UNIVERSITY OF CALIFORNIA, SAN DIEGO

Source parameter estimation and related uncertainties of small earthquakes in southern California

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

in

Earth Sciences

by

Deborah Lynn Kane

Committee in charge:

Frank Vernon, Chair
Joel Conte
Yuri Fialko
Jeff Gee
Peter Shearer

2011

The Dissertation of Deborah Lynn Kane is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2011

To my parents, who always encouraged
and supported my interests in science.

Table of Contents

List of Figures

Acknowledgements

My time in graduate school was greatly influenced by a number of people. Their collective support and encouragement helped make this dissertation possible, and often added to the fun along the way.

Frank Vernon has been a supportive advisor throughout my graduate school career while providing me with ample independence to explore and figure things out on my own. In addition to providing feedback on and advice on my research, I've also had opportunities to attend a wide range of professional meetings, write a few proposals, and become involved with a large collaborative group studying the San Jacinto Fault Zone. Some of these experiences do not always fall within the graduate education norm, and I appreciate the opportunities I've had to participate and get a feel for research careers beyond graduate school. I am grateful that he was frequently available for discussion, whether in person or via long-distance forms of communication.

Peter Shearer helped expand my interest in rupture directivity, and I have enjoyed working with him to look for evidence of this at Parkfield. His feedback on my research over the past several years has been greatly helpful and his proofreading skills as a coauthor are unrivaled. Knowing that he was just a building away for discussion when needed was often reassuring.

I am fortunate to have had Yuri Fialko, Jeff Gee, and Joel Conte serve on my committee. Their collective knowledge of geophysics, geosciences, and structural engineering helped provide interesting points of discussion along the way.

Debi Kilb has been an invaluable colleague and friend since we first met at the 2004 AGU meeting. I never would have predicted at the time that running into someone with such a similar name would so greatly influence my life over the next several years! I've enjoyed our research collaborations and expect we'll continue finding interesting projects to work on together. She has always been available for discussing research, sending drafts back and forth with sometimes entertaining edits, and chatting about life in general.

My research interests as an undergraduate student were already focused on seismology, and I thank Larry Hutchings and Guy Masters for their guidance as undergraduate advisors. I also thank my coauthors and former IGPP graduate students German Prieto and Bettina Allmann for their contributions to our research and for productive and social discussions at various conferences.

The ANZA group provided massive amounts of data for me to explore, vital Antelope support and advice, and the opportunity to see what collecting and processing all of this data really involves. I'm particularly grateful that I've avoided spending too much time picking! Witnessing the media response to large earthquakes and the sorts of questions that are asked by the public has also been an entertaining aspect of being part of the group.

I underwent my first experiences as a teaching assistant under the guidance of Gabi Laske, and I thank her for the opportunity to lead some fun beach walks and get a view of the classroom from a different perspective.

I greatly enjoyed my time as a Scripps Classroom Connection fellow, and I thank Cheryl Peach and Hubert Staudigel for creating the program at SIO. I had a lot of fun working with Dave Van Dusen, and he frequently reminded me quite a bit of my own high school physics teacher.

My colleagues and classmates at IGPP contributed a lot of fun to the whole process. I'm particularly grateful to the entering classes of 2004, 2005, and 2006 for being there all along the way. I've made some great friends and look forward to meeting up with everyone at future conferences! Friends at UCSD and beyond, especially the summer weekly BBQ crowd, provided fantastic distraction from my research.

My parents have always encouraged my interest in science, and were there for the first earthquake I can remember (Loma Prieta, 1989). Their love and support has been central to many major academic decisions made along the way. My sister provided entertaining brief interruptions throughout many work days.

Finally, I thank Alex James for his love and support throughout this whole process. Having another scientist in the house kept me motivated and provided all kinds of random science-related topics to talk about. We've had a lot of fun over the past years.

The chapters presented in this dissertation are reformatted versions of work either already published or being prepared for publication. Chapter 2, in full, is a reformatted version of the material as it appears in the Bulletin of the Seismological Society of America (Kane, D.L., G.A. Prieto, F.L. Vernon, and P.M. Shearer (2011), Quantifying

earthquake source property uncertainties, Bull. Seism. Soc. Am., v. 101; no. 2; p. 535-543; DOI: 10.1785/0120100166) and is reprinted with permission. I was the primary investigator and author of this paper. Chapter 3, in part, is currently being prepared for submission to Seismological Research Letters. Chapter 4, in full, is currently being prepared for submission to the Journal of Geophysical Research. Chapter 5, in full, is currently being prepared for submission to the Bulletin of the Seismological Society of America. In all cases I was the primary author and investigator.

VITA

2005	Bachelor of Science, University of California, San Diego
2008	Teaching Assistant, Scripps Institution of Oceanography University of California, San Diego
2009 - 2010	Scripps Classroom Connection Fellow Scripps Institution of Oceanography University of California, San Diego
2010	Teaching Assistant, Scripps Institution of Oceanography University of California, San Diego
2005 – 2011	Research Assistant, Scripps Institution of Oceanography University of California, San Diego
2011	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

Kane, D.L., G.A. Prieto, F.L. Vernon, and P.M. Shearer (2011), Quantifying earthquake source property uncertainties, Bull. Seism. Soc. Am., v. 101; no. 2; p. 535-543; DOI: 10.1785/0120100166.

Kane, D.L., D. Kilb, A.S. Berg, and V.G. Martynov (2007), Quantifying the remote triggering capabilities of large earthquakes using data from the ANZA seismic network catalog (southern California), J. Geophys. Res., 112, B11302, doi:10.1029/2006JB004714.

ABSTRACT OF THE DISSERTATION

Source parameter estimation and related uncertainties

of small earthquakes in southern California

by

Deborah Lynn Kane

Doctor of Philosophy in Earth Sciences

University of California, San Diego, 2011

Frank Vernon, Chair

Understanding the physics of earthquake rupture is critical to providing accurate estimates of seismic hazard and for effectively mitigating these hazards. Matching physical models to seismic data in order to better understand the earthquake rupture processes requires accurate and precise estimates of earthquake source properties. Measuring source properties such as rupture size and stress drop must include accounting for the effects of seismic wave propagation and making proper assumptions about the rupture process. This thesis focuses on the methods of estimating source properties of

small earthquakes and on the application of these methods to earthquakes in two distinctly different seismogenic regions of California.

In the San Jacinto Fault Zone, earthquakes recorded by a small aperture array allow quantification of source parameter uncertainties using empirical Green's functions and frequency-domain techniques. These uncertainties are frequently overlooked in source parameter estimation, and this study constrains them to ~30% of estimate values. A non-parametric time-domain method using a set of empirical Green's functions is described and applied to a series of example earthquakes. This approach minimizes the assumptions regarding the rupture process and can be used to study less simple ruptures. Correcting for the effects of seismic wave propagation is an important aspect of techniques used in source parameter estimation, and the conditions necessary to effectively use nearby earthquakes as path corrections are tested and quantified.

At the San Andreas Fault near Parkfield, the high degree of waveform similarity among closely spaced earthquakes is used to apply spatially averaged propagation path corrections and search for rupture directivity effects. This analysis shows that the population of small earthquakes in this region does not have a consistent unilateral rupture direction, but 70% of M>3 earthquakes exhibit characteristics of southeast-directed rupture. Computational models featuring a fault interface separating two materials for Parkfield-like conditions agree with the preferential southeast-directed rupture and present potential implications for earthquake source physics.

Combined, these studies of earthquake source parameter estimation can be used to improve future source parameter estimates, offer appropriate metrics for establishing uncertainty bounds, and contribute to the study of earthquake source physics.

Chapter 1: Introduction

Earthquakes have devastated communities throughout human history. The past decade has witnessed a seemingly high number of great, massively destructive events both in regions of relative wealth and in regions of poverty. These events have served to remind those living in areas of relatively high seismic hazard of the potential dangers they face, and public interest in seismology has increased considerably. In southern California, where several faults are capable of generating sizeable and destructive earthquakes, the public is fairly knowledgeable of the basics of earthquake science. However, there are frequent misconceptions of the true hazards present and what seismology can explain.

Understanding the physics underlying earthquake occurrence is key to ultimately achieving a better knowledge of the hazard potential of earthquakes at a given location and time. Much progress has been made in this area in the past several decades, but much also remains to be understood. We know that large earthquakes occur on pre-existing faults in Earth's lithosphere, but sometimes earthquakes occur on previously unknown faults. We know that large earthquakes are much less common than smaller earthquakes, and that the occurrence of a smaller event does little to reduce the stress in the lithosphere compared to a single large earthquake. We also know that there are limits to the depths of the seismogenic zone along faults accommodating strike-slip motion and that earthquake magnitude is linearly related to the area ruptured, but hazard forecasts are not always correct in estimating the magnitude of a potential earthquake.

In this study, I focus on estimating properties of smaller earthquakes (M 2 to M 5) in two regions of southern California. I apply common methods of analyzing earthquake source parameters to complex datasets along with more unique approaches to source parameter studies with more simple datasets. I explore how seismic source parameters are estimated, what sort of uncertainties these estimates entail, how such methods of estimation might be improved, and what we can ultimately learn about earthquake source physics by studying two very different regions of seismicity.

1.1 Earthquake source physics

Earthquakes are generated when the stresses due to motion of tectonic plates exceed the ability of Earth's lithosphere to resist this motion. When an earthquake occurs, the two sides of the fault slide past each other to allow stress to relax from the initial stress to the final stress. The difference of the initial stress and the final stress is referred to as the coseismic static stress drop (*e.g.*, Madariaga, 1976). This stress is not released instantaneously.

Simple models of stress release with slip may call for stress to decrease linearly as the fault slips, while more complex models incorporate features such as non-linear dependence of stress on slip and the possibility of the stress descending below the final stress before reaching the final stress. The process during which stress is released converts potential energy into three main components: fracture energy (released in the process of breaking the rock), frictional energy (released as heat during sliding), and radiated energy (released as seismic waves propagating away from the earthquake

source). Earthquake source physics seeks to constrain this breakdown of the energy budget in order to determine the relative sizes of each, how each scales with earthquake magnitude, how the proportional contributions change under different conditions, and how this budget is affected under the assumptions of various slip models (*e.g.*, Kanamori and Heaton, 2000; Abercrombie and Rice, 2005).

Of these three contributions to the total energy released during earthquake rupture, we are only capable of directly measuring radiated energy. This energy arrives at Earth's surface as seismic waves, and can thus be quantified through the study of recorded ground motion. Because of incomplete data coverage, however, quantifying the true radiated energy as well as other measureable parameters from recorded ground motion remains difficult.

We are also capable of indirectly measuring the coseismic static stress drop (hereafter referred to as 'stress drop') using recorded ground motion by assuming an analytical model relating stress drop to measureable source parameters. An area of frequent debate has focused on determining how the stress drop (and radiated energy) varies as a function of earthquake magnitude. A common simple model of earthquake rupture is self-similar with linear slip-weakening friction: as the radius of the earthquake rupture area increases linearly in size, the stress drop remains constant and the radiated energy increases linearly in response (Aki, 1967). Departures from self-similarity have been suggested by several studies and would have considerable implications for estimates of seismic hazard. Several studies have found stress drop to be constant over many orders of earthquake magnitude, while others have observed non-constant behavior of stress drop (*e.g.*, Abercrombie, 1995; Mayeda and Walter, 1996; Ide and Beroza, 2001; Ide *et*

al., 2003; Mori *et al.*, 2003; Prieto *et al.*, 2004; Abercrombie and Rice, 2005; Imanishi and Ellsworth, 2006; Allmann and Shearer, 2009). A major complication of these studies lies in the common tactic of combining multiple sets of results obtained from different research groups studying different regions with different methods. Comparing such studies requires much care.

1.2 Seismic source parameters

As our ability to understand the conversion of potential energy into fracture, frictional, and radiated energy is limited to observations of radiated energy and stress drop, seismologists have focused on quantifying these properties by relating ground motion to analytical rupture models. These models are used to infer properties of the earthquake source by quantifying recorded ground motion.

A simple model used to consider small earthquakes is the Haskell (1964) source model. The Haskell model is a dislocation source, in which fault slip rate increases linearly along a ramp function over a duration τ_1, continues slipping at a constant rate for a duration of τ_2, and decreases slip rate linearly over a duration τ_3 (Figure 1.1). This model produces a near-field displacement step in which the displacement recorded close to the epicenter exhibits an increase from the initial position to a final position that takes place over duration τ_1 (Figure 1.1). The final displacement is permanent. In the far-field, at distances much further from the epicenter, the recorded displacement exhibits a pulse of duration τ_1 and the final offset in displacement is zero. The amplitude of this simple pulse is related to near-field displacement and duration.

Recorded ground motion can be fit to these simple models to offer insight into the properties (*e.g.*, dimension, duration, or slip) of the earthquake source.

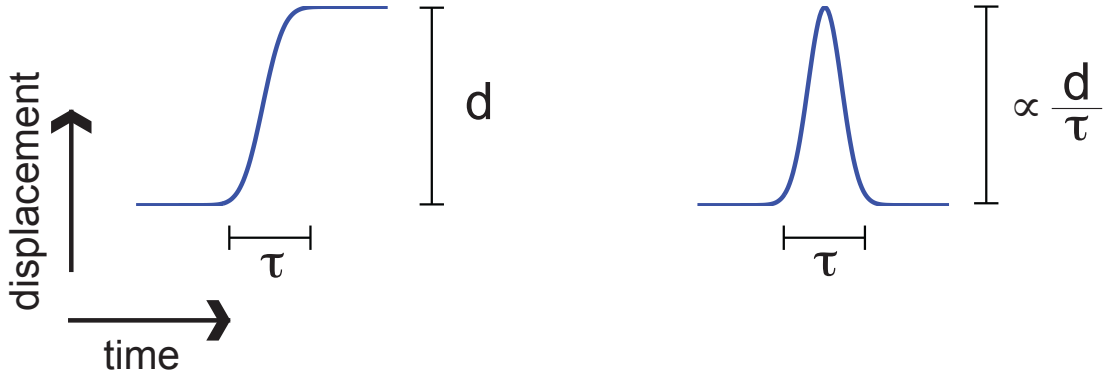

Figure 1.1: Near-field (left) and far-field (right) displacements with time. The displacement in the near-field, d, is a permanent offset and develops over a duration τ. In the far-field, this displacement occurs temporarily over the same duration.

1.2.1 Correcting for the complications of seismic wave propagation using empirical Green's functions

In order to effectively study earthquake sources, it is necessary to first consider how the propagation of seismic waves through Earth affects ground motion recorded at a seismic station. This is sometimes addressed in seismology by measuring the attenuation of seismic waves as they travel through structural variations in Earth over a given distance. In the cases considered in this study, the distances between the earthquakes and the seismic stations are relatively small and the waves propagate through highly varying structural features in Earth's lithosphere.

In such situations, the complications of seismic wave propagation can be treated through the use of empirical Green's functions (EGFs). This technique was first introduced by Hartzell (1978) and is frequently used in seismic studies today (*e.g.*, Frankel and Kanamori, 1983; Hutchings and Wu, 1990; Hough, 1997; Prieto *et al.*, 2004). The EGF technique can be used when an earthquake of interest also has a nearby, smaller earthquake recorded by the same seismic station. This method makes several underlying assumptions including the approximation of the smaller earthquake as a point-source relative to the larger earthquake and identical hypocentral locations of the two earthquakes (hence the assumption of identical propagation paths of the seismic waves to a given station).

The ground motion recorded for the larger earthquake at a given seismic station can be expressed as a convolution of several signals:

$$m_1(t) = s_1(t) * p(t) * i(t) \qquad < 1.1 >$$

Here, $m_1(t)$ represents the ground motion recorded for an earthquake with source signal $s_1(t)$, propagation path $p(t)$, and near-surface site effects and instrument response $i(t)$. The second, smaller earthquake can be represented similarly as:

$$m_2(t) = s_2(t) * p(t) * i(t) \qquad < 1.2 >$$

Because the two earthquakes share an assumed common hypocentral location, the signals due to the propagation path between the hypocenters and a specified recording station, $p(t)$, can be assumed to be identical (Figure 1.2). The local site-effects at the recording station and the signal contributions due to instrumentation are also identical. The approximation of the smaller earthquake as a point-source means that $s_2(t)$ is not a function of time, but rather a delta function of amplitude s_2. Subsequently, equation < 1.2 > can be rewritten as:

$$\frac{m_2(t)}{s_2} = p(t) * i(t) \qquad < 1.3 >$$

And this $p(t) * i(t)$ equivalence can be substituted into equation < 1.1 >, giving:

$$m_1(t) = s_1(t) * \frac{m_2(t)}{s_2} \qquad < 1.4 >$$

Rearranging,

$$m_1(t) = \frac{s_1(t)}{s_2} * m_2(t) \qquad < 1.5 >$$

$$m_1(t) = s(t) * g(t) \qquad < 1.6 >$$

Here, the relative source time function, $s(t)$, is redefined as $s_1(t)/s_2$. The recording of the smaller earthquake, $m_2(t)$, is now represented by $g(t)$ and called the empirical Green's function.

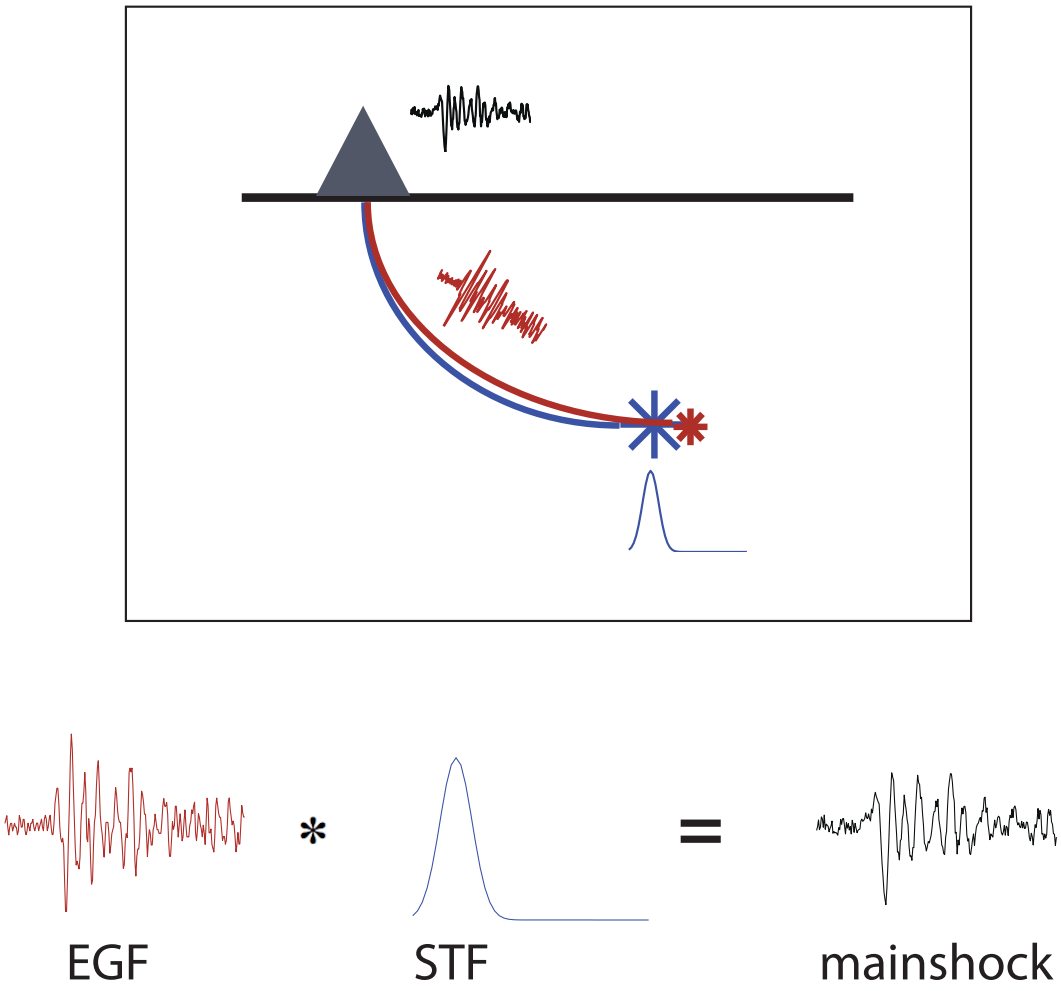

Figure 1.2: Cartoon depicting the EGF technique. In the EGF technique, two earthquakes of different sizes (shown with asterisks) occur very close in space to each other and are recorded by the same seismic station (triangle). If the source mechanisms are similar, then the ground motion recorded for the smaller earthquake (red waveform) can be used to correct for the propagation path between the mainshock source and the station. The mainshock source time function (STF) can be determined by inverting the equation.

Thus the EGF technique provides a means of removing the complications of seismic wave propagation through a medium of unknown structural variations while minimizing the number of assumptions required.

1.2.2 Estimation in the frequency domain

Seismic sources are frequently quantified by considering ground motion in the frequency domain and using expected frequency characteristics of the Haskell (and other) models in combination with EGF techniques. The most widely used of these models is the Brune source spectrum (Brune, 1970):

$$S(f) = \frac{M_0}{1 + \left(f / f_c \right)^2} \qquad < 1.7 >$$

The Brune source spectrum model (Figure 1.3) gives the analytical displacement amplitude spectrum as a function of frequency, f, for a given seismic moment (M_0) and corner frequency (f_c). The corner frequency of the source spectrum can be related to a rupture radius in the case of a circular shear rupture. Although it is a simple rupture model, many studies have demonstrated that it fits seismic data remarkably well for small earthquakes even though true events are unlikely to occur as simply as circular, shear ruptures.

In the frequency domain, the convolution of the EGF with the source term (equation 1.6) can be expressed as a simple multiplication of their spectra. Thus, it is a relatively simple process to obtain an earthquake source spectrum free from complications of propagation path by computing the displacement spectrum at a given

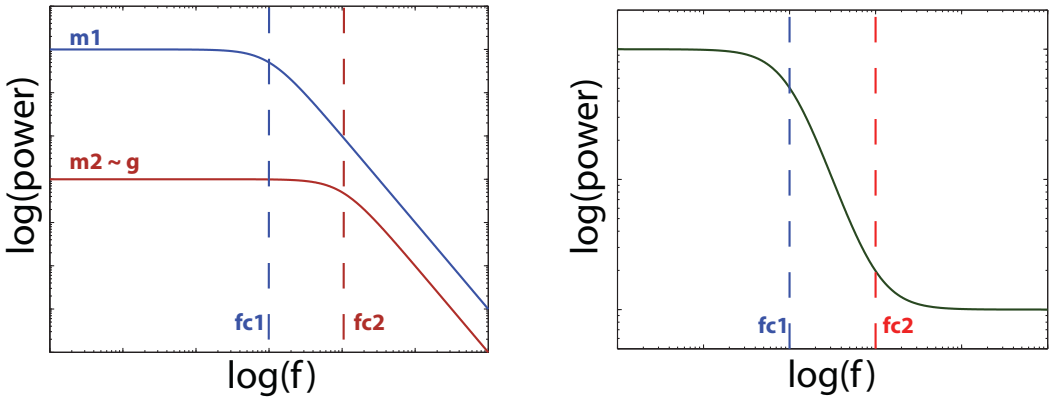

Figure 1.3: Cartoons illustrating use of spectral division with the Brune source spectrum. On the left, the mainshock (*m1*) and the EGF earthquake (*m2*) have spectra with different corner frequencies (*fc1* and *fc2*, respectively) and different amplitudes. By computing the spectral ratio, the Brune source spectrum can be fit to the result to estimate the corner frequencies. In many cases, *fc2* is neglected due to lack of high frequency data with adequate signal to noise ratios and the spectrum is assumed to continue decaying rather than leveling off at higher frequencies.

station for the larger earthquake (the 'mainshock', $M_1(f)$) and for the smaller EGF earthquake, $G(f)$, and calculating the ratio of these two spectra to obtain the source spectrum, $S(f)$:

$$S(f) = \frac{M_1(f)}{G(f)} \qquad < 1.8 >$$

The spectrum defined by the Brune model (equation 1.7) can then be fit to this spectral ratio to obtain estimates of the model parameters. This method can also be used with velocity or acceleration spectra as long as the mainshock and EGF spectra are of the same type.

1.2.3 Estimation in the time domain

In the time domain, the seismic source is represented by the source time function (STF). The STF for the Brune model is a simple pulse of duration τ. Past studies of small earthquakes have occasionally considered STFs but this method is not as common as the frequency domain method described above. Such studies have often computed the STF using the frequency domain by estimating the source spectrum, $S(f)$, as described above, and then transforming this spectrum back into the time domain. As discussed in Chapter 3, this introduces some complications. We address these complications and present a simple alternative method for estimating STFs along with several examples of using this method to obtain consistent results.

Similarly to the source spectrum in the frequency domain, the resulting STF in the time domain can be used to estimate source parameters by measuring the duration of the

STF and the area under the STF. These can be directly related to the duration of rupture and the seismic moment released.

1.2.4 Inferring static stress drop

Interest in the source scaling problem in the past decade has resulted in a number of studies addressing the process of estimating the coseismic static stress drop. The static stress drop can be directly related to the radius of the earthquake rupture (*e.g.*, Kanamori and Anderson, 1975; Madariaga, 1976):

$$\Delta\sigma = \frac{7}{16}\frac{M_0}{r^3} \qquad < 1.9 >$$

Here, r refers to the radius of the circular shear rupture assumed. Dynamic rupture simulations assuming a rupture velocity equivalent to 90% of the S-wave velocity, β, provide a relationship between radius and corner frequency (Madariaga, 1976):

$$r = 0.32\frac{\beta}{f_c} \qquad < 1.10 >$$

Methods described by Boatwright (1980) and Mori *et al.* (2003) can be used to determine the radius in the time domain; here the sine of the takeoff angle, $\theta_{takeoff}$, can be assumed to be an average of 0.64 and α represents the P-wave velocity:

$$r = \frac{1}{2}\tau\frac{v_r}{1-\frac{v_r}{\alpha}\sin\theta_{takeoff}} \qquad < 1.11 >$$

These formulae for rupture radius provide a means of estimating the stress drop in each of the two approaches discussed above.

1.2.5 Inferring rupture directivity

The final source parameter to consider in this study is rupture directivity. In our previous discussion of source parameters, the models used have always assumed a circular rupture initiating at the center and rupturing outward radially at constant speed. However, as we consider particularly in Chapter 4, such symmetric rupture propagation is rarely observed in large earthquakes and asymmetric rupture propagation may be an overlooked feature of smaller earthquakes. If an earthquake ruptures from one side of a rupture patch to the other, rather than from the center outwards, then the resulting STF pulse and Brune spectrum will vary with azimuth (Figure 1.4). These azimuthal variations lead to variations in both the pulse duration, τ, and the corner frequency, $f_{c,app}$, and can be respectively expressed as a function of azimuth (θ), rupture velocity, and seismic wave speed (c):

$$\tau(\theta) = \frac{L}{v_r} - \frac{L\cos\theta}{c} = \tau_0(1 - \frac{v_r}{c}\cos\theta) \qquad < 1.12 >$$

$$f_{c,app}(\theta) = f_c \frac{1}{1 - \frac{v_r}{c}\cos\theta} \qquad < 1.13 >$$

This situation violates the assumptions of the Brune source model in which a radially outward rupture of constant rupture velocity is generally used.

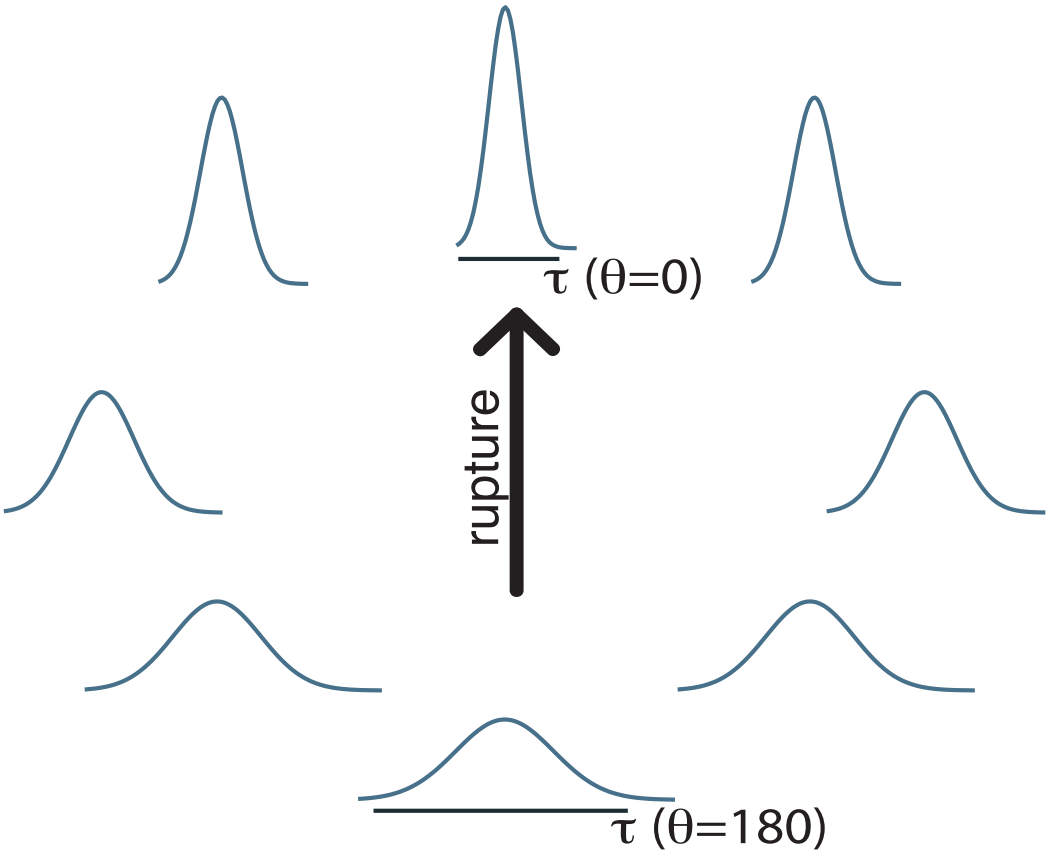

Figure 1.4: Cartoon illustrating the effects on source time function amplitude and duration as a function of azimuth when strong unilateral rupture directivity is present. In the direction of rupture, the source time function (STF) is greatly amplified and is much shorter in duration.

1.3 Interpreting source parameters and considering uncertainties

Source parameters of earthquakes over a range of sizes are routinely estimated in studies using various methods, but combining the results of multiple studies to characterize the process of earthquake rupture has proven to be a complicated task (*e.g.*, Ide and Beroza, 2001; Allmann and Shearer, 2009). Differences in processing techniques can affect the stability of results, and the possibility of different datasets exhibiting different trends remains unresolved. An objective of this study is to quantify the uncertainties using these methods. This information is necessary for simultaneously considering the results of events over orders of magnitudes from different datasets processed using different assumptions and techniques.

1.4 Study areas

This study features data from two seismogenic regions in California. The first region considered is the San Jacinto Fault Zone (SJFZ) near the Anza seismic gap. The second region considered is the extremely well-studied Parkfield section of the San Andreas Fault.

1.4.1 San Jacinto Fault Zone

The SJFZ is one of several right-lateral strike-slip fault zones in southern California accommodating plate boundary motion between the Pacific and North American plates. The one to two cm/year of total slip rate for the SJFZ system makes it one of the most active faults in southern California (Sharp, 1967; Fialko, 2006). Cumulative slip along the SJFZ is estimated to be 24 to 29 km over 2 million years (Sharp, 1967). The SCEC Phase II report identifies a magnitude 7.5 characteristic earthquake on the Anza section; paleoseismic data indicates that the area's last significant earthquake occurred about 1795 (WGCEP, 2008). The slip rate estimate coupled with the paleoseismic data provide for a current slip deficit of ~3 meters. This area of the fault also exhibits lower seismicity rates than other sections of the SJFZ and is known as the Anza seismic gap (Figure 1.5). The latest Working Group on California Earthquake Probabilities report (2008) forecast a 31% probability of an M > 6.7 earthquake occurring on the San Jacinto fault in the next 30 years.

Several recent M 4 to M 5+ earthquakes have been the largest earthquakes to occur in this area in many years and serve to remind the residents of southern California of the hazard presented by the SJFZ. These earthquakes all occurred to the southeast of the Anza seismic gap. The two M 5+ earthquakes in 2001 and 2005 triggered 7500+ cataloged aftershocks with magnitudes complete to ML 0. The four ML 4-5 earthquakes in 2008 and 2010 occurred to the south of the Trifurcation cluster.

