TU Delft

Delft University of Technology

General-purpose Inverse Modeling Framework for Energy Transition Applications Based on Adjoint Method and Operator-Based Linearization

Tian, X.

DOI
10.4233/uuid:54dd845a-93eb-4fa8-9797-83e468ecbce5

Publication date
2023

Document Version
Final published version

Citation (APA)
Tian, X. (2023). *General-purpose Inverse Modeling Framework for Energy Transition Applications Based on Adjoint Method and Operator-Based Linearization*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:54dd845a-93eb-4fa8-9797-83e468ecbce5

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

General-purpose Inverse Modeling Framework for Energy Transition Applications Based on Adjoint Method and Operator-Based Linearization

General-purpose Inverse Modeling Framework for Energy Transition Applications Based on Adjoint Method and Operator-Based Linearization

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.dr.ir. Tim van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
donderdag, 7 september, 2023, om 12:30 uur

door

Xiaoming Tian

Master of Engineering in Oil & Gas Field Development Engineering,
China University of Petroleum, Beijing, China,
geboren te Hunan, China.

Dit proefschrift is goedgekeurd door de

promotor: Dr. D. V. Voskov
promotor: Prof. dr. D. F. Bruhn

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Dr. D. V. Voskov,	Technische Universiteit Delft
Prof. dr. D. F. Bruhn,	Technische Universiteit Delft

Onafhankelijke leden:

Prof. dr. J. D. Jansen	Technische Universiteit Delft
Dr. J. O'Sullivan	The University of Auckland
Dr. H. Hajibeygi	Technische Universiteit Delft
Dr. O. Volkov	Stanford University
Dr. A. Daniilidis	Technische Universiteit Delft

Keywords: Inverse modeling, Energy transition, Adjoint method, DARTS, Geo-energy, Numerical simulation, History matching, Geothermal energy

Printed by: GildePrint

Front & Back: Designed by Xiaoming Tian. The figures at back cover are modified from "Illustrated Account of Well Salt Production in the Qing Dynasty" by Zigong Municipal Bureau of Culture, China

Copyright © 2023 by Xiaoming Tian

ISBN 978-94-6366-727-2

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

The Master, standing by the river, said,
"Time passes by like this, flowing away day and night."

Confucius

CONTENTS

Summary ix

Samenvatting xi

Preface xiii

1 Introduction 1
- 1.1 Social Relevance . . . 1
- 1.2 Inverse Modeling and Adjoint Method . . . 2
- 1.3 Reduced Dimension Techniques . . . 3
- 1.4 Reservoir Simulation Software . . . 4
- 1.5 Thesis Objectives . . . 5
- 1.6 Thesis Outline . . . 6

2 Forward Simulation and Operator-Based Linearization (OBL) 7
- 2.1 Governing Equations . . . 8
- 2.2 Operator-Based Linearization . . . 9
- 2.3 Well Treatment . . . 11
- 2.4 Linear Solver . . . 12

3 Inverse Modeling and Adjoint Gradient Method 15
- 3.1 Inverse Modeling Problems . . . 16
  - 3.1.1 Optimization Theory . . . 16
  - 3.1.2 Model Parameters . . . 17
- 3.2 Adjoint Method . . . 19
  - 3.2.1 Formulation of Adjoint Method . . . 19
- 3.3 Linear Solver for Adjoint Method . . . 21

4 Implementation and Validation of Adjoint Method in DARTS 23
- 4.1 One-Dimensional Example in MATLAB . . . 24
- 4.2 Implementation in C++/Python . . . 25
  - 4.2.1 Assembly of Derivatives in Adjoint Method . . . 25
  - 4.2.2 Implementation in Engines of DARTS . . . 29
  - 4.2.3 Python API for Inverse Modeling Feature . . . 31
- 4.3 Validation of Adjoint Method in DARTS . . . 31

5 Efficient Proxy Models Based on Adjoint Method 33
- 5.1 Discrete Well Affinity (DiWA) Data-Driven Proxy Model . . . 34
  - 5.1.1 Connectivity Graph for Proxy Model . . . 35
  - 5.1.2 Model Parameters And Objective Function . . . 36
  - 5.1.3 Validation of DiWA Proxy Model . . . 36

5.2 Stochastic DiWA Proxy Model . . . 42
5.2.1 Methodology . . . 43
5.2.2 The Application of Stochastic DiWA Proxy Model . . . 45
5.3 Conclusion And Discussion . . . 53

6 Efficient Inverse Modeling for Energy Transition Applications 57
6.1 Bayes' Theorem . . . 58
6.2 Application of Geothermal Project . . . 59
6.2.1 Introduction . . . 59
6.2.2 PCA-based Parameterization . . . 61
6.2.3 Modeling of Electromagnetic Response . . . 63
6.2.4 Geothermal Reservoir . . . 66
6.2.5 Conclusion And Discussion . . . 74
6.3 Application to CO_2 Storage Project . . . 74
6.3.1 Introduction . . . 75
6.3.2 FluidFlower Benchmark Description . . . 76
6.3.3 History Matching Based on Zonated Model and Simple Poisson-like Solver . . . 77
6.3.4 History Matching of FluidFlower Benchmark Using Two-stage Approach . . . 83
6.3.5 History Matching Results . . . 86
6.3.6 Regularization and RML Based on the Concentration Interpreted From Images . . . 88
6.3.7 Conclusion And Discussion . . . 89

7 Recapitulation and Conclusions 91
7.1 Adjoint Method And the Inverse Modeling Feature of DARTS . . . 91
7.2 Applications of the Inverse Capabilities of DARTS in DiWA Model . . . 92
7.3 Inverse Modeling for Energy Transition Projects . . . 93
7.4 Future Perspectives . . . 94

References 106

Acknowledgements 107

Curriculum Vitæ 111

List of Publications 113

SUMMARY

Inverse modeling is an important technique used in numerical modeling for geo-energy applications, for example, petroleum, geothermal, and CCS projects. It helps to calibrate models and improve the predictive capabilities of the numerical model for real-world projects. This dissertation focuses on inverse modeling for geo-energy systems, especially the implementation and application of the inverse modeling framework in DARTS. The framework was developed using the adjoint method and is designed to facilitate the assembly of the associated derivatives based on Operator-Based Linearization (OBL). The adjoint method is chosen in this study due to its efficiency and the fact that it provides an analytical solution for the gradient evaluation of the model parameters used in inverse modeling.

The forward simulation is the basis of inverse modeling, and it is briefly introduced in Chapter 2. The concept of the OBL is explained and its advantages of assembling the Jacobian and the associated derivatives for the adjoint method are also discussed.

Starting from Chapter 3, we will shift from the description of the forward simulation to inverse modeling, which is the main focus of this study. The definition of the objective function employed in geo-energy problems and optimization theory is presented in this chapter. A detailed explanation of the adjoint method will be provided to show the process of solving the Lagrangian multiplier and the final adjoint-based gradient. The associated derivatives that relate to the adjoint method will be explicitly explicated in the form of matrices and vectors by utilizing a one-dimensional model example in conjunction with the architecture of DARTS. Additionally, we will also discuss the selection of model parameters, observation types, and the Dirac measurement function.

Chapter 4 shows the development of the prototype of inverse modeling using the adjoint method in MATLAB. It was validated by using a two-parameter inverse modeling problem and compared with numerical gradient methods. This framework was later translated into C++ to enhance its efficiency and scalability. We focused on two types of engines, namely, the "super engine" and "thermal engine," and designed the adjoint method to solve the inverse modeling problems associated with these engines. We conducted a comparative analysis between the adjoint-based gradient and the numerical gradient, and the results showed that the angle between the two gradients was insignificant. Furthermore, the CPU time required for the adjoint method was considerably lower than that of the numerical method. The evaluation of the gradient using the adjoint method led to an efficiency improvement of two to three orders of magnitude in that comparative example.

Once the inverse modeling framework is done, it was applied to various types of projects. Chapter 5 presents the application example of the data-driven proxy model. Based on this framework, the Discrete Well Affinity (DiWA) data-driven model was proposed and validated. The DiWA model was further extended to stochastic DiWA to consider geological uncertainties. The results show that DiWA model can provide good pro-

duction forecast results, while the training of DiWA model is highly efficient because of the application of adjoint method and the utilization of the coarse grid.

To further advance the application of this framework in energy transition projects, a wider range of observations is considered and integrated into the objective function. This integration allows for a more diverse range of observations to be incorporated into the framework for various geo-energy projects. For example, Chapter 6 describes a geothermal inverse modeling project that incorporates the time-lapse reservoir temperature observations inferred from electromagnetic data. This example also illustrates the framework's potential for characterizing geological information. Similarly, another energy transition project focused on CO_2 storage incorporated time-lapse tracer concentration data to perform history matching based on this framework.

SAMENVATTING

Inverse modelleren is een belangrijke techniek die wordt gebruikt in numerieke modellering voor geo-energietoepassingen, zoals bijvoorbeeld aardolie-, geothermische en CCS-projecten. Het helpt om modellen te kalibreren en de voorspellende mogelijkheden van het numerieke model te verbeteren voor projecten in de echte wereld. Deze dissertatie richt zich op inverse modellering voor geo-energiesystemen, met name de implementatie en toepassing van het inverse modelleren binnen het kader van DARTS. Het kader is ontwikkeld met de adjoint-methode en is ontworpen om de samenstelling van de bijbehorende afgeleiden op basis van Operator-Based Linearization (OBL) te vergemakkelijken. De adjoint-methode is gekozen in dit onderzoek vanwege de efficiëntie en het feit dat het een analytische oplossing biedt voor de gradiëntbepaling van de modelparameters die worden gebruikt in inverse modellering.

De voorwaartse simulatie vormt de basis van inverse modellering en wordt kort geïntroduceerd in Hoofdstuk 2. Het concept van OBL wordt uitgelegd en de voordelen voor het samenstellen van de Jacobian en de bijbehorende afgeleiden voor de adjoint-methode worden besproken.

Vanaf Hoofdstuk 3 gaan we door met de beschrijving van invers modelleren, wat de belangrijkste focus van dit onderzoek is. De definitie van de objective function die wordt toegepast bij geo-energieproblemen en optimalisatietheorie wordt in dit hoofdstuk voorgesteld. Een gedetailleerde uitleg van de adjoint-methode zal worden gegeven om het proces van het oplossen van de Lagrange-vermenigvuldiger en de uiteindelijke adjoint-gebaseerde gradiënt te laten zien. De bijbehorende afgeleiden die verband houden met de adjoint-methode zullen expliciet worden uitgelegd in de vorm van matrices en vectoren door middel van een ééndimensionaal modelvoorbeeld in combinatie met de architectuur van DARTS. Daarnaast zullen we ook de selectie van modelparameters, observatietypen en de Dirac-meetfunctie bespreken.

In Hoofdstuk 4 wordt de ontwikkeling van het prototype van inverse modellering met behulp van de adjoint-methode in MATLAB getoond. Dit werd gevalideerd door gebruik te maken van een tweeparameter inverse modelleringsprobleem en vergeleken met numerieke gradiëntmethoden. Dit kader werd later vertaald naar C++ om de efficiëntie en schaalbaarheid te verbeteren. We concentreerden ons op twee soorten motoren, namelijk de "super engineën "thermal engine,ën ontwierpen de adjoint-methode om de inverse modelleringsproblemen geassocieerd met deze motoren op te lossen. We voerden een vergelijkende analyse uit tussen de adjoint-gebaseerde gradiënt en de numerieke gradiënt, en de resultaten toonden aan dat de hoek tussen de twee gradiënten verwaarloosbaar was. Bovendien was de CPU-tijd die nodig was voor de adjoint-methode aanzienlijk lager dan die van de numerieke methode. De evaluatie van de gradiënt met behulp van de adjoint-methode leidde tot een efficiëntieverbetering van twee tot drie ordes van grootte in dat vergelijkende voorbeeld.

Zodra het inverse modelleringskader gereed is, wordt het toegepast op verschillende

soorten projecten. Hoofdstuk 5 presenteert het toepassingsvoorbeeld van het op data gebaseerde proxy-model. Op basis van dit kader werd het Discrete Well Affinity (DiWA) op data gebaseerde model voorgesteld en gevalideerd. Het DiWA-model werd verder uitgebreid naar een stochastisch DiWA-model om geologische onzekerheden in overweging te nemen. De resultaten tonen aan dat het DiWA-model goede productievoorspellingen kan leveren, terwijl de training van het DiWA-model zeer efficiënt is vanwege de toepassing van de adjoint-methode en het gebruik van de grove raster.

Om de toepassing van dit kader in energietransitieprojecten verder te bevorderen, worden een breder scala aan waarnemingen overwogen en geïntegreerd in de doelfunctie. Deze integratie maakt het mogelijk om een meer diverse reeks waarnemingen in het kader op te nemen voor verschillende geo-energieprojecten. Bijvoorbeeld, Hoofdstuk 6 beschrijft een geothermisch invers modelleringsproject dat de tijdreeksen van reservoirtemperatuurwaarnemingen opneemt die zijn afgeleid van elektromagnetische gegevens. Dit voorbeeld illustreert ook het potentieel van het kader voor het karakteriseren van geologische informatie. Op vergelijkbare wijze heeft een ander energieovergangsproject dat gericht is op CO_2-opslag tijdreeksen van tracerconcentratiegegevens opgenomen om history matching uit te voeren op basis van dit kader.

PREFACE

I can still vividly recall the day I asked Mr. Cees Timmers to forward my resume and motivation letter to Denis. I was not sure if my description of developing geothermal energy to address the energy crisis in my motivation letter would convince Denis. Fortunately, after an interview about my research work during my master's studies, Denis officially sent me an invitation letter from Delft University of Technology.

After a year of dealing with paperwork and making preparations, I finally boarded the flight to the Netherlands. As the plane touched down at Schiphol airport, the weather outside with its strong winds and heavy rain immediately exposed me to the typical Dutch weather that I have heard before. It was quite a challenge, considering that just over ten hours ago, I had departed from the scorching hot Changsha airport, wearing only a T-shirt.

Upon reaching TU Delft's campus, I quickly immersed myself in the work on the inverse modeling framework. However, the initial process did not go smoothly. For about three months, I made little progress and even doubted the validity of the adjoint method. When I inquired about it with people around me, it appeared that not everyone heard or understood the adjoint method, and there were hardly any who could provide detailed technical insights. So, I knew I was in trouble and had to rely on myself. After engaging in multiple discussions with Denis and thoroughly studying Jan-Dirk's material on the adjoint method, I finally succeeded just before Christmas when I observed that the optimization trajectories formed by the adjoint gradients and numerical gradients almost completely overlapped. This success filled me with excitement, and I immediately shared the good news with Denis.

However, the happy times were short-lived. In 2020, it was the first time I couldn't celebrate the Chinese New Year with my family back home. Around the same time, COVID was first reported and rapidly spreading in China. I closely followed its development and made donations through the Chinese embassy. During that period, most people didn't perceive it as a severe infectious disease, and Europe hadn't yet become vigilant. Little did we know the immense impact COVID would have on people's lives worldwide.

Fortunately, even during the work-from-home phase, my research progress on the adjoint method wasn't significantly hindered. After initially implementing the adjoint method in MATLAB, I successfully integrated it into DARTS using C++, largely improving its computational performance. While gradually applying the adjoint method to the DiWA model, geothermal projects, and CCS projects, I further optimized the C++ of the adjoint method to enhance its flexibility, robustness, and generality for inverse modeling. Throughout this process, several master students I supervised utilized the adjoint method in DARTS for history matching and inverse modeling. Nowadays, DARTS has become an open-source, efficient subsurface flow simulator, and the source code of DARTS' adjoint method is also open-source and available to the public. I am delighted

to see more users using the inverse modeling feature of DARTS based on the adjoint method in the future.

I am glad that my four years of work on the adjoint method have made a small contribution to the open-source community. Accessibility to simulator source code is a crucial requirement of the adjoint method, and DARTS' open-source nature has facilitated its implementation for inverse modeling. The adjoint method's accuracy and efficiency have also contributed to DARTS becoming a more feature-rich and efficient simulator. This positive feedback loop is precisely why the open-source community continues to thrive. Now, I have successfully achieved the goal I mentioned in my motivation letter to Denis four years ago: to return to China after obtaining my Ph.D. degree and continue my academic career in the field of geothermal energy. My aspiration is to bring back this open-source spirit to the geothermal energy development community in China.

Xiaoming Tian
Delft, July 2023

1

INTRODUCTION

1.1. SOCIAL RELEVANCE

Geo-energy resources and their management have significant implications for human societies and industrial activities. Hydrocarbon resources are a crucial component of the global energy supply chain, contributing significantly to the economy and employment in many countries [99]. However, the consumption of fossil fuels has a significant impact on global climate change, and people need to explore clean and sustainable energy supply for the better development of human society and Earth. Geothermal energy is another subsurface resource that has the potential to contribute to a sustainable energy mix [33]. Geothermal energy systems utilize the heat stored in subsurface reservoirs to generate electricity or provide heating and cooling. Apart from looking for a cleaner geo-energy supply, we also need to think about how to reduce the existing carbon emission in the atmosphere. Escalating greenhouse gas concentrations in the atmosphere and their effect on climate have become an urgent and global concern. Large reductions in CO_2 emissions are mandated by the most recent IPCC reports, and virtually all the conceivable pathways require engineered removal of new carbon pollution sources [43, 42]. Carbon capture and storage (CSS) has been proven to be one of the most promising solutions to this environmental issue. Typically, CCS can reduce 85–90% CO_2 emissions from large point emission sources [60], for example, power plants, cement kiln plants, etc.

However, the successful implementation of the above hydrocarbon, geothermal, and CCS projects requires a detailed understanding of subsurface geology, petrophysical properties, and complex physics. Numerical modeling can assist in the design, optimization, and risk assessment of these geo-energy systems. Inverse modeling, in particular, is an important technique used in numerical modeling to calibrate models and improve their predictive capabilities. It uses field observations to infer model parameters, which is crucial for ensuring the accuracy and reliability of numerical models, especially when they are used to make predictions for real-world projects. Therefore, the development of efficient and accurate inverse modeling techniques is crucial for the successful imple-

mentation of geo-energy systems. This dissertation will focus on inverse modeling for geo-energy systems.

1.2. INVERSE MODELING AND ADJOINT METHOD

Essentially, inverse modeling is an optimization problem where the misfit between the observations and model response is minimized. There is a bunch of approaches to solving such optimization problems. Based on the presence of gradients in the course of optimization, the optimization methods are classified into gradient-free and gradient-based methods. For the gradient-free method, many of them directly aim at finding the global optimum without the gradient information provided, for example, simulated annealing [57], genetic algorithms [36], particle swarm algorithm [21] and so on. This kind of method is suitable for the highly nonlinear problem with many local extrema. For the gradient-based method, the idea is iteratively updating a set of model parameters to search for smaller objective function values (or larger objective function values. It depends on the problem one studies) along the direction of the gradient, which corresponds to the steepest descent. Therefore, the gradient-based method is known for its efficiency of search for the optimum.

This study will mainly focus on the gradient-based method for inverse modeling. It is critical to efficiently evaluate the gradients with respect to the model parameters. A very straightforward approach for evaluating the gradient is to calculate the partial derivatives with each of the model parameters numerically using the Finite Difference Method (FDM). To be more specific, the gradient value is equal to the difference of objective function values before and after applying a feasible perturbation on model parameters. It measures the sensitivity of the objective function to the model parameters. However, this procedure usually requires large amounts of computational effort. This computational issue is more pronounced when the degrees of freedom of the model parameters are very high. Apart from the computational issue, the gradient evaluation based on FDM possibly generates unreasonable gradient values if the selection of perturbation magnitude is not suitable for the given form of the objective function [73].

A promising method to solve this problem is applying the adjoint method to calculate gradients, especially in the case of the model parameters with high dimensionality. This method is particularly useful in situations where it is computationally expensive to compute the gradient of the objective function directly, but relatively inexpensive to solve the original differential equations, for example, the governing equation of the geo-energy system. It allows for the efficient calculation of the gradient by reusing the solution of the governing equations. In the field of geo-energy reservoir engineering, with the development of computer technology, the adjoint method was applied in numerical reservoir simulation to do history matching [11], petroleum recovery process [71, 23, 85], thermal recovery process [113] and so on. Later, as the conceptions of "smart well" and "smart field" arose, adjoint-based optimization was widely studied and applied in the field of reservoir engineering and reservoir management [7, 91, 93, 107].

1.3. REDUCED DIMENSION TECHNIQUES

With the increasing complexity of the reservoir simulation problem mentioned above, more and more high-resolution models with complicated physics are needed, as it is believed that the high-dimension model is capable of describing complex geological features through a set of grid blocks and associated rock and fluid properties. However, in many cases, the reliability of geological information is questionable or it is even not available. Even though it is possible to develop a high-fidelity model on a reliable basis of reservoir geology, a high-resolution model can exceed a few million blocks and may take hours or even days to simulate. It is still not computationally feasible to perform history matching or production optimization at such resolution since it involves a large number of simulation runs.

Different methods have been developed to overcome this issue. Those methods fall into two categories: simplified full-field models, and data-driven approaches. Methods such as upscaling, multi-scale methods and streamline simulation fall into the first category. The upscaling method uses a coarser grid model to mimic a high-resolution reservoir model [20]. The multi-scale method takes the fine-scale grid and locally transfers it to the global representation using the basis function concept [46, 31]. Next, the simulation is performed at a coarse grid with a fine grid solution (e.g. pressure) reconstructed based on pre-processed basis functions. Later, de Moraes et al. [74] pointed out that if restricting the mass balance equations from fine-scale to coarse-scale and then prolonging the solution back to the original scale, the derivatives of the restriction and prolongation operators help with an accurate estimation of the gradients on the coarse grid, even when the domain is highly heterogeneous. The streamline method is an Eulerian-Lagrangian approach that translates a real fluid transport problem into a one-dimensional problem solved along streamlines [3].

All methods in the first category require an underlying geological characterization as a basis for construction. However, in many cases, this information is not available or its reliability is questionable. Does it mean we cannot solve optimization or history-matching problems efficiently? Methods from the second category resolve this issue. The data-driven method assumes the building of a proxy model with a sufficient amount of degree of freedom to accurately mimic a realistic reservoir response based on its calibration to the production data. With frequent, sustained, and accurate data being fed into a reliable regression framework, data-driven models can provide an accurate forecast for the given reservoir [102].

There are many data-driven approaches available in the industry including the statistical data-driven model proposed by [44]; reduced-order models [10]; Capacitance Resistance Model [116, 2, 34]; flow-network model [59, 86, 5, 56] where a complex 3D flow is represented as a set of 1D finite-difference reservoir models; Interwell Numerical Simulation Model (INSIM) [119, 120] and INSIM-FT-3D [30] approach which applies a new Riemann solver based on a convex-hull method that helps to solve the Buckley-Leverett problem with gravity and allows for the inclusion of wells with arbitrary trajectories with multiple perforations. In the DARTS framework, Tian et al. [100] implemented the adjoint method and proposed a physics-based data-driven model called Discrete Well Affinity (DiWA) model. This model was further extended as a stochastic DiWA proxy model method for efficient history matching and production forecast [102, 103].

Apart from the above methods of "physically" coarsening the model grid cells, mathematically speaking, the degrees of freedom of the model can also be reduced by using fewer "key" components in the space, while these "key" components are mostly independent of each other and they represent and capture the most information of the original space. Principal Component Analysis (PCA) is a powerful tool to quantify the dependency among the components and boil down the "key" components under a specific criterion. PCA has been successfully applied in many history matching problems for multi-Gaussian fields [26, 79, 93]. However, Sarma et al. [92] pointed out that the direct use of PCA on non-Gaussian fields may lead to "Gaussian-looking" models after history matching. Therefore, they introduced kernel principal component analysis (KPCA) to solve this problem. Ma and Zabaras [67] later refined the KPCA approach. Although the KPCA approach focuses more on representing the multiple "feature" space in the reservoir, this method is essentially strongly nonlinear and brings challenging numerical issues. Vo and Durlofsky [106] proposed an optimization-based PCA method for the low-dimensional parameterization of complex geological models. Bukshtynov et al. [8] also introduced PCA parameterization in their gradient-based optimization framework for the closed-loop reservoir management problem. In their study, the adjoint method is applied and incorporated with the PCA parameterization to compute the new gradient in the reduced-dimension space.

1.4. RESERVOIR SIMULATION SOFTWARE

Gradient evaluation, history matching, and production optimization technologies mentioned in the previous sections are done based on forward numerical simulation. A general process of performing numerical reservoir simulation includes discretization and linearization of governing equations. The reservoir simulators are designed to linearize and solve the governing partial differential equations in discrete form. Some commercial software was developed for simulating the fluid flow underground. For example, Eclipse is a commercial simulator extensively used in the oil and gas industry [22]. CMG is another commercial software widely used in the petroleum industry [14]. It is known for its various simulators about different topics such as IMEX (black oil and unconventional simulator), STARS (thermal simulator), GEM (compositional simulator), etc.

However, these commercial simulators are not open source and therefore unfriendly for the users to add or customize their own models and physics into the simulations. In this case, the open-source simulators provide a more flexible and free platform for the users to implement and test their innovative methods and ideas. There are many open-source simulators that provide various capabilities for the simulations of geo-energy problems. The MATLAB Reservoir Simulation Toolbox (MRST) [61] is a free open-source simulator for reservoir modeling and simulation. It has lots of capabilities for simulating conventional geo-energy problems such as black oil and compositional models. It also provides some tools for discretization, upscaling, diagnostics, and so on. GEOSX [27] is another open-source simulator that models both the conventional underground and the energy transition problems. It is especially known for its capability of modeling geomechanical effects. DuMux [19] is a simulator written in C++ based on Dune. It is designed for groundwater management, petroleum engineering, and geothermal energy production. PorePy [50] was initially developed using Python code for the study of multiphysics

processes in deformable porous media. Now it is also extended to more general porous media problems.

Recently, Delft Advanced Reservoir Terra Simulator (DARTS) was officially released as a high-performance open-source simulator framework aiming at simulating multi-phase multi-component fluid flow problems considering complex physical effects [17]. The physics calculation of DARTS utilizes the Operator-Based Linearization (OBL) technique. This technique was proposed by [108] for addressing the challenges of balancing the accuracy of the numerical model and the performance of the simulator. The main idea of the OBL approach is parameterizing the state-dependent operators at the pre-processing stage which simplifies and accelerates the assembly of Jacobian. Later, the OBL technique was improved by adaptive parameterization [55] and implementation at GPU architecture [52]. Currently, DARTS framework has been utilized for modeling of advanced petroleum [53, 64], geothermal [55, 112] and CO_2 sequestration [47, 65] applications. Apart from the feature of efficient forward simulation, DARTS also has an inverse modeling capability based on the adjoint gradient method. Adjoint-based gradients were first implemented in DARTS for the multi-component multi-phase system of the petroleum-related reservoir simulation problem [100]. Later, adjoint capabilities have been extended for a more general spectrum of energy transition applications. This development and its application will be described and explained in detail in this dissertation.

1.5. THESIS OBJECTIVES

The research objectives of this dissertation can be summarized in the following points:

- Implementing the inverse modeling capabilities of DARTS based on the adjoint method. This procedure includes the development of the prototype of the inverse modeling framework in MATLAB and the final implementation in C++ for higher computational efficiency.

- Developing the Discrete Well Affinity (DiWA) proxy model method based on the adjoint method for the efficient data-driven model and production forecast.

- Making the inverse modeling framework of DARTS more flexible to use various field observations. In the application to a geothermal project, the definition of the objective function is designed to be very flexible for incorporating various kinds of observations, for example, the time-lapse temperature interpreted from the electromagnetic data.

- Extending and investigating the application of the inverse modeling framework to more energy transition topics. DARTS is capable of solving various types of energy transition problems. Accordingly, the adjoint method is utilized in DARTS to develop the inverse modeling framework for these energy transition problems such as CO_2 storage project.

1

1.6. THESIS OUTLINE

This dissertation proceeds as follows. Chapter 2 briefly introduces the governing equations of forward simulation. It describes the general form of the governing equation that considers various physical effects in the subsurface flow such as capillarity, gravity, and chemical reactions. The OBL method for linearizing the governing equations and the treatments of the boundary conditions at wells are demonstrated. In Chapter 3, the inverse modeling and the adjoint method are explained in detail. This chapter starts with a discussion of the inverse modeling problem and the model parameters that are adjusted in the course of history matching. Then, the adjoint method is introduced in order to improve the efficiency of the gradient evaluation used in inverse modeling. To clearly present the associated derivatives in the adjoint method, a one-dimensional model is used to demonstrate the detailed procedure of assembling those derivatives in DARTS.

The first prototype of inverse modeling using the adjoint method is implemented in MATLAB and then implemented in C++ for higher computational performance. Once the framework is implemented, it is validated and compared with the gradient evaluation using the conventional numerical method. The validation and the comparison can be found in Chapter 4. Next, Chapter 5 presents the application of the inverse modeling framework of DARTS to the data-driven proxy model. The Discrete Well Affinity (DiWA) data-driven model is proposed and validated. It is further extended to the stochastic DiWA proxy model method.

Chapter 6 provides more applications of the inverse modeling framework on energy transition projects, including geothermal and CO_2 storage projects. In both projects, uncertainty quantification is considered for real geo-energy problems. Finally, Chapter 7 concludes the work and raises the future perspectives of the development and extension of the inverse modeling feature of DARTS.

2

FORWARD SIMULATION AND OPERATOR-BASED LINEARIZATION (OBL)

Summary

This chapter briefly introduces the governing equations of forward simulation and the Operator-Based Linearization (OBL) proposed in [108]. The forward simulation is the basis of the inverse modeling, as the "model response" is directly generated from the forward simulation. The governing equations are first defined in the operator form and then linearized using the OBL method. Most of the property evaluations from OBL will be re-used in the inverse modeling for the efficient assembly of the associated matrices in the adjoint method. The well treatment is a critical part of defining the inner boundary conditions (i.e. well controls) of governing equations, which also affects the treatment of well observations in the adjoint approach [103].

Parts of this chapter have been published in Society of Petroleum Engineers, **SPE-212169-MS**, (2023) [103].

2.1. GOVERNING EQUATIONS

The fluid flow problem in porous media involves solving the mass and energy governing equations. In this section, the governing equations coupling both energy and mass terms are described. The energy and mass conservation equations describe a flow dynamic system bounded in the domain with volume Ω and surface Γ. The conservation equation can be written as:

$$\frac{\partial}{\partial t}\int_{\Omega} M^c d\Omega + \int_{\Gamma} \boldsymbol{F}^c \cdot \boldsymbol{n} d\Gamma = \int_{\Omega} Q^c d\Omega, \tag{2.1}$$

where M^c is the accumulation term for the c^{th} component ($c = 1,\ldots,n_c$, index of the mass components [e.g., water, CO_2] and $c = n_c + 1$, index of the energy quantity), $\boldsymbol{F}_c$ is the flux term of the c^{th} component, $\boldsymbol{n}$ is the unit normal direction pointing outward to the domain boundary, and Q_c is the source/sink term of the c^{th} component.

The accumulation term M^c needs to be split into two parts if both the mass and energy accumulations are considered. For the mass accumulation of a given component c, it can be written as:

$$M^c = \phi \sum_{j=1}^{n_p} x_{cj}\rho_j s_j, \quad c = 1,\ldots,n_c, \tag{2.2}$$

where ϕ is porosity, s_j is the saturation of phase j, ρ_j is the density [kmol/m^3] of phase j, and x_{cj} is the molar fraction of c component in j phase. For the energy accumulation term, it includes the internal energy of fluids and rock:

$$M^{n_c+1} = \phi \sum_{j=1}^{n_p} \rho_j s_j U_j + (1-\phi)U_r, \tag{2.3}$$

where U_j is the internal energy [kJ] pf phase j and U_r is rock internal energy [kJ]. The porosity of the compressible rock is given as:

$$\phi = \phi_0\big(1 + c_r(p - p_{\text{ref}})\big), \tag{2.4}$$

where ϕ_0 is the initial porosity, c_r is the rock compressibility [1/bar], and p_{ref} is the reference pressure [bar].

The flux term $\boldsymbol{F}_c$ should also be split into two parts: mass flux and energy flux. The mass flux of given component c is given as:

$$\boldsymbol{F}^c = \sum_{j=1}^{n_p} \big(x_{cj}\rho_j \boldsymbol{v}_j + s_j\phi \boldsymbol{f}_{cj}\big), \quad c = 1,\ldots,n_c. \tag{2.5}$$

Here $\boldsymbol{v}$ is the velocity and follows Darcy's law considering gravity effect for given phase j:

$$\boldsymbol{v}_j = -\mathbf{K}\frac{k_{rj}}{\mu_j}(\nabla p_j - \boldsymbol{\gamma_j}\nabla z), \tag{2.6}$$

where $\mathbf{K}$ is the permeability tensor [mD], k_{rj} is the relative permeability of phase j, μ_j is the viscosity of phase j [mPa$\cdot$s], p_j is the pressure of phase j [bar], $\boldsymbol{\gamma_j} = \rho_j \bar{g}$ is the

specific weight [N/m^3] (where $\bar{g}$ is the gravity acceleration), z is the depth vector [m]. The diffusion term $\boldsymbol{f}_{cj}$ is following Fick's law where

$$\boldsymbol{f}_{cj} = -\rho_j \mathbf{D}_{cj} \nabla x_{cj} \tag{2.7}$$

and $\mathbf{D}_{cj}$ is the diffusion coefficient [m^2/day]. The energy flux includes the thermal convection and conduction terms:

$$\boldsymbol{F}^{n_c+1} = \sum_{j=1}^{n_p} h_j \rho_j \boldsymbol{v}_j - \kappa \nabla T. \tag{2.8}$$

Here h_j is phase enthalpy [kJ/kg] and κ is effective thermal conductivity [kJ/m/day/K] defined as

$$\kappa = \phi \sum_{j=1}^{n_p} \kappa_j s_j + (1-\phi)\kappa_r, \tag{2.9}$$

where κ_j and κ_r are thermal conductivity of fluid phase j and rock respectively.

The source/sink term Q^c mainly includes the terms related to the chemical reaction. It can be present as:

$$Q^c = \sum_{k=1}^{n_k} \nu_{ck} r_k, \quad c = 1, \ldots, n_c, \tag{2.10}$$

where ν_{ck} is the stoichiometric coefficient associated with chemical reaction k for the component c and r_k is the rate for the reaction. For the energy equation, it can be written as:

$$Q^{n_c+1} = \sum_{k=1}^{n_k} y_{ek} r_{ek}, \tag{2.11}$$

where $\nu_{e\theta}$ is the stoichiometric coefficient associated with kinetic reaction θ for the energy and $r_{e\theta}$ is the energy rate for kinetic reaction.

To solve this mass and energy conservation equation numerically, it needs to be discretized spatially and temporally. Based on the finite volume method using two-point flux approximation, the discretized form of Equation (2.1) for the i^{th} reservoir gridblock can be written as:

$$g_i^c = V_i\Big(M_i^c(\omega_i) - M_i^c(\omega_{i(k-1)})\Big) - \Delta t\Big(\sum_l a_l F_l^c(\omega) + V_i Q_i^c(\omega)\Big) = 0, \quad c = 1, \ldots, n_c + 1. \tag{2.12}$$

Equation (2.12) shows the residual form of discretized Equation (2.1). We introduce V_i as the volume of the i^{th} gridblock and ω_i as the state variables at the current time step. In addition, $\omega_{i(k-1)}$ is the state variables at the previous time step and Δt is the time step. The term a_l is the contact area of the interface l between neighboring grids.

2.2. OPERATOR-BASED LINEARIZATION

The operator-based linearization (OBL) approach, as described in [108], is employed in DARTS to linearize the governing equation of Equation (2.1). The key concept behind OBL is discretizing the physical parameter space to improve the efficiency of the evaluations of the physical properties and their associated derivatives. This approach was

proposed and validated for complex multi-phase flow problems coupled with the thermal process. With the OBL method, the governing equation can be written in the form of state-dependent operators. The operator values and derivatives are evaluated and interpolated based on supporting points under the different resolutions of physical parameter space. A more advanced adaptive parameterization technique for OBL calculation was later proposed in [53] to largely reduce the computational time of parameterization in high-dimension physical parameter space. Compared with the conventional Jacobian assembly procedure in numerical simulation, the OBL method makes the Jacobian assembly procedure more flexible and efficient for highly complex physical problems. The inverse modeling with the adjoint method also benefits from the OBL's efficient assembly of relevant vectors and matrices.

