Title

Physics-based hydrogeologic models of fluid-fault interactions: implications of natural and anthropogenic poroelastic effects on induced seismicity

Author

Hill, Ryley

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO
SAN DIEGO STATE UNIVERSITY

Physics-based hydrogeologic models of fluid-fault interactions: implications of natural and anthropogenic poroelastic effects on induced seismicity

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Geophysics

by

Ryley Hill

Committee in charge:

University of California San Diego

Professor Yuri Fialko
Professor Julian McAuley

San Diego State University

Professor Matthew Weingarten, Chair
Professor Xiaobai Liu
Professor Thomas K. Rockwell

2023

The dissertation of Ryley Hill is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

_____

_____

_____

_____

_____

Chair

University of California San Diego

San Diego State University

2023

DEDICATION

To my mentors, friends, family, and Eliza. Your constant love and support is what made this possible.

EPIGRAPH

All models are wrong, but some are useful.

George E.P. Box

It is said by the Eldar that in water there lives yet the echo of the Music of the Ainur more than in any substance that is in this Earth...

J.R.R. Tolkien

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, I would like to express my deepest appreciation to my advisor, Dr. Matthew Weingarten, who was always patient, kind, and supportive academically and personally throughout my entire time as his first ever doctoral student. As his first student I was expecting there to be much more turmoil, but we worked extremely well together and I hope we can continue our professional and social relationships as time progresses. He was considerably lenient to the large amount of remote work and travel that I went on and allowed me the freedom and respect to work in the ways most efficient for me which were often not a typical 9 to 5 schedule. I think we both would agree we turned into a great team by the end. Many thanks to Dr. Yuri Fialko who guided me through many challenging and technical aspects of my studies. Yuri was always willing to help with whatever questions I had and many times had the perfect and swift response to very difficult questions. I would also like to thank Dr. Thomas K. Rockwell, Dr. Robert Guyer, Dr. Daniel Trugman, and Dr. Andrea Hampel who were paramount to the success of the projects I worked on and were always willing to lend aid or discussion related to the science I was a part of. I want to extend my academic appreciation to all the professors and mentors at IGPP who were always so kind and humble in their approaches to teaching. From other students to professors there is a great environment at both the UCSD campus and SDSU campus that is conducive to a healthy working environment.

Over the years at SIO and SDSU, I have made many friendships. Thank you to the 2018 cohort (Nick Lau, Yao Yu, Yue Du) and the rest of the 2018 incoming class from the other departments. We fostered a great culture that we all thrived in. Special thanks to Nick Lau and the many long hours we spent working on homework in the Keller just the two of us. We have an amazing friendship and bond that I only regret not taking more time out of my busy life to cultivate more. Similarly, Ruei-Jiun Hung was my mentee in my second year and in many ways he is more my mentor often teaching me about concepts. I appreciate all our venting moments together and friendship. Other friends like Drake, Maya, Susheel, Wes, Jacob, Zoe, Chloe, Derick,

Jianhua, Andrea, and Zeyu. I will cherish all our time we have spent together. I hope we can continue to be friends our whole lives.

I must also acknowledge my friends and family who make up a large part of my life. Thank you to my teammates on the ultimate frisbee teams I played for while in San Diego. A special thanks to Goose Helton who effectively saved my life with nutritional advice and was always an amazing life coach. All of the Reno ultimate frisbee crew- the real ones who know who they are. Special thanks to Gabe, Pete, and Nicole for our daily group chat memes and jokes. Thank you to my mother, father, and sister who have never once wavered in their unconditional support of me. Truly, I would have struggled financially without the support of my parents. I am so grateful for my parents generosity. My dad gave me his car when mine broke down and never once said no to a request for money if I asked.

Last but not least, I could not have gone through this journey without my partner, Eliza McCall. I could write about the support and companionship she has provided for me through all the struggles that would be longer than this thesis. I reserve myself to say simply this, "You must allow me to tell you how ardently I admire and love you".

Chapter 2, in full, is a reprint of the material as it appears in Nature 2023.: Hill, R.G., Weingarten, M., Rockwell, T.K., Fialko, Y., 2023. Major southern San Andreas earthquakes modulated by lake-filling events. Nature 618, 761–766. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reformatted version of a paper currently submitted for submission for publication. Hill, R.G., Weingarten, M., Langenbruch, C., Fialko, Y., 2023. Mitigation and optimization of induced seismicity using physics-based forecasting. Journal of Geophysical Research: Solid Earth, submitted. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is currently being prepared for submission for publication of the material. Hill, R.G., Trugman, D., Weingarten, M., 2023. Deciphering earthquake triggering mechanisms

with a fully coupled poroelastic model and machine learning analysis: application to the case of Paradox Valley Unit, Colorado. Geophysical Research Letters, in prep. The dissertation author was the primary investigator and author of this paper.

VITA

2015	B.S. in Geophysics, Physics, Applied Math, University of Nevada, Reno
2018	M.S. in Geophysics, University of Nevada, Reno
2023	Ph.D. in Geophysics, University of California San Diego & San Diego State University

PUBLICATIONS

Hill, R.G., Weingarten, M., Rockwell, T.K., Fialko, Y., 2023. Major southern San Andreas earthquakes modulated by lake-filling events. Nature 618, 761–766

Hill, R.G., Weingarten, M., Langenbruch, C., Fialko, Y., 2023. Mitigation and optimization of induced seismicity using physics-based forecasting. Journal of Geophysical Research: Solid Earth. *submitted.*

Hill, R.G., Trugman, D., Weingarten, M., 2023. Deciphering earthquake triggering mechanisms with a fully coupled poroelastic model and machine learning analysis: application to the case of Paradox Valley Unit, Colorado. *In preparation for publication*.

ABSTRACT OF THE DISSERTATION

Physics-based hydrogeologic models of fluid-fault interactions: implications of natural and anthropogenic poroelastic effects on induced seismicity

by

Ryley Hill

Doctor of Philosophy in Geophysics

University of California San Diego, 2023
San Diego State University, 2023

Professor Matthew Weingarten, Chair

In this dissertation, I use the finite element method to solve the fully coupled poroelastic time-dependent pore pressure and stress changes within the Earth's crust due to hydrological loads and wastewater injection in order to have an improved understanding of fluid-fault interactions, hazard associated with induced seismicity, and earthquake triggering processes. Chapter 1 is an introduction to the fundamental concepts of the role of fluids in faults, linear poroelasticity, and the primary means of characterizing fault instability for induced seismicity, which are all incorporated throughout the following chapters. Chapter 2 studies the Coulomb stress change

induced by the hydrological loading of ancient Lake Cahuilla on the Southern San Andreas Fault. Chapter 3 studies induced seismicity associated with the the wastewater injection at the Raton Basin, CO while developing a universal hazard mitigation and fluid optimization methodology. Chapter 4 is another wastewater injection study for the Paradox Valley Unit, CO that incorporates machine learning techniques to decipher earthquake triggering mechanisms associated with induced seismicity.

Chapter 1

Introduction

1.1 Motivation: the role of fluid in seismicity

Fluid-fault interaction in the Earth's crust can cause earthquakes (Ellsworth, 2013). Subsurface fluid changes can cause spatial and temporal changes in pore pressure and stress that contribute to fault (in)stability (Segall and Lu, 2015). Increasing fluid pore pressure on faults reduces the effective normal stress (Terzaghi *et al*., 1943) and makes earthquakes more likely. Only small changes in stress (0.01-0.1 MPa) (Reasenberg and Simpson, 1992; Stein, 1999) are needed to reactivate faults and cause earthquakes. A variety of natural and anthropogenic pore pressure and stress changes can reactive faults. A non-exhaustive list includes natural tectonic plate motion, surface hydrologic loads, production/injection of fluids, and a host of industrial activity (e.g. hydraulic fracturing, CO_2 sequestration, and enhanced geothermal exploration).

With increasing industrial activities across the globe, the rise of human-induced earthquakes also increased substantially over the past two decades and is projected to increase further (Foulger *et al*., 2018; Keranen and Weingarten, 2018). A considerable amount of seismicity was observed in historically aseismic regions, particularly in regions with large volumes of wastewater disposal (Ellsworth, 2013). Scientifically, injection-induced seismicity created a

natural laboratory that provided scientists with invaluable data to better understand the earthquake process. Societally, the moderate magnitude induced events (M5+) created risk in historically aseismic regions. Several induced earthquakes have had damaging effects, including the 2011 M5.6 Prague, Oklahoma earthquake, the 2011 M5.3 Trinidad CO earthquake and the 2012 M4.8 Timposon Texas earthquake (Keranen *et al*., 2013; Rubinstein *et al*., 2014; Frohlich *et al*., 2014).

Moderate magnitude events associated with industrial activities spurred an area of ongoing research investigating how to best mitigate the seismic hazard/risk associated with industrial activities (Schultz *et al*., 2021). Currently, traffic light protocols are implemented in order to manage and reduce risks of induced seismicity (Bommer *et al*., 2006; Zoback, 2012). However, these protocols remains have been implemented with varying degree of success (Woo *et al*., 2019; Ellsworth *et al*., 2019). Recent efforts have redefined improved traffic light protocols through a variety of different methods (Douglas and Aochi, 2014; Mignan *et al*., 2017; Baisch *et al*., 2019; Cremen and Galasso, 2020; Langenbruch *et al*., 2020).

Part of the difficulty in choosing what strategies are best to mitigate induced hazard/risk is the ambiguity in forecasting the magnitude of the events caused by human-induced changes. For example, it remains somewhat in contention whether or not the magnitude of earthquakes are a function of the total volume of fluid injected or as large as statistically expected (McGarr, 2014; Van der Elst *et al*., 2016). Additionally, it is clear that the rate of fluid pressure increase also contributes to increasing seismicity (Toda *et al*., 2002; Qin *et al*., 2022). However, models agree generally that continuing fluid injection increases the chance of a larger earthquake. The fundamental mechanism, in most instances, is not that the industrial activities are increasing the total energy of the system, but instead are accelerate preexisting and critically stressed faults towards failure (Walsh III and Zoback, 2015; Zoback and Townend, 2001).

Therefore, in nearly all cases of human-induced seismicity, there exists stress changes that trigger seismicity on preexisting and critically stressed faults (McGarr *et al*., 2002). The fluid effects that promote and cause failure of a particular fault may not be the same effects that cause

failure in successive earthquakes (Glasgow *et al*., 2021a; Zaliapin and Ben-Zion, 2016). The discrimination of events that were caused by the release of tectonic energy and those triggered by human-induced stress changes is capable of revealing important characteristic of the earthquake process (Eyre *et al*., 2019; Glasgow *et al*., 2023). Furthermore, efforts have extended to detailed simulations of induced seismicity to explore the physical conditions that may result in larger magnitude events, like runaway or unbounded ruptures that extend outside the perturbable region of a fault (Gischig, 2015; Norbeck and Horne, 2018; Kroll and Cochran, 2021).

Outside of the realm of industrial activities, hydrological loads from natural and anthropogenic sources can contribute to increases in pore pressure, flexural stresses, and poroelastic effects that stimulate seismicity (Simpson, 1976; Talwani, 1997; Gupta, 1992; Roeloffs, 1988b; Gupta, 2002; Luttrell *et al*., 2007; Brothers *et al*., 2011). Stimulated seismicity may occur shortly after reservoir impoundment of the hydrologic load from poroelastic coupling, or at longer timescales due to the increase of pore pressure at depth as a result of permeability dependent fluid diffusion (Talwani, 1997; Rajendran and Talwani, 1992). Permeability is enhanced within fault damage zones, which decays with distance away from the fault core (Fialko *et al*., 2002a; Cochran *et al*., 2009; Mitchell and Faulkner, 2009). The increased permeability act as fluid pathways that can facilitate fluid transport to seismogenic depths.

It is clear that our understanding of the role of fluid in seismicity is extensive, ongoing, and growing (Ge and Saar, 2022). In all instances of fluid induced seismicity discussed, the accurate characterization of fluid-fault interactions is dependent on calculations of pore pressure and stress through space and time. Therefore, the necessity of physics-based models that can accurately resolve the spatio-temporal pore pressure and stress distribution is critical. To evaluate pore pressure and stress conditions requires time-dependent coupling between the porous media and fluid pressure as well as the fluid flow within the media addressed by the theory of poroelasticity.

1.2 Linear Poroelasticity

The theory of linear poroelasticity requires the linear constitutive relations described by Biot (1941) coupled to fluid flow of Darcy's law (Darcy, 1856). Two formulations of Biot poroelasticity widely used today are Rice and Cleary (1976a) and Wang (2000). The geophysics community has widely adopted Rice and Cleary (1976a), with the rock mechanics community widely adopting Wang (2000). I prefer to use Wang's formulation as it more clearly resembles the equations solved in the finite-element method employed in this dissertation (Wang, 2000). The constitutive equations, six for stress (or strain), and one for pore pressure (or increment of fluid content) (the latter depending on what is chosen as the independent or dependent variable) are combined with three (one for each dimension) force equilibrium equations and an additional pressure diffusion equation by merging Darcy's law with the requirement of conservation of fluid mass (Wang, 2000).

Accordingly, the governing equations for a fully coupled linear poroelastic 3D medium are *definable* as (Wang, 2000):

$$G\nabla^2 u_i + \frac{G}{1-2\nu}\frac{\partial^2 u_j}{\partial x_i \partial x_j} = \alpha\frac{\partial P}{\partial x_i} - F_i \tag{1.1}$$

$$\alpha\frac{\partial \varepsilon_{ii}}{\partial t} + S_\varepsilon\frac{\partial P}{\partial t} = \frac{k}{\mu}\nabla^2 P + Q \tag{1.2}$$

Where G is the shear modulus, u the displacement, ν the Poisson's ratio, α the Biot-Willis coefficient, F the body force, k the permeability, μ the fluid viscosity, S_ε the constrained specific storage, ε_{ii} the dilatancy (trace of the strain tensor), P the pore pressure, and Q the fluid source (Wang, 2000). These equations are nearly identical to the classic equations for linear elasticity and diffusion of pore pressure, except for the coupling of pore pressure in the conservation of momentum equation (1.1) and the fluid flow coupled to strain by the requirement of fluid continuity (1.2).

There are some often mistaken similarities, yet important differences, with these equations and variables compared with hydrogeologic definitions of fluid flow. The hydrogeologic specific storage (S_s, [m^{-1}]) is defined as the volume of water release per unit decline of head per unit bulk volume while maintaining the representative element volume (REV) in a state of zero lateral strain and constant vertical stress:

$$S_s = \frac{1}{V_a}\frac{dV_w}{dh} = \frac{1}{V_a}\frac{dV_w}{dP}\frac{dP}{dh} = \frac{1}{V_a}\frac{dV_w}{dP}\gamma_w \tag{1.3}$$

The hydrogeologic one-dimensional (uniaxial) specific storage S_s can also be represented in terms of pressure ($S = \frac{S_s}{\rho_f g}$). With the proper substitutions (Wang, 2000) we arrive at Jacob's equation (Jacob, 1950), a common mechanical representation of hydrogeologic specific storage:

$$S = \frac{\alpha^2}{K_v} + \phi\left(\frac{1}{K_f} - \frac{1}{K_\phi}\right) \tag{1.4}$$

$$S = \frac{1}{K_v} + \phi\frac{1}{K_f} \tag{1.5}$$

Where K_v is the uniaxial drained bulk modulus, K_f the fluid bulk modulus, K_ϕ the unjacketed pore bulk modulus, and K_s' is the unjacketed bulk modulus (the solid grain bulk modulus). Equation (1.5) is under the assumption that solid grains and pores are incompressible ($\frac{1}{K_s'} = \frac{1}{K_\phi} = 0$). Note that $\alpha = 1 - \frac{K}{K_s'}$. It is important to realize that S here is not the same as S_ε found in Equation (1.2) The major difference being that S_ε is the volume of water released per unit decline of head per unit volume *while holding the control volume constant*. Therefore, S_ε will always be smaller than S since the changing volume can not contribute to the fluid released. It is helpful to represent the

constrained specific storage in terms of compressibilities:

$$S_\varepsilon = \frac{1}{K_s'}(1-\frac{K}{K_s'})+\phi(\frac{1}{K_f}-\frac{1}{K_\phi}) \tag{1.6}$$

$$S_\varepsilon = \phi\frac{1}{K_f} \tag{1.7}$$

Equation (1.7) is under the assumption that solid grains and pores are incompressible ($\frac{1}{K_s'} = \frac{1}{K_\phi} = 0$). Comparing (1.5) and (1.7) yields an important definition of the storage coefficients $S_\varepsilon \leq S$ which through more rigorous means can be shown to always be the case without the assumption of incompressible grains or fluids when $\nu > -1$ (Wang, 2000).

Under uniaxial strain and constant vertical stress, ie. S is the specific storage variable, then equation (1.2) can be simplified to the classic hydrogeology groundwater flow equation (Wang, 2000):

$$S\frac{\partial P}{\partial t} = \frac{k}{\mu}\nabla^2 P + Q \tag{1.8}$$

$$S_s\frac{\partial h}{\partial t} = K_H\nabla^2 h + Q \tag{1.9}$$

Compared with equation (1.2), the major differences are the change of specific storage and the coupling of the strain field, ie. $\frac{\alpha\partial\varepsilon_{kk}}{\partial t}$ into the fluid flow equation. The assumptions of uniaxial strain and constant vertical stress misrepresent the strain field at all REV in the domain since the fluid flow also distorts the strain field.

The reason for elaborating on the differences between the fully coupled poroelastic equations and the groundwater flow equations is due to use of these equations in modeling fluid flow in geological settings. Different simplifications to the choice of governing equations will result in different pore pressure and stress calculations through the physical model domain. Due to the coupling found in equations (1.1 and 1.2), analytical solutions are limited and numerical

models using the finite-element method are the only viable solution for 3D heterogeneous real-Earth problems.

In this dissertation, I resolve the fully coupled poroelastic equations (1.1 and 1.2) by building and solving numerical models with the finite-element software Abaqus (Dassault Systemes, 2020). Large Earth models are computationally expensive and we therefore required resources from the CSRC high-performance computing cluster at San Diego State University (Computational Science Research Center, 2023). The spatial evolution of pore pressure and stress is often post-processed in Matlab in order to characterize the fluid effects on faults and characterize fault instability.

1.3 Characterizing Fault Instability

An important development in visualizing fault instability is the graphical tool called the Mohr circle (Mohr, 1882). All possible orientations of shear and normal stress (for plane stress) acting on different angled failure planes can be described by a semicircle, called a Mohr circle (Figure 1.1).

Combining the the Mohr circle with a failure threshold allows for the characterization of fault instability. The Mohr-Coulomb failure criterion (Coulomb, 1773) can be represented as the linear envelope obtained from a plot of the shear stress versus the effective normal stress (Figure 1.1 - red line). This envelope represents a critical shear stress, that when reached by the material, will cause failure. Any configuration of stress below this threshold is stable and anything tangent or greater is unstable. This line is expressed as:

$$|\tau_{max}| = C + \mu_f \sigma_n \tag{1.10}$$

Where $|\tau_{max}|$ is the maximum shear strength limit, C is the cohesion, μ is the coefficient of friction

(Byerlee, 1978a), and σ_n is the effective normal stress.

The theory of effective stress was first developed by Terzaghi (Terzaghi *et al.*, 1943) to describe how pore pressure changes the state of stress. The effective stress σ_n is defined as (Nur and Byerlee, 1971):

$$\sigma_n = \sigma + P \tag{1.11}$$

$$\sigma_n = \sigma_{ij} + \delta_{ij}\alpha P \tag{1.12}$$

Pore pressure (P) only effects the normal stress but not the shear stress. Increasing pore pressure will reduce the effective normal stress, which shifts the Mohr-circle closer to the failure envelope. Thus, increasing pore pressure will always reduce the effective stress and promote failure. It is important to note that poroelastic stresses will change the shape of the circle as well e.g., Fig. 1 Keranen and Weingarten, 2018.

In all chapters of this thesis, increasing pore pressure from hydrologic loading or the injection of fluid is critically important to the change in effective stress. Decreasing effective stress due to increasing pore pressure always contributes to the destabilization of faults. The change in pore pressure often represents the largest component of the Coulomb stress change in our models. When this is not the case, poroelastic stress contributes more to failure on the fault and is frequently important in the far-field (Segall and Lu, 2015). In our models, we define Coulomb stress changes as the relative change from background conditions (King *et al.*, 1994a; Cocco, 2002):

$$\Delta CFS = \Delta\tau_s + \mu(\Delta\sigma_n + \Delta P) \tag{1.13}$$

in which $\Delta\tau_s$ is the change in shear stress, μ is the coefficient of friction (Byerlee, 1978a; Sibson, 1994; Fialko and Jin, 2021b), $\Delta\sigma_n$ is the change in normal stress (increases in compression are deemed negative) and ΔP is the change in pore-fluid pressure (increases in fluid pressure are deemed positive). The sign convention indicates that a positive change in the effective stress $\Delta\sigma_n$,

as well as a positive change in shear stress $\Delta\tau_s$, promote failure.

1.4 Dissertation Summary

This dissertation investigates how fully coupled poroelastic models can improve our understanding of fluid-fault interactions from hydrological loads and wastewater injection. This dissertation is structured follows: In Chapter 2, we use a numerical model and paleoseismic data to study the poroelastic stress induced by ancient Lake Cahuilla on the Southern San Andreas Fault over the past millennium. We find that stress and stressing rate changes were likely sufficient for triggering major earthquakes in the paleoseismic record. In Chapter 3, we combine statistical methods with a poroelastic model of to produce a map of fluid injection to forecast seismic hazard associated with wastewater injection in the Raton Basin. We also build an optimization framework on top of this forecasting model in order to maximize wastewater injected while reducing the seismic hazard. In Chapter 4, we combine machine learning random forests with a poroelastic model of fluid injection to forecast seismicity rate at the Paradox Valley Unit, CO. The induced seismicity is further treated with a game theory analysis to procure features that contributed the most to help decipher earthquake triggering mechanisms of the induced earthquakes. The analysis is complemented by a nearest neighbor distance clustering analysis to gain further insight into the type of earthquakes that are triggered more predominantly from injection stress stress direction or earthquakes triggered predominantly by prior earthquakes.

Tables and Figures

Table 1.1: Symbols (by order of appearance).

G	shear modulus
u	displacement
ν	Poisson's ratio
α	Biot-Willis coefficient
k	permeability
μ	fluid viscosity
S_ε	constrained specific storage
P	pore pressure
S_s	hydrogeologic uniaxial specific storage
V_a	bulk volume from which water is released
V_w	volume of water released from storage
h	hydraulic head
γ_w	specific weight of water
S	uniaxial specific storage
ρ_f	fluid density
K_v	uniaxial drained bulk modulus
K_f	fluid bulk modulus
K_ϕ	unjacketed pore bulk modulus
K_s'	unjacketed bulk modulus
K_H	hydraulic conductivity
$\lvert\tau_{max}\rvert$	maximum shear strength limit
C	cohesion
μ_f	coefficient of friction
σ_n	effective normal stress
CFS	Coulomb failure stress
τ_s	shear stress

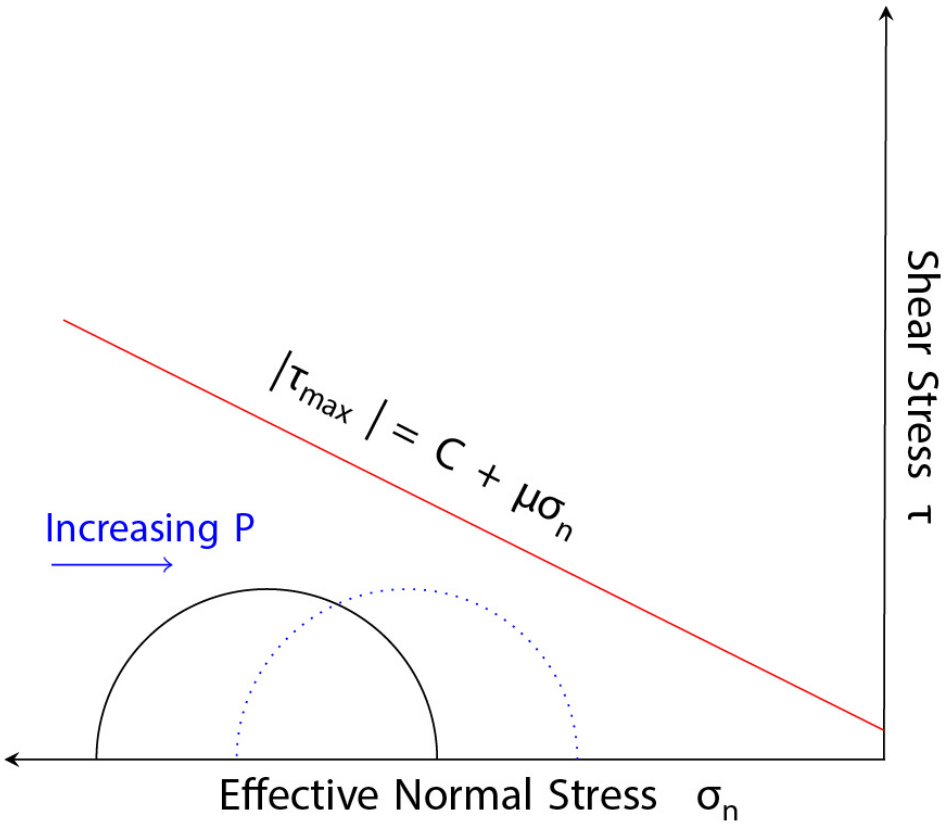


Figure 1.1: Mohr-Coulomb Diagram illustrating the effect of pore on the effective normal stress. The leftmost intersections of the Mohr circle and the x-axis is the σ_{max} and the rightmost intersection of the Mohr circle is the σ_{min}. Both are equally reduced by the increasing pore pressure which shifts the Mohr circle closer to the failure envelope. The dash circle represents the change from the initial state with increasing pore pressure. During injection or hydrological loading the shift towards failure increases the chance of an earthquake.

Chapter 2

Major Southern San Andreas Earthquakes Modulated by Lake Filling Events

Hydrologic loads can stimulate seismicity in the Earth's crust (Talwani, 1997), however evidence for triggering of large earthquakes remains elusive. The Southern San Andreas Fault (SSAF) in Southern California lies next to the Salton Sea (Tostrud, 1997), a remnant of ancient Lake Cahuilla that periodically filled and desiccated over the last millennium (Waters, 1983; Philibosian *et al*., 2011; Rockwell *et al*., 2018). Here, we use new geologic and paleoseismic data to demonstrate that the past 6 major earthquakes on the SSAF likely occurred during highstands of Lake Cahuilla (Rockwell *et al*., 2018; Rockwell *et al*., 2022). To investigate possible causal relationships, we computed time-dependent Coulomb stress changes (King *et al*., 1994b; Cocco, 2002) due to variations in the lake level. Using a fully coupled model of a poroelastic crust (Rice and Cleary, 1976b; Wang, 2000; LaBonte *et al*., 2009) overlying a viscoelastic mantle (Segall, 2010; Barbot and Fialko, 2010), we find that hydrologic loads increased Coulomb stress on the SSAF by several hundred kilopascals, and fault stressing rates by more than a factor of 2, likely sufficient for earthquake triggering (King *et al*., 1994b; Cocco, 2002). The destabilizing effects of lake inundation are enhanced by a non-vertical fault dip (Fialko, 2006; Lin *et al*., 2007b; Fuis

et al., 2012; Lindsey *et al*., 2014), presence of a fault damage zone (Fialko *et al*., 2002b; Cochran *et al*., 2009), and lateral pore pressure diffusion (Caine *et al*., 1996; Bense *et al*., 2013). Our model may be applicable to other regions where hydrologic loading, either natural (Cocco, 2002; Nof *et al*., 2012) or anthropogenic (Talwani, 1997; Gupta, 1992), was associated with significant seismicity.

2.1 Main

The Southern San Andreas Fault (SSAF) is the only historically quiescent seismogenic section of the San Andreas Fault system (Fig. 1), believed to pose the largest seismic hazard in California (Weldon *et al*., 2005; Fialko, 2006; Field *et al*., 2014). The last major earthquake on the SSAF occurred around 1726 (Fumal, 2002). The modern open interval of $\sim$300 years is well in excess of the average recurrence interval of 180$\pm$40 years over the last millennium (Fumal, 2002; Philibosian *et al*., 2011; Rockwell *et al*., 2018). Previous studies suggested that at least some large events on the SSAF may have occurred during highstands of ancient Lake Cahuilla (Gurrola and Rockwell, 1996; Thomas and Rockwell, 1996; Waters, 1983). Lake Cahuilla was a prehistoric lake that formed from periodic meandering of the Colorado River north into the Salton Trough, and had an estimated maximum volume of 236 km^3, ref. (Rockwell *et al*., 2018). Filling and draining of the lake subjected the SSAF system to surface loads corresponding to $\sim$100 meter-deep body of water (Gurrola and Rockwell, 1996; Thomas and Rockwell, 1996; Waters, 1983). Prior studies suggested that loading due to ancient Lake Cahuilla inhibited failure on most of the SSAF, except maybe for a small portion within the lake boundary (Luttrell *et al*., 2007), or that triggering of large events on the SSAF was mediated by slip on nearby extensional stepover faults (Brothers *et al*., 2011). Here, we combine new paleoseismic and geologic data with state-of-the-art numerical modeling to investigate the relationship between lake loading and major earthquakes on the SSAF over the last $\sim$1100 years.

Paleoseismic evidence is critically important in determining both the timing of large earthquakes and lake history. Previous work at the Coachella paleoseismic site, just below the ancient Lake Cahuilla shoreline (Fig. 1), suggested five probable and two possible earthquakes over the last millennium (Philibosian *et al*., 2011). Recently, Rockwell et al. (Rockwell *et al*., 2022) reinterpreted the stratigraphy at the Coachella site, building upon the results developed for the past two full lake inundations, which consider both the historical record and drowned stump ages (Rockwell *et al*., 2018). Three different lake level models, varied by different sampling assumptions, were constrained by over 122 radiocarbon dates across several paleoseismic sites in the ancient Lake Cahuilla inundation zone. The resulting nearly 2000 year history of ancient Lake Cahuilla is based on an extensive compilation and review of radiocarbon samples throughout the lake basin, and on the unique character of some stratigraphic unit interpretations. Distinct organic layers, measuring 1-2 cm in thickness and originally described as "soils", were reinterpreted to be the result of lake inundations. These organic-rich layers record where organic materials floated to near the shoreline and were incorporated into a lake sequence during rising lake waters. Supporting this interpretation is the fact that organic soils are not present in the hyper-arid environment of the Lake Cahuilla basin and that such organic layers have only been identified at or below the shoreline of Lake Cahuilla (Sieh, 1986; Gurrola and Rockwell, 1996; Rockwell *et al*., 2022). Similar observations have also been made along the Laguna Salada Fault further to the south (Mueller, 1984).

Based on these insights into the paleo-lake chronology, we interpret that 6 of the past 7 SSAF earthquake horizons lie within lacustrine sediments. This is critical, as new paleoseismic data that refine the history of lake ages (Rockwell *et al*., 2022) also place tighter constraints on the timing of past earthquakes. We converted the ^{14}C radiocarbon ages of organic material to calendar years C.E. using the OxCal algorithm v. 4.4 and the history of atmospheric ^{14}C concentrations (Bronk Ramsey, 2009; Reimer *et al*., 2020). The OxCal algorithm produces probability density functions (PDFs) of each date by convolving uncertainties in the radiocarbon measurements.

An ordered sequence based on the stratigraphic layering provides additional constraints to the radiocarbon age PDFs. The statistically consistent earthquake PDFs are calculated based on the interpretation of their timings in conjunction with the full sequence. The details of our interpretation of the paleoseismic record and OxCal model are provided in the Supplementary Materials. Our interpretations of where earthquake horizons lie in the stratigraphy are mostly unchanged from Philibosian et al. (2011). The major differences in earthquake ages are a result of re-classifying which sediments are lacustrine, as well as assuming most charcoal and all organic layers are from material that grew within the lake perimeter, and therefore have ages corresponding to the preceding dry periods between lakes (Rockwell *et al*., 2022). Figure 2 shows the calculated PDFs for major seismic events that occurred on the SSAF over the last $\sim$1100 years (see Extended Data Table 1 for earthquake dates). Also shown in Fig. 2 is the preferred history model of Lake Cahuilla (Rockwell *et al*., 2022) that includes the following six highstands: 1731-1733 CE (Lake A), 1618-1636 CE (Lake B), 1486-1503 CE (Lake C), 1192-1241 CE (Lake D), 1007-1070 CE (Lake E), 930-966 CE (Lake F). Highstands are preceded by a 13 yr-long inundation period and followed by a 50 yr-long dessication period consistent with prior lake models (Rockwell *et al*., 2018). The same radiocarbon dates are used to inform the timing of the earthquake and lake ages, but the lake timings are further constrained by climate data (Rockwell *et al*., 2022). While previous studies were inconclusive about temporal correlation of the earthquakes and lake episodes over the last millennium (Philibosian *et al*., 2011), the new earthquake history strongly suggests that all lake filling events were accompanied by large earthquakes (Fig. 2).

Such a correlation between the earthquake and lake timings is unlikely coincidental. A Monte Carlo statistical test of the lake and earthquake timings confirms a non-random relationship with $>$97% confidence (Extended Data Fig. 1). Also, hydrologic loads are known to stimulate seismicity (Roeloffs, 1988a; Segall, 1989; Gupta, 1992; Gupta, 2002; Talwani, 1997). Possible mechanisms include increases in pore pressure, flexural stresses, and poroelastic effects (Luttrell *et*

al., 2007; Brothers *et al*., 2011). Seismicity may be triggered shortly after reservoir impoundment by the poroelastic load of the lake at the Earth's surface, as well as on longer time scales due to the time-dependent diffusion of pore pressure to greater depths (Roeloffs, 1988a; Rajendran and Talwani, 1992; Talwani, 1997). Increases in pore fluid pressure, in general, bring faults closer to failure, and may potentially advance the timing of large events (Simpson *et al*., 1988; Talwani, 1997; Tao *et al*., 2015). Another factor relevant to triggering of seismicity is the presence of fluid pathways due to rock damage. Faults are commonly associated with damage zones, (Fialko *et al*., 2002b; Cochran *et al*., 2009) resulting in enhanced permeability that decays with distance away from the fault core (Mitchell and Faulkner, 2009; Dor *et al*., 2006; Rockwell *et al*., 2009; Morton *et al*., 2012; Rempe *et al*., 2013). Fault damage zones can therefore substantially facilitate fluid transport to seismogenic depths (Caine *et al*., 1996; Bense *et al*., 2013). There is evidence for an extensive SSAF damage zone from field observations at surface exposures, borehole data, and Earth tidal analyses (Morrow *et al*., 2014; Xue *et al*., 2016). Models of lake loading therefore need to account for the fully coupled poroelastic response and the heterogeneous hydro-mechanical properties of the fault zone and ambient crust.

To quantify the effects of lake loading, we built a fully-coupled, three-dimensional finite element model of the Salton Trough that accounts for time-varying surface loads, a realistic fault geometry, crustal poroelasticity, and viscoelastic relaxation in the ductile substrate (see Methods and Extended Data Fig. 2). We performed numerical simulations in which we varied material properties of the host rocks and damage zone to constrain a plausible range of pore pressure and stress evolution due to ancient Lake Cahuilla (Extended Data Fig. 3 ; Figs. S1-S4). The temporal evolution of the lake load, which is ascribed as both a pore pressure and vertical stress boundary condition at the Earth's surface, is constrained by the latest geologic data (Rockwell *et al*., 2022) (Fig. 2). Unless otherwise noted, models presented below assume the lake depth of 97 m, consistent with previous studies (Luttrell *et al*., 2007), but we also considered a spatially variable surface load controlled by bathymetry. Each lake episode consists of a unique filling,

highstand, desiccation, and consolidation history. We resolve the coupled deformation transient from the refined ∼1100 year lake history with ∼200 adaptive time steps across our numerical model domain. The increase in pore pressure within the basement is highly dependent on the assumed parameters of the fault zone, in particular the damage zone permeability (Extended Data Table 2). For a given loading history, higher values of permeability give rise to greater pore pressures at depth (Extended Data Fig. 3). Additionally, similar to the analytic solution (see Methods; Extended Data Fig. 4), we observe the "memory" effect of pore pressure at depth, whereby subsequent lakes can contribute to higher pore pressure due to the diffusive time lag of a previous lake superimposing on the next (Fig. 3).

The evolution of faults toward (or away from) failure is commonly described in terms of the Coulomb failure stress (ΔCFS) (Roeloffs, 1988a; King *et al.*, 1994b; Cocco, 2002) (also, see Methods). The calculated values of ΔCFS resolved on a potential failure plane (Figs. S3 and S6) exhibit a strong sensitivity to the fault geometry, and in particular the fault dip angle. We assume an average strike of 313^o based on the fault trace (Fig. 1) (USGS and California Geological Survey, 2019) and a fault dip of 60^o NE, which is constrained by geodetic and seismic data (Fialko, 2006; Lin *et al.*, 2007b; Fuis *et al.*, 2012; Schulte-Pelkum *et al.*, 2020). ΔCFS values are inversely proportional to the fault dip angle: a steeper fault experiences a smaller change in the Coulomb stress, and a more shallowly dipping fault experiences a larger change in the Coulomb stress (Fig. S1). Another factor affecting ΔCFS is a time-dependent flexure due to the varying surface load and the associated viscoelastic relaxation in the underlying ductile substrate. For a range of possible substrate viscosities suggested by previous work (Luttrell *et al.*, 2007), the effects of viscoelastic relaxation on ΔCFS are relatively minor (Fig. S2). Our simulations indicate that variations in the pore pressure are the dominant contributor to the total ΔCFS (Figs. S3 and S4). Models that used a simplified representation of only the elastic crust suggested higher values of ΔP but lower values of the effective ΔCFS (Figs. S3 and S4).

We find that the magnitude of shear stress ($\Delta\tau_s$), effective normal stress ($\Delta\bar{\sigma}_n$), pore

pressure (ΔP), and ΔCFS increase throughout the sequence of flooding cycles (Fig. 4a). Most importantly, ΔCFS remains predominantly positive across the fault surface during each cycle (Fig. 4a and Video S1-S3). Positive ΔCFS outside the lake boundary, albeit smaller in magnitude, is a consequence of lateral pore pressure diffusion (Fig. 3 and Video S1-S3). When the lake is at highstand, shear stress is positive (i.e., encouraging failure) in parts of the fault closest to the lake center, and slightly negative (i.e., inhibiting failure) outside of the lake. In contrast, the normal stress is negative (i.e., inhibiting failure) directly below the lake due to the increased vertical compression during lake impoundment. However, the effective normal stress stays positive on much of the fault due to increases in pore pressure. The modeled compression directly below the lake simply moderates the magnitude of the Coulomb stress in that region. Models with lower permeability display a similar evolution, although the spatial extent of pore pressure diffusion is smaller.

Figure 4 depicts evolution of ΔCFS for a single point on the fault plane near the center of the lake at the representative seismogenic depth of 7km (Point 21 in Fig. S5) for a range of models (M1-M5, see Extended Data Table 2) with different permeabilities of the fault damage zone (Lin *et al*., 2007b; Brothers *et al*., 2011). In each model, as the lake fills, the ΔCFS increases rapidly due to the initial lake impoundment (i.e., undrained response). During the desiccation phase of the lakes, the overall ΔCFS decreases, but remains mostly positive except for the current dry period from $\sim$1733 to the present, and dry period between lakes C and D. The models at the high end of the assumed damage stays positive in most parts of the fault zone permeability reveal a relatively rapid diffusion of pore pressure from the surface to seismogenic depths. The increased rate of ΔCFS during the lake highstands is indicative of the higher permeability. Additionally, the results depict how the short time interval between highstands in the first 2-3 lakes contributes to an overall cumulative increase in the Coulomb stress for the second (Lake E) and sometimes the third (Lake D) lakes, depending on the modeled fault permeability and depth. The remnant diffusion of pore pressure due to the first (F) lake coincides with the impoundment of the subsequent lake,

producing a larger overall ΔCFS. This additive effect is even more pronounced for lower fault zone permeabilities. We acknowledge that shorter lake highstands would reduce the additive effect on ΔCFS discussed in this paper as we use the maximum durations for the lake highstands (Rockwell *et al*., 2022). We also performed simulations in which the surface load is controlled by a local bathymetry. The results are similar to the reference model assuming a constant depth of 97 m, except the calculated stress changes and stressing rates are reduced by $\sim$40%. This is because the average lake depth is less than the maximum depth of 97 m. Predictions of the reference model therefore should be considered an upper bound. Depending on the rate of diffusion, depth of interest, and time interval between lakes, the effect of previous lakes may persist during subsequent impoundments. This memory effect is more pronounced for Model 3 and Model 4 due to lower permeability at depth (Extended Data Fig. 5). All models, including Model 5, which does not include a fault damage zone, produce positive ΔCFS values greater than $\sim$0.1 MPa, likely large enough for earthquake triggering (King *et al*., 1994b; Goebel *et al*., 2017; Verdecchia *et al*., 2021).

In addition to the magnitude of stress changes, earthquake triggering may be affected by variations in the stressing rate (Toda *et al*., 2002; Qin *et al*., 2022). We evaluate the ratio of the calculated stressing rate from the lake load to the average tectonic loading rate through time (Fig. 4b). The latter is calculated using a geodetically constrained dislocation model assuming a secular slip rate of 18 mm/yr (Lindsey and Fialko, 2013), which is in agreement with the most recent constraints on the long-term geologic slip rate (Blanton *et al*., 2020) (and refs. therein). The estimated tectonic stressing rate at seismogenic depth is $\dot{\sigma}_{13}$ = 16 kPa/yr (see Methods). We find that the largest perturbations in stressing rates occur, as expected, during the early flooding stages of each of the past six lakes (Fig. 4b). For 6 of the past 7 earthquake events on the SSAF ($\pm 2\sigma$ uncertainties; Extended Data Table 1), the resolved stressing rate from lake loading was greater than the tectonic loading rate. The respective perturbations remain positive through the lake highstands until dessication begins, during which the stressing rates drop below the average

interseismic stressing rate. Over the past ∼1100 years, the fault stressing rate during lake loading increased by more than a factor of 2 (Fig. 4b). These high stressing rates from lake loading persisted for multiple decades during each lake cycle. The relationship between lake cycles and earthquakes is potentially known at higher precision exclusively from the stratigraphy. There is evidence to suggest that events Coa-5 and Coa8+ may have occurred during lake filling phases while Coa-1, Coa-2 ,Coa-3&4, and Coa-6 occurred during highstands (see Methods).

Loading from ancient Lake Cahuilla induced stress changes capable of triggering events on the SSAF, increased ΔCFS across most of the fault, and on average increased peak stressing rates during lake flooding well above the tectonic stressing rate (Fig. 4b; Video S1-S3). Larger fault locking depths (Lindsey *et al*., 2014), smaller earthquake nucleation depths (Jin and Fialko, 2020), and elastic moduli of the host rocks (Eissa and Kazi, 1988) would result in a lower tectonic stressing rate, and a correspondingly higher contribution of the lake loading. Our model does not account for possible memory effects from earthquakes that occurred more than ∼1000 years ago, as such effects are likely negligible if the interseismic interval exceeds several hundred years (see Fig. 4a; also ref. (Salditch *et al*., 2020)). We also do not account for stress transfer due to slip on subsidiary faults that could potentially amount to ΔCFS on the order of 1 MPa (Brothers *et al*., 2011). Our models thus provide a conservative estimate for the modulation of stress on the SSAF by the filling and desiccation of ancient Lake Cahuilla.

Our results bear on the current 300 yr-long quiescent period on the SSAF. Variations in stress and stressing rate on the SSAF due to the lake load suggest that a gradual decline in water level since ancient Lake Cahuilla's last highstand may affect the timing of the SSAF's next event. Indeed, our model predicts predominantly negative ΔCFS for sufficiently long dry periods. The negative ΔCFS during dry periods is primarily due to the flexure of the upper crust caused by a decreased lake level, and, to a smaller degree, viscoelastic relaxation below the brittle-ductile transition. A gradual decrease in the ΔCFS since the last highstand was briefly reversed in 1905-1907 due to an overflow of the Colorado River into Salton Basin, which led

to formation of the present-day Salton Sea (Brothers *et al.*, 2011). Our model predicts a sharp increase in ΔCFS, on the order of 0.1 MPa, for most of the fault during the filling phase of the Salton Sea from 1905-1907 (Figs. S11-S16). This increase in ΔCFS was accompanied by the 18 April 1906 M_I 6.1 (MMI 8) event in the Salton Trough near Brawley, CA. The latter event was previously argued to be triggered by the great 1906 California Earthquake near San Francisco that occurred 11 hours before and 700 km away from the M_I 6.1 event (Meltzner and Wald, 2003). It is possible that such distant triggering was facilitated by the lake load. After 1907, the model predicts a subsequent decrease in ΔCFS as the Salton Sea started to decline- below the 0.1 MPa triggering threshold (Figs. S5-S11; Videos S1-S3). Further desiccation of the Salton Sea in the future will continue to have a stabilizing effect on the SSAF, but potentially increase the stress to be released in a future event (or a sequence of events).

Our results demonstrate how improved paleoseismic and paleo-lake records, together with advanced models of hydrologic loading of a heterogeneous poroelastic crust, provide new insights into the relationship between water level variations of ancient Lake Cahuilla and the past 7 major earthquakes on the SSAF. We find that increases in lake level result in positive Coulomb stress changes on most of the SSAF with stressing rates as high as 2-3 times the tectonic loading rate. Positive ΔCFS values are also seen on sections of the SSAF that are outside of the lake due to lateral diffusion of pore pressure along a permeable fault zone. This indicates that our model predictions are not strongly sensitive to uncertainties in the location of nucleation sites of past (as well as future) large events on the SSAF (See Figs. S10-S16 and Videos S1-S3). The pore pressure "memory effect" amplifies the contribution of successive lakes provided that intervals between inundations do not significantly exceed the characteristic diffusion time within the seismogenic layer. Our model may also be applicable to other areas where seismogenic faults are subject to hydrologic loads from natural or anthropogenic sources (Nof *et al.*, 2012; Gupta, 2002; Talwani, 1997).

2.2 Online Methods

2.2.1 Calculation of the Coulomb stress changes

The change in ΔCFS is defined as (Roeloffs, 1988a; King *et al.*, 1994b; Cocco, 2002):

$$\Delta CFS = \Delta\tau_s + \mu(\Delta\sigma_n + \Delta P) \tag{2.1}$$

where $\Delta\tau_s$ is the change in shear stress, μ is the coefficient of friction, $\Delta\sigma_n$ is the change in normal stress (increases in compression are deemed negative) and ΔP is the change in pore fluid pressure (increases in fluid pressure are deemed positive). The sign convention indicates that a positive change in the effective stress $\bar{\sigma}_n = \sigma_n + P$, as well as a positive change in shear stress $\Delta\tau_s$, promote failure. In this paper, we assume a typical value for static friction of μ=0.6 (Byerlee, 1978b; Sibson, 1994; Fialko and Jin, 2021a; Fialko, 2021), although it may vary considerably depending on a rock type and ambient conditions (Lockner *et al.*, 2011; Mitchell *et al.*, 2013; Mitchell *et al.*, 2015). The Coulomb stress changes ΔCFS are computed by resolving the normal and absolute shear stress components using a full stress tensor, the prescribed fault geometry, and sense of slip. Calculations were performed using a Matlab toolbox Abaqus2Matlab (Papazafeiropoulos *et al.*, 2017).

Both the magnitude and sign of the estimated ΔCFS is sensitive to several parameters in the context of poroelastic deformation, including: (i) the fault zone permeability, width, and connectivity to regions where large seismic events nucleate. Higher permeability values allow for pore pressure to diffuse faster, increasing pore pressure and ΔCFS at depth (Durham, 1997; Miller, 2013). A larger fault zone width and pore connectivity may also increase pore pressure and ΔCFS on the fault at greater depths (Chang and Segall, 2016). (ii) The fault dip angle (Fialko, 2006; Fuis *et al.*, 2012), the varying geometry of the lake, and the concomitant surface load (Luttrell *et al.*, 2007). The lake load produces regions of relative compression and extension

within the seismogenic crust (Extended Data Fig. 6). Regardless of the fault attitude and sense of slip, increases in the fault-normal compressive stress are expected to inhibit failure, and decreases in compressive stress are expected to promote failure. (iii) The undrained pore pressure effect, without which the total pore pressure will be underestimated (Ge, 2011; Tao *et al*., 2015) (Extended Data Fig. 6).

2.2.2 Fully-coupled 3-D finite element models

The governing equations for a fully coupled linear poroelastic three-dimensional medium are defined as (Biot, 1941; Rice and Cleary, 1976b; Wang, 2000):

$$G\nabla^2 u_i + \frac{G}{1-2\nu}\frac{\partial^2 u_j}{\partial x_i \partial x_j} = \alpha\frac{\partial P}{\partial x_i} - F_i \tag{2.2}$$

$$\alpha\frac{\partial \varepsilon_{ii}}{\partial t} + S_\varepsilon\frac{\partial P}{\partial t} = \frac{k}{\mu}\nabla^2 P + Q \tag{2.3}$$

Where G is the shear modulus, u the displacement, ν the Poisson's ratio, α the Biot-Willis coefficient, F the body force, k the permeability, μ the fluid viscosity, S_ε the constrained specific storage, ε_{ii} the dilatancy (trace of the strain tensor), P the pore pressure, and Q the fluid source (Wang, 2000). Equations (2)-(3) are nearly identical to the classic equations for linear elasticity and diffusion of pore pressure, except for the coupling of pore pressure in the conservation of momentum equation (2) and the fluid flow coupled to strain by the requirement of fluid continuity (3). Analytic solutions to the system (2)-(3) are restricted to a few highly idealized cases. We solve the respective equations numerically using the three-dimensional finite element software Abaqus-Simulia (Dassault Systemes, 2020; LaBonte *et al*., 2009).

As a preliminary cross-check, we considered a one-dimensional analytic solution for a periodic fluctuation of a surface load in a poroelastic half-space (Roeloffs, 1988a) (see Methods). For a reasonable choice of model parameters, the analytic solution reveals a cumulative increase

in pore pressure at depth through multiple inundation events. This "memory" effect is due to the fact that the maximum pore pressure increase at depth is delayed due to diffusion and does not completely vanish before the next flooding event. The timing of the superposition depends on the depth, permeability, and wavelength of the lake load. One important caveat is that the memory effect only emerges when one considers multiple lake cycles.

To account for spatial heterogeneity and anisotropy of the hydraulic properties of the lake sediments, upper crust and the fault zone, we construct a three-dimensional hydro-mechanical model based on parameters constrained by previous studies. Tompson et al. (Tompson *et al*., 2008) developed and arranged the hydro-stratigraphy of the Salton Trough sediments into three broad classes in order of increasing depth: Brawley, Palm Springs and Imperial Formations. The composition of these formations vary from sandstones to shales, shaley sandstones, conglomeratic and arkosic sandstones (see Extended Data Table 2). The Salton Trough sediments directly overlie a basement formation and upper mantle layer. The elastic parameters of the basement formation are based on inference from seismic wave velocities, and hydrological parameters are based on crustal depth dependencies which produce material properties close to those of Westerly granite (Extended Data Table 2) (Allam and Ben-Zion, 2012; Shmonov *et al*., 2003a; Wang, 2000). A crustal thickness of 18 km is assumed based on seismic observations of the region (Richards-Dinger and Shearer, 1997). We modeled the upper mantle both as an elastic (Sup. S6,S7) and a viscoelastic layer. The viscoelastic mantle is modeled similarly to previous studies as a simple linear Maxwell material (Luttrell *et al*., 2007; Lundgren *et al*., 2009; Pearse and Fialko, 2010; Johnson, 2013). We use a non-linear geometry option in Abaqus (NLGEOM) based on the results of Hampel et al. (2019), who recommended it for models involving isostatic effects (Hampel *et al*., 2019).

