# INTEGRATING NUMERICAL SIMULATIONS, MACHINE LEARNING AND BAYESIAN INVERSION IN INVESTIGATIONS OF HYDRAULIC FRACTURING AND WASTEWATER DISPOSAL OPERATIONS, AND ASSOCIATED SEISMIC ACTIVITY

A Dissertation

by

DAWID SZAFRANSKI

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | Benchun Duan |
| Committee Members, | Mark Everett |
| | Patrick Fulton |
| | Marcelo Sanchez |
| Head of Department, | Julie Newman |

May 2022

Major Subject: Geophysics

# ABSTRACT

The “shale revolution” which started in the United States in the mid-2000s significantly increases in the hydrocarbon production from unconventional reservoirs. As a result, wastewater disposal and hydraulic fracturing has become commonly conducted operations. Thus, understanding of fault and fracture mechanics in the presence of fluid flow is of a great importance to operators, regulators and scientists. In this study, we focus on applications of numerical simulations, machine learning and stochastic inversion in investigations of these processes and associated seismic activity. We first develop a forward modeling method by integrating fluid flow in a poroelastic medium and dynamic rupture on faults to estimate time of earthquake triggering and its magnitude. We explore the parameter space and find that formation permeability and elastic properties, along with fault-well distance have a big impact on accumulation of pressure and stress perturbations and eventually earthquake triggering. We apply the methodology to a real case study of the 2012 $M_w$ 4.8 Timpson (TX) induced earthquake. We reproduce not only the size of the mainshock but also main features of the aftershock sequence, building a direct physical link between wastewater injection and the earthquake. Combining the forward modeling method with machine learning regression and Bayesian inversion, we develop a methodology to better constrain fault frictional parameters and background stress states, which translates to improvement in simulation results. We further develop our methodology to simulate complex hydraulic fracture propagation and its interaction with natural fractures or bedding planes by introducing a

dual permeability model for flow simulations, mixed-mode failure criterion and cohesive elements along the fracture path. We investigate fracture propagation, activation, and interaction behaviors, and analyze the relationship between producing reservoir volume and microseismic cloud extent. We find that in many cases the Stimulated Reservoir Volume (SRV) does not correspond well with the Drained Reservoir Volume (DRV). We apply the methodology to the HFTS-1 (Hydraulic Fracturing Test Site – 1; Midland Basin) experiment case study and discuss implications of our results on field data analyses.

# DEDICATION

I want to dedicate this dissertation to my family who was with me and for me at any time during my work towards doctorate degree.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Benchun Duan, and my committee members, Dr. Mark Everett, Dr. Patrick Fulton, and Dr. Marcelo Sanchez for their guidance and support throughout the course of this research.

I also want to show my appreciation to Dr. Mukul Bhatia and Dr. Richard Gibson for their help and advice in spite of not being on my committee.

Thanks also go to my friends, colleagues and department faculty and staff for making my time at Texas A&M University a good and unique experience. Among those, special thanks to Eric Peavey, Abhishek Prakash, Matthew Dorsey, Qingjun Meng, Matthew Couchman, Bin Luo and Dunyu Liu with whom I had multiple valuable discussions and developed lasting friendships.

Last but not least, I would like to thank my wife, Maureen, my parents and sister, my parents-in-law and the rest of my family and close friends for their encouragement, help, patience and love during challenging times.

# CONTRIBUTORS AND FUNDING SOURCES

## Contributors

This work was supervised by a dissertation committee consisting of Professor Benchun Duan (Advisor), Professors Mark Everett and Patrick Fulton of the Department of Geology and Geophysics and Professor Marcelo Sanchez of the Department of Civil Engineering.

The EQdyna software used for part of the analysis in Chapters 2 and 3 were provided by Professor Benchun Duan. The data used in Chapters 2 and 3 are publicly available datasets from United States Geological Survey (USGS). Part of the data analyzed in Chapter 5 was provided by Neil Stegent from Halliburton.

All other work conducted for the dissertation was completed by the student independently.

## Funding Sources

This Ph.D. study was supported by the Department of Geology and Geophysics in the form of Teaching Assistantship, and the Crisman Institute-Berg Hughes Center fellowship with project number HF.11.19B.

The work was also supported by funds from Francesco Paolo do Gangi/Heep Endowed Professorship in Theoretical Geophysics that was awarded to Professor Benchun Duan, and scholarships from Berg-Hughes Center and College of Geoscience including Carolyn & Dan Pedrotti '53 Fellowship, Leonard Gage Larsen Memorial Fellowship, M.T. Halbouty ’30/AAPG Foundation Endowed Fellowship, Marianne W. & Willard R. Green Scholarship, and Sarah & John Hastings '84 Scholarship.

The research got also external support in form of Charlie and Jean Smith Scholarship, and Anadarko/SEG Scholarship awarded by Society of Exploration Geophysicists. Computational resources to conduct research were provided by Texas A&M High Performance Research Computing (https://hprc.tamu.edu). Center.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

# 1. INTRODUCTION

## 1.1. Research background and motivation

In the last two decades induced seismicity has become a real problem, especially in the areas where historically there was a low level of natural seismicity and are not prepared for occurrence of high-energy ground motion. We can observe a correlation between that increase of seismicity rate and the growth in unconventional hydrocarbon exploitation and waste water disposal operations. However, only small number of wells actually causes the induced earthquakes, even smaller number the felt ones. In some areas, such as Arkansas or California, the mixture of induced and natural seismicity makes it even more problematic to analyze what is the primary cause of certain earthquakes. Therefore, there is a need for deeper and better understanding of fault and fracture mechanics during either hydraulic fracturing or salt water disposal operations. In order to do that, there is a need for development of tools that can establish a reliable causation link between operations and registered seismicity. This is especially important for regulatory bodies who are granting permissions and supervising the operations in a given region. Besides regulatory bodies, the operating companies need a tool that allows for estimation of seismic hazard related to potential seismic activity and properly planning future operations.

The best way to meet those needs is to have software for physics-based modeling which can allow to reproduce phenomena associated with the operations. However, for the

simulations to be reliable, they require well-constrained model parameters. The problem lays in the availability of reliable values for certain medium properties, because of difficulty to get in-situ measurement of them. The solution of that problem is an indirect inversion of these parameters by matching available observations.

Another topic related directly or indirectly to induced seismicity is hydraulic fracturing. It has become a popular and crucial technology used in hydrocarbon extraction from unconventional reservoirs (such as shales or tight sands) or in development of Enhanced Geothermal Systems (EGS). Hydraulic fracturing involves a lot of complex processes, such as hydraulic fracture interactions with natural fractures, bedding planes and previously stimulated hydraulic fractures. To properly design the operations, one needs to understand these phenomena, be able to predict them and have verification tools to track how good the prediction is. Current industry standard tool to track hydraulic fracture propagation and evaluation of stimulated reservoir volume (SRV) is microseismic monitoring. However, SRV estimated based on microseismic cloud extent might not be a good proxy for drained reservoir volume (DRV) and actual extent of hydraulic fracture.

Thus, there is a need for tools that can reliably simulate hydraulic fracture propagation, its interactions with natural fractures and bedding planes, and reproduce associated microseismicity. Such software would be useful for both evaluation of past operations and monitoring effectiveness, and planning of future operations.

## 1.2. Overview on induced seismicity caused by water injection

### 1.2.1. Historical and recent cases of induced seismicity

nduced seismicity caused by high-pressure water injection has been well-recognized phenomenon for several decades (e.g. Healy et al., 1968; Kisslinger, 1976). However, increase of seismic activity in previously aseismic regions after nascent of hydrocarbon exploitation in unconventional reservoirs (e.g. Elsworth, 2013) brought more attention from not only academic researchers but also public authorities and petroleum companies. The peak of that increase was in 2015, with significant drop of seismic rates afterwards (e.g. Keranen and Weingarten, 2018), which can be a cumulative effect of low oil prices, better field practices and stricter regulations.

Waste water (or salt water) is a by-product of hydrocarbon production (e.g. Murray, 2013). Because of logistical and economic reasons, waste water is usually disposed in nearby Class II disposal wells, which in some cases can lead to induced seismic activity. The historical cases of such induced earthquakes have been reported since 1960s, including Rocky Mountain Arsenal (e.g. Healy et al., 1968) and Rangely Oil Field (Raleigh et al., 1976) in Colorado, and Baldwin Hills (Hamilton and Meehan, 1971) in California. More recently, seismicity triggered by salt water disposal, geothermal injections or hydraulic fracturing operations has been reported in the areas not only in US, but all around the world. In Texas, seismicity related to wastewater disposal wells was reported in highly populated Dallas-Fortworth area (Frohlich et al., 2011), in

moderately populated vicinity of Timpson in east Texas (Frohlich et al., 2014) or in Delaware basin nearby Pecos in western Texas (Skoumal et al., 2020; Skoumal and Trugman, 2021). Outside of Texas, there were multiple seismic swarms observed in Guy-Greenbrier region in Arkansas (Horton, 2012), which brought concerns because of vicinity to the New Madrid Fault Zone. Other examples include induced seismicity in Paradox Valley in Colorado (Ake et al., 2005; King et al., 2014), Youngstown in Ohio (Kim, 2013) or most recently reported association of seismic activity with water disposal in San Ardo in Central California (Goebel and Shirzaei, 2021). The most notable case, however, is the seismicity in multiple places in Oklahoma (Kim, 2013; Keranen et al., 2014; Manga et al., 2016), which in 2015 made Oklahoma the most seismically active state, surpassing even California. There are also cases of injection-induced earthquakes outside of United States. Some of them are associated with Enhanced Geothermal Systems projects such as seismicity in Basel, Switzerland (Catalli et al, 2016) or nearby Helsinki (Finland) (Kwiatek et al., 2019). Some cases have been associated directly with hydraulic fracturing in hydrocarbon exploitation, including Horn River Basin case (BC, Canada) (Farahbod et al., 2014) and seismicity in Blackpool (UK) (Clarke et al., 2014).

The presence of induced seismicity in most cases does not cause any significant problems. However, in some cases the earthquakes escalate to record-high magnitudes for a given region causing serious threat to local infrastructure due to shallow hypocenter locations. Among those are Mw5.8 September 2016 Pawnee earthquake (Manga et al., 2016) and Mw5.7 November 2011 Prague earthquake (Keranen et al., 2013) in

Oklahoma, Mw4.8 May 2012 Timpson earthquake (Frohlich et al., 2014) in Texas, and Mw5.4 August 2011 Trinidad earthquake (Barnhart et al, 2014) in Colorado. Studying and getting an understanding of processes leading to these earthquakes are important. It can lead to improvement of seismic hazard assessment, improvement of operation practices and more science-based regulations.

### 1.2.2. Methods used to study induced seismicity.

A very important question that people who analyze potentially induced earthquakes want to address is whether registered seismicity can be associated with a particular disposal or production well and volumes injected into it. Over the decades, there has been a big evolution in the methodology applied to answer this question.

The easiest, simplest and quickest way to provide such an assessment might be using simple questionaries similar to the one proposed by Davis and Frohlich (1993). Those questionnaires take into account spatial proximity of the registered earthquakes to wells in question, both horizontally and vertically, temporal correlation between operations and seismicity, plus background seismic activity in the area prior to beginning of the injection. That approach has been also employed in some more recent works including Frohlich et al. (2011) or Weingarten et al. (2015). While easy and quick to use, that methodology does not provide a causal link between seismicity and injections, and is arbitrary and qualitative by nature.

The evolution of this approach is quantitative calculation of temporal cross-correlation between number of registered earthquakes and injection rates, providing also a measure of time delay between the increase or decrease in injected volumes and the system response in increase or decrease in number of earthquakes. Such analyses have been done, e.g. for Central Arkansas (Horton, 2012), Enola Field in Oklahoma and Paradox Valley in Colorado (Oprsal and Eisner, 2014). Similarly, quantitative correlation can be calculated spatially. As an example, Hincks et al. (2018) established a relation between injection depth and cumulative released seismic moment released during earthquakes located in the given depth interval.

Other studies tried to establish some other types of qualitative correlations between certain characteristics of reported seismicity and their probability of being induced by injection. Again, comparisons of current and historical levels of seismicity in regional studies, such as Ellsworth (2013) for Central and Eastern United States or Atkinson et al. (2016) for Canada give a robust and simple answer that the excess earthquakes are indeed induced but fail to answer which wells, and how and why these wells cause earthquakes while others do not. Skoumal et al. (2015) tried to quantify "swarminess" of seismicity, claiming it to be an indicator of induced earthquakes. However, it is not difficult to find examples of potentially induced earthquakes that have normal mainshock-aftershock sequences. Among them are the Pawnee earthquake in Oklahoma (Manga et al., 2016) or the Timpson earthquake in Texas (Frohlich et al., 2014). There were also attempts of using the characteristics of seismic source (e.g. Huang et al., 2016;

Sumy et al., 2017) or ground motion (e.g. Atkinson, 2015; Yenier et al., 2017) to distinguish between natural and induced earthquakes. Goebel et al. (2016) also proposed using b-value taken from Gutenberg-Richter distribution for a catalog as a measure to distinguish induced seismicity, since higher b-value is associated with dominance of smaller-magnitude events and occurrence of seismicity in swarms.

A problem with all the methods described so far is the fact that they are generally looking for some form of data-based correlation between injection and produced seismicity, but they fail to provide a causation link between the two. A solution for this issue is implementation of physics-based modeling.

### 1.2.3. Numerical modeling in induced seismicity investigations.

Numerical modeling is always a form of approximation and simplification of complexity to most important phenomena playing a role in a given process. In case of induced seismicity caused by water injection, it is necessary to simulate water flow in a medium, how it affects stress on the fault and how those changes impact the potential for earthquake triggering.

The simplest approximation of such problem is using simulations of pressure diffusion in porous media (e.g. Shapiro et al., 2005; Keranen et al., 2014; Keranen and Weingarten, 2018; King et al., 2016). Even though this approach might be useful in

certain cases of induced seismicity, using just pressure diffusion disregards some important phenomena and the results may lead to erroneous conclusions.

Usually, we consider a medium, which is at least partially deformable and fluid which is compressible to a certain degree. Therefore, it is important to incorporate hydromechanical coupling into modeling. The most common form of such a coupling is theory of linear poroelasticity (Biot, 1941), assuming that pressure and stress can elastically deform solid medium and fluid. The result can be a far-field stress perturbation, which can contribute to fault activation. Figure 1.1 illustrates different ways the fault approaches failure due to either pore pressure or stress perturbations. In the recent literature we can find multiple examples. Segall and Lu (2015) implemented analytic solutions for an isotropic homogenous model to investigate the influence of poroelastic effects on the seismicity rate. Fan et al. (2016) investigated influence of permeability on fault activation with 2D numerical modeling of the Timpson (TX) earthquake. Juanes et al. (2016) probably disproved the possibility of 2012 Cavone (Italy) earthquakes to be induced by the injection and/or extraction in the area. Goebel et al. (2017) investigated whether Farview (OK) earthquakes triggering was predominantly caused by poroelastic stress transfer or pressure diffusion using both analytical and numerical approaches.

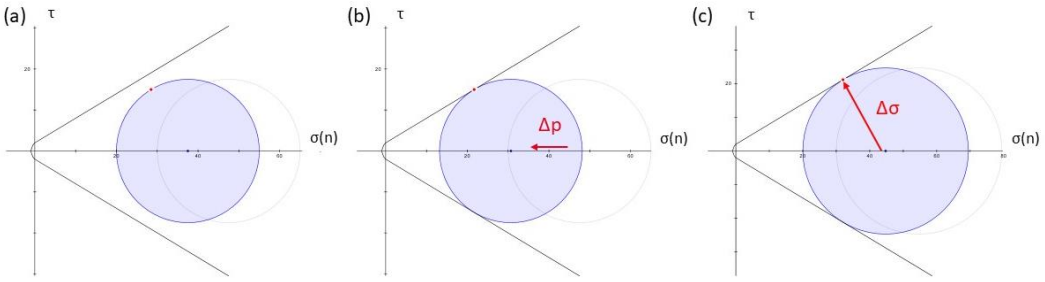


**Figure 1.1 Mohr circles and failure envelopes for (a) initial stress state, (b) increased pore pressure, and (c) increased stress.**

There are further possible extensions of modeling of poroelastic hydromechanical coupling. One of them is implementation of rate-and-state friction laws in order to investigate not only seismic but also aseismic slip caused by water injection (e.g. Alghannam and Juanes, 2020). Such approach was implemented in some real case studies, e.g. analysis of 2011 Mw 5.6 Prague (OK) sequence (Norbeck and Horne, 2016).

Another way, the quasi-static modeling of fault activation can be extended by inclusion of dynamic phenomena. Appropriate simulation of the coseismic phase allows to reproduce rupture size, earthquake magnitude and associated wavefield and ground motion. There are two main approaches to include these phenomena – quasi-dynamic or fully dynamic rupture simulations. While the first one is simpler to implement, the approximations made affect rupturing pattern during an earthquake and give different answers than fully dynamic solutions (Lapusta et al., 2000).

Because of higher computational cost and multiple challenges related to dynamic rupture simulations, there were only limited number of studies where they were used for investigation of induced seismicity. Most of them utilize simple 2D generic models to understand basic phenomena. For example, Jin and Zoback (2018) investigated rupture patterns and velocities with series of 2D simulations for homogenous plane-strain model. More recently, Yang and Dunham (2021) investigated injection-induced aseismic slip using 2D model with rate-and-state friction, and how it is affected by the evolution of porosity and permeability. Even fewer works combines fluid flow and rupture propagation simulations and applies it to real case study. Besides my work presented in this dissertation (Chapters 3 and 4) on the Timpson (TX) earthquake, Urpi et al. (2016) constructed a 2D plane strain model and simulated idealized CO2 injection with multiphase flow and combined it with rate-and-state friction law.

### 1.2.4. Objectives of this study

In our study we further develop and improve methodology of induced earthquakes simulations. We propose an integrated framework of fluid flow simulations in deformable porous media and dynamic rupture propagation modeling to study induced earthquakes. In comparison to previous studies, we go beyond 2D simulations for simple models. We identify that removing one dimension we lose some effects that are only observable in 3D medium.

In our first analysis, we test methodology on hypothetical models with homogenous medium to explore parameter space, establish importance of different parameters on the results, and observe basic effects connected to 3-dimensional geometry of the system. In the second analysis, we go beyond hypothetical models and attempt to apply improved modeling methodology to real case study of Timpson (TX) earthquake. In this study our aim was to confirm the link between reported seismicity and wastewater disposal, and reproduce not only timing but also a reported size of an earthquake(s) utilizing unique capabilities of our software. It creates also an opportunity to get a better insight into changes of specific model parameter values on the results of simulation in a realistic medium model.

### 1.3. Overview on hydraulic fracturing technology development and application.

#### 1.3.1. History of hydraulic fracturing and key technological developments.

Hydraulic fracturing is not a new technology and has been in use since its introduction by Stanolind Oil company and patented by Halliburton Oil Well Cementing Company (Montgomery and Smith, 2010). In its early days treatments were done with refined and crude oils as fracturing fluids and screened river sand as a proppant.

By 2008 over 50,000 stages were completed and many new inventions introduced. In 1950s and 1960s, water mixed with gelling agents replaced oils as fracturing fluid. At that time also river sand was replaced with US-standard-mesh sand as proppant and its concentrations were substantially increased. Another important invention related to

hydraulic fracturing done in 1950s and 1960s was introduction of simple, but powerful engineering models, including KGD (Khristianovic and Zheltov, 1955; Geertsma and de Klerk, 1969) and PKN (Perkins and Kern, 1961) models, and their implementations on early computers. Even though relatively simple, these models were far superior to manual calculations, and using charts and nomographs, which were very time consuming and very ineffective (Montgomery and Smith, 2010). In 1970s we observed the introduction of metal-based crosslinking agents, which were temperature resistant and had higher viscosity.

Despite long history of hydraulic fracturing, it was not as popular as it is today until the first decade of XXI century. Zoback and Kohli (2019) point out three key inventions that combined together, led to dramatic increase in production – horizontal drilling, multi-stage fracturing and slickwater fracturing fluid. In horizontal drilling, first section of the well is still vertical, but slightly above the target zone, well starts to gradually curve (from so called "kick-off" point) and changes direction until it turns into horizontal section landing in a target formation. The change of direction during drilling is allowed thanks to the use of two modes of drilling – sliding and rotating. The horizontal leg can extent up to several kilometers from the vertical section allowing for extensive depletion of the reservoir during multiple stages of hydraulic fracturing. It is possible to stay in the zone using mud-logging techniques and another important invention – Logging While Drilling (LWD), in which geophysical measurements are performed by instruments mounted behind the drill bit. In the well design stage, engineers aim to orient the

horizontal portion of the well along the direction of minimum horizontal stress ($\sigma_{hmin}$). In modern hydraulic fracturing projects, a single vertical well can be eventually diverted into multiple horizontal wells with possibility to hit multiple targets. A second key invention – multi-stage hydraulic fracturing allows for targeted depletion of the reservoir around selected sections of the well. There are two main techniques used for well completion – "plug-and-perf" method for the boreholes with casing and "sliding sleeve" method for open boreholes (e.g. Burton, 2016; Zoback and Kohli, 2019). For better efficiency of the completion, typically multiple stages of hydraulic fracturing are scheduled simultaneously. Finally, the invention of slickwater, which has much lower viscosity than gels, allows for fracturing fluid to travel much further away and stimulate natural fractures, which has been recognized as equally important to tensile propagation of hydraulic fracture itself (Zoback and Kohli, 2019). However, using of slickwater compromises the ability to transport proppant, resulting in much smaller fractures conductivities at greater distances.

### 1.3.2. Methods used to study hydraulic fracture propagation.

To efficiently design hydraulic fracturing operations, it is necessary to understand processes happening in the subsurface during stimulation, have in-depth knowledge about the geology and use reliable tools for scenarios evaluation. There are three main ways these goals can be met.

First way is performing field experiments. These experiments provide direct observations at the field scale, which can be easily applied to similar operations planned in the future. During these experiments a plentiful of data is collected, including well logs, core samples, rock samples, geophysical monitoring recordings (e.g. microseismic, fiber optics, etc.), or production data. A historical example of such experiment is 3rd GRI Staged Field Experiment conducted in east Texas (e.g. Robinson et al., 1992). More recent examples are Hydraulic Fracturing Test Site - 1 (HFTS-1) project in Midland Basin (e.g. Ciezobka et al., 2018) or Hydraulic Fracturing Test Site - 2 (HFTS-2) project in Delaware Basin (e.g. Pudugramam et al., 2021). These experiments provide also a great way to test new ideas, methods and tools. However, they are also very expensive, often require wide collaboration between different entities and might be applicable to only a particular part of the basin.

Second way to study hydraulic fracturing is performing laboratory experiments on cores and rock samples. This approach provides an opportunity to perform various tests in controlled environment and make observations on specific rock samples. Some of these experiments may also serve as a useful benchmark for numerical models. Well-setup experiments provide an important insight into basic phenomena and may lead to improvements of theoretical models. As an example, Wu et al. (2008) performed experiments on fracture height growth in layered media. Other researchers (e.g. Xing et al., 2017) proved theoretically predicted multi-scale nature of near tip region of hydraulic fractures. There are also experiments showing hydraulic fracture curving in the

vicinity of weak planes (e.g. Bunger et al., 2011). However, there are also several limitations of laboratory experiments. First of all, results and observations might not be directly applicable because of hydraulic fracture propagation scale dependence. Secondly, reservoir conditions might be very difficult to replicate in the laboratory. Third, obtaining statistically important sample of observations might be very time consuming for some of the experiments if one wants to generalize the conclusions.

Third way of investigation of hydraulic fracturing process is to use computer simulations. This approach has multiple advantages. First, calculations can be scaled to test scenarios in both laboratory and field scales. Secondly, in comparison to field and laboratory experiments they are cheap and relatively fast to perform. With availability of powerful computation clusters and even good personal computers, engineers or researchers can test a lot of different possible scenarios. Moreover, besides retrospective analysis of past operations, they can serve as a predictive tool for hydraulic fracture design. The downside of any modeling or simulation is the fact that we make different assumptions and get either better or worse approximation of actual processes. Thus, understanding of most important phenomena for a given case is crucial in selecting the type of modeling. Moreover, the more sophisticated the model is, the more input data it requires. Therefore, without well-constrained parameters, the results of simulation might present less of a value from practical standpoint for real case analysis or prediction. In the next section we elaborate more on this approach.

Before moving on to numerical methods, let us look at one specific type of field measurements which is commonly used in industry practice – microseismic monitoring. Detection and location of very weak (often times with negative magnitudes) seismic events is a popular way of SRV estimation (e.g. Zoback and Kohli, 2019), under the assumption that microseismic events are a result of hydraulic fracture propagation and its hydraulic interactions with preexisting natural fractures.

Similar to hydraulic fracturing, microseismic monitoring is not a recent invention. Its applications can be dated back to early XX century when seismic monitoring started to be used in mining. For example, in a long-term study of stresses and microseismic events in underground mines, Obert and Duvall (1957) made several important observations, such as a correlation between high stress and microseismic events concentrations, origin of microseismic events on faults and fractures or requirement of multi-channel recording for correct event location.

Another important application of microseismic monitoring are observations of waste fluid injection (discussed earlier) and monitoring of geothermal energy extraction. For the latter, Fenton Hill geothermal experiment in 1971-1995 is a staple example (e.g. Brown et al., 2012), which resulted in development of high-temp 3C downhole sensors, inversion for location of events from arrival times, and application of continuous monitoring of induced seismicity.

Finally, microseismic monitoring was used in different hydraulic fracturing projects and experiments. In 1973-1974 El Paso Natural Gas Company used surface monitoring in San Juan Basin (New Mexico) and Green River Basin (Wyoming) (e.g. Power et al., 1976). The M-site experiment (1992-1995) (e.g. Warpinski et al., 1996) resulted in introduction of automatic event detection, 3C broadband downhole sensors in multi-level arrays, and using perforation shots for orienting receivers. In Cotton Valley experiment in Texas (1997-1999) (e.g. Urbancic and Rutledge, 2000) high quality (even for modern standards), multi-well microseismic data was acquired, which allowed invention of relative-event location method or automation in event location.

There are two main types of microseismic monitoring. First one is surface monitoring, which was popularized by Microseismic Inc and their star-like arrays (e.g. Duncan and Eisner, 2010). With surface monitoring high volume of data is registered with simple 1C geophones and event location can be done using only a single wave arrival visible on multiple receivers. However, the biggest limitation is the detectability of small events in presence of a noise. The other type is downhole monitoring (e.g. Grechka and Heigl, 2017) with either single well (more typical) or multiple well arrays. This type of monitoring emerged during the Fenton Hill experiment. It provides a very good sensitivity and detectability of even very small events. However, deployment of downhole arrays is more complicated and more expensive. Additionally, for a single well array, there is a problem of azimuthal uncertainty of microseismic event location.

### 1.3.3. Numerical modeling in studies of hydraulic fracture propagation.

As mentioned in one of the previous sections about hydraulic fracturing technology, computer calculations and simulations introduced in 1950s and 1960s greatly improved the effectiveness of operation design. Although very useful, KGD (Khristianovic and Zheltov, 1955; Geertsma and de Klerk, 1969) and PKN (Perkins and Kern, 1961) models have some major limitations. First, those are solutions for 2D homogenous models, which in real world never exist. Secondly, the assumption of constrained height prevents from studying of vertical hydraulic fracture propagation.

One of the popular extensions of PKN model is Pseudo-3D model (e.g. Warpinski et al., 1993; Lecampion et al., 2018), in which 3D problem is reduced to 2D elastic mechanical problem and 1D fluid flow problem. This allows for studying also vertical growth of hydraulic fracture and extending homogenous models to piece-wise homogenous layered models. However, height predictions provided are sometimes inaccurate (Peshcherenko and Chuprakov, 2021) and the fracture geometry is still very limited.

The basic hydraulic fracture model consists of several important elements (Lecampion et al., 2018). First one is using of Linear Elastic Fracture Mechanics (LEFM) for solid mechanics part. Second one is approximation of fluid flow in the fracture with cubic law, which is derived based on Darcy's Law. Third element is describing fluid escape from the fracture to the surrounding medium by Carter Leak-Off model (Howard and

Fast, 1957). Finally, there is an assumption of zero pressure and zero width at the fracture tip during the propagation.

That basic model has been extended in different ways by the researchers, by implementation of more realistic physics. First improvement might be replacing the elastic medium with poroelastic medium as in work of Carrier and Garnet (2012). Moreover, linear fluids can be replaced by non-linear fluids (e.g. Dontsov and Pierce, 2015) and linear solids with non-linear solids (e.g. Needleman, 2014). Finally, simple constant injection assumption might be upgraded to non-linear injection conditions (e.g. Lecampion and Desroches, 2015), which can allow for testing different pumping schedules and their effects.

In their review, Lecampion et al. (2018) categorizes simulation methods into two families. First family focuses on explicit tracking of fracture front at each time step (most LEFM-based methods). Second family puts the emphasis on capturing location of fracture front (e.g. cohesive zone models).

Two most popular numerical approaches of solving these hydromechanical problems are Boundary Element Methods (BEM) and Finite Element Methods (FEM). First approach is very computationally efficient, but is good only for homogenous or simple layered models and provides solution only on the fracture surface. Second approach, even though computationally expensive, is much more versatile when it comes to complexity

of models that can be handled. Within that approach there are different variations including fracture growth with remeshing (e.g. Salimzadeh et al., 2016), fracture growth along predefined path (e.g. Chen, 2012), extended FEM method (e.g. Gupta and Duarte, 2014), phase field methods (Miehe et al., 2015), or meshless methods (e.g. Samimi and Pak, 2016).

### 1.3.4. Objectives of this study

We identify that there is a major disconnection between efforts to simulate realistic hydraulic fracture propagation and simulate microseismicity associated with that propagation. Usually, experts in petroleum engineering focus on the first aspect, while geophysicists put emphasis on the latter. There are only single works, in which both phenomena are modeled altogether (e.g. He and Duan, 2021).

One of the objectives of this study is to develop methodology and tool for integrated modeling of fluid flow, hydraulic fracture propagation and generation of microseismic events. We derive equations and implement them to simulate fluid flow in dual-porosity, dual-permeability, deformable medium (flow through porous rock and through fracture network and hydromechanical coupling; e.g Mehrabian and Abousleiman, 2014). We also implement complex fracture model, where fracture can propagate in almost any direction (limited by mesh) and cohesive zone elements adaptively added along the propagation path to emulate local plastic deformations. Through quasi-static analysis of

stress and pressure transfers and mixed-mode failures, our method tracks the location and size of micro-tremors, along with associated stress changes.

In the analysis part, we take advantage of our software capabilities and analyze hydraulic fracture propagation in 3D layered media, its interactions with natural fractures and bedding planes and associated microseismicity generation. This lays a foundation for proposed framework of establishing the extent of Drained Reservoir Volume based on data-derived Stimulated Reservoir Volume by incorporating our simulation methodology.

## 1.4. Overview on inversion methods and machine learning techniques with applications in geosciences.

### 1.4.1. Geophysical inversion methods and applications.

As discussed in the section about modeling, complex numerical simulations require well-constrained model parameters for the results to be useful in practice. Performing geophysical inversion is a way to constrain parameters that are hard to be measured directly. Permeability of the rock formations is a great example of such parameter, which is necessary and very important in fluid flow modeling (e.g. Fan et al., 2016; Shirzaei et al., 2019). In dynamic rupture modeling, background stress state and fault frictional properties are also essential model parameters (e.g. Gallovic et al., 2019).

In a very broad sense, inversion techniques can be classified into one of two main categories – deterministic and stochastic. One of the popular deterministic techniques is grid-search (or uniform search) technique (e.g. Sambridge and Kennett, 1986). It is simple and covers whole parameter space but is very inefficient in case of multiparameter inversions. Another popular techniques are gradient methods, including steepest descent and Newton-Rhapson algorithms (e.g. Gill et al., 1981). They are computationally very efficient, but do not perform well for strongly non-linear problems and final solution might indicate local, not global minimum. There are also advanced deterministic techniques trying to tackle mentioned problems, such as Simplex Method (Amoeba search) (e.g. Nelder and Mead, 1965) or Tabu search method (Vinther and Mosegaard, 1996). Furthermore, all deterministic techniques lack the ability to reliably estimate the uncertainty of the solution.

In the second category, there are also different algorithms and techniques available. One of them is Importance Sampling (e.g. Sambridge and Mosegaard, 2002), in which Monte Carlo algorithm is designed in the way that sampling density is proportional to probability density. Most well-known Importance Sampling algorithm is Metropolis-Hastings algorithm (e.g. Metropolis et al., 1953). Another group of stochastic methods are Simulated Annealing Methods (Rothman, 1985, 1986; Kirkpatrick et al., 1983), which are non-uniform Monte Carlo methods for global optimization. Then, there are also Genetic Algorithms (e.g. Goldberg, 1989) and Evolutionary Algorithms (e.g. Minster et al., 1995), which are fully non-linear and do not use derivatives of objective

function that is being minimized. Finally, the Neighborhood Algorithm (e.g. Sambridge, 1999) shares the similarities with Genetic Algorithms and Simulated Annealing, generating samples based of the results obtained from previous samples. There are three main advantages of stochastic inversion methods. They provide a simple way to understand uncertainties, they are resistant to be trapped in local minima and they are able to handle any non-linear problems. The biggest disadvantage, however is high computational cost required to generate tens of thousands of samples.

In this work, we combine physics-based simulations with the inversion of model parameters. There are only limited examples of these simulation-based inversions in studies of earthquakes to constrain medium parameters. In one example, Shirzaei et al. (2019) inverted for the hydraulic diffusivity of the injection formations using the InSAR data and poroelastic 3D model for Mw 4.8, May 2012, Timpson (TX) earthquake with the grid-search technique. Another example is the inversion of fault frictional parameters based on the data for Mw6.2, August 2016, Amatrice (Italy) tectonic earthquake using Monte Carlo inversion technique by Gallovic et al. (2019). In this study we use modified Metropolis-Hastings algorithm and Bayesian approach to the inversion, because of non-linearity of the problem and multi-dimensional parameter space. However, we go one step further and instead of generating all our samples with physics-based simulations, we generate statistically significant set (several hundred samples) of the simulation results, which then is used to train and optimize machine learning regression algorithm (see short overview in the next section). Finally, in inversion sampling we use trained

regression algorithm and have a result validation procedure described in Chapter 4. Replacing simulations with advanced regression keeps the link between physical parameters and the outcome, and greatly reduces computation time of the inversion.

### 1.4.2. Machine learning techniques and their application in geosciences.

Machine learning is a subset of artificial intelligence methods. In recent years, with the increase of data volumes, it become necessary to automatize some processes and have a tool, which can help in analysis of the datasets. There are three major categories of machine learning algorithms – supervised, unsupervised and reinforced learning (e.g. Sharma and Kumar, 2017). Supervised learning is a task of inferring a function from a labeled data(e.g. Sharma and Kumar, 2017). With the supervised learning one can either perform classification (assigning each sample to a certain class based on similarity between input and output relation in comparison to training samples) or regression (assigning a sample a certain output parameter(s) value(s) based on the function revealed during training on labeled data). Unsupervised learning, on the other side, is a task of inferring a function to depict concealed structure from unlabeled data (e.g. Sharma and Kumar, 2017). The most common unsupervised learning task is clustering, i.e. finding subsets in data with similar behavior/properties. In reinforced learning, algorithm interacts with the environment and receives feedback on rewards and penalties as the task progresses (e.g. Sharma and Kumar, 2017).

In this work, the focus is on the supervised learning. Among algorithms popularly used for classification jobs, there are Logistic Regression (e.g. Wright, 1995), Support Vector Machines (SVM) (e.g. Cortes and Vapnik, 1995), K-Nearest Neighbors (e.g. Nguyen et al., 2016), Neural Networks (e.g. Murtagh, 1991). For regression jobs, a user can apply technically the same algorithms (with some exceptions). The main difference is that output variable in classification is discrete (number representing specific class, such as formation flag), while in regression continuous (value of e.g. density stochastically predicted). The list of popular regression algorithms includes Linear Regression, Ensemble Methods such as Random Forest (e.g. Liaw and Wiener, 2002) or Bagging (e.g. Breiman, 1996), Decision Trees (e.g. Belson, 1959), K-Nearest Neighbors (e.g. Nguyen et al., 2016), Neural Networks (e.g. Murtagh, 1991). Each of those algorithms has its pros and cons and sometimes simpler algorithms can outperform the more complex ones. In multitude of single-variable regressions, there is no need to search beyond Linear or Polynomial Regression. However, as the number of input and output parameters increases the more advanced algorithms such as Random Forest or Neural Networks start to prevail due to parameter space complexity and non-linearity.

In recent decade, there was a rapid increase of applications where machine learning has been used. To narrow down to the realm of geoscience, machine learning is applied to tasks such as event location (e.g. Dodge and Harris, 2016), fracture modeling (e.g. Valera et al., 2018), fault failure prediction (e.g. Rouet-Leduc et al., 2018), seismic interpretation (Di et al., 2017) and salt detection (Gramstad and Nickel, 2018), well-log analysis (Saporetti et al., 2018), 3D timeshift extraction from 4D data (Dramsch et al.,

2019), seismic event classification of volcanic activity (Titos et al., 2018) or prediction of rupture interaction with heterogeneities (Ahamed and Daub, 2019). For more information regarding applications and advances of machine learning techniques in Earth sciences we refer to a review by Dramsch (2020).

### 1.5. Overview of the dissertation.

In this research, we have several main objectives we want to meet. First, we want to provide a methodology and a toolbox to simulate and analyze seismic activity induced by high-pressure injection operations, including wastewater disposal and hydraulic fracturing. Our goal is to develop software which can be applied not only to simple hypothetical models and theoretical studies, but also to real case scenarios.
Secondly, our aim is to replicate field operations in order to establish causational relationship between the injection and produced induced earthquakes or microseismicity.

We want to reproduce not only time of earthquake triggering but also its size (magnitude), which we identify as a lacking component in previous studies.
We also aim to take advantage of partially controlled environment of earthquake triggering and improve our understanding of the subsurface processes and properties.
One of our goals is to combine numerical modeling with Bayesian approach to inversion and field observations to improve geomechanical models of selected areas.

Last but not least, we want to show that proposed methodology has practical aspects in estimation of producing reservoir volume or level of seismic hazard, either in retrospective analysis of past pertains or design of the future ones.

All our projects are linked together and show the evolution of methodology and understanding of studied phenomena. We start from less complex simulations using simple hypothetical models, moving to adding complexity and applying it to real case study in the first two projects. During that time, we realize that simulations can be supplemented with inversion in order to better constrain some of the parameters used in modeling. In the last part, we modify and adapt existing tools to more complex and more difficult problem of hydraulic fracturing.

In Chapter 2, we present the project in which we implement analytical solutions of linear poroelasticity for point source in a 3D homogenous medium with a constant injection rate by Rudnicki (1986), and integrate it with a code for finite-element simulations of dynamic rupture propagation (EQdyna). Through parameter space exploration, we investigate the influence of different properties of the rocks and fluids, as well as the geometry of the operations on the potential for earthquake triggering. We find that three main driving factors are the distance between the injection well and the fault, permeability of the medium and shear modulus of the rocks. The novelty of this project lays in combination of poroelasticity and dynamic rupture modeling in 3D model laying a bridge to more advanced modeling.

In Chapter 3, we show a work in which we implement finite-difference solution of system of partial differential equations for stress tensor and pore pressure perturbations due to fluid flow in 3D heterogenous, deformable, porous medium, and integrate it with dynamic rupture propagation code EQdyna adopted for study of induced earthquakes. We apply the methodology to 17 May 2012, Mw 4.8 Timpson, TX earthquake case study. We were able to reproduce main shock (magnitude and focal mechanism) and approximately constrain fault frictional parameters. We show that uniform effective normal stress is probably the background stress pattern within the fault zone. Additionally, we manage to roughly reproduce main features of the aftershock sequence. The novelty of this study is application of integrated simulation of fluid flow in 3D poroelastic medium and dynamic rupture simulation to a real case study.

Chapter 4 covers the integration of numerical modeling, machine learning and Bayesian inversion into one framework and its application to the Timpson (TX) earthquake by inverting for background stress state and fault frictional properties. We also developed inversion module implementing modified Metropolis-Hastings algorithm (Monte Carlo type of inversion) and a module for machine learning regression algorithms training, optimization and comparison. We present how to form a training dataset from several hundred simulations of Timpson earthquake with variation of selected parameters in dynamic rupture model. We find that Random Forest, Bagging and K-Neighbors regression algorithms are most promising to be used as replacement of physics-based

simulations in inversion. We also get better constrains on stress state and frictional parameters with associated uncertainties by matching the observations and successfully cross-validate the results of inversion by running actual physics-based simulations.

In Chapter 5, we extend the fluid flow code to dual-porosity, dual permeability poroelastic media (flow through formations and within fractures) and develop two new modules. First module handles the fracture propagation in 3D heterogenous, poroelastic medium with the use of cohesive zone method (material degradation, time dependent aperture and permeability) and mixed-mode failure criterion. Second module is used for the generation of microseismicity due to quasi-static failure and calculation of stress changes caused by that failure. We show code verification against PKN model solution and published results for numerical modeling (Peshcherenko and Chuprakov, 2021) in heterogenous layered medium. In this chapter, we investigate the hydraulic fracture interactions with natural fractures and bedding plane for hypothetical models and for simplified model of HFTS-1 stimulation of well SUGG 171-6SU Stage 2 (Midland Basin, Upper Wolfcamp). We also investigate the relation between hydraulic fracture extent (Drained Reservoir Volume; DRV) and microseismic cloud extent (Stimulated Reservoir Volume; SRV). We find that presence of natural fractures can affect the hydraulic fracture propagation pattern even without direct hydraulic connection between the two. We show that there might be a big discrepancy between SRV and DRV, where SRV can either underestimate or overestimate the DRV, dependent on medium and

natural fractures properties. Finally, we find that both activation of bedding plane and natural fractures are equally possible mechanisms of microseismicity generation.

Chapter 6 contains the summary of the main findings of the research presented in this dissertation. It also shows the possible future directions of research on induced seismicity and microseismicity generated by high-pressure injection operations.

### 1.6. References

Ahamed, S., & Daub, E. G. (2019) 'Machine learning approach to earthquake rupture dynamics', arXiv preprint arXiv:1906.06250.

Alghannam, M., & Juanes, R. (2020) 'Understanding rate effects in injection-induced earthquakes', *Nature communications*, 11(1), pp. 1-6. https://doi.org/10.1038/s41467-020-16860-y

Ake, J., K. Mahrer, D. O'Connell, and L. Block (2005) 'Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado', *Bull. Seismol. Soc. Am.*, 95(2), pp. 664–683.

Atkinson, G. M. (2015) 'Ground-motion prediction equation for small-to-moderate events at short hypocentral distances, with application to induced-seismicity hazards', *Bull. Seismol. Soc. Am.*, 105, pp. 981–992.

Atkinson, G. M., et al. (2016) ‘Hydraulic fracturing and seismicity in the Western Canada Sedimentary Basin’, *Seismol. Res. Lett.*, 87, pp. 631–647.

Barnhart, W. D., et al. (2014) ‘Seismological and geodetic constraints on the 2011 Mw5.3 Trinidad, Colorado earthquake and induced deformation in the Raton Basin’, *J. Geophys. Res.*, 119, pp. 7923–7933

Belson, W. A. (1959) ’Matching and Prediction on the Principle of Biological Classification’, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 8(2), pp. 65–75. https://doi.org/10.2307/2985543

Biot, M. A. (1941), ‘General theory of three-dimensional consolidation’, *J. Appl. Phys.*, 12(2), pp. 155–164.

Breiman, L. (1996) ‘Bagging predictors.’, *Machine learning*, 24(2), pp. 123-140. https://doi.org/10.1007/BF00058655

Brown, D. W., et al. (2012) *Mining the earth's heat: hot dry rock geothermal energy*, Springer Science & Business Media.

Bunger, A. P., et al. (2011) ‘Experimental investigation of the interaction among closely spaced hydraulic fractures’, *45th US Rock Mechanics/Geomechanics Symposium*, June 2011.

Burton, W.A. (2016) ‘Multistage completion systems for unconventionals’, *U. Ahmed & D.N. Meehan (eds.), Unconventional oil and Gas Resources: Exploration and Development*, Boca Raton, FL: CRC Press

Carrier, B., & Granet, S. (2012) ‘Numerical modeling of hydraulic fracture problem in permeable medium using cohesive zone model’, *Engineering fracture mechanics*, 79, pp. 312-328.

Catalli, F., et al. (2016) ‘The importance of earthquake interactions for injection-induced seismicity: Retrospective modeling of the Basel Enhanced Geothermal System’, *Geophys. Res. Lett*., 43(10), 4992-4999.

Chen, Z., et al. (2009) ‘Cohesive zone finite element-based modeling of hydraulic fractures’, *Acta Mechanica Solida Sinica*, 22(5), pp. 443-452.

Ciezobka, J., Courtier, J., & Wicker, J. (2018) ‘Hydraulic fracturing test site (HFTS)-project overview and summary of results’, *SPE/AAPG/SEG Unconventional Resources Technology Conference*, July 2018.

Clarke, H., et al. (2014) ‘Felt seismicity associated with shale gas hydraulic fracturing: the first documented example in Europe’, *Geophys. Res. Lett.*, 41, pp. 8308–8314.

Cortes, C., & Vapnik, V. (1995) ‘Support-vector networks’, *Mach. Learn.*, 20, pp. 273–297. https://doi.org/10.1007/BF00994018

Davis, S. D., & Frohlich, C. (1993) ‘Did (or will) fluid injection cause earthquakes? Criteria for a rational assessment’, *Seismol. Res. Lett.*, 64(3-4), pp. 207–224.

Di, H., Shafiq, M. A., & AlRegib, G. (2017) ‘Seismic-fault detection based on multiattribute support vector machine analysis’, *2017 SEG International Exposition and Annual Meeting*, September 2017.

Dodge, D. A., & Harris, D. B. (2016) ‘Large-scale test of dynamic correlation processors: Implications for correlation-based seismic pipelines’, *Bulletin of the Seismological Society of America*, 106(2), pp. 435-452.

Dontsov, E. V., & Peirce, A. P. (2015), ‘Proppant transport in hydraulic fracturing: crack tip screen-out in KGD and P3D models’, *International Journal of Solids and Structures*, 63, pp. 206-218.

Dramsch, J. S. (2020) ‘70 years of machine learning in geoscience in review’, *Advances in Geophysics*, 61, 1.

Dramsch, J. S., et al. (2021) ‘Deep Unsupervised 4-D Seismic 3-D Time-Shift Estimation With Convolutional Neural Networks’, *IEEE Transactions on Geoscience and Remote Sensing*.

Duncan, P. M., & Eisner, L. (2010) ‘Reservoir characterization using surface microseismic monitoring’, *Geophysics*, 75, pp. 75A139-75A146.

Ellsworth, W. L. (2013) ‘Injection-induced earthquakes’, *Science*, 341, 1225942.

Fan, Z., Eichhubl, P., & Gale, J. F. W. (2016) ‘Geomechanical analysis of fluid injection and seismic fault slip for the Mw4.8 Timpson, Texas, earthquake sequence’, *J. Geophys. Res. Solid Earth*, 121, pp. 2798–2812.

Farahbod, A. M., et al. (2015) ‘Investigation of regional seismicity before and after hydraulic fracturing in the Horn River Basin, northeast British Columbia’, *Can. J. Earth Sci.*, 52, pp. 112–122.

Frohlich, C., et al. (2011) ‘The Dallas-Fort Worth earthquake sequence: October 2008 through May 2009’, *Bull. Seismol. Soc. Am.*, 101(1), pp. 327–340.

Frohlich, C., et al. (2014) ‘The 17 May 2012 M4.8 earthquake near Timpson, East Texas: An event possibly triggered by fluid injection’, *J. Geophys. Res. Solid Earth*, 119(1), pp. 581-593.

Gallovič, F., et al. (2019) ‘Bayesian dynamic finite-fault inversion: 2. Application to the 2016 Mw 6.2 Amatrice, Italy, earthquake.’, *Journal of Geophysical Research: Solid Earth*, 124(7), pp. 6970-6988.

Geertsma, J., & De Klerk, F. (1969) ‘A rapid method of predicting width and extent of hydraulically induced fractures.’, *Journal of petroleum technology*, 21(12), pp. 1571-1581.

Gill, P. E., Murray, W., & Wright, M. H. (1981) *Practical Optimization*. California: Academic San Diego.

Goebel, T. H. W., et al. (2016) ‘Wastewater disposal and earthquake swarm activity at the southern end of the Central Valley, California.’, *Geophys. Res. Lett.*, 43, pp. 1092–1099.

Goebel, T. H., & Shirzaei, M. (2021) ‘More than 40 yr of potentially induced seismicity close to the San Andreas Fault in San Ardo, Central California.’, *Seismological Research Letters*, 92(1), pp. 187-198.

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, Mass.: Addison-Wesley.

Gramstad, O., & Nickel, M. (2018) ‘Automated top salt interpretation using a deep convolutional net.’, *80th EAGE Conference and Exhibition*, June 2018.

Grechka, V. I., & Heigl, W. M. (2017) *Microseismic monitoring*. Tulsa, OK: Society of Exploration Geophysicists.

Gupta, P., & Duarte, C. A. (2014) ‘Simulation of non-planar three-dimensional hydraulic fracture propagation.’, *International Journal for Numerical and Analytical Methods in Geomechanics*, 38(13), pp. 1397-1430.

Hamilton, D. H., & Meehan, R. L. (1971) ‘Ground rupture in the Baldwin Hills’, *Science*, 172, pp. 333–344.

Healy, J. H., et al. (1968) ‘The Denver earthquakes’, *Science*, 161(3848), pp. 1301–1310.

He, Z., & Duan, B. (2021) ‘Study on the microseismic clouds induced by hydraulic fracturing.’, *Geomechanics and Geoengineering*, 16(5), pp. 400-416.

Hincks, T., et al. (2018) ‘Oklahoma's induced seismicity strongly linked to wastewater injection depth’, *Science*, 359(6381), pp. 1251-1255.

Horton, S. (2012) ‘Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake’, *Seismol. Res. Lett.*, 83(2), pp. 250–260.

Howard, G. C., & Fast, C. R. (1957) ‘Optimum fluid characteristics for fracture extension.’, *Drilling and production practice*.

Huang, Y., Beroza, G. C., & Ellsworth, W. L. (2016) ‘Stress drop estimates of potentially induced earthquakes in the Guy-Greenbrier sequence’, *J. Geophys. Res.*, 121, pp. 6597–6607.

Jin, L., & Zoback, M. D. (2018) ‘Modeling Induced Seismicity: Co-Seismic Fully Dynamic Spontaneous Rupture Considering Fault Poroelastic Stress’, *52nd US Rock Mechanics/Geomechanics Symposium*, June 2018.

Juanes, R., et al. (2016) ‘Were the May 2012 Emilia-Romagna earthquakes induced? A coupled flow-geomechanics modeling assessment.’, *Geophys. Res. Lett.*, 43(13), pp. 6891–6897.

Keranen, K. M., et al. (2013) ‘Potentially induced earthquakes in Oklahoma, USA: links between wastewater injection and the 2011 Mw 5.7 earthquake sequence’, *Geology*, 41, pp. 699–702.

Keranen, K. M., et al. (2014) ‘Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection’, *Science*, 345(6195), pp. 448-451.

Keranen, K. M., & Weingarten, M. (2018) ‘Induced Seismicity’, *Annu. Rev. Earth Planet. Sci.*, 46, pp. 149-174.

Khristianovic, S. A., & Zheltov, Y. P. (1955) ‘Formation of vertical fractures by means of highly viscous liquid’, *World Petroleum Congress Proceedings*, pp. 579-586.

Kim, W. Y. (2013) ‘Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio.’, *J. Geophys. Res. Solid Earth*, 118(7), pp. 3506–3518.

King, V. M, et al. (2014) ‘Geological structure of the Paradox Valley Region,Colorado, and relationship to seismicity induced by deep well injection’, *J. Geophys. Res.*, 119, pp. 4955–4978.

King, V. M., Block, L. V., & Wood, C. K. (2016) ‘Pressure/flow modeling and induced seismicity resulting from two decades of high-pressure deep-well brine injection, Paradox Valley, Colorado.’, *Geophysics*, 81(5), B119-B134.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983) ‘Optimization by simulated annealing’, *Science*, 220(4598), pp. 671-680.

Kisslinger, C. (1976) ‘A review of theories of mechanisms of induced seismicity’, *Engineering Geology*, 10(2-4), pp. 85-98.

Kwiatek, G., et al. (2019) ‘Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in Finland.’, *Science Advances*, 5(5), eaav7224.

Lapusta, N., et al. (2000) ‘Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate-and state-dependent friction’, *J. Geophys. Res.: Solid Earth*, 105(10), pp. 23765-23789.

Lecampion, B., Bunger, A., & Zhang, X. (2018) ‘Numerical methods for hydraulic fracture propagation: a review of recent trends.’, *Journal of natural gas science and engineering*, 49, pp. 66-83.

Lecampion, B., & Desroches, J. (2015) ‘Simultaneous initiation and growth of multiple radial hydraulic fractures from a horizontal wellbore.’, *Journal of the Mechanics and Physics of Solids*, 82, pp. 235-258.

Liaw, A., & Wiener, M. (2002) ‘Classification and regression by randomForest.’, *R news*, 2(3), pp. 18-22.

Manga, M., Wang, C. Y., & Shirzaei, M. (2016) ‘Increased stream discharge after the 3 September 2016 Mw 5.8 Pawnee, Oklahoma earthquake’, *Geophys. Res. Lett.*, 43, pp. 11588–11594.

Metropolis, N., et al. (1953) ‘Equation of state calculations by fast computing machines.’, *The journal of chemical physics*, 21(6), pp. 1087-1092.

Miehe, C., Schaenzel, L. M., & Ulmer, H. (2015) ‘Phase field modeling of fracture in multi-physics problems. Part I. Balance of crack surface and failure criteria for brittle crack propagation in thermo-elastic solids.’, *Computer Methods in Applied Mechanics and Engineering*, 294, pp. 449-485.

Minster, J.-B. H., et al. (1995) ‘Application of evolutionary programming to earthquake hypocenter determination’, *Proceedings of 4th Annual Conference of Evol. Prog.*, pp. 3–17.

Montgomery, C. T., & Smith, M. B. (2010) ‘Hydraulic fracturing: History of an enduring technology.’, *Journal of Petroleum Technology*, 62(12), pp. 26-40.

Murray, K. E. (2013) ‘State-scale perspective on water use and production associated with oil and gas operations, Oklahoma, US’, *Environmental science & technology*, 47(9), pp. 4918-4925.

Murtagh, F. (1991) ‘Multilayer perceptrons for classification and regression.’, *Neurocomputing*, 2(5-6), pp. 183-197.

Needleman, A. (2014) ‘Some issues in cohesive surface modeling.’, *Procedia IUTAM*, 10, pp. 221-246.

Nelder, J. A., R. Mead (1965) ‘A simplex method for function minimization.’, *The computer journal*, 7(4), pp. 308-313.

Nguyen, B., Morell, C., & De Baets, B. (2016) ‘Large-scale distance metric learning for k-nearest neighbors regression.’, *Neurocomputing*, 214, pp. 805-814.

Norbeck, J. H., & Horne, R. N. (2016) ‘Evidence for a transient hydromechanical and frictional faulting response during the 2011 Mw 5.6 Prague, Oklahoma earthquake sequence’, *J. Geophys. Res.: Solid Earth*, 121(12), pp. 8688-8705.

Obert, L., & Duvall, W. I. (1957) *Micro-seismic method of determining the stability of underground openings*. Washington, DC: US Government Printing Office.

Oprsal, I., & Eisner, L. (2014) ‘Cross-correlation – an objective tool to indicate induced seismicity’, *Geophys. J. Int.*, 196(3), pp. 1536-1543.

Rudnicki, J. W. (1986) ‘Fluid mass sources and point forces in linear elastic diffusive solids’, *Mech. Mater*. 5, pp. 383–393.

Perkins, T. K., & Kern, L. R. (1961) ‘Widths of hydraulic fractures.’, *Journal of petroleum technology*, 13(09), pp. 937-949.

Peshcherenko, A., & Chuprakov, D. (2021) ‘An ultrafast simulator for 3D propagation of a hydraulic fracture with rectangular shape.’, *Engineering Fracture Mechanics*, 243, 107512.

Power, D. V., et al. (1976) ‘Detection of hydraulic fracture orientation and dimensions in cased wells.’, *Journal of Petroleum Technology*, 28(09), pp. 1116-1124.

Pudugramam, V. S., et al. (2021) ‘Analysis and Integration of the Hydraulic Fracturing Test Site-2 (HFTS-2) Comprehensive Dataset.’, *SPE/AAPG/SEG Unconventional Resources Technology Conference*, July 2021.

Raleigh, C. B., Healy, J. H., & Bredehoeft, J. D. (1976) ‘An experiment in earthquake control at Rangely, Colorado’, *Science,* 191, 1230–1237.

Robinson, B. M., et al. (1992) ‘Hydraulic fracturing research in East Texas: third GRI staged field experiment.’, *Journal of Petroleum Technology*, 44(01), pp. 78-87.

Rothman, D. H. (1985) ‘Nonlinear inversion statistical mechanics, and residual statics corrections’, *Geophysics*, 50, pp. 2784–2796.

Rothman, D. H. (1986) ‘Automatic estimation of large residual statics corrections’, *Geophysics*, 51, pp. 332–346.

Rouet-Leduc, B., et al. (2018) ‘Estimating fault friction from seismic signals in the laboratory.’, *Geophysical Research Letters*, 45(3), pp. 1321-1329.

Rudnicki, J. W. (1986) ‘Fluid mass sources and point forces in linear elastic diffusive solids’, *Mech. Mater*., 5, pp. 383–393.

Salimzadeh, S., & Khalili, N. (2016) ‘Fully coupled XFEM model for flow and deformation in fractured porous media with explicit fracture flow.’, *International Journal of Geomechanics*, 16(4), 04015091.

Sambridge, M. (1999) ‘Geophysical inversion with a Neighbourhood algorithm, I, Searching a parameter space’, *Geophys. J. Int.*, 138, pp. 479–494.

Sambridge, M., & Kennett, B. L. N. (1986) ‘A novel method of hypocentre location’, *Geophys. J. R. Astron. Soc.*, 87, pp. 679– 697.

Sambridge, M., & Mosegaard, K. (2002) ‘Monte Carlo methods in geophysical inverse problems.’, *Reviews of Geophysics*, 40(3), 3-1.

Samimi, S., & Pak, A. (2016) ‘A fully coupled element-free Galerkin model for hydro-mechanical analysis of advancement of fluid-driven fractures in porous media.’, *International Journal for Numerical and Analytical Methods in Geomechanics*, 40(16), pp. 2178-2206.

Saporetti, C. M., et al. (2018) ‘Machine learning approaches for petrographic classification of carbonate-siliciclastic rocks using well logs and textural information.’, *Journal of Applied Geophysics*, 155, pp. 217–225.

Segall, P., & Lu, S. (2015) ‘Injection-induced seismicity: poroelastic and earthquake nucleation effects.’, *J. Geophys. Res. Solid Earth*, 120 (7), pp. 5082–5103.

Shapiro, S., Rentsch, S., & Rothert, E. (2005) ‘Characterization of hydraulic properties of rocks using probability of fluid-induced microearthquakes’, *Geophysics*, 70(2), F27—F34.

Shirzaei, M., et al. (2016) ‘Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas’, *Science*, 353, pp. 1416–1419.

Sharma, D., & Kumar, N. (2017) ‘A review on machine learning algorithms, tasks and applications.’, *International Journal of Advanced Research in Computer Engineering & Technology*, 6(10), pp. 1548-1552.

Shirzaei, M., Manga, M., & Zhai, G. (2019) ‘Hydraulic Properties of injection formations constrained by surface deformation’, *Earth Planet. Sci. Lett.*, 515, pp.125-134.

Skoumal, R. J., Brudzinski, M. R., & Currie, B. S. (2015) ‘Distinguishing induced seismicity from natural seismicity in Ohio: demonstrating the utility of waveform template matching.’, *J. Geophys. Res.*, 120, pp. 6284–6296.

Skoumal, R. J., et al. (2020) ‘Induced seismicity in the Delaware basin, Texas.’, *Journal of Geophysical Research: Solid Earth*, 125(1), e2019JB018558.

Skoumal, R. J., & Trugman, D. T. (2021) ‘The Proliferation of Induced Seismicity in the Permian Basin, Texas.’, *Journal of Geophysical Research: Solid Earth*, e2021JB021921.

Sumy, D. F., et al. (2017) ‘Low stress drops observed for aftershocks of the 2011 Mw 5.7 Prague, Oklahoma, earthquake.’, *J. Geophys. Res.*, 122, pp. 3813–3834.

Titos, M., et al. (2018) ‘Detection and classification of continuous volcano-seismic signals with recurrent neural networks.’, *IEEE Transactions on Geoscience and Remote Sensing*, 57(4), pp. 1936-1948.

Urbancic, T. I., & Rutledge, J. (2000) ‘Using microseismicity to map Cotton Valley hydraulic fractures.’, *SEG Technical Program Expanded Abstracts*, pp. 1444-1448.

Urpi, L., et al. (2016) ‘Dynamic simulation of CO2-injection-induced fault rupture with slip-rate dependent friction coefficient’, *Geomechanics for Energy and Environment*, 7, 47-65.

Valera, M., et al. (2018) ‘Machine learning for graph-based representations of three-dimensional discrete fracture networks.’, *Computational Geosciences*, 22(3), pp. 695-710.

Vinther, R., & Mosegaard, K. (1996) ‘Seismic inversion through Tabu Search’, *Geophys. Prospect.*, 44(4), pp. 555–570.

Warpinski, N. R., et al. (1993) *Hydraulic fracture model comparison study: Complete results (No. SAND-93-7042; GRI-93/0109)*. Albuquerque, NM: Sandia National Labs; Chicago, IL: Gas Research Inst.

Warpinski, N. R., et al. (1999) ‘Microseismic monitoring of the B-Sand hydraulic-fracture experiment at the DOE/GRI multisite project.’, *SPE Journal*, 4(03), pp. 242-250.

Weingarten, M., et al. (2015) ‘High-rate injection is associated with the increase in US mid-continent seismicity’, *Science*, 348(6241), pp. 1336-1340.

Wright, R. E. (1995) ‘Logistic regression.’, in Grimm, L.G., & Yarnold, P. R. (Eds.), *Reading and understanding multivariate statistics*. American Psychological Association, pp. 217–244.

Wu, R., et al. (2008) ‘A comparison of numerical and experimental results of hydraulic fracture growth into a zone of lower confining stress.’, *The 42nd US rock mechanics symposium (USRMS)*, June 2008.

Xing, P., et al. (2017) ‘Laboratory measurement of tip and global behavior for zero-toughness hydraulic fractures with circular and blade-shaped (PKN) geometry.’, *Journal of the Mechanics and Physics of Solids*, 104, pp. 172-186.

Yang, Y., & Dunham, E. M. (2021) ‘Effect of Porosity and Permeability Evolution on Injection-Induced Aseismic Slip.’, *Journal of Geophysical Research: Solid Earth*, 126(7), e2020JB021258.

Yenier, E., Atkinson, G. M., & Sumy, D. F. (2017) ‘Ground motions for induced earthquakes in Oklahoma.’, *Bull. Seismol. Soc. Am.*, 107, pp. 198–215.

Zoback, M. D., & Kohli, A. H. (2019) *Unconventional reservoir geomechanics*. Cambridge University Press.

# 2. INTEGRATING POROELASTIC EFFECTS OF WASTEWATER INJECTION AND RUPTURE DYNAMICS TO UNDERSTAND INDUCED SEISMICITY*[1]

## 2.1. Abstract

Seismicity induced by the waste water injection has become an important issue in the recent years, including several regions in United States, such as Oklahoma and Texas. However, many techniques that are used to evaluate whether a particular earthquake is induced or natural are not based on the law of physics, and as a consequence these techniques do not provide much insight into nature of that type of seismicity. In this study we combine the analytic solution of coupled hydromechanical problem with numerical dynamic rupture simulation to study the influence of different properties of the rocks and fluids, as well as the geometry of the operations on the potential for earthquake triggering. We find three main driving factors are the distance between the injection well and the fault, permeability of the medium and shear modulus of the rocks. We also observe that keeping injecting the fluid at the same rate after first occurrence of seismic event, escalates the size of earthquakes with time.

---

[1] The following URTeC paper with minor modifications, (Szafranski, D., & Duan, B. (2018) 'Integrating poroelastic effects of wastewater injection and rupture dynamics to understand induced seismicity.', *Unconventional Resources Technology Conference, July 2018*, pp. 2557-2576.), is reprinted with permission from Unconventional Resources Technology Conference, whose permission is required for further use.

## 2.2. Introduction

Induced seismicity due to water injection is not a new phenomenon and has been recognized in the literature for several decades (e.g. Healy et al., 1968; Kisslinger, 1976). However, in recent years a significant increase in the level of seismicity in the regions that were considered aseismic was observed (e.g. Elsworth, 2013). This increase of seismicity correlates very well with the increase of the exploitation of hydrocarbons from unconventional oil and gas reservoirs. The most widely-known example is seismicity in Oklahoma where the number of felt earthquakes (M3.0 or above) in year 2014 overcame the number of earthquakes in California, the most seismically active region in the North America.

Most often large volume of waste water is obtained as a by-product during production from unconventional reservoirs (e.g. Murray, 2013). The waste water is then injected in Class II disposal wells. There is an ongoing public dispute about which wells cause induced seismicity and which do not. The problem lies in the reliability of the assessment. In their paper, Davis and Frohlich (1993) proposed a questionnaire for distinguishing between natural and induced earthquakes. Criteria are based on the past seismic activity, the spatial (both in horizontal and vertical direction) proximity of the earthquakes to the injection well, temporal correlation of seismicity and operations as well as potentially hazardous geological structures in the area. This type of approach was used in multiple studies in the following years as the way to classify earthquakes as induced by particular waste water disposal well (e.g. Frohlich et al., 2011; Weingarten et

al., 2015). This methodology could be further extended by calculation of cross-correlation between two time series – injected volume and number of seismic events, which was used by Oprsal and Eisner (2014) to evaluate time correlation in Rocky Mountain Arsenal (CO, USA), Blackpool (UK), Enola Field (OK, USA), Paradox Valley (CO, USA) and Basel (Switzerland), and by Horton (2012) for seismicity in Central Arkansas. These methods are easy to understand, but they do not provide physical explanation of the processes acting in the subsurface.

