Stress Drop Variability: Implications for Earthquake Rupture Processes, Scaling Relations, and Seismic Ground Motions

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

am Fachbereich Geowissenschaften
der Freien Universität Berlin

vorgelgt von
Aglaja Blanke

Berlin, 2021

Eidesstaatliche Erklärung:

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

Berlin, den 26.02.2021

Referent:	Prof. Dr. Marco Bohnhoff
Korreferent:	Prof. Dr. Frederik Tilmann
Datum der Disputation:	10.05.2021

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

Acknowledgements

First of all, I would like to express my sincere gratitude to my academic supervisor, Prof. Marco Bohnhoff, for giving me the chance to do great research at GFZ. I am very grateful for all the support and trust that I received during my PhD to push forward my research career. Furthermore, I am very thankful for the opportunity to continue my work at GFZ. Mt. Etna is calling.

I would also like to thank Prof. Frederik Tilmann for taking over the position of the 2nd examiner of this thesis and for evaluating my work. In addition, I would like to thank all further committee members Dr. Stine Gutjahr, Prof. Serge A. Shapiro and Prof. Eline Le Breton for their precious time.

I wish to express my sincere gratitude to Grzegorz Kwiatek, who always stood by my side with his great expertise. Without you, I would not have been able to finish this thesis. Thank you for all your patience, support and cheering ups during my whole PhD. Thank you also for taking me on board of the EPOS project, which opened new doors for me and let me meet many great researchers in Europe.

I thank my co-authors Grzegorz Kwiatek, Thomas Goebel, Georg Dresen, Patricia Martínez-Garzón, and Marco Bohnhoff for supporting my work and for all discussions that finally helped me to complete this PhD thesis.

Accordingly, I am thankful to be co-author of the here presented EPOS article. I wish to express my gratitude to the whole EPOS TCS AH team for their collaboration and all the nice meetings we had at wonderful places. I am very happy to be part of our group and moreover, I am proud of what we achieved in the past years.

I would also like to thank Peter Malin for the fruitful discussion on coda waves and for showing interest in our research topics.

I thank all my former and present colleagues at GFZ for the pleasant atmosphere, nice conversations on science or beyond, and several funny after-work-meetings. I appreciate a lot to be part of our section. Without all of you, it would not have been the same.

Special thanks to Maria, who shared the office with me for a long time. I thank you for all our conversations, coffee breaks, shopping trips and visits at the hair dresser after work. I hope we can continue some of these things in the future.

Dear Ilaria, I am very happy that I met you in 2018 at ESC. Sharing my thoughts about fears and sorrows but also about the excitement for our work and progress of our personal skills always cheered me up. I am very grateful that I got the opportunity to be your guest at OGS-CRS in Udine and that I got the great chance to present my work to you and your colleagues.

My gratitude also goes to my very good friend Dennis Pietsch that has always been

there for me and accepted all my moods without judging me. I am also very thankful to all my other friends that have been very patient with me during the past years and haven't seen me a lot. I promise you to work on that issue.

Many thanks to my cousin Felicitas Macgilchrist, who spend her precious time on proofreading the new parts of my thesis. Many thanks also to Valerian and Maria for shortening my very nested German sentences in the German summary.

Finally, I wish to express my deepest gratitude to my family. We went through emotional times but always had the support of each other. Thank you for always being there for me.

Summary

The debate on self-similarity of source physics between small and large earthquakes has been high on the agenda in earthquake science for decades. Seismic source parameters, such as corner frequency, seismic moment, and stress drop, to name only a few, have been analyzed and compared across the entire earthquake magnitude range, considering large natural events, induced seismicity, and laboratory acoustic emissions. Stress drop is an important source parameter that substantially affects ground motions and is often used as an earthquake scaling parameter. The majority of stress drop studies show on average independence on earthquake size and thus scale-invariance is assumed. Scaling breakdown has, however, also been reported for individual datasets in which stress drop scales with earthquake magnitude. Regardless of the debate on self-similarity, a large global stress drop scatter of 0.01-100 MPa has been observed throughout all earthquake sizes, which is not yet fully understood. The reasons for this scatter need to be investigated in more detail to better understand the similarities and differences between earthquake source characteristics and the corresponding physical processes in the earthquake source.

This doctoral thesis focuses on small earthquakes, including induced seismicity and laboratory acoustic emission events, and addresses two groups of factors whose effects are assumed to cause scatter in stress drop estimates: 1) Non-physical factors that cover error propagation, and inappropriate assumptions made during the analysis of seismic data, can lead to bias in source parameter estimates. 2) Physical factors, which reflect the actual properties of earthquakes foci and rupture processes, are assumed to naturally influence the resulting static stress drop. In this thesis, non-physical effects on stress drop estimates are suppressed by considering only high-quality data, adequate methodologies, and carefully evaluated and selected parameters for analysis. One aspect that this thesis studies in more detail, is the effect of high-frequency wave attenuation on seismic records. Attenuation is expressed as the quality factor Q, which describes the quality of the medium that is in general unknown. Q is important for the correction of amplitude source spectra while estimating source parameters. Incorrect assumptions of Q can bias the high-frequency spectral fall-off leading to error propagation (a non-physical effect) during further data analysis.

The first study analyzes the time-dependent decay of S-coda waves of induced seismicity at *The Geysers* geothermal field, California, to better assess attenuation effects for high-frequency seismic signals. The application of the *moving window method* enables us to estimate stable coda quality factors (Q_C) for pre-defined frequency bands in the frequency domain. An additional sensitivity analysis is conducted beforehand to evaluate

the impact of different parameters used in the coda technique, which are the magnitude range, signal-to-noise ratio, maximum permitted uncertainties of Q_C estimates, moving window width, lapse time effect, and total coda length. Findings of the sensitivity analysis emphasize the relevance of parameterization and the usage of high quality data. Waveforms of in total 717 shallow earthquakes with duration magnitudes $1 < M_D < 3$ of two spatially separated locations in the northwestern and southeastern part of *The Geyser*s were investigated. Both areas show clear differences in attenuation properties dependent on locally varying geological, structural and geothermal characteristics. The coda quality factor Q_C was estimated for frequencies up to 70 Hz, exceeding previous field data considerations. Q_C obtained from the northwestern dataset was also investigated in the context of temporal injection variations. The temporal stability of Q_C suggests that either the local fault network is on average constant over time or the injection-induced fractures are not detectable with the frequencies considered. Additionally, Q_C from the northwestern dataset is compared to the quality factor obtained by direct S-waves (Q_β). Q_C shows on average higher stability in mean estimates compared to Q_β. Therefore, it is suggested that Q_C leads to fewer uncertainties of further estimated source parameters such as stress drop.

The second study focuses on acoustic emission (AE) events of two laboratory triaxial stick-slip experiments on oven-dried Westerly granite samples. Source parameters, in particular static stress drop, of acoustic emission events from a rough and a smooth fault were analyzed taking advantage of the *spectral ratio method* based on a multi-eGf (empirical Green's function) approach. Here, attenuation effects (path- and site effects) of linked, co-located events were suppressed and source parameters were directly obtained from the source spectra using the quasi-dynamic *Madariaga source model*. For the first time, a span of more than three orders of magnitude ($-9 < M_W < -5.6$) was evaluated for laboratory AE events providing a reasonable base for scaling analysis. Obtained AE stress drops are comparable to the globally observed estimates. However, a clear scaling breakdown was observed for both faults putting the global self-similarity assumption into question. Here, the physical aspects of fault surface roughness, source radius, and rupture velocity were scrutinized. AE stress drop shows mainly no dependence on source size and only very little variation with rupture velocity changes. The observed stress drop–magnitude scaling therefore suggests differences of slip over rough and smooth fault surfaces. The complexity of rough faults might inhibit larger slips leading to lower stress drops, whereas smooth faults can result in larger slips and thus higher stress drops due to fault simplicity.

These two studies provide new insights into the analysis of high-frequency seismic signals. More comparative studies based on similar approaches are necessary for further

earthquake cross-scale investigations. The variety of contemporary analytical methods and models render the comparison of studies challenging. Therefore, data- and interdisciplinary expertise exchange, data standardization, careful data quality evaluation, and preservation of methodological applications are needed to facilitate and accelerate future research. Drawing on these crucial requirements for future investigations, the third part of this thesis complements the first two scientific studies. It shows the innovative development of an online research platform headed by the Thematic Core Service Anthropogenic Hazards (TCS AH) within the framework of the Implementation Phase of the European Plate Observing System (EPOS-IP). The TCS AH team introduces the genesis of the novel online platform *IS-EPOS* (IS = Induced-Seismicity) that consists of a large number of complete, comprehensive datasets and applications related to the exploration and exploitation of georesources and the underground storage of liquids and gases. The pioneering work of the EPOS project provides an example of how enhanced research could deal with the similarities and differences between individual studies, and thus lead to a more comprehensive understanding of seismic data and its sources.

Zusammenfassung

Bereits seit Jahrzenten ist die Selbstähnlichkeit kleiner und großer Erdbeben eine vielfältig diskutierte Frage. Im Rahmen dessen werden häufig Erdbebenquellparameter, wie beispielsweise die Eckfrequenz, das seismische Moment und der seismische Spannungsabfall, des gesamten Magnitudenspektrums von großen natürlichen Beben, über induzierte Seismizität, bis hin zu kleinsten akustischen Emissionen aus Laborexperimenten analysiert und miteinander verglichen. Einer der wichtigsten Parameter ist dabei der seismisch bedingte Spannungsabfall, der die Bodenbewegung beeinflusst. Aufgrund der häufig angenommenen Unabhängigkeit zwischen Erdbebenmagnitude und Spannungsabfall, wird letzterer bevorzugt für die Skalierung von Erdbeben genutzt und gilt im Allgemeinen als skaleninvariant. Einige individuelle Datensätze weisen jedoch auch eine Abweichung der Skaleninvarianz auf, wobei der Spannungsabfall mit größer werdender Erdbebenmagnitude ansteigt. Neben der noch ungeklärten Skaleninvarianz wird außerdem eine große Schwankung der seismischen Spannungsabfälle (0.01-100 MPa) im gesamten Magnitudenspektrum beobachtet. Mögliche Ursachen dieser Streuung müssen eingehend untersucht werden, um unser Verständnis über die Ähnlichkeiten und Unterschiede der Eigenschaften und der zugehörigen physikalischen Prozesse von Erdbebenquellen verbessern zu können.

Der Schwerpunkt dieser Doktorarbeit liegt auf kleinen Erdbeben mit besonderem Fokus auf induzierte Seismizität und akustischen Emissionen aus Laborversuchen. Es werden zwei Gruppen von Faktoren sowie dessen Auswirkungen auf die Berechnung seismischer Parameter untersucht, die als mögliche Ursachen für die schwankenden seismischen Spannungsabfälle gelten. Dabei werden folgende Faktoren unterschieden: 1) nicht-physikalische Faktoren umfassen die Fehlerfortpflanzung und ungeeignete Annahmen während der seismischen Datenanalyse, die folglich zu unerwünschten Verzerrungen der berechneten Erdbebenquellparameter führen können so wie auch 2) physikalische Faktoren, welche die tatsächlichen Eigenschaften von Erdbebenherden und Bruchvorgängen beschreiben und somit reale Unterschiede in seismischen Spannungsabfällen hervorrufen können. Innerhalb dieser Arbeit sind mögliche Effekte nicht-physikalischer Faktoren durch eine strenge Kontrolle der Datenqualität, die Anwendung angemessener Methoden sowie der sorgsamen Auswahl geeigneter Analyseparameter auf ein Minimum reduziert worden. In dieser Dissertation wird besonders der Effekt durch die Dämpfung hochfrequenter seismischer Wellen innerhalb der seismischen Datenanalyse näher betrachtet. Die Dämpfung seismischer Wellen wird durch den Gütefaktor Q beschrieben, welcher die meist unbekannte Beschaffenheit des Untergrundgesteins widerspiegelt. In der seismischen Analyse wird Q für die Korrektur des seismischen Amplitudenspek-

trums und den daraus resultierenden Berechnungen der Erdbebenquellparameter genutzt. Ungenaue Annahmen in Bezug auf die Dämpfung seismischer Wellen können den hochfrequenten Bereich des Amplitudenspektrums verzerren und somit zu einer erhöhten Fehlerfortpflanzung (nicht-physikalischer Effekt) in der weiteren Quellparameteranalyse führen.

In der ersten Studie dieser Doktorarbeit wird der zeitliche Amplitudenabfall von S-Coda-Wellen induzierter Seismizität im kalifornischen Geothermalfeld *The Geysers* analysiert, um die Dämpfungseffekte hochfrequenter seismischer Signale besser einschätzen zu können. Die *Schiebefenstermethode* ermöglicht hierbei innerhalb des Frequenzbereiches die Abschätzung der Coda-Gütefaktoren (Q_C) für ausgewählte Frequenzbänder. Vorab wird eine zusätzliche Sensitivitätsanalyse durchgeführt, um eine mögliche Beeinflussung der Analyseparameter innerhalb der angewandten Methode abschätzen zu können. Überprüft werden die betrachtete Magnitudenspanne, das Signal-Rausch-Verhältnis, der Einfluss der Schwankung der finalen Q_C-Werte, die Größe der Schiebefenster, die Auswahl des zeitlichen Analysestartpunktes der Coda im Seismogramm und die Länge des untersuchten Coda-Ausschnittes. Die Sensitivitätsanalyse zeigt, dass eine geeignete Auswahl der Parametrisierung sowie die Datenqualität in der angewandten Coda-Analyse von Relevanz sind. Für die Coda-Analyse werden Wellenformen von insgesamt 717 flachen Erdbeben mit Magnituden zwischen 1-3 von zwei räumlich voneinander getrennten Lokationen im nordwestlichen und südöstlichen Bereich des Geothermalfeldes untersucht. Im Vergleich weisen beide Lokationen deutliche Unterschiede der Dämpfungseffekte auf, welche auf die variierenden geologischen-, strukturellen- und geothermalen Eigenschaften innerhalb des Untersuchungsgebietes zurückzuführen sind. Der Gütefaktor Q_C wird dabei bis zu einer Frequenz von 70 Hz erzielt und übertrifft somit den gängig betrachteten niedrigeren Frequenzbereich seismischer Felddaten. Der Gütefaktor Q_C des nordwestlichen Datensatzes wird zusätzlich im Kontext zeitlich variierender Reservoir-Injektionen analysiert. Die zeitliche Stabilität von Q_C deutet darauf hin, dass das Störungsnetzwerk im Untergrund entweder zeitlich konstant ist oder dass die Änderungen im Störungsnetzwerk, bedingt durch die Reservoir-Injektionen, nicht anhand der hier betrachteten Frequenzen auflösbar ist. Zusätzlich wird der Gütefaktor Q_C desselben Datensatzes mit dem Gütefaktor der direkten S-Welle (Q_β) verglichen. Die Coda-Analyse zeigt hierbei eine höhere Messstabilität als der üblicherweise berechnete Gütefaktor basierend auf der direkten S-Welle. Dies deutet somit darauf hin, dass die Coda-Analyse zu einer höheren Stabilität weiterer berechneter Erdbebenquellparameter, wie beispielsweise dem seismischen Stressabfall, führen kann.

Der zweite Teil dieser Doktorarbeit befasst sich mit akustischen Emissionen aus zwei triaxialen "Stick-Slip" Laborexperimenten an ofengetrockneten Westerly-Granit

Proben. Erdbebenquellparameter, im Speziellen der statische Spannungsabfall, von akustischen Emissionen eines rauen- und eines ebenen Bruches werden mit Hilfe einer Variante der *Spektralverhältnis-Methode* basierend auf dem Konzept einer multiplen empirischen Greenschen Funktion (multi-eGf) analysiert. Dabei werden Dämpfungseffekte (Pfad- und Standorteffekte) miteinander verlinkter, gleichlokalisierter Laborerdbeben supprimiert, sodass die Quellparameter direkt vom Quellspektrum unter Annahme des quasi-dynamischen *Madariaga-Modells* erfasst werden können. Zum ersten Mal werden dabei drei Größenordnungen von Magnituden (-9 < $M_W < -5.6$) innerhalb einer Laborstudie ausgewertet, was eine angemessene Basis für eine Skalierungsanalyse darstellt. Die berechneten Spannungsabfälle der akustischen Signale sind vergleichbar mit den globalen Beobachtungen innerhalb des gesamten Magnitudenspektrums. Dennoch stellt sich eine deutliche Abweichung von der allgemein angenommenen Skaleninvarianz innerhalb beider Experimente heraus, was die Selbstähnlichkeit über alle Erdbebenmagnituden in Frage stellt. In diesem Teil der Arbeit werden die physikalischen Aspekte, wie die Beschaffenheit der Bruch-oberfläche, der Erdbebenquellradius und die Bruchgeschwindigkeit untersucht. Der statische Spannungsabfall der akustischen Laborsignale weist dabei im Durchschnitt keine Abhängigkeit bezüglich der Quellgröße und nur eine geringe Variation mit Änderung der Bruchgeschwindigkeit auf. Die beobachtete Spannungsabfall-Magnituden-Skalierung lässt jedoch Unterschiede des Bruchversatzes über eine raue und eine ebene Bruchoberfläche vermuten. Die Komplexität des rauen Bruches könnte die Entstehung größerer Bruchversätze verhindern, welches die niedrigeren Spannungsabfälle erklären könnte. Der ebene Bruch hingegen kann möglicherweise durch seine geringere Komplexität größere Scherversätze generieren und somit größere Spannungsabfälle verursachen.

Beide Studien bieten neue Einblicke in die Analyse hochfrequenter seismischer Signale. Dementsprechend werden weitere Vergleichsstudien basierend auf ähnlichen Grundlagen für eine genauere skalenübergreifende Untersuchung von Erdbeben empfohlen. Der Vergleich von Studien wird jedoch häufig durch eine Vielzahl unterschiedlich angewandter Methoden und Grundlagenmodellen erschwert. Daher ist der Austausch von Daten- und interdisziplinärem Fachwissen, Datenstandardisierung, eine sorgfältige Auswertung der Datenqualität und der öffentliche Zugang methodischer Anwendungen essentiell für die Verbesserung zukünftiger Forschungsansätze. Der dritte Teil dieser Dissertation basiert auf diesen aufgelisteten notwendigen Ansprüchen zukünftiger Analysen und soll als Ergänzung zu den ersten beiden wissenschaftlichen Studien gesehen werden. Im dritten Teil wird die innovative Entwicklung einer Online-Forschungsplattform vorgestellt, welche durch die TCS AH (thematische Kerndienstleistung basierend auf anthropogenen Risiken) innerhalb der Implementierungsphase des

Europäischen Platten-Überwachungssystems (EPOS-IP) geleitet wird. Hierbei wird auf die Entwicklung einzelner Bestandteile der neuartigen Onlineplattform *IS-EPOS* (IS = Induzierte Seismizität) eingegangen, die auf einer großen Anzahl vollständiger und umfangreicher Datensätze bezogen auf Ressourcenausschöpfung, Ressourcenerforschung und Untergrundendlagerung von Gasen und Fluiden sowie zugehöriger Anwendungen basiert. Das EPOS Projekt gilt als Vorreitermodell für zukünftige Forschungsarbeiten, in denen Ähnlichkeiten und Unterschiede zwischen individuellen Studien besser erfasst und einfacher untersucht werden können. Dies soll langfristig zu einem besseren Verständnis der seismischen Daten und der zu Grunde liegenden seismischen Quellen führen.

Statement and list of publications

This doctoral thesis consists of three published articles in scientific journals and two related data publications. I am first author of the two scientific studies 1) and 2) that constitute the core of this dissertation. I am co-author of publication 3), which was developed in close collaborative work and describes an open data infrastructure of anthropogenic seismicity. Both additional data publications (Blanke et al., 2018, 2020) were prepared fully by myself and are freely accessible via the GFZ Data Services following the links in the listed references 1b) and 2b) below.

1a) **Blanke, A.**, G. Kwiatek, P. Martínez-Garzón, and M. Bohnhoff (2019). Sensitivity and stability analysis of coda quality factors at The Geysers Geothermal Field, California, ***Bulletin of the Seismological Society of America***, 109(3), 959-975, doi:10.1785/0120180219.

1b) **Blanke, A.**, G. Kwiatek, P. Martínez-Garzón, and M. Bohnhoff (2018). Supplementary material to "Sensitivity and stability analysis of coda quality factors at The Geysers geothermal field, California", V. 1.0 ***GFZ Data Services***, https://doi.org/10.5880/GFZ.4.2.2018.002 (data publication).

2a) **Blanke, A.**, G. Kwiatek, T. W. Goebel, M. Bohnhoff, and G. Dresen (2021). Stress drop–magnitude dependence of acoustic emissions during laboratory stick-slip, ***Geophysical Journal International***, 224, 1371-1380, doi:10.1093/gji/ggaa524.

2b) **Blanke, A.**, T. W. Goebel, and G. Kwiatek (2020). Acoustic emission source parameters of laboratory triaxial stick-slip experiments on two Westerly granite samples, ***GFZ Data Services***, https://doi.org/10.5880/GFZ.4.2.2020.008 (data publication).

3) Orlecka-Sikora, B., S. Lasocki, J. Kocot, T. Szepieniec, J.-R. Grasso, A. Garcia-Aristizabal, M. Schaming, P. Urban, G. Jones, I. Stimpson, S. Dineva, P. Sałek, K. Leptokaropoulus, G. Lizurek, D. Olszewska, J. Schmittbuhl, G. Kwiatek, **A. Blanke**, G. Saccorotti, ... and T. Fischer (2020). An open data infrastructure for the study of anthropogenic hazards linked to georesource exploitation, ***Scientific Data***, 7. 89, doi:10.1038/s41597-020-0429-3.

In publication 1), I functioned as main author. I performed the waveform data analysis, the sensitivity analysis and evaluated the results. I am responsible for the coordination and the writing of the manuscript and also the creation of all figures. I conducted the

data analysis based on a code provided by Grzegorz Kwiatek that I modified as part of my thesis. Figures were generated using MATLAB®, Generic Mapping Tool (GMT), and Corel Draw.

I am also the main author of publication 2). This article combines two fields of expertise that required a close collaborative evaluation and interpretation of the results. I am responsible for the main part of the writing, in particular the seismological aspects. Helpful support for the interpretation of geomechanical data was given by co-authors Thomas Goebel and Georg Dresen. I conducted the whole data analysis based on a code written by Grzegorz Kwiatek. Necessary modifications for the parameter calculations were added by myself as part of this thesis. Figures were generated and modified using MATLAB®, Generic Mapping Tool (GMT), POV-Ray and Corel Draw.

Publication 3) is the outcome of 5 years of collaborative work during the EPOS Implementation Phase (EPOS-IP) project under the auspices of the EU Research and Innovation programme Horizon 2020. Between 2015 and 2019 the whole team developed a scientific data platform related to anthropogenic hazards. This article presents the work that was done by all co-authors within the project period. I was actively involved in the project phase and actively contributed to the data gathering, data preparation, maintenance of communication with the data providers, data sets evaluation, data applications evaluation and the quality control of new provided data sets on the project website (https://tcs.ah-epos.eu/). My main responsibilities included the quality control of new data sets and the support and supervision of data providers. In addition, I was responsible for the adaption of the required data formats and the quality control the dataset "The Geysers Prati-9 and Prati-29 cluster" published on the project website.

Contents

1 Introduction

1.1 General Introduction

In antiquity, a shaking of the ground was thought to be an act of God or the result of battles between Gods and giants in the earth's interior. The occurrence of the devastating Lisbon earthquake in 1755 saw the first attempts at reconsidering this largely theological interpretation of sudden ground shaking. The Lisbon event is the largest reported historical earthquake (magnitude $\approx$ 8–9) in Europe, with effects felt also in many other parts of Europe, North Africa and even North America (see Fuchs, 2006). It astonished the whole world and evoked a debate on the underlying processes of earthquakes (e.g. Fuchs, 2006; Udías and Arroyo, 2009). Whereas many people continued to treat earthquakes as supernatural events caused by God, philosophers such as Immanuel Kant, but also other anonymous authors and theologians (see Udías and Arroyo, 2009) started to consider earthquakes as *natural phenomenon*. For the first time, surveys were carried out in order to assess the strength and consequences of the Lisbon event. This is considered as the "birth of modern seismology" and disaster management (e.g. Fuchs, 2006) and similar approaches of quickly gathering information on earthquake effects are still in use today.

First formulations of the propagation of elastic waves, including body-and surface waves, through solid rocks were investigated and introduced in the nineteenth century by renowned scientists such as Stokes, Cauchy, Poisson and Rayleigh (e.g. Shearer, 2009). We have only been fully aware of the *natural* restlessness of the earth and the origins of earthquakes since the early twentieth century after the memorable violent San Francisco earthquake of 1906. Henry Fielding Reid analyzed the deformation and displacements from geodetic data, together with observations from previous local earthquakes and additional deformation experiments and thus developed the theory of elastic rebound (e.g. Reid, 1910; Zoback, 2006). Only since that time did the terminology *natural disaster* become established in science and, later, more widely in the broader society (e.g. Meier, 2007).

The 1906 San Francisco earthquake helped to pioneer systematic studies on earthquakes in relation to tectonic forces. Around the same time, in 1915, Alfred Wegener provided evidence for the theory of continental drift, which only found broad acceptance later in the 1960's (e.g. Wegener, 1966). Thanks to Reid and Wegener, we now know that the brittle crust of the earth is subject to stress build-up, and that rupture failure occurs along faults. If the accumulated shear stress exceeds the local strength of the rocks, the fault sides move relative to each other. The accumulated energy is, in addition to inelastic

processes, released in propagating elastic waves that set the earth into vibration (e.g. Kanamori and Brodsky, 2001; Shearer, 2009). That violent process of mass movement and associated elastic wave propagation we call a tectonic earthquake.

Historical and recent large earthquakes have shown us that earthquakes are serious natural disasters that can cause devastating consequences including the destruction of buildings and infrastructure. These in turn evoke secondary hazards such as tsunamis, landslides, fires causing deaths. To date, we cannot make any deterministic predictions on *when*, *where* and with *which strength* an earthquake will occur. Probabilistic earthquake forecasting can only be made for very well-studied areas. Therefore, there is a pressing need to identify further factors and parameters, to access historical recorded data, and to conduct long-term observations.

Nonetheless, our knowledge about the physics of earthquake nucleation and propagation has improved remarkably over the recent decades as a result of the development of seismic sensors and new recording techniques, the improved quality of datasets, enhanced computational methods, and the application and analysis of laboratory experiments and man-made earthquakes (i.e. induced seismicity). To limit social, financial and material losses, ground motion assessment is crucial for any inhabited area. Ground motion prediction equations are based on, *inter alia*, seismic source parameters such as magnitude (e.g. moment magnitude M_W) or seismic stress drop σ (e.g. Cotton et al., 2013) that affects the near-field ground motions (e.g. Spottiswoode, 1993). The growth of seismic databases enhances the evaluation of possible seismic hazards and risks, whereas continuous, real-time recording of time- and space variations of seismic activity is indispensable for the improvement of early warning systems and strategies (e.g. Picozzi et al., 2017).

Unfortunately, earthquake source processes that occur inside the earth are not easily traceable. Most of the devastating earthquakes we record at the surface occur within several kilometers of depth without direct access. Analysis of passive seismic measurements combined with additional geophysical and geodetic research information can shed light on the origin and the extent of an earthquake. However, many uncertainties remain, especially those related to the properties of materials in the heterogeneous earth through which the transmitted seismic waves travel (e.g. path- and site effects). It is essential to reduce these uncertainties as much as possible to evaluate potential hazards and risks at the earth's surface.

1.2 Motivation

For a realistic hazard and risk assessment, it is crucial to understand physical source processes. Material properties, stress heterogeneities as well as fault sizes and fracture

density in the vicinity around the source play an important role and can enhance or lower the risk of damage at the surface. Because we are usually unable to drill into the earthquake source area, effort is primarily directed toward modeling, investigating induced earthquakes, and designing laboratory experiments to unravel the physical processes inside faults and to relate them to seismic hazard.

A long-lasting and still ongoing debate concerns the scale invariance of small and large earthquakes (e.g. Aki, 1967; Mayeda and Walter, 1996; Ide and Beroza, 2001; Kanamori and Rivera, 2004; Prieto et al., 2004; Abercrombie and Rice, 2005; Baltay et al., 2011; Malagnini et al., 2014). It is commonly assumed that earthquakes display self-similar behavior across a broad magnitude range (e.g. Ide et al., 2003; Imanishi and Ellsworth, 2006). This is expressed by scale-invariance of source parameters, such as static stress drop (e.g. Fig. 1.1) that is often used to correlate earthquake source scales (e.g. Cotton et al., 2013; Bohnhoff et al., 2016a).

However, the interpretation of earthquake constant-stress-drop-scaling in the literature is often not straight forward (e.g. Cocco et al., 2016). In general, earthquake self-similarity assumes that independent of the size, earthquakes are controlled by the same source physics. This suggests that knowledge gained from the analysis of small earthquakes can be scaled up (e.g. Walter et al., 2006) which could help understanding the processes of large earthquakes and associated seismic hazard. To date, there has been little research into magnitude ranges above $M_W > 7$ and below $M_W < -4$. Large events usually occur more rarely (due to long re-occurrence rates of > 100 years) as shown in the Gutenberg-Richter relation (Gutenberg and Richter, 1944). The ensuring lack of sufficient data make these more difficult to evaluate, especially for a single study area. By contrast, laboratory experiments can produce a large number of events, so-called acoustic emissions, within a short period of time. The difficulties are related to struggles with the technical requirements like the calibration of the acoustic sensors under high pressure, the control over laboratory boundary conditions, and the exact repetition of experiments necessary for the reliable analysis of acoustic emission data. It is also not known whether laboratory events can be directly scaled up or if they differ from larger events.

Nevertheless, previous studies provided evidence that several characteristics of larger scale earthquakes can be reproduced in laboratory experiments (e.g. precursor and aftershock event occurrence, stick-sip mechanisms and b-value behavior), making them a powerful tool to investigate fracture processes under laboratory conditions that are parameter-controlled to a certain degree (see Mogi, 1962; Scholz, 1968; Lockner, 1993; Thompson et al., 2009; Kwiatek et al., 2014a; Goebel et al., 2017). Moreover, laboratory experiments have been used to test requirements and applications for the exploration and

exploitation of georesources such as the production of geothermal energy or hydrocarbon (e.g. Villeneuve et al., 2018) and to evaluate suitable deposits for e.g. CO_2 storage or nuclear waste disposal (e.g. Yasuhara et al., 2011; Othman et al., 2018).

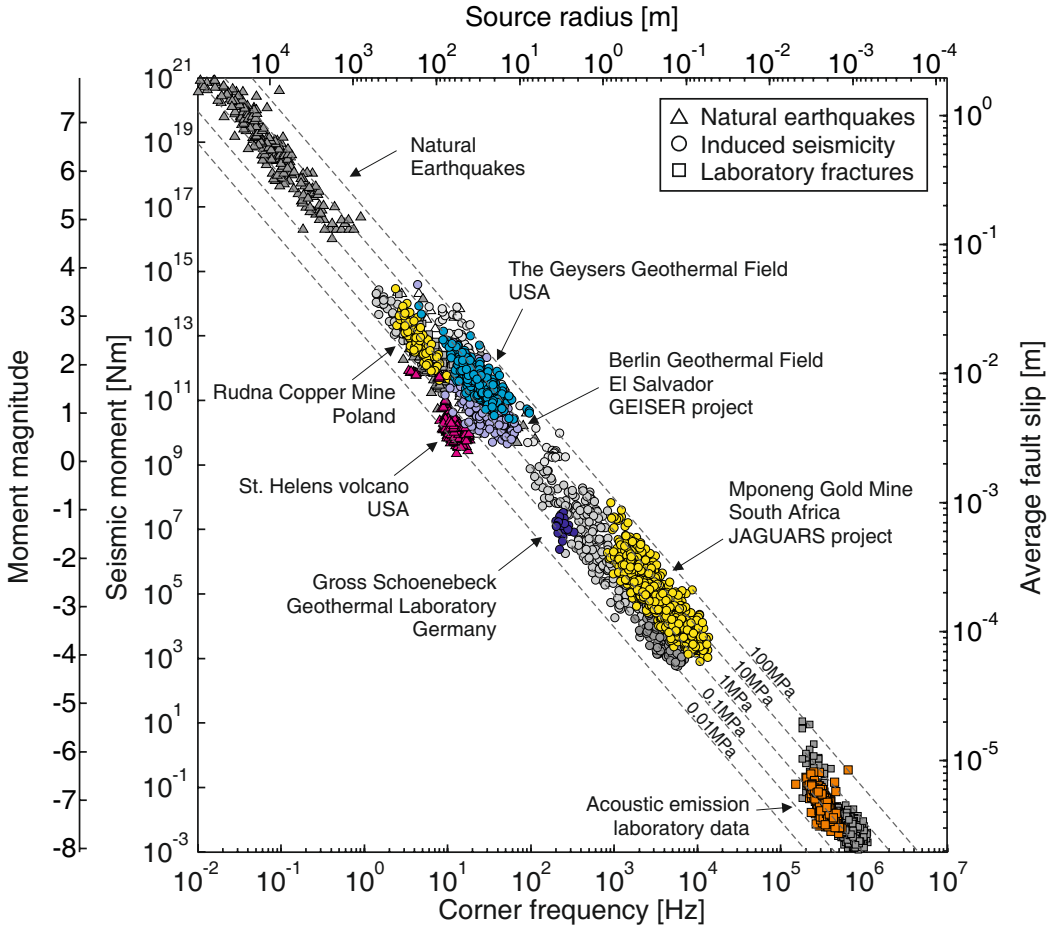

Figure 1.1: Coseismic stress drop plotted with key earthquake source parameters over the entire bandwidth of observed rupture processes extending from acoustic emissions in the laboratory to large natural earthquakes (modified from Bohnhoff et al. (2016a)). Individual datasets shown here are natural earthquakes (see Leonard, 2010), induced seismicity in mines and reservoirs (Kwiatek et al., 2011, 2015), volcano seismicity (Harrington et al., 2015). Acoustic emission data are unpublished results from the rock-deformation laboratory in GFZ-Section 4.2.

What is clearly seen when comparing earthquake source parameters from different studies (e.g. Fig. 1.1), is a global variation of stress drop between 0.01 and 100 MPa (e.g. Cocco et al., 2016). This variability cannot be associated with specific magnitude ranges and is observed over a broad span of earthquake sizes. On average, stress drop variation globally is not dependent on earthquake size and therefore supports the assumption of scale-invariance. Nonetheless, the variability covers five orders of stress drop magnitude.

This significant observed variability means that scaling up is not recommended. The large scatter of stress drop values might lead to under-or overestimated ground motions at larger scale range since we do not yet know which factors are responsible for the observed stress drop variation.

Closer inspection reveals that *single* datasets, when considered individually, sometimes also feature a significant stress drop variation and may even show a dependence on magnitude (M_W–σ relation). This increase of stress drop with growing magnitude for a *single* dataset is then interpreted as scaling breakdown (e.g. Mayeda and Walter, 1996; Cocco et al., 2016; Imanishi and Uchide, 2017; Baltay et al., 2019; Wang et al., 2019). This shows that the term *scale-invariance* is being applied rather vaguely in earthquake science; its use mainly depends on the study's point of view. What remains omnipresent is, however, the stress drop variability, which complicates the accurate prediction of ground motion and related potential hazards and risks (e.g. Cotton et al., 2013). Therefore, the fundamental question that is of particular interest in the context of earthquake scaling relations is: *What causes the variability of about five orders of stress drop magnitude observed for a broad span of earthquake sizes?*

To address this question, two main categories that affect stress drop in different ways can be scrutinized: 1) non-physical factors and 2) physical factors. Considerable effort has been put into the investigation of non-physical factors that can influence stress drop estimates (e.g. Ide and Beroza, 2001; Ide et al., 2003; Allmann and Shearer, 2009; Abercrombie, 2013). Non-physical factors cover observational uncertainties (e.g. Kaneko and Shearer, 2015), which can be e.g. poor station coverage, wrong corrections of site- and path effects, low signal-to-noise ratio (SNR), bandwidth limitations, or inaccurate source model assumptions. By repeating single studies but correcting inaccurate presumptions, applying different analysis methods, or removing of potential error sources (e.g. bandwidth limitation and applied methodology), it has been shown that effects of non-physical factors in the data analysis can bias the estimates leading to apparent scaling breakdown (e.g. Ide et al., 2003; Abercrombie, 2013).

Physical factors reflect the individual characteristics of earthquake sources. These include, e.g., differences of fault surface roughness, rupture- and slip variations, rupture type, directivity effects, or fault geometry (e.g. Sagy et al., 2007; Candela et al., 2011). These physical factors can vary between individual earthquake sources and may cause scatter and deviations between earthquake source studies that cannot be refuted by reviewing only the non-physical aspects (e.g. Mayeda and Walter, 1996; Malagnini et al., 2014). Differences in physical source properties could play an important role in the stress drop variation of individual earthquakes within an individual study site leading to the observed stress drop fluctuation within a dataset that is on average $\sim$ 2 orders of stress

drop magnitude (see Fig. 1.1 and Abercrombie and Rice, 2005; Abercrombie, 2015).

The uncertainty of whether stress drop estimates reflect actual physical processes, or may be a result of various artificial factors complicates the discussion on self-similarity and the study on physical earthquake source processes (e.g. Abercrombie et al., 2016; Cocco et al., 2016). This thesis investigates both non-physical and physical factors and their effects on source parameter estimates. Critical non-physical effects on seismic wave recordings, such as inaccurate correction of wave attenuation, and critical physical factors, such as the influence of fault surface roughness on static stress drop, were analyzed in detail. A high-resolution catalog for induced seismicity from The Geysers geothermal field, California, and two datasets of laboratory stick-slip experiments on Westerly granite serve as the research base for the main purposes of this study.

The quality factor Q can be seen as a physical factor that describes the quality (heterogeneity) of the medium with respect to structural defects and rock composition. In general, sub-surface structures and rock properties are unknown, which complicates the determination of Q (e.g Ide et al., 2003; Abercrombie and Rice, 2005; Sonley et al., 2006). These unknown properties of the medium and structural changes influence seismic waves on their way from the seismic source to the receiver. Therefore, Q has also been used to correct attenuation effects, such as intrinsic- and scattering attenuation, and to determine, typically in the amplitude-frequency domain, seismic source parameters (e.g. Ide et al., 2003; Kwiatek and Ben-Zion, 2016; Eaton, 2011; Wang et al., 2019). Wrong assumptions of the correction factor Q, can lead to bias in estimates of static stress drop and other source parameters (e.g. Ide et al., 2003).

This study demonstrates the importance of the selection of analysis parameters and knowledge about structures and faults specifically when covering a broader range of earthquake magnitudes. Especially when entering the high-frequency domain ($f > 1$ Hz), a generalized Q cannot ensure that high-frequency signals are sufficiently corrected to enable further source analysis. This can result in bias in further source parameter estimates indirectly dependent on Q (e.g. Ide et al., 2003). In this thesis, the analysis of the coda quality factor (Q_C) of induced seismicity covers a magnitude span of $1 \leq M_D \leq 3$ including high frequencies. The study's outcome contributes to a better understanding of the effects of attenuation on the analysis of small earthquakes that occur in areas characterized by different heterogeneity. It further shows that unknown attenuation effects on high-frequency waves in seismic analysis, dependent on the quality of the sub-surface media, may increase the uncertainties of resulting source parameter estimates, which creates a challenge for accurate hazard assessments.

Furthermore, the in this study presented laboratory datasets enable the analysis of the lower bound of earthquake scaling relations covering a wide moment magnitude

span ($-9 \leq M_W \leq -5.6$) that significantly extends previous studies of acoustic emission (AE) source parameters. A spectral ratio based technique was applied, which allows the suppression of undesirable effects from a number of non-physical parameters. Here, the main focus is set to physical aspects of stress drop variability that are evaluated together with mechanical and rock-mechanical data usually not available in field analysis. The AE stress drop estimates were studied in the context of the influence of fault heterogeneity (fault surface roughness, damage zone width), rupture velocity, source radius- and slip size. The results contribute to the debate on self-similarity at laboratory magnitude scale, about which there still exists relatively little research.

The physical analysis carried out in this thesis, and in general, require careful data processing and datasets of high quality. The development and refinement of methodologies to solve physical problems are also often time-consuming and challenging. Study results aiming to provide new insights into basic physical aspects, which should be applicable to the whole range of earthquake sizes, therefore need to be made available. Additionally, as the main part of this thesis (chapters 3 and 4) will show, single studies with restricted focus on a scientific topic cannot be evaluated, interpreted, or compared without knowledge of further complementary research information (e.g. seismic, mechanical, geological, structural, industrial, etc.). Exactly these issues constitute the core of the work of EPOS-IP (European Plate Observing System–Implementation Phase). The joint project consists of several thematic core services specialized on various topics of earth science aiming to create a platform available for the scientific-, academic, and industrial community to facilitate research in Europe. The TCS-AH (Thematic Core Service–Anthropogenic Hazards) team, a section of EPOS-IP, established an online research infrastructure including a high number of datasets related to anthropogenic hazards. Interdisciplinary data is gathered, standardized, quality controlled, and complemented by specifically developed applications to combine knowledge and to improve future hazard and risk assessments. The collaborative publication Orlecka-Sikora et al. (2020) demonstrates how to create such a data base and complements the empirical essence of this thesis.

1.3 Outline

The thesis focuses on two research articles (Blanke et al., 2019, 2021) that are complemented by a third publication (Orlecka-Sikora et al., 2020). Supplementary published material such as tables and figures were incorporated into the corresponding main chapters to better fit to the presentation of publications in this dissertation. Additional data publications to chapters 3 (Blanke et al., 2018) and 4 (Blanke et al., 2020) can be accessed via the GFZ Data Services following the provided links in Statement of Contributions.

Chapter 2 gives an overview of the theory and methodology relevant for the studies presented in chapter 3 and 4. It introduces seismic wave attenuation with focus on the S-wave coda quality factor. Furthermore, it presents the relation of the earthquake source parameters utilized in this study and the applied analysis methods.

In **chapter 3**, the scattering attenuation of coda S-waves (Q_C) is analyzed in particular by using high-quality seismic records from induced seismicity at The Geysers geothermal field, California. Incorrect assumptions of the quality factor can lead to inaccurate and imprecise source parameter estimates which in turn leads to under- or overestimated hazard and risk assessments. The moving window method of Phillips (1985) is applied to investigate the stability of Q_C estimates in comparison to direct S-wave Q_{beta} estimated by Kwiatek et al. (2015). A sensitivity analysis of applied parameters is conducted to exclude potential influences on the results. The study outcome is evaluated in the context of temporal injection variations and spatial differences in recording- and earthquake origin sites within the geothermal study area.

In **chapter 4** laboratory acoustic emissions from two stick-slip experiments on Westerly Granite samples of different fault surface roughness are analyzed. A variant of the spectral ratio method including a multi-eGf linkage of event pairs (e.g. Kwiatek et al., 2011, 2014a, 2015; Harrington et al., 2015) is used to determine relative corner frequencies, seismic moments and static stress drops. A Simulated Annealing algorithm based on the non-stationary Metropolis-Hastings Random Walk algorithm (e.g. Sen and Stoffa, 1995) is applied to solve the multidimensional inversion problem. Non-physical factors and their effects on estimated source parameters are discussed and possible influences of physical factors on static stress drop are analyzed in detail.

Chapter 5 presents the work carried out by the Thematic Core Service Anthropogenic Hazards (TCS AH) of the long-term project EPOS-IP (European Plate Observing System - Implementation Phase). The outcome is based on the collaborative effort of different work packages over several years that aimed to build up a new online infrastructure that is relevant and accessible for the scientific- and academic communities and the industrial partners. Comprehensive high-quality datasets covering different impacting factors related to exploration and exploitation of georesources were gathered, standardized and their quality was checked. By combining research related and industrial data and by making the data easily accessible, this project contributes to an advanced understanding and assessment of potential anthropogenic hazards and risks within different environments. This new infrastructure has the overarching goal of combining publications from different areas of expertise to facilitate hazard-related research in Europe.

Chapter 6 summarizes and discusses the key outcomes of chapters 3-5. An outlook including suggestions for further research concludes this thesis.

2 Theory

The lithosphere of the earth is in contrary to many common model assumptions heterogeneous and not perfectly elastic (e.g. Lay and Wallace, 1995). Inhomogeneities of various length scales exist. As long as they are of larger scale (e.g. subduction zones, mantle plumes, etc.) than the dominant seismic wavelengths, it is possible to detect them with classical methods like, for instance, seismic travel time analysis and seismic tomography (e.g. Lay and Wallace, 1995). It becomes difficult when inhomogeneities become smaller than typical dominant wavelengths. Different properties of rocks, temperature changes, density- and velocity anomalies, varying stress field orientations at e.g. volcanoes, cracks and local faults are only some parameters that build up the complexity of the earth's crust (e.g. Sato, 1978). Especially small-scale obstacles in the earth's crust can have strong influence on radiated seismic waves and contribute to local variations in ground shaking.

2.1 Seismic Attenuation

In seismic analysis, the impacts on propagating waves are described as path- and site effects. From seismic recordings, it is difficult to distinguish between elastic and inelastic processes that cause seismic wave amplitudes to decay and broaden with time and distance from the source. Seismologists refer to them as attenuation effects resulting from interactions of waves with different (usually unknown) structures and properties of the propagation medium. Due to this unknown complexity of the subsurface, it is also challenging to separate the seismic source term from attenuation effects that bias seismic records (e.g. Shearer, 2009).

