CRUSTAL SHAPING OF HIGH FREQUENCY GROUND MOTION IN THE WASHINGTON-OREGON, NORTHERN AND CENTRAL CALIFORNIA REGIONS

Ali Fatehi, B.S., M.S.

Advisor: Robert B. Herrmann, Ph.D.

Abstract

Large data sets of vertical and horizontal seismograms from the Pacific Northwest Seismic Network, Northern California Seismic Network, and Berkeley Digital Seismic Network were used to study the high frequency (0.25-16 Hz) ground motion characteristics in Washington-Oregon, northern, and central California. I have split the data set of central California into eastern and western parts to investigate regional variations of ground motion parameters. I performed regressions of Fourier velocity spectra and peak filtered ground velocities to obtain the propagation, excitation, and site terms.

The regression results for the propagation term show a simpler from of geometrical spreading in central east California compared to other regions. For northern California this term shows a rapid decrease of amplitude with distance especially for frequencies higher than 4 Hz. The propagation term for the Pacific Northwest and central west California show a typical trend of attenuation with some effects of supercritical reflections.

For the Pacific Northwest, the best model that fits the observation is expressed by $Q(f) = 280\ f^{0.55}$. The geometrical spreading is $r^{-1.2}$ for a distance range of 40 to 70 km. For other distance ranges the decrease of amplitude is less rapid. The central east California and central west California results are parameterized with $Q(f) = 280\ f^{0.50}$ and $Q(f) = 250\ f^{0.30}$ respectively. The northern California results are not easy to model and require using a frequency dependent η and geometrical spreading. The average model for two frequency ranges is $Q(f) = 210\ f^{0.50}$. The geometrical spreading effect for the frequencies higher than 4 Hz is very strong at distances less than 70 km.

The excitation terms were modeled using the Brune's source model. An average stress drop ($\Delta\sigma$) of 30 bars was obtained for the Pacific Northwest. The northern

California average value for stress drop is 90 bars. The observations of central east California were modeled with a stress drop of 45 bars while the best fit for the western part required $\Delta\sigma = 125$ bars.

The wide range of the values obtained for Q, $g(r)$, and $\Delta\sigma$, indicates that the ground motion parameters of a specific region may not be used for another region. For example, using the stress drops of central California may overestimate the ground motion level predictions in the Pacific Northwest.

Regional attenuation relationships can be developed using small earthquake data which are available in many regions. These relationships reflect the regional wave propagation characteristics and give more accurate results in seismic hazard analysis.

CRUSTAL SHAPING OF HIGH FREQUENCY GROUND MOTION IN THE WASHINGTON-OREGON, NORTHERN AND CENTRAL CALIFORNIA REGIONS

Ali Fatehi, B.S., M.S.

Advisor: Robert B. Herrmann, Ph.D.

A Dissertation Presented to the Faculty of the Graduate
School of Saint Louis University in Partial
Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2006

i

Professor Robert B. Herrmann,

Chairperson and Advisor

Assistant Professor Keith D. Koper

Assistant Professor Lupei Zhu

Dedication

To my mother and my wife

Acknowledgments

I would like to express my deep appreciation to my advisor, Dr. Robert Herrmann for his advise and encouragement throughout this work, for his support which made the hard times easier for me, and for his great contribution to my graduate education.

I am grateful to Dr. Kieth Koper who provided his genetic algorithm routines, critically reviewed this dissertation, and by spending too much time, contributed to some parts of my work. I would like to thank Dr. Lupei Zhu for the dissertation review, useful comments, and kind assistance in computational and theoretical problems. I would like to thank Dr. Brian J. Mitchell and Dr. David Crossley for their contribution to my study.

I thank Eric Haug and Bob Wurth for their help in providing technical assistance for computer problems. I also thank Cynthia Wise and Laurie Hausmann for their office related assistance.

I should specially thank Hongyi Li and Alemayehu Jemberie who helped me to settle down when I arrived in Saint Louis. I had many useful discussions with Felipe Leyton and I am grateful to him for his constructive suggestions and his helps. I wish to thank Young Soo Jeon for his assistance and help during my work. I would like to thank all other graduate students for being nice and helpful during my study

years.