Some investigators did one step further trying to infer pore pressure perturbations due to injection and their possible relation to triggering the earthquakes. For example, Hornbach et al. (2016) analyzed the connection between cumulative injected volumes, formation pressures and occurring seismicity for the Ellenburger formation in north Texas. Other authors performed simulations of pore pressure diffusion in the medium assuming that it is the main driver leading to occurrence of induced seismicity in geological formations during waste water injection (e.g. Shapiro et al., 2005; Keranen et al., 2014).

Few studies also take hydromechanical coupling into consideration (e.g. Jha and Juanes, 2014; Segall and Lu, 2015; Fan et al., 2016; Urpi et al., 2016). Fan et al. (2016) showed that neglecting coupling effects leads to slower diffusion of fluids and higher accumulation of pore pressure. Segall and Lu (2015) pointed out that not including these effects may lead to overestimation or underestimation of seismicity rate (dependent on

geometry). They also raised the point that for further distances from injection point, influence of stress changes dominates over pore pressure changes. A majority of previous studies use a quasi-static approach, while induced earthquakes are fundamentally dynamic phenomena, in particular their final rupture extents and sizes. Recently there were few people who started studying dynamic phenomena caused by waste water or CO2 injection (e.g. Urpi et al., 2016; Ampuero et al., 2017). We integrate the theory of linear poroelasticity with dynamic rupture simulations to study the influence of different parameters on the induced earthquake initiation, propagation and arrest. Integrated simulations give us some insight how uncertainties of their estimation may affect the potential for perceivable induced seismicity. These fundamental principles are important to understand before constructing complex models used for seismic hazard estimation.

### 2.3. Theory of linear poroelasticity

Hydro-mechanical coupling should play an important role in the process of triggering earthquakes. There are two phenomena that define the medium behavior (Wang, 2000) – solid-to-fluid coupling and fluid-to-solid coupling. The first one is connected with changes in fluid pressure caused by the changes in applied stress, while the second one is linked to changes in stress (or rock volume) due to applied pore pressures. These changes modify Mohr circles in a different way. Changes of pore pressure shift the circle along the normal stress axis closer to the failure envelope. Perturbations of stress enlarge the circle without shifting it, which eventually also result in crossing the failure

envelope. Neglecting changes in stress caused by applied pore pressure would only be justifiable when the fluid in the rock is highly compressible (e.g. air). Since geological media at the depth of seismogenic zone are most of the time saturated with water, it is worth to include two-way coupled problem while analyzing induced seismicity.

To describe the two-way hydro-mechanical coupling we use theory of linear poroelasticity that (Biot, 1941), which is an extension of Terzaghi's theory of consolidation (1925) from one to three dimensions. In theory of poroelasticity, the behavior of the medium is fully described by seven equations – six equations to obtain either complete stress ($\sigma$) or strain ($\varepsilon$) tensor and one to obtain either pore pressure ($p$) or increment of fluid content ($\zeta$). For this study, we use so called pure stiffness formulation, in which stress tensor and pore pressure are calculated. For simple case of homogenous, isotropic, three-dimensional elastic medium it is possible to get analytic solution of the problem. In this study we implemented solution proposed by Rudnicki (1986) for a point source of fluid mass injected into the volume:

$$p(\boldsymbol{x},t) = \frac{1}{(4\pi)^{3/2}\rho_0 r^3}\frac{(\lambda_u-\lambda)(\lambda+2G)}{\alpha^2(\lambda_u+2G)}\int_0^t q(t')\xi^3 \exp\left(-\frac{\xi^2}{4}\right)dt', \tag{2.1}$$

$$\sigma_{ij}(\boldsymbol{x},t) = \frac{1}{2\pi\rho_0 r^3}\frac{\mu(\lambda_u-\lambda)}{\alpha(\lambda_u+2G)}\int_0^t q(t')\left[\delta_{ij}(g-\xi g') + \frac{x_i x_j}{r^2}(\xi g' - 3g)\right]dt', \tag{2.2}$$

where $\rho_0$ is a fluid density in the reference state, $x_i$ and $x_j$ are distances along $i^{th}$ and $j^{th}$ direction between source and given point in the medium, $r=||x||_2$, G is a shear modulus, $\lambda$ and $\lambda_u$ are Lame parameters for drained and undrained conditions respectively, $\alpha$ is a Biot's coefficient, $q(t)$ is the injection history, $\delta_{ij}$ is a Kronecker delta, $\xi$ is similarity factor defined as:

$$\xi(t') = \frac{r}{\sqrt{K(t-t')}}, \tag{2.3}$$

where $K$ is hydraulic diffusivity, which depends on permeability $k$ and dynamic fluid viscosity $\eta$:

$$K = \frac{k}{\eta}\frac{(\lambda_u-\lambda)(\lambda+2G)}{\alpha^2(\lambda_u+2G)}. \tag{2.4}$$

In equation (2) $g$ is a function of $\xi$:

$$g(\xi) = erf\left(\frac{\xi}{2}\right) - \frac{\xi}{\sqrt{\pi}}exp\left(-\frac{\xi^2}{4}\right), \tag{2.5}$$

where *erf* is an error function defined as:

$$erf\left(\frac{\xi}{2}\right) = \frac{2}{\sqrt{\pi}}\int_0^{\frac{\xi}{2}} exp(-z^2)dz. \tag{2.6}$$

Then, function $g'$ is just a derivative of $g$ in respect to $\xi$.

If the flux in the source is constant, i.e. $q(t)=q=const.$, then equations (1) and (2) can be exactly integrated and simplified to the form of

$$p(\boldsymbol{x},t) = \frac{q}{4\pi\rho_0 r}\frac{\eta}{k}erfc\left(\frac{\xi}{2}\right), \tag{2.7}$$

$$\sigma_{ij}(\boldsymbol{x},t) = -\frac{q(\lambda_u-\lambda)G}{4\pi\rho_0 rK\alpha(\lambda_u+2G)}\left\{\delta_{ij}\left[erfc\left(\frac{\xi}{2}\right) - \frac{2g}{\xi^2}\right] + \frac{x_i x_j}{r^2}\left[erfc\left(\frac{\xi}{2}\right) + \frac{6g}{\xi^2}\right]\right\}, \tag{2.8}$$

where $erfc(\xi/2) = 1\text{-}erf(\xi/2)$. In our simulations instead of using Lame's parameters $\lambda$ and $\lambda_u$ and Biot's coefficient $\alpha$ as input parameters, we prefer to use Poisson's ratios for drained ($\nu$) and undrained conditions ($\nu_u$), and Skempton's coefficient $B$, which are related to prior ones

$$\lambda = \frac{2G\nu}{1-2\nu}, \tag{2.9a}$$

$$\lambda_u = \frac{2G\nu_u}{1-2\nu_u}, \tag{2.9b}$$

$$\alpha = \frac{3(\nu_u-\nu)}{B(1+\nu_u)(1-2\nu)}. \tag{2.10}$$

Theory of linear poroelasticity assumes that we have Darcy's type of flow in the medium. In our case, this requirement is met, because waste water is usually injected into formation of relatively high permeability.

## 2.4. Dynamic rupture propagation

To simulate triggering of the induced earthquake, including its initiation, propagation and arrest phases, we have to solve elastodynamic equation of motion:

$$\nabla \cdot \boldsymbol{\sigma} + \rho \boldsymbol{f_b} = \rho \frac{\partial \boldsymbol{v}}{\partial t}, \tag{2.11}$$

where $\boldsymbol{f_b}$ is body force vector and $v$ is the velocity vector. An important factor in getting accurate spontaneous rupture solution is the fault boundary condition choice. Based on results of comparison done by Dalguer and Day (2006), we use the traction-at-split-node (TSN) scheme in the formulation proposed by Day et al. (2005). To describe frictional behavior, we implemented the slip-weakening law (e.g. Ida, 1972), in which the friction coefficient linearly drops from static ($\mu_s$) to dynamic ($\mu_d$) value within the critical slip distance ($D_0$),

$$\mu = \begin{cases} \mu_s - (\mu_s - \mu_d)\frac{|s|}{D_0}, |s| < D_0 \\ \mu_d, |s| \geq D_0 \end{cases}, \tag{2.12}$$

where $|s|$ is the slip distance. In order to check whether there is a failure on the fault plane, we resolve the tractions on the fault plane and use the Mohr-Coulomb failure criterion to check its stability:

$$\tau \leq c + \mu \sigma_n', \tag{2.13}$$

where $\tau$ is shear stress on the fault plane, $c$ is the cohesion of the fault plane and $\sigma_n'$ is effective normal stress which we define as:

$$\sigma_n' = \sigma_n - \alpha p. \tag{2.14}$$

## 2.5. Numerical implementation

For solution of equation (11) we use standard formulation of explicit finite element method (FEM) (e.g. Hughes, 2000), implemented in the in-house software called EQdyna (e.g. Duan and Oglesby, 2006; Duan and Day, 2008; Duan, 2010, 2012, 2016).

The software was successfully verified in SCEC (Southern California Earthquake Center) code verification exercise on multiple benchmark models (Harris et al., 2009; 2011; 2018). In our computations we use cubic, non-overlapping elements and PML boundary conditions (e.g. Berenger, 1994).

We solve the integrated poroelastic and dynamic rupture problem within two distinct time scales. Time steps for fluid diffusion are in order of either days or weeks and dynamic rupture problem has time steps in order of milliseconds. In each time step for fluid diffusion we calculate the perturbations of pore pressure (equation (1)) as well as normal and shear stresses on the fault plane (equation (2)), after which we switch to dynamic mode and check the failure criterion. If the criterion is met, we simulate rupture propagation that leads to partial stress relaxation of the fault and changes in pore pressure field Then, we move to the next time step and continue until reaching specified limit of time steps.

### 2.6. Model setup

To investigate the influence of different parameters on the potential for triggering an induced earthquake and the process of its initiation, propagation and arrest, we design a model, geometry of which is shown in Figure 2.1. In Table 2.1 we summarize the parameters used in our tests. First, simulations are run for 60-day period. In case there is no slip event in that time frame, we run the simulation again, but for 60-week period. In the simulations we used constant injection rate of 4320 m3/day. In our models we use

hydrostatic gradient of pore pressures and lithostatic gradient for vertical stress, assuming the density of water equal to 1000 kg/m3 and bulk density of the rocks to be 2320 kg/m3, which corresponds to quartz-rich water-saturated sandstone with ~20% porosity. Then, we assume that stress normal to the fault plane is equal to 80% of vertical stress and shear stress is defined in terms of percentage value of shear strength (the right-hand side of inequality (13)).

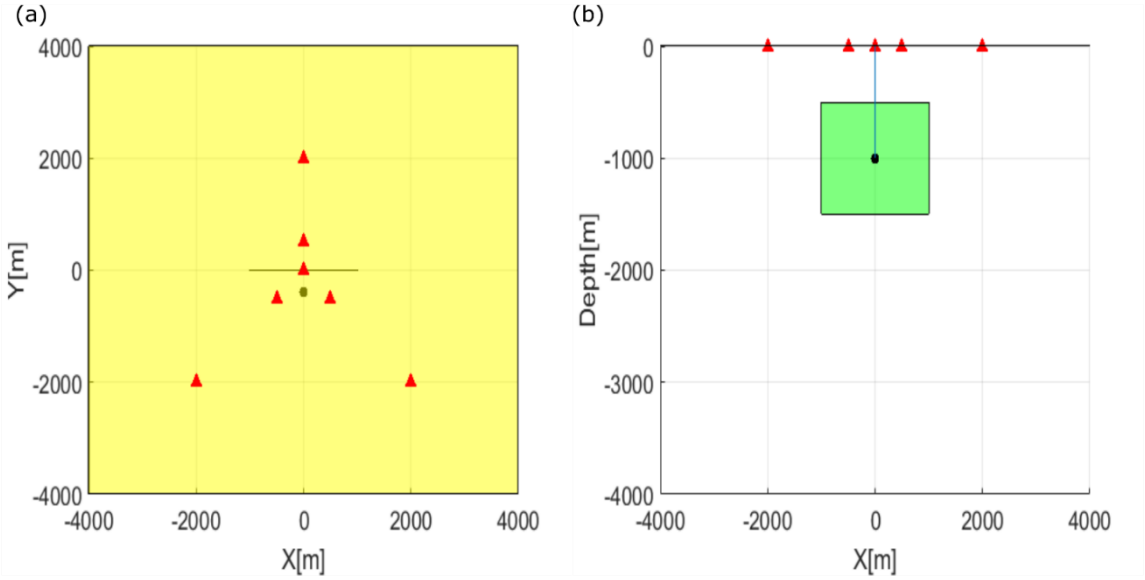


**Figure 2.1 Geometry of the model used in the simulations - (a) map view and (b) cross-section view. Red triangles represent seismic stations; blue line with black (olive) dot represent injection well with injection point; green rectangle/olive line represent fault**

**Table 2.1 Summary of parameters used in tests**

| Parameter | Reference value | Other values tested |
|---|---|---|
| Well-fault distace [m] | 400 | 200, 800, 1600 |
| Depth of injection [m] | 1000 | 500, 1500 |
| Well position along strike [m] | 0 | -1000, 1000 |
| Permeability [$m^2$] | $10^{-16}$ | $10^{-15}$, $10^{-14}$ |
| Shear modulus [GPa] | 15 | 1.5, 30 |
| Undrained - drained Poisson ratio | 0.05 (0.3 – 0.25) | 0.1, 0.15 |
| Cohesion [MPa] | 0 | 0.5, 2 |
| Static coefficient of friction | 0.6 | 0.5, 0.7 |
| Dynamic coefficient of friction | 0.4 | 0.2, 0.5 |
| Skempton's Coefficient | 0.7 | 0.5, 0.9 |
| Dynamic fluid viscosity [Pa·s] | 0.002 | 0.0002, 0.02 |
| Initial shear stress level [% of shear strains] | 95 | 90, 98, 99 |

### 2.7. Results

We run multiple sets of simulations changing different parameters and keeping the others as in the reference model. Then, we analyze the dataset and perform postprocessing during which for each single case we identify the potential of the system to produced triggered earthquake, as well as calculate several parameters that characterize the rupture behavior and size.

Firstly, we identify after how many days of injection we obtain the first slip on a group of neighboring elements with slip above 1 mm. Second parameter of interest is the size of initiation patch. By initiation patch we define the initial area that fails, before further rupture propagation. Next, we compute moment magnitude for a given event using definition of scalar seismic moment $M_0$ proposed by Aki (1966) and moment magnitude $M_w$ definition proposed by Hanks and Kanamori (1979):

$$M_0 = GD_{ave}A, \tag{2.15}$$

$$M_w = \frac{2}{3}(log_{10}M_0 - 9.1), \tag{2.16}$$

where $D_{ave}$ is average slip on the fault and $A$ is the ruptured area, both of which we have from simulations. We also use popular formula for calculation of the stress drop (e.g. Shearer, 2009):

$$\Delta\sigma = C_f G \frac{D_{ave}}{L_c}, \tag{2.17}$$

where $C_f$ is a constant dependent on fault geometry, that we assume to be equal to unity and $L_c$ is characteristic rupture dimension, which in our case is equal to square root of rupture area. As an output from simulations we also get maximum and average slips, and maximum and average slip rates. We analyze properties not only of the first slip but also of slips that occurred one week and two weeks after first slip occurrence. Influence of model parameter variations on occurrence of first slip and parameters of the rupturing process is presented in Figures 2.2 – 2.13.

First order observation that can be derived from the results of simulations is the fact that if we continue pumping the fluid into formation with the same injection rate, the magnitude of the earthquakes keeps increasing with time. It means that the stress drops

due to the rupture are actually smaller than the increase of stress and pore pressure perturbations and the fault is not able to relax.

For varying distance between fault plane and injection well (Figure 2.2) the time to first observed seismic event increases as distance squared, changing from 2 days for 200 m to 84 days for 800 m. For 1600 m there is no slip registered. Moreover, the fastest increase of the size and dynamics of the rupture is observed for shorter distances. Also, first slip size does not decrease with the distance. In fact, it is the smallest for the shortest distance, which can be accounted for the smallest area with perturbations of pore pressures and stresses that are high enough to cause a slip. However, for ruptures that we observe one and two weeks after first seismic event the decrease in size with distance is clearly visible from the graphs. To sum up, the potential for triggering larger magnitude earthquakes drops rapidly, when we locate the well further away from the fault plane.

Analyzing the influence of the depth of injection (Figure 2.3), we observe that for shallower injection first slip occurred earlier than for deeper well locations, which could be explained by generally lower stress levels in the shallower part of the medium. However, the largest sizes of the earthquakes occur if the well is located at the depth correlated with the middle section of the fault. Also, the increase of events size with time is the highest for the intermediate depth. At the same time, the deepest injection causes nascent of seismicity the latest and leads to smaller earthquakes with slow development of seismicity.

The positioning of the injection well along the strike of the fault also has an influence on the triggering potential and behavior of the rupture (Figure 2.4), because of the antisymmetric character of shear stress perturbations. If the well is placed in the way that the highest positive perturbations of shear stress are outside of the fault plane the latest nascent of the seismicity and the weakest earthquakes are observed. An exception is initial slip because it might be actually larger due to the choice of initial stress conditions on the fault plane. If these high positive perturbations occur on the fault plane, then location does not matter much anymore.

Permeability of the medium (Figure 2.5) is one of the most interesting parameters to study. Looking at its influence on the timing of the first slip, size of the rupture and the development of seismicity with time, we do not see any differences for relatively low permeabilities ($10^{-16}$ and $10^{-15}$ $m^2$). A further increase of permeability by an order of magnitude (i.e. $10^{-14}$ $m^2$), the behavior of the system completely changes and no seismic event can be observed even in the extended period of time. These results might be explained by the lower potential for high pore pressure build-up when fluid can easily flow.

Another very interesting parameter of the medium is shear modulus (Figure 2.6) because of its highly non-linear influence on the fault seismogenic potential. For relatively low value of shear modulus (1.5 GPa), the first slip occurs much later than for its higher

values (15 and 30 GPa). Also, lower value of shear modulus leads to significantly smaller ruptures and slower development of seismicity. These phenomena can be explained by the fact that weaker skeleton of the rock has lower potential for high stress build-up that might lead to larger earthquakes. For stronger rocks (i.e. higher shear modulus values) the difference in timing of the first slip is smaller (only few days), but the size of the rupture along with the pace of seismicity development continues to linearly increase, as the strength of the rock increases.

The other elastic moduli that we use to describe the medium is Poisson ratio, at both drained and undrained conditions (fluid can and cannot escape from the pore space respecitvely). The difference between these two also has some influence on the result of simulation (Figure 2.7). As we increase that difference, the shorter time is necessary to trigger the first event. Also, larger difference between its values for drained and undrained conditions corresponds to increased magnitudes (except for first slip) and pace of seismicity development. As the value of $\nu_u$ approaches the value of $\nu$, the solution comes closer to purely elastic medium.

Next parameter that we analyze is cohesion (Figure 2.8). Surprisingly in the range that we use in our test, the change in cohesion has virtually no influence on timing of first slip and almost no influence on the rupture size or development of the seismic activity. It is hard to explain because higher cohesion should correspond to stronger rock and higher shear stresses required to break its integrity, but it might be just connected with the fact

that fault zone has generally much lower cohesion than the intact rock surrounding the damaged zone (e.g. Sibson, 1977).

Out of two parameters connected with fault stability the static friction coefficient (Figure 2.9) has more significant influence on the triggering of the earthquakes. Since the higher values of friction coefficient correspond to the surface with more asperities, more time is required to build shear stress which is high enough to break these asperities and cause the earthquake. Moreover, higher static friction coefficient might correlate with smaller slips and corresponds to slower pace of seismicity development, because after nucleation of rupture, it loses its momentum more quickly, because more energy is required to propagate.

In dynamic formulation of problems, besides static friction coefficient, it is also necessary to consider dynamic friction coefficient (Figure 2.10) which affects only rupture propagation and arrest phase. Increasing the value of dynamic friction coefficient results in a slight decrease of the rupture size and slower development of the seismicity, which corresponds to less weakening of the damaged rock after rupture starts to propagate. What is interesting, is the increase of the nucleation patch size with the increase of the coefficient, which might be explained by the more uniform reaction of larger volume when weakening is smaller.

Next parameter, which in turn is connected with behavior of deformable porous medium is Skempton's coefficient (Figure 2.11). It describes the mutual influence of pore pressures and stresses, indicating pore pressure change due to rapid change of mean stress in the limiting condition of undrained rock. If the value approaches zero, then the model approximates purely elastic medium, but rocks and soils usually have the value of *B* between 0.5 and 1. With the increase of the coefficient value, the time till first slip elongates. Higher values also correspond to larger earthquakes (except for the initial slip) and higher pace of seismicity development. It may be explained by higher values corresponding to less compressible fluid (Wang, 2000) that helps the rock to support part of the stresses but at the same time leads to higher stress build-up and larger ruptures as an effect of that.

It is also worth to look at the influence of the injected fluid properties on the earthquake triggering potential (Figure 2.12). Based on our observations it has no influence on the time needed for causing the first slip event. It is also relatively unimportant when it comes to size of the rupture and the pace of seismicity development, causing only small increase of both with quite significant increase of fluid viscosity.

Last but not least, we analyze the influence of initial stress conditions, more precisely of the initial shear stress level in terms of its percentage value of the fault shear strength (Figure 2.13). As might be expected, the higher the initial shear stress, which means that the fault is closer to failure, the sooner we observe the first slip event. Beside the first

slip, the sizes of earthquakes as well as the pace of seismicity development increase with the increase of initial shear stress level. It happens because generally the entire fault plane is more critically stressed and as rupture starts to propagate, the dynamic perturbations of stress does not need to be very high to break the integrity of large part of the fault plane

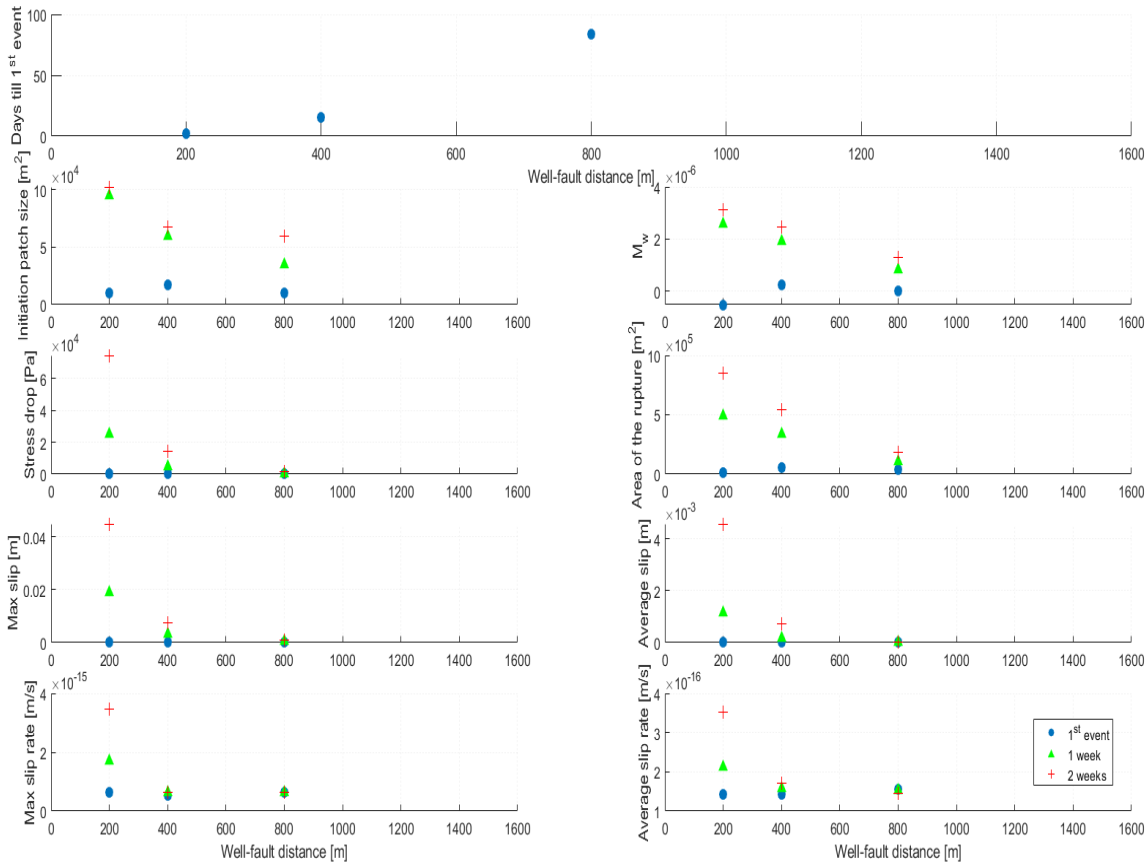


**Figure 2.2 Variation of earthquake parameters with changes of well-to-fault horizontal distance. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

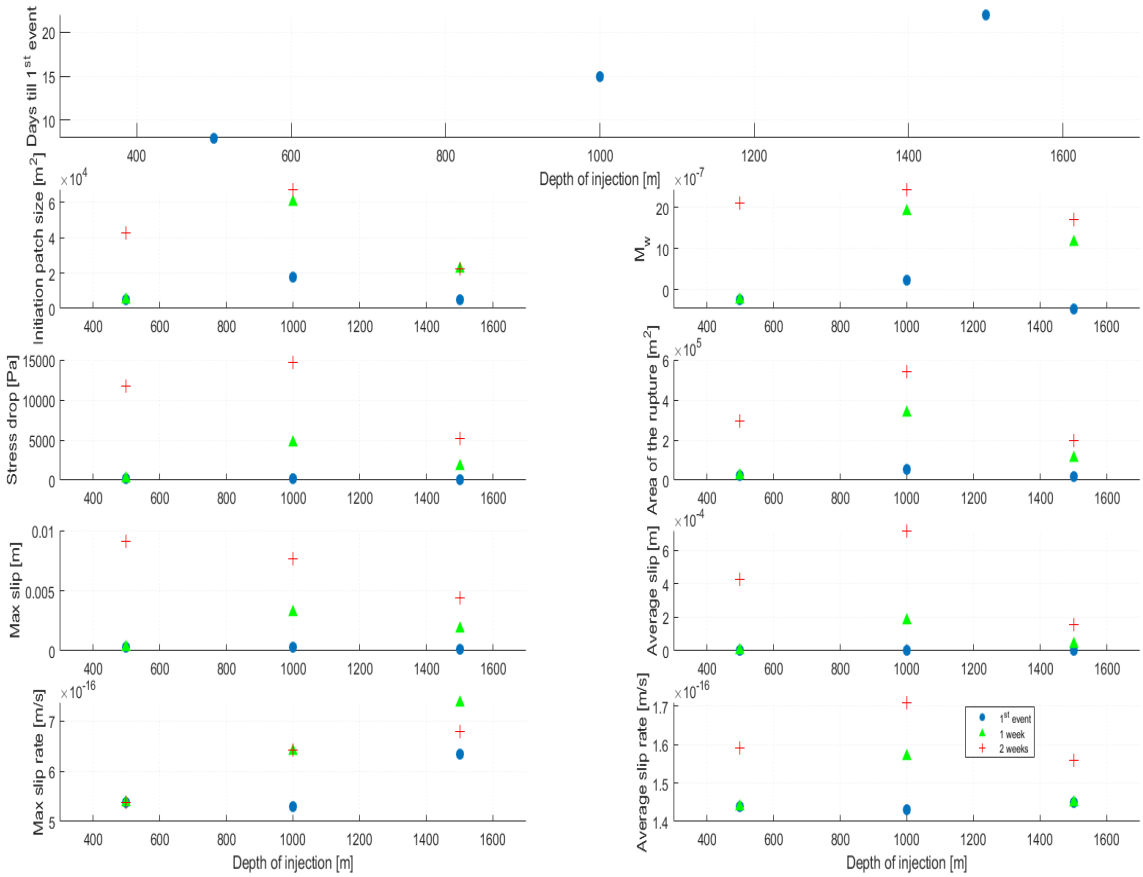


**Figure 2.3 Variation of earthquake parameters with changes of injection depth. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

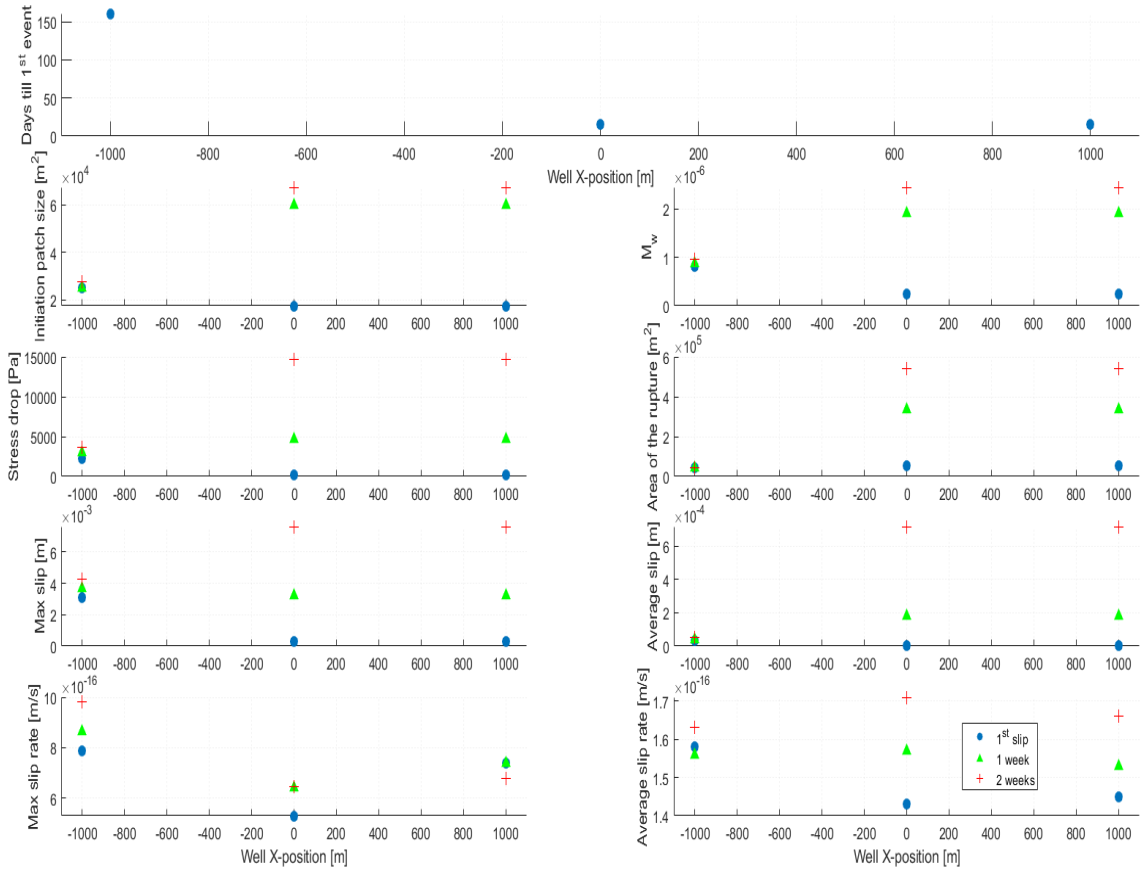


**Figure 2.4 Variation of earthquake parameters with changes of injection well position along fault strike. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

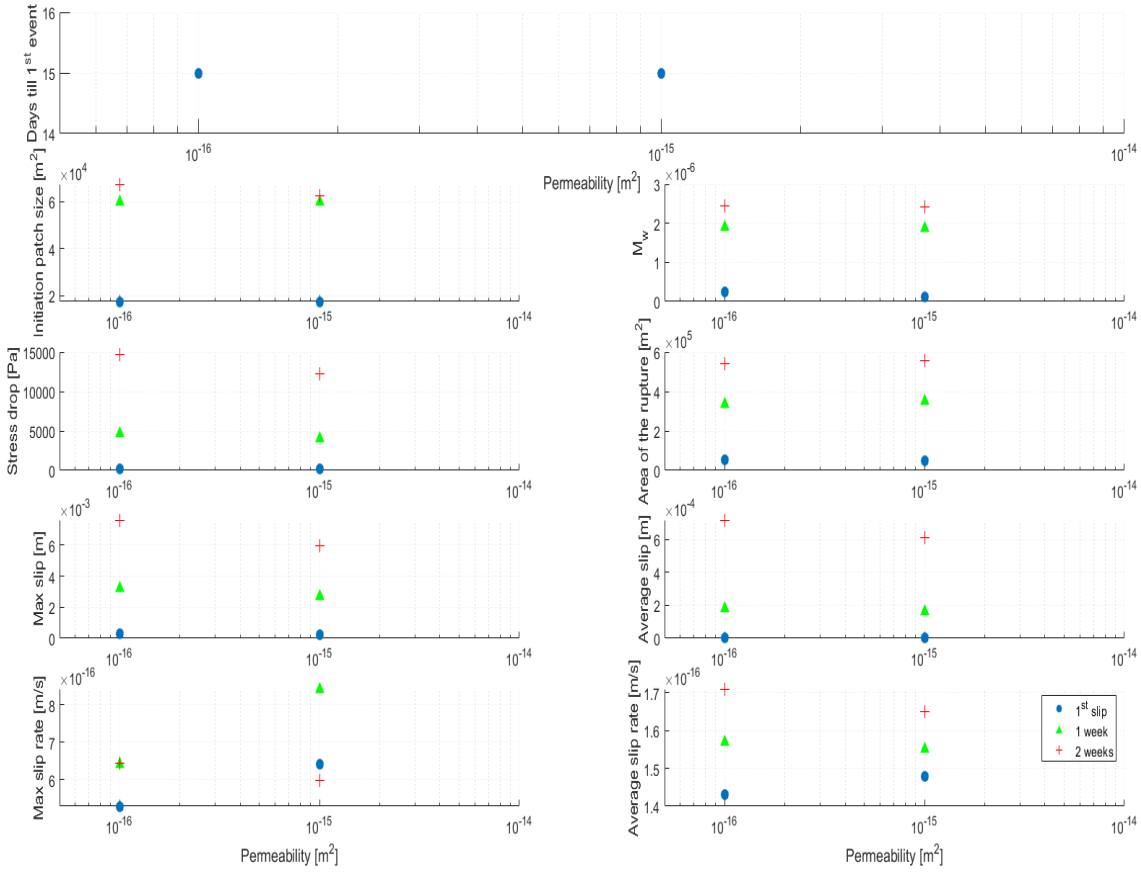


**Figure 2.5 Variation of earthquake parameters with changes of permeability. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

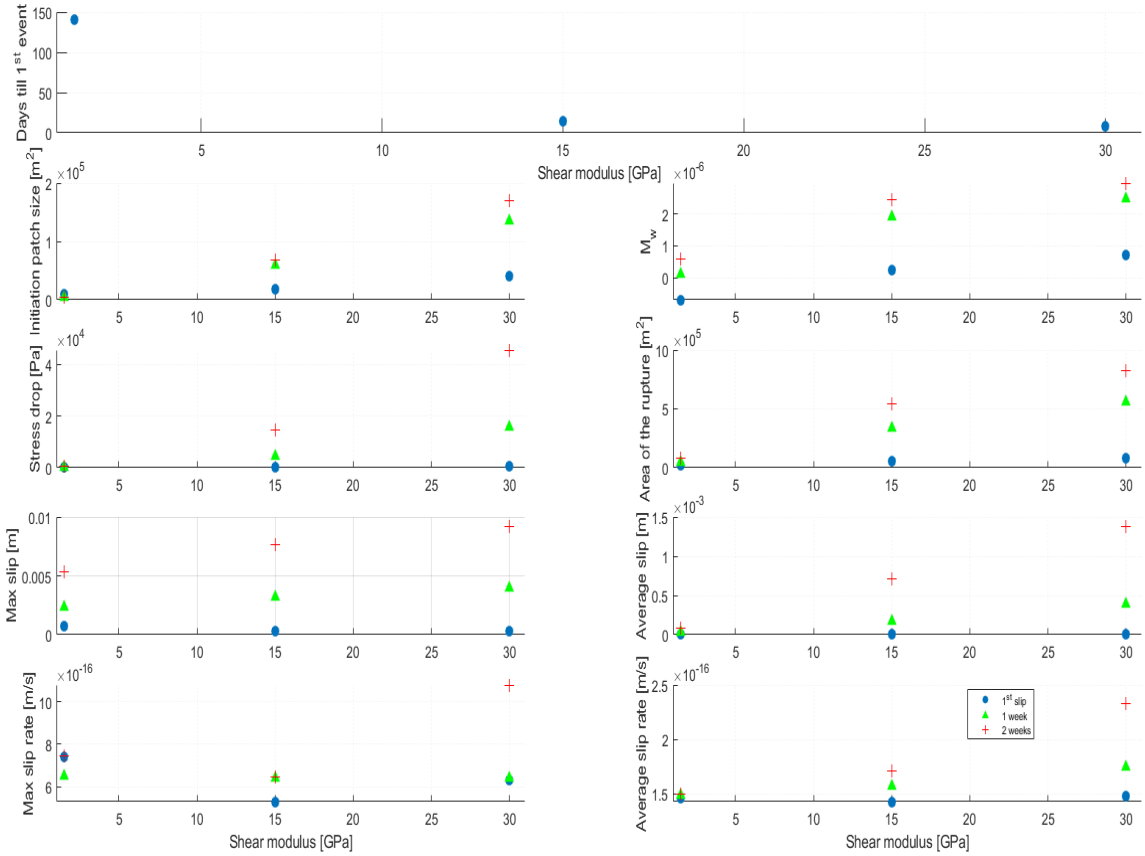


**Figure 2.6 Variation of earthquake parameters with changes of shear modulus value. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

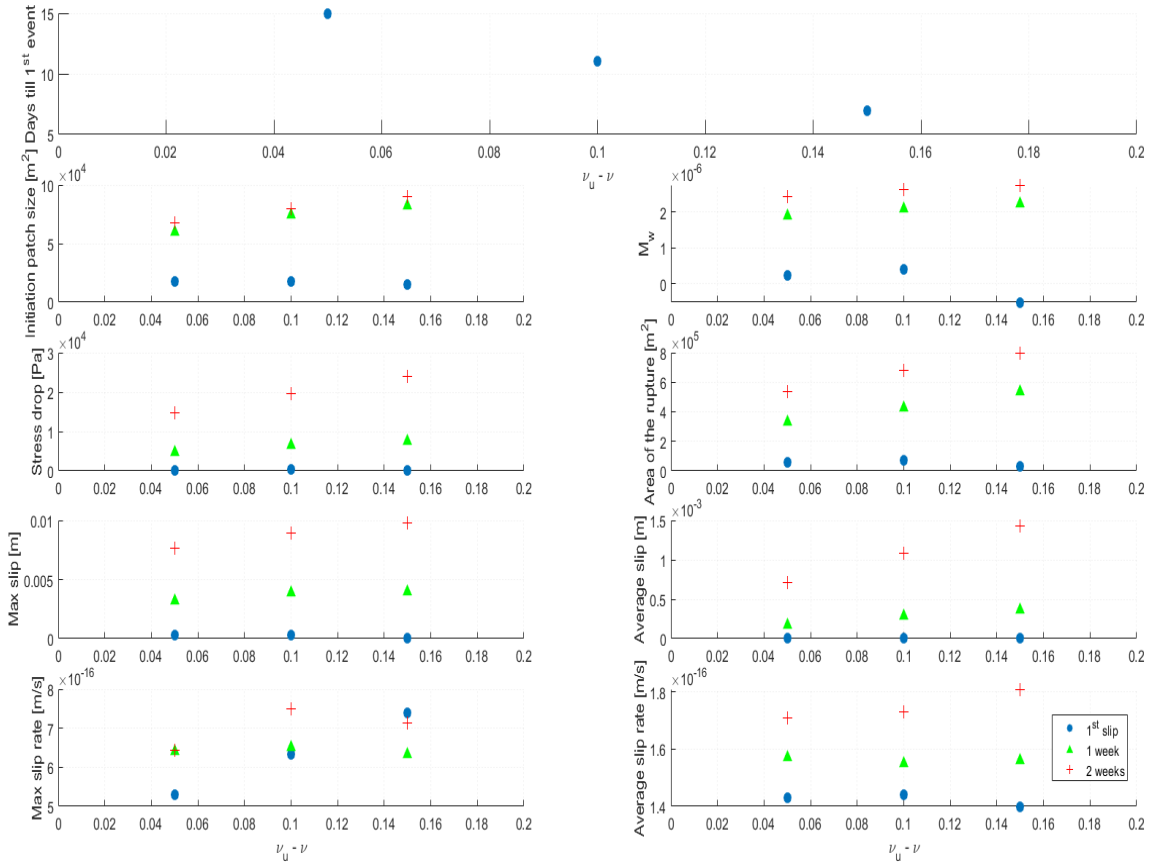


**Figure 2.7 Variation of earthquake parameters with changes of difference between Poisson ratio for undrained and drained rock. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

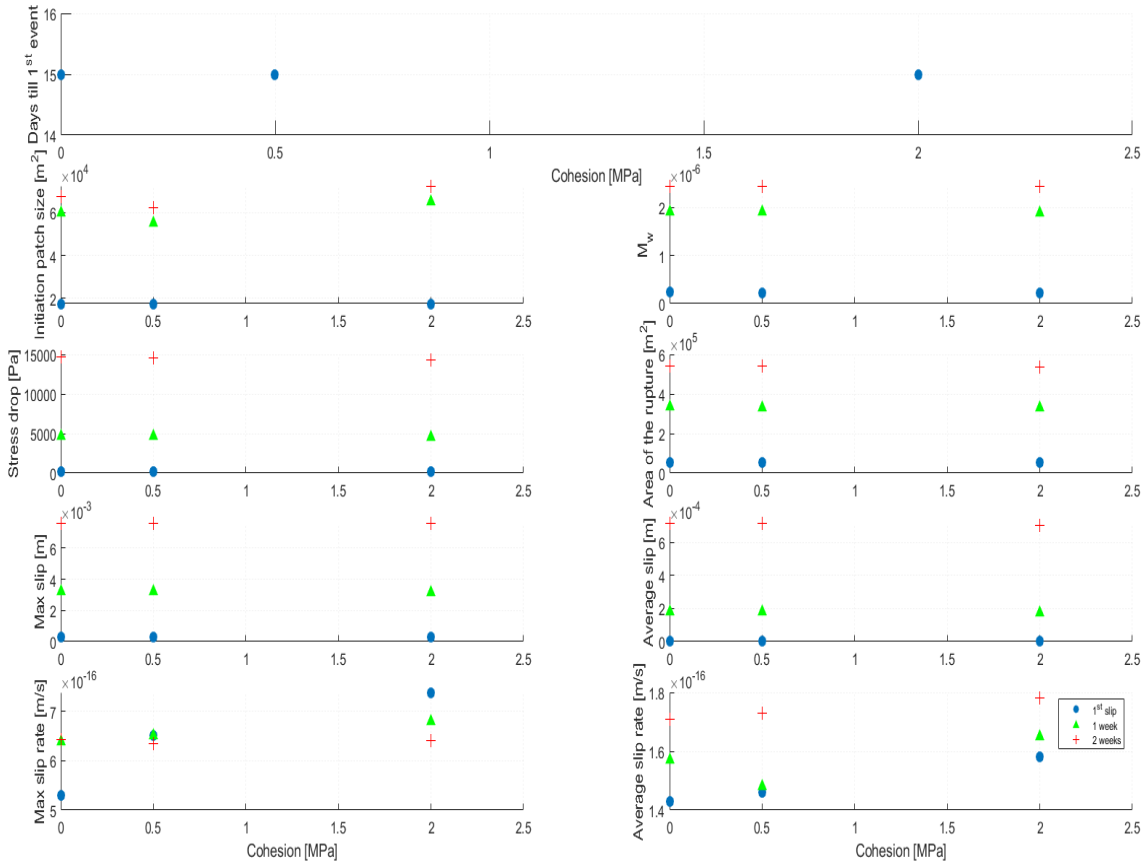


**Figure 2.8 Variation of earthquake parameters with changes of fault plane cohesion. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

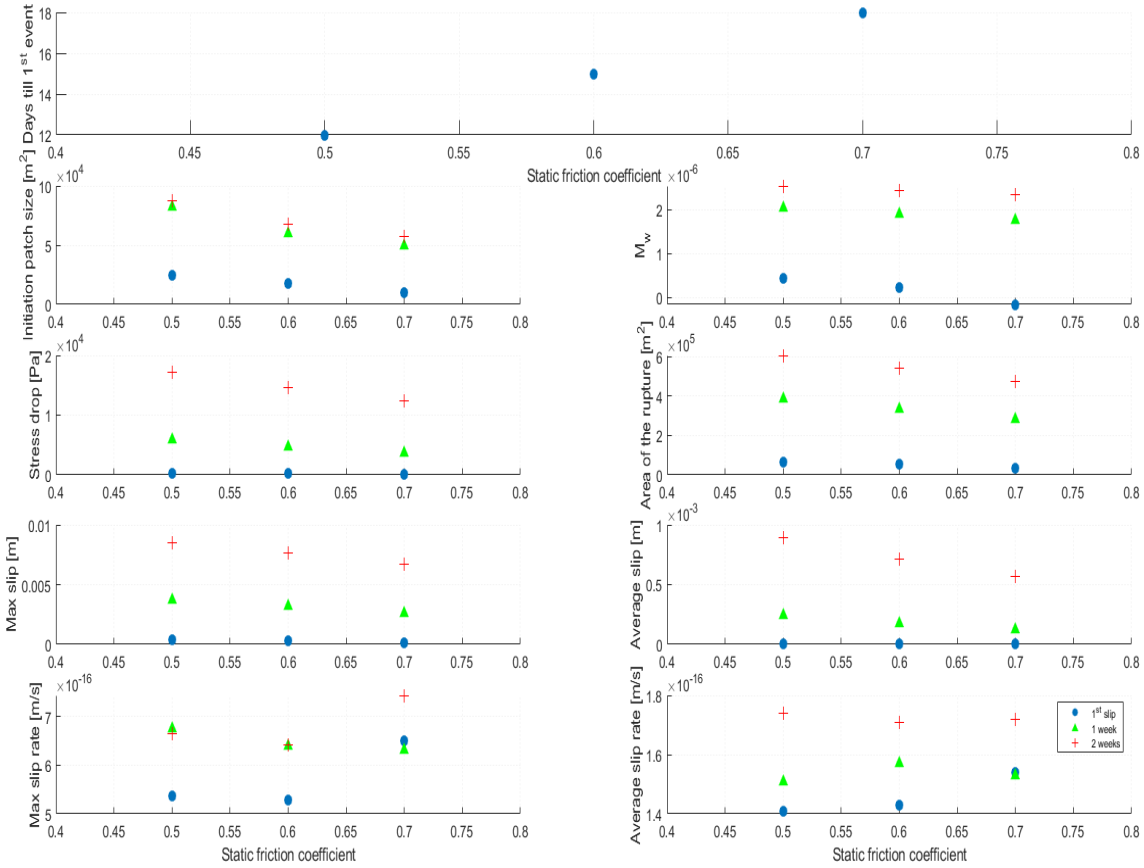


**Figure 2.9 Variation of earthquake parameters with changes of fault plane static coefficient of friction. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

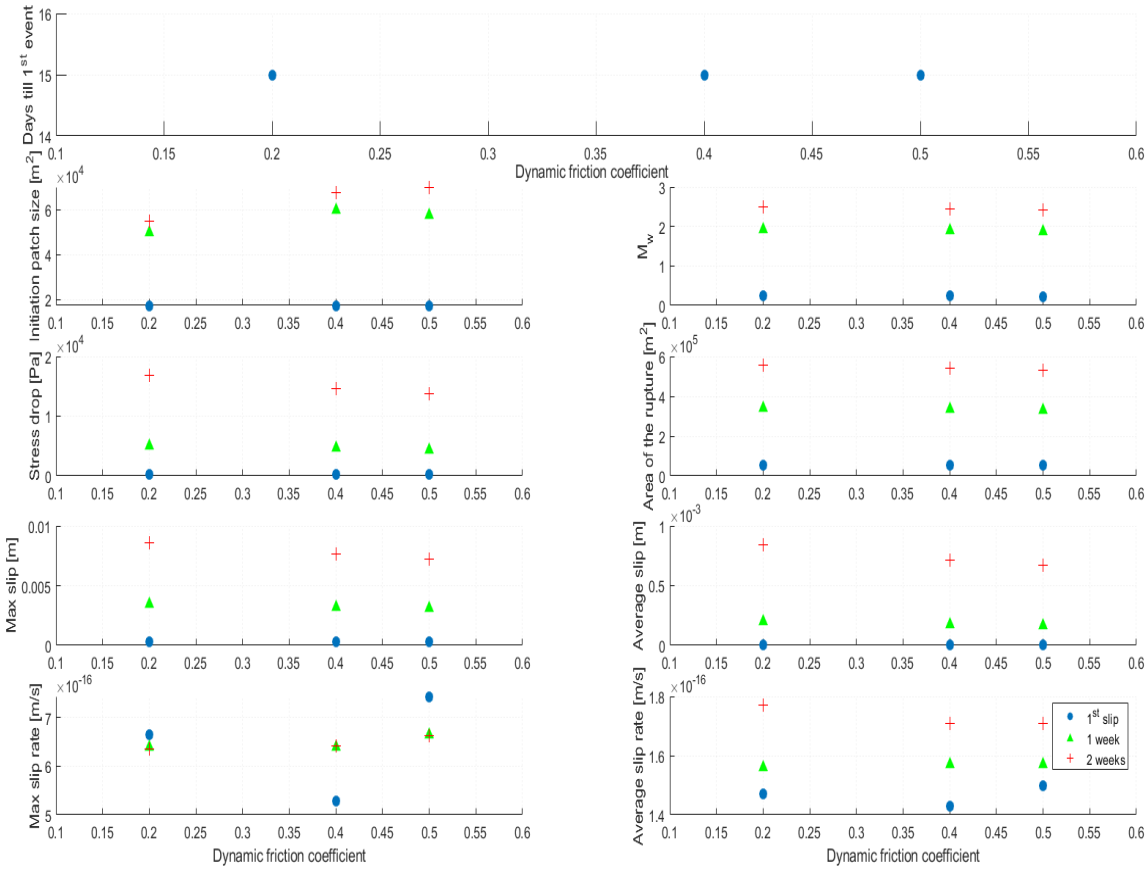


**Figure 2.10 Variation of earthquake parameters with changes of fault plane dynamic coefficient of friction. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

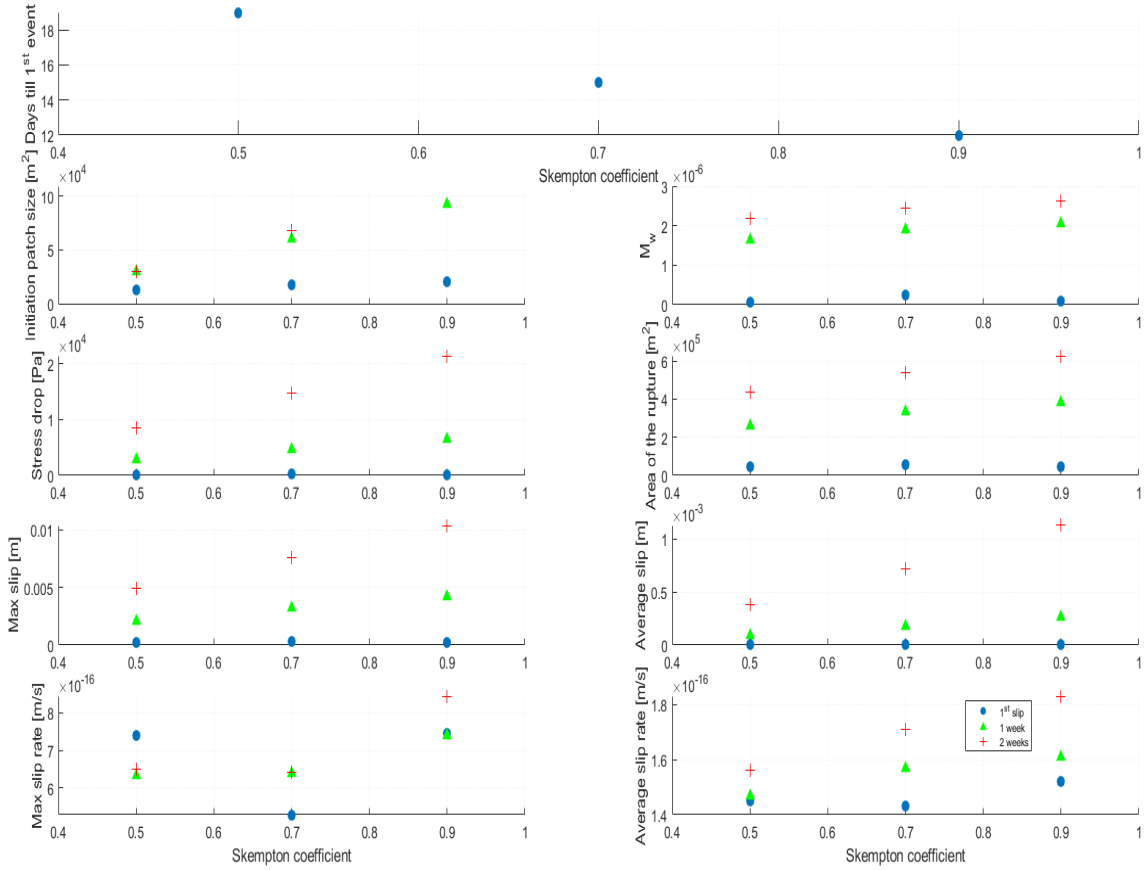


**Figure 2.11 Variation of earthquake parameters with changes of Skempton's coefficient. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

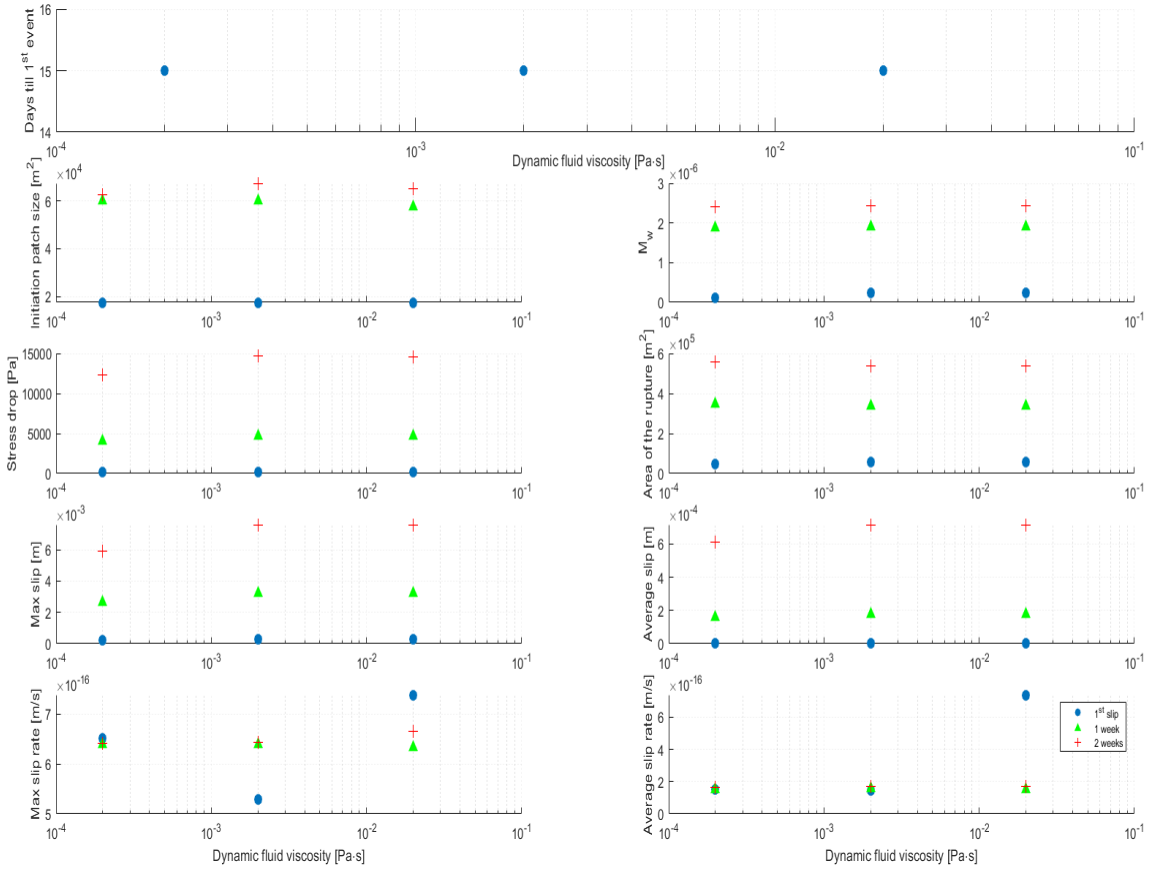


**Figure 2.12 Variation of earthquake parameters with changes of dynamic fluid viscosity. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

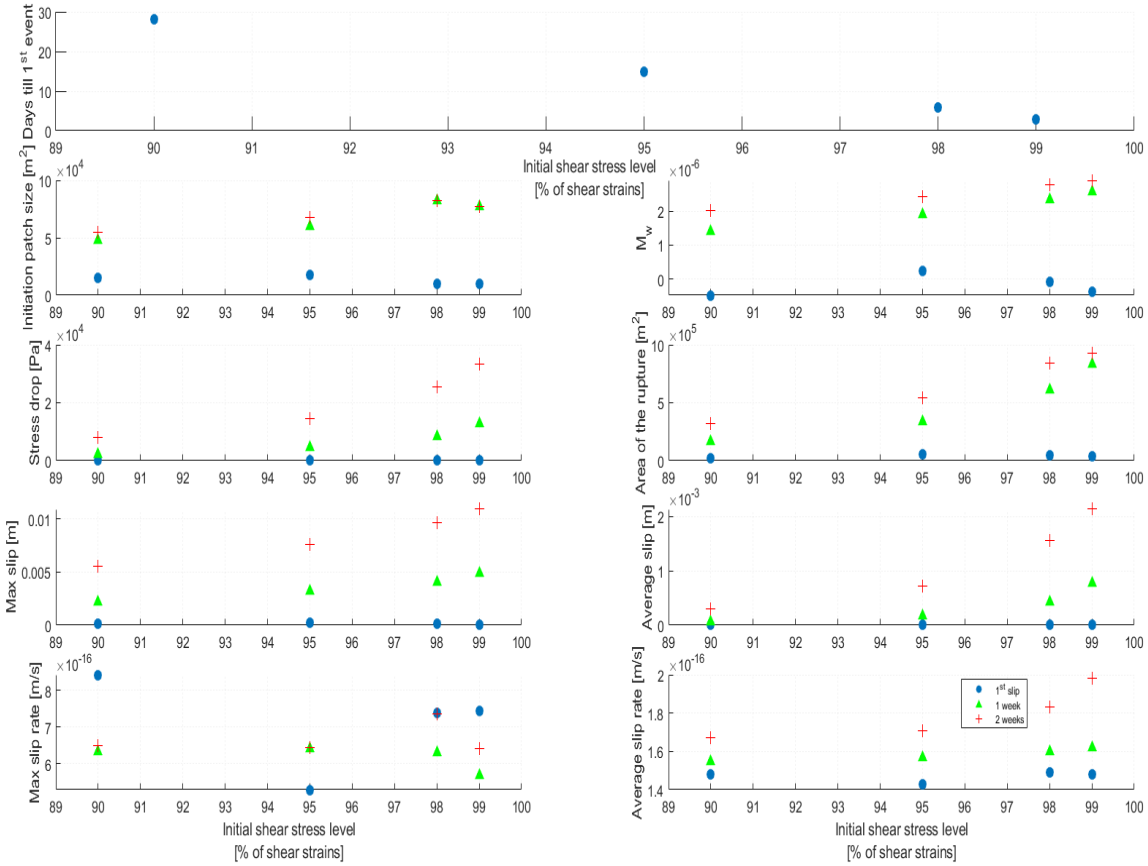


**Figure 2.13 Variation of earthquake parameters with changes of initial shear stress level expressed as a percentage of shear strains. Each panel represents different parameter (from top to bottom and left to right: No. days till first slip, Initiation patch size, Moment magnitude, Stress drop, Total area of the rupture, Maximum slip, Average slip, Maximum slip rate, Average slip rate). Blue circles correspond to first event, green triangles to event occurring one week after first event and red crosses correspond to event occurring two weeks after first event.**

## 2.8. Discussion

Results of this study have significant implications for studies of the induced seismicity. First and foremost, we know which parameters are more important for modeling and analysis and which could be neglected. That tells us about for which parameters we should get good estimates of to decrease the uncertainty of the solution. While the

geometry of operations (relative position of the fault plane and injection well) can be almost fully controlled and is very well-constrained, the parameters such as permeability or elastic modulus, which have significant influence, are much harder to precisely estimate. Another very important factor, which has a big impact on the timing and size of the failure, yet is usually not well-constrained is the level of fault proximity to failure. While it is common to determine the orientation of principal stresses and the ratio representing the relationship between all of them, the absolute values of effective normal stress and shear stress are usually unknown. Therefore, one needs to be cautious and besides most probable solution, specify the range of other possible scenarios.

Secondly, continuing waste water injection at the same rate, may cause size escalation of earthquakes in the area. Thus, studies of different injection schemes should be continued to determine the most effective practices for different cases to mitigate the seismic hazard and still keep the operation effective. Establishing efficient practices can lead to development of appropriate regulations and more specific rules connected with waste water disposal. Observed patterns can also enhance the monitoring of the operation sites and help set up proper criteria for different levels of alerts.

Besides very simplified model that is used in this study, simulations provide interesting insight into nature of induced seismicity and lay a strong understanding of basic principles before performing tests on more complex models, for which we might have problems to understand the results. The use of isotropic, homogenous poroelastic

medium instead of more realistic setup, makes it difficult to apply it directly to particular field cases, but allows for insight into basic principles connected with induced earthquakes triggering.

## 2.9. Conclusions

We performed multiple simulations of poroelastic effects of fluid injection into subsurface and their influence on potential for induced earthquakes and properties of the rupturing process. We learn which parameters have high and low importance for the obtained solution. Among the first group we include the distance of the injection well from the fault plane, permeability of the formation and shear moduli of the rocks. These three parameters have a great and non-linear impact on the potential for triggering of the earthquakes and then the size of these seismic events if they occur. If all major faults in the area are mapped we should be able to very precisely know the distance between the fault and injection well. Concurrently, information about permeability or shear moduli of the rocks in situ is usually quite vague, which significantly increases the uncertainty of prediction.

There are also parameters of very low importance, including the cohesion and viscosity of injected fluid. Low significance of their influence on the solution is actually very important, since we usually know very little about the cohesion of the fault zone and possibility of having only very rough estimate of its value serve our purpose. Other parameters of the model, for which we perform simulations, influence the solution in the

way that should not be neglected, but lack of precise knowledge about their values introduces acceptable level of uncertainty into evaluations of seismic hazard.

## 2.10. References

Aki, K. (1966) ‘4. Generation and propagation of G waves from the Niigata earthquake of June 14, 1964. Part 2. Estimation of earthquake moment, released energy and stress-strain drop from G wave spectrum’, *Bulletin of the Earthquake Research Institute*, 44, pp. 73–88.

Ampuero, J. P., et al. (2017) ‘Insights on Induced Seismicity Derived From Dynamic Models of Earthquake Nucleation, Propagation and Arrest.’ *AGU Fall Meeting Abstracts*, December 2017.

Berenger, J. P. (1994) ‘A perfectly matched layer for the absorption of electromagnetic waves’, *J. Comput. Phys.*, 114(2), pp. 185-200.

Biot, M. A. (1941) ‘General theory of three-dimensional consolidation’, *J. Appl. Phys.*, 12, pp. 155–164.

Dalguer, L. A., & Day, S. M. (2006) ‘Comparison of fault representation methods in finite difference simulations of dynamic rupture’, *Bull. Seismol. Soc. Am.*, 96, pp. 1764–1778.

Davis, S. D., & Frohlich, C. (1993) 'Did (or will) fluid injection cause earthquakes? - criteria for a rational assessment', *Seismol. Res. Lett.*, 64(3-4), pp. 207-224.

Day, S. M., et al. (2005) 'Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture', *J. Geophys. Res.*, 110, B12307.

Duan, B. (2016) 'Spontaneous rupture on natural fractures and seismic radiation during hydraulic fracturing treatments', *Geophys. Res. Lett*., 43(14), pp. 7451–7458.

Duan, B. (2012) 'Dynamic rupture of the 2011 Mw 9.0 Tohoku-Oki earthquake: Roles of a possible subducting seamount', *J. Geophys. Res.*, 117, B05311.

Duan, B. (2010) 'Role of initial stress rotations in rupture dynamics and ground motion: A case study with implications for the Wenchuan earthquake', *J. Geophys. Res*., 115, B05301.

Duan, B., & Day, S. M. (2008) 'Inelastic strain distribution and seismic radiation from rupture of a fault kink', *J. Geophys. Res.*, 113, B12311.

Duan, B., & Oglesby, D. D. (2006) 'Heterogeneous fault stresses from previous earthquakes and the effect on dynamics of parallel strike-slip faults', *J. Geophys. Res.*, 111, B05309.

Ellsworth, W. L. (2013) 'Injection-induced earthquakes', *Science*, 341, 142–149.

Fan, Z., Eichhubl, P., & Gale, J. F. W. (2016) 'Geomechanical analysis of fluid injection and seismic fault slip for the Mw4.8 Timpson, Texas, earthquake sequence', *J. Geophys. Res. Solid Earth*, 121, pp. 2798–2812.

Frohlich, C., et al. (2011) 'The Dallas-Fort Worth earthquake sequence: October 2008 through May 2009', *Bull. Seismol. Soc. Am.*, 101, pp. 327–340.

Hanks, T., & Kanamori, H. (1979) 'A moment magnitude scale', *J. Geophys. Res*., 84(B5), pp. 2348-2350.

Harris, R. A., et al. (2009) 'The SCEC/USGS dynamic earthquake rupture code verification exercise', *Seismol. Res. Lett.*, 80(1), pp. 119–126.

Harris, R. A., et al. (2011) 'Verifying a computational method for predicting extreme ground motion', *Seismol. Res. Lett.*, 82(5), pp. 638-644.

