and 3, two sets of silica sands were used to approximate two target lnK normal distributions mimicking natural sedimentary conditions (Figures 3.3a and 3.3b).

Figure 3.3 (a, b) Target lnK distributions (dashed lines) overlaying histograms of the five sands used for packing the selected realizations for Zones 1 and 3 (lnK); (c, d) Semi-variogram plots showing the selected experimental (black solid-line) and theoretical variograms (red solid-line) compiled with seven other realizations that also exhibited smooth variograms (gray dashed lines). It should be noted that “Real” stands for realization.

The sequential indicator simulator algorithm in SGeMS (Remy et al., 2009; Deutsch, 1998) was used to generate two sets of 200 unconditional realizations with correlated lnK fields for Zones 1 and 3. The final packing configuration was selected for each zone based on a two-stage selection criterion. First, realizations exhibiting normal distributions of the lnK-fields and smooth semi-variograms were drawn from the full set. These realizations were then tested in a flow and transport
model to identify which configuration provides the most nonuniform streamlines in Zone 1 and widest distributed plume in Zone 3. Figures 3.3b and 3.3c present the experimental and theoretical variograms for the selected realizations for packing Zones 1 and 3, compiled with seven other realizations that also exhibited smooth variograms for the sake of comparison. The selected discrete configurations for packing were found to reasonably approximate the target continuous distributions of \( \ln K \) designed for Zones 1 and 3, as shown in Figures 3.3a and 3.3b.


<table>
<thead>
<tr>
<th>Sand no.</th>
<th>#110</th>
<th>#70</th>
<th>#40/50</th>
<th>#20/30</th>
<th>#20</th>
<th>#16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Porosity ( [m^3 m^{-3}] )</td>
<td>0.340</td>
<td>0.413</td>
<td>0.334</td>
<td>0.406</td>
<td>0.410</td>
<td>0.397</td>
</tr>
<tr>
<td>( d_{50} [mm] )</td>
<td>0.103</td>
<td>0.2</td>
<td>0.35</td>
<td>0.54</td>
<td>0.70</td>
<td>0.88</td>
</tr>
<tr>
<td>Dry Bulk Density ( [g/cm^3] )</td>
<td>1.75</td>
<td>1.56</td>
<td>1.55</td>
<td>1.57</td>
<td>1.56</td>
<td>1.60</td>
</tr>
<tr>
<td>Saturated ( K^4 [m/day] )</td>
<td>5.52</td>
<td>12.10</td>
<td>44.93</td>
<td>101.95</td>
<td>181.44</td>
<td>476.93</td>
</tr>
<tr>
<td>( \ln(K) [-] )</td>
<td>1.70</td>
<td>2.5</td>
<td>3.80</td>
<td>4.6</td>
<td>5.20</td>
<td>6.2</td>
</tr>
</tbody>
</table>

\(^{1}\text{US Silica Co.}\)
\(^{2}\text{Unimin, Corp. (Granusil Sand)}\)
\(^{3}\text{Unimin, Corp. (Accusand Sand)}\)
\(^{4}\text{K values are based on column-scale laboratory testing}\)

The mean of \( \ln K \) distribution (\( \mu_{\ln K} \)) of the deep storage formation (Zone 1) was assumed to be lower than the shallow aquifer (Zone 3) due to soil compaction effects. Thus, the five designated sand categories for packing Zone 1 (sand #20 through #110) have \( \mu_{\ln K} \) lower than Zone 3 (sand #16 through #70). The heterogeneity of Zone 1 had a \( \ln K \) normal distribution with a mean (\( \mu_{\ln K} \)) of 3.8, variance (\( \sigma^2_{\ln K} \)) of 1.1, and correlation scale (\( \lambda_{Hz}/\lambda_{V} \)) of 3.16. While, Zone 3 had a \( \mu_{\ln K} \), \( \sigma^2_{\ln K} \), and \( \lambda_{Hz}/\lambda_{V} \) values of 4.6, 2.0, and 1.6, respectively. These geostatistical parameters were comparable to field data at well-characterized sites (Garabedian et al., 1991; Hess et al., 1992; Rehfeldt et al., 1992). It should be noted that a part of Zone 1 located immediately next to the
inclined caprock (encircled by red line in the inset part of Figure 3.2) was homogeneously packed out of necessity given the limited access to this confined area in the tank.

In Zone 2, a nonstationary structure was assumed due to the extreme depth being emulated by this zone (>1.0 km) and the expected uncorrelated discrete geological layering associated with the natural sequential deposition and weathering conditions in this zone. Therefore, this Zone was packed variably with sand ranging from #20 to #70 (Figure 3.2). This packing configuration was inspired by the nearly perfect layered stratigraphy of Utsira Formation, a target CGS reservoir located at the Sleipner field in the North Sea (Torp & Gale, 2004; Hermanrud et al., 2009; Zweigel et al., 2004).

3.1.5 Packing Procedures

Before starting the packing process, the selected sand configurations were drawn on the glass walls of the soil tank. Vertical metal dividers (22 cm tall) were placed at the boundaries of each sand type to create individual cells. The total height of the soil tank (123 cm) was packed in six stages, where six horizontal sand layers with 20 cm height each were packed sequentially along the full length of the tank, as show in Figure 3.4.

The standard wet-packing procedures were followed to fill the tank; the water level was constantly maintained at a height greater than the sand surface to avoid any entrapment of air bubbles. The sand was added to the individual cells, formed by the dividers and tank walls, then tamped with a metal rod and vibrated gently by hammering the tank walls using a rubber mallet. Upon the completion of a layer the dividers were raised, and the hydraulic connection between the packed cells was restored. The mass of sand used to fill each cell was determined in advance to ensure uniformity of density. The depth of the sand was compared with the design every two layers
to monitor the media settlement so that the packing procedure could be adjusted as necessary. The final packed heterogenous sand in the three zones of tank is presented in Figure 3.5.

![Figure 3.4 The process of tank packing.](image)

![Figure 3.5 Projected view of the final packed heterogeneity in the three zones of the soil tank.](image)
3.1.6 Tracer Properties

Sodium bromide (NaBr), a conservative tracer under laboratory conditions (i.e., pH, temperature, and pressure), was selected as a surrogate for the leaking brine. The total volume of tracer solution, specific to each experiment, was prepared before the experiments and stored in a plastic drum, equipped with in-house developed stirring and cooling systems to maintain a consistent tracer concentration. The electrical conductivity of the stored tracer in the drum was monitored with a conductivity meter (OAKTON-CON 100 series) to identify when the concentration of the solution is uniform and ready to be injected. In the experiments, the NaBr tracer was doped with either a blue/food dye (Envision DYEBLU-GL) or florescence dye (ACROS-Organics 17324-5000) to visually track the plume migration from the transparent side of the tank and develop sampling strategies. The decision of using florescence dye in the later experiments was taken because a high retardation of the blue/food dye was observed compared to NaBr tracer transport.

The concentration of the NaBr solution was designed based on non-dimensional density instability relationships (Oostrom et al., 1992; Barth et al., 2001a). The results of this non-dimensional analysis were further tested and validated in density-dependent FEFLOW simulations to determine what concentration triggers the onset of plume instability. A hypothetical problem including the worst settings for triggering plume instability was designed for this analysis to determine the highest possible NaBr concentration can be used without density effects. According to Barth (2001a), the density-driven instability effects on plume migration can be maximized by minimizing both (1) the overall head gradient across the problem domain and (2) the variation of the local head gradients across the heterogeneity facies of the permeability field. Thus, the sand configuration of the developed problem was designed to include three homogeneous zones along
the tank length (with sand #40/50, sand #70, and sand #110) to guarantee a limited variation in the local head gradients. In addition, a mild overall head gradient (~0.006) was imposed using the hydraulic head BCs at both ends of the modeled tank. Figure 3.6a illustrates the settings of the used hypothetical problem in this analysis.

![Figure 3.6a](image)

Figure 3.6 (a) Schematic of the developed testing hypothetical problem showing the locations where the concentration profiles were drawn; (b) concentration profiles of the 10,000 ppm NaBr tracer under the vertical (solid lines) and horizontal (markers only) gravity directions; (c) concentration profiles of the 50 ppm NaBr tracer (solid lines and solid markers) and 100 ppm NaBr tracer (dashed lines and hollow markers); In this graph, all solid lines are for gravity in v-direction, markers are for gravity in Hz-direction and colors reflect the type of sand.

Plume instability was explored through comparing the vertical profiles of the plume concentrations at the centers of the three homogeneous zones under different gravitational directions (Figure 3.6a). Two scenarios of gravity directions were considered: (1) the normal downward gravity direction representing the experimental settings (denoted as Z-gravity), and (2)
negative X-axis gravity direction, towards the storage zone, representing natural field conditions (denoted as X-gravity). Identical concentration profiles under the two scenarios of gravity directions reflect negligible density effects on plume migration. In this analysis, NaBr concentrations of 10000, 1000, 500, 100, and 50 ppm, respectively were tested. Only the results of tracer concentrations of 10,000 ppm, 100 ppm, and 50 ppm are presented in Figures 3.6b and 3.6c. Results showed that using NaBr-solution concentration in a range of 50-100 ppm will create negligible density driven flow conditions for the plume moving in the horizontal direction in the tank. Therefore, the concentration of the NaBr tracer used in all the experiments of this study was less than 100 ppm but higher than 50 ppm to ensure that the collected samples from the tank will have detectable levels of Br⁻ at the plume fringe. Detailed discussions of the prepared tracer concentration for each experiment and preparation procedures are provided under Chapters 4, 5, and 6.

3.1.7 Sampling Strategy

Aqueous sampling was determined in real-time during the experiments by visually monitoring the plume location; sampling was performed within the plume to a 20 cm extent from its boundary as observed from the dye. A sampling volume of 0.5 ml was extracted from each targeted septa port with a syringe, as shown in Figure 3.7a. The small extraction volume ensured that the flow field was not disturbed by the extraction process. Given the total volume of tracer used during each experiment, the mass removed from the system during sampling can be considered negligible, roughly 0.5% or less of the total amount of injected tracer. The collected samples were stored in 0.9 ml vials, sealed with parafilms, and placed in a refrigerator for preservation. These samples were later chemically analyzed using an ion chromatography (IC)
system (Dionex Aquion combined with autosampler Dionex As-Dv in Figure 3.7b), as will be discussed in section 3.1.7.

![Figure 3.7](image)

**Figure 3.7** (a) Aqueous Sampling from the soil tank through the self-sealing septa ports and (b) the used IC for chemical analysis.

### 3.1.8 Concentration Measurements

The developed approach for the chemical analysis using the IC was rigorously based on the guidelines provided by Tartari et al., (1995), Hautman & Munch (1997), Christison (2016) and EPA (2016). The approach included the following steps: (1) standard samples with known Br concentrations were used to estimate the calibration curves of the IC through finding a relationship between the Br concentration and the corresponding area under the ionic EC peak (Figure 3.8). (2) The developed relationship was then reversely used to determine the unknown Br concentrations of the collected samples during the experiments (Figure 3.9). Nine calibration solutions (0.1, 0.5, 1, 5, 10, 25, 50, 75, 100 ppm of Br) were prepared by diluting stock standards with 1000 and 50 ppm of Br (Inorganic Ventures IV-STOCK-59/64) to develop the multipoint calibration curve for the IC before each run of the machine (Tartari et al., 1995). Table 3.2 presents the retention times, peak heights, areas under the peak curves of the Br standard solutions prepared for one of the IC analyses. The average retention time for Br\(^-\) according to this analysis was 7.66 min, and the
detection limits for this calibration curve was estimated to be equal to 0.1 ppm. The results of several reproducibility tests for the IC showed that an error range of ~1-2% can exist in the chemical analysis and at very rare occasions this error reached ~5%, which was considered acceptable based on the used NaBr tracer concentrations in this study (50-100 ppm). Following the recommendations by Tartari et al., (1995), a quadratic function was used to fit the calibration points (Figure 3.9). The regression coefficient of determination (denoted as $R^2$) of the developed calibration curve, in the analysis presented in Figure 3.9, was 99.9995%, which was similar to the obtained $R^2$ value in most of the performed IC runs.

![Figure 3.8 IC response for the 75 ppm standard solution.](image)

Table 3.2 Retention times, peak heights, areas under the peak curves for the used Br standard solutions.

<table>
<thead>
<tr>
<th>Standard Conc.</th>
<th>Retention Time (min)</th>
<th>Fitted Conc. (ppm)</th>
<th>Area (uS*min)</th>
<th>Height (uS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stan. 1 - 0.1ppm</td>
<td>7.661</td>
<td>0.1044</td>
<td>0.0062</td>
<td>0.02</td>
</tr>
<tr>
<td>Stan. 2 - 0.5ppm</td>
<td>7.667</td>
<td>0.4987</td>
<td>0.0333</td>
<td>0.11</td>
</tr>
<tr>
<td>Stan. 3 - 1.0ppm</td>
<td>7.671</td>
<td>0.988</td>
<td>0.0671</td>
<td>0.23</td>
</tr>
<tr>
<td>Stan. 4 -5ppm</td>
<td>7.664</td>
<td>4.8873</td>
<td>0.3384</td>
<td>1.18</td>
</tr>
<tr>
<td>Stan. 5 -10ppm</td>
<td>7.661</td>
<td>9.8976</td>
<td>0.6928</td>
<td>2.43</td>
</tr>
</tbody>
</table>
Table 3.3 Continued

<table>
<thead>
<tr>
<th>Standard Conc.</th>
<th>Retention Time (min)</th>
<th>Fitted Conc. (ppm)</th>
<th>Area (uS*min)</th>
<th>Height (uS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stan. 6 -25ppm</td>
<td>7.654</td>
<td>24.8361</td>
<td>1.7881</td>
<td>6.4</td>
</tr>
<tr>
<td>Stan. 7 -50ppm</td>
<td>7.651</td>
<td>50.6121</td>
<td>3.814</td>
<td>13.97</td>
</tr>
<tr>
<td>Stan. 8 -75ppm</td>
<td>7.641</td>
<td>74.9497</td>
<td>5.8848</td>
<td>21.94</td>
</tr>
<tr>
<td>Stan. 9 -100ppm</td>
<td>7.631</td>
<td>99.8151</td>
<td>8.159</td>
<td>30.75</td>
</tr>
</tbody>
</table>

Figure 3.9 IC calibration curve to calculate NaBr concentrations of unknown samples using area of the Br- peak.

Dual-stage data correction was considered in this study to increase the accuracy of the generated datasets. First correction process was made to fix the effects of samples’ evaporation in the autosampler; samples reside in the autosampler for more than 12 hours during the process of chemical analysis, which affects the accuracy of concentration data due to sample evaporation. To
resolve this, five dummy vials were filled with deionized water and placed in the autosampler during all performed IC runs, so that a linear relationship between the volume loss from the samples due to evaporation and the vial retention time in the autosampler can be developed (Figure 3.10). This relationship was then used to estimate the volume lost from each vial based on its retention time in the autosampler and accordingly a correction factor was applied. The second correction process was against the sample evaporation inside the fridge, where the same method was employed, but instead of using dummy vials, randomly selected actual samples were used as references.

Figure 3.10 Example of the developed correction relationships for vial evaporation.

### 3.1.9 Water Head and Flow Rate Observation

The total water head inside the experimental tank was quantified using a set of 30 manometers as shown in Figure 3.11a. The accuracy of this method was examined through connecting one of the manometer tubes to a static water column and evaluate whether the water
level in the tube is affected by any capillary rise (Figure 3.11b). The results of this analysis revealed that a minor error of ±0.5 mm may exist in the head readings.

The tank outflow was quantified by collecting the effluent water from the constant head devices (CHD) over fixed-time intervals and massing it out. Given the tap water density (997 kg/m³), evaluated in the lab, and the weight of collected effluent water, the tank outflow volumetric rate was calculated. For measuring the tank inflow, the weight of the pumped water to the CHD and the drained water from the CHD were measured over a specific time period and then subtracted from each other to calculate the net volumetric rate of the tank inflow using the lab-evaluated tap water density. The tracer injection flow rate was also quantified using the same concept.

Figure 3.11 (a) Three sets of manometers connected to the pressure ports in the tank and (b) the comparison between water levels in the manometers and the water column.

3.1.10 Performed Experiments

The performed experiments in the developed soil tank were time consuming and costly; setting-up one experiment can take 2 weeks of preparation, around 1-2 weeks of execution besides at least one month of data logging and IC analysis depending on the number of collected aqueous samples (Table 3.3). In total, two preliminary experiments and five final experiments were performed in this soil tank to generate the required data and conduct the experimental testing
related to the investigations covered under this study. The preliminary experiments were carried out to examine the developed and adapted methods besides assessing the accuracy of the measuring devices.

Table 3.4 Summary of the Performed Experiments.

<table>
<thead>
<tr>
<th>Experiment Number</th>
<th>Served Paper - Hypothesis</th>
<th>Experiment Start Date*</th>
<th>Experiment End Date*</th>
<th>Duration*</th>
<th>Collected Data</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Total Collected Data for the Study</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aqueous: 3,159 Head: 3,220 Flow: 277</td>
<td></td>
</tr>
</tbody>
</table>

*Experiment dates and durations include the preparation period, actual running time, and chemical analysis time.

Three out of the five final experiments were conducted to generate high-resolution spatiotemporal data for calibrating (using two datasets) and validating (using one dataset) a transport groundwater model that was then used to evaluate the impact of uncertainty in source conditions on the prediction of brine leakage migration in Chapter 4 (Paper I). The fourth experiment was mainly conducted to collect data from a designed monitoring system by the developed optimal design framework discussed and presented in Chapter 5 (Paper II). The collected data was then used to calibrate a groundwater flow and transport model to make a set of
target predictions, which were eventually compared to experimental measurements. Based on this comparison the applicability of the developed optimal design framework was assessed. The fifth experiment was conducted to examine the performance of a leakage control system designed by using the global optimizer Genetic Algorithm with a transport model, as will be presented in Chapter 6 (Paper III).

More details about the data collected during these experiments and applied experimental BCs are provided in Chapters 4, 5, and 6. The major challenge in performing these experiments was the sampling effort; as collecting the aqueous samples is tied to the plume propagation (Section 3.1.7), the sampling work was on a continuous pace over the 24 hours of the day for the entire experiments' durations.

3.2 Numerical Methods

A large number of transport codes use the conventional coupling of Darcy’s law (Darcy, 1856) with the ADE, based on the Fickian assumption (Bird et al., 1960; Taylor, 1953), to simulate solute transport. The ADE has been shown to be inadequate in some cases however, as it can fail to capture the transport processes due to scale-dependent dispersion coefficients, heavy tailing of the BTCs and directional/time-variant porosity (Willmann et al., 2008). This has led to modify the ADE to describe non-Fickian solute transport (Benson et al., 2000a; 2000b; Schumer et al., 2001; 2009; Zhang et al., 2007; 2008; 2009), which in natural systems, can be caused by the presence of complex geological settings and soil heterogeneity. In deep subsurface regimes (e.g., brine leakage problem), modeling the groundwater flow and transport is even more sophisticated due to the existence of many interplaying physical processes at multiple layers that creates multidimensional flow fields; these processes include fracture flow, density effects, varying material properties and
water saturation levels (Ramadas et al., 2015). Comprehensive prior information and field data are therefore required to validate models used for addressing problems located at deep system.

In the context of brine leakage modeling, Saiers & Barth (2012) and Cohen et al., (2013) questioned the applicability of a MODFLOW-based model (Zheng & Wang, 1999) used by Myers (2012) to evaluate the potential risk of formation brine leakage based on oversimplifying the system flow field and specifying unrealistic model boundary conditions – all of which could not be adequately validated with field data. Yamamoto et al. (2009) further noted that models simulating pressure field perturbations caused by CO2 injection (i.e., leakage onset) should include an accurate description of the shallow and deep subsurface system characteristics. The use of ADE-based codes that can account for density-dependent flow (e.g., MODFLOW/MT3DMS or FEFLOW) to simulate leaking brine migration across a multilayered system is predicated on their ability to capture complex multidimensional flow fields. Model validation requires extensive long-term field data collected at high spatiotemporal resolutions for sites whose geology has been well characterized (Tsang et al., 2008; Ramadas et al., 2015; Jiang, 2011). Given that deep subsurface datasets of this nature may not be available, the laboratory data generated in the well-controlled intermediate-scale experiments in this study were used to validate a FEFLOW transport model developed to simulate the leaking brine.

3.2.1 Forward Simulation

The flow and transport code FEFLOW (V.7.1), originally developed by Diersch (2014) and currently owned by the DHI group, was chosen to develop the contaminant flow and transport models that simulate the brine leakage in the laboratory and numerical experiments of this study. FEFLOW is a FEM-based code that discretizes the solution domain and solves the coupled groundwater flow and transport equations (i.e., ADE) in a continuous FEM scheme. To ensure
local mass conservation in FEFLOW, the continuous FEM is enhanced by local smoothing techniques and consistent velocity approximator (Diersch, 2014). FEFLOW is capable of solving the ADE, coupled with variable density flow effects, through modifying the Darcy flux vector in Equation 1 as follows:

\[ q = -K \left( \nabla \bar{h} + \chi e \right) \]  

(8)

where \((\chi)\) is the buoyancy coefficient including fluid density effects and \((e)\) is the gravitational unit vector. Such features make FEFLOW a suitable candidate for testing the hypotheses of this study and conduct the required theoretical analyses. However, FEFLOW was never validated to capture solute transport process through multidimensional flow fields under complex geological settings similar to the conditions in the brine leakage problem. Thus, an essential component of this study was the validation of the FEFLOW-based model using the high resolution data generated in the first three final laboratory experiments (Chapter 4). Otherwise, the findings of this study will be questionable similar to Myers (2012).

3.2.2 Model Calibration

FePEST, a graphical user interface (GUI) that links the open-source inversion algorithm of PEST (Parameter ESTimation) to FEFLOW, was used to optimize the forward model parameters to fit the observed data in this study. PEST is an independent non-linear parameter estimation algorithm that was developed by Doherty (1994; 2010; 2015). The algorithm relies mainly on minimizing an objective function of least-squares, defined by the weighted residuals between the simulated and observed data, for quasi-linear/linearized forward models (Zhou et al., 2014; Doherty, 2003). Least-squares objective function, or the maximum likelihood method, allows for estimating parameters in complex non-linear models (e.g., transport problems), as discussed in
Section 2.7. Equation 9 presents the parameter upgrade-function applied by PEST to adjust model parameters and accordingly minimize the data-to-model misfit.

\[
p - p_0 = (J^TWJ)^{-1}J^Wr \tag{9}
\]

where \((p - p_0)\) is the upgrade vector from the previous parameters \((p_0)\) to the new parameter set \((p)\), \((r)\) is the objective function residuals using the previous parameter set \((p_0)\), \((W)\) is the weighting matrix, and \((J)\) is the Jacobian matrix. The Jacobian matrix (i.e., sensitivity matrix) is a non-linear regression relationship between model parameters and observations, which approximates/linearizes the non-linear forward model during the inversion process that reduces the computational burden of model calibration. PEST incorporates the Gauss–Marquardt–Levenberg Algorithm (GMLA) as a minimum search engine that efficiently converge under linear and non-linear problems (Doherty, 2015; Levenberg, 1944; Marquardt, 1963). The algorithm also provides the option of zonal parametrization that can contain a set of scattered pilot points if needed (Doherty, 2015). BeoPEST is a special edition of PEST that allows for parallel computing through adapting the protocols of TCP/IP or MPI to communicate between the master and slave computers. With the advances of BeoPEST, an in-house parallel computing system (8PCs with CPU @ 3.4GHz each) was used in this study to reduce the computational time of the model calibration process.

PEST suite of utilities includes the sub-programs of PREDUNC and GENLINPRED that can be used to evaluate the worth of different observation points through estimating their corresponding reduction in the post-calibration uncertainties of model parameters and predictions. These utilities were used in Chapter 5 to select the best monitoring locations for brine leakage. Further details about the mathematical foundation of these utilities are discussed in Chapter 5.
PEST is an inverse modeling code that can deal with highly parameterized models while maintaining reasonable stability of the calibration process through a wide spectrum of regularization techniques that can be applied by the user (Hunt et al., 2007). There are two main regularization methods available in PEST, which are Tikhonov, and subspace regularizations (Tonkin & Doherty, 2005; Doherty, 2015). Tikhonov regularization allows to define preferred values or differences for the target parameters, which constraints the minimum search of the objective function. However, the subspace method is a mathematical operation that aims to subdivide the parameter space (i.e., include all target parameters) into a Null Space, that comprises all non-sensitive and non-identifiable parameters towards observed data, and a Solution Space where all combinations of the identifiable parameters reside (Doherty, 2015). PEST freezes all Null space parameters while adjusting the Solution space parameters to fit the data, which reduces the dimensions of the searching space and the number of unknowns.