The fault damage zone is modeled in accordance with field hydrogeologic estimates of SAF permeability (Xue *et al*., 2016). Xue et al. (2016) estimated a relatively high permeability (10^{-14} m^2) constrained by borehole tidal responses using earth tidal analysis compared with

inferred permeability in the region and laboratory values (Brace, 1980; Morrow *et al*., 2014; Ross *et al*., 2020). We use this estimate of permeability as a high-end value and explore a range of fault damage zone permeabilities between the high-end value and permeability of the ambient crust (i.e., no contrast in permeability between the fault zone material and the host rocks; Extended Data Table 2). Therefore, we test the sensitivity of our model results to a range of fault permeabilities from $k = 10^{-14}$ m^2 (Model 1) to values consistent with the intact crystalline basement $k = 10^{-18}$ m^2 (Model 5) (Shmonov *et al*., 2003a). The fault permeability in Model 1 is considered high when compared to other fault zones worldwide or laboratory measurements. Thus, our preferred model (Model 2) uses $k = 10^{-15}$ m^2 which is more consistent with regional measurements and laboratory data (Brace, 1980; Morrow *et al*., 2014; Ross *et al*., 2020). Observations also show that the effective shear modulus of rocks within a fault zone may be lower than that of the host rock by as much as a factor of 2 (Fialko *et al*., 2002b; Cochran *et al*., 2009; Jeppson *et al*., 2010; Lindsey *et al*., 2014). In our model the shear modulus of rocks in the fault zone is taken to be a factor of 2 smaller than the shear modulus of the host rocks (Extended Data Table 2).

The numerical domain was developed and discretized in Abaqus/CAE (Dassault Systemes, 2020). The numerical domain has horizontal dimensions of 600 km x 600 km and a depth of 50 km, with the y-axis corresponding to North for the Universal Transverse Mercator (UTM) 11 North zone (Extended Data Fig. 2). The finite element mesh consists of nearly 2 million first-order linear tetrahedral elements. Characteristic element sizes vary from 30,000 m in the far-field to less than 100 m on the fault and lake edges. The SSAF fault trace is based on the USGS Quaternary fault database (USGS and California Geological Survey, 2019). We model a 60° northeast dip for the SSAF, which is consistent with geodetic slip models, seismicity, and seismic tomography data (Lindsey and Fialko, 2013; Schulte-Pelkum *et al*., 2020). We consider a 200 meter-wide damage zone centered on the SSAF slip interface (Cochran *et al*., 2009; Jeppson *et al*., 2010; Morrow *et al*., 2014). The lake shoreline is based on coordinates used by Luttrell et al. (2007) and a similar assumption in lake depth (97.2 m) across the entire lake

footprint to maintain consistency in model comparisons (Luttrell *et al*., 2007; Rockwell *et al*., 2018). Models that considered spatially variable lake depth used the SRTM Digital Elevation Model (Farr and Kobrick, 2000) and bathymetry of the Salton Sea to assign the vertical stress and pressure boundary conditions at the nodes within the lake boundaries.

The evolution of the poroelastic effects requires a careful consideration of the model boundary and initial conditions. We assume initial conditions of equilibrium stress and pore pressure (chapter 9 in ref. Segall, 2010). These conditions imply that the model only considers the effect of the lake and does not include any loading from tectonic stress. Gravity is included, with stress and pore pressure calculated as relative change from geostatic equilibrium. This is an important consideration as neglecting gravity in the viscoelastic model may overestimate surface displacements on time scales that exceed the characteristic relaxation time (Pollitz and Sacks, 1997; Barbot and Fialko, 2010). The bottom and sides of the model are fixed in the surface normal direction, and free of shear stress (a "roller" condition). The bottom of the poroelastic domain is considered to be insulated (zero flux boundary condition), while the sides and top of the model are considered to be permeable, and subject to vanishing pore pressure ($P = 0$). The top surface of the model is stress-free outside of the lake area. The lake extent is the only boundary condition that varies through time.

2.2.3 Tectonic Loading

We approximate the secular tectonic loading using a model of a screw dislocation in an elastic half-space (Savage and Burford, 1973) (see equation 2.25 in ref. (Segall, 2010)). The respective interseismic shear strain rate is given by:

$$\dot{\varepsilon}_{13} = \frac{-\dot{s}}{4\pi}\left[\frac{x_2 - d_2}{(x_2 - d_2)^2 + x_1^2} - \frac{x_2 + d_2}{(x_2 + d_2)^2 + x_1^2}\right]$$

where $\dot{s}$ is the fault slip rate (taken to be 18 mm/yr, ref. (Lindsey and Fialko, 2013)), x_1 is the horizontal coordinate with respect to the center of the fault damage zone, x_2 is the assumed depth of nucleation of large events, and $d = 10$ km is the estimated locking depth of the SSAF (Lin *et al.*, 2007a; Lindsey and Fialko, 2013).

For typical values of the shear modulus of the upper crust $G = 28 - 32$ GPa (Lindsey and Fialko, 2013), the stressing rate predicted by equation (6) at the bottom of the seismogenic zone (x_2=6-8 km) is $\dot{\sigma}_{13} = 2G(\dot{\varepsilon}_{13}) = 25 - 51$ kPa/yr. However, the distribution of stressing rates throughout the seismogenic layer depends on details of a transition between the locked and creeping parts of the fault(e.g., Lindsey and Fialko, 2016), as well as the degree of strain localization below the brittle-ductile transition(e.g., Takeuchi and Fialko, 2012). A dislocation model over-predicts the stressing rate, especially at the bottom of the seismogenic zone, due to a strong stress singularity. Also, the depth distribution of earthquakes in California shows that most earthquakes nucleate at depth of 3-4 km (Jin and Fialko, 2020). Therefore, we use the stressing rate at the free surface ($x_2 = 0$), $\dot{\sigma}_{13} \sim 16$ kPa/yr as a conservative lower bound, for the entire seismogenic layer.

2.2.4 San Andreas Fault, Coachella Site Earthquake History

The Coachella paleoseismic site (33.72722°N, 116.16976°W) trenches provide evidence of 6 or 7 lake phases and 5 to 7 earthquakes on the southern San Andreas fault over the last millenium (Rockwell *et al.*, 2022). Four types of stratigraphy have been identified at the Coachella site: lake clay/silt, recessional sand, aeolian sand, and fluvial alluvium. The distinction between aeolian sand and recessional sand may be difficult, as the aeolian sand can be derived entirely from the recessional and shoreline sand deposits. Please refer to Philibosian et al. (2011) Figure 5

for a generalized visual of the stratigraphic section and brief descriptions of each unit (Philibosian *et al*., 2011).

An "organic-rich layer" is described near the top of unit 1L which is strikingly similar to other organic mats from near shoreline sites that are interpreted as the debris that accumulates near the shoreline from the inundation of the lake basin. Unit 1L is described as a massive silt layer of lacustrine origin. An alternative interpretation is that unit 1L represents deltaic sedimentation during flooding of the site. In either case, 1L and 1S collectively represent regional Lake A at the Coachella site (Rockwell *et al*., 2022) [see their Supplementary Material]. The most recent event occurred during the highstand of Lake Cahuilla in 1726±7 C.E. (Sieh, 1986; Rockwell *et al*., 2018).

The penultimate lake earthquake, Coa-2, clearly occurs during deposition of unit 2L, which represents the lacustrine phase of Lake B (Philibosian et al., 2011; Rockwell et al., 2022), based on extensive liquefaction (only possible with water present) and upward fault terminations (Philibosian *et al*., 2011; Rockwell *et al*., 2022). That places Coa-2 to have occurred during Lake B.

Coa-3 was interpreted to have occurred during Lake C based on weak evidence of sediment (unit 3L) filling a structural trough. Coa-4 was interpreted as occurring before Lake C but between two organic layers in unit 4S. However, the reinterpretation is that these organic layers were deposited during the rising stages of Lake C along with fine sand deposits (probably of deltaic origin) implying that Coa-4 occurred during the early phase of Lake C and that the structural trough formed by Coa-4 was filled by the same lake phase, Lake C (Rockwell *et al*., 2022). Belle Philibosian agrees with this reinterpretation (Rockwell, personal communication, 2022). This means that the evidence for Coa-3 and Coa-4 are the same and that there was only one rupture, which occurred during the filling phase of Lake C but when the water had essentially reached the highstand shoreline (as the Coachella site lies at 9 m, only 4 m below the highstand).

Coa-5 was interpreted by Philibosian et al. (2011) as having occurred between deposition

of units 5S and 4L, with the units 5S and 5L folded and capped by undeformed unit 4L (Philibosian *et al*., 2011). Rockwell et al. (2022) interpret the organic unit at the top of unit 5S to represent the inundation and accumulation of organics associated with the drowning phase of Lake D (Rockwell *et al*., 2022). Coa-5 deforms the organic layer along with the underlying lake deposits, implying that the earthquake occurred after its deposition during the rising stages of Lake D. Supporting this interpretation, the strata of units 5L and 5S are plastically deformed, which in this arid environment, almost certainly required the presence of water which will only be present during a lake stand. Hence, Coa-5 is interpreted to have occurred during Lake D, and probably during the late filling stage as the event is capped by fine-grained deposition associated with the highstand of Lake D.

Coa-6 is interpreted to have occurred during deposition of unit 5L (Lake E) based on apparent upward terminations and a possible colluvial wedge of material derived from unit 5L. This evidence was considered weak, but the presented images look convincing, so we confirm the occurrence of a rupture during the highstand of Lake E.

Event Coa-7 is interpreted to have occurred between units 6S and 5L based on upward terminations and displaced strata capped by undeformed strata. This would potentially place this earthquake to between lakes E and F as there is no evidence of the presence of water at that time (no documented liquefaction).

A seventh event, Coa8+, is interpreted to have occurred in unit 7S, which Philibosian et al. (2011) describe as "coarse sand, gravel, and rounded cobbles form cross-cutting channel deposits" which are obviously fluvial in origin (Philibosian *et al*., 2011). Unit 7L is bedded fine sand and is only exposed at the very base of the section. It is interpreted by Philibosian et al. (2011) as "likely lacustrine" but a lake affiliation is not clearly demonstrated. An alternative interpretation is that unit 7L is a deltaic section that was buried by a large fluvial flood event (not a lake-filling flood event) during rising lake level. The rationale for this is that unit 7S is several meters thick and a date a charcoal sample from unit 7S returned an age of 1185±30 rcyBP (sample Sb4m14A-c),

which is consistent with Lake F or the dry period immediately preceding Lake F and not the significantly older Lake G.

Furthermore, the unit does not display the oxidation that is characteristic for sediments of Unit G, which were exposed at the surface for nearly 1 ka. Hence, we interpret the charcoal date as indicating that L7 may actually be the lower part of Lake F (deltaic rising lake phase). If correct, an earthquake occurred during the early depositional phase of lake F, followed by a large depositional event and then more lacustrine sediments. The well-sorted sand of the base of unit 7S appears to extend down through unit 7L and may represent a feeder pipe, with the lowest part of unit 7S possibly a liquefaction deposit.

2.2.5 Periodic loading of a poroelastic half-space

As a preliminary, we consider an analytical solution for a periodic fluctuation of water level over a poroelastic half-space (Roeloffs, 1988a). For a half-space extending in the positive direction of z (vertical) axis from $z = 0$ to $z = \infty$ with a periodic water load of amplitude p_s at the surface given as both vertical stress and pore pressure boundary conditions $\sigma_{zz}(z = 0,t) = -p(z = 0,t) = p_s exp(iwt)$ leads to the following solution(Wang, 2000):

$$\bar{p}_z = \gamma p_s + (1-\gamma) p_s exp(-z\sqrt{(w/2c)}) exp(-iz\sqrt{(w/2c)}) \tag{2.4}$$

Where $\gamma = \frac{B(1+v_u)}{3(1-v_u)}$ is the loading efficiency and c is the diffusivity. If we assume incompressible grains and fluid ($\gamma = 1$) then the response is entirely the vertical stress associated with water load and constant at all depth (undrained response) (Extended Data Fig. 6). The other end-member ($\gamma = 0$) assumes that the pore pressure is uncoupled from the applied stress and is given by the pore pressure diffusion solution (drained response) (Turcotte and Schubert, 2002; Wang, 2000) (equation 4-14) (Extended Data Fig. 6). Moreover, equation (1) is linear with respect to p_s. If

$p(0,t)$ is known, the solution for $p(z,t)$ may be parameterized in the frequency domain with the Green's function that carries a pressure disturbance for frequency w at $z = 0$ to $z > 0$:

$$p(z,t) = \int G(z,w)H(w)e^{-iwt}dw, \tag{2.5}$$

where

$$G(z,w) = \gamma + (1-\gamma)exp(-z\sqrt{w/2c})exp(-iz\sqrt{w/2c}),$$

and $H(w)$ is the Fourier transform of the surface ($z = 0$) history of the pressure:

$$p(0,t) = \int H(w)e^{-iwt}dw$$

Therefore, if we know the lake loading history at the surface, $p(0,t)$, we can simply take its Fourier transform, multiply this pore pressure distribution with the Green's function (3) in the frequency domain, and take the inverse Fourier transform of the product to arrive at the solution to the pore pressure distribution for all depths and times (Extended Data Fig. 4).

Given a quasi-periodic lake loading history (Rockwell *et al*., 2022), the distribution of pore pressure reveals an interesting effect. While the shape of the pore pressure distribution beyond the Earth's surface is entirely dependent upon the poroelastic constants in Eq. (1), for reasonable values of poroelastic constants of rocks comprising the upper ctust (Extended Data Table 2), the high frequency of the surface lake levels reveal a gradual increase of pore pressure magnitude. This effect would not be present in a "single lake" simulation, as the amplitude of the maximum P would just decay exponentially. Instead, with multiple lakes, the "memory" pore pressure effect is the result of a superposition of each independent Fourier component of $p(0,t)$. For example, the diffusive (drained) component of the poroelastic response of Lake F (Fig S5) lags

its surface response at depth. This is readily observed at greater depth (9 km) where essentially only the instantaneous (undrained) response to the surface load is felt through the entire load of a single lake cycle. Consequently, the lag of the maximum (drained) response of pore pressure can coincide with the maximum (undrained) response of pore pressure for subsequent lakes thus leading to larger pore pressures compared to a single lake response. Therefore, with multiple lakes there is an inherent temporal dependence of the maximum pore pressure. A point in our finite element model (model 2) below the lake at a depth of 7.2 km that lies within our fault is more coincident to the one-dimensional model at shallower 1 km depth (Extended Data Fig. 5). This 1D solution serves to conceptualize the important delayed memory effects of variable pore pressure. While the analytical 1-D solution is a useful first-order approximation, especially for locations directly below the lake, it fails to account for other potentially important factors affecting stress on the fault such as crustal heterogeneities, viscoelastic relaxation, and lateral pore fluid diffusion.

Acknowledgements

This work was supported by the Southern California Earthquake Center (grant 21091) to M.W. and NSF (EAR-1841273), NASA (80NSSC22K0506) and USGS (G20AP00051) to Y.F. This research benefited from correspondence with R. Guyer. This project used Quaternary fault data from the USGS. We acknowledge use of the CSRC high-performance computing cluster at San Diego State University.

Chapter 2, in full, is a reprint of the material as it appears in Nature 2023.: Hill, R.G., Weingarten, M., Rockwell, T.K., Fialko, Y., 2023. Major southern San Andreas earthquakes modulated by lake-filling events. Nature 618, 761–766. The dissertation author was the primary investigator and author of this paper.

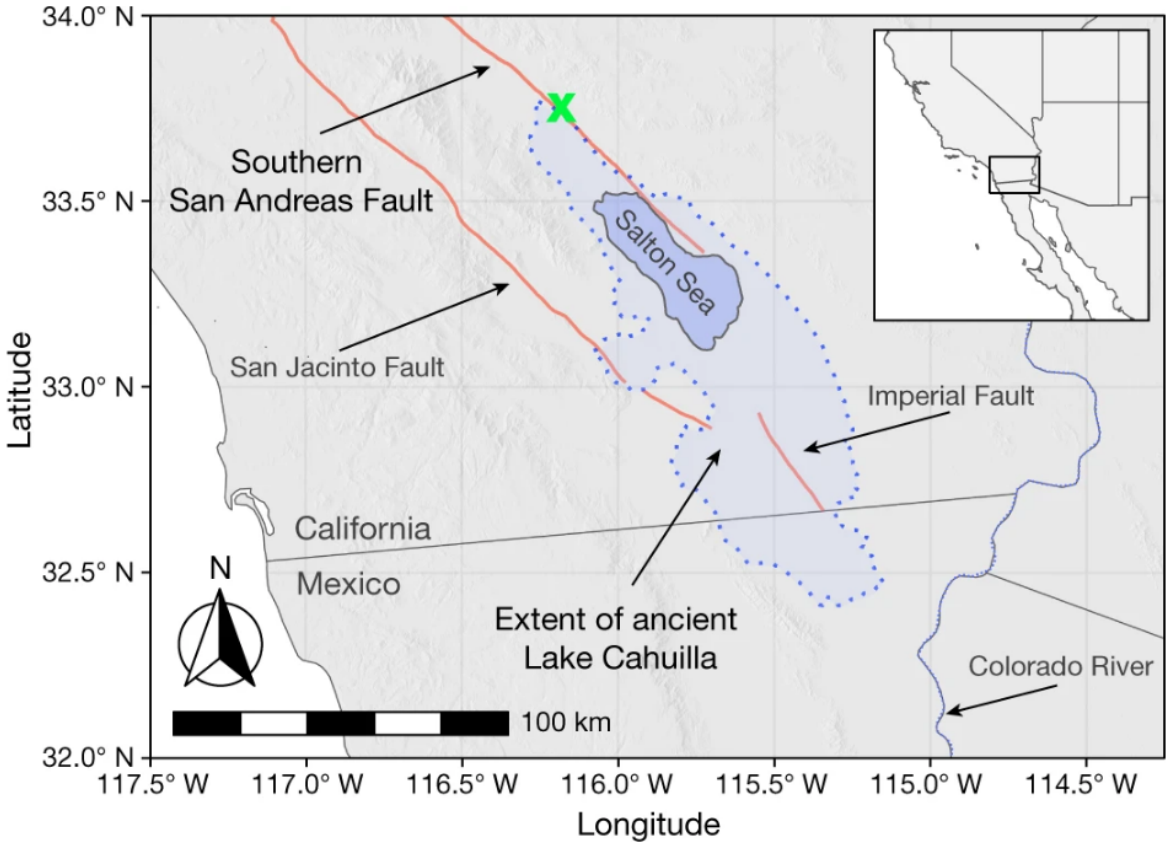
Tables and Figures

Table 2.1: Earthquake dates. Best estimates (to the nearest decade) and 95% confidence ranges of earthquake dates from the OxCal model.

Earthquake	Year	95% Confidence Range
Coa-1	1730	1721-1731
Coa-2	1610	1586-1636
Coa-3/4	1480	1459-1503
Coa-5	1200	1165-1244
Coa-6	1020	995-1040
Coa-7	970	923-1022
Coa-8	930	905-961

Table 2.2: Material properties. Best estimates (to the nearest decade) and 95% confidence ranges of earthquake dates from the OxCal model. *The mantle is modeled as a simple linear viscoelastic material with $3{\cdot}10^{19}$ Pa·s viscosityLuttrell *et al.*, 2007. Bulk modulus of permeating fluid (water) is assumed to equal $2.2{\cdot}10^{9}$ Pa. Several simulations were performed to evaluate the model sensitivity to the assumed parameters. Models 1-4 considered the effect of fault zone permeability for the following assumed permeability values: [10^{-14},10^{-15},10^{-16},10^{-17}] m^2, respectively. Model 5 corresponds to a case of no damage zone (permeability and Young's modulus are the same as those of the host rock). Model 2a uses the same material properties as Model 2, but assumes a spatially variable surface load controlled by a local bathymetry.

Hydrostratigraphy Unit	Post Brawley / Brawley	Palm Springs / Borrego	Imperial	Crystalline Basement	Mantle*	Fault Zone
Depth Range [km]	0-0.5	0.5-3	3-7	7-18	18-50	0-18
Poisson's Ratio	0.25	0.25	0.25	0.25	0.36	0.25
Young's Modulus [GPa]	50	50	50	58.4	140	1/2 host rock
Permeability [m^2]	$2{\cdot}10^{-13}$	$1.2{\cdot}10^{-12}$	$9{\cdot}10^{-14}$	10^{-18}	/	10^{-14},10^{-15},10^{-16},10^{-17}
Bulk Modulus of Solid Grains [GPa]	33	33	33	39	/	host rock
Reference	Tompson *et al.*, 2008	Tompson *et al.*, 2008	Tompson *et al.*, 2008	Shmonov *et al.*, 2003a	Turcotte and Schubert, 2002	Fialko *et al.*, 2002b; Lindsey and Fialko, 2013



.

Figure 2.1: Map of the Salton Trough and present-day Salton Sea (9.5 m average water depth). Blue dotted line denotes the historical extent of ancient Lake Cahuilla (13 m above sea level; 97.2 m maximum water depth) (Luttrell *et al*., 2007). Red lines denote traces of several major faults (USGS and California Geological Survey, 2019). Green 'X' symbol marks the location of the Coachella paleoseismic site (Philibosian *et al*., 2011)

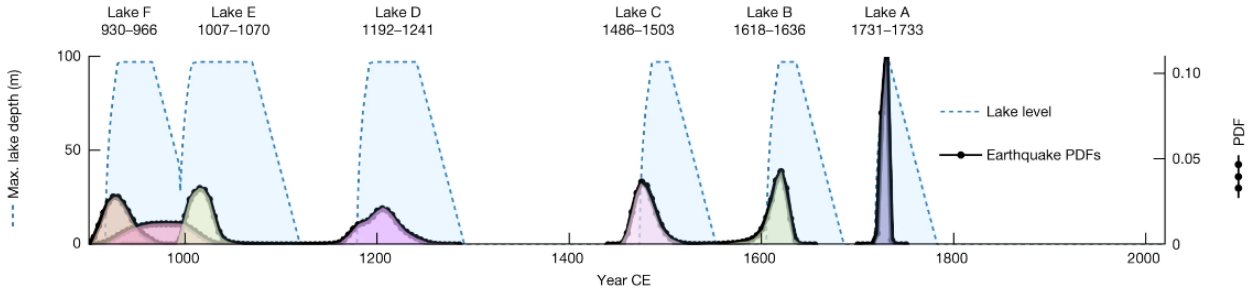


Figure 2.2: Earthquake PDFs (black dotted lines, right axis) calculated in this study (see Methods) and the relative lake loading history (blue dashed lines, left axis) from ref. (Rockwell *et al*., 2022). PDFs of pre-historic earthquakes (see Extended Data Table 1) are color-coded as follows: Coa-1 is blue, Coa-2 is green, Coa-3&4 is pink, Coa-5 is magenta, Coa-6 is yellow, Coa-7 is red, and Coa-8+ is orange.

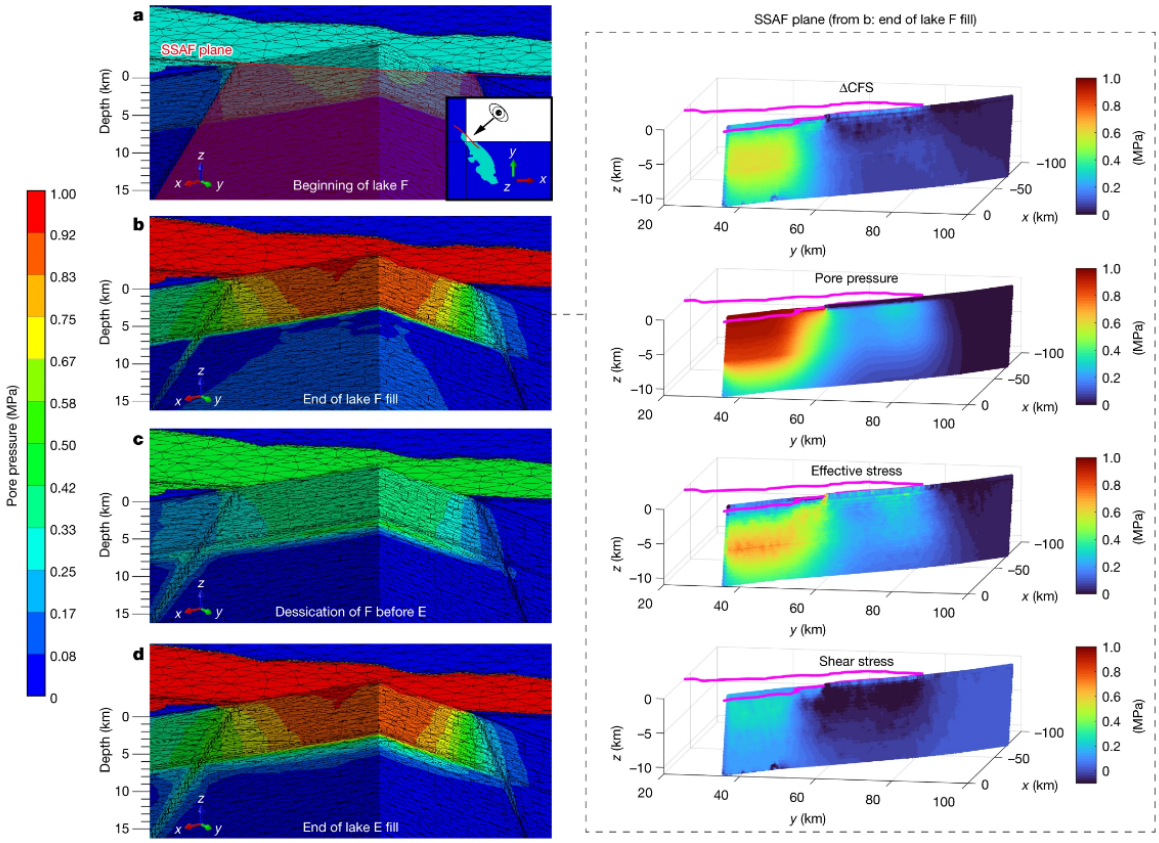


Figure 2.3: *Left*: Spatiotemporal evolution of pore pressure for a single lake cycle of ancient Lake Cahuilla. Calculations assume a constant lake depth of 97.2 m. The cross section cuts through the northwestern end of the lake and through the embedded dipping fault. A) Pore pressure immediately increases as the lake begins to inundate. B) At the time of lake highstand the pore pressure has diffused down much of the higher permeable fault but is more slowly varying elsewhere due to the surrounding low permeability basement. C) As the lake desiccates higher pore pressures remain in and around the fault. D) Subsequent lakes further increase the previously elevated pore pressure at basement depth. *Right*: The Coulomb stress change (ΔCFS), pore pressure (P), effective stress ($\bar{\sigma}_n$), and absolute shear stress (τ_s) perturbations on the SSAF fault plane for a single time frame in the lake cycle. The magenta line represents the outline of ancient Lake Cahuilla at the surface.

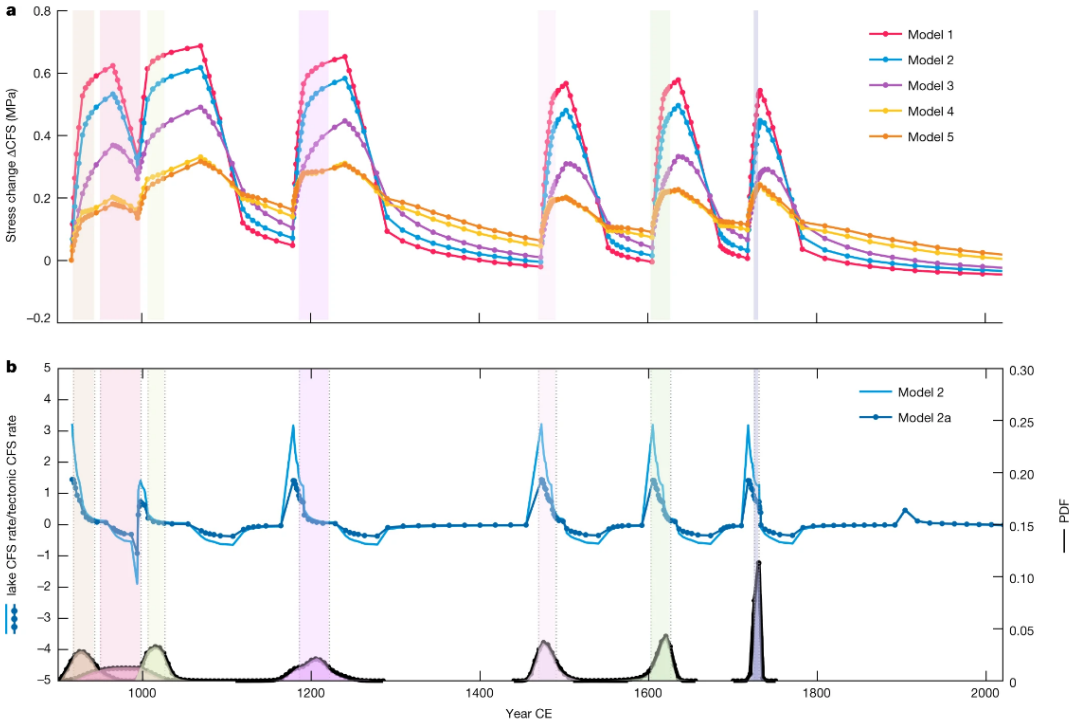


Figure 2.4: *A)* ΔCFS (MPa) on the SSAF as a function of time C.E. (years) at 7 km depth for location 21- a point on the fault near the center of the lake (See Fig. S5). Color lines correspond to models assuming different permeability of the fault zone (see Methods, Extended Data Table 2), from highest (Model 1) to lowest (Model 5, no permeability contrast with the host rocks). Same plot for a point farther from the lake center also shows predominantly positive ΔCFS, albeit of lower magnitude (Extended Data Fig. 5). *B)* The Coulomb stressing rate due to the lake loading normalized by the tectonic loading rate (16 kPa/yr, see Methods). Light blue line represents predictions of Model 2 (same as in panel A), and dark blue dotted line represents predictions of Model 2a that considers a spatially variable surface load based on a local bathymetry. The small peak in 1905 is due to the inundation of the Salton Sea (see Figs. S5-S11). Models 1-5 do not include Salton Sea inundation. Vertical color bars denote ± 2 standard deviations of the estimated earthquake dates.

Appendix

2.A Appendix to Chapter 2

2.A.1 Supplementary figures for Chapter 2

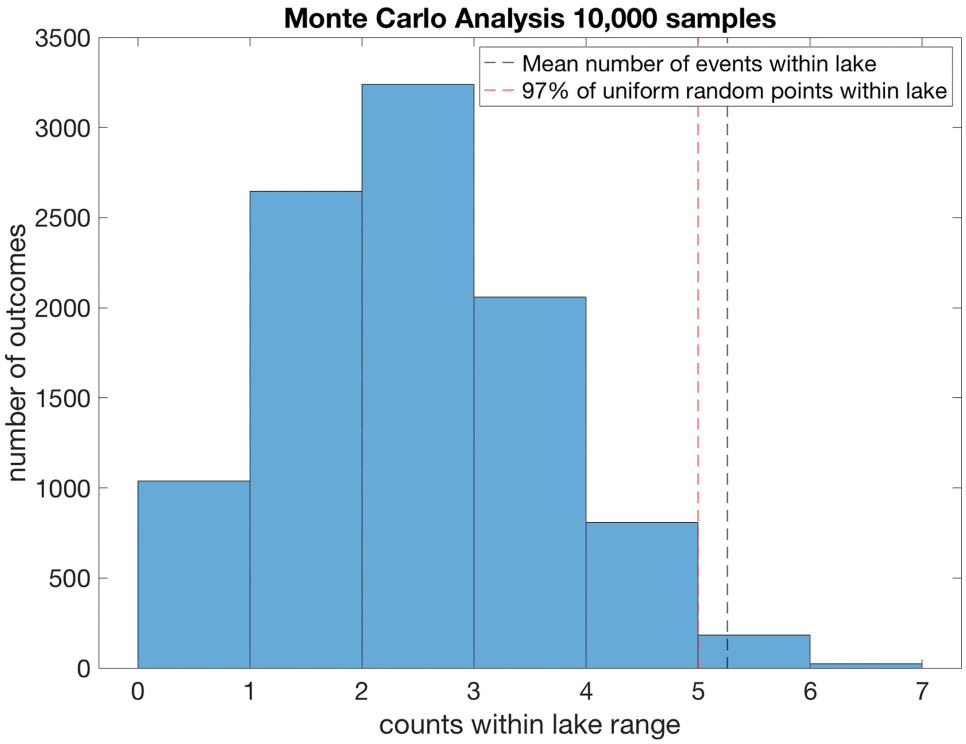


Figure 2.A.1: Extended Data Figure 1: Results of Monte Carlo statistical testing (10,000 samples) based on sampling earthquake PDF distributions and lake timings. After sampling the earthquake PDFs we determine how many fall inside the lake timings when the lake was greater than 70% full. We compare these timings to a uniform random distribution of 7 times across the same lake loading time range. We find that the mean timings that occur within lakes is >97% of the earthquake timings of a uniform random distribution that occur within lakes.

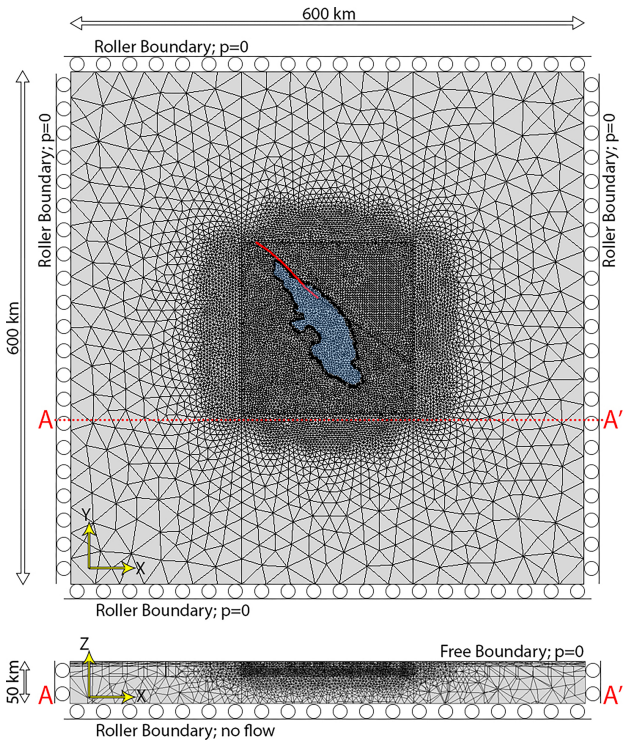


Figure 2.A.2: Extended Data Figure 2: 3D FEM model domain. The model mesh contains $\sim$2 million tetrahedron elements. The light blue color represents the extent of Ancient Lake Cahuilla. The prescribed vertical load is hydrostatic, to the lake maximum water head (97.2 m). The solid red line is the SSAF fault trace. The fault zone is modeled as a slab dipping to the north-east at 60^o (Lindsey and Fialko, 2013), with the assumed thickness of 200 m (Cochran *et al.*, 2009; Jeppson *et al.*, 2010; Morrow *et al.*, 2014).

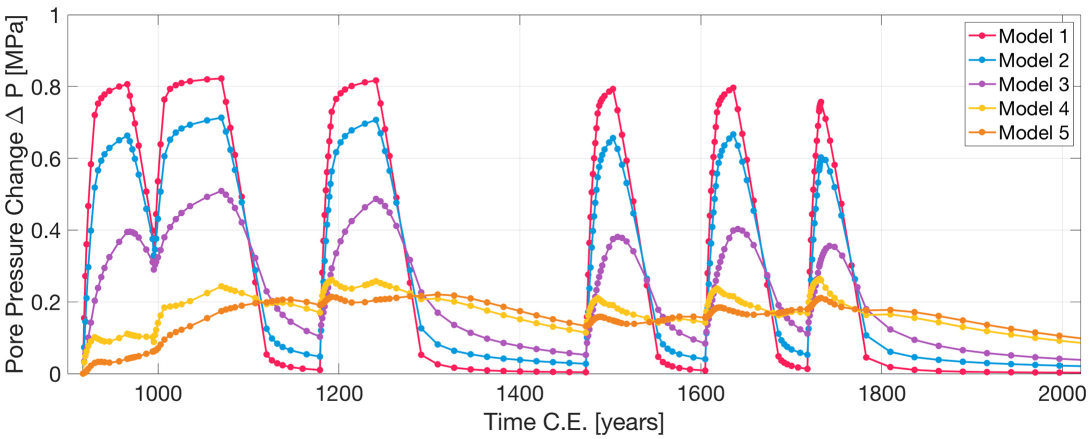


Figure 2.A.3: Extended Data Figure 3: Pore Pressure (MPa) on the SSAF as a function of time C.E. (years) at 7 km depth for location 21- a point on the fault near the center of the lake (See Fig. S5). Each model is based on the variable fault permeability with Model 1 as the highest permeable and Model 5 as no damage zone (Extended Data Table 2).

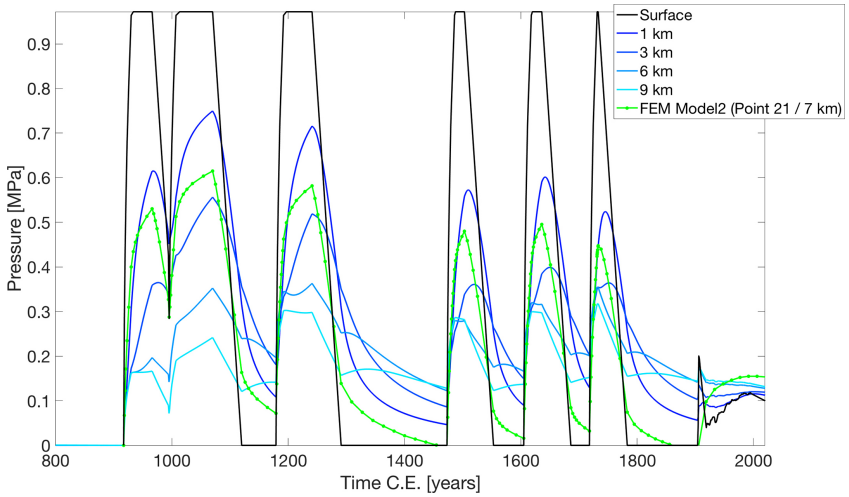


Figure 2.A.4: Extended Data Figure 4: 1D analytical model of pore pressure for a variety of different depths (blue) with surface lake level pore pressure (black). The smaller surface profile from 1905 to the present is the Salton Sea (Tostrud, 1997). FEM Model2 at 7.2 km depth (green line) shows the effect of 3D diffusion with a high-permeability fault damage zone embedded in a lower permeability host rock. The FEM model at 7.2 km resembles pore pressure in the 1D analytical case at 1 km, demonstrating how a fault damage zone can transmit pore pressure to depth effectively. ($\gamma = 0.1685$; k_{fault} = 1e-15 [m^2]; $k_{host/1Dmodel}$ = 1e-18 [m^2]).

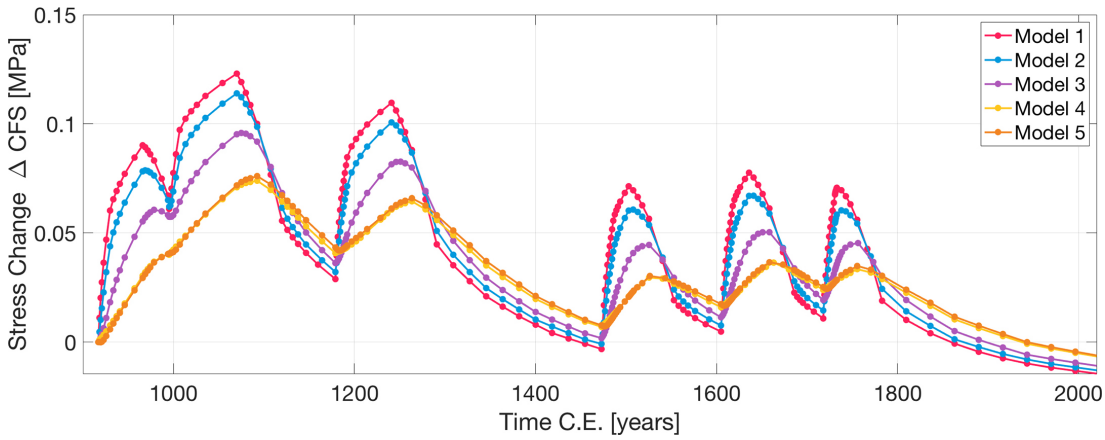


Figure 2.A.5: Extended Data Figure 5: Similar to Figure 4A from main text, but for a point farther away from the lake center (Point 24 in Fig. S5).

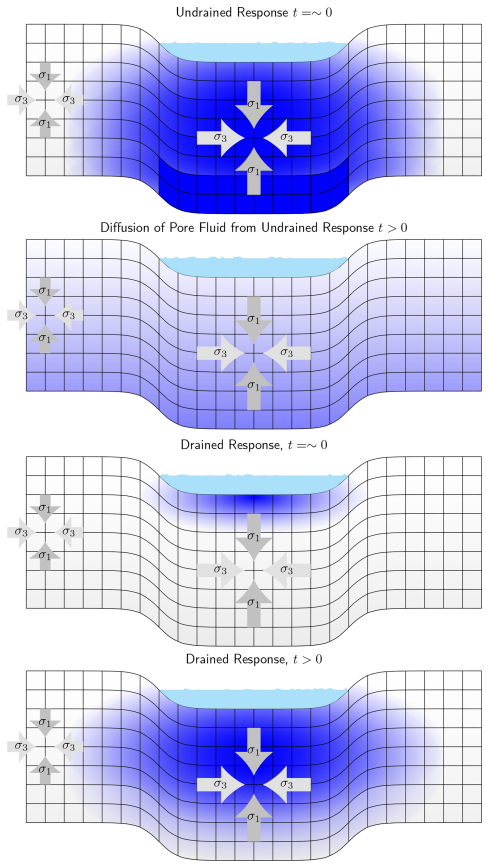


Figure 2.A.6: Extended Data Figure 6: The instantaneous and transient effects of the undrained and drained effect. At $t = 0$ the undrained effect is felt nearly instantaneous throughout the poroelastic medium beneath the lake. As time progresses this effect attempts to equillibriate at depth. At $t = 0$ the drained effect is not felt except for thesruface poroelastic medium and the bottom of the lake. As time progresses this effect increases pore pressure as diffusion drives fluid from the surface down. Furthermore, as the lake load is applied areas of compression form immediately beneath lake while areas of extension are formed near the edges.

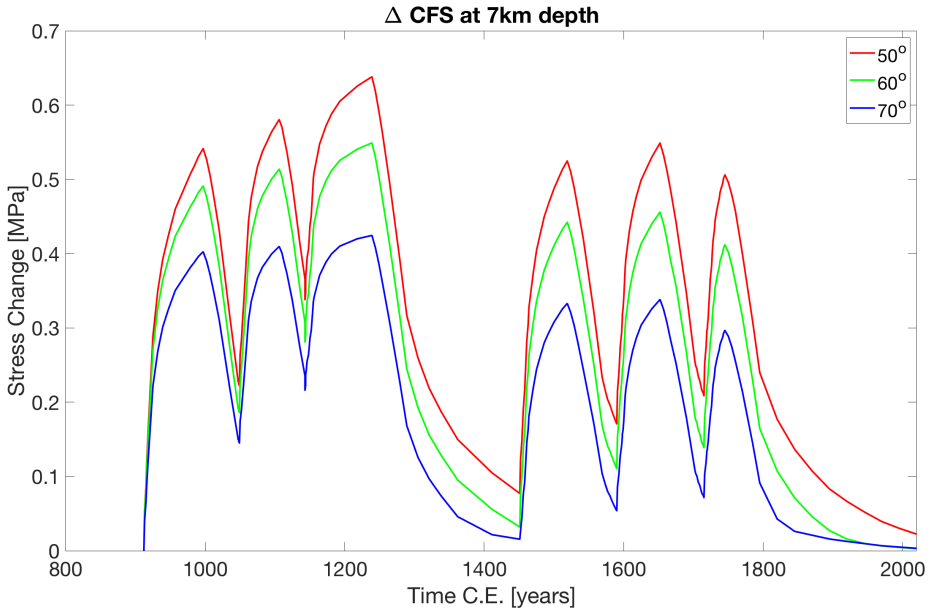


Figure 2.A.7: Figure S1: ΔCFS is dependent on different Fault dip angles. Based on the location of our lake geometry a steeper dip angle yields lower overall ΔCFS while a shallower dip angle yields higher overall ΔCFS. A steeper dip accommodates more compressional stress and is therefore lower while a shallower fault accommodates more extensional stress. Plot presented is based on a preliminary lake model that contains minor differences in ages.

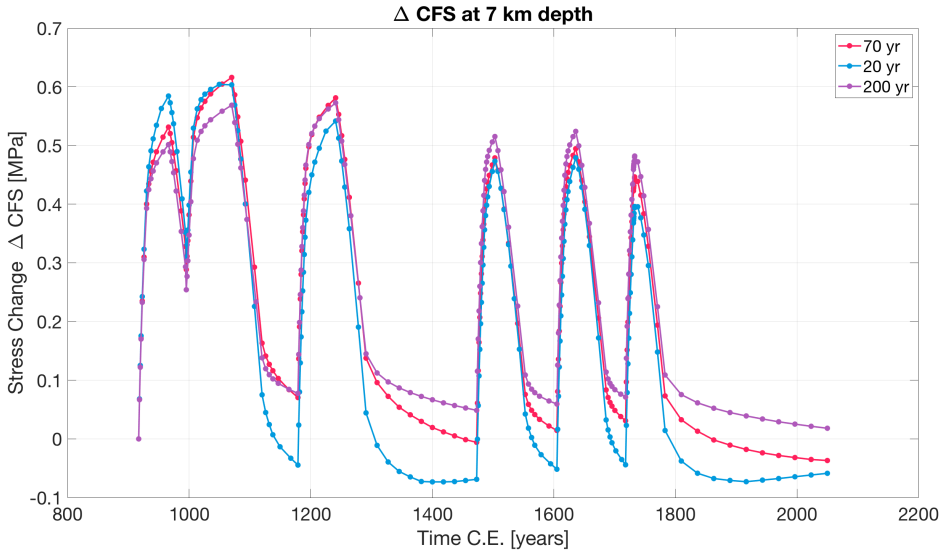


Figure 2.A.8: Figure S2: ΔCFS at 7 km depth for location 21- a point on the fault near the center of the lake (See Figure S10)ΔCFS for the fastest relaxation time endmember (20 years), the slowest (200 years), and preffered model viscosity (70 years) (Model 2) (Luttrell *et al.*, 2007). Changing the viscosity of the mantle results in minor changes to the maximum ΔCFS.

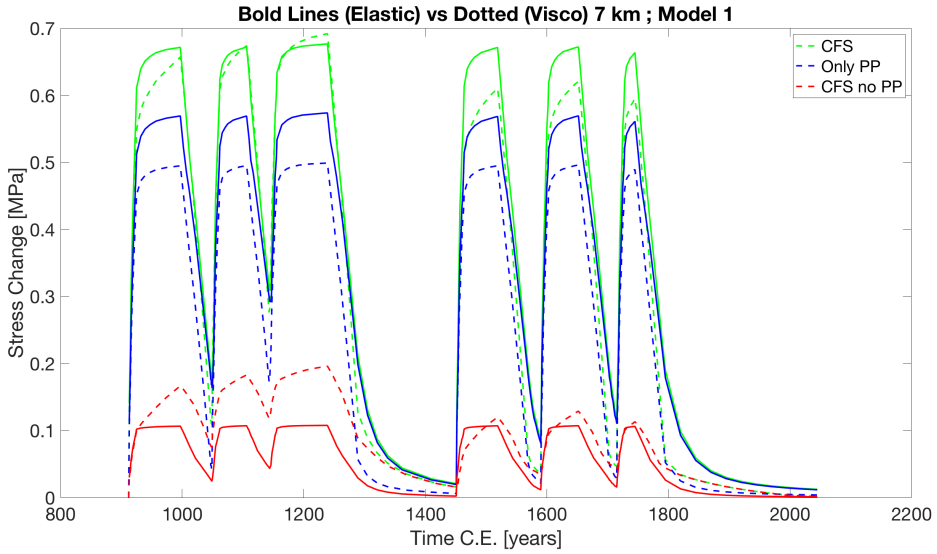


Figure 2.A.9: Figure S3: Model 1, 7 km depth for viscoelastic mantle and elastic mantle. Pore pressure is a dominant signal of the total ΔCFS. The elastic mantle model resolves larger pore pressure values, but lower $\Delta CFS - \Delta P$ values. Plot presented is based on a preliminary lake model that contains minor differences in ages.

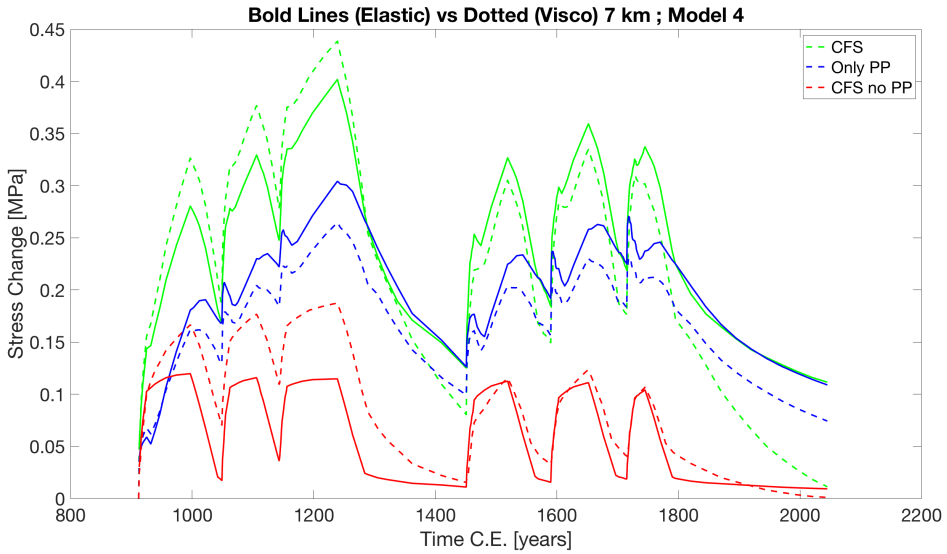


Figure 2.A.10: Figure S4: Model 4, 7 km depth for viscoelastic mantle and elastic mantle. Pore pressure is still the dominant signal of the total ΔCFS despite having a significantly less permeable damage zone. The elastic mantle model resolves larger pore pressure values, but lower $\Delta CFS - \Delta P$ values. Plot presented is based on a preliminary lake model that contains minor differences in ages.