Based on OBL, the residual form of governing Equation (2.12) can be re-arrange to (here we omit the gridlock index i):

$$V\phi_0[\alpha_c(\omega)-\alpha_c(\omega_{(k-1)})]-\Delta t\sum_{l\in\Gamma(i)}\sum_{j=1}^{n_p}[\bar{T}^l\beta^l_{cj}(\omega^{up})\Delta\psi^l_j+\bar{T}^l_d\gamma^l_j(\omega)\Delta\chi_{cj}]+\Delta tV\eta_c(\omega)=0. \tag{2.13}$$

Here, the pressure, temperature, and overall composition are combined and represented as a single physical state variable ω for a given gridblock i at the current timestep, $\omega_{(k-1)}$ is the physical state at the previous timestep, ω^{up} denotes the physical state of the upstream gridblock, $\bar{T}^l$ and $\bar{T}^l_d$ are the fluid and diffusive transmissibilities respectively and $\Gamma(i)$ is a set of interfaces for gridblock i. The Greek letters α, β, γ, χ, and η denote the state-dependent operators, which are defined as:

$$\alpha_{cj}(\omega) = \Big(1+c_r(p-p_{ref})\Big)\sum_{j=1}^{n_p}x_{cj}\rho_j s_j,\ c=1,\dots,n_c; \tag{2.14}$$

$$\beta_{cj}(\omega) = x_{cj}\rho_j k_{rj}/\mu_j,\ c=1,\dots,n_c,\ j=1,\dots,n_p; \tag{2.15}$$

$$\gamma_j(\omega) = \Big(1+c_r(p-p_{ref})\Big)s_j,\ j=1,\dots,n_p; \tag{2.16}$$

$$\chi_{cj}(\omega) = D_{cj}\rho_j x_{cj},\ c=1,\dots,n_c,\ j=1,\dots,n_p; \tag{2.17}$$

$$\eta_c(\omega) = \sum_{j=1}^{n_p}y_{cj}r_j(\omega),\ c=1,\dots,n_c. \tag{2.18}$$

The overall pressure difference between two neighboring blocks 1 and 2 at their interface l is computed as follows:

$$\Delta p^l_j = p_1-p^c_j(\omega_1)-\Big(p_2-p^c_j(\omega_2)\Big)-\frac{\rho_j(\omega_1)+\rho_j(\omega_2)}{2}\bar{g}(z_2-z_1), \tag{2.19}$$

where p^c_j is the capillary pressure. Equation (2.19) actually represents the phase-potential-upwinding (PPU) strategy applied in DARTS [55, 66].

The operator form of the energy equation is given as:

$$\begin{aligned}&\phi_0 V\left[\alpha_{ef}(\omega)-\alpha_{ef}(\omega_{k-1})\right]\\&\quad+\left(1-\phi_0\right)VU_r\left[\alpha_{er}(\omega)-\alpha_{er}(\omega_{k-1})\right]+\sum_l\Delta t\Gamma^l\Delta p^l_j\beta_e(\omega)\\&\quad+\sum_l\Delta t\Gamma^l_e\Delta T(\omega)\left[\phi_0\gamma_{ef}(\omega)+\left(1-\phi_0\right)\kappa_r\gamma_{er}(\omega)\right]=0,\end{aligned} \tag{2.20}$$

where the subscript e, f, and r denote the properties or operators related to energy, fluid, and rock, respectively. Γ_e^l is the coefficient related to thermal transmissibility. The operators are defined as:

$$\alpha_{ef}(\omega) = \left(1 + c_r\left(p - p_{ref}\right)\right) \sum_{j=1}^{n_p} \rho_j s_j U_j, \tag{2.21}$$

$$\alpha_{er}(\omega) = \frac{1}{1 + c_r\left(p - p_{ref}\right)}, \tag{2.22}$$

$$\beta_e(\omega) = \sum_{j=1}^{n_p} h_j \rho_j \frac{k_{rj}}{\mu_j}, \tag{2.23}$$

$$\gamma_{ef}(\omega) = \left(1 + c_r\left(p - p_{ref}\right)\right) \sum_{j=1}^{n_p} s_j \kappa_j, \tag{2.24}$$

$$\gamma_{er}(\omega) = \alpha_{er}(\omega). \tag{2.25}$$

Similar to the governing equation of mass in Equation (2.13), the energy equation is also formed by the accumulation and flux term. Equation (2.21) and Equation (2.22) represent the energy accumulation for fluid and rock, respectively. Equation (2.23) and Equation (2.24) describe the energy flux for the fluid and rock, respectively.

Finally, the complex evaluations of the physical properties and their derivatives with respect to the unknown variables are encapsulated into the evaluation of these operators and their derivatives, which can be performed very efficiently in the pre-processing stage or adaptively with a limited number of supporting points. For those points located in the physical regions that are defined by the supporting points, they are interpolated based on multi-linear interpolation. The discussions about the non-linearity of the system and the OBL resolution can be found in [55, 53, 108].

2.3. WELL TREATMENT

The well controls and constraints describe the inner boundary conditions of the governing equations. In DARTS, a connection-based multi-segment well model is used to simulate the flow in the wellbore [54]. Multiple well blocks representing the perforations are connected to the reservoir girdblocks. Those well blocks (i.e. well body) are then connected to a ghost well block (i.e. wellhead). This ghost well block is actually the placeholder where the well controls and constraints are applied.

For bottom hole pressure (BHP), a target pressure value is defined at the ghost well block:

$$p - p^{target} = 0. \tag{2.26}$$

The volumetric rate control is implemented through the volumetric rate operator ζ_j^{vol}:

$$Ts\,\zeta_j^{vol}(\omega)\Delta p - Q^{target} = 0, \tag{2.27}$$

where

$$\zeta_j^{vol} = \frac{\hat{s}_j(\omega)\sum_c \beta_{cj}(\omega)}{\hat{\rho}_t(\omega)}, \tag{2.28}$$

Q^{target} is the target volumetric flow rate at separator conditions [m^3/day], $\hat{s}_j$ and $\hat{\rho}_t(\omega)$ are the saturation and total fluid density respectively at separator conditions, and Ts is the well segment transmissibility between the wellhead and well body.

The aforementioned BHP and volumetric rate control can be coupled with thermal well controls. Specifically, the temperature well control in DARTS is written as:

$$T(\omega) - T^{\text{target}} = 0, \tag{2.29}$$

where T^{target} is the target temperature of injected fluid, $T(\omega)$ is dependent on the thermodynamic state ω.

The well control is a type of special operator that can be customized based on the production (or injection) of interests. In the inverse modeling problem, the adjoint method requires the derivatives of these well controls with respect to unknown variables. Since these well controls are now defined in the form of operators, the assembly of their derivatives in the adjoint method can also take advantage of the high efficiency of the OBL approach. More details can be found in the Section 4.2.1.

2.4. LINEAR SOLVER

In reservoir numerical simulation, the Jacobian matrix frequently exhibits a large number of zero entries due to the limited connectivity of the target cell with other cells in the domain. This sparsity property makes the Jacobian matrix a type of sparse matrix. To effectively store the sparse matrix in a contiguous manner in computer memory and to achieve high memory utilization and fast data transfer speed between memory and CPU, Compressed Sparse Row (CSR) format [89] has been designed.

To further enhance memory access in DARTS, an advanced Compressed Sparse Row (CSR) format known as Block Compressed Sparse Row (BCSR) is employed. Unlike CSR, BCSR stores a series of data in a block-wise manner instead of a single value entry. Each block has dimensions of $n \times n$, where n represents the number of unknown variables for a given cell. Additionally, BCSR enables block-wise sparse matrix operations, such as matrix-vector multiplication, which leverages modern CPU architectures to provide more efficient usage of register and cache storage [51]. The use of the BCSR format in DARTS enables the Jacobian assembly procedure to be more generalizable regardless of the type of mesh grid (e.g. structured or unstructured grid) employed for discretizing the domain.

Then, selecting an appropriate linear solver to solve these linear equations is vital, considering that the linear solution occupies a significant portion of the CPU time for both forward and inverse models. Preconditioning techniques are frequently employed in the linear system to enhance the convergence rate of the iterative solver considerably. Examples of such techniques include the Algebraic Multigrid Method (AMG) and Incomplete LU factorization.

The Constrained Pressure Residual (CPR) approach was specifically developed to resolve reservoir simulation equations, utilizing a two-stage preconditioning technique [110, 111]. Firstly, the pressure system is decoupled from the entire system and solved independently using an AMG-based scheme. Usually, a single V-cycle is enough for efficient preconditioning at the first stage. Secondly, the full system undergoes processing

via an ILU(0) preconditioner, employing the pressure solution obtained from the initial stage. This strategy has demonstrated remarkable robustness and efficiency, particularly in scenarios involving highly heterogeneous reservoirs and large time steps.

A brief description of CPR preconditioner is provided following the notations from [32]. Let A represent the coefficient matrix of the linear equation system $Ax = b$ for the reservoir:

$$A = \begin{bmatrix} A_{ss} & A_{sp} \\ A_{ps} & A_{pp} \end{bmatrix}, \tag{2.30}$$

where subscript p refers to pressure and subscript s to the remaining unknowns, x is the unknown variables, and b is the residuals. To solve this equation system, the pressure reduction is performed using the matrix R:

$$R = \begin{bmatrix} I & 0 \\ -A_{ps}{A_{ss}}^{-1} & I \end{bmatrix}, \tag{2.31}$$

then

$$\tilde{A}x = \tilde{b}, \tag{2.32}$$

where $\tilde{A} = RA$ and $\tilde{b} = Rb$. The pressure matrix is

$$A_p = C^T \tilde{A} C, \tag{2.33}$$

where C is the $(n_{cell} \times n_{eqn}) \times n_{cell}$ pressure-prolongation matrix, n_{cell} is the number of cells, n_{eqn} is the number of equations per cell. The CPR preconditioning matrix for $\tilde{A}$ is given by [110]:

$$M_{CPR}^{-1} = \tilde{M}^{-1}\left(I - \tilde{A}CA_p^{-1}C^T\right) + CA_p^{-1}C^T, \tag{2.34}$$

where $\tilde{M}$ is an approximate factorization of $\tilde{A}$.

The linear system in DARTS is effectively solved by employing the Flexible Generalized Minimum Residual (FGMRES) iterative method [88]. To enhance computational efficiency, a two-stage CPR preconditioner strategy is utilized. This strategy involves the decoupling of the pressure system from the Fully Implicit Method (FIM) system through a True-IMPES reduction approach, directly implemented using the BCSR storage format. Subsequently, a single V-cycle of AMG solver is executed to obtain an approximate solution for the pressure. The resulting approximation is then substituted back into the complete system, and the block ILU(0) preconditioner is applied to further improve convergence.

3

INVERSE MODELING AND ADJOINT GRADIENT METHOD

Summary

This chapter shifts focus from the forward simulation detailed in Chapter 2 to the discussion of the inverse modeling problem, which is the main focus of this dissertation. We will start with the definition of the objective function used in geo-energy problems and the optimization theory. Then the adjoint method is explained in detail. Combined with the architecture of DARTS and a one-dimensional model example, the associated derivatives related to the adjoint method are explicitly explained in the forms of matrices and vectors. Moreover, the selection of the model parameters and the types of observations are also discussed.

Parts of this chapter have been published in SPE Journal, (2021) [100] and in Geothermics (2023) [101].

3.1. INVERSE MODELING PROBLEMS

The "inverse problem" can be considered as the opposite of the "forward problem" as it is described in Chapter 2. In the forward simulation problem, the main goal is to compute the values of the unknown variables by solving the governing equation system. Those unknown variables are pressure, composition, temperature, etc. In a real geo-energy problem, these unknown variables from the numerical model correspond to the observations from the field. Assuming that the model parameters (e.g. porosity, permeability, capillary pressure, and so on) are available, the observations can be denoted as:

$$\mathrm{d}_{\mathrm{obs}} = G(u) + \epsilon, \tag{3.1}$$

where $\mathrm{d}_{\mathrm{obs}}$ is the observation, $G(u)$ is the model response, u is the model variable, and ϵ is the error.

However, the model parameters obtained from the field are usually unreliable or insufficient to describe the reservoir system correctly. The "inverse problem" is to solve Equation (3.1) for the model variable u, so that the obtained u can be used to determine this reservoir system and forecast the production.

3.1.1. OPTIMIZATION THEORY

Consider an unconstrained optimization problem (minimization problem for history matching)

$$\min_{u} J(u), \tag{3.2}$$

where J is the objective function to be optimized. The first-order sufficient and necessary optimality condition is given by:

$$\left.\frac{\partial J}{\partial u}\right|_{u=u^0} = 0 \tag{3.3}$$

where u^0 is the optimal model parameter where the optimal point is located. The sign of the eigenvalues of the second-order derivative matrix of J can be used to determine whether the optimum is minimum (all eigenvalues are larger than zero), maximum (all eigenvalues are smaller than zero), or saddle point (some eigenvalues are larger while some are smaller than zero). In this study, we will use the words "minimize", "minimum", or the associated phrases unless otherwise specified.

However, the history-matching problem of reservoir engineering is usually constrained by the governing equation and well controls. When solving Equation (3.2), it is not easy to incorporate the constraints at the same time, as the governing equation of the system is quite complex. In this case, the Lagrangian multiplier method can be introduced to solve the constrained problem. The modified objective function, therefore, can be written as:

$$\bar{\mathscr{J}}(u,\lambda) = J(u) + \lambda^{\mathrm{T}} g(u). \tag{3.4}$$

λ^{T} is the transposed form of Lagrangian multiplier. $g(u)$ represents the residual form of the governing equation. In most commercial or open-source simulators, the well controls are already incorporated into the linearized governing equations g. Therefore, the constraints of well controls are not explicitly shown in Equation (3.4).

Note that once the governing equation g is satisfied in the forward simulation, the term $\lambda^T g(u)$ is equal to zero. This means that $\tilde{\mathcal{J}}$ holds the same shape as J. In other words, $\tilde{\mathcal{J}}$ does not change the optimal point of J. More details about the Lagrangian multiplier will be presented in Section 3.2. Now, Equation (3.4) becomes an unconstrained optimization problem so that the first-order optimality condition can be used to compute the gradient if the gradient-based method is used to search for the optimum.

Once the objective function is defined, it is essential to find an appropriate optimization method to solve the optimization problem. Generally, the optimization methods can be classified into two types based on the presence of the gradient in the course of history matching. They are called gradient-based and gradient-free methods. For the highly nonlinear problem with multiple local minima, it is natural to choose the gradient-free method. Because this method has a higher possibility to help the optimizer randomly "jump" out of the local minima, therefore, try to find the global minimum. However, this method suffers from low efficiency in searching for the minimum. The gradient-based method features high efficiency in determining the "steepest" descending direction of the objective function for the optimizer. Although this method has the tendency to find the local minimum, this can be mitigated or avoided by starting from different initial guesses or perturbing the gradient direction. This study will use the gradient-based method to conduct the history matching.

3.1.2. MODEL PARAMETERS

It is essential to choose appropriate parameters as the model parameters to conduct the history matching. Intuitively, one may think that all parameters should be modified simultaneously in order to match the system's "history". In reality, however, it is impractical and unnecessary to handle all of the parameters in the course of history matching. This is because some parameters are not very sensitive to the model response, or some parameters are considered to be highly reliable to be directly applied in the system. Another reason is that the limited computational capacity may not allow the processing of large amounts of model parameters in history-matching problems.

It is also worth mentioning that the magnitude of different types of model parameters can be very different. For example, the values of the permeability can be hundreds to thousands mili-Darcy, while the porosity is bounded in the range of $(0, 1)$. This may result in very bad performances of the optimizer. Generally, the optimization algorithms are sensitive to the scale of the problem. Therefore, it is critical to scale different types of model parameters in similar and reasonable ranges of values. There are no standard rules to define what is the "reasonable" value range of the model parameters. A good rule of thumb is to scale them in (or around) the range of $(0, 1)$, as this range is usually suitable for the existing code packages of optimization algorithms.

In this study, transmissibility, well index, and rock-fluid interaction parameters are selected as the model parameters for the history matching and production forecast. It should be noted that not all types of model parameters need to be included in a given inverse modeling problem. The choice of model parameters can be based on the specific problem being addressed. Scaling strategies are employed for these parameters to suit specific problems.

TRANSMISSIBILITY

In reservoir simulation, an unstructured grid is commonly represented by a spatial connectivity graph known as a connection graph [62]. This involves defining the connections between gridblocks and their associated transmissibilities. The transmissibility $\bar{T}_{ij}$ between gridblocks i and j can be expressed for a general unstructured grid [49] as:

$$\bar{T}_{ij} = \frac{\bar{\alpha}_i \bar{\alpha}_j}{\bar{\alpha}_i + \bar{\alpha}_j}, \bar{\alpha}_i = \frac{\bar{A} k_i}{\bar{D}_i}. \tag{3.5}$$

Here, $\bar{A}$ represents the interface area between two gridblocks, $\bar{D}_i$ is the distance from the pressure node to the interface, calculated along the line connecting two pressure nodes, and k_i is the permeability of the gridlock. The transmissibility has a direct impact on the flow dynamics within the reservoir. Note that the transmissibility is a linear parameter with respect to the flow rate, but it also exhibits significant nonlinearity with respect to the objective function. As the initial guess, transmissibilities can be defined based on a two-point flux approximation approach, utilizing initial data on reservoir properties (such as rock permeability).

WELL INDEX

In reservoir simulation, a well index was introduced by [15] to connect the pressure/rate of a gridblock to the wellbore flowing pressure/rate. The equation that links a well and a reservoir gridblock under the assumption of single-phase flow can be expressed as follows:

$$q = \frac{WI_i}{\mu}(p_i - BHP) \tag{3.6}$$

where q is the well rate, WI_i is the well index of the grid block i intersected with the well, p_i is the pressure within the grid block i, and BHP is the well bottom hole pressure.

Once the well index is determined under single-phase assumptions, it can also be extended to accommodate the multiphase flow. Instead of relying on an arbitrary initial guess for the well index during the process of history matching, the well index is calculated using the equation proposed by Peaceman [82] (Equation (3.7)) and applied as the initial guess. While the well index values determined through the Peaceman equation may not be entirely accurate for unstructured grids, they can be refined and updated as the history matching progresses.

$$WI_i = \frac{2\pi\Delta z\sqrt{k_x k_y}}{\ln r_o/r_w + S}, \tag{3.7}$$

where

$$r_o = 0.28\frac{\left[\left(k_y/k_x\right)^{1/2}\Delta x^2 + \left(k_x/k_y\right)^{1/2}\Delta y^2\right]^{1/2}}{\left(k_y/k_x\right)^{1/4} + \left(k_x/k_y\right)^{1/4}}, \tag{3.8}$$

r_w represents the radius of the well, S represents the skin factor, k_x and k_y respectively denote the permeability in the x and y directions, and Δx and Δy respectively signify the block size in the x and y directions. In a manner similar to transmissibility, a well index can be considered as a representation of the linear relationship between the pressure difference and the flow rate between a wellbore and a reservoir block.

ROCK-FLUID INTERACTION PARAMETERS

In multi-phase problems, relative permeability is often utilized as an indicator to describe the flow capacity in porous media. The relative permeability is dependent on various factors such as phase saturation and rock wettability. In combination with the phase density and viscosity, these parameters can be consolidated as rock-fluid interaction parameters for the purpose of history matching. Taking the oil/water system as an example, the modified version of the Brooks-Corey model [6] is often utilized to depict the relative permeability of water and oil under varying saturation conditions. The modified Brooks-Corey model can be represented as follows:

$$k_{ro} = k^e_{ro}\left(1 - S^*_w\right)^{n_o}, \tag{3.9}$$

$$k_{rw} = k^e_{rw}\left(S^*_w\right)^{n_w}, \tag{3.10}$$

$$S^*_w = \frac{S_w - S_{wc}}{1 - S_{wc} - S_{or}}. \tag{3.11}$$

where S^*_w represents the normalized or effective water saturation, k_{rw} refers to the water relative permeability, k^e_{rw} and k^e_{ro} denote the endpoint water and oil relative permeabilities, respectively, n_w and n_o are the exponents for water and oil, respectively, S_w represents the water saturation, S_{wc} stands for the residual or connate water saturation, and S_{or} represents the residual oil saturation.

The relative permeability is incorporated into the β_{cj} operator (as seen in Equation (2.15)), where it is multiplied by $\frac{\rho_j}{\mu_j}$. Therefore, we can define six rock-fluid interaction parameters:

$$P_{fr} = \{S_{or},\ S_{wc},\ n_o,\ n_w,\ k^e_{rw}\frac{\rho_w}{\mu_w},\ k^e_{ro}\frac{\rho_o}{\mu_o}\}. \tag{3.12}$$

The entire reservoir can be divided into multiple flow regions, where it is assumed that within each region, fluid flow is governed by the same set of rock-fluid interaction parameters. As a result, there may be multiple sets of rock-fluid interaction parameters in a given inverse modeling problem. For other fluid flow systems, such as a water/oil/gas system, the rock-fluid interaction parameters can be defined in a similar manner as depicted in Equation (3.12).

3.2. ADJOINT METHOD

As discussed in Section 3.1.1, the utilization of gradient-free methods can prove to be highly ineffective when applied to practical inverse modeling problems. As a result, in this study, we will only employ the gradient-based method for the purpose of inverse modeling. Instead of utilizing the conventional numerical gradient-based method for inverse modeling, the adjoint method is applied in order to improve the efficiency of gradient calculation. In this section, we first provide an overview of the general form of the adjoint method. Then, we demonstrate its application through a 1D model example to clarify the underlying derivative terms involved in the adjoint method.

3.2.1. FORMULATION OF ADJOINT METHOD

In the adjoint method, rather than directly computing the first-order derivatives of the objective function, we combine it with the governing equations of the reservoir and then

calculate the derivatives. The resulting augmented objective function is as follows:

$$\bar{\mathscr{J}}(\omega, u, \lambda) = J(\omega, u) + \lambda^{\mathrm{T}} g(\omega, u). \tag{3.13}$$

Note that we have previously presented the concept of combining the original objective function with constraints, as outlined in Equation (3.4). Apparently, the constraint in this scenario is represented by the governing equation g. Here in this equation, the state variable ω is explicitly included, while the well control notation (e.g. ζ in Equation (2.28)) is omitted for the purposes of simplicity. As per the first-order optimality conditions shown in Equation (3.3), the following equation must be satisfied:

$$\bar{\mathscr{J}}_{\lambda} = g(\omega, u) = 0, \tag{3.14}$$

$$\bar{\mathscr{J}}_{\omega} = \lambda^{\mathrm{T}} g_{\omega}(\omega, u) + J_{\omega}(\omega, u) = 0, \tag{3.15}$$

$$\bar{\mathscr{J}}_{u} = \lambda^{\mathrm{T}} g_{u}(\omega, u) + J_{u}(\omega, u) = 0. \tag{3.16}$$

where the subscripts λ, ω, and u denote the derivatives with respect to the respective parameters. It is evident that Equation (3.14) is satisfied as it is the residual form of the governing equation and is solved during the forward simulation. The newly introduced λ^{T} is computed from Equation (3.15). Finally, $\mathscr{J}_{\boldsymbol{u}}$ is constructed and can be employed in the gradient-based method.

In numerical simulations, the governing equations are discretized both in space and time. To apply the adjoint method, Equation (3.14), Equation (3.15), and Equation (3.16) must also be transformed into their discretized forms accordingly. The deduction procedure will begin with the discretized form of the governing equation. At the k^{th} time step, the governing equation is expressed as:

$$g_k(\omega, \omega_{(k-1)}, u) = 0, \tag{3.17}$$

where $\omega_{(k-1)}$ represents the state variables at previous time step. The discretized objective function J can be formulated as::

$$J = \sum_{k=1}^{\mathrm{K}} L_k, \tag{3.18}$$

where K represents the total number of simulation time steps. The misfit between the model response and the observation data at a given simulation time step k is represented by L_k. This misfit term can be defined as follows:

$$L_k = (f_1 + f_2 + f_3 + f_4)\delta_t(\tau_{\mathrm{obs}}), \tag{3.19}$$

where $\delta_t(\tau_{\mathrm{obs}})$ is the Dirac measure function, which can be defined as:

$$\delta_t(\tau_{\mathrm{obs}}) = \begin{cases} 1 & \text{if } \mathrm{t} \in \tau_{\mathrm{obs}} \\ 0 & \text{if } \mathrm{t} \notin \tau_{\mathrm{obs}} \end{cases}, \tag{3.20}$$

where t is the time point at the endpoint of a given simulation time interval, τ_{obs} is a set of the observation time points, which means it is a subset of the simulation time points.

The expressions for the variables f_1, f_2, f_3, and f_4 are as follows:

$$\begin{aligned} f_1 &= \sum_{w=1}^{\mathrm{n_w}} \sum_{j=1}^{\mathrm{n_p}} C_{w,j}^{rate} \left(q_{w,j} - q_{w,j}^*\right)^2 \\ f_2 &= \sum_{w=1}^{\mathrm{n_w}} C_w^{BHP} \left(p_w - p_w^*\right)^2 \\ f_3 &= \sum_{w=1}^{\mathrm{n_w}} C_w^{T_{well}} \left(T_w - T_w^*\right)^2 \\ f_4 &= \sum_{m=1}^{\mathrm{n_{block}}} C_m^{\Psi_{res}} \left(\Psi_m - \Psi_m^*\right)^2, \end{aligned} \tag{3.21}$$

f_1, f_2, and f_3 denote the misfit between the model response and the observation data obtained from the wells. These terms correspond to the well phase flow rate, well bottom hole pressure (BHP), and well temperature, respectively. f_4 represents the misfit between the modeled response of the reservoir block time-lapse data and the effective reservoir block time-lapse observations interpreted from various geophysical data sources, such as electromagnetic (EM) data or seismic data. $\mathrm{n_w}$, $\mathrm{n_p}$, and $\mathrm{n_{block}}$ are the number of wells, phases, and reservoir blocks, respectively. Here $q_{w,j}$ is the model response of the rate of well w and phases j, p_w is the BHP of the well w, T_w is the temperature of well w. The term Ψ_m is the time-lapse data of reservoir block m. The superscript $*$ denotes the corresponding observation data of the terms in Equation (3.21). The notation C with its corresponding subscripts and superscripts defines the inverse of the diagonal covariance matrices for the measurement errors.

The discretized form of equation Equation (3.15) can now be derived, given the discretized forms of g_k and L_k, as follows:

$$\lambda_{k+1}^{\mathrm{T}} \frac{\partial g_{k+1}}{\partial \omega_k} + \lambda_k^{\mathrm{T}} \frac{\partial g_k}{\partial \omega_k} + \frac{\partial L_k}{\partial \omega_k} = 0, \tag{3.22}$$

$$\lambda_{\mathrm{K}}^{\mathrm{T}} \frac{\partial g_{\mathrm{K}}}{\partial \omega_{\mathrm{K}}} + \frac{\partial L_{\mathrm{K}}}{\partial \omega_{\mathrm{K}}} = 0. \tag{3.23}$$

The discretized expression for equation Equation (3.16) is as follows:

$$\bar{\mathscr{J}}_u = \sum_{k=1}^{\mathrm{K}} \left(\lambda_k^{\mathrm{T}} \frac{\partial g_k}{\partial u_k} + \frac{\partial L_k}{\partial u_k} \right). \tag{3.24}$$

Based on the well treatment outlined Section 2.3 and the time-lapse data, all the misfit terms in equation Equation (3.21) are independent of the model parameter u (i.e. transmissibility and well index). This implies that the derivative $\frac{\partial L_k}{\partial u_k}$ is a zero vector. The remaining derivatives shown in equations Equation (3.22), Equation (3.23), and Equation (3.24) will be derived in Section 4.2.1.

3.3. LINEAR SOLVER FOR ADJOINT METHOD

One of the crucial procedures in the evaluation of the adjoint-based gradient involves the solution of the Lagrangian multipliers, denoted by λ, as presented in Equations (3.22)

and (3.23). These two equations entail a sequence of linear systems of equations at each time step, which can be denoted as:

$$\bar{\bar{J}}^T \lambda = d \tag{3.25}$$

where $\bar{\bar{J}}^T$ is the transposed Jacobian collected from forward simulation, and

$$d = \begin{cases} -\left[\frac{\partial L_k}{\partial \omega_k} + \lambda_{k+1}^{\mathrm{T}} \frac{\partial g_{k+1}}{\partial \omega_k}\right]^T & \text{the rest of time steps k} \\ -\left[\frac{\partial L_{\mathrm{K}}}{\partial \omega_{\mathrm{K}}}\right]^T & \text{final time step K} \end{cases}. \tag{3.26}$$

However, the preconditioners such as CPR, which exploits the properties of the Jacobian of forward simulation to accelerate convergence, are not directly applicable to the linear equation system in the adjoint gradient method. This is due to the "transposed" Jacobian matrix in Equation (3.25), which necessitates significant modifications to the preconditioners [91]. In [32], an adaptation of the CPR preconditioner, referred to as CPRA, was proposed for the adjoint equation, which is also a linear equation system (i.e., Equation (3.25)). Following the notations from Equation (2.30) to Equation (2.34), and denoting the $A = \bar{\bar{J}}$ and $\tilde{A} = RA$, this linear system is written as:

$$(\tilde{A})^T y = d, \tag{3.27}$$

where $y = (R^T)^{-1}\lambda$. So the CPRA preconditioner M_{CPRA}^{-1} can be derived from M_{CPR}^{-1}:

$$M_{CPRA}^{-1} = M_{CPR}^{-T} = \left[I - C\left(A_p^T\right)^{-1} C^T (\tilde{A})^T\right] (\hat{M})^{-1} + C\left(A_p^T\right)^{-1} C^T, \tag{3.28}$$

where $\hat{M}$ is an approximate factorization of $(\tilde{A})^T$. By re-arranging the above equation, we obtain

$$M_{CPRA}^{-1} = C\left(A_p^T\right)^{-1} C^T \left[I - (\tilde{A})^T (\hat{M})^{-1}\right] + (\hat{M})^{-1}. \tag{3.29}$$

By utilizing the factorization M_{CPRA}^{-1}, the incomplete LU and AMG preconditioner, in conjunction with the GMRES algorithm, can be utilized to solve the linear equation system in the adjoint method. The results presented in [32] demonstrate that the CPRA/AMG solver takes roughly the same amount of time as the CPR/AMG forward simulations, highlighting the high efficiency of the CPRA preconditioner. Integrating this preconditioner into DARTS to solve the linear equation system in the adjoint method could be the future work within the inverse modeling framework of DARTS.

In the current version of the DARTS framework, a direct solver is utilized to resolve the linear system described in Equation (3.25). The design of the CPRA precondition would be the future work to be incorporated into this inverse modeling framework.

4

IMPLEMENTATION AND VALIDATION OF ADJOINT METHOD IN DARTS

Summary

The development of an inverse modeling prototype based on the adjoint method was initially implemented in MATLAB. A two-parameter inverse modeling problem was validated and compared with numerical methods. Subsequently, the framework was translated into C++ to increase efficiency and scalability. Two types of engines, the "super engine" and "thermal engine," were considered, and the adjoint method was designed to address the inverse modeling problems related to these engines. A comparative analysis was conducted between the adjoint-based gradient and the numerical gradient, and the results showed that the angle between the two gradients was negligible. Additionally, the CPU time required for the adjoint method was significantly lower than that of the numerical method. The gradient evaluation using the adjoint method yielded a two to three orders-of-magnitude improvement in efficiency. This efficiency improvement was found to be even larger for models with higher degrees of freedom.

Parts of this chapter have been published in SPE Journal, (2021) [100].

4.1. ONE-DIMENSIONAL EXAMPLE IN MATLAB

In contrast to the conventional numerical method of gradient evaluation, which can be conveniently added to existing simulators, the implementation of the adjoint method requires access to the simulator source code and substantial implementation efforts. Nevertheless, despite these implementation difficulties, the adjoint method has been demonstrated to be significantly more efficient in terms of performance. Thus, it is worth implementing the adjoint method to improve the performance of the inverse modeling feature of DARTS.

The first prototype of DARTS was implemented in MATLAB as a standalone simulator [51]. This prototype was designed to estimate the performance advantage of the OBL approach. The OBL was implemented in the AD-GPRS framework as a nonlinear formulation. First, *MolarVariableFormulation* of AD-GPRS was modified, as the DARTS framework is designed based on the overall-composition variable. The *fluidPropertiesCalculation* and *computeMassFluxTerm* functions were altered to apply OBL to residual equation computations. The evaluation of state-dependent operators was performed through multilinear interpolation procedures, replacing the ADETL-based computations. The state-dependent operators were interpolated adaptively and stored individually.

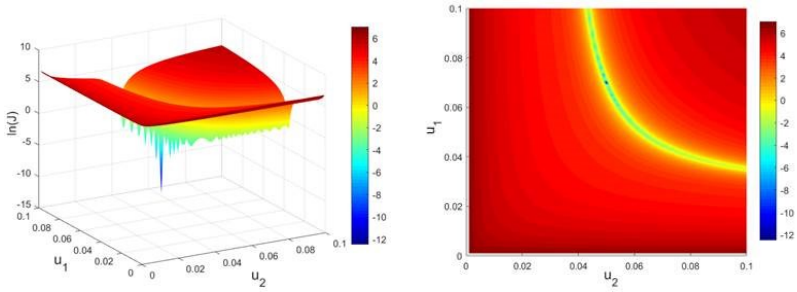
Accordingly, in this study, the inverse modeling feature of DARTS utilizing the adjoint method was initially implemented in MATLAB. In this implementation, the operator values are generated by AD-GPRS and provided to MATLAB via text files. The adjoint method is validated through the examination of a one-dimensional reservoir model, which is discretized into 40 blocks with Dirichlet boundary conditions on both sides. The reservoir is comprised of three components and is considered to have two phases in its system. The initial pressure in the reservoir is set to 85 bar and a water injection with a pressure of 100 bar is performed through the injection well. The model runs for 50 days. The objective function for the first phase of the producer is defined as:

$$J = \sum_{k=1}^{\mathrm{K}} \left(q_k - q_k^*\right)^2 \tag{4.1}$$

where k is the index of the simulation time step, K is the total number of the simulation time steps, q is the model response of phase rate, and q^* is the observation of phase rate.

For the purpose of visualization, this reservoir model is limited to two permeability regions. This is because the limited dimensionality of the model parameters, with only two dimensions, enables straightforward representation on graphical displays. However, it should be noted that this framework is capable of accommodating a greater number of model parameters. Figure 4.1 shows the plot of the objective function value (in logarithmic scale) as a function of two model parameters applied in the model. The model parameters u_1 and u_2 correspond to the transmissibility of the two permeability regions described previously.

As it can be seen from Figure 4.1, a distinctive banded area is present, which corresponds to the global minimum of the objective function. Additionally, there is a cluster of local extrema located in proximity to the global minimum. This observation highlights that even in inverse modeling problems with a limited dimensionality of model parameters, the presence of multiple local extrema can pose a challenge for the optimizer in its



sideview topview

Figure 4.1: The logarithmic values of the objective function in side view (left figure) and top view (right figure)

search for the global minimum. To validate the adjoint method, the results of adjoint-based gradients and numerical gradients are compared. A graphical representation of the reduction in the objective function value as a function of the iteration number is also provided, as shown in Figure 4.2. Both methods behave identically and require a significant amount of computational effort in the banded area containing multiple local extrema before converging to the global minimum.

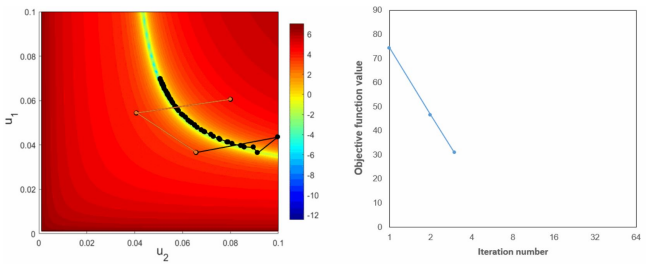
4.2. IMPLEMENTATION IN C++/PYTHON

To enhance efficiency and facilitate modularity and scalability, DARTS was later integrated as a standalone simulator in both C++/Python. C++ is utilized in DARTS for performance-critical parts such as linearization and solution of a linear system. Nonetheless, end-users are not required to work in the C++ environment as the DARTS simulation can be accessed via its Python API, offering a more user-friendly and flexible interface [51]. Similarly, the adjoint method implementation in DARTS is divided between C++ and Python. The computation of matrices and derivatives, solution of linear equations, and model response assembly are performed in C++, while pre- and post-processing tasks, such as optimization configuration and observation preparation, are implemented in Python.