PEST was chosen for the current study because it is capable of estimating the parameters and calibrating the models of non-linear problems (i.e., solute transport) and it includes an uncertainty analysis tools in its suite of utilities. For the same reasons, a considerable number of transport and flow modeling-based studies incorporated PEST in their analyses to estimate unknown environmental properties (Bravo et al., 2002; Tonkin & Doherty, 2005; Ma et al., 2012; Wijngaard et al., 2017; Sarris et al., 2018; Sullivan et al., 2019). One indirect benefit of PEST is its ability to access, read, and edit the input and output files of FEFLOW through FePEST. This feature was utilized in the present study for bridging between the closed-source code FEFLOW and the Genetic Algorithm (GA) rooted in MATLAB 2019b (MathWorks, 2020) to find the best brine leakage monitoring and controlling locations (Chapters 5 and 6). This bridge was established by setting up PEST to calculate the objective function residuals only (no-calibration process),
where PEST runs FEFLOW using the initial system parameters to obtain model results and estimate the misfit between simulated and observed data. Thereby, PEST is basically acting as a linear operator of FEFLOW code, through which model parameters can be modified by editing PEST control files, while model results can be obtained by running PEST using the modified parameters. Such bridge allowed for incorporating the GA with FEFLOW to find the best model parameters that minimize a penalty function formulated to optimize the monitoring locations (Chapter 5) and the placement of the leakage control wells (Chapter 6). The optimization process under the GA can also be regularized, using non-linear constraints, which was an important feature used in developing the design approaches in Chapters 5 and 6, as will be discussed.
CHAPTER 4

EXPLORING THE IMPACTS OF SOURCE CONDITION UNCERTAINTIES ON FAR-FIELD BRINE LEAKAGE PLUME PREDICTIONS IN GEOLOGIC STORAGE OF CO₂: INTEGRATING INTERMEDIATE-SCALE LABORATORY TESTING WITH NUMERICAL MODELING

Reproduced with permission from the Journal of Water Resources Research, slight modifications have been made to fit the dissertation style.

2021 American Geophysical Union. Ahmad H. Askar*, Tissa H. Illangasekare¹, Andrew Trautz¹,², Jakub Solovský³, Ye Zhang⁴, and Radek Fučík³.

4.1 Introduction

CO₂ geological sequestration (CGS) is a promising technique that is being applied to alleviate the atmospheric loading of greenhouse gases contributing to global warming. Deep saline formations are identified as efficient potential repositories for captured CO₂ due to their high storage capacity and availability around the world (IPCC, 2005; Lackner, 2003; Bachu, 2003; U.S. DOE, 2015). The pressure buildup associated with CO₂ injection is shown to displace native brines
within the storage formation over significant spatial areas (Van der Meer, 1992; Nicot, 2008; Birkholzer & Zhou, 2009; Birkholzer et al., 2011a; Cavanagh & Wildgust, 2011). Natural discontinuities (i.e., faults, fractures, etc.), pressure-induced pathways (i.e., activated faults and new fractures) and/or local abandoned deep wells within the local and regional storage formation caprock can act as leakage pathways from which far-field native and/or CO2-laden brine can escape (Figure 1.2) (Gasda et al., 2004; Rutqvist et al., 2007; 2008; Celia et al., 2011). In the occurrence of brine leakage, the quality of water in overlying aquifers can be detrimentally impacted and underground sources of drinking water (USDWs) can be contaminated through increased salinity, elevated TDS concentrations and/or the introduction of toxic trace metals (such as Pb, Cd, etc.) (Birkholzer et al., 2009; Jones et al., 2015; Qafoku et al., 2017; Gupta & Yadav, 2020).

In response to the potential contamination risk of CGS operations on shallow aquifers, the US Environmental Protection Agency (EPA) established the underground injection control (UIC) program to provide regulatory framework for the protection of USDWs. Under this program, a set of requirements regarding the planning and construction of CGS projects must be fulfilled to ensure the safe and efficient long-term trapping of injected CO2 (U.S. EPA, 2010); additional CGS project risk analysis and management guidelines were released by the Department of Energy (NETL, 2017b). Quantifying the risk of USDW contamination by injected CO2 has become a key consideration at all stages of project development (Sun et al., 2013b; Pawar et al., 2015). One of the goals of CGS risk assessment involves the prediction of brine leakage scenarios that allow areas that will likely be impacted to be identified and pre-leakage mitigation strategies developed (Pawar et al., 2015; Li & Liu, 2016).

Several field-scale numerical studies highlighted that the main controllers of the leakage rate from a CGS storage formation are pressure buildup magnitude, extent of spread (Nicot, 2008;
Birkholzer & Zhou, 2009; Yamamoto et al., 2009; Zhou et al., 2010; Person et al., 2010; Lions et al., 2014) and the permeability of the storage formation and caprock damage zones (Bricker et al., 2012; Zhao et al., 2012; Wainwright et al., 2013; Xie et al., 2015a; 2015b). Jeanne et al., (2013) and Vialle et al., (2016) found that the permeability distribution of the caprock damage zone can significantly affect the predicted CO₂ leakage pathways. Hyman et al., (2020) similarly showed that the leakage of supercritical carbon dioxide from a storage formation is sensitive to the permeability, connectivity, and structure of the caprock fracture network (i.e., fracture density); the effect of variation in the network structure was found to be stronger than that of the permeability magnitude variation. All of these studies indicate that the spatiotemporal evolution of a leaking brine is dependent on the source conditions including: (1) sizes, locations and permeabilities of caprock fractures and (2) permeability field and boundary conditions of the CO₂ storage formation.

Identifying source conditions becomes more complicated after the commencement of the injection phase of CGS operations as the geomechanical stresses and geochemical reactions induced by CO₂ injection will continuously alter the hydraulic and structural properties of caprock fractures. Geomechanical responses of the caprock fractures can include sliding, dilating, closing, opening, and intersecting with other fractures. Such structural alteration in the caprock fractures can affect their permeability, mass transport and dominant fluid pathways (Lei et al., 2017; Detwiler & Morris, 2019). Geochemical reactions, such as dissolution and precipitation process, can also occur if the mixture of the formation brine and injected CO₂ are not at chemical equilibrium with caprock minerals (Ross et al., 1981; Andreani et al., 2008; Rathnaweera et al., 2016; Morris et al., 2016).
Characterizing leakage source conditions with the level of detail necessary to ensure the accuracy of brine leakage prediction in CGS risk assessments poses several technical and financial challenges. To evaluate the impact of source condition uncertainties a transport model validated using high-resolution data on the system hydrogeological settings and the spatiotemporal development of the brine plume is required. Due to field data paucity, as discussed in Chapter 3, the required data for the analysis of this chapter was generated from three intermediate-scale experiments conducted in the soil tank described in Chapter 3. The experimental data was then used to calibrate and validate a transport model developed using FEFLOW (Diersch, 2014) to evaluate the impact of multiple uncertainty scenarios of source conditions. The results of these numerical experiments provided new insights into the impacts of leakage source conditions on the evolution of a far-field brine plume from a storage formation as a result of CGS activities.

4.2 Experimental Approach

This section presents the fundamental aspects and features of the testing system briefly and discuss the applied procedures to perform the three experiments pertain to the current exploration.

4.2.1 Testing System Setup

The soil tank described in Section 3.1.3 (Figure 3.2) was used to perform three independent experiments to simulate different leakage scenarios. A horizontal soil tank was used to simulate the vertical brine leakage to enable laboratory experimentation and overcome space limitations discussed in Section 3.1.2. The brine surrogate concentration, NaBr tracer, was designed carefully to avoid any density-driven flow instability during the plume migration in the horizontal soil tank (Section 3.1.6). Three distinct zones were represented in the soil tank including: (1) the CO₂ storage zone (denoted as Zone 1), (2) intermediate overlying formations (denoted as Zone 2), and
(3) shallow aquifer (denoted as Zone 3). The brine migration from Zone 1 to Zone 3 throughout Zone 2 was simulated in this soil tank under multidimensional flow fields created by four BCs at both ends of the tank (Figure 3.2). Zones 1 and 3 were separated by a perforated acrylic thick sheet to mimic a fractured caprock overlying a CO₂ storage zone. Six external tubes were installed at the downstream exit of the sheet openings to inject the brine surrogate (i.e., NaBr tracer). For more details about the tank instrumentation and media heterogeneity design, please refer to Sections 3.1.3 and 3.1.4. It should be noted that the terms “vertical” and “horizontal” were used throughout the chapter to describe the flow and system directions according to natural settings. However, the terms “longitudinal” and “transversal” were used mainly to describe the plume movement and spread with respect to the principal direction of plume migration in the tank.

4.2.2 Experimental Procedures

NaBr tracer solution, brine surrogate, was prepared for Experiments 1, 2, and 3 by mixing 5.6, 7.24 and 11.22 grams of granular NaBr salt (102.9 NaBr/mol – Mallinckrodt 0535) with 102 liters of free-bromide tap-water (water quality was evaluated internally), respectively. The NaBr tracer was doped with a 0.03% concentration (by volume) of a blue dye (Envision DYEBLU-GL) in the 1st experiment and florescence dye (ACROS-Organics 17324-5000) in the 2nd and 3rd experiments to visually track the plume migration from the transparent side of the tank and develop sampling strategies. For more details about the adapted sampling strategy and chemical analysis procedures, please refer to Sections 3.1.7 and 3.1.8.

In the first two experiments, a series of constant-head reservoirs was connected to the tank inlets and outlets at Zones 1 and 3 to adjust the head levels imposing Dirichlet BCs. While, in the third experiment, a long well was installed at the upper boundary of Zone 3; from which water was extracted at a fixed rate by a peristaltic pump to impose a Neuman BC; this BC results in a wider
distribution of the plume across Zone 3. Table 4.1 and Figure 4.1 present the assigned BCs to develop the ambient flow field in which different unique scenarios of plume migration were created by varying the injection rates and locations. The maximum hydraulic gradients induced by the assigned BCs were 0.018, 0.018, and 0.009 for Experiment 1, 2, and 3, respectively. It should be noted that while steady-state flow was the objective in these experiments, some minor transient flow was still observed.

Table 4.1 Summary of Experimental BCs and Collected Data statistics.

<table>
<thead>
<tr>
<th>Exp. No.</th>
<th>Inj. Port**</th>
<th>Boundary Conditions*</th>
<th>NaBr Conc. (ppm) (count)</th>
<th>Injection Rate (mL/min) (count)</th>
<th>Collected Data Density</th>
<th>Aqueous samples</th>
<th>Head Records</th>
<th>Inflow and Outflow***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp-1</td>
<td>3</td>
<td>BC1=165.44cm, BC2=164.47cm, BC3=161.15cm, BC4=150.50cm</td>
<td>42.4 ± 2.4 (17)</td>
<td>7.45 ± 0.04 (17)</td>
<td>423</td>
<td>570</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Exp-2</td>
<td>1</td>
<td>BC1=165.44cm, BC2=164.47cm, BC3=161.15cm, BC4=150.50cm</td>
<td>54.5 ± 2.6 (18)</td>
<td>7.9 ± 0.06 (19)</td>
<td>582</td>
<td>480</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Exp-3</td>
<td>6</td>
<td>BC1=165.10cm, BC2=No-Flow, BC3=8.3ml/min, BC4=157.9cm</td>
<td>85 ± 4.7 (22)</td>
<td>7.9 ± 0.1 (24)</td>
<td>1767</td>
<td>540</td>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>

* BC numbers are identified in Figure 3.2.
** Injection ports #3, #1 and #6 are located in the middle, top, and bottom of tank height, respectively.
*** The number of measurements of the system inflows and outflows at Zones 1 and 3 are usually equal.

A summary of the used injection ports, average concentrations of injected NaBr tracer, average injection rates, and collected data density can be found in Table 4.1. Upon adjustment of the constant head devices and the establishment of steady flow, the tracer solution was injected using a peristaltic pump (Cole-Parmer Masterflex L/S Peristaltic Pump 7526-60) at a rate that varied between 7.5 and 8.0 ml min⁻¹. Hydraulic heads, system inflow, outflow, injection rate, and tracer concentration were measured at least twice a day to monitor the temporal variations. A large
number of tracer samples were collected during the three experiments as shown in Table 4.1. These data were used to generate high-resolution spatiotemporal concentration breakthrough curves (BTCs) that were used for model calibration and validation. The temporal durations of Experiments 1, 2, and 3 were 204, 194, and 190 hours, respectively. Figure 4.1 shows the developed plumes in the testing tank under the three considered leakage scenarios.

Figure 4.1 Projected view of NaBr distribution in the three experiments. Contour lines denote the extents of the plume boundaries at specific times. Note that the tank elbow (i.e., turn) is not included in this figure.

4.3 Numerical Modeling

4.3.1 Model Setup

A 2D flow and transport model was developed in FEFLOW (Diersch, 2014) for the packing configuration presented in Figure 3.2 to simulate tracer transport in the steady-state flow field of the test tank; steady-state flow conditions were assumed so as to reduce the computational burden
of the optimization process. To validate the 2D assumption for the problem, a local 3D model was developed for the vicinity of the injection ports and across the openings' system (the most critical part for the plume evolution) and the resulted water and mass fluxes were compared to the 2D model. A difference less than 1% between the results of the two models was obtained. Isothermal conditions were also assumed in the model due to the limited temperature fluctuation observed in the experiments (± 2°C). Reducing model dimensions from 3D to 2D and the assumed isothermal conditions helped in limiting the computational cost of the model calibration process.

The assigned boundary and initial conditions were varied between simulations following those applied during each experiment (Figure 4.2). Note that injection rates and tracer concentrations were slightly adjusted, within the variability of experimental measurements and the standard error in the chemical analysis, to provide better fit with observations. The Péclet number was larger than 10² in Zones 2 and 3 (transport domain), indicating that the plume transport was advection dominated. Based on high-resolution digitized photographs of the packed sands, hydraulic conductivity (K) heterogeneity was explicitly represented in the flow model, where model domain was discretized by 167,207 triangular elements with average nominal sizes of 0.6 ± 0.16 cm². Local mesh refinement was implemented at the tracer source area to enhance model convergence, which resulted in an element size as small as 7 x 10⁻⁴ cm².

Within the computed steady-state velocity field, tracer transport was modeled by solving the ADE with a constant averaged NaBr mass flux assigned at the prescribed source location. Neglecting the observed slight fluctuation in the measured tracer injection rate led to obtain a better fit to observation data; Thornton et al., (2013) recommended a similar simplification for simulating contaminant source discharge with short-term fluctuation conditions. Loss of solute due to tracer sampling was considered negligible and not accounted for the modeling efforts. The predictor-
corrector time integrator introduced by Gresho et al., (1979) and improved by Bixler et al., (1989) was used to discretize the time and progressively optimize the length of time steps – this scheme reduced the transport model computational burden with minor effect of numerical dispersion. The model was queried at the same times that the experimental samples were taken so as to produce equivalent concentration fields.

Figure 4.2 The hydraulic conductivities, boundary conditions, injection rates and tracer concentrations that were assigned in the model to generate the presented plume distributions.

The column-scale K values presented in Table 3.1 were initially assigned as the soil properties in the model. Longitudinal dispersion coefficients ($\alpha_L$) of 0.12 cm and 0.05 cm were respectively specified for Zones 2 and 3 based on the empirical formulations of Neuman (1990b). Transverse dispersion coefficients ($\alpha_T$) were assumed to be 10% of the longitudinal dispersion coefficients (Gelhar et al., 1992). While Zone 1 was assigned $\alpha_L$ and $\alpha_T$ values of 0.1 cm and 0.07
cm, respectively. Despite the good fit to hydraulic head data (error ~1mm), significant discrepancies in the shape and position of the simulated plume were observed. Possible explanations include: (1) upscaling errors associated with the use of column-scale K values (Rovey II, 1998; Barth et al., 2001b), (2) the six caprock fractures may have had nonidentical K values as their screens rusted non-uniformly over time and (3) the permeability of the four gravel sections forming the BCs (see Section 3.1.1), used to control the BCs, may not be identical. These initial simulation results therefore prompted the detailed model calibration described below.

4.3.2 Model Calibration and Validation

FePEST was employed to calibrate the model parameters through a parallelized inverse problem on multiple-PCs computing system (see Section 3.2.2). Data from Experiments 2 and 3 were used for calibration and Experiment 1 for final validation. A linear sensitivity analysis revealed that K, dispersion coefficients, and porosity (ϕ) were the most influential parameters on model outputs - in agreement with previous studies (Sarris et al., 2018; Yoon & McKenna, 2012). Thus, FePEST was set to run forward models of Experiments 2 and 3 simultaneously and iteratively to reduce the overall objective function by finding the optimum identical values of these parameters that enhances the model fit to both datasets. In the inversion problem, zonal parametrization was used to explicitly represent the spatial distribution of target parameters (K, α_L,T, and ϕ) according to the experimental packing. Details about zones and bounds that were considered for each parameter is presented in Table 4.2.

Table 4.2 Summary of the assigned initial values and bounds for parameters subjected to optimization process in PEST for each experiment.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial Value</th>
<th>Priori Bounds</th>
<th># of zones</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>K (sand #110)</td>
<td>5.52</td>
<td>0.5</td>
<td>10</td>
<td>1 m/d</td>
</tr>
<tr>
<td>K (sand #70)</td>
<td>12.1</td>
<td>6</td>
<td>50</td>
<td>3 m/d</td>
</tr>
</tbody>
</table>
Table 4.3 Continued

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial Value</th>
<th>Priori Bounds</th>
<th># of zones</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>K (sand #40/50)</td>
<td>44.93</td>
<td>30</td>
<td>150</td>
<td>3</td>
</tr>
<tr>
<td>K (sand #20/30)</td>
<td>101.95</td>
<td>80</td>
<td>250</td>
<td>3</td>
</tr>
<tr>
<td>K (sand #20)</td>
<td>181.44</td>
<td>150</td>
<td>500</td>
<td>3</td>
</tr>
<tr>
<td>K (sand #16)</td>
<td>476.93</td>
<td>250</td>
<td>1200</td>
<td>1</td>
</tr>
<tr>
<td>K (BCs gravel chamber)</td>
<td>1000</td>
<td>0.1</td>
<td>10000</td>
<td>4</td>
</tr>
<tr>
<td>K (Openings)</td>
<td>50</td>
<td>1</td>
<td>1000</td>
<td>6</td>
</tr>
<tr>
<td>φ (sand types in zones #2, #3)</td>
<td>0.5</td>
<td>0.2</td>
<td>0.6</td>
<td>9*</td>
</tr>
<tr>
<td>αL (for tank zones #2 and #3)</td>
<td>0.12 &amp; 0.05</td>
<td>0.01</td>
<td>5</td>
<td>2**</td>
</tr>
<tr>
<td>αT (for tank zones #2 and #3)</td>
<td>0.012 &amp; 0.005</td>
<td>0.01</td>
<td>5</td>
<td>2**</td>
</tr>
</tbody>
</table>

* Based on the number of sand types in zones #2 and #3.
** Macro-dispersivity was considered for the entire tank zones #2 and #3 to avoid over parameterization, as the actual heterogeneity was explicitly represented in the model.

The calibrated model’s performance was appraised using different metrics. Scatter plots were used to compare simulated concentrations with observed values. Percent absolute error (PAE) quantified the deviation of simulated system state variables (i.e., hydraulic head and inflows/outflows) from the experiments. Normalized root mean square error (NRMS) similarly was calculated to evaluate the error of simulated concentration data. PAE and NRMS are defined respectively as in Equations 10 and 11:

\[
P_{AE} = \left| \frac{V_m - V_s}{V_m} \right| \times 100
\]  

\[
NRMS = \sqrt{\frac{\sum(C_m - C_s)^2}{n}} / C_0
\]

where \(V_m\) is a measured state variable of flow (i.e., transient hydraulic head and outflows), \(V_s\) is the simulated state variable of flow (i.e., steady state heads and inflows/outflows), \(C_m\) is measured concentration, \(C_s\) is simulated concentration, \(C_0\) is tracer concentration at the source, and \(n\) is sample size. Note that the error was defined as the difference between observed and simulated values at the same location and time in both the experiments and the simulations.
Temporal and spatial moment analyses (Freyberg, 1986; Farrell et al., 1994a; Fetter et al., 2017) were also conducted so as to test the model’s ability to capture the most critical characteristics of the plume (e.g., plume mean arrival time, center of mass, and spread). For this analysis, the first temporal and spatial moments normalized by the zeroth moments were used to determine the mean arrival times ($T_c$) and the mass center ($X_c$ and $Y_c$) of the simulated and observed plumes. The normalized second spatial moments were used to determine the longitudinal and transverse plume spreads from their centers of mass ($XX_c$ and $YY_c$). $T_c$, $X_c$, $Y_c$, $XX_c$ and $YY_c$ can be calculated as:

\[
T_c = \frac{M_{10}}{M_{00}} = \frac{\int_{\Omega} t^i \cdot C(t) dt}{\int_{\Omega} C(t) dt}
\]

\[
X_c = \frac{M_{10}}{M_{00}} = \frac{\int_{\Omega} x^i \cdot C(x,y) dx dy}{\int_{\Omega} C(x,y) dx dy}
\]

\[
Y_c = \frac{M_{01}}{M_{00}} = \frac{\int_{\Omega} y^j \cdot C(x,y) dx dy}{\int_{\Omega} C(x,y) dx dy}
\]

\[
XX_c = \frac{M_{20}}{M_{00}} = \frac{\int_{\Omega} (x^i - X_c)^2 \cdot C(x,y) dx dy}{\int_{\Omega} C(x,y) dx dy}
\]

\[
YY_c = \frac{M_{02}}{M_{00}} = \frac{\int_{\Omega} (y^j - Y_c)^2 \cdot C(x,y) dx dy}{\int_{\Omega} C(x,y) dx dy}
\]

where $C$ is the tracer concentration at that spatial and temporal coordinates of $t^i$, $x^i$, and $y^j$. $\Omega$ is the computational domain, $M_{00}$ is the zeroth moment (total mass), and $M_{1-2}$ are higher order moments.
4.4 Results and Discussion

4.4.1 Experimental Results and Model Performance

Although the model used in this analysis assumes steady-state flow conditions, it was compared to the experimental transient data collected on the system state variables including: the twice-daily measured hydraulic heads, inflows, and outflows. Figure 4.3a presents the maximum, minimum, and median of PAEs associated with each system state variable as measured over the duration of the experiments. Simulated hydraulic heads showed a very good fit compared to observed data with a maximum PAE ≤0.7%. Median PAEs in the predicted inflow and outflow of Zones 1 and 3 in Experiments 2 and 3 were ≤15%, except for Zone 1 outflow in Experiment 2. The outflow from Zone 1 in Experiment 2 was affected by the accumulation of rust clogging BC2 (Figure 3.2). Therefore, this outlet was subsequently disabled in Experiment 3. During Experiment 1, the effluent rate from Zone 3 was observed to decrease gradually over time as a result of a partial clogging in the outlet connected to BC3 (Figure 3.2). Analysis of the simulation results showed that a transient flow-type boundary condition should be assigned to Zone 3 outlet in the forward model (Figure 4.2) to improve the hydraulic head predictions. It should be noted that the effluent rates from Zone 1 were not measured during Experiment 1.

Goodness-of-fit between simulated and observed plume concentrations are represented in Figures 4.3b, 4.3c, and 4.3d. The plume concentrations were well predicted as can be seen in the NRMS values of 0.098 in Exp.1, 0.105 in Exp.2, and 0.083 in Exp.3. Despite having large datasets and a well characterized system to work with, reproducing observed transport data remains difficult. Barth et al., (2001) and Yoon and McKeena (2012) described such model-to-measurement discrepancies as inevitable because of the uncaptured sub-grid scale processes introduced by micro-heterogeneity in the tank packing. From our observations, other reasons can
be included; variations in the upstream and downstream effluent rates, variability in the tracer injection rate, possible non-steady flow in the immediate vicinity of the caprock fractures which enhances transverse dispersion, standard error margin of the IC analysis, the nonpoint-like measurements of plume concentrations (average concentration over the 0.5 ml sample volume), and potential error in the defined observation locations in the model (1-2 mm).

Figure 4.3 (a) calculated percent absolute error (PAE) for each calibrated flow variable with the median (50th percentile) value shown and (b, c, d) observed versus calibrated concentration data for each experiment. Table includes Normalized Root Mean Square (NRMS), sample size, mean and standard deviation (STD) of the residuals.
The slight variability in plume position over time due to the non-steady flow at the leakage source (Table 4.1), distorted the BTCs of the septa ports located at the plume fringe and front. Although, extremely distorted data points were excluded from the calibration dataset to avoid failure in the optimization process, remained data used for calibration included some slightly distorted BTCs. This distortion was most pronounced in the data of Experiment 2 because of the partial clogging occurred in BC2 (Figure 3.2), which induced non-steady conditions at the source and caused large NRMS.

