Characterization of CO₂ plumes in deep saline formations using fluid pressure tomography

Author(s): Hu, Linwei

Publication Date: 2017

Permanent Link: https://doi.org/10.3929/ethz-a-010884142

Rights / License: In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Characterization of CO₂ plumes in deep saline formations using fluid pressure tomography

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

(Dr.sc. ETH Zurich)

Presented by
LINWEI HU

M.Sc., Georg-August-Universität Göttingen
born on November 19, 1986
citizen of Hainan, China

accepted on the recommendation of

Prof. Dr. Simon Löw, examiner
Prof. Dr.-Ing. Rainer Helmig, co-examiner
Prof. Dr. Peter Bayer, co-examiner

2017
# Table of Contents

Acknowledgement ........................................................................................................... 5  
Nomenclature .................................................................................................................... 7  
Abstract .............................................................................................................................. 10  
Zusammenfassung ................................................................................................................ 12  

1 Introduction ..................................................................................................................... 14  
   1.1 CO₂ capture and storage (CCS) ................................................................................. 14  
   1.2 Geophysical methods .............................................................................................. 17  
   1.3 Pressure-based methods ......................................................................................... 19  
   1.4 Hydraulic tomography (HT) .................................................................................... 20  
   1.5 Objectives and structure of this work ...................................................................... 23  

2 Fundamentals of CO₂ sequestration in saline aquifers .................................................. 25  
   2.1 Trapping mechanism ............................................................................................... 25  
   2.2 Two-phase flow in porous media ............................................................................. 26  
      2.2.1 Mass and energy conservation equations ......................................................... 26  
      2.2.2 Relative permeability and capillarity ................................................................. 27  
   2.3 Fluid properties ....................................................................................................... 28  
      2.3.1 CO₂ properties ................................................................................................ 28  
      2.3.2 Brine properties ............................................................................................. 29  

3 Time-lapse pressure tomography for characterizing CO₂ plume evolution in a deep saline aquifer ............................................................................................................. 31  
   Abstract .......................................................................................................................... 31  
   3.1 Introduction .............................................................................................................. 32  
   3.2 Methodology ............................................................................................................ 34  
      3.2.1 Overview: cross well testing and inversion ....................................................... 34  
      3.2.2 Pressure-based tomographical inversion ......................................................... 35  
      3.2.3 Clustering and zonal calibration ...................................................................... 37  
      3.2.4 Forward modeling ......................................................................................... 37  
      3.2.5 Influence of CO₂ on fluid properties .............................................................. 39
4 Detection of carbon dioxide leakage during injection in deep saline formations by pressure tomography ..............................................................60

Abstract ..............................................................................................60

4.1 Introduction ........................................................................................61

4.2 Methodology ......................................................................................62
  4.2.1 Problem set-up ............................................................................62
  4.2.2 Leaky cases and model parameters ............................................63
  4.2.3 Fluid interference tests ...............................................................64
  4.2.4 Inversion in two-phase system ....................................................66
  4.2.5 Clustering ..................................................................................67

4.3 Result and discussions ......................................................................67
  4.3.1 Head changes ...........................................................................67
  4.3.2 Early time diagnostics ...............................................................69
  4.3.3 Inversion and clustering .............................................................69

4.4 Conclusions .....................................................................................73

4.5 Supporting information ....................................................................73
  4.5.1 Introduction ...............................................................................73
  4.5.2 Relationship between CO$_2$ saturation and mixed-phase diffusivity 73
  4.5.3 Inversion procedure ..................................................................74
  4.5.4 Results of configuration with 100 m well spacing .......................75
  4.5.5 Reliability map ..........................................................................78
  4.5.6 Clustering results .....................................................................78

5 Characterizing CO$_2$ plumes in deep saline formations: comparison and joint evaluation of time-lapse pressure and seismic tomography ........................................80

Abstract ..............................................................................................80

5.1 introduction .......................................................................................81

5.2 methodology .....................................................................................83
5.2.1  Overview of the methodology .......................................................... 83
5.2.2  Problem set-up .................................................................................. 84
5.2.3  Forward simulation ........................................................................... 90
5.2.4  Clustering and zonal calibration ....................................................... 92
5.3  Results .................................................................................................. 96
  5.3.1  PT and ST travel times ..................................................................... 96
  5.3.2  Diffusivity and velocity tomograms ................................................ 96
  5.3.3  1-D and 2-D clustering structure .................................................... 98
  5.3.4  Zonal calibration and calculated saturations .................................... 99
5.4  Discussion ............................................................................................ 101
  5.4.1  PT and ST travel times ..................................................................... 101
  5.4.2  Diffusivity and velocity tomograms ................................................ 101
  5.4.3  1-D and 2-D clustering structure .................................................... 102
  5.4.4  Saturation errors ............................................................................. 103
5.5  Conclusions ........................................................................................... 104
5.6  Appendices ............................................................................................ 105
  5.6.1  Appendix A. Discretization of two-phase flow simulation model ...... 105
  5.6.2  Appendix B. Gassmann-Wood rock physics model ......................... 106
  5.6.3  Appendix C. Full velocity (difference) tomograms ......................... 106
5.7  Supporting information ........................................................................ 107
  5.7.1  Introduction .................................................................................... 107
6  Conclusions and outlook ........................................................................ 113
  6.1  Concluding remarks ........................................................................... 113
  6.2  Outlook: current challenges and future perspective .......................... 114
Appendix 1  Rapid field application of hydraulic tomography for resolving aquifer heterogeneity in unconsolidated sediments .................................................. 116

Abstract ...................................................................................................... 116
A1.1 Introduction .......................................................................................... 117
A1.2 Field data processing and inversion methodology ............................... 120
  A1.2.1 Travel time inversion of pumping tests data .................................. 120
  A1.2.2 Travel time diagnostics ................................................................. 121
  A1.2.3 Attenuation inversion of pumping test data .................................... 123
  A1.2.4 Derivation of the Conversion factor for a Heaviside Source .......... 123
A1.3 Numerical example of the attenuation inversion based on an analogue outcrop study ................................................................. 125
A1.4 Field application of the travel time and attenuation tomography........ 127
A1.5 DP injection logging ........................................................................................................................................ 131
A1.6 Acknowledgement .................................................................................................................................................. 132
A1.7 Conclusions ............................................................................................................................................................ 132
A1.8 Supporting information ..................................................................................................................................... 134

Appendix 2  Prediction of solute transport in a heterogeneous aquifer utilizing hydraulic conductivity and specific storage tomograms ........................................ 135

Abstract ........................................................................................................................................................................ 135
A2.1 Introduction .......................................................................................................................................................... 136
A2.2 Material and methods ......................................................................................................................................... 138
   A2.2.1 Field site and experiments .......................................................................................................................... 138
   A2.2.2 Tomographic inversion .............................................................................................................................. 141
   A2.2.3 Numerical modeling ..................................................................................................................................... 144
A2.3 Results ................................................................................................................................................................. 145
   A2.3.1 Eikonal based inversion of hydraulic tests ......................................................................................... 145
   A2.3.2 Conceptual map and pilot points configuration ................................................................................... 146
   A2.3.3 Simulation of pumping tests .................................................................................................................. 148
   A2.3.4 Calibrated hydraulic parameter fields ................................................................................................. 149
   A2.3.5 Validation of the reconstructed aquifer with tracer test data ............................................................. 152
A2.4 Conclusions ......................................................................................................................................................... 155
A2.5 Acknowledgements ............................................................................................................................................ 156

Bibliography .................................................................................................................................................................. 158

Curriculum Vitae .......................................................................................................................................................... 175
Acknowledgement

I would like to give great thanks to my supervisor Prof. Dr. Simon Löw for offering me the opportunity to study as a PhD student in ETH Zurich, which greatly contributes to my life and my future career.

I appreciate greatly for the supervision of Prof. Dr. Peter Bayer and Dr. Ralf Brauchler, especially for their endless help and patience, for their efforts on discussing with me, and for their constructive advices. I would like to thank Prof. Dr. Peter Bayer for helping me to write scientific papers, which greatly improves my English.

I am grateful to all of my colleagues at the group of Engineering Geology for the pleasant work atmosphere. I would like to thank in particular Dr. Santo Jimenez, Mark Somogyvari, Jaime Rivera, Dr. Joseph Doetsch, Dr. Reza Mohammad Jalali, Peter Achtziger, Mohammad Afshari Moein, for the great help in solving my tedious problems. I also thank Dr. Fanny Leuenberger-West for teaching me how to work in a hydrochemical lab.

I would like to thank Chia-Yu Chen, Rita Shih, Dr. Reza Mohammad Jalali and Dr. Andrea Wolter for sharing fun in daily life in Zurich, as well as in every journey with me.

Last but not least, I wish to thank my parents for their encouragement and endless love. Without their support, it is impossible for me to finish this work.
Nomenclature

A  Tomographical matrix
A  “true” plume
B  Inverted plume
c_n  CO_2 compressibility (1/\text{pa})
c_w  Water compressibility (1/\text{pa})
c_{r, p}  Heat capacity (j/kg)
D  Diffusivity (m^2/s)
D_n  Diffusion coefficient of CO_2 in brine (m^2/s)
D_{post}  Diffusivity at post-injection (m^2/s)
D_{pre}  Diffusivity at pre-injection (m^2/s)
\Delta D  Diffusivity difference (-)
d  Aquifer thickness (m)
f_{a,d}  Conversion factor (-)
G_{dry}  Bulk modulus of dry frame rock (Pa)
G_f  Bulk modulus of mixing pore fluid (Pa)
G_m  Bulk modulus of rock matrix (Pa)
G_{sat}  Saturated bulk modulus (Pa)
g  Gravity acceleration (m/s^2)
H  Enthalpy (j/kg)
h  Head (m)
h_{f}(r, t)  First time-derivative of head (m/s)
h_{d}(r, t_{peak})  Maximum first time-derivative value of head data (m/s)
I_x  Integrated length in \textit{x}-direction (m)
I_z  Integrated length in \textit{y}-direction (m)
i  Component
K  Mixed-phase hydraulic conductivity (m/s)
K_w  Single-phase hydraulic conductivity (m/s)
k  Intrinsic permeability (m^2)
k_{cap}  Intrinsic permeability of caprock (m^2)
k_{ref}  Intrinsic permeability of reference media (m^2)
k_{nw}  Relative permeability of non-wetting phase (-)
k_{rw}  Relative permeability of wetting phase (-)
k_{seal}  Intrinsic permeability of bottom seal (m^2)
M  Local reliability indicator (-)
N_{dry}  Shear modulus of dry frame rock (Pa)
N_{sat}  Saturate shear modulus (Pa)
P  Global pressure (Pa)
P_0  Datum pressure (Pa)
P_c  Capillary pressure (Pa)
P_d  Entry pressure (Pa)
\( P_{d,\text{ref}} \)  
Entry pressure of reference media (Pa)

\( P_w \)  
Pressure of wetting phase (Pa)

\( P_n \)  
Pressure of non-wetting phase (Pa)

\( Q \)  
Source/sink term of mass (mol/s)

\( Q_h \)  
Source/sink term of heat (W)

\( Q_c \)  
Mass injection rate of CO\(_2\) (kg/s)

\( Q_w \)  
Mass injection rate of water (kg/s)

\( q_n \)  
Volumetric injection rate of CO\(_2\) (m\(^3\)/s)

\( q_w \)  
Volumetric injection rate of brine (m\(^3\)/s)

\( r_w \)  
Well radius (m)

\( S_n \)  
Saturation of non-wetting phase (-)

\( S_{\text{cal}} \)  
Calculated CO\(_2\) saturation (-)

\( S_{\text{true}} \)  
“true” CO\(_2\) saturation (-)

\( S_{nr} \)  
Residual saturation of non-wetting phase (-)

\( S_w \)  
Saturation of wetting phase (-)

\( S_{rw} \)  
Residual saturation of wetting phase (-)

\( S_s \)  
Mixed-phase specific storage (1/m)

\( S_{sw} \)  
Single-phase specific storage (1/m)

\( s \)  
Propagation path (m)

\( T \)  
Temperature (K)

\( T_0 \)  
Initial temperature (K)

\( T_{\text{inj}} \)  
Duration of injection at CO\(_2\) sequestration stage (h)

\( T_s \)  
Duration of injection for multi-level CO\(_2\) injection (h)

\( T_{\text{rec}} \)  
Duration of recovery at CO\(_2\) sequestration stage (h)

\( t_{a,d} \)  
Early time diagnostic (s)

\( t_{\text{peak}} \)  
Peak travel time (s)

\( U \)  
Internal energy (J)

\( U \)  
Left singular vector

\( V \)  
Right singular vector

\( V \)  
P-wave velocity (m/s)

\( V_{\text{post}} \)  
P-wave velocity at post-injection (m/s)

\( V_{\text{pre}} \)  
P-wave velocity at pre-injection (m/s)

\( \Delta V \)  
Velocity difference (m/s)

\( v \)  
Darcy flow velocity (m/s)

\( W \)  
Lambert’s \( W \) function

\( \mathbf{W} \)  
Diagonal matrix

\( X_{i,n} \)  
Molar fraction (-)

\( x_1 \)  
Injection point (m)

\( x_2 \)  
Observation point (m)

Greek symbols
α  Overestimation rate (-)
β  Underestimation rate (-)
ε  Total misclassification rate (-)
κ  Thermal conductivity coefficient (w/mk)
κ_{dry}  Dry rock thermal conductivity (w/mk)
κ_{wet}  Wet rock thermal conductivity (w/mk)
λ  Pore size distribution (-)
λ_T  Total mobility (1/pa s)
λ_w  Mobility of wetting phase (1/pa s)
λ_n  Mobility of non-wetting phase (1/pa s)
μ_w  Viscosity of wetting phase (pa s)
μ_n  Viscosity of non-wetting phase (pa s)
ξ  Saturation error (-)
ρ  Molar density (kg/m₃)
ρ_d  Dynamic mixed-phase density (kg/m₃)
ρ_m  Rock matrix density (kg/m₃)
ρ_n  Density of non-wetting phase (kg/m₃)
ρ_r  Rock density (kg/m₃)
ρ_s  Static mixed-phase density (kg/m₃)
ρ_w  Density of wetting phase (kg/m₃)
τ  Tortuosity (-)
ϕ  Porosity (-)
Abstract

Geological carbon dioxide storage (GCS) is a promising technique for cutting down anthropogenic carbon dioxide (CO\textsubscript{2}) emissions by storing CO\textsubscript{2} in natural geological media. Short- and long-term monitoring techniques are required for tracking the CO\textsubscript{2}-induced changes of the “invisible” subsurface, following up the fate of the disposed CO\textsubscript{2}. Monitoring techniques are also crucial for being able to immediately react to potential CO\textsubscript{2} leakage from the geological reservoir and for formulating remediation strategies. For this purpose, geophysical methods are extensively used. However, a major pitfall of most geophysical methods is that they merely provide indirect information on flow properties. This yields uncertainties especially for the estimation of CO\textsubscript{2} saturation. This PhD thesis introduces an alternative approach, fluid pressure tomography, which has potential to overcome this by direct linkage of the observed signals and the inversion procedure to the flow regime.

Due to the complications of CO\textsubscript{2} properties, CO\textsubscript{2} sequestration involves more complexities than single-phase flow. The primary concern of this work is to develop a single-phase proxy, which can significantly reduce the computing burden of full multiphase simulation and accelerate the inversion procedure. In this proxy, CO\textsubscript{2} and brine are assumed as a phase mixture, neglecting the secondary processes, such as thermal and chemical processes. Disposal of CO\textsubscript{2} in the brine-rich formations alters the mixed-phase flow properties. The mixed-phase specific storage increases greatly with increased CO\textsubscript{2} saturation, since CO\textsubscript{2} is much more compressible than brine. In contrast, variations in the mixed-phase conductivity are relatively small. Similar to the ratio of the mixed-phase conductivity and specific storage, mixed-phase diffusivity can change by up to two orders of magnitude, which can be recognized by fluid pressure tomography. Implementation of pressure tomography involves brine or CO\textsubscript{2} injections as sources, and pressure measurements in different locations as receivers. Pressure transients at the observations are utilized for travel-time based inversion, which yields the structural information of the subsurface. Plume development is inferred by comparing and clustering the inverted diffusivity tomograms acquired at different times. The CO\textsubscript{2} saturation of the identified plume is then derived by calibrating the measured pressures based on the single-phase proxy. A synthetic homogeneous case is used for demonstrating the feasibility of the method.

Applying pressure tomography not only to the storage formation, but also to the above aquifer, can detect potential CO\textsubscript{2} leakage occurred at different times. A no-leakage case is simulated as a reference to be compared with various leaky cases. It is demonstrated that pressure responses and hydraulic travel times in storage formation and the aquifer above provide a first insight in the leakage type. Comparison of the diffusivity tomograms in both storage formation and above aquifer among no-leakage and leaky cases can localize the leakage. Furthermore, the influence of data noise and well distance is examined. Results indicate that the noise has an impact on the inversion results and leakage detectability. Increase of well distance also weakens the detectability of CO\textsubscript{2} leakage, since it reduces the inversion resolution.
In the last part of this work, pressure tomography is conducted in heterogeneous formations, in comparison with crosswell seismic tomography under comparable conditions. Hydraulic travel times show much larger relative spread than seismic tomography, which allows pressure tomography to better resolve the more complicated geometries. Moreover, from the inverted tomograms, these two approaches show different capability for resolving the aquifer structure and the CO$_2$ plume. Pressure tomography delineates the structure of the initial CO$_2$-free formation better than seismic tomography, because it directly relates to formation permeability. For the post-injection periods, however, seismic tomography can always depict the main part of the plume, while pressure tomography is more influenced by the heterogeneity of the aquifer. Joint clustering of pressure and seismic tomography results combines the advantages of these two approaches. The plume shape is better identified, and also the estimation error of plume CO$_2$ saturation is reduced.

This work reveals the theoretical potential of the new concept of “time-lapse pressure tomography” to further adapt the single-phase hydraulic tomography to a two-phase flow system in a time-lapse manner. Based on the single-phase proxy, the tomographical inversion and calibration of the flow properties become rapid and computationally efficient. With the promising theoretical results, the methodology introduced here is ready for field applications future.
Zusammenfassung


Die Anwendung von Druck-Tomographie nicht nur auf das eigentliche Reservoir, sondern auch auf einen Grundwasserleiter im Hangenden, kann ein möglicherweise auftretendes CO_2-Leck erkennen. Hierzu wird als Referenz vorab ein Fall ohne Leck simuliert. Es wird gezeigt, dass Druck-Messungen und hydraulische Laufzeiten in Reservoir und Grundwasserleiter einen ersten


Diese Arbeit zeigt das theoretische Potenzial dieses neuen Konzepts der multitemporalen Druck-Tomographie, über das eigentlich einphasige hydraulische Tomographie angewandt wird, um die Dynamik eines zweiphasigen Fliesssystems abzubilden. Durch einphasige Näherung werden tomographische Inversion und Kalibrierung der Fliesseigenschaften stark beschleunigt. Mit den gezeigten vielversprechenden theoretischen Ergebnissen ist die hier eingeführte Methodik bereit für eine zukünftige Feldanwendung.
1 Introduction

1.1 CO₂ capture and storage (CCS)

Greenhouse gases (GHG), mainly including carbon dioxide (CO₂), methane (CH₄) and nitrous oxide (N₂O), has a significant impact on global warming and long-lasting climate change, which causes a severe and irreversible influence on human beings and the integrated ecosystem. The fifth Assessment Report (AR5) on climate change points out that the recent anthropogenic emissions of greenhouse gases (GHG) reach the highest value in history [IPCC, 2014]. In 2010, around 45±4.5 GtCO₂-eq/yr of anthropogenic GHG has been emitted. Among these GHG, CO₂ emissions accounts for a large part. The global anthropogenic CO₂ emissions have significantly increased since 1950 due to the great use of fossil fuels (Figure 1.1), and possibly reached around 37 Gt CO₂/yr in 2015 [Le Quéré et al., 2014; Celia et al., 2015]. Cumulative anthropogenic CO₂ emissions into the atmosphere climbed up to 2040±310 Gt over the last 260 years (between 1750 and 2011) [IPCC, 2014]. Although the emitted CO₂ can be stored in plants and soil, or in the ocean by the global carbon cycle, still around 40% of the emissions have stayed in the atmosphere (880±35 Gt CO₂). The remained CO₂ leads to a warm-up effect on the global temperature. By 2010, the globally average combined land and ocean surface temperature anomaly is about 0.2 °C [IPCC, 2014].

![Global anthropogenic CO₂ emissions](image)

Figure 1.1 Annual amount of global anthropogenic CO₂ emissions (from 1850 to 2010) from forestry and other land use vs. fossil fuels, cement and flaring [IPCC, 2014]

Despite various measures being suggested for reducing the CO₂ emissions, such as improving the energy usage efficiency and replacing fossil fuels by several alternative “clean” energy sources (e.g., solar and wind energy, hydropower), carbon dioxide capture and storage (CCS) is still reckoned as the only countermeasure for mitigating the industrial-scale anthropogenic CO₂ emissions [Global CCS Institute, 2015]. Generally, a CCS project includes extraction of CO₂ from industrial or energy-related sources (e.g., a power plant), transport to a geological storage site and
long-term isolation and storage from the atmosphere [IPCC, 2005]. The geological storage of CO₂ is the most important part of an entire CCS project. Favorable conditions are required for selecting suitable geological formations [Celia et al., 2015]:

- Sufficient storage capacity to dispose of the produced CO₂
- Injectivity which can retain the CO₂ injection rate as the supplied rate from the emission source
- Large containment to prevent CO₂ leakage to the shallow groundwater and soil, as well as to the atmosphere

Potential geological storage formations mainly contain depleted oil/gas reservoirs [Winter and Bergman, 1993; Hawkes et al., 2005; Li et al., 2006; Nogueira and Mamora, 2008; Godec et al., 2011; Underschultz et al., 2011; Whittaker et al., 2011; Jenkins et al., 2012; Tambach et al., 2015; Ampomah et al., 2016; Ojo and Tse, 2016], enhanced oil fields [Walsh and Lake, 1989; Malik and Islam, 2000; Aycaguer et al., 2001; Shaw and Bachu, 2002; Bachu et al., 2004; Gozalpour et al., 2005; Gorecki et al., 2012; Godec et al., 2013; Liao et al., 2016], gas recovery sites (EGR) [Clemens and Wit, 2002; Mamora and Seo, 2002; Al-Hashami et al., 2005; Schepers, 2009; Wu et al., 2011; Dahaghi, 2013], deep unmineable coal seams [Reeves, 2001; Gale et al., 2004; Shi and Durucan, 2005; Smith et al., 2005; Kronimus et al., 2008], and deep saline formations [Bachu and Adams, 2003; Bentham and Kirby, 2005; Eccles et al., 2009; Kharaka et al., 2009; Bauer et al., 2012; Celia et al., 2015; Rathnaweera et al., 2016]. Usually, at depleted oil/gas reservoirs, there usually exists a sufficient number of geological profiles from the previous site investigation. Formulating and optimizing a CO₂ storage strategy could benefit from these data. One major disadvantage of these storage reservoirs is that pre-existing wells could be a potential source for CO₂ leakage. Despite this, disposal of CO₂ in an enhanced oil or gas field can increase the recovery of oil or gas production. This is called “Enhanced Oil Recovery (EOR)” or “Enhanced Gas Recovery (EGR)”. EOR or EGR has a marked improvement in economic efficiency, and the techniques have been well-established in petroleum industry. However, previous studies show that only 60% of the injected CO₂ can be stored in the reservoir, with 40% of the CO₂ escaped from the production well [Shaw and Bachu, 2002; Gozalpour et al., 2005]. The storage efficiency is greatly limited to the exploration techniques. CO₂ storage in unmineable coal seams is even more challenging. Due to the strong absorption of CO₂ in coal bed, it could displace the host CH₄ to enhance its production [Gale et al., 2001; Mazzotti et al., 2009]. Nevertheless, the permeability of the host formation is usually very low, resulting in a small storage capacity. Moreover, the absorbed CO₂ even decreases the permeability. So far, no successful operations are achieved, thus this concept has already been abandoned in practice [Celia et al., 2015].

All aforementioned potential storage reservoirs have a common drawback, that is, the relatively small storage capacity, except for deep saline aquifers. Deep saline aquifers usually offer a higher storage efficiency due to their common occurrence and large containment. Most of the saline aquifers are deeper than 800 m with high temperature. Under these conditions, CO₂ is injected and stored in a supercritical state, increasing the available storage amount compared to other formations [Bachu and Adams, 2003]. To date, there exists several worldwide ongoing and
planned CCS projects, utilizing deep saline aquifers for sequestering CO$_2$. Here several major storage sites are exemplarily listed:

**Frio Pilot** [e.g., Kharaka et al., 2006; Xu et al., 2010; Daley et al., 2011; Hovorka et al., 2011]. The Frio brine pilot site is located in southeast Texas, USA. The target injection reservoir, Frio formation, is made of Oligocene Frio sandstone. The initial pressure and temperature range from 15-16.6 MPa and 53-60 °C, respectively. Two small-scale CO$_2$ injection experiments were conducted at different parts of the formation. Frio-I brine pilot is located at 1541-1546 m, composed of Frio Formation “C” sandstone (23 m thick). Permeability and porosity of this formation are 2-3 Darcy and 32%, respectively. 1600 t CO$_2$ was injected at a rate of about 3 kg/s in 2004, which lasted for 10 days. The second injection was operated in 2006. CO$_2$ was injected in the Frio-II brine pilot, at which the injection formation is Blue Sand of Frio formation. It is located at 1657 m, and its thickness is around 17 m. Permeability and porosity are similar to Formation “C”, which are 1-4 Darcy and 30%. Only 380 t CO$_2$ was injected at this time, and the injection was operated for 5 days.

**Ketzin** [e.g., Förster et al., 2010; Norden and Frykman, 2013; Götz et al., 2014; Bergmann et al., 2016]. The Ketzin onshore site is located in the North German Basin. CO$_2$ is injected in the Stuttgart Formation, which was deposited at Upper Triassic age. The target reservoir is around 71-74 m thick, composed of sandstone of the Middle Keuper. The overlying caprock consists of mudstone, with a thickness of 160 m. The reservoir is located at a depth of 651-633 m, with a broad range of permeability (0.02-2700 mD). Porosity is around 2%-26%. The reservoir pressure and temperature are 6.2 MPa and 33 °C. From 2008 until 2013, more than 67 kt CO$_2$ was injected. The injection rate varied at different phases. On average, it is around 10 kt/yr.

**Sleipner** [e.g., Arts et al., 2005, 2008; Chadwick et al., 2010, 2016; Cavanagh and Haszeldine, 2014]. The Sleipner offshore site is located in the North Sea, and the host aquifer is Utsira Sand of late Cenozoic age, overlain by thick shale seals. This aquifer is at a depth of around 800-1000 m, with a thickness of more than 200 m. The high permeability (>1 Darcy) and porosity (35%-40%) on average provides good injection quality and a storage capacity (estimated storage efficiency is around 3×10$^5$ Mt). Pressure and Temperature range from 8-8.6 MPa and 29-37 °C, respectively. The industrial-scale CO$_2$ injection was first operated in 1996, and around 16 Mt CO$_2$ has been sequestered by 2015. CO$_2$ is injected 200 m below the aquifer top, at a rate of nearly 1 Mt/yr. The injection is planned to be continued until around 2020, remaining the injection rate at ~ 0.9 Mt/yr.

**Snøhvit** [e.g., Eiken et al., 2011; Hansen et al., 2013; Buscheck et al., 2016; Tasianas et al., 2016]. The Snøhvit offshore site is located in the Hammerfest basin in the Barents Sea. The target injection formation is the Tubåen Formation, which is located at a depth of 2600-2700 m below sea level, overlain by shale seals. The aquifer is composed of a deltaic to fluvial sand sequence, formed in the Early Jurassic. Its permeability is more than 500 mD, and porosity is around 10%-20%. The formation pressure and temperature are about 28 MPa and 80 °C. The CO$_2$ injection started from 2008, and by 2011, 1.09 Mt CO$_2$ was injected, at an average rate of 0.7 Mt/yr.

These four injection sites vary from small pilot scale to large commercial scale, and till now, abundant experiences are gained from their successful operations. However, potential hazards still
can be raised by CO₂ injection due to complicated geological conditions (e.g., high heterogeneity of the reservoir, complex faulting and fracturing). The injected CO₂ into saline aquifers can leak towards the shallower aquifers through the seal imperfections, such as pre-existing abandoned or failed-sealed wells and the transmissive faults or fractures present in the subsurface system (Figure 1.2) [Lemieux, 2011; Walter et al., 2012]. CO₂ will escape through the caprock because of the buoyancy effect and change to gas phase as it reaches shallow aquifers. Dissolved CO₂ in the natural fresh groundwater can lower the pH of the environment [Keating et al., 2010; Trautz et al., 2013], reducing the groundwater quality by enhancing the mobilization of trace elements [Apps et al., 2010]. Furthermore, formation water of saline aquifers is usually with high concentration of total dissolved solids (TDS) and other toxic substances [Kharaka et al., 2009]. Therefore, the displaced brine also possibly contaminates fresh/potable water, and it even alters the discharge behavior of the near-surface waterbody (e.g., lakes, streams or springs). Although the technologies required for CCS already exist, new concepts are still needed for long-term monitoring of the injection system [Nordbotten and Celia, 2011], to ensure that the CO₂ sequestration process is controllable and prompt remediation strategies can be formulated as soon as leakage occurs.

![Figure 1.2 Potential CO₂ leakage through the leaky well and connected fractures](image)

**1.2 Geophysical methods**

Geophysical methods have been developed and applied to most of the CO₂ sequestration sites for years. There exists several well-established approaches, for instance, seismic survey [e.g., Eiken et al., 2000; Arts et al., 2004a, 2004b, 2005; Saito et al., 2006; Lumley, 2010; Daley et al., 2011; Ajo-Franklin et al., 2013; Chadwick et al., 2016; Commer et al., 2016], electrical resistivity/conductivity survey [e.g., Eiken et al., 2000; Christensen et al., 2006; Giese et al., 2009; Zemke et al., 2010; Bergmann et al., 2012; Wagner et al., 2015] and gravity monitoring [Arts et al., 2004a, 2008; Gasperikova and Hoversten, 2008; Alnes et al., 2011]. Among these methods, seismic surveys are more extensively used due to their broad application scale. Seismic surveys usually
comprise well log and core measurements (sonic logging) [Xue et al., 2006; Azuma et al., 2011], vertical seismic profile (VSP) methods [e.g., Daley et al., 2007, 2011; Urosevic et al., 2011; Götz et al., 2014], cross-hole seismic methods [Hoversten et al., 2002; Saito et al., 2006; Daley et al., 2007; Xue et al., 2009; Ajo-Franklin et al., 2013], and surface seismic methods [Arts et al., 2004b; Juhlin et al., 2007; Chadwick et al., 2010; White et al., 2011; Ivandic et al., 2015]. These methods can be applied to delineate a CO₂ plume on different scales and with different resolution. When CO₂ phase exists, logging approaches can measure CO₂ saturation in the near borehole region. However, these cannot provide spatial information on CO₂ plume extension. The investigation scale of VSP and cross-hole seismic approaches is around 1-100 m, and that of surface seisms can be up to kilometers. Application of seismic tomographic configurations can derive spatial distribution of seismic properties, which are related to the spatial distribution of rock and fluid properties through certain rock physics models [e.g., Gassmann, 1951; Batzle and Wang, 1992; Mavko and Mikerji, 1998].

Seismic surveying applied for CO₂ plume detection is based on the propagation of elastic waves (P-waves, S-waves) in the subsurface. Seismic waves travel more slowly through a partly CO₂-saturated environment than through a pure water saturated formation (pushdown effect), particularly the P-wave velocity is reduced. Existing CO₂ saturation will cause a drop of the P-wave velocity by 5~60% [Xue and Lei, 2006; Lumey 2010; Vera, 2012]. On the contrary, S-wave velocity only slightly decreases. According to Vera [2012], it decreases by around 0.64%. Aside from the effect of velocity reduction, reflectivity is increased, which is sensitive to sharp saturation contrasts [Eiken et al., 2000]. Therefore, in order to detect CO₂ phase, two or three-dimensional P-wave velocity distributions can be mapped to determine the velocity field prior to injection. During CO₂ injection, arrival times of P-wave are determined at different point of times and are inverted to derive tomographic images of P-wave velocity distributions at certain time steps. The differences of the velocity fields determined between the measurements, performed at different times, reflect the evolution of CO₂ plume. This procedure is called four-dimensional seismic imaging.

The conversion of measured or inverted P-wave velocity into CO₂ saturation is a challenging task. First, there exists no direct general petrophysical model between P-wave velocity and CO₂ saturation. Available models are usually site-specific and non-linear. Second, seismic data quality is crucial to derive significant velocity tomograms. A good four-dimensional seismic monitoring program requires a very high signal to noise level, which is based on a high fluid compressibility contrast. Third, for the inversion of P-wave velocity tomograms, a relatively high and spatially regularly distributed data density is required. All the above mentioned issues are potential sources that can lead to artefacts during the inversion of the seismic data or during the conversion of seismic velocity into values of CO₂ saturation.
1.3 Pressure-based methods

In comparison with the geophysical approaches, pressure-based methods have several advantages. First of all, they are directly connected to the flow properties of the subsurface, especially the permeability. Variability of permeability is normally much larger than porosity, which is the key parameter to control the subsurface flow. Second, theoretically, the application scale of pressure-based methods can also be up to kilometers [e.g., Birkholzer et al., 2009; Zha et al., 2014, 2015] if the pressure stimulation is sufficiently large.

Therefore, pressure-based methods are deemed appealing for CO₂ sequestration problems. Previous studies indicate that pressure data can be used for:

**Estimating the integrated flow properties.** Single-well or inference hydraulic tests conducted prior to CO₂ injection could give a baseline information of the reservoir [e.g., Doughty et al., 2008; Wiese et al., 2010]. At the beginning of CO₂ injection, pressure measurements at the injection and monitoring wells can also be utilized for evaluating the permeability of the storage formation [Doughty et al., 2008]. In this case, the CO₂ is injected in small amounts, and it does not significantly change the initial permeability and porosity.

**Deriving the residual brine and CO₂ saturation.** Residual saturation is a crucial parameter which affects the long-term pressure evolution. [Doughty et al., 2008] estimated the residual brine and CO₂ saturations from the pressure measurements through a field CO₂ injection test at Frio brine pilot. Moreover, pressure measurements can be coupled with other available data to invert for the residual brine and CO₂ saturation, such as tracer breakthroughs and thermal responses [Rasmusson et al., 2014].

**Delineate the CO₂ plume size.** [Martinez-Landa et al., 2013] and [Mishra et al., 2013] derived several analytical solutions to roughly estimate the CO₂ front by analyzing the transient data at the injection or observation well, regardless the heterogeneity of the reservoir.

**Detecting brine/CO₂ leakage and calculating the leaky rate.** This is a main focus of recent research efforts on pressure-based methods. The brine/CO₂ leaky rate or amount at pre-existing abandoned wellbore or fault zones can be estimated by historical matching of pressure data using analytical or semi-analytical solutions [Nordbotten et al., 2004; 2008; Celia and Nordbotten, 2009; Zhou et al., 2009; Cihan et al., 2011; Meckel et al., 2013; Hosseine et al. 2014]., or by inverting the pressure anomalies derived from forward simulations [Sun and Nicot, 2012; Jung et al., 2013; Lee et al., 2015] or field data [Sun et al., 2016]. Additionally, detectability of the leakage events is also evaluated by several sensitivity analyses on the pressure data, considering a different degree of model uncertainties [Chabora and Benson, 2009; Azzolina et al., 2014; Wang and Small, 2014].

However, most of these pressure-based approaches are limited for providing the spatial information about flow properties, as well as the CO₂ plume extent under complicated geological conditions. This could yield incorrect prediction of the CO₂ plume migration and wrong estimation of the trapped CO₂ mass in the storage formation. In the following section, a novel pressure-based method, hydraulic (or pressure) tomography (HT), is introduced as an alternative approach to depict
spatial distribution of flow properties. The disadvantages of traditional hydraulic tests will be discussed first, and details on HT and the commonly used inversion algorithms will be given later.

1.4 Hydraulic tomography (HT)

In order to directly acquire hydraulic subsurface properties, different hydraulic approaches can be applied. Conventional hydraulic tests, such as single-well injection or pumping tests, allows for the derivation of bulk properties of an aquifer. However, interference hydraulic tests can only provide integrated information about hydraulic parameters at certain scale and dimension [Wu et al., 2005]. Spatially integrated hydraulic aquifer information is very often not sufficient to successfully predict solute transport in the subsurface [Freyberg, 1986; Yeh, 1992, 1998]. Hence, heterogeneity of the formation plays an important role in deep CO₂ injection projects, particularly with respect to plume migration [Liu and Kitanidis, 2011].

Another approach is to use information obtained from a dense well network and to reconstruct the subsurface structure utilizing geostatistical interpolation approaches. The problem of this procedure includes mainly two aspects. First, it requires information from many wells, which is usually not available especially for the costly deep boreholes. Second, geostatistical data analysis usually yields small scale information about medium properties, which are not sufficient and indicative for larger scale flow and transport problems [Illman et al., 2008].

Over the last twenty years, hydraulic tomography (HT) was first proposed by Gottlieb and Dietrich [1995], and has been developed as an innovative hydrogeological characterization approach. It offers high potential to provide spatial information about hydraulic subsurface heterogeneity. Up to now, HT has already been successfully applied to numerical studies [e.g., Yeh and Liu, 2000; Zhu and Yeh, 2005, 2006; Jiménez et al., 2013], to lab experiments [e.g., Liu et al., 2002, 2007; Illman et al., 2007, 2010; Berg and Illman, 2011], and to field studies [e.g., Bohling et al., 2007; Straface et al., 2007; Cardiff et al., 2009, 2013; Hu et al., 2011; Brauchler et al., 2013a; Jiménez et al., 2015] It can also be employed to unconfined aquifers [e.g., Cardiff and Barrash, 2011; Mao et al., 2013] as well as fractured aquifers [e.g., Hao et al., 2008; Illman et al., 2009; Zha et al., 2015].

The usual implementation of hydraulic tomography consists of performing the hydraulic tests at different locations (sources) and measuring the pressure signals at different observations (receivers). Two representative hydraulic tests are slug tests [Brauchler et al., 2010; Paradis et al., 2015] and pumping tests [e.g., Yeh and Zhu, 2007; Hu et al., 2011]. Due to their different features, the application scale of slug tests is usually smaller than pumping tests. However, slug tests are associated with relatively lower costs and shorter experimental duration. Moreover, no water is extracted during the experiments and it does not require any post-treatment. Cardiff et al. [2013a] proposed using oscillatory pumping tests at different frequencies to generate the information for the tomographic inversion. Similar to slug tests, no water needs to be handled during and after the experiments. Compared to constant-rate pumping tests, signal processing is more robust, since it is
less sensitive to the data noise. Both slug tests and pumping tests can be considered as “active” methods, as the stimulations are generated and controlled artificially. Recently, an extended concept of HT, “river tomography”, was proposed by Yeh et al. [2004]. The “passive” natural sources are considered as hydraulic stimuli, such as river stage variations. Even though the sources are unpredictable in this case, river tomography is promising for reconstructing basin-scale subsurface structure, which greatly widens the application window of HT [Yeh et al., 2009].

In the following a summary of the most popular inversion schemes is given. Yeh et al. [1996] developed a successive linear estimator (SLE) based inversion scheme of hydraulic tomography. The general algorithm of SLE is to generate an initial cokriged, mean-removed logarithm transmissivity and hydraulic head map based on observed hydraulic conductivity values and head data using a classical cokriging approach. Subsequently, a successive linear estimator is applied by solving the steady state groundwater flow equation to update the transmissivity and head field until the covariances and cross-covariance of the estimated transmissivity and head values between two successive iterations is sufficiently small. The SLE method improves the estimates on conductivity field compared to the classic cokriging approach [e.g., Hoeksema and Kitanidis, 1984, 1989; Ahmed and De Marsily, 1993]. Furthermore, the SLE is demonstrated by Hughson and Yeh [1998, 2000] as a more computationally efficient inversion scheme than other classical inverse methods (summarized in Yeh [1986]). However, the SLE uses the data set from one single hydraulic test or simultaneously includes all the data from a series of sequential hydraulic tests. The algorithm of this method can become ill-conditioned and unstable. Yeh and Liu [2000] further developed the SLE based inversion approach to a sequential successive linear estimator (SSLE) approach. The main improvement of the SSLE is that all the obtained pumping data sets are applied sequentially during iteration. The estimated hydraulic conductivity field and covariances are conditioned on previous sets of head data, and they are employed to the next estimation step according to a new set of pumping data. The iteration continues until all the data sets are fully utilized. The SSLE can increase the computing efficiency by remaining the small covariance matrix size. Moreover, it provides an unbiased conditional mean estimate which can reduce the solution non-uniqueness. Liu et al. [2002, 2007] has employed the SSLE method in laboratory experiments. Furthermore, Illman et al. [2008] tested the SSLE approach in a laboratory aquifer using different pumping rates and signal-to-noise (S/N) ratios.

Zhu and Yeh [2005] further developed the SSLE approach to invert for both hydraulic conductivity and specific storage using transient drawdown data. Straface et al. [2007] applied this transient SSLE method successfully in the field. However, using the transient data leads to significant computational burden. To solve this problem, an inversion method of matching temporal moments of full transient drawdown curves by the SSLE was proposed by Zhu and Yeh [2006] which is similar to the suggestions by Li et al., [2005]. With this approach, the selected temporal moments of drawdown data are applied to the inversion instead of the full drawdown curves. Yin and Illman [2009] tested this temporal moment approach in a laboratory sandbox.