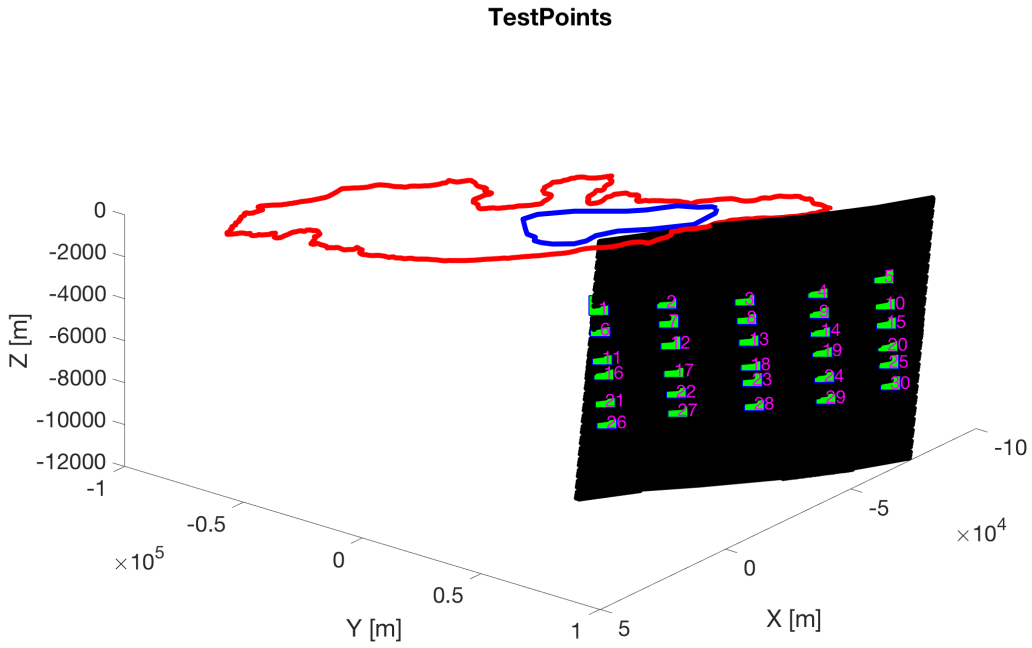


Figure 2.A.11: Figure S5: Test points along the SSAF equivalent to 6 rows at (3,4,5,6,7,8 km depth) and 5 columns as shown as green squares on a reduced cloud of black points of the fault. Red line is the perimeter of Ancient Lake Cahuilla and the blue line is the perimeter of the Salton Sea.

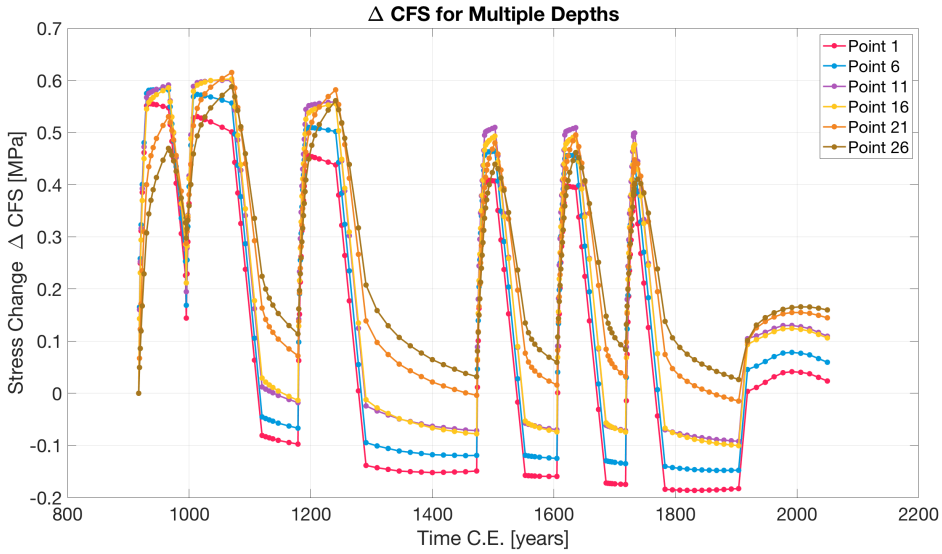


Figure 2.A.12: Figure S6: 6 row points for (3,4,5,6,7,8 km depth) associated with column 1 from S5 (Model 2).

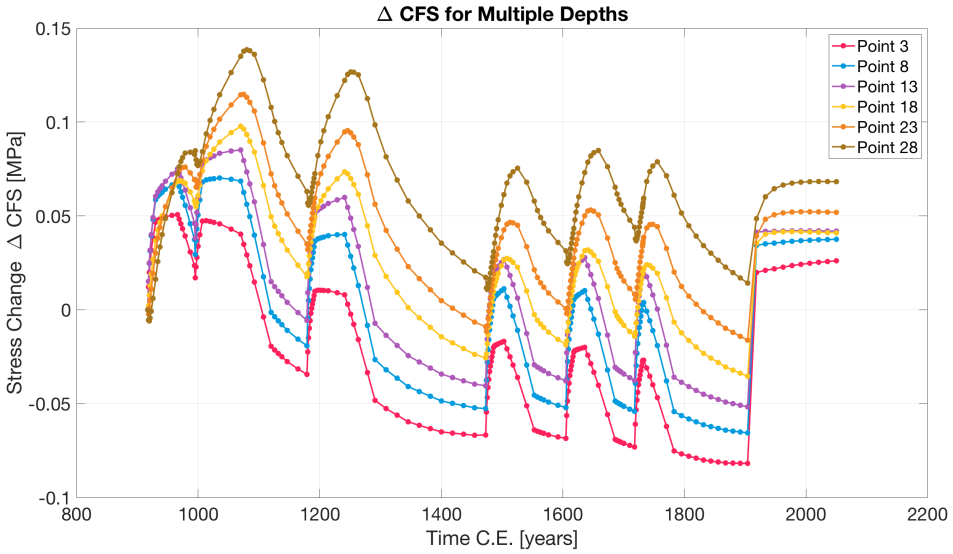


Figure 2.A.13: Figure S7: 6 row points for (3,4,5,6,7,8 km depth) associated with column 3 from S5 (Model 2).

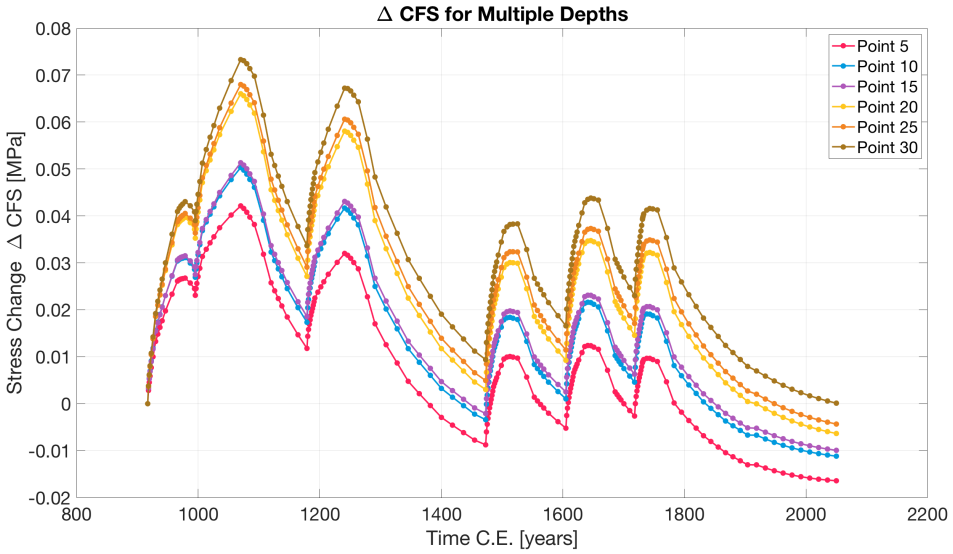


Figure 2.A.14: Figure S8: 6 row points for (3,4,5,6,7,8 km depth) associated with column 5 from S5 (Model 2).

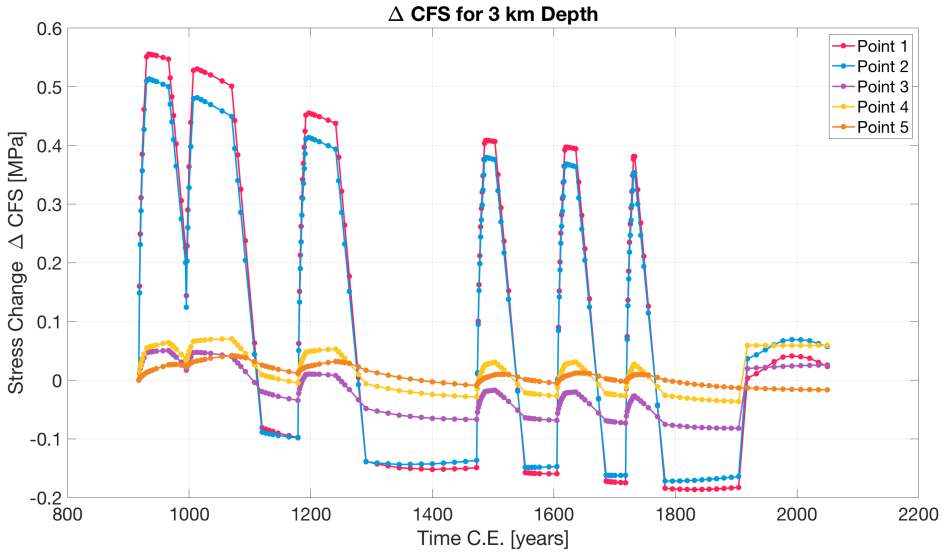


Figure 2.A.15: Figure S9: 5 column points associated with row 1 (3 km depth) from S5 (Model 2). Notice Point 3 has lower ΔCFS than points 4 and 5 which are farther away from the center of the lake which may seem counter intuitive. The reason for a lower magnitude is due to the lake geometry. The lake bottlenecks above point 3 and then increases in size allowing for more pore pressure diffusion near the end, then in the middle. Watch Video S1 to see the pore pressure increasing in this somewhat bimodal distribution.

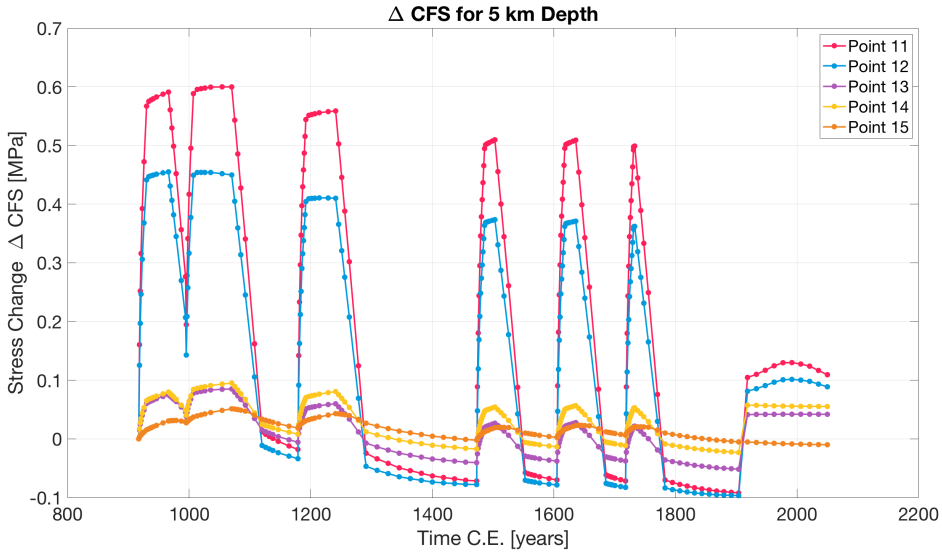


Figure 2.A.16: Figure S10: 5 column points associated with row 3 (5 km depth) from S5 (Model 2).

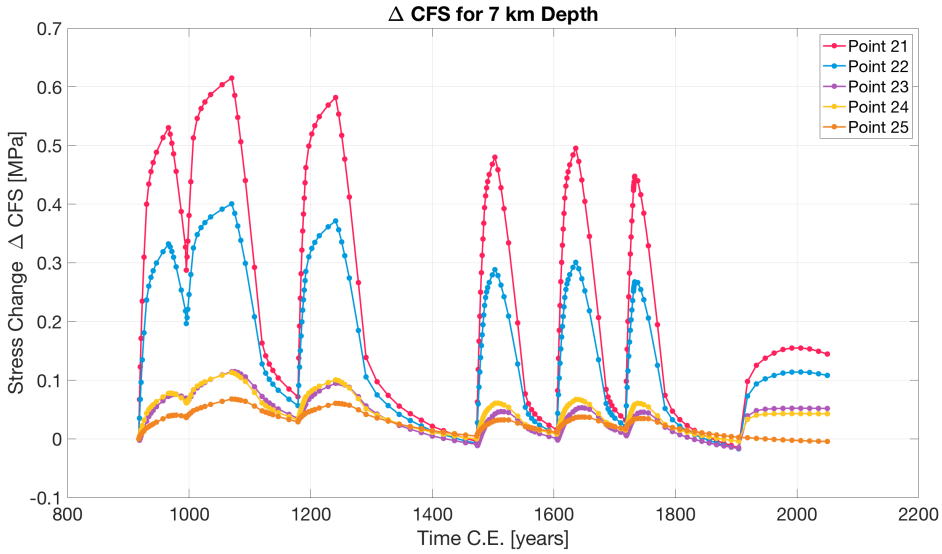


Figure 2.A.17: Figure S11: 5 column points associated with row 4 (7 km depth) from S5 (Model 2).

Chapter 3

Mitigation and optimization of induced seismicity using physics-based forecasting

It is well recognized that underground fluid injection can induce seismicity by altering stresses on pre-existing faults. Here, we investigate minimizing induced seismic hazard by optimizing injection operations in a physics-based forecasting framework. We built a 3D finite element model of the poroelastic crust for the Raton Basin, Central US, and use it to estimate time dependent Coulomb stress changes due to $\sim$25 years of wastewater injection in the region. Our finite element model is complemented by a statistical analysis of the seismogenic index (SI), a proxy for critically stressed faults affected by variations in the pore pressure. Forecasts of seismicity rate from our hybrid physics-based statistical model suggest that induced seismicity in the Raton Basin, from 2001 - 2022, is still driven by wastewater injection. Our model suggests that pore pressure diffusion is the dominant cause of Coulomb stress changes at seismogenic depth, with poroelastic stress changes contributing about 5% to the driving force. Linear programming optimization for the Raton Basin reveals that it is feasible to reduce seismic hazard for a given amount of injected fluid (safety objective) or maximize fluid injection for a prescribed seismic hazard (economic objective). A common theme across the scenarios presented shows the

optimization tends to spread out high-rate injectors and shift them to regions of lower SI. This intuitive result has practical importance: managing injection rate per unit field area may be a useful tool to reduce induced seismic hazard. Our optimization framework is both flexible and adaptable to mitigate induced seismic hazard in other regions and for other types of subsurface fluid injection.

3.1 Key Points

1. Hybrid fully-coupled poroelastic physical and statistical model suggests induced seismicity in Raton Basin is still primarily driven by wastewater injection.

2. Linear-programming optimization can reduce seismic hazard for a given amount of injected fluid (safety objective) or maximize fluid injection for a prescribed seismic hazard (economic objective).

3. Optimization tends to spread out higher rate injection wells, thus managing injection rate per unit field area may be a useful tool to reduce basin-scale induced seismic hazard.

3.2 Plain Language Summary

The Raton Basin, in the central United States, has had a remarkable increase in seismicity coincident with large wastewater injection since 2001. This seismicity primarily occurs at depths greater than several kilometers where preexisting faults in the crystalline basement are reactivated by fluid percolation. The spatial extent and rate of the induced earthquakes can inform hazard maps which display the probability of an earthquake occurrence within a specific time period. We use the physics-based and statistical models to develop an optimization framework that may help inform well operations. The proposed method allows for the maximization of injected fluid

(the economic objective) and the reduction of seismic hazard (the safety objective).

3.3 Introduction

Induced seismicity is a growing problem world-wide as it accompanies a variety of industrial activities, including hydraulic fracturing (Rutqvist *et al*., 2015; Bao and Eaton, 2016) and wastewater disposal (Ellsworth, 2013; Keranen *et al*., 2014; Shirzaei *et al*., 2016), extraction and storage of natural gas (Grasso and Wittlinger, 1990; Thienen-Visser and Breunese, 2015; Zbinden *et al*., 2017), CO_2 sequestration (Goertz-Allmann *et al*., 2014; White and Foxall, 2016), and renewable geothermal energy exploitation (Fialko and Simons, 2000; Giardini, 2009; Majer and Peterson, 2007; Mignan *et al*., 2015). Within the last decade, a dramatic increase in seismic activity in the Central and Eastern United States (CEUS) was caused by deep injection of water that was co-produced with oil (Keranen *et al*., 2014; Walsh III and Zoback, 2015; Langenbruch and Zoback, 2016; Langenbruch *et al*., 2018). Several moderate (M5+) events were induced in historically aseismic regions (Ellsworth, 2013; Weingarten *et al*., 2015; Foulger *et al*., 2018). Like natural tectonic earthquakes, induced events occur on pre-existing critically stressed faults, primarily in the crystalline basement (Townend and Zoback, 2000).

The occurrence of induced seismicity is attributed to various physical mechanisms, including pore pressure diffusion, poroelastic coupling and stress changes caused by seismic or aseismic fault slip (Segall and Lu, 2015; Keranen and Weingarten, 2018; Ge and Saar, 2022). In general, all mechanisms may contribute to the triggering of seismicity, because induced earthquakes can be triggered by stress changes just above stress perturbations caused by the Earth's tides (1-10 kPa) (Bachmann *et al*., 2012; Cacace *et al*., 2021; Wang *et al*., 2022; Stokes *et al*., 2023). Modelling studies at well-characterized injection locations show that the relative significance of these mechanisms varies from site to site depending on the physical rock properties, reservoir structure, fault geometry, seismotectonic conditions, and distance from injection among others.

Pore pressure diffusion and poroelastic stress changes are considered primary mechanisms for induced seismicity (Segall and Lu, 2015; Keranen and Weingarten, 2018; Zhai *et al*., 2019; Ge and Saar, 2022; Stokes *et al*., 2023).

Understanding and mitigating the seismic response to fluid injection is still a major challenge, not just for wastewater disposal, but for other types of subsurface fluid injection: CO2 sequestration, enhanced geothermal systems and hydraulic fracturing. In each region where subsurface fluid injection occurs, it is paramount to future operations to find an optimal balance of efficient yet safe injection practices. The field of hydrogeology has long used coupled groundwater simulations and management models to optimize pressure changes in multiple wells for a certain benefit (Gorelick, 1983; Gorelick and Zheng, 2015). For example, (Gorelick and Remson, 1982) sought the optimal solution that maximized pollutant disposal while meeting spatial water quality standards at the wells over time. A similar approach in the case of wastewater injection and induced seismicity could be to maximize injection while meeting spatial fault reactivation constraints.

Here, we present a framework that seeks to optimize the amount of wastewater injected at the basin-scale with a fully-coupled poroelastic model combined with a statistical seismicity forecasting model. Optimization is performed under a spatially varying Coulomb failure stressing rate constraint dependent on faulting orientation (King *et al*., 1994a; Cocco, 2002; Jin *et al*., 2022). We first demonstrate the hybrid model's effectiveness at forecasting the observed seismicity in the Raton Basin of Colorado and New Mexico – a long-standing and well-documented case of induced seismicity. We then demonstrate the feasibility of future induced seismicity management using optimization of injection under various constraint scenarios.

For our simulation and management models, we take advantage of the linearity in the fully coupled poroelastic equations as well as the linearity in the Coulomb stress equation. Coupled poroelastic calculations are performed using a 3D finite element hydromechanical model (Dassault Systemes, 2020). Our statistical seismicity model follows the methodology of prior

work performed in Oklahoma and Kansas, where spatiotemporal variations of induced seismic hazard are calculated from pore pressure changes and spatial variations of the subsurface's susceptibility to induced earthquakes (Langenbruch *et al.*, 2018). The susceptibility is described by the spatially varying seismogenic index (SI), a proxy for the number and stress state of pre-existing basement faults affected by stress changes (Langenbruch and Zoback, 2016; Shapiro *et al.*, 2010). Note that the SI model applied in Oklahoma and Kansas only considered pore pressure changes, while we consider the fully coupled problem by including poroelastic stress changes in the Coulomb stress analysis. We then form a management model using a response matrix for rate dependent model constraints provided by the SI.

The management models considered are three 5-year prospective scenarios that use the remnant pore pressure and stress conditions from prior injection in the Raton Basin. In each scenario, the optimization chooses which injection wells to operate and at which monthly rate of injection. The first scenario optimizes induced seismic hazard for an injection strategy that tapers the overall injection by 70% from the 2022 levels (reduction objective). The second scenario minimizes the seismic hazard for the current Raton Basin injection rate, thus optimizing seismic hazard for a given injected volume (safety objective). The third scenario maximizes the total injected volume while holding constant Raton Basin's currently forecasted seismic hazard (economic objective). The total framework serves as a flexible platform by which the optimization of injection activities are drafted to reduce the seismic hazard and maximize an economic objective.

3.3.1 Raton Basin

The Raton Basin, a $\sim$150 km long by $\sim$75 km wide sedimentary basin situated along the border between Colorado and New Mexico, has shown a remarkable seismic rate increase coincident with the beginning of industrial-scale wastewater injection in 2001 (Rubinstein *et al.*,

2014) (Figure 3.1). The rate increase was punctuated by the August 23rd, 2011 M5.3 Trinidad, Colorado earthquake, which caused structural damage in the nearby town of Trinidad, as well as 17 M4+ events, the most recent of which occurred on March 10th, 2023 [ANSS Comprehensive Catalog] (Figure 3.1). Previous studies have linked seismicity and wastewater injection wells operating in the basin using observational evidence and physical modeling (Rubinstein *et al.*, 2014; Barnhart *et al.*, 2014; Nakai *et al.*, 2017b). The time-dependent seismic hazard associated with these induced events can change based on the pumping rates associated with the injection wells. Understanding both the spatial and temporal change of past seismic hazard is critical to mitigating future hazard.

Injection induced seismicity began in 2001 and peaked in late 2011 with the August 23rd, 2011 M5.3 Trinidad, Colorado earthquake (Figure 3.2). Since 2011, regional injection rates have declined more than $\sim$33%, but the basin continues to exhibit an elevated seismicity rate with several recent M4+ events (Glasgow *et al.*, 2021a). The regional stress field is heterogeneous, with a substantial rotation of the maximum horizontal stress from predominantly north-south to east-west directions (Snee and Zoback, 2022). The earthquake focal mechanisms indicate a mixture of normal and strike-slip earthquakes (Wang *et al.*, 2020; Glasgow *et al.*, 2021a).

Geologic and hydrogeologic data indicate that the injection reservoir, the Dakota-Purgatoire Formation, a fractured sandstone reservoir, and underlying sedimentary units are permeable and hydraulically connected over a large lateral extent of the basin (Geldon, 1989; Nelson *et al.*, 2013). The injection reservoir is also well-confined from the shallower stratigraphy within the basin by more than 700 m of poorly-permeable Pierre Shale. Additionally, the western boundary is characterized by the Sangre de Cristo Mountain thrust fault system, a complex of west-dipping, Laramide-age thrust faults that show dip-slip offsets of 0.6 to 3 km (Clark *et al.*, 1966). The observed seismicity in the Raton Basin is primarily found within the crystalline basement at average depths of 5 - 7 km below surface (Nakai *et al.*, 2017b; Glasgow *et al.*, 2021a). There is also strong evidence to suggest three prominent zones of seismicity: Tercio, Vermejo Park, and

Trinidad (Figure 3.1) (Macartney and O'Farrell, 2010; Higley, 2007; Barnhart *et al*., 2014).

3.4 Physics-based Forecasting Model

3.4.1 Methods

Linear poroelasticity

To understand how injection across the Raton Basin is changing stress on pre-existing basement faults, we develop a fully coupled poroelastic model and compute the Coulomb stress changes at depth. Linear poroelasticity is essential to understanding the time-dependent coupling between the deformation of, and fluid flow in, hydrogeologic units within the Earth. The governing equations for a fully coupled linear poroelastic three-dimensional medium are defined as (Biot, 1941; Rice and Cleary, 1976b; Wang, 2000):

$$G\nabla^2 u_i + \frac{G}{1-2\nu}\frac{\partial^2 u_k}{\partial x_i \partial x_k} = \alpha\frac{\partial p}{\partial x_i} - F_i, \tag{3.1}$$

$$\alpha\frac{\partial \varepsilon_{kk}}{\partial t} + S_\varepsilon\frac{\partial p}{\partial t} = \frac{k}{\mu}\nabla^2 p + Q, \tag{3.2}$$

where G is the shear modulus, u the displacement, ν the Poisson's ratio, α the Biot-Willis coefficient, F the body force, k the permeability, μ the fluid viscosity, S_ε the constrained specific storage, ε_{kk} the volumetric strain, and Q the fluid source (Wang, 2000). Equations (3.1) are nearly identical to the classic equations for linear elasticity except for the coupling of pore pressure in the conservation of linear momentum equations (3.1) and the fluid flow coupled to strain by the requirement of fluid continuity (3.2). However, the system (3.1)-(3.2) is more difficult to solve,

with analytic solutions restricted to a few highly idealized cases. We solve the respective equations numerically using the three-dimensional finite element software Abaqus FEA (Dassault Systemes, 2020; LaBonte *et al*., 2009; Pearse and Fialko, 2010; Hill *et al*., 2023).

The pore pressure diffusion is governed by an inhomogeneous diffusion equation Eq. (3.2). Because the fluid flow is coupled with the strain field pore pressure changes have direct effects on the stress and changes in the strain have direct effect on the fluid pressure. Under different assumptions, the stress field will uncouple from the pore pressure field and the diffusion equation resembles its hydrogeologic counterpart; the ground water flow equation $S\frac{\partial p}{\partial t} = \frac{k}{\mu}\nabla^2 p + Q$ (where $S = S_\varepsilon \frac{K_v^{(u)}}{K_v}$) (Detournay and Cheng, 1993; Wang, 2000).

Following (Gorelick and Remson, 1982) and (Gorelick *et al*., 1993), we use a physics-based numerical model to generate a unit source response matrix (see section 3.5.1). The key difference is that our simulation model incorporates the fully coupled poroelastic response (3.1-3.2), calculated using a finite element model, and generates a unit source response matrix of Coulomb stress (3.3) which is only possible due to the linearity in all the equations. The Coulomb stress is also dependent on fault geometries (SM Figure 3.B.1).

Stressing rate and earthquake probability

Triggering of seismic events due to fluid injection can be adequately described by equations (3.1-3.2) and changes in Coulomb stress (Wang, 2000; Cocco, 2002). We define Coulomb stress τ as:

$$\tau = \tau_s + \mu(\sigma_n + P), \tag{3.3}$$

where τ_s is the shear stress on a fault plane, σ_n is the normal stress (compression is deemed negative), P is the pore pressure, and μ is the coefficient of friction. An increase in pore pressure reduces the absolute value of the effective stress $(\sigma_e = \sigma_n + P)$ such that the Coulomb stress

increases, corresponding to promotion of failure. In the presence of a regional stress field even modest perturbations in pore pressure may encourage slip on preexisting critically stressed faults. The diffusion of pore pressure is highly dependent on hydraulic properties. Furthermore, depending on fault geometries, the poroelastic coupling of the fluid may play a significant role in promotion or inhibition of fault failure, especially in the far field where the effects of fluid percolation are negligible (Segall and Lu, 2015).

Similar to previous work (Langenbruch *et al*., 2018), which was carried out in the region of north-central Oklahoma and southernmost Kansas, seismicity data in the Raton Basin also shows the expected increase of earthquake probability with the rate of stress increase (Supplementary Methods). These observations can be used to describe the monthly earthquake rates $R_{\geq M}(\boldsymbol{r},t)$ according to a modified Gutenberg-Richter law for induced earthquakes (Langenbruch *et al*., 2018):

$$R_{\geq M}(\boldsymbol{r},t) = 10^{a(\boldsymbol{r},t)-bM} = \left[\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r},t)\right]^2 10^{\Sigma_\tau(\boldsymbol{r})-\boldsymbol{b}M}, \tag{3.4}$$

Here, we replaced the pore pressure rate, used by Langenbruch *et al*. (2018) by the monthly Coulomb stressing rate $\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r},t)$ in space and time to add the effect of poroelastic coupling. $\Sigma_\tau(\boldsymbol{r})$ is the spatially varying Seismogenic Index (SI). The SI and b values are evaluated through a specific calibration period (see section 3.4.3). The calibrated parameters are then used to forecast expected earthquake rates and to initialize the management model (see section 3.5) for optimization. An important distinction from previous studies (Langenbruch *et al*., 2018) is the use of Coulomb stressing rate $\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r},t)$ as opposed to pressure rates. While pore pressure rates are still the dominant signal (SM Figure 3.B.2), the fully coupled numerical model takes into account the stress field.

3.4.2 Numerical Domain

The numerical domain was developed and discretized in Abaqus CAE (Complete Abaqus Environment, Dassault Systemes, 2020). The domain has horizontal dimensions of 120 km x 200 km and a depth dimension of 14 km, with the *y* axis corresponding to north in the Universal Transverse Mercator coordinates (Figure 3.3). The finite-element mesh consists of nearly 1.5 million first-order hexahedral elements. Characteristic element sizes vary from 5,000 m in the far field to less than 500 m near the injection wells and in the vicinity of the central basin. The depth domain is partitioned into the 5 distinct hydrogeologic layers of the basin. The heterogeneous hydrogeologic properties of the model are summarized in Table 1. Permeability and storage parameters of the primary injection formations, the Dakota-Purgatoire and Morrison-Glorieta, were calibrated from analysis of injection step-rate tests (see Supplementary Materials). The permeability k of the Dakota-Purgatoire formation and the Morrison-Glorietta formation is taken to be 6.4 - 6.8 $\times 10^{-14}$ and 5.8 - 8.9 $\times 10^{-14}$ m^2, respectively. While no wells penetrate the crystalline basement for diagnostic analysis of basement permeability, we chose a crystalline basement permeability ($k = 1 \times 10^{-15}$ [m^2]) that results in the best correlation between the observed seismicity rates and modelled pressure rates (Figure 3.2). While this permeability is slightly higher than that inferred from small-scale field measurements of basement in other regions, it is similar to large-scale measurements made in regions of induced seismicity. In addition, it is also consistent with depth-dependent permeability models for continental crust at the mean depth of seismicity ($k \approx 3.35 \times 10^{-15}$ [m^2]) (Shmonov *et al.*, 2003b), and constraints on in situ hydraulic diffusivity of the upper crust from observations of post-seismic deformation(e.g., Fialko, 2004). The increased permeability is chosen to capture the basin-scale permeable faults that transmit fluid pressure to seismogenic depths.

We assume initial conditions of equilibrium stress and pore pressure (Segall, 2010; chapter 9). Therefore, the model only considers the perturbing effects of the wastewater injection and does not include any tectonic loading. The bottom and sides of the model are fixed only in

the surface normal direction (the roller boundary condition). The top surface of the model is stress-free. We model the Sangre de Cristo Mountain complex of thrust faults as barriers to cross-fault fluid flow and use an insulating condition at the western boundary of the model. We use the same injection depth of 1,500 m for all wells as the former is the middle depth of the modelled Dakota-Purgatoire injection reservoir. We record pore pressure and stress perturbations at the mean seismogenic depth of ∼7,040 m which is equivalent to ∼38,000 observation points for each time step. Generation of the SI map requires the full 29 well injection profile data ranging from November 1994 to December 2017, giving rise to 331 time steps, while the 5 year response matrix models require only 61 time steps.

3.4.3 Seismogenic Index (SI)

The SI map is a map of the seismo-tectonic state controlled by the number and stress state of pre-existing faults in the crystalline basement affected by Coulomb stress changes (Figure 3.4) (Langenbruch *et al.*, 2018). The SI ($\Sigma_\tau(r)$) is determined in local regions of 7 km radius at ∼25,000 seed points. The seed points represent the interpolated Coulomb stress changes produced by the model at the mean seismogenic depth within the crystalline basement. The higher the SI ($\Sigma_\tau(r)$) at each seed point, the higher the earthquake rate caused by a given Coulomb stress increase, because a higher number of (or more critically stressed) preexisting faults are affected by the Coulomb stress increase (see Eq. 3.4).

Calibration of the SI is set based on a calibration time period. In this way, future modelled Coulomb stressing rates are used to forecast expected spatiotemporal earthquake rate. We set the calibration time (Nov 1994 to July 2016) of our SI map prior to the Glasgow et al., 2020 study and find that forecasted earthquakes (July 2016 to July 2020) are well explained by basin Coulomb stressing rate, despite lowered injection rates at this time (Figure 3.6).

Calibration of SI follows closely to previous methods (Langenbruch *et al.*, 2018). The

following steps are performed to calibrate the SI maps:

1. Monthly Coulomb stressing rates $\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r_n},t)$ at all n seed points with a radius of 7-km around a selected seed point up to a given calibration time t_c (we use Nov-1994 to July-2016) are extracted, squared, and summed $\sum_n \left[\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r_n},t \leq t_c)\right]^2$

2. The total number $N_{M\geq M_c}(t \leq t_c)$ ($M_c = 2.5$, see Supplementary Figure 3.B.3) of earthquakes within a 7-km radius around the current seed point observed up to the given calibration time is summed.

3. Estimate of the b-value is computed using all $M \geq M_c$ earthquakes recorded through the calibration time t_c in the complete study area.

4. The SI at location $\boldsymbol{r}$ is evaluated:

$$\Sigma_\tau(\boldsymbol{r}) = \log_{10} N_{M\geq M_c}(t \leq t_c) - \log_{10}\left\{\sum_n \left[\frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r_n},t \leq t_c)\right]^2\right\} + b(t_c)M \tag{3.5}$$

Due to to the occurrence of singular earthquakes outside of the local areas of elevated seismicity one can get outlier SI values. These events are often attributed to Coulomb stressing rates that are quite low which results in significantly larger than average SI at those locations. Prior work found that as soon as two earthquakes occurred within the chosen radius of any given seed point a good estimate of the SI can be obtained (Langenbruch *et al.*, 2018). Our region uses a smaller radius and calibration magnitude. Therefore we precondition the SI to only be evaluated when there are more than 3 earthquakes. We evaluate the sensitivity of the SI for a smaller 5-km radius and removal of the "more than 3 earthquakes" precondition. These changes produce an SI map that appears different, as outliers are now included, but the overall seismicity rate remains very similar (SM Figure 3.B.4-3.B.6).

Within the central basin region, we find that the SI varies by about 1.5 units (Figure 3.4). A one unit increase in SI is the equivalent of expecting 10 times more earthquakes for the same

CFS rate change at that location. A higher SI in the central basin corresponds spatially with the well known zones of seismicity: Tercio, Vermejo Park, and Trinidad.

The SI is dependent on the spatial density of the observed seismicity and the radius of inclusion. This implies that seed points without observed seismicity in a 7-km radius will not produce SI. For the purpose of forecasting seismicity and optimizing injection rates for the entire basin we use an inverse distance weighting interpolation (power=2, radius=∞) (Figure 3.4) in areas that have no observed seismicity during the calibration period. The interpolated map helps inform the Coulomb stressing constraints in the SI dependent response matrix models.

3.4.4 Results & Discussion: Forecast Performance (2016 - 2020)

The results of the time dependent pore pressure evolution and associated seismicity during our calibration time are shown in Figure 3.5. The pore pressure continues to increase at depth within the basin due to the diffusion of fluid pressure despite lowered injection rates during 2016-2022. The total pore pressure increases, but the rate of increase declines (Figure 3.2). Returning to Eq. (3.4), we can now forecast seismicity rate beyond our calibration time using both the SI map and Coulomb stress perturbations from the numerical model. Figure 3.6 depicts the seismicity rate forecasts from a variety of calibration time periods and the resulting projected seismicity rate between 2016 and 2020. There is little sensitivity of the modelled earthquake rates to the calibration time. We find that the observed seismicity rate from 2016 to 2020 is fit well by our calibrated SI model and the computed Coulomb stress changes.

Furthermore, assuming the occurrence of induced earthquakes follows a Poisson process (Langenbruch *et al*., 2011; Langenbruch and Zoback, 2016; Shapiro *et al*., 2010), the probability of exceeding a magnitude M, that is the probability to observe one or more events of magnitude

M or larger, is given by (Langenbruch *et al.*, 2018):

$$Pr(M) = 1 - Pr(0, M, N_{\geq M}) = 1 - \exp(-N_{\geq M}) \tag{3.6}$$

Where, ($N_{\geq M}$) is the expected number of events of magnitude M or larger in a considered time interval (see Eq. 3.4).

Based on our calibrated model, we compute the annual expected number of events in the range from M 2.5-6.5 and determine magnitude exceedance probabilities using Eq. 3.6 (Figure 3.7). Our results suggest that between 2016-2020 there was a $\sim$85% probability to observe one or more M$\geq$4+ earthquakes and a $\sim$18% probability to observe one M$\geq$5+. We find that Coulomb stress rates at seismogenic depth continued to trigger seismicity between 2016-2020 although injection rates declined. Therefore, induced seismicity was still driven by wastewater injection during this time period. Declining injection rates alone are not necessarily an indicator of decreased seismic hazard as one must also consider diffusion-driven time delays in the induced seismicity process.

3.4.5 Results & Discussion: Business As Usual Forecast (2022 - 2027)

In this section we explore the seismicity forecasted by our calibrated model from 2022 through 2027 under a 'business as usual' (BAU) injection scenario. The BAU scenario uses the last observed monthly injection rate for each well from May 2022 and holds them constant until May 2027 (Figure 3.8). This scenario serves as the baseline comparison for the optimization scenarios presented in Section 3. We list the following important results of the BAU forecast:

- The BAU forecast from 2022-2027 shows that the probability to exceed a M$\geq$5+ event is $\sim$15% and a M$\geq$4+ event is $\sim$75% (Figure 3.9).

- Spatially, higher rate injection wells are clustered in the central portion of the basin near

the Vermejo Park cluster. Injection wells in this area, just south of the CO-NM border, on average inject at rates higher than 20,000 m^3 per month (Figure 3.10 (B)).

- Seismic hazard is also mostly elevated in this same region for the BAU forecast (Figure 3.10 (A)). Within this region of clustered injection, the spatial probability to exceed a M≥4+ is ∼20% over the 5-year BAU forecast.

- Seismic hazard in the North of the basin is proportionally smaller. We interpret this as a result of lower injection rates, largely below ∼10,000 m^3 per day, and lower SI in this region.

- The two observed M4+ events that have occurred from May 2022 to September 2023 occur within the zone of elevated seismic hazard forecasted by our model (Figure 3.10).

- In comparison to a complete shut-in of injection in May 2022, BAU injection increases the likelihood of an M≥4+ event by 150% (from 30% to 75%) and a M≥5+ by more than 200% (from 5% to 15%) (Figure 3.9).

SM Figures 3.B.7-3.B.8 show the seismicity rate forecasts resulting from the BAU projected injection rates. The forecasted seismicity rates are used to produce magnitude exceedance probabilities from our calibrated SI model (Figure 3.9). Figure 3.9 also includes the lower bound on any optimization we can achieve, the shut-in scenario, which represents the post-diffusion pore pressure and stress effects from the full injection history (ie. blue line in Figure 3.8). The 5 year hazard for the shut-in scenario is also characterized spatially for a probability of exceeding a M≥4+ (Figure 3.11). Given enough prior seismicity to produce a SI map and a physical model to produce Coulomb stress rate any future injection scenarios can be considered in our model. We elaborate on three management models in the following sections.

3.5 Physics-Based Forecasting with Optimization

3.5.1 Methods

The previous sections describe the methods to construct the simulation model built from two data sets: (1) the physics-based poroelastic model and (2) the statistical seismicity model or SI map (Figure 3.12). In this section we describe the additional methods required to frame our problem as a management model that allows for varied optimizations. In our optimization model, the objective function allows for the maximization of a desired objective, i.e. total injection rate, using decision variables (monthly injection rates) subject to constraints, such as CFS rate at a particular location. In order to solve this optimization problem, we must build a response matrix of the system and use mixed-integer and linear programming to resolve our objective. An overview of the simulation-optimization procedure, including the construction of the simulation model, is provided in Figure 3.12.

Objective Function

In our study of the Raton Basin, the objective function is framed to maximize a desired objective over the 5-year management period. This objective function is maximized subjected to specific constraints, i.e. Coulomb stress or Coulomb stress rate $\dot{\tau}$, below a threshold at chosen locations. Linear programming employs the unit-source solutions of the response matrix by linear superposition to acquire the optimal injection rates at each of the 29 wells in our model. The general framework of the linear program is represented as:

$$\min_{q} f^T q \tag{3.7}$$

subject to

$$Rq \leq x \tag{3.8}$$

$$0 \leq q \leq ub \tag{3.9}$$

where q is the injection rate at each of the wells for each time step (i.e. monthly), f^T is a row vector of negative ones [-1,$\cdots$,-1] so that the objective function seeks to maximize the cumulative injection, R is the response matrix (see section 3.5.1), x is the constraint vector ($\dot{\tau}$) at each of the model output locations, and ub is the upper bound on the monthly injection rate for each well. For all optimization scenarios presented, the upper bound for a single well injection rate is 1500 m^3/day, which represents the threshold of high-rate well injection nationwide (Weingarten *et al.*, 2015). We solve the linear program using the linprog() function in MATLAB which generates optimal values of q, i.e. the injection rates, for each well that does not exceed the constraints at the model output points. This objective function subject to various constraints is flexible and adaptable to a wide variety of adjustments within linear programming optimization. In section 3.5.1, we elaborate on different ways to alter the management model constraints and provide a selection of controls that may be of interest to real-world injection practices.

Response Matrix

Given any linear system used to describe a given simulation model, a management model can be built with a response matrix. Construction of the response matrix requires individual unit-source solutions for each well operating within the management model. A unit-source solution is generated by producing an impulse from an individual well (i.e. unit flow rate) and measuring its response at all model output locations for the duration of the management period. The impulse has a fixed value for a specified period and a value of zero thereafter. The response of the system are changes in pore pressure and stress. Due to the linearity of the Coulomb stress

equation (Eq. 3.3), Coulomb stress and Coulomb stress rate are derived from this response (see Appendix for rate response matrix construction).

In our model, the Raton Basin contains 29 wells. Therefore, we must generate 29 independent, unit-source impulses (one for each well) and record the unit response at all model output locations. We must record each response for the entire 5-year management period (ie. June-2022 to June-2027). Each time step in the model is 30 days. Hence, the unit-source response is a single flow rate equivalent to 100 m^3/day for the first time step and then zero for the 60 months after. The result of this procedure is the unit-source response matrix of CFS rate produced by each well at every model output location (SM Fgure 3.B.9). An example of this procedure is provided in the supplement (SM Methods 3.B.3; SM Figure 3.B.10).

Considering Injection Prior to Management Time Period

Our optimization management model optimizes injection rates under a set of given constraints for a prescribed management time period. It does not, inherently, consider injection prior to the management time period. We solve this issue by taking the difference of Coulomb stress between two simulations: (1) an ABAQUS simulation which considers all injection from Nov 1994 - 2027 (BAU rates) and (2) a response matrix simulation which considers only injection from 2022 - 2027 (BAU rates). The resulting Coulomb stressing rates represent the contribution of all prior injection during the management time period. This could be considered a 'complete shut-in' scenario from 2022 - 2027.

We calculated seismicity rates and a probability of exceedance curve expected from this shut-in scenario (Figure 3.9). SM Figure 3.11 depicts the spatial distribution of hazard for yearly time steps. If wells were to have suddenly shut-off in May 2022 our model predicts that there would still be a $\sim$35% probability of exceeding a M$>$4+ earthquake in the next 5 years. The shut-in Coulomb stress rate perturbations are added to the Coulomb stress rate constraints of the optimization results prior to the seismicity rate and seismic hazard calculations, thus serving as

the initial conditions or starting point in the optimizations. This step is essential, otherwise the seismic hazard is underestimated by the optimizations alone.

Mixed Integer Programming

Mixed-integer programming (MIP) allows the optimization manager to impose constraints that simulate real-world injection practices (Gorelick and Remson, 1982; Hsu and Yeh, 1989). Without MIP, the optimization solution is free to produce large swings in injection rate at individual wells. In reality, large injection wells have tolerances for injection rate changes over time. MIP allows the optimization manager to place controls what wells are operating and how the wells operate (independent or dependent on one another) through time. Injection rates can be constrained within a running average of past injection at a particular well, or monotonically increase or decrease injection through time, or exclude certain wells during certain periods.

The process of applying different types of MIP constraints is similar for most scenarios. First, a mixed-integer matrix is constructed R^* such that $R^*q \leq x^*$, where q is the corresponding injection well location for each management period and x^* is a vector of additional constraints. Both R^* and x^* are concatenated with original response matrix equation, Eq. (3.8), and the objective function is maximized subject to these combined constraints (R and R^*). A simplified example is provided in SM Section 3.B.3, and further description of applying each type of MIP constraint in the management model is provided in SM Section 3.B.4.

Setting a Desired Seismic Hazard

The optimization problem described above is setup to constrain only CFS rate at specified locations through time. However, the optimization manager may still use our methodology to achieve a desired seismic hazard. This is performed by combining the calculated CFS rates with the SI model to produce seismicity rate forecasts. Optimization is still possible without coupling to a SI map if desired (See Supplementary Methods 3.B.3; SM Figure 3.B.13-3.B.14).

For a desired magnitude exceedance probability $Pr(M)$ (Eq. 3.6), a user can solve for the total number of earthquakes expected during the management period ($N_{\geq M}$). This $N_{\geq M}$, in combination with spatially varying SI map $\Sigma_\tau(r)$, can be used to calculate desired Coulomb stress rate constraints $x_{\dot{\tau}}$ for the management model:

$$x_{\dot{\tau}} = \frac{\partial}{\partial t}\tau_\tau(\boldsymbol{r},t) = \sqrt{\frac{N_{\geq M}}{P \cdot T} 10^{-\Sigma_\tau(\boldsymbol{r})+bM}} \tag{3.10}$$

where P now refers to the total number of constraint points in the SI model and T refers to the to total time chosen for the management period. This initialization assumes that each point in the model will carry a scaled portion of the total earthquake probabilistic hazard- ie. $\frac{N_{\geq M}}{P \cdot T} 10^{bM}$ which is scaled by the SI (ie. $10^{-\Sigma_\tau(r)}$). In our case, the total number of model points exceeds the computational limitation of the linear program and a subset of the total model points must be chosen. For example, the output of our model contains >30,000 points across the basin, but we reduce this total to 500 constraint locations for the management model. The chosen points are based on a uniform random distribution of points within a circle that contains all of the seismicity (SM Figure 3.B.15).

In practice, we have found that the CFS rate constraints provided by equation 3.10 always produce a basin-wide $Pr(M)$ lower than the desired threshold $Pr(M)$. The desired threshold $Pr(M)$ would only be met if the CFS rate constraint threshold is met at all points P for all time T. To resolve this issue, we iteratively solve the optimization model while increasing the CFS rate constraints at locations within the model that reached that threshold at any time during the management period. In this way, the constraints slowly increase based on which locations require a higher CFS rate in order to produce the desired $Pr(M)$ in the basin. For our study, we set a goal of achieving the desired $Pr(M)$ in the basin to within $\pm 0.2\%$ (See Methods 3.B.5).

The following steps describe the methodology, generalized for application to other studies:

1. Choose a desired exceedance probability for an arbitrary magnitude threshold and solve for

$N_{\geq M}$ (Eq. 3.6).

2. Calculate CFS rate constraints for the management model (Eq. 3.10).

3. Find optimal injection rates for calculated CFS rate constraints.

4. Calculate exceedance probabilities $Pr(M)$ across the basin for the optimized solution.

5. Check if exceedance probabilities $Pr(M)$ are within $\pm 0.2\%$ of desired $Pr(M)$.

6. If yes, skip steps 7 and 8.

7. If no, adjust CFS rate constraints dependent on too high or too low of threshold.

8. Return to step 3.

3.5.2 Prospective Case 'Reduction' - Reduce the Seismic Hazard

The first prospective case we consider is called 'Reduction' (Figure 3.8 - Prospective Case Reduction). Prospective case 'Reduction' is the management solution for a hypothetical well operation that seeks to reduce the overall injection and maintain the hazard within a chosen threshold. We include a constraint that the overall injection must be reduced by at least 80% from May 2022 levels by the end of the 5 year management window. Additionally, we constrained seismic hazard such that the probability of exceeding a M$\geq$ 4+ event is 40% lower than the BAU forecast (Figure 3.9). The optimization and iterative method arrive at a solution to these constraints while maximizing the amount of fluid injected.

In order to achieve a smooth tapering of injection from the BAU initial injection rate of $\sim$10,000 m^3 per day we incorporate a MIP constraint to the management model. The constraint is a monotonic decrease of at least 2% each month for all injection wells (see 3.B.4) (Figure 3.8 - yellow line). This constraint smoothly reduces the overall injection rate and therefore the Coulomb stress rate by the end of the five year management period.

We find that there are several wells in the optimization that are never injecting, and that the algorithm preferentially chooses injectors towards the northeast more than other locations (Figure 3.13b). The northeast portion of the basin is a relatively low SI area (Figure 3.4). The west-central portion of the basin, which contains the highest SI hazard, does not have large amounts of injection during the management period. The optimization preferentially chooses to spread out large injectors from one another and to regions of lower SI (Figure 3.13b).

Another important observation is that prior injection still drives significant hazard due to the time delay of pressure diffusion continuing to elevate the Coulomb stress rate in the periphery of the basin (Figure 3.13a). Hazard is elevated in the west-central and western portion of the basin by prior injection, despite the optimization lowering injection in these areas. Our iterative technique still slowly reduces injection at wells and areas associated with high prior hazard if hazard thresholds are not initially met. In this way, our method takes into account prior injection through iterative forward solutions without direct inclusion in the optimization constraint vector (see Section 3.B.5).

The enhanced hazard to the west in all of our models does not consider previously mapped faults unless they were captured by the SI map. This hazard is primarily driven by continued Coulomb stress rate increase from prior injection. The inclusion of known faults is currently a limitation to our method. However, additional spatial constraints from known faults could be implemented as additional rows/elements in the response matrix/constraint vector prior to optimization. Constraint thresholds of Coulomb stress or Coulomb stress rate could be applied to these known faults.

Visualizing the optimization at each time step is informative to the evolution of hazard and how each individual well injects over time (SM Video 1). For the prospective case ‘Reduction’, wells inject continuously in the northeast - a low SI area - for the entire management period. Higher SI areas still receive injection but the optimization tends to spread the overall hazard across the basin.

3.5.3 Prospective Case 'Safety'

Our second prospective case consider how the optimization algorithm might disperse BAU injection rates in order to minimize seismic hazard (i.e. 'Safety') (see Section 3.4.5 and Figure 3.8).

The second optimization solution, which we call prospective case 'Safety', seeks an optimized solution that lowers the overall seismic hazard while the basin-wide injection rate is constrained at May 2022 levels for the 5 year management period. The optimization will preferentially increase volume in wells where SI is lower, because the Coulomb stress rate constraints will be relaxed in these areas (see Equation 10). By moving injection volume to wells and areas with lower SI, the forecasted seismic hazard is reduced. The solution therefore produces an overall annual exceedance curve that is lower for the same total injection volume (Figure 3.8 - pink line).

Figure 3.14 describes the optimization results across the basin for prospective case 'Safety'. When the spatial distribution of injection is compared to the Business As Usual case, we find that the optimization spread injection volume out more evenly throughout the basin, instead of clustering injection in the central region. At the same, seismic hazard increases on the peripheries of basin away from the higher SI zones in the central basin. In the central basin, forecasted haard is reduced greatly, with less than 2% probability to exceed an M≥4+ within 7 km. This is compared to nearly 20% probability to exceed an M≥4+ within 7 km in the Business As Usual case in the central basin. Forecasted hazard is highest in the northeast portion of the basin, with 10% probability to exceed an M≥4+ within 7 km.