The spatial moment analysis of simulated and observed plumes revealed an average PAE of 0.43% in the location of the plumes’ centers and 10% in their widths. The temporal moment analysis showed that the model was able to reproduce the mean arrival times of the three plumes at the sampling locations of Experiments 1, 2 and 3 with an average PAE of 0.4%, 2.7%, and 1.8%, respectively. Figure 4.4 presents a visual comparison between observed and simulated plume arrival times as contour lines, which overlay spatial maps of the final calibrated parameters. As shown in Figure 4.4, a high porosity value (0.6) was assigned to the #70 fine sand in Zone 2 to match the experimental arrival times. This was caused by packing errors in Zone 2; when dividers were raised to pack the next level in the tank, small portions of sand were blended with the neighboring cells (e.g., sand #20 and sand #70 cell mixed at their interfaces). This created thin (~4mm thick) layers with different permeability which as a result, introduced unaccounted local micro-heterogeneity. Plumes encountered this micro-heterogeneity branched out (Figure 4.6), varying the arrival times of the plumes at ports located in the same sand type. This in turn misinformed PEST, causing the algorithm to unrealistically adjust the porosity to fit the data. However, it is worth mentioning that this porosity field led to a reasonable prediction of the mean arrival times of the validation Experiment #1.
Figure 4.4 Visual comparison between observed and simulated plume distributions which overlay spatial maps of the final calibrated parameters (hydraulic conductivity ($K$), Porosity ($\phi$), and transverse dispersion coefficients ($\alpha_T$)) in (a) Experiment 1, (b) Experiment 2, and (c) Experiment 3.

Ensuring that the model was able to capture the flow field of Zone 1 is an important consideration as the primary focus of this analysis is the exploration of how inaccurate information about the storage formation impacts simulation results. This was addressed by comparing the simulated and observed flow fields in Zone 1 and at the source vicinity. In this test, a bromide-free dye tracer was injected at a few locations in Zone 1 and around the fractures so that the
experimental flow field could be delineated. The resulting tracer trajectories were compared to the streamlines of the model in the same zones (Figure 4.5). The good fit between observed and simulated flow fields in Zone 1 and source vicinity revealed the ability of the calibrated model to reproduce the physical processes controlling NaBr migration in the tank while fitting observed data. Based on above discussion, the validated flow and transport model was verified as an analog for the physical experiment and could therefore be used to explore the impacts of different scenarios of uncertainty in source conditions.

![Figure 4.5: One example of the conducted visual comparison between the model and experimental flow fields at Zone 1: (a) the flow pathways and their trajectories (red dashed lines) of the bromide-free florescence tracer in Zone 1 during the flow field delineation experiments, (b) the uncalibrated model streamlines (yellow dashed lines) in Zone 1 and (c) the calibrated model streamlines (yellow dashed lines) in Zone 1.](image)

4.4.2 Numerical Experiments

In order to explore the impact of source condition uncertainties, the validated model was used to create a hypothetical reference scenario (Figure 4.6a) based on the physical dimensions of the soil tank and calibrated model parameters (Figure 4.4). In this scenario, the model domain was rotated back by 90° to simulate the brine migration under the density contrast effects expected in
the field, which was manipulated in the tank to enable laboratory experimentation. The authors recognize that the brine leakage plume predicted under the chosen modeling scale is likely to be specific to the simplified experimental settings of this research. Much like other CGS intermediate-scale studies, e.g., Trevisan et al., (2017) and Plampin et al., (2017), this one can still provide important insights for modeling the full-scale problem. Moreover, upscaling the model parameters that were calibrated using experimental data to field dimensions was considered invalid to avoid introducing scaling errors. Nevertheless, this analysis can still shed the light on the important source conditions that need to be well-characterized to reasonably predict the brine leakage migration pathways and areas likely to be impacted in the shallow aquifer.

In the analysis presented below, the effects of altering source conditions on the relative change of the brine plume distribution is quantified by comparing the simulation results against those of the developed reference scenario. This scenario applied the same boundary conditions as Experiment 1, except the hydraulic head at BC3 that was lowered to 160 cm so as to ensure that the plume does not approach the boundaries of Zone 2 after including a lateral flow of 0.005 m/day at Zone 2 (Figure 4.6a). It should be noted that the lateral flow in Zone 2 was applied only in the numerical experiments. The densities of the native brine and overlying groundwater were assumed to be 1100 and 1000 Kg/m$^3$, respectively (Zhou et al., 2010; Birkholzer et al., 2011a; Sun et al., 2013b). Steady-state flow and transport simulations were considered satisfactory for this analysis as risk assessment is typically informed by the eventual distribution of the brine plume.
Figure 4.6 (a) The simulated steady state plume of the reference scenario (“reference plume”) including assigned BCs and (b) Zone 1 heterogeneity used to develop the new reference scenarios. Note the direction of flow is from bottom to top and all outer boundaries were assigned no-flow BCs unless otherwise stated.

The overall uncertainty analysis was composed of 46 different scenarios, categorized under five possible cases of uncertainty, in which various hydraulic and geometrical parameters were adjusted. Descriptions of different types of tested scenarios are provided in Table 4.3. The scenarios were divided into two primary groups (denoted A and B) according to the location where the system parameters were adjusted; parameters included the hydro-structural settings of the storage zone (group A) and caprock fractures (group B). The impact of the changes associated with each scenario was quantified in terms of the relative error, which accounts for errors in the
predicted spread and pathway of the brine plume. A Cartesian (X, Y) coordinate system was used, in which X corresponds to the streamwise longitudinal direction and Y corresponds to the transverse (relative to primary flow) direction (Figure 4.4 and 4.6a). The plume mass center (Xc and Yc) and spread (XXc and YYc) in the uncertainty scenarios, calculated using Equations 13-16, were compared to the reference plume using Equation 17:

\[
\text{Relative Error for } Z (RE_Z) = \left| \frac{Z_{\text{Reference plume}} - Z_{\text{Uncertainty scenario}}}{Z_{\text{Reference plume}}} \right| \times 100
\]

where \((Z)\) is an arbitrary variable that can represent any of the normalized spatial moments, Xc, Yc, XXc and YYc. To recognize changes in the geostatistical parameters in scenarios A7 to A27 in Table 4.3, it is important to recall the original geostatistical parameters of the implemented heterogeneity in Zone 1 in the soil tank and the reference scenario, which were \(\lambda_V / \lambda_{Hz} = 35/11\), \(\sigma^2_{lnK} = 1.0\), \(\mu_{lnK} = 3.8\). The adapted heterogeneity realizations of Zone 1 in the uncertainty scenarios from A4 to A27 can be found in Figure 4.7.

Owing to the packing limitations described in Section 3.1.4, the implemented heterogeneity in Zone 1 included a homogenous part upstream of the fractures (highlighted red in Figure 3.2). This part can create a uniform flow prior to the fractures and impact our evaluation of the errors associated with assuming a homogeneous storage formation or a single leakage pathway (Scenarios A1 and B4). To ensure efficient evaluation of these errors, two new reference scenarios underlaying more heterogeneous lognormal hydraulic conductivity field (lnK) and occupying the full area of Zone 1 were developed (Figure 4.6b) and used to address the impact of these equivalent assumptions in additional scenarios (A1\(^\prime\), A1\(^\prime\)\(^\prime\), B4\(^\prime\), and B4\(^\prime\)\(^\prime\)).
<table>
<thead>
<tr>
<th>Uncertainty Location</th>
<th>Scenario ID</th>
<th>Scenario Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case 1</strong>: Assuming Homogeneous Storage Formation (5 scenarios)</td>
<td></td>
<td>A1 to A3: Independently, assigning well-estimated, halved, and doubled equivalent lnK in Zone 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A1' &amp; A1``: Assuming homogenous lnK to represent two more heterogeneous permeability fields of Zone 1</td>
</tr>
<tr>
<td><strong>Case 2</strong>: Storage Formation Geostatistical Parameter Uncertainties (24 scenarios)</td>
<td></td>
<td>A4 to A6: Identical geostatistical parameters to original reference scenario ((\lambda\sqrt{\lambda} = 35/11, \sigma^2\lnK = 1.0, \mu\lnK = 3.8))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A7 to A9: Uncertainty introduced through decreasing the mean lnK ((\mu\lnK)) ((\lambda\sqrt{\lambda} = 35/11, \sigma^2\lnK = 1.0, \mu\lnK = 2.5))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A10 to A12: through increasing the variance of lnK distribution ((\sigma^2\lnK)) ((\lambda\sqrt{\lambda} = 35/11, \sigma^2\lnK = 3.5, \mu\lnK = 3.8))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A13 to A15: through decreasing the (\mu\lnK) and increasing the (\sigma^2\lnK) ((\lambda\sqrt{\lambda} = 35/11, \sigma^2\lnK = 3.5, \mu\lnK = 2.5))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A16 to A18: through decreasing the transverse correlation length ((\lambda_v)) ((\lambda\sqrt{\lambda} = 11/11, \sigma^2\lnK = 1.0, \mu\lnK = 3.8))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A19 to A21: through decreasing both (\lambda_v) and (\mu\lnK) ((\lambda\sqrt{\lambda} = 11/11, \sigma^2\lnK = 1.0, \mu\lnK = 2.5))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A22 to A24: through increasing (\lambda_v) ((\lambda\sqrt{\lambda} = 80/11, \sigma^2\lnK = 1.0, \mu\lnK = 3.8))</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A25 to A27: through decreasing both (\lambda_v) and (\mu\lnK) ((\lambda\sqrt{\lambda} = 80/11, \sigma^2\lnK = 1.0, \mu\lnK = 2.5))</td>
</tr>
<tr>
<td><strong>Case 3</strong>: Uncertainty in Storage Formation Boundary Conditions (3 scenarios)</td>
<td></td>
<td>A28 to A30: Independently, increasing BC1 by 2%, lowering BC2 by 2%, and lastly considering both errors (varying BC1 by +2% and BC2 by -2%)</td>
</tr>
<tr>
<td><strong>Case 4</strong>: Uncertainty in Fracture Hydraulic Conductivities ((K_{fracs})) (8 scenarios)</td>
<td></td>
<td>B1 to B3: Independently, doubling (K_{fracs}), halving (K_{fracs}), and reversing the order of (K_{fracs})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B4 to B6: Once at a time, representing all fractures by an equivalent single pathway ((K_{eq.fracs})), halving (K_{eq.fracs}) and doubling (K_{eq.fracs}).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B4' &amp; B4``: Assuming an equivalent single leakage pathway when two more heterogeneous permeability fields were considered in Zone 1.</td>
</tr>
<tr>
<td><strong>Case 5</strong>: Poor Characterization of Fracture Structural Settings (6 scenarios)</td>
<td></td>
<td>B7 to B9: Independently, moving all fractures 5cm upward, 5cm downward, and moving the two middle ones 5cm apart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B10 to B12: Independently, doubling, halving all fracture sizes, and doubling the size of the middle ones only</td>
</tr>
</tbody>
</table>

* Details in each line in the table respectively describe the induced uncertainty in each scenario stated in the second column of the table. Adjusted variables are Zone 1 lnK, \(\sigma^2\lnK\), \(\mu\lnK\), \(\lambda\), BCs, \(K_{fracs}\), \(K_{eq.fracs}\). Scenarios A1’, A1``, B4’, and B4`` (grey-shaded) relied on different reference scenarios as shown in Figure 4.6.
**4.4.3 Relative Error Analysis**

The following discussion is divided into five primary subsections, each addressing a different practical case of uncertainty and its respective scenarios (Table 4.3). The errors in plume predictions associated with all 46 scenarios are presented in Figure 4.8. When viewed collectively, the results showed significantly lower relative errors (REs) in the longitudinal position of the predicted plume center and spread ($X_c$ and $XXc$) compared to the REs in its transverse location and...
spread (Yc and YYc), except scenarios B3 to B6. Therefore, in the following discussion a particular focus will be given to the REs in plume transverse center and spread (RE_Yc and RE_YYc), and to the RE_Xc for scenarios B3 to B6 only. It should be noted that errors in predicting the transverse position and spread of the plume are more critical for identifying the likely impacted areas in the shallow zone, as the errors in the longitudinal direction are more related to the plume arrival times.

**Assuming Homogeneous Storage Formation (Case 1):**

Under most standard field operations, only low-resolution data regarding the storage reservoir are readily available. Modelers, therefore, often represent the heterogeneous storage zone with an equivalent (homogeneous) K value [e.g. Class et al. (2009), Keating et al. (2013), and Vermeul et al. (2014)]. In this context, Scenarios A1 to A3 investigated how the accuracy of the estimated equivalent K for the heterogeneous storage zone can impact the simulation results. Except for Scenario A2, these scenarios showed a reasonable prediction of the plume transverse position and spread (RE_Yc<4.5% and RE_YYc<6.5%). Scenario A2 resulted in a relatively high error in the plume width (RE_YYc =15.4%) due to a reduction in the leakage flux caused by the wrongly estimated, low equivalent K, assigned in Zone 1. Such lowered leakage flux allowed the transverse flow in Zone 2 to enhance plume spreading. The reasonable prediction of the plume migration in Scenario A1 can be due to the homogeneously packed part in Zone 1 that made the inflow to the fractures uniform.

Unlike Scenario A1, the impact of homogenizing a fully heterogeneous storage formation was evaluated through Scenarios A1‘ and A1‘‘. Using an equivalent K under those scenarios resulted in the largest RE in plume transverse position among all tested scenarios (19%) and caused a RE in plume spread up to 49%. It was noted that the main cause of these errors is the inaccurate representation of the permeability field underling the fractures by homogenizing the entire storage
zone. Therefore, the homogeneity assumption for the storage formation should only be considered when the layer underling the damage zone is homogeneous, and the equivalent $K$ is representative for this layer. Therefore, it is important to first identify the potential leakage pathways in the Caprock (Ingram & Urai, 1999; Ligtenberg, 2005) and then characterize the local heterogeneity underling them with high-resolution methods, such as cross-well seismic imaging (Yu et al., 2008; Nalonnil & Marion, 2012).

**Storage Formation Geostatistical Parameter Uncertainties (Case 2):**

The availability of high-resolution data to accurately estimate the geostatistical parameters of the storage formation heterogeneity is typically rare. This raises a question of what our prediction accuracy would be if such data were available. In order to explore this case, different realizations of ln$K$ distribution were considered in Scenarios A4 to A6 while maintaining the geostatistics of the storage zone heterogeneity in the original reference scenario. Results of these simulations revealed a RE in the plume size and the transverse location of the plume center less than 3.5% and 1.2%, respectively. This suggests that knowing the actual $K$ distribution relative to fracture locations is still important for obtaining acutely high accuracy of plume predictions.

Scenarios A4 to A6 exhibited less REs than of those calculated under scenarios incorporating geostatistical parameter uncertainty (A7 to A27). Results from these scenarios showed that inaccurate estimate of the mean ln$K$ ($\mu_{\ln K}$) can lead to a significant error in the predicted brine leakage compared to the ln$K$ variance ($\sigma_{\ln K}^2$) or the correlation length ($\lambda_L$). The error in the predicted transverse location of plume center (RE$_{Yc}$) was shown to increase by 4-17% when $\mu_{\ln K}$ was decreased from 3.8 to 2.5. Lowering the $\mu_{\ln K}$ of Zone 1 decreased the leakage flux through the fractures and across Zone 2. This allowed Zone 2 transverse flow to control the plume position and cause a large spreading of the mass (average RE$_{YYc}$=51%) compared to scenarios with
correct $\mu_{\ln K}$ (average RE$_{Yc}$=9%). Uncertainty in $\sigma^2_{\ln K}$ or $\lambda_L$ (Scenarios A10-12, A16-18, A22-24) yielded moderate to high RE$_{Yc}$ values (~0.4% to 9%). Therefore, the $\mu_{\ln K}$ of the storage formation should be well estimated and defined in the model even if some uncertainty is expected in our estimates of the lnK distribution variance and correlation lengths.

Figure 4.8 (a, b) Relative error (RE) in the location of the plume center and (c, d) plume width relative to the reference plume under the studied 5 cases. The graph includes a cartoon illustrating what is meant under each type of RE and the coordinates adapted in the estimation of the RE (Figure 4.6). Note, y-axis changes dramatically from a, b to c, d.

The largest REs in the predicted plume position and size under case 2 were observed in Scenarios A21, A25, and A27. In these scenarios, the plume center was largely deviated (RE$_{Yc}$ ≈ 12-17%) as a result of the presence of low K inclusions blocking fracture numbers 3 to 6 (Figure 4.6), which reduced the overall leakage flux and affected the hydraulic head distribution at the source. Under this head distribution, the plume migrated along different pathway from the leakage
source. A large spreading of the plume (RE\textsubscript{YY}c between 63-105\%) was also observed in these scenarios because the transverse flow in Zone 2 became more dominant after leakage flux was reduced. These simulations emphasize again the importance of local heterogeneity characterization discussed in Case 1.

**Uncertainty in Storage Formation Boundary Conditions (Case 3):**

Accurately evaluating the hydraulic head in a storage formation can be challenging in the field. In Scenarios A28 to A30, a 2\% error was assumed in the constant head BCs of Zone 1 to mimic limitations of field measurements. Results showed that the plume width and its transverse position were more sensitive to lowering the downstream head (BC2) than increasing the upstream head (BC1) in Zone 1. Increasing BC1 by 2\% resulted in a RE\textsubscript{Yc} of 0.3\% and a RE\textsubscript{YY}c of 16\%, while decreasing BC2 by 2\% resulted in a RE\textsubscript{Yc} of 9.5\% and a RE\textsubscript{YY}c of 59\%. Lowering the head at BC2 (Scenario A29) redirected a significant portion of the background flow in Zone 1 towards this downstream boundary reducing the leakage flow from the fractures. Such reduction in the leakage flow allowed the transverse flow in Zone 2 to displace the plume and enhance its spreading. However, in Scenario A28, elevating the upstream head (BC1) increased the hydraulic gradient over the entire problem domain (between Zones 1 and 3). Therefore, the longitudinal flow in Zone 2 was increased, which limited the impact of the transverse flow of this zone on the plume.

By including both errors in Zone 1 BCs (Scenario A30), opposing effects were promoted on the leakage flow, which maintained the plume velocity in Zone 2 and limited the error in its spread. However, under this scenario the RE in the plume transverse position was relatively high (RE\textsubscript{Yc} \sim 6\%) due to the induced changes in the head distribution at the source by the inaccurate BCs. This suggests that both BCs of the storage formation, particularly the downstream one, should be accurately defined in the model to make reasonable predictions.
Uncertainty in Fracture Hydraulic Conductivities (Case 4):

In Scenarios B1 and B2, the hydraulic conductivities of the fractures ($K_{fracs}$) were multiplied by factors of 2 and 0.5, respectively, to investigate the impact of this parameter on model results. These scenarios resulted in small REs in plume transverse position and spread ($RE_{yc} < 3.0$ and $RE_{yyyc} < 2.2$). However, in Scenario B3, where accurate $K_{fracs}$ were estimated at false locations in the fracture network (i.e., reversing the order of the six $K_{fracs}$ while maintaining the fracture locations), $RE_{yc}$ and $RE_{yyyc}$ were up to 8.8% and 27%, respectively. This highlights that the spatial distribution of the inferred fracture hydraulic conductivities can be more important than their magnitudes.

By assuming an equivalent single leakage pathway (Scenario B4), the RE in plume position was significantly lower than in plume spread ($RE_{yc} = 0.1\%$ and $RE_{yyyc} = 20\%$). Notably that by adding a margin of error to the used equivalent $K$ of this single pathway (Scenarios B5 and B6), the REs almost did not change, which again shows the low sensitivity of model predictions to the magnitude of $K_{fracs}$. In addition, comparing Scenarios B3 and B4 shows that assuming a single leakage pathway to represent discrete fractures can result in more accurate predictions than assigning wrong distribution of the fracture permeabilities.

The above resulted low RE in the plume transverse position when a single leakage pathway was assumed (Scenario B4) can be attributed to the effect of streamlining the inflow to the fractures by the packed homogeneous part in Zone 1. However, when a fully heterogenous storage formation was considered (Scenarios B4′ and B4′′), assuming a single leakage pathway resulted in a reasonably high RE in plume transverse position and spread (max $RE_{yc} = 7\%$ and max $RE_{yyyc} = 29.5\%$). It should be noted that the maximum RE in the predicted plume transverse position under
these scenarios was still less than Scenario B3, where the inferred $K_{\text{fracs}}$ were assigned wrongly to the accurately identified/located fractures in the caprock.

Only Scenarios B3 to B6 showed a noticeable error in the longitudinal location of the plume mass center ($R_{E_{Xc}}$ ranged between 4.8% and 15.5%). This can be explained by the augmented focused leakage flow through the assumed single pathway. Therefore, assuming a single leakage pathway can affect the prediction of the plume arrival times.

**Poor Characterization of Fracture Structural Settings (Case 5):**

As shown in Figure 4.8, the impact of slightly altering the fracture structural settings was minor on the prediction plume pathway and spread. This can be due to two reasons; (1) the homogeneously packed part in Zone 1, and (2) the limited variation in the fracture sizes and locations in the tested scenarios (Table 4.3) because of the relatively short transverse length-scale of the soil tank (123 cm). Given the findings of Hyman (2020), discussed in the introduction, it is recommended to further explore the impact of uncertainty in fracture network structure on brine migration using a larger scale of the problem domain.

**4.5 Summary**

In this chapter, the impact of source condition uncertainties on the accuracy of leaking brine plume predictions was investigated (**Hypothesis 1**). Prediction models should be able to simulate brine leakage and transport in complex multilayered geologic systems with interacting regional natural and leakage flows. As field datasets are not readily available for model testing and validation, three comprehensive intermediate-scale laboratory experiments were used to generate high-resolution spatiotemporal data on brine plume development under different leakage scenarios. Experimental data was used to validate a flow and transport model developed using
existing code FEFLOW to simulate brine plume under varying source conditions. Spatial moment analysis was conducted to evaluate how uncertainty in source conditions impacts brine migration predictions. Results showed that inaccurately prescribing the permeability field of storage formation and caprock fractures in models can cause errors in leakage pathway and spread predictions up to ~19% and ~100%, respectively. Moreover, the analysis presented herein pointed out which of the characterization parameters result in the largest errors in the predictions. This knowledge will help to cost-efficiently investigate source conditions and thus improve site selection and design of CGS operations. This is particularly critical for cases where the storage zone caprock is overlaid by USDWs and can potentially leak under high injection stresses. The findings of this analysis also highlighted that the risk assessment studies developed based on poor characterization of source conditions can be misleading and unreliable.
5.1 Introduction

CO₂ geological sequestration (CGS) into deep saline formations is promoted as a key technology to reduce the loading of anthropogenic greenhouse gases into the atmosphere to mitigate global warming (IPCC, 2005; Lackner, 2003; Bachu, 2003; U.S. DOE, 2015). During and post-injection of supercritical CO₂ (ScCO₂), existing faults or pressure-induced fractures in the confining caprock of the storage formation pose a potential leakage risk of formation brine and CO₂ (Gasda et al., 2004; Rutqvist et al., 2007; 2008; Celia et al., 2011). Leakage of CO₂/brine can drastically degrade the quality of underground sources of drinking water (USDWs) by elevating their total dissolved solids (TDS) and toxic metal concentrations (Birkholzer et al., 2009; Jones et al., 2015; Qafoku et al., 2017; Gupta & Yadav, 2020). In the face of such possible adverse impacts, the U.S. Environmental Protection Agency (EPA) and International Energy Agency (IEA) released
a regulatory framework that requires submitting a monitoring, reporting, and verification plan before CO₂ injection. This plan should include a detailed description of monitoring methods, technologies, locations, and observed parameters (i.e., pressure and salinity) (U.S. EPA, 2013; IEA, 2010).