The ANZA seismic network is designed to provide good azimuthal coverage of earthquakes within the SJFZ and has recorded more than 3 million waveforms from

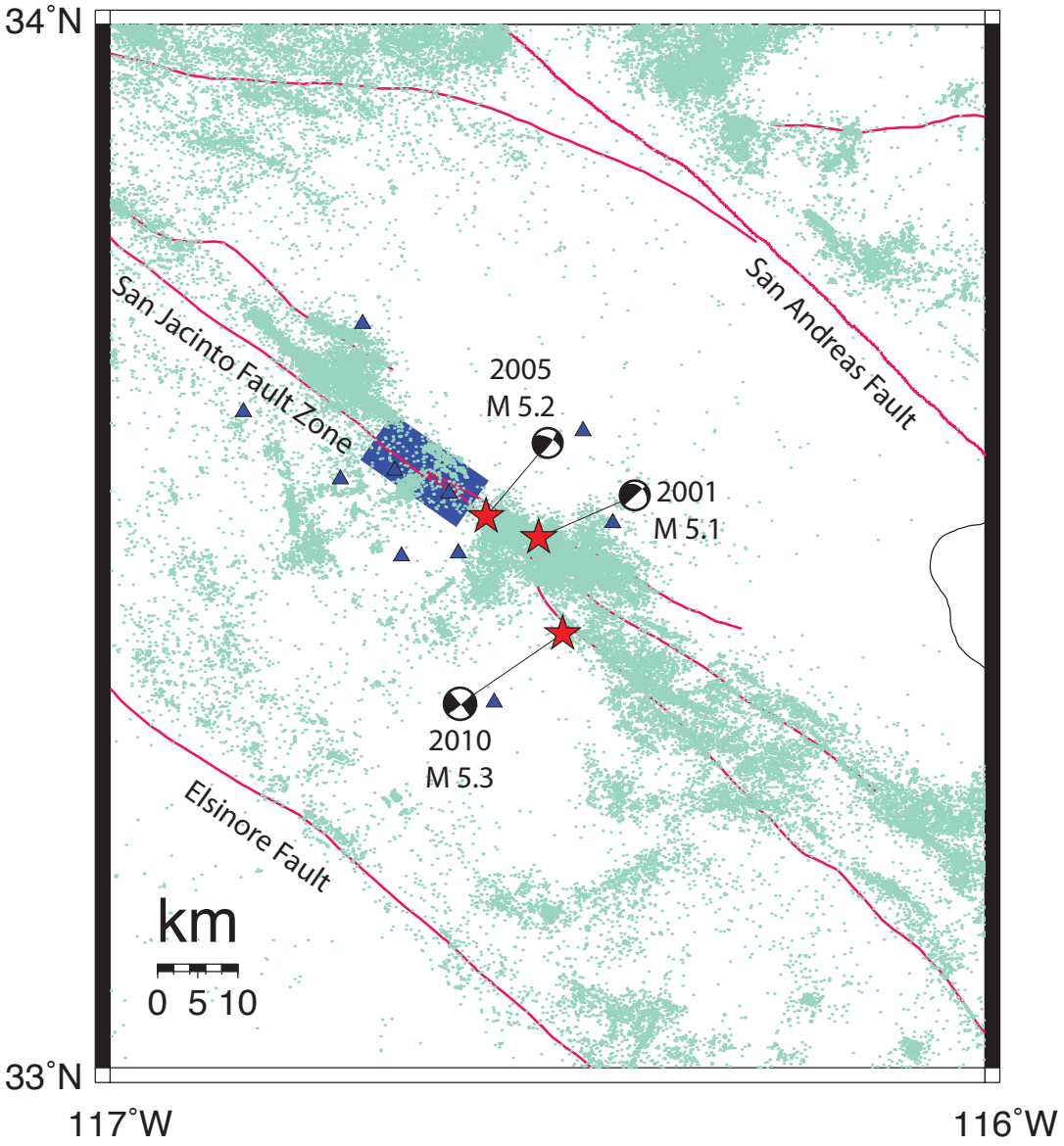

Figure 1.5: Map of the San Jacinto Fault Zone and surrounding area. Green dots represent earthquakes from 1981 to 2005 (locations from Lin *et al.*, 2007) and blue triangles mark the sites of the ANZA stations. The 2001, 2005, and 2010 M > 5 earthquakes are highlighted with stars and the corresponding focal mechanism solutions are shown. The blue shaded region marks the approximate location of the Anza seismic gap.

65000+ local earthquakes since 1982. Previous studies of earthquakes within the SJFZ have suggested that complex fault structure in the region could be responsible for the complicated source property results (Hartse *et al.*, 1994; Kilb *et al.*, 2007). Several parallel- and cross-faults are apparent, with seismicity tending to concentrate in the two clusters on the edges of the Anza seismicity gap (Figure 1.5) identified by Sanders and Kanamori (1984). The Hot Springs cluster to the northwest of the Anza gap lies between the mapped traces of the Hot Springs faults at depths of 15 to 22 km. A more diffuse zone of seismicity is present in the Trifurcation cluster southeast of the gap where three strands of the San Jacinto fault (Buck Ridge, Clark, and Coyote Creek) diverge as the fault zone continues southward. Earthquakes in the Trifurcation cluster generally have depths of 7 to 17 km. Within the strands of the SJFZ, earthquake locations are much less localized than those occurring along more mature faults that have accommodated greater overall slip (Powers and Jordan, 2010). Aftershocks from the 2001, 2005, and 2010 M ~5 earthquakes fall within a volume rather than aligning along a single plane outlining the fault. Research following the 2001 and 2005 earthquakes concluded that both events likely ruptured on the conjugate plane (Walker *et al.*, 2005; Kilb *et al.*, 2005), providing further evidence of heterogeneous faulting mechanisms in the SJFZ.

1.4.2 Parkfield

The San Andreas Fault (SAF) at Parkfield, California, presents a unique opportunity to seek a better understanding of earthquake source physics. This region (Figure 1.6) is defined as a distinct section of the fault by boundaries separating the

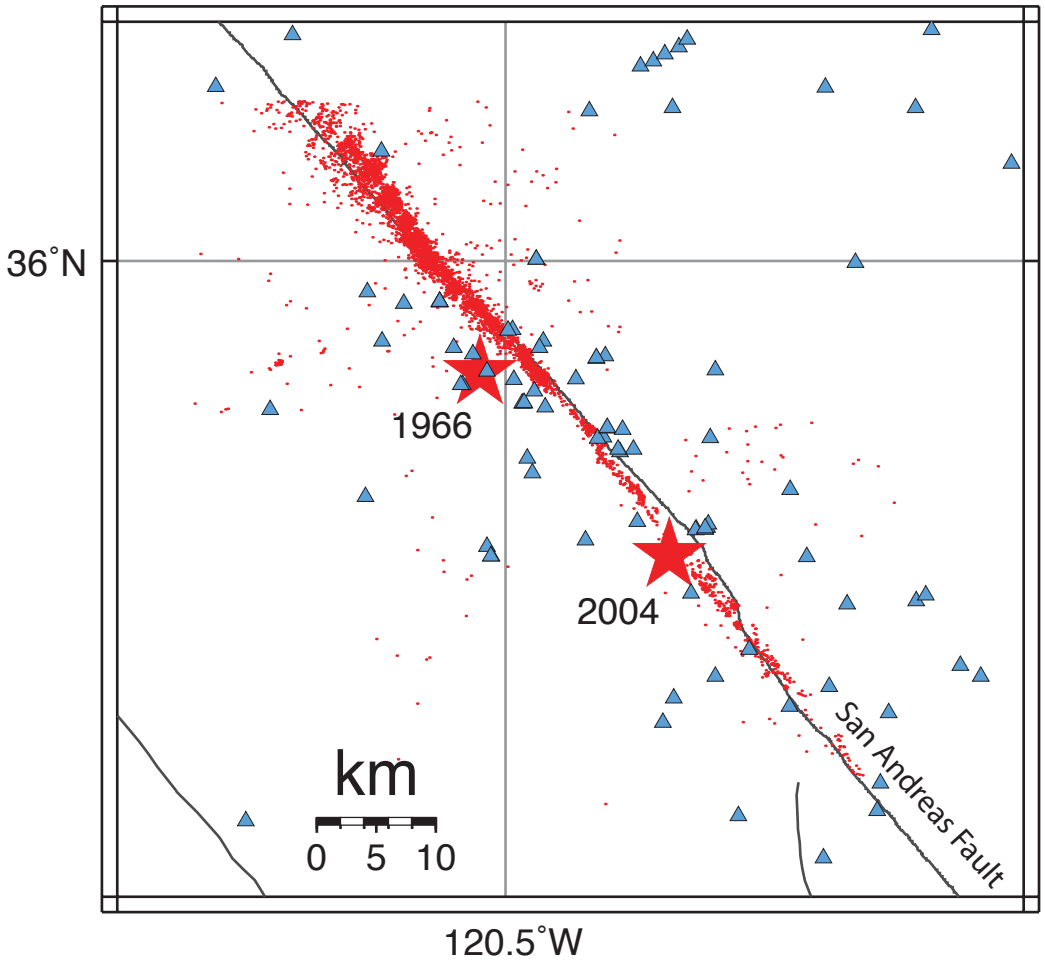

Figure 1.6: Map of the Parkfield section of the San Andreas Fault. Red dots mark seismicity from 1984 to 2005, and triangles show a set of recording stations operating during this time period. The locations of the 1966 and 2004 M 6 earthquakes are highlighted with stars.

locked section towards the southeast from the creeping section towards the northwest (Wallace, 1990). Seismicity primarily occurs on a vertically oriented plane (Eberhart-Phillips and Michael, 1993; Thurber *et al.*, 2006). The recurrence of M ~6 earthquakes at Parkfield has resulted in extensive studies of the seismicity in the region and the fault zone structure, and Parkfield is one of the best instrumented faults on Earth.

Previous studies have concluded that the fault represents a distinct, natural barrier between two blocks of different properties (Eberhart-Phillips and Michael, 1993; Thurber *et al.*, 2006). This juxtaposition of two different blocks is due to the sizeable amount of slip the fault has accommodated. Seismicity at Parkfield is much more spatially confined than seismicity within the SJFZ. Application of various earthquake relocation techniques has produced detailed images of fine fault structure and studies have noted the strong clustering of event hypocenters, including sets of 'repeating' earthquakes that appear to rupture the same area of the fault with each subsequent small earthquake (Waldhauser *et al.*, 2004).

1.5 This study

In this thesis, I explore aspects of source parameter estimation and the related uncertainties as described above.

In Chapter 2, I use frequency domain techniques with data from a small aperture array in the SJFZ to quantify the uncertainties introduced in source parameter estimates by near-surface scattering effects. These uncertainties are what might be expected as minima for a 'good' seismic station. The corresponding uncertainties in stress drop

estimates have implications for comparing source parameters across multiple studies (in which uncertainties are generally neglected).

In Chapter 3, I address source parameter estimation in the time domain. The time domain has not been used as extensively as the frequency domain in studies of small earthquakes. Here, I describe a simple method of estimating source time functions that relies on fewer assumptions than the frequency domain methods. This advantageous method has been used in previous studies, and I provide an extension of this method to work with multiple EGF waveforms simultaneously to stabilize the results and provide a more robust estimate of an STF than might otherwise be obtained. I demonstrate this with several examples from synthetic data and from SJFZ data.

In Chapter 4, I explore the prevalence of unilateral rupture directivity towards the southeast or towards the northwest along the SAF at Parkfield. This is done in the context of a computational model that suggests a preferred rupture direction may be due to the structural properties of the fault (Shi and Ben-Zion, 2006). Whether or not this model is relevant for natural faults has been extensively debated (*e.g.*, Harris and Day, 2005; Ben-Zion, 2006; Harris and Day, 2006). I use frequency domain analysis to consider azimuthal variations in ground motion without directly assuming a source model.

In Chapter 5, I use a large dataset of waveforms from the SJFZ to investigate how the choice of EGF earthquake affects source parameter results and to provide general constraints for EGF event selection in the region and in other regions of relatively heterogeneous distributions of seismicity.

Acknowledgements

I thank Frank Vernon and Luciana Astiz for comments on this chapter.

References

Abercrombie, R. (1995), Earthquake source scaling relationships from -1 to 5 ML using seismograms recorded at 2.5-km depth, *J. Geophys. Res.*, 100 (B12), pp.24015-24036.

Abercrombie, R., and J. Rice (2005), Can observations of earthquake scaling constrain slip weakening?, *Geophys. J. Int.*, 162 (2), pp.406-424.

Aki, K. (1967), Scaling Law of Seismic Spectrum, *J. Geophys. Res.*, Vol. 72, No. 4, pp.1217-1231.

Allmann, B. P., and P. M. Shearer (2009), Global variations of stress drop for moderate to large earthquakes, *J. Geophys. Res.*, 114, B01310, doi:10.1029/2008JB005821.

Ben-Zion, Y. (2006), Comment on "Material contrast does not predict earthquake rupture propagation direction" by R.A. Harris and S.M. Day, *Geophys. Res. Lett.*, Vol. 33, L13310, doi:10.1029/2005GL025652

Boatwright, J. (1980), A spectral theory for circular seismic sources; simple estimates of source dimension, dynamic stress drop, and radiated seismic energy, *Bull. Seism. Soc. Am.,* Vol. 70, No. 1.

Brune, J. (1970)., Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, 75 (26), 4997-5009.

Eberhart-Phillips, D., and A. Michael (1993), Three-Dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California, *J. Geophys. Res.*, Vol. 98, No. B9, pp.15,737-15,758,

Fialko, Y. (2006), Interseismic strain accumulation and the earthquake potential on the southern San Andreas fault system, *Nature*, 441, doi:10.1038/nature04797, pp.968-971.

Frankel, A., and H. Kanamori (1983), Determination of rupture duration and stress drop for earthquakes in southern California, *Bull. Seism. Soc. Am.*, Vol. 73 (6), pp.1527-1551.

Harris, R.A., and S.M. Day (2005), Material contrast does not predict earthquake rupture propagation direction, *Geophys. Res. Lett.*, Vol. 32, L23301, doi:10.1029/2005GL023941

Harris, R.A., and S.M. Day (2006), Reply to comment by Y. Ben-Zion on "Material contrast does not predict earthquake rupture propagation direction", *Geophys. Res. Lett.*, Vol. 33 (13), L13311, doi:10.1029/2006GL026811

Hartse, H.E., M.C. Fehler, R.C. Aster, J.S. Scott, and F.L. Vernon (1994), Small-scale stress heterogeneity in the Anza seismic gap, southern California, *J. Geophys. Res.*, 99(B4), pp.6801-6818.

Hartzell, S. (1978), Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.*, 5(1).

Haskell, N. (1964), Total energy and energy spectral density of elastic wave radiation from propagating faults, *Bull. Seism. Soc. Am.*, Vol. 54, No. 6, pp.1811-1851.

Hough, S.E. (1997), Empirical Green's function analysis: taking the next step, *J. Geophys. Res.,* Vol. 102, No. B3, pp. 5369-5384.

Hutchings, L. and F. Wu (1990), Empirical Green's Functions From Small Earthquakes: A Waveform Study of Locally Recorded Aftershocks of the 1971 San Fernando Earthquake, *J. Geophys. Res.*, Vol. 95, No. B2, pp.1187-1214.

Ide, S., and G. Beroza (2001), Does apparent stress vary with earthquake size?, *Geophys. Res. Lett.*, Vol. 28 (17), pp.3349–3352.

Ide, S., G. Beroza, S. Prejean, and W. Ellsworth (2003), Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. Geophys. Res.,* Vol. 108(B5), doi:10.1029/2001JB001617

Imanishi, K., and W. Ellsworth (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 Abercrombie, McGarr, Kanamori, and di Toro (Editors), *AGU Geophys. Monograph*, 170, 81-90.

Kanamori, H. and D.L. Anderson (1975), Theoretical basis of some empirical relations in seismology, *Bull. Seism. Soc. Am.*, Vol. 65, No. 5, pp.1073-1095.

Kanamori, H. and T.H. Heaton (2000), Microscopic and macroscopic physics of earthquakes, in Geocomplexity and the Physics of Earthquakes, Rundle, Turcotte and Klein (Editors), *AGU Geophys. Monograph*, 120, pp. 147-163.

Kilb, D., D.C. Agnew, B. Allmann, D.L. Kane, G. Lin, G.A. Prieto, P.M. Shearer, F.L. Vernon, and K. Walker (2005), The 2001 and 2005 Anza Aftershock Sequences: Examination of Mainshock Rupture Direction, Post-Seismic Creep, Focal Mechanisms and Spatial/Temporal Behavior of Relocated Aftershocks, *Eos Trans. AGU*, 86(52), Fall Meet. Suppl., Abstract S23B-0245.

Kilb, D., V.G. Martynov, and F.L. Vernon (2007), Aftershock Detection Thresholds as a Function of Time: Results from the ANZA Seismic Network following the 31 October 2001 ML 5.1 Anza, California, Earthquake, *Bull. Seism. Soc. Am.*, 97 (3), pp.780-792, doi:10.1785/0120060116

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

Madariaga, R. (1976), Dynamics of an expanding circular fault, *Bull. Seism. Soc. Am.*, Vol. 66 (3), pp.639-666.

Mayeda, K. and W. Walter (1996), Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes, *J. Geophys. Res.*, 101 (B5), pp.11195-11208.

Mori, J., R. Abercrombie, and H. Kanamori (2003), Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake, *J. Geophys. Res.*, Vol. 108 (B11), 2545, doi:10.1029/2001JB000474

Powers, P.M. and T.H. Jordan (2010), Distribution of seismicity across strike-slip faults in California, *J. Geophys. Res.*, 115, B05305, doi:10.1029/2008JB006234.

Prieto, G., P.M. Shearer, F.L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking P and S spectra, *J. Geophys. Res.*, 109 (B08310).

Sanders, C.O., and H. Kanamori (1984), A seismotectonic analysis of the Anza seismic gap, San Jacinto fault zone, southern California, *J. Geophys. Res.*, 89, 5873—5890.

Sharp, R.V. (1967), San Jacinto fault zone in the Peninsular Ranges of Southern California, *GSA Bulletin*, Vol. 78; No. 6; pp.705-729.

Shi, Z., and Y. Ben-Zion (2006), Dynamic rupture on a bimaterial interface governed by slip-weakening friction, *Geophys. J. Int.*, doi: 10.1111/j.1365-246X.2006.02853.x

Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006), Three-Dimensional Compressional Wavespeed Model, Earthquake Relocations, and Focal Mechanisms for the Parkfield, California, Region, *Bull. Seism. Soc. Am.*, 96 (4B), S38.

Waldhauser, F., W. Ellsworth, D. Schaff, and A. Cole (2004), Streaks, multiplets, and holes: high-resolution spatio-temporal behavior of Parkfield seismicity, *Geophys. Res. Lett.*, Vol. 31, L18608, doi:10.1029/2004GL020649

Walker, K., D. Kilb, and G. Lin (2005), The 2001 and 2005 Anza Earthquakes: Aftershock Focal Mechanisms, SCEC 2005 Meeting.

Wallace, R. (1990). The San Andreas Fault System, California, *U.S. Geol. Surv. Prof. Pap.*, Vol. 1515.

Working Group on California Earthquake Probabilities (2008), The Uniform California Earthquake Rupture Forecast, Version 2 (UCERF 2), USGS Open File Report 2007-1437.

Chapter 2: Quantifying Seismic Source Parameter Uncertainties

We use data from a small aperture array in southern California to quantify variations in source parameter estimates at closely spaced stations (distances ranging from ~7-350 meters) to provide constraints on parameter uncertainties. Many studies do not consider uncertainties in these estimates even though they can be significant and have important implications for studies of earthquake source physics. Here, we estimate seismic source parameters in the frequency domain using empirical Green's function (EGF) methods to remove effects of the travel paths between earthquakes and their recording stations. We examine uncertainties in our estimates by quantifying the resulting distributions over all stations in the array. For coseismic stress drop estimates, we find that minimum uncertainties of ~30% of the estimate can be expected. To test the robustness of our results, we explore variations of the dataset using different groupings of stations, different source regions, and different EGF earthquakes. Although these differences affect our absolute estimates of stress drop, they do not greatly influence the spread in our resulting estimates and these sensitivity tests show that they are not the primary contribution to the uncertainties in our parameter estimates for single stations. We conclude that establishing reliable methods of estimating uncertainties in source parameter estimates (including corner frequencies, source durations, and coseismic static

stress drops) is essential, particularly when the results are used in the comparisons among different studies over a range of earthquake magnitudes and locations.

2.1 Introduction

The scaling of earthquake source parameters with magnitude has been a point of debate in recent years, and is motivated by the implications for understanding earthquake source physics. The issue is whether large earthquakes can be adequately approximated by linearly scaling the attributes of smaller earthquakes, or whether something fundamentally differs in the rupture physics of different size earthquakes. Many studies have looked for evidence (or lack thereof) of self-similar source scaling (*e.g.,* Abercrombie, 1995; Mayeda and Walter, 1996; Ide and Beroza, 2001; Ide, *et al.*, 2003; Mori, *et al.*, 2003; Prieto *et al.*, 2004; Abercrombie and Rice, 2005; Imanishi and Ellsworth, 2006) by using a range of datasets and techniques, but many of these studies lack quantitative estimates of parameter uncertainties. Understanding the similarities and differences of earthquakes over a range of magnitudes is critical for earthquake source physics. Smaller earthquakes occur much more frequently than larger earthquakes, make up a much greater portion of the data collected, and allow a statistical consideration of source properties not achievable with individual large earthquakes. If there is something fundamentally different in the rupture physics of different size events, then our understanding of the hazard presented by large earthquakes will be inadequate. Accurately quantifying differences in earthquake source parameters should include consideration of the uncertainties in each measurement (Abercrombie and Rice, 2005; Prieto *et al.*, 2006).

The heart of the source scaling question currently lies in how researchers estimate seismic source parameters, how these estimates are subsequently combined over varied datasets, and how magnitude scaling is evaluated (Abercrombie and Rice, 2005). Here, we focus on the issue of possible magnitude scaling by considering the often overlooked uncertainties in the source parameter estimates. We approximate the uncertainties by studying the distribution of estimates over an array of closely spaced stations (distances ranging from ~7 to 350 m). Differences in the propagation paths between a given source and all stations are very small because the spacing between the stations is small relative to the distance between the earthquake locations and the array (smallest source-station separation is 6.4 km). Azimuthal variations due to rupture directivity effects can likewise be ignored. The ground motions associated with an earthquake should be similar at each station because of this geometry, but local heterogeneities in near-surface structure could produce somewhat different waveforms across the array. This unique dataset allows us to constrain the uncertainties of our parameters at each station individually and gain some insight as to the predicted variability for other data by analyzing the resulting estimates and their distribution.

2.2 The San Jacinto Fault Zone and the Small Aperture Array (SAA) experiment

We use data from a high frequency array experiment installed in 1990 at Pinyon Flat Observatory (PFO) near the San Jacinto Fault Zone (SJFZ) in southern California (Al-Shukri, *et al.*, 1995; Vernon *et al.*, 1991, Vernon *et al.*, 1998, Wagner, 1998). Pinyon

Flat is located on a pluton at the northern end of the Peninsular Ranges batholith in an area of nearly flat topography. Instruments were installed at the base of the uppermost weathered granodiorite layer, on top of the slightly more solid bedrock (Vernon *et al.*, 1998, and references therein). This site was chosen as a location to test coherence of seismic waves over an array. The benefits of the site include easy accessibility, a hard rock region with relatively uniform geology, and minimal local topographic variation. Previous studies suggest that the weathered surface layer is heterogeneous with variable depth, and that it acts as a waveguide for incident P-waves (Al-Shukri *et al.*, 1995; Vernon *et al.*, 1998; Wagner, 1998). Wilson and Pavlis (2000) additionally showed that the variation in site responses across the array occurs on small distance scales of approximately the size expected for weathered granitic rocks.

With the aim of studying coherence on a scale not previously attempted, the SAA experiment (Owens *et al.*, 1991) was installed for one month. The array consisted of 58 surface stations and two borehole stations. Thirty-six of the surface stations were placed in a square grid with ~7 m spacing between adjacent stations. The remaining twenty-two surface stations extended in arms away from the square grid to the south and east using ~21 m spacing between stations. The two borehole stations were placed near the center of the square grid at 150m and 275m depths in separate boreholes (Figure 2.1). Three-component L22-D geophones (2-Hz response) and 6-channel PASSCAL REFTEK RT72A-02 dataloggers were used at each of the SAA stations. The network was event-triggered at the deepest borehole station, and each station recorded ground velocity at 250 samples per second. These closely located stations recorded ground motion with visually similar waveforms across the array (Figure 2.2).

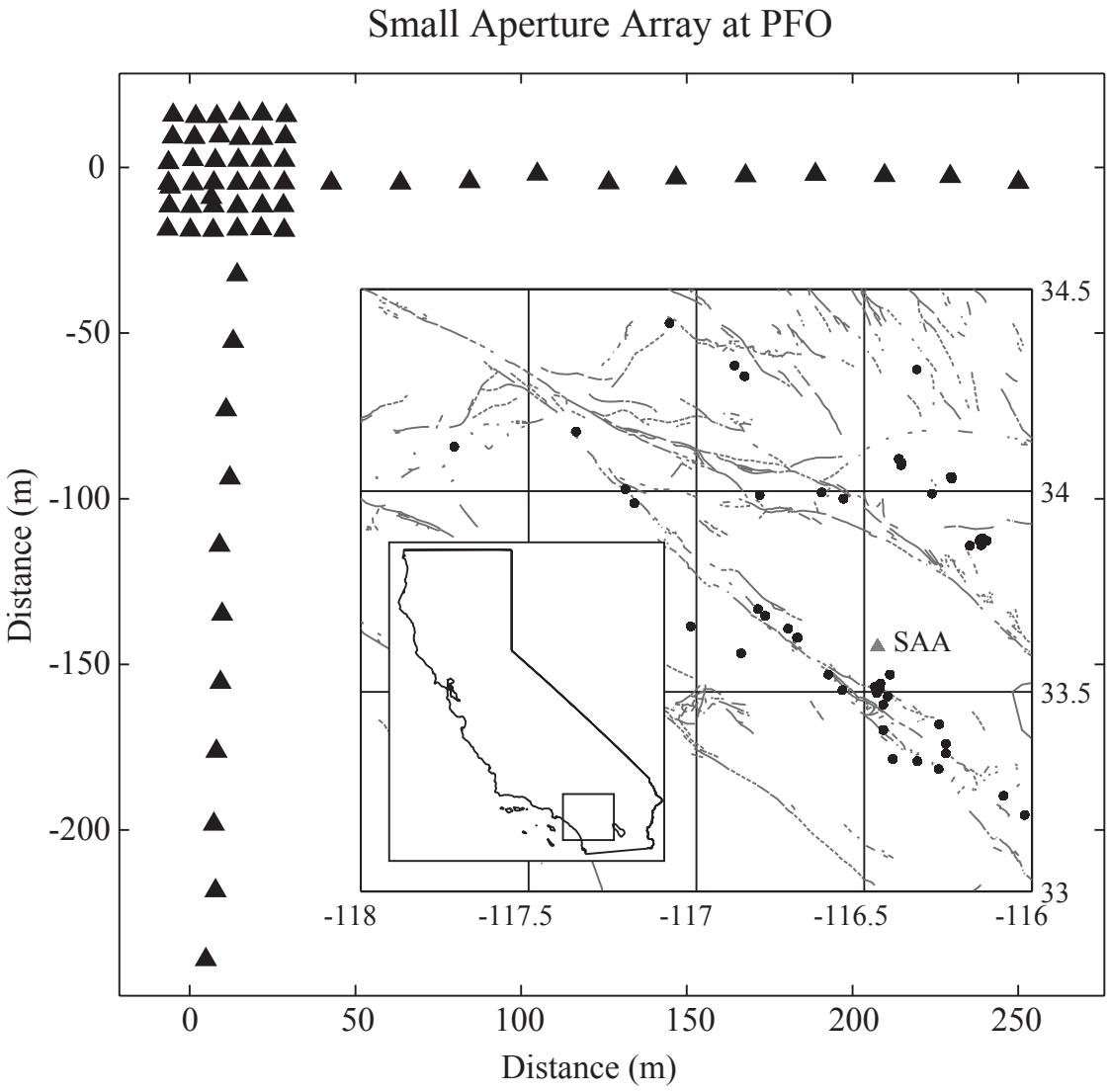

Figure 2.1: Map showing the Small Aperture Array (SAA) station distribution (triangles), study area, and relocated seismicity (black circles). The SAA was installed along the SJFZ; mapped fault traces are shown as gray lines. Several small clusters of earthquakes are apparent in the map.

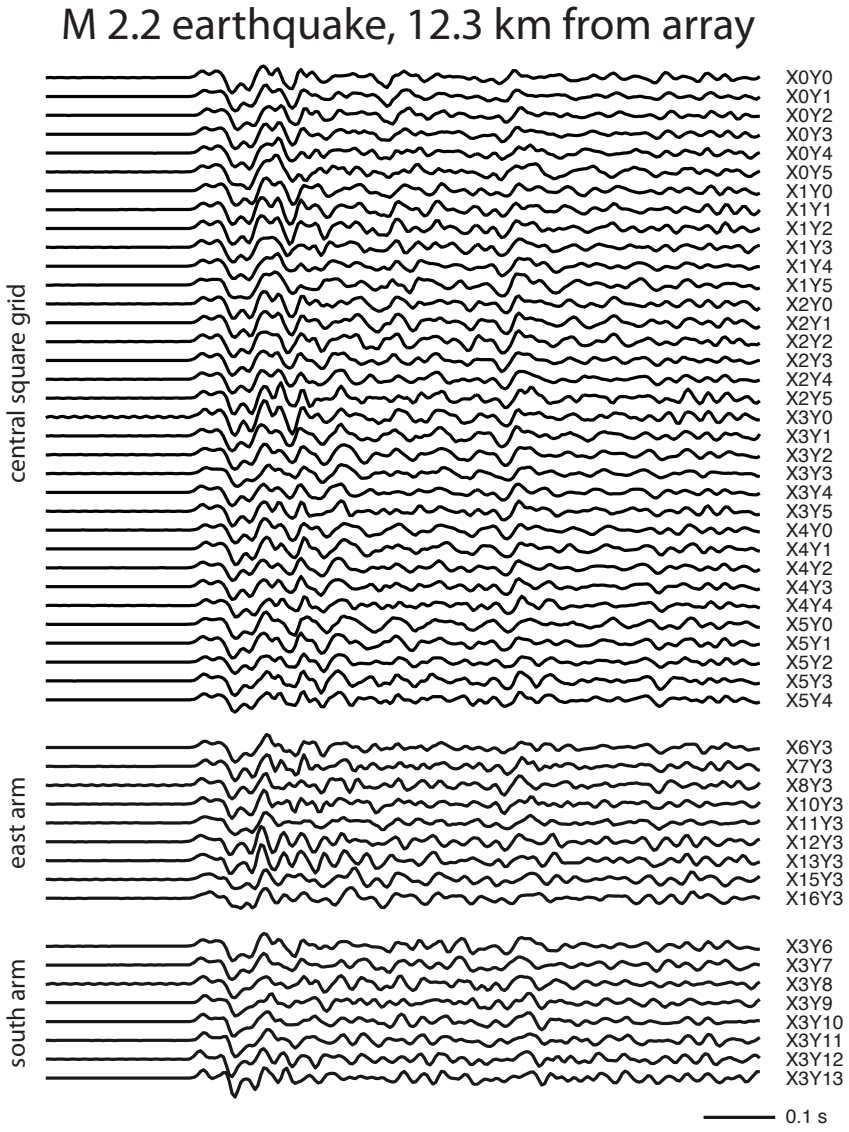

Figure 2.2: Waveforms from a M 2.2 earthquake recorded by the SAA at an epicentral distance of 12.3 km. The closely spaced stations within the central square grid record very similar waveforms while the stations in the linear arms record somewhat less similar waveforms. Focal sphere effects can be neglected because the stations are very close to each other.

Using the array data in conjunction with data from the local ANZA network, 156 local earthquakes (M0.8 to M4.0) were initially identified and located. Many of these earthquakes occurred in clusters to the south and west of the array along the trace of the SJFZ. We estimate location uncertainties to be at least 1 km. Additional earthquakes were recorded but not initially located or assigned magnitudes. We estimate locations for these events using arrival lag times determined by waveform cross-correlation of the orthogonal arm stations in conjunction with S – P arrival times. To assign magnitudes to these additional events, we determine a local scaling of magnitude as a function of source-station distance and peak amplitude. This process nets 55 additional earthquakes, bringing our catalog to a total of 211 local events (M0.7 to M4.0).

Previous analysis of data recorded by SAA shows that coherence among stations rapidly decreases above 15 Hz at almost all length scales (Vernon *et al.*, 1998). P- and S-wave coda analysis suggests the presence of localized heterogeneities in the near-surface structure (Wagner *et al.,* 1998). Al-Shukri *et al.* (1995) reports evidence of significant frequency-dependent variations in the temporal and spectral domains, and suggests that the differences can be explained by variations in the near-surface conditions.

2.3 Estimating source parameters

We use frequency domain methods with empirical Green's function (EGF) techniques to estimate corner frequencies and coseismic static stress drops from P-wave data.

2.3.1 Methodology

Seismic source parameters are routinely estimated in the frequency domain. We fit a Brune (1970) source spectrum model to relative source spectra:

$$u(f) = \frac{\Omega_0}{1 + \left(f / f_c \right)^2} \qquad < 2.1 >$$

In some studies, the value of the exponent in the denominator is allowed to vary between 1 and 3 to control how quickly the signal decays above the corner frequency, f_c. Here, the value of the exponent is set, and f_c and the long-period amplitude, Ω_0, are fit to the data. Estimates of these spectral parameters can subsequently be used to estimate parameters not directly measurable in the data.

One of the common parameters considered in the source scaling controversy is the coseismic stress drop, first formulated by Eshelby (1957):

$$\Delta\sigma = \frac{7}{16} \frac{M_0}{r^3} \qquad < 2.2 >$$

Here, r is the radius of the earthquake's assumed circular rupture patch. We can combine equation 2.2 with the predicted relationship between source radius and P-wave corner frequency given by Madariaga (1976):

$$f_c = 0.32 \frac{\beta}{r} \qquad < 2.3 >$$

where β is the S-wave velocity (the rupture velocity is assumed to be 0.9β). Doing so allows us to derive a relationship between stress drop and P-wave corner frequency:

$$\Delta\sigma = M_0 \left(\frac{f_c}{0.42\beta} \right)^3 \qquad < 2.4 >$$

Of special note in this formula is the cubic relationship of corner frequency with stress drop. We cannot measure stress drop for small earthquakes directly and instead we must estimate it from other source parameters; significant uncertainties in the corner frequency estimates will result in subsequent large uncertainties in estimates of stress drop.

2.3.2 Empirical Green's function methods

The ground motion, m, recorded at each station can be represented by a convolution of signals from the seismic source, s_0, from the effects of the travel path between the source and the recording station, p, and from the site and instrumentation effects at the station, i:

$$m = s_0 * p * i \qquad < 2.5 >$$

We use a small earthquake as an EGF to isolate the source term from the signal recorded for a larger, nearby earthquake (Hartzell, 1978). In doing so, we assume that: (1) the two earthquakes have the same radiation patterns, (2) the path and site effects are identical for both earthquakes because the earthquakes are approximately colocated, and (3) the EGF earthquake source can be treated as a point source because it is both smaller in size and shorter in rupture duration than the larger magnitude 'mainshock' earthquake. Under these assumptions, we can rewrite equation 2.5 as:

$$m = g * s \qquad < 2.6 >$$

Here, g is the ground motion recorded for the EGF earthquake and s is the relative source contribution of the two earthquakes.