Xiang et al. [2009] applied the SLE method using the head data sets simultaneously (SimSLE) instead of sequential application. This avoids the loop iteration of SSLE by the estimates of the
adjoint state equation. However, it requires data sets with good quality, and it also demands large memory to compute the covariance matrix. Mao et al. [2013] extended the application of the SimSLE to a variable-saturated, unconfined aquifer. They concluded that sequential pumping tests can identify the spatial hydraulic conductivity in both saturated and unsaturated zone, and distribution of specific storage can only be characterized in the saturated zone.

Zha et al. [2014] jointly inverted drawdown data and flux measurements to derived the hydraulic conductivity distribution of a two-dimensional synthetic fracture media based on the SSLE algorithm. The results show a good agreement with the “true” hydraulic conductivity field, and the inversion is not sensitive to the data noise level (has been examined at 2%-20% noise level in that work). Further, in Zha et al. [2015], drawdown data obtained from the field pumping tests at the Mizumi site are used for inverting the hydraulic conductivity and specific storage using the SSLE algorithm, as a case study for validating the feasibility of HT for characterizing large scale fracture formation.

Bohling et al. [2002, 2007] proposed an inversion approach named steady shape flow regime. They found that although drawdown is changing with time, the head difference between observation points, which is related to hydraulic conductivity, do not vary in steady state. Hence, hydraulic gradient between observation points at steady state are employed for estimates of hydraulic conductivity. Steady shape can be established quickly according to the work of Hu et al. [2011]. They applied this method to both synthetic dataset and field data. The drawdown data can be analyzed in a computationally efficient way.

Quasi-linear Bayesian geostatistical method [Kitanidis, 1995] is another inversion approach used for hydraulic tomography [e.g., Snodgrass and Kitanidis, 1998; Nowak et al., 2003; Nowak and Cirpka, 2004; Fienen et al., 2008; Li et al., 2008; Cardiff et al., 2009; Cardiff and Barrash, 2011]. Bayes’ theorem is the basis of this approach. Cardiff et al. [2009] and Cardiff and Barrash [2011] utilized it in an unconfined aquifer to get the reconstruction of the depth-integrated hydraulic conductivity field. Further, Zhou et al., [2016] applied the quasi-linear geostatistical method to invert for the hydraulic conductivity and specific storage fields based on not the drawdown data, but on the measured sinusoidal and cosinusoidal coefficients obtained by Fast Fourier Transformation (FFT). Correspondingly, the forward model used for the inversion is based on a reformulated phasor-based (steady-periodic) groundwater flow equation, instead of a standard groundwater model.

Ensemble Kalman filter (EnKFs) is applied for characterizing flow in subsurface hydrology for years [e.g., Chen and Zhang, 2006; Hendricks Franssen and Kinzelbach, 2008; Tong et al., 2010]. It is a stochastic inverse modeling approach to estimate the state variables and model parameters by using a Bayesian updating scheme. The EnKFs are efficient when variables are characterized by a multivariate Gaussian dependence. However, in most situations this assumption is not valid. A non-linear, monotonous transformation is proposed by Schöniger et al. [2012] and Nowak et al. [2013] to invert synthetic drawdown data. The drawdown data were simulated using multi-Gaussian log-conductivity fields. The transformed drawdown data are linearly dependent on log-conductivity values. The linearized dependence enhances the processing quality of the
available information, and increases the accuracy of parameter identification, which is the key to predict flow and transport. Moreover, it offers the potential to solve high-dimensional problems with a high computational efficiency [Hendricks Franssen and Kinzelbach, 2009].

An alternative tomographical inversion algorithm is an eikonal based method. The transient groundwater flow equation is approximated through an eikonal equation by applying an asymptotic solution [Virieux et al., 1994]. Pressure travelling along trajectories is calculated based on the eikonal equation, which is solved commonly by ray-tracing or particle tracking techniques [e.g., Kulkarni et al., 2000; Vasco et al., 2000; Datta-Gupta et al., 2011; Brauchler et al., 2003, 2007, 2010; He et al., 2006; Vasco and Karasaki, 2006]. The “speed” of pressure propagation depends on the hydraulic diffusivity (i.e., the ratio of hydraulic conductivity and specific storage) field between the testing and observation wells. Brauchler et al. [2003] proposed the travel-time based approach to invert the spatial distribution of hydraulic diffusivity. The inversion procedure can be used for Dirac pulse (slug tests) or Heaviside pulse (pumping test) at the origin, which is based on a travel time line integral relating the square root of the peak travel time to the inverse square root of the hydraulic diffusivity. In analogy to this work, Brauchler et al. [2011] derived an attenuation integral, which relates the attenuation of a pressure signal, originating from a Dirac source, to the specific storage distribution as a function of arc-length along the propagation path. Brauchler et al. [2013] further developed this method and adapted it to the requirements of pumping test data (Heaviside signal). Compared to the other hydraulic tomography approaches, this travel-time based approach does not invert full pressure signals. Therefore, the inversion scheme is computationally fast. Furthermore, the structural information gained from the diffusivity tomogram can be used as a starting model for a sequential inversion scheme, such as a steady shape analysis [Hu et al., 2011] or a pilot point based approach [Jiménez et al. 2013, 2015].

1.5 Objectives and structure of this work

This thesis aims to developing and applying a two-phase pressure tomography method for characterizing an evolving CO₂ plume in deep saline formations. This implies in detail:

- Utilize a travel-time based inversion scheme to reconstruct the subsurface structure, and characterize the CO₂ plume shape and saturation at different times based on a new single-phase proxy
- Design a novel test sequence of interference fluid injections to examine the feasibility of pressure tomography for detecting CO₂ leakage
- Compare pressure tomography to existing geophysical approaches, evaluate capability and limitations for identifying the CO₂ plume at different times
- Couple pressure tomography with geophysical approaches to improve saturation estimation

In the following, Chapter 2 gives a brief description of the fundamentals of CO₂ sequestration in deep saline aquifers. The complexities of a CO₂-brine system are elaborated point by point. Chapter 3 shows the development of two-phase pressure tomography and how it can be used for
identifying a CO$_2$ plume. Chapter 4 is a case study of pressure tomography employed for detecting CO$_2$ leakage. Chapter 5 examines the feasibility of pressure tomography in heterogeneous formations, while cross-hole seismic tomography is implemented under the identical conditions to compare the inversion performance of the two tomographical approaches. Additionally, the results of the two methods are jointly evaluated.
2 Fundamentals of CO₂ sequestration in saline aquifers

2.1 Trapping mechanism

During and after CO₂ injection into deep saline aquifers, it can be trapped in the reservoir in four ways (Figure 2.1):

- CO₂ can be hydrodynamically trapped as a gas or in a supercritical state in the aquifer [Bachu, 2000; Xu et al., 2003]. This includes stratigraphic/structural trapping and residual trapping processes. Migration of CO₂ at the early injection stage is mainly dominated by the pressure gradient and buoyancy. CO₂ is anticipated to be trapped by impervious stratigraphic or structural seal units (e.g., a very low permeability caprock and a deactivated fault barrier) at this stage. Structural traps sometimes play a less important role than caprock seals, in the case that CO₂ is injected far away from the reservoir boundaries [Vishal and Singh, 2016].

- During the stratigraphic/structural trapping processes, a part of CO₂ enters into the pore spaces by capillary pressure. This part of CO₂ is stored by residual fluid trapping, remaining immobile for a long time.

- CO₂ can dissolve in the brine. The solubility of CO₂ depends on formation pressure, temperature and salinity. This is called solubility trapping and is important for the determination of potential for CO₂ capture through geochemical reactions [Bachu and Adams, 2003]. In most deep saline formations, solubility of CO₂ into brine decreases with the increased formation temperature and salinity, and increases with the augmented pressure [e.g., Duan and Sun, 2003]

- The dissolved CO₂ in water can react with minerals in the geologic formation, leading to the mineral trapping process, through which CO₂ is precipitated to secondary carbonates [Class et al., 2002].

![Figure 2.1 Conceptual scheme for the four major CO₂ trapping mechanisms (modified from Saeedi [2012])](image-url)
During the entire lifecycle of a CO$_2$ sequestration process, stratigraphic/structural trapping and residual trapping are dominated for the early times (Figure 2.2). On the contrary, solubility and mineral trapping take a relatively long time during the sequestration. If they take place, then leakage events occur most likely at the stratigraphic/structural and residual trapping phases. The solubility and mineral trapping mechanisms mainly depend on chemical reactions, with minor associated potential risks.

![Figure 2.2 Trapping mechanisms and time scale effect of CO$_2$ sequestration [IPCC, 2005]](image)

### 2.2 Two-phase flow in porous media

#### 2.2.1 Mass and energy conservation equations

The general mass conservation equation for multiphase and multicomponent problem can be expressed as [Helmig, 1997]:

$$
\frac{\partial}{\partial t} \left[ f \left( \chi_n \rho_n S_n + \chi_{nw} \rho_w S_w \right) \right] = \nabla \cdot \left[ \frac{\chi_n \rho_n k k_{mn}}{\mu_n} \left( \nabla P_n - \rho_n g \nabla z \right) \right] + Q_n + Q_{nw} + F_{nw}
$$

(2.1)

where $\chi_{in}$ and $\chi_{iw}$ are the mass fraction of component $i$ in wetting and non-wetting phase respectively. $\rho$, $S$, $f$, $k$, $k_r$, $P$ and $Q$ refer to phase density, saturation, rock porosity, rock instinct permeability, phase pressure and production/injection amount. $F_{nw}$ is the diffusion of non-wetting phase in wetting phase. The subscript $n$ and $w$ indicate non-wetting phase and wetting phase, respectively.

The energy conservation equation is formulated as [Lu and Lichtner, 2005]:

$$
\frac{\partial}{\partial t} \left[ f \sum_j S_j \rho_j U_j + (1-f) \rho_c c_p T \right] + \nabla \cdot \left[ \sum_j \left( q_j \rho_j H_j - \kappa \nabla T \right) \right] = Q_e
$$

(2.2)
where $Q_e$ is the sink/source term of heat, $U_j$ is the internal energy of phase $j$ ($j = n, w$), and $H_j$ is the enthalpy of phase $j$. $\rho_r$, $c_p$, $\kappa$ refer to rock density, heat capacity and thermal conductivity of the porous rock.

In order to solve these two equations, two closure conditions are utilized:

$$S_u + S_w = 1 \quad (2.3)$$

$$P_e(S_u) = P_u - P_w \quad (2.4)$$

### 2.2.2 Relative permeability and capillarity

Two alternative models for calculating the values of relative permeability are widely used, i.e. the Brooks and Corey model [Brooks and Corey, 1964] and the van Genuchten model [van Genuchten, 1980].

The relationship between the relative permeability and the effective saturation in the Brooks and Corey model is determined by following Equations (2.5) to (2.8). In these equations, $k_{rw}$ and $k_{rn}$ are relative permeability of wetting phase and non-wetting phase, respectively. $S_e$ is the effective saturation, defined by the phase saturation ($S_w$ and $S_n$) and residual saturation of wetting phase and non-wetting phase ($S_{wr}$ and $S_{nr}$). The pressure difference between two phases, capillary pressure $P_c$, is determined by entry pressure $P_d$ and effective saturation $S_e$. $\lambda$ is the pore size distribution parameter, which varies from materials and specific site conditions. Usually, it varies from 0.2 to 3 [Helmig, 1997]. Small values indicate a more heterogeneous material, and large values are characteristic for homogeneous material.

$$k_{rw} = S_e^{2+3\lambda} \quad (2.5)$$

$$k_{rn} = (1 - S_e)^2 \left(1 - S_e^{2+3\lambda} \right) \quad (2.6)$$

$$S_e = \frac{S_w - S_{wr}}{1 - S_{wr} - S_{nr}} \quad (2.7)$$

$$P_c = P_d S_e^{\frac{1}{\lambda}} \quad (2.8)$$

The relative permeability and capillary pressure based on the van Genuchten model are calculated by Equations (2.9) to (2.12):

$$\alpha = \frac{1}{P_e(S_e = 0.5)} k_{rw} = \sqrt{S_e \left[1 - (1 - S_e^{1/\lambda})^m \right]^2} \quad (2.9)$$
\[ k_m = (1 - S_e)^{1/3} \left(1 - S_e^{1/n}\right)^{2m} \]  \hspace{1cm} (2.10)

\[ P_e = \frac{1}{\alpha} \left(S_e^{-1/m} - 1\right)^{1/\alpha} \]  \hspace{1cm} (2.11)

\[ \alpha = \frac{1}{P_e \left(S_e = 0.5\right)} \]  \hspace{1cm} (2.12)

where \( m \) and \( n \) are geometry parameters describing the pore connectivity [Mualem, 1976], and \( m \) equals \( \left(1 - \frac{1}{n}\right) \).

### 2.3 Fluid properties

#### 2.3.1 CO\(_2\) properties

In the following, a summary of the most common models to calculate a) density, b) viscosity, c) solubility, and d) diffusion of CO\(_2\) are given:

- **Density of CO\(_2\)** is increasing with increasing pressure, and decreasing with an increase in temperature. The density of the gas phase is lower than 200 kg/m\(^3\), while the liquid phase is around 600~1200 kg/m\(^3\). The density of supercritical CO\(_2\) is around 200~1000 kg/m\(^3\), which is close to the liquid state. The density of CO\(_2\) can be calculated based on a real gas state equation. Common models used for the calculation of CO\(_2\) density are the Soave-Kwong-Redlich model [Soave, 1972], the Peng-Robinson model [Peng and Robinson, 1976], and the Span and Wagner model [Span and Wagner, 1996]. In this work, the Span and Wagner model (Figure 2.3a) is applied.

- **The change of CO\(_2\) viscosity as a function of temperature and pressure** is similar to the change of density. The viscosity of the gas phase is smaller than 0.02 mPa s, while that of the liquid phase is around 0.04~0.18 mPa s. Viscosity of supercritical CO\(_2\) is around 0.02~0.14 mPa s. Several models exist to calculate the viscosity of CO\(_2\), such as the LBC model [Lohrenz et al., 1964], the JST model [Jossi et al., 1962], the models by [Chung et al., 1984], and by [Fenghour et al., 1998]. In this work, the variant by Fenghour et al. [1998] is applied (Figure 2.3b).
• Solubility is a function of temperature, pressure and salinity. In most cases, the solubility is increasing with rising pressure, and decreasing with increase of temperature and salinity. The solubility of CO$_2$ is around 8% (mass fraction) in fresh water. The solubility of CO$_2$ in fresh water and brine can be calculated using Henry’s Law, the Spycher model [Spycher and Pruess, 2004] and the Duan and Sun model [Duan and Sun, 2003]. In our case, the model by Duan and Sun [2003] is applied to estimate the CO$_2$ solubility in brine, since it is more accurate than other models (the error between the calculated and experimental data is about 3%).

• The diffusion coefficient of CO$_2$ in water is influenced by the temperature, pressure and the content of the mixture. The diffusion coefficient in the gas phase is around $10^{-5}$~$10^{-4}$ m$^2$/s, in the liquid phase it is around $10^{-10}$~$10^{-9}$ m$^2$/s. The diffusion coefficient can be calculated through the Wilke-Chang equation [Wilke and Chang, 1955]. Independent from pressure, the diffusion coefficient of CO$_2$ is increasing with temperature and salinity. However, the effect of diffusion is slight, even over long geological time scales [Vishal and Singh, 2016].

2.3.2 Brine properties

In this section, a summary of the most popular models to calculate a) density and b) viscosity of the brine is given:

• Brine density is related to the salinity and the amount of dissolved CO$_2$, which are determined by pressure and temperature. Density increases with the decreased temperature, and also with the increased amount of dissolved CO$_2$. Pressure has minor effects on the density change, as the compressibility of the brine is relatively small. In general, variations in density can cause brine convection, and by this accelerate the dissolution of CO$_2$. Thus it has impact on the long-term CO$_2$ migration and storage processes. Several models can be applied for calculating the
brine density [e.g., Batzle and Wang, 1992], and in this study, the one by Duan et al. [2008] is used.

- The viscosity of the brine is mainly influenced by the temperature and salinity, whereby the influence of pressure can be neglected. CO$_2$ dissolution increases the viscosity of the brine at constant temperature. The formulation by IFC [1967] is utilized in this work for calculating the brine viscosity.
3 Time-lapse pressure tomography for characterizing CO₂ plume evolution in a deep saline aquifer


Abstract

A time-lapse pressure tomography inversion approach is applied to characterize the CO₂ plume development in a virtual deep saline aquifer. Deep CO₂ injection leads to flow properties of the mixed-phase, which vary depending on the CO₂ saturation. Analogous to the crossed ray paths of a seismic tomographic experiment, pressure tomography creates streamline patterns by injecting brine prior to CO₂ injection or by injecting small amounts of CO₂ into the two-phase (brine and CO₂) system at different depths. In a first step, the introduced pressure responses at observation locations are utilized for a computationally rapid and efficient eikonal equation based inversion to reconstruct the heterogeneity of the subsurface with diffusivity (D) tomograms. Information about the plume shape can be derived by comparing D-tomograms of the aquifer at different times. In a second step, the aquifer is subdivided into two zones of constant values of hydraulic conductivity (K) and specific storage (Sₜ) through a clustering approach. For the CO₂ plume, mixed-phase K and Sₜ values are estimated by minimizing the difference between calculated and “true” pressure responses using a single-phase flow simulator to reduce the computing complexity. Finally, the estimated flow property is converted to gas saturation by a single-phase proxy, which represents an integrated value of the plume. This novel approach is tested first with a doublet well configuration, and it reveals a great potential of pressure tomography based concepts for characterizing and monitoring deep aquifers, as well as the evolution of a CO₂ plume. Still, field-testing will be required for better assessing the applicability of this approach.
3.1 Introduction

Despite the rising awareness of global environmental changes and increased efforts in establishing renewable energy sources, global carbon dioxide (CO$_2$) emissions are still on a rise. Therefore, in the low-carbon energy strategies of many countries, carbon capture and storage (CCS) plays a prominent role. Instead of releasing CO$_2$ to the atmosphere, CCS implies the capturing CO$_2$ at the site of large producers and the storage in deep geological formations, mainly saline aquifers. Even though numerous test sites have been developed, only in a few cases large, industrial-scale amounts of CO$_2$ are injected into the subsurface, for instance, in former oil or gas reservoirs. Much attention has been paid to injecting CO$_2$ into deep saline aquifers. These aquifers potentially offer enormous storage capacities [Bachu and Adams, 2003]. However, large-scale applications are still scarce and thus long-term experience is missing. Injection of CO$_2$ into the subsurface involves many unknowns. Perhaps the most critical among these is the uncontrolled migration of CO$_2$. The injection of CO$_2$ in deep saline aquifer generates a plume, which may leak into overlying strata along existing geological fault/fractures or along abandoned wells or leaky well completion [Lemieux, 2011]. Moreover, possible hydro-fracture propagation during long-term pressure evolution can create new leakage pathways. This underlines the importance of monitoring CO$_2$ plume behavior during and after short- and long-term injection periods.

Common approaches to characterize CO$_2$ plumes are based on established geophysical techniques, such as seismic surveys [e.g., Ajo-Franklin et al., 2013], electrical resistivity/conductivity surveys [e.g., Bergmann et al., 2012], and gravity monitoring [e.g., Chadwick, 2006]. Additionally, distributed acoustic sensing is considered as an innovative means for CO$_2$ monitoring [e.g., Daley et al., 2013]. All of these techniques represent indirect ways of characterizing the CO$_2$ plume and its migration. None of these offers a direct relationship between CO$_2$ saturation and the hydraulic regime in the deep aquifer. For instance, in seismic surveys the relationship between seismic waves and CO$_2$ saturation is site-specific, non-linear, and associated with a high signal-to-noise ratio. Moreover, for seismic tomography and electrical resistance tomography (ERT), the coarse spatial resolution may be a limiting factor, given that residual CO$_2$ often forms meter-scale plumes [Martinez-Landa et al., 2013].

Apart from geophysical approaches, fluid injection/extraction tests (water, CO$_2$ or water-CO$_2$ mixture) can be used to derive two-phase flow properties. Two-phase flow properties are different from single-phase flow properties because injected CO$_2$ will introduce a high compressibility to the flow system [Vilarrasa et al., 2010]. Further, the total mobility (summation of the separate phase mobility) of the phase mixture (CO$_2$, water or brine) in the aquifer varies due to the lower viscosity of the CO$_2$ phase. All these effects directly influence the pressure responses from fluid injection tests.

Previous studies have shown that both single-phase and two-phase flow properties for can be derived from fluid injection/extraction tests by pressure data analysis. Interference pumping tests prior to CO$_2$ injection were proposed by Doughty et al. [2004] for characterizing the single-phase flow properties of the potential storage formation. Wiese et al. [2010] pointed out that pressure data
obtained from single-well and crosswell pumping tests conducted before CO$_2$ injection can be used to identify the hydraulic connectivity and the boundary type of the formation at the Ketzin site, Germany. The pressure evolution during CO$_2$ injection can also be used for estimating flow properties. When CO$_2$ is injected in small amounts, pressure data from the injection well and a monitoring well can be used to constrain the permeability of the formation [Sminchak et al., 2009]. Doughty et al. [2008] estimated residual brine and CO$_2$ saturations from transient pressure data derived from a CO$_2$ injection test at Frio Brine Pilot, Texas, United States. Sun and Nicot [2012] developed a pressure anomaly inversion procedure for detecting CO$_2$ leakage locations and rates from abandoned wells. Martinez-Landa et al. [2013] proposed a single-well CO$_2$ saturated brine injection test in a brine-CO$_2$ system with a low injection rate to identify the residual CO$_2$ saturation and the width of the CO$_2$ plume near the well by pressure data analysis. Mishra et al. [2013] derived an analytical solution to track a CO$_2$ front by analyzing the transient pressure data at the observation well. However, none of these approaches can delineate the spatial extent of the CO$_2$ plume based on pressure data.

Detailed knowledge of aquifer properties necessary for tracking the CO$_2$ plume is typically hampered by the small number of available exploration boreholes. Conventional injection/extraction tests can only provide integral information of natural or induced flow properties of an aquifer. Crosswell injection/extraction tests that work with well pairs can, to some extent, provide averaged flow parameters between wells [Wu et al., 2005]. In the field of petroleum reservoir engineering, interference tests and pulse tests are commonly used to characterize the reservoir [e.g., Fokker et al., 2012], as well as to inspect two-phase properties [e.g., Finsterle and Pruess, 1996; Fokker and Verga, 2011]. These tests, however, cannot delineate the spatial heterogeneity of the subsurface [Freyberg, 1986]. Geostatistical interpolation can be applied to reconstruct flow properties of the subsurface based on the data from a dense borehole network. However, because boreholes penetrating into deep aquifers are scarce, borehole-based geostatistical data analysis cannot provide accurate information of medium properties for large-scale flow and transport problems [Illman et al., 2007].

During the last twenty years, hydraulic tomography evolved as an alternative to conventional hydraulic field testing, because it is specifically suited for reconstructing the spatial distribution of hydraulic aquifer parameters [e.g., Butler et al., 1999; Yeh and Liu, 2000; Bohling, 2007; Brauchler et al., 2007, 2011; Mao et al., 2013]. This approach is based on a series of hydraulic tests with a single fluid, i.e. water. The method utilizes a series of induced pressure changes, which are used to reconstruct spatial heterogeneity of single-phase flow properties by an inversion scheme. In analogy to seismic tomography, a series of space-filling streamline patterns can be generated by a combination of several different injection intervals (source) and observation points (receiver) [Butler et al., 1999]. With sufficiently recorded data, an appropriate inverse model can then be used to obtain a reliable image of the subsurface.

One variant of hydraulic tomographical inversion is based on the approximation of the transient groundwater flow equation by an eikonal equation [Virieux et al., 1994]. Ray tracing or particle tracking techniques are commonly applied to solve the eikonal equation by the calculation
of pressure propagation paths [e.g., Kulkarni et al., 2000; Vasco et al., 2000; Brauchler et al., 2003, 2010]. A core element in hydraulic tomography is that hydraulic diffusivity (i.e., the ratio of hydraulic conductivity and specific storage) is directly related to measured travel times of pressure perturbation. In analogy to seismic travel time inversion, spatial distribution of hydraulic diffusivity thus can be obtained in a computationally rapid and efficient way. However, the inverted diffusivity tomogram can only serve as a proxy mapping of the true diffusivity structure of the aquifer. Also specific values of the hydraulic conductivity and the specific storage cannot be derived from inverted tomograms. Hence, the reconstructed structure is usually employed as the starting model for a second inversion step based on full signal calibration with a forward groundwater flow model [Hu et al., 2011; Jiménez et al., 2013].

Similar to the travel-time based hydraulic tomography approach in shallow aquifers, pressure tomography can be used to identify, characterize and monitor the spatial extent of the CO₂ plume. The underlying idea is that through injection, a mixed brine-CO₂ phase forms in the deep aquifer, which induces an apparent transient heterogeneity in the hydraulic properties of the system. Short-term fluid or gas injection in combination with a tomographical set-up can facilitate a rapid and efficient travel-time based inversion. This will provide qualitative aquifer diffusivity maps that can be utilized for tracking the migrating CO₂ plume. However, quantitative estimates of the spatial distribution of CO₂ saturation are only possible by calibration with an appropriate multi-phase forward model. Also, in comparison to hydraulic tomography in shallow aquifers, practical limitations will be the costs for drilling and instrumentation of multiple deep boreholes, which are needed for obtaining a spatial resolution.

The main objective of this paper is to develop a fundamentally new pressure-based tomographical approach to delineate the CO₂ plume. In contrast to other geophysical methods this approach has potential to directly compute spatially averaged CO₂ saturations by the two-phase flow properties. In the following, we shortly describe the travel-time based procedure for deriving diffusivity tomograms. The structural information from these tomograms is extracted by zonal clustering, which in a homogeneous aquifer can be used to determine the shape of the plume. We then discuss ways of how to simplify computationally demanding multi-phase flow models. We propose a streamlined single-phase model, which emulates the multi-phase system, while substantially speeding up the ultimate calibration of the forward model. As an example for a real case application, we choose a two-dimensional simplified virtual site to develop our approach.

3.2 Methodology

3.2.1 Overview: cross well testing and inversion

For the injection model used in this study it is assumed that CO₂ is injected through a source well in a deep aquifer, and pressure changes are monitored in one nearby observation well that acts as a receiver. These pressure changes originate from injection at different depths at the source and are recorded at different levels along the observation well screen. A packer system allows
partitioning of the injection interval at the source well. This facilitates a high-resolution tomographic analysis of multiple source-receiver combinations. The pressure pulses may be introduced by water or brine injection ahead of the CO₂ injection in order to characterize the pre-injection hydraulic conditions in the aquifer. Alternatively, it is also possible to use depth-dependent short-term injection of CO₂ at the source intermittently to the sequestration procedure. Regardless of the injected compound and operation stage, the same procedure of pressure tomography will be applied for characterizing the reservoir.

By comparing the tomograms from different points in time, the propagation of the CO₂ plume can be recorded. In a first step, the inversion procedure of this time-lapse pressure-based tomography utilizes a common travel-time based hydraulic tomographical approach for structural imaging. Travel times can then be used to gain first order insights into the aquifer structure and structural changes. In a second step, the structured aquifer models are calibrated to the full pressure-response curves (zonal calibration). For reducing simulation time and increasing computational efficiency, the complex multi-phase conditions in the aquifer are approximated by a novel single-phase based emulator.

### 3.2.2 Pressure-based tomographical inversion

Our travel time inversion approach to obtain hydraulic tomograms from crosswell pressure tests is based on the work by Vasco et al. [2000]. With their approach, the groundwater flow equation is approximated by an eikonal equation, efficiently solved by ray tracing or particle tracking techniques. Consider a transient pressure curve, resulting from a Dirac signal generated at point \( x_1 \) (source), traveling along the propagation path \( s \), and recorded at point \( x_2 \) (observation). The relationship between hydraulic travel time and hydraulic diffusivity is expressed by a line integral:

\[
\sqrt{t_{\text{peak}}} = \frac{1}{\sqrt{6}} \int_{s_1}^{s_2} \frac{ds}{D(s)} \tag{3.1}
\]

where \( t_{\text{peak}} \) is the peak travel time of the recorded pressure curve, and \( D \) is the hydraulic diffusivity.

For our experiments, Heaviside type sources (i.e., the continuous injection) are applied in order to obtain more pronounced and far-reaching pressure signals compared to those from a Dirac type source. To use the inversion scheme to the Heaviside type injection tests, the first time-derivative of the pressure readings is required [Vasco et al., 2000]. As found in the hydraulic tomography study by Hu et al. [2011], as well as in our work shown below, the peak times are not distinct, which is indicated by small variation of the time-derivative at the peak. Hence, to improve resolution early travel time diagnostic \( (t_{\text{e},d}) \) is used for the inversion. In fact, higher conductive parts of the subsurface can be resolved better by early travel time than by peak time diagnostic [Cheng et al., 2009]. Further, the early travel times can be more accurately obtained than the peak times, because the curve of the time-derivative of pressure has a much sharper slope before the
peak time. The early travel time diagnostic is computed by introducing a conversion factor \( f_{a,d} \) [Brauchler et al., 2003]:

\[
\sqrt{t_{a,d}} = \frac{1}{\sqrt{6f_{a,d}}} \int_{s_0}^{s_f} \frac{ds}{D(s)}
\]  

(3.2)

with

\[
f_{a,d} = -W\left(\frac{a_{2/3}}{e}\right)
\]  

(3.3)

where \( W \) is Lambert’s \( W \) function, and \( a_d \) corresponds to the hydraulic head ratio term defined as \( \frac{h_d(r,t)}{h_d(r,t_{peak})} \), where \( h_d(r,t_{peak}) \) is the maximum value of the first time-derivative of pressure data for a Heaviside pulse. \( h_d(r,t) \) is the first time-derivative of pressure as a function of time and space. In this study, we applied the 20% travel time diagnostic (\( t_{20\%} \), i.e. the time at which the pressure pulse rises to 20% of its ultimate peak value) for the inversion procedure [Hu et al., 2011]. In our simulation study, wellbore storage is neglected. However, in practice, it can cause a delay of travel times [Prats and Scott, 1975]. The influence of this delay decreases with distance between source and receiver. It can be estimated analytically and included in the travel time based inversion [Brauchler et al., 2007].

Following the travel-time based inversion, reliability of the inverted tomograms is evaluated by using the null space energy as indicator. For the concept of null space energy maps, we utilize the procedure proposed by Böhm and Vesnaver [Böhm and Vesnaver, 1996]. They estimated the ray distribution in pixels of a tomographical model to assess the stability of seismic velocity tomograms. A tomographical matrix \( A \) is decomposed into two orthonormal matrices (\( U \) and \( V \)) and a diagonal matrix \( W \) by singular value decomposition (SVD):

\[
A = UVV^T
\]  

(3.4)

where the elements of matrix \( A \) represent the total length of propagation paths in each pixel of the tomogram. The reliability of the inversion result can be assessed by the singular values, which are the elements of matrix \( W \). Small or zero values of the elements in matrix \( W \) indicate large instability in the solution of a tomographical system. However, the singular values can only be considered as a global indicator for reliability estimation. Each singular value relates to a column of the matrix \( V \), and it cannot provide detailed reliability at each pixel. Thus, a local reliability indicator, \( M \), is defined to describe the reliability of each pixel for the tomogram. The summation of the squared column element (\( v_i \)) of matrix \( V \) implies the orthonormal basis of the model domain:

\[
M = \sum_i v_i
\]  

(3.5)
The value of $M$ ranks from 0 to 1. Here, we use $M = 0.5$ as the threshold to judge the reliability of inverted results. Inverted diffusivity values with $M$ smaller than 0.5 are considered as reliable, and if $M$ is greater than 0.5, the values per pixel are considered as unreliable.

3.2.3 Clustering and zonal calibration

The inverted tomograms provide estimates of the cross-sectional diffusivity distribution at different times during sequestration, including pre-injection and post-injection conditions. To derive the plume distribution, the diffusivity data sets are partitioned into two clusters, one representing the plume and the other the ambient aquifer. k-means clustering is applied and two zones with constant conductivity and specific storage values are obtained. Through clustering and by comparison with the tomogram derived for the pre-injection stage, the plume can be identified at different operational times. Aside from this, the clusters can be used to constrain the spatial parameter distribution in a numerical flow model. Zones of equal hydraulic properties are determined by clustering of reliable diffusivity values in the inverted tomograms. The hydraulic properties are estimated by calibration to the pressure signals.

The saturation changes are assumed relatively small during a multilevel CO$_2$ injection test. This allows us to use a single-phase simulator ("emulator") that solves the pressure equation decoupled from the saturation equation. This emulator is then applied to calibrate the hydraulic properties of inverted zonal structure by minimizing the residuals between the observed and calculated pressure responses. For the zonal calibration, we use the parameter estimator PEST [Doherty, 2010]. Ultimately, the calibrated specific storage will be used for calculating CO$_2$ saturations.

3.2.4 Forward modeling

The mass, energy conservation and moment equations of a two-phase two-component system are as follows [Lichtner, 1996; Helmig, 1997]:

$$\frac{\partial}{\partial t} \left[ f \sum_a S_a \rho_a X_i^a \right] + \nabla \cdot \left( u_a \rho_a X_i^a - f \tau S_a \rho_a D_a \nabla X_i^a \right) = Q_i$$

(3.6)

$$\frac{\partial}{\partial t} \left[ f \sum_a S_a \rho_a U_a + (1-f) \rho_c c_T \right] + \nabla \cdot \left( \sum_a u_a \rho_a H_a - \kappa \nabla T \right) = Q_h$$

(3.7)

$$u_a = -\frac{kk_a}{\mu_a} (\nabla P_a - \rho_a g z)$$

(3.8)

where $f$ is porosity, $\tau$ is tortuosity, $S_a$ is phase saturation, $\rho_a$ is molar density of phase $a$, $\rho_a$ is mass density of phase $a$, $D_a$ is phase diffusivity coefficient, and $X_i^a$ is molar fraction of component $i$ in phase $a$ ($i = w$, wetting phase; $i = n$, non-wetting phase). Parameter $u_a$ is the Darcy velocity, $Q_i$ and $Q_h$ are source/sink terms of mass and heat, respectively, $U_a$ is internal
energy, $H_a$ is enthalpy, $\rho_r$ is rock density, $\kappa$ is the thermal conductivity coefficient of the rock, $c_r$ is the heat capacity of the rock, $k_{ra}$ is the relative permeability of phase $a$, $k$ is the intrinsic permeability, and $T$ is the temperature.

The initial and side boundary conditions are defined as:

$$P_w = P_0 + \Delta P / \Delta z \quad (3.9)$$

$$S_{ra} = 0 \quad (3.10)$$

where $P_w$ is the brine pressure at different vertical positions, calculated by initial datum pressure ($P_0$) and pressure gradient ($\Delta P / \Delta z$).

Usually, a CO$_2$ injection in the saline reservoir is controlled by several hydraulic, thermal, mechanical and thermal processes. We simulate the early time, that is, the first months of CO$_2$ injection. At this stage, stratigraphic trapping dominates residual, solubility and mineral trapping mechanisms [IPCC, 2005]. This allows us to develop an approximate simulation procedure that neglects secondary trapping mechanisms, given the following assumptions:

- The two-phase flow in porous media (geo-reservoir) obeys the generalized Darcy’s law.
- Mass and phase transfer between the two phases are considered.
- Dissolution of CO$_2$ into the brine is in thermodynamic equilibrium.
- Gravity and capillary effects are taken into account.
- Diffusion of CO$_2$ in brine is described by Fick’s Law.
- The initial temperature gradient is set to zero given the small thickness of the aquifer.
- The interactions between fluids and rock are neglected. These include chemical reactions between fluids and rock, mechanical process, and heat transfer between fluids and rock.
- The pressure introduced from injection tests will not lead to brittle or ductile deformation of the rock.

For calculating the values of relative permeability and capillary pressure, the Brooks-Corey-Burdine model [Burdine, 1953; Brooks and Corey, 1964] is applied. Juanes et al. [2006] and Doughty et al. [2008] discussed in detail the residual trapping mechanism in CO$_2$ storage reservoirs. During the CO$_2$ injection phase (drainage) change in residual saturation is small and therefore the hysteretic behaviors in relative permeability-saturation and capillary pressure-saturation relationships are further neglected in our study. Additionally, variations in wettability of rock surfaces and in mineral composition are also considered negligible. In order to reduce model complexity, the residual saturation of two phases is assumed to be zero for the forward simulation:

$$k_{rw} = S_w^2 \quad (3.11)$$

$$k_{rw} = (1-S_w)^2 \left(1-S_w^{2+\lambda} \right) \quad (3.12)$$
where $k_w$ is the relative permeability of the wetting phase, and $k_n$ is the relative permeability of the non-wetting phase; $S_w$ is the saturation of wetting phase, and $\lambda$ denotes the pore distribution parameter. The capillary pressure ($P_c$) is calculated from:

$$P_c = P_d S_w^\frac{1}{r}$$

(3.13)

where $P_d$ is the entry pressure. The thermal conductivity is determined after [Somerton et al., 1974]:

$$\kappa = \kappa_{dry} + S_w (\kappa_{wet} - \kappa_{dry})$$

(3.14)

where $\kappa_{dry}$ and $\kappa_{wet}$ represent the dry and fully-saturated rock thermal conductivities.

### 3.2.5 Influence of CO$_2$ on fluid properties

For the stage prior to CO$_2$ injection, the pressure equation describing the single-phase system (fully brine saturated) can be written as:

$$(f c_w) \frac{\partial P_w}{\partial t} - \nabla \left[ \left( \frac{k}{\mu_w} \right)(\nabla P_w - \rho_w g) \right] - q_w = 0$$

(3.15)

where $q_w$ is the volumetric brine injection rate, and $c_w$ is the compressibility of brine:

$$c_w = \frac{1}{\rho_w} \frac{d\rho_w}{dP_w}$$

(3.16)

$\rho_w$ represents brine density, and $\mu_w$ is the dynamic viscosity of brine. Hydraulic conductivity ($K_w$) and specific storage ($S_{sw}$) are formulated as:

$$K_w = \rho_w g \left( \frac{k}{\mu_w} \right)$$

(3.17)

$$S_{sw} = \rho_w g (f c_w)$$

(3.18)

and the head change ($\Delta h$) is derived from the pressure change ($\Delta P_w$) by:

$$\Delta h = \frac{\Delta P_w}{\rho_w g}$$

(3.19)

Unlike for shallow aquifers, in deep aquifers, density, viscosity and compressibility are affected by pressure and temperature changes, and are therefore not constant.

The injected CO$_2$ will change the flow properties of the two-phase mixture due to its high compressibility. The flow properties of the mixture are therefore dependent on the saturation of
each phase. The flow properties of the mixed-phase can be derived from a mixed-phase global pressure equation [Chen and Ewing, 1997]:

\[
\left( f S_w c_w + f S_n c_n \right) \frac{\partial P}{\partial t} - \nabla \cdot \left[ \lambda_t k \left( \nabla P - \rho_n g \right) \right] - q_a = 0 \tag{3.20}
\]

where \( c_n \) is the compressibility of CO\(_2\) written as:

\[
c_n = \frac{1}{\rho_n} \frac{d \rho_n}{d P} \tag{3.21}
\]

\( \lambda_t \) is total mobility, determined by summation of individual phases’ mobility (\( \lambda_w, \lambda_n \)):

\[
\lambda_t = \lambda_w + \lambda_n = \frac{k_w}{\mu_w} + \frac{k_n}{\mu_n} \tag{3.22}
\]

and \( \rho_d \) is the dynamic density of mixed-phase, which is defined as:

\[
\rho_d = \frac{\lambda_w}{\lambda_t} \rho_w + \frac{\lambda_n}{\lambda_t} \rho_n \tag{3.23}
\]

In addition to the dynamic density, a static density of a mixed-phase [Wang and Beckermann, 1993] is defined to describe the static parameters:

\[
\rho_s = S_a \rho_w + S_d \rho_n \tag{3.24}
\]

In analogy to the single-phase flow equation, we can readily obtain the mixed-phase conductivity (\( K \)) and specific storage (\( S_s \)) through:

\[
K = \rho_d g \left( \lambda_t k \right) \tag{3.25}
\]

\[
S_s = \rho_s g \left( f S_w c_w + f S_n c_n \right) \tag{3.26}
\]

The diffusivity of the mixed-phase (\( D \)) is given by

\[
D = \frac{\lambda_t k \left( \rho_d \right)}{f S_w c_w + f S_n c_n \left( \rho_s \right)} \tag{3.27}
\]

According to Span and Wagner [1996], the compressibility of CO\(_2\) is one to two orders of magnitude larger than that of the brine, which leads to a significant increase of the storage term (\( f S_w c_w + f S_n c_n \)). Moreover, the compressibility of CO\(_2\) is nonuniform within the plume and depends on pressure and temperature. The mobility of CO\(_2\) is almost one magnitude larger than that of the brine, which may cause the conductive term (\( \lambda_t k \)) to vary non-monotonically due to the nonlinear changes of the relative permeability as a function of CO\(_2\) saturation. Thus the relationship between mixed-phase diffusivity to CO\(_2\) saturation depends on the total mobility of the
mixed-phase, on the compressibility of the two fluids, as well as on the ratio of dynamic mixed-phase density to static mixed-phase density. These aspects have to be accounted for in the following, when the mixed phase conditions are approximated by a fast standard groundwater flow model (emulator). Our hypothesis is that, through the conversion of a two-phase flow system involving two separate phases to an approximate mixed-phase flow system, transient pressure response curves can be delineated. In order to confirm this, we will compare pressure responses obtained from a full two-phase flow model and the mixed-phase emulator.