I wish to thank my mother, sisters, and brothers for their unconditional love and support. Finally I should thank my wife, Jamileh, for her endless love, support, and encouragement during many years of my studies.

Table of Contents

List of Tables

x

List of Figures

Chapter 1

Introduction

Engineering seismology deals mostly with strong ground motions from earthquakes. These are the motions capable of affecting people, structures and the environment. Evaluation of these effects requires a quantitative description of strong ground motion (*Kramer*, 1996). The results of such quantifications are used in seismic hazard assessment to mitigate the effects of damaging earthquakes. Seismic hazard analysis begins with determining and quantifying all potential seismic sources in the study area. The wave propagation (attenuation) properties are then used to estimate the level of ground motion at a particular distance from the composite of all sources (*Kramer*, 1996; *Reiter*, 1991). A specific attenuation relationship is needed for each region because different regions have different wave propagation characteristics. The source scaling parameters such as stress drop are also different in different tectonic environments (*Kanamori and Anderson*, 1975).

A large number of attenuation relationships, based on the regression analysis of strong ground motion data, have been proposed for different regions. Availability of sufficient strong ground motion data has been a big limitation for deriving the predictive relationships in many parts of the world including the United States, because these data are only available in places which have both large earthquakes and sufficient numbers of strong motion seismographs. A different approach has been taken for regression of small earthquakes to constrain high frequency ground motion. This approach characterizes wave propagation by analyzing small earthquake motions and then uses the random vibration theory (RVT) to estimate peak motions for large earthquakes (*Herrmann and Malagnini*, 2005). Unlike strong ground motion records, data from small earthquakes are available for many parts of US where local seismic networks operate there. Pacific Northwest and Northern California are good

examples of regions with dense local seismic networks and large numbers of small earthquakes. Significant ground motion levels are expected in the Pacific Northwest and surrounding areas from a possible future (North America-Juan de Fuca) subduction zone earthquake. Many parts of northern and central California have high probability of occurrence of large earthquakes in future. Studies in these regions have focused on such possible large earthquakes and their resultant strong ground motions in Cascadia (*Atkinson*, 1995).

While some ground motion studies of the U.S. Cascadia region have focused on the Washington-Oregon region, no detailed study of ground motion scaling has been done for the northernmost part of the California. Studies for Washington-Oregon have used data sets with a sparse distribution of seismic stations. The motivation of this study is to characterize wave propagation and source spectra in Washington-Oregon, northern and central California to provide the state-of-the-art ground motion scaling for use in seismic hazard analysis. I will use short period and 3-component broadband digital seismograms of Pacific Northwest Seismic Network (PNSN), Northern California Seismic Network (NCSN), and Berkeley Digital Seismic Network (BDSN) to estimate the excitation, propagation, and site terms in the frequency range of 0.25-16 Hz. To do this I will use a regression technique and the results will be interpreted in terms of a predictive model to estimate the source and propagation parameters. All three networks have thousands of digital waveforms from small, moderate, and a few large earthquakes of the study areas, which provide an excellent data set for this study. I will also use the coda normalization method (*Frankel et al.*, 1990) to have an independent estimate of the propagation term and to compare with the regression results. The current hazard maps of the study regions use the attenuation relationships based on regression of a sparse strong motion data set. My results can be used in future hazard maps since they cover a wide range of frequencies and distances, some of which were not covered by previous studies due to poor coverage and limited frequency range of strong ground motion instruments.

Chapter 2

Review of Related Literature

2.1 Earthquakes in Pacific Northwest

The Cascadia subduction zone is the convergent boundary between Juan de Fuca and North American plates. This zone extends from southwestern British Columbia to northern California. It is one of the major sources of earthquakes in the Pacific Northwest (Figure 2.1). These earthquakes occur in the Benioff zone and shallow crust. The Pacific Northwest Seismic Network locates more than 1000 earthquakes with magnitude 1 and larger each year (http://wwww.pnsn.org/ INFO-GENERAL/faq.html).