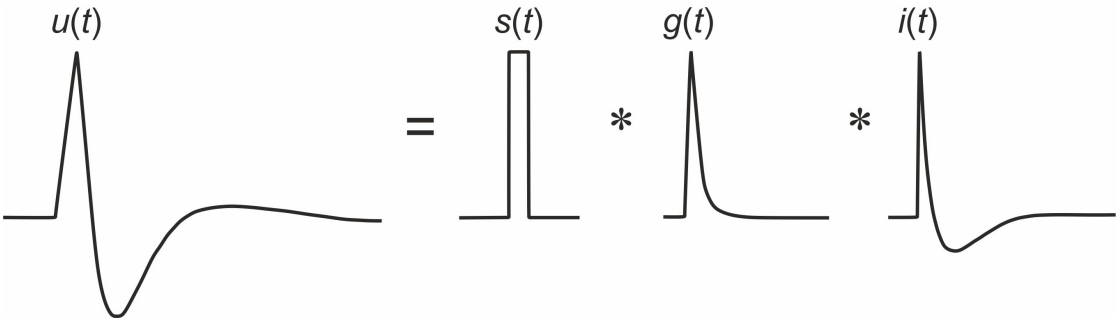

Figure 2.1: A simple approximation of a recorded seismic signal $u(t)$ in a seismogram (modified from Stein and Wysession (2003)). It can be described by the convolution of three terms, which are the seismic source $s(t)$, effects of the earth's structure $g(t)$ that influence propagating seismic waves, and the response of the instrument $i(t)$.

Observed ground motions $u(t)$ recorded in a seismogram can be expressed as:

$$u(t) = s(t) * g(t) * i(t), \tag{2.1}$$

where the source term s is convolved ($*$) with the elastic Green's function g, which describes site- and path effects (including attenuation), and with the instrument response i (e.g. Lay and Wallace, 1995; Stein and Wysession, 2003). In the frequency domain, equation (2.1) can be written as the product of the transfer functions (e.g. Stein and Wysession, 2003):

$$U(f) = S(f)G(f)I(f). \tag{2.2}$$

Attenuation is inversely proportional to the dimensionless quality factor Q that describes the energy loss (ΔE) of the seismic wave per oscillation cycle (e.g. Lay and Wallace, 1995; Shearer, 2009):

$$Q = \frac{2\pi E}{\Delta E}. \tag{2.3}$$

Here, E is the initial total energy. A small Q indicates high attenuation (large energy loss and fast amplitude decay) whereas a large Q implies low attenuation (low energy loss and slow amplitude decay) (see Fig. 2.2). Travel time variations are therefore also reflected in Q because seismic waves in highly attenuating areas display lower propagation velocities (e.g. Lay and Wallace, 1995). The quality factor Q is also assumed to change with depth in the earth, showing smaller values in the heterogeneous uppermost crust and larger values in the less heterogeneous lower crust (e.g. Shearer, 2009). Furthermore, Q is sensitive to temperature changes (e.g. Lay and Wallace, 1995) in the earth with larger values for colder and smaller values for heated regions (e.g. geothermal fields or tectonically active regions).

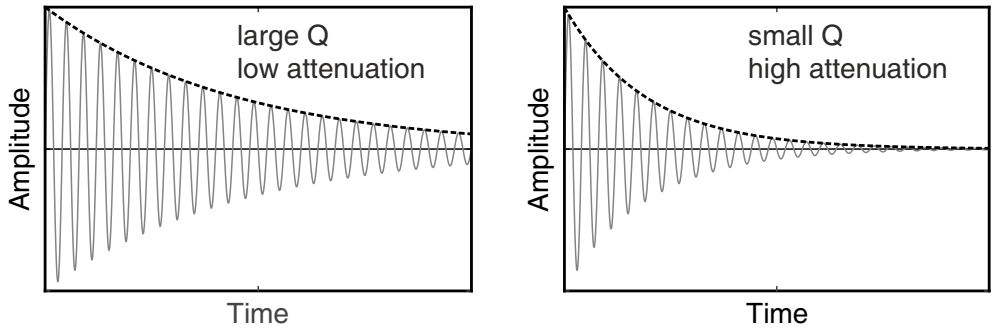

Figure 2.2: Schematic Illustration of a harmonic sinus function with envelope (dashed line) indicating the amplitude decay with time for large Q (left) and small Q (right).

The amplitude (A) decay with time (t) of harmonic waves can be written as:

$$A(t) = A_0 e^{\frac{-\omega_0 t}{2Q}}, \tag{2.4}$$

with A_0 as the initial amplitude and $\omega_0 = 2\pi f$ as the angular frequency (e.g. Lay and Wallace, 1995). Here, Q is assumed to decay as in eq. (2.3). The decay with distance (x) can be expressed as:

$$A(x) = A_0\, e^{\frac{-\omega x}{2QV}}, \tag{2.5}$$

where V is either the P-wave (V_α) or S-wave velocity (V_β) with the respective quality factor Q_α or Q_β (e.g. Lay and Wallace, 1995; Shearer, 2009).

Geometrical spreading describes the decrease of wave energy per unit area due to the spherical expanding wave front with distance (R) to the source (e.g. Stein and Wysession, 2003). For body waves, the amplitude decreases as $1/R$, which is stronger than for surface waves ($1/\sqrt{R}$). The two most interesting seismic attenuation processes, however, are intrinsic and scattering attenuation (Sato et al., 2002). The total quality factor (Q_t) is very often assumed to be composed of both (e.g. Dainty and Toksöz, 1981) and is given as:

$$Q_t^{-1} = Q_i^{-1} + Q_{sc}^{-1}. \tag{2.6}$$

Intrinsic attenuation Q_i is the inelastic loss of energy that reduces wave amplitudes and changes pulse shapes due to the conversion of elastic energy into heat (e.g. Lay and Wallace, 1995; Stein and Wysession, 2003). The driving force of intrinsic attenuation is internal friction (e.g. Lay and Wallace, 1995), which arises along mineral dislocations, at grain boundaries, and due to water or gas filled pores and cracks inside of rocks. Intrinsic attenuation is presumed to be the dominant process for damping of direct waves (P- and S-waves) at short distances as they travel along the simplest pathway and are often assumed to have wavelengths larger than the crustal inhomogeneities (e.g. Gibowicz and Kijko, 1994; Lay and Wallace, 1995). Especially for low frequencies 0.001 Hz $< f <$ 1 Hz, inelastic energy loss is assumed to be frequency independent (e.g. Lay and Wallace, 1995; Shearer, 2009).

Scattering attenuation (Q_sc) is an elastic process where the energy of direct waves is redistributed and partially shifted to later arriving wave portions, which decreases the amplitude of phases (e.g. Lay and Wallace, 1995; Sato et al., 2012). These later arrivals are referred to as coda waves (cf. Fig. 2.3) that follow the primary wave records and build the "tail" in the seismogram. Scattering of waves result from diffraction, reflection and refraction processes due to interaction of seismic waves with small-scale inhomogeneities (also called scatterers) such as cracks and faults (e.g. Aki, 1969; Lay and Wallace, 1995; Sato et al., 2012). These scatterers are most dominant in the upper crust and affect in particular high-frequency waves due to their shorter wavelengths (e.g. Stein and Wysession, 2003). With later coda portions, a larger volume of the medium between source and receiver is sampled (see Fig. 2.3). Thus, the coda is assumed to

average path- and source (radiation pattern) effects, leading to extremely stable source parameter estimates (e.g. Mayeda and Walter, 1996).

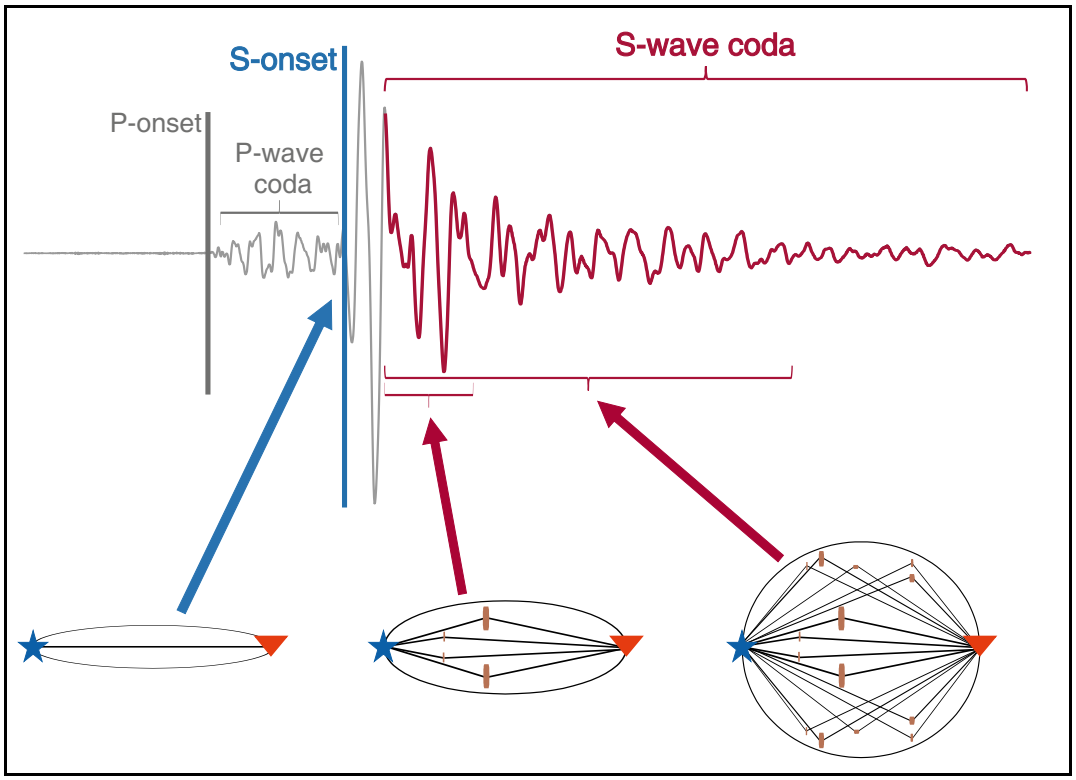

Figure 2.3: Seismic waveform record at local distance (R « 100 km) with highlighted P- and S-wave onsets and P-and S-wave codas. Simplified sketches at the bottom visualize the increasing sampled volume between source (star) and receiver (triangle) with time. S-waves in the seismic coda are generated due to scattering processes at randomly but homogeneously distributed scatterers (short bars).

2.1.1 S-wave coda quality factor (Q_C)

The analysis of the seismic coda is a popular method to approximate attenuation effects on high-frequency waves and to assess the small-scale complexity of the crust. Without inhomogeneities, no coda waves would be excited and recorded in seismograms (e.g. Sato et al., 2012). Therefore, the record of coda waves is an indirect proof of the heterogeneity of the earth's crust. A seismic coda can be recorded after the direct P- and the S-waves (cf. Fig. 2.3). In this thesis, I only concentrate on the S-wave coda and corresponding coda quality factor Q_C for earthquakes recorded at local distances $R < 20$ km.

The S-coda follows the direct S-wave in the seismogram and is assumed to consist mostly of redirected S-waves scattered at various and randomly but uniformly distributed small-scale inhomogeneities (e.g. Aki and Chouet, 1975). Observations made in previous studies show that envelopes of local seismic coda recorded at different stations in the

same seismogenic area are comparable after the elapsed time (lapse time) of twice the direct S-wave onset time (e.g. Rautian and Khalturin, 1978). Thus, the coda seems to be independent on the source-to-receiver distance and location (e.g. Phillips, 1985; Su et al., 1996), whereas the direct S-wave decays with epicentral distance (e.g. Rautian and Khalturin, 1978). The primary S-wave also strongly depends on focusing and defocusing effects in the source (e.g. Campillo et al., 1999). That makes measurements of direct S-wave amplitudes difficult in comparison to coda amplitudes, for which, in theory, source radiation pattern can be neglected (e.g. Aki, 1969; Aki and Chouet, 1975; Mayeda and Walter, 1996; Stein and Wysession, 2003). Taking the independence on epicentral distance into account, occurring irregularities in coda records can provide information on local site effects (e.g. Aki, 1969; Su et al., 1996). Thus, the seismic coda can serve as a proxy for spatial attenuation changes (e.g. Phillips et al., 1988). In addition, the seismic coda of local earthquakes can be used to estimate a stable local earthquake magnitude, known as duration magnitude M_D. Therefore, the total duration of a seismogram or the total coda duration is considered (e.g. Herrmann, 1975; Aki and Chouet, 1975; Mayeda, 1993). However, caution needs to be taken, when comparing coda studies with each other. Depending on the analyzed coda portion and related sampled volume of the underground, Q_C is dependent on lapse time (e.g. Rautian and Khalturin, 1978).

It has been suggested that the seismic coda includes both scattering and intrinsic attenuation but it is very difficult to separate them from each other. Different assumptions were made trying to understand the impact of inelastic- and elastic processes influencing the seismic coda (e.g. Aki and Chouet, 1975; Aki, 1980b; Frankel and Wennerberg, 1987; Mayeda et al., 1992). The quality factor Q is often considered to be frequency independent and mainly intrinsic for low frequencies (e.g. Lay and Wallace, 1995; Shearer, 2009). However, a constant averaged Q is also frequently used in studies dealing with higher frequencies > 1Hz and solely with first direct phase arrivals (e.g. Ide et al., 2003) because the assessment of accurate attenuation for high frequencies is in general rather difficult (e.g. Abercrombie, 1995) and requires modeling. Depending on the model assumption, the coda quality factor Q_C is presumed to contain larger scattering processes for frequencies higher than 1 Hz, in particular when strong tectonic heterogeneities exist (e.g. Aki and Chouet, 1975; Aki, 1980b, 1981). Especially at shorter wavelengths (higher frequencies), Q_C shows a dependence on f. Scattering becomes very strong when scatterer (x) and wavelength (λ) are of similar size (e.g. Sato et al., 2012) which might lead to the observed frequency-dependence. Low-frequency waves are usually of larger scale than the scatterer sizes ($x << \lambda$) and thus rather unaffected by the small-scale obstacles in the earth's crust.

The frequency dependence can be described in form of a power law (e.g. Phillips et al.,

1988) that is used to compare quality factors between source studies:

$$Q_C = Q_0 f^n, \tag{2.7}$$

where Q_0 is the quality factor at f = 1 Hz and n describes the relation of attenuation and frequency (e.g. Aki, 1980b; Morozov et al., 2008; Mukhopadhyay et al., 2008; Havskov et al., 2016). The power law indicates that seismic records need to be bandpass-filtered to obtain quality factors for different wavelengths. The more heterogeneous a study area is, the stronger is the dependence of Q_C on f and the exponent n increases.

However, the coda quality factor is still not yet understood in detail (e.g. Ibanez et al., 1990; Calvet and Margerin, 2013) and different assumptions lead to different results of that we should be aware of, when comparing individual studies. For natural earthquakes with epicentral distances $\leq$ 100 km, Q_C (using body-waves) is commonly investigated in frequency ranges of approximately 1 – 30 Hz (e.g. Aki, 1980a; Phillips and Aki, 1986; Phillips et al., 1988; Novelo-Casanova and Lee, 1991; Gibowicz and Kijko, 1994; Mukhopadhyay et al., 2008; Calvet and Margerin, 2013). This thesis considers strictly selected high-quality seismic records at very short epicentral distances ($\leq$ 19 km), facilitating the Q_C analysis up to approximately 70 Hz for the closest sensors.

2.1.2 A brief overview of coda theory

The scattering theory is an approach to statistically resolve small-scale heterogeneities in the crust (Lay and Wallace, 1995). To explain wave scattering effects responsible for the observed amplitude decay of local seismic coda records, two primary models (e.g. Aki and Chouet, 1975) can be distinguished: 1) the single isotropic back-scattering model, and 2) the diffusion model. Depending on the model assumption, Q_C can be differently interpreted, which shows that the composition of coda waves is not yet well understood in detail.

The single isotropic back-scattering model is the simplest approach (e.g. Aki and Chouet, 1975; Sato et al., 2012) and is assumed to be valid for short lapse times (valid in the near field). This model assumes that radiated waves of an isotropic source are incoherently scattered only once on randomly but uniformly distributed point-like inhomogeneities in an otherwise homogeneous medium between source and receiver. P-waves and conversions between S- and P-waves are not considered (e.g. Aki, 1980a; Herrmann, 1975). This observation was supported by studies showing that Q_C and the quality factor of the direct S-wave (Q_β) are usually comparable (e.g. Aki, 1981). Therefore, the coda is presumed to consist only of S-to-S back-scattered waves (e.g. Aki, 1980b) traveling with the same constant velocity as the direct S-wave (e.g. Aki

and Chouet, 1975). Here, source and receiver are theoretically placed at the same location, allowing to only analyze the coda after twice the S-onset time. Further, the loss of energy by scattering of primary waves and multiple scattering are neglected in the single back-scattering theory, making this model a simple approach but violating the conservation of energy at the same time (e.g. Aki and Chouet, 1975). On the contrary to multiple-scattering approaches, this method is not able to distinguish between intrinsic and scattering attenuation (e.g. Frankel and Wennerberg, 1987) and consists of both (e.g. Aki, 1980b):

$$Q_C^{-1} = Q_t^{-1} = Q_i^{-1} + Q_{sc}^{-1}. \tag{2.8}$$

The diffusion model (Aki and Chouet, 1975) covers strong multiple-scattering effects and anisotropic scattering at randomly distributed heterogeneities in the earth. This allows to also consider P-waves and conversions between P- and S-waves, which is likely closer to reality. Wave diffusion requires much stronger scattering properties of the medium (e.g. Dainty and Toksöz, 1981) than single-scattering. The idea of strong scattering effects in the diffusion model by Aki and Chouet (1975) may not be straight forward due to the attenuation characteristics in the terrestrial crust that would prevent the required long scattered wave paths (e.g. Dainty and Toksöz, 1981). However, diffusion processes are assumed to be valid in the far field at long lapse times after the wavelets have been scattered multiple times or in highly inhomogeneous media. The application of the diffusion model based on strong multiple scattering in a homogeneous half-space assumes that Q_C primarily describes intrinsic attenuation (e.g. Shapiro et al., 2000):

$$Q_C = Q_i. \tag{2.9}$$

This is often confirmed by studies that use modified models based on multiple-scattering assumptions (e.g. Gao et al., 1983; Fehler et al., 1992). These approaches allow to separate Q_i and Q_sc by analyzing the coda amplitude decay (e.g. Frankel and Wennerberg, 1987; Fehler et al., 1992; Del Pezzo et al., 2001).

In general, the single back-scattering model by Aki and Chouet (1975) seems to be a good approach for coda studies with short epicentral distances. Nevertheless, multiple scattering effects are of importance at long lapse times. Both basic models were further developed within the last decades and many different new assumptions, such as double or triple scattering, anisotropic scatterers, random media, and others, were added and new methods were developed (e.g. Sato, 1977, 1982; Gao et al., 1983; Wu and Aki, 1985; Frankel and Wennerberg, 1987; Fehler et al., 1992; Wang et al., 2019). In spite of new models, scattering processes are still not well understood and it is still a matter of debate which model assumption should be applied leading to the most realistic interpretation of

coda Q in the earth's crust (e.g. Aki, 1980b,a; Frankel and Wennerberg, 1987; Calvet and Margerin, 2013).

Analyzing coda Q, is a preferred routine not only to assess attenuation effects but also source parameters. It was shown that using the seismic coda provides the most stable seismic source parameter estimates due to its characteristics (e.g. Aki, 1969; Mayeda and Walter, 1996; Mayeda et al., 2007; Yoo et al., 2011):

- no dependence on epicentral distances after twice the S-onset time
- insensitivity on the seismic source-radiation anisotropy
- also single station analysis provides stable estimates
- seismograms with clipped first arrivals can be analyzed using the unclipped coda portion
- the duration magnitude is not sensitive to crustal structures
- Q_C is similar to Q_β and can be a stable proxy of Q_β

Therefore, besides different coda models, a variety of methods exist to investigate seismic sources and different specific properties of coda waves (e.g. Aki, 1980b; Mayeda, 1993; Mayeda and Walter, 1996; Mayeda et al., 2007).

2.1.3 Coda decay analysis by Phillips (1985)

The analysis of seismic coda decay can be done in time and frequency domain. Seismic source analysis is preferably conducted in the frequency domain because spectral amplitudes are less affected by wave propagation effects (Madariaga, 2015). Following the single isotropic scattering theory, the power spectral density (PSD) of coda waves can be used and written as a function of time (e.g. Aki and Chouet, 1975; Phillips, 1985; Phillips et al., 1988):

$$P(f,t) = C(f)\,t^{-m} exp\left(\frac{-2\pi f t}{Q_C(f)}\right), \tag{2.10}$$

where $C(f)$ is assumed to be constant for short epicentral distances and describes the coda source factor at defined frequencies, t is the lapse time, and $m = 2$ represents the geometrical spreading of body wave amplitudes with time. The exponent, which includes the quality factor Q_C, describes the attenuation of S-coda waves with increasing time for center frequencies of defined frequency bands.

In this thesis, the extension of Sato (1977) is applied that allows sensor and receiver to move apart, making the single isotropic scattering model proposed by Aki and Chouet

(1975) more realistic:

$$P(f,t) = C(f)\,K(\alpha)\,exp\left(\frac{-2\pi ft}{Q_C(f)}\right), \tag{2.11}$$

with

$$K(\alpha) = \frac{1}{\alpha} ln\left(\frac{\alpha+1}{\alpha-1}\right) \tag{2.12}$$

and

$$\alpha = \frac{t}{t_S}. \tag{2.13}$$

Here, $K(\alpha)$ replaces the factor t^{-2} and α is the geometrical spreading factor that considers the S-wave travel time t_S. The dataset used in this study features very short coda durations. Thus, the required lapse time of two times the S-wave onset time in the single scattering theory cannot be respected. The extension of Sato (1977) allows to analyze the coda at shorter lapse times closer to the direct S-wave recording. However, the direct S-wave should not be included in the selected coda windows.

Following Phillips (1985), two steps are necessary to analyze the S-wave coda in terms of the frequency dependence of Q_C: 1) the moving window method, and 2) a regression analysis. A general scheme of the procedure steps applied in chapter 3 of this thesis is illustrated in Figure 2.4.

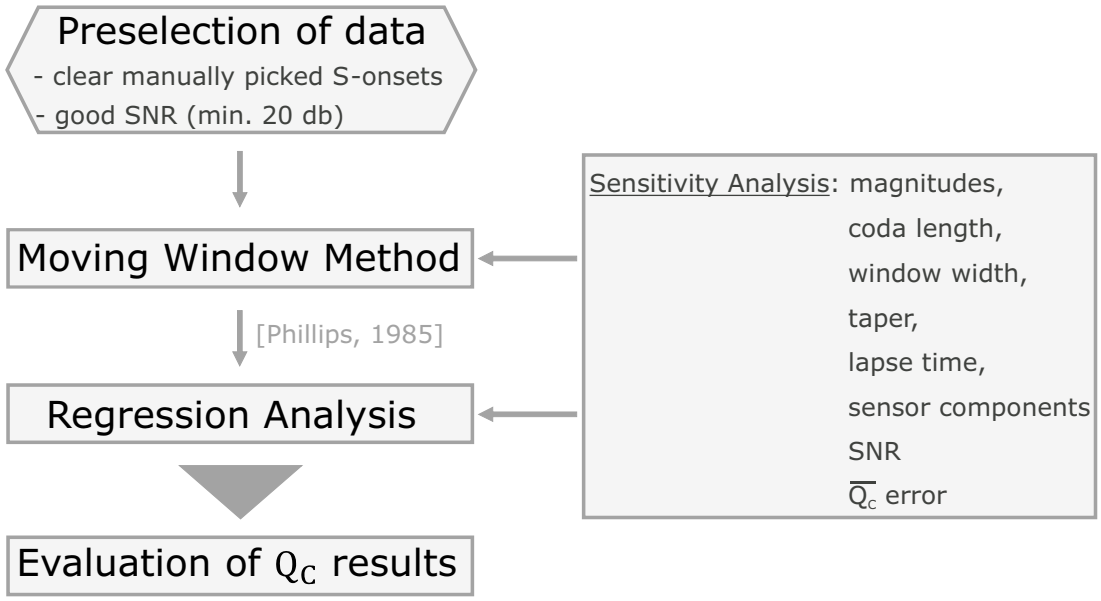

Figure 2.4: Schematic workflow of coda analysis using the approach of Phillips (1985). A sensitivity analysis was added in this thesis to evaluated potential influence of analysis parameters and data quality and will be described in detail in chapter 3.4.

Moving window method[1]

To guarantee a good SNR, a reference window of the noise is selected ahead of the seismic signal. The S-onset is manually picked to indicate the direct S-wave. Then, a lapse time is chosen that specifies the start of the coda analysis. The lapse time should exclude the direct S-wave and should start after the occurrence of largest amplitudes in the seismogram to not bias the coda decay measurement. In a next step, the moving window width is defined, which depends on the size of the earthquake. The window width needs to include some signal oscillations to represent the actual coda decay. The PSD (eq. 2.10) is calculated for both signal and reference noise windows for each frequency band and each moving window. The total coda length, which is the coda portion that is chosen to be investigated, depends mainly on the SNR. According to the power law (eq. 2.7), the seismogram needs to be filtered for narrow (typically octave-width) frequency bands. The moving windows defined above, move then in predefined step sizes through the filtered seismogram, determining the amplitude decay with time in the frequency domain. The moving windows always overlap 40 % to guarantee uncorrelated residuals (e.g. Phillips and Aki, 1986).

Regression analysis

The coda decay is now measured from the temporal decay of measured spectral amplitudes. Mathematically, this is done by taking the logarithm of both sides of eq. (2.11):

$$ln\left(\frac{P(f,t)}{K(\alpha)}\right) = ln(C(f)) - \left(\frac{2\pi f}{Q_C(f)}\right)t, \tag{2.14}$$

which defines a set of linear regression problems with dependent variable t for each considered center frequency f of a defined frequency band. Then, for each center frequency, the coda quality factor Q_C can be calculated from the slope (b) of measurements of the regressive relation of eq. (2.14):

$$Q_C(f) = \frac{2\pi f}{b}. \tag{2.15}$$

The whole procedure is conducted for each event at each station. For each single station an average Q_C for each center frequency is then calculated out of all analyzed earthquakes. The resulting mean-Q_C curves ($\overline{Q_C}$) can be compared and evaluated with respect to characteristics of the specific station site and the wave travel path (e.g. geology, known faults, injection rates, etc.) in the investigation area.

[1] Please see Figure 3.2a in chapter 3 for a better visualization.

2.2 Seismic Source Parameters

Since the 1960s, studies of earthquake sources have been conducted for the broad range of large natural events down to tiny acoustic emissions produced during laboratory experiments (e.g. Eshelby, 1957; Brace and Byerlee, 1966). To analyze earthquake sources, seismic waveform recordings are typically used. As discussed in the previous sections, attenuation is responsible for damping of emitted waves. Site- and path effects bias observed ground displacements and related displacement spectra (see Figs 2.5a and b). Thus, corrections for attenuation need to be applied to obtain true source spectra and accurate source parameter estimates. In this section, only the source parameters that are essential for the here presented research articles are introduced.

Spectral source parameters

The earthquake far-field displacement source spectrum (see Fig. 2.5) can be expressed as the omega-squared source model (e.g. Brune, 1970):

$$\Omega(f) = \frac{\Omega_0}{\left[1 + \left(\frac{f}{f_0}\right)^{\gamma n}\right]^{\frac{1}{\gamma}}}, \tag{2.16}$$

where $\Omega(f)$ describes the far-field displacement amplitude spectrum, f_0 indicates corner frequency, and Ω_0 [m*s] is the low-frequency spectral displacement amplitude. The omega-squared model by Brune (1970) uses $n = 2$ for the high-frequency fall-off rate and $\gamma = 1$, which describes the sharpness of the corner between the frequency plateau and the fall-off (see Fig. 2.5b). Boatwright (1978) uses values of $n = 2$ and $\gamma = 2$ instead, which produces a sharper corner and sometimes a better fit between the model and the observed spectrum (e.g. Abercrombie et al., 2016), facilitating the identification of corner frequency.

In ideal circumstances, the low-frequency plateau (the spectral level, Ω_0) is equal to the area below the observed ground displacement amplitude (which corresponds to the seismic moment rate) and thus proportional to M_0 (Fig. 2.5). The low-frequency spectral level is less affected by attenuation which is generally stronger for higher frequencies (cf. eq. (2.4)). Therefore, seismic moment M_0 is assumed to be a stable source parameter. Seismic moment M_0 and corner frequency f_0 can be directly assessed from the far-field displacement spectrum (e.g. Fig. 2.5). Mathematically, M_0 can be obtained by:

$$M_0 = \frac{4\pi\rho V_C^3 R \Omega_0}{R_C S_C F_C}, \tag{2.17}$$

where R is the source to receiver distance and V_C is the respective wave-velocity. R_C, S_C and F_C are correction factors of the respective free-surface amplification, site effects and wave radiation pattern (e.g. Gibowicz and Kijko, 1994).

The intersection of the high-frequency spectral fall-off and low-frequency plateau (Fig. 2.5) indicates f_0, which is reversely proportional to the source size, once a specific source model is assumed. However, due to the attenuation affected high-frequencies, the determination of accurate f_0 is challenging. Depending on the attenuation correction factor (e.g. the exponent in eq. (2.10) for the PSD of coda waves), the spectral fall-off is shifted and f_0 can vary by a factor of 2 (e.g. Kaneko and Shearer, 2015). Therefore, the quality factor needs to be assessed very carefully.

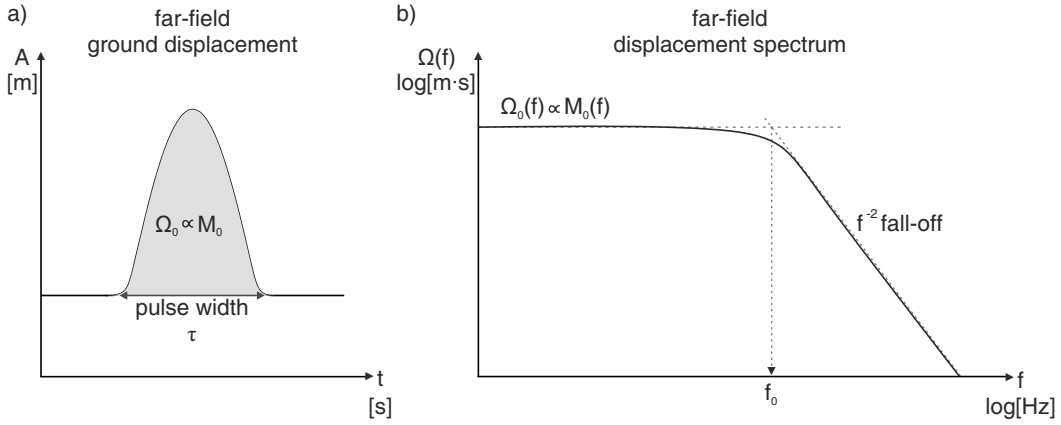

Figure 2.5: (a) Far-field ground displacement in time domain derived from the integrated velocity record. The pulse width τ is inversely proportional to the corner frequency (f_0). The area below the pulse (gray) is conform with the low-frequency spectral level (Ω_0) that is proportional to the seismic moment (M_0). (b) Idealized theoretical model of a far-field shear displacement spectrum obtained by the Fourier transformation of e.g. P- or S-wave of the ground displacement seismogram. f_0 indicates the transition of constant low-frequency spectral level (Ω_0) and high-frequency fall-off. Frequencies higher than f_0 are affected by attenuation processes.

Moment magnitude

M_0 is further used to calculate the moment magnitude M_W (Kanamori, 1977), which characterizes the earthquake size:

$$M_W = \frac{log_{10}(M_0) - 9.1}{1.5}. \tag{2.18}$$

By using M_0, the moment magnitude is directly related to the physical source properties and does not saturate for larger earthquakes (e.g. Shearer, 2009) in comparison to other magnitude types (e.g. local magnitude, body-wave magnitude, surface-wave magnitude).

Source radius

Using M_0 and f_0, further source parameters can be estimated, that are needed to characterize the earthquake source. These parameters usually depend on further model assumptions. According to the model, parameter estimates can vary significantly. Therefore, the information on the applied model is necessary to compare studies with each other. Otherwise, the lack of this information often leads to ambiguities in the interpretation of multiple datasets processes due to different model assumptions.

Different types of theoretical source models, with static, dynamic or quasi-dynamic properties, have been developed (e.g. Brune, 1970; Sato and Hirasaw, 1973; Madariaga, 1976) and are in use. In the far-field, a point source is typically assumed as the wavelengths are much larger than the source radius. In these models, the source size is estimated using the wave-velocity and the corner frequency. For small earthquakes the kinematic source model of Brune (1970) or the quasi-dynamic model of Madariaga (1976) are often applied. Both models consider a spectral fall-off close to the omega-squared model. Brune's source model describes a simple circular fault without taking the rupture velocity (V_r) into account. It uses a constant factor of $C = 2.34$ that is independent on the observation angle and valid for S-waves:

$$r_0 = \frac{CV_\beta}{2\pi f_\beta}, \tag{2.19}$$

with

$$V_\beta = \frac{V_\alpha}{\sqrt{3}} \quad or \quad V_\beta = \sqrt{\frac{\mu}{\rho}}, \tag{2.20}$$

where indices α and β indicate P- and S-wave related parameters, respectively.

By contrast, Madariaga's model describes a plane circular frictional fault, where the slip nucleates in the center:

$$r_0 = \frac{CV_\beta}{2\pi f_C} \tag{2.21}$$

The rupture propagates outwards assuming a constant rupture velocity equal to 0.9 times the shear-wave velocity (V_β) until it suddenly stops. This model considers correction factors of $C = 2.01$ for P-waves and $C = 1.32$ for S-waves and is dependent on the angle between the normal to the fault and the take-off direction of the considered wave. Here, f_C is f_α or f_β, respectively. The difference in the S-wave correction factor of both models, Brune and Madariaga, is about 1.7 times larger for Brune, which leads to differences in the source radius and related further estimates. Stress drops, for example, are about 5 times larger using the Madariaga source model (e.g. Shearer, 2009). Due to the known differences, each of both models can be convert into the other.

Seismic stress drop

Different parameters related to stress measurements, such as e.g. apparent stress, radiated seismic energy, scaled energy and seismic stress drop, can be estimated. In this thesis, the focus lies on static stress drop only and will be discussed in the following.

Seismic stress drop is a fundamental earthquake source parameter and plays an essential role in earthquake hazard evaluation. Over time, the shear stress builds-up on a fault. By exceeding the critical strength of the fault, a sudden shear stress drop occurs and the fault planes move relative to each other, releasing the accumulated energy (e.g. Lay and Wallace, 1995). Seismic stress drop is an approximation of shear stress release during an earthquake and can be estimated from seismological data records. Large stress drops can thereby result in high ground motions hazardous for nearby populated areas. Figure 2.6 illustrates the schematic shear stress change during an earthquake. Two categories of stress drop are usually distinguished: 1) dynamic stress drop (σ_d) and 2) static stress drop ($\Delta\sigma$). Dynamic stress drop is the difference between the initial stress and frictional shear stress on the fault during the slip and is also often called effective stress (e.g. Gibowicz and Kijko, 1994). It addresses radiated seismic energy, reflecting the faulting dynamics that affect ground motions (e.g. Kanamori and Rivera, 2004; Cotton et al., 2013).

Due to complex attenuation and site effects especially for smaller earthquakes in the upper crust (e.g. Kanamori and Rivera, 2004) and also due to the application of static or quasi-dynamic source models, static stress drop is often used in studies that focus on seismic scaling relations (e.g. Abercrombie, 1995; Allmann and Shearer, 2009; Cocco et al., 2016; Shearer et al., 2019, and references therein). Static stress drop ($\Delta\sigma$) (e.g. Aki, 1967; Shearer, 2009) is a measure of the difference between the shear stress before (σ_0) and after an earthquake (σ_1) on a fault, averaged over the rupture fault surface (S)

$$\Delta\sigma = \frac{1}{A}\int_S [\sigma_0 - \sigma_1] dS, \tag{2.22}$$

with A being the total fault area.

For crustal earthquakes, observed static stress drops range between 0.01 and 100 MPa (cf. Fig. 1.1). It is a general assumption that stress drop is independent on earthquake size (self-similarity) because the slip on the fault scales with the rupture area. Static stress drop can be calculated from the seismic source spectrum using M_0 and r_0 estimates. Assuming a circular crack-type source (Eshelby, 1957), the static stress drop is expressed by:

$$\Delta\sigma = \frac{7\pi\mu\overline{D}}{16r} = \frac{7M_0}{16r^3}, \tag{2.23}$$

where $\overline{D}$ is the average displacement and μ is the shear modulus. Here, source radius

(r), which depends on corner frequency (eq. 2.19), is cubed. As mentioned earlier, f_0 can be biased by attenuation that affects the spectral fall-off. In addition to possible wrong model assumptions, estimate uncertainties can therefore also influence estimates of r_0 and even more significantly $\Delta\sigma$ by a factor of up to 8 (e.g. Kaneko and Shearer, 2014). Difficulties in the estimation of stress drop due to error propagation (non-physical parameters), unknown fault dimension, possible heterogeneous slip distribution, and further individual fault characteristics (physical parameters), can strongly influence stress drop estimates as it will be discussed in chapter 4. Taking the latter into account, it needs to be highlighted that measuring absolute stress drop in the earth's crust is unfeasible and we are only able to refer to approximated estimates that complicate the accurate assessment of hazards and risks.

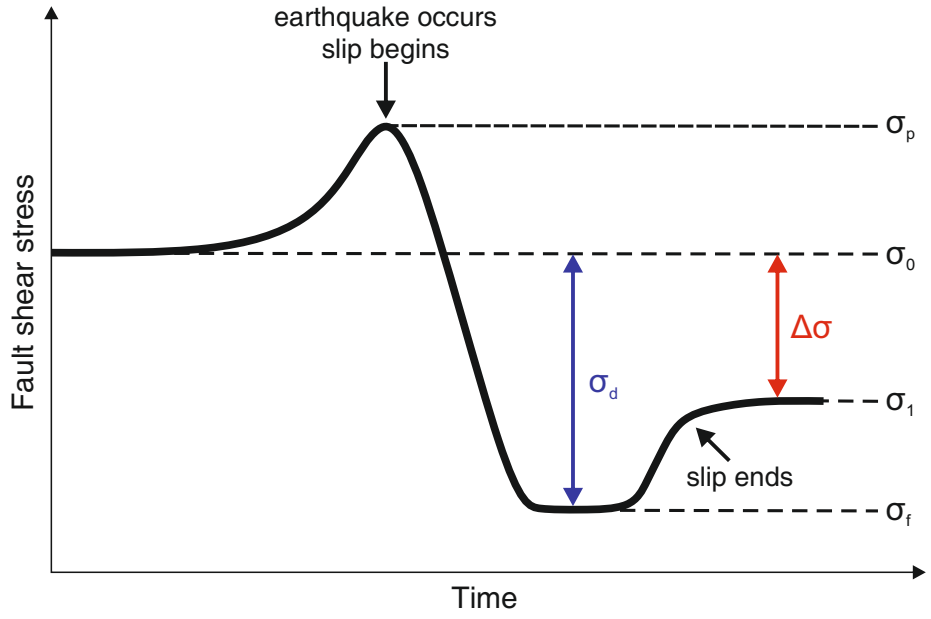

Figure 2.6: Simplified representation of shear stress–time evolution on the fault and seismic stress drop measures (e.g. Causse et al., 2014). Black curve shows shear stress increase and sudden stress release with time. $\Delta\sigma$ indicates static stress drop and σ_d dynamic stress drop. Different stress levels occurring during faulting processes are indicated by σ_p (strength of the fault), σ_0 (initial stress before the earthquake), σ_1 (final stress after the earthquake), and σ_f (frictional stress).

Rupture velocity

Rupture velocity is another important source parameter that potentially influences stress drop estimates. As shown for source models of Brune (1970) and Madariaga (1976), rupture velocity assumptions differ between source models. Further, it is not known whether the assumption of stable rupture velocity is appropriate for any types of earthquakes or

not. On average, V_r is assumed to range between 75 – 95 % of V_β (e.g. Kanamori and Rivera, 2004) for the majority of shallow tectonic earthquakes. However, exemptions have been reported (e.g. Kanamori and Rivera, 2004, and references inside), questioning several possible reasons behind observed variations in V_r.

One possible reason could be related to roughness, suggesting variations of seismic slip over a complex rough fault and a simple smooth fault. It might be that during complex rupture and crack processes more energy is transferred into heat and crack surface energy (e.g. Kanamori and Rivera, 2004) leading to spatially lower rupture and slip velocities (e.g. Mai and Beroza, 2002). Instead of assuming a constant rupture velocity, Sato and Hirasaw (1973) discussed the implication of changing rupture velocity on further source parameters. If an earthquake has a lower rupture velocity, the correction term C (eqs (2.19) and (2.21) needs to be adjusted accordingly. Here, a reduction of V_r leads to an increase of seismic stress drop (cf. eq. (2.23)). Table 2.1 shows correction factors usable in equations (2.19) and (2.21). The correction factors were experimentally determined by Sato and Hirasaw (1973) by assuming an omega-squared model and rupture velocity reductions up to 50 %. The impact on stress drop for laboratory AE events is demonstrated in chapter 4.

Table 2.1: Correction factors for reduced V_r.

V_r	C_P	C_S
$0.5V_\beta$	1.14	1.57
$0.6V_\beta$	1.24	1.70
$0.7V_\beta$	1.41	1.81
$0.8V_\beta$	1.50	1.90
$0.9V_\beta$	1.60	1.99

[1] Note: Excerpt of P- and S-wave correction factors (C_P, C_S from Sato and Hirasaw (1973)).

2.3 The Spectral Ratio Method - A Multi-eGf Variant

To improve the assessment of local seismic hazards, it is important to estimate the seismic source parameters and in specific seismic stress drop as accurately as possible. The spectral ratio technique is a popular method applied in seismic source analysis to avoid making assumptions on propagation-path effects. It is based on the eGf (empirical Green's function) technique, in which a small earthquake serves as an eGf to deconvolve the propagation effects from a co-located larger event by dividing their source spectra and to finally obtain the source time function of the larger event (e.g. Hartzell, 1978;

Mueller, 1985; Frankel et al., 1986). The resolved relative source time functions can then be used to estimate e.g. rupture velocity or directivity effects, or to invert for the spatial distribution of slip over the fault plane (e.g. seismic source tomography) (e.g. Frankel et al., 1986). Therefore, certain requirements such as similar focal mechanisms and a certain magnitude difference of $\sim$1 – 2.5 magnitude units (e.g. Abercrombie et al., 2016) between the two co-located events need to be fulfilled to adequately determine the source time function.

The spectral ratio method aims to resolve the spectral source parameters. Here, the listed requirements for an event pair are not as strict as for the eGf method (e.g. Kwiatek et al., 2011). However, in the spectral ratio method, seismic event pairs also need to be co-located and recorded at the same sensor to allow for assuming similar path-, site-, and sensor effects (e.g. eq. (2.2)) for both considered events. The observed Fourier source spectrum of the large event is divided by the small event as follows:

$$\Psi^{jk}(f) = \frac{S_j(f)\, G_{ji}(f)\, I_i(f)}{S_k(f)\, G_{ki}(f)\, I_i(f)}, \tag{2.24}$$

where $\Psi(f)$ is the displacement spectral amplitude ratio of the event pair j and k observed at sensor i, $G(f)$ is the Green's function (including attenuation, the geometrical spreading, path-, and site effects), I is the instrument recording function, and $S(f)$ is the source spectrum. The spectral ratio of the two co-located events is assumed to suppress the unknown path-, site-, and sensor effects and only the information on the seismic sources of both events remains. Consequently, the seismic source spectra of each event using eq. (2.16) can replace $S(f)$ in eq. (2.24). By considering the Boatwright source model, the spectral ratio is then given by:

$$\Psi^{jk}(f) = \frac{S_j(f)}{S_k(f)} = \frac{M_0^j}{M_0^k} \left(\frac{1 + \left(\frac{f}{f_0^k}\right)^4}{1 + \left(\frac{f}{f_0^j}\right)^4} \right)^{\frac{1}{2}}. \tag{2.25}$$

With equation (2.25), corner frequencies of events forming the spectral ratio pair as well as their seismic moment ratio can be determined. In a variant of the spectral ratio method, in which multiple events fulfill the eGf criteria, a large number of event pairs can be inverted at once and source parameters can be constrained (e.g. Kwiatek et al., 2011; Harrington et al., 2015). In this multi-eGf spectral ratio method each event is linked to a number of other events that fulfill pre-defined eGf requirements (c.f. chapter 4). In this way, a network of event connections is generated. For each linked event pair the theoretical spectral ratio is formed (following eq. (2.25)) and then compared to the

corresponding observed spectral ratio. The proposed inversion approach facilitates to fit all theoretical spectra with the observed spectra at once. Thereby, the inversion problem relies on finding an optimum combination of seismic moment and corner frequency for each event within the multiple-event-network that provides the best fit between theoretical and observed data. The multiple linking thereby aims to stabilize the inversion for source parameters of each event due to the reduction of the impact of outliers. As reported by, e.g., Abercrombie (2013), the position of an event in the spectral ratio (event j or k in eq. (2.25)) might influence the results. It was shown that the smaller event (k) shows less well constrained source parameters as it would show acting as the larger event (j) in the spectral ratio of a linked event pair. In the presented multi-eGf spectral ratio approach an event, which is linked to several other events with different event magnitudes, can be both the larger- but also the smaller event in the spectral ratios during the inversion. Thus, the reported issue by Abercrombie (2013) can be significantly reduced. To find the best fitting solutions of f_0 and M_0, the inversion problem must be converted into an optimization problem.