3.2.6 Case study

Virtual site

To explore the suitability of pressure tomography for the characterization of a CO$_2$ plume, a test case is developed which is loosely oriented at the Heletz site in Israel (Table 3.1). The site conditions adopted in our model are mainly taken from Rasmusson et al. [2014]. The proposed CO$_2$ storage formation is made of sandstone layers with a thickness of 15 m. It is located at a depth of 1600 m and sealed by a shale layer as the caprock. The formation pressure is 14.76 MPa with a formation temperature of 340.15 K at the aquifer bottom. CO$_2$ is in a supercritical state at these pressure and temperature conditions. The sandstone represents a deep saline aquifer and the salinity is assumed of 67 g/l. The rock density, intrinsic permeability and effective porosity of the sandstone are set to 2550 kg/m$^3$, $1 \times 10^{-13}$ m$^2$ and 0.2, respectively. These values are representative for deep CO$_2$ storage formations composed of sandstone [Hovorka et al., 2005]. Entry pressure of sandstone can cover a broad range [Krevor et al., 2011], here we choose a moderate value of 4000 Pa. The thermal properties of the rock, including heat capacity and thermal conductivity, are estimated according to [Robertson, 1988].

Numerical model

The sandstone formation is simulated in a two-dimensional numerical model as a homogeneous, isotropic aquifer with a thickness ($d$) of 15 m (Table 3.1). The lateral dimension ($x$ - direction) of the model is 580 m. A simplified two-dimensional model is appropriate to demonstrate the effectiveness of the inversion techniques as has been demonstrated by Yeh and Zhu [2007] or Xiang et al. [2009], but it neglects the three-dimensional nature of the flow pattern in a real injection system. For characterizing three-dimensional non-uniform plume spreading in the field, at least three wells would be required.

The initial brine pressure distribution in the model is hydrostatic and brine is the only liquid phase. The supercritical CO$_2$ is injected at the center of the model. The observation well is placed at a distance of 50 m, which is a common distance for well pairs at CO$_2$ pilot sites [Wiese et al., 2010]. The thickness of each layer is set to be 0.6 m and held constant throughout the simulation. Between the two wells, a horizontal discretization of 0.5 m is chosen, with a progressive refinement to 0.09 m towards the injection well. Outside of the target area, the grid size is increased.
exponentially with the largest grid size of 40 m at the distant boundaries. This yields a discretization of 287 and 25 grid cells in horizontal and vertical directions respectively, with 7175 grid cells in total.

Figure 3.1 Configuration of the source (injection well) and the receiver (observation well) in a cross-sectional numerical model.

The open source code PFLOTRAN [Hammond et al., 2014] is employed for the two-phase non-isothermal supercritical CO$_2$-brine flow simulation. The density of the brine/CO$_2$ mixture is calculated based on the state equation from Duan et al. [2008], and the viscosity of brine computation follows the IFC [1967]. The solubility equation for describing CO$_2$ in brine is taken from Duan and Sun [2003]. We apply the state equation developed by Span and Wagner [1996] to obtain the density and the method proposed by Fenghour et al. [1998] for the viscosity of supercritical CO$_2$.

Table 3.1 List of parameter values of numerical model based on a virtual site.

<table>
<thead>
<tr>
<th>parameter</th>
<th>value</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>intrinsic permeability ($k$)</td>
<td>$1 \times 10^{-13}$ m$^2$</td>
<td>Rasmusson et al. [2014]</td>
</tr>
<tr>
<td>porosity ($f$)</td>
<td>0.2</td>
<td>estimated</td>
</tr>
<tr>
<td>initial temperature ($T_0$)</td>
<td>340.15 K</td>
<td>Rasmusson et al. [2014]</td>
</tr>
<tr>
<td>salinity</td>
<td>67 g/l</td>
<td>estimated</td>
</tr>
<tr>
<td>gravity acceleration ($g$)</td>
<td>9.81 m$^2$/s</td>
<td></td>
</tr>
<tr>
<td>initial datum pressure ($P_0$)</td>
<td>14.76 MPa</td>
<td>Erlström et al. [2011]</td>
</tr>
<tr>
<td>diffusion coefficient of CO$_2$ in brine ($D_n$)</td>
<td>$3 \times 10^{-9}$ m$^2$/s</td>
<td>Wilke and Change [1995]</td>
</tr>
<tr>
<td>rock specific heat capacity ($c_r$)</td>
<td>930 J/kg K</td>
<td>Robertson [1988]</td>
</tr>
<tr>
<td>dry thermal conductivity ($\kappa_{dry}$)</td>
<td>3 W/m K</td>
<td>Robertson [1988]</td>
</tr>
<tr>
<td>wet Thermal conductivity ($\kappa_{wet}$)</td>
<td>4.5 W/m K</td>
<td>Robertson [1988]</td>
</tr>
<tr>
<td>entry pressure ($P_d$)</td>
<td>4000 Pa</td>
<td>estimated</td>
</tr>
<tr>
<td>pore-size distribution ($\lambda$)</td>
<td>0.76</td>
<td>Dana and Skoczylas [2002]</td>
</tr>
<tr>
<td>initial averaged brine density ($\rho_w$)</td>
<td>1052.5 kg/m$^3$</td>
<td>Duan et al. [2008]</td>
</tr>
<tr>
<td>initial averaged brine viscosity ($\mu_w$)</td>
<td>$4.2 \times 10^4$ Pa s</td>
<td>IFC [1967]</td>
</tr>
<tr>
<td>initial averaged compressibility of brine ($c_w$)</td>
<td>$3.8 \times 10^{-10}$ Pa$^{-1}$</td>
<td>calculated</td>
</tr>
<tr>
<td>initial hydraulic conductivity ($K_w$)</td>
<td>$2.46 \times 10^{-6}$ m/s</td>
<td>calculated</td>
</tr>
<tr>
<td>initial specific storage ($S_{sw}$)</td>
<td>$9.1 \times 10^{-7}$ 1/m</td>
<td>calculated</td>
</tr>
</tbody>
</table>
Crosswell testing at different stages

In the numerical model, several crosswell experiments are conducted in a tomographic configuration to derive space filling pressure response curves. These experiments are all configured equally, that is 5×5 screens at the source and the receiver, the length of each injection interval is 0.6 m. Injection of CO₂ occurs at different times at each of these intervals (Figure 3.2 and Table 3.2). The injection of CO₂ over discrete time intervals facilitates time-lapse analysis and is intended to reveal how broad the application window of pressure tomography is. We also distinguish three different durations (short, medium, long) of these stages, since the period of injection is an important factor in the injection model and its role for the inverted results unknown. The four studied stages are characterized as follows:

Stage 1: Crosswell multilevel brine injection tests. Prior to CO₂ injection, five multilevel brine injection tests at different depths are simulated and the pressure response curves recorded at the observation well. The formation is fully brine-saturated without any CO₂. The brine is injected from bottom to top sequentially at a rate of \( Q_w = 0.001 \text{ kg/s} \). Between two subsequent injections (interval of 2 h each), a recovery period of 15 h is simulated until the pressure has recovered to initial conditions. Thus, the total duration of this stage is 70 h (\( \Delta t \)). Pressure tomography at this pre-injection stage provides a reference, which can be compared with the pressure signals obtained during CO₂ injection at a later stage.

Stage 2: CO₂ injection. At this stage, CO₂ sequestration is initiated and conducted for a short period of \( \Delta t_z = 120 \text{ h} \), a medium period of 240 h or long period of 360 h. CO₂ is injected at a rate of \( Q_c = 0.02 \text{ kg/s} \) over the entire depth at the source well, which creates a two-phase system in the deep aquifer. No crosswell experiments are carried out at this stage.

Stage 3: Shut-in after CO₂ injection. This stage represents a recovery period after the previous injection stage. The pressure recovers for \( \Delta t_z = 240 \text{ h} \) to its initial hydrostatic state. Thus, hydrostatic pressure conditions are reinstated for the calibration procedure at a later stage.

Stage 4: Crosswell multilevel CO₂ injection tests. Analogous to the initial multilevel brine injection tests (stage 1), and with the same set-up, CO₂ is now used as a medium for crosswell testing (\( Q_c = 0.02 \text{ kg/s} \)). In contrast to the brine injection, shorter recovery periods are applied and thus the formation pressure does not fully recover at this stage. This is to avoid long relaxation phases and the associated transient effects induced by CO₂ migration (e.g. changes in plume shape) during the tomographic analysis. In our example, we set equal injection and recovery periods and adjusted them to the injected volume of CO₂ (stage 2), with 4 h (short), 5 h (medium) and 6 h (long) (Table 3.2). Pressure responses derived during this stage are used to reveal the modified in-situ flow properties and by this localize the CO₂ plume.
Table 3.2 Duration of the four stages (\(\Delta t_1\) to \(\Delta t_4\)) given short, medium and long CO\(_2\) storage periods.

<table>
<thead>
<tr>
<th>injection time</th>
<th>(\Delta t_1) (h)</th>
<th>(\Delta t_2) (h)</th>
<th>(\Delta t_3) (h)</th>
<th>(\Delta t_4) (h)</th>
<th>total duration (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>short</td>
<td>70</td>
<td>120</td>
<td>240</td>
<td>36</td>
<td>466</td>
</tr>
<tr>
<td>medium</td>
<td>70</td>
<td>240</td>
<td>240</td>
<td>45</td>
<td>595</td>
</tr>
<tr>
<td>long</td>
<td>70</td>
<td>360</td>
<td>240</td>
<td>54</td>
<td>724</td>
</tr>
</tbody>
</table>

Figure 3.2 Time sequence of the four sequentially modeled operational stages with crosswell testing before and after CO\(_2\) injection.

3.3 Results and discussion

3.3.1 Forward modeling results

Figure 3.3 shows the simulated CO\(_2\) plume development for the four stages assuming short, medium and long injection in stage 2. The snapshots of CO\(_2\) saturation distribution were taken at the end of the respective stage. The maximum CO\(_2\) saturation within the plume reaches around \(S_n\) = 0.6. As expected, the crosswell-testing by injecting brine does not show any effect in stage 1. As soon as CO\(_2\) is injected, a plume evolves and migrates laterally and upwards towards the top of the
aquifer due to buoyancy effects. Because of the high pressure at the source well during injection the lateral plume extension is most pronounced there. During recovery in stage 3, gravity forces become the main driver for CO₂ and the plume develops towards the top of the aquifer. At stage 4, little change in the plume shape and in saturation is observed. This is an important observation, since it demonstrates that during the final crosswell CO₂ injection tests the plume shape can be assumed to be constant.

Figure 3.3 CO₂ plume evolution according to its saturation during short, medium and long-term CO₂ injection, illustrated for stages 1 to 4 (see Figure 3.2).

In the following, we focus on the multilevel-injection experiments for tomographic analysis. The pressure responses recorded at the observation well at the initial and last stage are presented in Figure 3.4. Since these responses are similar for the five receiver levels, we only show the results for the central screen. The sequential multilevel brine injection yields uniform responses, which are all similar for the homogeneous aquifer (Figure 3.4a). The recovery period of 15 h is sufficient to recover to the initial system pressure distribution such that no response is influenced by the previous one.

Figure 3.4b depicts the recorded pressure during all following stages, ending with the final multilevel CO₂ injection. It shows that after the start of CO₂ injection, a maximum pressure of $P_\text{w} = 15.8$ MPa is reached within the aquifer. The same pressure is obtained for different CO₂ storage times (short, medium, long). During recovery (stage 3), the pressure returns to the initial hydrostatic level. During subsequent multilevel injection, the pressure recovery is incomplete and each new injection contributes to the recorded pressure. Due to the increased extension of CO₂ plume and high compressibility of CO₂, pressure responses are less well pronounced after longer injection and they decline during sequential multilevel testing.
3.3.2 Comparison of pressure responses from proxies and full models

Our strategy is to approximate the complex full two-phase model with a simplified single-phase proxy and by this facilitate fast iterative model calibration. The suitability of the proxy (implemented in MODFLOW) is assessed by comparing results with those of the two-phase simulations with PFLOTRAN. In the following, the simulated pressure changes at the observation well are transferred to head changes via Equation (3.19).

First, the head changes from the brine injection tests at stage 1 prior to CO$_2$ storage will be compared. In the proxy, the hydraulic parameters are fixed to the initial values of the full model as specified in (Table 3.1). Potential density and viscosity changes due to pressure and temperature variations are neglected. Comparison between both models in Figure 3.5 shows only small differences, and thus the proxy is considered a viable tool for the inversion.

The effect of viscosity, density and compressibility on hydraulic conductivity as defined in Equation (3.16) and (3.17) can explain the small discrepancies between both curves. The two-phase simulation showed minor changes on the order of $1.8 \times 10^{-9}$ Pa s in the viscosity of water. Therefore, a mean value of $4.2 \times 10^{-9}$ Pa s was chosen. In contrast, density and compressibility of the brine vary more strongly due to induced pressure and temperature changes. These also in turn affect hydraulic conductivity and specific storage. According to Equation (3.18), specific storage is not determined by brine density, but relates to the term $\frac{d\rho_w}{dP_w}$. Referring to the state equation of Duan et al. [2008], this term equals to $4 \times 10^{-7}$ kg/(m$^3$ Pa) for our simulation conditions. Thus, specific storage can be treated as a constant. In contrast, hydraulic conductivity is determined by brine density. The pressure increase leads to a higher brine density in the full two-phase model. Consequently, hydraulic conductivity increases as well, which could cause the larger head changes in the full model compared to the proxy.
 Analogous to the single-phase proxy of stage 1, a mixed phase proxy (emulator) for simulating the head changes from multilevel CO\textsubscript{2} injection (stage 4) is introduced. It is based on the following assumptions: The process is assumed isothermal. A single integrated value of CO\textsubscript{2} compressibility $c_n$ exists, which can minimize the discrepancies between the results from full two-phase model and mixed-phase proxy. The two phases in the aquifer, brine and CO\textsubscript{2}, are slightly compressible. Therefore, density values of brine and CO\textsubscript{2} ($\rho_w$ and $\rho_n$) are set constant according to the values at the beginning of the injection. In the emulator, compressibility of brine and CO\textsubscript{2} ($c_w$ and $c_n$) does not change, calculated from the state equation of two fluids. Capillary pressure ($P_c$) is neglected, so that the liquid pressure ($P_w$) equals the gas pressure ($P_n$) and to the global pressure ($P$). Furthermore, saturation changes are ignored, so that pressure changes are calculated solely from the global pressure equation (Equation (3.20)). Finally, viscosity of the two phases ($\mu_w$ and $\mu_n$) is set constant.

The mixed-phase flow properties derived from Equation (3.25) and Equation (3.26) utilized as the input parameters for the emulator. At stage 4, the state equation of CO\textsubscript{2} reveals that the density of CO\textsubscript{2} as a function of the pressure ($\frac{d\rho_n}{dP_n}$) changes from $4\times10^{-5}$ to $5.3\times10^{-5}$ kg/(m\textsuperscript{3}·Pa) for a pressure range from 14.01 to 15.51 MPa at 340.15 K, respectively. Through the transformation from Equation (3.25) to Equation (3.27), the gas saturation values at the start of the tests were transferred to mixed-phase conductivity, specific storage and diffusivity.

The transferred mixed-phase flow parameters vs. CO\textsubscript{2} saturation are plotted in Figure 3.6. Figure 3.6a shows that the maximum mobility of CO\textsubscript{2}, $\lambda_n$, is one order of magnitude higher than that of water, which is due to the difference in viscosity of these two phases. The mixed-phase hydraulic conductivity $K$ spanning a range from $9.8\times10^{-7}$ to $1.3\times10^{-5}$ m/s (Figure 3.6d) is controlled by the total mobility $\lambda_f$ and dynamic mixed-phase density $\rho_d$ (Figure 3.6b, c). The
resulting conductivity does not change monotonously, and shows minor variability within the saturation ($S_n$) range of 0-0.5. When CO$_2$ saturation rises from 0 to 0.22, the conductivity declines due to a decrease in brine mobility and a dynamic mixed-phase density. When CO$_2$ saturation is larger than 0.22, the conductivity increases along with the growing CO$_2$ saturation.

Specific storage ($S_r$) rises with CO$_2$ saturation, since the introduced compressibility is much larger than the initial value for the brine. The range of specific storage varies from $9.1 \times 10^{-7}$ to $9.1 \times 10^{-5}$ l/m (Figure 3.6e, blue line) when $\frac{dp_n}{dP_n} = 4 \times 10^{-5}$ kg/(m$^3$·Pa) and it changes from $9.1 \times 10^{-7}$ to $1.2 \times 10^{-4}$ l/m (Figure 3.6e, red line) when $\frac{dp_n}{dP_n} = 5.3 \times 10^{-5}$ kg/(m$^3$·Pa). Being the ratio of conductivity to specific storage, the curve delineating the mixed-phase diffusivity in Figure 3.6f does not show a monotonous behavior either. Overall, the diffusivity is lower than that of the CO$_2$-free aquifer, which is 2.7 m$^2$/s.
Figure 3.6 Flow properties vs. CO₂ saturation: a) mobility of two phases; b) mobility fraction of two phases; c) mixed-phase density; d) mixed-phase conductivity; e) mixed-phase specific storage; f) mixed-phase diffusivity (blue lines: $\frac{d\rho_n}{dP_n} = 4 \times 10^{-5}$ kg/(m$^3$ Pa); red lines: $\frac{d\rho_n}{dP_n} = 5.3 \times 10^{-5}$ kg/(m$^3$ Pa)).

The key point when designing the emulator is to determine a robust integrated value of $\frac{d\rho_n}{dP_n}$, which shows the same pressure responses as the full model. In the pressure tomographic approach, especially the early parts of these responses (early time diagnostics) are examined. Therefore, we
tested in a plausible range of $\frac{dp}{dP_n}$ from $4 \times 10^{-5}$ kg/(m$^3$ Pa) to $5.3 \times 10^{-5}$ kg/(m$^3$ Pa) and compared the stage 4 pressure response curves for variable injection times and injected CO$_2$ volumes (Figure 3.7). For the short, moderate and long scenarios, values of $4.9 \times 10^{-5}$ kg/(m$^3$ Pa), $5.1 \times 10^{-5}$ kg/(m$^3$ Pa) and $5.2 \times 10^{-5}$ kg/(m$^3$ Pa) respectively, were considered to be most suitable (Figure 3.7, black dotted line).

![Figure 3.7 Head comparison of full model and emulator for different periods of CO$_2$ injection during stage 4: a) short, b) medium, c) long time injection.](image)

In the following, the results from modelling the evolution of mixed-phase conductivity, specific storage and diffusivity within the plume is shown for the four stages considered. This is of particular interest, because the emulator is only capable of simulating a pseudo-mixed single phase. In Figure 3.8-Figure 3.10, the forward modelling data of the full model are visualized, which are derived from the CO$_2$ saturations at different times (see Figure 3.3). In this simulation, the variability of mixed-phase conductivity (Figure 3.8) is within $1 \times 10^{-6}$ and $4.5 \times 10^{-6}$ m/s and hence, is quite small. The largest values occur at the top of the plume where the saturation of CO$_2$ is at a maximum (Figure 3.3). Consistent with the transformation function (Figure 3.6), the smallest values appear where the CO$_2$ saturation is below 0.22.

Similar to the hydraulic conductivity, the mixed-phase specific storage is correlated with the CO$_2$ saturation (Figure 3.9) but computed values span over two orders of magnitude from $1 \times 10^{-6}$
to $1 \times 10^{-4} \text{ l/m}$.

The diffusivity (Figure 3.10) can be determined from the subsequent travel-time based inversion. As inferred from Figure 3.6f, in the plume, the mixed-phase diffusivity declines in contrast to the ambient aquifer ($D = 2.7 \text{ m}^2/\text{s}$). However, towards the top where CO$_2$ saturation exceeds 0.22, the diffusivity slightly increases again, following its relationship with the conductivity.

Figure 3.8 Mixed-phase hydraulic conductivity evolution during short, medium and long-term CO$_2$ injection, illustrated for stages 1 to 4 (see Figure 3.2).

Figure 3.9 Mixed-phase specific storage evolution during short, medium and long-term CO$_2$ injection, illustrated for stages 1 to 4 (see Figure 3.2)
3.3.3 Eikonal based inversion

Early time diagnostics

With the full numerical model of the virtual site, we cannot only predict the evolution of the CO₂ plume, but also the pressure responses from the tomographic tests with brine (stage 1) and CO₂ injection (stage 4). In the following we will focus on the travel time diagnostics used for the eikonal inversion. Afterwards, the emulator is used for the calibration of the response curves. In the discussion of the results, we name the five source intervals at the central injection well S1-S5 from bottom to top, and the five receiver intervals at the observation well R1-R5.

In a first step, the head changes and their first time-derivatives are derived from the multilevel brine injection tests during stage 1 (Figure 3.11e). The results are identical in each source-receiver combined pattern and thus only the result of one single injection test is presented. As shown in Figure 3.11a, the head increases by 1.62 m during the 2 hours of injection but does not reach steady state. The peak travel time appears after 411 s, with some oscillation in the derivative. The values of the more robust early time diagnostics, $t_{-10\%}$, $t_{-20\%}$, $t_{-30\%}$, $t_{-40\%}$ and $t_{-50\%}$, are 21 s, 45 s, 72 s, 90 s and 113 s, respectively.

In a second step, travel times from stage 4 were computed by the first time-derivative of the head changes. The trends in the head changes and associated first time-derivatives are similar at all observation points (R1-R5). Figure 3.11b-d illustrates the head changes from the multilevel CO₂ injection tests at the bottom of the observation screen (R1), and Figure 3.11f-h shows the corresponding first time-derivatives of the head changes. From S1 to S5, the decreasing head
changes reflect the growing influence of CO$_2$ accumulating at the top of the aquifer. For the inversion, considering the initially steep slope of the derivative curves, the early travel time diagnostics seem favorable. These diagnostics are smallest at the bottom observation screen (R1) and increase towards R5 (not shown). A comparison of the head change and first derivative curves after the three different CO$_2$ injection periods reveals that the developed plume also lowers the head changes and their first derivative values for a given source-receiver configuration. Accordingly, the early travel time diagnostics show a delay when the CO$_2$ injection period increases.

With the insight from Figure 3.11, we focus on the early time diagnostics, which are more informative than peak times. We selected a moderate value of $t\,\cdot\,20\%$ for the inversion process. In fact, the other early time diagnostics can offer similar structural information of the inverted tomograms since they equally show a consistent change for each source-receiver pattern. The main difference from using different diagnostics would be seen in the absolute values of inverted diffusivity and thus it is not further discussed here.
Figure 3.11 Head changes and first derivatives of five injection period in R1 before CO\textsubscript{2} injection, and during short, medium and long-term CO\textsubscript{2} injection.
Diffusivity tomograms

With the recorded travel time diagnostics (1-20%), eikonal based inversion can be employed to obtain a diffusivity tomogram. We applied the Geotom code [Jackson and Tweeton, 1996] for the inversion by implementing the SIRT [Gilbert, 1972]. A staggered grid technique [Vesnaver and Böhm, 2000] was applied to increase the nominal resolution of the tomograms. Without this technique only a coarse tomogram with a grid of 5 columns and 4 rows would be feasible based on the number of source and receiver intervals. The number of unknowns is adapted to the number of available source-receiver configurations, and thus the inversion problem can be approximately considered as an “Even-Determined Problem” [Brauchler et al., 2003, 2007]. In the staggered grid approach, instead of one grid, multiple shifted grids were applied for separate inversion, which together can be arithmetically averaged into a high-resolution tomogram. The shifting step sizes in \( x \)-direction and in \( z \)-direction were 2 m and 1.875 m respectively, and a final diffusivity tomogram with 25 columns and 8 rows was obtained from 10 individual inversions. The inversion was stopped once the error between calculated and measured travel times was convergent. This procedure was applied to the multilevel brine-injection experiment in stage 1, and to the multilevel \( \text{CO}_2 \) injection (stage 4).

As a reference, the diffusivity tomogram of stage 1 was constructed. Figure 3.12a shows that the inverted diffusivity \( (D) \) varies only within a small range from 2.84 to 2.94 m\(^2\)/s. This means the homogeneity of the aquifer is nicely reflected by the tomogram. As it is common for such eikonal inversion, absolute values are not well reproduced [Jiménez et al., 2013]. Here the mean value of 2.9 m\(^2\)/s is higher than the “true” value for the aquifer (2.7 m\(^2\)/s). Also, the null space energy was calculated, and the obtained reliability map shows that the values at the upper boundary are the least reliable. The unreliable values and the slight variability of \( D \) are caused by the non-uniform trajectory density. In such crosswell tests, trajectory density is larger in the model center than at the top and bottom [Hu et al., 2011] and this is reflected by the minor variability of \( D \) in the central area of the tomogram.

The diffusivity tomograms of stage 4 after short, medium and long injection periods are depicted in Figure 3.12b-d. A common feature in all tomograms is the low diffusivity area at the injection well, which is consistent with the \( \text{CO}_2 \) plume zone simulated by the full two-phase model (PFLOTRAN simulation). It can be seen that the tomograms after different injection periods show the evolution of the \( \text{CO}_2 \) plume. Again, absolute diffusivity values deviate substantially from the simulated “truth”. The inverted range of \( D \) values is about half of the true range. Particularly in the \( \text{CO}_2 \)-free ambient aquifer, which appears in red color in the true model, \( D \) is strongly underestimated. In this part of the aquifer, the arithmetic mean values of \( D \) are 0.97, 0.55, and 0.14 m\(^2\)/s for short, medium and long injection periods, respectively, which are all below the “true” value \( (D = 2.7 \text{ m}^2/\text{s}) \). These results demonstrate that the derived tomograms are suitable to capture structural information and thus to identify the shape of the plume. However, the inverted values of \( D \) are not appropriate for computing the \( \text{CO}_2 \) saturation. This requires full pressure response calibration, which is pursued below.
The computed null space energy maps (Figure 3.12, third column) show larger values (greater than 0.5) at the upper right corner and the bottom of the cross section for all injection periods. However, these unreliable pixels or zones are outside the plume area, and hence a localization of the plume was possible.

![Figure 3.12 True and inverted diffusivity tomograms with reliability maps after stage 4 in different times: a) pre-injection; b) short injection; c) medium injection; d) long injection.](image)

**Clustering and full signal calibration**

As a standard data partitioning method, k-means clustering is applied to cluster the tomograms obtained after CO$_2$ injection, and by this, to distinguish between plume and ambient aquifer pixels. k-means clustering classifies the data according to their distance to the nearest centroid. The two centroids of each cluster were determined by histogram analysis, and kept constant for all injection periods (\( D = 0.055 \text{ m}^2/\text{s}, 0.13 \text{ m}^2/\text{s} \)). In order to account for data reliability, the cluster analysis was only based on pixels with null space energy smaller than 0.5. The positions with greater values were filled up by nearest neighbor interpolation of the adjacent cluster, which was here always the one representing the CO$_2$-free ambient aquifer.

At this point, we set up one homogeneous aquifer model for stage 1, representing the pre-injection conditions, and for each injection period a clustered model with zonal structure is available. This facilitates a two-step full pressure signal calibration (using PEST), assuming that the zones represent homogeneous areas. From calibrating the pre-injection model, a hydraulic conductivity of \( 2.24 \times 10^{-6} \text{ m/s} \) and specific storage of \( 8.6 \times 10^{-7} \text{ l/m} \) are determined. These values
are also valid for the ambient aquifer after CO$_2$ injection, and thus assigned to this zone in the stage 4 models.

Recalling the relationships between flow properties and CO$_2$ saturation (Figure 3.6), calibrating the mixed-phase specific storage of the emulator is best suited for estimating CO$_2$ saturation due to the monotonous relationship. Moreover, the introduced CO$_2$ phase disturbs the mixed-phase specific storage much more than mixed-phase conductivity. As an additional argument, Wu et al. [2005] pointed out that the early head changes are more sensitive to specific storage compared to hydraulic conductivity. Thus, only the specific storage of the plume was calibrated, assuming that the hydraulic conductivity can be set to be uniform for the entire model. In this case, relative permeability curves have no impact on the inversion.

The first three stress periods of the head changes were utilized for the calibration since they display minor differences between the results from the full model and the proxy. The calibrated specific storage values for the plumes of three injection periods are shown in Table 3.3. The calibrated values indicate that the specific storage of the plume is almost half a magnitude larger than that of the original aquifer. Referring to the changes of $\frac{dp_n}{dn}$, the CO$_2$ saturation inferred from specific storage can vary from 0.18 to 0.25, 0.3 to 0.43, and 0.27 to 0.39 for short, medium and long injection periods, respectively. If we take the mean value of $\frac{dp_n}{dn}$ to calculate the CO$_2$ saturation, then the final values of CO$_2$ saturation are 0.21, 0.35, and 0.32. These values agree very well with the arithmetic mean of the “true” saturation values within the clustered structure (0.22, 0.33 and 0.32) (Table 3.3). In a last step, we visualize the plume shapes and compare the calculated CO$_2$ saturation within the plume with the “true” saturation distribution (Figure 3.13). The figure shows that the plume, especially the edge of the plume, can be localized by the inversion and clustering approaches.

Table 3.3 Calibration result and transferred CO$_2$ saturation.

<table>
<thead>
<tr>
<th></th>
<th>calibrated $S_n$ (1/m)</th>
<th>calibrated $K$ (m/s)</th>
<th>$S_n$ (in range) (-)</th>
<th>$S_n$ (calculated) (-)</th>
<th>$S_n$ (true) (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-injection</td>
<td>$8.6 \times 10^{-7}$</td>
<td>$2.24 \times 10^{-6}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>short-injection</td>
<td>$4.02 \times 10^{-5}$</td>
<td>-</td>
<td>0.18-0.25</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>medium-injection</td>
<td>$6.27 \times 10^{-5}$</td>
<td>-</td>
<td>0.3-0.43</td>
<td>0.35</td>
<td>0.33</td>
</tr>
<tr>
<td>long-injection</td>
<td>$5.83 \times 10^{-5}$</td>
<td>-</td>
<td>0.27-0.39</td>
<td>0.32</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Figure 3.13. Calculated saturation compared with “true” values after stage 4 in different times: a) short injection; b) medium injection; c) long injection. The grey line in “true” saturation graphs delineates the inverted plume shape.

3.4 Conclusions

Monitoring techniques are essential for improving the safety and optimizing the operation of CO₂ storage sites. A novel approach, the time-lapse pressure tomography inversion, is developed and successfully demonstrated for monitoring the CO₂ plume evolution in underground storage reservoirs. This method is based on principles of hydraulic tomography, which is most commonly used for the characterization of shallow aquifers. The latter utilizes hydraulic pressure signals from multiple sources and receivers for a spatial reconstruction of heterogeneity. Since a CO₂ plume directly influences the hydraulic properties and induces an apparent heterogeneity, a concept similar to hydraulic tomography can be adopted for reconstructing the shape of the plume. By inspecting the transient behavior of the plume through repeated pressure applications, a time-lapse pressure tomography is obtained.

We have shown the feasibility of the proposed investigation method by simulating a realistic deep CO₂ storage site in a homogeneous saline aquifer. Pressure signals are simulated by both brine and CO₂ injection and recorded between two wells at different levels. The signals are inverted by a fast but approximate travel time based tomographic procedure, which offers a first insight into the plume shape and how it evolves over time. We demonstrated that the information from the pressure signals could be used for full calibration of a process-based numerical model. This provides the opportunity for resolving not only the shape of the plume, but also for directly determining the spatial distribution of the non-wetting phase saturation. In the case study presented here, the estimated values of saturation are consistent with those computed explicitly in simulated CO₂ injection scenarios. This good agreement holds even though a simplified single-phase model
or “emulator” was used for the calibration. The approximation of multi-phase processes with a simplified single-phase numerical model appears sufficient to capture the conditions relevant for the tomographic inversion. By this, a computationally efficient full signal based inversion is possible without requiring a complex multi-phase simulation.

One fundamental advantage of the new method is that it requires only a doublet well configuration, i.e. an injection and an observation well, and, therefore, can be applied to a large range of geological reservoirs. Another advantage is that the method is fast (from a few minutes to hours) and rather inexpensive from a computational point of view, as it requires the injection of only small fluid (water, brine, CO$_2$) volumes. Note that no fluids are extracted and have to be disposed of. Nevertheless, for multilevel injection sequences, packers must be installed which may entail additional costs. In principle, the application window of the proposed pressure tomography is not restricted to homogeneous and isotropic aquifers only. This will be further explored in future work focusing on the transient changes in the reservoir in the time-lapse framework. A crucial point for feasibility in practice will be that the pressure signals reach good spatial resolution, and that a sufficiently high signal-to-noise ratio is achieved. This will determine the application scale, which is expected to be smaller than, for example, that of seismic tomography, but with better resolution at the small scale. Ideally, for real-time monitoring of a CO$_2$ plume, the time-lapse pressure tomography approach is combined with complementary tracer testing or geophysical techniques.
4 Detection of carbon dioxide leakage during injection in deep saline formations by pressure tomography


Abstract

CO₂ injected into storage formations may escape to the overlying permeable layers. Mixed-phase diffusivity, namely the ratio of hydraulic conductivity and specific storage of the phase mixture, declines with increasing CO₂ saturation. Thus it can be an indicator of CO₂ leakage. In this study, we perform interference brine or CO₂ injection tests in a synthetic model, including a storage reservoir, an above aquifer, and a caprock. Pressure transients derived from an observation well are utilized for a travel-time based inversion technique. Variations of diffusivity are resolved by inverting early travel time diagnostics, providing an insight of plume development. Results demonstrate that the evolution of CO₂ leakage in the above aquifer can be inferred by interpreting and comparing the pressure observations, travel times and diffusivity tomograms from different times. The extent of the plume in reservoir and upper aquifer are reconstructed by clustering the time-lapse data sets of the mixed-phase diffusivity, as the diffusivity cannot be exactly reproduced by the inversion. Furthermore, this approach can be used to address different leaky cases, especially for leakage occurring during the injection.
4.1 Introduction

Deep saline aquifers are deemed most suitable for geological CO$_2$ storage [Bachu and Adams, 2003]. However, injection of CO$_2$ in deep reservoirs is not free of risks. An evolving plume might migrate to overlying strata along unknown seal imperfections, faults or fractures in caprocks, or pre-existing abandoned wells [Lemieux, 2011]. Also, the high pressures during injection can create new pathways for escape. When CO$_2$ emanating from storage formations reaches shallow aquifers, it contaminates the groundwater environment [Apps et al., 2010]. Therefore, identification of possible CO$_2$ leakage is a crucial task during sequestration. Monitoring needs to localize weak zones in a caprock and migration pathways to be able to plan countermeasures.

Pressure-based techniques are appealing for early stage leakage detection, since pressure perturbations propagate much faster than a CO$_2$ plume itself. In previous work, mainly synthetic models are utilized to assess the impact on pressure buildup due to brine/CO$_2$ leakage [e.g., Birkholzer et al., 2009] and detectability of brine/CO$_2$ leakage through leaky wells or permeable caprock based on pressure transients in the storage reservoir and its overlying aquifer [e.g., Chabora and Benson, 2009; Nogues et al., 2011; Azzolina et al., 2013; Wang and Small, 2014]. Forward modeling can be done by fully coupled numerical simulators, or by simplified analytical or semi-analytical solutions [e.g., Nordbotten et al., 2004; 2008; Celia and Nordbotten, 2009; Zhou et al., 2009; Cihan et al., 2011]. By comparison to simulation results, pressure observations at a field site can be interpreted [e.g., Park et al., 2012].

Alternatively, leakage locations, rates and permeability of leaky wells are inferred by inverting pressures and calibrating forward models. Most studies in this field utilize the pressure anomalies in the storage formation or the above aquifer for the inversion, which requires knowledge of pressure responses in a no-leakage case [e.g., Sun and Nicot, 2012; Jung et al., 2013; Lee et al., 2015]. The leakage amount can be estimated by historical matching of pressure data to analytical solutions [e.g., Meckel et al., 2013; Hosseini, 2014]. However, such simplified approaches may not be satisfactory for complex geometries. Also, none of the available pressure-based inversion techniques can provide spatial information of an evolved secondary CO$_2$ plume in the above-zone, which is fundamental for the planning of remediation strategies. To date, only seismic reflection tomography is utilized for mapping a leaky plume [e.g., Arts et al., 2005; Chadwick et al., 2014]. However, its spatial resolution and coverage is limited.

Hydraulic, or more generally, pressure tomography has been developed during the last decades as an approach to reconstruct the subsurface heterogeneity. Its applicability in single-phase flow field has been demonstrated successfully by numerical models [e.g., Yeh and Liu, 2000], lab experiments [e.g., Illman et al., 2010], and field-scale tests [e.g., Brauchler et al., 2013]. [Hu et al., 2015] proposed a pressure-based tomographical approach to track the evolution of a CO$_2$ plume in a homogeneous storage reservoir. They characterized the CO$_2$-induced heterogeneity of hydraulic properties and delineated the plume shape by the comparison of diffusivity tomograms in a time-lapse strategy. The objective of this paper is to introduce this pressure-based tomographical approach to characterize potential CO$_2$ leakages in a multilayer system with a storage formation,
an above aquifer, and an intervening caprock. Repetition of pressure tomography prior to, and after CO₂ injection, enables continuous monitoring of the CO₂ plume and potential leakage between the well pair. Contrary to other pressure-based inversion methods, this approach requires merely short-duration pressure transients and it can be applied for multiple leakage cases occurred after CO₂ injection. The multilayer system is initially saturated with brine, and brine displacement is not considered in this work. In the following sections, a novel design of interference fluid tests applied to the storage formation and the above aquifer is proposed. The observed pressure data at the monitoring well is utilized for inverting the spatial distribution of diffusivity in tomograms and evolving plumes in the two permeable layers are determined by clustering the tomograms.

4.2 Methodology

4.2.1 Problem set-up

A synthetic case is set up in a numerical model. We utilize an open source code, PFLOTRAN [e.g., Hammond et al., 2014], for the forward simulation of the two-phase, two-component, non-isothermal transport processes. For the forward model, the CO₂ storage formation, the overlying monitoring zone, and the caprock are considered as an entirely confined system (Figure 4.1). Following the suggestions by [Sun et al., 2015], the storage formation is named the “reservoir” and the monitoring zone above the caprock is the “aquifer”. The bottom of the simulated example is 1600 m below the ground surface. Reservoir and aquifer are composed of sandstone, whereas the material of the caprock is shale or siltstone. Each layer is assumed homogeneous and isotropic, with an equivalent thickness of 15 m. Pressure at the bottom is assigned a value of 14.7 MPa, with a constant vertical pressure gradient of 0.01 MPa/m. No-flow conditions are assigned to the top and bottom of the system, while the grids at the distant boundaries are set to a constant hydrostatic pressure. The initial system is stagnant and fully saturated with brine, with a constant temperature of 65°C. Under these conditions, CO₂ is in supercritical state.

The entire system is simulated in a simplified two-dimensional cross-sectional model with non-uniform rectangular grids. A three-dimensional and non-uniform plume could be identified by increasing the number of the wellbores and performing the injection tests in different wellbores. The lateral extent (x-direction) of the model is 560 m. An injection well with a radius of 0.09 m is placed at the center of the model domain, completely penetrating the three formations. Two observation wells are located 50 and 100 m away from the injection well to monitor the pressure responses, which are comparable to the distance of well pairs in several pilot sites [e.g., Wiese et al., 2010; Niemi et al., 2012]. The model comprises 45 layers, each with a constant thickness of 1 m. The lateral discretization is set to 1 m between the wells, with a progressive refinement to 0.09 m toward the injection well. Outside of the area constrained by the well pair, the horizontal grid resolution is increased exponentially. The largest grid size is at the distant boundaries with a value of 30 m. Ultimately, a discretization of $322 \times 45 = 14,490$ grid cells in total is created.
4.2.2 Leaky cases and model parameters

On the premise that the structure of the three-layer system is characterized by previous site investigation before the CO₂ injection campaign starts, we distinguish three different cases: “case N”, “case F” and “case D”. The presumptive model parameters of three cases are summarized in Table 4.1. Case N refers to a no-leakage case with a caprock of permeability $k = 1 \times 10^{-19}$ m². Simulation results of this case serve as a reference for comparison with the two other leaky cases. Case F implies a case of unanticipated fracture leakage in the caprock, which is caused by any potential excitation, such as overpressure by CO₂ injection, or a seismic event. We suppose the creation of a leaky pathway occurred at 350 h after the depth-integrated CO₂ injection. New pathways evolve, reflected in a discontinuity in the caprock. For convenience, a rectangle gap is assigned as a leaky path (Figure 4.1) with a moderate permeability value of $1 \times 10^{-11}$ m². Case D simulates diffusive leakage. In this case, CO₂ may migrate through the seal unit, of which the sealing capacity prior to CO₂ injection is assumed to be overestimated. Effective permeability of the seal unit is set at a relatively high value of $1 \times 10^{-14}$ m². This case can be used for CO₂ leakage through small intra-formation shale layers as a more realistic condition.

