Mitigating the Problem of Non-uniqueness in Fluid-Flow Modeling

by

Saleh Mohammed Al Nasser

Submitted to the Department of Earth Atmospheric and Planetary Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2023

© Massachusetts Institute of Technology 2023. All rights reserved.

Author.....
Department of Earth Atmospheric and Planetary Sciences
January 13, 2023

Certified by.....
Frank Dale Morgan
Professor of Geophysics
Thesis Supervisor

Accepted by.....
Robert D. Van der Hilst
Schlumberger Professor of Earth Sciences,
Head of the Department of Earth, Atmospheric, and Planetary Sciences

Mitigating the Problem of Non-uniqueness in Fluid-Flow Modeling

by

Saleh Mohammed Al Nasser

Submitted to the Department of Earth Atmospheric and Planetary Sciences
on January 13, 2023, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Modeling fluid flow in porous media is a valuable and essential tool for developing underground resources such as hydrocarbon reservoirs, groundwater aquifers, or CO_2 sequestration projects. The modeling, if done accurately, can provide a reliable forecast of future fluid behavior. However, the properties of the porous media and the correct solutions to the physics equations describing the macroscopic fluid flow are essential to ensure accurate modeling and, consequently, reliable forecasts. Therefore, the need to discretize the porous mediums into a large number of grids is often crucial to capture the observed data's behavior. And because the data has a low abundance spatially, it is impossible to model the fluid flow uniquely. In the thesis, we study ways to transform the modeling of fluid flow in porous media into a less non-unique problem by exploring different models and data spaces. By reducing the number of grids, we quantitatively demonstrate the possibility of producing more accurate representations of reservoirs. Also, through the resolution matrix analysis and the use of Shannon information entropy, we developed a method to acquire data adaptively for an optimum survey design. Additional data sets from self-potential or seismic surveys have complemented the fluid flow data in different joint inversion methods. Using self-potential data allows the detection of fractures with higher confidence. The seismic data was used in a cross-gradient joint inversion scheme to constrain the inversion of fluid flow data. The joint inversion helped in around 16% reduction in the seismic velocity root-mean-square-error (RMSE) and almost 26% decrease in the permeability RMSE.

Thesis Supervisor: Frank Dale Morgan
Title: Professor of Geophysics

Acknowledgments

I would like to convey my gratitude and appreciation to my thesis advisor, Prof. Frank Dale Morgan. His continuous support and care are beyond measure. Prof. Dale treated me not only as his student but also as his son. The door to his office is always open whenever I need academic or personal advice. I owe him tremendous respect for the person I became in the years I spent at MIT.

My appreciation is extended to my thesis committee members, Prof. Nafi Toksoz, Prof. Matej Pec, Prof. Julien de Wit, and Dr. Mustafa Al Ali. Each one of them has influenced me in unique ways. Their suggestions and feedback made my research much richer, and my dissertation is something I can be proud of having written. Nafi's wisdom and powerful comments have always provided me with a solid push to excel. I am deeply thankful to Matej for his continuous encouragement. I had the pleasure of working with Matej on many experiments at his lab, which contributed greatly to my understanding of experimental science. Julien sees problems from a different perspective, and his creative thoughts broaden my thinking process. Mustafa has a unique way of forcefully lifting my spirit and raising my confidence. He was one of the earliest supporters of my study at MIT.

I would also like to thank and acknowledge the professors who taught me at MIT; Rob van der Hilst, Tom Herring, Bradford Hager, Sai Revela, Ruben Juanes, Laurent Demanet, and others.

I acknowledge my colleagues at EAPS and ERL for their support and kindness. Each one has given me a lot of time and effort. Thanks to EAPS and ERL administrations for all the help they provided.

I would like to thank Saudi Aramco for the full scholarship and financial support. My thanks are extended to the members of the Geophysical Technology Team at EXPEC Advanced Research Center.

I am grateful to my wife, Hawra Alhashem, for being part of the journey. She believed in me during the moments when I doubted myself. Her enormous love, care, and support thorough out the years have made the experience magnificent.

I owe my most profound appreciation and gratitude to my parents (Mohammed Al Nasser & Fahima Al Shilaty) and my siblings for their limitless love and care. My only hope is to make them proud. This work is dedicated to loved ones who I lost and who contributed a lot to the development of my life.

Contents

List of Figures

List of Tables

Chapter 1

Non-uniqueness in Fluid-flow Model Representation

Even as the world transitions toward cleaner and renewable energy resources, there will be a continued dependence on fossil fuels. Because fossil fuels will serve as primary energy sources over the next decade, we, as scientists, are on a mission to provide reliable forecasts of the future performance of hydrocarbon reservoirs to drive policies. Further scientific findings regarding reservoirs will aid policymakers in making informed decisions about reservoir development and performance.

Hydrocarbons still contribute to almost 80% of the world's energy resources. Figure 1-1 displays the global energy consumption by source as of 2021. The consumption of oil alone in 2021 contributed to almost $51,170$ terawatt-hours. It is helpful to understand this number within the context of our daily lives. According to the U.S. Energy Information Administration, the average consumption for a house in the U.S. was around 30 kilowatt-hours per day.

Even with the increasing dependence on renewable and green resources in recent years, the demand for fossil fuel energy is not showing decreasing sign. Given the continued high reliance on hydrocarbons, it is evident that policymakers and the public

need reliable estimates and forecasts of supply and demand. Therefore, there is a need to continue analyzing and improving the processes of hydrocarbon production and development.

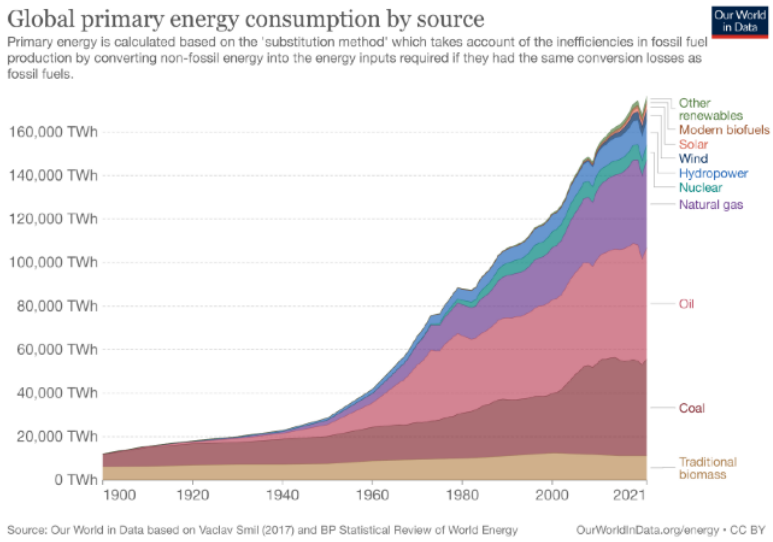

Figure 1-1: Primary energy consumption by sources. In 2021, oil alone contributed to almost 51,170 terawatt-hours.

The exploration and development of underground resources, including hydrocarbons, frequently involve multidisciplinary partnerships. Typically, the process begins with a geologist mapping the structural geology and assessing the subsurface for the presence of hydrocarbons, water, or minerals. Then, geophysicists will use various techniques (e.g., seismic, gravity, and resistivity) to image the subsurface. Each technique will aid in the comprehension of subsurface characteristics. Seismic data, for example, can provide structural information about subsurface geology and the existence of potential hydrocarbon traps. In contrast, surface gravity measurements can yield information regarding subsurface density. Higher density, for instance, may indicate the presence of heavy minerals, whereas lower density may indicate the presence of fluid.

Although most surveys begin with a geologist identifying potential exploration re-

gions in the field, the order in which we conduct surveys has flipped in recent years. Due to advancements in geophysics, geophysical surveys and geological interpretations have become standard operating procedures. However, geoscience investigation is frequently the initial step in exploring underground resources.

Engineers and scientists rely on the subsurface images constructed by geoscientists to drill the first appraisal well for hydrocarbon exploration. If hydrocarbon is discovered in the first well, multiple wells will be drilled to delineate the reservoir's extent. The reservoir will then enter the phase of production and monitoring. At this stage, the reservoir's production rate and pressure are constantly monitored to optimize its performance, which requires knowledge of the reservoir's physical properties and fluid flow behavior. In addition to monitoring the reservoir, the ultimate goal is to forecast its performance in the future. Predicting reservoir behavior requires knowledge of the fluid flow's physics and the reservoir's geological properties. Consequently, we must first model the fluid flow in reservoirs to comprehend the historical production rate and then predict future performance.

Fluid flow modeling of a hydrocarbon reservoir is a powerful tool for predicting its future performance. Modeling reservoirs requires an understanding of the fluid behavior in porous media. For an engineer to achieve a prediction of a certain reservoir's future performance, a set of requirements are essential. For instance, an improved understanding of the geological nature of the reservoir is needed to develop a model with hydraulic parameters such as porosity and permeability. Also, there is a need to develop a stable and accurate solution to the partial differential equations that describe the fluid flow in porous media. If such an improved understanding is achieved, engineers can run multiple models of the fluid flow in a reservoir to match the performance of the existing one, thereby introducing the possibility of predicting future performance.

Forecasting a reservoir's performance is usually achieved with the standard history-matching approach. This approach relies on determining the values of those parameters, such as porosity and permeability, that will reproduce the observed behavior

of a reservoir. By adjusting such parameters, the simulator will match the historical production of the reservoir, thus identifying one set of parameters for the intended outcome. More specifically, the reservoir is discretized numerically into grid blocks. Each grid block has physical properties such as permeability. Assuming there is a stable solution to the fluid flow equations, the simulator can numerically simulate the fluid production rate and pore pressure changes within the reservoir. The physical properties of each grid block are adjusted until the simulation of the production rate matches the historical observation. Once the historical performance is matched, the future production rate is predicted.

A significant issue in the history matching approach is raised when exploring the differences between the simulated model and actual geological data. For example, there may be cases in which the misfit is such that the differences between the observed data and the data generated by the model are not found to be significantly different, but the inverted model is far from the geological realization. In such cases, history matching would be confirmed successful, despite being potentially inaccurate.

The inverse problem of fluid flow characterizes the reservoir hydraulic parameters (e.g., porosity, permeability, and saturation) using pressure readings obtained from the boreholes. Despite several proposed solutions and approaches, the solution to the fluid flow inverse problem is still unclear. For example, Carrera and Neuman (1989) have attributed the discrepancy in solving the inverse problem to non-uniqueness, non-identifiability, and instability according to Hadamard's identification of ill-posed mathematical models. In 1902, Hadamard stated that mathematical models of physical phenomena would be ill-posed if any of the following properties were not satisfied: 1) The problem has a solution. 2) The solution is unique. 3) The solution depends continuously on the data. The fluid-flow inverse problem appears ill-posed under at least one of Hadamard's three criteria. Geophysical surveys tend to result in discontinuous data from measurement errors, which suggest that the solution is unstable. In 1973, Neuman attributed an infinite number of possible solutions for the inverse problem to

errors in the data and the governing equations. The lack of data or the discontinuous nature of the data can lead to non-unique solutions.

It is also helpful to consider the analogies between the inverse problem's solution and a linear system of equations where the matrix is singular or rank-deficient. From a mathematical perspective, consider a linear function (F) that operates on a parameter vector (m) to generate data in vector d (Fm=d). If m«d, there exists a vector x in the null space of F that satisfies Fx=0. Then, a model (m+x) satisfies the data as follows: F(m+x)=Fm+Fx=d. Therefore, there are many combinations of models that can fit the data.

Others, including McLaughlin and Townley (1996), have attributed the ill-posed quality of the inverse problem of the flow parameters to the low sensitivity of the data to the parameters. The researchers stated that the steady-state potential data is relatively insensitive to the spatial variation in the hydraulic parameters. However, Nelson (1960) has demonstrated that it is still possible to uniquely obtain the hydraulic parameter through inversion if the spatial data is known at every stream tube under a steady flow.

A potential solution to the inverse problem features the application of the parameterization approach (Neuman, 1973; Sun and Yeh, 1985; Eppstein and Dougherty, 1996; McLaughlin and Townley, 1996), where parameters are assumed to be constant at specific blocks or segments of the aquifer. The parameters are then estimated through an optimization algorithm that provides a good fit for the measured data. However, a single model may not produce representative data, a notion that motivated Hernánez et al. (1997) to use conditional simulation methods. The methods are intended to identify a set of probable parameters rather than a single unique solution. Moreover, Yeh (1986) and Carrera et al. (2005) have provided a detailed introduction to solving the inverse problem for groundwater hydrology.

Each chapter of the current thesis presents a unique method for solving the inverse modeling of fluid flow in reservoirs. We first introduce the basic methodology for

forward and inverse fluid flow modeling in porous media. We hope that by the end of this chapter, the reader will understand the basic methods and the relevant issues in inverting the reservoir parameters.

1.1 Forward Modeling

Forward modeling is the process of simulating the behavior of fluid as it flows in porous media. In the 19th century, Henry Darcy was the first to work on describing the fluid flow in porous media mathematically. Darcy studied the volumetric flow rate out of a tank filled with sand and water and concluded that the rate is proportional to the pressure gradient. With a slight alteration to Darcy's notations, Darcy's fluid velocity for a single-phase fluid can be written as the following:

$$v = -\frac{k}{\mu}\nabla P \tag{1.1}$$

where v is the fluid velocity in (m/s), k is permeability in (m^2), μ is the viscosity in (cp), and P is pressure in (psi).

Generally, writing a fluid mass balance equation for a cube of rock with the Darcy velocity leads to a diffusion-type equation describing the single-phase fluid flow in porous media. For a slightly compressible fluid, the pressure diffusion equation in oil fields units can be written as follows:

$$\frac{C_1}{\mu}\nabla \cdot ([K]\nabla P(x,y,z,t)) = \frac{\phi C_t}{C_2}\frac{\partial P(x,y,z,t)}{\partial t} + q(x,y,z,t) \tag{1.2}$$

where μ is the fluid viscosity in (cp), $[K]$ is the permeability tensor in milliDarcy(mD), P is pressure in (psi), ϕ is porosity, C_t is the total compressibility in psi^{-1}. $C_1 = 1.127 \times 10^{-3}$ and $C_2 = 5.615$ are the conversion factors to oil fields units. The source/sink term q has a unit of (STB/d). A full definition of the quantities used in the flow equation

is found in Table1.1

Parameters	Oilfield Unit	SI Unit	Conversion to Oilfield Unit (multiply SI unit)
Permeability (K)	mD	m^2	9.9×10^{16}
Compressibility (C_t)	psi^{-1}	Pa^{-1}	6897
Viscosity (μ)	cp	$Pa \cdot s$	1000
Rate (oil) (q)	STB/d	m^3/s	5.4×10^5
Pressure (P)	psi	Pa	1.45×10^{-4}
Dimensions (x,y,z)	ft	m	3.3

Table 1.1: The conversion table from SI unites to oilfield units.

Different numerical methods, such as finite methods, can solve the parabolic partial differential eq1.2 (PDE). In this thesis, we adopt an implicit finite volume method to approximate the solution to the PDE. The finite volume method discretizes the physical reservoir models into voxels (Figure 1-2), with each one representing the fluid flow parameters (K,μ,ϕ,q, and C_t), and solves for the time-dependent pressure $P(x, y, z, t)$ given the initial and boundary conditions. The complete discretized equation for each voxel is as follows:

$$\begin{aligned} C_t\phi\frac{\delta x\delta y\delta z}{\delta t}(P^{n+1}_{i,j,k} - P^n_{ijk}) - T^x_{i-\frac{1}{2},j,k}P^{n+1}_{i-\frac{1}{2},j,k} - T^y_{i,j-\frac{1}{2},k}P^{n+1}_{i,j-\frac{1}{2},k} - T^z_{i,j,k-\frac{1}{2},k}P^{n+1}_{i,j-\frac{1}{2},k} \\ +(T^x_{i-\frac{1}{2},j} + T^x_{i+\frac{1}{2},j} + T^y_{i,j-\frac{1}{2}} + T^y_{i,j+\frac{1}{2}} + T^z_{i,j,k-\frac{1}{2}} + T^z_{i,j,k+\frac{1}{2}})P^{n+1}_{i,j,k} \\ +T^x_{i+\frac{1}{2},j}P^{n+1}_{i+\frac{1}{2},j,k} - T^y_{i,j+\frac{1}{2}}P^{n+1}_{i,j+\frac{1}{2},k} - T^z_{i,j,k+\frac{1}{2}}P^{n+1}_{i,j,k+\frac{1}{2}} + q\delta x\delta y\delta z = 0 \end{aligned} \tag{1.3}$$

where T is the transmissivity at the edges of the voxels and can be estimated by:

$$T^x_{i-\frac{1}{2},j,k} = \frac{K_{i-\frac{1}{2},j,k}}{\mu}\frac{\delta y\delta z}{\delta x} \tag{1.4}$$

The K in the transmissivity equation is permeability at the edges and estimated by the harmonic average of the permeability of two adjacent voxels:

$$K_{i-\frac{1}{2},j,k} = \frac{2\delta x K_{i-1,j,k}K_{i+1,j,k}}{\delta x K_{i-1,j,k} + \delta x K_{i+1,j,k}} \tag{1.5}$$

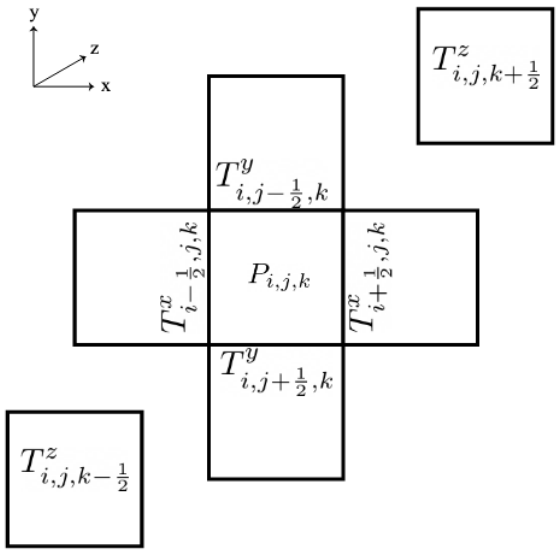

Figure 1-2: A diagram showing the distribution of grids used for discretization.

The forward modeling code solves a system of equations like the one written in eq1.3 at each time step. The pore pressure for each grid at multiple time steps is the output of the forward modeling. The solution to the equation system can be computationally expensive, depending on the model's size. For the purposes of the current thesis, our models were minimized in size by utilizing the backslash feature in MATLAB to solve the system efficiently. For a more extensive system representing a large number of grids ($\approx 10^6$), the system of the equations can be solved by iterative methods such as the Gauss-Seidel method and the Jacobie method.

Here, we ran a forward modeling problem to illustrate the solution and cross-validate the solution with an approximate analytical solution to the PDE. Consider a homogenous reservoir model with an extent of 1000 m in x and y dimensions and a uniform thickness of 10 m. Solving the fluid flow equation on a 51 by 51 grid numerically simulates the fluid flow. The reservoir has homogenous isotropic permeability, $K_x = K_y = 60$ mD, and constant porosity at 0.15. The fluid is at a single phase at all pressures. The system's total compressibility is $c_t = 12e - 6$ 1/psi. The fluid's viscosity is $\mu = 1.5$ cp. One producing well at the model's center is located at the grid

block (26,26). The producing well has a constant pumping rate of 100 STB/day. The reservoir's initial pressure is P=4000 psi. Boundaries are held at a constant pressure on all sides, P=4000 psi. A summary of the parameters and values is presented in Table 1.2. In Figure 1-3, the pressure distribution in the reservoir is shown as a result of simulating the fluid flow for 100 hours. The pressure is lowest at the production well and radially increases toward the reservoir's boundaries. The curve in Figure 1-3b illustrates the pressure drawdown as a function of distance away from the well. We expect the pressure to stabilize at a far distance from the well at the initial pressure.

Parameters	value
K(mD)	60
phi	0.15
μ(cp)	1.5
C_t(1/psi)	12×10^{-6}
q(m^3/day)	50
Initial P	4000 psi
Boundary P	4000 psi
Time (hours)	100
Dimensions[x,y,z] (m)	[1000,1000,10]

Table 1.2: The values of the parameters used for simulating the fluid flow.

To validate our simulation solution, we compared the drawdown curve to an approximate solution to the pressure drawdown away from the wells. Matthews & Russull, in 1967, estimated that the following equation could solve the pressure curve away from a well:

$$P(x,t) = P_i + [\frac{70.6Q\mu}{KH}]Ei[\frac{-948\phi\mu C_t x^2}{Kt}] \tag{1.6}$$

Figure 1-3.c displays the approximate solution in a solid red line on top of the simulation results. There is a slight discrepancy between the two solutions because of the size of the grids. We expect the difference between the two solutions to decrease as the grid size is reduced and the number of grid blocks increases. The spatial error analysis for the numerical solution is estimated to be of second-order accuracy. Therefore, we have performed the same simulation on different grids to see how the differences

vary. Figure 1-4 represents the root mean square error (RMSE) between the simulated pressure (P_s) and the approximated pressure (P_a) as the grid size changes.

$$RMSE = \sqrt{\frac{\sum(P_s - P_a)^2}{N}} \tag{1.7}$$

Given the initial and boundary conditions, the forward modeling of fluid flow in porous media is a well-posed Cauchy problem, meaning it has a unique and stable solution.

1.2 Inverse Modeling

Inverse modeling is the opposite of forward modeling. Rather than simulating the fluid pressure in a reservoir due to known parameters, we estimated the model's parameters from observed pressure data in inverse modeling. The model parameters $\boldsymbol{m}$ describing the system are related to the observed data d through the forward model operator $\mathbf{F}$ such that:

$$\mathbf{F}(\boldsymbol{m}) = d. \tag{1.8}$$

The main objective of inversion is to determine $\boldsymbol{m}$ from the observed data d. In our cases, $\boldsymbol{m}$ is a vector (length m) of the parameters' types multiplied by the total number of grids, and d is a vector (length n) of observed data. The numbers m and n depend on the scale of the model and the number of data points. Consider a reservoir model that has 100 grids, and there is an unknown permeability and porosity for each grid. Therefore, $\boldsymbol{m}$ will be a vector of length 200. If the same reservoir has two boreholes to measure pressure every day for ten days, then vector d will have a size of 20.

A solution to the inverse problem exists when a minimum residual exists between the observed and the 'model' data. A commonly used measure of the misfit is the $\mathbf{L2} - norm$

$$O(m) = \| \left(F(m) - d\right) \|^2 \tag{1.9}$$

However, the inverse problem is typically ill-posed, resulting in many models fitting the data. The objective function that only relies on the quantitative misfit as in eq1.9 is not enough to describe the reservoir models, especially for many parameters. A simple example demonstrating the non-uniqueness in the fluid flow modeling is shown in Figure 1-5. The two models in Figure 1-5 have shuffled permeability values but resulted in an exact data match. Therefore, we often adjust the objective function by

adding a regularization term to constrain the inverse process.

$$O(m) = \| \ (F(m) - d) \ \|^2 + \alpha^2 \ \| \ L\delta m \ \|^2 \tag{1.10}$$

Here, L is a Tikhonov matrix (Tikhonov,1977), and λ is a penalization term. For our fluid flow problem, we have found that the second difference matrix is the best choice for L to resolve an accurate reservoir model. Many methods can be used to minimize the square data misfit with regularization. Each method is unique, based on the specific data and solution. Oliver and Chen (2011) reviewed the algorithms used to solve the history-matching problem. Each algorithm is used for different purposes. Manual history matching can be used to create multiple realizations for uncertainty quantification, but it most likely results in unrealistic geological realism and, therefore, limits the prediction power. The evolutionary methods, such as a genetic algorithm, are considered the standard approach for assisted history matching, with the number of variables being small. However, it has a slow rate of convergence. Another method, such as the neighborhood algorithm, approximates the posterior probability density function but also has a slow convergence. In this thesis, we have adopted a gradient-based method, an efficient method for finding the local minimums. The minimization of the objective function in eq1.10 is achieved by an iterative least-squares method to estimate the parameter perturbation vector $\delta \boldsymbol{m}$:

$$\delta \boldsymbol{m} = (\mathbf{J}^T \mathbf{J} + \lambda^2 L^T L)^{-1} \mathbf{J}^T (F(m) - d). \tag{1.11}$$

The Jacobian "$\mathbf{J}$" matrix is numerically calculated as the partial derivatives of the data for the parameters:

$$\mathbf{J} = \frac{\partial d}{\partial m} \tag{1.12}$$

For the rest of this thesis, we have rescaled the inversion problem on a logarithmic

scale as it has outperformed the non-logarithmic approach. The logarithmic rescaling transforms the forward problem in eq1.8 to:

$$\mathbf{F}(\ln(\boldsymbol{m})) = d \tag{1.13}$$

With this transformation, the Jacobian "**J**" has slight alteration and can be computed numerically by:

$$\mathbf{J} = \frac{m}{d}\frac{\partial d}{\partial m} \tag{1.14}$$

Finally, in each iteration, the new parameter vector m_{n+1} is updated by the perturbation vector($\delta \boldsymbol{m}$) where n is the iteration number

$$m_{n+1} = m_n e^{\delta m} \tag{1.15}$$

We followed the Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1964) method to update each iteration's regularization factor λ.

Here, we demonstrate the inverse modeling process on a synthetic example to verify and evaluate the method. The reader should notice that any reference to the "true" model in the following examples refers to the synthetic model used to simulate the pressure data.

The first step was building a synthetic reservoir model to simulate the observed data. The areal dimension of the reservoir model is 1000 m $\times$ 1000 m and has a uniform thickness of 10 m. The discretized space has 31 $\times$ 31 grids in x and y dimensions. The model has one producing well at the central grid, pumping at a rate of 100 m^3/day, and two observation wells at grids (5,5) and (26,26), respectively. As shown in Figure 1-6, we have designed the synthetic model to have two permeability anomalies with a constant background permeability. The model has a homogenous porosity. Simulating the fluid flow in the model results in the pressure change at the three wells, as illustrated in Figure 1-6. Each pressure curve has 11 data points in time with one measurement

per 5 days. The data displays a pressure decrease at all well locations as the production well is pumped at a constant rate. The initial pressure at the reservoir is 6000 psi. The four boundaries around the reservoir are at constant pressure- at 6000 psi.

The pressure curves are then used in the inversion algorithm to estimate the permeability structure. We started with a homogenous structure as a starting guess for the permeability. We performed the inversion algorithm for multiple iterations until the solution converged to an unchanged root mean square error between the observed pressure and model curves. The inversion is terminated when the RMSE difference between two sequential iterations is small ($< 10^{-6}$). Figure 1-6 summarizes the result of inversion for the permeability structure. The solution has converged in almost 120 iterations resulting in an inverted image showing the two permeability anomalies without delineating the structural shapes. We can visually compare the final inverted image with the actual model to verify our results. However, we need to compute the model resolution matrix to quantitatively estimate the inverted model's accuracy. The following chapters of this thesis will discuss the use of the resolution matrix in multiple settings to quantify and improve the quality of the inversion.

1.3 Thesis Outline

The current thesis explores the non-uniqueness associated with fluid flow in porous media and how non-uniqueness can be reduced by adopting various techniques. A range of scenarios and examples will be presented throughout the thesis to illustrate the effectiveness of the discussed methods.

In Chapter 2, we will demonstrate the effect of model space and data space on inverting the pressure data for the parameters. The synthetic examples presented there will show how the fluid flow inverse problem is highly non-unique in that multiple settings result in different realizations of the reservoir parameters. In Chapter 3, we will attempt to reduce the complexity of reservoir modeling by reducing the number

of grids and also covers optimizing the wellbore locations to improve the inversion and reduce the non-uniqueness. In Chapter 4, we utilize additional geophysical data to map the permeability structure and image fractures in reservoirs. Chapter 5 discusses applying the cross-gradient method to invert the reservoir permeability jointly from the seismic and fluid flow data. Finally, Chapter 6 will summarize and discuss the overall results.

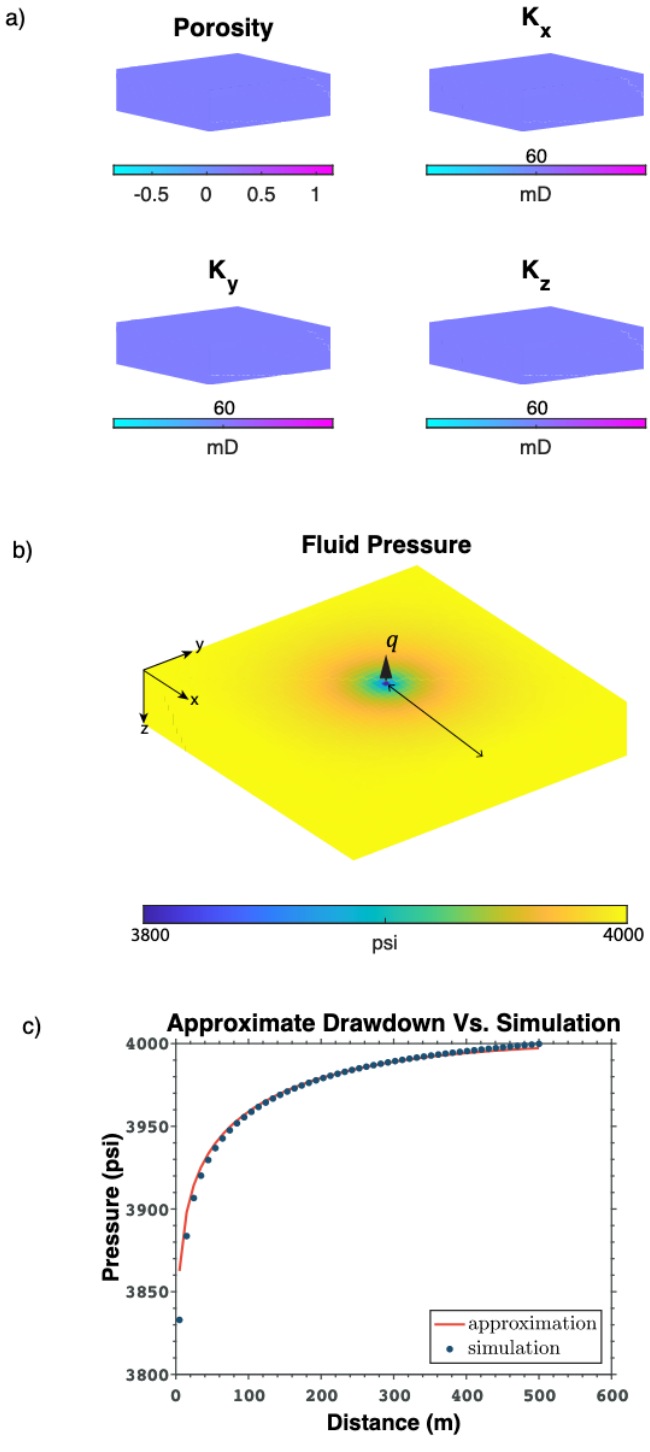

Figure 1-3: a) a 3D model with a homogenous permeability and porosity. b) fluid pressure at the surface after simulating fluid flow for 100 hours. q is the location where the reservoir model is produced at a constant rate. The black line is placed where the pressure drawdown is plotted with distance in Figure (c). c) a pressure drawdown curve for the approximate solution and the simulation.