2.4 Sampling and Optimization Algorithm

The optimization algorithm aims to scan the pre-defined n-dimensional space of modeled parameters and minimize the difference (ε) between the observed (obs) data and their modeled predictions (th). The outcome is saved in a cost function:

$$\varepsilon\left(\Psi_{th}, \Psi_{obs}\right) = \sum_{i} \sum_{j,k} ||\Psi_{th}^{jk}\left(f\right) - \Psi_{obs}^{jk}\left(f\right)||_{L1} = min. \tag{2.26}$$

Here, the L1-norm is used, which measures the absolute differences of the vector components and which is relative robust against outliers. However, depending on the distribution of inverted parameter values, other norm types can be applied.

The multi-eGf spectral ratio inversion includes n-numbers of events, each with two free parameters of f_0 and M_0. The present inversion therefore features a 2 x n-dimensional model space, describing a multi-dimensional non-linear inverse problem that needs to be solved. An optimal combination of two vector elements is searched that minimizes the errors between the theoretical and observed spectral ratios of all event pairs forming the linked network. Due to the non-linearity of the inverse problem and the large number of parameters to invert for, the global optimization algorithm Simulated Annealing is used. A naive grid search approach in that case would be inefficient regarding the high number of parameters and the computational time needed to sample the whole grid space. On the other hand, "greedy" optimization algorithms (cf. Sen and Stoffa, 1995) such as simplex

or a gradient method do not assure that the achieved minimum of the cost function posed in eq. (2.26) is the global one. The chance to only find a local minimum and to miss the global minimum is thus present. Therefore, a Markov Chain Monte Carlo based method that quasi-randomly samples the 2 x n-dimensional model space can be applied to find the global minimum of the cost function.

2.4.1 Metropolis-Hastings Random Walk

The Metropolis-Hastings Random Walk (MHRW) is the simplest version of Markov Chain Monte Carlo algorithms and based on Bayesian statistics that are used to intelligently sample the multi-dimensional space of parameters. This method performs a dependent (short-memory) sampling of the model parameter space in which the next sample step only depends on the current position and not on the steps done before:

$$m_j = m_i + \Delta m_i. \tag{2.27}$$

The current model (m_i) is perturbed by Δm_i to obtain a new model m_j (e.g. Sen and Stoffa, 1995). The perturbation depends on a pre-defined step size that is randomly changed with each iteration of the inversion to allow for larger and smaller step sizes and thus to guarantee a sampling of the whole parameters space. However, the step size should not be too large or too small to avoid too chaotic sampling or stagnation in the sampled space of model parameters. In the next step, the difference between cost functions estimated for the perturbed and the current model is calculated:

$$\Delta\varepsilon_{ij} = \varepsilon\left(m_j\right) - \varepsilon\left(m_i\right). \tag{2.28}$$

The acceptance of the new model depends on two criteria. 1) If $\Delta\varepsilon_{ij} \leq 0$, which means the new model has a lower cost function than that based on the previous model parameters, the new model is always accepted (cf. Sen and Stoffa, 1995). 2) If $\Delta\varepsilon_{ij} > 0$, the Metropolis criterion decides whether the new model is accepted or not. This criterion allows the algorithm to randomly accept models worse (in a sense of cost function) than the previous model (i.e. to allow also moving in the uphill direction). This property of MHRW prevents the algorithm to get stuck in the parameters space around a local minimum of the cost function (Sen and Stoffa, 1995). The probability of the model acceptance is expressed by:

$$P = exp\left(-\frac{\Delta\varepsilon_{ij}}{T}\right), \tag{2.29}$$

with temperature T. To decide whether the new model is accepted or not, $\Delta\varepsilon_{ij}$ is compared

with a uniformly distributed random number U between 0 and 1, and the following applies:

$$exp\left(-\frac{\Delta\varepsilon_{ij}}{T}\right) > U\,[0,1]\,. \tag{2.30}$$

The new model is accepted only if the left side is larger than U (e.g. Fig. 2.7).

2.4.2 Simulated Annealing

The MHRW described in section 2.4.1 is a sampling algorithm that efficiently probe the multi-dimensional space of parameters. To achieve the global minimum of the cost function, the MHRW algorithm needs to be replaced with the Simulated Annealing technique (see Sen and Stoffa, 1995) as described in the following.

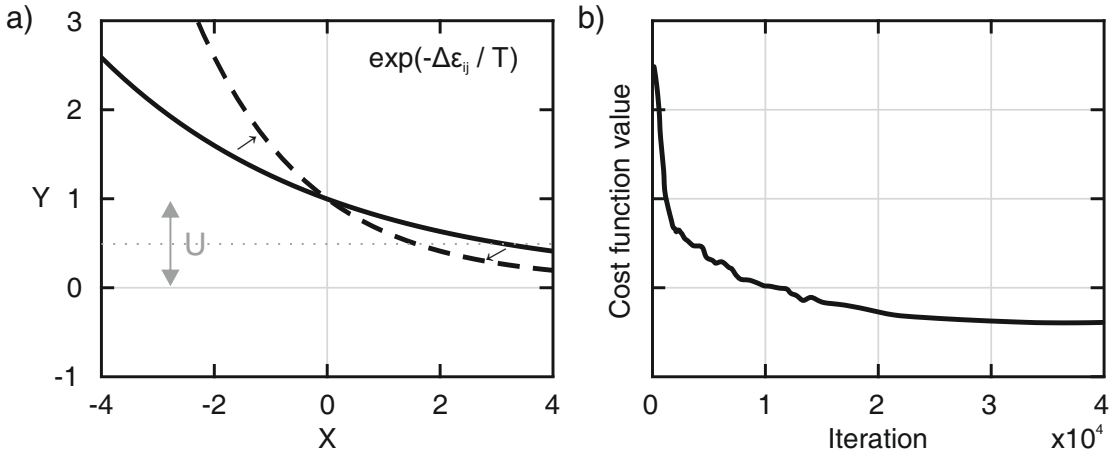

Figure 2.7: Illustration of exponential function (eq. (2.30)). Dashed line and small arrows show the change of function with reduced temperature T. Gray arrow indicates U vector. Dotted gray line indicates the center of U vector. With decreasing T, the e-function decreases in the positive X direction. Thus, U is more likely larger than the function and poor models are less often accepted. b) Conversion of cost function towards the minimum model error (best fit) with increasing iteration step.

After a pre-defined number of iteration steps (e.g. every 200 iterations) of the MHRW algorithm, the temperature T in equation (2.29) is slightly reduced to "cool down" the algorithm. This procedure is inspired by nature, in which usually an energetic equilibrium is tried to be reached. It can be applied to many processes such as crystallization in slowly cooling magma or hardening of metals (e.g. Kirkpatrick et al., 1983). In a heated environment, atoms move around chaotically in space, needing high energy. While cooling the system, the atoms reduce their mobility and try to find a new optimum energetic position in the new (cooled down) environment. For the presented inversion

problem, it is assumed that after a certain number of iterations the parameters space is efficiently sampled and the probability of wrongly sampling local minima is reduced. Therefore, the need to randomly accept poorer models becomes less important. By progressively reducing T, the exponential function in equation (2.30) becomes steeper and, thus, the left term becomes smaller and models with $\Delta\varepsilon_{ij} > 0$ are less often accepted (Fig. 2.7). Consequently, with progress of the optimization procedure, only a model better than the previous model is accepted and the algorithm turns more and more into a gradient algorithm while T –> 0 (e.g. Sen and Stoffa, 1995) until the cost function converges towards the minimum (see Fig. 2.7 and eq. (2.26)).

3 Sensitivity and Stability Analysis of Coda Quality Factors at The Geysers Geothermal Field, California[1]

Abstract

Over 700 induced seismic events recorded between June 2009 and March 2015 at different parts of The Geysers geothermal field, California, are used to estimate local S-wave coda quality factors (Q_C). Recorded by the 31-station short-period Berkeley-Geysers seismic network, the events have duration magnitudes $1 < M_D < 3$, depth ranges of 1 and 4 km, and epicentral distance ranges of 0.7–19 km. We apply the coda analysis technique of Phillips (1985) to find Q_C. Using a sequence of overlapping time windows, the average power spectral density of the coda is calculated. We extract the signal amplitudes at fixed octave-width frequency bands, measure their decay with time and fit for $Q_C(f)$ estimates with associated uncertainties. We investigate the sensitivity of the Q_C results to different input parameters, including lapse time, magnitude range, moving window width, total coda length, and seismic-sensor components. The choice of quality criteria, signal-to-noise ratio, and coda Q uncertainties ($2\sigma(Q_C(f))$) are found to be most sensitive factors. Testing different window lengths and lapse times results in relatively small variations of Q_C. The final, highest quality mean coda Q ($\overline{Q_C}(f)$) estimates are further tested in the context of their spatio-temporal behavior in the reservoir. We found that distance and azimuthal dependence of $\overline{Q_C}(f)$ are related to the observed crack-induced reservoir anisotropy, lithological, and structural features. At the northwestern The Geysers about 50 % larger $\overline{Q_C}(> 40$ Hz) estimates are obtained compared with the southeast. In contrast, geothermal production rate variations, analyzed for a tight cluster in the northwest, do not influence the $Q_C(f)$ estimates. Moreover, we compare the $Q_C(f)$ results with previous estimates of direct S-wave quality factors (Q_β). A match for Q_C results at 7 Hz center frequency with Q_β estimates is observed. However, Q_C estimates show lower scattering and thus a higher stability.

[1]This research article is published as Blanke, A., G. Kwiatek, P. Martínez-Garzón, and M. Bohnhoff (2019). Sensitivity and Stability Analysis of Coda Quality Factors at The Geysers Geothermal Field, California. *Bulletin of the Seismological Society of America*, 109(3):959-975. https://doi.org/10.1785/0120180219.

3.1 Introduction

The seismic quality factor Q is required for earthquake source studies that aim at investigating seismic source properties. Q describes the attenuation characteristics of the propagation medium and its impact on seismic waves. It is used to correct the seismic source spectrum to determine seismic source parameters such as corner frequency, seismic moment, radiated energy, static stress drop and apparent stress. It is also used for the prediction of ground motion (e.g. Stork and Ito, 2004) in hazard and risk quantification. These parameters are, inter alia, sensitive to Q. Therefore, the inaccurate estimation of Q leads to uncertainties in the calculation of source parameters and ground motion estimations (Eaton, 2011).

Traveling towards the earth's surface, seismic waves lose energy due to geometrical spreading and intrinsic- and scattering attenuation. Intrinsic attenuation describes the inelastic energy loss due to the composition of the medium in general or in particular due to the absorption at grain boundaries and mineral dislocations. On the other hand, scattering attenuation provides information about the elastic energy loss caused by the interaction of seismic waves with structural inhomogeneities. These include cracks, fractures and faults of different orientations and lengths, or variations in rock type and formation, all causing diffractions, reflections and refractions (e.g. Aki, 1980b; Shapiro and Kneib, 1993).

In general, the seismic quality factor can be determined by using direct seismic phases and applying a variety of methods (e.g. spectral fitting). In a homogeneous medium, it is assumed that a direct wave travels straight from source to receiver mostly unaffected by attenuation at short epicentral distances (e.g. Abercrombie, 2013). However, in geothermal areas direct waves are influenced at least by the quality of the medium (intrinsic attenuation) and high temperatures. Therefore, we assume that Q obtained by the direct phase (Q_β) only specifies intrinsic and not scattering attenuation, and that scattering attenuation for this phase is small enough to be negligible in the short paths that we consider in our analysis. Spectral fitting usually provides only a single average Q estimation and we take Q_β as frequency-independent.

Scattering attenuation, as a result of various scatterer sizes distributed in the earth (Frankel, 1991), is frequency-dependent. This type of attenuation can be obtained from analyzing the coda of a seismogram. Seismic waves hit scatterers in the medium and get scattered on their way to the surface resulting in later arrivals in the seismogram. S-waves are affected more than P-waves due to their propagation characteristics. In this study, we focus on the analysis of the S-wave coda to obtain the frequency-dependent coda quality factor (Q_C).

Q_C analysis was first used by Aki (Aki, 1969). Since then, several different approaches have been developed using different types of waves (surface waves, body waves), seismic networks and sensors, varying epicentral distances (regional, local, teleseismic) and various mathematical concepts to estimate Q_C (e.g. Sato, 1977; Aki and Chouet, 1975; Phillips, 1985; Novelo-Casanova and Lee, 1991; Margerin et al., 1998).

The main advantage of analyzing the seismic coda in comparison to direct phases is the elimination of critical parameters. These are the source radiation pattern, source directivity, path effects and source-receiver distance that have to be estimated first when applying, for example, the spectral analysis. Depending on source-receiver distance and wave type, the seismic coda spectrum is almost independent on the latter parameters (e.g. Aki, 1969; Aki and Chouet, 1975; Sato, 1977; Phillips and Aki, 1986). Although there is a variety of different coda Q techniques such as the back-scattering model (Aki and Chouet, 1975) or the single isotropic-scattering model (Sato, 1977), usually similar values of Q_C are obtained (e.g. Novelo-Casanova and Lee, 1991). Nonetheless, the parameterization (e.g. total coda length, starting point of coda analysis, width of moving analyzing windows, to name a few) used within Q_C analysis techniques but also seismic network characteristics need to be considered and must be assessed carefully to allow comparison between different studies (e.g. Mukhopadhyay et al., 2008; Calvet and Margerin, 2013).

The isotropic single scattering model by Phillips (1985) was identified to provide stable Q_C estimates over a wide frequency range allowing to use the early coda portion immediately following the direct S-wave arrival (e.g. Sato, 1977; Novelo-Casanova and Lee, 1991; Gibowicz and Kijko, 1994). The Geysers geothermal field in California exhibits short coda durations due to strong attenuation and small earthquake magnitudes. Therefore, we apply the approach of Phillips (1985) to gain information on scattering attenuation. By applying a single scattering model, it is generally assumed that the coda quality factor is composed of intrinsic and scattering attenuation because they are not distinguishable in this case (e.g. Gibowicz and Kijko, 1994).

The Geysers geothermal field (Fig. 3.1) is located in the Mayacamas Mountains approximately 110 km northeast of San Francisco, California, within the northwest-striking San Andreas transform fault system (Ross et al., 1999; Erten et al., 2001; Sadowski et al., 2016). The geothermal field is the world's largest steam-dominated geothermal reservoir and bound by two regional right-lateral strike-slip faults (Martínez-Garzón et al., 2016). It is in operation since the 1960s and several hundred thousand induced seismic events up to magnitude ~ 5 have occurred exhibiting a clear spatiotemporal relationship to the water injection and steam production (Ross et al., 1999; Erten et al., 2001; Elkibbi and Rial, 2003; Martínez-Garzón et al., 2014).

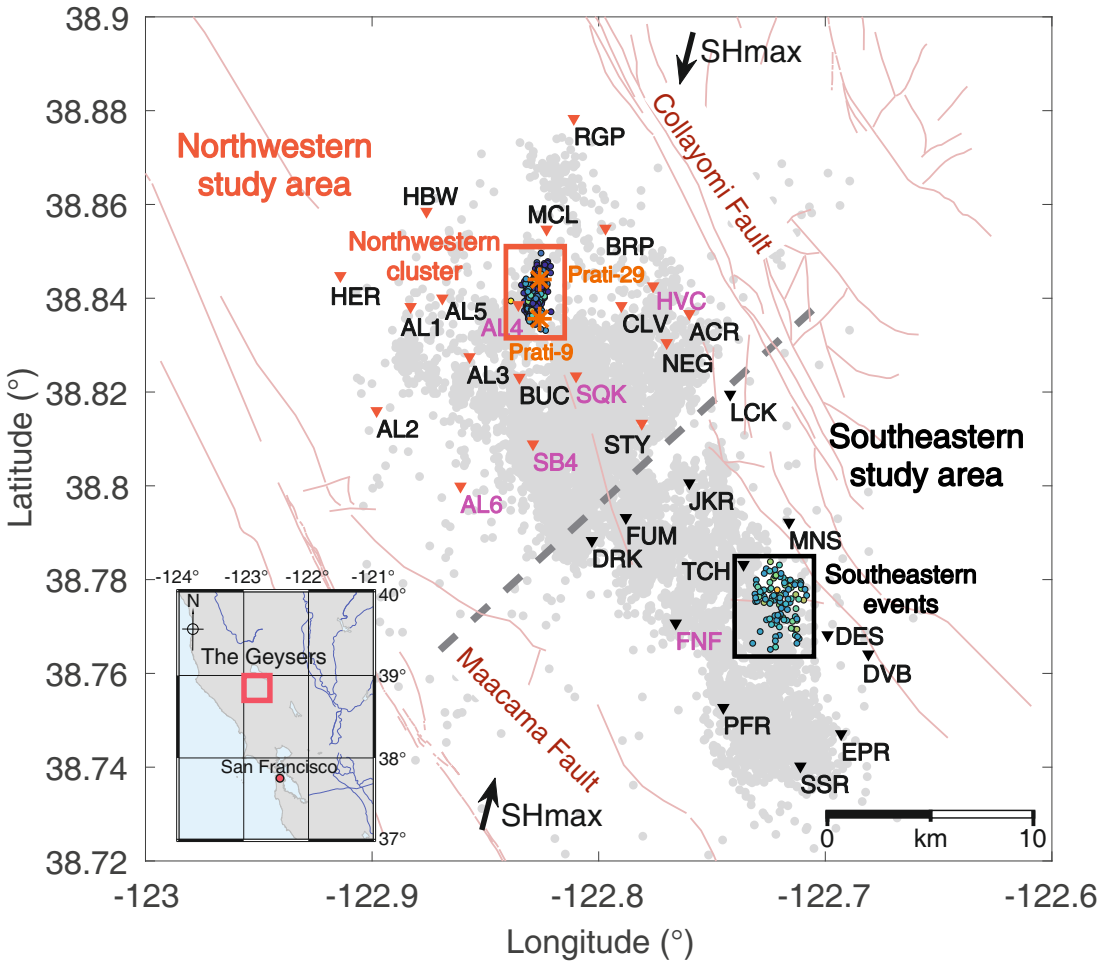

Figure 3.1: Overview of The Geysers geothermal field. Map inset shows the location of The Geysers approximately 110 km northeast of San Francisco, California. Boxes indicate selected event sites. Inverted triangles indicate seismic stations with station names. Light shaded triangles feature northwestern stations and dark shaded triangles reflect southeastern stations. Shadings of triangles are also utilized in Figures 3.8-3.10 for distinguishing northwestern and southeastern stations and for showing the connection between these figures. Arrows indicate the direction of maximum horizontal stress (SHmax). Light gray dots reflect the seismicity recorded between 2009 and 2015 at The Geysers. Asterisks show location of injection wells Prati-9 and Prati-29. Dashed line indicates the imaginary division of the study area into a northwest and southeast part (e.g. Erten et al., 2001; Beall and Wright, 2010). Thin lines show main faults. Highlighted stations are used as examples in sections 3.4 and 3.5, Figures 3.3-3.8 and 3.11 and Table 3.1.

Seismic activity at The Geysers varies spatially with higher seismicity rates and moment release (in part magnitudes $M \geq 3.0$) in the northwest (Beall and Wright, 2010) compared with the southeast. There, the microseismicity typically occurs down to 5 km, reaching the high temperature zone (up to 360° C) (Elkibbi and Rial, 2003; Beall and Wright, 2010). The extend of the high-temperature zone to the southeast is unknown.

In the southeast, the seismicity is on average observed at shallower depths (Beall and Wright, 2010), which might be related to the bedrock geology. The whole area lies within the Mesozoic Franciscan Complex composed of sedimentary rocks, fractured low-grade metamorphic rocks (e.g. metagraywacke) and igneous rocks. The entire area is underlain by two subsurface lithologies, a hornfels and a granitic intrusion, locally known as felsite that intruded the hydrothermal system and enhances the reservoir's rock quality (Ross et al., 1999; Elkibbi and Rial, 2003; Sadowski et al., 2016). The intrusion shows an elevation to the southeastern edge of the field and subsides towards the northwest (e.g. Hartline et al., 2015; Sadowski et al., 2016). Nevertheless, the rocks are permeated by a predominantly vertical, or steeply dipping minor fracture network, primarily oriented parallel to the maximum horizontal stress orientation (see Fig. 3.1 and Elkibbi and Rial, 2003). This fracture network induces anisotropy into the otherwise isotropic rocks (Erten et al., 2001; Elkibbi and Rial, 2003; Elkibbi et al., 2005). Predominant fractures in the northwest are aligned N10°E - N50°E and in the southeast N40° - 50°E and N40° - 50°W. There is also evidence that some cracks have shallower dips and different orientations, making the investigation area more complex (Elkibbi and Rial, 2003). Moreover, shear-wave splitting is observed at The Geysers which is assumed to result from the crack-induced anisotropy (Erten et al., 2001; Elkibbi and Rial, 2003; Elkibbi et al., 2005). The presence of fractures, shear-wave splitting as well as sometimes delayed S-wave amplitude peaks, that we observed during manual phase picking, support the idea of scattered S-waves at The Geysers and motivate our analysis of the frequency-dependent Q_C. We conduct the coda analysis for the southeastern and northwestern part (see Fig. 3.1), which differ in their geological and geothermal character. We also analyze our final results with epicentral- and azimuthal distance dependence at The Geysers. The close relation of injection phases and occurring seismicity additionally motivate us to evaluate the temporal behavior of coda Q with changing injection activities at the injection wells Prati-9/-29 (see Fig. 3.1) in the northwestern part of The Geysers.

In addition to the coda analysis, we also apply a sensitivity analysis including all involved parameters (e.g. Mukhopadhyay et al., 2008; Havskov et al., 2016). We therefore compare results obtained by using different parametrizations to investigate their impact on the outcome.

Using the seismic coda to gain information of the seismic source is simple and provides the most stable estimates of Q and source parameters (e.g. Mayeda and Walter, 1996; Yoo et al., 2011). Kwiatek et al. (2015) estimated the local quality factor for the northwestern Geysers by applying an initial source parameters estimation using direct S-waves and spectral fitting. Because coda analysis provides more stable source parameters than general common methods (e.g. spectral analysis), we use the available intrinsic non-

frequency-dependent Q_β estimates of Kwiatek et al. (2015) for comparison with our obtained frequency-dependent Q_C estimates at low frequencies. In addition, shear-wave splitting and partially observed delayed peak amplitudes also motivate us to have a closer look at both obtained Q estimates (see tables S3 and S4 of Blanke et al., 2018).

3.2 Data Base

For this study, we selected a spatially confined cluster (Fig. 3.1) from the northwestern part of the field (Kwiatek et al., 2015; Martínez-Garzón et al., 2016, 2017). The catalog consists of 869 events covering an area of approximately 1 x 2 km around the waste water injection wells Prati-9 and Prati-29 (Martínez-Garzón et al., 2017) for which full moment tensors were determined based on manual phase and amplitude picks. The events occurred between January 2006 and June 2015 forming a tight spatial cluster close to the injection locations. The event rate reflects the injection flow rate that in general was higher during the winter than during the summer seasons (Kwiatek et al., 2015; Martínez-Garzón et al., 2016). All microseismic events were first relatively relocated by applying the double-difference relocation technique (Waldhauser and Ellsworth, 2000) to render more precisely the hypocenter relocations down to about 50 m (Kwiatek et al., 2015). The three-component seismograms were filtered using a 1 Hz high-pass Butterworth filter. Therefore, only high-quality waveforms from seismic receivers with a signal-to-noise ratio (SNR) of at least 20 dB were considered (Kwiatek et al., 2015).

In this study, we use about 617 microseismic events of the earlier refined seismic catalog recorded between June 2009 and March 2015 with in total 31 short-period three-component surface instruments (OYO-GS11D) of the local Berkeley-Geysers (BG) network, sampling data at a rate of 500 Hz. These events were selected from the overall catalog using a reduced time period in which all 31 seismic stations were in operation. The selected events cover duration magnitudes from M_D 1 to 3 and occurred at depths of 1–3 km, that is, mostly within the high temperature zone of the northwestern geothermal study area. We refer to this event dataset as northwestern cluster in the following (cf. Fig. 3.1). The epicentral distances range from 0.7 to 15 km, providing high SNR.

For the coda analysis, a first preselection of the available events was made to facilitate the later procedures and to reduce computation time. Only events with reliable manually picked S-onsets and a reasonably high SNR were chosen for the coda Q analysis using 40-second-long waveform windows extended by 10 s added prior to the detected event (Kwiatek et al., 2015).

Because of the tentative structural subdivision of the entire geothermal field into a northwestern and southeastern part, we additionally analyzed 100 manually picked

events from the southeastern part of The Geysers (see Fig. 3.1). The selection of those additional events is based on the location within the network to ensure enough seismic measurements for the coda Q analysis which could not be made by using only the northwestern cluster. We chose the area around the southeastern station TCH (see Fig. 3.1) and selected the largest events of M_D 2–3. The southeastern events are not clustered but occurred within the same depth range of 1–3 km. Here, the epicentral distances range between 1.3 km and 18.5 km and the data was processed in the same way as the northwestern dataset. Q_C obtained from events occurred in different parts of the study area provides us a more objective interpretive approach of our results.

3.3 Methodology

To estimate the frequency dependency of Q_C, we follow the coda analysis technique of Phillips (1985). It uses a single isotropic scattering model as an approximation of S-wave scattering in the earth's crust. It is assumed that point source and receiver are spatially separated from each other and inhomogeneities are randomly but homogeneously distributed.

The frequency-dependent Q_C calculation was performed in two steps.

3.3.1 Moving window method

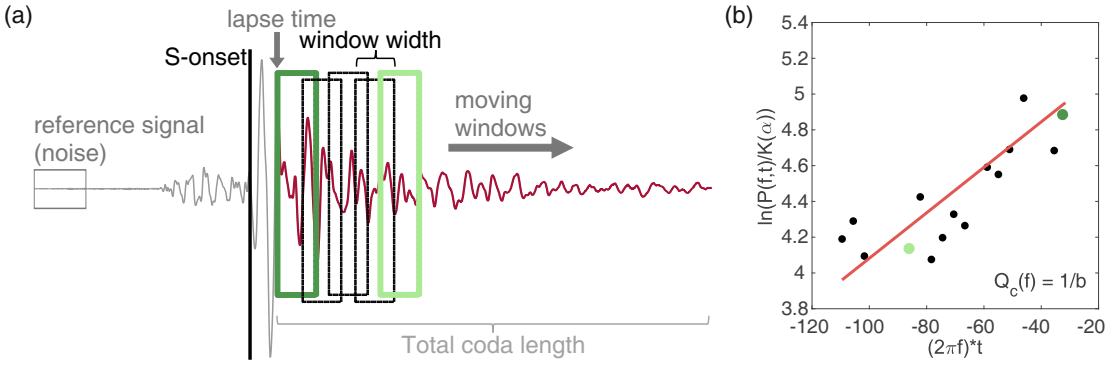

Figure 3.2: Graphical illustration of (a) moving window analysis and (b) regression analysis of Phillips (1985) using S-wave coda. (a) A reference signal is selected prior to the seismic signal of interest. The lapse time is chosen after the S-onset pick to avoid including the source. Overlapping moving windows are defined. The seismogram is filtered in different octave-width frequency bands to obtain the power spectral amplitude decay for each center frequency of each frequency band. (b) Rescaled amplitudes of narrowband-filtered seismograms as a function of related time. Highlighted moving windows in (a) are related to highlighted dots in (b). Q_C is estimated from the slope of regression line for each frequency band.

First, we calculated the Power Spectral Density (PSD) for moving time windows of predefined size (Fig. 3.2). This procedure was performed for each event at each station component in specified octave-width frequency bands (Aki and Chouet, 1975; Phillips, 1985; Phillips and Aki, 1986). Therefore, we predefined center frequencies for each band, ranging from 1 to 69 Hz. Because of a fast amplitude decay down to the level of noise in the investigation area, we started from the very early coda part and tested different lapse times (time elapsed since the origin earthquake occurred).

The PSDs were calculated for overlapping moving windows (Fig. 3.2) applying a cosine taper. A window overlap of 40 % guaranteed uncorrelated residuals (Phillips and Aki, 1986). We tested different lapse times and total coda lengths to check the stability of Q_C estimates. To ensure a reasonable SNR, a seismogram part (the noise) prior to the event onset was selected as a reference level for the seismic signal. A minimum coda amplitude was set at two times the noise level in a considered frequency band.

3.3.2 Regression analysis

In a second step, the coda decay was modeled in frequency domain for a total coda length of 6 and 10 s, following Aki and Chouet (1975):

$$P(f,t) = C(f)t^{-2}exp\left(\frac{-2\pi ft}{Q_C(f)}\right), \tag{3.1}$$

in which *P(f,t)* describes the PSD of the coda wave displacement amplitude, $C(f)$ is the coda source factor and is assumed to be a constant for near-field narrow bandpass filtered seismograms (Novelo-Casanova and Lee, 1991; Gibowicz and Kijko, 1994), t is the lapse time (starting time of coda window analysis), t^{-2} describes the geometrical spreading of body wave amplitudes with elapsed time (Jin and Aki, 1989) and Q_C is the coda quality factor. Because of the fast amplitude decay at The Geysers and a short coda duration, we replaced the factor t^{-2} with $K(\alpha)$, introduced by Sato (1977)

$$K(\alpha) = \frac{1}{\alpha} ln\left(\frac{\alpha+1}{\alpha-1}\right), \tag{3.2}$$

with

$$\alpha = \frac{t}{t_S}, \tag{3.3}$$

which allows to assume a local separation of source and receiver, taking the geometrical spreading factor α, in which t_S is the S-wave travel time, into account (see Sato, 1977;

Gibowicz and Kijko, 1994). By taking the logarithm of equation 3.1 of both sides

$$ln\left(\frac{P(f,t)}{K(\alpha)}\right) = ln\left(C(f)\right) - \left(\frac{2\pi f}{Q_C(f)}\right), \tag{3.4}$$

we obtained measurements for each moving window within a selected frequency band. These measurements were fitted with a regression line, describing the decay of coda wave envelopes with time per frequency band, and the slope of the line is inversely proportional to Q_C. Thus, we received coda quality factors for each center frequency $Q_C(f)$ for each single event at each station and each sensor component (two horizontal, one vertical). Finally, we averaged our $Q_C(f)$ estimates at each station and calculated mean coda Q ($\overline{Q_C}(f)$) curves and their standard deviations (st. dev.) for each station.

3.4 Sensitivity Analysis of Coda *Q*

We first checked whether $Q_C(f)$ depends on the choice of a particular magnitude group. Furthermore, we tested whether the choice of different analyzed lapse times, taper sizes, window widths, and total coda lengths influence our $\overline{Q_C}(f)$ estimates. We also compared $\overline{Q_C}(f)$ between the three sensor components. Simultaneously to the sensitivity analysis, we applied quality criteria to remove the most striking outliers of the $\overline{Q_C}(f)$ curves. For this, we tested a set of SNRs and maximum allowed uncertainties $2\sigma\left(Q_C(f)\right)$ (95 % confidence interval) of all single estimations that in the next step were used to calculate the final $\overline{Q_C}(f)$ curves for each station. The results of our sensitivity analysis are demonstrated using only the northwestern cluster. The same final results are received by applying the analysis on the southeastern events.

3.4.1 Magnitude dependence

We divided the dataset by three predefined magnitude groups: 1) $M_D = 1.0$–1.4, 2) $M_D = 1.5$–2.0, and 3) $M_D = 2.1$–3.0. Single $Q_C(f)$ (Fig. 3.3a) and additionally $\overline{Q_C}(f)$ curves (Fig 3.3b) were estimated for each magnitude range at each single station component. We applied, based on visual inspection of Figure 3.3a, a loose quality criteria to the mean curves in Figure 3.3b to exclude unrealistically outlying curves. We restricted the accepted uncertainties of single $Q_C(f)$ estimates to 180 ($2\sigma(Q_C(f)) < 180$) and plotted the mean curves with error bars.

It is evident that larger magnitudes ($M_D = 2.0$–3.0) result in more stable $Q_C(f)$ curves (Fig. 3.3a) in comparison with smaller magnitudes ($M_D = 1.0$–1.4). For magnitudes between 1.0 and 2.0, the most characteristic magnitude range of our dataset, the curves

show the highest variability related to a lower SNR of small earthquakes. However, in Figure 3.3b it is clearly shown that the calculated mean curves exhibit similar shapes, regardless of considered magnitude ranges. Here, only the SNR seems to be a factor that influences the uncertainties in coda Q estimates, especially at very high frequencies. The stability and similarity of the mean curves lead us to conclude that the tested magnitude groups do not significantly affect the $\overline{Q_C}(f)$ estimates and can be neglected in the presented magnitude range.

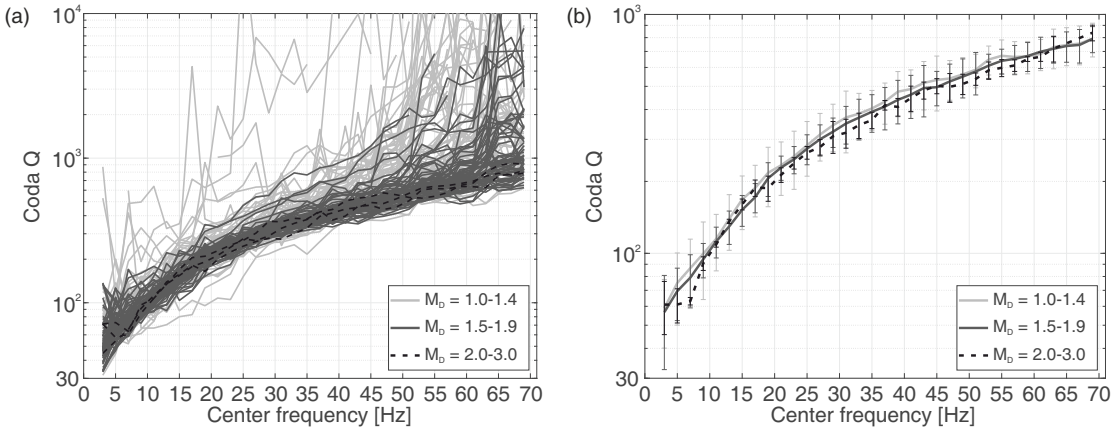

Figure 3.3: Coda Q estimates for station FNF obtained from the analysis of NW cluster. The curves are divided and shaded respectively by magnitude intervals of M_D 1.0–1.4, M_D 1.5–1.9, and M_D 2.0–3.0. (a): single $Q_C(f)$ curves of all used events. (b) $\overline{Q_C}(f)$ curves of the three defined magnitude ranges restricted by $2\sigma\,(Q_C(f)) < 180$. Error bars indicate standard deviation (st. dev.) of $\overline{Q_C}(f)$ estimates.

3.4.2 Quality criteria

After merging all events, regardless of event magnitudes, we additionally checked whether restrictions to even more high-quality SNRs and to accepted $Q_C(f)$ uncertainties (2σ st. dev. of single estimated coda Q values) affect the $\overline{Q_C}(f)$ results. The SNR in this part of the study is calculated by dividing the integral of the band-filtered signal by the integral of the band-filtered noise. In this way, we achieve a SNR value for each single coda Q estimate. We compared $\overline{Q_C}(f)$ curves with same parameterizations testing different SNRs and accepted $2\sigma\,(Q_C(f))$ of single estimated $Q_C(f)$ values that have been used to obtain the ultimate $\overline{Q_C}(f)$. We tested SNRs of ≥ 5000 and ≥ 10.000 and maximum accepted $2\sigma\,(Q_C(f))$ of ≤ 60, ≤ 30 and ≤ 10.

Table 3.1 shows $\overline{Q_C}(f)$ estimates with respective quality criteria for selected stations of the whole study area. Different applied SNRs do not have an obvious influence on $\overline{Q_C}(f)$, which proves that our data are already of high quality. But by restricting the SNR to ≥ 10.000, minor features become obvious that might be of interest (see bold

numbers in Tab. 3.1). $\overline{Q_C}(f)$ estimates of some stations such as AL4, SB4 and AL6 show anomalies at some frequencies without imposed quality criteria. By applying higher SNRs, $Q_C(f)$ outliers with lower SNRs are removed and do not influence $\overline{Q_C}(f)$ for the respective frequency band anymore. Nonetheless, site effects are still visible and these still results in slightly different $\overline{Q_C}(f)$ estimates at some frequencies (see Tab. 3.1 stations AL4, SB4 and AL6). This observation is important for the later discussion of coda Q results.

Although investigating constrains to the allowed $2\sigma\left(Q_C(f)\right)$ at each center frequency, we observe significant differences in $\overline{Q_C}(f)$ (see Tab. 3.1). We receive no estimates at the highest frequencies while accepting only small errors (e.g. Tab. 3.1 stations AL6 and FNF). That indicates that $Q_C(f)$ estimates at very high center frequencies have higher uncertainties and have therefore a stronger influence on $\overline{Q_C}(f)$. At the same time, $Q_C(f)$ at lower frequencies remain stable at all stations. We do not list results of $2\sigma\left(Q_C(f)\right) \leq 10$, because the restriction seems to be too radical ending up with almost no evaluable $\overline{Q_C}(f)$ estimates for most of the stations.

Table 3.1: Impact of SNR and $2\sigma\left(Q_C(f)\right)$ on $\overline{Q_C}(f)$ Estimates.

Site	f (Hz)	$\overline{Q_C}(f)$ for SNR $\geq 5,000$ $2\sigma\left(Q_C(f)\right) \leq 60$	$\overline{Q_C}(f)$ for SNR $\geq 5,000$ $2\sigma\left(Q_C(f)\right) \leq 30$	$\overline{Q_C}(f)$ for SNR $\geq 10,000$ $2\sigma\left(Q_C(f)\right) \leq 60$	$\overline{Q_C}(f)$ for SNR $\geq 10,000$ $2\sigma\left(Q_C(f)\right) \leq 30$
AL4	5	68	66	68	66
	19	174	167	175	168
	33	257	243	258	243
	47	**301**	**278**	**306**	**281**
	63	**348**	**301**	**357**	**311**
HVC	5	68	67	68	67
	19	179	174	179	174
	33	302	288	302	288
	47	397	372	397	371
	63	478	457	476	457
SB4	5	58	57	58	57
	19	163	160	165	161
	33	273	266	273	266
	47	342	320	343	320
	63	404	**353**	407	**364**
AL6	5	74	71	74	71
	19	203	194	204	194
	33	303	282	303	282
	47	364	316	363	316
	63	**394**	-	**409**	-
FNF	5	72	71	72	71
	19	208	202	209	202
	33	359	353	361	353
	47	489	461	488	461
	63	599	-	599	-

[1] Note: Excerpt of $\overline{Q_C}(f)$ estimates obtained from the northwestern cluster at different seismic stations for selected center frequencies (f). Different signal-to-noise ratios (SNRs) and coda Q uncertainties ($2\sigma\left(Q_C(f)\right)$) are applied. Horizontal bars indicate no measurements for the related center frequency and above. Bold numbers highlight estimate changes with varying SNRs at high frequencies.

3.4.3 Sensor component dependence

The previously determined $\overline{Q_C}(f)$ values were estimated per station per component (in total three components). This enables us to compare the $\overline{Q_C}(f)$ measurements between the components. At stations with continuously good SNRs, all three components display very similar $\overline{Q_C}(f)$ curves (Fig. 3.4). We show examples for two stations, HVC and FNF, which are in the closer vicinity of the northwestern cluster 4.5 km, and 10 km away in the southeast of the study area, respectively.

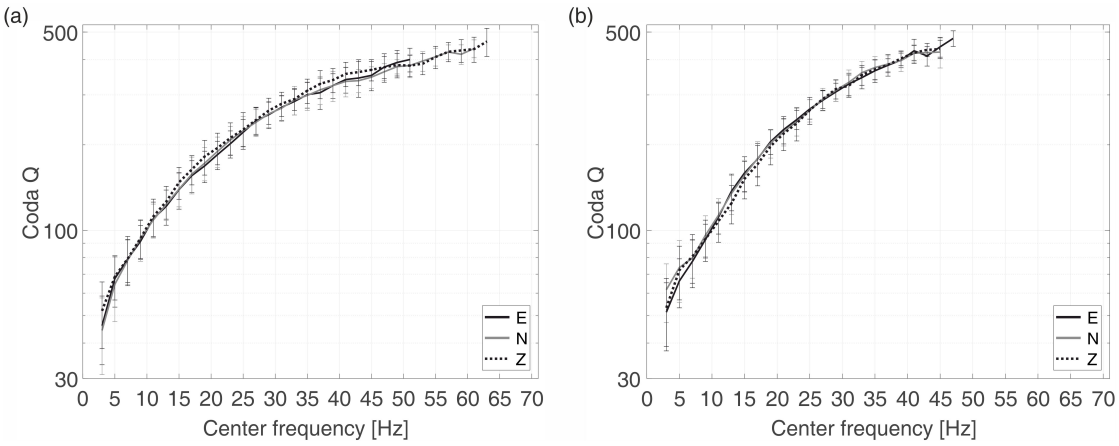

Figure 3.4: $\overline{Q_C}(f)$ curves with st. dev. (error bars) of each component for the closer station (a) HVC and the more distant station (b) FNF. Different line styles and shadings indicate the horizontal components E and N, and the vertical component Z.

In general, we observe that the vertical component provides longer curves and more estimates at high frequencies (e.g. Fig. 3.4a). Station FNF (Fig. 3.4b) features shorter curves related to the larger epicentral distance to the northwestern cluster. The lack of estimates at high frequencies at more distant stations could result in less stable mean curves and small differences between sensor components, due to fewer $Q_C(f)$ estimations in total. Nevertheless, the overall agreement of $\overline{Q_C}(f)$ estimates and their st. dev. between sensor components in Figure 3.4 allows to estimate a single $\overline{Q_C}(f)$ curve at each station. Because of more included measurements for each frequency band, the resulting curves become more stable in shape (less scattering at high frequencies). For all our stations, we therefore decided to use all sensor components since there is no dependence visible in our study.

3.4.4 Moving window width and related taper dependence

We tested the influence of different moving time-window widths on our $\overline{Q_C}(f)$ estimates. We chose 128 and 256 samples (= 0.256 and 0.512 s time window) and related tapers of 8 and 16 samples for the moving window size.

Regardless of the selected station, we observe that the choice of moving window width influences mainly the very high frequencies, however, in inconsistent manner (Fig. 3.5). At station HVC (Fig. 3.5a), for instance, we observe small differences in curve shapes but still within the error bars (st. dev. $\left(\overline{Q_C}\right) \approx \pm 30$), whereas at station FNF (Fig. 3.5b) no differences can be seen. Taking into account the uncertainties (Fig. 3.5), the choice of moving window width and related taper does not influence the estimates in our study when starting the analysis in the early coda. The smallest width of 128 samples was used for the final $\overline{Q_C}(f)$ curves.

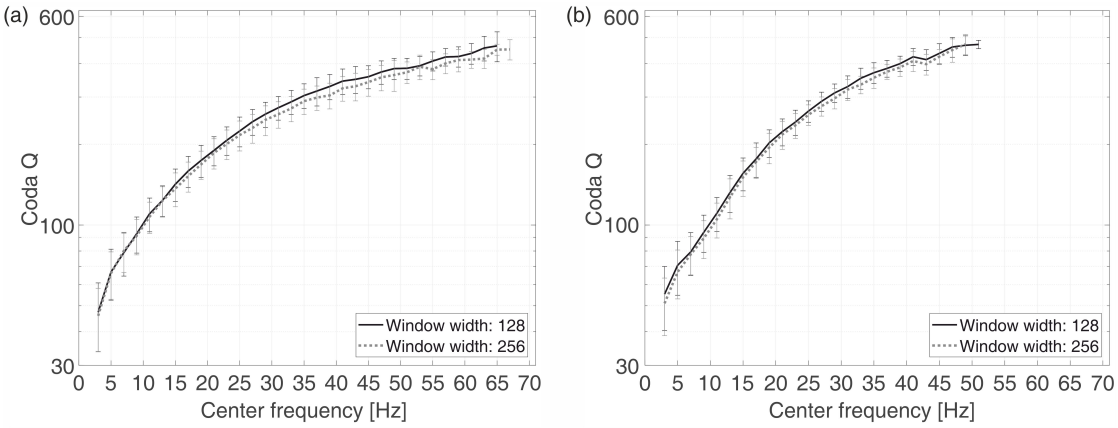

Figure 3.5: $\overline{Q_C}(f)$ curves with st. dev. (error bars) per center frequency for station (a) HVC and (b) FNF. Different line styles indicate tested window widths of 128 and 256 samples.