<table>
<thead>
<tr>
<th>Table 4.1 Summary of model parameter values for three cases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter</td>
</tr>
<tr>
<td>intrinsic permeability</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>entry pressure</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>effective porosity</td>
</tr>
<tr>
<td>pore size distribution</td>
</tr>
<tr>
<td>specific heat capacity</td>
</tr>
<tr>
<td>dry thermal conductivity</td>
</tr>
<tr>
<td>wet thermal conductivity</td>
</tr>
<tr>
<td>rock density</td>
</tr>
<tr>
<td>initial temperature</td>
</tr>
<tr>
<td>salinity</td>
</tr>
<tr>
<td>gravitational acceleration</td>
</tr>
<tr>
<td>residual saturation of wetting phase</td>
</tr>
<tr>
<td>residual saturation of non-wetting phase</td>
</tr>
<tr>
<td>initial datum pressure</td>
</tr>
<tr>
<td>diffusion coefficient of dissolved CO₂ in brine</td>
</tr>
</tbody>
</table>

For all cases, the intrinsic permeability of the reservoir and aquifer equals $1 \times 10^{-13}$ m². Rock density and effective porosity of the two permeable layers and the leaky gap are 2600 kg/m³ and 0.25 respectively, and they are set to 2700 kg/m³ and 0.05 for the caprock. The heat capacity and
the dry and wet rock thermal conductivity of the permeable units (reservoir, aquifer and leaky gap) are 860 J/kg, 2.6 W/mK, 4.6 W/mK, and 830 J/kg, 1 W/mK, 1.8 W/mK for the caprock. A value of 67 g/l specifies the initial salinity of the reservoir. The Brooks-Corey-Burdine model [Brooks and Corey, 1964; Burdine, 1953] is employed for calculating the relative permeability and capillary pressure according to the CO₂ saturation. For the scope of this study, residual saturation of the two phases is neglected to minimize model complexity. The pore size distribution is assigned values of 1.5 for the permeable units and 0.5 for the caprock respectively. The entry pressure of the low-permeability caprock is at a high value of $3.9 \times 10^6$ Pa, and equals $1.2 \times 10^4$ Pa for the high-permeability caprock. The entry pressure of the two permeable formations and the leaky gap is $8 \times 10^3$ Pa and $1 \times 10^3$ Pa, respectively.

Figure 4.1 Schematic sketch of the conceptual model. Distance between two wells is 50 m, and the thickness of each layer is 15 m. AS1-AS3 and RS1-RS3 imply the installed sources in the aquifer and reservoir respectively, while AR1-AR3 and RR1-RR3 indicate the corresponding receivers in the monitoring well. Red dashed line sketches the area of fracture leakage (case F).

4.2.3 Fluid interference tests

Crosswell fluid injection tests are conducted in a tomographic configuration. In the reservoir and aquifer, three intervals (sources) with a screen of 1 m are set up for multilevel brine/CO₂ injection tests (Figure 4.1). The observation points (receivers) are placed at the same depths and record the transient pressure signals continuously. The distance between two sources or receivers is 5 m. For better distinction, the sources are named as RS1-RS3 in the reservoir and AS1-AS3 in the aquifer. Correspondingly, the receivers are named as RR1-RR3 and AR1-AR3. In analogy to the work of [Hu et al., 2015], the fluid injection tests are designed in four stages as following:

Stage 1: Crosswell multilevel brine injection tests. At this stage, we conduct six brine injection tests in the initially CO₂-free formations. First, brine is injected from AS1 to AS3 sequentially, and
then from RS1 to RS3 with a constant injection rate of 0.01 kg/s. The injection sequence is not influential to the inversion. Duration of each injection test is 0.2 h, following a recovery period of 10 h between two subsequent injections. Pressure returns to the initial conditions after each recovery test. Note that we do not consider different salinities in the reservoir and aquifer. This could be simulated but is not crucial for the presented inversion procedure.

Stage 2: CO$_2$ injection. After stage 1, the CO$_2$ sequestration phase is simulated. Here we assume the injected CO$_2$ is in supercritical state, and it is injected at a constant rate of 0.012 kg/s over the entire thickness of the reservoir to create a two-phase system. We distinguish two different injection durations referred to as “short injection” and “long injection”, lasting 250 h and 500 h, respectively.

Stage 3: Shut-in period. At this stage, the injection is suspended and pressure falls off to its original status. The shut-in duration for short and long injection is 120 h and 250 h, respectively. It is possible to calibrate saturations by adding this stage, and in reality, it can be considered as a period for preparing the following hydraulic or other geophysical tests.

Stage 4: Crosswell multilevel brine/CO$_2$ injection tests. Brine is injected firstly in the aquifer from AS1 to AS3. Subsequently, CO$_2$ is injected in the reservoir from RS1 to RS3. Pressures derived at this stage are used for inverting the plumes in both reservoir and aquifer. All brine or CO$_2$ injection tests are executed at a constant rate of 0.01 kg/s. The duration of each injection and recovery test is defined based on two principles:

- The injection rate has to be sufficiently large and the injection period has to be long enough to generate a pressure response, with a sufficient signal-to-noise ratio. However, a premise of this is that the injection should not induce large disturbance in the reservoir or aquifer.

- Travel times can be delayed by nonuniform pressure buildup or heterogeneity of the formation. In the reservoir, travel times are mainly determined by the CO$_2$-induced heterogeneities. On the contrary, the travel times in the aquifer are influenced by both nonuniform pressures and potential leaks. For resolving small leaks with lower CO$_2$ saturations, the recovery period in the aquifer should be longer than in the reservoir. However, complete recovery of the pressure is almost impossible since full recovery needs much more time. Hence, here the recovery in the aquifer is terminated as the pressure reaches a quasi-steady state.

Hence, different injection and recovery durations are assigned at this stage, and the injection is halted as soon as the travel times (i.e. the time relates to the maximum pressure changes) can be noticeably identified. The duration of the injections depends also on the changes of the initial system, i.e., CO$_2$ enters and expands in the reservoir and aquifer. During short injection, for the three cases, durations of injection tests in the aquifer and reservoir are set to 0.2 and 3.5 h, and recovery tests last 8 and 3.5 h. During long injection, injection tests of brine sustains 0.2, 1 and 1 h for each of the cases N, D, and F, respectively. Injection and recovery tests of CO$_2$ in the reservoir are assigned equivalent durations, such as 5, 4 and 4.5 h for cases N, D and F.
4.2.4 Inversion in two-phase system

Pressure responses derived from stages 1 and 4 are used for the inversion. We utilize a single-phase proxy to apply a travel-time based inversion method [Vasco et al., 2000; Brauchler et al., 2003]. In a CO₂-brine system, considering CO₂ and brine as a phase mixture, the mixed-phase diffusivity, \( D \), is defined as [Hu et al., 2015]:

\[
D = \frac{K}{S_s} = \left( \frac{k}{f} \right) \left[ \frac{k_{rw} \rho_w + k_{rn} \rho_n}{\mu_n} \right] \left( S_w \rho_{sw} + S_n \rho_{sn} \right)
\]

(4.1)

Where \( K \) and \( S_s \) are the mixed-phase conductivity and specific storage. \( k \) is the intrinsic permeability, \( f \) is the effective porosity. \( k_r, \rho, \mu, S, c \) represent the relative permeability, density, viscosity, saturation and compressibility. The subscripts \( w \) and \( n \) indicate wetting phase (brine) and non-wetting phase (CO₂) respectively. \( S_s \) can rise by two orders of magnitude as \( S_n \) increases, whereas \( K \) varies nonmonotonically with \( S_n \), and its variance is within one magnitude [Hu et al., 2015]. Hence, \( D \) rapidly decreases according to the augmented \( S_n \) at the beginning, and then increases slowly as \( S_n \) reaches a certain value. (see section 4.5.2 of the supporting information). Overall, \( D \) during the post-injection phase is smaller than its initial value, and it can be inverted by travel time diagnostics. The travel times relate to the diffusivity by a line integral:

\[
\sqrt{t_{a,d}} = \frac{1}{\sqrt{6 f_{a,d}}} \int_{x_1}^{x_2} ds \frac{D(s)}{D(s)}
\]

(4.2)

Where \( s \) is the propagation path, \( x_1 \) and \( x_2 \) are the injection and observation point respectively. \( f_{a,d} \) is a conversion factor defined as \( -W \left( \frac{\alpha_d^{2/3}}{e} \right) \). \( W \) is Lambert’s \( W \) function, and \( \alpha_d \) relates to the ratio of the first time-derivative of the space-time pressure data to the first time-derivative of the maximum value. If \( f_{a,d} \) equals to 1, then the travel time is the peak time corresponding to the maximum time-derivative value. In analogy to seismic travel time inversion, spatial diffusivity can be resolved by an eikonal solver in a computationally efficient way. Considering that the early time diagnostics resolve the preferential flow better, the 20% travel time diagnostics (\( t_{20\%} \)) are used in this study for the inversion following [Hu et al., 2015]. Derived travel time diagnostics are perturbed with 1%-Gaussian noise for assessing the impact of noise associated with field data. The noise level is assumed the same as that of the seismic travel times, which are derived from a comparable set-up [Ajo-Franklin et al., 2013]. Potential noise can be caused for example by measurement errors due to different pressure devices and techniques.
4.2.5 Clustering

The limits of the inversion approach discussed in [Hu et al., 2015] show that derived absolute values of diffusivity are often not comparable among different tomograms. This applies also to tomograms taken at the same time but separately for reservoir and aquifer. Therefore, following the travel-time inversion, we partition the diffusivity tomograms of the reservoir and aquifer separately into two clusters by k-means clustering. Ideally, one cluster represents the plume and the other, the ambient formation. In contrast to the traditional k-means approach, the centroid of each cluster is determined by fitting the inverted logarithm diffusivities with a mixture of two Gaussian functions. Clustering will define the plume shape in the reservoir, as well as the extent of the secondary plume in the aquifer. Furthermore, the clustered structure can be utilized for acquiring saturations by a sequential and zonal calibration procedure, with the knowledge of initial flow properties. Further details can be found in [Hu et al., 2015].

4.3 Result and discussions

In this section, we only show and discuss the results of the 50 m well pair. Results and discussion of a 100 m doublet can be found in the supplementary document (section 4.5.4).

4.3.1 Head changes

In this section, we examine direct observations and insight from forward modelling by comparing the different responses during stages 2 to 4, given one intact and two leaky caprocks. Pressure data in stage 1 are merely available in case N for deriving the reference travel times and inverting the CO$_2$-free formations, thus it is not presented. For inspection of results, we transfer the pressure data to head change ($\Delta h$) by:

$$\Delta h = \frac{\Delta P_w}{\rho_w g}$$

where $\Delta P_w$ is the difference between the transient pressure data and the initial hydrostatic pressure, $\rho_w$ is the density of brine, and $g$ denotes gravitational acceleration.

The simulated head changes are almost equivalent at RR1-RR3, as well as those among AR1-AR3. Therefore, only the observations in RR2 and AR2 are displayed. As shown in Figure 4.2a and Figure 4.2c, at the onset of depth-integrated CO$_2$ injection in the reservoir, the head change $\Delta h$ in the reservoir rises sharply and then reaches a quasi-steady state. In case N, as expected, pressure does not propagate through the very low-permeability caprock, and there are no responses observed in the aquifer. In case F, $\Delta h$ presents similar changes as in case N during short injection (Figure 4.2a). However, the fracture leakage cannot be recognized yet since it occurs almost at the end of stage 3 when pressure falls to a low level. In reality, this acute leakage may be due to a seismic event happening at the recovery period or when CO$_2$ injection is suspended. During long
injection, $\Delta h$ shows a sudden drop-off in the reservoir when the leaky gap is created due to the high pressure. Simultaneously, $\Delta h$ increases rapidly in the aquifer. After the acute event, $\Delta h$ arrives at a new steady state again. In case D, due to the pressure released permanently across the permeable seal, head changes are observed in both the reservoir and aquifer. Thus, comparison of results for the different cases could indicate the type of leakage.

Figure 4.2 Head changes in the reservoir and aquifer at: a) stage 2-3 during short injection; b) stage 4 during short injection; c) stage 2-3 during long injection; d) stage 4 during long injection. The red dot indicates the time when an acute fracture leakage event occurred.

Figure 4.2b and Figure 4.2d show the head changes at stage 4. $\Delta h$ rises first in the aquifer as brine is injected. In case N, injections in the aquifer or reservoir do not cause fluctuations in the reservoir for two injection periods. In contrast, in the two leakage cases, the leakage can be recognized by the responses in the aquifer. $\Delta h$ then falls back to quasi-steady state for all cases, without a full recovery. After a certain period, CO$_2$ is injected into the reservoir, and $\Delta h$ increases rapidly. Similarly, there are no perturbations in the aquifer in case N, whereas responses are observed in the aquifer in cases F and D. However, in the reservoir, the given time is not sufficient for recovering to steady state.
4.3.2 Early time diagnostics

We computed the early travel time diagnostics ($t_{-20\%}$) by differentiating pressures in stage 1 and 4 of three cases, and compare them respectively (Table 4.2). Note, the diagnostics we present here have been multiplied by the conversion factor $f_{a,d}$ and then squared. At stage 1, $t_{-20\%}$ are equivalent for all the source-receiver configurations with a value of $39.01 \text{ s}^{0.5}$, which indicates the initial homogeneity of reservoir and aquifer. At stage 4, diagnostics in the reservoir show a similar trend for the three cases. Values of $t_{-20\%}$ are small when CO$_2$ is injected at RS1, and they are largest at RS3. Further comparison among travel time diagnostics during short and long injections in each case indicates that the CO$_2$ saturation in the reservoir increased the fastest in case N and the slowest in case F. During the short-injection, t-20% are smaller in the leakage cases when compared to case N. This is because the entire system is different for the three cases. In case N, the seal unit has a very low permeability and in case D, its permeability is even higher. In case F, at the beginning of multilevel brine injection tests, the leaky gap appears in the caprock, changing its effective permeability. Diagnostics in the aquifer also provide a first insight about leakage. In case N, observed minor changes in $t_{-20\%}$ implies that the aquifer can be considered as homogeneous. On the contrary, once CO$_2$ enters the aquifer in cases F and D, values of $t_{-20\%}$ show significant variations during long injection. Since the travel time diagnostics show a relatively large variance when compared to the noise, the diagnostics derived from the reservoir when CO$_2$ is injected are comparable to noise-free data. In contrast, in the aquifer where no CO$_2$ exists, the noise level is more relevant (see the supporting information, Table 4.3).

4.3.3 Inversion and clustering

The inversion is merely based on the travel times rather than on the simulated pressures. Consequently, there is no additional flow simulation required during the inversion. Diffusivity was derived by solving Equation (4.2) utilizing a stagger grid technique [Böhm and Vesnaver, 1996]. This technique can improve the resolution of the final inversion tomogram and alleviate the inversion anomalies caused by grid positions. A model with three rows and three columns was initially set up for the inversion according to the available source-receiver configurations. Sequentially, 23 inversion models were created by shifting the initial model in the horizontal and vertical directions by 7 and 2 times, respectively. The final tomogram was determined by averaging the values of each inversion result with a better resolution of 24 rows and 12 columns. Details on the inversion procedure are shown in the supplement (section 4.5.3).
Table 4.2 Computed values of $t$-20% of different source-receiver configurations in aquifer and reservoir.

<table>
<thead>
<tr>
<th>Source Receiver</th>
<th>Pre Injection</th>
<th>Short Injection</th>
<th>Long Injection</th>
<th>Short Injection</th>
<th>Long Injection</th>
<th>Short Injection</th>
<th>Long Injection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aquifer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS1 AR1</td>
<td>39.01</td>
<td>38.81</td>
<td>38.84</td>
<td>36.67</td>
<td>43.19</td>
<td>37.22</td>
<td>45.50</td>
</tr>
<tr>
<td>AS2 AR1</td>
<td>39.01</td>
<td>38.99</td>
<td>38.84</td>
<td>36.67</td>
<td>43.19</td>
<td>37.22</td>
<td>45.50</td>
</tr>
<tr>
<td>AS3 AR1</td>
<td>39.01</td>
<td>38.99</td>
<td>38.84</td>
<td>36.67</td>
<td>43.19</td>
<td>37.22</td>
<td>45.50</td>
</tr>
<tr>
<td><strong>Reservoir</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS1 RR1</td>
<td>39.01</td>
<td>58.13</td>
<td>63.75</td>
<td>57.61</td>
<td>60.10</td>
<td>53.21</td>
<td>53.90</td>
</tr>
<tr>
<td>RS2 RR1</td>
<td>39.01</td>
<td>96.53</td>
<td>132.23</td>
<td>103.36</td>
<td>113.47</td>
<td>95.81</td>
<td>116.57</td>
</tr>
<tr>
<td>RS3 RR1</td>
<td>39.01</td>
<td>121.86</td>
<td>155.04</td>
<td>131.92</td>
<td>139.27</td>
<td>125.38</td>
<td>145.72</td>
</tr>
<tr>
<td>RS3 RR2</td>
<td>39.01</td>
<td>121.87</td>
<td>155.11</td>
<td>131.94</td>
<td>139.34</td>
<td>126.68</td>
<td>147.19</td>
</tr>
</tbody>
</table>

From forward simulations, the derived saturation distributions are considered as the “truth” (Figure 4.3a). The maximum saturation of CO$_2$ in case N is 0.61 during short injection and 0.67 during long injection. The CO$_2$ plume does not breach the caprock in this case. During the long injection in case F, the leaky path is created. The tip of the plume migrates towards the highly permeable path, and CO$_2$ assembles at the boundary between the leaky gap and the aquifer. The maximum saturation at the top of the leaky gap reaches 0.86, and CO$_2$ is released to the aquifer, driven by buoyancy effects. Simultaneously, the plume in the reservoir extends slower than in case N. In case D, CO$_2$ migrates upwards through the permeable seal, with a maximum saturation value of 0.62 for short and 0.68 for long injection. The plume size in the reservoir is smaller than that in cases N and F during the short injection, and it is between the size observed in cases N and F during the long injection (because of abrupt leakage in case F). During the long injection in case D, the plume reaches the upper aquifer, inducing a secondary plume.

Figure 4.3b presents the mixed-phase diffusivity transferred by Equation (4.1). Original diffusivity of the reservoir and aquifer is 2.46 m$^2$/s, and the low-permeability seal is reflected by $1.2 \times 10^{-5}$ m$^2$/s. Diffusivity value of the leaky gap is 246 m$^2$/s, and it is 1.22 m$^2$/s for the relatively high-permeability seal unit in case F. Injected and leaky CO$_2$ alters the diffusivity by two orders of magnitude. The minimum diffusivity values are not at the front of the plume, since diffusivity does
not change monotonically with CO\textsubscript{2} saturation [Hu et al., 2015]. In cases N and F, the minimum diffusivity is around 0.02 m\textsuperscript{2}/s, corresponding to a saturation of 0.31. In case D, the minimum diffusivity is around 0.01 m\textsuperscript{2}/s, appearing in the caprock where the saturation is 0.25. CO\textsubscript{2} entered into the seal unit in this case due to its unexpectedly high permeability, thus reducing the diffusivity of the seal by a factor of around 2.

Based on the travel times derived from stage 1 and 4, the inverted diffusivity tomograms are shown in Figure 3c. Diffusivity of the caprock is not determined as there is no pressure information from the caprock. The blue areas in the tomograms (Figure 4.3c) indicate a decrease of diffusivity caused by higher CO\textsubscript{2} saturations. Overall, inverted diffusivity varies from 0.01-2.9 m\textsuperscript{2}/s, and these values are not within the range of the “true” values. This discrepancy is expected to be mainly caused by inaccuracies of the single-phase proxy and by the nonuniform trajectory distribution. Locally, small trajectory density masks small information content, and thus results in a poorly posed inversion problem. However, the small diffusivity values in the tomograms still provide structural information and thus insight into plume evolution. In the three cases, the inverted diffusivity values in the reservoir are much larger than those in the aquifer due to the larger variance of travel times derived in the reservoir. The inverted diffusivity values of the aquifer show minor changes during the short injection, which indicates a homogeneous distribution and no disturbance. During the long injection, the diffusivity in the aquifer shows a decrease when the leakage occurs in cases F and D. Figure 4.3e and Figure 4.3f show the reliability maps, which are discussed in the supporting document (section 4.5.5). In this study, only three sources and receivers are applied, leading to a relatively high null space energy distribution. Therefore, these reliability maps are utilized mainly as a reference for the inverted flow field, but not for the following clustering procedure.

For obtaining the plume shape, a k-means clustering method was applied (Figure 4.3d). The plumes identified in cases N and D grow with time, while in case F, plume shapes are similar for different injection periods. This is consistent with the conditions simulated for the “truth”. As the most striking feature, the secondary plumes in the aquifer are also characterized by the clustering. Although the results imply that the plume extents are slightly overestimated, they still could provide information on the leakage type, potential position and migration paths. The inversion results and reliability maps of 1%-noise data are similar to the noise-free data for delineating the plume shape. Slight differences due to this noise level can thus be neglected. However, in case F, increase of the distance between the two wells reduces the trajectory density of the travel time data set, weakening the detectability of the leaky plume in the upper aquifer by adding noise (Figure 4.3, lines d and h). This can be overcome by increasing the number of measurements. In addition, the secondary CO\textsubscript{2} plume in the aquifer above can be diluted by the ambient groundwater flux or hydraulic activities in the aquifer, which may hamper detecting especially small leaks.
Figure 4.3. a) “true” CO$_2$ saturation; b) “true” diffusivity transferred from “true” CO$_2$ saturation; c) inverted diffusivity (noise-free); d) inverted diffusivity (1% noise); e) reliability map (noise-free); f) reliability map (1% noise); g) clustered structure of pre-, short- and long- injection periods of three cases (noise-free); h) clustered structure of the pre-, short- and long- injection periods of three cases (1% noise).
4.4 Conclusions

A rapid pressure-based inversion approach is proposed for detecting CO$_2$ leakage through the innovative tomographical set-up of fluid injection tests. It can be adapted to various cases of CO$_2$ leakage without previous determination of possible leaky locations. CO$_2$ leaking into the caprock and upper aquifer retards the plume development in the storage formation, which can be reflected through the comparison of inverted diffusivity tomograms. In comparison to previous pressure-based approaches, which mainly focus on the pre-existing leaky paths, it is now also possible to localize the leaky location and approximately delineate the spreading area of leakage in different times by the inverted plumes in the above aquifer. Results indicate that pressure tomography is more suitable for identifying small-scale leaks, which might occur near the CO$_2$ injection well. Still, the influence of formation heterogeneity will be explored in future work.

4.5 Supporting information

4.5.1 Introduction

In the following supporting information, we present how to transfer CO$_2$ saturations to diffusivities in section 4.5.2 and Figure 4.4. Section 4.5.3 and Figure 4.5 show the details of the inversion procedure. We also simulate an additional scenario in which the distance between the injection and observation wells is 100 m. Results (section 4.5.4 and Figure 4.6, Figure 4.7) are shown in a similar manner as Figure 4.3 in the main manuscript, and section 4.5.4 gives a brief discussion on these results. We do not intend to discuss these results in the main manuscript, since they are comparable to the results of 50 m well spacing, and they are provided as a sensitivity analysis of our experimental set-up. In section 4.5.5, the significance of reliability maps is briefly elaborated. Section 4.5.6 explains the clustering approach with an example shown in Figure 4.8.

Table 4.3 shows the early travel time diagnostics ($t$-$20\%$) of 50 m well spacing, perturbed by 1%-Gaussian noise.

4.5.2 Relationship between CO$_2$ saturation and mixed-phase diffusivity

Diffusivity is the ratio of $K$ and $S_j$, which are calculated by intrinsic permeability, effective porosity, fluid density, viscosity and compressibility. Density, viscosity and compressibility vary along with the pressure and temperature. For converting saturations to diffusivities, we utilize the mean values of these three parameters during the injections at stage 4:
- Density of brine ($\rho_w$): 1057.5 kg/m$^3$
- Density of CO$_2$ ($\rho_s$): 531 kg/m$^3$
- Viscosity of brine ($\mu_w$): $4.3\times10^{-4}$ Pa·s
Viscosity of CO$_2$ ($\mu_n$): $4 \times 10^{-5}$ Pa·s

Compressibility of brine ($\frac{dP_n}{dP_w} \frac{1}{\rho_w}$): $3.8 \times 10^{-10}$ 1/Pa

Compressibility of CO$_2$ ($\frac{dP_n}{dP_n} \frac{1}{\rho_n}$): $8.8 \times 10^{-8}$ 1/Pa

Figure 4.4 CO$_2$ saturation vs. mixed-phase diffusivity of three components.

### 4.5.3 Inversion procedure

The inversion program we used in this work is Geotom [Jackson and Tweeton, 1996]. It uses a least squares solution for solving a linear inverse problem. The diffusivity distribution is inverted by an algorithm called SIRT (simultaneous iterative reconstruction technique). More details can be found in Jackson and Tweeton [1996]. In general, the diffusivity is calculated based on the observed travel time diagnostics by a ray-tracing technique. It is first assigned to an initial distribution. Sequentially, the root-mean-square (RMS) residual is calculated. Here the residual is defined as the difference between a calculated diagnostic ($\sqrt{6t_{cal,i}f_{a,d}}$) and an observed diagnostic ($\sqrt{6t_{obs,i}f_{a,d}}$).

The RMS residual is the square root of the mean of the squared residuals. For each iteration, the diffusivity is updated to generate new diagnostics. The RMS residual is minimized through the iterative process, based on the objective function ($\Psi$):

$$\Psi = \min \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \sqrt{6t_{obs,i}f_{a,d}} - \sqrt{6t_{cal,i}f_{a,d}} \right)^2}$$

(4.4)

where $t_{obs,i}$ and $t_{cal,i}$ are the $i$th observed and calculated peak times, respectively. $N$ is the total number of measurements of travel times. Figure 4.5 presents an example of how the RMS residual converges. Inversion cost less than 1 min to a maximum of 10 min, depending on the number of iterations.
4.5.4 Results of configuration with 100 m well spacing

Figure 4.6 and Figure 4.7 show the results of a 100 m distant well pair for cases N, F, and D, respectively. Inverted plumes in the reservoir (see Figure 4.6g and Figure 4.6h, Figure 4.7g and Figure 4.7h) are slightly larger compared to the results of the 50 m well pair (see Figure 4.3g and Figure 4.3h). In the aquifer, the size of the plumes is also larger than those of the 50 m well pair (see Figure 4.3g and Figure 4.3h, Figure 4.6g and Figure 4.6h, Figure 4.7g and Figure 4.7h). This is because we only increased the lateral distance between the two wells, but without changing the aquifer thickness. Consequently, for larger well distance, the tomographical arrays are constrained with smaller angles and this aggravates the resolution of the plume. This however, is not only critical to pressure tomography, but also to any other crosswell tomographical approaches, such as crosswell seismic tomography.

Increasing the distance also reduces the detectable saturation contrast in the aquifer. Small leaks can hardly be resolved if the noise is added to the travel times (Figure 4.7b, long-injection, case F, line h). Therefore, for resolving the small-scale leaks, the distance between the two wells should not be too large compared to the aquifer thickness. This allows the available tomographical arrays resolving more complicated plume geometries.
Figure 4.6 a) “true” CO₂ saturation; b) “true” diffusivity transferred from “true” CO₂ saturation; c) inverted diffusivity (noise-free); d) inverted diffusivity (1% noise); e) reliability map (noise-free); f) reliability map (1% noise); g) clustered structure of pre-, short- and long- injection periods of three cases (noise-free); h) clustered structure of the pre-, short- and long- injection periods of three cases (1% noise).
Figure 4.7 a) “true” CO₂ saturation; b) “true” diffusivity transferred from “true” CO₂ saturation; c) inverted diffusivity (noise-free); d) inverted diffusivity (1% noise); e) reliability map (noise-free); f) reliability map (1% noise); g) clustered structure of pre-, short- and long- injection periods of three cases (noise-free); h) clustered structure of the pre-, short- and long- injection periods of three cases (1% noise).
4.5.5 Reliability map

Reliability of the inversion results is estimated by the so-called “null space energy” (NSE), which is originally an indicator of the ray-coverage for seismic tomography (Böhm and Vesnaver [1996]). In pressure tomography, the NSE indicates the “trajectory coverage”. The value of NSE ranges from 0 to 1, reflecting large to small trajectory coverage. Low NSE values (large trajectory coverage) indicate the preferential flow pathways, whereas high NSE values reflect the parts with less flow pathways in the formation. Hence, high NSE values could be considered as an indicator for plumes, since the plumes lower the diffusivity of initially CO₂-free formations.

In case N, prior to CO₂ injection, NSE shows higher values near the boundaries because of the crosswell set-up. Less trajectories exist near the upper and bottom boundaries, as well as near the injection and observation wells. After CO₂ injection, in all cases, the values of NSE in the reservoir increase at the left due to the evolving plume. Correspondingly, the low NSE values at the center and at the right indicate dense trajectories. The higher diffusivity values at the right part are consistent with the location where the trajectories are most intensive.

In case F and D, the NSE distribution during short-injection in the aquifer is comparable to that of the no-leakage case. The leaks in the aquifer appear during long injection. In comparison with the NSE maps of short-injection, NSE shows higher values near the leaky locations (Figure 4.3e and Figure 4.3f, Figure 4.7e and Figure 4.7f, red line).

4.5.6 Clustering results

Clustering is based on the k-means method. A dataset of inverted diffusivity is partitioned into two clusters, representing plume and ambient aquifer, respectively. In lieu of choosing the initial centroids arbitrarily, here the centroids are determined by the probability density function (pdf) of the log transformation of diffusivity. The pdf is fitted with the summation of two 1-D Gaussian functions:

\[
f(x) = a_1 \exp\left(-\frac{(x-b_1)^2}{c_1^2}\right) + a_2 \exp\left(-\frac{(x-b_2)^2}{c_2^2}\right)
\]

(4.5)

The centers, \(b_1\) and \(b_2\), are utilized as the centroids. Afterwards, the points within the dataset are clustered by calculating their distances to the centroids. In a time-lapse process, the centroids of the long-injection are then utilized for clustering the diffusivities derived from short-injection. Figure 4.8 shows exemplarily the pdf and fitted Gaussian functions.
Figure 4.8 The pdf and fitted Gaussian functions of the logarithm diffusivities in case D. a) in the reservoir; b) in the aquifer

Table 4.3 Computed values of t-20% of different source-receiver configurations in aquifer and reservoir (the diagnostics are perturbed with 1%-Gaussian noise)

<table>
<thead>
<tr>
<th>source receiver</th>
<th>pre-injection</th>
<th>short-injection</th>
<th>long-injection</th>
<th>short-injection</th>
<th>long-injection</th>
<th>short-injection</th>
<th>long-injection</th>
</tr>
</thead>
<tbody>
<tr>
<td>aquifer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS1</td>
<td>38.60</td>
<td>38.47</td>
<td>38.87</td>
<td>36.70</td>
<td>42.94</td>
<td>37.57</td>
<td>45.04</td>
</tr>
<tr>
<td>AR1</td>
<td>38.70</td>
<td>38.77</td>
<td>38.88</td>
<td>36.85</td>
<td>42.61</td>
<td>37.15</td>
<td>45.09</td>
</tr>
<tr>
<td>AR2</td>
<td>37.86</td>
<td>38.68</td>
<td>38.58</td>
<td>36.65</td>
<td>42.55</td>
<td>36.73</td>
<td>45.28</td>
</tr>
<tr>
<td>AR3</td>
<td>39.57</td>
<td>38.53</td>
<td>38.92</td>
<td>36.25</td>
<td>41.68</td>
<td>37.08</td>
<td>35.92</td>
</tr>
<tr>
<td>AS2</td>
<td>39.14</td>
<td>38.89</td>
<td>39.03</td>
<td>36.39</td>
<td>40.92</td>
<td>37.58</td>
<td>35.91</td>
</tr>
<tr>
<td>AR1</td>
<td>38.72</td>
<td>38.88</td>
<td>39.10</td>
<td>36.49</td>
<td>41.74</td>
<td>36.47</td>
<td>35.84</td>
</tr>
<tr>
<td>AR2</td>
<td>39.55</td>
<td>39.56</td>
<td>38.77</td>
<td>36.99</td>
<td>40.61</td>
<td>37.32</td>
<td>34.37</td>
</tr>
<tr>
<td>AR3</td>
<td>38.34</td>
<td>38.86</td>
<td>38.82</td>
<td>36.09</td>
<td>40.82</td>
<td>36.42</td>
<td>34.44</td>
</tr>
<tr>
<td>AS3</td>
<td>38.97</td>
<td>38.55</td>
<td>39.70</td>
<td>36.66</td>
<td>40.50</td>
<td>37.41</td>
<td>33.65</td>
</tr>
<tr>
<td>reservoir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS1</td>
<td>38.68</td>
<td>58.06</td>
<td>64.10</td>
<td>56.17</td>
<td>60.30</td>
<td>53.46</td>
<td>54.12</td>
</tr>
<tr>
<td>RR1</td>
<td>39.04</td>
<td>58.27</td>
<td>65.11</td>
<td>57.92</td>
<td>59.63</td>
<td>52.76</td>
<td>54.56</td>
</tr>
<tr>
<td>RR2</td>
<td>38.54</td>
<td>57.92</td>
<td>62.67</td>
<td>56.91</td>
<td>60.09</td>
<td>53.79</td>
<td>55.12</td>
</tr>
<tr>
<td>RR3</td>
<td>38.58</td>
<td>97.37</td>
<td>133.37</td>
<td>103.09</td>
<td>113.87</td>
<td>96.44</td>
<td>115.76</td>
</tr>
<tr>
<td>RS2</td>
<td>39.01</td>
<td>95.24</td>
<td>132.75</td>
<td>103.55</td>
<td>111.47</td>
<td>96.51</td>
<td>117.69</td>
</tr>
<tr>
<td>RR1</td>
<td>39.61</td>
<td>97.00</td>
<td>130.73</td>
<td>104.19</td>
<td>114.78</td>
<td>97.98</td>
<td>118.36</td>
</tr>
<tr>
<td>RR2</td>
<td>38.71</td>
<td>120.83</td>
<td>154.37</td>
<td>131.56</td>
<td>142.64</td>
<td>126.64</td>
<td>146.93</td>
</tr>
<tr>
<td>RR3</td>
<td>39.16</td>
<td>121.46</td>
<td>155.65</td>
<td>134.01</td>
<td>140.63</td>
<td>125.00</td>
<td>147.00</td>
</tr>
<tr>
<td>RR3</td>
<td>38.92</td>
<td>122.54</td>
<td>160.74</td>
<td>131.31</td>
<td>138.90</td>
<td>127.00</td>
<td>148.52</td>
</tr>
</tbody>
</table>
5 Characterizing CO$_2$ plumes in deep saline formations: comparison and joint evaluation of time-lapse pressure and seismic tomography


Abstract

Monitoring the migration of sequestered CO$_2$ in deep heterogeneous reservoirs is inherently difficult. Geophysical methods have been successfully used, but flow conditions are only indirectly linked to the measured properties. Pressure tomography (PT) has been proposed as an alternative method to depict the structure of deep saline formations for CO$_2$ sequestration and to continuously delineate CO$_2$ plumes. In contrast to more established geophysical measurements, pressure transients are directly related to flow conditions, which allows for the calibration of permeability. We investigate the influence of aquifer heterogeneity on PT performance, and compare PT results to crosshole seismic traveltime tomography (ST). Multilevel fluid injections and high-frequency P-wave pulses are induced in a simulated deep borehole, and the recorded signals at another well are processed by a traveltime inversion scheme. The reservoir structure is inferred by clustering the inverted hydraulic diffusivity prior to CO$_2$ injection, and the plume distribution is determined by clustering the tomograms of the inverted mixed-phase diffusivity difference and P-wave velocity difference. The clustered structures are then utilized for zonal calibration to acquire the saturation within the plumes. Modeling results indicate that PT provides a clearer structural information of the CO$_2$-free aquifer due to its direct linkage to the permeability. However, the plume depicted by PT can be ambiguous, whereas ST is less sensitive to the prevailing heterogeneity of permeability at post-injection and can thus image the plume more clearly. PT and ST can be complementary to each other through the joint clustering to improve plume shape identification and estimation of spatial CO$_2$ saturation.
5.1 introduction

Recently, the report “Global status of CCS: 2015” [Global CCS Institute, 2015] pointed out that carbon dioxide capture and storage (CCS) is the only countermeasure to lessen greenhouse gas emissions in a significant scale from industrial processes. Among various geological storage media, deep saline aquifers are considered sound formations for sequestering CO₂, in which CO₂ is injected and stored in a supercritical state, with a large storage capacity. Safe disposal of CO₂ in saline aquifers demands favorable storage conditions, such as high porosities in the storage medium and an impermeable caprock. In case of unfavorable conditions, such as seal imperfections, pre-existing faults or overpressure-induced fractures, an evolving CO₂ plume can escape towards shallower formations and contaminate potable aquifers. For minimizing the risk of CO₂ leakage and for formulating effective remediation strategies, appropriate site investigation and CO₂ plume monitoring techniques are required.

Geophysical exploration methods are extensively applied for depicting stratigraphy and CO₂ plume geometry in deep saline formations. Their applicability has been demonstrated at several CO₂ storage sites, such as Sleipner [e.g., Chadwick et al., 2010], Snøhvit [e.g., Shi et al., 2013], In Salah [e.g., Ringrose et al., 2009], Ketzin [e.g., Zhang et al., 2012], Cranfield [e.g., Doetsch et al., 2013], Frio brine pilot [e.g., Daley et al., 2011], and Nagaoka [e.g., Nakajima et al., 2014]. The most common geophysical approaches for monitoring of CO₂ plumes are seismic surveys, electrical methods, and gravity measurements. The induced CO₂ phase alters the effective physical properties of the storage formation (e.g., seismic velocity, electrical resistivity and density) over time, which is examined by time-lapse data sets.

Among established geophysical techniques, active seismic surveys are most commonly used. Usually, seismic measurements are conducted prior to CO₂ injection to obtain baseline information, and then measurements are repeated multiple times after CO₂ injection. CO₂ plume evolution is monitored by the traveltime delay or amplitude anomalies from different vintages. Depending on the configurations of sources and receivers, seismic-based approaches can be classified into: a) Surface seismic survey. Typically, a 2-D or 3-D surface seismic survey is conducted for large-scale problems, and its spatial resolution is limited. In the case that the CO₂ layer thickness is less than the resolution, the uncertainty for evaluating the CO₂ mass becomes significant [Ivandic et al., 2015]. b) Surface-downhole monitoring. This approach includes two different configurations, vertical seismic profiling (VSP) and moving-source-profiling (MSP). It is employed for estimating the vertical expansion of the CO₂ plume and for improving the vertical resolution to complement surface seismic methods. c) Crosshole measurements. Here, the seismic sources and receivers are installed in different boreholes, and the experiments are performed for obtaining insight in the reservoir and mapping CO₂ plumes between the borehole pair. The crosshole variant can provide high resolution information between the boreholes, which are typically separated by a distance of tens to hundreds of meters. The main inversion algorithms for crosshole seismic tomography include traveltime based and full waveform inversion [e.g., Pratt and Shipp, 1999].
While successful at imaging migrating CO₂, two difficulties have been identified for established geophysical approaches. First, petrophysical models have to be formulated for analyzing the stored CO₂ in the subsurface, which is challenging mainly due to the uncertainty in model parameters. For instance, for a petrophysical model that couples a patchy fluid saturation model, the patch size of CO₂ is difficult to determine precisely in practice [Daley et al., 2011]. Second, favorable conditions are required for the repeatability of geophysical experiments, such as high contrast of reservoir properties, identical experimental set-ups at different times, and low noise levels. Not having these ideal conditions can lead to artefacts during inversion of time-lapse data sets, or during conversion of geophysical parameters into CO₂ saturation.

As a complementary method for characterizing reservoir and monitoring CO₂ plumes during the early time of storage, as well as for detecting CO₂ leakage, pressure-based methods have recently been suggested. Since pressure directly relates to flow conditions and travels much faster than a CO₂ plume, pressure-based methods are recognized as an appropriate approach for evaluating the flow properties before and during CO₂ injection [e.g., Doughty et al., 2008; Wiese et al. 2010], as well as for early detection of CO₂ leakage [e.g., Birkholzer et al., 2009; Sun et al., 2016]. However, few pressure-based methods can provide a spatial image of CO₂ plumes. Several available techniques can identify the plume shape to a certain degree. For instance, the uprising of a CO₂ plume near the well or a CO₂ front can be inferred by some analytical or semi-analytical solutions [e.g., Nordbotten et al., 2004; Cihan et al., 2011]. The approach of Martinez-Landa et al. [2013] can estimate the proximal width of a CO₂ plume by analyzing pressure measurements in a single borehole. Nevertheless, most of these methods are based on the assumption that the reservoir is homogeneous, neglecting most of the involved physicochemical CO₂ transport processes and potential reservoir heterogeneity.

For resolving spatial variability of flow properties, hydraulic tomography (HT) was proposed by Gottlieb and Dietrich [1995] and has seen significant development since its introduction. Compared to conventional hydraulic/pressure tests, HT can delineate the spatial distribution of hydraulic parameters. Feasibility of HT has been studied in both porous [e.g., Yeh and Liu, 2000; Hu et al., 2011] and fractured media [e.g., Illman et al., 2009; Zha et al., 2015] by numerical simulations [e.g., Jiménez et al., 2013], laboratory experiments [e.g., Brauchler et al., 2007] and field tests [e.g., Brauchler et al. 2013; Jiménez et al. 2015]. The application scale of HT can vary from several meters to kilometers. The rationale of HT is conceptually analogous to geophysical tomographic techniques. In lieu of using geophysical sources (e.g., active or passive seismic excitations), HT or more generally pressure tomography (PT), requires a series of pressure stimulations in a tomographic array. Tomographical data sets are derived by conducting fluid injection or extraction tests in different intervals at one well (sources), with pressures measured at the response well at different observation levels (receivers). The pressure measurements are used for reconstructing hydraulic parameter heterogeneity by different inversion techniques, such as sequential successive linear estimator (SSLE) [Yeh and Liu, 2000], quasi-linear Bayesian geostatistical method [e.g., Nowak and Cirpka, 2004], ensemble Kalman filter (EnKFs) [e.g., Schöniger et al., 2012], and the travel-time based approach [Brauchler et al., 2003]. Among these
techniques, the travel-time based inversion approach requires minor computational effort as it is a deterministic approach without geostatistical bias, and the inversion process does not involve flow modeling. Instead, the groundwater flow equation is approximated to an eikonal equation, which can be solved by the ray-tracing technique [e.g., Jackson and Tweeton, 1996]. Structural information of the aquifer is then inferred from reconstructed diffusivity tomograms.