Our solution, during the 5 year management window, reduces the basin-wide annual exceedance probability M≥4+ from 75% to 71%. This optimized result is a relatively small reduction in the annual exceedance probabilities. However, we found that injection prior to the management period contributes to a large portion of the overall hazard observed during the 5 year window. If the prospective case 'Safety' is run without prior injection, the optimization can

reduce the annual exceedance probability M≥4+ from 75% to 58% (Figure 3.9 - green line). This reduction in seismic hazard is due to the optimization shifting injection to areas of lower SI.

Simply excluding prior injection does not, in and of itself, reduce the overall exceedance probabilities. We ran a seismic hazard forecast for the Business As Usual case excluding prior injection and found the annual exceedance probability for a M≥4+ earthquake increased from 75% to 80% (Figure 3.9 - BAU without prior injection line). The reason for this increase in overall seismic hazard when excluding prior injection is that prior injection was on a long-term decline, especially in areas with high SI. These declining injection rates prior to the management time period actually reduce the Coulomb stress rate in areas where the BAU injection is high. Therefore, counter intuitively, excluding prior injection increases the seismic hazard in the BAU case and decreases in the 'Safety' case.

The results from the 'Safety' case reveal that prior injection can have a large influence on how much the optimization method reduces overall seismic hazard. Furthermore, it highlights the importance of optimizing injection as early as possible in the course of an induced seismic sequence. In the case of Raton Basin, injection and induced seismicity have been ongoing for multiple decades, which reduce the positive safety effects of minimizing seismic hazard during the management period.

3.5.4 Prospective Case 'Economic'

The third optimization solution, which we call prospective case 'Economic', seeks to increase the overall injection rate but maintain the same basin-wide seismic hazard as the BAU case (see Section 3.4.5 and Figure 3.8). In this case, we allow the optimization freedom to increase the overall volume that can be injected in any month of the 5 year management window. An optimal solution is found when the basin-wide annual exceedance probabilities are within $\leq 2\%$ of the BAU probability of exceedance for M≥4+ ($\sim$75%). We include two constraints on

individual wells in this solution: (1) no individual well injection rate can exceed 1,500 m^3/day, and (2) an MIP constraint that limits individual well injection rates to within a 6-month running average so that the optimization cannot drastically front-load or back-load the management period with injection volume. Again, the Coulomb stress rate constraints derived from the SI map force the optimization to preferentially increase volume in areas away from the largest seismic hazard (i.e. lower SI).

An optimal solution was found for the 'Economic' case, which increased the overall injection rate basin-wide compared to the BAU case (Figure 3.8 - green line). The solution shows a gradual increase in basin-wide injection rate from $\sim$300,000 m^3/month in 2022 to $\sim$375,000 m^3/month in 2027. The increase in cumulative volume injected in the 'Economic' case is more than 1,080,000 m^3 ($\sim$6,750,000 barrels) when compared to the BAU case.

The spatial distribution of injection in the 'Economic' case shows a substantial change in the how the field would be operated during the 5 year management period (Figure 3.15b). Of the 29 potential injection wells, the optimization chooses to inject at only 12 wells, while the remaining 17 are completely shut-in. Of the 12 wells which operate during the 5 year window, only 6 inject at rates higher than 20,000 m^3/month. These 6 injectors, where the vast majority of fluid is injected, are spread out across the entirety of the well field and to regions of lower SI. These 6 wells inject at a more or less a constant rate for the entire management time (SM Video 2). Clustering of injection is held to a minimum when compared to the 'Reduction' or 'Safety' case.

This case highlights what the optimization method ultimately attempts achieve: spatially distributed injection across regions of lower SI. By spreading out injectors, the basin-wide Coulomb stress rate is reduced by minimizing superposition of clustered injectors. By concentrating injection in regions of lower SI, the Coulomb stress rate that is created by injection results in lower induced seismicity. This combination of effects – spatially distributed injection in regions of lower SI – allows for the highest basin-wide injection rates (and largest cumulative injected

volume) for a given seismic hazard.

3.6 Discussion

The combination of physics-based forecasting with optimization management shows promise for future work in mitigating induced seismic hazard at the basin-scale. The optimization framework allows a user to maximize a particular objective (i.e. reduction, safety or economic) while maintaining a specified induced seismic hazard. Our method is also flexible and adaptable to other regions or other types of fluid injection that induce seismicity. The main components are the following:

1. **Physics-based model of pressure and/or stress change.** First, a physics-based model of injection must be built of the region that has good estimates of the relevant reservoir flow parameters. Here, we have built a fully coupled, poroelastic numerical model using the finite-element method calibrated using injection data from reservoir step-rate tests. However, a finite-difference model could also work (e.g. MODFLOW). Any linear system is the key. Depending on whether the poroelastic stress effects are marginal to the pore pressure effects may influence this decision.

2. **Seismogenic Index (SI) Map.** Second, a SI map (see Section 3.4.3) must be calibrated from the empirical relationship of seismic response to injection. Thus, some degree of prior injection and earthquake history are required for forecasting. Without the SI map, optimization is still possible, but will not be constrained by desired seismic hazard.

3. **Response Matrix.** Third, a response matrix of system is built from impulse-responses of the system to a unit injection at each prospective injection site (see Section 3.5.1). The response matrix allows the optimization to scale injection rates of individual wells to find the combination which both satisfies the constraints and maximizes the objective function.

4. **Optimization Framework.** Lastly, an optimization framework of an objective function, constraints and decision variables are input. The model then seeks the optimized solution that will satisfy either a reduction, safety or economic objective and maximize fluid injected.

The adaptability of this method to other regions is possible through the gathering of required basin-specific input data on reservoir flow parameters, injection and seismicity response. In addition, the method is flexible enough to consider any fluid injection that produces a linear poroelastic response. Listed below are some of the potential improvements and limitations of the current framework:

1. **Real-time optimization and forecasting:** Once the physics-based model and SI map are initially calibrated the user could develop an optimal injection strategy and continuously update the SI map if seismicity evolves in new areas. The response matrix method allows for quick integration of new constraints without the need to re-run elaborate physical models continuously. Therefore, rapid adjustments in well optimization are possible as the SI adjusts and improves in new areas of the basin.

2. **Stacked optimization for model uncertainty:** As described in Section 3.B.4, stacked optimization allows the user to find one set of optimal injection rates that explicitly account for the uncertainty in the physical model. The existing framework contains uncertainty in the seismic hazard due to the Poisson distribution within the SI model. However, stacked optimization allows the user to consider uncertainty within the physical model (i.e. a distribution of flow parameters). Stacked optimization does require more computational power as it requires N (where N is the number of wells) additional model runs for each uncertain distribution to be appended to the response matrix.

3. **Non-linear programming:** Non-linear programming allows optimization of non-linear objective functions and constraints. Currently, our linear program cannot explicitly optimize injection using seismic hazard (R) as a constraint because R is non-linearly related to CFS

rate. Therefore, we rely on an iterative approach to optimize injection to a desired seismic hazard (see Section 3.B.5). Non-linear programming may be able to address the issue of local-minima in the optimal solution where currently non-unique solutions may be found by a linear program. Our iterative method slowly adjusts the constraint locations one at a time to prevent any over saturation in hazard and injected fluid at any one location in the solution. Non-linear programming may be able to save computational time as compared to the iterative approach.

4. **Incorporating known fault maps:** A key piece of future work is the integration of known fault maps within the optimization framework. Known faults would serve as additional constraint locations appended to the response matrix and constraint vector, where pressure and/or stress change would be limited. From a practical point of view, known faults in many cases of induced seismicity are not the primary drivers of induced seismic hazard (i.e. Oklahoma), but users may desire to avoid stressing faults when optimizing basin-scale injection. This optimziation framework would allow the consideration of both an SI map and fault maps.

5. **Incorporating risk for policy:** While we looked at the total hazard in the region, it would be possible to constrain hazard spatially depending on seismic risk (Schultz *et al.*, 2021). For example, agreement might be met with industrial well operations that maximizes the fluid injected while restricting hazard in an area with high risk, like a densely populated area. A scientifically informed policy, for example one that limits the probability of exceeding a M$\geq$5+ earthquake within a high risk zone, could be met while still reaching the economic objective of the well operators.

3.7 Conclusions

Here, we investigated the relationship between wastewater injection and seismicity in the Raton Basin of Colorado and New Mexico using a physics-based forecasting framework. First, a 3D finite element model of a poroelastic crust is used to estimate time dependent Coulomb stress changes over the more than two decades of Raton Basin injection. The outputs of Coulomb stress rate from our finite element model were combined with a seismogenic index (SI) model to forecast induced seismicity in space and time throughout the basin. Using this hybrid physics-statistical forecasting model we found the following conclusions:

1. The recent and ongoing induced seismicity within the Raton Basin is well explained by our physics-based forecasting model. Declining seismicity rates between 2016 - 2022 are forecasted well by the decline in basin-wide injection rate. Despite injection rate declines, modeled Coulomb stress rate is still increasing in several regions of the basin, suggesting that induced seismic hazard is still ongoing. Our model also shows that induced seismicity is driven primarily by the pore pressure component of the poroelastic stresses, with poroelastic stress changes accounting for about 5% of the driving force.

2. Using our physics-based forecasting model, we estimated the induced seismic hazard produced by continued Raton Basin injection at May 2022 levels through 2027 (Business As Usual case). Our 5 year forecast estimates the probability to exceed a M$\geq$4+ event is 75% and M$\geq$5+ event 14%.

3. Linear-programming optimization using the response matrix method is implemented successfully using a safety objective framework that reduces seismic hazard for given amount of fluid injection (safety objective) or (b) maximizes fluid injection for a prescribed seismic hazard (economic objective).

4. Across the different objectives tested, the optimization algorithm tends to spread injection

out across the field when compared to the Business As Usual case. In the safety and economic objective cases, we observed the algorithm spreading out higher rate injection wells from one another and to regions lower seismogenic index (SI). We also demonstrate that injection prior to the optimization management period may have differing effects on seismic hazard during the management period. In the reduction and safety cases, we show that prior injection enhanced seismic hazard during the management period, thus decreasing the impact of injection optimization. We conclude that optimization of injection earlier in an induced sequence will allow for better control of seismic hazard during the management period.

Acknowledgements

Chapter 3, in full, is a reformatted version of a paper currently submitted for submission for publication. Hill, R.G., Weingarten, M., Langenbruch, C., Fialko, Y., 2023. Mitigation and optimization of induced seismicity using physics-based forecasting. Journal of Geophysical Research: Solid Earth, submitted. The dissertation author was the primary investigator and author of this paper.

Tables and Figures

Table 3.1: Material Properties. Hydrogeologic material values for different units and their corresponding depths in the numerical model. Note that the model begins at 1 km depth below the surface.

Unit	Pierre-Benton-Niobrara	Dakota-Purgatoire	Morrison-Entrada-Gloreita	Sangre De Cristo	Crystalline Basement
Depth (km)	1-1.4	1.4-1.6	1.6 - 2	2 - 2.8	2.8 - 15
Permeability (m^2)	$1{\cdot}10^{-20}$	$6.7{\cdot}10^{-14}$	$8.9{\cdot}10^{-14}$	$8{\cdot}10^{-15}$	$1{\cdot}10^{-15}$
E (GPa)	0.22	38	32	40.74	60
v	0.3	0.287	0.13	0.15	0.25
K_s (GPa)	0.34	33.8	26.6	36.6	42
ϕ	0.38	0.25	0.07	0.06	0.01

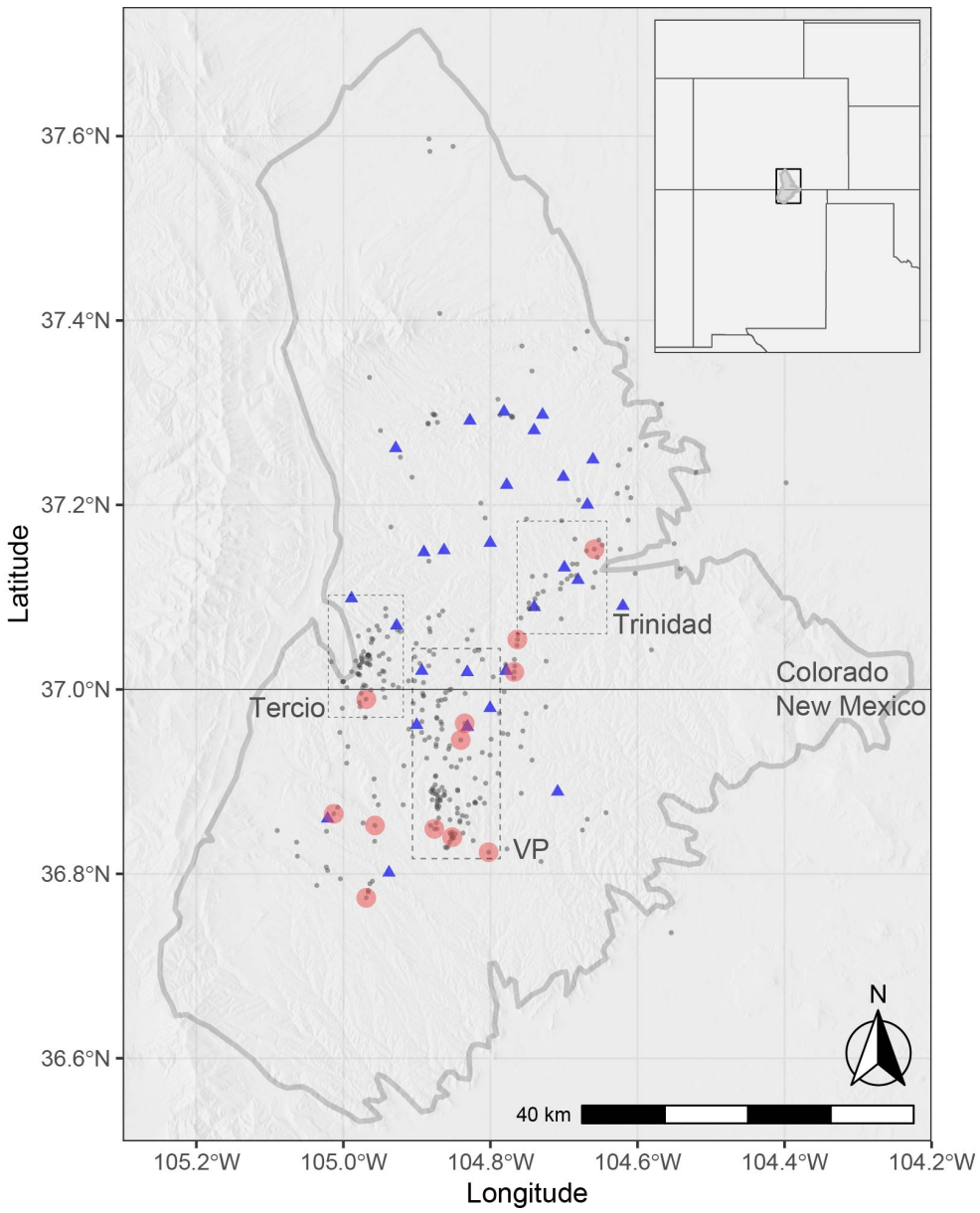


Figure 3.1: **Regional Context.** Light grey outline is the Raton Basin. Blue triangles are the 29 injection wells. Grey dots are earthquakes with M≥2.5 and red dots are earthquakes with M≥4 from Nov-2001 to July-2020. Boxed regions represent zones of seismicity: Tercio, Vermejo Park, and Trinidad.

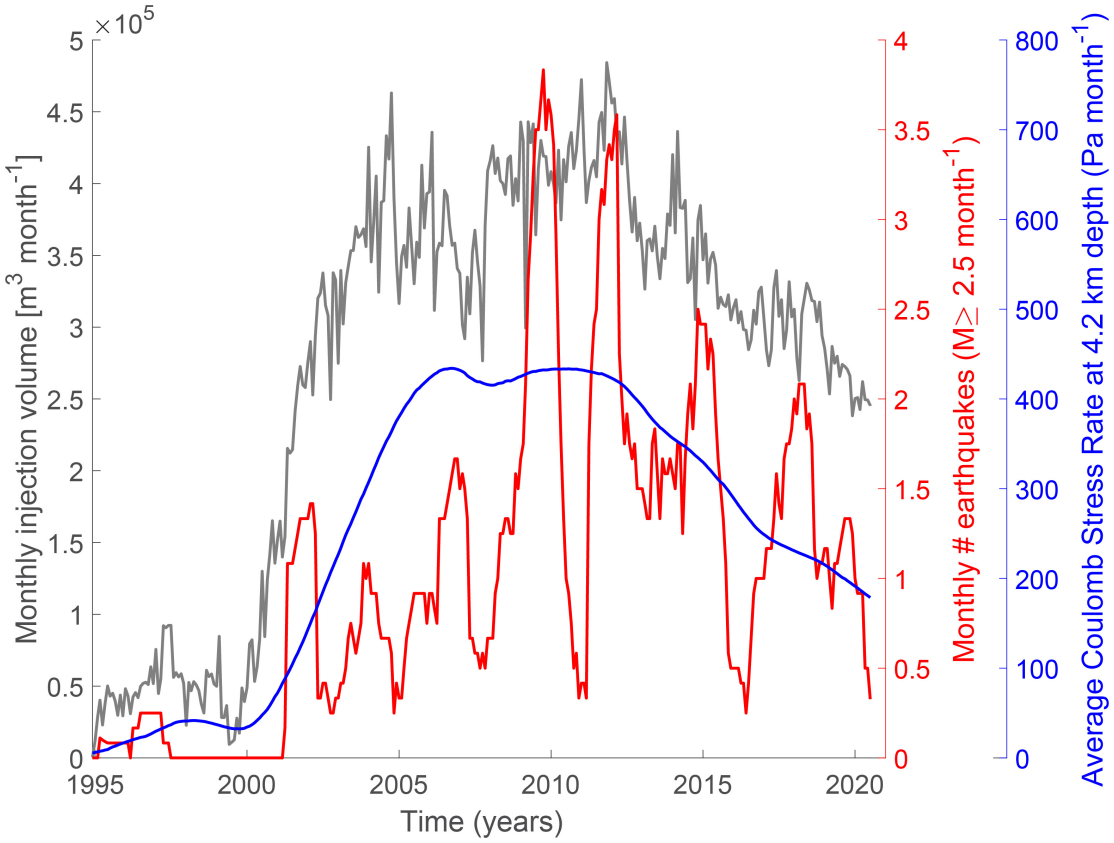


Figure 3.2: **Injection, induced earthquakes, and Coulomb stress rate.** Total monthly injection volume (grey), observed earthquakes M≥2.5 (1 year moving mean), and the average modelled Coulomb stress rate in the study area. The Coulomb stress rate lags the injection rate due to the diffusion of pore pressure into the crystalline basement. A correlation between increased stress at depth and seismicity is observed.

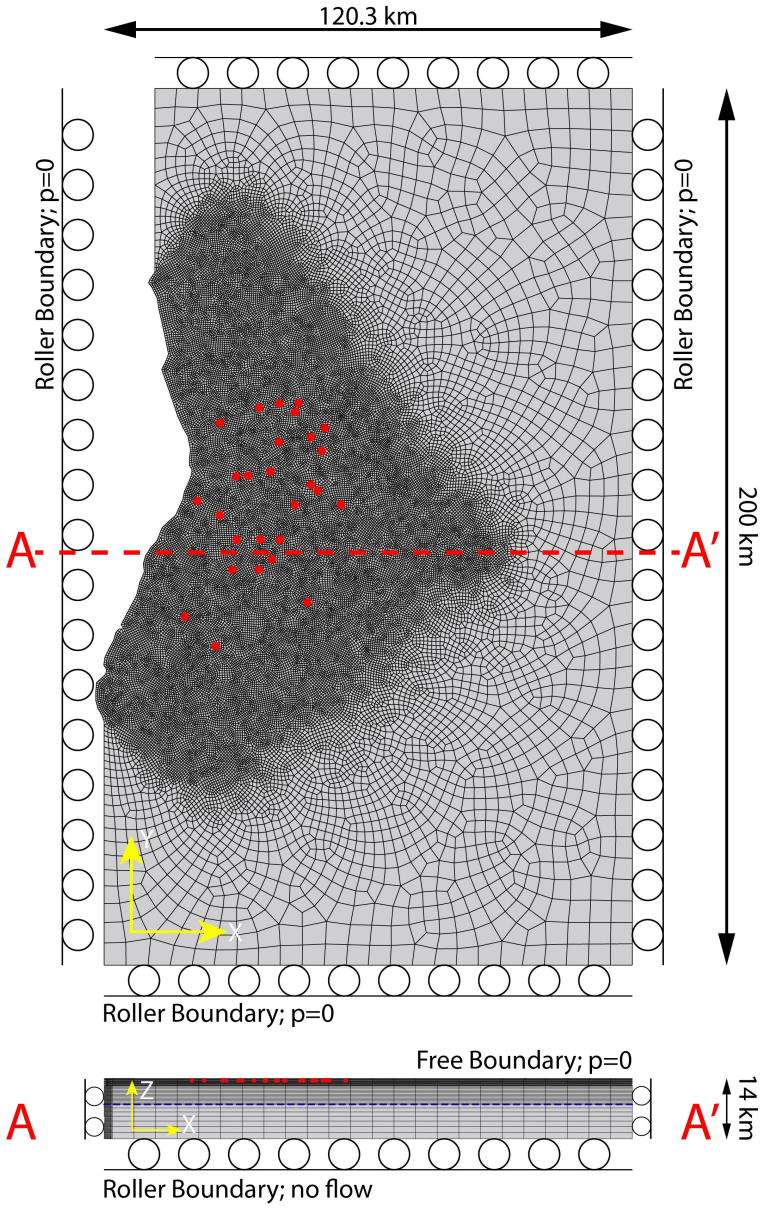


Figure 3.3: **Numerical Domain.** Three-dimensional finite-element model domain. The model mesh contains about 1.5 million hexahedron elements. The Red dots represent the well injection locations. The blue dotted line represents pore pressure and stress output location at the mean seismogenic depth (∼7 km depth or 4240 m below the top of the crystalline basement).

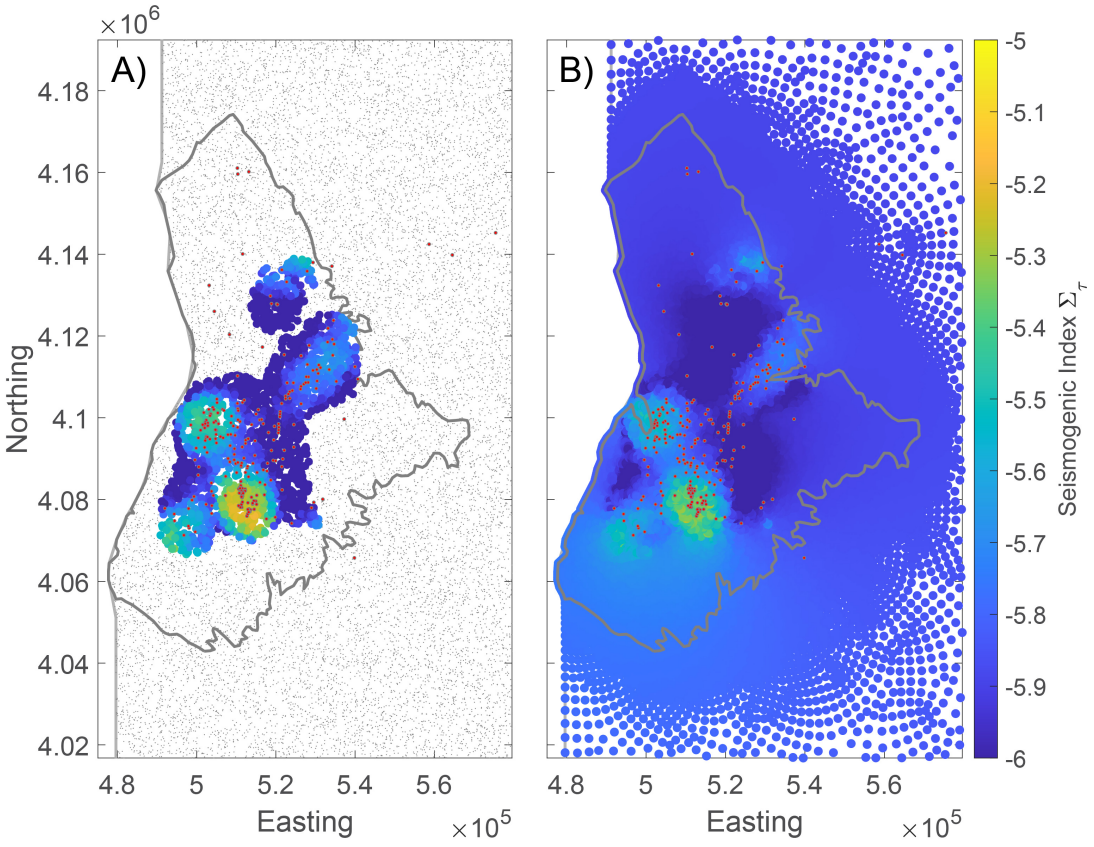


Figure 3.4: **Seismogenic Index Σ_τ Maps.** Mapped spatial variability of the SI in the Raton Basin. The SI is computed in local regions of 7-km radius around the 25,000 seed points (grey dots in panel A)). The calibration time is between Nov-1994 and July-2016. See Methods for additional details. Red dots represent earthquakes M$\geq$2.5 used in calibration. Panel B) represents the inverse distance weighted interpolation of the SI to the model points used in the forward model management solutions.

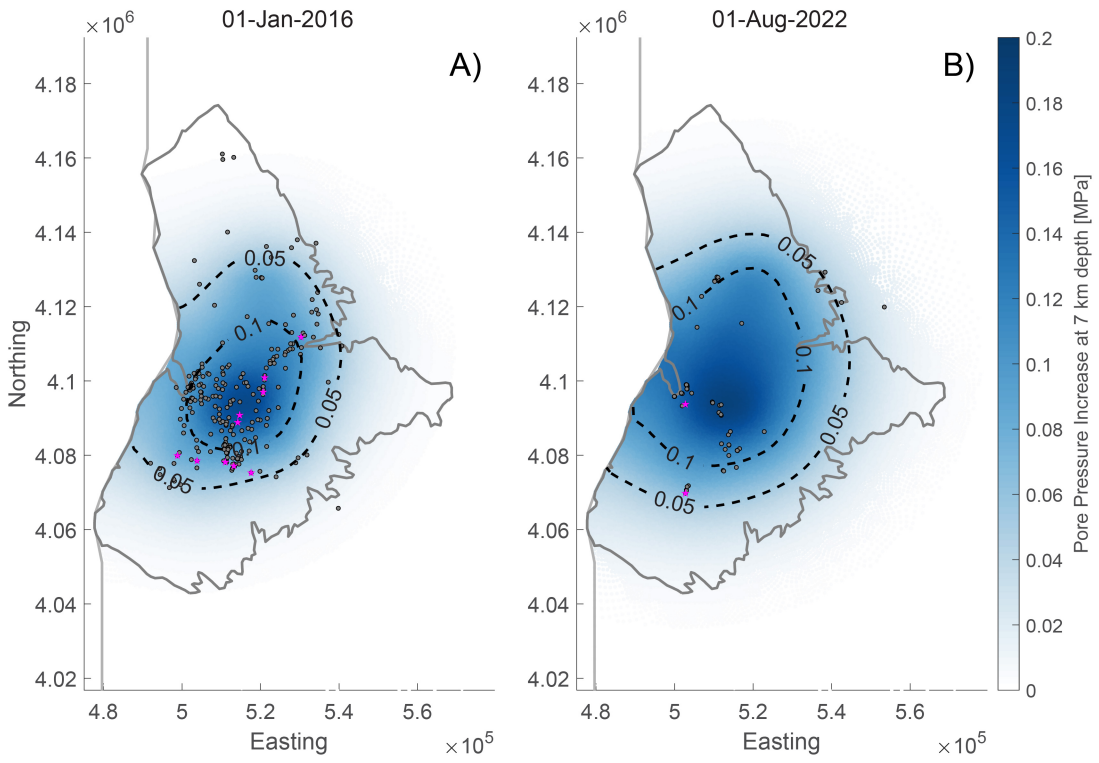


Figure 3.5: **Pore Pressure Increase.** A)) Pore pressure increase at mean seismogenic depth across the basin including seismicity from Dec 1994 through Jan 2016. Black dots represent earthquakes with M≥2.5+ and magenta stars are earthquakes with M≥4+. B) Pore pressure increase at mean seismogenic depth across the basin including seismicity between July 2016 to July 2022.

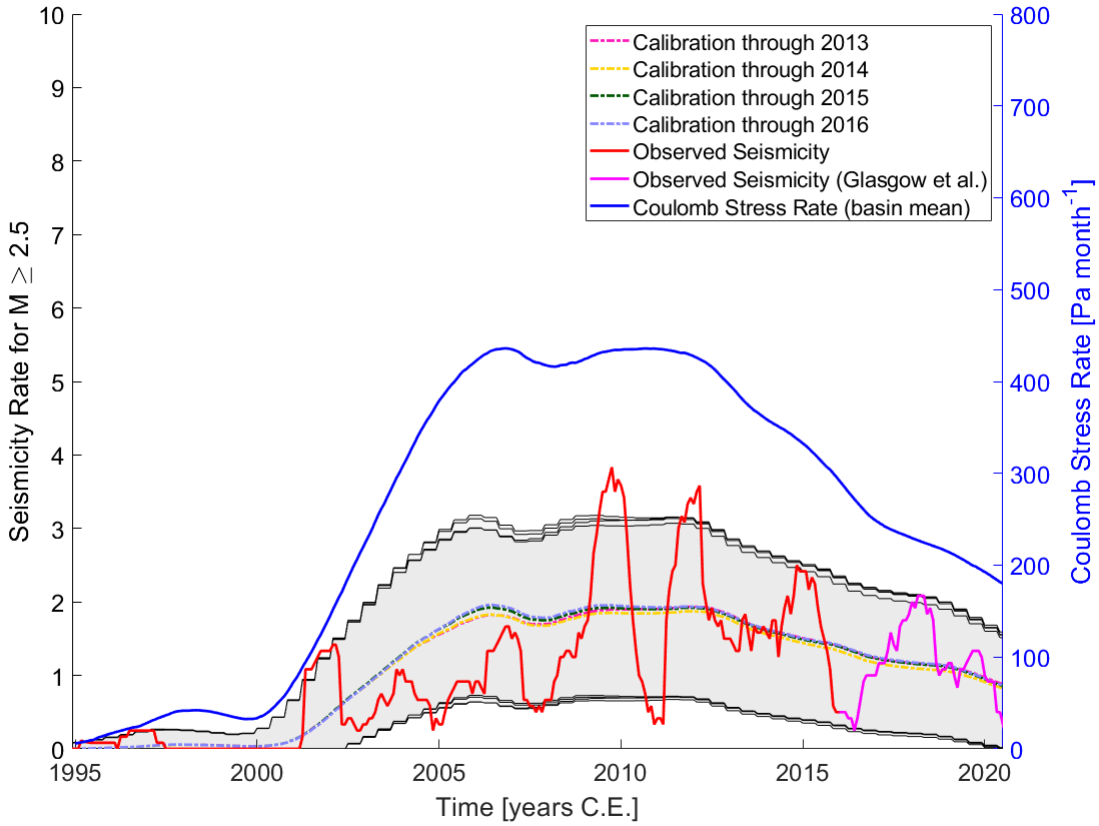


Figure 3.6: **Seismicity Rate Forecast.** Seismicity rate forecasts, above our completeness magnitude M≥2.5, compared to observed seismicity rate (1 year moving mean). Calibration period is from Nov 1994 through 2013, 2014, 2015, and 2016 prior to the Glagow et al., 2021 study (Glasgow *et al*., 2021a). The earthquakes and longest calibration time period used to calibrate the SI model is represented by the red line. The varying dashed lines and grey boundaries are the 95% confidence bounds forecasted by the seismicity rate produced from the SI model that includes the inverse distance weighted interpolation (right panel of Figure 3.4). Magenta line represents the observed seismicity from Glasgow et al., 2021 which is well explained by the seismicity rate forecasted by our model.

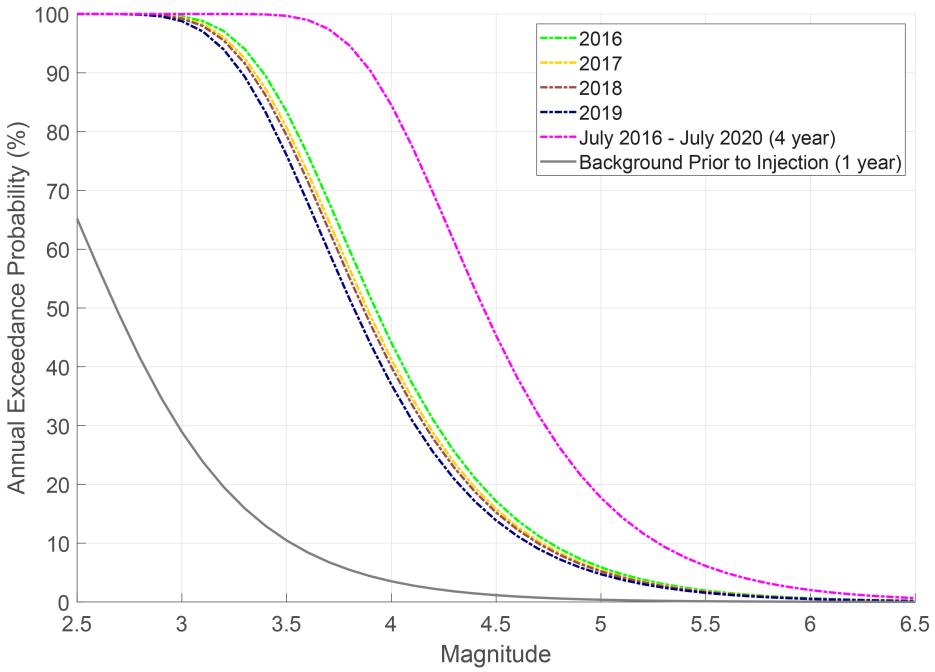


Figure 3.7: **Forecasted Magnitude Exceedance Probabilities.** Exceedance probabilities for magnitudes M≥2.5-6.5 from our physic-based forecasting model. Each line represents the probability forecasted by our model based on the calibrated SI map and computed Coulomb stress model outputs. The forecasted probability from 2016-2020 is significantly higher than the tectonic background (grey line) and is highest in 2016. Background probabilities are derived from prior work (Rubinstein *et al.*, 2014). Each year from 2016 to 2019 the the magnitude exceedance probabilities or decreasing, but still above the tectonic background level. From 2016 to 2020 the potential to trigger a M≥5+ increases to ∼18%.

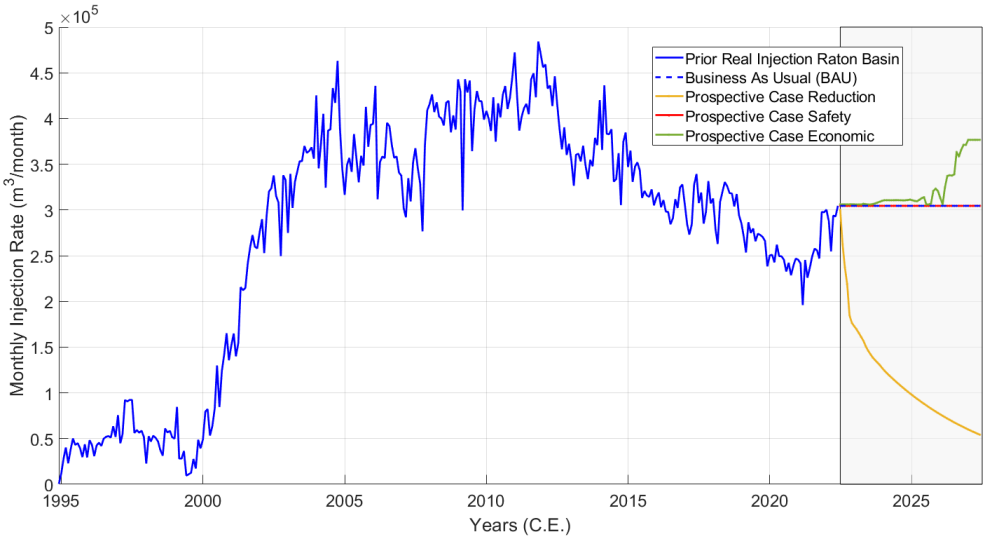


Figure 3.8: **Different Optimization Scenarios.** Plot shows the monthly injection rate (total of all 29 wells) for the observed data (blue). At June-01-2022, the next 5 year window (gray box) represent the forecasted injection rates. The business-as-usual rate takes the last known injection rates and holds them constant for the five years (blue-dash). The prospective case 'Reduction' is the optimized injection rates subject to reducing the overall injection by 70% in 5 years as well as a taper in individual well rates (yellow). The prospective case 'Safety' is the optimized injection rates subject to the constraint that the total fluid injected must be the same as the BAU, but reduces the overall hazard (Figure 3.9) (red). The prospective case 'Economic' is the optimized injection rates subject to the constraint that the overall 5 year hazard must be the same as the BAU, but increases the overall injection (green).

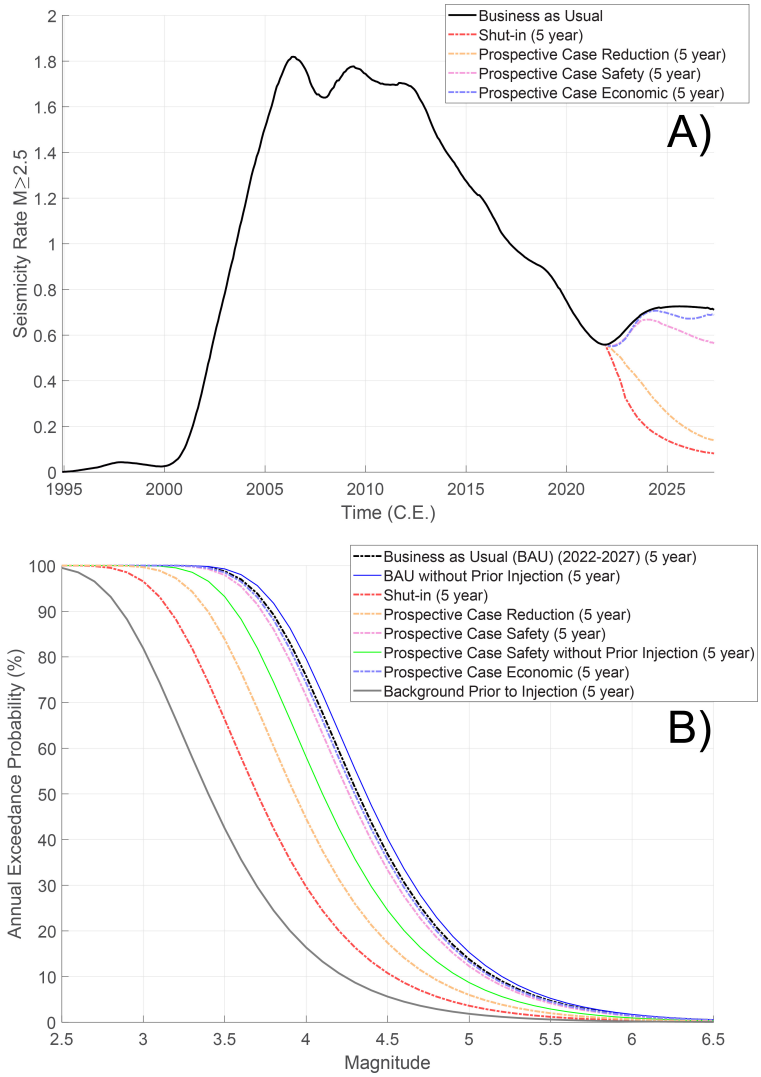


Figure 3.9: **Seismicity Rate Forecasts and Forecasted Magnitude Exceedance Probabilities (Optimizations).** A) Seismicity rate for M≥2.5 from beginning of injection until beginning of optimization management period. Each of the 5 year optimizations have an associated exceedance probability in the next panel. B) Exceedance probabilities for scenarios projected into the future (see main text). The Business as Usual (BAU) forecast is determined by extrapolating the last observed injection well data into the next 5 years. The shut-in forecast is determined in a similar way, but for immediate shut-in of all wells in June-2022. Prospective Case 'Reduction' considers reducing overall injection volume by 80% while not allowing the probability of exceeding a M≥4+ to be over 45%. Prospective case 'Safety' considers the same amount of fluid as the BAU case, but a more spatially optimized strategy based on the SI map. Prospective case 'Economic' optimizes to a solution for much more fluid for the same seismic hazard as the BAU case.

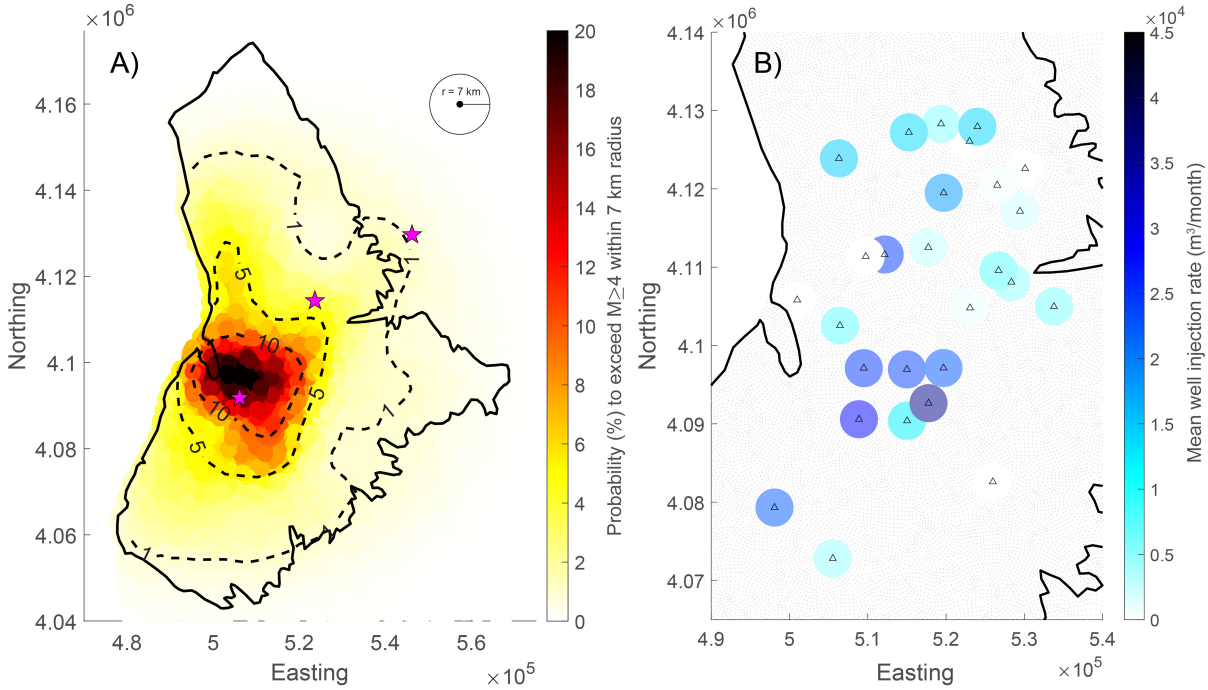


Figure 3.10: **BAU Hazard and Mean Injection Rate**. A) Magnitude exceedance hazard map for M≥4+ for the 5 year management window. Each location is taken as the sum in a 7 km radius. Magenta stars (3) represent the locations of actually observed M≥4+ earthquakes between June-2022 and Sept-2023. B) The Mean well injection rate (m^3/month) for all 29 wells (triangles) in the BAU extrapolation. Grey dots represent model nodes.

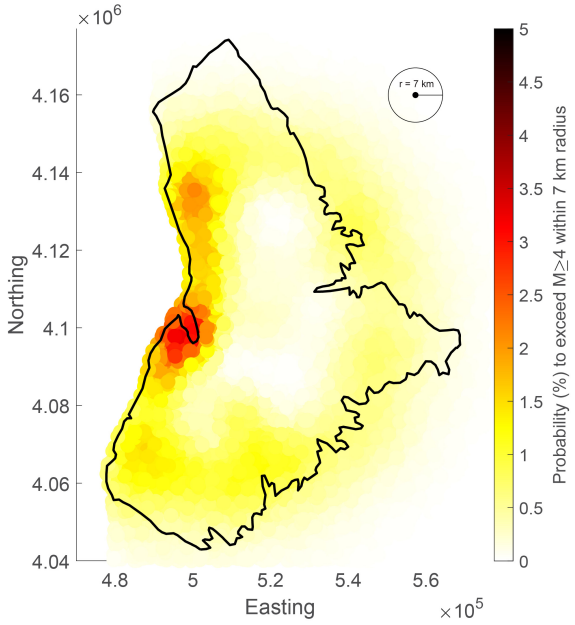


Figure 3.11: **Shutin Hazard**. The 5 year hazard for the shut-in scenario (all wells cease injection in May 2022 and stay off for 5 years) is also characterized spatially for a probability of exceeding a M≥4+. Shut-in represents the post-diffusion pore pressure and stress effects from the full injection history that continue to linger through the model and contribute to perturbations. Note that the colorbar axis is lower (5%) compared to all other maps which use 20% to clearly show the spatial distribution of the hazard.

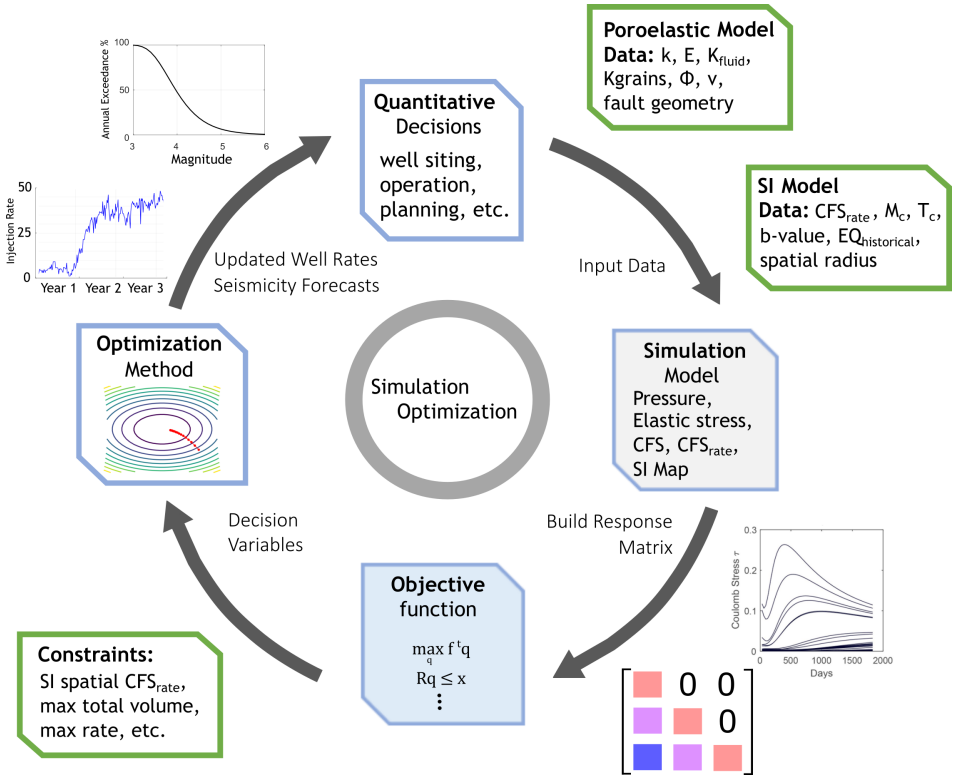


Figure 3.12: **Simulation Optimization Schematic**. Beginning at the top, operations consider quantitative decisions in well placing and operation prior to injection. By developing a numerical model and SI map from current injection a simulation model is built. The simulation model is used to build a response matrix which through linear programming solves a desired objective function (maximize the fluid injected). Additional constraints further inform the optimization which arrives at informed injection rates and spatial hazard maps to then advise future operation practices.

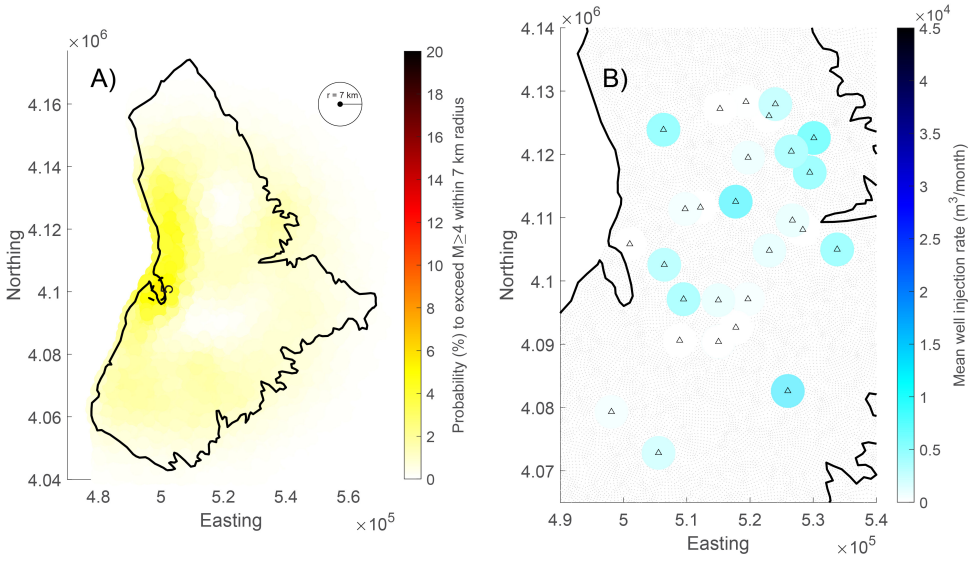


Figure 3.13: **Prospective Case 'Reduction' Results.** A) Magnitude exceedance hazard map for M≥4+ for the 5 year management window. Each location is taken as the sum in a 7 km radius. B) Mean injection rate (m^3/month) at each well location (triangles). There are several locations where the optimization chooses not to inject. The grey dots represent the model nodes.

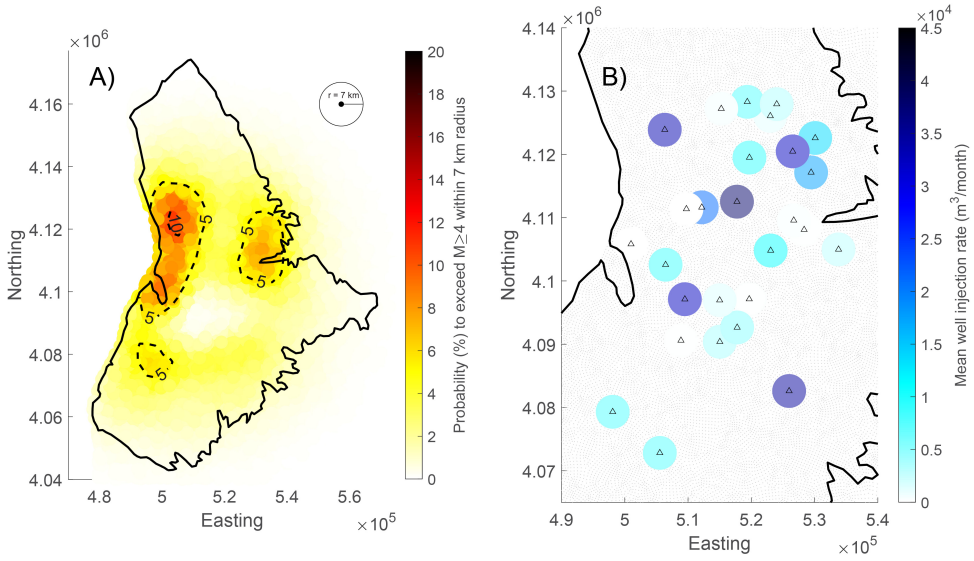


Figure 3.14: **Prospective Case 'Safety' Results.** A) Total probability of exceeding a M≥4+ earthquake across the entire basin during the total 5 year management window. Hazard is spread more evenly throughout the model and in less than the BAU case in areas that contribute to high hazard. B) Mean injection rate in m^3/month at each well location (triangles). There are several locations where the optimization chooses not to inject. The grey dots represent the model nodes.