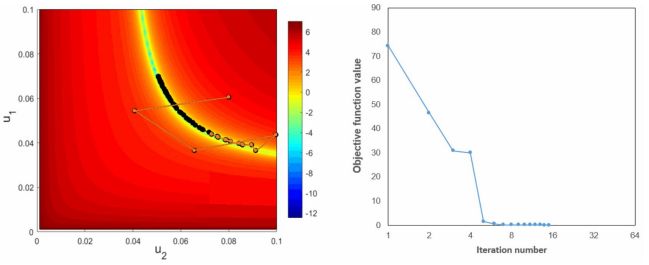
Initially, DARTS featured a suite of engines specifically designed to address a range of geo-energy problems. Recently, the framework underwent significant refactoring to enhance its generality and accommodate the more complex physics involved in geo-energy problems. The current version of DARTS consists of a "super engine", capable of addressing a broad range of problems, and a "thermal engine", dedicated to geothermal problems. The next few sections will provide detailed explanations of the adjoint method implementation.

4.2.1. ASSEMBLY OF DERIVATIVES IN ADJOINT METHOD

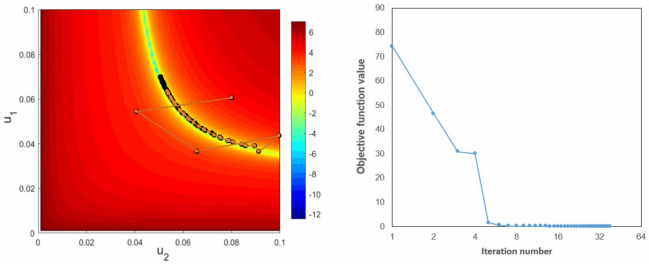
This section takes a 1D reservoir model as an example to represent the typical structures of the associated matrices and vectors in the adjoint method following the Operator-Based Linearization framework. The relevant C++ variables of DARTS can be found in Section 4.2.2.



(a) The 3rd iteration



(b) The 15th time step



(c) The 40th time step

Figure 4.2: Iteration plot (the second column) and optimization trajectories (the first column) with black and orange lines corresponding to numerical and adjoint-based gradients respectively at different iteration steps

THE DERIVATIVES RELATED TO GOVERNING EQUATION

The partial derivatives $\frac{\partial g_{k+1}}{\partial \omega_k}$ and $\frac{\partial g_k}{\partial \omega_k}$, as represented by Equation (3.22) and Equation (3.23), are relevant to the governing equation g. The derivative $\frac{\partial g_k}{\partial \omega_k}$ is specifically the Jacobian matrix associated with the governing equation. Therefore, it can be conveniently obtained through the forward simulation. The term $\frac{\partial g_{k+1}}{\partial \omega_k}$ is the derivative of the governing equation g with respect to the state variable ω at the "previous" time step (i.e. the k^{th} time step). To provide a clearer explanation of these derivatives, we present a matrix formulation of them using a 1D reservoir model as an example.

Figure 4.3 illustrates a 1D reservoir model, displaying the index numbers of the associated state variables ω for each block and the transmissibility Tb (or well index WI) of

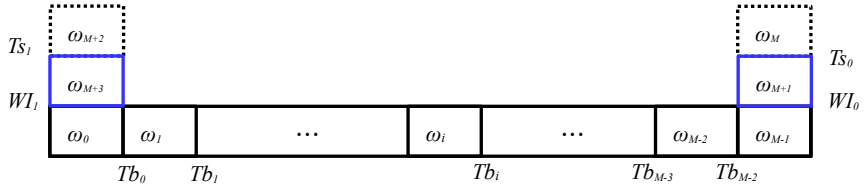


Figure 4.3: The 1D reservoir model with multiple reservoir blocks and two wells (W_0 and W_1) at the boundaries.

each block interface. *Tb* and *WI* are designated as the model parameters in this study. The blocks with solid line boundaries represent the reservoir blocks, with the blue blocks being the well bodies. The blocks with dashed line boundaries are the well heads, also referred to as ghost well blocks. It is worth noting that the index number of the well body is always greater than that of the well head within a single well. The well segment transmissibility between the well head and well body is denoted as *Ts*.

Based on the OBL method and Equation (2.13), the discretized governing equation in residual form for the block i at the k^{th} time step can be written as (here we omit the component index c and phase index j):

$$g_i = V_i\phi_i\left(\alpha_i(\omega) - \alpha_i\left(\omega_{(k-1)}\right)\right) - \bar{T}_i\Delta t\beta_i(\omega)\left(p_{i+1} - p_i\right) - \bar{T}_{i-1}\Delta t\beta_{i-1}(\omega)\left(p_{i-1} - p_i\right). \tag{4.2}$$

To simplify the expression of equations, the reservoir transmissibility *Tb*, well index *WI*, and well segment transmissibility *Ts* are grouped into a single notation $\bar{T}$ shown in Equation (4.2), as they all characterize the fluid flow capability of the block interface. Similarly, a single notation β is used to denote both the volumetric well rate operator ζ_j^{vol} from Equation (2.28) and the mass flux operator β_{cj} from Equation (2.15).

Following the notation in Equation (4.2), the term $\frac{\partial g_{k+1}}{\partial \omega_k}$ of Equation (3.22) in matrix form can be written as:

$$\frac{\partial g_{k+1}}{\partial \omega_k} = \begin{bmatrix} -V_0\phi_0 A_0 & & & & & & & & \\ & \ddots & & & & & & & \\ & & -V_i\phi_i A_i & & & & & & \\ & & & \ddots & & & & & \\ & & & & -V_{M-1}\phi_{M-1}A_{M-1} & & & & \\ & & & & & 0 & & & \\ & & & & & & -V_{M+1}\phi_{M+1}A_{M+1} & & \\ & & & & & & & 0 & \\ & & & & & & & & -V_{M+3}\phi_{M+3}A_{M+3} \end{bmatrix}, \tag{4.3}$$

where

$$A_i = \frac{\partial \alpha_i}{\partial \omega_i}. \tag{4.4}$$

As noted from Equation (4.2), the term $\alpha_i\left(\omega_{(k-1)}\right)$ is the only component in the equation that depends on the state variable at the previous time step. This is why the derivatives are located along the diagonal of Equation (4.3). The well head serves as the location for

well controls and, therefore, does not include any accumulation terms in its block. As a result, the derivatives at the well head index are 0. The derivatives of the well body are denoted in blue text in Equation (4.3).

Another derivative associated with the governing equation is the derivative of g_k with respect to u_k (or $\bar{T}_k$ in this 1D example model), which is represented in Equation (3.24) as:

$$\frac{\partial g_k}{\partial \bar{T}_k} = \Delta t \begin{bmatrix} -\beta_0(p_1 - p_0) & & & & & & & & & -\beta_{M+3}(p_{M+3} - p_0) \\ -\beta_0(p_0 - p_1) & -\beta_1(p_2 - p_1) & & & & & & & & \\ & \ddots & \ddots & & & & & & & \\ & & -\beta_{i-1}(p_{i-1} - p_i) & -\beta_i(p_{i+1} - p_i) & & & & & & \\ & & & \ddots & \ddots & & & & & \\ & & & & -\beta_{M-3}(p_{M-3} - p_{M-2}) & -\beta_{M-2}(p_{M-1} - p_{M-2}) & & & & \\ & & & & & -\beta_{M-2}(p_{M-2} - p_{M-1}) & -\beta_{M-1}(p_{M+1} - p_{M-1}) & & & \\ & & & & & & 0 & & & \\ & & & & & & -\beta_{M-1}(p_{M-1} - p_{M+1}) & & & \\ & & & & & & & 0 & & \\ & & & & & & & & -\beta_{M+3}(p_0 - p_{M+3}) \end{bmatrix} \tag{4.5}$$

In Equation (4.5), the number of columns is equal to the number of transmissibility plus well index. This fact indicates that Equation (4.5) pertains to a non-square matrix. Each column comprises two elements that correspond to the two neighboring blocks at a particular interface. The last two columns display the derivatives with respect to the well index *WI* of the injector and producer. The blue-colored notation represents the pressure of the well body block.

THE DERIVATIVES RELATED TO WELL TREATMENT

Based on the definition of the objective function, which is specified from Equation (3.18) to Equation (3.21), the misfit terms f_1, f_2, and f_3 from Equation (3.21) are related to the observations at wells. Their respective derivatives are defined as follows:

(1) ***The derivatives of well rate misfit*** f_1

In order to account for the various combinations of well types (i.e. injector and producer) and well controls (i.e. BHP control and rate control), a new notation β_{up} is introduced to represent the flux operator at the upstream well block. The upstream well block is determined based on the pressure difference between the well body and well head. For instance, if the pressure of the well head is greater than the pressure of the well body, the well head is considered as the upstream well block. The definition of β_{up} can be found in Equation (2.28). As a result, the value of the i^{th} element in the derivative vector $\frac{\partial L_k}{\partial \omega_k}$ can be expressed as follows:

$$\left[\frac{\partial L_k}{\partial \omega_k}\right]_i = \begin{cases} 0 & \text{if i} \in \text{well head blocks with BHP control} \\ B\beta_{up}Ts & \text{if i} \in \text{well head blocks with rate control} \\ -B\beta_{up}Ts & \text{if i} \in \text{well body blocks} \\ B\frac{\partial \beta_{up}}{\partial \omega}Ts\Delta p & \text{if i} \in \text{upstream well blocks} \\ 0 & \text{otherwise} \end{cases}, \tag{4.6}$$

where

$$B = 2C_{w,j}^{rate}\left(q_{w,j} - q_{w,j}^{*}\right)\delta_t(\tau_{\text{obs}}). \tag{4.7}$$

In this expression, the subscripts w and j are taken from Equation (3.21). The subscripts of B, β_{up}, Ts, ω, and Δp are omitted for brevity. These values can be derived from the state and well control information for the wells.

(2) ***The derivatives of well pressure misfit*** f_2

Since the model response of well pressure in DARTS is directly obtained from the well head block, the expression of $\frac{\partial L_k}{\partial \omega_k}$ for the i^{th} element is given as:

$$\left[\frac{\partial L_k}{\partial \omega_k}\right]_i = \begin{cases} 0 & \text{if i} \in \text{well head blocks with BHP control} \\ 2C_w^{BHP}\left(p_w - p_w^{*}\right)\delta_t(\tau_{\text{obs}}) & \text{if i} \in \text{well head blocks with rate control} \\ 0 & \text{otherwise} \end{cases}, \tag{4.8}$$

where the subscript w is taken from the Equation (3.21).

(3) ***The derivatives of well temperature misfit*** f_3

The model response of well temperature is dependent on the thermodynamic state of the upstream block. Therefore, the expression of $\frac{\partial L_k}{\partial \omega_k}$ at the given i^{th} element reads:

$$\left[\frac{\partial L_k}{\partial \omega_k}\right]_i = \begin{cases} 2C_w^{T_{well}}\left(T_w - T_w^{*}\right)\delta_t(\tau_{\text{obs}})\frac{\partial T(\omega)}{\partial \omega} & \text{if i} \in \text{upstream well blocks} \\ 0 & \text{otherwise} \end{cases}, \tag{4.9}$$

where the subscript w is taken from the Equation (3.21).

THE DERIVATIVES RELATED TO TIME-LAPSE DATA

In the context of reservoir blocks, time-lapse data can be defined as various physical properties such as component concentration, phase saturation, temperature, etc. These values can be derived through flash calculation, which can be readily obtained from OBL. Here a new operator, denoted by Ψ as mentioned in Equation (3.21), is introduced to represent the relationship between the time-lapse data and the state variable ω. The expression for $\frac{\partial L_k}{\partial \omega_k}$ at the i^{th} element can thus be formulated as follows:

$$\left[\frac{\partial L_k}{\partial \omega_k}\right]_i = \begin{cases} 2C_m^{\Psi_{res}}\left(\Psi_m - \Psi_m^{*}\right)\delta_t(\tau_{\text{obs}})\frac{\partial \Psi(\omega)}{\partial \omega} & \text{if i} \in \text{reservoir blocks} \\ 0 & \text{if i} \in \text{well blocks} \end{cases}, \tag{4.10}$$

where the subscript m is taken from the Equation (3.21).

4.2.2. IMPLEMENTATION IN ENGINES OF DARTS

The super engine incorporates the effects of gravity, capillarity, and chemical reactions in its governing equations, and as a result, these effects must be accounted for when constructing the derivatives in the adjoint method. The thermal engine possesses a code structure that is comparable to the super engine. The source code for the adjoint method can be accessed through the Open-DARTS repository [81].

ASSEMBLY OF DERIVATIVE MATRICES

Following the Equation (4.3), the assembly of $\frac{\partial g_{k+1}}{\partial \omega_k}$ requires consideration solely of the accumulation terms of the governing equation, as this derivative is only dependent on the term at the previous time step. The relevant C++ variables of DARTS are shown:

- ***dg_dx_n_temp*** — $\frac{\partial g_{k+1}}{\partial \omega_k}$ in Block Compressed Sparse Row (BCSR) format
- ***PV*** — array of pore volumes
- ***op_ders_arr*** — array of the derivatives of the operator α_{cj}
- ***RV*** — array of initial rock volumes
- ***kin_fac*** — array of kinetic rate constants

Another important derivative matrix to be assembled is $\frac{\partial g_k}{\partial u_k}$. In accordance with Equation (4.5) and taking into account the aforementioned effects, the relevant C++ variables are as follows:

- ***dg_dT_general*** — $\frac{\partial g_k}{\partial u_k}$ in BCSR format
- ***p_diff*** — pressure difference between two neighboring blocks
- ***avg_density*** — averaged phase density between two neighboring blocks
- ***grav_coef*** — gravity coefficient
- ***trans_mult_exp*** — exponent used for transmissibility multiplier
- ***dt*** — simulation time step
- ***op_vals_arr*** — array of the operator β_{cj}, including the term of flux and capillarity

ASSEMBLY OF DERIVATIVE VECTORS

The derivative vector is related to the model response. They consists of various forms of $\frac{\partial L_k}{\partial \omega_k}$ as specified in Equations (4.6) and (4.8) to (4.10). The relevant C++ variables are as follows:

- ***Temp_dj_dx*** — array of $\frac{\partial L_k}{\partial \omega_k}$
- ***rates*** — array of the operator β_{up}
- ***rates_derivs*** — array of $\frac{\partial \beta_{up}}{\partial \omega}$
- ***segment_transmissibility*** — well segment transmissibility Ts
- ***p_diff*** — pressure difference between two neighboring blocks
- ***op_ders_arr_customized*** — array of derivatives of time-lapse operator, i.e. $\frac{\partial \Psi(\omega)}{\partial \omega}$

4.2.3. PYTHON API FOR INVERSE MODELING FEATURE

The pre-processing phase of the inverse modeling procedure in Python, including the assembly of observations, covariance matrices, and scaling factors, must be completed prior to passing the values to the DARTS engine and calling the engine function to compute the adjoint-based gradient. The C++ engine of DARTS offers relevant Python APIs for these purposes. The following are examples of these APIs:

- ***calc_adjoint_gradient_dirac_all*** — function of computing adjoint-based gradient
- ***push_back_to_Q_*** — function of passing the values of observations
- ***push_back_to_cov_*** — function of passing the values of covariance matrices
- ***push_back_to_wei_*** — function of passing the values of weighted scaling factors
- ***dirac_vec*** — variable of Dirac measure function array
- ***derivatives*** — variable of adjoint-based gradient array

4.3. VALIDATION OF ADJOINT METHOD IN DARTS

In this section, the implementation of the inverse modeling feature based on the adjoint method within the DARTS framework is validated through a comparison of gradient evaluation results calculated using the adjoint method and numerical method. The validation is performed using more general and complex cases, such as unstructured grids and multiple realizations, to test the efficacy and efficiency of the framework in these situations.

Five unstructured grid models with varying resolutions (as shown in Figure 4.4) are utilized as the basis for generating ensembles of realizations. For each grid, 100 different realizations are created through the uniform sampling of cell permeabilities ranging from 1 to 100000 mD. The adjoint gradients and numerical gradients are subsequently calculated and compared for all 100 realizations of each grid. The comparison of the two gradient evaluations is performed using the angle between these two gradient vectors, which is computed using the following equation:

$$\theta = \arccos \frac{\boldsymbol{grad}_a \cdot \boldsymbol{grad}_{num}}{\|\boldsymbol{grad}_a\| \cdot \|\boldsymbol{grad}_{num}\|} \tag{4.11}$$

where $\boldsymbol{grad}_a$ and $\boldsymbol{grad}_{num}$ are the vectors of adjoint-based gradient and numerical gradient, respectively.

The mean values of the angle and CPU time for each grid model in the calculation of both adjoint-based and numerical gradients are presented in Table 4.1.

The third column of Table 4.1 displays the degrees of freedom of the model parameters for each grid model, which are comprised of the transmissibility and well index. The mean CPU time of computing the gradients by using the adjoint method ($t_{\boldsymbol{grad}_a}$) and the numerical method ($t_{\boldsymbol{grad}_{num}}$) are presented. As it can be seen from Table 4.1, the mean values of the angle θ for five grids are approximately 3.6°, which is small and

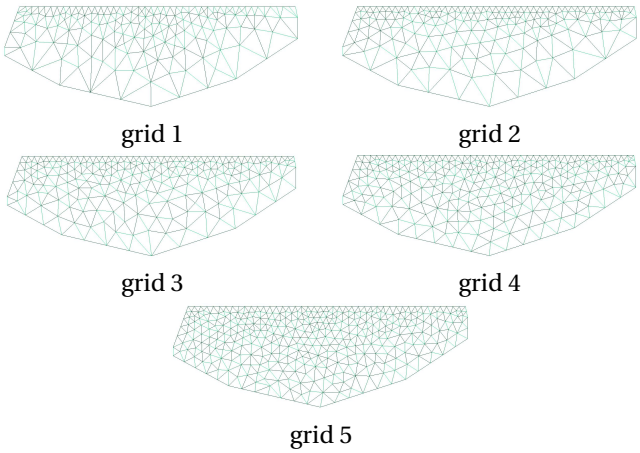


Figure 4.4: Five grid models with different grid resolutions

Table 4.1: Angles between adjoint gradients and numerical gradients and CPU time of computing gradients

Model	Angle θ (°)	DoF	$t_{\boldsymbol{grad}_{num}}$ (sec.)	$t_{\boldsymbol{grad}_a}$ (sec.)	$t_{\boldsymbol{grad}_{num}} / t_{\boldsymbol{grad}_a}$
grid 1	3.49	424	1711.28	3.75	456
grid 2	3.33	478	2214.77	4.23	524
grid 3	3.60	599	3483.75	5.43	642
grid 4	3.24	741	5376.97	6.79	791
grid 5	4.34	983	10312.31	9.69	1064

supports the validity of the adjoint gradient. This also indicates that the adjoint gradient can be utilized to assess the suitability of the small perturbation applied in the calculation of the numerical gradient using the finite difference method. This is because the adjoint method circumvents the need for introducing a small perturbation. The relationship between the number of model parameters and the CPU time, as displayed in Table 4.1, is plotted in Figure 4.5. This figure suggests that with the introduction of more model parameters, the adjoint method presents a higher gradient calculation efficiency in comparison to the numerical method.

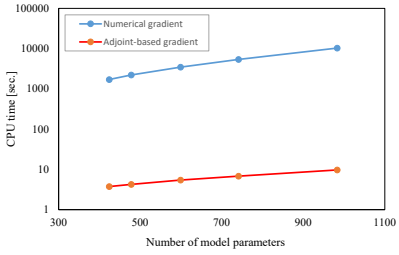


Figure 4.5: The CPU time for computing the adjoint-based gradient and numerical gradient

5

EFFICIENT PROXY MODELS BASED ON ADJOINT METHOD

Summary

The application of the inverse modeling framework of DARTS to the data-driven proxy model method is presented in this chapter. The Discrete Well Affinity (DiWA) data-driven model was initially proposed and validated, followed by the introduction of the adjoint method with the aim of improving the performances of history matching and production forecasting capabilities of the DiWA proxy model. Finally, a stochastic variation of the DiWA proxy model is proposed and applied to real oil reservoir models. The results show that the computational time of training the DiWA model can be largely reduced because of the utilization of the coarse grid and adjoint method. The final trained DiWA model is capable of providing good production forecast results.

Parts of this chapter have been published in SPE Journal, (2021) [100] and in Journal of Petroleum Science and Engineering, (2022) [102].

5.1. DISCRETE WELL AFFINITY (DIWA) DATA-DRIVEN PROXY MODEL

The rapid progression of computer technologies has presented many opportunities for subsurface applications to perform complex numerical simulations utilizing high-resolution three-dimensional geological models. These models are crucial for project management and decision-making of reservoirs, as they illustrate complex geological features through the representation of grid blocks and associated rock and fluid properties. However, the reliability of geological information utilized in the modeling process may sometimes be questionable or unavailable. Despite the possibility of developing a high-fidelity model on a reliable geological foundation, the high resolution of these models can result in a simulation consisting of several million blocks, requiring substantial computational resources and time, often measured in hours or days. The current state of computational feasibility does not permit the conduction of history matching or reservoir development optimization at such a resolution, due to the large number of forward simulation runs required.

The issue of computational feasibility in high-resolution models can be addressed through the development of various methods, which can be categorized into two groups: simplified full-field models, and data-driven approaches. Simplified full-field models include techniques such as upscaling, multi-scale methods, and streamline simulation. Upscaling involves the numerical homogenization of the high-resolution model, representing it as a set of coarser grid blocks with assigned effective properties that replicate the response of the high-fidelity model [20]. Multi-scale methods are designed to efficiently capture the large-scale behavior of the solution while avoiding the resolution of small-scale features [38, 46]. The streamline simulation [4, 16] is an Eulerian-Lagrangian approach with implicit pressure explicit saturation (IMPES) time approximation, which translates a full 3D transport solution into a set of one-dimensional equations that are solved along streamlines.

All methods in the first group, which is based on simplified full-field models, rely on an underlying geological characterization as a foundation for their implementation. However, in many cases, the availability and reliability of such information may be uncertain. In such cases, the utilization of methods from the second group, data-driven approaches, offers a solution. The data-driven method involves the construction of a proxy model with a sufficient number of degrees of freedom to accurately replicate the response of a realistic reservoir through calibration to production data. With the continuous acquisition of accurate data fed into a reliable regression framework, data-driven models can provide an accurate prediction for the given reservoir.

There are many data-driven approaches applied in the industry for reservoir management and decision-making, including the statistical data-driven model proposed by [44]; reduced-order models [10]; the Capacitance Resistance Model [116, 2, 34]; the flow-network model [59, 86, 5, 56], in which a complex 3D flow is represented as a set of 1D finite-difference reservoir models; the Interwell Numerical Simulation Model (INSIM) [119, 120] and INSIM-FT-3D [30] approach, which uses a new Riemann solver based on a convex-hull method to solve the Buckley-Leverett problem with gravity and allows for the inclusion of wells with arbitrary trajectories with multiple perforations; and various

other alternative methods that employ Artificial Intelligence (AI) [72] and data fitting [121]. Each of these approaches has specific advantages and limitations.

In this section, we present the DiWA data-driven proxy model method that blends the benefits of both physical-based and data-driven modeling methods to achieve accurate oil reservoir forecasts based on historical field data while maintaining adherence to fundamental physical processes. This method does not require precise grid parameters or a comprehensive understanding of reservoir geology. We will focus on the general applicability of the DiWA model in simulation scenarios with limited heterogeneities and gravitational effects. The inverse modeling feature of DARTS based on the adjoint method will be utilized in this study to conduct the history matching for the DiWA model.

5.1.1. CONNECTIVITY GRAPH FOR PROXY MODEL

In order to establish a connection between well locations and production data, it is necessary to represent the reservoir domain of interest in a discrete form. The geometrical discretization of the reservoir is usually accomplished by means of control volume partitioning. In this study, unstructured partitioning and finite-volume discretization [49] are employed to create a proxy model. This approach results in a spatial connectivity graph that provides a discrete representation of the proxy reservoir model in terms of the connections between control volumes and their associated transmissibilities [62]. The proposed data-driven approach uses this technique to partition the reservoir domain at a coarse resolution.

The discretized model is defined using boundaries that are gridded using the hierarchical approach: a volume (convex polyhedra) is bounded by a set of surfaces (convex polygons), a surface is bounded by a series of curves (segments), and a curve is bounded by two endpoints (nodes). To grid the model, we utilized the automatic open-source meshing software package GMSH [28]. Since the exact grid parameters and rock properties, such as permeability and exact layer geometry, are not available, initial guesses for model parameters were computed based on averaged values from initial evaluations of the reservoir parameters. When these values were not known, they were estimated either using an analog field or based on common sense. Figure 5.1 illustrates a typical unstructured mesh with corresponding well locations.

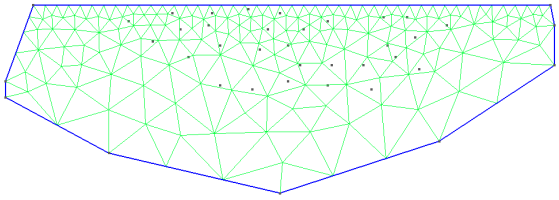


Figure 5.1: Typical grid for connectivity graph generated for proxy model with corresponding border lines and well positions.

5.1.2. MODEL PARAMETERS AND OBJECTIVE FUNCTION

In this study, the history matching is carried out using constrained optimization, whereby each model parameter is bounded by minimum and maximum values. Combining with the bound, each model parameter is scaled to the interval $[0,1]$, as some of the regression algorithms are sensitive to the scale of a problem. The initial guess and constrained intervals for the model parameters are determined based on the available data and are incorporated into the data-driven model. Corresponding to the model parameters from Section 3.1.2, we present the lower and upper bounds for those model parameters.

The number of the model parameters of transmissibility is equal to the number of interfaces between grid blocks. During the regression process, these parameters are subject to constraints, specifically $\bar{T}_{min} = 1$ cP$\cdot$m^3/d/bar and $\bar{T}_{max} = 50{,}000$ cP$\cdot$m^3/d/bar. The number of the well index is equal to the number of well perforations. They are bounded between 1 cPm^3/d/bar and 1000 cPm^3/d/bar. Apart from the transmissibility and well index, the rock-fluid interaction parameter is also included in the model parameters. They are subject to the following constraints.

The model parameters for transmissibility in a reservoir simulation are determined by the number of interfaces between grid blocks. During the history matching process, these parameters are subject to constraints, specifically $\bar{T}_{min} = 1$ cP$\cdot$m^3/d/bar and $\bar{T}_{max} = 50{,}000$ cP$\cdot$m^3/d/bar. Similarly, the well index is determined by the number of well perforations and is bounded between $WI_{min} = 1$ cPm^3/d/bar and $WI_{max} = 1000$ cPm^3/d/bar. In addition to the transmissibility and well index, the model parameters also include the rock-fluid interaction parameter, which is subject to the following constraints:

$$P_{fr,min} = \{0.0, 0.0, 0.00001, 0.00001, 100, 10\}, \tag{5.1}$$

$$P_{fr,max} = \{0.49, 0.49, 5, 5, 3000, 2000\}. \tag{5.2}$$

The first four parameters are unitless; the units of the last two are $[s \cdot m^{-2}]$. These constraints were obtained based on the physical interpretation of model parameters.

In this example of DiWA model of oil reservoir, the objective function is constructed using the phase flow rates of the wells, represented by f_1 in Equation (3.21). To improve the efficiency of the gradient evaluation used in the history matching, the adjoint method is applied for the model parameters of transmissibility and well index. As for the rock-fluid interaction parameters, their associated gradients are evaluated using the finite difference method.

5.1.3. VALIDATION OF DIWA PROXY MODEL

To validate the proposed DiWA proxy model method, we utilized two ensembles of stochastic fluvial models. For each ensemble, we constructed two reduced-order models: the first model was generated using conventional flow-based upscaling [12], while the second model was generated using the proposed DiWA proxy model. The high-fidelity models were considered as reference models for generating the observed true data of the oil rate. The DiWA models were trained based on the observed true data of the high-fidelity model using constant transmissibilities and well indexes as initial guesses. Next, we compared the results of the upscaled model and the DiWA model with the high-fidelity model for the entire ensemble to assess their accuracy. All models were under the assumption that fluid and rock physics were known and fixed, except for transmissibility

and well indexes. The same set of parameters was used for all other model tests, which are listed in Table 5.1.

Table 5.1: Main physical properties used for comparison among high fidelity, upscaled and DiWA models.

Phase	Oil	Water
Fluid compressibility, [1/bar]	1.34×10^{-4}	4.35×10^{-5}
Fluid densities, [kg/m^3]	897.0	1002.8
Residual saturation	0.15	0.225
Endpoint relative permeability	0.4	1.0
Saturation exponent	3.0	3.0
Viscosity, [cP]	1.294	0.320

GENERATION OF ENSEMBLES OF HIGH-FIDELITY FLUVIAL MODELS

The stochastic ensemble of fluvial models was used to illustrate the accuracy of the proposed methodology of DiWA proxy model. High-fidelity model ensembles were generated by two different modeling approaches:

- FLUMY [29]: Process-based models using FLUMY software, see example in Figure 5.2(a),
- MPS [98]: Multiple Point Statistics (MPS) models, see example in Figure 5.2(b).

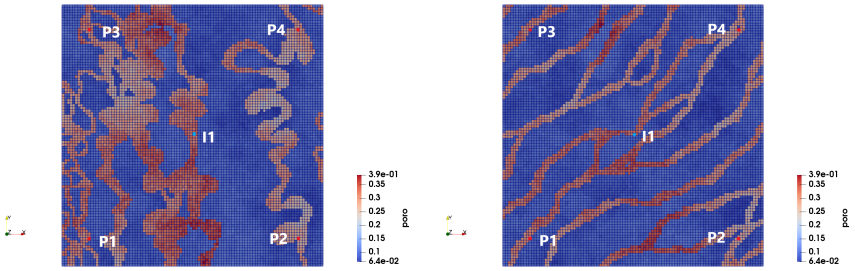


Figure 5.2: Porosity distributions of a typical high-fidelity model realization generated by process-based modeling approach with FLUMY software (a) and (b) stochastic modeling approach using Multiple Point Statistics (MPS) (b).

This results in completely different model complexity between the ensembles. Each high-fidelity ensemble model has a size of 100 by 100 grid cells (cell dimensions are $10 \times 10 \times 10$ m). The models use a simple 5-spot vertical well set-up. One injector (blue dot "I1" in Figure 5.2) is located in the middle of the reservoir and surrounded by four producers (red dots "P1", "P2", "P3" and "P4" in Figure 5.2) located at reservoir edges. Injection wells are modeled by setting a BHP control of 250 bar plus uniformly distributed random perturbations ranging from 1 to 30 every 200 days, and production wells are set at a fixed BHP control of 100 bar. The simulation was limited to 4000 days. The main

difference between models generated by MPS and FLUMY is the main paleoflow orientation ranging from SW-NE to W-E. Besides, FLUMY model has a limited statistical variability in comparison to the MPS model [37].

The typical porosity distribution of the high-fidelity realizations from the MPS and FLUMY ensembles can be observed in Figure 5.2. It is evident that the model generated by MPS is more complex, as the phase can only flow through distinct channels, which are generally smaller in size than the coarse grid block. Conversely, the FLUMY-generated model has numerous overlapping channels that provide multiple possible flow paths. Once the coarser proxy models are trained to match the historical field data of the FLUMY or MPS model, it becomes apparent that the trained proxy model of FLUMY is better suited for capturing the reservoir dynamics on a coarser scale. Further details about high-fidelity ensemble generation, upscaling, and simulation properties can be found in [37].

GENERATION OF UPSCALED PROXY MODELS

Upscaled proxy models were generated using a global flow-based upscaling technique, which involves solving the fine-scale incompressible single-phase pressure equation and using it to derive the coarse-scale transmissibility:

$$-\nabla \cdot \left(\frac{K}{\mu} \nabla \left(p - \rho \bar{g} \right) \right) = q_{well}. \tag{5.3}$$

Under the two-point flux approximation, Equation (5.3) can be written in the discrete form, in which the coarse properties can be evaluated by:

$$\left(q_x^c\right)_{i+1/2,j} = \left(\bar{T}_x^c\right)_{i+1/2,j} \left(p_{i,j}^c - p_{i+1,j}^c \right). \tag{5.4}$$

Here, $\left(q_x^c\right)_{i+1/2,j}$ is the coarse flux across the interface $i+1/2, j$ simply defined as the integrated fine scale fluxes across the coarse interface, $\left(\bar{T}_x^c\right)_{i+1/2,j}$ is the coarse transmissibility and ($p_{i,j}^c$ and $p_{i+1,j}^c$) are the coarse pressures obtained by arithmetic averaging the fine-scale pressures contained in each coarse-scale block respectively.

A similar approach for a flow-based upscaled well index of well w can be derived, given by the following equation

$$WI_w = \frac{q_w}{p_{i,j}^c - p_w}. \tag{5.5}$$

The big advantage of the global flow-based upscaling technique is its computational efficiency and accuracy. However, for highly heterogeneous reservoirs, the resulting transmissibility values may be large or even negative [12]. In this case, we correct the value of transmissibility at the nearest bound (either the upper or lower bound) and then assign new pressure values to the neighbor blocks by requiring that Darcy's law is fulfilled. However, the correction of transmissibility may cause new pressure differences between the neighbor blocks. Hence, the upscaling procedure of computing the transmissibity must be applied in an iterative fashion until the criteria are satisfied. This iterative upscaling procedure is generally used to obtain a positive definite transmissibility matrix, which is typically reached within 5 iterations, see [35] for more details.

Each high-fidelity model was upscaled laterally by 100 times (10×10) in the x and y directions, which means the upscaled model has the size of 10 by 10 grid cells (cell dimensions are $100 \times 100 \times 10$ m). The global upscaling with specific well conditions used in the reference model was applied, details can be found in [37]. The resulting upscaled transmissibility, porosity and well index was used to initialize the upscaled proxy model. Well controls and simulation time were kept identical to the high-fidelity model. Moreover, the same upscaled porosity was used in the data-driven proxy model to ensure pore volumes match between models.

GENERATION OF DIWA DATA-DRIVEN MODELS

DiWA Data-driven proxy models were generated with the same grid and well configuration as for upscaled proxy models. The initial guess for transmissibilities and well indexes of the data-driven proxy model was chosen as a uniform distribution of 100 $\mathrm{cP\,m^3/d/bar}$ and 200 $\mathrm{cP\,m^3/d/bar}$.

Then, DiWA data-driven models are trained based on the observed true data. The history matching was limited to 100 iterations. Some cases converged before reaching the imposed maximum. The training of a single realization takes from 20 minutes to one hour on a single cluster node with four Intel Xeon CPU E5-2650 v3 processors.

COMPARISON BETWEEN UPSCALED AND DIWA DATA-DRIVEN PROXY MODELS

To ensure a proper comparison between the DiWA data-driven and upscaled models, we maintained consistency in terms of model volume, physical properties, and well controls. This was essential since the upscaling procedure is sensitive to boundary conditions. As such, all parameters were kept identical. In order to validate the accuracy of the proxy modeling methodology, we compared the stochastic response of the trained proxy results for two data-driven models against conventional flow-based upscaled models.

Figure 5.3 illustrates the total water rate of all 100 realizations for high-fidelity, DiWA data-driven proxy, and upscaled proxy cases. It can be seen that the stochastic response of the high-fidelity and DiWA data-driven models have a reasonably good agreement for both mean and individual realization water rates, whereas, the upscaled model rates matched worse. Figure 5.4 shows the P10, P50 and P90 quantile response from Figure 5.3. It shows that the DiWA data-driven model has a better agreement with the response of the high-fidelity model compared with the response of the upscaled model. The oil and water mean error for DiWA data-driven and the upscaled model are shown in Table 5.2.