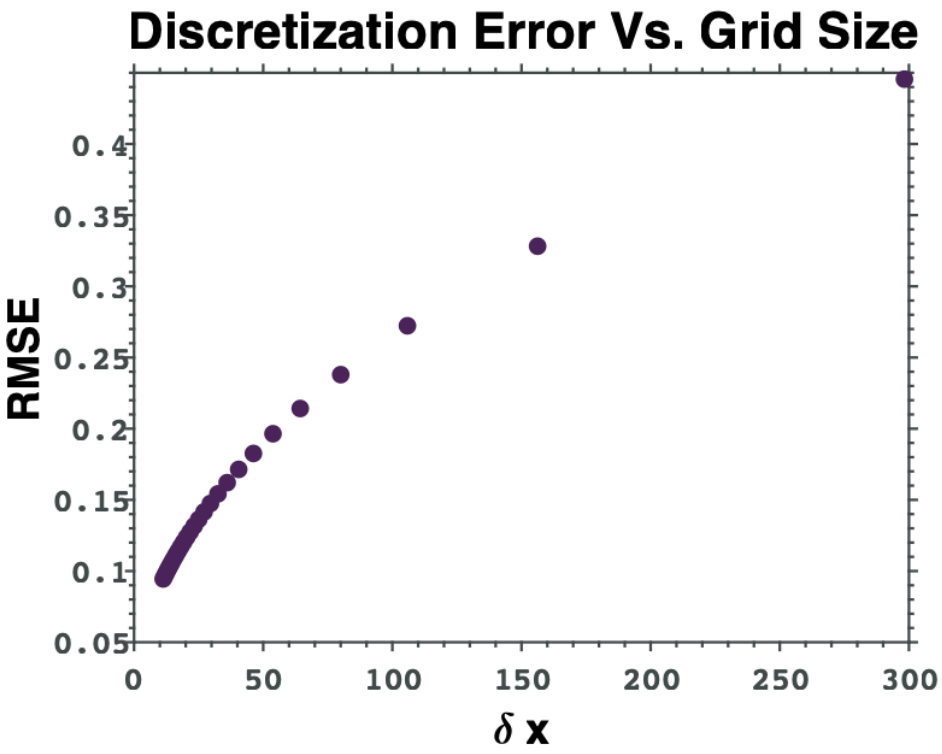

Figure 1-4: The RMSE error displayed at different discretization grid sizes.

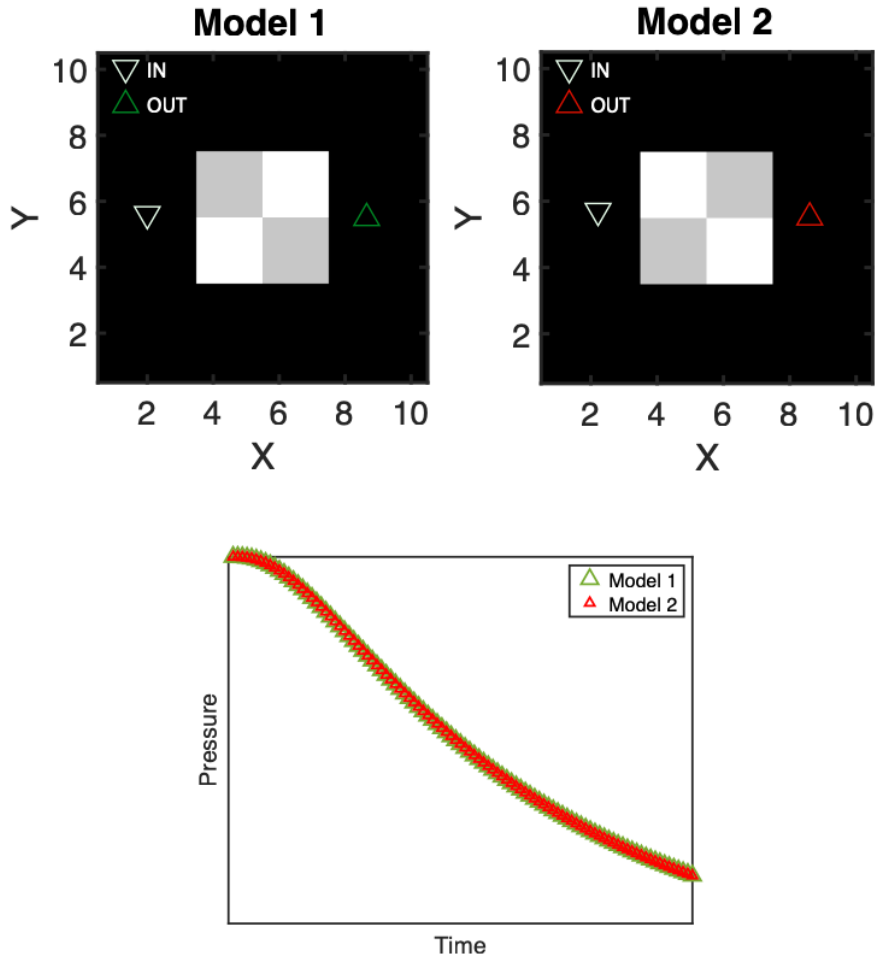

Figure 1-5: Two models with different permeability values. Downward and white triangles represent fluid injecting, and upward triangles represent fluid pumping out, colored differently for model 1 and model 2. The curves at the bottom figure show the pressure measured at the red and green triangles.

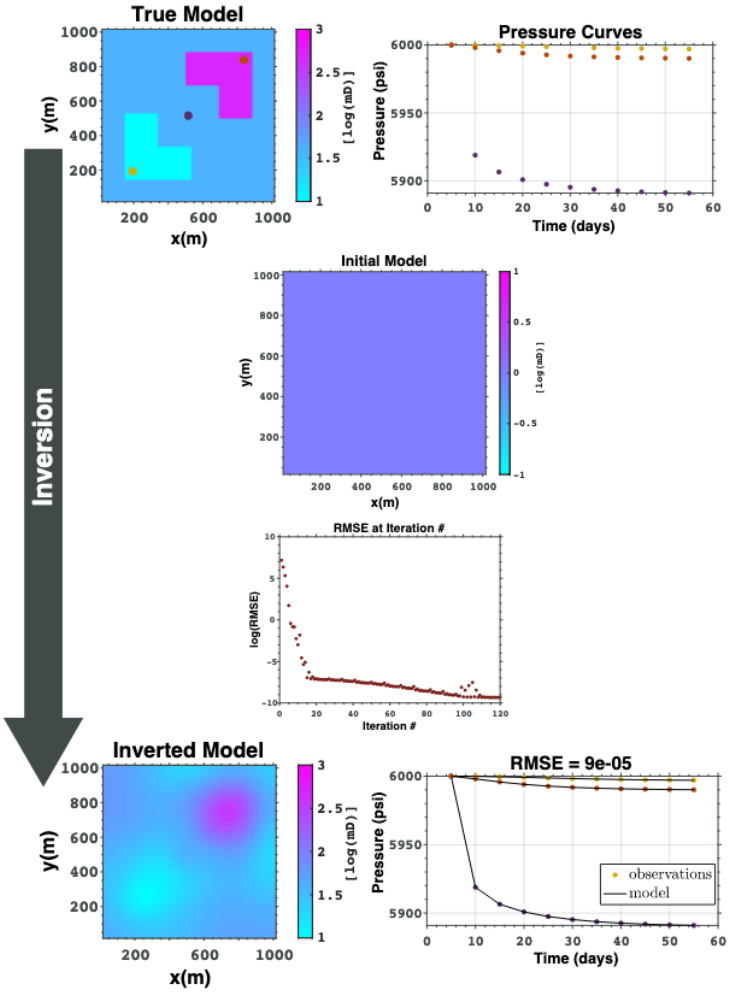

Figure 1-6: A summary of an inversion example. The true model is shown on the top with three colored circles where the pressure curves are observed. The central well acts as a production well; the other two are observation wells. The pressure curves on the top are colored based on the location of the well. The middle figures show the initial model used in the inversion and, below it, the RMSE at each iteration in the inversion. The final image at the bottom displays the inverted image at the last iteration with the fitted pressure curves beside it.

Chapter 2

Effects of Data and Model Spaces on Fluid-flow Parameters Inversion

2.1 Abstract

The inversion of the reservoir parameters is highly non-unique due to many factors related to the data and models. Because the number of observation wells where the pressure data is measured is spatially sparse, there can be many realizations of the reservoir that fit the data. One way to quantify the quality and the accuracy of the invasion is by computing the resolution matrix. The resolution matrix is implicitly computed from the generalized inverse matrix. We can see in different cases that the measured quality of the resolution matrix varies with changing the data's location, quality, and number. The number of models is also affecting the accuracy of the inversion results.

2.2 Introduction

The forward modeling of fluid flow in porous media simulates the pressure distribution across a reservoir resulting from producing or injecting fluid. In contrast, inverse modeling estimates the fluid flow model by inverting the pressure distribution data. The quality of the estimated inverted model is quantified by computing the model resolution matrix. Wiggins, in 1972, applied parameter resolution analysis to surface wave problems and claimed that the inversion studies are incomplete if such analysis is not considered. The resolution matrix R indicates how resolved the inverted model is compared to the actual model. The resolution matrix is estimated by:

$$R = J^{-g}J \tag{2.1}$$

where J^{-g} is the generalized inverse solution. R has a dimension m x m, where m is the number of parameters. If R is an identity matrix, we have a well-resolved solution; otherwise, it might indicate a bad resolution (Figure 2-1). The resolution matrix in Figure 2-2, displays different batches of value concentrations. The concentration of values found near well locations indicates the most resolved parameters. The parameters close to the location where the data is observed have higher resolution than the rest. We can measure the overall quality of the resolution matrix by computing the Backus-Gilbert spread function (Backus and Gilbert, 1967, 1968).

$$Spread(R) = \sum_{i=1}^{m}\sum_{j=1}^{m}[R_{ij} - I_{ij}]^2 \tag{2.2}$$

On the other hand, the information matrix is a qualitative measure of the predicted data. The information matrix has a size of n $\times$ n and displays the significance of data points. The computation of the information matrix is as follows:

$$N = JJ^{-g} \tag{2.3}$$

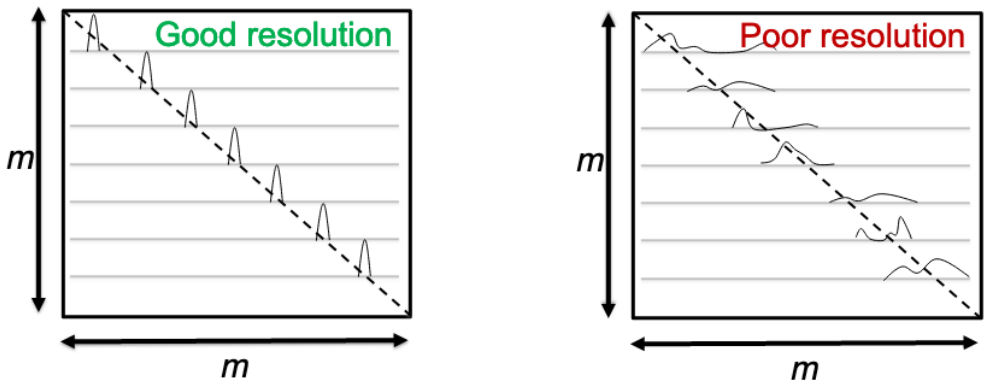

Figure 2-1: This figure is a schematic diagram to differentiate between a good and poor resolution matrix. The good matrix has most of its values at the diagonal, while the poor matrix has more spread-out values around the diagonal.

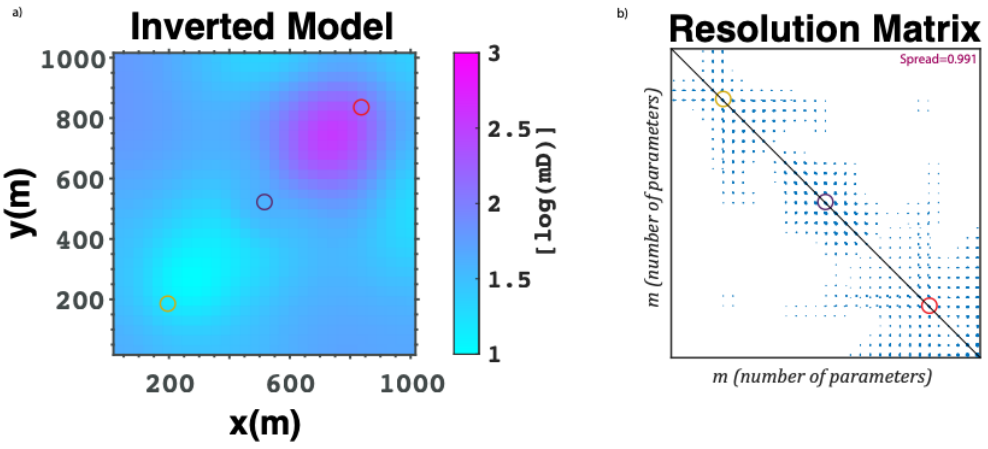

Figure 2-2: a) the inverted image with the location of the three wells highlighted with circles. b) the resolution matrix at the final iteration. Each circle on the R corresponds to a location of a well, colored accordingly.

The interpretation of the information matrix evaluates the quality of the modeled data compared to the observed data. Each row in the matrix is a weighted average of the neighboring observed data. The highest values are expected on the diagonal and around it. For example, in the first row with the highest value near the diagonal, the interpretation is that the quality of the predicted data to the observed data is high. Figure 2-3 displays the information matrix for the problem presented in Figure 2-3.

The N matrix displays three featured boxes. Each one of the boxes is related to data from one observation well. We plotted the pressure curves for reference on the side of the N matrix. Considering each box separately, we notice that the low values occur near the middle of the curves (center of the box), indicating that the prediction is poor near the intermediate values of the curves. An ideal information matrix that illustrates a perfect resolution to the predicted data is if the matrix N is the identity matrix. The spread calculation quantifies the quality of the information matrix. The lower the spread, the better predicted the data.

$$Spread(N) = \sum_{i=1}^{m}\sum_{j=1}^{m}[N_{ij} - I_{ij}]^2 \tag{2.4}$$

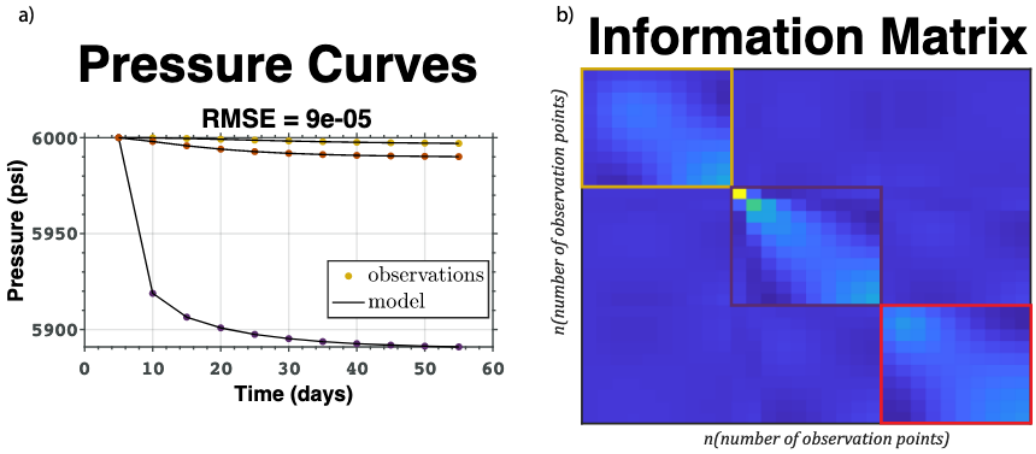

Figure 2-3: a) Pressure curves for three wells. b) the information matrix at the final iteration. Each box on the N corresponds to the pressure data of a well, colored accordingly.

After establishing the inversion and forward modeling techniques, we investigated the effect of limited data and model space on the inversion results. The inversion for reservoir parameters is found to be non-unique, leading to an examination of the non-uniqueness from the data and model space aspects.

2.3 Effects of data space

2.3.1 Case 1

The number of wells where pressure data is observed is critical to the history-matching process. Thus, we want to examine this closely and understand how this could affect reservoir modeling. We start with a 2-D base model with a constant porosity(ϕ) and a varying permeability (K in milliDarcy). The dimension of the reservoir is 1000 ft by 1000 ft, and it has a constant thickness of 50 ft. It is subjected to a constant production rate at the center of the reservoir. The initial condition of the reservoir is set at homogenous pressure. The boundaries are set at constant pressure, too.

Various vertical monitoring wells are randomly distributed across the reservoir. The synthetic pressure data from these wells are computed by discretizing the flow equation on 11 x 11 grids. Figure 2-4 displays the inversion result for the case where the number of observation wells matches the number of parameters. In the first case, where the number of wells equals the number of parameters, the inversion resulted in a model that matched the true model. The resolution matrix verifies that the inversion is almost perfect, shown by a zero spread. Figure 2-4 displays the inversion result for different cases where the number of wells varies. As the number of wells decreases, the resolution of the inverted image becomes less accurate. By examing the resolution matrix for the other cases in Figure 2-5, we notice that the diagonal shape of the matrix gets distorted as the number of wells decreases. When there is only data from a single well, neither the inverted image nor the resolution matrix has a coherent structure.

2.3.2 Case 2

In this case study, we want to examine the effect of the location of the monitoring wells on the inversion results. We use the same model presented in Case 1 to compute the synthetic data. The model has 11×11 grids with the same number of parameters.

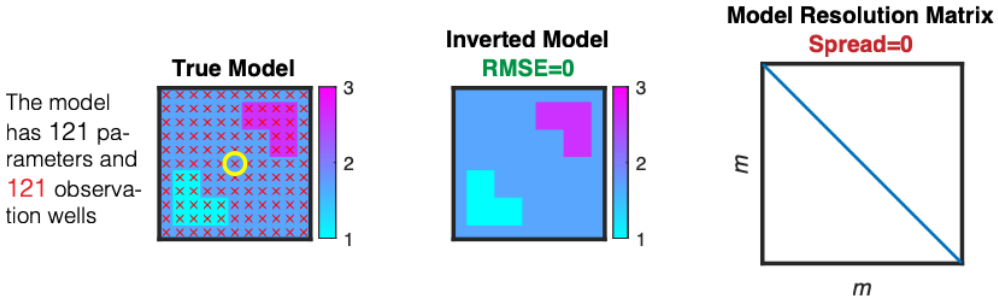

Figure 2-4: This figure illustrates inverting the pressure data from 121 wells for the permeability structure. The image on the left is the true permeability model. The yellow circle is the location where the production well is located. The red crosses are the locations where observation wells are placed.

One producing well is at the center, with the pressure change at this well being used as a data point. The presented examples in Figure 2-6 display different inversions for different random distributions of four observation wells in addition to the producing well. This case is very similar to Case 1, with the exception that the number of monitoring wells is fixed this time. For each example presented in this case, four wells are distrusted randomly across the reservoir. The results of the inversion are shown in Figure 2-6.

2.3.3 Case 3

Data can contain noise and could affect the inversion. The noise in the pressure data can be related to the measurement tool or fluctuation in the pressure over a long period. In this example, we add a random gaussian noise with a standard deviation, $\sigma = 1$. In Figure 2-7, we used the same true model with data from four wells in addition to the producing well to invert for the permeability. The inversion was stopped after 30 iterations to prevent it from diverging. The final inverted model does not resemble anything from the true model.

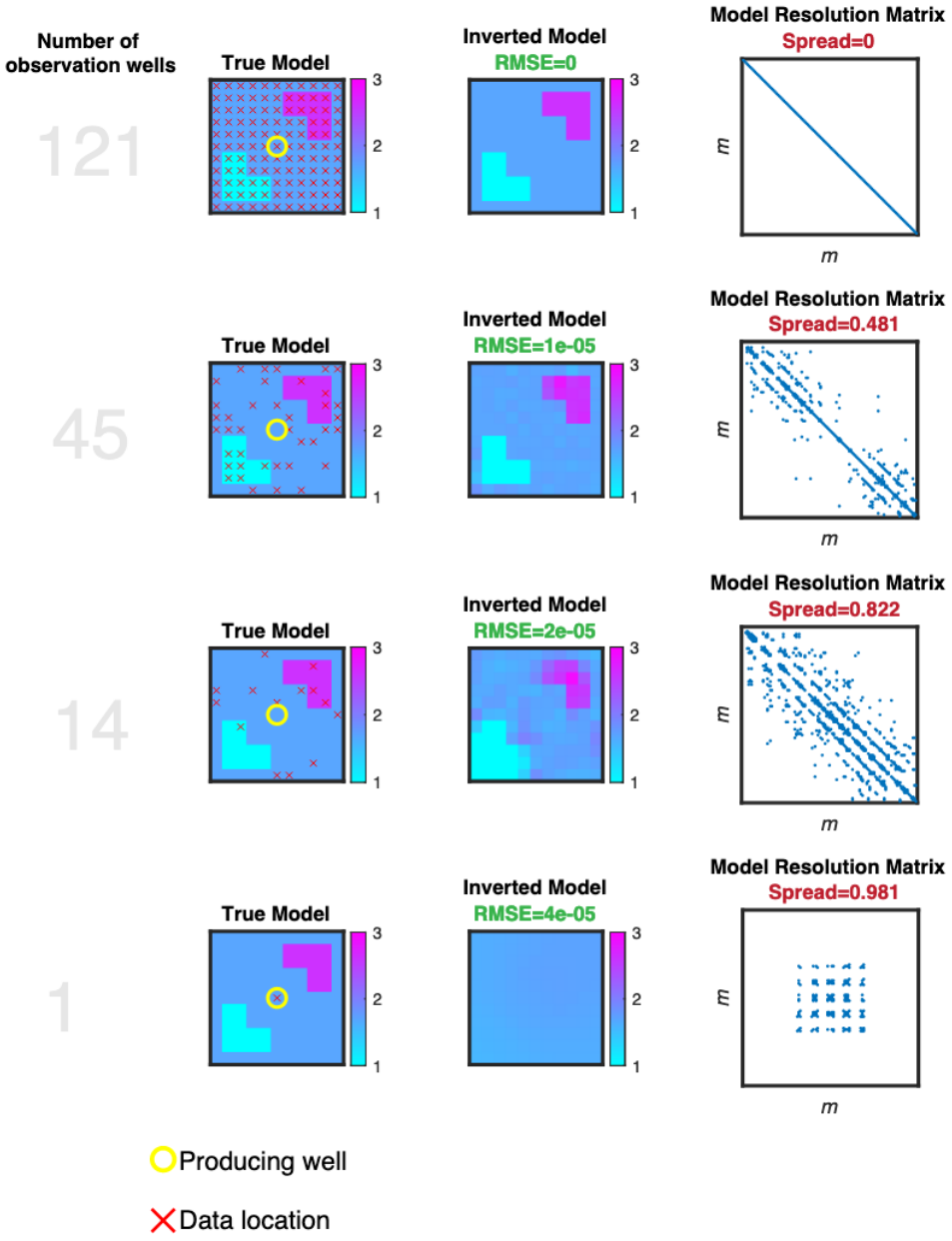

Figure 2-5: The first column of images displays the true model used to simulate the pressure data. Each row displays different red crosses where the pressure data is observed, ranging from 121 wells to only one well. The middle column represents the image from inverting the data observed at the red marker. On top of each inverted image is the final RMSE at the last iteration. The third column is the resolution matrix for each inverted image. The spread of the resolution matrix is displayed on top of each resolution matrix.

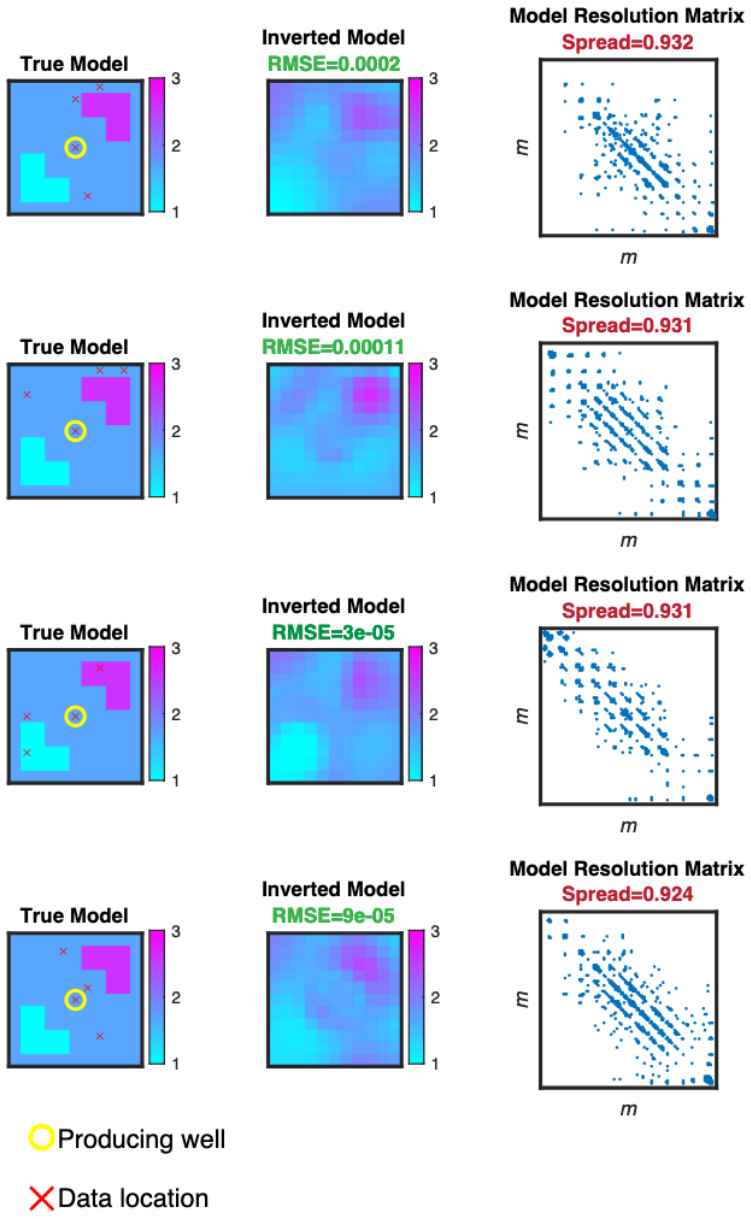

Figure 2-6: The figures show the inversion results from inverting different scenarios of four randomly distributed wells. Each inversion image resolves various structures in the model. The resolution matrices also have various structures depending on the location of the observation wells.

2.3.4 Case 4

The pressure data observed at a well location has multiple features, such as its gradient. In the previous cases, the transient pressure decreased with time, following

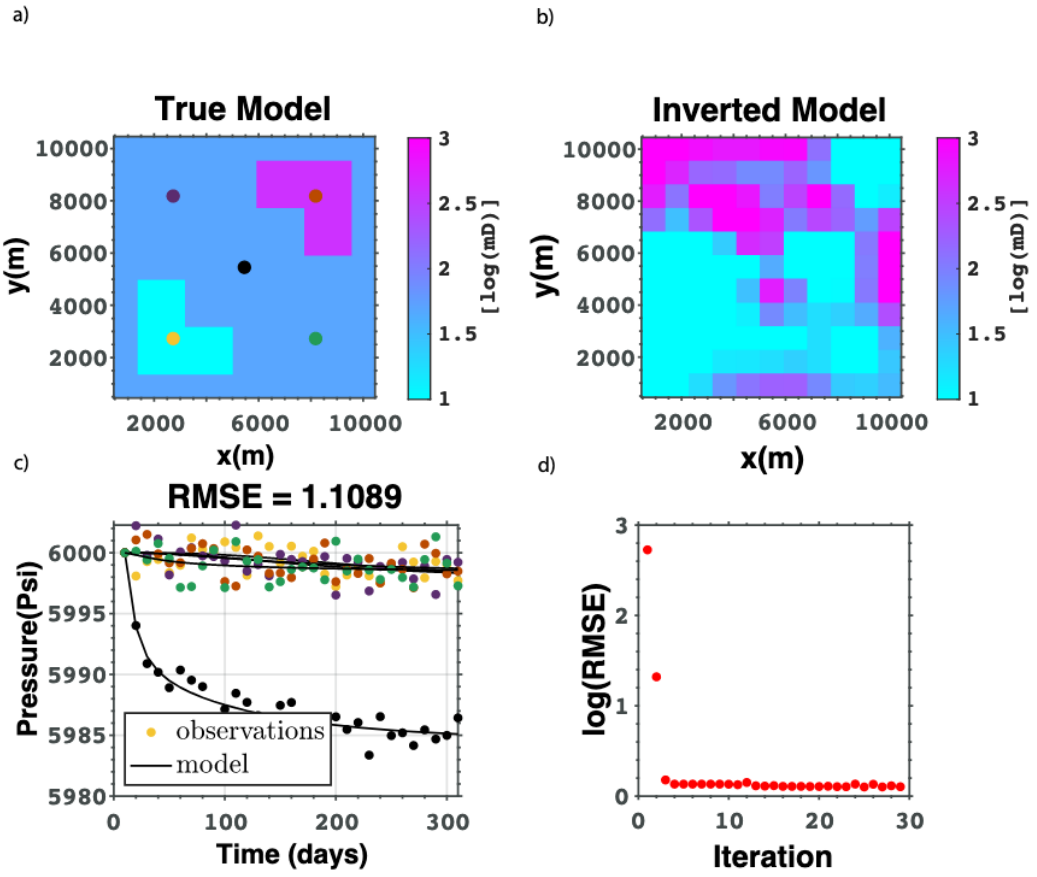

Figure 2-7: Inversion of noisy pressure data. The pressure data are measured at five wells colored differently, including the central producing wells. A random noise of one standard deviation is added to the pressure data. The different colors of the curves correspond to the location of the wells. The inverted in Figure (b) is the result after 30 iterations.

a similar trend in discharging a transistor in an electrical circuit. Identical to charging a transistor, where voltage builds up, the pressure in reservoirs can also be built up by injecting fluid into it. We sought to examine whether the different curves and response times for the pressure depletion and build-up would affect the inversion results differently. New synthetic pressure data are generated using the model presented in Case 1. The curves in Figure 2-8 are the pressure data at the central wellbore, which is partially used for production and later for injection. Inverting the new pressure

curves does not change the result or improve the resolution matrix. By comparing the resolution matrix in Figure 2-5 and the matrix in Figure 2-8, we do not see any difference in the result.

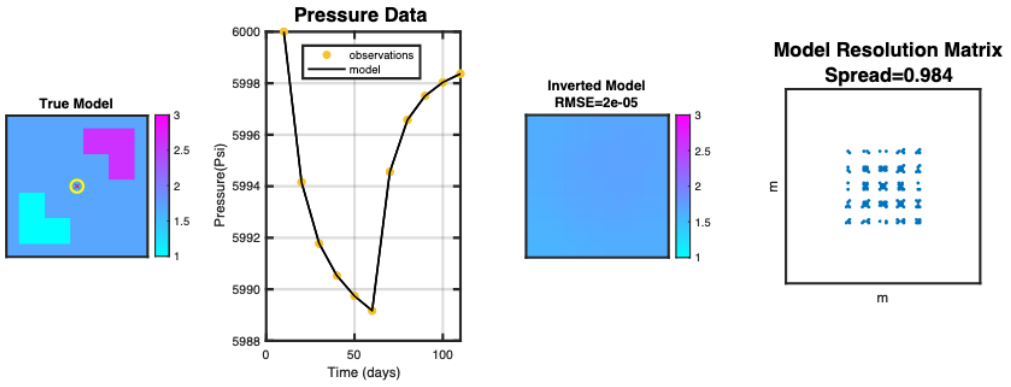

Figure 2-8: The figure displays the inversion of pressure data at the central well. The well was produced for the simulation's first half, then shut off, and fluid was injected for the second half. The far-right image is the resolution matrix.

2.4 Effects of model space

The model space is referred to as the parameter space under investigation. Therefore, the type and the number of parameters define the model space. In previous cases, the model space consisted of one parameter type, permeability, under the assumption that it was isotropic for 2-D cases($K_x = K_y$). We attempted to invert the data for a model with one parameter type.

However, in this section, we want to expand the model space and demonstrate how the effect of expanded model space can result in different inversion results. The model space was, thus, broadened to include more parameters. The model includes the permeability in the x and y dimensions and the porosity. The first attempt, in Figure 2-9, was to invert the three-parameter types (K_x, K_y, and ϕ). The result of inversion is not ideal, even with pressure data at as many wells as the number of grids. Figure

2-10 displays the inverted image for the different parameters in four scenarios, where the number of observation wells changes.