Harris, R. A., et al. (2018) ‘A Suite of Exercises for Verifying Dynamic Earthquake Rupture Codes’, *Seismological Research Letters*, 89(3), pp. 1146-1162.

Healy, J., et al. (1968) ‘The Denver earthquakes’, *Science*, 161, pp. 1301–1310.

Hornbach M., et al. (2016) ‘Ellenburger wastewater injection and seismicity in North Texas’, *Physics of the Earth and Planetary Interiors*, 261, pp. 54-68.

Horton, S. (2012) ‘Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake’, *Seismol. Res. Lett.*, 83, pp. 250–260.

Hughes, T. J. R. (2000) *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover, Mineola, N. Y.

Ida, Y. (1972) ‘Cohesive force across the tip of a longitudinal-shear crack and Griffith’s specific surface energy’, *J. Geophys. Res.*, 77, pp. 3796-3805.

Jha, B., & Juanes, R. (2014) ‘Coupled multiphase flow and poromechanics: A computational model of pore pressure effects on fault slip and earthquake triggering’, *Water Resour. Res.*, 50, pp. 3776–3808.

Keranen, K. M., et al. (2014) ‘Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection’, *Science*, 345(6195), pp. 448-451.

Kisslinger, C. (1976) ‘A review of theories of mechanisms of induced seismicity’, *Engineering Geology*, 10(2-4), pp. 85-98.

Murray, K. E. (2013) ‘State-scale perspective on water use and production associated with oil and gas operations, Oklahoma, US’, *Environmental science & technology*, 47(9), pp. 4918-4925.

Oprsal, I., & Eisner, L. (2014) ‘Cross-correlation – an objective tool to indicate induced seismicity’, *Geophys. J. Int.*, 196(3), pp. 1536-1543.

Rudnicki, J. W. (1986) ‘Fluid mass sources and point forces in linear elastic diffusive solids’, *Mech. Mater.*, 5, pp. 383–393.

Segall, P., & Lu, S. (2015) ‘Injection-induced seismicity: Poroelastic and earthquake nucleation effects’, *J. Geophys. Res. Solid Earth*, 120, pp. 5082–5103.

Shapiro, S., Rentsch, S., & Rothert, E. (2005) ‘Characterization of hydraulic properties of rocks using probability of fluid-induced microearthquakes’, *Geophysics*, 70(2), F27—F34.

Shearer, P. M. (2009) *Introduction to seismology*. Cambridge University Press.

Sibson, R. H. (1977) 'Fault rocks and fault mechanisms', *Journal of the Geological Society*, 133(3), pp. 191-213.

Terzaghi, K. (1925) *Erdbaumechanik auf bodenphysikalischer Grundlage*. F. Deuticke.

Urpi, L., et al. (2016) 'Dynamic simulation of CO2-injection-induced fault rupture with slip-rate dependent friction coefficient.', *Geomechanics for Energy and Environment*, 7, pp. 47-65.

Wang, H. (2000) *Theory of Linear Poroelasticity With Applications to Geomechanics and Hydrogeology*. Princeton, N. J: Princeton Univ. Press.

Weingarten, M., et al. (2015) 'High-rate injection is associated with the increase in US mid-continent seismicity', *Science*, 348(6241), pp. 1336-1340.

# 3. EXPLORING PHYSICAL LINKS BETWEEN FLUID INJECTION AND NEARBY EARTHQUAKES – THE 2012 $M_W$4.8 TIMPSON (TX) CASE STUDY*[2]

## 3.1. Abstract

In this work, we integrate a fluid flow model of 3D deformable porous media with a dynamic rupture model of earthquakes in 3D heterogeneous geologic medium. The method allows us to go beyond fault failure potential analyses and to examine how big an earthquake can be if part of a fault reaches failure due to fluid injection. We apply the method to the May 17, 2012, $M_w$4.8 Timpson (TX) earthquake as a case study. The simulated perturbations of pore pressure and stress from wastewater injection at the time of the mainshock are high enough (several MPa) to trigger an earthquake. Dynamic rupture modeling could reproduce the major observations from the $M_w$4.8 event, including its size, focal mechanism and aftershock sequence, and thus building a more convincing physical link between fluid injection and the $M_w$4.8 earthquake. Furthermore, parameter space studies of dynamic rupture modeling allow us to place some constraints on fault frictional properties and background stresses. For the Timpson

[2] The following paper with minor modifications, (Szafranski D., and Duan, B. (2020). Exploring physical links between fluid injection and nearby earthquakes: The 2012 M 4.8 Timpson, Texas, case study, *Bulletin of the Seismological Society of America*, *110 (5)*, 2350–2365. https://doi.org/10.1785/0120200090), is reprinted with permission from Seismological Society of America, whose permission might be required for further use.

case, we find that a dynamic friction coefficient of ~0.3, a value of ~0.1m for the critical slip distance in the slip-weakening friction law, and uniform effective normal stress are associated with the Timpson earthquake fault. By reproducing main features of the aftershock sequence of the mainshock, we also demonstrate that the method has potential to become a predictive tool for fluid injection design in the future.

### 3.2. Introduction

Induced seismicity related to fluid injection and extraction has been studied since the mid-20th century (e.g. Healy et al., 1968). Over the past 15 years, the number of felt earthquakes in the central and eastern United States at first has significantly increased (e.g. Ellsworth, 2013) with the peak of number of events in 2015 and then dropped by ~70% (e.g. Keranen and Weingarten, 2018). Even though there were strong natural earthquakes in the history of that region, including devastating earthquakes in New Madrid Fault Zone in 1811-1812 (e.g. Johnston, 1996), the average level of seismic activity is much lower than in regions located nearby the active plate boundaries, such as California. The increase in seismicity rate coincides with the burgeoning exploration and exploitation of unconventional, on-shore hydrocarbon resources, and there is a strong evidence that registered earthquakes are induced by the ongoing operations.

Most of the induced seismicity is associated with the water injection into salt water disposal (SWD) wells in the United States, including examples in Oklahoma (e.g. Kim, 2013; Keranen et al., 2014; Manga et al., 2016), east and central Texas (e.g. Frohlich et

al., 2011; Frohlich et al., 2014), the Guy-Greenbrier area in Arkansas (e.g. Horton, 2012) or Paradox Valley in Colorado (e.g. Ake et al., 2005). Over the years, the methods of establishing the potential link between wastewater injection and seismicity have evolved significantly. The simplest method for induced seismicity discrimination uses simple questionnaires (e.g. Davis and Frohlich, 2013) in which researchers semi-qualitatively take into account spatial proximity of earthquakes to the injection wells, temporal correlation between the operations and seismicity, as well as past background seismic activity.

A more quantitative way to find a link between seismic activity and injection operations is to calculate cross-correlation between injection rates and seismicity rates. This approach was applied in multiple cases, including Central Arkansas (Horton, 2012) and the Rocky Mountain Arsenal in Colorado (Oprsal and Eisner, 2014). Not only temporal, but also spatial correlation between injection and seismicity can be quantified. For example, Hincks et al. (2018) showed statistical correlation between the depth of injection and seismic moment released during earthquakes. While these methods provide a statistical link between injection and seismicity, a causal relation cannot be convincingly established.

To establish a causal relation between injection and seismicity, researchers resort to physics-based models. One type of models solves pressure diffusion in a porous media (e.g., Shapiro et al., 2005; Keranen and Weingarten, 2018). Another important type of

models incorporates poroelastic effects, which are the result of hydromechanical coupling between fluid and rocks. For example, Segall and Lu (2015) used poroelastic homogenous model in investigation of influence of poroelastic effects on seismicity rate. Juanes et al. (2016) using quasi-static simulation possibly disproved a causal relationship between Cavone (Italy) earthquakes and the water injection and hydrocarbon production in that area. Fan et al. (2016) performed 2D simulation of fluid flow and fault activation to study the influence of formation permeability on the timing of the Timpson (Texas) earthquake. Another example is the work of Goebel et al. (2017), who used both numerical and analytical solutions of poroelastic equations to analyze whether poroelastic stress or pore pressure was the dominant factor in triggering of the Fairview (Oklahoma) earthquakes.

By quantifying the perturbations of the pore pressure and the stresses from fluid injection, and then evaluating fault failure potentials by calculating some quantities such as the Coulomb failure stress (e.g., Fan et al., 2016), these physics-based models can build a causal relation between injection and seismicity on nearby faults. Typically, when part of a nearby fault reaches failure condition due to fluid injection, one may infer that an earthquake occurred on the fault is induced. Some of these studies further integrate these perturbations with a rate-and-state friction model of earthquake nucleation and seismicity (Dieterich, 1994) to study seismicity rate evolution associated with fluid injection operations (Segall and Lu, 2015) or time delay in earthquake triggering of the 2011 $M_w$5.6 Prague (Oklahoma) sequence (Norbeck and Horne, 2016).

One of next steps along this line of research is to build physics-based models that can reproduce the size of a main shock and possibly the timing of its aftershock sequence. The models discussed above do not address these questions. Failure of part of a fault does not tell the final extent of an earthquake rupture and thus the size of the earthquake, as earthquake rupture propagation is a dynamic process and the final rupture extent depends on not only initial, static stress conditions at the beginning of rupture as obtained in the above models for failure potential analyses, but also dynamic stress perturbations during rupture propagation (e.g., Andrews, 1976; Day, 1982; Scholz, 2002). To build a more convincing physical link between a fluid injection operation and a nearby earthquake, the scientific community needs to move beyond failure potential analyses and to further quantify the rupture extent (and thus earthquake size). Strictly speaking, a undisputable physical link between a fluid injection and a nearby earthquake can be built only after the observed earthquake size can be reproduced. Ideally, predictive models of seismic hazards from fluid injections should also be able to assess earthquake sizes in addition to failure potentials to better help operators design fluid injections. To achieve these goals, we need to examine the co-seismic rupture process.

Among earthquake rupture models, the spontaneously dynamic rupture model (i.e. the spontaneous rupture model), in which dynamic rupture propagation is governed by initial stress conditions, frictional evolution, and dynamic stress perturbations (e.g., Andrews, 1976; Day, 1982; Harris et al., 2009), is probably the most appropriate

physics-based model to examine the rupture extent (and thus the earthquake size). Recently, some researchers have started to use the spontaneous rupture model in their studies of induced seismicity. Urpi et al. (2016) used a 2D plane-strain model and a rate-and-state friction law in study of earthquakes associated with $CO_2$ injection. Jin and Zoback (2018) investigated rupture patterns and velocities with a 2D homogenous plane-strain model. Szafranski and Duan (2018) combined an analytical solution for a homogenous hydromechanical model of fluid injection with a spontaneous rupture model to explore effects of rock and fluid properties on induced seismicity in a generic faulting model.

In this study, we develop a methodology that integrates a numerical, poroelastic model of fluid flow in a heterogeneous deformable porous medium with a spontaneous rupture model of earthquakes on faults embedded in the medium to explore physical links between fluid injection and nearby earthquakes, focusing on the earthquake size issue that is beyond failure potential evaluation. We apply the method to the 2012 May 17 $M_w$ 4.8 Timpson (Texas) earthquake, demonstrating its applicability to real cases, and associated challenges and opportunities. The advantage of selecting this earthquake as a case study is its relative simplicity. There is a single predominantly strike-slip fault, only few injection wells and a single main shock. There are also multiple publications on this earthquake, including a study by Frohlich et al. (2014) which provided hypocentral locations of the main shock, preshocks and aftershocks and an analysis on earthquake intensity and injection data, hypothesizing that the earthquake might be induced by

wastewater disposal. Fan et al. (2016) performed a geomechanical analysis based on the borehole images, which can serve as a reference point in building our models. Shirzaei et al. (2016) and Shirzaei et al. (2019) provide additional constraints on the permeability of the geologic formations near the 2012 Timpson earthquake fault.

### 3.3. Theory and Methods

In this study we integrate numerical modeling of fluid flow and dynamic rupture propagation to build direct physical links between fluid injection and nearby earthquakes. We describe the fluid flow part as the two-way hydro-mechanical coupling, for which we use theory of linear poroelasticity (Biot, 1941), which is the extension (from 1D to 3D) of Terzaghi's theory of consolidation (1925). In theory of linear poroelasticity, the behavior of the medium is fully described by seven partial differential equations – six equations to obtain the complete stress tensor (σ) and one to obtain pore pressure (p). As a fluid flow governing equation, we use the combined fluid continuity equation and Darcy's Law expressing it in terms of pore pressure and mean stress (e.g. Wang, 2000):

$$\frac{\alpha}{KB}\left[\frac{B}{3}\frac{\partial \sigma_{kk}}{\partial t}+\frac{\partial p}{\partial t}\right]-\frac{1}{\mu}(\nabla k\cdot\nabla p+k\nabla^2 p)=Q, \tag{3.1}$$

where $\alpha$ is the Biot-Willis coefficient, $K$ is the bulk modulus, $B$ is the Skempton's coefficient, $p$ is pore pressure, $\sigma_{kk}$ is a trace of the stress tensor, $k$ is an intrinsic permeability, $\mu$ is fluid dynamic viscosity and $Q$ is volume of fluid per unit bulk volume per unit time (external source). We treat our wells as point sources located at certain depths. The Biot-Willis coefficient is defined as (e.g., Segall and Lu, 2015):

$$\alpha = \frac{3(\nu_u - \nu)}{B(1+\nu_u)(1-2\nu)}, \tag{3.2}$$

where $\nu$ *and* $\nu_u$ are drained and undrained Poisson's ratios.

To get stress caused by fluid injection we use a set of Beltrami-Mitchell equations (e.g. Wang, 2000):

$$\nabla^2\sigma_{ij} + \frac{1}{1+\nu}\frac{\partial^2\sigma_{kk}}{\partial x_i \partial x_j} + 2\eta\left[\frac{1-\nu}{1+\nu}\frac{\partial^2 p}{\partial x_i \partial x_j} + \delta_{ij}\nabla^2 p\right] = -\frac{\nu}{1-\nu}\delta_{ij}\,\nabla\cdot\boldsymbol{F} - \frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}, \tag{3.3}$$

where $\sigma_{ij}$ is a stress tensor, $\eta$ is the poroelastic stress coefficient, $\delta_{ij}$ is the Kronecker delta and $\boldsymbol{F}$ is external body force. The poroelastic stress coefficient is related to Poisson ratio and Biot-Willis coefficient (Wang, 2000):

$$\eta = \frac{1-2\nu}{2(1-\nu)}\alpha. \tag{2.4}$$

Equations (3.3) take into account the compatibility equations, i.e. the corresponding displacement field is continuous and single-valued. In practice, we assume that all the terms on the right-hand side of equations (3.3) are equal to zero, because we are mostly interested not in total values of stresses but in values of their perturbations, which greatly simplifies the calculations.

To solve the system of equations (3.1) and (3.3), we use a finite-difference method (FDM) implemented within our in-house code called PyFluFlow. In the code, we approximate spatial partial derivatives with central differences and time derivatives with backward differences. For stress tensor components and pore pressure, we use the Dirichlet boundary conditions (i.e., zero values for them on the boundaries are assigned).

The initial conditions of the simulations are $\sigma_{ij}(t_0) = 0$ and $p(t_0) = 0$. In the solution phase we perform the following steps:

1. At a given time step we solve equation (3.1) for pore pressure perturbation in the entire volume, assuming that the time derivative of mean stress is initially equal to zero. This method encompasses a “fixed-stress” sequential coupling (e.g. Kim et al., 2011).
2. After obtaining pore pressure, we solve the system of equations (3.3) to obtain the distribution of perturbations of all stress tensor components in the entire volume.
3. After stress perturbations are calculated, equation (3.1) is solved again. This time, the time derivative of mean stress is non-zero.
4. We perform several iterations of steps 2 and 3 to ensure a convergent solution of poroelastic effects.
5. For grid points associated with the fault plane we recalculate the stress tensor perturbations to obtain perturbations of normal and shear stresses.
6. We move to the next time step and repeat all the steps until the time limit of the simulation is reached.

PyFluFlow performance was validated in the benchmark test, where we compared a finite-difference solution with an analytic solution for a 3D homogenous model with a point source of injection and the constant injection rate proposed by Rudnicki (1986). Both solutions were calculated with the same number of time steps and the same cubic grid (see ”Description of the benchmark problem” in the supplemental material for more

details). Figure 3.1 presents the comparison of pore pressure (Figures 3.1a and 3.1c) and one of the components ($\sigma_{11}$) of the stress tensor (Figures 3.1b and 3.1d) from the analytical and numerical solutions. The differences in solutions were minor, which suggests that the code works properly.

The simulation of dynamic rupture propagation allows for capturing the processes of induced earthquake initiation, rupture propagation and arrest. Here, we solve the elastodynamic equation:

$$\nabla \cdot \boldsymbol{\sigma} + \rho_b \boldsymbol{f_b} = \rho_b \frac{\partial \boldsymbol{v}}{\partial t}, \tag{2.5}$$

where $\boldsymbol{f_b}$ is body force vector, $\rho_b$ is a bulk density of the rock, and $v$ is the velocity vector. As a fault boundary condition in our solution, we use traction-at-split-node (TSN) scheme in the formulation of Day et al. (2005). For the frictional behavior description, we implement the linear slip weakening law (e.g. Ida, 1972), in which friction coefficient drops from the static ($\mu_s$) to dynamic ($\mu_d$) value along the critical slip distance ($D_0$):

$$\mu(s) = \begin{cases} \mu_s - (\mu_s - \mu_d)\frac{|s|}{D_0}, |s| < D_0 \\ \mu_d, |s| \geq D_0 \end{cases}, \tag{2.6}$$

where $|s|$ is the slip distance and $\mu(s)$ is friction coefficient (not fluid dynamic viscosity as above). Finally, as a fault failure criterion, we use the Mohr-Coulomb failure criterion:

$$\tau \leq c + \mu\sigma_n', \tag{2.7}$$

where $\tau$ is shear stress on the fault plane, $c$ is the cohesion of the fault plane and $\sigma_n'$ is effective normal stress as defined:

$$\sigma_n' = \sigma_n - \alpha p. \tag{2.8}$$

In our numerical implementation of equation (2.5), we use a standard formulation of an explicit finite-element method (FEM) (e.g. Hughes, 2000). The in-house software for dynamic rupture simulation called EQdyna (e.g. Duan and Oglesby, 2006; Duan, 2012; Liu and Duan, 2018) was successfully validated within the Southern California Earthquake Center (SCEC) code verification exercise (Harris et al., 2009; 2011; 2018). For the model boundaries, we use Perfectly Matched Layer (PML) boundary conditions (e.g. Berenger, 1994; Liu and Duan, 2018). In both fluid flow and dynamic rupture simulations, we use hexahedral, non-overlapping elements.

PyFluFlow and EQdyna are coupled together through transferring relevant, calculated results of pore pressure and stresses. First, the perturbations of pore pressure and stresses in the model due to a certain history of fluid injection are calculated by PyFluFlow. Then, these perturbations are added to the initial stress and pore pressure values prescribed as inputs to EQdyna (see the next section for details). After dynamic events simulated by EQdyna, stress changes (drop or increase) from EQdyna are used as inputs to PyFluFlow. Then, the above processes repeat as needed to simulate fluid injection and a sequence of earthquakes.

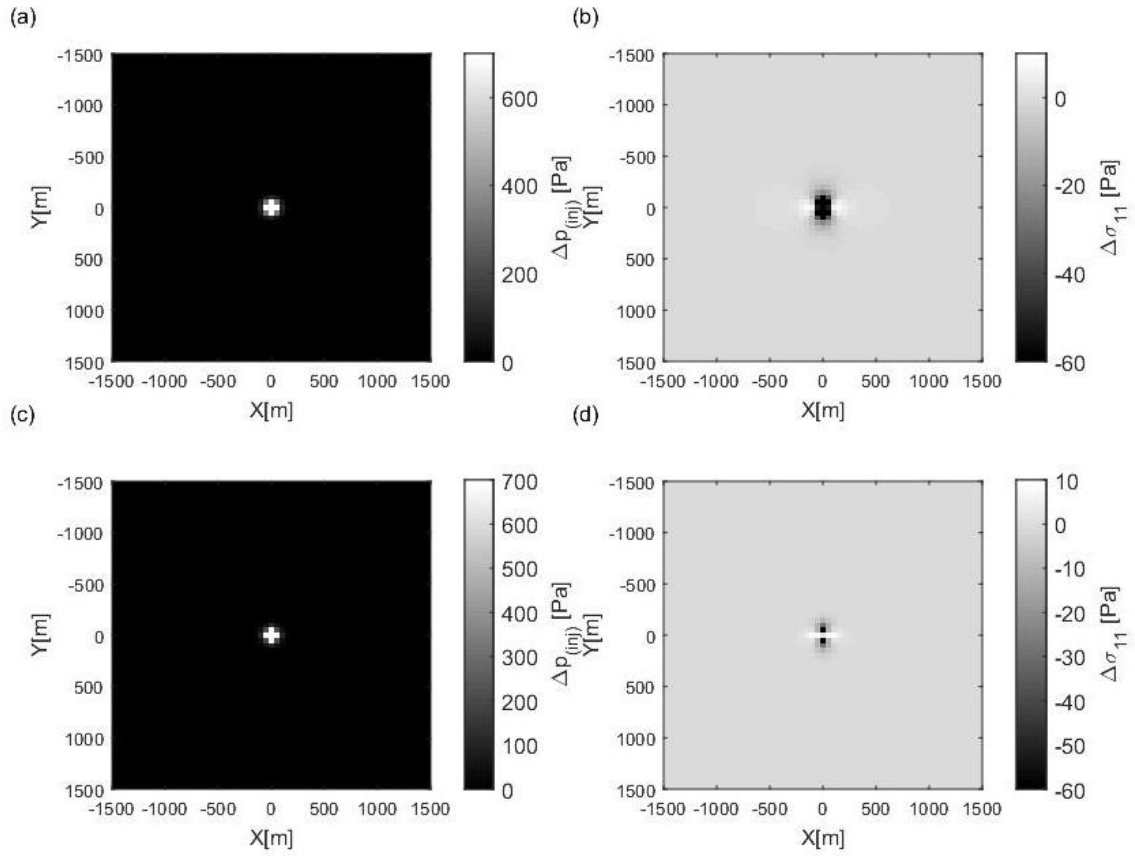


**Figure 3.1 Comparison of analytical ((a) and (b)) with numerical ((c) and (d)) results of benchmark tests for pore pressure perturbations ((a) and (c)) and one component of stress tensor (σ11) perturbations ((b) and (d)).**

### 3.4. Models

To build fluid flow models of the Timpson case, we combine pieces of information from several sources. We collect the injection data (The Railroad Commission of Texas reports) for the period between Jan 2007 and Dec 2017 for 4 injection wells analyzed by Frohlich et al. (2014). The changes of injected volumes over time are presented in Figure 3.2a. Note that reports contain monthly injected volumes, while in our simulations we use 1-day time step, and therefore we evenly split the volumes between each day of the

month. The surface locations of the wells (from The Railroad Commission of Texas) and epicenters of the earthquakes, which we compiled from both the USGS seismic catalog and Frohlich et al. (2014), are presented in Figure 3.2b. All the considered wells are injecting into Rodessa and Pine Island Formations at the depths between 1783 m and 1859 m. The Rodessa Formation is overlain by the impermeable Ferry Lake Formation. Several other sedimentary formations occur between the Pine Island Formation and crystalline basement, the oldest of which are Lower Jurassic in age (Shirzaei et al., 2016).

The hydromechanical model we use for the subsurface is a simplified horizontally-layered model similar to the one used by Shirzaei et al. (2019) (see Table 3.1 with modified model from Shirzaei et al., 2019). Parameter values presented here are constrained under the assumption of undrained rock conditions, i.e. the rate of stress change is high and the fluid cannot enter or leave the pore space fast enough to equilibrate (low to moderate permeabilities). In this model we neglected the gentle southwestward dip (0.6°; Fan et al., 2016) of the rock layers. The model domain has the size of 22.2 x 22.2 x 6 km (Northing x Easting x Depth) with grid spacing of 250 x 250 x 500 m for fluid flow simulations. Fault strike and dip we use are 330 (NW-SE trending) and 63 degrees towards SW, respectively. We pick these values based on the aftershock hypocenters distribution (Frohlich et al., 2014 and USGS Earthquake Catalog available at https://earthquake.usgs.gov/earthquakes/search/). The fault plane extends vertically between 1500 m and 5000 m in the subsurface (also based on aftershocks distribution

from Frohlich et al., 2014) , which might be larger than the activated segment considering relatively high uncertainty of the events vertical locations In our model, the fault trace is straight and is 11 km long, where the length is limited to the segment of fault which we expect to be most significantly affected by the water injection for the sake of computational efficiency. We remark that the fault dimensions we choose here are larger than the rupture extent of the 2012 $M_w$4.8 earthquake, as we explore spontaneous rupture of that main shock (and thus the rupture extent is not predefined). The fault zone thickness in our fluid flow calculation is equal to one grid spacing (i.e. 250 m) The fault zone in our simulations has different properties ($\nu = 0.25$; $\nu_u = 0.45$; $B = 0.75$; $k = 10\times10^{-16}$ m$^2$; $K = 3$ GPa) from the surrounding intact rocks (Table 3.1).We assume that the fault zone is less permeable than the injection formation (though not completely impermeable as in Fan et al., 2016) and that the rock within the fault zone is damaged and more easily compressible (here lower bulk modulus). For the Poisson ratio and Skempton's coefficient of the fault zone, we assume they are the same as the deepest layer in our model (see Table 3.1). The value of injected fluid dynamic viscosity is equal to $1.1\times10^{-3}$ Pa·s based on Shirzaei et al. (2019).

There are three main ingredients in a spontaneous rupture model (e.g., Harris et al., 2009), including geologic structure (e.g. fault geometry) and rock properties, fault frictional properties (and friction law), and initial fault stresses (shear and normal components). The slip-weakening law (i.e., Equation (3.6)) is the most widely used friction law in rupture dynamics and supported by laboratory experiments (e.g.,

Dieterich, 1980; Ohnaka and Yamashita, 1989). In this friction law, there are three important frictional parameters (on the fault plane), including static and dynamic frictional coefficients, $\mu_s$ and $\mu_d$, and the critical slip distance $D_0$. Based on many observations of fault activation in sedimentary basins (Zoback, 2010), we use a static friction coefficient value of $\mu_s = 0.6$, which aligns well with the Byerlee's law in rock friction experiments (Byerlee, 1978). $D_0$ *and* $\mu_d$ are relatively poorly constrained. In general, $\mu_d$ should be smaller than $\mu_s$ for fault sliding, and it largely dominates the stress drop level for a given initial stress condition. The stress drop level dictates fault slip and plays an important role in determining the earthquake magnitude (with the rupture extent and rock shear modulus as other two parameters). Therefore, we explore a range of $\mu_d$ values (between 0.2 and 0.5) by using the spontaneous rupture model and the observed earthquake size. Laboratory experiments of rock friction shows values of $D_0$ on the order of μm (e.g., Dieterich, 1980) and seismic inversions shows values of $D_0$ on the order of tens of cm (e.g., Mikumo et al., 2003). The difference may be caused by 1) scale-dependence of $D_0$ (cm-scale sample sizes in laboratory versus tens to hundreds of km faults in seismic studies), and 2) purely frictional behavior in laboratory experiments (thus very small $D_0$ values) versus off-fault energy dissipation processes such as nearby-fault damage factored into $D_0$ values in seismic studies besides frictional energy dissipation, among other factors. In this study, we choose $D_0$ values that align well with seismic observations and use the spontaneous rupture model and the observed earthquake size to explore a range of $D_0$ values (between 0.05 m and 0.3 m). The fault

cohesion in the Mohr-Coulomb failure criterion (Equation 7) is chosen to be zero in this study, which is consistent with Byerlee's law (Byerlee, 1978).

Another important ingredient in a spontaneous rupture model is the initial shear and normal stresses on the fault plane before a dynamic event (earthquake). The perturbations calculated from the above fluid flow model are only part of the initial stresses in this study. The other part is a background stress level that needs to be set up. As a guide, we make use of the orientation of the maximum horizontal stress (*SHmax*) as well as the magnitude of the minimum horizontal stress (*Shmin)* from Fan et al. (2016), which were obtained from the breakouts and drilling-induced fractures orientations. Based on their work we use an azimuth of 079 degrees for *SHmax*. To start, we calculated the hydrostatic pore pressure profile and the vertical stress based on the density profile. However, the magnitude of the maximum horizontal stress is not well constrained. Although one may test a range of values for *SHmax* and resolves the principal stresses onto the fault plane to obtain the initial fault normal and shear stresses, we adopt a different approach in this study. We directly set up two ratios, the normal-to-vertical stress ratio and the initial-to-yield shear stress ratio. The yield stress is the product of the effective normal stress and the static friction coefficient. We find that a value of 1.4 for the normal-to-vertical stress ratio and a value of 0.6 for the initial-to-yield shear stress ratio are consistent with the orientation of *SHmax* and the values of pore pressure and *Shmin* reported by Fan et al. (2016; see Figure 3.3). These values of stresses and pore pressures are used as the starting point for our dynamic rupture

simulations, and then we introduce some complexities in their distributions to mimic heterogeneities within the fault zone in the mechanical, hydraulic and frictional properties. Besides the depth-dependent stress condition as shown in Figure 3.3, we also test cases with the uniform effective normal stress condition. The uniform effective normal stress condition can be caused by overpressurization at depth, which has been observed in many sedimentary basins (e.g. Zoback, 2010).

Beside homogenous fault models, we also construct heterogenous models with random distributions of selected model parameters to capture unknown heterogeneities in the fault zone. In each of these models, we assign specific normal distributions (average and standard deviation) of selected parameters. For each element in our model, we pick random values based on these specific normal distributions for those selected parameters. Furthermore, in some of these models, we also apply a smoothing Gaussian filter to increase the average size of heterogeneities (from one-element size to multiple-element size). We present how the Gaussian filter works in the ”Gaussian smoothing explanation” chapter in the supplemental material. The model parameters with randomized distributions include normal-to-vertical stress ratio ($\sigma_n/\sigma_v$) (reflecting local stress reorientations), initial-to-yield shear stress ratio ($\tau_0/\tau_u$) (indicating differences in shear stress accumulation levels), static and dynamic friction coefficients ($\mu_s$ and $\mu_d$), critical slip distance ($D_0$) and the pore pressure deviation from a given background trend ($\Delta p$). For these parameters, we also impose some restrictions, such as the requirement for

initial shear stress value to be smaller than yield stress. Table 3.2 gives some examples of the models.

**Table 3.1 Hydromechanical layered model of the subsurface (Modified from Shirzaei et al., 2019).**

| Layer top depth [m] | $\nu$ | $\nu_u$ | $B$ | $k$ [m$^2$] | $K$ [GPa] | $\rho_b$ [kg/m$^3$] |
|---|---|---|---|---|---|---|
| **0** | 0.33 | 0.45 | 0.80 | $6.80 \cdot 10^{-16}$ | 18.5 | 2293 |
| **500** | 0.33 | 0.45 | 0.80 | $1.00 \cdot 10^{-14}$ | 24.8 | 2405 |
| **1000** | 0.23 | 0.45 | 0.80 | $4.51 \cdot 10^{-16}$ | 32.1 | 2470 |
| **1500** | 0.23 | 0.45 | 0.90 | $1.00 \cdot 10^{-14}$ | 40.6 | 2572 |
| **2000** | 0.25 | 0.45 | 0.75 | $3.00 \cdot 10^{-16}$ | 52.0 | 2622 |
| **3000** | 0.25 | 0.45 | 0.75 | $2.92 \cdot 10^{-17}$ | 53.5 | 2650 |

**Table 3.2 Normal distributions (average +/- standard deviation) used for generation of models for dynamic rupture simulation along with indication whether parameter distribution was smoothed and type of background stress/pore pressure trend.**

| | $\sigma_n/\sigma_v$ | $\tau_0/\tau_u$ | $\mu_s$ | $\mu_d$ | $D_0$ [m] | $\Delta p$ [% of p] | Stress pattern and frictional parameters |
|---|---|---|---|---|---|---|---|
| **Model 1** | 1.4 ± 0.0 | 0.6 ± 0.0 | 0.6 ± 0.0 | 0.3 ± 0.0 | 0.1 ± 0.0 | 0 ± 0 | $\sigma_{neff}$ const and homogenous |
| **Model 2** | 1.4 ± 0.3 | 0.6 ± 0.1 | 0.6 ± 0.2 | 0.3 ± 0.2 | 0.05 ± 0.05 | 0 ± 10 | $\sigma_{neff}$ const and heterogenous |
| **Model 3** | 1.4 ± 0.3 | 0.6 ± 0.1 | 0.6 ± 0.2 | 0.3 ± 0.2 | 0.1 ± 0.05 | 0 ± 10 | $\sigma_{neff}$ const and heterogenous |
| **Model 4** | 1.4 ± 0.3 | 0.6 ± 0.1 | 0.6 ± 0.2 | 0.3 ± 0.2 | 0.1 ± 0.05 | 0 ± 10 | $\sigma_{neff}$ const and smooth heterogenous |
| **Model 5** | 1.4 ± 0.3 | 0.6 ± 0.1 | 0.6 ± 0.2 | 0.4 ± 0.2 | 0.05 ± 0.05 | 0 ± 10 | $\sigma_{neff}$ const and smooth heterogenous |

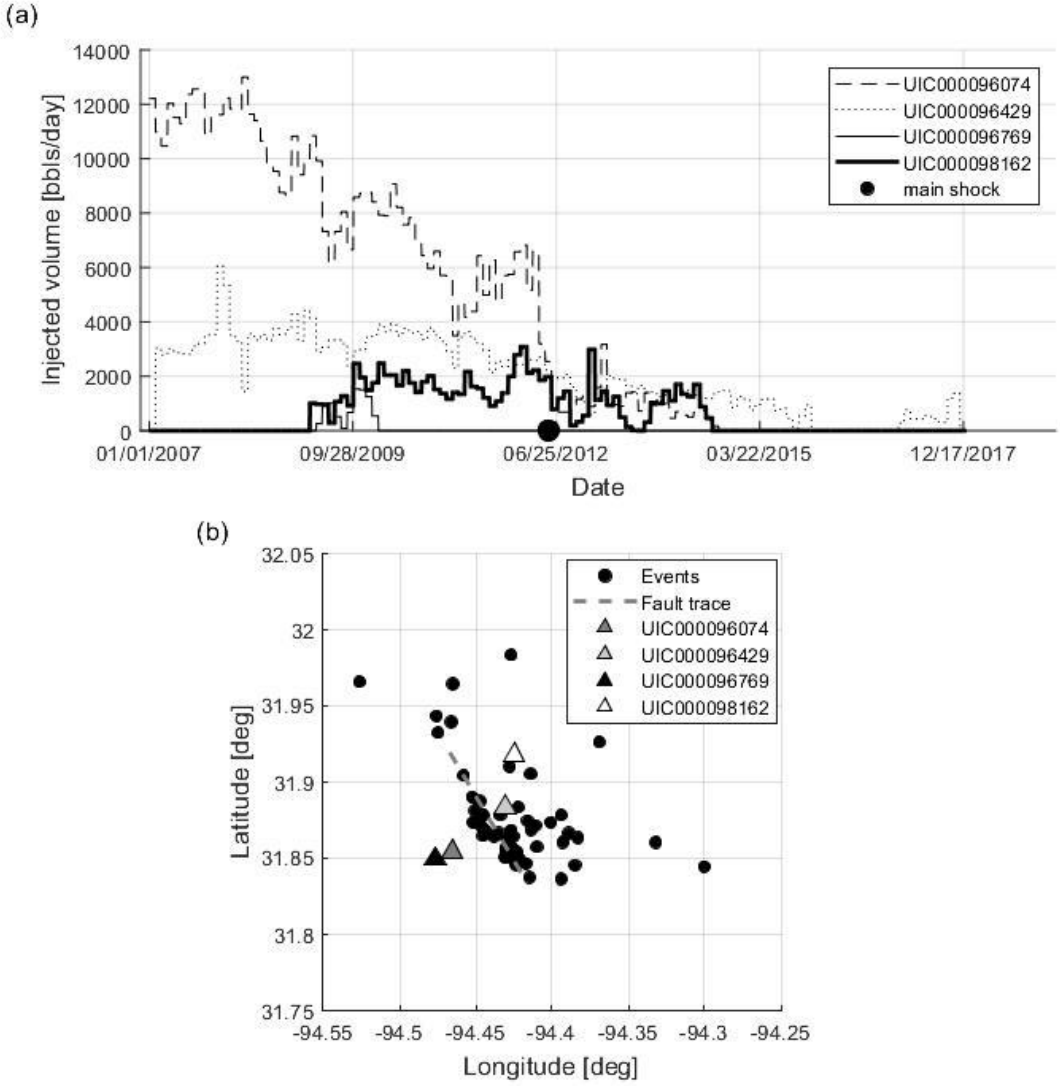


**Figure 3.2 (a) Injected volume changes in the period of 01/01/2007 till 12/31/2017. Filled circle shows the time of the main shock occurrence and (b) map of injection well locations (triangles) and earthquake epicenter locations (circles) with marked fault trace (dashed line).**

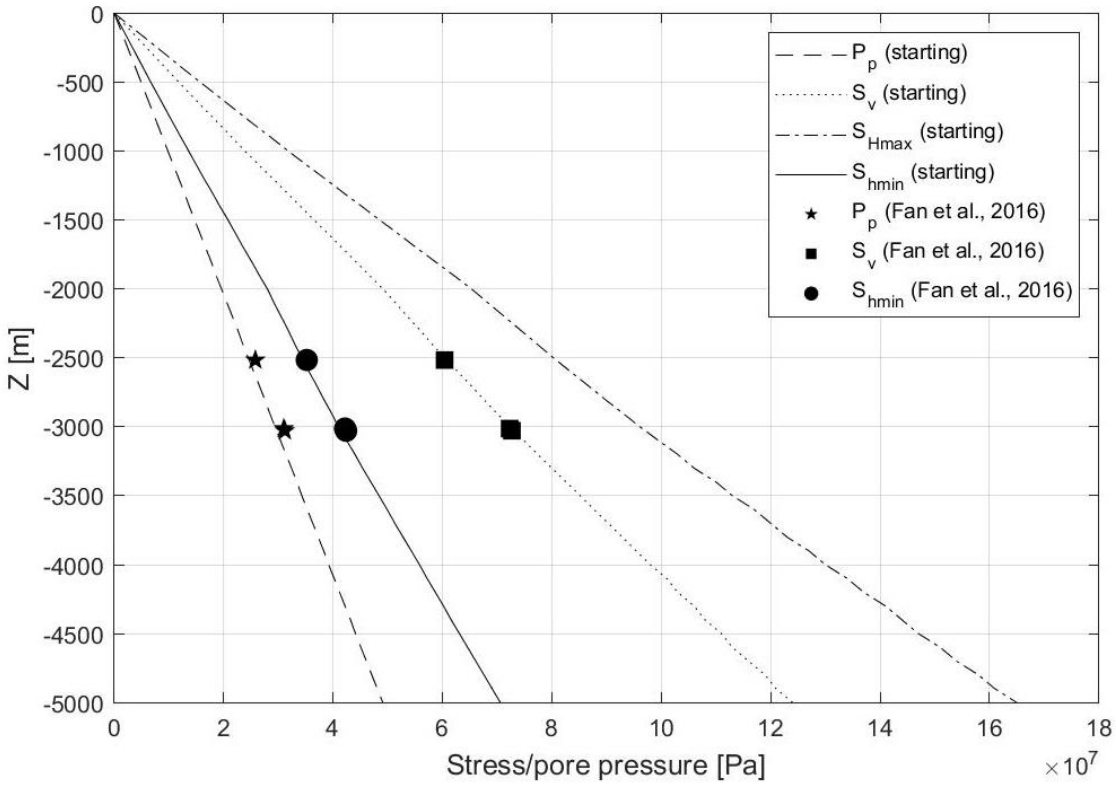


**Figure 3.3 Vertical profiles of stresses and pore pressures used as a starting point of the analysis (lines), as well as values reported by Fan et al. (2016).**

## 3.5. Results

### 3.5.1. Results of fluid flow simulations

Using the model described in the previous section, including the wastewater injection volume data (Figure 3.2a), we run a simulation of fluid flow to obtain an evolution of pore pressure and stress perturbations. Distributions of the pore pressure (Figure 3.4a), normal stress (Figure 3.4b) and shear stress (Figure 3.4c) perturbations on the fault are for the day of main shock occurrence (05/17/2012) due to fluid injection since 2007 from the four wells (Figure 3.2).

Perturbations are mainly localized in the shallow portion of the fault, and are slightly shifted towards the southeast. The peak values of perturbations reach ~4.7 MPa, ~8.1 MPa and ~4.8 MPa for pore pressure, normal stress and shear stress respectively. The magnitudes of the perturbations are approximately in an agreement with the values reported by Fan et al. (2016).

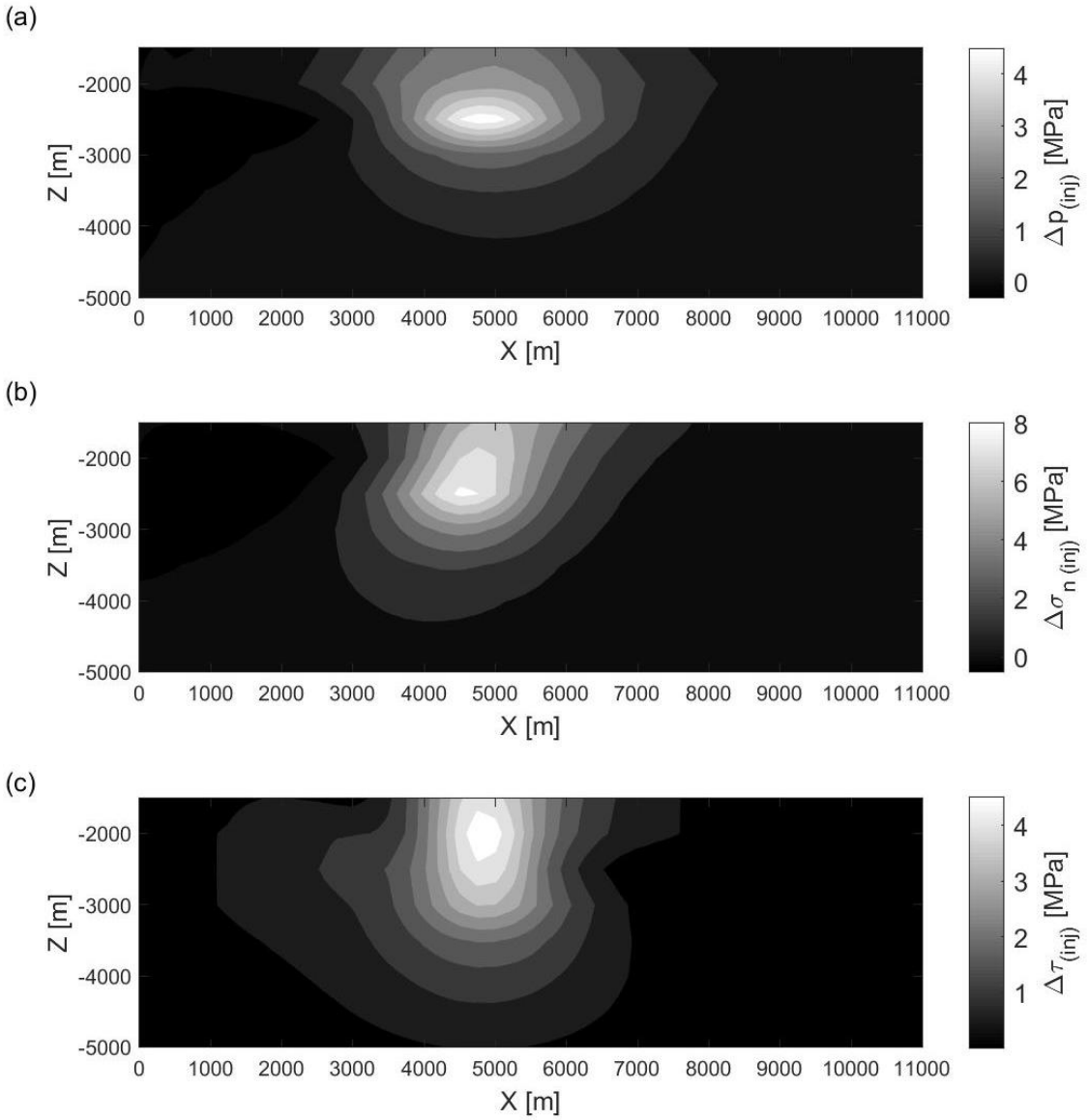


**Figure 3.4 Distribution of (a) pore pressure, (b) normal stress and (c) shear stress perturbations on the fault plane (SE-NW from left to right) on the day of the main shock (05/17/2012). The perturbations are the result of injection from 4 wells presented in Figure 3.2.**

### 3.5.2. Results of dynamic rupture simulations

To build physical links between the Timpson earthquake and the wastewater injection operations in the area, it is necessary to match the timing, size and mechanism of the

earthquake. We assume that the produced perturbations of pore pressure and stress at the time of the earthquake are just large enough to initiate rupture on the fault, and there were no major earthquakes between the beginning of the simulated injection (01/01/2007) and the day of the mainshock occurrence (05/17/2012), and our hydromechanical model of the subsurface is correct. This assumption allows us to focus on the second part of the task, which is matching the size and mechanism of the earthquake by spontaneous rupture modeling. Because of a lack of seismic monitoring stations nearby the activated fault segment, the kinematic inversion of slip, stress drop, or slip rate is not feasible. Due to sparsity of the monitoring network, the hypocenter depth is also very poorly constrained. The only benchmark information obtained about the earthquake is the moment magnitude and focal mechanism (Frohlich et al., 2014), even though there are non-negligible uncertainties associated especially with the focal mechanism inversion. Therefore, our goal of dynamic rupture simulations is to reproduce the earthquake with a similar magnitude as reported for the main shock ($M_w$4.8) with predominantly strike-slip motion (not trying to exactly match the reported mechanism). Our dynamic rupture models produce the results of rupture area and slip on the fault. Together with shear modulus of the rock, we can calculate the moment magnitude from these models. Additionally, for the models that match the observations for the main shock, we made an attempt to reproduce the aftershocks which were at least M2.0 and with hypocenter locations lying in the vicinity of the main shock (+/- several kilometers), which gave 10 aftershocks happening between May, 2012 and September, 2013.

We explore three fault friction conditions: uniform frictional parameters, heterogeneous frictional parameters and smoothed heterogeneous parameters. For each friction condition, we explore two stress conditions: hydrostatic pore pressure with depth-dependent stress, and overpressure with uniform effective normal stress. That results with six friction and stress conditions. For each condition, we explore parameter space of critical slip distance ($D_0$) and dynamic friction coefficient parameter ($\mu_d$). The results are summarized in Figure 3.5, and then we show detailed slip and rupture time distributions (Figures 3.6 to 3.10) for 5 selected representative models listed in Table 3.2.

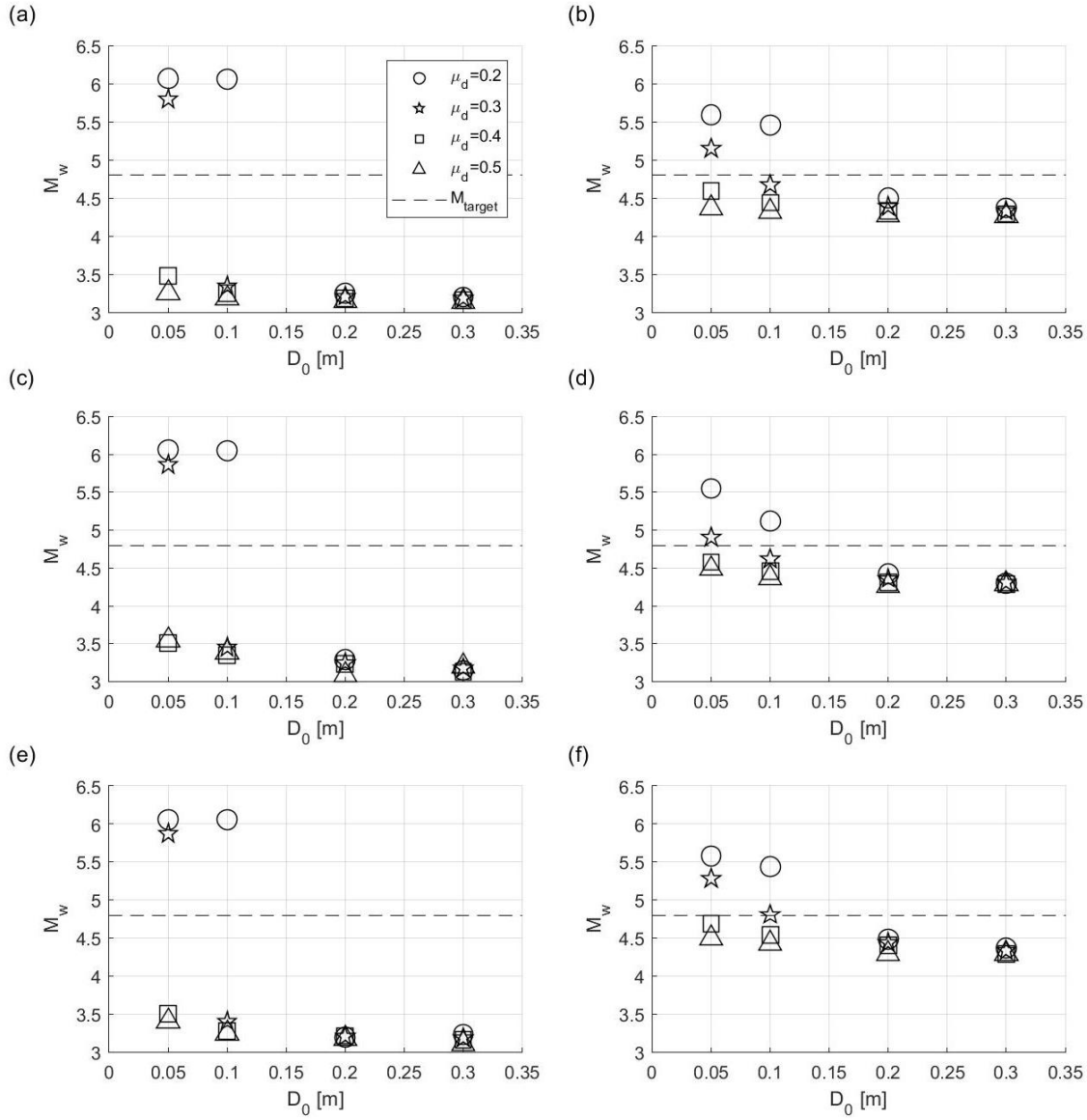


**Figure 3.5 Changes of simulated earthquake magnitude with changes of critical slip distance and dynamic friction coefficient values for models with: (a) constant values of frictional parameters, depth-dependent stress and hydrostatic pore pressure; (b) constant values of frictional parameters and uniform effective normal stress; (c) normal distribution of frictional parameters values (heterogenous), depth-dependent stress and hydrostatic pore pressure; (d) normal distribution of frictional parameters values (heterogenous) and uniform effective normal stress; (e) normal distribution of frictional parameters values (smoothed heterogenous), depth-dependent stress and hydrostatic pore pressure; (f) normal distribution of frictional parameters values (smoothed heterogenous) and uniform effective normal stress.**

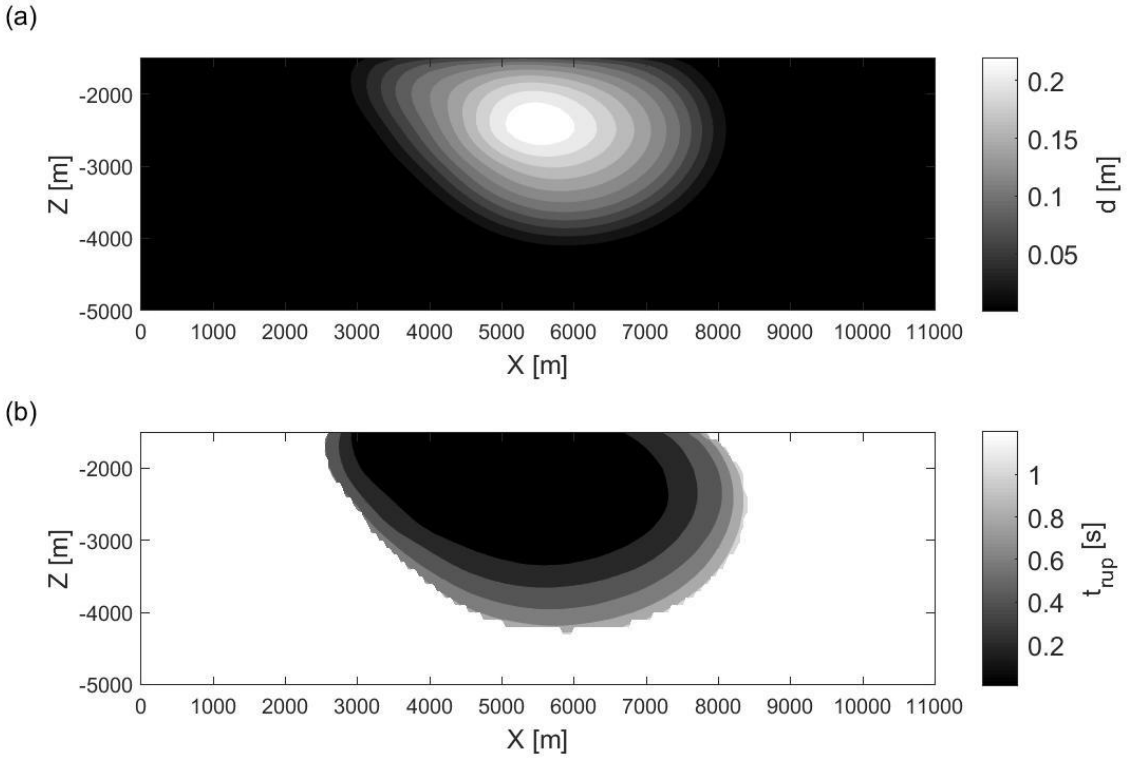


**Figure 3.6 Distributions of (a) total slip and (b) rupture time on the fault plane (SE-NW from left to right) for Model 1 from Table 3.2, where we have constant values of the frictional parameters (D0 = 0.1 m and μd = 0.3) and the initial stress state with the uniform effective normal stress. The simulated magnitude is $M_w$4.7, which is close to reported $M_w$4.8 for Timpson (TX) earthquake.**

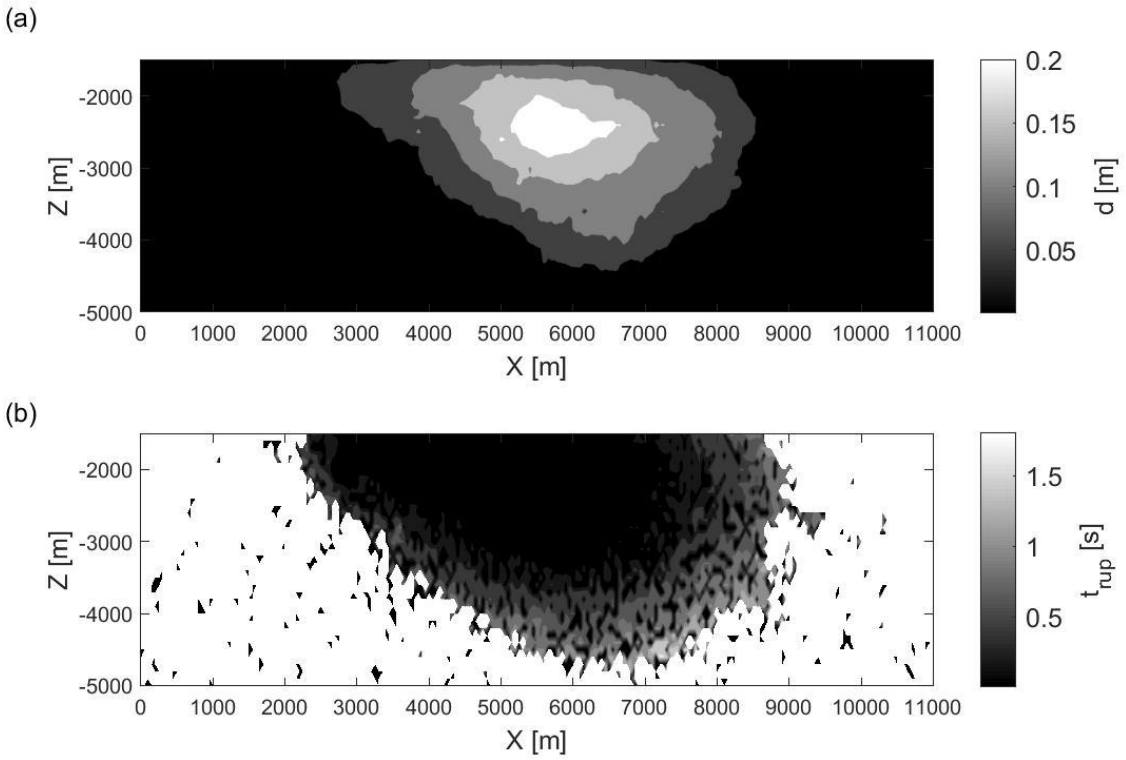


**Figure 3.7 Distributions of (a) total slip and (b) rupture time on the fault plane (SE-NW from left to right) for Model 2 from Table 3.2, where we have normal distributions of the frictional parameters (D0 = 0.05 ± 0.05 m and μd = 0.3 ± 0.2) and the initial stress state with the uniform effective normal stress. The simulated magnitude is $M_w$4.9, which is close to reported $M_w$4.8 for Timpson (TX) earthquake.**

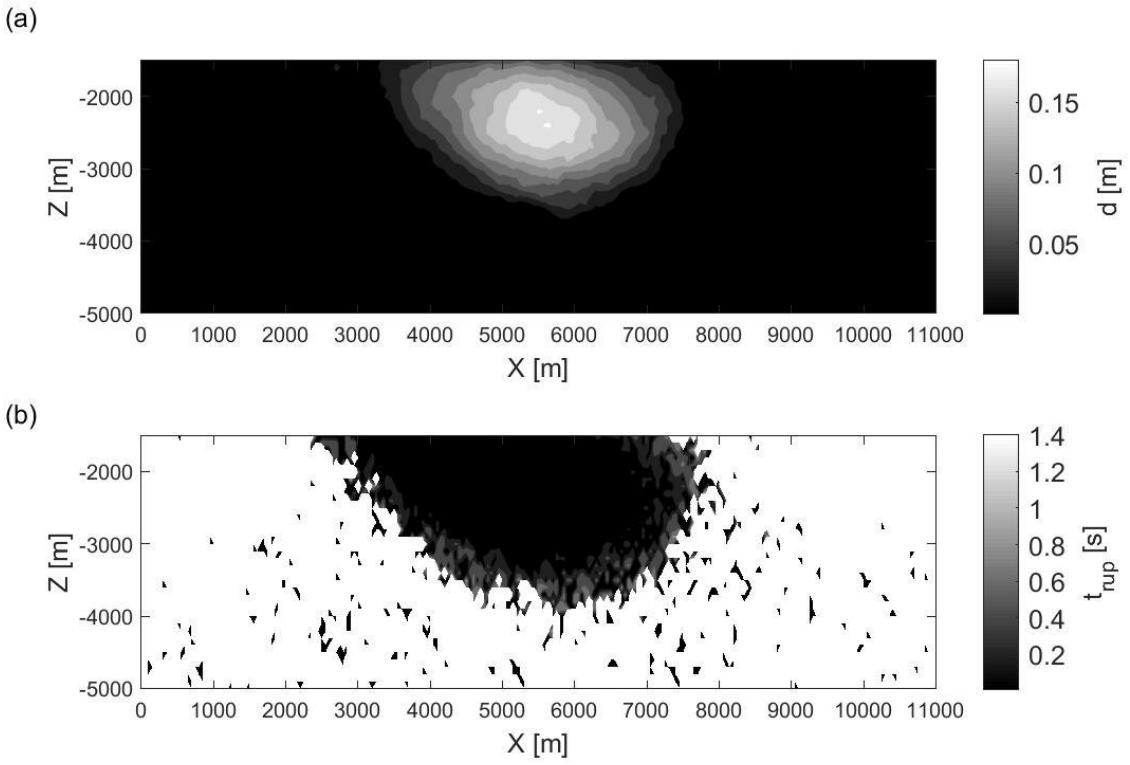


**Figure 3.8 Distributions of (a) total slip and (b) rupture time on the fault plane (SE-NW from left to right) for Model 3 from Table 3.2, where we have normal distributions of the frictional parameters (D0 = 0.1 ± 0.05 m and μd = 0.3 ± 0.2) and the initial stress state with the uniform effective normal stress. The simulated magnitude is $M_w$4.6, which is close to reported $M_w$4.8 for Timpson (TX) earthquake.**

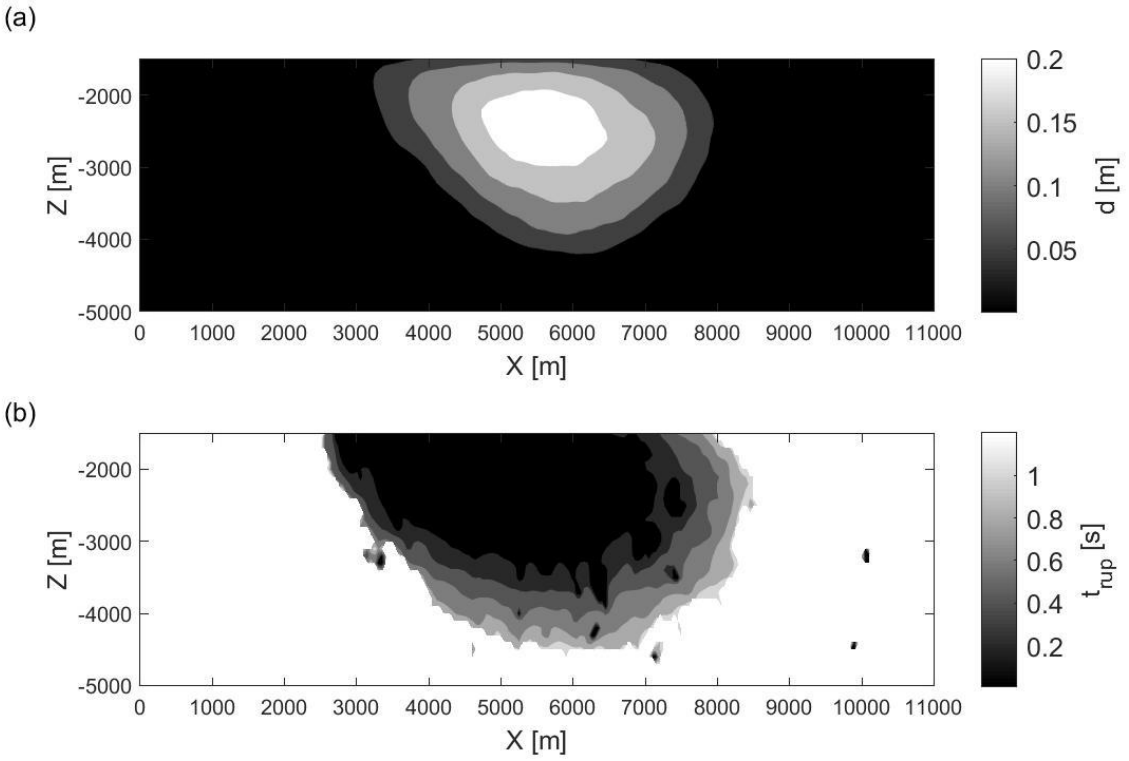


**Figure 3.9 Distributions of (a) total slip and (b) rupture time on the fault plane (SE-NW from left to right) for Model 4 from Table 3.2, where we have normal distributions of the frictional parameters (D0 = 0.1 ± 0.05 m and μd = 0.3 ± 0.2), applied Gaussian filter, and the initial stress state with the uniform effective normal stress. The simulated magnitude is $M_w$4.8, which is our best match to the reported $M_w$4.8 for Timpson (TX) earthquake.**

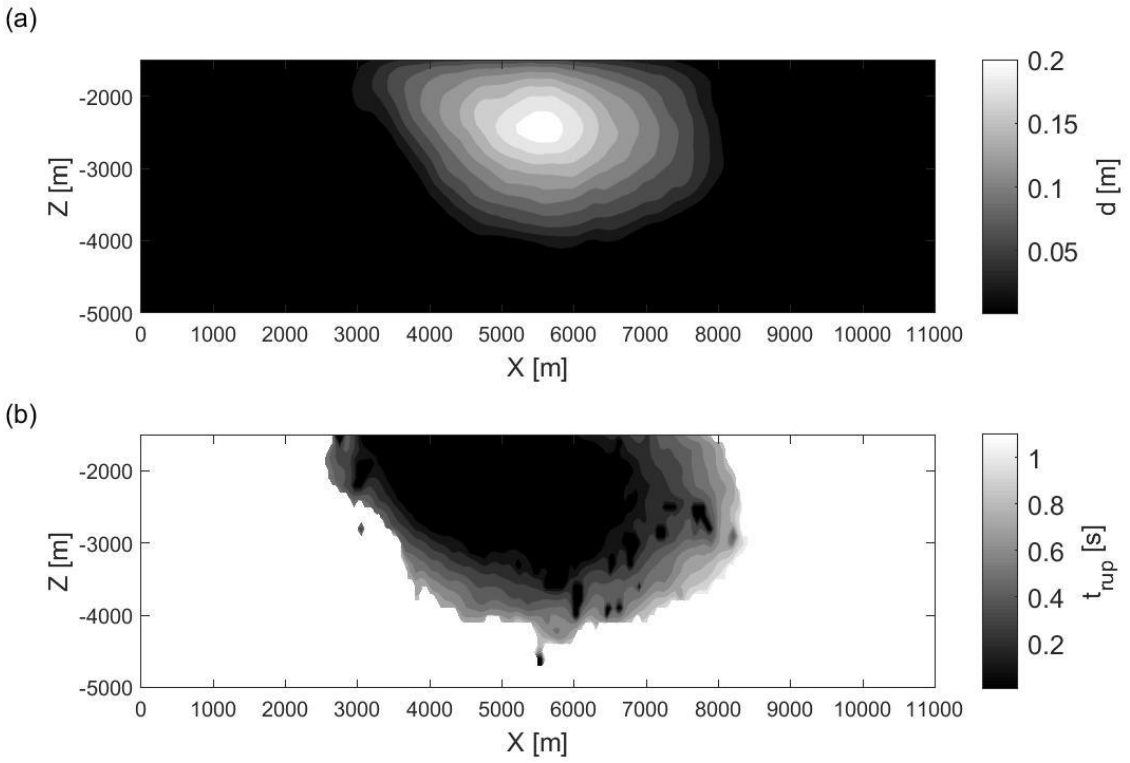


**Figure 3.10 Distributions of (a) total slip and (b) rupture time on the fault plane (SE-NW from left to right) for Model 5 from Table 3.2, where we have normal distributions of the frictional parameters (D0 = 0.05 ± 0.05 m and μd = 0.4 ± 0.2), applied Gaussian filter, and the initial stress state with the uniform effective normal stress. The simulated magnitude is $M_w$4.7, which is close to reported $M_w$4.8 for Timpson (TX) earthquake.**

### 3.5.3. Results of dynamic rupture simulations for homogenous fault

We start by testing the initial model (uniform frictional properties, depth-dependent stress and hydrostatic pore pressure) to explore parameter space for $D_0$ and $\mu_d$. In all cases, we obtain the activation on the fault. We make two major observations based on these simulations. First, the critical slip distance and the difference between static and dynamic friction coefficients have a great impact on the final size of ruptures. Second, the change of rupture size is non-linear. Figure 3.5a illustrates the changes of moment

magnitude $M_w$ for the selected range of critical slip distance ($D_0$) and dynamic friction coefficient ($\mu_d$) values. For higher values of $D_0$ and $\mu_d$, the magnitudes of produced earthquakes do not exceed $M_w$3.5 which is caused by the fact that the rupture barely propagates outside of the initiation patch. For lower values of $D_0$ and $\mu_d$ there is a large jump in simulated magnitudes to ~$M_w$6.0. For the cases of $M_w$6.0, the entire fault plane is ruptured. We remark that ~$M_w$6.0 events from this parameter space exploration do not mean this fault can produce ~$M_w$6.0 earthquakes, which may be too alarming. In particular, we notice that the simulated slip for these events exceeds 1.6 m in some parts of the fault plane, which is unrealistic. As a comparison, the coseismic slip inversion results for 2004 $M_w$6.0 Parkfield (CA) earthquake indicate the maximum slip of ~0.3 m (Wang et al., 2012). Therefore, we consider ~$M_w$6.0 events as an indication that the combinations of $D_0$ and $\mu_d$ values in these models are unrealistic and should be ruled out for feasible values for this earthquake. It is also worth noticing that increasing $D_0$ can counterbalance the effect of decreasing $\mu_d$ and vice-versa, suggesting a tradeoff between these two model parameters.