Table 5.2: The mean errors of water and oil rate of FLUMY model

Model	Water rate	Oil rate
DiWA data-driven model	1.5%	5.8%
Upscaled model	13.0%	10.6%

Then, the same test was performed for a more complicated model ensemble build with Multi-Point Statistics (MPS) modeling approach. The water rates and the corresponding quantile response of all 100 realizations for high-fidelity, DiWA data-driven proxy, and upscaled model are shown in Figure 5.5 and Figure 5.6, respectively. The oil

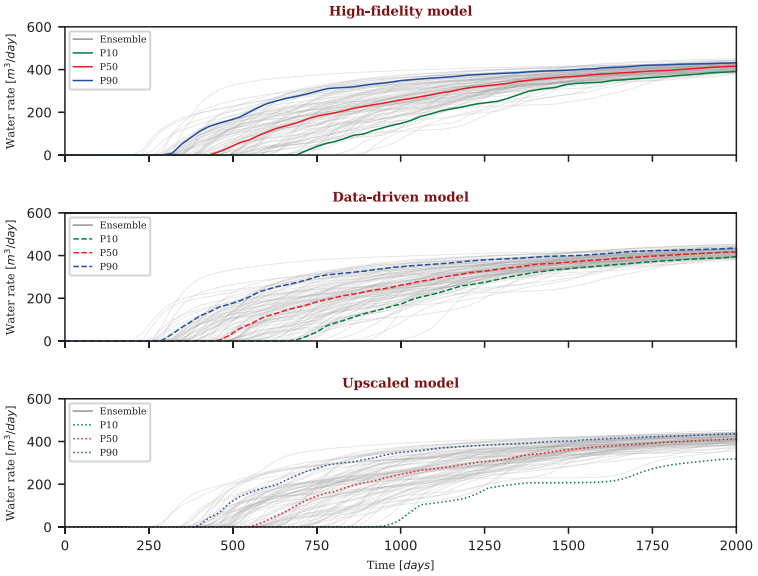


Figure 5.3: The total water rate for the high-fidelity (reference) model with the size of the 100x100 grid block for the hundred FLUMY realizations, together with DiWA data-driven, and upscaled models (10x10) response. Grey lines indicate rates from a single model realization, whereas the red, blue, and green lines indicate the quantile response of the ensemble i.e the P10, P50, and P90.

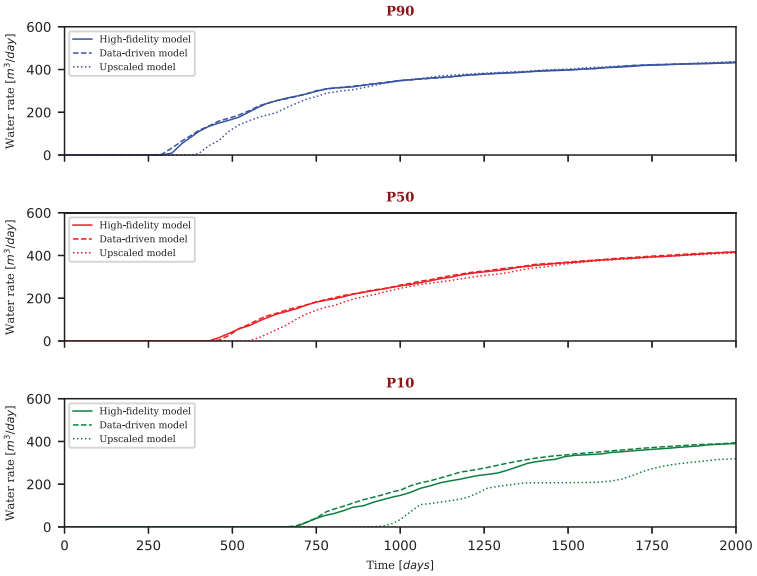


Figure 5.4: P10, P50 and P90 quantile response of the high-fidelity (reference), DiWA data-driven, and upscaled models for 100 FLUMY realizations.

and water mean errors for DiWA data-driven and the upscaled model are shown in Table 5.3. Based on those results, one can conclude that errors for both types of proxy models are higher for the MPS ensemble than the FLUMY ensemble.

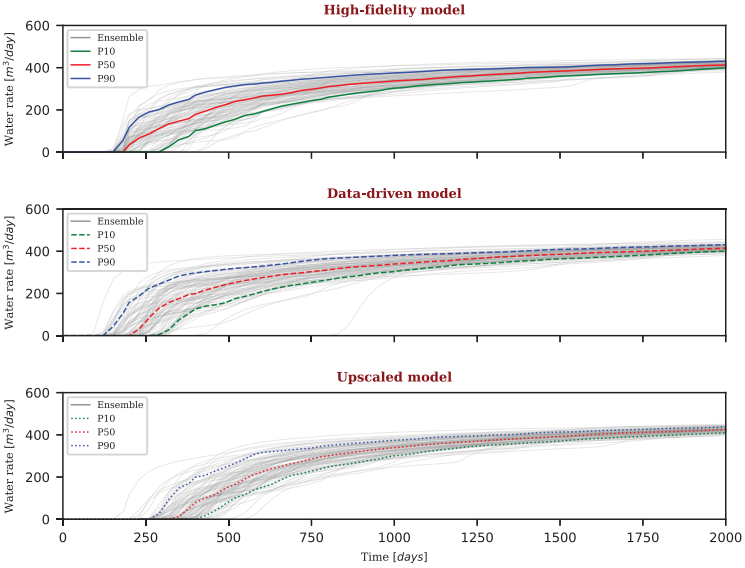


Figure 5.5: The total water rate for the high-fidelity (reference) model with the size of the 100x100 grid block for the hundred MPS realizations, together with DiWA data-driven, and upscaled models (10x10) response.

Table 5.3: The mean errors of water and oil rate of MPS model

Model	Water rate	Oil rate
DiWA data-driven model	4.5%	5.9%
Upscaled model	19.6%	12.0%

It is an expected result as it is much more difficult to find a value for the effective property on a coarse scale that will accurately represent fine-scale features (e.g., small and poorly connected channels, which can be seen in Figure 5.2,b). On the contrary, the channels in the FLUMY model overlap each other creating more distinct and rough flow paths, which are easier to capture on a coarse scale. The overall accuracy of the DiWA data-driven proxy model is still significantly higher than that of the upscaled proxy model. This is because the upscaled model uses the reduced order approximate properties, while the DiWA data-driven model is directly trained by the response of the reservoir. For example, the upscaling of some reservoir properties (e.g. porosity and permeability) is based on averaging methods. In this case, it may lose some accuracy using a more homogenized model to characterize the original heterogeneity of the model. Besides, we only apply well-established single-phase upscaling which has limited accuracy in multiphase flow situations. The examples of FLUMY and MPS confirm the applica-

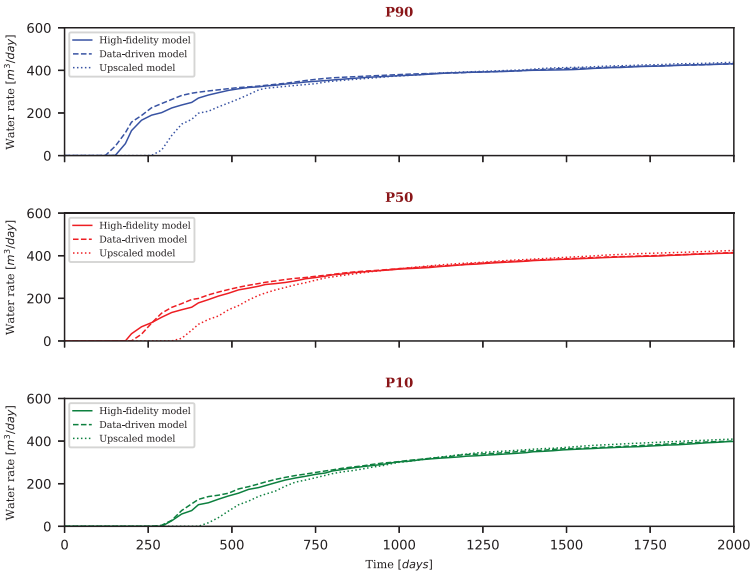


Figure 5.6: P10, P50 and P90 quantile response of the high-fidelity (reference), DiWA data-driven, and upscaled models for 100 MPS realizations.

bility of the DiWA data-driven approach for uncertainty quantification analysis when a reliable and accurate high-fidelity model is not available.

5.2. STOCHASTIC DIWA PROXY MODEL

In Section 5.1, the DiWA proxy model was introduced and tested on two large ensembles of fluvial proxy models (i.e. FLUMY and MPS model). In this section, we extend the DiWA proxy model method to the stochastic DiWA proxy model method, which randomly samples the permeability of each cell of the stochastic proxy model based on the probability density of the field. Prior to model training, a filtering procedure is applied to remove outliers while retaining the most promising candidates under a specific threshold. The retained candidates are then history-matched to the observation data.

The application of these stochastic DiWA proxy models is essential because it enables to search for the optimum solution in a wider range of the parameter space, therefore it avoids being trapped in local minima or losing the generality, especially in the case of lacking field geological information. Unlike traditional methods for generating prior geological models, which require detailed petrophysical data, the stochastic proxy model uses a much coarser grid to represent connectivity between wells and reservoirs while incorporating the field's basic geological statistics.

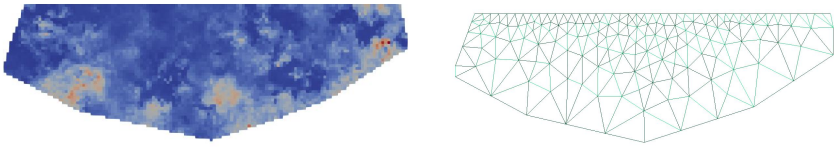
The use of the adjoint method and the coarse grid of the proxy model makes it possible to perform history matching for a large ensemble of realizations within hours. The trained proxy model can be further used in production optimization problems with high efficiency due to its low degrees of freedom. This section also includes the characteriza-

tion of geological information using a proxy model, which is demonstrated by testing on a three-phase 3D proxy model. The results show that the proxy approach can be used to characterize the petrophysical information of the reservoir if the initial guess is sufficiently close to the true model.

5.2.1. METHODOLOGY

DIWA MODEL CONSTRUCTION

We utilize the Brugge field model to demonstrate the procedure of generating a DiWA model and to apply it in our framework. Brugge is a benchmark model for the optimization of reservoir production [84], which is used to generate observation data for the DiWA model training. The structure of the Brugge model consists of an East-West elongated half-dome with a large boundary fault at the northern edge. There are 20 production wells surrounded by 10 injection wells in this field. The realization encoded as FY-SS-KP-8-73 is used in this study to prepare the observation data of oil production. The permeability distribution of the realization FY-SS-KP-8-73 can be found in Figure 5.7. This model runs for 3720 days with BHP control changing every 120 days. For the DiWA model, unstructured partitioning and finite-volume discretization [48] are applied to generate a coarse-resolution DiWA model.



Realization: FY-SS-KP-8-73 Unstructured DiWA model

Figure 5.7: The realization FY-SS-KP-8-73 and the unstructured DiWA model. Note that the realization FY-SS-KP-8-73 is a structured grid model with 44 550 active blocks, while there are only 283 cells in the unstructured DiWA model.

Notice that this unstructured DiWA model can only be considered as a proxy model since it is based on very basic information about the original field. For example, the reservoir boundaries are approximated by several piece-wise linear segments as can be seen in Figure 5.7. The large boundary fault is not included in the unstructured DiWA model. The well locations are obtained based on projecting the real well locations to the nearest unstructured control volumes. The entire 2D domain of this field is meshed and then extruded in the vertical direction by an average constant thickness of the reservoir. Different characteristic lengths are set to generate a grid with coarser and finer meshing in the outer and inner boundaries, respectively. The finer inner resolution is explained by the main well locations and corresponding flow dynamics while the rest of the domain is located in the peripheral water drive with no significant flow. This helps to largely reduce the degrees of freedom of the proxy model. However, the DiWA model can always be further refined or elaborated simply by adding more piece-wise linear segments inside or at the boundary of the reservoir domain when it is necessary, for example, introducing the fault or extra cells around wells as shown in [100].

As for the discretization of the well, it is conducted by adding the well to the connection list in the form of a well head block and a well body block. The well head, where the well control or constraint is assigned, is only connected with the well body, then the well body is further connected to the neighboring reservoir block. This procedure above shows the flexibility of discretizing the reservoir and wells, because we only need to update the existing connection list of the reservoir by adding two extra connectivities (i.e. the well head to the well body, and the well body to the reservoir block) for a well. One may also notice that the newly added well blocks may not be necessarily located in the center of the reservoir block, so the well index can not be calculated correctly. This problem can be solved by introducing an equivalent square block for the triangular cell where the well is located, and plugging in the equivalent Δx and Δy into Equation (3.8). Although this method can not provide an accurate value of the well index for the given well, the approximated well index will be further updated in the course of history matching.

THE PSEUDOCODE FOR THE CONSTRUCTION AND APPLICATION OF DIWA MODEL

The construction procedure of the DiWA model is shown in Algorithm 1. It should be noted that in this study the sampling of the permeability K is based on the probability density of the high-fidelity model. However, the high-fidelity model is not the prerequisite in practice because this sampling procedure can be done as long as there is information on the probability density of the reservoir.

Algorithm 1 The construction of DiWA model candidates

1: Use several piece-wise linear segments to represent the reservoir boundaries
2: Discretize the reservoir with very coarse unstructured 2D cells
3: Extrude the 2D mesh in vertical direction by an average constant reservoir thickness to form a 3D model
4: $seed = 1$, $N = 1$, and $list_{seed}$
5: **while** $N \leq 1000$ **do**
6: Do the sampling of permeability K for each cell
7: Generate the connection list to represent the connectivity between the neighboring cells
8: Calculate the transmissibility $\bar{T}$ and add them to the connection list
9: Add the well cells to the reservoir mesh
10: Calculate the well index WI for each well and add them to the connection list
11: Run a forward simulation for the proxy model
12: Calculate the misfit between the model response and the observation data
13: **if** misfit ≤ threshold **then**
14: Collect this model as candidate and save the $seed$ in $list_{seed}[N]$
15: $N = N + 1$
16: **end if**
17: $seed = seed + 1$
18: **end while**

Once the collection of the DiWA model candidates is finished, they will be trained

to fit the observation data by changing the values of the transmissibility $\bar{T}$, well index WI, and several parameters related to relative permeability. These parameters are the model parameters in history matching and are already explained in Section 3.1.2. The trained DiWA models can be used to predict production because they are calibrated by the observation data. Production optimization is another application of these trained DiWA models. The pseudocode of the training and the forecast application of the DiWA models are shown in Algorithm 2.

Algorithm 2 The training and the forecast application of the DiWA model

for N=1,2,3... **do**
  $seed \Leftarrow list_{seed}[N]$ ▷ $list_{seed}$ is generated from Algorithm 1
  Use $seed$ to reproduce the DiWA model candidate
  Re-scale $\bar{T}$, WI, and $\{S_{or},\ S_{wc},\ n_o,\ n_w,\ k^e_{rw}\rho_w/\mu_w,\ k^e_{ro}\rho_o/\mu_o\}$ in the range of (0, 1)
  while $misfit > tol_{misfit}$ and $stepsize > tol_{stepsize}$ **do**
    Calculate the gradients with respect to $\bar{T}$ and WI using adjoint method
    Calculate the gradients with respect to $\{S_{or},\ S_{wc},\ n_o,\ n_w,\ k^e_{rw}\rho_w/\mu_w,\ k^e_{ro}\rho_o/\mu_o\}$ using finite difference method
    Update $\bar{T}$, WI, and $\{S_{or},\ S_{wc},\ n_o,\ n_w,\ k^e_{rw}\rho_w/\mu_w,\ k^e_{ro}\rho_o/\mu_o\}$ using the gradients
    Calculate the misfit between the model response and the observation data
  end while
  Set the total simulation time as $t = t_{training} + t_{forecast}$ to do the forecast
  Re-run the trained model using the optimized model parameters
  Calculate the errors between the model response and the observation data
  Plot and compare the model response and the observation data
end for

5.2.2. THE APPLICATION OF STOCHASTIC DIWA PROXY MODEL

In this section, we start with an example that proves the applicability of the DiWA approach for truly 3D models with a three-phase flow. Besides, this example demonstrates that the trained proxy model can recover the petrophysical information when the initial guess is close enough to the true solution. Finally, we will show an efficient approach for the generation of the DiWA proxy models which can be used in uncertainty quantification or robust optimization.

DETERMINISTIC RECONSTRUCTION OF PERMEABILITIES

To test the framework, we take a refined SPE 1 model [78] for the generation of true data. In this model, there are mainly three different layers with constant permeability. Apart from the oil and water phase, the gas phase is also considered in this model and the gas is injected from the top layer at one of the reservoir corners. We build a coarse model and train it based on the observation data generated from the modified SPE 1 model. After the coarse model is trained, we interpret the regressed transmissibility back into

permeability and compare them with the permeability distribution of the original SPE 1 model.

According to the paper SPE-9723-PA [78], the SPE 1 model is a three-phase three-dimensional black-oil reservoir where there is dissolved gas existing in the reservoir fluids. The densities of oil, water, and gas at the surface condition are 786.5 kg/m^3, 1037.8 kg/m^3, and 0.97 kg/m^3, respectively. The rock compressibility is 4.35×10^{-5} bar^{-1} when the reference pressure is 277.0 bar. The water formation volume factor, compressibility, and viscosity at the same reference pressure are 1.029, 4.54×10^{-5} bar^{-1}, and 0.31 cp, respectively. The parameters related to the relative permeability and PVT data of oil and gas phases can be found in Tables 5.4 to 5.6.

Table 5.4: The parameters of relative permeability used in SPE 1 model

Phase	Oil	Water	Gas
Residual saturation	0.103	0.197	0.013
End point relative permeability	0.584	0.910	0.830
Saturation exponent	2.54	1.01	1.62

Table 5.5: The formation volume factor and viscosity of the gas phase in SPE 1 model

Pressure (bar)	Formation volume factor (-)	Viscosity (cP)
1	166.666	0.008
18	12.093	0.0096
35	6.274	0.0112
69	3.197	0.014
137	1.614	0.0189
171	1.294	0.0208
205	1.08	0.0228
273	0.811	0.0268
341	0.649	0.0309
613	0.386	0.047

Table 5.6: The formation volume factor, gas oil ratio, and viscosity of the oil phase in SPE 1 model

Pressure (bar)	Formation volume factor (-)	Gas oil ratio (-)	Viscosity (cP)
1	1.062	0.001	1.04
18	1.15	0.0905	0.975
35	1.207	0.18	0.91
69	1.295	0.371	0.83
137	1.435	0.636	0.695
171	1.5	0.775	0.641
205	1.565	0.93	0.594
273	1.695	1.270	0.51
613	1.579	1.270	0.74

In our modified version of the SPE 1 model, there are 37500 grid blocks in total

[50 × 50 × 15]. The block dimension in both the x and y directions is 60 m. In the z direction, each layer is divided into 5 grid blocks, and the block sizes for the three layers in the z direction are 1.2 m, 1.8 m, and 3 m, respectively. The porosity is 0.3. The permeabilities for three layers from top to bottom are 500 mD, 50 mD, and 200 mD, respectively. There are two wells (one injector and one producer) located in the opposite corners of the reservoir, see Figure 5.8. Gas is injected into the top layer and oil is produced from the bottom layer. The initial pressure of this model is 330 bar. The initial molar fraction of gas, oil, and water are 0.001, 0.648, and 0.351, respectively. The injector is set at a constant BHP control of 400 bar, and the producer is set at a constant BHP of 150 bar. The total simulation time of this model is 2000 days.

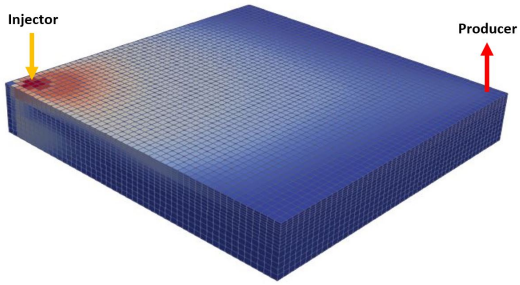


Figure 5.8: SPE 1 model with two wells located at the opposite corners. The injector and its perforations are located at grid block (3,3,1-5). The producer and its perforations are located at grid block (48,48,11-15). The dimension in the z-direction is exaggerated 15 times for better visualization.

A DiWA proxy model with 300 grid blocks [10 × 10 × 3] is constructed and trained based on the observation data generated from the high-fidelity model. To make sure the well positions of the coarse model are identical to the high-fidelity model, we set the injector and producer at the grid block (1,1,1) and (10,10,3), respectively. Since three phases are considered in this problem, Stone 1 Model [96] is implemented to calculate the three-phase relative permeability and the same sets of the parameters of the Stone 1 Model are applied in both high-fidelity and proxy model and they are not adjusted in the course of model training. Only the model parameters of transmissibility and well index are adjusted in the course of model training using the adjoint gradients. The initial guess for transmissibilities and well indexes are generated by adding a random perturbation around the true transmissibility of the high-fidelity model. The scaling factors for transmissibilities and well indexes are 10000 and 1000, respectively. The error between the trained parameters and the true data is calculated using the following equation:

$$Err = \frac{\delta_t(T_{obs}) \sum_{k=1}^{K} \sum_{w=1}^{n_w} \sum_{j=1}^{n_p} \left(\boldsymbol{q}_{k,w,j} - \boldsymbol{Q}_{k,w,j}\right)^2}{\sum_{k=1}^{K} \sum_{w=1}^{n_w} \sum_{j=1}^{n_p} \left(\boldsymbol{Q}_{k,w,j}\right)^2}. \tag{5.6}$$

where k, w, and j are the index of time steps, wells, and phases, respectively. q denotes the model response of the DiWA model. Q is the observations. K, n_w, and n_p are the total number of time steps, wells, and phases, respectively.

The training results are shown in Table 5.7 and Figure 5.9.

Table 5.7: The iterations and the errors between the rates of the DiWA and the high-fidelity model before and after training under different perturbations around the true transmissibility

Perturbation	Error-initial, %	Error-trained, %	Iteration
2%	0.1167	0.0525	3
5%	0.0846	0.0501	4
10%	0.3057	0.0414	5
20%	2.0281	0.0436	2
40%	10.4036	0.0429	3
80%	45.8739	0.0450	12
100%	82.5188	0.0197	34

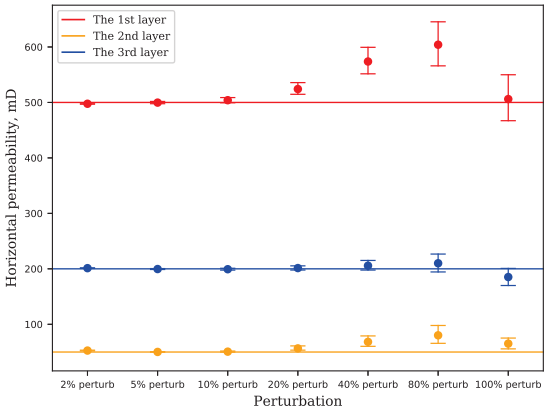


Figure 5.9: The mean and the deviation of the horizontal permeability of the DiWA model after training. The dots and the error bars represent the mean values and the deviation for the whole grid blocks in each layer, respectively.

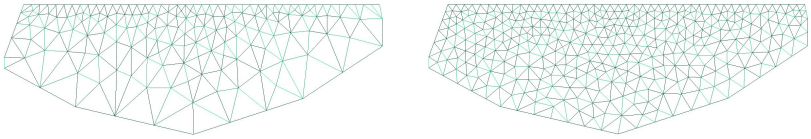
It can be seen from Table 5.7 that as the increasing perturbation to the true transmissibility, the errors between the proxy and high-fidelity model increase from 0.1% to 83%. After the training, all proxy models can achieve a very small error value. It is noticed that the error after training for the 100%-perturbation case is 0.0197%, which is smaller than the error of 0.0525% of the 2%-perturbation case. This can be explained by the non-uniqueness of the minimization procedure since the errors after training of all cases are close to zero.

However, a smaller error between the DiWA model and the high-fidelity model does not guarantee that the true solution is recovered due to the ill-posedness of the inverse problem. In Figure 5.9, the mean values and deviations of the permeability values for each layer of the proxy model demonstrate significant divergence for the perturbation larger than 20%. This indicates that the selection of initial guesses can largely affect the characterization results of geological information. When the initial guess of the proxy model is generated based on highly reliable geological information, the trained model can recover the true petrophysical characteristics. However, when the geological information is missing or not sufficiently constrained, we need to apply a sampling proce-

dure to generate the proxy model candidates. An example of the sampling procedure and training results is described next.

THE TRAINING AND PRODUCTION FORECAST OF THE DIWA PROXY MODEL

The unstructured model Case 1 and Case 5 from Figure 4.4 are selected to test the proposed framework. Both grids are shown in Figure 5.10. The model parameters of the coarser model consist of 394 transmissibility, 30 well indexes, and 6 rock-fluid interaction parameters. As for the finer model, there are 953 transmissibility, 30 well indexes, and 6 rock-fluid interaction parameters.



Case 1 (coarse DiWA model) Case 5 (fine DiWA model)

Figure 5.10: Two DiWA models with different grid resolutions

We will first show the training and production forecast results based on the coarser model. The reservoir initial pressure, initial oil saturation, and BHP control in the DiWA model are kept identical to the high-fidelity model. Unlike the conventional process of the generation of prior geological models that requires detailed well logging data, seismic data, etc., we randomly sample the permeability for each cell of the DiWA model based on the probability density of the realization FY-SS-KP-8-73 shown in Figure 5.11. The generation of these stochastic DiWA models can be done in the situation of very limited geological information, while it incorporates the basic geological statistics of the field.

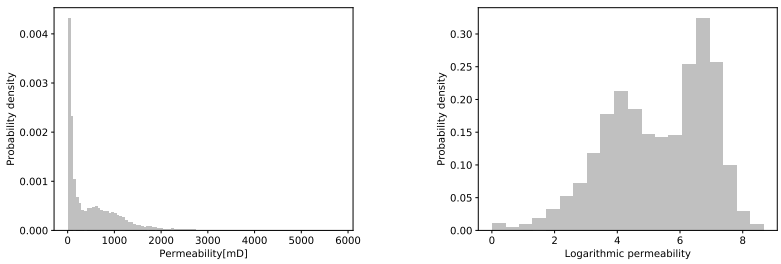


Figure 5.11: The probability density for the permeability of the realization FY-SS-KP-8-73. The x-axis of the right figure represents the logarithmic permeability.

Equation (3.5) is then applied to compute the connection list and the corresponding transmissibilities for the coarser model. These transmissibilities are taken as the initial guess of transmissibility parameters and are bounded in the range of (0.01, 50000). The scaling factor for the transmissibility is 50000. As for the initial guess for well indexes,

Equations (3.7) and (3.8) provide an approximation of well index for the structured grid. Here we introduce an equivalent length of the triangular cell in the unstructured grid to replace the structured block size Δx and Δy in Equation (3.8). They are given by:

$$\Delta x = \Delta y = \sqrt{A_{tri}}, \tag{5.7}$$

where A_{tri} is the area of the given triangular cell. Equation (5.7) assumes that the triangular cell in the unstructured grid has the same effect on the well index as the square block with the same area in the structured grid. The well index is bounded in the range of (0.001, 10000), and they are normalized by the scaling factor 10000. The initial guess of 6 rock-fluid interaction parameters are uniformly sampled from the range of (0.00001, 5), (0.00001, 5), (0, 0.49), (0, 0.49), (100, 3000), and (10, 2000), respectively. These 6 rock-fluid interaction parameters will be applied in the whole reservoir, and they are normalized by the scaling factors 5, 5, 0.49, 0.49, 3000, and 2000, respectively. The parameters of relative permeability and the PVT data of oil used in the Brugge high-fidelity model can be found in Tables 5.8 and 5.9. There are only oil and water phases existing in the reservoir throughout the production period. The density of oil and water are 897.0 kg/m^3 and 1002.8 kg/m^3, respectively. The rock compressibility is 5.08×10^{-5} bar^{-1} when the reference pressure is 170.0 bar. The water formation volume factor, compressibility, and viscosity at the same reference pressure are 0.9927, 4.35×10^{-5} bar^{-1}, and 0.32 cp, respectively.

5

Table 5.8: The parameters of relative permeability used in Brugge high-fidelity model

Phase	Oil	Water
Residual saturation	0.15	0.225
End point relative permeability	0.4	1.0
Saturation exponent	3.0	3.0

Table 5.9: The formation volume factor and viscosity of oil in the Brugge model

Pressure (bar)	Formation volume factor (-)	Viscosity (cp)
1	1.0007	1.294
170	0.9780	1.294
500	0.9337	1.294

The high-fidelity model runs for 3720 days to generate true observation data. For every 120 days, the BHP controls of injection wells are set as 170 bar plus uniformly distributed random perturbations ranging from 1 to 30, and the BHP controls of production wells are set as 130 bar minus uniformly distributed random perturbations ranging from 1 to 30, see Figure 5.12. We choose the first 3000 days as the training period to train the DiWA model. After that, the DiWA model continues to run for 720 more days in order to test the accuracy of the forecasting period compared with the true response.

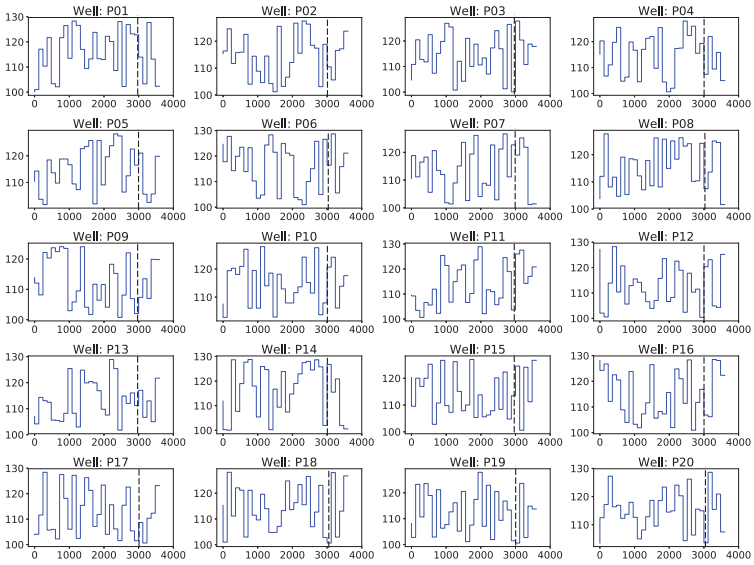
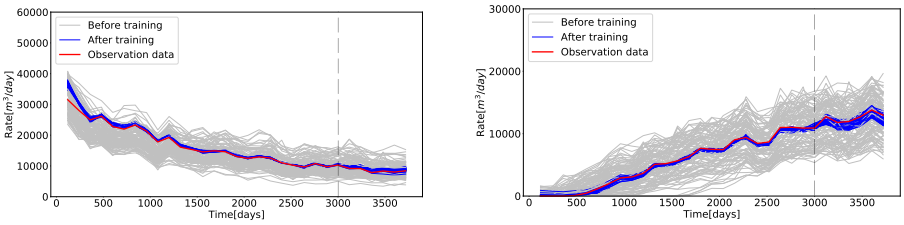


Figure 5.12: The BHP control of production wells; the x-axis shows the production time (days), the y-axis shows the BHP of production wells (bar); The vertical dashed lines separate the training periods and the prediction periods

Before the training of the DiWA model, we first sample for many different realizations and then check their misfit of the total oil and water rate between the model response and observation data. The threshold for the rate variability is taken as 60% of the true rate, and then we use this threshold to filter out the outliers. Only the realization that meets this threshold will be retained otherwise discarded. This sampling procedure would not stop until 1000 candidates of realization are collected. Later, these 1000 realizations are set as the initial guesses to train the DiWA model. Note that the training procedure for the ensemble of proxy models can be finished within a limited time thanks to the high efficiency of gradient calculation using the adjoint method and the utilization of computer clusters. The total training time of the 1000 realizations took below 12 hours on 10 cluster nodes with 40 Intel Xeon CPU E5-2650 v3 processors in total.

The results of the total oil and water production rate of 100 best realizations are shown in Figure 5.13. The grey and blue curves represent the rate before and after training, respectively. The red curves show the true data. The vertical dashed lines separate the training periods and the prediction periods. It is clearly seen that proxy models cannot capture the characteristics of the first stage (i.e. transient stage that is from the beginning to around 400 days) of oil production due to the coarse representation of well connectivity in the proxy methodology. However, the later period is captured quite well and the deviation of the model response from the true solution is very limited. This deviation increases when the model switch from training to forecast period which is expected.

More detailed comparisons of oil rates for 20 production wells are plotted in Figure 5.14. The results show that the mean error of the training period and prediction pe-



Total oil rate Total water rate

Figure 5.13: The total oil (left) and water (right) production rate of 100 best fit realizations before and after training of Case 1 (coarse model)

riod of the oil rate are 0.59% and 0.62%, respectively. As it can be seen from Figure 5.14, the oil rates of production wells have a good match with the true data in the training period (before the vertical dashed lines). It can be seen that only a few production wells have a relatively larger spread of oil rates in the forecast period compared with the spread in the training period. The water production curves of these 20 production wells are shown in Figure 5.15. As it can be seen that most of the wells also have a good match with the true data of water production rates. There are three wells (Well:P01, P04, and P05) that have water breakthrough after the training period. We plotted these three wells in Figure 5.16 separately. Note that the magnitudes of the y-axis (water production rate) of these three wells are smaller than the rest of the wells. Most of the realizations capture the timing of the water breakthrough. But the water rates of these realizations after the water breakthrough have a relatively larger spread compared with the curves before the water breakthrough. Again, these results are expected since proxy models have a limited forecast capability.

Similarly, we implement the identical filtering and training strategy for the finer model in Figure 4.4, which is a much finer unstructured grid and has more degrees of freedom compared with the coarser model. The total production rates of oil and water are shown in Figure 5.17. The oil rates of 20 production wells are plotted in Figure 5.18. It can be seen that the spread of the oil rates has been reduced for some wells. The resulting mean error of the training period and prediction period are 0.44% and 0.17%, respectively. The total training time took below 23 hours on the same cluster setup. The water production curves of these 20 production wells are plotted in Figure 5.19. The forecast results of the water breakthrough of Well:P01, P04, and P05 are shown in Figure 5.20. A similar conclusion can be drawn that the spread of the water rates is reduced compared with the results of Case 1 in Figure 5.15, and most of the realizations are able to capture the timing of water breakthrough. The forecast after the water breakthrough also has a relatively larger spread compared with the curves before the water breakthrough. This phenomenon is expected because the nature of the severe non-linearity in water breakthrough makes the model hard to perfectly match the flow behavior in the reservoir.

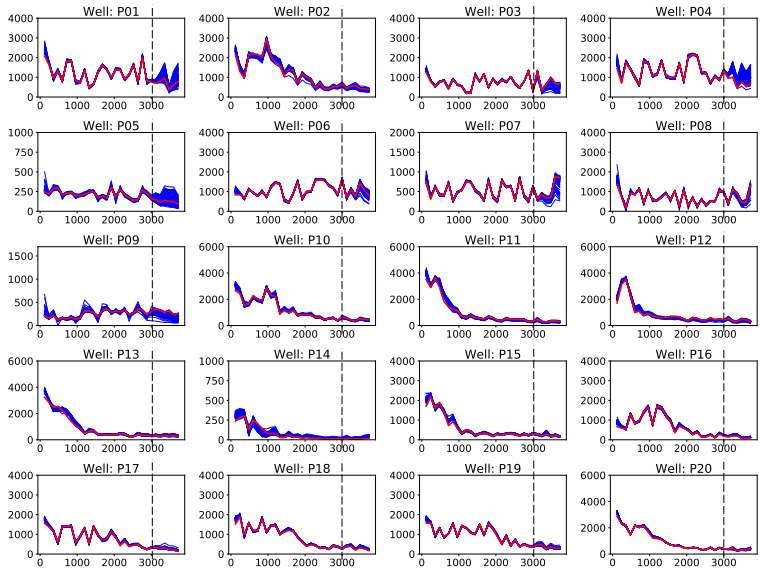


Figure 5.14: The oil rates of the 100 best-fit realizations (blue curves) and the observation data (red curves); the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day); the mean error of training period and prediction period between these 100 realizations of DiWA model of Case 1 (coarse model) and the observation data are 0.59% and 0.62%, respectively

5.3. CONCLUSION AND DISCUSSION

In this chapter, the Discrete Well Affinity (DiWA) data-driven proxy model is proposed and validated. The adjoint framework is implemented in DARTS and applied for calculating the gradients with respect to transmissibility and well indexes. Later, these adjoint gradients are combined with the numerical derivatives with respect to the rock-fluid parameters for the training of proxy models based on gradient optimization. We constructed the Discrete Well Affinity (DiWA) data-driven proxy model with a limited number of degrees of freedom and impressive forward simulation performance. The main idea of the DiWA model is that it uses very coarse gird and very basic geological information to represent a complex reservoir structure and the fluid flow in the reservoir. The coarse grid can be either structured or unstructured. The DiWA model includes geological properties like the average reservoir thickness, porosity, etc. But more geological properties can also be added to the DiWA model if they are available.