3.4.5 Lapse time effect

Many authors (e.g. Mukhopadhyay et al., 2008; Calvet and Margerin, 2013; Havskov et al., 2016) report that the coda analysis is significantly dependent on lapse time. The observation of receiving higher $\overline{Q_C}(f)$ estimates using a larger lapse time is explained by the assumption that later arriving waves travel deeper through the earth or pass through a larger volume between source and receiver. In general, it is suggested to start after at least two times the S-wave onset time (t_S) to ensure the validity of single backscattered S-waves (e.g. Rautian and Khalturin, 1978; Phillips, 1985).

Using equation 3.2 allowed us to start earlier in the coda. We account for this decision due to the rapid amplitude decay observed in our seismograms that makes it impossible to start after two times the S-wave onset time. Thus, we tested lapse times of 1.0, 1.1, 1.2 and 1.5 times t_S to see whether it affected the $\overline{Q_C}(f)$ estimates.

Figure 3.6 shows $\overline{Q_C}(f)$ curves, again for stations HVC and FNF, for four tested lapse times. The lapse time does not seem to affect the $\overline{Q_C}(f)$ curve significantly at station HVC. At lapse time 1.5 t_S, the shortest curve ends at 61 Hz. The $\overline{Q_C}(f)$ curves at station

FNF are more affected by increasing lapse time. An increase in $\overline{Q_C}(f)$ for lapse time 1.5 t_S (light dotted curves) between 5 and 31 Hz is obvious in comparison to smaller lapse times but still within the given st. dev. This small increase is sometimes also observed at other stations with a larger distance to the northwestern cluster. Almost no estimates could be obtained for the most distant stations such as DES or TCH at lapse time 1.5 t_S. The shortest curve (ending at 45 Hz) at station FNF is also obtained for the largest lapse time. Therefore, using 1.5 t_S is not advisable as it shows the shortest mean curves.

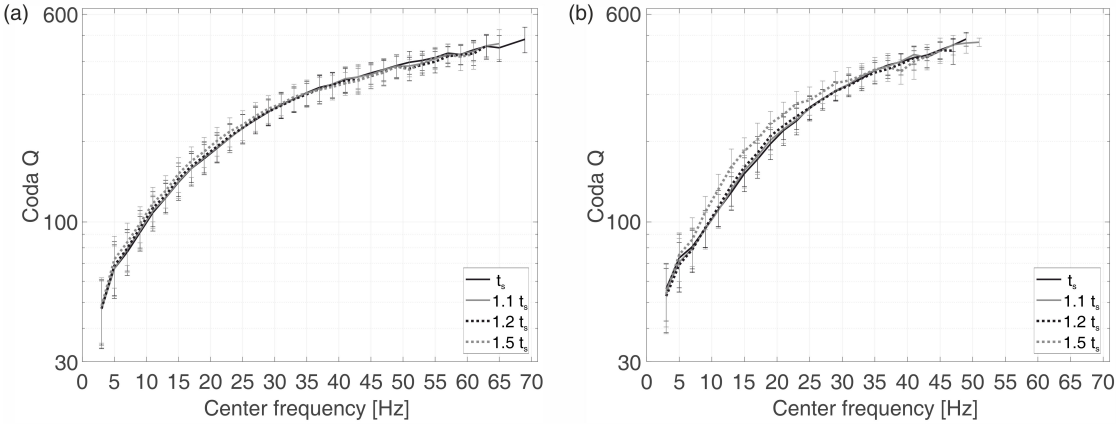

Figure 3.6: $\overline{Q_C}(f)$ estimates with st. dev. (error bars) per center frequency for station (a) HVC and (b) station FNF. Different line styles and shadings indicate tested lapse times of 1, 1.1, 1.2 and 1.5 times t_S.

Excluding 1.5 t_S lapse time, at both stations no substantial variations can be observed. With respect to observations at the very high frequencies with larger lapse times (less measurements) and with the intention of not including the direct S-wave source information but conserving the scatter information of the early coda, we decided to use a lapse time of 1.1 t_S.

3.4.6 Total coda length

Due to the existent fast amplitude decay, it was difficult to vary the total length of coda analysis window. We tested a length of 10 s duration and compared the results with a shorter length of 6 s. Smaller total coda lengths could not be tested which is explained by including events of magnitudes up to ~ 3 that do not provide enough oscillations for the analysis. On the other hand, using a larger coda length than 10 s is not advisable for our dataset as there is a risk of including further events (e.g. buried in the coda signals aftershocks and other events).

At all station sites, we observe lower $\overline{Q_C}(f)$ estimates using a 6-seconds-long total coda length (Fig. 3.7). It also seems that $\overline{Q_C}(f)$ estimates for more distant stations such

as FNF show larger differences by using different total coda lengths (Fig. 3.7b).

It is assumed that a longer total coda length also samples a larger volume of the underground (Mukhopadhyay et al., 2008). Because the whole coda includes scattering information, and to get as much information about the scattering attenuation at The Geysers as possible, we chose the longest feasible total coda length of 10 s.

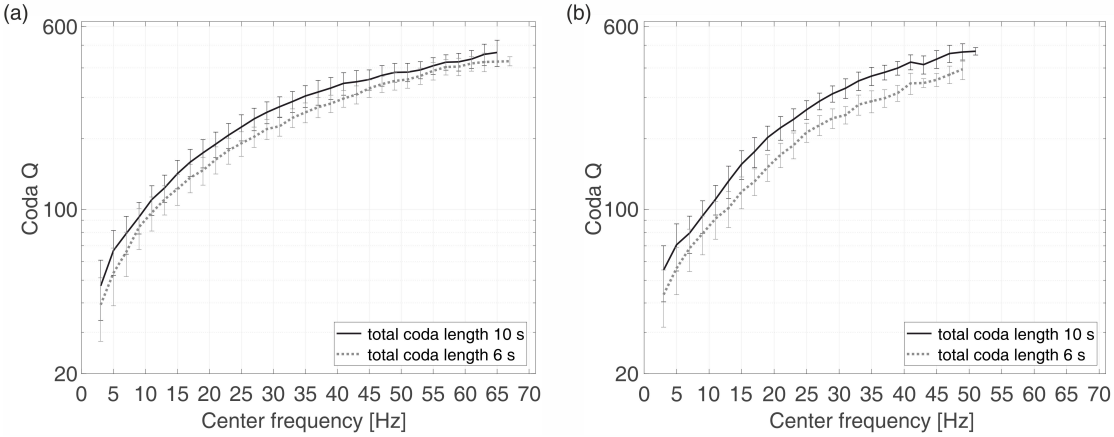

Figure 3.7: $\overline{Q_C}(f)$ estimates with st. dev. (error bars) per center frequency for (a) station HVC and (b) station FNF. Different line styles indicate tested total coda lengths of 6 and 10 s duration.

The preferred parameterization of our sensitivity analysis is summarized and stated in Table 3.2. We mention here that we use the same parameterization for both analyzed event sites.

Table 3.2: Tested Parameters of Sensitivity Analysis.

Parameters	Tested Values[†]
Signal-to-noise ratio	$\geq$ 5,000; $\geq$ **10,000**
$2\sigma\left(Q_C(f)\right)^*$	$\leq$ 10; $\leq$ **30**; $\leq$ 60
Window width (samples)	**128**; 256
Taper (samples)	**8**; 16
t_S (s)[‡]	1.0 t_S; **1.1** t_S; 1.2 t_S; 1.5 t_S
Coda length (s)	6; **10**

[†] Bold numbers highlight preferred parameterization used for the final $\overline{Q_C}(f)$ estimates.
[*] Coda Q uncertainty within 95 % confidence interval.
[‡] Lapse time.

3.5 Results of Coda Q Analysis

The final $\overline{Q_C}(f)$ curves (see also tables S1 and S2 of Blanke et al., 2018) of all stations and of both datasets are shown in Figure 3.8. The curves are shaded in accordance with station location in the northwestern (faint dotted lines) and southeastern part (dark solid lines) of The Geysers. Stations with larger epicentral distances to the respective dataset are characterized by much shorter curves due to a lower SNR.

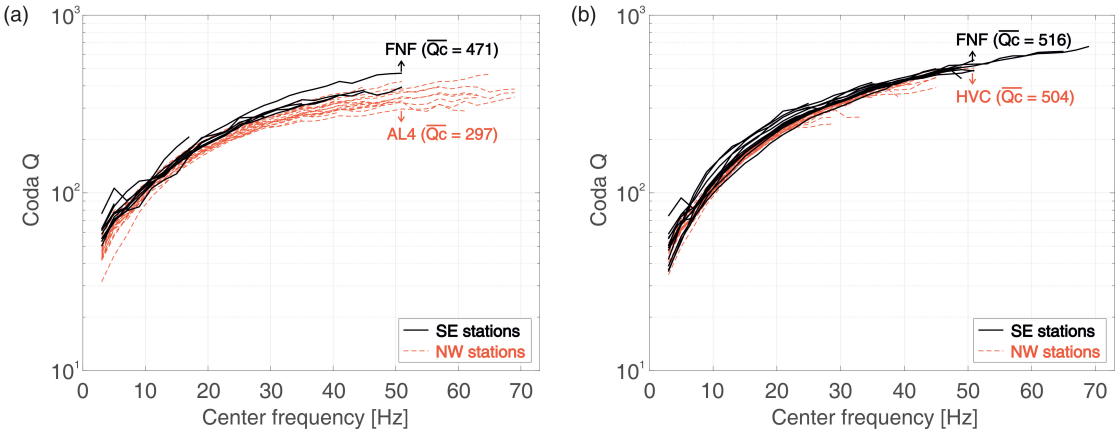

Figure 3.8: $\overline{Q_C}(f)$ curves of available stations at The Geysers with center frequencies of octave-width frequency bands. Stations located in the northwest are shown as faint dashed lines and southeastern stations are displayed as dark solid lines. (a) $\overline{Q_C}(f)$ cures of the northwestern cluster; (b) $\overline{Q_C}(f)$ curves of the southeastern events. At 52 Hz, $\overline{Q_C}(f)$ estimates show the differences between southeastern and northwestern stations for the (a) northwestern cluster and (b) southeastern events.

Using the northwestern cluster, it is clearly visible that estimates of stations towards the southeast exhibit on average higher values over the whole frequency range than stations in the northwest (Fig. 3.8a). At approximately 30 Hz, the curves start to separate from each other and show larger variations of $\overline{Q_C}(f)$ through the entire study area towards southeast as it is shown for stations FNF and AL4 in Figure 8a at 52 Hz. In contrast, $\overline{Q_C}(f)$ obtained from southeastern events (Fig. 3.8b) display comparable estimates over the whole frequency range. For lower frequencies the spread of $\overline{Q_C}(f)$ between all stations is quite narrow for both event locations (Figs 3.8a, b). However, it is obvious that $\overline{Q_C}(f)$ estimates received from southeastern events exhibit much larger values at the very high center frequencies (up to approx. 50 % larger estimates). Station FNF shows a $\overline{Q_C}$(52 Hz) of 471 by using the northwestern cluster and a $\overline{Q_C}$(52 Hz) of 516 using the events from the southeast.

Coda Q studies usually provide a power law for estimated frequency-dependent

Table 3.3: Fitted Power Law* $Q_C = bf^n$ for different parts of The Geysers.

	Northwestern Cluster		Southeastern Cluster	
Station	b	n	b	n
ACR	23.39	0.70	16.04	0.87
AL1	21.35	0.72	16.85	0.83
AL2	21.11	0.74	19.18	0.82
AL3	16.34	0.80	16.72	0.83
AL4	21.29	0.69	16.73	0.83
AL5	19.50	0.74	16.70	0.84
AL6	20.70	0.75	17.26	0.85
BRP	18.52	0.75	15.06	0.87
BUC	11.28	0.89	10.22	1.00
CLV	17.28	0.79	11.74	0.96
DES	35.81	0.52	12.14	0.97
DRK	19.64	0.76	18.51	0.83
DVB	32.61	0.60	14.44	0.98
EPR	16.72	0.95	10.32	1.04
FNF	17.26	0.84	18.65	0.85
FUM	21.90	0.73	18.22	0.84
HBW	17.86	0.77	15.51	0.86
HER	19.22	0.76	17.60	0.82
HVC	17.97	0.78	14.19	0.89
JKR	22.93	0.71	15.84	0.89
LCK	25.71	0.67	18.55	0.82
MCL	19.03	0.73	14.09	0.90
MNS	27.37	0.63	16.76	0.83
NEG	20.39	0.76	11.85	0.97
PFR	23.96	0.70	17.14	0.89
RGP	19.88	0.73	19.09	0.80
SB4	16.14	0.79	15.89	0.86
SQK	15.91	0.80	12.51	0.95
SSR	21.65	0.74	15.04	0.93
STY	27.00	0.62	18.12	0.82
TCH	26.93	0.58	12.49	0.96

* valid for our Q_C estimates starting from 3 Hz. Curves were extrapolated to 1 Hz.

attenuation curves in the style of

$$Q_C = Q_0 f^n, \tag{3.5}$$

to make single studies comparable. Q_0 is Q_C at 1 Hz, f is the center frequency and n is a

constant describing the frequency relationship (e.g. Aki, 1980b; Mukhopadhyay et al., 2008; Morozov et al., 2008; Havskov et al., 2016). Because we receive coda Q estimates starting at 3 Hz, our given Q_0 at 1 Hz is only an approximation fitted to the power law. To highlight that Q_0 reflects an adjusted value, we call it b instead (Table 3.3).

To assess the possible reasons for observed discrepancies between the stations and two event location sites, we present our $\overline{Q_C}(f)$ estimates in different diagrams related to the idea of a potential dependence on distance or azimuth (Figs 3.9 and 3.10).

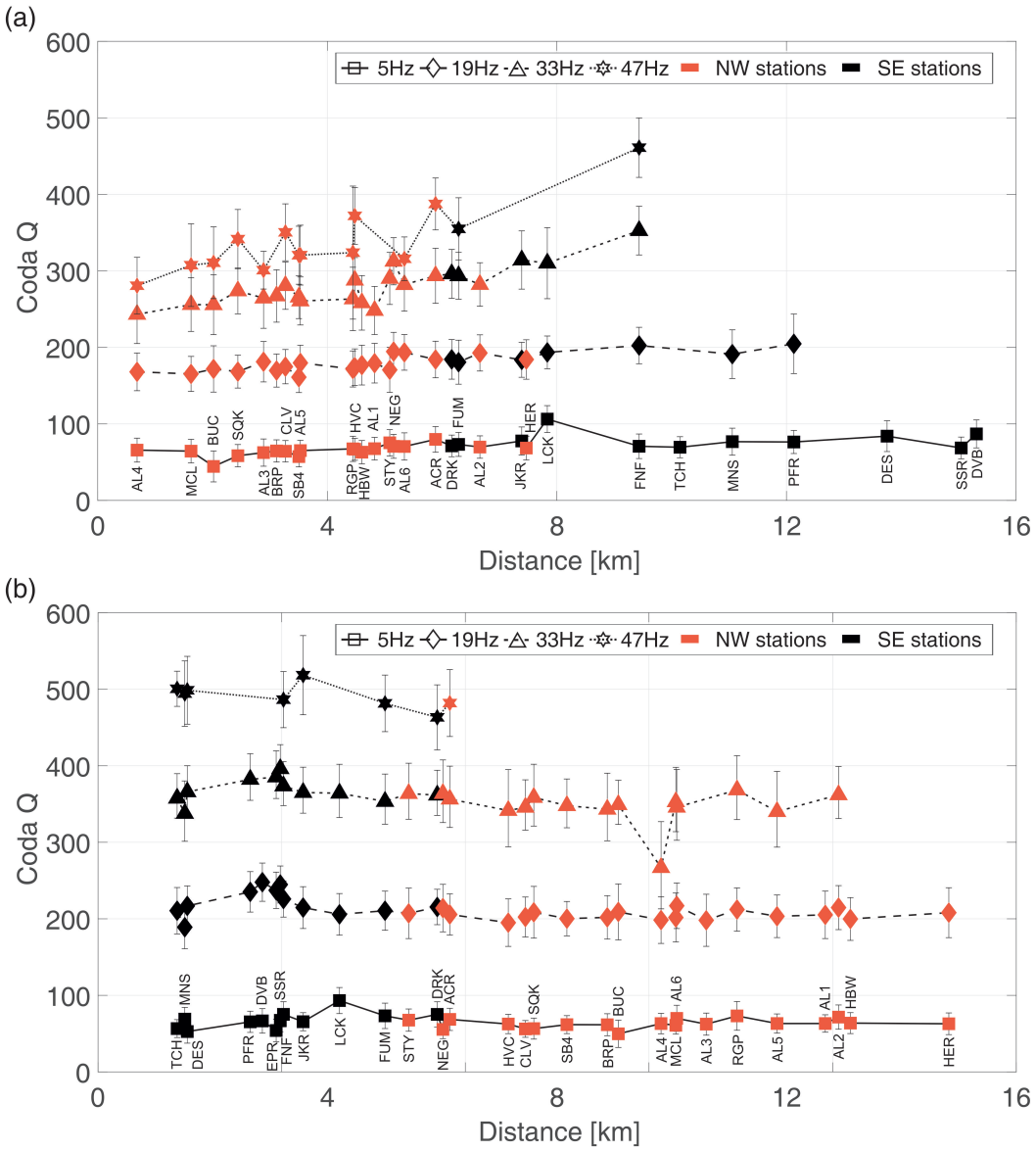

Figure 3.9: $\overline{Q_C}(f)$ estimates with st. dev. (error bars) of selected center frequencies at each station ordered by distance to the center of event locations. Symbols are related to selected center frequencies of 5, 19, 33, and 47 Hz. Bright symbols indicate northwestern stations. Dark symbols indicate southeastern stations. (a) Results of the northwestern cluster and (b) results of southeastern events.

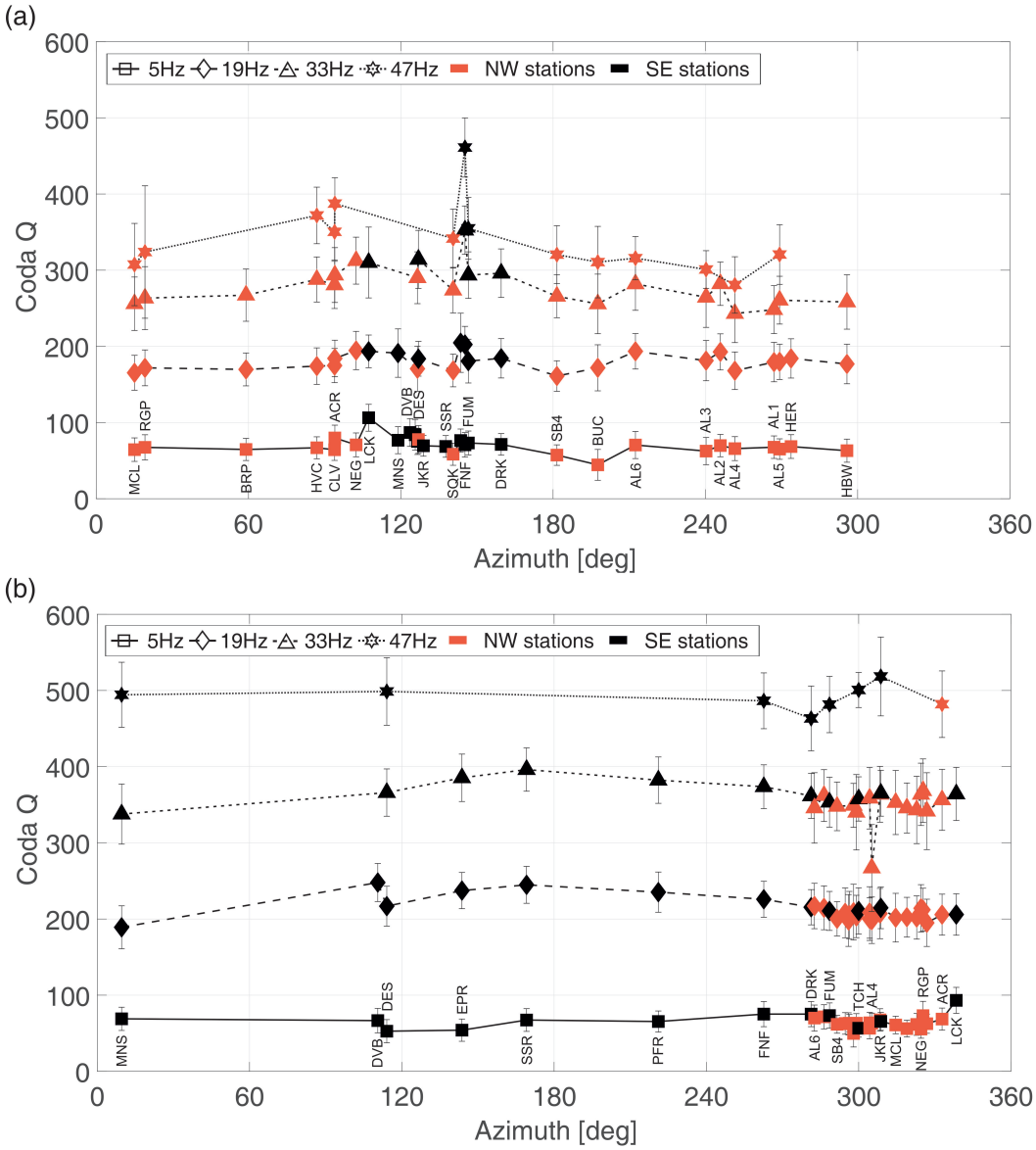

Figure 3.10: $\overline{Q_C}(f)$ with st. dev. (error bars) of selected center frequencies at each station ordered by azimuth angle to the center of event locations. Symbols are related to selected center frequencies of 5, 19, 33, and 47 Hz. Bright symbols indicate northwestern stations. Dark symbols indicate southeastern stations. (a) $\overline{Q_C}(f)$ estimates obtained from the northwestern cluster and (b) $\overline{Q_C}(f)$ estimates of southeastern events.

Figure 3.9 shows $\overline{Q_C}(f)$ estimates from all stations for four selected center frequencies (5, 19, 33 and 47 Hz) as a function of distance to visualize variations in coda Q between all stations within different frequency bands. For the lowest center frequency (5 Hz), our $\overline{Q_C}(f)$ estimates only show a very slight increase with distance for the northwestern cluster (Fig. 3.9a). However, few stations such as BUC and LCK show lower or higher $\overline{Q_C}(f)$ values in between. Looking at the higher frequencies (19, 33 and 47 Hz), the increase with distance is more significant for the northwestern cluster but still some

$\overline{Q_C}(f)$ values at some stations have lower or higher quality factors in between (e.g. at 47 Hz stations AL3, RGP, AL6, etc.). Conversely, we see a quite stable course of $\overline{Q_C}(f)$ estimates with a small jump in $\overline{Q_C}(f)$ at around 4 km for the southeastern events (Fig. 3.9b). These observations led us to also plot the quality factors for the same frequency bands with azimuth angles from the center of the respective event location site (see Fig. 3.10).

Looking at the azimuthal distributions (Fig. 3.10), it becomes obvious that $\overline{Q_C}(f)$ estimates increase at azimuths between 100° and 150° for the northwestern cluster at all center frequencies. Combining the observations of Figures 3.8–3.10, it seems that for events in the northwest of The Geysers, coda Q increases towards southeast with distance from the cluster center. The lowest $\overline{Q_C}(f)$ estimates at 33 and 47 Hz (Fig. 3.10a) are observed for stations north-northeast to northwest of the northwestern cluster at azimuths between 240° and 50°. Looking at $\overline{Q_C}(f)$ estimations of the southeastern events (Fig. 3.10b), we also see a slight increase toward the south around 170° for 19 and 33 Hz center frequencies. Stations located between azimuths of 270° and 360° do not show a preference of $\overline{Q_C}(f)$ estimates into a certain direction. Instead, $\overline{Q_C}(f)$ estimates are close to each other and seem to exhibit randomly minor ups and downs. Only at very high frequencies, a variation in $\overline{Q_C}(f)$ is visible around 300°.

In addition to the spatial analysis, we also tested whether we observe temporal coda Q changes in response to injection operations at the wells Prati-9 and Prati-29 between 2009 and 2015, using only the northwestern cluster. We compared $\overline{Q_C}(f)$ graphs of minor active injection phases (primarily during summer season) and active injection phases (generally during winter season). Because of the small event magnitudes, it was not easy to make an unambiguous comparison of mean curves consisting of data with a time span of only few months. Not all events were recorded consistently at all stations or did not fulfill the quality criteria we applied, which led to unstable or even no estimates in some cases. For very few closer stations, we obtained mean curves as it is shown, for instance, for station SQK in Figure 3.11. SQK was chosen due to the fact that it consistently shows estimates that can be considered as an average of all estimates at all available other recording sites and features a quite stable behavior up to high frequencies. It therefore seems to be an acceptable representative station for this study outcome.

In Figure 3.11a, we plotted a $\overline{Q_C}(f)$ graph for each winter season (active injection periods) between 2009 and 2014 to see if there is a change in $\overline{Q_C}(f)$ over time. The light dashed line is plotted as a reference and represents the general $\overline{Q_C}(f)$ curve of SQK that was already shown in Figure 3.8a. All curves are similar up to 45 Hz and do not indicate an apparent change in $\overline{Q_C}(f)$ over time. At higher frequencies, $\overline{Q_C}(f)$ estimates of winter 2012-2013 feature slightly higher values than the others. The small irregularities in shape

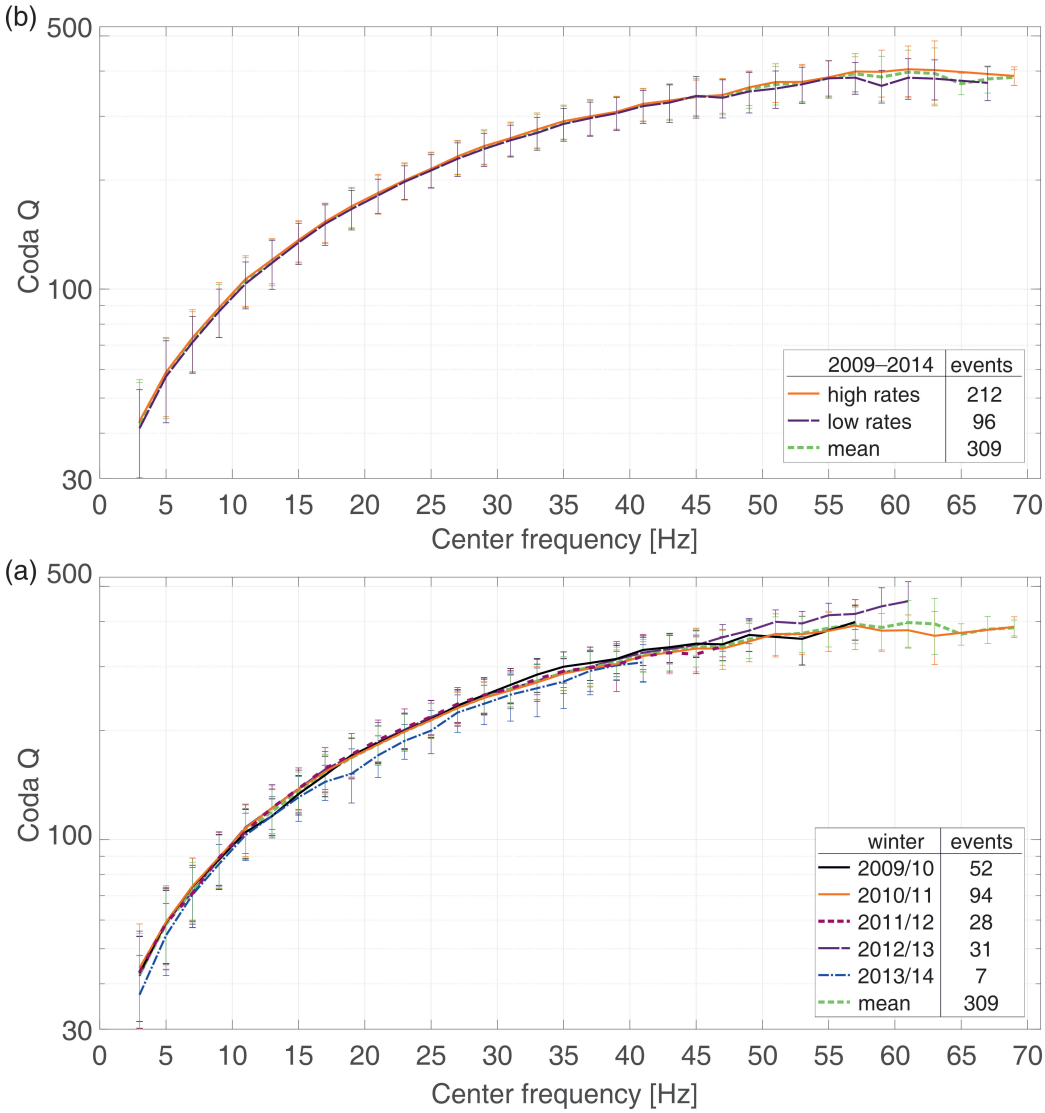

Figure 3.11: (a) $\overline{Q_C}(f)$ with st. dev. (error bars) of all active injection phases between winter 2009 and winter 2014 at station SQK. Different line styles and shadings indicate injection periods. (b) $\overline{Q_C}(f)$ curves for all active injection periods (light solid line) and all less active injection periods (dark dashed lines) between 2009 and 2014 at station SQK. Light dashed line is plotted as a reference showing the $\overline{Q_C}(f)$ values of station SQK already shown in Figure 3.7. Legend shows number of events available for the estimation of plotted curves.

for in general all shown periods and also the not existing mean estimates for winter 2012-2013 and 2013-2014 at the very high frequencies result from only few events that were available for calculating a $\overline{Q_C}(f)$ curve (Fig. 3.8a). Therefore, at this point we have to keep in mind that $Q_C(f)$ estimates at the very high center frequencies are less reliable than at lower frequencies due to higher st. dev. and lower recording quality. Comparing the dashed reference curve with the other curves, all of them follow the same shape and

error bars are overlapping. Because of the low number of available events per winter period we additionally gathered data of all minor active phases (summer periods), and calculated two mean curves for 2009-2014 (Fig. 3.11b). The curves show the same shape with slightly higher $\overline{Q_C}(f)$ estimates for active injection phases (high rates), which is however not very striking because the curves are within the st. dev. Both graphs follow the shape of the general $\overline{Q_C}(f)$ curve (light green) of station SQK.

As a last step, we compared intrinsic S-wave Q (Q_β) estimates obtained by Kwiatek et al. (2015) from spectral fitting with our final frequency-dependent Q_C results (see tables S3 and S4 of Blanke et al., 2018). Both studies use the same dataset of the northwestern cluster what makes a direct comparison interesting. Therefore, we compared $Q_C(f)$ estimates at different single low center frequencies with the Q_β estimates and plotted them together as probability density function bar diagrams in Figure 3.12. Interestingly, we see a match of mean values (at approximately 80) of both analyses while choosing the coda Q estimates at 7 Hz center frequency (Fig. 3.12). The high variation of Q_β estimates (31–587) is noticeable. In comparison, Q_C(7 Hz) estimates show noteworthy smaller variations (11–154).

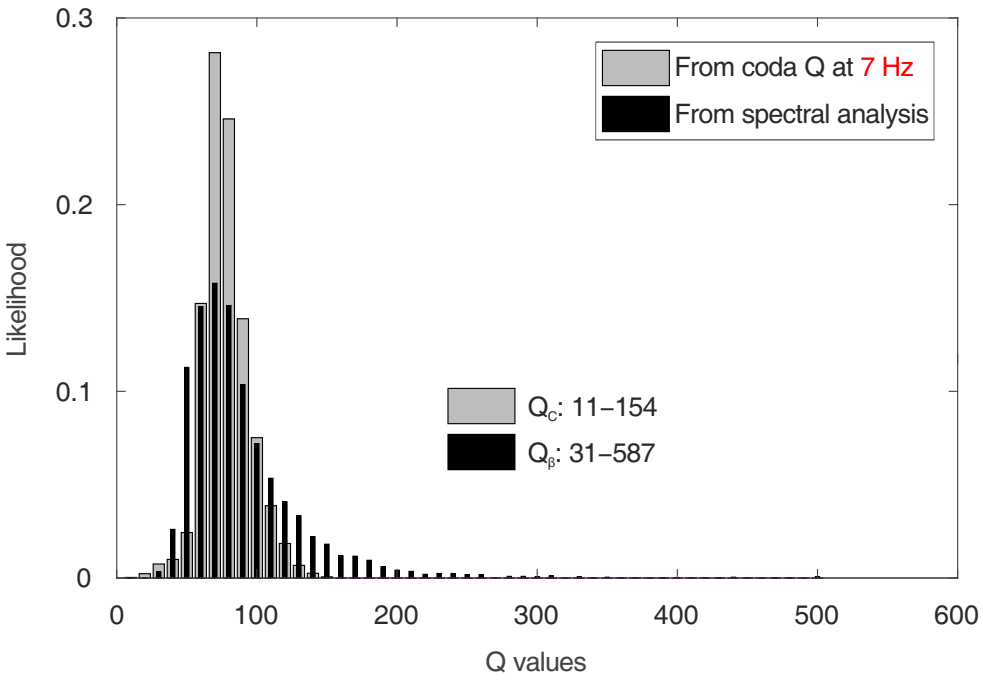

Figure 3.12: Probability density function of Q_C(7 Hz) and Q_β estimates. Wide bright bars show results of coda Q analysis at 7 Hz for the whole study area, estimated from the northwestern cluster. Thin dark bars show results of spectral analysis by Kwiatek et al. (2015), estimated from S-phases. $\overline{Q_C}(7Hz)$ and $\overline{Q_\beta}$ are at approximately 80. Ranges of numbers for Q_C(7 Hz) and Q_β next to the bar plot reflect the variations in Q results for both analysis techniques.

3.6 Discussion

Seismic waveforms recorded at The Geysers are characterized by a fast amplitude decay and very often by a delayed S-wave amplitude peak. These are indicators of a highly attenuating subsurface structure and provoke scattering effects, induced by small cracks, fractures, and features of the geothermal field itself. We used a seismic cluster from the northwest Geysers and 100 selected events from the southeastern part (see Fig. 3.1) to analyze the local S-wave coda to investigate the frequency-dependent coda quality factor based on the approach of Phillips (1985). This method requires a set of different parameters that have to be fixed for the $Q_C(f)$ calculation. For this reason, we performed a parameter sensitivity analysis to have a better control over the procedures and accordingly a more comprehensive view at our findings. We investigated our final coda Q results with regard to its spatiotemporal characteristics by considering the epicentral distance, azimuth, and injection periods. Additionally, we compared our $Q_C(f)$ results with the frequency-independent Q_β estimates of Kwiatek et al. (2015).

3.6.1 Sensitivity of coda analysis parameters

A set of different parameters has to be defined for the coda analysis technique that has been applied here. Therefore, we analyzed the impact of moving window width with related taper, lapse time, sensor components, and total coda length on coda Q. With respect to duration magnitudes of $1 < M_D < 3$, we additionally tested whether all available events can be smoothly used simultaneously for the analysis. Furthermore, to enhance the quality of the waveform recordings, we applied different lower limits of SNRs and upper limits coda Q uncertainties to obtain reliable representative $\overline{Q_C}(f)$ curves for The Geysers.

In particular, the restrictions for the allowed uncertainties seem to have a major impact on the results (see Tab. 3.1). Depending on the accepted $Q_C(f)$ uncertainties, differences in $\overline{Q_C}(f)$ estimates are observed. By setting the maximum acceptable $2\sigma\left(Q_C(f)\right) \geq 30$ and taking the respective single $Q_C(f)$ estimates to calculate $\overline{Q_C}(f)$ values, we obtain smoothed curves up to the very high center frequencies for most of the stations (see Fig. 3.8). Some irregularities still remain at some stations such as, for instance, LCK, SB4 or AL4 (see Figs 3.9a,b), which cannot be smoothed using SNRs higher than 10,000. Therefore, we assume that our final estimates provide also information on very local site effects at the respective station. Because we assume an isotropic scattering model, forward scattering is dominant for wavelengths smaller than scatterer sizes and back scattering would be dominant for wavelengths lager than scatterer sizes (e.g. Sato, 1977). Increased density of cracks, for instance, very close to a station site could cause a drop or

a rise in $\overline{Q_C}(f)$ at certain frequencies related to crack sizes and interacting wavelengths. Thus, nearby stations might not show similar results. Therefore, by using the single station $\overline{Q_C}(f)$ curves instead of one single mean curve for the whole study area, it might not be necessary to pay attention to site effects at all when correcting the seismic source spectra at all recoding sites separately.

In general, the decrease in $\overline{Q_C}(f)$ estimates with smaller accepted uncertainties can be related to single $Q_C(f)$ estimates for smaller earthquakes of magnitudes $1 < M_D < 2$ (see Fig. 3.3). These are predominantly recorded at The Geysers. Single curves partially present a wide spreading of $Q_C(f)$ (see Fig. 3.3a). If this spreading is related to a poor SNR or caused by accidentally including a subsequent earthquake in the later coda portion, some of the estimated $Q_C(f)$ values may only just fell into the accepted restrictions and change the shape of $\overline{Q_C}(f)$ curve. These distorted estimates will be excluded when using tighter restrictions to the SNR level and uncertainty level acceptance. Consequently, $\overline{Q_C}(f)$ values will slightly decrease at the highest frequencies.

We found that magnitudes of $1 < M_D < 3$ do not influence our $\overline{Q_C}(f)$ determination. However, it is advisable, as an outcome of the quality analysis, to accept only high SNRs because it is possible that the results will be biased by outliers or extreme estimations that apparently may occur at the microseismic scale.

In many other coda studies it has been observed that Q_C increases with lapse time (e.g. Mukhopadhyay et al., 2008; Calvet and Margerin, 2013; Havskov et al., 2016). In case of The Geysers, we also find an increase of $\overline{Q_C}(f)$ with increasing lapse time, especially for more distant stations (see Fig. 3.6), but not as remarkable as it is usually reported. As a negative side effect, larger lapse times lead to deterioration of our estimates at higher frequencies and for distant stations. This is explained by the faster achievement of noise level for higher frequencies with time. The very fast amplitude decay let us assume that we do not include the source itself even if we start early in the coda. The decision of using a lapse time of 1.1 t_S is therefore justified by the latter assumption. Rautian and Khalturin (1978) found that coda envelopes of stations at different epicentral distances start to overlap after two times the S-onset time or even later. In case of The Geysers, the conformity of single sensor components (see Fig. 3.4) as well as a certain similarity of $\overline{Q_C}(f)$ curves at smaller center frequencies for the northwestern cluster and even up to the very high-frequency bands for the southeastern events (see Fig. 3.8) also confirm that starting at 1.1 t_S is reasonable here.

The total coda length was, however, difficult to test due to the limited lengths of coda signals. We confirm to see at least a slight dependence on coda Q, in particular at the distant stations. We tested two end-member cases of 6 and 10 s lengths and see a decrease in $\overline{Q_C}(f)$ estimates using a 6 s total duration (Fig. 3.7). By choosing a 6 s total coda

length, the decrease of amplitudes, from lapse time to the end of the defined length, is in total faster and hence steeper compared to a 10 s coda duration, in which the later smaller and much slower decreasing coda part is included. Therefore, using a longer total coda length overall leads to flattening of regression lines and higher $Q_C(f)$ estimates in general.

Because it is assumed that a larger volume of the study area is sampled by using a longer coda duration, we consequently decided to use a 10 s coda duration. Thus, we include more information about local attenuation properties at The Geysers. We conclude that this parameter influences the analysis of The Geysers data, but keep in mind the limited test conditions.

The selection of a suitable moving window width determines how many oscillations of the recorded signal we analyzed at a time, and it can therefore influence the slope of regression lines in the analysis. We irregularly observed slightly higher $\overline{Q_C}(f)$ estimates for decreased window widths (e.g. Fig. 3.5a). Based on the fact that most of the analyzed events from the northwest and southeast are characterized by small duration magnitudes of M_D 1–2 and with respect to the sampling rate of 500, we found that 128 samples per window are enough. For larger event magnitudes ($M_D > 3$), we suggest to choose a respective larger window width.

Overall, the sensitivity analysis results emphasize the relevance of parameterization. In case of The Geysers, lapse time, SNRs, and coda Q uncertainty restrictions must be handled with caution to obtain reliable Q_C estimations. Knowledge about magnitudes, sensor characteristics, and recording qualities are very helpful for the parameters determination.

3.6.2 Influence of injection rate changes on coda Q in the northwestern area of The Geysers

Two wastewater injection wells, Prati-9 and Prati-29, were in operation between 2007 and 2014 (Kwiatek et al., 2015) leading to the occurrence of seismicity in the northwestern cluster (see Fig. 3.1). We investigated the influence of injection rate variations on Q_C. The intensification of fluid injection (up to 200 l/s in both wells) coinciding with the peak in seismic activity suggest a development of the local fracture network in winter months, and a decrease of Q_C is expected to occur. However, careful analyses of low- and high-injection seasons suggest no difference in $\overline{Q_C}(f)$ estimates. We also do not see any long-term temporal decrease of $\overline{Q_C}$ values (Fig. 3.11a and b). Accordingly, the medium's fracture network seems to be persistent over time on average. On the other hand, it might be the case that provoked fractures are too small to be captured by the

frequencies we analyzed.

3.6.3 Comparison of northwest- and southeast-derived coda Q Results

The resulting $\overline{Q_C}(f)$ curves of Figure 3.8 (see also tables S1 and S2 Blanke et al., 2018) feature significant differences in $\overline{Q_C}(f)$ estimates for the northwestern cluster. Taking into account the results of distance and azimuth analyses, $\overline{Q_C}(f)$ increases toward southeast. This is in accordance with the orientation and uplift of the basement felsite. Results obtained from southeastern events show a higher conformity among each other (Fig. 3.8b). Also much larger $\overline{Q_C}(f)$ estimates (starting to rise faster at 17 Hz up to the highest frequency bands) were obtained by using the southeastern events. For both datasets, the highest values are received for station FNF (compare with Fig. 3.8) where the felsite approximately reaches its maximum uplift (Hartline et al., 2015).

A small similarity between results of the northwestern cluster and southeastern events is observed in the azimuthal plots (Fig. 3.10). Also the curves of southeastern events show at least a very small increase towards the southernmost stations. On one hand, this can be explained by the location at the edge of the active operational geothermal field. This area might not be as fractured as the northwestern part. On the other hand, the felsite uplift reaches the southern edge of the geothermal field (see Hartline et al., 2015) and high $Q_C(f)$ estimates might be related to this feature. A change in $\overline{Q_C}(f)$ with distance from southeastern events towards the north is not visible.

Merging all observations received by analyzing both datasets, we conclude that the felsite influences seismic waves on their way to the surface (see Gritto et al., 2013) and therefore, also Q_C estimations. Events triggered in the northwest and also recorded in the northwestern part of The Geysers travel through higher fractured and metamorphic media. Waves reaching the southeastern study area travel deeper and are influenced by the higher quality factor of the uplifted felsite. Therefore, in the southeast, seismic waves are not as much affected by inhomogeneities as waves that directly reach the earth's surface in the northwest (see Gritto et al., 2013). Consequently, coda Q reaches higher estimates calculated at the southern stations than at the northwestern stations.

The felsite might also serve as an explanation for coda Q observations obtained by the analysis of southeastern events. Because of the felsite uplift towards the southeast close to the surface, the events consequently occur in a more homogeneous medium. This feature explains the higher and more consistent $\overline{Q_C}(f)$ estimates. Here, seismic waves mainly propagate directly through the felsite to the surface which causes the higher quality factors. Seismic waves triggered in the southeast and traveling towards the north

also pass through the felsite for quite a long path till they reach the surface. This effect gets obvious in Figure 3.8.

In general, the observations indicate that low-frequency waves are not as influenced as high-frequency waves. This is confirmed by comparing $\overline{Q_C}(f)$ estimates at low frequencies (e.g. Fig. 3.8). Coda Q seems to be stable for low frequencies up to approximately 17 Hz at both analyzed earthquake sites in the study area. Therefore, we conclude that larger wavelengths do not significantly interact with different lithologies and the local fracture network at The Geysers. Accordingly, small variations in coda Q curves (e.g. Figs 3.9 and 3.10) observed at different stations and higher center frequencies, looking at both event sites results, can be interpreted as local site effects caused by different, very local fracture network characteristics.

For future studies, we recommend to take into consideration the subsurface geology for the analysis of high-frequency waves at increasing epicentral distances. S-waves are sensitive to changes in the medium at high frequencies, and hence it is important to be aware of the local geology for an appropriate interpretation of Q_C estimates.

3.6.4 Frequency-dependent Q_C versus Frequency-Independent Q_β

The advantage of using the coda of local earthquakes is the elimination of the influence of the source itself. It has been reported that parameters obtained from seismic codas have lower uncertainties than received from direct phases (e.g. Mayeda and Walter, 1996; Yoo et al., 2011). In this study, we compared $Q_C(f)$ estimates at low center frequencies with Q_β obtained from direct S-phases by Kwiatek et al. (2015), using the same dataset of the northwestern cluster. The extremely fast amplitude decay (high-attenuation properties), crack-induced anisotropy, S-wave splitting, and sometimes observed delayed S-wave peaks made us suspicious whether the calculated assumed frequency-independent Q_β really reflects intrinsic Q. In addition, the previously listed observations and also the very local anomalies noted at different stations and different frequencies (e.g. Fig. 3.10) suggest that in case of The Geysers coda Q is predominantly caused by scattering attenuation.

