Earthquake Nucleation Processes Across Different Scales and Settings

---

Dissertation
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
(Dr. rer. nat)
im Fach Geophysik

am Fachbereich Geowissenschaften
der Freien Universität Berlin

vorgelegt von
Stephan Bentz

Berlin, 2020

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Berlin, 05/09/2020

Referent:	Prof. Dr. Marco Bohnhoff
Korreferent:	Prof. Dr. Serge Shapiro
Tag der Disputation:	17.08.2020

Freie Universität Berlin
Fachbereich Geowissenschaften
Institut für Geologische Wissenschaften
Malteserstr. 74 - 100
12249 Berlin

Acknowledgements

First, I would like to thank my supervisor Marco Bohnhoff for inviting me to perform research in his working group and all the continuous support and funding throughout the different stages of my thesis. I want to especially thank you for introducing me to seismological fieldwork, especially in Turkey and entrusting me afterwards with complete autonomy to follow one of my favorite aspect of our scientific work.

To Pati goes all my gratitude, you have been one of the kindest, nicest, and most fun person to work with. You guided me from being a somewhat lost Bachelor student at a strange place called GFZ, to my Masters and throughout the PhD, until today. None of this would have been possible without you! As I already wrote in my MSc thesis, which became true back then (so fingers crossed again): I hope that this was once again just the beginning of many more productive years of friendship to come.

Many thanks also to Grzegorz, for all those times I was lost in the physics somewhere along the way and you took time and patience to drag me out of the mess that is theoretical seismology (and also for dragging me out of the snow in Switzerland). Thanks for all the nerdy discussions, ranging from spectral analysis, earthquake locations, and machine learning, to which graphic novel to read next.

Georg, I really enjoyed the last year working together on some *crazy ideas* about induced seismicity, trying to find out what the hell is actually going on there. Sometimes we (those nasty wigglefreaks) get lost in our favorite waveforms and we need someone to remind us that the earth is not just an elastic half-space, and waves travel through those pesky things called rocks. Thanks for bringing me back to my geomechanical roots and for reminding me, that after all, we have indeed the right to speculate.

Thanks to Peter for introducing me to borehole seismology, teaching me a lot about the necessary instrumentation and methodology. I always enjoyed the spirit of curiosity you bring to our group, as well as all the controversial yet immensely interesting ideas.

Rita, thank you for helping me settle at GFZ and with basically everything else. You are the beating heart of our section, please never change!

Without wanting to sound too melodramatic, but I consider the last years some of the most difficult ones of my life. Along the road, I struggled more than just scientifically, also personally, existentially, mentally, and physically. I am certain that I wouldn't have made it this far without the unwavering support of all my friends and colleagues.

I owe an unmeasurable debt to Vanessa and Livia, whose uncompromising strength, unconditional support, and love carried me through the roughest of patches. It's very hard for me to find words describing how grateful I am to have met you. When I first came to Potsdam, I would have never imagined to find such warmth and kindness. Though we had our ups and downs, you guys became my second family here, and I will always love you for that.

Thanks goes to many more friends and colleagues, who have been my pillars of strength

during the past three years: Hermi with all our very work related discussions on groundbreaking topics, Audrey for the countless fun nights (that's what she said), Lea for sharing my passion for crocodiles, Vivi for making all our Turkey trips much more entertaining by teaching me the fighting style of the native Corsicans, Bernhard for taking me to this amazing beergarden at Ostkreuz, Valerian for not stabbing me (yet), Maria for providing all those temporary escape from our PhD lifes during our coffee breaks, Claudius and Christopher for their jolly spirits and questionable taste in music, Amandine and Bita for pushing Karaoke on all of us, Io for letting me win some pool matches, Michael and Erik for all those fun and entertaining lunch breaks, and all the others who have made these difficult years also some of the most entertaining ones: Aglaja, Lei, Tina, Digdem, Taylan, Recai, Wade, Dave, Emrah, and all the countless fellow students I met at all those conferences and workshops.

Finally, I would like to thank my family for their continuous support getting me to this stage: from my very first thoughts about university subjects, to my first lecture concerning some dead cockroaches, till the end of my PhD.

Thank you all.

Abstract

Extended nucleation phases of earthquakes have been regularly observed, yet the underlying mechanisms governing the initiation phase of rupture are yet to be understood in detail. Currently two end member models exist to explain earthquake nucleation: one model claiming that the nucleation phase of a small earthquake is indistinguishable from that of a large one, while the other proposes fundamental differences in the underlying process. Previous studies have been using the same seismological observations to argue for either model, leaving the need of further investigations into the nucleation behavior of earthquakes across scales and different settings. The thesis at hand contributes to the current discussion on earthquake nucleation by providing additional observational evidence for extended nucleation phases, complex rupture interaction and growth across a number of different scales and settings. Here, earthquake nucleation is investigated for three different scenarios, each with varying degrees of complexity: 1) the *controlled* case of induced seismicity in hydraulic stimulations of geothermal reservoirs, where rupture growth is assumed to be primarily governed by anthropogenic activity, 2) the *partly-controlled* setting of a geothermal field with a long history of fluid injection and production, and 3) the *uncontrolled* case of natural seismicity in the central Sea of Marmara, where earthquake nucleation is purely governed by the regional tectonics.

First, the temporal evolution of seismicity and the growth of observed moment magnitudes for a range of past and present hydraulic stimulation projects associated with the creation of enhanced geothermal systems are analyzed. They reveal a clear linear relation between injected fluid volume/hydraulic energy and cumulative seismic moments. For most projects studied, the observations are in good agreement with existing physical models that predict a relation between injected fluid volume and maximum seismic moment of induced events. This suggests that seismicity results from a stable, pressure-controlled rupture process at least for an extended injection period. Overall evolution of seismicity is independent of tectonic stress regime and is most likely governed by reservoir specific parameters, such as the preexisting structural inventory. In contrast, a few stimulations reveal unbound increase in seismic moment suggesting that for these cases evolution of seismicity is mainly controlled by stress field, the size of tectonic faults and fault connectivity. The uncertainty over whether or not a transition between behaviour is likely to occur at any point during the injection is what motivates the need for a next generation monitoring and traffic-light system accounting for the possibility of unstable rupture propagation from the very beginning of injection by observing the entire seismicity evolution at high resolution for an immediate reaction in injection strategy. Furthermore, the majority of pressure-controlled stimulations shows the potential of actively controlling the size of induced earthquakes, if an injection protocol is chosen based on continuous feedback from a near-real-time seismic monitoring system.

Second, moderate sized earthquakes at The Geysers geothermal field (California), where

years of injection and production across hundreds of wells have led to a unique physical environment, are studied. While overall seismicity at The Geysers is generally governed by anthropogenic activities, contributions of individual wells or injection activities are hard to distinguish, thus making detailed managing of occurring magnitudes challenging. New high-resolution seismicity catalogs framing the occurrence of 20 $M_{\mathrm{L}} > 2.5$ earthquakes were created. The seismicity catalogs were developed using a matched filter algorithm, including automatic determination of P and S phase onsets and their inversion for absolute hypocenter locations with corresponding uncertainties. The selected 20 sequences sample different hypocentral depths and hydraulic conditions within the field. Seismic activity and magnitude frequency distributions displayed by the different earthquake sequences are correlated with their location within the reservoir. Sequences located in the northwestern part of the reservoir show overall increased seismic activity and low b values, while the southeastern part is dominated by decreased seismic activity and higher b values. Periods of high injection coincide with high b values and vice versa. These observations potentially reflect varying differential and mean stresses and damage of the reservoir rocks across the field. Additionally, a systematic search for seismicity localization using a multi-step cross-correlation analysis was performed. No evidence for increased correlation between the occurring seismicity and the mainshock for any of the 20 sequences could be seen, indicating that each main nucleation spot was seismically silent prior to the main rupture. However, a number of highly inter-correlated earthquakes for sequences below the reservoir and during high injection activity is observed. Under these conditions, the seismicity surrounding the future mainshock source region is more concentrated and might be evidence for a cascading nucleation process. About 50 % of analyzed sequences exhibit no change in seismicity rate in response to the large main event. However, we find complex waveforms at the onset of the main earthquake, suggesting that small ruptures spontaneously grow into or trigger larger events, consistent with a cascading type nucleation.

Third, the spatiotemporal evolution of seismicity during a sequence of moderate ($M_{\mathrm{W}}4.7$ and $M_{\mathrm{W}}5.8$) earthquakes occurring in September 2019 at the transition between a creeping and a locked segment of the North Anatolian Fault in the central Sea of Marmara (Turkey) was analyzed. A matched filter technique was applied to continuous waveforms from the regional network, substantially reducing the magnitude threshold for detection. Sequences of foreshocks preceding the two mainshocks are clearly seen, exhibiting different behaviors: a migration of the seismicity along the entire fault segment on the long-term and a concentration around the epicenters of the large events on the short-term. Suggesting that both seismic and aseismic slip during the foreshock sequences change the stress state on the fault, bringing it closer to failure. Furthermore, the observations also suggest that the $M_{\mathrm{W}}4.7$ event contributed to weaken the fault as part of the preparation process of the $M_{\mathrm{W}}5.8$ earthquake.

Combining the results obtained from different settings, it becomes apparent that, regardless of the tectonic setting and degree of *anthropogenic control* over the seismicity, there is a wide range of complex nucleation behaviours not yet explained by any of the current models of earthquake nucleation. A simplistic view of earthquake nucleation as either a deterministic or a stochastic process seems inconsistent with the obtained results and fails to account for a more complex nucleation behaviour. Observations from The Geysers and the western Sea of Marmara earthquake sequence, suggest that both cascade triggering and aseismic slip can play major roles in the nucleation of moderate sized earthquakes. Both mechanisms seem to jointly contribute to fault initiation, even within the same rock volume. A separation of the two mechanisms can potentially be thought of at The Geysers, where cascade triggering seems to dominate in highly damaged parts of the reservoir, suggesting that the anthropogenic activity can at least partially influence the nucleation behavior of the occurring seismicity. This would be in agreement with the results obtained from analysis of hydraulic stimulations, where during the pressure-controlled phase of injection rupture growth is controlled by the injected fluid.

Zusammenfassung

Nukleationsphasen von Erdbeben sind seit mehreren Jahrzehnten beobachtet worden, die zugrunde liegenden Mechanismen sind allerdings bis heute noch nicht vollständig verstanden. Derzeit existieren zwei Modelle, die die Nukleation von Erdbeben erklären. Dem ersten Modell zu Folge, ist der Beginn von kleinen und großen Erdbeben nicht zu unterscheiden, während das zweite Modell fundamentale Unterschiede in den jeweilig erforderlichen Prozessen sieht. Da die gleichen seismologischen Beobachtungen in früheren Studien dafür herangezogen wurden, für beide Modelle zu argumentieren, besteht weiterhin der Bedarf detaillierter Untersuchungen des Nukleationsverhalten von Erdbeben verschiedenster Skalen und Rahmenbedingungen. Die vorliegende Dissertation leistet einen Beitrag zur Diskussion über Prozesse, die zur Entstehung von Erdbeben führen unter Berücksichtigung weiterer Beobachtungen von verlängerten Nukleationsphasen und der komplexen Interaktion zwischen Störungen und Bruchausbreitung über verschiedene Skalen und strukturelle Bedingungen. Dabei werden drei unterschiedlich komplexe Szenarien untersucht: 1) der Fall induzierter Seismizität während der hydraulischer Stimulierungen von geothermalen Reservoiren, wobei die Bruchausdehnung primär durch Injektionsaktivitäten gesteuert wird, 2) ein Geothermiefeld mit einer langen und komplexen Injektions- und Produktionsgeschichte und 3) der Fall tektonischer Seismizität im Marmarameer, wo Erdbeben-Nukleation allein durch die lokalen und regionalen tektonischen Gegebenheiten bestimmt wird.

Die Betrachtung fluid-induzierter Seismizität, im Rahmen der Stimulierung geothermischer Reservoire, zeigt eine lineare Beziehung zwischen dem injizierten Fluidvolumen bzw. der eingebrachten hydraulischen Energie und dem kumulativen seismischen Moment. Für den überwiegenden Teil der untersuchten Projekte stimmen die Beobachtungen mit existierenden Modellen überein, die den Zusammenhang zwischen injiziertem Volumen und maximalem seismischen Moment vorhersagen. Dies suggeriert, dass ein Großteil der seismischen Aktivität während der Injektion das Resultat eines stabilen, druck-kontrollierten Bruchprozesses ist. Die Gesamtentwicklung von induzierten Erdbeben scheint unabhängig von den tektonischen Rahmenbedingungen zu sein, sondern wird wahrscheinlich von reservoir-spezifischen Parametern, wie etwa dem lokalen Störungsinventar, gesteuert. Im Gegensatz dazu zeigen einige wenige Stimulationen einen überproportionalen Anstieg des seismischen Moment. Dies deutet darauf hin, dass in diesen Fällen Seismizität durch das tektonische Spannungsfeld, die Größe von Störungen und das Störungsnetzwerk kontrolliert wird. Ein Übergang zwischen den beiden Fällen kann jederzeit auftreten, oder ganz ausbleiben. Diese Beobachtungen indizieren den Bedarf für eine neue Generation von Überwachungs- und Kontrollsystemen, welche die Möglichkeit einer instabilen Bruchausbreitung von Beginn der Stimulation an berücksichtigen. Da die meisten Projekte einen druck-kontrollierten Verlauf zeigen, deutet dies tendenziell auf die Möglichkeit hin, dass induzierte Seismizität kontrollierbar sein kann, wenn das Injektionsprotokoll durch seis-

mische Nah-Echtzeit-Systeme überwacht wird.

Im zweiten Fall werden moderate Erdbeben ($M_\mathrm{L} > 2.5$) im The Geysers Geothermie Field (Kalifornien, USA) untersucht, wo jahrzehntelange Injektion und Produktion über mehrere hundert Brunnen eine physikalisch einzigartige Umgebung geschaffen haben. Während die gesamte Seismizität in The Geysers hauptsächlich durch die anthropogenen Aktivitäten bestimmt wird, sind die Beiträge von einzelnen Bohrungen und Injektionsphasen schwierig zu differenzieren. Neue Erdbebenkataloge um das Auftreten von 20 $M_\mathrm{L} > 2.5$ Erdbeben sind mit Hilfe eines Matched Filter erstellt worden. Die 20 ausgewählten Sequenzen stammen gleichmäßig aus verschiedenen Tiefen und hydraulischen Bedingungen im Reservoir. Seismische Aktivität und die Magnituden-Frequenz Verteilung der verschiedenen Sequenzen werden mit ihrer Lage im Reservoir verglichen. Sequenzen im nordwestlichen Teil des Reservoirs zeigen erhöhte Erdbebenaktivität und niedrige b Werte, wohingegen der südöstliche Teil von niedriger Erdbebenaktivität und hohen b Werten dominiert wird. Perioden mit hohen Injektionsvolumen decken sich mit Zeiten hoher b Werte und umgekehrt. Diese Beobachtungen reflektieren potenziell unterschiedliche Differenzialspannungen und unterschiedlich stark deformiertes Reservoirgestein. Eine potentielle Erdbebenkonzentrierung wurde durch die systematische Anwendung von Kreuzkorrelation untersucht, dabei konnte jedoch für keine der 20 Sequenzen eine starke räumliche Konzentrierung von Mikroseismiziät um den Nukleationspunkt des Hauptbebens gefunden werden. Allerdings konnte eine kleine Zahl von untereinander sehr ähnlichen Wellenformen für Sequenzen unterhalb des Reservoirs, sowie während starker Injektionsaktivität beobachtet werden. Unter diesen Bedingungen scheint sie Seismizität konzentrierter aufzutreten, was auf eine Kaskadenwirkung im Entstehungsprozess der Erdbeben hindeutet. Ungefähr 50 % der ausgewerteten Sequenzen zeigen keine signifikante Abweichung von der seismischen Hintergrundaktivität als Reaktion auf das Hauptbeben. Die Hauptbeben dieser Sequenzen zeigen allerdings komplexe Wellenformen, was darauf hindeutet, dass kleine Brüche spontan in größere Erdbeben wachsen können oder diese auslösen.

Für den *unkontrollierten* Fall, bei dem Seismizität allein durch die tektonischen Gegebenheiten kontrolliert wird, ist die räumlich-zeitliche Entwicklung von Seismizität während einer Sequenz vor, während und nach zwei moderaten ($M_\mathrm{W}4.7$ und $M_\mathrm{W}5.8$) Erdbeben im September 2019 untersucht worden. Diese Sequenz trat an dem vermuteten Übergang zwischen einem seismisch verhakten und einem kriechenden Teil der Nordanatolischen Verwerfungszone auf. Ein Matched Filter Algorithmus ist verwendet worden, um die Detektionsschwelle des regionalen Breitbandseismometer Netzwerks signifikant zu reduzieren. Für die beiden größten Erdbeben sind Sequenzen von Vorbeben mit unterschiedlichem Verhalten identifiziert worden: eine langfristige Migration der Seismizität entlang des gesamten Störungssegment, sowie eine kurzfristige Konzentration um die Epizentren der großen Erdbeben. Dies suggeriert, dass sowohl aseismischer als auch koseismischer Versatz während der Vorbebensequenz das Spannungsfeld der Störung verändert haben und diese näher zum Versagen gebracht haben. Weiterhin deuten die Beobachtungen daraufhin, dass das $M_\mathrm{W}4.7$ Beben zu einer Schwächung der Störung führte und somit die Entwicklung des $M_\mathrm{W}5.8$ Hauptbebens begünstigte.

Die vergleichende Analyse zeigt, dass eine Vielzahl von komplexen Nukleationsverhalten für Erdbeben verschiedenster Skalen beobachtet werden kann, unabhängig der tektonischen Rahmenbedingungen und dem anthropogenen Einfluss auf die Seismizität. Ein auf existierenden Modellen basierter Ansatz, um die Nukleation von Erdbeben als simplen deterministischen oder stochastischen Prozess zu erklären, widerspricht den Beobachtungen, welche eine weitaus komplexere Interaktion unterschiedlicher Mechanismen vermuten lassen. Die Ergebnisse von The Geysers und aus dem Marmarameer lassen darauf schließen, dass die Entstehung von Erdbeben sowohl durch Kaskadenwirkung als auch durch aseismische Deformation bestimmt werden können. Weiterhin zeigen lokale Unterschiede im Nukleationsverhalten von moderaten Erdbeben in The Geysers, dass anthropogene Aktivitäten das Entstehungsverhalten von Erdbeben beeinflussen können. Dies ist auch im Einklang mit der Analyse von hydraulischen Stimulationen, wo in der druck-kontrollierten Phase der Injektion die Bruchausdehnung primär durch das injizierte Volumen kontrolliert wird.

Statement of contributions

The main part of the thesis presented here consists of three peer-reviewed papers that have been published in scientific journals:

1. "Analysis of Microseismicity Framing $M_L > 2.5$ Earthquakes at The Geysers Geothermal Field, California." *Journal of Geophysical Research: Solid Earth* 124 (2019).
   Authors: **S. Bentz**, P. Martínez-Garzón, G. Kwiatek, G. Dresen, M. Bohnhoff

2. "Seismic Moment Evolution during Hydraulic Stimulations." *Geophysical Research Letters* 47(5) (2020).
   Authors: **S. Bentz**, G. Kwiatek, P. Martínez-Garzón, M. Bohnhoff, G. Dresen

3. "A two-scale preparation phase preceded a $M_W 5.8$ Earthquake in the Sea of Marmara offshore Istanbul, Turkey." *Seismological Research Letters.*
   Authors: V. Durand, **S. Bentz**, G. Kwiatek, G. Dresen, C. Wollin, O. Heidbach, P. Martínez-Garzón, F. Cotton, M. Nurlu, M. Bohnhoff

In publication (1) (Bentz et al., 2019) I performed all the waveform processing, data analysis, and developed most of the used software during the study. Furthermore, I created all figures, wrote the major part of the manuscript, and coordinated the writing of the discussion and paper in general. The absolute location procedure was developed and written by Grzegorz Kwiatek and implemented and applied by me.

In publication (2) (Bentz et al., 2020) I functioned as lead author, collected and analyzed all datasets, created all figures, and wrote major parts of the manuscript. The discussion of the results was a collaborative effort by all authors.

For publication (3) (Durand et al., 2020) I was responsible for the waveform processing (including manual phase onset determination and template matching/detection location), waveform cross-correlation search for earthquake repeaters, source-time function inversion, computed the final earthquake catalogs, and compilation of several figures. Grzegorz Kwiatek performed the aboslute and relative template localization using my provided phase onsets. Both, Grzegorz Kwiatek and me, jointly wrote the data and methodology section of the manuscript. Virginie Durand performed the spatio-temporal analysis of the earthquake sequence, led the interpretation of the results in collaboration with all authors, wrote major parts of the manuscript, created most of figures, and merged individual contributions from the coauthors.

Declaration

The PhD project presented here was partly funded via the Helmholtz Centre Potsdam, German Research Centre for Geosciences GFZ and the Helmholtz Association through the Young Investigators Group "Seismic and Aseismic Deformation in the brittle crust: implications for Anthropogenic and Natural hazard" (VH-NG-1232, led by Dr. Patricia Martínez-Garzón).
Analysis throughout the thesis was mostly performed using the Matlab® and/or Python environment. Codebase developed for the matched filter algorithm was written in C/C++, uses some parts of the SAC I/O IRIS library, and was based on older versions of the time domain CPU matched filter code by Zhigang Peng. The high-performance cluster facilities of the Helmholtz Centre Potsdam, German Research Centre for Geosciences GFZ were used to run the matched filter algorithm.

Contents

1 Introduction

A dark room, a scientist looking closely at the monitors, the ground begins to rumble, first slow and steady, followed by intense shaking, buildings collapsing, widespread panic, and finally the earth splitting apart... this is how earthquakes are typically shown in popular disaster movies. This grotesque depiction is built on the general fear but also on the fascination for these poorly understood phenomena. How far are these scenes, however, from reality? As seismologists, we are occasionally sitting in our dark offices, looking at waveforms and following the news on another monitor, although we rarely see earthquakes unfolding before our eyes. Past and present civilizations have been literally shaken to their foundations by devastating earthquakes, costing countless lives and inflicting unimaginable pain on those who were lucky enough to survive. Yet, out of the deadliest natural disasters, earthquakes are for sure one of the least understood or managed. While we can anticipate secondary effects (such as tsunamis) after the quake, as well as occasionally provide warnings in the range of seconds or minutes to shut down critical infrastructure before the waves hit (so called *early warning*), earthquakes remain in a deterministic sense unpredictable.

Our only defense in place against earthquakes is a long-term forecast based on a probabilistic assessment of the seismic hazard present at a given place, indicating which areas need to be "prepared" for the potential occurrence of large earthquakes, through stricter building codes, public outreach and education, and preparation of relief efforts. This hazard assessment is based on the knowledge of present faults accumulating strain. Even though subduction zones are capable of hosting some of the largest magnitude earthquakes ever recorded (M ∼9), they are typically occuring at greater distances to cities and critical infrastructure, translating to reduced risk from the earthquake itself. Intercontinental transform faults, on the other hand, can produce significantly smaller earthquakes (M ∼7) at substantially higher risk to human lifes, as several of these faults lie below major population centers (such as San Francisco, Los Angeles and Istanbul). Even though, the largest earthquakes are considered overdue in the seismic cycle of these major fault zones, the ground can by no means be considered seismically quite. Thousands of small earthquakes are registered each year by large surface and borehole seismometer networks, deployed with the aim of imaging the subsurface structure and deriving the state of stress on the faults. Extensive observations have made it clear that small earthquake sequences follow very much the same rules and statistical patterns as larger ones (e.g. Papazachos, 1975; Jones & Molnar, 1979; Savage et al., 2017; Malin et al., 2018; Trugman & Ross, 2019). Thus, using small and moderate-sized earthquake sequences as a representation for large earthquakes not yet experienced, provides a way to understand how earthquakes initiate, propagate, and interact with each other, topics heavily debated among the scientific community, with immense implications for the entire human society. An important point still to address is the question of the predictability of earthquakes: are they truly the result of a random unpredictable process, or do they exhibit a systematic pattern that can be predicted? Recent years have witnessed an incredible advancement in the seismological observation of precursory processes and extended preparation phases. These observations have fueled an *earthquake nucleation renaissance* in the seismological community, providing us with much needed information to further our knowledge of the underlying processes associated with earthquake initiation.

1.1 Motivation

What mechanism drives the nucleation of large magnitude earthquakes? A simple question with a wide range of complex answers. While we know that earthquakes result from the release of accumulated elastic energy along a preexisting fault plane in the earth's brittle crust (wherever the stored strain is sufficiently high), how the earthquake initiates, and what factors allow it to grow to a certain size is still heavily debated (Ellsworth & Beroza, 1995; Bouchon et al., 2011; Ellsworth & Bulut, 2018; Gomberg, 2018). Currently two models have been proposed for the nucleation of earthquakes, both able to explain seismic observations (Ellsworth & Beroza, 1995): 1) the *cascade* model, suggesting that earthquake initiation is a random cascading triggering process of spontaneous small subevents growing into larger ones, eventually leading to a delayed failure of the entire fault. 2) The *preslip* model, where earthquake initiation is preceded by aseismic slip on parts of the faults, and individual subevents are constrained to the area of preslip. Preferring one model over the other has significant implications, not only for the physics driving earthquake initiation, but also for the potential interpretation and usage of foreshocks and other precursory signs for any kind of earthquake forecasting or prediction. In the cascade model, there is no inherent difference between the beginning of small and large earthquakes, and any small earthquake has the potential of triggering a cascade of increasingly larger slip events (Ellsworth & Beroza, 1995). Following this interpretation, the seismic nucleation phase is a stochastic process of successive failure of growing fault elements, and the final rupture size is determined by the last jump in the size of fault elements. This interpretation precludes any possibility to anticipate how large an earthquakes will be, as sequences not producing large events are indistinguishable from sequences that do. The preslip model on the other hand suggests that the beginning of small and large earthquakes is different and failure for large events begins with a period of slow and stable sliding over a limited area. Once the slipping patch reaches a critical size, the process becomes unstable and the fracture propagates away from the nucleation spot. In this deterministic model, the seismic nucleation phase is a mechanical response to a process already in progress, thus providing the opportunity for at least short-term earthquake forecasting, if that underlying process can be effectively identified and measured.

Over the last decade, the existence of slow slip during tectonic earthquake sequences has been well established (e.g., Lohman & McGuire, 2007; Peng & Zhao, 2009; Roland & McGuire, 2009; Kato & Nakagawa, 2014; Vuan et al., 2018), but whether it also governs the nucleation phase is still an open question. Distinguishing different mechanisms related to the nucleation of earthquakes is largely based on the temporal and spatial evolution of foreshocks in comparison to the hypocenter of the mainshock. However, favoring either the cascade or preslip model using high-resolution waveform data from large earthquakes has proven to be difficult, as even for the same dataset the interpretation given in different studies vary (e.g., Bouchon et al., 2011; Ellsworth & Bulut, 2018). In that regard, Gomberg (2018) highlights the current non-uniqueness of plausible interpretations, arguing that arguments brought in favor of one model can easily be turned into arguments for the other model... So where does this lead us?

By now, these relatively simple models trying to explain earthquake nucleation are more than 25 years old, yet no additional model has been suggested or broadly discussed since then and the controversy between the two continues. Is earthquake nucleation purely governed by one of the two models and their related processes? Is juxtaposition of the two a more realistic possibility? Or are we still considering a rather simplistic view on earthquake nucleation as either a deterministic or stochastic process, and missing a more complex interaction between several processes, local and global, coming together? In this thesis I use high-resolution earthquake catalogs for moderate sized earthquakes in different tectonic and anthropogenic settings allowing to investigate whether smaller magnitude earthquakes exhibit the same complex failure process as large magnitude earthquakes (contrary to what the *preslip* nucleation model postulates). I also intend to analyze which factors influence the characteristic nucleation phase of earthquakes, and whether precursory signs, of any type, are pervasive across scales and settings. Finally, I also aim to adress the question of whether earthquake nucleation processes can be identified as either stochastic, deterministic, both, or neither.

These key objectives are complemented by observations of rupture evolution in a more controlled environment, where the conditions driving seismicity are artificially constrained, thus bridging the gap from laboratory/in-situ experiments to the tectonic field scale. By considering single fluid-injection projects, where the hydraulic energy inserted into the system seems to be the major driving force of seismicity evolution and growth, it becomes possible to directly relate failure initiation and rupture growth with the underlying driving mechanism. Besides shedding light on the fundamental physical processes governing fault initiation and the early nucleation phase of earthquakes, investigation of large induced ruptures has broad socio-economic implications. In recent years this topic has received wide spread public attention as large induced earthquake related to the hydraulic stimulation of several enhanced geothermal systems (EGS) have led to the termination or suspension of these latter, most prominently in the case of the geothermal projects in Basel (Giardini, 2009) and St. Gallen (Diehl et al., 2017), and most recently for the EGS site near Pohang, South Korea (Ellsworth et al., 2019). Understanding how earthquakes related to the injection activity grow poses a fundamental question in controlling injection induced seismicity, as well as a significant challenge in the global application of geothermal energy as a possible replacement for conventional hydrocarbons. Providing an overview of seismicity growth for prominent EGS sites, identifying fundamental differences between projects that resulted in large magnitude events and those that did not could potentially provide the basis for the next generation of near-real time monitoring systems for induced seismicity related to fluid injection.

1.2 Outline

This thesis is mostly based on two peer-reviewed publications in ISI journals (Bentz et al., 2019; 2020) and one manuscript currently under consideration for publication (Durand et al., submitted). Whenever applicable electronic supplements to the publications have been included into the main body of the corresponding chapter in order to fit better into the frame of joint publication in this thesis.

Chapter 2 presents a general framework of the theory and implementation on waveform cross-correlation techniques largely used in chapters 4 and 5 to improve detection capabilities of seismic networks. More extended applications of cross-correlation, as well as shortcomings and alternatives are explained in the later part of this section. The application of the methodology to specific dataset is further discussed in each respective chapter.

Chapter 3 discusses the seismic moment evolution during different hydraulic stimulation projects related to the development of enhanced geothermal systems and one scientific drilling site. Maximum observed seismic moment and the cumulative seismic moment were temporally related to operational parameters such as the cumulative volume injected, the hydraulic energy applied, and injection efficiency. Furthermore, the observations are discussed in the frame of current models trying to relate seismicity with injection parameters. [Published as Bentz et al., 2020]

Chapter 4 presents an extensive analysis of microseismicity patterns surrounding the occurrence of moderate sized induced earthquakes at The Geysers geothermal field, California. Cross-correlation template matching was used to extend the local catalog in the timeframes surrounding 20 $M_{\mathrm{L}} > 2.5$ events. Different sequences were statistically compared across the reservoir, depending on injection parameters and location in the reservoir. A systematic search for precursory signs indicating extended nucleation phases of the mainshocks was performed, including increased seismic activity or spatial correlation, b value variations, and foreshock evolution. [Published as Bentz et al., 2019]

Chapter 5 investigates the spatial and temporal evolution of seismicity during a moderate sized earthquake sequence in the central Sea of Marmara, Turkey. The matched filter was used to significantly increase the seismic catalogs related to the sequence, enhance the spatial resolution of fore- and aftershock patterns, and search for transient repeater events. The obtained seismic catalog was used to investigate the early nucleation phase of the mainshock, its possible driving mechanisms, as well as the interevent interaction between the two largest earthquakes.

Chapter 6 synthesizes the results presented in the previous chapters by summarizing and discussing the combined implications of the individual studies presented here, and provides an outlook on future research directions building upon the results of this thesis.

2 Maximizing Earthquake Detection Through Computationally Efficient Cross-Correlation

It is has been long established, that in signal processing one of the most effective ways of isolating and identifying an incoming signal in a stream of noise polluted data is to cross-correlate the entire time-series with a highly similar signal (e.g., van Trees, 1968). The resulting cross-correlation function of the continuous data stream will display a high value at the onset of the target signal, indicating the high degree of similarity. This method is referred to as *North filter* or more commonly *Matched filter*. Because of the necessity of an a priori well-known signal to detect another, the technique has only found its way into a seismological application in recent decades (e.g., Gibbons & Ringdal, 2006; Shelly et al., 2007). The basic concept, however, has been proven and well established since the middle of the 20*th* century (Turin, 1960). Matched filter techniques have been, and are currently still being used regularly in radar application, where a known signal is sent out by the antenna and the received reflected signal is searched for common characteristics with the outgoing signal (Woodward, 1953).

Wide implementation into seismology has only recently begun, since most recorded signals have an apriori unknown origin. Thus, most initial detectors for worldwide or local seismometer networks are based on some kind of energy or signal envelope. The most common one being the short-term average (STA) and long-term average (LTA) trigger (e.g., Freiberger, 1963), where the energy in a short time-window is compared against a long time-window moving through the continuous data. After exceedance of a fixed ratio threshold, an event detection is declared. Application of energy based detectors are often a good initial processing step, but typically suffer in the detection of low signal-to-noise ratio (SNR) events (e.g., Kwiatek & Ben-Zion, 2016). Weaker signals may be identified by array-techniques such as the stacking (beamforming) of data recorded by closely located sensors, resulting in coherent stacked higher SNR recordings (e.g., Rost, 2002; Oye et al., 2004; Ben-Zion et al., 2015). However, densely spaced arrays constitute only a small part of the global seismic networks, thus severely limiting the ability to detect low SNR events using traditional energy detectors.

If a region exhibits a sufficient number of earthquakes, which have been previously detected and well located, their application as templates in a matched filter can significantly improve the resolution of the associated seismic catalog. Typically this procedure is performed across an entire seismic network, where the recorded data share similar durations and frequency content at all stations (e.g., Gibbons & Ringdal, 2006; Frank & Abercrombie, 2018). Initially, Shearer (1994) used a matched filter technique on long-period seismograms to detect missing large magnitude earthquakes in the southern ocean, which were not detected due to the operational gap of most global networks. Peng and Zhao (2009) used template matching to detect 11 times more aftershocks than listed in the standard catalogue related to the 2004 $M_{\mathrm{W}}6.0$ Parkfield earthquake, identifying aftershock migration patterns as primarily caused by propagating afterslip of the mainshock. Shelly et al. (2007) investigated the mechanisms of non-volcanic tremor and low-frequency earthquake swarms beneath Shikoku, Japan. Using a catalog created with a matched filter algorithm, they suggested that tremor and slow-slip are different proxies for the same process. Meng et al. (2012) detected approximately 70 times the number of earthquakes listed in the Southern California Earthquake Catalog surrounding the $M_{\mathrm{W}}7.2$ El Mayor-Cucapah earthquake, capturing significant seismicity rate changes due to both dynamic and static stress changes. In another study, Aiken et al. (2018) investigated potential dynamic triggering at The Geysers geothermal field by applying a matched filter analysis, lowering the magnitude of completeness by one full order. Using recent advancement in computing power and the increase usage of GPU over CPU clusters Z. E. Ross et al. (2019) performed a template matching on the entire Southern California Seismic Network, utilizing over 300,000 template waveforms, resulting in a more than tenfold increase of earthquake detections. The resulting catalog likely constitutes the most detailed regional earthquake catalog to date. Using this catalog, Trugman and Ross (2019) were able to identify pervasive foreshock activity across Southern California, suggesting that extended nucleation phases including foreshock activity are much more common than previously thought.

A large number of different implementations of the matched filter algorithm currently exist (e.g., Meng et al., 2012; Beaucé et al., 2017; Senobari et al., 2018). In the following, I present the underlying theory behind matched filter, also known as *template matching*, as well as the practical and computational implementation developed and applied in the different parts of the thesis. Furthermore, I discuss the necessary post processing steps, which have been efficiently streamlined during the thesis, following the application of a matched filter in order to obtain a detailed seismicity catalog. Finally, I present further scientific application of the correlation calculations performed during the detection process, as well as, discuss shortcomings and alternatives to template matching.