This chapter started with the testing of two synthetic fluvial model ensembles FLUMY and MPS. The results showed good prediction accuracy for a significantly reduced model size. Both training and prediction accuracy is within a satisfactory level. In addition, the data-driven proxy methodology was compared with a conventional flow-based upscaling technique and demonstrated an improved accuracy within both fluvial model ensembles.

Next, we tested a simple 3D proxy model with multiple layers and three flowing phases. This model was trained to match observation data generated from a modified

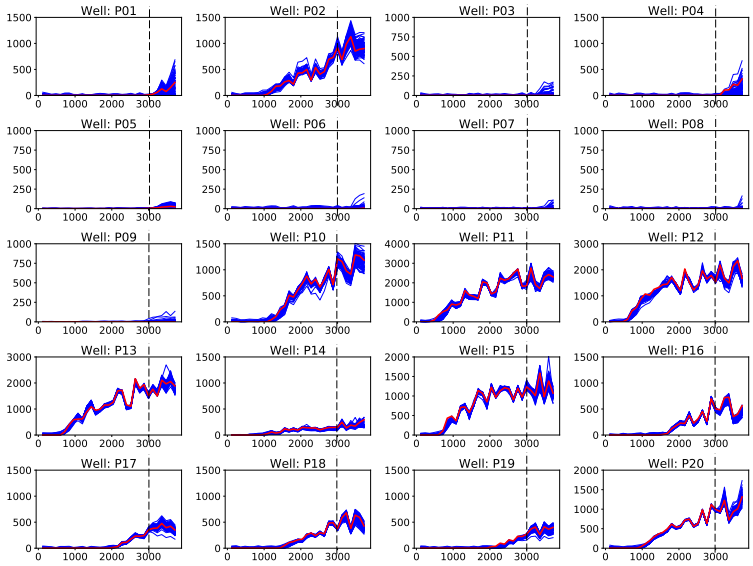


Figure 5.15: The water production rates of the 100 best-fit realizations (blue curves) and the observation data (red curves) for Case 1; the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day)

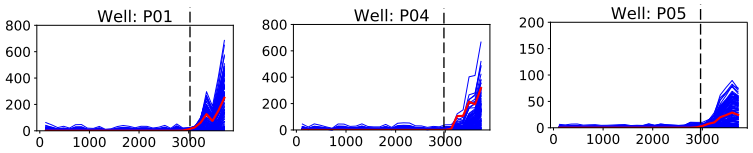
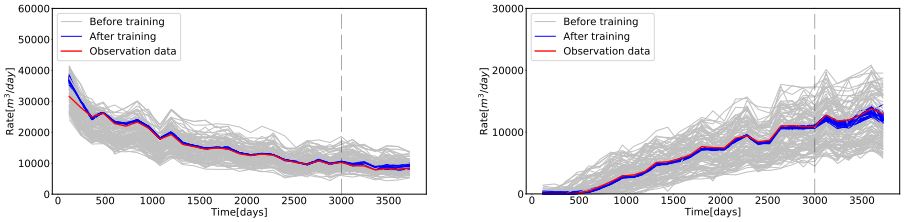


Figure 5.16: The water production rates of the 100 best-fit realizations (blue curves) and the observation data (red curves) for Well: P01, P04 and P05 of Case 1; the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day)



Total oil rate Total water rate

Figure 5.17: The total oil (left) and water (right) production rate of 100 best fit realizations before and after training of Case 5 (fine model)

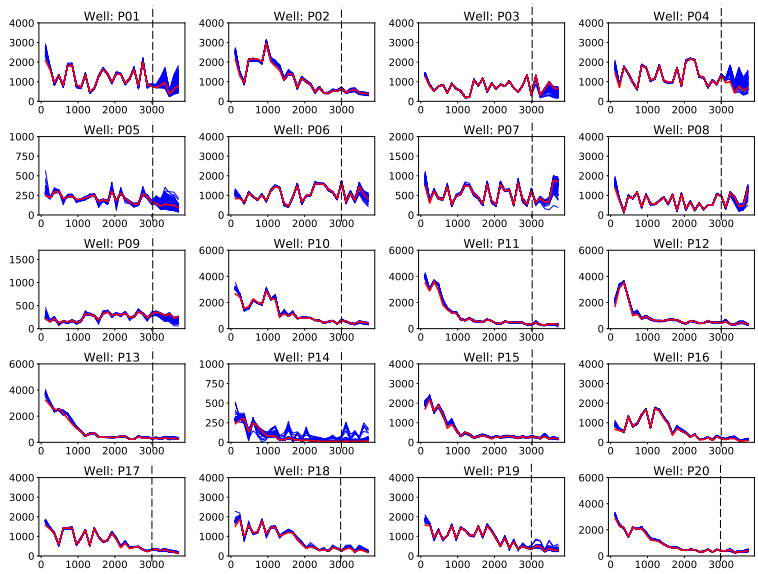


Figure 5.18: The oil rates of the 100 best-fit realizations (blue curves) and the observation data (red curves); the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day); the mean error of training period and prediction period between these 100 realizations of DiWA model of Case 5 (fine model) and the observation data are 0.44% and 0.17%, respectively

SPE 1 model. The results showed that the proxy model can be used to characterize geological information of the reservoir when the initial guess is generated based on reliable geological information. But this may not be true when the initial guess for the model is far from the true geology. This inspires us to introduce an efficient sampling approach where each statistical member has been trained to match the true data.

A high-fidelity Brugge model was utilized to generate true data for model training and test the performance of the proposed approach. In the procedure of true data generation, uniformly distributed random perturbations were added to BHP control. The generated true data were used then to train the DiWA model. The proxy model used in this study only contains some basic geological information like the contour and the average thickness of the original reservoir model. However, more information can be added to this framework to improve the training process with appropriate regularization.

The proposed framework has several benefits. Firstly, the OBL technique helps to increase the efficiency of the assembly of the residual and the Jacobian in both forward simulation and adjoint gradients calculation. Secondly, the performance of the training improved significantly since the cost of the adjoint gradient calculation is almost equivalent to a single forward run, which makes feasible a proposed stochastic DiWA proxy methodology. Furthermore, if more model parameters are introduced to the model, a higher convergence efficiency of the adjoint method can be obtained compared with the numerical gradients approach.

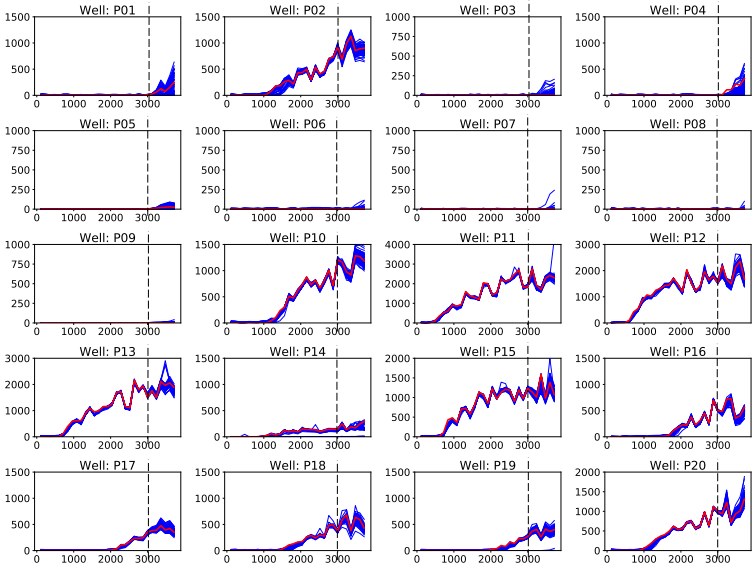


Figure 5.19: The water production rates of the 100 best-fit realizations (blue curves) and the observation data (red curves) for Case 5; the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day)

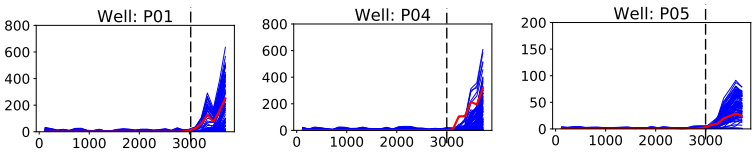


Figure 5.20: The water production rates of the 100 best-fit realizations (blue curves) and the observation data (red curves) for Well: P01, P04 and P05 of Case 5; the x-axis shows the production time (days), the y-axis shows the oil production rate (m^3/day)

6

EFFICIENT INVERSE MODELING FOR ENERGY TRANSITION APPLICATIONS

Summary

In this chapter, the extension of the inverse modeling feature of DARTS is presented, aimed at enhancing its versatility in energy transition projects. Specifically, two applications are discussed in detail. The first application concerns the history matching of the geothermal reservoir. In this project, diverse observations, including well rates, well temperature, and reservoir temperature inferred from electromagnetic data, are incorporated into the objective functions. The second example involves the history matching of the CO_2 storage project. Both examples highlight the significance of uncertainty quantification and the employment of inverse modeling strategies for complex systems.

Parts of this chapter have been published in Geothermics (2023) [101] and in Transport in Porous Media (2023) [104].

6.1. BAYES' THEOREM

The inverse problem is almost never well posed [80], especially for reservoir engineering problems. This is because uncertainties, noises, or even errors could be from anywhere in the field measurement, physical assumptions, theoretical models, and so on. For example, the pressure measurement in the field is usually subject to noise. The measurement of the fluid velocity may be a bit more reliable than the pressure measurement, but the difference between the reservoir and surface separator conditions may still cause the low quality of the velocity measurement. The interpretation of the petrophysical data (e.g. permeability, porosity, etc.) is another source of uncertainties. Differing from the uncertainty generated in the field, the inexact description of the system caused by the physical assumptions and theoretical models results from the limitations of the computational capacity or our knowledge about the underground system.

Therefore, a real inverse modeling problem can be described as follows: find the optimal model parameters given the observations with noise (e.g. the pressure data with noise), preset assumptions (e.g. neglecting capillary pressure or assuming isothermal condition), and the inexact theoretical model (e.g. different choices of Equations of State). As it is mentioned previously, the inverse problem is usually ill-posed. This means that there will be multiple solutions of the model parameters describing the system under given tolerance. Apparently, we will not just choose one of the solutions and simply drop the rest of them. But how to solve and quantify them? From Bayes' theorem's perspective and point of view, the prior information should be incorporated with the observations when inferring those model parameters. It follows that

$$f(u \mid \mathrm{d_{obs}}) = \frac{f(\mathrm{d_{obs}} \mid u)\, f(u)}{f(\mathrm{d_{obs}})} \tag{6.1}$$

where $f(\mathrm{d_{obs}})$ is the probability density function (PDF) of d_{obs}, also known as evidence. $f(u)$ is the PDF of prior information. $f(\mathrm{d_{obs}} \mid u)$ is the likelihood given the condition of model parameter u. $f(u \mid \mathrm{d_{obs}})$ is the PDF of posterior given the observation d_{obs}.

By modifying the prior PDF $f(u)$, the likelihood $f(\mathrm{d_{obs}} \mid u)$ is gradually maximized under the given criteria. The modified prior PDF is called the posterior PDF $f(u \mid \mathrm{d_{obs}})$. From the viewpoint of Bayesian estimation, explicit knowledge of $f(\mathrm{d_{obs}})$ is unnecessary [80], as it is independent of u. In practical situations, however, it is still difficult to explicitly describe such PDFs mathematically. Instead, we usually estimate the posterior uncertainty of the model by sampling from the ensemble of realizations after history matching. From Bayesian extensions to the usual theory of parametric bootstrapping, it can be shown that a sample may be generated by solving

$$u_c = \arg\min_u E(u) \tag{6.2}$$

where $E(u)$ is defined by

$$E(u) = \frac{1}{2}(u - u_{ref})^T C_M^{-1}(u - u_{ref}) + \frac{1}{2}(G(u) - \mathrm{d_{obs}})^T C_D^{-1}(G(u) - \mathrm{d_{obs}}), \tag{6.3}$$

where C_D^{-1} is the inverse of the covariance matrix of the combined measurement and forward modeling error process, and u_{ref} is a sample from the geostatistical prior. Note

that compared with the Equation (3.18) that describes the misfit between the model response and observations, Equation (6.3) has an extra term of $\frac{1}{2}(u-u_{ref})^T C_M^{-1}(u-u_{ref})$ that describes the difference between the updated model parameter and prior information, which is also known as regularization in history matching problem. In the next two sections about the geothermal and CO_2 storage projects, Bayes' theorem will be applied to conduct the uncertainty quantification.

6.2. APPLICATION OF GEOTHERMAL PROJECT

In this section, an efficient and flexible adjoint-based framework for history matching and forecasting geothermal energy extraction is presented. In this framework, we applied the Principal Component Analysis to reduce the parameter space for representing the complex geological model. The adjoint method is implemented for gradient calculation to speed up the history-matching iteration process. Operator-based linearization (OBL) used in this framework makes the calculation of the physical state and its derivatives very efficient and facilitates the matrix assembly in the adjoint method. This study primarily focuses on history matching based on combined observation of well production and in-situ electromagnetic measurements to predict the temperature front. However, different types of misfit terms can be added to the objective function based on practical considerations. For example, our history-matching case studies include model misfit terms applied for regularization purposes. The measurement data is extracted from the true model, and realistic measurement errors are considered. Also, in this work, we propose an optimal weighting strategy for the terms of the objective function to balance their sensitivity with respect to the model parameters. The high efficiency of the framework is demonstrated for the geothermal doublet model implemented at the heterogeneous field model with multiple realizations. The framework allows for generating posterior Randomized Maximum Likelihood (RML) estimates of the entire ensemble of the realizations with a reasonable computational cost. Results show that the framework can achieve reliable history-matching results based on the doublets production data and the reservoir electromagnetic measurement.

6.2.1. INTRODUCTION

Numerical simulation is an important and essential tool for developing geothermal resources. The most accurate resource estimations are achieved with high-fidelity simulations performed with physical models based on the first principles of mass and energy conservation. In order to calibrate such physical models, their response should be matched to geothermal field observations, which includes the observed well temperature, well flow rate, time-lapse data, etc. This calibration process is therefore called history matching and comprises the adjustment of uncertain geological model parameters, e.g., permeability, porosity, and thermal conductivity. Regarding geological uncertainty, the Bayesian approach is often adopted, taking into account all available prior knowledge, such as the hard data obtained from samples, geological formation scenarios, and assumptions on the statistical distribution. In this context, history matching has the goal to assimilate observed data and the prior uncertain model to generate a posterior estimate of the uncertain model. Typically, the observed data comprises the well water rates,

bottom hole pressure (BHP), and well temperature. To the best of our knowledge, limited research has been conducted on the history matching of reservoir temperature distribution as interpreted from electromagnetic data, though there is some research about the history matching for the electromagnetic data in hydrocarbon reservoirs, for example, [117, 118].

Data assimilation can be performed in different ways, with methodologies having varying complexity and cost in terms of the number of numerical simulations required. Some methodologies, such as posterior sampling with Monte Carlo simulations, are straightforward to implement but are prohibitively expensive in conjunction with geothermal physical models. A more accessible way is to formulate and solve the history matching as an optimization problem where the observed historical data misfit is minimized at the same time as the statistical likelihood of the model parameters is maximized. This methodology has been applied in numerous works for the full statistics in the form of the Randomized Maximum Likelihood (RML) approach see for example [80, 97] as well as for the reduced statistics in the form of Maximum A-Posteriori (MAP) estimate approach, see for example [8, 106]. In this paper, we adopt the RML approach for Electro-Magnetic (EM) monitoring and make it more efficient by coupling it with parameter space reduction as well as fast optimization and simulation techniques.

When considering the prior information in the history matching, the covariance matrix inversion is a big issue. This issue is even more pronounced in high-fidelity geological models. In this case, the dimension of the history matching problem needs to be reduced so that the problem can be solved in a lower dimensional space. This is reasonable because the history matching problem is usually over-parameterized for large models [8], as the amount of observation data is usually much smaller than the number of model parameters.

Many methods for reducing the dimension have been widely investigated and applied in the field of geothermal, petroleum, and groundwater hydrology engineering. Upscaling is one of the most often used methods to reduce the number of model parameters. In this method, a coarse grid model is applied to represent the high-resolution model by assigning the effective properties in each coarse cell [20]. [87] investigated the upscaling of thermal conductivity of sedimentary formations for geothermal exploration. The results show that the harmonic averaging strategy is more accurate than other averaging methods. More upscaling approaches of hydraulic conductivity, permeability, and other petrophysical properties are analyzed in the reviews and studies of [90, 75], and [114]. However, consistent upscaling of geothermal models with convective and conductive flow still poses significant challenges [83]. Similarly, [38] and [46] introduced the multi-scale method designed to efficiently capture the large-scale behavior of the solution without resolving all small-scale features.

All of the aforementioned methods achieve high computational efficiency because of the reduced dimension that physically coarsens the model grid cells. Mathematically speaking, the degrees of freedom of the model can also be reduced by using fewer "key" components in the space, while these "key" components are mostly independent of each other and they represent and capture the most information of the original space. Principal Component Analysis (PCA) is a powerful tool to quantify the dependency among the components and boil down the "key" components under a specific criterion. PCA has

been successfully applied in many history matching problems for multi-Gaussian fields [26, 79, 93]. However, [92] pointed out that the direct use of PCA on non-Gaussian fields may lead to "Gaussian-looking" models after history matching. Therefore, they introduced kernel principal component analysis (KPCA) to solve this problem. [67] had later refined the KPCA approach. Although the KPCA approach focuses more on representing the multiple "feature" space in the reservoir, this method is essentially strongly nonlinear and brings challenging numerical issues. [106] proposed an optimization-based PCA method for the low-dimensional parameterization of complex geological models. [8] also introduced PCA parameterization in their gradient-based optimization framework for the closed-loop reservoir management problem. In their study, the adjoint method is applied and incorporated with the PCA parameterization to compute the new gradient in the reduced-dimension space.

In this project, the geothermal history matching and prediction using the reduced dimension technique are investigated. The PCA is applied to determine the dimension size of the reduced parameter space and find out the principal components to represent the reservoir model. We implement this framework in DARTS, which has the feature of simulating the geothermal developing process [112]. We also extend the DARTS geothermal engine with the inversion feature of the adjoint method for gradient calculation in this study. To make the history-matching framework more flexible, various types of misfit terms can be added to the objective function based on practical considerations. We also propose an optimal weighting strategy for the terms of the objective function to balance their sensitivity with respect to the model parameters. The heterogeneous geological models of the Egg field with 100 realizations [45] are utilized in this study to demonstrate the performance and high efficiency of this framework.

6.2.2. PCA-BASED PARAMETERIZATION

In the previous section, we mentioned that the adjoint method is introduced as an efficient approach for the evaluation of gradient. Nevertheless, the high dimensionality and uncertainty inherent in geological models still present notable obstacles to the optimizer's search for the global minimum. As previously noted, these obstacles can be addressed through the utilization of PCA approach, which enables the transformation of the original space into a reduced space. The uncertainty of the reservoir is usually quantified by an ensemble of geological realizations. Initially, the ensemble is generated based on all available prior geophysical information (e.g., hard data from the rock samples, seismic measurements, etc.). Then, the history matching process reduces further uncertainty. The uncertainty and the correlations among these realizations are described by a covariance matrix of the model parameters. Mathematically speaking, as long as the covariance (i.e., the off-diagonal elements in the covariance matrix) is nonzero, it means those model parameters depend on each other and, therefore, can be represented using fewer "independent" model parameters in the parameter space. From the perspective of the history matching problem, the amount of observation data is much smaller than the number of unknown model parameters, so the history matching problem is over-parameterized for large models [8]. This also indicates that the history matching problem is nonconvex, and the optimal solution is nonunique. The history matching, therefore, requires regularization term as shown in Equation (6.3), and it

writes:

$$R(u) = \alpha_R (u - u_{ref})^T C_M^{-1} (u - u_{ref}), \tag{6.4}$$

where α_R is the scaling coefficient, C_M^{-1} is the inverse of the covariance matrix, and u_{ref} is the reference of the model parameters.

As mentioned above, the model parameter space can be represented by using fewer "independent" variables in the new space. Here we utilize the Principal Component Analysis (PCA) to project the original ensemble of prior realizations into a new space characterized by a set of linearly uncorrelated variables. We denote the model parameters in the new space as ξ. The procedure of mapping the original model parameters u to ξ is described below. More details can be found in [106, 92].

The covariance matrix can be calculated by using the following equation:

$$C_M = \frac{XX^T}{N_r - 1}, \tag{6.5}$$

where N_r is the total number of realizations, and X is given by:

$$X = [u_1 - \bar{u}, ..., u_{N_r} - \bar{u}]. \tag{6.6}$$

X is a $N_u \times N_r$ matrix, where N_u is the number of the model parameters in a single realization. The $\bar{u}$ in Equation (6.6) is the mean of the ensemble of the realization, and here we assume a Gaussian distribution for the model parameters u. Instead of directly decomposing C_m, we perform Singular Value Decomposition (SVD) on the matrix $y = X/\sqrt{N_r - 1}$ because of the decomposition efficiency. The factorized y using SVD is given:

$$y \approx \tilde{U}_{N_\xi} \tilde{\Sigma}_{N_\xi} \tilde{V}_{N_\xi}^T, \tag{6.7}$$

where $\tilde{\Sigma}$ is a diagonal matrix that contains the singular value of y, $\tilde{U}$ and $\tilde{V}$ are the unitary matrices that contain the left- and right-singular vectors of y respectively, the suffix N_ξ denotes the dimension size of the new parameter space (i.e. ξ-space) after the truncation to $\tilde{U}$, $\tilde{\Sigma}$, and $\tilde{V}$. The method of the prescribed portion of the variance (energy) contained in eigenvalues is used to determine N_ξ:

$$\frac{\sum_{k=1}^{N_\xi} \sigma_k^2}{\sum_{k=1}^{N_\Sigma} \sigma_k^2} \geq 90\%, \tag{6.8}$$

where σ_k is the square root of the eigenvalue (i.e. the element along the diagonal of $\tilde{\Sigma}$), N_Σ is the total number of the elements along the diagonal of $\tilde{\Sigma}$. Here we prescribe 90% as the energy portion value when determining the dimension of ξ-space. Once the dimension size N_ξ is determined, several important transformation matrices can be defined:

$$\Phi = \tilde{U}_{N_\xi} \tilde{\Sigma}_{N_\xi}, \tag{6.9}$$

$$\hat{\Phi}^{-1} = \tilde{\Sigma}_{N_\xi}^{-1} \tilde{U}_{N_\xi}^T, \tag{6.10}$$

$$\Psi = \tilde{\Sigma}_{N_\xi} \tilde{U}_{N_\xi}^T. \tag{6.11}$$

The formulations of mapping between the original space and ξ-space are then written:

$$u = \Phi\xi + \bar{u}, \tag{6.12}$$

$$\xi = \Phi^{-1}(u - \bar{u}) \approx \hat{\Phi}^{-1}(u - \bar{u}), \tag{6.13}$$

$$\nabla_\xi \mathcal{J} = \Phi^T \cdot \nabla_u \mathcal{J} = \Psi \cdot \nabla_u \mathcal{J}, \tag{6.14}$$

where u and ξ are the model parameters in the original space and the reduced-dimension ξ-space, respectively. $\nabla_\xi \mathcal{J}$ and $\nabla_u \mathcal{J}$ are the gradients of misfit term $\mathcal{J}$ with respect to the model parameter in the reduced-dimension ξ-space and the original space, respectively.

The history matching based on Randomized Maximum Likelihood (RML) takes different samples of u as the u_{ref}. Therefore, the regularization terms Equation (6.4) in the reduced-dimension ξ-space should be written as:

$$\begin{aligned} \alpha_R(u - u_{ref})^T C_M^{-1}(u - u_{ref}) &= \alpha_R[\tilde{U}\tilde{\Sigma}(\xi - \xi_{ref})]^T(\tilde{U}\tilde{\Sigma}^{-2}\tilde{U}^T)[\tilde{U}\tilde{\Sigma}(\xi - \xi_{ref})] \\ &= \alpha_R[(\xi - \xi_{ref})^T\tilde{\Sigma}\tilde{U}^T](\tilde{U}\tilde{\Sigma}^{-2}\tilde{U}^T)[\tilde{U}\tilde{\Sigma}(\xi - \xi_{ref})] \\ &= \alpha_R(\xi - \xi_{ref})^T(\xi - \xi_{ref}), \end{aligned} \tag{6.15}$$

where

$$\xi_{ref} = \hat{\Phi}^{-1}(u_{ref} - \bar{u}). \tag{6.16}$$

We will also check how the regularization affects the history-matching results if the model is constrained to the ensemble prior mean. In such case, the model reference u_{ref} is fixed as $\bar{u}$. Therefore the regularization term is:

$$\alpha_R(u - u_{ref})^T C_M^{-1}(u - u_{ref}) = \alpha_R \xi^T \xi. \tag{6.17}$$

6.2.3. MODELING OF ELECTROMAGNETIC RESPONSE

In this section, we will present the application of the proposed framework on a 2D fluvial geothermal model. The electromagnetic monitoring data will be included in the objective function for history matching. One of the synthetic geological layers from the West Netherlands Basin [18, 95] is taken as a 2D fluvial model. This 2D model is discretized into 2400 blocks ($60 \times 40 \times 1$) with a block size of 30 m × 30 m × 2.5 m. The permeability and porosity maps are shown in Figure 6.1.

The model runs for 40 years with an injector and a producer in the reservoir. The reservoir has a uniform initial pressure of 200 bars and a uniform initial temperature of 348.15 K. The thermal conductivity and the volumetric heat capacity of the reservoir rock are 200 kJ/m/day/K and 2200 kJ/m^3/K, respectively. The water is injected with a BHP control of 300 bars at 308.15 K. The producer has the BHP control of 50 bars.

The real time-lapse electromagnetic (EM) monitoring data is not available for this case study. Instead, we create an approximation of the temperature response interpreted from EM monitoring data. We will call this approximation of the temperature response as "EM data" in the following sections. The EM data will be used in the objective function

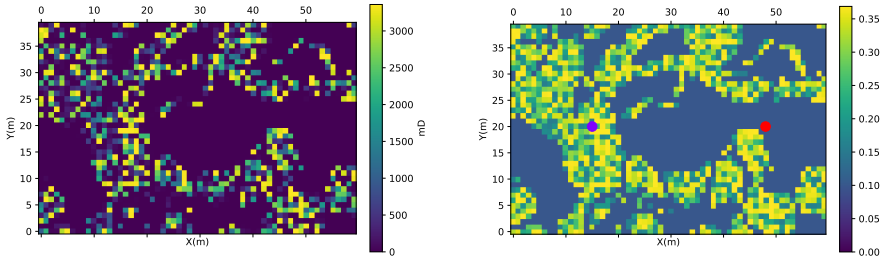


Figure 6.1: The permeability map in the x direction (left) and the porosity map (right) of 2D fluvial model. The purple and red dot in the right figure represent the well locations of the injector and the producer, respectively

f_4 of Equation (3.21). For that, we collect the simulated temperature data and blur it by applying a spatial filtering kernel of the size that is consistent with the resolution scale of typical EM measurements. This synthetic EM data will be treated as the observations d_{obs}. The steps are:

- Generate the true temperature data d_{true} based on the true model
- Compute the spatial filtering kernel size based on the r^2 to the EM receivers, where r is the distance between the given point and the EM receiver location. In this 2D model, we set the EM receiver at the location of the production well.
- Apply the spatial filtering strategy to the true temperature data based on the kernel size at different locations in the reservoir.

The schematic of generating time-lapse EM data d_{obs} is shown in Figure 6.2. The measurement error matrix C_D of EM data is computed based on the square of the difference between d_{true} and d_{obs}.

Since the 2D model uses the BHP control as the well control, the misfit term of the objective function Equation (3.18) consists of f_1, f_3, and f_4 from Equation (3.21). The model is history matched based on the first 20 years of production and forecast for the next 20 years. The training and forecast curves are shown in Figure 6.3. The results of the time-lapse temperature data are demonstrated in Figure 6.4. As it can be seen from Figure 6.3, the trained well rates and temperature match with the observation data very well, though the well temperature of the forecast period deviates a bit from the observation data. Figure 6.4 also indicates a good history matching result of the time-lapse temperature data.

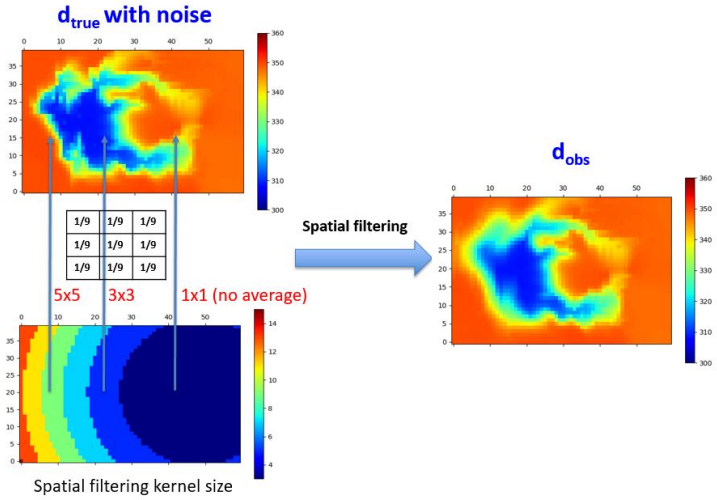


Figure 6.2: The schematic of generating synthetic time-lapse EM data

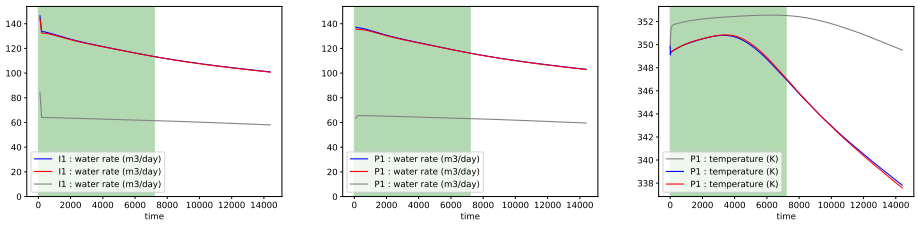


Figure 6.3: The training and forecast results of the injection well rate (left), production well rate (middle), and production well temperature (right), respectively. The red curves are the observation data. The gray and blue curves are the results before and after model training. The light green areas demonstrate the training periods while the rest areas are the forecast periods.

Note that the CPU time (Intel CPU i7-8556U) for a single forward simulation takes 2.1 seconds, while the gradient calculation using the adjoint method only takes 2.5 seconds for 4702 model parameters in a single history-matching iteration. This means that the computational time used for computing gradient is comparable to the CPU time of a single forward simulation, which is much less than the computational time of 4702 forward simulations used in the conventional numerical gradient calculation. This shows that the proposed framework based on the adjoint method has a significant improvement in the efficiency of the history matching problem. In the next section, we will present a more complex 3D example of the Egg reservoir and introduce the dimension reduction technique based on the Principal Component Analysis method.

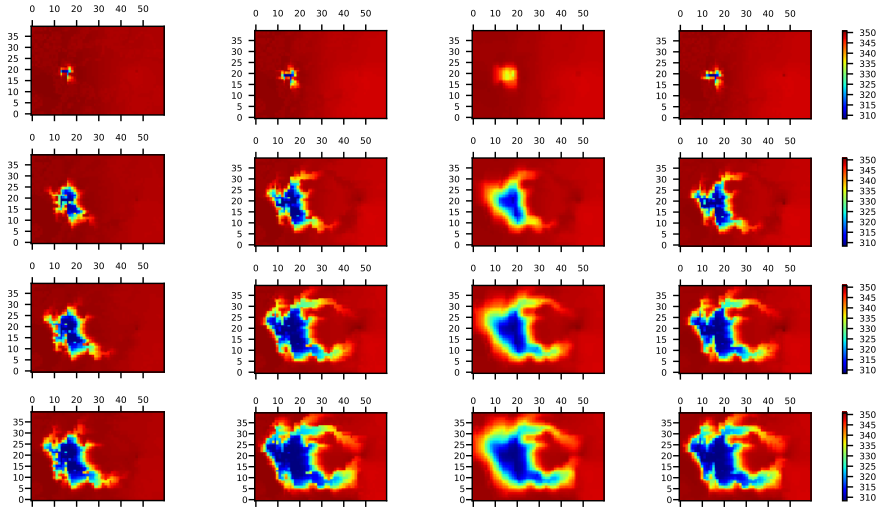


Figure 6.4: The temperature distributions at different time steps. From the top to the bottom rows, they represent the temperature distributions at the 100, 2000, 4000, and 6000 days respectively. From the left to right columns, they are the result before history matching, result after history matching, EM data (observation), and true data, respectively.

6.2.4. GEOTHERMAL RESERVOIR

In this section, a large ensemble of realizations of the Egg model [45] (but without egg shape) will be used in this study to demonstrate the history matching framework with PCA-based parameterization. Generally, when the degrees of freedom of the model are not so high, we can directly solve the history matching problem without using the dimension reduction technique. However, the high-resolution model usually brings a large number of degrees of freedom in the history matching problem. In this case, it may introduce redundant degrees of freedom to the problem. Moreover, if we try to conduct the history matching directly on the high-resolution model, it is often limited by the computer resources and capacities. With the concerns of the redundant degrees of freedom and the limitation of the computer capacity in the complex history matching problem, we introduce the dimension reduction technique into the proposed history matching framework.

We will use the Egg model to demonstrate the process of dimension reduction in history matching. The Egg model is an open-access geological reservoir ensemble [45]. This model consists of an ensemble of 100 permeability realizations of a three-dimensional channelized reservoir. The term "realization #n" hereafter denotes the nth realization of the original set. The realization #1 will be considered a "true" model to generate the observation data. The rest of the 99 realizations will be used as priors to train the Egg model. There are 60×60×7 blocks with the block size of 30m×30m×12m in the model. The permeability data are imported from the ensemble realizations. The permeability in the x and y directions are identical, while the permeability in the z direction is 0.1 times the horizontal permeability. The uniform properties of porosity, pressure, initial reser-

voir temperature, the volumetric heat capacity of rock, and the thermal conductivity of rock are 0.2, 200 bar, 348.15 K, 2200 kJ/m^3/K, 181.44 kJ/m/day/K, respectively. A doublet (one injector and one producer) is set in the reservoir with all seven layers perforated, see Figure 6.5. The true model "realization #1" will run for 40 years to generate the observation data. The first 30 years are taken as the training period, and the rest 10 years are the forecast period to test the performance of the proposed method.

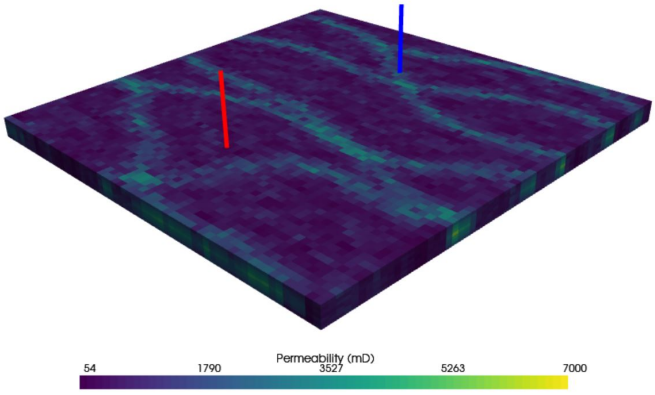


Figure 6.5: The permeability distribution of realization #1. The blue and red bars represent the location of the injector and producer, respectively

Following the dimension reduction process described in Section 6.2.2 and Equation (6.8), the dimension of the model parameter space of transmissibility is reduced to 82. The total number of the dimension is therefore 82 plus 14 (well indexes). This is much less than the original dimension of the transmissibility (i.e., 71160 block interfaces) and well indexes (i.e., 14 perforation positions) model parameter space. The time-lapse electromagnetic data are also included in the objective function to train the model. The history matching process of most of the priors can be finished within 8 hours. The total training time of 99 realizations can be finished within 10 hours on DelftBlue cluster [1] with multiple nodes and cores with 2x Intel XEON E5-6248R 24C 3.0GHz processors.

THE TRAINING BASED ON MAXIMUM LIKELIHOOD ESTIMATION

In the Maximum Likelihood Estimation (MLE) method, we will show the history matching results without the consideration of regularization. This means, compared with Bayes' theorem, only the likelihood will be maximized, while the prior information is not considered.