Pairs of closely located earthquakes are required for the EGF method to be successful and not all earthquakes in our data catalog have an appropriate EGF earthquake. We restrict our 'mainshocks' to earthquakes M ≥ 2 and pair these with EGF earthquakes that are at least one unit of magnitude smaller and within a hypocentral distance of 3 km of the mainshock event. These limits were chosen to accommodate as many mainshock-EGF pairs as possible given the geometry of the clustered events in our dataset and the event location uncertainties. We additionally constrain our data by only using mainshock and EGF earthquake pairs with similar S-to-P maximum amplitude ratios (mean value over the array must be within a factor of 2). This functions as a simple test of source mechanism similarity because significant differences in nodal plane orientation will produce large differences in these ratios. We require all seismograms used in our analysis to satisfy a minimum mean signal-to-noise ratio of 3 for frequencies between 4 and 45 Hz. These requirements limit our useable data to a total of 7 mainshocks with 23 possible mainshock-EGF pairs.

2.3.3 Frequency domain

We compute velocity spectra for 1 second windows by using multitaper spectrum estimation with six tapers (Park, *et al.*, 1987; Prieto, *et al.*, 2009). The frequency domain representation allows us to write equation 2.6 as the multiplication of two spectra rather than the convolution of two time series:

$$M(f) = G(f)S(f) \qquad < 2.7 >$$

We subsequently compute the spectral ratio of the mainshock to the EGF records at each station in order to remove the path and site effects and isolate the source spectrum, *S(f)*, of the larger earthquake. We do not filter the waveforms or smooth the spectra prior to computing the spectral ratio and fitting the model parameters. For each spectral ratio, we use unweighted least-squares estimation to fit a source model (equation 2.1) to the logarithm and determine the relative scalar seismic moment and corner frequency of the mainshock spectrum (Figure 2.3). These fits are performed over the same frequency points at each station with the frequencies used in the inversion determined by signal-to-noise ratio (SNR) tests. We require a minimum mean SNR of 3 as measured over 15 Hz bandwidths below 45 Hz. Estimates above 45 Hz are included when the SNR requirement is met over 5 Hz bandwidths at all stations satisfying the original 45 Hz criterion. Events with fewer than 5 stations meeting SNR requirements are excluded. The spectral fits to the source models are relatively good for the individual spectral ratio measurements, and we find that the n=2 model is adequate (Figure 2.3). We do not fit the higher corner frequency of the EGF event because we expect these frequencies to be near or above the limitations imposed by SNR constraints, and we assume that deviations from a flat EGF source spectrum will be small in this frequency range. The mainshock corner frequencies estimated at each individual station exhibit variations across the array, although the spectral fits visually appear similar in log-space.

While it is often a good practice to apply an inverse weighting with frequency to this inversion (*e.g.*, Ide *et al.*, 2003; Prejean and Ellsworth, 2001) to account for the large number of samples at high frequencies, this method does not produce reliable spectral fits for this dataset. The spectral ratios tend to exhibit a characteristic dip between 10 and 20

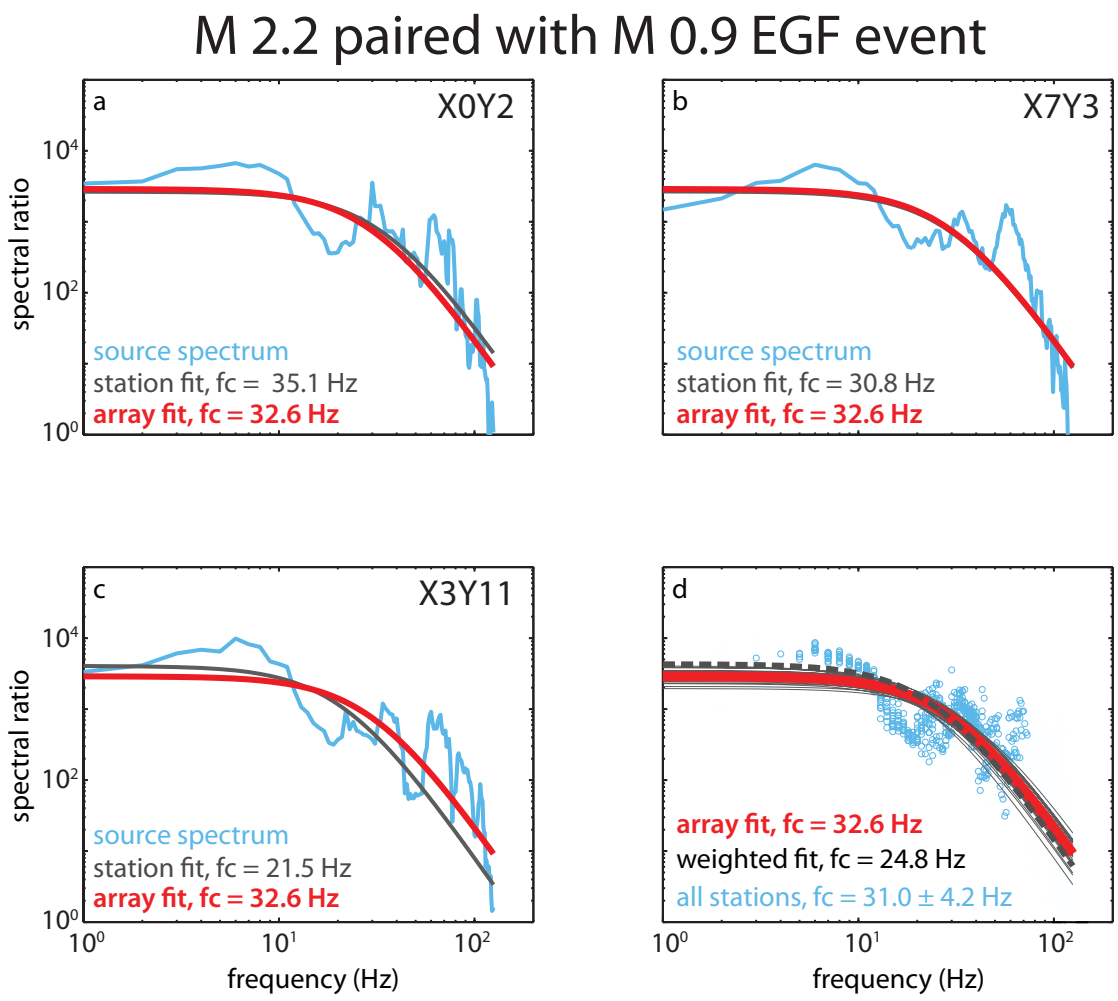

Figure 2.3: Example source spectrum fitting for one earthquake pair. Spectra plotted in the lightest shade show the mainshock/EGF spectral ratio at a single station in (a) the central square grid, (b) the east arm, and (c) the south arm. The thin gray curve shows the spectral fit for that single station with corner frequencies given in the legends. These fits are performed over the same band at each station with the maximum frequency determined by SNR limits. The spectral fit obtained using all stations simultaneously without weighting is plotted as the thick black line in each subplot. In (d), the spectral ratios used in the estimates (circles) and individual fits at all stations (thin gray lines) are shown. The dashed line in (d) shows the simultaneous spectral fit obtained when using weighted least-squares, with the weighting given by the variance of the amplitudes at each frequency. The spectral ratio shown in gray is from the deeper borehole sensor; the shape is comparable to those of the surface stations but the borehole sensor records are excluded from the inversion.

Hz for most event pairs and the inverse frequency weighting prioritizes this dip over the roll-off at higher frequencies. This spectral feature is consistent across the dataset and appears larger in some cases then in others.

To test the robustness of individual corner frequency fits, we measure the spectral misfit while varying the corner frequency and relative seismic moment around each original estimate. Viegas *et al.* (2010) used a similar grid search technique to constrain source parameter estimates within a 5% increase in fit variance. We apply a comparable limit by measuring corner frequencies at a spectral misfit increase of 5% from the original fit to obtain confidence intervals for each estimate.

We obtain a more stable estimate of these source parameters by also fitting for relative moment and corner frequency simultaneously over all stations in the array using an unweighted least-squares inversion. This method produces array fits very similar to the individual station fits. Performing the array inversion using weighted least-squares, with weighting given by the variance of the amplitudes at each frequency point, results in consistently higher relative moment estimates and lower corner frequency estimates (Figure 2.3). This is due to the smaller variation of amplitudes exhibited at lower frequencies as compared to the noisier high frequency amplitudes. The stations are very close together relative to the distance between the earthquakes and the array, allowing us to ignore any potential complications of azimuthal variations in the source when considering all stations in the array simultaneously. Thus we focus on the differences in the results obtained for the different stations while recognizing that the single takeoff angle sampled is an incomplete representation of the source. With an azimuthally distributed array, such simultaneous spectral fitting should average any variations due to

rupture directivity and differences in mainshock and EGF radiation patterns. The change to the overall spectral fit and subsequent change in stress drop estimates due to using weighted least-squares demonstrates that differences in fitting methods may introduce strong biases in the results.

Our corner frequency estimates are consistent with those from previous studies for earthquakes in this magnitude range and region (Prieto *et al.*, 2004). The distribution of individual station estimates has a consistent spread over the range of magnitudes considered after normalizing by the mean corner frequency estimate for each event. The variations in these estimates are not due to the contribution of a single event alone. We do not draw conclusions on the presence or lack of earthquake self-similarity due to the limited range of magnitudes in our selection of mainshock earthquakes ($2.0 < M < 3.4$) and narrow azimuthal coverage.

These frequency domain methods are commonly used to estimate earthquake source parameters, but they suffer from several limitations and require assumptions that are not always testable. The spectral model assumes a circular, shear rupture at the source. The EGF technique assumes that mainshock and EGF earthquakes are colocated (rarely the case in most datasets) and that the event pairs chosen have identical source radiation patterns. Finally, these methods require the difference in corner frequencies between the mainshock and EGF earthquakes to be large enough to resolve. If the corner frequencies are too close to each other, then the mainshock corner frequency estimates will be biased.

2.3.4 Stress drop

We treat the cataloged local magnitudes as moment magnitudes to determine seismic moments for computing stress drops (equation 2.4). While these magnitude scales are likely not equivalent in this magnitude range (*e.g.*, Shearer *et al.*, 2006), we assume that the appropriate scaling is the same for all events. As our primary focus is on examining the distribution of estimates over the array for each mainshock earthquake, any differences from the absolute moments will not affect our results. Indeed, we find that the stress drops determined from these assumptions are higher than might be expected from other studies with a median value of 115 MPa. These differences could be attributed to our treatment of local magnitudes as moment magnitudes, truly higher stress drops in the region (Prieto, *et al.*, 2004; Frankel and Kanamori, 1983), or source effects (*e.g.*, rupture directivity) that affect the corner frequency results due to the narrow source-array azimuth we are considering. Stress drops computed from the array spectral fits are somewhat lower than those resulting from the individual station estimates. The small range of earthquake magnitudes in our dataset limits our abilities to determine if stress drop scaling is constant or varying with earthquake magnitude.

2.4 Uncertainties in source parameter estimates

Our primary goal in this study is to measure station-to-station variations in source parameter estimates in the small aperture array and assign an appropriate estimate of uncertainty for source parameter estimates at a single station. We limit our analysis to the surface stations, and quantify variability in estimates of corner frequency by considering

the percent deviation from the mean estimate over all stations for each pair of events. This normalization is necessary to reasonably compare variations in parameters estimated over a range of earthquake magnitudes. The distribution of percent deviation from the mean for corner frequency exhibits somewhat long tails (Figure 2.4). We measure the width of this distribution with the interquartile range (IQR), which gives the difference between the third and first quartile. The IQR estimate is 23% for the corner frequency distribution. This means that if our distribution is exactly symmetric, then half of our data falls within the +/- 11.5% of the mean corner frequency estimate for each pair.

Looking in further detail at each of the individual 23 pairs of events shows that some pairs exhibit a much wider distribution of corner frequency estimates than the others. These pairs are not limited to any particular mainshock or EGF earthquake and are likely due to matching mainshock earthquakes with EGF earthquakes that do not meet all of our EGF technique assumptions. Pairs have been excluded from our dataset based on the separation distance between the mainshock and EGF earthquake, signal-to-noise ratio limitations, and source similarity requirements; these remaining pairs successfully passed each of these tests. Examination of the spectra for the most anomalous pairs of events shows that the spectra are not consistently fit over all stations due to complexity in the spectral shape, and the corner is fit in different locations depending on the station considered. To examine the degree to which including such pairs affects our results, we remove pairs of events that have IQR values greater than twice the mean IQR of the entire dataset; this process limits the dataset further to 21 pairs out of the total 23 considered previously. After removing these pairs from our dataset, the resulting distribution of corner frequency deviations for all remaining pairs decreases in width

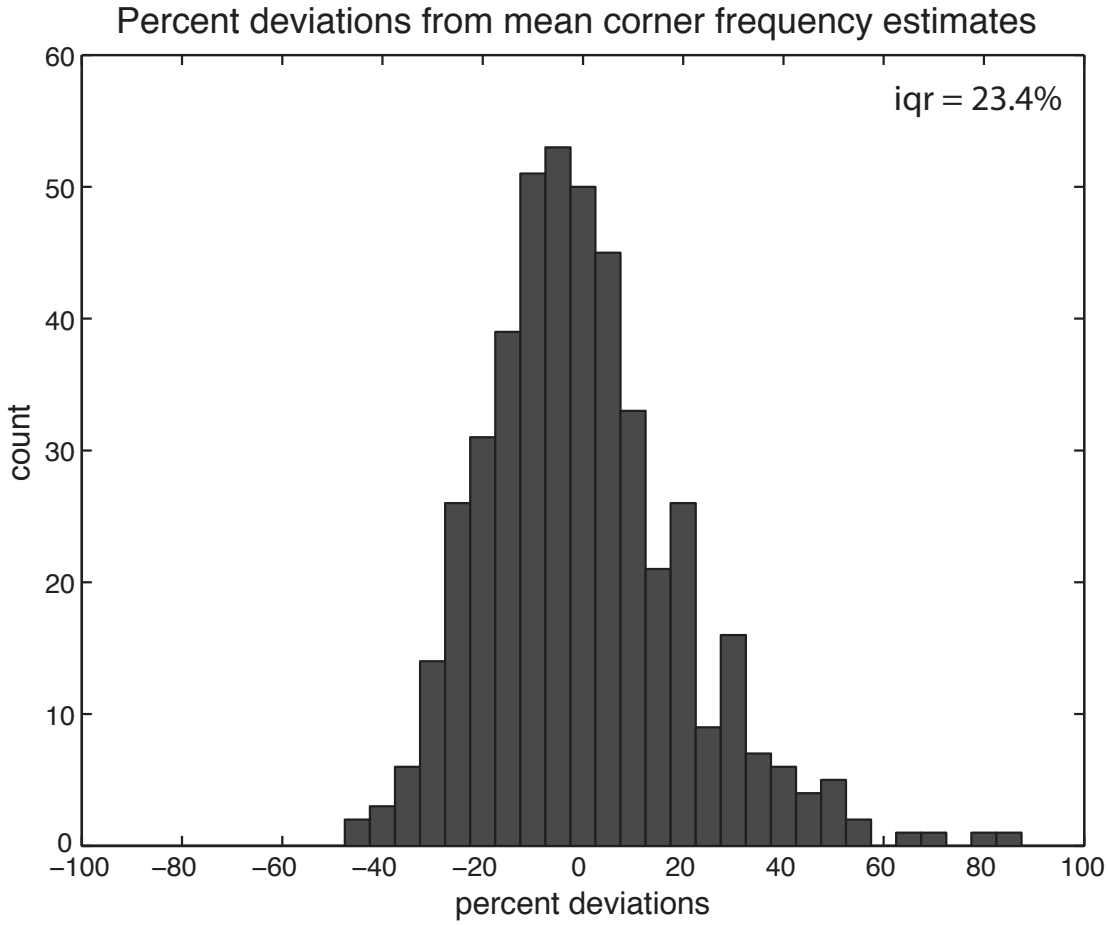

Figure 2.4: Distribution of percent deviations from the mean for corner frequency estimates.

slightly from 23% to 22.5%. All following estimates include the outliers unless otherwise specified. Our corner frequency distributions generally fit a lognormal distribution regardless of inclusion or exclusion of the outlier pairs; the apparent skewness in the distribution is due to plotting the percent deviation.

We do not observe any systematic dependence of distribution width on either mainshock magnitude or mainshock-EGF magnitude differential; this suggests that the variability we are measuring is due to variations at the stations rather than the sources. The sizes of the confidence intervals also show no strong dependence on event magnitude after normalizing by the corner frequency estimates. We find that the confidence intervals for each individual corner frequency estimate are generally larger than the width of the distribution describing the variation in estimates across the array. This suggests that estimating confidence intervals using the grid search method may provide a reasonable means of approximating the source parameter uncertainties in some cases.

We consider stress drop deviations from the mean in the log-domain rather than by percent. For stress drop estimates determined from the frequency domain results, the log-domain IQR is 0.31 (Figure 2.5). If we assume a symmetric distribution, then this corresponds to an uncertainty range of ±0.15 on each log stress drop estimate to contain the middle half of the data (e.g., an estimate of 1 MPa stress drop would be equivalent to 0 ± 0.15 in the log-domain, or error bars from 0.7 to 1.4 MPa in the linear domain). This value scales logarithmically when applied to absolute estimates and corresponds to an uncertainty range of 0.07 to 0.14 MPa for a mean stress drop estimate of 0.1 MPa, and 7 to 14 MPa for a mean estimate of 10 MPa. Excluding the outlier pairs decreases the IQR by less than 0.01, demonstrating that our results are robust. These values suggest we

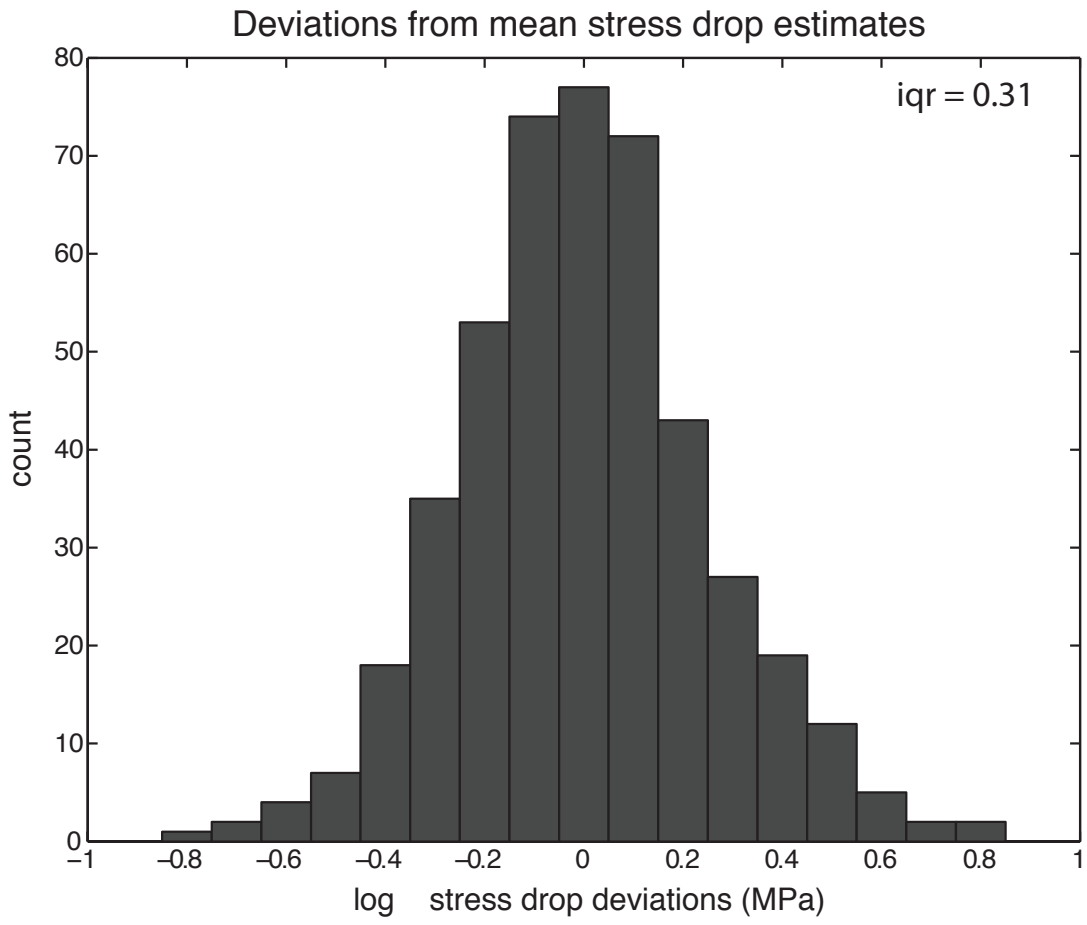

Figure 2.5: Distribution of deviations from the mean stress drop estimates (in log-space) using frequency domain methods.

should consider minimum uncertainties of ±~30% for individual station estimates of stress drop, particularly because these uncertainty estimates are based on IQR estimates which are more conservative than one standard deviation.

2.5 Variations due to data choice

The variations we observe in source parameter estimates could have several origins. We verify that the variations seen in the source parameter estimates are not dominated by unusual source locations, large local site effect variations at a subset of the stations in the array, or choices of mainshock and EGF earthquake pairs. We review the variations of estimates in each of these populations to confirm that our results are not biased by a subset of our data. We do not observe a correlation between mainshock magnitudes and the subsequent distribution widths of the estimates.

2.5.1 Station location

The location of individual stations in the array could contribute to the variations we see in the source parameter estimates if there are consistent variations in the very localized rock structure near a specific station. Previous work with this dataset has demonstrated a loss in coherence across the array above 15 Hz and suggested that small-scale variations in local structure affect the propagation of ground motion significantly on the scale of station spacing in this array (Vernon *et al.*, 1998).

We compare the distribution of source parameter estimates from three different station groupings: (1) the central square grid, (2) the south arm, and (3) the east arm (Figure 2.6). All three sets display relatively symmetric distributions of estimates, confirming that the distribution of results from the full array is not significantly affected by one of these groupings. The central square grid contains the highest number of stations, generating a much smoother distribution of estimates.

We consider variations from the mean estimate for each mainshock-EGF earthquake pair on a station-by-station basis to look for site effects on a smaller scale than the footprint of the central grid and of the two arms (Figure 2.6). Groupings of positively and negatively biased stations are apparent for each pair of events but this pattern is not consistent when comparing pairs among each other. In addition, we find that variations in source parameter estimates are consistent over all station separation distances rather than being a function of separation distance. Although local structural variations may affect the estimates at each station, these variations should not be dominant when combining deviations from the mean estimate over all stations and all mainshock-EGF earthquake pairs.

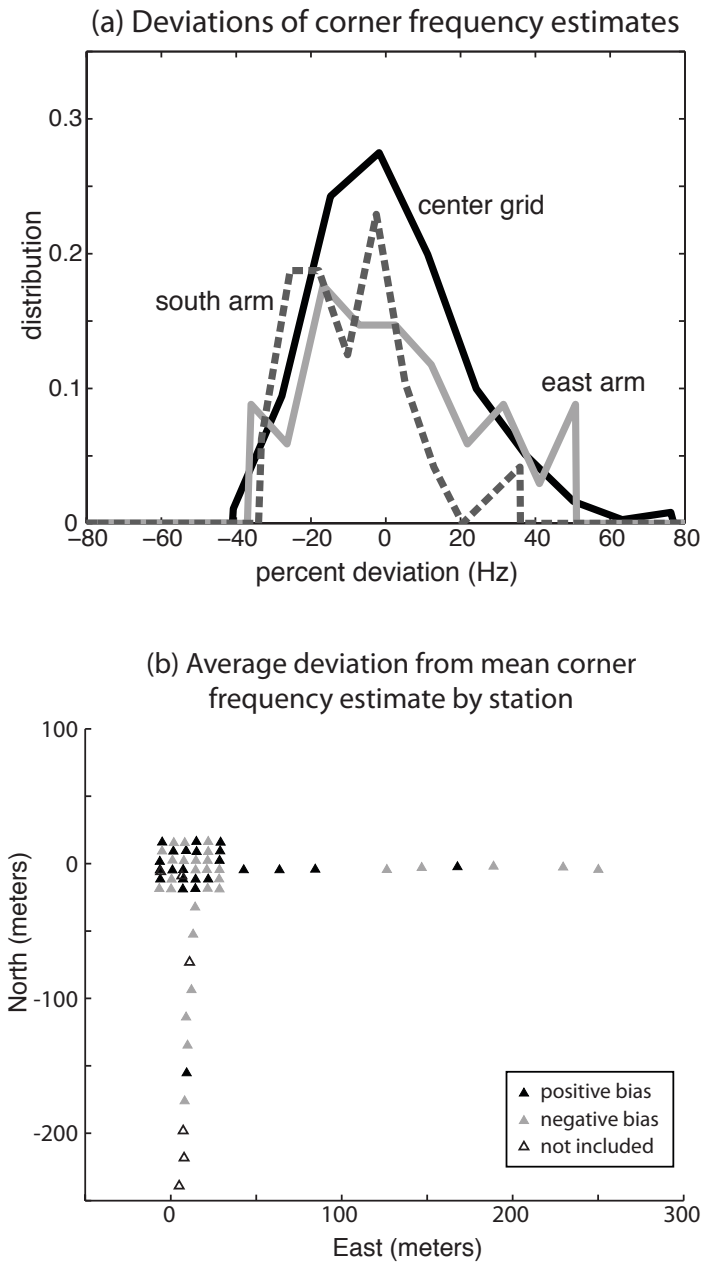

Figure 2.6: Analysis of effects of grouping stations. In (a), PDFs of each of 3 groups are shown. In (b), mean deviations of each station for all event pairs are coded by color. Although clustering of negatively and positively biased stations are shown, these patterns are not stable and vary with each pair of events.

2.5.2 Earthquake pair location

We next check to see if variations in the distribution of source parameter estimates are related to a particular set of earthquake sources. We first look at the distribution of estimates from nearby earthquakes (within 30km of the array) and compare these estimates with those from more distant earthquakes. We do not find a consistent difference between the distributions of these two populations. Additionally, the closest and furthest events from the array do not appear to give substantially different results, confirming that we are not merely measuring uncertainties due to signal quality.

We also look at the distribution of estimates based on source-station azimuth to check if earthquakes from one source location produce results differing considerably from those in a different source region, which would skew our resulting distribution. Again, we do not find consistent variations in the distribution of estimates from these groups. The location of the earthquake pair does not significantly affect the distribution of estimate deviations observed.

2.5.3 Choice of EGF earthquake

The EGF method assumes that the mainshock earthquake and the EGF earthquake are colocated and have the same source mechanism. This is commonly addressed by requiring the potential EGF hypocenters to be within a specified distance of the mainshock hypocenter. Here, we require mainshock and EGF earthquakes to be within 3 km of each other. This hypocentral separation is larger than the suggested limits of some

studies (*e.g.*, Mori and Frankel, 1990) but is reasonable to consider given the location uncertainties of events recorded by this network. We confirm that the mainshock and EGF events in each pair are closely located to each other by comparing P and S arrival time differences. Analysis of these arrival time differences shows that 4 of the 23 pairs may have mainshock-EGF hypocentral separation distances of at least 2 km. Of these 4 pairs, we previously suggested one pair might be an outlier given the wide distribution of corner frequency estimates. The remaining 3 pairs do not have larger distribution widths than the remaining data.

We additionally require the two earthquakes to have similar ratios of S-wave amplitude to P-wave amplitude. If the focal mechanisms of the two events are significantly different, then the S- to P-wave amplitude ratios should differ as well, whereas earthquakes with the exact same focal mechanism should have the same S- to P-wave amplitude ratio. This requirement removes any potential EGF earthquakes with sources differing significantly from the mainshock source.

We find that the choice of EGF earthquake for each mainshock earthquake is one of the largest contributors to producing variations in absolute estimates. We consider an event of M 3.4 with 6 EGF pairs (M 1.4 to M 2.3) to test how the EGF choice affects our results. The distributions of deviations from the mean of the results for each of these possible EGFs show similar widths of uncertainties of stress drop estimates regardless of hypocentral separation distance or EGF magnitude (Figure 2.7). However, the actual values of corner frequency (or duration or stress drop) estimated for each of these earthquake pairs vary. This effect could be due to the earthquakes in each pair being too far apart from each other (therefore having different propagation paths between the

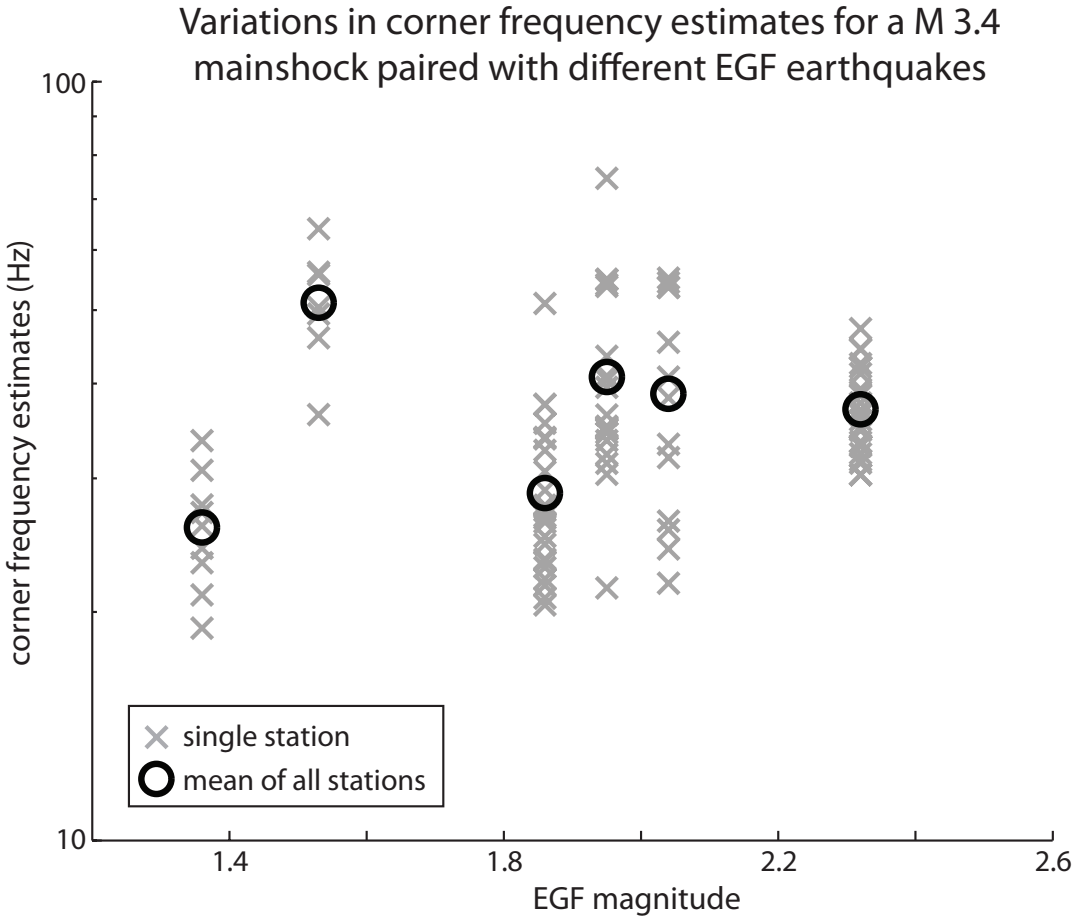

Figure 2.7: Corner frequency estimates at each station plotted versus the magnitude of the corresponding EGF earthquake for a single M 3.4 earthquake.

source and station), having sources with significantly different properties, or being too close in magnitude to each other (hence the smaller EGF earthquake having a corner frequency close enough to the mainshock corner frequency that the fitting routine is unable to effectively separate between the two corners). We do not observe any systematic relationship between corner frequency estimates and hypocentral separation distances. In this dataset, the similarity of the earthquake source mechanisms may have the greater effect on absolute corner frequency estimates; we are unable to verify this with focal mechanisms due to lack of data.

2.6 Discussion

Whether earthquakes scale linearly with magnitude over the full range of rupture sizes is an important question to resolve, as it has considerable implications for studies of earthquake rupture physics and seismic hazards in large, damage-producing earthquakes. If the conditions produced during rupture differ along the range of earthquake magnitudes, then our abilities to forecast ground motion from large earthquakes could be limited in locations with only small recorded earthquakes. Self-similarity over all earthquake sizes permits the consideration of smaller earthquake waveforms, which are recorded much more frequently and have populations of sufficient size for reasonable statistical analysis.

Uncertainties in source parameter estimates have largely been ignored in many past source parameter studies, perhaps because it is difficult to sufficiently quantify these uncertainties with only a limited distribution of seismic stations. These uncertainties can

be sizeable, however, when we consider that static stress drop scales with the cube of corner frequency. A comparison of the variation in corner frequency estimates across the array of stations with the confidence intervals determined by the grid search technique suggests that such techniques may be useful for approximating source parameter uncertainties in general. Additional methods for assessing uncertainties of source parameter estimates may include use of multiple EGF events (*e.g.*, Prieto, *et al.*, 2006) and use of different source spectral ratios (*e.g.*, Malagnini and Mayeda, 2008).

We have quantified the distribution of stress drop deviations from the mean values using frequency domain techniques, and we find that the resulting distribution suggests that uncertainties of at least ~30% of the absolute measurement are appropriate to consider for estimates made at a single station. Additionally, our work shows that any small-scale site effects due to a heterogeneous surface layer do not consistently bias our results.

The largest contribution to variations in the absolute estimates of source parameters appears to be due to the choice of EGF earthquake. Our method for choosing potential mainshock-EGF earthquake pairs relies on a simple hypocentral separation distance constraint and a comparison of the S and P amplitude ratios of the two waveforms. This may not be sufficient, however, as we find that the estimates of source parameters for a given mainshock can vary significantly depending on the choice of EGF earthquake. This likely results because of differences in the source mechanisms of the two events (as the EGF technique assumes that the source radiation patterns are identical), from too large of a distance between the hypocenters, or due to the separation

in seismic moment and corner frequency between the two events not being large enough to be individually resolved.

Our results suggest that much care needs to be taken when comparing source parameter estimates among various mainshocks as the uncertainties in estimates could be large. Averaging over several stations, as is done in many studies, may reduce the uncertainties of the estimates but introduces potential complications related to different path effects and rupture directivity effects that vary with azimuth. The resulting estimates may be heavily influenced by the choice of method, model, or EGF earthquake. Combining results obtained by various methods should be approached with caution and interpretations of past studies may need to take such uncertainties into account. Devising reliable techniques for estimating source parameter uncertainties will be necessary to draw robust conclusions regarding variations in the seismic source in future studies.