2.1 Theoretical Background

For two functions f and g their cross-correlation is defined as:

$$(f \star g)(\tau) \triangleq \int_{-\infty}^{\infty} \overline{f(t)} g(t+\tau) dt, \tag{2.1}$$

where $\overline{f(t)}$ indicates the complex conjugate of $f(t)$ and τ the time displacement, known as the correlation lag (Papoulis, 1962). For discrete functions equation (2.1) changes to (Rabiner & Schafer, 1978):

$$(f \star g)(n) \triangleq \sum_{m=-\infty}^{\infty} \overline{f(m)} g(m+n). \tag{2.2}$$

Equation (2.1) and (2.2) are very similar to the convolution of the same two functions f and g (Damelin & Miller, 2011), where the convolution may be expressed through:

$$(f * g)(t) \triangleq \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau, \tag{2.3}$$

and for discrete functions as:

$$(f * g)(n) \triangleq \sum_{m=-\infty}^{\infty} f(m)g(n-m). \tag{2.4}$$

Thus from equation (2.1) and (2.3) it becomes apparent that the cross-correlation of f and g is equal to the convolution of the time-reversed complex conjugate of f and g:

$$(f(t) \star g(t))\,(t) = \left(\overline{f(-t)} * g(t)\right). \tag{2.5}$$

Now, consider $u_{N,\Delta t}(t_u)$ a vector of the time series u with N samples, where t_u is the starting time and Δt is the sample interval:

$$u_{N,\Delta t}(t_u) = \left(u(t_u), u(t_u + \Delta t), ..., u(t_u + (N-1)\Delta t)\right). \tag{2.6}$$

Following the convention of Gibbons and Ringdal (2006), the inner product between $u_{N,\Delta t}(t_u)$ and any other arbitrary non-zero time series $v_{N,\Delta t}(t_v)$ is defined by

$$\begin{aligned} \langle u(t_u), v(t_v)\rangle_{N,\Delta t} &= \langle u_{N,\Delta t}(t_u), v_{N,\Delta t}(t_v)\rangle \\ &= \sum_{i=0}^{N-1} u(t_u + i\Delta t)v(t_v + i\Delta t). \end{aligned} \tag{2.7}$$

Then from equation (2.2) and (2.7) it follows, that the normalized cross-correlation between $u_{N,\Delta t}(t_u)$ and any other arbitrary time series $v_{N,\Delta t}(t_v)$ is expressed through

$$C(u(t_u), v(t_v))_{N,\Delta t} = \frac{\langle u(t_u), v(t_v)\rangle_{N,\Delta t}}{\sqrt{\langle u(t_u), u(t_u)\rangle_{N,\Delta t}\langle v(t_v), v(t_v)\rangle_{N,\Delta t}}}. \tag{2.8}$$

$C(u(t_u), v(t_v))_{N,\Delta t}$ is called the cross-correlation coefficient of $u_{N,\Delta t}$ at time t_u and $v_{N,\Delta t}$ at time t_v, i.e. at zero lag to each other (Gibbons & Ringdal, 2006). Due to the normalization through the denominator in equation (2.8) the coefficient will always lay in the interval between -1 and +1. Both extremes can only be achieved if $u_{N,\Delta t}$ is equal to $v_{N,\Delta t}$, which is called the auto-correlation of $u_{N,\Delta t}$. For real seismological data, depending on the frequency content and length of the compared time-series, a value larger than 0.7 can already indicate very similar waveforms, whereas a value larger than 0.9 suggests that the two waveforms originate from nearly the same location, caused by a rupture with almost the same mechanism and duration of the source-time function (Israelsson, 1990). Geller and Mueller (1980) recognized that the degree of similarity, which is attributed to the source location of two waveforms, is controlled by the dominant frequency band used in the correlation. They observed that for similar waveforms, their respective origin may not be separated by more than a quarter of the dominant wavelength. This suggests, that at least partially the frequency content of the signal controls the degree of waveform similarity obtained by cross-correlation. Formally, this means that $u_{N,\Delta t}(t_u)$ and $v_{N,\Delta t}(t_v)$ should only contain the frequency bands that are to be compared. In practice, this often results in applying either a high-pass or band-pass filter to both time-series before correlating in the frequency band of interest. Note that, if an insufficient filter is applied, any unwanted underlying linear or constant trend of both $u_{N,\Delta t}(t_u)$ and $v_{N,\Delta t}(t_v)$ should be removed before the correlation can be calculated without errors, e.g. $u_{N,\Delta t}(t_u) = u_{N,\Delta t}(t_u) - \frac{1}{N}\sum_{i=1}^{N} u_i(t_i)$.

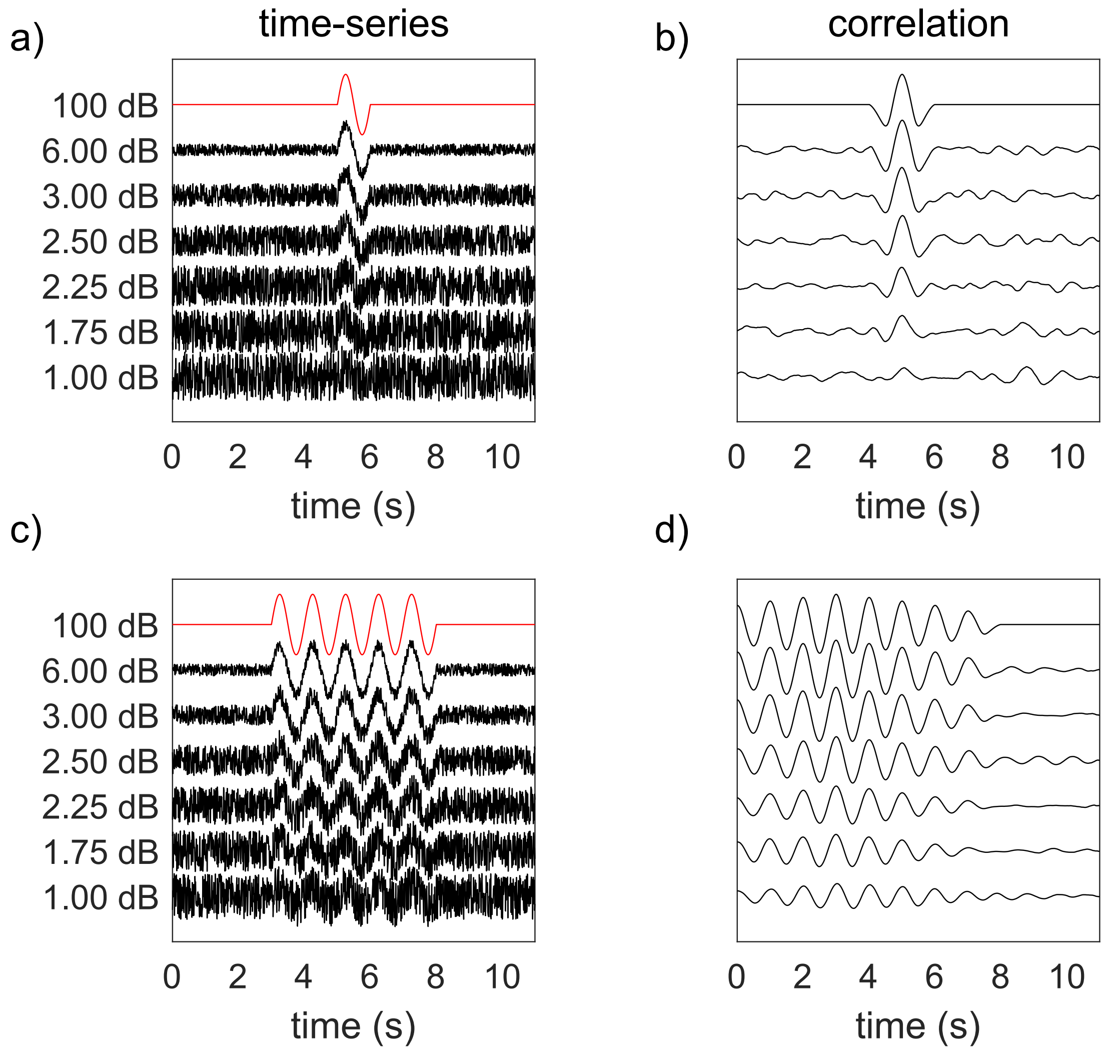

Figure 2.1: Comparison of the influence of different input signal SNR and length on the correlation. (a) Input time-series of one period 1 Hz sine wave with different SNR ratios. (b) Correlation time-series of each trace in (a) with the original sinusoid (red line in (a)). (c) Input time-series of (a) now extended to contain 5 periods. (d) Correlation time-series of each trace in (c) with the original sinusoid (red line).

The application of cross-correlation as a signal detector is conceptually shown in Figure 2.1. Consider one period of a simple sinusoid embedded in continuous noise (Figure

2.1a): The corresponding normalized cross-correlation is computed using equation (2.8), where $u_{N,\Delta t}$ represents the template sine waveform, and $v_{N,\Delta t}$ represents N samples of the continuous noisy data. A time series containing cross-correlation coefficients, called a *correlogram*, is obtained by shifting the correlated part of the data $v_{N,\Delta t}$ sample by sample, until all the continuous data has been correlated with the template (Figure 2.1b). For the single sinusoid, the correlation can help identifying the signal buried in noise, up until very low SNR. Now, considering more periods of the same signal buried in approximately the same random noise (Figure 2.1c), the detection capabilities of template matching further increases. For a buried signal of 5 periods of 1 Hz sine waves, the presence and location of the signal can accurately be matched down to a SNR of 1 dB (Figure 2.1d), whereas a single period signal stops to be visible at larger SNR (Figure 2.1b). This highlights the influence of frequency and time window selection for correlation in general, and for template matching specifically, which needs to be carefully evaluated, as an arbitrary choice may have significant impact on the results obtained.

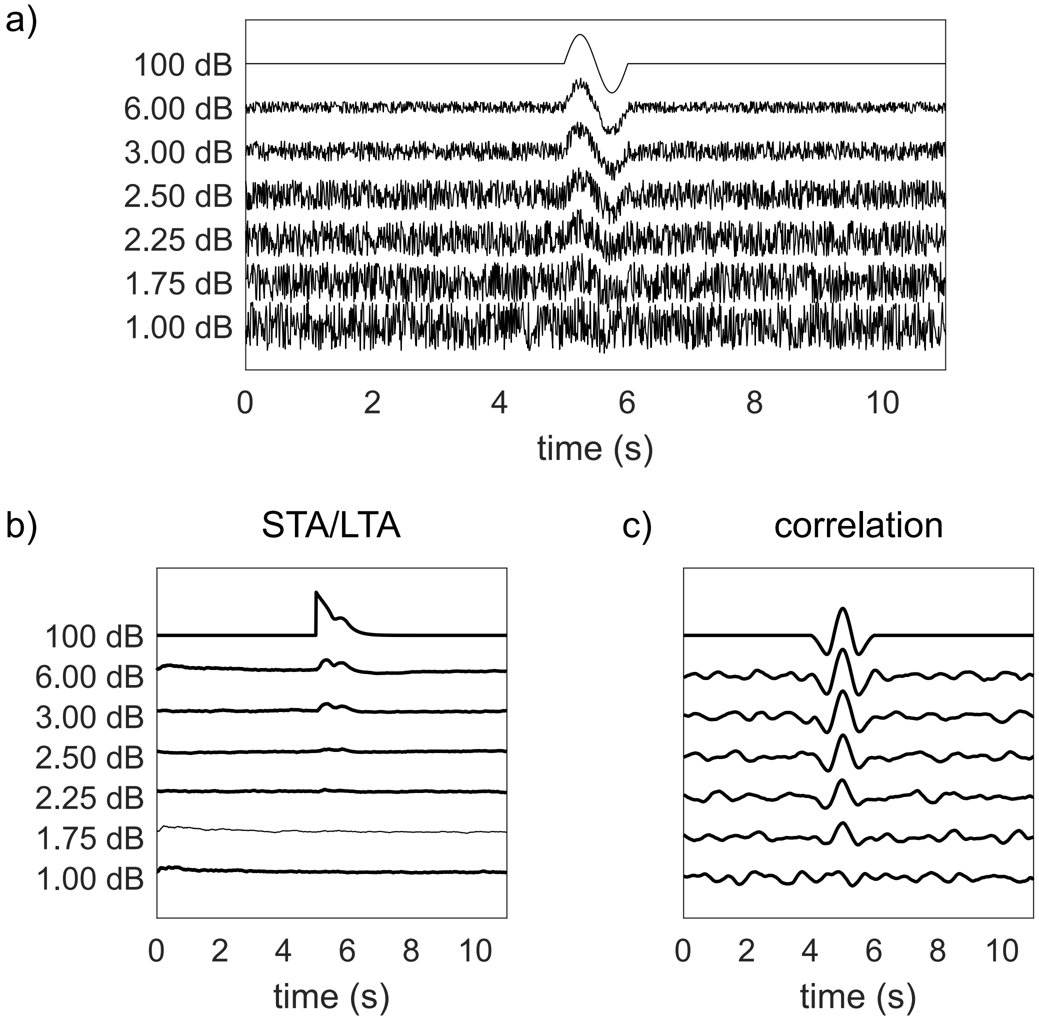

Figure 2.2: Comparison of a traditional energy based STA/LTA detector and the template matching. (a) Input time-series of one period 1 Hz sine wave with different SNR ratios. (b) The STA/LTA ratio when applying an STA of 0.25 s and LTA of 2 s on the input data. (c) Correlation time-series of each trace in (a) with the original sinusoid as template.

Comparing the performance of template matching against an energy based STA/LTA trigger on a simple example shows the clearly superior detection capabilities of the matched filter (Figure 2.2). For the same sinusoid as in Figure 2.1a, again buried in increasing levels of noise (Figure 2.2a) the application of an STA/LTA detector, already carefully calibrated to the dominant frequency of the signal, on the raw waveform results in a clear detection only when the noise level is very low (Figure 2.2b). A SNR of 6 dB already significantly reduces the ratio between STA and LTA, in our example, below a chosen threshold of 3, which would no longer trigger the detection. Using a template waveform to cross-correlate across the noisy data (Figure 2.2c), a clear peak in the correlogram is visible down to much lower SNR.
The example in Figure 2.2c brings up a crucial problem in the matched filter algorithm: when do we declare the detection of a new signal based on the correlogram? For an energy based detector, a simple threshold value is sufficient, as we expect the signal to contain a certain amount of energy. A combination of both systems could be imagined, where first template matching is performed and then an STA/LTA trigger is run over the signals exceeding a certain correlation coefficient. However, not all events that are similar to any given template contain enough energy to make them detectable through energy or stand out from the background noise, which constitutes the basic reasoning behind template matching. Applying the STA/LTA trigger on the correlogram rather than the waveform itself would be a more reasonable approach, but fine-tuning the trigger would proof difficult, as there are no a priori information available on how similar certain signals in a trace are, or how wide the correlation peak in any given correlogram might be. Choosing a correlation coefficient as a fixed threshold to declare detection, would circumvent the problem of the unknown correlogram complexion. However, as shown in Figure 2.1, the correlation coefficient can be substantially influenced by the choice of frequency band and time length of the two correlated signals. The significance of a high correlation coefficient depends primarily on the complexity of the chosen template. A seismogram recorded in the near-field for a small magnitude earthquake compared to a larger earthquake with the same hypocenter and source mechanism will potentially result in low raw cross-correlation coefficients, nevertheless there would be significant peak in the correlation function. Therefore, detection declaration should depend on each template and its correlation coefficient to all data. Meaning, that if the correlation coefficient of a template waveform with a certain part of the data is substantially elevated compared to the correlation coefficients of the surrounding data this represents a significant change in the data, independent of the numerical value assigned to that change. Most commonly, a multiple x of the median absolute deviation (MAD) is used to identify such dynamic changes in cross-correlation coefficient (Shelly et al., 2007), where the MAD is defined by:

$$\mathrm{MAD} = \mathrm{median}(|C_i - \mathrm{median}(C)|). \tag{2.9}$$

The determination of the multiple x of the MAD depends largely on the data quality and noise level present, but values between 7 and 9 are typical (e.g., Meng et al., 2012; Meng & Peng, 2014; Bentz et al., 2019). The probability of a normally distributed random variable exceeding x times the MAD can be calculated as:

$$f(x) = \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{0.67449x}{\sqrt{2}}\right)\right), \tag{2.10}$$

which results in a theoretical false detection probability of $\sim 3.3 \times 10^{-8}$ for a detection threshold of 8 times the MAD.

2.2 Computational Implementation

The Matched Filter Algorithm (MFA) is applied in seismology by cross-correlating a known waveform, i.e. the template, with a continuous trace of ground-motion recordings. This may be done directly on the incoming trace from an online sensor, that is, live detection, or through applying it on already recorded data. In this thesis, I developed and applied the MFA for the latter case, utilizing high-performance computing power to scan through continuous data. Before this action can be undertaken, however, there are a number of crucial pre-processing steps to prepare the data and reduce future computation time (Figure 2.3).

In order to be valid for the normalized cross-correlation, equation (2.8) requires the so called zero-mean input, meaning that the data needs to be detrended and the overall mean should be close to zero. This condition must be fulfilled by both the template and the targeted continuous waveform. Additionally, a consistent bandpass filter is applied to all input data, where the frequency band should be chosen according to the dominant frequency content of events that are to be detected as well as the templates. Typically, template magnitudes are at least 2 orders larger than the events to be detected, making it potentially difficult to find a consistent frequency band between templates and detections. Nevertheless, for small magnitude events ($-1 < M_\mathrm{W} < 0.5$) I chose a second-order Butterworth bandpass filter between $4 - 40\,\mathrm{Hz}$, which has proven to work efficiently in resolving microseismicity patterns previously missed (Bentz et al., 2019). For larger magnitudes ($0.5 < M_\mathrm{W} < 2$), I decreased the band to $1 - 10\,\mathrm{Hz}$. However, for coherency, the entire workflow should be repeated several times with different frequency bands each time, ensuring that detection limitations are not imposed by the frequency band. Following detrending and filtering, the data may be down sampled, depending on the frequency band chosen and the sample rate of the sensors (Figure 2.3). Equation (2.8) needs to be calculated as often as there are samples in the continuous trace, thus using data sampled with a high frequency but filtered below the Nyquist frequency will artificially increase computational time, with no additional gain through the high sample rate.

Furthermore, the length of the template has not only a significant influence on the correlation output (Figure 2.2), but also plays a role in the computing costs. Restricting the correlation to only include the seismic phases which exhibit the most energy and coherency across a network, helps to reduce computational time and maximize potential cross-correlation coefficients. As primary (P) and secondary (S) waves are typically the most energetic phases with clear onsets that separate them well from the noise, the template matching is performed using the S wave arrivals on horizontal sensors and the P wave on vertically orientated sensors. Thus, each template needs to have associated phase onsets. In the studies performed during this thesis, I picked all template phase onsets manually. First, this ensured that all templates have precise phase picks, which is important for use in location/relocation techniques and later correlation analysis, as well as to have consistent manually picks by not distributing the task among several operators, which would result in potentially differently picked waveforms within the same dataset. Thus, I restrict the template waveform to only include data surrounding the respective phase onset. The length of the time window can vary depending on the targeted magnitude range, but typical values range from 0.5 – 1.5 s before to 3 – 5 s after the phase onset. Following these pre-processing steps, which are performed on a local machine, template and continuous data are saved in binary files, where one template file corresponds to a single template at one component of the network, and one continuous file corresponds to 24 hours of recordings at one component.

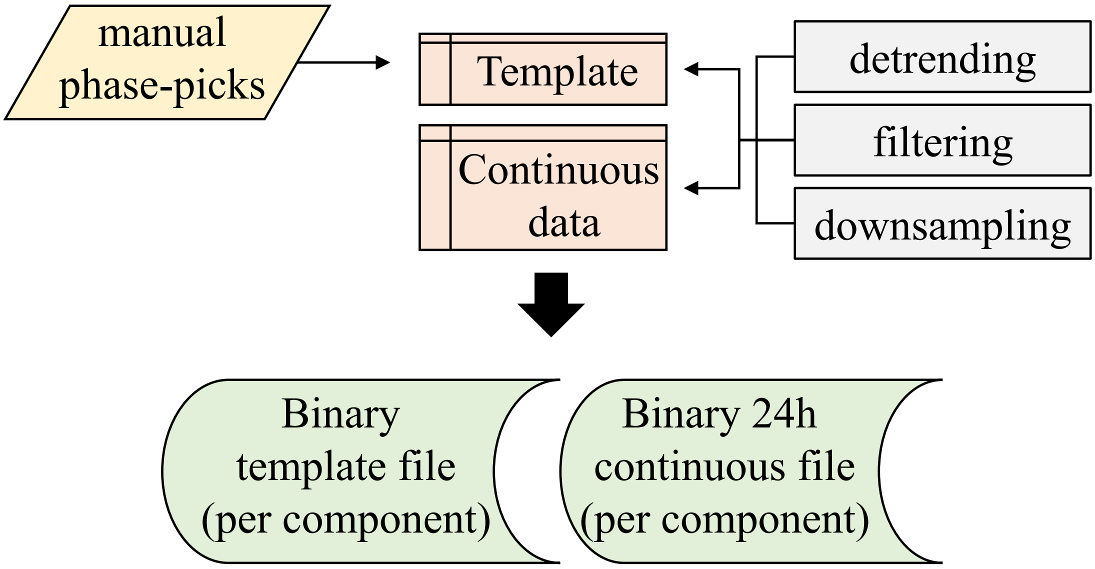

Figure 2.3: Schematic workflow of the pre-processing steps required before the template matching is performed.

The core of the practical implementation of the Matched Filter Algorithm is computing equation (2.8) for all time steps, all templates, and the entire analyzed data as quickly as possible. This base problem exhibits parallelism on multiple levels, which can significantly reduce computation time if properly exploited (Meng et al., 2012; Bentz et al., 2019). For each component at all stations, multiple templates need to be correlated with multiple days of continuous recordings. Performing all these independent calculations in parallel is an efficient way of reducing computational time and costs, when memory allocation and sharing are carefully considered.

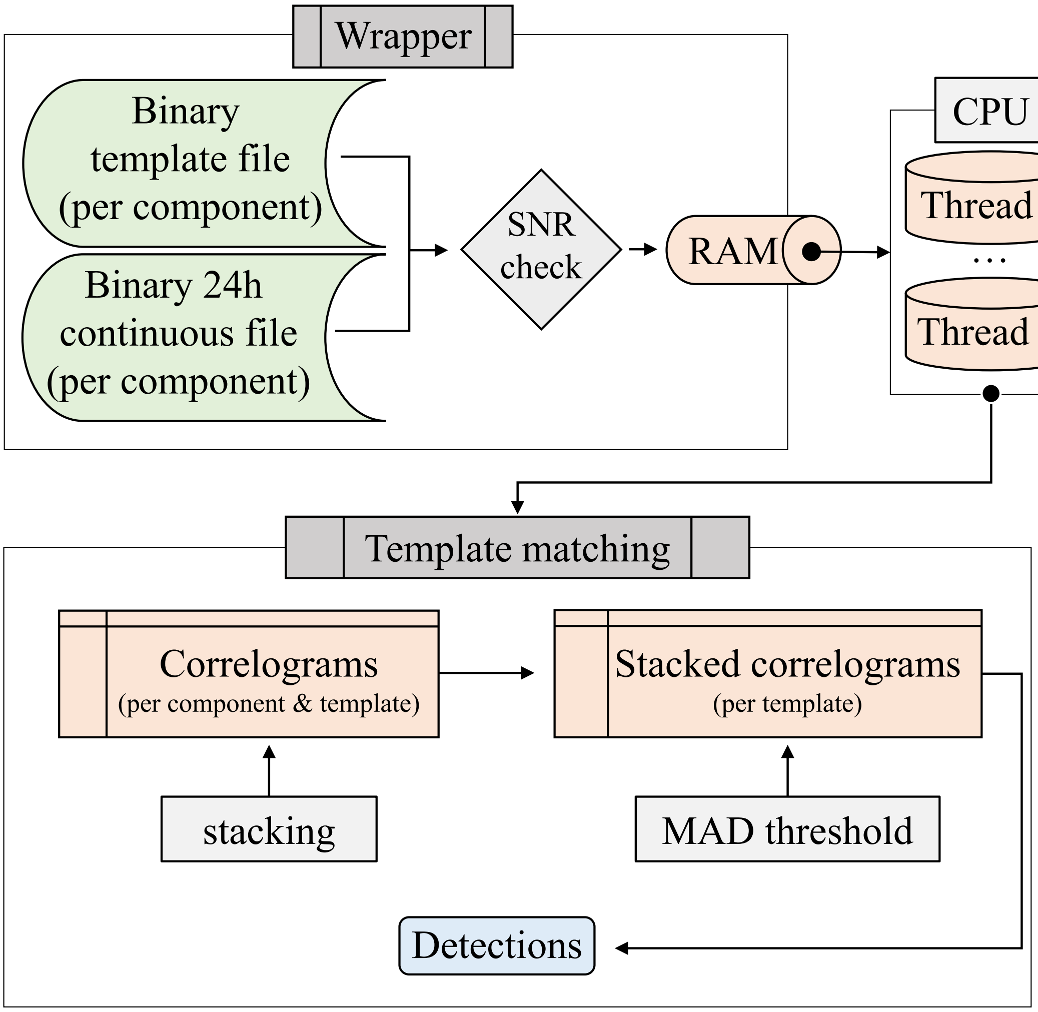

Figure 2.4: Schematic workflow of the computational implementation of template matching.

I implemented the MFA as follows (Figure 2.4):

1. A local wrapper program (bash/python/Matlab$^{®}$ scripts are available, depending on the platform running the computation) performs the data input operations, loading the continuous and template binary files into separate arrays in the memory. This step also includes reading in the phase onset information, which are contained in the binary header of the template file.

2. SNR test is performed on all template waveforms, disregarding any templates where the SNR between the data before and after the phase onset (P wave on vertical, S wave on horizontal components) does not exceed a threshold of at least SNR = 3. This ensures that no noisy signals are accidentally considered as templates, resulting in false detections.

3. The wrapper determines the optimum number of independent processes, and within each process determines the optimum number of threads, which is depending on the availability of CPU nodes and/or the local CPU architecture and system memory. Then, the computation jobs are dispatched to different CPU cores and nodes.

4. There are two options of numerically implementing and calculating equation (2.8). First let's consider the time-domain, where each thread computes the cross-correlation of one template with one day of continuous data. Further parallelizing by separating that computation into as many time steps as there are cases of overlapping samples between the two waveforms would in theory be possible, however, such a massively extended parallelizing would require the use of GPU computation over CPU. The implementation can be represented through pseudocode as:

```
int    n;    // Number of samples to be correlated
float  *temp;   // n-dimensional array containing the template
float   *cont;     // n-dimensional array containing the continuous data

float crosscorrelation_norm(n, temp, cont){
  int    i;
  double cor = .0;
  double acor1 = .0;
  double acor2 = .0;
  float  ccc;

  for(i = 0; i < n; i++){
    cor  += (temp[i] * cont[i]); // Correlation between template and
      continuous data
    acor1 += (temp[i] * temp[i]); // Autocorrelation of template
    acor2 += (cont[i] * cont[i]); // Autocorrelation of continuous
      data
  }
  ccc = cor/sqrt(acor1*acor2); // Normalized cross correlation
  return(ccc);
}
```

Here, equation (2.8) is directly put into a numerical function: the three correlation sums are calculated for one template containing n samples with n samples of the continuous data (Meng et al., 2012). This calculation needs to be repeated as often as there are samples in the continuous trace, by shifting the correlated part of the continuous data by one sample further until the end of the array. Therefore, this method is only viable for computation on a high-performance computing cluster, as the execution of all parallel jobs costs enormous CPU time.

For that, I also implemented the matched filter to be performed on regular systems by shifting the cross-correlation calculation into the frequency domain. That procedure could be represented by:

```
float  *temp;   // n1-dimensional array containing the template
float  *cont;   // n2-dimensional array containing the continuous data

float crosscorrelation_norm(temp, cont){
    float cor[sizeof(cont)];
    float acor1[sizeof(temp)];
    float acor2[sizeof(cont)];
    float ccc[sizeof(cont)];

    cor  = ifft(fft(temp) * fft(flipud(cont))); // Correlation between
       template and continuous data
    acor1 = ifft(fft(temp) * fft(flipud(temp))); // Autocorrelation of
       template
    acor2 = ifft(fft(cont) * fft(flipud(cont))); // Autocorrelation of
       continuous data
    for(i = 0; i < sizeof(cont); i++){
        ccc[i] = cor[i]/sqrt(sum(acor1[max]) * sum(acor2[max]))
    }
    ccc = cor/sqrt(acor1*acor2); // Normalized cross correlogram
    return(ccc);
}
```

Where the cross-correlation is now computed as the convolution, calculated by summation of the Fourier spectra, with the time reversed signal (equation 2.5). The advantage of this approach is, that the entire correlogram is calculated in a single step and afterwards normalized in one loop, which accounts for most of the required CPU time in this technique.

5. For each template, the correlograms of all traces at one day are stacked. As only the relevant surrounding part of the phase onset of the template was correlated, during stacking each correlogram has to be shifted back according to the move out of the station relative to the origin time (i.e. the travel time).
6. The stacked correlogram is scanned through, whenever the MAD threshold is exceeded a detection is declared and the onset time is saved. Note that, due to the

shift back to origin time from the individual correlograms, the detection times correspond to the origin times, not the times of highest correlation of the detections. In case of multiple detections in a 2 − 5 s time window (the value depends on the data), only the detection with the highest cross-correlation coefficient is kept and the remaining are rejected.

7. An initial detection catalog is saved with origin times, MAD, and correlation coefficients. Detections with the lowest MAD and cross-correlation coefficients are manually plotted and examined, to eliminate the possibility of any false detections (Figure 2.5).

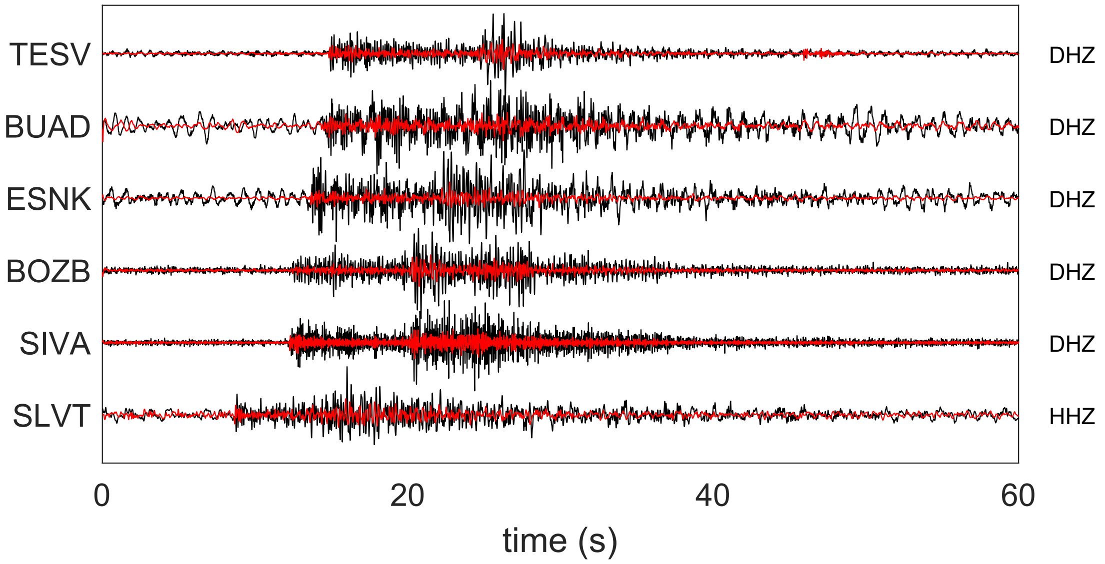

Figure 2.5: Example waveforms at six different channels for the matched filter algorithm detection in the western Sea of Marmara (Detection 20190922153045 as black line) using the template 20190922120709 (red line). Each trace is normalized and the template is scaled to one third of the detection.

2.3 Magnitude Estimation of Detections

The magnitude for each new detected event is calculated by comparing the amplitudes between the detection waveforms and the templates used to find it, assuming that a tenfold increase in amplitude corresponds to a 1-unit increase in magnitude:

$$M_{\text{det}} = M_{\text{temp}} + \log_{10}\left(\frac{A_{\text{det}}}{A_{\text{temp}}}\right) c, \tag{2.11}$$

where M_{temp} is the template magnitude, c is a constant, and A_{det} and A_{temp} are the peak amplitudes of detection and template, respectively (Peng & Zhao, 2009; Schaff & Richards, 2014; Vuan et al., 2018). In most cases c equals 1, assuming that the template and detection are located at the same spot and have had the same wave paths to the station where the amplitude ratio is calculated. Nevertheless, small differences in locations may already result in significant scatter in this relation (Figure 2.6). Therefore, I perform a site-specific linear regression to obtain detection magnitudes: first performing a linear fit between template magnitude and the logarithm of the maximum amplitude for each component of each station. Then I test the validity of this relation by recalculating the template magnitudes based on the linear fit and compare it with the initial catalog magnitudes. This step helps to identify outlier templates, and remove them from the linear fitting (see Figure 2.6, there are distinct events not fitting to equation 2.11). Afterwards, the linear fit is recalculated and relative magnitudes are calculated for each detection. The standard deviation of the fit across all components is taken as a measure of error for the magnitude determination (Bentz et al., 2019).

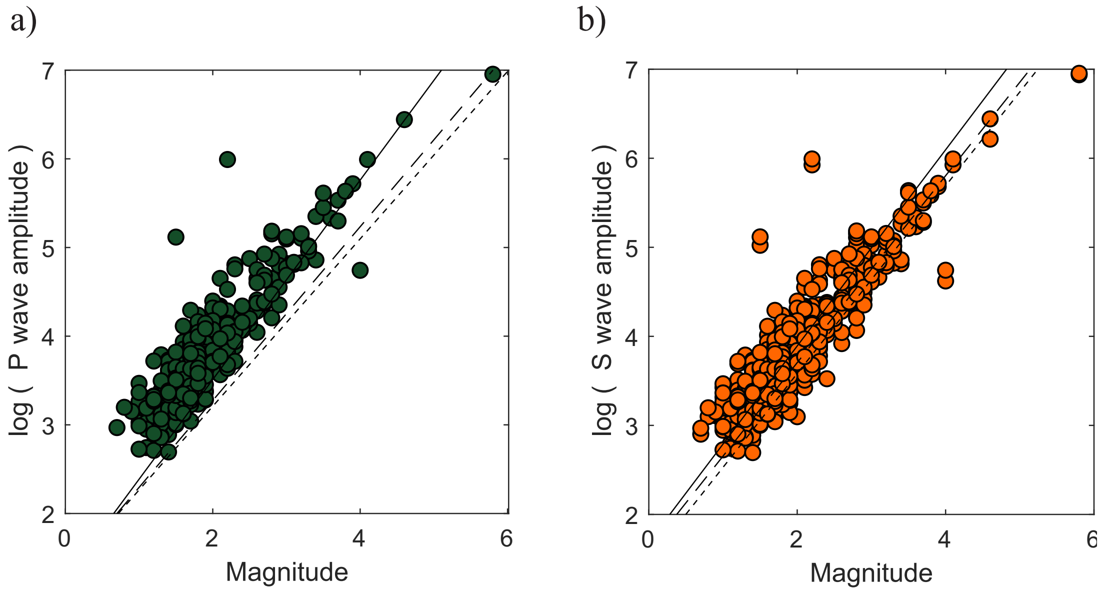

Figure 2.6: Relation between (a) P wave amplitude, (b) S wave amplitude and the event magnitude for an earthquake sequence in the central Sea of Marmara, Turkey. The lines represent different linear fits: Solid line shows a linear fit to all events, dashed only to $M > 2.5$, dotted only to $M > 3$.