Next, we repeat the test with exploration of the critical slip distance and dynamic friction coefficient parameter space for assumption of the uniform normal effective stress across the fault. Figure 3.5b shows the changes of moment magnitude $M_w$ as we change $D_0$ and $\mu_d$. Again, in all cases the fault is being activated and we see a big impact of these frictional parameters on the final result of simulation. However, the non-linearity is less prominent, because the difference between the lowest simulated magnitude (~M4.3 for

high $D_0$ and high $\mu_d$) and the highest one (~M5.6 for low $D_0$ and low $\mu_d$) is much smaller than in the previous test. In fact, for the values of $D_0 = 0.1$ m and $\mu_d = 0.3$ (see Model 1 in Table 3.2), we obtain the earthquake with M4.7, as well as reasonable slip and rupture time distributions (Figures 3.6a and 3.6b respectively). The peak value of the slip is ~23 cm and the rupture propagation process takes ~1.2 s. The motion on the fault is predominantly a strike-slip and there is a visible directivity in the rupture propagation, where it propagates faster northwards (~ 2250 m/s) than southwards (~1650 m/s).

### 3.5.4. Results of dynamic rupture simulations for heterogenous fault (non-smooth)

Even though, we approximately reproduced the Timpson $M_w$4.8 earthquake with uniform distribution of frictional parameters and relatively simple background stress pattern, realistic faults are heterogenous. To represent these heterogeneities, we use random distributions of model parameters (element-to-element heterogeneities). We perform a similar set of tests as for the uniform parameters distributions, varying the averages in normal distributions of critical slip distance and dynamic friction coefficient, while keeping the same standard deviations (see e.g. Model 2 in Table 3.2). Figure 3.5c shows the changes of $M_w$ with changing $D_0$ and $\mu_d$ for the case of hydrostatic pore pressure (with appropriate random variations), while Figure 3.5d presents the results for the case with uniform effective normal stress. Similarly, to cases with uniform distributions of fault properties, the simulations with hydrostatic background pore pressure do not reproduce the size of the Timpson earthquake ($M_w$4.8) and an impact of the $D_0$ and μd changes on resulting magnitude is highly non-linear. At the same time,

with heterogenous fault and uniform effective normal stress we have 2 instances where we were able to approximately match the target magnitude. Using Model 2 (Table 3.2) we got $M_w$4.9 and with Model 3 (Table 3.2) simulation results in Mw4.6 earthquake. The frictional parameters and stress distributions on the fault plane for both cases are presented in the figures in ”Distributions of frictional parameters, pore pressure and stresses in models used in simulations” chapter in the supplemental material. For Model 2 rupture time distribution (Figure 3.7b) indicates the propagation time of ~1.7 s. The peak slip (see Figure 3.7a) reaches the value of ~22 cm located at the depth of ~2.5 km. For Model 3 it takes rupture ~1.4 s to propagate (see Figure 3.8b) with the peak slip value of ~18 cm (see Figure 3.8a). The dominant motion of the fault is strike-slip. Due to very complicated rupture pattern it is difficult to confidently estimate the rupture velocities but qualitatively we can say that in both cases the propagation northwards is dominant.

### 3.5.5. Results of dynamic rupture simulations for heterogenous fault (smooth)

Next, we tested models where beside using randomized distribution of fault parameters, we also applied Gaussian filter to smooth the model, effectively increasing the size of the heterogeneities (asperities). Again, we perform the suite of simulations for different values of $D_0$ and $\mu_d$, with either hydrostatic pore pressure or constant effective normal stress. The changes of $M_w$ for cases with initial hydrostatic pore pressure (see Figure 3.5e) are very similar to what was observed in previous tests for uniform and non-smooth randomized distributions of parameters (strong non-linearity and no simulation

with reproduced target magnitude). For background uniform effective normal stress, the only major difference between models without (Figure 3.5d) and with (Figure 3.5f) smoothing is higher magnitude for average $D_0$ of 0.05 m and average $\mu_d$ of 0.3 (for smooth cases). Among the smoothed models, we also got 2 cases where we either exactly reproduced the earthquake magnitude ($M_w$4.8 for Model 4 from Table 3.2) or were very close ($M_w$4.7 for Model 5 from Table 3.2). The frictional parameters and stress distributions for both simulations are shown in the figures in ”Distributions of frictional parameters, pore pressure and stresses in models used in simulations” chapter in the supplemental material. Figure 3.9b presents the rupture time distribution for Model 4. The total propagation time of the rupture is ~1.3 s, with visible rupture directivity and rupture velocity of ~2400 m/s northwards and ~1300 m/s southwards. The peak value of slip on the fault for that model is ~22 cm (see Figure 3.9a). For Model 5 the total propagation time for the rupture is ~1.1 s (see Figure 3.10b) with northwards rupture velocity of ~1900 m/s and much slower southwards propagation with velocity of ~900 m/s. The total slip reaches ~21 cm (see Figure 3.10a). For both models the dominant motion is strike-slip.

### 3.5.6. Results of aftershock reproduction analysis

Additionally, we check whether using some of these models that reproduce the main shock we are able to also reproduce the observed aftershocks sequence (Figure 3.2b). We define aftershocks as seismic events which take place after the mainshock and are located on the same fault (or fault system). We pick only the aftershocks with M2.0 or

higher, which gives 10 aftershocks. For the 10 events, we check whether perturbations caused by continued injection together with the stress perturbations from the previous earthquakes can lead to fault activation at the points in time when these seismic events took place. For this test we use Models 1, 2 and 4. Figures 3.11a and 3.11b show the comparison of the magnitudes and released seismic moment for each event for the three models and the real observation. First and foremost, we were able to get fault activations for all points in time for all three models. Secondly, we were able to match the size of the large M4.1 aftershock in September, 2013 quite well. However, for the events which happened between December, 2012 and January, 2013, the simulated magnitudes are obviously higher than those of the reported events. Thirdly, for Model 2 we were able to get good magnitude match between the immediate aftershocks (up to 1 month from main shock). Overall, Model 4 gives the smallest value of $L_1$ norm for the released seismic moment. Model 4 is a model with a smoothed, random distribution of the parameters, suggesting heterogeneous fault properties with large sizes of the asperities better reproduce the aftershock sequence. The differences between actual aftershock sequences and the modeled ones can be the result of different factors we did not take into account, such as possibility of aseismic slip in between seismic events or not changing of the fault properties from one rupture to another.

Finally, if we look at the distribution of the aftershocks along the strike (see Figure 3.2), we can observe that the majority of them is localized in southern and central portion of the fault. We can explain it using the results of the simulations. First, the shear stress

perturbations are higher in the southern portion of the fault (Figure 3.4c). Secondly, the rupture of the main shock propagated mostly northwards (Figures 3.6-3.10), leading to significant release of shear stress in that segment while shear stress release is smaller in the southern portion. Furthermore, there is a general match between the reported and simulated locations, though exact locations differ to a certain degree.

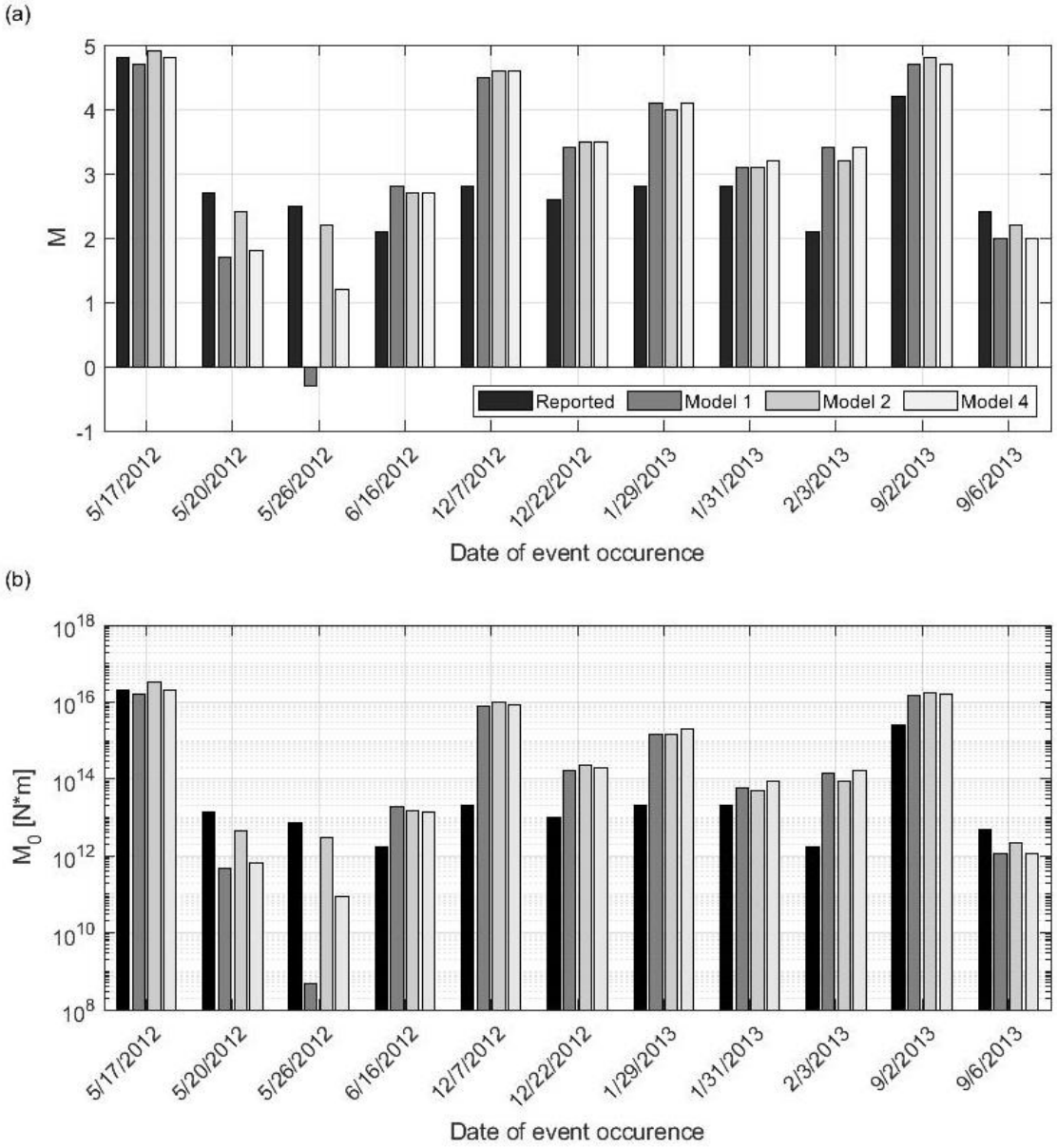


**Figure 3.11Comparison of the reported and modeled magnitudes (a) and released seismic moments (b) for the main shock and 10 aftershocks with reported magnitudes at least M2.0 which occurred within the area of interest.**

### 3.6. Discussion

To build more convincing links between fluid injection and nearby earthquakes, the scientific community needs to move further to address how big an earthquake can be

from examining mainly failure potentials of nearby faults. This requires the examination of earthquake rupture propagation.Spontaneous rupture modeling provides a physics-based means to perform this analysis. In this study, we develop a numerical method to integrate a fluid flow model and a spontaneous rupture model in 3D and use the 2012 $M_w$4.8 Timpson (TX) earthquake as a case study to demonstrate its applicability for real earthquakes. Given the fluid injection history (Figure 2) and the relatively well-constrained hydromechanical model parameters (Tables 1) of the Timpson case, we obtain the perturbations of pore pressure and stresses on the fault (Figure 4). Geological structure, including fault geometry and rock properties, is also relatively well constrained in the Timpson case. Then, we use spontaneous rupture models to explore two major factors, fault frictional properties (mainly frictional coefficients and the critical slip distance in the slip-weakening friction law) and initial background stresses. We were able to reproduce the major features of the 2012 $M_w$4.8 earthquake, mainly the earthquake size and the focal mechanism (general sense of motion), using model parameter values that are reasonable. Therefore, the integrated method provides a more direct physical link between the fluid injection and the mainshock in the Timpson case than previously failure potential analyses (e.g., Fan et al., 2016).

Furthermore, the method also largely reproduces the aftershock sequence, suggesting its potential to become a predictive tool with more efforts to improve it in the future. Historically matched models (i.e., calibrated models with validated methods and model parameter values) may be used to help future fluid injection designs in active operation

fields. By simulating different scenarios of fluid injections, operators may find a range of fluid injection designs that should not induce significant earthquakes and rule out other fluid injection scenarios that may induce significant earthquakes. Although this ability still needs many efforts to be fully robust, this study may serve as a starting effort toward this direction.

Improved understanding of earthquake physics is one of critical directions in earthquake science studies and seismic hazard analyses. Induced seismicity associated with industry activities in the recent decade generated relatively more abundant data with some known controls on earthquake generation (such as fluid injection data), and thus provides a good opportunity to study earthquake physics. To seize this opportunity, the scientific community need to develop tools and models that can assimilate these data. Our effort in this study is one example. Fault friction properties and initial stress conditions are two very important factors that dictate dynamic earthquake ruptures and thus earthquake sizes, while they are relatively poorly constrained. By matching available observations of the Timpson earthquake (mainly the earthquake size and the focal mechanism), we were able to place some constraints on these two factors. Given $\mu_s$ of 0.6, we find that $\mu_d$ of 0.3 and $D_0$ of 0.1 m could reproduce the observed mainshock size and its aftershock sequence. Between the depth-dependent effective normal stress condition (associated with hydrostatic pore pressure) and the uniform effective normal stress condition (due to over-pressurization in the fault zone), we find that the latter could reproduce the observations while the former cannot.

One of future directions in the methodology development is to add the automatic inversion of frictional parameters, initial stress and pore pressure magnitude, with quality control of the results and then evaluation of the best fitting models regarding their geological consistency. During inversion the target function can be either magnitude match, seismogram correlation, or both. In cases where local seismic network is available and the studies of rupture directivity or kinematic inversion of the slip can be done, these additional results can be used as an additional cross-check.

Another direction of the methodology improvement is implementation of more complex friction laws such as rate-and-state friction laws to capture the spontaneous nucleation phase of an earthquake as in earthquake cycle simulations (e.g., Lapusta et al., 2000; Lapusta and Liu, 2009; Liu and Duan, 2019). This will also require a more complex integration of a fluid model and an earthquake cycle model that captures both spontaneous nucleation and rupture propagation.

### 3.7. Conclusions

We develop a new method to study induced seismicity. By integrating a fluid flow model with a dynamic rupture model, we are able to go beyond failure potential analyses and build more direct links between fluid injection and nearby earthquakes by reproducing major available observations associated with the earthquakes. The

application of the method to the 2012 Timpson earthquake demonstrates that the method works well to real cases.

We get additional knowledge on how different model parameters used in dynamic rupture simulations affect the behavior of the fault and final size of the rupture. We observe that the dependence of the simulated earthquake size (magnitude) on the changes of ($D_0$) and ($\mu_d$) are moderately to highly non-linear dependent on the initial stress conditions. Both $D_0$ and $\mu_d$ affect the rupture size in a similar way (the larger the values, the smaller the rupture), suggesting a tradeoff between the two properties. For the initial stress conditions with simple hydrostatic pore pressure, we either got too small (M3.5 or smaller) or unrealistically large rupture (~$M_w$6.0). On the other hand, we find that uniform effective normal stress (i.e., overpressurization of the fault zone at depth) could reproduce the observed $M_w$4.8 earthquake. Therefore, we conclude that uniform effective normal stress is likely the stress state of the Timpson earthquake fault.

We also show that there is a potential in using integrated fluid flow and dynamic rupture simulations to analyze not only the main shock but also entire aftershock sequences. We provide a plausible explanation of the aftershock spatial distribution based on the distribution of stress perturbations and rupture pattern. In the Timpson case, most of aftershocks are located in southern and central portion of the fault which coincides with higher shear stress perturbations and mostly northward propagation of the rupture, which may be associated with lower shear stress release in the southern portion of the fault. We

also tried more quantitative analysis trying to replicate the occurrence and magnitudes of the larger (M2.0 and higher) aftershocks. It ended up with partial success. We were able to get the fault activation at all required points in time. For heterogenous models we were also able to approximately match the magnitudes of aftershocks happening during 1$^{st}$ month after main shock as well as the largest aftershock. We believe that models calibrated by history-matching present a potential for a future predictive tool for injection operations planning.

### 3.8. Data and resources

Injection data (volumes, injection depth and wells locations) used in this study was acquired from The Railroad Commission of Texas (https://www.rrc.texas.gov/about-us/resource-center/; last accessed on Jan 27, 2020). Part of the combined earthquake catalog was downloaded from the USGS Earthquake Catalog (https://www.usgs.gov/natural-hazards/earthquake-hazards/earthquakes; last accessed on Jan 27, 2020). Provided supplemental material contains the description of the benchmark problem, explanation of Gaussian smoothing, and the figures with distributions of the frictional parameters, pore pressure and stresses for models used in simulations.

### 3.9. Acknowledgements

We would like to thank Texas A&M High Performance Research Computing Center for providing the computational resources utilized for this study. We also want to thank the Crisman Institute for Petroleum Research, the Berg-Hughes Center for Petroleum and

Sedimentary Systems, and Gangi/Heep Endowed Professorship at Texas A&M University for financial support of the project through scholarships. We also wish to acknowledge Eric Peavey, Qingjun Meng, Abhishek Prakash, Guest Editor Heather DeShon and our three anonymous reviewers for valuable feedback which helped to improve this manuscript.

### 3.10. References

Ake, J., Mahrer, K., O'Connell, D., & Block, L. (2005) 'Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado', *Bull. Seismol. Soc. Am.*, 95(2), pp. 664–683.

Andrews, D. J. (1976) 'Rupture velocity of plane strain shear cracks', *J. Geophys. Res.*, 81, pp. 5679–5687.

Berenger, J. P. (1994) 'A perfectly matched layer for the absorption of electromagnetic waves', *J. Comput. Phys.*, 114(2), pp. 185-200.

Biot, M. A. (1941) 'General theory of three-dimensional consolidation', *J. Appl. Phys.*, 12(2), pp. 155–164.

Byerlee, J.D. (1978) 'Friction of rocks', *Pure Appl. Geophys.*, 116, pp. 615-626.

Davis, S. D., & Frohlich, C. (1993) ‘Did (or will) fluid injection cause earthquakes? Criteria for a rational assessment’, *Seismol. Res. Lett.*, 64(3-4), pp. 207–224.

Day, S. M. (1982) ‘Three-dimensional simulation of spontaneous rupture: The effect of nonuniform prestress’, *Bull. Seismol. Soc. Am.* 72, pp. 1881– 1902.

Day, S. M., et al. (2005) ‘Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture’, *J. Geophys. Res. Solid Earth*, 110, B12307.

Dieterich, J. (1980) ‘Experimental and model study of fault constitutive properties’, *Solid Earth Geophys. and Geotech.*, 42, pp. 21-29.

Dieterich, J. (1994) ‘A constitutive law for rate of earthquake production and its application to earthquake clustering’, *J. Geophys. Res.*, 99(B2), pp. 2601 – 2618.

Duan, B. (2012) ‘Dynamic rupture of the 2011 Mw 9.0 Tohoku-Oki earthquake: Roles of a possible subducting seamount’, *J. Geophys. Res. Solid Earth*, 117, B05311.

Duan, B., & Oglesby, D. D. (2006) ‘Heterogeneous fault stresses from previous earthquakes and the effect on dynamics of parallel strike-slip faults’, *J. Geophys. Res. Solid Earth*, 111, B05309.

Ellsworth, W. L. (2013) ‘Injection-induced earthquakes’, *Science*, 341, 1225942.

Fan, Z., Eichhubl, P., & Gale, J. F. W. (2016) ‘Geomechanical analysis of fluid injection and seismic fault slip for the Mw4.8 Timpson, Texas, earthquake sequence’, *J. Geophys. Res. Solid Earth*, 121, pp. 2798–2812.

Frohlich, C., et al. (2011) ‘The Dallas-Fort Worth earthquake sequence: October 2008 through May 2009’, *Bull. Seismol. Soc. Am.,* 101(1), pp. 327–340.

Frohlich, C., et al. (2014) ‘The 17 May 2012 M4.8 earthquake near Timpson, East Texas: An event possibly triggered by fluid injection’, *J. Geophys. Res. Solid Earth*, 119(1), pp. 581-593.

Goebel, T. H. W., et al. (2017) ‘The 2016 Mw5.1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at> 40 km from fluid disposal wells’, *Earth Planet. Sci. Lett.,* 472, pp. 50-61.

Harris, R. A., et al. (2009) ‘The SCEC/USGS dynamic earthquake rupture code verification exercise’, *Seismol. Res. Lett.,* 80(1), pp. 119–126.

Harris, R. A., et al. (2011) ‘Verifying a computational method for predicting extreme ground motion’, *Seismol. Res. Lett.*, 82(5), pp. 638-644.

Harris, R.A., et al. (2018) ‘A Suite of Exercises for Verifying Dynamic Earthquake Rupture Codes’, *Seismol. Res. Lett.*, 89(3), pp. 1146-1162.

Healy, J. H., et al. (1968) ‘The Denver earthquakes’, *Science*, 161(3848), pp. 1301–1310.

Hincks, T., et al. (2018) ‘Oklahoma's induced seismicity strongly linked to wastewater injection depth’, *Science*, 359(6381), pp. 1251-1255.

Horton, S. (2012) ‘Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake’, *Seismol. Res. Lett.*, 83(2), 250–260.

Hughes, T. J. R. (2000) *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover, Mineola, N. Y.

Ida, Y. (1972) ‘Cohesive force across the tip of a longitudinal-shear crack and Griffith’s specific surface energy’, *J. Geophys. Res.*, 77(20), pp. 3796-3805.

Jin, L., & Zoback, M. D. (2018) 'Modeling Induced Seismicity: Co-Seismic Fully Dynamic Spontaneous Rupture Considering Fault Poroelastic Stress', *52nd US Rock Mechanics/Geomechanics Symposium*, June 2018.

Johnston, A. C., & Schweig, E. S. (1996) 'The enigma of the New Madrid earthquakes of 1811–1812.', *Annu. Rev. Earth Planet. Sci.*, 24(1), pp. 339-384.

Juanes, R., et al. (2016) 'Were the May 2012 Emilia-Romagna earthquakes induced? A coupled flow-geomechanics modeling assessment.', *Geophys. Res. Lett.*, 43(13), pp. 6891–6897.

Keranen, K. M., et al. (2014) 'Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection', *Science*, 345(6195), pp. 448-451.

Keranen, K. M., & Weingarten, M. (2018) 'Induced Seismicity', *Annu. Rev. Earth Planet. Sci.*, 46, pp. 149-174.

Kim, J., Tchelepi, H. A., & Juanes, R. (2011) 'Stability and convergence of sequential methods for coupled flow and geomechanics: Fixed-stress and fixed-strain splits', *Comput. Method. Appl. M.*, 200(13-16), pp. 1591-1606.

Kim, W. Y. (2013) ‘Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio.’, *J. Geophys. Res. Solid Earth*, 118(7), pp. 3506–3518.

Lapusta, N., & Liu, Y. (2009) ‘Three-dimensional boundary integral modeling of spontaneous earthquake sequences and aseismic slip’, *J. Geophys. Res.: Solid Earth*, 114(B9).

Lapusta, N., et al. (2000) ‘Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate-and state-dependent friction’, *J. Geophys. Res. Solid Earth*, 105(10), pp. 23765-23789.

Liu, D., & Duan, B. (2018) ‘Scenario Earthquake and Ground-Motion Simulations in North China: Effects of Heterogeneous Fault Stress and 3D Basin Structure’, *Bull. Seis. Soc. Am.*, 108(4), pp. 2148–2169.

Liu, D., Duan, B., & Luo, B. (2019) ‘EQsimu: A 3D finite element dynamic earthquake simulator for multicycle dynamic of geometrically complex faults governed by rate- and state-dependent friction’, *Geophys. J. Int.*, 220, pp. 598-609.

Manga, M., Wang, C. Y., & Shirzaei, M. (2016) ‘Increased stream discharge after the 3 September 2016 Mw 5.8 Pawnee, Oklahoma earthquake’, *Geophys. Res. Lett.*, 43, pp. 11588–11594.

Mikumo, K., Olsen, K. B., Fukuyama, E., & Yagi, Y. (2003) 'Stress-breakdown time and slip-weakening distance inferred from slip-velocity functions on earthquake faults', *Bull. Seismol. Soc. Am.*, 93(1), pp. 264-282.

Norbeck, J. H., & Horne, R. N. (2016) 'Evidence for a transient hydromechanical and frictional faulting response during the 2011 Mw 5.6 Prague, Oklahoma earthquake sequence', *J. Geophys. Res.: Solid Earth*, 121(12), pp. 8688-8705.

Ohnaka, M., & Yamshita, T. (1989) 'A cohesive zone model for dynamic shear faulting based on experimentally inferred constitutive relation and strong-motion source parameters', *J. Geophys. Res. Solid Earth*, 94(B4), pp. 4089-4104.

Oprsal, I., & Eisner, L. (2014) 'Cross-correlation – an objective tool to indicate induced seismicity', *Geophys. J. Int.*, 196(3), pp. 1536-1543.

Rudnicki, J. W. (1986) 'Fluid mass sources and point forces in linear elastic diffusive solids', *Mech. Mater.*, 5, pp. 383–393.

Scholz, C. H. (2002). *The mechanics of earthquakes and faulting*, 2nd edn, Cambridge University Press.

Segall, P., & Lu, S. (2015) ‘Injection-induced seismicity: poroelastic and earthquake nucleation effects.’, *J. Geophys. Res. Solid Earth*, 120(7), pp. 5082–5103.

Shapiro, S., Rentsch, S., & Rothert, E. (2005) ‘Characterization of hydraulic properties of rocks using probability of fluid-induced microearthquakes’, *Geophysics*, 70(2), F27—F34.

Shirzaei, M., et al. (2016) ‘Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas’, *Science*, 353, pp. 1416–1419.

Shirzaei, M., Manga, M., & Zhai, G. (2019) ‘Hydraulic Properties of injection formations constrained by surface deformation’, *Earth Planet. Sci. Lett.*, 515, pp. 125-134.

Szafranski, D., & Duan, B. (2018) ‘Integrating Poroelastic Effects of Wastewater Injection and Rupture Dynamics to Understand Induced Seismicity’, *Unconventional Resources Technology Conference*, July 2018, pp. 2557-2576.

Terzaghi, K. (1925) *Erdbaumechanik auf bodenphysikalischer Grundlage*. F. Deuticke.

Urpi, L., et al. (2016) ‘Dynamic simulation of CO2-injection-induced fault rupture with slip-rate dependent friction coefficient’, *Geomech. Energy Envir.* 7, pp. 47-65.

Wang, H. (2000) *Theory of Linear Poroelasticity With Applications to Geomechanics and Hydrogeology*. Princeton, N. J: Princeton Univ. Press.

Wang, L., et al. (2012) 'Stress-and aftershock-constrained joint inversions for coseismic and postseismic slip applied to the 2004 M6. 0 Parkfield earthquake', *J. Geophys. Res. Solid Earth*, 117(B7).

Zoback, M. D. (2010) *Reservoir geomechanics*. Cambridge University Press.

# 4. INTEGRATED WORKFLOW WITH NUMERICAL SIMULATION, MACHINE LEARNING REGRESSION AND BAYESIAN INVERSION – PRINCIPLES AND CASE STUDY OF INDUCED SEISMICITY

## 4.1. Abstract

The main objective of this work is to develop a new workflow integrating numerical simulations of fluid flow and dynamic rupture propagation, machine learning regression techniques and Bayesian inversion of subsurface model parameters. We present the theory behind each step as well as practical application of the proposed methodology on May 2012, $M_w$4.8 Timpson, TX, earthquake. Numerical simulations show that the triggering of the earthquake is related to the wastewater disposal with the dominant role of poroelastic stress changes. Dynamic rupture simulations allow us to reproduce the size of the earthquake. Using the results from a set of simulations, we form a training dataset and compare the performance of different regression algorithms. Random Forest, Bagging and K-Neighbors regression algorithms are the most promising and we use them in the inversion procedure to replace numerical simulations. We test multiple inversion scenarios and cross-validate them with the results of corresponding simulations. Because of this, we can constrain stress state and fault frictional parameters by matching the observations. We also discuss the limitations of the current methodology and propose further extensions in the future.

## 4.2. Introduction

In studies of either natural or induced earthquakes, using numerical simulations is a common practice, leading to better understanding of the physical processes associated with these seismic events (e.g. Shapiro et al., 2005; Juanes et al., 2016; Jin and Zoback, 2018; Duan, 2019). However, the biggest challenge in getting meaningful results from the simulation of a real case study is to have a well-constrained values of all or most of the model parameters (e.g. Weng and Yang, 2018; Shirzaei et al., 2019; Szafranski and Duan, 2020). Therefore, there is a trade-off between using less complex modeling with robustly known inputs but very simplified medium and many uncaptured phenomena, and sophisticated modeling but with some poorly-constrained parameters.

In this work we focus on induced seismicity caused by water injection into the subsurface, which is or was a problem in certain areas across the United States, including Oklahoma (e.g. Yeck et al., 2017; Goebel et al., 2017; Johann et al., 2018), Texas (e.g. Frohlich et al., 2011; Frohlich et al., 2014; Skoumal et al., 2020) and Ohio (e.g. Kim, 2013; Skoumal et al., 2015). One of the trends in studies addressing this issue is to use numerical modeling mimicking fluid flow in the medium in order to evaluate the influence of the injection on the fault slip potential (e.g. Fan et al., 2016). The complexity of the modeling can vary from simple pressure diffusion in homogenous medium (e.g. King et al., 2016), through implementing poroelastic stress changes (e.g. Deng et al., 2016), to including the rate-and-state friction (e.g. Alghannam and Juanes, 2020) or dynamic rupture propagation (e.g. Jin and Zoback 2018; Szafranski and Duan, 2020).

When conducting numerical studies, it is important to identify the parameters that have a major impact on the final solution and then to assign the best possible values for the modeling. However, sometimes it is very difficult or impossible to directly measure these model parameters. In fluid flow modeling, one of such parameters is permeability of the rock formations (e.g. Fan et al., 2016; Szafranski and Duan, 2018; Shirzaei et al., 2019). In dynamic rupture modeling, stress state and frictional parameters are among these parameters (e.g. Gallovic et al., 2019; Szafranski and Duan, 2020).

One of the possible solutions for this issue is to perform the inversion of the unknown or poorly known parameters. There are only limited attempts in performing these inversions. In their study of $M_w$4.8, May 2012, Timpson (TX) earthquake, Shirzaei et al. (2019) inverted for the hydraulic diffusivity of the injection formations using the InSAR data (surface deformations) and poroelastic 3D model. They used a grid-search method of inversion, which is simple and easy to apply, but is not efficient in case of a multi-parameter inversion, and has limited capability of providing uncertainty estimate. Gallovic et al. (2019) performed the inversion of fault frictional parameters based on the data for $M_w$6.2, August 2016, Amatrice (Italy) tectonic earthquake. They implemented Bayesian framework with Monte Carlo inversion algorithm, which gives a good estimation of the uncertainty and works well also for multi-parameter inversion. However, Monte Carlo type algorithms, even though good for inversion of many model parameters, still require running thousands of simulations to generate statistically reliable sample (e.g. Fichtner and Simute, 2018).

In this study, we show that one of the possible solutions to optimize the inversion procedure can be combination of using numerical modeling and machine learning supervised regression algorithms. The inputs and outputs from numerical modeling serve as a training dataset for tuning the regression algorithms. Ahamed and Daub (2019) used a similar optimization idea combining dynamic rupture simulations and machine learning classification algorithms, Random Forest (RF; e.g. Liaw and Wiener, 2002) and Artificial Neural Network (ANN; e.g. Hassoun, 1995), in order to predict whether the dynamic rupture can or cannot break the heterogeneities on the fault.

We combine numerical modeling, Bayesian inversion and machine learning supervised regression into a single framework and apply it in the case study of $M_w$ 4.8, May 2012, Timpson (TX) earthquake (Frohlich et al., 2014). Extended dataset from our previous study (Szafranski and Duan, 2020) is used to train different regression algorithms, which serve as a sample generation tool in our Bayesian inversion. We compare performance of multiple popular regression algorithms and select ones that work best in this particular problem. Then, we invert for several model parameters along with uncertainty evaluation. Additionally, we compare performance of different inversion setups and propose a method to evaluate the quality of the inversion results.

Even though we show a specific case study of an earthquake induced by the wastewater disposal, the presented methodology can be adapted to other types of problems involving

time-consuming modeling of subsurface processes, both in a regional scale (e.g., large earthquakes) and local/small scale (e.g., reservoir modeling).

### 4.3. Methods

In this paper, we develop a workflow to study fluid-injection induced seismicity, which integrates numerical modeling, data analysis and application of supervised machine learning regression techniques, and Bayesian inversion of selected model parameters, along with the evaluation of the inversion performance. Figure 4.1 presents a graphical representation of the proposed workflow.

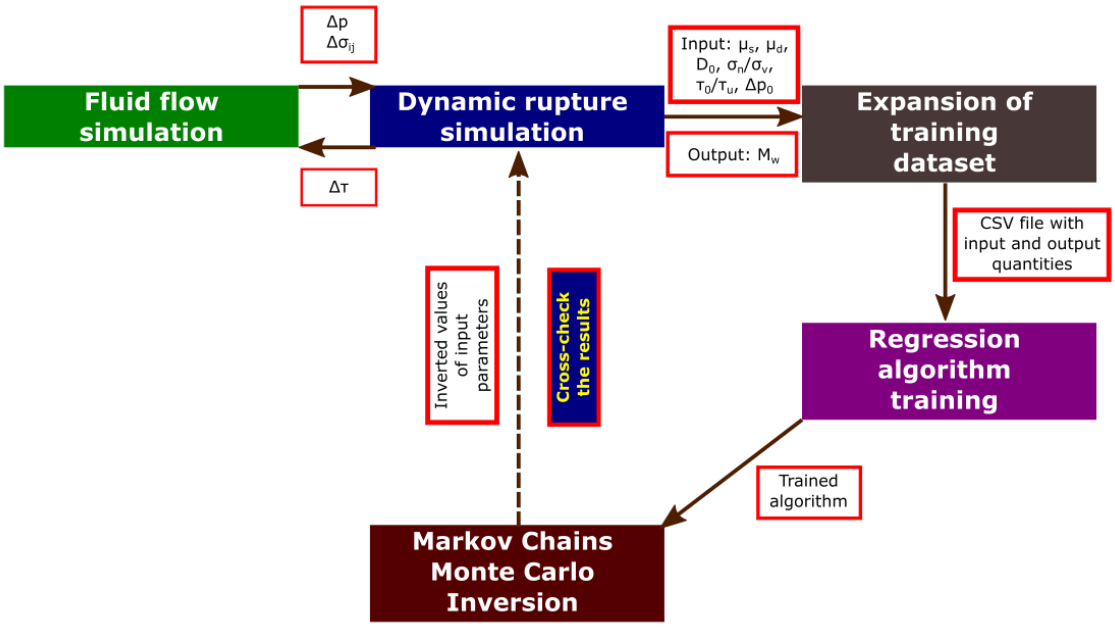


**Figure 4.1 Proposed workflow with major elements of the methodology and main input/output data passed between different modules.**

### 4.3.1. Fluid flow in a linear poroelastic model

The first component of the methodology involves numerical modeling of fluid flow in a deformable porous medium. Using the theory of poroelasticity, we simulate perturbations of pore pressure and stress caused by hydromechanical coupling during fluid movement within the medium. To obtain desired quantities we solve a system of 7 partial differential equations. For fluid flow, we use the diffusion equation expressed in terms of pore pressure and stress (e.g. Wang, 2000):

$$\frac{\alpha}{KB}\left[\frac{B}{3}\frac{\partial \sigma_{ss}}{\partial t}+\frac{\partial p}{\partial t}\right]-\frac{1}{\mu}(\nabla k\cdot\nabla p+k\nabla^2 p)=Q \tag{4.1}$$

where $K$ is a bulk modulus, $B$ represents Skempton's coefficient, $\alpha$ is Biot-Willis coefficient, $p$ is pore pressure, $\sigma_{kk}$ denotes a trace of the stress tensor, $k$ is an intrinsic permeability, $\mu$ is fluid dynamic viscosity and $Q$ represents a volume of fluid per unit bulk volume per unit time (external source of fluid). We define Biot-Willis coefficient following e.g., Segall and Lu (2015):

$$\alpha=\frac{3(\nu_u-\nu)}{B(1+\nu_u)(1-2\nu)}, \tag{4.2}$$

where $\nu$ and $\nu_u$ are drained and undrained Poisson's ratios.

We calculate stress caused by fluid injection using a set of Beltrami-Mitchell equations (e.g. Wang, 2000):

$$\nabla^2\sigma_{ij}+\frac{1}{1+\nu}\frac{\partial^2\sigma_{ss}}{\partial x_i\partial x_j}+2\eta\left[\frac{1-\nu}{1+\nu}\frac{\partial^2 p}{\partial x_i\partial x_j}+\delta_{ij}\nabla^2 p\right]=-\frac{\nu}{1-\nu}\delta_{ij}\,\nabla\cdot\boldsymbol{F}-\frac{\partial F_i}{\partial x_j}-\frac{\partial F_j}{\partial x_i} \tag{4.3}$$

where $\sigma_{ij}$ is a stress tensor, $\delta_{ij}$ denotes the Kronecker delta, $\eta$ is the poroelastic stress coefficient, and $\boldsymbol{F}$ represents the external body force. The poroelastic stress coefficient is defined as (Wang, 2000):

$$\eta = \frac{1-2\nu}{2(1-\nu)}\alpha. \tag{4.4}$$

In our implementation, we assume that all the terms on the right-hand side of equations (4.3) are equal to zero, since we are mainly interested in values of pore pressure and stress perturbations. This assumption greatly simplifies the calculations.

To solve the system of equations (4.1) and (4.2) we use the in-house finite-difference code called PyFlyFlow. The detailed description and verification of the code can be found in our previous publication (Szafranski and Duan, 2020).

### 4.3.2. Fault failure and dynamic rupture propagation

The second component of our workflow is the modeling of dynamic rupture propagation. To capture rupture initiation, propagation and arrest, we solve the elastodynamic equation:

$$\nabla \cdot \boldsymbol{\sigma} + \rho_b \boldsymbol{f_b} = \rho_b \frac{\partial \boldsymbol{v}}{\partial t}, \tag{4.5}$$

where $\rho_b$ is bulk density of the rock, $\boldsymbol{f_b}$ denotes body force vector, and $\boldsymbol{v}$ represents the velocity vector. We use traction-at-split-node scheme in the formulation of Day et al. (2005) as our fault boundary condition. To simulate friction on the fault, we implement the linear slip weakening law (e.g. Ida, 1972):

$$\mu(s) = \begin{cases} \mu_s - (\mu_s - \mu_d)\frac{|s|}{D_0}, |s| < D_0 \\ \mu_d, |s| \geq D_0 \end{cases}, \tag{4.6}$$

where $|s|$ is the slip distance, $\mu(s)$ represents slip-dependent friction coefficient (not fluid dynamic viscosity as above), $D_0$ is critical slip distance, and $\mu_s$ and $\mu_d$ are static and

dynamic friction coefficient values respectively. We also use the Mohr-Coulomb failure criterion as the fault failure criterion:

$$\tau \leq c + \mu\sigma_n', \tag{4.7}$$

where $\tau$ is shear stress on the fault plane, $c$ is the cohesion of the fault plane and $\sigma_n'$ is effective normal stress. In theory of poroelasticity it has the following form:

$$\sigma_n' = \sigma_n - \alpha p. \tag{4.8}$$

To solve equation (4.5) we use also in-house, finite-element code called EQdyna. It has been described and applied in multiple previous publications (e.g. Duan and Oglesby, 2006; Duan, 2012; Liu and Duan, 2018; Szafranski and Duan, 2018; Szafranski and Duan, 2020). It was also successfully validated within the Southern California Earthquake Center code verification exercise (Harris et al., 2009; 2011; 2018).

### 4.3.3. Dataset evaluation and machine learning regression algorithms

As shown in Figure 1, the third element of the workflow involves building the training dataset and data analysis before training the regression algorithms. Beside our observations from previous work (Szafranski and Duan, 2020) and visual inspection of input and output data plots, we also calculate Pearson's cross-correlation coefficients $R_{ij}$ between different input features to evaluate whether some of them can be potentially omitted in the training process:

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}}. \tag{4.9}$$

In equation (4.9) $C$ represents the covariance matrix. Number of features can be reduced if there are some of them which show high correlation with one another.

The next step in the workflow involves the training of the regression algorithm. In this study, we select a suite of 8 different algorithms. We tune the hyperparameters of each algorithm to reduce bias coming from the arbitrary selection. The first algorithm is Ridge Regression (e.g. Marquard and Snee, 1975; McDonald, 2009). It is an extension of the ordinary least-squares regression, in which we penalize high weights of the coefficients $\boldsymbol{w}$ in the linear model, by minimizing the function:

$$min(w): \|Xw - y\|_2^2 + \beta\|w\|_2^2, \tag{4.10}$$

where $\beta$ is a complexity parameter, $X$ is an array of input features values and $y$ is an output feature value. With higher $\beta$, there is a larger shrinkage of weights. The second algorithm is Lasso Regression (e.g. Tibshirani, 1996), which is very similar to Ridge, but instead of $L_2$ norm in penalty term, it uses $L_1$ norm in minimizing the function:

$$min(w): \|Xw - y\|_2^2 + \beta\|w\|_1. \tag{4.11}$$

As a result, unlike Ridge regression where $\alpha$ is always positive, Lasso Regression can reduce weights $\boldsymbol{w}$ to 0, effectively eliminating certain features from the model. The key in both regressions is to select the correct value of $\alpha$, which can be done with the grid-search method.

The third method that we test is Theil-Sen Regression (or Theil-Sen Estimator; Theil, 1950; Sen, 1968). It belongs to so-called outlier-robust methods. In Theil-Sen Estimator, we find least-squares solutions for subsets of samples and calculate spatial median for all

the solutions. The method is robust to multivariate outliers, but the robustness diminishes with increasing number of dimensions.

We also test several other methods, which combine results of a certain base regressors, and therefore are called ensemble methods. First of them is Random Forest Regression (e.g. Breiman, 2001; Liaw and Wiener, 2002), where the base estimator is a decision tree. We run multiple decision trees at once and we take the average of these results. Averaging combined with introduced randomness reduce the variance. With Random Forest Regressor, it is possible to perform a feature selection.

Another averaging, ensemble method we test is Bagging Regression (e.g. Breiman, 1996). In this meta-estimator, we fit a base regressor on a subsets of dataset, and in the end aggregate all results by averaging or voting procedure. Even though any basic estimator might be used, we apply here again the Decision Tree regressor.
The last ensemble method that we use is AdaBoost Regressor (e.g. Freund and Schapire, 1997; Kummer and Najjaran, 2014). Instead of averaging, it utilizes a boosting technique. It starts with fitting the base regressor on the original dataset and estimates the error of the prediction. Then, based on the error, it adjusts weights for different instances and performs fitting on the copy of the original dataset. In further iterations, the estimator focuses on the poorly predicted (more difficult) cases. In the end, the final prediction is obtained via majority vote or the sum of all predictions.

The next of algorithm that we test is K-Neighbors Regressor (e.g. Nguyen et al., 2016). The prediction of the parameter output value of a point located within input parameter space is estimated based on the values of K-nearest neighbors of that point. We use here distance-based weights, meaning that closer neighbor has bigger influence than further neighbor. To speed up the algorithm, instead of brute-force calculation of distances between all pairs of points, we use Ball Tree algorithm (e.g. Omohundro, 1989), which divides the parameter space into hyper-spheres, calculates the distance of one selected point to the boundary of the hyper-sphere, in which it is located, and assigns the same distance to all the points within that hyper-sphere.

The last algorithm we test in this study is Multi-Layer Perceptron (MLP) Regressor (e.g. Murtagh, 1991). This regression technique uses a concept of neural networks. The algorithm learns the non-linear function:

$$f(\cdot): R^m \rightarrow R^o, \tag{4.12}$$

where *m* is number of input dimensions and *o* is number of output dimensions. The neural network is constructed in the way that it consists of layers and neurons. In each neural network there are input and output layers, plus at least one hidden layer in between. Within hidden layer the values passed from the previous layer are transformed into weighted linear summation, followed by the use of non-linear activation function:

$$g(\cdot): R \rightarrow R. \tag{4.13}$$

In the output layer the activation function is the identity function (effectively no activation). In other layers function from equation (14) had a form of rectified linear unit function:

$$g(z) = \begin{cases} 0\,, z < 0 \\ z\,, z \geq 0 \end{cases}. \tag{4.14}$$

The neural network is trained by backpropagation technique with the regression loss function being a square error:

$$Loss(\hat{y}, y, w) = \frac{1}{2}\|\hat{y} - y\|_2^2 + \frac{\beta}{2}\|w\|_2^2, \tag{4.15}$$

where $\boldsymbol{w}$ represents weights within the layers, $y$ are actual values of samples and $\hat{y}$ are predicted values of the sample. As a network optimization algorithm, we use Adam Gradient Descent algorithm (Kingma and Ba, 2014), which is an extension of the Stochastic Gradient Descent. Unlike Stochastic Gradient Descent, where learning rate is a single value for all weights updated and does not change during learning, Adam Gradient Descent adapts learning rates for different parameters using first and second moments of gradients. To properly use the MLP regressor it is necessary to tune its hyperparameters such as number of hidden layers, number of neurons per layer or number of iterations in training.

To compare and select best algorithms for further study we mix quantitative and qualitative approaches. To quantify algorithm performance, we calculate the coefficient of determination ($R^2$), explained variance score and mean squared error (MSE) for both training and test dataset predictions. We calculate coefficient of determination as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \tag{4.16}$$

where $\bar{y}$ is the average of the sample. The explained variance score is calculated as:

$$Explained\ variance = 1 - \frac{Var\{y - \hat{y}\}}{Var\{y\}}, \tag{4.17}$$

where *Var{*}* means variance of the given quantity. Finally, the mean squared error is defined as:

$$MSE = \frac{1}{N_{samp}} \sum_{i=0}^{N_{samp}-1} (y_i - \hat{y}_i)^2. \tag{4.18}$$

Beside quantitative measures, we also base our judgement on the qualitative evaluation of the cross-plots between moment magnitudes from numerical simulation and regression predictions.

### 4.3.4. Bayesian inversion of medium properties

After training of the regression algorithm, we move on to inversion part of the workflow. In this work we use statistical inversion, with Bayesian approach in particular. In Bayesian approach, we construct the probability function for particular parameter using a priori knowledge, and during inversion process we update the probability as new information becomes available.

A subclass of Bayesian are Markov Chains Monte Carlo (MCMC) methods (e.g. Brooks, 1998). In MCMC we perform Monte Carlo simulations, which have Markov property, i.e., the next simulation/state is dependent only on the current state in the chain of simulations (lack of memory). In these methods we also use acceptance-rejection

sampling, which serves as a guidance whether simulations are going into right direction (estimation of posterior distribution) or not by accepting or discarding results of a given simulation in construction of the posterior distribution. In this work we implement a relatively simple but powerful Metropolis-Hastings algorithm (Metropolis, 1953; Hastings, 1970). Figure 4.2 shows how the Metropolis-Hastings algorithm has been implemented in this study.

The inversion procedure itself requires tuning of several parameters including number of samples drawn or the maximum length of the step from the current state to the next one. To evaluate the performance of a given inversion, we run a numerical simulation, extract the moment magnitude of an earthquake, which served here as an output parameter in training of regression algorithms, and then calculate a measure that we propose and call Moment Magnitude Stacked Difference:

$$Stacked\ difference = \left|M_w^{ML} - M_w^{sim}\right| + \left|M_w^{ML} - M_w^{reg}\right|, \qquad (4.19)$$

where $M_w{}^{ML}$, $M_w{}^{sim}$ and $M_w{}^{reg}$ are moment magnitudes from machine learning based inversion, numerical simulation of dynamic rupture propagation and registered (actual) earthquake. Since we replace running actual simulations for each Monte Carlo sample, we not only compare "the best" result from given inversion to actual reported magnitude, but also the consistency between the machine learning statistical prediction and deterministic prediction from the numerical simulation. The lower the stacked difference, the better.

Finally, it is important to define what the result of inversion is. We obtain the probability distributions of inverted parameters, which in classical understanding of the result might seem to be a bit cryptic. However, having probability distribution we can not only have a single, “best” value of the parameter, but also evaluate whether the parameter can actually be resolved and estimate the uncertainty of the solution. For inversion evaluation purpose, from each posterior distribution, we extract both average and highest probability solutions. Average works well for distributions which are unimodal and relatively symmetrical (such as normal distribution). If the distribution is strongly asymmetric or multimodal, average of the distribution would provide a meaningless information. Highest probability at the same time works with asymmetric and multimodal distributions. However, it might give a misleading result when the posterior distribution has a “flat top” (almost uniform interval) or is “spiky” due to low number of accepted answers.

The results of simulations used in cross-check can later be included in extension of training dataset and the procedure might be iteratively repeated until satisfactory level of success is reached.

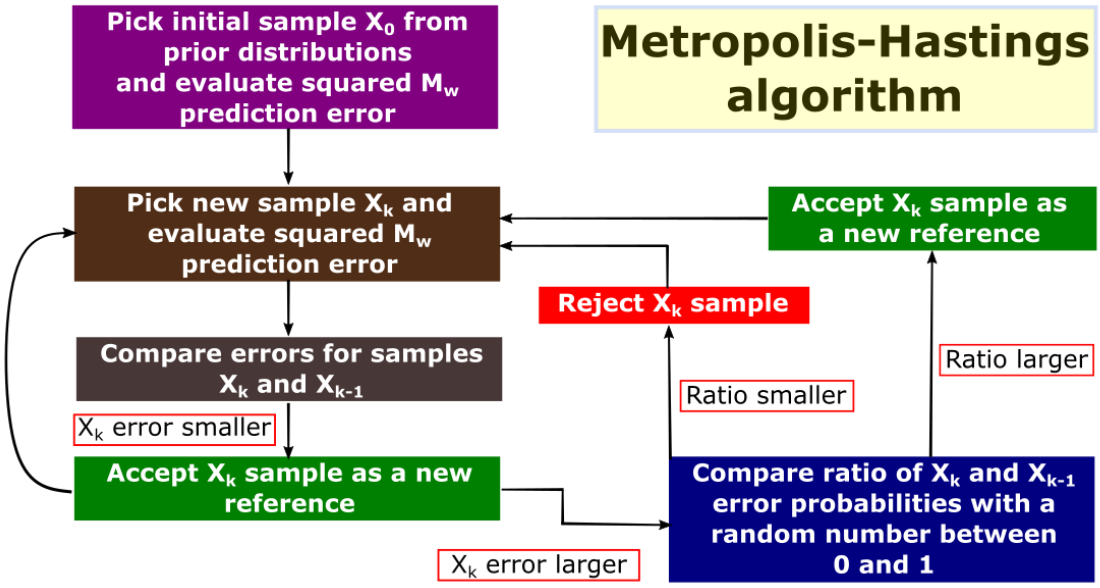


**Figure 4.2 Flowchart of the Metropolis-Hasting algorithm.**

### 4.4. Models

This study focuses on the application of the described methodology to the case of May 2012, $M_w$4.8 Timpson (TX) earthquake. As shown in multiple studies (e.g. Frohlich et al, 2014; Fan et al., 2016), the mainshock was most probably caused by the wastewater disposal at nearby Class II Injection wells. Figure 4.3 shows the location of our study site, as well as the geometry of the simulation model. The ruptured fault has strike of 330° and dip of 63.5° SW (e.g. Fan et al., 2016; Szafranski and Duan, 2020). The source mechanism inversion indicates predominantly strike-slip movement on the fault during the mainshock.

For fluid flow simulations, we use the same model as in Szafranski and Duan (2020) (see Table 4.1), which is modified model from Shirzaei et al. (2019), where authors

constrained the permeabilities of the injection layer. The fault zone is treated in the simulation as a part of the medium with properties different from the surrounding formation, including K = 3 GPa; $\nu = 0.25$; $\nu_u = 0.45$; $k = 10\times10^{-16}$ $m^2$; B = 0.75.

In construction of models for dynamic rupture simulation we also use similar approach as in Szafranski and Duan (2020), where we specify three parameters describing initial (before injection) stress state in the medium: normal-to-vertical stress ratio ($S_n/S_v$), shear-to-yield stress ratio ($\tau_0/\tau_u$) and deviation of pore pressure from the main background trend (e.g. from hydrostatic pressure) as the fraction of that pore pressure ($\Delta p_b^{fcrt}$). Another three parameters describe fault frictional behavior: static friction coefficient ($\mu_s$), dynamic friction coefficient ($\mu_d$) and critical slip distance ($D_0$). To mimic the random heterogeneity within the fault zone, for each of these parameters we define its normal distribution in terms of mean value and standard deviation. Additionally, we specify whether the main background trend is simple hydrostatic pore pressure or the fault zone has uniform effective normal stress (i.e., it is overpressured). Finally, we allow the smoothing of the random heterogeneities if we want to represent larger scale asperities on the fault plane.

In the previous study (Szafranski and Duan, 2020) as well as in this work, we run many dynamic rupture simulations. The main output is the moment magnitude of the mainshock. For the purpose of training regression algorithms, we form dataset, in which the model parameters are training inputs. Figure 4.4 shows the statistical distributions of

different stress state and frictional parameters within smaller (original) and larger (extended) datasets.

**Table 4.1 Modified Shirzaei et al. (2019) hydro-geomechanical model of the medium in Timpson (TX) case study**

| Depth of the layer top [m] | $\rho_b$ [kg/m$^3$] | $K$ [GPa] | $\nu$ | $\nu_u$ | $B$ | $k$ [m$^2$] |
|---|---|---|---|---|---|---|
| **0** | 2293 | 18.5 | 0.33 | 0.45 | 0.80 | $6.80 \cdot 10^{-16}$ |
| **500** | 2405 | 24.8 | 0.33 | 0.45 | 0.80 | $1.00 \cdot 10^{-14}$ |
| **1000** | 2470 | 32.1 | 0.23 | 0.45 | 0.80 | $4.51 \cdot 10^{-16}$ |
| **1500** | 2572 | 40.6 | 0.23 | 0.45 | 0.90 | $1.00 \cdot 10^{-14}$ |
| **2000** | 2622 | 52.0 | 0.25 | 0.45 | 0.75 | $3.00 \cdot 10^{-16}$ |
| **3000** | 2650 | 53.5 | 0.25 | 0.45 | 0.75 | $2.92 \cdot 10^{-17}$ |

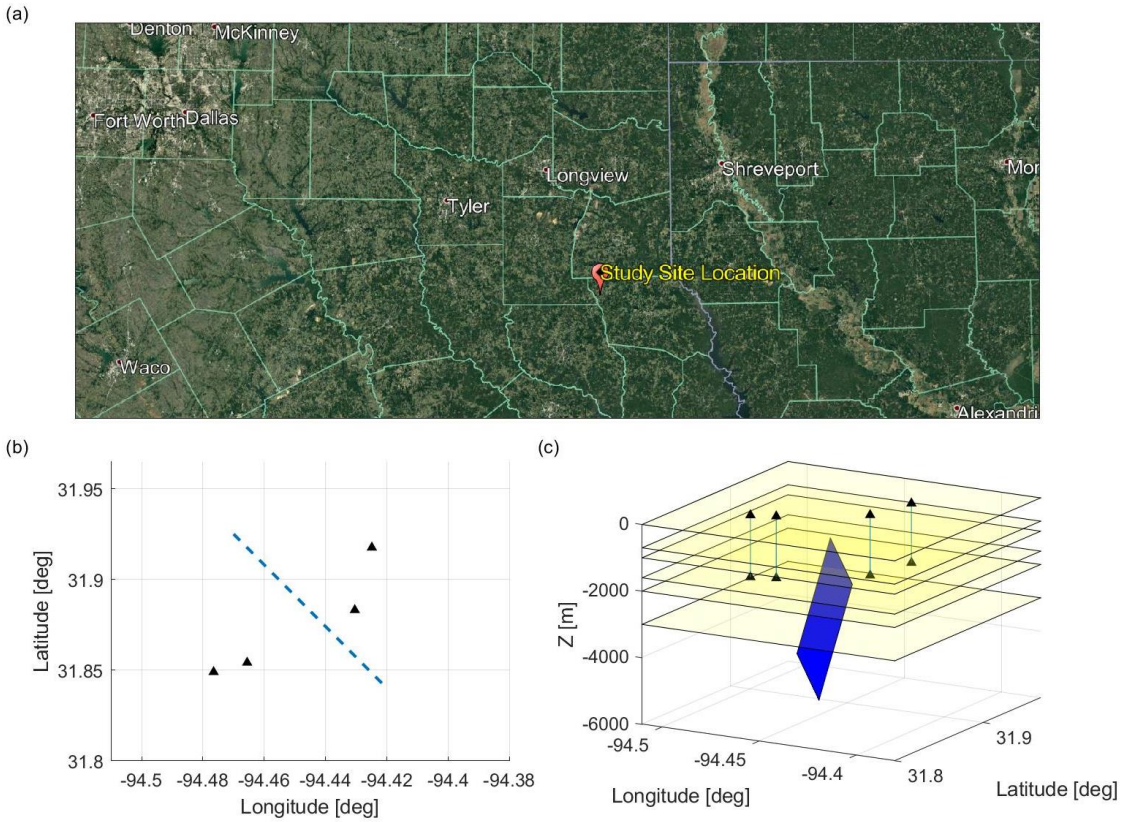


**Figure 4.3 (a) Google Earth aerial photo with marked Timpson (TX) earthquake study site location; (b) map view showing the geometry of the system, including fault trace (blue dashed line) and injection wells (black triangles); (c) the slanted view at the geometry of the system including fault plane (blue plane), layers tops (transparent yellow planes) and injection wells (black triangles connected with blue lines).**

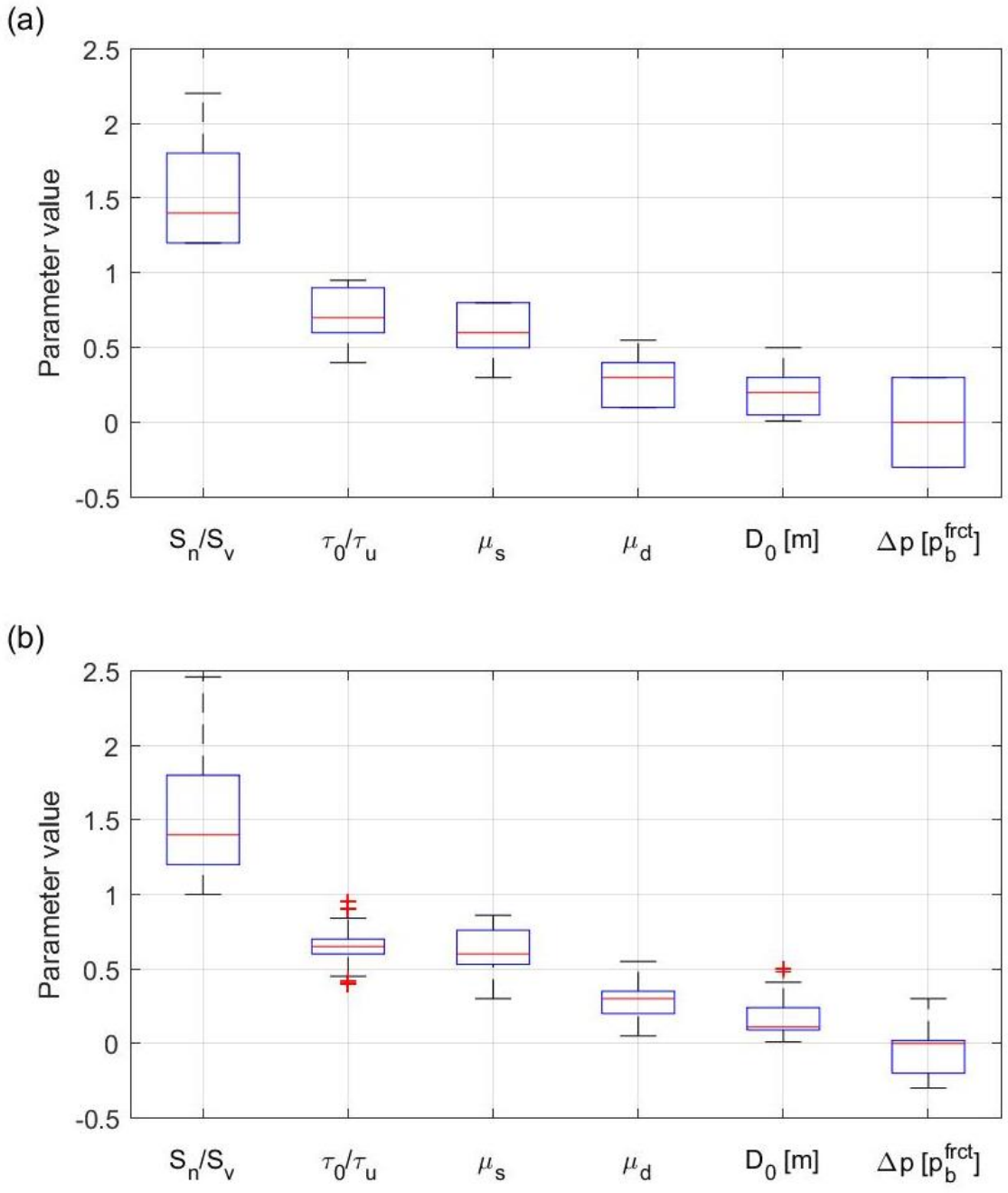


**Figure 4.4 Box-plots presenting the content of the original/smaller (a) and extended (b) dataset (input parameter values). Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25th and 75th percentiles respectively and whiskers show the dataset boundaries after excluding statistical outliers.**

## 4.5. Results

### 4.5.1. Numerical simulations

The first and very important part of the results are the distributions of the pore pressure and stress perturbations on the fault plane due to fluid injection (see Figure 4.5). It is apparent that normal stress perturbations are dominant over pore pressure and shear stress perturbations. The peak pore pressure perturbations are ~4.7 MPa, normal stress ~8.1 MPa and shear stress ~4.8 MPa. They are mostly localized in the shallow portion of the fault, slightly towards SE from the center. These values are similar to the ones reported in Fan et al. (2016). For more information, one can refer to Szafranski and Duan (2020).

The next part in our workflow is dynamic rupture modeling. The output we get from a single simulation includes the slip and slip rate distribution on the fault, the fault rupture time distribution, the size of the initiation patch and final rupture area, seismic moment and seismic magnitude (as derivatives of total rupture area and average slip within it) and shear stress drop on the fault plane. One can also analyze generated synthetic seismograms. However, in this study we do not use them since comparison to registered seismograms would require better velocity and attenuation models of the subsurface.

Figure 4.6 shows an example of the dynamic rupture simulation results that produces a synthetic earthquake matching the Timpson mainshock with moment magnitude ($M_w$ 4.8). First important observation is that not the entire fault is ruptured during this event

and it is important to distinguish between the rupture area and the fault segment area. Secondly, the majority of the rupture occurs within first 0.5 second after which the rupture slows down and is finally halted shortly after ~1 second of duration. Third, the peak slip occurs in the shallow, central portion of the fault segment and is equal to ~0.2 m. Finally, we see that on majority of the fault the stress drop is within the range of several MPa, peaking at ~10 MPa, and on the edge of rupture area there is a small stress increase which can later lead to further seismic activity.