Numerous monitoring methods have been developed for CGS applications (Ajayi et al., 2019; NETL, 2017a; Vermeul et al., 2016). These methods can be grouped under three general categories: (1) air-borne, (2) near-surface, and (3) deep subsurface systems. Although monitoring systems in Categories 1 and 2 are relatively less costly to implement, their reliability as standalone systems is always questionable due to several reasons (Yang et al., 2011). Their signals are sensitive to environmental background noise, which can lead to false warnings due to leakage-unrelated mechanical and hydraulic stresses in the shallow zones. Moreover, they have limited ability to detect the CO₂/brine leakage early due to their substantial separation distances from the source (Winthaegen et al., 2005; NETL, 2017a). Integrated monitoring systems consisting of shallow observation wells and deep zone sensors (Category 3) allow early leakage detection and can provide data for model calibration, which will improve leakage-related predictions (Bourne et al., 2014; Ajayi et al., 2019). Installing a large number of sensors in deep overlying formations of a CO₂ storage zone (1-2 km deep) is expected to be cost-prohibitive (Hovorka et al., 2013; Nordbotten et al., 2004; Vermeul et al., 2016; Tsang et al., 2008) and might also pose safety risks (Kelessidis, 2009; Liu et al., 2013). Various methods have been developed to optimize the number of needed observation points to reduce the monitoring cost (Yang et al., 2011; Seto & McRae, 2011a; 2011b; Dai et al., 2015; 2016; Yonkofski et al., 2016; Yang et al., 2017; Chen et al., 2017; 2018; Jeong et al., 2019).
Despite the adverse impacts of brine leakage on USDWs, it can be treated as an early alert for CO\textsubscript{2} leakage from deep zones (Sun et al., 2013; Réveillère & Rohmer, 2011). Thus, models simulating the migration of brine leakage to the shallow aquifer can be used as numerical testbeds to identify the optimal monitoring locations within the extended footprint of CO\textsubscript{2} injection (Birkholzer et al., 2011a; Cavanagh & Wildgust, 2011; Celia et al., 2011; Wunsch et al., 2013). Predictions using these models are usually uncertain due to the typically poor knowledge about source conditions in deep geological settings (refer to Chapter 4). Hence, a key focus when using models to design CGS monitoring systems should be to reduce model prediction uncertainty. Predictive uncertainty is an accumulated uncertainty resulting from multiple sources (e.g., model structure and parameter errors) at different knowledge levels (e.g., data-statistical and scenario uncertainty) of various nature (e.g., epistemic uncertainty) (Walker et al., 2003). These three dimensions of uncertainty (i.e., sources, levels, and natures) will be systematically highlighted throughout the chapter. Researchers have developed both linear and non-linear approaches to estimate predictive uncertainty. These approaches have been employed to find the best monitoring locations so that predictive uncertainty of calibrated models is reduced significantly (Dausman et al., 2010; Herckenrath et al., 2011; Brunner et al., 2012; Wöhling et al., 2016; Dai et al., 2016; Chen et al., 2017; Vilhelmsen & Ferré, 2018; Zell et al., 2018; Safi et al., 2019). These approaches are usually called optimal design (OD) frameworks. It is worth mentioning that monitoring data has a specific reducibility degree to the predictive uncertainty through model calibration (Funtowicz et al., 1990; Van der Sluijs et al., 2002). Incorporating extra data in model calibration does not always guarantee better predictions, as the new data may introduce processes that were not considered in the problem conceptualization and model development, which can increase the structure and scenario uncertainties in the problem under study (Walker et al., 2003).
Generally, using non-linear Monte-Carlo (MC) based methods to estimate predictive uncertainty under an OD framework can be computationally burdensome due to the need for a large number of forward runs. Examples of such methods are the Generalized Likelihood Uncertainty Estimation, and Markov-Chain MC methods (Jin et al., 2010; Chen et al., 2017). To reduce the computational cost of non-linear approaches, Tonkin and Doherty (2009) introduced the calibration-constrained MC method that was later evaluated by Tavakoli et al., (2013). Oladyshkin et al., (2011) and Pasetto et al., (2014) used less computationally-demanding surrogate models to minimize the cost of relying on MC simulations. Others such as Chen and Oliver (2013) and White [2018] incorporated the ensemble smoother technique to lower the effort needed for finding the parameter posterior probability density function. Despite the above approaches, non-linear methods still pose different degrees of numerical challenges in determining model prediction uncertainties compared to linear techniques. Moreover, the uncertainty analysis computational cost tends to be directly proportional to model non-linearity and complexity (Safi et al., 2019). Thus, the applicability of such methods to complex groundwater transport models is tied to how much computational capacity is available to the user.

In contrast, the linear first-order second-moment method (FOSM), also called the “linear error propagation” method (Goldstein & Wooff, 2007; Doherty, 2015; White et al., 2016) is more computationally efficient as it needs running the forward model only once to generate the required sensitivity matrices for estimating predictive uncertainty. Moreover, a comparison between the results of FOSM and MC-based methods showed that the FOSM can reasonably estimate predictive uncertainty even for highly non-linear models (Kunstmann et al., 2002; Dausman et al., 2010; Herckenrath et al., 2011). Zell et al., (2018) pointed out that FOSM overestimates predictive uncertainty, which enables FOSM-based OD frameworks to generate conservative designs for
monitoring systems. It is noteworthy that the uncertainty quantification tool in the PEST suite of utilities (Doherty et al., 2010; 2015), called PREDUNC is developed based on the theory of linear error propagation (i.e., FOSM) within a Bayesian context (Moore & Doherty, 2005; Christensen & Doherty, 2008).

In this chapter, the applicability of a developed OD framework for designing CGS monitoring systems with a minimum number of deep sensors is validated. In this OD framework, PREDUNC and Genetic Algorithm (Deep et al., 2009) were combined with a FEFLOW-based transport model to select the best monitoring locations for the potential brine leakage from CO₂ geological storage. The selection was based on the informativity of data collected from these locations to model predictions of our interest; similar approaches were adapted by Wöhling et al., (2016), Vilhelmsen & Ferré (2018), and Saif et al. (2019) to design monitoring systems for shallow groundwater regimes. Validating such an OD framework becomes challenging in the field, as initiating a leakage event to collect data from the optimum monitoring locations to test the applicability of the method is neither practically nor financially feasible. In addition, the reliability of using synthetic data from numerical models would always be conditioned by the accuracy of the models in reproducing the complex processes associated with brine migration from the deep to shallow zones.

In order to resolve these challenges, the soil tank described in Section 3.1.3 was used to validate the developed framework. The validation included four main steps: (1) applying the OD framework to develop a monitoring system for a simulated leakage event in the soil tank, (2) collecting data using this system to calibrate the FEFLOW-based transport model, (3) comparing model predictions with experimental observations, and (4) evaluating data-to-prediction informativity and thus the OD framework efficacy in selecting the most critical monitoring
locations for leakage detection and model calibration. Moreover, as a part of the OD framework application, the worth of different data types collected from the deep and shallow zones toward the predictions of interest to CGS operators was discussed; the considered predictions in this analysis included the leakage rate, storage zone flow, and plume concentrations in the shallow aquifer, in addition to determining the leakage source and fracture permeabilities.

5.2 Global Optimal Design Framework (OD Framework)

In this section, a brief review of the principles of linear uncertainty analysis is provided along with a detailed description of the modified OD framework (Figure 5.1) after Wöhling et al., (2016) and Vilhelmsen & Ferré (2018).

5.2.1 Linear Predictive Uncertainty Analysis

Under the assumptions of local model linearity and Gaussian variability of the noise associated with model parameters, observations, and predictions, the posterior uncertainty of a certain prediction ($s$) can be determined using Equation 18 as described by Dausman et al., (2010):

$$\sigma_s^2 \approx \mathbf{y}^t \mathbf{C}(\mathbf{k}) \mathbf{y} - \mathbf{y}^t \mathbf{C}(\mathbf{k}) J^t [J \mathbf{C}(\mathbf{k}) J^t + \mathbf{C}(\mathbf{\epsilon})]^{-1} J \mathbf{C}(\mathbf{k}) \mathbf{y}$$  \hspace{1cm} (18)

where $J$ is the linearized form of the forward model, or as referred to in PEST “Jacobian matrix”, where the sensitivity of model outputs to parameters during the calibration process is expressed, $\mathbf{y}$ is the sensitivity of predictions to model parameters, $\mathbf{C}(\mathbf{\epsilon})$ is a square covariance matrix of model-to-data misfit errors (this encapsulates structure, geometry, and measurement errors that all fall under the category of epistemic uncertainty), and $\mathbf{C}(\mathbf{k})$ is a square covariance matrix of model innate parameters ($\mathbf{k}$) variability, representing the statistical uncertainty in the prior information about the system hydraulic properties. In both $\mathbf{C}(\mathbf{k})$ and $\mathbf{C}(\mathbf{\epsilon})$ matrices, the
diagonal elements represent the variance of parameters or observations, while the off-diagonal elements represent their inter-correlations.

The $C(\epsilon)$ matrix is usually assumed to be diagonal in recognition of the probability of measurements being statistically independent (Dausman et al., 2010) and is due to the difficulty of predicting the true covariance of errors between model outputs and data. Vilhelmsen and Ferré (2018) argued that an assumption about $C(\epsilon)$ off-diagonal elements is important, even if the assumed correlation structure was highly biased, to eliminate obtaining redundant data from the monitoring systems. Based on the previous chapter results, where the correlation between model-to-data misfit errors was found close to zero, $C(\epsilon)$ was assumed diagonal.

In the proposed approach, $J$ matrix and $y$ vector were calculated using FePEST (Diersch, 2014). $\sigma^2_s$ was calculated using Equation 18 through the utility of PREDUNC under PEST (Figure 5.1). Note that Equation 18 does not need either field observations, parameter values, or model outputs to be known before application but needs only the expected uncertainties in them to be assumed. Therefore, this equation suites the class of problems where available data about the system are sparse and imperfect (e.g., brine leakage problem). By default, PREDUNC assumes that the pre-calibration variances of observations represented in the $C(\epsilon)$ diagonal elements, are proportional to the inverse of observation weights (Doherty, 2015). Thus, by assigning zero weight to a particular observation in the PEST control file, PREDUNC is instructed to disregard this observation while developing $C(\epsilon)$ and calculating $\sigma^2_s$ (Doherty, 2015).

5.2.2 Data Worth

Despite the underlying assumption of model linearity, a computationally efficient data worth method can still be formulated using Equation 18 by evaluating the increase in post-
calibration predictive uncertainty due to data exclusion. This can be achieved by calculating the

data worth ($DW$) as follows:

$$DW = \frac{\sigma_{sub}^2 - \sigma_{full}^2}{\sigma_{pre}^2 - \sigma_{full}^2}$$  \hspace{1cm} (19)$$

where $\sigma_{full}^2$ is the lowest achievable predictive uncertainty representing the case where

measurements (observations) from all the potential monitoring locations are available, $\sigma_{sub}^2$ is the

predictive uncertainty after disregarding a specific subset of monitoring points, and $\sigma_{pre}^2$ is the pre-

calibration uncertainty of model predictions based on our prior knowledge. Equation 19 evaluates

the mere contribution of tested data (i.e., data disregarded from the entire dataset) in the total

reduction of predictive uncertainty. $DW$ can be calculated for individual points or a combination

of monitoring locations. Notably, in the case of evaluating the worth of an individual observation

point (i.e., only one observation location was disregarded), the uncertainty reduction caused by the

interdependency between data information contents is discounted by subtracting $\sigma_{full}^2$ from $\sigma_{sub}^2$

as shown in Equation 19. This minimizes the potential redundancy in the data gathered by the

designed monitoring system.

Equation 19 assumes that all potential monitoring locations can be predefined, which is

practical because one cannot place a sensor or an observation well anywhere in the problem

domain due to both functional and regulatory constraints. Such well-defined solution space renders

the search algorithm high reliability in terms of finding the most optimum “applicable” monitoring

scheme. Moreover, the reduced dimensionality of the solution space by predefining the potential

monitoring locations in the system domain minimizes the computational burden of the entire

optimization process. Notably, Matrix $J$ and $\sigma_{full}^2$ are computed once for using Equation 19.
However, $\sigma^2_{sub}$ has to be iteratively calculated for each monitoring design evaluated. In each iteration, tested data have to be assigned zero weights in the PEST control file so that their contribution to the uncertainty reduction is unaccounted by PREDUNC. Components A and B in Figure 5.1 show the applied procedures to calculate $\sigma^2_{full}$ and $\sigma^2_{sub}$.

5.2.3 DW to Multiple Predictions

The goal of CGS monitoring should not be limited to provide an early alert of contamination risk; but also to generate useful information that helps to identify the leakage source and improve model predictive uncertainty (Winthaegen et al., 2005; Jung et al., 2013; Bourne et al., 2014). Multiple predictions, such as source location, leakage flux, fracture permeabilities, storage zone lateral flow, and plume distribution in the shallow aquifer, might be of interest to operators and remedial system designers of CGS projects. However, the calculated $DW$ in Equation 19 is a prediction-specific value because the informativity of data to model predictions varies with the type (e.g., concentration, head, flow, permeability, etc.) and location of each prediction (Vilhelmsen & Ferré, 2018). To combine the data-worth of different monitoring locations toward multiple predictions, Vilhelmsen and Ferré (2018) introduced the term of Value Index (VI), which is defined as:

$$VI_j = \sum_{i=1}^{n} \omega_i DW_{i,j}$$  \hspace{1cm} (20)

where $VI_j$ is the Value Index of a proposed monitoring location/s $j$ toward $n$ predictions, and $\omega_i$ is the weight of a prediction $i$. $DW_{i,j}$ is the data worth of the proposed monitoring location/s $j$ considering prediction $i$. The prediction weight ($\omega_i$) is used to prioritize a specific prediction over another based on its value to a set of stakeholders (e.g., regulators and site operators). In this
chapter, the $\omega_i$s were assigned based on the importance of each prediction to CGS operators. Therefore, the leakage location and mass flux predictions were assigned the highest weights, followed by fracture permeabilities, and then plume distribution in the shallow aquifer for reasons discussed in Section 5.3.4.

Figure 5.1 Schematic of the main components of global OD framework.

**5.2.4 Sequential Evaluation**

Sequential evaluation (SE) is a classical technique used in designing monitoring systems through ranking monitoring locations based on their individual utility in reducing predictive uncertainty (Carrera et al., 1984; Wallis et al., 2014). This technique ignores the interdependency
between the information content of data from different monitoring locations, which can cause data redundancy (Wöhling et al., 2016). Despite the drawbacks of the SE technique, Vilhelmsen & Ferré (2018) argued that SE is preferred when forecasts are driven by the same physical processes and not linearly related to model inputs. In addition, this technique reduces the impact of assuming uncorrelated model-to-measurements misfit errors (Vilhelmsen & Ferré, 2018).

The proposed approach uses the SE as a qualitative tool to preliminary identify the optimum monitoring locations and prescribe them to the global search algorithm as an initial design guess (Component B in Figure 5.1). Thus, the global search algorithm can start from a location presumably near the global minimum of the solution space, which will help to reduce the convergence time of the optimization process. The efficacy of integrating the SE technique with the global search algorithm is discussed in Section 5.4. This approach allows providing a balance between the advantages and disadvantages of the SE technique.

### 5.2.5 Global Optimization

To find the optimum set of monitoring locations, the Genetic Algorithm (GA) modified by Deep et al., (2009) and rooted in the global optimization toolbox in MATLAB 2019b (MathWorks, 2020) was employed. The GA was selected for this analysis because it was successfully applied in similar OD frameworks developed by Wöhling et al., (2016) and Safi et al., (2019) to design monitoring systems for shallow groundwater systems. In addition, the GA can search the global minimum of a discontinuous, non-convex solution space with multi-attribute variables (Wöhling et al., 2016), as is the case in the leakage problem. The selected optimization function is given as:

\[
X_{opt} = \arg \min_{x \in X} [f(1/VI(x))] 
\]  

(21)
where $\mathbf{X}_{\text{opt}}$ is a vector of the best monitoring locations that exhibit the minimum of the Value Index inverse ($1/VI$), $\mathbf{X}$ is a vector of all possible monitoring locations, and sub-vector $\mathbf{x}$ is a combination of selected monitoring locations to be evaluated. The main steps of the GA process to perturb $\mathbf{x}$ and find the global minimum of Equation 21 are presented in component C of Figure 5.1.

The first population matrix ($P_1$) of the proposed monitoring designs generated by the GA is composed of $u$ rows combining different monitoring locations ($\mathbf{x}$) (Figure 5.1). Sub-vector $\mathbf{x}$ is randomly sampled from $\mathbf{X}$ and its size ($k$) depends on the desired monitoring density, while the number of sampled combinations denoted as $u$, depends on the user-defined population size ($u=50$ in this analysis). The first row of the population matrix was set in this analysis to be equal to the initial guess of the optimum monitoring locations found by the SE technique. Proposed monitoring designs ($\mathbf{x}_k$) are then ranked based on the VI of each design (Equation 20). Ranked designs in $P_1$ are then processed by the GA standard producing functions of selection, mutation, and cross-over to generate the next population ($P_2$) (Deep et al., 2009).

Similar to Wöhling et al., (2016), Vilhelmsen & Ferré (2018), and Safi et al., (2019), 40% of the current population was retained as elite candidates in the next population ($P_2$), top two ranked designs were treated as parents for producing 15% of the offspring using Laplace crossover function, and remaining individuals were resampled randomly outside the admissible range from vector $\mathbf{X}$ using a Power mutation function (Deep et al., 2009). A relatively high percentage of mutated individuals was considered in this work to intensively explore the solution space, particularly after assigning an initial guess of the best locations. The last step compares the ranks of previous and current proposed designs ($\mathbf{x}_u$ and $\mathbf{x}_{u+1}$) to select the new elite candidates for generating the next population ($P_2$) and so forth. The GA was set to progress until either the
selected best monitoring locations are the same over the last 50 trials or the maximum allowed number of trials (100 in this analysis) is reached.

Briefly, the proposed OD framework steps (Figure 5.1) are: (1) adapting the sequential evaluation technique through applying Equation 19 on each potential monitoring location to calculate its $DW$ to each prediction, (2) estimating the VI of each monitoring location using Equation 20 that represent its $DW$ toward all concerned predictions, (3) developing the initial guess of the optimum monitoring locations using the estimated VIs and define it for the GA, and (4) running the GA to find the final set of the optimum monitoring locations that provide the most informative data to reduce the uncertainty of concerned predictions.

5.3 Monitoring Scheme Design

The proposed OD framework in this chapter was validated by implementing the designed monitoring system in a soil tank and then evaluating the informativity of collected data to the concerned predictions. In this section, a brief description of the soil tank is provided, in addition to a comprehensive discussion about the monitoring system design considerations and the applied OD framework procedures. More details of the tank construction, heterogeneity design, and packing procedures can be found in Chapter 3.

5.3.1 Testing System Design

The horizontal tank described in Section 3.1.3 was used in this analysis to evaluate the ability of data collected from the monitoring system designed using the OD framework to inform the model predictions of interest. In this section, the main features of the testing tank and boundary conditions are briefly recalled. As discussed in Section 3.1.5, the tank was heterogeneously packed using six well characterized silica sands with properties presented in Table 3.1. The sand
configuration was chosen to create three distinct zones: (1) a confined saline formation where \( \text{CO}_2 \) is injected at far distance (Zone 1), (2) the overlaying intermediate zone between the caprock and the shallow aquifer (Zone 2), and (3) the unconfined shallow zone (Zone 3).

Figure 5.2 Schematic depicting (a) migration process of a leaking brine from a \( \text{CO}_2 \) storage formation under natural field conditions, (b) experimental setup, packing heterogeneity, and instrumentation. Dominant flow direction occurs from the right to the left side of the tank.
In order to reproduce the multidimensional flow field expected in this problem, multiple boundary conditions (BCs) were used in Zones 1 and 3 of the tank. Figure 5.2 presents a schematic of the conceptualized brine migration under field conditions and a sketch of the rotated experimental setup including BCs, tank instrumentation, and packing configuration.

For the current analysis, three boundaries of the tank, located at Zone 1 and the lower end of Zone 3, were connected to constant head reservoirs to obtain fixed water heads. While, the fourth boundary, located at the top side of Zone 3, contained a horizontal well that was connected to a peristaltic pump to assign a constant flow BC (Figure 5.2). With these BCs, the flow in Zone 2 was controlled by the hydraulic head difference between Zones 1 and 3. Additional 40 septa ports were installed in Zone 2 to give more flexibility in the selection of the observation points across the overlying formations (red dots in Figure 5.2). Thirty manometers were installed across the length of the tank for hydraulic head measurement. Fifteen grams of NaBr salt (Mallinckrodt 0535-102.9 NaBr/mol) was mixed with 113.5 liters of tap-water and a 0.03% (by volume) of fluorescence dye (ACROS-Organics 17324-5000) to prepare the ~94ppm Br-tracer solution (i.e., brine surrogate) of the experiment. For more details about the tank design, instrumentation, and media heterogeneity, please refer to Sections 3.1.3 and 3.1.4.

5.3.2 Base-Model Rationale and Assumptions

To obtain the sensitivity matrices required for estimating predictive uncertainty and designing the monitoring system, a numerical model that simulates the flow field and solute transport across the geological formations spanning the depth from the storage zone to the shallow aquifer is needed. This model herein is referred to as the base-model because it relies on the available basic information on the hydrogeological regime (e.g., BCs and sink/sources) and preliminary site investigation data (Safi et al., 2019). In the brine leakage problem, the monitoring
system has to be designed and implemented before the injection phase of CO$_2$ starts. Therefore, only sparse low-resolution data on the geologic structure of the system, porous media heterogeneity, and other parameters needed to develop such a base-model will be available.

Ideally, detailed information about source conditions, storage zone permeability field, and fracture hydraulic settings are required to simulate the migration pathways of a leaking brine (refer to Chapter 4). However, under field conditions, mapping the permeability field of a storage formation can be costly and challenging (Vialle et al., 2016; Lei et al., 2017). In the literature, it is generally assumed that this zone can be represented as a homogenous layer using an effective permeability value [e.g. Class et al., (2009), Keating et al. (2013b), and Vermeul et al (2014)]. However, as shown in Chapter 4, assuming homogeneity in the storage zone at the vicinity of the fractures can result in large errors in the prediction of the plume pathway. In the soil tank used for this analysis, the part of Zone 1 located below the slanting caprock (highlighted in red in Figure 5.2) was packed homogeneously due to limitations of access. This allowed us to assume an effective permeability value to represent the storage zone in the base-model, which imposes a structure uncertainty to the leakage problem addressed in this study as expected in the field.

Most of the reported CGS monitoring system designs relied on predefining the expected leakage sources (Yonkofski et al., 2016; Chen et al., 2018; Jeong et al., 2019). The approach investigated here also assumes that some prior information on the potential fracture and damage zone locations in the caprock, where brine leakage can occur, is available. Different geophysical-based methods were reported in the literature to identify the potential locations of the pressure-driven fractures in the top seal of confined hydrocarbon reservoirs and caprocks of carbon storage formations (Ingram & Urai, 1999; Ligtenberg, 2005; Lohr et al., 2008; Krawczyk et al., 2015; Ziesch et al., 2019). Recently, Feng et al., (2018) predicted the apertures, porosities, and
permeabilities of stress-induced fractures in a brittle confined reservoir with an error range of 11-15%. Briefly, their approach relied on enhancing the geomechanical model predictions by defining accurate rock properties obtained from detailed X-ray and CT core scans. Incorporating the pre-existing caprock discontinuities in the geomechanical model is essential to accurately predict the locations of the potential stress-induced fractures (Min et al., 2004; Rutqvist et al., 2007; 2008; Pan et al., 2013; Figueiredo et al., 2015). As above methods can be applied to identify the potential zones, where fractures can develop in the caprock under CO$_2$ injection stresses, predefining the leakage sources in the *base-model* simulating the tank setup was considered a valid approach.

Under field conditions, multiple potential damage zones in the caprock may be predicted. To mimic these conditions in the experiments, the injection ports numbered 1, 3, and 5 were defined as three potential leakage locations in the *base-model* (Figure 5.3). This allows the OD framework to select the best monitoring locations based on three anticipated leakage plumes in the system, which ensures the detection of the leakage event if any of these sources was eventually activated. In addition, to account for the possible uncertainty in the identified locations of the potential damage zones (i.e., structure and leakage-scenario uncertainties), three separate *base-models* representing three different realizations of the caprock fractures and injection ports (Figure 5.3) were developed to be run simultaneously under the optimization process. The first *base-model* included the actual locations of the fractures and injection ports (#1, 3, and 5) in the tank [denoted as Base-Frac(actual)], while the two other *base-models* included shifted locations of the fractures and injection ports (#1, 3, and 5) by 7 cm in the upward and downward directions of the tank [denoted as Base-Frac(up) and Base-Frac(Down), respectively]. It is noteworthy that the 14% imposed error in fracture locations ($\pm$7 cm shift) was chosen within the error range ($\sim$11-15%) obtained by Feng et al. (2018) while predicting the hydraulic and geometrical settings of the stress-induced fractures.
in Keshen reservoir. Moreover, the limited length of the confining caprock layer (100 cm) in the tank constrained the maximum allowed shift of the fractures and injection points in the base-models to 7 cm. In summary, three base-models, each includes different locations of the caprock fractures and injection ports, were developed to mimic the expected uncertainty in our predictions of the potential leakage sources in the field (i.e., leakage-scenario uncertainty). However, a large number of base-models might be required to represent the expected uncertainty in model structure (e.g., storage zone heterogeneity) and leakage scenarios (e.g., plume distributions) in actual CGS sites. A limited number of base-models were considered in this analysis to demonstrate and validate the developed OD framework with manageable computational resources.

5.3.3 Model Setup and Parametrization

The three base-models introduced in the previous section were developed as 2D transient flow and transport models using FEFLOW. The solution domain of the base-models was discretized with a high-resolution mesh (~190,000 triangular elements) to represent the heterogeneous sand packing of Zone 3 and the sand layers of Zone 2. The notion of explicitly representing the sand configurations in Zones 3 and 2 in the base-models is centered around the capability of seismic geophysical surveys in the field to map the deep overlying formation interfaces and the shallow aquifer heterogeneity with reasonable accuracy (Torp & Gale, 2004; Hermanrud et al., 2009; Zweigel et al., 2004). While Zone 1 was represented by a homogenous layer with an effective hydraulic conductivity value due to the difficulty of mapping its heterogeneity in the field (i.e., structure uncertainty), as discussed in Section 5.3.2. Figure 5.3 presents the three developed base-models in FEFLOW and the specified boundary and initial conditions to generate the three hypothetical leakage plumes seen in each sub-figure.
Figure 5.3 Developed base-models including the generated hypothetical plumes overlying soil properties; model boundary and initial conditions are presented.