Subduction zone (interplate) earthquakes can have large magnitudes. Geological studies in the Cascadia subduction zone show that large interplate earthquakes in the Pacific Northwest have a recurrence interval of 400–600 years and that the last great earthquake occurred about 300 years ago (http://wwww.pnsn. org/HAZARDS/CASC ADIA/cascadia-zone.html). Some examples of significant shallow crustal earthquakes in Washington-Oregon are: the 1962 north of Portland, the 1981 Mount St. Helen's seismic zone, and the 1997 Bremerton. All three events had magnitudes about 5-5.5 (http://wwww.dnr.wa.gov /geology/hazards/equakes.html).

The Benioff zone earthquakes are also called intraplate earthquakes because they occur inside the subduction plate and not at the plate boundary. The largest of the recorded Benioff zone events in the Pacific Northwest in the last century were the M=7.1 Olympia earthquake in 1949, the M=6.5 Seattle-Tacomma earthquake in 1965, and the M=5.1 Sastop earthquake in 1999. The M=6.8 Nisqually, Washington earthquake in 2001 was the first major seismic event of the region in this century. It occurred in the Wadati-Benioff zone of the Juan de Fuca plate at a depth of 51 km
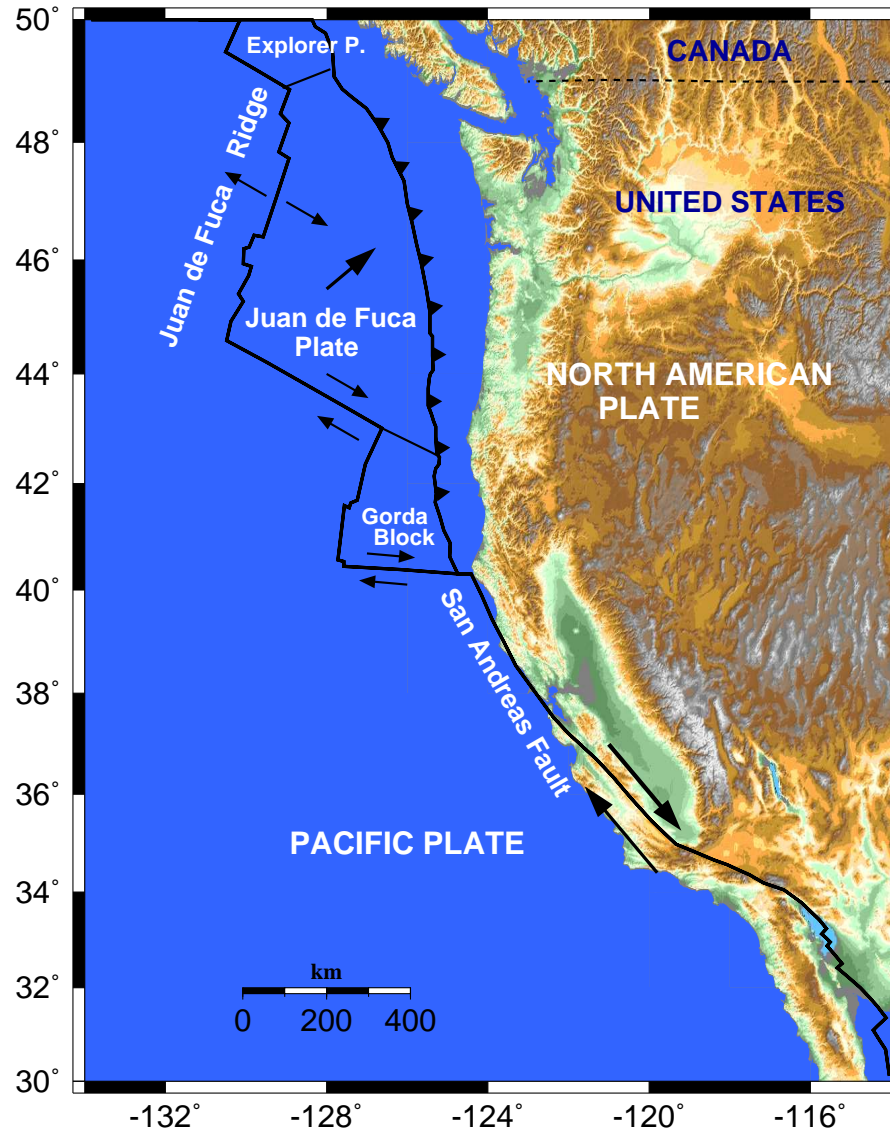
Figure 2.1: Tectonic map of Cascadia subduction zone and California.

beneath the Puget Sound region.