To date, HT methods are mainly applied to the “static” single-phase flow condition (i.e., hydraulic parameters are considered not varying with time) and shallow aquifers. Hu et al. [2015] introduced the concept of “time-lapse pressure tomography”. It is proposed to utilize the travel-time based inversion strategy for identifying an evolving CO\textsubscript{2} plume in a homogeneous deep saline aquifer. In this case, replacement of the local brine by CO\textsubscript{2} induces an effect on the observed flow properties over time. Considering CO\textsubscript{2} and brine as a phase mixture, the mixed-phase diffusivity can be changed by up to two orders of magnitude due to the high compressibility of CO\textsubscript{2}. Spatial diffusivity variations can be inferred from inspecting pressure transients at different times. Thus, the inversion of traveltimes derived from pressures offers time-lapse information of the plume. Furthermore, CO\textsubscript{2} saturation can be estimated by including a storativity-saturation model.

The main objective of this study is to explore the inversion performance of PT in heterogeneous formations, and to compare the results with those from crosshole traveltime seismic tomography. In contrast to seismic traveltime tomography, two-phase PT is unusual for such conditions. Due to the direct and more deterministic relationship between CO\textsubscript{2} saturation and hydraulic properties such as storativity, it is anticipated that pressure tomography can complement existing seismic approaches through the improved estimation of saturation. In the following, we first briefly introduce the two tomographic inversion concepts utilized in a time-lapse manner. These are examined together in scenarios with different degrees of heterogeneity. For simplification, the expression “seismic tomography” mentioned in this study refers to the traveltime based seismic tomography method.

5.2 methodology

5.2.1 Overview of the methodology

In this work, the inversion methodology is tested using real site-based synthetic models. As a basis, we set up three scenarios for this synthetic model, which differ with respect to heterogeneity and model parameters. Subsequently, we investigate the changes of the mixed-phase diffusivity and P-wave velocity induced by CO\textsubscript{2} injection. The relationship between either diffusivity or velocity and CO\textsubscript{2} saturation provides the basis for the inversion.

Figure 5.1 presents the general flowchart of the forward simulation, inversion and calibration procedures. The CO\textsubscript{2} sequestration process is simulated with a fully-coupled two-phase simulator, PFLOTRAN [Hammond et al., 2014]. PT and ST data acquisition are simulated prior to and after CO\textsubscript{2} injection. The derived hydraulic and seismic traveltimes are inverted separately to reconstruct the spatial distribution of diffusivity and velocity. The structure of the reservoir and the geometry
of the CO₂ plume are obtained by individual or joint clustering of the tomograms at different times. Ultimately, we acquire the CO₂ saturation through the calibrated specific storage based on the clustered structure, and the calibration is conducted in a single-phase emulator. In the following sections, each step is explained in detail.

![Figure 5.1 Schematic flowchart of the forward simulation, inversion and calibration strategy](image)

### 5.2.2 Problem set-up

#### Virtual site and three scenarios

A simplified 2-D cross-sectional model based on a virtual site is utilized for testing our method (as in Hu et al., 2015; 2016). The simulated regime is located at a depth of 1600 m, constituting three components: a storage reservoir, an overlying caprock and an underlying bottom seal (Figure 5.2). The thickness of the reservoir and the two low permeability components are 15 m and 30 m, respectively. The reservoir is composed of sandstone, and the caprock and seal bottom are shale formations with very low permeability. For improving computing efficiency, the caprock and seal bottom are assumed to be impervious and considered as no-flow boundaries during flow simulation. The lateral extension of the entire model is 580 m, which is bounded with constant hydrostatic pressure at the east and west sides. The pressure at the model bottom is 14.76 MPa, with a pressure gradient of 0.01 MPa/m. The aquifer is initially fully saturated with brine, and its temperature and
salinity are 67 °C and 67 g/l, respectively. Under these conditions, CO$_2$ is injected and sequestered in a supercritical state.

Fluid is injected at the center of the model, with an accompanying observation well located 50 m away. The distance between the well pair is comparable to several practical injection sites, such as Ketzin [Wiese et al., 2010] and Heletz [Niemi et al., 2015]. The lateral grid size of the model increases telescopically, ranging from 0.09 m at the injector to 40 m at the side boundaries. The vertical discretization of the model is 0.6 m. Finally, the model is discretized by 287 and 75 grid cells in the horizontal and vertical direction, with 21525 grid cells in total (see Appendix A).

For PT, 5 sources and 5 receivers are used. These are screened at injection and observation wells in the reservoir (Figure 5.2, red circles and crosses). The length of each source (fluid injection interval) is 0.6 m, and the distance between two adjacent sources or receivers is 3.6 m. Sources and receivers for ST are also assumed to be located at the two wells (sources are in the injection well and receivers are in the observation well). The distribution of source-receiver configurations for ST experiments is much denser (Figure 5.2, blue circles and crosses). Here, 76 sources and 76 receivers are installed not only in the reservoir, but also in the caprock and bottom seal. This yields larger-angle tomographic arrays in order to improve spatial resolution. The vertical distance between two adjacent seismic sources or receivers is 1 m.

Figure 5.2 Conceptual model and source-receiver configurations of PT and ST. Three variants of aquifer heterogeneity, one homogeneous and two two-layer scenarios are considered. Note that sources and receivers for ST are illustrated schematically, and their true numbers are much higher (76).
Three different scenarios are defined for exploring the influence of reservoir heterogeneity. In the first scenario, the reservoir is perceived as a homogeneous and isotropic aquifer. In the second and third scenarios, the reservoir consists of two “perfect”, homogeneous and isotropic layers (Figure 5.2). The thickness of the upper and lower layers is 5 m and 10 m, respectively. We discriminate these two scenarios by assigning different values of model parameters (permeability, porosity, and other parameters calculated by these) for the two layers. In the next sections, we refer to these three scenarios as “homogeneous”, “2layers_A” and “2layers_B”, with 2layers_A having the higher permeability in the upper layer and 2layers_B in the lower layer.

Model parameters

Model parameters are summarized in Table 5.1 and Table 5.2. The values of model parameters for the homogeneous scenario are the same as in Hu et al. [2015]. The intrinsic permeability (k), porosity (φ) and rock density (ρr) are 1×10^{-13} m^2, 0.23 and 2550 kg/m^3, respectively. The characteristic functions of permeability-saturation and capillary pressure-saturation are based on the Brooks-Corey-Burdine model [Brooks and Corey, 1964; Burdine, 1953]. The pore size distribution (λ) and entry pressure (P_a) are set to 0.76 and 4000 Pa.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intrinsic permeability of the caprock and seal bottom (k_cap and k_seal)</td>
<td>1×10^{-19} m^2</td>
<td>Wang and Small [2014]</td>
</tr>
<tr>
<td>Porosity of the caprock and seal bottom (φ_cap and φ_seal)</td>
<td>0.05</td>
<td>Wang and Small [2014]</td>
</tr>
<tr>
<td>Salinity</td>
<td>67 g/l</td>
<td>estimated</td>
</tr>
<tr>
<td>P-wave velocity of the caprock and seal bottom (V_cap and V_seal)</td>
<td>3500 m/s</td>
<td>Markov et al. [2009]</td>
</tr>
<tr>
<td>P-wave velocity of CO2 (V_n)</td>
<td>292.9 m/s</td>
<td>NIST</td>
</tr>
<tr>
<td>Initial datum pressure (P_0)</td>
<td>14.76 MPa</td>
<td>Erlström et al. [2011]</td>
</tr>
<tr>
<td>Initial average brine density (ρ_w)</td>
<td>1052 kg/m^3</td>
<td>Duan et al. [2008]</td>
</tr>
<tr>
<td>Initial average brine viscosity (μ_w)</td>
<td>4.2×10^{-4} Pa s</td>
<td>Span and Wagner [1996]</td>
</tr>
<tr>
<td>Average isothermal compressibility of brine (c_w)</td>
<td>3.8×10^{-10} 1/Pa</td>
<td>Duan et al. [2008]</td>
</tr>
<tr>
<td>Average CO2 density (ρ_n)</td>
<td>520 kg/m^3</td>
<td>Span and Wagner [1996]</td>
</tr>
<tr>
<td>Average isothermal compressibility of CO2 (c_n)</td>
<td>9×10^{-8} 1/Pa</td>
<td>Span and Wagner [1996]</td>
</tr>
<tr>
<td>Average CO2 viscosity (μ_n)</td>
<td>3.9×10^{-5} Pa s</td>
<td>Fenghour et al. [1998]</td>
</tr>
<tr>
<td>Bulk modulus of dry frame rock (G_{dry})</td>
<td>3×10^9 Pa</td>
<td>Caspari et al. [2011b]</td>
</tr>
<tr>
<td>Bulk modulus of rock matrix (G_m)</td>
<td>3.7×10^{10} Pa</td>
<td>Markov et al. [2009]</td>
</tr>
<tr>
<td>Shear modulus of saturated rock (N_{sat})</td>
<td>2.5×10^{10} Pa</td>
<td>Markov et al. [2009]</td>
</tr>
</tbody>
</table>
In the scenario 2layers_A, the permeability of the top layer is $1 \times 10^{-12}$ m$^2$ and the bottom layer is $1 \times 10^{-13}$ m$^2$. The corresponding $\phi$ is set to 0.28 and 0.22, which is calculated from $k$ according to an empirical equation for sandstone [Schön, 2011]

$$k = 0.04\phi \exp(35.77\phi)$$  \hspace{1cm} (5.1)

The pore size distribution $\lambda$ is set to a constant value of 1.5 for the entire model. $P_d$ is calculated by the Leverett scaling function [Leverett, 1941], which explains the difference in $k$, $\phi$ and $P_d$ between a reference and an unknown media

$$P_d = \sqrt{\frac{k_{ref}}{k}} \phi_{ref} P_{d,ref}$$  \hspace{1cm} (5.2)

where the values of the reference permeability ($k_{ref}$), porosity ($\phi_{ref}$) and entry pressure ($P_{d,ref}$) are $1 \times 10^{-13}$ m$^2$, 0.25 and 8700 Pa, respectively [Rasmusson et al., 2014]. Rock density ($\rho_r$) is estimated by the rock matrix density ($\rho_m$) and brine density ($\rho_w$). In the scenario 2layers_B, the values of the model parameters are switched between the layers.

Table 5.2 Parameter values of the reservoir in the three scenarios. Hydraulic conductivity, specific storage, diffusivity, entry pressure and velocity are calculated based on intrinsic permeability and porosity.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Zone</th>
<th>$k$ (m$^2$)</th>
<th>$\phi$</th>
<th>$K_w$ (m/s)</th>
<th>$S_{sw}$ (1/m)</th>
<th>$D$ (m$^2$/s)</th>
<th>$P_d$ (Pa)</th>
<th>$V$ (m/s)</th>
<th>$c_p$ (J/kg·K)</th>
<th>$\kappa_{wet}$ (W/mK)</th>
<th>$\kappa_{dry}$ (W/mK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>homogeneous</td>
<td>1</td>
<td>$1 \times 10^{-13}$</td>
<td>0.23</td>
<td>2.46$\times 10^{-6}$</td>
<td>9.1$\times 10^{-7}$</td>
<td>2.7</td>
<td>4000</td>
<td>4465</td>
<td>930</td>
<td>3</td>
<td>4.5</td>
</tr>
<tr>
<td>2layers_A</td>
<td>1</td>
<td>$1 \times 10^{-12}$</td>
<td>0.28</td>
<td>2.46$\times 10^{-5}$</td>
<td>1.1$\times 10^{-6}$</td>
<td>22.36</td>
<td>2932</td>
<td>4579</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$1 \times 10^{-13}$</td>
<td>0.22</td>
<td>2.46$\times 10^{-6}$</td>
<td>8.6$\times 10^{-7}$</td>
<td>2.86</td>
<td>8143</td>
<td>4454</td>
<td>860</td>
<td>2.5</td>
<td>4.5</td>
</tr>
<tr>
<td>2layers_B</td>
<td>1</td>
<td>$1 \times 10^{-13}$</td>
<td>0.22</td>
<td>2.46$\times 10^{-6}$</td>
<td>8.6$\times 10^{-7}$</td>
<td>2.86</td>
<td>8143</td>
<td>4454</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>$1 \times 10^{-12}$</td>
<td>0.28</td>
<td>2.46$\times 10^{-5}$</td>
<td>1.1$\times 10^{-6}$</td>
<td>22.36</td>
<td>2932</td>
<td>4579</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Relationship between flow properties and fluid saturation**

The initial single-phase hydraulic conductivity ($K_w$) and specific storage ($S_{sw}$) of the aquifer are related to the intrinsic permeability ($k$), porosity ($\phi$), brine density ($\rho_w$) and viscosity ($\mu_w$) by Equations (5.3) and (5.4):

$$K_w = \rho_w g \left( \frac{k}{\mu_w} \right)$$  \hspace{1cm} (5.3)
\[ S_{sw} = \rho_w g (f c_w) \]  \hspace{1cm} (5.4)

where \( \rho_w, \mu_w \) and \( c_w \) are assumed constant by averaging the values within the reservoir. Furthermore, hydraulic diffusivity of the initial CO\(_2\)-free formation \( D_{pre} \) is defined as the ratio between \( K_w \) and \( S_{sw} \):

\[ D_{pre} = \frac{k}{f c_w \mu_w} \]  \hspace{1cm} (5.5)

The transferred values of \( D_{pre} \) for the three scenarios are listed in Table 5.2.

After CO\(_2\) injection, considering CO\(_2\) and brine as a phase mixture, the mixed-phase diffusivity \( D_{post} \) is determined by the intrinsic permeability and porosity, the fluid properties of brine and CO\(_2\) (density, viscosity, and compressibility), the relative permeability, and the saturation of the two phases [Hu et al., 2016]:

\[ D_{post} = \frac{K}{S_b} = f \left( \frac{k}{\mu_w} \right) \left( \frac{k_{rw} \rho_w + k_{rn} \rho_n}{\mu_w \mu_n} \right) \left( \frac{S_w c_w + S_n c_n}{S_w \rho_w + S_n \rho_n} \right) \]  \hspace{1cm} (5.6)

where \( k_{rw} \) and \( k_{rn} \) are the relative permeability of brine and CO\(_2\), respectively. Equations (5.5) and (5.6) indicate that the diffusivity varies with CO\(_2\) saturation. Figure 5.3 displays five different \( D-S_n \) models for the three scenarios. Model 1 (black solid line) is for the homogeneous scenario. Models 2 (red solid line) and 3 (blue solid line) are for the scenarios 2layers_A and 2layers_B, utilizing the average values of the model parameters. Model 4 (green dash line) and 5 (pink dash line) are also for the scenarios 2layers_A and 2layers_B, relating to the two layers.

Figure 5.3 Diffusivity \( (D) \) vs. CO\(_2\) saturation \( (S_n) \). Model 1: homogeneous scenario; model 2: integrated model for 2layers_A; model 3: integrated model for 2layers_B; model 4: layer with small permeability, \( k \); model 5: layer with large \( k \).
**Relationship between P-wave velocity and fluid saturation**

In order to infer CO$_2$ saturations and their temporal or lateral variation from seismic velocity tomograms or to calculate the change in seismic velocity due to a change in CO$_2$ saturation, a specific relationship between velocity and CO$_2$ saturation needs to be assumed. Usually, such a relationship is determined by lab tests of field samples [e.g., Vanorio et al., 2011]. In this study, we apply the Gassmann-Wood theory [Gassmann, 1951; Wood, 1941], which is valid at low seismic frequencies, to characterize seismic velocity in a rock that is saturated with a fluid mixture [Caspari et al., 2011b]. The Gassmann-Wood rock physics model was chosen as one possibility out of the several realistic theories for different CO$_2$ patch sizes and local CO$_2$ distribution. While the choice of the rock physics model and its calibration is critical for field studies in order to get a realistic estimation of CO$_2$ saturation, the choice of the rock physics model is somewhat arbitrary for synthetic studies. The parameters used for the calculation of the velocities are given in Table 5.1 and for the equations used, see Appendix B.

Initial velocity of the storage formation (Table 5.2) is calculated based on Equation (B-4) by assuming a CO$_2$-free reservoir ($S_n = 0$). $S_n = 0$. The parameter values are summarized in Table 1. The velocity difference $\Delta V$ is defined as the difference between the velocity before ($V_{pre}$) and after ($V_{post}$) CO$_2$ injection:

$$\Delta V = V_{post} - V_{pre} \quad (5.7)$$

Analogues to the relationship of $D - S_n$ (Figure 5.3), $\Delta V - S_n$ is plotted in Figure 5.4 for five different models (see above).

![Figure 5.4 Velocity difference ($\Delta V$) vs. CO$_2$ saturation ($S_n$). Model 1: homogeneous scenario; model 2: integrated model for 2layers_A; model 3: integrated model for 2layers_B; model 4: layer with small $k$; model 5: layer with large $k$;](image)

Similar to $D - S_n$, the relationship between $\Delta V$ and $S_n$ is not monotonic. Slight differences among these models are because of the different porosities for each of them. However, changes of $V$ along with $S_n$ are obviously smaller than those of $D$.
5.2.3 Forward simulation

Single-phase and two-phase flow simulation

An open source code, PFLOTRAN [Hammond et al., 2014], is employed for the forward simulation of the single-phase and two-phase flow processes. For the forward simulation, we focus on the early-stage injection procedure. Residual trapping, chemical reactions with the rock matrix, and any geomechanical processes are neglected. The entire simulation is divided into four stages: baseline study (stage 1), CO₂ sequestration (stage 2), shut-in period (stage 3), and repetition of interference fluid injection tests (stage 4).

Stage 1: Baseline study. At this stage, brine is injected from the bottom to the top sources in the borehole. For all the three scenarios, each injection lasts for 2 h following a recovery period of 15 h, allowing the pressures falling back to the initial hydrostatic condition. Pressure transients during the injections are used for depicting the structure of the aquifer, as well as the initial hydraulic conductivity and specific storage.

Stage 2: CO₂ sequestration. At this stage, for each scenario, CO₂ is injected over the entire aquifer with different injection durations to generate two plumes of different size. We discriminate the two injections by naming them “short-injection” and “long-injection”, respectively. The injection rate for the homogeneous scenario and 2layers_A is 0.01 kg/s, and for 2layers_B it is 0.015 kg/s. The durations of the short- and long-injections for the three scenarios are listed in Table 3.

Stage 3: Shut-in period. This stage is included for recovering the pressure to a quasi-steady state. Durations are summarized in Table 3. In applications in practice, the experiments of the next stage can be prepared now.

Stage 4: Repetition of interference fluid injections (Stage 1). As soon as the pressure recovers to a quasi-steady state, the CO₂ is injected sequentially comparably to previous brine injections. Pressure fluctuations are recorded at the observations for tomographic inversion. The injection rate for the homogeneous scenario, 2layers_A and 2layers_B is 0.02, 0.01, and 0.01 kg/s, respectively. Durations of one injection and the following recovery period are set to be equivalent (Table 5.3).

| Table 5.3 Fluid injection rate and duration of the test sequence |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| scenario          | time              | \( Q_c \) (stage 2) | \( Q_c \) (stage 4) | \( T_{inf} \) (h) | \( T_{rec} \) (h) | \( T_{s1} \) (h) | \( T_{s2} \) (h) | \( T_{s3} \) (h) |
| homogeneous       | short             | 0.01              | 0.02              | 120              | 240              | 4               | 4               | 4               | 4               | 4               | 4               |
|                   | long              |                   |                   | 360              | 240              | 6               | 6               | 6               | 6               | 6               | 6               |
| 2layers_A         | short             | 0.01              | 0.01              | 120              | 40               | 0.7             | 0.7             | 0.7             | 0.4             | 0.4             | 0.4             |
|                   | long              |                   |                   | 320              | 110              | 1.1             | 1.1             | 1.1             | 0.7             | 0.7             | 0.7             |
| 2layers_B         | short             | 0.015             | 0.01              | 90               | 40               | 0.4             | 0.4             | 0.4             | 0.7             | 0.7             | 0.7             |
|                   | long              |                   |                   | 230              | 60               | 0.8             | 0.8             | 0.8             | 1.1             | 1.1             | 1.1             |
Hydraulic travel time calculation

However, here we employ the early traveltime diagnostic for the inversion, since it resolves the preferential flowpaths better than the late diagnostic. We use 20% traveltime diagnostic ($t_{20\%}$) for the inversion as suggested by Hu et al. [2015], which refers to the time at which the pressure derivative rises to 20% of its peak amplitude. The diagnostic is then converted to “traveltime” ($\hat{t}$) as input for an eikonal solver based on the following equation

$$\hat{t} = \sqrt{6t_{a,d}} \times \sqrt{-W\left(\frac{a_d^{2/3}}{e}\right)}$$

(5.8)

where $t_{a,d}$ is $t-20\%$, and $W$ is Lambert’s $W$ function. $a_d$ is the pressure-derivative ratio term, representing the quotient of the temporal and spatial pressure derivative and the maximum pressure derivative. Note, in order to compare with seismic travel times, the phrase “hydraulic travel times” hereafter refers to the converted term $\hat{t}$, and thus the unit is $s^{0.5}$. The hydraulic travel times are perturbed with 1% of Gaussian noise prior to the inversion, which is a realistic noise level for field studies [see e.g., Ajo-Franklin et al. 2013].

Seismic travel time calculation

Based on the flow simulations discussed in section 5.2.3, seismic travel times are calculated for stage 1 (baseline) and stage 4 (after CO$_2$ sequestration). The seismic velocity distribution for the different scenarios at the two stages are calculated using the equations discussed in section 5.2.2 and the parameters of Table 5.1 and Table 5.2. The travel times are computed in the high-frequency limit using the finite-difference eikonal solver of Podvin and Lecomte [1991] on a mesh with 240 × 340 cells and a cell size of 0.25 m. Travel times between all 76 sources and 76 receivers are calculated, so that 5776 data points are available for each data set. Similar to the hydraulic travel times, all seismic travel times are also contaminated with 1% Gaussian noise before inversion.

Inversion of hydraulic travel times

The inversion scheme of the hydraulic travel times is based on Brauchler et al. [2003]. The hydraulic travel time ($\hat{t}$) is related to the reciprocal value of the square root of hydraulic diffusivity ($\sqrt{D}$) by a line integral:

$$\hat{t} = \int_{s_1}^{s_5} \frac{ds}{\sqrt{D(s)}}$$

(5.9)

A staggering technique is applied to the inversion procedure for improving the resolution of the final tomogram, and to weaken the effect of positioning [Vesnaver and Böhm, 2000; Somogyvári et al., 2016]. The base model used for the inversion is discretized by 5 columns and 4
rows. By shifting the underlying model 9 times in the horizontal direction and 3 times in the vertical direction during the inversion, the final tomogram reaches a resolution of 50×16 cells.

Repetition of the inversion to the time-lapse data sets delineates the diffusivity distribution at different times. For removing the influence of the pre-injection structure, we define the diffusivity difference ($\Delta D$) as following:

$$\Delta D = \log_{10} D_{\text{post}} - \log_{10} D_{\text{pre}} = \log_{10} \left[ \frac{c_s \mu_w \left( k_{rw} \rho_w / \mu_w + k_{rn} \rho_n / \mu_n \right)}{(S_w c_w + S_n c_n)(S_w \rho_w + S_n \rho_n)} \right]$$

(5.10)

Note, the unit of $\Delta D$ is non-dimensional, since the right side of Equation (5.10) can be formulated as $\log_{10} \left( \frac{D_{\text{post}}}{D_{\text{pre}}} \right)$, which can be considered as the logarithm of the normalized diffusivity. By assuming the state variables (the density, viscosity, and compressibility) of the brine and CO$_2$ are constant, Equation (5.10) shows that the $\Delta D$ is merely determined by the fluid saturation and the pore size distribution ($\lambda$). In the supporting information, Figure 5.10 presents the change of $\Delta D$ along with the CO$_2$ saturation $S_n$. The model of the homogeneous scenario shows a slight difference from the other four models due to their varying $\lambda$ values (model for the homogeneous scenario: $\lambda = 0.76$; the other four models: $\lambda = 1.5$).

**Inversion of seismic travel times**

The seismic travel time data sets that are simulated for the different scenarios and at the different times (stage 1 and stage 4) are inverted using the algorithm of Doetsch et al. [2010b]. It implements an Occam’s type inversion with stochastic regularization (chosen integral scales of $I_x = 16$ m, $I_z = 8$ m), in which the regularization strength is decreased until the data is fit to the error level. The assumed error on the travel times is 1%, in accordance with the noise contamination level, and all inversion results fit the data to that assumed error level.

The inversion results of the baseline inversion are used as starting and reference models for the time-lapse inversions of data acquired during the multilevel CO$_2$ injections. A difference inversion scheme is being used that inverts for the changes to the baseline inversion result [Doetsch et al., 2010a], so that even small changes to the baseline model can be resolved and inversion artifacts are minimal. The results of the time-lapse inversions are analyzed and shown as change in velocity compared to the baseline inversion result.

**5.2.4 Clustering and zonal calibration**

**Time-lapse 1-D and 2-D clustering**

The goal of clustering is to determine the baseline structure and to delineate the extent of CO$_2$ plumes at different times. The unlabeled data of inverted tomograms are partitioned into
homogeneous groups where the hydraulic conductivity and specific storage are constant. We first apply individual clustering (i.e., 1-D clustering) to the diffusivity tomograms of the CO$_2$-free aquifer. Spatial heterogeneity is determined by the clusters, which are used in the following calibration. Individual and joint clustering (i.e., 2-D clustering) are employed to the tomograms of the diffusivity difference and velocity difference obtained after CO$_2$ injection.

The clustering approach is a modified k-means method. Unlike the usual k-means method [e.g., McQueen, 1967], the centroids of the clusters are determined by fitting the data histogram with the summation of multiple 1-D or 2-D Gaussian functions. Both 1-D and 2-D clustering are performed in a time-lapse strategy. The centroids are determined by the data histogram at long-injection, and then applied at short-injection as well. The inverted shapes are compared to the “true” plumes and their ambient aquifers by analyzing the dissimilarity in their shape. Three metrics can be utilized for obtaining this goal: overestimation rate, underestimation rate, and total pixel misclassification error rate. These metrics can loosely depict the dissimilarity of two binary images, which in our study are the “true” plume and inverted plume. They are detection performance measures, regardless of the pixel positions and intensity. Definitions of these metrics are given in the following [Baddeley, 1992].

Let $A$, $B$ and $X$ be the “true” plume, inverted plume and the pixel raster (i.e., the entire inversion model), respectively. Pixels that belong to $B$ but not $A$ are called Type I errors; Pixels that belong to $A$ but not $B$ are called Type II errors.

Type I error ($\alpha$), or also called “overestimation rate”, is calculated by:

$$\alpha(A,B) = \frac{n(B \setminus A)}{n(X \setminus A)} \quad (5.11)$$

Type II error, or “underestimation rate”, is calculated by:

$$\beta(A,B) = \frac{n(A \setminus B)}{n(A)} \quad (5.12)$$

where $n(A)$ and $n(B)$ are the number of pixels in the “true” plume $A$ and inverted plume $B$. $n(X)$ is the total number of the pixels in the raster. Based on these two error rates, the total pixel misclassification error rate ($\varepsilon$) for binary images is defined as:

$$\varepsilon(A,B) = \frac{n(A \Delta B)}{n(X)} = \alpha(1-r) + \beta r \quad (5.13)$$

Zonal calibration and estimation of saturation

For reducing the model complexity and accelerating calibration, we use MODFLOW [Harbaugh, 2005] to run the forward simulation during the inversion procedure, considering CO$_2$
and brine as a mixed phase. Figure 5.5 presents a summary of two different zonal calibration steps carried out sequentially:

First, for the pre-injection, we merely calibrate the effective hydraulic conductivity and specific storage assuming the aquifer is homogeneous (Figure 5.5a). The calibrated values of $K_w$ and $S_w$ are assigned to the clustered aquifer zone for the post-injection (Figure 5.5b). As pointed out in Hu et al. [2015], the mixed-phase conductivity $K$ of the plume zone shows only minor changes during CO$_2$ injection, and thus this value is kept the same as $K_w$ for the post-injection. Only the mean mixed-phase specific storage $S_s$ of the plume zone needs to be calibrated (Figure 5.5b). The mean $S_s$ of the plume is then transferred from $S_s$ (Figure 5.5c) by the following equation:

$$S_s = \frac{f}{g} \left( \rho_s S_w + \rho_w S_s \right) \left( c_w S_w + c_s S_s \right)$$  \hspace{1cm} (5.14)

Second, the calibration is conducted considering the heterogeneity of the reservoir. For the two layered scenarios, the aquifer structure is determined by clustering the inverted $D_{pre}$, and the $K_w$ and $S_w$ of each cluster are calibrated primarily for the pre-injection (Figure 5.5d). For the post-injection, the plume zone can be divided into two secondary plumes based on the aquifer structure. $K_w$ and $S_w$ of each cluster in the aquifer zone are fixed, as well as $K$ in the plume zone ($K = K_w$) (Figure 5.5e). The mixed-phase specific storage $S_s$ of the two secondary plumes are calibrated and converted to the $S_s$ of the two plumes by Equation (5.14).

In order to reduce the potential non-uniqueness of the calibration results in the second step, the acquired $S_s$ in the first step are utilized as prior information for the following calibration. Here we make an assumption that the $S_s$ in the high-permeability layer is larger than the low-permeability layer. The mean $S_s$ (Figure 5.5c) is transferred back to the two $S_s$ values for the layered structure, considering the porosities are different for the two layers. These two values are used for constraining the calibration. For instance, presuming $K_{w,1}$ is larger than $K_{w,2}$ in Figure 5.5d, parameters $S_s$ of the two layers transferred by the mean $S_s$ are expressed as $\bar{S}_{s,1}$ and $\bar{S}_{s,2}$, respectively. Then the calibrated $S_{s,1}$ should be larger than $\bar{S}_{s,1}$, and $S_{s,2}$ is smaller than $\bar{S}_{s,2}$. The validity of this assumption depends on whether the initial aquifer structure can be identified properly.
Figure 5.5 Conceptual scheme for zonal calibration. Parameters in red color are calibrated, and those in blue color are fixed during the calibration. Parameters in green color are converted by the specific storage-saturation models.

Figure 5.6 shows five different models for converting $S_e$ back to $S_n$. The models 1 to 3 (black, blue and red solid lines) show the relationship between the effective $S_e$ and $S_n$ for 1-plume structure in three different scenarios. Models 4 and 5 (green and pink dash lines) are used for obtaining the saturation within each secondary plume for the two layered scenarios.

Figure 5.6 Mixed-phase specific storage ($S_i$) vs. CO$_2$ saturation ($S_n$). Model 1: homogeneous scenario; model 2: integrated model for 2layers_A; model 3: integrated model for 2layers_B; model 4: layer with small $k$; model 5: layer with large $k$.

The performance of the calibration is evaluated by calculating the saturation error ($\zeta$). It is estimated by the difference between the calculated saturation ($S_n^{\text{cal}}$) and the arithmetic mean of the “true” saturations within the same inverted plume or secondary plume structure (note, not the “true” plume extent) normalized by $S_n^{\text{true}}$:
\[ \xi(\%) = \frac{S_{\text{cal}} - S_{\text{true}}}{S_{\text{true}}} \times 100 \]  

(5.15)

5.3 Results

The procedure described above is tested step-by-step on the three reservoir scenarios in order to compare the performance of pressure (PT) and seismic tomography (ST).

5.3.1 PT and ST travel times

For eliminating the effect of different source-receiver configurations for PT and ST, we only compare the travel times of the tomographic arrays in the horizontal direction. Note, the seismic inversion is based on the full data set of the travel times (i.e., the total 5776 travel times). Table 5.4 (see the supporting information) lists the statistics of these travel times (noise-free and with noise) in the horizontal direction derived for PT and ST. The horizontal hydraulic travel times vary from 12.7 s\(^{0.5}\) (in scenario 2layers_B) to 173.4 s\(^{0.5}\) (in the homogeneous scenario). The hydraulic travel times increase by 154% to 486% after CO\(_2\) injection. The most effected horizontal seismic travel times through the reservoir increase from 11.2 ms before CO\(_2\) injection to 11.76 ms after injection, corresponding to an increase of 5%. Overall, changes of the seismic travel times are much smaller than for hydraulic travel times, with changes ranging from 0% to 5%.

In order to compare the results for PT and ST, the relative spread of the horizontal travel times (i.e., the standard derivation normalized by the mean) is examined (see Table 5.4). At pre-injection, the relative spread of the noise-free travel times in the homogeneous scenario is 0. For the travel times with noise, the relative spread is consistent with the noise level (1%). In the two layered scenarios, the relative spread at pre-injection of the hydraulic travel times (0.23 for 2layers_A and 0.1 for 2layers_B) is much greater than the seismic travel times (around 0.01 to 0.02). At post-injection, the relative spread for the seismic travel times is around 0.01 to 0.02 for all the three scenarios, while for the hydraulic travel times, it reaches from 0.1 to 0.7. In the following, we only present and discuss the results derived from the data with noise, thus considering the more realistic cases.

5.3.2 Diffusivity and velocity tomograms

The results obtained from two-phase flow simulation for the different scenarios are considered as the “truth”, and these serve as a reference for assessing tomographic inversion. The values of “true” diffusivity, velocity and their differences for pre-, short-, and long-injections are calculated based on the simulated “true” CO\(_2\) saturation, according to Equations (5.3) to (5.7), and (5.10). Both true profiles and inverted tomograms are depicted in Figure 5.7, and a complete list can be found in the supporting information (Table 5.5 and Table 5.6).
In the homogeneous scenario, the true $D$ and $V$ at pre-injection are 2.7 m$^2$/s and 4465 m/s, respectively (Figure 5.7I-a and Figure 5.7I-g). The inverted $D$ varies from 2.8 to 3 m$^2$/s (Figure 5.7I-d), which is slightly different to previous results by Hu et al. (2015) for the same scenario due to the noise for hydraulic traveltimes that is included here. The seismic tomography includes the caprock above and below the reservoir, and the reservoir can be clearly identified in the tomograms (see Appendix C, Figure C-1). Inside the reservoir, the inverted $V$ range from 4064 to 4755 m/s (Figure 5.7I-j). The relative spread is greater than that observed for the inverted $D$. The inverted velocity in caprock/bottom seal is much less than the reservoir, which varies from 3413 to 3800 m/s. In the scenarios 2layers_A and 2layers_B, the true $D$ and $V$ of the initial CO$_2$-free formation are 22.4 m$^2$/s and 4579 m/s for the high permeability layer, and 2.9 m$^2$/s and 4454 m/s for the low permeability layer (Figures 5.7II-a, Figure 5.7II-g, Figure 5.7III-a and Figure 5.7III-g). The inverted $D$ for 2layers_A and 2layers_B is within a range smaller than the true values (3.6 to 10.2 m$^2$/s for 2layers_A and 7.2 to 12.2 m$^2$/s for 2layers_B, Figures 5.7II-d and Figure 5.7III-d). The inverted $V$ shows a similar range for the two scenarios, which is 3998 to 4930 m/s and 4067 to 4854 m/s, respectively (Figures 5.7II-j and Figure 5.7III-j).

For the two post-injection periods, both of the true $D$ and $V$ follow a non-monotonic change along with $S_n$ (Figure 5.3 and Figure 5.4). The true $D$ decreases by up to two orders of magnitude, resulting in a minimum $\Delta D$ (i.e., the difference of logarithm $D$ at pre- and post-injection) value of around -2 for all three scenarios (Figure 5.7I-b and Figure 5.7I-c, Figure 5.7II-b and Figure 5.7II-c, Figure 5.7III-b and Figure 5.7III-c). The minimum $\Delta V$ is -340.9 m/s for the homogeneous scenario, and it is -356.4 m/s for 2layers_A and 2layers_B. It is noticeable that for each scenario, the minimum $\Delta D$ and $\Delta V$ are the same during the short- and long-injection since they do not correspond to the maximum $S_n$ value (Figure 5.4 and Figure 5.7). Likewise, in the true profiles of the flow and seismic parameters, the smallest values are shown within the plume where $S_n$ has a moderate value. The inverted $\Delta D$ for the homogeneous scenario and 2layers_A span a smaller range compared to the truth (Figure 5.7I-e and Figure 5.7I-f, Figure 5.7II-e and Figure 5.7II-f). In contrast, the inverted $\Delta D$ has a larger range than the truth for 2layers_B (Figure 5.7III-e and Figure 5.7III-f). The absolute values of the inverted $\Delta V$ are in general smaller than the true values for all the three scenarios.
Figure 5.7 “true” profiles (subgraphs a to f) vs. inverted tomograms (subgraphs g to l) in the three scenarios. For ST, the model extends 30 m above and below the reservoir (not shown) to include the caprock and bottom seal ($V=3500$ m/s). CO$_2$ is injected at the left of the model and the migrating plume can be seen in both PT and ST by the inverted tomograms. The internal structure of the aquifer can only be resolved using PT, because internal variations in seismic velocity are too small compared with the contrast to caprock.

5.3.3 1-D and 2-D clustering structure

Prior to the zonal calibration, clustering was implemented based on the inverted results to obtain the plume shape at different times. First, according to the inversion performance of the baseline, the inverted $D$ derived prior to CO$_2$ injection was clustered to determine the structure of the reservoir. Subsequently, the inverted $\Delta D$ and $\Delta V$ were clustered individually (1-D clustering), and then jointly (2-D clustering). In order to judge the different approaches, the plume
shapes derived from both 1-D and 2-D clustering processes were compared to the “true” plumes according to the three aforementioned metrics (overestimation rate, underestimation rate, and total misclassification rate, calculated by Equations (5.11) to (5.13)). Figure 5.8 depicts the clustering results, with the metrics shown as the numbers in the same figure. In terms of misclassification, ST outperforms both PT and JT (i.e., joint clustering), but the numbers are generally similar.

![Diagram](image)

Figure 5.8 “true” clustering structures (subgraphs a to c) vs. 1-D (subgraphs d to i) and 2-D (subgraphs j to l) clustering structures in the three scenarios. Numbers in blue color: overestimation rate (α); numbers in green color: underestimation rate (β); numbers in red color: total misclassification rate (ε). The performance among PT, ST and JT is comparable in the homogeneous scenario and 2layers_A, while in 2layers_B, JT shows a combination of the 1-D clustering results from both PT and ST.

### 5.3.4 Zonal calibration and calculated saturations

Based on the identified aquifer and plume zones by 1-D and 2-D clustering, zonal calibration was conducted. The effective $K_w$ and $S_{rw}$ of the original formation (ignoring its heterogeneity) were calibrated primarily in MODFLOW with PEST (Doherty, 2010), using the pressure observations derived from the full model. Because of the proxy, the calibrated $K_w$ and $S_{rw}$ at pre-
injection are slightly smaller than the true values (the errors are around 10%). The calibrated $K_w$ and $S_{sw}$ were then used as the prior information for the aquifer zone, and the calibrated $K_w$ for the plume zone was also fixed for the post-calibration. The calibrated $S_e$ of the plume zone (see supporting information, Figure 5.8) was then converted to $S_n$ in the homogeneous scenario, as well as 2layers_A and 2layers_B by model 1, 2, and 3 (Figure 5.6), respectively. The calibration results of the 1-plume and 2-plume structures for the three scenarios are presented in Figure 5.9, Figure 5.12, Table 5.7 and Table 5.9, respectively.

Calibration quality was evaluated by calculating the error of $S_n$ ($\xi$) using Equation (5.15) (see Table 5.8). For 1-plume structure, in the homogeneous scenario, $\xi$ varies from -3% to 11% except for the short-injection case in which the plume structure was derived from ST ($\xi$ is 89%). In scenario 2layers_A, $\xi$ ranges from 36% to 51% for short-injection, which is generally higher than for long-injection (23%-29%). In scenario 2layers_B, $\xi$ remains low (<10%) except for the case of PT for the long injection (46%). Overall, the joint clustering results are not always the best, but they are robust. The estimation errors are reduced, maintaining errors below 36% for all the scenarios. For 2-plume structure, $\xi$ changes from -24% to 127% in 2layers_A, and which varies from 0% to 63% in 2layers_B (see Table 5.9).

![Figure 5.9 “true” vs. calibrated saturations of three scenarios. The white dot line indicates the boundary of two “true” or inverted layer boundaries. Calibrated CO$_2$ plume saturations are assumed homogeneous for each calibration and thus represent an average value within the plume.](image-url)
5.4 Discussion

5.4.1 PT and ST travel times

Changes in PT and ST travel times are a direct measure of the sensitivity to relevant changes in the reservoir. Results indicate that variability of hydraulic travel times is generally much larger than seismic travel times. A great variability of travel times is favorable, as this potentially allows for better resolution of the subsurface. Through the comparison of relative spread in three scenarios, it is implied that relative spread correlates with the degree of heterogeneity. No relative spread indicates noise-free homogeneous conditions. Prior to CO₂ injection, the highest relative spread is obtained for 2layers_A, where small-scale contrasts in permeability are simulated by a relatively thin conductive layer. After CO₂ injection, the highest relative spread is in 2layers_B, indicating the largest contrasts in permeability between the relatively thick conductive layer and the CO₂ plume. In addition, it is remarkable that the imposed noise has a small impact on the relative spread of the hydraulic travel times, since the spread of the noise free data is much larger than the noise level.