Zonal parametrization technique was used to represent the spatial distribution of model parameters in the inverse problem formulated to calculate the sensitivity matrices ($J$ and $y$) required for solving Equation 18 in the OD framework. On the basis of a sensitivity analysis conducted by SENSAN utility in PEST/FePEST, hydraulic conductivity ($K$), transverse dispersivity ($\alpha_T$), and porosity ($\phi$) were found to have the largest influence on the hydraulic head and plume concentration observations in the system. Given that the absolute values of parameters do not contribute to the uncertainty analysis conducted by PREDUNC, average $K_s$ equal to 115 and 110 m/d were assigned to Zone 1 and the layers of Zone 2, respectively. However, externally determined $K$ values for each sand type, using column experiments (Table 3.1), were used to define the soil properties of Zone 3. In addition, the empirical formulas suggested by Neuman (1990b)
were used to specify the initial longitudinal and transversal dispersion coefficients in the inverse problem (\(\alpha_L = 0.1/0.12/0.12\) cm and \(\alpha_T = 0.07/0.05/0.05\) cm for Zones 1/2/3, respectively). The sand porosities were given a value of 0.4 (21 zones), and the six fractures were given similar \(K_s\) of 50 m/d. Table 5.1 presents the used model parameters with the specified high and low bounds (i.e., parameter uncertainty). To alleviate the computational burden of inverse modeling, BeoPEST under FePEST was used to parallelize the inverse problem over 8 PCs with i7-CPPUs @ 3.4GHz/4cores.

Table 5.1 Summary of parameter zones and assigned bounds.

<table>
<thead>
<tr>
<th>ID</th>
<th>Parameter</th>
<th>Zones</th>
<th>Parameter Bounds</th>
<th>Comment</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>K – Storage Zone</td>
<td>1</td>
<td>0.5 10</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>2</td>
<td>K – Overlying Formations</td>
<td>16</td>
<td>0.01 500</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>3</td>
<td>K – Shallow Aquifer (Sand #70)</td>
<td>1</td>
<td>6 50</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>4</td>
<td>K – Shallow Aquifer (Sand #40/50)</td>
<td>1</td>
<td>30 150</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>5</td>
<td>K – Shallow Aquifer (Sand #20/30)</td>
<td>1</td>
<td>80 250</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>6</td>
<td>K – Shallow Aquifer (Sand #20)</td>
<td>1</td>
<td>150 500</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>7</td>
<td>K – Shallow Aquifer (Sand #20)</td>
<td>1</td>
<td>250 1200</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>8</td>
<td>K – Gravel Section (BC1)</td>
<td>1</td>
<td>0.01 10000</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>9</td>
<td>K – Gravel Section (BC2)</td>
<td>1</td>
<td>0.01 10000</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>10</td>
<td>K – Gravel Section (BC3)</td>
<td>1</td>
<td>0.01 10000</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>11</td>
<td>K – Gravel Section (BC4)</td>
<td>1</td>
<td>0.01 10000</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>12</td>
<td>K – Fractures</td>
<td>6</td>
<td>0.01 10000</td>
<td>Log transformed</td>
<td>m/d</td>
</tr>
<tr>
<td>13</td>
<td>Injection Fluid Rate / 2D Elemental Area</td>
<td>3</td>
<td>1.00E-08 100000</td>
<td>-</td>
<td>m³/m²/d</td>
</tr>
<tr>
<td>14</td>
<td>Injection Mass Rate / 2D Elemental Area [(C_{initial} \times ) Injection Fluid Rate (ID #13)]</td>
<td>3</td>
<td>1.00E-08 100000</td>
<td>Tied to 13</td>
<td>g/m²/d</td>
</tr>
<tr>
<td>15</td>
<td>(\alpha_T) (Tank Zones #2 and #3)</td>
<td>21</td>
<td>0.0001 0.1</td>
<td>-</td>
<td>m</td>
</tr>
<tr>
<td>16</td>
<td>(\alpha_L) (Tank Zones #2 and #3)</td>
<td>21</td>
<td>0.0001 0.5</td>
<td>-</td>
<td>m</td>
</tr>
<tr>
<td>17</td>
<td>(\phi) (Tank Zones #2 and #3)</td>
<td>21</td>
<td>0.1 0.6</td>
<td>-</td>
<td>m</td>
</tr>
</tbody>
</table>

* Zones in this context refer to the zonal parameters of model variables defined for PEST and they are not related to the three zones of the tank.

The bounds of parameters presented in Table 5.1 were used by PEST to calculate the pre-calibration parameter uncertainties. Moreover, these bounds reflect our confidence on the available prior knowledge about each parameter. For example, the tight bounds given to the \(K_s\) of the sand
in Zone 3 were based on the fact that the shallow aquifer can be characterized more accurately than the deep formations. This is because many characterization methods that generally can be used in the shallow zones are not feasible and cost-prohibitive to be used in the deep formations. It should also be noted that as the injection concentration was fixed at $C_{\text{initial}}$, the injection mass rate was tied to the fluid injection rate through the relation, $C_{\text{initial}} = \frac{\text{injection mass rate/area (g/m}^2\text{/d)}}{\text{injection fluid rate/area (m/d)}}$. It should be noted that in the field, the relatively uniform concentration of native brine in the storage formation (i.e., $C_{\text{initial}}$) can be quantified during the preliminary site investigations.

5.3.4 Model Predictions

The chosen target predictions under the OD framework process, are the necessary parameters and model outputs to be known for developing an appropriate remedial system and identifying the likely areas to be impacted in the shallow zone. Fifteen different predictions were selected, including: (1) plume concentrations at five points in the shallow aquifer (Zone 3), (2) mass flux emanating from the three anticipated leakage sources to identify which of them was activated, creating the leakage event, (3) inflow to the storage layer (Zone 1), which is important for designers to be captured by the model to make accurate predictions of the plume pathway as discussed in Chapter 4, and (4) six effective $K_s$ of the caprock fractures. The outflow from Zone 1 was not included in the prediction list because the tank outlet at this zone (Figure 5.2) was switched off during the experiment due to some accumulated rust. Henceforth, the inflow in Zone 1 will be denoted as the storage zone flow.

In the validation process of the OD framework, these predictions were compared with experimental measurements to evaluate their accuracy and thus the selected monitoring locations’
informativity. Since the fracture $K_s$ cannot be directly measured in the soil tank, their estimated $K_s$ were evaluated indirectly by comparing the simulated and measured heads at four points located 15 cm downstream from the fractures (denoted as “validation heads”, henceforth). Note that these validation heads were not used in model calibration.

5.3.5 OD Framework Application

The OD framework was applied to essentially answer the question: “If the number of sensors that is feasible to be placed in the deep overlying formations was limited due to cost factors, what is the minimum number of sensors and their locations that can help to make the most accurate model predictions?”. To answer this question, eight testing scenarios were developed. A fraction that defines the number of monitoring locations in Zones 1 and 2 (hard-to-gather) with respect to the number of locations in Zone 3 (easy-to-gather) was used to define these scenarios as follows:

$$\frac{\text{Number of monitoring locations in Zones 1 and 2}}{\text{Number of monitoring locations in Zone 3}} = \frac{5 \text{ to } 15}{30, 40, 30, 40, 30, 40, 30, 40}$$

The maximum number of allowed monitoring sensors in the whole system under these scenarios (i.e., 55 sensors) was determined based on a trade-off analysis between the target predictive uncertainty reduction and the required total number of sensors (presented in Section 5.4.1). In the first six scenarios (from $\frac{5}{30}$ to $\frac{15}{40}$), the number of the costly deep sensors in Zones 1 and 2 (numerator) is significantly less than the shallow observation points in Zone 3 (denominator). However, the last two scenarios ($\frac{15}{0}$ and $\frac{20}{0}$) were included in this analysis to find the most important monitoring locations in the deep formations (Zones 1 and 2) for predicting the plume concentrations in the shallow aquifer before the plume reaches its bottom boundary. This means
no transport data will be available at the shallow aquifer (Zone 3) in these two scenarios. Such scenarios emulate the case when authorities want to identify the likely areas to be impacted in the shallow aquifer before the plume arrives at these zones for taking early remedial actions. This is an important design scenario to be considered to ensure the capability of the monitoring system to provide informative data to such early prediction of the plume distribution in the shallow aquifer.

The selection pool of monitoring locations ($X$) made available for the OD process contained data collected from 629 potential sampling ports and 30 manometers. The number of sampling ports considered in the OD process is less than the total number of sampling ports in the tank because all ports located outside of the transport domain (e.g., in the gravel sections and Zone 1) were disregarded. Concentration and hydraulic head data were assigned standard deviations of 5 ppm and 0.2 cm, respectively. Since the absolute measurements do not take a part in the uncertainty analysis by PREDUNC, the preliminary observation data assigned to PEST to calculate the sensitivity matrices used in Equation 18 were extracted from the numerical simulation of the hypothetical leakage events in the base-models. Predictions were assigned different weights ($\omega$) as follows: $\omega = 1.0$ for the shallow aquifer concentrations (5 predictions), $\omega = 1.8$ for injection mass fluxes (3 predictions), $\omega = 1.2$ for storage zone flow (1 prediction), and $\omega = 1.5$ for fracture hydraulic conductivities (6 predictions). Predicting the plume concentrations at the shallow aquifer was given the lowest weight as site characterization of this zone can always be revisited to enhance model predictions. The higher weights given to the mass flux and storage zone flow reflect the fact that identifying the leakage source and the hydraulic settings of the storage zone and fractures can be more critical for designing and developing remedial solutions.
5.3.6 Experimental Procedures

A brine leakage experiment was conducted in the intermediate-scale test system described in Chapter 3 for the validation of the design framework. Monitoring data collected during the leakage simulation was used for the transport model calibration. The model was then used to make the 15 predictions discussed in Section 5.3.4, and then they were compared with experimental measurements to evaluate their accuracies. Same BCs to those used in the base-models were applied in the experiment, except the injection rate and concentration of the NaBr tracer. Note that using the same BCs does not produce identical experimental and numerical results because the base-models were not calibrated yet. The applied injection mass flux and tracer concentration in the experiment were respectively higher than of those assigned in the base-models by 14.2% and 27.5% to mimic field conditions in terms of uncertainty in the predicted leakage settings using the geophysical and geomechanical modeling approaches discussed in Section 5.3.2.

A peristaltic pump (Cole-Parmer Masterflex L/S 7526-60) was used to inject the tracer solution through injection port #5 (Figure 5.2) at an average rate of 7.9 ml min$^{-1}$. Using one port for NaBr tracer injection imitates as if only one location from the three anticipated leakage sources was activated, creating the leakage event. Table 5.2 summarizes the applied BCs and the statistics of collected data from the monitoring locations including head and concentration observations. Concentration data was generated through chemically analyzing the collected aqueous samples from the selected septa-ports by the OD framework using an ion chromatography (IC) system (Dionex Aquion combined with autosampler Dionex As-Dv). The hydraulic heads were measured using the manometers selected by the OD framework as well, twice a day. It should be noted that the rapid increase in the water levels in the manometers right after injection was recorded by four high-resolution cameras that provide 1080p HD videos at 30 frames per second. This transient
head data is important for identifying the activated source through model calibration (will be discussed in Section 5.4.3).

Since the framework validation process relies on comparing the 15 model predictions discussed in Section 5.3.4 to their corresponding experimental measurements, the injection flow rate, inflow to the storage layer, and the validation heads were measured twice a day besides sampling the five prediction points in Zone 3. It is worth mentioning that extra aqueous samples were collected from the injection line to ensure the uniformity of the tracer concentration.

Table 5.2 Summary of boundary conditions and collected data.

<table>
<thead>
<tr>
<th>Experimental BCs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone 1 (BC1)</td>
<td>165.5 cm (Constant Head)</td>
</tr>
<tr>
<td>Zone 1 (BC2)</td>
<td>No-Flow (Outlet was switched off)</td>
</tr>
<tr>
<td>Zone 3 (BC3)</td>
<td>3.05 ± 0.08 mL/min (Horizontal Well)</td>
</tr>
<tr>
<td>Zone 3 (BC4)</td>
<td>156.6 cm (Constant Head)</td>
</tr>
<tr>
<td>NaBr Initial Conc. - No. Of Samples</td>
<td>93.7 ± 1.6 ppm - 23 Samples</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Collected Data</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NaBr Inj. Rate - No. of Measurements</td>
<td>7.9 ± 0.02 mL/min - 21 Measurements</td>
</tr>
<tr>
<td>Zone 1 Constant Head Inflow</td>
<td>16.61 ± 0.99 mL/min - 16 Measurements</td>
</tr>
<tr>
<td>Zone 2 Constant Head Outflow</td>
<td>20.6 ± 1.06 mL/min - 16 Measurements</td>
</tr>
<tr>
<td>No. of Head Readings</td>
<td>750 (330 transient data* and 420 after steady head)</td>
</tr>
<tr>
<td>No. of Aqueous Samples</td>
<td>239 samples analyzed by IC + 161 Tracer-free samples*</td>
</tr>
</tbody>
</table>

* Transient head readings are the values extracted from the HD video recorded at the beginning of injection. Tracer-free Samples mean samples with no color of the fluorescein dye.

5.4 Results and Discussion

In this section, the worth of data provided by the potential monitoring locations toward each prediction is discussed. In addition, the design procedures of the monitoring system and the OD framework validation results are presented.
5.4.1 Data Worth and Value Index Calculations

The analysis discussed in this section represents the steps #1, #2 and #3 of the OD framework application. The $DW$ of every potential monitoring location to each of the 15 predictions (Section 5.3.4) was estimated using Equation 19 within the sequential evaluation technique presented in Figure 5.1. To ease the presentation of $DW$ results, the 15 predictions were then categorized into four forecast classes based on their types and locations, as shown in Figure 5.4. The overall worth of every monitoring location to each forecast class was then calculated as follows:

$$DW_{overall,j} = \left( \frac{\sum_{1}^{m} \sum_{1}^{n} DW_{i,j}}{DW_{max}} \right) / 3$$

(23)

where $DW_{i,j}$ is the worth of data collected from a specific monitoring location/s $j$ to make prediction $i$, $n$ is the number of predictions under each forecast class and $m$ is the number of base-models. $DW_{overall,j}$ ranges from zero to one. In Figure 5.4, all potential concentration (dots) and head (crosses) observation points were color-coded based on the calculated $DW_{overall,j}$ to identify the most critical group of monitoring locations to each forecast class.

$DW_{s overall,j}$ presented in Figures 5.4a and 5.4b show that concentration data is more important than head data to predict the leakage mass flux at the source and plume distribution in the shallow aquifer. However, $DW_{s overall,j}$ in Figure 5.4c illustrate that the prediction of the storage zone flow relies merely on head data from both Zone 1 and the leakage vicinity. The forecast of fracture $K$s depend on both head and concentration data in Zone 2 as indicated by the $DW_{s overall,j}$ in Figure 5.4d. This can be attributed to the sensitivity of the hydraulic heads in Zone 2 to the fracture $K$s, which in turn controls the velocity field that drives the advection transport of
the plume (refer to Chapter 4). Figures 5.4b and 5.4d show that data from relatively shallow depths (i.e., the deepest part of Zone 3) are still important for making predictions in deep formations such as fracture $K$s and leakage mass flux. In contrast, the prediction of plume distribution at the shallow aquifer requires data near the prediction points (Figures 5.4a). Moreover, slight differences were found between the worth of the shallow monitoring locations (in Zone 3) to predict plume concentrations in the shallow aquifer.

(a) Shallow Aquifer Concentrations (5 Predictions)

(b) Injection/Leakage Mass Flux (3 Predictions)

(c) Storage Zone Flow (1 Prediction)

(d) Fracture Hydraulic Conductivities (6 Predictions)

Figure 5.4 The overall worth of potential monitoring locations to the four combined categories of predictions calculated using the sequential selection technique. Indicative contour lines were included to delineate the highest DW zones.
In order to design a cost-effective monitoring system, we need to find the minimum number of sensors that can provide an acceptable reduction in the uncertainty of all predictions. For this task, the monitoring locations were sequentially disregarded from the complete potential observation set, based on their worth to each prediction, while computing the corresponding gradual increase in the predictive uncertainty. Figure 5.5 shows three plots of the increased predictive uncertainty as a response to the disregarded monitoring data in the three base-models. These plots revealed that using only 7-10% from all potential monitoring locations (i.e., 45 to 65 ports) can lead to 60-80% of uncertainty reduction in most predictions, as highlighted by the green and orange boxes. Therefore, the maximum allowable number of monitoring locations in the developed eight testing scenarios in Section 5.3.5 was set equal to 55.

Figure 5.5 The increase in uncertainty of all predictions, as a result of sequentially disregarding monitoring location data in the three base-models. Conc, Inj, and Frac, respectively, refer to the five concentration prediction points in the shallow aquifer, the injection flux at the three considered leakage sources, and the six fracture permeabilities.
The plots in Figure 5.5 suggest that the leakage mass fluxes and fracture $Ks$ (black and purple lines in Figure 5.5) are the highest data-demanding predictions; their uncertainties increase slowly with data removal from multiple depths. That is due to the fact that both predictions are being informed by the head and concentration data over almost the entire system domain (see Figure 5.4b and 5.4d). While the prediction of the plume concentrations in the shallow aquifer (red lines) showed a considerably low density of the required observation points. The difference in the steepness of the uncertainty variation with reduced monitoring locations in the three plots indicates the significant impact of the plume distribution relative to monitoring locations on the informativity of collected data. For example, the prediction of the leakage/injection rate (black lines) under the Base-Frac$_{\text{actual}}$ model required a larger number of data points than that of those needed under the Base-Frac$_{\text{up}}$ and Base-Frac$_{\text{down}}$ models. Finally, the calculated $DWs$ using Equation 19 were substituted in Equation 20 to calculate the Value Index (VI) of each monitoring location toward all predictions for preparing the initial guess vector of the best monitoring locations to be used in the GA.

### 5.4.2 Monitoring System Design

The GA was then applied as the fourth and final step of the OD framework to select the best monitoring locations that maximize the system VI without exceeding the allowed number of sensors at each zone defined by the eight testing scenarios discussed in Section 5.3.5. For each testing scenario, four optimization runs were performed using the three base-models simultaneously. The GA in two of these runs was supported by the initial guess (IG) prepared by applying the sequential evaluation technique. However, the other two runs were based on randomly sampling the initial monitoring locations from the $X$ vector (Section 5.2.5). The four monitoring system designs proposed by the OD framework under these four optimization runs were evaluated.
by calculating the ratio between their VIs and the VI obtained by employing the data from all potential monitoring points in the uncertainty analysis (Figure 5.6a). When viewed collectively, the results showed that a VI ratio of at least 0.40 was achieved by all proposed designs under these four runs, as shown in Figure 5.6a. Notably, the high VI ratio obtained under the Base-Frac\textsubscript{(up)} model compared to the other two base-models can be attributed to the larger number of monitoring locations encountered by the core of the plume. The proximity between monitoring locations and the plume core results in high observation-to-parameter sensitivity and thus high data-to-prediction informativity, which leads to a large reduction in predictive uncertainty.

Figure 5.6 (a) Value Index of the selected set of monitoring locations relative to the full dataset, and (b) the final implemented design in the experiment. GA and IG refer to the Genetic Algorithm and Initial Guess, respectively.
Figure 5.6a shows, in general, that using the GA with an initial guess leads to a slightly higher performance, except in scenarios $\frac{15}{0}$ and $\frac{20}{0}$. This indicates that the GA, under the number of trials used in this analysis is efficient in exploring small-sized solution spaces without the need of initial guess, as in scenarios $\frac{15}{0}$ and $\frac{20}{0}$, where the number of potential monitoring locations was limited to 267. However, in the large-sized solution spaces, the rest of the scenarios where the number of potential monitoring locations reaches 629, providing the GA with an initial guess enhanced its performance by approximately 10%.

The final design was a blend of all selected monitoring locations under scenarios $\frac{15}{30}$ and $\frac{20}{0}$ (defined in Equation 22). The criteria for this design included: (1) minimizing the total number of monitoring points by sacrificing the minor VI increase under scenario $\frac{15}{40}$ (i.e., selected locations under scenario $\frac{15}{30}$ were considered), and (2) acquiring data capable of informing the prediction of plume concentrations in the shallow aquifer at an early stage of leakage development (i.e., selected locations under scenario $\frac{20}{0}$ were considered). Figure 5.6b shows the layout of the final design implemented in the soil tank. It is worth mentioning that the VI percent ratio of this final design ranged between 50-77%.

The impact of the assigned prediction weights, design constraints, and data informativity (see Figure 5.4) on the selected monitoring locations by the OD framework can be clearly seen. For example, the selected monitoring locations in Zone 3 under such final design have the highest informativity to predictions that were given the highest weights, such as injection mass flux and fracture permeabilities (see Figure 5.4). While the selected monitoring locations in Zone 3 were
relatively less informative to the prediction of plume concentrations in the shallow aquifer that was given the lowest weights.

5.4.3 Leakage Source Conditions Analysis

To identify the activated leakage source conditions and the most probable fracture locations, the three base-models were employed in three separate inverse problems that only included the parameters of leakage/injection mass fluxes and fracture permeabilities (9 parameters in total). In this analysis, the Base-Frac(actual) model exhibited the lowest error in the objective function for concentration data (45% lower than the other base-models), which indicated that this model represents the most likely correct locations of fractures and injection ports. Therefore, the Base-Frac(actual) model was then used exclusively in the further steps of the validation process (i.e., to make the predictions used in the framework validation). Moreover, in the same analysis, the highest injection/leakage mass flux was predicted for injection port #5 in the three base-models (3 orders of magnitude higher compared to injection ports #1 and #3, which were given almost zero mass flux values by PEST). Sources with significantly low predicted mass fluxes (ports #1 and #3) were then considered inactive and were omitted from the inverse problem pertained to the validation process. By omitting these parameters from the inverse problem, the total number of considered predictions in the validation analysis became eleven. Such analysis shows that PEST can identify the most likely leakage sources and fracture locations in the caprock using data collected from the optimum monitoring locations found by the OD framework. This indicates that monitoring data can be used to eliminate the unrealistic base-models from the solution ensemble (Tarantola et al., 2006), which will reduce the structure and scenario uncertainties in the problem and thus enhancing the prediction accuracy.
5.4.4 OD Experimental Validation

The comparison between model prediction and experimental measurements are presented in Figure 5.7 as box plots showing the maximum, minimum, 25th, 75th percentiles and the mean of the normalized percent errors (NPE) of model predictions; NPE = \[ \frac{P_m - P_s}{\max(P_m)} \times 100 \], where \( P_m \) and \( P_s \) are the measured and simulated predictions, respectively. Figure 5.7 also presents the root mean square of model-to-data errors (RMS).

To expand this validation analysis, prediction accuracy was tested under three data availability scenarios, reflecting different stages of plume development: the first scenario represents the early leakage stage where concentration data would be only available from the monitoring locations close to the source at Zones 1 and 2 (denoted as Scenario 1), the second scenario mimics the later stage of plume development when it reaches the lowest part of Zone 3 as highlighted in Figure 5.6b (denoted as Scenario 2), and finally, the third scenario represents the stage at which the plume reaches the near-surface shallow areas (denoted as Scenario 3). It should be noted that a full hydraulic head dataset from zones 1, 2, and 3 was used under these three scenarios because leakage pressure signals can transmit extremely fast, resulting in head changes in the shallow aquifers at the early stages of the leakage.

Predictions of Fracture Permeabilities:

Results presented in Figure 5.7 show that under the three data availability scenarios, the calibrated model was able to predict the hydraulic head of the four manometers located close to the source with significantly low mean NPE (<0.1%). This indicates a low sensitivity of the head predictions to data density, which was confirmed by Figure 5.4 as well. However, the calibrated
fracture $K_s$ under the three scenarios were found to be non-unique, indicating the importance of both shallow and deep data to find the optimum values of fracture $K_s$ (Figure 5.4).

![Normalized Percent Error (NPE) in model predictions under the three data availability scenarios.](image)

**Figure 5.7** Normalized Percent Error (NPE) in model predictions under the three data availability scenarios.

**Predictions of Plume Concentrations at the Shallow Aquifer:**

The best prediction of plume concentrations at the shallow aquifer (mean NPE ≈10-30% except for the prediction of the plume concentrations at point #5) was made under Scenario 3, in which all collected shallow and deep data were used to calibrate the model. Despite the larger number of data points used in Scenario 2 (data from Zone 2 and lowest part of Zone 3) compared to Scenario 1 (data from Zone 2 only), the prediction of the plume concentrations at the shallow aquifer made under Scenario 1 was found to be more accurate. That can be attributed to two
reasons; first, the low to moderate informativity of the additional data used in Scenario 2 to the shallow aquifer concentration predictions (refer to Figure 5.4), which propagated more errors than reducing predictive uncertainty. Moreover, such additional data may have affected the degree of interdependency between the information content of the present data, which decreased the data informativity. Second, there were two distorted breakthrough curves among the additional data used under Scenario 2 caused by local sand-packing imperfections, which enlarged the margin of added errors to predictions. The prediction accuracy of the plume concentrations at prediction point #5 (Figure 5.4) was the lowest in the three data availability scenarios due to the large distance between point #5 and the shallowest data used in this analysis (Figure 5.6b).