The quality of a seismic catalog is typically judged by the magnitude of completeness M_C, that is the lowest magnitude until the catalog fulfills a Gutenberg-Richter power law distribution:

$$\log_{10}(N) = a - bM, \tag{2.12}$$

where N is the cumulative number of earthquakes having magnitudes larger than M and a and b are constants. M_C is calculated following the goodness-of-fit test by Wiemer and Wyss [2000]:

$$R(a, b, M_i) = 100 - \left(\frac{\sum_{M_i}^{M_\mathrm{max}} |B_i - S_i|}{\sum_i B_i} \times 100 \right), \tag{2.13}$$

computing a synthetic distribution of magnitudes using the same b, a, and M_i values and calculating the absolute difference, R, of the number of events in each magnitude bin between the observed, B_i, and synthetic, S_i, distributions. M_C is defined as the point at which a power law can fit 90 % of the seismicity or more, that is, at $R = 90\,\%$. The obtained b values can be further analyzed, depending on the frame of the study. However, it is important to assess the stability of the b value, should it be further interpreted (Raub et al., 2017). In that case, a Monte Carlo experiment is performed, where the magnitudes are assigned an error drawn from a normal distribution with a maximum value determined by the variations in the relative magnitude calculations (Bentz et al., 2019).

2.4 Phase Associations and Absolute Localization

Assuming that new earthquakes are detected at a sufficient number of stations and components, an absolute localization through travel time inversion is possible. The exact arrival times for each new detection are based on cross-correlation lag times (Figure 2.7). First, the detection waveform is extracted from the continuous recordings. Then, the detection is cross-correlated again with all templates at all stations. This secondary cross-correlation is performed in an even narrower time window around the phase onsets, that is, $< 0.5\,\mathrm{s}$ before and $< 1\,\mathrm{s}$ after the manual phase pick on the template. Vertical components are used to determine P wave arrival times, and horizontal components for the S wave. Afterwards, for the template showing the highest correlation to the detection, the phase picks are computed as the time lag between the cross-correlation maximum and the template pick. This procedure was extensively tested whenever applied, by automatically repicking all template phase onsets using the correlation and lag times with all remaining templates. These new automatic onsets can be used to locate the templates again and compare this location with the previous locations from the manual reference picks. If the locations deviate from each other for more than the uncertainty of the initial location, the corresponding template should be removed from application to the new detections.

Absolute hypocenter locations are resolved by use of the maximum intersection method, where the earthquake hypocenter is determined independently of the origin time through the use of the equal differential time (EDT) surface (Font et al., 2004). For any pair of arrivals, such as the same phase on two different stations or two different phases on the same station, an EDT surface is defined as the collection of all spatial points in the subsurface that satisfy the time difference between the observed arrivals for a given velocity model (Zhou, 1994). The intersection of all possible EDT surfaces is equal to the global minimum of the cost function used:

$$\mathrm{CF} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\delta T_{\mathrm{obs},i} - \delta T_{\mathrm{th},i})^2}, \tag{2.14}$$

where $\delta T_{\mathrm{obs},i}$ is the travel time difference between the observed i-th arrival pair, $\delta T_{\mathrm{th},i}$ the corresponding theoretical difference, and N the total number of arrival pairs. The concept of this method can be demonstrated very easily (Figure 2.8). The addition of more EDT surfaces (Figure 2.8a), which is equal to introducing more differences in equation 2.14, provides a more and more detailed picture of the global minimum of the cost function, and therefore narrows down the location of the hypocenter (Figure 2.8b).

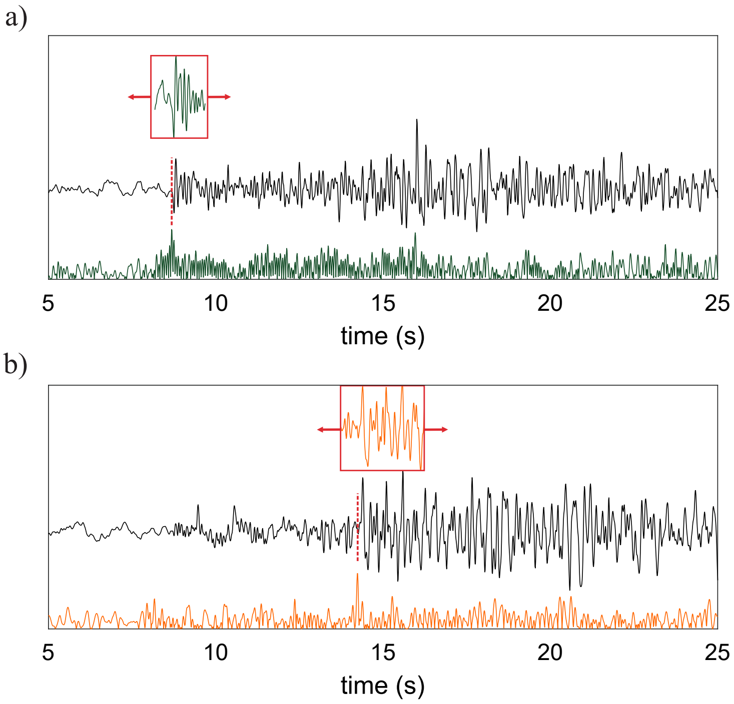

Figure 2.7: Example of the determination of phase onsets for a matched filter detection. (a) The P wave onset at the vertical component (black line) is determined by cross correlating a small time window around the P wave arrival of the template (red box) with the detection. The onset is picked at the peak of correlation function (green curve). (b) The S wave onset at one horizontal component (black line) is determined by cross correlating a small time window around the S wave arrival of the template (red box) with the detection. The onset is picked at the peak of correlation function (orange curve).

The global minimum is solved for by using the hypocenter of the highest correlated template as an initial location in a Metropolis-Hastings random walk to sample the three-dimensional space of hypocenter locations (Hastings, 1970). For a fixed number of iterations, the error function is evaluated at randomized values based on the current best fit solution, as follows: assuming that XYZ_{old} is the current event location, then the new trial location XYZ_{new} will be equal to XYZ_{old} plus a random error drawn from a normalized distribution. The new location will be accepted if it fulfills a Boltzmann-type probability distribution:

$$\exp\left(-\frac{\mathrm{CF}(XYZ_{\text{new}})-\mathrm{CF}(XYZ_{\text{old}})}{\sigma}\right) > rand, \tag{2.15}$$

where σ is the width of the distribution, set to 10^{-3} here, and $rand$ is a random number drawn from a uniform distribution between -1 and +1. If the probability is not fulfilled the location is rejected and a new step is performed based on the old values. An initial maximum step of 100 m per iteration is allowed in a burn-in period of up to 10^5 samples, where the step range is increased or decreased to ensure that the acceptance ratio is equal to 50 ± 2%. Then, the main algorithm is run for an additional 10^6 steps. This approach allows to not only find the absolute hypocenter location, but also to determine its probability density function by examining the misfit for each iteration independent from whether the step was accepted or rejected. Then the average location uncertainty (classical error ellipsoid) can be estimated by a principal component analysis of all sampled locations.

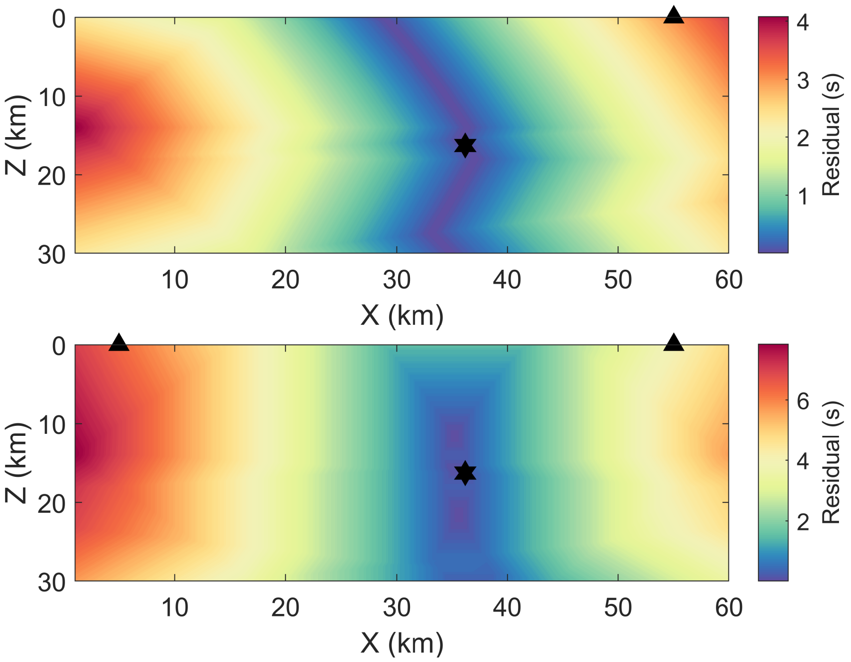

Figure 2.8: Synthetic EDT surfaces calculated from full wavefield modelling using a 1D velocity model of the eastern Sea of Marmara, Turkey (Karabulut et al., 2011). Triangles indicate seismic stations, black star the earthquake hypocenter. (a) The cost function of one single station plotted using both *P* and *S* wave. (b) The combined cost function of two stations using *P* and *S* waves at each station.

2.5 Relative Localization

Depending on the seismic network coverage, aperture, and instrumentation, it might happen that for certain earthquakes detected by the matched filter no sufficient amount of stations will detect the earthquake with enough energy for phase onsets to be determined. Without a good and large set of phase picks, a classical travel time inversion as described earlier is not possible, or would result in locations with substantial uncertainties. However, as typically the templates used in the matched filter have well constrained locations, the cross-correlation coefficients between template and detection can help to locate new earthquakes relative to these templates.

Relative localization is by now a standard post-processing step in creating detailed earthquake catalogs, and is most commonly based on the double-difference technique introduced by Waldhauser and Ellsworth (2000). Where the differential travel times between *P* wave and *S* wave arrivals for closely located earthquakes are used to relocate them relative to each other. This way, poorly known path effects between source and receiver are eliminated, improving hypocenter accuracy by a factor of up to 10. Further improvement can be achieved by including the cross-correlation values between event pairs on top of the double differences. However, this technique still requires a sufficiently large set of differential arrival times and cross-correlation pairs at multiple stations.

Considering low SNR events, detected at a handful of stations, or even only a single station, a relative location using template cross-correlation is still possible in a weighted location scheme:

$$XYZ_{\text{det}} = \frac{\sum_{i=1}^{N}(\tanh^{-1}(ccc_i)XYZ_i)}{(\tanh^{-1}(ccc_i))} \tag{2.16}$$

where XYZ are the three-dimensional coordinates, ccc_i the mean correlation coefficient between the detection and the i-th template, and N the total number of templates which exceeded the MAD limit for the located detection at one station. As the same detection is triggered N times by N different templates, this approach allows to locate new detections based on all templates that are more similar to the detection than to uncorrelated noise. Equation 2.16 can be used on a single station or in combination at multiple stations. The weight of each template is determined by the Fisher transform of the cross-correlation coefficient (Fisher, 1921), to ensure that high correlation values close to one have much larger weights compared to low or intermediate ones (Figure 2.9).

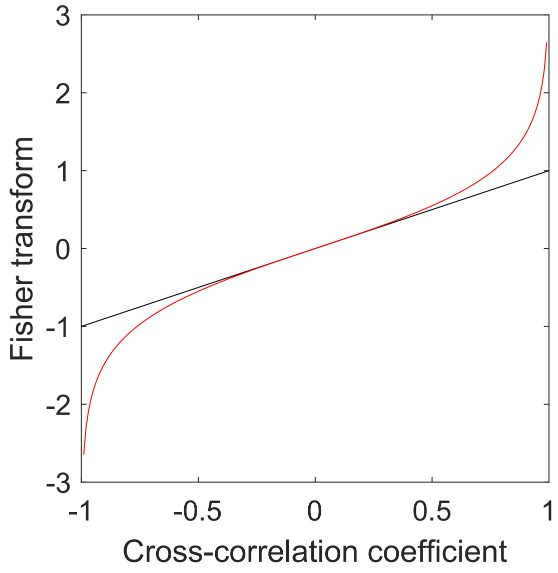

Figure 2.9: Comparison between weights of the cross-correlation coefficient (black line) and its Fisher transform (red line).

2.6 Additional Processing Based on the Obtained Cross-Correlations

Following all steps described in the previous sections, the final result of the matched filter approach is a classical earthquake catalog, containing locations, origin times, and magnitudes. This new catalog can substantially improve earthquake detection capabilities, by lowering the magnitude detection threshold and thus the magnitude of completeness compared to public network catalogs. In addition, many earthquakes of similar magnitude than already present in the existing catalogs may be detected, providing a more detailed picture of underlying processes (Figure 2.10). We often see that seismic catalogs lack events with magnitudes above their magnitude of completeness when compared to matched filter catalogs (e.g. Figure 2.10a). This can be caused by fitting a power law to the frequency-magnitude-distribution of the catalog, but falsely assuming wrong constants, that is the b and a value of the Gutenberg-Richter law. However, these constants are important input parameters in a series of seismic hazard assessments and scenarios. Therefore, providing better assessments of the true magnitude distribution using template matching detected events is an important step in improving our understanding of earthquake processes and hazard assessments.

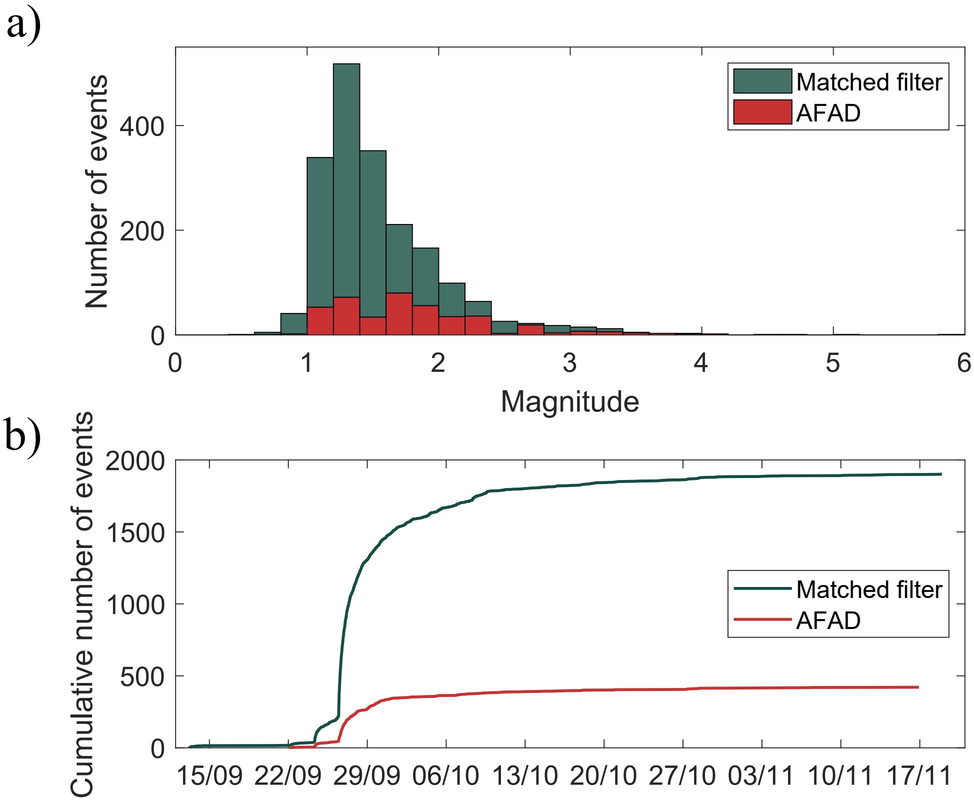

Figure 2.10: Comparison between the national seismic catalog provided by the Turkish Disaster and Emergency Management Presidency for Emergency AFAD and the results obtained from template matching in Chapter 5. (a) Number of events against magnitude. (b) Cumulative number of events against time.

The cross-correlation analysis performed during the template matching can be further processed, directly addressing seismological and geomechanical questions. One such application is the search and identification of similar events (*earthquake repeaters*). These are events that exhibit extremely similar waveforms, typically having much larger cross-correlation coefficients than needed for template matching, i.e. cross-correlation coefficients larger than 0.9. These earthquakes are expected to originate from the same spot and rupturing with the same mechanism (Rice, 1993). Generally, two types of earthquake repeaters are observed, similar earthquakes that occur in rapid succession of each other, either in the frame of earthquake swarms or fore-/main-/aftershock sequences (called *transient repeaters*), and so called *long lasting repeater earthquakes* with periodicity of several months or even years. Both types are attributed to the repeating and persistent reloading and rerupturing of small asperities, but the mechanisms causing the potential strain energy surrounding the asperities are very different.

Long-lasting repeaters can be attributed to the tectonic background loading assuming it is sufficiently large (e.g., Lengliné & Marsan, 2009), or more generally to slow and aseismic slip of the fault surrounding the asperity (e.g., Igarashi et al., 2003; Bourouis & Bernard, 2007). Similar events with much shorter periodicity are more likely related to an accelerating nucleation process (e.g., Bouchon et al., 2011; Kato et al., 2012) or the reloading of asperities caused by afterslip of larger earthquakes (e.g., Peng & Zhao, 2009). The observation of long-lasting repeaters can first help to identify creeping sections of a fault and secondly, their reoccurrence times can provide valuable inside into the slip rate of large faults (Chen et al., 2007). Furthermore, estimating the amount of creep based on the coseismic displacement of repeater pairs allows a first order assessment of the accumulation of stress for particular fault segments, which has substantial implications for seismic hazard assessments (e.g., Bohnhoff et al., 2017).

One example of such applications of cross correlation for further interpretation is the earthquake sequence presented in chapter 5 of this thesis. During a moderate sized earthquake sequence in September 2019 in the Central Basin of the Sea of Marmara (Turkey), the importance and presence of slow slip in the fore-/main-/aftershock earthquake interaction was not obvious and presented one key challenge in the analysis. Schmittbuhl et al. (2016a) had previously investigated several clusters of long-term repeaters west of the Central Basin. To identify if large aseismic slip on nearby fault segment was playing a major role in the earthquake triggering, cross-correlation was used to investigate a potential acceleration in reoccurrence time of these long-term repeaters prior to the September 2019 sequence (Figure 2.11). Extending the repeater catalog of Schmittbuhl et al. (2016a) to the last 11 years, no substantial change in periodicity (Figure 2.11a) or accumulated moment in the repeater clusters (Figure 2.11b) could be observed. This may indicate that any slow slip or aseismic component in the earthquake triggering process was rather a short-term phenomenon, or that the detection capabilities of the seismic network were not adequate to capture all pattern changes, or that slow slip did not play a significant role in triggering this earthquake sequence at all (see also chapter 5 for further in-depth discussion on this topic).

The occurrence of similar events with short recurrence times, so called transient repeaters, was also investigated for the September 2019 Central Basin sequence, using the entire template matching catalog. A density-based spatial clustering algorithm was used to identify groups of events that exhibit cross-correlation coefficients larger than 0.75 to each other (Ester et al., 1996). The threshold was purposely set lower than typical repeater studies (> 0.9), as the goal was to identify similar events indicating rupture of the same broader area, rather than find events located at exactly the same spot. A total of 6 transient repeater clusters before the $M_W5.8$ earthquake were found (Figure 2.12), which would have not been detected through the regular regional seismic catalog. The occurrence and the locations these events highlights the possible influence of afterslip of the initial $M_W4.7$ earthquake in triggering the $M_W5.8$ earthquake, or they could also indicate a cascading triggering mechanism (Ellsworth & Bulut, 2018) (see chapter 5 for further in-depth discussion).

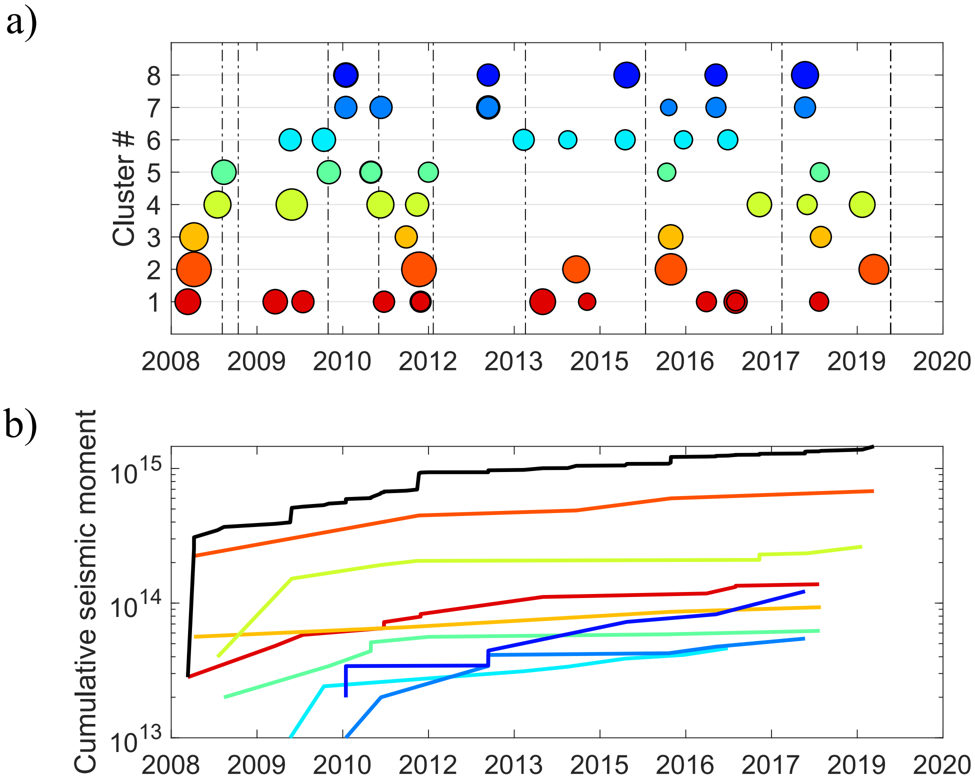

Figure 2.11: (a) Occurrence time of long-term repeater earthquakes in the western Sea of Marmara first investigated by Schmittbuhl et al. (2016a), color coded by the different clusters, and size encoded with magnitudes. The vertical dashed lines indicate $M_W > 4$ earthquakes on the surrounding fault segments. (b) The cumulative seismic moment evolution per repeater cluster (color coded by cluster) and the total cumulative seismic moment of all clusters (black line).

Furthermore, cross-correlation can also be applied continuously across detected events to see if a series of subsequent events shows a trend of localization, which should be apparent in the average cross-correlation. Previous studies have found increasing average correlations prior to the initiation of a mainshock, making it a potential precursory signal before large earthquakes (e.g., Malin et al., 2018). This would be in agreement with mechanical laboratory studies, who have found an increased spatial correlation before main failure on rock samples (e.g., Hirata et al., 1987; Zang et al., 1998; Goebel et al., 2012). In this respect, the cross-correlation coefficients already calculated during template matching can help to identify localization trends in seismicity. For the September 2019 central Sea of Marmara earthquake sequence, a distinct increase in average correlation coefficients was observed right before the $M_W5.8$ earthquake. This increase was accompanied by a migration of earthquakes towards the future rupture spot. Cross-correlation can therefore provide an additional observation of the detailed failure process related to large and moderate sized earthquakes.

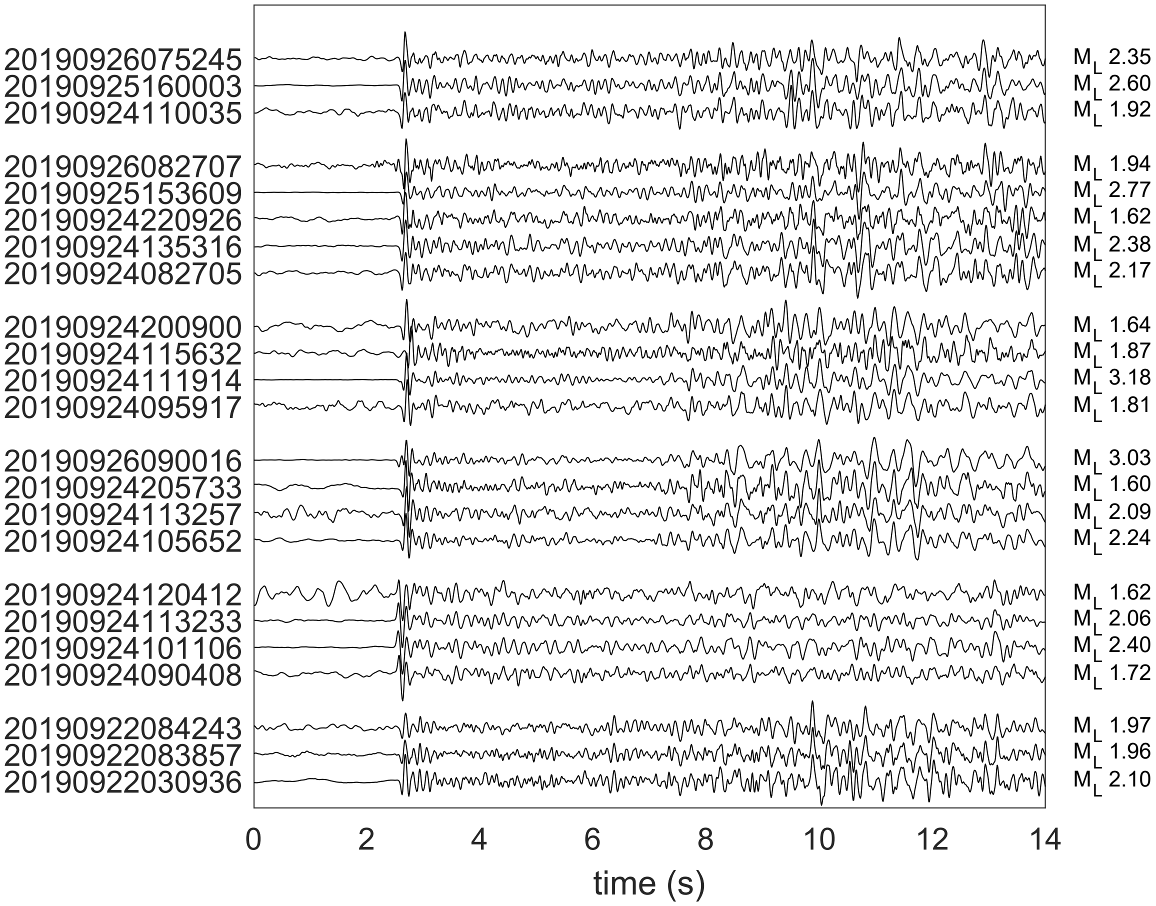

Figure 2.12: Vertical component recordings of the *SLVT* surface broadband station in the central Sea of Marmara for all similar earthquake clusters before the September 2019 $M_\mathrm{W}5.8$ mainshock. Event IDs are indicated on the left, and the respective magnitude on the right.

2.7 Drawbacks and Alternatives to Template Matching

Without a doubt template matching is one of the most powerful techniques to find earthquakes in continuous data recordings (e.g., Z. E. Ross et al., 2019). It's widely used to overcome the limitations of energy based detectors, such as STA/LTA, for low SNR conditions. Given long standing deployments of most seismic networks, utilizing the entire waveform instead of only impulsive onsets, can help to identify signals from seismic sources that repeat in time (Geller & Mueller, 1980), assuming that path effects are essentially constant over seismological times scales (Poupinet et al., 1984). Waveform cross-correlation scores high on the most important metrics of seismic detectors, that is, in detection sensitivity and computational efficiency (Yoon et al., 2015), both when applied to sparse seismic networks (e.g., Shearer, 1994) and dense local networks (e.g., Meng & Peng, 2014; Wang et al., 2015). However, the obvious drawback of applying known waveforms to find similar earthquakes in noisy data is that an a priori waveform template needs to be known, thus severely limiting its general applicability (Yoon et al., 2015).

Subspace detectors have been proposed as one alternative to matched filter, because they generalize template matching to similar, nonrepeating sources (e.g., Maceira et al., 2010). Instead of comparing raw waveforms to each other, subspace detection is based on the singular value decomposition of a set of waveforms to identify an orthonormal representation that captures the most important similarities. These representations can be used in a similar manner than template matching, to find events which are similar above a certain threshold. This approach can especially help to identify similar events that exhibit slight offsets in earthquake locations, or find overlapping events (Barrett & Beroza, 2014). Nevertheless, subspace detectors also require an a priori knowledge about the searched for signals, resulting in the same limitations as the template matching but at higher computing costs (Yoon et al., 2015).

Cross-correlation can be generalized to detect unknown signals, through the autocorrelation approach. Assuming that the target signal is only of short duration, the continuous signal can be split up into N overlapping short windows. Then the cross-correlation of all possible window pairs is calculated, and any pairs exceeding a given detection threshold are labeled as possible events. This technique has proven valuable to find low-frequency earthquakes within tectonic tremors (Brown et al., 2008). Unlike template matching, no a priori knowledge of the waveform is required, thus making it in theory possible to apply it blindly on large sets of continuous data. In reality, however, autocorrelation has several disadvantages. First, cross-correlation pairs marked as possible events need to be further processed and analyzed to identify them as true earthquake detections and not as correlated noise, potentially requiring additional cross-correlation or stacking. Furthermore, autocorrelation is one of the computationally most expensive detection techniques, making it impractical for large quantities of data. For N set of windows, $N(N-1)/2$ cross-correlations need to be calculated, accounting for all possible event pairs. Thus, computing time scales with N^2, where most of it is wasted performing redundant operations on cross-correlating uncorrelated noise (Yoon et al., 2015).

Recently, Yoon et al. (2015) introduced a new earthquake detection technique called Fingerprint and Similarity Thresholding (FAST) utilizing advanced data mining algorithms originally designed to identify similar audio clips within large databases. This new approach combines the computational efficiency of traditional template matching with the general applicability of autocorrelation, as no a priori information about the waveforms have to be provided. FAST extracts the most important features of the signals through the wavelet transform, and then utilizes the sign of the standardized Haar wavelet coefficient of the spectral image to associate a binary fingerprint to any given window in the data. A locality-sensitive hashing algorithm is used to identify probable pairs of similar fingerprints, without actually comparing dissimilar pairs. FAST is a scalable approach, capable of substantially lowering the detection threshold of seismic networks and identify microseismic patterns (Yoon et al., 2019). Nevertheless, FAST suffers from a larger false detection rate than those of classical template matching or autocorrelation techniques. In addition, as the algorithm requires pairs of similar fingerprints, any waveform that is not repeated or similar to any other waveform in a given dataset will be completely missed.

For the most part, current algorithms in seismology can't keep up with the increasing density of seismic networks and the associated increase in large continuous datasets. The shear amount of data stored by many institutions takes many of the described methods to their limits. Cross-correlation and subspace detectors remain one of the strongest contenders for standard application on massive datasets, through their computational efficiency and extremely low false detection rate. Nevertheless, it is important that our capabilities to process large quantities of data grows alongside our observational capabilities. Interaction between seismology and computer science utilizing state of the art data mining techniques will become most crucial in the future. Algorithms such as FAST are certainly a step in the right direction. Additionally the advances in ultrafast computers with the ever growing number of edited/labeled datasets will potentially make the application of machine learning one of the largest growing fields in seismology, and geosciences in general (Bergen et al., 2019). The strength of machine learning lies in the data-driven discovery, that is, the ability to extract new information from data, which aligns perfectly with seismology, a largely data-driven science. Neural networks and other classifiers are perfect candidates to improve our categorization process, i.e. to differentiate noise from signal (e.g., Paitz et al., 2018), or to detect phase onsets and locate earthquakes (e.g., Perol et al., 2018; Z. E. Ross et al., 2018). However, it should also be noted that many datasets represent complex physical systems with nonstationary components, potentially providing significant challenges to machine learning algorithms. Regardless of the approach, new data-driven methods need to be developed in order to not let the growth of data outrun our capabilities to analyze them.

3 Seismic Moment Evolution During Hydraulic Stimulations

Summary

Recent results from an EGS project in Finland suggest a possibly successful physics-based approach in controlling stimulation-induced seismicity in geothermal projects. We analyzed the temporal evolution of seismicity and the growth of maximum observed moment magnitudes for a range of past and present stimulation projects. Our results show that the majority of the stimulation campaigns investigated reveal a clear linear relation between injected fluid volume, hydraulic energy and cumulative seismic moments. For most projects studied, the observations are in good agreement with existing physical models that predict a relation between injected fluid volume and maximum seismic moment of induced events. This suggests that seismicity results from a stable, pressure-controlled rupture process at least for an extended injection period. Overall evolution of seismicity is independent of tectonic stress regime and is most likely governed by reservoir specific parameters, such as the preexisting structural inventory. In contrast, there are few stimulations that reveal unbound increase in seismic moment suggesting that for these cases evolution of seismicity is mainly controlled by stress field, the size of tectonic faults and fault connectivity. Transition between the two states may occur at any time during injection, or not at all. Monitoring and traffic light systems used during stimulations need to account for the possibility of unstable rupture propagation from the very beginning of injection by observing the entire seismicity evolution in near-real time and at high resolution for an immediate reaction in injection strategy.

Published as Bentz, S., Kwiatek, G., Martínez-Garzón, P., Bohnhoff, M., Dresen, G (2020). "Seismic moment evolution during hydraulic stimulations" in *Geophysical Research Letters*, 47(5), doi: 10.1029/2019GL086185.

3.1 Introduction

The effort to reduce production of greenhouse gases, especially in Europe, has fueled the search for clean and renewable energy sources. Geothermal energy has long been investigated as a potential complement and long-term replacement for traditional fossil fuels in electricity and heat production, with its significant baseload capacity. However, unlike in hydrothermal systems in regions of elevated heat flow, projects aiming to extract fluids hot enough to produce electricity need to reach large depths, typically $>$4,000 m in crystalline basement or sedimentary basins. In order to develop deep geothermal reservoirs in low-permeability rocks, the formation needs to be hydraulically stimulated. Creation of Enhanced Geothermal Systems (EGS) opens fluid flow paths by injection of large quantities of water, which is typically accompanied by seismicity (E. Majer et al., 2012; Ellsworth, 2013).

Large induced earthquakes have led to the termination or suspension of several EGS projects, such as the deep heat mining project Basel in Switzerland (Giardini, 2009) and St. Gallen (Diehl et al., 2017). The occurrence of a $M_{\mathrm{W}}5.5$ earthquake in 2017 near Pohang, South Korea, has been linked to a nearby located EGS project (Ellsworth et al., 2019). In densely populated urban areas substantial public concern about EGS projects now exists and is mainly related to stimulation-induced seismicity. In contrast, two large scale EGS projects in Australia have been operational for years and their remote location has prevented any acceptance issue arising from induced seismicity (Baisch et al., 2006; Albaric et al., 2014). However, most recently Kwiatek et al. (2019) reported on the successful and safe stimulation of the world's deepest EGS project so far, located near Helsinki. The authors attributed their success to the high-precision, almost real-time monitoring and analysis of seismic data feeding into a traffic light system and guiding the adaptive stimulation strategy (Ader et al., 2019).