Our comparison reveals that $\overline{Q_\beta}$ corresponds very well to our $\overline{Q_C}(f)$ estimates received at 7 Hz center frequency (see Fig. 3.12, and table S3 Blanke et al., 2018). Because Q_β is estimated by using the same dataset and also S-waves (direct S-wave phases), it should be considered that at least some of the observations like delayed S-wave peaks could reveal that the generally assumed to be unaffected direct S-wave has experienced an influence of subsurface conditions present at The Geysers. Because of the obvious match of Q_β with Q_C(7 Hz) and the above mentioned observations, we are inclined

to consider that also Q_β might reflect scattering Q in case of The Geysers. However, this speculation suggests a new investigative study to resolve responsible processes at The Geysers to understand and better explain the local energy dissipation in a highly attenuating geothermal study area.

The second interesting observation is the extreme spreading of Q_β values (see Fig. 3.12). Q_β varies between 31 and 587, whereas Q_C(7 Hz) shows a smaller spread of estimates between 11–154. The larger variation in Q_β estimates might be related to instability of the inversion due to the trade-off between corner frequency and quality factor (Kwiatek et al., 2015). In addition, the spectral fitting method samples a smaller volume due to the analysis of the direct S-phase. On the contrary, the coda analysis samples a larger volume of the underground structure due to its starting point at a later lapse time and the analysis of a larger window of the seismogram. Sampling a larger volume averages the Q of the medium and results in more stable estimates. With respect to the observations of the Q_β and $Q_C(f)$ comparison, we conclude that coda Q is more stable than Q_β which was also reported by some other studies (e.g. Mayeda and Walter, 1996; Yoo et al., 2011).

The conformity of $\overline{Q_\beta}$ and $\overline{Q_C}$(7 Hz) leads us to make assumptions related to seismic source parameters, particularly the corner frequency. Usually, Q obtained by the spectral analysis is used for the correction of the direct seismic source spectra to estimate the seismic moment and corner frequency more accurately. The predominant magnitude range of our analyzed events is $1 < M_D < 2$, which on average is related to corner frequencies around 10 Hz (e.g. Kwiatek et al., 2015). Because of an overlap of frequencies of the direct phase and early coda, the match of $\overline{Q_\beta}$ and $\overline{Q_C}$ could therefore be a hint of a dominant S-wave corner frequency for used events, namely approximately 7 Hz. Nonetheless, the latter assumption of resolving the dominant corner frequency by finding a match of $\overline{Q_\beta}$ and $\overline{Q_C}(f)$ requires further investigations. Nonetheless, the match of $\overline{Q_\beta}$ and $\overline{Q_C}(f)$ confirms that using the more stable coda Q can lead to more stable source parameter estimates and ground motion assessments.

3.7 Conclusion

Applying the approach of Phillips (1985), we used the coda analysis to obtain the frequency-dependent coda quality factor at The Geysers geothermal field, California. For a better control over the parameterization of the moving window method and regression analysis, we first applied a sensitivity analysis to obtain the most stable and reliable $\overline{Q_C}(f)$ curves for each available recording station at The Geysers.

The outcome of our sensitivity analysis reveals that the choice of signal quality criteria,

especially restrictions of Q_C uncertainties, need to be treated with caution. Hints of very local changes in $Q_C(f)$ could be eliminated or extreme estimates could bias the results. We found that a high SNR is required to resolve small anomalies within the $\overline{Q_C}(f)$ curves at some of the stations, not visible at other stations in the vicinity. These local anomalies in mean curves could be interpreted as site effects caused by different, very local fracture network characteristics.

Moving window width and total coda length by contrast only show small variations in $Q_C(f)$, whereas sensor components and magnitude ranges of M_D [1, 3] do not seem to have an influence on the mean estimates.

Our results highlight differences in $\overline{Q_C}(f)$ depending on the event location at The Geysers. The different basement structures in the northwestern and southeastern Geysers seem to influence the scattering quality factor at high frequencies ($f > 17$ Hz). The selected datasets reveal an azimuthal dependence on $\overline{Q_C}(f)$ likely caused by the dominantly felsite basement uplift towards the south. Therefore, a single $\overline{Q_C}(f)$ curve for the whole Geysers area might not be useful. Depending on the respective area of interest at The Geysers, an appropriate $\overline{Q_C}(f)$ estimation has to be considered.

Conversely, we cannot observe a relation between seasonal geothermal production rate activities and coda Q estimates that were analyzed using the northwestern cluster close to injection wells Prati-9 and Prati-29. It is conceivable that the fracture network in the vicinity of both injection sites is on average constant with time. Alternatively, the injection-induced fractures are too small to be captured in the analyzed frequency range. In the latter, a possible dependence on high-frequency coda Q estimates would remain undetected in this study.

From the comparison of our Q_C estimates with Q_β estimates of Kwiatek et al. (2015), we see that Q_C is more stable. In addition, we show that $\overline{Q_\beta}$ is equivalent to $\overline{Q_C}(f)$ at 7 Hz that leads to the suggestion that also Q_β represents scattering attenuation at The Geysers. However, future studies are needed that aim at investigating possible energy dissipation processes at The Geysers. Because of the match of the differently obtained Q types, Q_C and Q_β, and the lower scattering of Q_C estimates we assume that using the more stable coda Q leads to more stable source parameters and ground motion assessments.

4 Stress Drop–Magnitude Dependence of Acoustic Emissions during Laboratory Stick-Slip[1]

Abstract

Earthquake source parameters such as seismic stress drop and corner frequency are observed to vary widely, leading to persistent discussion on potential scaling of stress drop and event size. Physical mechanisms that govern stress drop variations are difficult to evaluate in nature and are more readily studied in controlled laboratory experiments. We perform two stick-slip experiments on fractured (rough) and cut (smooth) Westerly granite samples to explore fault roughness effects on acoustic emission (AE) source parameters. We separate large stick-slip events that generally saturate the seismic recording system from populations of smaller AE events which are sensitive to fault stresses prior to slip. AE event populations show many similarities to natural seismicity and may be interpreted as laboratory equivalent of natural micro-seismic events. We then compare the temporal evolution of mechanical data such as measured stress release during slip to temporal changes in stress drops derived from AEs using the spectral ratio technique. We report on two primary observations: (1) In contrast to most case studies for natural earthquakes, we observe a strong increase in seismic stress drop with AE size. (2) The scaling of stress drop with magnitude is governed by fault roughness, whereby the rough fault shows a more rapid increase of the stress drop–magnitude relation with progressing large stick-slip events than the smooth fault. The overall range of AE sizes on the rough surface is influenced by both the average grain size and the width of the fault core. The magnitudes of the smallest AE events on smooth faults may also be governed by grain size. However, AEs significantly grow beyond peak-roughness and the width of the fault core. Our laboratory tests highlight that source parameters vary substantially in the presence of fault zone heterogeneity (i.e. roughness and narrow grain size distribution), which may affect seismic energy partitioning and static stress drops of small and large AE events.

[1]This research article is published as Blanke, A., G. Kwiatek, T. H. W. Goebel, M. Bohnhoff, and G. Dresen (2021). Stress Drop–Magnitude Dependence of Acoustic Emissions during Laboratory Stick-Slip. *Geophysical Journal International*, 224, 1371-1380, https://doi.org/10.1093/gji/ggaa524.

4.1 Introduction

Since the late 1960s, there is a debate on whether earthquake source parameters, and thus rupture processes, are self-similar across magnitude scales or show a scale dependence. A plethora of studies exists that analyzed stress drop across a range of earthquake magnitudes (e.g. Aki, 1967; Ide and Beroza, 2001; Ide et al., 2003; Allmann and Shearer, 2009; Baltay et al., 2011; Abercrombie, 2013; Cocco et al., 2016). The static stress drop is the difference between initial and final shear stress spatially averaged over the rupture surface (e.g. Aki, 1967; Brune, 1970; Candela et al., 2011; Cotton et al., 2013). A constant stress drop indicates a self-similar source process irrespective of the event magnitude.

The static stress drop is an important source parameter that relates to the energy budget of seismic events (e.g. Kanamori and Rivera, 2006) and also affects near-field ground motions (e.g. Spottiswoode, 1993). Hence, it is an important parameter in seismic hazard and risk assessment and it is commonly used to establish ground motion prediction equations (e.g. Cotton et al., 2013; Huang et al., 2017; Baltay et al., 2019).

It is difficult to estimate stress drop accurately from seismic data mainly due to uncertainties in the assessment of corner frequency, model-dependence and error propagations. Many studies discuss potential stress drop dependence on seismic moment or rupture dimension (e.g. Aki, 1967; Abercrombie, 1995; Ide et al., 2003; Kwiatek et al., 2011; Cocco et al., 2016). Indeed, stress drop estimates averaged across the entire bandwidth of earthquake magnitudes ($-9 < M < 8$) suggest a global average value of a few MPa (e.g. Cocco et al., 2016). However, the individual datasets display varying average stress drops spanning over some decades (e.g. Kanamori and Brodsky, 2004; Allmann and Shearer, 2009; Kwiatek et al., 2011; Cotton et al., 2013; Cocco et al., 2016; Huang et al., 2016). In some cases, observed deviations from constant stress drop may be related to non-physical factors including site-, path-, and/or sensor-related effects, limited recording bandwidth, insufficient station coverage, inadequate sensor-rock coupling, incorrect or specific source model assumptions, and error propagation in parameter estimates (e.g. Ide and Beroza, 2001; Ide et al., 2003). In general, stress drop estimates can be improved by using high-quality seismic data recorded over a wide frequency band and appropriate waveform processing techniques (e.g. Shearer et al., 2019). However, studies clearly indicate a scatter of estimated stress drops ranging from 0.01 MPa to 100 MPa, thus significantly exceeding internal uncertainties originating from non-physical factors. This suggests stress drop varies possibly also due to different physical processes in the earthquake source (e.g. Cocco et al., 2016). Fault properties (e.g. composition, roughness, and geometry) and heterogeneous stress conditions may cause stress drops to vary (e.g.

Candela et al., 2011; Sagy et al., 2007; Cocco et al., 2016; Baltay et al., 2019).

Current data quality of natural and induced seismic waveforms likely limits the resolution of the derived physical parameters governing source mechanics. Since the pioneering work of Brace and Byerlee (1966), laboratory experiments are known to contribute to a better understanding of the physics of natural earthquakes (Beeler, 2006). That is, because the boundary conditions in the laboratory can be controlled and experiments can be repeated. Recent studies, analyzing source and statistical properties of natural and laboratory seismic activity (e.g. McGarr and Fletcher, 2003; Thompson et al., 2009; Goebel et al., 2013; Yoshimitsu et al., 2014), indicate similarity of physical earthquake processes across a broad range of scales. Thus, seismic rupture processes down to the sub-millimeter scale can be studied by analyzing high-frequency acoustic emission events (AEs) (e.g. Bohnhoff et al., 2010).

Only few laboratory studies exist that analyze stress drops of AEs in detail (e.g. Goodfellow and Young, 2014; Yoshimitsu et al., 2014; McLaskey et al., 2014). An exact sensor calibration to determine absolute event magnitudes for AEs is complicated, because sensor response is affected by confining pressure, sensor coupling and orientation (e.g. Plenkers et al., 2011). Consequently, a comparison of absolute AE stress drops in the laboratory with stress drops of natural tectonic earthquakes can be difficult. Nevertheless, stress drop differences between experiments with similar sensors can be compared with much higher confidence. Such a comparison is the main focus of the present study.

We analyzed AE events of two laboratory stick-slip experiments on Westerly granite samples to resolve the impact of fault surface roughness and fault core width (cf. Goebel et al., 2014a) on seismic source characteristics. In general, Westerly granite is chosen for laboratory experiments because it represents abundant crystalline rocks of the continental upper crust. Its isotropic texture (e.g. Goebel et al., 2012) facilitates the analysis of seismic AE data. The recorded AEs cover a large moment magnitude range ($-9 < M_W < -5.6$) compared to previous AE studies mentioned above. Both seismic and mechanical data were measured throughout the experiments. The mechanical data provides a direct estimate of the differential stress released during stick-slip, that is, the mechanical stress drop ($\Delta\sigma_m$) (we distinguish mechanical ($\Delta\sigma_m$) and seismic ($\Delta\sigma_{AE}$) stress drop, based on how the measurement was made, that is, force cell for $\Delta\sigma_m$ vs. AE sensors for $\Delta\sigma_{AE}$). The seismic waveform recordings were clipped during the large slip events, preventing detailed source parameter studies of the associated largest AEs. However, unclipped waveforms were recorded for smaller AE events during the inter-slip periods. These are thought to represent the analogue of earthquakes during the inter-seismic period between two large tectonic earthquakes. We used these waveforms to determine corner frequencies and seismic moments of AE populations prior to larger slip events

within the rock specimen by applying the spectral ratio method (e.g. Ide et al., 2003; Abercrombie and Rice, 2005; Kwiatek et al., 2011; Harrington et al., 2015). Changes in source characteristics of these AE populations are thought to image changes in fault stress state as detailed below. It should be noted that the seismic stress drop ($\Delta\sigma_{AE}$) estimates for AE events are not directly connected to mechanical stress drop $\Delta\sigma_m$. Furthermore, $\Delta\sigma_{AE}$ is primarily a measure of the proportion of high-frequency energy radiation for events of a given magnitude that reflects stress release on the micro-scale using several assumptions about rupture velocity and aspect ratio (e.g. Aki, 1967; Brune, 1970).

4.2 Experimental Setup and Data

AE waveform data was recorded during laboratory triaxial stick-slip experiments performed at room temperature on two cylindrical oven-dried Westerly granite samples (cf. Goebel et al., 2012; Kwiatek et al., 2014c). Westerly granite exhibits grain sizes between 0.05 mm and 2.2 mm with an average grain size of 0.75 mm (e.g. Tullis and Yund, 1977; Goebel et al., 2014b). The experiments were performed in a servo-controlled MTS loading frame equipped with a pressure vessel and samples were placed in a rubber jacket to prevent intrusion of hydraulic confining oil. To monitor AE activity, 16 piezoceramic transducers with a resonance frequency of about 2 MHz were placed in brass casings, which were attached directly to the sample surface and in top and bottom pistons, respectively. Full waveform data was recorded using a transient recording system (DAX-Box, Prökel, Germany) in triggered mode at a sampling frequency of 10 MHz and an amplitude resolution of 16 bits. In addition, horizontal and vertical P-wave velocities were measured at regular 30 seconds intervals during the experiments, using active ultrasonic pulse transmissions. Time-dependent velocity measurements were used to locate AE events and assess the evolution of damage in the sample.

We report on two different experiments: (1) Experiment S12 was done on a saw-cut sample of 105 mm height and 50 mm diameter and cut at 30° to the vertical axis across. The saw-cut surface was roughened with a coarse silicon carbide powder (60 grit). Prior to loading, sample S12 was subjected to a constant confining pressure (P_c) of 133 MPa and then loaded in axial compression (σ_1) using a strain rate of $3 \cdot 10^{-4}$ mm/s. (2) Experiment W5 was conducted on a cylinder (107 mm x 40 mm) with Teflon-filled saw-cut notches of 1.9 cm length at 30° inclination to the vertical axis. Sample W5 was then fractured at 75 MPa resulting in a complex rough fault. The fault was locked by increasing P_c to 150 MPa and subsequently activated in a series of stick-slip events at an axial strain rate of $3 \cdot 10^{-6}$ mm/s.

4.3 Method

4.3.1 Data pre-processing

P-wave onsets of AEs and first P-wave amplitudes were automatically picked using the Akaike information criterion (e.g. Stanchits et al., 2006). AE hypocenters were estimated using a hybrid grid search-simplex algorithm assuming a time-dependent anisotropic velocity model derived from ultrasonic transmission data. The estimated location accuracy of AE hypocenters is about ±2 mm (e.g. Stanchits et al., 2006; Goebel et al., 2014a,b). P-wave amplitudes were corrected for sensor coupling quality using ultrasonic transmission measurements (Kwiatek et al., 2014b). Relative AE magnitudes were estimated as

$$M_{AE} = log_{10}\sqrt{\frac{1}{n}\sum_{i-1}^{n}(A_i R_i)^2}, \tag{4.1}$$

where A_i and R_i are wave amplitude and source-to-receiver distance to the sensor i, respectively (e.g. Zang et al., 1998). Following Dresen et al. (2020), we converted M_AE into the seismic moment and from that we calculated the moment magnitude M_W. In the following, the P-wave amplitudes were used to derive full moment tensors of selected events of highest quality (Kwiatek et al., 2014c; Goebel et al., 2015, 2017).

4.3.2 Data preparation

We analyzed AE waveform data from 10 laboratory stick-slip failures that occurred during testing of samples W5 and S12. To estimate AE static stress drops ($\Delta\sigma_\mathrm{AE}$), we applied a variant of the spectral ratio technique (e.g. Ide et al., 2003; Abercrombie and Rice, 2005) developed and tested for different seismic datasets including volcanic and induced seismicity (e.g. Kwiatek et al., 2011, 2014a, 2015; Harrington et al., 2015). The estimated corner frequencies (f_0) and seismic moments (M_W) were used to calculate relative static stress drops of AEs recorded in both experiments.

We selected a window of 256 (25.6 μs) samples from the AE waveforms starting shortly before the P-onset, removed the linear trend, and applied a two-sided von Hann's taper. Signal and noise windows were transferred to the Fourier domain using the multitaper method of Percival et al. (1993). To meet the quality criteria, we selected spectra with a signal-to-noise ratio of at least 3 within a frequency band (f_min, f_max), with a maximum range between $f_\mathrm{min} \approx 7.8 \cdot 10^4$ and $f_\mathrm{max} \approx 5.0 \cdot 10^6$ Hz. For further analysis, we only considered AE events with at least six available spectra that met the criteria given above.

AE events with a sufficient number of spectra were further quality-constrained. First,

we selected events with a maximum location accuracy residuum ≤ 0.2 ($\pm$ 2 mm hypocenter location accuracy) and for which a seismic moment tensor solution was available (e.g. Kwiatek et al., 2014c). Considering the location uncertainties, we restricted the catalog to AE events occurring inside of the sample to exclude miss-located events. For sample W5, we also restricted AE locations close to the specimen main fault surface to better compare results with sample S12. Finally, we grouped AEs to large AE event populations according to stick-slip cycles and discarded AEs associated with large slip events (cf. Figs 4.2a and c).

4.3.3 Spectral ratio technique

For each grouped AE event population we computed spectral ratios based on linked AE event pairs (e.g. Kwiatek et al., 2011; Harrington et al., 2015). The event linkage was based on three criteria: 1) maximum allowed inter-event distance d (i.e. co-located events), 2) minimum accepted magnitude difference ΔM between pairs of events, and 3) minimum number of AE neighbors (number of links), $N_{\min}$. These quality parameters were coordinated carefully to maximize the empirical Green's function (eGf) criteria (e.g. Shearer et al., 2019). Due to the linking of several AE events, each AE could have numerous eGfs. Such multi-eGf links yielded more stable inversion results. The low AE activity per stick-slip in sample S12 restricted d to 15 mm. The high AE number in sample W5 allowed using $d \leq 3.7$ mm. The minimum magnitude difference for event pairs was $\Delta M \geq 0.3$ for S12 and $\Delta M \geq 0.4$ for W5. Finally, the AE events were linked by connecting each AE to at least five events within d and ΔM, which ensured inversion-stability at moderate computational cost. The inversion problem relied on optimizing the cost function in the form (e.g. Kwiatek et al., 2015)

$$Cost\left(\Psi_{th},\Psi_{obs}\right)=\sum_{i}\sum_{j,k}w_{ijk}\left(f\right)||\Psi_{th}^{jk}\left(f\right)-\Psi_{obs}^{jk}\left(f\right)||_{L1}=\min, \tag{4.2}$$

with

$$\Psi_{th}^{jk}\left(f;M_0^j,M_0^k,f_0^j,f_0^k\right)=\frac{S^j\left(f;M_0^j,f_0^j\right)}{S^k\left(f;M_0^k,f_0^k\right)}=\frac{M_0^j}{M_0^k}\left(\frac{1+\left(\frac{f}{f_0^k}\right)^4}{1+\left(\frac{f}{f_0^j}\right)^4}\right)^{\frac{1}{2}}, \tag{4.3}$$

to find relative seismic moment M_0 and corner frequency f_0 for each event in a population. We used the Boatwright source model (Boatwright, 1978) in equation (3). Here, cubed measured velocities and spectral levels were used to estimate initial seismic moments

M_0. Differences between theoretical and observed spectral ratios, Ψ_{th} and Ψ_{obs}, of two linked AE event pairs (j, k) at sensor i, were weighted equally over all frequencies, using the coefficient w_{ijk}, as the original Fourier spectra $S(\cdot)$ in equation (3) were interpolated to logarithmic domain beforehand. The weights were set to 0 whenever the signal-to-noise ratio from any observed spectra forming the pair (j, k) at station i was below 3 and therefore outside of the common frequency interval $(f_{\text{min1}}, f_{\text{max1}}) \cap (f_{\text{min2}}, f_{\text{max2}})$. The multidimensional inversion problem expressed by equation (2) was solved using a Simulated Annealing approach based on non-stationary Metropolis-Hastings Random Walk (MHRW) algorithm (e.g. Sen and Stoffa, 1995). The application of MHRW algorithm separately allowed calculating AE stress drop uncertainties. Representative samples of displacement spectra uncertainties are shown as heat-maps for small, medium and large AE events in Fig. 4.1.

To obtain results comparable to seismological studies at larger scales, we followed Kanamori (1977) to calculate the seismic moment magnitude:

$$M_W = \frac{log_{10}(M_0 - 9.1)}{1.5}. \tag{4.4}$$

AE stress drop estimates ($\Delta\sigma_{\text{AE}}$) were obtained following Eshelby (1957)

$$\Delta\sigma_{\text{AE}} = \frac{7}{16}\frac{M_0}{r^3}. \tag{4.5}$$

The source radius (rupture dimension) r was estimated from corner frequency measurements assuming the dynamic circular source model of Madariaga (1976):

$$r = \frac{CV_\beta}{2\pi f_0}, \tag{4.6}$$

with C = 1.32. The S-wave velocity V_β was obtained from measured P-wave velocity V_α using:

$$V_\beta = \frac{V_\alpha}{\sqrt{3}}. \tag{4.7}$$

4.4 Results

4.4.1 Mechanical stress drop and AE activity relations

Each experiment resulted in six stick-slip events clearly indicated by large measured stress drops (LSDs) in the differential stress curve (Figs 4.2a and c). The LSDs observed

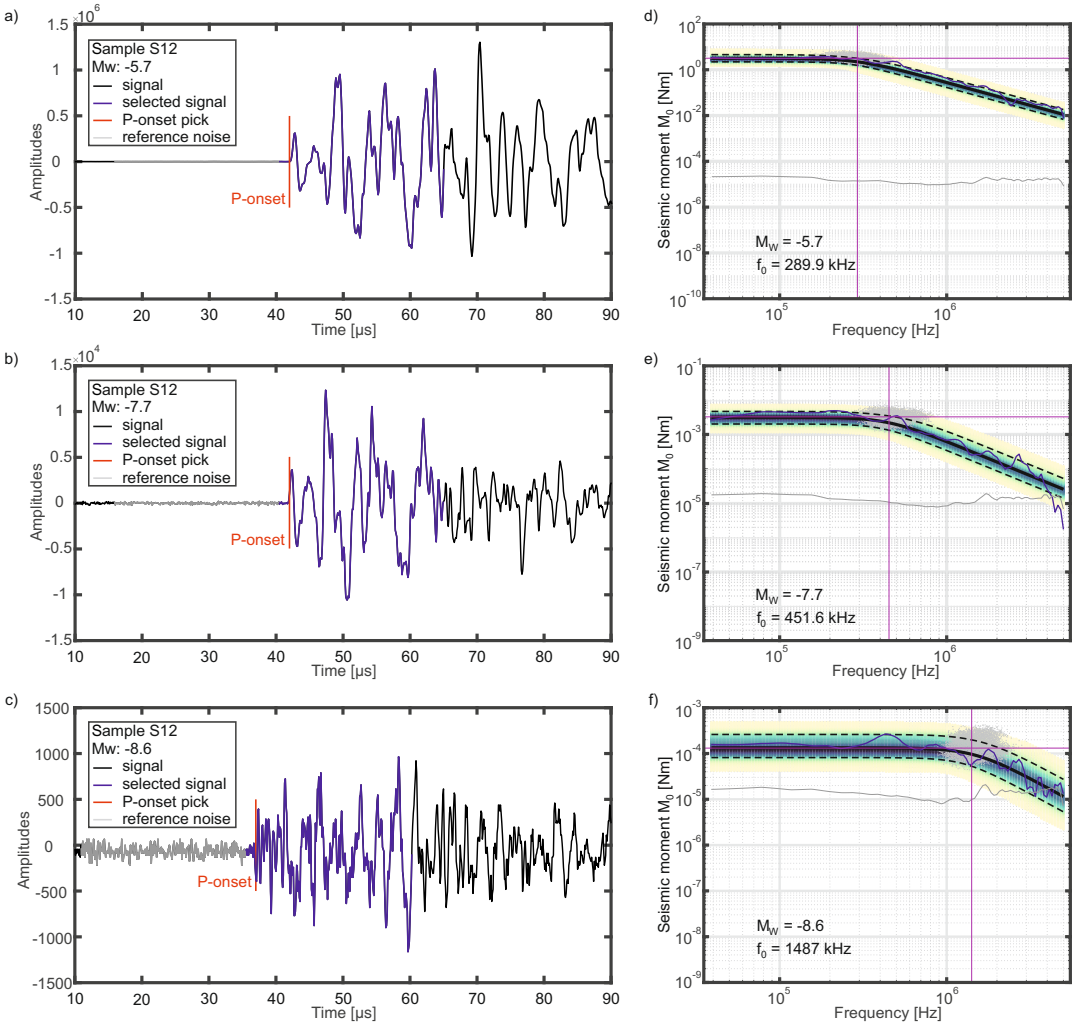

Figure 4.1: AE waveforms of a) a large (M_W = -5.7), b) a medium (M_W = -7.5) and c) a small (M_W = -8.7) AE event. Highlighted are the P-onset pick (red), the selected waveform window (blue) and the reference noise signal (gray) used for the inversion. Subfigures d) – f) visualize the corresponding average observed spectra (blue line) with average noise level (gray line) and the source parameter fitting. Gray dots show the MHRW-sampled corner frequency–seismic moment pairs representing uncertainties of estimated source parameters. Pink solid vertical and horizontal lines indicate maximum likelihood solution. Black thick solid line shows the corresponding optimum spectral fit. Dark and light shaded parts indicate heat map of spectral fits showing optimal and non-optimal fitting areas, respectively. Black dashed lines constrain the area of acceptable spectral fittings (95 % confidence interval).

in sample S12 first increase progressively with multiple slips but then remain constant. The post-slip minimum stress varies. Peak differential stresses of 400 MPa with maximum LSDs of approx. $\Delta\sigma_m \approx 300$ MPa are reached for stick-slips four and five. In contrast, maximum peak stress level for sample W5 stays roughly constant at about 290 MPa (25 % lower than in S12). Here, the LSDs increase progressively as the post-slip

minimum decreases. A maximum stress drop of $\Delta\sigma_m \approx 190$ MPa (~ 35 % less compared to S12) is reached for stick-slip number 5.

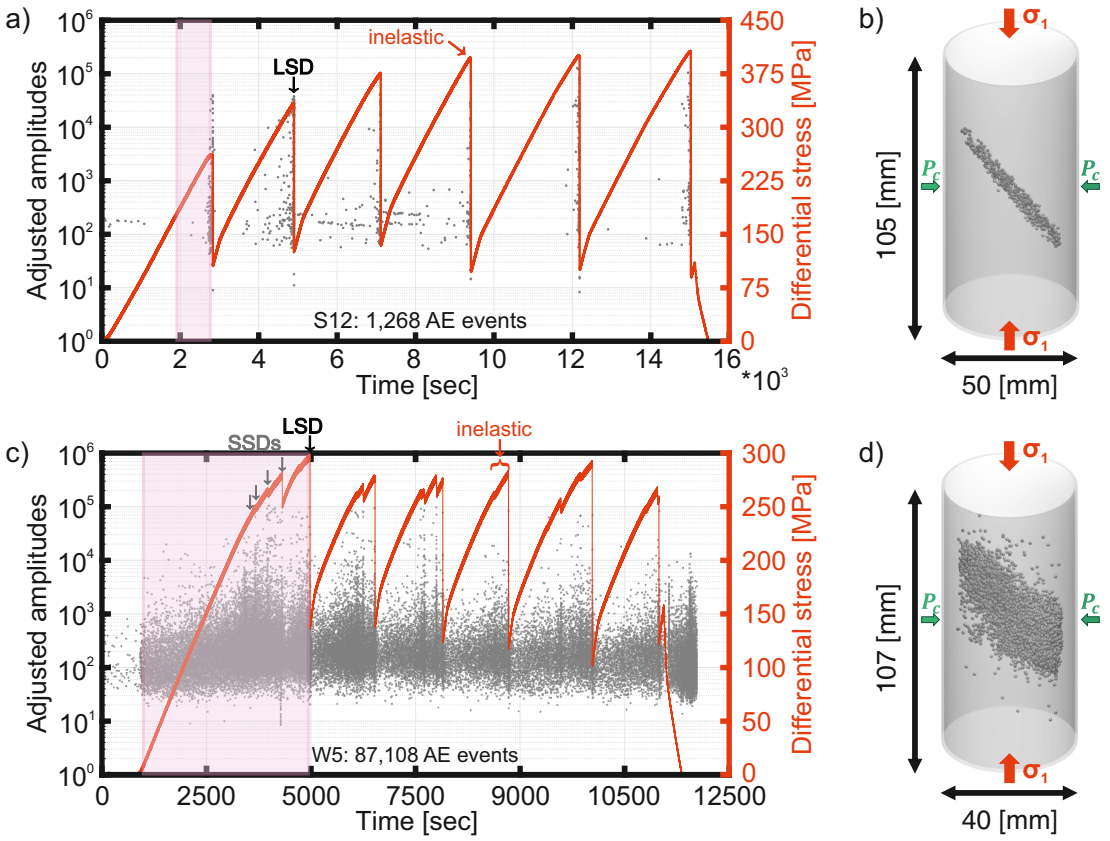

Figure 4.2: Differential stress (red curve) and AE amplitude (gray dots) as a function of time from start of experiment S12 (a) and W5 (c). Examples of inelastic deformation phases are marked and labeled in red. Gray arrows indicate examples of SSDs (small mechanical stress drops) and black arrows indicate examples of LSDs (large mechanical stress drops). Shaded rectangles exemplify event populations used for demonstrating the regression analysis in Figs 4.3a and b. (b) and (d) Spatial distribution of AEs (dots) within samples S12 and W5, respectively. Locations of AEs reflect a simple fault zone for S12 and a complex fault zone for W5. Bold arrows indicate axial stress (σ_1) and confining pressure (P_c).

We observe differences in the AE event occurrence in relation with differential stress and time between both experiments. Sample W5 features a broader damage zone (cf. Goebel et al., 2014a) compared to specimen S12 (Figs 4.2b and d). Beyond yield stress, small mechanical stress drops (SSDs) are frequently observed prior to LSDs of the rough fault (Fig. 4.2c), clearly reflecting the larger fault complexity. Sample W5 shows continuously occurring AEs during the entire stick-slip cycle with an increased AE event rate related to large and small mechanical stress drops and reduced AE activity in the aftershock sequences (cf. Goebel et al., 2015). For this sample, we located a total number

of 87,108 AEs. In contrast, sample S12 shows an increased AE activity only close to peak stress and LSDs without prominent aftershock sequences (cf. Kwiatek et al., 2014c). In total, 1,268 AEs were located along the simple saw-cut fault plane (Fig. 4.2b). The stress-time curve (Fig. 4.2a) shows higher yield points, higher peak stress and no precursory SSDs as found in sample W5.

4.4.2 AE stress drops

We obtained source parameters of 688 AEs for sample S12 and 1,882 AEs for sample W5 from the first five stick-slip cycles. The determined range of AE stress drops is between 0.01 and 100 MPa, which is comparable to the range in other studies (Fig. 4.3a). For sample S12, the estimated M_W and $\Delta\sigma_{AE}$ are slightly larger (-9 < M_W < -5.6 and 0.01 < $\Delta\sigma_{AE}$ < 100 MPa) than for W5 (-9.1 < M_W < -6.6 and 0.01 < $\Delta\sigma_{AE}$ < 10 MPa). However, in contrast to induced and natural seismicity shown in Fig. 4.3a, we observe a systematic increase in $\Delta\sigma_{AE}$ with magnitude (Fig. 4.3a, inset). This dependence of $\Delta\sigma_{AE}$ on M_W is very pronounced in both lab datasets. Our data extend the magnitude ranges analyzed in previous laboratory studies usually spanning only 1-2 orders of magnitude (see Fig. 4.4). Noteworthy are the comparable stress drop estimates between this and previous studies for respective comparable magnitude ranges (Fig. 4.4). We find that the observed M_W–$\Delta\sigma_{AE}$ relation is not affected by bandwidth limitation problems. The resolved corner frequencies are below 50 % of the maximum fitting frequency ($f_{max} \approx 2 \cdot 10^6$ Hz), which is typically assumed to be 80 % of the Nyquist frequency ($f_{Ny} \approx 5 \cdot 10^6$ Hz) (e.g. Ruhl et al., 2017). In addition, even by excluding the smallest events, the observed dependency between M_W and $\Delta\sigma_{AE}$ remains (cf. Figs 4.3b and c).

We performed a least-squares linear regression analysis for both experiments on each individual inter-slip period (Figs 4.3b and c). Data points within a 95 % confidence interval (CI) were considered for the regression analysis and lower and upper bounds of the 95 % confidence interval of the regression slope variation were calculated to visualize its possible variance. The sensitivity of $\Delta\sigma_{AE}$ to changes in M_W increases with consecutive stick-slip cycles, leading to an increase in slope (b) from 1.12 to 1.31 between the first and the last inter-slip period for S12 and an increase from b = 0.82 to 1.26 for sample W5 (Figs 4.3d and e). Consequently, while the average value of b is substantially larger for the smooth fault, the change in b with slip events is a factor of 2 larger for the rough fault. Sample S12 indicates a saturation of slope-growth in the latest stick-slip cycles, analogous to the trend in mechanical stress drop $\Delta\sigma_m$ (Fig. 4.3d). Experiment W5 also shows similar behavior between $\Delta\sigma_m$ and $\Delta\sigma_{AE}$ estimates but does not indicate saturation.

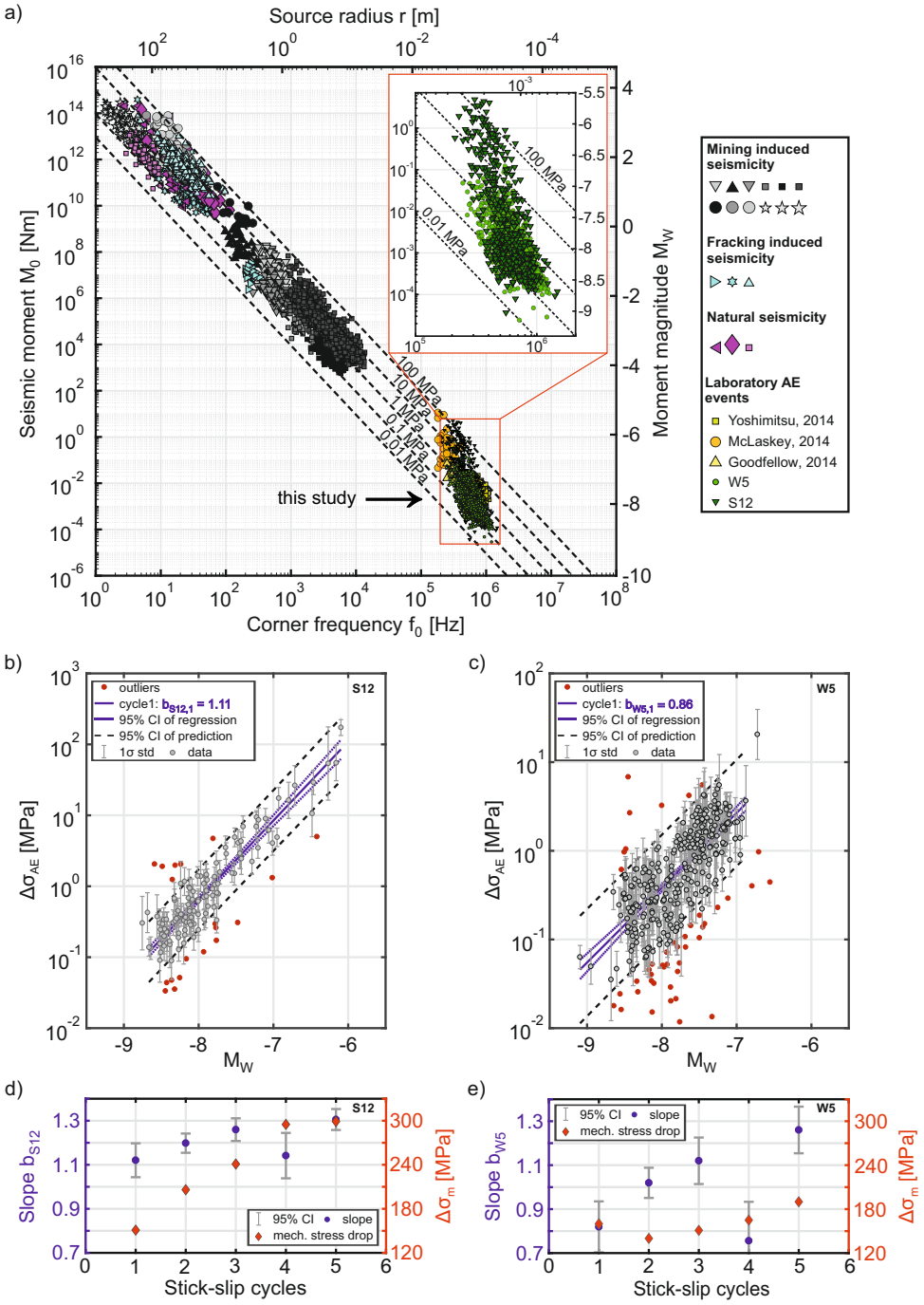

Figure 4.3: (a) Stress drop and source parameter relations, modified after Kwiatek et al. (2011), using the Madariaga (1976) source model. Colors separate 23 different studies of natural-, mining induced- and fracking induced seismicity, and laboratory AE events. This study results (green triangles and circles) are magnified in the inset. Dashed lines indicate constant stress drops of 0.01-100 MPa. (b) and (c) Regression analysis for interslip periods before the first stick-slip events (cf. highlighted rectangles in Figs 4.2a and c) of S12 and W5, respectively. Regression line (solid) is shown with 95 % CI (confidence interval) of slope variance (dotted) and considered estimates used for slope calculation (dashed). (d) and (e) M_W–$\Delta\sigma_{AE}$ slope changes per stick-slip cycle (circles). Diamonds represent mechanical stress drops for each stick-slip event (cf. Figs 4.2a and c).

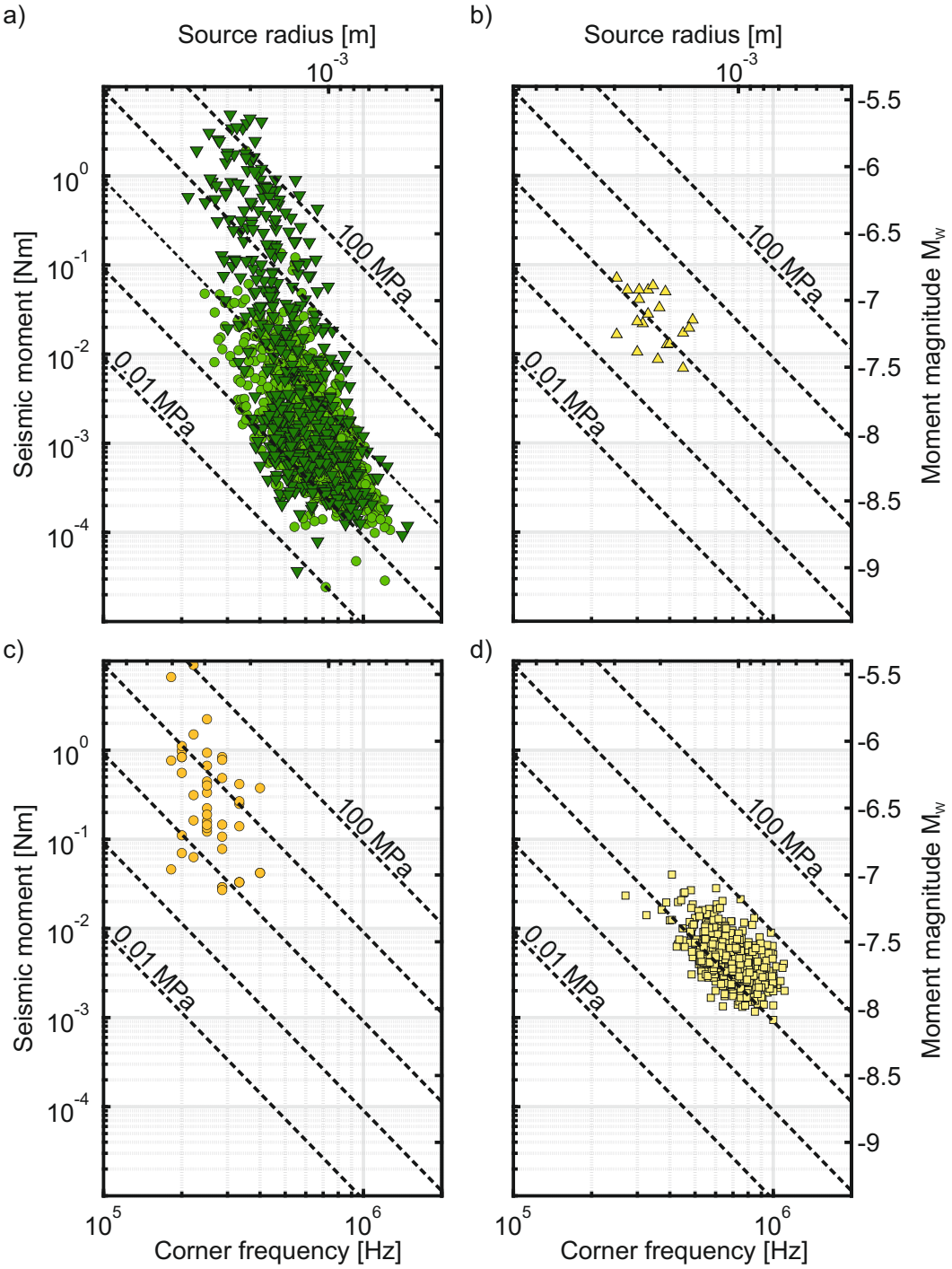

Figure 4.4: Magnification of laboratory AE study results from Fig. 4.3a. Seismic source parameter relations from a) this study with circles indicating results of W5 and triangles showing results of S12, b) Yoshimitsu et al. (2014), c) McLaskey et al. (2014) and d) Goodfellow and Young (2014). Dashed lines indicate constant stress drops ($\Delta\sigma_{AE}$) of 0.01-100 MPa. Note this study (subfigure a) shows a broader span of magnitudes and corner frequencies and a larger number of AE events. A static stress drop dependence on event size is visible. This study covers the magnitude ranges of the other shown laboratory studies. Comparable static stress drop estimates for same magnitude ranges are obtained.