It is important to note that even with the resultant prediction errors of the plume concentrations in the shallow aquifer, the model was able to qualitatively identify the likely areas to be impacted by the leakage under all scenarios. This prediction error can always be minimized by providing the model with more accurate prior information about the shallow aquifer heterogeneity, hydraulic and transport properties from relatively inexpensive field investigations. Moreover, there is an unreducible margin of predictive uncertainty due to the inherent randomness in all natural processes (Walker et al., 2003), that should be factored in during the evaluation of data informativity.

Predictions of Storage Zone Flow and Leakage Mass Flux:

The predictions of the longitudinal flow in the storage zone and leakage flux showed increasing accuracy when using calibration data obtained at locations closer to the source. This can be attributed to the worth of data used to make these predictions. For example, the prediction of the storage zone flow is informed merely by the head and concentration data at the source vicinity (deep zone data), see Figure 5.4. Since the three tested scenarios included the same deep
zone data, the accuracy of the flow prediction is controlled mainly by the margin of added error by the non-informative data in each scenario (shallower zones data). Thus, Scenario 1, where the lowest density of non-informative data resided, exhibited the highest prediction accuracy of the storage zone flow and leakage flux (mean NPE ≈ 6% and 5.7%, respectively).

Based on the above results, it is recommended to adopt an iterative stage-wise calibration technique for base-models as follows: (1) calibrate the model based on all available data to estimate initial values for model parameters; and (2) for each prediction of interest, recalibrate the model using the data that yields the highest worth to this specific prediction. Moreover, the above analysis indicates that randomly adding observation points to the monitoring system does not guarantee higher prediction accuracy, particularly if the added data may yield a considerable margin of error, particularly if the added data may yield a considerable margin of error.

In actual CGS sites, data on deep geological and hydraulic settings will always be limited, which can affect the optimal selection of the monitoring locations by the OD framework. Such data scarcity is typically addressed using probabilistic approaches where multiple realizations of the system properties are generated based on the retrieved geo-statistical prior knowledge from the available data. In the current analysis, such data limitation was represented by developing three base-models to account for the uncertainty in the fracture structural settings. However, under actual field conditions, numerous realizations might be needed to represent the large-variance multidimensional probability function of the system properties. That, in turn, will boost the computational burden of the optimization process, especially for the sophisticated models simulating the leakage transport in multiphase flow fields, because at least one forward model run is needed for each realization of the system properties to derive the predictive uncertainty for each model. In this context, Aitokhuehi & Durlofsky (2005), Van Essen et al., (2009) and Wang et al.,
(2012) suggested different methods to handle multiple geological realizations under a single optimization problem. Similarly, in the current analysis, the three base-models were run simultaneously under a single inverse problem to ensure that the calculated sensitivity matrices are informed by all possible system properties. That is an important consideration to develop a monitoring system that can potentially work under the uncertain conditions of field data.

5.5 Summary

This chapter presented the development and testing of an optimal design framework that combines the uncertainty analysis tool in the inversion code PEST with the Genetic Algorithm, to find the best locations of a minimum number of deep sensors for brine leakage monitoring (Hypothesis 2). Due to practical and financial challenges, testing the design framework in the field is not feasible. Thus, the framework was experimentally validated in an intermediate-scale testing tank for potential field implementation. The validation involved the following steps: (1) acquiring data from the monitoring system installed in the soil tank, where brine leakage was simulated from the storage zone to the shallow aquifer, (2) calibrating the transport numerical model using the acquired data, and (3) comparing the calibrated model predictions with experimental measurements. The predictions of interest to this research included the leakage rate, storage zone flow, and plume concentrations in the shallow aquifer, besides identifying the leakage source and fracture permeabilities. The results demonstrated that the developed monitoring system using the proposed framework was able to detect the leakage and provide useful data to identify the source location and make plume predictions with 30% normalized error.
6.1 Introduction

$\text{CO}_2$ geological sequestration (CGS) into deep saline formations is a promising technology applied for minimizing the amount of released $\text{CO}_2$ to the atmosphere causing global warming (IPCC, 2005; Lackner, 2003; Bachu, 2003; U.S. DOE, 2015). The pressure build-up in the storage formation during $\text{CO}_2$ injection can affect the caprock integrity, which can lead to pressure-driven leakage of native brine or $\text{CO}_2$, posing a contamination risk on the underground sources of drinking water (USDWs). This leakage can occur via existing or pressure-induced conductive pathways, such as activated faults, caprock fractures, abandoned wells, and small rock (Gasda et al., 2004; Rutqvist et al., 2007; 2008; Celia et al., 2011). The pressure increase in the storage formation can reach far distances from the injection site due to the rapid propagation of the pressure waves (Birkholzer et al., 2009), which makes investigating the potential leakage pathways complex and costly (Oldenburg et al., 2016).
In order to reduce the potential geomechanical risks of the storage zone pressure build-up on the caprock integrity (e.g., induced seismicity and fault activation), and the potential contamination of the USDWs, a brine extraction technique was proposed in different studies as an efficient pressure management method (Buscheck et al., 2011; Bergmo et al., 2011; Birkholzer et al., 2012; Buscheck et al., 2014). Brine extraction can also increase the future CO₂ storage capacity, limits the area of site characterization and risk assessment, decouple neighboring CGS operations by minimizing pressure interference and enhance CO₂ trapping (Birkholzer & Zhou, 2009; Buscheck et al., 2011; Court et al., 2012; Cameron & Durlofsky, 2012). Extracted brine can be used as a feedstock for desalination plants and cooling/heating facilities or can be reinjected in different over-or underlying saline formations for disposal purposes (Buscheck et al., 2011; Court et al., 2012).

In recognition of the high cost of brine handling process (Sullivan et al., 2013), Birkholzer et al., (2012) and Cihan et al., (2015) explored different optimization approaches to minimize the extraction volume of brine to maintain the pressure at stressed faults below critical values. The variables in their optimization problem were the locations and extraction rates of the wells. González-Nicolás et al., (2019) emphasized that the lack of accurate information about the subsurface geological settings can affect the optimization results leading to a poor design of the extraction system and thus not achieving the needed pressure control in the storage zone. Such uncontrolled pressure build-up can cause different modes of failures in the caprock, including fault activation, fracture opening, dilating, and intersecting with other fractures (Lei et al., 2017; Detwiler & Morris, 2019), which poses a high leakage risk on the USDWs.

Based on the above, designing the extraction system to control the potential leakage from the storage formation instead of just managing the pressure build-up can be more critical for the
environmental safety of CGS operations (Court, 2011). Moreover, developing an emergency remedial response strategy for CGS projects is one of the EPA requirements (U.S. EPA, 2010). Leakage control in this context refers to partially reversing the advective leakage flux back to the storage formation and utilizing the dilution capacity of the overlying formations to dilute the residual fraction of the leakage mass. To design such system, the optimization problem should be formulated to constrain the leakage concentrations in the shallow aquifer on the maximum allowed limits by the EPA (Siirila et al., 2012). The question that might be raised here is, “can the leakage-control wells be utilized to manage the pressure build-up in the storage reservoir prior to the onset of the leakage event?”. Buscheck et al., (2011) and Court et al., (2012) highlighted the relative flexibility in locating the pressure relief wells in the storage formation due to the rapid propagation of the pressure waves in the subsurface systems. Therefore, leakage-control wells have the potential to be used in reducing the reservoir pressure even if they were not placed in the most optimal locations for this purpose.

Numerous simulation-based optimization approaches can be found in the literature for resolving different groundwater problems, such as optimizing the number of observation wells (Wöhling et al., 2016; Vilhelmsen & Ferré, 2018; Zell et al., 2018; Safi et al., 2019) or finding the best design of a groundwater supply scheme (Ahlfeld & Heidari, 1994; Mantoglou et al., 2004). In the context of CGS management, Birkholzer et al., (2012) applied a derivative-based method to find the optimum extraction scheme to control the reservoir pressure. Cihan et al., (2015) argued that the irregularity of the objective function in the optimal well placement problems can lead to suboptimal solutions if derivative-based methods were employed. Therefore, Cihan et al., (2015) and González-Nicolás et al., (2019) relied on the global optimizer of the Differential Evolution Algorithm to find the best locations and rates of the pressure relief wells. Cameron and Durlofsky
(2012) employed the algorithm of Hooke-Jeeves Direct search to place the injection and production wells in optimal locations that enhance the brine cycling process needed for minimizing the mobile fraction of \( \text{CO}_2 \) in the storage formation. Zhang & Agarwal (2012) and Liu et al., (2016) incorporated the Genetic Algorithm (GA) with the multi-phase simulator TOUGH2 to optimize the injection pressure, rate, and depth to reach the maximum \( \text{CO}_2 \) storage capacity with minimum environmental adverse impacts. Recently, Liu et al., (2020) used the GA to optimize the pumping rate and well placement of one \( \text{CO}_2 \) injection well and four production wells to enhance the CGS operation in a \( \text{CO}_2 \) enhanced gas recovery site. In other applications, the GA was also successfully used for well placement optimization problems related to oil and gas development (Bangerth et al., 2006; Janiga et al., 2019a; 2019b).

Field testing and validation of design approaches, including the one presented here, are technically and financially unfeasible. While also, relying solely on using numerical tools to test these approaches can make the study findings inconclusive due to the disregarded inherent uncertainties in field measurements and system prior-knowledge in numerical testing. Thus, an integrated approach that combines between experimental and numerical methods was used for the validation and testing effort in this chapter. An intermediate-scale laboratory experiment was conducted to examine the applicability of using the developed design approach under field-like uncertainty conditions. This approach incorporates the GA with a FEFLOW-based transport model (Diersch, 2014) to find the best well locations in the \( \text{CO}_2 \)-storage formation for controlling the potential brine leakage by extracting the minimum brine volume. After experimental validation, a numerically generated study was used to evaluate the practicality of employing the brine extraction technique as a mean to control brine leakage under field-relevant scenarios.
The validation experiment of this analysis was conducted in the soil tank described in Section 3.1.3. In this tank, the migrating brine from the storage formation to the shallow aquifer was controlled by an extraction system designed based on the GA optimization analysis. The validity of the approach was tested by evaluating whether the plume concentrations observed in the experiment matched with predefined constraining limits in the optimization problem. However for the numerical testing, a hypothetical brine leakage plume from the Pond fault at the Vedder formation in the southern San Joaquin Basin, California, was controlled by an extraction well selected and operated based on the GA results. The Vedder formation is a potential CO₂ storage reservoir with a target injection rate of 5 MT CO₂/year for a period of 50 years (Zhou & Birkholzer, 2011; Birkholzer et al., 2011; Wainwright et al., 2013). Deep extraction is considered a practical technique for field implementation if the extraction rate required to control the potential brine leakage is less than the CO₂ injection rate needed to meet the target storage plan.

6.2 Optimization Method

The GA modified by Deep et al., (2009) and rooted in the global optimization toolbox in Matlab 2019b (MathWorks, 2020) was used in this work. Simulation-based optimization relies on running the forward model several times until the global minimum of an objective function is found, and a set of constraints is satisfied. The optimization problem entails finding the best locations and pumping rates of the extraction wells that can control the potential leakage from a CO₂ storage zone by extracting the minimum amount of formation brine. Therefore, the objective function including the problem constraints used for this analysis can be expressed as:

\[ \min_{u \in \Omega} f(u), \quad \text{subject to} \quad w. [C(u) \leq C_T] \]  

(24)
where, $J$ is the objective function to be minimized representing the total pumping rate from all selected wells, and $u$ is a matrix of the optimization variables including well locations ($W$) and pumping rates ($Q$), where $\Omega$ refers to the variable bounds. $C$ is a vector of the simulated leakage concentrations at specific areas in the shallow aquifer, subjected to a weighted constraint of the target leakage concentrations ($C_T$) at the shallow aquifer. $C_T$ can be defined based on the maximum allowed concentration of a regulated contaminant by EPA (Siirila et al., 2012). Notably, the constraint weights ($w$) give us the flexibility of assigning different priorities of any specific zones of the shallow aquifer (e.g., close to drinking water wells) to meet the concentration constraints. This option is vital for distinguishing between the potable and non-potable shallow aquifers in the optimization problem.

To limit the solution space dimensions, all potential locations of the extraction wells in the storage zone are assumed known and can be predefined in the model domain. This assumption was considered practical because one cannot place a deep well at any location across the system domain due to technical and regulatory constraints. In addition, the reduced dimensionality of the solution space will minimize the computational burden of the optimization process. Therefore, the selection pool, made available for the GA, is composed of integer values representing the well IDs (e.g., from 1 to 100), where every well ID corresponds to a certain well location in the system. In contrast, the pumping rates were treated as non-integer variables to find the lowest possible pumping rates. Moreover, a non-linear constraint of the brine plume concentrations at the shallow aquifer was added to the optimization problem. Such constraint was enforced by considering any suggested solution that violates the predefined concentration limits infeasible and should be penalized. For more details on the adapted penalty function, refer to Deb (2000). Similar to Cihan...
et al., (2015), the costs of extraction, handling, and treatment of the brine were assumed proportional to the total pumping rates of the extraction system.

The key steps of the GA procedure to perturb \( u \) and find the global minimum of Equation 1 while satisfying the assigned constraints are provided in Figure 6.1. The first population matrix \( (P_1) \) of the proposed solutions \( (u_v) \) generated by the GA has a dimension of \( v \times k \), where \( v \) is the population size (\( v=50 \) in this analysis) and \( k \) is the number of variables. Notably, \( k = 2 \) means one extraction well is desired, as each well is defined by one location and one pumping rate. The variables \( W \) and \( Q \), composing the \( u \) matrix, are randomly sampled by the GA within the defined bounds \( (\Omega) \) for Equation 1 to construct the \( P_1 \). The proposed solutions \( (u_v) \) are ranked based on the value of the objective function \( J(u_v) \), which is basically a penalty function to distinguish between feasible and infeasible solutions (Deb, 2000). Ranked solutions are then processed via the standard producing function of the GA, including the selection, mutation, and crossover functions, to construct the next population matrix \( (P_1) \). The GA setup assumptions used in this chapter are as follows: 40% of the prior population is maintained in the next population as elite candidates, the two top-ranked candidates are treated as parents for the production of 15% of the offspring using Laplace crossover function, while the rest of the offspring is produced using a Power mutation function (Deep et al., 2009). Every time the GA generates a new population matrix, the solution ranks of the two previous populations are compared to select the elite candidates and parents of the new population matrix. The GA keeps on looping until one of the predefined stopping criteria is met; either the resulted optimum values are almost the same over the last 20 trials or the number of generated populations reaches the maximum defined limit (500 in this analysis).
6.3 Experimental Validation

In this section, the optimization method was evaluated and then experimentally validated through the following steps: (1) a pre-constructed soil tank, where brine leakage events can be mimicked, was modified for this analysis, (2) a transport model was developed in FEFLOW to simulate such brine leakage events and act as a testing bed for the extraction schemes explored by the GA, (3) an optimization problem was formulated by applying Equation 1 on the tank system through the GA using the developed transport model as a forward simulator, (4) a sensitivity analysis of the optimization results towards the variation in the system parameters was conducted, (5) the optimality degree of the GA results was tested using the forward simulations, and (6) finally the selected extraction system by the GA was implemented in the tank and its performance was evaluated by checking whether the controlled leakage plume violates the constraining concentrations at the shallow aquifer.

The section starts by a brief description of the laboratory test system and the formulated optimization problem to design the brine leakage control wells. Thereafter, the results of the optimization process, sensitivity analysis, optimality check and the experimental validation are presented and discussed.
6.3.1 Experimental Setup

The horizontal tank described in Section 3.1.3 was used in this analysis as a testing bed to evaluate the efficacy of the designed extraction system to control the simulated brine leakage in the tank. As discussed in Section 3.1.3, three distinct zones were included in the soil tank representing the storage zone (Zone 1), intermediate overlying formations (Zone 2), and the shallow aquifer (Zone 3) in the field. Zones 1 and 3 were packed to include spatially correlated permeability fields, while Zone 2 was packed as a layered system. Finally, four boundary conditions (BCs) were assigned at the four ends of the tank to generate multi-dimensional and interactive flow field between the three district zones of the tank as expected under field conditions. Figure 6.2 illustrates a schematic of the conceptualized upward brine leakage in the field and the horizontal soil tank used to simulate the vertical brine leakage. Figure 6.2 also shows the packing configuration, placed instrumentation and the experimental BCs.

The top boundary of Zone 1 was connected to a peristaltic pump (Neuman BC), while both boundaries of Zone 3 were connected to constant head reservoirs (Dirichlet BC). The bottom boundary of Zone 1 was assigned no-flow BC. By adjusting these BCs, the flows in the tank and across Zone 2 were controlled; the flow in Zone 2 is defined by the head difference between Zones 1 and 3. Out of the 30 manometers in the tank, the nine manometers located in Zone 1 (Section 3.1.2) were designed to be used for both hydraulic measurements and fluid extraction if selected by the GA. Seventeen additional extraction wells were placed in Zone 1 so that each of them was within a different sand type in the random heterogeneous packing. This makes the total number of potential extraction wells located at Zone 1 equal to 26 (as shown in the inset part of Figure 6.2).
Figure 6.2 Schematic depicting (a) brine leakage migration from a CO₂ storage formation under natural settings, (b) experimental setup including packing heterogeneity, tank instrumentation and the potential locations of extraction wells. Note the flow occurs from right to left.

The tracer was prepared using 20 grams of NaBr salt (Mallinckrodt 0535-102.9 NaBr/mol) with 0.03% (by volume) of fluorescence dye (ACROS-Organics 17324 -5000) and mixing with ~171.6 liters of tap water. This resulted in a concentration of NaBr tracer around 101 ppm. For
more details on the problem conceptualization, experimental design, tank construction, and packing procedures, refer to Sections 3.1.3 and 3.1.4.

6.3.2 Modeling the Simulated Leakage in the Tank

In the application of simulation-based optimization, a numerical model is needed in conjunction with the GA for performing forward simulations. The finite element transport code of FEFLOW (V 7.1) was used to develop a 2D flow and transport model to simulate the leakage control experiment. The tank domain was discretized in the developed model using 281,591 elements to obtain a high-resolution velocity field in Zone 1 and source vicinity for convergence enhancement of the optimization problem. Steady-state (SS) injection and transport conditions were assumed in the developed model for two reasons: (1) minimizing the computational burden of the optimization process; the forward model must be run thousands of times by the GA to find the optimum solution, and (2) focusing on reproducing the most critical distribution of the plume over Zones 2 and 3, which happens when the plume occupies the entire area of the tank after reaching SS conditions (refer to Chapter 4).

Birkholzer et al., (2012), Cihan et al., (2015), and González-Nicolás et al., (2019) recommended designing the extraction wells after implementing the CGS monitoring system so that more accurate data can be collected for model calibration and system characterization. By assuming the same degree of data availability in this analysis, it is valid to explicitly represent the sand heterogeneity in the three tank zones in the model. The initial hydraulic conductivities (Ks) of the different sand types and caprock fractures were retrieved from the column-experiment values presented in Section 3.1.4. The sand porosities (ϕ) in the three zones were assigned an average value of 0.4. The longitudinal and transversal dispersion coefficients (α_L=0.14 cm and α_T= 0.01 cm for the entire domain) were specified in the model based on the empirical formulas.
of Neuman (1990). The $K$-field of the model was then calibrated using the pre-injection data of hydraulic heads, inflows, and outflows (i.e., mimicking the collected data prior to onset of leakage) using the inversion code of PEST (Doherty et al., 2010; 2015). The calibrated model exhibited a good fit in the simulated heads and flows with low root mean square errors (RMS) (i.e., RMS=2.87 mm for hydraulic heads and RMS= 0.02 m$^3$/d for the tank in- and outflows).

![Image](image_url)

**Figure 6.3** The three forward models, used by the GA in the optimization problem, including the boundary conditions, potential extraction wells, calibrated $K$-field and the location of the concentration constraining points.

The latest advances in geophysical techniques have enabled us to predict the hydraulic properties of pressure-driven fractures in a rock system using seismic data and calibrated geomechanical models (Ingram & Urai, 1999; Ligtenberg, 2005; Feng et al., 2018). By applying similar techniques in CGS sites, the location of the future damage zones in the caprock can be
predicted, and the potential leakage rate via these fractures can be inferred. Therefore, in this analysis, the potential leakage sources (i.e., the injection ports in the experiment) were predefined in the forward model as mass influx elements with an assumed injection rate of 7.9 mL/min and tracer concentration of ~78 ppm. However, it should be noted that the used tracer concentration in the experiment was slightly higher than what was used in modeling to introduce some uncertainty in the mass flux described in the model as expected in actual leakage problems.

Multiple potential damage zones and leakage sources can be predicted during geophysical and geomechanical investigations. Due to the high installation cost of deep extraction systems, the optimization algorithm need to find the best well locations and extraction rates that can control the leakage from any of the identified potential leakage sources. Therefore, in this analysis, three forward models with different hypothetical leakage scenarios were set to run simultaneously under the GA to find the optimum extraction scheme that can control any of these anticipated leakage scenarios. Figure 6.3 presents the three hypothetical leakage scenarios considered in the optimization process. The figure also shows the locations of all potential extraction wells, assigned BCs, and the calibrated $K$-field using the pre-injection hydraulic head and flow data.

### 6.3.3 Optimization Problem Setup for the Experimental Problem

The GA was applied to select the best two wells out of the 26 potential well locations that can control any leakage occurring from the three considered leakage sources (i.e., injection ports number 1, 3, or 5). The GA was also set to allow for selecting only one well if this solution exhibited the lowest pumping rate to control the leakage plume. A total number of 32 constraining points were prescribed in Zone 3 to regularize the objective function by defining the maximum allowed contaminant concentrations in the shallow aquifer; every 16 points were vertically distributed at $X=5.76$ and $X=6.83$ m (marked by the dashed blue line in Figure 6.3). These
constraining points are denoted as “Constraining Group #1 and #2” in Figure 6.2 and henceforth in the text. The locations of these constraining groups were chosen to represent the deepest and mid-level strata of the shallow aquifer in the field. Various target concentrations ($C_T$) of the leakage plume at the shallow aquifer with different constraining weights ($w$) were assigned to these constraining groups; $w = 0.5$ was given to the points in the constraining group #1, while $w = 2.0$ was given to the points in the constraining group #2. Points in constraining group #2 were given a higher weight in the objective function because of their proximity to the shallow aquifer, which makes them more critical to be satisfied by the GA solution. Figure 6.3 shows the ID of each potential extraction well and the location of the 32 concentration constraining points relative to the three leakage sources considered in the forward models.

### 6.3.4 Sensitivity Analysis of the Optimization Results

Six testing cases underlining multiple scenarios were considered in this analysis to examine the sensitivity of the optimization results to fracture permeabilities and storage zone heterogeneity, due to the typical lack of information about these parameters in the field (Table 6.1). In case #1, the calibrated K-field on the pre-injection hydraulic data was used reflecting some knowledge about the system heterogeneity and fracture settings. However, from the second to the fourth case, the K-fractures were assumed uncertain; cases #2 and #3 included K-fractures equal 1 and 500 m/d, while the twelve scenarios under case #4 included K-fractures that were randomly sampled from 1 to 500 m/d. The 20 scenarios under cases #5 and #6 were developed to address the effect of uncertainty in storage zone heterogeneity on well selection and optimal pumping rates. To develop these 20 scenarios, conditional and unconditional geostatistical simulations of Zone 1 heterogeneity were performed using the sequential indicator simulator (SIS) in SGeMS (Deutsch,
2006; Remy et al., 2009). Notably, a constraint of $C_T \geq 0.0$ ppm was imposed on all the above scenarios.

Table 6.1 Studied Case Scenarios and Optimization Results.