5.4.2 Diffusivity and velocity tomograms

In Figure 5.7, two-phase simulations show that the fronts of the plumes in the two heterogeneous scenarios have more complicated geometries (Figure 5.7II-b, Figure 5.7II-c, Figure 5.7III-b and Figure 5.7III-c) in comparison to the true plumes in the homogeneous scenario (Figure 5.7I-b and Figure 5.7I-c). The high permeability layer largely controls the plume. CO₂ migrates preferentially within the highly conductive layer, while the migration is also controlled by buoyancy, complicating the plume geometry. In 2layers_A, the plume distribution is representative for a multi-layer system (i.e., two or more continuous layers), in which the top layer has the highest permeability. Here the plume ultimately assembles at the top of the reservoir, which is caused by a combined effect of the high permeability of the upper layer and buoyancy (Figure 5.7II-b and Figure 5.7II-c). However, the less permeable lower layer hinders expansion of the plume. The scenario 2layers_B exemplifies conditions where a highly conductive channel exists between the two wells. Here, the plume travels faster at the bottom layer, forming a striking finger-like shape at the boundary of the two layers (Figure 5.7III-b and Figure 5.7III-c). The finger is not delimited strictly below the boundary because of buoyancy effects.

At pre-injection, the small range of inverted $D$ still nicely reflects the homogeneous properties of the aquifer in homogeneous scenario. The strong velocity variation (3500 m/s versus ~4500 m/s) between caprock/bottom seal and reservoir enables identification of the reservoir (Appendix B, Figure B-1), but makes it difficult to judge if the reservoir itself is homogeneous or heterogeneous. In the two layered scenarios, the inverted $D$ generally has a higher value and less data spread in 2layers_B, which, in this scenario, is due to the larger high permeability area and also to shorter travel times. In addition, the inverted $D$ tomograms in the two scenarios display a layered distribution, which is consistent with the true aquifer structure to some extent. However,
the “perfectly” horizontal boundary between the two layers was not accurately reconstructed due to non-horizontal tomographic rays as well as regularization, which both cause smearing between the inversion cells. Comparison of the three scenarios indicates that PT resolves the internal structure and especially the hydraulic properties of the aquifer better, since it is related to both permeability and porosity (Equations (5.3) to (5.5)). Usually, permeability shows much larger spatial variability than porosity. In contrast, ST is able to delineate the structure of the reservoir, but fails to identify additional variations within the reservoir. Seismic velocity mainly depends on the porosity and rock density, which might only have slight variations within an aquifer. Velocity variations within the reservoir are only about 120 m/s prior to CO$_2$ injection, which is difficult to recover simultaneously with the 1000 m/s variation between reservoir and caprock/bottom seal.

At post-injection, the true profiles show that the contrast of $\Delta V$ caused by CO$_2$ is much smaller compared to $\Delta D$. Inversion results of $\Delta D$ and $\Delta V$ are not consistent with the true values. Nevertheless, the information about the plume can still be inferred by the small values in the $\Delta D$ and $\Delta V$ tomograms. The $\Delta D$ tomograms of the two layered scenarios indicate that the variability of permeability has an adverse impact as PT is applied to the post-injection. PT resolves the secondary plume in the lower permeability layer better. In the case that the contrast of diffusivity is sufficiently large at pre-injection, the secondary plume in higher permeability layer can still be identified (e.g., in 2layers_A, Figure 5.7II-e and Figure 5.7II-f). Conversely, in 2layers_B, the finger at the layer boundary is masked in the $\Delta D$ tomograms (Figure 5.7III-e and Figure 5.7III-f). This is mainly because the variations of the inverted $D$ at pre-injection are comparably small. ST can capture the main front of the plume from the $\Delta V$ tomograms of all the scenarios, since it is not influenced by the permeability. However, the less relative spread of the seismic traveltimes limits the capability of ST to identify the small-size plumes in the low permeability layer (Figure 5.7III-k and Figure 5.7III-l).

Overall, the inverted values from the three scenarios at both pre- and post-injection indicate that neither diffusivity nor velocity values can be precisely reproduced by the inversions. Direct transformation of inverted values to CO$_2$ saturation leads to an incorrect estimation. There are several reasons that can explain this. First, for both PT and ST, the loose density of the trajectories or rays in the low diffusivity or velocity parts masks the small content in the tomograms. Second, these errors can also be attributed to inaccuracies introduced by using the single-phase proxy.

### 5.4.3 1-D and 2-D clustering structure

In the homogeneous and 2layers_A scenarios, clustering of $D$ before CO$_2$ injection shows a homogenous and two-layer aquifer structure (Figure 5.8I-d and Figure 5.8II-d). The clustering results after CO$_2$ injection from both 1-D and 2-D clustering are of comparable quality. Because of the similar distribution of the inverted $\Delta D$ and $\Delta V$, they show a strong correlation (see supporting information, Figure 5.11a). Even 2-D clustering does not significantly improve the results. In some cases, the 1-D clustering results based only on PT or ST are better than the 2-D clustering results. For instance, if we compare the total misclassification rate (Figure 5.8, red
numbers) to assess the quality of the results, in the homogeneous scenario, the clustering result based on PT at short-injection provides the best agreement with the “true” plume (Figure 5.8I-e). On the contrary, the clustering result from ST (Figure 5.8I-h) is the worst, thus it has a negative impact on the final joint clustering result (Figure 8I-k). Additionally, in 2layers_A, during the short-injection, the clustered ΔD and ΔV display a discontinuous distribution near the injection location (Figure 5.8II-e and Figure 5.8II-h), yet this continuity vanishes in the jointly clustered plume (Figure 5.8II-k).

In 2layers_B, the two-layer structure also can be indicated by the 1-D clustering of D (Figure 5.8III-d). However, for the two post-injection periods, the plumes show a significant difference from the 1-D clustering results. The geometries of the plumes delineated by PT are more vertical (Figure 5.8III-e and Figure 5.8III-f), while those derived from ST show a more lateral distribution (Figure 5.8III-h and Figure 5.8III-i). ΔD and ΔV show less correlation during the 2-D clustering process (see supporting information, Figure 5.11b). The dissimilarities between ΔD and ΔV hamper the acquisition of the 2-D centroids for clustering the time-lapse data sets. Therefore, we cluster ΔD and ΔV in another way. The 2-D histogram was fitted by a model composed of multiple Gaussian functions (the aquifer zone) and a uniform distribution (the plume zone). The cutoff of the aquifer and plume zones was at the edge of the Gaussian functions where the values are equal to the mean value of the uniform distribution. This was applied for both the short- and long-injection runs. Consequently, the 2-D clustered structures are deemed to be a superposition of the plumes from both PT and ST (Figure 5.8III-k and Figure 5.8III-l), but in a more systematic way.

Clustering works best for the homogeneous scenario according to the total misclassification rate ε (smaller than 0.1 for all the three scenarios). The heterogeneous scenarios show higher misclassification due to the additional complexity. Values of ε indicate that the clustering performance in 2layers_B (ε : 0.07-0.15) is better than 2layers_A (ε : 0.13-0.18). The overestimation rate α shows a similar trend as ε. The plume extents are most overestimated in 2layers_A compared to the other scenarios. The underestimation rate β of the plumes is relatively high as the true plumes near the injection well or if the plume fronts are not accurately characterized (e.g., Figure 5.8I-h and Figure 5.8III-f). In general, joint clustering in general reduces the underestimation of the plume extent (e.g., Figure 5.8I-k, Figure 5.8II-k, Figure 5.8III-k and Figure 5.8III-l).

### 5.4.4 Saturation errors

The underestimation rate β is considered the most crucial criterion to assess proper spatial classification. In Figure 5.10, the saturation error ξ is plotted with β. It is clearly shown that, for each scenario, that increased β generally provokes a higher saturation error. This is because the hydraulic conductivity and specific storage in the underestimated part of the plume is assigned the same values as the aquifer. As discussed in section “Model parameters”, the specific storage of a plume can be around 1-2 orders of magnitude larger than the ambient aquifer. Therefore, the
underestimation of the plume size leads to a higher value of specific storage within the inverted plume.

In comparison with homogeneous and 2layers_B, 2layers_A shows higher estimation errors. This might be due to the overestimation of the plume extent. The “true” saturation $S_n^{true}$ is derived by averaging the saturations within an inverted plume, that is, it is the arithmetic mean of these saturations. As $S_f$ increases non-linearly with $S_n$ (Figure 5.6), the transferred $S_n^{cal}$ might be different from the averaged $S_n^{true}$, even though they correspond to the same calibrated $S_n$.

For 2-plume structures in 2layers_A and 2layers_B, in general, the estimated $S_n$ values are consistent with the fact that $S_n$ is larger in the high permeability layer. However, the estimation errors span a broader range in comparison to the previous results from the 1-plume structure. As discussed above, these errors can be due to the misclassification of the secondary plume in each layer. Moreover, they can be also attributed to the pressure discrepancy between the full model and the proxy. In order to obtain the specific storage of each secondary plume, more pressure measurements were used for the 2-plume structure than the 1-plume structure, and thus the calibration involved more pressure errors. This can be improved in the future work by quantifying the errors between the two-phase forward simulation and the single-phase proxy under different conditions.

5.5 Conclusions

We investigate the feasibility of pressure tomography (PT) and compare the inversion performance with cross-hole seismic tomography (ST) for homogeneous and heterogeneous reservoirs. Relations between the inverted parameters (the mixed diffusivity and P-wave velocity) and the CO$_2$ saturation as used by these two methods are comparable. Both are fast and computationally efficient as they are eikonal-based. However, since different signals are processed, these two approaches can be complementary to each other for characterizing an evolving CO$_2$ plume shape and for evaluating the CO$_2$ saturation.

In our scenarios, the upper and lower boundary of the reservoir can only be detected using ST. PT cannot be used in the impermeable caprock. ST is less suitable to resolve the smaller internal contrasts in the layered reservoir. Better results when reconstructing the heterogeneity of the reservoir can be obtained by PT as it directly links to aquifer permeability. The capability for resolving the plume shape is distinct for PT and ST due to the different features of the travel times. First of all, values and ranges of hydraulic travel times are much larger than those of seismic travel times. This gives PT a better sensitivity to the CO$_2$ plume. However, for PT, reservoir heterogeneity can alleviate the diffusivity contrast caused by CO$_2$ injection, and thus the front of the plume is hard to delineate by the diffusivity difference tomogram. ST can better resolve the lateral spreading part of the plume. Consequently, best results are generally derived from the presented joint inversion.
By clustering and subsequent zonal calibration, the mean saturation of the plume can be determined. We demonstrate by sequential calibration strategies how the saturation of different layers can be distinguished. Again, jointly clustered structures provide best results for the various scenarios and conditions examined. However, it is not surprising that improper spatial delineation of the plume makes it difficult to properly estimate the saturation.

This study provides an insight into the capability of PT for application in heterogeneous formations, and its potential for complementing the geophysical approaches. One crucial point remains the transfer of this approach to the field. The main challenges of a field application include the technical implementation of sources and receivers in deep reservoirs, conducting interference injection tests during the course of CO\textsubscript{2} injection, and interpreting PT or joint PT-ST results given non-ideal conditions in the field. From the field injection tests in several CO\textsubscript{2} storage sites (e.g., Ketzin and Cranfield), it is clear that it is possible to conduct multilevel CO\textsubscript{2} injection tests and to obtain useful pressure signals. Application of such tests in tomographic arrangements to complement seismic measurements is thus a promising area for future study. Also, more geophysical approaches used for CO\textsubscript{2} sequestration will be considered for comparing and being combined with pressure tomography, such as seismic full waveform inversion and electrical resistance tomography.

5.6 Appendices

5.6.1 Appendix A. Discretization of two-phase flow simulation model

![Figure A-1. Discretization of two-phase flow simulation model. a) Full model (Not to scale). b) Zoom-in area between injector and responder.](image-url)
5.6.2 Appendix B. Gassmann-Wood rock physics model

The low-frequency Gassmann equations (Gassmann, 1951) are widely used for calculating rock and fluid elastic properties in fully saturated media. Saturated bulk modulus \( G_{\text{sat}} \) is given by

\[
G_{\text{sat}} = G_{\text{dry}} + \frac{\left(1 - \frac{G_{\text{dry}}}{G_m}\right)^2}{\phi + \frac{1 - \phi}{G_f}} \quad \text{(B-1)}
\]

where \( G_{\text{dry}} \) and \( G_m \) are the bulk modulus of dry frame rock (drained of pore fluid) and rock matrix. \( G_f \) is the bulk modulus of the pore fluid, which can be single-phase or multiphase.

Bulk modulus of mixing pore fluid can be calculated by Wood’s equation (Wood, 1941)

\[
G_f = \left(\frac{S_w}{G_w} + \frac{S_n}{G_n}\right)^{-1} \quad \text{(B-2)}
\]

where \( G_w \) and \( G_n \) are the bulk modulus of brine and CO\(_2\), respectively. \( G_n \) equals the product of the density and P-wave velocity of CO\(_2\). The P-wave velocity in saturated rock is then estimated by the following equations

\[
V = \sqrt{\frac{G_{\text{sat}} + \frac{4}{3} N_{\text{sat}}}{\rho_r}} \quad \text{(B-3)}
\]

Where \( \rho_r \) is rock density. For a CO\(_2\)-brine system, it is calculated through the linear relationship

\[
r_r = f \left( S_w r_w + S_n r_n \right) + (1 - f) r_m \quad \text{(B-4)}
\]

\( r_w \) and \( r_n \) are brine and CO\(_2\) densities, and \( r_m \) is rock matrix density.

5.6.3 Appendix C. Full velocity (difference) tomograms

Figure C-1 shows the full velocity or velocity difference tomograms derived from ST inversion (an example of homogeneous scenario, corresponding to Figures 7I-j to 7I-l). The grid size of inversion model is 1 m \( \times \) 1 m. The dashed line delineates the reconstructed reservoir. Note that, the results depicted here are slightly different in comparison to Figure 7. This is because the model grid shown in Figure 7 (ST results) is interpolated to 1 m \( \times \) 0.9375 m, to be consistent with the PT results and to implement the joint clustering.
Figure C-1. Full velocity (pre-injection) and velocity difference (post-injection) tomograms in homogeneous scenario. The area outlined by dashed line indicates the reconstructed storage reservoir. Note that sources and receivers are illustrated schematically, and their true numbers are much higher (76).

5.7 Supporting information

5.7.1 Introduction

In the following supporting information, we provided additional information for complementing the manuscript. Table 5.4 presents a comparison hydraulic and seismic travel times, which are derived from the horizontal tomographic arrays in the storage reservoir.

Table 5.5 and Table 5.6 show the travel-time inversion results of pressure tomography and seismic tomography. Table 5.7 compares the “true” and calibrated hydraulic conductivity and specific storage of the initial reservoir.

Table 5.8 and Table 5.9 are the results of zonal calibration. Figure 5.10 shows the relationship between diffusivity difference and CO₂ saturation. Figure 5.11 presents the correlations between the mixed-phase diffusivity difference and velocity difference. Figure 5.12 provides the calibrated CO₂ saturation of the inverted plumes in the two layered scenarios.
### Table 5.4 Travel times statistics for PT and ST

<table>
<thead>
<tr>
<th>scenario</th>
<th>time</th>
<th>pressure tomography (PT)</th>
<th>seismic tomography (ST)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$t_{\text{min}}$ (s$^{0.5}$)</td>
<td>$t_{\text{max}}$ (s$^{0.5}$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>noise-free</td>
<td>noise-free</td>
</tr>
<tr>
<td>homogeneous</td>
<td>pre</td>
<td>29.6</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td>short</td>
<td>42.1</td>
<td>126.3</td>
</tr>
<tr>
<td></td>
<td>long</td>
<td>114.9</td>
<td>171.8</td>
</tr>
<tr>
<td>2layers_A</td>
<td>pre</td>
<td>15.3</td>
<td>25.2</td>
</tr>
<tr>
<td>2layers_B</td>
<td>short</td>
<td>40.8</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>long</td>
<td>50.5</td>
<td>81.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1% noise</td>
<td></td>
</tr>
<tr>
<td>homogeneous</td>
<td>pre</td>
<td>29.2</td>
<td>29.6</td>
</tr>
<tr>
<td></td>
<td>short</td>
<td>43.1</td>
<td>125.8</td>
</tr>
<tr>
<td></td>
<td>long</td>
<td>114.5</td>
<td>173.4</td>
</tr>
<tr>
<td>2layers_A</td>
<td>pre</td>
<td>15.4</td>
<td>25.4</td>
</tr>
<tr>
<td>2layers_B</td>
<td>short</td>
<td>40.8</td>
<td>64.8</td>
</tr>
<tr>
<td></td>
<td>long</td>
<td>50.2</td>
<td>81.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2layers_A</td>
<td></td>
</tr>
<tr>
<td>homogeneous</td>
<td>pre</td>
<td>14.6</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>short</td>
<td>12.6</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>long</td>
<td>14.2</td>
<td>93.5</td>
</tr>
</tbody>
</table>

### Table 5.5. Summary of “true” and inverted diffusivity/diffusivity difference for the three scenarios

<table>
<thead>
<tr>
<th>scenario</th>
<th>time</th>
<th>“true” $S_u$ (-)</th>
<th>“true” $D$ (m$^2$/s)</th>
<th>inverted $D$ (m$^2$/s)</th>
<th>“true” $\Delta D$ (-)</th>
<th>inverted $\Delta D$ (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pre-injection</td>
<td>0</td>
<td>2.7</td>
<td>2.8~3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>homogeneous</td>
<td>short-injection</td>
<td>0~0.6</td>
<td>0.03~2.7</td>
<td>0.01~1.7</td>
<td>-2~0</td>
<td>-2.4~0.2</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0~0.67</td>
<td>0.03~2.7</td>
<td>0.02~0.25</td>
<td>-2~0</td>
<td>-2.2~1.1</td>
</tr>
<tr>
<td></td>
<td>pre-injection</td>
<td>0</td>
<td>22.1, 2.9</td>
<td>3.6~10.2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2layers_A</td>
<td>short-injection</td>
<td>0~0.72</td>
<td>0.03~22.4</td>
<td>0.03~3</td>
<td>-2~0</td>
<td>-2.1~0.3</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0~0.8</td>
<td>0.03~22.4</td>
<td>0.03~1.7</td>
<td>-2~0</td>
<td>-2.1~0.6</td>
</tr>
<tr>
<td></td>
<td>pre-injection</td>
<td>0</td>
<td>2.9, 22.1</td>
<td>7.2~12.2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2layers_B</td>
<td>short-injection</td>
<td>0~0.68</td>
<td>0.03~22.4</td>
<td>0.01~31.1</td>
<td>-2~0</td>
<td>-2.7~0.47</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0~0.72</td>
<td>0.03~22.4</td>
<td>0.01~30.5</td>
<td>-2~0</td>
<td>-2.8~0.43</td>
</tr>
</tbody>
</table>
Table 5.6 Summary of “true” and inverted velocity/velocity difference for the three scenarios

<table>
<thead>
<tr>
<th>scenario</th>
<th>time</th>
<th>“true”</th>
<th>inverted</th>
<th>“true”</th>
<th>inverted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$S_n$</td>
<td>$V$</td>
<td>$\Delta V$</td>
<td>$\Delta V$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-)</td>
<td>(m/s)</td>
<td>(m/s)</td>
<td>(m/s)</td>
</tr>
<tr>
<td>homogeneous</td>
<td>pre-injection</td>
<td>0</td>
<td>4465</td>
<td>4064-4755</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>short-injection</td>
<td>0-0.6</td>
<td>4124-4465</td>
<td>-</td>
<td>-340.9-0</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0-0.67</td>
<td>4124-4465</td>
<td>-</td>
<td>-340.9-0</td>
</tr>
<tr>
<td>2layers_A</td>
<td>pre-injection</td>
<td>0</td>
<td>4579, 4454</td>
<td>3998-4930</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>short-injection</td>
<td>0-0.72</td>
<td>4097-4579</td>
<td>-</td>
<td>-356.4-0</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0-0.8</td>
<td>4097-4579</td>
<td>-</td>
<td>-356.4-0</td>
</tr>
<tr>
<td>2layers_B</td>
<td>pre-injection</td>
<td>0</td>
<td>4454, 4579</td>
<td>4067-4854</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>short-injection</td>
<td>0-0.68</td>
<td>4097-4579</td>
<td>-</td>
<td>-356.4-0</td>
</tr>
<tr>
<td></td>
<td>long-injection</td>
<td>0-0.72</td>
<td>4097-4579</td>
<td>-</td>
<td>-356.4-0</td>
</tr>
</tbody>
</table>

Table 5.7 “true” vs. calibrated hydraulic conductivity and specific storage prior to CO$_2$ injection

<table>
<thead>
<tr>
<th>scenario</th>
<th>zone</th>
<th>“true”</th>
<th>cal.</th>
<th>error of</th>
<th>“true”</th>
<th>cal.</th>
<th>error of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$K_w$</td>
<td>$K_w$</td>
<td>($\cdot$)</td>
<td>$S_{sw}$</td>
<td>$S_{sw}$</td>
<td>($\cdot$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(m/s)</td>
<td>(m/s)</td>
<td></td>
<td>(1/m)</td>
<td>(1/m)</td>
<td></td>
</tr>
<tr>
<td>homogeneous</td>
<td>zone 1</td>
<td>2.46×10$^8$</td>
<td>2.24×10$^8$</td>
<td>-0.09</td>
<td>9.10×10$^{-7}$</td>
<td>8.60×10$^{-7}$</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>entire aquifer</td>
<td>9.54×10$^6$</td>
<td>8.49×10$^6$</td>
<td>-0.11</td>
<td>9.41×10$^{-7}$</td>
<td>8.62×10$^{-7}$</td>
<td>-0.08</td>
</tr>
<tr>
<td>2layers_A</td>
<td>zone 1</td>
<td>2.46×10$^5$</td>
<td>2.19×10$^5$</td>
<td>-0.11</td>
<td>1.11×10$^{-6}$</td>
<td>8.62×10$^{-7}$</td>
<td>-0.22</td>
</tr>
<tr>
<td></td>
<td>zone 2</td>
<td>2.46×10$^6$</td>
<td>2.18×10$^6$</td>
<td>-0.11</td>
<td>8.59×10$^{-7}$</td>
<td>8.58×10$^{-7}$</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>entire aquifer</td>
<td>1.75×10$^7$</td>
<td>1.56×10$^7$</td>
<td>-0.11</td>
<td>1.03×10$^{-6}$</td>
<td>9.42×10$^{-7}$</td>
<td>-0.08</td>
</tr>
<tr>
<td>2layers_B</td>
<td>zone 1</td>
<td>2.46×10$^6$</td>
<td>2.44×10$^6$</td>
<td>-0.01</td>
<td>8.59×10$^{-7}$</td>
<td>8.95×10$^{-7}$</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>zone 2</td>
<td>2.46×10$^5$</td>
<td>2.18×10$^5$</td>
<td>-0.11</td>
<td>1.11×10$^{-6}$</td>
<td>9.36×10$^{-7}$</td>
<td>-0.15</td>
</tr>
</tbody>
</table>
## Table 5.8 Zonal calibration results (1-plume)

<table>
<thead>
<tr>
<th>scenario</th>
<th>time</th>
<th>structure</th>
<th>cal. ( S_z )</th>
<th>cal. ( S_n )</th>
<th>true ( S_n )</th>
<th>( \xi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>short-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>homogeneous</td>
<td>PT</td>
<td>4.7\times10^{-5}</td>
<td>0.25</td>
<td>0.24</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>9.3\times10^{-5}</td>
<td>0.64</td>
<td>0.34</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>4.2\times10^{-5}</td>
<td>0.22</td>
<td>0.22</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>long-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT</td>
<td>6.8\times10^{-5}</td>
<td>0.39</td>
<td>0.36</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>6.9\times10^{-5}</td>
<td>0.40</td>
<td>0.37</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>7.3\times10^{-5}</td>
<td>0.43</td>
<td>0.39</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2layers_A</td>
<td>short-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT</td>
<td>2.5\times10^{-5}</td>
<td>0.12</td>
<td>0.08</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>2.9\times10^{-5}</td>
<td>0.14</td>
<td>0.09</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>2.3\times10^{-5}</td>
<td>0.11</td>
<td>0.08</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>long-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT</td>
<td>3.6\times10^{-5}</td>
<td>0.17</td>
<td>0.14</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>4.2\times10^{-5}</td>
<td>0.21</td>
<td>0.16</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>3.9\times10^{-5}</td>
<td>0.19</td>
<td>0.15</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>2layers_B</td>
<td>short-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT</td>
<td>2.6\times10^{-5}</td>
<td>0.11</td>
<td>0.11</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>3.8\times10^{-5}</td>
<td>0.17</td>
<td>0.15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>2.6\times10^{-5}</td>
<td>0.11</td>
<td>0.11</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>long-injection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PT</td>
<td>5.7\times10^{-5}</td>
<td>0.27</td>
<td>0.19</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>4.3\times10^{-5}</td>
<td>0.19</td>
<td>0.18</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JT</td>
<td>3.9\times10^{-5}</td>
<td>0.17</td>
<td>0.17</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.9 Zonal calibration results (2-plume)

<table>
<thead>
<tr>
<th>scenario</th>
<th>time</th>
<th>structure</th>
<th>secondary plume</th>
<th>cal. $S_n$ (1/m)</th>
<th>cal. $S_{n'}$ (-)</th>
<th>true $S_{n'}$ (-)</th>
<th>$\xi$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>short-injection</td>
<td>ST</td>
<td>plume1</td>
<td>2.54×10^{-5}</td>
<td>0.10</td>
<td>0.092</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>1.48×10^{-5}</td>
<td>0.072</td>
<td>0.071</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>3.47×10^{-5}</td>
<td>0.14</td>
<td>0.10</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>1.24×10^{-5}</td>
<td>0.06</td>
<td>0.077</td>
<td>-24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>2.34×10^{-5}</td>
<td>0.089</td>
<td>0.091</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>1.46×10^{-5}</td>
<td>0.071</td>
<td>0.068</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>4.47×10^{-5}</td>
<td>0.18</td>
<td>0.22</td>
<td>-15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>1.81×10^{-5}</td>
<td>0.090</td>
<td>0.051</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>5.02×10^{-5}</td>
<td>0.21</td>
<td>0.25</td>
<td>-17</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>2.35×10^{-5}</td>
<td>0.12</td>
<td>0.053</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>4.96×10^{-5}</td>
<td>0.21</td>
<td>0.24</td>
<td>-13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>1.85×10^{-5}</td>
<td>0.092</td>
<td>0.052</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>long-injection</td>
<td>ST</td>
<td>plume1</td>
<td>2.15×10^{-5}</td>
<td>0.11</td>
<td>0.09</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>3.31×10^{-5}</td>
<td>0.13</td>
<td>0.12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>3.15×10^{-5}</td>
<td>0.17</td>
<td>0.12</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>5.13×10^{-5}</td>
<td>0.22</td>
<td>0.20</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>2.18×10^{-5}</td>
<td>0.11</td>
<td>0.10</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>3.30×10^{-5}</td>
<td>0.13</td>
<td>0.12</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>6.19×10^{-5}</td>
<td>0.27</td>
<td>0.21</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>4.77×10^{-5}</td>
<td>0.27</td>
<td>0.17</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>3.60×10^{-5}</td>
<td>0.19</td>
<td>0.15</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>5.42×10^{-5}</td>
<td>0.23</td>
<td>0.22</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume1</td>
<td>3.26×10^{-5}</td>
<td>0.17</td>
<td>0.17</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>plume2</td>
<td>4.65×10^{-5}</td>
<td>0.19</td>
<td>0.18</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.10 Diffusivity difference ($\Delta D$) vs. CO2 saturation ($S_n$). Model 1: homogeneous scenario; model 2: integrated model for 2layers_A; model 3: integrated model for 2layers_B; model 4: layer with small permeability, k; model 5: layer with large k;
Figure 5.11 2-D clustering examples of a) homogeneous scenario; b) 2layers_B

Figure 5.12 “true” saturations vs. calculated saturations (2-plume structure)
6 Conclusions and outlook

6.1 Concluding remarks

In this thesis, pressure tomography is tested in the CO$_2$-brine system. A single-phase proxy is introduced for facilitating fast travel-time based inversion and accelerating the calibration. Overall, it has been demonstrated that the pressure tomography can be used for characterizing an evolving CO$_2$ plume in saline aquifers, and it has the potential to be coupled with more established geophysical methods.

In Chapter 3, a synthetic homogeneous saline aquifer is simulated to assess the feasibility of the proposed pressure tomography method. CO$_2$-induced variations of the flow properties are first examined. By approximating the two-phase separate flow by a single-phase flow problem, it is revealed that the mixed-phase diffusivity and specific storage can vary by two orders of magnitudes of the original single-phase diffusivity. In contrast, the mixed-phase conductivity shows a relatively small change due to the CO$_2$ injection. Variations of the mixed-phase diffusivity along with the CO$_2$ saturation are not monotonic, which makes the direct transformation from the inverted diffusivity to the saturation difficult. Moreover, low trajectory density appearing in the low diffusivity areas causes inaccuracy and nonuniqueness in the tomograms during the eikonal-based inversion, which hampers the reproduction of the “true” values. Therefore, a time-lapse clustering approach is utilized for determining the plume extent. Clustering results show good agreement with the “true” plume development over time. Saturation of the clustered plume is estimated by calibrating the full pressure signals in a single-phase emulator, greatly improving the computing efficiency. Calibrated saturation is consistent with the “true” value, which gives a first insight of pressure tomography for identifying the storage reservoir and the CO$_2$ plume.

Chapter 4 shows an application of pressure tomography for detecting and characterizing CO$_2$ leakage into an aquifer above the storage reservoir. This method is extended to be used in a typical multi-layer CO$_2$ sequestration system, which is composed of a storage formation, a caprock, and an above aquifer. Pressure tomography is conducted in the storage formation and above monitoring aquifer by alternative CO$_2$ and brine injections, respectively. Different leaky cases are assumed, including a fractured leaky case and a diffusive leaky case. Pressure responses appear in the above aquifer when the leaks occurred. An acute change indicates fractured leakage, and on the contrary, pressure changes gradually if CO$_2$ leaks through the caprock diffusively. Travel times are also an important sign for leakage. CO$_2$ has a delay effect on travel times. Leakage of CO$_2$ into the upper aquifer leads to a reduction of the CO$_2$ mass in the reservoir, which provokes that the travel times in the storage reservoir are smaller in two leakage cases. In contrast, the travel times in the upper aquifer are larger than in the no-leakage case, which can be used for inverting the leaky plume. Results imply that the spreading of the leaky plume can be delineated by the diffusivity tomograms, which is however influenced by the data quality and the configuration of tomographic arrays.
In Chapter 5, the inversion performance of pressure tomography is further evaluated in heterogeneous formations and compared to crosswell seismic tomography. Similar to pressure tomography, P-wave velocity also decreases with the increased CO$_2$ saturation, which is the basis for inverting the CO$_2$ plume. However, the CO$_2$-induced contrast of P-wave velocity are much smaller than the mixed-phase diffusivity. This explains why seismic travel times in general have smaller variance than hydraulic travel times. Capacity for resolving the plume extent of these two approaches are different. It has to be pointed out that both of the two approaches cannot retrieve the “true” mixed-phase diffusivity nor P-wave velocity. Specifically, the aquifer structure can be determined by clustering the inverted diffusivity before CO$_2$ injection, which is attributed to the large variability of the initial reservoir permeability. After CO$_2$ injection, pressure tomography is constrained by the observed heterogeneity of the initial aquifer. Although this can resolve more complicated plume geometries, the main part of the plume may be hardly detected from the diffusivity tomograms. In the opposite, seismic tomography gives more robust results on depicting the major part of the plume. Combination of the two approaches by joint clustering shows great potential for making improvement on the plume identification. The plume extents show better agreement with the “true” plumes, and the estimation error of CO$_2$ saturation is also reduced.

6.2 Outlook: current challenges and future perspective

As shown in this thesis, pressure tomography is tested in simplified 2-D models, and also with several simplifications of the physical processes. There still remain several challenges of the method, which require further exploration and improvement:

- The validity of the single-phase proxy has to be examined with different site conditions. In this work, I only selected a realization in which the pressure and temperature are the same at a specific site. For demonstrating the robustness of the proxy, it has to be examined with more different site conditions, for instance, with realistic pressure and temperature gradients. Moreover, the inversion bias caused by the assumptions (e.g., zero capillary pressure, no residual trapping, etc.) has to be further analyzed to outline the application scope of the proxy.
- The numerical models used in this thesis are limited to two dimensions. In particular, 2-D cross-sectional model are used, which could greatly reduce the simulating time. However, this type of set-up refers to a quasi 3-D condition by neglecting the flow in the third dimension. Therefore, implementation of the method in 2-D radial or 3-D models is more realistic and can be oriented at more realistic site conditions. A complete 3-D inversion requires not only increasing the number of spatially distributed observations, but also the pressure stimulation performed in different wellbores will be more demanding. Pressure tomography is sensitive to the heterogeneity of the host formation. The inversion algorithm needs to be improved in the future to resolve the plume shape. Based on the single-phase proxy, other single-phase based pressure tomographical methods can be applied to address this issue, such as the pilot-point
method shown in Appendix 2. Furthermore, pressure tomography needs to be tested in more heterogeneous cases, such as in an aquifer analogue, and also for large-scale problems.

- Combination of pressure tomography and geophysical approaches by utilizing available field data. At the CO$_2$ injection site Ketzin, several hydraulic tests, CO$_2$ injection tests and geophysical tests have been conducted. Data from these tests can be used for evaluating the CO$_2$ saturation. It will be a significant step for validating the single-phase proxy.

- Although pressure tomography in this work is only used for CO$_2$ sequestration problem, it is still feasible for other multiphase flow conditions, such as DNAPL-water (two-phase) and DNAPL-water-gas (three-phase) regimes. Furthermore, theoretically, any significant variations of flow properties can be characterized by implementing pressure tomography in different vintages. Hence, time-lapse pressure tomography can not only be utilized for early-time stage, but also for long-time scale problem. Future work is considered for adopting this method to broaden its application field.
Appendix 1  Rapid field application of hydraulic tomography for resolving aquifer heterogeneity in unconsolidated sediments


Abstract

A new framework is introduced for hydraulic tomography application and validation in the field. Our motivation is the need for methods that are both efficient and expressive for resolving the spatial distribution of heterogeneous hydraulic properties in aquifers. The presented strategy involves time-efficient field experiments and a computationally efficient inversion scheme. By exploiting the early travel time diagnostics of the hydraulic pressure pulses recorded during tomographic crosswell tests, and new application of attenuation inversion, only short-term pumping tests are required. Many of these can be conducted in one day. The procedure is developed by a numerical experiment with a highly heterogeneous aquifer analogue and then applied to a field case with a shallow, unconsolidated sedimentary aquifer, the Stegemühle site in Germany. It is demonstrated that the performance of a suite of tomographic short-term pumping tests, data processing and inversion for the reconstruction of heterogeneous diffusivity and specific storage distribution is possible within one day. Additionally, direct-push injection logging is performed at the field site, and the obtained field data is utilized for successful validation of the hydraulic tomograms. We also compare both methods with respect to the necessary requirements, time demand in the field and complexity of interpretation.
A1.1 Introduction

Hydraulic tomography and direct-push (DP) methods were identified as the next generation of hydraulic characterization technologies [Butler, 2005]. Hydraulic tomography involves the combined interpretation of multiple cross well hydraulic tests. The recorded hydraulic pulses are inverted to characterize the hydraulic conductivity (K) and specific storage (S_v) distributions with a high resolution and accuracy. DP can also serve this purpose, as it allows for low-cost vertical profiling of K at several nearby positions in a shallow aquifer. Since 2005, both methods were further developed and tested at different field sites. The centerpiece of DP profiling is the probe, which is attached to a steel pipe string. It is advanced into the ground by the weight of the DP unit, or, if probe design allows, depth of penetration can be augmented by application of a hydraulic percussion hammer. Liu et al. [2012] give a comprehensive overview of different DP probes.

They distinguish between probes for absolute and relative K-profiling. Absolute K-profiling is performed with the direct push permeameter (DPP), which consists of a probe with a short screen and two pressure transducers attached to the probe surface. After the probe is pushed to the target depth, a short hydraulic test is performed by injecting water through the screen and recording the pressure response at the attached pressure transducers, while the flow is controlled at the surface. A complete test series at one depth level can be performed within 15 min. The DPP probe was proposed and successfully tested at the well-documented Geohydrologic Experimental Monitoring Site [e.g., Butler, 2005] in Lawrence, Kansas, and at the Nauen site located 40 km west of Berlin, Germany [Yaramanci et al., 2002]. They recorded several K-profiles and validated them by multilevel slug tests and by information obtained from core samples.

Relative K-profiling is performed with the direct push injection logger (DPIL), which consists of a probe with a short screen for the injection of water. Pressure and flow are recorded and regulated at the surface using a pressure transducer and a flow controller. Dietrich et al. [2008] proposed a DPIL-probe, which can be used only in a discontinuous mode. This means that the probe advancement has to be stopped at each test interval (depth level). They reconstructed a relative K-profile over a depth of 18 m consisting of 35 measurement points in a fine- to medium-grained sand aquifer. The profile was obtained within 3 h and validated by core data and slug tests. An alternative DPIL probe, which allows for continuous profiling with a vertical resolution of 0.015 m was proposed by McCall et al. [2009]. This DPIL probe was tested within a sand and gravel aquifer, where several 30 m deep relative K-profiles with a vertical resolution of 0.015 m and an advancement rate of 0.02 m/s could be recorded.

Liu et al. [2009] proposed a new type of probe, which combines the advantages of both the DPIL probe designed for continuous K-profiling and the DPP probe designed for absolute K-profiling. The variant is called “high-resolution K probe” (HRK) and it facilitates sampling of a 10 m profile with a vertical resolution of 0.015 m in less than 30 min. Bohling et al. [2012] applied these hydraulic DP tools to obtain several profiles from the heavily studied MAcro Dispersion Experiment (MADE) site [Boggs et al., 1992] to characterize detailed K-variations at the site. The data are compared with those from the flowmeter profiles that have served as the primary basis

117
for characterizing the heterogeneous aquifer at the site [Rehfeldt et al., 1992]. Overall, the patterns of variation acquired by DP are quite similar to those in the flowmeter data.

Similar to the rapid development of DP profiling, hydraulic tomography has substantially evolved. Since 2005, several new and improved inversion schemes have been introduced, and the number of hydraulic tomography field studies is continuously increasing. This history documents the transition of hydraulic tomography as theoretic concept that is approved for characterizing synthetic “digital” aquifers to a robust field method to reconstruct hydraulic properties of real aquifers. In addition to $K$-fields, spatial diffusivity and/or $S_r$-distributions can be obtained.

Bohling et al. [2007a] proposed a field assessment of hydraulic tomography in unconsolidated sediments utilizing the steady shape flow regime [Wenzel and Fishel, 1942]. The latter enables evaluation of transient data with the computational efficiency of a steady state model. They reconstructed the spatial K-distribution assuming a given structure of homogeneous horizontally arranged layers that are persistent within the entire model domain. The reconstructions were validated by tracer and DPP tests. For the technical field implementation, the continuous multichannel tubing (CMT) system [Einarson and Cherry, 2002] was applied. This allowed for recording of pressure response at seven different depth levels without using multipacker systems in the observation wells.

Straface et al. [2007a] published the first field hydraulic tomography application that exploits the potential of geostatistically based inversion. The applied procedure relies on the sequential successive linear estimator (SSLE), as proposed by Zhu and Yeh [2005]. The SSLE inversion scheme successively includes the transient head data from different pumping tests, such that the size of the covariance matrix is small and the calculation demand can be reduced. To account for the non-uniqueness issue the hydraulic parameter field is treated as an outcome of a stochastic spatial process, whereby the mean parameter distribution is reconstructed by matching the observations from the pumping test responses. Straface et al. [2007a] used this inversion scheme for depth integrated reconstruction of hydraulic conductivity and specific storage fields. However, the reconstructed hydraulic conductivity and specific storage tomograms showed no correlation among each other, which is expected in natural sedimentary aquifers [e.g., Bayer et al., 2011], and no additional field data were provided to support the tomograms. Meanwhile, SSLE-based inversions have successfully been applied in other studies as well. Berg and Illman [2011b] performed a three-dimensional (3-D) transient hydraulic tomographic field assessment in highly heterogeneous till and glaciofluvial material. The database of their hydraulic tomographic investigations comprises four pumping and up to 41 observation intervals with a pumping time between 6.5 and 24 h. The large number of observation intervals was realized by utilizing a CMT system. The reconstructed 3-D hydraulic conductivity and storage tomograms were validated by a high number of permeameter data.

A large scale field assessment (>500 m) was performed by Illman et al. [2009b]. They analyzed two large-scale cross-hole pumping tests in a granite aquifer to compute hydraulic conductivity and storage tomograms. The reconstructions show several distinct zones characterized by high-hydraulic conductivity and low-specific storage values that are continuous over hundreds of
meters. The authors interpreted these features as fault zones. Huang et al. [2011b] presented a field study of steady state hydraulic tomography in unconsolidated sediments based on 110 pressure response curves, whereby each pumping test was performed for at least 72 h to reach steady state flow conditions. For the inversion, they used the scheme of Xiang et al. [2009], which is an improved SSLE allowing for the simultaneous inversion of all drawdown curves. The reconstructed tomograms displayed a depth integrated distribution of hydraulic conductivity and storage, and they were validated by matching the drawdown of pressure response curves that were not used for the inversion. Li et al. [2008] also employed a geostatistical approach to jointly invert data from steady state pumping tests and flowmeter measurements to estimate hydraulic conductivity in three dimensions. The pumping tests were performed in 29 fully screened wells and it took about 2 h to reach steady state conditions.