Since moment magnitude is a relatively simple quantity to be matched by the simulations, we observe the equivalency between solutions from multiple models. We assume that good match is $M_w$ 4.8 ± 0.2 and making that assumption in the smaller dataset we have 16 simulations which match the condition and in the extended dataset we have 94 matching simulation results. The statistics of the matching models from the extended dataset are presented in Figure 4.7. Although, we want to establish the set of model parameters producing the desired size of an earthquake, to properly train regression algorithm, we include population of inputs that produce as big as possible range of different magnitudes, so that the algorithm can also predict non-matching outputs by interpolation rather than extrapolation, which in machine learning practices should be avoided if possible.

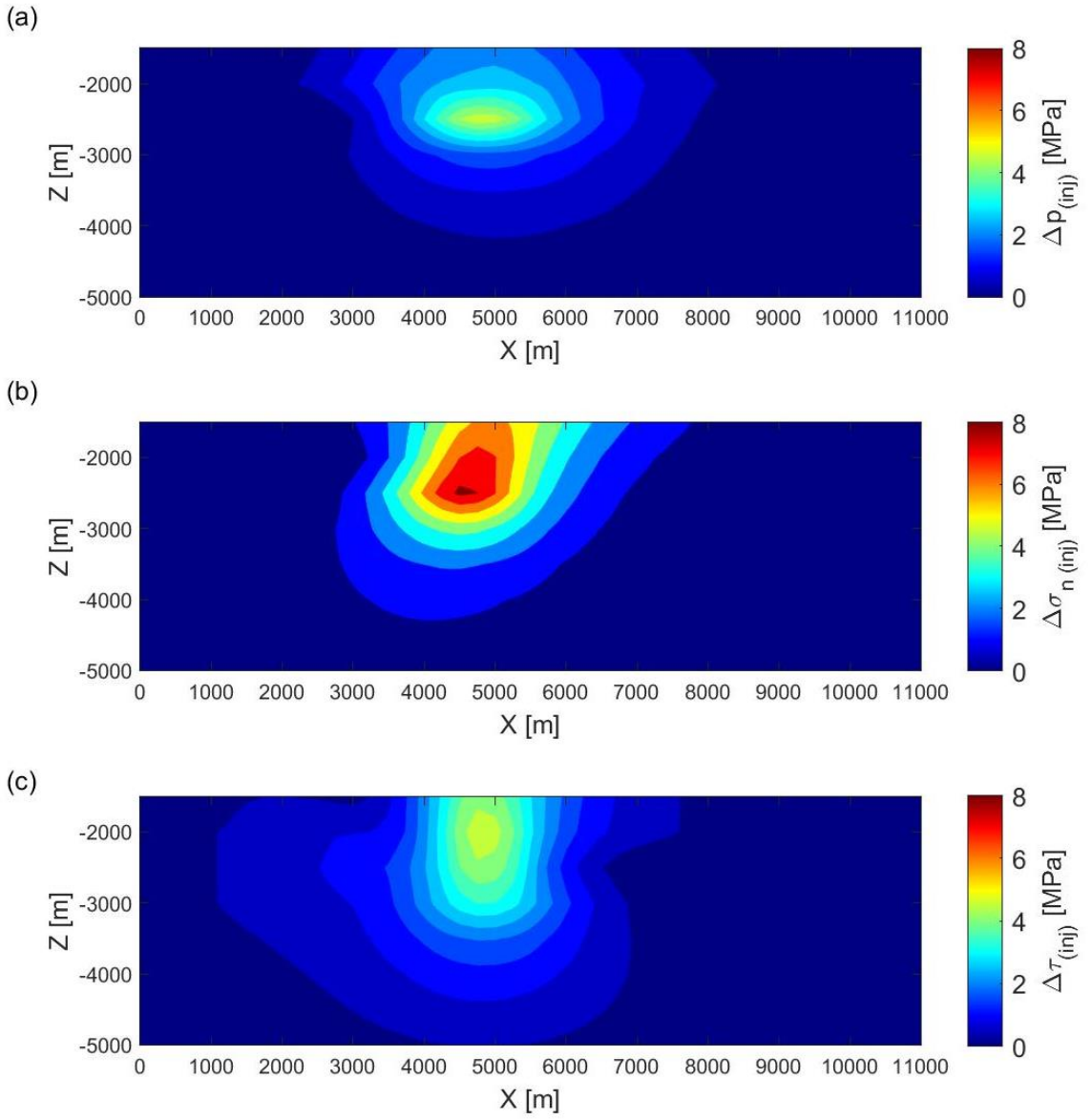


**Figure 4.5 Distribution of the (a) pore pressure, (b) normal stress and (c) shear stress perturbations on the fault plane (SE-NW from left to right), on the day of the Timpson earthquake mainshock simulated for the injection of 4 disposal wells, for model in Table 4.1.**

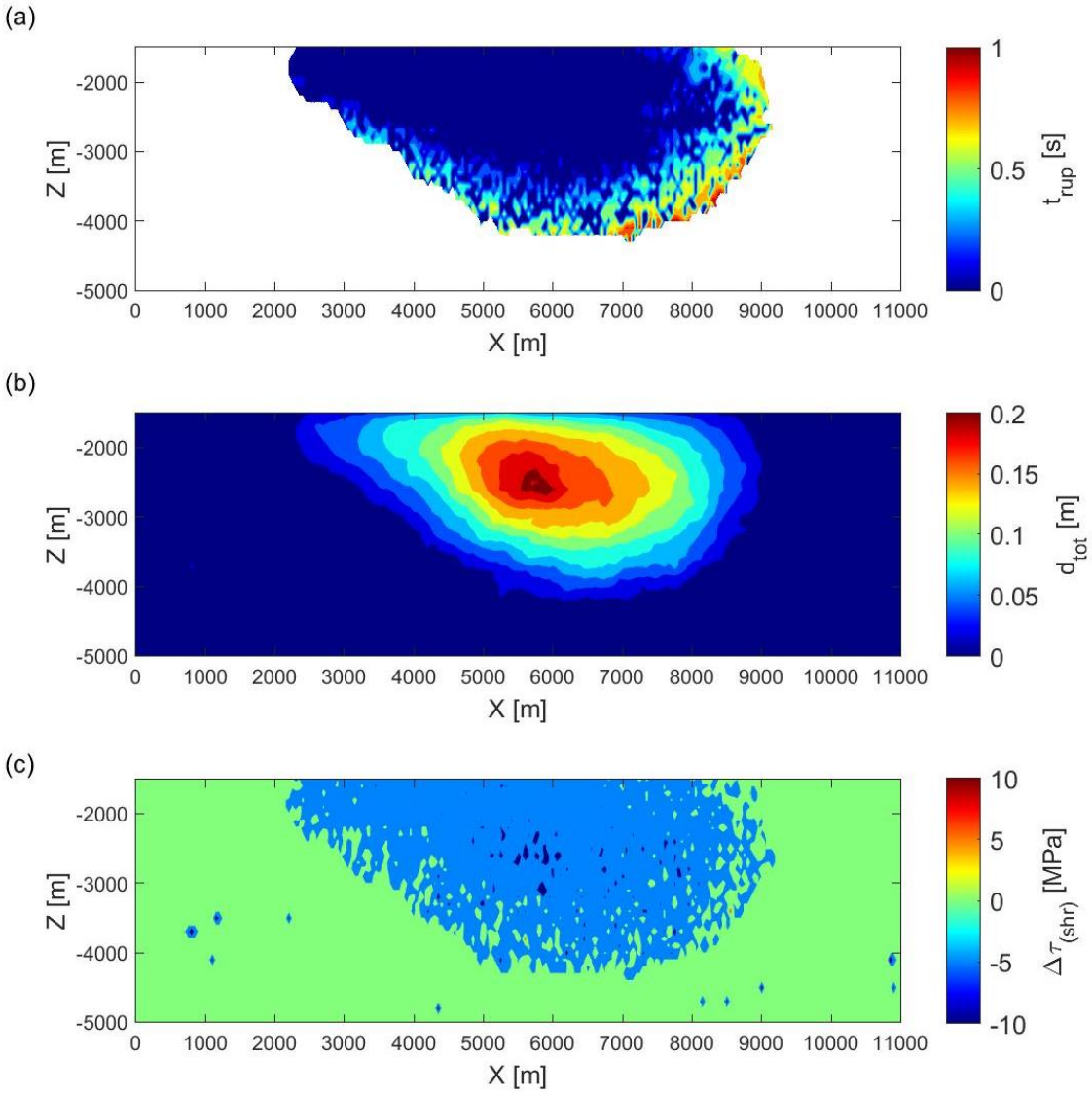


**Figure 4.6 Distribution of (a) rupture time, (b) total slip and (c) stress drop on the fault plane (SE-NW from left to right) for the time of the Timpson earthquake mainshock for one of the models included in the training dataset. Simulated magnitude is Mw4.8 which is one of the possible exact matches of reported $M_w$.**

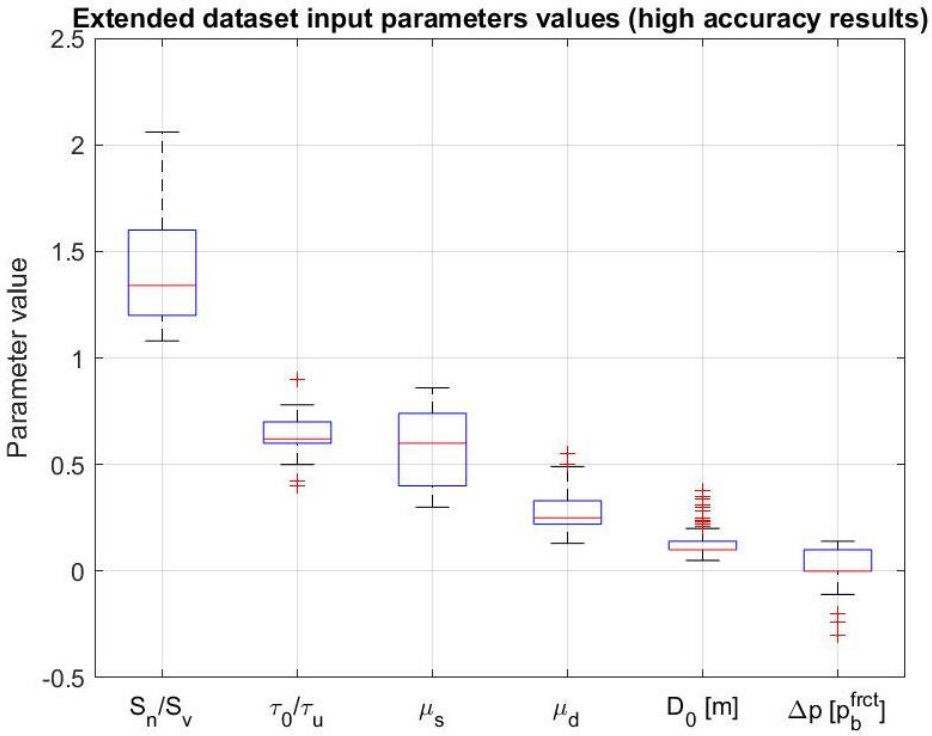


**Figure 4.7 Box-plots presenting the ranges of input parameter values in extended dataset which provided closely matching Mw of 4.8 +/- 0.2. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25th and 75th percentiles respectively and whiskers show the dataset subset boundaries after excluding statistical outliers.**

### 4.5.2. Tuning of regression algorithms

Next element of the workflow after building dataset for machine learning is training and tuning of the regression algorithms. However, to maintain the balanced weighting of different input parameters in optimizing algorithms, it is usually necessary to scale or normalize the data. Even though our input parameters are not vastly different in their values ranges, we could still expect $S_n/S_v$ ratio to play the dominant role without scaling the data. In this study we normalize the dataset, which scales each input vector to unit norm.

As mentioned in the *Methods* section, we select 8 different regression algorithms and for each of them we perform optimization of selected, important hyperparameters specific for a given method. For example, in Ridge regression we optimize the value of parameter $\beta$, in Random Forest we optimize the number of estimators (decision trees), while in Multi-Layer Perceptron we focus on number of hidden layers, number of neurons per layer and parameter $\beta$. In order to do that, we do the grid-search across a discrete and constrained parameter space for each algorithm. Then we calculate $R^2$ based on the performance on training dataset and test dataset. Finally, we pick optimum values of hyperparameters looking at the mean value of $R^2$ for both training and test datasets, which should assure the balance between in-sample and out-of-sample performance and prevent overfitting.

To further evaluate and compare different regression algorithms, beside the coefficient of determination (equation 4.16) we also calculate the explained variance score (equation 4.17) and the mean squared error (equation 4.18) for the optimized algorithms. Generally, the higher the $R^2$ or explained variance score, and the lower the mean squared error, the better is the performance of the regression algorithm. Figures 4.8 and 4.9 present the results of these calculations for the original and extended datasets, respectively. First major observation is that Ridge, Lasso and Theil-Sen regression algorithms significantly underperform and score poorly even after optimization. The best performers in this comparison are Random Forest and Bagging algorithms. MLP, K-

Neighbors and AdaBoost got quite high scores as well, though they perform a bit worse on test subsets, especially for the extended dataset. However, qualitative look at the tuning results and $M_w^{sim}$ vs $M_w^{ML}$ cross-plots force us to eliminate MLP and AdaBoost from further consideration. Especially in AdaBoost, the magnitudes are overestimated for lower magnitudes and underestimated for higher magnitudes which falsely gives high scores in quantitative measures. K-Neighbors algorithm on the other hand seems more promising giving sensible and good match on the cross-plot and being much faster than Random Forest or Bagging in prediction (which plays an important role in the inversion procedure). Figure 4.10 gives better insight in the tuning of Random Forest, K-Neighbors and Bagging algorithms. The biggest variation of scores dependent on hyperparameter value is for the K-Neighbors algorithm, favoring rather smaller number of the nearest neighbors involved in prediction. For Random Forest and bagging the variation is much smaller and technically any number of estimators would work well. However, the higher the number of estimators, the longer it takes to make a prediction. Figure 4.11 shows the $M_w^{sim}$ vs $M_w^{ML}$ cross-plots for these three algorithms. Beside high magnitudes in the dataset ($M_w$~5.8 and higher), where some magnitudes are underestimated by all regression algorithms, the fit is very good. The underperformance in the upper range can be the result of the limit of maximum magnitude that can be observed on the given segment of the fault.

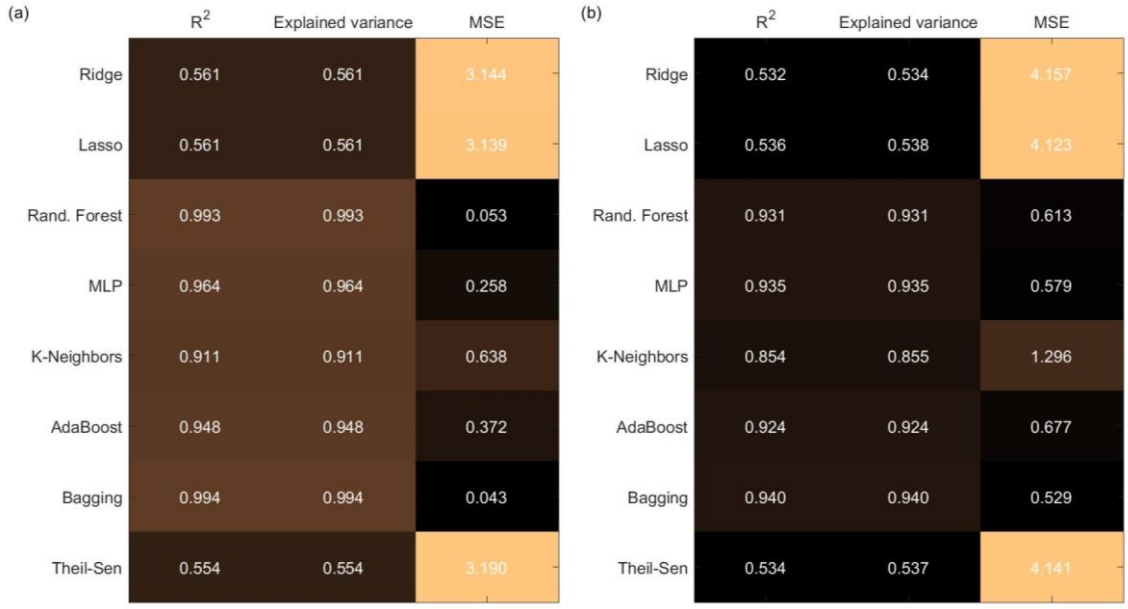


**Figure 4.8 Matrices showing regression model fit measures for optimally-tuned (within explored hyperparameter space) algorithms on original/smaller dataset. (a) shows the results for the training dataset and (b) for the test dataset.**

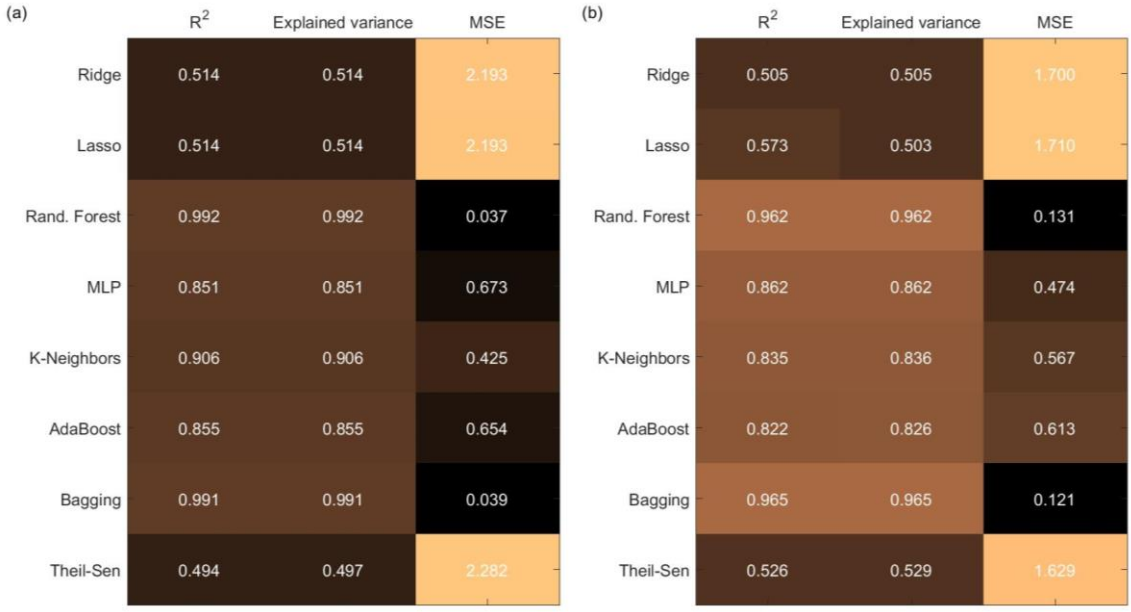


**Figure 4.9 Matrices showing regression model fit measures for optimally-tuned (within explored hyperparameter space) algorithms on extended/larger dataset. (a) shows the results for the training dataset and (b) for the test dataset.**

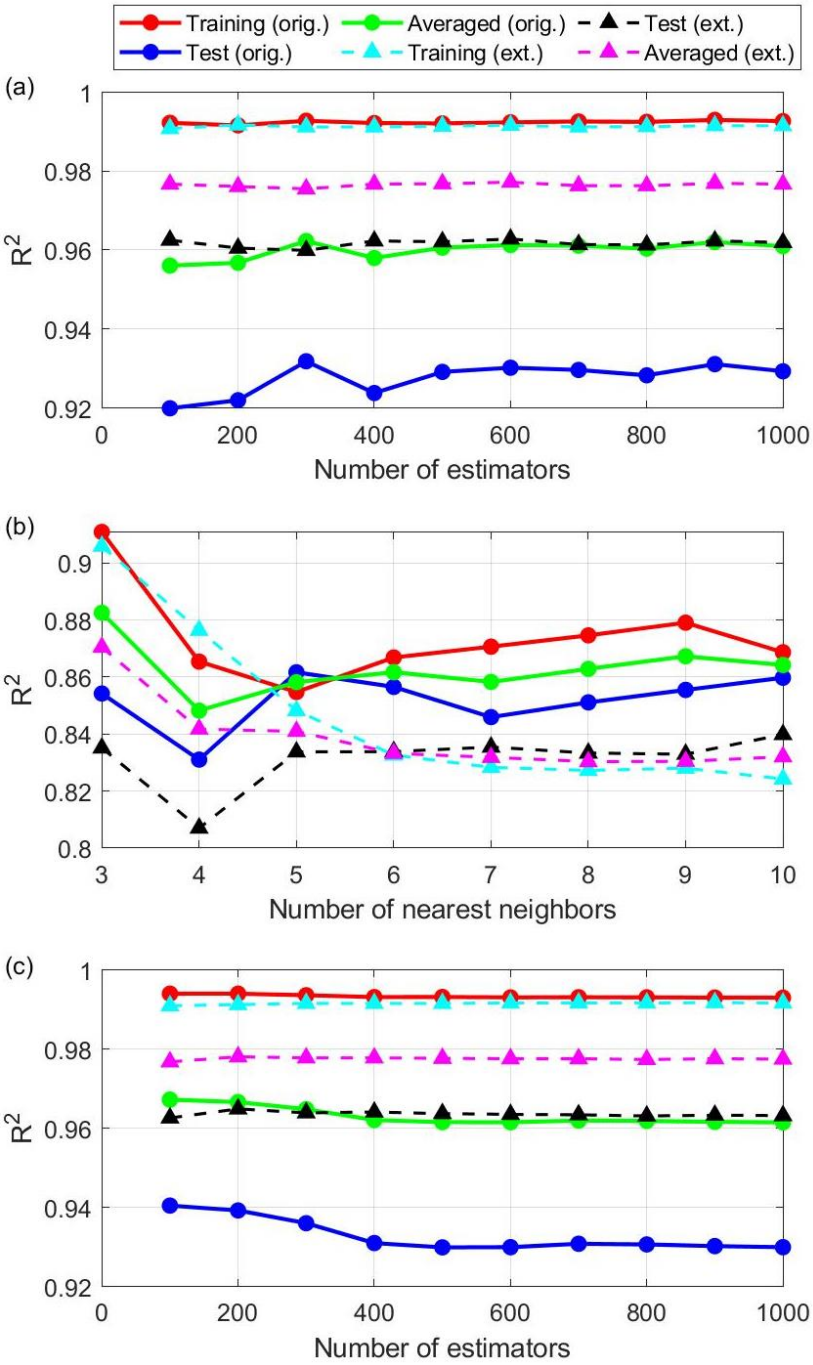


**Figure 4.10 The results of the grid-search within parameter spaces for tuning of (a) Random Forest, (b) K-Neighbors and (c) Bagging algorithms on original (circles) and extended (tringles) datasets with R2 distinguished for training and test parts of these datasets as well as the average of $R^2$ for training and test datasets (see legend above for specific references).**

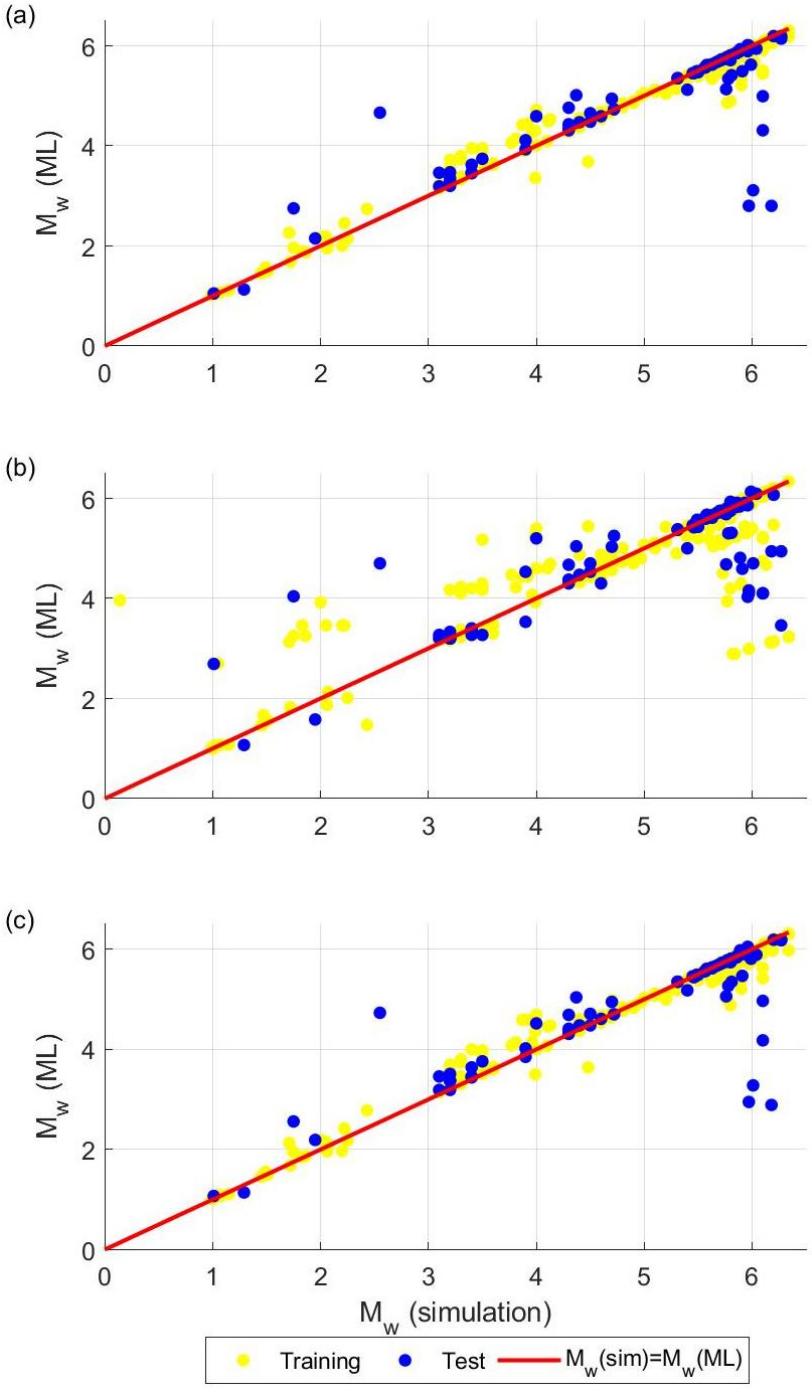


**Figure 4.11 Comparison of Mw for numerical simulation and regression prediction for (a) Random Forest, (b) K-Neighbors and (c) Bagging algorithms optimally-tuned on original dataset. Yellow dots represent samples from training and blue dots from test part of dataset. Red line corresponds to situation where magnitudes from simulation and regression would be equal.**

### 4.5.3. Inversion and its cross-validation

The penultimate part of our workflow is running the inversion using the Metropolis-Hastings algorithm. In total, we run inversion 351 times testing different values of its hyperparameters, different regression algorithms, different sizes of training datasets on which the regression was trained and different sets of inverted parameters. For each one, we pick highest probability and average values from posterior distributions, run cross-validation simulations (in total 702 simulations) and calculate stacked absolute moment magnitude difference (equation 20). Figure 4.12 shows how inversion results depend on the moment magnitude difference acceptance standard deviation, which translates to how is the tolerance for results giving worse results than previous MCMC sample. Median in the plot would indicate almost no preference for a particular standard deviation value, but the minimum and maximum show that lower standard deviation is preferred. Figure 4.13 presents how inversion results depend on the other Metropolis-Hastings algorithm hyperparameter which is sample move range standard deviation, which represents how much the value of the next MCMC sample can change relative to the current sample. The minimum value promotes smaller or intermediate sample moves, while the median indicates intermediate or larger sample move range. However, number of the inversions with extremely wrong results is smaller for larger move ranges, which reflects smaller probability of "local minimum trap".

Next, we analyze the dependence of the inversion quality on the regression algorithm used in generation of the samples (see Figure 4.14). Minimum errors (values of absolute

stacked $M_w$ difference) are obtained for Random Forest and Bagging regression techniques (though by just a small margin), while median error is favorable for K-Neighbors regression algorithm. At the same time, using Random Forest regression algorithm results in the smallest number of the inversions with very high errors (outliers), which overall makes it the best pick.

We also run inversion for different sets of parameters being inverted at the same time – either only frictional fault properties, only stress state related parameters or both at the same time. Figure 4.15 shows that the minimum error is obtained when we allow the inversion of all parameters at the same time. The median, however would indicate that inverting for only one group of parameters is preferrable. Furthermore, the smallest number of high-error inversions was for inversion of fault frictional parameters. Thus, it seems that more degrees of freedom allow to minimize the error in inversions better but requires more careful approach.

Additionally, we look at the influence of number of MCMC samples we use in the inversion on the quality of the inversions (see Figure 4.16). The minimum and median of the inversion errors seems to promote smaller number of samples, which may seem counterintuitive. However, we need to take into account that in this plot we have significant underrepresentation of inversions run with 50000 samples and the statistics might be skewed.

We also compare the results of inversions with different sizes of the dataset on which the regression algorithm is trained (see Figure 4.17). Median and minimum of error indicate that smaller training dataset is preferred, while inversions with algorithms trained on the larger dataset give less very wrong inversion results. These results can be explained by the fact that in the extended dataset the additional data points are not covering as broad parameter space as the data points in the original training dataset. Therefore, in the training process weights might be skewed more towards certain areas of the parameter space and algorithms will get worse scores in tuning procedure.

Finally, we compare two types of result that can be taken from inversion as "the best" result – highest probability and average of posterior distribution (see Figure 4.18). First, the minimum of inversion error does not favor neither of these two types. Second, median of the error slightly favors average results. Third, the highest probability results are generally more resistant to very significant errors. Our advice is however, that the person analyzing the results need to make a decision whether taking average (e.g. posterior distribution close to normal or truncated uniform uniform) or highest probability (e.g. multimodal or log-normal posterior distribution) is better in a given case.

Beside doing statistical analysis of all the inversion runs, we also select two runs with the lowest stacked moment magnitude difference and examine them in more details. Unfortunately, the dynamic rupture simulation for statistically best inversion result gives

a rupture pattern that from geologic point of view is very unlikely, because it has multiple equivalent initiation patches spread around different parts of the fault segment. The 2nd best result was much more geologically sound and thus we present more details about it. Figure 4.19 shows the posterior distributions resulting from the inversion. We can see that higher $S_n/S_v$ ratio values are preferred but not exclusive (Figure 4.19a). The peak of probability is well-defined in that posterior distribution with highest probability for $S_n/S_v = 2.4$. Results also indicate a preference for intermediate $\tau_0/\tau_u$ ratio (Figure 4.19b), also with well-defined probability peak and highest probability for $\tau_0/\tau_u = 0.76$. The peak of probability in the posterior distribution of $\mu_s$ (Figure 4.19c) is visible but not as well-defined. The highest probability indicates relatively low static friction coefficient value of $\mu_s = 0.36$. The posterior distribution for dynamic friction coefficient looks almost like a truncated uniform distribution and does not have a defined peak (Figure 19d). The highest probability indicates $\mu_d = 0.24$, which in relation to $\mu_s$ is relatively high. However, because of the shape of the distribution, the value is poorly resolvable and we should have limited trust towards it. The highest probability value is $D_0 = 0.13$ m, which is relatively low and the peak in posterior distribution is defined but not very prominent (Figure 19e). Finally, the posterior distribution for the pore pressure modification in relation to general background trend has well-defined peak and the highest probability value indicates quite significant overpressure with $\Delta p_b^{frct} = 0.16$. In this particular inversion we use the Random Forest regression algorithm which is tuned on the extended training dataset. The values for the moment magnitude difference acceptance standard deviation and sample move standard deviation are equal to 0.2 and

2, respectively. We run 10000 MCMC samples and took highest probability solutions for simulation cross-validation. We predict $M_w$ 4.77 and simulate $M_w$ 4.74, which gives stacked absolute moment magnitude difference of 0.06.

The inverted values mentioned in the previous paragraph are actually averages of normal distributions we use to generate initial stress state and spatial distributions of fault frictional parameters. Additionally, we assume standard deviations of 0.3 for $S_n/S_v$ ratio, 0.1 for $\tau_0/\tau_u$ ratio, 0.2 for both static and dynamic friction coefficients, 0.1 m for critical slip distance, and 0.1 for $\Delta p_b^{frct}$. Figure 20 presents one of the realizations of these normal distributions as spatial distribution of frictional coefficients difference (Figure 4.20a), critical slip distance (Figure 4.20b), as well as background pore pressure (Figure 4.20c), normal stress (Figure 4.20d) and shear stress (Figure 4.20e) on the fault. Because of the assumption of uniform normal effective stress, the shear stress distribution on the fault does not have the depth-dependence pattern.

Finally, Figure 4.21 shows the results of the dynamic rupture simulation for the model presented in Figure 4.20. Majority of the ruptured segment is located in the shallow, central part of the fault. There are some small co-existing minor failures in the deeper portion of the fault, which are side effects of the introduced heterogeneity. However, the major trend is not affected by them and the rupture is propagating mainly along strike with minor component along dip, which agrees with the observations. Maximum slip on fault reaches ~0.21 m, while an average slip is ~0.03 m. The peak stress drop is reaching

~17 MPa in the area of maximum slip but the average stress drop is ~2 MPa. On the edge of the rupture area there is actually a small stress increase.

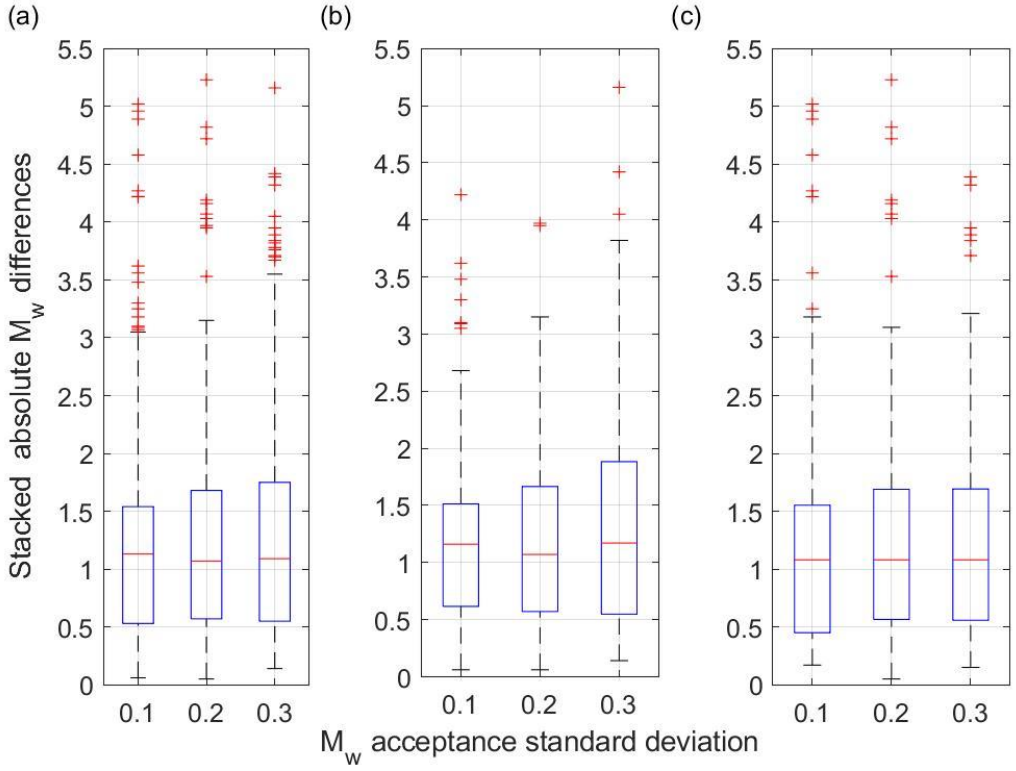


**Figure 4.12 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different values of Mw acceptance standard deviation used in Metropolis-Hastings algorithm. (a) shows all combined results, (b) presents results for highest probability (of posterior distributions) parameters results and (c) shows results for average (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

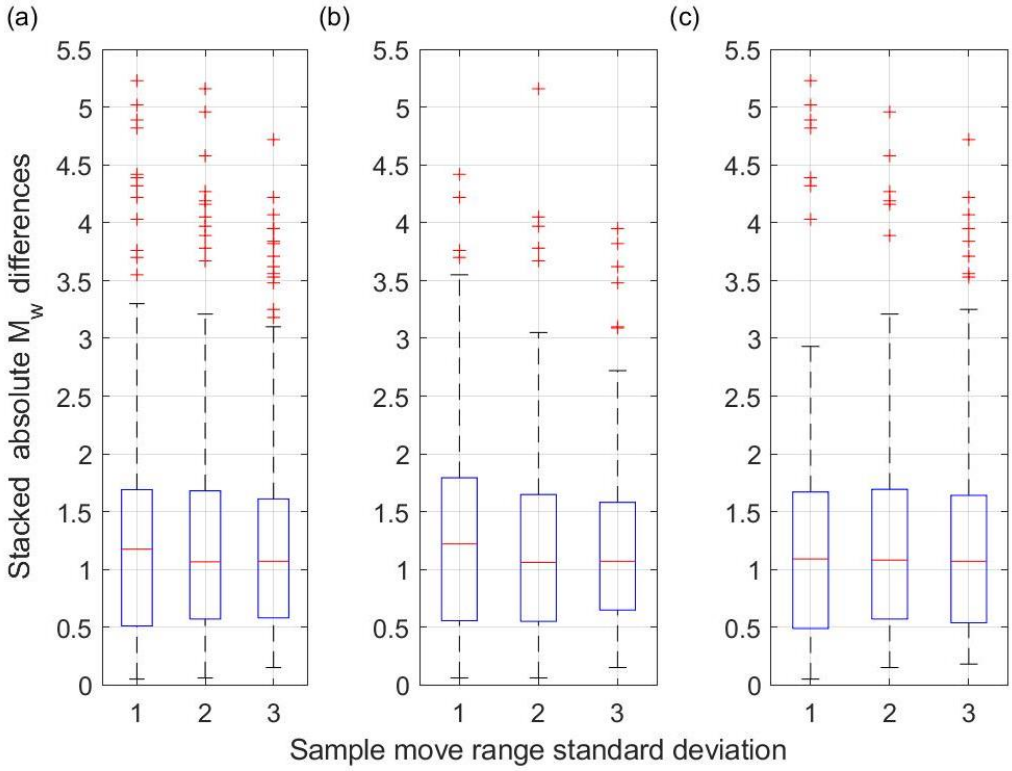


**Figure 4.13 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different values of sample move range standard deviation used in Metropolis-Hastings algorithm. (a) shows all combined results, (b) presents results for average (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

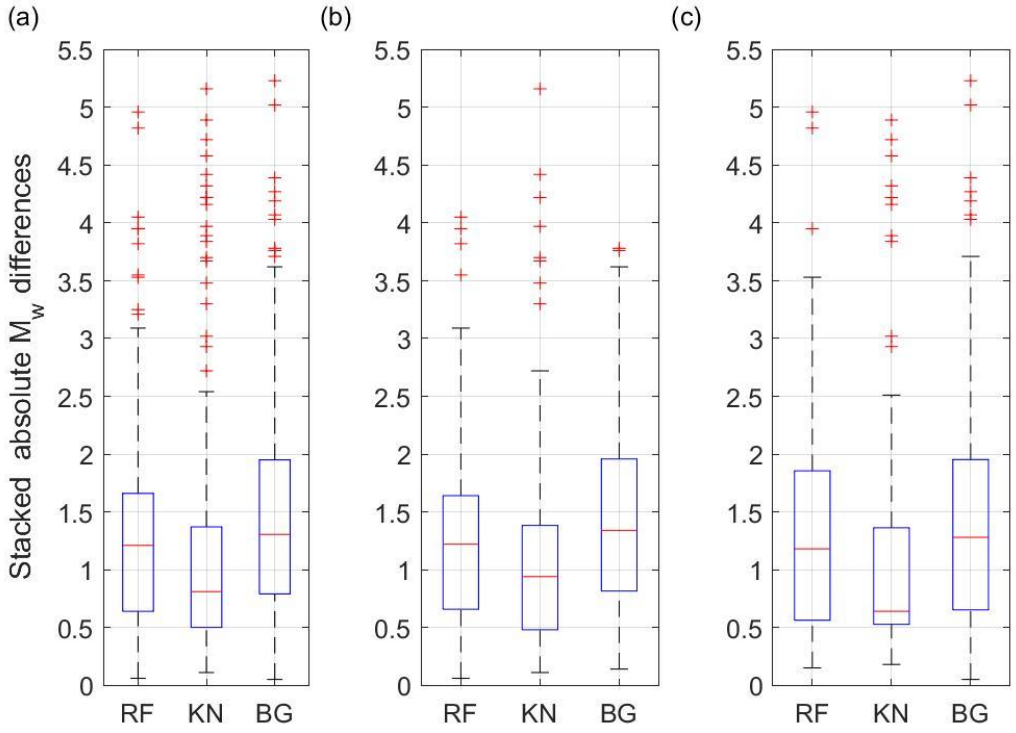


**Figure 4.14 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different regression algorithms (RF – Random Forest; KN – K-Neighbors; BG - Bagging). (a) shows all combined results, (b) presents results for highest probability (of posterior distributions) parameters results and (c) shows results for average (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

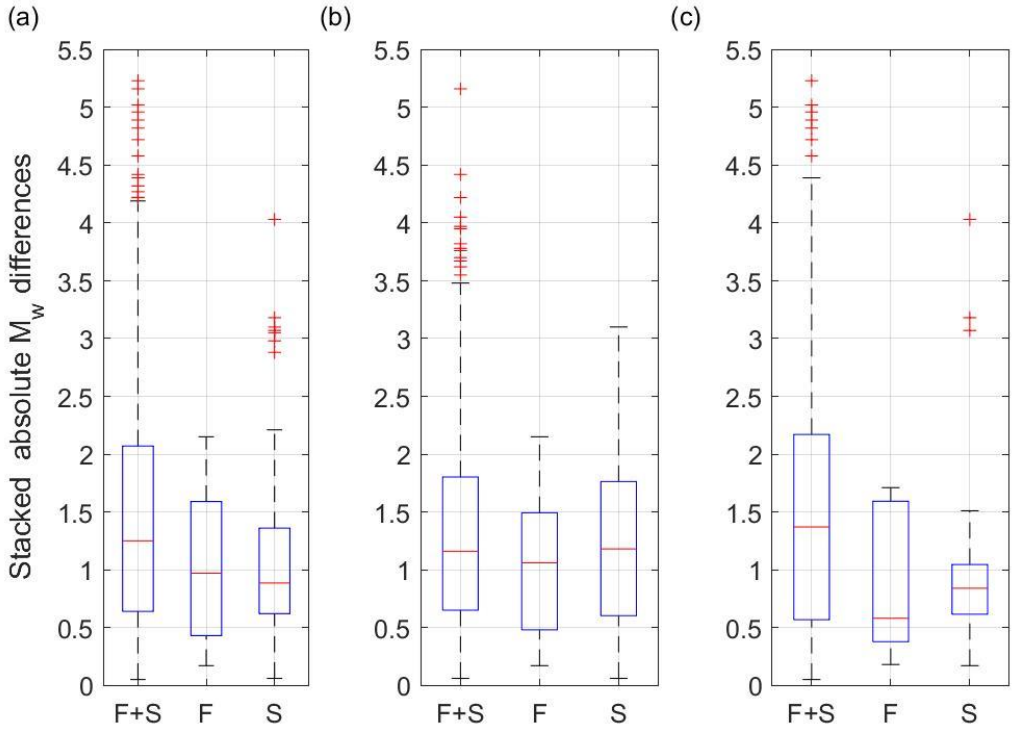


**Figure 4.15 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different sets of inverted parameters (F – fault friction parameters; S – stress state parameters). (a) shows all combined results, (b) presents results for highest probability (of posterior distributions) parameters results and (c) shows results for average (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

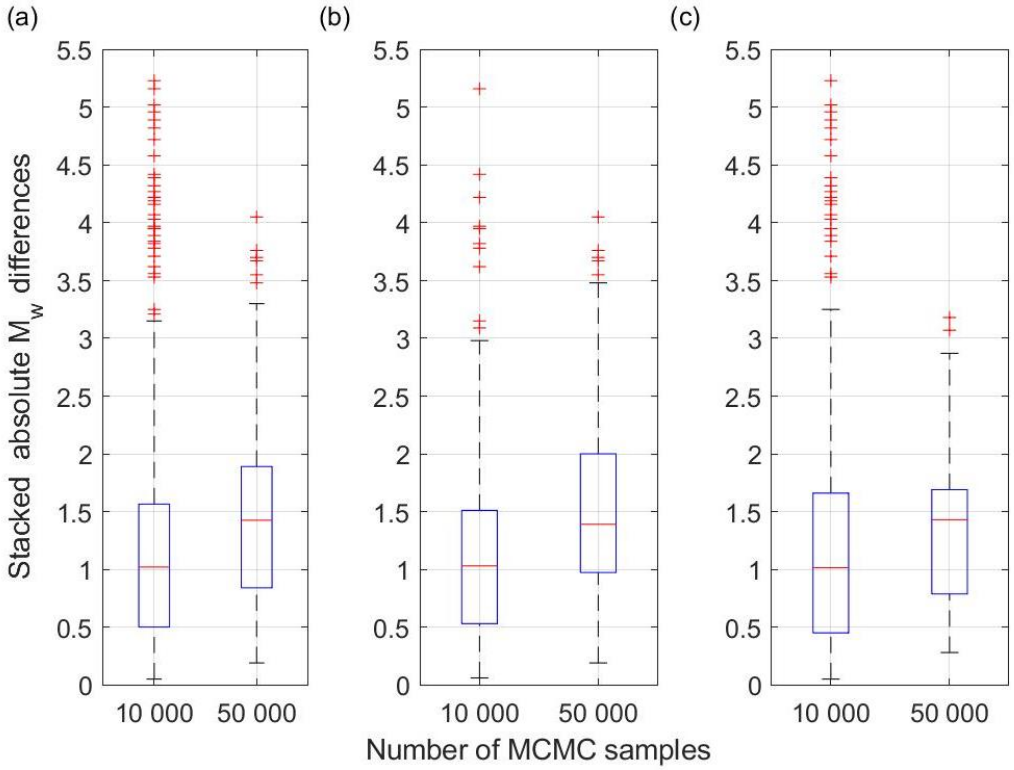


**Figure 4.16 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different number of samples used in Metropolis-Hastings algorithm. (a) shows all combined results, (b) presents results for highest probability (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

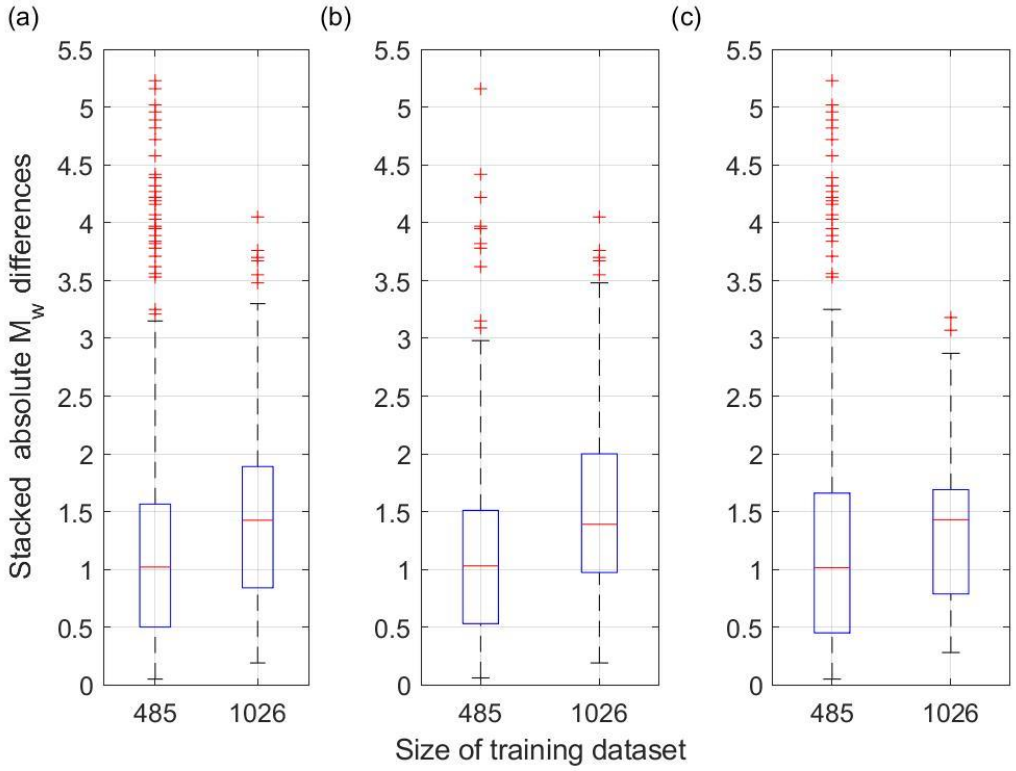


**Figure 4.17 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for different number of samples in the training datasets. (a) shows all combined results, (b) presents results for highest probability (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

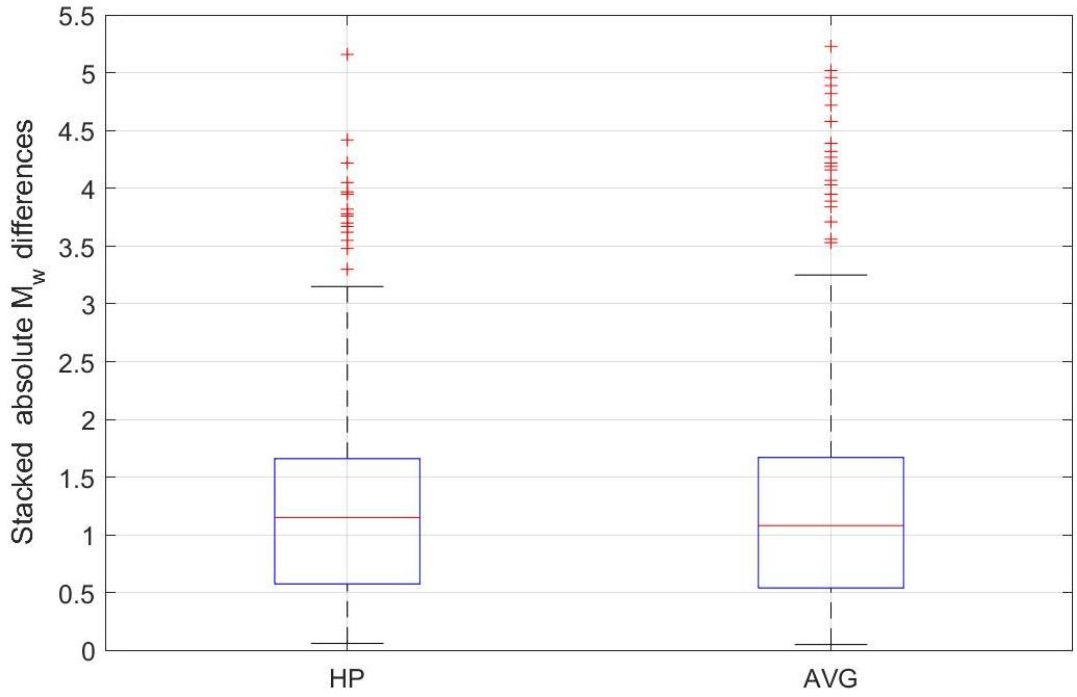


**Figure 4.18 Box-plots summarizing the results of inversion in terms of stacked absolute Mw differences ranges for highest probability (of posterior distributions) parameters results and average (of posterior distribution) results. Red crosses denote outliers, red line represents median, bottom and top of the blue boxes represent 25$^{th}$ and 75$^{th}$ percentiles respectively and whiskers show the inversion results boundaries after excluding statistical outliers.**

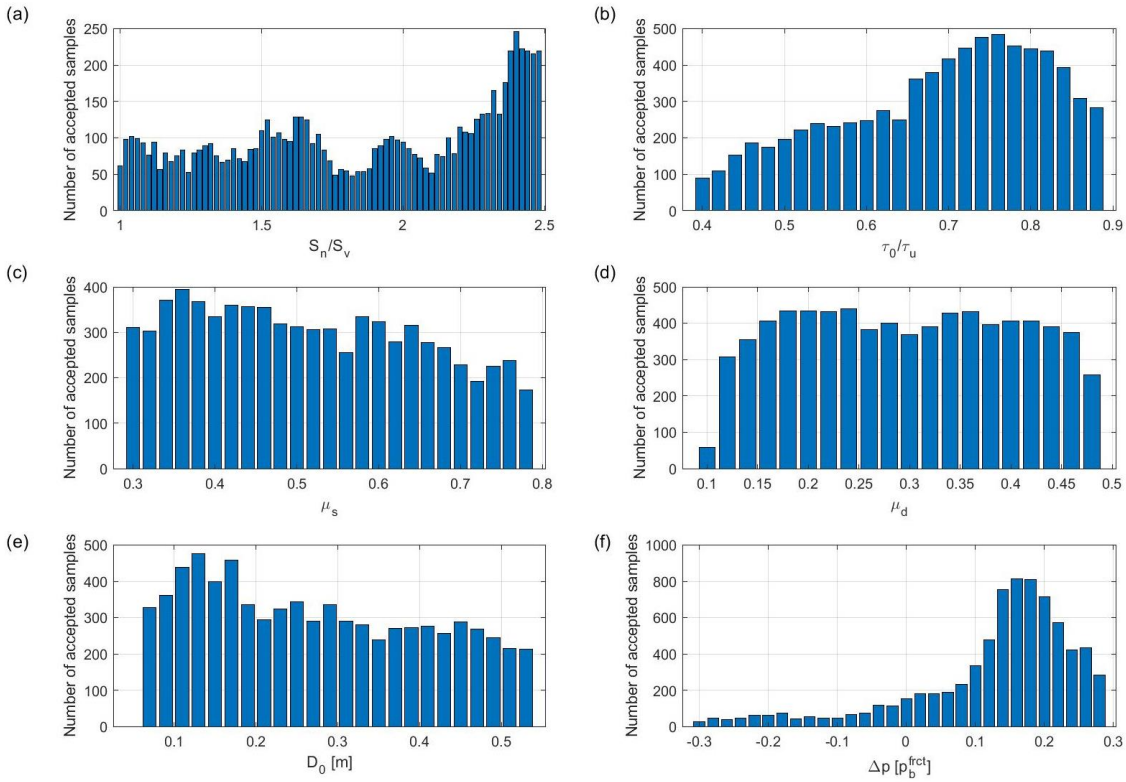


**Figure 4.19 Discrete posterior distribution of the inverted parameters for the 2nd best inversion results. (a) is distribution for normal-to-vertical stress ratio, (b) for shear-to-yield stress ratio, (c) and (d) for static and dynamic friction coefficients, (e) for critical slip distance and (f) for background pressure modification.**

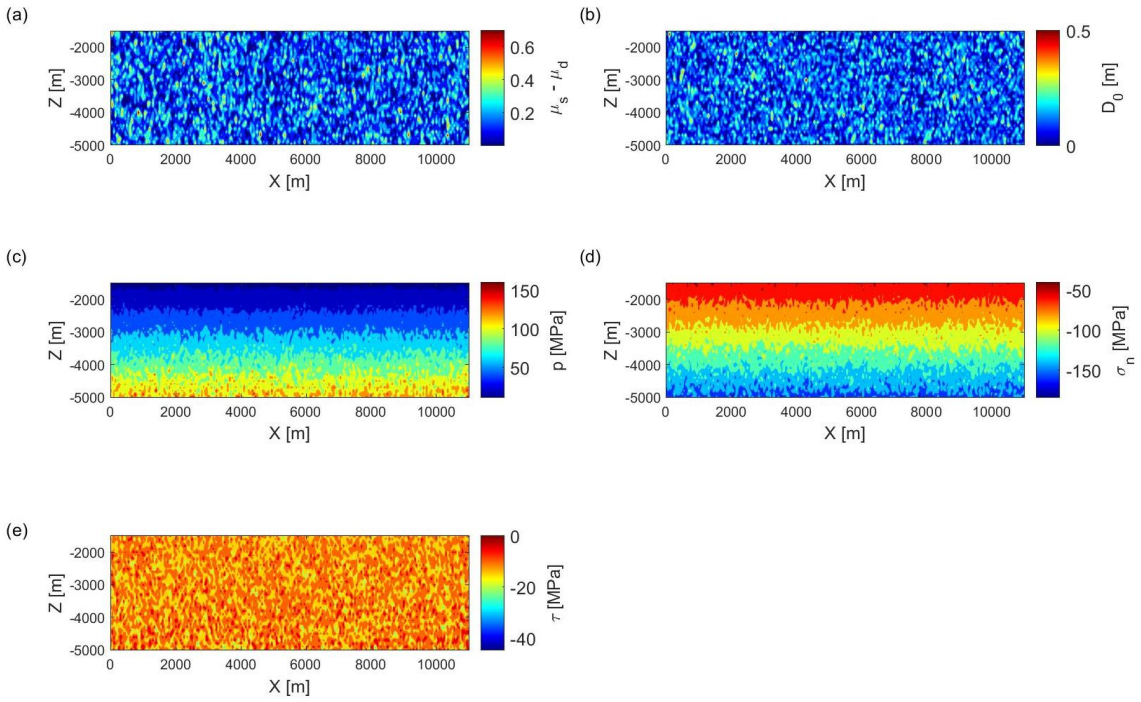


**Figure 4.20 Distribution of (a) static – dynamic friction coefficient difference, (b) critical slip distance, background (c) pore pressure, (d) normal stress, (e) shear stress on the fault plane (SE-NW from left to right) as one of the random realizations of the highest probability 2nd best inversion results, where $S_n/S_v$=2.4 ± 0.3, $\tau_0/\tau_u$ = 0.76 ± 0.1, $\mu_s$ = 0.36 ± 0.2, $\mu_d$ = 0.24 ± 0.2, $D_0$ = 0.13 ± 0.1, $\Delta p$ = 0.16 ± 0.1.**

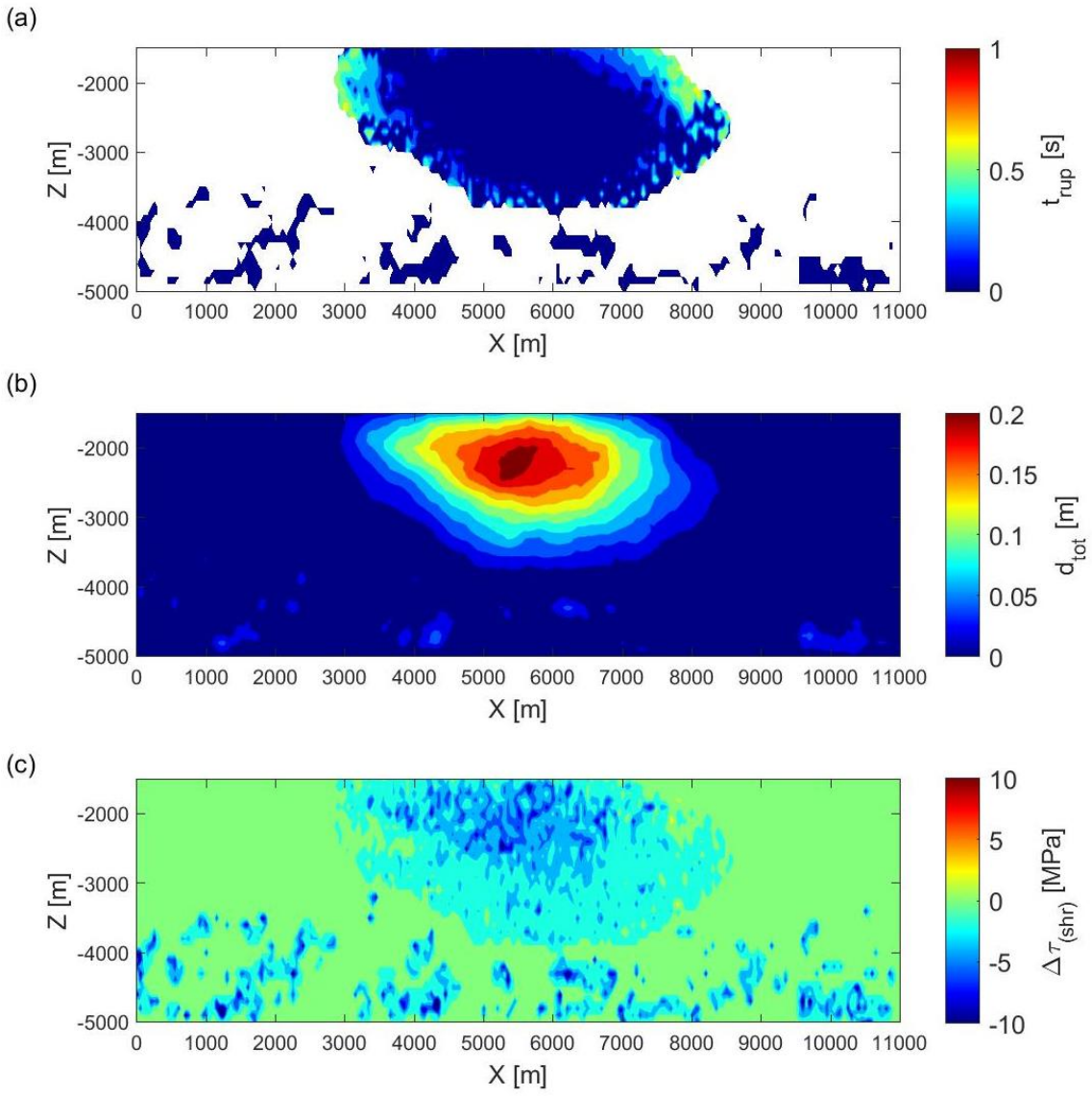


**Figure 4.21 Distribution of (a) rupture time, (b) total slip and (c) stress drop on the fault plane (SE-NW from left to right) for the time of the Timpson earthquake mainshock for the 2nd best inversion result model. Simulated magnitude is Mw4.8 which is one of the possible exact matches of reported $M_w$.**

## 4.6. Discussion

In the previous sections, we present how the entire workflow from Figure 4.1 can be implemented in the real case scenario. Here, we discuss several important but nonessential topics.

First one is about the construction of the training dataset and the selection of parameters we would like to use in the training of the algorithm and later in the inversion process. In Figure 4.22, we show the cross-correlation coefficient matrix between the input parameters we use in training and inversion (based on the original dataset). In between most of the input parameters there is very low to low correlation, which indicates that they are not strongly related and can provide valuable extra information in the process of regression algorithm training. The only two input parameters that have a correlation of some statistical significance are $\mu_s$ and $\mu_d$, which makes sense also from a physical point of view. The rupture process after initial failure is controlled by the difference between these two (plus critical slip distance) and therefore we shall expect some degree of correlation between them. That creates a potential to reduce the number of input parameters.

We conduct an experiment where we perform training and cross-validation of the selected regression algorithms after replacing $\mu_s$, $\mu_d$ and initial $\tau_0/\tau_u$ with the single parameter that is dependent on the results of the fluid flow simulation for a particular

moment in time. In one scenario, we use a popular metrics in the dynamic rupture modeling community, which is *S-value*. It can be defined as:

$$S(t) = \frac{\tau_u(t) - \tau_0(t)}{\tau_0(t) - \tau_f(t)}, \tag{4.20}$$

where $\tau_u(t)$, $\tau_0(t)$ and $\tau_f(t)$ are the values of yield stress, shear stress and frictional stress at the time *t*. In our case the time is the moment right before the rupture happens. In the second scenario instead of S-value, we use $\tau_0/\tau_u$ ratio at the time of right before the rupture, which includes changes of shear and yield stress due to injection process. In both scenarios we take an average value on the fault plane and use that as one of the inputs in training of the regression algorithms. We perform tuning of Random Forest, K-Neighbors and Bagging algorithms on both original and extended datasets. In Figure 4.23 we compare the performance of each algorithm for each scenario with the reference $R^2$ value for both training and test subsets of both small and large dataset. When dataset is smaller, K-Neighbors performs better on both training and test subsets for both scenarios. K-Neighbors also performs better for the scenario with $\tau_0/\tau_u$ ratio on the day of rupture for larger dataset for both training and test subsets. However, with introduced *S-value*, K-Neighbors performs worse on the larger dataset. Random Forest and bagging algorithms perform worse on training subsets of either smaller or larger datasets for both scenarios. They also perform worse on the test subset of the larger dataset for both scenarios. However, if we consider test subset of the smaller dataset, these algorithms increase their performance. Overall, the replacement in the second scenario seems to be a better option if we would like to reduce the number of the input parameters.

Additionally, the better performance of K-Neighbors might be correlated with the fact that this algorithm performs better for smaller parameter spaces.

There are also some limitations of this work. First, we want to point out the limited information available about the earthquake. For example, there is lack of the local (close) seismic stations, which makes slip inversion on the fault plane or using the registered waveforms not feasible. Second limitation is using only magnitude of the mainshock in regression training and the inversion procedure. This limitation can be actually addressed in further methodology development, by e.g. using magnitudes of the entire mainshock – aftershock sequence. In Figure 4.24 we present a result of entire sequence simulation using statistically best and 2$^{nd}$ best models and comparing modeling results with the reported magnitudes. For practical purposes we only use aftershocks which are at least M2.0. We can see that for the main shock and immediate aftershocks we get a good match between reported and simulated earthquakes. For the aftershocks that took place between the end of 2012 and beginning of 2013, the magnitudes and released seismic moment are not matching and are generally higher in the simulated events. For the late aftershocks from the second half of 2013, we again get pretty good match between registered and modeled seismic events. Beside the differences in values, we can observe an interestingly good match in general trend of event-to-event changes of moment magnitude and released seismic moment. The problem in matching the values can be the result of not capturing all physical phenomena present in the

subsurface, such as possible creeping motion on the fault which could release some of the stress accumulated between early and intermediate aftershocks.

Finally, in the inversion procedure we invert for only general distribution characteristics of parameters on the fault, via inverting for normal distribution averages and using sensible standard deviations. This choice is again driven by the limited data available. Possibly having slip distribution on fault or multiple seismograms from close stations we might be able to invert for precise distribution of fault parameters.