<table>
<thead>
<tr>
<th>Case Scenarios (C_T \geq 0.0 ppm)</th>
<th>Number of Scenarios</th>
<th>1st Well# (Freq %)</th>
<th>2nd Well# (Freq %)</th>
<th>µ.Q1* (SD) (mL/min)</th>
<th>µ.Q2* (SD) (mL/min)</th>
<th>µ.QT** (SD) (mL/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Calibrated K-Frac.</td>
<td>1</td>
<td>24 (--)</td>
<td>--</td>
<td>25.64 (--)</td>
<td>--</td>
<td>25.64 (--)</td>
</tr>
<tr>
<td>2. Six K-Frac. = 1 m/d</td>
<td>1</td>
<td>24 (--)</td>
<td>--</td>
<td>25.47 (--)</td>
<td>--</td>
<td>25.47 (--)</td>
</tr>
<tr>
<td>3. Six K-Frac. = 500 m/d</td>
<td>1</td>
<td>1 (--)</td>
<td>24 (--)</td>
<td>9.95 (--)</td>
<td>16.69 (--)</td>
<td>25.64 (--)</td>
</tr>
<tr>
<td>4. Random K-Fractures (1-500 m/d)</td>
<td>12</td>
<td>24 (75%)</td>
<td>13 (17%)</td>
<td>26 (8%)</td>
<td>13.16 (7.16)</td>
<td>12.43 (7.3)</td>
</tr>
<tr>
<td>5. Conditioned Realizations for Zone 1</td>
<td>10</td>
<td>24 (60%)</td>
<td>1 (30%)</td>
<td>26 (10%)</td>
<td>15.44 (7.79)</td>
<td>10.38 (7.86)</td>
</tr>
<tr>
<td>6. Unconditioned Realizations for Zone 1</td>
<td>10</td>
<td>24 (80%)</td>
<td>1 (10%)</td>
<td>25 (10%)</td>
<td>15.37 (7.44)</td>
<td>10.07 (7.42)</td>
</tr>
</tbody>
</table>

* $\mu.Q1$ and $\mu.Q2$, refers to the mean pumping rate of the first and second selected wells by the GA.

** The total extraction ($Q_T$) was always higher than the sum of the tank inflow and tracer injection (24.35 mL/min).

Table 6.1 presents the optimization results of all tested scenarios including the frequency of well selection (Freq) besides the mean ($\mu$) and standard deviation (SD) of the chosen pumping rates. The GA selected well #24 as the best pumping location for case #1 (the experiment scenario). Well #24 is located upstream of the fractures close to BC1 at moderately permeable sand type (sand 40/50 see Figure 6.3). This was found to be in a good agreement with the finding of Cihan et al., (2015) that the lowest extraction rates will always occur in the permeable zones of the storage formation. It should be noted that well #1, located 12 cm downstream well #24, was the second most selected candidate under all tested scenarios after well #24. The reciprocal selection between
these two wells as optimum solutions, despite their proximity, highlights the important role of the storage zone heterogeneity and fracture permeabilities in altering the local flow field at the source vicinity. This affects the best pumping location and rate to control the leakage as well. Moreover, Cihan et al., (2015) emphasized that the consistency in well selection (i.e., well #1 and #24) under different problem settings can be a truthfulness indicator of the optimization results.

Well selection was found to be more sensitive to changes in Zone 1 heterogeneity than that of the fracture $K_s$. Minor differences were noticed in the GA results between the cases of using conditioned or unconditioned realizations to define the random field of the storage zone heterogeneity (cases #5 and #6). However, when these cases were compared to the actual calibrated $K$-field scenario (case #1), a major difference in well selection was found due to the variation of the heterogeneity structure. This indicates the importance of mapping the permeability field of the storage formation accurately for optimum well selection. Varying the fracture $K_s$ affected the values of the optimal pumping rates more than well selection (cases #2-4). That is because the fracture $K_s$ control the leakage flow rate, which in turn must be minimized through brine extraction. The effect of fracture $K_s$ on the optimum pumping rate was more pronounced in case #2, where the reduction in the fracture $K_s$ decreased the required pumping rate but did not affect the well selection. The important role of storage zone heterogeneity and fracture $K_s$ in determining the leakage pathways was highlighted in Chapter 4, which supports the above findings. The results show an overall high sensitivity of the optimal solution to the changes in storage zone heterogeneity and fracture permeability because both parameters control the local flow field at the vicinity of the source. This analysis clarifies the importance of running the optimization algorithm several times under all possible realizations of fracture permeabilities, distributions and storage zone heterogeneities to find the optimal solution under field information uncertainties.
6.3.5 Optimality Check for the Tank Extraction System

To test the optimality of the GA results within the solution space provided by the numerical model, a set of forward simulations were conducted using the same calibrated-model employed in the optimization process. In these simulations, the optimal results of the GA were slightly altered to evaluate the corresponding impact on the plume concentrations in the shallow aquifer and thus the optimality of the GA results. In the first three simulations, the optimum pumping rate \( Q_T \) found under case #1 (for well #24) was prescribed to other wells that were not selected by the GA for this case (i.e., wells #13, 14, 15, and 16). These wells were chosen for this analysis due to their locations at high permeable sands close to the caprock fractures, which can be seen as an indication for their potential ability to control the leakage plume. In the fourth simulation, the optimum pumping rate selected for well #24 under case #1 was reduced by 0.1% only. Figure 6.4 compares the plume concentration profiles along the constraining points in Zone 3 generated under all tested scenarios with the optimal GA solution (i.e., zero concentration profile drawn by a solid line).

The results reveal that using any well other than the selected well by the GA (well #24) or even reducing the optimal \( Q_T \) (25.64 mL/min) by only 0.1%, would produce a leakage plume that violates the assigned concentration constraints (i.e., \( C_T \gg 0.0 \) ppm) in the shallow aquifer. The shown sensitivity of the plume concentrations in the shallow aquifer to the well location and extraction rate in this analysis highlights the fact that finding the optimal leakage-control system is not a “commonsense” problem due to the high non-convexity and dimensionality of the solution space. Given that this analysis was performed under the assumption of perfect knowledge about the system (i.e., as the same model used to find the best solution was employed in this analysis), the GA results should also be tested under field-like uncertainty conditions to validate the proposed design approach for practical implementation. It should also be noted that experimentally testing
all of the above scenarios (18 scenarios) using the soil tank was not practical considering the time it takes to setup and complete an experiment.

6.3.6 Experimental Procedures for Testing the Optimal Extraction System

The main motivation of the experimental work in this study was to test the selected GA solution under an expected margin of uncertainty in field information. Such uncertainty conditions was mimicked using a model that incorporates (1) uncalibrated transport parameters, (2) K-field that was optimized on pre-injection hydraulic data only, (3) constant porosity value for all sand types, (4) imperfect representation of soil micro-heterogeneity, and (5) inaccurate tracer concentration, to design the extraction system in the tank.

Owing to the high cost and long time needed for the laboratory experiments (~1.5 months), a single leakage control experiment was conducted in the soil tank to test the GA results under case #1 presented in Table 6.1. Injection port #5 was selected for this experiment as it provides the
widest plume and highest leakage rate compared to the other sources, as learned from the analysis in Chapter 4. Such plume characteristics offer critical testing conditions for the leakage control system efficacy, which supports the reliability of the validation analysis. Using a single injection port mimics a scenario where one leakage source was activated by the induced injection pressure in the storage zone. Based on the GA results for case #1 presented in Table 6.1, well #24 was operated at a pumping rate equal to $25.62 \pm 0.2$ mL.min$^{-1}$ during the experiment. An injection rate of $7.9 \pm 0.2$ mL.min$^{-1}$ and Zone 1 inflow of $16.4 \pm 0.2$ mL.min$^{-1}$ were applied. Three peristaltic pumps (Cole-Parmer Masterflex L/S 7526-60) were used to create the injection and extraction conditions. The hydraulic head BCs presented in Figure 6.3 were also defined in the experiment through the constant head reservoirs for generating the flow field shown in Figure 6.2.

The experiment lasted for 13.2 days. In the first 12 days, aqueous samples were collected from the two most critical points in the constraining groups #1 and #2; these points were identified to be critical based on numerical simulations that revealed their early response to the leakage plume generated from port #5. However, on the last day of the experiment (13th day), aqueous samples were collected from all constraining points in Zone 3 besides 41 ports located at Zone 2 to map the plume extent after control. Collected samples were then analyzed using an ion chromatography system (Dionex Aquion combined with autosampler Dionex As-Dv) following the procedures presented in Sections 3.1.7 and 3.1.8. The hydraulic heads, tank inflow, and outflow were recorded twice a day. A video camera was used to take time-lapse images of the fluorescein-dyed-tracer every minute to record the plume propagation. The validation analysis was performed by checking whether the prescribed zero concentrations at the constraining points in Zone 3 were violated by the controlled leakage.
6.3.7 Results of Experimental Validation

The leakage experiment started by injecting the NaBr tracer through port #5 after reaching steady flow in the tank. The extraction phase was not initiated until the plume generated from the leakage approached the closest sampling port to the injection point, located 12 cm downstream. This procedure was applied to mimic an event of early leakage detection by deep monitoring sensors that alerted the operators to start the brine extraction from the storage zone. Although having a continuous extraction over the entire experimental time was intended, the extraction process was accidentally stopped on the 5th day for 3.5 hours due to a pump failure. As this is potential scenario that may occur in the field, the data was used to answer the question: “under unfavorable site conditions of possible failure of the extraction system, can the leakage plume still be controlled and reversed back to the storage formation, if the extraction was reactivated?”

Figure 6.5a shows the development of plume boundaries with time during the control period of the simulated leakage in the tank (including the pump failure time). The images show that a full control of the plume was achieved by restarting the control system after pump failure: most of the leakage flow was reversed to Zone 1, while only a small portion of the plume propagated in Zone 2 by dispersion. The plume then reached equilibrium after ~6 days at ~35 cm from the caprock (red dashed line in Figure 6.5a). Collected data over the entire ~13.2 days of the experiment showed undetectable levels of Br concentrations with the IC (<0.1 ppm) at the two most critical constraining points in the shallow aquifer layer. Figure 6.5b shows an interpolated concentration map of the leakage plume based on the 13th-day data collected from the 41-ports in Zone 2 and the constraining points in Zone 3. As shown in this figure, the plume was controlled and maintained in the vicinity of the source for entire time of the experiment. These results validate
the applicability of using the GA to find the optimum locations and pumping rates of the extraction wells used to control the potential brine leakage from CO₂ storage formations.

Figure 6.5 (a) Time-lapse photos of the controlled plume and (b) measured concentrations of the plume at the last day of the experiment.

Despite the ability to control conditions in the intermediate-scale system mimicking field physical process, there were limitations to assess the practicality of the extraction technique using the experiment only. Birkholzer et al., (2012), Cihan et al., (2015), and González-Nicolás et al., (2019) suggested a metric, called brine extraction ratio, to assess the practicality of the extraction
system for field implementation. This ratio describes the volume of native brine needs to be extracted relative to the volume of CO$_2$ planned to be injected. Such ratio should be less than one to consider the extraction technique to be practical. In the experiment, the leakage flow through the constantly opened six fractures was controlled mainly by the head gradient between Zones 1 and 3. Thus, to control the leakage and contain the plume in Zone 1, the hydraulic gradient should be reversed towards Zone 1. Since this condition cannot be accomplished unless the pumping rate from Zone 1 exceeds the sum of the tank inflow (through BC1) and tracer injection, the extraction ratio will always be greater than one (see Table 6.1). Therefore, it was essential to evaluate the practicality of the extraction technique in the field, where only a portion from the lateral flow in the storage layer (i.e., native brine) leaked out through caprock discontinuities. This evaluation was performed using numerical modeling as will be discussed in the next section.

6.4 Numerical Testing

As the experimental validation was limited to a system downscaled from field setting, additional analysis was conducted using numerical modelling at realistic scales of CGS applications. For this hypothetical analysis, the potential CO$_2$ storage site at the Vedder formation in San Joaquin Basin, California was selected. Extensive information on the hydraulic properties and geologic settings of the San Joaquin Basin and the Vedder formation are available from past studies associated with oil and gas exploration (USGS, 2007) and CGS assessment (Zhou & Birkholzer, 2011; Birkholzer et al., 2011; Wainwright et al., 2013; Jeanne et al., 2017). The Vedder formation consists of permeable sandstone layers that dip upward towards the Sierra Nevada mountain-range located at the eastern boundary of the San Joaquin Basin. The Vedder formation is overlaid by a double caprock shale layers (i.e., Freeman Shale and Fruitvale Shale) with average thickness of each about 300 m. Figure 6.6a shows a simplified cross-section of the San Joaquin
The Vedder formation storage capacity for a CO$_2$ injection scenario with a rate of 5 Mt CO$_2$/year for a period of 50 years was tested numerically by several researchers (Birkholzer et al., 2011; Wainwright et al., 2013; Pawar et al., 2014). Most of these studies emphasized that the existing faults within the project area of review (EPA, 2010) can be subject to a significant pressure increase, which may cause fault activation, induced seismicity, and thus brine/CO$_2$ leakage. Therefore, Cihan et al., (2015) and González-Nicolás et al., (2019) studied the possibility of implementing a brine extraction system to manage the build-up pressure in the Vedder formation. On the other hand, Jeanne et al., (2016; 2017) focused on assessing the stability of the Pond fault and overlaying formations under such an injection scenario. They found that the occurrence of a large earthquake as a consequence of this injection is unlikely, but low-magnitude frequent seismic events can be expected to happen. Micro-seismic events may not represent a significant geomechanical risk on the shallow zones, but during such events, fault permeability can increase and act as a continuous hydraulic pathway (Zoback & Gorelick, 2012; Gassiat et al., 2013; Lei et al., 2017; Detwiler & Morris, 2019). In contrast, Jeanne et al. (2018) showed that a succession of fault-slip events due to a series of micro-seismic activities will not always generate a fully activated conductive pathway. They also found that during fault activation, the permeability behind the rupture propagation is higher than in front of it, which results in more conductive interval of the fault close to the pressure source (i.e., storage formation). In the same context, the numerical studies of Guglielmi et al. (2015), Cappa et al. (2018), Rutqvist et al. (2020), and Park et al. (2020) showed a fault opening that occurs near the pressure source due to a reduction in the effective
normal stresses on the fault after CO₂ injection. Such opening can impose a considerable variation in the permeability of the deep and shallow parts of the activated fault. Moreover, the migrating high-velocity fluids through the fault, immediately after activation, can fluidize the sand of the fault internal surfaces in the overlying sandstone formations of San Joaquin Basin (Palladino, et al., 2018). This process will fill the upper part of the fault with native sand media that, in turn, will limit the effective conductive depth of the created pathway to the shale caprock thickness. A comprehensive review on the potential of fault activation and hydraulic pathway creation in CGS sites can be found in Guglielmi et al. (2021).

Based on above, the Pond fault may act as a continuous leakage pathway from the storage formation to the shallow aquifer (Shapiro, 2018), or as a buried pathway (e.g., buried thrust fault (Boston, et al., 2016)) with an effective conductive length that extends only to the top boundary of the caprock (i.e., Fruitvale Shale). Leakage through a continuous pathway is expected to be advection dominated with minimal lateral dispersion in the system. While the propagation of a leakage plume from a buried pathway will be governed by both advection and dispersion transport mechanisms due to its interaction with the overlying groundwater regimes and surrounding units. Advection-dominated leakage can be controlled by reversing the upward migrating flow through the pathway back to the storage zone without relying on the dilution capacity of the overlying formations. Therefore, to examine the benefit of considering such dilution capacity in the control system design, both of the potential leakage pathways from the Vedder formation (i.e., continuous, and buried pathways) were considered in this analysis. In the present chapter, the GA was used to find the optimum location and pumping rate of an extraction well that can control the hypothetical brine leakage event from the Vedder formation so that the plume concentrations in the shallow aquifer do not exceed a specified limit.
6.4.1 Modelling Simplifications and Assumptions

Large-scale 3D multiphase models are computationally intensive and can present a significant obstacle to be used as forward models in the optimization problems (Celia & Nordbotten, 2009). But the question is “whether we need that degree of modeling complexity to simulate a far-field brine leakage from a CO₂ storage zone?” Generally, researchers apply various simplifications to reduce the computational burden without significantly affecting the model prediction accuracy. In deep leakage problems, the large variability in the geological settings along the vertical direction of the system is central for plume migration to the shallow aquifer compared to variations in the lateral directions. Thus, Gassiat et al., (2013), Jeanne et al., (2013), Huang et al., (2015), and Vialle et al., (2016) used 2D cross-sectional vertical models to study CO₂ and brine leakage problems. Since this chapter focuses also on the vertical migration of far-field brine leakage reaching the shallow aquifer, a 2D cross-sectional vertical model was used to simulate the hypothetical leakage problem.

Regarding the different phases of flow that need to be considered for modeling brine leakage, Cihan et al., (2013) showed the suitability of using a single-phase flow model to simulate the far-field flow and brine leakage. In addition, a review article by Bandilla et al., (2015) reported the validity of using a single-phase model to simulate the pressure response and flow dynamics at distant areas away from the supercritical CO₂ plume. Examples of single-phase models used in exploring CGS-related problems include Nicot (2008) and Zhou et al., (2009). It is worth mentioning that when Cihan et al., (2015) simulated the pressure build-up at the Vedder formation, they applied the multiphase flow equations within a square of 15×15 km² which was smaller than the full domain size of 84×112 km². This is because a few previous studies on the same site showed that the CO₂ plume migration distance in the Vedder formation is less than this square size...
as mentioned by Cihan et al., (2015). Therefore, single-phase flow conditions were assumed in the 2D model used to simulate the Vedder formation pressure build-up and the upward brine leakage that follows.

Figure 6.6 (a) simplified cross-section of San Joaquin Basin stratigraphy and (b) transient simulation of the leakage plume propagation after fault activation by the equivalent injection of 5 Mt CO₂/year for 50 years.

To enable the single-phase flow assumption, an equivalent brine injection to CO₂ injection needs to be considered. Cihan et al., (2013) found that the corresponding changes in the volumetric CO₂ injection rate due to pressure-driven variations in the fluid properties (CO₂ density) are minor. Therefore, they applied a constant-rate fluid injection to represent the CO₂ injection in a single-phase model. Their single-phase model resulted in sufficiently accurate pressure build-up in the storage formation and brine leakage rate compared to the multiphase model. Thus, constant-rate brine injection was used in this analysis as well to represent the CO₂ injection. To ensure the
reproduction of comparable pressure fields between the 2D model of this analysis and the 3D model of Cihan et al., (2015), the prescribed equivalent injection rate was adjusted until a similar pressure build-up under the Pond fault was obtained. By that, a reasonable simulation of the pressure-driven brine leakage at the vicinity of the Pond fault due to the injection scenario of 5 Mt CO₂/year was guaranteed.

The Vedder formation has moderate salinity levels that range between 100 to 29,000 ppm, with an average brine density around 989 Kg/m³ (Zhou & Birkholzer, 2011; Cihan et al., 2015). Due to the minimal difference between the densities of the formation brine and overlying groundwater (998 Kg/m³), free-density flow conditions were assumed in the 2D model of this analysis. All the above model simplifications were necessary to reduce the computational burden of the forward simulation underlying the optimization process.

6.4.2 San Joaquin Basin Model Setup

The simulation code FEFLOW was used to develop the 2D flow and transport model of the San Joaquin Basin with the domain dimensions presented in Figure 6.6a. A total of 675,025 mesh elements with an average size of ~77 m² were used to spatially discretize the problem domain. A mesh refinement was applied at the fault vicinity to improve model convergence; the smallest mesh size in this zone reached 0.001 m². The basin stratigraphy shown in Figure 6.6a was explicitly represented in the model. The values of the formation specific storages, Ks, and porosities (Figure 6.6a) were retrieved from Cihan et al., (2015) and Jeanne et al., (2017). Following Jeanne et al., (2017), constant hydraulic heads at the model boundaries were prescribed while assuming initial hydrostatic pore pressure in the flow model. An overpressure of ~96 KPA was imposed in the Vedder formation through the assigned BCs to mimic the typical confinement conditions of the storage formations in the field (Osborne & E., 1997; Deming, 2001). Adverse
lateral hydraulic gradients of 25E-7 were assigned to the overlying formations and storage zone in the model. Given the opposing effects of temperature and pressure on the brine density with depth, isothermal conditions were assumed in this model similar to Birkholzer et al., (2011). The longitudinal and transversal dispersion coefficients were prescribed as 246 m and 24.6 m, respectively, based on the formulas suggested by Neuman (1990) and Gelhar et al., (1992).

Widths of 1m and 100m were used to describe the fault core and its surrounding damage zone, respectively. Notably, the geomechanical deformations in high-porosity sandstone result in a permeable fault core, acting as a flow conduit, and a less permeable surrounding damage zone (Tondi et al., 2006; Tondi, 2007; Fossen et al., 2007). Since most of the geological formations of the San Joaquin Basin are composed of grained sandstone (Wagoner, 2009), the K of the Pond fault core after activation was assumed two orders of magnitude higher than the average of Ks in the system. While the K of the damage zone was assumed lower than the adjacent formations’ Ks by two orders of magnitudes, following Gassiat et al., (2013). It is worth mentioning that this configuration of the K values was applied to represent the Pond fault as a continuous leakage pathway in the model (the 1st leakage mechanism) (Figure 6.6a). However, to represent the Pond fault as a buried leakage pathway (the 2nd leakage mechanism), the formations overlaying the Fruitvale Shale caprock were connected across the fault zone as shown in the inset part of Figure 6.6a. Reservoir salinity was prescribed by interpolating the TDS concentrations between the eastern and western sides of the study area (i.e., 100 ppm and 29,000 ppm, respectively).

6.4.3 Optimization Problem Description

Brine leakage through a continuous pathway can be faster and riskier than from a buried pathway, as the plume from the later will encounter multiple overlying formations before reaching the shallow aquifer, which will reduce its velocity and contamination load. Therefore, to explore
the extent of the potential areas to be impacted by brine leakage in the shallow zone of San Joaquin Basin (i.e., Kern River formation), a 2D model representing the fault as a continuous leakage pathway was run under transient flow and transport conditions for 5,000 years. Figure 6.6b illustrates that the brine leakage can reach the shallow aquifer with a concentration higher than 300 ppm in less than 50 years (before the end of the injection phase). Therefore, implementing a leakage control system that uses brine extraction wells can be a feasible solution to maintain a continuous injection process and protect the shallow aquifer from contamination risk. For optimizing the locations and extraction rates of these control wells for both potential leakage mechanisms in the system (i.e., through a continuous or buried leakage pathway), a steady-state flow and transport model was used under the GA to avoid the computational capacity required for running the transient model a large number of times (in some cases more than thousand). Under steady-state conditions, the reduction in plume concentrations caused by the infinite dispersion and diffusion process of the plume, is opposed by the enhanced advective transport of the leakage due to the applied continuous injection in this model. This results in a more distributed leakage plume across the overlaying formations but with slightly lower maximum concentration compared to the transient model results (20% lower). Even though this model can be seen unrealistic, it was considered satisfactory for the purpose of the current analysis, which is investigating the practicality of the leakage control technique (i.e., the extraction ratio) under field-scale conditions and the sensitivity of the extraction rates towards the target concentrations in the shallow aquifer.

A total of 97 potential well locations were predefined along the stretch of the Vedder formation for the GA selection (Figure 6.6a). The potential well locations were separated by 200 m. The optimization problem was constrained by a set of target concentrations ($C_T$ from 0.1 to 20 ppm) of the leakage at the interface between the shallow aquifer and its underling formation
(marked by a thick red line in Figure 6.6a). Different $C_T$ were tested to explore the impact of increasing the allowable concentration limits at the shallow aquifer on the optimization results. Regulators may increase the $C_T$ when the shallow aquifer contains brackish water (e.g., coastal aquifers) or is already affected by other contamination sources, which reduces the strategic value of the underground water resource. The maximum allowed concentrations of the leakage plume at the shallow aquifer should also be tied to the quality of the brine in the storage formation. For example, if the native brine contains or is expected to dissolve significant amounts of hazardous trace metals (e.g., As, Pb, Cd, etc.) (Wunsch et al., 2013; Zheng et al., 2014; Qafoku et al., 2017), very low target concentrations should be considered in the shallow aquifer.

6.4.4 Optimal Extraction System

Table 6.2 presents the selected extraction wells by the GA with their corresponding extraction ratios to control the leakage under both of the potential leakage mechanisms in the system (i.e., through a continuous or buried leakage pathway). To recall the definition of the extraction ratio, it is a ratio that describes the required extraction rate to control the leakage relative to the equivalent injection rate of CO$_2$. Since a steady-state model was used in this analysis, the extraction and injection rates are directly related to the extracted and injected volumes of brine and CO$_2$, respectively. Following Birkholzer et al., (2012), Cihan et al., (2015), and González-Nicolás et al., (2019), such extraction ratio was used as a metric to evaluate the practicality of the extraction technique to control a brine leakage event. As shown in Table 6.2, the GA selected well #73 for all tested scenarios of plume target concentrations under both of the potential leakage mechanisms in the system. Such replicated selection of the extraction well suggests that with a single pathway in the caprock and a homogeneous permeability field of the storage zone, the closest well to the leakage pathway becomes the optimal solution. This finding shows again the importance of
accurately representing the storage zone heterogeneity in the forward model underlying the optimization problem to find the best extraction location.