Data and Resources

Data from the Pinyon Flat High Frequency Array are available through the IRIS DMC under network code 1990 XA. Additional information about the experiment can be found in PASSCAL Data Report 91-002.

Acknowledgements

We thank those associated with the collection, analysis, and cataloging of the Small Aperture Array data. Discussion with Gary Pavlis and his ongoing interest in the

Small Aperture Array project were helpful in completing this work. We appreciate early reviews from and invaluable discussion with Debi Kilb. Bob Parker provided useful advice on techniques in inverse theory. We thank Rachel Abercrombie and an anonymous reviewer for their helpful comments on this manuscript. This material is based upon work supported by the National Science Foundation under Grant Numbers EAR-0908042, EAR04-17983, and DGE 0841407, and by the Incorporated Research Institutions for Seismology under their Cooperative Agreement No. EAR-0733069 with the National Science Foundation. This work (in full) is a reformatted version of the material as it appears in the Bulletin of the Seismological Society of America (Kane, D.L., G.A. Prieto, F.L. Vernon, and P.M. Shearer (2011). Quantifying seismic source parameter uncertainties, Bull. Seism. Soc. Am., Vol. 101, No. 2, pp. 535–543, doi: 10.1785/0120100166) and is reprinted with permission. Deborah Kane is the primary investigator and author of this paper.

References

Abercrombie, R. (1995), Earthquake source scaling relationships from -1 to 5 ML using seismograms recorded at 2.5-km depth, *J. Geophys. Res.*, 100 (B12), 24015-24036.

Abercrombie, R., and J. Rice (2005), Can observations of earthquake scaling constrain slip weakening?, *Geophys. J. Int.*, 162 (2), 406-424.

Al-Shukri, H., G. Pavlis, and F. Vernon (1995), Site effect observations from broadband arrays, *Bull. Seism. Soc. Am.*, 85 (6), 1758-1769.

Brune, J. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, 75 (26), 4997-5009.

Eshelby, J. (1957), The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. Roy. Soc. Lond. Math. Phys. Sci.*, 241 (1226), 376-396.

Frankel, A., and H. Kanamori (1983), Determination of rupture duration and stress drop for earthquakes in southern California, *Bull. Seism. Soc. Am.*, 73 (6), 1527-1551.

Hartzell, S. (1978), Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.*, 5(1).

Ide, S., and Beroza, G. (2001), Does apparent stress vary with earthquake size?, *Geophys. Res. Lett.*, 28 (17), 3349–3352.

Ide, S., G. Beroza, S. Prejean, and W. Ellsworth (2003), Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. Geophys. Res*., 108(B5) (2271), 029.

Imanishi, K., and W. Ellsworth (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 Abercrombie, McGarr, Kanamori, and di Toro (Editors), *AGU Geophys. Monograph*, 170, 81-90.

Madariaga, R. (1976), Dynamics of an expanding circular fault, *Bull. Seism. Soc. Am.*, 66 (3), 639-666.

Malagnini, L. and K. Mayeda (2008), High-stress strike-slip faults in the Apennines: An example from the 2002 San Giuliano earthquakes (southern Italy), *Geophys. Res. Lett.*, 35, L12302, doi:10.1029/2008GL034024.

Mayeda, K. and W. Walter (1996), Moment, energy, stress drop, and source spectra of western United States earthquakes from regional coda envelopes, *J. Geophys. Res.*, 101 (B5), 11195-11208.

Mori, J. and A. Frankel (1990), Source parameters for small events associated with the 1986 North Palm Springs, California, earthquake determined using empirical Green's functions, *Bull. Seism. Soc. Am.*, 80 (2), 278-295.

Mori, J., R. Abercrombie, and H. Kanamori (2003), Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake, *J. Geophys. Res.*, 108 (B11, 2545).

Owens, T., P. Anderson, and D. McNamara (1991), 1990 Pinyon Flat High Frequency Array Experiment: An IRIS Eurasian Seismic Studies Program Passive Source Experiment, PASSCAL Data Report 91-002 .

Park, J., C. Lindberg, and F. Vernon (1987), Multitaper spectral analysis of high-frequency seismograms, *J. Geophys. Res.*, 92 (B12), 12675-12684.

Prejean, S.G. and W.L. Ellsworth (2001), Observations of Earthquake Source Parameters at 2 km Depth in the Long Valley Caldera, Eastern California, *Bull. Seism. Soc. Am.*, 91 (2), 165-177.

Prieto, G. A., R. L. Parker, F. L. Vernon, P. M. Shearer and D. J. Thomson (2006), Uncertainties in earthquake source spectrum estimation using empirical Green functions, Earthquakes: Radiated Energy and the Physics of Faulting, Abercrombie, McGarr, Kanamori, and di Toro eds. *AGU Geophys. Monograph*, 170, pp. 69-74. doi : 10.1029/170GM08.

Prieto, G., P.M. Shearer, F.L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking P and S spectra, *J. Geophys. Res.*, 109 (B08310).

Prieto, G., R. Parker, and F. Vernon (2009), A Fortran 90 library for multitaper spectrum analysis, *Comput. Geosci.*, 35, 1701-1710.

Shearer, P.M., G.A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in southern California, *J. Geophys. Res*., 111, B06303, doi:10.1029/2005JB003979.

Vernon, F., G. Pavlis, T. Owens, D. McNamara, and P. Anderson (1998), Near-surface scattering effects observed with a high-frequency phased array at Pinyon Flats, California, *Bull. Seism. Soc. Am.*, 88 (6), 1548-1560.

Vernon, F., J. Fletcher, L. Carroll, and A. Chave (1991), Coherence of seismic body waves from local events as measured by a small-aperture array, *J. Geophys. Res.*, 96 (B7), 11981-11996.

Viegas, G., R.E. Abercrombie, and W.-Y. Kim (2010), The 2002 M5 Au Sable Forks, NY, earthquake sequence: Source scaling relationships and energy budget, *J. Geophys. Res.*, 115, B07310, doi:10.1029/2009JB006799.

Wagner, G. (1998), Local wave propagation near the San Jacinto fault zone, southern California: Observations from a three-component seismic array, *J. Geophys. Res.*, 103 (B4), 7231-7246.

Wilson, D.C. and G.L. Pavlis (2000), Near-Surface Site Effects in Crystalline Bedrock: A Comprehensive Analysis of Spectral Amplitudes Determined from a Dense, Three-Component Seismic Array, *Earth Interactions*, Vol. 4, No. 2.

Chapter 3: Application of Time-Domain Deconvolution and Empirical Green's Function Techniques to Obtain Source Time Functions of Small Earthquakes

Quantifying source parameters of earthquakes is fundamental to understanding the physics of earthquake rupture. Researchers commonly estimate parameters such as seismic moment, corner frequency, and stress drop, but the variability in estimates can be large due to the use of different methods and source models. Most source parameter estimates are made using frequency-domain model-based methods, and efforts to use both time-domain and non-parametric methods have been less utilized. We describe a non-parametric method for estimating earthquake source time functions (STFs). This method uses empirical Green's functions and a non-negative least-squares inversion in the time domain to compute the STF. It offers the advantage of being able to incorporate a series of EGF waveforms simultaneously to generate a more robust STF solution. This simultaneous inversion approach makes it possible to include EGF waveforms from a given event recorded at multiple azimuths or to include multiple EGF events recorded at a single station. We demonstrate the application of this method under various situations with synthetic data and with data from two arrays in southern California.

3.1 Introduction

The source time function (STF) of an earthquake represents the coseismic motion on the fault plane as a function of time, and is proportional to the earthquake's moment rate function. STFs provide a time domain view of earthquake rupture and can be used to estimate the duration of an event, the seismic moment released, the rupture velocity of the earthquake, and any temporal changes associated with these parameters. STFs are routinely computed as a component of the moment tensor solution for large earthquakes ($M > 5$) but are rarely determined for small events.

Models of earthquake rupture used to estimate seismic source parameters often treat smaller earthquakes as simple shear ruptures occurring within planar circular patches (*e.g.*, Haskell, 1964; Brune, 1970; Madariaga, 1976). STFs corresponding to these models exhibit simple, symmetric pulse shapes (*e.g.*, Haskell, 1964 model) with defined rupture rise time and duration. These STFs can be represented by simple analytical models in the frequency domain. Larger magnitude earthquakes rarely exhibit such simple STF shapes, however, and more commonly feature asymmetrical STF pulses with multiple peaks. Developing methods to resolve such features in STFs of small earthquakes while incorporating minimal assumptions is important for drawing conclusions about the rupture processes of small earthquakes and for further developing theories of earthquake source physics beyond relatively simple models.

Such small local earthquakes can be studied by using empirical Green's function (EGF) techniques to correct for the effects of the propagation path between the earthquake and the recording instrument (*e.g.*, Hartzell, 1978; Frankel and Kanamori, 1983). This method requires ground motion records of a target 'mainshock' earthquake

and of a smaller colocated earthquake with a similar source mechanism. EGF techniques effectively separate the contributions of the source and path propagation in the recorded ground motion and can be used to isolate the source signal. Here, we assume that the ground motion produced by an earthquake is equivalent to the earthquake's STF convolved with the ground motion of this smaller, colocated EGF earthquake.

Mathematically, this can be represented by a simple formula:

$$m(t) = g(t) * s(t) \qquad < 3.1 >$$

Here, *m(t)* is the ground motion time series recorded for the larger earthquake (hereafter referred to as the 'mainshock'), *g(t)* is the ground motion time series recorded for the smaller EGF earthquake, *s(t)* is the STF, and * represents convolution. One common method of computing the STF uses the frequency domain, where the spectrum of the resulting STF is determined by computing the spectral ratio of the mainshock record to the EGF record (*e.g.*, Mori and Frankel, 1990; Ide *et al.*, 2003; Prejean and Ellsworth, 2001). This method of spectral division can be problematic in frequency bands with low EGF spectral amplitudes, and the use of approximations such as waterlevel regularization are sometimes included to stabilize the result. The STF is then obtained by transforming the spectral ratio back to the time domain.

More recently, several studies have used the projected Landweber deconvolution (PLD) method (Bertero *et al.*, 1997) to estimate STFs (*e.g.*, Bertero, *et al.*, 1998; Lanza, *et al.*, 1999; Vallee, 2004; McGuire, 2004; Harrington and Brodsky, 2009; Oth *et al.*, 2011). The PLD method is more stable than simple spectral deconvolution and makes use of the computationally efficient frequency domain techniques described in Bertero et al. (1997). However, stable results are best obtained with user interaction in the algorithm to

subjectively select misfit values for the final result, and only a single pair of records can be addressed at a time.

In this study, we reconsider the methods of computing STFs. We describe a simple method of computing STFs that does not require heavy regularization, is computationally simple and is non-parametric with robust results. We demonstrate this approach with a series of synthetic examples and with earthquakes recorded in southern California. Similar techniques have been used successfully in previous studies (see references discussed below). We additionally describe an extension of this method in which waveforms from multiple EGF earthquakes can be used simultaneously to provide increased stability to the result and robust estimates of STFs.

3.2 Method

We use time domain deconvolution with EGF techniques to estimate the STF of an earthquake. This method assumes that the ground motion due to an earthquake is equivalent to the earthquake's source time function convolved with the ground motion due to a smaller, colocated earthquake with the same focal mechanism orientation. We perform discrete deconvolution in the time domain by creating a Green's function matrix, G, out of the EGF waveform, $g(t)$, recorded at a given station, and we express the convolution as a multiplication of matrices (Bracewell, 2000):

$$m = Gs \qquad < 3.2 >$$

where

$$G = \begin{bmatrix} g(t_1) & 0 & 0 & \cdots \\ g(t_2) & g(t_1) & 0 & \cdots \\ g(t_3) & g(t_2) & g(t_1) & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ g(t_M) & g(t_{M-1}) & g(t_{M-2}) & \cdots \end{bmatrix}$$

Here, m and s are vectors containing the mainshock waveform with M samples and a resulting STF of length L samples. G includes the EGF waveform expressed as a lower triangular matrix with the first sample of $g(t)$ repeated along each diagonal element; G has dimensions of M by L. We use least-squares inversion to solve for the STF, s. This method requires prescribing a maximum allowed length for the STF. We apply a nonnegativity constraint because we do not expect the resulting STFs to be negative (which would indicate motion opposite the direction of rupture). The nonnegativity constraint additionally acts as a simple filter on the resulting STF without necessitating filtering of the waveforms prior to inversion. Additional regularization introduced by minimizing the second derivative of the STF produces much smoother results when working with a single pair of waveforms (mainshock and EGF earthquakes recorded at a single station).

A significant advantage of this method lies in the ability to simultaneously invert multiple pairs of records of the same mainshock/EGF event pair for a single event by extending the waveform vectors and matrices:

$$\begin{bmatrix} [m_1] \\ [m_2] \\ \\ [m_N] \end{bmatrix} = \begin{bmatrix} [G_1] \\ [G_2] \\ \\ [G_N] \end{bmatrix} s \qquad < 3.3 >$$

Here, the set of matrix equations includes the mainshock and EGF waveforms for N record pairs. The vector of mainshocks has length MN while the vector for the STF waveform remains length L. The matrix of EGF waveforms is now composed of several G_i matrices and has dimensions MN by L. Inversion for the STF is performed identically to that of the single record pair method using these extended matrices.

This approach provides additional stability to the STF estimation because the resulting STF contains the signal common to all mainshock waveforms and gives less influence to signals present at individual stations only. Applying a simultaneous minimization of the STF's second derivative can be further used to stabilize the STF result, although the results without this additional step are generally more robust than the single record pair results and this minimization is not always necessary.

Variations on this inversion for STFs have been described in previous studies (*e.g.*, Velasco *et al.*, 1994; Gurrola *et al.*, 1995). However, these methods are not in common use for studies of smaller earthquakes, perhaps due to the ubiquitous use of the spectral ratio method. While the spectral ratio method is sometimes used with multiple waveform pairs to achieve a similar simultaneous result (*e.g.*, Hough, 1997; Prieto *et al.*, 2004), it brings the distinct disadvantage of assuming a circular shear source model and does not allow deviation from these simple parameters.

A novel way of applying this inversion technique to robustly estimate the STF for a given event is to use multiple EGF events simultaneously, in a combination analogous to the multiple EGF method sometimes used in the frequency domain (Hough, 1997; Prieto *et al.*, 2004). Using multiple EGF events in the deconvolution procedure lessens any effects due to assumptions not being met appropriately for a given EGF event (*e.g.*,

mismatched focal mechanism orientation) by averaging over several events, produces smoother solutions, and allows for a consideration of azimuthal variability of the STF if rupture directivity is present.

3.3 Examples

We apply this technique to several synthetic and recorded waveforms to demonstrate its effectiveness in various scenarios.

3.3.1 Single event pair at one station

We start with the most simple case: a single pair of waveforms. We create a synthetic dataset by convolving a synthetic STF with an actual small earthquake record (here, the EGF) and adding uniformly distributed noise with a maximum amplitude of 10% of the peak waveform amplitude to create a synthetic mainshock waveform. We then compute the STF for the mainshock/EGF record pair and compare the results to the input STF (Figure 3.1).

The resulting STF obtained using the most simple inversion parameters (nonnegative least-squares without further constraints) matches the input synthetic STF fairly well in overall shape but exhibits large spikes in amplitude between consecutive sample points. No additional signals beyond the added noise have been introduced in this synthetic dataset. Thus, the resulting oscillations in the STF must be related to the added noise. We convolve the resulting STF with the EGF to determine how well the mainshock waveform is fit, and we find that the RMS waveform misfit is only 3%. This

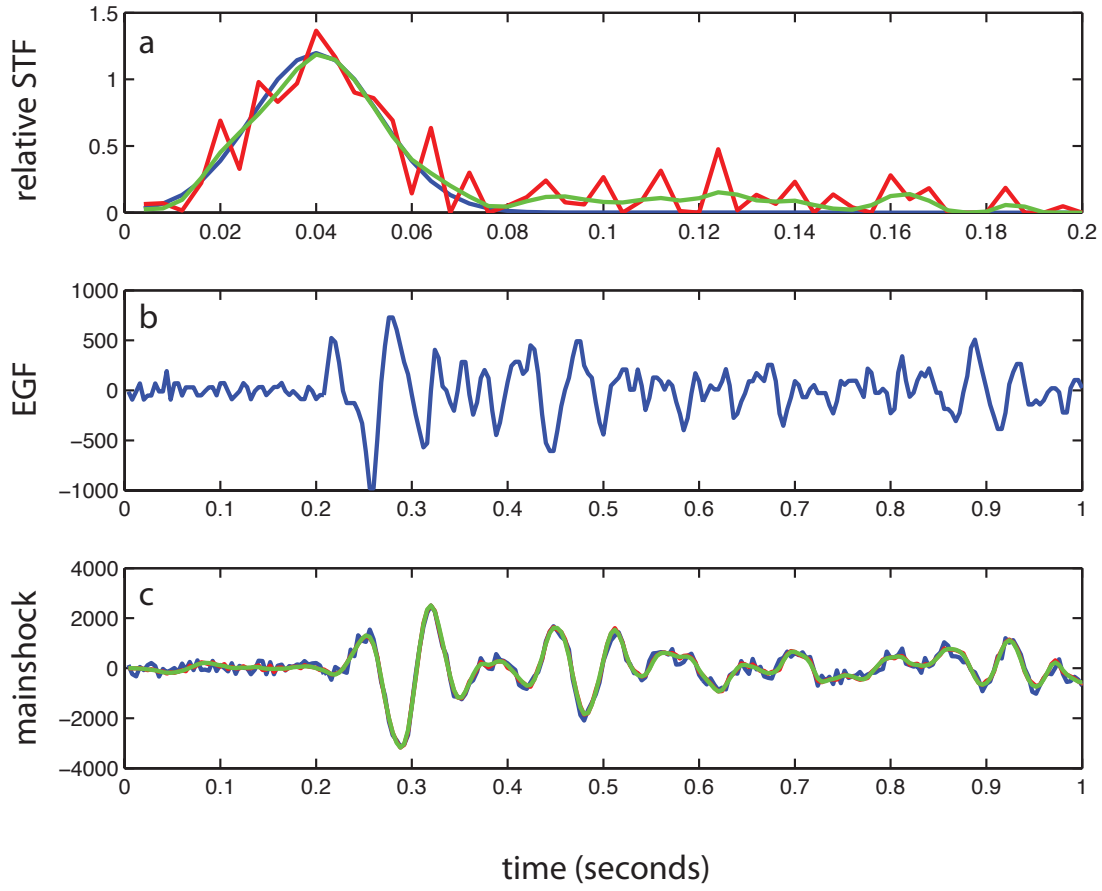

Figure 3.1: Synthetic STF example. In (a), a comparison of the synthetic input STF (blue), the resulting STF from the inversion (red) and the resulting STF using second derivative regularization (green) are shown. The EGF waveform used to generate the synthetic mainshock waveform is shown in (b). The synthetic input mainshock waveform (blue) is plotted in (c), with the waveform fits using the basic inversion (red) and the regularized inversion (green) overlaid for comparison.

simple inversion produces results that fit the input mainshock waveform well, but the STF result is not well matched to the synthetic input due to signal variability introduced with the added noise.

Performing the inversion while minimizing the second derivative of the STF produces a much smoother result with more consistent estimates of STF duration and shape (Figure 3.1) without changing the mainshock waveform RMS misfit value significantly from the unconstrained inversion (change < 0.1%). The smoother STF solution also provides a more consistently bounded pulse and makes it easier to determine the duration of the initial pulse. Increasing the amplitude of the uniform noise added to the synthetic mainshock waveform increases the misfit relatively little. Doubling the amplitude of the noise to a maximum of 20% of the peak waveform amplitude increases the misfit to a mean of 6%, and adding noise with peak amplitudes comparable in size to the input waveform amplitudes produces reasonable waveforms with average misfit of 18%. Increased noise amplitude makes it more difficult to determine consistent estimates of STF duration and shape, however, and this indicates that caution is needed when measuring these parameters in noisy datasets.

An important limitation of this method is imposed by the sampling rate of the data because the resulting STF duration must be an integer number of samples. The minimum duration is two samples, which corresponds to a maximum resolvable corner frequency of ~25 Hz for 100 sps data (0.02 s duration) and ~60 Hz for 250 sps data (0.08 s), when assuming a Brune-type source model and using the relationships between corner frequency and duration described in Mori *et al.* (2003). This limitation acts as a lowpass filter on the waveforms by effectively excluding signals at frequencies above this limit.

Higher sampling rates are necessary to resolve STFs of smaller earthquakes but could also introduce complications due to decreased signal-to-noise ratios at higher frequencies. These limitations are not a problem in the frequency domain because the source spectrum model fitting allows for the equivalent of partial samples, and the somewhat higher Nyquist frequency (50 Hz for 100 sps data and 125 Hz for 250 sps data) permits potentially higher corner frequency estimates for the same data.

We test the inversion method with recorded seismic data using small earthquakes recorded by a small aperture array in southern California (Owens *et al.*, 1991). The waveforms recorded by this array for a given event are very similar at each station due to close station spacing (~7 to 350 m spacing between stations). Because the stations are so closely spaced, any azimuthal variations in the radiation pattern or any effects of rupture directivity will not produce significant variability in waveform similarity between stations. Testing the single record pair case with this dataset is advantageous because we can compare the resulting STFs at each station; similarity among the results verifies that the inversion is behaving as expected. Previous studies using this dataset found that localized geological differences result in a loss of coherence across the array for frequencies above ~15 Hz (Vernon *et al.*, 1998). These localized differences result in mainshock and EGF waveforms that are similar at each station but not identical (Figure 3.2).

Figure 3.2 shows the input waveforms and the estimated STFs at five stations for a M 2.2 mainshock paired with a M 0.9 EGF earthquake; we use P-wave velocity waveforms from 26 available stations in the following analysis. The epicenters of these two earthquakes are separated by 1.8 km and the mainshock is 12.3 km from the array.

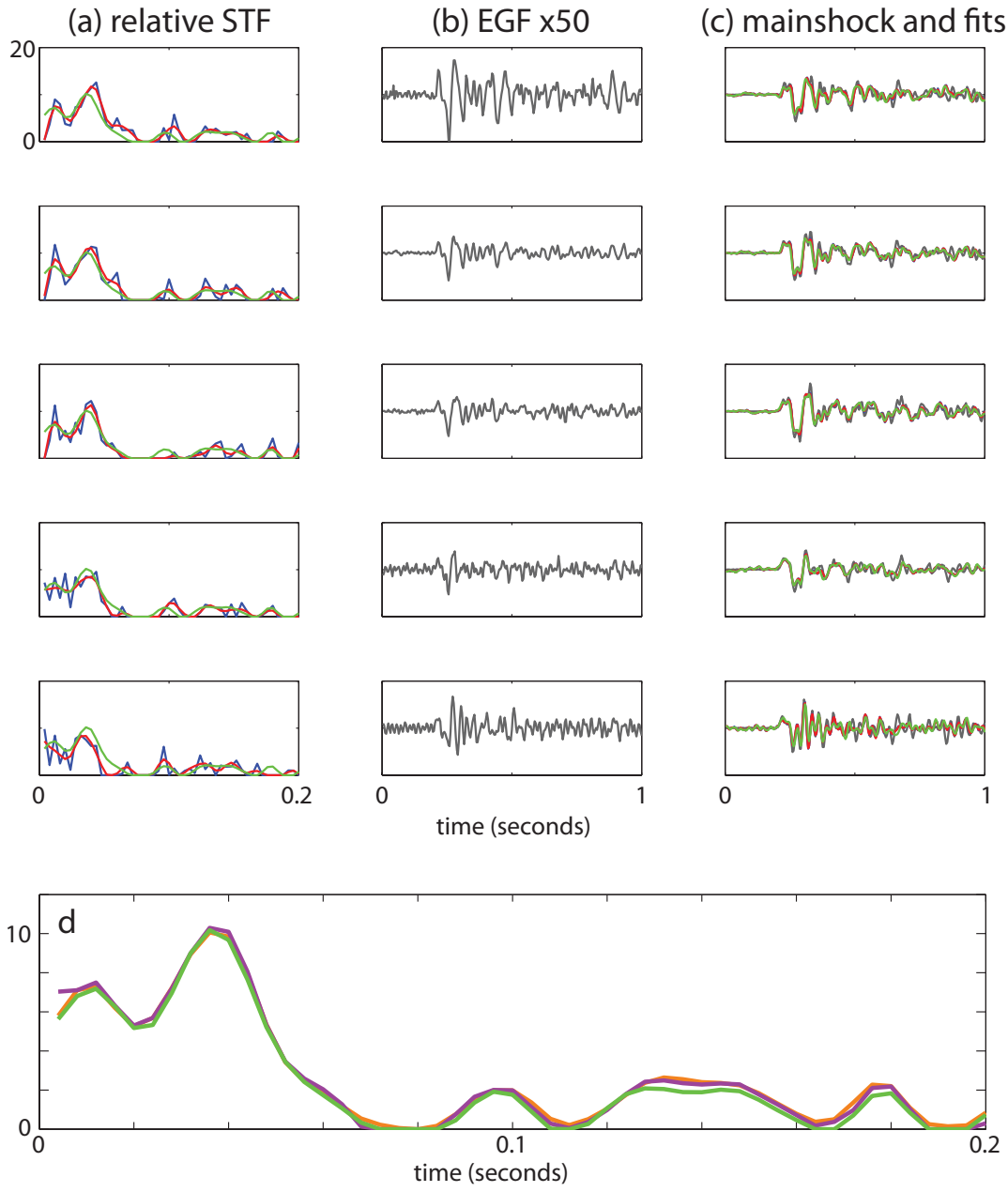

Figure 3.2: Example STF results from the small aperture array data. In (a), a set of STF results is shown for a series of stations. The individual station results without additional regularization are plotted in blue, the results obtained while minimizing the second derivative are plotted in red, and the array STF (identical in all plots in (a)) is plotted in green. In (b), a set of EGF waveforms is shown. Mainshock waveforms (gray) and synthetic fits are plotted in (c), with colors corresponding to STF inversion types from (a). In (d), the mean (orange) and median (purple) individual station STFs are shown along with the array STF result from the simultaneous inversion (green).

Location uncertainties for this dataset are estimated to be at least 1 km due to the limited azimuthal distribution of recording stations. We cannot accurately determine either the interevent spacing between the two earthquakes or the similarity of the focal mechanism orientations to confirm the validity of the EGF method assumptions. In this specialized case, however, any related effects will minimally affect the spatial variability of the STF result across the array and we can neglect such effects for the purpose of establishing how well a similar STF result is resolved at each station.

We compute the STFs separately at each station while minimizing the second derivative (Figure 3.2). The resulting STFs have similar durations and similar shapes at each station across the array, as anticipated, and we confirm that the inversion works well with recorded seismic data. The variability in STF shape is similar to the oscillating spikes observed in the synthetic dataset and the shape does not appear to be entirely constrained using the single stations individually. Measurements of the duration of the initial pulse have a mean value of 16.2 samples (0.065 s) and a standard deviation of 1.7 samples (0.007 s). Analysis of the overall mean or median STF over the array does not produce a comparable duration, however. The shifts in time of individual STF pulses cause the mean and median STF to be much longer than any of the individual results; these demonstrate that computing STFs at individual stations and then stacking the results will not produce a robust result.

3.3.2 Single event pair at multiple stations

We expand this synthetic test dataset to explore the reliability of the inversion method using multiple stations simultaneously. We convolve the synthetic input STF

with a set of EGF records from the small aperture array. When we invert simultaneously for a single STF over all stations in the array, the resulting 'array STF' matches the input STF quite well and averages out the amplitude oscillations present in the single pair inversion (Figure 3.3) without needing to minimize the second derivative. We use 100 trials with varying synthetic STFs and find that the average waveform misfit between the input mainshock waveforms and those created by convolving the array STF solution with each EGF waveform is 3.1%. This is comparable to the misfits obtained using the STFs estimated at each station individually. The stability of the resulting STF shape also greatly improves our ability to effectively measure the rupture duration.

We test the multiple station simultaneous inversion with the small aperture array waveforms from the earthquake pair used in the single pair example. Performing the simultaneous inversion over multiple stations requires precise arrival picks to prevent misaligned waveforms from shifting the STF in time at each station and complicating the array STF results. We use waveform cross-correlation to align waveforms and then assign a single pick to all records for an event prior to performing the inversion.

The resulting array STF for the SAA data exhibits a much smoother shape than the individual station results (Figure 3.2). This confirms that the spikes visible in the individual results are likely complications related to noise. The 0.072 s duration of the STF is also comparable to the values obtained using the individual stations. This similarity confirms that the array STF method produces a more reasonable estimate than the mean or median STFs, which are unable to clearly define the start and end of the source pulse. Waveform comparisons between the recorded mainshock waveforms and

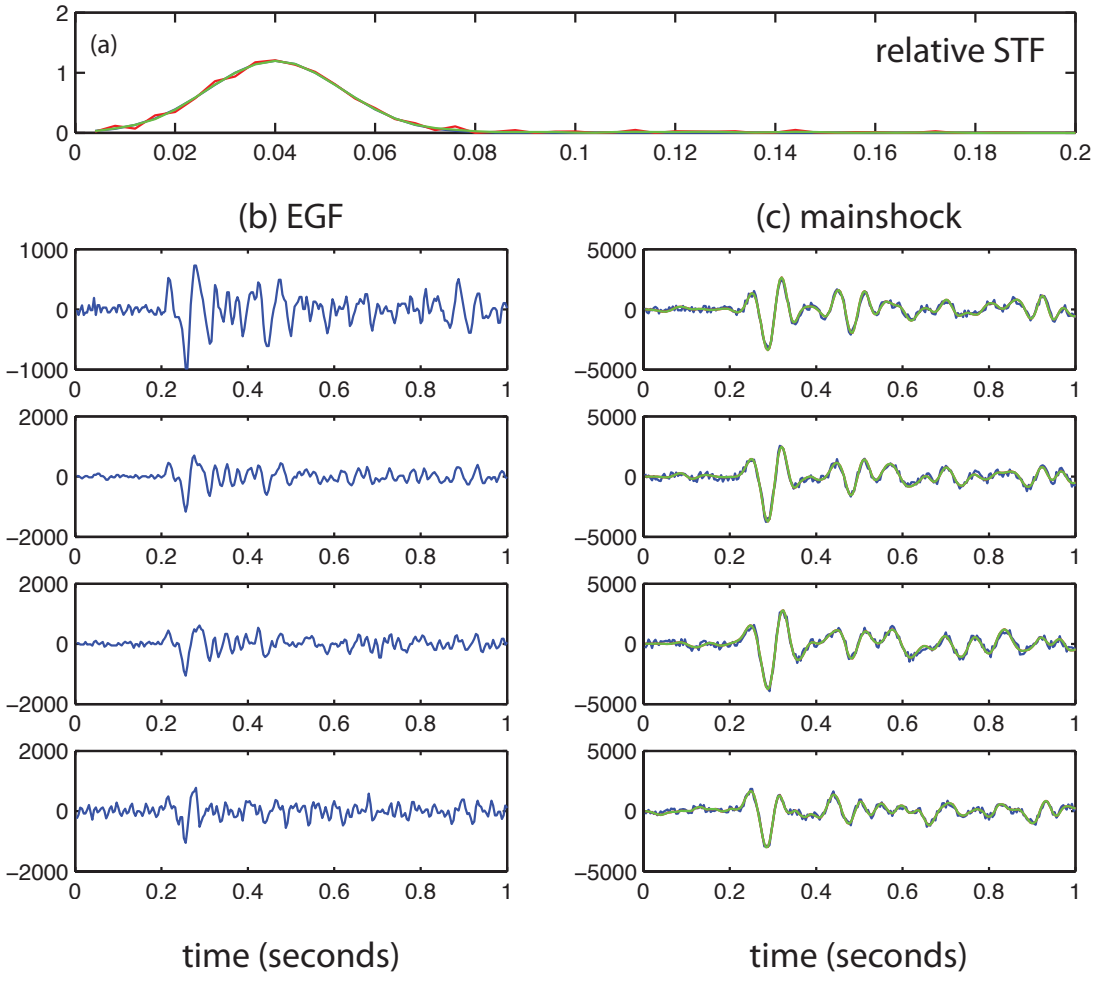

Figure 3.3: Synthetic STF examples for simultaneous inversion. In (a), a comparison of synthetic input STF (blue), the resulting STF from the simultaneous inversion (red) and the resulting STF using second derivative regularization with the simultaneous inversion (green) are shown. A subset of the EGF waveforms used for generating the synthetic mainshock waveforms are shown in (b). A subset of synthetic input mainshock waveforms (blue) are plotted in (c), with the waveform fits using the basic simultaneous inversion (red) and the regularized simultaneous inversion (green) overlaid for comparison.

synthetics computed by convolving the array STF result with the recorded EGF waveforms show that the STF result produces solutions with small misfits.

This method can also be applied to arrays with stations distributed azimuthally around earthquake sources if both the mainshock and EGF earthquake exhibit bilateral rupture. We use a pair of events from the ANZA (southern California) catalog to explore the results of the simultaneous inversion. We choose an event pair from the southern California relocated catalog (Lin *et al.*, 2007) by searching for closely located events. This catalog uses waveform cross-correlation to relocate hypocenters. Events with high degrees of waveform similarity are automatically clustered in space. We choose a mainshock earthquake of M 3.3 and an EGF earthquake of M 1.8. This pair of events has a cataloged hypocentral separation of only 130 m and has particularly highly correlated waveforms at seven out of nine stations (Figure 3.4).

We compute the single station STFs at each station (Figure 3.4). The STFs appear to have similar durations at each station and do not show obvious effects of unilateral rupture directivity that would present as azimuthally-dependent changes in duration. The waveform fits are quite good at seven of the stations, but the two remaining stations exhibit larger misfits.

The simultaneous inversion over all stations produces a relatively simple STF pulse shape with a duration of 0.14 s (Figure 3.4). The waveform misfits using the array STF are slightly higher than the individual STFs, as expected. Waveform fits using the simultaneous array STF show clear differences across the stations in the array; this suggests that some care must be taken in EGF event selection to achieve a robust result even when averaging over multiple stations.