2.1.1 Nisqually earthquake

On February 28, 2001 a significant (Mw=6.8) earthquake occurred in the Puget Sound area of western Washington. The USGS located the event at 47.1°N and 122.72°W at a 51 km depth. It produced ground shaking over a wide area and caused damage in the cities of Olympia, Seattle, and Tacoma. The maximum recorded ground acceleration was 0.3 g (www.neic.usgs.gov/neis/eq_depot/2001/eq_010228/).

The earthquake had a normal, dip-slip mechanism. Bending of subducting plate in subduction zones causes the normal faulting in subduction zones. Juan de Fuca is a young plate with a moderate rate of convergence which indicates that the Cascadia subduction zone has a warm slab environment (*Choy and Kirby*, 1994). The Nisqually earthquake was an unusual event compared to other deep events of the same magnitude in the Pouget Sound region. It had a low stress drop and lacked high frequency energy. Furthermore , the Nisqually earthquake was followed by only a few aftershocks.

2.2 Earthquakes in northern and central California

The San Andreas and Hayward faults are two major faults in northern and central California. The San Andreas fault is the main boundary between the Pacific plate and North America plate (Figure 2.1). The Berkeley Digital Seismic Network and the Northern California Seismic Network record a large number of earthquakes in California each year. The earthquake record history of California goes back to 1769 when an earthquake happened near Los Angeles (http://neic.usgs.gov/neis/ states/california/california_his tory.html). The 1906 San Fransisco earthquake (estimated M_W=7.8) was one of the most destructive events in the California. It ruptured a branch of the San Andreas fault. Several other large events have struck different parts of California which have caused extensive damage and loss of lives. The most notable recent events include: the 1989 Loma Prieta earthquake (Mw=6.9), the 2003 San Simeon earthquake (Mw=6.5), and the 2004 Parkfield earthquake (Mw=6.0).

2.3 Previous ground motion studies

Previous ground motion studies in the world including the Pacific Northwest and California have focused on evaluation of source, propagation and site effects from observed earthquake data. Estimation of source, propagation, and site effects has

been performed using linear regression and nonlinear inversion of observed ground motion data from earthquakes (*Andrews*, 1982, 1986; *Boatwright et al.*, 1991; *Chin and Aki*, 1991; *Moya et al.*, 2000; *Herrmann and Malagnini*, 2005). Most of these kinds of ground motion studies have used direct P- and S-wave data. The separability and trade-off between different effects are still an active area in seismological research.

Coda waves have also been used to estimate the source, path, and site effects. *Aki* (1969) and *Aki and Chouet* (1975) used seismic coda to study the source and path effects from small local earthquakes. Coda waves are very useful when the direct phases are saturated on the seismogram. Since they sample the whole medium in which they propagate, coda waves can be used to derive an average of the medium properties. Because of this property, unlike the direct phases, there is no need for path correction for coda waves.

While some studies involve empirical regressions, more attempts have been recently made toward using model-based or parametric inversions. Model-based predictions also use parametric models to represent source and propagation effects. The stochastic method is a model based simulation technique, which is widely used in engineering seismology (*Boore*, 1983; *Boore et al.*, 1992a; *Atkinson and Boore*, 1997; *Gregor et al.*, 2002).

2.3.1   Earthquake source scaling

The Brune point source model (*Brune*, 1970) has been widely used in earthquake source scaling. The omega-squared Brune point source model is characterized by a stress drop parameter and seismic moment. The Brune model, however, has some simplifying assumptions and neglects finite fault effects (*Haddon*, 1996). *Boatwright and Choy* (1992) observed that the spectra of large earthquakes depart from the predictions of Brune point source model. *Atkinson* (1993) introduced an empirical two-corner Brune source model for the earthquake source spectrum. A finite fault model was then used to characterize some ground motion features of large earthquakes such as long duration and directivity (*Silva et al.*, 1990; *Silva and Stark*, 1992; *Atkinson and Silva*, 1997).