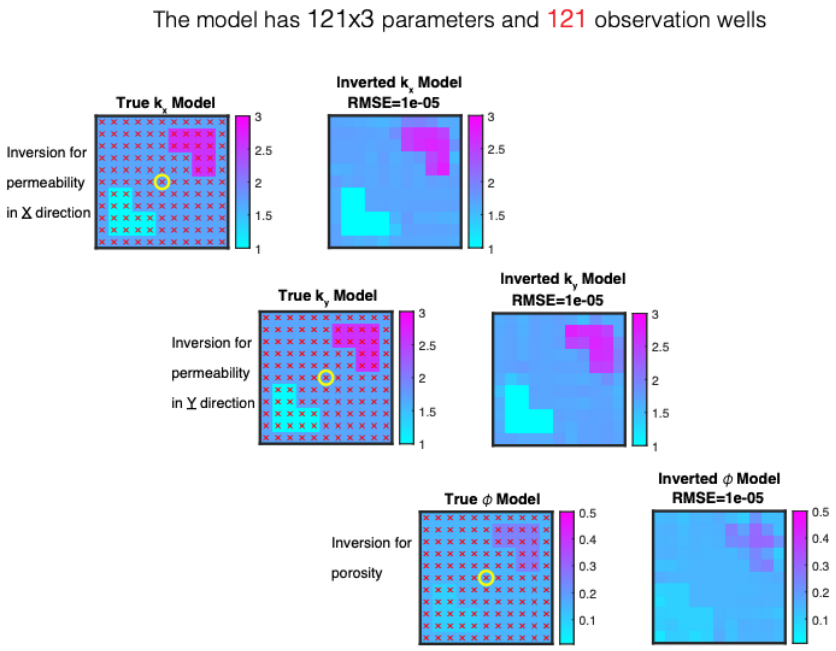

Figure 2-9: This figure illustrates inverting the pressure data from 121 wells for the porosity, in addition to the permeability structures in X and Y.

2.5 Discussion

The forward and inverse modeling methods for fluid flow in porous media were validated and proven effective. Although the forward problem always has at least one solution, we show that inverse modeling does not guarantee a unique answer. The final inverted image of the reservoir parameter depends on factors in the data space and the model space. The presented cases illustrate the difficulty in modeling reservoir parameters. It is easy to understand the effect of the amount of data on the inversion. With less data available, the problem becomes underdetermined. The regularization of the inverse algorithm has improved the inversion results and has made the problem

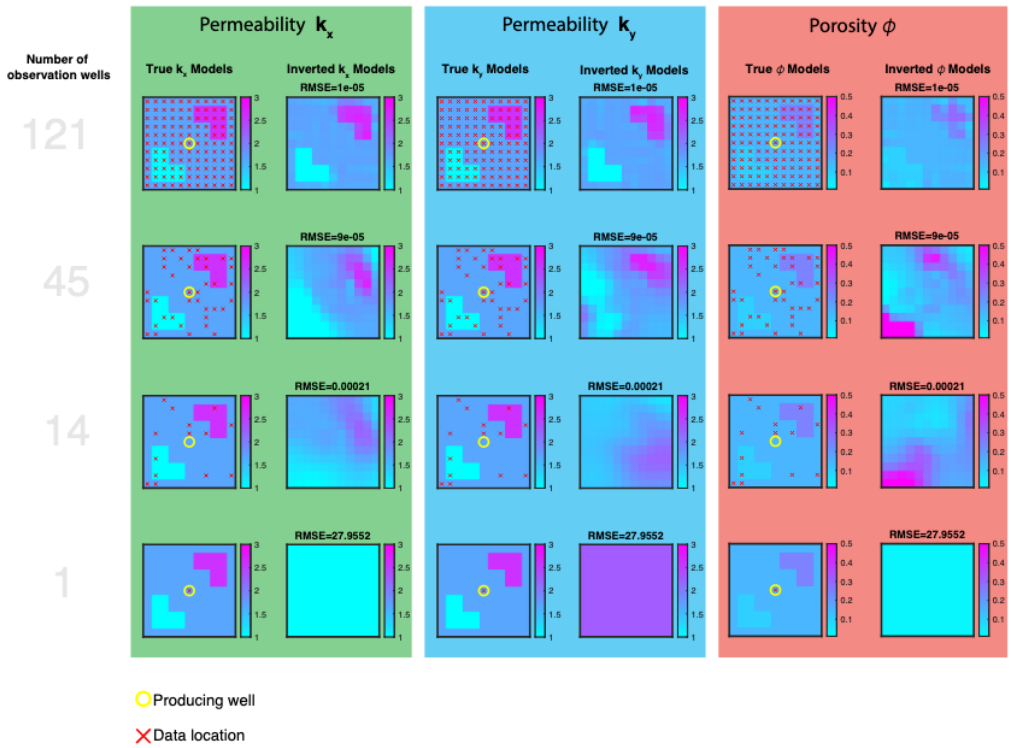

Figure 2-10: Each color box illustrates the inversion of one type of parameter. The green box displays the true permeability in the X direction and the inversion of the data obtained at the data locations. The blue box is the inversion of the permeability in the Y direction. Finally, the red box is the inversion of the porosity. Each row displays different red crosses where the pressure data is observed, ranging from 121 wells to only one well.

less ill-posed. The inversion scheme fits the data excellently, as indicated by the RMSE in figure 1-6. In the first example of case 1, where the number of observation wells matches the number of parameters, the inversion result perfectly matches the true model. Because the inversion utilizes the same forward modeling method used to create the synthetic data, we expected the result to be close to perfect. The resolution matrix confirms the accuracy of the inversion, which is an exact identity matrix. The number of observation wells does not affect the match between the observed and the modeled data (RMSE). However, the recovered reservoir models show a significant difference as the number of wells decreases. The spread of the resolution matrices emphasizes how

the recovered models are affected as the number of wells changes.

The examples in Case 2 indicate that the number and distribution of data are significant for resolving the parameters. Because fluid flows in spatial space, we could estimate parameters accurately depending on where we measured the pressure. Even though the resolution matrices in Figure 2-1 do not show a significant difference, they have slightly different spread values. In the fourth chapter of this thesis, we investigate the optimum distribution of wells to maximize the resolution matrix and the information content.

In addition to the effects of data quantity, the noise in the data can cause the inversion method to diverge from the solution. The inversion for the parameters with noise did not properly image the reservoir. As a result, the method got stuck in the local minimum or diverged from the solution. We, therefore, have changed the algorithm to stabilize the inversion and ensure convergence to the problem. The new algorithm uses singular value decomposition of the jacobian matrix J from the equation 1.11. Singular Value Decomposition (SVD), which is mathematically robust and numerically stable, can be used to solve eq1.11. The Jacobian "**J**" matrix is decomposed into the product of three matrices:

$$\mathbf{J} = \mathbf{USV}^T. \tag{2.5}$$

where $\mathbf{U}(n \times m)$ is the left orthognal vectors representing the data space, $\mathbf{V}$ is a matrix of size $(m \times m)$ containing right orthogonal vectors representing the parameter space, and $\mathbf{S}(m \times m)$ is a diagonal matrix containing mostly non-zero eigenvalues of J. The diagonal entries of $\mathbf{S}$ are $(\lambda_1, \lambda_2, ...\lambda_m)$, where λ_i are the eigenvalues, ordered from highest to lowest. By substituting eq2.5 in eq1.11,

$$\delta \boldsymbol{m} = (\mathbf{VS}^2\mathbf{V}^T + \alpha^2 I)^{-1}\mathbf{VSU}^T\delta d. \tag{2.6}$$

Adding the regularization factor to the diagonal matrix and taking the inverse will lead

to the following:

$$\delta \boldsymbol{m} = \mathbf{V} diag\left[\frac{1}{\lambda_i^2 + \alpha^2}\right] \mathbf{V}^T \mathbf{V} \mathbf{S} \mathbf{U}^T \delta d. \tag{2.7}$$

$$\delta \boldsymbol{m} = \mathbf{V} diag\left[\frac{\lambda_i}{\lambda_i^2 + \alpha^2}\right] \mathbf{U}^T \delta d. \tag{2.8}$$

eq2.8 provides a final solution of damped least-squares inversion with SVD. The solution has become more stable but could not resolve the exact parameters (Figure 2-11). We think that with proper initial guesses, the inversion of the data with noise can be more accurate.

We also illustrated the effects of the number of properties in interest on the inversion. Whenever the number of parameters became larger, the inversion proved unsuccessful in resolving the exact parameters. Even though observation data is at each grid, the inversion method could not obtain accurate estimates of the parameters (Figure 2-10). Furthermore, as we increase the number of model parameters, we tend to make the inverse problem underdetermined, leading to higher model uncertainty and deterioration of the final resolution. In chapter three, we will discuss the idea of reducing the model space and its complexity through inversion.

For uncertainty quantification, the model covariance matrix is computed. It gives us a quantitative measure of how uncertainty in the true model, the reference model, and the data propagate into model uncertainty. The model covariance depends on the covariance of the data and how the error is mapped from data to model parameters. A unit covariance matrix can characterize the degree of error amplification in the mapping. If the data are assumed to be uncorrelated and to have uniform variance, the unit covariance matrix is given by

$$Cov_m = J^{-g} J^{gT} \tag{2.9}$$

The matrix Cov_m has a dimension similar to the resolution matrix (m x m). We can examine the different covariance matrices for a different solution to evaluate the

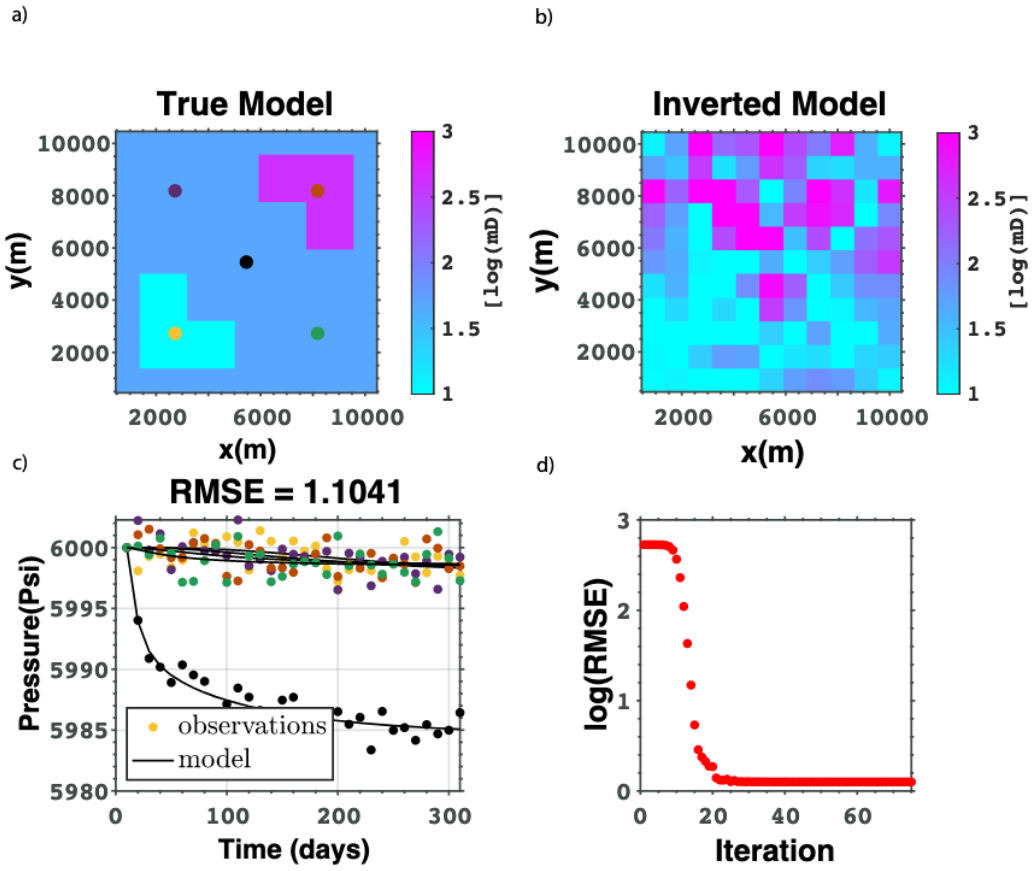

Figure 2-11: This figure shows the inversion of noisy pressure data using the SVD method. The pressure data are measured at five wells colored differently, including the central producing wells(black). A random noise of one standard deviation is added to the pressure data. The different colors of the curves correspond to the location of the wells. b)The inverted model after 78 iterations. c) pressure curves in psi at each well location. d) the log of the RMSE at each iteration.

uncertainty. The figure (Figure 2-12) displays the results from case one with the covariance matrices added. The diagonal of the covariance matrix represents the variance of the parameters. The size of the covariance matrix can be estimated by computing the trace of the matrix.

$$size(Cov_m) = \sum_{i=1}^{m} [cov_m]_{ii} \tag{2.10}$$

The large size of the covariance matrix indicates a low certainty (high variance) in the

parameter estimation.

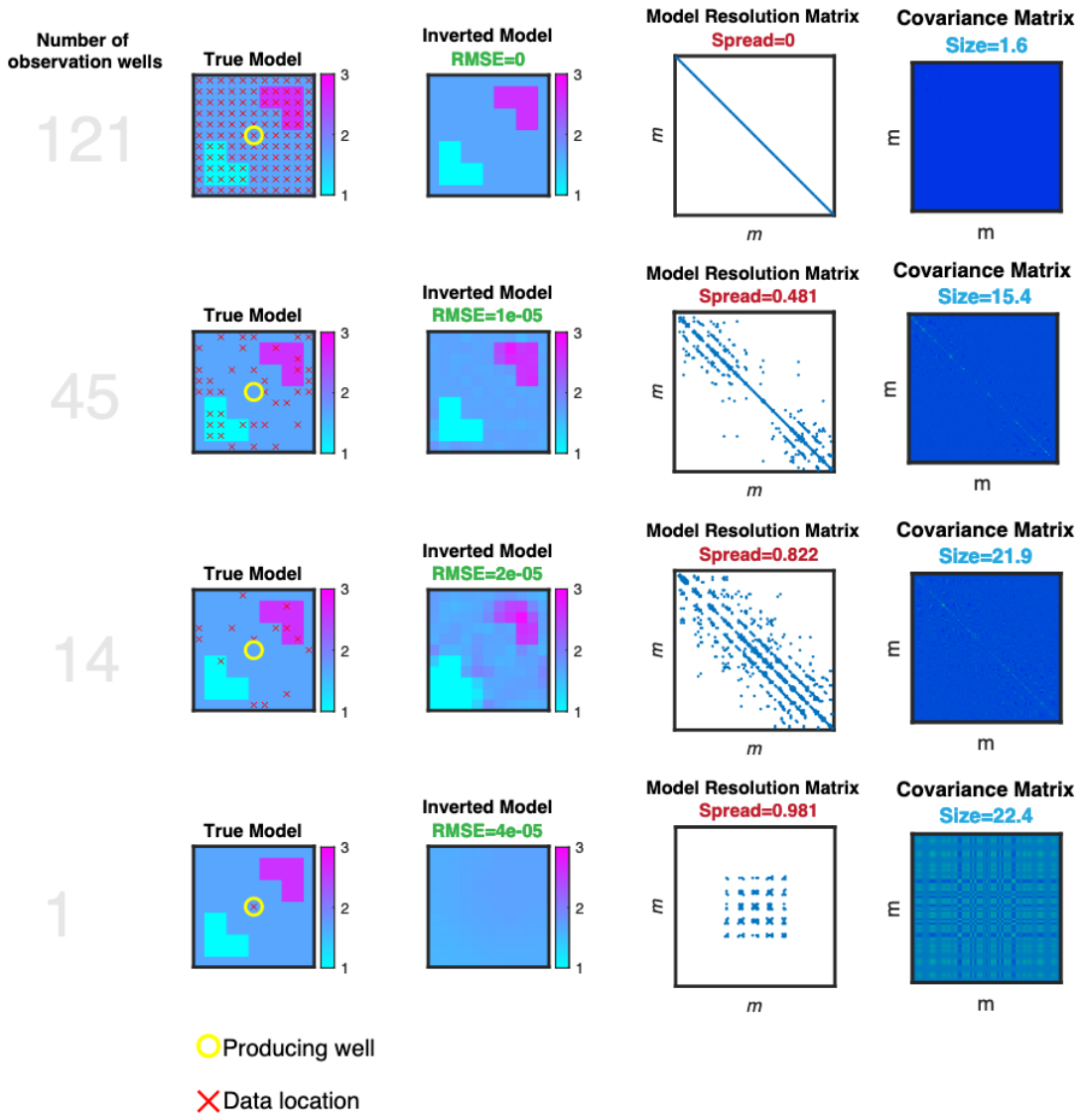

Figure 2-12: The covariance matrices highlight the uncertainty of using a different number of wells to invert the permeability structure.

In summary, the synthetic examples presented here indicate that the fluid flow inverse problem is highly non-unique in that multiple settings result in different realizations of the reservoir parameters. The objective of the geophysical data inversion process is to accurately approximate the geological structure in the subsurface. The

sources of non-uniqueness in fluid flow modeling may stem from any of the following issues: 1) The data quality and quantity will allow different ranges of models to fit the data. 2) Different combinations of parameters can result in the same identified misfit. 3) The inversion algorithm can lead to a local minimum, never reaching the global minimum solution.

Chapter 3

Reducing Complexity and Optimizing Drilling Locations to Better Represent Reservoirs

3.1 Abstract

History matching and predicting the future performance of hydrocarbon reservoirs are considered significant challenges that engineers face. Predicting the production of a hydrocarbon reservoir requires an accurate characterization of its fluid-flow properties, such as permeability and porosity, which can be used to simulate future production. The challenge in characterizing these properties is a function of the type and quality of the available data. Borehole data, for instance, is costly to acquire, which makes optimizing well placement a critical task for obtaining appropriate data for accurate reservoir modeling. Also, the simulation's complexity makes evaluating the reservoir performance expensive. The conventional reservoir history matching procedure usually requires a trial-and-error process of altering various reservoir parameters and simulating the pressure distribution and field production. This article proposes 'Inversion mod-

eling' to obtain a reduced-parameter structure and represent reservoirs with unique, simplified models. The simplified model is an effective medium with a few permeability values that are spatially distributed while maintaining the model's accuracy. In a later section, we perform a two-stage approach for adaptive wellbore placement with the end target of accurately resolving the reservoir parameters. In the first stage, we look for the distribution of wells that minimizes a measure of a model's non-uniqueness. The second stage uses the Shannon information entropy and the resolution matrix to find the best well locations in an adaptive manner. We test this approach on a synthetic simulation to show the effectiveness of the proposed method.

3.2 Introduction

Fluid in the subsurface flows in porous media due to physical parameters, fluid properties, and driving forces. Simulating the fluid flow in the porous media can help us analyze the subsurface and characterize its properties. Interpreting the reservoir characteristics, such as permeability and porosity, can enable us to perform history matching of the reservoir and ultimately predict its future performance. The simulation is often achieved by forward modeling, which refers to constructing a model whose behavior matches an actual reservoir. Forward modeling aims to estimate and optimize field performance under several production conditions. This challenging task requires a good understanding of a given reservoir's geology and fluid flow physics. The required time and cost for simulation are proportional to the reservoir's structure, size, and complexity. The current approach to determining the reservoir properties approximates the spatial static parameters of reservoirs by a finite discretization method by subdividing the model into cells, each cell characterized by constant parameters. Often, a large (up to 10^9) number of cells is considered to reproduce the reservoir's performance with high resolution. The parameter of each cell is altered until the output of the forward modeling matches the production data; "History Matching." In some

cases, the forward modeling is done on an upscaled, coarse-scaled model of a finer one to reduce computation cost.

An alternative method to manual history matching is Inverse Modeling, which refers to estimating the parameters that fit the measured data by an automatic algorithm. In groundwater hydrology or the petroleum industry, inversion is related to estimating the permeability and porosity from the well production data. Because the data is much less than the number of parameters, mathematicians may refer to the problem as an under-determined problem (Carrera, Alcolea, Hidalgo & Slooten, 2005). Regularization is one method to address this problem, transforming the inversion into an approximate unique problem. The inverse problem can be formulated as a solution to an operator equation:

$$d = F(m) \tag{3.1}$$

where m is a vector of the model parameters from a model space, d is a data vector from the data space, and F is the forward operator.

Regardless of the approach taken toward history matching, the problem is ill-posed due to the failure to meet one or more of the following conditions: 1) solution m exists, 2) the solution m is unique, and 3) the solution m depends continuously on the data d. The data we obtain from the field is spatially limited to drilled wellbores. Thus, many models can fit the data. The resulting model is a non-unique representation of the reservoir. The problem can, though, become well-posed by regularization. A classical principle of regularization theory is discussed in (Tikhonov and Arsenin, 1977). Instead of considering an infinite set of solutions with regularization, we narrow the class of models that solve the inverse problem. Figure 3-1 illustrates a definition of an approximate solution $\hat{m}$ to the inverse problem eq3.1. The modeled data $\hat{d}$ resulting from the solution $\hat{m}$ is the closest to the observed data d. For an approximate solution, which can be achieved by Tikhonov regularization, both the solution and the modeled

data are part of a larger subset in the model and data space.

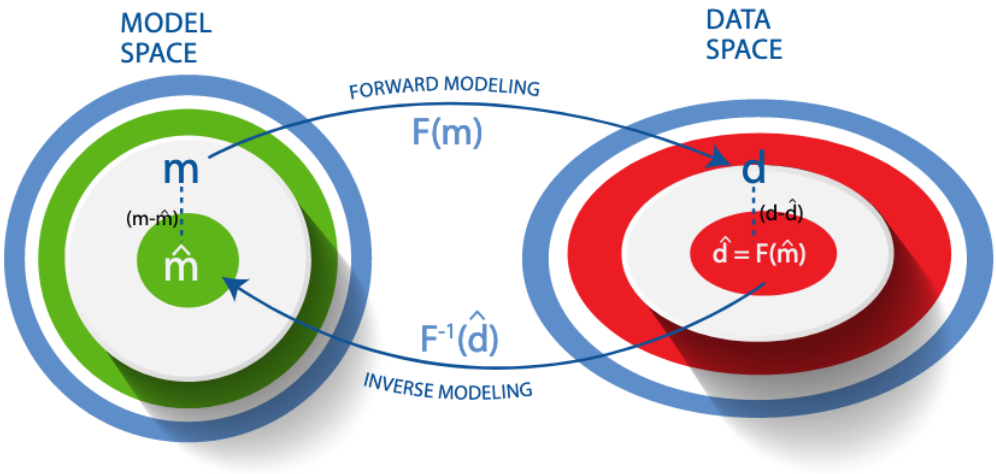

Figure 3-1: Schematic representation of the model and data spaces.

The solution to the inverse problem should find the best $\hat{m}$, in which its modeled data minimizes the distance between the d_{obs} and the $\hat{d}$. This chapter studies the concept of narrowing the model subspace and carefully choosing a data subspace to reduce the non-uniqueness that arises by having many possible solutions. The first section of this chapter (Reducing Reservoir Complexity) studies the reduction of the model space. In the second section (Optimization of Drilling Location for an Improved Reservoir Representation), we are developing a method to choose data locations that reduce non-uniqueness and increase information gain.

3.3 Reducing Reservoir Complexity

In this section, an inversion is used to estimate a reduced-parameter structure of reservoirs. The flow structure is described by two hydraulic parameters: permeability and porosity. Permeability is defined as the ability of the medium to transport the

fluid through it under pressure gradients, and porosity describes the storing properties of the porous medium. The parameters are identified from piezometric pressure data in a transient flow condition. We have gathered synthetic well data from modeling a complex reservoir with a structure as large as ten thousand cells. Then, we performed the inversion on a smaller grid to reach a structured reservoir as small as one hundred cells. The significance of this study is that doing reservoir modeling with a very large number of cells is expensive, while a much simpler model is just as accurate.

3.3.1 Methodology

The reduction of the model complexity of a reservoir depends on the existing forward and inverse modeling techniques discussed in Chapter 1. In the forward modeling, we simulated the pore pressure resulting from fluid flowing in porous media under different initial and boundary conditions. The forward modeling method solves the pressure diffusion equations numerically. For a simple case, we are modeling the pressure changes in time as we are producing/injecting fluid into a reservoir. The reservoir is discretized into grids where each grid has a specific parameter value.

The inversion algorithm attempts to minimize the L_2 norm between the observed data d and the modeled data resulting from a reduced model $m_{reduced}$ (eq3.2). The last term in the equation is the regularization term. The α controls the rate of convergence to a solution. L is a regularization matrix.

$$O(m) = \| \left(F(m_{reduced}) - d\right) \|^2 + \alpha^2 \| L\delta m_{reduced} \|^2 \tag{3.2}$$

3.3.2 Numerical Experiments

Two different numerical experiments with synthetic data are presented to demonstrate the reduction in the modeling of reservoirs. Each experiment has the same objective: to reduce the number of parameters in the model while maintaining reason-

able accuracy. The first experiment examines complexity reduction by reducing the grid into a coarse, regular grid. The second experiment attempts to find a further reduced irregular grid.

Experiment 1

In this experiment, we reduce the complexity of a reservoir on a regular grid. Consider the two-dimensional reservoir shown in Figure 3-2; the reservoir has a random permeability structure constructed using a spectral density function of an exponential autocovariance function (Gelhar & Axness, 1983). We obtain synthetic pressure data from the well locations by simulating a single-phase fluid flow on a large grid (130x130). The boundaries of the model are constant pressure boundary conditions, and the model has an initial constant pressure everywhere. One production well is placed at the center of the model, which pumps fluid out of the reservoir at a constant rate of 1000m^3/day. Also, there are four wells used as monitoring wells. Details of the simulation parameters are presented in table 3.1. The simulation runs for 40 days, and the pressure data are collected at each monitoring well with an increment of one data point per day (Figure 3-2).

We used the pressure data to invert for the permeability structure on a reduced grid. The reduced grid, which has a size of (50x50), is almost 15% of the original grid (number of parameters) used to generate the data. Inverting the pressure data on the reduced grid appears comparable to the true model but with a slightly lower resolution (Figure 3-3). If we compare the actual data obtained for the original model to the data from the reduced model, the RMSE (root mean square error) is insignificantly low.

The attempt to further reduce the number of grids has also resulted in low RMSE values. Figure 3-4 demonstrates the inversion results on various grid sizes. No matter how small the number of grids, the inversion result produces very low data RMSE. The overall structure of the inverted images is slightly similar to the actual structure. In the case where the number of grids is (10x10), the inverted structure appears to

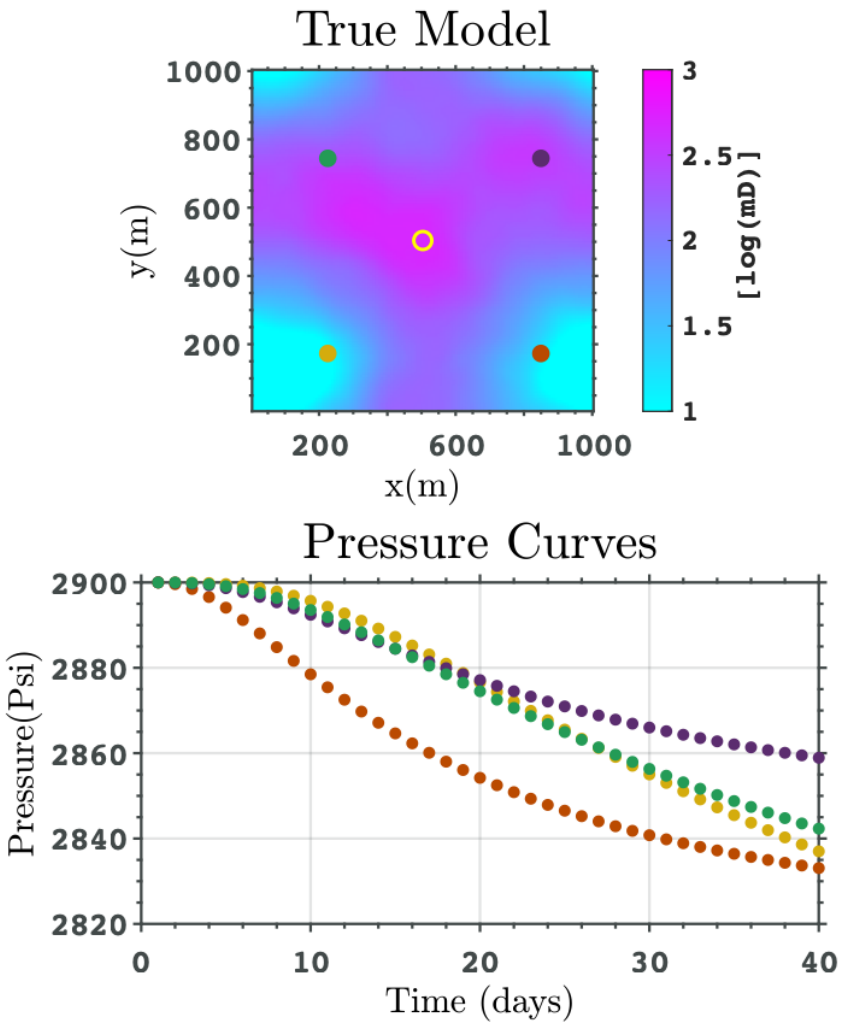

Figure 3-2: The true permeability model with the location of the monitoring wells colored differently. The bottom figure is the pressure curves with time at each monitoring well (colored accordingly).

be different from the true model, but it still highlights the high and low permeability regions.

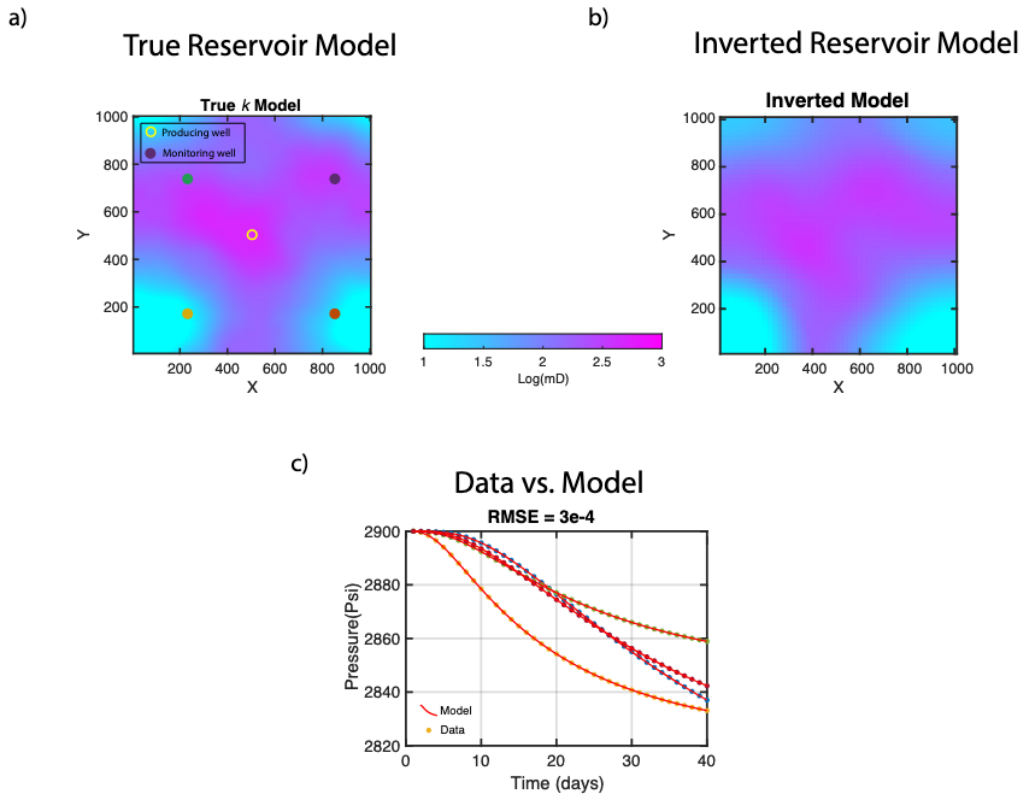

Figure 3-3: a) The true permeability model. b) the inverted model on a grid 50x50. c) Observation pressure data with the best fit.

The least complex model we have achieved is the model with grids(10x10). We attempted to reduce that grid to an even less complex one that has (5x5) grids. However, such a small number of grids has not enabled us to devise a plausible structure that fits the data. In the next experiment, we developed a scheme to reduce the complexity with a combination of variable grid sizes.