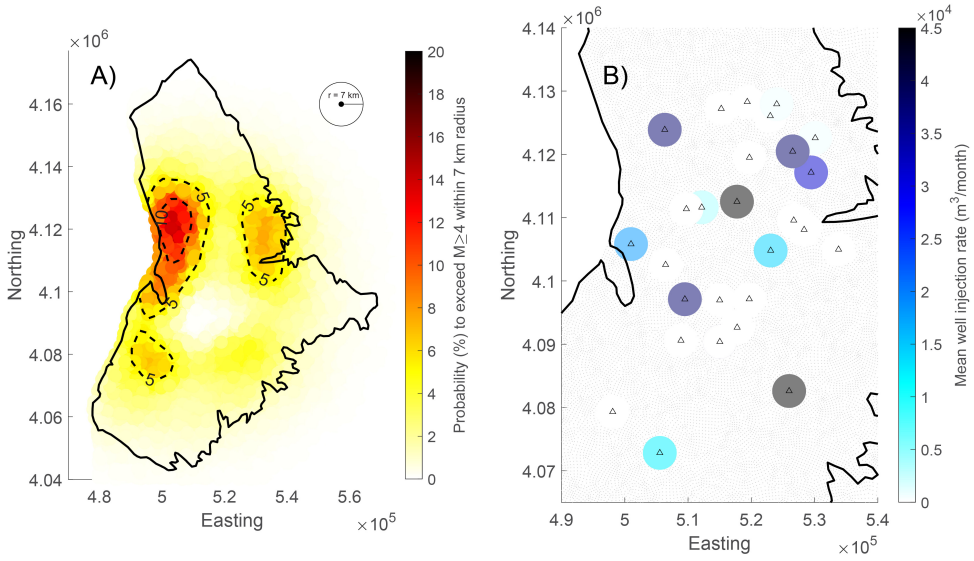


Figure 3.15: **Prospective Case 'Economic' Results.** A) Total probability of exceeding a M≥4+ earthquake across the entire basin during the total 5 year management window. The highest probability western part of the basin is associated with the large fluid injection. B) Mean injection rate in m^3/month at each well location (triangles). There are several locations where the optimization chooses not to inject. The grey dots represent the model nodes.

Appendix

3.A Appendix to Chapter 3

3.B Supplementary

3.B.1 Data

Wastewater injection well data for Las Animas County, Colorado was retrieved from Colorado Oil and Gas Corporation Commission Website (https://ecmc.state.co.us/#/home), (Accessed: 20223-10-10). Wastewater injection well data for Colfax County, New Mexico was retrieved from New Mexico Oil Conservation Division Permitting Website (https://wwwapps.emnrd.nm.gov/OCD/OCDPermitting/Data/Wells.aspx) (Accessed: 20223-10-10). In this study we convert injection well data from bbl/month to m^3/day across 29 wells from Nomber 1994 to May 2022 (See Supplementary Data).

Multiple seismic studies have taken place in the Raton Basin. We leverage these combined data sets to form a comprehensive catalog of earthquakes up to July-2020. Earthquakes from 1963–2013 are given from Rubinstein et al., 2014, which include recorded earthquakes by the USGS temporary seismic networks from 2001-2011 (Rubinstein *et al*., 2014). Earthquakes from 2008-2010 were recorded by the EarthScope Transportable Array (Nakai *et al*., 2017a). Past

2013, we rely on cataloged earthquakes from USGS National Earthquake Information Center (NEIC). Furthermore, from July 2016 to July 2020 Earthquakes are provided from a combined broadband seismometer and geohpone node study available from the International Seismological Centre (Glasgow *et al*., 2021a; Glasgow *et al*., 2021b).

3.B.2 Step Rate Tests

Prior work calibrated reservoir permeability in the main injection reservoirs, the Dakota Formation and the Entrada Formation, from injection-recovery step rate tests (Hernandez and Weingarten, 2019). A step rate test determines how pressures within a formation change as a result of small-scale injection. The pressure changes can be converted to input parameters for AQTESOLV which utilizes a Theis step-drawdown test to approximate hydraulic properties (Duffield, 2007). SM Table 1 provides the permeability values obtained from AQTESOLV for different cases: case 1 considered the lowest values of psi from each step, case 2 considered the highest values of psi, case 3 used an incremental increase per minute, and case 4 was simply the recovery data. Case 4 provided the lowest mean residual for both reservoirs (SM Figure 3.B.11-3.B.12) and was chosen as the preferred permeability for the model. We include plots from within AQTESOLV of the data and transmissivity solution The calculated permeability are within those reported by previous studies (Belitz and Bredehoeft, 1988; Nakai *et al*., 2017b).

3.B.3 Simplified Optimization Example (no SI map required)

There is strong evidence to suggest that stressing rate and accumulated stress, the latter which is related to the total injected volume ($\int_V \Delta\tau(P) \sim \Delta V$) (Van der Elst *et al*., 2016), are key factors that influence the occurrence of induced seismicity (McGarr, 2014; Weingarten *et al*., 2015; Toda *et al*., 2002; Qin *et al*., 2022). As an example, the following management model uses the prior total Coulomb stress τ and Coulomb stress rate $\dot{\tau}$ at locations in the Raton Basin

that were associated with injection induced M≥ 4+ events. We make the assumption that all former M≥ 4+ events occurred at the mean seismogenic depth where model results are output. This management model can be thought of as a retroactive example since we exclusively let the previous stress conditions of past large earthquakes inform the management model solution. Therefore, this method does not require an SI map to forecast the hazard, although the solution to the injection rates q can be used to forward solve the hazard if desired. The following steps describe the methodology, generalized for application to other studies:

1. Resolve the stress and pore pressure spatiotemporal evolution from the numerical domain based on the full well injection history.

2. Record the τ and $\dot{\tau}$ at each M≥4+ earthquake location in the numerical domain during the time step it occurred. These will provide the constraints for the x_τ and $x_{\dot{\tau}}$ respectively.

3. Generate response matrix R_τ and $R_{\dot{\tau}}$ (See Appendix) for both τ and $\dot{\tau}$ then stack them vertically; This requires running Q individual models based on Q wells for the length of the management period desired.

4. Solve the linear program management model: $\begin{bmatrix} R_\tau \\ R_{\dot{\tau}} \end{bmatrix} q \leq \begin{bmatrix} x_\tau \\ x_{\dot{\tau}} \end{bmatrix}$

SM Figure 3.B.13 describes the derived constraints at each of the earthquake locations and the resulting optimization of the τ and $\dot{\tau}$ at each of the locations during the management period. Note that the total Coulomb stress and Coulomb stressing rate thresholds are never exceeded. The cumulative injection rate is also reduced. Another important feature of the optimization is the shape of the τ and $\dot{\tau}$ at each of the locations during the management period. Notice that τ steadily increases and that $\dot{\tau}$ increases near the end. The optimization only considers the 5 year management period, and therefore does not consider what ramping the injection rates and subsequent τ and $\dot{\tau}$ near the end of the management period would do for the months following the management period. We present the solution this way to introduce the response matrix method

and reveal the inherent flaws in the optimization since this exact solution would not be ideal for practical use. However, there are a variety of solutions that makes use of mixed-integer programming to control the behavior of the injection wells to avoid this type of solution which we elaborate on in the main text and incorporate for prospective case 'Reduction'.

3.B.4 Mixed-Integer Programming and Additional Constraints

Monotonic decreasing/increasing is an injection scenario by which the injection for all the wells is only ever decreasing/increasing and never increasing/decreasing. The construction of the mixed-integer R^* matrix for a monotonically decreasing scenario is simple. If we consider q_{jk} to represent the injection rate for well j at managment period k, then for all k the constraint $q_{j,k+1} \leq q_{j,k}$ must be satisfied for monotonically decreasing rates. To ensure that this constraint is met x^* must equal a column vector of zeros with length m, and the integer matrix R^* would contain -1s across the diagonal and 1s offset from the diagonal by the number of wells. Similarly, for the monotonic increasing scenario the constraint that $q_{j,k} \leq q_{j,k+1}$ must be satisfied. To achieve this the integer matrix R^* would contain 1s along the diagonal and -1s offset from the diagonal by the number of wells, but with the important inclusion that the diagonals associated with the last time step at all well locations is 0 because otherwise $q_{j,k} \leq 0$ which would result in zero injection rates for all time. We include the monotonically decreasing constraint for Prospectice case 'Reduction'.

Running average constrains the injection rates to to be equivalent to an average over t management periods such that the constraint $\frac{q_{j,k+1}+q_{j,k+2}+\cdots q_{j,k+t}}{t} \leq q_{j,k}$ is satisfied. The running average is useful if smoothing of injection rates through time is desired. The mixed integer construction still results in a column vector of zeros with length m for x^*. The integer matrix R^* therefore contains diagonal integer values equivalent to $-t$ and t 1s offset from the diagonal by the number of wells times t.

Exclusion of certain wells is another constraint that is necessary for typical injection management practices. The construction of the integer matrix R^* is similar to the monotonic scenario. In order to satisfy the constraint for specific wells such that $q_{j,k} \leq 0$ wells at specified management periods in the R^* matrix are represented with 1s since Eq. (3.8) limits the injection rate q as nonnegative. The combination of monotonic, running average, and exclusion of wells allows for a wide variety of variable injection scenarios that are all possible to optimize for.

Furthermore, uncertainty in the simulation model is also possible to incorporate into the management solution. While not included in this study, the concept is similar to the previous management model controls. For example, in our model of the Raton Basin, if there was significant uncertainty in the fault permeability structure we could recreate an entirely new response matrix based on an altered simulation model where the fault zone permeability in the model was changed. This would require 29 (each well) different unit-source solutions ie. model runs. The newly formed response matrix is appended with the primary response matrix and also the constraint vector is appended. The linear program will find an optimal solution again, but with the inclusion that the uncertainty in permeability is accounted for. Uncertainties in any of the material parameters is accountable for different model realizations which are 'stackable' *ad infinitum*. It is important to note that solving the linear program in this way means that the solution finds the optimal injection solution *to* the uncertainty instead of *with* uncertainty. The only 'free' uncertainty that does not require additional simulation model realizations is that of the fault geometry. Additional τ response matrices are calculable for different receiver fault geometries and concatenated in the same way as any other uncertainties.

3.B.5 Iterative Method

The iteration technique is designed to slowly adjust the rate constraints at the subset of model output locations such that the forward solution of the constraints and subsequent seismicity

rate and seismic hazard across the entire basin arrives at the desired threshold. The technique is not exhaustive or optimized, but was found to work adequately for our efforts.

1. Given the forward solution of rate constraints x from the optimized injection rates q resolve the total seismicity rate and subsequent hazard across the basin. If within the tolerance of the desired threshold finish the iteration. If not within the tolerance of the desired threshold continue to the next step.

2. Find the locations l used in the optimization (ie. the 500 subset of points used in optimization (SM Figure 3.B.15)) that for all time during management period (ie. 5 years) reached their constraints, even one time step.

3. For the specific locations l, increase their constraints (for all time) by a small amount. That is to say use a multiplier that increases the constraint. The amount is based on how far way from the desired solution the current total probability is. If close, then the scaling is low, but if far the scaling can be larger if desired by the user. Otherwise if the probability is too high reduce all constraints by an adjustable percentage.

4. Solve the optimization again with the adjusted constraints which will produce a new q array.

5. Forward solve a solution for the rate constraints x given the new q.

6. If you are incorporating previous remnant stress fields, add those stress rates to x now. This is for Prospective Case #2.

7. Return to step 1.

3.B.6 Supplementary figures for Chapter 3

Table 3.B.1: Permeability Calibration. Calculated permeabilities in each step rate case test for the Dakota and Entrada formations obtained from AQTESOLV (Hernandez, 2020).

	Permeability m^2	
Test Case	Dakota	Entrada
Low Displacement	$6.825 \cdot 10^{-14}$	$5.892 \cdot 10^{-14}$
High Displacement	$6.415 \cdot 10^{-14}$	$6.164 \cdot 10^{-14}$
Increasing Displacement	$6.607 \cdot 10^{-14}$	$5.836 \cdot 10^{-14}$
Recovery	$6.667 \cdot 10^{-14}$	$8.924 \cdot 10^{-14}$

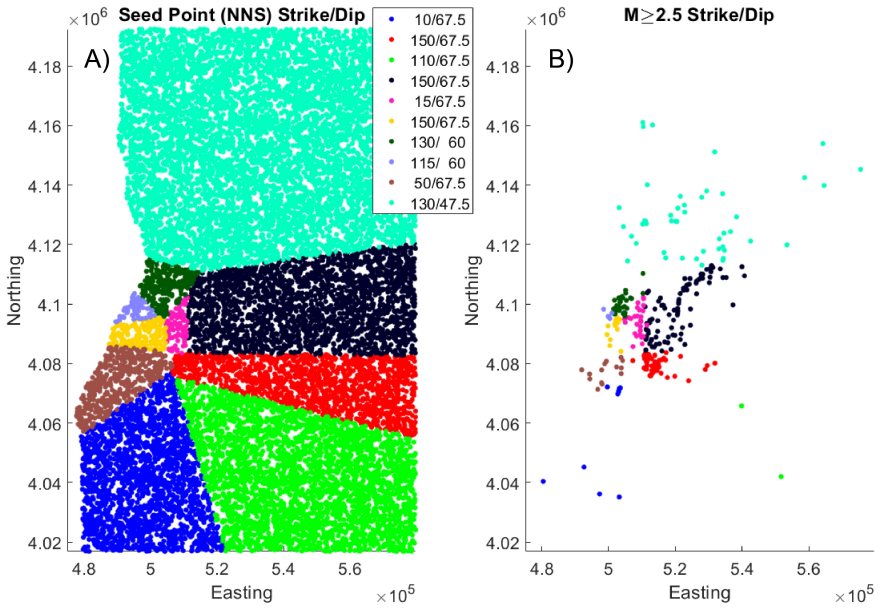


Figure 3.B.1: Map of the ∼25,000 seed points (A) and M≥2.5+ earthquakes (B) with their associated fault geometries (strike/dip). The strikes were determined by a nearest neighbor search (NNS) across the basin by choosing several varying locations of strikes given by previous work (Glasgow *et al*., 2021a) (their Fig. 5). Dips were determined by taking the closest large event focal mechanisms. Regardless of inaccuracies in our fault geometry assumptions, the fault geometries play a minimal role in the overall Coulomb stress rate calculations since the pore pressure rate is the largest component which is independent of fault geometry (SM Figure 3.B.2).

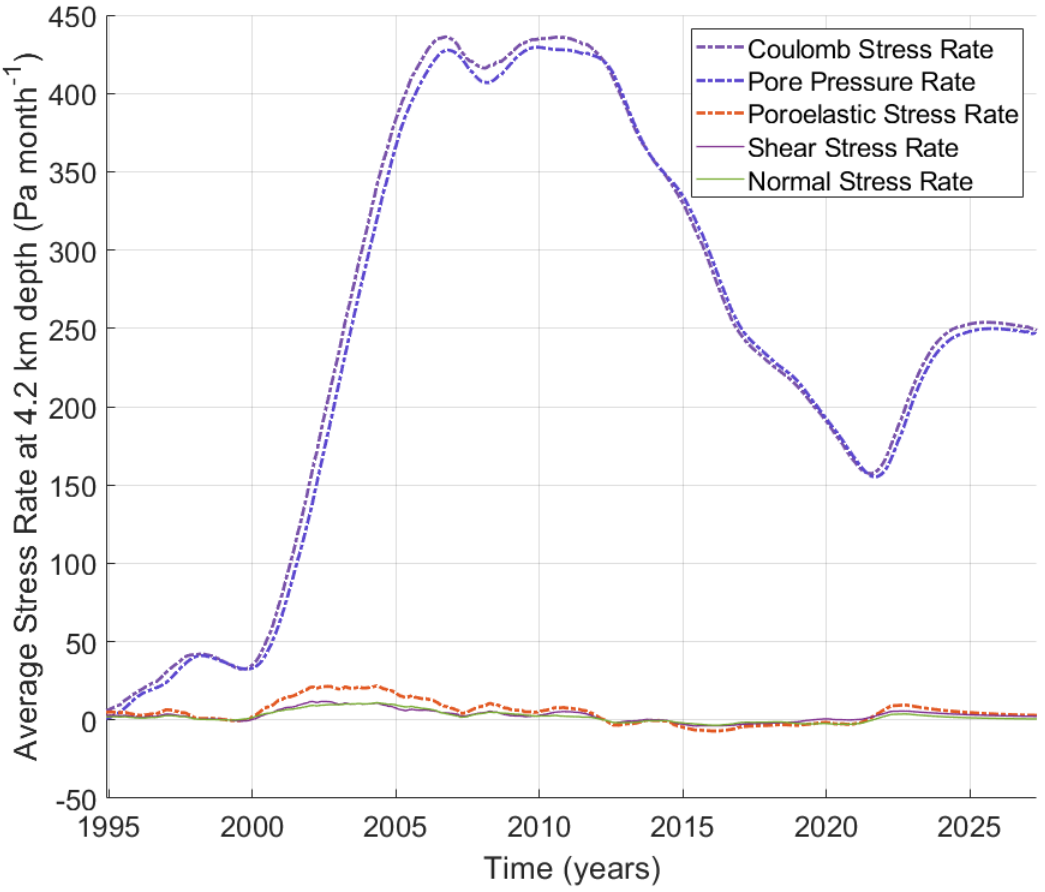


Figure 3.B.2: Average stressing rates for the central Raton basin at 4.2 km depth. The dominant signal of the Coulomb stress rate is the pore pressure rate.

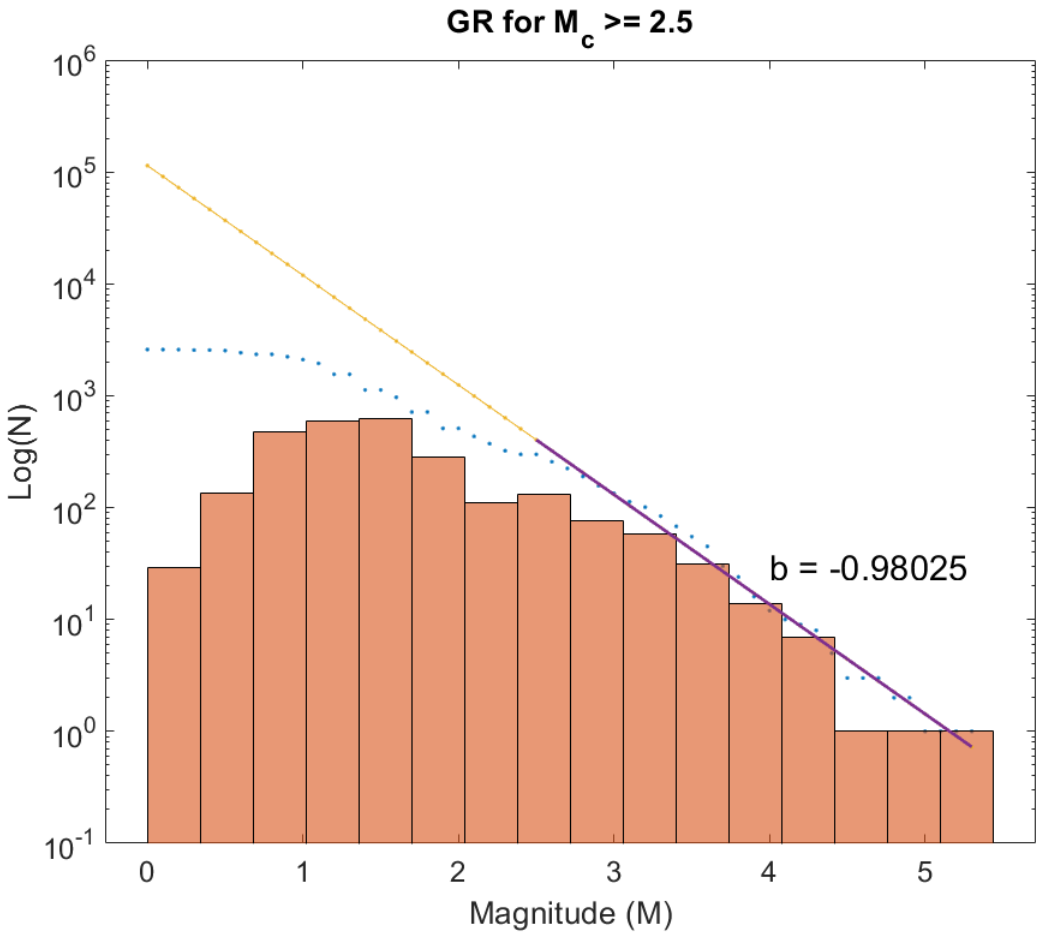


Figure 3.B.3: GR law of earthquake catalog prior to higher resolution data from Glagow et al., 2021 (Glasgow *et al*., 2021a) (see Data). A magnitude cut-off of M_c=2.5 is chosen from visual inspection where the frequency of events experience 'roll-off' from b-value estimate.

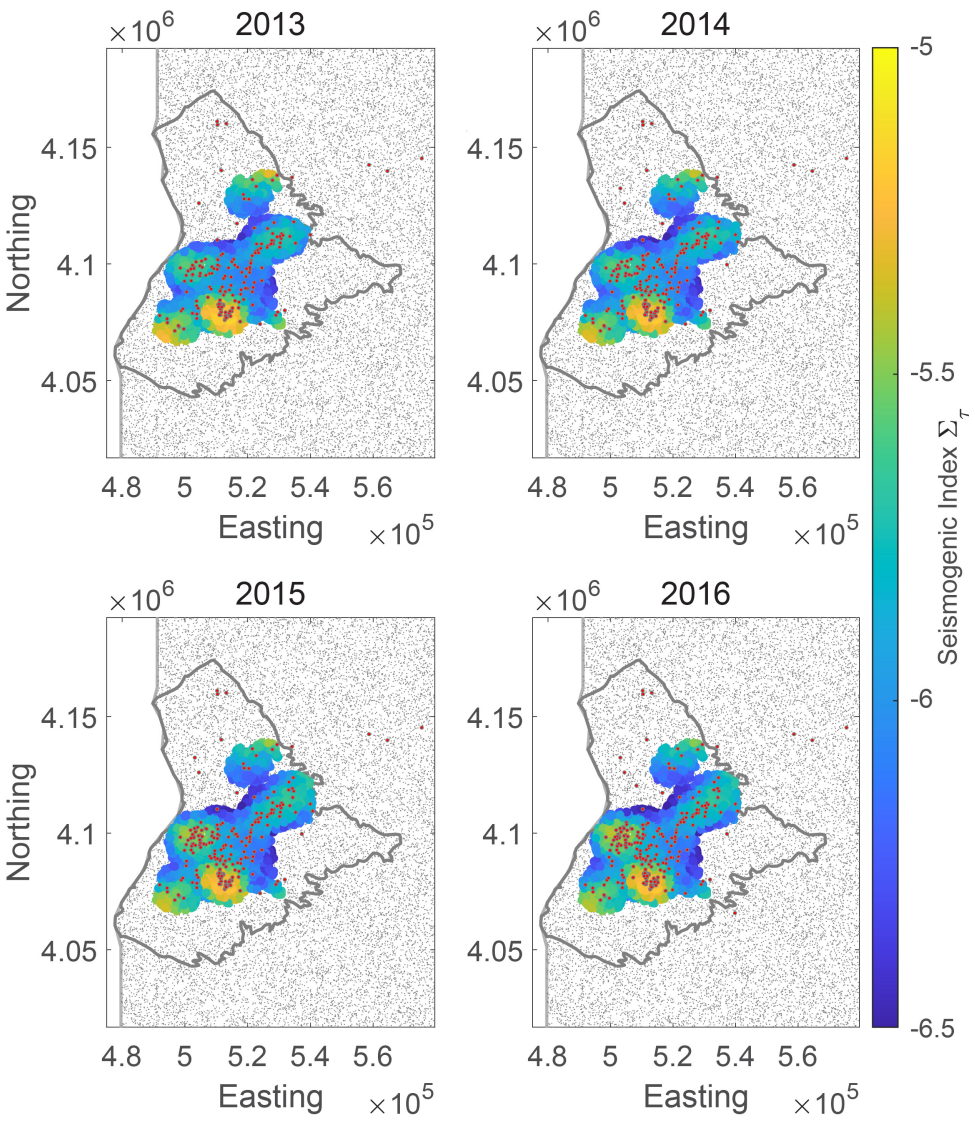


Figure 3.B.4: SI map for varying calibration period 2013-2016. With increasing earthquake count the SI improves in spatial resolution, but there is little change among the different calibration years.

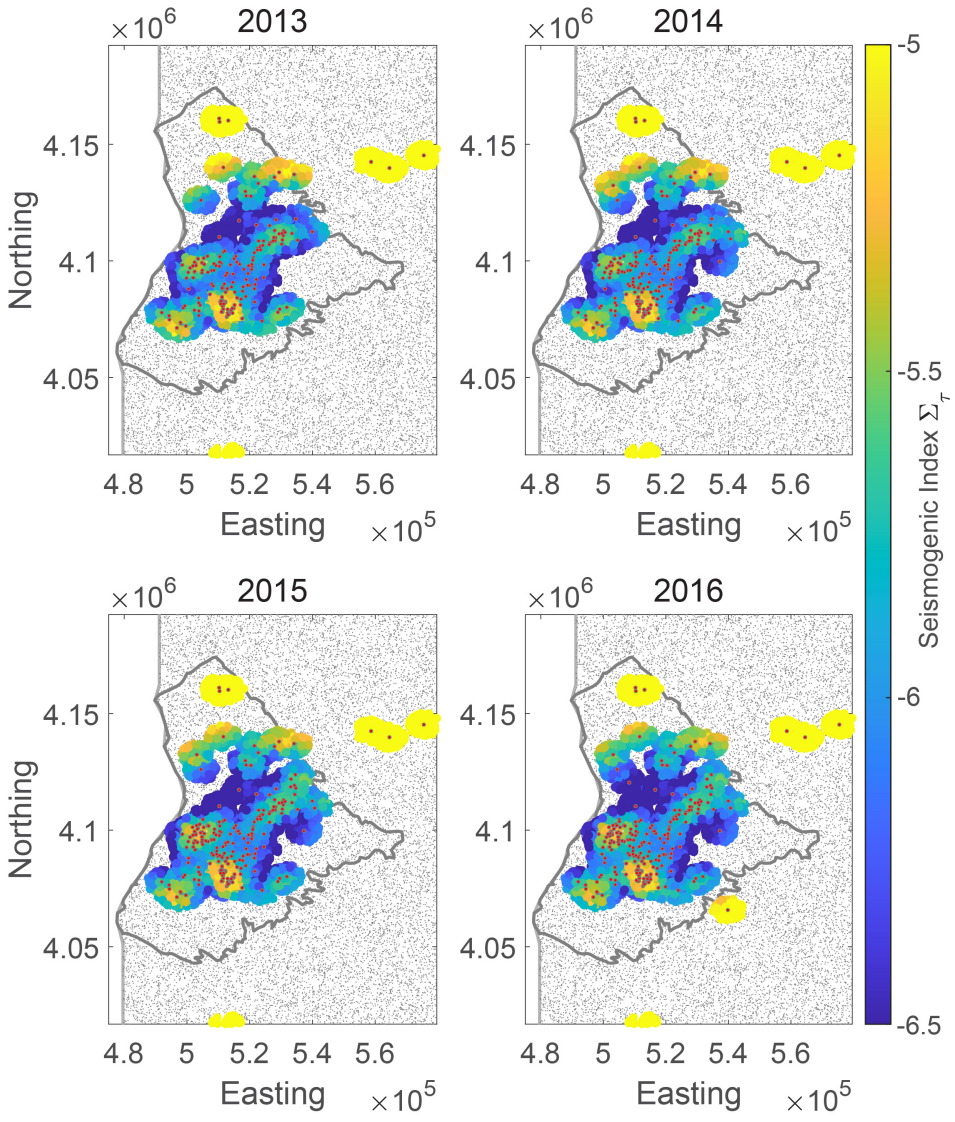


Figure 3.B.5: SI map for varying calibration period 2013-2016 for the model that uses a 5-km search radius and removes the >3 earthquake precondition. Outliers away from the basin provide high localized areas of enhanced SI. Notice that within the basin though, the overall structure and features of enhanced SI do not change.

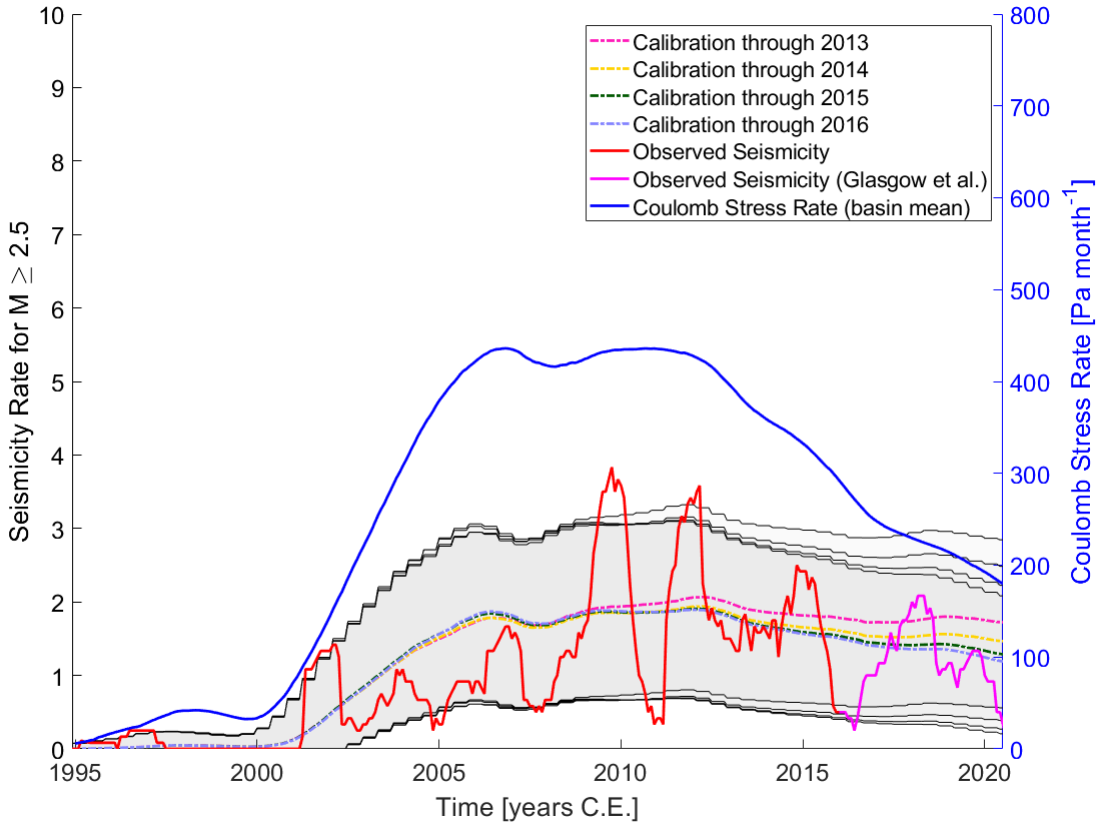


Figure 3.B.6: Seismicity rate forecasts, above our completeness magnitude M≥2.5, compared to observed seismicity rate (1 year moving mean). Calibration period is from Nov 1994 to July 2016, prior to the Glagow et al., 2021 study. The earthquakes and time period used to calibrate the SI model is represented by the red line. The grey areas are the 95% confidence bounds for the different calibration time periods for the the forecasted seismicity rate produced from the SI model that includes the inverse distance weighted interpolation (right panel of Figure 3.4). Magenta line represents the observed seismicity from Glasgow et al., 2021 which is well explained by the seismicity rate forecasted by our model.

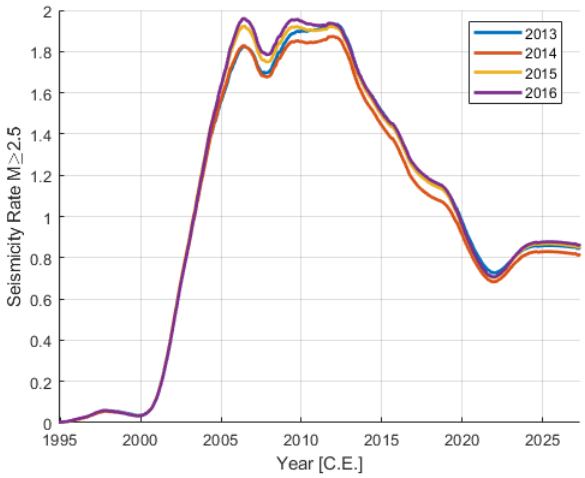


Figure 3.B.7: Seismicity rate for four different calibration periods including the BAU forecast after May 2022. This is the forecast based on our SI model shown in Figure 3.B.4.

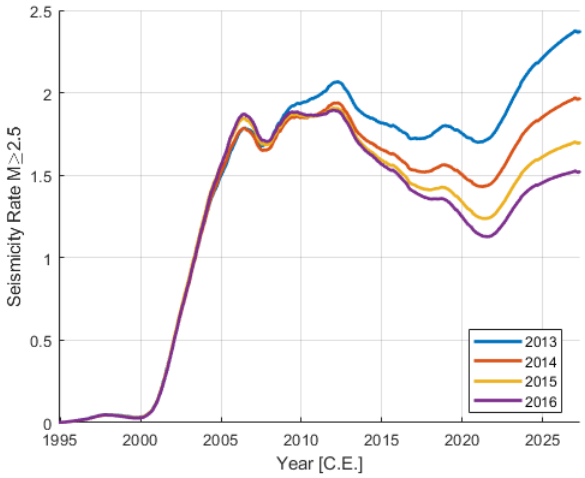


Figure 3.B.8: Seismicity rate for four different calibration periods including the BAU forecast after May 2022. This is the forecast based on our SI model shown in SM Figure 3.B.5. Notice that the seismicity rate increases much more than the prior model in SM Figure 3.B.7. The reason is that the large outliers of SI now experience elevated rates of Coulomb stress rate which contribute to the overall seismicity rate considerably more.

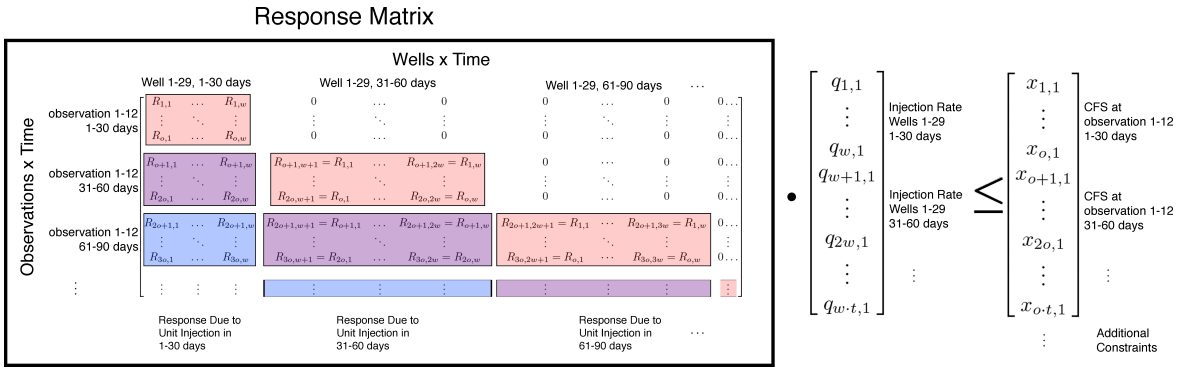


Figure 3.B.9: **Response Matrix.** As an example, we denote the Coulomb stress response matrix as R_{mn}, where m=732 is the number of rows that equals the number of model output locations (12) times the number of time steps (61), and where n=1769 is the number of columns that equals the number of wells (29) times the number of time steps (61). Steps to form the response matrix for the Coulomb stress *rate* are provided in the Appendix. If we denote q as the injection rates at each of the 29 wells for all time steps (61), we can multiply Rq to produce the resulting Coulomb stress at each of the observed locations for each time step. This is the foundation for the management model and linear program optimization. An example of using 12 model output locations is presented in the Supplementary Methods 3.B.3).

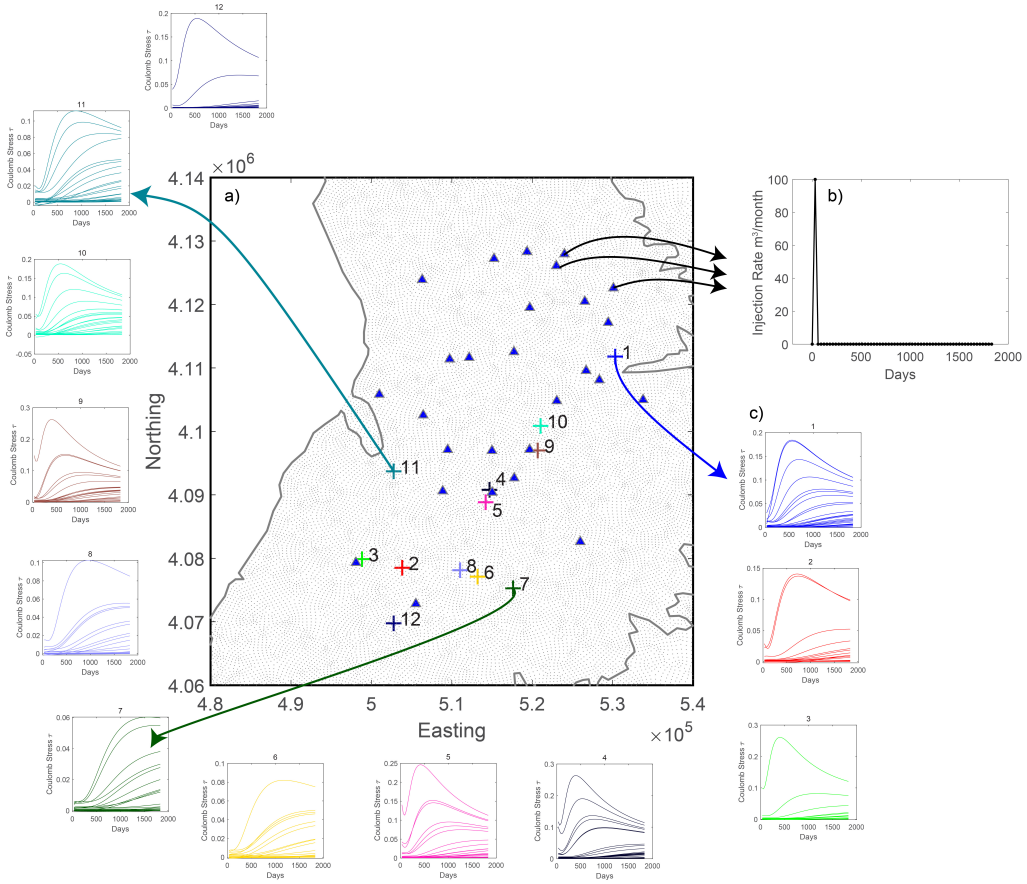


Figure 3.B.10: **Visualization of Response Matrix Generation.** Panel a) shows a zoomed in portion of the model with corresponding well locations (blue triangles) and 12 model output locations where prior M≥4+ events occured in the basin. Panel b) is the unit impulse injection rate. We create 29 separate models that follow this injection profile for each well. The impulse response is an injection of 100 m^3 immediately followed by zero injection rate with no injection at the other well locations. Note that the unit impulse response shares both the number of time steps and total time length of the management model. We then record the response at the entire basin (model points in grey). For this example we choose 12 points associated with prior earthquakes. Panel c) shows the 29 responses that each unit response has on each given location. These response values are combined in the response matrix (Figure 3.B.9).

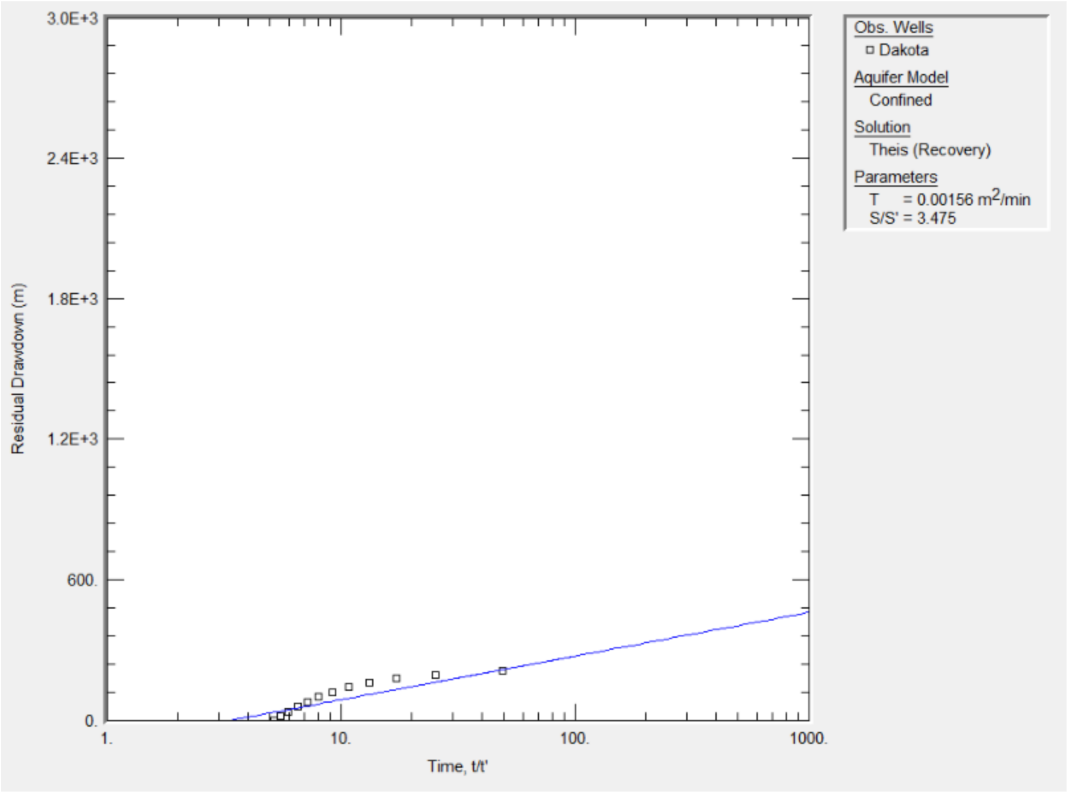


Figure 3.B.11: **Dakota Formation (Recovery).** Dakota formation residual drawdown over $log(t/t')$ calculated fit from AQTESOLV. Transmissivities were converted to permeability (Table 1).

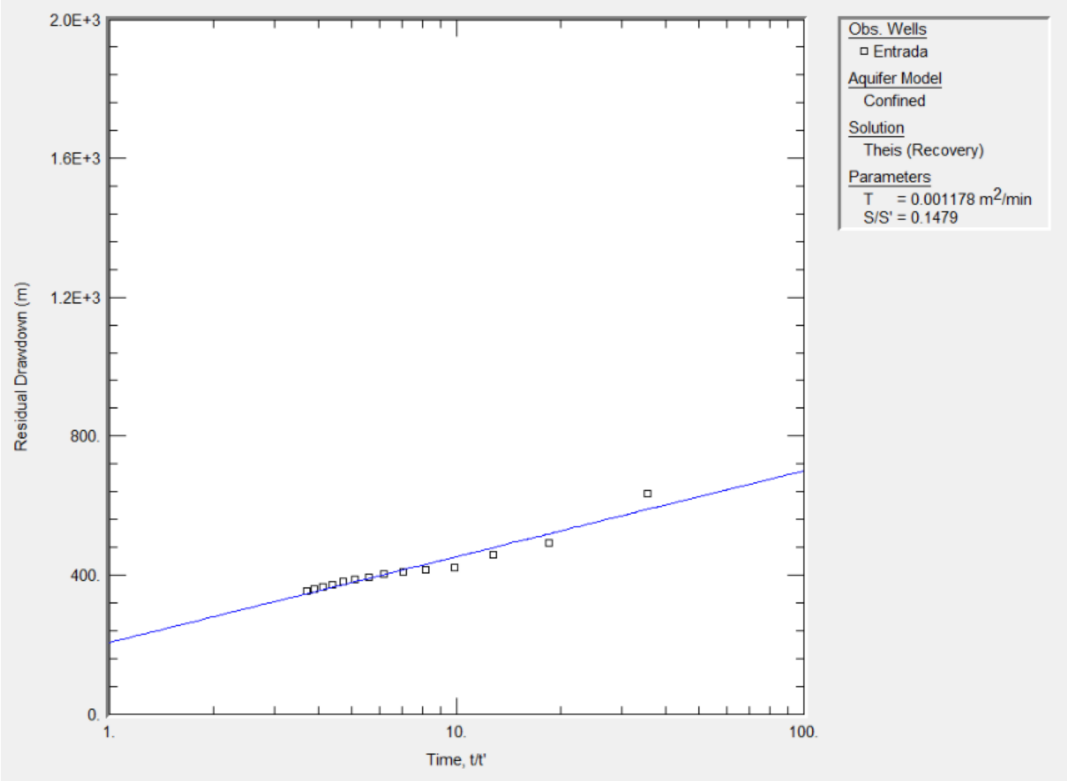


Figure 3.B.12: **Entrada Formation (Recovery).** Entrada formation residual drawdown over $log(t/t')$ calculated fit from AQTESOLV. Transmissivities were converted to permeability (Table 1).

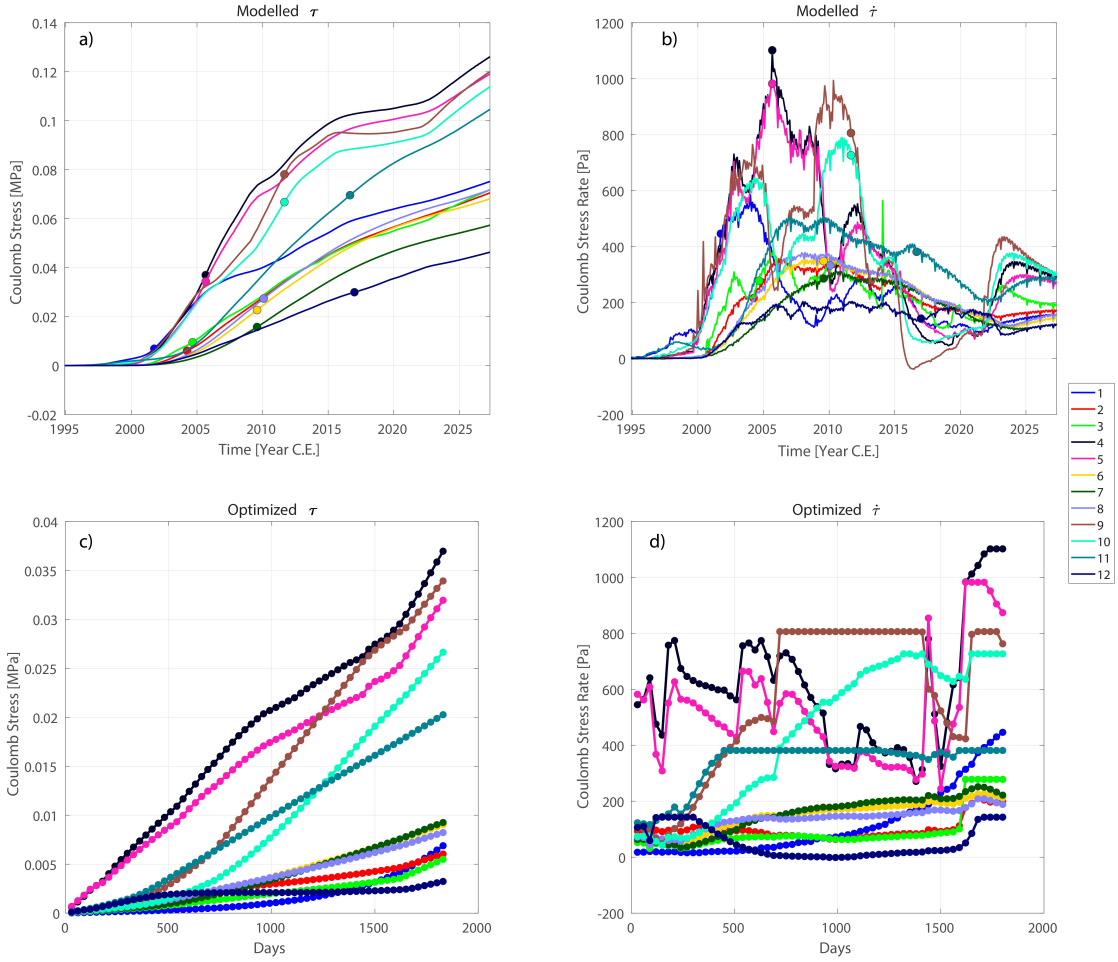


Figure 3.B.13: **Simple Optimization Example (12 points)** Panel a) is the modelled Coulomb stress at the 12 model output points with dots at the time of previously recorded M≥4+ earthquakes at that location. Panel b) is the modelled Coulomb stress rate at the 12 model output points with dots at the time of the M≥4+ earthquake at that location. Notice how the Coulomb stress rate at the model output points 4,5,8,6,7,10 coincide when rates were peaking indicating good, and entirely independent, agreement between Coulomb stress rate and timing of seismicity. Panel c) is the optimized Coulomb stress which is considerably lower than the modelled stress. Panel d) is the optimized Coulomb stress rate. Notice how some locations clearly reach the maximum allowed rate for some time steps. Individual model output locations compared to the overall and rate constraints through time are provided in the Supplementary (SM Figure 3.B.14).