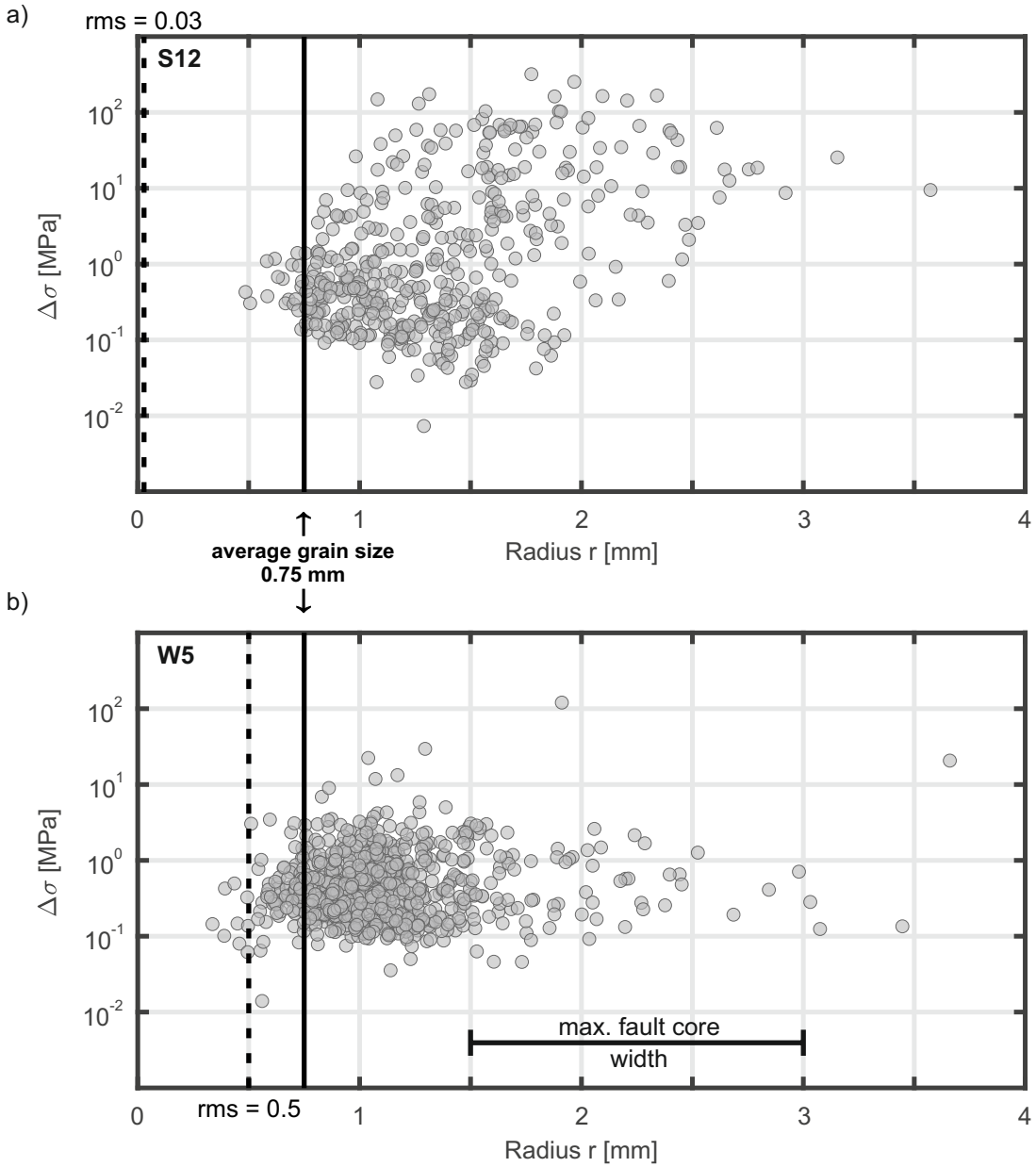

Figure 4.5: Source radii and static stress drop relation of (a) S12 and (b) W5. Vertical solid line indicates average grain size of Westerly granite. Dashed line shows rms of roughness from 2 D white-light interferometry measurement and horizontal bar indicates maximum width of the fault core along the slip direction for sample W5.

The overall size distribution of AEs can be linked to dominant length scales in the experiments. Minimum AE source radii, estimated from equation (4.6), are approximately limited by the average grain sizes of Westerly granite (see Fig. 4.5). Differences in surface roughness expressed as rms of vertical topography (cf. Goebel et al., 2014b) in Fig. 4.5 seem to only play a secondary role in controlling minimum AE size.

The maximum size of events on the rough surface in sample W5 is similar to the width

of the core deformation zone (maximum fault core width = 1.5 to 3 mm), providing here a plausible upper bound for AE event growth. However, AEs on the smooth fault in sample S12 also grow to similar sizes, which are clearly beyond peak roughness wavelength and fault core width on S12. Such large events on S12 also result in substantially higher $\Delta\sigma_{AE}$ and show a dominance of earthquake-like double couple source mechanisms (Fig. 4.6). Thus, while minimum event sizes in both experiments can potentially be linked to grain size distribution, the maximum seismic event size appears more sensitive to fault roughness with smoother faults producing large magnitude events with higher $\Delta\sigma_{AE}$.

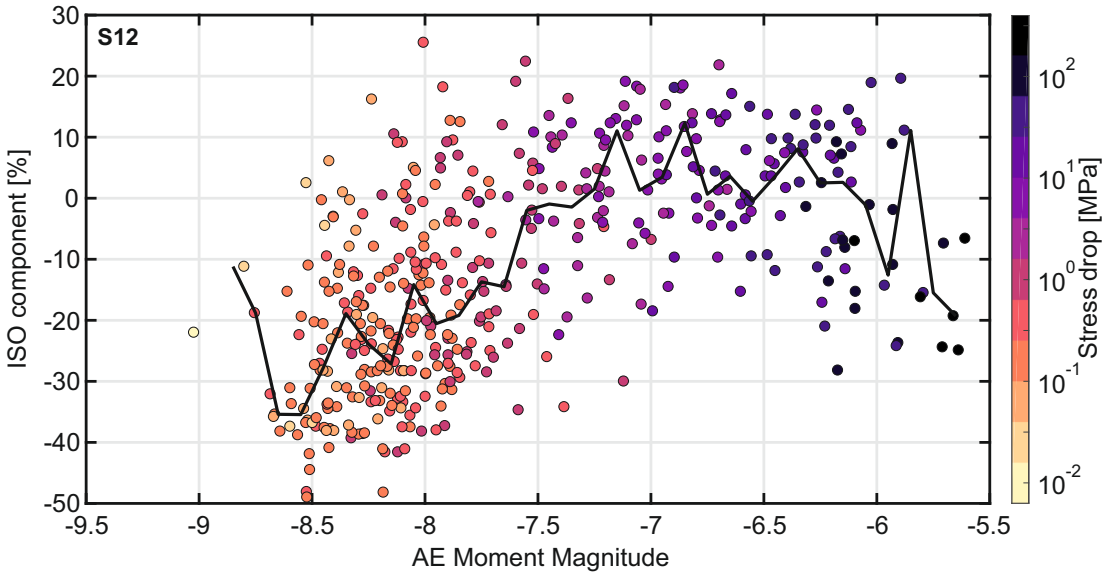

Figure 4.6: Relation of estimated AE moment magnitudes and ISO component percentages of full moment tensor solutions of experiment S12 (modified after Kwiatek et al. (2014c)). Data shown here, are gathered from all stick-slip cycles and color-coded by their stress drop estimates ($\Delta\sigma_{AE}$). Solid line indicates median ISO component estimated per 0.1 magnitude bin. Note the relationship between low stress drop estimates, small AE magnitudes and larger negative amount of ISO component and vice versa.

4.5 Discussion

Whether seismic stress drop of natural and induced seismicity varies with earthquake magnitude or remains constant is still a matter of debate. Although, stress drop estimates fall within a similar range of 0.01-100 MPa for events between M_W -9 and 4 (Fig. 4.3), individual well resolved datasets show variability potentially due to differences in underlying source processes. The present laboratory study clearly highlights a M_W–$\Delta\sigma_{AE}$ dependence (Figs 4.3d and e) beyond the commonly observed range of static stress drop variation (e.g. Abercrombie, 2015) in individual datasets. This is in contrast to a majority

of field studies that found a constant static stress drop independent of event magnitude (e.g. Ide and Beroza, 2001; Prieto et al., 2004). In the following, we discuss potential effects, non-physical and physical, that may control the presented M_W–$\Delta\sigma_{AE}$ relation.

4.5.1 Non-physical effects

Non-physical parameters, related to data quality and analysis, may cause apparent dependence of stress drop on magnitude and errors in source parameter calculations (e.g. Ide and Beroza, 2001; Abercrombie, 2013). The list of potential factors includes large azimuthal gaps, inaccurate correction of attenuation (Q), site effects, a limited magnitude range, bandwidth limitations, or an inaccurate source model.

To minimize the error-term, we followed several instrumental and analytical steps. (1) The laboratory set-up allows us to design a seismic network that prevents large azimuthal gaps. This is accomplished by placing sensors all around the sample and the expected source region and not just along the surface, which is a common limitation for natural events. (2) We use the spectral ratio method that, if applied carefully, eliminates path and site effects. In addition, we explore the effect of using several eGfs for each event, which adds to the robustness of source spectra inversions. (3) Furthermore, our data covers a reasonably large magnitude range (-9.1 < M_W < -5.6), substantially expanding the magnitude span of comparable previous AE studies (e.g. McLaskey et al., 2014; Yoshimitsu et al., 2014). (4) Lower and upper frequency bounds f_{min} and f_{max} define the range of resolvable frequencies of the AE spectra. Considering the careful selection of analyzed AE events including a good signal-to-noise level, a large number of available recorded waveforms and estimated spectra, and the application of multi-eGf links for each analyzed event, we consider a potential bandwidth problem unlikely. Even by excluding the smallest recorded events with magnitudes M_W < -8.4, the clear stress drop–magnitude dependence remains. (5) We use the commonly accepted circular shear source model of Madariaga (1976) to make our study comparable to previous work. However, we obtain stress drop estimates clearly exceeding the general observed variability.

4.5.2 Physical effects

Stress drop may be affected by changes in seismic source processes. Suggested factors influencing these processes address a strongly heterogeneous stress field (e.g. Candela et al., 2011; Wang et al., 2019), e.g. induced by a complex fault geometry (roughness), and changes in rupture velocity (e.g. Kanamori and Rivera, 2004). Rupture and also slip velocity may be affected by structural and material heterogeneities in the fault zone. However, potential effects from these factors have not yet been sufficiently investigated.

Impact of fault surface roughness

The main macroscopic difference between the experiments analyzed in this study is the degree of fault roughness, which in turn leads to notable differences in AE rates, peak stresses, yield point and the extent of the work-hardening (inelastic) phase (Figs 4.2a and c). The smooth fault S12 (Fig. 4.2b), characterized by a simple homogeneous fault zone, indicates strong coupling of the saw-cut blocks. The fault remained fully locked up to almost peak stress, which is reflected by the extended linear mechanical stress increase (Figs 4.2a and b). During the elastic loading phase, AE activity remained very low with precursory activity only occurring after differential stresses exceeded the yield point (Fig. 4.2a). The resulting high peak stresses were released in slip events with LSDs.

In contrast, the rough fault W5 (Fig. 4.2d) exhibits a broader damage zone and an anastomosing network of slip surfaces (Goebel et al., 2012). The higher roughness likely results in relatively poor coupling of the fault blocks. The yield point occurred at lower stresses within the inter-slip periods and AE activity was high during loading and SSDs (Fig. 4.2c). Peak stresses at LSDs were 25 % lower than on the smooth surface. Similarly, $\Delta\sigma_{\mathrm{m}}$ was about 35 % smaller.

The differences due to fault roughness are also reflected in AE source parameters. On the rough surface the maximum AE moment magnitudes ($M_{\mathrm{Wmax}} \sim -6.6$) and stress drops ($\Delta\sigma_{\mathrm{AEmax}} \sim 10$ MPa) are about one order smaller than on the smooth surface ($M_{\mathrm{Wmax}} \sim -5.6$, $\Delta\sigma_{\mathrm{AEmax}} \sim 100$ MPa). The smooth fault S12 shows a weak slope increase of M_{W}–$\Delta\sigma_{\mathrm{AE}}$ regression lines as peak stress and $\Delta\sigma_{\mathrm{m}}$ increase for the first four slip events (Fig. 4.3d). This may be related to conditioning of the saw-cut in combination with increasing formation of fault gouge. That likely causes the larger AE stress drops and greater moment magnitudes. The increase of $\Delta\sigma_{\mathrm{m}}$ and M_{W}–$\Delta\sigma_{\mathrm{AE}}$ regression slopes may saturate once a thin but stable gouge layer covers the slip surface. The rough fault W5 shows an increase of $\Delta\sigma_{\mathrm{m}}$ and M_{W}–$\Delta\sigma_{\mathrm{AE}}$ regression slopes possibly in response to a continuous destruction and smoothing of asperities. With large stick-slip events, asperities may be reduced as the slip planes become progressively smoother.

Interestingly, in both experiments the source radii of AEs are roughly bounded by the average grain size (Fig. 4.5). This suggests that $\Delta\sigma_{\mathrm{AE}}$ is slip-controlled and may not scale with grain size or source radius. We discuss this hypothesis in more detail in the following section.

Structural heterogeneity and grain size effects

Grain scale structural heterogeneities may affect AE source radii that in turn affect static stress drops as expressed in equations (5) and (6). In general, initial crack size

in undeformed rock samples scales with grain size (e.g. Dresen and Evans, 1993). The source radii of AEs observed in both samples range between 0.4 mm and 3.5 mm, which is on the order of grain size variation (Fig. 4.5). We do not observe a significant dependence of stress drop on source radius, with the exception of a weak trend for specimen S12. In contrast, the change in $\Delta\sigma_{AE}$ is large compared to the change in source radius. Therefore, we posit that observed $\Delta\sigma_{AE}$ variations may be mainly attributed to increasing slip over fracture surfaces correlated to grain size along grain boundaries for small events or across multiple grains for larger events (e.g. Candela et al., 2011). This is supported by the observed increase in complexity of source processes for smaller AE events (cf. Kwiatek et al., 2014c) displaying a significant contribution of non-double couple components of AE moment tensors (Fig. 4.6). This again suggests that an increase of the mechanical stress drop, related to progressively increasing slip over the fault surface, translates to microscale fracture propagation and AE source characteristics (cf. Figs 4.3d and e, and 4.5).

Rupture Velocity

Coalescence of cracks during rupture propagation may also result in higher rupture velocities as ruptures accelerate. For example, double-couple (simple shear motion) events occurring in sample S12 show high stress drops (Fig. 4.6). In contrast, smaller events caused by crack propagation along grain boundaries and kinks may display only limited slip and are characterized by larger non-double-couple components. It is conceivable that during complex crack growth more energy is spent in crack surface energy and heat at the expense of radiated energy (Kanamori and Brodsky, 2004). This could cause spatially variable seismic slips (e.g. Mai and Beroza, 2002) independent of source radius but dependent on roughness (cf. Candela et al., 2011).

To test whether changes in rupture velocity (V_r) significantly affect our stress drop observations, we reduced V_r with decreasing M_W, using correction factors provided by Sato and Hirasaw (1973). We first adjusted source radii and stress drops using equations (6) and (5), respectively. We then calculated a mean M_W–$\Delta\sigma_{AE}$ regression slope for each experiment by gathering AEs from all analyzed AE populations. We reduced V_r to 50 % for the smallest AEs. By reducing AE rupture velocities from 90 % to 50 %, V_r shows only a small effect on M_W–$\Delta\sigma_{AE}$ regression slopes for sample S12 (Fig. 4.7). The effect is slightly more visible for sample W5, due to an increase of stress drop for small events (cf. Kanamori and Rivera, 2004). The still existing scaling clearly shows only a limited effect of V_r changes on the M_W–$\Delta\sigma_{AE}$ scaling relation. This suggests that the observed changes in stress drop may be related to complex rupture along a rough slip plane, possibly leading to varying radiated energy and energy partitioning during slip.

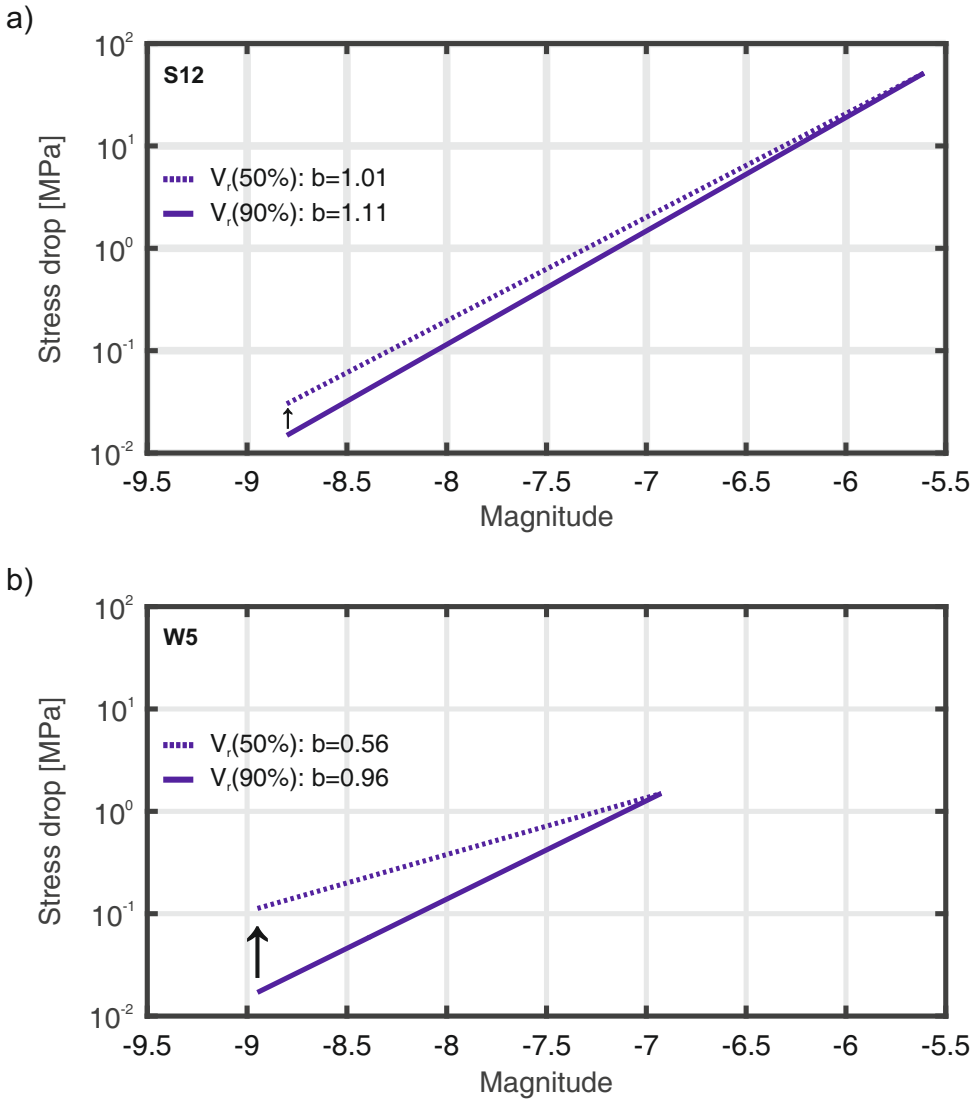

Figure 4.7: Regression slope changes caused by a rupture velocity reduction for small AEs. AE events of all AE populations were gathered to show the mean slope for (a) S12 and (b) W5. Solid lines are calculated using Sato and Hirasaw (1973) constants for 0.9 V_r rupture velocity. Dashed lines show the changed slope for 0.5 V_r. Arrows indicate a shift (increase) in static stress drop ($\Delta\sigma_{AE}$).

4.5.3 Implications for tectonic earthquakes along fault zones

Few field studies have investigated the effect of fault maturity and related roughness evolution on earthquake source parameters (e.g. Ben-Zion and Sammis, 2003; Şengör et al., 2005; Sagy et al., 2007; Brodsky et al., 2011; Bohnhoff et al., 2016b). Geometric complexity is thought to decrease with increasing fault length of fault segments, finite displacement and age. Martínez-Garzón et al. (2015) showed that for strike-slip faults maximum observed earthquake magnitudes generally scale with total slip and fault

length. Dynamic rupture models indicate that geometric complexity and associated stress field heterogeneity affect the entire source process from nucleation to rupture arrest and therefore also affect stress drop, rupture velocity and radiated energy (e.g. Ripperger et al., 2007).

Comparing our tests to field observations, a fresh fractured surface in the laboratory may be a good proxy for an 'immature' and complex natural fault zone (e.g. Kwiatek et al., 2014c; Goebel et al., 2017). A smooth saw-cut fault may serve as a laboratory analogue of a 'mature' fault that experienced extensive surface smoothing due to multiple seismic events over a long time-span. Nevertheless, up-scaling laboratory observations to natural earthquakes remains challenging, not least due to the fact that characteristic length scales are very different between field and laboratory tests. Up-scaling not only involves formulating appropriate constitutive laws for key physical processes observed in experiments but also some renormalization procedure allowing to formulate, for example, an effective friction law (Campillo et al., 2001). Furthermore, the actual roughness of natural faults at seismogenic depth is yet to be determined, which is an extremely challenging task. Thus, laboratory experiments that focus on key-parameters related to seismic slip and stress drop variations will remain the only analogue of hazardous earthquakes along tectonic faults.

4.6 Conclusion

We investigated seismic and mechanical stress drop variations during triaxial stick-slip experiments on two faulted Westerly granite samples with different roughness. We cover a broad magnitude range of -9 < M_W < -5.6, which is substantially larger than in previous related AE studies. Using the spectral ratio method, we calculated AE stress drops which are within a comparable range of observations from induced and natural earthquakes up to M_W = 4. A key observation is the strong increase of AE static stress drop with AE magnitude, which differs from most source studies analyzing natural seismicity and seems to be much more pronounced for AE events. This increase is most pronounced for AEs on smooth surfaces. The slope of the corresponding relationship between $\Delta\sigma_{AE}$ and M_W increases with consecutive stick-slips on rough surfaces. Mechanical stress drops also increase with successive slip events on rough and smooth surfaces potentially due to progressive surface smoothing. Average grain size and the width of the fault core provide approximate lower and upper bounds of AE sizes on rough surfaces. Large AE events on smooth fault surfaces exhibit the highest stress drops and are dominated by double-couple moment tensors. Our results indicate a direct coupling between the scale of heterogeneity (grain size, roughness and damage zone width) and seismic event characteristics in the

laboratory. The experiments suggest that larger stress drops require mature, smooth faults, whereas rougher faults promote lower stress drop events. Similarly, crustal heterogeneity may be an important factor that governs earthquake stress drop variations in nature. This needs to be further investigated, in particular for hazard-prone plate-bounding earthquakes, given the large implications for the resulting risk to near-fault population centers and infrastructure.

5 An open data infrastructure for the study of anthropogenic hazards linked to georesource exploitation[1]

Abstract

Mining, water-reservoir impoundment, underground gas storage, geothermal energy exploitation and hydrocarbon extraction have the potential to cause rock deformation and earthquakes, which may be hazardous for people, infrastructure and the environment. Restricted access to data constitutes a barrier to assessing and mitigating the associated hazards. Thematic Core Service Anthropogenic Hazards (TCS AH) of the European Plate Observing System (EPOS) provides a novel e-research infrastructure. The core of this infrastructure, the IS-EPOS Platform (tcs.ah-epos.eu) connected to international data storage nodes offers open access to large grouped datasets (here termed episodes), comprising geoscientific and associated data from industrial activity along with a large set of embedded applications for their efficient data processing, analysis and visualization. The novel team-working features of the IS-EPOS Platform facilitate collaborative and interdisciplinary scientific research, public understanding of science, citizen science applications, knowledge dissemination, data-informed policy-making and the teaching of anthropogenic hazards related to georesource exploitation. TCS AH is one of 10 thematic core services forming EPOS, a solid earth science European Research Infrastructure Consortium (ERIC) (www.epos-ip.org).

5.1 Introduction

The exploitation of georesources and underground storage of liquids and gases can pose environmental hazards as they can induce seismicity and cause deformation of the ground surface. They can create new fractures that change rock-mass permeability and may cause groundwater contamination and/or air pollution from the emission of

[1]This research article is published as Orlecka-Sikora, B., S. Lasocki, J. Kocot, T. Szepieniec, J.-R. Grasso, A. Garcia-Aristizabal, M. Schaming, P. Urban, G. Jones, I. Stimpson, S. Dineva, P. Sałek, K. Leptokaropoulus, G. Lizurek, D. Olszewska, J. Schmittbuhl, G. Kwiatek, A. Blanke, G. Saccorotti, K. Chodzińska, Ł. Rudziński, I. Dobrzycka, G. Mutke, A. Barański, A. Pierzyna, E. Kozlovskaya, J. Nevalainen, J. Kinscher, J. Sileny, M. Sterzel, S. Cielesta, and T. Fischer (2020). An open data infrastructure for the study of anthropogenic hazards linked to georesource exploitation. *Scientific Data*, 7. 89, https://doi.org/10.1038/s41597-020-0429-3.

fugitive gas and particulate matter (Gupta, 2002; Etiope and Martinelli, 2002; Osborn et al., 2011; Davies et al., 2013; Rivard et al., 2014; Vengosh et al., 2014; Grigoli et al., 2017; Foulger et al., 2018; Orlecka-Sikora et al., 2019; Lee et al., 2019; Lasocki and Orlecka-Sikora, 2020). Transient strong motions from induced earthquakes can cause infrastructure losses, human injuries and fatalities (Gupta, 2002; Lasocki et al., 2017). The socio-economic impacts of such anthropogenic hazards (AH) are significant. These risks can threaten or prohibit the development of the associated industries, including those instrumental for the transition to a low-carbon future such as geothermal energy and carbon dioxide sequestration. Vital technological activities may have to cease without mitigation of the accompanying hazards. This is a particularly sensitive topic in the densely populated areas of Europe where such technological activities often take place close to inhabited areas (Deichmann and Giardini, 2009; Sintubin, 2018).

These AH are poorly understood despite intensive world-wide research. Among the reasons for this are the complexities of geological and geophysical processes, the diversity and time-variability of the industrial processes responsible for generating these hazards, the complex link with natural hazards (e.g. in the distinction between triggered and induced seismicity) and the commonly non-public, proprietary nature of the data. However, AH research requires a holistic and interdisciplinary approach, as well as access to integrated and standardized data. Comprehensive science-industry collaborative efforts to monitor effectively, analyze and evaluate anthropogenic seismicity and the resulting hazards are underpinned by data sharing and can have tangible benefits for both industry and society. Access to 'big' and 'open' data from numerous case studies across different georesource sectors is needed to facilitate deeper scientific insight, enable retrospective research and to improve transparency. To facilitate this, a community of European scientists have created the Thematic Core Service Anthropogenic Hazards (TCS AH), utilizing the framework of the European Plate Observing System program (EPOS).

TCS AH is a new cooperative research undertaking developed by European representatives from science and industry, with a transnational governance framework covering implementation, best practice and sustainability strategies, outreach and dissemination. TCS AH provides the framework for global-scale investigation of AH related to georesource exploitation, achieved through open access data and applications in accordance with FAIR (Findable, Accessible, Interoperable, Reusable) data principles (Wilkinson et al., 2016).

TCS AH, one of the ten thematic core services of EPOS, has been developed within the framework of the European Commission's infrastructural projects FP7-INFRASTRUCTURES-2010-1 (the EPOS preparatory phase, EPOS-PP), and H2020-INFRADEV-1-2015-1 (implementation phase, EPOS-IP) (Fig. 5.1). TCS AH coordinates the main-

tenance and development of the IS-EPOS Platform, initially established as a working prototype by the IS-EPOS POIG.02.03.00-14-090/13-00 project.

Figure 5.1: Landing web page of the TCS AH with access to IS-EPOS portal (tcs.ah-epos.eu).

Data is gathered in thematic episodes. Each episode forms a comprehensive set of geophysical, industrial and environmental data related to induced seismicity originating from exploitation of a particular georesource. The platform is an innovative e-research environment for researchers, integrating the data and software applications, and providing a flexible virtual laboratory workspace for data processing, analysis and visualization. The platform supports collaborative functionalities, e.g. the sharing of user workspaces. The access to the integrated resources of the IS-EPOS Platform is open to all. Visitors to the IS-EPOS portal are able to preview the available resources without registration. Registered users, however, can utilize all the functionality of the platform.

5.2 Methods

5.2.1 System architecture

The TCS AH research infrastructure is integrated and provided to the community through the IS-EPOS Platform. The platform (Fig. 5.2) provides access to:

- data and metadata gathered in the form of episodes with associated tools for data and metadata search and visualization;
- applications for online data processing;

- user-organized workspaces for storing and sharing data and applications.

In addition to the direct access via the IS-EPOS Platform, TCS AH will also be integrated with the EPOS-ERIC IT infrastructure. This will enable all EPOS users to access the data and functionality of the IS-EPOS Platform and to use them in multidisciplinary projects conducted with tools provided by the wider EPOS-ERIC IT Platform.

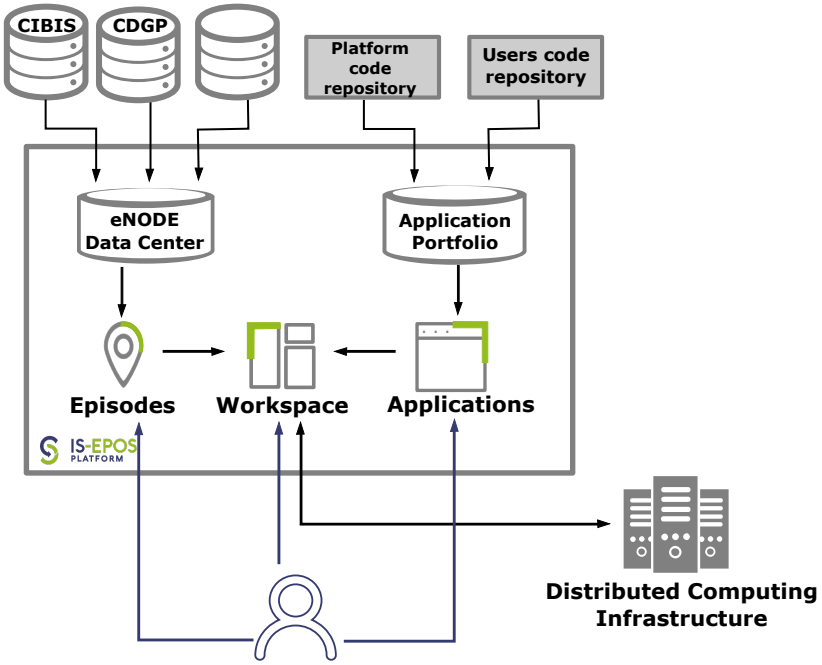

Figure 5.2: General architecture of IS-EPOS Platform with its main components: Episodes (data and metadata), Workspace and Applications. Episode data are stored in eNodes, fed into the eNode Data Center and then offered to the user. On their request, the data is loaded to a user's Workspace. Application codes are stored in the code repositories, gathered in an Application Portfolio, and then shown to the user. These can then be loaded into a Workspace and executed on Distributed Computing Infrastructure.

Data records and acquisition

Episode datasets related to investigations of particular anthropogenic hazard phenomena are maintained by participating institutions forming separate eNodes (e.g. CIBIS and CDGP eNodes). Data in eNodes are formatted in accordance with IS-EPOS Platform data-format specifications and described with associated metadata. The data are provided in unified formats, described in the IS-EPOS Platform documentation (docs.cyfronet.pl/display/ISDOC/Data + formats). The data formats are based, as far as possible, on existing standard formats (e.g. seed, mseed, QuakeML, InventoryXML), but

for the types of data for which there are no established standard formats, the platform uses custom Matlab-based formats: catalog, GDF (Generic Data Format) and MDDF (Multi-Dimensional Data Format). The GDF and MDDF files have been designed as smart structures to store diverse scientific data, for example, water quality, air quality, industrial production data and geological/geophysical data. These data formats are very widely used and easy to use for further processing. GDF and MMDF files contain the identity of the geographic coordinate system in which data are stored, the time zone in which the time is determined, and information about the stored data, such as units, data type, and names of variables with descriptions. These metadata are stored in the CRF field (Coordinate Reference System) of the Matlab structure. Geographical data are usually stored using the WGS-84 ellipsoid reference coordinate system. Coordinate conversion to WGS-84 is a part of the IS-EPSOS Platform data integration process before the data undergoes thorough quality control and made available to users. There are also a number of format converter applications on the IS-EPOS Platform, which allow conversion from the internal formats to user-selected formats when the data is downloaded.

The associated metadata make the data discoverable and searchable, provide additional information about the content and origin of a specific data item and provide information on any access policy. Both data and metadata are subject to quality control (see: Technical Validation section below). Every episode published on the IS-EPOS Platform has a unique DOI for citation and direct access to the original data source, which is crucial under FAIR data principles. All the metadata coming from different eNodes are collected and indexed in the eNode Data Center.

Applications

Applications allow the processing of data from the IS-EPOS Platform as well as that uploaded by the user. Applications range from simple data management routines to advanced services for specialized data analysis. The latter are software packages developed, maintained and published by researchers. Future development of the IS-EPOS Platform should enable users to run their own data-analysis scripts within the platform (Fig. 5.3).

Workspace

The workspace provides the user with a framework to organize data and applications into integrated scientific projects. Data in workspaces are represented as files and they can be organized within a hierarchy of directories (Fig. 5.4). Each data file has its own metadata

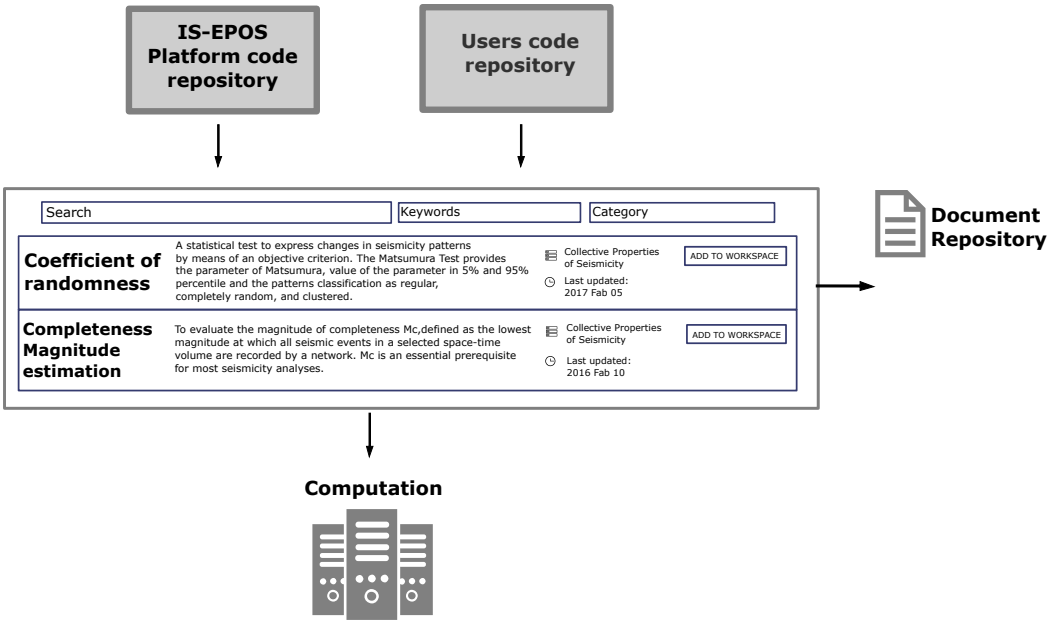

Figure 5.3: IS-EPOS Platform applications, combined from the official code repository, as well as from a custom user code repository. The applications have references to associated publications stored in the document repository.

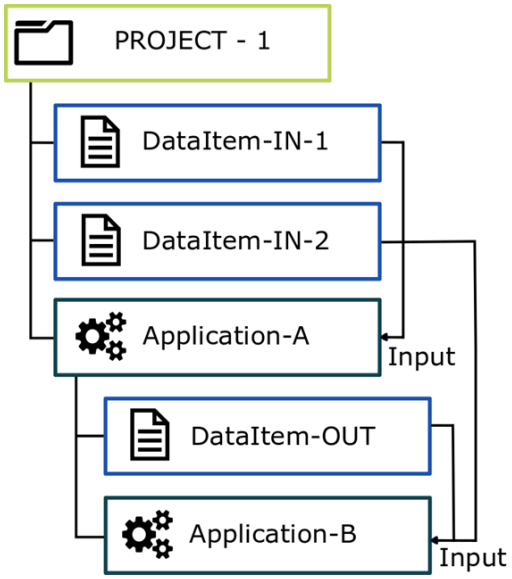

Figure 5.4: Schema of a sample organization of data and applications in workspace.

information allowing the user to match compatible applications with it. Users can add applications to their personal workspace in order to execute them on selected data and capture the resulting output. An application in a workspace is represented by a special kind of directory. This directory both organizes the data and provides an interface to handle data processing parameters, trigger the computation, as well as visualize, preserve and export the processed output (see Fig. 5.5). The processing itself is delegated to a distributed computing infrastructure (cloud or high-performance computers). The output of an application execution is stored in the application directory and may become input for another application. Consequently, a series of applications with their resultant data may be combined to form a workflow.

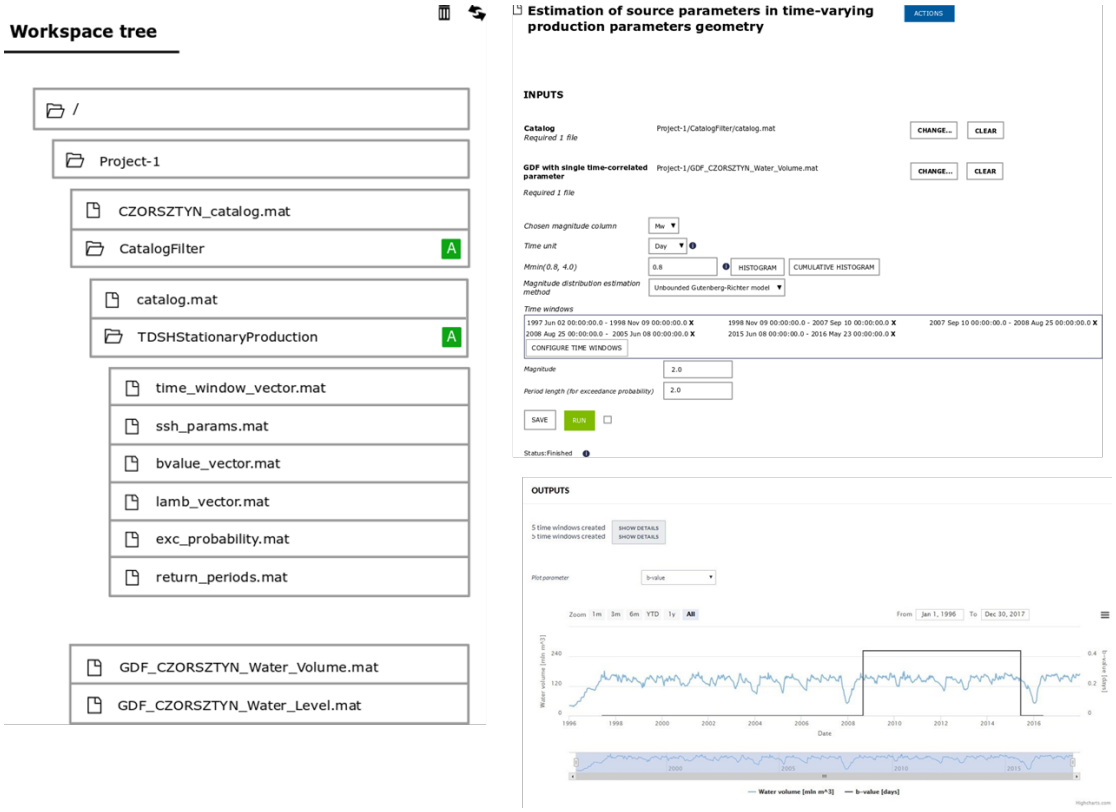

Figure 5.5: Sample application localized in a workspace, displaying the form of parameters, computation status and resulting visualization. The resulting data are stored in the application directory within a workspace tree, on the left.

All data stored in the user's workspace can be shared with other users of the platform, either in an editable or read-only mode. Similarly, the application settings and processing workflows can be shared.

Integration with EPOS-ERIC IT

Ultimately, all ten thematic core services of EPOS, including TCS AH, will be integrated into the EPOS-ERIC IT infrastructure known as Integrated Core Services, ICS, where users will be able to discover, access and process multidisciplinary datasets available from the individual TCSs. TCS AH is integrated with the EPOS Authentication, Authorization and Accounting Infrastructure (EPOS AAAI). This allows TCS AH users to log on to the ICS platform with their existing credentials and wider EPOS users to access TCS AH data and applications. There is also an authorization mechanism to recognize various user roles and attributes through the EPOS AAAI in order to grant access to specific resources within the platform. This is implemented using standard protocols such as OpenId Connect and OAuth 2.0. The interactions involved in these protocols are depicted in Fig. 5.6.

Another aspect of integration with the ICS is related to the metadata that describes data maintained by the individual TCSs. These metadata are harvested and converted by the ICS to make the data discoverable and usable in a uniform way across the different EPOS TCS platforms.

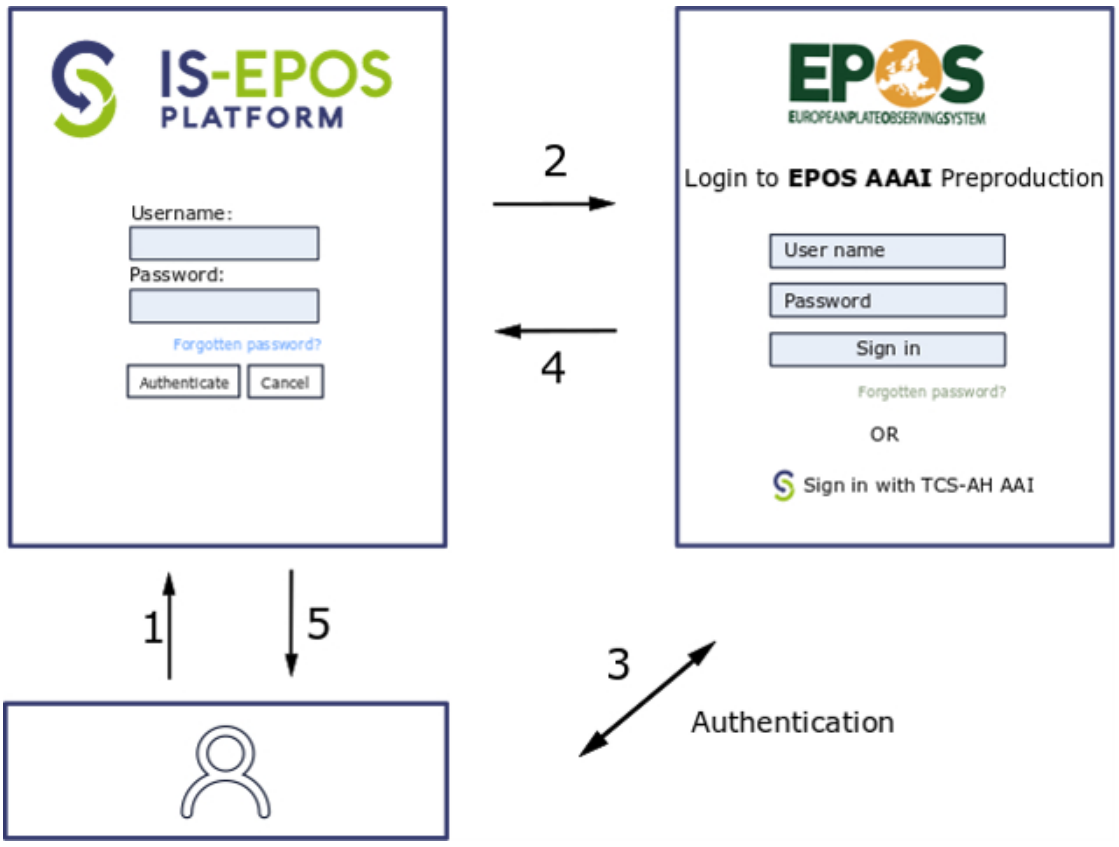

Figure 5.6: The process of user authentication to the IS-EPOS portal using an EPOS AAAI account. Numbers refer to the flow of actions.

5.2.2 User engagement and science-industry partnership

The AH community is comprised of stakeholders from academia, industry, educational institutions, NGOs, the public and government organizations. Spreading awareness of AH phenomena and the services offered by TCS AH has involved building an online presence (project website content, social media communications, online newsletters), face-to-face interaction (outreach events, scientific workshops, IS-EPOS Platform training workshops) and scientific dissemination (peer reviewed publications, conference presentations and participation at events in, and related to, the field of anthropogenic hazards). Approximately 10 training workshops have taken place with over 200 AH community stakeholders from the scientific community being trained to use the IS-EPOS Platform across 7 countries during the course of the EPOS-IP project. In addition to gaining feedback on the IS-EPOS Platform, the training events organized by TCS AH and important face-to-face stakeholder engagements have fostered the development of new ideas for the use of the integrated research infrastructure. The IS-EPOS Platform has also been used as a teaching tool in an educational project aimed at junior high and high schools called ERIS – Exploitation of Research Results in School practice, funded with support from the European Commission within the ERASMUS + Programme. It has also been integrated into university teaching as an educational and effective research resource at several of the TCS AH consortium institutions.

By August 2019, the IS-EPOS Platform had more than 1000 registered users (an increase of 720 % since the launch of the EPOS-IP project in 2015). To stimulate further the use of the platform and the provision of services, the TCS AH consortium is open to new collaborations. Data and/or software from new potential partners, from on-going or past projects are welcome. Collaboration is also sought with companies producing equipment for seismic activity monitoring or accompanying processes. The ambition is not only to be a provider of an e-research environment but also to perform inter-community social functions like project brokering and developing opportunities for new interdisciplinary and international collaborations.

5.2.3 Governance and future perspectives

The TCS AH community is organized around a TCS AH consortium of 13 partnering institutions (suppliers of data and/or services) from 8 European countries. The work of the consortium is governed by a Consortium Board consisting of representatives of all the partners and with a director elected for 5 years by a majority of the board. Five sections have been established within TCS AH: (a) Implementation of services; (b) Administration, law & accounting; (c) Episodes integration and application implementation;

(d) Promotion and dissemination; (e) Projects & partnership. There are two external committees, the Data Provider and User committees which are composed of the main representatives of the respective communities, selected by the majority of the board to advise the director. An Innovation Advisory Committee consisting of stakeholders from academia, industry, local and central administration bodies, society and others, invited by the Consortium Board, consults on the decision-making processes behind TCS AH developments. Membership of the TCS AH consortium is open to new partners.

The plan is to establish TCS AH as a coherent coordinating framework for the AH community, together with a robust research infrastructure to develop strategic research capacity for addressing AH challenges. The goals are to facilitate new discoveries, connect the stakeholder community, boost public understanding of AH, develop outreach materials for the public and future scientists, stimulate innovation through knowledge transfer, provide solutions to industrial partners and engage in a three-way transfer of knowledge between industry, science and society.