Experiment 2

The result of the first experiment has demonstrated a way to reduce the modeling of the permeability structure of a reservoir with less complexity using a regular fewer

Parameters	value
K(mD)	(1-1000)
phi	0.15
μ(cp)	1.5
C_t(1/psi)	12×10^{-6}
q(m^3/day)	1000
Initial P	2900 psi
Boundary P	2900 psi
Time (days)	40
Dimensions[x,y,z] (m)	[1000,1000,30]

Table 3.1: The values of the parameters used for simulating the fluid flow for experiment one.

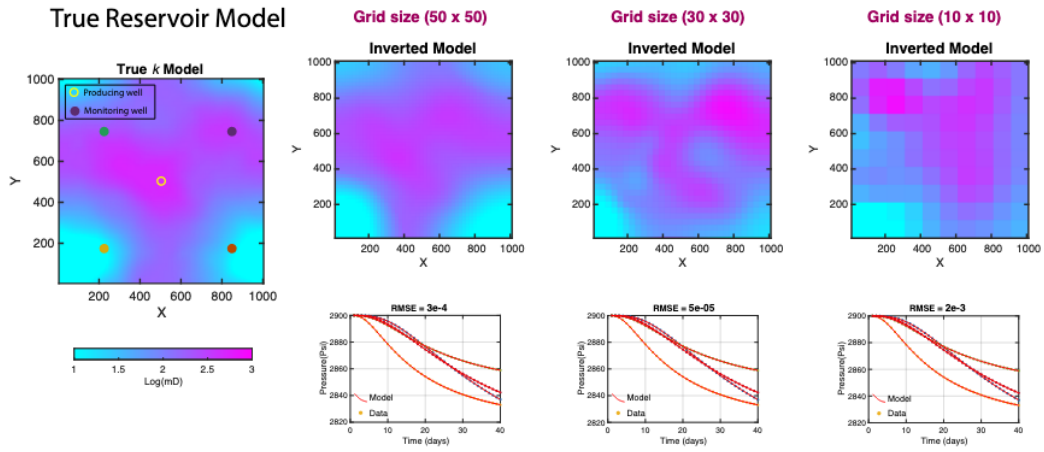

Figure 3-4: The result of inverting the observation data at a different number of grids. The displayed RMSE is the final error at the last iteration.

number of grids. It has also shown that the method can produce an upscaled model-an averaged permeability model of the complex on-with an inversion method. However, we feel the complexity can be reduced further with a different approach. When we attempted to reduce the grids to a smaller grid, there was no simple structure that would fit the data.

Our alternative way to reduce the structure is to refine the grid irregularly. Hence, we will construct a grid with variable block sizes. The designs of the various blocks follow multiple steps. A summary of the steps is illustrated in Figure 3-5: Step 1:

Start with partitioning a fine grid into a coarse grid with large blocks. For example, the (30x30) model is overlayed by a (5x5) grid. Then, all values in the fine grid (30x30) that fall in the course grid (5x5) are averaged. The result should be an averaged model with the size of (5x5) grids. Step 2: Find a coarse grid that has a high permeability compared to the pre-determined cutoff permeability value and refine the course grid into a fine grid of a specified size. Step 3: Invert the pressure data on the new irregular grid.

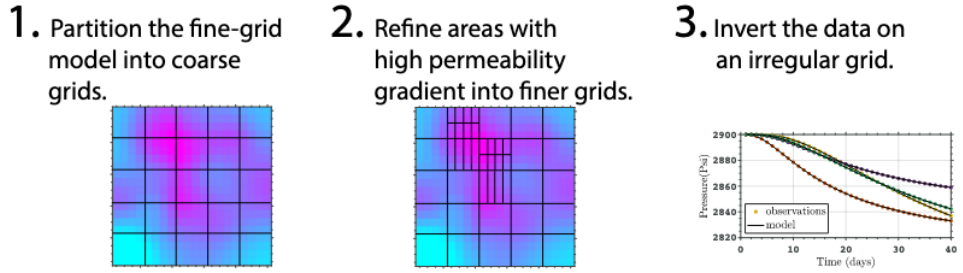

Figure 3-5: a) The true permeability model. b) the inverted model on a refined grid. c) Observation pressure data with the best fit.

The described method allows us to find a unique grid that we can use for the inversion process. However, this method still relies on having a regular grid obtained by the method in the first experiment. Once the reduced grid in experiment one is created, we can experiment with different partitioning and refining until we find the simplest irregular model grid that fits the data.

We now revisit the first experiment's results and reduce the complexity to demonstrate the described method. We followed the above steps to find a grid between the 10x10 and 5x5 that will allow us to resolve a structure with a low RMSE. We started by partitioning the 30x30 inverted model into 24 large blocks, and then we refined the large blocks into smaller grids based on the permeability value. The inversion algorithm finds the best permeability values for each grid that result in a low data RMSE. Figure

3-6 shows the result of inverting the pressure data on a refined grid. The inverted model shows two distinct regions.

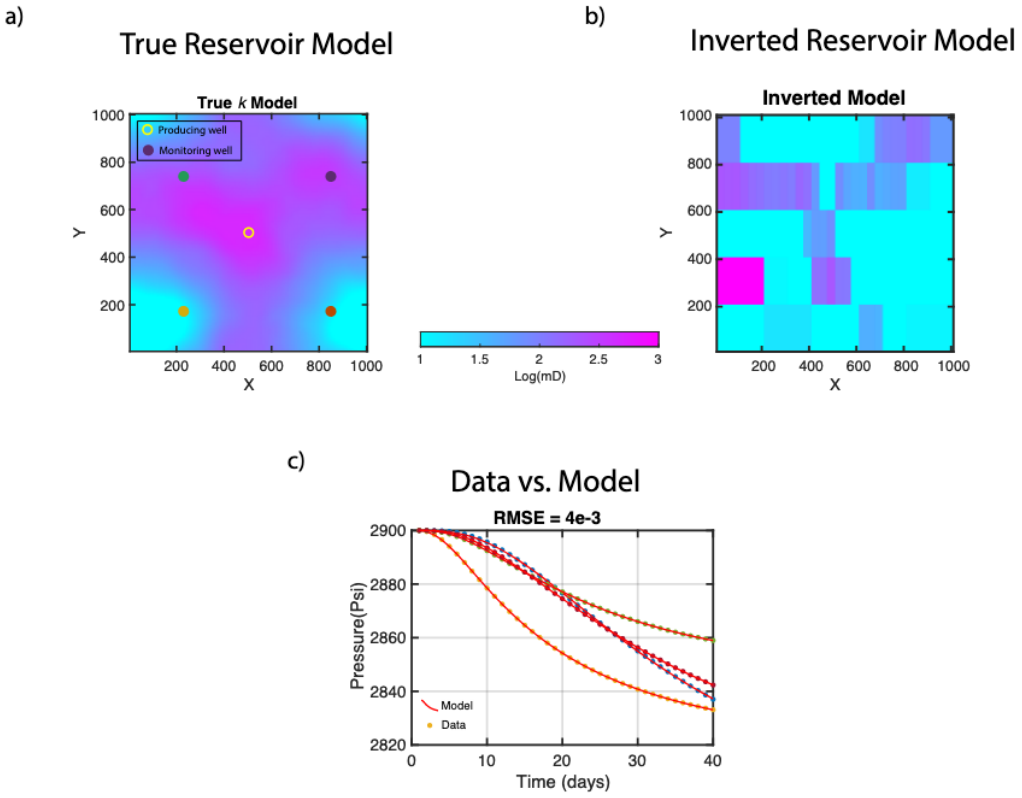

Figure 3-6: An overview of the steps adopted to refine the model into irregular grids.

3.4 Optimization of Drilling Location for an Improved Reservoir Representation

The spatial distribution of wells can be critical for optimizing the production of a hydrocarbon reservoir. During the development stage of a hydrocarbon field for an optimum production rate, wells are ideally located in areas of high reservoir porosity and permeability, assuming an accurate model of these properties is available. An alternative, and more practical optimization problem, is maximizing parameter resolution, which can aid in optimizing the production rate.

The optimum design for a geophysical survey focuses on quantitatively obtaining a more accurate subsurface model. Kijko (1977) was the first to attempt to optimize a geophysical experimental design, with the type of survey being global seismic monitoring for earthquake-location problems. The problem of optimal geophysical designs has been pursued in many other applications, from mapping the ocean bottom (Barth and Wunsch (1990)) and soil testing (Crisp et al., (2020)), to magnetotelluric investigations (Jones and Foster (1986)) and electrical resistivity tomography (Uhlemann et al., (2018)). Maurer et al., (2010) provide a solid background and development for survey design in geophysical methods.

This study focuses on finding the spatial distribution of wells that can help resolve the permeability structure of a hydrocarbon reservoir. The optimization objective in this study is different from the previously published work. Wang et al., (2007) and Zandvliet et al., (2008), for instance, implemented a gradient optimization algorithm to optimize the wellbore locations that optimize the net present value (NPV). Their method depends on the pre-existing knowledge of the reservoir properties (permeability and porosity), assuming that it is accurate. Others, such as Montes et al., (2001), optimized the NPV problem using a genetic algorithm. Nogueira and Schiozer (2009) proposed a methodology to optimize the number and placement of horizontal and

vertical wells, which optimizes the NPV in a field with the lowest number of flow simulations. In contrast, we optimize the locations of wellbores to maximize our information and reduce inaccuracy in resolving the permeability structure of a hydrocarbon reservoir. The pressure data from wells are used in an inversion scheme to characterize the reservoir permeability, which can then be ultimately used for optimizing production or NPV.

We adopt a two-stage adaptive method by utilizing the forward and inverse modeling of fluid flow in porous media. In each iteration, we look for the next optimum wellbore location to maximize the information gain and reduce model uncertainty. Various publications discuss the use of an adaptive method for optimizing survey designs. For instance, Stummer et al., (2004) implemented an adaptive approach to find electrode configurations that provide the most new information. Cole and Morgan (2009) also looked at an adaptive experimental setup using a determinant-based objective function and showed an improvement over the randomly generated surveys in a Monte Carlo simulation.

This study adopts a technique previously developed by Winterfors and Curtis (2008) to quantify and detect system non-uniqueness. This study also applies a uniquely developed measure of data information to find an optimal well placement.

3.4.1 Method

Our optimization scheme aims to find the optimum survey design that would yield the most crucial data for accurate reservoir modeling. In the first stage, we will estimate the best distribution of wells that minimizes a measure of non-uniqueness. This stage does not depend on available data and can be performed before the start of the survey.

In the second stage, we utilize the resolution matrix (Menke, 2012) and the Shannon information (Shannon, 1948) to choose the optimum location for boreholes in an adaptive fashion. After every newly drilled borehole, we search within the survey do-

main for the next optimum borehole site, the one whose predicted data best reduces the spread of the resolution matrix and increases the Shannon information entropy.

Stage 1

The main challenge with the reservoir characterization inverse problem, as is common with many geophysical problems, is that it is ill-posed, making the data explainable by many different models. However, some datasets can provide better constraints on the model, making the problem appear less non-unique than other datasets. We, therefore, attempt to find a dataset that will reduce the non-uniqueness of the system under study, which means we first need a way to quantify non-uniqueness for different possible well-placement designs. Winterfors and Curtis (2008) discuss detecting non-uniqueness for a nonlinear system if the forward model is continuous and differentiable. In their approach, they would start with initial guesses for two models, m_{1i} and m_{2i}, and iteratively update the models until converging towards the final m_{1f} and m_{2f}, that yield the most similar predicted data. We formulate the problem below to find the two models for the reservoir characterization problem.

Consider the two different initial models, m_{1i} and m_{2i}. The fluid flow modeling would result in two different data sets:

$$F(m_{1i}) = d_1, F(m_{2i}) = d_2 \tag{3.3}$$

where F is the forward operator for modeling fluid flow in reservoirs. d_1 and d_2 are the pressure data resulting from modeling the fluid flow for the two models. Using a finite volume scheme, the forward operator solved the pressure diffusion equation to simulate fluid flow in porous media.

$$\frac{1}{\mu}\nabla \cdot ([K]\nabla P(x,y,t)) = \phi C_t \frac{\partial P}{\partial t} + q(x,y,t) \tag{3.4}$$

where μ is the fluid viscosity in (cp), $[K]$ is the permeability tensor in millidarcy(mD), P is pressure in (psi), ϕ is porosity, and C_t is the total compressibility in (psi^{-1}). The source/sink term q has a unit of (volume/time).

The forward operator (F), in eq3.3, has information about the boundary and initial conditions. It also has details about the source/sink term and the survey designs, such as the locations of wellbores. m_{1i} and m_{2i} are parameters vectors containing reservoir permeability. Because the property of interest in this study is permeability, other properties, such as porosity and compressibility, can be embedded in the forward operator F.

We want to find two models m_{1f} and m_{2f} that are unequal and have an insignificant difference between their modeled data. Hence, we will perturb the initial models until we match the data. We can expand eq3.3 using Tayler expansion to determine the model's perturbation for comparable data.

$$F(m_{1i}) + J_1 \delta m_1 = F(m_{2i}) + J_2 \delta m_2 \tag{3.5}$$

where J_1 and J_2 are the sensitivity matrices for two different linear systems with different initial models. δm_1 and δm_2 are the perturbations of models one and two, respectively. We can reformulate eq3.5 in a matrix form as follows:

$$\begin{bmatrix} J_1 & -J_2 \end{bmatrix} \begin{bmatrix} \delta m_1 \\ \delta m_2 \end{bmatrix} = F(m_{2i}) - F(m_{1i}) \tag{3.6}$$

Using the SVD factorization of the $[J_1 - J_2]$ matrix (Hansen, 1999), we will solve for perturbed models in the least-squares solution. When the solution converges, the difference between the two models, m_{1f} and m_{2f}, is the measure of the non-uniqueness. Figure 3-7 illustrates the basic steps for determining the measurement of non-uniqueness.

If we start the problem with one wellbore placed anywhere, we aim to find the next

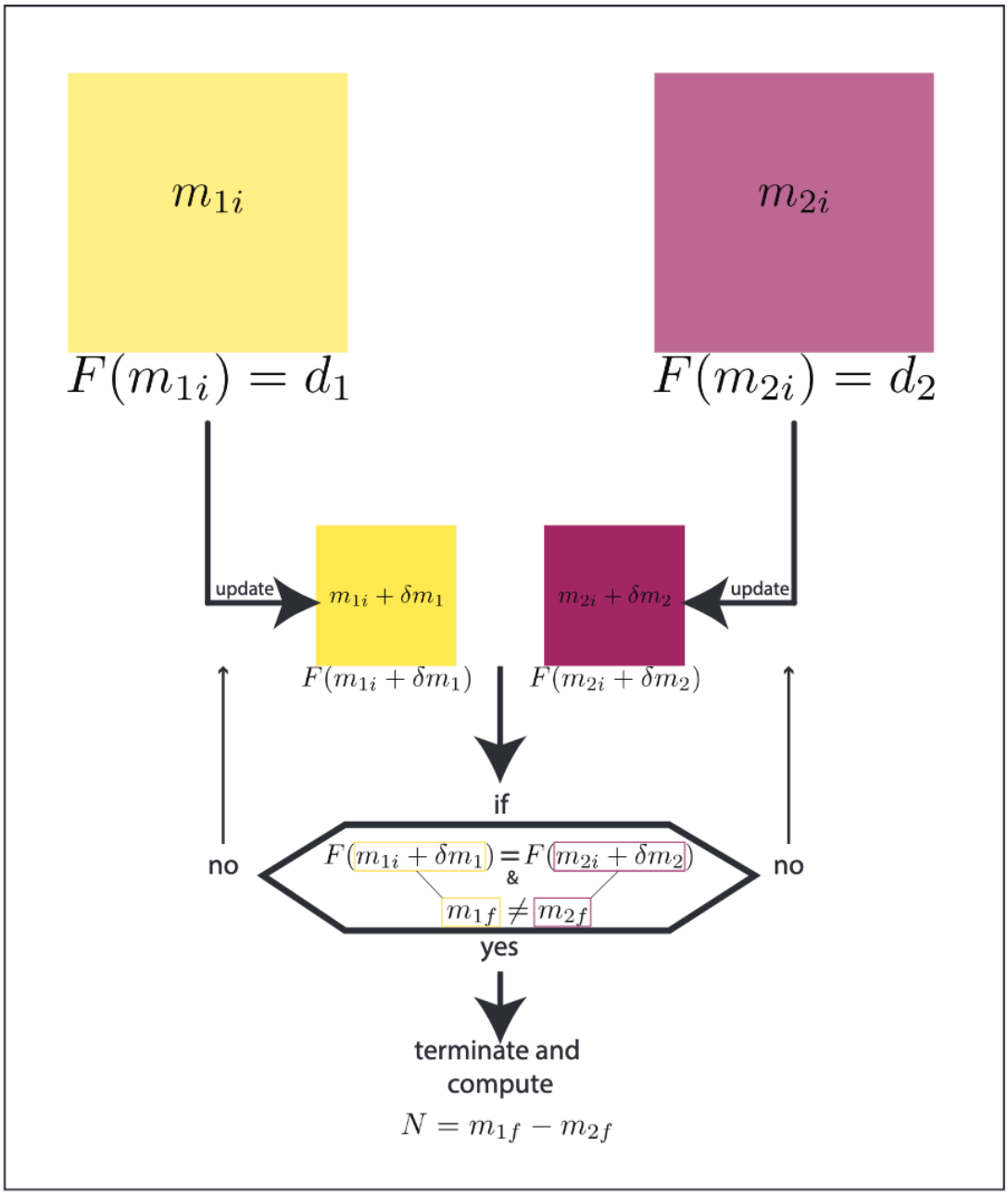

Figure 3-7: a general overview of the non-uniqueness detection scheme.

optimum well location by exploring the reduction in the problem's non-uniqueness. We formulate an optimization problem as follows:

$$\theta(\alpha) = ||G(\alpha) - N|| \tag{3.7}$$

where θ is an objective function, and G is a forward operator developed to measure the non-uniqueness between two models for well placement design α. N is the non-uniqueness measure. We use a grid search method for this minimization problem since we are searching for only one additional well placement, which would not require extensive computation. This allows us to find designs that give a minor difference between the two estimated models: m_{1f} and m_{2f}. The data from the optimal well location is then added to the available dataset, the starting point for stage 2 of the proposed method for optimal well placement. The process in this stage can be applied to construct a list of wellbores' locations; however, in this study, we kept it to one for simplicity.

Stage 2

The second stage in the survey design method is an adaptive scheme that depends on previously acquired data to choose a new well location. At first, the available data from previously obtained surveys is inverted for the reservoir parameters, and the data covariance matrix is computed. We, therefore, can estimate the information gained for possible wellbore locations using Shannon information entropy.

Shannon information entropy was developed to characterize the information content transmitted between a source and a receiver (Shannon, 1948 ; Cover & Thomas, 1991). Shannon, in his paper, described a metric for the information delivered from a source to a receiver using only binary decisions (yes or no). We can explain the Shannon information with a typical example regarding sending a word through binary digits and estimating the required number of bits. If each letter in the English language has the same probability (1/26), then the Shannon information I associated with one letter is $I = -log(\frac{1}{26}) = 4.7$. To understand this number, if someone can ask yes/no questions to guess any letter in the English language, they would require an average of 4.7 questions to guess the letter. Shannon information entropy estimates the level of information associated with the possible outcomes of a variable. Consider a random variable x

that has N number of possible values $x \in (x_1, x_2, \ldots, x_N)$ with the corresponding probabilities $p \in (p_1, p_2, \ldots, p_N)$, the Shannon information entropy is written as follows:

$$E = -\sum p \; ln(p) \tag{3.8}$$

where $p(x)$ is the probability of an event to occur and has to be a value $0 < p < 1$. The above equation is expressed in natural units. It can be expressed in binary units (bits) by substituting the ln to log_2. In our calculation, we decided to use the natural units as they seemed appropriate for our problem. The higher the entropy is, the higher the information gain and the less the uncertainty.

To put this in context for our problem, we want to estimate the entropy for each possible Wellbore location. High entropy indicates high information content. The p_d is the probability of the data center on the zero data-RMSE and has a standard deviation estimated from the model covariance matrix. The assumption here is that the probability of the data has a Gaussian distribution. Finally, we compute the Shannon information entropy for each possible location for the new well.

The locations with maximum entropy are chosen as potential well locations to perform the following model resolution matrix analysis. Computing the resolution matrix is an expensive process that requires running the inversion algorithm, in which we invert the existing real data and the predicted data at each potential location. The predicted data comes from modeling fluid flow on the estimated model parameters from the inversion at the beginning of this stage. We then compute the resolution R_m matrix, which measures how close the inverted model is relative to the accurate model, using the equation below:

$$R_m = J^{-g} J \tag{3.9}$$

where J^{-g} is the generalized inversed matrix, and J is the jacobian matrix. R_m will equal the identity matrix if the inverted model is very close to the true model. The

spread (S) of the resolution matrix is then computed by the L_2 norm of the difference between the resolution matrix and the identity matrix:

$$S = -\|R_m - I\| \tag{3.10}$$

where I is the identity matrix. The minus sign is there because we are interested in the lowest spread value. The optimum location for the new well is the one that maximizes the entropy and minimizes the spread resolution matrix, which is done by maximizing the quantity:

$$[E(\alpha) + S(\alpha)] \tag{3.11}$$

The data from the optimal well location is added to the current dataset, and the process is repeated to find the next optimum well location. We summarize the adaptive data acquisition in three steps, as illustrated in Figure 3-8:
1) Acquire new data. 2) Estimate a permeability model using the pressure data from the existing wells. 3) Estimate the current model & data uncertainty 4) Explore different well placement designs that solve the following optimization equation and repeat from step one:

$$maximize[E(\alpha) + S(\alpha)] \tag{3.12}$$

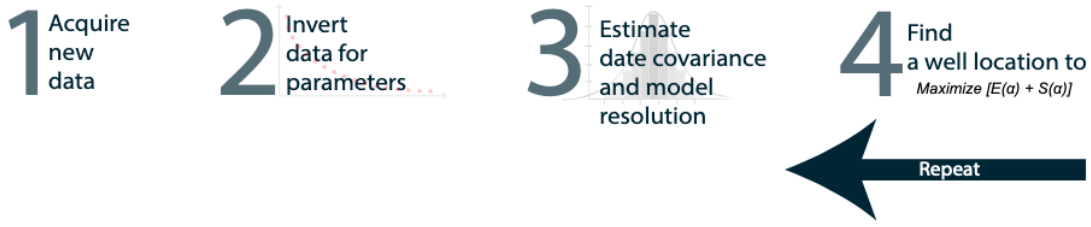

Figure 3-8: Flowchart illustrating the main steps for the adaptive placement of wells in Stage 2. 1) pressure data is acquired at new wellbore locations. 2) pressure data are inverted for the permeability structure. 3) estimation of the data and model covariance. 4) compute the $E(\alpha)$ and $S(\alpha)$. Finally, find the location that maximizes equation 3.12.

Before examining this two-stage approach for improving our wellbore placement on

synthetic simulations, we want to present a simple example to illustrate the method described in stage 2. Consider a nonlinear model that maps a parameter vector into data. The model is as follows:

$$d(x) = m_1 x + m_2 e^{-m_3 x} \tag{3.13}$$

The model can be defined by $F(m) = d + e$. F is the nonlinear operator that solves eq3.13. m is a vector of unknown parameters. d is the observed data, and e is a random data noise. Through a least-squares method, we can solve the objective function that minimizes the L_2 norm between observed data and a modeled solution:

$$\theta(m) = \|F(m) - d\| \tag{3.14}$$

To illustrate the survey design, we will start with two observation locations. The first inversion for the parameter will result in a fit to the data Figure 3-9a. We can, therefore, estimate the model covariance at all possible locations and ultimately evaluate the probability of the data using a Gaussian distribution, assuming it has a Gaussian distribution. The estimation of the model covariance matrix Cov_m can be explicitly computed from the generalized inverse matrix J^{-g} through:

$$Cov_m = J^{-g} J^{-g^T} \tag{3.15}$$

The probability of the data enables us to estimate the Shannon information entropy for all possible locations. The spread of the resolution matrix is also estimated for all possible locations. Both $E(\alpha)$ and $S(\alpha)$ are normalized between zero and one, as graphed in Figure 3-9b. The lowest points in the graph represent the two starting observation locations. The highest points will determine the new survey locations. Figure 3-9 illustrates the adaptive approach to surveying this problem.

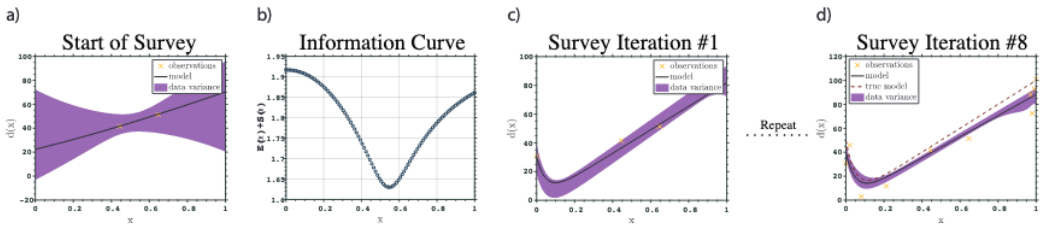

Figure 3-9: a) The two crosses are the observation points and their associated $d(x)$ values. The solid line is the best fit model for the observation locations. The shaded purple area highlights the data variance. b) The curve is the $E(\alpha) + S(\alpha)$ at different observation locations. c) The new observation point appeared at the beginning of the survey line with the new estimated model and the data variance. d) The previous steps are repeated eight times until we finalize a total of 10 survey locations. On top of the best fit model, we drew the true model in a dashed line.

3.4.2 Experiments

We now present the results from a synthetic experiment to illustrate our method. Here, we assume that permeability is the only parameter of concern. Hence, the survey design will place wells in locations that minimize the non-uniqueness in the inverse problem and improve the imaging of the permeability structure. We assume that the forward fluid flow operator F is fully known, in addition to the boundary and initial conditions. The 2-dimensional reservoir permeability model has a square shape and size of 10^4x10^4m, as shown in black in Figure 3-10. Boundaries and initial conditions are set at constant pressure, and the producing well is fixed at the center (marked in red). The model is discretized into 11x11 grids to solve the finite volume fluid flow problem. Because we assume the production well is fixed at the center of the model, we want to explore the problem's non-uniqueness if we are to add another wellbore for monitoring. Figure 3-10 labels the 121 possible well locations with question marks. Which grid will result in the least measure of non-uniqueness if a wellbore is placed in it? Figure 3-10b color codes the measure of non-uniqueness if the wellbores were to be placed at each possible location. We choose the grid with the minimum non-uniqueness value to be the optimum location of the second well (marked by a green circle). The

simulated pressure data is then computed at these locations over several days.

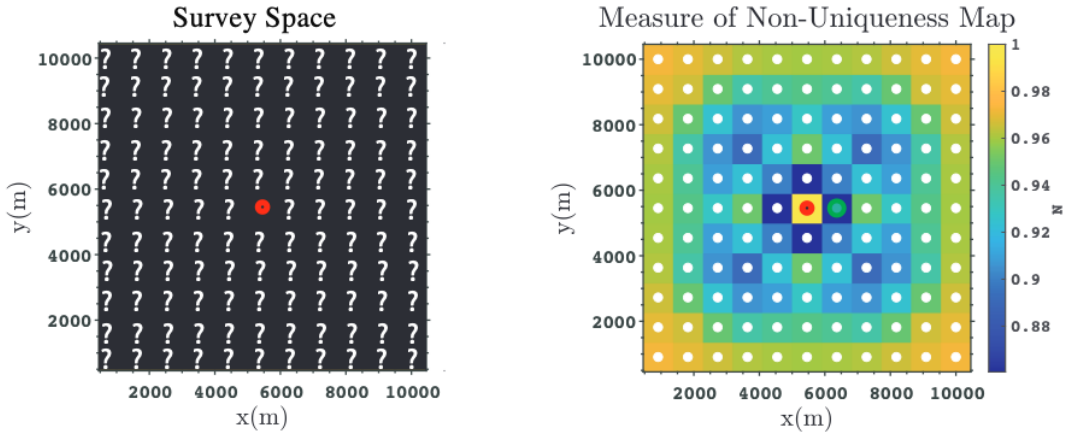

Figure 3-10: a) Reservoir space with the fixed red circle highlighting the production well location. Question marks label the possible additional wellbore locations. b) the measure of non-uniqueness of the inverse problem if the second well was placed at any survey grid locations. The green circle represents the location with the least value of the non-uniqueness measure chosen as a borehole location.

Afterward, we adopt the method described in the previous section to find the optimum well sites that maximize equation 3.12. We aim to find a total of six wells for our synthetic experiment, including the two well locations chosen using the non-uniqueness measurements. There is no particular reason for deciding to optimize six well sites. Six is not too large a number to run the synthetic experiment. Therefore, the second stage of the method attempts to find the optimum locations of four additional wells. The beginning of stage two relies on having an initial image of the permeability structure. Therefore, we inverted the pressure data obtained from the starting two wells for the permeability structure. The permeability can enable us to simulate the fluid flow in the structure and, therefore, predict the pressure data for all discretized grids of the model. Assuming the pressure data has Gaussian distribution, we can also estimate the data's probability for all grids. We can then compute the Shannon information entropy and run the inversion multiple times to assess the spread of the resolution matrix. Figure 3-11 illustrates how we achieved the optimum wellbores locations in

several iterations. The top image displays the wells' starting locations with the true permeability reservoir model in the background. The curves in the central figures are the observation data overlayed by the modeled data. The third column in the figure presents the inverted images. After each inversion, we construct an information image that displays the value of the $E(\alpha) + S(\alpha)$ at each grid. The highest value in the information map determines the location of the new wellbore for the next iteration. We repeated the process four times to find the optimum location for six wellbores. At the end of each survey, we invert for the permeability structure with the measure of the model-RMSE (root mean square error between the inverted model and the true model) and the data-RMSE (root mean square error between true data and predicted data). A summary of the iterative survey is presented in Figure 3-11

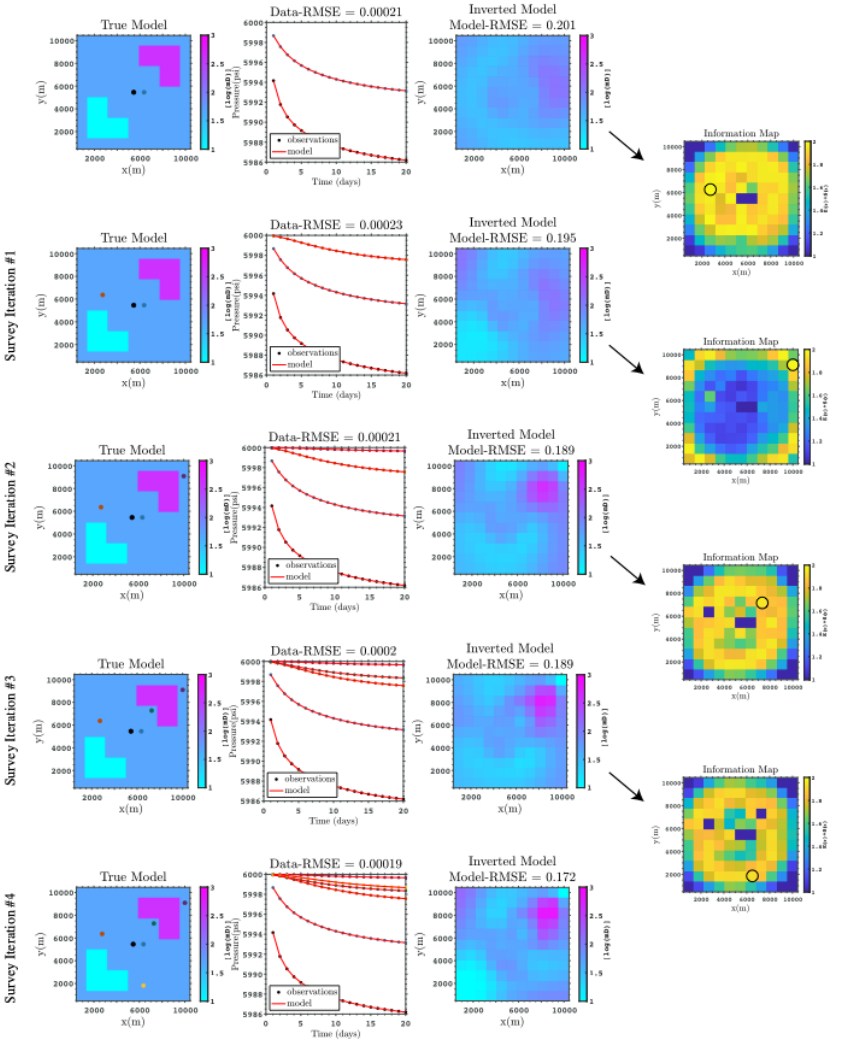

Figure 3-11: Summary of the four adaptive survey iterations. The first column of figures on the left shows the true permeability structure, highlighting wellbore locations in different colors. The number of survey iterations increases from the starting survey at the top until the final iteration at the bottom. The second column graphs the observed pressure data colored by the wellbore dot fitted by the best-inverted model. The third column highlights the inverted model, given the pressure data at the wellbores in column one. Finally, column four is the information map for the $E(\alpha) + S(\alpha)$ at each grid. The computation of the information maps happens after each inversion. The black circles on the information maps mark the location of the newly added wellbore for the next survey iteration.