The history matching results of well temperature are shown in Figure 6.6. Except for a few realizations that are relatively far from the observation data (red curve), most of the trained realizations have good history matching results (the green area in the figure) to the observation data. Although the forecasting results (the white area in the figure) demonstrate a wider range of well temperature, this is expected and acceptable for the forecasting period. Here we introduce Root Mean Square Error (RMSE) to quantify the

error of the results. The expression of RMSE is given as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (G - \mathrm{d_{true}})^2}{N}} \tag{6.18}$$

where G is the model response and $\mathrm{d_{true}}$ is the true data. For each realization, the RMSE can be calculated based on Equation (6.18). The ensemble average RMSE for MLE is 0.253.

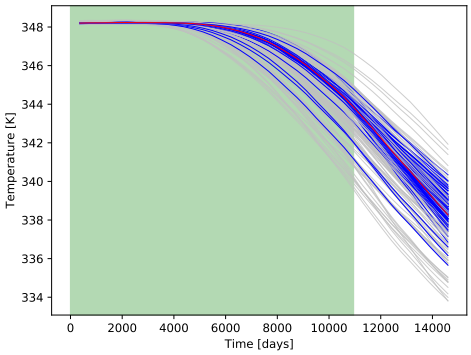


Figure 6.6: The history matching results of the well temperature for the entire ensemble of the Egg model based on MLE. The red curve is the observation data of well temperature. The gray and blue curves are the results before and after history matching, respectively. The green area shows the training period, while the white area is the forecasting period.

The history matching results of reservoir temperature data are shown in Figure 6.7. This figure shows the temperature results of the 4th layer of realization #66. It can be seen that the temperature results after the history matching have a good match with the EM observation data, even though the original temperature distribution is very different from the observed EM data. This indicates that the optimizer might have largely changed the original permeability distribution of realization #66 to try to match the EM observation data. Therefore, we plot the permeability distribution of this realization and the "true" permeability distribution in Figure 6.8.

As it can be seen from figure (d) in Figure 6.8, the optimizer tries to block the original fluvial channel (blue channelized area), while re-construct the channel of the "true" model (red channelized area). This is because, with the information given by observed EM data, the adjoint gradient is able to capture the characteristics of the "true" permeability field, and therefore re-construct the temperature map that is similar to the observed EM data.

In the next section, we will repeat this history-matching procedure, while considering the prior information of the realization permeability based on Bayes' theorem. The reduced-dimension technique based on PCA is also introduced.

THE TRAINING BASED ON RANDOMIZED MAXIMUM LIKELIHOOD

In the Randomized Maximum Likelihood method (RML), the prior information is considered and the regularization term chooses each realization from the ensemble as the

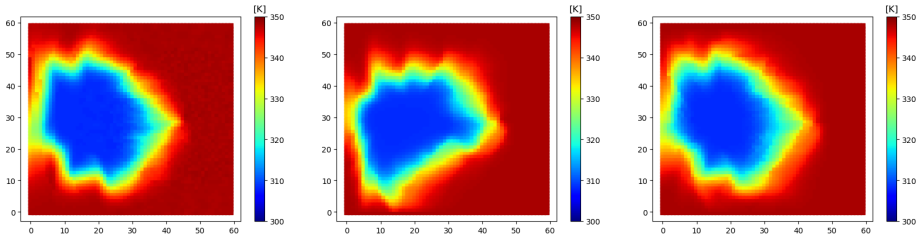


Figure 6.7: The history matching results of the reservoir temperature of the 4th layer of the Egg model. The left figure shows the EM observation data of the "true" model. The middle and right figures are the results of realization #66 based on MLE.

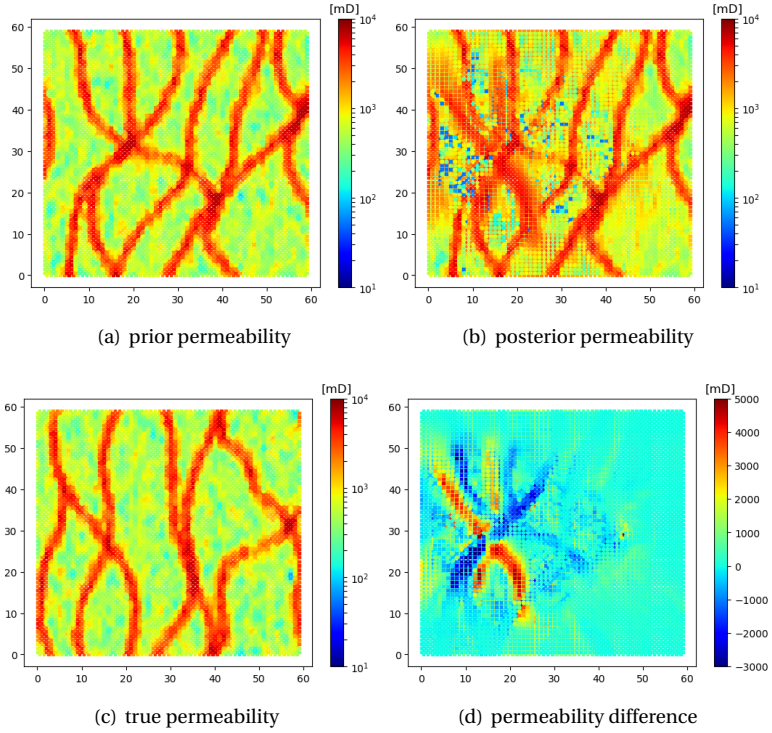


Figure 6.8: The permeability distribution of the 4th layer of Egg model. Figure (a) demonstrates the prior permeability of realization #66; (b) is the posterior permeability; (c) shows the "true" permeability; and (d) demonstrates the permeability difference between posterior and prior based on MLE.

reference model. Therefore, the regularization term in the reduced-dimension ξ-space is written as Equation (6.15).

Firstly, we conducted a comparison of RML in the original full space and reduced-dimension ξ-space to check the feasibility and the accuracy of the reduced-dimension

technique based on PCA. As it is shown in Equation (6.4), the RML in the original full space requires the inversion of covariance matrix C_M of model parameters. In the Egg reservoir model, the model parameters have a dimension of 71160 (i.e. the transmissibility of the interfaces of reservoir blocks). This indicates large amounts of computational cost and memory consumption for the matrix inversion. To inverse a non-positive definite covariance matrix, the pseudo-inverse function from Python Numpy linear algebra package is chosen to inverse C_M. This inversion procedure needs around 500 GB of memory and takes more than one day to finish the computation. Finally, the output inversed covariance matrix C_M^{-1} has a size of 39.6 GB and can be stored for further use in each realization in the ensemble.

The history matching results of well temperature are shown in Figure 6.9. It is evident to see that the history matching has a more divergent result in the original full space, compared with the result in the reduced-dimension ξ-space. Their ensemble average RMSE is 0.564 and 0.314, respectively. It indicates that the reduced-dimension technique based on PCA is capable of generating more accurate history-matching results. Moreover, this technique requires much less machine memory and computational time to evaluate the regularization term.

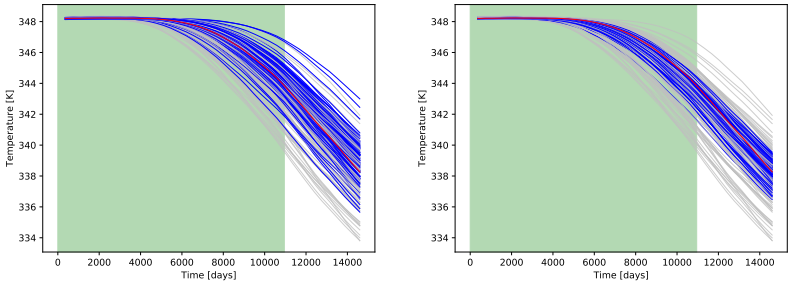


Figure 6.9: The history matching results of the well temperature based on RML under the original full space (left figure) and reduced-dimension ξ-space (right figure). The red curve is the observation data of well temperature. The gray and blue curves are the results before and after history matching, respectively. The green area shows the training period, while the white area is the forecasting period.

Compared with the history matching based on MLE shown in the previous section, the well temperature curves of the entire ensemble based on RML (the right figure in Figure 6.9) have a wider distribution range around the observation data curve (red curve) and a larger ensemble average RMSE. This is because the regularization term forces the optimizer to search the optimum around the prior, instead of searching in an area that is far away from the prior in the parameter space. From Bayes' theorem perspective, this history-matching procedure considered both the prior knowledge and the likelihood maximization.

To compare with the history-matching results of the previous section, we also take the 4th layer of realization #66 as an example to show the history-matching results of reservoir temperature data and permeability map based on RML. They are demonstrated in Figure 6.10 and Figure 6.11.

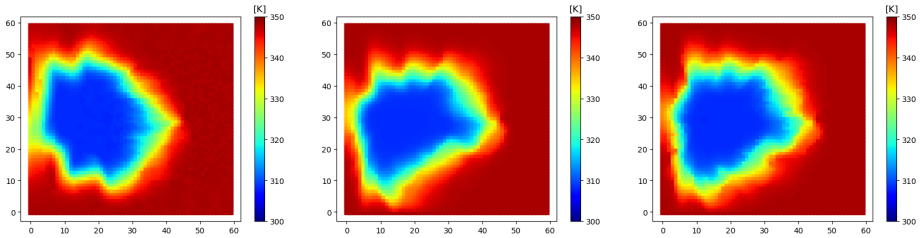


Figure 6.10: The history matching results of the reservoir temperature of the 4th layer of Egg model. The left figure shows the EM observation data of the "true" model. The middle and right figures are the results of realization #66 based on RML.

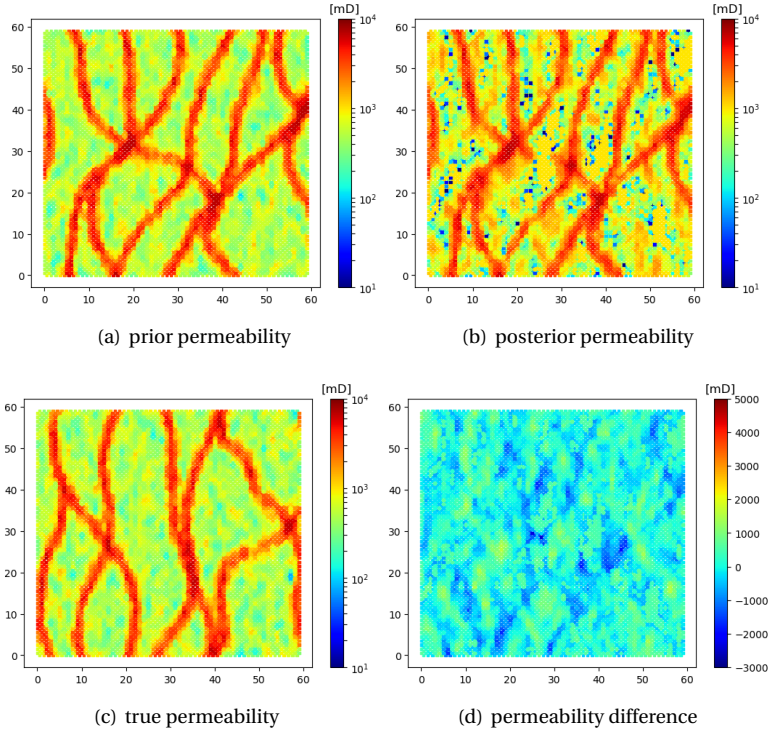


Figure 6.11: The permeability distribution of the 4th layer of Egg model. Figure (a) demonstrates the prior permeability of realization #66; (b) is the posterior permeability; (c) shows the "true" permeability; and (d) demonstrates the permeability difference between posterior and prior based on RML.

It can be seen from Figure 6.10 that the history matching result of reservoir temperature based on RML has less similarity to the results of the previous section (i.e. MLE method). This is because, with regularization, the optimizer tends to preserve the prior information instead of severely reforming the prior to maximize the likelihood. This is

also reflected in the permeability difference map (i.e. figure (d)) in Figure 6.11. The permeability distribution is not largely changed compared with the permeability difference map in Figure 6.8, though there are still some slight permeability modifications along the prior fluvial channel (i.e. blueish color in the permeability difference map).

THE TRAINING CONSTRAINED TO ENSEMBLE PRIOR MEAN

We also conducted the history matching to explore the impact of regularization on history matching outcomes when the model is constrained to the ensemble prior mean. In this case, the reference model used in regularization is fixed with the average of the whole realization ensemble. The regularization term in the reduced-dimension ξ-space is therefore shown as in Equation (6.17).

The history matching results of the well temperature are shown in Figure 6.12. The ensemble average RMSE is 0.312. Similarly, we take the 4th layer of realization #66 as an example to show the results of reservoir temperature data and permeability map in Figure 6.13 and Figure 6.14.

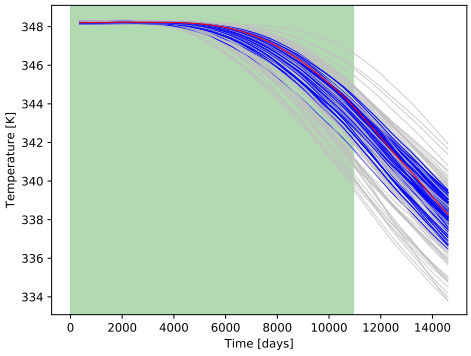


Figure 6.12: The results of the history matching are constrained to the ensemble prior mean under the reduced-dimension ξ-space. The red curve is the observation data of well temperature. The gray and blue curves are the results before and after history matching, respectively. The green area shows the training period, while the white area is the forecasting period.

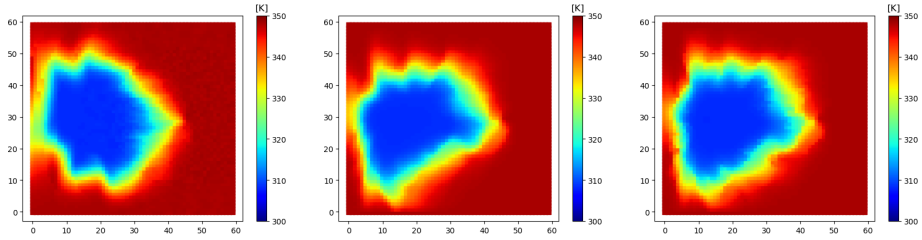
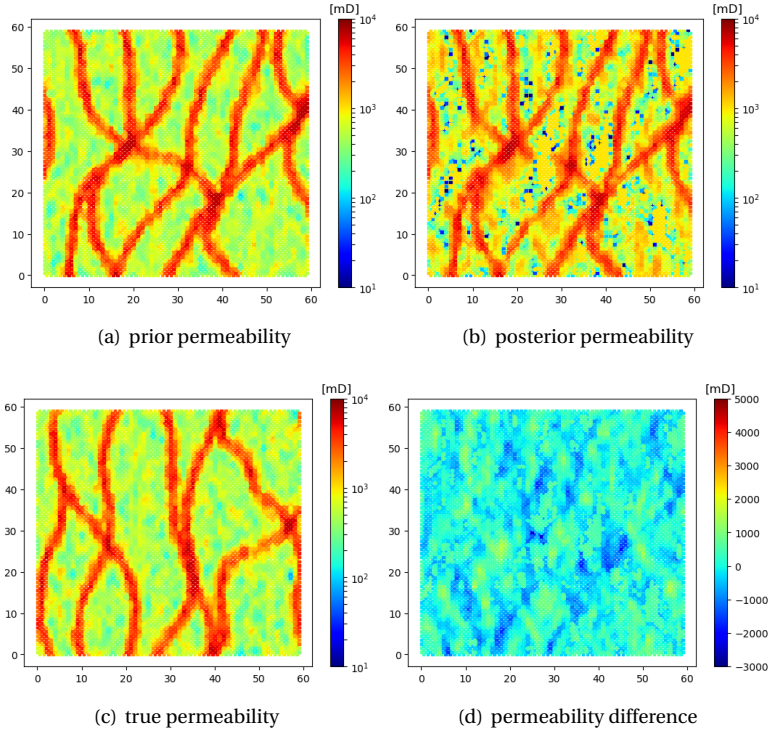


Figure 6.13: The history matching results of the reservoir temperature of the 4th layer of Egg model. The left figure shows the EM observation data of the "true" model. The middle and right figures are the results of history matching constrained to the ensemble's prior mean.



(a) prior permeability (b) posterior permeability

(c) true permeability (d) permeability difference

Figure 6.14: The permeability distribution of the 4th layer of Egg model. Figure (a) demonstrates the prior permeability of realization #66; (b) is the posterior permeability; (c) shows the "true" permeability; and (d) demonstrates the permeability difference between posterior and prior based on the history matching constrained to ensemble prior mean.

	MLE	RML(full space)	RML	Constrained to $\bar{u}$
RMSE	0.253	0.564	0.314	0.312
CPU time	10 hours	> 1 day	10 hours	10 hours
DoF of space	96	71174	96	96

Table 6.1: The RMSE, CPU time, and degrees of freedom of the model parameters for each method.

We do not observe the evident difference between this method and RML. This might be because, with the regularization term, the optimizer is prone to search for the optimum around the initial guess therefore trapped in the local minimum that is close to the initial guess. This phenomenon is as expected, considering the high heterogeneity and the fluvial channels existing in the Egg model. The RMSE, CPU time, and degrees of freedom (DoF) of each method are demonstrated in Table 6.1.

6.2.5. CONCLUSION AND DISCUSSION

In this project, an efficient and flexible adjoint-based history-matching framework for geothermal reservoirs is proposed and developed. To ensure scalability, high-performance optimization and space reduction techniques are introduced and implemented in this framework.

The adjoint method is successfully implemented into the DARTS simulator for the geothermal engine. With the application of the adjoint method, the calculation efficiency and accuracy of the gradient used in the history matching iterations are largely improved. This allows the framework to conduct the history matching with thousands of model parameters. In this project, the history matching of the Egg model with the model parameter number of 71160 can be finished within 10 hours running on an Intel XEON E5-6248R processor.

To represent the complexity of the model while keeping the uncertainty of the model ensemble, the Principal Component Analysis is utilized to transform the model from the original space to the reduced-dimension ξ space, and vice versa. This procedure involves the transformation of the model parameters, gradients, and regularization terms.

The electromagnetic data measurements have been also considered in the proposed history matching framework. Combined with the conventional observations of well rates and BHP, the trained model is able to predict the temperature front. However, it is not necessary to include all types of observations in this history-matching framework. The framework provides the flexibility of adding or removing different types of observations in the objective function based on the availability and necessity of the observation data.

The comparison of the history matching based on MLE, RML, and constrained to ensemble prior mean, are presented in this study to illustrate the performance of the proposed framework. The results show that the adjoint gradient is able to capture the characteristics of the "true" permeability field and re-construct the temperature map that is observed from the EM data. This phenomenon is especially pronounced in the history matching based on MLE. In the history matching based on RML and the case constrained to the ensemble prior mean, the history matching considers both the likelihood maximization and the prior knowledge information. Therefore, the model response has relatively less similarity to the observation, while the prior permeability information is preserved.

6.3. APPLICATION TO CO_2 STORAGE PROJECT

In this section, the inverse modeling framework of DARTS is applied to a new energy transition project of CO_2 storage. we conduct a comprehensive history matching study for the FluidFlower benchmark model, which is an international benchmark study of CO_2 storage project [77]. History matching is first performed based on a low dimensional zonated structured model using a simple Poisson-like solver. The permeability of six facies and two faults are inferred in this stage to match the digitized concentration data. The history matching is then further enhanced to consider the spatial variation of permeability and buoyancy effects. In this stage, the model is switched to an unstructured grid. Efficient adjoint methods are used to evaluate the gradient used in the optimization of data misfits or equivalent Bayesian log-likelihoods. With efficient optimization methods available for both maximum aposteriori model inference and Random-

ized Maximum Likelihood methods for model uncertainty, we perform history matching using both binary and continuous concentration observations. The results show that the tracer plumes match the experimental plumes more accurately compared with the results obtained from the initial stage of model tuning for the parsimonious zonated model. History matching result based on the concentration observations provides more similar shapes of plumes compared with the case based on the binary observations. The permeability difference between the model before and after history matching reveals that the tracer plume zone and the high permeable zone are the regions of high sensitivity in terms of data misfit between the model response and observations.

6.3.1. INTRODUCTION

Escalating greenhouse gas concentrations in the atmosphere and their effect on climate have become an urgent and global concern. Extensive CO_2 emissions are a pervasive aspect of modern industrialized society and its dependence on fossil fuels. Large reductions in CO_2 emissions are mandated by the most recent IPCC reports, and virtually all the conceivable pathways require engineered removal of new carbon pollution sources [43, 42]. Carbon capture and storage (CSS) has been proven to be one of the most promising solutions to this environmental issue. Typically, CCS can reduce 85–90% CO_2 emissions from large point emission sources [60], for example, power plants, cement kiln plants, etc. It is also very likely that carbon-removal technologies – direct air capture or bioenergy with CCS - will be required because emissions of CO_2 cannot otherwise be reduced in time to avoid dangerous global warming [41]. The IEA models require a rapid scale-up of CCS, from 40 Mt p.a. now to a Gt p.a. by 2030 [40].

Geological storage approaches for CO_2 can be classified into several types based on the geological sites and formations: CO_2 injection for enhanced oil recovery [58, 9, 13], unmineable coal bed storage [115], storage in saline aquifers [25, 63, 65], deep ocean storage [39], and in-situ carbonation [69, 94]. Amongst these five CO_2 storage techniques, CO_2 storage in saline aquifers is considered to have enormous potential for the storage of CO_2. This is because saline aquifers can be widely found in both onshore and offshore areas. The usual four CO_2 trapping mechanisms, with varying characteristic time scales, are expected to apply: structural, residual, solubility, and mineral trapping, in increasing order of time scale.

Some CO_2 storage projects that have been conducted in Europe, especially in the region of North Sea [24, 70, 105]. Most of these storage projects are located offshore because these regions are considered to have the majority of the capacity of CO_2 storage in Western Europe. To understand and quantify the flow dynamics of CO_2 storage process, numerical simulation is crucial for modeling the process based on the governing equations and equations of state. However, the numerical simulation of CO_2 storage is challenging because of the complexity of the fluid thermodynamics and the unfolding of multiscale physics.

In this study, we will utilize the DARTS framework designed for modeling energy transition applications [53, 112, 65]. DARTS uses an operator-based linearization (OBL) approach for dealing with highly nonlinear problems resulting from the fully implicit approximation of governing equations [108]. The OBL approach can also help to reduce the computational time of solving complex physical problems by interpolations based

on a few supporting points in the parameter space. Apart from the efficient forward reservoir simulation, DARTS also has an inverse modeling capability based on the adjoint gradient method. Adjoint gradients were first implemented in DARTS for the multi-component multi-phase system of the petroleum-related reservoir simulation problem [100]. This inverse modeling framework was then combined with the Discrete Well Affinity (DiWA) data-driven proxy model. This was further developed to a stochastic DiWA proxy model approach in [102]. Recently, both the forward and inverse modeling capabilities of DARTS were extended into generic energy transition applications.

Here, we use the inverse modeling of DARTS to perform history matching on the experimental results of the tracer test of the FluidFlower benchmark project. The tracer test results are first digitized into component concentration data. Gradient-based history matching is then performed to minimize the difference between the numerical model response and the digitized experimental results. An unstructured grid is used to build the geological model of the experimental deposition, which includes layering and fault interpretations. To represent the uncertainty of the model inferences, the Randomized Maximum Likelihood (RML) method is introduced and a spatial model of heterogeneity in lognormal permeability is added to the geological model. The result shows a good match between the model response and the experimental observations. With the adjoint gradient information, the optimizer is able to capture the characteristics of the fluid flow pattern, especially in the region of tracer plumes where the fluid dynamics are very pronounced.

In the following sections, we will first give a short description of the FluidFlower benchmark setup. Next, a history matching is conducted based on a simple Poisson-like solver. To further improve the history matching results and consider the permeability uncertainty and buoyancy effect, a history matching using two-stage approaches is presented. The results based on the binary and gradual reconstruction of experimental observations are also compared.

6.3.2. FLUIDFLOWER BENCHMARK DESCRIPTION

The FluidFlower experimental rig [77] comprises an engineered heterogeneous sand pack assembled within a thin (25mm) vertical Hele-Shaw cell, about 2.8m wide by 1.3m high. The layered heterogeneous structure is filled using sands filtered into different grain sizes: the sand is placed in a layered fashion between the front glass panel and a sealed back panel with perforations for ports that can be used as injectors, producers, or pressure gauge locations. A sketch of the experimental rig with colored layers is shown in Figure 6.15.

Experimentally, the sand facies have been sieved into groups labeled ESF, C, D, E, F, and G, and there are 3 "fault" regions manufactured in the model. All measurements for permeability and of different types of sand have been performed in several independent sand-pack experiments. However, the in-situ values of sand layer parameters can be quite different due to non-uniform sand distribution and boundary effects, which is why a history-matching workflow is necessary.

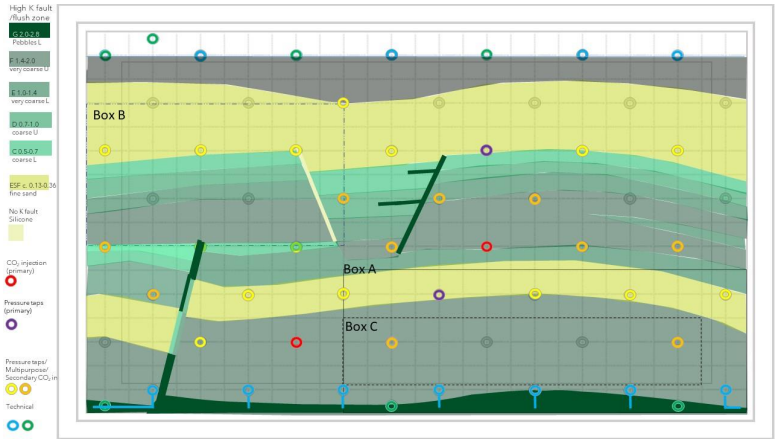


Figure 6.15: The sketch of the experimental rig geometry following FluidFlower benchmark description

6.3.3. HISTORY MATCHING BASED ON ZONATED MODEL AND SIMPLE POISSON-LIKE SOLVER

In this section, we describe baseline history matching performed by the CSIRO team using a simple single-phase flow model. Some results and ideas from this are taken over to the DARTS–based history-matching study of section 6.3.4.

GOVERNING FORWARD MODEL

For the tracer tests, flow is presumed to be single phase, isotropic, the density-independent of tracer, and governed by the averaged thickness (2D model) diffusion equation

$$c_t \phi h \frac{\partial p}{\partial t} - \nabla.(k_w/\mu_w) h(r) \nabla p = Q(r,t) \tag{6.19}$$

where p is the pressure, h the cell thickness, ϕ the porosity, k_w the endpoint permeability of the media for water, μ_w the water viscosity, c_t the total (rock plus water) compressibility, and Q the volume rate of water injection at the ports as a function of time. The porosity ϕ and thickness h are taken as "known" - from sample measurement and thickness measurement bilinear interpolations. The sample measurements for (endpoint) permeability, for the labeled facies, are notated $k_{w,l}$ in the below, for facies l.

The digitized and rasterized model of the facies is shown in Figure 6.16, with labels $1,\ldots 9$ for the facies {1=ESF, 2=C, 3=D, 4=E, 5=F, 6=G, 7=Fault1, 8=Fault2, 9=Fault3}. Fault 2 (facies 8) is designed to be impermeable, and no inferences for its properties are performed in the below.

From the digitized and interpreted images of facies labels $F_i \in \mathscr{L} = \{1,2,..7,9\}$, in voxel i, for facies {ESF, C, D, E, F, G, Fault1, Fault3}, the absolute permeabilities are modeled by $k_{w,i} = u_{F_i} k_{F_i}$ in gridblock i, where the model is $u = \{u_l\}$, $l \in \mathscr{L}$. The supplied "core" sample absolute permeabilities k_l are taken as per Table 6.2.

The model parameters u_l, $l \in \mathscr{L}$ are expected to be $O(1)$ correction factors, and the rescaling is designed to make the inversion as well-scaled as practically reasonable. The

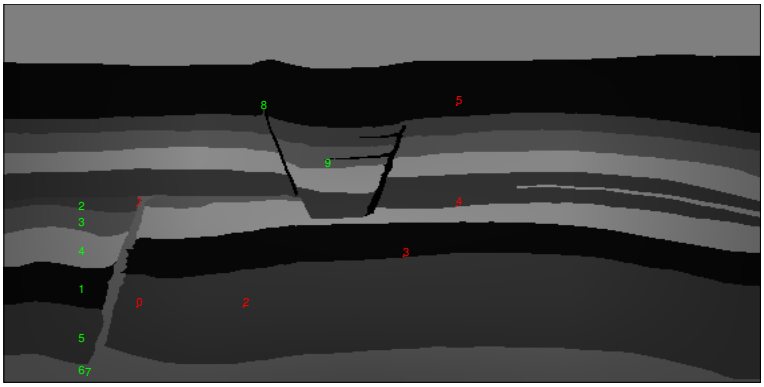


Figure 6.16: Digitised and interpreted facies-label (green codes) image used as the basis for the parameterization. The labeled ports (red) 0,1,...5 are used in the inversion, corresponding to ports {5_3(0),5_7(1),9_3(2),15_5(3),17_7(4),17_11(5)}.

Table 6.2: Endpoint permeabilities and fixed porosities for modeling

Facies	Index l	Permeability[D] (k_l)	Porosity ϕ_l
ESF	1	39	0.44
C	2	293	0.43
D	3	424	0.44
E	4	708	0.45
F	5	258	0.43
G	6	488	0.46
Fault1	7	488	0.46
Fault2	8	0	0
Fault3	9	488	0.46

characteristic time scale for diffusion $t = \frac{\phi\mu c_t L^2}{k_{sc}}$, for the experiment box size L, is very short, typically $t \approx 5$ms using scaling $\phi = 0.45$, $\mu_w = 10^{-3}$ Pa.s, $c_t \approx 0.7310^{-9}$ Pa^{-1}, $L = 1m$, $k_{sc} = 6.9874 \times 10^{-10} m^2$. All pressures are rescaled to dimensionless pressure using a scaling $P_{sc} = Q_{sc}\mu_w/(k_{sc}h_0)$, with additional chosen constants $h_0 = 0.025$m (typical cell width), $Q_{sc} = 6.25 \times 10^{-7}$m^3/s (the water injection rate) yielding scaling pressure $P_{sc} = 35.778$Pa.

For this reason, the forward modeling has approximated the response as steady–state on the time scale of the experiment, with the pressure corresponding to the steady-state solution of Equation (6.19) at each instant of time. During the experiment, the rates Q are sustained as constant over 30-minute interval chunks, so the overwhelming majority of the data are collected with the system equilibrated in terms of pressure. In the experimental data, we see the measured pressures respond virtually instantly to the applied Q when it changes.

The implication of the assumption of constant density is that the fluid configuration remains fixed as long as there is no water injection. The triple-injection tracer experi-

ment was split into 3 successive injections/tranches, each of three 30 min constant-rate injections with rest intervals:

(Port 17_7)	5.19-5.49pm (clear)	6.20-6.50pm(clear)	7.20-7.50pm(clear)	(Day 1)
(Port 9_3)	8.52-9.22pm (clear)	9.52-10.22pm(clear)	10.52-11.22pm(clear)	(Day 1)
(Port 9_3)	2.48-3.18pm (blue)	3.48-4.18pm (blue)	4.48-5.18pm (blue)	(Day 2)

The upshot of this is that for the first 6 hours of the tracer experiment, about 3 hours are injecting, and 3 hours waiting. After Day 1, there is a $\approx$ 16 hour wait till the third tranche of tests on Day 2, during which buoyancy rise in the tracer becomes evident in the comparison of the images at the end of tranche 2 and the start of tranche 3. Since the gravity effects are slow, we expect this effect in the first two tranches to be relatively weak, but the effect is clear in the 3rd tranche images, especially around the 9_3 port injection. The Poisson-like model neglects this buoyancy effect, but for small buoyant drift we expect the residual error at the tracer front to end up symmetrically distributed around the prediction of the constant-density model, so we believe the approximation is reasonable, especially for profiles formed around newly-injected ports. Roughly speaking, we expect the approximation to increase the variance of the predictive model to the leading order, but not the bias.

Thus, on the basis that time-stepping the diffusion equation was unnecessary, we solve the steady state Poisson equation with Q at two different places (ports 9_3, 17_7), corresponding to the experiment performed. The resulting fixed pressure fields and velocities are then steady over the period where the injection rate is held steady, and are sufficient to compute the advection of the tracers over this interval and the observed pressure fields. The model was based on a 5mm grid ($N_x = 568$, $N_y = 300$), with upper boundary condition set as fixed (atmospheric) pressure in the water column, and no–flow Neumann conditions on the left, right, and lower edges of the model. A typical pressure solution for injection at port 17_7 is shown in Figure 6.17. To model the tracer

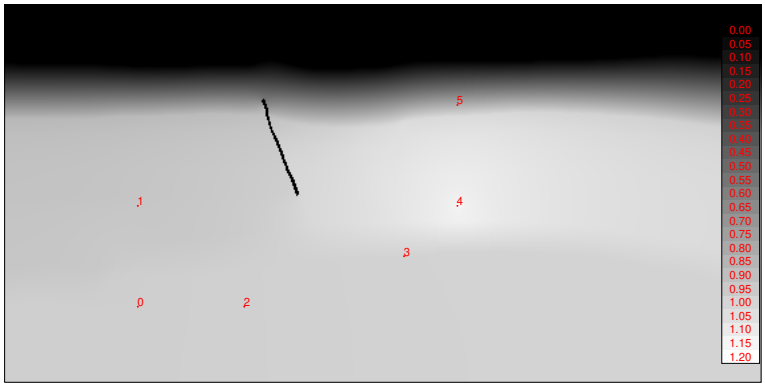


Figure 6.17: Typical (dimensionless) pressure solution of steady-state Poisson problem for injection at 17_7 port (indexed 4).

movement, tracers were advected along streamlines using the fixed velocities computed from the Poisson solves, continuing for the correct time corresponding to each tracer color. The tracer advection was implemented using the upwinding scheme described

in [109]. The time-stepping in the tracer computation was adjusted so the number of time steps divided the total tracer time exactly, which has the merit that any computed quantities from the tracer image are a smooth differentiable function of the parameters in the PDE.

HISTORY MATCHING

For inversion, the data available consisted of the injection port and monitor port pressures during the injections, and color tracer images at the end of the injection periods. The cross–port pressures were conspicuously noisy and clearly close to the noise floor of the instruments. Significant drift was evident in these measurements, and even the stable values showed the curious property of the amplitude not diminishing in a consistent way with the distance from the injection port. This has implications in the inversion if this data is weighted very heavily. The injection pressures were strong signals but measured some way away (20cm or more) from the actual injection face and subject to unknown frictional and other losses in the feed plumbing. This makes them not very useful for inversion. By contrast, the tracer images were very clearly interpretable, rich in spatial content, and not obviously contaminated by an experimental artifact of any significant kind.

The cell is initially filled with a blue tracer. Injections were modeled at $Q = 2250\text{ml/h}$ ($6.25 \times 10^{-7}\text{m}^3/\text{s}$) for 3×30 minutes at port 17_7 with clear tracer, then 3×30 minutes at port 9_3 with more clear tracer, then 3×30 minutes at 9_3 with blue tracer. The forward modeling operation, which implements standard tracer advection under an upwinding scheme, is expressed below as a function $f_t(u)$ which generates concentration profiles, which are very close to unity inside the swept region, and fall rapidly to zero at the tracer front. The tracer advection is stepped forward for precisely the number of time steps needed for the injection, and numerical integrations of the total tracer mass over the modeling grid at the end of the simulation agree very closely with the mass known to be injected from Q in the tracer source. Under the assumptions of the single-phase PDE and the fast equilibration time, the experimental 30 min wait time between injections does not need to be modeled, as nothing happens in the Poisson model if the sources are switched off since the velocities are then instantly zero and no advection occurs. The modeled tracer positions at the end of each 30min injection period are compared to digital image experimental data for inversion.

The inversion was couched as a Bayesian inverse problem with a likelihood $P(\mathrm{d_{obs}}|\mathrm{u})$ formed as a joint probability using pressure and tracer data $\mathrm{d_{obs}} = \{\mathrm{p_{obs}}, \mathrm{c_{tracer}}\}$. The model was taken to be multiplier modifiers of the permeability parameters, per facies, and applied in a "paint-by-numbers" fashion over the labeled facies model. The thickness and porosity data were considered to be sufficiently precise and experimentally stable to be fixed for the purposes of model prediction and inversion. The Bayesian framework was completed with the provision of a weak prior for the modifier parameters, of Gaussian form $P(u_l) \sim N(1, \sigma^2)$ with $\sigma = 5$ for each parameter. The associated prior covariance is $C_p = \mathrm{diag}\{\sigma^2\}$. The model point estimate at the global maximum of the posterior probability is referred to as the MAP (maximum a posteriori) inversion.