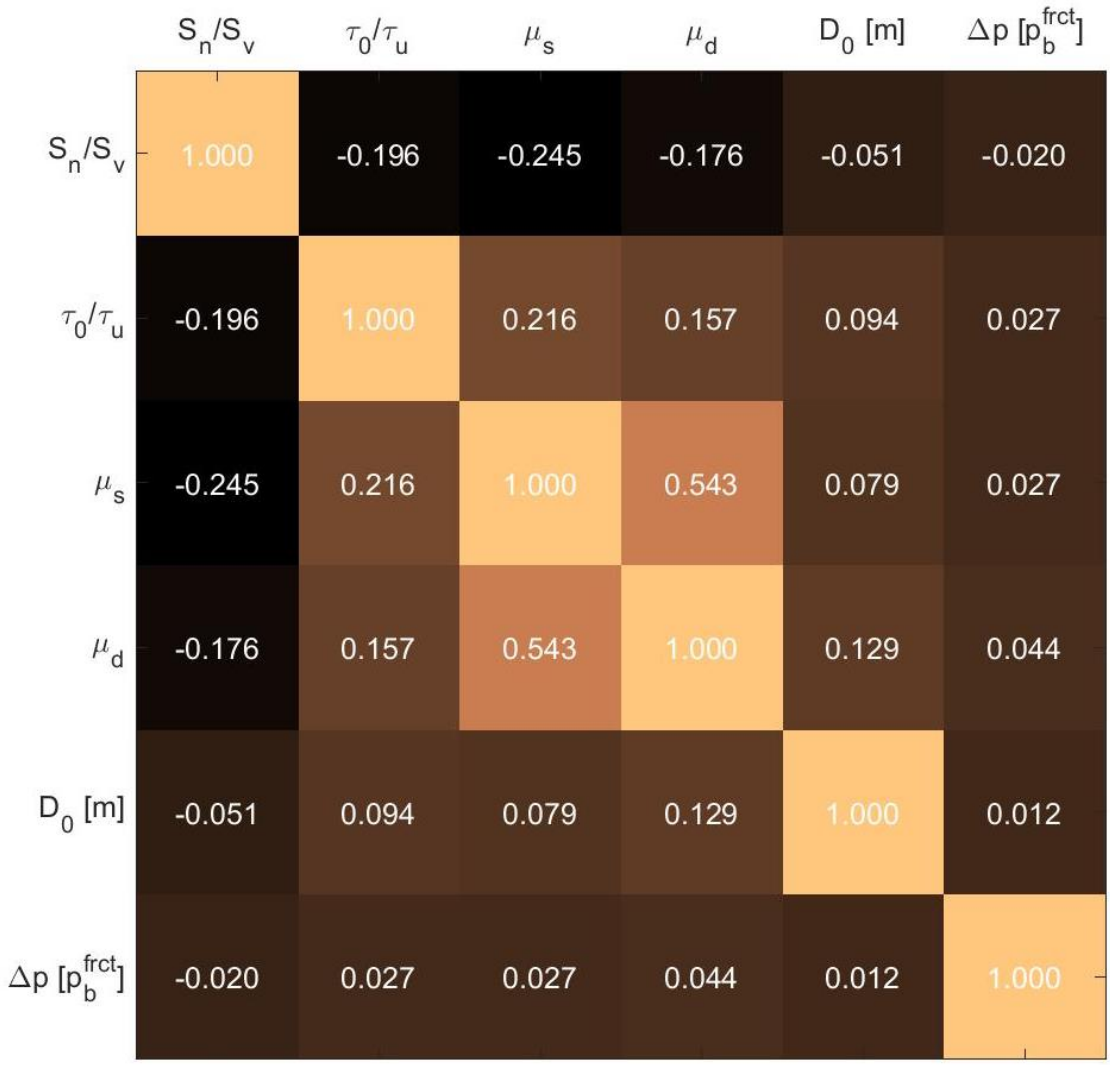


**Figure 4.22 Matrix of cross-correlation coefficients between inverted input parameters.**

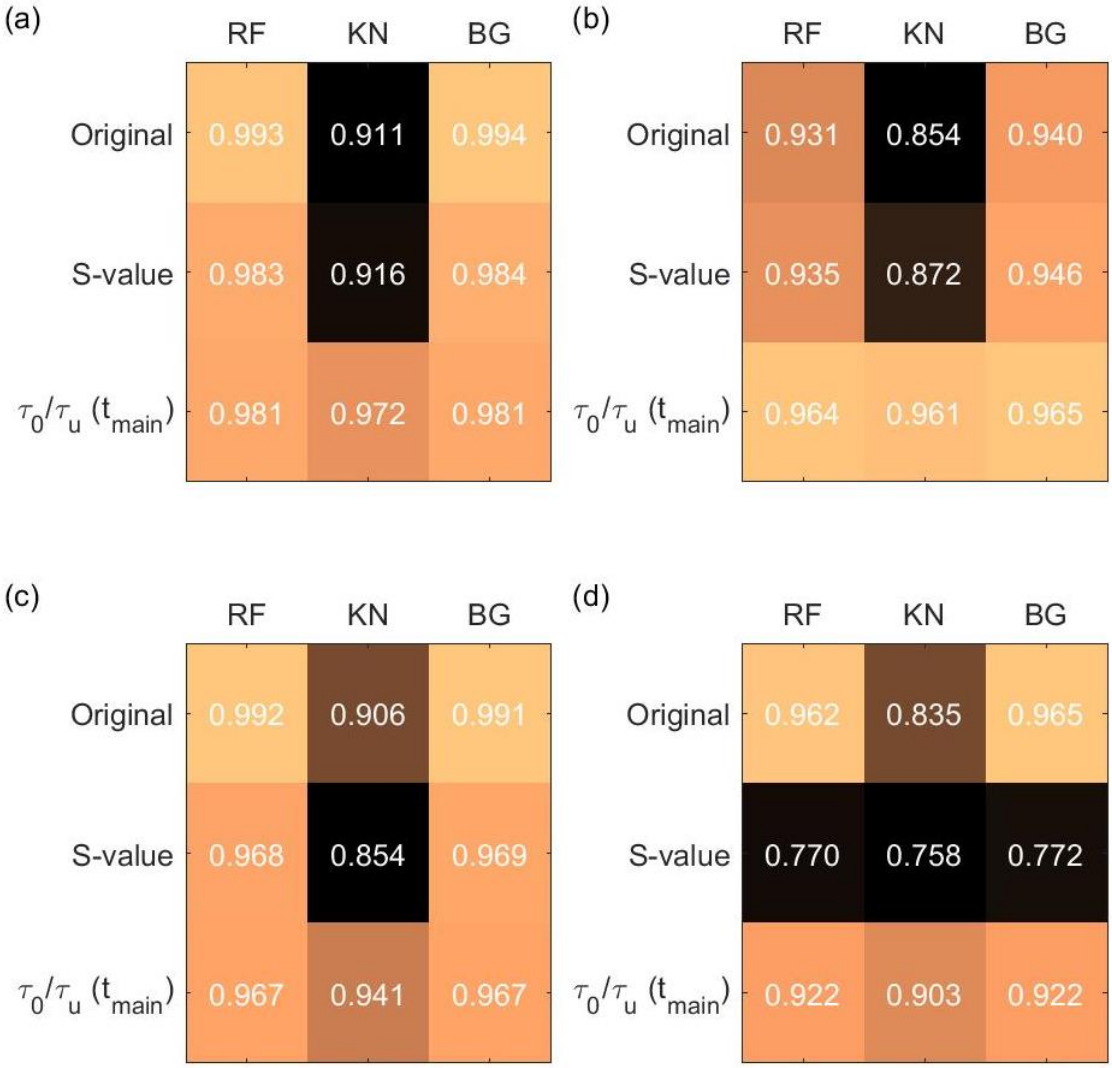


**Figure 4.23 Comparison between tuning results (in terms of R2) of the selected regression algorithms (RF – Random Forest; KN – K-Neighbors; BG - Bagging) on small ((a) and (b)) and large ((c) and (d)) for training ((a) and (c)) and test ((b) and (d)) subsets for either original set of input parameters, or replacing background $\mu_s$, $\mu d$ and $\tau_0/\tau_u$ with the average S-value or average of $\tau_0/\tau_u$ coefficient on the fault plane at the moment right before the rupture (main shock).**

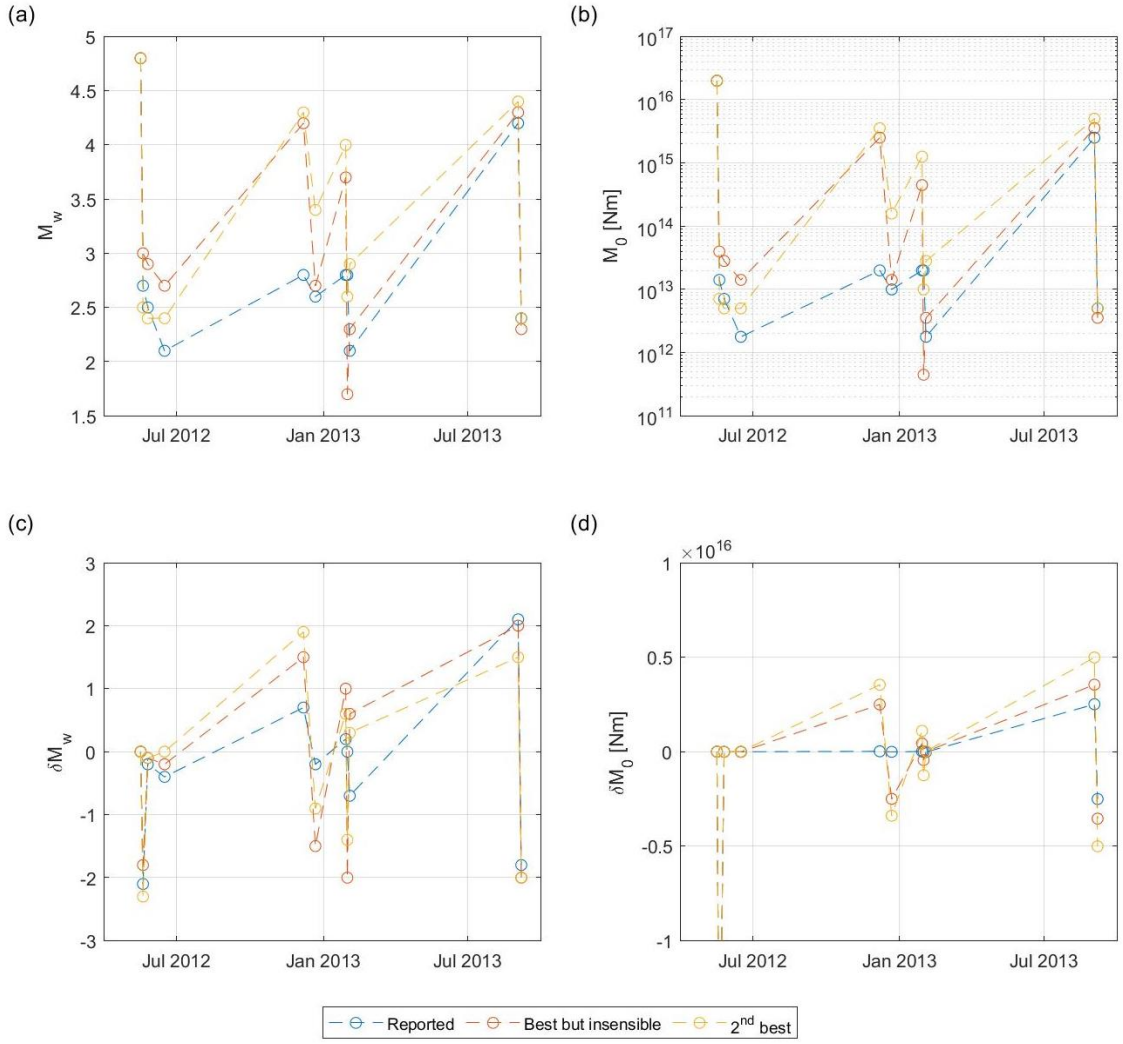


**Figure 4.24 Comparison of reported and simulations for two top inversion results (according to stacked absolute Mw differences measure) in terms of (a) moment magnitude, (b) released seismic moment, (c) event-to-event change of moment magnitude, and (d) event-to-event released seismic moment for the Timpson mainshock-aftershock sequence.**

## 4.7. Conclusions

In this work, we develop a workflow, which connects modeling of fluid flow in deformable porous medium, simulation of dynamic rupture propagation, tuning and

training of the machine learning regression algorithms and Bayesian inversion. We applied the methodology to May 2012 $M_w$ 4.8 Timpson, TX, earthquake.

From fluid flow simulation we get the perturbations on the fault on the order of several MPa with the dominant driving force being the effective normal stress change. These results are used in further simulations of dynamic rupture propagation from which we derive synthetic earthquakes and calculate moment magnitudes. The magnitudes are ranging from ~$M_w$ 0.0 to ~$M_w$ 6.5, and we got 16 and 94 earthquakes with magnitudes falling into range of $M_w$ 4.8 ± 0.2 in the original and extended datasets, respectively.

Using results of dynamic rupture simulations, we form two training datasets, containing 485 and 1026 samples. With these datasets we perform tuning of selected hyperparameters for the suite of regression algorithms. Through this process, we determine that the most promising results are for Random Forest, Bagging and K-Neighbors regression algorithms, with performance indicator $R^2$ above 0.9 for the first two and above 0.85 for the last algorithm. On original dataset, Random Forest performs best with 300 estimators, Bagging with 100 estimators and K-Neighbors with 3 nearest neighbors involved in the computation.

Trained regression algorithms are then used in the Bayesian inversion replacing running numerical simulations with the machine learning predictions. Dependent on the selection of training dataset size and number of MCMC samples in inversion, we can save

between 85 and 97% of computation time on a single inversion. However, we need to consider the necessity of inversion calibration and validation. As shown in the example above, it is necessary to do reality check of the results, which can help to avoid erroneous interpretations and discard geologically improbable solutions. Our selected solution, which scores 2$^{nd}$ according to evaluation statistic and physically sensible at the same time, is obtained using Random Forest regression trained on the extended dataset incorporated in the Metropolis-Hastings algorithm with intermediate values for both moment magnitude acceptance and sample move standard deviations, and only 10000 samples generated in the Markov Chains. The highest probability values of parameters taken from their posterior distributions indicate relatively high initial normal stress ($S_n/S_v$ = 2.4) and pore pressure ($\Delta p_b^{frct}$ = 0.16) on the fault, with intermediate level of initial shear stress ($\tau_0/\tau_u$ = 0.76). At the same time, we get relatively low values of critical slip distance of 0.13 m, and low values of static and dynamic friction coefficients of 0.36 and 0.24, respectively.

We also show that the original input parameters in training datasets can be potentially replaced by some alternatives. Dependent on a particular regression algorithm or size of training dataset, we can get either slightly better or slightly worse results of regression algorithms tuning.

Finally, the study could be further expanded by building training dataset based on numerical simulations of entire sequence of seismic events instead of using only

mainshock. In this study we only present such simulation for two models and compare it with reported events. The general trend is captured well, but the values of magnitudes for several aftershocks are visibly too large in comparison to reality, which would require further investigation.

The proposed methodology can be applied to real scenarios, but the user needs to be aware of the limitations presented by a certain case. It cannot be used as a simple black-box and requires knowledge not only of the algorithms, but also of the earthquake physics and geology.

### 4.8. References

Ahamed, S., & Daub, E. G. (2019) 'Machine learning approach to earthquake rupture dynamics', *arXiv preprint arXiv:1906.06250.*

Alghannam, M., & Juanes, R. (2020) 'Understanding rate effects in injection-induced earthquakes.', *Nature communications*, 11(1), pp. 1-6.

Breiman, L. (1996) 'Bagging predictors.', *Machine learning*, 24(2), pp. 123-140.

Breiman, L. (2001) 'Random forests.', *Machine learning*, 45(1), pp. 5-32.

Brooks, S. (1998) ‘Markov chain Monte Carlo method and its application.’, *Journal of the royal statistical society: series D (the Statistician)*, 47(1), pp. 69-100.

Day, S. M., et al. (2005) ‘Comparison of finite difference and boundary integral solutions to three-dimensional spontaneous rupture.’, *Journal of Geophysical Research: Solid Earth*, 110(B12).

Deng, K., Liu, Y., & Harrington, R. M. (2016) ‘Poroelastic stress triggering of the December 2013 Crooked Lake, Alberta, induced seismicity sequence.’, *Geophysical Research Letters*, 43(16), pp. 8482-8491.

Duan, B. (2012) ‘Dynamic rupture of the 2011 Mw 9.0 Tohoku-Oki earthquake: Roles of a possible subducting seamount.’, *Journal of Geophysical Research: Solid Earth*, 117(B5).

Duan, B. (2019) ‘Multicycle dynamics of the Aksay Bend along the Altyn Tagh fault in Northwest China: 1. A simplified double bend.’, *Tectonics*, 38(3), pp. 1101-1119.

Duan, B., & Oglesby, D. D. (2006) ‘Heterogeneous fault stresses from previous earthquakes and the effect on dynamics of parallel strike-slip faults.’, *Journal of Geophysical Research: Solid Earth*, 111(B5).

Fan, Z., Eichhubl, P., & Gale, J. F. (2016) 'Geomechanical analysis of fluid injection and seismic fault slip for the Mw4. 8 Timpson, Texas, earthquake sequence.', *Journal of Geophysical Research: Solid Earth*, 121(4), pp. 2798-2812.

Fichtner, A., & Simutė, S. (2018) 'Hamiltonian Monte Carlo inversion of seismic sources in complex media.', *Journal of Geophysical Research: Solid Earth*, 123(4), pp. 2984-2999.

Freund, Y., & Schapire, R. E. (1997) 'A decision-theoretic generalization of on-line learning and an application to boosting.', *Journal of computer and system sciences*, *55*(1), pp. 119-139.

Frohlich, C., et al. (2011) 'The Dallas–Fort Worth earthquake sequence: October 2008 through May 2009.', *Bulletin of the Seismological Society of America*, 101(1), pp. 327-340.

Frohlich, C., et al. (2014) 'The 17 May 2012 M4.8 earthquake near Timpson, East Texas: An event possibly triggered by fluid injection.', *Journal of Geophysical Research: Solid Earth*, 119(1), pp. 581-593.

Gallovič, F., et al. (2019) ‘Bayesian dynamic finite-fault inversion: 2. Application to the 2016 Mw 6.2 Amatrice, Italy, earthquake.’, *Journal of Geophysical Research: Solid Earth*, 124(7), pp. 6970-6988.

Goebel, T. H. W., et al. (2017) ‘The 2016 Mw5. 1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at> 40 km from fluid disposal wells.’, *Earth and Planetary Science Letters*, 472, pp. 50-61.

Harris, R. A., et al. (2009) ‘The SCEC/USGS dynamic earthquake rupture code verification exercise.’, *Seismological Research Letters*, 80(1), pp. 119-126.

Harris, R. A., et al. (2011) ‘Verifying a computational method for predicting extreme ground motion.’, *Seismological Research Letters*, 82(5), pp. 638-644.

Harris, R. A., et al. (2018) ‘A suite of exercises for verifying dynamic earthquake rupture codes.’, *Seismological Research Letters*, 89(3), pp. 1146-1162.

Hassoun, M. H. (1995) *Fundamentals of artificial neural networks*. MIT press.

Hastings, W. K. (1970) ‘Monte Carlo sampling methods using Markov chains and their applications.’, *Biometrika*, 57(1), pp. 97-109.

Ida, Y. (1972) ‘Cohesive force across the tip of a longitudinal-shear crack and Griffith's specific surface energy.’, *Journal of Geophysical Research*, 77(20), pp. 3796-3805.

Jin, L., & Zoback, M. D. (2018) ‘Fully dynamic spontaneous rupture due to quasi-static pore pressure and poroelastic effects: An implicit nonlinear computational model of fluid-induced seismic events.’, *Journal of Geophysical Research: Solid Earth*, 123(11), pp. 9430-9468.

Johann, L., Shapiro, S. A., & Dinske, C. (2018) ‘The surge of earthquakes in Central Oklahoma has features of reservoir-induced seismicity.’, *Scientific reports*, 8(1), pp. 1-14.

Juanes, R., et al. (2016) ‘Were the May 2012 Emilia-Romagna earthquakes induced? A coupled flow-geomechanics modeling assessment.’, *Geophysical Research Letters*, 43(13), pp. 6891-6897.

Kim, W. Y. (2013) ‘Induced seismicity associated with fluid injection into a deep well in Youngstown, Ohio.’, *Journal of Geophysical Research: Solid Earth,* 118(7), 3506-3518.

King, V. M., Block, L. V., & Wood, C. K. (2016) ‘Pressure/flow modeling and induced seismicity resulting from two decades of high-pressure deep-well brine injection, Paradox Valley, Colorado.’ *Geophysics*, 81(5), B119-B134.

Kingma, D. P., & Ba, J. (2014) ‘Adam: A method for stochastic optimization.’, *arXiv preprint arXiv:1412.6980*.

Kummer, N., & Najjaran, H. (2014) ‘Adaboost. MRT: Boosting regression for multivariate estimation.’, *Artif. Intell. Research*, 3(4), pp. 64-76.

Liaw, A., & Wiener, M. (2002) ‘Classification and regression by randomForest.’, *R news*, 2(3), pp. 18-22.

Liu, D., & Duan, B. (2018) ‘Scenario Earthquake and Ground-Motion Simulations in North China: Effects of Heterogeneous Fault Stress and 3D Basin StructureScenario Earthquake and Ground-Motion Simulations in North China.’, *Bulletin of the Seismological Society of America*, 108(4), pp. 2148-2169.

Marquardt, D. W., & Snee, R. D. (1975) ‘Ridge regression in practice.’, *The American Statistician*, 29(1), pp. 3-20.

McDonald, G. C. (2009) ‘Ridge regression.’, *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), pp. 93-100.

Metropolis, N., et al. (1953) ‘Equation of state calculations by fast computing machines.’, *The journal of chemical physics*, 21(6), 1087-1092.

Murtagh, F. (1991) ‘Multilayer perceptrons for classification and regression.’, *Neurocomputing*, 2(5-6), pp. 183-197.

Nguyen, B., Morell, C., & De Baets, B. (2016) ‘Large-scale distance metric learning for k-nearest neighbors regression.’, *Neurocomputing*, 214, pp. 805-814.

Omohundro, S. M. (1989) *Five Balltree Construction Algorithms*. Berkeley: International Computer Science Institute.

Segall, P., & Lu, S. (2015) ‘Injection-induced seismicity: Poroelastic and earthquake nucleation effects.’, *Journal of Geophysical Research: Solid Earth*, 120(7), pp. 5082-5103.

Sen, P. K. (1968) ‘Estimates of the regression coefficient based on Kendall's tau.’, *Journal of the American statistical association*, 63(324), pp. 1379-1389.

Shapiro, S. A., Rentsch, S., & Rothert, E. (2005) ‘Characterization of hydraulic properties of rocks using probability of fluid-induced microearthquakes.’, *Geophysics*, 70(2), F27-F33.

Shirzaei, M., Manga, M., & Zhai, G. (2019). Hydraulic properties of injection formations constrained by surface deformation. *Earth and Planetary Science Letters*, 515, 125-134.

Skoumal, R. J., et al. (2020) 'Induced seismicity in the Delaware Basin, Texas.', *Journal of Geophysical Research: Solid Earth*, 125(1), e2019JB018558.

Skoumal, R. J., Brudzinski, M. R., & Currie, B. S. (2015) 'Earthquakes induced by hydraulic fracturing in Poland Township, Ohio.', *Bulletin of the Seismological Society of America*, 105(1), pp. 189-197.

Szafranski, D., & Duan, B. (2018) 'Integrating Poroelastic Effects of Wastewater Injection and Rupture Dynamics to Understand Induced Seismicity', *Unconventional Resources Technology Conference*, July 2018, pp. 2557-2576.

Szafranski D., and Duan, B. (2020) 'Exploring physical links between fluid injection and nearby earthquakes: The 2012 M 4.8 Timpson, Texas, case study', *Bulletin of the Seismological Society of America*, 110 (5), 2350–2365.

Theil, H. (1950) 'A rank-invariant method of linear and polynomial regression analysis.', *Indagationes Mathematicae*, 12(85), 173.

Tibshirani, R. (1996) ‘Regression shrinkage and selection via the lasso.’, *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), pp. 267-288.

Wang, H. (2000) *Theory of Linear Poroelasticity With Applications to Geomechanics and Hydrogeology*. Princeton, N. J: Princeton Univ. Press.

Weng, H., & Yang, H. (2018) ‘Constraining frictional properties on fault by dynamic rupture simulations and near-field observations.’, *Journal of Geophysical Research: Solid Earth*, 123(8), pp. 6658-6670.

Yeck, W. L., et al. (2017) ‘Oklahoma experiences largest earthquake during ongoing regional wastewater injection hazard mitigation efforts’, *Geophys. Res. Lett*., 44.

# 5. INTEGRATED MODELING OF FLUID FLOW, HYDRAULIC FRACTURE PROPAGATION AND ASSOCIATED MICROSEISMICITY

## 5.1. Abstract

Hydraulic fracture propagation is commonly tracked using microseismic monitoring during the stimulation. However, not all microseismic events can be associated with Drained Reservoir Volume (DRV). We develop an integrated modeling methodology, verify our software, perform a series of numerical experiments on hypothetical models, and simulate a real case scenario in the Midland basin. We show that not only the differences in formations' properties but also presence of natural fractures can affect hydraulic fracture propagation, even without direct hydraulic connection. We demonstrate that the preferred direction of fracture propagation can change with time, which leads to complex fracture networks. We also present cases in which using Stimulated Rock Volume (SRV) as a proxy of DRV can either overestimate or underestimate that volume. We also show that natural fracture and bedding plane failures are equally possible mechanisms of microseismicity.

## 5.2. Introduction

Hydraulic fracturing technology, along with studies of fracture propagation, has been around for many decades. It finds its applications in many fields, including oil and gas production (e.g. Arthur et al., 2009; Pudugramam et al., 2021), geothermal energy production (e.g. Fehler, 1989; Legarth et al., 2005), rock burst prevention (e.g. Fan et al.,

2012) or water well production enhancement (e.g. Adams and Rowe, 2013). However, over past two decades it has been mostly associated with unconventional hydrocarbon exploitation from low permeability formations, such as tight sandstones or shales.

Rapid development of unconventional oil and gas fields goes along with the need to proper understanding of hydraulic fracture propagation process. The first way of getting more insight is doing extensive field experiments, in which the whole procedure is monitored with multiple instruments, and core and rock samples are extensively collected for detailed analysis, such as in the Hydraulic Fracturing Test Site - 1 (HFTS-1) experiment conducted in the Midland Basin (Texas, USA) (e.g. Ciezobka et al., 2018). While this approach provides a lot of information, it is very expensive, time consuming and might apply only to a certain basin. Second way to get better insight is laboratory experiments done on small rock samples from different fields (e.g. Bunger, 2008; Bunger and Lecampion, 2017). They are generally less expensive than field experiments but require a lot of time and are performed on very small samples, which presents a challenge when scaling up to the entire field development.

Third widely used approach in studies of hydraulic fracture propagation is modeling of the fracturing process using a variety of more or less sophisticated techniques (e.g. Lecampion et al., 2018). Fracture propagation simulations are used by companies in order to optimize the treatment setup, including well orientation, fracturing fluid selection, or pumping pressures and rates (e.g. Kong et al., 2015; Ghassemi, 2017).

Simulations allow for preliminary testing efficiency of different setups before the production phase even starts, improving cost-efficiency of the operations. Beside planning, the modeling can also provide a tool for retrospective analysis of the treatment, by matching with some field observations, thus providing feedback on the possible reasons for success or failure of the well.

Among hydraulic fracture modeling techniques there are three major groups – analytical, semi-analytical and numerical. Analytical models are represented by formulas derived for homogenous models with simple geometries of the fractures, describing basic features of the hydraulic fracture. Three most known models are KGD fracture model (Khristianovic and Zheltov, 1955; Geertsma and de Klerk, 1969), PKN fracture model (Perkins and Kern, 1961) and radial fracture model (e.g. Lecampion and Desroches, 2015). KGD fracture is described by a set of formulas for 2D plane-strain conditions, assuming the fracture height is much larger than the fracture length. PKN fracture is height-constrained fracture, with an assumption of much larger length than height. Unlike the previous two geometries, radial fracture is axisymmetric and penny-shaped. Until today they can be used to get quick, preliminary predictions on fracture geometry development in certain cases. However, for modern-day applications they assume overly-simplified subsurface model and predictions might be very misleading (e.g. Zhou et al., 2016). On the other hand, they work well as a verification tool for new, more sophisticated methods (e.g. Lecampion et al., 2018).

Among semi-analytical models it is worth to mention the pseudo 3D (P3D) model (Settari and Cleary, 1986), which is an extension of PKN model. In the P3D model, the fracture is divided into cells, each of which has assigned width and height based on local pore pressure and 3D problem is technically reduced to 2D. These models allow for fracture height growth prediction, but they often overestimate height when height containment is poor (Peshcherenko and Chuprakov, 2021). Despite not very good prediction accuracy, the P3D is often used in the industry due to its good computational efficiency.

The last and probably the biggest category are numerical methods. First subset is Boundary Element Methods (BEM) (e.g. Hossain and Rahman, 2008), which have an advantage of generally faster computations than Finite Element Methods (FEM). One of the most popular models of that type is the planar 3D model (e.g. Zia and Lecampion, 2020), which allows for quite accurate prediction of height and length of the fracture in a 3D layered medium. However, it is limited to layered media with homogenous layers and homogenous background stress.

The other large subset of numerical methods is Finite Element Methods (e.g. Hughes, 2000). Unlike BEM methods, they allow for a medium with any type of properties heterogeneity, background stress heterogeneity and geometry of the fracture (e.g. Lecampion et al., 2018). Due to high computational cost of these methods, researchers initially focused only on 2D simulations (e.g. Mendelsohn, 1984). Even nowadays many

studies of hydraulic fracture propagation limit the modeling to two dimensions with an assumption of a plane-strain model (e.g. Barba and Picano, 2020). However, 2D modeling misses an influence of multiple important factors that are important in proper simulation of hydraulic fracturing, e.g., height containment due to layering of the medium.

In some studies, computational cost of 3D FEM modeling is reduced by predefining the fracture path (e.g. Golovin et al., 2015). However, in practice the hydraulic fracture is rarely oriented exactly in the direction of maximum horizontal stress ($\sigma_{Hmax}$), and the geometry of the fracture path will be curved (e.g. Feng and Gray, 2018). Additionally, some stress heterogeneities caused by presence of other fractures or poroelastic effects might modify the initial trajectory of the fracture growth. There are four main ways for researchers to adapt simple FEM to handle unknown fracture path. First solution is the application of adaptive remeshing (e.g. Salimzadeh et al., 2016) that follows the geometry of the fracture. While they can replicate fracture geometry very accurately, they are computationally extremely expensive. Moreover, some researchers bring up the problem of potential mass conservation issues while using this technique (e.g. Secchi et al., 2007). Another possible solution is keeping original mesh but introducing some form of plastic deformations around the tip of the fracture. Most typical realization of that idea is the Cohesive Zone Model (e.g. Chen et al., 2009), allowing for capturing the process zone. While much more efficient computationally than adaptive remeshing and allowing for complex fracture geometry without mass conservation problem, that approach

presents a mesh dependency problem of the fracture geometry evolution. It can be alleviated by reduction of element size, but at the cost of computational efficiency. Third possible way of handling the problem is by using Extended Finite Element Methods (XFEM) (e.g. Gupta and Duarte, 2014). The idea is to enrich the elements located nearby discontinuity, by increasing the number of degrees of freedom, followed by the interpolation of the displacement field. It allows the fracture to propagate across the element, without being constrained to model grid. There are different methodologies that allow for coupling with fluid flow (e.g. Feng and Gray, 2017), or adding poroelasticity and cohesive zone (e.g. Wang, 2015). While that scheme works well in 2D simulations, it is a big challenge to extend it to 3D, and it requires more computational resources than simple Cohesive Zone Method. Next available approach is the implementation of Phase Field Methods (PFM) (e.g. Miehe et al., 2015). In these methods, continuous scalar variable (e.g., crack field) is introduced and the displacement field is smeared across several elements. It reduces the propagation path dependence on the mesh but at the same time have some major challenges. First, additional treatment is needed to distinguish between tensile and compressive fields. Second, for coupling with fluid flow, fracture width needs to be reconstructed which is a challenging task in PFM (e.g. Lecampion et al., 2018). Also, mesh requirements are harsh, which sometimes leads to significant refinement and increased computational cost. Finally, it is worth to mention one more recent trend to use some meshless methods, such as Smooth Particle Hydrodynamics (e.g. Douilett-Grellier et al., 2016), which were developed for complex

fluid dynamics but are still in development phase when it comes to application in hydraulic fracture modeling.

From the physics/mechanics perspective, the most classical models are based on Linear Elastic Fracture Mechanics (e.g. Rice and Drucker, 1967), in which the driving force is energy release rate, and both solid and fluid parts of the model are linear. Fluid flow is usually simulated only within a fracture using the cubic law (e.g. Bear,1972), assuming that the flow is between two parallel plates (fracture walls) with constant width. To handle fluid leak-off from the fracture due to certain level of surrounding formation permeability and fracture wall permeability, it is common practice to apply Carter's leak-off model (Howard and Fast, 1957), assuming small diffusion velocity relative to fracture propagation velocity, as well as small fracture net pressure in relation to magnitude of far field effective stress.

That classical model can be extended in multiple different ways. For example, Darcian flow and cubic law can be replaced by more sophisticated Reynolds equation for non-Darcian flow (e.g. Zhang et al., 2017). The flow only inside fracture might be extended to handle also a flow through the porous rock around by introduction of dual-porosity, dual-permeability model (e.g. Gerke and Van Genuchten, 1993). In this model, flow is separately handled within two different types of porosity – pore space in the rock and fractures. The flow between fractures and formation pore space is described as interporosity exchange flux, dependent on the difference of pressures and a constant

proportional to permeability. Introduction of dual-porosity, dual-permeability medium removes the need to use Carter's leak-off model. Another possible extension is introduction of poroelastic medium (e.g. Mehrabian and Abousleiman, 2014), in which the solid is deformable by the increased pressure/volume of the fluid, and the fluid has non-negligible compressibility (e.g., water, oil, or proppant). Finally, the elastic solid can be replaced by non-linear plastic solid, either locally around fracture as in Cohesive Zone Models (e.g. Needleman, 2014) or in the entire model.

Another important aspect connected with hydraulic fracturing operations and modeling of them is a way of model verification through some field observations. The most popular way of tracking hydraulic fracture extent during the treatment is monitoring of microseismicity (e.g. Grechka and Heigl, 2017). Most routinely, after detection of microseismic events, they are located with the use of derived velocity model for a given field with calculation of their magnitude (e.g. Warpinski, 2009). Next, the volume that they occupy is estimated, which is called Stimulated Reservoir Volume (SRV) and the half-length of the fracture is measured based on the distance from the perforation point (e.g. Mayerhofer et al., 2010). However, there are two problems with that approach. First, the microseismic events can occur not only on hydraulic fracture or natural fractures that hydraulic fracture hit, but also on natural fractures that were activated due to poroelastic stress change relatively far away from the hydraulic fracture (e.g. Tan et al., 2014). As a result, the volume that actually produces hydrocarbons, often referred to as Drained Reservoir Volume (DRV) (e.g. Nandlal and Weijermars, 2019), is

overestimated. Secondly, the mechanism of failure might make a difference in how conductive a given fracture is. Predominantly shear fractures might not have a significant aperture to transfer fluids from reservoir back to the well. The solution to these problems could be modeling of microseismicity and wave propagation. There are some works which focus on modeling of microseismic events standalone, without including modeling of hydraulic fracture propagation (e.g. Hobro et al., 2016). There are very limited attempts of microseismic clouds simulations using some geomechanical modeling of hydraulic fracture propagation including mapping the fracture failures around the fracture (He and Duan, 2021).

In this work, we develop an integrated methodology of coupled modeling of fluid flow in dual-porosity, dual-permeability poroelastic media with 3D hydraulic fracture propagation using the Cohesive Zone Method and generation of microseismicity. Within that framework we can simulate complex fracture systems including branching of hydraulic fracture, interactions with natural fractures and surrounding formations. The method also allows for testing any pumping schedule for multiple hydraulic fractures stimulated at the same time. We present here the method, benchmark problems against results of analytical and published numerical models, simulations on hypothetical, synthetic models and modeling for a model mimicking operations during the HFTS-1 project.

### 5.3. Methods

In this work, we integrate numerical modeling of fluid flow in a deformable porous and fractured medium, hydraulic fracture propagation and interaction with natural fractures and bedding planes in a 3D subsurface model, and microseismicity generation with a mixed-mode failure criterion. Figure 5.1 presents the workflow developed in this study.

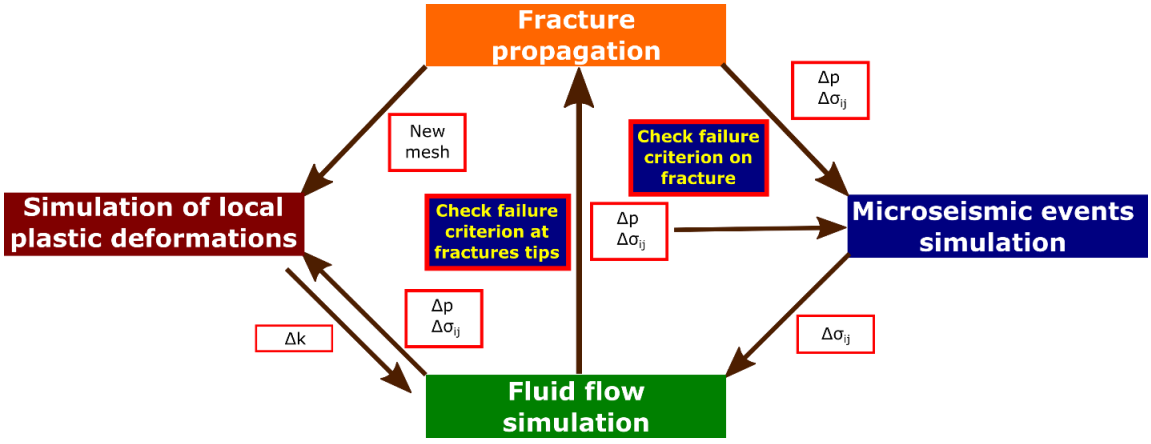


**Figure 5.1 Conceptual presentation of the integrated workflow.**

#### 5.3.1. Fluid flow modeling

For simulations of fluid flow, we combine flow through fracture with flow in porous formations into a single framework by introducing a dual-porosity, dual-permeability deformable medium. To fully describe the stress state in the subsurface, we need to solve a system of 8 partial differential equations – 2 equations for pore pressure inside pore space and fractures, and 6 equations for all independent stress tensor components. We start with general equations for fluid flow in a multi-porosity medium (e.g. Mehrabian and Abousleiman, 2014):

$$\frac{\overline{\alpha_i}}{K_i B_i}\left[\frac{B_i}{3}\frac{\partial \sigma_{ss}}{\partial t}+\frac{\partial p_i}{\partial t}\right]-\frac{1}{\mu}(\nabla k_i\cdot\nabla p_i+k_i\nabla^2 p_i)=Q+\sum_{j=1}^{N}\Gamma_{ij}(p_j-p_i), \qquad (5.1)$$

where $N$ corresponds to number of porosity types (it equals to 2 for dual-porosity), indices $i$ and $j$ indicate to which porosity a given quantity applies ($i = 1 \ldots N$), $\alpha_i$ are Biot-Willis coefficients, $B_i$ are Skempton's coefficients, $K_i$ are bulk moduli, $\sigma_{ss}$ is a trace of the stress tensor, $p_i$ are respective pore pressures, $\mu$ is fluid dynamic viscosity, $k_i$ are intrinsic permeabilities, $Q$ denotes a volume of fluid per unit bulk volume per unit time (external source of fluid), and $\Gamma_{ij}$ represent interporosity flows (for example from fractures to primary rock pore space), which are dependent on the pore pressure differences between different porosity networks. We define Biot-Willis coefficient same as Segall and Lu (2015):

$$\alpha = \frac{3(\nu_u - \nu)}{B(1+\nu_u)(1-2\nu)}, \tag{5.2}$$

where $\nu$ and $\nu_u$ are Poisson ratios for drained and undrained conditions respectively. In this study, we assume that medium has 2 types of porosity – primary pore space of a rock formation and fractures. Therefore, in our case (5.1) will take the following form for primary pore space:

$$\frac{\overline{\alpha_{mat}}}{K_{mat}B_{mat}}\left[\frac{B_{mat}}{3}\frac{\partial \sigma_{ss}}{\partial t} + \frac{\partial p_{mat}}{\partial t}\right] - \frac{1}{\mu}(\nabla k_{mat} \cdot \nabla p_{mat} + k_{mat}\nabla^2 p_{mat}) = Q - c_l\big(p_{mat} - p_{frac}\big), \tag{5.3}$$

where subscripts *"mat"* and *"frac"* correspond to values of given parameter/quantity for primary matrix porosity and fractures respectively, and $c_l$ is interporosity exchange coefficient. Similarly, for fractures the equation (1) takes the form of

$$\frac{\overline{\alpha_{frac}}}{K_{frac}B_{frac}}\left[\frac{B_{frac}}{3}\frac{\partial \sigma_{kk}}{\partial t} + \frac{\partial p_{frac}}{\partial t}\right] - \frac{h_f}{12\mu}\big(\nabla(e^3) \cdot \nabla p_{frac} + (e^3)\nabla^2 p_{frac}\big) = Q - c_l\big(p_{frac} - p_{mat}\big) + \frac{\partial e}{\partial t}, \tag{5.4}$$

where $h_f$ is local fracture height and $e$ denotes fracture aperture.

We also solve six equations for stress tensor components using Beltrami-Mitchell equations (e.g. Wang, 2000) adapted for multi-porosity media:

$$\nabla^2\sigma_{ij} + \frac{1}{1+\nu}\frac{\partial^2\sigma_{ss}}{\partial x_i \partial x_j} + 2\eta\sum_{m=1}^{N}\eta_m\left[\frac{1-\nu}{1+\nu}\frac{\partial^2 p_m}{\partial x_i \partial x_j} + \delta_{ij}\nabla^2 p_m\right] = -\frac{\nu}{1-\nu}\delta_{ij}\,\nabla\cdot\boldsymbol{F} - \frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i}, \tag{5.5}$$

where $\sigma_{ij}$ denotes a stress tensor, $\delta_{ij}$ is the Kronecker delta, $\eta_m$ represents poroelastic stress coefficients, $p_m$ are pore pressures for different pore networks, $N$ denotes the number of different pore networks, and $\boldsymbol{F}$ is an external body force. In our case, $N$ is equal to 2 since we consider a dual-porosity medium. Moreover, because we are interested mostly in perturbations of pore pressures and stresses, the right-hand-side term in equation (5) is equal to zero, which greatly simplifies the calculations. Thus, the final form of the equation for stress tensor has the following form:

$$\nabla^2\sigma_{ij} + \frac{1}{1+\nu}\frac{\partial^2\sigma_{kk}}{\partial x_i \partial x_j} + 2\left[\eta_{mat}\left(\frac{1-\nu}{1+\nu}\frac{\partial^2 p_{mat}}{\partial x_i \partial x_j} + \delta_{ij}\nabla^2 p_{mat}\right) + \eta_{frac}\left(\frac{1-\nu}{1+\nu}\frac{\partial^2 p_{frac}}{\partial x_i \partial x_j} + \delta_{ij}\nabla^2 p_{frac}\right)\right] = 0. \tag{5.6}$$

In general, the poroelastic stress coefficient $\eta$ is defined as (Wang, 2000):

$$\eta = \frac{1-2\nu}{2(1-\nu)}\alpha. \tag{5.7}$$

### 5.3.2. Failure criterion

When integrating fluid flow with hydraulic fracture propagation, an essential part of the process is a selection of the failure criterion, which indicates when the stress or pore pressure level on a fracture reaches its yield point. In this study, we assume that the fracture can reach its failure via tensile opening, shear or a combination of both. To capture that phenomenon, we use a mixed-mode quadratic failure criterion (e.g. Comanho and Davila, 2002; Feng and Gray, 2017):

$$\left[\frac{\langle t_n \rangle}{t_n^0}\right]^2 + \left[\frac{t_s}{t_s^0}\right]^2 + \left[\frac{t_d}{t_d^0}\right]^2 \leq 1, \tag{5.8}$$

where $t_n^0$, $t_s^0$ and $t_d^0$ are tensile and two shear strengths (that most commonly we assume to be the same) of the rock, and $t_n$, $t_s$ and $t_d$ are tractions in normal, shear and 2nd shear directions (along-strike and along-dip). $\langle t_n \rangle$ means that only positive (tensile) effective normal stress is considered in the opening failure process. In this paper, we assume that the tensile stress needs to overcome fracture toughness $K_{IC}$ to cause the opening and for shearing we use Mohr-Coulomb criterion:

$$\tau \leq c + \mu_s(\sigma_n - \alpha p), \tag{5.9}$$

where $\tau$ is shear stress on the plane at given orientation, $c$ is cohesion of the material, and $\mu_s$ is static friction coefficient.

Additional benefit of using the mixed-mode criterion is a possibility of source mechanism reconstruction, which can be useful if we have source mechanism solutions available for registered microseismic events.

### 5.3.3. Hydraulic fracture propagation modeling

One of the biggest challenges in numerical modeling of hydraulic fracture propagation is dealing with moving boundary of the fracture. Every time we reach a yield point of the given rock at the tip of the fracture, we add the new element or grid point for the next time step of calculations. Unlike analytical and semi-analytical methods, where we often assume certain velocity of fracture propagation, in numerical methods, typically hydraulic fracture can propagate with discrete steps.

To limit the problem of too rapid jumps in fracture length and width (from zero to fully opened), we use a Cohesive Zone Method, in which we treat a region around fracture as plastic material. In the process of material degradation, we track the changes of its stiffness:

$$K_d = (1 - D)K_0, \tag{5.10}$$

where $K_d$ is material stiffness at the given point in time, $K_0$ is material initial stiffness, and $D$ is damage variable defined as

$$D = \frac{\delta_m^f(\delta_m^{max} - \delta_m^0)}{\delta_m^{max}\left(\delta_m^f - \delta_m^0\right)}, \tag{5.11}$$

where $\delta_m{}^f$, $\delta_m{}^0$ and $\delta_m{}^{max}$ are effective displacements at complete failure, at initiation damage, and at maximum damage during loading respectively. The effective displacement $\delta_m$ we define as

$$\delta_m = \sqrt{\langle\delta_n\rangle^2 + \delta_s^2 + \delta_d^2}, \tag{5.12}$$

where $\delta_n$, $\delta_s$ and $\delta_t$ are displacement in normal, 1st and 2nd shear directions respectively. Notice, we only take into account tensile displacements to prevent non-physical effects (fracture walls interpenetration). Besides degradation of material stiffness, the displacement on fracture also affects the traction/stress on that fracture, which can be mathematically described by a piece-wise linear function

$$t_k^{max} = \begin{cases} (1-D)\bar{t_k}\,, \bar{t_k} \geq 0 \\ \bar{t_k}\,, otherwise \end{cases}, \tag{5.13}$$

where $k$ represents a component of traction (normal or shear), $\bar{t_k}$ is fracture interface stress/traction at a given time and $t_k^{max}$ is interface stress/traction at the maximum opening or shear displacement.

In our fracture propagation approach, we allow a fracture to propagate in 3 different fashions:

- Propagate along dip direction
- Propagate horizontally along strike direction
- Propagate horizontally at the direction different from the strike (across the elements)

To check whether fracture continues propagation along strike or changes direction, at the tip of the fracture we test failure criterion at 3 different angles. If more than one direction would allow for failure, we pick the one in which stress state is the closest to failure. When the fracture branches, we treat new segment as a new fracture which shares nodes

with original fracture, which allows to treat fluid flow within the fracture system as a piece-wise 2D problem.

Since the fracture is allowed to propagate along the path that is not predefined, we are adding new cohesive elements as the fracture system develops, i.e., a new cohesive region is added after it reaches yield point.

### 5.3.4. Numerical implementation

In this study, we solve equations (5.3), (5.4) and (5.6) using finite-difference method implemented in in-house code PyFluFlow (v. 2.0). The detailed description of the single-porosity version of the code, along with the code verification are in Szafranski and Duan (2020). In the new version, however, there are some additional steps:

- After calculation of pore pressure distribution in the formations (primary pore network), we transfer that information on local mesh and calculate pore pressure distribution within the fracture. The calculation is done in 2D (along-strike and along-dip).
- Having pore pressure perturbation distributions from both formations and fractures, we calculate the stress changes in the whole medium.
- After that we iterate the process until convergence, calculating pore pressures and stress perturbations

Fracture propagation simulation is implemented also in a finite-difference scheme as an extension to PyFluFlow. After computation of pore pressure and stress perturbations, we

resolve total stresses on each fracture and check failure criterion on the fracture tips. If the criterion is met, we modify the mesh for pore pressure calculations within the fractures.

After checking failure criterion, we also calculate changes of fracture displacements (e.g. fracture aperture), resulting stress changes and material strength drop. If there is a failure, we first add new cohesive zone elements. Finally, using failure area and average displacement on the fracture, we calculate released seismic moment and moment magnitude of generated microseismic events.

### 5.4. Code verification

#### 5.4.1. Models

To verify our code, we use two benchmark problems. First test is against PKN and KGD analytical models for the boundary case of M-vertex (viscosity-storage-dominated), assuming high viscosity (here equal to 0.1 Pa·s), negligible fracture toughness (we use $K_{IC}$ = 10 Pa) and very negligible leak-off effect (we assume equivalent of fracture wall permeability of $10^{-25}$ $m^2$). We consider an elastic medium with Poisson ratio of 0.25 and Young's modulus of 30 GPa. We also assume constant injection rate of 0.053 $m^3/s$. In the simulation we constrained fracture height at 50 m and we compare development of fracture half-length.

For the second benchmark we pick a more complex 3-layer model ("Test 7" model) modified from Peshcherenko & Chuprakov (2021), where the injection is conducted

within a middle layer which has lower fracture toughness than the surrounding layers. Other physical properties of these layers are in Table 5.1. The injection rate is constant and equal to 0.053 $m^3/s$ with fluid viscosity of 0.2 Pa·s. Injection point is located in the center of middle layer.

**Table 5.1 Physical properties of the layers in 3-layer benchmark model ("Test 7" model) modified from Peshcherenko & Chuprakov (2021)**

| Layer | Layer thickness [m] | ν | $\nu_u$ | B | k [$m^2$] | Y [GPa] | $P_p$ [MPa] | $\sigma_{hmin}$ [MPa] | $K_{IC}$ [MPa·$m^{1/2}$] |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 210 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 30.0 | 50.0 | 50.0 | 4.5 |
| **2** | 60 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 30.0 | 50.0 | 50.0 | 1.5 |
| **3** | 100 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 30.0 | 50.0 | 50.0 | 4.5 |

### 5.4.2. Results

Comparison of numerical solution to analytical solution is one of the ways to verify that a code works properly. Figure 5.2 shows the comparison of PyFluFlow solution against PKN and KGD solutions obtained for a homogenous model described in the previous section. Even though our simulator predicts longer fracture in the first few hundred seconds, the final half-length lies in between PKN and KGD predicted half-lengths. The difference might be the result of imperfect conformity of our model with the assumptions of analytical solutions, such as much longer length than height for PKN solution or non-zero fracture toughness.

The second conducted comparison is performed on the 3-layer model with surrounding layers having larger toughness than middle layer. This time we compare both fracture half-

length and height changes with time (see Figure 5.3). We get very similar half-length to Peshcherenko and Chuprakov (2021) result from their simulator Razor. However, there is a visible fracture height difference. Our result is more affected by higher fracture toughness values in surrounding formations (good containment). However, after enough time for pressure accumulation, the seal is being broken and fracture continue to propagate.

The difference can be explained by the back-stress effect which was observed by Carrier and Garnet (2012), who show that fracture extent can be several times smaller if the pore pressure-stress coupling is taken into account. In our case the difference is not as significant since we use values of poroelastic parameters which are a proxy for elastic model used by Chuprakov and Peshcherenko (2021). A complete reduction to purely elastic model would be difficult and time consuming since we would need to change the equations to be solved.

Looking at the verification results, we can say that PyFluFlow simulator provides comparable estimates to analytical solutions and solutions for more complicated models published in the literature, with differences coming from additional physics effects uncaptured by simpler simulators and approximation of some parameters non-existent in these simulators.

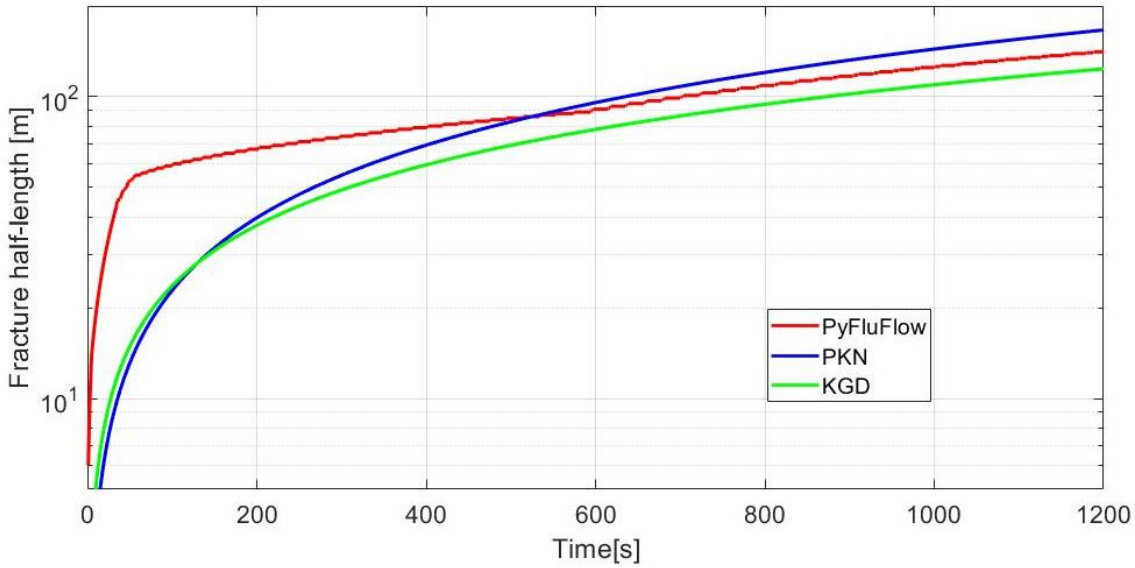


**Figure 5.2 Comparison of fracture half-length development in time between our numerical simulation, and analytical PKN and KGD models.**

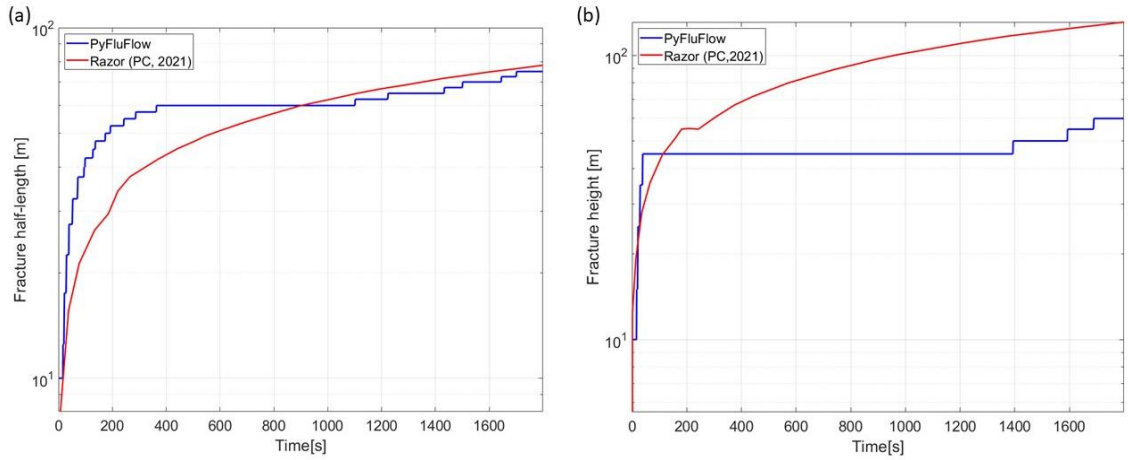


**Figure 5.3 Comparison of fracture (a) half-length and (b) height development in time between our numerical simulation PyfluFlow, and results from simulation in Razor in Peshcherenko and Chuprakov (2021).**

## 5.5. Hydraulic and natural fractures interactions

### 5.5.1. Models

Beside code verification models, we also tested performance of our code on multiple simple models in which hydraulic fracture interacts with natural fractures present in the medium. By default, we use a layered model with 3 layers, where the hydraulic fracture is initiated in the center of the middle layer. Most important properties of these layers are summarized in Table 5.2.

We assume azimuth of maximum horizontal stress to be at 90° and strike-slip faulting regime ($\sigma_{Hmax}/\sigma_v$=1.2 and $\sigma_{hmin}/\sigma_v$=0.8). For simplicity, pore pressures and stresses are uniform within each layer. We also assume that cohesion of intact rock in each layer (which affects the shearing strength) is much higher than the tensile strength so that we avoid propagation of fracture by pure shearing. All existing fractures, however, have relatively low cohesion of 1MPa, which allows for their activation in a shearing mode. Permeability of the fractures is aperture-dependent, which in turn is dependent on the stress state at a given time. We also assume that leak-off from the fractures to the medium is almost negligible and corresponding to permeability of $10^{-25}$ $m^2$. The viscosity of the fracturing fluid in our models is 0.1 Pa·s and we assume constant injection rate similar to the benchmark models of 0.053 $m^3/s$.

In Setup 1, we consider a system with a single hydraulic fracture and 2 sets of natural fractures – with strikes of 0° and 90° (perpendicular and parallel to the hydraulic fracture orientation). Figure 5.4 shows that setup in map (a) and 3D (b) views. In case of hydraulic fracture propagation along $\sigma_{Hmax}$ azimuth, none of the natural fractures is

going to be intersected by the hydraulic fracture. Therefore, any microseismic events would be considered “dry” events. In Setup 2, we have a set of natural fractures with strikes of 45°, which is more favorable orientation for triggering shear events. Figure 5.5 shows that setup in a map (a) and 3D (b) view. Finally, in the Setup 3 we have natural fractures with different strike orientations, which can be intersected by hydraulic fracture. This allows us to see how frac hit against natural fracture can influence the growth of the hydraulic fracture. Figure 5.6 presents the geometry of that setup in map view (a) and 3D view (b).

**Table 5.2 Physical properties of the layers in 3-layer synthetic models.**

| Layer | Depth range [m] | $\nu$ | $\nu_u$ | B | k [$m^2$] | K [GPa] | $\rho$ [kg/$m^3$] | $P_p$ [MPa] | $\sigma_{hmin}$ [MPa] | $\mu_s$ | $K_{IC}$ [MPa·$m^{1/2}$] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 1650-1845 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 20.0 | 2650 | 50.0 | 50.0 | 0.6 | 5.0 |
| **2** | 1845-1895 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 20.0 | 2650 | 50.0 | 50.0 | 0.6 | 1.0 |
| **3** | 1895-2000 | 0.25 | 0.26 | 0.01 | $10^{-25}$ | 20.0 | 2650 | 50.0 | 50.0 | 0.6 | 5.0 |

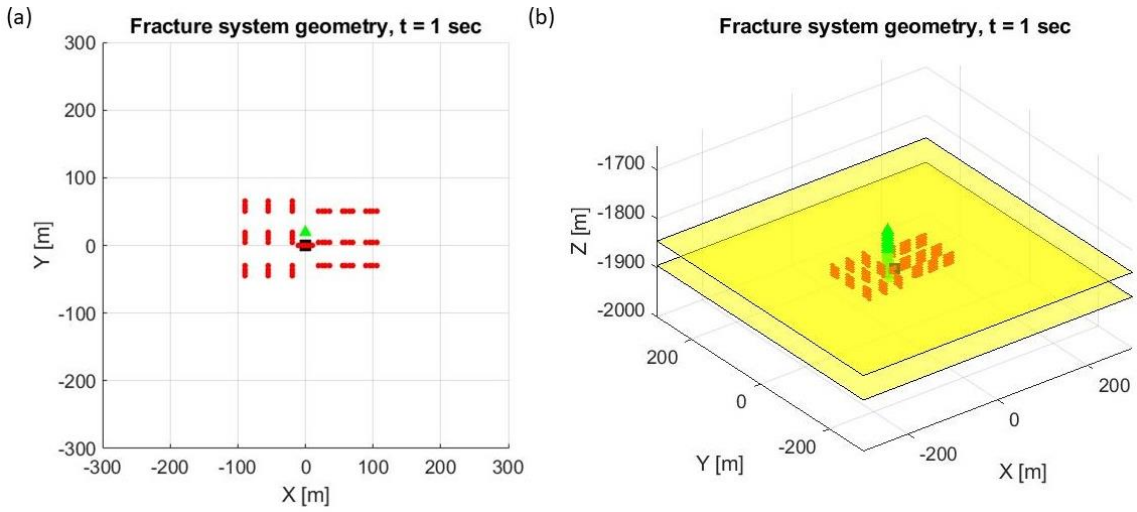


**Figure 5.4 (a) Map view and (b) 3D view on the Setup 1 of fracture system geometry. Red planes are fractures, black square is the perforation point, green triangles are simulated monitoring stations and yellow planes are layer boundaries.**

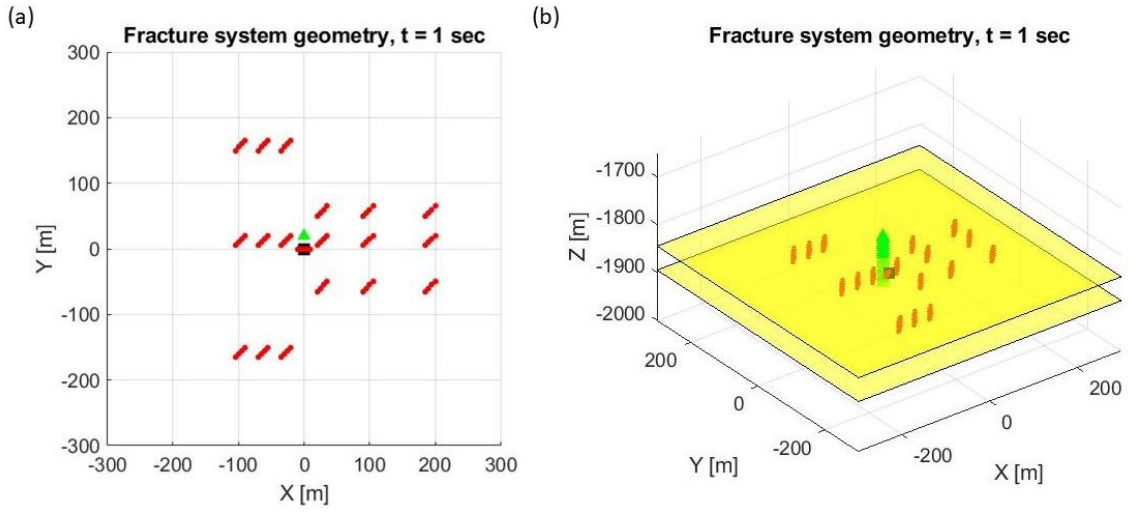


**Figure 5.5 (a) Map view and (b) 3D view on the Setup 2 of fracture system geometry. Red planes are fractures, black square is the perforation point, green triangles are simulated monitoring stations and yellow planes are layer boundaries.**

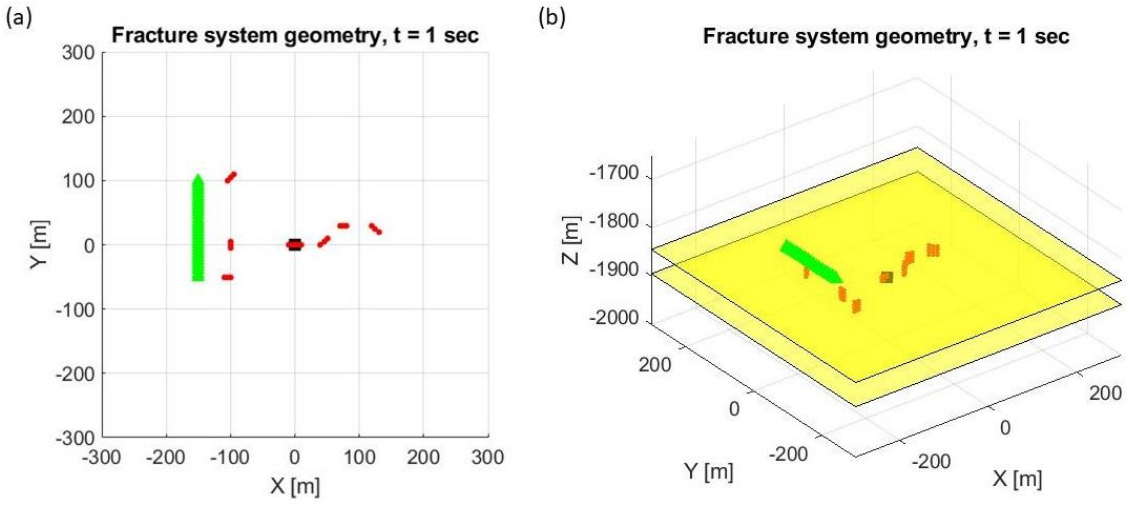


**Figure 5.6 (a) Map view and (b) 3D view on the Setup 3 of fracture system geometry. Red planes are fractures, black square is the perforation point, green triangles are simulated monitoring stations and yellow planes are layer boundaries.**

### 5.5.2. Results

One of the main objectives of this study is to investigate the interactions between hydraulic fracture and natural fractures present in the medium. Even though in fracture system geometry Setup 1 we do not have any natural fractures directly on the expected pathway of the hydraulic fracture, we can see from Figure 5.7 that there is a significant asymmetry between eastern and western side of the hydraulic fracture. If we look at the fracture growth and associated microseismic cloud development (Figures 5.8 and 5.9), we notice that the longer limb of the fracture is on the side where we have natural fractures that are parallel to the hydraulic fracture and are activated during its propagation. On the other hand, shorter limb is on the side where natural fractures are perpendicular to hydraulic fracture and are not activated until later time. If activated natural fractures are close to propagating hydraulic fracture (as in Setup 1), a positive

feedback loop can be formed between hydraulic fracture and natural fractures and stress changes caused by failure on one can result in big enough perturbations on the other so that it also fails.