6

The Brune point source model has been used to estimate the stress drop from earthquake spectra. *Choy and Boatwright* (1990) reported a dynamic stress drop of 51 bars for the Loma Prieta earthquake. *Dewberry and Crosson* (1995) used *S*-coda amplitudes for source scaling and moment estimation of Pacific Northwest earthquakes. They inferred an average stress drop of 32 bars for the region. *Atkinson and Silva* (1997) used a stochastic finite fault model to reduce the discrepancy between observed source spectra and the Brune model prediction. They observed a decrease of stress drop with increasing magnitude for California earthquakes assuming a point source model and stated that the apparent decrease may indicate a saturation effect due to point source distance rather than actual stress on the fault.

In most studies the source spectra are empirically derived using a linear regression technique. Non-linear inversion methods are also used to derive source spectra. *Moya et al.* (2000) used a genetic algorithm technique and obtained a range of stress drops from 10 to 100 bar for some aftershocks of the 1995 Kobe, Japan, earthquake assuming a Brune source model. The idea of an average constant stress drop (*Kanamori and Anderson*, 1975) is under debate. For example, the moderate shallow earthquakes of California are more consistent with an average stress drop of 100 bars (*Atkinson and Silva*, 1997, 2000), while shallow Cascadia events are reported to have a lower stress drop (*Atkinson*, 1995). In a recent study *Malagnini et al.* (2004) estimated the dynamic stress drop for some regional earthquakes recorded by the seismic network of eastern Alps and concluded that events in that region show a non-constant dynamic stress drop. They attributed the difference in stress drops to the difference in rupture dynamics between small and large events.

2.3.2   Attenuation

Attenuation relationships express a ground motion parameter (velocity, acceleration, etc.), as a function of earthquake magnitude and distance from epicenter. Proposed relationships include some other parameters such as local site effects and focal mechanisms. An attenuation relationship is one of the important inputs of a seismic hazard assessment procedure (*Cornell*, 1968; *Reiter*, 1991; *Kramer*, 1996).

Many researchers in the past have used the regression of available strong motion data sets from earthquakes to empirically derive attenuation relationships for use in seismic hazard studies. *Campbell* (1981) used worldwide data to predict an attenuation relationship for the mean peak horizontal acceleration (PHA):

$$\ln PHA(g) = -4.141 + 0.868M - 1.09 \ln[R + 0.606 \exp(0.7M)] \qquad (2.1)$$

He used magnitude 5.0 to 7.7 earthquakes for sites with maximum distance of 50 km from the earthquake. As more data became available, the attenuation relationships became more refined. *Campbell and Bozorgnia* (1994) used worldwide accelerograms from earthquakes of moment magnitude 4.7 to 8.1 to develop a new attenuation relationship. Their predictive relationship included some parameters to account for the earthquake mechanism and site conditions. Those parameters however, were constant numbers. *Boore et al.* (1992a) derived a relationship for western North America using earthquakes of magnitude 5.0 to 7.7. *Toro et al.* (1995) developed a relationship for eastern North America. *Joyner and Boore* (1998) derived a relationship for peak horizontal velocity (PHV). Most of those relationships were derived for local distances using strong ground motion data of earthquakes with magnitudes larger than 4.5. As a result these types of relationships were developed for only a few regions of the US and the world. One solution to the problem was to apply a world-wide-based relationship to another specific region. *Atkinson and Boore* (2003) used a large global data set of subduction zone earthquakes to derive empirical ground motion relations and concluded that applying the results of their study to a specific region is not well justified.

Using small earthquake motions in the regions where enough strong motion data are not available is useful to study wave propagation properties and regional attenuation. The results then can be used with random vibration theory to predict the peak motions from large earthquakes (*Herrmann and Malagnini*, 2005). The procedure includes the regression of ground motion data for each frequency in the frequency band of interest to yield the excitation, propagation, and the site terms. The results

can be interpreted using a theoretical model to estimate source and propagation parameters. The methodology which was first developed for New Madrid, Missouri, was then applied to some other regions of the US and the world for strong motion scaling. *Jeon* (2003) applied the methodology to the Utah region. *Malagnini et al.* (2000b) used the same technique for data from Germany. The same method was used to constrain high frequency ground motion in southern California, Appennines, Italy, and central Mexico (*Malagnini et al.*, 2000a; *Ortega*, 2000; *Raoof et al.*, 1999).