Table 6.2 San Joaquin Basin Optimization Results

<table>
<thead>
<tr>
<th>Target Concentration at Shallow aquifer (CT)</th>
<th>Continuous Pathway</th>
<th>Buried Pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Selected Well Number</td>
<td>Extraction ratio*100 (%)</td>
</tr>
<tr>
<td>CT ≳ 0.1 ppm</td>
<td>73</td>
<td>47.31</td>
</tr>
<tr>
<td>CT ≳ 0.5 ppm</td>
<td></td>
<td>47.26</td>
</tr>
<tr>
<td>CT ≳ 1.0 ppm</td>
<td></td>
<td>47.22</td>
</tr>
<tr>
<td>CT ≳ 5.0 ppm</td>
<td></td>
<td>47.02</td>
</tr>
<tr>
<td>CT ≳ 10 ppm</td>
<td></td>
<td>46.85</td>
</tr>
<tr>
<td>CT ≳ 20 ppm</td>
<td></td>
<td>46.61</td>
</tr>
</tbody>
</table>

The results presented in Table 6.2 reflect below-unity values of the extraction ratios (i.e., less than 50%) that indicate a reasonable practicality of the extraction technique to be field implemented. Moreover, the extraction ratios were found decreasing with higher target concentrations (\(CT\)) at the shallow aquifer. Such decrease was more pronounced under the buried pathway leakage mechanism compared to the continuous leakage pathway, where increasing the \(C_T\) from 0.1 to 20 ppm reduced the required extraction rate by ~1.5% under the later mechanism versus ~8% under the former mechanism. This difference can be attributed to the fact that the leakage transport through a continuous pathway is advection dominated. While the propagation of a brine plume that leaked from a buried pathway is governed by both advection and dispersion transport mechanisms, which allows for efficiently utilizing the dilution capacity of the overlaying formations. To extend this analysis, the extraction ratio data was fitted to a 1st order polynomial that resulted in \(R^2\) higher than 0.95 (Figure 6.7a). This regression analysis showed that with increasing the \(C_T\) to 50 ppm only, the extraction ratio can be reduced in the buried pathway scenario 5 times more than it is in the continuous pathway scenario. Since the extraction ratio is just a
different expression of the pumping rate required to control the leakage, the results indicate that managing a leakage from a buried pathway can be less expensive than a leakage from a continuous pathway. However, when the required $C_T$ at the shallow aquifer is extremely low (0.1 ppm), the extraction ratios under both leakage mechanisms become similar (Figure 6.7a and Table 6.2). This is because in the case of low $C_T$, the leakage flow should be totally eliminated from the source, which can be only accomplished by reversing the whole leakage flow to the storage zone, regardless the pathway setting.

Figures 6.7b – 6.7e illustrate the plume concentration profiles at the bottom boundary of the shallow aquifer (marked by a red line in Figure 6.6a) when the extraction well #73 was operated at the optimum extraction rate determined by the GA under both potential leakage mechanisms. As shown, the GA was able to find the minimum extraction rates while satisfying the concentration constraints in all tested scenarios. By comparing the concentration profiles in Figures 6.7b and 6.7c with Figures 6.7d and 6.7e, a more dispersed plume by the lateral flow of the overlying formations can be noticed in the scenario of a leakage from a buried pathway. This confirms the efficient utilization of the dilution capacity of these formations under this scenario. However, the sharp concentration spikes observed in the plume profiles under the continuous leakage pathway scenario indicates a minimal dispersion effect on the controlled plume (Figures 6.7b and 6.7c). It is worth mentioning that the dilution capacity of the overlying formations is expected to change with time, formation depths and soil type. Moreover, the findings related to the buried pathway leakage mechanism can be applicable in fracture leakage scenario, as both generate leakage plumes that experience similar transport conditions.
Figure 6.7 (a) the sensitivity of the extraction ratios to the change in the target concentrations. (b-e) the concentration profiles of the plume along the interface between Kern River and Etchegoin formations (marked by a red line in Figure 6.6a) for different target concentration scenario under the leakage mechanisms of (b, c) continuous or (d, e) buried pathway.

6.5 Discussion

The findings of this chapter showed that designing an optimal extraction scheme to control the potential brine leakage from CO2 storage formations is not a trivial problem as it comes with a high-dimensional, complex solution space. A global optimizer is necessary to be incorporated in this problem to deal with such non-convex, multivariant, discontinuous solution space. The GA
was successfully employed in this analysis to find the most optimal extraction control system for the experimental and numerical leakage problems that were considered. The GA results elucidated that the permeability architecture of the storage zone can highly alter the optimal extraction location, while the fracture permeabilities can have more impact on the extraction rates. It was also found that errors in well placements or the used pumping rates during operation can lead to a leakage plume that violates the maximum allowed concentrations in the shallow aquifer. These findings suggest that the extraction well locations can be determined, prior to CO₂ injection, using high-resolution data on storage zone heterogeneity collected during site investigation. It is worth mentioning that this high-resolution data can be obtained by applying the cross-well seismic imaging method between non-equipped injection and monitoring wells before operation (Nalonnil & Marion, 2012; Yu et al., 2008). However, the preliminary extraction rates, found prior CO₂ injection, should be revisited after leakage onset to be adjusted based on the inferred leakage pathways’ permeabilities using transient monitoring data on pressure changes and plume development.

The lack of accurate knowledge about deep geological settings and its transport properties in CGS sites (e.g., storage zone heterogeneity and dispersivity) is typically addressed in leakage modeling through considering different realizations of the system possible characteristics (Yonkofski et al., 2016; Chen et al., 2018; Jeong et al., 2019). Moreover, caprock damage zones would not even exist at the time of designing the extraction system, which requires developing more realizations to represent all the likely combinations of pathway locations, distributions and permeabilities. Dealing with all of these realizations under the same optimization problem can be challenging. Aitokhuehi & Durlofsky (2005), Van Essen et al., (2009), and Wang et al., (2012) introduced various approaches to integrate multiple system realizations under a single optimization
problem to find the overall optimal solution. However, under such conditions the optimization process can be computationally intensive. The experimental analysis of this chapter showed that uncertainty in specific system properties (e.g., transport parameters, leakage initial concentration, sand porosities, etc.) do not affect the extraction system design. Such findings can be used to reduce the dimensions of the parameter space that need to be sampled to develop the system realizations, which will decrease the required computational capacity for the optimization process.

The performed numerical experiments in this chapter showed an inverse relationship between the target concentrations in the shallow aquifer and the required extraction rates from the deep storage zone to control the leakage plume, particularly in the case of leakage through a buried pathway. This finding can be used to optimize the necessary extraction rate of the control system by allowing higher concentrations of the contaminant in the shallow aquifer. Such decision should be tied to the regulatory requirements by EPA, brine quality in the storage zone, and the expected geochemical reactions along the travel distance to the shallow aquifer. Applying integrated control systems is another approach to reduce the required extraction rates to control the leakage plume. In these systems, the deep extraction wells are responsible for alleviating the plume concentrations at the shallow aquifer to a certain level that can be further handled by shallow remediation methods. By that, the operational cost of the control process can be reasonably decreased.

### 6.6 Summary

In this chapter, the approach of using brine extraction to control the potential far-field brine leakage from CO2 repositories was evaluated. For this purpose, the extraction wells are designed to reverse the brine leakage flow and minimize the plume concentrations at the shallow aquifer by utilizing the dilution capacity of the overlying formations. Since handling large amounts of brine is costly, the Genetic Algorithm was incorporated with a transport model to optimize well-
placements and extraction-rates. An ~8 m long intermediate-scale tank designed to mimic expected field conditions was used to validate this approach as field data is not available. The approach was further tested numerically for a hypothetical leakage scenario at the Vedder storage formation in San-Joaquin basin to assess the technique’s practicality in the field. The results showed that storage zone heterogeneity and fracture permeabilities can significantly affect the optimum locations and pumping rates of the extraction wells. Brine leakage can be controlled by extracting a native-brine volume less than 50% of the injected CO₂. The target concentrations in the shallow aquifer have high impact on the extraction rates required to control a leakage through a fracture compared to that of a fault.
CHAPTER 7

SUMMARY, CONCLUSIONS AND FUTURE WORK

The research conducted under this dissertation had the ultimate goal of improving the management of CGS applications in general and far-field brine leakage in particular. The findings presented in this dissertation will help both operators and modelers to develop reliable risk assessment studies and design cost-effective monitoring networks and emergency leakage control systems. This chapter summarizes the key outcomes from hypotheses testing besides a number of specific conclusions that were derived based on the conducted analyses in Chapters 4, 5, and 6 (i.e., Papers I, II, and III, respectively) and finally, several avenues for future work are discussed\(^1\).

7.1 Summary of Hypotheses Testing Outcomes

In order to alleviate the adverse environmental impacts of CGS applications, the EPA released a set of federal requirements that should be met to ensure safe operation of CGS facilities. These requirements include developing and periodically reviewing: (1) a risk assessment study, (2) a monitoring system design and (3) an emergency leakage control strategy. Risk assessment relies mainly on the prediction of the leakage pathways in the overlying formations to identify the likely areas to be impacted in the shallow aquifer. Such predictions are expected to be sensitive to the accuracy of data on deep geological settings and source conditions. It is crucial for modelers to know which source parameters yield the highest influence on model predictions so that

\(^1\) The conclusions presented in this chapter are reproduced from the three papers listed in section 1.2. For this reproduction, proper permissions were taken from the publisher of Paper I, and the co-authors of the three papers.
particular attention can be given to them during site investigation. To explore these parameters as a part of testing the 1st hypothesis (Section 2.8), a validated transport model with experimental data was used to evaluate the impact of field data uncertainty on the accuracy of brine leakage predictions in 46 different scenarios, presented in Chapter 4 (Paper I). The results of this analysis showed that using inaccurate information about storage zone’s heterogeneity and pressure field to develop a model for leakage simulation can lead to a significant error in the predicted plume pathway and distribution. In turn, such errors will adversely impact the reliability of our risk evaluation. Additional specific conclusions were derived along the testing analysis of the 1st hypothesis that can be found in Section 7.2.

CGS monitoring is essential to protect the shallow aquifer against contamination. Owing to the high cost of placing sensors in deep formations, many of the currently used observation strategies depend on airborne and near-surface monitoring techniques (e.g., ground deformation). In contrast to deep monitoring, such techniques can give false leakage signals because of their sensitivity to data noises in the shallow zones. In Chapter 5 (Paper II), a design framework of cost-effective monitoring systems through optimally integrating both shallow and deep observations was developed and validated using an intermediate-scale soil tank to address the 2nd hypothesis (Section 2.8). The validation results showed that such an integrated system allowed for early leakage detection and provided valuable leakage data to improve model predictions that are necessary for developing remediation strategies. Additional specific conclusions were derived along the testing analysis of the 2nd hypothesis that can be found in Section 7.3.

Large-scale pressure build-up in the storage formation due to CO₂ injection poses geomechanical hazards on the caprock integrity and thus a contamination risk on the shallow aquifer. Multiple researchers suggested using deep extraction wells to manage the pressure
increase in the storage formations. Chapter 6 (Paper III) introduced the concept of using the same technique to control the potential brine leakage from CO₂ repositories so that the plume concentrations in the shallow aquifer do not exceed regulatory levels. In this technique, the extraction wells are designed to reverse the brine leakage flow and minimize the plume concentrations at the shallow aquifer by utilizing the dilution capacity of the overlying formations. Since handling large amounts of brine can be very costly, a global optimizer was employed to find the best well locations that give the lowest pumping rates to control the leakage. To address the 3rd hypothesis (Section 2.8), the applicability of the suggested design method and the practicality of the control technique were tested using experimental and numerical approaches, respectively. The results revealed that the extraction wells, designed using the suggested method, were able to maintain the leakage plume at close proximity to the source and retain the plume concentrations in the shallow aquifer at levels below the predefined limit in the optimization problem. Additional specific conclusions were derived along the testing analysis of the 3rd hypothesis that can be found in Section 7.4.

To conduct the above analyses, comprehensive knowledge about the deep geological settings and high-resolution data on leakage development are needed. Since such data is not available from the field, an ~8m long soil tank, where brine migrates from the storage zone to the shallow aquifer throughout multiple intermediate layers, was developed to generate the required data. More details about the soil tank design, instrumentation, and sand configuration can be found in Section 3.1.3. Despite all of the effort devoted to mimic natural settings in the soil tank, downscaling the problem length-scale and reducing system complexity may have impacted the research findings. Thus, it is important to note that the results from this laboratory-based study are not directly applicable to different field settings that may be encountered at potential CGS sites.
However, applying the intermediate-scale testing approach gave us the opportunity to assess the potential impact of uncertainty on leakage predictions and validate design approaches of leakage monitoring and control systems, which is not possible in the field due to cost and technical constraints.

7.2 Specific Conclusions from Chapter 4 (Paper I)

Predicting brine leakage pathways is essential to assess the contamination risk to underground sources of drinking water overlaying deep saline formations used in CO\textsubscript{2} geologic sequestration. Due to the difficulty of characterizing leakage source conditions with high accuracy in the field, model predictions of these potential leakage pathways can have a significant margin of uncertainty. Analysis presented in Chapter 4 based on intermediate-scale testing and numerical experiments allowed to draw the following conclusions that will enhance the practice of modelling brine leakage in CGS operations:

1. The calibrated flow and transport model developed using FEFLOW, a conventional ADE-based code, was validated to simulate the process of brine migration through a multilayered heterogeneous aquifer system represented in an intermediate-scale laboratory tank that simulated the interacting regional natural and leakage flows.

2. The analysis conducted in that chapter, using an intermediate-scale tank setting, showed that the lack of accurate representation of the storage zone heterogeneity can cause errors of up to 19\% in the predicted plume pathway and 100\% in its spread. The significance of this uncertainty impact will vary depending on the real field conditions, but it can be concluded that heterogeneity of the deep storage zone, particularly in the vicinity of the potential caprock damage zone, should be well defined in the model for making reasonable predictions of brine leakage plume.
3. Plume predication is highly sensitive to the mean value of the hydraulic conductivity (K) field of the storage zone compared to the other geostatistical parameters defining K distribution (variance and correlation lengths); uncertainty in the mean K can elevate the error in the predicted plume pathway by 17%.

4. High-resolution characterization of the storage zone permeability field can be challenging. Therefore, using an equivalent (homogeneous) K value in the simulations has been often applied. The scenarios analysis conducted showed that homogeneity assumption in the vicinity of the leakage fracture produces errors up to 19% and 49% in the plume pathway and spread, respectively. These findings highlight the importance of accurately characterizing the local heterogeneity underneath the potential fracture zones instead of mapping the entire permeability field of the storage formation.

5. The leakage pathway in the caprock may not be a single fracture but a damaged zone through which brine escapes to the overlying formations. Due to prohibitive investigation costs, modelers represent the damage zone by a single equivalent leakage pathway. The analysis showed that not representing the leakage zone as an equivalent leakage pathway can result in slightly less errors compared to assigning inaccurate discrete fracture permeabilities. However, both cases revealed an error around 8% in the predicted plume pathway, which will vary depending on site conditions.

6. When a leakage event occurs due to the sudden opening of a fracture, the vertical pressure gradients resulting from the high confining pressure in the storage zone will produce a vertical flow. In the deep overlaying formations located close to the fractures, the flow velocities are expected to be high. As the leakage driven-flow travels upward, it interacts with the regional natural flow field. The analysis showed that prescribing the lateral (or
regional) flow field is essential for the prediction of the plume pathway and spread at the entry of the shallow aquifers.

7. In the simulations, it is necessary to define the boundary conditions in the storage formation as accurately as possible to predict the flow field contributing to the development of the brine plume. The analysis showed that a small error in the storage zone boundary conditions (around 2%), particularly the downstream one, can significantly affect the accuracy of plume prediction (error up to 9.5% and 59% in the plume pathway and spread). This finding suggests the importance of obtaining data to define the storage zone boundary conditions during site characterization.

The above findings pointed out which of the characterization parameters result in the largest errors in the predictions. This knowledge will help to cost-efficiently investigate source conditions and thus improve site selection and design of CGS operations. This is particularly critical for cases where the storage zone caprock is overlaid by USDWs and can potentially leak under high injection stresses. The results also highlight that risk assessment studies developed based on poor characterization of source conditions can be misleading and unreliable. Thus, CGS stakeholders, such as local communities, industry, government, NGOs, and society can evaluate the accuracy of identified areas and activities to be impacted by a CGS operation in a risk assessment study using above findings. In addition, findings of this analysis could be extended to shed light on possible errors in predicting CO₂ leakage, other gas migration problems, and brine leakage in other deep injection applications like hydraulic fracturing and deep waste disposal.

7.3 Specific Conclusions from Chapter 5 (Paper II)

A proposed design framework that incorporates linear uncertainty analysis with the global optimizer of the Genetic Algorithm (GA) was presented and validated in Chapter 5 to be used in
developing the monitoring systems of CO₂ geological storage operations with a minimum number of observation points in the deep geological formations for logistical and cost considerations. To validate this framework, an intermediate-scale experiment was conducted in the soil tank described in Chapter 3, where a leakage event was simulated, and data were collected from the best monitoring locations, enhancing model predictions. Finally, the identified source location and the obtained predictions of leakage rate, storage zone flow, and plume distribution in the shallow aquifer, by the calibrated model with collected data, were compared to experimental measurements to validate the framework applicability. This analysis showed that the proposed OD framework was able to identify a set of informative monitoring locations that helps reducing the uncertainty in important predictions to CGS operators (a mean normalized error of less than 30% was obtained) while maintaining an early detection of the leakage. In addition, a set of conclusions that was driven from the optimal design analysis and the validation work performed in this study can be summarized as follows:

1. Both head and concentration data from the shallow and deep formations are crucial for identifying the leakage source, rate, and likely areas to be impacted in the shallow aquifer.

2. Shallow zone concentration data have high worth to predict the leakage mass flux and effective fracture permeabilities in the caprock; both of which were found to be the highest data-demanding predictions as well.

3. The Genetic Algorithm performance can be enhanced by incorporating an initial guess of the best design found by a sequential selection technique.

4. The iterative stage-wise calibration technique increases the accuracy of model predictions; for each prediction, the model is iteratively re-calibrated using the subset of data that yields the highest informativity to this specific prediction.
5. With deep zone concentration data capturing only the early development of brine leakage plume, the calibrated model was able to qualitatively identify the likely areas to be impacted in the shallow aquifer.

6. Considering additional random data in model calibration without conducting data worth analysis can increase the error of model predictions by propagating extra error from the inaccurate low-to-moderate informative data to predictions.

7.4 Specific Conclusions from Chapter 6 (Paper III)

In Chapter 6, using deep extraction wells as an emergency remediation strategy to control the potential brine leakage was evaluated. The Genetic Algorithm (GA) was incorporated with the experimentally validated FEFLOW-based transport model to find the optimum locations and minimum pumping rates of the extraction wells. An intermediate-scale leakage control experiment was conducted in the soil tank described in Chapter 3 to validate the applicability of the develop approach. In this experiment, the leakage plume was controlled by an extraction well located in the storage zone so that the plume concentrations in the shallow aquifer do not exceed a predefined limit. Moreover, the practicality of the leakage control technique was examined to be applied in the field under a realistic scale of the leakage problem using numerical simulations. For this analysis, a hypothetical leakage event from the Vedder formation in the San Joaquin Basin, California, was controlled using a single extraction well designed based on the results of the GA. The conclusions based on the conducted experimental and numerical analysis in this chapter are:

1. Finding the optimal extraction location and rate to control a brine leakage plume is not a trivial problem as it includes high-dimensional, non-convex solution space.

2. The Genetic Algorithm was able to deal with this solution space and find the optimal well locations and pumping rates for the tested experimental and numerical leakage problems.
3. Accurate mapping of the storage zone heterogeneity and reasonable predictions of the future fracture permeabilities are required for designing the leakage control system as both parameters can significantly affect the optimal locations and pumping rates of the extraction wells.

4. By implementing the selected extraction well and pumping rate in the soil tank, the leakage flow was reversed to the storage formations and the plume dispersion was maintained at close proximity to the caprock fractures.

5. The extraction system was able to control the leakage plume even after a pump failure that lasted for 3.5 hours.

6. The experimental analysis showed that the inevitable field data uncertainty in the system transport parameters, soil porosities, and leakage initial concentration did not affect the performance of the designed control system.

7. The numerical experiments showed that by extracting brine volumes less than half of the injected amount of CO₂ in the Vedder formation, the plume concentrations at the shallow aquifer mass maintained below a concentration of 0.1 ppm; such results indicate the practicality of using the extraction technique for brine leakage control.

8. An inverse relationship was found between the leakage target concentrations in the shallow aquifer and the extraction rates from the storage formation; this relation can be used to reduce the control system cost by allowing for higher contaminant concentration in the shallow aquifer if this was acceptable by authorities.

9. This inverse relationship was more pronounced when a leakage through a buried pathway or a caprock fracture was controlled compared to a continuous leakage pathway.
mechanism due to the efficient utilization of the dilution capacity of the overlying formations in the former scenarios.

10. The EPA requirement of developing an emergency remedial response strategy will be satisfied by designing a leakage control system using the proposed approach in this study; such system will be initially utilized for pressure build-up management, until a leakage event happens, only then, it will be used as an emergency response tool for leakage control.

7.5 Recommendations for Further Research

Further investigations are required to apply the findings of this research on actual CGS operations to improve the risk assessment analysis of far-field brine leakage. For example, more work should be performed to explore the impact of fracture sizes and heterogeneity of caprock discontinuities on the leakage pathway predictions. This will help enhancing leakage simulations that ensure better site-selection of CGS facilities and accurate identification of the likely areas and activities to be impacted in the shallow aquifer. The sensitivity of plume prediction to these fracture settings was not technically feasible for investigation under this study due to the limited length and width of the caprock surrogate used in the test tank. However, as the transport code of FEFLOW was validated to properly simulate brine leakage development and migration from CO₂ repositories, the impact of uncertainty in these parameters can be investigated numerically using a large scale CGS system domain or even under a hypothetical leakage problem at actual CGS sites.

On a different note, the developed and validated framework in this study, to optimally select the best observation points in the shallow and deep formations for monitoring far-field brine leakage, was based on incorporating the linear uncertainty analysis tool in PEST (i.e., FOSM) with the global optimizer of Genetic Algorithm (GA). To avoid the linearity assumption between model predictions and parameters in this approach, this framework can be modified to incorporate non-
linear uncertainty methods (e.g., ensemble-smoother-based tools or null-space Monte Carlo method). However, the feasibility of such modification should be assessed based on a trade-off analysis between the extra benefits that can be gained in the design of the monitoring system (e.g., more accurate predictions) and the added computational burden to the optimization process using this approach. These benefits can be evaluated by comparing the data informativity of monitoring systems designed using linear- and non-linear-based frameworks. This work will provide the designers and modelers with some guidance on the best approach they can use based on the computational resources they have. In addition, exploring the sensitivity of the monitoring system design to the prediction weights is recommended to help providing guidelines on how to assign such weights.

Finally, one additional avenue of future work could be exploring the differences in results between using steady-state and transient forward models while designing the control system of brine leakage. This effort can lead to useful knowledge about the necessity and advantages of employing transient forward models under the optimization algorithms, despite their high computational cost, to find the locations and pumping rates of the extraction wells. This knowledge is important for designer and modelers to decide about the most appropriate design approach based on the available computational resources and site-conditions.
REFERENCES

Agartan, E., 2015. Fundamental study on the effects of heterogeneity on trapping of dissolved CO\textsubscript{2} in deep geological formations through intermediate-scale testing and numerical modeling, PhD Dissertation. Golden: Colorado School of mines.


Deutsch, C. V., 2006. A sequential indicator simulation program for categorical variables with point and block data: BlockSIS. Computers & Geosciences, 32(10), pp. 1669-1681.


Gelhar, L. W., 1986. Field-measured values of longitudinal dispersivity as a function of the scale of measurement. Water Resources Research, 22(9), pp. 135S-145S.


Gurevich, A. E. & Chilingarian, G. V., 1993. Subsidence over producing oil and gas fields, and


with sensitivity analysis using a Latin-hypercube sampling technique. Groundwater, 34(5),
pp. 811-818.


big data and smart field technology for detecting leakage in a CO₂ storage project. In SPE
annual technical conference and exhibition. New Orleans, Louisiana, USA, Society of
Petroleum Engineers.

Hanna, R. B. & Rajaram, H., 1998. Influence of aperture variability on dissolutional growth of

storage in depleted or depleting oil and gas fields: What can we learn from existing

Harlow, I. F., 1939. Waste problems of a chemical company. Industrial & Engineering Chemistry,
31(11), pp. 1346-1349.

projects (No. ANL/EVS/R-11/1), Argonne, IL: Argonne National Lab (ANL).

Harvey, C. & Gorelick, S. M., 2000. Rate-limited mass transfer or macrodispersion: Which
dominates plume evolution at the Macrodispersion Experiment (MADE) site?. Water


quantify parameter uncertainty and its effect on predictions of a groundwater flow model.
Environmental Modelling & Software, 24(6), pp. 749-763.


Pawar, R., Bromhal, G., Carroll, S., Chu, S., Dilmore, R., Gastelum, J., Oldenburg, C., Stauffer, P., Zhang, Y. and Guthrie, G., 2014. Quantification of key long-term risks at CO\textsubscript{2}


245


APPENDIX A

MATERIAL REPRODUCTION PERMISSIONS

Permissions to reproduce Figures (1.1, 2.1, 2.2, 2.3, 2.5, and 2.6), Table 1.1, and permissions to include Chapter 4 (Paper I) may be found in the Supplemental File named “Permissions”.