Nucleation and propagation of earthquake ruptures related to fluid injection pose fundamental questions related to understanding and controlling injection-induced seismicity. Early models such as McGarr (1976) assumed that volume changes induce local changes in deviatoric stresses, which are then relaxed by induced earthquakes. Shapiro et al. (2010) defined a seismogenic index characterizing the level of seismic activity expected from injecting fluids into the rock formation, expanding on earlier work on the occurrence probability of events above a certain magnitude threshold (Shapiro et al., 2007). McGarr (2014) defined a critical pore pressure change caused by fluid injection inducing seismic events. The predicted cumulative seismic moment and maximum event magnitude are expected to increase with total fluid volume injected. Shapiro et al. (2011) postulated that the expected maximum magnitude may be related to the minimum axis of an ellipsoid covering the spatial distribution of earthquake hypocenters representing the fluid-stimulated volume. However, Norbeck and Horne (2018) recently analyzed numerically potential factors influencing rupture growth beyond the stimulated volume affected by pressure perturbations. Also, Galis et al. (2017) used a fracture mechanics model to analyze the relation between injected fluid volume and stable rupture propagation. The authors differentiate between stable self-arrested ruptures and unstable “runaway” ruptures, which propagate out of the stimulated volume on pre-stressed faults. Furthermore, Galis et al. (2017) found a fracture mechanics-based scaling relationship between maximum expected magnitude of self-arrested events and injected volume. These models predicting maximum magnitudes of induced events implicitly assume a stable rupture propagation process (Galis et al., 2017). In contrast, van der Elst et al. (2016) argued that induced earthquakes occurring along tectonic faults favorably oriented with respect to the tectonic stress field are unstable. The maximum expected magnitude is then only limited by regional tectonics and fault connectivity, which also holds for induced earthquakes.

Here, we compare the evolution of different parameters derived from seismic catalogs with injection parameters from several EGS projects and one scientific drilling project (Table 3.1). Unlike what was previously done, we treat each project as a dynamic experiment, rather than constraining the analysis to the final arrested state. The seismicity evolution with progressive fluid injection from the stimulations is compared to current models of injection-induced seismicity during the entire stimulation period. Our analysis provides insight about how seismic moment and related parameters of induced events evolve during injection and how these parameters are affected by tectonic faulting regime.

3.2 Data and Methods

We analyzed data collected from different studies investigating a total of eight stimulation projects, all located in crystalline basement rock (Table 3.1). These include the most prominent European EGS projects in Basel, Switzerland (BAS) (Häring et al., 2008) and Soultz-sous-Forêts (STZ), France. In Soultz, three different stimulations over the course of 10 years were performed in different wells and different depths. Therefore, we differentiate between the injections in 1993 (STZ93), 2000 (STZ00), and in 2003 (STZ03) (EMC., 2017, 2018a, 2018b). We also included the deepest EGS Project to date (St1), located in Helsinki, Finland (Kwiatek et al., 2019). Furthermore, we included the fluid-injection experiment from the German super deep scientific drilling hole (KTB) (Baisch et al., 2002), two Australian EGS projects, located at Paralana (Para) (Albaric et al., 2014) and the 2003 Cooper Basin (CBN) injection (Baisch et al., 2006), and the EGS project near Pohang, South Korea (Ellsworth et al., 2019). Finally, we also considered a single well injection period at the Berlín geothermal field (BGF), El Salvador, representing the only hydrothermal site considered here (Bommer et al., 2006; Kwiatek, Bulut, et al., 2014). The projects are located at sites dominated by strike-slip (e.g. KTB), normal faulting (e.g. Soultz) and transpressive (e.g. St1) tectonic stress regimes in a depth range between 3.5 and 9.0 km (Table 3.1).

Table 3.1: Summary of the analyzed injection projects.

Project	M_{max}	M_0	$\Delta V(\mathrm{m}^3)$	M_{C}	b value	Rock type	Depth (km)	Tectonic regime	Reference
Basel (BAS)	3.0	$3.9 \cdot 10^{13}$	$1.2 \cdot 10^{4}$	0.8	1.48	Crystalline	4.6	Strike-Slip	Häring et al. (2008), Kraft and Deichmann (2014)
Cooper Basin (CBN)	3.7	$4.5 \cdot 10^{14}$	$2.0 \cdot 10^{4}$	-0.9	0.77	Crystalline	4.3	Thrust	Baisch et al. (2006), Asanuma et al. (2005)
Soultz 1993 (STZ93)	0.3	$3.6 \cdot 10^{9}$	$4.5 \cdot 10^{4}$	-1.8	1.06	Crystalline	3.5	Normal	EMC. (2017)
Soultz 2000 (STZ00)	2.5	$7.1 \cdot 10^{12}$	$2.7 \cdot 10^{4}$	0.1	0.94	Crystalline	4.5	Strike-Slip	EMC. (2018a)
Soultz 2003 (STZ03)	2.9	$2.5 \cdot 10^{13}$	$3.4 \cdot 10^{4}$	0.0	1.24	Crystalline	4.5	Strike-Slip	EMC. (2018b)
Helsinki (St1)	1.9	$8.9 \cdot 10^{11}$	$1.9 \cdot 10^{4}$	-0.6	1.35	Crystalline	5.7	Strike-Slip / Thrust	Kwiatek et al. (2019)
Paralana (Para)	2.5	$7.6 \cdot 10^{12}$	$3.1 \cdot 10^{3}$	-0.3	0.94	Crystalline	3.6	Strike-Slip / Thrust	Albaric et al. (2014)
KTB	0.7	$1.4 \cdot 10^{10}$	$3.8 \cdot 10^{3}$	-0.8	1.0	Crystalline	5.4/9.0	Strike-Slip	Baisch et al. (2002)
Pohang	5.5	$2.2 \cdot 10^{17}$	$1.2 \cdot 10^{4}$	1.1	0.7	Crystalline	4.0	Strike-Slip	Ellsworth et al. (2019)
Berlín GF (BGF)	1.47	$2.2 \cdot 10^{11}$	$3.1 \cdot 10^{5}$	1.0	1.96	Crystalline	1.8/2.2	Strike-Slip	Kwiatek, Bulut, et al. (2014)

The main parameters analyzed in this study are the maximum observed moment magnitude during and following the stimulation $M_{\mathrm{max,obs}}$, and cumulative seismic moment $M_{0,\mathrm{cum}}$ with injected fluid volume and hydraulic energy E_{hyd}, i.e. the product of wellhead pressure and injected volume integrated over time. In order to compare the cumulative seismic moment released during each injection project, we have to take into account any possible bias introduced by different magnitudes of completeness (M_{C}) in the catalogs. Therefore, following Wiemer and Katsumata (1999) we calculate using the maximum curvature of the frequency-magnitude distributions (Table 3.1). Assuming a Gutenberg-Richter power law distribution of magnitudes (Gutenberg & Richter, 1956):

$$\log_{10} N = a - bM, \tag{3.1}$$

where N is the cumulative number of earthquakes having magnitudes larger than M, and a and b are constants. We calculate the b values using a maximum likelihood estimate (Bender, 1983). Then, with the maximum magnitude, b value, number and total moment of events above M_{C} we computed synthetic catalogs for each sequence extending down to the overall lowest calculated M_{C} (-1.8) using the upper bounded magnitude-frequency probability density function of Page (1968). From each synthetic catalog we calculated the ratio of moment missing from the real catalog between the magnitude of completeness and M_{W} -1.8, and added this to the cumulative seismic moment of the sequence. Missing events accounted for a maximum of 3 % of the total released moment at St1, for other sequences the ratio was typically below 1 %.

Furthermore we compare seismic injection efficiency I_{E}, i.e. the ratio of cumulative radiated energy and hydraulic energy for the different stimulation projects (Maxwell et al., 2008). The radiated seismic energy E_0 for each event is estimated from the seismic moment M_0 and stress drop $\Delta\sigma$ following Hanks and Kanamori (1979) as:

$$E_0 = \Delta\sigma \frac{M_0}{2G} \eta_{\mathrm{R}}, \tag{3.2}$$

where G is the shear modulus calculated as $G = \rho V_{\mathrm{S}}^2$, with density ρ and shear wave velocity V_{S}, and η_{R} is the radiation efficiency. When available we used already calculated catalogs of radiated energy (Kwiatek, Bulut, et al., 2014; Kwiatek et al., 2019) or computed radiated energy using average stress drops available (Jost et al., 1998; Charléty et al., 2007; Baisch et al., 2009; Goertz-Allmann et al., 2011), and assumed a median radiation efficiency of 0.46 (McGarr, 1999). The KTB data were corrected for the finite bandwidth applied in previous studies (Ide & Beroza, 2001). For the Pohang and Paralana stimulations no stress drop or radiated energy calculations were available and so we assumed an average stress drop of 3 MPa (Cocco et al., 2016).

3.3 Results

3.3.1 Maximum Magnitude With Injected Volume

For most of the respective stimulation period, all analyzed datasets show an increase of maximum observed moment magnitude $M_{\mathrm{max,obs}}$ with injected volume that roughly fits to a linear trend in a double logarithmic plot (Figure 3.1a). However, the slopes of the respective evolutions differ significantly between sites. The increase of $M_{\mathrm{max,obs}}$ with injected fluid volume follows a slope between 1 and 1.5 as predicted by McGarr (2014) and Galis et al. (2017), respectively, for most of the projects analyzed here. The trends observed for KTB, Paralana and St1 projects are best predicted by the model of Galis et al. (2017). In contrast, Soultz93 and Soultz00 stimulations show a slope which is better modeled by McGarr (2014) with an even smaller slope for BGF. In striking contrast, stimulations performed at the Cooper Basin and Pohang sites show a much steeper increase of $M_{\mathrm{max,obs}}$ with injected fluid volume, which for Pohang finally led to the occurrence of the largest known event induced by stimulation of a geothermal reservoir ($M_{\mathrm{W}}5.5$). We note that these two steep trends are not predicted by any of the existing models.

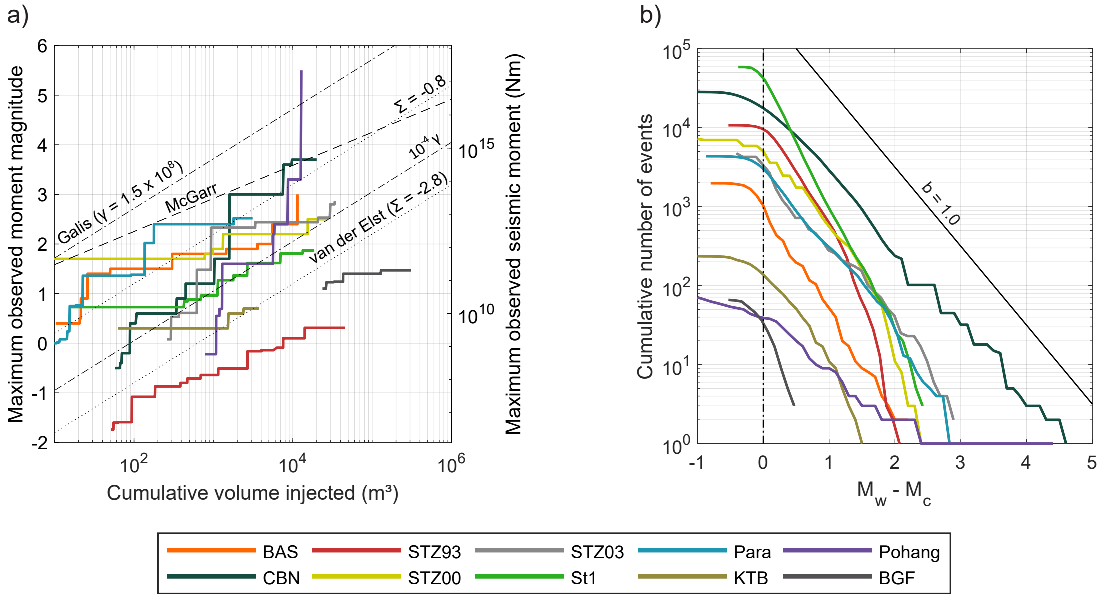

Figure 3.1: (a) Temporal evolution of maximum observed moment magnitude $M_{\mathrm{max,obs}}$ with cumulative volume injected. Dashed and dotted lines indicate maximum predicted magnitudes using different models with varying parameters (McGarr, 2014; van der Elst et al., 2016; Galis et al., 2017). (b) Magnitude-frequency distribution for each project, aligned at their magnitude of completeness.

Basel, Soultz03, and Paralana initially show a relatively steep increase of $M_{\mathrm{max,obs}}$ before settling to a more modest trend fitting reasonably well McGarr (2014) and Galis et al. (2017) models (Figure 3.1a). For Basel the occurrence of a $M_{\mathrm{W}}3.0$ just after terminating the injection (shut-in) is a large deviation from the previous linear trend. Soultz03 displays a steep increase of $M_{\mathrm{max,obs}}$ with injected volume at the beginning of the stimulation, which then decreases onto a slope close to unity after 1,000 m^3 of injected fluid.

Interestingly, even injection at the same site does not yield the same evolution of $M_{\text{max,obs}}$, which might be relevant for predicting $M_{\text{max,obs}}$ for subsequent injections at the same site. The reservoir in Soultz-sous-Forêts was stimulated during three different periods, in 1993 at 3.5 km depth and in 2000 and 2003 at 4.5 - 5 km depth. In 1993, seismic activity was limited to very small magnitudes, with $M_{\text{max,obs}}$ not exceeding 0.3. Seismicity evolution was mostly linear and is well fitted by the model of McGarr (2014). This stands in contrast to the later injections in 2000 and 2003, resulting in seismic events of much larger magnitudes. In 2000, within a few hours after stimulation started a $M_{\text{W}}1.8$ event occurred, which was followed by a linear increase in magnitudes to $M_{\text{W}}2.5$ at the end of injection. Likewise, the stimulation in 2003 is characterized by first a steep increase of maximum magnitude at the onset of injection followed by a more modest increase for most of the remaining injection period.

3.3.2 Seismic Moment Evolution

Comparing the cumulative seismic moment $M_{0,\text{cum}}$ to the cumulative volume of fluid injected we find a clear linear relationship (Figure 3.2a). Most of the seismic sequences display a stable long-term increase of $M_{0,\text{cum}}$ with a slope of about one, consistent with the model of McGarr (1976, 2014). This suggests that for the dominant part of the injection period there exists a linear response between seismic energy and injected fluid volume. Many sequences are characterized by an initial phase of rapid increase in released seismic moment followed by an almost linear slope. The steep initial increase is especially pronounced in the Soultz03 injection.

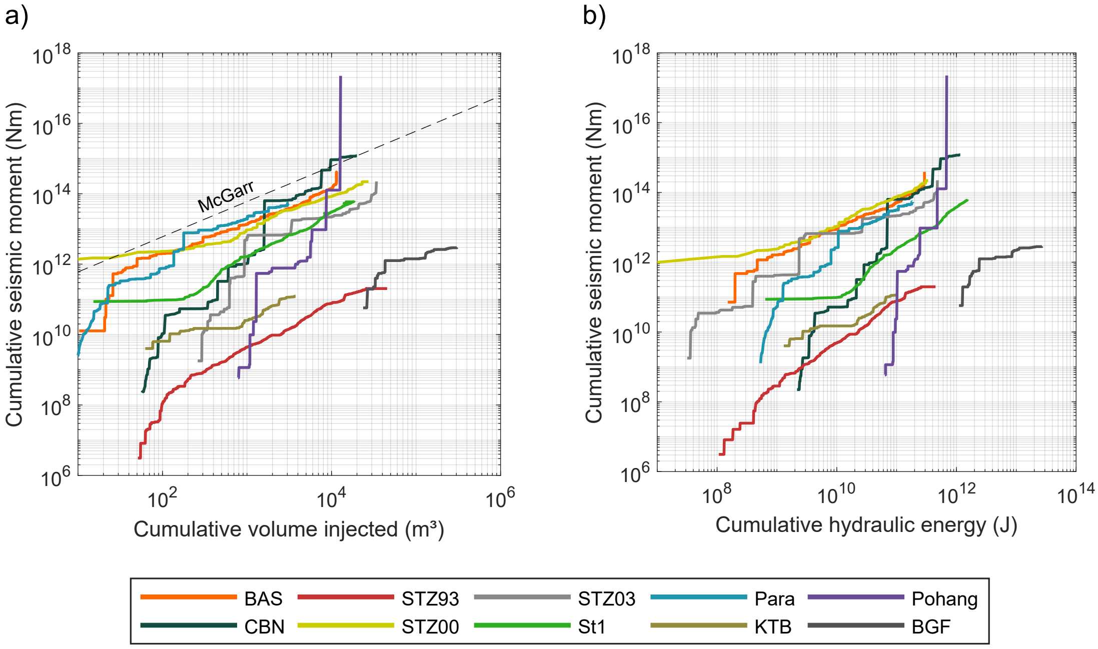

Figure 3.2: Temporal evolution of cumulative seismic moment $M_{0,\text{cum}}$ with (a) cumulative fluid volume injected and (b) cumulative hydraulic energy E_{hyd}. The theoretical limit on cumulative seismic moment with injected volume following McGarr (2014) assuming a shear modulus of 30 GPa is plotted as a dashed line.

In striking contrast to most EGS projects analyzed here, Cooper Basin and Pohang show a much steeper increase of $M_{0,\mathrm{cum}}$ with volume injected (Figure 3.2a). Both slopes are similar to the initial slopes of the other projects, yet remain high for the entire stimulation, not switching to the linear slope observed for the other seismic sequences. Interestingly, the Paralana dataset shows a jump in $M_{0,\mathrm{cum}}$ during an otherwise linear evolution, caused by the occurrence of several large magnitude earthquakes ($2 < M_\mathrm{W} < 2.5$) compared to the remaining seismicity during the same time period. Following the sudden increase, the trend reverts back to the previous linear slope until shut-in. For the Basel sequence the evolution of $M_{0,\mathrm{cum}}$ follows a linear increase with injected volume for the entire injection period, only to deviate shortly after shut-in due to the occurrence of a $M_\mathrm{W}3.0$ earthquake.

We find a linear relationship between $M_{0,\mathrm{cum}}$ and hydraulic energy E_hyd comparable to the relation between $M_{0,\mathrm{cum}}$ and injected volume. (Figure 3.2b) However, several of the projects collapse to a similar trend or are separated by a distinct offset in seismic moment. The slope of the relation of $M_{0,\mathrm{cum}}$ and E_hyd is about unity for the most part. Different trends are found for the final stimulation phase of Cooper Basin and the post shut-in phase of Basel. Although punctuated by several large events, the Cooper Basin stimulation follows mostly a similar trend as the other projects. In general, most of E_hyd is applied during the period of linear moment increase. Only during the Pohang stimulation $M_{0,\mathrm{cum}}$ increased much more rapidly than for all the other projects.

Note that about the same total amount of E_hyd was applied in several of the projects but resulted in very different cumulative seismic moment release. The later injections into the deep part of the Soultz-sous-Forêts reservoir in 2000 and 2003 display very similar relations between $M_{0,\mathrm{cum}}$ and E_hyd. In contrast, the 1993 stimulation at shallower depth, with comparable fluid volumes injected and hydraulic energy exhibits a similar evolution trend yet more than three orders of magnitude lower total seismic moment release.

3.3.3 Seismic Injection Efficiency

We analyzed seismic injection efficiency I_E for each of the stimulation projects (Figure 3.3). Injections performed at Soultz93, KTB, and BGF display very low I_E, on the order of 10^{-5}. Soultz93 shows a steady increase, while KTB exhibits a steady decrease of I_E during the active stimulations. Efficiency at BGF first increases and then slowly decreases again. In Pohang and Cooper Basin the injection started with relatively low efficiencies, about 10^{-6}. At Pohang the efficiency increased very rapidly during stimulation, interrupted by small periods of decreasing I_E coinciding with pauses in the injection activity. The final value reached at Pohang after the occurrence of the $M_\mathrm{W}5.5$ event is larger than 1, indicating that more energy was radiated seismically than hydraulically injected. Cooper Basin exhibits a very irregular growth of I_E, increasing during stimulation in two major steps to a final value of $2 \cdot 10^{-2}$. Basel and Soultz00 show an overall decreasing efficiency during the course of the stimulation, although the initial trend at Soultz00 is dominated by the early $M_\mathrm{W}1.8$ event within hours of injection start. Paralana and Soultz03 first show an increase of I_E up to 10^{-2}, followed by a slow decrease for the main part of the injection activity. Similarly, for the majority of projects, I_E converges toward relatively stable values between 10^{-3} and 10^{-2}. It is interesting, that a steady decrease of I_E at Basel, still resulted in the large project stopping $M_\mathrm{W}3.0$ event.

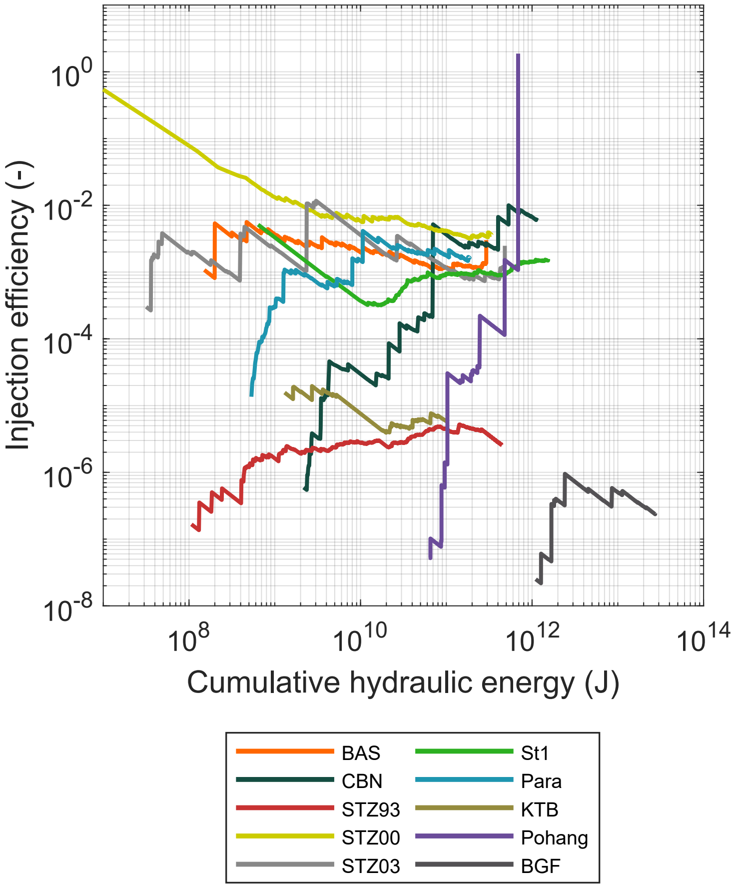

Figure 3.3: Temporal evolution of injection efficiency I_E, i.e. radiated energy per hydraulic energy, against cumulative hydraulic energy.

3.4 Discussion and Conclusion

Recently, Kwiatek et al. (2019) reported on a successful attempt controlling the evolution of seismic magnitudes during an EGS operation by adjusting the stimulation protocol. But whether or not maximum magnitudes of injection-induced seismicity can indeed be controlled and if events have the same magnitude limit as tectonic earthquakes is currently still a matter of debate (van der Elst et al., 2016; Hofmann et al., 2018). It is generally assumed that small perturbations of effective stresses are sufficient to initiate failure to critically stressed faults (Zoback & Townend, 2001). The ensuing rupture is expected to propagate until stored elastic energy is consumed at the rupture tip or the rupture hits a sufficiently strong barrier. However, as for tectonic earthquakes, the conditions controlling rupture arrest and final magnitude of induced earthquakes are still poorly understood (Ripperger et al., 2007; Garagash & Germanovich, 2012; Galis et al., 2017). Controlling induced event magnitudes during a stimulation campaign amounts to successfully maintaining event size within the limits of stable rupture propagation (maximum arrestable magnitude of Ripperger et al. (2007); Galis et al. (2017), pressure-constrained ruptures of Norbeck and Horne (2018)). Existing models relating maximum expected event magnitude to injected fluid volume assume an initial phase of stable growth of cumulative seismic moment before runaway ruptures occurs (McGarr, 2014; Galis et al., 2017). In this study we examined prominent examples of EGS stimulations with regard to maximum observed moment magnitude, seismic moment evolution and injection efficiency in response to injection operations.

A striking observation is that the general properties and evolution of seismicity are largely independent of the tectonic stress regime. The analyzed projects cover a range of tectonic settings from dominantly strike-slip and normal faulting to transpressive regimes (Table 3.1), yet no relationship between $M_{\mathrm{max,obs}}$, seismic moment evolution, or injection efficiency and background stress regime could be observed. Also, we did not observe a substantial variation in b value between different projects where stimulations were performed in different stress regimes (Figure 3.1b) as has been observed previously (Schorlemmer et al., 2005). This is in agreement with the analysis of 41 case studies of fluid injection projects (Evans et al., 2012) and supported by observations of changes in the local stress state as a function of depth (e.g. in Soultz, Cuenot et al., 2006). In addition, it may be expected that reservoir stresses are considerably modified due to poroelastic stress changes caused by fluid injection (Martínez-Garzón et al., 2013; Schoenball et al., 2014). Instead, we suggest that local tectonic features of the reservoir, such as characteristic fault length, orientation, and frictional properties, govern the seismic response of the reservoir. This may be illustrated by comparing different stimulation phases in Soultz. Here the seismic activity level and $M_{\mathrm{max,obs}}$ change substantially between stimulations performed at 3.5 km depth in 1993 and 4.5 – 5 km depth in 2000 and 2003 (Figure 3.1). During later stimulations, much larger event magnitudes are observed for similar fluid volumes injected, possibly related to a strongly perturbed stress state from previous injections and/or activation of a different structural inventory.

Existing models predicting maximum magnitudes with injected volume fit the data reasonably well. Accounting for site-specific parameters, such as b value, some data may better be approximated by the rupture physics-based model of Galis et al. (2017) or the model of van der Elst et al. (2016), while others follow the relationship of McGarr (2014). However, what is important to note is, that for most of the sites the temporal evolution of maximum observed magnitudes with the injected volume generally follows a constant slope, at least for substantial parts of the injection. While the maximum expected magnitude is considered a key parameter for each injection site, we find that continuous tracking of temporal changes of cumulative seismic moment in relation to injection parameters provides insight into site-specific seismicity evolution and potentially provides a first-order seismic hazard information. The data show two distinct populations among the described projects. A larger group displays a stable evolution of cumulative seismic moment across the largest part of injection activity: this includes 8 out of 10 analyzed data sets. We interpret this behavior following Galis et al. (2017) as an indication for stable growth of self-arrested ruptures. In this scenario the size of the earthquakes is likely governed by the induced pore-pressure perturbation (Norbeck & Horne, 2018). If the pore pressure, that is, the injected volume or applied hydraulic energy, grows, so does the rupture area. This would result in constantly increasing magnitudes and cumulative moment with injected volume, a behavior we can clearly observe (Figure 3.2). Interestingly, the slope characterizing increase in cumulative seismic moment with injected fluid volume and with hydraulic energy input is close to unity for most stimulations. This is in excellent agreement with the prediction of McGarr (2014) suggesting a linear relation between volumetric strain in the reservoir and cumulative seismic moment. Most of the sequences start with a short phase of steeply increasing seismic moment with injected fluid, which stabilizes to a constant slope afterward (Figure 3.2a). This may be seen as an initial adaption phase of fracture opening and fault activation around the well with a subsequent phase of stable rupture propagation following the expanding pore pressure front.

The second group of projects, including Pohang and Cooper Basin, shows a steep increase of maximum magnitude and cumulative seismic moment, which appears to grow unbound and does not stabilize. These trends are not captured by any of the models and suggest activation of runaway ruptures, where rupture size is only limited by the size of tectonic faults. For Pohang, immediately after onset of injection, seismic moments start to grow seemingly unbound. This is in contrast to the Basel and Cooper Basin project. The substantial part of Basel injection remains stable only to produce a large $M_\mathrm{W}3.0$ event after shut-in. This highlights the problem that exceeding the maximum size of arrestable ruptures may occur at any point during injection, resulting in unstable rupture propagation Galis et al. (2017). In addition, typically a significant time delay may exist between seismic response and change of injection parameters due to hydraulic diffusion characteristics of the reservoir. The Cooper Basin injection displays time periods of stable growth, punctuated by substantially larger events that may represent unstable ruptures. This explanation of the behavioral change is also indirectly contained in the assumptions of the current models. Following McGarr (2014) and Kostrov (1974), the ratio between average stress drop and the volume affected by injection should scale linearly with cumulative seismic moment. Any observed change in this scaling would indicate a considerable change in either stress drop or stimulated volume, pointing at a fundamental property change of the injected system.

Seismic injection efficiencies observed for the different sites vary between 7 orders of magnitude. KTB, Soultz93 and BGF show very low efficiencies, below 10^{-4}, indicating a seismically very inefficient process, such as fracture creation (Maxwell et al., 2008) or thermal diffusion in a high-heat environment in the case of BGF (Kwiatek, Bulut, et al., 2014). Most of the other projects converge toward 10^{-3} to 10^{-2}, values typical for tectonic earthquakes (Kanamori, 2001). Suggesting that the same energy balance holds between tectonic and most injection-induced earthquakes in crystalline basement. The large seismic injection efficiency at Pohang, suggesting more energy radiated seismically than injected hydraulically, most likely reflects that the energy released through the large $M_\mathrm{W}5.5$ event was dominantly elastically stored on an already critically pre-stressed fault. While we assumed a median radiation efficiency η_R of 0.46 (McGarr, 1999), the percentage of energy associated with stress drop that is radiated could potentially vary from site to site. However, McGarr (1999) reported values of radiation efficiency between 0.08 and 0.92, which would lead only to very small changes in the computed injection efficiencies compared to the overall trends (Figure 3.3).

In agreement with suggestions of Galis et al. (2017), our observations of a potential transition from stable to unstable growth of seismic moment of injection-induced seismicity has important implications for the design of traffic light systems aiming at controlling the seismic hazard and risk imposed by hydraulic stimulations. The physical model predictions appropriately describe the pressure-controlled injection phase in terms of seismic moment evolution. Any potential trends of seismic moment evolution observed during injection that are completely at odds with the models should be considered seriously, possibly leading to a reassessment of further injection operations. For example, a behavior pointing at unstable rupture growth from the start of injection as found for Pohang, needs to be identified by monitoring not only the exceedance of a critical maximum magnitude but also by near-real-time detection of spatio-temporal evolution of event magnitudes or seismic moment with injected volume. It would be crucial to differentiate between stimulations resulting in runaway ruptures right from the beginning of the injection activities and stimulations that may become stable after an initial adaption phase. So far, the only observed clue to differentiate the two cases is the injected fluid volume up until the steep seismic moment increase persists. In a multistage stimulation project these observations could be utilized by analyzing the seismic moment evolution during each injection stage and modify subsequent injection strategies based on the observed trends. Such a strategy was successfully tested in the recent St1 stimulation project in Helsinki (Kwiatek et al., 2019). Second, the case of a self-arrested rupture becoming potentially unstable with further fluid injection needs to be identified early enough for operators to change injection strategies or cease stimulation. In practice this may prove difficult, as the Basel stimulation showed. Seismic moment and maximum magnitude growth were stable and injection efficiency was decreasing during main stimulation: nevertheless, the stimulation resulted in the project-stopping $M_\mathrm{W}3.0$ event after shut-in. Thus, trends in the seismic efficiency are not necessarily a reliable test for safe operations. However, the recent success in managing maximum magnitudes in the St1 stimulation in Helsinki, through close interaction of operator and seismic monitoring, highlights the potential of a new generation of traffic light systems combining near-real-time seismic monitoring and subsequently adapted injection strategies. Optimizing such rapid-response traffic light systems may be key to safe development of geothermal reservoirs even in urban areas. This may assist in geothermal energy as a valuable and economic contribution to a mix of renewable energies in implementing the ongoing energy transition.

4 Earthquake Nucleation at The Geysers Geothermal Field, California

Summary

Preparatory mechanisms accompanying or leading to nucleation of larger earthquakes have been observed at both laboratory and field scale, but conditions favoring the occurrence of observable preparatory processes are still largely unknown. In particular, it remains a matter of debate why some earthquakes occur spontaneously without noticeable precursors as opposed to events that are preceded by an extended failure process. In this study, we have generated new high-resolution seismicity catalogs framing the occurrence of 20 $M_\mathrm{L} > 2.5$ earthquakes at The Geysers geothermal field in California. To this end, a seismicity catalog of the 11 days framing each large event was created. We selected 20 sequences sampling different hypocentral depths and hydraulic conditions within the field. Seismic activity and magnitude frequency distributions displayed by the different earthquake sequences are correlated with their location within the reservoir. Sequences located in the northwestern part of the reservoir show overall increased seismic activity and low b values, while the southeastern part is dominated by decreased seismic activity and higher b values. Periods of high injection coincide with high b values, and vice versa. These observations potentially reflect varying differential and mean stresses and damage of the reservoir rocks across the field. About 50 % of analyzed sequences exhibit no change in seismicity rate in response to the large main event. However, we find complex waveforms at the onset of the main earthquake, suggesting that small ruptures spontaneously grow into or trigger larger events.

Published as Bentz, S., Martínez-Garzón, P., Kwiatek, G., Dresen G., Bohnhoff, M. (2019). "Analysis of Microseismicity Framing $M_\mathrm{L} > 2.5$ Earthquakes at The Geysers Geothermal Field, California" in *Journal of Geophysical Research: Solid Earth*, 124, doi: 10.1029/2019JB017716.

4.1 Introduction

Characteristic spatial and temporal patterns of seismicity indicating preparation and nucleation of large magnitude earthquakes would allow to significantly improve short-term earthquake forecasting and subsequent hazard and risk mitigation. For decades, intensive research efforts have been made with the goal of detecting earthquake preparation processes, such as statistical analysis of seismicity catalogs, analysis of foreshocks or identification of seismic coalescence (e.g., Jones, 1985; Eneva & Ben-Zion, 1997; Mignan, 2012, 2014). Since the 1980s, the occurrence of foreshocks accompanied by accelerated seismic moment release (e.g., Jones & Molnar, 1979; Papazachos, 1975; Shaw et al., 1992), increase in the seismicity rates (e.g., Ellsworth et al., 1981) or an increase in Benioff strain (e.g., Ben-Zion & Lyakhovsky, 2002) have been examined as potential signatures to forecast the short-term occurrence of a larger earthquake. Detailed analysis of seismic events involving foreshocks or weak precursor phases as observed for some earthquakes, has motivated classical cascade and pre-slip nucleation models (Ellsworth & Beroza, 1995).

However, the occurrence of such potentially precursory signatures is not systematic in space or time and the conditions favoring them are still a matter of debate (e.g., Kanamori, 1981; Mignan, 2011, and references therein). Some recent large earthquakes have been identified to be preceded by foreshock sequences, such as the 1999 $M_\mathrm{W}7.4$ Izmit earthquake (Bouchon et al., 2011), the 2010 $M_\mathrm{W}7.2$ El Mayor-Cucapah earthquake (Hauksson et al., 2010) or the 2011 $M_\mathrm{W}9.1$ Tohoku-Oki earthquake (e.g., Kato et al., 2012). In contrast, others have shown that foreshocks do not exist systematically (e.g., Zechar & Zhuang, 2010), or that they are indistinguishable from background seismicity (Hardebeck et al., 2008). Furthermore, Jones (1985) estimated that only 6 % of the earthquakes in southern California are followed by an earthquake of larger magnitude within 5 days and a distance of less than 10 km. This is in contrast to the more recent work of Bouchon et al. (2013) and Brodsky and Lay (2014) who suggest that foreshocks are more common. This is in agreement with many laboratory studies showing foreshock activity prior to stick slip events (Lei et al., 2003; Goebel et al., 2012; Selvadurai & Glaser, 2015). Zaliapin and Ben-Zion (2013b) identified seismicity clusters using nearest neighbor distributions in a space-time and magnitude domain and found that an increased occurrence of foreshocks occurs in areas displaying high heat flow. In laboratory experiments, acoustic emissions during rock deformation are frequently identified as precursors for the main rupture. A distinct drop in b value preceding the failure of the rock specimen was first described by Scholz (1968) and subsequently confirmed in many studies (Meredith et al., 1990; Zang et al., 1998). Furthermore, Goebel et al. (2012) found that geometric asperities identified in CT scan images were connected to regions of low b values, increased event densities and moment release over multiple stick-slick cycles. Most recently, Rouet-Leduc et al. (2017) used machine learning to predict the time remaining until failure in shear laboratory experiments by analyzing acoustic emission signals emitted from the fault zone.