The coda normalization method (*Aki*, 1980; *Frankel et al.*, 1990; *Sato and Fehler*, 2000) has been used to study the attenuation of shear waves. At a single station, the method which removes the effect of the source, the instrument, and the site amplification, is useful where the instrument response is not available or correct. The method can combine data from many earthquakes to find a stable estimate of attenuation.

2.3.3   Site effects

Different soil types have different responses to the ground motion from earthquakes. Compared to the hard rock and stiff soils, softer soils usually amplify the ground motion from earthquakes (*Kramer*, 1996; *Stiedl et al.*, 1996). The effect of amplification at a site due to the surface geology and subsurface structure should be considered in ground motion evaluation at that site. To estimate the site amplification at a site, usually a nearby rock site is used as the reference site. In this method it is assumed that reference rock site has a flat response of unit amplitude (or a negligible response) so that it doesn't distort the incoming signal (*King and Tucker*, 1984; *Field and Jacob*, 1995). Other authors, however, pointed out that rock sites can disturb the seismic waves in some frequency ranges and the assumption of the method is not always valid (*Stiedl et al.*, 1996; *Moya et al.*, 2000). Rock sites and soft soil sites can also deamplify the amplitude of the seismic waves. Amplification and deamplification due to surface geology of a particular site is a complex phenomenon which also depends on the frequency and the level of the ground motion.

Significant amplification of seismic waves was observed during the 1985 Michocan,

9

Mexico (*Singh et al.*, 1988), 1988 Armenian (*Borcherdt et al.*, 1989), and 1989 Loma Prieta earthquakes (*Hough et al.*, 1990). Study of structural damage caused by those events showed that local site effects are more important than previously assumed, and that assessing the effects of the sedimentary deposits (amplification) is very important in the earthquake resistant design of critical facilities.

Field and Jacob (1995) reviewed several methods of site-response estimation. The spectral ratio method (*Borcherdt*, 1970) is the most common method in which the spectrum at the site of interest is divided by the spectrum at a nearby hard rock (reference) site. The method gives a good estimate of site response if the two sites have similar source and path effects and if the reference site has a negligible response. Receiver function type estimates are another method of site response estimation in which one takes the horizontal to vertical component ratios of the shear wave spectrum (*Langston*, 1979; *Lermo et al.*, 1993). *Andrews* (1986) applied a generalized inversion approach to simultaneously solve for the source and site terms. The non-reference-site methods, such as the receiver function type method, are useful in regions that do not have good reference sites.

Spectral ratios of coda amplitudes have also provided reliable estimates of site effects. The site amplification for coda waves is assumed to be the same as the direct S-waves. This was first demonstrated by *Tsujiura* (1978) who showed that coda amplitude ratios between two sites are comparable with their S-wave amplitude ratios. Since then many researchers have used the same technique to estimate S-wave site amplifications using coda waves (*Phillips and Aki*, 1986; *Mayeda et al.*, 1991; *Kato et al.*, 1995).

2.4 Parameterization and modeling of strong ground motion

Aki (1967) derived a theoretical scaling law for the seismic spectrum based on a dislocation model (*Haskell*, 1964) of the seismic source. The similarity assump-

tion in Aki's work implied a constant stress drop model. *Brune* (1970) modeled an earthquake source by applying a stress pulse to the dislocation surface neglecting the fault propagation effects. Aki's scaling (based on the Haskell's kinematic dislocation model) and the Brune's dynamic model have been used by many seismologists for earthquake ground motion prediction (*Boore*, 2003).

Stochastic modeling of strong ground motion (*McGuire and Hanks*, 1980; *Hanks and McGuire*, 1981; *Boore*, 1983) is widely used in engineering seismology. It combines the seismological models of predicted seismic amplitude spectrum with random processes based on the fact that large parts of observed strong motions on the seismogram are random and incoherent. Using the stochastic method, the ground motion at a site can be obtained by either time domain simulations or estimates of peak motions (using random vibration theory) for a given amplitude spectrum (*Boore*, 1983, 2003).