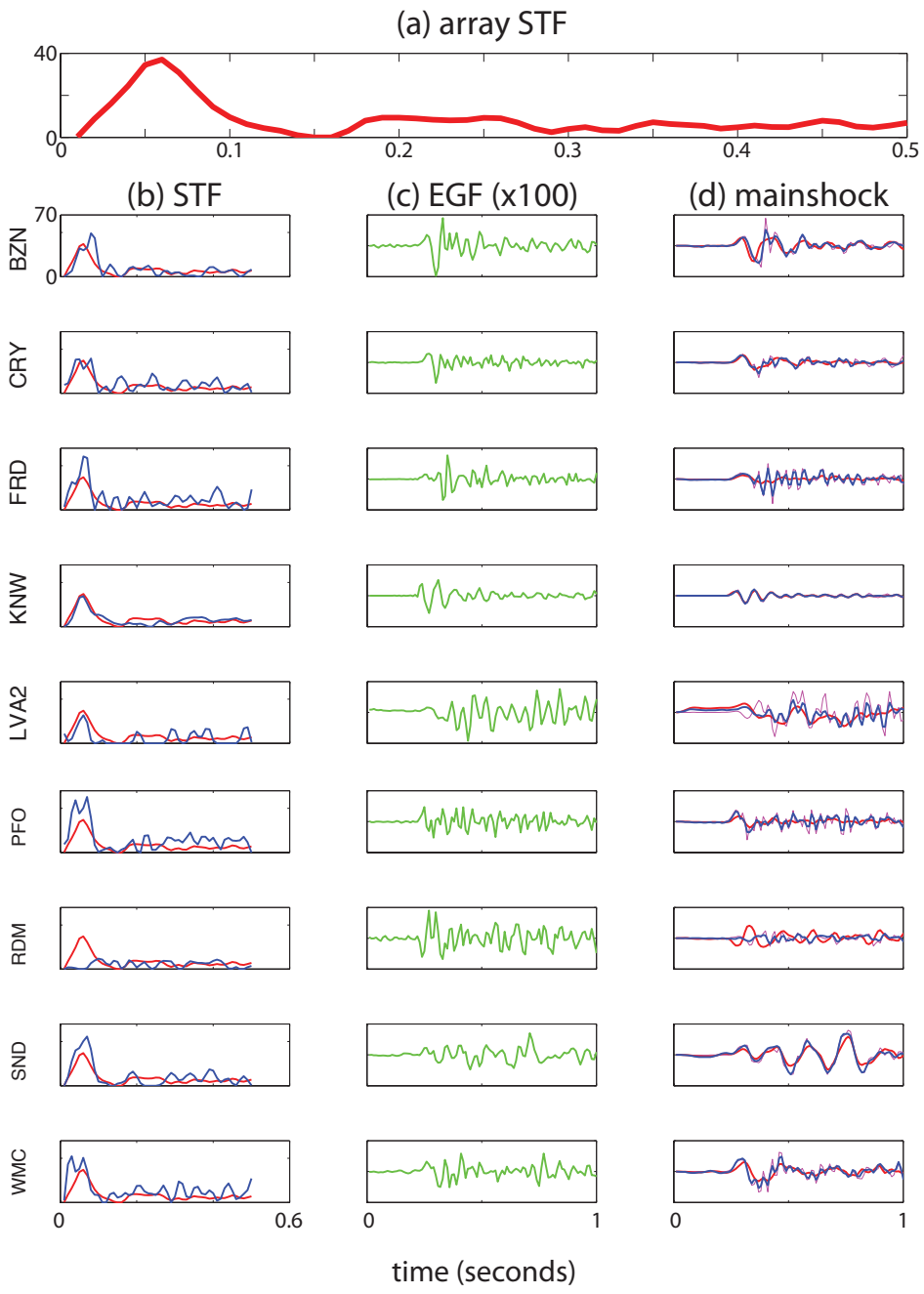

Figure 3.4: STF inversion example for a M 3.3 mainshock earthquake with a M 1.8 EGF earthquake. In (a), the array STF is plotted in red. STF results in (b) show the array STF (red) along with the individual station results. The set of EGF waveforms used is shown in (c), and the mainshock waveforms (magenta) along with waveform fits are shown in (d).

3.3.3 Single mainshock with multiple EGFs at one station

The ability to perform a robust, simultaneous inversion over a set of waveform pairs is additionally possible using a series of EGF events and inverting for an STF result at each station individually. This approach offers the advantages of the simultaneous inversion (*e.g.*, smooth STF solutions and less noisy results) while allowing consideration of sources with more complicated ruptures. Unilateral rupture directivity can be resolved using this approach, whereas the previous simultaneous inversion over a single mainshock/EGF earthquake pair at a series of stations requires assuming a bilateral rupture. We demonstrate this application of the STF inversion method using a M 3.3 earthquake from the ANZA network with a series of M 1.2 to M 1.8 EGF events (Figure 3.5). We show the results of two inversion approaches. First, we invert for the STF simultaneously using each EGF event at all stations, as in the previous example. These results show that the STF solutions have similar shape and amplitude regardless of EGF event selection. Second, we invert for a single STF result for each station using the four EGF events simultaneously, after scaling the EGF events to uniform amplitude to account for variability in EGF magnitude. The series of resulting STFs have somewhat variable durations and amplitudes at the different stations. The synthetic mainshock waveforms created by convolving each of the STF solutions with the EGF waveforms generally match the true mainshock waveform well, although some stations are not as well matched.

Ultimately, this approach offers the ability to consider more complex ruptures (*e.g.*, unilateral rupture directivity) while obtaining relatively smooth STF solutions.

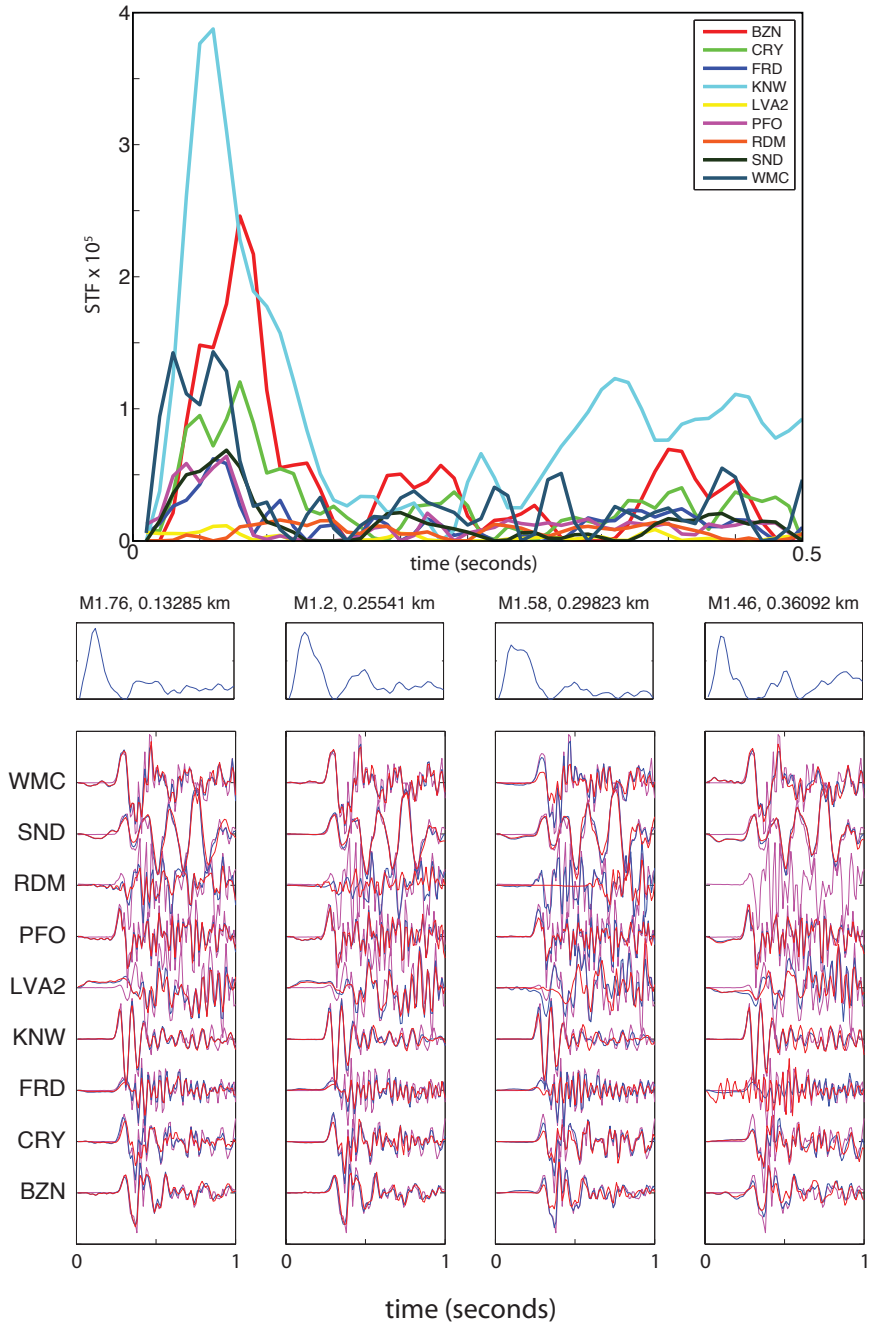

Figure 3.5: Example STF inversion results using multiple EGF events simultaneously. In the top plot, each STF result is obtained using all four EGF events simultaneously at individual stations. In the center, the STF results using all stations simultaneously for each individual EGF are plotted with EGF magnitudes and distance from mainshock hypocenter. Mainshock waveforms (magenta) are shown at the bottom, with waveform fits for individual EGF events in blue and simultaneous EGF at individual stations in red.

However, complications introduced by strong directivity signals in the EGF ruptures, local site effects, and other violations of the EGF assumptions may obfuscate the true source effects.

3.4 Discussion

Using time-domain waveform inversion with EGF events to obtain robust estimates of STFs offers a non-parametric approach to estimating non-parametric source properties. This is potentially advantageous compared to the more common frequency domain methods because such approaches generally require assumption of a rupture model. This method is useful in general single waveform pair mainshock – EGF studies, but has the potential to be of much greater value in situations with many potential EGF events, with ruptures showing effects of unilateral rupture directivity, and with more complex source time function shapes.

Incorporating the minimization of the second derivative of the result (or another regularization scheme not discussed here) helps to smooth the STF result but is not as necessary in situations where multiple waveforms can be used simultaneously (*e.g.*, multiple stations for one EGF event or multiple EGF events for one station). Selection of EGF events remains critical, however, because EGF events that violate the assumptions of the method will contribute spurious signals in the resulting STF.

3.4.1 Uncertainties of estimates

We quantified the variability of duration estimates using synthetic data created from the small aperture array, but such uncertainty estimates represent only the method's variability due to noise and do not represent the true variability measured in estimates of such source parameters. Previous analysis of the small aperture array data quantified the variations of corner frequency estimates across the array of closely spaced stations, and related this variability to localized site variations (Kane *et al.*, 2011). These localized site variations occur on the scale of the station spacing, and cause a loss of signal coherence above ~15 Hz across the array (Vernon *et al.*, 1998).

We repeat the analysis of Kane *et al.* (2011) in this study with estimates of source duration instead of corner frequency. For each mainshock/EGF event pair meeting our selection requirements, we perform the STF inversion while minimizing the second derivative of the result at each station individually. We compute the mean pulse duration across the array, and then determine the percent deviation of each estimate from the array mean. This normalization makes it reasonable to combine percent deviations for all event pairs into a single distribution without needing to further account for differences in event magnitudes. By combining percent deviations for all event pairs into a single distribution, we can measure the variability of the estimates for a larger dataset (Figure 3.6). We measure the width of this distribution using the interquartile range (IQR), which gives the difference between the third and first quartiles of the distribution. The IQR includes half of the total data and hence gives a smaller uncertainty range than the standard deviation.

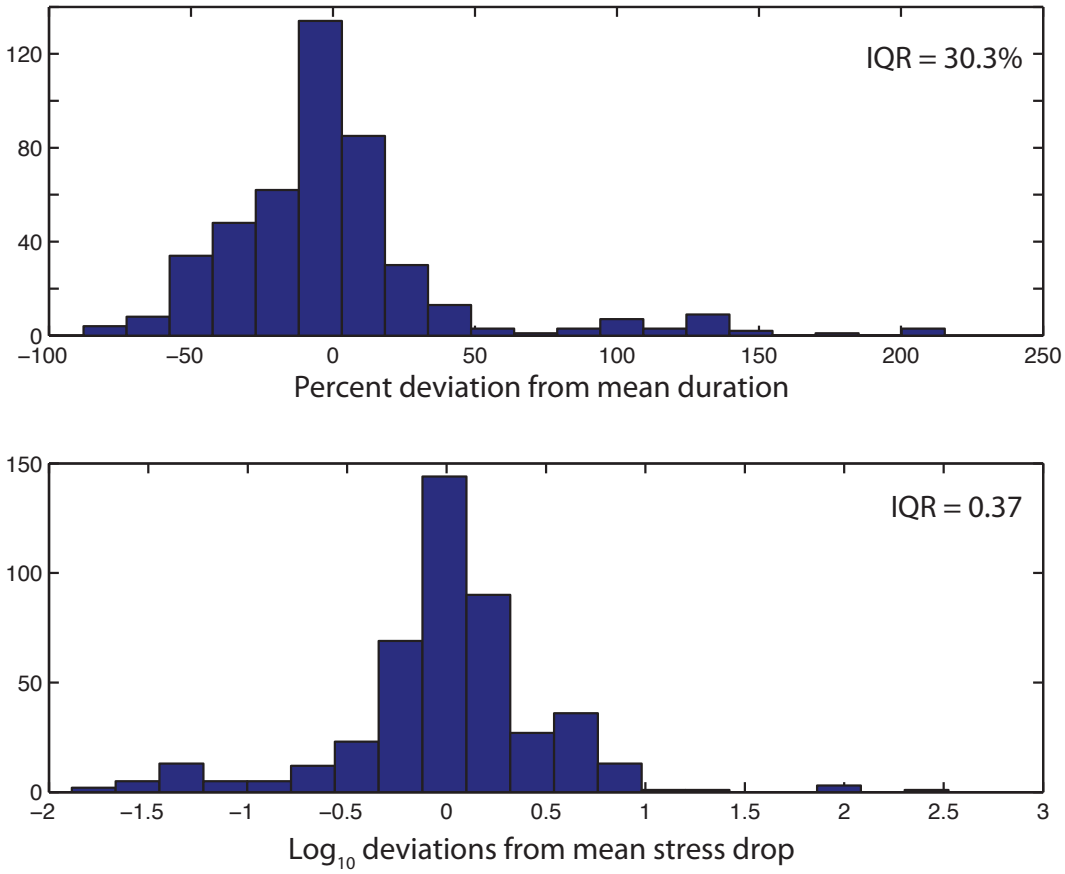

Figure 3.6: Distribution of percent deviations from the mean for duration estimates and distribution of $\log_{10}$ deviations from the mean for stress drop estimates.

The IQR of the duration estimates is 30%, which is somewhat larger than the 23% IQR of the corner frequency estimates (Kane *et al.*, 2011). If the distribution of percent deviations was symmetric, than this IQR could be interpreted as having half of the total data falling within ±15% of the array means.

We compute stress drop for each duration estimate using the relationship between rupture radius and rupture duration given in Mori *et al.* (2003). We repeat the uncertainty analysis in the log-domain using the deviations across the array from the mean stress drop for each event pair, and we measure the IQR of the distribution as 0.37. This is again somewhat larger than the value of 0.31 obtained using the corner frequency estimates. For an estimate of stress drop uncertainty over the array using the duration measurements, this results in error bars of ±0.18 in the log-domain. This corresponds to error bars from 0.7 MPa to 1.5 MPa for a stress drop estimate of 1 MPa, and scales logarithmically to error bars from 7 MPa to 15 MPa for an estimate of 10 MPa. These values are comparable to those measured in the frequency domain.

3.5 Conclusions

We have described a simple method for computing earthquake STFs using matrix inversion and EGF techniques. Our method assumes mainshock and EGF earthquake source similarity, as well as STF nonnegativity, but does not make any further assumptions regarding the source. Synthetic tests and verification with data from a small aperture array and the ANZA network show that this method gives robust results when used to invert waveforms from a pair of events simultaneously over a number of stations. This method gives much smoother results without any further assumptions.

In addition to demonstrating this method with an array of closely spaced stations (thereby ignoring any potential directivity effects and quantifying the measurement uncertainties), we have also demonstrated the inversion's validity for use in larger arrays in the case of bilateral rupture without directivity effects. This method can be extended to use multiple EGF earthquakes with a single mainshock earthquake and to simultaneously invert for the mainshock STF at single stations while smoothing the noise effects. The application of the array STF inversion offers the potential for observing rupture directivity of small earthquakes while minimizing assumptions regarding rupture characteristics.

Acknowledgements

We thank those associated with the collection, analysis, and cataloging of the Small Aperture Array data and the ANZA network data. Early discussions with Glenn Ierley and Bob Parker provided useful feedback. This work, in part, is in preparation for submission to Seismological Research Letters (Kane, D.L. and F.L. Vernon, Application of Time-Domain Deconvolution and Empirical Green's Function Techniques to Obtain Source Time Functions of Small Earthquakes). Deborah Kane is the primary investigator and author of this paper.

References

Bertero, M., D. Bindi, P. Boccacci, M. Cattaneo, C. Eva and V. Lanza (1997), Application of the projected Landweber method to the estimation of the source time function in seismology, *Inverse Problems*, 13, pp. 465–486.

Bertero, M., D. Bindi, P. Boccacci, M. Cattaneo, C. Eva and V. Lanza (1998), A novel blind-deconvolution method with an application to seismology, *Inverse Problems*, 14, pp. 815–833.

Bracewell, R. (2000), The Fourier Transform and Its Applications, McGraw-Hill.

Brune, J. (1970), Tectonic stress and the spectra of seismic shear waves from earthquakes*, J. Geophys. Res.*, 75 (26), pp. 4997-5009.

Frankel, A., and H. Kanamori (1983), Determination of rupture duration and stress drop for earthquakes in southern California, *Bull. Seism. Soc. Am.,* 73 (6), 1527-1551.

Gurrola, H., G.E. Baker and J.B. Minster (1995), Simultaneous time-domain deconvolution with application to the computation of receiver functions, *Geophys. J. Int.*, Vol. 120 (3) pp. 537-543.

Harrington, R.M. and E.E. Brodsky (2009), Source Duration Scales with Magnitude Differently for Earthquakes on the San Andreas Fault and on Secondary Faults in Parkfield, CA, *Bull. Seism. Soc. Am.*, 99, 23-23-2334.

Hartzell, S. (1978), Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.*, 5(1).

Haskell, N. (1964), Total energy and energy spectral density of elastic wave radiation from propagating faults, *Bull. Seism. Soc. Am.*, Vol. 54, No. 6, pp.1811-1851.

Hough, S.E. (1997), Empirical Green's function analysis: taking the next step, *J. Geophys. Res.,* Vol. 102, No. B3, pp. 5369-5384.

Ide, S., G. Beroza, S. Prejean, and W. Ellsworth (2003), Apparent break in earthquake scaling due to path and site effects on deep borehole recordings, *J. Geophys. Res.,* Vol. 108(B5), doi:10.1029/2001JB001617

Kane, D.L., G.A. Prieto, F.L. Vernon, and P.M. Shearer (2011), Quantifying seismic source parameter uncertainties, *Bull. Seism. Soc. Am.*, Vol. 101, No. 2, pp. 535–543, doi:10.1785/0120100166

Lanza, V., D. Spallarossa, M. Cattaneo, D. Bindi, and P. Augliera (1999), Source parameters of small events using constrained deconvolution with empirical Green's functions, *Geophys. J. Int.*, Vol 137, pp.651-662.

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

Madariaga, R. (1976), Dynamics of an expanding circular fault, *Bull. Seism. Soc. Am.*, 66 (3), 639-666.

McGuire, J. (2004), Estimating finite source properties of small earthquake ruptures. *Bull. Seism. Soc. Am.*, 94 (2), 377-393.

Mori, J. and A. Frankel (1990), Source parameters for small events associated with the 1986 North Palm Springs, California, earthquake determined using empirical Green's functions, *Bull. Seism. Soc. Am.*, 80 (2), 278-295.

Mori, J., R. Abercrombie, and H. Kanamori (2003), Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake, *J. Geophys. Res.*, Vol. 108 (B11), 2545, doi:10.1029/2001JB000474

Oth, A., S. Parolai, and D. Bindi (2011), Spectral Analysis of K-NET and KiK-net Data in Japan, Part I: Database Compilation and Peculiarities, *Bull. Seism. Soc. Am.*, Vol. 101, No. 2, pp. 652–666, doi: 10.1785/0120100134.

Owens, T., P. Anderson, and D. McNamara (1991), 1990 Pinyon Flat High Frequency Array Experiment: An IRIS Eurasian Seismic Studies Program Passive Source Experiment, PASSCAL Data Report 91-002 .

Prejean, S.G. and W.L. Ellsworth (2001), Observations of Earthquake Source Parameters at 2 km Depth in the Long Valley Caldera, Eastern California, *Bull. Seism. Soc. Am.*, 91 (2), 165-177.

Prieto, G., P.M. Shearer, F.L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking P and S spectra, *J. Geophys. Res.*, 109 (B08310).

Vallee, M. (2004), Stabilizing the Empirical Green Function Analysis: Development of the Projected Landweber Method, *Bull. Seism. Soc. Am.*, Vol. 94 (2), pp. 394-409.

Velasco, A.A., C.J. Ammon, and T. Lay (1994), Empirical green function deconvolution of broadband surface waves: Rupture directivity of the 1992 Landers, California (Mw= 7.3), earthquake, *Bull. Seism. Soc. Am.*, Vol. 84, No. 3, pp. 735-750.

Vernon, F., G. Pavlis, T. Owens, D. McNamara, and P. Anderson (1998), Near-surface scattering effects observed with a high-frequency phased array at Pinyon Flats, California, *Bull. Seism. Soc. Am.*, 88 (6), 1548-1560.

Chapter 4: Rupture Directivity of Small Earthquakes at Parkfield

We use small earthquakes (M < 5) from Parkfield to look for evidence of consistent rupture directivity along the San Andreas Fault. We analyze azimuthal variations in source spectra after applying an iterative process to correct for wave propagation effects. We avoid using common source spectrum models in our analysis because these models generally assume symmetric rupture; instead, we look for azimuthal variations in the amplitudes of the source spectra over specified frequency bands.

Our results show similar proportions of events exhibiting rupture directivity characteristics towards either the southeast or northwest but spatial variability is apparent. We observe a higher proportion of northwest directivity ruptures following the 2004 M 6 Parkfield earthquake. We also observe a possible preference for southeast rupture directivity among the larger magnitude earthquakes, with 70% of events M > 3 exhibiting southeast rupture characteristics.

4.1 Introduction

Fault rupture during an earthquake does not occur instantaneously. Rupture propagates from a site of nucleation with rupture velocities commonly estimated to be a substantial fraction of the shear wave velocity. When rupture propagates predominately in a single direction from nucleation, the resulting ground motion can be subject to dramatic azimuthal effects (Ben-Menahem, 1961).

Commonly used models for small earthquake sources often assume simple radially symmetric rupture at a constant rupture velocity without allowing for more complex rupture propagation (*e.g.,* expanding circular crack model and variations described by Eshelby, 1957; Brune, 1970; and Madariaga, 1976). However, observations of large earthquakes have shown that ruptures often propagate asymmetrically (McGuire *et al.,* 2002). Several methods have been used to compare seismological estimates of rupture propagation with independently determined hypocenter locations; these studies have established the predominance of unilateral rupture propagation for large earthquakes (*e.g.*, McGuire *et al.,* 2002; Henry and Das, 2001; Mai *et al.,* 2005). Smaller earthquakes, however, have generally been overlooked in such studies due to insufficient seismic data resolution and azimuthal coverage (Boatwright, 2007); hypocenters and rupture propagation patterns cannot be conclusively determined using these same methods. Such smaller events are more often treated with simple models.

One hypothesized mechanism of unilateral rupture propagation is due to the presence of a bimaterial interface at the fault. Theoretical models of strike-slip (mode II) ruptures along a bimaterial interface, where two blocks of different rheological properties

are positioned adjacently due to long-term slip along the fault, feature asymmetric ruptures with a preferred rupture direction controlled by the properties of the two blocks (*e.g.*, Shi and Ben-Zion, 2006). Whether such a model applies to the three-dimensional interfaces of natural faults has yet to be conclusively determined. Some researchers have argued that this model does not apply to natural faults because the geometry of the fault and the localized stress heterogeneities will exert greater control over rupture propagation (Harris and Day, 2005; and references therein). Earthquake relocation studies have produced much more clearly defined faults highlighted by locations of smaller earthquakes, but whether these faults can be adequately modeled by a planar surface or whether a more complex surface is dominant remains unclear.

The San Andreas Fault (SAF) at Parkfield, California, is a close natural approximation to a bimaterial interface. The seismicity along the fault is distributed along a vertical plane (Figure 4.1) and previous studies at Parkfield concluded that the fault represents a distinct, natural barrier between two blocks of different characteristics (Eberhart-Phillips and Michael, 1993; Thurber *et al.,* 2006). Applying the bimaterial interface model to this region of the SAF predicts a preferred rupture direction along the fault strike towards the southeast (Harris and Day, 2005). The SAF is therefore a good candidate fault for testing whether the computational model is applicable to the complex 3-D geometry exhibited in real faults, and for determining if this model offers an explanation for variability in rupture direction. The observation of a preferred rupture direction and the subsequent effects on ground motion would have widespread implications for both earthquake source physics and for earthquake hazard analysis on

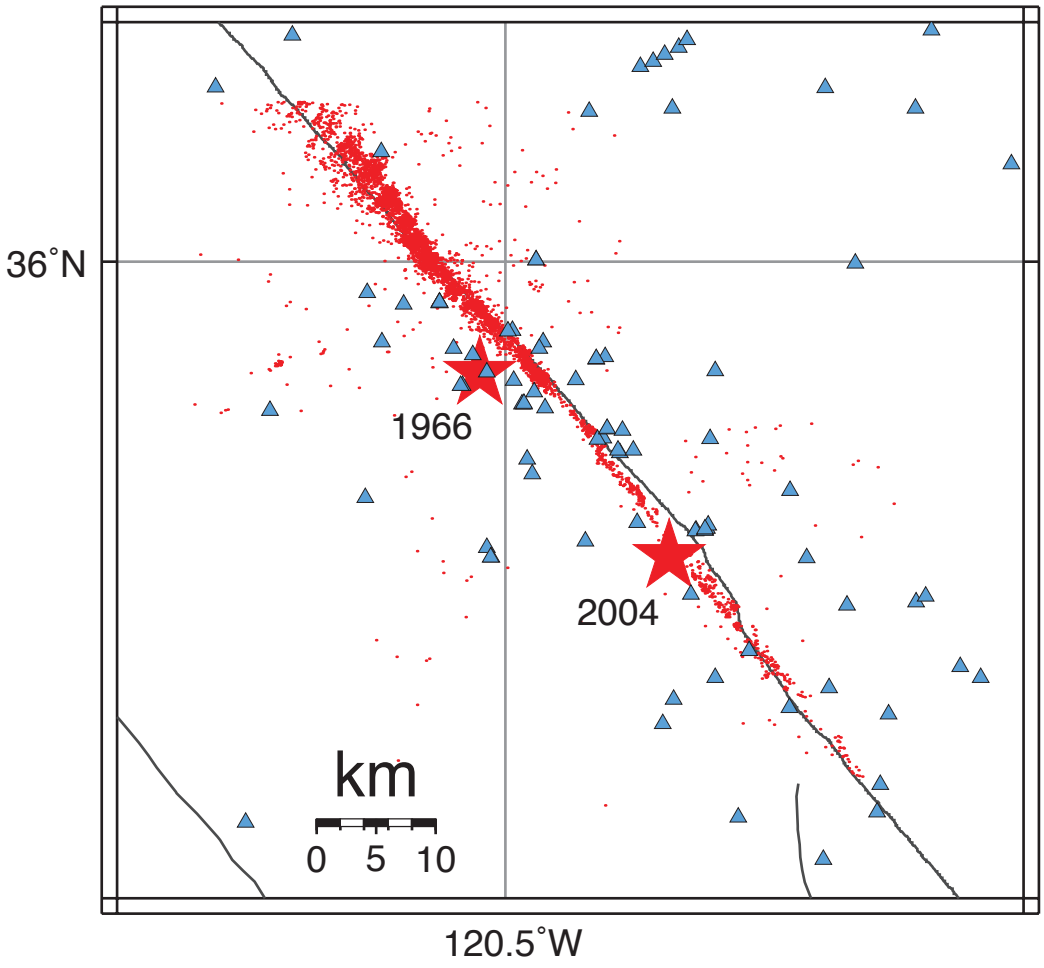

Figure 4.1: Stations (triangles) and earthquakes from 1984-2005 (dots) along the Parkfield segment of the San Andreas fault (SAF) used in this analysis. The locations of the 1966 and 2004 M 6 events are indicated with stars.

mature faults because ground motion amplified by unilateral rupture directivity could produce greater damage than might otherwise be expected.

The SAF at Parkfield had its long-awaited M 6 earthquake in September 2004 after 38 years following the previous 1966 M 6 earthquake. This recent large event at Parkfield sparked considerable interest and debate in the bimaterial interface model. Harris and Day (2005) reviewed studies of prior M 4 to M 6 earthquakes at Parkfield and summarized the rupture directions of this set of events: the two previous M 6 Parkfield earthquakes in 1934 and 1966 ruptured towards the southeast and matched the prediction of the bimaterial interface model, but five M 4 to M 5 events did not match the prediction. The 2004 event nucleated near the southeast end of the Parkfield section and propagated towards the northwest (Fletcher and Spudich, 2004; Langbein *et al.,* 2005; Bakun *et al.,* 2005). Harris and Day (2005, 2006) presented this observational evidence along with a review of numerical simulations and concluded that the bimaterial interface was likely not the primary effect. A subsequent comment by Ben-Zion (2006) attributed the earthquake nucleation to a sub-fault and suggested that the rupture direction matched the preferred direction for this sub-fault. The SAF at Parkfield has been extensively studied, but it is clear that the complexities of the fault either obscure a rupture directivity preference due to smaller scale structure or allow other effects to control rupture direction.

Additional studies of smaller earthquakes at Parkfield have contributed further to this debate through observations of asymmetric distribution of small earthquake locations (Rubin and Gillard*,* 2000; Rubin*,* 2002) and asymmetric aftershock locations (Zaliapin and Ben-Zion, 2011). A recent study (Lengline and Got, 2011) used repeating earthquake

sequences to observe a predominance of southeast rupture propagation in the region of the fault with the largest velocity contrasts across the fault boundary. Another study (Wang and Rubin, 2011) estimated rupture directivity of earthquakes at Parkfield by modeling synthetic spectral ratios and found that of the best-resolved events, ~40% exhibited bilateral rupture characteristics and more than 80% of the remaining unilateral events showed SE rupture directivity.

In this study, we seek to constrain the presence or absence of a preferred rupture direction along the fault by studying smaller earthquakes at Parkfield while minimizing the assumptions in analysis. We look for evidence that small ($M < 5$) earthquakes at Parkfield exhibit characteristics of unilateral rupture directivity along the strike of the fault. We do not use common point-source earthquake models in our analysis because such models generally assume symmetric rupture. Instead, we look for azimuthal differences in amplitudes of displacement spectra to constrain the rupture direction (Figure 4.2). For a model earthquake with unilateral rupture directivity along the fault strike, we expect to see higher spectral amplitudes at higher frequencies in the direction of rupture and lower spectral amplitudes in the opposite direction. We use displacement spectra of events in the region to measure a single scalar value representing these relative spectral amplitudes at frequencies in a specified frequency band. By reducing each earthquake's recorded ground motion to a single value, we can consider the population overall and analyze statistical variations among a larger group of events instead of being limited to analysis of individual events. For events with evidence of unilateral rupture directivity along the fault strike, we then determine whether or not a preferred rupture

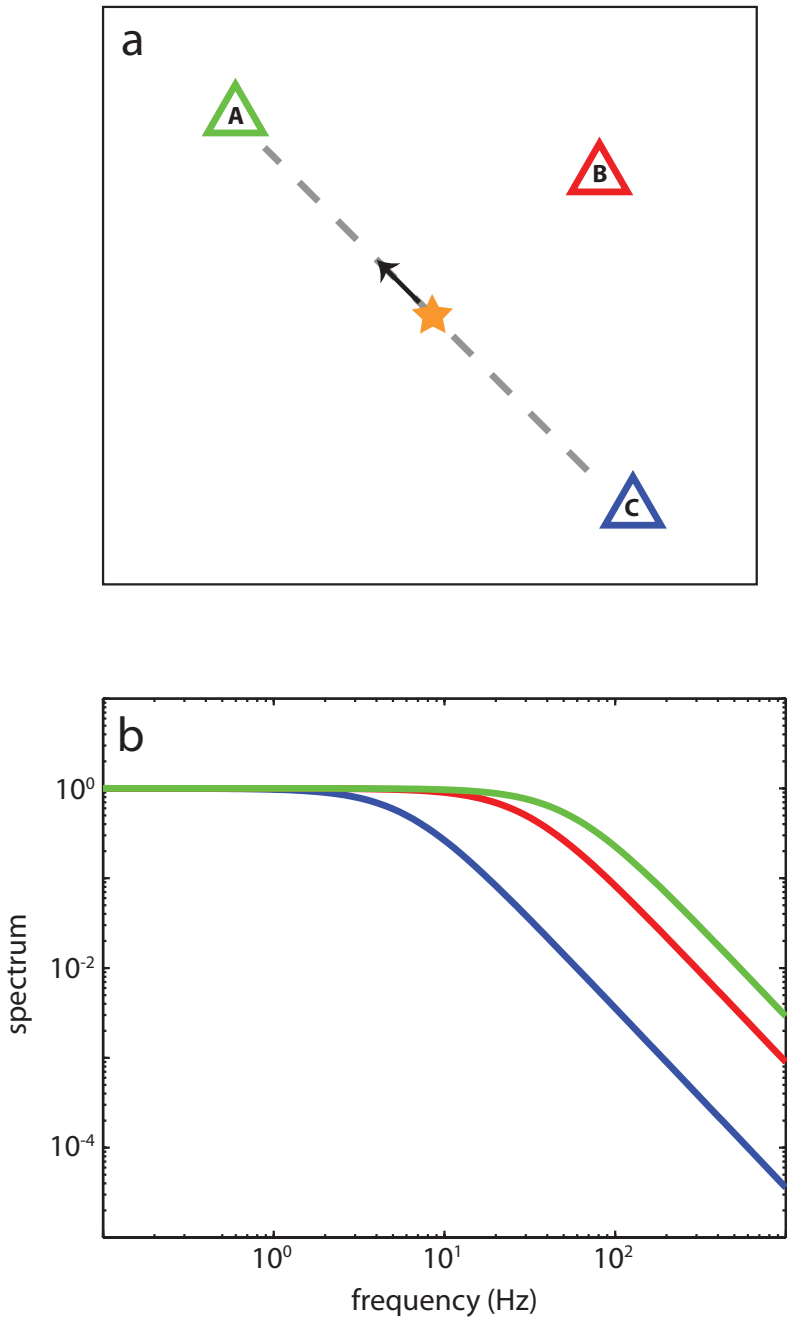

Figure 4.2: (a) Cartoon of earthquake rupture towards the northwest and (b) the resulting azimuthal behavior of the source spectra recorded by stations at various azimuths. Station coloring in (a) indicates corresponding spectrum in (b).

direction is present. Finally, we compare any preferred directions to the direction indicated by the bimaterial interface model.