The inversion is performed using a Levenberg–Marquardt routine [76, 68], which requires the Jacobian $J = \{\partial f / \partial u\}$ of the forward response with respect to the unknown

model parameters. Since the model dimensionality was very low and the forward model speed very fast (measured in seconds), this was computed using simple forward differences.

The negative log posterior $E(u) \sim -\log(P(\mathrm{d_{obs}}|\mathrm{u})P(\mathrm{u}))$ used in the optimisation step was written as a standard l_2 misfit energy

$$E(u) = E_{\text{pressure}}(u) + E_{\text{tracer}}(u) + E_p(u) \tag{6.20}$$

where the cross–port pressure misfit, accumulated over only stable average measurements at ports p5.3_1 , p5.7_1 , p9.3_1 , p15.5_1 , p17.7_1 , p17.11_1 is

$$E_{\text{pressure}}(u) = \tfrac{1}{2}\lambda_p ||\mathrm{p_{obs}} - f_{\mathrm{p}}(\mathrm{u})||_2^2 \tag{6.21}$$

and the trace image mismatch is written as

$$E_{\text{tracer}}(u) = \tfrac{1}{2}\lambda_t ||\mathrm{c_{tracer}} - f_{\mathrm{t}}(\mathrm{u})||_2^2 \tag{6.22}$$

The weights λ_p, λ_t are adjusted so the tracer data is dominant in the likelihood as this data is much more abundant and artifact-free. The prior Bayesian term amounts to

$$E_p(u) = -\log(P(u)) \sim \tfrac{1}{2}\sum_{l\in\mathscr{L}} (u_l - 1)^2/\sigma_p^2 \tag{6.23}$$

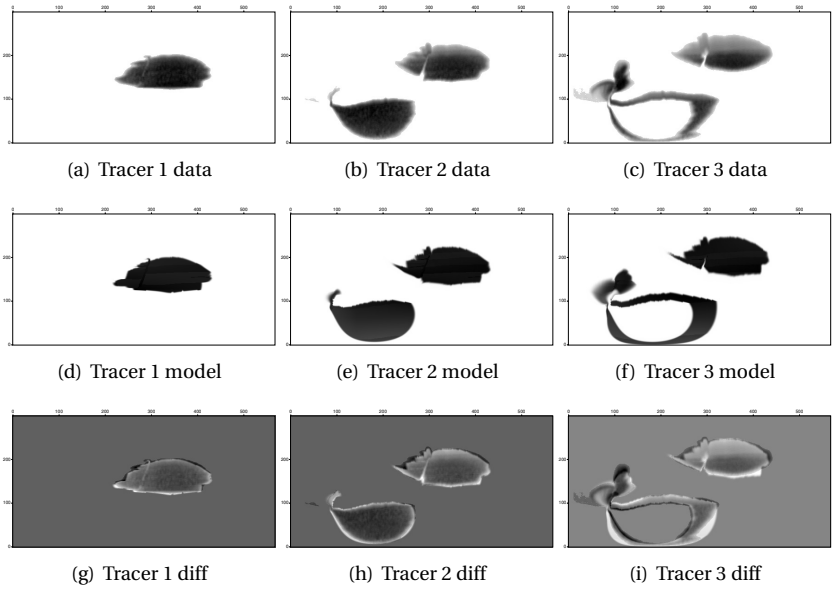
and has a very benign influence on the inversion, except that the likelihood term from the tracer image is expected to be sensitive to permeability ratios only, i.e. has a null space associated with a global scalar multiplier. If no pressure data are used, the weak prior will have the effect of producing a MAP point as the nearest point in the likelihood null space to the prior mean point $u = \mathbf{1}$, i.e. the core measurements. In dimensionless units, the pressure data $\mathrm{p_{obs}}$ are $O(1)$ numbers, but rather noisy, so setting $\lambda_p = 1$ seems appropriate. The tracer data c are processed from the digital images to have concentration values ranging over [0, 1]. Since the associated l_2 norm has a very large number of voxels, λ_t is scaled such that the tracer misfit energy is $E_{\text{tracer}}(u) = 1000/2$ for a model that produces no tracer concentration ($f_t(u) = \mathbf{0}$), i.e. the information content is equivalent to 1000 measurements. In practice since the volume in which experimental and forward–modeled concentrations differ is only a small fraction of the image, the misfit energy from this term ends up being $O(10)$, perhaps equivalent to putting 10-fold the emphasis on the tracer images as the crosswell pressure data.

One possible approach was to assert that the inconsistencies in the crosswell pressure data were too problematic to warrant their inclusion, and that the inversion should be performed on the basis of the 3–injection tracer data alone. It was also considered reasonable to merge the parameters for regions 5 and 6, since region 6 is at the edge of the modeling region and will have a more fragile permeability inference.

The corresponding parameter inferences are as per Table 6.3. The table shows the dimensionless scaled model inference and corresponding actual unitized values. The final column is the dimensionless uncertainty estimate $\sigma_l^2 = H_{ll}^{-1}$ for each parameter formed from the inverse Hessian matrix at the final optimum, where $H = J^T J + C_p^{-1}$. The final inversion forward model images and associated data snapshots are depicted in Figure 6.18.

Table 6.3: Inversion results from 3–tracer inversion, regions 5 and 6 merged to a common parameter.

Facies	Index l	u_l	Permeability[D] (k_l)	σ_l
ESF	1	0.67	26.2	0.235
C	2	0.60	176	0.65
D	3	0.73	309	0.78
E	4	0.96	678	1
F	5	4.18	1080	0.61
G	6	4.18	1080	0.61
Fault1	7	2.74	1340	0.84
Fault2	8	-	0	-
Fault3	9	3.23	1575	0.84



(a) Tracer 1 data (b) Tracer 2 data (c) Tracer 3 data

(d) Tracer 1 model (e) Tracer 2 model (f) Tracer 3 model

(g) Tracer 1 diff (h) Tracer 2 diff (i) Tracer 3 diff

Figure 6.18: Tracer data image and associated MAP forward models. Greyscale is [0,1]=[white,black].

A more optimistic approach was to include the crosswell data, but down-weighted in the sense previously described, and allow independence in zone 5 and 6 scaled permeabilities. The result of this inversion is shown in Figure 6.18. One sees that the parameters differ from the previous inversion, but within the estimated standard deviation associated with the estimated "statistical power" of the data embedded in the likelihood weightings.

Table 6.4: Inversion results from 3–tracer inversion, with crosswell data, and facies 5,6 independent.

Facies	Index l	u_l	Permeability[D] ($u_l k_l$)	σ_l
ESF	1	0.89	34.6	0.24
C	2	1.08	317	0.65
D	3	0.78	332	0.78
E	4	1.2	840	1.0
F	5	5.7	1477	0.61
G	6	2.43	1186	0.84
Fault1	7	4.27	2082	0.84
Fault2	8	-	0	-
Fault3	9	2.41	1176	0.84

6.3.4. HISTORY MATCHING OF FLUIDFLOWER BENCHMARK USING TWO-STAGE APPROACH

In this section, we describe a history-matching approach combining the best practices of FluidFlower benchmark studies performed by CSIRO and TU Delft teams. We will use the numerical framework based on DARTS since it allowed us to include some advanced physics (e.g., the buoyancy of a plume) and numerical capabilities (e.g., adjoint gradients). Several ingredients used in the original approach of the TU Delft team have been enhanced by results and ideas from CSIRO's history-matching study.

INITIAL MODEL TUNING

Based on interpreted high-resolution images of the rig, we built an unstructured mesh grid model as depicted in Figure 6.19. Following the experimental description of uncon-

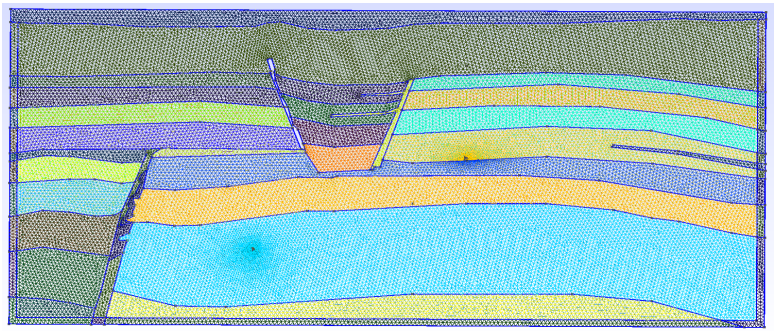


Figure 6.19: The unstructured mesh grid of FluidFlower model

solidated layering in the rig, the mesh grid model is divided into multiple layers filled with nine different types of sand in total. Different layers are assigned to different permeability and porosity values. The anisotropy of the permeability is also modelled by the introduction of a vertical/horizontal anisotropy ratio. This anisotropy corresponds to the distribution of sand grains into laminated layers clearly visible in high-resolution images and mostly related to fine-grain sand. Within the same layer, the same type of

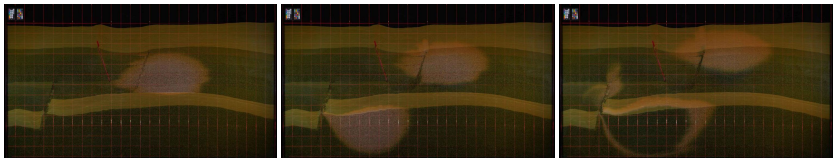
sand is maintained, and the petrophysical values of the grid cells are kept constant.

The petrophysical properties of the layers in Figure 6.15 including permeability are initially estimated from the sand-pack experiments. However, these measurements are different from the in-situ properties of sand layers due to experimental variations in sand deposition/assembly, loading, and boundary effects. For the fitting procedure, we keep the porosity unchanged while the permeability and the anisotropy of the layers are taken as the free parameters. A significant decision was taken to also introduce the density of the clear tracer as an additional parameter. The optimization was again based on gradient methods in low dimensions from finite differences of efficient forward simulations. Specifically, the objective function is defined as:

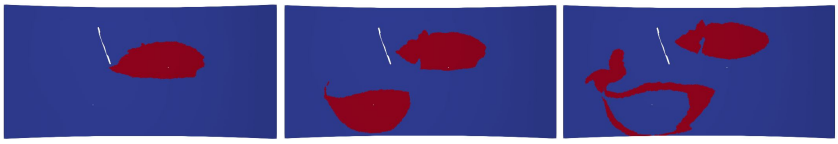
$$E(u) = ||G(u) - \mathrm{d}_{\mathrm{rmobs}}||_2^2 \tag{6.24}$$

where E is the objective function, u are the model variables, G is the model response, and $\mathrm{d}_{\mathrm{obs}}$ is the observation data.

Instead of using the original tracer concentration of the experimental images as the observations, we process the experimental tracer images into several binary images, corresponding to the end of each 3×30 minute tracer injection group; see Figure 6.20. The first row of this figure demonstrates the experimental images of the tracer plumes at the 3rd, 6th, and 9th "macro" time steps. The second row shows the associated digitized binary map of each time step. The value of the red color is set as 1, while the blue color is 0. A threshold of 7×10^{-5} tracer concentration is used to binarise the response $G(\boldsymbol{u})$ to either 0 or 1. Although the objective function defined by the binarised model response and observations is mathematically non-differentiable, numerically it is not an issue to compute the gradient and the optimizer managed to overcome it in this modeling tuning stage. These binary maps delineate the boundaries of the tracer plumes and will be utilized as the observations in the initial stage of model tuning. This simple image processing technique will be later compared with accurate image recognition performed by CSIRO shown in Figure 6.18.



(a) The 3rd time step (b) The 6th time step (c) The 9th time step



(d) The 3rd binary map (e) The 6th binary map (f) The 9th binary map

Figure 6.20: The experimental images of tracer plumes and the associated digitized binary maps.

The initial model parameters are obtained from Table 6.4. They are the permeability

of layers C, D, E, F, Fault1, Fault3, G, and the anisotropy factor of the layer ESF in the z-direction. These eight parameters are updated in the history matching iterations. The optimal values of these eight parameters are shown in Table 6.5. Note that Table 6.5 also includes placeholder information for the water layer W.

Table 6.5: The petrophysical properties of the model layers

Facies	Porosity[-]	Permeability[D]	Anisotropy[x, y, z]
ESF	0.43	34.6	[1, 1, 0.316]
C	0.44	302	[1, 1, 1]
D	0.44	1016	[1, 1, 1]
E	0.45	549	[1, 1, 1]
F	0.45	1976	[1, 1, 1]
G	0.44	1743	[1, 1, 1]
Fault1	0.44	2554	[1, 1, 1]
Fault3	0.44	739	[1, 1, 1]
W	0.44	10000	[1, 1, 1]

Figure 6.21 shows the corresponding forward model of the last time step (i.e. the 9th time step in Figure 6.20) based on the tuned petrophysical properties of Table 6.5. In Figure 6.21, the left figure is the tracer concentration map of the model response. The middle figure is the binary plot based on a given threshold of the tracer concentration (i.e. red color if larger than 7×10^{-5}; otherwise, blue color). The right figure is the experimental result of the tracer test at the last time step. It is evident that the position of the simulated tracer plume (the left and middle figure) is systematically shifted from the experimental results. Assuming the forward physics is adequate, and that the optimization has found a global minimum, this indicates that the spatial model needs to be enriched in order to better match the experimental results. In the next section, we enlarge the model with additional parameters to address the underfitting problem. Spatial variation of the petrophysical properties will also be introduced as an attempt to model the variation caused by the manual sand deposition process in the experimental rig.

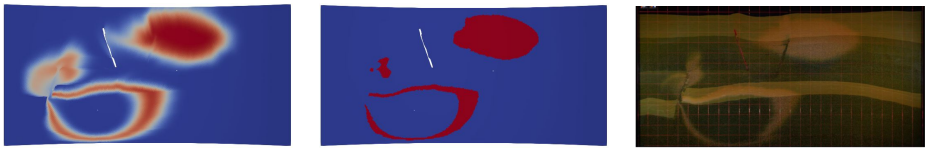


Figure 6.21: The tracer test results at the last time step. The left and middle figures are the results of numerical model response. The right figure is the experimental results.

THE GENERATION OF THE PRIOR ENSEMBLE

In the previous sections, nine types of sand are "painted" into the model to represent the different layers and faults. This forms homogeneous petrophysical properties within a single layer. However, since the sand is manually filled into the rig, it is difficult to maintain homogeneous properties within the same layer. Moreover, compaction is observed

in the course of pre-injection flushing, environmental temperature fluctuations, water injection, etc. It is impossible to have identical compaction everywhere in the layers. Therefore, it is necessary to include spatial uncertainty in the model.

To this end, the high dimensional permeability field $\boldsymbol{K}$ is introduced to the model, and equipped with a prior distribution $\log(\boldsymbol{K}) \sim N(\log(\boldsymbol{K}_t), C_p)$, where $\boldsymbol{K}_t$ is the spatially mapped tuned permeability from Table 6.5, and C_p is a Gaussian variogram based correlation model of sill $\text{diag}\{C_p\} = \sigma^2$, and $\sigma = 0.02$. Samples may be drawn from this model in many standard ways, such as sequential Gaussian simulation, as depicted in Figure 6.22.

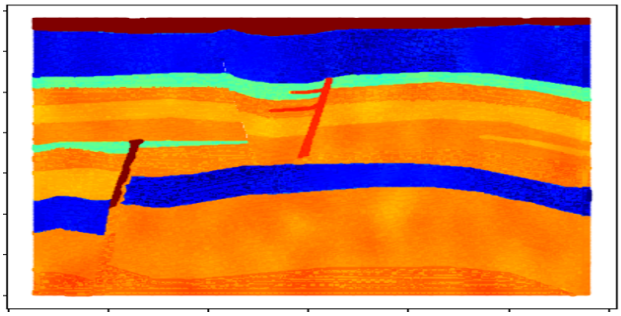


Figure 6.22: A sample from the prior distribution for $\boldsymbol{K}$ with spatial lognormal variation

HISTORY MATCHING FRAMEWORK USING RANDOMIZED MAXIMUM LIKELIHOOD (RML)

The enriched model is now equipped with some regularization apparatus, which behaves in a very similar way to explicit Bayesian prior declarations. Further, the high dimensionality of this model ($n = 72262$ parameters) means that dense Jacobian methods are no longer possible, and efficient gradient methods must be sought. The new objective function reads

$$E(u) = ||G(u) - \mathrm{d}_{\mathrm{obs}}||_2^2 + \mathrm{R(u)}, \tag{6.25}$$

where R is the regularization term. Posterior uncertainties of this model associated with the Gibbs distribution $P(u|d_{\mathrm{obs}}) \sim \exp(-E(u)/2)$ are approximated using the Randomized Maximum Likelihood (RML) approach, which is described in Section 6.1.

6.3.5. HISTORY MATCHING RESULTS

In this section, we describe how the history matching framework was applied for the FluidFlower benchmark and show several results.

HINGE LOSS FUNCTION

As per the approach in Section 6.3.4, instead of using the original tracer plumes of the images as the observations, we digitized the original images into binary maps, e.g. the second row of Figure 6.20. Based on the information of the binary maps, we replaced the misfit term of Equation (6.25) (i.e. the first term at the right-hand side in this equation) with the hinge loss function. The updated objective function reads:

$$E(u) = ||H(u)||_2^2 + R(u), \tag{6.26}$$

where $H(u)$ is the hinge loss function. If the given cell is located at the red region in Figure 6.20, the hinge loss function is defined as:

$$H(u) = \begin{cases} 0 & \text{if G(u) > threshold} \\ G(u) - \mathrm{d_{obs}} & \text{if G(u) } \leq \text{ threshold} \end{cases}. \tag{6.27}$$

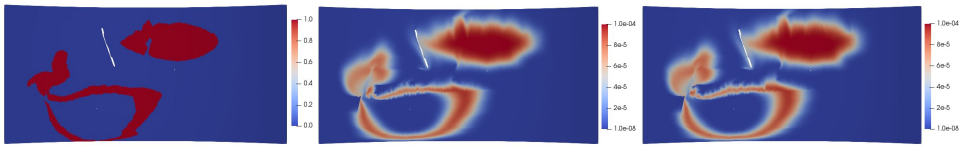
Similarly, when the given cell is located at the blue region in Figure 6.20, the hinge loss function is defined as:

$$H(u) = \begin{cases} 0 & \text{if G(u) < threshold} \\ G(u) - \mathrm{d_{obs}} & \text{if G(u) } \geq \text{ threshold} \end{cases}. \tag{6.28}$$

REGULARIZATION AND RANDOMIZED MAXIMUM LIKELIHOOD

The regularization term (u) in Equation (6.25) is defined as Equation (6.4). Note that we will only keep the elements at the diagonal in the covariance matrix in Equation (6.4), while the off-diagonal elements are all zero. The reason is that the covariances describe the petrophysical correlation between different positions in the reservoir induced by the sedimentary assembly process. However, the sand in this experimental rig is artificially filled, with few of the usual depositional or consolidation characteristics.

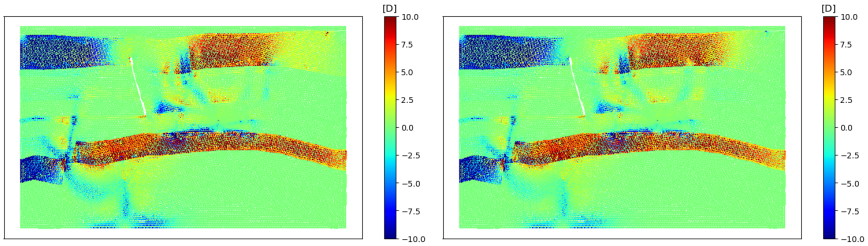
In the Randomized Maximum Likelihood (RML) method, $\boldsymbol{u}_{ref}$ refers to different prior realizations with different spatial lognormal variations. The details of the generation of these priors are described in Section 6.3.4. With the digitized binary observations and the RML method, we implemented multiple history matchings based on different references $\boldsymbol{u}_{ref}$. Figure 6.23 demonstrates two examples of the history-matching results of the tracer concentration at the last time step. Their associated changes of the permeability after history matching are shown in Figure 6.24.



(a) binary observation (b) case 1 (c) case 2

Figure 6.23: The binary observation and two examples of the history matching results based on the binary observation. The figures demonstrate the model response of the inferred model at the last time step

Compared with the tracer concentration result in Figure 6.21, the results in Figure 6.23 match the experimental plume observations more accurately, especially for the top right plume. This history-matching result keeps the original pattern of the tracer plumes obtained from the initial stage of model tuning for eight homogeneous parameters. At the same time, the spatial lognormal variations of permeability are introduced to represent the spatial uncertainties in the model. The total misfit error of the inferred model ensemble is reduced by around 39%. Note that there are already two stages of model history matching and model tuning implemented before this stage of history matching using adjoint methods.



(a) (b)

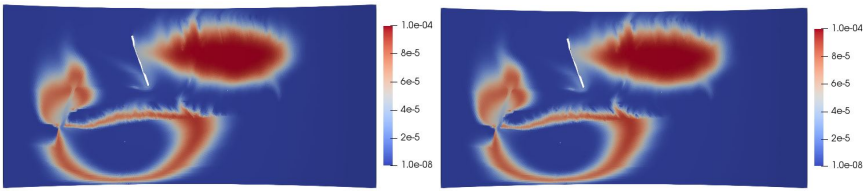
Figure 6.24: Two examples of the changes of the permeability distribution after the history matching based on the binary observations

Figure 6.24 demonstrates two examples of the permeability difference between the model before and after history matching. It is clear from Figure 6.24 that the permeabilities in the tracer plume zone and the high permeable zone have more permeability adjustment compared with the rest of the region. This is because the tracer plume zone and the high permeable zone are the regions of high sensitivity in terms of data misfit.

6

6.3.6. REGULARIZATION AND RML BASED ON THE CONCENTRATION INTERPRETED FROM IMAGES

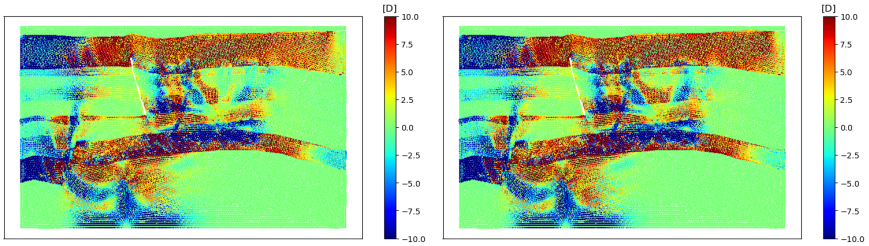
An alternative strategy of history matching was also conducted based on the concentration data interpreted from experimental images, as per the Tracer 1, 2, and 3 data in Figure 6.18. In this case, we can directly use the original objective function Equation (6.25) to infer the model, instead of introducing the Hinge loss function for the treatment of binary observations in Equation (6.26). Based on the same realization ensemble $\boldsymbol{u}_{ref}$ and RML method, the history matching results of the tracer concentration and the changes of permeability are shown in Figure 6.25 and Figure 6.26.



(a) (b)

Figure 6.25: Two examples of the history matching results based on the tracer concentration observations. The figures demonstrate the model response of the inferred model at the last time step

The error of the inferred model ensemble is reduced by around 16%. Again, this error reduction is achieved using the high dimensional spatial model and adjoints. The misfits associated with the parsimonious layer models are not discussed here. Figure 6.25 has more similar plume shapes to the experimental images compared with the results in



(a) (b)

Figure 6.26: Two examples of the changes of the permeability distribution after the history matching based on the tracer concentration observations

Figure 6.23.

6.3.7. CONCLUSION AND DISCUSSION

A comprehensive history matching for FluidFlower experimental results has been described in this study. Several stages of model tuning and inference have been implemented either sequentially or in parallel in order to achieve a good history matching result for the FluidFlower tracer tests. In general, the pressure data were of very limited use in model fitting, being noisy and shifted by unknown frictional losses. However, tracer images were of high quality and rich in information.

The history matching was first conducted based on single-phase physics without buoyancy effects, using a simple Poisson-like solver. A structured grid and digitized concentration data were utilized in this stage to infer the permeability of six facies types and two faults. The numerical framework of this approach is a simple finite difference code with upwind tracer advection, using Marquardt methods for gradient-based optimization.

To further improve the history matching result and account for the spatial uncertainty of the permeability field, two extra layers of model enrichment and history matching were added. For initial forward model refinement, the forward physics was switched to an unstructured grid with a facility for modeling buoyancy. Eight petrophysical parameters (including permeability and anisotropy) were chosen as model variables, and simplified binary tracer images were used. Inversions were performed using Gauss-Newton methods and explicit gradients obtained by finite differences. This low-dimensional inversion was then taken as the reference model of the next stage of history matching using the adjoint-based gradient method as the workhorse in a Randomized Maximum Likelihood (RML) approach, with a richer and high-dimensional spatial model to account for heterogeneity. The coarse-scale morphology of the tracer images was remarkably well predicted by the low dimensional models, but fine detail could only be well reproduced by significant inflation of the spatial model space.

7

RECAPITULATION AND CONCLUSIONS

7.1. ADJOINT METHOD AND THE INVERSE MODELING FEATURE OF DARTS

In this study, the inverse modeling capabilities of DARTS have been developed based on adjoint methods. This framework fulfills the inverse capabilities of DARTS as a powerful and efficient reservoir simulator. The OBL approach was proposed to facilitate the Jacobian assembly in the forward simulation. Furthermore, the OBL approach also provides a convenient way to obtain the associated derivatives used in adjoint assembly during the inverse modeling based on the adjoint method. To cater to the broad spectrum of applications in energy transition projects, we designed and implemented the adjoint method in both the super engine and thermal engine of DARTS. These efforts pave the way for more extensive applications of DARTS in energy transition projects.

In this inverse modeling framework, the transmissibility and well index are taken as the model parameters to be adjusted in the course of history matching. More gradients related to other types of model parameters (e.g. the rock-fluid interaction parameters) can also be combined with the adjoint-based gradient of transmissibility and well index to conduct the history matching. Since the degrees of freedom corresponding to permeability (transmissibility) are usually very high, this makes the gradient evaluation procedure very prohibitive if the numerical method is applied. Therefore, the adjoint method is implemented for transmissibility and well indexes in order to achieve a high efficiency to evaluate their gradients. The validation and the comparison of the adjoint-based gradient with the numerical gradient show that the CPU time of adjoint-based gradient evaluation can be several orders of magnitude faster than the numerical gradient. This efficiency improvement can be even higher when the degrees of freedom of the geological model are larger. Apart from the efficiency improvement, the gradient evaluation based on the adjoint method does not need to introduce an infinitesimal quantity to evaluate the gradient as required in the numerical Finite Difference Method. This expels

a source error in the gradient evaluation caused by the introduction of such an infinitesimal quantity. In other words, the adjoint method provides an analytical solution for the gradient evaluation of the model parameters used in inverse modeling.

To address the issue of the time measurement mismatch between the field observation timestep and the simulation timestep, the Dirac delta measurement function is introduced. Various types of misfits between the observations and the model response such as well phase rate, BHP, well temperature, and time-lapse data, can be easily incorporated into the objective function. However, these misfit terms can differ significantly in magnitude and units. This would cause big problems for the optimizer to search for the optimal solution, as the performance of the optimizer is very sensitive to the magnitude of the model parameters and objective function values. Therefore, a weighting strategy is applied in order to normalize these misfit terms into a reasonable range.

The assembly of the derivatives in the adjoint method is presented in Section 4.2.1. A one-dimensional model example is used to present the matrix structures of the derivatives of the residuals. Corresponding to different types of well treatments in DARTS, the adjoint-based derivatives of well phase rate, BHP, well temperature, and time-lapse data of the reservoir are also demonstrated. With the assembly of those derivatives, the adjoint-based gradient can be solved by using the Equations (3.22) to (3.24). Then, the resulting adjoint-based gradient can then be utilized by the optimizer to efficiently search for the optimum.

Initially, the prototype of the inverse modeling framework based on the adjoint method was developed and validated in MATLAB. Subsequently, the code was fully implemented in C++ and Python. Similar to the forward simulation framework of DARTS, the performance-critical aspects of the adjoint method are implemented in C++ (e.g., the assembly of derivatives and solving the linear equations). For the pre- and post-processing tasks such as the preparation of the objective function, the code is implemented in Python. Corresponding to the "super engine" and "thermal engine", the adjoint method is also implemented in both engines.

7.2. APPLICATIONS OF THE INVERSE CAPABILITIES OF DARTS IN DIWA MODEL

Once the inverse modeling framework is implemented, it is applied to several geo-energy applications. In Chapter 5, the Discrete Well Affinity (DiWA) data-driven proxy model is proposed and validated. The DiWA model is designed to represent the complex reservoir structure and fluid flow using very coarse grids and basic geological information, such as average reservoir thickness and porosity. More geological properties can be added to the model if available. To calibrate the model, the adjoint method is employed to evaluate the gradient used in the training procedure. Since the DiWA model uses very coarse grids, it is necessary to include and calibrate rock-fluid parameters, such as the Brooks-Corey model, during training.

The DiWA model was first validated by using two synthetic fluvial model ensembles, namely FLUMY and MPS. The results demonstrate that the DiWA model achieves good prediction accuracy, despite a significantly reduced model size, with both training and prediction accuracy at satisfactory levels. Moreover, a comparison between the data-

driven proxy methodology and a conventional flow-based upscaling technique was conducted. The data-driven proxy methodology exhibits improved accuracy compared to the conventional approach within both fluvial model ensembles.

Next, a simple 3D proxy model with multiple layers and three flowing phases was tested using observation data generated from a modified SPE 1 model. The aim was to determine whether the proxy model could accurately characterize the geological information of the reservoir when the initial guess is based on reliable geological information. Our results indicate that the proxy model performed well under these conditions. However, it was noted that the model's performance was significantly degraded when the initial guess was far from the true geology. This finding led us to propose an efficient sampling approach that involves training of each statistical member to match the true data.

To evaluate the proposed approach, a high-fidelity Brugge model was utilized to generate true data for model training. To ensure variability in the data, uniformly distributed random perturbations were added to BHP control during the data generation process. Subsequently, the DiWA model was trained using the generated true data. The proxy model employed in this study only incorporated basic geological information, such as the contour and average thickness of the original reservoir model. However, the framework can be extended to include additional information and regularization techniques to enhance the training process.

7.3. INVERSE MODELING FOR ENERGY TRANSITION PROJECTS

In addition to its applications in data-driven DiWA proxy modeling, this framework is also capable of solving inverse modeling problems in the field of energy transitions, including geothermal energy and CO_2 storage projects. To accommodate a broader range of observations in these applications, the framework was modified to include various types of observation data, such as well temperature, time-lapse temperature distribution, and time-lapse component concentration. The resulting structure is highly flexible and can be adapted to suit the specific requirements of each individual project.

In the geothermal energy project, we successfully completed the history matching of the Egg geothermal model, which comprises 71160 model parameters, within 10 hours using Intel XEON E5-6248R processors. To represent the complexity of the model while preserving the uncertainty of the model ensemble, we employed Principal Component Analysis (PCA) to transform the model from its original space to a reduced-dimension ξ space and vice versa. This procedure involves the transformation of the model parameters, gradients, and regularization terms.

In this geothermal project, electromagnetic data measurements were incorporated as an additional type of observation data. When combined with conventional data such as well rates and BHP, the trained model was able to accurately predict the temperature front. It is worth noting that not all types of observations need to be included in the history-matching framework. Our framework is designed to be flexible, allowing for the addition or removal of various types of observation data in the objective function based on their availability and relevance to the specific project requirements.

The comparison of the history matching based on MLE, RML, and constrained to ensemble prior mean, are presented in this study to illustrate the performance of the

proposed framework. The results show that the adjoint gradient is able to capture the characteristics of the "true" permeability field and re-construct the temperature map that is observed from the EM data. This phenomenon is particularly evident in the MLE-based history-matching approach. In contrast, the history matching based on RML and constraint to ensemble prior mean considered both the likelihood maximization and prior knowledge information, resulting in a model response that is relatively dissimilar to the observation while preserving the prior permeability information.

The CO_2 storage project, specifically the FluidFlower experiment, also utilized the proposed framework with an extension to include time-lapse concentration measurements of the injected tracer. A comprehensive history-matching process was conducted to achieve a good match with the experimental results. This involved multiple stages of model tuning and inference, implemented either sequentially or in parallel, to achieve the desired outcome. The initial stage of history matching in the FluidFlower experiment involved single-phase physics without considering buoyancy effects. A Poisson-like solver was employed using a structured grid and digitized concentration data to infer the permeability of six facies types and two faults. The numerical framework utilized for this approach was a simple finite difference code with upwind tracer advection, with Marquardt methods used for gradient-based optimization.

To further improve the history matching result and account for the spatial uncertainty of the permeability field, two extra layers of model enrichment and history matching were added. For initial forward model refinement, the forward physics was switched to an unstructured grid with a facility for modeling buoyancy. Eight petrophysical parameters (including permeability and anisotropy) were chosen as model variables and simplified binary tracer images were used. Inversions were performed using Gauss-Newton methods and explicit gradients obtained by finite differences. In the next stage, this low-dimensional inversion was then taken as the reference model of the next stage of history matching using the adjoint-based gradient method as the workhorse in a Randomized Maximum Likelihood (RML) approach, with a richer and high-dimensional spatial model to account for heterogeneity. The coarse-scale morphology of the tracer images was remarkably well predicted by the low dimensional models, but fine detail could only be well reproduced by significant inflation of the spatial model space.

7.4. FUTURE PERSPECTIVES

The adjoint method is a powerful tool for gradient evaluation in optimization problems. It provides an analytical solution for the gradient evaluation by combining the objective function with the constraints such as the residual form of governing equation. Instead of computing the partial derivatives one by one in the numerical method, the adjoint method simultaneously solves the partial derivatives with respect to the model parameters. However, the adjoint method requires access to the simulator source code and extensive code implementation efforts, which may not be feasible for black box software like commercial simulators. Furthermore, the implementation of the adjoint method may require an exclusive implementation for specific model parameters to assemble the associated derivatives. Additionally, for highly nonlinear and implicitly included model parameters, implementing the adjoint method can be challenging. Therefore, careful consideration of model parameters is crucial in selecting those that significantly affect

geo-energy problems or dominate fluid mass and energy.

In the inverse modeling feature of the DARTS platform, we have implemented the adjoint method for the transmissibility and well index as the primary model parameters. However, as we continue to explore different fluid flow effects and complex physics, additional model parameters can be incorporated using the adjoint method in the future. Currently, the inverse modeling feature of DARTS is focused on hydro-thermal-chemical problems. However, implementing the adjoint method on geomechanical problems would be an interesting and challenging area of research. Induced seismicity, in particular, has garnered increased attention in the society. Therefore, it is crucial to explore the potential of the adjoint method for reducing the risks of induced seismicity.

The adjoint method may also be designed and implemented for production optimization problems. In this case, the code for solving the Lagrangian multiplier (see Equation (3.22) and Equation (3.23)) keeps unchanged, while only the derivatives with respect to the new model parameters (see Equation (3.24)) need to be adjusted. Furthermore, the iterative linear solver can be re-designed for the adjoint method in the future. So far, the conventional preconditioner is not applicable to the backward linear equation system in the adjoint method.

REFERENCES

[1] Delft High Performance Computing Centre (DHPC). *DelftBlue Supercomputer (Phase 1)*. https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase1. 2022.

[2] A. Albertoni and L.W. Lake. "Inferring interwell connectivity only from well-rate fluctuations in waterfloods". In: *SPE Reservoir Evaluation and Engineering* 6.1 (2003), pp. 6–15.

[3] R.P. Batycky, M.J. Blunt, and M.R. Thiele. "A 3D Field-Scale Streamline-Based Reservoir Simulator". In: *SPE Reservoir Engineering* 12.04 (Nov. 1997), pp. 246–254. DOI: 10.2118/36726-pa.

[4] R.P. Batycky, M.J. Blunt, and M.R. Thiele. "A 3D Field-Scale Streamline-Based Reservoir Simulator". In: *SPE Reservoir Engineering (Society of Petroleum Engineers)* 12.4 (1997), pp. 246–253.

[5] M. Borregales et al. "Data-Driven Models Based on Flow Diagnostics". In: *ECMOR XVII*. European Association of Geoscientists & Engineers, 2020. DOI: 10.3997/2214-4609.202035122.

[6] RH Brooks and AT Corey. "Hydraulic properties of porous media, hydrology papers, no. 3". In: *Colorado State University, Collins, Colo* (1964).