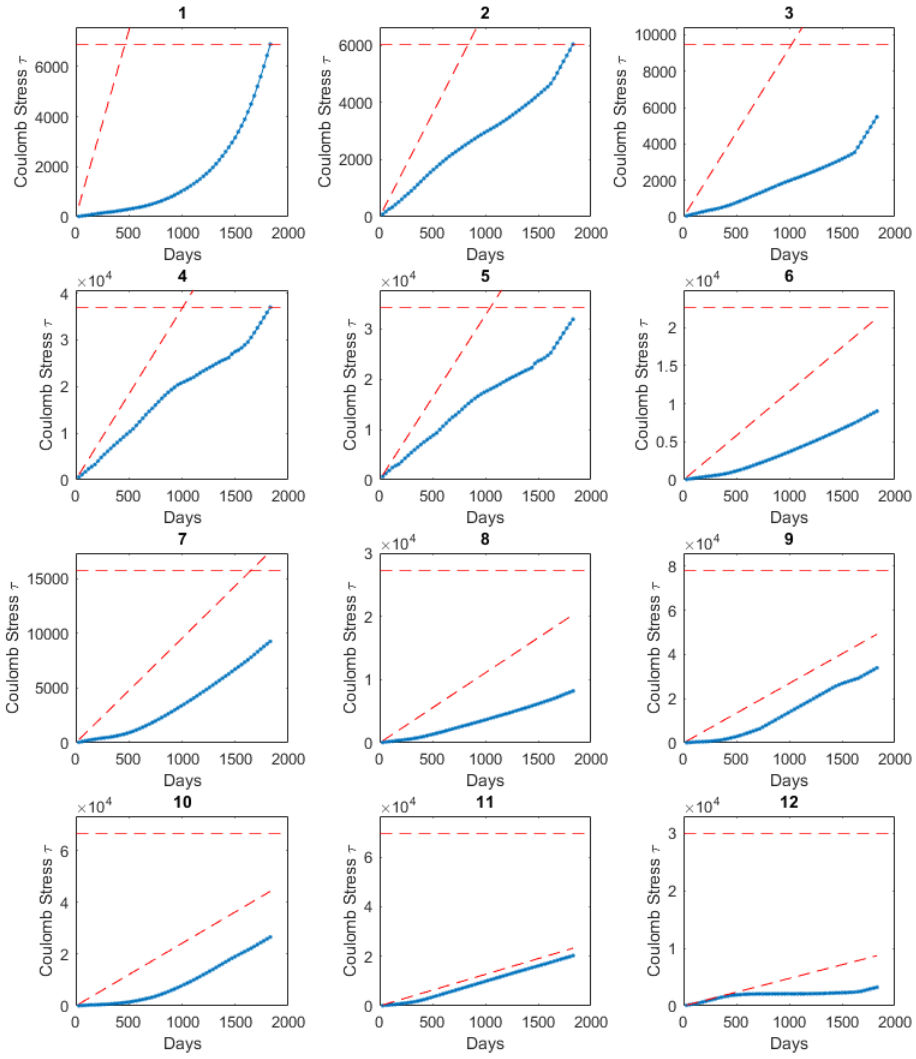


Figure 3.B.14: Simplified management model example results that does not require an SI map. The (blue line) is the optimized Coulomb stress results for the 12 model output locations compared to the maximum Coulomb stress allowed (horizontal red dash line) and compared to the maximum Coulomb stress rate (angled red dash line) allowed at each of the model output locations.

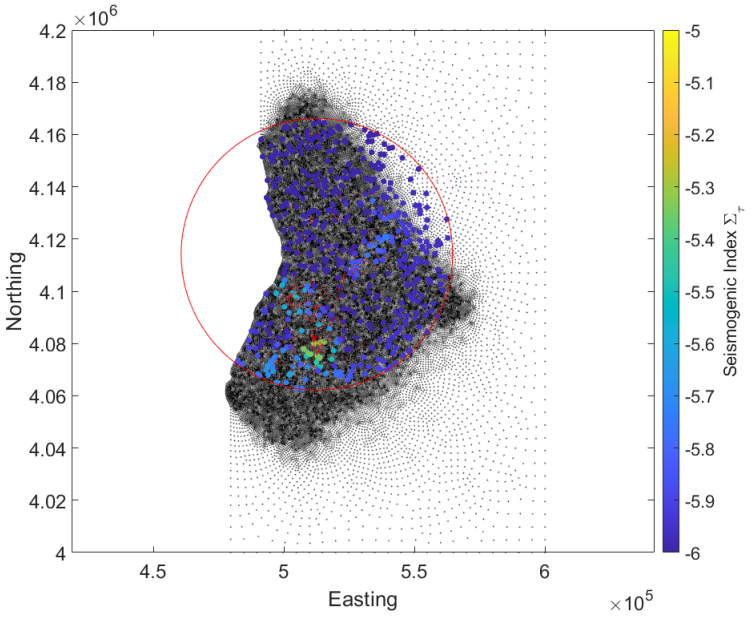


Figure 3.B.15: Uniform random distribution of points (500) used for management model 2 examples. The red dots represent the earthquakes in the basin with M$\geq$2.5. The red circle represents the subset of the model points used such that all seismicity is within it ensuring that the random points chosen for the initialization of the optimization are not irrelevant.

Chapter 4

Deciphering earthquake triggering mechanisms with a fully coupled poroelastic model and machine learning analysis: application to the case of Paradox Valley Unit, Colorado

In areas of induced seismicity, both stress changes from fluid injection and from static deformation caused by slip along faults can trigger earthquakes. Distinguishing the difference between injection-driven independent events and earthquake-driven events will improve our understanding of the spatio-temporal behavior of induced seismicity. Here, we explore earthquake triggering mechanisms by combining (1) a calibrated fully coupled poroelastic model of wastewater injection, (2) static stress from fault dislocations, (3) a machine learning random forest regression model composed of varying stress metrics, (4) Shapley values to determine feature importance, and (5) temporal clustering techniques to validate our results. Our case study

area, the Paradox Valley Unit, Colorado, presents a relatively simple, idealized case to test our methodology: a single, high-pressure injector that has induced >7000 earthquakes between 1991 and 2012 with a detailed earthquake catalog, injection history, and well defined hydrogeologic heterogeneities. Our poroelastic model, built with Abaqus, resolves time dependent stress change at ∼3000 hypocentral locations within 8 km of the well for 284 monthly time steps.

4.1 Introduction

A variety of anthropogenic industrial activities, including wastewater disposal, cause induced seismicity (Ellsworth, 2013; Keranen *et al*., 2014; Shirzaei *et al*., 2016). Similar to natural events, induced seismicity occurs on pre-existing critically stressed faults (Townend and Zoback, 2000). Generating induced seismicity from the reactivation of faults is attributed to several physical mechanisms: increasing pore pressure diffusion (Keranen and Weingarten, 2018; Weingarten *et al*., 2015; Langenbruch *et al*., 2018), poroelastic coupling (Segall and Lu, 2015), and stress changes caused by seismic or aseismic fault slip (Ge and Saar, 2022; Brown and Ge, 2018).

These physical mechanisms for induced seismicity contribute to the triggering potential of each earthquake. Since induced earthquakes can be triggered by small stress changes (1-10 kPa) (Bachmann *et al*., 2012; Cacace *et al*., 2021; Stokes *et al*., 2023) a large difficulty arises in deciphering which mechanism was responsible for triggering each earthquake. We are particularly interested in discerning which earthquakes were driven by fluid stress changes and which earthquakes were driven by stress changes from prior earthquakes. Furthermore, site-to-site differences in physical rock properties, reservoir structure, fault geometry, and remnant tectonic stress could contribute relative differences in mechanism dependencies.

A major difficulty in deciphering induced earthquake triggering mechanisms is uncertainty in the absolute stress within the Earth's crust (Delorey *et al*., 2021). Relative stress changes from

fluid injection require analytical or numerical models to resolve the spatio-temporal evolution of pore pressure and poroelastic stress. To capture the fully-coupled poroelastic stress changes (Biot, 1941; Rice and Cleary, 1976b; Wang, 2000) induced from the fluid sources requires detailed knowledge of the hydrogeologic properties of the region. Point measurements of absolute stress magnitudes are possible from overcoring, hydraulic fracturing, or the focal mechanism stress inversion method (Bredehoeft *et al*., 1976; Fama and Pender, 1980; Tanaka *et al*., 1998; Zoback and Healy, 1992; Gephart and Forsyth, 1984). These measurements allow for estimation of the principal stress directions and typical faulting geometry. The fault geometry is critical for resolving fault plane stress tractions that characterize fault stability (King *et al*., 1994a; Cocco, 2002). Hence, any attempt at discerning induced earthquakes requires an accurate and comprehensive hydrogeological model, detailed injection well data, precise fault geometries, and high-resolution earthquake catalog.

Here, we build a three-dimensional (3D) fully-coupled poroelastic model of Paradox Valley Unit, CO (PVU) and resolve time dependent pore pressure and stress changes due to wastewater injection in an attempt to discern earthquake triggering mechanisms. The primary goal is defining what earthquakes are predominantly triggered by stress changes from the injection and what earthquakes are predominantly triggered by earthquake-earthquake interaction. To inform the contribution of our earthquake triggering mechanisms, we use a random forest regression machine learning analysis and SHapley Additive exPlanations (SHAP), a game theoretic approach to explain the output of any machine learning model (Lundberg and Lee, 2017). We corroborate our results with an independent induced seismicity cluster analysis, which reveals strong evidence that the physics-based machine learning method provides novel insight into discerning triggering mechanism not previously captured. This model explores the induced earthquake triggering process and could help discern what regions are more or less susceptible to stress changes from anthropogenic sources. The methodology presented here is applicable not just for wastewater disposal, but to other types of subsurface injection: CO_2 sequestration, enhanced geothermal

systems, and hydraulic fracturing.

4.2 Paradox Valley Unit (PVU) Data

The PVU is a program run by the U.S. Bureau of Reclamation which has been disposing deep brine into a confined aquifer between 4.3 and 4.6 km depth in Paradox Valley, Colorado since 1995 (Ake *et al*., 2005; Denlinger and RH O'Connell, 2020) (Figure 4.1). The high-pressure fluid injection induced $>$7000 earthquakes between 1991 and 2012, which have all been documented as induced seismicity (Ake *et al*., 2005; Block *et al*., 2015; King *et al*., 2016; Denlinger and RH O'Connell, 2020). Most seismicity within 5 km of the injection well were induced within the first 10 years and nearly all within the high permeability injection reservoir known as the Leadville formation. This zone is highly pressurized from decades of continuous pumping and dictates the lateral migration of seismicity away from the wellbore. These carefully studied events support the notion of a $\sqrt{t}$ diffusion model for pressurization from the well (Block *et al*., 2015; King *et al*., 2016) (Figure 4.2). Additional ancillary data also make this an ideal study region: numerous wells that extend into deeper formations than just the Leadville aquifer, 3D seismic tomography, logs of P-wave velocity, density and porosity from the near surface to basement in the injection well, and logging of geologic units in other wells in the area (Denlinger and RH O'Connell, 2020).

Most importantly, previous work has already compiled a detailed, fully coupled poroelastic model (Denlinger and RH O'Connell, 2020). This model is given by a grid description of nodes with corresponding parameter values which we validate in Abaqus with improved meshing near the well where pressure and stress gradients are highest (see SM 4.A.1). Figure 4.2 includes a plot of the earthquake distribution overlayed on a cross-section of the numerical model mesh.

4.3 Methods

Here, we resolve time dependent pore pressure and stress changes throughout the PVU. The pore pressure and stress perturbations are taken to produce von Mises stress features that are fed into the ML/SHAP analysis. An additional feature, which we call the earthquake feature, is created and also fed into the ML/SHAP analysis. The earthquake feature is calculated from prior earthquakes that may have perturbed the current earthquake. The ML/SHAP analysis seeks to find the model contribution of both the stress and earthquake features at each earthquake for all time. We support our interpretations of triggering mechanisms from the ML/SHAP with results from a nearest neighbor distance cluster analysis.

4.3.1 Numerical Model

We model the relative increase in pore pressure ΔP (scalar) and poroelastic stress ΔS (2nd order tensor) for the PVU using a model with one injection well in the center of the model domain (SM Figure 4.A.1). The hydrogeologic structure is based on a unique nodal distribution of parameters that we reduced down to 1000 unique unit formations and use Abaqus to resolve the linear poroelastic equations (Hill, 2021) (see SM 4.A.1). The model dimensions are 50 km by 50 km with a 18 km depth. Figure 4.2 shows a cross-section through the well injection zone. The injection is divided across three perforated zones consistent with prior modeling and uses the entire injection history as 7952 unique daily rates in our model from 10-July-1991 to 16-April-2013 (Denlinger and RH O'Connell, 2020) (Figure 4.2). We output ΔP and ΔS from these daily steps across the entire domain at 284 monthly time steps. We do not include earthquakes in our study that occur outside of our model domain time despite the earthquake catalog extending until 31-December-2019 (Figure 4.2).

4.3.2 Stress Features

The Abaqus outputs of ΔP and ΔS were post-processed in Matlab using abaqus2matlab (Papazafeiropoulos *et al*., 2017). The stress features of ΔP and ΔS represent the relative change induced from the fluid injection and are resolved at the closest value in the domain to each $\sim$3000 earthquakes during our study time. The stress features are direct input parameters in the ML/SHAP analysis and therefore we looked at a variety of different stress features during the preliminary stages of this work consistent with prior forecasting methodologies (DeVries *et al*., 2018; Sharma *et al*., 2020; Qin *et al*., 2022) (see Supplementary). We found that von Mises stress and von Mises stress rate were often the best stress features for forecasting the seismicity rate and are the only two stress features we consider in our current model. We make the assumption that the von Mises stress is resolved uniformly using a strike azimuth of 260^0 and vertical dip consistent with the most common faulting structure present from the earthquakes locations (Denlinger and RH O'Connell, 2020).

4.3.3 Earthquake Feature

Static stress transfer modeling can successfully resolve stress transfer between faults in an elastic half space with homogeneous isotropic elastic properties (Lin and Stein, 2004; Toda *et al*., 2005). Stress transfer can promote or reduce the potential of earthquake triggering, depending on the coefficient of friction, fault geometry, and sense of slip (King *et al*., 1994a; Stein, 1999). Since the exact geometries of every earthquake in our model are unknown, we choose to develop an earthquake feature that is based on prior perturbable earthquakes.

We use 'cutde' (Thompson, 2021) to resolve stress transfer produced from fullspace triangle dislocation elements (Nikkhoo and Walter, 2015). Several assumptions are required for the static stress transfer modeling: (1) We assume a uniform stress drop for every event of 3 MPa, (2) a shear modulus of 30 GPa, and (3) a Poisson ratio of 0.25. Under this framework we

show that the von Mises stress is self similar for both parallel and perpendicular receiver receiver planes at a given distance from the event (SM Figure 4.A.11). By varying event magnitude, we calculate a radius from the center of the dislocation that can increase the potential of failure up to a distance that intersects the 10 kPa triggering threshold (Reasenberg and Simpson, 1992; Stein, 1999). Then, for every earthquake, we create an earthquake-to-earthquake feature, which counts the number of earthquakes that could have perturbed it. The earthquake count is represented by $\ln(N+1)$, where N is the number of perturbable earthquakes to have occurred prior to each event. Higher N indicates a higher likelihood of earthquake-earthquake interaction.

4.3.4 ML/SHAP Analysis

We use the machine learning technique of random forest regression (RFR) to fit our observed seismicity (Ho *et al*., 1995; Ho, 1998). The RFR model makes a prediction on the target variable, which are one-hot encoded occurrences of the observed earthquakes. Contrary to typical applications of machine learning , we do not split our dataset into training and test sets. Instead, we fit the entire data set since we are interested in the feature importance during each event.

We use a set of input features from both the stress and earthquake feature as well as their time lags. The time lags are introduced to capture any potential anisotropy or hydromechanical heterogeneity that the numerical model is not capturing or time delayed effects that former earthquakes may have when perturbing the current earthquake. We avoid overfitting and optimize hyper-parameters by using 5-fold cross validation. We find that including more lags improves the overall fit of our model, up to $\sim$50 lags, but is likely over-fitting and unrealistic. We assume that the physical meaning of the lags are unreasonable beyond $\sim$1 year before the actual earthquake timing and reserve our total lags to the local minimum of 5 lags (SM Figure 4.A.12). In other words, a model can contain the current stress/earthquake feature (+0 lag), the time period prior (+1 lag), and the time periods before that (+2-+5 lag etc..) or any combination of that set (SM

Figure 4.A.12).

To assess our feature-importance we avoid permutation-based feature importance in random forest analyses. This method randomly shuffles each feature and computes the model's change in performance with the features that impact the performance the most having higher importance. Instead, we use SHAP, which explains the predictions of our target variable (earthquake or no earthquake) by computing the contribution of each feature to the prediction (Shapley *et al.*, 1953; Lundberg and Lee, 2017). The major difference between permutation based and SHAP based feature importance is that SHAP represents an additive feature attribution method and quantifies the magnitude of feature attributions. The amount that a feature i contributes to the model prediction is a linear function of binary variables given as:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i \tag{4.1}$$

Where $z' \in \{0,1\}^M$ and M is the number of simplified input features and $\phi_i \in \Re$. Here, f is the original prediction model to be explained by the explanation model g. An effect ϕ_i matches to each feature, and the sum of all feature attributions approximates the output $f(x)$ of the original model. Since the Shapely value is represented as an additive feature, it is a linear model and the contributions of each feature can be added to describe the contribution that the stress features have compared to the earthquake features. This is preferred compared to permutation feature importance which chooses importance based on the decrease in model performance.

4.3.5 Cluster Analysis

As an independent test of earthquake behavior, we investigate how the PVU seismicity is distributed in magnitude, space, and time using a traditional cluster analysis. We use the nearest neighbor distance (NND) space-time-magnitude domain (Baiesi and Paczuski, 2004) for each

pair of events i and j using the following equation:

$$\eta_{ij} = \begin{cases} t_{ij}(r_{ij})^d 10^{-bm_i}, & t_{ij} > 0; \\ \infty, & t_{ij} \leq 0 \end{cases} \tag{4.2}$$

Where, t_{ij} is the interevent time (year), r_{ij} is the inter event distance (km), d is the dimension of the earthquake hypocenter distribution (1.6), b is the b-value (1.0), and m_i is the ith event magnitude (Zaliapin and Ben-Zion, 2013; Schoenball *et al.*, 2015). The NND is separable into rescaled distance (R_{ij}) and rescaled time (T_{ij}) where (Zaliapin *et al.*, 2008; Zaliapin and Ben-Zion, 2013):

$$\eta_{ij} = R_{ij} T_{ij} \tag{4.3}$$

$$R_{ij} = (r_{ij})^d 10^{-bm_i/2} \tag{4.4}$$

$$T_{ij} = (r_{ij})^d 10^{-bm_i/2}, \tag{4.5}$$

An advantage of this form of NND is that the clustering style of seismicity can be displayed by a joint 2D distribution of rescaled time $\log_{10} T_{ij}$ and rescaled distance $\log_{10} R_{ij}$ (Zaliapin *et al.*, 2008; Zaliapin and Ben-Zion, 2013; Zaliapin and Ben-Zion, 2016). The distribution helps to describe the type of earthquake clustering style since observed seismicity often shows a bimodal joint distribution divided by a constant line and chosen nearest-neighbor threshold n_0. Events below this threshold are classified as clustered (i.e. earthquake-earthquake) and the events that are above this threshold are classified as background (i.e. stress-driven or independent) (Zaliapin and Ben-Zion, 2016). We use the NND distributions for the PVU as an independent test of the physical mechanism driving each earthquake in the sequence. We hypothesize that our ML/SHAP model will preferentially separate stress-driven vs earthquake-earthquake driven events as identified by Zaliapin and Ben-Zion (2016).

4.4 Results

4.4.1 Numerical Model Results

The fully-coupled poroelastic model shows that areas with seismicity experience pore pressure increase from 0.005 MPa to 9 Mpa. Most pore pressure increase occurs within within an 8 km radius around the injection well (SM Figures 4.A.5-4.A.10). Most seismicity occurs in close vicinity of the injection well and the ΔP is highest in early 1999 ($\sim$9 MPa). The pressure changes near the well mimic injection rate changes as the temporal delay of diffusion is negligible. Elsewhere, the diffusion process dominates the pressure changes and therefore the increase in pore pressure is more gradual through time (SM Figure 4.A.8-4.A.9). Across the domain, seismicity occurs during the highest rates of pressure increase. This observation is consistent with other instances of wastewater induced seismicity (Langenbruch *et al.*, 2018; Qin *et al.*, 2022). The increasing pore pressure diffuses laterally through the highly permeable Leadville formation. Low permeability confining units above and below the reservoir restrict vertical pressure migration (SM Video 1).

4.4.2 Cluster Analysis Results

Results of the NND cluster analysis show that a larger portion of the earthquakes are classified as the clustered mode (Zaliapin and Ben-Zion, 2016; Goebel *et al.*, 2019)(Figure 4.3). The constant threshold value $\eta_0 = -5.4$ is chosen based on a 1D Gaussian mixture model analysis (Zaliapin *et al.*, 2008; Zaliapin and Ben-Zion, 2016). The clustering behavior is similar to other cases of wastewater induced seismicity (Zaliapin and Ben-Zion, 2016; Glasgow *et al.*, 2021a). The cluster distribution is marginally bimodal with the dominant clustered events occurring at short space-time distances. These results are also dissimilar from other cases of induced seismicity where background-to-cluster proportions are higher, albeit different mechanical processes are

occurring (e.g., The Geysers, Zaliapin and Ben-Zion, 2016). A large portion of the background domain is characterized by low R_{ij} and large T_{ij}, which often characterizes these events as repeaters (Zaliapin and Ben-Zion, 2016). These events make sense in the context of single well injection. The cyclical nature of the injection means repetitive changes in stress occur at the same locations. This is observed in the pore pressure results at different clusters near the well where the pore pressure closely follows the flux of the injection (SM Figures 4.A.5-4.A.8).

4.4.3 ML/SHAP Model Results

Our best-fit model uses the following: 1000 total trees, 10 tree maximum depth, 10 minimum sample split, and 4 minimum samples for a leaf node. The best-fit model was decided from a 5-fold cross-validation analysis grid searched over the hyperparameters. Figure 4.3 shows the fit of our random forest model for two different model types. One model uses only the von Mises stress rate and earthquake feature while the other model uses both the von Mises stress and the von Mises stress rate as well as the earthquake feature (including lags). We find that the mean squared error (MSE) is slightly lower for the model that includes both stress features. However, we choose to present the parsimonious solution of one stress feature and refer the reader to the supplementary for the results including both stress features, which contains small differences to the main results (SM Figures 4.A.16-4.A.19).

The SHAP analysis results are summarized in SM Figure 4.A.13. We output the results exclusively at the time when the earthquakes occur since we are only interested in discerning the contribution of the stress features at that time. A summary of the SHAP contributions for all time, not just when the earthquakes occur, is presented in the supplementary material (SM Figure 4.A.14). The feature with the higher overall impact on the model is the perturbable earthquake feature. This feature represents the number of earthquakes that occurred during the chosen time step that potentially perturbed the earthquake in question. The next most important features,

with nearly equal importance, are the lagged von Mises stress rate. These stress features are considerably less important on average compared with the earthquake feature.

To assess the total contribution of the stress features vs the earthquake features, we compare the cumulative feature results. Separating which earthquakes are dominated by cumulative feature importance, SM Figure 4.A.15 shows that the ratio of earthquakes that have a higher stress feature contribution compared to earthquakes that have a higher total earthquake feature contribution is about 1:5. We examined the sensitivity of this since it would be expected that increasing lags may contribute to higher contribution to stress. While the stress contribution does increase for models that include 0,+1,+2 lags, after the model reaches +3 lags, earthquakes that are considered to have a higher total stress contribution increase marginally. For example, from +3 lags to +5 lags the ratio has a percent increase of only $\sim$0.5% (SM Figure 4.A.20). We do not pursue sensitivity past +5 lags as the SHAP analysis is computationally expensive with increasing features. It is important that when testing increasing lag sensitivity the ratio of the total number of stress features to earthquake features remains the same.

4.5 Discussion

We compare the earthquakes total stress feature contribution percent to the NND (η_{ij}) in Figure 4.4. We find that earthquakes the ML/SHAP model identify as having a $>$50% stress feature contribution tend to be events characterized as background mode in the NND model. Furthermore, in the NND model, the background events are mostly the independent Poisson mode (Zaliapin *et al.*, 2008; Zaliapin and Ben-Zion, 2016). This comparison shows many of the stress driven earthquakes behave as parent earthquakes that were likely induced by pore pressure and stress changes first that then trigger further seismicity. We compare the earthquake distributions that are $>$50% stress feature contribution compared to the overall seismicity to confirm whether these two sets of values are from the same distribution. We perform a two sample

Kolmogorov–Smirnov test and reject the null hypothesis that the two distributions come from the same distribution with 99% confidence (SM Figure 4.A.21).

We also compare the spatial distribution of the earthquakes that have a >50% stress feature contribution (Figure 4.5). We find that many of the stress driven earthquakes are clustered near the injection well. This would be expected, since these earthquakes are occurring at sharp changes in the stress changes produced in the immediate vicinity of the well. There are some clusters away from the well that show large amounts of seismicity, but only a few stress dominated earthquakes. We would expect that these earthquakes would occur earlier in the onset of each cluster if they were to cause the subsequent triggering of other offspring earthquakes. Often, clusters do behave in this manner. This observation is consistent with the machine learning process since earthquakes that had no prior earthquakes would not be expected to have a strong prior earthquake feature contribution. However, some stress dominated earthquakes do not occur earlier than their surrounding earthquakes as well as within a cluster there is sometimes more than one stress driven earthquake. Some areas that are not entirely discernible as clusters appear to behave as multiple stress driven earthquakes.

Results of this study indicate that the physics-based model combined with the machine learning and SHAP analysis can discern a large portion of the independent background mode events identified by the NND cluster analysis. These earthquakes are mostly parent earthquakes that result in further seismicity in either the same area or help start sequences of clustered events. This result is critical for future wastewater practices. If an injection produces seismicity driven dominantly by earthquake-earthquake interaction that area may be considered as highly triggerable and less susceptible to control via well operation best practices. On the contrary, if most earthquakes are stress driven, then the site may be more susceptible to well operational control.

4.6 Conclusion

We combined a 3D fully-coupled poroelastic model of the PVU with a RFR and SHAP analysis to discern earthquake triggering mechanisms. We compared these results with NND cluster analysis to support our claims of earthquakes driven by the pore pressure and stress perturbations induced from the injection well as opposed to earthquakes driven by static stress transfer of prior earthquakes. The novel method, which incorporates game theory methods not previously explored, reveals that there is good agreement between cluster style and stress contribution from the SHAP analysis. Our results indicate that approximately $\sim$20% of the earthquakes are driven by stress changes from the well while the remaining earthquakes are driven by static stress changes from prior earthquakes. We expect this ratio of injection-driven vs earthquake-earthquake driven seismicity to vary by geologic region, stress state, distribution of preexisting faults, and injection style. Deciphering the earthquake triggering process at candidate injection sites will discern good candidates for a variety of anthropogenic activities as areas that produce many triggered clusters from induced industrial activities could make worse candidates. We foresee this method applied to data with higher temporal and spatial resolution could improve the RFR and SHAP results.

Acknowledgements

Chapter 4, in full, is currently being prepared for submission for publication of the material. Hill, R.G., Trugman, D., Weingarten, M., 2023. Deciphering earthquake triggering mechanisms with a fully coupled poroelastic model and machine learning analysis: application to the case of Paradox Valley Unit, Colorado. Geophysical Research Letters, in prep. The dissertation author was the primary investigator and author of this paper.

Tables and Figures

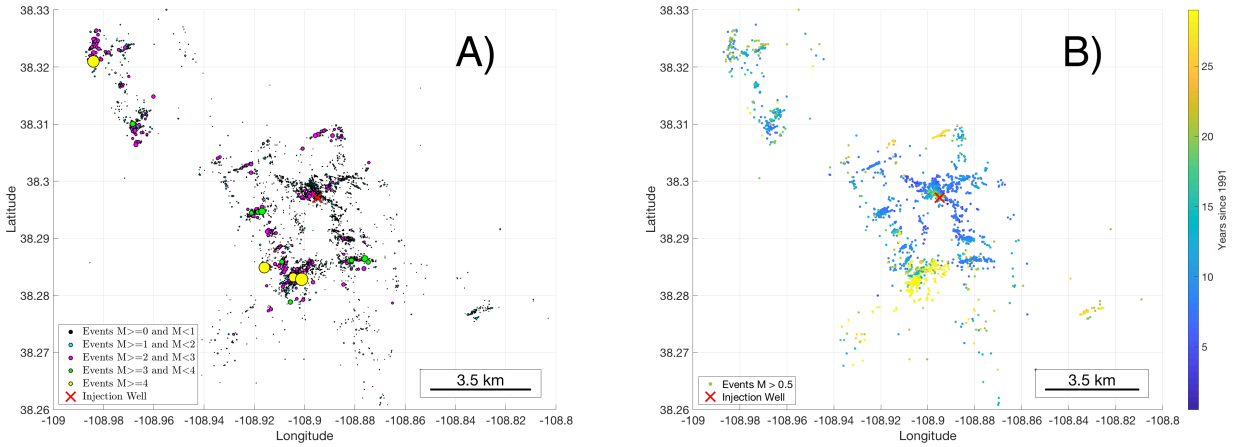


Figure 4.1: A) Regional setting of the Paradox Valley Unit, CO (PVU). The well is denoted by a red 'X' on the map. The deep brine injection began in 1991 at a depth of 4.3 km. Most seismicity is clustered near the well where stress perturbations are largest and fluctuate the most. B) Same view as (A), but temporal timing of events. There are more than 7000 earthquakes in the catalog, but within the 8 km radius around the well which we use for analysis includes only 3000.

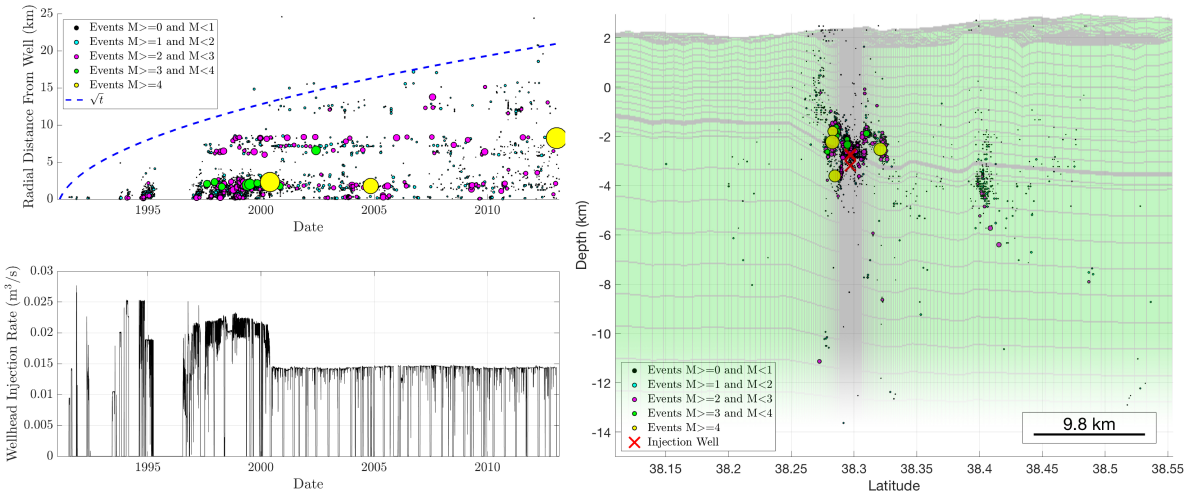


Figure 4.2: Left) Earthquakes plotted as their radial distance from the well and time. Most earthquakes behave in a typical $\sqrt{(t)}$ diffusion rate away from the well consistent with progressive lateral migration of seismicity through the permeable Leadville (Ake *et al.*, 2005; Block *et al.*, 2015; Denlinger and RH O'Connell, 2020). Flow tests were performed prior to 1995. Notice injection is highest during peak injection rates $\sim$1997. Our model records pore pressure and stress perturbations from 10-July-1991 to 16-April-2013. Right) Numerical model cross section with earthquake and well depth superimposed. The model a fully-coupled poroelastic model based on prior work (Denlinger and RH O'Connell, 2020). We increase the grid discretization near the well to capture large changes in pressure gradients (see Supplementary Methods).

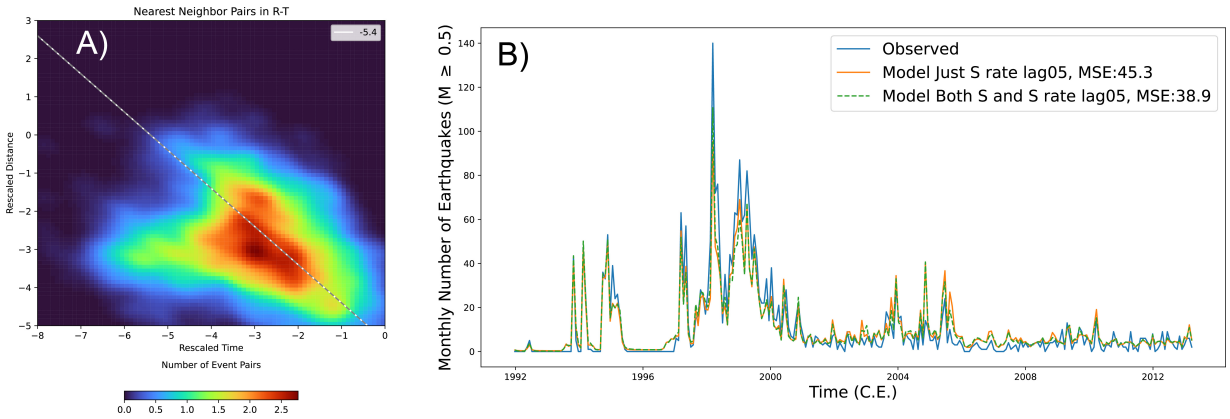


Figure 4.3: A) Nearest neighbor time-distance distributions for the seismicity of the PVU. The color bar represents the number of event pairs. The total number of earthquakes used in this analysis is 2927. The diagonal dashed line is the η_0 background (above) and clustered (below) mode threshold. The value is a constant value determined by the 1D Gaussian mixture model and is -5.4. B) Forecasted seismicity rate across for all time steps. Orange line represents the best fit model that includes only the von-Mises stress rate. The green line includes von-Mises stress and has slightly better fit.

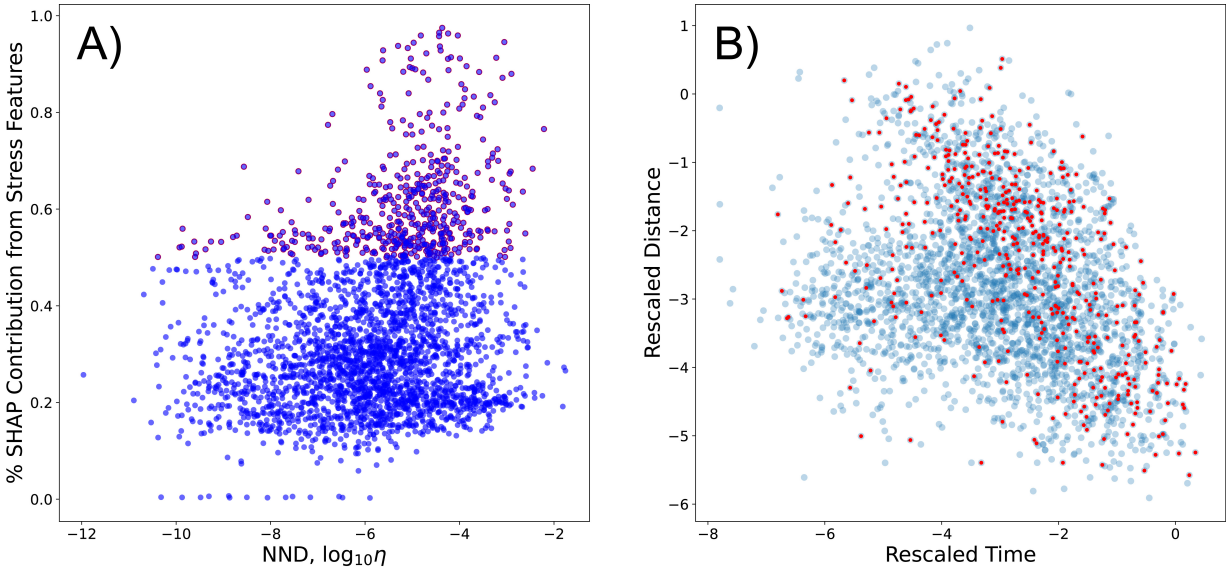


Figure 4.4: A) The SHAP stress feature contribution vs. the nearest neighbor distance value. Many of the earthquakes cluster below the 50% stress feature contribution indicating and to the left of the -5.4 cluster threshold. However, earthquakes that have >50% stress feature contribution, denoted as red circles on both panels, tends to fall on the 'background' mode of the NND (to the right of -5.4). These results are consistent with what we might expect for earthquakes driven by stress from the injection since they would act as initial parent earthquakes that might trigger subsequent seismicity in a region that has experienced stress changes high enough to begin seismicity. B) Comparing the earthquakes that have a >50% stress feature contribution on the rescaled distance rescaled time plot. Many of the earthquakes cluster in the independent background mode with a second distribution towards the repeater mode and a few earthquakes spread out in the cluster mode.

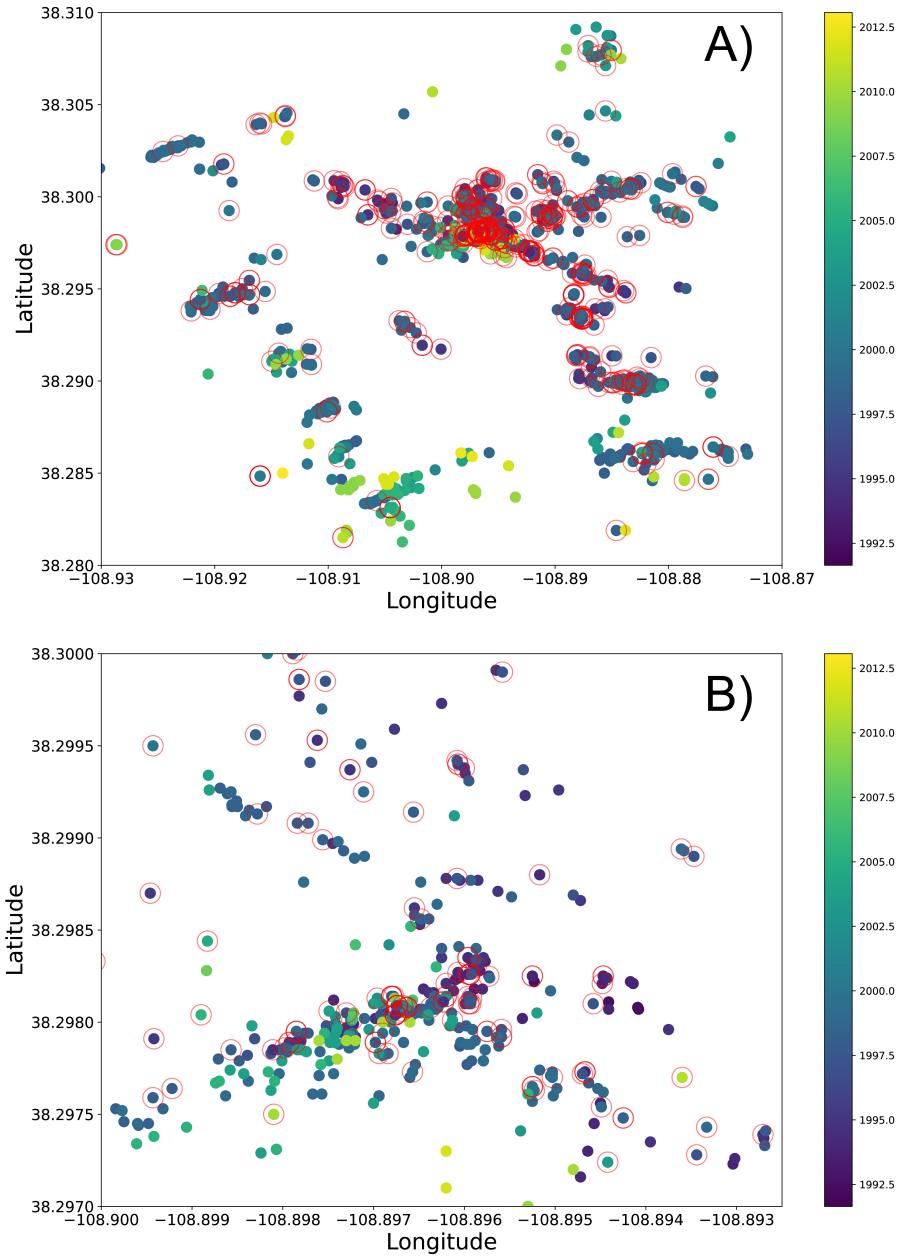


Figure 4.5: A) Map view of most earthquakes used in our study and denoted in color by the time inwhich they occured. The red circled events represent those circled in red in Figure 4.3 (i.e. earthquakes that had >50% stress feature contribution). B) same as above panel, but zoomed in near well. The earthquakes strongly stress driven near the injection well, but also appear at different clusters throughtout the domain. Often those away from the well have early times compared to the other earthquakes in their cluster suggesting they may be starting the seismicity in those areas. There are some examples of earthquakes that are close but nearly stress driven as opposed to earthquake driven as well.

Appendix

4.A Appendix to Chapter 4

4.A.1 Model Pre-processing

A variety of issues and subsequent solutions arose in the model preprocessing that is important to elaborate on. As mentioned, previous work already compiled resources into a comprehensive, fully coupled poroelastic model of the PVU (Denlinger and RH O'Connell, 2020). However, this model was not easily portable to Abaqus and lacked sufficient discretization to capture large pressure gradients near the well. The methodology used to transfigure the initial model are presented here. We compare the model to a well known analytical solution and observed wellhead pressures to confirm its robustness.

Material Parameters and Meshing

The first difficulty with the Denlinger and O'Connel (D&O) model (Denlinger and RH O'Connell, 2020) is that the poroelastic material parameters are all defined at the nodes of the mesh. In Abaqus, there are a few material parameters defined at the nodes (pore pressure, void ratio, and saturation), but the elements (hexahedrons defined spatially by 8 nodes) are assigned other material parameters (ie. Young's modulus and bulk modulus of solid grains). After simple

conversions of the given material parameters in the D&O model to the values used in Abaqus, we thought the best way to solve the issue of defining the *node only* values to elements would be to average the 8 nodal coordinates that make up a hexahedron element to the value at that element.

However, the averaging proved ineffective for a variety of reasons. First, the D&O model near the region of the well head experiences strong changes in material values. The Leadville formation, the high permeable injection formation, is embedded in low permeable material. The nodal change between these materials was actually only 1 node thick in some instances so by taking the average of 8 nodes resulted in significantly reducing the order of magnitude of material permeability for areas near fluid injection. Second, the strong changes in material values coupled with the large spatial discretization of the D&O model near the wellhead resulted in unrealistic gradients and convergence issues.

Therefore, in order to solve the issues present with the conversion of the D&O model to Abaqus, we decided to make several adjustments to our model that we believe make it a stronger model overall. First, we decided to reduce the spatial discretization near the well head. The well head is actually composed of 3 separate perforated injection zones and creates strong pressure gradients that require smaller spatial sampling in order to capture the large and rapid changes there. This is difficult to do based on the previous mesh since preserving spatial features such as dipping beds and down scaling material features is not straight forward. Thankfully, the vertical discretization was already well defined by the D&O model so the only change to the discretization was the horizontal directions. We solved this problem by preserving the number of elements whilst changing the horizontal spacing to grow exponentially from the location of the well head. Then, the vertical spacing and material parameters of the D&O model are preserved in the smaller spacing by using a nearest point search measured in Euclidean distance. The spatial meshing changes between the D&O model and ours are shown in Figure 4.A.1.

The second adjustment we made was in the determination of material parameters throughout the model. As previously mentioned the D&O model allows for entirely unique material

parameters at every node, which caused difficulties in convergence for Abaqus. Using the newly discretized mesh of nodes/elements, and their associated material parameters, we applied a k-medoids clustering algorithm to cluster the nodes/elements based on similar material metrics across the combined set of materials. K-medoids is similar to k-means clustering, but instead of choosing the average from the kth cluster it chooses an actual data point as the center of the cluster. We worked with several different material cluster values, but ultimately decided on 1000. At this number, the model preserves many of the naturally occurring geological features such as the layered beds and salt domes whilst also maintaining a high level of material contrast near the wellhead without generating drastic gradients.

FEM Results Compared to Analytical Solution/Observation

A well known analytical solution exists to describe the spatial and temporal evolution of pore pressure due to continuous fluid injection into a poroelastic full space (Rudnicki, 1986). In order to gauge the success of the model, we first compare this solution to the 3D model using homogeneous material parameters. Additionally, we reduce the 3 injection nodes to a single node to better reflect the analytical solution. The radial analytical solution of pore pressure is compared with the closest radial axis given by the nodes shown in Figure 4.A.2. The solution for pore pressure matches well to the analytical solution after 10 days of constant injection using a typical bulk value of the crust as shown in Figure 4.A.3.

One thing to note is that the solution of pore pressure increases rapidly closer to the point of injection. The strong pressure gradients at this location require smaller elements then the horizontal discretization in the D&O model (200 m).

With the model now confirmed in the simplest case it was time to test a variety of k-medoid models, as previously described, and compare them to the observed wellhead pressures to confirm that the model was capable of capturing the observations. It is important to note that any complex model will result in overfitting of the wellhead data, and thus poor predictive ability for future

data.

There has been a plethora of previous work from observational drilling to pressure-flow modeling designed to capture the reservoir permeability structure (King and Block, 2019). These different observations and modeling have provided a sizeable range of permeability values. For example, the permeability of intact limestone and dolomite varies from 0.01 to 0.1 mD (Bear, 1988). Fracturing is expected to increase permeability outside of this laboratory setting. Drill stem tests gave an original permeability of 7.97 mD, yet at the same time additional analysis indicated permeability between 1.3 and 1.5 mD. Samples from a well 4.6 km to the northeast yielded permeability ranges of 0.03 to 1.3 mD (Harr, 1988). An earlier model by Denlinger and Roeloffs (Roeloffs and Denlinger, 2009) arrived at a permeability in the injection zone of 28 mD, with significantly lower values for the other formations. Additional pressure-flow models also arrive at ranges of 9.06 to 29.2 mD for certain injection phases (King and Block, 2019). The current *best* model (the D&O model) throughout the entire model domain, only has a maximum permeability of 1.97 mD. The final 1000 k-medoids model, modeled at constant injection rate (typical daily average from PVU injection data), is compared with several hypothetical analytical solutions for constant injection rate for a range of bulk permeabilities in Figure 4.A.4.

The final 3D heterogeneous model compares well with a range of typical observational values and observed wellhead pressures. In the near-field, the permeability matches the higher permeability analytical solutions as expected since there is likely fractured media in this location (King and Block, 2019). In the far-field, where the permeability structure is expected to decrease, the model approaches the lower permeability analytical solution. For the future, it will likely be important to test a variety of physics based models to understand the sensitivity introduced in the machine learning. However, we are confident in the evidence presented that our current model, adopted from the D&O model, is robust enough to continue with the primary goal of this work.

4.A.2 Supplementary figures for Chapter 4

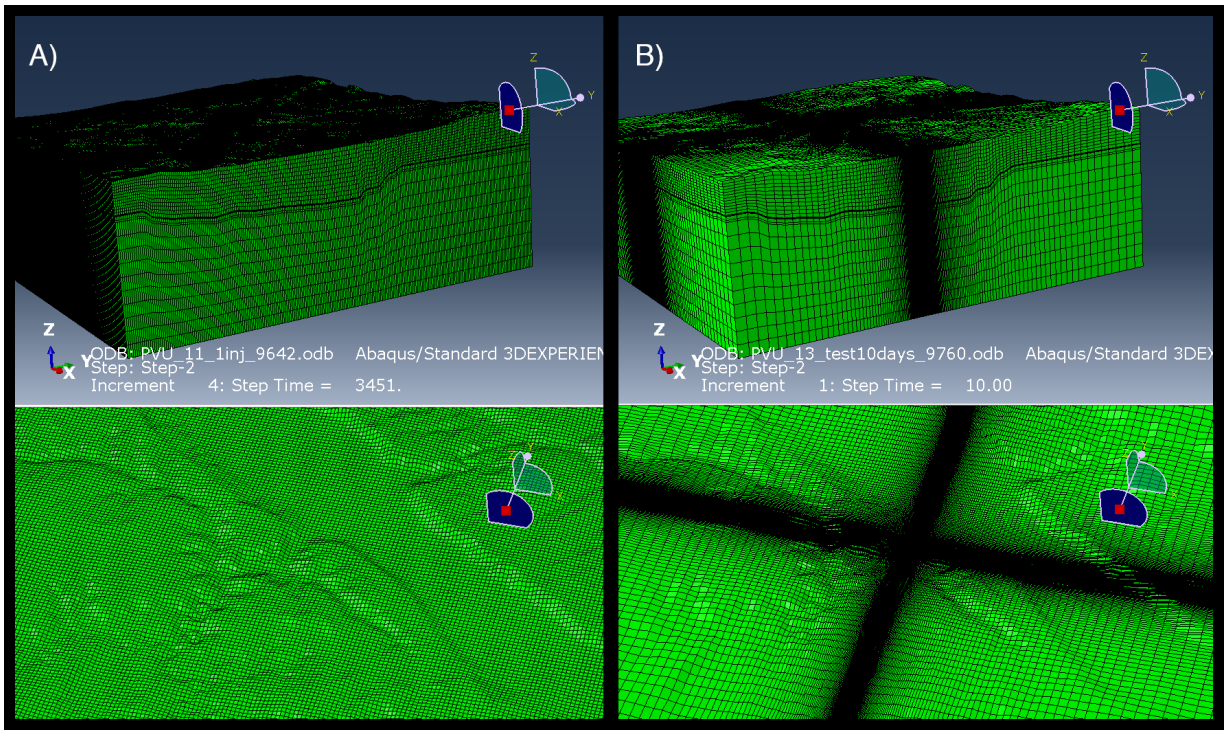


Figure 4.A.1: Previous model mesh from D&O model (Panel A) with surface view of well location compared to (Panel B) our smaller discretized model with similar surface view.

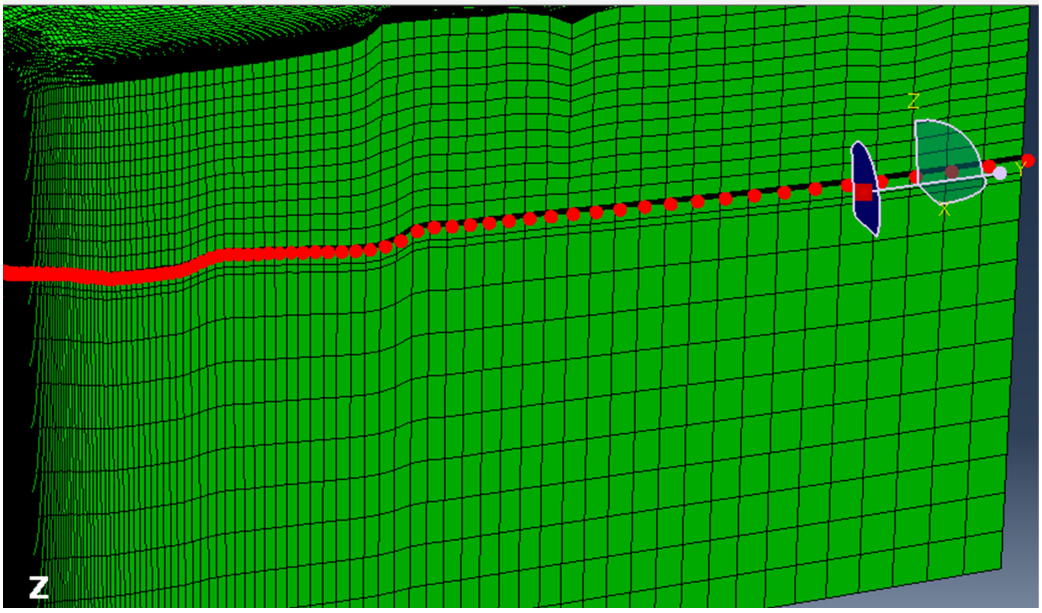


Figure 4.A.2: Nodes used in comparison with analytical solution. Well is located on the left and extends to the far field on the right.