4.2 Data: The San Andreas Fault at Parkfield

At Parkfield, a distinct section of the SAF is defined by boundaries separating the locked section towards the southeast from the creeping section towards the northwest (Wallace, 1990). The recurrence of M ~6 earthquakes at Parkfield has led to extensive studies of local seismicity and fault structure. Earthquakes at Parkfield generally occur along a vertical plane (Eberhart-Philipps and Michael, 1993; Thurber *et al.,* 2006). Application of various earthquake relocation techniques has produced detailed images of fine fault structure while highlighting relatively shallow (~5 km depth) earthquake clusters broadly distributed to the northwest and distinctly separate bands of seismicity at somewhat greater depths (5 to 12 km) to the southeast (Waldhauser *et al.,* 2004; Thurber *et al.,* 2006).

The fault aligns with a velocity contrast of ~5-20% (Eberhart-Phillips and Michael, 1993; Thurber *et al.,* 2006). Velocity variations along the strike of the fault are also present. Seismicity appears to occur along a single plane at depth. However, two surface traces have been observed and the connection to the structure at depth remains unresolved (Thurber *et al.,* 2006).

We use a dataset of 2263 earthquakes recorded at 108 stations in central California to look for evidence of rupture directivity of small earthquakes on the SAF at Parkfield (Figure 4.1). This waveform dataset is from the Northern California Earthquake

Data Center (NCEDC) and represents a subset of earthquakes recorded from 1984 to 2005. Allmann and Shearer (2007) previously used this dataset to investigate spatial variability in coseismic stress drop prior to and following the 2004 M 6 Parkfield earthquake. We use the multitaper displacement P-wave spectra computed by Allmann and Shearer (2007) over 1.28 s windows for 100 sps data, and we maintain their defined signal-to-noise ratio constraints. The full dataset consists of 31,432 displacement spectra.

4.3 Methods

4.3.1 Iterative separation of source and path spectral contributions

The close spacing of events, combined with a wide range of source-station distances (< 1 km to 100 km) and source-station azimuths in this dataset, present an opportunity to separate the effects of seismic wave propagation from the signal of the seismic sources. We assume that the travel paths between each of two closely located earthquakes and a given station will be approximately identical (Hartzell, 1978). Thus, any effects of scattering and attenuation will also be similar along these paths. This assumption can be extended to a cluster of events with similar hypocentral locations recorded by any station, as long as the separation distance between the sources and a given station is sufficiently large compared to the interevent separation distances of the cluster. We can use these similarities to our advantage by employing an iterative separation process to identify and isolate contributions from the individual earthquake sources and from the path effects. Previous studies have used this technique to determine

average path effects for correcting source spectra (*e.g.*, Warren and Shearer*,* 2002; Prieto *et al.,* 2004; Shearer *et al.,* 2006).

The spatial distribution of events in this region is not sufficiently small to approximate all events as being from a single source location, and we must divide the events spatially to meet the conditions of our assumptions. We use the k-means clustering algorithm (MATLAB R2009B function KMEANS) to split the cataloged events into twenty clusters based on hypocentral locations by minimizing the distance separating a given event hypocenter from the centroid of a cluster of events (Figure 4.3). We initiate the clustering routine with cluster centroids chosen from a trial clustering of a subset of the data. K-means clustering tends to produce clusters of similar size.

We perform the iterative separation of average source, path, and residual effects for each cluster of events following the method described in Prieto *et al.* (2004). This process results in decomposing each log-displacement frequency spectrum, u_{ij}, into three spectral components:

$$u_{ij}(f) = s_i(f) + p_j(f) + r_{ij}(f) \qquad < 4.1 >$$

Here, s_i represents the average source spectrum for event i; p_j represents the average effects due to travel path, local site effects, and instrument response at station j; and r_{ij} represents the remaining residual spectrum. The travel path terms between each cluster and a given station are similar for adjacent clusters. We remove the resulting path terms, p_j, from the recorded displacement spectra, u_{ij}, to correct the records for path effects and leave only the source and residuals for further analysis (hereafter referred to as 'corrected' displacement spectra). These corrected spectra should retain signals due to rupture directivity, unless all events in a cluster rupture in the same direction. In this case,

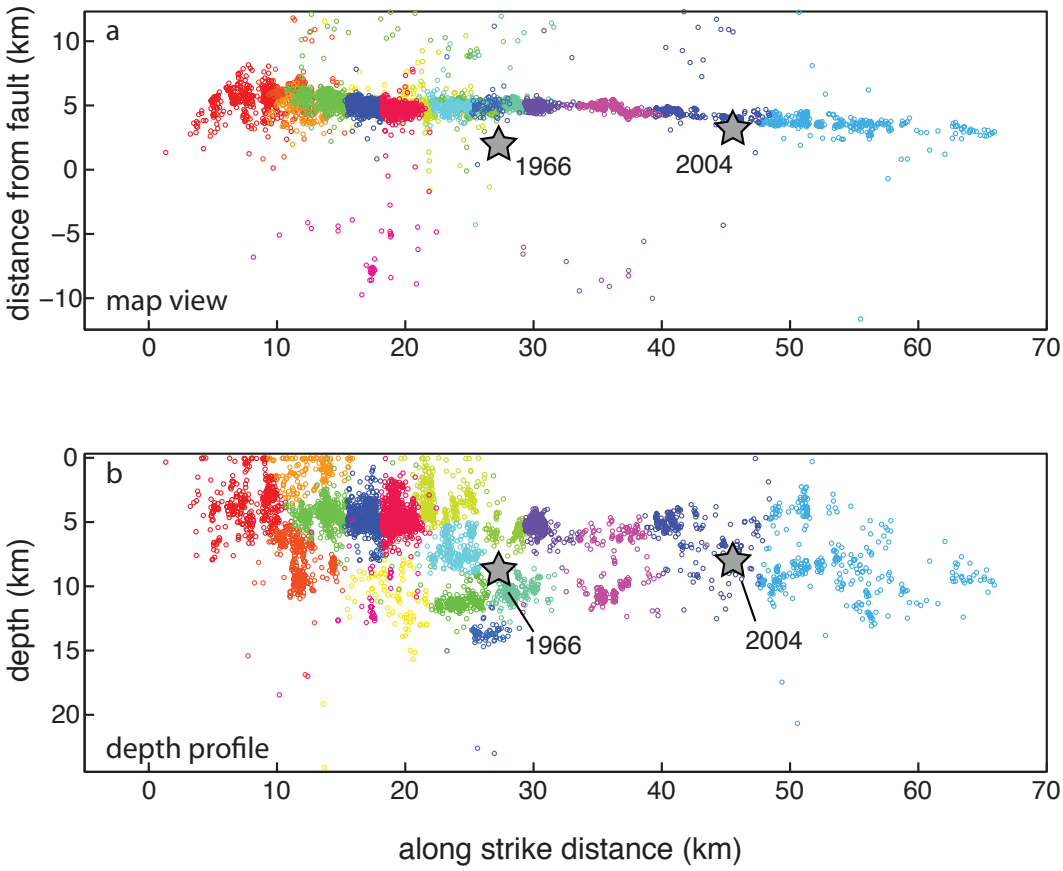

Figure 4.3: (a) Map view and (b) profile view of events considered in this study with spatial clustering indicated by color. The locations of the 1966 and 2004 events are included for reference.

much of this signal would be partitioned into the path term during the iterative separation because it would be similar for all events, and it would be difficult to discern this special case from that of random directivity using this method.

4.3.2 Measuring directivity

Unilateral rupture directivity will produce earthquake source pulses and source spectra that vary with azimuth (Ben-Menahem, 1961). This effect can be observed as shorter duration, higher amplitude source time functions in the direction of rupture, and longer duration, lower amplitude source time functions in the opposite direction. This dependence of duration on azimuth can be formulated using a simple model (*e.g.*, Haskell, 1964) as:

$$\tau(\theta) = \frac{L}{v_r} - \frac{L\cos\theta}{c} = \tau_0\left(1 - \frac{v_r}{c}\cos\theta\right) \qquad < 4.2 >$$

Here, τ represents the duration of the source pulse and θ is the azimuth measured between the direction of rupture and the recording station. τ_0 is equivalent to the rupture length, L, divided by the rupture velocity v_r; this reference source pulse duration corresponds to the case where θ is zero or the rupture is bilateral. The seismic wave velocity is represented by c.

In the frequency domain, this azimuthal variation appears as higher spectral amplitudes at higher frequencies in the direction of rupture and a lack of such high frequency signal in the opposite direction (Figure 4.2). Low frequency amplitudes of source spectra remain unchanged with azimuth. This effect on corner frequency is

inversely related to the effect on source pulse duration and can be described by the azimuthal variation in apparent corner frequency of the source spectrum:

$$f_{c,app} = f_c \frac{1}{1 - \frac{v_r}{c}\cos\theta} \qquad < 4.3 >$$

Here, the apparent corner frequency of the source spectrum $f_{c,app}$, is a function of the true corner frequency, f_c, the rupture velocity, the seismic wave velocity, and the angle between the direction of rupture and the direction of the recording station. If we assume that each earthquake ruptures in a direction along the strike of the fault, then we can determine rupture direction by comparing the source spectra observed in each along-strike direction.

We apply this concept to the Parkfield dataset using the corrected P-wave displacement spectra for each cataloged earthquake. For each event, we select records from stations within a ±45° window of the SAF trace. The ±45° window will include apparent corner frequencies of ~0.75f_c or less in the opposite direction of rupture, and ~1.5f_c or greater in the direction of rupture (assuming a constant rupture velocity of 80% of the S-wave velocity and a P-wave velocity equivalent to $\sqrt{3}$ times the S-wave velocity). This window includes 180° of the total azimuthal directions, and preserves a large portion of the data for further analysis while assuring that several stations are likely to have adequate signal-to-noise ratios for each event. For each azimuthal window (to the southeast and to the northwest), we compute the mean corrected spectrum (Figure 4.4). We quantify the directivity as the log-difference of the mean spectral amplitude to the southeast direction and the mean spectral amplitude to the northwest direction over a specified frequency band (e.g., 20-25 Hz). This processing converts a set of spectra for

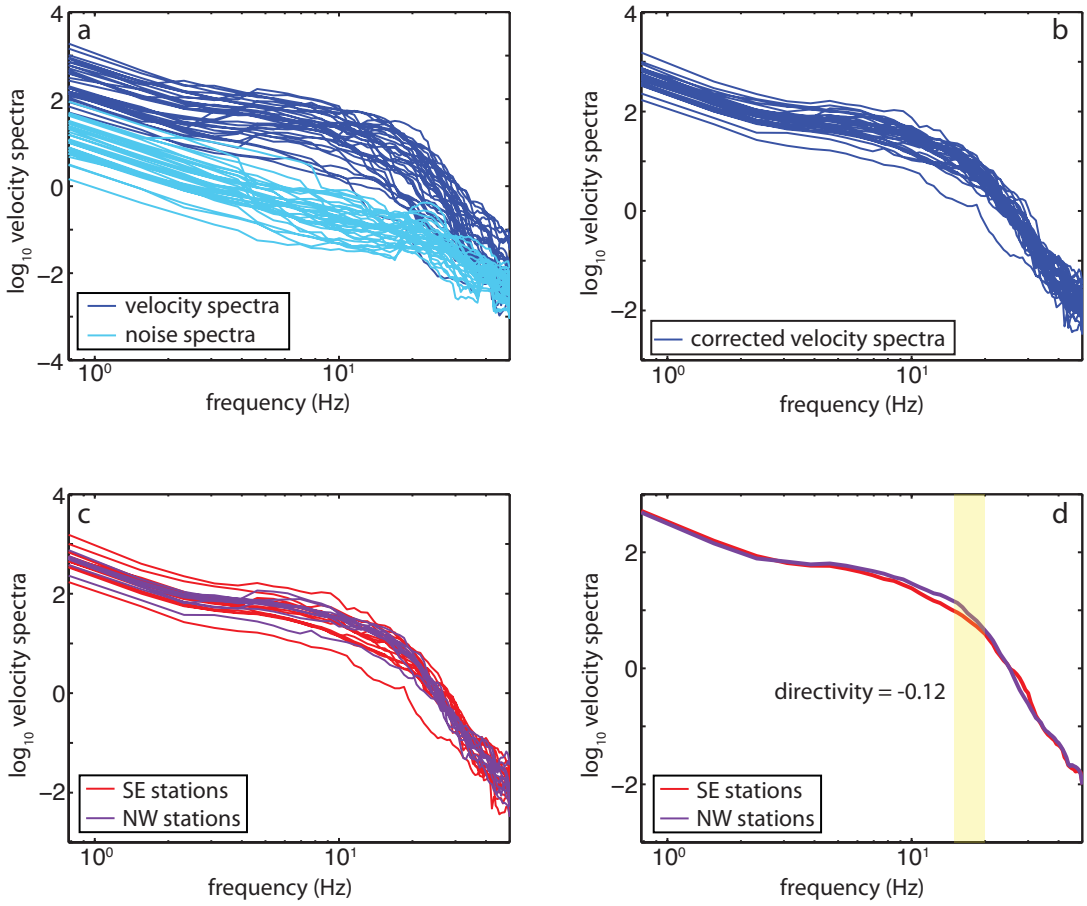

Figure 4.4: Example processing for an event. In (a), recorded displacement amplitude spectra (dark blue) are shown with noise (light blue). In (b), the spectra have been corrected by the iteratively computed path terms. In (c), the spectra from stations falling within the defined azimuthal wedges from the event are color coded, with spectra recorded at stations to the southeast plotted in red and spectra recorded at stations to the northwest plotted in purple. In (d), we plot the mean spectrum in each of the two azimuthal wedges and we highlight the frequency band used to estimate the directivity.

each event into a single scalar value describing the rupture direction based on the assumptions of along-strike unilateral rupture. A positive value indicates higher average spectral amplitudes to the southeast while a negative value indicates higher average spectral amplitudes to the northwest. To obtain uncertainty estimates for these directivity measurements, we use statistical bootstrapping to resample the spectra for each event 100 times with replacement.

This method has a few limitations due to data availability and initial assumptions. We assume unilateral rupture directivity will occur in a horizontal direction aligned with the trace of the fault. Any component of vertical rupture directivity or directivity misaligned from the fault strike will result in values that do not correspond with the true rupture direction. This method may not work well if there are too few records in either azimuthal direction because the averaging of the spectra may not result in sufficiently smooth spectra for measuring mean differences in spectral amplitudes. We limit our study to events with a minimum of three records towards the northwest and three records towards the southeast to minimize such effects.

We use a single frequency band for analysis for earthquakes over a range of magnitudes, and the resulting directivity values will exhibit a dependence on magnitude because the true corner frequencies are related to magnitude (*e.g.*, an event of magnitude 4 will have a higher directivity value than an event of magnitude 3 if all other conditions are equal because the true corner frequency of the M 4 event will be much lower than the band used to measure directivity). This effect can make it difficult to choose an appropriate frequency band to use for all events because the band needs to be beyond the corner frequencies of the events in order to resolve a separation of the spectra in each

direction while staying within the limits of adequate signal-to-noise ratios. Variations in rupture velocity can also affect the resulting measurements because lower rupture velocities will produce smaller variations in corner frequency and directivity estimates at a given frequency band than would be expected at higher rupture velocities.

Our method measures mean values of spectra above the event corner frequencies, and these spectra often have lower signal-to-noise ratios in these bands. We confirm that the directivity value results are consistent across several frequency bands and not dependent on the band chosen for analysis by repeating our analysis using three different frequency bands of 5 Hz width (15 to 20 Hz, 20 to 25 Hz, and 25 to 30 Hz) and verifying that the results are consistent. In the following results, we focus on the 15 to 20 Hz band.

4.4 Results

Of the 2263 events in our dataset, 839 events met the processing criteria. Overall, this set of 839 earthquakes has similar proportions of events with directivity estimates towards the southeast and northwest. We observe 424 (51%) exhibiting rupture towards the southeast and 415 (49%) exhibiting rupture towards the northwest.

The applied path correction between each cluster and each station removes the average propagation effects for the events in the cluster, but it does not account for the differences in event location within a given cluster. We fit a simple one-dimensional along-strike model to the directivity results in each cluster to estimate a constant attenuation term (Q_P) for within the cluster. We do not estimate a frequency-dependent Q_P because the directivity results are measured over a narrow frequency band. Not all clusters show effects of local attenuation, and the estimates of Q_P range from ~70 to

~3000. Most Q_P values are consistent with the 3-D Q_P model presented by Bennington *et al.* (2008); the higher values describe clusters with minimal trend in directivity estimates within the cluster. We apply the resulting attenuation corrections to the clusters. These corrections do not change the resulting rupture directivity in any of the events.

We define a subset of our results for further analysis by selecting events in which the directivity to the northwest or southeast is significant at the 90% confidence level based on the uncertainty limits obtained with the bootstrap technique. Of these 378 significant events (45% of the total 839 directivity estimates) we find that 183 events (48%) exhibit SE rupture and 195 events (52%) exhibit NW rupture. These proportions change minimally if we decrease the significance cutoff to the 85% confidence level or if we increase the cutoff to the 95% confidence level, confirming that our results are robust. These results indicate that a preferred rupture direction of smaller events is not a dominant effect overall in the seismogenic zone at Parkfield. We do not observe a strong preference for rupture direction in this dataset, but it is necessary to further investigate characteristics of these events to find or rule out directivity on smaller spatial scales or over subsets of the data.

4.4.1 Effect of location

Figure 4.5 displays a map and profile view of the locations of these 378 significant events. Spatial variability in rupture direction is apparent, and rupture direction seems to show a preference within small groupings of event locations. These

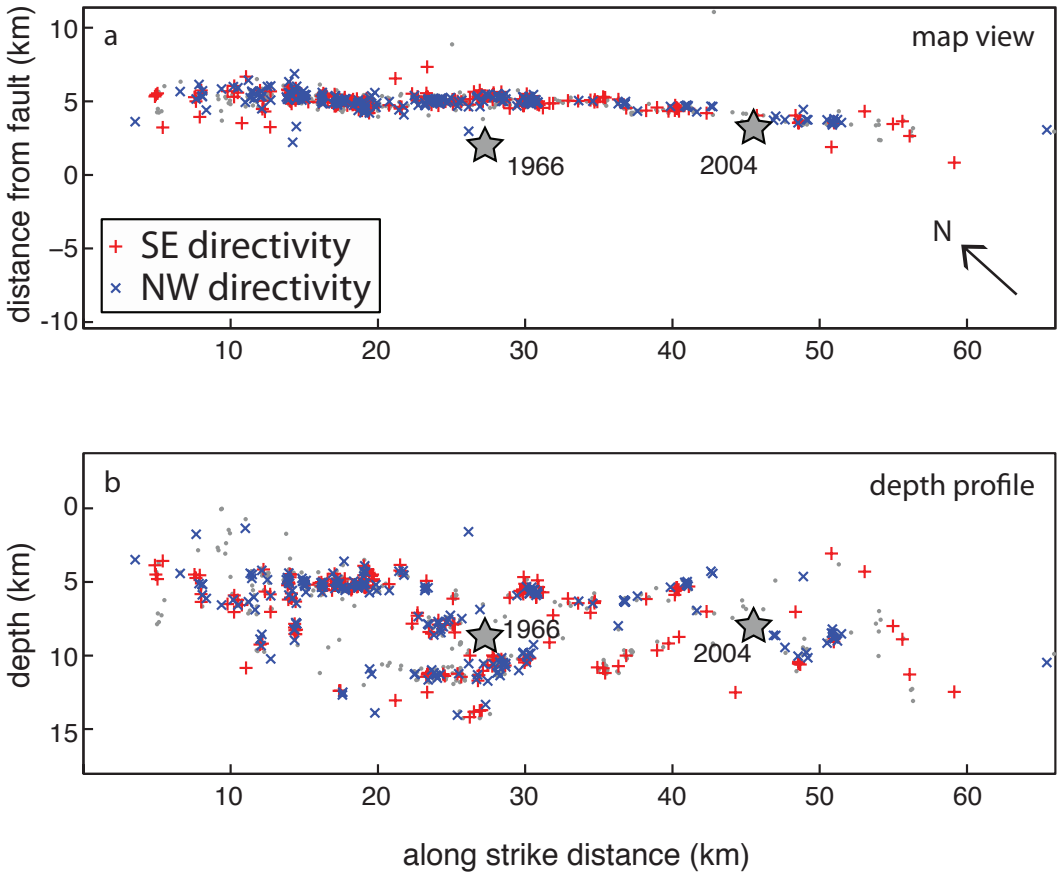

Figure 4.5: Map (a) and profile (b) views showing rupture directivity results. Events without significant results are indicated with a gray dot. Events with significant results are plotted as red '+' symbols (SE rupture) and as blue 'o' symbols (NW rupture).

variations may represent true spatial differences in directivity properties related to smaller scale rheological or stress variations along the fault.

To confirm any true spatial variability in event rupture directivity, we first establish the independence of our results from the initial subgroups obtained in the clustering process. We investigate this visually by looking at the directivity results for individual clusters of events and comparing these with adjacent clusters. Within individual clusters, the directivity of events appears to be spatially grouped rather than randomly distributed. Because path corrections are uniformly applied to all events in each cluster and the interevent spacing differences are accounted for by the Q_P correction, such a signal is due to either inappropriate path corrections for some events or due to true variability in rupture directivity.

We perform two simple tests to verify the independence of our results from the clustering process. First, we repeat the analysis using thirty smaller clusters instead of the initial twenty. We find that the results of the overall dataset change only minimally by decreasing the size of the clusters. Second, for two sets of adjacent clusters, we combine the adjacent clusters into a single larger cluster and repeat the analysis. We find that of the events processed in both the original twenty cluster analysis and in the combined clusters analysis, no events show a change in rupture directivity result. These tests demonstrate that the spatial variability we observe in the rupture directivity results is not due to the clustering process and could represent true spatial variability in preferred rupture directivity.

4.4.2 Effect of the 2004 M 6 Parkfield earthquake

We compare the data from events prior to the 2004 M 6 earthquake with the data from events following the mainshock to determine if any rupture directivity preference exists that may be related to the interseismic period or to the aftershock sequence (Figure 4.6). The events prior to the 2004 earthquake exhibit a distribution of rupture direction similar to the overall dataset, with 49% of events rupturing towards southeast. Notably, the events following the 2004 earthquake exhibit a stronger preference for NW directivity, with only 38% of events rupturing towards southeast and the remaining 62% towards northwest. The aftershocks seem to preferentially match the rupture direction of the mainshock, not unlike comparisons of other mainshock source properties (*e.g.*, focal mechanism orientation) with those of aftershocks.

4.4.3 Effect of earthquake magnitude

Our method of measuring rupture directivity includes an inherent bias with event magnitude due to the limitations of using a single frequency band for all events. Smaller magnitude events will have higher spectral corner frequencies than larger magnitude events, and a smaller difference between SE and NW station spectra will subsequently occur. Because of this bias, it is necessary to consider the effects of earthquake magnitude on the results of the directivity estimates and to determine if any such effect is due to the measurement bias or is a true difference among earthquake rupture size.

Larger earthquakes are more likely to produce directivity estimates in our analysis because these events are generally recorded by more stations at greater distances and at

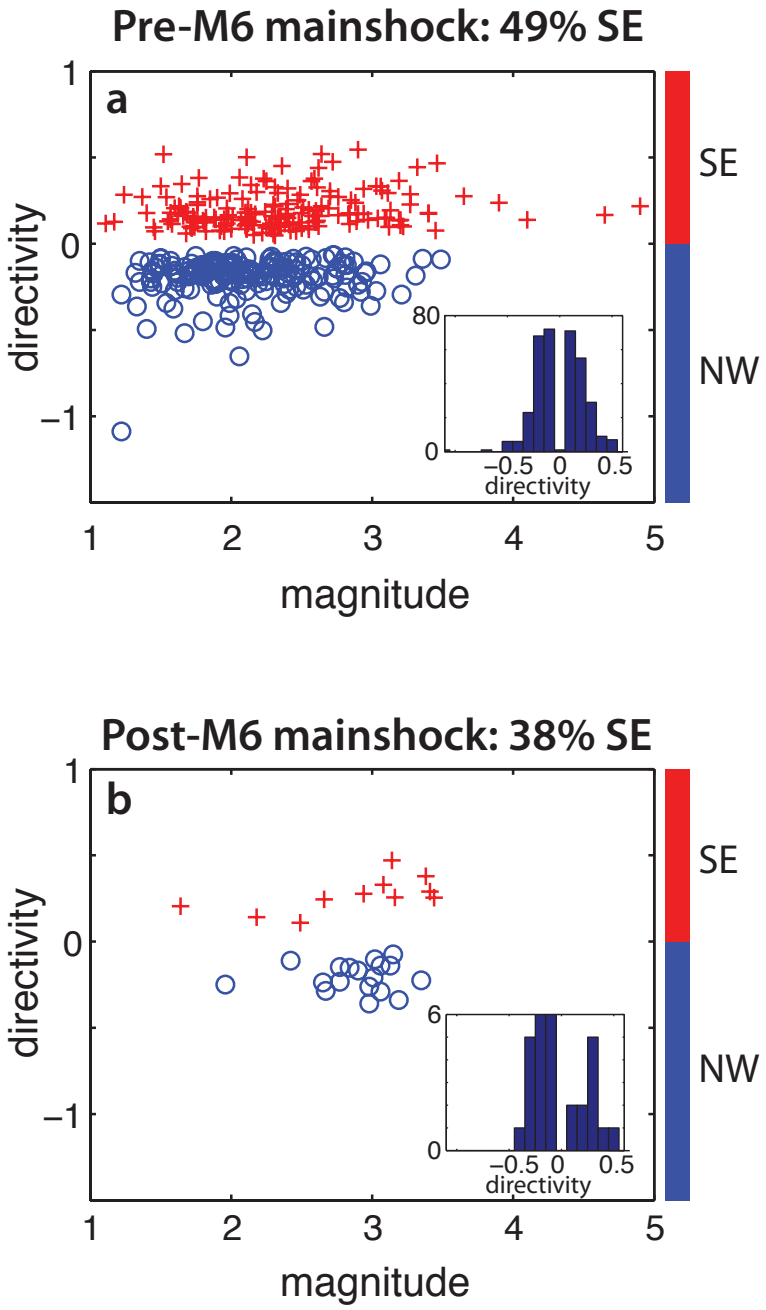

Figure 4.6: Comparison of directivity estimates for events prior to the 2004 M 6 mainshock with those following the mainshock; estimates are plotted by magnitude. Inset histograms show the overall distribution of estimates in each population. We observe a temporal change in the proportion of events with northwest rupture directivity.

higher signal to noise ratios than comparably located smaller earthquakes. The greater rupture area of larger magnitude events will also produce an effectively lower corner frequency in the source spectrum as compared to a smaller earthquake, and this effect will make the difference in spectral amplitudes greater at a given higher frequency than can be observed for a smaller earthquake. The effect of such difference is to obtain a higher number of significant directivity estimates for larger earthquakes.

In Figure 4.7, we again show the profile plot of events with significant results. We overlay this plot with a separate representation showing the rupture direction for events of magnitude greater than M 3. We observe a clear discrepancy in rupture directivity for events at these magnitudes. Instead of comparable proportions of events with rupture in either direction, we observe 70% of events with M ≥ 3 exhibiting rupture towards the southeast and only 30% with rupture towards the northwest. This subset of data contains 46 events. Fourteen of these events occurred in the aftershock sequence of the 2004 M 6 Parkfield earthquake. For the events following the 2004 earthquake, 6 out of 14 exhibit rupture towards southeast and the remaining 8 exhibit rupture towards northwest.

4.5 Discussion and Conclusions

We use a simple comparison of displacement spectra to estimate rupture directivity of small earthquakes at Parkfield. This method is best suited for cases of unilateral rupture directivity along the strike of the fault towards the northwest or the southeast. It is somewhat limited in its ability to resolve rupture directivity because it cannot differentiate between bilateral rupture and events without significant rupture directivity results. The lack of availability of data with adequate frequency content and

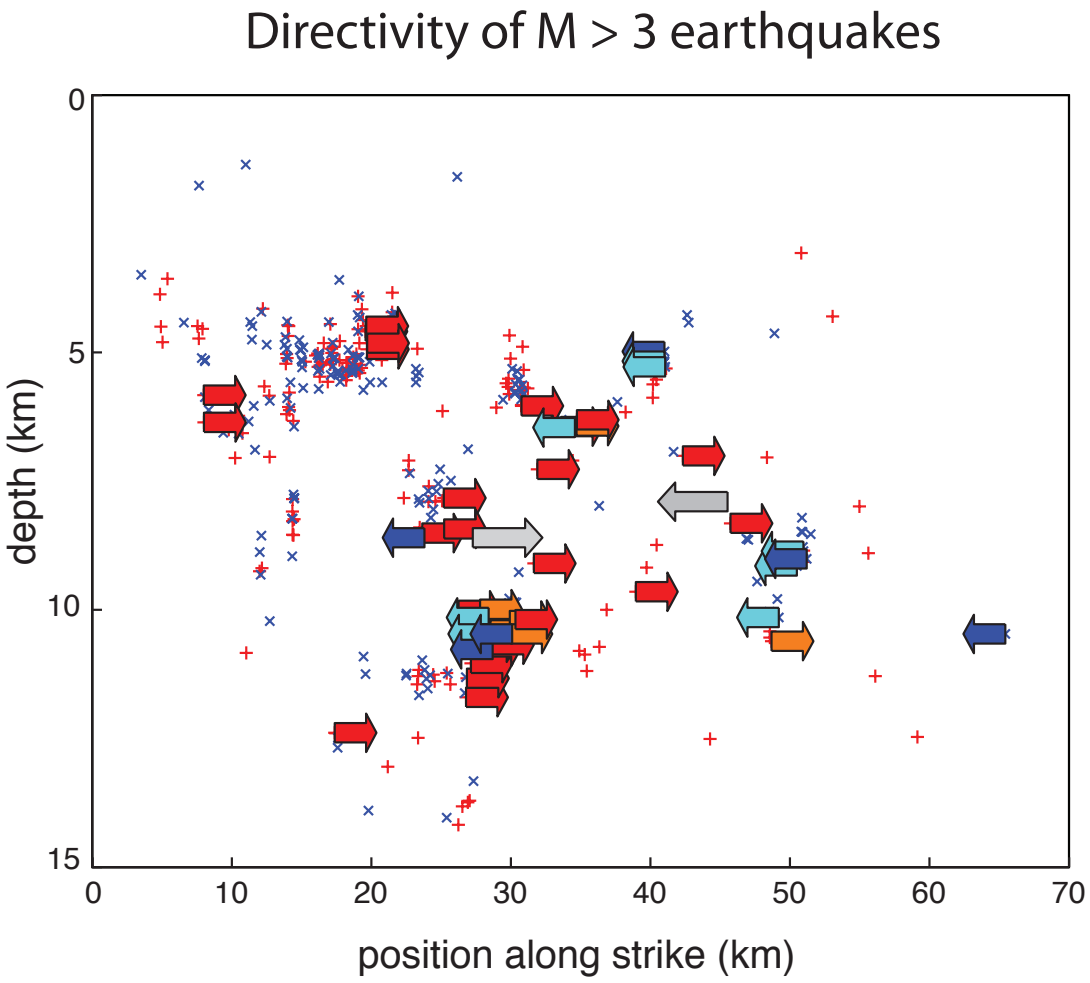

Figure 4.7: Profile view showing rupture direction of events with significant directivity results. Red '+' symbols indicate southeast rupture while blue 'x' symbols indicate northwest rupture. We overlay a set of arrows indicating rupture directivity of events with M > 3. Lighter shaded arrows indicate events in the 2004 aftershock sequence, and the larger gray arrows mark the location and the rupture direction of the 1966 and 2004 mainshocks (not analyzed in this study).

geographic distribution of stations can also obscure rupture directivity effects within the overall dataset. Although the effects of these limitations are important to consider for rupture directivity estimates of individual events, we expect this simple method to be sufficient when working with a large set of data.

Our analysis considers 2263 events at Parkfield and produces 839 directivity estimates after removing events with insufficient azimuthal data coverage, applying corrections for propagation paths, and applying corrections for attenuation within each cluster. 378 of these results are significant at the 90% confidence level. We do not observe a strong preference for rupture directivity towards either the southeast or towards the northwest for this subset of results with significant estimates. The bimaterial model predicts a preference for southeast rupture directivity when assuming that the interface divides two blocks of different rheological properties. However, this simple view of the model does not generally account for the variability of rheological properties on the scales observed in tomographic studies (*e.g.*, Thurber *et al.,* 2006). Our relatively simple analysis hides important details in the full dataset. For example, we observe variability in rupture directivity patterns within spatial groupings of events which could correspond to smaller scale rheological or stress variability. The 461 results that do not meet the significance criterion may correspond to ruptures with bilateral or vertical rupture characteristics, or may not have strong enough directivity to be resolved using this method. A recent study of earthquakes at Parkfield performed finite source inversions for a set of small earthquakes and found sizeable vertical rupture components for some ruptures (Uchide and Ide, 2010); our method cannot resolve these features.