Another observation that we can derive from Figures 5.8 and 5.9 is discrepancy between the extent of hydraulic fracture and the microseismic cloud. On one hand, if natural fractures are favorably oriented and close to failure, they might be activated by propagation of hydraulic fracture by far-field stress transfer. As a result, SRV would be large, when DRV is actually smaller. On the other hand, hydraulic fracture might propagate through some volume and produce undetectable microseismicity if the natural fractures are not favorably oriented and/or are far from failure.

We also run another simulation for Setup 1, this time using homogenous model with properties of the middle layer to observe the effect of heterogeneity on horizontal and vertical propagation of the hydraulic fracture. Figure 5.7 shows that there is no difference in fracture length, but pronounced difference in height. The period when hydraulic fracture is contained within the middle layer causes in the end a shorter vertical extent by few tens of meters.

In results for fracture system geometry Setup 2, where again we do not expect hydraulic fracture to hit any natural fracture but the fractures have the same strike on both sides of the perforation, we see almost perfect symmetry between the eastern and western limbs

of the hydraulic fracture (see Figure 5.10 (a)). Vertical propagation is again close to symmetrical as well. Looking at Figure 5.11 and 5.12 we see however that more microseismic events are generated on the western side of perforation, even though the distance to some of the fractures is larger there than that on the eastern side. This shows that relative location of natural fractures and hydraulic fracture (including strike angle) may have effect on the density of microseismic cloud and its shape. It adds additional complexity in interpretation of microseismic because we often have also a bias and asymmetry of the locations distribution due to positioning of monitoring stations.

In fracture system geometry Setup 3 we include fractures of different orientations, including two which are located along the expected path of hydraulic fracture propagation at different distances. However, from Figure 5.13 we cannot see any significant and long-lasting asymmetry in hydraulic fracture propagation either horizontally or vertically. This is due to the fact that the hydraulic fracture continues its growth along original strike even when it hits natural fractures on its way. From Figure 5.14 we see that after being hit by the hydraulic fracture, the two natural fractures rapidly grow vertically. With time the whole system follows vertical growth of the main hydraulic fracture. Another interesting observation is also small horizontal growth of one of the natural fractures after hydraulic fracture reaches quite significant length and its pace of growth slows down. It means with time natural fractures can become a preferred pathway of system growth.

For Setup 3 we run also two modifications of the original model. First, assuming a homogenous model with properties of the middle layer. Second, making the surrounding layers weaker (lower toughness than the middle layer), which can mimic a “runaway” fracture scenario. In Figure 5.15 we can see that while there is no significant difference in fracture half-length between the homogenous and original high $K_{IC}$ models, the length in low $K_{IC}$ case is much larger. The final height difference is even more pronounced for the end member models. In Figures 5.16 and 5.17 we have a confirmation of our observations that hydraulic fracture hits both natural fractures on its way but after short-term halt, it continues its propagation without changing direction. In Figure 5.16 (c) the propagation of Fracture 2 in NE direction is also visible. Possibly for a longer simulation it might eventually hit another natural fracture and increase the DRV size. Even though in this model we mostly focus on different aspects of system behavior, it is worth to notice that all but one fracture is activated early in the stimulation process. The fracture with 0° strike did not produce any shearing events even after being hit by hydraulic fracture and growing vertically. Additionally, in Figure 5.17 (f) we see increase in events’ magnitudes which agree with typical observations that the level of microseismic activity increases later during stimulation.

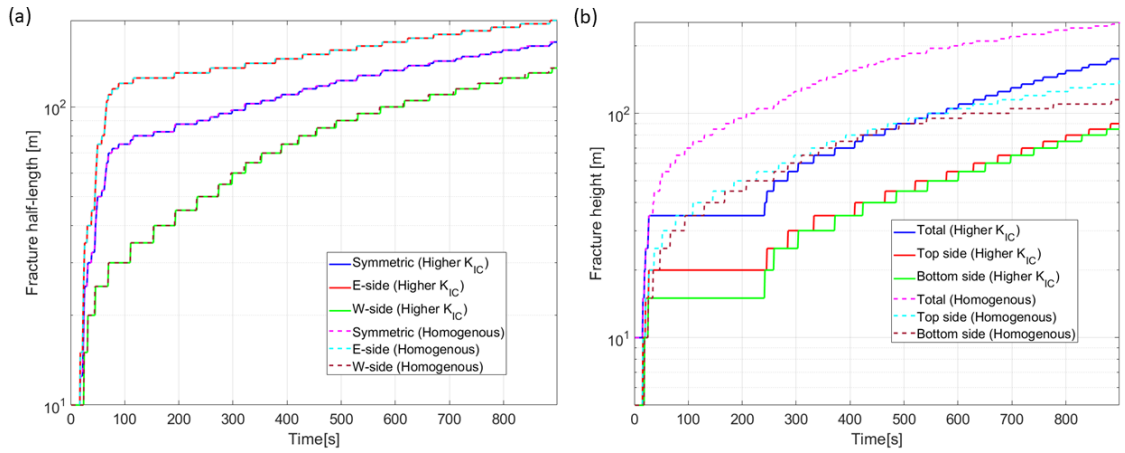


**Figure 5.7 (a) Fracture half-length changes with time (fracture geometry system Setup 1) for assumption of symmetric growth, as well as asymmetric growth of eastern and western side of the hydraulic fracture for default 3-layer model as well as homogenous model with properties of middle layer; (b) fracture height changes with time (total height as well as "semi-height" for top and bottom halves) for default 3-layer model as well as homogenous model with properties of middle layer.**

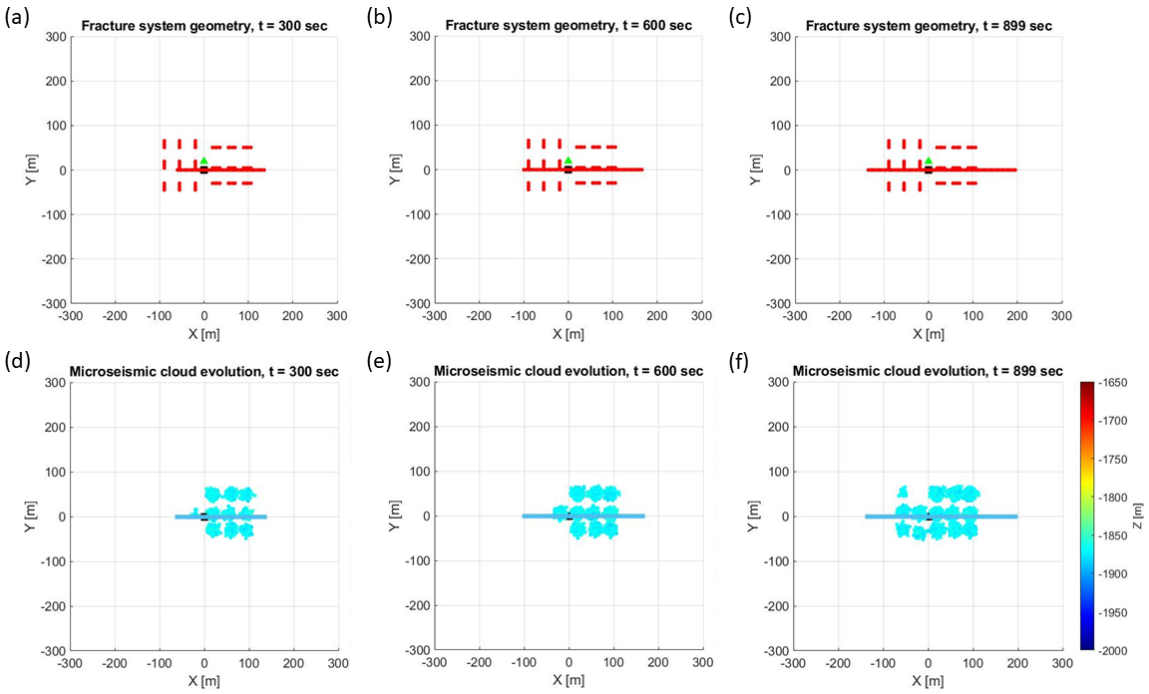


**Figure 5.8 Map view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment ((d) – (f)) for fracture system geometry Setup 1. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red lines are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) blue line represents hydraulic fracture, black square is the perforation point and colorful dots represent microseismic events with depth indicated by their color.**

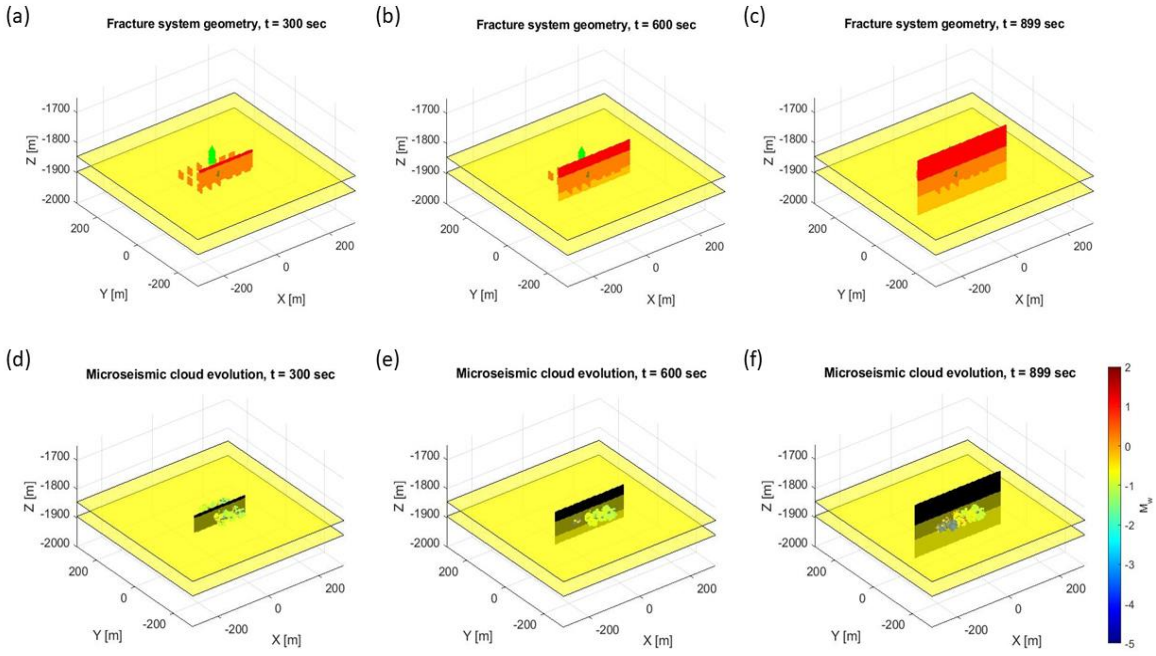


**Figure 5.9 3D view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment ((d) – (f)) for fracture system geometry Setup 1. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red planes are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) black plane represents hydraulic fracture and colorful dots represent microseismic events with magnitude indicated by their color.**

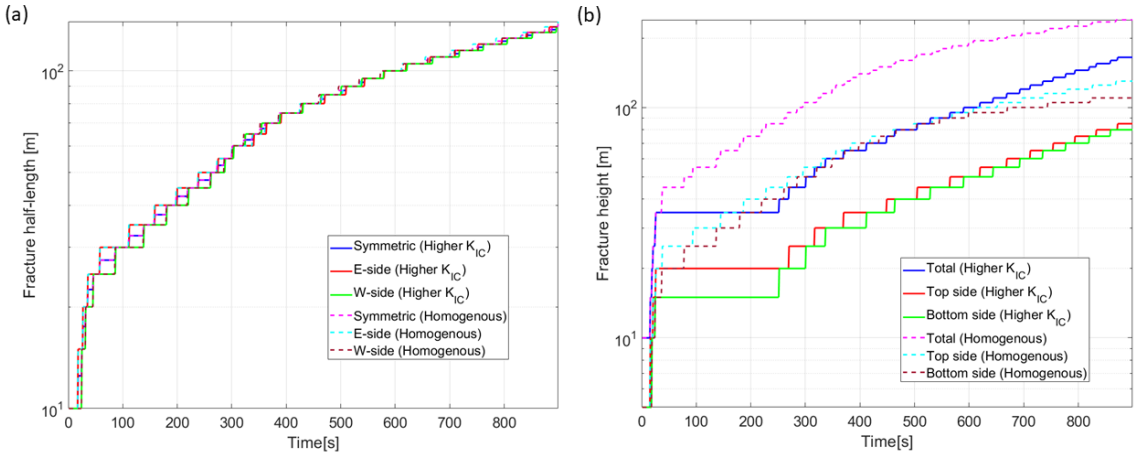


**Figure 5.10 (a) Fracture half-length changes with time (fracture geometry system Setup 2) for assumption of symmetric growth, as well as asymmetric growth of eastern and western side of the hydraulic fracture for default 3-layer model as well as homogenous model with properties of middle layer; (b) fracture height changes with time (total height as well as "semi-height" for top and bottom halves) for default 3-layer model as well as homogenous model with properties of middle layer.**

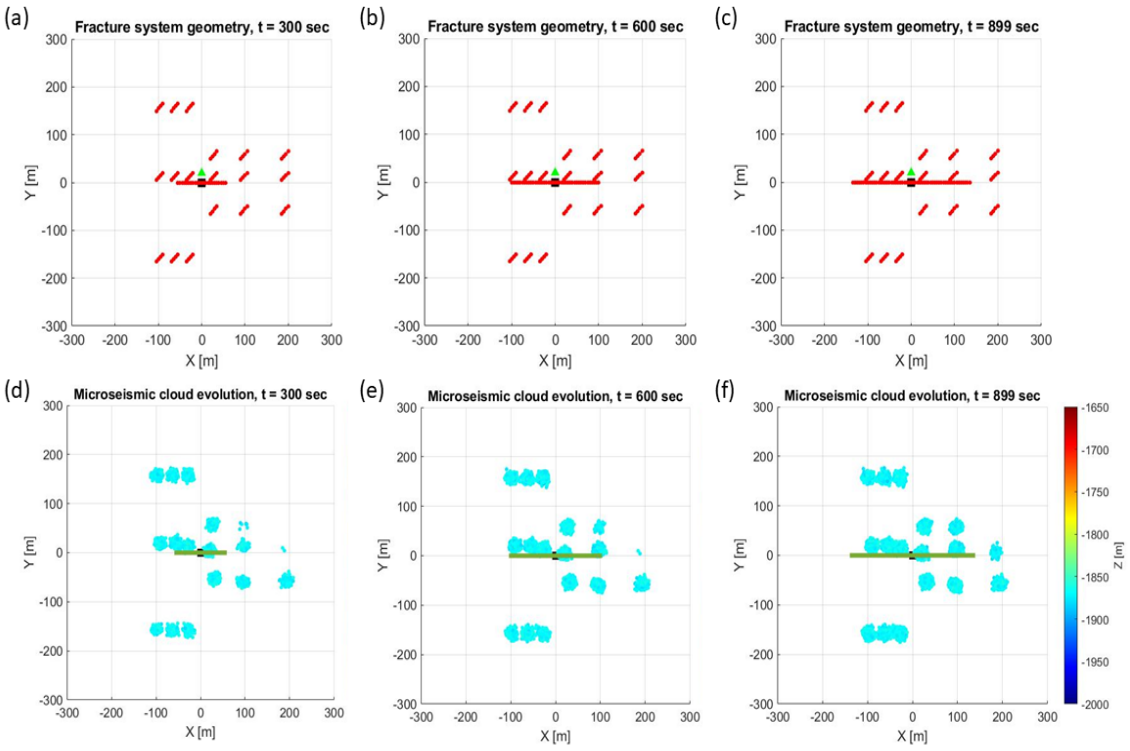


**Figure 5.11 Map view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment ((d) – (f)) for fracture system geometry Setup 2. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red lines are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) green line represents hydraulic fracture, black square is the perforation point and colorful dots represent microseismic events with depth indicated by their color.**

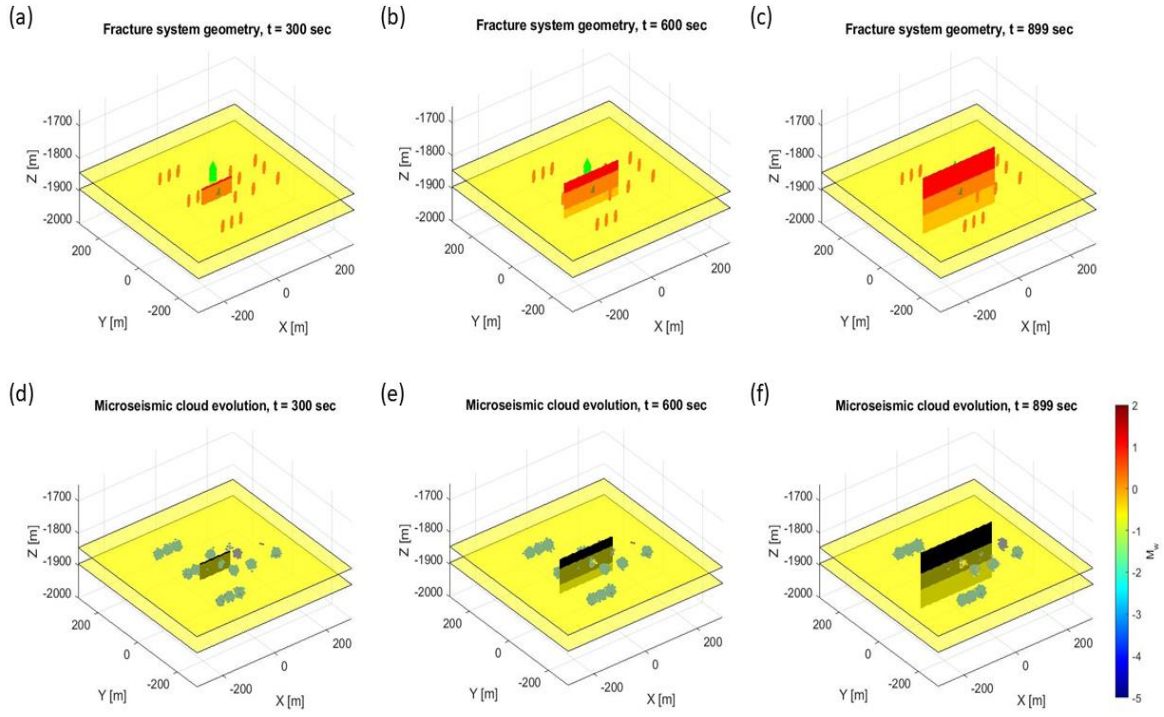


**Figure 5.12 3D view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment ((d) – (f)) for fracture system geometry Setup 2. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red planes are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) black plane represents hydraulic fracture and colorful dots represent microseismic events with magnitude indicated by their color.**

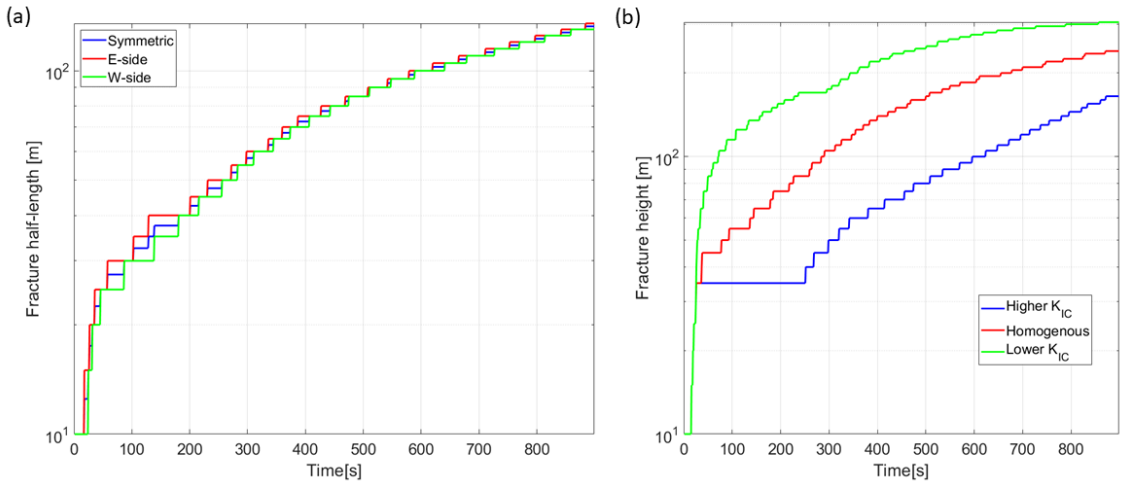


**Figure 5.13 (a) Fracture half-length changes with time (fracture geometry system Setup 3) for assumption of symmetric growth, as well as asymmetric growth of eastern and western side of the hydraulic fracture for default 3-layer model; (b) fracture height changes with time (total height as well as "semi-height" for top and bottom halves) for default 3-layer model.**

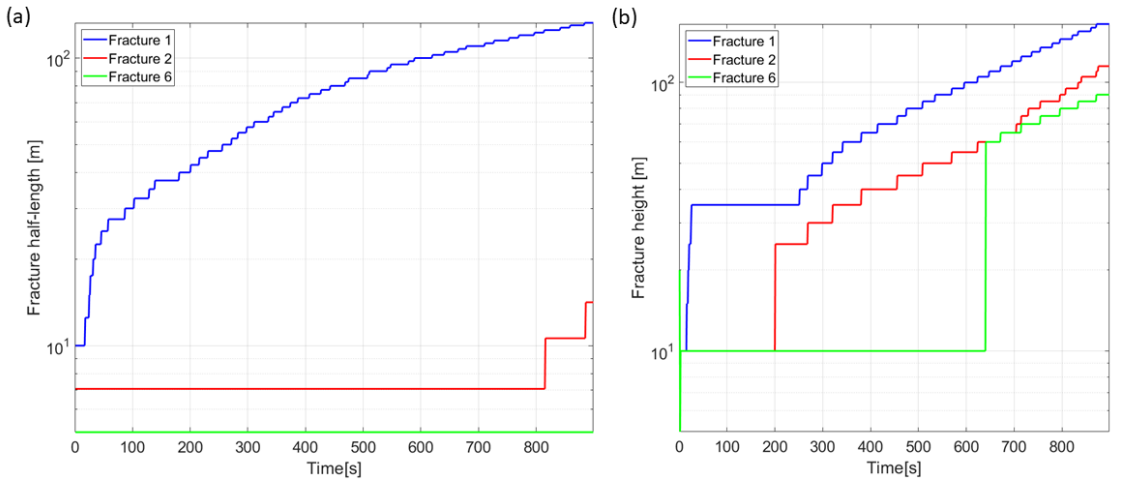


**Figure 5.14 (a) Fractures half-length and (b) fractures height changes with time (fracture geometry system Setup 3) under assumption of symmetric growth for hydraulic fracture (Fracture 1) and two natural fractures (Fracture 2 and Fracture 6) which were hit by the hydraulic fracture during its propagation.**

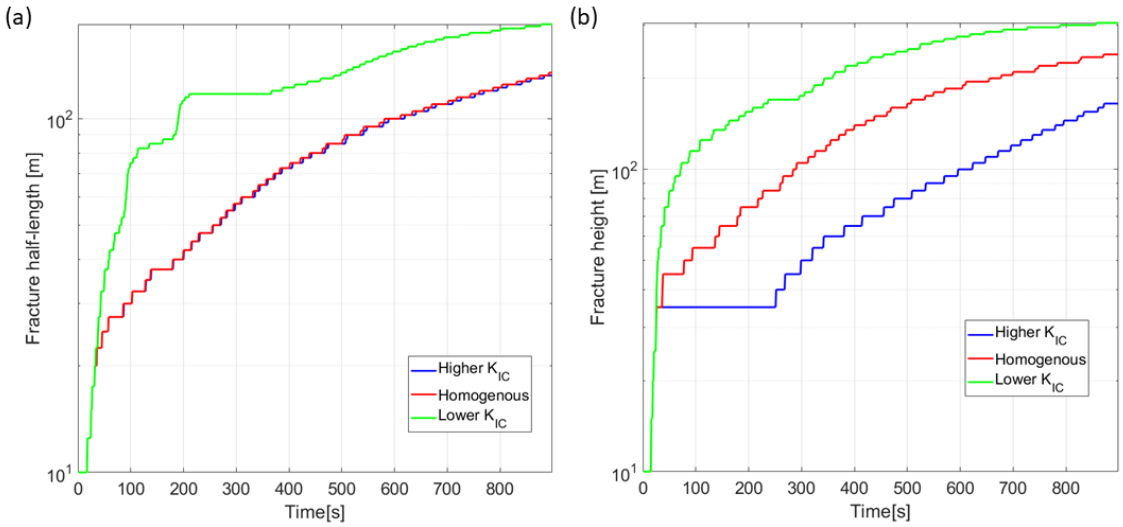


**Figure 5.15 Comparison of hydraulic fracture (a) half-length (under symmetric growth assumption) and (b) height for default 3-layer model with higher fracture toughness for surrounding layers, 3-layer model with lower fracture toughness for surrounding layers (0.1 MPa·m$^{1/2}$) and homogenous model with properties of middle layer.**

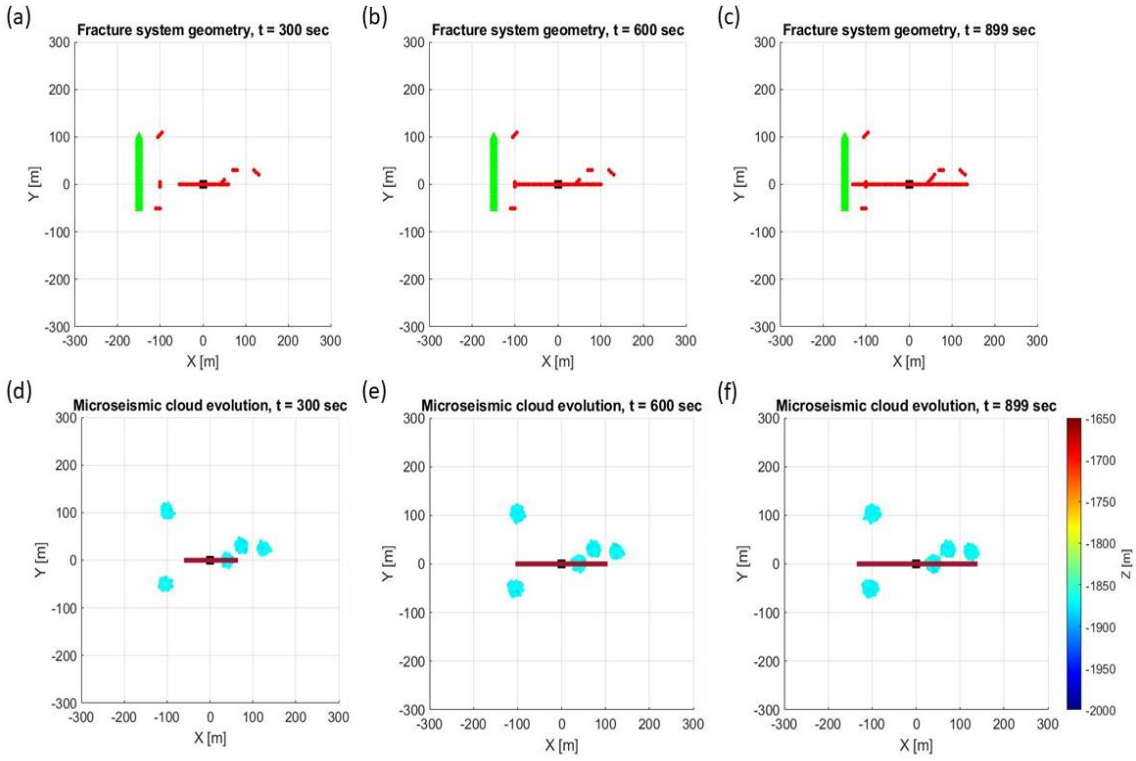


**Figure 5.16 Map view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment((d) – (f)) for fracture system geometry Setup 3. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red lines are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) – (f) maroon line represents hydraulic fracture, black square is the perforation point and colorful dots represent microseismic events with depth indicated by their color.**

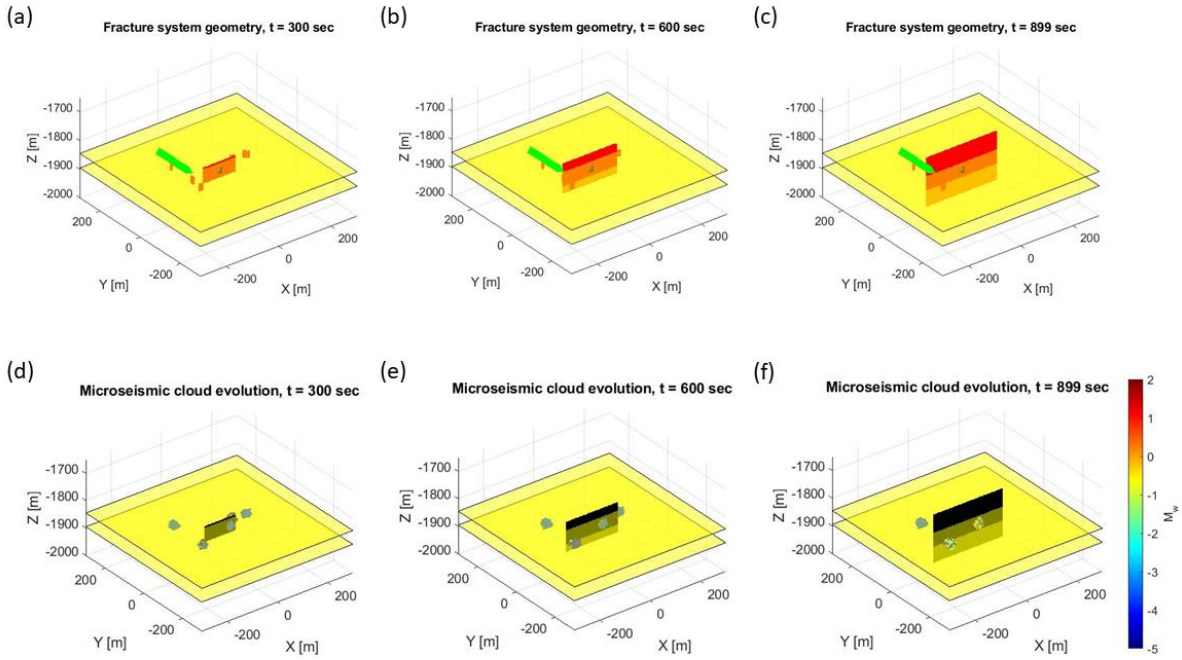


**Figure 5.17 3D view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment((d) – (f)) for fracture system geometry Setup 3. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red planes are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) black plane represents hydraulic fracture and colorful dots represent microseismic events with magnitude indicated by their color.**

## 5.6. Application toa field case of HFTS-1

### 5.6.1. Model

In this section, we apply the developed method to a field case. We construct a simplified model mimicking the geological model of the Hydraulic Fracturing Test Site - 1 (HFTS-1) located in Reagan County (TX) in Midland Basin. In this comprehensive experiment, there was a stimulation of 11 lateral wells with over 400 stages in Upper and Middle Wolfcamp formations (Ciezobka et al., 2018). During the experiment, a lot of valuable pieces of data were acquired, including microseismic data, well-logs, long section of

core or samples from side-wall cores. We use different parts of that dataset to build our simplified geomechanical model for simulation and to better understand the characteristics of microseismicity. We also utilized works published by the team participating in that collaboration, including analysis of microseismic dataset (Stegent and Candler, 2018; Wang et al., 2019; Maity, 2018) or analysis of core data from slanted well (Maity et al., 2018).

Our simplified model has 6 layers (see Table 5.3). The richest data coverage is available for Upper and Middle Wolfcamp. Based on well-log and laboratory measurements on rock samples we constrained elastic moduli, fracture toughness and cohesion for intact rocks. Permeability information was sparse but consistently indicating very impermeable rock. Using the dataset and supplementing with previous studies on Permian Basin (e.g. Friedrich and Monson, 2013; Snee and Zoback, 2018) we approximately constrain the stress state in the target interval. We know that Upper, Middle and Lower Wolfcamp are significantly overpressured, whereas Lower Spraberry formation is underpressured due to many years of exploitation (Friedrich and Monson, 2013). The azimuth of $\sigma_{Hmax}$ is N87°E and parameter describing faulting regime $A_{\varphi}$=0.81 indicating normal faulting (Snee and Zoback, 2018). Based on these pieces of information we define stress ratios of $\sigma_{Hmax}/\sigma_v = 0.9335$ and $\sigma_{hmin}/\sigma_v = 0.65$. Since we do not have any information on poroelastic parameters, we make an assumption of model close to elastic ($\nu_u$ close to $\nu$ and $B$ close to 0).

Based on available data, we also know that the dominant strike of natural fractures in the subsurface is ~N45°E and that these fractures are subvertical. In this study we replicate part of the stimulation of Stage 2 for well SUGG 171-6SU, which was the first well stimulated in the experiment. We picked Stage 2 since it generated more microseismic events which form well-defined cloud but is still close to the beginning of the operations. In our simulation we populated model with 19 natural fractures, covering fractures at the extreme locations to the east, west, north, south, as well as deepest and shallowest ones plus additional fractures in between those. All of them have strike and dip in agreement with dataset. We assume that initially these fractures are almost closed (very small aperture), and that fracture plane has low cohesion (0.1 MPa). Since we do not have information about the coefficient of friction, we assign $\mu_s = 0.6$, which is typical value for sedimentary basin faults and fractures (e.g. Zoback, 2010). The geometry of the system is presented in Figure 5.18.

**Table 5.3 Physical properties of the layers in simplified HFTS-1 model (LSP+D – Lower Spraberry and Dean; UWC1 – Upper Upper Wolfcamp; UWC2 – Lower Upper Wolfcamp; MWC1 – Upper Middle Wolfcamp; UWC2 – Lower Middle Wolfcamp; LWC – Lower Wolfcamp)**

| Layer | Depth range [m] | ν | $\nu_u$ | B | k [$m^2$] | K [GPa] | ρ [kg/$m^3$] | $P_p$ [MPa] | $\sigma_v$ [MPa] | C [MPa] | $K_{IC}$ [MPa·$m^{1/2}$] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LSP+D | 2000-2264 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 28.9 | 2561 | 16.4 | 47.5 | 9.11 | 2.56 |
| UWC1 | 2264-2371 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 28.4 | 2513 | 36.6 | 52.5 | 6.26 | 2.50 |
| UWC2 | 2371-2404 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 28.6 | 2503 | 36.8 | 54.0 | 4.60 | 2.53 |
| MWC1 | 2404-2463 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 28.7 | 2509 | 37.2 | 55.3 | 10.50 | 2.54 |
| MWC2 | 2463-2530 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 28.4 | 2509 | 38.1 | 56.8 | 11.00 | 2.51 |
| LWC | 2530-2750 | 0.25 | 0.26 | 0.01 | $6.15\cdot10^{-23}$ | 23.5 | 2509 | 38.5 | 57.7 | 8.05 | 2.08 |

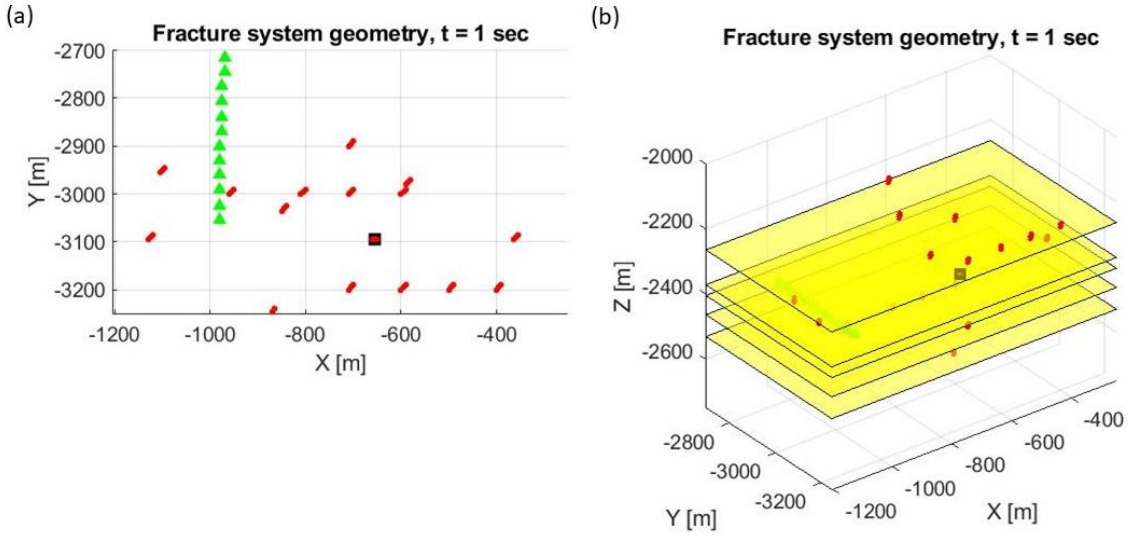


**Figure 5.18 (a) Map view and (b) 3D view on the fracture system geometry for simplified HFTS-1 model. Red planes are all the fractures, black square is the perforation point, green triangles are simulated monitoring stations and yellow planes are layer boundaries.**

### 5.6.2. Results

In this section, our aim was not to directly match the observations, which would be extremely difficult considering the level of details that would need to be captured. Rather than that, we focus on big-picture observations of hydraulic fracture growth in a multi-layered heterogenous medium and fracture extent relation with the volume occupied by microseismic clouds.

From Figure 5.19 we can see that the bottom part of the hydraulic fracture grows and quickly hits a mechanical barrier. Since the downward growth was a path of the least resistance until this point, upward and horizontal growth stops for a short period of time because they require more energy to fracture the rock. Then, after accumulation of enough stress and pressure the fracture continues its growth horizontally and slightly vertically. However, vertical upward growth is being halted by a strong mechanical barrier (the interface between Lower Spraberry and Upper Wolfcamp). Horizontal and downward growth continues until along-strike extension reaches another plateau which lasts until the end of our simulation. After ~20 min of stagnation, hydraulic fracture overcomes the mechanical barrier and continues growing upwards into Lower Spraberry. Possibly the horizontal extension plateau would be broken after longer time but it would require more energy. The growth of hydraulic fracture is not continuous as we can expect in real-case scenarios. During time we observe strong vertical asymmetry but not much of horizontal asymmetry, which might be a result of horizontal layers (which locally might be a good enough approximation).

Looking at microseismic cloud evolution (Figures 5.20 and 5.21), we can see that the fractures at the similar depth to well SUGG 171-6SU are activated relatively early on, which can indicate a strong far-field stress transfer in horizontal (especially along-strike) direction. As a result, we are able to reproduce E-W extent of the microseismic cloud. The deeper and shallower fractures are not activated until later in simulation when hydraulic fracture breaks through the mechanical barrier. In the time of simulation, the shallowest and southern-most fractures are not activated. In the first case, it can be a result of later activation or activation via bedding slip. In the second case, the event is very significantly off-plane from other events and might be possibly mislocated or there is some heterogeneity we did not capture in our simulation. As simulation progresses, especially after breaking mechanical barrier on Lower Spraberry interface, we can observe an increase in microseismic events magnitudes.

Finally, in Figure 5.22 we show how the simulated fracture extent would look against the reported microseismic data from different perspectives. Even filtering out later microseismic events we can conclude that the extent of microseismic cloud which is a typical proxy for SRV is almost twice as large as the extent of the hydraulic fracture, which is consistent with observations made on core data from the slanted well (Maity et al., 2018) indicating smaller DRV than SRV.

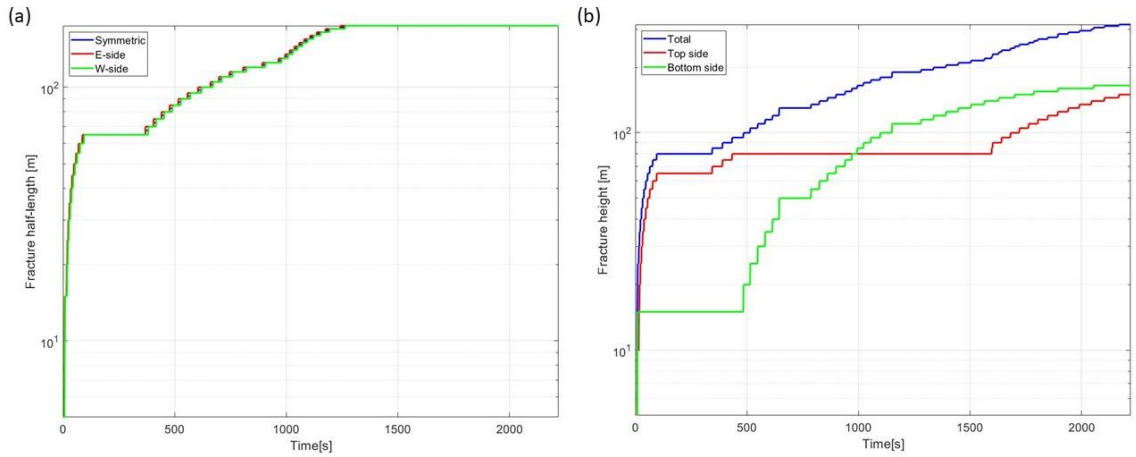


**Figure 5.19 (a) Fracture half-length changes with time (fracture geometry system for HFTS-1 model) with assumption of symmetric growth, as well as asymmetric growth of eastern and western side of the hydraulic fracture (b) fracture height changes with time (total height as well as "semi-height" for top and bottom halves).**

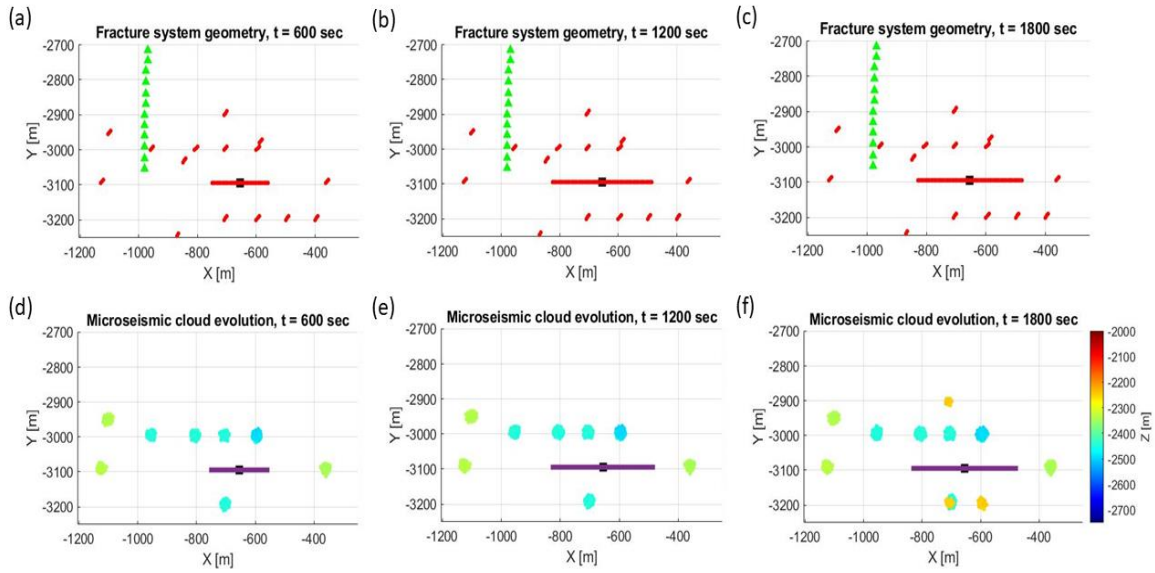


**Figure 5.20 Map view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment((d) – (f)) for fracture system geometry in HFTS-1 model. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red lines are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) – (f) purple line represents hydraulic fracture, black square is the perforation point and colorful dots represent microseismic events with depth indicated by their color.**

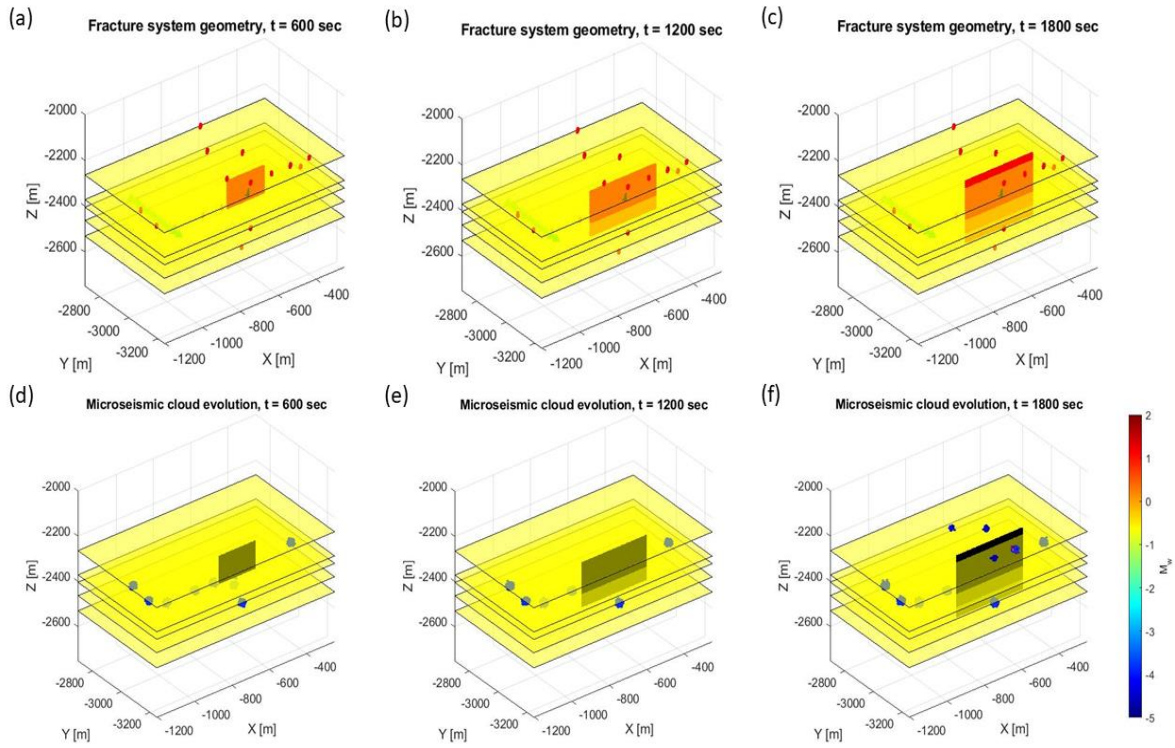


**Figure 5.21 3D view on time snapshots fracture system geometry development ((a) – (c)) and associated microseismic cloud development up to the given moment((d) – (f)) for fracture system geometry in HFTS-1 model. We do not include here tensile events caused by hydraulic fracture propagation, since they are usually not detectable. In (a) – (c) red planes are fractures, black square is the perforation point and green triangles are simulated monitoring stations. In (d) –(f) black plane represents hydraulic fracture and colorful dots represent microseismic events with magnitude indicated by their color.**

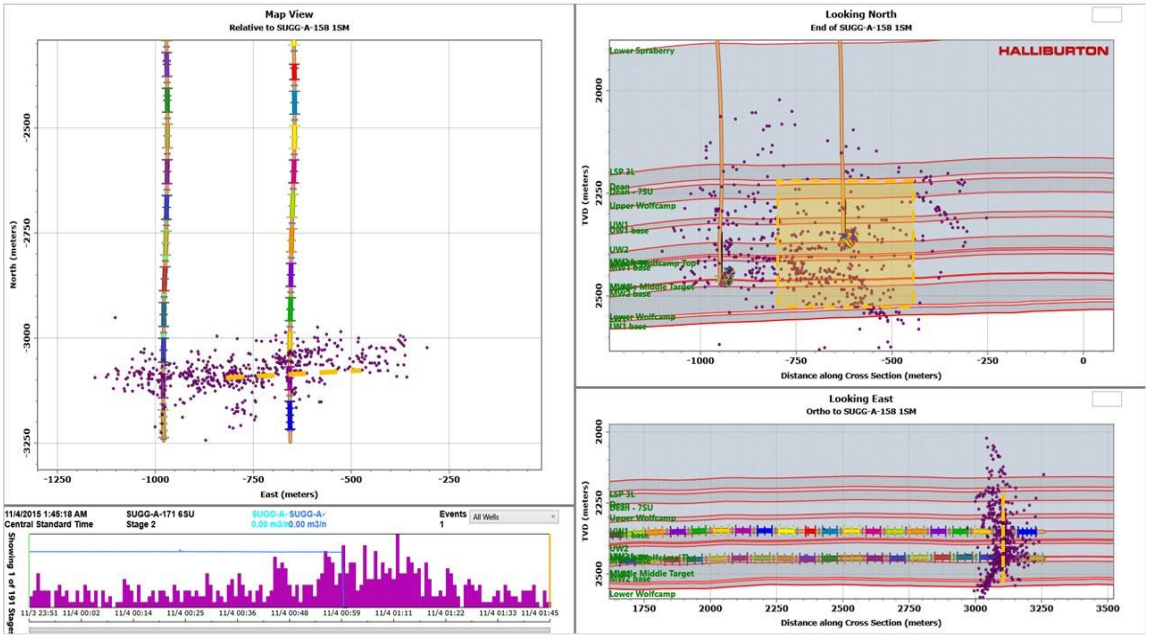


**Figure 5.22 Screenshot with displayed microseismic data for Stage 2 of well SUGG 171-6SU shown from different perspectives – map view in top left, "barrel view" in top right, N-S profile in bottom right. Panel in bottom left is time series with injection history (line plot) and bar plot with number of microseismic events. Drawn orange dashed lines and orange polygon are representation of the hydraulic fracture extent according to the simulation on HFTS-1 simplified model in this study after the first 37 minutes of stimulation.**

## 5.7. Discussion

In the presented models above, we mostly focus on hydraulic fracture propagation in a 3D layered medium and its interaction with natural fractures either via far-field stress transfer through the medium (all models) or direct mechanical intersection (Setup 3 in Synthetic models). However, multiple studies (e.g. Stanek and Eisner, 2017; Tan et al., 2021) show that bedding plane slip could be an important mechanism of microseismicity generation. To cause shear slip along sub horizontal bedding it is necessary to overcome very high confining stress equal approximately to vertical stress. As a result, the extent

of such microseismic cloud would be limited to the vicinity of hydraulic fracture unless given bedding is directly intersected by hydraulic fracture. Thus, we propose that in reality the generated microseismicity is a combination of hydraulic fracture interaction with favorably-oriented natural fractures and with bedding planes.

Here, we explore bedding plane slip activation process. In order to do that, we use the same 3-layer model as in our Synthetic Models section. However, this time instead of treating the layer interface as a jump in mechanical properties, we treat it as a "natural horizontal fracture", allowing for slip and aperture change due to stress and pressure perturbation. Figure 5.23 shows the pore pressure and stress distribution on one of the bedding planes around the time of its activation (Figure 5.23 (a) – (c)), the along-strike profile showing the extent of the hydraulic fracture (Figure 5.23 (d)) and a map view with microseismic events and marked hydraulic fracture (Figure 5.23 (e)). The bedding plane is activated when hydraulic fracture hit the bedding plane and then propagated along strike. Activation is asymmetric (on the eastern side) and is a cumulative effect of mainly pore pressure and tensile stress perturbation. We show that bedding plane activation is a viable explanation of part of the registered microseismic activity and should be further investigated in the future.

Another important aspect that we want to bring attention to is how to improve the monitoring of hydraulic fracturing process. It seems that microseismic monitoring provides only partial answer to the question about the actual extent of stimulated

volume, since many microseismic events can be purely associated with far-field stress transfer and these fractures might not contribute to production of hydrocarbons. One of the recent solutions is supplementing conventional microseismic monitoring with Distributed Acoustic Sensing (DAS), through which we can to some extent monitor microseismicity, but also look at the slow strain changes with time. For our fracture geometry system Setup 3 with a horizontal set of stations we extract the strain tensor changes with time (see Figure 5.24). Because of the cable N-S orientation we mainly pay attention to components $\varepsilon_{22}$, $\varepsilon_{12}$ and $\varepsilon_{23}$. We can see that as the hydraulic fracture approaches monitoring cable, component $\varepsilon_{22}$ changes from initially compression, through no strain up to increasingly tensile strain. Unfortunately, in that tested setup the hydraulic fracture did not reach the cable and we cannot see what signature fracture hit would produce.

The methodology we present in this study can also supplement field microseismic monitoring in estimation of the volume contributing to hydrocarbon production (i.e., DRV). Having the knowledge about the subsurface geology and formations mechanical properties, along with information about natural fractures orientations and/or density, one can construct the model populated with preexisting fractures for the simulations of past, ongoing or future treatments. From the results, it is possible to distinguish between microseismic events that are associated with hydraulically connected fractures versus events triggered by far-field stress. If we already have inverted microseismic events locations from monitoring, they can serve as a benchmark to simulation. If the actual

cloud is reproduced within an acceptable margin of error, DRV might estimated as a volume occupied by hydraulic fracture and all fractures hydraulically connected to it. If microseismic clouds do not match, then it creates an opportunity to run more simulations and improve the predictive model, e.g. using the integrated methodology we proposed in our previous work (Chapter 4).

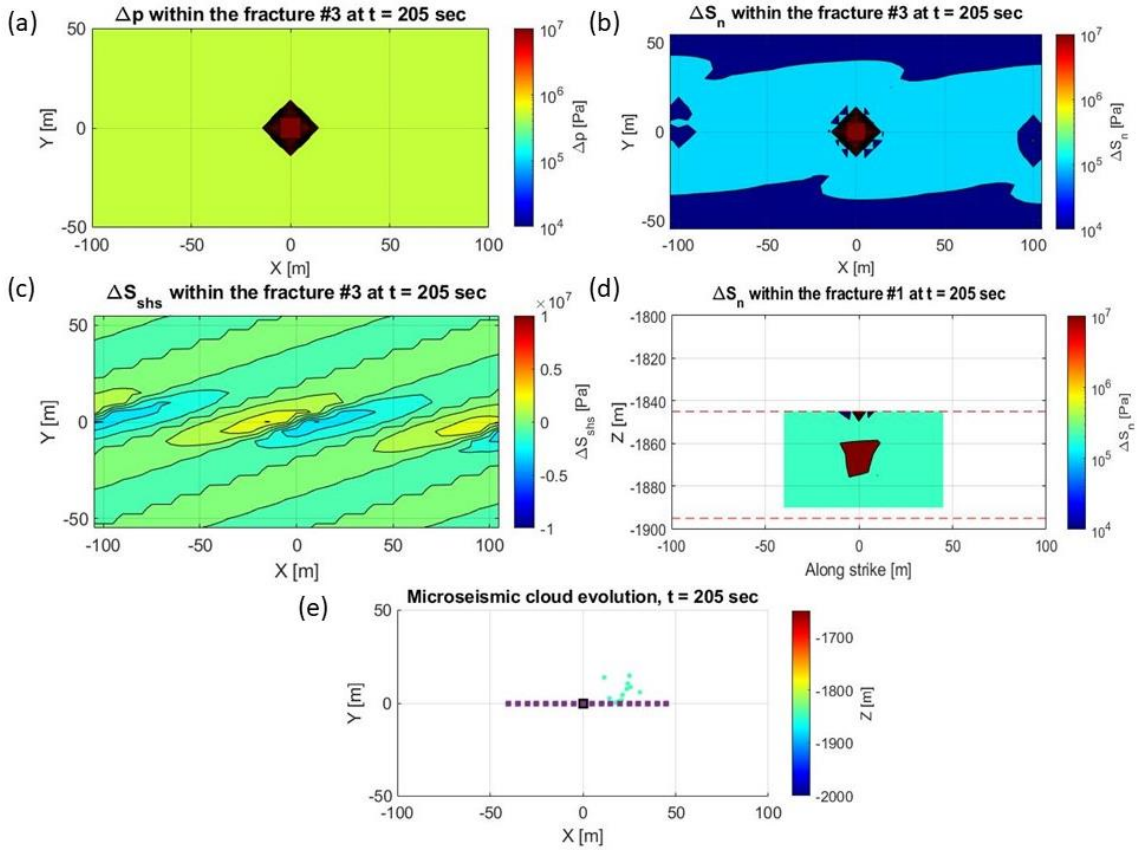


**Figure 5.23 (a) – (c) distributions of pore pressure, normal stress and shear stress on the top bedding plane shortly after the nascent of microseismic activity; (d) the geometry of the hydraulic fracture with normal stress distribution at the same moment with red dashed lines representing bedding locations; (e) microseismic cloud distribution at the same moment (color dots coded by the depth of an event), along with marked perforation point (black square) and extent of hydraulic fracture (purple dashed line).**

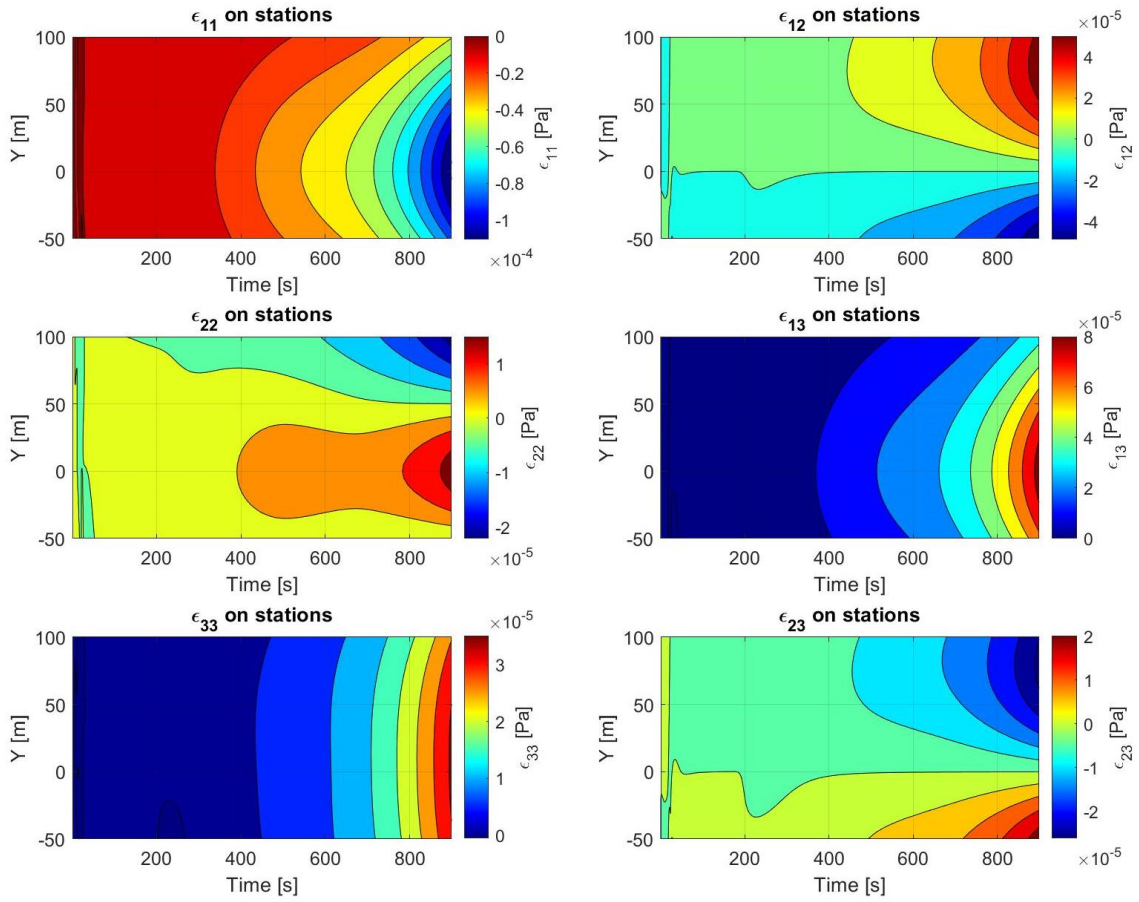


**Figure 5.24 All components of strain tensor registered on the simulated DAS cable for simulation on fracture system geometry Setup 3 with a N-S oriented array (See Figure 5.6 for its location).**

## 5.8. Conclusions

In this study we present an integrated methodology for numerical simulations of fluid flow, hydraulic fracture propagation and its interaction with natural fractures and bedding planes, and generation of microseismicity in a 3D deformable, heterogenous medium. We verify the code against analytical solutions (classical PKN and KGD models) as well as numerical simulation results of Razor simulator from Peshcherenko and Chuprakov (2021). We get good agreement with analytical solutions and replicate

the fracture length for the more complex model. We explain fracture height differences as a result of using poroelastic instead of elastic model (in spite of creating relatively close proxy) and back-stress effects.

We also run simulations for 3 different setups of fracture networks to investigate hydraulic fracture propagation in presence of natural fractures, its interaction with them via direct intersection or by far-field stress perturbation. We also look into the hydraulic fracture extent in comparison to the volume occupied by microseismic cloud to distinguish between “wet” and “dry” events. We observe that in certain situations microseismic cloud might be much larger than actual drained volume, while in other conditions hydraulic fracture can penetrate a volume without generation of almost any microseismicity.

Presence of nearby natural fractures can affect the propagation of the hydraulic fracture, and either hinder or propel it via stress feedback, which depends on natural fractures’ predominant orientation. This might result in fracture asymmetry, although the preferred orientation of natural fractures on one side of perforation is rather unlikely. We also show that in situation when hydraulic fracture intersects natural fracture it can cross it without change of direction, but later in the process the propagation path might be diverted towards that natural fracture.

In the simulation for a simplified model of HFTS-1 experiment, we show that the extent of hydraulic fracture is probably smaller than the extent of microseismic cloud, thus SRV is not a good proxy of DRV. This result is in agreement with previous analyses of the core from a slanted well. We also obtain a horizontally symmetric fracture. Therefore, the asymmetry of microseismic cloud is most probably the effect of the monitoring array locations and bias in number of events detected on the limb closer vs further from the seismic stations.

By conducting a small experiment, we also show that bedding plane slip is a viable alternative mechanism of microseismicity generation. However, further study involving all interactions at once need to be done to evaluate what percentage of reported events corresponds to natural fracture vs bedding plane slip.

Last but not least, the methodology developed in this work can be used in estimation of DRV based on registered and simulated microseismic clouds (and associated SRVs). This approach can be useful both in hydraulic fracturing design and retrospective analysis of stimulations.

## 5.9. References

Adams, J., & Rowe, C. (2013, May) ‘Differentiating applications of hydraulic fracturing’, *ISRM International Conference for Effective and Sustainable Hydraulic Fracturing*, May 2013.

Arthur, J. D., Bohm, B., & Layne, M. (2009) ‘Hydraulic fracturing considerations for natural gas wells of the Marcellus Shale’, *Gulf Coast Association of Geological Societies Transactions*, 59, pp. 49-59.

Dalla Barba, F., & Picano, F. (2020) ’A novel approach for direct numerical simulation of hydraulic fracture problems’, *Flow, Turbulence and Combustion*, *105*(2), pp. 335-357.

Bear, J. (1972) *Dynamics of Fluids in Porous Media*. New York: American Elsevier Publishing Company.

Bunger, A. P., Jeffrey, R. G., & Detournay, E. (2008) ‘Evolution and morphology of saucer-shaped sills in analogue experiments’, *Geological Society, London, Special Publications*, *302*(1), pp. 109-120.

Bunger, A., & Lecampion, B. (2017) ‘Four critical issues for successful hydraulic fracturing applications’, in X.-T. Feng (eds), *Rock mechanics and engineering*., 5(16), pp. 1-34.

Carrier, B., & Granet, S. (2012) ‘Numerical modeling of hydraulic fracture problem in permeable medium using cohesive zone model’, *Engineering fracture mechanics*, *79*, pp. 312-328.

Camanho, P. P., & Dávila, C. G. (2002) *Mixed-mode decohesion finite elements for the simulation of delamination in composite materials*. Hampton, VA: NASA.

Chen, Z., et al. (2009) ‘Cohesive zone finite element-based modeling of hydraulic fractures’, *Acta Mechanica Solida Sinica*, *22*(5), pp. 443-452.

Ciezobka, J., Courtier, J., & Wicker, J. (2018) ‘Hydraulic fracturing test site (HFTS)-project overview and summary of results’, *SPE/AAPG/SEG Unconventional Resources Technology Conference*, July 2018.

Douillet-Grellier, T., et al. (2016) ‘Mixed-mode fracture modeling with smoothed particle hydrodynamics’, *Computers and Geotechnics*, *79*, pp. 73-85.

Fan, J., et al. (2012) ‘Directional hydraulic fracturing to control hard-roof rockburst in coal mines’, *International Journal of Mining Science and Technology*, *22*(2), pp. 177-181.

Fehler, M. C. (1989) ‘Stress control of seismicity patterns observed during hydraulic fracturing experiments at the Fenton Hill hot dry rock geothermal energy site, New Mexico’, *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, July 1989, 26(3-4), pp. 211-219.

Feng, Y., & Gray, K. E. (2017) ‘Parameters controlling pressure and fracture behaviors in field injectivity tests: a numerical investigation using coupled flow and geomechanics model’, *Computers and Geotechnics*, 87, pp. 49-61.

Feng, Y., & Gray, K. E. (2018) ‘Modeling of curving hydraulic fracture propagation from a wellbore in a poroelastic medium’, *Journal of Natural Gas Science and Engineering*, 53, pp. 83-93.

Friedrich, M., & Monson, G. (2013) ‘Two practical methods to determine pore pressure regimes in the Spraberry and Wolfcamp formations in the Midland basin’, *Unconventional Resources Technology Conference*, August 2013, pp. 2475-2486.

Geertsma, J., & De Klerk, F. (1969) ‘A rapid method of predicting width and extent of hydraulically induced fractures’, *Journal of petroleum technology*, 21(12), pp. 1571-1581.

Gerke, H. H., & Van Genuchten, M. T. (1993) ‘A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media’, *Water resources research*, 29(2), pp. 305-319.

Ghassemi, A. (2017) ‘Application of rock failure simulation in design optimization of the hydraulic fracturing’, *Porous rock fracture mechanics*. Woodhead Publishing, pp. 3-23.

Golovin, S. V., et al. (2015) ‘Hydraulic fracture numerical model free of explicit tip tracking’, *International Journal of Rock Mechanics and Mining Sciences*, 76, pp. 174-181.

Grechka, V. I., & Heigl, W. M. (2017) *Microseismic monitoring*. Tulsa, OK: Society of Exploration Geophysicists.

Gupta, P., & Duarte, C. A. (2014) ‘Simulation of non-planar three-dimensional hydraulic fracture propagation’, *International Journal for Numerical and Analytical Methods in Geomechanics*, 38(13), pp. 1397-1430.