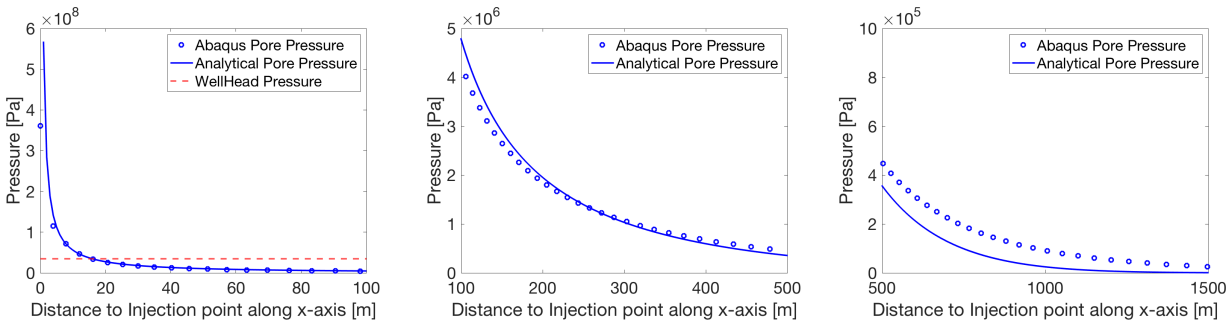


Figure 4.A.3: Analytical solution compared to the homogeneous 3D model. Dashed red line represents the average well head pressure of the observed PVU.

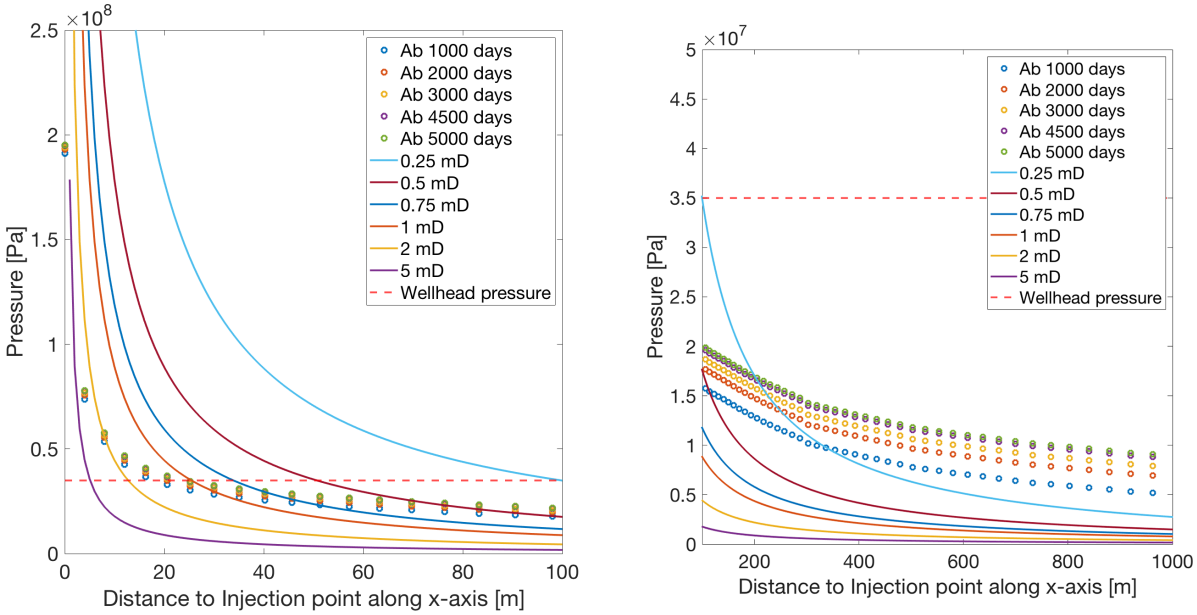


Figure 4.A.4: Final 1000 k-medoids model compared to several analytical solutions for a variety of constant rate injection times.

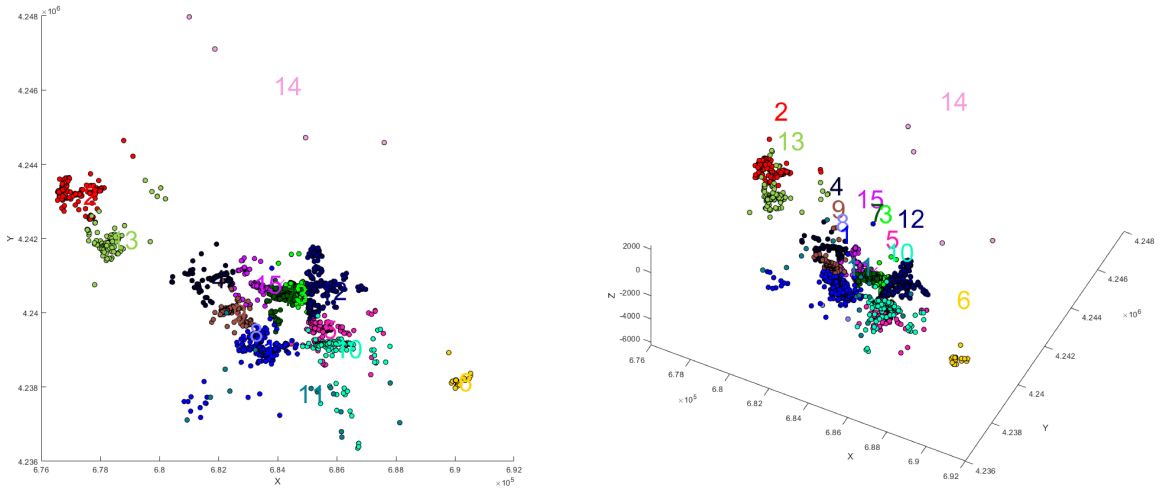


Figure 4.A.5: Different k-means cluster locations (1-15) of seismicity for the PVU. We extract the pore pressure at the center of each seismicity cluster from the numerical model in the subsequent figures. We include results for the near well cluster (7), two further regions with more diffuse responses (4) and (10) as well as farther distance (2) and (6).

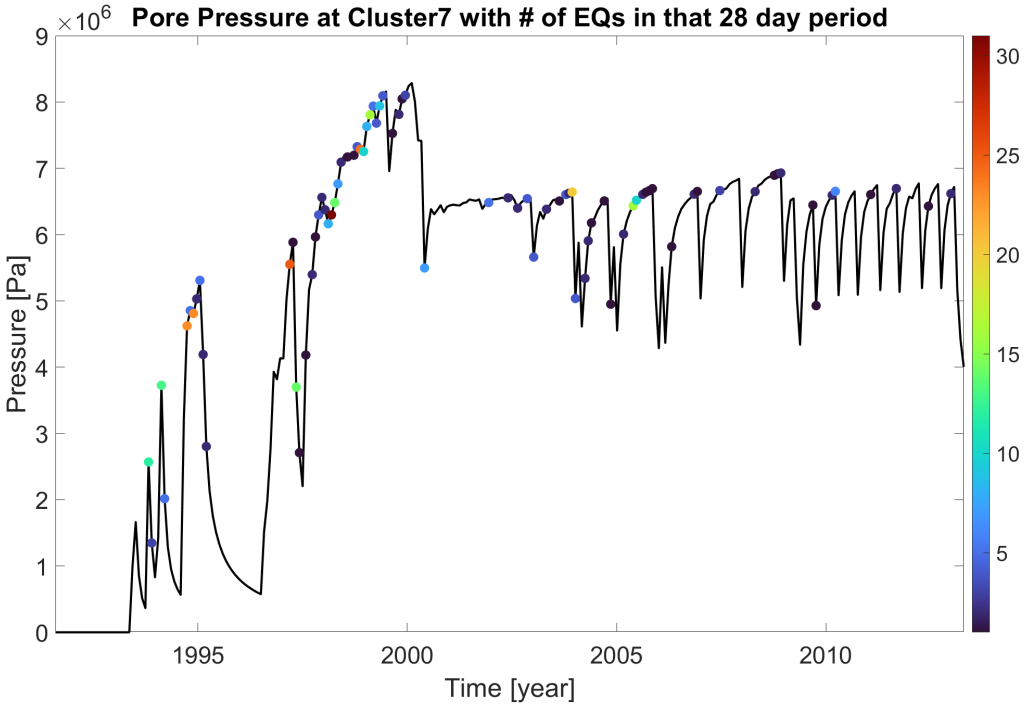


Figure 4.A.6: Cluster 7 near the well and pore pressure profile at the center of cluster. The pore pressure mimics the injection well rates due its close vicinity to the well.

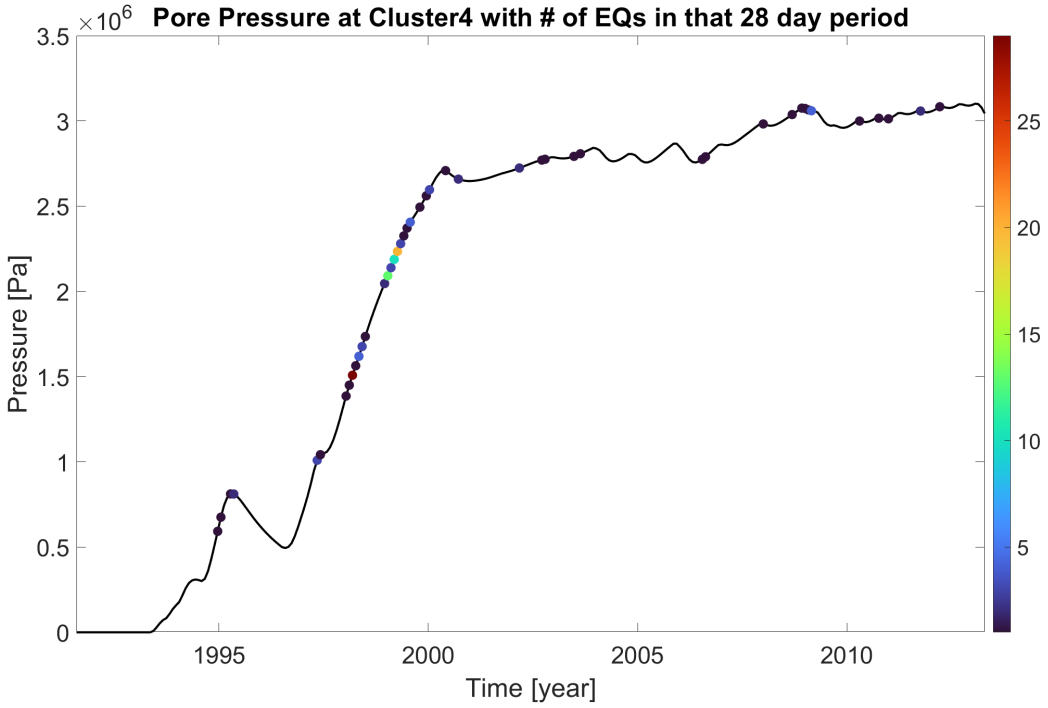


Figure 4.A.7: Cluster 4.

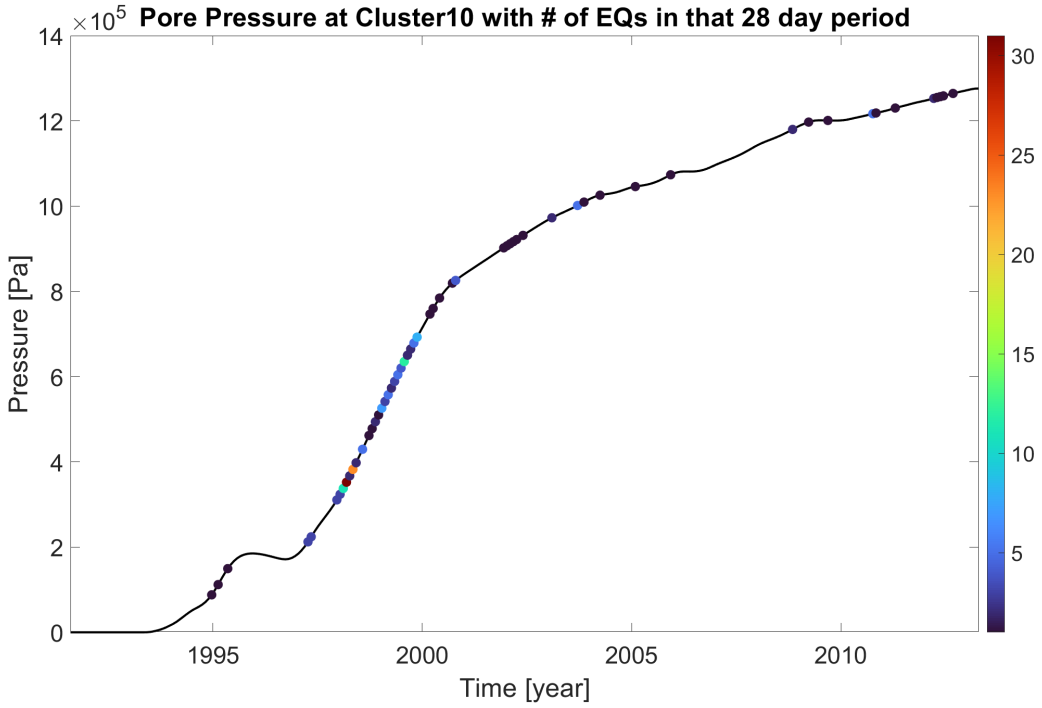


Figure 4.A.8: Cluster 10.

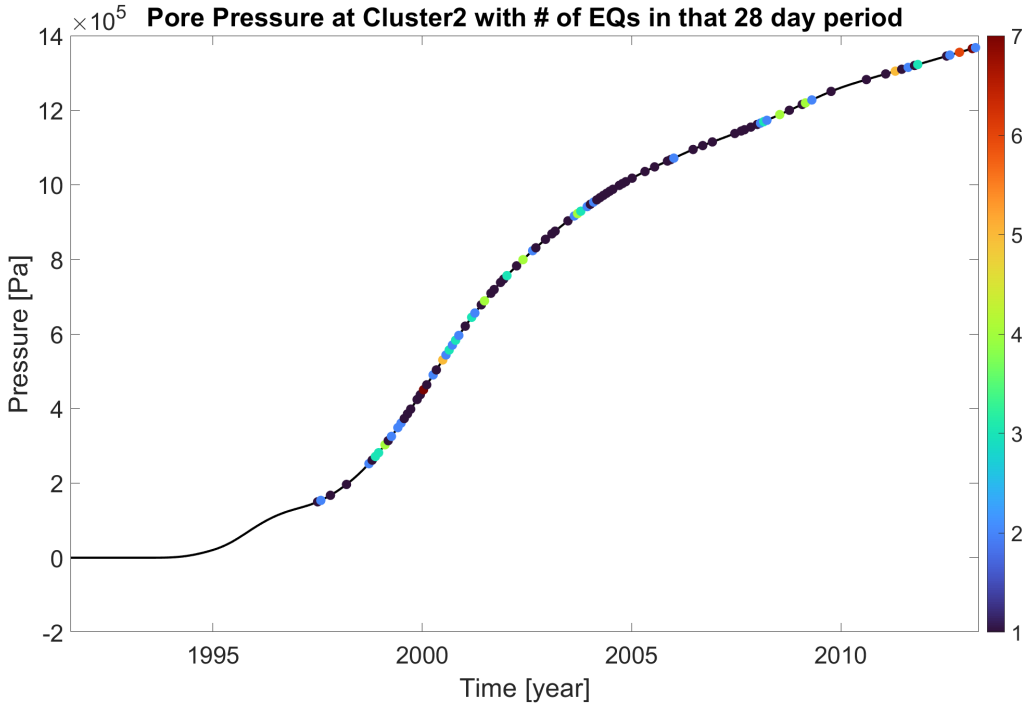


Figure 4.A.9: Cluster 2.

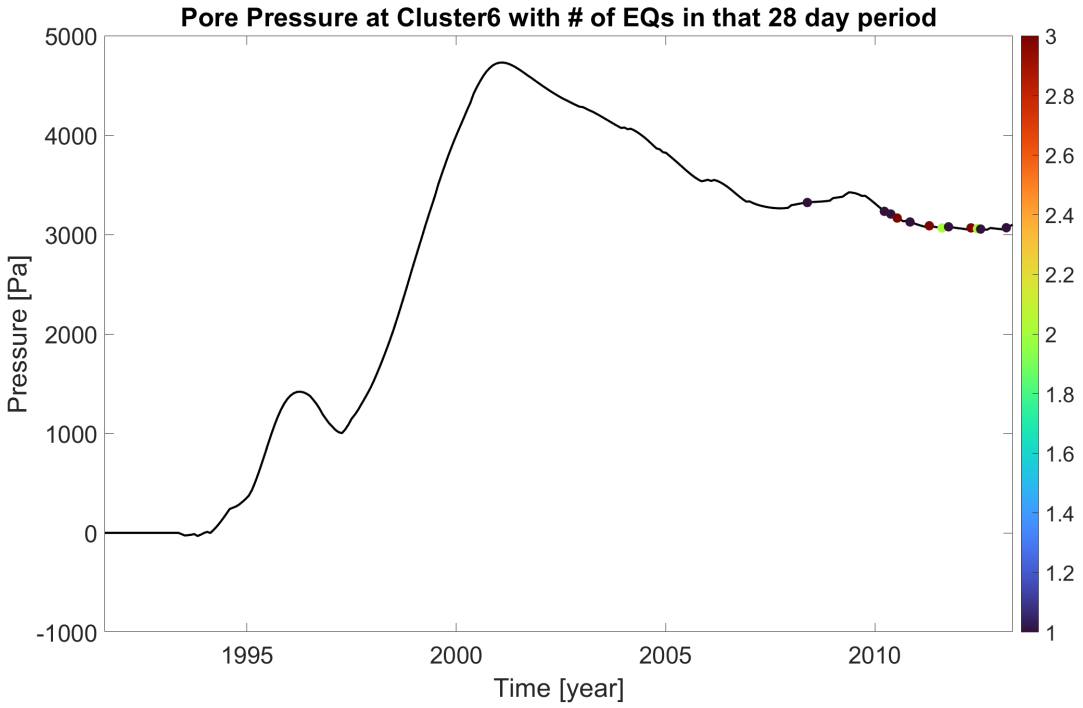


Figure 4.A.10: Cluster 6.

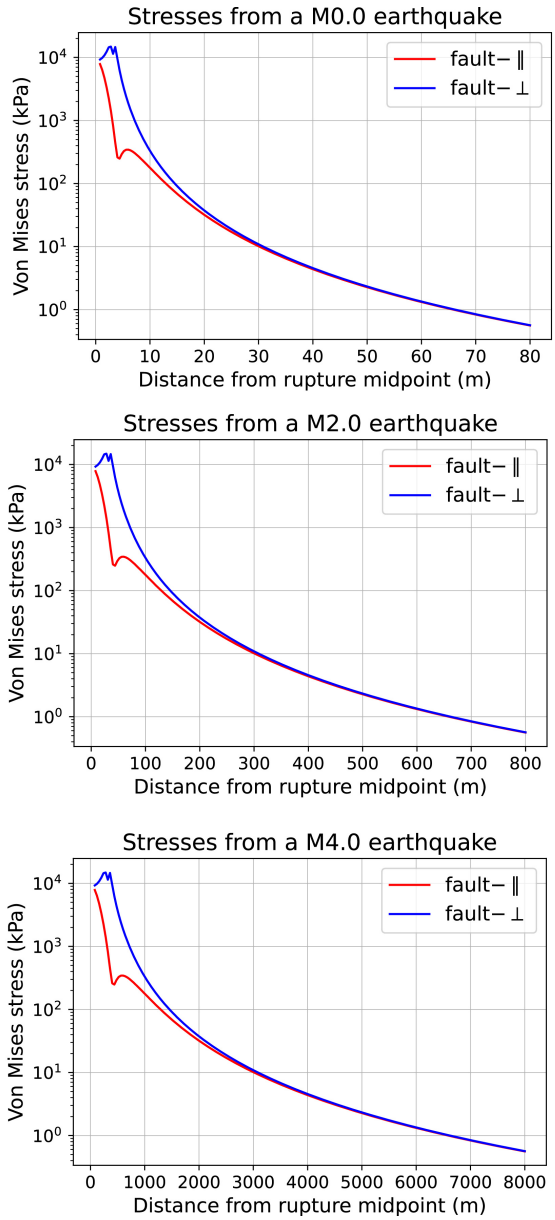


Figure 4.A.11: The von Mises stress in kPa for the three varying earthquake magnitudes (0.0, 2.0, and 4.0). We use cutde (Thompson, 2021) to resolve stress transfer produced from fullspace triangle dislocation elements assuming a uniform stress drop of 3 MPa, a shear modulus of 30 GPa, and a Poisson ratio of 0.25. We show that the von Mises stress is self similar for opposite receiver planes at certain distances, dependent on the magnitude, produced by the dislocation. We use thetriggering threshold of 10 kPa (Reasenberg and Simpson, 1992; Stein, 1999) which increases depending on the magnitude size. This distance is our perturbable radius used for the earthquake feature.

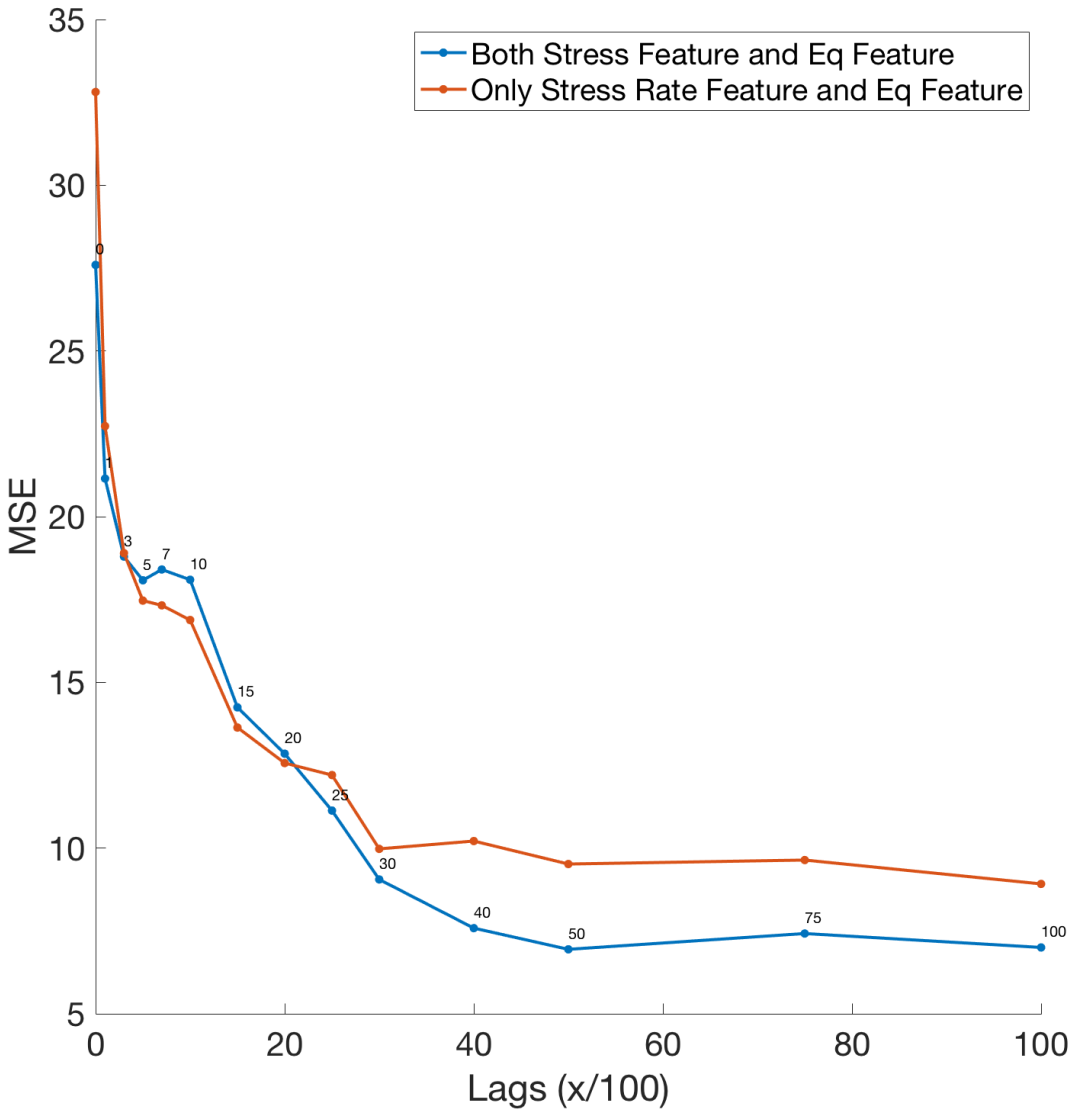


Figure 4.A.12: A sensitivity test to increasing and the overall MSE fit to the seismicity rate. We find that there is a local minimum near 5 lags. The fit does not improve after approximately 50 lags.

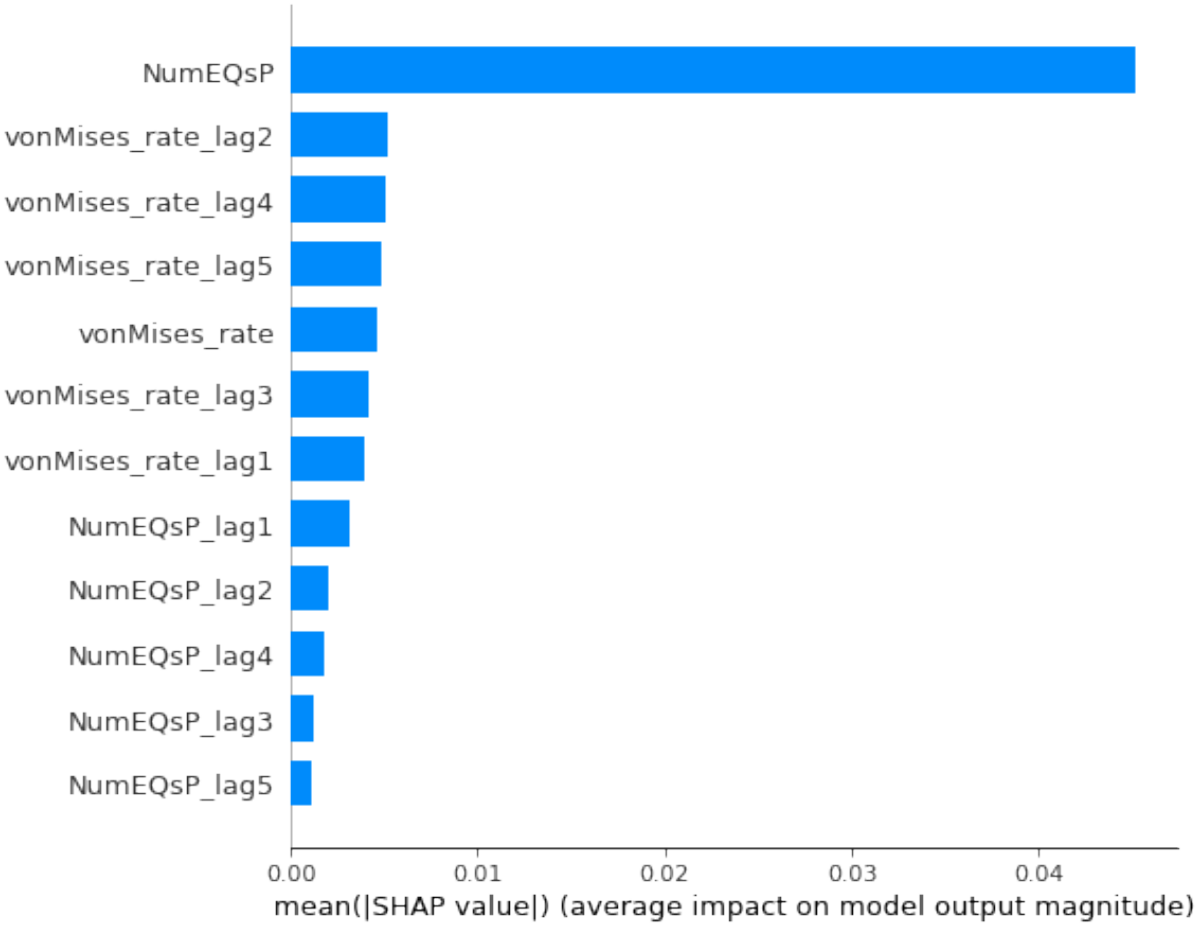


Figure 4.A.13: Mean absolute SHAP value for times in the model that an earthquake actually occured. This represents 2927 total events. The most important feature is the number of perturbable earthquakes (NumEQsP) that occurred during that same time step as the earthquake in question. The next 6 variables are all the stress rate from the stress change from the injection.

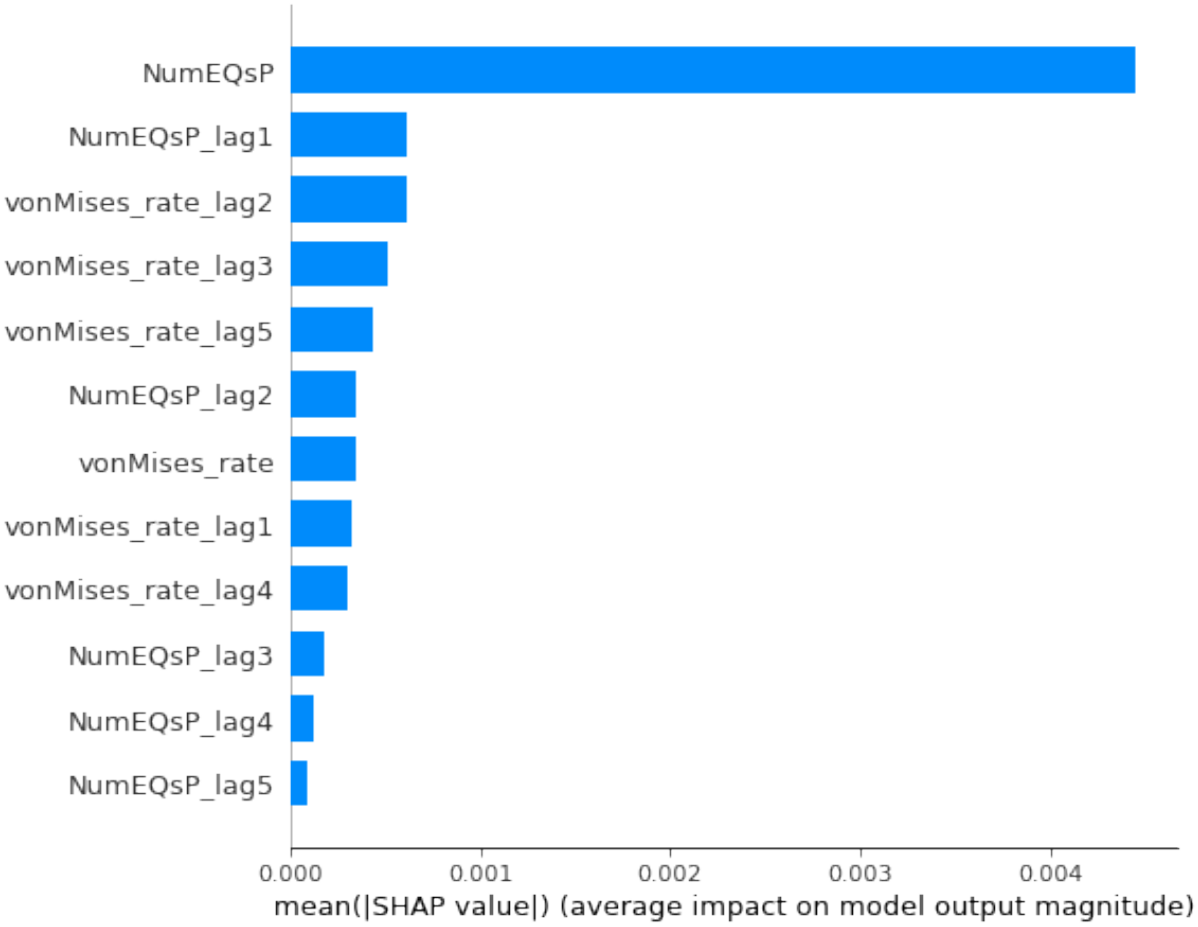


Figure 4.A.14: Simlar to SM Figure 4.A.13 except for all time steps in the model which includes the time steps when an earthquake is not occurring ($2927 * 284 = 831,268$ total samples).

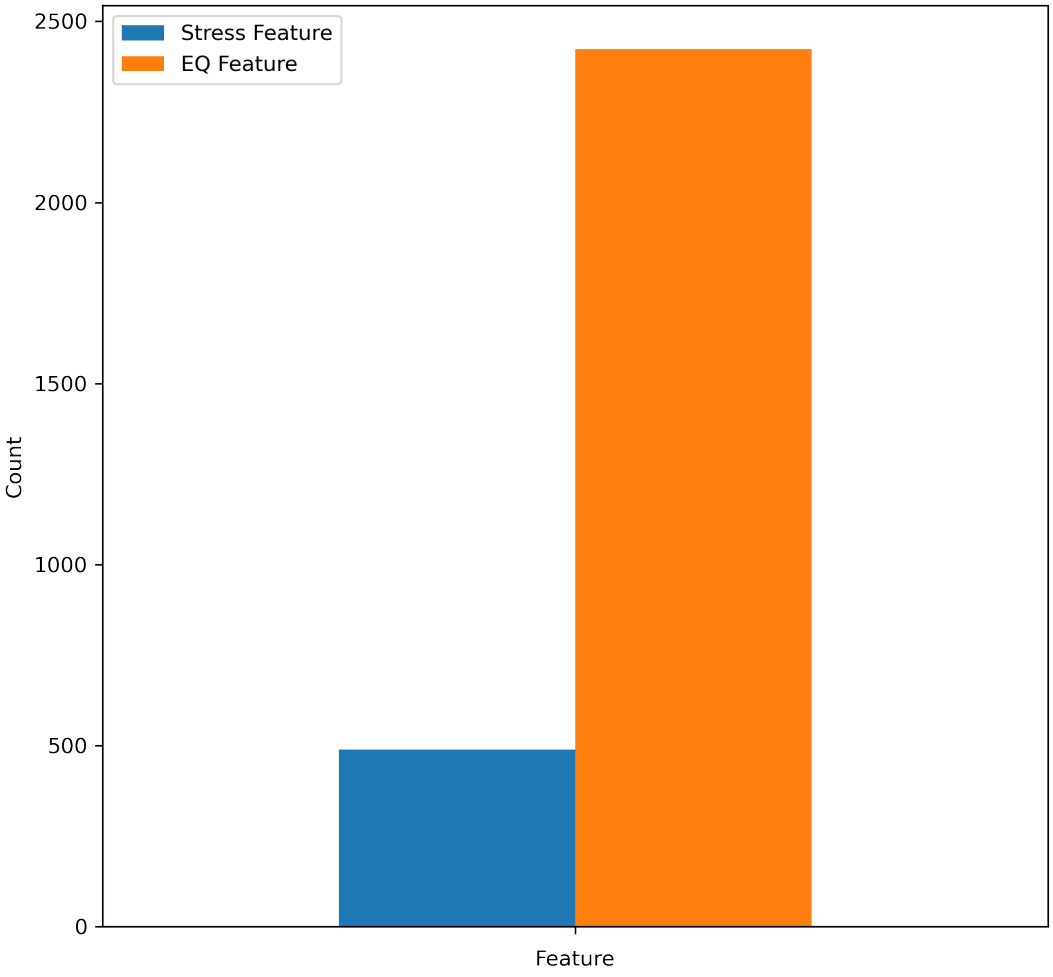


Figure 4.A.15: Ratio of the earthquake stress contribution totals for both the stress features and the earthquake features. For our model of including +5lags the stress feature to earthquake feature ratio approximately 1:5.

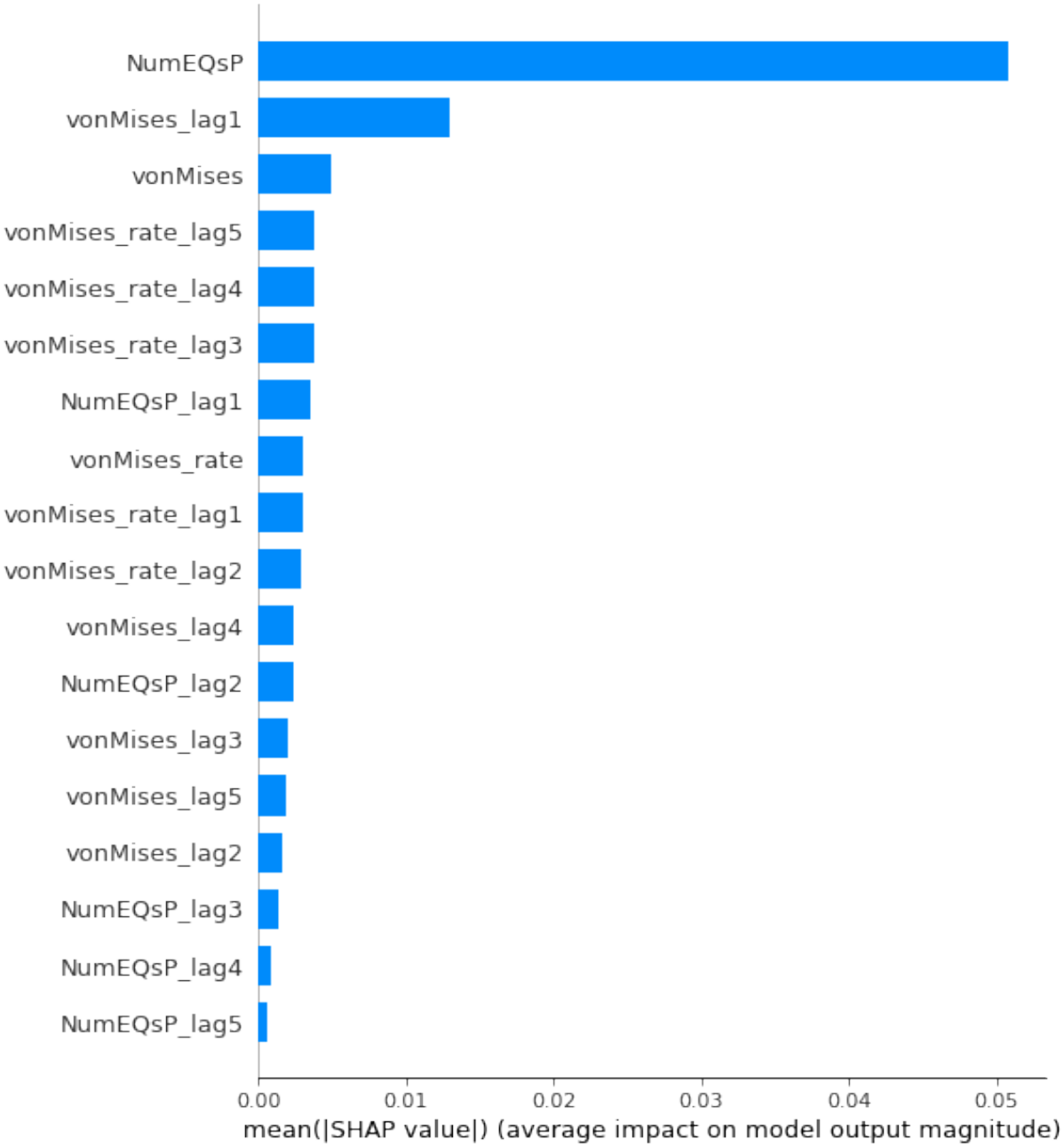


Figure 4.A.16: Similar to SM Figure 4.A.13 but for the model that includes both the von Mises stress and the von Mises stress rate. This represents 2927 total events. The most important feature is the number of perturbable earthquakes (NumEQsP) that occurred during that same time step as the earthquake in question. The next 65 variables are a mix of the von Mises and von Mises rate.

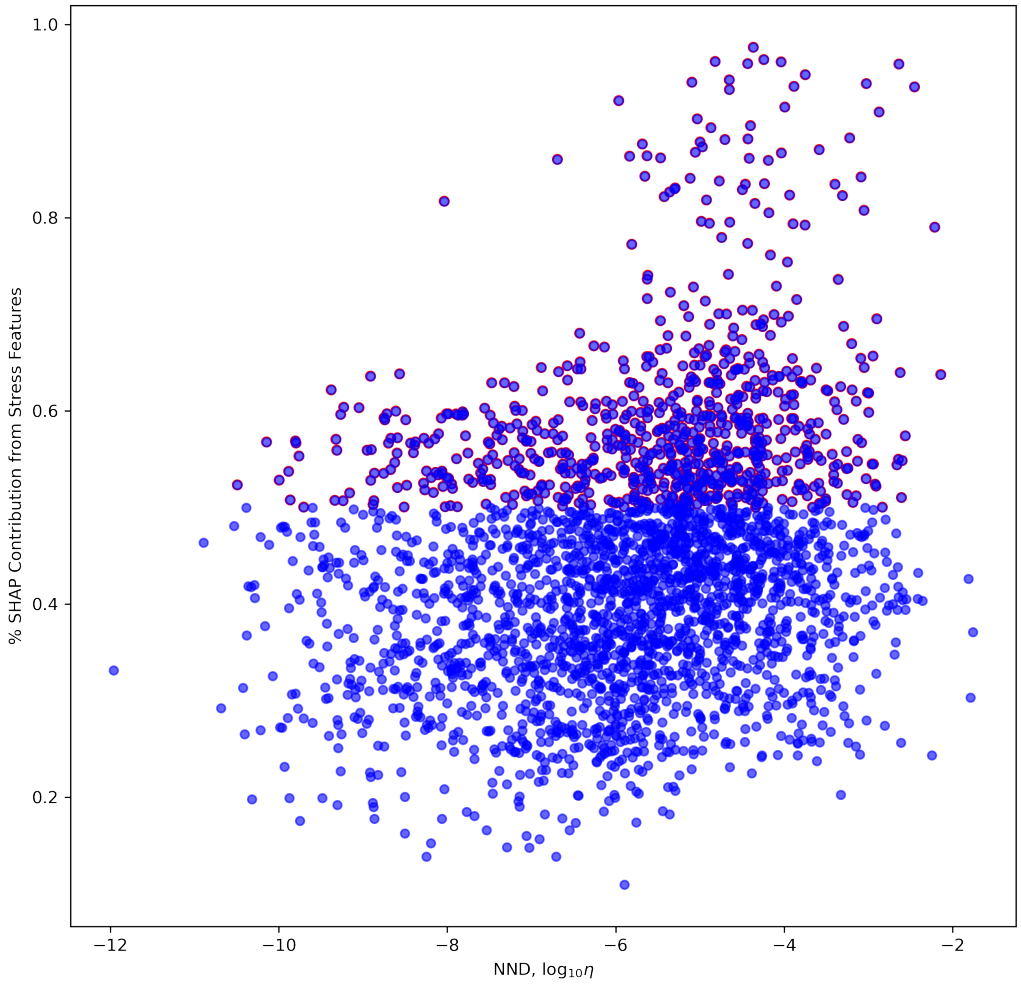


Figure 4.A.17: Similar to SM Figure 4.4 but for the model that includes both the von Mises stress and the von Mises stress rate. There is more earthquakes associated with the clustered mode, but still a large amount of background mode earthquakes.

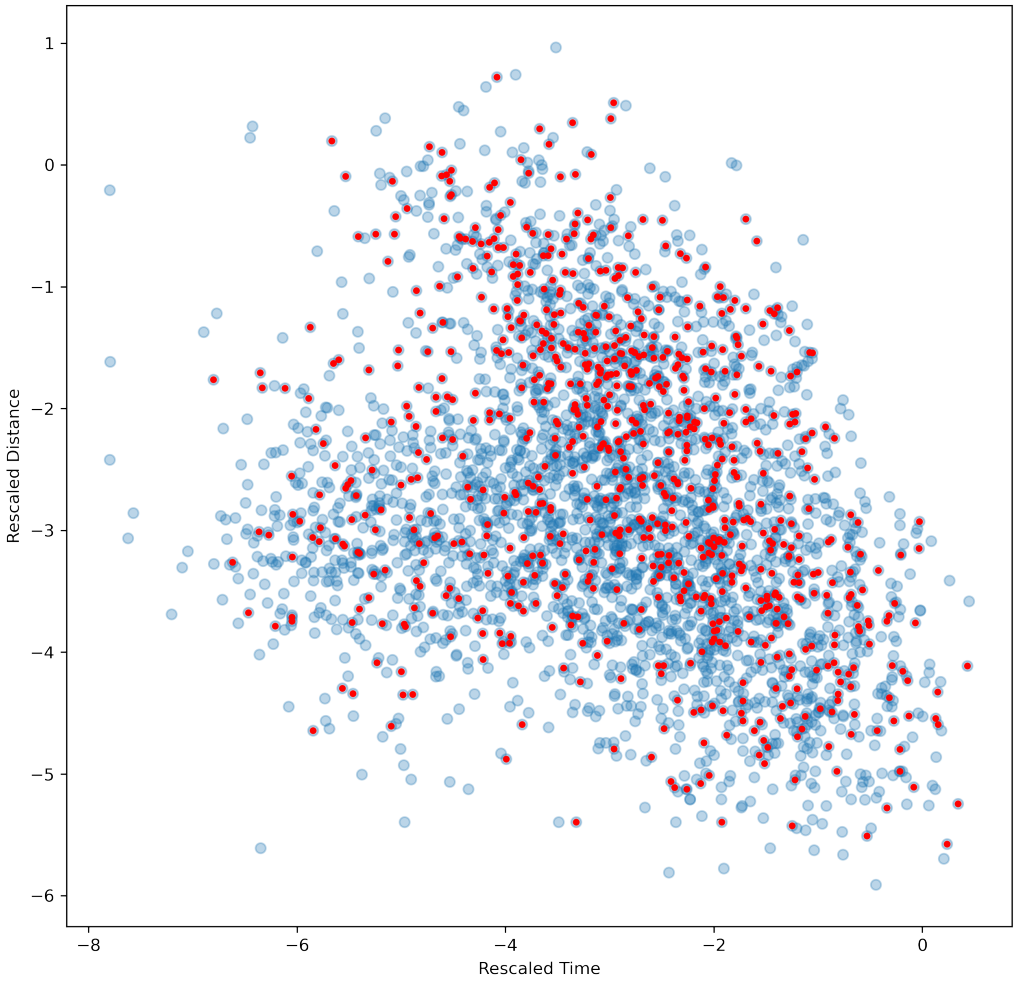


Figure 4.A.18: Similar to SM Figure 4.4 but for the model that includes both the von Mises stress and the von Mises stress rate. There is more earthquakes associated with the clustered mode, but still a large amount of background mode earthquakes.

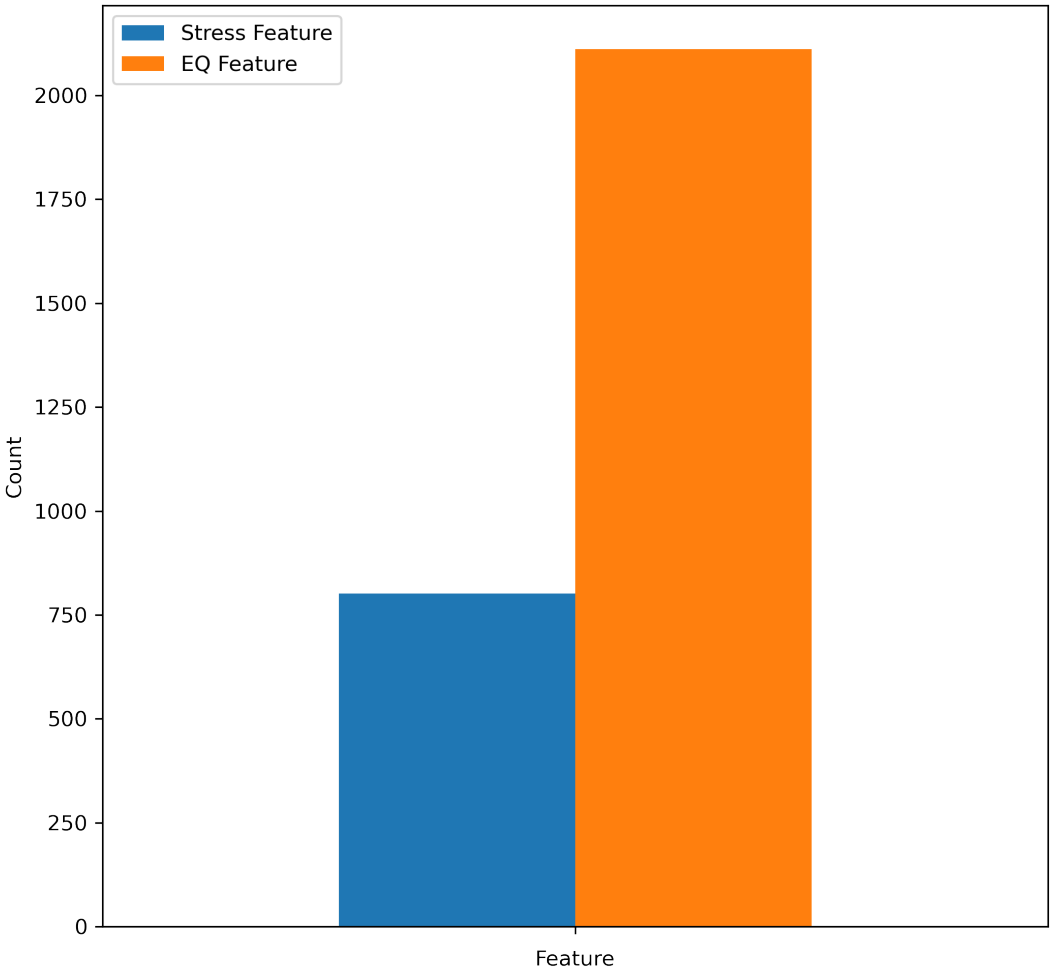


Figure 4.A.19: Similar to SM Figure 4.A.15 but for the model that includes both the von Mises stress and the von Mises stress rate. Ratio of the earthquake stress contribution totals for both the stress features and the earthquake features. For our model of including +5lags the stress feature to earthquake feature ratio approximately 1:3 which is must higher than the (1:5) ratio seen in the model that only has one stress feature.