The first hydraulic tomographic field study in an unconfined aquifer was presented by Cardiff et al. [2009]. The steady state inversion is based on nine dipole pumping tests performed in unconsolidated gravel and sand sediments. The tests were designed in a way that a full set of drawdown and recovery data can be collected in one working day. For the reconstruction of the depth integrated hydraulic conductivity field data, the Bayesian geostatistical inversion method proposed by Kitanidis [1995] was utilized.

An alternative inversion approach is based on a transformation of the transient groundwater flow equation into the eikonal equation, using an asymptotic approach [Virieux et al., 1994]. The eikonal equation can be solved with less computational effort using ray tracing techniques, i.e., calculation of trajectories, or particle tracking methods that allow for the calculation of transient pressure propagation along trajectories [Vasco et al., 2000]. He et al. [2006] further developed the approach by Vasco et al. [2000] through matching the amplitudes in addition to travel time. For the inversion, they used an iterative sparse matrix solver [Paige and Saunders, 1982]. In their application, two pumping tests were performed in naturally fractured limestone. They showed that the reconstruction of a two dimensional (2-D) permeability tomogram enables one to image an orthogonal fracture pattern, which could be validated by seismic tomographic measurements.

Brauchler et al. [2010] applied a similar travel time based inversion scheme to crosswell interference slug tests at the Stegemühle site in Germany which consists of a shallow, 2 m thick sand and gravel aquifer overlain by a confining clay layer with a thickness of 2–3 m. They inverted the pressure response of 196 crosswell interference slug tests performed between five wells. Based on a travel time inversion scheme proposed by Brauchler et al. [2003], they re-constructed the 2-D diffusivity distribution of the aquifer. Brauchler et al. [2011] further developed the inversion approach and inverted crosswell interference slug tests with a travel time and attenuation based inversion scheme. They obtained the diffusivity and specific storage distribution between the wells in two and three dimensions.

The hydraulic tomographic field applications so far have shown that the method is suitable for mapping hydraulic subsurface features and for estimating the hydraulic parameters hydraulic conductivity and storage in 2-D and 3-D. However, obtaining a high-spatial resolution by
hydraulic tomographic investigation is still accompanied by relative complex analysis [Bohling et al., 2002b], partly high-computational costs, and relative long pumping times.

In this paper, we propose a field strategy for hydraulic tomography that can be (a) analyzed and (b) performed with a similar speed as DP profiling. (a) Therefore, we further developed the computationally efficient travel time and attenuation based inverse scheme proposed by Brauchler et al. [2011]. The inverse scheme allows for the inversion of Dirac signals and was successfully applied to the inversion of crosswell interference slug tests. In this work we further develop the inverse scheme and adapt it to the requirements of Heaviside signals and apply it to the inversion of short term pumping tests. Beyond this, we introduce the concept of null space energy maps for the validation of the hydraulic tomograms, which was originally developed for geophysical ray tomography. The further developed inverse scheme is tested numerically with data from a hydraulic tomography analogue outcrop study. (b) The field implementation is realized in a way that a suite of tomographic measurements can be recorded in 1 day. Therefore, we limit the pumping time to 300 s, which permits us to record 30 transient pressure response curves between a 200 well (pumping well) and a CMT (observation well) in a few hours. The reconstructed tomograms are compared and evaluated by means of four DPIL logs performed between the pumping and the observation well.

A1.2 Field data processing and inversion methodology

In the following, we introduce the fundamental concept of travel time based inversion as shown in detail by Brauchler et al. [2011], and discuss the requirements for field data processing. Then, a novel adaptation of the attenuation based inversion scheme to the inversion of pumping tests signals is proposed. The methodology is first applied to a numerical example. The gained experience is exploited for a field demonstration and validation with DPIL.

A1.2.1 Travel time inversion of pumping tests data

The proposed travel time inversion is based on the work of Vasco et al. [2000]. They developed an inversion scheme based on the transformation of the groundwater flow equation into the eikonal equation [Virieux et al., 1994], which can be solved with ray tracing or particle tracking methods in a computationally efficient way. The keystone of the procedure is the following line integral proposed by Vasco et al. [2000]:

\[
h_h(r, t_{\alpha,h}) = \alpha_h h_{\text{max}}, \quad 0 < \alpha_h < 1 \quad (A1.1)
\]

where \(t_{\text{peak}}\) is defined the source-to-observation-point travel time of the peak of a transient pressure curve, resulting from a Dirac signal generated at point \(x_1\) (source), traveling along the propagation path \(s\), and recorded at point \(x_2\) (observation point). \(D\) is the diffusivity. In a homogeneous media Equation (A1.1) can be expressed as follows [Vasco et al., 2000]:

120
\[ t_{\text{peak}} = \frac{S_r r^2}{6K} \]  
(A1.2)

where \( S_r \) is the specific storage, \( r \) is the distance between the source and the observation point, and \( K \) the hydraulic conductivity. Equation (A1.1) is only valid for an impulsive type source (Dirac source) and was successfully applied to invert travel times derived from crosswell interference slug tests by Brauchler et al. [2010]. Vasco et al. [2000] showed that the pressure response of a pumping test signal, usually described with a Heaviside function, can be analyzed as a response to an impulsive (Dirac delta) type source by considering the time-derivative of head data. This differentiation when applied to field data comprises three steps: (1) Wavelet de-noising [e.g. Xiang et al., 2009] is utilized to smooth the transient head data. (2) A polynomial regression is applied to the smoothed transient head data. (3) The first derivative of the polynomial fit is estimated. The three steps are depicted in Figure A1.1.

### A1.2.2 Travel time diagnostics

For travel time inversion, we use an early travel time diagnostic rather than the peak travel time. Brauchler et al. (2007) defined the travel time diagnostic as “the time of occurrence of a certain feature of the transient pressure pulse”. The \( t-10\% \) diagnostic in this study is the time at which the time derivative of the pressure pulse rises to 10\% of its ultimate peak (Figure A1.1a). In this study we decided to use the \( t-10\% \) diagnostic for the numerical example and the \( t-50\% \) diagnostic for the field example. The choice of the travel time diagnostics is a compromise between data quality (early time noise) and the findings of Cheng et al. [2009] that early travel times are better suited to resolve the diffusivity distribution of an aquifer. Figure 1 shows that the travel time diagnostic \( t_{\text{peak}} \) can be determined with a high level of confidence. In order to determine the uncertainty associated with the choice of the 50\% travel time diagnostics, used in the field example, we performed a Monte Carlo analysis. The filtered signal of the pressure response curve displayed in Figure A1.1 was fitted with a Gaussian function consisting of 4 terms and 12 coefficients. For 2800 fits the travel time diagnostic \( t_{50\%} \) and \( t_{\text{peak}} \) were calculated. The 2800 calculated travel time diagnostics \( t_{50\%} \) are characterized by a mean of 1.16 s and an associated standard deviation of 0.076 s. The comparison with the 2800 calculated travel time diagnostics \( t_{\text{peak}} \) characterized by a mean value of 4.03 s and an associated standard deviation of 0.84 indicates that the travel time diagnostic \( t_{50\%} \) can be determined with a comparable accuracy as the travel time diagnostic \( t_{\text{peak}} \).

In Figure A1.2, additionally, the statistical parameters median, 25th and 75th percentiles, and the range are displayed as box-whisker plot for the travel time diagnostic \( t_{50\%} \). With respect to the mathematical derivation of the transformation factor for the inversion of additional travel time diagnostics besides the peak time the reader is referred to Brauchler et al. [2003].
Figure A1.1 Field data processing. (a) Processing of the pressure response recorded at the observation interval. (b) Processing of the pressure response recorded at the pumping interval.

Figure A1.2 The statistical parameters of the Monte Carlo analysis, median, 25th and 75th percentiles, and the range are displayed for the travel time diagnostic t-50%.
A1.2.3 Attenuation inversion of pumping test data

Brauchler et al. [2011] presented an attenuation integral which relates the attenuation of a Dirac source signal to the inverse of the aquifer parameter specific storage $S_s$. The line integral reads:

$$\left( \frac{h_d(x_2)}{H_{0,d}} \right)^{-1/3} = B^{-1/3} \int_{h_1}^{h_2} \left( \frac{1}{S_s(s)} \right)^{-1/3} ds$$  \hspace{1cm} (A1.3)

The attenuation of the Dirac source signal is expressed by the initial displacement $H_0$ and the hydraulic head $h(x_2)$ at the observation point as a function of the length of the propagation path $s$. The subscript d stands for a Dirac source and the parameter $B$ summarizes all test specific parameters that can be treated as constants during the inversion and is defined as follows:

$$B = \frac{\pi r_c^2}{\sqrt{\left( \frac{2\pi}{3} \right)^3}} \exp \left( \frac{3}{2} \right)$$ \hspace{1cm} (A1.4)

Here $r_c$ is the casing radius of the test well.

The attenuation integral is only valid for a Dirac type source signal. For the determination of the right hand side of Equation (A1.3) the head $h(x_2)$ and $H_0$ have to be determined. $h(x_2)$ can be easily estimated utilizing the first derivative of the pressure response at the observation interval (Figure A1.1a). For the determination of the initial displacement $H_0$, it is not possible to refer to the first derivative of the pumping signal, because the velocity of the water movement in the well is largest at $t_0$, (the point of time when the pumping test is initiated) and this point of time is camouflaged by noise. Hence, we utilize a conversion factor, which allows for relating Heaviside and Dirac signals in order to estimate the initial displacement $H_0$. In the following, we summarize the derivation of the conversion factor proposed by Brauchler et al. [2003].

A1.2.4 Derivation of the Conversion factor for a Heaviside Source (after Brauchler et al., 2003)

The solution of the flow equation using a Heaviside source for an infinite domain can be expressed according to Häfner et al. [1992] as follows:

$$h_h(r,t) = \frac{Q}{4\pi r K} \text{erfc} \left( \frac{r}{\sqrt{4\pi K t / S_s}} \right)$$ \hspace{1cm} (A1.5)

where $Q$ denotes the flow rate, $K$ the hydraulic conductivity and $h_h(r,t)$ is the hydraulic head as function of space and time. Spherical coordinates are used here because each injection port represents a point source and the signal can be assumed to spread radially due to the test intervals.
The subscript $h$ stands for a Heaviside source and the maximum drawdown $h_{\text{max}}$ is equivalent to
\[
\frac{Q}{(4\pi rK)}.
\]

The key element of the transformation factor is the introduction of a head ratio $\alpha_h$, which enables the conversion of a signal originating from a Heaviside source into a signal originating from a Dirac source. $\alpha_h$ is introduced as follows:
\[
h_h(r,t_{a,h}) = \alpha_h h_{\text{max}}; \quad 0 < \alpha_h < 1 \quad (A1.6)
\]

Inserting Equation (A1.5) in Equation (A1.6) yields:
\[
\frac{Q}{4\pi rK} \text{erfc} \frac{r}{\sqrt{4Kt_{a,h}/S_y}} = \alpha_h \frac{Q}{4\pi rK} \quad (A1.7)
\]

Utilizing Equation (A1.2) to simplify the left hand side of Equation (A1.7) results in:
\[
\frac{Q}{4\pi rK} \text{erfc} \frac{3t_{\text{peak}}}{2t_{a,h}} = \alpha_h \frac{Q}{4\pi rK} \quad (A1.8)
\]

Introducing the conversion factor $f_{a,h} = \frac{t_{\text{peak}}}{t_{a,h}}$ leads to:
\[
\alpha_h = \text{erfc} \frac{3}{2} f_{a,h} \text{erfc} \frac{3}{2} \quad (A1.9)
\]

The conversion factor $f_{a,h}$ is 1, if the height of the Heaviside signal is around 8.36 % of $h_{\text{max}}$. The time and the respective hydraulic head are equivalent to the peak time and the amplitude of a signal with a Dirac source at the origin. By replacing $r$ by $r_c$ in Equation (A1.6) this conversion factor can be applied to the signal recorded in the pumping interval if $h_{\text{max}}$ is known. In this case $h_{\text{max}}$ is defined as drawdown under steady state conditions. In theory, steady state conditions can only be established if a constant head boundary is reached. For an efficient field implementation, we define steady state conditions to be established when the drawdown curve can be approximated by a straight line with a slope of one-tenth of a percent, whereby the slope is defined in meters per second. Figure A1.1b displays the straight line and $h_{\text{max}}$. In conclusion, the amplitude $H_0$ assuming an impulsive type source is defined by 8.36 % of $h_{\text{max}}$.

Note, the estimation of $h_{\text{max}}$ is an approximation because the derivation of the conversion factor is based on the assumption of a homogeneous hydraulic parameter distribution. However, Brauchler et al. [2003] applied this conversion factor successfully to characterize a fractured sandstone sample. In the following, we perform a numerical study based on an aquifer analogue outcrop study in order to show the potential and limits of the proposed hydraulic attenuation inversion.
A1.3 Numerical example of the attenuation inversion based on an analogue outcrop study

This section builds upon the hydraulic tomography outcrop analogue study performed by Hu et al. [2011]. The outcrop analogue, consisting of fluvial unconsolidated sediments, was developed by Bayer et al. [2011]. The study comprises six parallel high-resolution photographs of an exposed quarry face that were taken every 2 m, as the gravel was excavated. The outcrop photographs were carefully interpreted to yield maps of lithology. For each representative lithological unit, measurements were performed in the laboratory providing porosity, as well as hydraulic conductivity (hydrofacies classification). The specific storage values, assigned to each hydrofacies group were taken from data reported in literature [Domenico and Mifflin, 1965]. Maji and Sudicky [2008] interpolated between the six profiles and translated the gathered information into a 3D hydraulic parameter distribution. Meanwhile an alternative interpolation based on multiple point statistics is also available [Comunian et al. 2011].

Hu et al. [2011] simulated a suite of short-term pumping tests using a tomographic measurement array, which we will utilize to investigate the possibilities of the attenuation inversion (Equations (A1.3)-(A1.9)). The tomographic set-up is displayed in Figure A1.3. For the attenuation inversion we utilized only source-receiver configurations with a trajectory angle \( \alpha \) smaller than 40°. The trajectory angle \( \alpha \) is defined as the angle between the horizontal and a straight line connecting source and receiver (Figure A1.3). Thereby we follow the suggestion of Hu et al. [2011] that data with smaller source-receiver angles are better suited for the reconstruction of horizontally arranged features.

![Figure A1.3 Spatial position of the pumping and observation intervals. (a) Cross section. (b) Plan view.](image)

Figure A1.4 shows the limitations of applying the conversion factor for the processing of data. For source-receiver combinations where the source is located in the high permeability zone between 2.5 and 4 m in vertical position (Figure A1.4a and c), we receive unreasonably high values for the attenuation \( h(t) / H_0 \). This is not surprising because the derivation of the conversion factor...
is based on homogeneity assumption, but the aquifer analogue data set is characterized by an up to 5 orders of magnitude range of hydraulic conductivity values (Figure A1.4a and e). For this reason, we neglected source-receiver combinations with a value larger than 0.5 for $h(t)/H_0$. This upper constrain was determined by Equation (A1.3). The distance between source ($x_1$) and receiver ($x_2$) position was appraised with the well distance. Inserting this parameter in Equation (A1.3) gives an attenuation value of 0.5 that corresponds to a specific storage value of $1.5 \times 10^{-5}$ 1/m, which represents a lower boundary of specific storage values in fluvial unconsolidated sediments.

The model domain for the inversion consists of 8 cells in $x$-direction (horizontal) and 10 cells in $y$-direction (vertical). To increase the nominal resolution of the inversion, the staggered grid method is applied [Vesnaver and Böhm, 2000]. It averages the velocity values obtained from different inversions of a regular coarse gridded model, slightly shifted in space, both in horizontal and vertical directions. We performed such a displacement of the initial grid, three times in $x$- and $y$-direction.

The inversion results for specific storage are displayed in Figure A1.4c, g. Both the reconstructions in West-East and North-South directions show that the most characteristic feature, located between $y = 1$ m and 3 m in vertical direction and characterized by higher specific storage:

Figure A1.4 Comparison of the aquifer analog data with the reconstructed specific storage values and the associated null space energy maps. (a–d) Vertical profiles in West-East direction. (e–h) Vertical profiles in South-North direction.
values, could be reconstructed reliably. The horizontal continuity of this zone as well as the absolute values could be determined with adequate precision. The comparison with the true aquifer analogue data set (Figure A1.4b, f) reveals that also the zone between \( y = 4 \) m and 5 m, characterized by slightly higher specific storage values compared to the background, is reproduced. However, comparison with the aquifer analogue also shows that small-scale variability (<20 cm in size) using the proposed attenuation inversion procedure cannot be resolved (see also Hu et al. [2011]). In particular, the zone between \( y = 3 \) m and 4 m can be reconstructed only close to the test well.

To assess the reliability of the tomographic model we computed the null-space energy map of the area of the model domain. The null-space energy map represents a measure of the reliability of a tomographic system, because it relates the trajectory distribution to the mesh used for the discretization of the investigated area. Further details about the calculation of the null space energy map are given in supporting information.

In Figure A1.4d, h we displayed the null space energy maps ranging from zero to one, whereby a value of zero indicates a high reliability and a value of one a low reliability. In general, the displayed null space energy maps are characterized by a high reliability, which can be explained with the numerical experimental set-up, i.e. high number of test and observation intervals and high trajectory density, respectively. However, at the boundaries of the model domain the trajectory density is decreasing and thus the reliability is decreasing. The influence of the low reliability can be seen in the reconstructed specific storage tomograms at the top of the model domain, which is characterized by low specific storage values. Within this zone three positions are marked with red ellipses that are characterized by higher values. These higher values are associated with a low reliability and, hence, could be potential artifacts.

### A1.4 Field application of the travel time and attenuation tomography

The field implementation was performed at the well-characterized Stegemühle test site located in the Leine valley, close to Göttingen, Germany. Currently, the infrastructure of the site consists of a network of 26 monitoring locations, comprising 1”, 2”, 6” and multi-chamber wells screened over the whole aquifer thickness (Figure A1.5). The 6” wells were drilled with a top drive drilling rig, whereas all other wells were installed using DP technology (e.g. Dietrich and Leven [2006]). Brauchler et al. [2010] give a detailed overview about the structural and hydraulic characterization of the site using conventional investigation techniques. The structural composition of the braided river sediments was characterized by surface refraction seismics, gamma ray logging and direct-push electrical conductivity logging. For selected wells, cores were recovered to calibrate the recorded logs. The aquifer has a thickness between 2-2.5 m, and it is built up by intercalated sand and gravel layers. A confining unit that is composed of silt and clay overlies the aquifer. The thickness of the confining unit varies between 3-3.5 m. Figure A1.6a,b shows the two DP EC logs recorded at the wells P0/M25 and P5/M17.5 (Figure A1.5) that were used as test and observation wells for the field implementation of hydraulic tomography. Hydraulic characterization is based on
single-well pumping and crosswell slug interference tests. Hydraulic conductivity estimates, derived from the analytical evaluation of multi-level single-well slug tests, performed over five to seven different depths in each 2” well, vary between $10^{-4}$ m$^{-1}$ and $1.2 \times 10^{-3}$ m$^{-1}$. Generally, the $K$-values increase with aquifer depth (Figure A1.7).

![Figure A1.5 Monitoring well network at the Stegemühle test site.](image)

![Figure A1.6 Geological interpretation of the subsurface derived from DP electrical conductivity logging.](image)
Figure A1.7 Hydraulic conductivity estimates derived from type curve analysis of multi-level slug tests.

The data base for the hydraulic tomographic experiment comprises 30 pressure responses that were recorded between the 2” well (P0/M25) and the multi-chamber well (P5/M17.5). The experimental set-up is displayed in Figure A1.8. The suite of tomographic pumping tests was recorded by employing a double packer system with a screened interval of 0.25 m and an interval tube inner diameter of 0.031 m in the test well. For the observation well, we utilized the Continuous Multi-channel Tubing (CMT) System [Einarson and Cherry, 2002]. This system was originally developed for multi-level sampling and consists of a pipe with seven continuous separate channels or chambers (ID = 0.014 m), which are arranged in a honeycomb shape. In each individual chamber a 0.08 m long opening, covered with a sand filter, was cut. In total five short-term pumping tests were carried out using a double-packer-system in the test well, and the respective pressure responses of each test were recorded at six different depth levels in the multi-channel well with a frequency of 50 Hz. The experimental set-up is displayed in Figure A1.8.

Figure A1.8 Experimental setup of the hydraulic tomography measurements.
A diffusivity and a specific storage tomogram were reconstructed using the new inversion procedure proposed in section 2. For the diffusivity reconstruction displayed in Figure A1.9a, the 50% travel time diagnostic was employed for the inversion. The prerequisite of using such an early travel time diagnostic is excellent data quality. Figure A1.1a clearly shows that it was possible to record pressure drawdown curves with a maximum drawdown below one centimeter with a very low noise level. The model domain consists of 45 cells for both tomograms. Additionally, we applied the method of staggered grid and shifted the mesh four times in the horizontal direction and three times in the vertical direction. As starting values, we used the parameter estimates derived from the analytical evaluation of the pumping tests data. The travel time and attenuation inversion, including the staggered grid calculation, took less than one minute on a conventional notebook. For the validation of the tomograms the null space energy maps were calculated (Figure A1.9b, d).

![Figure A1.9 (a, b) Reconstructed diffusivity tomogram and the associated null-space energy map. (c, d) Reconstructed specific storage tomogram and the associated null-space energy map. (e) Computed hydraulic conductivity tomogram utilizing the equation $D = K / S_s$. (f) Comparison of DPIL-logs with an interpolated image of the hydraulic conductivity tomogram.](image)

Both diffusivity and specific storage tomograms are characterized by a horizontal layering (Figure A1.9a, c). The reconstructed diffusivity distribution shows the lowest value, approximately 2 m$^2$/s close to aquifer top, and the highest values between 10 m$^2$/s to 20 m$^2$/s close to aquifer bottom (Figure A1.9a). The specific storage distribution is characterized by values between $3 \times 10^{-5}$ and $10^{-4}$ 1/m, whereby the lowest values are close to aquifer bottom and the highest ones close to the top (Figure A1.9c). The hydraulic conductivity tomogram, depicted in Figure A1.9e, is computed from the equation $D = K / S_s$ and shows a similar parameter distribution as the diffusivity tomogram. The estimated values range from $10^{-4}$ to $10^{-3}$ m/s and are representative for sand and gravel aquifers. Furthermore, the range, and the general spatial distribution of the derived
hydraulic conductivity values, agrees with the values derived from multi-level slug testing (Figure A1.7).

The calculated null space energy maps associated with the diffusivity and specific storage tomogram show both a similar pattern (Figure A1.9b, d). The maps show the highest values (close to 1) indicating a low reliability, at the top of the aquifer between \( x = 3 \) m and 9 m. in a depth between 147.4 and 147.0 m.a.s.l. The low reliability of these sections can be explained with a low trajectory density and the low trajectory density is caused again by the observation position, which is located 0.2 m below the top of the model domain. Note, the applied inversion technique allows only for the determination of the parameter space between test and observation interval. The only way to improve the significance of this part is to install a further observation interval at the top of the aquifer.

### A1.5 DP injection logging

The direct push injection logger (DPIL) is a small diameter tool that consists of a short screen located just behind a drive point, which is attached to the lower end of a pipe string [Dietrich et al., 2008]. The probe is advanced in the ground by using the weight of the DP unit supported by a hydraulic jack hammer. During advancement in the ground, water is injected through the screen in order to avoid clogging. As soon as the desired test depth is reached, further advancement is stopped and a series of tests are performed. In this field study, at the Stegemühle site, testing was performed with three different injection (flow) rates. Flow rate and pressure were measured on-site. Both can be transformed into a relative hydraulic conductivity estimate \( K_{DPIL} \), which can be used as a proxy for \( K \). For details on the derivation of \( K_{DPIL} \) the reader is referred to Dietrich et al. [2008].

We recorded five profiles, each comprising between 15 and 20 measurement intervals. The profiles were recorded between the test and observation well used for the short-term pumping tests (Figure A1.5). However, only four profiles could be evaluated because one log exhibited a strong dependence between flow rate and \( K_{DPIL} \), which indicates technical problems with the flow controller or trunk line. In Figure A1.9e, the profiles are displayed in comparison to the reconstructed hydraulic conductivity tomogram.

The comparison between the DPIL profiles with the reconstructed hydraulic conductivity tomogram shows an overall agreement that the highest hydraulic conductivity values are below 146 m.a.s.l. That means the top of the higher permeability zone can be determined in the DPIL logs as well as in the hydraulic conductivity tomogram. Beyond this, the DPIL logs recorded at \( x \) -directions 6.2 m and 7.5 m agree over the whole thickness of the aquifer. In particular, the agreement between the DPIL log recorded at the position \( x = 7.5 \) m, characterized by the lowest \( K_{DPIL} \) in a depth 146.8 m.a.s.l. and highest \( K_{DPIL} \) values of all logs in a depth of 146.2 m.a.s.l agrees in all details with the reconstructed hydraulic conductivity tomogram. Note, the positions of the DPIL profiles are shown by the black arrow in Figure A1.9f.
The upper part of the DPIL logs, above 147 m.a.s.l., at x-directions 3 m and 4.5 m indicates changes in hydraulic conductivity not reconstructed in the hydraulic conductivity tomograms. One explanation for the lack of dynamics in the hydraulic conductivity tomogram could be the different resolution of the two characterization techniques. The DPIL logs display the hydraulic properties in the vicinity of the open screen section. That means small scale heterogeneities on the sub decimetre scale strongly influences the DPIL results. However, the numerical example has shown that the applied tomographic inversion scheme is not able to resolve such small scale heterogeneities. Another explanation could be the low reliability of the hydraulic conductivity tomogram above 147 m.a.s.l. indicated by the calculated null space energy maps (Figure A1.9b, d). The decline of the $K_{DPIL}$ values below 146.4 m.a.s.l can be explained by the transition from the aquifer to the underlying aquitard. This transition zone is interpreted as a mix of aquifer material and weathered marl stone. Note this zone could not be reconstructed by the hydraulic tomography, because all source and receiver positions were located within the aquifer. Still, the agreement of the DPIL profiles and the reconstructed $K$-tomogram supports the reliability of the estimated spatial distribution of the hydraulic properties estimated with the travel time and attenuation tomography, as well as with the DPIL.

A1.6 Acknowledgement

The investigations were conducted with the financial support of the Swiss National Science Foundation to the project “A field assessment of high-resolution aquifer characterization: An integrated approach combining hydraulic tomography and tracer tomography” under grant number 200021_140450/1 and the financial support of the German Research Foundation to the project “High resolution aquifer characterization based on direct-push technology: An integrated approach coupling hydraulic and seismic tomography” under grant no. BR3379/1-2. We would like to thank Gualtiero Böhm for his help with the null space energy maps. This manuscript greatly benefited from comments by Jungfeng Zhu and two anonymous reviewers.

A1.7 Conclusions

We presented a field assessment of high-resolution aquifer characterization of unconsolidated sediments based on hydraulic tomography and DP injection logging. The results show that both, the hydraulic tomograms and the DP injection logging provide information about hydraulic subsurface parameters with a spatial resolution that would not be feasible with conventional hydraulic investigation and evaluation technologies. However, for practical implementation, the following aspects have to be considered: (a) equipment requirements, (b) time demand for field implementation, (c) complexity or easiness of the evaluation.
(a) The equipment requirements for both methods are comparable: Both methods require a DP unit in order to access the subsurface i.e. for probe advancement and well installation. Beyond this standard pumping test equipment, a double packer system is needed to record a suite of hydraulic tomography measurements. For the hydraulic profiling the DP injection logger consisting of a screened metal cone, trunk-line, flow controller and pressure transducer, is needed. We think in comparison to conventional hydraulic field testing, which relies on existing wells, the requirements for both investigation technologies can be justified with respect to the gained spatial high resolution parameter estimates.

(b) The DPIL, as well as the short-term pumping test, can be performed in one day including test set-up. In this field study, we pumped for five minutes, which is more than sufficient for the applied travel time and attenuation based inversion scheme. Short pumping time can be a limitation for other inversion schemes that are based on solving the groundwater flow equation, because the early times are mainly determined by the specific storage, $S_s$, distribution, and later times or even steady state conditions are better suited for reconstructing the $K$-field [Wu et al., 2005; Sun et al., 2012].

For the hydraulic tomography measurements, an additional day is needed for well installation. However, an experienced field technician team can install five DP-wells with a depth of 10 m easily in one day. Note, by installing a larger number of observation wells more than one tomographic profile can be recorded without substantially increasing the workload.

(c) We adapted the travel time and attenuation based inverse scheme proposed by Brauchler et al. [2011] to the requirements needed for the tomographic inversion of short-term pumping tests in order to minimize the complexity and time requirements for hydraulic tomography inversion. The main advantages of the inversion scheme are the low computational requirements of eikonal solvers and that no information about the hydraulic boundary conditions of the investigated area is needed. However, the applied inversion scheme can only reconstruct the parameter reconstruction between test and observation well. The processing of the short-term pumping test data can be largely automatized using any script language, and for the inversion user-friendly eikonal solvers are commercially available. The short pumping time in combination with the straightforward inversion technique allows for the reconstruction of preliminary $K$ and $S_s$ distribution already in the field, which is particularly useful for an adaptive site investigation approach. In this context it has also to mention that inversion schemes that solve the groundwater flow equation might reconstruct hydraulic conductivity fields that are closer to the true parameter distribution. However, as mentioned above short-term pumping test are not well suited for these inversion approaches because fitting later times of a transient pressure curve or even steady state conditions lead to better $K$-field reconstructions.

The evaluation of the DPIL is based on a simple spread sheet calculation and the relative $K_{DPIL}$ values can be estimated directly in the field. The shortcomings of the DPIL providing relative $K_{DPIL}$ values can be overcome by using complementary measurements such as DP slug tests or by using other probes such as the high resolution $K$-profiling tool (HRK).
The field assessment of the new hydraulic tomography framework and comparison with DPIL showed that both methods provide information about hydraulic conductivity on a scale and accuracy which goes beyond the expressiveness of conventional hydraulic testing. The further developed hydraulic attenuation inversion scheme additionally allows for the reconstruction of specific storage distribution. At the Stegemühle site it was revealed that a spatially high resolution of the parameter estimates derived from hydraulic tomography can be recorded and evaluated in a short amount of time, which even allows for an adaptive site characterization, i.e. deciding immediately on-site about measurements to be performed next, based on prior investigation results. Beyond this, hydraulic parameter estimates with such a spatially high resolution have the potential to strongly increase the significance of hydrogeophysical investigations [Brauchler et al., 2012].

### A1.8 Supporting information

In the following we give a short summary of the calculation of the of the null-space energy map, which is based on the work of Böhm and Vesnaver, 1996. The null space energy map was originally developed for the validation of geophysical ray tomograms. The null space energy map comprises a singular value decomposition of the tomographic matrix $A$, where $a_{ij}$ of the matrix $A$ are the lengths of the $i_{th}$ trajectory path in the $j_{th}$ cell. This matrix can be factorized into three components:

$$A = U W V^T$$  \hspace{1cm} (A1.10)

The squared matrices $U$ and $V$ are orthonormal:

$$U U^T = I$$  \hspace{1cm} (A1.11)

$$V V^T = I$$  \hspace{1cm} (A1.12)

$U$ and the elements $w_j$ of the diagonal matrix $W$ are the singular values corresponding to the square of the eigenvalues. The stability of the tomographic inversion is defined by the elements $w_j$, whereby small singular values show instabilities of the inversion process. Note, the columns of the matrix $V$ of the decomposition (Equation (A1.10)) display an orthonormal basis of the model domain. Hence, the local reliability, $R$, of each pixel can be defined as the summation of the squared elements $v_i$ of the matrix $V$:

$$R = \sum_i {v_i}^2$$  \hspace{1cm} (A1.13)
Appendix 2 Prediction of solute transport in a heterogeneous aquifer utilizing hydraulic conductivity and specific storage tomograms


Abstract

A sequential procedure of hydraulic tomographical inversion is applied to characterize at high resolution the spatial heterogeneity of hydraulic conductivity and specific storage at the field test site Stegemühle, Germany. The shallow aquifer at this site is examined by five short-term multilevel pumping tests with 30 pumping-observation pairs between two wells. Utilizing travel time diagnostics of the recorded pressure response curves, fast eikonal-based inversion is shown to deliver insight into the sedimentary structures. Thus, the structural information from the generated travel time tomogram is exploited to constrain full calibration of the pressure response curves. Based on lateral extrapolation from the measured inter-well profile, a three-dimensional reconstruction of the aquifer is obtained. It is demonstrated that calibration of spatially variable specific storage in addition to hydraulic conductivity can improve the fitting of the model while the structural features are only slightly changed. At the field site, two tracer tests with uranine and sodium-naphthionate were also performed and their concentrations were monitored for 2 months. The measured tracer breakthrough curves are employed for independent validation of the hydraulic tomographical reconstruction. It is demonstrated that major features of the observed solute transport can be reproduced, and structures relevant for macrodispersive tracer spreading could be resolved. However, for the mildly heterogeneous aquifer, the tracer breakthrough curves can also be approximated by a simplified homogeneous model with higher dispersivity. Therefore, improved validation results that capture specific characteristics of the breakthrough curves would require additional hydraulic measurements.
A2.1 Introduction

Nearly all hydraulic tomographic field studies are driven by the need to provide spatial high-resolution parameter fields for solute transport predictions. In fact, tomographic approaches, among others [e.g., Mariethoz et al., 2010], can resolve sedimentary structures or fractures that control preferential flow. Their potential and superiority to traditional field investigation techniques was demonstrated in several previous studies [Gottlieb and Dietrich, 1995; Yeh and Liu, 2000; Illman et al., 2010; Berg and Illman, 2011a, 2015]. However, the effort of data collection and data evaluation is higher for tomographic investigation methods in comparison to conventional methods that avoid spatial assignment of estimated hydraulic parameters. This motivates a strong interest for enhanced tomographic field and inversion techniques [Bohling et al., 2002; Zhu and Yeh, 2006; Lochbühler et al., 2013]. Naturally, the development of new field technologies and field data collection strategies is delayed in time in comparison to the computer-based development of inversion schemes. Numerical studies with virtual aquifers are essential means for motivating, developing, and testing new schemes, but their viability can only be approved by often laborious field experiments. Therefore, especially during the last few years, the number of field studies has been catching up. These started from simplified two-dimensional, depth-integrated characterizations [e.g., Straface et al., 2007b] to arrive at full three-dimensional reconstructions [e.g., Illman et al., 2009; Berg and Illman, 2011b; Cardiff et al., 2013] based on a large number of interference tests.

For field investigations, one of the important and at the same time most challenging tasks is the evaluation of the significance and reliability of the reconstructed hydraulic parameter fields. Independent from the inversion technique, all field studies use the residual error from data fitting as a first measure for the quality of their inversion results. Unfortunately, this information is not sufficient because a large number of parameter distributions might exist that equally honor the measured data. Hence, independent information and measures have to be exploited to evaluate the quality of the reconstructed parameter fields. Geological information such as deposition information or fault information based on a detailed structural geological study is utilized by Straface et al. [2007a] and Illman et al. [2009] to support reconstructed parameter fields. Berg and Illman [2012] used a large number of permeameter and grain size tests in combination with multilevel slug tests to interpret the estimated hydraulic conductivity (K) fields. The tests were performed at the North Campus Research Site (NCRS), Waterloo, Canada, in a heterogeneous confined aquifer built up by tills and glaciofluvial deposits. The high vertical resolution of multilevel slug tests and direct-push injection-logs were exploited by Brauchler et al. [2010, 2013] to interpret reconstructed diffusivity fields estimated at the Stegemühle Site, Göttingen, Germany. This site is characterized by a shallow confined aquifer consisting of fluviatile sediments. In comparison to the conditions at the NCRS [variance of log conductivity, $\sigma_K^2 = 1.72$, Alexander et al., 2011], the aquifer at the Stegemühle Site is less heterogeneous ($\sigma_K^2 = 0.2$).
Cardiff et al. [2012] and [2013a] compared porosity logs and multi-level slug tests with reconstructed parameter fields. They performed a three-dimensional (3-D) transient hydraulic tomographic field experiment utilizing highly flexible packer systems at the Boise Hydrogeophysical Research Site (BHRS), USA. The BHRS site [e.g., Straface et al., 2011; Cardiff et al., 2013a] is characterized by a mixed sand/gravel/cobble facies and in comparison to the other two test sites it shows unconfined conditions.

Hydraulic tomographical measurements were performed by Vasco and Karasaki [2001] and [2006a] to reconstruct preferential flow paths at the Raymond Field Site, California, USA. Intensive geological experiments allowed for a comparison of identified preferential flow paths, imaged in the hydraulic tomograms, with borehole conductivity logs and seismic tomograms. Such utilization of independently collected data for comparison with the reconstructed hydraulic tomograms can be a challenge due to different observation scale and mismatch in resolution [e.g., Brauchler et al., 2012]. Huang et al. [2011a] successfully applied independent validation pumping tests to field data recorded at the test site of the National Yunlin University of Science and Technology in Taiwan. The site consists of fluvialite sediments with mean hydraulic conductivity values of around $10^{-4}$ m/s, which are comparable to those found at the Stegemühle Site. A more direct way to validate tomograms is to use direct visual comparisons between inverted hydraulic conductivity and laboratory experiments [e.g., Illman et al., 2007; 2010].

Although the main motivation of hydraulic tomographic field studies is to provide high-resolution information for solute transport predictions, only a small number of field studies were published that employ tracer test data to interpret or validate reconstructed tomograms. Bohling et al. [2007b] utilized a solute tracer test to evaluate the capability of steady-shape tomography. They show that the tracer test could support the existence of a highly conductive layer, which was reconstructed by hydraulic tomography. For further verification, a large number of small-scale hydraulic tests were performed at the Geohydrologic Experimental and Monitoring Site (GEMS), USA, but none of these revealed the presence of this high-conductivity zone. The GEMS site is a heavily studied alluvial confined aquifer that consists of 11 m of sand and gravel overlain by silt and clay. Another field example utilizing tracer test data and flow data for inversion was presented by Vasco and Finsterle [2004] at the Grimsel Rock Laboratory in Switzerland. In contrast to the work of Bohling et al. [2007b], the tracer test data was not used for validation. Instead, the different data types were inverted together to improve the sig- nificance of the reconstructed hydraulic tomograms. Illman et al. [2012b] showed that estimated hydraulic conductivity tomograms predicted better tracer distribution patterns during a dipole tracer test than other traditional methodologies (i.e., effective parameter/macrodispersion approach or heterogeneous approach using ordinary kriging based on core samples). They also emphasize the difficulties of capturing details of the tracer breakthrough due to intrinsic methodological limitations, such as effects of noise in head measurements and “the less diffusive nature of the tracer which demands a much higher resolution mapping of the $K$-field”. Ni et al. [2009] compared the predictive capabilities of a 2-D tomographic reconstruction to that of a homogeneous model. In their theoretical study, they showed that the tomographic variant was capable of reproducing tracer
breakthrough curves (BTCs) independently of the transport distance. In contrast, a homogeneous advection-dispersion model with empirically estimated dispersivity could not match the BTC. When calibrated to the BTC, the homogeneous model would properly reproduce the BTC, but obtained dispersivity would need to be raised with transport distance.

In this paper, our main objective is to validate subsurface reconstructions from hydraulic tomographic inversion for predicting solute transport in the field. Therefore, we choose the hydraulic tomography procedure developed by Jiménez et al. [2013], which combines eikonal and pilot point based inversion approaches. The procedure was originally presented for the reconstruction of a \( K \)-field and theoretically assessed by application to a virtual aquifer. Here, we further develop and adapt it to the requirements of the simultaneous 3-D reconstruction of specific storage and hydraulic conductivity. We apply it to the inversion of short-term pumping tests at the Stegemühle site. Between the same wells originally used for the hydraulic tomographic investigations by Brauchler et al. [2013a], a forced gradient tracer test is performed. Two fluorescent tracers are injected in two different depths of the aquifer and the BTCs are recorded in the observation well over the entire thickness of the aquifer. The tracer test results are contrasted with the new findings from hydraulic inversion, and we reveal, to what extent and at which accuracy it is feasible to reconstruct structures relevant for subsurface transport. Similarly to Ni et al. [2009], we also compare the BTC prediction by reconstructed model to that of a homogenous one.

A2.2 Material and methods

A2.2.1 Field site and experiments

Stegemühle site

The experimental field test site Stegemühle is located south of the city of Göttingen, Lower Saxony, Germany (Figure A2.1). In order to carry out hydrogeological and hydrochemical field research under controlled natural conditions, five 1”-, twenty-one 2”- (five of them are multi-chamber wells) and three 6”-observation wells were installed during the period of 2006–2011. The composition of the shallow subsurface was determined by a variety of methods such as inspection of sediment cores, grain size analysis, direct-push electrical conductivity logging, borehole gamma-ray logging, electrical resistivity tomography, and seismic travel time inversion [e.g., Hu, 2007; Vogt, 2007; Meyer et al., 2014; Hu, 2011]. The aquifer is composed of unconsolidated fluviatile sediments (sand and gravel) of Quaternary age (Weichsel Glaciation). These sediments have a varying thickness of 1.0–3.3 m and are overlain by alluvial clay. The aquifer bottom is at a depth of 1.9-7.0 m below land surface with erosional contact to the underlying claystone formation of Middle Keuper Age. In the middle of the field site, which is the focus area of this study, the aquifer exhibits confined conditions. Here, Hu [2011] and Brauchler et al. [2013b]
applied multi-level slug tests and observed vertically varying hydraulic conductivities with higher values at the bottom of the aquifer.