5.3 Resource Description

5.3.1 Data records and acquisition–cross-national episode eNodes

IG PAS eNode – CIBIS

The IG PAS eNode "Induced Seismicity Research Infrastructure Center" (CIBIS) is located in Krakow, Poland and managed by the Institute of Geophysics Polish Academy of Sciences, where more than 30 AH episodes from various projects are stored and maintained (Figs. 5.2 and 5.7), covering a broad variety of industrial activities. The majority of the episodes relate to induced seismicity and ground instability from underground mining (10 episodes). Water reservoir impoundment is the second most represented industrial activity (6 episodes). Next are conventional hydrocarbon extraction, underground fluid storage and geothermal energy production episodes (5 datasets, including one complex episode involving CO_2 storage). Episodes on unconventional hydrocarbon extraction related to fracking (4 datasets) are also stored in CIBIS with data included from the H2020 projects SHEER (Shale gas Exploration and Exploitation Induced Risks) and S4CE (Science 4 Clean Energy). The integrated episodes are located mainly in Europe (24), but also in Northern America (3) and Asia (3) (Fig. 5.7).

Data within the episodes mostly include geophysical observations related to the anthropogenic hazard acquired directly from instrumental measurements as well as industrial data describing the processes potentially causing the hazard. Geophysical data include seismological event catalogs, ground motion catalogs and waveform data, but

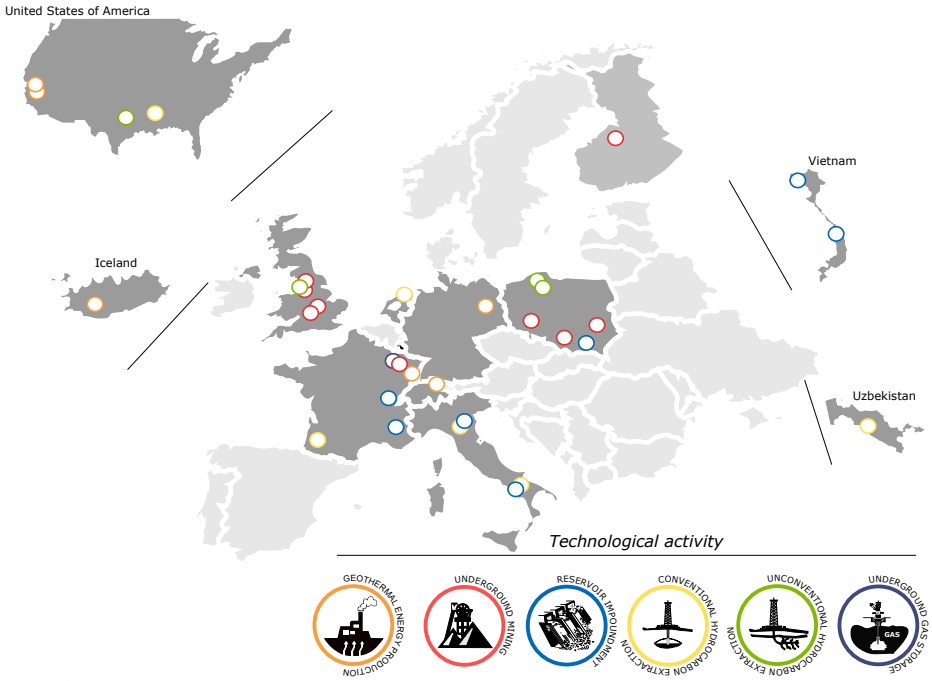

Figure 5.7: Geographical distribution of episodes stored in CIBIS and CDGP.

also water and air quality data from in situ and laboratory measurements. Geological data, including bedrock geology and tectonic features are also included, as well as geospatial and geotemporal data related to the causal activity. Every episode also contains a brief description of the locality and potentially hazard-inducing processes together with references and complementary documentation. A typical episode summary page is shown in Fig. 5.8.

CDGP eNode

The "Data Center for Deep Geothermal Energy" (Centre de Données de Géothermie Profonde - CDGP) archives data collected at the Upper Rhine Graben geothermal sites and distributes them to the scientific community, taking relevant intellectual property rights into account. It is located at the École et Observatoire des Sciences de la Terre (EOST, Université de Strasbourg, CNRS), and was created by the LabEx G-eau-thermie Profonde, a research program on deep geothermal energy founded by the French Ministry of Research and Education under the “Laboratories of Excellence” initiative. It operates as a standalone data center (cdgp.u-strasbg.fr) but has also been incorporated as an eNode of TCS AH, available through the IS-EPOS Platform.

Data distributed by the CDGP consist of seismological (catalogs, waveforms) and

hydraulic data that were acquired during stimulation or circulation phases at the Soultz-sous-Forêts pilot plant during the research period (1988–2010). Other geophysical data (gravimetric, magnetic, InSAR, geodesy, velocity model, etc.) will be also added to the data store in the future, as well as other data from different projects (Rittershoffen, Illkirch, Vendenheim). Agreements with industrial partners allow the CDGP to distribute particular data to the academic community. The CDGP infrastructure directly shares as much information as possible such as episode metadata with the IS-EPOS Platform to avoid work duplication, however, some special tasks are necessary to complete the metadata information needed by the IS-EPOS Platform as well as the reformatting of certain data.

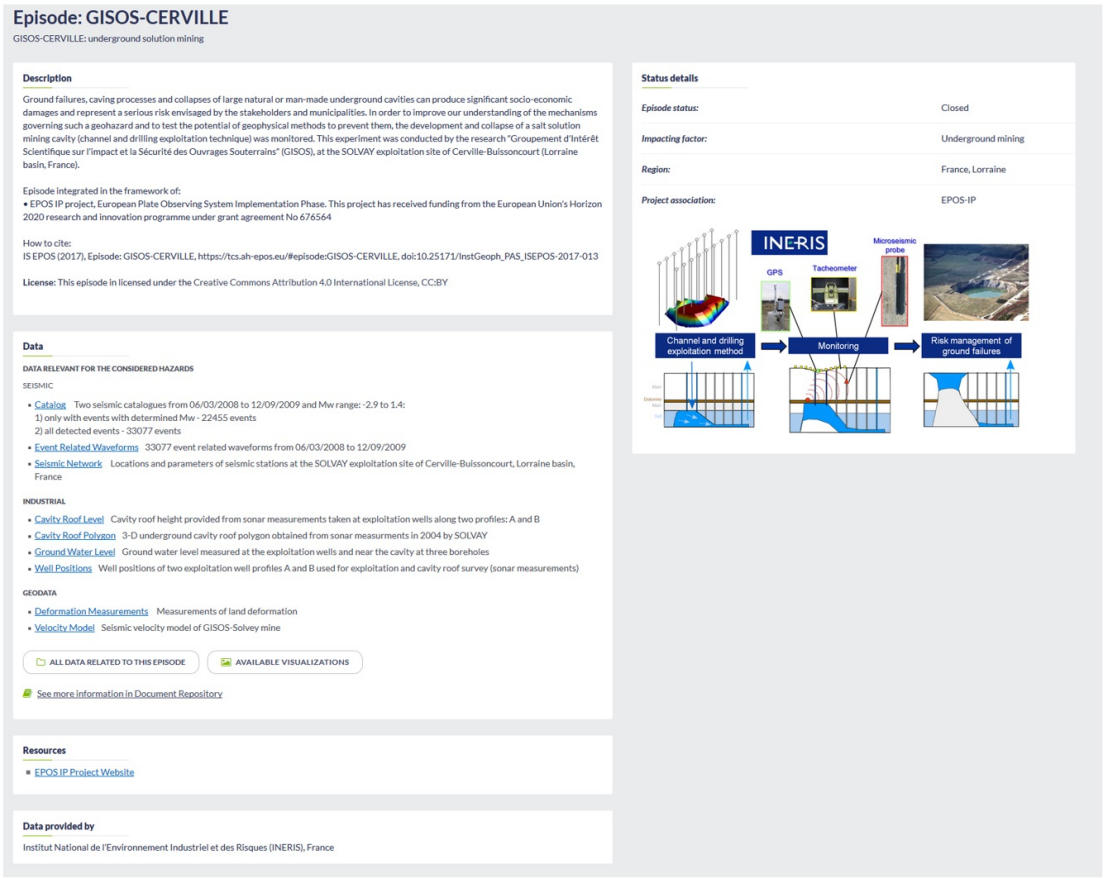

Figure 5.8: Example of data organization within an episode on IS-EPOS Platform.

5.3.2 Applications

The IS-EPOS Platform provides users with a range of data analysis and manipulation routines. These are mostly (but not exclusively) Matlab or Python programs that allow analysis, processing and visualization of data available on the platform or imported into it (Rudziński et al., 2017; Leptokaropoulos et al., 2019). Currently, 44 applications are

integrated into the IS-EPOS Platform grouped into 13 thematic categories (Table 5.1) in order to facilitate particular research analyses related to anthropogenic hazards. Some programs can be freely downloaded under the terms of a GNU General Public License.

The integrated applications comprise software tools which implement peer-reviewed techniques designed for dealing with specific scientific issues. These tools enable the user to perform research operations and analyses based on either seismic catalog data (e.g. correlation analysis, probabilistic seismic hazard assessment, earthquake interactions, stress field modeling and collective properties of seismicity) or waveform recordings (e.g. seismic event detection algorithms, event location, source parameter estimation including moment tensor inversion, and spectral analysis). In order to achieve more flexible workflows and a user-friendly environment, these applications are supported by a variety of data handling and visualization applications for performing basic parameter operations and data filtering, converting catalog and waveform data formats and graphically visualizing seismic as well as operational data. The results of the applications can be downloaded, reformatted, visualized, and used as input for further analysis. The user can create individual projects within their personal workspace, which can then be shared with other colleagues or research teams.

New applications are being developed and added in the Continuous Integration and Continuous Delivery (CI/CD) scheme bringing additional scientific potential to the platform, improving the user experience and increasing the ability to conduct advanced studies in the field of anthropogenic hazards. Advanced applications (template matching and multi-scale array-based algorithms for microseismicity detection and location, statistical toolboxes for magnitude complexity analysis and seismic event clustering) are being developed and are ultimately planned to be integrated into the IS-EPOS Platform as part of the project SERA —H2020-INFRAIA-2016-1.

In order to make the IS-EPOS Platform more inclusive and user friendly, future development plans for the platform include the introduction of the choice of user language. The ultimate goal is to provide the users with the ability to upload their own code scripts and to modify the source code of existing applications according to their specific requirements.

Table 5.1: Applications available on the IS-EPOS Platform.

Category	**Application**
Collective Properties of Seismicity	Anderson-Darling test for exponentiality of inter-event time
	Coefficient of randomness
	Completeness Magnitude estimation
	Magnitude conversion
	Priestley-Subba Rao (PSR) test

Converters	CSV to Catalog converter
	Catalog to ASCII converter
	Catalog to Vectors converter
	GDF to Vectors converter
	GDF to XLS converter
	Catalog to XLS converter
	Ground Motion Parameters Catalog builder
	Time Series builder
	Seed converter
Correlation Analysis	Auto-correlation
	Cross-correlation
Data Processing Applications	Basic Vector Operations
Download Tools	Signal download tool
	Waveform download tool
Earthquake Interactions	Earthquake interactions: Georesource scale
	Earthquake interactions: Mainshock scale
	Earthquake swarm (reshuffling analysis)
	Time correlated earthquakes (Seasonal trends)
Event Detection Algorithms	Template-matching based detection algorithm
Filtering Tools	Catalog filter
	Estimation of source parameters in time-varying production parameters geometry
	Ground Motion Prediction Equations
	MERGER: Dynamic risk analysis using a bow-tie approach
Source Parameter Estimation	Effective stress drop estimate
	Estimation of source parameters in time-varying production parameters geometry
	FOCI
	Localization
	Mechanism: Full Moment Tensor
	Mechanism: Shear Slip
	Spectral Analysis
	Waveform-based seismic event location
Stress Field Modeling	Stress inversion
Visualization	Estimate of maximum possible magnitude for reservoir triggered seismicity
	Fracture Network Models - Mechanical Stresses
	Front Advance histograms
	Integrated Google Maps data visualization
	Seismic Activity with Front Advance

5.3.3 Document repository

The IS-EPOS Platform provides a repository of documents associated with episodes and applications. This e-repository is powered by EPrints 3 (GLU license, www.eprints.org) which has been developed by the School of Electronics and Computer Science at the University of Southampton, UK. Thus the IS-EPOS Platform documents support OAI 2.0 with a base URL of tcs.ah-epos.eu/eprints/cgi/oai2. The front page is organized so that legal requirements, the user guide, guidance on how to cite the IS-EPOS Platform and other references to episodes and applications, can be directly accessed. The platform legal regulations contain documents regarding data policy, the data management plan

and TCS AH governance services.

The user guide section consists of technical documents with instructions to assist using the IS-EPOS Platform, with a step-by-step tutorial on how to access and make use of particular applications. Links to the user guide for specific applications are also available from the application level. The user guide also includes guidance about data formats and the vocabulary of the IS-EPOS Platform.

The IS-EPOS Platform documents are thematically organized and classified, consisting of links to relevant articles, books, book sections, conference and workshop items or reports. There are currently 858 items in the e-repository. The internal search engine allows browsing using filters and search criteria. Items are classified according to episode, type of hazard-inducing phenomenon, monitoring technology, and geographic region. Consequently, the elements of the document repository can be readily accessed by applying specific filters. In addition, references relating to particular episodes and applications are also available as a link from their respective sections on the platform. There is also a collection of links to documents that reference the IS-EPOS Platform itself (tcs.ah-epos.eu/eprints/id/saved_search/5).

5.3.4 Integration of new episode data–quality control procedure

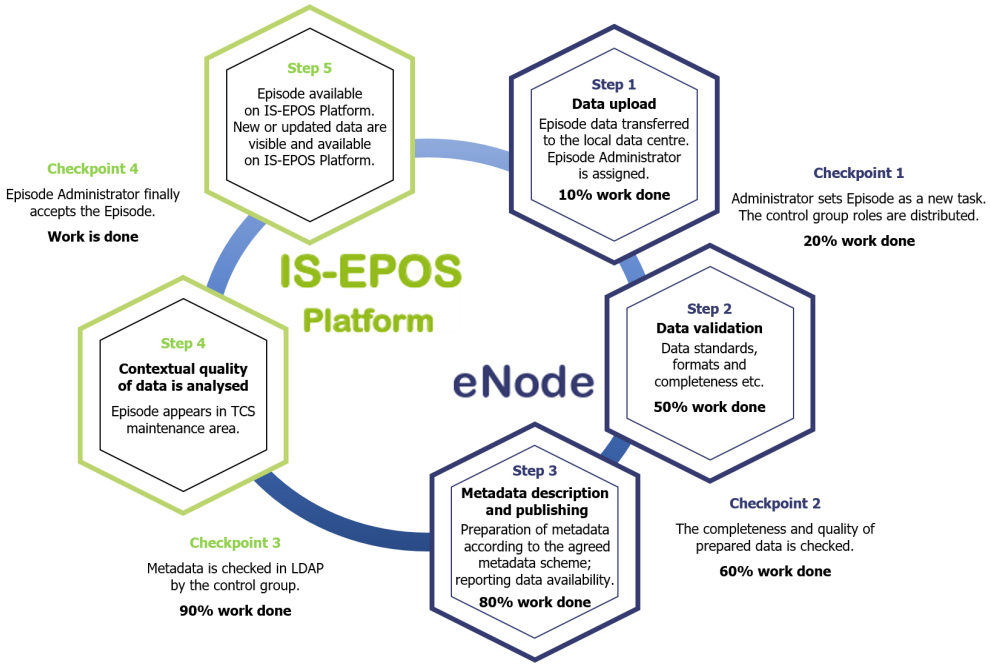

Figure 5.9: Quality Control Workflow of the AH Episode Access Service.

Quality control of the integrated and published episode data is carried out according to the diagram shown in Fig. 5.9 which illustrates the data processing steps and relevant control

points. An open-source project web system is used to manage the episode integration process (Redmine, www.redmine.org). This allows the assignment and management of personnel responsible for the episode integration, as well as tracking the progress of that integration. The uploaded data are prepared, checked and validated to ensure that they conform to commonly used standards and formats. Any inconsistencies in the data structure are identified by the quality control team. Once the data and metadata are approved, they are made available on the IS-EPOS Platform test portal from the external source. The episode quality and integrity is then reviewed together with the data owner before the data and associated metadata is made public on the IS-EPOS Platform.

5.4 Application Validation

Any code intended to be integrated as an IS-EPOS Platform application has to pass a two-step validation before it is released in the application portfolio. First, before the integration process starts, a technical validation aims to determine whether the code meets the technical requirements of integration. The program is checked to ascertain if it can be run on distributed computing infrastructure, if the licenses (of the code itself or the libraries it uses) allow for its use on the IS-EPOS Platform and if the input and output file formats are compatible with the platform. The second step of validation includes thorough testing of the integrated application before releasing it to the public. This is an evaluation of the interface and the application potential by testing the application with various data samples available on the platform, an assessment of the application potential to build chains with other integrated applications to form a workflow, an evaluation of the usefulness of the application documentation, and the identification of bugs. This stage of validation may be an iterative process, including improvements to the code and reassessment of its functionality.

5.5 Usage Examples

5.5.1 Example of application available on IS-EPOS platform

Complementary to the data-handling, processing and visualization tools, the IS-EPOS Platform also provides a set of tools for tackling complex multi-hazard risk (MHR) scenarios usually found in activities related to the development of georesources. A service called “Simulator for Multi-hazard risk assessment in ExploRation/exploitation of GEoResources” (MERGER) has been designed to provide tools for probabilistic multi-hazard analyses (Garcia-Aristizabal et al., 2019).

The main objective of this tool is to provide a quantitative model for performing highly specialized MHR assessments in which risk pathway scenarios are structured using a bow-tie approach, which implements the integrated analysis of fault trees and event trees. The methodology implemented in this service is suitable for performing dynamic environmental risk assessments. This is characterized by the bow-tie structure coupled with a wide range of probabilistic models flexible enough to consider different typologies of phenomena, the Bayesian implementation for data assimilation, the handling and propagation of modeling uncertainties and the possibility of integrating data derived from integrated assessment modeling.

Once MHR scenarios have been identified and structured according to a bow-tie logic, both the scenario structure and the data can be loaded through a user-friendly graphical interface which guides the user through this process (Fig. 5.10).

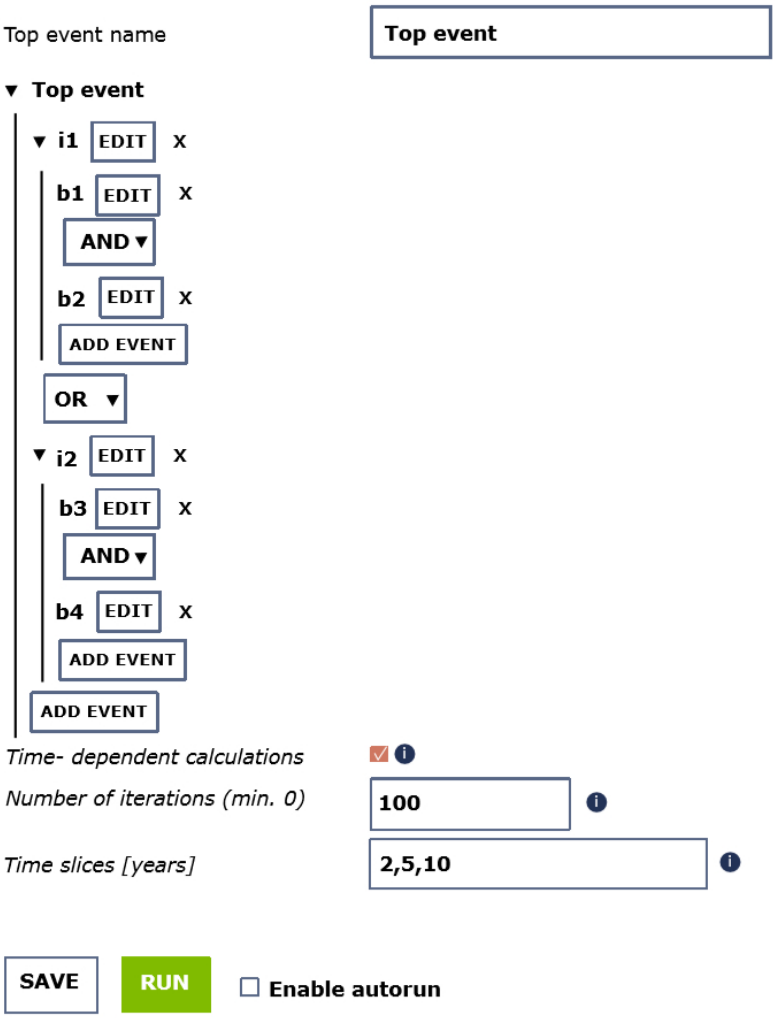

Figure 5.10: View of the graphical user interface available for the input of a fault tree (data and logic structure) in MERGER.

The output produced by this tool provides probabilistic assessments of the fault trees and event trees that were implemented for analyzing a given MHR scenario. Data output is provided numerically (in a log file) and graphically (e.g., through histograms) for each event of interest (as for example shown in Fig. 5.11).

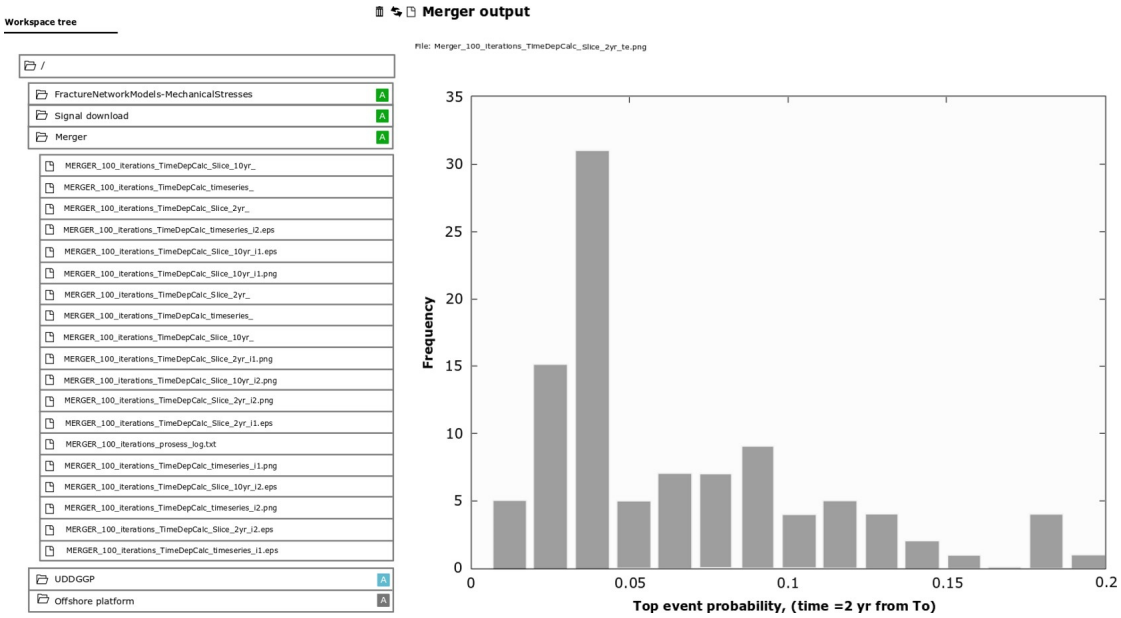

Figure 5.11: Example of the output produced by MERGER and shown on the IS-EPOS Platform for the top event of a fault tree. The selected output is displayed directly in the workspace, other results can be selected in the workspace tree on the left.

5.5.2 Use case 1: Studying correlation between injection rate and seismicity rate

Discovering the nature of the relationships between induced seismicity and the industrial factors that are its cause is probably the most important goal of anthropogenic seismicity studies because only the understanding of such relationships can lead to the development of methods to mitigate the anthropogenic seismic hazard. The example in Fig. 5.12 presents the steps in investigating if and how the seismicity induced by geothermal energy production correlates with the rate of fluid injection. The data for this use case has been taken from the episode “The Geysers Prati 9 and Prati 29 cluster” (Kwiatek et al., 2015) on the IS-EPOS Platform (`tcs.ah-epos.eu/\#episode:THE_GEYSERS_Prati_9_and_Prati_29_cluster`, doi:10.25171/InstGeoph_PAS_ISEPOS-2017-011) and the analysis makes use of applications integrated on the platform.

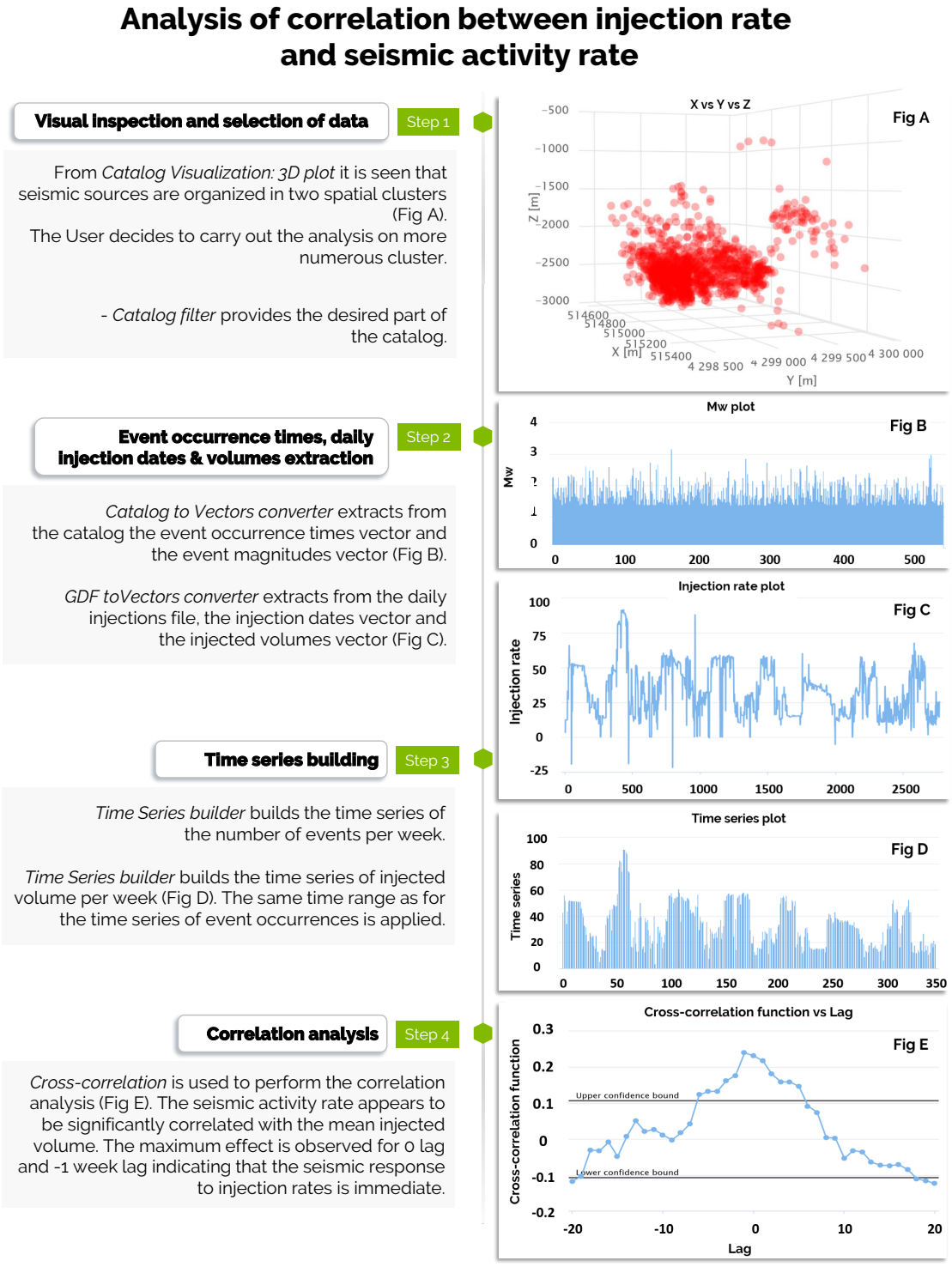

Figure 5.12: Workflow of analysis of correlation between injection rate and seismic activity rate during geothermal energy production.

5.5.3 Use case 2: Integrated visualization of artificial lake water level changes and triggered seismicity

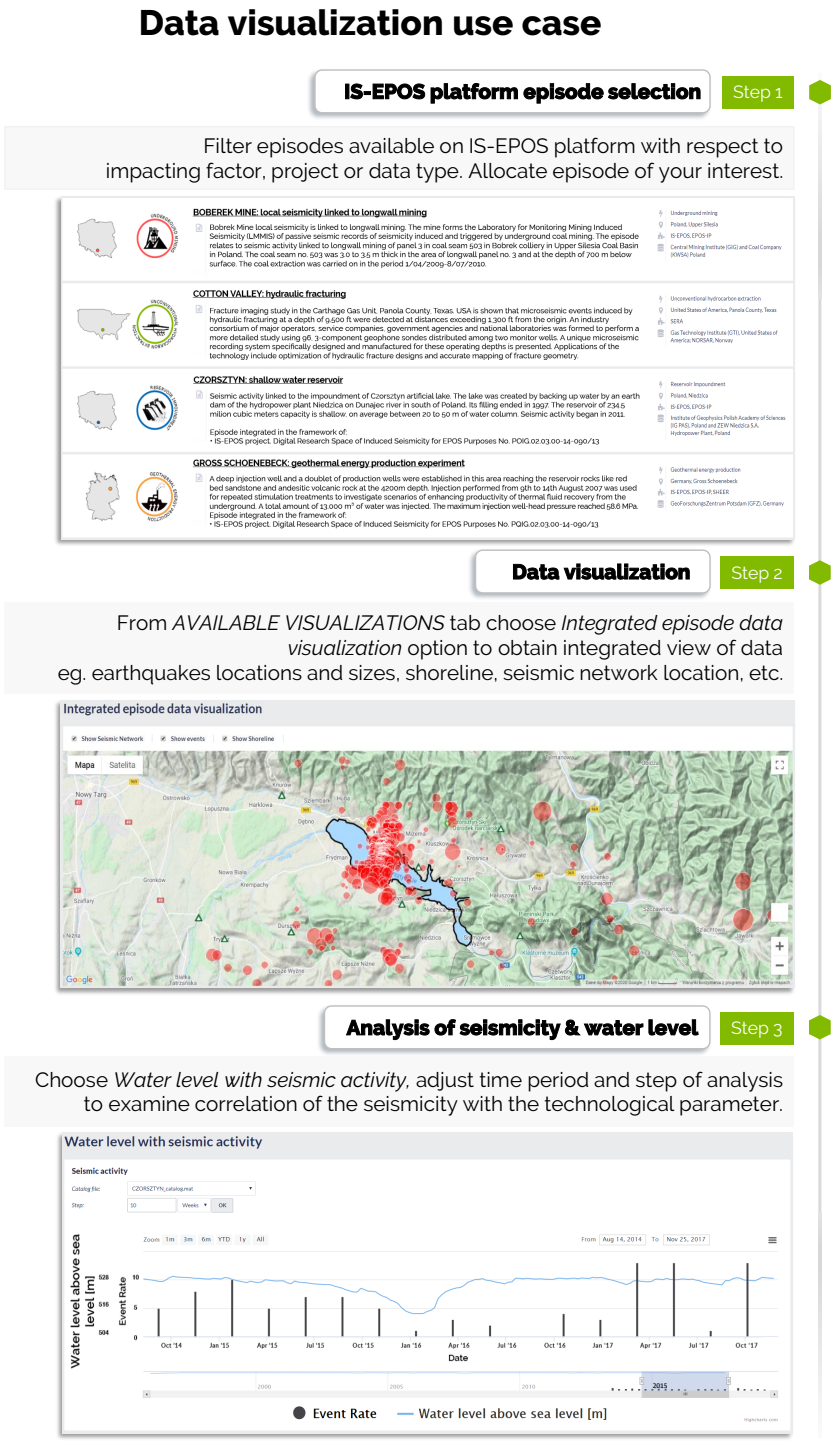

Figure 5.13: Example of the integrated visualization of water reservoir triggered seismicity and the triggering technological operations.

Comprehensive visualization tools available on the IS-EPOS Platform enable integrated visual inspection of multidisciplinary data. The example presented in Fig. 5.13 is taken from the episode "Czorsztyn" of seismicity triggered by the impoundment of an artificial lake https://tcs.ah-epos.eu/#episode:CZORSZTYN, doi:10.25171/InstGeoph_PAS_ISEPOS-2017-004. It shows the steps leading to the visualization of the triggered earthquake locations and magnitudes, the location of seismic recording stations, the lake shoreline on a topographic map background, and the occurrence times and sizes of the earthquakes superimposed on the lake water level time series. The IS-EPOS Platform maps data in WGS-84 ellipsoid coordinates, so datasets from various regions of the globe can be directly compared. Detailed information about the coordinate system used for a particular dataset is included in a file structure description and in the metadata.

5.5.4 Use case 3: Seismic hazard assessment

The IS-EPOS Platform features enable complex scientific analyses. The workflow diagram in Fig. 5.14 shows steps of an example analysis towards seismic hazard assessment. The presented analysis starts with loading a seismic catalog in csv format into the user's workspace. After converting the catalog into the internal IS-EPOS (.mat) format, the user carries out the analysis using various applications available on the platform. The user can execute advanced statistical applications that support both parametric and nonparametric methods, and perform analyses leading to the probabilistic seismic hazard assessment in a stationary or time-dependent mode. Applications implemented on the platform also allow performing a range of alternative analyses based on adopted scenarios. The user can conduct ground motion prediction, source parameters estimation, etc.

The datasets of episodes are available on the IS-EPOS Platform of Thematic Core Service Anthropogenic Hazards: tcs.ah-epos.eu. In accordance with the EPOS Data Policy which is available at www.epos-ip.org; and in accordance with TCS AH Data Policy, which is available at www.tcs.ah-epos.eu; datasets and applications are licensed under the Creative Commons Attribution 4.0 International License, CC:BY. Applications are available on the IS-EPOS Platform of Thematic Core Service Anthropogenic Hazards: tcs.ahepos.eu. The source code of the platform itself is available at bit.ly/ISEPOSsourcecode.

Figure 5.14: Illustration of the workflow for seismic hazard analysis. (Left) the workflow, (right) platform screen snapshots.

6 Conclusion and Outlook

6.1 Conclusion

Independent on the origin of earthquakes (induced, triggered or natural), strong ground motions can be evoked, resulting in hazardous consequences at the earth's surface. Large urban conglomerations such as San Francisco, Los Angeles or Istanbul, closely located to fault zones capable of producing large earthquakes, are particularly endangered by the potentially devastating consequences of earthquakes such as ground shaking and the destruction of buildings and infrastructure, and secondary hazards such as floods and fires. In addition, anthropogenic seismic hazards may result from underground mining- and research facilities, geothermal plants, or (future planned) underground storages that can provoke unexpected seismicity, which may lead to damage at nearby local towns and cities.

Studying large earthquakes is a challenge because they only occur rarely and we lack information about variations and similarities of physical processes in seismic sources. Especially for individual hazard-prone areas, this lack of information causes considerable difficulties in the assessment of ground motions and related risks expected on the earth's surface. On the contrary, earthquakes of smaller sizes occur statistically more often (Gutenberg and Richter, 1944) making them useful for investigating the physics of the seismic source and associated seismic hazard in a more detailed way. Enhanced data resolution and new data processing techniques allow current studies to overcome the limitations of past investigations related to seismic parameters. This facilitates the analysis of smaller events with very high-frequency content. Nonetheless, the debate on similarities and dissimilarities between small and large seismic events regarding possible options for scaling up the physical processes is still ongoing.

A frequently discussed source parameter used for the analysis of earthquake scaling relations is seismic stress drop, which affects the near-field ground motions (e.g. Spottiswoode, 1993). Several studies have shown that any size of earthquake on average shows a scale-invariant behavior regarding stress drop. However, as discussed in the previous chapters, the interpretation of earthquake source self-similarity over a wide range of earthquake magnitudes in the literature is not straight-forward (e.g. Cocco et al., 2016). Several unknown parameters exist complicating the assessment of potential scaling trends and related uncertainties, and affecting the evaluation of seismic source parameters. By comparing global stress drop estimates based on similar model assumptions over a wide range of earthquake sizes, a variation in stress drop magnitude of 0.01 - 100 MPa has

been observed. This large variation needs to be better investigated to improve earthquake hazard and risk assessments. The fundamental question therefore is: *What causes the fluctuation of about five orders of stress drop magnitude?*

In general, two groups of factors can be distinguished that are assumed to cause the variance of stress drop estimates in different ways: 1) Non-physical factors such as bandwidth limitation, attenuation correction effects, and incorrect model assumptions can bias seismic source parameter estimates, leading to artificial scatter in stress drop observations. Usually, these spurious variations can be reduced or eliminated by the application of an appropriate analytical method, as presented in this thesis. 2) Physical factors reflect variations of earthquake source characteristics like rupture- and slip velocity, or properties of the fault surface and its evolution during the rupture process. These factors and their effects on source parameter estimates are still barely understood and require further investigation.

This thesis took advantage of induced earthquakes and laboratory acoustic emissions to address both non-physical and physical factors important for a better understanding of seismic source and for improving future hazard and risk assessments. When analyzing small earthquakes, we are dealing with high-frequency waves that are much more sensitive to irregularities in the earth than low-frequency waves. Local small-scale structures and material heterogeneities of rocks are still not easily detectable with seismic recordings. The crustal inhomogeneities in particular cause complex attenuating effects on the high frequencies of small earthquakes' propagating waves. Consequently, the high-frequency spectral fall-off of the earthquake source spectrum is distorted, making it difficult to investigate absolute source parameters such as corner frequency, which is used for stress drop estimation and ground motion assessment.

The quality factor Q, used to correct high frequency attenuation effects in the displacement source spectrum, was handled in different ways in chapters 3 and 4. Chapter 3 focused on determining S-wave coda quality factors (Q_C) of injection induced earthquakes at The Geysers geothermal field in California. The moving window method by Phillips (1985) was applied to investigate local spatio-temporal scattering attenuation variations. Strict requirements on data quality and maximum Q_C estimate uncertainties were investigated within an additional sensitivity analysis, which enabled the analysis of very high-frequency coda waves between 3 and 70 Hz, exceeding commonly considered frequency ranges of seismic field data. Mean Q_C estimates at lower frequencies (center frequency = 7 Hz) show very similar but more robust values compared to the Q_β estimates of direct S-waves determined by spectral fitting. This leads to the assumption that coda quality factor estimates enhance the reliability and stability of source parameter calculations (e.g. corner frequency) critical for static stress drop estimates and further

source parameters. Furthermore, this study suggests spatial heterogeneity of Q_C estimates within only short epicentral distances (< 19 km). Geology, structural complexity and the geothermal reservoir itself seem to influence (attenuate) very high frequencies differently, depending on the earthquake source location in the northwestern or southeastern part of The Geysers. This should be taken into account while calculating source parameters. Another interesting finding is the temporal stability of Q_C with changing injection periods in the northwestern The Geysers. Here, the reservoir exhibits on average either a temporally constant fracture network around the two considered injection wells, or alternatively it can be assumed that changes within the fracture network due to injections are of smaller scale and cannot be detected within the analyzed frequency range.

The coda analysis is only one of several possible approaches to address wave attenuation effects while characterizing seismic sources. In case of a large number of co-located events and when analyzing small earthquakes, the application of the spectral ratio method based on a multi-eGf approach is appropriate, as this method avoids making any detailed assumptions about wave attenuation effects. The multi-eGf spectral ratio method was applied in chapter 4 to laboratory acoustic emission (AE) events of two triaxial stick-slip experiments on Westerly Granite samples characterized by different fault roughness. The main focus lay on the estimation of reliable AE static stress drops for a wide magnitude range (-9 < M_W < -5.6), thus significantly extending previous studies (Yoshimitsu et al., 2014; McLaskey et al., 2014; Goodfellow and Young, 2014). Based on the Madariaga source model, final AE stress drop estimates were compared to the previous AE studies and to the global stress drop scatter observed across a broad span of earthquake magnitudes. Due to the broad magnitude range of about 3 orders it was possible for the first time to reasonably evaluate the scaling relations of laboratory earthquakes. The estimates cover in total the span of previously estimated AE stress drops and fall into the globally observed static stress drop variation of about five orders of stress drop magnitude. In comparison to larger earthquakes, however, the laboratory AE stress drops of both experiments show a clear dependence on event magnitude (M_W-$\Delta\sigma$ relation) which increases with each large stick-slip failure, indicating scaling breakdown.

The knowledge of both laboratory fault surfaces in combination with direct measurements of mechanical stress data facilitated the analysis of the impact of different fault surface roughness on AE derived source parameters. AE events occur on the level of grain scale and fault core width (mm-scale) in both experiments, and the sizes of micro-fractures are size-limited. Larger AE stress drops and magnitudes arise on the smooth fault surface but are absent at the rough fault. Source specific parameters (physical factors) such as rupture velocity, source radius and full moment tensor solutions were investigated in the context of the observed M_W-$\Delta\sigma$ relations. The observations

suggest that micro-slips over a rough fault surface are hampered by the high number of asperities on the fault (increased roughness leads to more complex source processes shown in high volumetric components in the moment tensor solutions of small AE events) while smooth faults can provoke larger slips and therefore larger stress drops due to the simplicity of the fault surface (smoothness leads to simple shear movements reflected in larger double-couple components in the moment tensor solutions for larger AE events). Similar observations were made in the mechanical stress data that show larger differential stresses and stress drops in the smooth sample and more complex stress measurements and smaller stress drops for the rough sample. Consequently, this study indicates that different fault surface conditions might influence the energy dissemination during faulting- and slip processes. With regard to natural faults, these laboratory results contribute to the idea that stress drop- but also magnitude differences might depend on fault roughness which is assumed to be lower for mature faults and higher for immature faults, depending on the long-term history of the individual fault (e.g. Ben-Zion and Sammis, 2003; Causse et al., 2014; Bohnhoff et al., 2016b).

Both scientific studies in chapters 3 and 4 showed that reducing the non-physical effects by, for instance, applying appropriate methods with respect to the data limitations are crucial for estimating stable source parameters from seismic records. Only if non-physical effects are removed from the data is it possible to investigate physical factors to resolve potential variations in individual seismic sources. Therefore, high-quality data records are essential to facilitate studies across the whole earthquake magnitude scale in order to better assess the observed variations in source scaling relations. In addition, both the studies presented here could only be interpreted with additional auxiliary information (e.g. industrial-, geological-, and mechanical data, sensor limitations, etc.). The lack of such auxiliary data usually complicates detailed research. Therefore, providing data access to verified, standardized, comprehensive datasets is important for the scientific community to solve problems related to earthquake physics and to reduce potential error sources during calculations. Furthermore, not only the storage and access of data but also the preservation of applications and methodologies facilitates comparative studies on earthquake source physics. These requirements on high-quality seismic research were addressed in the complementary publication presented in chapter 5. Here, the development of the new IS-EPOS online research platform, operated by the TCS AH team in the frame work of the EPOS-IP project phase, was presented. It aims to enhance future research by implementing the aforementioned requirements on data on the basis of anthropogenic hazard related studies. The remarkable diversity of the platform includes 1) quality checked and standardized full datasets, 2) specifically developed application tools, and 3) a private but also public workspace to process and exchange further data

products. The platform aims to connect different communities (academic, industrial, public), which currently work in isolation from one another. The project serves as an excellent example of bringing together the expertise from multiple technical and scientific fields necessary to overcome our present limitations and to achieve outstanding future research.

Summarizing this thesis, the main conclusions are the following:

1) The reliable assessment of quality factors is essential when small earthquakes are analyzed. Local changes in geology and structure, expressed in path- and site effects, have to be assessed to correct seismic source spectra and to calculate seismic source parameters more accurately. Careful selection of those analysis parameters which best meet the data characteristics is one way to reduce estimation variations and thus uncertainties.

2) Local coda S-wave quality factors were analyzed at very high frequencies up to 70 Hz which exceeds most of the previously considered frequency bands in field studies. Using high frequencies, local crustal small-scale obstacles can be detected, which helps to better evaluate the heterogeneity of the study area and to detect very local site effects close to the seismic sensors. Thus, attenuation of high-frequency waves can be better assessed, necessary to refine seismic source parameter estimates of small earthquakes

3) The multi-eGf spectral ratio approach enables for the estimation of spectral seismic source parameters when local path effects are unknown or too complex to model, but a large number of co-located events is available. Multiple linkage of events helps to stabilize the inversion for source parameters, and reduces the influence of possible outliers. This method was successfully applied for the first time on laboratory AE data from two stick-slip experiments.

4) The AE study covers more than 3 orders of magnitudes and significantly extends previous individual AE studies of narrower magnitude ranges. The broadband character of the study enabled us to address static stress drop variations for the smallest earthquakes.

5) AE stress drops fit well to the globally observed estimates but show a clear dependence on event magnitude, indicating an apparent scaling breakdown for the lower end of the seismic scale.

6) As AE event sizes are roughly bounded by the average grain size and fault core width, the observed stress drop-magnitude dependence was found to be related to

varying slip sizes over micro-fractures. Frictional properties of the fault surfaces (roughness and fault complexity) might be responsible for the observed variation in slip and thus stress drop estimates between different study areas or on very heterogeneous faults.