In fluid-induced seismicity, Walter et al. (2017) identified decreasing b values leading up to the $M_{\mathrm{W}}5.8$ Pawnee earthquake in Oklahoma, suggesting that differential stress was slowly building up along the fault in the months prior to the earthquake (Scholz, 2015). Additionally, Savage et al. (2017) observed an increase in localization of foreshocks of the $M_{\mathrm{W}}5.0$ Prague earthquake in Oklahoma, initially seen in the entire damage zone and later localizing into a 100 m thick narrow zone close to the mainshock.

In order to fully understand the presence or lack of these precursory signals, we have studied entire earthquake sequences, namely, foreshocks, mainshock, aftershocks, and their relation and interplay from The Geysers geothermal field in California. In this context, Martínez-Garzón et al. (2018) previously investigated the relationship between aftershock productivity and injection activity at different geothermal fields in California (including The Geysers) and found a positive correlation of aftershock productivity with net-produced volume. Their results indicate that anthropogenic activity might have a significant influence on foreshock and aftershock patterns in induced seismicity environments.

In addition, high-resolution seismic monitoring allowing for the detection and location of small seismic events may also promote the identification of changing patterns surrounding large event sequences. A recent meta-analysis based on 37 independent studies showed that the nucleation process of earthquakes tends to be detected if the minimum magnitude of the available earthquakes in the catalog is at least three orders of magnitude smaller than the selected mainshock (Mignan, 2014), typically constraining the detection of earthquake foreshocks (if any) to the largest events included in a seismicity catalog. This is supported by laboratory observations, where high detection capabilities allow anticipating the pending main event before the macroscopic failure occurs (Goebel et al., 2012; P. A. Johnson et al., 2013; Kwiatek, Goebel, & Dresen, 2014). Still, several foreshocks signifying the preparation process of a $M_{\mathrm{W}}4.2$ earthquake in the eastern Sea of Marmara (Turkey) could be successfully detected using borehole instrumentation (Malin et al., 2018). Yoon et al. (2019) utilized recently developed data mining techniques to investigate the nucleation phase of the 1999 $M_{\mathrm{W}}7.1$ Hector Mine earthquake, identifying 50 foreshocks in the 20 hours before the mainshock. Given already existing data and deployed surface instrumentation, extracting the smallest possible events using state of the art detection methods, is key in identifying physical processes related to large ruptures. That way, Meng et al. (2012) detected 70 times the number of earthquakes listed in the official catalog, when applying a matched filter technique to seismic recordings at the Salton Sea geothermal field surrounding the time of the $M_{\mathrm{W}}7.2$ El Mayor-Cucapah earthquake. This allowed capturing significant seismicity rate changes due to both dynamic and static stress changes.

The Geysers geothermal field in northern California shows high seismicity rates, good azimuthal coverage by a local seismic network, and a long history of continuous local high-quality seismic monitoring. The excellent monitoring conditions there allowed collecting one of the most extensive datasets worldwide of fluid-induced seismicity. Here, we analyze the microseismicity surrounding the occurrence of 20 moderate to large magnitude ($2.5 < M_\mathrm{L} < 4.5$) earthquakes at The Geysers geothermal field to investigate how the properties of large earthquake sequences are influenced by ongoing anthropogenic activity and whether large events are preceded by a visible preparation process and how systematic such observations are. To do so, we utilize a matched filter algorithm to extract the maximum amount of data from the continuous seismic recordings from a local high density broadband network. Merging with existing catalogs and restricting to events located close to the location of each main event, frequency-magnitude distributions are determined aiming to characterize the influence of different geomechanical properties and reservoir characteristics on the behavior of each earthquake sequence. Additionally, a systematic search for potential precursory patterns for each of the sequences is performed. The findings are discussed and related to local stress variations and different damage states within the reservoir rock formation.

4.2 Study Region

The Geysers geothermal field is located in the Mayacamas Mountains roughly 110 km north of the San Francisco bay area in northern California (Figure 4.1). With an installed capacity of more than 1,500 MW across 22 power plants, it is currently the largest geothermal field in the world. Steam is extracted from the vapor-dominated reservoir from approximately 322 active production wells. The field has been producing geothermal energy since the 1960s. Following its peak in 1987, however, production has been declining ever since (Gunasekera et al., 2003), most likely caused by a combination of decreasing reservoir pressure and cooling (Mossop & Segall, 1997). To stabilize production and achieve a net fluid balance within the reservoir, large volumes of treated waste water are injected across some 54 wells (E. L. Majer & Peterson, 2007).

Tectonically, this region is dominated by right-lateral strike-slip motion accommodating the relative movement between the North American Plate and the Pacific Plate (Lisowski et al., 1991). The geothermal field is bounded by two regional faults (Figure 4.1). Based on GPS-derived slip rates below 1 mm/yr the Collayomi Fault is considered to be currently inactive, while the Maacama Fault is considered active based on an average slip rate of about 13 mm/yr (Murray et al., 2014). The reservoir is not capped by a geologically well-defined formation, but rather by silica deposition along fluid flow paths resulting in self-sealing of the hydrothermal system (Facca & Tonani, 1965). Today's reservoir is well explored through the large number of wells drilled, and is primarily composed of low porosity (1 – 2 %) but highly fractured greywacke (Lipman et al., 1978). Temperatures across the field vary significantly, between 240 °C at 2 km depth in the southeastern and more than 350 °C at 2.75 km in the northwestern part of the reservoir (Rutqvist et al.,

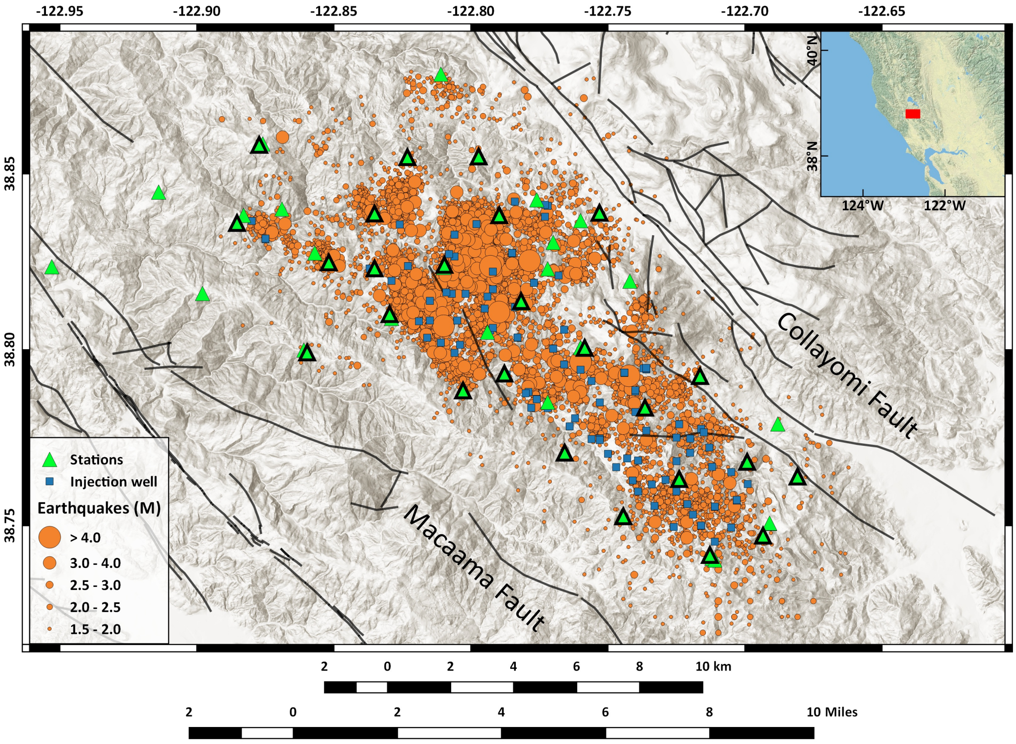

Figure 4.1: Overview map of The Geysers geothermal field with the main faults(black lines), injection wells (blue squares), occurring seismicity ($M_{\mathrm{L}} > 1.5$) recorded during 2012 and 2013, and the deployed seismic stations (green triangles; black frames indicate the position of the temporary broadband instruments).

2016).

During the first half of the 20th century, The Geysers region was characterized by low seismic activity. Since the active exploitation of the field began, an increasing number of earthquakes have been reported (e.g., Eberhart-Phillips & Oppenheimer, 1984; Oppenheimer, 1986). A dense local short-period network was deployed in 2003, and ever since allows to capture an annual seismicity rate of more than 4,000 earthquakes above magnitude $M_{\mathrm{L}} > 1$ (Figure 4.1). However, seismic activity is not homogeneous across the field, with variations in focal depths (C. W. Johnson et al., 2016; Trugman et al., 2016) and frequency magnitude distributions (Convertito et al., 2012; Trugman et al., 2016) between the northwestern and southeastern part of the field, as well as strong seasonal fluctuations in seismicity rate (C. W. Johnson et al., 2016; Trugman et al., 2016). Inversion of seismic data for the local maximum horizontal stress shows a general agreement with the NNE-SSW trending regional tectonic structures (Eberhart-Phillips & Oppenheimer, 1984; Martínez-Garzón et al., 2013). The increased seismicity since the start of production has been associated with the ongoing anthropogenic activity through a number of possible physical mechanisms, such as reduced effective stresses due to changes in pore pressure (E. L. Majer & Peterson, 2007; Martínez-Garzón et al., 2014). Additionally, temperature contrasts between the injected water and the hot reservoir rock can lead to thermal fracturing close to the injection points (Segall & Fitzgerald, 1998; Jeanne et al., 2014) or cooling induced geochemical alterations causing microseismic activity (Allis, 1982). Finally, steam production and the resulting reservoir depletion, observed through increased subsidence at the surface (Gunasekera et al., 2003), could further modify the effective stress within the formation.

4.3 Data

Seismicity across The Geysers geothermal field is recorded by a permanent surface network of 31 short-period stations operating since 2003, run by the Lawrence Berkeley National Laboratory (Berkeley-Geysers seismic network, IRIS code BG). In the frame of the European Commission GEISER project (EC grant number 241321), this network was supplemented with 26 broadband stations between June 2012 and July 2013, all collocated with selected short-period stations. Each site was equipped with 60 s Guralp or 120 s Trillium sensors, sampled at 200 Hz. The recorded broadband, high-density, and good azimuthal coverage continuous recordings resulted in one of the highest quality datasets for geothermally induced seismicity. Main noise components were found to be microseism (0.16 Hz) and anthropogenic noise (10 – 20 Hz) (Yu et al., 2018).

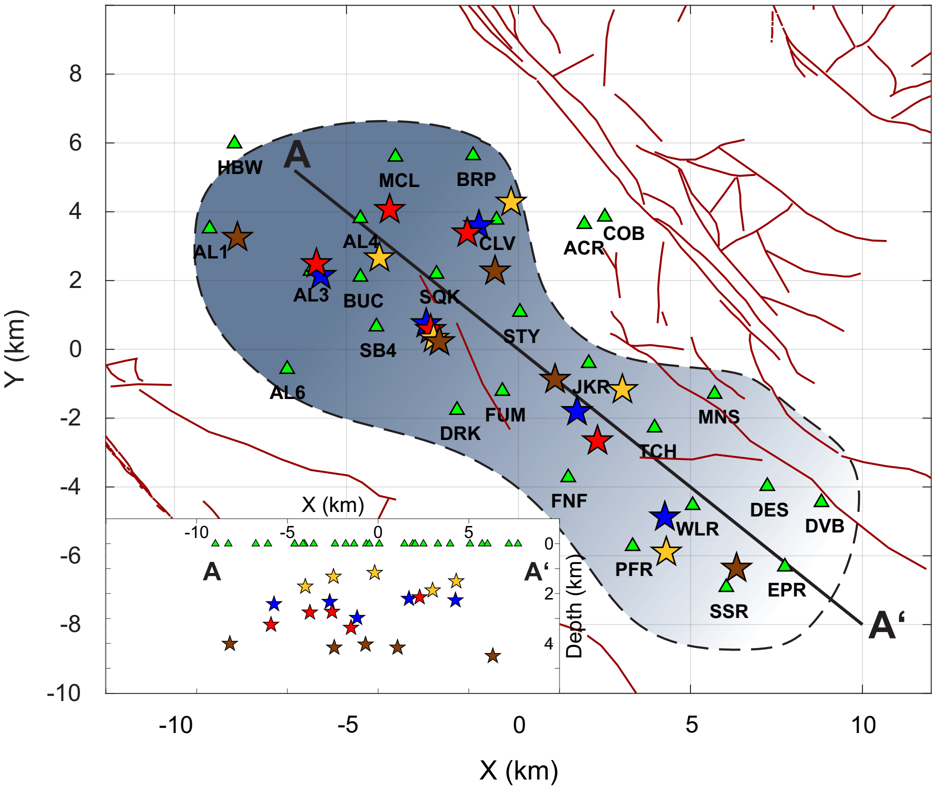

Figure 4.2: Location of the 20 master events with $M_L > 2.5$ analyzed in this study, as well as the broadband station configuration (green triangles), and the local faults (brown lines). The geothermal field outline is shown by the dashed area, and the dipping reservoir is indicated by the changing shades. Event projections onto the cross section A-A' are plotted in the bottom left. Color codes are used to distinguish the different earthquake groups: red and blue indicate events located within the reservoir during low and high injection periods, respectively, and yellow and brown indicate events located above and below the reservoir, respectively.

Based on automatic processing of the short-period network, two different earthquake catalogs are available through the North California Earthquake Data Center (NCEDC). We utilized the local EGS catalog and corresponding phase picks between June 2012 and July 2013 produced by the Lawrence Berkeley National Laboratory (LBNL), which represents the most detailed local catalog for the geothermal field (Denllinger et al., 2017). During this time period a total of 120 earthquakes with $M_{\mathrm{L}} > 3$ were observed. Based on this catalog, 20 master events were selected for our analysis (Figure 4.2). The locations of these events were chosen as to sample different mean stresses (i.e. by covering the entire possible depth range) as well as fluid pore pressures (i.e. by selecting events occurring during periods of relatively low and high injection flow rates). We selected 20 events starting with the largest magnitudes. If the next potential event was not sampling a new setting of the field, we progressed further down the magnitude list. We chose 10 events located between 2 – 4 km depth, representing the overall reservoir depth range across the majority of the field (although the reservoir tends to be shallower and deeper in the SE and NW, respectively). Out of these 10 events within the reservoir, we took five each during high and low injection periods, respectively. Then, we chose five events from above and five from below the reservoir. During the selection process, we ensured to spatially sample the field as evenly as possible. For each of the 20 master events, between 50 and 120 adjacent events located within their source region, that is, 200 – 500 m hypocentral distance, were taken as templates for the matched filter technique, supplementing the mainshock waveforms. Both P and S wave onsets were manually re-picked for each template. Templates and continuous waveforms were detrended and filtered between 4 and 40 Hz using a second-order Butterworth bandpass filter. For each master event and all available stations, we sorted 5 days of continuous recording before and after the origin day of the master event, together with the day of each master event into 24 hour long sections.

4.4 Methodology

4.4.1 Matched Filter Algorithm Implementation

The core of the matched filtering algorithm (MFA) is a network wide cross-correlation between a time-series and a known reference signal (e.g., Shearer, 1994; Shelly et al., 2007),

$$c_i(t) = \frac{\sum_{n=1}^{N} v_i(t_n) u_i(t_n)}{\sqrt{\sum_{n=1}^{N} v_i(t_n)^2 u_i(t_n)^2}}, \tag{4.1}$$

where $c_i(t)$ is the normalized cross-correlation coefficient at the i-th station between the continuous data $u_i(t)$ and the template waveform $v_i(t)$ (Gibbons & Ringdal, 2006). For big datasets with small sampling intervals this task is computationally expensive, usually requiring a high-performance computing cluster (Meng et al., 2012; Beaucé et al., 2017). Fortunately the base problem exhibits parallelism on multiple levels, which can significantly reduce computation time (Meng et al., 2012). For each component, multiple templates need to be correlated with the continuous recordings on all stations and components, which can be split up into multiple sub-series. Parallelizing all these steps is an obvious way of reducing computational time when memory allocation and sharing is considered.

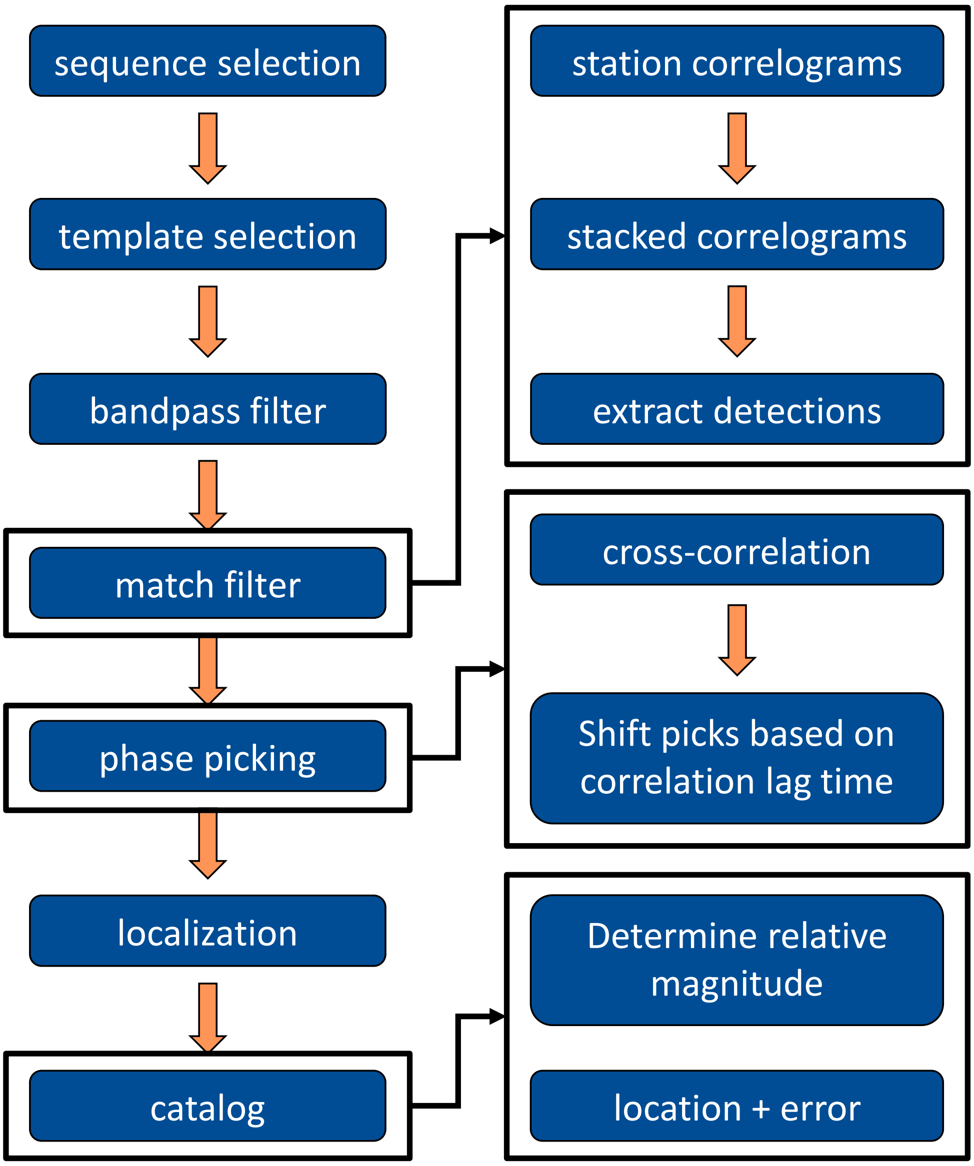

Figure 4.3: Flow chart of the analysis steps performed in this study. On the left are all individual steps in chronological order, and on the right there is detailed information on the core steps.

We implemented the MFA as follows: 1) a single data input operation is performed that loads the continuous and template data into the memory as separate arrays; 2) the autocorrelation sum of the continuous data and template is precomputed once for later use in the normalization of the correlation calculation (denominator in equation 4.1); 3) the correlation is parallelized for each component and template into a maximum number of threads; 4) for each thread the calculation of the cross-correlogram, that is, the correlation coefficient with time, is performed in the frequency domain, as a convolution of the continuous data with the time reversed template; 5) the cross-correlogram is normalized by the precomputed values from 2) (Figure 4.3).

After accounting for the station moveout the cross-correlograms are stacked across the stations for each template, positive peaks on the stacked cross-correlogram now represent detections with approximately the same location as the used template event. Detections are extracted by applying a threshold of 7.5 times the median absolute deviation (MAD). For a normally distributed random variable the probability of exceeding 7.5 times the MAD is 2.1×10^{-7}. Testing the use of different MAD thresholds and component combinations on small test datasets showed that using only the P wave on vertical component of each station and a template time interval of about 1 s prior to and 3 s after the phase arrival yielded the most robust detection results. We allowed only for a single detection within a 3 s window, in case of multiple detections we kept the detection with the highest mean correlation coefficient across the stations and rejected the remaining.

4.4.2 Phase Picking and Hypocenter Determination

Determination of the exact arrival times for each new detection is based on cross-correlation lag time. We extracted 10-s-long waveforms framing the detection time and cross-correlated them once more with the data from each template at the different stations and components in a narrower time window based on expected travel times along the network, that is, 0.3 s before and 1 s after the manual pick. Vertical components were used to determine the P wave arrival, horizontal components for the S wave. For the template showing the highest cross-correlation, the picks were computed as the time lag between the cross-correlation maximum and the template pick. We constrained further processing to detections that have at least 6 picks with at least 2 picks for each phase (P and S). For each sequence all template phase onsets were first automatically re-picked using the correlation with all remaining templates. The new onsets were used to locate each template again. If the location deviated horizontally for more than 150 m from the catalog location that particular template was excluded from the correlation with the detections. Hypocenter locations of events were determined using maximum intersection method, where the earthquake hypocenter is determined independently of the origin time through the use of the equal differential time (EDT) surface (Font et al., 2004). For any pair of arrivals, such as the same phase on two different stations or two different phases on the same station, an EDT surface is defined as the set of all points in the subsurface that satisfy the time difference between the observed arrivals, assuming a fixed velocity model (Zhou, 1994). The intersection of all possible EDT surfaces is equal to the global minimum of the cost function used:

$$\mathrm{CF} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\delta T_{\mathrm{obs},i} - \delta T_{\mathrm{th},i}\right)^2}, \tag{4.2}$$

where $\delta T_{\mathrm{obs},i}$ is the travel time difference between the observed i-th arrival pair, $\delta T_{\mathrm{th},i}$ the corresponding theoretical difference, and N the total number of arrival pairs. Theoretical arrival times were calculated using a local one-dimensional velocity model (Eberhart-Phillips & Oppenheimer, 1984). In the following, we used the Metropolis-Hastings random walk algorithm (Hastings, 1970) to sample the three-dimensional space of hypocenter locations. The standard deviation for the probability density function (PDF) for each location is on average less than 200 m (Figure 4.4). Horizontal location differences for events already located are usually within 150 m, but vertical locations may vary significantly, up to 400 m. It is important to note, that the new locations occasionally shift seismic clusters as a whole, but do not modify the relative locations of events to each other.

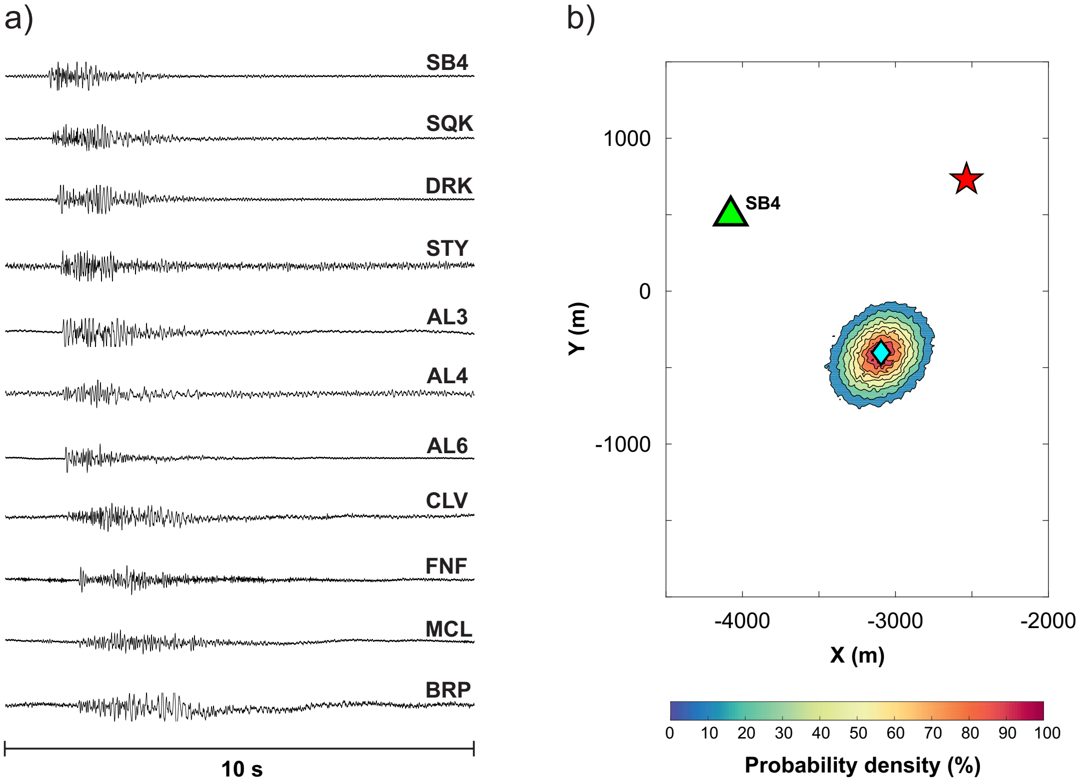

Figure 4.4: Example detection on the 13/01/2013 at 17:24:44 GMT with M_{L} 0.7. (a) 10 s of waveforms across the network. (b) The result of the Metropolis-Hastings localization algorithm. The initial location based on the template is indicated by the red star, while the best fit solution including its probability density function is shown surrounding the blue diamond.

4.4.3 Magnitude of Completeness and b Value Estimation

We calculate the magnitude of each new detected event by comparing the amplitudes between the waveforms and the templates used to find it, assuming that a tenfold increase in amplitude corresponds to 1-unit increase in magnitude:

$$M_{\mathrm{det}} = M_{\mathrm{temp}} + \log_{10}\left(\frac{A_{\mathrm{det}}}{A_{\mathrm{temp}}}\right), \tag{4.3}$$

where M_{temp} is the template magnitude, and A_{det} and A_{temp} are the peak P wave amplitudes of detection and template, respectively (Peng & Zhao, 2009; Schaff & Richards, 2014; Vuan et al., 2018). However, equation 4.3 assumes the same location and wave paths for template and detection. To avoid dispersion due to small differences in location we use a site specific regression: first performing a linear fit between template magnitude and the logarithm of the maximum amplitude for each station, and then calculating a relative magnitude based on that fit. The standard deviation is taken as a measure of error for the relative magnitude determination.

Assuming a Gutenberg-Richter power law distribution of magnitudes:

$$\log_{10} N = a - bM, \tag{4.4}$$

where N is the cumulative number of earthquakes having magnitudes larger than M and a and b are constants; we calculate the b and a values as a function of the minimum magnitude, M_i, using a maximum likelihood estimate (Bender, 1983). Following the goodness-of-fit test of Wiemer and Wyss (2000), we compute a synthetic distribution of magnitudes using the same b, a, and M_i values, and calculate the absolute difference, R, of the number of events in each magnitude bin between the observed and synthetic distribution:

$$R(a, b, M_i) = 100 - \left(\frac{\sum_{M_i}^{M_{\max}} |B_i - S_i|}{\sum_i B_i} \times 100 \right), \tag{4.5}$$

where B_i and S_i are the observed and predicted cumulative number of events in each magnitude bin, respectively. We define the magnitude of completeness, M_{C}, at $= 90\,\%$, that is, the point at which a power law can fit $90\,\%$ of the seismicity or more. For comparison, M_{C} is also calculated using the maximum curvature of the frequency-magnitude distribution (Wiemer & Katsumata, 1999). To assess the stability of the obtained b values we perform a Monte Carlo experiment, where the magnitudes are assigned an error drawn from a normal distribution with a maximum value determined by the variations in the relative magnitude calculation. This resampling is performed 10^5 times and provides a reliable probability density function for the frequency-magnitude distribution.

4.5 Results

4.5.1 Earthquake Catalogs

Applying the matched filter technique to seismic sequences around the 20 selected master events at The Geysers yielded a significant improvement in the detection threshold compared to the existing high-quality local catalogs. We were able to detect a total of more than 63,000 earthquakes in the 11-day periods surrounding 20 master events analyzed (Figure 4.5 and Table 4.1). Application of the criteria defined for earthquake locations (see the Methodology section) resulted in 47,333 located earthquakes, representing a 75 %

success ratio between detections and accurate locations. Most of the events that we were unable to locate were part of a small number of sequences, resulting in much higher location ratio (around 85 %) for the remaining sequences. This is visible in the congruence between the frequency magnitude distributions from the existing catalogs and the newly derived catalog for the analyzed time periods (Figure 4.5). Compared to the 32,099 seismic events already existing in the catalog, our new detections almost doubled the number of events. This increased the number of successfully located events by 50 % in total. However, this improvement was not uniform across the different sequences, ranging from more than a twofold increase to an actual decrease in located events compared to the local catalog in three cases (Table 4.1). This decrease in the number of detections was most likely caused by a limited number of templates available for three sequences located near the outer edges of the geothermal field, which resulted in a reduced waveform variability. Comparison of the frequency-magnitude distributions between the original local catalog and a new merged catalog, including all new detections as well as the already located earthquakes included in the Lawrence Berkeley National Laboratory catalog, showed very similar trends (Figure 4.5). Both catalogs fulfill a unimodal Gutenberg-Richter relation very well. The distributions tend to diverge for the low magnitudes, demonstrating the catalog improvement achieved.

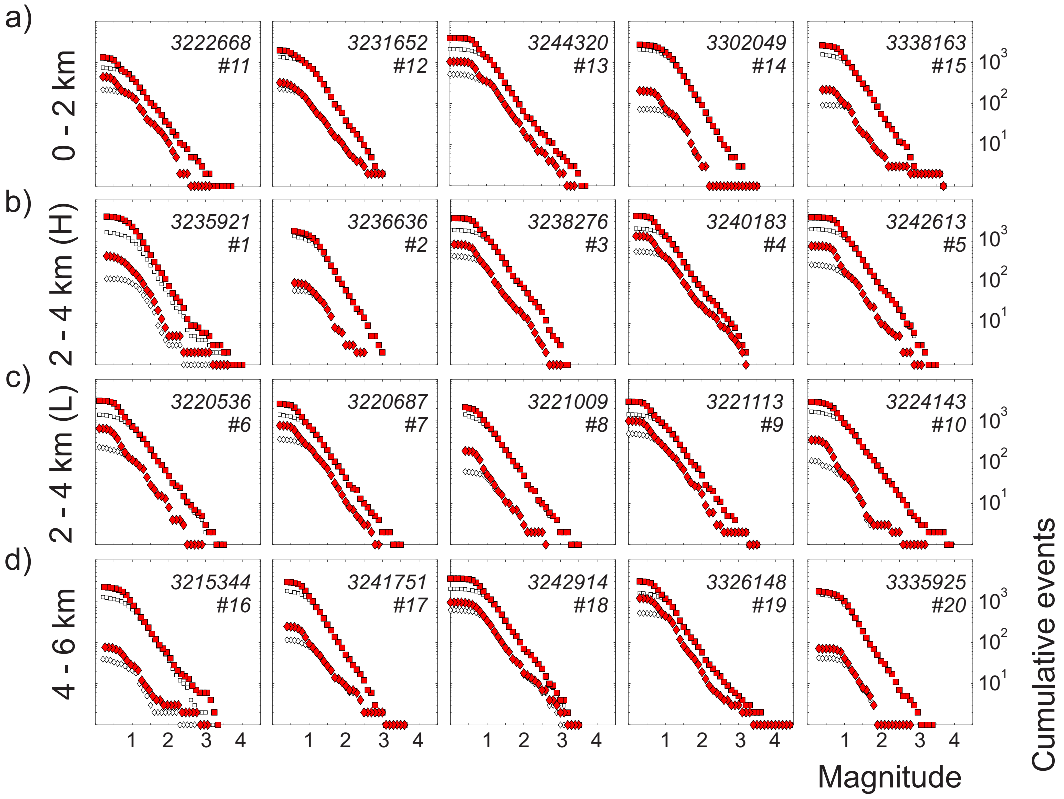

Figure 4.5: Comparison of magnitude frequency distributions for each event group between our detections and the existing local catalog. Squares show the cumulative number of all located events, diamonds only events withon 2,000 m of the large-magnitude event. Red symbols indicate our merged catalog, white symbols only the LBNL EGS catalog. (H) and (L) in the caption denote sequences during high and low injection periods, respectively.

Table 4.1: Detailed information for the 20 mainshocks and their sequences analyzed.

	Origin Time (UTC)					Matched filter			
No.	dd/mm/yyyy	HH:MM:SS.SS	Event ID	M_L	Depth (km)	No. of templates	No. of detections	No. of locations	Catalog improvement (%)
1	17/12/2012	00:03:27.32	3235921	3.6	2.2	117	4,650	2,417	146
2	23/12/2012	06:45:16.55	3236636	2.5	2.3	63	2,540	911	68
3	03/01/2013	10:51:44.51	3238276	3.0	2.9	82	3,832	3,367	182
4	13/01/2013	15:38:52.47	3240183	3.2	2.3	97	5,456	3,778	187
5	28/01/2013	02:26:43.56	3242613	3.1	2.4	55	3,806	3,497	179
6	31/08/2012	23:03:28.23	3220536	2.9	2.8	92	3,636	2,938	207
7	01/09/2012	23:32:52.22	3220687	2.9	2.7	86	2,626	2,339	164
8	04/09/2012	03:19:24.29	3221009	2.6	2.1	72	2,289	1,509	106
9	05/09/2012	01:32:58.80	3221113	3.6	3.4	93	3,644	2,734	190
10	27/09/2012	05:58:58.94	3224143	3.2	3.2	81	3,153	2,554	152
11	16/09/2012	18:54:00.41	3222668	3.1	1.3	70	1,303	1,166	155
12	18/11/2012	10:55:53.40	3231652	3.0	1.2	46	1,698	1,563	115
13	06/02/2013	04:05:30.67	3244320	3.1	1.7	58	3,915	3,525	168
14	08/03/2013	07:59:45.01	3302049	3.5	1.5	68	3,102	1,584	76
15	01/06/2013	23:00:45.91	3338163	3.5	1.9	71	2,751	2,198	140
16	20/07/2012	04:08:57.21	3215344	2.7	4.0	54	3,369	1,886	151
17	22/01/2013	17:58:40.99	3241751	3.6	4.2	62	2,754	2,507	143
18	29/01/2013	16:45:25.40	3242914	3.5	4.0	59	3,433	3,296	166
19	13/04/2013	22:22:04.14	3326148	4.4	4.2	61	3,046	2,816	179
20	17/05/2013	07:25:47.56	3335925	2.8	4.5	44	2,038	748	51

Note. The number of locations refers to the detections that passed through our strict quality criteria and were located using the maximum intersection method. Catalog Improvement is the ratio between the number of locations and the number of events already in the Lawrence Berkeley National Laboratory Enhanced Geothermal System (LBNL EGS) catalog.