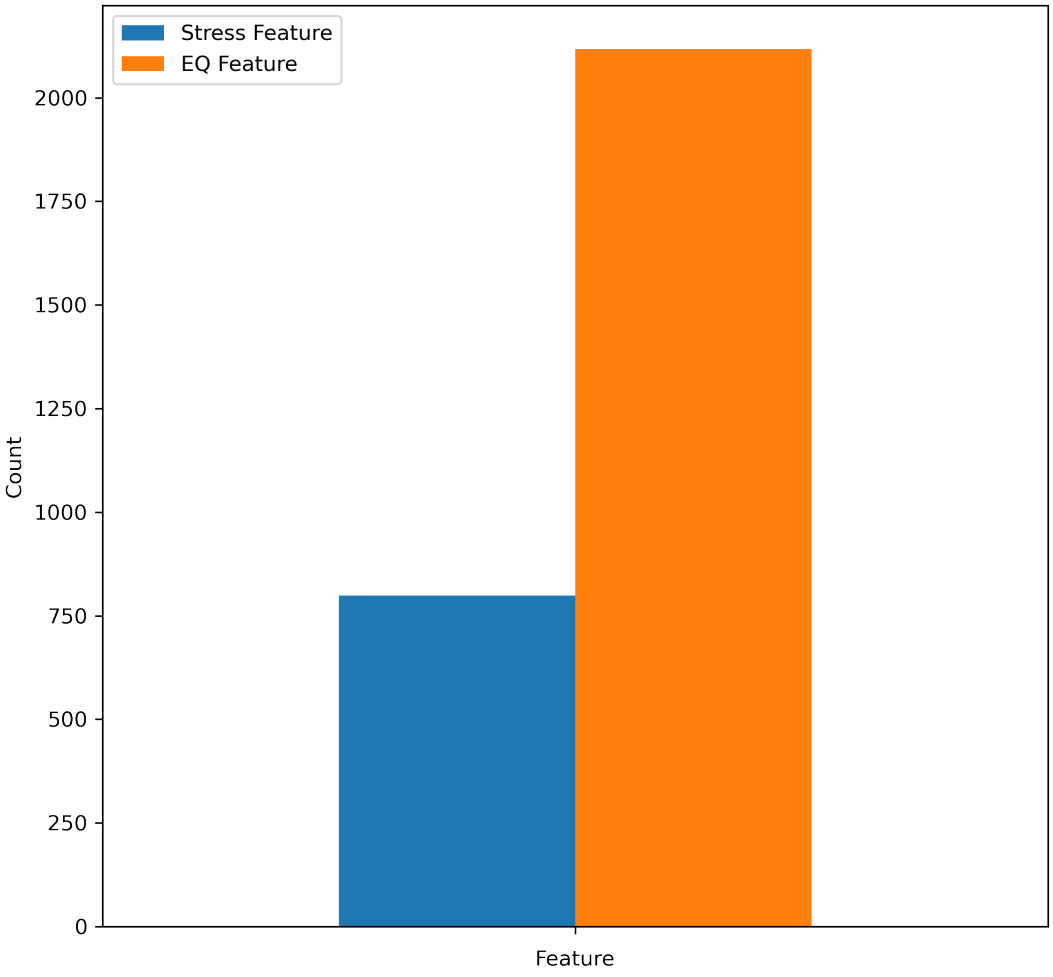


Figure 4.A.20: Similar to SM Figure 4.A.15 but for the model that includes both the von Mises stress and the von Mises stress rate and only +3 lags. The ratio is (0.3774) compared with the ratio at +5 lags (0.3794) shown in Figure SM 4.A.19.

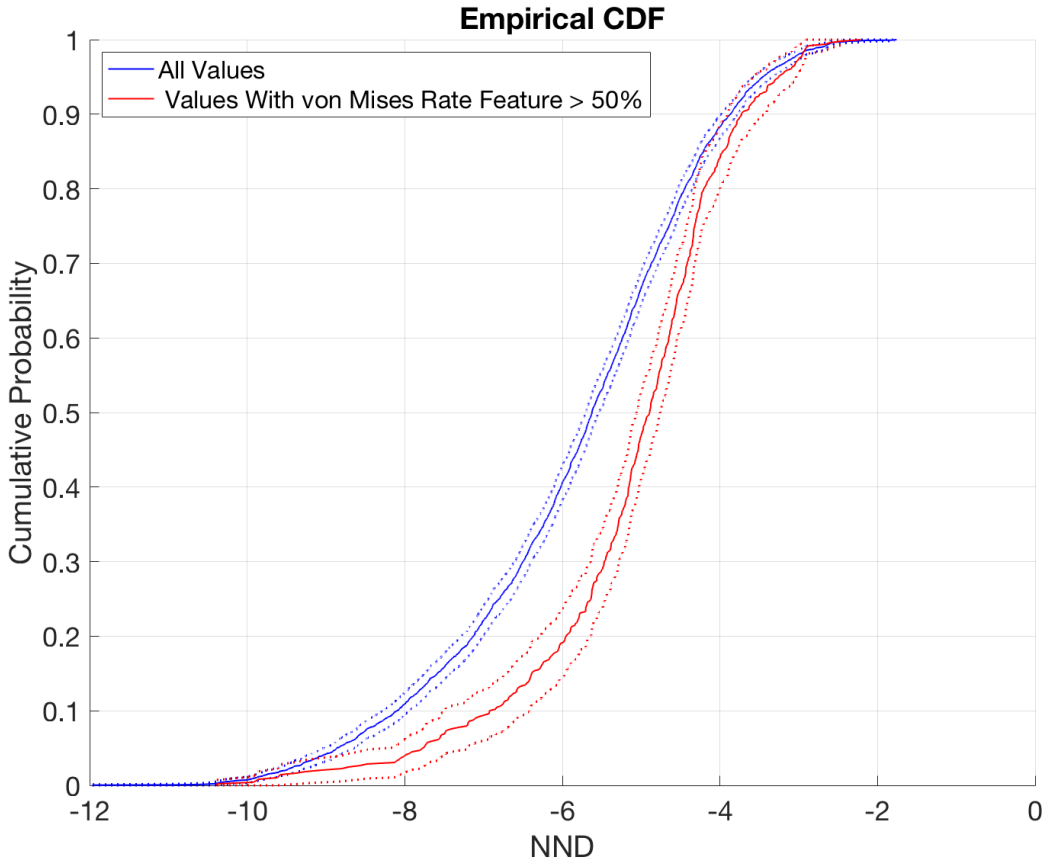


Figure 4.A.21: Empirical cumulative density functions of the two sample Kolmogorov–Smirnov test. We show that the distribution for the earthquakes with stress contribution $>$50% are not drawn from the same distribution as the total earthquakes with 99% confidence. Dashed line represents lower and upper confidence bounds for each distribution.

Bibliography

Ake, J., Mahrer, K., O'Connell, D., and Block, L. (2005). "Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado". In: *Bulletin of the Seismological Society of America* 95.2, pp. 664–683.

Allam, A. A. and Ben-Zion, Y. (2012). "Seismic velocity structures in the southern California plate-boundary environment from double-difference tomography". In: *Geophysical Journal International* 190.2, pp. 1181–1196.

Bachmann, C. E., Wiemer, S., Goertz-Allmann, B., and Woessner, J. (2012). "Influence of pore-pressure on the event-size distribution of induced earthquakes". In: *Geophysical Research Letters* 39.9.

Baiesi, M. and Paczuski, M. (2004). "Scale-free networks of earthquakes and aftershocks". In: *Physical review E* 69.6, p. 066106.

Baisch, S., Koch, C., and Muntendam-Bos, A. (2019). "Traffic light systems: To what extent can induced seismicity be controlled?" In: *Seismological Research Letters* 90.3, pp. 1145–1154.

Bao, X. and Eaton, D. W. (2016). "Fault activation by hydraulic fracturing in western Canada". In: *Science* 354.6318, pp. 1406–1409.

Barbot, S. and Fialko, Y. (2010). "A unified continuum representation of post-seismic relaxation mechanisms: semi-analytic models of afterslip, poroelastic rebound and viscoelastic flow". In: *Geophys. J. Int.* 182, pp. 1124–1140.

Barnhart, W. D., Benz, H. M., Hayes, G. P., Rubinstein, J. L., and Bergman, E. (2014). "Seismological and geodetic constraints on the 2011 Mw5. 3 Trinidad, Colorado earthquake and induced deformation in the Raton Basin". In: *Journal of Geophysical Research: Solid Earth* 119.10, pp. 7923–7933.

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

Belitz, K. and Bredehoeft, J. D. (1988). "Hydrodynamics of Denver Basin: Explanation of subnormal fluid pressures". In: *AAPG bulletin* 72.11, pp. 1334–1359.

Bense, V., Gleeson, T., Loveless, S., Bour, O., and Scibek, J. (2013). "Fault zone hydrogeology". In: *Earth-Science Reviews* 127, pp. 171–192. ISSN: 00128252.

Biot, M. A. (1941). "General Theory of Three-Dimensional Consolidation". In: *J. Appl. Phys.* 12.2, pp. 155–164. ISSN: 0021-8979, 1089-7550.

Blanton, C. M., Rockwell, T. K., Gontz, A., and Kelly, J. T. (2020). "Refining the spatial and temporal signatures of creep and co-seismic slip along the southern San Andreas Fault using very high resolution UAS imagery and SfM-derived topography, Coachella Valley, California". In: *Geomorphology* 357, p. 107064. ISSN: 0169555X.

Block, L. V., Wood, C. K., Yeck, W. L., and King, V. M. (2015). "Induced seismicity constraints on subsurface geological structure, Paradox Valley, Colorado". In: *Geophysical Journal International* 200.2, pp. 1172–1195.

Bommer, J. J., Oates, S., Cepeda, J. M., Lindholm, C., Bird, J., Torres, R., Marroquín, G., and Rivas, J. (2006). "Control of hazard due to seismicity induced by a hot fractured rock geothermal project". In: *Engineering geology* 83.4, pp. 287–306.

Brace, W. F. (1980). "Permeability of crystalline and argillaceous rocks". In: *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* Vol. 17 Issue 5. Elsevier, pp. 241–251.

Bredehoeft, J., Wolff, R., Keys, W., and Shuter, E. (1976). "Hydraulic fracturing to determine the regional in situ stress field, Piceance Basin, Colorado". In: *Geological Society of America Bulletin* 87.2, pp. 250–258.

Bronk Ramsey, C. (2009). "Bayesian Analysis of Radiocarbon Dates". In: *Radiocarbon* 51.1, pp. 337–360. ISSN: 0033-8222, 1945-5755.

Brothers, D., Kilb, D., Luttrell, K., Driscoll, N., and Kent, G. (2011). "Loading of the San Andreas fault by flood-induced rupture of faults beneath the Salton Sea". In: *Nature Geoscience* 4.7, pp. 486–492. ISSN: 1752-0894, 1752-0908.

Brown, M. R. and Ge, S. (2018). "Small earthquakes matter in injection-induced seismicity". In: *Geophysical Research Letters* 45.11, pp. 5445–5453.

Byerlee, J. (1978a). "Friction of rocks". In: 116, p. 12.

Byerlee, J. (1978b). "Friction of rock". In: *Pure Appl. Geophys.* 116, pp. 615–626.

Cacace, M., Hofmann, H., and Shapiro, S. A. (2021). "Projecting seismicity induced by complex alterations of underground stresses with applications to geothermal systems". In: *Scientific Reports* 11.1, p. 23560.

Caine, J. S., Evans, J. P., and Forster, C. B. (1996). "Fault zone architecture and permeability structure". In: *Geology*, p. 4.

Chang, K. W. and Segall, P. (2016). "Injection-induced seismicity on basement faults including poroelastic stressing". In: *J. Geophys. Res.* 121.4, pp. 2708–2726.

Clark, K., Northrop, S., and Read, C. (1966). "Geology of the Sangre de Cristo Mountains and adjacent areas, between Taos and Raton, New Mexico". In: *Taos-Raton-Spanish*

Peaks Country (New Mexico and Colorado): Geological Society 17th Annual Fall Field Conference Guidebook, pp. 56–65.

Cocco, M. (2002). "Pore pressure and poroelasticity effects in Coulomb stress analysis of earthquake interactions". In: *J. Geophys. Res.* 107, p. 2030. ISSN: 0148-0227.

Cochran, E. S., Li, Y.-G., Shearer, P. M., Barbot, S., Fialko, Y., and Vidale, J. E. (2009). "Seismic and geodetic evidence for extensive, long-lived fault damage zones". In: *Geology* 37, pp. 315–318.

Computational Science Research Center, 2. (2023). *CSRC*.

Coulomb, C. A. (1773). "Essai sur une application des regles de maximis et minimis a quelques problemes de statique relatifs a 1'architecture". In: *Mem. Div. Sav. Acad.*

Cremen, G. and Galasso, C. (2020). "Earthquake early warning: Recent advances and perspectives". In: *Earth-Science Reviews* 205, p. 103184.

Darcy, H. (1856). *Les fontaines publiques de la ville de Dijon: exposition et application des principes à suivre et des formules à employer dans les questions de distribution d'eau*. Vol. 1. Victor dalmont.

Dassault Systemes, 2. (2020). *ABAQUS (version 2019)*.

Delorey, A. A., Bokelmann, G. H., Johnson, C. W., and Johnson, P. A. (2021). "Estimation of the orientation of stress in the Earth's crust without earthquake or borehole data". In: *Communications Earth & Environment* 2.1, p. 190.

Denlinger, R. P. and RH O'Connell, D. (2020). "Evolution of faulting induced by deep fluid injection, Paradox Valley, Colorado". In: *Bulletin of the Seismological Society of America* 110.5, pp. 2308–2327.

Detournay, E. and Cheng, A. H.-D. (1993). "Fundamentals of poroelasticity". In: *Analysis and design methods*. Elsevier, pp. 113–171.

DeVries, P. M., Viégas, F., Wattenberg, M., and Meade, B. J. (2018). "Deep learning of aftershock patterns following large earthquakes". In: *Nature* 560.7720, pp. 632–634.

Dor, O., Ben-Zion, Y., Rockwell, T., and Brune, J. (2006). "Pulverized rocks in the Mojave section of the San An- dreas Fault Zone". In: *Earth Planet. Sci. Lett.* 245, pp. 642–654.

Douglas, J. and Aochi, H. (2014). "Using estimated risk to develop stimulation strategies for enhanced geothermal systems". In: *Pure and Applied Geophysics* 171, pp. 1847–1858.

Duffield, G. (2007). $AQTESOLV^{TM} Version 4.5 User's Guide$.

Durham, W. B. (1997). "Laboratory observations of the hydraulic behavior of a permeable fracture from 3800 m depth in the KTB pilot hole". In: *Journal of Geophysical Research: Solid Earth* 102 (B8), pp. 18405–18416. ISSN: 01480227.

Eissa, E. and Kazi, A. (1988). "Relation between static and dynamic Young's moduli of rocks". In: *International Journal of Rock Mechanics and Mining & Geomechanics Abstracts* 25.6.

Ellsworth, W. L. (2013). "Injection-induced earthquakes". In: *Science* 341.6142.

Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., and Shimamoto, T. (2019). "Triggering of the Pohang, Korea, earthquake (M w 5.5) by enhanced geothermal system stimulation". In: *Seismological Research Letters* 90.5, pp. 1844–1858.

Eyre, T. S., Eaton, D. W., Garagash, D. I., Zecevic, M., Venieri, M., Weir, R., and Lawton, D. C. (2019). "The role of aseismic slip in hydraulic fracturing–induced seismicity". In: *Science advances* 5.8, eaav7172.

Fama, M. D. and Pender, M. (1980). "Analysis of the hollow inclusion technique for measuring in situ rock stress". In: *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*. Vol. 17. 3. Elsevier, pp. 137–146.

Farr, T. and Kobrick, M. (2000). "Shuttle Radar Topography Mission produces a wealth of data". In: *AGU Eos* 81, pp. 583–585.

Fialko, Y. (2004). "Evidence of fluid-filled upper crust from observations of post-seismic deformation due to the 1992 M_w7.3 Landers earthquake". In: *J. Geophys. Res.* 109, B08401, 10.1029/2004JB002985.

Fialko, Y. and Jin, Z. (2021a). "Simple shear origin of the cross-faults ruptured in the 2019 Ridgecrest earthquake sequence". In: *Nature Geoscience*.

Fialko, Y., Sandwell, D., Agnew, D., Simons, M., Shearer, P., and Minster, B. (2002a). "Deformation on nearby faults induced by the 1999 Hector Mine earthquake". In: *Science* 297, pp. 1858–1862.

Fialko, Y. and Simons, M. (2000). "Deformation and seismicity in the Coso geothermal area, Inyo County, California: Observations and modeling using satellite radar interferometry". In: *J. Geophys. Res.* 105, pp. 21781–21793.

Fialko, Y. (2006). "Interseismic strain accumulation and the earthquake potential on the southern San Andreas fault system." In: *Nature* 441.7096, pp. 968–971. ISSN: 1476-4687.

— (2021). "Estimation of absolute stress in the hypocentral region of the 2019 Ridgecrest, California, earthquakes". In: *J. Geophys. Res.*

Fialko, Y. and Jin, Z. (2021b). "Simple shear origin of the cross-faults ruptured in the 2019 Ridgecrest earthquake sequence". In: *Nature Geoscience* 14.7, pp. 513–518.

Fialko, Y., Sandwell, D., Agnew, D., Simons, M., Shearer, P., and Minster, B. (2002b). "Deformation on nearby faults induced by the 1999 Hector Mine earthquake." In: *Science* 297.5588, pp. 1858–1862. ISSN: 1095-9203.

Field, E. H., Arrowsmith, R. J., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., Johnson, K. M., Jordan, T. H., Madden, C., *et al*. (2014). "Uniform California Earthquake Rupture Forecast, version 3 (UCERF3) The time-independent model". In: *Bull. Seism. Soc. Am.* 104.3, pp. 1122–1180.

Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., and Davies, R. J. (2018). "Global review of human-induced earthquakes". In: *Earth-Science Reviews* 178, pp. 438–514.

Frohlich, C., Ellsworth, W., Brown, W. A., Brunt, M., Luetgert, J., MacDonald, T., and Walter, S. (2014). "The 17 May 2012 M4. 8 earthquake near Timpson, East Texas: An event possibly triggered by fluid injection". In: *Journal of Geophysical Research: Solid Earth* 119.1, pp. 581–593.

Fuis, G. S., Scheirer, D. S., Langenheim, V. E., and Kohler, M. D. (2012). "A new perspective on the geometry of the San Andreas fault in Southern California and its relationship to lithospheric structure". In: *Bull. Seism. Soc. Am.* 102, pp. 236–251.

Fumal, T. E. (2002). "Timing of Large Earthquakes since A.D. 800 on the Mission Creek Strand of the San Andreas Fault Zone at Thousand Palms Oasis, near Palm Springs, California". In: *Bulletin of the Seismological Society of America* 92.7, pp. 2841–2860. ISSN: 0037-1106.

Ge, S. (2011). "Comment on "Evidence that the 2008 Mw 7.9 Wenchuan Earthquake Could Not Have Been Induced by the Zipingpu Reservoir" by Kai Deng, Shiyong Zhou, Rui Wang, Russell Robinson, Cuiping Zhao, and Wanzheng Cheng". In: *Bulletin of the Seismological Society of America* 101.6, pp. 3117–3118. ISSN: 0037-1106.

Ge, S. and Saar, M. O. (2022). "Induced seismicity during geoenergy development—A hydromechanical perspective". In: *Journal of Geophysical Research: Solid Earth* 127.3, e2021JB023141.

Geldon, L. (1989). "Ground-water hydrology of the central Raton Basin, Colorado and New Mexico". In.

Gephart, J. W. and Forsyth, D. W. (1984). "An improved method for determining the regional stress tensor using earthquake focal mechanism data: application to the San Fernando earthquake sequence". In: *Journal of Geophysical Research: Solid Earth* 89.B11, pp. 9305–9320.

Giardini, D. (2009). "Geothermal quake risks must be faced". In: *Nature* 462.7275, pp. 848–849.

Gischig, V. S. (2015). "Rupture propagation behavior and the largest possible earthquake induced by fluid injection into deep reservoirs". In: *Geophysical Research Letters* 42.18, pp. 7420–7428.

Glasgow, M., Schmandt, B., Wang, R., Zhang, M., Bilek, S. L., and Kiser, E. (2021a). "Raton Basin induced seismicity is hosted by networks of short basement faults and mimics tectonic earthquake statistics". In: *Journal of Geophysical Research: Solid Earth* 126.11, e2021JB022839.

Glasgow, M., Schmandt, B., Wang, R., Zhang, M., Bilek, S., and Kiser, E. (2021b). *Raton Basin 2016-2020 earthquake catalog*.

Glasgow, M. E., Schmandt, B., and Bilek, S. L. (2023). "Cascading multi-segment rupture in an injection-induced earthquake sequence with a Mw 5.3 mainshock". In: *Earth and Planetary Science Letters* 620, p. 118335.

Goebel, T., Weingarten, M., Chen, X., Haffener, J., and Brodsky, E. (2017). "The 2016 Mw5.1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at >40 km from fluid disposal wells". In: *Earth Planet. Sci. Lett.* 472, pp. 50–61. ISSN: 0012-821X.

Goebel, T., Rosson, Z., Brodsky, E., and Walter, J. (2019). "Aftershock deficiency of induced earthquake sequences during rapid mitigation efforts in Oklahoma". In: *Earth and Planetary Science Letters* 522, pp. 135–143.

Goertz-Allmann, B. P., Kühn, D., Oye, V., Bohloli, B., and Aker, E. (2014). "Combining microseismic and geomechanical observations to interpret storage integrity at the In Salah CCS site". In: *Geophysical Journal International* 198.1, pp. 447–461.

Gorelick, S. M. (1983). "A review of distributed parameter groundwater management modeling methods". In: *Water Resources Research* 19.2, pp. 305–319.

Gorelick, S. M., Freeze, R. A., Donohue, D., Keely, J. F., *et al.* (1993). *Groundwater contamination: optimal capture and containment.* Lewis Publishers Inc.

Gorelick, S. M. and Remson, I. (1982). "Optimal dynamic management of groundwater pollutant sources". In: *Water Resources Research* 18.1, pp. 71–76.

Gorelick, S. M. and Zheng, C. (2015). "Global change and the groundwater management challenge". In: *Water Resources Research* 51.5, pp. 3031–3051. eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014WR016825.

Grasso, J.-R. and Wittlinger, G. (1990). "Ten years of seismic monitoring over a gas field". In: *Bulletin of the Seismological Society of America* 80.2, pp. 450–473.

Gupta, H. K. (1992). *Reservoir Induced Earthquakes*. Elsevier.

— (2002). "A review of recent studies of triggered earthquakes by artificial water reservoirs with special emphasis on earthquakes in Koyna, India". In: *Earth-Science Reviews* 58.3, pp. 279–310. ISSN: 00128252.

Gurrola, L. D. and Rockwell, T. K. (1996). "Timing and slip for prehistoric earthquakes on the Superstition Mountain Fault, Imperial Valley, southern California". In: *Journal of Geophysical Research: Solid Earth* 101 (B3), pp. 5977–5985. ISSN: 01480227.

Hampel, A., Lüke, J., Krause, T., and Hetzel, R. (2019). "Finite-element modelling of glacial isostatic adjustment (GIA): Use of elastic foundations at material boundaries versus the

geometrically non-linear formulation". In: *Computers & Geosciences* 122, pp. 1–14. ISSN: 00983004.

Harr, C. (1988). "Final geological well report". In: *US Bureau of Reclamation Test Well* 1.

Hernandez, R. (2020). "Fluid pressure modeling of faults and simulation-optimization of wastewater injection in the Raton Basin, CO-NM". MA thesis. San Diego State University.

Hernandez, R. and Weingarten, M. (2019). "Step-rate Test Calibration of Primary Injection Reservoir Permeability in a Case of Injection Induced Seismicity, Raton Basin, CO-NM". In: *AGU fall meeting abstracts*. Vol. 2019, S13E–0493.

Higley, D. K. (2007). *Petroleum systems and assessment of undiscovered oil and gas in the Raton Basin–Sierra Grande Uplift Province, Colorado and New Mexico—USGS Province 41*. 69-N. US Geological Survey.

Hill, R. (2021). In: forthcoming.

Hill, R. G., Weingarten, M., Rockwell, T. K., and Fialko, Y. (2023). "Major southern San Andreas earthquakes modulated by lake-filling events". In: *Nature* 618, pp. 761–766.

Ho, T. K. *et al*. (1995). "Proceedings of 3rd international conference on document analysis and recognition". In: *Proceedings of 3rd international conference on document analysis and recognition*.

Ho, T. K. (1998). "The random subspace method for constructing decision forests". In: *IEEE transactions on pattern analysis and machine intelligence* 20.8, pp. 832–844.

Hsu, N.-S. and Yeh, W. W.-G. (1989). "Optimum experimental design for parameter identification in groundwater hydrology". In: *Water Resources Research* 25.5, pp. 1025–1040.

Jacob, C. (1950). "Flow of groundwater". In: *Engineering hydraulics*, pp. 321–386.

Jeppson, T. N., Bradbury, K. K., and Evans, J. P. (2010). "Geophysical properties within the San Andreas Fault Zone at the San Andreas Fault Observatory at Depth and their relationships to rock properties and fault zone structure". In: *Journal of Geophysical Research* 115 (B12), B12423. ISSN: 0148-0227.

Jin, Z. and Fialko, Y. (2020). "Finite slip models of the 2019 Ridgecrest earthquake sequence constrained by space geodetic data and aftershock locations". In: *Bull. Seism. Soc. Am.* 110, pp. 1660–1679.

Jin, Z., Fialko, Y., Zubovich, A., and Schöne, T. (2022). "Lithospheric deformation due to the 2015 M7.2 Sarez (Pamir) earthquake constrained by 5 years of space geodetic observations". In: *J. Geophys. Res.* 127, e2021JB022461.

Johnson, K. (2013). "Slip rates and off-fault deformation in Southern California inferred from GPS data and models". In: *J. Geophys. Res.* 118, pp. 5643–5664.

Keranen, K. M., Savage, H. M., Abers, G. A., and Cochran, E. S. (2013). "Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence". In: *Geology* 41.6, pp. 699–702.

Keranen, K. M. and Weingarten, M. (2018). "Induced seismicity". In: *Annual Review of Earth and Planetary Sciences* 46, pp. 149–174.

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., and Ge, S. (2014). "Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection". In: *Science* 345.6195, pp. 448–451.

King, G. C. P., Stein, R. C., and Lin, J. (1994a). "Static Stress Change and the Triggering of Earthquakes". In: *Bull. Seism. Soc. Am.* 84, pp. 935–953.

King, G. C. P., Stein, R. S., and Lin, J. (1994b). "Static stress changes and the triggering of earthquakes". In: *Bull. Seismol. Soc. Am.* 84, pp. 935–953.

King, V. and Block, L. (2019). *Paradox valley seismic and injection data - AMERIGEOSS community platform DataHub. (beta)*.

King, V. M., Block, L. V., and Wood, C. K. (2016). "Pressure/flow modeling and induced seismicity resulting from two decades of high-pressure deep-well brine injection, Paradox Valley, Colorado". In: *Geophysics* 81.5, B119–B134.

Kroll, K. A. and Cochran, E. S. (2021). "Stress controls rupture extent and maximum magnitude of induced earthquakes". In: *Geophysical Research Letters* 48.11, e2020GL092148.

LaBonte, A., Brown, K., and Fialko, Y. (2009). "Hydrogeologic detection and finite-element modeling of a slow-slip event in the Costa Rica prism toe". In: *J. Geophys. Res.* 114, B00A02.

Langenbruch, C., Dinske, C., and Shapiro, S. (2011). "Inter event times of fluid induced earthquakes suggest their Poisson nature". In: *Geophysical Research Letters* 38.21.

Langenbruch, C., Ellsworth, W. L., Woo, J.-U., and Wald, D. J. (2020). "Value at induced risk: Injection-induced seismic risk from low-probability, high-impact events". In: *Geophysical Research Letters* 47.2, e2019GL085878.

Langenbruch, C., Weingarten, M., and Zoback, M. D. (2018). "Physics-based forecasting of man-made earthquake hazards in Oklahoma and Kansas". In: *Nature communications* 9.1, pp. 1–10.

Langenbruch, C. and Zoback, M. D. (2016). "How will induced seismicity in Oklahoma respond to decreased saltwater injection rates?" In: *Science advances* 2.11, e1601542.

Lin, G., Shearer, P. M., and Hauksson, E. (2007a). "Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005". In: *J. Geophys. Res.* 112, B12309, doi:10.1029/2007JB004986.

Lin, G., Shearer, P. M., and Hauksson, E. (2007b). "Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate southern California seismicity from 1981 to 2005". In: *J. Geophys. Res.* 112.B12. B12309, pp. 1–14. ISSN: 2156-2202.

Lin, J. and Stein, R. S. (2004). "Stress triggering in thrust and subduction earthquakes and stress interaction between the southern San Andreas and nearby thrust and strike-slip faults". In: *Journal of Geophysical Research: Solid Earth* 109.B2.

Lindsey, E. and Fialko, Y. (2013). "Geodetic slip rates in the Southern San Andreas Fault System: Effects of elastic heterogeneity and fault geometry". In: *J. Geophys. Res.* 118, pp. 689–697.

Lindsey, E. O., Sahakian, V. J., Fialko, Y., Bock, Y., Barbot, S., and Rockwell, T. K. (2014). "Interseismic strain localization in the San Jacinto fault zone". In: *Pure Appl. Geophys.* 171.11, pp. 2937–2954.

Lindsey, E. O. and Fialko, Y. (2016). "Geodetic constraints on frictional properties and earthquake hazard in the Imperial Valley, Southern California". In: *J. Geophys. Res.* 121, pp. 1097–1113.

Lockner, D. A., Morrow, C., Moore, D., and Hickman, S. (2011). "Low strength of deep San Andreas fault gouge from SAFOD core". In: *Nature* 472.7341, pp. 82–85.

Lundberg, S. M. and Lee, S.-I. (2017). "A unified approach to interpreting model predictions". In: *Advances in neural information processing systems* 30.

Lundgren, P. E., Hetland, A., Liu, Z., and Fielding, E. J. (2009). "Southern San Andreas-San Jacinto fault system slip rates estimated from earthquake cycle models constrained by GPS and interferometric synthetic aperture radar observations". In: *J. Geophys. Res.* 114, B02403, doi:10.1029/2008JB005996.

Luttrell, K., Sandwell, D., Smith-Konter, B., Bills, B., and Bock, Y. (2007). "Modulation of the earthquake cycle at the southern San Andreas fault by lake loading". In: *J. Geophys. Res.* 112.B8, B08411. ISSN: 0148-0227.

Macartney, H. and O'Farrell, C. (2010). "A Raton Basin geothermal prospect". In: *AAPG, Durango, CO*.

Majer, E. L. and Peterson, J. E. (2007). "The impact of injection on seismicity at The Geysers, California Geothermal Field". In: *International Journal of Rock Mechanics and Mining Sciences* 44.8, pp. 1079–1090.

McGarr, A., Simpson, D., Seeber, L., and Lee, W. (2002). "Case histories of induced and triggered seismicity". In: *International Geophysics Series* 81.A, pp. 647–664.

McGarr, A. (2014). "Maximum magnitude earthquakes induced by fluid injection". In: *Journal of Geophysical Research: solid earth* 119.2, pp. 1008–1019.

Meltzner, A. J. and Wald, D. J. (2003). "Aftershocks and triggered events of the great 1906 California earthquake". In: *Bull. Seism. Soc. Am.* 93, pp. 2160–2186.

Mignan, A., Broccardo, M., Wiemer, S., and Giardini, D. (2017). "Induced seismicity closed-form traffic light system for actuarial decision-making during deep fluid injections". In: *Scientific reports* 7.1, pp. 1–10.

Mignan, A., Landtwing, D., Kästli, P., Mena, B., and Wiemer, S. (2015). "Induced seismicity risk analysis of the 2006 Basel, Switzerland, Enhanced Geothermal System project: Influence of uncertainties on risk mitigation". In: *Geothermics* 53, pp. 133–146.

Miller, S. A. (2013). "The Role of Fluids in Tectonic and Earthquake Processes". In: *Advances in Geophysics*. Vol. 54. Elsevier, pp. 1–46. ISBN: 978-0-12-380940-7.

Mitchell, E., Fialko, Y., and Brown, K. M. (2013). "Temperature dependence of frictional healing of Westerly granite: experimental observations and numerical simulations". In: *Geochemistry, Geophys., Geosystems* 14, pp. 567–582.

Mitchell, E., Fialko, Y., and Brown, K. (2015). "Frictional properties of gabbro at conditions corresponding to slow slip events in subduction zones". In: *Geochemistry, Geophys., Geosystems* 16.11, pp. 4006–4020.

Mitchell, T. and Faulkner, D. (2009). "The nature and origin of off-fault damage surrounding strike-slip fault zones with a wide range of displacements: A field study from the Atacama fault system, northern Chile". In: *Journal of Structural Geology* 31.8, pp. 802–816. ISSN: 01918141.

Mohr, O. (1882). "Über die Darstellung des Spannungszustandes und des Deformationszustandes eines Körperelementes und über die Anwendung derselben in der Festigkeitslehre". In: *Der Civilingenieur* 28.2, pp. 113–156.

Morrow, C., Lockner, D., Moore, D., and Hickman, S. (2014). "Deep permeability of the San Andreas Fault from San Andreas Fault Observatory at Depth (SAFOD) core samples". In: *Journal of Structural Geology* 64, pp. 99–114. ISSN: 01918141.

Morton, N., Girty, G. H., and Rockwell, T. K. (2012). "Fault zone architecture of the San Jacinto fault zone in Horse Canyon, southern California: A model for focused post-seismic fluid flow and heat transfer in the shallow crust". In: *Earth Planet. Sci. Lett.* 329, pp. 71–83.

Mueller, K. (1984). "Neotectonics, alluvial history and soil chronology of the southwestern margin of the Sierra de Los Cucapas". In: *Baja California Norte [Master's thesis]: San Diego, California, San Diego State University*.

Nakai, J., Sheehan, A., and Bilek, S. (2017a). "Seismicity of the rocky mountains and Rio Grande Rift from the EarthScope Transportable Array and CREST temporary seismic networks, 2008–2010". In: *Journal of Geophysical Research: Solid Earth* 122.3, pp. 2173–2192.

Nakai, J., Weingarten, M., Sheehan, A., Bilek, S., and Ge, S. (2017b). "A possible causative mechanism of Raton Basin, New Mexico and Colorado earthquakes using recent seismicity patterns and pore pressure modeling". In: *Journal of Geophysical Research: Solid Earth* 122.10, pp. 8051–8065.

Nelson, P. H., Gianoutsos, N. J., and Anna, L. O. (2013). "Outcrop control of basin-scale underpressure in the Raton Basin, Colorado and New Mexico". In.

Nikkhoo, M. and Walter, T. R. (2015). "Triangular dislocation: an analytical, artefact-free solution". In: *Geophysical Journal International* 201.2, pp. 1119–1141.

Nof, R., Ziv, A., Doin, M., Baer, G., Fialko, Y., Wdowinski, S., Eyal, Y., and Bock, Y. (2012). "Rising of the lowest place on Earth due to Dead Sea water-level drop: Evidence from SAR Interferometry and GPS". In: *J. Geophys. Res.* 117, B05412, doi:10.1029/2011JB008961.

Norbeck, J. H. and Horne, R. N. (2018). "Maximum magnitude of injection-induced earthquakes: A criterion to assess the influence of pressure migration along faults". In: *Tectonophysics* 733, pp. 108–118.

Nur, A. and Byerlee, J. (1971). "An exact effective stress law for elastic deformation of rock with fluids". In: *Journal of geophysical research* 76.26, pp. 6414–6419.

Papazafeiropoulos, G., Muñiz-Calvente, M., and Martínez-Pañeda, E. (2017). "Abaqus2Matlab: A suitable tool for finite element post-processing". In: *Adv. Eng. Softw.* 105, pp. 9–16.

Pearse, J. and Fialko, Y. (2010). "Mechanics of active magmatic intraplating in the Rio Grande Rift near Socorro, New Mexico". In: *J. Geophys. Res.* 115, B07413.

Philibosian, B., Fumal, T., and Weldon, R. (2011). "San Andreas Fault Earthquake Chronology and Lake Cahuilla History at Coachella, California". In: *Bulletin of the Seismological Society of America* 101.1, pp. 13–38. ISSN: 0037-1106.

Pollitz, F. F. and Sacks, I. S. (1997). "The 1995 Kobe, Japan, earthquake: A Long-Delayed Aftershock of the Offshore 1944 Tonankai and 1946 Nankaido Earthquakes". In: *Bull. Seism. Soc. Am.* 87, pp. 1–10.

Qin, Y., Chen, T., Ma, X., and Chen, X. (2022). "Forecasting induced seismicity in Oklahoma using machine learning methods". In: *Scientific Reports* 12.1, p. 9319.

Rajendran, K. and Talwani, P. (1992). "The role of elastic, undrained, and drained responses in triggering earthquakes at Monticello Reservoir, South Carolina". In: *Bull. Seism. Soc. Am.* 82.4, pp. 1867–1888.

Reasenberg, P. A. and Simpson, R. W. (1992). "Response of regional seismicity to the static stress change produced by the Loma Prieta earthquake". In: *Science* 255.5052, pp. 1687–1690.

Reimer, P. J., Austin, W. E., Bard, E., Bayliss, A., Blackwell, P. G., Ramsey, C. B., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., *et al.* (2020). "The IntCal20 Northern

Hemisphere radiocarbon age calibration curve (0–55 cal kBP)". In: *Radiocarbon* 62.4, pp. 725–757.

Rempe, M., Mitchell, T., Renner, J., Nippress, S., Ben-Zion, Y., and Rockwell, T. (2013). "Damage and seismic velocity structure of pulverized rocks near the San Andreas Fault". In: *J. Geophys. Res.* 118.6, pp. 2813–2831.

Rice, J. R. and Cleary, M. P. (1976a). "Some basic stress-diffusion solutions for fluid-saturated elastic porous media with compressible constituents". In: *Rev. Geophys.* 14, pp. 227–241.

Rice, J. R. and Cleary, M. P. (1976b). "Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents". In: *Rev. Geophys.* 14.2, p. 227. ISSN: 8755-1209.

Richards-Dinger, K. B. and Shearer, P. M. (1997). "Estimating crustal thickness in southern California by stacking *PmP* arrivals". In: *Journal of Geophysical Research: Solid Earth* 102 (B7), pp. 15211–15224. ISSN: 01480227.

Rockwell, T., Sisk, M., Girty, G., Dor, O., Wechsler, N., and Ben-Zion, Y. (2009). "Chemical and physical characteristics of pulverized Tejon Lookout granite adjacent to the San Andreas and Garlock faults: Implications for earthquake physics". In: *Pure Appl. Geophys.* 166.10, pp. 1725–1746.

Rockwell, T. K., Meltzner, A. J., Haaker, E. C., and Madugo, D. (2022). "The late Holocene history of Lake Cahuilla: Two thousand years of repeated fillings within the Salton Trough, Imperial Valley, California". In: *Quaternary Science Reviews* 282, p. 107456.

Rockwell, T. K., Meltzner, A. J., and Haaker, E. C. (2018). "Dates of the Two Most Recent Surface Ruptures on the Southernmost San Andreas Fault Recalculated by Precise Dating of Lake Cahuilla Dry Periods". In: *Bulletin of the Seismological Society of America* 108.5, pp. 2634–2649. ISSN: 0037-1106, 1943-3573.

Roeloffs, E. (1988a). "FAULT STABILITY CHANGES INDUCED BENEATH A RESERVOIR WITH CYCLIC VARIATIONS IN WATER LEVEL". In: *J. Geophys. Res.* 93, pp. 2107–2124.

Roeloffs, E. and Denlinger, R. (2009). "An axisymmetric coupled flow and deformation model for pore pressure caused by brine injection in Paradox valley, Colorado: Implications for the mechanisms of induced seismicity". In: *Preliminary Report to the Bureau of Reclamation*.

Roeloffs, E. A. (1988b). "Fault stability changes induced beneath a reservoir with cyclic variations in water level". In: *Journal of Geophysical Research* 93 (B3), p. 2107. ISSN: 0148-0227.

Ross, Z. E., Cochran, E. S., Trugman, D. T., and Smith, J. D. (2020). "3D fault architecture controls the dynamism of earthquake swarms". In: *Science*, p. 6.

Rubinstein, J. L., Ellsworth, W. L., McGarr, A., and Benz, H. M. (2014). "The 2001–present induced earthquake sequence in the Raton Basin of northern New Mexico and southern Colorado". In: *Bulletin of the Seismological Society of America* 104.5, pp. 2162–2181.

Rudnicki, J. W. (1986). "Fluid mass sources and point forces in linear elastic diffusive solids". In: *Mechanics of Materials* 5.4, pp. 383–393. ISSN: 0167-6636.

Rutqvist, J., Rinaldi, A. P., Cappa, F., and Moridis, G. J. (2015). "Modeling of fault activation and seismicity by injection directly into a fault zone associated with hydraulic fracturing of shale-gas reservoirs". In: *Journal of Petroleum Science and Engineering* 127, pp. 377–386.

Salditch, L., Stein, S., Neely, J., Spencer, B. D., Brooks, E. M., Agnon, A., and Liu, M. (2020). "Earthquake supercycles and long-term fault memory". In: *Tectonophysics* 774, p. 228289.

Savage, J. and Burford, R. (1973). "Geodetic determination of relative plate motion in central California". In: *J. Geophys. Res.* 78, pp. 832–845.

Schoenball, M., Davatzes, N. C., and Glen, J. M. (2015). "Differentiating induced and natural seismicity using space-time-magnitude statistics applied to the Coso Geothermal field". In: *Geophysical Research Letters* 42.15, pp. 6221–6228.

Schulte-Pelkum, V., Ross, Z. E., Mueller, K., and Ben-Zion, Y. (2020). "Tectonic Inheritance With Dipping Faults and Deformation Fabric in the Brittle and Ductile Southern California Crust". In: *Journal of Geophysical Research: Solid Earth* 125.8. ISSN: 2169-9313, 2169-9356.

Schultz, R., Beroza, G. C., and Ellsworth, W. L. (2021). "A risk-based approach for managing hydraulic fracturing–induced seismicity". In: *Science* 372.6541, pp. 504–507.

Segall, P. (1989). "Earthquakes triggered by fluid extraction". In: *Geology* 17, pp. 942–946.

Segall, P. (2010). *Earthquake and Volcano Deformation*. Princeton University Press.

Segall, P. and Lu, S. (2015). "Injection-induced seismicity: Poroelastic and earthquake nucleation effects". In: *Journal of Geophysical Research: Solid Earth* 120.7, pp. 5082–5103.

Shapiro, S. A., Dinske, C., Langenbruch, C., and Wenzel, F. (2010). "Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations". In: *The Leading Edge* 29.3, pp. 304–309.

Shapley, L. S. *et al*. (1953). "A value for n-person games". In.

Sharma, S., Hainzl, S., Zöeller, G., and Holschneider, M. (2020). "Is Coulomb stress the best choice for aftershock forecasting?" In: *Journal of Geophysical Research: Solid Earth* 125.9, e2020JB019553.

Shirzaei, M., Ellsworth, W. L., Tiampo, K. F., González, P. J., and Manga, M. (2016). "Surface uplift and time-dependent seismic hazard due to fluid injection in eastern Texas". In: *Science* 353.6306, pp. 1416–1419.

Shmonov, V., Vitiovtova, V., Zharikov, A., and Grafchikov, A. (2003a). “Permeability of the continental crust: implications of experimental data”. In: *Journal of Geochemical Exploration* 78-79, pp. 697–699. ISSN: 03756742.

— (2003b). “Permeability of the continental crust: implications of experimental data”. In: *Journal of Geochemical Exploration* 78, pp. 697–699.

Sibson, R. H. (1994). “An assessment of field evidence for ‘Byerlee’friction”. In: *Pure and Applied Geophysics* 142, pp. 645–662.

Sieh, K. (1986). “Slip rate across the San Andreas fault and prehistoric earthquakes at Indio, California”. In: *Eos Trans. AGU* 67.44, p. 1200.

Simpson, D., Leith, W., and Scholz, C. (1988). “Two types of reservoir-induced seismicity”. In: *Bull. Seism. Soc. Am.* 78, pp. 2025–2040.

Simpson, D. W. (1976). “Seismicity changes associated with reservoir loading”. In: *Engineering Geology* 10.2-4, pp. 123–150.

Snee, J.-E. L. and Zoback, M. D. (2022). “State of stress in areas of active unconventional oil and gas development in North America”. In: *AAPG Bulletin* 106.2, pp. 355–385.

Stein, R. S. (1999). “The role of stress transfer in earthquake occurrence”. In: *Nature* 402.6762, pp. 605–609.

Stokes, S. M., Ge, S., Brown, M. R., Menezes, E. A., Sheehan, A. F., and Tiampo, K. F. (2023). “Pore Pressure Diffusion and Onset of Induced Seismicity”. In: *Journal of Geophysical Research: Solid Earth* 128.3, e2022JB026012.

Takeuchi, C. and Fialko, Y. (2012). “Dynamic models of interseismic deformation and stress transfer from plate motion to continental transform faults”. In: *J. Geophys. Res.* 117, B05403, doi:10.1029/2011JB009056.

Talwani, P. (1997). “On the Nature of Reservoir-induced Seismicity”. In: *Pure appl. geophys.* 150, p. 20.

Tanaka, Y., Fujimori, K., and Otsuka, S. (1998). “In-situ stress measurement and prediction of great earthquake”. In: *Earthquake* 50.2, pp. 201–208.

Tao, W., Masterlark, T., Shen, Z.-K., and Ronchin, E. (2015). “Impoundment of the Zipingpu reservoir and triggering of the 2008 M_w 7.9 Wenchuan earthquake, China”. In: *J. Geophys. Res.* 120.10, pp. 7033–7047. ISSN: 2169-9313, 2169-9356.

Terzaghi, K., Mechanics, T. S., and Wiley, J. (1943). “Sons”. In: *New York*.

Thienen-Visser, K. van and Breunese, J. (2015). “Induced seismicity of the Groningen gas field: History and recent developments”. In: *The Leading Edge* 34.6, pp. 664–671.

Thomas, A. P. and Rockwell, T. K. (1996). "A 300- to 550-year history of slip on the Imperial fault near the U.S.-Mexico border: Missing slip at the Imperial fault bottleneck". In: *Journal of Geophysical Research: Solid Earth* 101 (B3), pp. 5987–5997. ISSN: 01480227.

Thompson, B. (2021). *cutde*.

Toda, S., Stein, R. S., Richards-Dinger, K., and Bozkurt, S. B. (2005). "Forecasting the evolution of seismicity in southern California: Animations built on earthquake stress transfer". In: *Journal of Geophysical Research: Solid Earth* 110.B5.

Toda, S., Stein, R. S., and Sagiya, T. (2002). "Evidence from the AD 2000 Izu islands earthquake swarm that stressing rate governs seismicity". In: *Nature* 419.6902, pp. 58–61.

Tompson, A., Demir, Z., Moran, J., Mason, D., Wagoner, J., Kollet, S., Mansoor, K., and McKereghan, P. (2008). *Groundwater Availability Within the Salton Sea Basin Final Report*. LLNL-TR-400426, 932394. LLNL, LLNL–TR–400426, 932394.

Tostrud, M. (1997). "The Salton Sea 1906-1996: computed and measured salinities and water levels". In: *Draft Report, Colorado River Board of California*, p. 84.

Townend, J. and Zoback, M. D. (2000). "How faulting keeps the crust strong". In: *Geology* 28.5, pp. 399–402.

Turcotte, D. L. and Schubert, G. (2002). *Geodynamics, 2nd ed.* New York, NY: 456 pp., Cambridge Univ.

USGS and California Geological Survey, C. (2019). *Quaternary fault and fold database for the United States*. https://www.usgs.gov/natural-hazards/earthquake-hazards/faults/. [Online; accessed 10-July-2019].

Van der Elst, N. J., Page, M. T., Weiser, D. A., Goebel, T. H., and Hosseini, S. M. (2016). "Induced earthquake magnitudes are as large as (statistically) expected". In: *Journal of Geophysical Research: Solid Earth* 121.6, pp. 4575–4590.

Verdecchia, A., Cochran, E. S., and Harrington, R. M. (2021). "Fluid-Earthquake and Earthquake-Earthquake Interactions in Southern Kansas, USA". In: *J. Geophys. Res.* 126.3, e2020JB020384.

Walsh III, F. R. and Zoback, M. D. (2015). "Oklahoma's recent earthquakes and saltwater disposal". In: *Science advances* 1.5, e1500195.

Wang, H. (2000). *Theory of Linear Poroelasticity: With Applications to Geomechanics and Hydrogeology*. Princeton, New Jersey: 287 pp., Princeton Univ. Press.

Wang, R., Schmandt, B., Zhang, M., Glasgow, M., Kiser, E., Rysanek, S., and Stairs, R. (2020). "Injection-induced earthquakes on complex fault zones of the Raton Basin illuminated by machine-learning phase picker and dense nodal array". In: *Geophysical Research Letters* 47.14, e2020GL088168.

Wang, W., Shearer, P. M., Vidale, J. E., Xu, X., Trugman, D. T., and Fialko, Y. (2022). "Tidal modulation of seismicity at the Coso geothermal field". In: *Earth and Planetary Science Letters* 579, p. 117335.

Waters, M. R. (1983). "Late Holocene Lacustrine Chronology and Archaeology of Ancient Lake Cahuilla, California". In: *Quaternary Research* 19.3, pp. 373–387. ISSN: 0033-5894, 1096-0287.

Weingarten, M., Ge, S., Godt, J. W., Bekins, B. A., and Rubinstein, J. L. (2015). "High-rate injection is associated with the increase in US mid-continent seismicity". In: *Science* 348.6241, pp. 1336–1340.

Weldon, R. J., Fumal, T. E., Biasi, G. P., and Scharer, K. M. (2005). "Geophysics - Past and future earthquakes on the San Andreas fault". In: *Science* 308, pp. 966–967.

White, J. A. and Foxall, W. (2016). "Assessing induced seismicity risk at CO2 storage projects: Recent progress and remaining challenges". In: *International Journal of Greenhouse Gas Control* 49, pp. 413–424.

Woo, J.-U., Kim, M., Sheen, D.-H., Kang, T.-S., Rhie, J., Grigoli, F., Ellsworth, W. L., and Giardini, D. (2019). "An in-depth seismological analysis revealing a causal link between the 2017 MW 5.5 Pohang earthquake and EGS project". In: *Journal of Geophysical Research: Solid Earth* 124.12, pp. 13060–13078.

Xue, L., Brodsky, E. E., Erskine, J., Fulton, P. M., and Carter, R. (2016). "A permeability and compliance contrast measured hydrogeologically on the San Andreas Fault". In: *Geochemistry, Geophysics, Geosystems* 17.3, pp. 858–871. ISSN: 15252027.

Zaliapin, I. and Ben-Zion, Y. (2013). "Earthquake clusters in southern California I: Identification and stability". In: *Journal of Geophysical Research: Solid Earth* 118.6, pp. 2847–2864.

— (2016). "Discriminating characteristics of tectonic and human-induced seismicity". In: *Bulletin of the Seismological Society of America* 106.3, pp. 846–859.

Zaliapin, I., Gabrielov, A., Keilis-Borok, V., and Wong, H. (2008). "Clustering analysis of seismicity and aftershock identification". In: *Physical review letters* 101.1, p. 018501.

Zbinden, D., Rinaldi, A. P., Urpi, L., and Wiemer, S. (2017). "On the physics-based processes behind production-induced seismicity in natural gas fields". In: *Journal of Geophysical Research: Solid Earth* 122.5, pp. 3792–3812.

Zhai, G., Shirzaei, M., Manga, M., and Chen, X. (2019). "Pore-pressure diffusion, enhanced by poroelastic stresses, controls induced seismicity in Oklahoma". In: *Proceedings of the National Academy of Sciences* 116.33, pp. 16228–16233.

Zoback, M. D. (2012). "Managing the seismic risk posed by wastewater disposal". In: *Earth* 57.4, p. 38.

Zoback, M. D. and Healy, J. H. (1992). “In situ stress measurements to 3.5 km depth in the Cajon Pass scientific research borehole: implications for the mechanics of crustal faulting”. In: *Journal of Geophysical Research: Solid Earth* 97.B4, pp. 5039–5057.

Zoback, M. D. and Townend, J. (2001). “Implications of hydrostatic pore pressures and high crustal strength for the deformation of intraplate lithosphere”. In: *Tectonophysics* 336.1-4, pp. 19–30.