3.4.3 Discussion

The current advancement in computational powers enables engineers to model fluid flow in a reservoir with higher accuracy and faster convergence. However, the inverse process to the forward modeling will still have accuracy issues. The inversion will mostly lead to a non-unique solution, especially for a model with many parameters. Hence, matching the observed data with a unique model is impossible and comes with the cost of the resolution.

With the described method in this study, we can expedite the characterization of the reservoirs' properties while aiming to find a unique solution. The demonstrated experiment shows that the higher the complexity of the reservoir, the more accurate the model and the less its data error. However, this comes with the cost of computation. Figure 3-12 compares a different number of grids versus the data-RMSE, model-RMSE, computation time, and parameter resolution. The model root means squares error (model-RMSE) measures the image difference between the reduced model and the true big model. The reduced model is scaled into the size of the big model to be able to compute the difference. The spread of the resolution matrix, described in Chapter 2, measures how well the parameters are resolved. The computation time increases exponentially as the number of grid increase. The data-RMSE also shows an exponential decrease in the error as the number of grids increases. However, we have noticed that for specific grid sizes (50, 100, and 300), where the models are not too small, the data-RMSE values are within the same error range. In contrast to the exponential change of the data-RMSE, the spread and the model-RMSE decrease linearly with the increase in the number of grids. Although in a real case scenario, we do not have access to estimating the model-RMSE, it provided insights for the presented synthetic experiments. Models with 1600 and 2500 grids are better at representing the complex reservoir than the model with 6400 grids based on the model-RMSE. In the end, we saw that most grids that are not very small and not too big are adequate and,

in some cases, more accurate than large models in representing the true models with a significant reduction in computation time.

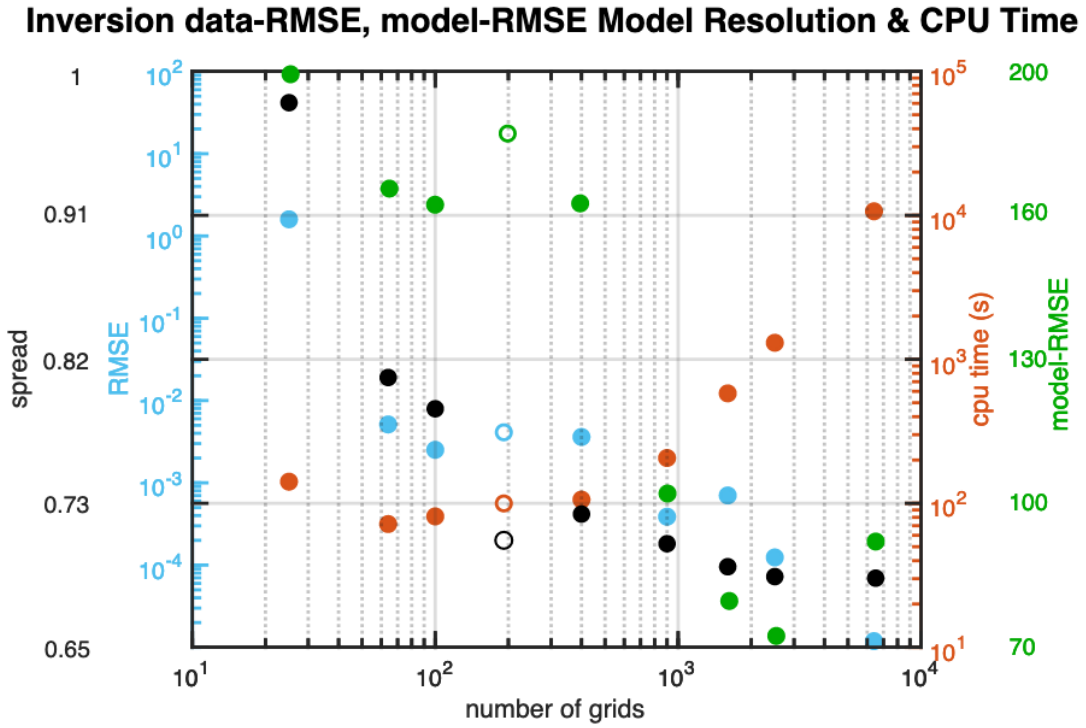

Figure 3-12: The data-RMSE, model-RMSE, spread, and CPU time resulted from inversion for the reservoir parameter at a different number of grids. The filled circles are the result of the inversion on a regular grid. The open circles are the results of inverting the data on the irregular refined grid.

In the second section of this chapter, we introduced a method to optimize wellbore locations. Our approach to optimizing wellbore locations is targeted toward better data sampling to assist in inverting for a more accurate reservoir structure. The first stage of the method can measure the non-uniqueness of any survey design. We explored the technique with one wellbore, making the problem more straightforward. However, it is still possible to perform the analysis for multiple wellbores. It would be, however, much more computationally expensive. The measure of non-uniqueness is directly related to the forward operator and its embedded features (e.g., boundary and initial conditions, size) and designs (e.g., location and number of wells). Also, the

method showed the concept of utilizing the Shannon information entropy and the resolution matrix to optimize the wellbore placement in a hydrocarbon field. Figure 3-13 demonstrates inversion results for various randomly generated survey designs. Most of these random designs have not outperformed our method. However, there are few cases where well placement has produced a better-resolved model than the solution obtained using our proposed method. Therefore, while the method does not find the global optimum survey, as expected by a greedy-adaptive optimization scheme, it finds a wellbore location configuration whose performance lies within the vicinity of the more optimal designs.

One of the most significant advantages of this scheme is the ability to scale and parallelize in terms of computation. The most time-consuming process is obtaining the information map, which requires the calculation of the resolution matrix multiple times. This process can be parallelized and reduce the computation time.

Finally, this method can be applied to any geophysical survey as long as we have an accurate forward model and a stable inversion scheme. The presented example assumes that the wellbore data follows a Gaussian distribution. This is not a limiting assumption, as applying this method using other distributions is still possible by modifying the probabilities in the computation of the Shannon information entropy. Although this approach was applied to one type of problem, it can still be applied to many geophysical and non-geophysical problems.

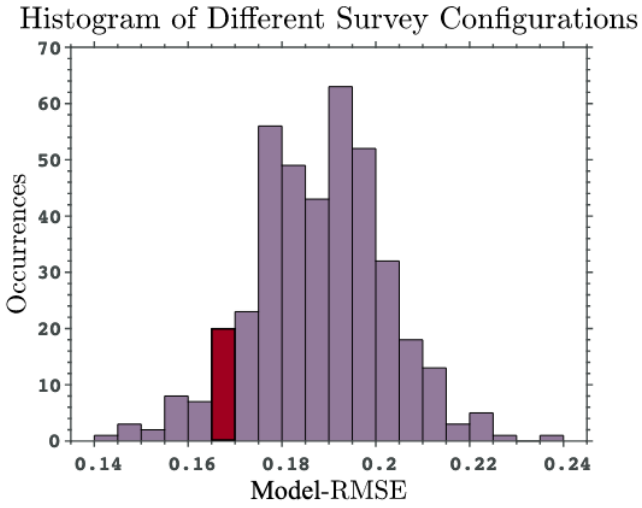

Figure 3-13: The graph displays the model-RMSE from different distributions of the four wells. The highlighted bar represents the model-RMSE for the design obtained using our proposed method.

Chapter 4

Joint Self-Potential and Fluid Flow Inversion for Imaging Permeability Structure and Detecting Fractures

4.1 Abstract

Detecting fracture locations and permeability structure of a subsurface reservoir accurately could aid in optimizing production performance. Optimizing the performance of subsurface groundwater or hydrocarbon reservoirs depends on having an accurate characteristic map of the reservoir. In pursuing the history matching process, someone must utilize the available well data, such as downhole pressure and flow rate, to estimate the reservoir characteristics that produce a response similar to the historical data. The simulators often repeat the forward modeling process until good simulation results match the historical data. However, matching the historical data is costly and will likely lead to a non-unique reservoir model. We developed an algorithm that jointly uses bottom-hole pressure data with surface Self-Potential (SP) measurements to invert for (permeability) structure. SP is an inexpensive geophysical survey that

enables one to invert for what is called subsurface source "geo-battery," knowing the resistivity structure in the subsurface. This article presents different synthetic experiments to show the improvement in characterizing reservoirs when different data sets (borehole pressure and surface self-potential) are used.

4.2 Introduction

Fractures play an essential role as fluid corridors in groundwater and energy reservoirs, which makes the characterization of fractures critical for developing subsurface fluid resources. The presence of clusters of fractures in an underground reservoir might result in a variation in the flow rate. From an economic point of view, the variation in flow rate can result in unstable economic returns. Interest in fracture characterization has grown in recent years due to the growth of unconventional resources (Cueto-Felgueroso and Juanes, 2013.) Characterizing the fractures can significantly aid in optimizing the production performance of a reservoir. Various surface geophysical surveys and underground logging methods can provide information about fractures in the subsurface. However, imaging fractures with these methods is very complicated due to the size of fractures, the contrast of the fracture properties to the surrounding materials, and the noise in the collected data. For instance, relying on seismic methods to interpret naturally occurring fractures in a reservoir has significant uncertainty (Burns et al., 2007; Alfred Lacazette et al., 2013). The high-resolution seismic survey might be able to image large fractures but does not provide a probability of the fractures being productive or permeable. Also, Imaging fractures with borehole logs lack spatial resolution far from the boreholes. The sparsity of the borehole data limits the imaging resolution of the fractures. Therefore, combining multiple measurements for characterizing fractured geological reservoirs is more applicable. Conventionally, seismic interpretation is independently used with fluid flow modeling to characterize fractures. However, improvement in imaging fractures has been achieved on synthetic data when

seismic scattered wavefield data is combined with flow data in a sequential framework (Kang et al., 2016). Although the joint seismic-flow framework works well, the seismic method is costly, and most migration and scattering imaging techniques have limited capability in detecting multiple fractures (Hao Hu and Yingcai Zheng, 2018). Thus, the alternative way is to obtain a different inexpensive geophysical measurement to aid in imaging the fractures.

The passive Magnetotelluric (MT) or transient electromagnetic (TEM) methods are methods that can be used besides the fluid flow data to explore the subsurface fluid system. The currents in the subsurface can be detected using MT and TEM at great depth at the cost of high resolution. An example of using TEM in characterizing near-surface structures is discussed in (Colombo et al., 2008). However, using these methods in a 2D or 3D environment is very expensive and time-consuming.

The self-potential method is a useful geophysical technique complementing the borehole flow data because of its sensitivity to fluid flow in porous media and its simple implementation. Joining two or multiple data sets in an inversion algorithm can fall under two categories. 1) The parameters obtained from the different methods are similar or linked by a relationship. 2) There is no apparent relationship between the different parameters, but there is a structure similarity (Hyndman et al., 1994; McKenna and Poeter, 1995; Linde et al., 2006; Irving and Singha, 2010). Both categories apply in the case of jointly inverting fluid flow and self-potential data, unlike the fluid flow and seismic, where only structure similarity could be applied. That is because There is no accurate or unified relationship linking the seismic parameters to the fluid flow parameters.

The observed surface self-potential (SP) signal is a function of the Streaming-Potential properties of the rock/fluid couple and the gross electrical conductivity structure. This phenomenon is related to an electrical double layer (EDL)(Davis et al., 1978). This double layer is formed by a Stern layer of weak or strong ions and a diffuse layer in which the ionic concentrations obey Poisson-Boltzmann statistics (Chapman,

1913). In fluid-saturated rocks, the mineral surface has net charges resulting from a chemical reaction between the mineral and the pore fluid. In the presence of a pore pressure gradient in porous media, fluid flow drags the electrical charges of the diffusive layer. The drag creates an electrical current called the streaming current, which is the source of an electrical field known as the streaming potential (Figure 4-1) (e.g., Ahmad, 1969; Morgan, 1989; Revil et al., 2003; Jardani et al., 2006). The streaming potential can be measured in saturated or unsaturated conditions using a network of nonpolarizable electrodes connected to a digital multichannel multimeter. (e.g., Alkafeef et al., 2001; Vichabian and Morgan, 2002; Vinogradov and Jackson, 2011; Revil and Mahardika, 2013; Morgan et al., 2019). Fluid behavior in a fractured medium results in variations in the Self-potential signals. The fluid flow behavior in a fractured medium directly affects the SP signals. Several experiments managed to record SP responses by injecting water into a fractured aquifer (DesRoches et al., 2018). And others used SP data to detect groundwater flow in fractures (e.g., Fagerlund and Heinson, 2003). In fact, (SP) method is the only geophysical method directly sensitive to the flow in the subsurface (Revil et al., 2005). Therefore, the SP method is very applicable for imaging fractures. The method of combining flow data with SP has been used in hydrogeophysics in a growing number of publications (e.g., Wurmstich and Morgan, 1994; Suski et al., 2006; Jardani et al., 2006, 2009; Revil and Mahardika, 2013; Soueid Ahmed et al., 2016). However, none of these studies have examined imaging fractures by combining the different data sets. This chapter describes a method to detect naturally occurring fractures in reservoirs and permeability structures by combining borehole pressure data and SP surface data in a joint inversion solution. An accurate method for forward modeling SP signals and fluid flow in fractures is essential. Several synthetic experiments were generated modeling the fluid response and SP signals in a fractured medium. The methods for modeling the synthetic data sets are explained in section 4.3.1. The final inversion results will demonstrate how these two data sets can aid in detecting the fractures and improve resolving of the permeability anomalies.

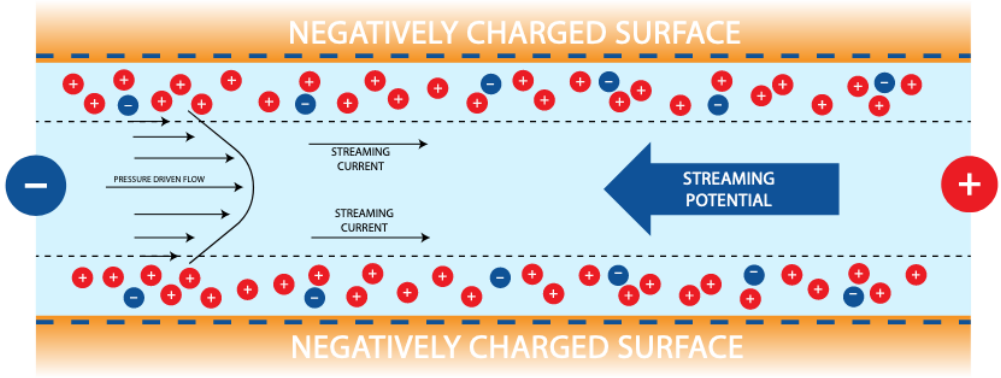

Figure 4-1: Saturated rocks with fluid naturally develop an electric double layer. The streaming current associated with the pressure gradient moves the positive charges toward the right end, creating a streaming potential in the opposite direction.

4.3 Method

4.3.1 Forward modeling

The forward modeling of SP signals relies first on solving the fluid flow problem. The fluid flow simulation solves the transient pressure diffusion equation in eq4.1, given the fluid and the static rock parameters.

$$\frac{1}{\mu}\nabla \cdot ([K]\nabla P(x,y,z,t)) = \phi C_t \frac{\partial P(x,y,z,t)}{\partial t} + q(x,y,z,t) \tag{4.1}$$

where μ is the fluid viscosity in (cp), $[K]$ is the permeability tensor in (m^2), P is pressure in (psi), ϕ is porosity, and C_t is the total compressibility in (psi^{-1}). The source/sink term q has a unit of m^3/s. The fluid flow problem is solved using an existing Matlab tool (MRST (Lie, 2019)) in addition to our own finite volume solver. The solver utilizes the embedded discrete fracture model (EDFM) method to model fluid flow in fractured media (Hajibeygi et al., 2011).

Once fluid pressure is solved for the media, the SP response is modeled by solving

coupled systems of equations (Sill, 1983). Pressure measurements are used as a Self-Potential signal generator. If we assume that the only thing that carries a charge in the subsurface is fluid, then the coupled problem is written as follows:

$$j = -\sigma \nabla U - A_{EK} \nabla P - A_{ED} \nabla C - A_{TE} \nabla T \tag{4.2}$$

where j is the current density $(A \cdot \mathrm{m}^2)$, σ is the rock electrical conductivity $(s \cdot \mathrm{m}^1)$, ∇U is the electrical potential gradient $(V \cdot \mathrm{m}^{-1})$, ∇P is the fluid pressure gradient $(\mathrm{Pa} \cdot \mathrm{m}^{-1})$, ∇C is the concentration gradient $(\mathrm{M} \cdot \mathrm{m}^{-1})$ and ∇T is the temperature gradient $(\mathrm{K} \cdot \mathrm{m}^{-1})$. The terms ($A_{EK}$, A_{ED} and A_{TE}) are the electrokinetic, exclusion-diffusion, and thermoelectrical cross-coupling terms resulted from Onsager's irreversible thermodynamic theorem (Onsager, 1931; Bear, 1988). These cross-coupling terms are functions of the electrical conductivity of saturated rock (σ) and a coupling coefficient denoted by (c):

$$A_{EK} = \sigma c_{EK}, A_{ED} = \sigma c_{ED}, A_{TE} = \sigma c_{TE} \tag{4.3}$$

The coupling coefficients introduced above are petrophysical properties measured in a lab that relate gradients of pressure, concentration, and temperature to conductivity by empirical relationships.

Following the continuity equation of an electric charge gives $\nabla \cdot J = 0$ (Sill, 1983). By combining eq4.2 and eq4.3 while considering the continuity equation, we can write:

$$\nabla \cdot (\sigma \nabla U) = -\nabla \cdot (A_{EK} \nabla P) - \nabla \cdot (A_{ED} \nabla C) - \nabla \cdot (A_{TE} \nabla T) \tag{4.4}$$

To simulate the Self-Potential (U), eq4.4 is discretized and solved for the given values of σ, coupling terms (c_{EK}, c_{ED}, c_{TE}), and (P, C, T). Pressure, concentration, and temperature are obtained by solving Darcy's law, Fick's law, and Foriour's law. In this article, the concentration and temperature effects are ignored for simplicity. Therefore,

eq4.4 can be reduced to the following:

$$\nabla \cdot (\sigma \nabla U) = -\nabla \cdot (A_{EK} \nabla P) \tag{4.5}$$

This is the Poisson equation of the self-potential field U. The right-hand side is the source term associated with the pore pressure from fluid flow in a reservoir. Eq4.5 can be solved analytically to compare the result with the finite volume method adopted in this study. By treating the right-hand side of eq4.6 as a source (s) the equation becomes:

$$\nabla \cdot (\sigma \nabla U) = s \tag{4.6}$$

Assuming a 2D space, the solution to $U(x, y)$ is solved by obtaining the green's function $G(\xi, \eta, x, y)$ that satisfies the following equation with the boundary conditions.

$$\nabla \cdot (\sigma \nabla G) = \delta(\xi - x, \eta - y) \tag{4.7}$$

where δ is a two-dimensional delta function with ξ and η as dummy variables. Solving for the G enables the solution for the potential U (eq4.8). Detailed presentations of green's functions in solving differential equations are found in textbooks (e.g., Riley et al 2002, Greenberg 2015)

$$U(x, y) = \int \int G(\xi, \eta; x, y) s(\xi, \eta) d\xi d\eta \tag{4.8}$$

A test model is considered to verify this study's numerical finite volume method to simulate the SP data. The result of the numerical method is compared to the analytical solution. Consider a steady-state problem of a circular-grid reservoir with homogenous fluid and electrical properties (Figure 4-2). Details of the reservoir parameters are presented in Table 4.1. The model is bound by Dirichlet boundary conditions and has a single source point in the middle. For this particular situation, the solution to the

potential is in the form:

$$U(r) = \frac{s}{2\sigma\pi} ln(r) + A \tag{4.9}$$

where $r = \sqrt{(\xi - x)^2 + (\eta - y)^2}$ is the distance from the source, and A is a constant dependent on the boundary conditions. The source term (s) can be obtained by considering the steady-state version of the fluid flow (eq4.1).

parameters		values	units
Model			
	dimension	x=1000, y=1000, z=10	m
	initial pressure	100	bar
	boundaries	pressure@100	bar
	source	1	m^3/day
Rock			
	permeability k	10	mD
	porosity	.20	
	conductivity	0.001	S/m
	Coefficient factor (c_{EK})	$1e^{-6}$	$\frac{S}{m}\frac{V}{bar}$
Fluid			
	density	1000	kg/m^3
	viscosity	1	cP
Total compressibility		$1e^{-4}$	1/bar

Table 4.1: The parameters used for modeling the validation example.

4.3.2 Inverse Modeling

The objective of the inverse modeling is to obtain a permeability structure of a reservoir from the surface-observed SP signals and downhole pressure measurements. The procedure involves subdividing the subsurface into prisms of equal size. Each prism will have constant permeability. The fractures are represented as embedded layers and not as grids. By altering the permeability at each grid and simulating the transient pressure and SP signal, it is hoped that there will be a permeability structure that matches some observed SP signals and pressure measurements. Figure 4-4 illustrates the main steps taken in forward and inverse modeling.

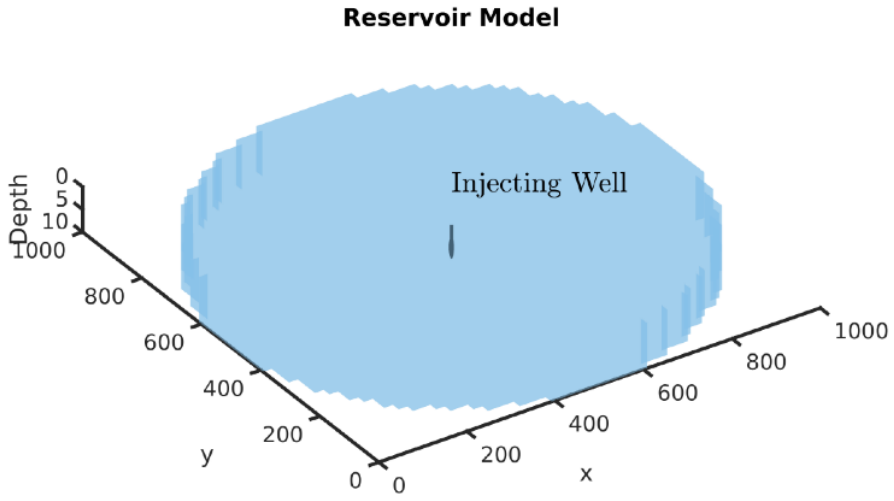

Figure 4-2: A two-dimensional circular homogenous reservoir model.

Assuming that the only unknown parameter in the subsurface is permeability k, the forward problems are simplified by nonlinear forward operators; F for pressure and

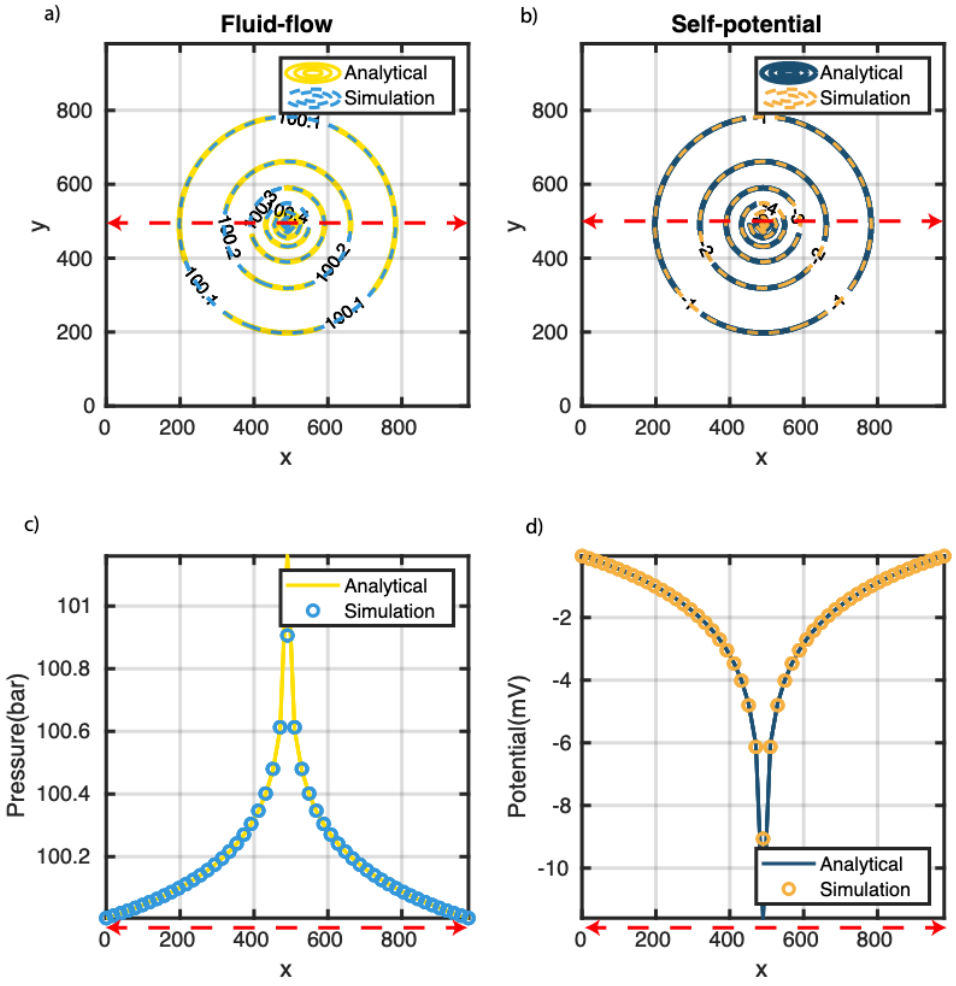

Figure 4-3: Results of solving the fluid-flow and self-potential problem analytically and numerically. a) A bird's view showing the contour of the pressure by the analytical solution (solid line) and numerical solution (dashed line). b) Contours of the potential solved analytically (solid line) and numerically (dash line). c) The analytical and numerical pressure solutions across the reservoir, where the (red line) is drawn. d) The potential across the reservoir (red line) is solved analytically and numerically.

S for the self-potential.

$$d_p = F(m), d_{sp} = S(m) \tag{4.10}$$

where m is a vector containing the subsurface's permeability values at each discretized grid. d_p and d_{sp} are the pressure data and SP data, respectively. The length of the

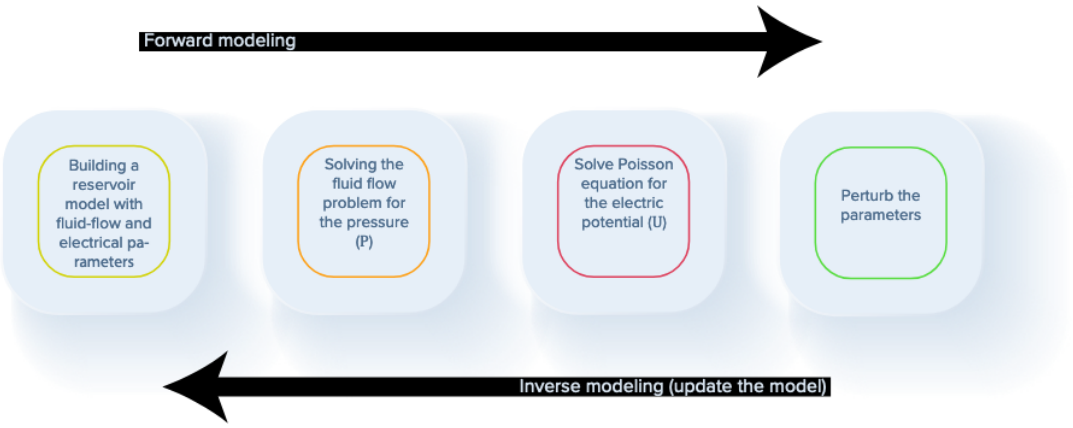

Figure 4-4: Illustration of the forward and inverse modeling workflow.

vector d_p is specified based on the number of observational wellbores multiplied by the time steps. The vector d_{sp} has a length defined by the number of SP data locations multiplied by the time steps. The two data vectors are not necessarily the same length. With the knowledge of the forward modeling of the two data sets, the goal of the joint inversion is to obtain (m) that minimizes the objective function in eq4.11:

$$\phi(m) = \| W_p(F(m) - d_p) \|^2 + \| W_{sp}(S(m) - d_{sp}) \|^2 + \alpha^2 \| Lm \|^2 \tag{4.11}$$

The objective function $\phi(m)$ is a summation of L_2 norms for the different data sets with a regularized term. The L_2 norm is usually a standard loss function in geophysics due to its stability in finding a unique solution compared to the L_1 norm. The first term $(\| W_p(F(m) - d_p) \|^2)$ is the norm between the modeled pressure data and the observed pressure data weighted by W_p. The second term $(\| W_{sp}(S(k) - d_{sp}) \|^2)$ is the norm between the modeled SP data and the observed data weighted by W_{sp}. The weight terms; W_p and W_{sp} are diagonal matrices with the reciprocal of the data uncertainty populating the diagonal. The final term $(\alpha^2 \| Lk \|^2)$ is a regularization term defined by the regularization matrix L and a penalty parameter α (Tikhonov,1977). The regularization matrix L consists of gradient fields in the $x, y,$ and$z$ directions to force

parameters' smoothing. The misfit guides the choice for the parameter α in the data to lead the inversion convergence into the objective function minimization. The data misfit is the root mean square error (RMSE) between the observed data d_{obs} and the modeled data d_m.

$$RMSE = \frac{1}{n}\sqrt{\sum (d_{obs} - d_m)^2} \tag{4.12}$$

Eq4.11 can be simplified further by combining the operators, the observed data, and the weights into large matrices and vectors such as $d_{obs} = [d_p d_{sp}]$, $G = [FS]$, and $W = [W_p W_{sp}]$.

$$\phi(m) = \| W(G(m) - d_{obs}) \|^2 + \alpha^2 \| Lm \|^2 \tag{4.13}$$

The minimizing problem is solved in an iterative process. The perturbation of the parameter vector m is estimated using the damped least square method. The Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1964) method is adapted to update (α)in each iteration.

4.4 Numerical Result

The synthetic experiments can provide verification and insight into the methods described in Section 2. The objective of each experiment is different, and the analysis of its result can assess the method's usefulness. There are a total of three experiments divided into two types of experiments; structure imaging and fracture detection. The overall objective is to examine the improvement in estimating the reservoir flow parameters by jointly inverting the borehole pressure data and the Self-potential data. The first experiment focused on estimating permeability anomalies in the reservoir. In the rest of the experiments, the joint inversion scheme images naturally occurring fractures in reservoirs.