7) High quality, standardized and multidisciplinary data are necessary to overcome potential difficulties related to seismic data processing, helping to focus the research effort on solving actual problems related to earthquake physics. A platform for data sharing and creating synergies across different areas of expertise is a valuable tool to e.g. combine and evaluate the physical and non-physical aspects required to assess observed and calculated parameters.

6.2 Outlook

Based on the conclusions, this thesis encourages further investigations of small earthquakes in both field studies and laboratory experiments. New recording techniques and refined sensors sensitivities allow us to record much smaller earthquakes (higher frequency content) that can help to reach considerably higher resolution of small-scale structures in the earth's crust. So far, it can be assumed that the choice of an averaged constant quality factor for an individual study site might be inappropriate when the area is large or very heterogeneous in sense of structures and geology. Especially when a wide magnitude range including very small earthquakes is inspected, it must be taken into account that smaller earthquakes with higher frequency content are more sensitive to local structural changes than larger events. Therefore, improving our understanding of attenuation effects on high-frequency waves could reduce uncertainties in the spectral analysis, leading to a decrease of error sources during the estimation of earthquake source parameters.

The M_W-$\Delta\sigma$ relation obtained for the laboratory AE study is so far only one of very few observations made within this magnitude range. Performing more experiments with variable boundary conditions could reveal whether the observations made in this study are valid for the laboratory scale in general or are unique and tight to the type of presented experiments. Boundary conditions incorporate for instance the type of rock sample (granite, sand stone, marble, basalt etc.), temperature changes, fluid involvement and mechanical stress changes (e.g. confining pressure).

An enlarged data catalog of laboratory experiments could also help to move us one step closer towards identifying physical key properties of the seismic source that may vary between individual earthquakes and study sites, and lead to discrepancies observed in

global and individual stress drop estimates. As it was shown in this thesis, both laboratory experiments suggest that roughness is a factor that influences static stress drop and event magnitude. The aspect of roughness influence on earthquake source scaling and stress drop can also be studied at larger scales. As roughness is assumed to decrease with time and slip history, mature faults are thought to be smoother than immature faults. Different to laboratory faults, natural faults are not directly accessible and it is very difficult to approximate their roughness and further properties. However, this study showed that the geometric complexity of fault plane surfaces can be indirectly derived from seismic data. The investigation of full moment tensor solutions, for instance, can give indication on fault complexity. It is suggested that smooth faults produce simpler shear motions than rough faults, resulting in larger magnitudes and static stress drops and consequently in larger ground motions. The hypocentral distributions of e.g. aftershock sequences could also provide information on fault zone width. Rough laboratory faults exhibit broader damage zones in which seismicity can occur off the main fault plane. Thus, a broad distribution of aftershocks might give indication on fault zone complexity. Another aspect coupled to roughness might be the complexity of slip distribution. Clustering of events around fault patches could reveal areas of increased local stresses (e.g. at asperities). This might provide indication that a fault is characterized by weaker and stronger areas and therefore consist of both rough and smooth fault segments.

As we increase our understanding of the factors controlling seismic stress drop and its variability, we can increasingly refine the hazard and risk assessments for future earthquakes.

Bibliography

Abercrombie, R. E. (1995). Earthquake source scaling relationships from -1 to 5 ML using seismograms recorded at 2.5-km depth. *Journal of Geophysical Research: Solid Earth*, 100(B12):24015–24036. doi:10.1029/95JB02397.

Abercrombie, R. E. (2013). Comparison of direct and coda wave stress drop measurements for the Wells, Nevada, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, 118(4):1458–1470. doi:10.1029/2012JB009638.

Abercrombie, R. E. (2015). Investigating uncertainties in empirical Green's function analysis of earthquake source parameters. *Journal of Geophysical Research: Solid Earth*, 120(6):4263–4277. doi:10.1002/2015JB011984.

Abercrombie, R. E., Bannister, S., Ristau, J., and Doser, D. (2016). Variability of earthquake stress drop in a subduction setting, the Hikurangi Margin, New Zealand. *Geophysical Journal International*, pages 306–320. doi:10.1093/gji/ggw393.

Abercrombie, R. E. and Rice, J. R. (2005). Can observations of earthquake scaling constrain slip weakening? *Geophysical Journal International*, 162(2):406–424. doi:10.1111/j.1365-246X.2005.02579.x.

Aki, K. (1967). Scaling law of seismic spectrum. *Journal of Geophysical Research*, 72(4):1217–1231. doi:10.1029/JZ072i004p01217.

Aki, K. (1969). Analysis of the seismic coda of local earthquakes as scattered waves. *Journal of Geophysical Research*, 74(2):615–631. doi:10.1029/JB074i002p00615.

Aki, K. (1980a). Attenuation of shear-waves in the lithosphere for frequencies from 0.05 to 25 Hz. *Physics of the Earth and Planetary Interiors*, 21(1):50–60. doi:10.1016/0031-9201(80)90019-9.

Aki, K. (1980b). Scattering and attenuation of shear waves in the lithosphere. *Journal of Geophysical Research: Solid Earth*, 85(B11):6496–6504. doi:10.1029/JB085iB11p06496.

Aki, K. (1981). Attenuation and scattering of short-period seismic waves in the lithosphere. In *Husebye E.S., Mykkeltveit S. (eds) Identification of Seismic Sources–Earthquake or Underground Explosion. Nato Advanced Study Institutes Series (Serie C–Mathematical and Physical Sciences)*, volume 74, pages 515–541. Springer. doi:10.1007/978-94-009-8531-5_28.

Aki, K. and Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, 80(23):3322–3342. doi:10.1029/JB080i023p03322.

Allmann, B. P. and Shearer, P. M. (2009). Global variations of stress drop for moderate to large earthquakes. *Journal of Geophysical Research: Solid Earth*, 114(B1). doi:10.1029/2008JB005821.

Baltay, A., Ide, S., Prieto, G., and Beroza, G. (2011). Variability in earthquake stress drop and apparent stress. *Geophysical Research Letters*, 38(6). doi:10.1029/2011GL046698.

Baltay, A. S., Hanks, T. C., and Abrahamson, N. A. (2019). Earthquake stress drop and Arias intensity. *Journal of Geophysical Research: Solid Earth*, 124(4):3838–3852. doi:10.1029/2018JB016753.

Beall, J. and Wright, M. (2010). Southern extent of The Geysers high temperature reservoir based on seismic and geochemical evidence. *Geothermal Resources Council Transactions*, 34:1199–1202.

Beeler, N. (2006). Inferring earthquake source properties from laboratory observations and the scope of lab contributions to source physics. *In: Earthquakes: Radiated Energy and the Physics of Faulting (eds R. Abercrombie, A. McGarr, G. Di Toro and H. Kanamori*, 170:99–119. doi:10.1029/170GM12.

Ben-Zion, Y. and Sammis, C. G. (2003). Characterization of fault zones. *Pure and Applied Geophysics*, 160(3-4):677–715. doi:10.1007/PL00012554.

Blanke, A., Goebel, T. H. W., and Kwiatek, G. (2020). Acoustic emission source parameters of laboratory triaxial stick-slip experiments on two Westerly granite samples. doi:10.5880/GFZ.4.2.2020.008.

Blanke, A., Kwiatek, G., Goebel, T. H. W., Bohnhoff, M., and Dresen, G. (2021). Stress drop–magnitude dependence of acoustic emissions during laboratory stick-slip. *Geophysical Journal International*, 224:1371–1380. doi:10.1093/gji/ggaa524.

Blanke, A., Kwiatek, G., Martínez-Garzón, P., and Bohnhoff, M. (2018). Supplementary material to "Sensitivity and stability analysis of coda quality factors at The Geysers geothermal field, California. doi:10.5880/GFZ.4.2.2018.002.

Blanke, A., Kwiatek, G., Martínez-Garzón, P., and Bohnhoff, M. (2019). Sensitivity and stability analysis of coda quality factors at The Geysers Geothermal

Field, California. *Bulletin of the Seismological Society of America*, 109(3):959–975. doi:10.1785/0120180219.

Boatwright, J. (1978). Detailed spectral analysis of two small New York State earthquakes. *Bulletin of the Seismological Society of America*, 68(4):1117–1131.

Bohnhoff, M., Dresen, G., Ellsworth, W. L., and Ito, H. (2010). Passive seismic monitoring of natural and induced earthquakes: Case studies, future directions and socioeconomic relevance. In *New Frontiers in Integrated Solid Earth Sciences*, pages 261–285. Springer. doi:10.1007/978-90-481-2737-5_7.

Bohnhoff, M., Kwiatek, G., and Dresen, G. (2016a). Von der Gesteinsprobe bis zur Plattengrenze: Skalenübergreifende Analyse von Bruchprozessen. *System Erde*, 6(1):50–55. doi:10.2312/GFZ.syserde.06.01.8.

Bohnhoff, M., Martínez-Garzón, P., Bulut, F., Stierle, E., and Ben-Zion, Y. (2016b). Maximum earthquake magnitudes along different sections of the North Anatolian fault zone. *Tectonophysics*, 674:147–165. doi:10.1016/j.tecto.2016.02.028.

Brace, W. and Byerlee, J. (1966). Stick-slip as a mechanism for earthquakes. *Science*, 153(3739):990–992. doi:10.1126/science.153.3739.990.

Brodsky, E. E., Gilchrist, J. J., Sagy, A., and Collettini, C. (2011). Faults smooth gradually as a function of slip. *Earth and Planetary Science Letters*, 302(1-2):185–193. doi:10.1016/j.epsl.2010.12.010.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of Geophysical Research*, 75(26):4997–5009. doi:10.1029/JB075i026p04997.

Calvet, M. and Margerin, L. (2013). Lapse-time dependence of coda Q: Anisotropic multiple-scattering models and application to the Pyrenees. *Bulletin of the Seismological Society of America*, 103(3):1993–2010. doi:10.1785/0120120239.

Campillo, M., Favreau, P., Ionescu, I. R., and Voisin, C. (2001). On the effective friction law of a heterogeneous fault. *Journal of Geophysical Research: Solid Earth*, 106(B8):16307–16322. doi:10.1029/2000JB900467.

Campillo, M., Margerin, L., and Shapiro, N. (1999). Seismic wave diffusion in the earth lithosphere. In *Fouque JP. (eds) Diffuse Waves in Complex Media. NATO Science Series (Series C: Mathematical and Physical Sciences)*, pages 383–404. Springer. doi:10.1007/978-94-011-4572-5_13.

Candela, T., Renard, F., Bouchon, M., Schmittbuhl, J., and Brodsky, E. E. (2011). Stress drop during earthquakes: Effect of fault roughness scaling. *Bulletin of the Seismological Society of America*, 101(5):2369–2387. doi:10.1785/0120100298.

Causse, M., Dalguer, L., and Mai, P. M. (2014). Variability of dynamic source parameters inferred from kinematic models of past earthquakes. *Geophysical Journal International*, 196(3):1754–1769. doi:10.1093/gji/ggt478.

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

Cotton, F., Archuleta, R., and Causse, M. (2013). What is sigma of the stress drop? *Seismological Research Letters*, 84(1):42–48. doi:10.1785/0220120087.

Dainty, A. M. and Toksöz, M. N. (1981). Seismic codas on the earth and the moon: A comparison. *Physics of the Earth and Planetary Interiors*, 26(4):250–260. doi:10.1016/0031-9201(81)90029-7.

Davies, R., Foulger, G., Bindley, A., and Styles, P. (2013). Induced seismicity and hydraulic fracturing for the recovery of hydrocarbons. *Marine and Petroleum Geology*, 45:171–185. doi:10.1016/j.marpetgeo.2013.03.016.

Deichmann, N. and Giardini, D. (2009). Earthquakes induced by the stimulation of an enhanced geothermal system below Basel (Switzerland). *Seismological Research Letters*, 80(5):784–798. doi:10.1785/gssrl.80.5.784.

Del Pezzo, E., Bianco, F., and Saccorotti, G. (2001). Separation of intrinsic and scattering Q for volcanic tremor: An application to Etna and Masaya volcanoes. *Geophysical Research Letters*, 28(16):3083–3086. doi:10.1029/2001GL013372.

Dresen, G. and Evans, B. (1993). Brittle and semibrittle deformation of synthetic marbles composed of two phases. *Journal of Geophysical Research: Solid Earth*, 98(B7):11921–11933.

Dresen, G., Kwiatek, G., and Ben-Zion, Y. (2020). Partitioning of seismic and aseismic preparatory processes before stick-slip failure. *Pure Applied Geophysics*. doi:10.1007/s00024-020-02605-x.

Eaton, D. W. (2011). Q determination, corner frequency and spectral characteristics of microseismicity induced by hydraulic fracturing. In *SEG Technical Program Expanded Abstracts 2011*, pages 1555–1559. Society of Exploration Geophysicists. doi:10.1190/1.3627499.

Elkibbi, M. and Rial, J. (2003). Shear-wave splitting: An efficient tool to detect 3D fracture patterns at The Geysers, California. In *Proceedings, 28th Stanford Workshop on Geothermal Reservoir Engineering*, pages 143–149. Stanford University, Stanford CA.

Elkibbi, M., Yang, M., and Rial, J. (2005). Crack-induced anisotropy models in The Geysers geothermal field. *Geophysical Journal International*, 162(3):1036–1048. doi: 10.1111/j.1365-246X.2005.02697.x.

Erten, D., Elkibbi, M., and Rial, J. (2001). Shear wave splitting and fracture patterns at The Geysers (California) geothermal field. In *Proceedings of the Workshop on Geothermal Reservoir Engineering*, volume 26, pages 139–147.

Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 241(1226):376–396. doi:10.1098/rspa.1957.0133.

Etiope, G. and Martinelli, G. (2002). Migration of carrier and trace gases in the geosphere: An overview. *Physics of the Earth and Planetary Interiors*, 129(3-4):185–204. doi: 10.1016/S0031-9201(01)00292-8.

Fehler, M., Hoshiba, M., Sato, H., and Obara, K. (1992). Separation of scattering and intrinsic attenuation for the Kanto-Tokai region, Japan, using measurements of S-wave energy versus hypocentral distance. *Geophysical Journal International*, 108(3):787–800. doi:10.1111/j.1365-246X.1992.tb03470.x.

Foulger, G. R., Wilson, M. P., Gluyas, J. G., Julian, B. R., and Davies, R. J. (2018). Global review of human-induced earthquakes. *Earth-Science Reviews*, 178:438–514. doi:10.1016/j.earscirev.2017.07.008.

Frankel, A. (1991). Mechanisms of seismic attenuation in the crust: Scattering and anelasticity in New York State, South Africa, and southern California. *Journal of Geophysical Research: Solid Earth*, 96(B4):6269–6289. doi:10.1029/91JB00192.

Frankel, A., Fletcher, J., Vernon, F., Haar, L., Berger, J., Hanks, T., and Brune, J. (1986). Rupture characteristics and tomographic source imaging of ML 3 earthquakes near Anza, southern California. *Journal of Geophysical Research: Solid Earth*, 91(B12):12633–12650. doi:10.1029/JB091iB12p12633.

Frankel, A. and Wennerberg, L. (1987). Energy-flux model of seismic coda: Separation of scattering and intrinsic attenuation. *Bulletin of the Seismological Society of America*, 77(4):1223–1251.

Fuchs, K. (2006). The great earthquakes of Lisbon 1755 and Aceh 2004 shook the world. Seismologists' societal responsibility. *European Review*, 14(2):207–219. doi:10.1007/978-1-4020-8609-0_4.

Gao, L., Lee, L., Biswas, N., and Aki, K. (1983). Comparison of the effects between single and multiple scattering on coda waves for local earthquakes. *Bulletin of the Seismological Society of America*, 73(2):377–389.

Garcia-Aristizabal, A., Kocot, J., Russo, R., and Gasparini, P. (2019). A probabilistic tool for multi-hazard risk analysis using a bow-tie approach: Application to environmental risk assessments for geo-resource development projects. *Acta Geophysica*, 67(1):385–410. doi:10.1007/s11600-018-0201-7.

Gibowicz, S. and Kijko, A. (1994). *An introduction to mining seismology, In International Geophysics*, volume 55. Academic Press, San Diego, California.

Goebel, T. H. W., Becker, T., Sammis, C., Dresen, G., and Schorlemmer, D. (2014a). Off-fault damage and acoustic emission distributions during the evolution of structurally complex faults over series of stick-slip events. *Geophysical Journal International*, 197(3):1705–1718. doi:10.1093/gji/ggu074.

Goebel, T. H. W., Becker, T., Schorlemmer, D., Stanchits, S., Sammis, C., Rybacki, E., and Dresen, G. (2012). Identifying fault heterogeneity through mapping spatial anomalies in acoustic emission statistics. *Journal of Geophysical Research: Solid Earth*, 117(B3). doi:10.1029/2011JB008763.

Goebel, T. H. W., Candela, T., Sammis, C., Becker, T., Dresen, G., and Schorlemmer, D. (2014b). Seismic event distributions and off-fault damage during frictional sliding of saw-cut surfaces with pre-defined roughness. *Geophysical Journal International*, 196(1):612–625. doi:10.1093/gji/ggt401.

Goebel, T. H. W., Kwiatek, G., Becker, T. W., Brodsky, E. E., and Dresen, G. (2017). What allows seismic events to grow big?: Insights from b-value and fault roughness analysis in laboratory stick-slip experiments. *Geology*, 45(9):815–818. doi:10.1130/G39147.1.

Goebel, T. H. W., Sammis, C., Becker, T., Dresen, G., and Schorlemmer, D. (2015). A comparison of seismicity characteristics and fault structure between stick–slip experiments and nature. *Pure and Applied Geophysics*, 172(8):2247–2264. doi:10.1007/s00024-013-0713-7.

Goebel, T. H. W., Schorlemmer, D., Becker, T., Dresen, G., and Sammis, C. (2013). Acoustic emissions document stress changes over many seismic cycles in stick-slip experiments. *Geophysical Research Letters*, 40(10):2049–2054. doi:10.1002/grl.50507.

Goodfellow, S. and Young, R. (2014). A laboratory acoustic emission experiment under in situ conditions. *Geophysical Research Letters*, 41(10):3422–3430. doi:10.1002/2014GL059965.

Grigoli, F., Cesca, S., Priolo, E., Rinaldi, A. P., Clinton, J. F., Stabile, T. A., Dost, B., Fernandez, M. G., Wiemer, S., and Dahm, T. (2017). Current challenges in monitoring, discrimination, and management of induced seismicity related to underground industrial activities: A European perspective. *Reviews of Geophysics*, 55(2):310–340. doi:10.1002/2016RG000542.

Gritto, R., Yoo, S.-H., and Jarpe, S. P. (2013). Three-dimensional seismic tomography at The Geysers geothermal field, CA, USA. *Proceedings of the Thirty-Eighth Work-shop on Geothermal Reservoir Engineering*, 1113:1–12.

Gupta, H. K. (2002). A review of recent studies of triggered earthquakes by artificial water reservoirs with special emphasis on earthquakes in Koyna, India. *Earth-Science Reviews*, 58(3-4):279–310. doi:10.1016/S0012-8252(02)00063-6.

Gutenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in California. *Bulletin of the Seismological Society of America*, 34(4):185–188.

Harrington, R. M., Kwiatek, G., and Moran, S. C. (2015). Self-similar rupture implied by scaling properties of volcanic earthquakes occurring during the 2004-2008 eruption of Mount St. Helens, Washington. *Journal of Geophysical Research: Solid Earth*, 120(7):4966–4982. doi:10.1002/2014JB011744.

Hartline, C. S., Walters, M. A., Wright, M. C., Forson, C. K., and Sadowski, A. J. (2015). Three-dimensional structural model building, induced seismicity analysis, drilling analysis, and reservoir management at The Geysers geothermal field, northern California. *Geothermal Resources Council Transactions*, 39:603–614.

Hartzell, S. H. (1978). Earthquake aftershocks as Green's functions. *Geophysical Research Letters*, 5(1):1–4. doi:10.1029/GL005i001p00001.

Havskov, J., Sørensen, M. B., Vales, D., Özyazıcıoğlu, M., Sánchez, G., and Li, B. (2016). Coda Q in different tectonic areas, influence of processing parameters. *Bulletin of the Seismological Society of America*, 106(3):956–970. doi:10.1785/0120150359.

Herrmann, R. B. (1975). The use of duration as a measure of seismic moment and magnitude. *Bulletin of the Seismological Society of America*, 65(4):899–913.

Huang, Y., Beroza, G. C., and Ellsworth, W. L. (2016). Stress drop estimates of potentially induced earthquakes in the Guy-Greenbrier sequence. *Journal of Geophysical Research: Solid Earth*, 121(9):6597–6607. doi:10.1002/2016JB013067.

Huang, Y., Ellsworth, W. L., and Beroza, G. C. (2017). Stress drops of induced and tectonic earthquakes in the central United States are indistinguishable. *Science Advances*, 3(8):e1700772. doi:10.1126/sciadv.1700772.

Ibanez, J., Del Pezzo, E., De Miguel, F., Herraiz, M., Alguacil, G., and Morales, J. (1990). Depth-dependent seismic attenuation in the Granada zone (Southern Spain). *Bulletin of the Seismological Society of America*, 80(5):1232–1244.

Ide, S. and Beroza, G. C. (2001). Does apparent stress vary with earthquake size? *Geophysical Research Letters*, 28(17):3349–3352. doi:10.1029/2001GL013106.

Ide, S., Beroza, G. C., Prejean, S. G., and Ellsworth, W. L. (2003). Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *Journal of Geophysical Research: Solid Earth*, 108(B5). doi:10.1029/2001JB001617.

Imanishi, K. and Ellsworth, W. L. (2006). Source scaling relationships of microearthquakes at Parkfield, CA, determined using the SAFOD pilot hole seismic array. *In Earthquakes: Radiated Energy and the Physics of Faulting (eds. R. Abercrombie, A. McGarr, G. Di Toro and H. Kanamori)*, 170:81–90. doi:10.1029/170GM10.

Imanishi, K. and Uchide, T. (2017). Non-self-similar source property for microforeshocks of the 2014 Mw 6.2 Northern Nagano, central Japan, earthquake. *Geophysical Research Letters*, 44(11):5401–5410. doi:10.1002/2017GL073018.

Jin, A. and Aki, K. (1989). Spatial and temporal correlation between coda Q-1 and seismicity and its physical mechanism. *Journal of Geophysical Research: Solid Earth*, 94(B10):14041–14059. doi:10.1029/JB094iB10p14041.

Kanamori, H. (1977). The energy release in great earthquakes. *Journal of geophysical research*, 82(20):2981–2987. doi:10.1029/JB082i020p02981.

Kanamori, H. and Brodsky, E. E. (2001). The physics of earthquakes. *Physics Today*, 54(6):34–40.

Kanamori, H. and Brodsky, E. E. (2004). The physics of earthquakes. *Reports on Progress in Physics*, 67(8):1429. doi:10.1088/0034-4885/67/8/R03.

Kanamori, H. and Rivera, L. (2004). Static and dynamic scaling relations for earthquakes and their implications for rupture speed and stress drop. *Bulletin of the Seismological Society of America*, 94(1):314–319. doi:10.1785/0120030159.

Kanamori, H. and Rivera, L. (2006). Energy partitioning during an earthquake. *In Earthquakes: Radiated Energy and the Physics of Faulting (eds R. Abercrombie, A. McGarr, G. Di Toro and H. Kanamori)*.

Kaneko, Y. and Shearer, P. (2014). Seismic source spectra and estimated stress drop derived from cohesive-zone models of circular subshear rupture. *Geophysical Journal International*, 197(2):1002–1015. doi:10.1093/gji/ggu030.

Kaneko, Y. and Shearer, P. (2015). Variability of seismic source spectra, estimated stress drop, and radiated energy, derived from cohesive-zone models of symmetrical and asymmetrical circular and elliptical ruptures. *Journal of Geophysical Research: Solid Earth*, 120(2):1053–1079. doi:10.1002/2014JB011642.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680. doi:10.1126/science.220.4598.671.

Kwiatek, G. and Ben-Zion, Y. (2016). Theoretical limits on detection and analysis of small earthquakes. *Journal of Geophysical Research: Solid Earth*, 121(8):5898–5916. doi:10.1002/2016JB012908.

Kwiatek, G., Bulut, F., Bohnhoff, M., and Dresen, G. (2014a). High-resolution analysis of seismicity induced at Berlín geothermal field, El Salvador. *Geothermics*, 52:98–111. doi:10.1016/j.geothermics.2013.09.008.

Kwiatek, G., Charalampidou, E.-M., Dresen, G., and Stanchits, S. (2014b). An improved method for seismic moment tensor inversion of acoustic emissions through assessment of sensor coupling and sensitivity to incidence angle. *International Journal of Rock Mechanics and Mining Sciences*, 65:153–161. doi:10.1016/j.ijrmms.2013.11.005.

Kwiatek, G., Goebel, T. H. W., and Dresen, G. (2014c). Seismic moment tensor and b value variations over successive seismic cycles in laboratory stick-slip experiments. *Geophysical Research Letters*, 41(16):5838–5846. doi:10.1002/2014GL060159.

Kwiatek, G., Martínez-Garzón, P., Dresen, G., Bohnhoff, M., Sone, H., and 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 of Geophysical Research: Solid Earth*, 120(10):7085–7101. doi:10.1002/2015JB012362.

Kwiatek, G., Plenkers, K., Dresen, G., and Group, J. R. (2011). Source parameters of picoseismicity recorded at Mponeng deep gold mine, South Africa: Implications for scaling relations. *Bulletin of the Seismological Society of America*, 101(6):2592–2608. doi:10.1785/0120110094.

Lasocki, S. and Orlecka-Sikora, B. (2020). Anthropogenic seismicity related to exploitation of georesources. *In: Gupta H. (eds) Encyclopedia of Solid Earth Geophysics. Encyclopedia of Earth Sciences Series, Springer, Cham.* doi:10.1007/978-3-030-10475-7_277-1.

Lasocki, S., Orlecka-Sikora, B., Mutke, G., Pytel, W. M., Rudziński, Ł., Markowski, P., and Piasecki, P. (2017). A catastrophic event in Rudna copper-ore mine in Poland on 29 November, 2016: what, how and why. *Rock Mass Response to Mining and Underground Construction*, pages 316–324. doi:10.1126/science.aax1878.

Lay, T. and Wallace, T. C. (1995). *Modern global seismology*, volume 58. Academic Press, San Diego.

Lee, K.-K., Ellsworth, W. L., Giardini, D., Townend, J., Ge, S., Shimamoto, T., Yeo, I.-W., Kang, T.-S., Rhie, J., Sheen, D.-H., et al. (2019). Managing injection-induced seismic risks. *Science*, 364(6442):730–732. doi:10.1126/science.aax1878.

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

Leptokaropoulos, K., Cielesta, S., Staszek, M., Olszewska, D., Lizurek, G., Kocot, J., Lasocki, S., Orlecka-Sikora, B., Sterzel, M., and Szepieniec, T. (2019). IS-EPOS: a platform for anthropogenic seismicity research. *Acta Geophysica*, 67(1):299–310. doi:10.1007/s11600-018-0209-z.

Lockner, D. (1993). The role of acoustic emission in the study of rock fracture. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 30(7):883–899. doi:10.1016/0148-9062(93)90041-B.

Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bulletin of the Seismological Society of America*, 66(3):639–666.

Madariaga, R. (2015). 4.02 - Seismic source theory. In Schubert, G., editor, *Treatise on Geophysics (Second Edition)*, pages 51 – 71. Elsevier, Oxford. doi:10.1016/B978-0-444-53802-4.00070-1.

Mai, P. M. and Beroza, G. C. (2002). A spatial random field model to characterize complexity in earthquake slip. *Journal of Geophysical Research: Solid Earth*, 107(B11):ESE–10. doi:10.1029/2001JB000588.

Malagnini, L., Mayeda, K., Nielsen, S., Yoo, S.-H., Rawles, C., Boschi, E., et al. (2014). Scaling transition in earthquake sources: A possible link between seismic and laboratory measurements. *Pure and Applied Geophysics*, 171(10):2685–2707. doi:10.1007/s00024-013-0749-8.

Margerin, L., Campillo, M., and Tiggelen, B. (1998). Radiative transfer and diffusion of waves in a layered medium: new insight into coda Q. *Geophysical Journal International*, 134(2):596–612. doi:10.1111/j.1365-246X.1998.tb07142.x.

Martínez-Garzón, P., Bohnhoff, M., Ben-Zion, Y., and Dresen, G. (2015). Scaling of maximum observed magnitudes with geometrical and stress properties of strike-slip faults. *Geophysical Research Letters*, 42(23):10–230. doi:10.1002/2015GL066478.

Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., and Dresen, G. (2016). Impact of fluid injection on fracture reactivation at The Geysers geothermal field. *Journal of Geophysical Research: Solid Earth*, 121(10):7432–7449. doi:10.1002/2016JB013137.

Martínez-Garzón, P., Kwiatek, G., Bohnhoff, M., and Dresen, G. (2017). Volumetric components in the earthquake source related to fluid injection and stress state. *Geophysical Research Letters*, 44(2):800–809. doi:10.1002/2016GL071963.

Martínez-Garzón, P., Kwiatek, G., Sone, H., Bohnhoff, M., Dresen, G., and Hartline, C. (2014). Spatiotemporal changes, faulting regimes, and source parameters of induced seismicity: A case study from The Geysers geothermal field. *Journal of Geophysical Research: Solid Earth*, 119(11):8378–8396. doi:10.1002/2014JB011385.

Mayeda, K. (1993). mb(LgCoda): A stable single station estimator of magnitude. *Bulletin of the Seismological Society of America*, 83(3):851–861.

Mayeda, K., Koyanagi, S., Hoshiba, M., Aki, K., and Zeng, Y. (1992). A comparative study of scattering, intrinsic, and coda Q- 1 for Hawaii, Long Valley, and central California between 1.5 and 15.0 Hz. *Journal of Geophysical Research: Solid Earth*, 97(B5):6643–6659. doi:10.1029/91JB03094.

Mayeda, K., Malagnini, L., and Walter, W. R. (2007). A new spectral ratio method using narrow band coda envelopes: Evidence for non-self-similarity in the Hector Mine sequence. *Geophysical Research Letters*, 34(11). doi:10.1029/2007GL030041.

Mayeda, K. and Walter, W. R. (1996). Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes. *Journal of Geophysical Research: Solid Earth*, 101(B5):11195–11208. doi:10.1029/96JB00112.

McGarr, A. and Fletcher, J. B. (2003). Maximum slip in earthquake fault zones, apparent stress, and stick-slip friction. *Bulletin of the Seismological Society of America*, 93(6):2355–2362. doi:10.1785/0120030037.

McLaskey, G. C., Kilgore, B. D., Lockner, D. A., and Beeler, N. M. (2014). Laboratory generated M-6 earthquakes. *Pure and Applied Geophysics*, 171(10):2601–2615. doi:10.1007/s00024-013-0772-9.

Meier, M. (2007). Zur Terminologie der (Natur-)Katastrophe in der griechischen Historiographie: Einige einleitende Anmerkungen. *Historical Social Research*, 32(3):44–56. doi:10.12759/hsr.32.2007.3.44-56.

Mogi, K. (1962). Study of elastic shocks caused by the fracture of heterogeneous materials and its relations to earthquake phenomena. *Bulletin of the Earthquake Research Institute, University of Tokio*, 40(1):125–173.

Morozov, I. B., Zhang, C., Duenow, J. N., Morozova, E. A., and Smithson, S. B. (2008). Frequency dependence of coda Q, part I: numerical modeling and examples from peaceful nuclear explosions. *Bulletin of the Seismological Society of America*, 98(6):2615–2628. doi:10.1785/0120080037.

Mueller, C. S. (1985). Source pulse enhancement by deconvolution of an empirical Green's function. *Geophysical Research Letters*, 12(1):33–36. doi:10.1029/GL012i001p00033.

Mukhopadhyay, S., Sharma, J., Massey, R., and Kayal, J. (2008). Lapse-time dependence of coda Q in the source region of the 1999 Chamoli earthquake. *Bulletin of the Seismological Society of America*, 98(4):2080–2086. doi:10.1785/0120070258.

Novelo-Casanova, D. A. and Lee, W. (1991). Comparison of techniques that use the single scattering model to compute the quality factor Q from coda waves. *Pure and Applied Geophysics*, 135(1):77–89. doi:10.1007/BF00877010.

Orlecka-Sikora, B., Cielesta, S., and Lasocki, S. (2019). Tracking the development of seismic fracture network from The Geysers geothermal field. *Acta Geophysica*, 67(1):341–350. doi:10.1007/s11600-018-0202-6.

Orlecka-Sikora, B., Lasocki, S., Kocot, J., Szepieniec, T., Grasso, J. R., Garcia-Aristizabal, A., Schaming, M., Urban, P., Jones, G., Stimpson, I., et al. (2020). An open data infrastructure for the study of anthropogenic hazards linked to georesource exploitation. *Scientific Data*, 7(1):1–16. doi:10.1038/s41597-020-0429-3.

Osborn, S. G., Vengosh, A., Warner, N. R., and Jackson, R. B. (2011). Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proceedings of the National Academy of Sciences*, 108(20):8172–8176. doi:10.1073/pnas.1100682108.

Othman, F., Yu, M., Kamali, F., and Hussain, F. (2018). Fines migration during supercritical CO2 injection in sandstone. *Journal of Natural Gas Science and Engineering*, 56:344–357. doi:10.1016/j.jngse.2018.06.001.

Percival, D. B., Walden, A. T., et al. (1993). *Spectral analysis for physical applications–Multitaper and Conventional Univariate Techniques*. Cambridge University Press, UK.

Phillips, W. S. (1985). *The separation of source, path and site effects on high frequency seismic waves: an analysis using coda wave techniques*. PhD thesis, Massachusetts Institute of Technology.

Phillips, W. S. and Aki, K. (1986). Site amplification of coda waves from local earthquakes in central California. *Bulletin of the Seismological Society of America*, 76(3):627–648.

Phillips, W. S., Lee, W., and Newberry, J. T. (1988). Spatial variation of crustal coda Q in California. In *Aki K., Wu RS. (eds) Scattering and Attenuations of Seismic Waves, Part I Pageoph Topical Volumes*, pages 251–260. Birkhäuser, Basel.

Picozzi, M., Oth, A., Parolai, S., Bindi, D., De Landro, G., and Amoroso, O. (2017). Accurate estimation of seismic source parameters of induced seismicity by a combined approach of generalized inversion and genetic algorithm: Application to The Geysers geothermal area, California. *Journal of Geophysical Research: Solid Earth*, 122(5):3916–3933. doi:10.1002/2016JB013690.

Plenkers, K., Schorlemmer, D., Kwiatek, G., and Group, J. R. (2011). On the probability of detecting picoseismicity. *Bulletin of the Seismological Society of America*, 101(6):2579–2591. doi:10.1785/0120110017.

Prieto, G. A., Shearer, P. M., Vernon, F. L., and Kilb, D. (2004). Earthquake source scaling and self-similarity estimation from stacking P and S spectra. *Journal of Geophysical Research: Solid Earth*, 109(B8). doi:10.1029/2004JB003084.

Rautian, T. and Khalturin, V. (1978). The use of the coda for determination of the earthquake source spectrum. *Bulletin of the Seismological Society of America*, 68(4):923–948.

Reid, H. F. (1910). The mechanics of the earthquake. *The California Earthquake of April 18, 1906, Report of the State Earthquake Investigation Commission*.

Ripperger, J., Ampuero, J.-P., Mai, P., and Giardini, D. (2007). Earthquake source characteristics from dynamic rupture with constrained stochastic fault stress. *Journal of Geophysical Research: Solid Earth*, 112(B4). doi:0.1029/2006JB004515.

Rivard, C., Lavoie, D., Lefebvre, R., Séjourné, S., Lamontagne, C., and Duchesne, M. (2014). An overview of Canadian shale gas production and environmental concerns. *International Journal of Coal Geology*, 126:64–76. doi:10.1016/j.coal.2013.12.004.

Ross, A., Foulger, G. R., and Julian, B. R. (1999). Source processes of industrially-induced earthquakes at The Geysers geothermal area, California. *Geophysics*, 64(6):1877–1889. doi:10.1190/1.1444694.

Rudziński, Ł. et al. (2017). IS-EPOS Platform for anthropogenic seismicity research: A modern e-tool of EPOS Thematic Core Service "Anthropogenic Hazards", designed to integrate infrastructures and facilitate research of human induced seismicity. In *Proc. 9th International Symposium on Rockbursts and Seismicity in Mines–RaSiM9, November 15-17, Santiago, Chile (Vallejos, J. A. ed)*, pages 15–17.

Ruhl, C., Abercrombie, R., and Smith, K. (2017). Spatiotemporal variation of stress drop during the 2008 Mogul, Nevada, earthquake swarm. *Journal of Geophysical Research: Solid Earth*, 122(10):8163–8180. doi:10.1002/2017JB014601.

Sadowski, A., Forson, C., Walters, M., and Hartline, C. (2016). Compilation surface geologic map for use in three-dimensional structural model building at The Geysers geothermal field, northern California. In *Proceedings: 41st Workshop on Geothermal Reservoir Engineering*, volume 41. Stanford University, Stanford, California.

Sagy, A., Brodsky, E. E., and Axen, G. J. (2007). Evolution of fault-surface roughness with slip. *Geology*, 35(3):283–286. doi:10.1130/G23235A.1.

Sato, H. (1977). Energy propagation including scattering effects sengle isotropic scattering approximation. *Journal of Physics of the Earth*, 25(1):27–41. doi:10.4294/jpe1952.25.27.

Sato, H. (1978). Mean free path of S-waves under the Kanto district of Japan. *Journal of Physics of the Earth*, 26(2):185–198. doi:10.4294/jpe1952.26.185.

Sato, H. (1982). Coda wave excitation due to nonisotropic scattering and nonspherical source radiation. *Journal of Geophysical Research: Solid Earth*, 87(B10):8665–8674. doi:10.1029/JB087iB10p08665.

Sato, H., Fehler, M., Wu, and Ru-Shan (2002). Scattering and attenuation of seismic waves in the lithosphere. *In Lee W.H.K., Kanamori H., Jennings P.C. and Kisslinger C. (eds): International Handbook of Earthquake & Engineering Seismology*, 81(A):195–208.

Sato, H., Fehler, M. C., and Maeda, T. (2012). *Seismic wave propagation and scattering in the heterogeneous earth: Second Edition*, volume 496. Springer. doi:10.1007/978-3-642-23029-5.

Sato, T. and Hirasaw, T. (1973). Body wave spectra from propagating shear cracks. *Journal of Physics of the Earth*, 21(4):415–431. doi:10.4294/jpe1952.21.415.

Scholz, C. (1968). Experimental study of the fracturing process in brittle rock. *Journal of Geophysical Research*, 73(4):1447–1454. doi:10.1029/JB073i004p01447.

Sen, M. and Stoffa, P. (1995). Global optimization methods in geophysical inversion. *Advances in Exploration Geophysics*, volume 4.

Şengör, A., Tüysüz, O., Imren, C., Sakınç, M., Eyidoğan, H., Görür, N., Le Pichon, X., and Rangin, C. (2005). The North Anatolian fault: A new look. *Annual Review of Earth Planetary Sciences*, 33:37–112. doi:10.1146/annurev.earth.32.101802.120415.

Shapiro, N., Campillo, M., Margerin, L., Singh, S., Kostoglodov, V., and Pacheco, J. (2000). The energy partitioning and the diffusive character of the seismic coda. *Bulletin of the Seismological Society of America*, 90(3):655–665. doi:10.1785/ 0119990021.

Shapiro, S. and Kneib, G. (1993). Seismic attenuation by scattering: Theory and numerical results. *Geophysical Journal International*, 114(2):373–391. doi:10.1111/j.1365-246X.1993.tb03925.x.

Shearer, P. M. (2009). *Introduction to Seismology (Second Edition)*. Cambridge University Press.

Shearer, P. M., Abercrombie, R. E., Trugman, D. T., and Wang, W. (2019). Comparing EGF methods for estimating corner frequency and stress drop from P wave spectra. *Journal of Geophysical Research: Solid Earth*, 124(4):3966–3986. doi:10.1029/2018JB016957.

Sintubin, M. (2018). The groningen case: When science becomes part of the problem, not the solution. *Seismological Research Letters*, 89(6):2001–2007. doi:10.1785/0220180203.

Sonley, E., Abercrombie, R. E., Abercrombie, R., McGarr, A., Di Toro, G., and Kanamori, H. (2006). Effects of methods of attenuation correction on source parameter determination. *Geophysical Monograph-American Geophysical Union*, 170:91. doi:10.1029/170GM11.

Spottiswoode, S. (1993). Seismic attenuation in deep-level mines, In: Proceedings of the 3rd International Symposium on Rockburstst and Seismicity in Mines (ed. Young, R. P.), Kingston, Canada. *Rockbursts and Seismicity in Mines*, pages 409–414.

Stanchits, S., Vinciguerra, S., and Dresen, G. (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics*, 163(5-6):975–994. doi:10.1007/s00024-006-0059-5.

Stein, S. and Wysession, M. (2003). *An introduction to seismology*. Blackwell Publishing, Oxford.

Stork, A. and Ito, H. (2004). Source parameter scaling for small earthquakes observed at the western Nagano 800-m-deep borehole, central Japan. *Bulletin of the Seismological Society of America*, 94(5):1781–1794. doi:10.1785/012002214.

Su, F., Anderson, J. G., Brune, J. N., and Zeng, Y. (1996). A comparison of direct S-wave and coda-wave site amplification determined from aftershocks of the Little Skull Mountain earthquake. *Bulletin of the Seismological Society of America*, 86(4):1006–1018.

Thompson, B., Young, R., and Lockner, D. A. (2009). Premonitory acoustic emissions and stick-slip in natural and smooth-faulted Westerly granite. *Journal of Geophysical Research: Solid Earth*, 114(B2). doi:10.1029/2008JB005753.

Tullis, J. and Yund, R. A. (1977). Experimental deformation of dry Westerly granite. *Journal of Geophysical Research*, 82(36):5705–5718. doi:10.1029/JB082i036p05705.

Udías, A. and Arroyo, A. L. (2009). The Lisbon earthquake of 1755 in Spanish contemporary authors. In *The 1755 Lisbon Earthquake: Revisited*, pages 7–24. Springer.

Vengosh, A., Jackson, R. B., Warner, N., Darrah, T. H., and Kondash, A. (2014). A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environmental Science & Technology*, 48(15):8334–8348. doi:10.1021/es405118y.

Villeneuve, M. C., Heap, M. J., Kushnir, A. R., Qin, T., Baud, P., Zhou, G., and Xu, T. (2018). Estimating in situ rock mass strength and elastic modulus of granite from the Soultz-sous-Forêts geothermal reservoir (France). *Geothermal Energy*, 6(1):1–29. doi:10.1186/s40517-018-0096-1.

Waldhauser, F. and Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, 90(6):1353–1368. doi:10.1785/0120000006.

Walter, W. R., Mayeda, K., Gok, R., and Hofstetter, A. (2006). The scaling of seismic energy with moment: Simple models compared with observations. *American Geophysical Union Monograph*, 170(UCRL-JRNL-218096):25–41.

Wang, H., Ren, Y., Wen, R., and Xu, P. (2019). Breakdown of earthquake self-similar scaling and source rupture directivity in the 2016–2017 central Italy seismic sequence. *Journal of Geophysical Research: Solid Earth*, 124(4):3898–3917. doi:10.1029/2018JB016543.

Wegener, A. (1966). *The origin of continents and oceans*. Dover Publications, Inc., New York.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9. doi:10.1038/sdata.2016.18.

Wu, R. and Aki, K. (1985). Elastic wave scattering by a random medium and the small-scale inhomogeneities in the lithosphere. *Journal of Geophysical Research: Solid Earth*, 90(B12):10261–10273. doi:10.1029/JB090iB12p10261.

Yasuhara, H., Kinoshita, N., Ohfuji, H., Lee, D. S., Nakashima, S., and Kishida, K. (2011). Temporal alteration of fracture permeability in granite under hydrothermal conditions and its interpretation by coupled chemo-mechanical model. *Applied Geochemistry*, 26(12):2074–2088. doi:10.1016/j.apgeochem.2011.07.005.

Yoo, S.-H., Rhie, J., Choi, H., and Mayeda, K. (2011). Coda-derived source parameters of earthquakes and their scaling relationships in the Korean Peninsula. *Bulletin of the Seismological Society of America*, 101(5):2388–2398. doi:10.1785/0120100318.

Yoshimitsu, N., Kawakata, H., and Takahashi, N. (2014). Magnitude- 7 level earthquakes: A new lower limit of self-similarity in seismic scaling relationships. *Geophysical Research Letters*, 41(13):4495–4502. doi:10.1002/2014GL060306.

Zang, A., Christian Wagner, F., Stanchits, S., Dresen, G., Andresen, R., and Haidekker, M. A. (1998). Source analysis of acoustic emissions in Aue granite cores under symmetric and asymmetric compressive loads. *Geophysical Journal International*, 135(3):1113–1130. doi:10.1046/j.1365-246X.1998.00706.x.

Zoback, M. L. (2006). The 1906 earthquake and a century of progress in understanding earthquakes and their hazards. *GSA TODAY*, 16(4/5):4. doi:10.1130/GSAT01604.1.

List of Figures

List of Tables