[7] D.R. Brouwer and J.D. Jansen. "Dynamic Optimization of Waterflooding With Smart Wells Using Optimal Control Theory". In: *SPE Journal* 9.04 (Dec. 2004), pp. 391–402. DOI: 10.2118/78278-pa.

[8] Vladislav Bukshtynov et al. "Comprehensive framework for gradient-based optimization in closed-loop reservoir management". In: *Computational Geosciences* 19.4 (May 2015), pp. 877–897. DOI: 10.1007/s10596-015-9496-5.

[9] Barbara Cantucci et al. "Geochemical modeling of CO2 storage in deep reservoirs: The Weyburn Project (Canada) case study". In: *Chemical Geology* 265.1-2 (July 2009), pp. 181–197. DOI: 10.1016/j.chemgeo.2008.12.029.

[10] M.A. Cardoso, L.J. Durlofsky, and P. Sarma. "Development and application of reduced-order modeling procedures for subsurface flow simulation". In: *International Journal for Numerical Methods in Engineering* 77.9 (2009), pp. 1322–1350.

[11] W.H. Chen et al. "A New Algorithm for Automatic History Matching". In: *Society of Petroleum Engineers Journal* 14.06 (Dec. 1974), pp. 593–608. DOI: 10.2118/4545-pa.

[12] Yuguang Chen and Louis J Durlofsky. "Adaptive local–global upscaling for general flow scenarios in heterogeneous formations". In: *Transport in Porous Media* 62.2 (2006), pp. 157–185.

[13] Laura Chiaramonte et al. “Fracture characterization and fluid flow simulation with geomechanical constraints for a CO2–EOR and sequestration project Teapot Dome Oil Field, Wyoming, USA”. In: *Energy Procedia* 4 (2011), pp. 3973–3980. DOI: 10.1016/j.egypro.2011.02.337.

[14] *CMG (Computer Modeling Group)*. https://www.cmgl.ca/. Accessed on March 15, 2023. 2023.

[15] K.H. Coats et al. “Three-dimensional simulation of steamflooding”. In: *Soc Pet Eng AIME J* 14.6 (1974), pp. 573–592.

[16] Akhil Datta-Gupta and Akhil Datta-Gupta. *Streamline Simulation: Theory and Practice*. Texas: Society of Petroleum Engineers, 2007.

[17] *Delft Advanced Research Terra Simulator (DARTS)*. https://darts.citg.tudelft.nl/. Accessed on March 15, 2023. 2023.

[18] Marinus E. Donselaar, Remco M. Groenenberg, and Douglas T. Gilding. “Reservoir Geology and Geothermal Potential of the Delft Sandstone Member in the West Netherlands Basin”. In: *Proceedings World Geothermal Congress*. 2015.

[19] *DuMux*. https://dumux.org/. Accessed on March 15, 2023. 2023.

[20] L.J Durlofsky. “Upscaling and gridding of fine scale geological models for flow simulation”. In: *8th International Forum on Reservoir Simulation Iles Borromees, Stresa, Italy*. Vol. 2024. 2005, pp. 1–59. DOI: 10.1016/j.advwatres.2016.07.019.

[21] R. Eberhart and J. Kennedy. “A new optimizer using particle swarm theory”. In: *MHS’95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*. IEEE, 1995. DOI: 10.1109/mhs.1995.494215.

[22] *ECLIPSE Simulators*. https://www.software.slb.com/products/eclipse. Accessed on March 15, 2023. 2023.

[23] Zohreh Fathi and Fred W. Ramirez. “Optimal Injection Policies for Enhanced Oil Recovery: Part 2-Surfactant Flooding”. In: *Society of Petroleum Engineers Journal* 24.03 (June 1984), pp. 333–341. DOI: 10.2118/12814-pa.

[24] Anne-Kari Furre et al. “20 Years of Monitoring CO2-injection at Sleipner”. In: *Energy Procedia* 114 (July 2017), pp. 3916–3926. DOI: 10.1016/j.egypro.2017.03.1523.

[25] S. Garcia, S. Kaminska, and M. Mercedes Maroto-Valer. “Underground carbon dioxide storage in saline formations”. In: *Proceedings of the Institution of Civil Engineers - Waste and Resource Management* 163.2 (May 2010), pp. 77–88. DOI: 10.1680/warm.2010.163.2.77.

[26] G.R. Gavalas, P.C. Shah, and J.H. Seinfeld. “Reservoir History Matching by Bayesian Estimation”. In: *Society of Petroleum Engineers Journal* 16.06 (Dec. 1976), pp. 337–350. DOI: 10.2118/5740-pa.

[27] *GEOSX*. https://www.geosx.org/. Accessed on March 15, 2023. 2023.

[28] C. Geuzaine and J.-F. Remacle. "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities". In: *International Journal for Numerical Methods in Engineering* 79.11 (2009), pp. 1309–1331.

[29] Benjamin Grappe et al. "Dynamic Modelling of Meandering Fluvial Systems at the Reservoir Scale, FLUMY Software". In: *Second Conference on Forward Modelling of Sedimentary Systems*. European Association of Geoscientists & Engineers. 2016, cp–483.

[30] Zhenyu Guo and Albert C. Reynolds. "INSIM-FT in three-dimensions with gravity". In: *Journal of Computational Physics* 380 (Mar. 2019), pp. 143–169. DOI: 10.1016/j.jcp.2018.12.016.

[31] Hadi Hajibeygi et al. "Iterative multiscale finite-volume method". In: *Journal of Computational Physics* 227.19 (Oct. 2008), pp. 8604–8621. DOI: 10.1016/j.jcp.2008.06.013.

[32] Choongyong Han et al. "Adaptation of the CPR Preconditioner for Efficient Solution of the Adjoint Equation". In: *SPE Journal* 18.02 (Jan. 2013), pp. 207–213. DOI: 10.2118/141300-pa.

[33] *Here's why geothermal systems could be better for storing renewable energy than batteries*. https://www.weforum.org/agenda/2022/11/geothermal-renewable-energy-storage/. Accessed on March 19, 2023. 2022.

[34] Rafael Holanda et al. "A State-of-the-Art Literature Review on Capacitance Resistance Models for Reservoir Characterization and Performance Forecasting". In: *Energies* 11.12 (Dec. 2018), p. 3368. DOI: 10.3390/en11123368.

[35] Lars Holden and Bjørn Fredrik Nielsen. "Global upscaling of permeability in heterogeneous reservoirs; the output least squares (ols) method". In: *Transport in Porous Media* 40.2 (2000), pp. 115–143.

[36] John H. Holland. "Genetic Algorithms and Adaptation". In: *Adaptive Control of Ill-Defined Systems*. Springer US, 1984, pp. 317–333. DOI: 10.1007/978-1-4684-8941-5_21.

[37] S. de Hoop et al. "Quantification of coarsening effect on response uncertainty in reservoir simulation". In: *16th European Conference on the Mathematics of Oil Recovery, ECMOR 2018*. 2018.

[38] Thomas Y. Hou and Xiao-Hui Wu. "A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media". In: *Journal of Computational Physics* 134.1 (June 1997), pp. 169–189. DOI: 10.1006/jcph.1997.5682.

[39] Kurt Zenz House et al. "Permanent carbon dioxide storage in deep-sea sediments". In: *Proceedings of the National Academy of Sciences* 103.33 (Aug. 2006), pp. 12291–12295. DOI: 10.1073/pnas.0605318103.

[40] IEA. *Carbon capture, utilisation and storage*. IEA, 2020. URL: https://www.iea.org/fuels-and-technologies/carbon-capture-utilisation-and-storage.

[41] IEA. *CCUS in the transition to net-zero emissions*. IEA, 2020. URL: https://www.iea.org/reports/ccus-in-clean-energy-transitions/ccus-in-the-transition-to-net-zero-emissions.

[42] IEA. *Energy Technology Perspectives, Special Report on Carbon Capture Utilisation and Storage: CCUS in clean energy transitions*. IEA, 2020. URL: https://www.iea.org/reports/ccus-in-clean-energy-transitions.

[43] IPCC. *IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press, 2022.

[44] F.E. Jansen and M.G. Kelkar. "Non-stationary estimation of reservoir properties using production data". In: *Proceedings - SPE Annual Technical Conference and Exhibition*. Vol. Omega. Pt 2. 1997, pp. 131–138.

[45] J. D. Jansen et al. "The egg model - a geological ensemble for reservoir simulation". In: *Geoscience Data Journal* 1.2 (Nov. 2014), pp. 192–195. DOI: 10.1002/gdj3.21.

[46] P. Jenny, S.H. Lee, and H.A. Tchelepi. "Multi-scale finite-volume method for elliptic problems in subsurface flow simulation". In: *Journal of Computational Physics* 187.1 (2003), pp. 47–67.

[47] K. Kala and D. Voskov. "Element balance formulation in reactive compositional flow and transport with parameterization technique". In: *Computational Geosciences* 24.2 (2020), pp. 609–624. DOI: 10.1007/s10596-019-9828-y.

[48] M. Karimi-Fard, L.J. Durlofsky, and K. Aziz. "An Efficient Discrete-Fracture Model Applicable for General-Purpose Reservoir Simulators". In: *SPE Journal* 9.02 (June 2004), pp. 227–236. DOI: 10.2118/88812-pa.

[49] M. Karimi-Fard, L.J. Durlofsky, and K. Aziz. "An efficient discrete-fracture model applicable for general-purpose reservoir simulators". In: *SPE Journal* 9.2 (2004), pp. 227–236.

[50] Eirik Keilegavlen et al. "PorePy: an open-source software for simulation of multiphysics processes in fractured porous media". In: *Computational Geosciences* 25.1 (Oct. 2020), pp. 243–265. DOI: 10.1007/s10596-020-10002-5.

[51] M. Khait. "Delft Advanced Research Terra Simulator: General Purpose Reservoir Simulator with Operator-Based Linearization". PhD thesis. 2019. DOI: 10.4233/UUID:5F0F9B80-A7D6-488D-9BD2-D68B9D7B4B87.

[52] M. Khait, D. Voskov, and R. Zaydullin. "High performance framework for modelling of complex subsurface flow and transport applications". In: *ECMOR 2020 - 17th European Conference on the Mathematics of Oil Recovery*. 2020. DOI: 10.3997/2214-4609.202035188.

[53] Mark Khait and Denis Voskov. "Adaptive Parameterization for Solving of Thermal/Compositional Nonlinear Flow and Transport With Buoyancy". In: *SPE Journal* 23.02 (Jan. 2018), pp. 522–534. DOI: 10.2118/182685-pa.

[54] Mark Khait and Denis Voskov. "Integrated Framework for Modelling of Thermal-Compositional Multiphase Flow in Porous Media". In: *Day 1 Wed, April 10, 2019*. SPE, Mar. 2019. DOI: 10.2118/193932-ms.

[55] Mark Khait and Denis Voskov. "Operator-based linearization for efficient modeling of geothermal processes". In: *Geothermics* 74 (July 2018), pp. 7–18. DOI: 10.1016/j.geothermics.2018.01.012.

[56] A. Kiærr et al. "Evaluation of A Data-Driven Flow Network Model (FlowNet) for Reservoir Prediction and Optimization". In: *ECMOR XVII*. European Association of Geoscientists & Engineers, 2020. DOI: 10.3997/2214-4609.202035099.

[57] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by Simulated Annealing". In: *Science* 220.4598 (May 1983), pp. 671–680. DOI: 10.1126/science.220.4598.671.

[58] Ronald W Klusman. "Evaluation of leakage potential from a carbon dioxide EOR/sequestration project". In: *Energy Conversion and Management* 44.12 (July 2003), pp. 1921–1940. DOI: 10.1016/s0196-8904(02)00226-1.

[59] P. Lerlertpakdee, B. Jafarpour, and E. Gildin. "Efficient production optimization with flow-network models". In: *SPE Journal* 19.6 (2014), pp. 1083–1095.

[60] Dennis Y.C. Leung, Giorgio Caramanna, and M. Mercedes Maroto-Valer. "An overview of current status of carbon dioxide capture and storage technologies". In: *Renewable and Sustainable Energy Reviews* 39 (Nov. 2014), pp. 426–443. DOI: 10.1016/j.rser.2014.07.093.

[61] Knut-Andreas Lie. *An Introduction to Reservoir Simulation Using MATLAB/GNU Octave*. Cambridge University Press, July 2019. DOI: 10.1017/9781108591416.

[62] Kok-Thye Lim. "A new approach for residual and Jacobian arrays construction in reservoir simulators". In: *SPE Computer Applications* 7.04 (1995), pp. 93–96.

[63] Qi Liu and M. Mercedes Maroto-Valer. "Parameters affecting mineral trapping of CO2 sequestration in brines". In: *Greenhouse Gases: Science and Technology* 1.3 (Aug. 2011), pp. 211–222. DOI: 10.1002/ghg.29.

[64] X. Lyu, M. Khait, and D. Voskov. "Operator-Based Linearization Approach for Modelling of Multiphase Flow with Buoyancy and Capillarity". In: *SPE Journal* (2021). DOI: 10.2118/205378-PA.

[65] X. Lyu, D. Voskov, and William R Rossen. "Numerical Investigations of Foam-assisted CO_2 Storage in Saline Aquifers". In: *International journal of greenhouse gas control* (2021).

[66] Xiaocong Lyu, Mark Khait, and Denis Voskov. "Operator-Based Linearization Approach for Modeling of Multiphase Flow with Buoyancy and Capillarity". In: *SPE Journal* 26.04 (Apr. 2021), pp. 1858–1875. DOI: 10.2118/205378-pa.

[67] Xiang Ma and Nicholas Zabaras. "Kernel principal component analysis for stochastic input model generation". In: *Journal of Computational Physics* 230.19 (Aug. 2011), pp. 7311–7331. DOI: 10.1016/j.jcp.2011.05.037.

[68] Kaj Madsen, Hans Bruun Nielsen, and Ole Tingleff. “Methods for Non-Linear Least Squares Problems (2nd ed.)” In: *Informatics and Mathematical Modelling*. Technical University of Denmark. 2004.

[69] Jürg M. Matter and Peter B. Kelemen. “Permanent storage of carbon dioxide in geological reservoirs by mineral carbonation”. In: *Nature Geoscience* 2.12 (Nov. 2009), pp. 837–841. DOI: 10.1038/ngeo683.

[70] L.G.H. Van der Meer. “The K12-B CO 2 injection project in the Netherlands”. In: *Geological Storage of Carbon Dioxide (CO2)*. Elsevier, 2013, 301–332e. DOI: 10.1533/9780857097279.3.301.

[71] Greg J. Mehos and W.Fred Ramirez. “Use of optimal control theory to optimize carbon dioxide miscible-flooding enhanced oil recovery”. In: *Journal of Petroleum Science and Engineering* 2.4 (July 1989), pp. 247–260. DOI: 10.1016/0920-4105(89)90002-8.

[72] S. et al. Mohaghegh. “Artificial intelligence and data mining: enabling technology for smart fields”. In: *The Way Ahead* 5.03 (2009), pp. 14–19.

[73] R.J.D. Moraes et al. “Multiscale gradient computation for flow in heterogeneous porous media”. In: *Journal of Computational Physics* 336 (2017), pp. 644–663. DOI: 10.1016/j.jcp.2017.02.024.

[74] Rafael J. de Moraes, Hadi Hajibeygi, and Jan Dirk Jansen. “A multiscale method for data assimilation”. In: *Computational Geosciences* 24.2 (June 2019), pp. 425–442. DOI: 10.1007/s10596-019-09839-2.

[75] Shlomo P. Neuman et al. “Recent Advances in Statistical and Scaling Analysis of Earth and Environmental Variables”. In: *Advances in Hydrogeology*. Springer New York, 2013, pp. 1–25. DOI: 10.1007/978-1-4614-6479-2_1.

[76] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer New York, 2006. DOI: 10.1007/978-0-387-40065-5.

[77] Jan M. Nordbotten et al. “Final Benchmark Description: FluidFlower International Benchmark Study”. In: (2022). DOI: 10.5281/ZENODO.6807102.

[78] Aziz S. Odeh. “Comparison of Solutions to a Three-Dimensional Black-Oil Reservoir Simulation Problem (includes associated paper 9741)”. In: *Journal of Petroleum Technology* 33.01 (Jan. 1981), pp. 13–25. DOI: 10.2118/9723-pa.

[79] Dean S. Oliver. “Multiple Realizations of the Permeability Field From Well Test Data”. In: *SPE Journal* 1.02 (June 1996), pp. 145–154. DOI: 10.2118/27970-pa.

[80] Dean S. Oliver, Albert C. Reynolds, and Ning Liu. *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press, 2008. DOI: 10.1017/cbo9780511535642.

[81] *Open-DARTS*. https://gitlab.com/open-darts. Accessed on March 19, 2023. 2023.

[82] D. W. Peaceman. “Interpretation of well-block pressures in numerical reservoir simulation with nonsquare grid blocks and anisotropic permeability”. In: *Society of Petroleum Engineers journal* 23.3 (1983), pp. 531–543.

[83] D. Perkins. “Reservoir Simulation for Play-based Development of Low Enthalpy Geothermal Resources.” MA thesis. 2019. URL: https://repository.tudelft.nl.

[84] Lies Peters et al. “Results of the Brugge Benchmark Study for Flooding Optimization and History Matching”. In: *SPE Reservoir Evaluation & Engineering* 13.03 (June 2010), pp. 391–405. DOI: 10.2118/119094-pa.

[85] W. Fred Ramirez, Zohreh Fathi, and Jean Luc Cagnol. “Optimal Injection Policies for Enhanced Oil Recovery: Part 1 Theory and Computational Strategies”. In: *Society of Petroleum Engineers Journal* 24.03 (June 1984), pp. 328–332. DOI: 10.2118/11285-pa.

[86] Guotong Ren et al. “Implementation of Physics-Based Data-Driven Models With a Commercial Simulator”. In: *SPE Reservoir Simulation Conference*. Society of Petroleum Engineers, 2019. DOI: 10.2118/193855-ms.

[87] Wolfram Rühaak et al. “Upscaling thermal conductivities of sedimentary formations for geothermal exploration”. In: *Geothermics* 58 (Nov. 2015), pp. 49–61. DOI: 10.1016/j.geothermics.2015.08.004.

[88] Youcef Saad. “A Flexible Inner-Outer Preconditioned GMRES Algorithm”. In: *SIAM Journal on Scientific Computing* 14.2 (Mar. 1993), pp. 461–469. DOI: 10.1137/0914028.

[89] Yousef Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Jan. 2003. DOI: 10.1137/1.9780898718003.

[90] Xavier Sanchez-Vila, Alberto Guadagnini, and Jesus Carrera. “Representative hydraulic conductivities in saturated groundwater flow”. In: *Reviews of Geophysics* 44.3 (2006). DOI: 10.1029/2005rg000169.

[91] P. Sarma, K. Aziz, and L.J. Durlofsky. “Implementation of Adjoint Solution for Optimal Control of Smart Wells”. In: *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2005. DOI: 10.2118/92864-ms.

[92] Pallav Sarma, Louis J. Durlofsky, and Khalid Aziz. “Kernel Principal Component Analysis for Efficient, Differentiable Parameterization of Multipoint Geostatistics”. In: *Mathematical Geosciences* 40.1 (Dec. 2007), pp. 3–32. DOI: 10.1007/s11004-007-9131-7.

[93] Pallav Sarma et al. “Efficient real-time reservoir management using adjoint-based optimal control and model updating”. In: *Computational Geosciences* 10.1 (Mar. 2006), pp. 3–36. DOI: 10.1007/s10596-005-9009-z.

[94] H.T. Schaef, B.P. McGrail, and A.T. Owen. “Carbonate mineralization of volcanic province basalts”. In: *International Journal of Greenhouse Gas Control* 4.2 (Mar. 2010), pp. 249–261. DOI: 10.1016/j.ijggc.2009.10.009.

[95] S. Shetty, D. Voskov, and D. F. Bruhn. “Numerical Strategy for Uncertainty Quantification in Low Enthalpy Geothermal Projects”. In: *Workshop on Geothermal Reservoir Engineering*. 2018. URL: https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2018/Shetty.pdf.

[96] H.L. Stone. “Probability Model for Estimating Three-Phase Relative Permeability”. In: *Journal of Petroleum Technology* 22.02 (Feb. 1970), pp. 214–218. DOI: 10.2118/2116-pa.

[97] Andreas S. Stordal and Geir Nævdal. “A modified randomized maximum likelihood for improved Bayesian history matching”. In: *Computational Geosciences* 22.1 (June 2017), pp. 29–41. DOI: 10.1007/s10596-017-9664-x.

[98] Sebastien Strebelle and Marjorie Levy. “Using multiple-point statistics to build geologically realistic reservoir models: the MPS/FDM workflow”. In: *Geological Society, London, Special Publications* 309.1 (2008), pp. 67–74.

[99] *The Oil and Gas Industry in Energy Transitions*. https://www.iea.org/reports/the-oil-and-gas-industry-in-energy-transitions. Accessed on March 19, 2023. 2020.

[100] X. Tian et al. “Discrete Well Affinity (DiWA) Data-Driven Proxy Model for Production Forecast”. In: *SPE Journal* (May 2021), pp. 1–17. DOI: 10.2118/205489-pa.

[101] Xiaoming Tian, Oleg Volkov, and Denis Voskov. “An advanced inverse modeling framework for efficient, flexible, and adjoint-based history matching of geothermal energy”. In: *Geothermics* (2023).

[102] Xiaoming Tian and Denis Voskov. “Efficient application of stochastic Discrete Well Affinity (DiWA) proxy model with adjoint gradients for production forecast”. In: *Journal of Petroleum Science and Engineering* 210 (Mar. 2022), p. 109911. DOI: 10.1016/j.petrol.2021.109911.

[103] Xiaoming Tian and Denis Voskov. “Efficient Inverse Modeling Framework for Energy Transition Applications Using Operator-Based Linearization and Adjoint Gradients”. In: *SPE Reservoir Simulation Conference, 2023*. Society of Petroleum Engineers, 2023. DOI: 10.2118/212169-ms.

[104] Xiaoming Tian et al. “The study of the history matching for FluidFlower experimental results”. In: *Transport in Porous Media* (2023).

[105] Vincent Vandeweijer et al. “Monitoring the CO2 injection site: K12-B”. In: *Energy Procedia* 4 (2011), pp. 5471–5478. DOI: 10.1016/j.egypro.2011.02.532.

[106] Hai X. Vo and Louis J. Durlofsky. “A New Differentiable Parameterization Based on Principal Component Analysis for the Low-Dimensional Representation of Complex Geological Models”. In: *Mathematical Geosciences* 46.7 (July 2014), pp. 775–813. DOI: 10.1007/s11004-014-9541-2.

[107] O. Volkov and D.V. Voskov. “Effect of time stepping strategy on adjoint-based production optimization”. In: *Computational Geosciences* 20.3 (2016), pp. 707–722. DOI: 10.1007/s10596-015-9528-1.

[108] Denis V. Voskov. “Operator-based linearization approach for modeling of multiphase multi-component flow in porous media”. In: *Journal of Computational Physics* 337 (May 2017), pp. 275–288. DOI: 10.1016/j.jcp.2017.02.041.

[109] C.B. Vreugdenhil and B. Koren. “A robust upwind discretization method for advection, diffusion and source terms”. English. In: *Numerical Methods for Advection-Diffusion Problems*. Notes on Numerical Fluid Mechanics. Germany: Vieweg, 1993, pp. 117–138.

[110] J.R. Wallis. “Incomplete Gaussian Elimination as a Preconditioning for Generalized Conjugate Gradient Acceleration”. In: *All Days*. SPE, Nov. 1983. DOI: 10.2118/12265-ms.

[111] J.R. Wallis, R.P. Kendall, and T.E. Little. “Constrained Residual Acceleration of Conjugate Residual Methods”. In: *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 1985. DOI: 10.2118/13536-ms.

[112] Y. Wang et al. “An efficient numerical simulator for geothermal simulation: A benchmark study”. In: *Applied Energy* 264 (2020). DOI: 10.1016/j.apenergy.2020.114693.

[113] Liu Wei, W. Fred Ramirez, and Yu Feng Qi. “Optimal Control of Steamflooding”. In: *SPE Advanced Technology Series* 1.02 (July 1993), pp. 73–82. DOI: 10.2118/21619-pa.

[114] Xian-Huan Wen and J.Jaime Gómez-Hernández. “Upscaling hydraulic conductivities in heterogeneous media: An overview”. In: *Journal of Hydrology* 183.1-2 (Aug. 1996), pp. ix–xxxii. DOI: 10.1016/s0022-1694(96)80030-8.

[115] Curt M. White et al. “Separation and capture of CO2 from large stationary sources and sequestration in geological formations–coalbeds and deep saline aquifers”. In: *Journal of the Air & Waste Management Association* 53.6 (June 2003), pp. 645–715. DOI: 10.1080/10473289.2003.10466206.

[116] A.A. Yousef et al. “A capacitance model to infer interwell connectivity from production and injection rate fluctuations”. In: *Proceedings - SPE Annual Technical Conference and Exhibition*. 2005, pp. 323–341.

[117] Yanhui Zhang and Ibrahim Hoteit. “Feature-Oriented Joint Time-Lapse Seismic and Electromagnetic History Matching Using Ensemble Methods”. In: *SPE Journal* 26.03 (Oct. 2020), pp. 1341–1365. DOI: 10.2118/203847-pa.

[118] Yanhui Zhang, Femke C. Vossepoel, and Ibrahim Hoteit. “Efficient Assimilation of Crosswell Electromagnetic Data Using an Ensemble-Based History-Matching Framework”. In: *SPE Journal* 25.01 (Nov. 2019), pp. 119–138. DOI: 10.2118/193808-pa.

[119] H. Zhao et al. “INSIM: A data-driven model for history matching and prediction for waterflooding monitoring and management with a field application”. In: *Society of Petroleum Engineers - SPE Reservoir Simulation Symposium 2015*. Vol. 1. 2015, pp. 431–461.

[120] Hui Zhao et al. “Flow-Path Tracking Strategy in a Data-Driven Interwell Numerical Simulation Model for Waterflooding History Matching and Performance Prediction with Infill Wells”. In: *SPE Journal* 25.02 (Apr. 2020), pp. 1007–1025. DOI: 10.2118/199361-pa.

[121] D.I. Zubarev. “Pros and cons of applying proxy-models as a substitute for full reservoir simulations”. In: *Proceedings - SPE Annual Technical Conference and Exhibition*. Vol. 5. 2009, pp. 3234–3256.

ACKNOWLEDGEMENTS

Doing my PhD at Delft University of Technology has been an incredible journey. I have had the pleasure of being surrounded by very professional teams of scholars and some of the most friendly and interesting colleagues. Engaging with the professors and fellow researchers here has been an amazing experience because we not only delve deeply into academic topics but also experience the collision of diverse cultural perspectives. I see culture shock as a representation of the world's rich diversity, rather than something negative. I am truly grateful to TU Delft for offering such an international and multicultural platform for us. These four years of doctoral experience will undoubtedly remain cherished and unforgettable memories in my life. I also want to take this opportunity to express my heartfelt gratitude to all the individuals and institutions that supported and guided me along the way.

First and foremost, I want to express my thanks to the China Scholarship Council (CSC). Without their financial support, I doubt I would have successfully completed my doctoral studies. Additionally, I am immensely thankful to my parents for their unwavering support and selfless love. While these four years of pursuing my doctorate have been a period of growth and learning for me as a young man, I am fully aware that for my elderly parents, it meant coping with the effects of aging and loneliness. I regret not being physically present with them during these years, but their comforting video calls during difficult times provided the encouragement I needed to persevere in my studies.

Dear Denis, I feel honored to have had you as my supervisor, guiding my academic work for four years. If given the chance to choose again, I would not hesitate to have you as my supervisor once more. I deeply admire your high passion for research and your wholehearted dedication to scientific research, working together with our entire DARTS team. You consistently made the effort to individually discuss research progress and challenges with each of us, always collaborating to find solutions. Four years ago, you introduced me to the challenging topic of adjoint method, from initially being puzzled by adjoint method, I gradually grasped its essence and even implemented them in C++, eventually sharpening the adjoint method framework in DARTS. I thoroughly enjoyed the process of facing challenges and successfully overcoming them. Moreover, your encouragement to actively participate in various academic conferences not only provided excellent platforms for exchanging ideas with scholars worldwide but also sparked academic collaboration and innovation. And of course, attending conferences in different parts of the world also allowed us to immerse ourselves in the local culture and customs. While we may have had disagreements on certain academic topics, they never hindered our continuous exploration and collaboration on our academic journey. I believe that the adjoint method based on DARTS will not be the end after my graduation, but a new beginning, as we have already opened up new directions for the adjoint method within the MPFA framework. I am confident that this will be a highly challenging and fascinating endeavor.

Dear David, I am grateful to have you as my co-promoter. Although our interactions have been limited, I have vivid memories of the few discussions we have had, such as our exploration of geothermal geological modeling along with your student. It is a shame that I could not join your course of Geothermal Field Trip because of COVID. I also want to express my gratitude for your participation in my go/no-go meetings, yearly progress meetings, and review of my doctoral dissertation.

Xiaocong, I truly appreciate all the help and support you have provided me, both academically and in daily life. Before I prepared to go to the Netherlands, I asked for your advice on numerous matters concerning life in the Netherlands and TU Delft, such as registration, housing, transportation cards, bank cards, and more. You patiently answered every one of my questions. After arriving at TU Delft, you introduced me to the Dutch local customs and shared your experiences of living in the Netherlands. Of course, you also showed me where are those Chinese supermarkets:) Academically, as I was initially unfamiliar with DARTS, I often needed to ask you about DARTS, and you always provided detailed explanations of its source code structure and which C++ functions corresponded to specific DARTS functionalities. Without your help, I do not think I would have made such rapid progress in implementing the adjoint method on DARTS.

Michiel, I enjoy discussing both academic and interesting life-related topics with you. We have attended many academic conferences together, including ECMOR, MATHIAS DAY, SPE conference, SIAM, and each time, our trips and conference discussions have been filled with engaging academic content and shared interesting local anecdotes. I appreciate the invitation to your apartment in Rotterdam when we bid farewell to Gabriel Miranda. You made a great spaghetti dish, and it was delightful to meet Miranda, an interesting friend from Brazil. Additionally, the experience in Norway's fjords and Trolltunga was really fantastic! I am grateful that you convinced me to take that trip after the SIAM conference. The hike to Trolltunga was an unforgettable experience. You often show curiosity about various Chinese and Asian snacks, and occasionally, I share some Chinese snacks with you. I am surprised that you can handle many spicy Chinese foods snacks. As I have mentioned many times, if you ever travel to China, I will definitely show you the local customs and some nice Chinese dishes.

Many thanks to all my colleagues in the DARTS group. I am delighted to have worked and spent four wonderful years with a group of smart and interesting guys. Ilshat, I completed my first "merge request" for the adjoint method code with your help. I greatly admire your computer knowledge and skills, and I learned a lot about version control and Git from you. Aleks, my next major task is to apply the adjoint method to your MPFA engine, and I am thankful for your detailed explanations of the MPFA and for sharing your MPFA model. Every discussion with you about MPFA and poromechanics has been highly informative. Mark, even though our interactions were limited, I appreciate you explaining the DARTS source code architecture and some basic ideas about CSR. Gabriel, thank you for sharing insights on data assimilation. Whenever I had questions about data assimilation, you were always willing to help clarify my doubts. Yuan, although we met four years ago, you only recently became my colleague and sit next to me in the office. Your DAP project sounds very interesting, and I look forward to seeing some interesting research results about the DAP well. George, you have always been a cool guy, even when discussing serious news, history, and politics. I hope you make break-

throughs in combining machine learning with DARTS. Additionally, I want to thank you for helping me proofread the Dutch versions of the Propositions and Summary in this dissertation. So, you literally "contributed" to this dissertation!:) Kiarash, thank you for bringing laughter to our office with your random talks. Stephen, even though we did not have much time to interact, I am thankful for your explanation of how to use the unstructured discretizer and its principles. Yang, thank you for explaining the thermal engine's source code and its usage in DARTS. Luisa and Artur, it has been a pleasure working with you on the open-darts project, and I sincerely hope open-darts continues to thrive.

Kai, we have known each other for many years. Now that you have embarked on a new journey as a postdoc, I wish you all the best in your academic career. Kishan and Willemijn, you always initiate interesting conversations in the office. I am happy to join in your conversations often. Sian, I am amazed by the variety of special and delicious cakes you always bring to us. Ziliang, although your office is in a different building, we still have many opportunities to interact, such as during fluidflower, SIAM, playing badminton together, and the surprise we prepared for your girlfriend at Scheveningen beach.

Due to COVID, there have been about two years when we were all working from home, and as a result, I did not have many chances to have in-depth conversations with more people in our section. Nonetheless, I feel deeply connected to the warm and harmonious atmosphere created by the Reservoir Engineering Section. I am grateful to be a part of this big family.

Mom and Dad, no words can express my love for you, so I hope to continually show my love for you through actions in the future. Special thanks also go to my cousin Lele. In fact, you have been like a sister to me, always supporting and assisting me in every way possible. Jianxin, I am excited about the future and hope to have you by my side all the time.

CURRICULUM VITÆ

Xiaoming TIAN

16-02-1995 Born in Hunan, China.

EDUCATION

2012–2016 Bachelor Degree, Petroleum Engineering
China University of Geosciences (Beijing)

2016–2019 Master Degree, Oil & Gas Field Development Engineering
China University of Petroleum (Beijing)

2019–2023 Doctor of Philosophy, Geoscience & Engineering
Delft University of Technology

Thesis: General-purpose Inverse Modeling Framework for Energy Transition Applications Based on Adjoint Method and Operator-Based Linearization

Promotor: Dr. Denis V. Voskov
Prof. dr. David F. Bruhn

List of Publications

Journal Articles

9. **X. Tian**, O. Volkov, and D. Voskov, *An Advanced Inverse Modeling Framework for Efficient, Flexible, and Adjoint-based History Matching of Geothermal Energy*, under review.

8. **X. Tian**, M. Wapperom, J. Gunning, S. Jackson, A. Wilkins, C. Green, J. Ennis-King, and D. Voskov, *The Study of the History Matching for FluidFlower Experimental Results*, under review.

7. **X. Tian** and D. Voskov, *Efficient Application of Stochastic Discrete Well Affinity (DiWA) Proxy Model with Adjoint Gradients for Production Forecast*, Journal of Petroleum Science and Engineering **210**, 109911 (2022).

6. **X. Tian**, A. Blinovs, M. Khait, and D. Voskov, *Discrete Well Affinity Data-Driven Proxy Model for Production Forecast*, SPE Journal (2020).

5. **X. Tian** and P. Dong, *Well-flow Behavior of Reorientation Fracture Considering Permeability Anisotropy*, Engineering Computations **35**, 2459-2479 (2018).

4. M. Wapperom, **X. Tian**, A. Novikov, and D. Voskov, *FluidFlower benchmark: lessons learned from the perspective of subsurface simulation*, Transport in Porous Media (2023).

3. B. Flemisch et al. **X. Tian**, *The FluidFlower International Benchmark Study: Process, Modeling Results, and Comparison to Experimental Data*, accepted.

2. L. Wang, C. Liu, P. Jiang, H. Zhang, **X. Tian**, N. Zhao, and B. Mahlalela, *Characterization of Two-phase Flow in Porous Media Using Global Mobility*, Journal of Petroleum Science and Engineering **177**, 188-197 (2019).

1. C. Liu, K. Li, **X. Tian**, G. Zhao, Y. Chen, and B. Mahlalela, *Experimental Studies on Production Performance of Oil Reservoirs with Bottom Water*, Journal of Petroleum Science and Engineering **172**, 527-537 (2019).

Conference Proceedings

4. **X. Tian** and D. Voskov, *Efficient Inverse Modeling Framework for Energy Transition Applications Using Operator-Based Linearization and Adjoint Gradients*, SPE Reservoir Simulation Conference, **SPE-212169-MS**.

3. **X. Tian**, M. Asadiesfahani, D. Voskov, and J. Groenenboom, *Stochastic Discrete Well Affinity (DiWA) Model for Data Quality Diagnostic and Production Forecast*, European Association of Geoscientists & Engineers **2022**, 1-16 (2022).

2. **X. Tian**, C. Liu, and K. Li, *Numerical Study of an Expandable Multi-layer Thermoelectric Generator System*, World Geothermal Congress (2020).

1. **X. Tian**, C. Liu, and K. Li, *Modeling of Geothermal Power Generation from Abandoned Oil Wells Using In-situ Combustion Technology*, 43rd Stanford Workshop on Geothermal Reservoir Engineering (2018).