Although the proportion of ruptures in either direction is similar when considering the dataset overall, we observe changes in rupture directivity related to the 2004 M 6 Parkfield. The aftershocks of the 2004 mainshock contain a higher proportion of events with rupture towards the northwest; this directivity matches the rupture directivity of the mainshock.

We observe a greater proportion of events with southeast rupture directivity when we limit the dataset to the largest earthquakes ($M > 3$). In this subset of data, 70% of events exhibited southeast directivity. This proportion differs within the 2004 aftershock sequence, with only 43% of these $M > 3$ earthquakes rupturing towards the southeast. It is of interest to note that this differs from the proportion of $M > 4$ events with southeast rupture cataloged by Harris and Day (2005).

The overall results of our analysis are highly variable: about half of the estimates do not produce significant values of directivity (indicating possible vertical or bilateral ruptures), and the remaining half are split between southeast ruptures and northwest ruptures. The degree of variability in apparent rupture direction in a region of relatively simple geometric fault configuration suggests that there may not be a strong preference for a particular rupture direction overall. However, given the observed spatial variability in our results, the possibility of smaller scale controls on rupture directivity due to fault geometry, presence of bimaterial interfaces, or heterogeneous stress distribution cannot be ruled out. The results of this study and other analyses of rupture directivity of earthquakes at Parkfield (*e.g.*, Harris and Day, 2005; Lengline and Got, 2011; Wang and Rubin, 2011) indicate that rupture directivity is likely due to several contributing factors

and that predicting the rupture direction of future earthquakes will be difficult even in regions that appear to have geometrically simple faulting.

Acknowledgements

We thank those associated with collection, analysis, and cataloging of Parkfield seismicity. Discussions with Yehuda Ben-Zion, Allan Rubin, and Enning Wang were helpful in completing this work. This work (in full) is being prepared for submission to the Journal of Geophysical Research (Kane, D.L., P.M. Shearer, B. Goertz-Allmann, and F.L. Vernon, Rupture Directivity of Small Earthquakes at Parkfield). Deborah Kane is the primary investigator and author of this paper.

References

Allmann, B., and P. Shearer (2007), Spatial and temporal stress drop variations in small earthquakes near Parkfield, California, *J. Geophys. Res*., Vol. 112, B04305, doi:10.1029/2006JB004395.

Bakun, W.H., B. Aagaard, B. Dost, W.L. Ellsworth, J.L. Hardebeck, R.A. Harris, C. Ji, M.J.S. Johnston, J. Langbein, J.J. Lienkaemper, A.J. Michael, J.R. Murray, R.M. Nadeau, P.A. Reasenberg, M.S. Reichle, E.A. Roeloffs, A. Shakal, R.W. Simpson, and F. Waldhauser (2005), Implications for prediction and hazard assessment from the 2004 Parkfield earthquake, *Nature*, Vol. 437, doi:10.1038/nature04067.

Ben-Menahem, A. (1961), Radiation of seismic surface-waves from finite moving sources. *Bull. Seism. Soc. Am.*, Vol. 51, No. 3, pp. 401-435.

Ben-Zion, Y. (2006), Comment on "Material contrast does not predict earthquake rupture propagation direction" by R.A. Harris and S.M. Day, *Geophys. Res. Lett.*, Vol. 33, L13310, doi:10.1029/2005GL025652

Bennington, N., C. Thurber, and S. Roecker (2008), Three-Dimensional Seismic Attenuation Structure around the SAFOD Site, Parkfield, California, *Bull. Seism. Soc. Am.*, Vol. 98 (6), 2934-2947.

Boatwright, J. (2007), The persistence of directivity in small earthquakes, *Bull. Seism. Soc. Am.*, Vol. 97, No. 6, pp. 1850–1861, doi: 10.1785/0120050228.

Brune, J. (1970)., Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, 75 (26), 4997-5009.

Eberhart-Phillips, D., and A. Michael (1993), Three-Dimensional velocity structure, seismicity, and fault structure in the Parkfield region, central California, *J. Geophys. Res.*, Vol. 98, No. B9, pp.15,737-15,758,

Eshelby, J. (1957), The determination of the elastic field of an ellipsoidal inclusion, and related problems, *Proc. Roy. Soc. Lond. Math. Phys. Sci.*, 241 (1226), 376-396.

Fletcher, J.B., P. Spudich, and L.M. Baker (2006), Rupture Propagation of the 2004 Parkfield, California, Earthquake from Observations at the UPSAR, *Bull. Seism. Soc. Am.*, Vol. 96, No. 4B, pp. S129–S142, doi: 10.1785/0120050812

Harris, R.A., and S.M. Day (2005), Material contrast does not predict earthquake rupture propagation direction, *Geophys. Res. Lett.*, Vol. 32, L23301, doi:10.1029/2005GL023941

Harris, R.A., and S.M. Day (2006), Reply to comment by Y. Ben-Zion on "Material contrast does not predict earthquake rupture propagation direction", *Geophys. Res. Lett.*, Vol. 33 (13), L13311, doi:10.1029/2006GL026811

Hartzell, S. (1978), Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.*, 5(1).

Haskell, N. (1964), Total energy and energy spectral density of elastic wave radiation from propagating faults, *Bull. Seism. Soc. Am.*, Vol. 54, No. 6, pp.1811-1851.

Henry, C. and S. Das (2001), Aftershock zones of large shallow earthquakes: fault dimensions, aftershock area expansion and scaling relations, *Geophys. J. Int.*, 147, 272-293.

Langbein, J., J. R. Murray, and H. A. Snyder (2006), Coseismic and Initial Postseismic Deformation from the 2004 Parkfield, California, Earthquake, Observed by Global Positioning System, Electronic Distance Meter, Creepmeters, and Borehole Strainmeters, *Bull. Seism. Soc. Am.*, Vol. 96, No. 4B, pp. S304–S320, doi: 10.1785/0120050823.

Lengline, O., and J.-L. Got (2011), Rupture Directivity of Micro-Earthquake Sequences near Parkfield, California, *Geophys. Res. Lett.*, Vol. 38, L08310, doi:10.1029/2011GL047303

Madariaga, R. (1976), Dynamics of an expanding circular fault, *Bull. Seism. Soc. Am.*, 66 (3), 639-666.

Mai, P., P. Spudich, and J. Boatwright (2005), Hypocenter Locations in Finite-Source Rupture Models, *Bull. Seism. Soc. Am.*, Vol. 95, No. 3, pp. 965–980, doi:10.1785/0120040111

McGuire, J., L. Zhao, and T. Jordan (2002), Predominance of unilateral rupture for a global catalog of large earthquakes, *Bull. Seism. Soc. Am.*, 92 (8), pp. 3309-3317.

Prieto, G., P.M. Shearer, F.L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking P and S spectra, *J. Geophys. Res.*, 109 (B08310).

Rubin, A. (2002), Aftershocks of microearthquakes as probes of the mechanics of rupture, *J. Geophys. Res.*, Vol. 107, No. B7, 2142, 10.1029/2001JB000496

Rubin, A., and D. Gillard, (2000), Aftershock asymmetry/rupture directivity among central San Andreas fault microearthquakes, *J. Geophys. Res.*, Vol. 105, No. B8, pp. 19,095-19,109.

Shearer, P.M., G.A. Prieto, and E. Hauksson (2006), Comprehensive analysis of earthquake source spectra in southern California, *J. Geophys. Res*., 111, B06303, doi:10.1029/2005JB003979.

Shi, Z., and Y. Ben-Zion (2006), Dynamic rupture on a bimaterial interface governed by slip-weakening friction, *Geophys. J. Int.*, doi: 10.1111/j.1365-246X.2006.02853.x

Thurber, C., H. Zhang, F. Waldhauser, J. Hardebeck, A. Michael, and D. Eberhart-Phillips (2006), Three-Dimensional Compressional Wavespeed Model, Earthquake Relocations, and Focal Mechanisms for the Parkfield, California, Region, *Bull. Seism. Soc. Am.*, 96 (4B), S38.

Uchide, T., and S. Ide (2010), Scaling of earthquake rupture growth in the Parkfield area: Self-similar growth and suppression by the finite seismogenic layer, *J. Geophys. Res.*, Vol. 115, B11302, doi:10.1029/2009JB007122.

Waldhauser, F., Ellsworth, W., Schaff, D., and Cole, A. (2004, Jan 1), Streaks, multiplets, and holes: high-resolution spatio-temporal behavior of Parkfield seismicity. Geophys. Res. Lett .

Wallace, R. (1990), The San Andreas Fault System, California (Vol. 1515).

Wang, E., and Rubin, A. (2010, Dec 30), Rupture directivity of microearthquakes on the San Andreas fault from spectral ratio inversion. 1-14.

Warren, L., and Shearer, P. (2002), Mapping lateral variations in upper mantle attenuation by stacking P and PP spectra. J. Geophys. Res , 107, 2342.

Zaliapin, I., and Ben-Zion, Y. (n.d.). Asymmetric distribution of aftershocks on large faults in California. GJI .

Chapter 5: Selecting Empirical Green's Functions in Regions of Fault Complexity: A Study of Data From the San Jacinto Fault Zone, Southern California

The San Jacinto Fault Zone (SJFZ) in southern California features a complex distribution of fault traces and seismic sources. Earthquakes in this region rarely occur along a linear, well-defined fault trace and instead tend to occupy a volume of the crust. We use recent seismicity in this region to examine the appropriateness of applying the empirical Green's function (EGF) technique to data from regions of heterogeneous faulting structures. Within regions of complex fault structure, it can be challenging to select appropriate EGFs (*i.e.*, smaller earthquakes with similar focal mechanisms and hypocentral locations) to constrain the source properties of target M > 3 mainshock events. The wealth of data within the SJFZ allows us to quantify the variability in mainshock source parameter estimates, which we obtain by testing a range of potential EGF events. We leverage these results to define restrictions required to obtain adequate EGF path correction. We find that selecting EGF events with hypocenters within a certain distance of the mainshock hypocenter is not always effective in this region. Requiring matching first motion polarities and high correlation between the mainshock and EGF waveforms recorded at a given station will significantly improve EGF event selection. These constraints can be transferable to other regions lacking substantial groups of homogeneous seismic source mechanisms to guide future EGF event selection.

5.1 Introduction

Obtaining robust estimates of earthquake source properties is key in resolving many questions concerning earthquake source physics. However, effectively accounting for signal contributions from the travel path between a given source and a recording station and from near-site scattering effects is challenging. These are typically referred to as 'path-' and 'site-' effects, respectively. The empirical Green's function (EGF) technique is commonly used to isolate the signal of the seismic source from the recorded ground motion (*e.g.*, Hartzell, 1978; Frankel and Kanamori, 1983; Hutchings and Wu, 1990; Hough, 1997; Prieto *et al.*, 2004). In this technique, for a given earthquake (hereafter referred to as the 'mainshock') the path- and site-effect contributions to the seismic waveform are approximated by the ground motion of a smaller earthquake sharing a similar source location and focal mechanism (Figure 5.1). This approach assumes that the smaller earthquake (hereafter referred to as the 'EGF earthquake') approximates a point-source relative to the larger earthquake, and that the waveform of the smaller earthquake primarily represents the propagation path and site effect signals. The mainshock source is thus obtained by deconvolving the EGF earthquake signal from the mainshock signal.

For the EGF technique to adequately identify source contributions within the recorded mainshock ground motion, several conditions must be met. The technique assumes that the EGF earthquake: (1) is sufficiently small to be approximated as a point source relative to the mainshock earthquake; (2) is close enough to the mainshock earthquake such that the propagation paths between each event and a given station are identical; and (3) has a focal mechanism solution that mimics the mainshock and

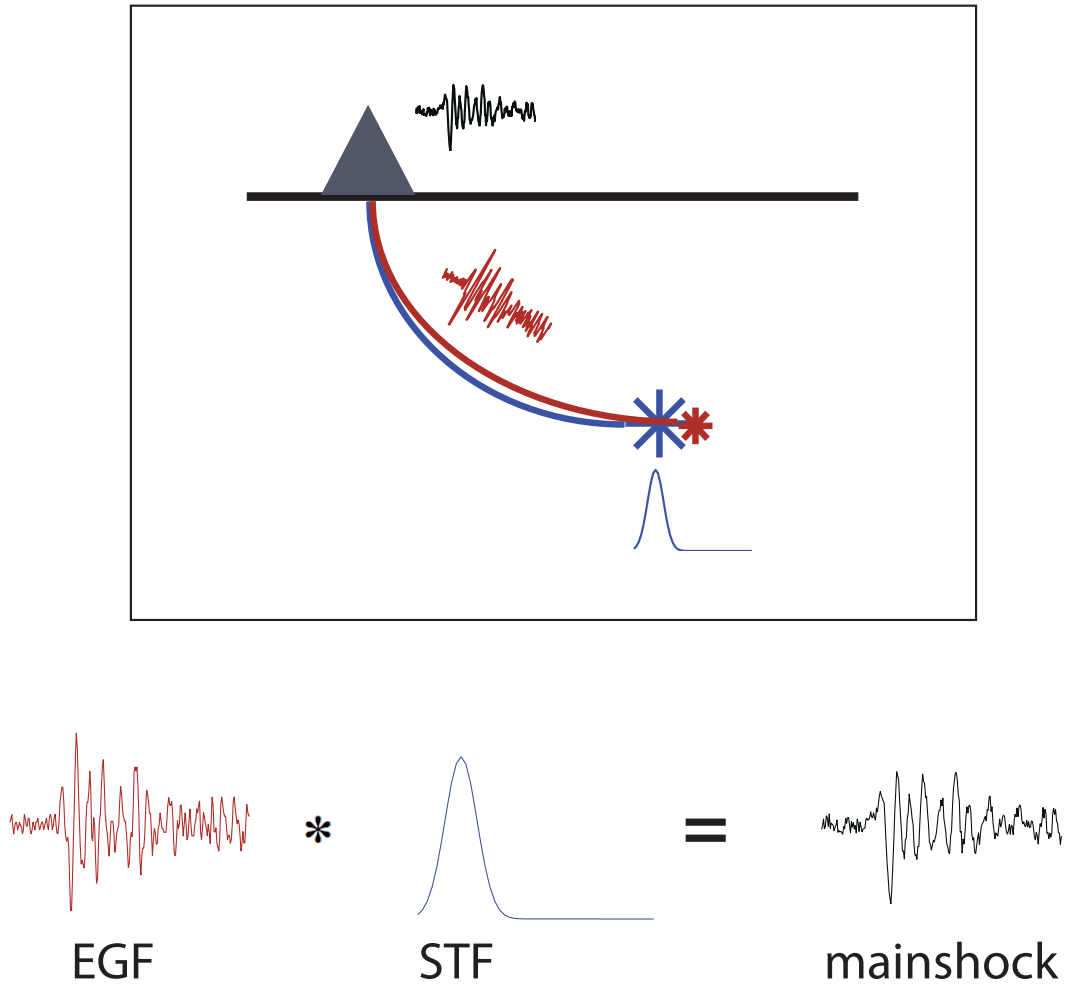

Figure 5.1: Cartoon illustrating the EGF method. Two earthquakes have similar hypocenter locations, are different in magnitude, and are recorded by the same station. The waveform recorded for the larger earthquake (black) is equivalent to a source time function (blue) convolved with the smaller earthquake waveform (red).

therefore the two events share an identical radiation pattern. Deviations from these conditions could result in errors in the mainshock source solution (Velasco *et al.*, 1994; Vallee, 2004). In practice, it is not practical to verify each condition is met. To satisfy the point-source approximation, the EGF earthquakes are often restricted to be at least one unit of magnitude smaller than the mainshock. The hypocenters of the two events are commonly limited to 1-2 km epicentral spacing, particularly for smaller datasets with few EGF options. Because focal mechanism solutions are typically not available for small earthquakes, similar focal mechanism orientation of the EGF and mainshock events is often assumed without confirmation. Any strong deviations in any of these EGF technique assumptions can contribute erroneous errors in the results.

More sophisticated processing techniques have been developed to refine EGF selection. For example, a set of possible EGF earthquakes can be normalized and stacked in the frequency domain to produce an average EGF in order to reduce errors introduced by variations in the propagation path (*e.g.*, Hough, 1997; Prieto *et al.*, 2004; Allmann and Shearer, 2007). However, this approach requires an adequate number of smaller EGF events located close to the mainshock, and many datasets lack this type of sufficient data. In another approach, if the EGF earthquake size and location assumptions are appropriately met but the focal mechanisms of the mainshock and EGF differ, then the EGF waveforms can be rotated to match the mainshock radiation pattern (Tan and Helmberger, 2010). These varied tactics can be effective in many situations (*e.g.*, regions of relatively homogeneous faulting mechanisms and high seismicity) but because the 'true' answer is unknown it is difficult to measure the technique's successes and failures.

If the selected EGF event is a poor choice from the start, applying these more sophisticated processing techniques will prove ineffective.

In this study, we use data from the San Jacinto Fault Zone (SJFZ) in southern California to test the EGF method and to estimate conditions necessary for obtaining success with the EGF method. The SJFZ provides a good testing ground because it has a high seismicity rate within a highly heterogeneous fault structure, and because it hosts earthquakes with a range of focal mechanism orientations, magnitudes, and locations (Figure 5.2). The variability in this unique dataset allows us to explore the parameter space of EGF event selection and to quantify the robustness of EGF results.

5.2 The San Jacinto Fault Zone and the ANZA Network

The SJFZ is one of several right-lateral strike-slip fault zones in southern California. Seismicity within the SJFZ has been monitored by the ANZA seismic network since 1982, and over 65,000 local earthquakes have been recorded and cataloged. These events are distributed in a broad swath along the SJFZ near Anza, California. Many of the event locations are concentrated within the Hot Springs cluster to the northwest of the Anza seismicity gap and within the Trifurcation cluster to the southeast of the gap (Figure 5.2). Bailey *et al.* (2010) analyzed the distribution of focal mechanism types within several fault zones in southern California, and suggested that the heterogeneity of focal mechanism types within the SJFZ in particular could be related to variability in fault zone structure. The heterogeneous distribution of source focal mechanisms (Figure 5.2) and broad distribution of hypocentral locations provides an

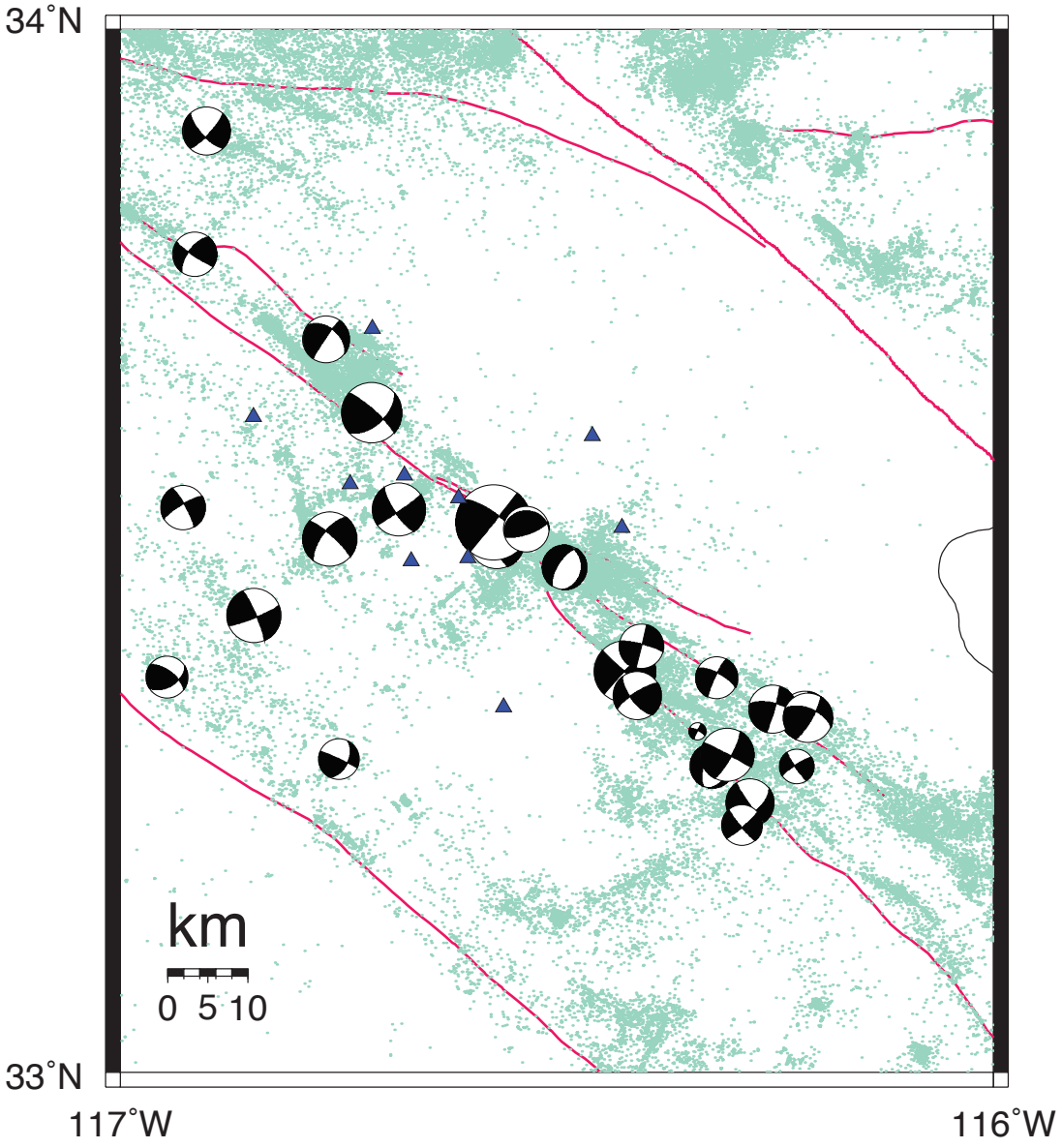

Figure 5.2: Map of SJFZ seismicity and M > 3 focal mechanisms for events in this study. Small dots indicate seismicity from 1982 to 2005, red lines outline mapped fault traces, and blue triangles indicate locations of ANZA stations. Available focal mechanism solutions (Hardebeck and Shearer, 2003) are plotted for a subset of M > 3 events used in this study.

ideal dataset for testing the EGF method and for determining what limitations should be imposed when selecting EGF events in these types of complex regions.

This work focuses on a set of 52 M > 3 mainshock earthquakes recorded by the ANZA network (Figure 5.2). For each mainshock, we create a catalog of potential EGF earthquakes which locate within ~10 km of the mainshock epicenter and are included in the relocated catalog by Lin *et al.* (2007). The hypocenter locations in the Lin *et al.* (2007) catalog are generally within ~3 km of the hypocenters in the original ANZA catalog (Figure 5.3). Because our analysis requires precise interevent locations, we use only events in the Lin *et al.* (2007) relocated catalog.

Our selection criteria net a dataset of 52 trial mainshocks, 56,196 trial EGF events, and a total of 183,953 P-wave velocity waveforms recorded by the ANZA network (Figure 5.2). For each waveform, we compute the spectrum from a one second time window beginning 0.25 seconds prior to the P-wave arrival. We also compute a noise spectrum using one second of data prior to the start of the P-wave. These spectra are calculated using the multitaper method (Thomson, 1982) with a time-bandwidth parameter of 3.5. We restrict each EGF record to have a minimum signal to noise ratio (SNR) of 3 within the 2 to 5 Hz band. We also require all data in our study to have a minimum SNR of 3 for at least half of the frequency points examined.

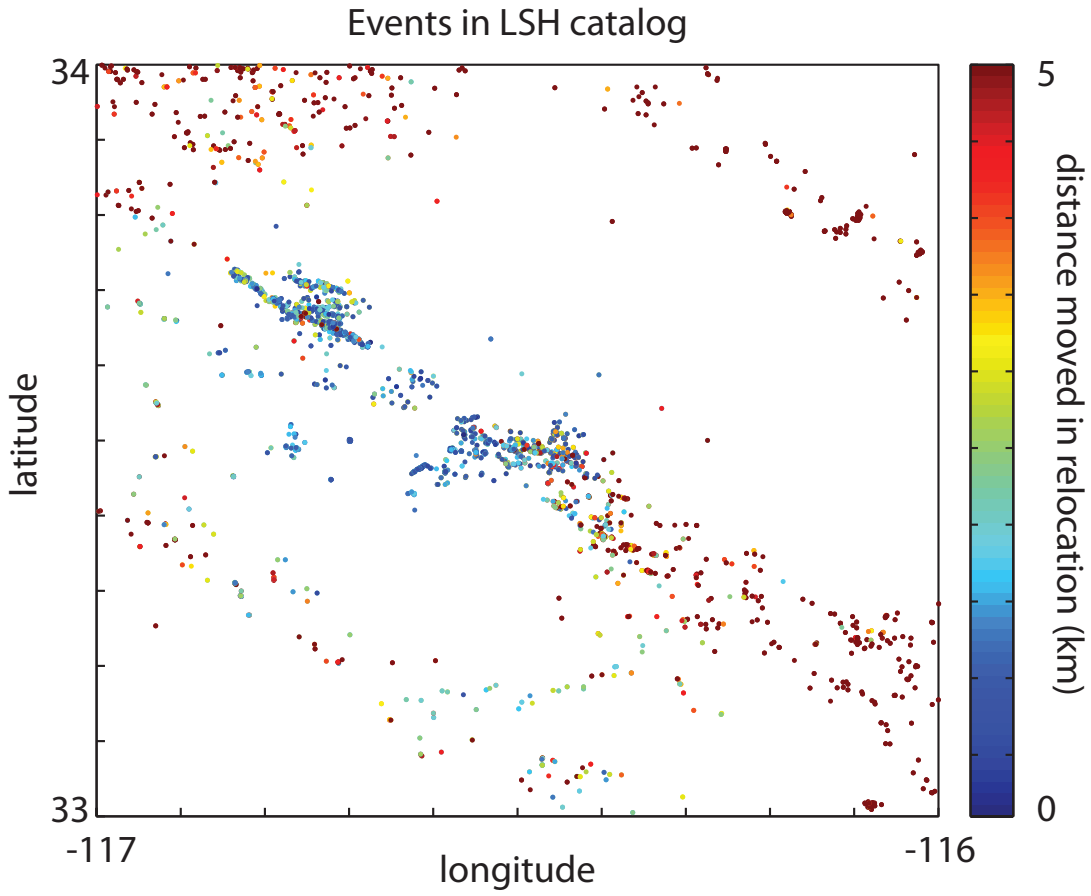

Figure 5.3: Map of relocated seismicity. Color scale indicates the distance between the original catalog location in the ANZA catalog and the location in the Lin *et al.* (2007) relocated catalog. Events located nearest to the ANZA stations generally moved smaller distances in the relocation process than events with less adequate station coverage.

5.3 Source spectra and source parameter estimates

The ground motion recorded at a given station can be represented as a convolution of signals from the source, *s(t)*, the propagation path between source hypocenter and recording station, *p(t)*, and any near-site surface scattering effects and contributions from the recording instrument, *i(t)*. For the mainshock earthquake (indicated by subscript *m*), we formulate this as:

$$u_m(t) = s_m(t) * p_m(t) * i_m(t) \qquad < 5.1 >$$

We represent the EGF earthquake similarly using a subscript of *egf*:

$$u_{egf}(t) = s_{egf}(t) * p_{egf}(t) * i_{egf}(t) \qquad < 5.2 >$$

In accordance with the EGF method, we treat the propagation paths and instrument effects for both the mainshock and EGF earthquakes as identical (e.g., $p_m(t) = p_{egf}(t)$ and $i_m(t) = i_{egf}(t)$). We also assume that the EGF earthquake source is a point-source in time, and hence s_{egf} is a scalar value rather than a function of time. Several studies have confirmed this approximation for EGF events of M < 2.5 when the mainshock is at least one unit of magnitude larger than the EGF event (*e.g.*, Frankel and Kanamori, 1983; Hutchings and Wu, 1990). These assumptions allow us to rewrite equation < 5.1 > as:

$$u_m(t) = s_{rel}(t) * u_{egf}(t) = \frac{s_m(t)}{s_{egf}} * u_{egf}(t) \qquad < 5.3 >$$

Here, $s_{rel}(t)$ is the relative source time function obtained by scaling the mainshock source by the amplitude of the EGF source. The waveforms $u_m(t)$ and $u_{egf}(t)$ are the recorded

ground motions. Obtaining $s_{rel}(t)$ requires inverting the recorded ground motion equation to solve for the source term.

In practice, it is common to approach this inversion in the frequency domain because the convolution can be simply represented as multiplication:

$$U_m(f) = S(f)U_{egf}(f) \qquad < 5.4 >$$

Then the relative source spectrum, $S(f)$, is determined through spectral division:

$$S(f) = U_m(f)/U_{egf}(f) \qquad < 5.5 >$$

Several source spectrum models have been developed for simple ruptures (*e.g.*, Brune, 1970; Boatwright, 1984) to estimate properties of the seismic source from the relative source spectrum. In this study, we use the Brune (1970) source spectrum to fit the relative seismic moment, M_0, and the corner frequency, f_c:

$$S(f) = \frac{M_o}{1 + \left(\frac{f}{f_c}\right)^2} \qquad < 5.6 >$$

Here, the corner frequency is inversely related to the source rupture duration.

Various inversion techniques can be used to obtain the best-fitting parameters of the Brune source spectrum. In this work, we use a nonlinear least squares approach with an applied weighting equal to the inverse of the frequency for each data point. We favor this approach because it minimizes bias from a greater number of data points at higher frequencies, and instead gives greater weight to the more stable, less noisy, lower frequency data points (Figure 5.4). We compute spectral ratios and identify the preferred corner frequency for all possible mainshock/EGF event pairs meeting the signal to noise ratio criteria.

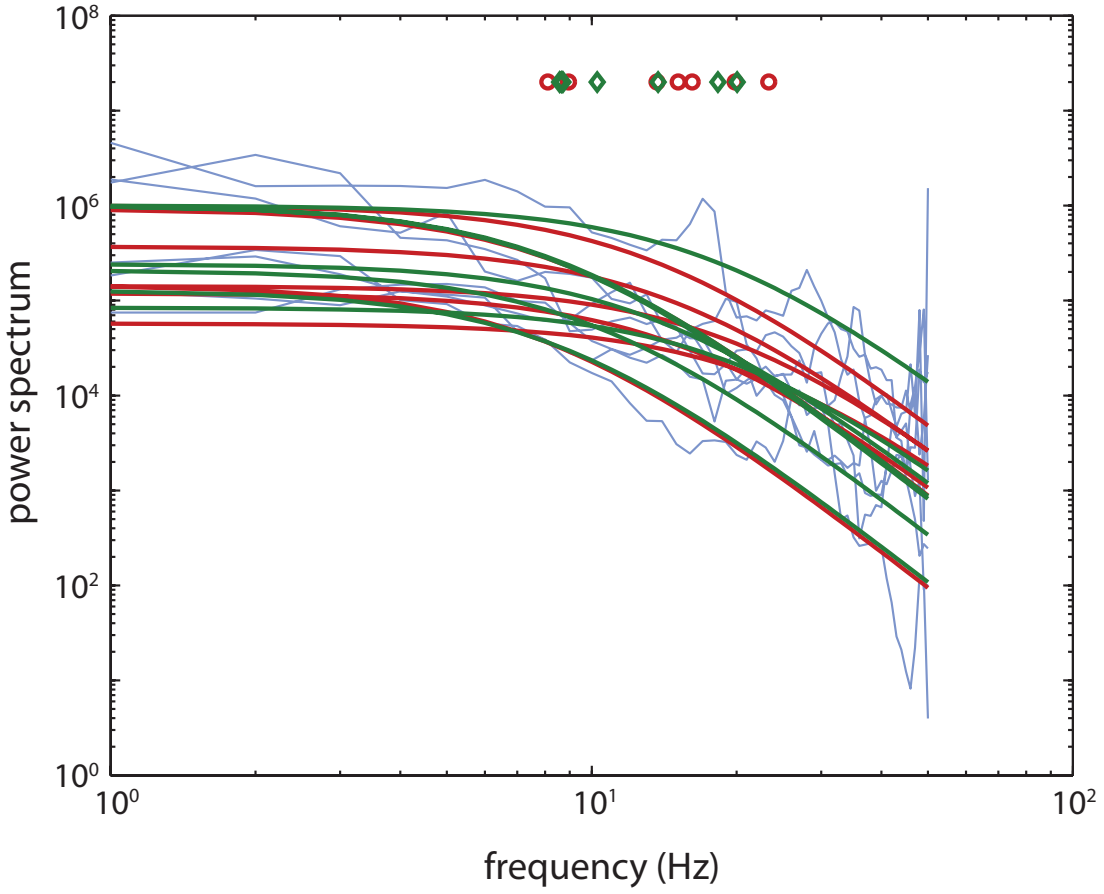

Figure 5.4: Sample source spectrum fitting for a M 3.6 mainshock and M 1.6 EGF earthquake. Spectral ratios for individual stations are plotted in blue. Brune source spectrum fits for each spectral ratio are plotted in red (no weighting) and green (weighted inversely to frequency). Markers towards the top of the plot indicate corner frequency measurements.

We assume that an ideal choice of mainshock and EGF event pairing will generate identical corner frequency estimates at all stations in the array. This approach overlooks the potential effects of complications such as unilateral rupture directivity, which would result in an azimuthal dependence in corner frequency measurements. To analyze how well each EGF earthquake represents the path effects for a given mainshock, we measure the standard deviation of the corner frequency estimates at all stations. Prior to measuring this standard deviation, we take the logarithm of the corner frequency estimates prior to measuring this standard deviation in order to normalize the effects of fitting the spectrum in the log domain (e.g., the difference in source spectrum fit for a shift in corner frequency from 10 to 15 Hz is significantly larger than the fit for a shift in corner frequency from 30 to 35 Hz). We define this measured quantity as the 'corner frequency variability'.

5.4 Variability in EGF waveforms

A key aspect of our analyses is that we use a suite of possible EGFs over a very broad range of characteristics. We consider trial EGF earthquakes that easily fall within normal limitations for EGF selection along with trial EGF events that fall far outside of these normal limits in order to sufficiently identify the transition region separating these populations.