The increase in detected and located events is much more distinct at close distances to the master event. For this reason, for further analysis we selected only seismic events occurring within a given radius of the master mainshock. Different distance thresholds were tested, ranging from 500 m to over 2 km. Assuming standard scaling relations between magnitude, and source radius (Kwiatek et al., 2011) and an average stress drop on the order of 0.1 – 10 MPa (Kwiatek et al., 2015), the source region for the largest earthquake analyzed was between 1 and 1.5 km. For consistency, we choose to fix the distance threshold across all sequences to a 2 km radius from the mainshock. Following this approach, *b* values with corresponding uncertainties were recalculated for each sequence, restricted to events located within 2 km of the mainshock epicenter (Figure 4.5, Table 4.2). Our Monte Carlo resampling which accounted for the possibility of magnitude errors produced generally stable *b* values (Table 4.2 and Figure 4.6).

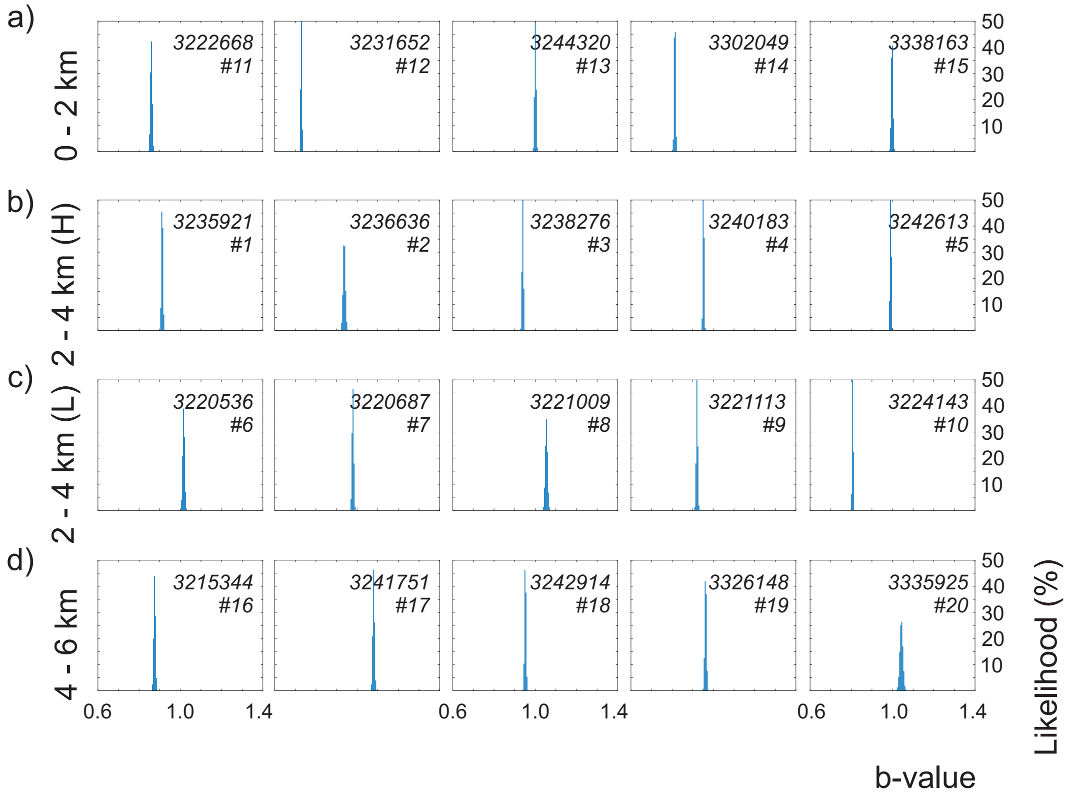

Figure 4.6: Comparison of *b* value stability obtained through the Monte-Carlo resampling accounting for errors in magnitude estimation. (H) and (L) in the caption indicates sequences during high and low injection periods, respectively.

In the following, we analyzed potential relations between the statistical properties of events from the 20 catalogs with reservoir characteristics and mechanical (e.g. mean stresses) and hydraulic parameters (e.g. flow rates).

Table 4.2: Results from fitting the magnitude frequency distribution to the Gutenberg-Richter Law.

No.	Event ID	M_L	Depth (km)	$M_{C,2000\,m}$ (merged)	$M_{C,2000\,m}$ (LBNL)	b value (all)	b value (2000 m)
1	3235921	3.6	2.2	0.9	1.2	1.07 ± .016	0.96 ± .044
2	3236636	2.5	2.3	0.9	1.1	1.23 ± .031	1.05 ± .095
3	3238276	3.0	2.9	0.6	1.1	1.01 ± .017	1.11 ± .039
4	3240183	3.2	2.3	0.6	1.1	1.03 ± .016	1.01 ± .027
5	3242613	3.1	2.4	0.8	1.1	1.11 ± .019	1.23 ± .049
6	3220536	2.9	2.8	0.6	0.9	1.04 ± .020	1.12 ± .057
7	3220687	2.9	2.7	0.6	0.8	1.00 ± .018	0.98 ± .034
8	3221009	2.6	2.1	0.8	0.8	1.07 ± .024	1.29 ± .125
9	3221113	3.6	3.4	0.5	0.8	0.93 ± .017	0.97 ± .032
10	3224143	3.2	3.2	0.5	0.3	1.01 ± .020	1.25 ± .073
11	3222668	3.1	1.3	0.5	1.1	0.88 ± .022	0.88 ± .039
12	3231652	3.0	1.2	0.5	0.9	0.74 ± .012	0.81 ± .038
13	3244320	3.1	1.7	0.7	1.2	1.00 ± .016	0.97 ± .033
14	3302049	3.5	1.5	0.7	1.2	1.21 ± .029	1.09 ± .079
15	3338163	3.5	1.9	0.7	1.1	1.12 ± .023	1.03 ± .080
16	3215344	2.7	4.0	0.6	1.1	1.13 ± .026	0.92 ± .117
17	3241751	3.6	4.2	0.9	1.2	1.11 ± .019	1.07 ± .082
18	3242914	3.5	4.0	0.9	0.9	1.02 ± .018	1.02 ± .033
19	3326148	4.4	4.2	0.7	1.1	0.98 ± .017	0.97 ± .027
20	3335925	2.8	4.5	1.0	0.9	1.28 ± .040	1.25 ± .156

Note. Errors given represent the standard deviation of the b value.

4.5.2 Evolution of b Value and Seismic Activity

The b Values and Seismic Activity for Different Mainshock Locations Within the Reservoir Clear differences in b values were found depending on the location of the sequence within the field (Figure 4.7a). At the outer NW boundary of The Geysers, lower b values (around 0.9) were observed, which increased to 1.1 – 1.2 within the NW half of the field. Toward the center of the field the b values reached a mean value around 1.0. In the SE part of the geothermal field, b values ranged from 1.0 to 1.2, with increasing values when approaching the most SE of the field. However, the b value uncertainties were also larger in this portion of The Geysers. While comparing b value from each sequence with the corresponding mainshock depth, an increase in b values from 0.85 to 1.1 was observed around 1.7 km, coinciding with the approximate upper boundary of the reservoir (Figure 4.7b). No scaling between b value of the sequence and the mainshock magnitude was seen (Figure 4.7c).

Interestingly, we also observed large spatial differences in total seismic activity, that is, the number of events, across the geothermal field. In the northwestern part of The Gey-

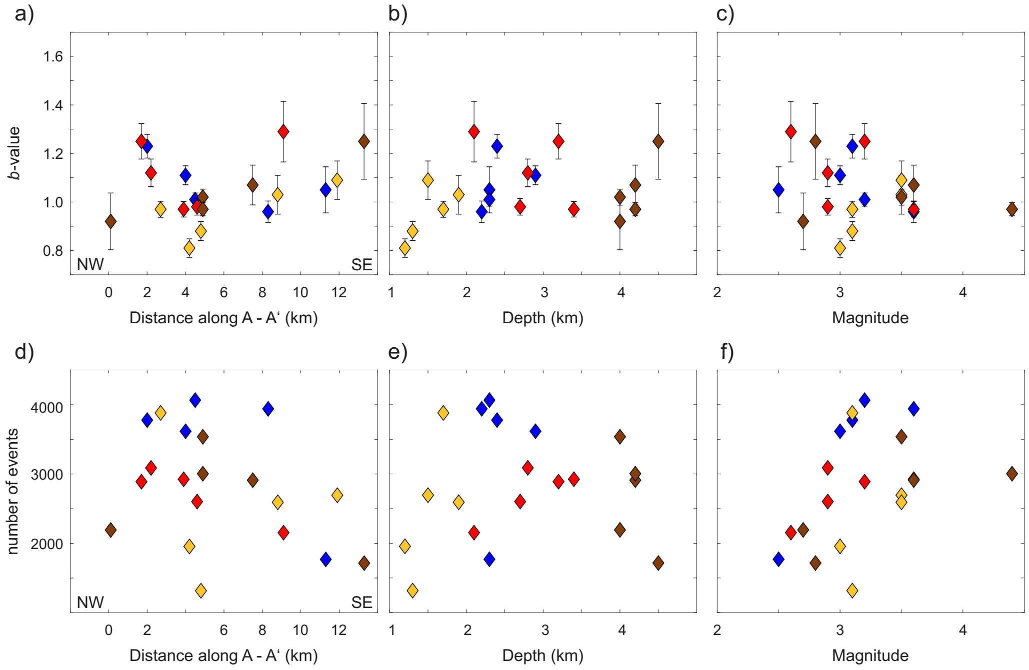

Figure 4.7: Comparison between the *b* value and (a) position across the geothermal field (see Figure 4.2), (b) depth, and (c) magnitude. The frequency magnitude distributions vary significantly across the geothermal field, but show lower *b* values above the reservoir. Relation of total seismic activity and (d) position across the geothermal field (see Figure 4.2), (e) depth, and (f) magnitude. For events located within or below the reservoir, increased seismic activity is observed in the NW Geysers. Yellow diamonds indicate shallow events between 0 and 2 km depth, blue and red show events at 2 – 4 km depth during high and low injection periods, respectively, and brown are events below the reservoir.

sers, the number of events was significantly higher than in the southeastern part (Figure 4.7d). In depth section the total seismic activity was the highest within the reservoir, with decreased activity above and below the reservoir (Figure 4.7e). The relationship between seismic activity and mainshock magnitude is ambiguous, however larger magnitude mainshocks tend to also have a higher number of events in their sequence (Figure 4.7f).

When comparing three horizontally collocated sequences with varying depth but in the same area of The Geysers, we observed a larger *b* value below the reservoir (Figure 4.8a). At shallow mainshock depth the *b* value was approximately 0.95, while below the reservoir it reached 1.15. The difference exceeded the standard deviation of the calculated *b* values (Figure 4.8a). At the same time, the total number of events increased with depth for these collocated sequences, from 1,457 at shallow depth to 3,004 at large mainshock depth. In summary, both *b* value and seismic activity of each sequence were found to be strongly dependent on the epicentral location of the mainshock event within the geothermal field, and slightly dependent on the mainshock depth. This suggests that seismicity patterns surrounding large magnitude events at The Geysers are more governed by their location within the field (i.e. the geological and tectonic setting or possibly density of injection wells and production history), than by their depth (i.e. mean stress and temperature).

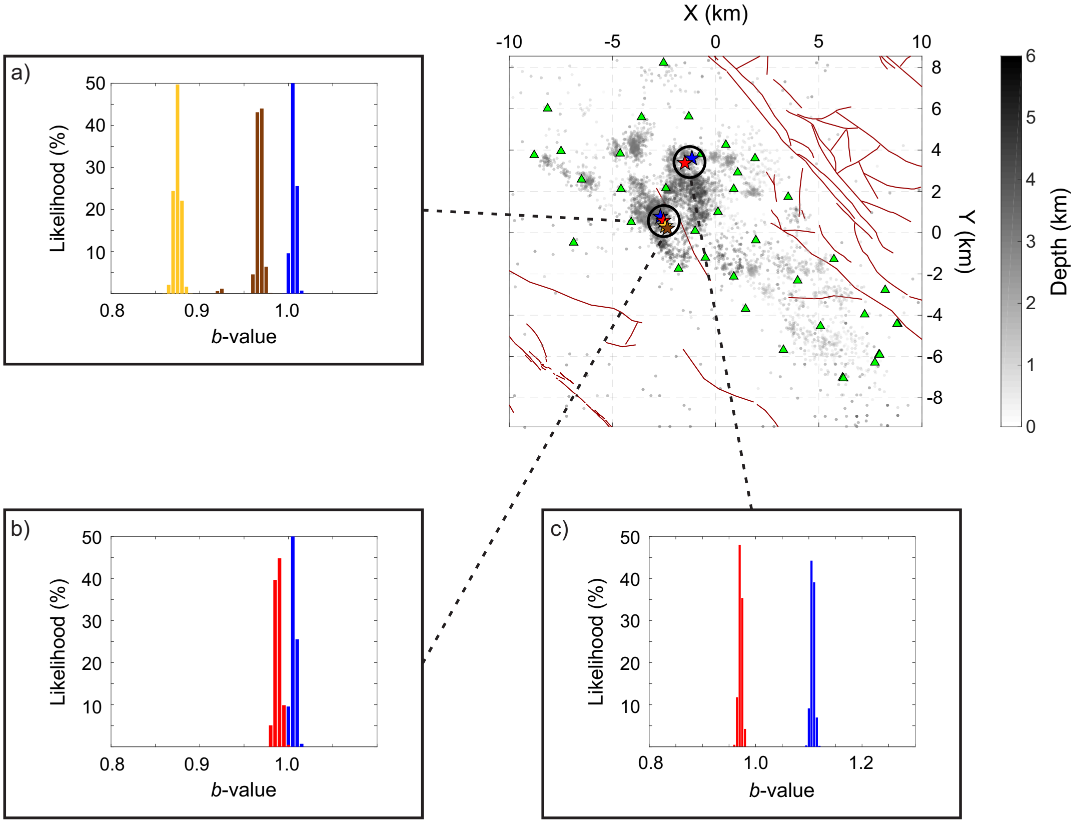

Figure 4.8: Comparison of *b* values: (a) between three collocated sequences with different depths, and between collocated sequences during high (b) and low (c) injection periods, respectively. In (a), yellow and brown color represent sequences above and below the reservoir, respectively. Blue and red indicate sequences during high and low injection periods, respectively. The location of the sequences on the field is given in the top right corner, as well as the seismicity encoded with hypocenter depths.

Changes of *b* Value and Seismic Activity Related to Injection Volume The amount of injected water into the reservoir is subjected to strong seasonal fluctuations, allowing us to investigate the influence of injection activity on the behavior of the studied seismic sequences. We identified two sequence pairs, where a large mainshock occurred during both high and low injection periods in close proximity to each other (Figures 4.8b and 4.8c). At times of high injection volume, an increase in *b* value could be observed, from 0.97 to 1.02 and from 0.96 to 1.1, respectively. In both cases, the difference in *b* value is considered significant from the Monte Carlo experiments (Figures 4.8b and 4.8c). The largest difference was observed for the sequence pair in the northern-most part of the field, compared to the smaller difference in the central part of the field.

Similarly, the seismic activity varies with the injection volume during the time of the sequence and between different sequence types. Considering the same collocated sequences as discussed above with the *b* value, we find that for a case located at the northern margin of the investigated area during times of high injection volume about 50 % more earthquakes occurred compared to the low injection period. In the central portion, this ratio was about 25 %. Considering the entire groups of high and low injection sequences, we found a general tendency to increased number of events in each sequence during times when larger fluid volumes were injected (Figure 4.7).

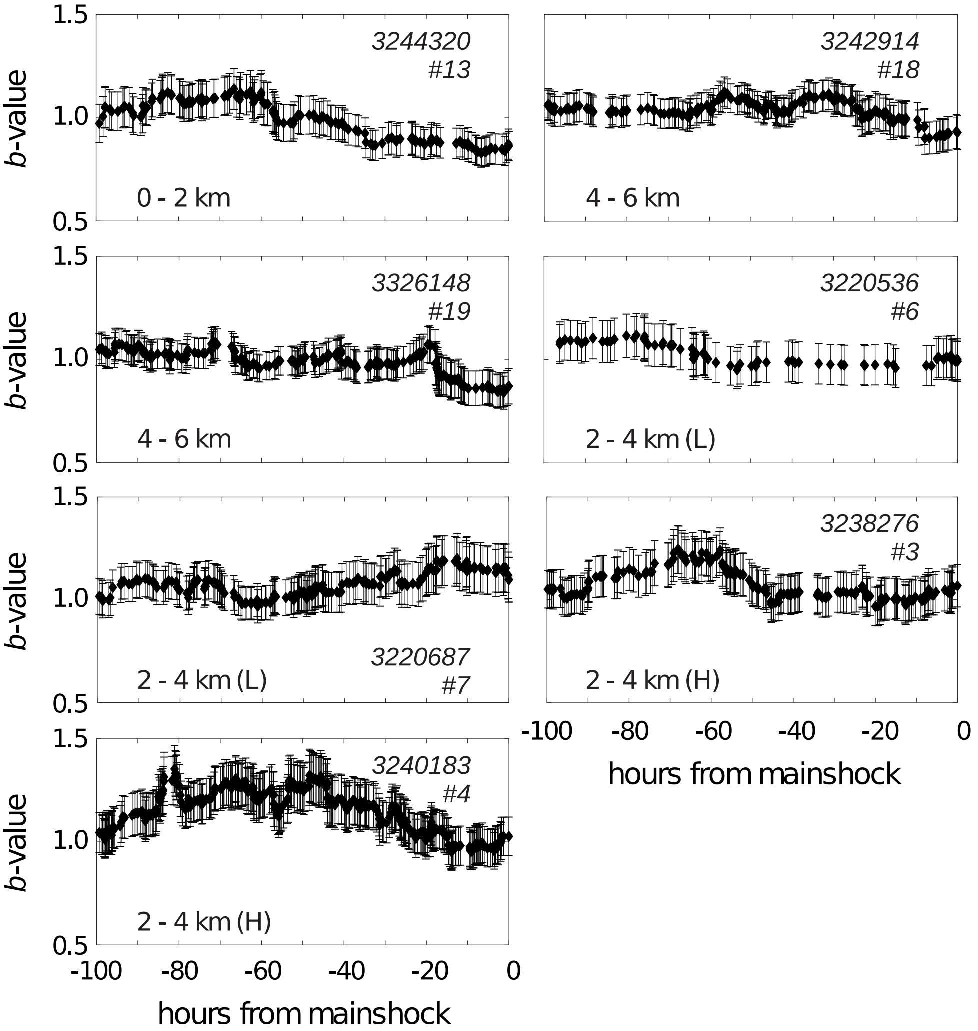

Figure 4.9: Temporal evolution of the b value leading up to the mainshock for sequences containing a sufficiently large number of earthquakes. The moving b value was calculated using sliding windows of 100 events. Error bars indicate one standard deviation. (H) and (L) in the caption signify high and low injection periods, respectively. The number in the upper right represents the mainshock ID (Table 4.1).

Temporal Evolution of b Value Within Each Sequence We additionally calculated the temporal evolution of b values dividing them into equally sized windows including 100 events. The magnitude of completeness was consistently recalculated for each bin of events. Varying window sizes lead only to smoother or noisier curves but did not affect the general trends. Only seven sequences contained a large enough number of events distributed adequately across time to compute reliable time series of b values, including one above the reservoir. For this latter sequence, a decrease of approximately 0.25 in b value starting 3 days before the mainshock could be observed, with the value stabilizing at a minimum centered on the time of the main event (Figure 4.9). Below the reservoir, b values showed a continuous decrease of about 0.2, starting about 5 days prior to the mainshock for each of the two calculated cases. During times of low injection, sequences within the reservoir exhibited a relatively stable b value over time, with a small increase of about 0.2, within 24 h before the main event. In contrast, the two analyzed high injection sequences displayed a similar trend as the shallow sequence, with a decrease of b value starting about three days prior to the mainshock (Figure 4.9).

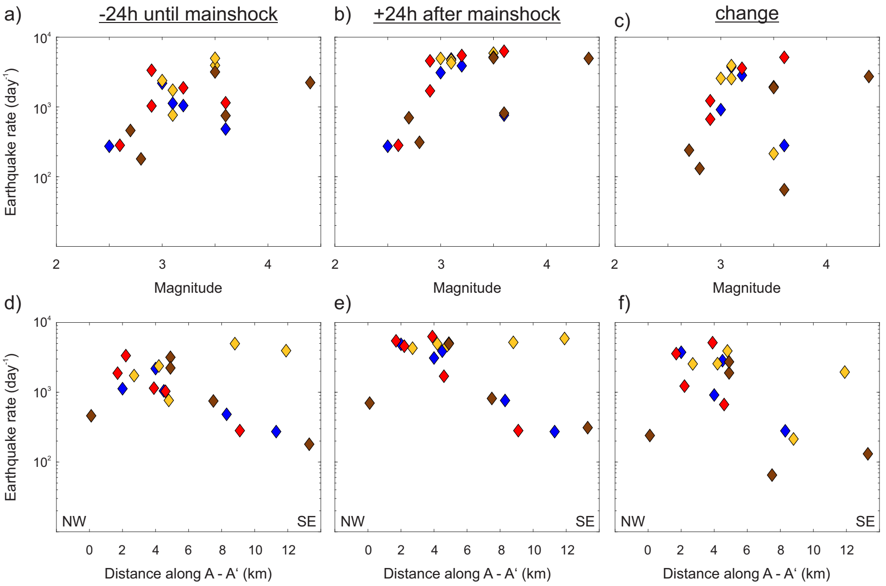

Figure 4.10: Average earthquake rates around the time of the mainshock for each sequence. Earthquake rate (a) 24 hours before and (b) 24 hours following the large magnitude event compared against mainshock magnitude, and (d) and (e) position across the geothermal field (see Figure 4.2). Yellow diamonds indicate shallow events between 0 – 2 km depth, blue and red show events at 2 – 4 km depth during high and low injection periods, respectively, and brown represents events below the reservoir. Independent of its magnitude, large magnitude earthquakes do not always trigger a significant earthquake rate change within the reservoir.

4.5.3 Seismicity Rates Surrounding the Mainshock Events

Temporal changes in the seismicity patterns were analyzed by means of the seismicity rate. Equal sized moving windows containing 10 events were used to calculate the seismicity rate by comparing the origin time of the first and last earthquake in each bin. Comparing the number of events around the mainshocks allowed separating two distinct types of behavior: 11 sequences displayed a significant increase of earthquake rate and seismic moment release after the time of the main event (Figure 4.10). However, we did not observe any systematic increase right before the mainshock. Almost all sequences within the reservoir during low injection period and above the reservoir were followed by some sort of seismicity rate increase. In contrast, a second group of events did not show any sign of increased or accelerated seismicity after the time of the main event, those were mostly events during high injection periods or located below the reservoir, i.e. times of high stress (Martínez-Garzón et al., 2013). The magnitude of the mainshock played an important role, since only events larger than $M_\mathrm{L}2.9$ showed a significant seismicity increase (Figure 4.10c). Additionally, the position of each sequence in the field had a visible influence on the rate change pattern. The majority of sequences showing a seismicity rate increase were located in the northwestern part of The Geysers, only a single sequence in the southeastern half displayed a significant rate increase around the time of the main rupture (Figure 4.10f). Importantly, changes to seismicity rate were observed only following the mainshock, but no sequence displayed an increase of seismicity leading up to the main event.

4.6 Discussion

The waveforms from the dense short-period network at The Geysers geothermal field constitute a unique dataset of high-quality continuous recordings in the research field of induced seismicity. Extension of the existing permanent short-period network at The Geysers with broadband instrumentation allows the study of earthquakes ranging more than 6 orders of magnitude, from microseismicity to moderately large magnitudes, in full detail. Previous studies on large earthquakes at The Geysers have primarily focused on the relationship between them and anthropogenic activity (e.g., Allis, 1982; A. Ross et al., 1999; Martínez-Garzón et al., 2017, 2018), its impact on the source physics of earthquakes (e.g., A. Ross et al., 1996; L. R. Johnson, 2014; Yu et al., 2018), as well as the physical mechanisms governing seismicity (e.g., Martínez-Garzón et al., 2014; Kwiatek et al., 2015). Due to the high resolution and optimal network coverage, seismic data from The Geysers can provide important insights about how reservoir conditions govern the initiation of ruptures.

In the following, we first discuss the resolution of our detection catalogs with respect to other studies. Then, we utilize our observations to discuss whether the large events at The Geysers do display a preparation process and if so, which conditions appear to govern it.

4.6.1 Resolution of the Employed Matched Filter Technique

Application of the matched filter technique to a particular dataset may result in a different decrease of the magnitude of completeness depending on the scope of each study. For example, Aiken et al. (2018) investigated potential dynamic triggering at The Geysers by applying a MFA to continuous vertical recordings half a day before and one day after 25 distant mainshocks. The authors were able to lower the magnitude of completeness by one full order of magnitude, down to $M_\mathrm{L} \sim 0$, by utilizing around 3,000 templates from the entire geothermal field. The goal in our study was not to increase the overall detection capabilities across the entire geothermal field, but to successfully create a detailed record of events occurring in temporal and spatial vicinity of 20 large magnitude mainshocks.

This allowed us to capture detailed seismicity patterns, as well as investigate potential preparation processes. Even though the overall magnitude of completeness improvement was only about 0.5 compared to the local catalog, almost 27,000 of our located events constitute previously undetected earthquakes related to the analyzed mainshocks. Therefore, direct quantitative comparison between an unbiased local catalog and our detections, which a priori assume a certain waveform, should be done with care.

Our observations highly depend on the quality of the obtained catalogs. The small detection threshold used ensured that we captured events even buried in noise, while at the same time our strict workflow and localization procedure ensured that no regional earthquake was falsely matched and included in our analysis. However, it may have occurred that events were not detected even if they were closely located to the source region of the master events, if their mechanism did not relate to any template mechanism. By selecting a large variety of templates with different waveforms and merging the catalog with the existing local catalog from Lawrence Berkeley National Laboratory this problem should be reduced, but cannot be eliminated entirely.

4.6.2 Sensitivity of Mainshock Sequence Statistics

Seismicity patterns surrounding moderate magnitude earthquakes at The Geysers are governed by a number of different factors. We found that the position of the mainshock within the field plays the largest role in the characteristics displayed by its seismic sequence. The northwestern part of The Geysers shows significantly increased seismic activity around the time of the mainshock, compared to the background level. This is reflected in an increase in the daily earthquake rate (Figure 4.10), as well as an increased total number of events (Figure 4.7). On the other hand, large magnitude mainshocks located in the southeastern part generally show no significant change of seismic activity accompanying the main event. Several differences exist between NW and SE of The Geysers, including reservoir depth and the depth of the felsite unit (Jeanne et al., 2014), seismic velocities (Gritto & Jarpe, 2014), coda-Q properties (Blanke et al., 2019), and seismic activity (C. W. Johnson et al., 2016), magnitude-frequency statistics of the events (Convertito et al., 2012) and stress field orientation (Martínez-Garzón et al., 2014). In addition, there are more wells injecting fluid in the NW part of the field and the seasonal fluctuations in the volume of water injected are stronger. These features suggest that reservoir geology and anthropogenic operations might play a dominant role controlling the seismicity associated with moderate magnitude earthquakes at The Geysers.

The local stress field at The Geysers varies laterally and with depth between normal faulting and strike-slip (Martínez-Garzón et al., 2013). However, we did not observe a direct relationship between mainshock depth and an accompanying earthquake rate change. The sequences which displayed enhanced seismic activity after the mainshock were located across the entire depth range, from 1.7 to 4.2 km, analyzed (Figures 4.7 and 4.10), but were restricted to the northwestern part of the field, where seismicity in general and the reservoir itself are located deeper than in the southeast. Additionally, the northwestern part of The Geysers field is characterized by much higher reservoir temperatures, possibly an indication for larger heat flow. Our observation of larger presence of aftershocks in the area of higher temperatures and fluid content is therefore, consistent with the findings at global and regional scale from Zaliapin and Ben-Zion (2013b, 2016).

Overall, we observe that the mainshocks appear to produce increased seismic activity only if the mainshock magnitude is $M_\mathrm{L} > 2.9$, independent of the magnitude of completeness and number of templates. This may indicate that a certain threshold of seismic moment release has to be overcome to transfer sufficient static stress as to trigger aftershocks, or to increase seismicity to a significant enough degree to be apparent about the already considerable background rate. Increasing number of triggered aftershocks with mainshock magnitude is commonly observed for different seismicity catalogs (Zaliapin & Ben-Zion, 2013a), and it was explained by a combination of the larger magnitude range available to detect aftershocks which are smaller than the mainshock and a constant magnitude of completeness. Considering the low magnitude of completeness, however, this cannot explain the complete lack of aftershocks for some of the sequences.

Martínez-Garzón et al. (2013, 2016) showed that the stress field orientation and geometry of reactivated faults additionally varies with time following changes in the fluid injection flow rates. Although, seismicity at The Geysers strongly correlate with injection rate and volume we did not see a strong dependence between the seismic productivity of a mainshock and the injected volume around its time of occurrence here (Figure 4.7). Similarly, Martínez-Garzón et al. (2018) using cluster analysis found that at The Geysers the number of aftershocks only partly correlates with the injection rates of the field, with the relation varying significantly with time and no continuous trend found.

Experimental results based on rock deformation laboratory experiments have shown that the b value from the frequency magnitude distribution of acoustic events depends inversely on the differential stress (Scholz, 1968, 2015). In agreement with these findings, Schorlemmer et al. (2005) found that b values in southern California were significantly different for the three different faulting regimes, with larger and lower b values observed for normal and reverse faulting, respectively.

According to poro-elasticity, differential stresses should change during time periods of larger fluid injection, particularly along the direction of the maximum horizontal stress (Schoenball et al., 2010), and consequently b values should change. However, we find no average difference between b values observed for sequences during high and low injection volume, respectively. Comparing collocated sequences during different seasons, and thus different injected volumes, we found that during high injection the b value is statistically higher than during times of low injection. These findings are consistent with earlier temporal analysis of b values at one isolated cluster of seismicity in the NW part of the field, where a positive correlation between injection rate and b value over the course of 7 years was observed (Leptokaropoulos et al., 2018). Injection induced increase in pore pressure and the associated reduction of effective stress could lead to weakening of the reservoir rock, resulting in failure before large stresses are able to accumulate. This is in agreement with our observed b value variation across the entire geothermal field, with large b values mostly constrained to the northwestern part, that is, the seismically most active part of The Geysers. Furthermore, the sudden increase of b value at reservoir depth could be explained by the ability to accommodate larger stresses in the unfractured rocks above the geothermal reservoir. Thus b values at The Geysers might be also influenced by the rock damage within the reservoir (Amitrano, 2003).

4.6.3 Do Preparation Processes Exist for Moderate-Magnitude Earthquakes at The Geysers?

If observed in nature, precursory patterns can be more than 3 orders of magnitude lower than the subsequent mainshock (Mignan, 2014), making them extremely difficult to detect and observe for moderate sized earthquakes. The dense instrumentation at The Geysers provides a low magnitude of completeness dataset, ideal for detecting and studying potential foreshock sequences. Here, standard based energy detectors provide a magnitude of completeness of about M_{L} 1.1, and application of our matched filter further reduced it by half a magnitude for selected sequences, with an even lower detection threshold. However, utilizing these high quality continuous recordings for 20 large sequences and statistically analyzing microseismicity four orders of magnitude lower than the respective mainshock, we could not observe any seismic precursory processes in the days preceding the main rupture. None of the sequences analyzed showed an increased spatial correlation of microseismic activity with the main event leading up to its rupture. However, most of the sequences showed localization between subsequent event pairs, but compared with larger groups of events no increased mean correlation could been observed. Decreases of b value leading up to the main failure were seen only for two sequences, thus cannot be generalized. However, similar systematic observations have been made in rock deformation experiments in the laboratory (Zang et al., 1998; Goebel et al., 2017), warranting more detailed analysis of this potential behavior at The Geysers in the future.

Only 11 sequences displayed any reaction in seismic activity to the large magnitude event, representing an increase of seismicity rate due to aftershock activity and not foreshocks. Similar to these results, Trugman et al. (2016) found anomalously low aftershock productivity when analyzing seismicity across The Geysers, suggesting that earthquakes are primarily induced by localized stress changes from injection, rather than coseismic stresses associated with other earthquakes occurring nearby. In accordance with that, 9 of the investigated sequences did not show a response in the seismicity rate to the occurrence of the large magnitude event. When examining the waveforms for these particular events, smaller events, occurring just seconds before the main failure, were identified on many of them (e.g. Figure 4.11). Manual picks locate those events within less than a hundred meters of the mainshock, indicating that a small area of initial failure propagated to a larger rupture. We observe these complex failures predominantly in the northeastern and southeastern part of The Geysers. This indicates that in those areas, which have the longest history of injection activity across the field, single small events may propagate into larger ruptures without showing a traditional earthquake sequence with fore-, main-, and aftershock. The occurrence of aftershocks was limited almost exclusively to the northwestern part of the geothermal field. Increased injection activity and high-temperatures have most likely led to a locally high stress environment, which is much more susceptible to aftershock triggering from stress transfer due to the shorter injection history. Assuming a less damaged formation in this part of the reservoir, the lack of precursory seismic signs of accumulating stress for large ruptures is surprising as the rocks should still be strong enough to display them. This might indicate, that stress buildup is not compensated for coseismically but may in fact be aseismic.

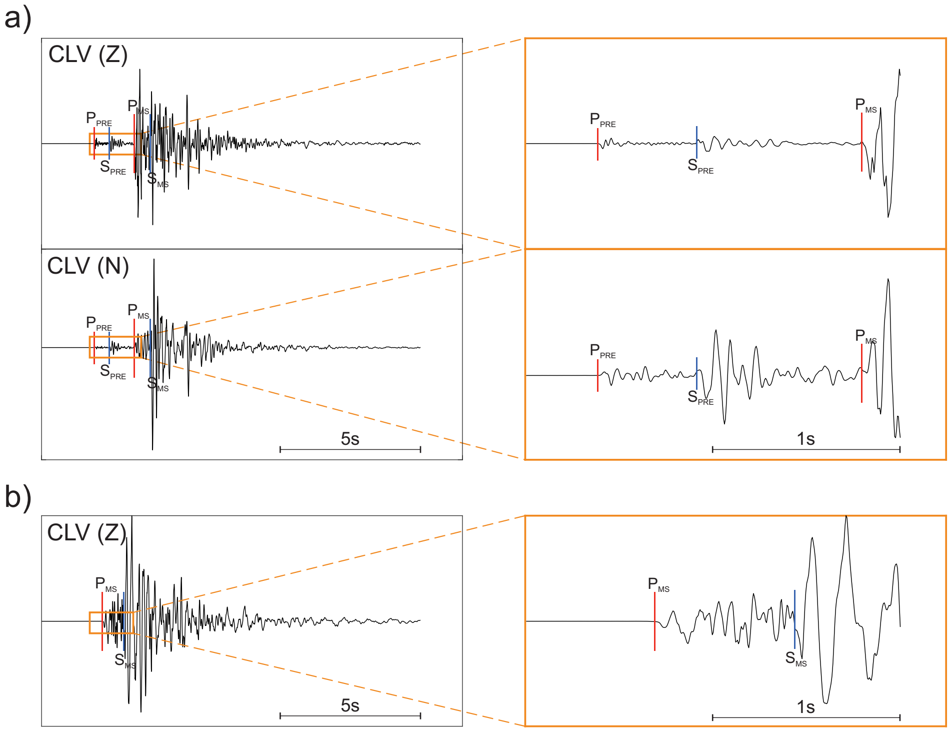

Figure 4.11: (a) Example of a preevent (Event ID: 3238276; # 3 in Table 4.1) at station CLV. (a) Top and bottom panel show the vertical and horizontal component, respectively. P and S phase arrivals are indicated for the preevent (PRE) and the mainshock (MS). (b) Example of a regular mainshock waveform (Event ID: 3240183; # 4 in Table 4.1).