4.4.1 Structure Imaging Experiment

This experiment will attempt to invert for permeability structure in a confined reservoir. For example, a realistic scenario to structure imaging might be identifying fluvial channels. At the end of this experiment, we expect to compare images obtained from inverting, single data sets, and two data sets. The synthetic data used for the experiment results from simulating the forward models on the fluid-flow and electrical models (Figure 4-5). The model is a 3D cubic model (dimension 1000 m in($x, y,$ and z) with a homogenous background permeability and two structural anomalies; one has a high permeability value, and the other is a low value. We introduced a producing well at the model's center and two monitoring wells on different sides. The depth of each well is 500 m. The model's top and bottom have a no-flow boundary condition, while the sides have constant pressure conditions. The initial condition is set as a constant pressure everywhere within the model, similar to the boundary pressure. The initial pressure is 200 bars everywhere. We simulated the problem on a discretized regular grid size (11x11x11) in $x, y,$ and z. We estimated the transient pressure data at the three wells using the parameter presented in Table 4.2. In addition, the fluid pressure at all grids of the domain has to be estimated as it will be an input for forward-modeling the Self-potential measurements.

The parameters for the self-potential physical model are defined for the conductivities and coefficient factor (present in Table 2). The conductivities and coefficient factor values are within the range observed for sandstone, similar to values used in (Sheffer & Oldenburg). The self-potential data were estimated at surface electrodes spaced at approximately 90 m, equal to the grid spacing. A reference electrode is placed far from the experiment region, so the fluid flow does not influence it. The domain's top and bottom boundaries are set as insulated boundary conditions, while a constant voltage is set at the domain's sides to match the fluid flow's constant pressure boundaries. The simulation for both models, fluid flow and SP, span for a period of 20 days with data

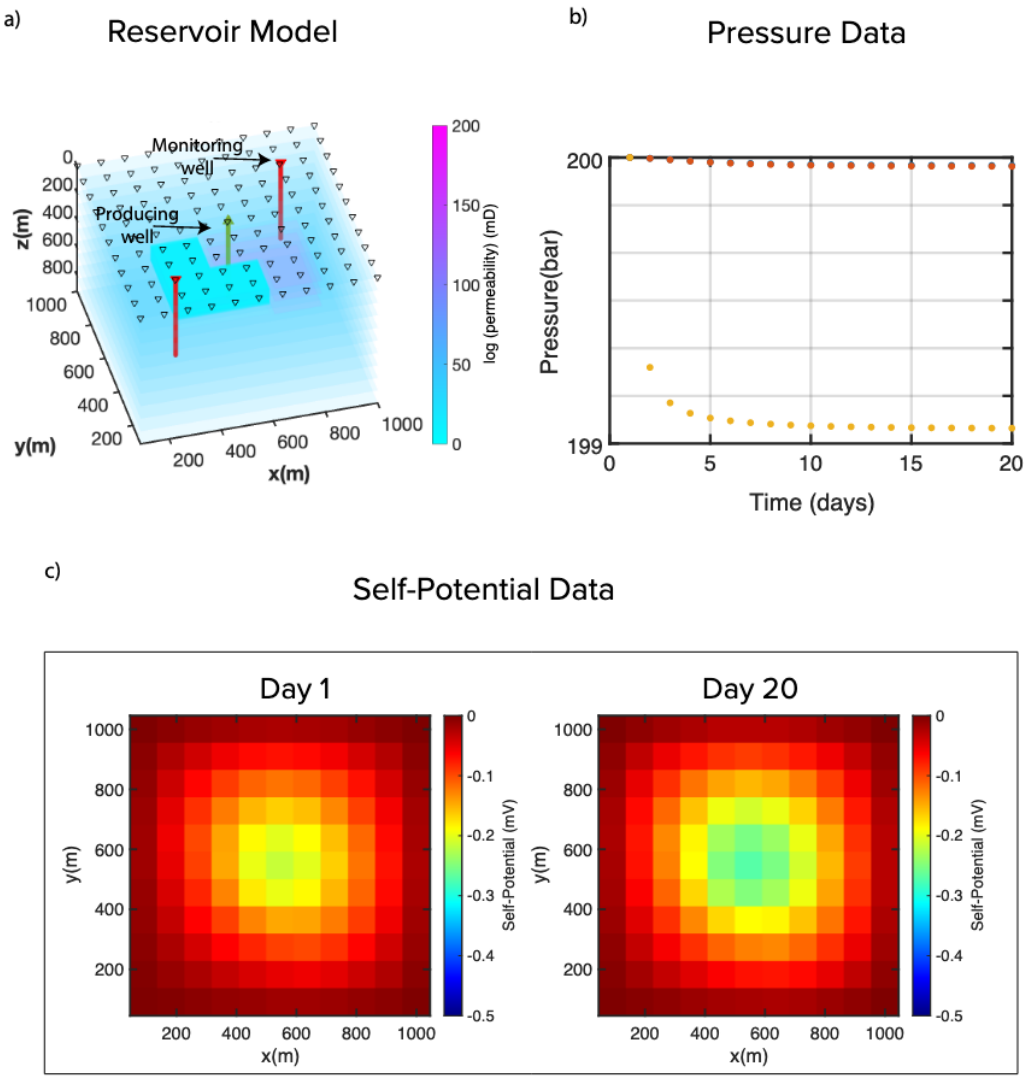

Figure 4-5: a) a 3D reservoir model with permeability anomalies. There are two monitoring wells (red) and one production well (green). The black triangles represent the locations of the SP electors. b) The pressure measurements with time at the three well locations. c) Self-Potential distribution at the ground level is shown at different simulation times. a) Self-Potential in mV after ten days of pumping fluid out of the reservoir at 100 $m^3 day^{-1}$. b) after 20 days of pumping.

collected with an increment of one data point per day. Pressure data and SP data are presented in Figure 4-5. The Self-potential data distribution is shown at two different simulation times.

parameters		value	units
Rock			
	Matrix permeability	100	mD
	Anomaly permeability	10,1400	mD
	porosity	.10	
	conductivity	0.001	S/m
	Coefficient factor (c_{EK})	$1e-6$	$\frac{\mathrm{S}}{\mathrm{m}}\frac{\mathrm{V}}{\mathrm{bar}}$
Fluid			
	density	1000	$\mathrm{kg/m^3}$
	viscosity	1	cP
Total compressibility		$1e-4$	1/bar

Table 4.2: Parameters details for the permeability structure experiment.

The method of jointly inverting the two data sets should aid in improving the resolution of the permeability anomalies; at least, that is the expectation. The joint inversion is compared against a result of inverting only the pressure data set. The objective function inverting one data set does not differ much from eq4.11. The first or the second terms in eq4.11 can be ignored depending on the data used. For this comparison, the fluid flow problem is considered the individual data set; therefore, the SP data norm is eliminated in eq4.11. The iterative process of inversion starts with a homogenous model. The final image is the algorithm's output stopped when the RMSE (eq4.12) difference is small ($< 10^{-6}$) between the previous and the current iteration. The RMSE resulting from inverting the pressure data only suggested an almost perfect fit, although the inverted image differs from the true model. The overall structure of the true permeability anomalies was not resolved, nor were their average values. The pressure data from the three wellbores is insufficient to resolve the model's permeability. However, the addition of the SP surface data has performed better in resolving a more accurate structure without adding more wellbores. The structure of low anomaly at a shallower depth (Figure 4-6) has been delineated with the help of the

SP data. Even though the deeper structure was not fully resolved, the joint inversion still estimated a high permeability value.

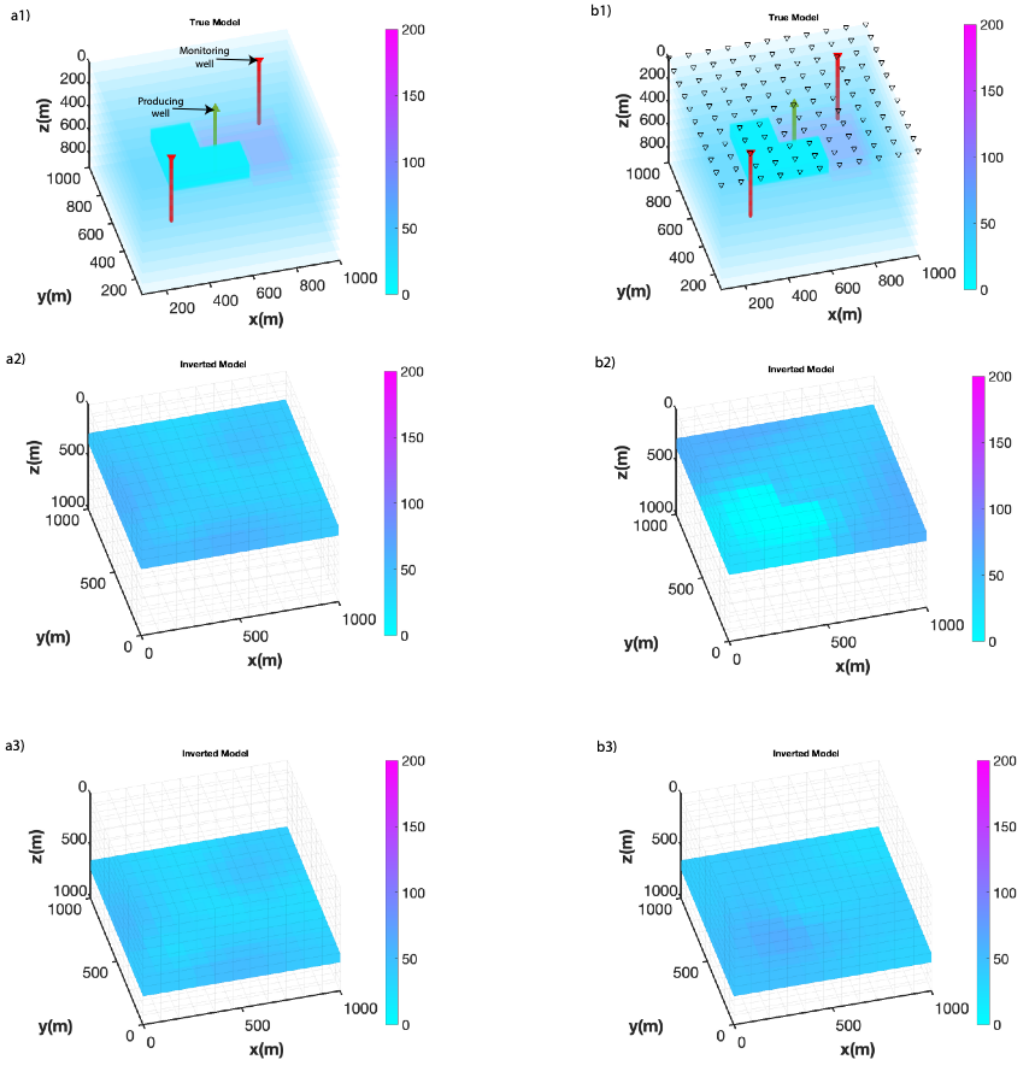

Figure 4-6: (a1) is a 3D synthetic model with two permeability anomalies at different depths. (a2) shows the inversion image of a shallow using data from the three wells. (a3) shows the inversion results at a deep layer. The locations where SP measurements were recorded are highlighted in triangles (b1). (b2) shallow inversion results from the joint inversion method. (b3) joint inversion results at the deep layer.

4.4.2 Fractures Detection Experiments

The presented experiments in this section investigate the effectiveness of adding a surface SP to the pressure borehole data to image or detect random natural fractures in reservoirs.

One Fracture Experiment

The simplest case to investigate fracture detection might be considering one fracture in a reservoir. A computer-generated model of a reservoir with one fracture is built to simulate the synthetic data. The model is populated with fluid flow and electrical parameters for simulating the data.

The reservoir model has a homogenous background matrix permeability with one fracture across the reservoir. The fluid flow parameters of the fracture are different from the background parameters. Details of the properties of the model are presented in Table 4.3. The relationship between the fracture permeability and its aperture is based on the parallel plate model:

$$k_f = \frac{a^2}{12} \tag{4.14}$$

where k_f is the fracture permeability and a is the fracture aperture. The other parameter in Table 4.3 are gathered from modeling experiments presented in (DesRoches et al., 2018)

The solution for the pressure field is solved on a discretized three-dimensional grid (21x21x3) with an extent of (500 m x 500 m x100 m) in the x, y, and z. The top and bottom are no flux boundaries, while the model sides are set at constant pressure. The transient pressure data are collected from three wells; two injecting wells at the corners and one producing well at the center, as illustrated in Figure 4-7. Injecting wells are labeled "I," while producing wells are labeled "P". The perforation depth of all wells is 50 m below the surface and set to pump fluid at constant rates. The two injection

wells, I_1 and I_2, inject fluid into the reservoir at a rate of 100 m^3/day and 300 m^3/day respectively.

parameters		value	units
Rock			
	Matrix permeability	30	mD
	porosity	.10	
	conductivity	0.001	S/m
	Coefficient factor (c_{EK})	$1e-2$	$\frac{\text{S}}{\text{m}}\frac{\text{V}}{\text{bar}}$
Fluid			
	density	1000	kg/m^3
	viscosity	1	cP
Fracture			
	permeability	$1e7$	mD
	porosity	1	
	Aperture	$3.4e-4$	m
	conductivity	$5e-2$	S/m
	Coefficient factor (c_{EK})	0.4	
Total compressibility		$1e-4$	1/bar

Table 4.3: Details for simulating the one-fracture model experiment.

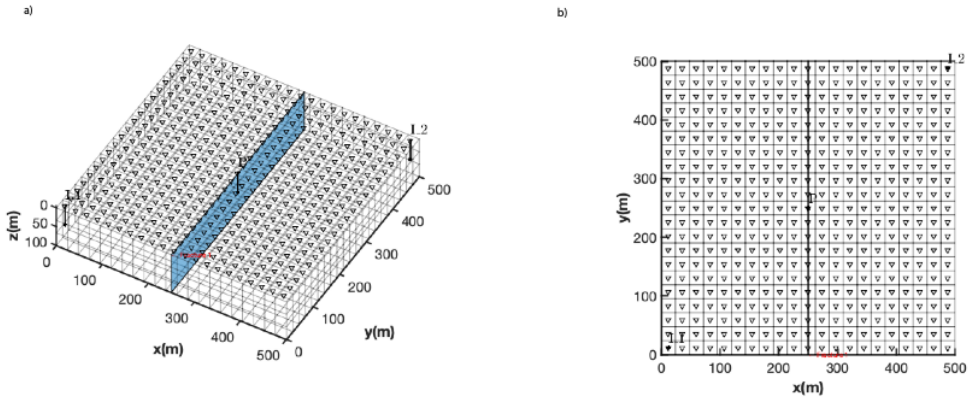

Figure 4-7: Two views of the 3-Dimensional survey area. a) is an angled view of the model showing the two vertical fractures and the locations of the wells. The production well is placed center of the fracture. The injunction wells are located at the corners. The permeability of the matrix is homogenous. The triangles are the location of the Self-Potential electrodes. b) is a bird's view of the survey area.

The SP measurements are collected at the surface grid (21x21), as illustrated by the triangles in Figure 4-7. SP electrodes are spaced at 25 m covering the surface grids.

Signals at each grid are referenced to a signal far away from the domain. Pressure and SP signals are collected over five days, with logarithmic increments to capture the pressure decrease at the first stage of the simulation. Figure 4-8 shows the pressure data over five days and the SP data over two days. The SP signals clearly show significant potential change at the location of the fractures.

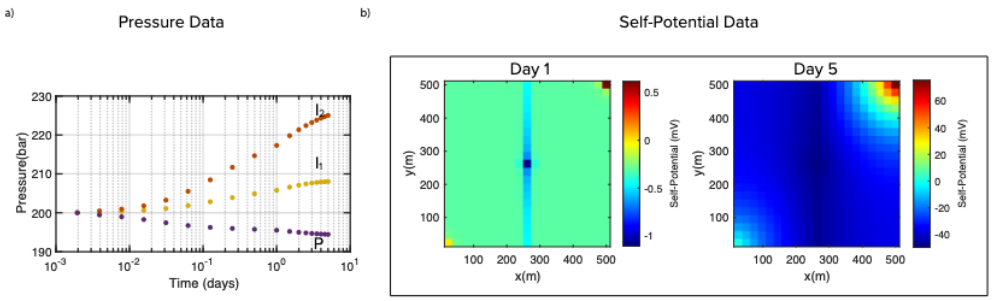

Figure 4-8: a) shows the pressure in bars at each well. Two wells are pumping at different constant rates. I_1 is injecting at 100 m^3/day and I_2 is injecting at 300 m^3/day. The production well is pumping at a 400 m^3/day rate. b) shows the recorded Self-Potential signals at two different periods.

The high magnitude of the SP is in the middle, where fluid pumping is an indication of streaming potential flowing in the opposite direction to the fluid flow. The inversion was performed on a regular grid similar to the original grid used for simulating the synthetic data. The only difference is in the fracture representation. When we simulated the true synthetic data, we embedded extra fine grids on top of the regular grid, to simulate the fractures' effect. On the other hand, the inversion grid consists of only the regular grid without the fine fracture grids. The first iteration in the inversion starts with an initial homogeneous permeability model. Following the same procedure as in the structure inversion (section 3.1), the iterative process was stopped when there was no large change in the RMSE. The fracture location was not resolved when attempting to detect the fracture with only the pressure data from the wells (Figure 4-9). However, with the added SP data, the high permeability regions where

the fracture is located were detected. Comparing the two solutions in terms of the RMSE between the observed data and the modeled shows a significantly lower error for the individual inversion of the pressure data. In the joint inversion, we are trying to match different data sets with one model, so it is expected to result in a higher RMSE. The two inversion algorithms performed their best to match the observed data despite the accuracy of the final inverted image.

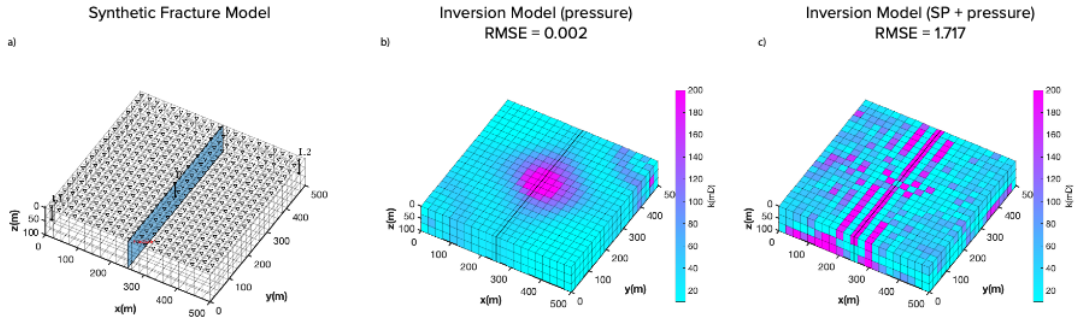

Figure 4-9: (a) A 3D synthetic reservoir model is designed with one vertical fracture. (b) Permeability structure resulted from the inversion using only the pressure data at well locations. (c) Final permeability structure utilizing pressure and SP data in a joint inversion. The Actual fracture locations are highlighted with a black line on the inverted images.

Multiple Fractures Experiments

The number of fractures in a reservoir can affect the inversion result. The fracture network in a reservoir can be complicated. In the previously presented experiment, there was only one fracture. Therefore, the evaluation of the result was easy. A more complex model is built to investigate the inversion approach (Figure 4-10).

The reservoir model used in this experiment to simulate the synthetic data is similar to the model used in the one fracture experiment. The only difference is in the number and the dimensions of the fractures. There are a total of ten randomly distributed fractures. Each fracture was placed by randomly picking its two endings from the grid.

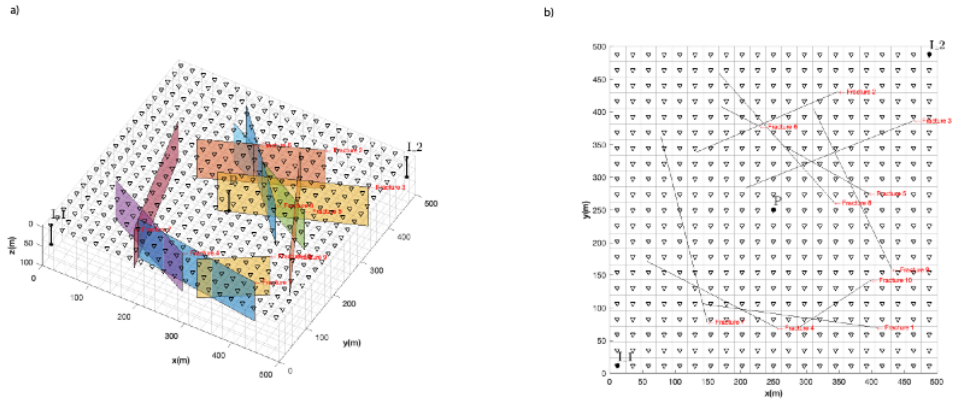

Figure 4-10: Two views of the 3-Dimensional survey area. a) Random fractures are presented in a humongous permeability background model. The three wells are highlighted by I_1, I_2 for Injection well number 1 and 2, and P for the production well. The fractures have different orientations and parameters (aperture, permeability). The triangles are the location of the Self-Potential electrodes. b) is a bird's view of the survey area.

The background matrix permeability is still homogenous. The Random fractures are introduced in the domain with different permeability and apertures. The permeability in the fractures ranges between (1 Darcy to 10e6 Darcy) and ranges in an aperture between (3.4e-6 to 1e-2). The SP signal in Figure 4-11 does not show any changes in the potential around the fractures as in the experiment (3.2.1). The existence of multiple fractures in the reservoir has distorted the SP signal, so it is impossible to distinguish different fractures by looking only at the SP data.

This experiment makes the difference between inverting single data sets and two data sets even more evident (Figure 4-12). An overlay of the fracture network is drawn on top of the final permeability image to help see the improvement in the detection. The joint inversion approach outperformed the inversion of individual data sets. The improvement in the result is judged based on the permeability value after the last iteration and not on the final RMSE. Inverting the two data sets has shown permeability variation near fractures despite the high RMSE compared to the inversion of the pressure data. Unfortunately, because the domain has a small number of grids,

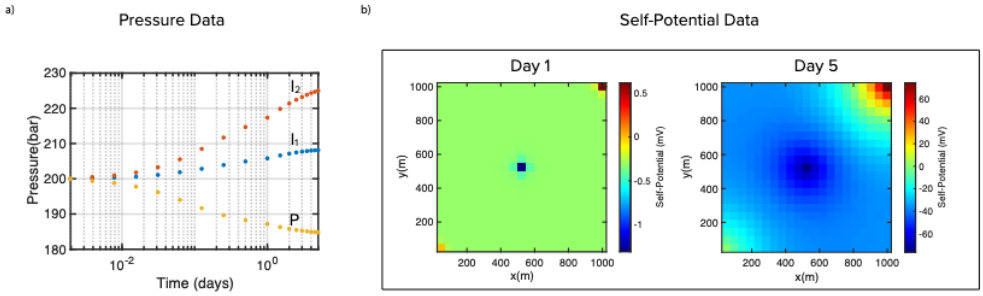

Figure 4-11: The transient pressure and self-potential data computed on the multiple fracture model. a) shows the pressure in bars at each well. Two wells are pumping at different constant rates. I_1 is injecting at 100 m^3/day and I_2 is injecting at 300 m^3/day. The production well is pumping at a 400 m^3/day rate. b) shows the recorded Self-Potential signals at two different periods.

the exact location of fractures cannot be located. Fractures are minor in size compared to the size of the grids. If the inversion is run on a finer grid, it would be possible to locate the fractures with an improved resolution. However, this would be at the cost of computation time and was not attempted in this experiment. The CPU time is approximately increasing exponentially as the number of grids increases.

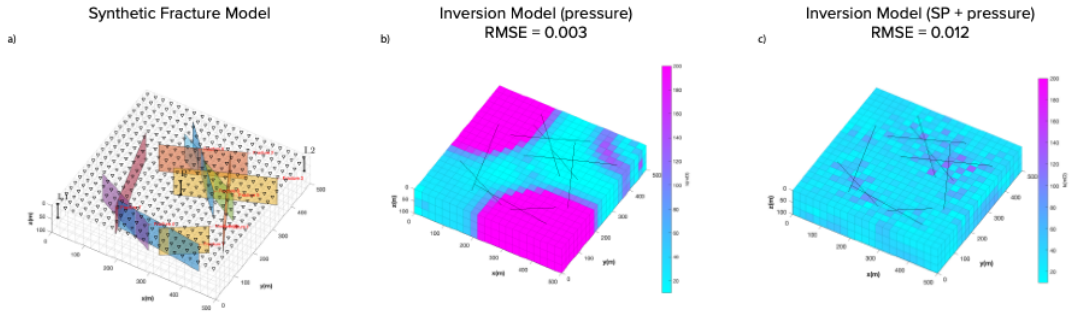

Figure 4-12: (a) A 3D synthetic reservoir shows multiple fractures. (b) Permeability structure resulted from the inversion using only the pressure data at well locations. (c) Final permeability structure utilizing pressure and SP data in a joint inversion. The Actual fracture locations are highlighted with black lines on the inverted images.

4.5 Discussion

The experiments demonstrated the inversion for the fluid flow parameters from a single and two data sets. The derivation of the forward modeling in section 2 shows the direct sensitivity between the fluid flow and the SP method. Therefore, the joint inversion of the fluid flow data with the SP is particularly applicable to characterize the fluid flow parameters. The inversion for the permeability using only pressure data from a limited number of wellbores could not resolve the fracture locations or the permeability structure of the reservoirs. The number and the distribution of wellbores are usually low and sparse. Therefore, inverting pressure data for the fluid-flow parameter lacks resolution, especially further away from the wellbore. With the SP data, in addition to the borehole pressure data, the imaging of the permeability structure is improved (Figure 4-6). The detection of high permeabilities at the fractures was demonstrated in the experiment. The self-potential method has illustrated its effectiveness in detecting fractures and imaging permeability anomalies. While our method is somewhat targeted to solve a similar problem, previously studied by Ahmed et al. (2016), the results shown in this chapter are extended to include fracture detection. They extensively studied the effect of hydraulic stimulation by injecting fluids at different rates on joint inversion. With insufficient stimulation, they could not improve the joint inversion results. On the contrary, the result in this article has improved significantly with minimum stimulation, which was in the form of a constant injection rate. The inversion algorithm performed its best to reduce the data residuals in all cases regardless of the data provided. The RMSE between the observed and modeled data is very small in all presented cases, even though the inverted images are different. Because the true synthetic models are already known, it was easy to compare the results of the two inverted images and decide which one was more accurate. In reality, this comparison will not be possible because we do not know the true earth

model. Therefore, a parameter resolution analysis is applied to estimate the accuracy of the inverted parameters (Wiggins, 1972). The computation of the resolution matrix R (eq4.15) can estimate how resolved the inverted parameter is compared to the true one.

$$R = G^{-g}G \tag{4.15}$$

where G^{-g} is the generalized inverse solution (Menke, 2010). R has a dimension $m \times m$, where m is the number of parameters. Ideally, the R would be an identity matrix for the unique results. If the resolution matrix shows a spread of values around the diagonal, it indicates inaccuracy and poor resolution (Figure 4-13).

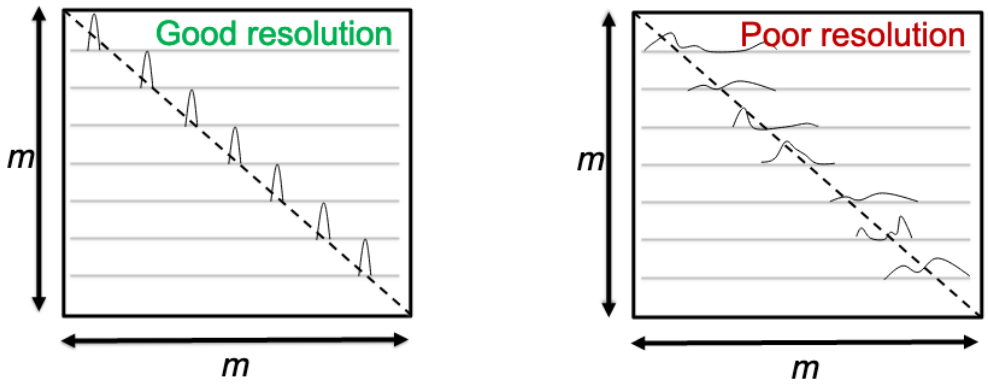

Figure 4-13: A schematic diagram showing two resolution matrices; a good resolution matrix on the right and a poor resolution matrix on the left. The good matrix has most of its values at the diagonal, while the poor matrix has more spread-out values.

The resolution matrices for the experiments are presented in Figure 4-14. These matrices are slightly different than what we have seen before in Chapter 2. To describe the main elements in the matrices, refer to Figure 4-15. Each row in the resolution matrix shows how each parameter is a weighted average of the other neighboring parameters. By examining the resolution matrix, we see nine boxes. The three boxes in the diagonal represent the three layers of the model. The off-diagonal boxes represent the different layers too. The reason that this resolution matrix has this shape is related

to the convention used in this study to order the grids. If we consider the middle row (dashed black box in Figure 4-15), which represents the parameter at the center of the model at Layer 2, it has weights from the neighboring parameters in Layers 1 and 3 in addition to Layer 2.

Examining the resolution matrix for the one-fracture experiments in Figure 4-14 shows a different batched concentration of values. The concentration of values is found near the location where the data is collected, especially at later 1 and 2. Consequently, the parameters close to the observed data location have higher resolution than the rest. The overall quality of the resolution matrix is computed by the Backus-Gilbert spread function (Backus and Gilbert, 1967, 1968). The spread (eq4.16), is an L_2 norm between the R and the identity matrix; I.

$$spread\,(R) = \sum_{i=1}^{m}\sum_{j=1}^{m}\left(R_{ij} - I_{ij}\right)^2 \tag{4.16}$$

In Figure 4-14, in addition to the image of the resolution matrices, we also write the spread on each one. The spread was lower in both cases when the two data sets were used for the inversion.

Overall, the described method is computationally very expensive as the big matrix G in eq4.12 has to be computed at each iteration. This is one limitation of the joint inversion problem, especially when the number of parameters increases. The G matrix will become bigger and bigger as we decide to improve the modeling resolution. Nevertheless, it is still faster than solving the fluid flow and seismic joint inversion problem because of the direct relationship between the fluid parameters and the SP. The problem with using seismic and flow data to model fluid properties is due to the different relationships coupling the seismic to the fluid-flow parameters. A sequential approach might be better used than a joint approach. The SP, on the contrary, is directly related to the flow properties. Another limitation that could complicate the inversion of the SP data is the complexity of the self-potential signals. It was assumed

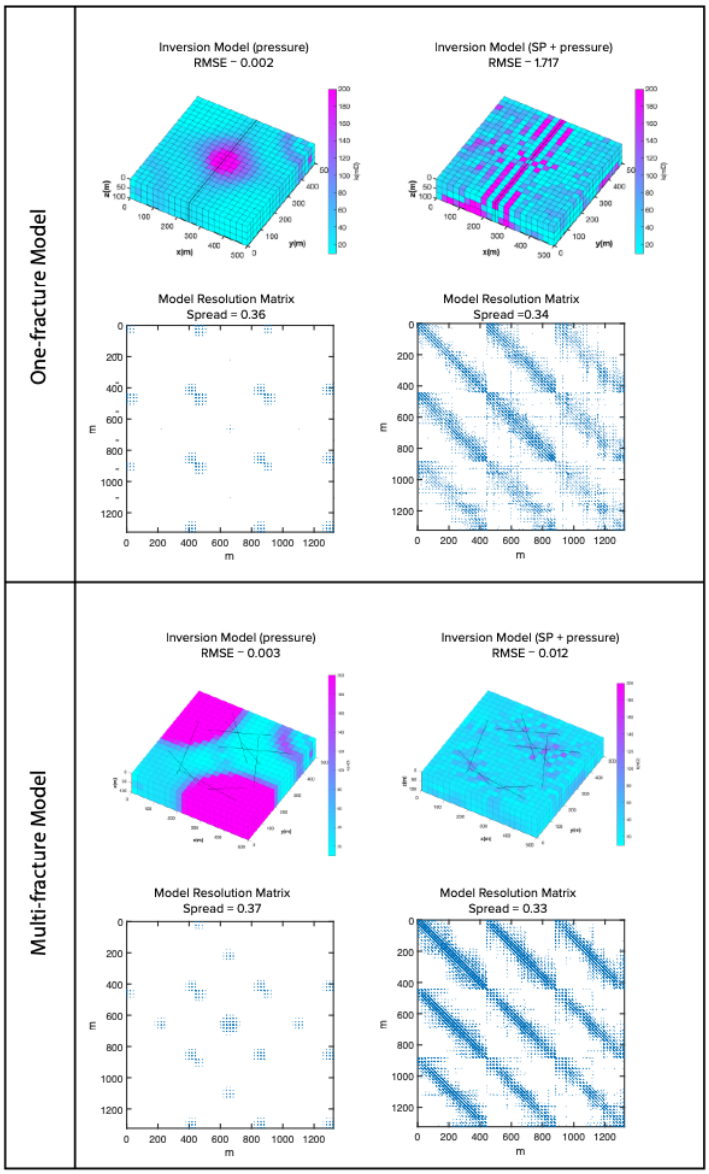

Figure 4-14: Summary of the inversions for experiment 2. The top figures represent the one-fracture case, and the bottom figures are the multi-fracture case. In addition to presenting the final inverted images, resolution matrix R is added below each inverted image. The calculated spread indicates how accurate the inversion is. The lower the spread-out the better-resolved the image.