Figure A2.1 Monitoring well network at the Stegemühle test site. The 2” wells are colored black, and multichamber wells are colored in blue. The plane between P5/M17.5 and P0/M25 corresponds to the eikonal inversion domain.

Field implementation of hydraulic tests

A series of crosswell multi-level pumping tests were performed at the test site Stegemühle, implementing a tomographic array along a straight line between a pumping well (P0/M25) and an observation well (P5/M17.5) (Table A2.1). The distance between these two wells is 9 m (Figure A2.1). During each pumping test, the water was partially pumped out of the pumping well P0/M25 by employing double packer systems with a screened interval of 0.25 m. The tube connected to the pump has an internal diameter (ID) of 0.031 m. The observation well P5/M17.5 (Figure A2.2) is a multi-chamber well-constructed with the Continuous Multi-channel Tubing (CMT) System [Einarson and Cherry, 2002]. This well consists of a pipe with six continuous separate channels (ID = 0.014 m), which are arranged in a honeycomb shape and lead to different depths. This design allows for the measurement of water level changes at different depths of the aquifer.

For the profile between the pumping well and the observation well, five short-term pumping tests were carried out. For every short-term pumping test and every pumping interval, the pressure changes in the six different depths of the multi-chamber wells were recorded at a frequency of 50 Hz with the pressure transducer (PDCR 35/D-8070) connected to a data logger (Campbell Scientific® CR 3000). By varying the pumping interval, a total number of 30 (5×6) drawdown curves for the profile were recorded [Brauchler et al., 2013]. The pumping tests in series produced a pattern of crossing trajectories between test and observation well, similar to the paths of a radar or seismic experiment. The travel times and hydraulic attenuations between the wells P0/M25 and P5/M17.5
thus can be utilized for eikonal based cross-sectional reconstruction of hydraulic parameter distributions.

Table A2.1 Basic information of the two wells P0/M25 and P5/M17.5 used for hydraulic tomography inversion and tracer testing

<table>
<thead>
<tr>
<th></th>
<th>P0/M25</th>
<th>P5/M17.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>single screen</td>
<td>multi-chamber</td>
</tr>
<tr>
<td>Aquifer thickness (m)</td>
<td>1.98</td>
<td>1.99</td>
</tr>
<tr>
<td>Well height (m)</td>
<td>0.87</td>
<td>0.65</td>
</tr>
<tr>
<td>Elevation of the well top (m.a.s.l)</td>
<td>152.23</td>
<td>151.54</td>
</tr>
<tr>
<td>Surface elevation (m.a.s.l)</td>
<td>151.36</td>
<td>150.89</td>
</tr>
<tr>
<td>Well bottom (m.a.s.l)</td>
<td>145.28</td>
<td>145.3</td>
</tr>
<tr>
<td>Aquifer top (m.a.s.l)</td>
<td>147.382</td>
<td>147.415</td>
</tr>
<tr>
<td>Aquifer bottom (m.a.s.l)</td>
<td>145.401</td>
<td>145.419</td>
</tr>
</tbody>
</table>

Field implementation of tracer tests

Tracer test data is used for independent validation of the derived aquifer model. Non-reactive tracer tests are effective means to identify preferential flow paths or integral transport parameters of the subsurface, such as porosity and dispersivity. They can be conducted under natural gradient or forced gradient conditions. Compared to the natural gradient tracer test, the forced gradient tracer test is hydraulically well controlled and not constrained to a given natural flow field. Moreover, higher mass recovery rates can be achieved [e.g., Ptak et al., 2004]. Consequently, forced gradient conditions were also favored at the field site.

The experiment was conducted between the same wells previously used for the pumping tests (P0/m25, P5/m17.5). We selected two different tracers with similar transport properties, the fluorescent dye tracers uranine and sodium-napthionate [Leibundgut et al., 2009], in order to allow for a robust analysis and to mitigate the effects of possible measurement errors, data noise or tracer-specific transport behavior.

Prior to tracer injection, a steady radial flow field was established by continuous water extraction from the single screen well at a constant rate of ~ 0.3 l/s. A mass of 15 g uranine and 150 g sodium-napthionate dissolved in water were injected into the multi-chamber well, each in a different depth (Figure A2.2). The injection periods of uranine and sodium-napthionate were 5 min and 3 min, respectively, and by this a pulse-like injection was realized. The different injection periods were employed due to technical reasons. They have negligible influence on the results as they are very small compared to the time of tracer arrival. At the monitoring well depth-integrated concentrations were measured every 30 s by a flow-through field fluorometer (type GGUN-FL30), calibrated to the specific tracers and local groundwater conditions. For quality control, at some time points the pumped-out water was also manually sampled and subsequently analyzed in the laboratory. The entire tracer experiment lasted two months, and the water levels at the two wells
were measured every few days to ensure the flow field was steady. Detailed information on experimental settings is summarized in Table A2.2.

![Figure A2.2 Setup of the tracer experiment at Stegemühle site with well configuration and illustration of injection levels for uranine and sodium-naphthionate. Additionally, electrical conductivity (EC) measurements are depicted (in red) for both wells, which delineate the aquifer boundaries.]

<table>
<thead>
<tr>
<th></th>
<th>Uranine</th>
<th>Sodium-naphthionate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pumping rate (l/s)</td>
<td>0.301-0.305</td>
<td></td>
</tr>
<tr>
<td>Injection start data and time</td>
<td>2011/10/18, 13:45</td>
<td>2011/10/18, 13:53</td>
</tr>
<tr>
<td>Injection mass (g)</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td>Injection chamber</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Height of chamber (m.a.s.l)</td>
<td>145.761</td>
<td>147.261</td>
</tr>
<tr>
<td>Injection duration (s)</td>
<td>300</td>
<td>180</td>
</tr>
</tbody>
</table>

**A2.2.2 Tomographic inversion**

The inversion procedure we use here encompasses the sequential application of an eikonal based and pilot point based inversion scheme (Figure A2.3). It was presented in detail by Jiménez et al. [2013] and makes use of the strengths of both inversions and minimizes drawbacks. On the
one side, eikonal based inversion is a fast, well tested methodology capable of providing insight into hydraulic parameters and aquifer structure. Strictly speaking it represents an approximation, since it treats the parabolic flow equation as a wave equation. In addition, the eikonal approach, as presented in this paper, only uses a time diagnostic from the whole pressure signal leaving the rest of the information unused. As a consequence, the eikonal based inversion is highly efficient for fast detection of structures, but as an approximation it is less accurate in estimation of hydraulic parameter values. On the other side, pilot point-based inversion commonly works by fitting the flow equation, using the full recorded pressure signal and thus making full use of the measured information. Therefore, hydraulic parameter values can be determined, but this is computationally demanding, especially for 3-D reconstructions.

![Diagram of sequential inversion procedure](image)

Figure A2.3 Main elements of sequential inversion procedure: the transient pressure signal is inverted by an eikonal-based approach to deliver a conceptual map. This structural information is used to constrain full pressure signal inversion based on pilot points.

Jiménez et al. [2013] showed how to link both schemes in a synergetic way. Eikonal based inversion is utilized to extract structures from reconstructed diffusivity fields ($D$-tomograms) but not parameter values. Since we are interested mainly in $K$, the original procedure is refined here by utilizing eikonal based estimates of $K$-distribution rather than diffusivity fields. As a bridging step, specific storage ($S_z$)-tomograms are developed by attenuation tomography [Brauchler et al., 2013]. Given $D = K / S_z$, a fully eikonal based $K$-tomogram can be derived from the $D$- and $S_z$-tomograms. The reliability of the tomographic models is assessed by means of null space energy maps. A null space energy map represents a measure of the reliability of a tomogram. It relates the
trajectory distribution to the mesh used for the discretization of the investigated area [e.g., Böhm and Vesnaver, 1996] and comprises a singular value decomposition of the tomographic matrix.

We consider the eikonal based $K$-distribution as a proxy, but carrying valuable insight into subsurface structures. For extracting this information, it is converted into a zonal image (conceptual map) using a k-means clustering algorithm. Pilot points serve as auxiliary variables that are often combined with regularization techniques to fill a model parameter field. During the pilot point based inversion procedure here, the information content of the conceptual maps is exploited as follows [Jiménez et al., 2013]:

Initial values: each individual pilot point is assigned an initial value equal to that of the corresponding cluster centroid.

Pilot point positioning: higher pilot point density is favored at locations where parametric variability is suspected, i.e. at cluster boundaries. This step is performed using a finite element mesh generator. For each cluster a mesh is designed, and at each node a fixed pilot point is assigned. This procedure automatically leads to a refinement at the cluster boundaries.

Regularization: Interpolation among the pilot points is based on the conceptual maps as well. Three conditions are proposed, which have to be fulfilled so that two pilot points are correlated: (i) both pilot points pertain to the same cluster; (ii) the distance between the pair of pilot points is smaller than the average length of the cluster in horizontal direction; and (iii) there must be no other pilot point from a different cluster within a given space of influence [Jiménez et al., 2013]. For the regularization implementation, i.e. the spatial relationships among the pilot points, a graph theoretical concept is adopted [e.g., Bhark et al., 2011a]. Initially, prior to the calibration, each pair of pilot points is examined and an adjacency matrix is developed:

$$
\begin{pmatrix}
0 & \ldots & a_{ij} \\
\vdots & \ddots & \vdots \\
a_{ij} & \ldots & 0
\end{pmatrix}
\quad \leftrightarrow 
\begin{cases}
1, \text{conditions fulfilled} \\
0, \text{otherwise}
\end{cases}
$$

(A2.1)

where $i = j = 1, \ldots$, number of pilot points, $a_{ij}$ is a boolean indicator, $p$ denotes a pilot point, and $C(\quad)$ is the adjacency matrix. The adjacency matrix dictates if two pilot points are connected in a graph or not, based on the three conditions listed above, and it is used to calculate the regularization function, $\Phi_r$.

Note that we want to arrive at a 3-D parameter field, but the conceptual map gives only insight into structures in a 2-D vertical slice between source and receiver well. For 3-D extrapolation, variograms along the horizontal and vertical axes are constructed from the eikonal-based diffusivity tomogram. Along the tomogram, pilot point values are considered as hard data, and the values for each cell of a given numerical model grid are derived from 3-D kriging.

The hydraulic parameters with spatial heterogeneity, which are addressed by the inversion procedure, are $K$ and $S_r$. On one hand, the resulting $K$- and $S_r$-field must honor the recorded pressure response data. This is evaluated by implementation in a flow model and comparison to the field data. On the other hand, the parameter fields are constrained by regularization. This is
described by an objective function, which is solved based on Lagrange multipliers [Doherty, 2010b; Jiménez et al., 2013].

### A2.2.3 Numerical modeling

#### Hydraulic test simulation

Inversion methods such as the pilot point-based approach use a large number of iterative flow model runs. Therefore, any possibilities for minimizing simulation time are of interest. An appealing option is using locally refined grids, and accordingly hydraulic crosswell simulations were performed with MODFLOW-LGR [Mehl and Hill, 2005; Vilhelmsen et al., 2012], a transient, three-dimensional groundwater flow code. MODFLOW-LGR allows local refinement of a finite difference grid, as an extension to the classical MODFLOW code. It couples two or more finite difference grids called parent and child. A parent grid can cover a large area in order to accommodate regional flow and boundaries. A much more refined child grid can be used to study more local phenomena, for instance the hydraulic effects in the vicinity of a pumping well.

#### Tracer test simulation

By numerical simulation of the tracer test with the reconstructed aquifer, the suitability of the hydraulic tomography approach for predicting solute transport can be evaluated. For this purpose, the reconstructed 3-D aquifer is implemented in a flow and transport model, and the simulated results are contrasted with those by a homogeneous case. The heterogeneous model represents the full heterogeneity in $K$ and $S_r$ obtained from the tomographical inversion procedure. In the homogeneous model, $K$ and $S_r$ are set constant, taking the mean of the estimated values. For flow modeling, the forward model used for pilot point-based full signal inversion is selected. The code MT3DMS [Zheng and Wang, 1999] is chosen for solving solute transport.

By comparing measured and, by these models, simulated tracer BTCs, we can assess the gain from resolving aquifer heterogeneity and also validate the inverted model. Still, additional parameters need to be specified before the transport models can be run. Crucial unknowns are dispersivity and effective porosity. Since these parameters cannot be determined separately with sufficient accuracy, we consider their possible value ranges, and estimate the most likely parameter ranges for homogeneous and heterogeneous models through a Bayesian approach, a Markov Chain Monte Carlo (MCMC) sampling procedure. It utilizes the Metropolis–Hastings algorithm [Metropolis et al., 1953; Hastings, 1970] to sample realizations of longitudinal dispersivity and effective porosity, separately for heterogeneous and homogeneous models. For simplification, transversal dispersivity is set 1/10 the value of the longitudinal one, which is a rough but common assumption in related work [Molina-Giraldo et al., 2011]. For generating new realizations within the MCMC framework, (i) one of the parameters is selected randomly, (ii) a
new parameter value is proposed using a Gaussian random walk, and (iii) the acceptance ratio \( \alpha \) is computed:

\[
\alpha = \min \left( 1, \frac{L(m_{\text{new}})}{L(m_{\text{old}})} \frac{g(m_{\text{old}} \rightarrow m_{\text{new}})}{g(m_{\text{new}} \rightarrow m_{\text{old}})} \right)
\]

(A2.2)

Where \( L \) is the likelihood function, \( g \) is the proposed distribution, and \( m \) denotes a model parameterization. Finally, a random number \( u \) is drawn from a uniform distribution on \([0,1]\) and the realization is accepted if \( a > u \), and rejected otherwise. As search criterion, the RMSE between measured and modeled tracer BTCs is selected. The function that maps from RMSE to likelihood is \( L = 10^{-\frac{\text{RMSE}}{2\sigma}} \) with \( \sigma \) equals to 0.2.

**A2.3 Results**

**A2.3.1 Eikonal based inversion of hydraulic tests**

Brauchler et al. [2013c] reconstructed a diffusivity (\( D \)) and specific storage tomogram (\( S_s \)) utilizing the eikonal based inversion approach. The derived tomograms displayed in Figure A2.4a-e are shortly discussed in the following; however, for more details we refer to Brauchler et al. [2013c].

For the eikonal based inversion, a starting 2-D model domain of 45 cells was applied. Utilizing the method of staggered grid, the mesh was shifted four times in the horizontal direction and three times in vertical direction, which led to the final resolution of 540 pixels imaged in Figure A2.4. For the diffusivity reconstruction displayed in Figure A2.4a, the 50\% travel time diagnostic was employed for the inversion. The travel time and attenuation inversion, including the staggered grid calculation, took less than one minute on a conventional notebook. For quantifying the reliability of the \( D \)- and \( S_s \)-tomograms (Figure A2.4a, c), null space energy maps are provided in Figure A2.4b, d. These maps illustrate the uncertainty associated with the eikonal based inversion. A value of 1 of the null space (black color) means the lowest possible confidence on the values obtained for a given cell in a tomogram, and a value of 0 (white color) indicates highest confidence.

The \( D \)- and \( S_s \)-tomograms indicate horizontal layering (Figure A2.4a, c), with values of \( D \) between 2 m\(^2\)/s and 20 m\(^2\)/s, and of \( S_s \) between 3 \times 10\(^{-5}\) and 10\(^{-4}\) m\(^{-1}\). The \( K \)-tomogram (Figure A2.4) is obtained by \( D = K / S_s \). It shows a similar structure, with moderate heterogeneity and values ranging from \( K = 10^{-4} \) to \( 10^{-3} \) m/s, which are considered typical for sand and gravel aquifers.
Figure A2.4 a and b) Reconstructed diffusivity tomogram and the associated null-space energy map. c and d) Reconstructed specific storage tomogram and the associated null-space energy map. e) Computed hydraulic conductivity tomogram using \( D = K / S_t \) [see Brauchler et al., 2013]. f) Resulting cluster distribution based on hydraulic conductivity tomogram. g) Pilot point distribution, cluster IDs, and Tikhonov regularization connections (gray lines). h) Resulting hydraulic conductivity from pilot points inversion using \( K \) and \( S_t \) as a parameter (longitudinal slice of the 3-D domain, Figure A2.8b, for comparison purposes).

### A2.3.2 Conceptual map and pilot points configuration

The conceptual map to support the subsequent pilot point inversion was developed from the eikonal based results and is displayed in Figure A2.4f. The number of clusters was set to four, which, after visual inspection, was considered the maximum possible to keep the main structures of the tomograms. Following Jiménez et al. [2013], cells associated with high null space energy values (here > 0.9) were ignored during clustering. These gap cells, which are mainly located close to the top boundary, were filled by nearest neighbor interpolation from adjacent cells.

The conceptual map is used to guide pilot point positioning and setting initial values for hydraulic parameter (\( K \) and \( S_t \)) calibration. This involves transforming the conceptual map into a vector graph in order to obtain a digital image of the cluster interfaces. The latter are displayed as black lines in Figure A2.4g. Then a mesh is assigned to the vector graph using a finite element mesh generator. The mesh nodes are translated into pilot point positions. Each pilot point an initial value equal to the value of the corresponding cluster centroid is assigned.

The selected mesh generator uses Delaunay triangulation. The utilization of the mesh generator leads to a higher pilot point density along the cluster boundaries. This distribution is favorable because along these boundaries the largest contrasts in hydraulic properties are expected. A maximum element size of 0.7 m and a resolution curvature of 0.3 are selected. Both parameters
control the mesh (i.e. pilot point) density and how fast it declines away from the cluster boundaries. The maximum element size determines how big each grid element can be and the resolution curvature limits the mesh size along a curved boundary [COMSOL, 2012]. Accordingly, the lower the values for each parameter are, the more pilot points are allocated. The maximum number of nodes is set to 1200. Table A2.3 lists the number of pilot points assigned to each cluster. The maximum element number, which controls the total number of pilot points, is a compromise between available computational resources, available observations and desired resolution.

Table A2.3 Number of pilot points and connections for each cluster

<table>
<thead>
<tr>
<th>Cluster ID</th>
<th>Number of pilot points</th>
<th>Number of connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>75</td>
<td>358</td>
</tr>
<tr>
<td>2</td>
<td>96</td>
<td>301</td>
</tr>
<tr>
<td>3</td>
<td>125</td>
<td>369</td>
</tr>
<tr>
<td>4</td>
<td>67</td>
<td>238</td>
</tr>
<tr>
<td>Total</td>
<td>363</td>
<td>1266</td>
</tr>
</tbody>
</table>

The regularization step for interpolation among the pilot points is also guided by the conceptual maps. Based on their neighborhood relationships, pilot points are connected in pairs, yielding separate networks that delineate the structures observed in the conceptual maps. In Figure 3-4g, these networks are illustrated as grey lines. For the maximum distance between a pair of pilot points three meters were chosen, which equals the average length of the clusters in horizontal direction, and for the calculation of the space of influence an angle of 30° was set [e.g., Jiménez et al., 2013].

For interpolating between the pilot points and lateral extrapolation in 3-D, ordinary kriging is applied. The diffusivity tomogram displayed in Figure A2.4a is utilized to derive the underlying semi-variograms depicted in Figure A2.5. As expected for fluviatile sediments, we find a larger range in the horizontal (2.3 m) than in the vertical direction (1 m). During the subsequent calibration, the $K$ and $S_v$ values at the pilot point cells are tuned, and by kriging the values of the other cells are filled. In the numerical model, pilot points exist only in the vertical tomogram slice between the source and receiver well. For lateral extrapolation, we assume stationary, horizontally isotropic geostatistical properties and consistently use kriging with the same semi-variograms. However, it is clear that due to the missing field data away from the well couple, the resulting 3-D field will lose reliability in lateral distance. Alternatively, several tomograms from different source and receiver wells may be collected from the field and combined for 3-D inversion, such as shown in Berg and Illman [2012a].
A2.3.3 Simulation of pumping tests

In the numerical groundwater flow model domain, the refined child grid (13 m×4 m×2 m) is embedded in the coarser parent grid (60 m×60 m×2 m). The selected refinement ratio between parent and child grid is 5:1, resulting in total around 1 million cells. Fixed hydraulic heads are implemented to enforce the regional hydraulic gradient (0.004) measured in the field.

The child grid simulates the source (P0/M25) and receiver well (P5/M17.5) from the field test. Along the source borehole, five vertical screens are implemented. This resembles those screens in the field at which the pumping tests were performed. At the receiver an array of six observation points is defined in the model to simulate the multi-chamber well configuration.

For the calibration, the repeated pumping tests at the source well are simulated by transient flow modeling. Each test simulates a 140 seconds pumping at 0.3 l/s. With five tests at different depths and six receivers, 30 pressure response curves are recorded in total. Each pressure signal is discretized by 300 points. Two choices are compared for fitting these curves to those measured at the field site. In the first one, only $K$ is calibrated, and $S_{\mu}$ is set a constant value of $S_{\mu,\text{mean}} = 5.5 \times 10^{-4}$ 1/m equal to the mean of the tomogram (Figure A2.4c). The initial values for $K$ are specified equal to the centroids of the clusters. In the second one, both parameters are independently adjusted, which means that the number of decision variables is doubled. However, to save computational time, here the initial values of $K$ are selected according to the results from previous calibration of $K$ only.

The selected optimization algorithm (Levenberg-Marquardt) is suited for parallelization, and a cluster with a desktop (Intel i7 3.4 GHz, 16 GB RAM) and a workstation (Intel Xeon E5 3.1 GHz, 64 GB RAM) was used. The number of parallel runs was 20, distributed on both machines. A single model run took approximately 5 minutes. A total of 1130 models runs were needed. The full pilot points inversion took approximately 5 hours.
A2.3.4 Calibrated hydraulic parameter fields

The hydraulic parameters at the Stegemühle site show low variability in comparison with the conditions at other test sites such as NCRS, where similar experiments were conducted [e.g., Berg and Illman, 2012b]. This is reflected in the pressure response curves (Figure A2.6) which all show a similar behavior. Despite that, in order to be able to resolve heterogeneous structures, we suggest to make full use of the measured information while exploiting the degrees of freedom in the hydraulic model. This means, the calibration procedure is applied to fit all and the complete pressure response curves, and this is achieved by not only calibrating the spatial distribution of $K$ but also $S_r$.

Figure A2.6 compares the 30 measured pressure response curves with those calibrated by $K$ adjustment only. The fitting error (root mean squared error, RMSE) is minimized to $5 \times 10^{-4}$ m, and it is shown that most curves are properly reproduced. This fitting error seems to be the lowest possible with the current parametrization. In some cases, the later stages are not ideally captured. This can be seen for most of the responses from the 5th interval at the receiver. This is improved by also including $S_r$ as a decision variable. The resulting modelled pressure curves, as depicted in Figure A2.7, fit better to the measurements, and we reduce the RMSE to $3 \times 10^{-4}$ m.

The resulting $K$-fields are visualized Figure A2.8. Kriging variance or estimation error of the hydraulic parameters assigned to cells increases rapidly from the vertical slice between source and receiver well, which contains the pilot points. Hence, to those cells surrounding the shown central region, mean values are assigned (also for Figure A2.9). Including $S_r$ in the inversion yields a very similar field, and only a central layer with higher $K$ found in Figure A2.8a appears less accentuated in Figure A2.8b. Figure A2.9 depicts the calibrated distribution of $S_r$. It is revealed that highest values of around $7 \times 10^{-5}$ m$^{-1}$ are characteristic for the upper part of the aquifer, whereas the $S_r$ in the lower ranges around $4 \times 10^{-5}$ m$^{-1}$. Comparison of Figure A2.8 and Figure A2.9 nicely shows how the structures are related. With regularization and kriging, these clusters stimulate the calibration of zones that can be interpreted as individual sedimentary hydrofacies [e.g., Bayer et al., 2011]. The latter are characterized by specific and fairly constant hydraulic properties, and this is reproduced here by the spatial correlation between the structures for $K$ and $S_r$ in Figure A2.8 and Figure A2.9.
Figure A2.6 Observed and modeled pressure responses by adjusting only hydraulic conductivity \( (K) \) during 3-D pilot point-based inversion.

Figure A2.7 Observed and modeled pressure responses by adjusting hydraulic conductivity \( (K) \) and specific storage \( (S_s) \) during 3-D pilot point-based inversion.

The vertical parameter distribution between source and receiver well in the 3-D aquifer model can be compared with the tomograms reconstructed based on travel time diagnostics (Figure A2.4). It is shown that the basic layer structure is maintained, with higher \( K \) values on the lower-right section and low values in the upper section. Same geometries can be recognized in both cluster map and pilot point based field (Figure A2.4f, h). In comparison with the travel time based \( K \) tomogram (Figure A2.4e), full signal inversion yields a decrease in the range of \( K \). This is also true for the specific storage ranges. In comparison with the \( S_s \)-tomogram (from \( 4 \times 10^{-5} \) to \( 8 \times 10^{-5} \) 1/m) (Figure A2.4c), a lower variability is observed in Figure A2.9 (from \( 4 \times 10^{-5} \) to \( 7 \times 10^{-5} \) 1/m).
Figure A2.8 Reconstructed hydraulic conductivity of aquifer using a) hydraulic conductivity and b) hydraulic conductivity and specific storage.
A2.3.5 Validation of the reconstructed aquifer with tracer test data

In order to validate the reconstructed hydraulic parameter fields for predicting solute transport, the uranium and sodium-naphthionate tracer tests are used. Applying tracer data to validate an inversion procedure based solely on hydraulic data poses evident challenges. Tracer transport is not fully predictable if only hydraulic data is used. Moreover, the tracer test was performed under a different hydraulic regime, and therefore comparison between tracer simulation and measurements will elucidate the robustness of the hydraulic inversion.

As explained, the two tracers were applied between the source and receiver wells used for the hydraulic inversion. The measured BTCs of the tracers are illustrated in Figure A2.10. The tracers were monitored for two months, but here only the time of breakthrough is shown. A first visual inspection reveals that the curves follow a nearly ideal shape with early steep increase of concentration and, after a peak is reached, tailing sets in. With a closer look, we also recognize non-uniformities in both curves. The sodium-naphthionate BTC shows a small step in the later phase after the peak has passed by. In contrast, before uranium reaches the maximum peak, an apparent local peak already appears which widens the period when the highest concentration is detected.

![Breakthrough curves](image)

Figure A2.10 Breakthrough curves (BTCs) measured at the pumping well during the tracer tests and MCMC realizations for the a) uranium and b) sodium-naphthionate.

Our main question is whether the reconstructed heterogeneity is accurate enough for predicting the transport of the tracers. Aside from this, we also ask if this complexity is needed at all. Therefore, results for the reconstructed aquifer are compared to the simplest reference, which is simulation with a homogeneous system. The main steps are implementation in a flow and transport model, specification of unknown transport parameters and comparison of both model results.
For simulation of the groundwater flow velocity field, the same flow model setup as for the pilot points based inversion was used. The models make use of a 3-D grid in order to accurately capture the heterogeneity of the aquifer and to account for potential transversal spreading of the tracer. Several authors recognize the importance of 3-D models [e.g., Liu et al., 2007a]. For example, Illman et al. [2008a] state “the knowledge of detailed 3-D distributions of $K$ is critical in prediction of contaminant transport”. Steady-state conditions are assumed according to the static hydraulic settings during the experiment. The fixed head boundary conditions establish a constant regional flow field, and the pumping well P0/M25 (Figure A2.3) is configured with an extraction rate of 0.3 l/s to simulate the forced gradient conditions generated in the field. In the heterogeneous model, the reconstructed $K$-field was implemented, in the homogeneous one, the arithmetic mean of $1.5 \times 10^{-4}$ m/s of the heterogeneous variant was chosen. By using steady-state models, the inverted $S_y$-fields are not utilized. However, the $K$-fields are derived by coupled inversion of $K$ and $S_y$ values under transient conditions. This way, the information content of the transient hydraulic experiment is exploited for the steady-state flow simulation during the tracer test.

The transport code MT3DMS [Zheng and Wang, 1999] is selected for solving solute transport. The transport model domain covers 10 m × 4 m × 2 m with a spatial discretization of 160 × 32 × 40 cells, summing up to 204,800 cells. Computational time on a 2015 desktop (i7, 16 GB RAM, 250 GB SSD) was approximately 5 minutes per model run using the method of characteristics.

The value ranges of two unknown parameters for transport modeling, dispersivity and effective porosity needed to be estimated. For this, two MCMC chains were run, one for the homogeneous model and other for the heterogeneous model. Both BTCs are adjusted and a summed up RMSE is computed. Each chain has a length of 3000 model runs, with a burn-in of 1500. BTCs measured at the pumping well during the tracer tests are compared with those resulting from the simulation with the obtained MCMC ensemble. In Figure A2.10, only those for the heterogeneous model are depicted. The BTCs simulated with the homogeneous model are comparable and not shown here. For both the homogeneous and the heterogeneous model, the BTCs obtained with different dispersivity and effective porosity realizations spread around the observed BTC, and no bias or other systematic error is observed.

The nearly ideal shape of the BTCs indicate that the aquifer exhibits only a moderate heterogeneity, and thus minor differences among homogeneous and heterogeneous model results exist. The homogeneous variant can appropriately capture the general form of the BTCs given properly tuned dispersivity and effective porosity values [see also Ni et al., 2009]. The heterogeneous variant is similarly suitable, which however also means that special characteristics of the BTCs are not resolved. This can be attributed to the limited resolution of the tomograms and the reconstructed fields in order to delineate local hydraulic heterogeneities relevant for solute transport. However, the limitation of an even very detailed reconstruction of aquifer heterogeneity based on a single source-receiver plane certainly plays a major role, because the tracer transport occurs in a 3-D domain. Therefore, the discrepancy between the BTCs and the
heterogeneous model simulation may also be caused by the applied lateral extrapolation from the vertical source-receiver plane. For example, the untypical spreading of the uranine BTC around the peak may be due to unseen lateral aquifer heterogeneity that strongly sidetracks the tracer. The irregular behavior of the measured sodium-naphthionate curve after around 5 days may indicate that a portion of the tracer is temporary split apart from the main plume and reaches the pumping well with a time lag. This is observed as local peak and accentuates the tailing of the earlier main tracer mass fraction. As our 3-D reconstruction is only based on extrapolation from one profile, more adjacent and ideally differently oriented profiles would be needed for capturing such lateral heterogeneities.

Lateral tracer loss is also supported by the mass recovery rates of 86% (uranine) and 50% (sodium-naphthionate). The relative low recovery observed for sodium-naphthionate in comparison to uranine during the test cannot be definitively attributed to a certain reason. As the shape of the tracer BTC does not support distinctive retardation of the tracer, sorption processes are unlikely the reason. Due to the relatively long duration of the test and moderate groundwater temperatures (10–12 °C), a microbiological degradation of sodium-naphthionate in the aquifer seems possible. Rapid microbiological sodium-naphthionate degradation was observed by Goldscheider et al. [2001] for water samples with a certain storage period, depending on storage temperature. Studies for those kinds of processes are scarce and more research effort in this direction is needed. Nevertheless, as normalized BTCs are employed for this study, a lower tracer recovery does not compromise the results.

Although, the reconstructed model is not superior to a simple homogeneous alternative in delineating the tracer BTCs, it appropriately resolves structures relevant for the transport of the tracers. This is revealed by comparison of the estimated value ranges for the MCMC ensembles. Figure 3-11 shows that the RMSE for both model variants are similar, and the effective porosities range between 0.1 to 0.27. This is the same for both, and also within the broad range of previously reported values of 0.10 to 0.25, [Meischner, 1985; Schlie, 1989; Hu, 2011; Meyer, 2011]. However, dispersions values need to be substantially higher when a homogeneous model is used. Best results are obtained for a longitudinal dispersivity of $\alpha_L = 2.67$ m for the homogeneous model. The optimal fit for the heterogeneous case is at $\alpha_L = 1.64$ m, which shows that macro-dispersive effects are simulated explicitly and correctly through the reconstructed macro-scale hydraulic heterogeneity. The value of $\alpha_L = 1.64$ m is still significant and this denotes that heterogeneities exist at a smaller scale than the resolution of the hydraulic tomography at this site and with this experimental configuration, which strongly influence the tracer spreading. By individual tracer BTC fitting, the estimated values of $\alpha_L = 1.57$ m for uranine and 1.72 m for sodium-naphthionate slightly deviate from the result of combined fitting. These differences are not judged as significant enough to identify clear differences in the tracer-specific transport or associated with the different injection levels.
A2.4 Conclusions

The presented work shows that the sequential travel-time and pilot point based approach can be applied to high resolution reconstruction of hydraulic parameters at a field site. It is demonstrated, for the only slightly heterogeneous field site that the presented procedure can identify sedimentary structures. However, comparison of model-based predictions with the tracer tests at the site reveals that the reliability of the derived aquifer model also exhibits limitations.

The tracer test delivered two BTCs, which show minor irregularities and this indicates the only moderate heterogeneity at the Stegemühle site. Therefore, even a homogeneous model can provide similarly good predictions as a heterogeneous variant with the reconstructed K-field. A main point is that crucial transport parameters, especially dispersivity, need to be set. We have not pre-specified these parameters but analyzed suitable value ranges by applying a MCMC based search. In other words, for minimizing any bias we examined model validity within these ranges. Within these degrees of freedom, the reconstructed model performs similarly well as a homogeneous one. This reflects that even though macro-scale heterogeneities are reconstructed, their combined effect on tracer spreading averages. Therefore, the tracer breakthrough curves can also be predicted by a higher integral dispersivity in a much simpler homogeneous model. However, as pointed out in the theoretical study by Ni et al. [2009], even if a homogenous model can provide an appropriate fit, it will not capture the scale-dependent increase of dispersion with transport distance [see e.g. Gelhar et al., 1992; Molina-Giraldo et al., 2011]. In contrast, by
reconstructing transport-relevant structures, their effect on macro-dispersion is explicitly simulated, and thus the heterogeneous model is more suited for predicting solute transport along shorter or longer distances.

Still, in our application case neither the homogeneous nor heterogeneous model variant perfectly predicts the recorded tracer concentrations. When measurement errors can be neglected, we interpret inconsistencies caused by unresolved lateral heterogeneity. The proposed sequential approach employs 3-D hydraulic simulation and inversion, but structures are constrained only by the vertical 2-D travel time tomograms. For improved structural reconstruction, additional tomograms between different source and receiver wells would be needed. With these, more reliable results from hydraulic parameter interpolation rather than the presented extrapolation can be expected. Aside from this, as Illman et al. (2012a) demonstrate in a sandbox experiment, the resolution by HT could be improved with a higher density of sources and receivers. As a result, so far unresolved micro-structures could be detected and the value of the dispersivity would be further decreased.

The presented coupled inversion procedure shows to further refine $K$- and $S_r$-tomograms between the investigated source and receiver wells in comparison to travel time-based inversion only. A main observation is that pilot point-based inversion reduces heterogeneity, although homogenization is not enforced through regularization. This may be due to the fact that the 2-D travel time tomograms are based on a diagnostic of early arrival times, which are accentuated by the existence of high $K$ zones or preferential flow paths. In contrast, the pilot point approach calibrates the full pressure response curves and calibrates a 3-D model, and by this a higher volume of the aquifer is referred to. Further insight could be obtained, for instance, by employing different parts of the response curves for pilot point-based inversion.

As an innovative step, it is shown that including $S_r$ in addition to $K$ in the pilot point-based inversion is beneficial for minimizing model misfit to field data. This has rarely been included in related work [e.g., Castagna and Bellin, 2009]. However, this means also doubling the number of decision variables for the optimization problem. This is potentially not desirable, as this eventually can over-parameterize the problem and enhance its ill-posedness. In fact, the generated heterogeneous aquifer model can be considered as one solution of many, and further alternative realizations fitting the data could be explored. As future work, we therefore plan to envisage the diversity of several equally probable realizations, based on $K$ with or without $S_r$ as free parameters.

A2.5 Acknowledgements

The investigations were conducted with the financial support of the Swiss National Science Foundation to the project “A field assessment of high-resolution aquifer characterization: An integrated approach combining hydraulic tomography and tracer tomography” under grant number 200021_140450 / 1, and the CCES funded project “RECORD Catchment”. The helpful
comments of the associate editor and three reviewers are greatly appreciated. Further thanks go to Gabi Moser for language corrections. All data are available by emailing the corresponding author (santos.jimenez@erdw.ethz.ch).
Bibliography


Cavanagh, A. J., and R. S. Haszeldine (2014), The Sleipner storage site: Capillary flow modeling of a layered CO₂ plume requires fractured shale barriers within the Utsira Formation, Int. J.


Global CCS Institute (2015), The Global Status of CCS: 2015, Summary Report, Melbourne, Australia,


Gorecki, C., J. Hamling, J. Ensrud, E. Steadman, and J. Harju (2012), Integrating CO$_2$ EOR and
Hovorka, S. D. et al. (2011), Monitoring a large volume CO₂ injection: Year two results from...


Jackson, M. J., and D. R. Tweeton (1996), 3DTOM: Three-dimensional geophysical tomography,


Lu, C., and P. C. Lichtner (2005), PFLOTRAN : Massively Parallel 3D Simulator for CO2 Sequestration in Geologic Media, Carbon N. Y.


McCall, W., T.M. Christy, T. Christopherson, and H. Isaacs (2009), Application of direct push methods to investigate uranium distribution in an alluvial aquifer, Ground Water Monitoring Remediation, 29(4), 65–76.


Nowak, W., S. Tenkleve, and O. A. Cirpka (2003), Efficient Computation of Linearized Cross-


Ringrose, P., M. Atbi, D. Mason, M. Espinassous, Ø. Myhrer, M. Iding, A. Mathieson, and I. Wright (2009), Plume development around well KB-502 at the In Salah CO2 storage site, First Break, 27, 85–89.


Sun, A. Y., and J. P. Nicot (2012), Inversion of pressure anomaly data for detecting leakage at
geologic carbon sequestration sites, Adv. Water Resour., 44, 20–29,
Sun, A. Y., J. Lu, B. M. Freifeld, S. D. Hovorka, and A. Islam (2016), Using pulse testing for
leakage detection in carbon storage reservoirs: A field demonstration, Int. J. Greenh. Gas
Sun, R., T.-C.J., Yeh, D. Mao, M. Jin, W. Lu, and Y. Hao (2012), A Temporal sampling strategy
Tambach, T. J., M. Koenen, L. J. Wasch, and F. van Bergen (2015), Geochemical evaluation of
CO2 injection and containment in a depleted gas field, Int. J. Greenh. Gas Control, 32, 61–80,
Tasianas, A., L. Mahl, M. Darcis, S. Buenz, and H. Class (2016), Simulating seismic chimney
structures as potential vertical migration pathways for CO2 in the Snøhvit area, SW Barents
Sea: model challenges and outcomes, Environ. Earth Sci., 75(6), 504, doi:10.1007/s12665-
016-5500-1.
Tong, J., B. X. Hu, and J. Yang (2010), Using data assimilation method to calibrate a heterogeneous
conductivity field conditioning on transient flow test data, Stoch. Environ. Risk Assess.,
Trautz, R. C. et al. (2013), Effect of Dissolved CO2 on a Shallow Groundwater System: A
CO2 storage in a depleted gas field: An overview of the CO2CRC Otway Project and initial
monitoring of CO2 injection into a depleted gas reservoir - Otway Basin Pilot Project,
Vanorio, T., A. Nur, and Y. Ebert (2011), Rock physics analysis and time-lapse rock imaging of
geochemical effects due to the injection of CO2 into reservoir rocks, Geophysics, 76(5), O23-
Vasco, D. W., and K. Karasaki (2001), Inversion of pressure observations: An integral formulation,
Vasco, D. W., and K. Karasaki (2006), Interpretation and inversion of low-frequency head
Vasco, D. W., and S. Finsterle (2004), Numerical trajectory calculations for the efficient inversion
10.1029/2003WR002362.
Vera, V. C. (2012), Seismic modelling of CO2 in a sandstone aquifer, Priddis, Alberta. Doctoral
dissertation, University of Calgary.
Exploration 3, 323-334.
Edge, 19(9), 944, doi:10.1190/1.1438762.
Vihelmsen, T. N., S. Christensen, and S. W. Mehl (2012), Evaluation of MODLFLOW-LGR in
connection with a synthetic regional-scale model, Ground Water, 50(1), 118-132.
Vilarrasa, V., D. Bolster, S. Olivella, and J. Carrera (2010), Coupled hydromechanical modeling
of CO2 sequestration in deep saline aquifers, Int. J. Greenh. Gas Control, 4(6), 910–919,
Virieux, J., C. Flores-Luna, and D. Gibert (1994), Asymptotic theory for diffusive electromagnetic
...
Yeh, T. J., K. Hsu, and C. Lee (2004), On the possibility of using river stage tomography to characterize the aquifer properties of the Choshuishi Alluvial Fan, Taiwan. AGU Fall Meeting Abstracts.


Curriculum Vitae

Personal
Name: Linwei Hu
Date of Birth: November 19, 1986
Citizenship: Hainan Province, China

Education
2013/01-present  PhD student of Department of Earth Sciences, ETH Zurich, Switzerland
Under supervision of Prof. Simon Low
2008/10-2011/11  M. Sc, University of Göttingen, Germany
Under supervision of Prof. Thomas Ptak
2003/09-2007/07  B. Sc, Nanjing University of Information Science and Technology

Journal papers


**Conferences**


**Award**

Outstanding Student Poster (OSP), in EGU 2016