4.7 Conclusion

We analyzed microseismicity framing the occurrence of 20 moderate to large magnitude earthquakes at The Geysers geothermal field, northern California. Detailed earthquake catalogs for the 11 days surrounding each mainshock were obtained using previously detected earthquakes at the nucleation spot as templates in a Matched Filter Algorithm. When considering events located within the rupture plane of each mainshock, the magnitude of completeness could be lowered to about $M_{\mathrm{C}} = 0.5$ in most cases. In total approximately 27,000 new unique earthquake locations were obtained.

Applying statistical tools and seismological parameters to the microseismic catalogs, we investigated the seismic behavior before as well as the immediate response of the reservoir following large magnitude earthquakes. Different parts of the field provide different feedback to the occurrence of large mainshocks. In the northwestern part of The Geysers, sequences show increased seismic activity following the main event, while in the southeastern portion of the field no deviation from the background seismicity rate could be observed. Similarly, larger b values for sequences in the northwest than southeast were seen. We interpret this behavior as variations in the local stress field and degree of damage in the reservoir formation. In the northwest high stressing rates result in higher background seismicity rates. Even though no clear relationship between mainshock depth and the sequence behavior was observed, the three shallowest sequences exhibit the smallest b values, most likely reflecting the lack of high damage zones just above the reservoir. Besides the fact of higher density in injection wells in the northwestern The Geysers, no clear temporal link between injection activity and the seismic productivity of each sequence or their b value could be found.

Using statistical analysis, no precursory patterns in the days preceding each large magnitude event were observed. Seismicity rates prior to main failure were relatively stable, temporal changes in b value leading up to the main event were seen in only two cases, representing an interesting but isolated observation. Events which did not trigger a change in seismicity rate in the following days show, however, distinct small events in the seconds before their onset, suggesting that in parts of The Geysers large magnitude events grow from smaller initial earthquakes.

The Geysers reservoir represents a very particular physical environment (e.g. high heat flow, long-lasting fluid injection and production cycles, and shallow reservoir depth). Thus our results may contribute a new perspective to the ongoing discussion of precursory seismic patterns potentially preceding large magnitude earthquakes, as well as illuminate specific reservoir responses to stress perturbations dependent on local tectonic and geological settings. Furthermore, we highlight the effectiveness of new earthquake detection and location methodologies on already existing high quality datasets and the improvement they provide. Additionally, the findings of changing seismic behavior across the geothermal field, may have influence on seismic hazard estimations, as they heavily depend on b value und seismicity rate.

5 The Two-Scale Preparation Phase of a M_W5.8 Earthquake in the Sea of Marmara Offshore Istanbul, Turkey

For copyright reasons, this chapter was removed from the electronic version. The corresponding article can be retrieved through the DOI link provided below.

Published as Durand, V., Bentz, S., Kwiatek, G., Dresen, G., Wollin, C., Heidbach, O., Martínez-Garzón, P., Cotton, F., Nurlu, M., Bohnhoff, M. (2020). "A two-scale preparation phase preceded a M_W5.8 Earthquake in the Sea of Marmara offshore Istanbul, Turkey." in *Seismological Research Letters*, doi: 10.1785/0220200110.

6 Synthesis and Outlook

Recently, advances in seismic instrumentation and data mining algorithms have changed the way we perceive extended nucleation phases of earthquakes. While a few years ago, they were still considered an exception to the rule, there now is an increasing number of earthquakes that show substantial precursory activity (e.g., Bouchon et al., 2013; Trugman & Ross, 2019). Prominent examples include the 1999 $M_\mathrm{W}7.4$ İzmit (Bouchon et al., 2011; Ellsworth & Bulut, 2018), the 2011 $M_\mathrm{W}9.1$ Tohoku (Kato et al., 2012; Mavrommatis et al., 2015), the 2014 $M_\mathrm{W}8.2$ Iquique (Schurr et al., 2014), and the 2015 $M_\mathrm{W}8.3$ Illapel earthquake (Huang & Meng, 2018). Recent studies have found substantial precursory signals even before moderate sized earthquakes, when near-fault and low-noise recordings were present (e.g., Malin et al., 2018). This is in agreement with observations from laboratory experiments on crustal rocks, where the main failure of the rock sample is frequently preceded by increased acoustic emissions and cross-correlation (Goebel et al., 2012). These insights have largely been possible thanks to novel techniques allowing to extract information from seismological recordings previously hidden in the noise (e.g., Z. E. Ross et al., 2019). Nevertheless, the prognostic value of these precursory signals is still heavily debated (e.g., Mignan, 2014). Understanding why some earthquakes display precursors and others don't is a fundamentally important question in modern seismology, with immense implications in terms of short-term operational earthquake forecasting and seismic risk assessment: If large earthquakes are the result of a mostly random triggering process of foreshocks (*cascade model*), we might not be able to anticipate them. On the other hand, if they are the product of an underlying measurable preparatory deformation process (*preslip model*), there is still a future potential to forecast large magnitude earthquakes (Ellsworth & Beroza, 1995).

The goal of this thesis was to investigate earthquake nucleation behavior across scales and in different settings, contributing to the overall challenge of understanding how earthquakes initiate, and how their nucleation behavior is influenced by tectonic or anthropogenic boundary conditions. While many previous studies have focused on high-resolution analysis of seismic patterns surrounding large earthquakes, detailed identification and analysis of preparation phases of moderate sized earthquakes is rarely done and faces significant observational challenges. Nevertheless, these small earthquake sequences are often the only source of information for regions where large earthquakes are less frequent, but expected in the future. In addition, it is especially moderate sized seismicity that typically poses substantial seismic hazard in anthropogenic fluid-injection operations, such as enhanced geothermal systems or wastewater injections. Thus, identifying the characteristic initiation phase of earthquakes in these tectonic and anthropogenic environments will ultimately help assess seismic hazard and potentially allow for the formulation of physics-based predictions on the nucleation behavior of hazardous events.

Hydraulic stimulations of reservoirs in the frame of enhanced geothermal systems (EGS) are of special interest when considering the occurrence and nucleation of moderate sized earthquakes. In the past, several projects have been terminated due to the occurrence of seismic events and the subsequent acceptance issues, such as the Basel deep heat mining project (Giardini, 2009). Consequently, EGS projects now rely on close interaction between industry, scientific community, and local stakeholders. In addition, detailed high resolution hydraulic and seismic data acquired during stimulations through dense downhole monitoring are nowadays required by law. Chapter 3 focused on the analysis of a collection of the most prominent hydraulic stimulations performed to date, using publicly available datasets. Contrary to chapters 4 and 5, the investigation was restricted to the statistical analysis of the observed seismicity, without extracting further information from seismic recordings. This was motivated by the need of an "easy to use" indicator proxy during stimulation to monitor rupture growth and to evaluate the potential for a large magnitude event triggered or induced by fluid injection. Unlike most current monitoring systems, the analysis showed that tracking the cumulative seismic moment in relation to injection parameters rather than the maximum observed magnitude provides a better insight into site-specific seismicity evolution and first-order seismic hazard assessment. The majority of the analyzed stimulations, 8 out of 10, showed a stable evolution of cumulative seismic moment, indicative for a stable growth of a self-arrested rupture (Galis et al., 2017). Seismicity was apparently governed by the induced pore-pressure perturbation, where an increase in injected volume or applied hydraulic energy led to an increase in rupture area. In agreement with McGarr (2014), the slope characterizing this growth of cumulative seismic moment with hydraulic energy is close to unity, suggesting a linear relation between volumetric strain in the reservoir and cumulative seismic moment. The key observations here, however, was that most stimulations exhibited a short stage of steeply increasing seismic moment, interpreted as an initial adaptation phase of fracture opening and fault activation around the well. A second group of projects, including the Pohang (South Korea) and Cooper Basin (Australia) stimulations, showed a steep increase of cumulative seismic moment and maximum observed magnitude. Pohang displayed a seemingly unbound growth of seismic moments, whereas during the Cooper Basin stimulation time periods of stable growth were punctuated by substantially larger events. These evolutions are not captured by any existing models for moment growth during fluid injection, and suggest activation of runaway ruptures, where rupture size is only limited by the size of the tectonic faults. During the Basel stimulation (Switzerland) the moment evolution remained extremely stable, yet a short time after shut-in the $M_\mathrm{W}3.0$ event ocurred which caused the project to stop immediately. In this case, no further control could have been imposed by any changes in injection protocols during the stimulation phase.

The obtained results in this study have far reaching implications for our basic understanding of earthquake nucleation and for the design of next-generation traffic-light systems aiming at controlling the seismic hazard imposed by hydraulic stimulations. First, the fast transition from a stable to an unstable rupture propagation in Basel highlights the problem that the exceedance of maximum arrestable rupture size may occur at times of otherwise stable seismic response. Additionally, a significant time delay may exist between seismic response and injection due to the characteristics of pore pressure diffusion. On the other hand, examination of the seismic moment evolution with hydraulic energy or volume allows us to identify cases, where rupture growth is unstable from the very beginning of injection, such as the case of Pohang. To be able to achieve this, a near realtime monitoring is necessary. Identifying these runaway ruptures early enough can significantly reduce the seismic hazard associated with hydraulic stimulations. Furthermore, the tectonic stress regime was found to have almost no influence on the general properties and evolution of seismicity. The b values between different projects where stimulations were performed in different stress regimes did not show substantial variations, suggesting that regional stress levels may not prominently affect the evolution of seismicity and the nucleation of large earthquakes during stimulation. It seems more likely, that local tectonic and structural features of the reservoir, such as fault length and orientation, as well as frictional properties dominate the seismic response.

In chapters 4 and 5 a waveform data-driven approach was chosen to investigate nucleation behavior of moderate sized earthquakes. In order to overcome limitations imposed by traditional waveform processing, a computational efficient matched filter algorithm and a scalable post-processing workflow was refined to two different datasets. Chapter 4 examined microseismicity patterns for the 11 days surrounding 20 $M_{\mathrm{L}} > 2.5$ earthquakes at The Geysers geothermal field in California. There, the goal was to identify 1) If extended preparation processes could be systematically observed; 2) Which factors influence the earthquake nucleation behavior; and 3) How the reservoir responds to the occurrence of larger events. In the frame of this study, a total of more than 27,000 previously undetected earthquakes were identified and located using the methodology described in chapter 2. The immediate response of the reservoir following the occurrence of large ($M_{\mathrm{L}} > 2.5$) mainshocks varied substantially within the geothermal field. In the northwestern part of The Geysers, large events were followed with an increased seismic activity, while in the southeastern portion of the field seismicity remained at the background level. Similar differences could be observed in the magnitude-frequency distribution, where generally larger b values for sequences in the northwest than for those in the southeast were observed. These different behaviors are most likely caused by variations in the local stress field and degree of damage in the reservoir rock, where higher stressing rates in the northwest lead to higher background seismicity rates. Small b values at shallow depth indicate the lack of high-damage zones just above the reservoir. However, using extensive statistical analysis, no pervasive precursory patterns could be observed in the days preceding each large-magnitude event. Seismic activity before the main shocks was relatively stable, and temporal changes in b value leading to the main failure were seen in only two cases. Events which were not followed by a change in seismicity rate showed distinct smaller events in the seconds before their onset. These event-doublets were mostly restricted to the northeastern and southeastern part of the reservoir, which corresponds to the areas with the longest history of injection activity, indicating that single small events in this highly damaged environment may propagate or trigger larger ruptures without showing a traditional earthquake sequence. In accordance with that, aftershock activity was limited to the northwestern part of The Geysers, where local high stress environment are more susceptible to aftershock triggering from coseismic stress transfer.

The Geysers has unique boundary conditions, thus generalization of the observations is difficult. Nevertheless, our observation of a heterogeneous behavior across the reservoir, which corresponds to different production cycles, may still contribute to a better understanding of earthquake initiation in similar settings. The degree of reservoir damage and local stress heterogeneities seems to govern nucleation behavior at The Geysers, where in certain areas large events seem to be triggered or grow out of stress perturbations induced by small earthquakes, in agreement with the cascading model of earthquake nucleation. Furthermore, increased seismic activity prior to the mainshocks is not observed for any part of the field, suggesting that already small stress perturbations rather than extended cascading chains are enough to trigger larger events. Additionally, the lack of transient repeater earthquakes makes a substantial contribution of aseismic slip to the triggering unlikely. On the other hand, aseismic contributions cannot be entirely excluded, as similarly to the coseismic stress transfer at The Geysers, small amounts of stress transfer may be enough to trigger large events without displaying the characteristic appearance of repeated ruptures of the same asperities.

Analysis of a moderate sized earthquake sequence during September 2019 in the western Sea of Marmara (Turkey), presented in chapter 5, showed a very different behavior. The Sequence consisted of two large events, $M_{\mathrm{W}}4.7$ and $M_{\mathrm{W}}5.8$, each preceded by substantial foreshock activity. Using the matched filter technique more than 2,100 earthquakes were detected and relatively located, compared to 577 in the regional catalog. The improvement was especially pronounced in the days before the first large event, where using cross-correlation some 40 foreshocks were found against the 4 present in the official catalog. Utilizing this high resolution catalog, a clear acceleration and migration of seismicity towards the epicenters of the mainshocks a few hours before their occurrence was observed. Furthermore, already several days before the $M_{\mathrm{W}}4.7$ event, the entire future $M_{\mathrm{W}}5.8$ rupture plane was illuminated by foreshocks. Combining the temporal and spatial evolution with the assumed surface trace of the fault branch suggests that the $M_{\mathrm{W}}4.7$ rupture was haltered by a local barrier imposed by the fault geometry. Propagation of afterslip triggered by the $M_{\mathrm{W}}4.7$ event may have contributed to weakening the fault enough for the $M_{\mathrm{W}}5.8$ to surpass that barrier and rupture the entire fault (Kato et al., 2012, 2016). Seismic migration and large-scale activation of the fault plane before main failure indicate a significant contribution of aseismic slip in loading the fault area (Peng & Zhao, 2009). Nevertheless, the migration of seismicity towards the source region could also be accounted for within the framework of the cascading source model (Ellsworth & Bulut, 2018). It seems that for the nucleation of the September 2019 sequence, a combination of both aseismic slip and cascading triggering due to stress transfer from smaller events substantially contributed to fault failure, and that neither model is to be categorically excluded.

Observations from The Geysers and the western Sea of Marmara earthquake sequence suggest that both cascade triggering and aseismic slip can play major roles in the nucleation of moderate sized earthquakes. Surprisingly, both mechanisms seem to jointly contribute to fault rupture initiation even within the same seismogenic volume. At The Geysers a separation of the two mechanisms can be inferred, where cascade triggering seems to dominate in highly damaged parts of the reservoir, indicating that the anthropogenic activity can at least partly influence the nucleation behavior of earthquakes. In the Sea of Marmara sequence, the observed seismicity patterns suggest an interplay between aseismic slip either prior to the sequence or induced by the initial $M_\mathrm{W}4.7$ (which of the two is really the case cannot be inferred from the available data as no aseismic slip was directly observed) and a migration of cascading foreshocks “preparing” the entire rupture plane for the $M_\mathrm{W}5.8$ mainshock. Thus, concerning the discussion between different earthquake nucleation models, the results in this thesis suggests that regardless of the setting there is a wide range of possible combinations between the two endmember models (*cascade* versus *preslip*). However, the observations also highlight the need for near fault observations and extensive data mining and post-processing to make these small scale changes in seismicity patterns visible.

Precise identification of the different mechanisms related to the nucleation of earthquakes requires an extensive examination of the spatiotemporal evolution of the surrounding seismicity. So far, studies about pervasive foreshock activity which have examined a large number of sequences, have rather focused on statistical relations between the events (Trugman & Ross, 2019). In the future, sequences exhibiting potential extended nucleation phases should be studied in more details, ideally following a consistent workflow (e.g. joint analysis of spatiotemporal microseismicity patterns, earthquake statistics, and GPS/strain data). As our understanding of earthquake initiation is largely data-driven, expanding our ability to detect small scale precursors is one crucial step towards operational short-term earthquake forecasting. Precisely identifying and locating precursory activity using data mining techniques could be supplemented with machine learning algorithms (Bergen et al., 2019), which have previously been proven to effectively predict the time of main failure in rock deformation experiments (Rouet-Leduc et al., 2017). At the moment, understanding earthquake nucleation boils down to analyzing the spatiotemporal evolution of fore- and aftershocks, to distinguishing cascade triggering from slow-slip induced mechanisms, or a combination of the two. Both mechanisms have been extensively observed in nature, including in this thesis, yet we do not know which factors favor either mechanism or none of them. One way to tackle the issue would be to analyse extended nucleation phases depending on fault and frictional properties. Very detailed seismic observation exist surrounding the San Andreas Fault Zone or the North Anatolian Fault zone, where we know from previous studies that segments of each fault can be quite different in regards to geological age and frictional behavior (e.g., Bilham & King, 1989; Sieh & Williams, 1990; Lisowski et al., 1991; Barka, 1992, 1996; Stein et al., 1997). Utilizing these near-fault observatories a direct relationship between nucleation behavior and fault properties could be investigated. Finally, a detailed spatiotemporal evolution of induced seismicity in EGS projects and laboratory experiments can further connect the anthropogenic influence of nucleation behavior with frictional rock properties.

References

Ader, T., Chendorain, M., Free, M., Saarno, T., Heikkinen, P., Malin, P. E., ... Vuorinen, T. (2019). Design and implementation of a traffic light system for deep geothermal well stimulation in finland [Journal Article]. *Journal of Seismology.* doi: 10.1007/s10950-019-09853-y

AFAD. (2019). National seismic network of turkey (tu) [Journal Article]. *Disaster and Emergency Management Presidency Turkey (AFAD).*

Aiken, C., Meng, X., & Hardebeck, J. (2018). Testing for the 'predictability' of dynamically triggered earthquakes in the geysers geothermal field [Journal Article]. *Earth and Planetary Science Letters*, *486*, 129-140. doi: 10.1016/j.epsl.2018.01.015

Albaric, J., Oye, V., Langet, N., Hasting, M., Lecomte, I., Iranpour, K., ... Reid, P. (2014). Monitoring of induced seismicity during the first geothermal reservoir stimulation at paralana, australia [Journal Article]. *Geothermics*, *52*, 120-131. doi: 10.1016/j.geothermics.2013.10.013

Allis, R. G. (1982). Mechanism of induced seismicity at the geysers geothermal reservoir, california [Journal Article]. *Geophysical Research Letters*, *9*(6), 629-632.

Amitrano, D. (2003). Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with thebvalue [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *108*(B1). doi: 10.1029/2001jb000680

Armijo, R., Pondard, N., Meyer, B., Uçarkus, G., de Lépinay, B. M., Malavieille, J., ... Sarikavak, K. (2005). Submarine fault scarps in the sea of marmara pull-apart (north anatolian fault): Implications for seismic hazard in istanbul [Journal Article]. *Geochemistry, Geophysics, Geosystems*, *6*(6). doi: 10.1029/2004gc000896

Asanuma, H., Soma, N., Kaieda, H., Kumano, Y., Izumi, T., Tezuka, K., ... Wyborn, D. (2005). Microseismic monitoring of hydraulic stimulation at the australian hdr project in cooper basin [Journal Article]. *Proceedings World Geothermal Congress 2005 Antalya, Turkey.*

Baisch, S., Bohnhoff, M., Ceranna, L., Tu, Y., & Harjes, H.-P. (2002). Probing the crust to 9-km depth: Fluid-injection experiments and induced seismicity at the ktb superdeep drilling hole, germany [Journal Article]. *Bulletin of the Seismological Society of America*, *92*(6), 2369-2380.

Baisch, S., Voros, R., Weidler, R., & Wyborn, D. (2009). Investigation of fault mechanisms during geothermal reservoir stimulation experiments in the cooper basin, australia [Journal Article]. *Bulletin of the Seismological Society of America*, *99*(1), 148-158. doi: 10.1785/0120080055

Baisch, S., Weidler, R., Voros, R., Wyborn, D., & de Graaf, L. (2006). Induced seismicity during the stimulation of a geothermal hfr reservoir in the cooper basin, australia [Journal Article]. *Bulletin of the Seismological Society of America*, *96*(6), 2242-2256. doi: 10.1785/0120050255

Barka, A. (1992). The north anatolian fault zone [Journal Article]. *Annales Tectonics*, *6*(Special Issue).

Barka, A. (1996). Slip distribution along the north anatolian fault associated with the large earthquakes of the period 1939 to 1967 [Journal Article]. *Bulletin of the Seismological Society of America*, *36*(5), 1238-1254.

Barka, A., Akyüz, S., Altunel, E., Sunal, G., Cakir, Z., Dikbas, A., … Page, W. (2002). The surface rupture and slip distribution of the 17 august 1999 izmit earthquake (m 7.4), north anatolian fault [Journal Article]. *Bulletin of the Seismological Society of America*, *92*(1), 43-60.

Barrett, S. A., & Beroza, G. C. (2014). An empirical approach to subspace detection [Journal Article]. *Seismological Research Letters*, *85*(3), 594-600. doi: 10.1785/0220130152

Batsi, E., Lomax, A., Tary, J., Klingelhoefer, F., Riboulot, V., Murphy, S., … Géli, L. (2018). An alternative view of the microseismicity along the western main marmara fault [Journal Article]. *Bulletin of the Seismological Society of America*, *108*(5A), 2650-2674. doi: 10.1785/0120170258

Beaucé, E., Frank, W. B., & Romanenko, A. (2017). Fast matched filter (fmf): An efficient seismic matched-filter search for both cpu and gpu architectures [Journal Article]. *Seismological Research Letters*, *89*(1), 165-172. doi: 10.1785/0220170181

Bender, B. (1983). Maximum likelihood estimation of b values for magnitude grouped data [Journal Article]. *Bulletin of the Seismological Society of America*, *73*(3), 831-851.

Bentz, S., Kwiatek, G., Martínez-Garzón, P., Bohnhoff, M., & Dresen, G. (2020). Seismic moment evolution during hydraulic stimulations [Journal Article]. *Geophysical Research Letters*. doi: 10.1029/2019gl086185

Bentz, S., Martinez-Garzon, P., Kwiatek, G., Dresen, G., & Bohnhoff, M. (2019). Analysis of microseismicity framing ml > 2.5 earthquakes at the geysers geothermal field, california [Journal Article]. *Journal of Geophysical Research: Solid Earth*. doi: 10.1029/2019JB017716

Ben-Zion, Y., & Lyakhovsky, V. (2002). Accelerated seismic release and related aspects of seismicity patterns on earthquake faults [Journal Article]. *Pure and Applied Geophysics*, *159*, 2385-2412.

Ben-Zion, Y., Vernon, F. L., Ozakin, Y., Zigone, D., Ross, Z. E., Meng, H., … Barklage, M. (2015). Basic data features and results from a spatially dense seismic array on the san jacinto fault zone [Journal Article]. *Geophysical Journal International*, *202*(1), 370-380. doi: 10.1093/gji/ggv142

Bergen, K. J., Johnson, P. A., de Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid earth geoscience [Journal Article]. *Science*, *363*(6433). Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/30898903` doi: 10.1126/science.aau0323

Bilham, R., & King, G. (1989). The morphology of strike-slip faults: Examples from the san andreas fault, california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *94*(B8), 10204-10216. doi: 10.1029/JB094iB08p10204

Blanke, A., Kwiatek, G., Martínez-Garzón, P., & Bohnhoff, M. (2019). Sensitivity and stability analysis of coda quality factors at the geysers geothermal field, california [Journal Article]. *Bulletin of the Seismological Society of America*, *109*(3), 959-975. doi: 10.1785/0120180219

Bohnhoff, M. (2017). Gonaf – the borehole geophysical observatory at the north anatolian fault in the eastern sea of marmara [Journal Article]. *Scientific Drilling*, *5*, 1-10. doi: 10.5194/sd-5-1-2017

Bohnhoff, M., Bulut, F., Dresen, G., Malin, P. E., Eken, T., & Aktar, M. (2013). An earthquake gap south of istanbul [Journal Article]. *Nat Commun*, *4*, 1999. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/23778720` doi: 10.1038/ncomms2999

Bohnhoff, M., Martínez-Garzón, P., Bulut, F., Stierle, E., & Ben-Zion, Y. (2016). Maximum earthquake magnitudes along different sections of the north anatolian fault zone [Journal Article]. *Tectonophysics*, *674*, 147-165. doi: 10.1016/j.tecto.2016.02.028

Bohnhoff, M., Wollin, C., Domigall, D., Küperkoch, L., Martínez-Garzón, P., Kwiatek, G., … Malin, P. E. (2017). Repeating marmara sea earthquakes: indication for fault creep [Journal Article]. *Geophysical Journal International*, *210*(1), 332-339. doi: 10.1093/gji/ggx169

Bommer, J. J., Oates, S., Cepeda, J. M., Lindholm, C., Bird, J., Torres, R., … Rivas, J. (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal project [Journal Article]. *Engineering Geology*, *83*(4), 287-306. doi: 10.1016/j.enggeo.2005.11.002

Bouchon, M., Durand, V., Marsan, D., Karabulut, H., & Schmittbuhl, J. (2013). The long precursory phase of most large interplate earthquakes [Journal Article]. *Nature Geoscience*, *6*(4), 299-302. doi: 10.1038/ngeo1770

Bouchon, M., Karabulut, H., Aktar, M., Özalaybey, S., Schmittbuhl, J., & Bouin, M. P. (2011). Extended nucleation of the 1999 mw 7.6 izmit earthquake [Journal Article]. *Science*, *331*(877-880). doi: 10.1126/science.1197341

Bourouis, S., & Bernard, P. (2007). Evidence for coupled seismic and aseismic fault slip during water injection in the geothermal site of soultz (france), and implications for seismogenic transients [Journal Article]. *Geophysical Journal International*, *169*(2), 723-732. doi: 10.1111/j.1365-246X.2006.03325.x

Brodsky, E. E., & Lay, T. (2014). Recognizing foreshocks from the 1 april 2014 chile earthquake [Journal Article]. *Science*, *344*, 700-702.

Brown, J. R., Beroza, G. C., & Shelly, D. R. (2008). An autocorrelation method to detect low frequency earthquakes within tremor [Journal Article]. *Geophysical Research Letters*, *35*(16). doi: 10.1029/2008gl034560

Bécel, A., Laigle, M., de Voogd, B., Hirn, A., Taymaz, T., Yolsal-Cevikbilen, S., & Shimamura, H. (2010). North marmara trough architecture of basin infill, basement and faults, from psdm reflection and obs refraction seismics [Journal Article]. *Tectonophysics*, *490*(1-2), 1-14. doi: 10.1016/j.tecto.2010.04.004

Charléty, J., Cuenot, N., Dorbath, L., Dorbath, C., Haessler, H., & Frogneux, M. (2007). Large earthquakes during hydraulic stimulations at the geothermal site of soultz-sous-forêts [Journal Article]. *International Journal of Rock Mechanics and Mining Sciences*, *44*(8), 1091-1105. doi: 10.1016/j.ijrmms.2007.06.003

Chen, K. H., Nadeau, R. M., & Rau, R.-J. (2007). Towards a universal rule on the recurrence interval scaling of repeating earthquakes? [Journal Article]. *Geophysical Research Letters*, *34*(16). doi: 10.1029/2007gl030554

Cocco, M., Tinti, E., & Cirella, A. (2016). On the scale dependence of earthquake stress drop [Journal Article]. *Journal of Seismology*, *20*(4), 1151-1170. doi: 10.1007/s10950-016-9594-4

Convertito, V., Maercklin, N., Sharma, N., & Zollo, A. (2012). From induced seismicity to direct time-dependent seismic hazard [Journal Article]. *Bulletin of the Seismological Society of America*, *102*(6), 2563-2573. doi: 10.1785/0120120036

Cuenot, N., Charléty, J., Dorbath, L., & Haessler, H. (2006). Faulting mechanisms and stress regime at the european hdr site of soultz-sous-forêts, france [Journal Article]. *Geothermics*, *35*(5-6), 561-575. doi: 10.1016/j.geothermics.2006.11.007

Damelin, S., & Miller, W. (2011). *The mathematics of signal processing* [Book]. Cambridge University Press.

[Dataset]. (2019). Retrieved from `http://doi.org/10.5880/GFZ.4.2.2019.002` doi: 10.5880/GFZ.4.2.2019.002

Denllinger, M., Majer, E., Freeman, K., Johnson, L. R., Jarpe, S., Nihei, K., … Walters, M. (2017). Final report : Monitoring the effect of injection of fluids from the lake county pipeline on seismicity at the geysers, california, geothermal field [Journal Article]. *United States*. doi: 10.2172/1410331

Diehl, T., Kraft, T., Kissling, E., & Wiemer, S. (2017). The induced earthquake sequence related to the st. gallen deep geothermal project (switzerland): Fault reactivation and fluid interactions imaged by microseismicity [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *122*(9), 7272-7290. doi: 10.1002/2017jb014473

Dieterich, J. H. (1978). Preseismic fault slip and earthquake prediction [Journal Article]. *Journal of Geophysical Research*, *83*(B8), 3940. doi: 10.1029/JB083iB08p03940

Durand, V., Bentz, S., Kwiatek, G., Dresen, G., Wollin, C., Heidbach, O., … Bohnhoff, M. (2020). A two-scale preparation phase preceded a mw5.8 earthquake in the sea of marmara offshore istanbul, turkey [Journal Article]. *Seismological Research Letters*.

Durand, V., Bouchon, M., Floyd, M. A., Theodulidis, N., Marsan, D., Karabulut, H., & Schmittbuhl, J. (2014). Observation of the spread of slow deformation in greece following the breakup of the slab [Journal Article]. *Geophysical Research Letters*, *41*(20), 7129-7134. doi: 10.1002/2014gl061408

Eberhart-Phillips, D., & Oppenheimer, D. H. (1984). Induced seismicity in the geysers geothermal area, california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *89*(B2), 1191-1207. doi: 10.1029/JB089iB02p01191

Ellsworth, W. L. (2013). Injection-induced earthquakes [Journal Article]. *Science*, *341*(6142), 1225942. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/23846903` doi: 10.1126/science.1225942

Ellsworth, W. L., & Beroza, G. C. (1995). Seismic evidence for an earthquake nucleation phase [Journal Article]. *Science*, *268*(5212), 851-855. doi: 10.1126/science.268.5212.851

Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 izmit earthquake by a triggered cascade of foreshocks [Journal Article]. *Nature Geoscience*, *11*(7), 531-535. doi: 10.1038/s41561-018-0145-1

Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., & Shimamoto, T. (2019). Triggering of the pohang, korea, earthquake (mw 5.5) by enhanced geothermal system stimulation [Journal Article]. *Seismological Research Letters*. doi: 10.1785/0220190102

Ellsworth, W. L., Lindh, A. G., Prescott, W. H., & Herd, D. G. (1981). The 1906 san francisco earthquake and the seismic cycle [Journal Article]. *Eearthquake Prediction: An International Review.* doi: 10.1029/ME004p0126

EMC., E. . G. (2017). Episode: 1993 stimulation soultz-sous-forêts [collection] [Journal Article]. *EOST - CDGP*. doi: 10.25577/SSFS1993

EMC., E. . G. (2018a). Episode: 2000 stimulation soultz-sous-forêts [collection] [Journal Article]. *EOST - CDGP*. doi: 10.25577/SSFS2000

EMC., E. . G. (2018b). Episode: 2003 stimulation soultz-sous-forêts [collection] [Journal Article]. *EOST - CDGP*. doi: 10.25577/SSFS2003

Eneva, M., & Ben-Zion, Y. (1997). Application of pattern recognition techniques to earthquake catalogs generated by model of segmented fault systems in three-dimensional elastic solids [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *102*(B11), 24513-24528. doi: 10.1029/97jb01857

Ergintav, S., Reilinger, R. E., Çakmak, R., Floyd, M., Cakir, Z., Doğan, U., … Özener, H. (2014). Istanbul's earthquake hot spots: Geodetic constraints on strain accumulation along faults in the marmara seismic gap [Journal Article]. *Geophysical Research Letters*, *41*(16), 5783-5788. doi: 10.1002/2014gl060985

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters [Journal Article]. *KDD-96 Proceedings*, 226-231.

Evans, K. F., Zappone, A., Kraft, T., Deichmann, N., & Moia, F. (2012). A survey of the induced seismic responses to fluid injection in geothermal and co2 reservoirs in europe [Journal Article]. *Geothermics*, *41*, 30-54. doi: 10.1016/j.geothermics.2011.08.002

Facca, G., & Tonani, F. (1965). The self-sealing geothermal field [Journal Article]. *IAV International Symposium on Volcanology, New Zealand.*

Fisher, R. A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample [Journal Article]. *Metron*, *1*(3-32), 205-235.

Font, Y., Kao, H., Lallemand, S., Liu, C.-S., & Chiao, L.-Y. (2004). Hypocentre determination offshore of eastern taiwan using the maximum intersection method [Journal Article]. *Geophysical Journal International*, *158*(2), 655-675. doi: 10.1111/j.1365-246X.2004.02317.x

Frank, W. B., & Abercrombie, R. E. (2018). Adapting the matched-filter search to a wide-aperture network: An aftershock sequence and an earthquake swarm in connecticut [Journal Article]. *Bulletin of the Seismological Society of America.* doi: 10.1785/0120170190

Freiberger, W. (1963). An approximate method in signal detection [Journal Article]. *Journal of Applied Mathematics*, *20*, 373-378.

Galis, M., Ampuero, J. P., Mai, P. M., & Cappa, F. (2017). Induced seismicity provides insight into why earthquake ruptures stop [Journal Article]. *Science Advances*, *3*(12). doi: 10.1126/sciadv.aap7528

Garagash, D. I., & Germanovich, L. N. (2012). Nucleation and arrest of dynamic slip on a pressurized fault [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *117*(B10). doi: 10.1029/2012jb009209

Geller, R. J., & Mueller, C. S. (1980). Four similar earthquakes in central california [Journal Article]. *Geophysical Research Letters*, *7*(10), 821-824.

Giardini, D. (2009). Geothermal quake risks must be faced [Journal Article]. *Nature*, *462*(17).