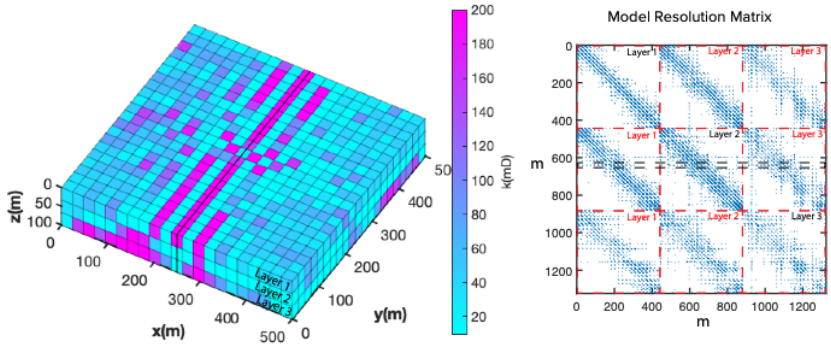

Figure 4-15: The inversion image of the one fracture model with the resolution matrix. The resolution matrix shows that each row, representing one parameter, is a weighted average of parameters from the three different layers. The black dashed box shows spread-out values around the parameter located at the center of the model.

throughout this study that the self-potential signals are only related to the fluid flow, disregarding the other forces explained in the equation 4.2. Including the other forces in the inversion process will require solving the problem for more parameters. This is still possible with the joint inversion scheme, but it is outside the scope of this thesis. For instance, including the other forces might be necessary for modeling SP in geothermal systems where thermoelectric is in effect. The presented method does not search for the location of the fractures, nor does it invert for their locations. It, however, aids in interpreting the inverted permeability images for the fracture locations. In the multi-fracture experiment, it was impossible to distinguish between the high and low permeability fractures (Figure 4-12). It appears that only high permeability fractures were detected. The multiple-fracture experiment was done again, with all fractures having similar parameters. It resulted in better illumination of the fracture regions (Figure 4-16). The method is susceptible to the permeability of the fractures. The inversion with only the fluid flow data proves that fracture detection is almost impossible with the available three wellbores in all cases. On the other hand, the added SP data to the existing pressure data improved the confidence in fracture detection.

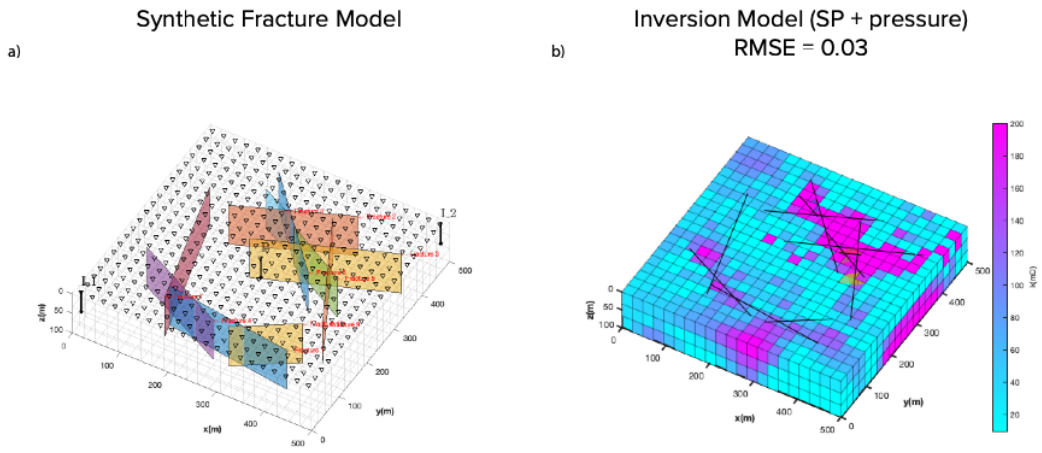

Figure 4-16: Results of running experiment 3.2.2 with all fractures having similar high permeability values. a) the 3D model with multiple fractures. b) the permeability image results from joint inverting of the SP and pressure data.

Chapter 5

Joint Fluid-flow Seismic Inversion Using Cross Gradient Method

5.1 Abstract

Optimizing a hydrocarbon reservoir's performance depends on having an accurate characteristic map of the reservoir. This map's construction requires a reasonably precise match to the historical reservoir performance, such as pressure and production rate. However, matching the historical data is costly and will lead to a non-unique reservoir model. This chapter developed an algorithm that jointly inverts for the reservoir permeability and seismic velocity using a cross-gradient method. The results are a more accurate reservoir model and an improved seismic image.

5.2 Introduction

A hydrocarbon-producing reservoir often goes through multiple stages as part of its life cycle, spanning from exploration to production. In early stages, geoscientists depend on field surveys such as seismic to explore hydrocarbon prospects. Interpretation

of the prospects could include the structure of the reservoir in addition to the fluid flow parameters (e.g., porosity & permeability). The flow parameters are an indication of the producing quality of a reservoir. If a reservoir is successfully explored and hydrocarbons are discovered, then the reservoir will be evaluated. At the evaluation stage, multiple boreholes are drilled to reduce uncertainty related to the reservoir geometry and characteristics. The drilled boreholes can also provide additional data from cores or logs, helping to describe the reservoir's complexity. Next, engineers and simulators collaborate to develop a field mainly using economically viable solutions such as the number of drilled production wells and their location. A preliminary reservoir model, including the rock properties, is built using seismic and borehole data. This model can aid in running many what-if scenarios to adopt the most appropriate strategy for field development. Finally, the field will undergo the production stage, which might be divided into different stages to ensure high hydrocarbon recovery.

During the production stage, the reservoir simulator can use the previously built reservoir model to match the reservoir's historical performance in the forward modeling process. In forward modeling, a simulator attempts to simulate the fluid flow behavior in a reservoir. The modeling process is repeated multiple times to update the reservoir model until a good simulation from an updated reservoir model matches the historically recorded data (Baker et al., 2006; Oliver & Chen, 2011).

There are some limitations to the sequential approach to history matching. There is often minimum interaction between the different disciplines during the stage of history matching. The seismic imaging used to obtain a preliminary model cannot resolve the reservoir down to the scale of the fluid-flow parameter, making the interpretation of the flow parameters difficult. Consequently, the model will not accurately capture the fluid behavior in the porous media. Not only that but also, the image acquired from the seismic data requires a stable and accurate algorithm with a relatively close initial guess to the actual model.

Another limitation that was discussed by (Kowalsky et al., 2006) is related to the

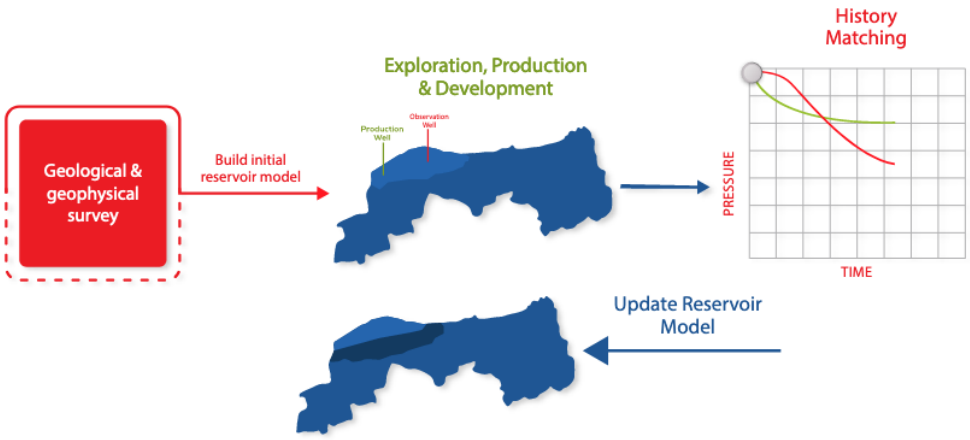

Figure 5-1: An overview diagram showing the stages leading to history matching

anisotropic nature of the seismic waves. The azimuthal variation in the traveling wave might indicate different relative velocities for the same region and, therefore, can be interpreted as different permeabilities.

The costly reservoir history matching process will also lead to a non-unique solution (Tavassoli Z. et al., 2004). The inverse problem is ill-posed and challenging to find a unique solution. Many reservoir models could result in an excellent match to the historical data. Additionally, in the sequential approach, the models obtained from the history match are never used to constrain or improve seismic imaging. There is a clear separation between the process done by geoscientists and engineers (Figure 5-1).

Therefore, in this chapter, we adopt a simultaneous inversion method instead of characterizing reservoirs with a conventional sequential process. Two different data sets, surface seismic and down-hole pressure data, are used in a simultaneous joint inversion algorithm to reduce the non-uniqueness and improve the resolution of the reservoir image. The objective of the joint inversion scheme is to minimize multiple data misfit functions while yielding a reasonable estimation of the model of interest.

In geophysics, there exist two techniques for joint inversion. One method depends on the direct relationship between two or three parameters, for example, seismic velocity and conductivity (Nielsen and Jacobsen,2000; Moorkamp et al., 2011). The second method, which we are adopting in this chapter, focuses on finding structural similarities as described by (Haber and Oldenburg, 1997; Gallardo and Meju 2004). Gallardo & Meju in (2007) used the cross-gradient to invert magnetotelluric and seismic data jointly. (Gao & Haijiang, 2018) applied the method to cross-borehole seismic and DC resistivity tomography. We will be testing a similar technique on synthetic seismic and pressure data to examine the effectiveness and accuracy of the method in detecting and imaging permeability structures.

There have been many studies done to understand the relationship between permeability and compressional seismic wave velocities. Kassab and Weller (2015) concluded that there is no clear relationship between the two parameters. Even establishing an empirical relationship has not been successful. However, some studies attempted to correlate the permeability to seismic velocity from various core data (e.g., Prasad 2003; Kitamura et al. 2010, Hamada H & Joseph, 2020). While these attempts illustrated a correlation, they are not applicable/relevant to all types of reservoirs. They only fit certain rock types, such as sandstones, under specific conditions (eq, porosity, saturation, mechanical properties of rocks). Therefore, we extended the cross-gradient coupling to help resolve the reservoir permeability structure by enforcing a structure similarity constraint rather than depending on a coupled relationship.

5.3 Method

The method is built on establishing accurate and stable forward and inverse modeling algorithms for the fluid flow in porous media and for seismic imaging problems. The fluid flow forward problem simulates the down-hole transient pressure in porous media. We discretize the reservoir into regular grids and solve the fluid transport

equation using a finite volume scheme (Peaceman, 2000).

Given the model parameters (porosity and permeability), we can formulate the fluid flow simulation by the non-linear system (5.1):

$$d_p = F(k) \tag{5.1}$$

Where d_p is a vector of the transient pressure data, F is the forward operator that solves the finite volume discretization with prescribed initial and boundary conditions, and k is a vector of the model parameters.

On the other hand, the seismic forward problem solves the acoustic wave equation using a second-order accurate finite difference approximation and can be simplified by the notation(5.2).

$$d_s = G(v) \tag{5.2}$$

Where d_s is a vector of the seismic signals, G is the forward operator, and v is a vector of the model parameters.

In addition to these two equations, we will have to estimate the cross-gradient structure constraint (eq5.3), which can be written as follows for a two-dimensional structure:

$$t(x,z) = \nabla k(x,z) \times \nabla v(x,z) \tag{5.3}$$

By computing $t(x,z)$, a 2-Dimensional image with cross-gradient values is estimated. If these cross-gradient values are small, they indicate that the two models are structurally similar. Even if there is a difference in the values of the parameters in the two models, the cross-gradient will result in a measure of the similarity of the structures of the two models. Figure 5-2 illustrates an example of the cross-gradient using two different models, one case where the structure is similar and the other where the structure is different

The adapted scheme in this study aims to minimize three objective functions; the pressure residuals, the seismic residuals, and the cross-gradient objective function. The objective functions can be written in permeability (k) and compressional wave velocity (v).

$$\alpha_f(k) = \|W_k(F(k) - d_p)\| \tag{5.4}$$

$$\alpha_s(v) = \|W_v(G(v) - d_s)\| \tag{5.5}$$

$$\alpha_t(k, v) = \|t(k, v)\| \tag{5.6}$$

Each minimization equation depends on separate algorithms to solve. The term $\|W_k(F(k) - d_p)\|$ is the L_2 norm between the modeled fluid flow and the observed pressure data weighted by W_k. The second objective function, $\|W_s(G(v) - d_s)\|$, is the L_2 norm between the modeled and the observed seismic signals weighted by W_s. The weight terms; W_k and W_s are diagonal matrices with the reciprocal of the data uncertainty populating the diagonal.

The last objective function $\alpha_t(k, v) = \|t(k, v)\|$ should disappear if the spatial gradient of the permeability and velocity models point in the same direction. If there is a difference in the structure of the two models, the values of the objective function will deviate from zero. To illustrate the computation of the cross-gradient, consider the different models presented in Figure 5-2. The cross-gradient of two similar structures should be zero even if the two structures vary in magnitude (Figure 5-2.a). The difference in the structure will result in variation in the cross-gradient values (Figure 5-2.b).

The joint inversion method in this study does not combine the different data (seismic and fluid flow) in a coupled system. Instead, we solve for each parameter k and

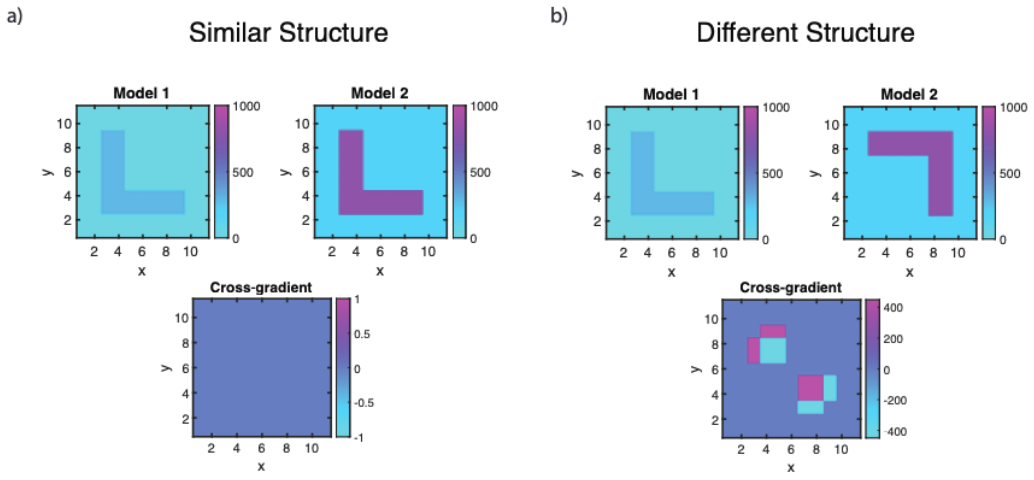

Figure 5-2: The cross-gradient function on different models. a) Models with similar structures but different magnitudes. b) Models with different structures and magnitudes.

v separately and constrain the results with the cross-gradient method in a separate step. Because of the non-linearity of the inverse problem, we solve for the perturbed parameters (δm & δv) in an iterative approach.

In Figure 5-3, we present an overview of the joint inversion algorithm. The joint inversion algorithm starts with separate inversions for each data set. We perform the full waveform inversion (FWI) for the seismic data to obtain an inverted velocity model using a pre-existing open-source python code (PySIT: Seismic imaging toolbox for Python (Hewett & Demanet, 2017)), and in parallel, we invert for a permeability model from the pressure data using a least squares method. The separate inversions are done iteratively until convergence to a solution is reached, in which the resolved models represent permeability (k) and velocity (v). We numerically estimate the cross-gradient sensitivity between the two models by slightly varying the structure and recalculating the cross-gradient. Finally, a coupled system of equations is formulated to solve for the parameter perturbation simultaneously while targeting to minimize the cross-gradient objective function. The main calculation happens in the joint inversion step, where

three systems of equations are solved (eq 4).

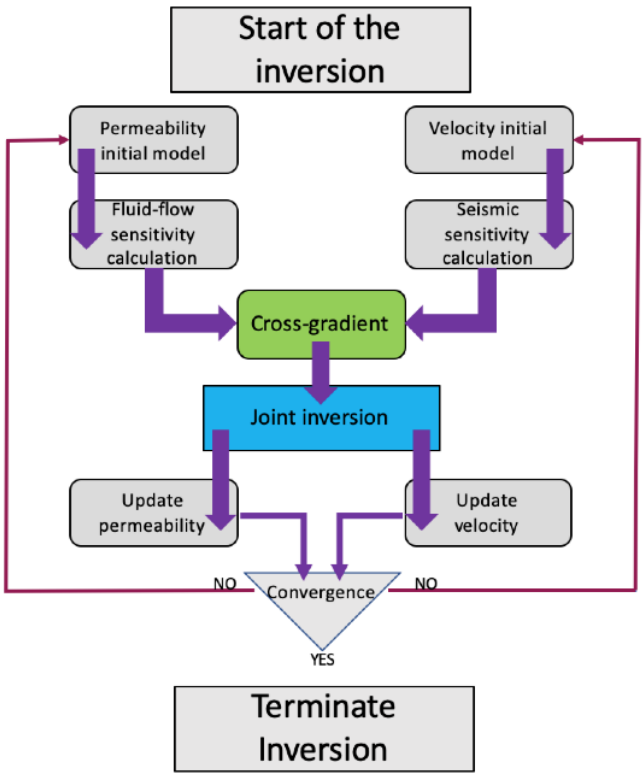

Figure 5-3: Overview flow-chart of the proposed joint inversion approach with the cross-gradient constraint.

$$\begin{bmatrix} w_f J & 0 \\ 0 & w_s I \\ w_t T_k & w_t T_v \end{bmatrix} \begin{bmatrix} \delta k \\ \delta v \end{bmatrix} = \begin{bmatrix} w_f \delta d_f \\ w_s \delta v_0 \\ -w_t t \end{bmatrix} \tag{5.7}$$

where J is the sensitivity matrix for the fluid flow populated by the partial derivatives of the pressure data with respect to the permeabilities, I is an identity matrix. $T_k = \frac{\partial t_{x,y}}{\partial k}$ and $T_v = \frac{\partial t_{x,y}}{\partial v}$ are the sensitivity matrices of the cross-gradient function with respect to the fluid-flow and the seismic parameters, respectively.

Unlike the formulation proposed by (Gao & Zhang, 2018), the formulation in this study takes advantage of the sensitivity matrix J, computed in the previous step, to improve estimations of the perturb permeability. In their formulation, the coupling system updates the parameters based on the values obtained from the previous iteration without considering the misfit in the data.

The different terms in the coupled system are weighted differently by the w_f, w_s, and w_t. The weights are chosen by testing various values and finding the one that minimizes the data misfits and the cross-gradient functions. The ranges of the weights might differ based on different scenarios, where the amount of data or the magnitude of the parameters are different.

The coupled system of the equation can be expressed by

$$A\delta m = \delta d \tag{5.8}$$

where A is a large matrix consisting of the sensitivity and identity matrices.

$$A = \begin{bmatrix} w_f J & 0 \\ 0 & w_s I \\ w_t T_k & w_t T_v \end{bmatrix} \tag{5.9}$$

δm is a vector containing (δk & δv).

$$\delta m = \begin{bmatrix} \delta k \\ \delta v \end{bmatrix} \tag{5.10}$$

And,

$$\delta d = \begin{bmatrix} w_f \delta d_f \\ w_s \delta v_0 \\ -w_t t \end{bmatrix} \tag{5.11}$$

In each iteration, the perturbation of the parameters is estimated using an SVD (Hansen, 1999) factorization of matrix A in a regularized damped least square method (Tikhonov & Vasiliy, 1977) as follows (Ren & Kalscheuer, 2020):

$$A = USV^T \tag{5.12}$$

$$\delta m = (VS^2V^T - \epsilon I)^{-1} VS^2U\delta d \tag{5.13}$$

Where V is a matrix having the right eigenvectors, U is a matrix with the left eigenvectors, S is the diagonal matrix with eigenvalues populating the diagonal in decreasing order. ϵ is a damping factor, and I is an identity matrix. ϵ controls the rate of convergence towards a solution. The Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1964) method alters the ϵ at each iteration.

5.4 Synthetic Results

We are testing our proposed method on numerical examples. The synthetic data used for the inversion results from forward modeling of the fluid flow in porous media and modeling the seismic reflection. Therefore, we need to set up two models; one model for permeability and another for velocity. Each one of the models has property

anomalies that are similar in terms of location and structure but different in value. The synthetic reservoir model has a constant thickness and is set up in 2-dimensional space. Also, it has three boreholes; two injecting wells and one producing at the center. The three wells act as the source/sink term for the fluid flow simulation and serve as the data point where the pressure is recorded. We simulated the fluid flow on a model of size 1100x1100 m by discretizing the flow equation on 11x11 grids. The reservoir has a constant thickness. Both initial and boundary pressure conditions are set to constant. The duration of the simulation is run for 20 days. Only the pressure data from the three wells will be used for the inversion. The inversion is performed using the least squares method. We assume that all the other fluid parameters are known except for permeability. The simulation result of fluid flow is presented in Figure 5-5. The reservoir in Figure 5-5 shows the permeability structure on a log scale. The pressure versus time curves are plotted, showing two curves with increasing pressure with time as they represent injecting wells. The two injecting wells do not show similar pressure curves due to the permeability anomaly's variability. The anti-symmetry in the permeability anomaly causes the pressure curves to behave differently on both sides of the reservoir. The other curve is a producing well showing a decreasing pressure with time.

For the synthetic seismic data, we generated a 3-dimensional velocity model. Within the model, we embedded a 2-Dimensional model, representing the fluid-flow model. A sketch diagram illustrating the two models is shown in Figure 5-4. We simulated the seismic acoustic signals in the model and recorded the signals on the surface. A recording station is placed at each grid on the surface, with a total of 121 stations. A Ricker wavelet with a peak frequency of 10 Hz is the source in the center of the model (Figure 5-6). The anomalies are placed in the exact locations where the permeability anomalies exist. As discussed earlier, there is no generalized mathematical equation correlating permeability to seismic velocity. Therefore, we randomly chose one of the linear relationships from (Prasad, 2003) to relate the permeability to velocity values.

Prasad (2003) has shown that many permeability-velocity relationships can be linearly correlated under certain conditions indicated by the flow zone indicator (FZI). For $FZI = 1$, the relationship is as follows:

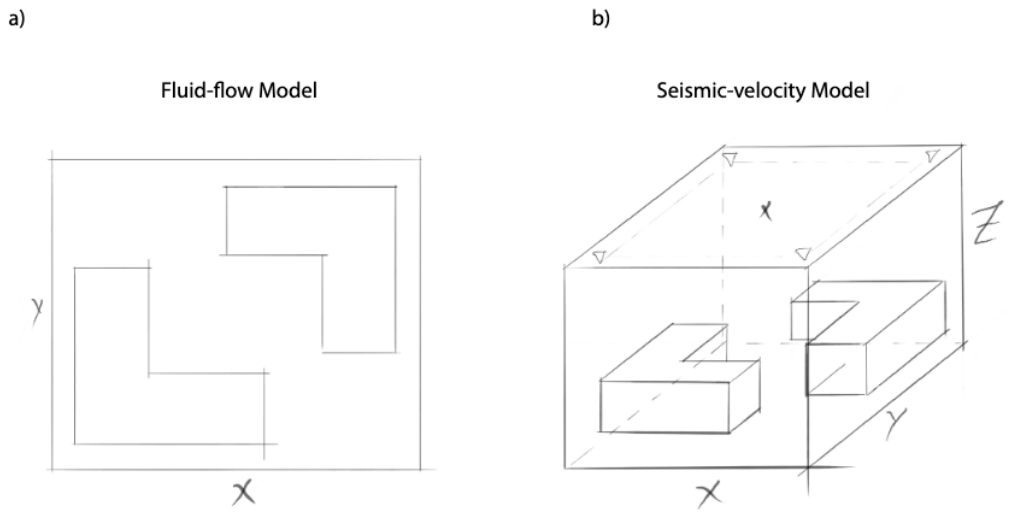

Figure 5-4: A sketch illustrating the dimension and the physical space of the models. a) The permeability model is built in a 2-Dimensional space. b) The velocity model is built in 3-Dimensional space. The x indicates the location of the seismic shot, and the small triangles are the location of the receivers.

$$log_{10}(k) = 4.05 - 0.75v \tag{5.14}$$

The seismic data is computed on a model of size 1100x1100x1100m on 11x11x11 grids. The forward and inversion of the seismic signals are performed using the PySIT python code. Figure 5-6 displays the shot gathers for all receivers. The length of the seismic record is three seconds, where the sampling rate is set at six milliseconds.

After the pressure and seismic data were simulated using the "true" models, we performed the proposed method Figure 5-3 to compare its result with an inversion done on single-type data. Figure 5-7 shows the summary of the joint inversion scheme.

Figure 5-7.b & Figure 5-7.e present the results of inverting pressure and seismic data for permeability and velocity individually. We did not apply the cross-gradient

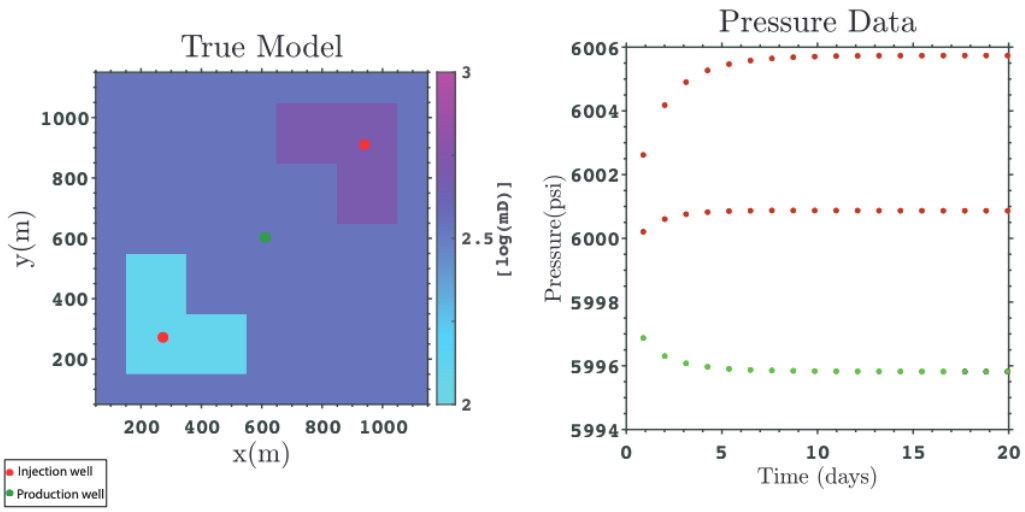

Figure 5-5: The true permeability model with the location of the injecting and producing wells colored differently. The right figure shows the pressure curves with time at each well (colored accordingly).

method in this individual inversion to constrain the structure. Clearly, the inverted image miss interprets the anomalies' value and shape. Both separate inversions have done a decent job of detecting the different anomalies.

In Figure 5-7.c & Figure 5-7.f, we present the result of using the cross-gradient joint inversion method. There is a definite improvement in the anomalies' final images for both models. In this particular case, we chose the following values for weights: $w_f = 0.1, w_s = 0.01, w_t = 200$ after examining various values for the weights.

The image root means square error (iRMSE) and data root means square error (RMSE) can assist in estimating the overall improvement of the inversion results. The true model used to create the synthetic data is known; thus, computing the error between the inverted and true images is possible. The images resulting from the individual inversions show slightly higher iRMSE than what was constructed by joint inversion, indicating that the joint inversion resulted in models being closer to the true one. The joint inversion contributed to an almost 7% reduction in the iRMSE for the seismic velocity model and an approximately 1% reduction for the permeability model. Even though there is improvement in the overall parameter estimation indicated by

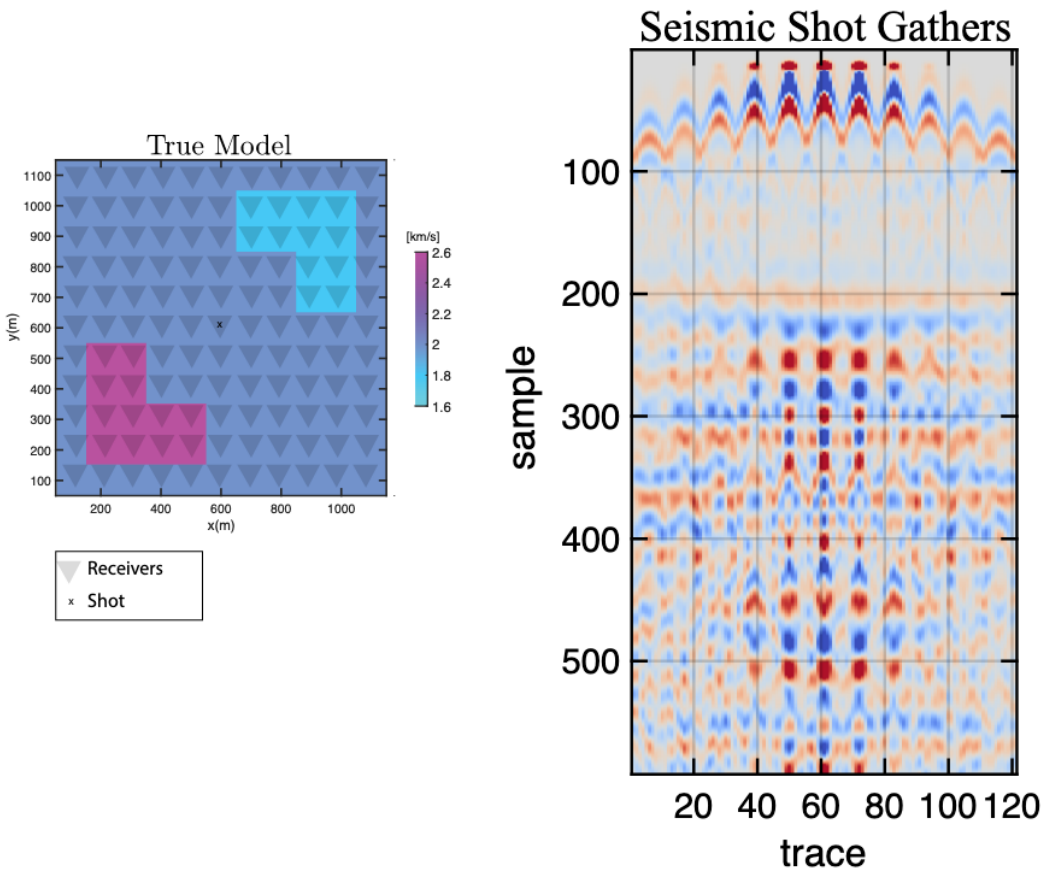

Figure 5-6: The true seismic velocity model. The location of the shot is marked by (X). The triangles are the locations of the receivers. b) The shot gathers, displaying all the traces.

iRMSE, the RMSE data did show a slight increase.

$$iRMSE = \sqrt{\frac{\sum(m_{true} - m_{inversion})^2}{M}} \tag{5.15}$$

$$RMSE = \sqrt{\frac{\sum(data_{true} - data_{model})^2}{N}} \tag{5.16}$$

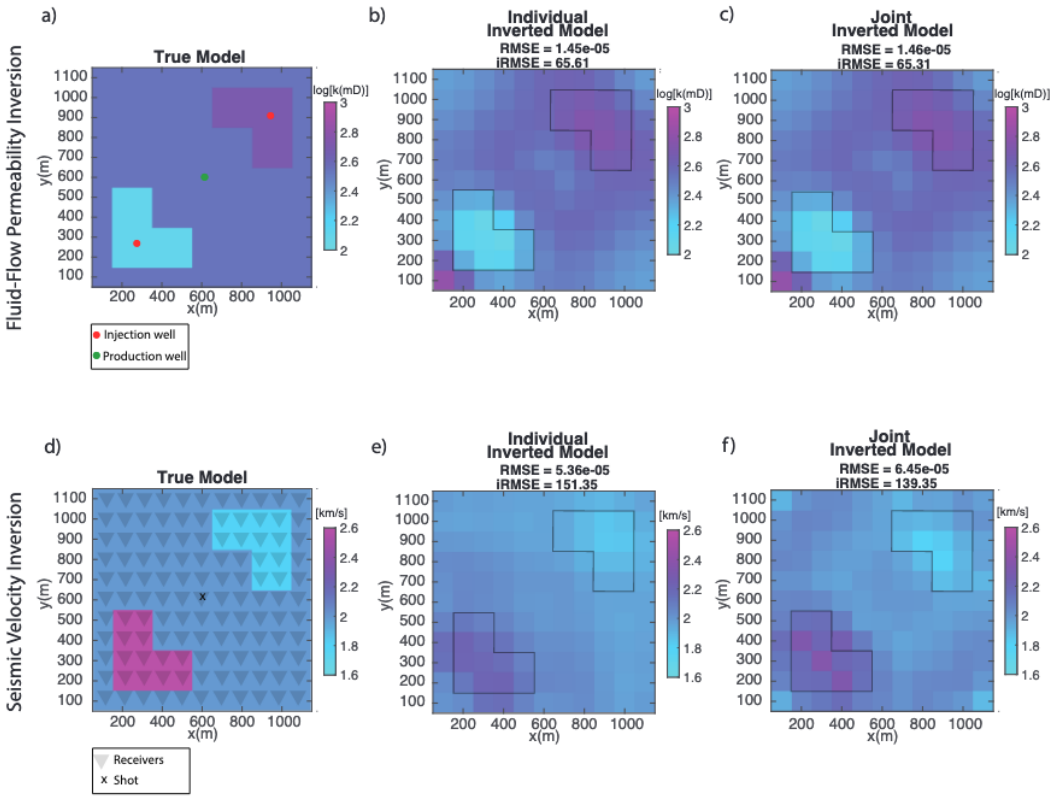

Figure 5-7: Comparing the individual inversion and the joint inversion method for fluid flow parameter (permeability) and the seismic parameter (velocity). a) true model in (x, y) plane with two permeability anomalies in milliDarcy. Pressure data are collected from the two injecting wells and one producing well. b) inversion for the permeability using only data from the three wells. c) joint inversion result for the permeability using the cross-gradient method. d) (x, y) section of the true seismic model used for the synthetic data. e) individual inversion for the seismic velocity. f) joint inversion for the seismic velocity.