Traditional EGF techniques require the EGF earthquake and mainshock earthquake hypocenters to be very close to each other. In Figure 5.5, we present an example mainshock and a series of potential EGF earthquakes at various hypocentral

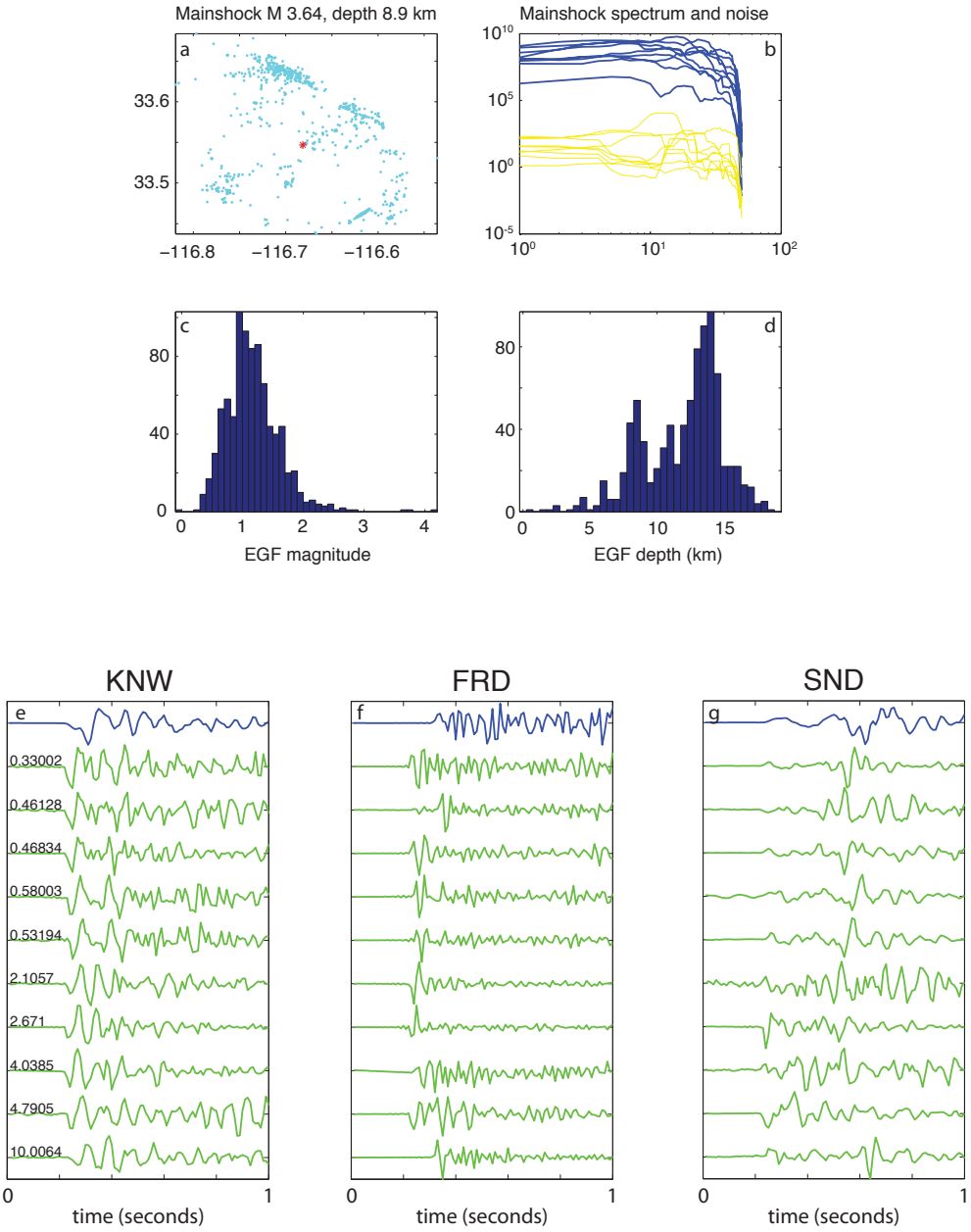

Figure 5.5: Example M 3.6 mainshock and set of M 0.8 to M 2.0 EGF events at different hypocentral separation distances. In (a), the locations of the mainshock and EGF events are shown on a map. In (b), the mainshock (blue) and noise (yellow) spectra are plotted. Histograms show the distribution of (c) magnitudes and (d) depths for potential EGF events. In (e), (f), and (g), the mainshock waveform for each station is plotted in blue. A subset of potential EGF waveforms are shown in green. These are ordered by distance between mainshock and EGF hypocenter locations (given by numbers).

separation distances. At the closest spacing between mainshock and possible EGF, the velocity waveforms vary considerably across the array but exhibit similar characteristics between mainshock and EGF records at each individual station. Such similarity would traditionally confirm the validity of the EGF assumptions and the acceptability of this mainshock/EGF pairing. As the spacing between mainshock and EGF hypocenters increases, these waveforms are more highly variable at each station because the assumption of identical path propagation effects is less appropriate.

In regions of relatively homogeneous focal mechanism distributions, it is reasonable to assume that a mainshock and EGF event pair chosen by location proximity will likely satisfy the requirement of identical focal mechanism solutions. In the SJFZ, where the distribution of focal mechanism orientations is heterogeneous relative to other regions in California (*e.g.*, the San Andreas Fault at Parkfield), this criterion is less likely to be satisfied. Another problem is that focal mechanism solutions are prone to substantial uncertainties and smaller magnitude events such as those used as EGF earthquakes often do not have reliable solutions (Hardebeck and Shearer, 2003). A reasonable proxy for focal mechanism similarity is thus waveform similarity (*e.g.*, Lin *et al.*, 2007). This can be quantified by the peak cross-correlation coefficient between waveform pairs at each station. If the mainshock and EGF events share a common location and hence propagation path, then we expect the only differences in the recorded waveforms to be from differences in the sources. If in addition the mainshock and EGF events share a common focal mechanism orientation, then we expect the first-motion polarity at each station to be identical for the two events. A notable difference in first-

motion polarity between the two events would indicate a corresponding difference in focal mechanisms.

To help assess focal mechanism similarity, we compute peak cross-correlation coefficients for each mainshock and EGF waveform pair at each station using 0.25 seconds of data following the P-wave arrival. This time window excludes the noise prior to the P-wave arrival while retaining several cycles of P-wave ground motion. In Figure 5.6, we provide an example of a mainshock and a set of possible EGF events that were selected based on waveform similarity. As expected, the EGF events with the highest peak cross-correlation coefficients indeed exhibit waveforms very similar to the mainshock. However, the peak cross-correlation coefficient can remain large even when the alignment of the waveforms is shifted by a cycle. This is a common problem especially when the first-motion polarities for the two waveforms do not match yet the rest of the waveform is very similar (see Figure 5.6 for an example). This example illustrates that using pairwise waveform similarity as a proxy for focal mechanism similarity does not always provide the correct result. To help resolve this problem, we suggest using waveform cross-correlation values in addition to assessing agreement in the waveforms' first-motion polarities.

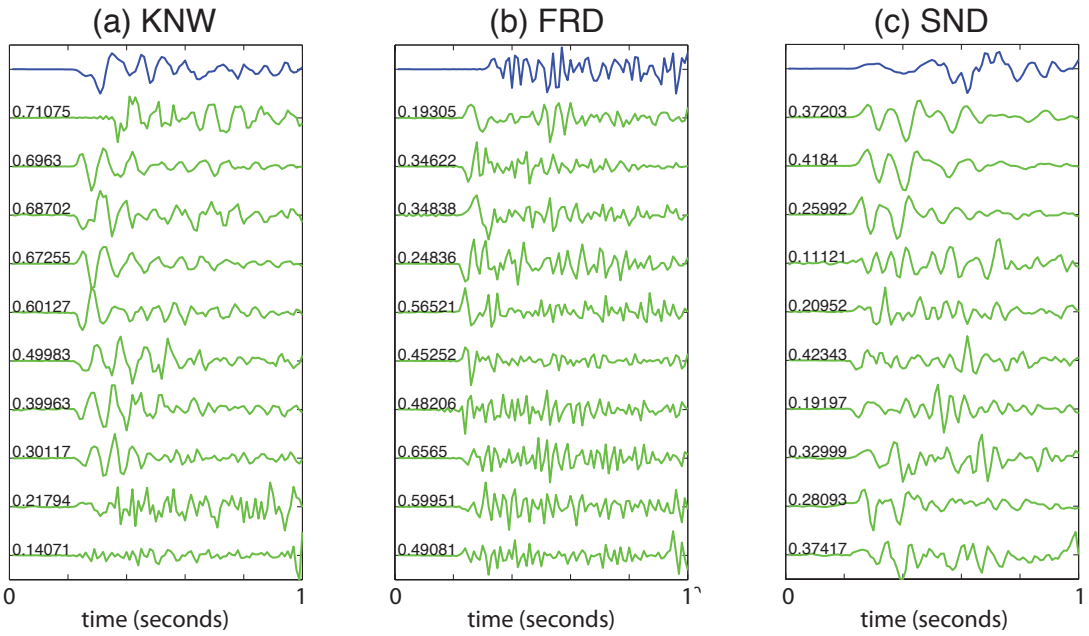

Figure 5.6: Example mainshock and set of EGF events with varying peak cross-correlation coefficients. EGF events are selected from the same set as in Figure 5.5; EGF events shown here are M 0.6 to M 2.4. Here, the events are chosen based on waveform cross-correlation and ordered from highest peak cross-correlation to lowest (values given at left) at station KNW.

5.5 EGF selection criteria

General guidelines for selecting an appropriate EGF for a given mainshock include finding a nearby earthquake with a magnitude at least one unit smaller than the mainshock. This approach should work fairly well in regions that contain minimal complexity and earthquakes with relatively similar source mechanisms and similar hypocentral locations. However, in practice, many regions do not have adequate station coverage to obtain robust hypocenter locations much less well-constrained focal mechanism solutions, making it difficult to assess if a particular EGF event is a good choice. In this section, we focus on each EGF selection criterion individually to determine what minimum requirements need to be satisfied when using data from the SJFZ. We assume that the preferred EGF event will produce the same corner frequency estimate at every station, and so we quantify the variability in corner frequency estimates for each mainshock and EGF earthquake pairing across the network.

5.5.1 EGF magnitude

Requirements for differences in magnitude between the mainshock and EGF events are among the easiest criteria to apply because magnitude is relatively easy to estimate. To determine how the magnitude differential between the mainshock and the EGF earthquake affects the resulting source parameter estimates, we focus on a set of possible EGF earthquakes for each mainshock that are located within 3 km (hypocentral distance) of the mainshock. For each mainshock and EGF event pair we measure the

magnitude differential and peak waveform cross-correlation value at each station. We define the expected ideal subset of these data pairs as those with interevent hypocentral separation distances of less than 500 m and median peak cross-correlation values over the array of 0.55 or greater. We analyze the resulting corner frequency variability by looking at the mean values in differential magnitude bins of 0.5 magnitude units. The resulting binned corner frequency variability estimates are slightly lower for the more ideal subset of data than for the overall set. Each set of data shows a weak decreasing trend of corner frequency variability with magnitude differential. It appears that once the EGF earthquake is sufficiently small (at least one unit of magnitude smaller than the mainshock) any smaller EGF events do not produce significantly different estimates. This is in agreement with previous studies (e.g., Hutchings and Wu, 1990) which found that EGF earthquakes at these magnitudes do not feature EGF source characteristics and the point-source assumption is valid. Our results indicate that the commonly used minimum magnitude differential of one unit of magnitude should be sufficient, and that resulting corner frequency estimates likely improve minimally as the magnitude differential increases.

5.5.2 Mainshock and EGF hypocentral separation distance

We next consider the effects of the separation distances between the mainshock and EGF hypocenter locations. We limit the EGFs considered to those with magnitudes of 1.5 to 2.5 units smaller than the mainshock magnitudes in order to remove magnitude differential effects. We measure the corner frequency variability for these event pairs and

then compute the median of these estimates over 0.1 km separation distance bins of the hypocentral locations (Figure 5.7). The resulting trend of the median corner frequency variability with changing interevent separation distance has two interesting features. First, the corner frequency variability is smallest at the closest separation distances and increases approximately linearly with separation distance up to ~2 km. This result validates the expectation that the closest events will produce the best estimates, and provides an upper bound on separation distance at which the variability in the results no longer deviates substantially from those at greater separation distances where we do not expect the data to meet the EGF assumptions. The second interesting feature is the relatively constant median corner frequency variability with separation distance from ~2 km separation to ~14 km separation. This relatively constant value is fairly low, and the minimal change with distance demonstrates that the waveforms must have fairly similar characteristics even as the interevent spacing increases.

5.5.3 Mainshock and EGF waveform similarity

Last, we consider the effects of selecting EGF events based on waveform similarity measured using peak cross-correlation coefficients alone. We compute the binned corner frequency variability over cross-correlation coefficient bins of 0.05 width. We use a subset of event pairs with EGF magnitudes 1.5 to 2.5 units smaller than the mainshocks; this subset considers potential EGFs at all distances. We explore the results by using the mean, maximum, and median values of cross-correlation coefficients for each mainshock and EGF event pair, because we have an estimate of cross-correlation for

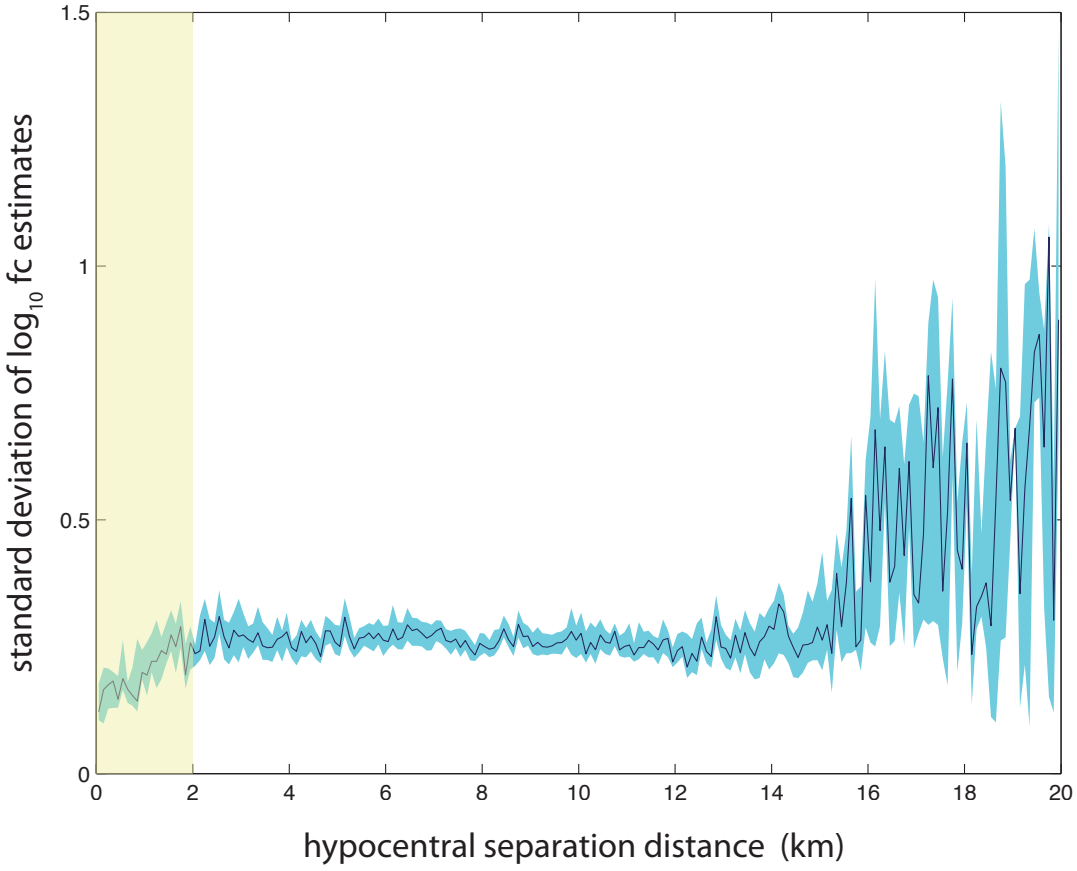

Figure 5.7: Median corner frequency variability plotted versus hypocentral separation distance. The results are shown using distance bins of 0.1 km width. The highlighted region marks 2 km for reference.

each station in the array rather than a single value for the event pair as in our previous analyses. With each of the three quantities considered (mean, median and maximum cross-correlation coefficient), a trend of corner frequency with cross-correlation is apparent. These trends all exhibit low corner frequency variabilities at high levels of correlation, with a transition to increasing corner frequency variability at lower levels of correlation. This observation reinforces the importance of waveform similarity for EGF event selection; it matches our expectation of similar corner frequency results across the array when an appropriate EGF is selected and more highly variable corner frequency results when a poor EGF choice is made. Our results suggest that an acceptable result can be achieved when the mean and/or median cross-correlation coefficient is greater than 0.35 or the maximum cross-correlation coefficient is greater than 0.5.

5.6 Discussion

Our results show that for data with relatively robust and small uncertainties in earthquake locations, selecting EGF events based on the separation distance between the mainshock and EGF hypocentral locations is more important in EGF choice than selection based on magnitude differentials or similarity in focal mechanism orientation. However, the focal mechanism solution similarity is critical for estimating robust results because event pairs that lack similarly oriented focal mechanisms will violate the assumptions of the EGF technique.

In Figure 5.8, we plot the hypocentral separation distance versus the median peak cross-correlation coefficient in 0.1 km distance bins. Importantly, it is clear that the cross-

correlations are best for events within ~1 km of each other and that events at greater distances have similar median cross-correlations regardless of separation distance. Our observations are consistent with those that highlight the importance of waveform coherency (Scott, 1992). Using SJFZ data, Scott (1992) observed that closely spaced events had highly coherent spectra even at separation distances larger than expected from previously described models. Some stations have higher coherence at higher frequencies and for events with greater interevent spacing distances than others. The relatively high coherence at the lower frequency bands for interevent spacing of greater than a few hundred meters indicates that the material can be assumed to be locally homogeneous at depth (Scott, 1992). This has the potential result of causing the effect of separation distance between mainshock and EGF event to not be as strong as what might be expected in regions with lower waveform coherence. As a result, the source spectra fits appear to give acceptable results even when our assumptions of collocation are violated by >1 km. This indicates that in complex regions such as along the SJFZ, a less strict EGF choice may appear acceptable because the corner frequency estimates behave as expected with azimuth, yet the results will be incorrect. This work suggests an appropriate EGF should be selected using relocated earthquake catalogs to get accurate separation distances, choosing separation distances of less than 1 km, and confirming waveform similarity between the mainshock and EGF waveforms at each station. Each of these three requirements must be met to obtain robust results and minimize erroneous results.

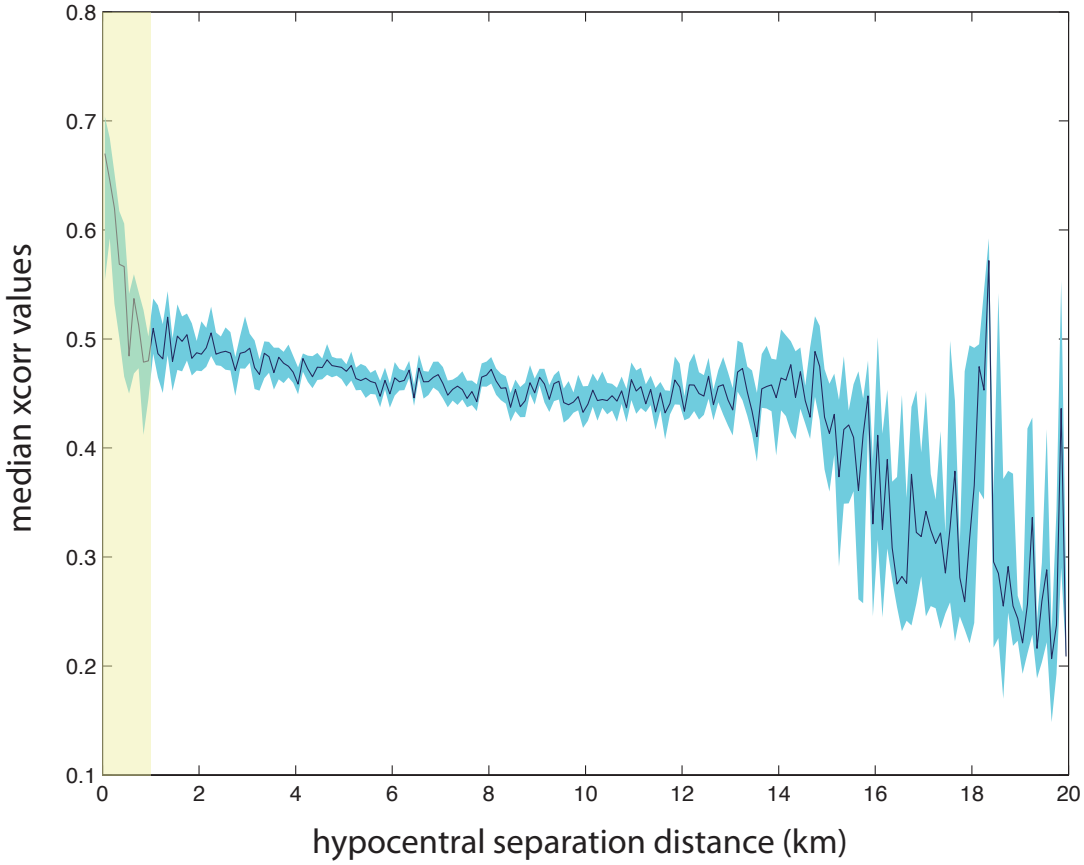

Figure 5.8: Median peak cross-correlation values plotted versus hypocentral separation distance. These results are shown using distance bins of 0.1 km width. The highlighted region marks 1 km for reference.

5.7 Conclusions

We have presented a simple analysis of the characteristics required for selecting an acceptable EGF earthquake for source parameter estimation in the SJFZ. In this region, it is clear that EGF selection can be greatly improved by precisely measuring the separation distances between mainshock and EGF earthquakes with a relocated hypocenter catalog and by comparing the waveforms at each station to confirm matching first motions and high waveform cross-correlation coefficients. Simply applying a pairwise separation distance requirement could net inappropriate EGF choices. We have quantitatively identified a crucial maximum event pair separation of ~2 km, above which we find results indistinguishable from pairs with much greater separation. This demarcation in separation distance likely results because of the high degree of waveform coherence at the lower frequencies likely generated by deeper homogenous structures (Scott, 1992). The difference in magnitude between the mainshock and EGF event is a less important factor as long as the commonly used minimum difference of one magnitude unit is met to satisfy the point-source approximation.

Our comparison of hypocentral separation distance with waveform cross-correlation coefficients suggests that separation distances of 1 km should be considered the maximum allowed, at least within regions of high complexity such as the SJFZ. At this interevent spacing, the waveforms are generally more highly correlated at any given station than at larger spacings. Ideally, the cross-correlation coefficients for any given pair would be measured to confirm an acceptable EGF choice.

The process of selecting appropriate EGF events should be approached with caution because source parameter estimates determined using loose EGF constraints may be contaminated by additional signals not matching the propagation path. Use of more sophisticated EGF techniques, including simultaneous use of multiple EGF events, will average any inconsistencies but may not sufficiently correct for such contributions in regions of highly heterogeneous seismicity.

Acknowledgements

We thank the ANZA group for collection, analysis, and cataloging of the San Jacinto Fault Zone data. This material is based upon work supported by the Southern California Earthquake Center (SCEC). This work (in full) is being prepared for submission to the Bulletin of the Seismological Society of America (Kane, D.L., D.L. Kilb, and F.L. Vernon, Selecting Empirical Green's Functions in Regions of Fault Complexity: A Study of Data From the San Jacinto Fault Zone, Southern California). Deborah Kane is the primary investigator and author of this paper.

References

Allmann, B., and P. Shearer (2007), Spatial and temporal stress drop variations in small earthquakes near Parkfield, California, *J. Geophys. Res*., Vol. 112, B04305, doi:10.1029/2006JB004395.

Bailey, I.W., Y. Ben-Zion, T.W. Becker, and M. Holschneider (2010), Quantifying focal mechanism heterogeneity for fault zones in central and southern California, *Geophys. J. Int.,* doi: 10.1111/j.1365-246X.2010.04745.x

Boatwright, J. (1984), The effect of rupture complexity on estimates of source size, *J. Geophys. Res.*, Vol. 89, No. B2, pp. 1132-114.

Brune, J. (1970)., Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.*, 75 (26), 4997-5009.

Frankel, A., and H. Kanamori (1983), Determination of rupture duration and stress drop for earthquakes in southern California, *Bull. Seism. Soc. Am.*, 73 (6), 1527-1551.

Hardebeck, J. L. and P. M. Shearer, (2003), Using S/P Amplitude Ratios to Constrain the Focal Mechanisms of Small Earthquakes, *Bull. Seismo. Soc. Am.*, 93, 2434-2444.

Hartzell, S. (1978), Earthquake aftershocks as Green's functions, *Geophys. Res. Lett.*, 5(1).

Hough, S.E. (1997), Empirical Green's function analysis: taking the next step, *J. Geophys. Res.,* Vol. 102, No. B3, pp. 5369-5384.

Hutchings, L. and F. Wu (1990), Empirical Green's Functions From Small Earthquakes: A Waveform Study of Locally Recorded Aftershocks of the 1971 San Fernando Earthquake, *J. Geophys. Res.*, Vol. 95, No. B2, pp.1187-1214.

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

Prieto, G., P.M. Shearer, F.L. Vernon, and D. Kilb (2004), Earthquake source scaling and self-similarity estimation from stacking P and S spectra, *J. Geophys. Res.*, 109 (B08310).

Scott, J. (1992), "Microearthquake studies in the Anza Seismic Gap", Ph.D. Dissertation, University of California San Diego.

Tan, Y. and D. Helmberger (2010), Rupture Directivity Characteristics of the 2003 Big Bear Sequence, *Bull. Seism. Soc. Am.*, Vol. 100(3), pp. 1089-1106.

Thomson, D.J. (1982), Spectrum estimation and harmonic analysis, *Proceedings of the IEEE*, Vol. 70 (9), pp. 1055-1096.

Vallee, M. (2004), Stabilizing the Empirical Green Function Analysis: Development of the Projected Landweber Method, *Bull. Seism. Soc. Am.*, Vol. 94 (2), pp. 394-409.

Velasco, A.A., C.J. Ammon, and T. Lay (1994), Empirical green function deconvolution of broadband surface waves: Rupture directivity of the 1992 Landers, California (Mw= 7.3), earthquake, *Bull. Seism. Soc. Am.*, Vol. 84, No. 3, pp. 735-750.

Chapter 6: Conclusions

Increasing our understanding of earthquake source physics and developing the potential to apply this knowledge to improving earthquake hazard studies requires being able to effectively measure earthquake sources of both large and small earthquakes. Such knowledge is critical for ultimately estimating seismic hazard in regions vulnerable to large earthquakes and for appropriately mitigating these hazards. This thesis has focused on methods of estimating source parameters of small earthquakes and particularly focuses on the intricacies of the empirical Green's function technique. I've considered several aspects of source parameter estimation, including:

1) complications due to shallow site effects and their contributions to source parameter measurement uncertainties;
2) using a non-parametric time-domain method to estimate source parameters; and
3) how to best select an appropriate EGF in the San Jacinto Fault Zone and in other regions of heterogeneous seismicity distribution.

I've also applied an EGF-like average propagation path correction to look for evidence of unilateral rupture directivity of small earthquakes.

In Chapter 2, we use data from a small aperture array to quantify uncertainties in estimates of corner frequency and stress drop. Because the array consists of 60 closely spaced stations and the events considered are at much greater distances from the array than the interstation spacing, these uncertainties are due primarily to near-surface site effects from wave propagation directly beneath the array. Our results suggest that

minimum uncertainties of stress drop estimates should allow for ±30% variation. This result has serious implications for studies of earthquake source scaling, in which stress drop results from many different regions and using many different methods are often combined to consider the scaling of stress drop with earthquake magnitude. We also find that although the distribution of uncertainties for a given mainshock is fairly consistent regardless of the choice of EGF earthquake used to correct for propagation effects, the EGF choice affects the absolute value of the source parameter results. It is necessary to take caution when combining source parameter estimates from a series of events or from different studies in order to prevent complications due to measurement uncertainty.

In Chapter 3, we apply a time-domain discrete deconvolution approach to estimate source time functions for small earthquakes. This approach uses EGF techniques but does not need a source model as the frequency-domain approach generally requires; it is a non-parametric method that only assumes the EGF assumptions and that the STF is nonnegative. We additionally describe how this method can incorporate multiple EGF waveforms simultaneously to produce a more robust result. In the simultaneous array STF approach, this allows for several waveform combination possibilities. First, a set of waveforms for a mainshock and EGF event recorded at an array can be used to estimate a single STF solution. This application is best used in a situation with either stations clustered in a single azimuthal direction relative to the event locations (*e.g.*, the small aperture array) or with azimuthally distributed stations around a mainshock without strong rupture directivity characteristics. In a second approach, multiple EGF events can be used with a single mainshock to achieve a more robust STF result at a single station. This presents the potential situation where effects of unilateral rupture directivity could

be observed and quantified by solving for a different STF solution at each source-station azimuth using multiple EGF events simultaneously.

In Chapter 4, we compare displacement spectra from earthquakes on the San Andreas Fault near Parkfield, California to look for signals of rupture directivity towards either along-strike azimuth. In this study, we do not use the traditional EGF techniques, but instead derive a path correction for each of twenty spatial clusters of earthquakes by iteratively separating out the mean source spectrum and path correction spectrum within the cluster. The results do not indicate an overall preference in rupture directivity when considering the full set of significant results over a range of earthquake magnitudes. Such a preference would have implications for the relevance of treating the fault as a bimaterial interface, because computational models predict that such an interface should produce a predictable rupture direction. We do observe a greater proportion of $M > 3$ earthquakes with rupture towards the southeast, which matches predictions of the bimaterial interface model under these conditions. We also observe a temporal effect on rupture directivity following the 2004 M 6 Parkfield earthquake; events following the mainshock have a higher proportion of northwest directed ruptures. We conclude that rupture directivity is likely influenced by many competing factors, and predicting the rupture directivity of large earthquakes may prove difficult even when the fault surface is close to a relatively simple geometrical interface.

In Chapter 5, we consider the effects of EGF earthquake choice on source parameter results for $M > 3$ earthquakes in the San Jacinto Fault Zone. For each mainshock, we use a series of possible EGF events to estimate corner frequencies from source spectral ratios at each station and then analyze the variability in corner frequency

estimates across the entire ANZA network. This analysis demonstrates that EGF earthquake selection must be approached with care because even poor choices for EGF earthquakes can produce results that appear to be reliable. A basic set of criteria for selecting an appropriate EGF event in the SJFZ include requiring the EGF magnitude to be at least one unit of magnitude smaller than the mainshock, requiring the EGF earthquake hypocenter to be within 1 km of the mainshock earthquake hypocenter, and requiring the EGF and mainshock waveforms at each station to have highly correlated waveforms and matching first motion polarities. More advanced EGF techniques such as using multiple EGF earthquakes simultaneously should also take these criteria into account to prevent poor EGF choices from biasing the resulting source parameter estimates.

These studies collectively contribute to increased understanding of the process of source parameter estimation. The results of Chapter 2 and Chapter 5 have implications for past and future studies utilizing EGF techniques to determine source parameters. In particular, previous studies combining analysis of source parameter results determined with several analysis methods could be reinterpreted with additional consideration of uncertainties. The method outlined in Chapter 3 offers an opportunity to measure source parameters while making fewer assumptions, and can also be used to look for signals of rupture directivity. The directivity analysis of events at Parkfield in Chapter 4 indicates that directivity on natural faults is likely controlled by several factors and deserves future study.

6.1 Future Work

The analyses of source parameter estimation techniques presented in this thesis provide guidelines for their effective use under the complicated situations often present in seismological research. Considering uncertainties of these source parameter estimates is a necessary step to resolving the questions of source parameter scaling, particularly when combining results from multiple studies to examine a wide range of earthquake magnitudes, and contributing to increased knowledge of earthquake source physics.

Recent seismic network expansions in the San Jacinto Fault Zone, along with the planned temporary installation of several linear arrays across the fault, will make it possible to much more carefully consider the appropriateness of a given EGF earthquake in this region. The additional azimuthal coverage contributed by these temporary stations will contribute greatly to the abilities to resolve moment tensors for confirming source orientation. Data from borehole stations installed as part of the Plate Boundary Observatory can be used to better determine how the near-surface effects contribute to measurement uncertainties through analysis of colocated borehole and surface stations and will hopefully provide smaller uncertainty bounds when borehole waveform data is incorporated into EGF analysis.

We plan to repeat the analysis of EGF choice presented in Chapter 5 using STF methods and compare the effects of time versus frequency domain approaches. Some studies of EGF selection have checked how well the EGF approximates a path correction by confirming that the STF solution has a reasonable shape; STF solutions that exhibit low amplitudes for much longer durations are unlikely due to true source effects are more likely related to poor EGF choice.

The method described here for determining STFs has potentially interesting applications for earthquake source studies. For example, this method could be used to compute STFs for a large set of earthquakes for which STFs are not routinely computed with moment tensor analysis. In addition to providing a time-domain consideration of source parameters in a region, such a large set of STF results could be used in studies of the shape of moment rate profiles (e.g., Mehta *et al.*, 2006) and could potentially offer insight into earthquake dynamics.

The STF method could possibly be expanded to simultaneously consider rupture directivity of the mainshock. We have performed some initial feasibility tests on synthetic data to try this application and found that it sufficiently resolves the STF and directivity if the shape of the STF can be assumed. Extending this inversion to resolve directivity without assuming STF shape would create an exciting opportunity to study both rupture duration and directivity of large populations of small earthquakes.

Our study of rupture directivity of small earthquakes at Parkfield could be repeated along similarly simple fault structures (*e.g.*, the North Anatolian Fault). Consideration of rupture directivity in more geometrically complicated regions such as the SJFZ can be approached using the STF methods, or by using spectral analysis similar to that of the Parkfield analysis but on individual mainshocks.

References

Mehta, A.P., K.A. Dahmen, and Y. Ben-Zion (2006), Universal mean moment rate profiles of earthquake ruptures, *Physical Review E*, Vol. 73 (5), doi:10.1103/PhysRevE.73.056104