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation [Journal Article]. *Geophysical Journal International*, *165*(1), 149-166. doi: 10.1111/j.1365-246X.2006.02865.x

Goebel, T. H. W., Becker, T. W., Schorlemmer, D., Stanchits, S., Sammis, C., Rybacki, E., & Dresen, G. (2012). Identifying fault heterogeneity through mapping spatial anomalies in acoustic emission statistics [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *117*(B3). doi: 10.1029/2011jb008763

Goebel, T. H. W., Kwiatek, G., Becker, T., Brodsky, E. E., & Dresen, G. (2017). What allows seismic events to grow big?: Insights from b-value and fault roughness analysis in laboratory stick-slip experiments [Journal Article]. *Geology*. doi: 10.1130/G39147.1

Goertz-Allmann, B. P., Goertz, A., & Wiemer, S. (2011). Stress drop variations of induced earthquakes at the basel geothermal site [Journal Article]. *Geophysical Research Letters*, *38*(9), n/a-n/a. doi: 10.1029/2011gl047498

Gomberg, J. (2018). Unsettled earthquake nucleation [Journal Article]. *Nature Geoscience*, *11*(7), 463-464. doi: 10.1038/s41561-018-0149-x

Gritto, R., & Jarpe, S. P. (2014). Temporal variations of vp/vs-ratio at the geysers geothermal field, usa [Journal Article]. *Geothermics*, *52*, 112-119. doi: 10.1016/j.geothermics.2014.01.012

Gunasekera, R. C., Foulger, G. R., & Julian, B. R. (2003). Reservoir depletion at the geysers geothermal area, california, shown by four-dimensional seismic tomography [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *108*(B3). doi: 10.1029/2001jb000638

Gutenberg, B., & Richter, C. F. (1956). Magnitude and energy of earthquakes [Journal Article]. *Annals of Geophysics*, *9*(1).

Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale [Journal Article]. *Journal of Geophysical Research*, *84*(B5), 2348. doi: 10.1029/JB084iB05p02348

Hardebeck, J. L., Felzer, K. R., & Michael, A. J. (2008). Improved tests reveal that the accelerating moment release hypothesis is statistically insignificant [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *113*(B8). doi: 10.1029/2007jb005410

Hastings, W. (1970). Monte carlo sampling methods using markov chains and their applications [Journal Article]. *Biometrika*, *57*. doi: 10.1093/biomet/57.1.97/284580

Hauksson, E., Stock, J., Hutton, K., Yang, W., Vidal-Villegas, J. A., & Kanamori, H. (2010). The 2010 m w 7.2 el mayor-cucapah earthquake sequence, baja california, mexico and southernmost california, usa: Active seismotectonics along the mexican pacific margin [Journal Article]. *Pure and Applied Geophysics*, *168*(8-9), 1255-1277. doi: 10.1007/s00024-010-0209-7

Hergert, T., & Heidbach, O. (2010). Slip-rate variability and distributed deformation in the marmara sea fault system [Journal Article]. *Nature Geoscience*, *3*(2), 132-135. doi: 10.1038/ngeo739

Hergert, T., & Heidbach, O. (2011). Geomechanical model of the marmara sea region-ii. 3-d contemporary background stress field [Journal Article]. *Geophysical Journal International*, *185*(3), 1090-1102. doi: 10.1111/j.1365-246X.2011.04992.x

Hirata, T., Satoh, T., & Ito, K. (1987). Fractal structure of spatial distribution of microfracturing in rock [Journal Article]. *Geophysical Journal of the Royal Astronomical Society*, *90*, 369-374.

Hofmann, H., Zimmermann, G., Zang, A., & Min, K.-B. (2018). Cyclic soft stimulation (css): a new fluid injection protocol and traffic light system to mitigate seismic risks of hydraulic stimulation treatments [Journal Article]. *Geothermal Energy*, *6*(1). doi: 10.1186/s40517-018-0114-3

Huang, H., & Meng, L. (2018). Slow unlocking processes preceding the 2015 mw 8.4 illapel, chile, earthquake [Journal Article]. *Geophysical Research Letters*, *45*, 3914–3922. doi: 10.1029/

Häring, M. O., Schanz, U., Ladner, F., & Dyer, B. C. (2008). Characterisation of the basel 1 enhanced geothermal system [Journal Article]. *Geothermics*, *37*(5), 469-495. doi: 10.1016/j.geothermics.2008.06.002

Ide, S., & Beroza, G. C. (2001). Does apparent stress vary with earthquake size? [Journal Article]. *Geophysical Research Letters*, *28*(17), 3349-3352.

Igarashi, T., Matsuzawa, T., & Hasegawa, A. (2003). Repeating earthquakes and interplate aseismic slip in the northeastern japan subduction zone [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *108*(B5). doi: 10.1029/2002jb001920

Israelsson, H. (1990). Correlation of waveforms from closely spaced regional events [Journal Article]. *Bulletin of the Seismological Society of America*, *80*(6), 2177-2193.

Jeanne, P., Rutqvist, J., Vasco, D., Garcia, J., Dobson, P. F., Walters, M., ... Borgia, A. (2014). A 3d hydrogeological and geomechanical model of an enhanced geothermal system at the geysers, california [Journal Article]. *Geothermics*, *51*, 240-252. doi: 10.1016/j.geothermics.2014.01.013

Johnson, C. W., Totten, E. J., & Bürgmann, R. (2016). Depth migration of seasonally induced seismicity at the geysers geothermal field [Journal Article]. *Geophysical Research Letters*, *43*, 6196-6204. doi: 10.1002/2016GL069546

Johnson, L. R. (2014). Source mechanisms of induced earthquakes at the geysers geothermal reservoir [Journal Article]. *Pure and Applied Geophysics*, *171*(8), 1641-1668. doi: 10.1007/s00024-014-0795-x

Johnson, P. A., Ferdowsi, B., Kaproth, B. M., Scuderi, M., Griffa, M., Carmeliet, J., ... Marone, C. (2013). Acoustic emission and microslip precursors to stick-slip failure in sheared granular material [Journal Article]. *Geophysical Research Letters*, *40*(21), 5627-5631. doi: 10.1002/2013gl057848

Jones, L. M. (1985). Foreshocks and time-dependent earthquake hazard assessment in southern california [Journal Article]. *Bulletin of the Seismological Society of America*, *75*(6), 1669-1679.

Jones, L. M., & Molnar, P. (1979). Some characteristics of foreshocks and their possible relationship to earthquake prediction and premonitory slip on faults [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *84*(B7), 3596-3608. doi: 10.1029/JB084iB07p03596

Jost, M. L., Büßelberg, T., Jost, O., & Harjes, H.-P. (1998). Source parameters of injection-induced microearthquakes at 9 km depth at the ktb deep drilling site, germany [Journal Article]. *Bulletin of the Seismological Society of America*, *88*(8), 815-832.

Kanamori, H. (1981). The nature of seismicity patterns before large earthquakes [Journal Article]. *Earthquake Prediction: An International Review*, *1*. doi: 10.1029/ME004p0001

Kanamori, H. (2001). Energy budget of earthquakes and seismic efficiency [Book Section]. In *Earthquake thermodynamics and phase transformations in the earth's interior* (Vol. 76, p. 293-305). International Geophysics.

Karabulut, H., Schmittbuhl, J., Özalaybey, S., Lengliné, O., Kömeç-Mutlu, A., Durand, V., . . . Bouin, M. P. (2011). Evolution of the seismicity in the eastern marmara sea a decade before and after the 17 august 1999 izmit earthquake [Journal Article]. *Tectonophysics*, *510*(1-2), 17-27. doi: 10.1016/j.tecto.2011.07.009

Kato, A., Fukuda, J., Nakagawa, S., & Obara, K. (2016). Foreshock migration preceding the 2016 mw 7.0 kumamoto earthquake, japan [Journal Article]. *Geophysical Research Letters*, *43*, 8945-8953. doi: 10.1002/2016GL070079

Kato, A., & Nakagawa, S. (2014). Multiple slow-slip events during a foreshock sequence of the 2014 iquique, chile mw 8.1 earthquake [Journal Article]. *Geophysical Research Letters*, *41*, 5420–5427. doi: 10.1002/

Kato, A., Obara, K., Igarashi, T., Tsuruoka, H., Nakagawa, S., & Hirata, N. (2012). Propagation of slow slip leading up to the 2011mw 9.0 tohoku-oki earthquake [Journal Article]. *Science*, *335*, 705-708. doi: 10.1126/science.1215141

KOERI. (2001). Bogazici university kandilli observatory and earthquake research institute [Journal Article]. *International Federation of Digital Seismograph Networks.* doi: 10.7914/SN/KO

Kostrov, V. V. (1974). Seismic moment and energy of earthquakes, and seismic flow of rock [Journal Article]. *Earth Physics*, *1*, 23-40.

Kraft, T., & Deichmann, N. (2014). High-precision relocation and focal mechanism of the injection-induced seismicity at the basel egs [Journal Article]. *Geothermics*, *52*, 59-73. doi: 10.1016/j.geothermics.2014.05.014

Kwiatek, G., & Ben-Zion, Y. (2016). Theoretical limits on detection and analysis of small earthquakes [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *121*(8), 5898-5916. doi: 10.1002/2016jb012908

Kwiatek, G., Bulut, F., Bohnhoff, M., & Dresen, G. (2014). High-resolution analysis of seismicity induced at berlín geothermal field, el salvador [Journal Article]. *Geothermics*, *52*, 98-111. doi: 10.1016/j.geothermics.2013.09.008

Kwiatek, G., Goebel, T. H. W., & Dresen, G. (2014). Seismic moment tensor and b value variations over successive seismic cycles in laboratory stick-slip experiments [Journal Article]. *Geophysical Research Letters*, *41*(16), 5838-5846. doi: 10.1002/2014gl060159

Kwiatek, G., Martínez-Garzón, P., Dresen, G., Bohnhoff, M., Sone, H., & Hartline, C. (2015). Effects of long-term fluid injection on induced seismicity parameters and maximum magnitude in northwestern part of the geysers geothermal field [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *120*(10), 7085-7101. doi: 10.1002/2015jb012362

Kwiatek, G., Plenkers, K., & Dresen, G. (2011). Source parameters of picoseismicity recorded at mponeng deep gold mine, south africa: Implications for scaling relations [Journal Article]. *Bulletin of the Seismological Society of America*, *101*(6), 2592-2608. doi: 10.1785/0120110094

Kwiatek, G., Saarno, T., Adler, T., Bluemle, F., Bohnhoff, M., Chendorain, M., ... Wollin, C. (2019). Controlling fluid-induced seismicity during a 6.1-km-deep geothermal stimulation in finland [Journal Article]. *Science Advances*, *5*.

Lange, D., Kopp, H., Royer, J. Y., Henry, P., Cakir, Z., Petersen, F., ... Geli, L. (2019). Interseismic strain build-up on the submarine north anatolian fault offshore istanbul [Journal Article]. *Nat Commun*, *10*(1), 3006. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/31285439` doi: 10.1038/s41467-019-11016-z

Lapusta, N., & Rice, J. R. (2003). Nucleation and early seismic propagation of small and large events in a crustal earthquake model [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *108*(B4). doi: 10.1029/2001jb000793

Lei, X., Kusunose, K., Satoh, T., & Nishizawa, O. (2003). The hierarchical rupture process of a fault: an experimental study [Journal Article]. *Physics of the Earth and Planetary Interiors*, *137*(1-4), 213-228. doi: 10.1016/s0031-9201(03)00016-5

Lengliné, O., & Marsan, D. (2009). Inferring the coseismic and postseismic stress changes caused by the 2004mw= 6 parkfield earthquake from variations of recurrence times of microearthquakes [Journal Article]. *Journal of Geophysical Research*, *114*(B10). doi: 10.1029/2008jb006118

Leonard, M. (2010). Earthquake fault scaling: Self-consistent relating of rupture length, width, average displacement, and moment release [Journal Article]. *Bulletin of the Seismological Society of America*, *100*(5A), 1971-1988. doi: 10.1785/0120090189

Le Pichon, X., Sengör, A. M. C., Demirbağ, E., Rangin, C., Imren, C., Armijo, R., ... Tok, B. (2001). The active main marmara fault [Journal Article]. *Earth and Planetary Science Letters*, *192*.

Leptokaropoulos, K., Staszek, M., Lasocki, S., Martínez-Garzón, P., & Kwiatek, G. (2018). Evolution of seismicity in relation to fluid injection in the north-western part of the geysers geothermal field [Journal Article]. *Geophysical Journal International*, *212*(2), 1157-1166. doi: 10.1093/gji/ggx481

Lipman, S. C., Strobel, C. J., & Gulati, M. S. (1978). Reservoir performance of the geysers field [Journal Article]. *Geothermics*, *7*, 209-219.

Lisowski, M., Savage, J. C., & Prescott, W. H. (1991). The velocity field along the san andreas fault in central and southern california [Journal Article]. *Journal of Geophysical Research*, *96*(B5), 8369. doi: 10.1029/91jb00199

Lohman, R. B., & McGuire, J. J. (2007). Earthquake swarms driven by aseismic creep in the salton trough, california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *112*(B4). doi: 10.1029/2006jb004596

Maceira, M., Rowe, C. A., Beroza, G., & Anderson, D. (2010). Identification of low-frequency earthquakes in non-volcanic tremor using the subspace detector method [Journal Article]. *Geophysical Research Letters*, *37*(6), n/a-n/a. doi: 10.1029/2009gl041876

Majer, E., Nelson, J., Robertson-Tait, A., Savy, J., & Wong, I. (2012). Protocol for addressing induced seismicity associated with enhanced geothermal systems [Journal Article]. *US Department of Energy*, *DOE EE-0662*.

Majer, E. L., & Peterson, J. E. (2007). The impact of injection on seismicity at the geysers, california geothermal field [Journal Article]. *International Journal of Rock Mechanics and Mining Sciences*, *44*(8), 1079-1090. doi: 10.1016/j.ijrmms.2007.07.023

Malin, P. E., Bohnhoff, M., Blümle, F., Dresen, G., Martínez-Garzón, P., Nurlu, M., … Yanik, K. (2018). Microearthquakes preceding a m4.2 earthquake offshore istanbul [Journal Article]. *Scientific Reports*, *8*(1). doi: 10.1038/s41598-018-34563-9

Martínez-Garzón, P., Bohnhoff, M., Kwiatek, G., & Dresen, G. (2013). Stress tensor changes related to fluid injection at the geysers geothermal field, california [Journal Article]. *Geophysical Research Letters*, *40*(11), 2596-2601. doi: 10.1002/grl.50438

Martínez-Garzón, P., Bohnhoff, M., Mencin, D., Kwiatek, G., Dresen, G., Hodgkinson, K., … Kartal, R. F. (2019). Slow strain release along the eastern marmara region offshore istanbul in conjunction with enhanced local seismic moment release [Journal Article]. *Earth and Planetary Science Letters*, *510*, 209-218. doi: 10.1016/j.epsl.2019.01.001

Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., & Dresen, G. (2017). Volumetric components in the earthquake source related to fluid injection and stress state [Journal Article]. *Geophysical Research Letters*, *44*(2), 800-809. doi: 10.1002/2016gl071963

Martínez-Garzón, P., Kwiatek, G., Sone, H., Bohnhoff, M., Dresen, G., & Hartline, C. (2014). Spatiotemporal changes, faulting regimes, and source parameters of induced seismicity: A case study from the geysers geothermal field [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *119*(11), 8378-8396. doi: 10.1002/2014jb011385

Martínez-Garzón, P., Vavryčuk, V., Kwiatek, G., & Bohnhoff, M. (2016). Sensitivity of stress inversion of focal mechanisms to pore pressure changes [Journal Article]. *Geophysical Research Letters*, *43*(16), 8441-8450. doi: 10.1002/2016gl070145

Martínez-Garzón, P., Zaliapin, I., Ben-Zion, Y., Kwiatek, G., & Bohnhoff, M. (2018). Comparative study of earthquake clustering in relation to hydraulic activities at geothermal fields in california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *123*(5), 4041-4062. doi: 10.1029/2017jb014972

Mavrommatis, A. P., Segall, P., Uchida, N., & Johnson, K. M. (2015). Long-term acceleration of aseismic slip preceding the mw 9 tohoku-oki earthquake: Constraints from repeating earthquakes [Journal Article]. *Geophysical Research Letters*, *42*(22), 9717-9725. doi: 10.1002/2015gl066069

Maxwell, S., Shemeta, J., Campbell, E., & Quirk, D. (2008). Microseismic deformation rate monitoring [Journal Article]. *SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers.*

McGarr, A. (1976). Seismic moments and volume changes [Journal Article]. *Journal of Geophysical Research*, *81*(8), 1487-1494. doi: 10.1029/JB081i008p01487

McGarr, A. (1999). On relating apparent stress to the stress causing earthquake fault slip [Journal Article]. *Journal of Geophysical Research*, *104*(B2), 3003-3011.

McGarr, A. (2014). Maximum magnitude earthquakes induced by fluid injection [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *119*, 1008-1019. doi: 10.1002/2013JB010597

Meng, X., & Peng, Z. (2014). Seismicity rate changes in the salton sea geothermal field and the san jacinto fault zone after the 2010 mw 7.2 el mayor-cucapah earthquake [Journal Article]. *Geophysical Journal International*, *197*(3), 1750-1762. doi: 10.1093/gji/ggu085

Meng, X., Yu, X., Peng, Z., & Hong, B. (2012). Detecting earthquakes around salton sea following the 2010 mw7.2 el mayor-cucapah earthquake using gpu parallel computing [Journal Article]. *Procedia Computer Science*, *9*, 937-946. doi: 10.1016/j.procs.2012.04.100

Meredith, P. G., Main, I. G., & Jones, C. (1990). Temporal variations in seismcitiy during quasi-static and dynamic rock failure [Journal Article]. *Tectonophysics*, *175*, 249-268. doi: 10.1016/0040-1951(90)90141-T

Mignan, A. (2011). Retrospective on the accelerating seismic release (asr) hypothesis: Controversy and new horizons [Journal Article]. *Tectonophysics*, *505*(1-4), 1-16. doi: 10.1016/j.tecto.2011.03.010

Mignan, A. (2012). Seismicity precursors to large earthquakes unified in a stress accumulation framework [Journal Article]. *Geophysical Research Letters*, *39*(21), n/a-n/a. doi: 10.1029/2012gl053946

Mignan, A. (2014). The debate on the prognostic value of earthquake foreshocks: a meta-analysis [Journal Article]. *Sci Rep*, *4*, 4099. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/24526224` doi: 10.1038/srep04099

Mossop, A., & Segall, P. (1997). Subsidence at the geysers geothermal field, n. california from a comparison of gps and leveling surveys [Journal Article]. *Geophysical Research Letters*, *24*(14), 1839-1842. doi: 10.1029/97gl51792

Murray, J. R., Minson, S. E., & Svarc, J. L. (2014). Slip rates and spatially variable creep on faults of the northern san andreas system inferred through bayesian inversion of global positioning system data [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *119*(7), 6023-6047. doi: 10.1002/2014jb010966

Norbeck, J. H., & Horne, R. N. (2018). Maximum magnitude of injection-induced earthquakes: A criterion to assess the influence of pressure migration along faults [Journal Article]. *Tectonophysics*, *733*, 108-118. doi: 10.1016/j.tecto.2018.01.028

Ohnaka, M. (1992). Earthquake source nucleation: a physical model for short-term precursors [Journal Article]. *Tectonophysics*, *211*, 149-178.

Oppenheimer, D. H. (1986). Extensional tectonics at the geysers geothermal area, california [Journal Article]. *Journal of Geophysical Research*, *91*(B11), 11463. doi: 10.1029/JB091iB11p11463

Oye, V., Chavarria, J. A., & Malin, P. E. (2004). Determining safod area microearthquake locations solely with the pilot hole seismic array data [Journal Article]. *Geophysical Research Letters*, *31*(12), n/a-n/a. doi: 10.1029/2003gl019403

Page, R. (1968). Aftershocks and microaftershocks of the great alaska earthquake of 1964 [Journal Article]. *Bulletin of the Seismological Society of America*, *58*(3), 1131-1168.

Paitz, P., Gokhberg, A., & Fichtner, A. (2018). A neural network for noise correlation classification [Journal Article]. *Geophysical Journal International*, *212*(2), 1468-1474. doi: 10.1093/gji/ggx495

Papazachos, B. C. (1975). Foreshocks and earthquake prediction [Journal Article]. *Tectonophysics*, *28*, 213-226.

Papoulis, A. (1962). *The fourier integral and its applications* [Book]. New York: McGraw-Hill.

Parsons, T. (2004). Recalculated probability of m > 7 earthquakes beneath the sea of marmara, turkey [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *109*(B5). doi: 10.1029/2003jb002667

Peng, Z., & Zhao, P. (2009). Migration of early aftershocks following the 2004 parkfield earthquake [Journal Article]. *Nature Geoscience*, *2*(12), 877-881. doi: 10.1038/ngeo697

Perfettini, H., & Avouac, J. P. (2007). Modeling afterslip and aftershocks following the 1992 landers earthquake [Journal Article]. *Journal of Geophysical Research*, *112*(B7). doi: 10.1029/2006jb004399

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location [Journal Article]. *Science Advances*, *4*.

Poupinet, G., Ellsworth, W. L., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the calaveras fault, california [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5719-5731. doi: 10.1029/JB089iB07p05719

Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals* [Book]. New Jersey: Prentice Hall.

Raub, C., Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., & Dresen, G. (2017). Variations of seismic b -value at different stages of the seismic cycle along the north anatolian fault zone in northwestern turkey [Journal Article]. *Tectonophysics*. doi: 10.1016/j.tecto.2017.05.028

Reilinger, R. E., McClusky, S. C., Oral, M. B., King, R. W., Toksoz, M. N., Barka, A. A., … Sanli, I. (1997). Global positioning system measurements of present-day crustal movements in the arabia-africa-eurasia plate collision zone [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *102*(B5), 9983-9999. doi: 10.1029/96jb03736

Rice, J. R. (1993). Spatio-temporal complexity of slip on a fault [Journal Article]. *Journal of Geophysical Research*, *98*(B6), 9885. doi: 10.1029/93jb00191

Ripperger, J., Ampuero, J. P., Mai, P. M., & Giardini, D. (2007). Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *112*(B4). doi: 10.1029/2006jb004515

Roland, E., & McGuire, J. J. (2009). Earthquake swarms on transform faults [Journal Article]. *Geophysical Journal International*, *178*(3), 1677-1690. doi: 10.1111/j.1365-246X.2009.04214.x

Ross, A., Foulger, G. R., & Julian, B. R. (1996). Non-double-couple earthquake mechanisms at the geysers geothermal area, california [Journal Article]. *Geophysical Research Letters*, *23*(8), 877-880. doi: 10.1029/96gl00590

Ross, A., Foulger, G. R., & Julian, B. R. (1999). Source processes of industrially-induced earthquakes at the geysers geothermal area, california [Journal Article]. *Geophysics*, *64*. doi: 10.1190/1.1444694

Ross, Z. E., Meier, M.-A., & Hauksson, E. (2018). P wave arrival picking and first-motion polarity determination with deep learning [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *123*(6), 5120-5129. doi: 10.1029/2017jb015251

Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in southern california [Journal Article]. *Science*, *364*(6442), 767-771. doi: 10.1126/science.aaw6888

Rost, S. (2002). Array seismology: Methods and applications [Journal Article]. *Reviews of Geophysics*, *40*(3). doi: 10.1029/2000rg000100

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes [Journal Article]. *Geophysical Research Letters*, *44*(18), 9276-9282. doi: 10.1002/2017gl074677

Ruiz, S., Metois, M., Fuenzalida, A., Ruiz, J., Leyton, F., Grandin, R., … Campos, J. (2014). Intense foreshocks and a slow slip event preceded the 2014 iquique mw 8.1 earthquake [Journal Article]. *Science*, *345*(6201), 1165-1169.

Rutqvist, J., Jeanne, P., Dobson, P. F., Garcia, J., Hartline, C., Hutchings, L., … Walters, M. (2016). The northwest geysers egs demonstration project, california – part 2: Modeling and interpretation [Journal Article]. *Geothermics*, *63*, 120-138. doi: 10.1016/j.geothermics.2015.08.002

Savage, H. M., Keranen, K. M., Schaff, D. P., & Dieck, C. (2017). Possible precursory signals in damage zone foreshocks [Journal Article]. *Geophysical Research Letters*, *44*, 5411-5417. doi: 10.1002/2017GL073226

Schaff, D. P., & Richards, P. G. (2014). Improvements in magnitude precision, using the statistics of relative amplitudes measured by cross correlation [Journal Article]. *Geophysical Journal International*, *197*(1), 335-350. doi: 10.1093/gji/ggt433

Schmittbuhl, J., Karabulut, H., Lengliné, O., & Bouchon, M. (2016a). Long-lasting seismic repeaters in the central basin of the main marmara fault [Journal Article]. *Geophysical Research Letters*, *43*(18), 9527-9534. doi: 10.1002/2016gl070505

Schmittbuhl, J., Karabulut, H., Lengliné, O., & Bouchon, M. (2016b). Seismicity distribution and locking depth along the main marmara fault, turkey [Journal Article]. *Geochemistry, Geophysics, Geosystems*, *17*(3), 954-965. doi: 10.1002/2015gc006120

Schoenball, M., Dorbath, L., Gaucher, E., Wellmann, J. F., & Kohl, T. (2014). Change of stress regime during geothermal reservoir stimulation [Journal Article]. *Geophysical Research Letters*, *41*(4), 1163-1170. doi: 10.1002/2013gl058514

Schoenball, M., Müller, T. M., Müller, B. I. R., & Heidbach, O. (2010). Fluid-induced microseismicity in pre-stressed rock masses [Journal Article]. *Geophysical Journal International*, *180*(2), 813-819. doi: 10.1111/j.1365-246X.2009.04443.x

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes [Journal Article]. *Bulletin of the Seismological Society of America*, *58*(1), 399-415.

Scholz, C. H. (2015). On the stress dependence of the earthquake b value [Journal Article]. *Geophysical Research Letters*, *42*(5), 1399-1402. doi: 10.1002/2014gl062863

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes [Journal Article]. *Nature*, *437*(7058), 539-42. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/16177788` doi: 10.1038/nature04094

Schurr, B., Asch, G., Hainzl, S., Bedford, J., Hoechner, A., Palo, M., ... Vilotte, J. P. (2014). Gradual unlocking of plate boundary controlled initiation of the 2014 iquique earthquake [Journal Article]. *Nature*, *512*(7514), 299-302. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/25119049` doi: 10.1038/nature13681

Segall, P., & Fitzgerald, S. D. (1998). A note on induced stress changes in hydrocarbon and geothermal reservoirs [Journal Article]. *Tectonophysics*, *289*, 117-128.

Selvadurai, P. A., & Glaser, S. D. (2015). Laboratory-developed contact models controlling instability on frictional faults [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *120*(6), 4208-4236. doi: 10.1002/2014jb011690

Senobari, N. S., Funning, G. J., Keogh, E., Zhu, Y., Yeh, C. M., Zimmerman, Z., & Mueen, A. (2018). Super-efficient cross-correlation (sec-c): A fast matched filtering code suitable for desktop computers [Journal Article]. *Seismological Research Letters*. doi: 10.1785/0220180122

Shapiro, S. A., Dinske, C., & Kummerow, J. (2007). Probability of a given-magnitude earthquake induced by a fluid injection [Journal Article]. *Geophysical Research Letters*, *34*(22). doi: 10.1029/2007gl031615

Shapiro, S. A., Dinske, C., Langenbruch, C., & Wenzel, F. (2010). Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations [Journal Article]. *The Leading Edge*, *29*, 304-309. doi: 10.1190/1.3353727

Shapiro, S. A., Krüger, O. S., Dinske, C., & Langenbruch, C. (2011). Magnitudes of induced earthquakes and geometric scales of fluid-stimulated rock volumes [Journal Article]. *Geophysics*, *76*(6), WC55-WC63.

Shaw, B. E., Carlson, J. M., & Langer, J. S. (1992). Patterns of seismic activity preceding large earthquakes [Journal Article]. *Journal of Geophysical Research*, *97*(B1), 479. doi: 10.1029/91jb01796

Shearer, P. M. (1994). Global seismic event detection using a matched filter on long-period seismograms [Journal Article]. *Journal of Geophysical Research*, *99*(B7), 13,713 - 13,725.

Shelly, D. R., Beroza, G. C., & Ide, S. (2007). Non-volcanic tremor and low-frequency earthquake swarms [Journal Article]. *Nature*, *446*(7133), 305-7. Retrieved from `https://www.ncbi.nlm.nih.gov/pubmed/17361180` doi: 10.1038/nature05666

Sieh, K. E., & Williams, P. L. (1990). Behavior of the southernmost san andreas fault during the past 300 years [Journal Article]. *Journal of Geophysical Research*, *95*(B5), 6629. doi: 10.1029/JB095iB05p06629

Socquet, A., Valdes, J. P., Jara, J., Cotton, F., Walpersdorf, A., Cotte, N., ... Norabuena, E. (2017). An 8 month slow slip event triggers progressive nucleation of the 2014 chile megathrust [Journal Article]. *Geophysical Research Letters*, *44*(9), 4046-4053. doi: 10.1002/2017gl073023

Stein, R., Barka, A., & Dieterich, J. H. (1997). Progressive failure on the north anatolian fault since 1939 by earthquake stress triggering [Journal Article]. *Geophysical Journal International*, *128*, 594-604.

Sugan, M., Vuan, A., Kato, A., Massa, M., & Amati, G. (2019). Seismic evidence of an early afterslip during the 2012 sequence in emilia (italy) [Journal Article]. *Geophysical Research Letters*, *46*, 625-635. doi: 10.1029/2018GL079617

Tape, C., Holtkamp, S., Silwal, V., Hawthorne, J., Kaneko, Y., Ampuero, J. P., ... West, M. E. (2018). Earthquake nucleation and fault slip complexity in the lower crust of central alaska [Journal Article]. *Nature Geoscience*, *11*(7), 536-541. doi: 10.1038/s41561-018-0144-2

Trugman, D. T., & Ross, Z. E. (2019). Pervasive foreshock activity across southern california [Journal Article]. *Geophysical Research Letters*, *46*(15), 8772-8781. doi: 10.1029/2019gl083725

Trugman, D. T., Shearer, P. M., Borsa, A. A., & Fialko, Y. (2016). A comparison of long-term changes in seismicity at the geysers, salton sea, and coso geothermal fields [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *121*(1), 225-247. doi: 10.1002/2015jb012510

Turin, G. L. (1960). An introduction to matched filters [Journal Article]. *IRE Transactions on Information Theory*, *6*(3), 311-329.

Uchida, N., Kalafat, D., Pinar, A., & Yamamoto, Y. (2019). Repeating earthquakes and interplate coupling along the western part of the north anatolian fault [Journal Article]. *Tectonophysics*, *769*, 228185. doi: 10.1016/j.tecto.2019.228185

van der Elst, N. J., Page, M. T., Weiser, D. A., Goebel, T. H. W., & Hosseini, S. M. (2016). Induced earthquake magnitudes are as large as (statistically) expected [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *121*(6), 4575-4590. doi: 10.1002/2016jb012818

van Trees, H. (1968). *Detection, estimation and modulation theory* [Book]. New York: John Wiley and Sons, Inc.

Vuan, A., Sugan, M., Amati, G., & Kato, A. (2018). Improving the detection of low-magnitude seismicity preceding the mw 6.3 l'aquila earthquake: Development of a scalable code based on the cross correlation of template earthquakes [Journal Article]. *Bulletin of the Seismological Society of America*, *108*(1), 471-480. doi: 10.1785/0120170106

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california [Journal Article]. *Bulletin of the Seismological Society of America*, *90*(6), 1353-1368.

Walter, J. I., Chang, J. C., & Dotray, P. J. (2017). Foreshock seismicity suggests gradual differential stress increase in the months prior to the 3 september 2016mw5.8 pawnee earthquake [Journal Article]. *Seismological Research Letters*, *88*(4), 1032-1039. doi: 10.1785/0220170007

Wang, J., Templeton, D. C., & Harris, D. B. (2015). Discovering new events beyond the catalogue—application of empirical matched field processing to salton sea geothermal field seismicity [Journal Article]. *Geophysical Journal International*, *203*(1), 22-32. doi: 10.1093/gji/ggv260

Wiemer, S., & Katsumata, K. (1999). Spatial variability of seismicity parameters in aftershock zones [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *104*(B6), 13135-13151. doi: 10.1029/1999jb900032

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from alaska, the western united states, and japan [Journal Article]. *Bulletin of the Seismological Society of America*, *90*(4), 859-869. doi: 10.1785/0119990114

Wollin, C., Bohnhoff, M., Martínez-Garzón, P., Küperkoch, L., & Raub, C. (2018). A unified earthquake catalogue for the sea of marmara region, turkey, based on automatized phase picking and travel-time inversion: Seismotectonic implications [Journal Article]. *Tectonophysics*, *747-748*, 416-444. doi: 10.1016/j.tecto.2018.05.020

Woodward, P. M. (1953). *Probability and information theory, with applications to radar* (2nd ed.) [Book]. Pergamon Press.

Yamamoto, R., Kido, M., Ohta, Y., Takahashi, N., Yamamoto, Y., Pinar, A., ... Kaneda, Y. (2019). Seafloor geodesy revealed partial creep of the north anatolian fault submerged in the sea of marmara [Journal Article]. *Geophysical Research Letters*, *46*, 1268-1275. doi: 10.1029/2018GL080984

Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search [Journal Article]. *Science*.

Yoon, C. E., Yoshimitsu, N., Ellsworth, W. L., & Beroza, G. C. (2019). Foreshocks and mainshock nucleation of the 1999 mw 7.1 hector mine, california, earthquake [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *124*(2), 1569-1582. doi: 10.1029/2018jb016383

Yu, C., Vavryčuk, V., Adamová, P., & Bohnhoff, M. (2018). Moment tensors of induced microearthquakes in the geysers geothermal reservoir from broadband seismic recordings: Implications for faulting regime, stress tensor and fluid pressure [Journal Article]. *Journal of Geophysical Research: Solid Earth*. doi: 10.1029/2018jb016251

Zaliapin, I., & Ben-Zion, Y. (2013a). Earthquake clusters in southern california ii: Classification and relation to physical properties of the crust [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2865-2877. doi: 10.1002/jgrb.50178

Zaliapin, I., & Ben-Zion, Y. (2013b). Earthquake clusters in southern california i: Identification and stability [Journal Article]. *Journal of Geophysical Research: Solid Earth*, *118*(6), 2847-2864. doi: 10.1002/jgrb.50179

Zaliapin, I., & Ben-Zion, Y. (2016). A global classification and characterization of earthquake clusters [Journal Article]. *Geophysical Journal International*, *207*(1), 608-634. doi: 10.1093/gji/ggw300

Zang, A., Wagner, C., Stanchits, S., Dresen, G., Andresen, R., & Haidekker, M. (1998). Source analysis of acoustic emissions in aue granite cores under symmetric and asymmetric compressive loads [Journal Article]. *Geophysical Journal International*, *135*, 1113-1130.

Zechar, J. D., & Zhuang, J. (2010). Risk and return: evaluating reverse tracing of precursors earthquake predictions [Journal Article]. *Geophysical Journal International*, *182*(3), 1319-1326. doi: 10.1111/j.1365-246X.2010.04666.x

Zhou, H.-w. (1994). Rapid three-dimensional hypocentral determination using a master station method [Journal Article]. *Journal of Geophysical Research*, *99*(B8), 15439. doi: 10.1029/94jb00934

Zoback, M., & Townend, J. (2001). Implications of hydrostatic pore pressures and high crustal strength for the deformation of intraplate lithosphere [Journal Article]. *Tectonophysics*, *336*, 19-30.

List of Figures

List of Tables

Curriculum Vitae

For reasons of data protection, the curriculum vitae is not included in the online version of the dissertation.

For reasons of data protection, the curriculum vitae is not included in the online version of the dissertation.