5.5 Discussion

The cross-gradient joint inversion method aims to minimize the structure difference between the models (seismic velocity and permeability) and the data misfit simultaneously. The method works well if there is a similarity in structure, as shown in the presented example. In real scenarios, this method would work to detect large-scale features such as faults and channels because the seismic inversion does not resolve the velocity to the scale of fluid flow parameters. Therefore, the cross-gradient method will

not be effective in characterizing the small-scale heterogeneity of a reservoir. We have demonstrated that the cross-gradient joint inversion method improves the results for the case where two data sets are utilized. The improvement of the imaging is illustrated by examining the (iRMSE).

Although this method is computationally costly, it shows that using a cross-gradient can improve imaging of the seismic velocity and, in some cases, improve the permeability estimation. The cost in computation is attributed to two main factors; 1) the computation of the big matrix (A) in eq 5.8 and its singular value decomposition 2) the estimation of the weighting values w_f, w_s and w_t which require multiple runs to choose the most appropriate values. However, the general method is very versatile because the initial guess is a homogeneous model and can still resolve a reasonable solution. The significance of the method is that it can be more useful in cases where there is more data coverage for one of the surveying methods. In Figure 5-8, we examined a test case with more observed pressure data from additional wells. The joint inversion result shows a better-resolved structure, especially for the seismic velocity. The increase in the number of wells enabled the algorithm to delineate the structure of the anomalies better. There is almost a 16% reduction in the seismic velocity iRMSE with more wells compared to 7% when there were only three wells. The permeability iRMSE also shows a significant reduction of 26% compared to the 1% achieved with fewer data.

In the inversion algorithm, the relationship that was assumed to generate true models was not taken into consideration. Even though we know the relationship (eq 5.14), we wanted to test the inversion for the two models regardless of such relationships. As discussed earlier, there is no clear relationship linking permeability to velocity. Using the cross-gradient inversion of the two models, we can estimate a relationship linking the two parameters. The method presented in this chapter can assist in finding such empirical equations. We used a non-parametric Gaussian process regression (Williams & Rasmussen, 1995) to predict a relationship between the two parameters and estimate permeability for a wide range of velocities. The cross plot of the velocity and

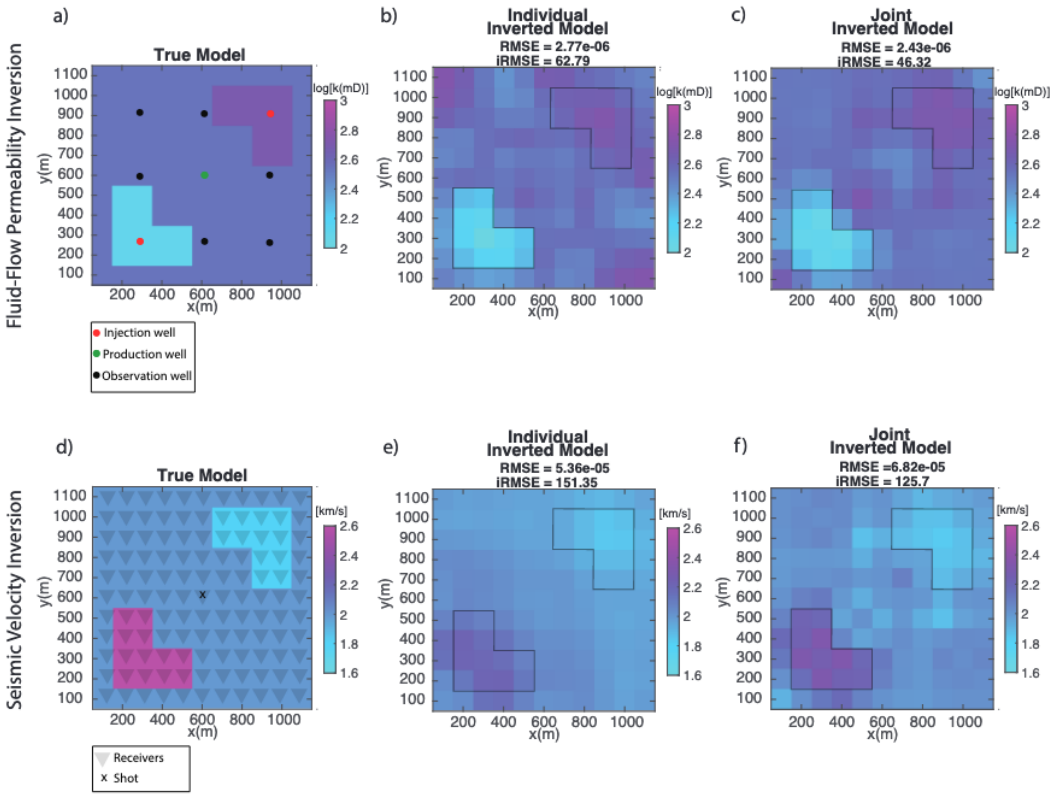

Figure 5-8: Comparing the individual and joint inversion methods for fluid flow parameter (permeability) and the seismic parameter (velocity). a) true model in (x, y) plane with two permeability anomalies in milliDarcy. Pressure data are collected from a total of nine wells. b) inversion for the permeability using only data from the wells. c) joint inversion result for the permeability using the cross-gradient method. d) (x, y) section of the true seismic model used for the synthetic data. e) individual inversion for the seismic velocity. f) joint inversion for the seismic velocity.

permeability shows an almost linear relationship, which is expected considering that the relationship used to create the models is linear. Let's compare the fit of the estimated parameters from the joint inversion to the fit from the individual inversion. The joint inversion fit is closer to the assumed relationship Figure 5-9.

Therefore, we propose that inverting for seismic velocity and permeability through the cross-gradient method can be used to determine a more accurate relationship between the two parameters. While a relationship might be deduced from well logs by correlating the sonic logs to the core permeability, the joint inversion scheme can allow

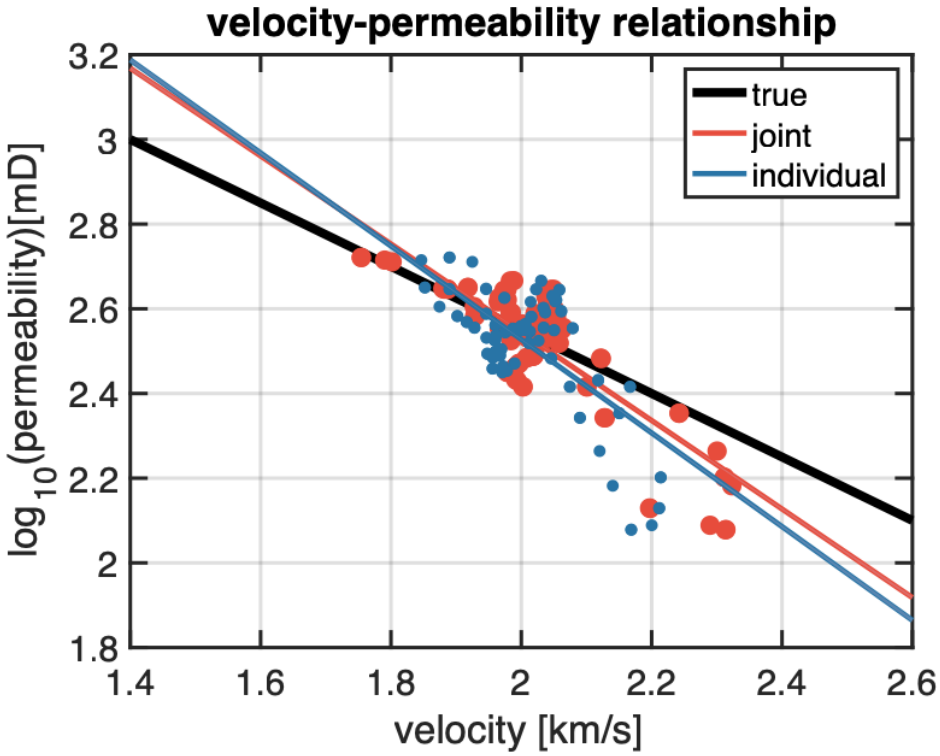

Figure 5-9: The velocity-permeability relationship from different inversion results. The black line is the true relationship from eq5.14 used to build true models. The blue line is the fit to the data from the individual inversion. The red line is the fit to the data from the joint inversion results.

us to infer a more general relationship of an entire field or a region.

Chapter 6

Conclusion

The fluid flow modeling in porous media is an essential tool for predicting the future performance of hydrocarbon reservoirs. Characterizing reservoirs' properties, such as porosity and permeability, is achieved with various geophysical and engineering methods. Geophysical surveys usually provide low-resolution characterizing of reservoirs but have broad spatial coverage. Engineering methods such as logs observed from wellbores can provide high-resolution information but are limited by the area around the wellbores. Reservoir-scale property characterization with high resolution can be achieved with history matching, where the parameters of the reservoirs are altered, so the simulated fluid flow data matches the observed data.

In the first and second chapters of the thesis, we discussed using downhole pressure data in inversion for the reservoir parameters. The use of gradient-based methods in inversion was an effective choice. Levenberg-Marquardt algorithm allowed combining the Newton's method with the steepest descent for a fast and stable convergence towards local or global minimum. The algorithm even provided better and more stable convergence when the problem was modified for a log-scaled solution. We have noticed the logarithmic scaling reduced the jaggedness convergence and made it more durable because it smoothens the misfit space. The non-log-scaled problem, on the

other hand, still converges to a solution but slowly and with significant fluctuation. A similar result was seen in (Weelock et al., 2015), where the inversion of log-scaled MT data outperforms the use of non-log-scaled data.

By the end of the second chapter, we demonstrated the problem associated with history matching. The solution to the inverse problem is highly non-unique and ill-posed. Even with the regularization of the objective function, where it added a smoothening constraint, the recovered permeability models are non-unique.

An added advantage to using the gradient-based inversion method is the ability to compute the generalized inverse matrix explicitly and consequently compute the resolution matrix (R). The model resolution matrix measures the closeness of the inverted preferred solution to the true model. Without the computation of the resolution matrix, inversion results would be incomplete. There are many ways to evaluate the quality of the R matrix. For no particular reason, we chose to use the spread (Backus and Gilbert, 1967, 1968), which computes the norm between the R matrix and an Identity matrix. Even though the R matrix can evaluate the quality of the inverted results, in some cases, it is used to compare different inversion results, as demonstrated in chapters; 2 & 3. Having the comparison as an objective, the various methods of evaluating the R matrix won't matter because the difference is relative. The study of the R matrix for different inversions provided insight into how non-unique the fluid flow inversion results. The first experiments were done to illustrate the non-uniqueness and show how the model space (number of models) and the data space (amount and quality of data) can result in different model resolutions of the same model. For example, in the case where the locations of four boreholes were randomly shuffled, inversion resulted in different resolved models. Thereby, the R matrix was used in later chapters as a lens, in which we can see the differences between the unknown true model and the best-inverted results.

Understanding the fact that non-uniqueness is an inherited problem in the inversion of the reservoir parameters makes simulating the reservoir with a large number of grids

\parameters counter-intuitive. Representing a reservoir with many parameters would result in a broader space of non-unique solutions. Therefore, in chapter three, we studied the reduction of the number of parameters to represent reservoirs with better accuracy and with less expensive computations. In some cases, the reduction in the number of parameters resulted in models that are as good as the complex model, if not better. Reducing the 133x133 true model into a model with a size of 50x50 has visually and quantitively (image-RMSE) represented the true reservoir better than the 80x80 model with an order of magnitude less in computation time. We demonstrated that there is a different method for reducing the complexity. Using a non-regular griding scheme demonstrated a slightly improved resolved model (less spread in the R matrix) over the regular grids of more parameters. Nevertheless, it is visually and quantitively not as good as the other higher models.

Not only does reducing the number of grids \parameters enable a better representation of the reservoir, but through survey design, we have also seen that the inversion can result in an improved resolved parameter. The two-stage approach for survey design allowed us to seek high-content information and high model resolution by examining various scenarios of wellbore placement. The method did not find a global optimum survey design, though; it was in the local minimum. Because the approach depends on an adaptive process, finding the global solution for all the wells at once was not possible.

We have not looked into pairing the two sections of the second chapter to examine the reduction of non-uniqueness. This could be the future direction towards which we move for possible improvement. A disadvantage to the adaptive method for this particular problem is related to the cost of acquiring new data from a new wellbore. In many cases, the choice to drill a new well is to optimize production and not for the sake of gathering new data. Alternatively, additional data from different geophysical surveys can be added to constrain the borehole data. Chapter four has results of complimenting the borehole pressure data with surface self-potential data.

The limited spatial coverage for the boreholes can result in a poor resolution for the reservoir parameters, especially away from the borehole locations. Combined with borehole pressure data, the SP data improved the detection of permeability structure not only for large anomalies but also for small features such as fractures. The sensitivity of the SP method to the fluid flow in porous media provided a unique way to detect fractures.

We implemented the resolution matrix analysis to differentiate the improvement in the result of the joint inversion. Presented different scenarios in the joint inversion study improved the final image of the permeability structure compared to inversion of only pressure data despite high data RMSE. As discussed, SP signals studied in this thesis are primarily the consequence of an electrokinetic mechanism. The other coupled forces described by Onsager resulted from the microscale mechanisms (such as thermoelectric and electrochemical effects) that may influence the signal but are not included in our study. This effect might not be important in analyzing macroscale features (such as fractures). Still, it may be useful for volcanic and geothermal settings where a hydrothermal system is in play.

Considering that the SP method can cover a large surface area with a fraction of the cost spent to acquire seismic data, it is suitable for real-case scenarios. Though that does not mean the use of seismic data is not useful. Seismic data is the most likely data to be available for most hydrocarbon fields. It just sounds reasonable to take advantage of the seismic data.

In chapter five of the thesis, we used the surface seismic data to structurally constrain the inversion in a joint inversion process. There is no specific relationship to map the permeability into seismic velocity and vice versa; therefore, applying a structure similarity for the inversion of the two data sets (pressure and seismic data) seemed plausible

We studied the use of a cross-gradient method in inversion for permeability and seismic velocity. In the joint inversion workflow, we simultaneously minimized three

objective functions related to the fluid-flow data residuals, seismic data residuals, and the cross-gradient function. The workflow of the joint inversion accounts for the sensitivity of the different systems with the proper weights and therefore avoids coupling the two systems. We have better mapped the structural anomalies by testing the method on synthetic data and comparing the result to the separated inversion. Throughout the study, we demonstrated that the strength of the algorithm is in cases when there is high data coverage of one of the data types. A crucial element that adds to the success of the method depends on the weights. We essentially looked for only one weighing value for each factor and kept them constant at all iterations. There can be an alternative way that we have not explored in this study, and that is examining different weights of the factors in each iteration. This will be obviously a time-consuming approach because, at each iteration, the different weight values are tested to find the optimum choice. At the end of this study, we showed that this method could be used to conclude the relationship between the velocity and the permeability of an entire field.

Even though this thesis has attempted to study the non-uniqueness in fluid flow through different approaches, combining the presented methods might result in a better solution. In future pursuits, we can look at combining the various aspects of each chapter into one problem. The joint inversion of SP data with pressure data can be combined with the joint seismic fluid inversion in one algorithm. Also, the survey design can be applied to improve the survey design for the SP and seismic data.

The real-world scenarios might pose more complicated cases than what was presented in this thesis. Most of the synthetic data used in our studies are noise-free and recorded with higher precision than what can be recorded in the fields. This complication from real data will worsen the problem of non-uniqueness and limit our inversion capabilities. Therefore, combining the different approaches is the best way to reduce the non-uniqueness.

Bibliography

Alkafeef, S., Gochin, R., and Smith, A. (2001). The effect of double layer overlap on measured streaming currents for toluene flowing through sandstone cores. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 195(1-3):77–80.

Backus, G. and Gilbert, F. (1968). The resolving power of gross earth data. *Geophysical Journal International*, 16(2):169–205.

Backus, G. E. and Gilbert, J. F. (1967). Numerical applications of a formalism for geophysical inverse problems. *Geophysical Journal International*, 13(1-3):247–276.

Baker, R. O., Chugh, S., Mcburney, C., and McKishnie, R. (2006). History matching standards; quality control and risk analysis for simulation. In *Canadian International Petroleum Conference*. OnePetro.

Barth, N. and Wunsch, C. (1990). Oceanographic experiment design by simulated annealing. *Journal of Physical Oceanography*, 20(9):1249–1263.

Bear, J. (1988). *Dynamics of fluids in porous media*. Courier Corporation.

Burns, D. R., Willis, M. E., Toksöz, M. N., and Vetri, L. (2007). Fracture properties from seismic scattering. *The Leading Edge*, 26(9):1186–1196.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology journal*, 13(1):206–222.

Carrera, J. and Neuman, S. P. (1986). Estimation of aquifer parameters under transient and steady state conditions: 1. maximum likelihood method incorporating prior information. *Water Resources Research*, 22(2):199–210.

Chapman, D. L. (1913). Li. a contribution to the theory of electrocapillarity. *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, 25(148):475–481.

Coles, D. A. and Morgan, F. D. (2009). A method of fast, sequential experimental design for linearized geophysical inverse problems. *Geophysical Journal International*, 178(1):145–158.

Colombo, D., Mantovani, M., Hallinan, S., and Virgilio, M. (2008). Sub-basalt depth imaging using simultaneous joint inversion of seismic and electromagnetic (mt) data:

a crb field study. In *SEG Technical Program Expanded Abstracts 2008*, pages 2674–2678. Society of Exploration Geophysicists.

Cover, T. M., Thomas, J. A., et al. (1991). Entropy, relative entropy and mutual information. *Elements of information theory*, 2(1):12–13.

Crisp, M., Jaksa, M., and Kuo, Y. (2020). Toward a generalized guideline to inform optimal site investigations for pile design. *Canadian Geotechnical Journal*, 57(8):1119–1129.

Cueto-Felgueroso, L. and Juanes, R. (2013). Forecasting long-term gas production from shale. *Proceedings of the National Academy of Sciences*, 110(49):19660–19661.

Davis, J. A., James, R. O., and Leckie, J. O. (1978). Surface ionization and complexation at the oxide/water interface: I. computation of electrical double layer properties in simple electrolytes. *Journal of colloid and interface science*, 63(3):480–499.

DesRoches, A. J., Butler, K. E., and MacQuarrie, K. T. (2018). Surface self-potential patterns related to transmissive fracture trends during a water injection test. *Geophysical Journal International*, 212(3):2047–2060.

Eppstein, M. J. and Dougherty, D. E. (1996). Simultaneous estimation of transmissivity values and zonation. *Water Resources Research*, 32(11):3321–3336.

Fagerlund, F. and Heinson, G. (2003). Detecting subsurface groundwater flow in fractured rock using self-potential (sp) methods. *Environmental Geology*, 43(7):782–794.

Gallardo, L. A. and Meju, M. A. (2004). Joint two-dimensional dc resistivity and seismic travel time inversion with cross-gradients constraints. *Journal of Geophysical Research: Solid Earth*, 109(B3).

Gallardo, L. A. and Meju, M. A. (2007). Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltime data for structural and lithological classification. *Geophysical Journal International*, 169(3):1261–1272.

Gao, J. and Zhang, H. (2018). An efficient sequential strategy for realizing cross-gradient joint inversion: method and its application to 2-d cross borehole seismic traveltime and dc resistivity tomography. *Geophysical Journal International*, 213(2):1044–1055.

Gelhar, L. W. and Axness, C. L. (1983). Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research*, 19(1):161–180.

Gómez-Hernánez, J. J., Sahuquillo, A., and Capilla, J. (1997). Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data—i. theory. *Journal of Hydrology*, 203(1-4):162–174.

Greenberg, M. D. (2015). *Applications of Green's functions in science and engineering*. Courier Dover Publications.

Haber, E. and Oldenburg, D. (1997). Joint inversion: a structural approach. *Inverse problems*, 13(1):63.

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52.

Hamada, G. and Joseph, V. (2020). Developed correlations between sound wave velocity and porosity, permeability and mechanical properties of sandstone core samples. *Petroleum Research*, 5(4):326–338.

Hansen, P. C. (1999). The l-curve and its use in the numerical treatment of inverse problems.

Hewett, R. J. and Demanet, L. (2017). Pysit: Seismic imaging toolbox for python. *Mass. Inst. Technol., Cambridge, MA, USA, Tech. Rep.*

Hu, H. and Zheng, Y. (2018). 3d seismic characterization of fractures in a dipping layer using the double-beam method. *Geophysics*, 83(2):V123–V134.

Hyndman, D. W., Harris, J. M., and Gorelick, S. M. (1994). Coupled seismic and tracer test inversion for aquifer property characterization. *Water Resources Research*, 30(7):1965–1977.

Irving, J. and Singha, K. (2010). Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities. *Water Resources Research*, 46(11).

Jardani, A., Dupont, J.-P., and Revil, A. (2006). Self-potential signals associated with preferential groundwater flow pathways in sinkholes. *Journal of Geophysical Research: Solid Earth*, 111(B9).

Jones, A. G. and Foster, J. H. (1986). An objective real-time data-adaptive technique for efficient model resolution improvement in magnetotelluric studies. *Geophysics*, 51(1):90–97.

Kang, P. K., Zheng, Y., Fang, X., Wojcik, R., McLaughlin, D., Brown, S., Fehler, M. C., Burns, D. R., and Juanes, R. (2016). Sequential approach to joint flow-seismic inversion for improved characterization of fractured media. *Water Resources Research*, 52(2):903–919.

Kassab, M. A. and Weller, A. (2015). Study on p-wave and s-wave velocity in dry and wet sandstones of tushka region, egypt. *Egyptian Journal of Petroleum*, 24(1):1–11.

Kijko, A. (1977). An algorithm for the optimum distribution of a regional seismic network—i. *pure and applied geophysics*, 115(4):999–1009.

Kitamura, K., Takahashi, M., Mizoguchi, K., Masuda, K., Ito, H., and Song, S.-R. (2010). Effects of pressure on pore characteristics and permeability of porous rocks as estimated from seismic wave velocities in cores from tcdp hole-a. *Geophysical Journal International*, 182(3):1148–1160.

Kowalsky, M. B., Chen, J., and Hubbard, S. S. (2006). Joint inversion of geophysical and hydrological data for improved subsurface characterization. *The Leading Edge*, 25(6):730–734.

Lacazette, A., Vermilye, J., Fereja, S., and Sicking, C. (2013). Ambient fracture imaging: A new passive seismic method. In *Unconventional Resources Technology Conference*, pages 2331–2340. Society of Exploration Geophysicists, American Association of Petroleum ....

Linde, N., Binley, A., Tryggvason, A., Pedersen, L. B., and Revil, A. (2006). Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data. *Water Resources Research*, 42(12).

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441.

Matthews, C. S., Russell, D. G., et al. (1967). *Pressure buildup and flow tests in wells*, volume 1. Henry L. Doherty Memorial Fund of AIME New York.

Maurer, H., Curtis, A., and Boerner, D. E. (2010). Recent advances in optimized geophysical survey design. *Geophysics*, 75(5):75A177–75A194.

McKenna, S. A. and Poeter, E. P. (1995). Field example of data fusion in site characterization. *Water Resources Research*, 31(12):3229–3240.

McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem. *Water Resources Research*, 32(5):1131–1161.

Menke, W. (2012). *Geophysical data analysis: discrete inverse theory: MATLAB edition*, volume 45. Academic press.

Montes, G., Bartolome, P., and Udias, A. L. (2001). The use of genetic algorithms in well placement optimization. In *SPE Latin American and Caribbean petroleum engineering conference*. OnePetro.

Moorkamp, M., Heincke, B., Jegen, M., Roberts, A. W., and Hobbs, R. W. (2011). A framework for 3-d joint inversion of mt, gravity and seismic refraction data. *Geophysical Journal International*, 184(1):477–493.

Morgan, F. D. (1989). Fundamentals of streaming potentials in geophysics: Laboratory methods. In *Detection of subsurface flow phenomena*, pages 133–144. Springer.

Morgan, F. D., Al Nasser, S., Jerry, R., and Verneuil, A. (2019). Investigations into groundwater flow towards a spring in the saphire area, soufriere, st lucia, west indies. In *SEG International Exposition and Annual Meeting*. OnePetro.

Nelson, R. W. (1960). In-place measurement of permeability in heterogeneous media: 1. theory of a proposed method. *Journal of Geophysical Research*, 65(6):1753–1758.

Neuman, S. P. (1973). Saturated-unsaturated seepage by finite elements. *Journal of the hydraulics division*, 99(12):2233–2250.

Nielsen, L. and Jacobsen, B. H. (2000). Integrated gravity and wide-angle seismic inversion for two-dimensional crustal modelling. *Geophysical Journal International*, 140(1):222–232.

Nogueira, P. D. B. and Schiozer, D. J. (2009). An efficient methodology of production strategy optimization based on genetic algorithms. In *Latin american and caribbean petroleum engineering conference*. OnePetro.

Oliver, D. S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221.

Onsager, L. (1931). Reciprocal relations in irreversible processes. i. *Physical review*, 37(4):405.

Peaceman, D. W. (2000). *Fundamentals of numerical reservoir simulation*. Elsevier.

Prasad, M. (2003). Velocity-permeability relations within hydraulic units. *Geophysics*, 68(1):108–117.

Ren, Z. and Kalscheuer, T. (2020). Uncertainty and resolution analysis of 2d and 3d inversion models computed from geophysical electromagnetic data. *Surveys in Geophysics*, 41(1):47–112.

Revil, A. and Mahardika, H. (2013). Coupled hydromechanical and electromagnetic disturbances in unsaturated porous materials. *Water resources research*, 49(2):744–766.

Revil, A., Naudet, V., Nouzaret, J., and Pessel, M. (2003). Principles of electrography applied to self-potential electrokinetic sources and hydrogeological applications. *Water Resources Research*, 39(5).

Riley, K., Hobson, M., and Bence, S. (1998). Mathematical methods for physics and engineering. *A Comprehensive Guide.[Quantum operators]*, pages 648–712.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423.

Sheffer, M. and Oldenburg, D. (2007). Three-dimensional modelling of streaming potential. *Geophysical Journal International*, 169(3):839–848.

Soueid Ahmed, A., Jardani, A., Revil, A., and Dupont, J.-P. (2016). Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. *Water Resources Research*, 52(9):6769–6791.

Stummer, P., Maurer, H., and Green, A. G. (2004). Experimental design: Electrical resistivity data sets that provide optimum subsurface information. *Geophysics*, 69(1):120–139.

Sun, N.-Z. and Yeh, W. W.-G. (1985). Identification of parameter structure in groundwater inverse problem. *Water Resources Research*, 21(6):869–883.

Suski, B., Revil, A., Titov, K., Konosavsky, P., Voltz, M., Dagès, C., and Huttel, O. (2006). Monitoring of an infiltration experiment using the self-potential method. *Water Resources Research*, 42(8).

Tavassoli, Z., Carter, J. N., and King, P. R. (2004). Errors in history matching. *SPE Journal*, 9(03):352–361.

Tikhonov, A. N. and Arsenin, V. Y. (1977). Solutions of ill-posed problems. *New York*, 1(30):487.

Uhlemann, S., Wilkinson, P. B., Maurer, H., Wagner, F. M., Johnson, T. C., and Chambers, J. E. (2018). Optimized survey design for electrical resistivity tomography: combined optimization of measurement configuration and electrode placement. *Geophysical Journal International*, 214(1):108–121.

Vichabian, Y. and Morgan, F. D. (2002). Self potentials in cave detection. *The Leading Edge*, 21(9):866–871.

Vinogradov, J. and Jackson, M. (2011). Multiphase streaming potential in sandstones saturated with gas/brine and oil/brine during drainage and imbibition. *Geophysical Research Letters*, 38(1).

Wang, C., Li, G., and Reynolds, A. C. (2007). Optimal well placement for production optimization. In *Eastern regional meeting*. OnePetro.

Wheelock, B., Constable, S., and Key, K. (2015). The advantages of logarithmically scaled data for electromagnetic inversion. *Geophysical Journal International*, 201(3):1765–1780.

Wiggins, R. A. (1972). The general linear inverse problem: Implication of surface waves and free oscillations for earth structure. *Reviews of Geophysics*, 10(1):251–285.

Williams, C. and Rasmussen, C. (1995). Gaussian processes for regression. *Advances in neural information processing systems*, 8.

Winterfors, E. and Curtis, A. (2008). Numerical detection and reduction of non-uniqueness in nonlinear inverse problems. *Inverse Problems*, 24(2):025016.

Wurmstich, B. and Morgan, F. D. (1994). Modeling of streaming potential responses caused by oil well pumping. *Geophysics*, 59(1):46–56.

Yeh, W. W.-G. (1986). Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water resources research*, 22(2):95–108.

Zandvliet, M., Handels, M., van Essen, G., Brouwer, R., and Jansen, J.-D. (2008). Adjoint-based well-placement optimization under production constraints. *Spe Journal*, 13(04):392–399.