He, Z., & Duan, B. (2021) ‘Study on the microseismic clouds induced by hydraulic fracturing’, *Geomechanics and Geoengineering*, 16(5), pp. 400-416.

Hobro, J., Williams, M., & Calvez, J. L. (2016) ‘The finite-difference method in microseismic modeling: Fundamentals, implementation, and applications.’, *The Leading Edge*, 35(4), pp. 362-366.

Hossain, M. M., & Rahman, M. K. (2008) ‘Numerical simulation of complex fracture growth during tight reservoir stimulation by hydraulic fracturing.’, *Journal of Petroleum Science and Engineering*, 60(2), pp. 86-104.

Howard, G. C., & Fast, C. R. (1957) ‘Optimum fluid characteristics for fracture extension.’ *Drilling and production practice*, January 1957.

Hughes, T. J. R. (2000) *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover, Mineola, N. Y.

Khristianovic, S. A., & Zheltov, Y. P. (1955) ‘Formation of vertical fractures by means of highly viscous liquid.’, *4th World Petroleum Congress Proceedings,* pp. 579-586.

Kong, B., Fathi, E., & Ameri, S. (2015) ‘Coupled 3-D numerical simulation of proppant distribution and hydraulic fracturing performance optimization in Marcellus shale reservoirs.’, *International Journal of Coal Geology*, 147, pp. 35-45.

Lecampion, B., Bunger, A., & Zhang, X. (2018) ‘Numerical methods for hydraulic fracture propagation: a review of recent trends.’, *Journal of natural gas science and engineering*, 49, pp. 66-83.

Lecampion, B., & Desroches, J. (2015) ‘Simultaneous initiation and growth of multiple radial hydraulic fractures from a horizontal wellbore.’, *Journal of the Mechanics and Physics of Solids*, 82, pp. 235-258.

Legarth, B., Huenges, E., & Zimmermann, G. (2005) ‘Hydraulic fracturing in a sedimentary geothermal reservoir: Results and implications.’, *International Journal of Rock Mechanics and Mining Sciences*, 42(7-8), pp. 1028-1041.

Maity, D. (2018) ‘Microseismicity analysis for HFTS pad and correlation with completion parameters.’, *Unconventional Resources Technology Conference,* July 2018, pp. 3810-3821.

Maity, D., Ciezobka, J., & Eisenlord, S. (2018) ‘Assessment of in-situ proppant placement in SRV using through-fracture core sampling at HFTS.’, *Unconventional Resources Technology Conference*, July 2018, pp. 3810-3821.

Mayerhofer, M. J., et al. (2010) ‘What is stimulated reservoir volume?’, *SPE Production & Operations*, 25(01), pp. 89-98.

Mehrabian, A., & Abousleiman, Y. N. (2014) ‘Generalized Biot's theory and Mandel's problem of multiple-porosity and multiple-permeability poroelasticity.’, *Journal of Geophysical Research: Solid Earth*, 119(4), pp. 2745-2763.

Mendelsohn, D. A. (1984) ‘A Review of Hydraulic Fracture Modeling—Part I: General Concepts, 2D Models, Motivation for 3D Modeling.’, *ASME. J. Energy Resour. Technol.*, 106(3), pp. 369–376.

Miehe, C., Schaenzel, L. M., & Ulmer, H. (2015) ‘Phase field modeling of fracture in multi-physics problems. Part I. Balance of crack surface and failure criteria for brittle crack propagation in thermo-elastic solids.’, *Computer Methods in Applied Mechanics and Engineering*, 294, pp. 449-485.

Nandlal, K., & Weijermars, R. (2019) ‘Impact on drained rock volume (DRV) of storativity and enhanced permeability in naturally fractured reservoirs: upscaled field case from hydraulic fracturing test site (HFTS), Wolfcamp Formation, Midland Basin, West Texas.’, *Energies*, 12(20), 3852.

Needleman, A. (2014) ‘Some issues in cohesive surface modeling.’, *Procedia IUTAM*, 10, pp. 221-246.

Perkins, T. K., & Kern, L. R. (1961) ‘Widths of hydraulic fractures.’, *Journal of petroleum technology*, 13(09), pp. 937-949.

Peshcherenko, A., & Chuprakov, D. (2021) ‘An ultrafast simulator for 3D propagation of a hydraulic fracture with rectangular shape.’, *Engineering Fracture Mechanics*, 243, 107512.

Pudugramam, V. S., et al. (2021) ‘Analysis and Integration of the Hydraulic Fracturing Test Site-2 (HFTS-2) Comprehensive Dataset.’, *Unconventional Resources Technology Conference*, 5229.

Rice, J. R., & Drucker, D. C. (1967) ‘Energy changes in stressed bodies due to void and crack growth.’, *International Journal of Fracture Mechanics*, 3(1), pp. 19-27.

Salimzadeh, S., Paluszny, A., & Zimmerman, R. W. (2017) ‘Three-dimensional poroelastic effects during hydraulic fracturing in permeable rocks.’, *International Journal of Solids and Structures*, 108, pp. 153-163.

Secchi, S., Simoni, L., & A. Schrefler, B. (2007) ‘Mesh adaptation and transfer schemes for discrete fracture propagation in porous materials.’ *International journal for numerical and analytical methods in geomechanics*, 31(2), pp. 331-345.

Segall, P., & Lu, S. (2015) ‘Injection-induced seismicity: Poroelastic and earthquake nucleation effects.’ *Journal of Geophysical Research: Solid Earth*, 120(7), pp. 5082-5103.

Settari, A., & Cleary, M. P. (1986) ‘Development and testing of a pseudo-three-dimensional model of hydraulic fracture geometry.’, *SPE Production Engineering*, 1(06), pp. 449-466.

Snee, J. E. L., & Zoback, M. D. (2018) ‘State of stress in the Permian Basin, Texas and New Mexico: Implications for induced seismicity.’, *The Leading Edge*, 37(2), pp. 127-134.

Staněk, F., & Eisner, L. (2017) ‘Seismicity induced by hydraulic fracturing in shales: A bedding plane slip model.’, *Journal of Geophysical Research: Solid Earth*, 122(10), pp. 7912-7926.

Stegent, N., & Candler, C. (2018) ‘Downhole microseismic mapping of more than 400 fracturing stages on a multiwell pad at the Hydraulic Fracturing Test Site (HFTS): Discussion of operational challenges and analytic results.’, *Unconventional Resources Technology Conference*, July 2018, pp. 3754-3781.

Szafranski D., and Duan, B. (2020) ‘Exploring physical links between fluid injection and nearby earthquakes: The 2012 M 4.8 Timpson, Texas, case study’, *Bulletin of the Seismological Society of America*, 110 (5), pp. 2350–2365.

Tan, Y., Chai, C., & Engelder, T. (2014) ‘Use of S-wave attenuation from perforation shots to map the growth of the stimulated reservoir volume in the Marcellus gas shale.’, *The Leading Edge*, 33(10), pp. 1090-1096.

Tan, Y., et al. (2021) ‘Mechanism of Microseismic Generation During Hydraulic Fracturing–With Evidence From HFTS 2 Observations.’ *Unconventional Resources Technology Conference*, July 2018, 5296.

Wang, H. (2000) *Theory of Linear Poroelasticity With Applications to Geomechanics and Hydrogeology*. Princeton, N. J: Princeton Univ. Press.

Wang, H. (2015) ‘Numerical modeling of non-planar hydraulic fracture propagation in brittle and ductile rocks using XFEM with cohesive zone method.’, *Journal of Petroleum Science and Engineering*, 135, pp. 127-140.

Wang, S., et al. (2019) ‘Learnings from the Hydraulic Fracturing Test Site (HFTS)# 1, Midland Basin, West Texas—A Geomechanics Perspective.’, *Unconventional Resources Technology Conference, July 2019*, pp. 2906-2922.

Warpinski, N. (2009) ‘Microseismic monitoring: Inside and out.’, *Journal of Petroleum Technology*, 61(11), pp. 80-85.

Zhang, W., et al. (2017) ‘A pore-scale numerical model for non-Darcy fluid flow through rough-walled fractures.’, *Computers and Geotechnics*, 87, pp. 139-148.

Zhou, J., Huang, H., & Deo, M. (2016) ‘Numerical study of critical role of rock heterogeneity in hydraulic fracture propagation.’, *50th US Rock Mechanics/Geomechanics Symposium*, June 2016.

Zia, H., & Lecampion, B. (2020) ‘PyFrac: A planar 3D hydraulic fracture simulator.’, *Computer Physics Communications*, 255, 107368.

Zoback, M. D. (2010). *Reservoir geomechanics*. Cambridge University Press.

# 6. CONCLUSIONS

## 6.1. Impact of model parameters on timing and size of the induced earthquake.

Based on our simulations with an analytical solution of poroelastic problems combined with dynamic rupture modeling, we find that permeability, elastic modulus of the medium play an important role in the size of the stress and pore pressure perturbations, which in turn affects the timing of event triggering, as well as the size of induced earthquake. We notice that lower values of elastic modulus promotes earlier fault activation but also smaller magnitudes of seismic events. Another observation is that intermediate to low permeability of formation promotes the occurrence of induced seismicity, whereas high permeability does not allow for any rupture. In addition, from our numerical simulations, in which we introduce layering with a strong permeability contrast between formations, the magnitudes and distribution of stress and pressure perturbations are different from those for homogenous model. For example, a peak accumulation of perturbations for the Timpson (TX) earthquake is located nearby the formations boundary.

Secondly, in dynamic rupture propagation studies of the Timpson (TX) earthquake, we show that the pattern of background stress is very important when it comes to the final size of the earthquake rupture. Simple depth-dependent stress distribution promotes either very small (or sometimes no failure) or very large induced earthquakes, whereas uniform normal effective stress within the fault zone (indicating overpressure) can

produce earthquakes with intermediate magnitudes and smoother rupture propagation patterns. Another aspect related to background stress is its magnitude. In this study, we show that the closer the initial proximity of the fault stress to the yield stress, the larger the earthquake can be triggered because activated patch can be larger. With dynamic acceleration of the rupture, breaking through large portion of the fault is not hard.

Besides the background stress, fault frictional parameters also play a major role in the time of triggering, as well as the size of simulated earthquake. Static friction coefficient mainly influences the timing of induced earthquakes. From parameter space exploration with dynamic rupture models, we observe the increase of the earthquake size with the increasing difference between static and dynamic friction coefficients. Inversely, we also notice the increase in earthquake magnitudes with the decrease of critical slip distance. Correlations between these parameters and the magnitudes of induced earthquakes are strongly non-linear. Moreover, without inversion of these parameters from observed seismic data, there are usually no direct observations that might constrain them for the purpose of modeling.

Finally, based on our results, the distance between the well and the fault is very important in consideration of the onset of induced seismicity. Generally, the further away the well is from the fault, the later induced seismicity occurs. If the well is far enough from the fault, even with high injection rates, the perturbations might never cause earthquake triggering. Furthermore, both with analytical and numerical solutions,

we observe that the smaller the distance of the well from the fault the more concentrated are perturbations and the smaller is the earthquake initiation patch and final rupture area. In the study of Timpson (TX) earthquake, we also introduce randomized heterogeneity resembling asperities observed in fault zone field studies. We find that the size of an earthquake can be decreased when these asperities are present in the model. This shows that even small propagation barriers can halt the rupture if it did not gain enough acceleration before reaching them.

### 6.2. Importance of simulating hydromechanical coupling and dynamic effects in seismic hazard analysis.

Based on results analysis from multiple simulations of fluid flow, we observe that in many cases far-field poroelastic stress transfer through the solid medium can be a main source of perturbations on the fault, hence causing fault activation and triggering of an earthquake. Therefore, neglecting poroelastic effects would lead us to inaccurate conclusions.

Moreover, dynamic rupture simulations allow for more reliable magnitude estimation. From our observations, the difference in failure size predicted with quasi-static versus dynamic modeling can be significantly different, with quasi-static failure area being equal or (more commonly) smaller than dynamic failure. Therefore, one of the possible applications of dynamic rupture modeling is the deterministic estimation of seismic hazard in the area.

As we show in the study of Timpson (TX) earthquake, by running dynamic rupture simulations for multiple different parameters setups we can estimate the realistic maximum possible magnitude that can be triggered by the injection, which is one of the most important challenges in the seismic hazard assessment. In fact, more popular stochastic method of seismic hazard analysis does not allow to find what is possible worst-case scenario considering a certain fault system in the area, because they rely on historical records of seismicity, which may not include large paleo-earthquakes or take into account human impact on the seismicity level.

Dynamic rupture simulations also allow for tracking the approximate Peak Ground Accelerations (PGA) or Velocities (PGV) on the earth surface which can be useful in estimation of seismic hazard, though we did not focus on this aspect in this dissertation.

Moreover, modeling of hydraulic fracturing using poroelastic model allows to reproduce a microseismic cloud with distinction between “dry” and “wet” events. Using poroelastic model we can also observe a “back-stress” effect, which is a feedback response between fractures or medium and fractures. Additionally, in hydraulic fracturing we also use a dual permeability model with stress-dependent aperture and fracture permeability. Thus, poroelastic stress changes can significantly influence fluid flow, hydraulic fracture propagation and its interactions with natural fractures.

In this research we use 3D models to capture nuances and effects that are not possible to replicate with 2D modeling both for induced seismicity and hydraulic fracture propagation. We also introduce different forms of heterogeneities to models, such as layering or fault zone asperities, which allows us to study realistic models and apply methodology to real case studies. Homogenous model can be useful to get initial understanding of some processes but the results of simulations are vastly different from models with layering or other types of heterogeneity.

Another important decision we need to make, when deciding on appropriate methodology are primary goals of a given analysis. In the first study, our focus was on general understanding of dynamic rupture propagation caused by water injection. Thus, we implemented appropriate analytical solutions. Among advantages of analytical approaches is very short computation time. Another advantage is model simplicity in understanding and implementation. As a result, these models are very useful for basic understanding of selected processes happening during water injection. On the other hand, analytical solutions are limited to simple models (usually homogenous) with a lot of assumptions (e.g. constant injection rate). Therefore, the biggest disadvantage of this approach is the lack of their applicability to thorough studies of real case examples of induced earthquakes.

In further studies we go beyond very simplistic models and want to show that we can actually apply integrated modeling to real case studies. Therefore, we switch from

analytical solutions to numerical solution of hydromechanical coupling problem. Numerical approaches, dependent on selected ways of implementation, can potentially handle heterogenous and very complex models. Thus, they can be applied to real case studies. Additionally, they are useful in investigations of complex interactions or addressing specific nuance problems. However, handling complexity comes with relatively slow computations. This issue can be partially resolved by using parallel computations or GPU on modern work stations, computation clusters or clouds. Numerical approaches have also some numerical limitations, e.g. minimum mesh size and maximum time step for which given medium or fracture permeability can be handled with stable solution.

### 6.3. Lessons learned from Timpson (TX) earthquake analysis.

Using an updated layered, poroelastic model for fluid flow simulations in Timpson, TX earthquake area published by Shirzaei et al. (2019), we are able to get magnitudes of pressure and stress perturbations in agreement with results from Fan et al. (2016) (around several MPa) that allows fault failure in the segment nearby the location of the mainshock hypocenter. Normal stress perturbations are twice as large as pore pressure perturbations and shear stress perturbations are very comparable to pore pressure perturbations in terms of their magnitudes. This result emphasizes the importance of poroelastic effects in fault activation analysis.

By doing parameter space exploration we are able to find dynamic rupture models that allow to reproduce mainshock moment magnitude ($M_w$ 4.8) and focal mechanism (predominantly strike-slip). Moreover, we are successful with reproduction of the aftershock sequence matching its main features for selected models.
Furthermore, with an initial exploration of parameter space, we find that models with lower critical slip distance, intermediate dynamic friction coefficient and uniform normal effective background stress pattern promote earthquake of reported magnitude $M_w$ 4.8. In some of dynamic rupture models we introduce randomized fault zone heterogeneity (representing asperities), which can produce different propagation patterns and, in some cases, halt ruptures earlier than in corresponding homogenous models.

We identify that establishing the values that we want to test might be a difficult task, especially if the relationships between parameter values and moment magnitudes are highly non-linear. Therefore, to better constrain the poorly-known parameters, it is necessary to perform a systematic parameter space exploration. In Chapter 4 we extend our investigation of the Timpson (TX) earthquake and develop a framework in which we combine numerical modeling, machine learning and stochastic inversion to better constrain model parameters.

Since we perform multi-parameter inversion with strongly non-linear problem, stochastic inversion is superior to deterministic algorithms, because of its better efficiency and resistance to the local minima traps. Moreover, we obtain not only the

best-matching values but also their uncertainties based on posterior distributions. In cases where multiple parameters do not have well-constrained values, we can explore whether we can in fact constrain them and what is the degree of model equivalency in our problem. From posterior distributions, we can also learn whether given parameter is actually important in simulation and small changes of its value result in significant changes in the outcome, or whether it is the opposite. Posterior distributions that are close to truncated uniform indicate lack of importance, whereas distributions with pronounce maximum or maxima emphasize the importance of the parameter.

We invert for stress and frictional parameters of the fault. We obtain relatively high initial normal stress and pore pressure on the fault, with intermediate level of initial shear stress (relative to yield stress level). At the same time, we get relatively low values of critical slip distance, and low values of static and dynamic friction coefficients. Thanks to using inversion we manage to improve a match of not only mainshock reported and simulated magnitudes, but also better reproduce the aftershock sequence. We also get better understanding of the scope of models that satisfy our matching criteria, which gives us an additional insight into model equivalency issue.

Overall, we were successful in applying our methodology to the Timpson (TX) earthquake case study, showing the value of the integrated modeling algorithm in investigations of seismicity induced by waste water disposal. We confirm and strengthen the causation link between these operations and the triggered earthquake sequence.

In this study, we also implement an innovative approach to improve computational efficiency of the inversion by replacing physics-based simulations with machine learning regression algorithm. However, we still keep physical link between model input parameters and parameter(s) summarizing simulation results. With proper training and optimization of selected machine learning regression algorithms, we are able to get very high accuracy of moment magnitude predictions relative to magnitudes predicted via dynamic rupture simulations ($R^2$ above 0.9). The best performing algorithms are Random Forest and Bagging algorithms (ensemble methods).

Dependent on the size of the training dataset (we use 485 or 1026 samples) and number of Monte Carlo samples in inversion, we manage to save between 85% to 97% of computation time on a single inversion run by replacing physics-based simulations with regression algorithms predictions. Even including cross-validation time and multiple runs of inversions, the efficiency of the hybrid stochastic-deterministic approach is much more efficient than using physics-based simulations alone in an inversion scheme.

### 6.4. Adaptation of numerical modeling of induced seismicity to simulation of hydraulic fracturing operations and associated microseismicity.

In both induced seismicity and hydraulic fracturing simulation problems, there are some common points. First, we need to simulate fluid flow in a deformable porous medium. Second, we resolve pore pressure and stress changes on a fault or fracture plane. Additionally, both phenomena are occurring at relatively shallow depth within a

sedimentary basin. Therefore, we observe only brittle deformations and temperature does not have significant impact on mechanical behavior of fracture or fault.

On the other hand, there are also several significant differences between induced seismicity and hydraulic fracturing problems. First of all, we operate on different scales. Simulating induced seismicity is associated with regional or local faults that are much larger than a hydraulic fracture or natural fractures. The difference is also in time scales, where hydraulic fracturing stimulations last minutes, hours or days if we want to reproduce multiple stages. At the same time, simulation of waste water disposal and induced earthquakes replicate months or years of operations. Secondly, in hydraulic fracturing with moving boundary problem (fracture propagation), some level of remeshing and special treatment of boundary conditions is necessary. Also, since water is being pumped directly into fracture (hydraulic fracturing), the distinction between flow through the fracture and through the intact rock pore space is more important than for wastewater disposal, where we try to avoid pumping into preexisting faults. Due to strongly anisotropic permeability of the fractures (higher along the strike than perpendicular to it), single porosity medium assumption is no longer a valid approximation of the reality.

In order to meet new requirements of the problem, we expand our code with new additions and modules. First, we replace equations for a single-porosity medium with newly-derived equations for a dual-porosity medium (separate treatment of flow through

fractures and pore space in the medium). Secondly, we develop a fracture propagation module, which changes the extent of predefined fracture if an appropriate failure criterion is met. We also develop a cohesive zone simulation module, which is a simple treatment of local plastic deformations, and introduce fracture aperture and permeability dependence on stress changes and deterioration of material with increase of fracture aperture. We also expand previously used Mohr-Coulomb failure criterion into mixed-mode quadratic failure criterion including also possibility of tensile failure. To accommodate possibility of very high permeabilities along the fractures, we add a possibility to have different time steps for fluid flow in the medium and in the fracture. Furthermore, for both tensile and shear deformations, we introduce a distinction between strengths for existing fractures (naturally weaker because of material damage) versus the strength of intact rock which need to be overcome to propagate a fracture. Last but not least, we add a microseismicity generation extension, which identifies failed fracture areas, calculates moment magnitude, hypocenter location and stress changes due to each failure.

**6.5. Interactions of hydraulic fracture with natural fractures and bedding planes.**

In our work, we identify several different ways of interaction between a hydraulic fracture and natural fractures. First observation is that the hydraulic fracture can change stress and/or pore pressure in natural fractures either via direct mechanical interaction when intersecting it or by far-field poroelastic stress transfer through the medium.

Secondly, when intersecting a natural fracture, the hydraulic fracture can cross it without change of propagation direction if orientation of $\sigma_{hmin}$ is favorable and it requires less energy to continue tensile opening along the current strike direction. However, after being intersected natural fracture can become new pathway of fracture system extension, e.g. in situation when hydraulic fracture tip is far from the injection point. This new pathway does not necessarily activate until later in stimulation.

Thirdly, if natural fractures have favorable orientation, they can be activated way before hydraulic fracture reaches the area. Moreover, if a natural fracture is close to hydraulic fracture, it can create feedback effect on hydraulic fracture, promoting its faster or slower propagation. This can happen even if the two are not intersected.

Besides interactions with natural fractures, a hydraulic fracture can also interact with weak bedding planes by opening or shearing it. In our simulations, we notice mixed-mode interaction between hydraulic fracture and bedding plane. The failure on the bedding plane occurs when there is a direct mechanical interaction between that plane and hydraulic fracture, which differs from interactions with natural fractures that can be activated with far-field stress. In addition, fluid redirection into the space between beddings can halt the vertical propagation of hydraulic fracture. However, further investigation is necessary on this topic.

### 6.6. Relationship between SRV and DRV, and recommendations for improvement of hydraulic fracturing monitoring.

Most commonly the microseismic cloud extent is used to estimate the SRV, which is supposed to be a proxy for producing volume. There is an underlying assumption that microseismic events occur on fractures connected hydraulically to hydraulic fracture. However, as we show in Chapter 5, some microseismic events can be associated with far-field stress transfers and failure on natural fractures that do not contribute to production because of lack of hydraulic connection.

DRV is the volume occupied by hydraulic fracture and all natural fractures that are hydraulically connected with hydraulic fracture, and SRV is supposed to approximate it. Unfortunately, SRV rarely reflects the actual DRV. In some situations, when natural fractures are not far from failure and are favorably oriented, the SRV can be much larger than actual DRV. Also, in specific cases when natural fractures are far from failure (e.g. unfavorably oriented), or generate only very small failures, we do not see microseismic events and SRV is underestimating actual DRV. Examples of both of these possible cases are presented in Chapter 5.

One of the solutions to those shortcomings of traditional microseismic monitoring is using increasingly popular fiber optic cables for monitoring of slow and fast strain changes (DAS), which can allow to better track hydraulic fracture extent. Both our models and observations of different researchers (e.g. Zhang et al., 2020) show that in

DAS recordings we can observe characteristic patterns indicating approaching, hitting and crossing hydraulic fracture. Fiber optics technology still has its limitations and does not provide as high-quality seismic data as standard microseismic arrays with 3C geophones. Therefore at the moment the best solution is probably the combination of two technologies to get the best insight into the geomechanical processes happening in the subsurface. Using our simulation methodology and adopting proposed inversion technique to specific problem can also supplement and help in interpretation of the geophysical monitoring recordings.

### 6.7. Value of the presented research and possible future directions.

The research comes with a lot of scientific and business-related values. First, we develop the methodologies to study induced earthquakes, microseismicity, as well as fault and fracture mechanics. Secondly, we provide a tool that can be used to analyze past cases of seismicity and hydraulic fracturing operations, or to plan and optimize future operations by testing different scenarios of injection. This tool also allows for seismic hazard analysis associated with induced seismicity.

We present how to use our tool and approach to analyze fractures interactions and generated microseismicity. We also discuss how to establish the relationship between SRV and DRV. Our insights on these phenomena can be useful in analyses of real case studies and interpretation of available microseismic data.

Another piece in the developed toolbox is a method for the inversion of stress and selected parameters of the faults or fractures using hybrid a deterministic-stochastic approach. Better constraints on model parameters from this method translate into more reliable simulation results and predictions of future scenarios. Moreover, understanding of associated uncertainties can help in making educated decisions and avoid possible overinterpretation of the results.

In this research, we go beyond application of the developed methods and tools to simple, hypothetical models and perform analyses of the selected case studies. First case study is on the Timpson (TX) induced earthquake which we present to showcase a performance of our methods in Chapter 3 and Chapter 4. In Chapter 5 we present simulation results focused on the interactions of hydraulic fracture with natural fractures and generation of corresponding microseismic cloud for HFTS-1 site. In all these studies, simulations or inversion provide additional information on medium properties and nature of subsurface processes.

Overall, we were successful in providing new insights, methods and tools to analyze induced seismicity, hydraulic fracture propagation and associated microsismicity. We also acknowledge that there are still questions that can be addressed or a room for methodology improvement.

In numerical methodology part, better meshing and computational procedures to handle more complex geometries will be needed. Moreover, improving connection between quasi-static and dynamic simulations, and finally merging into fully dynamic scheme would allow to explore additional interesting and complex phenomena.
Also, more physics can be added to current schemes. Physics-based simulations are an art of identification of the most impactful processes and approximation of the reality without compromising the necessary complexity of the phenomena. Some of the possible improvements may include thermoporoelasticity, poroelastoplasticity, turbulent flows within fractures or using rate-and-state friction laws.

Another direction in future research is to apply the developed methodologies and tools to more case studies of wastewater disposal or hydraulic fracturing. The tools we develop are very powerful and we only show a part of their capabilities in presented studies. Along the same line, different operation scenarios can be explored and compared, such as different pumping schedules, multiple hydraulic fractures, different fracturing fluids etc. This type of the scenario testing is a way to help in improvements in operation design optimization, which translates in the efficiency of the stimulation, profit increase and potentially seismic hazard reduction.

## 6.8. References

Fan, Z., Eichhubl, P., & Gale, J. F. W. (2016), 'Geomechanical analysis of fluid injection and seismic fault slip for the Mw4.8 Timpson, Texas, earthquake sequence', *J. Geophys. Res. Solid Earth,* 121, pp. 2798–2812.

Shirzaei, M., Manga, M., & Zhai, G. (2019) 'Hydraulic Properties of injection formations constrained by surface deformation', *Earth Planet. Sci. Lett.*, **515**, pp. 125-134.

Zhang, Z., et al. (2020) 'Modeling of fiber-optic strain responses to hydraulic fracturing', *Geophysics*, *85*(6), pp. A45-A50.

# APPENDIX A

## DESCRIPTION OF THE BENCHMARK PROBLEM

The verification of the PyFluFlow code we benchmark its performance against the results that we get from the analytical solutions mentioned in the main text (Rudnicki, 1986). For that purpose, we setup simple homogenous, isotropic, 3D model with a point source and the following parameters of the medium and fluid:

- Bulk modulus of 25 GPa
- Drained Poisson ratio of 0.25 and undrained Poisson ratio of 0.3
- Skempton's coefficient of 0.7
- Fluid viscosity of 0.002 Pa·s
- Intrinsic permeability of $10^{-16}$ $m^2$

Benchmark model has cubic grid with 60 x 60 x 40 elements and spacing of 50 m in between them. There are 9 time steps with time step of 1 day. The injection rate is of 0.05 $m^3/s$ is constant. Program performs updates of pore pressures and stresses in 5 iterations and the fluid source is located in the central point of the model. Using the same parameters, we calculated the analytic solution in exactly the same grid of points and compared both solutions.

# APPENDIX B

## GAUSSIAN SMOOTHING EXPLANATION

Gaussian smoothing (filtration) is done by using 2D convolutional operator with kernel of the Gaussian curve shape (“bell curve”). The filter is defined by specifying the standard deviation of the filter, where larger standard deviation give more smoothing of the array. It has similar effect to mean filter but while averaging, it assigns higher weights to the points closer to the central point around which the mean is calculated. The example of Gaussian smoothing application is presented in Figure B.1.

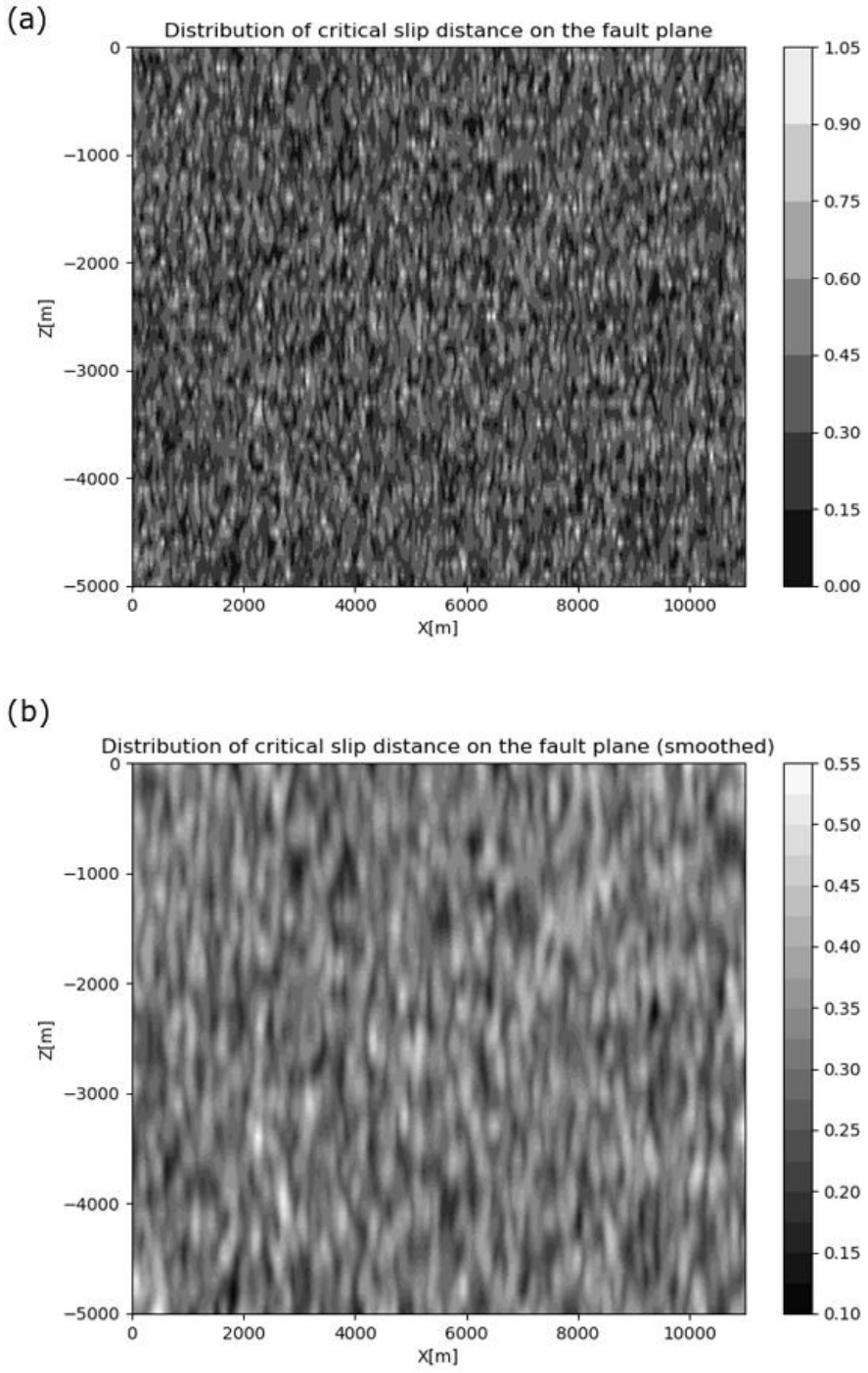


**Figure B.1 Comparison of the distributions of critical slip distance values picked from normal distribution without (a) and with (b) applied Gaussian smoothing filter. The size of heterogeneities is increased but the values contrasts are smaller after filtration.**

APPENDIX C

DISTRIBUTIONS OF FRICTIONAL PARAMETERS, PORE PRESSURE AND

STRESSES IN MODELS USED IN SIMULATIONS

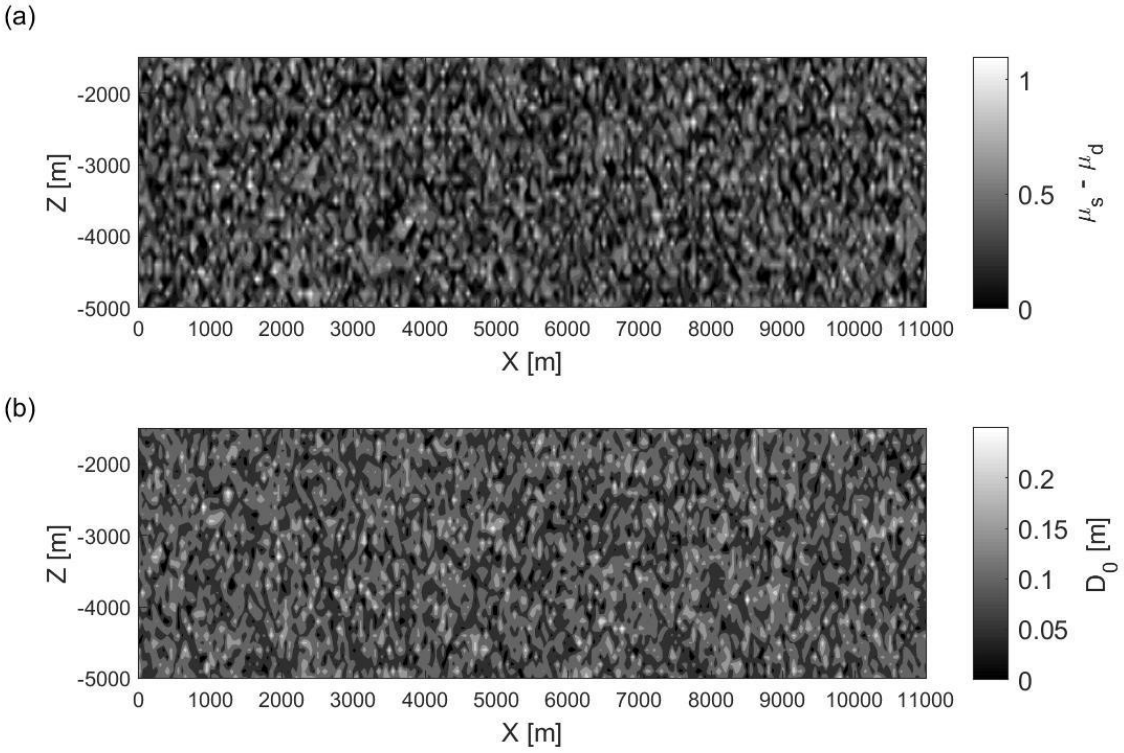


**Figure C.2 Distribution of (a) static and dynamic friction coefficient differences and (b) critical slip distance values on the fault plane (SE-NW from left to right) for Model 2 from Table 3.2. Normal distribution from which the values of friction coefficients are picked have the average of 0.6 and standard deviation of 0.2 for static friction coefficient and average of 0.3 and standard deviation of 0.2 for dynamic friction coefficient. Normal distribution from which the values of critical slip distance values are picked have the average of 0.05 m and standard deviation equal to 0.05.**

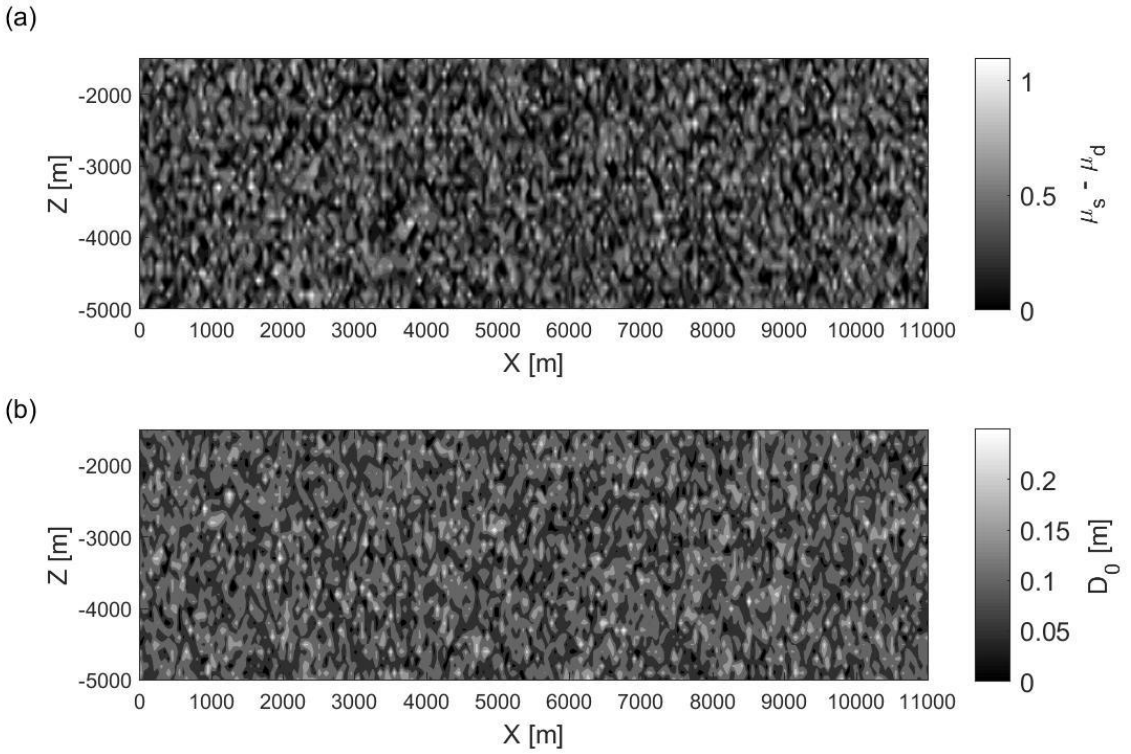


**Figure C.3 Distribution of (a) static and dynamic friction coefficient differences and (b) critical slip distance values on the fault plane (SE-NW from left to right) for Model 3 from Table 3.2. Normal distribution from which the values of friction coefficients are picked have the average of 0.6 and standard deviation of 0.2 for static friction coefficient and average of 0.3 and standard deviation of 0.2 for dynamic friction coefficient. Normal distribution from which the values of critical slip distance values are picked have the average of 0.1 m and standard deviation equal to 0.05.**

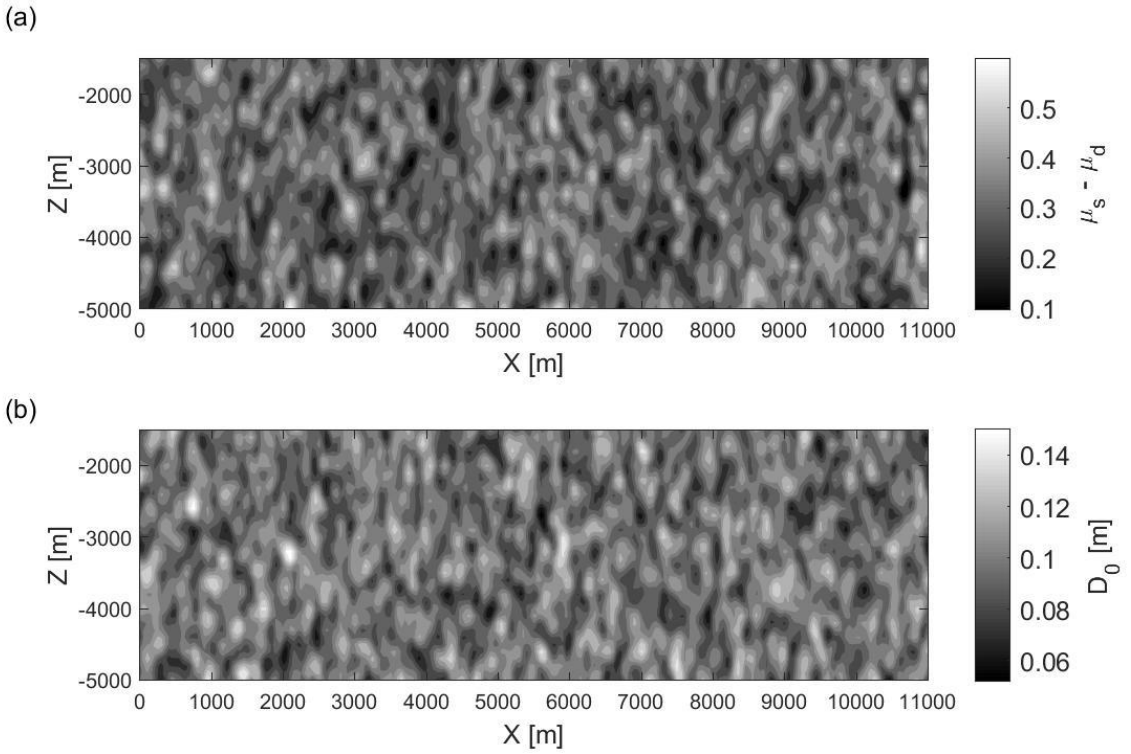


**Figure C.4 Distribution of (a) static and dynamic friction coefficient differences and (b) critical slip distance values on the fault plane (SE-NW from left to right) for Model 4 from Table 3.2. Normal distribution from which the values of friction coefficients are picked have the average of 0.6 and standard deviation of 0.2 for static friction coefficient and average of 0.3 and standard deviation of 0.2 for dynamic friction coefficient. Normal distribution from which the values of critical slip distance values are picked have the average of 0.1 m and standard deviation equal to 0.05. The obtained distribution is smoothed with the Gaussian filter.**

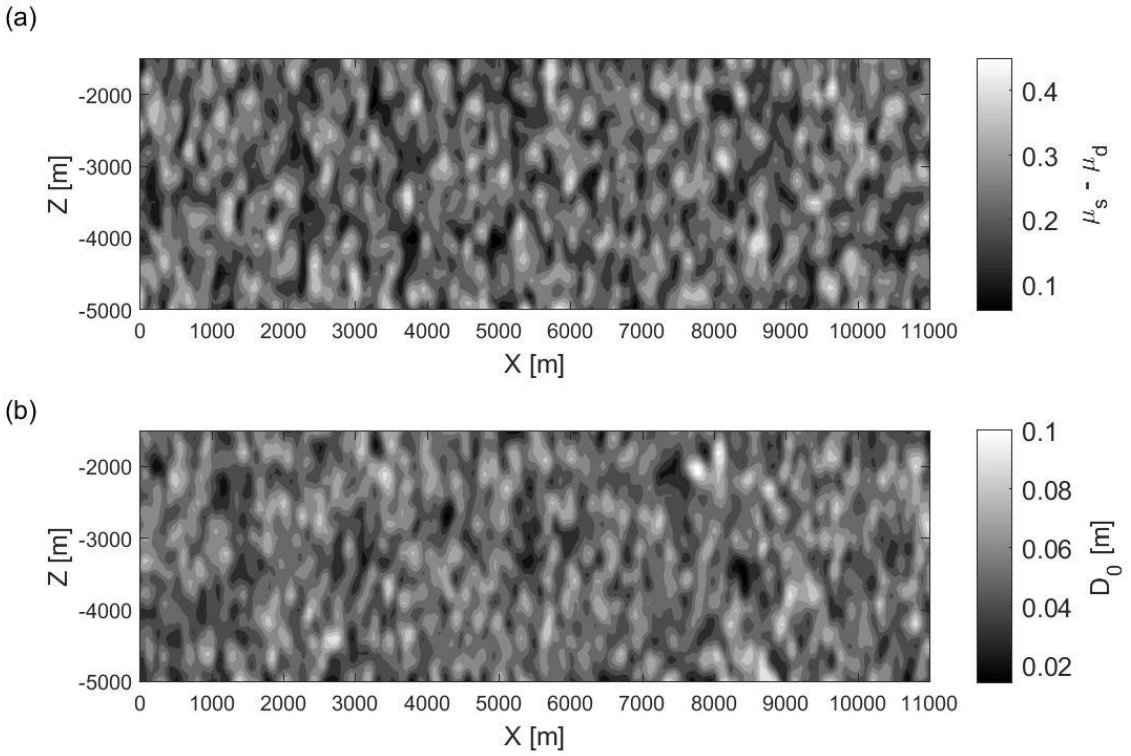


**Figure C.5 Distribution of (a) static and dynamic friction coefficient differences and (b) critical slip distance values on the fault plane (SE-NW from left to right) for Model 4 from Table 3.2. Normal distribution from which the values of friction coefficients are picked have the average of 0.6 and standard deviation of 0.2 for static friction coefficient and average of 0.4 and standard deviation of 0.2 for dynamic friction coefficient. Normal distribution from which the values of critical slip distance values are picked have the average of 0.1 m and standard deviation equal to 0.05. The obtained distribution is smoothed with the Gaussian filter.**

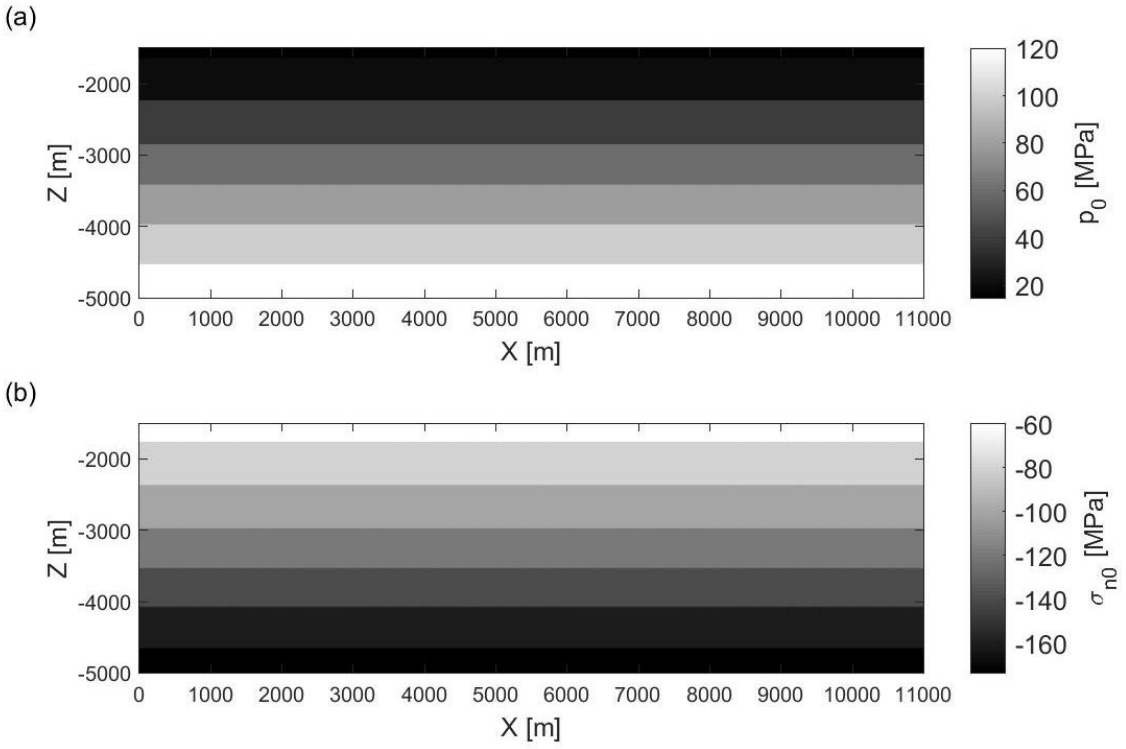


**Figure C.6 Distribution of initial (before the start of injection) (a) pore pressure and (b) normal stress (SE-NW from left to right) for Model 1 from Table 3.2. The distribution was obtained by combination of keeping constant effective normal stress for the entire depth interval with the fault (which means an overpressure) and multiplication of vertical lithostatic stress value at a certain depth by the value of normal-to-vertical stress ratio value of 1.4 in (b) and multiplication of yield stress at a given point by the value of initial-to-yield shear stress ratio value of 0.6 which results in uniform initial shear stress of 13.1 MPa.**

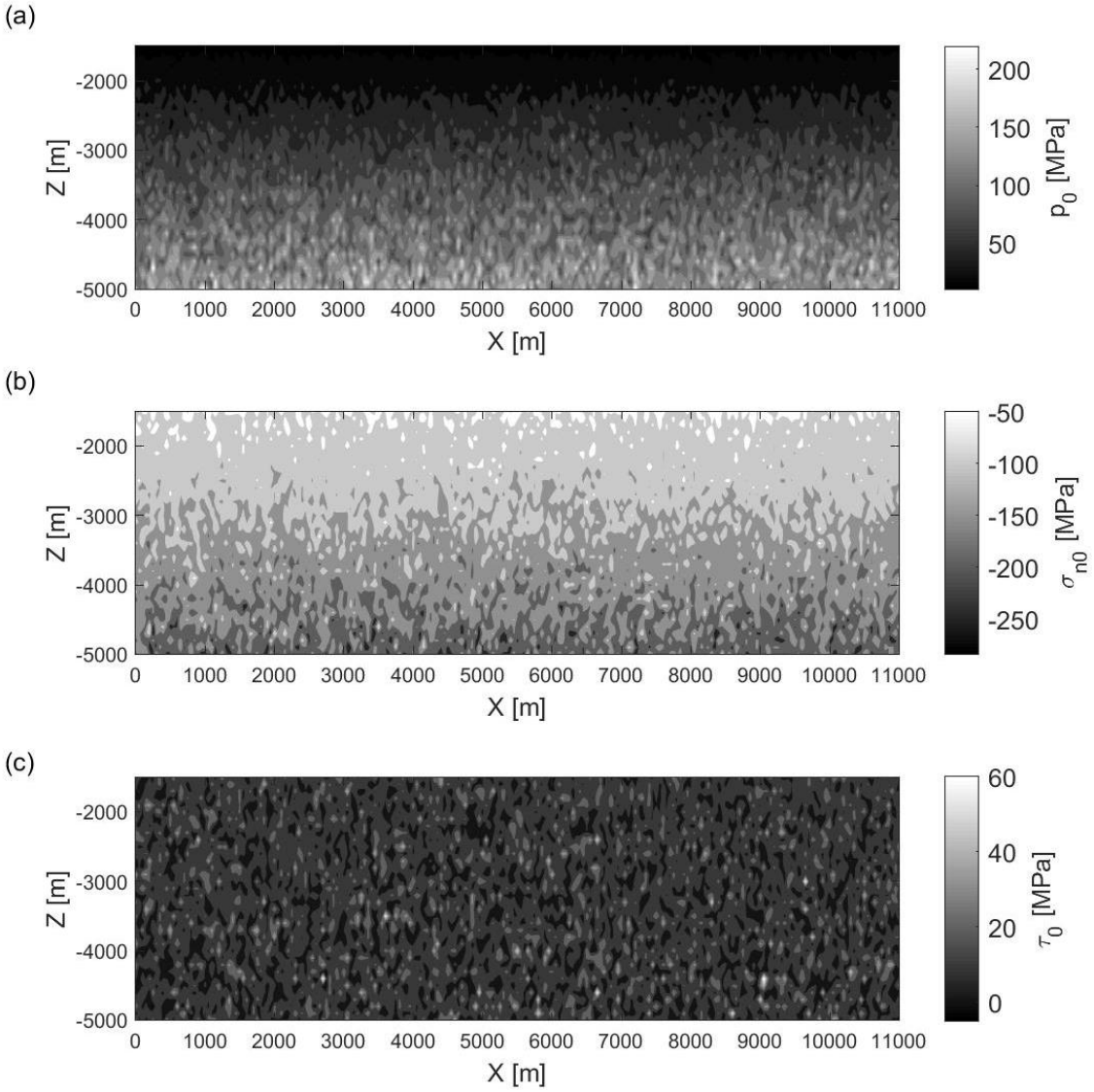


**Figure C.7 Distribution of initial (before the start of injection) (a) pore pressure, (b) normal stress and (c) shear stress on the fault plane (SE-NW from left to right) for Model 2 from Table 3.2. The distribution was obtained by combination of keeping constant effective normal stress for the entire depth interval with the fault (which means an overpressure) and randomly picked deviations from the trend with normal distribution with average of 0 and standard deviation of 0.1 (the fraction of value at a given depth) in (a), multiplication of vertical lithostatic stress value at a certain depth by the value of normal-to-vertical stress ratio picked randomly from the normal distribution with average of 1.4 and standard deviation of 0.3 in (b) and multiplication of yield stress at a given point by the value of initial-to-yield shear stress ratio picked randomly from the normal distribution with average of 0.6 and standard deviation of 0.1 in (c).**

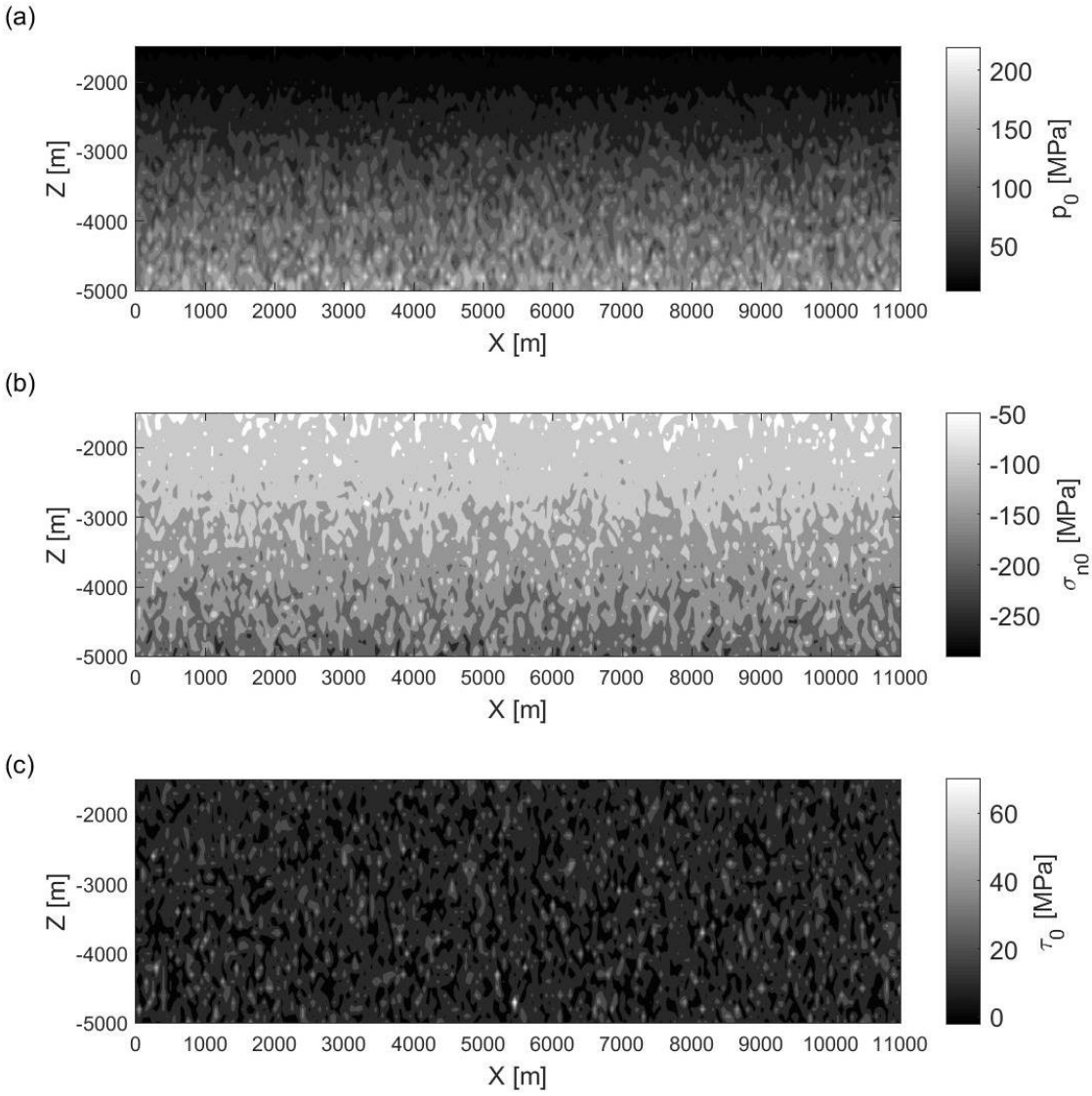


**Figure C.8 Distribution of initial (before the start of injection) (a) pore pressure, (b) normal stress and (c) shear stress on the fault plane (SE-NW from left to right) for Model 3 from Table 3.2. The distribution was obtained by combination of keeping constant effective normal stress for the entire depth interval with the fault (which means an overpressure) and randomly picked deviations from the trend with normal distribution with average of 0 and standard deviation of 0.1 (the fraction of value at a given depth) in (a), multiplication of vertical lithostatic stress value at a certain depth by the value of normal-to-vertical stress ratio picked randomly from the normal distribution with average of 1.4 and standard deviation of 0.3 in (b) and multiplication of yield stress at a given point by the value of initial-to-yield shear stress ratio picked randomly from the normal distribution with average of 0.6 and standard deviation of 0.1 in (c).**

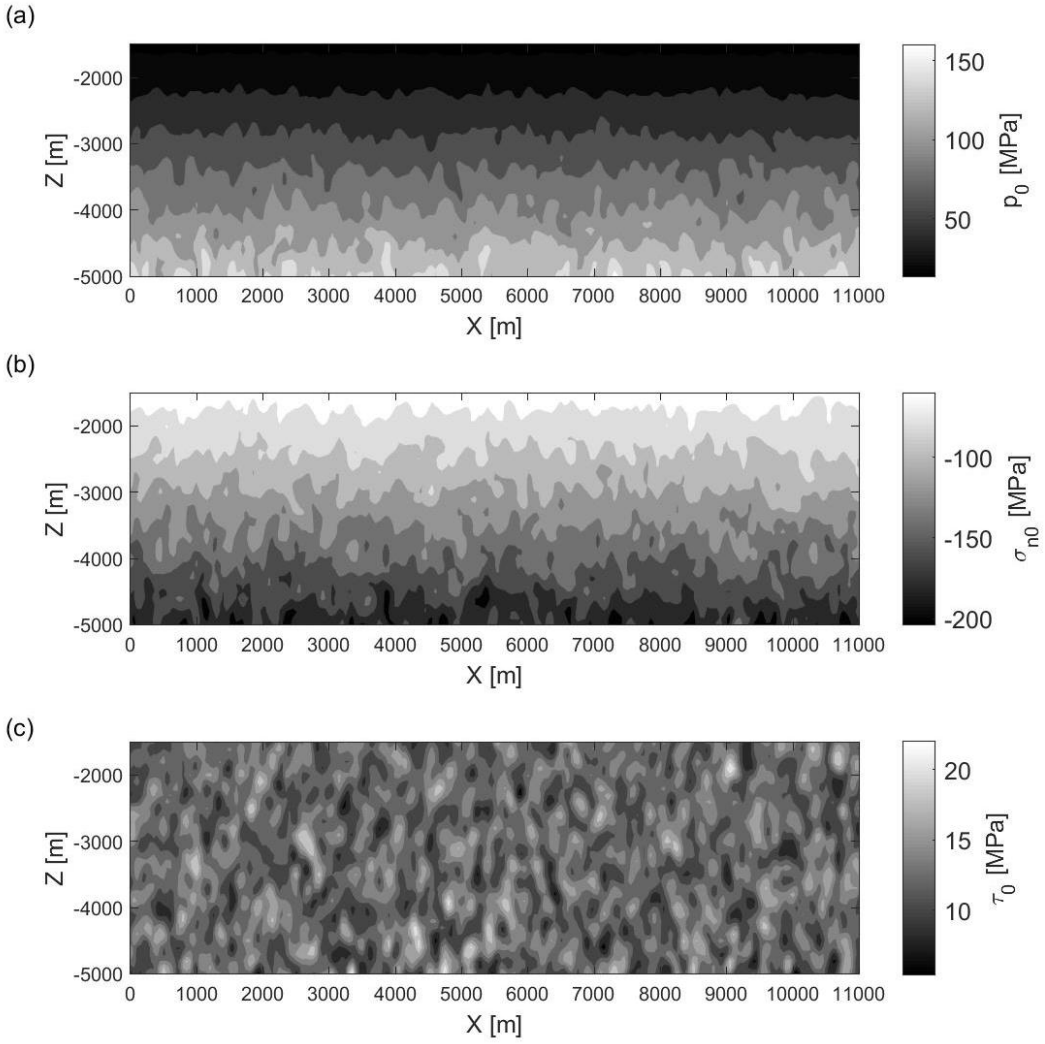


**Figure C.9 Distribution of initial (before the start of injection) (a) pore pressure, (b) normal stress and (c) shear stress on the fault plane (SE-NW from left to right) for Model 4 from Table 3.2. The distribution was obtained by combination of keeping constant effective normal stress for the entire depth interval with the fault (which means an overpressure) and randomly picked deviations from the trend with normal distribution with average of 0 and standard deviation of 0.1 (the fraction of value at a given depth) in (a), multiplication of vertical lithostatic stress value at a certain depth by the value of normal-to-vertical stress ratio picked randomly from the normal distribution with average of 1.4 and standard deviation of 0.3 in (b) and multiplication of yield stress at a given point by the value of initial-to-yield shear stress ratio picked randomly from the normal distribution with average of 0.6 and standard deviation of 0.1 in (c). The distributions of all randomly picked parameters were smoothed with Gaussian filter before multiplications.**

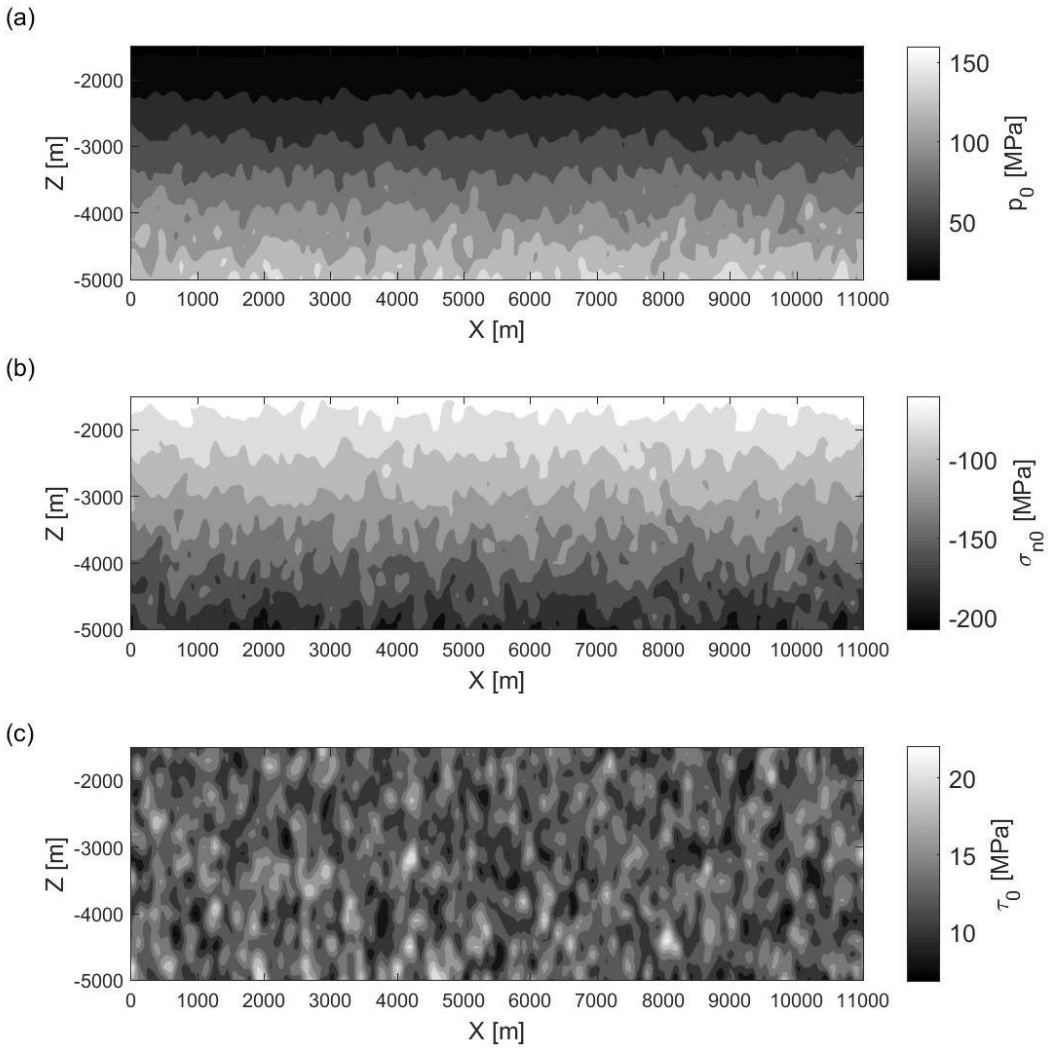


**Figure C.10 Distribution of initial (before the start of injection) (a) pore pressure, (b) normal stress and (c) shear stress on the fault plane (SE-NW from left to right) for Model 5 from Table 3.2. The distribution was obtained by combination of keeping constant effective normal stress for the entire depth interval with the fault (which means an overpressure) and randomly picked deviations from the trend with normal distribution with average of 0 and standard deviation of 0.1 (the fraction of value at a given depth) in (a), multiplication of vertical lithostatic stress value at a certain depth by the value of normal-to-vertical stress ratio picked randomly from the normal distribution with average of 1.4 and standard deviation of 0.3 in (b) and multiplication of yield stress at a given point by the value of initial-to-yield shear stress ratio picked randomly from the normal distribution with average of 0.6 and standard deviation of 0.1 in (c). The distributions of all randomly picked parameters were smoothed with Gaussian filter before multiplications.**