In summary, deriving empirical predictive relationships for seismic wave attenuation requires a large data set of strong motion from moderate to large earthquakes. Those data sets are not available for many regions because of low seismic activity and poor distribution of seismic stations. Applying a relationship developed for a specific region to another region may overestimate or underestimate the ground motion level. This is also true for stress drop, since the stress drops derived from data in a region may significantly differ from those of another region. Large data sets of locally recorded small earthquakes in places with dense distribution of seismic stations can be used to characterize the source, propagation, and site terms. The results can be modeled to estimate the source and propagation parameters such as stress drop and seismic Q. Since the seismic networks in the study areas have both short period and broad band instruments, the study results can cover a wide frequency range.

Chapter 3

Theory and Method

3.1 Earthquake source models

Stress drop is the most important parameter in the faulting process during an earthquake. It is defined as the difference between the stress at a point on the fault before and after the earthquake (*Lay and Wallace*, 1995). Stress drop can be determined from the seismic spectrum. Models of seismic source spectrum include kinematic and dynamic models.

3.1.1 Haskell's model–rectangular fault

A simple kinematic model of a rectangular fault was developed by *Haskell* (1964). Slip (Δu) propagates only along the length L with a constant velocity. Fracture can be modeled either as unilateral or bilateral. The displacement amplitude spectrum based on the Haskell's model consists of a flat part at low frequencies and a decreasing (ω^{-2}) part after a certain corner frequency. The source time function in ithe Haskell's model includes a rise time; the amplitude spectrum shape is a combined effect of the source dimensions and the rise time. Finite dimension of the source and rise time have been validated by studying the observed spectra of seismic sources (*Aki*, 1967; *Lay and Wallace*, 1995; *Udias*, 1999).

3.1.2 Brune's model–circular fault

Brune (1970) developed a simple model of an extended circular seismic source known as the Brune's model. Brune's model consists of a circular fault in a homogeneous material. An initial stress σ_0 is applied to the fault surface instantaneously so

there is no fracture propagation and after time t_1 stress reaches a new value σ_1. So $\Delta\sigma = \sigma_0 - \sigma_1$. The shear pulse generates a propagating shear wave. The direction of propagation is perpendicular to the fault plane. The stress pulse at a distance x is defined as (*Udias*, 1999)

$$\Delta\sigma(x,t) = \Delta\sigma H(t - \frac{x}{\beta}) \,, \tag{3.1}$$

where β is the shear-wave propagation velocity and $H(t)$ is the Heaviside unit step function. At time t, the shear wave pulse will propagate a distance βt into the medium. So the instantaneous strain is

$$\varepsilon = \frac{u(t)}{\beta t} = \left(\frac{\Delta\sigma}{\mu}\right) \,, \tag{3.2}$$

and the shear displacement

$$u(t) = \frac{\Delta\sigma\beta t}{\mu} \tag{3.3}$$

The Fourier transform of (3.3) is

$$U(\omega) = \int_0^\infty \frac{\Delta\sigma\beta t e^{i\omega t}}{\mu} \, dt = -\frac{\Delta\sigma\beta}{\mu\omega^2} \,, \tag{3.4}$$

It is obvious from (3.3) that particle velocity $\dot{u}(t)$ is directly proportional to the stress drop: $\dot{u}(t) = \Delta\sigma(\beta/\mu)$. This indicates that when $\Delta\sigma$ is constant, $\dot{u}(t)$ should also be constant and independent of earthquake. The fault propagation is neglected in Brune's model. In fact the rupture propagates with velocity v_r which results in a reduced particle velocity instead.

The far field source spectrum predicted by Brune's model is

$$U(\omega) = \frac{\Delta\sigma\beta}{\mu}\frac{1}{\omega^2 + b^2} \,, \tag{3.5}$$

where $b = 2.33\beta/a$ and a is the radius of the fault. Equation (3.5) does not include the radiation pattern and distance effects. Like the Haskell's model the spectrum (3.5) has a flat part at low frequencies and decreases as ω^{-2} at higher frequencies.

Fault dimensions are usually obtained from the spectra of S waves. For this

purpose, the Brune's circular model is a good approximation for small to moderate earthquakes. For larger earthquakes which the length of the fault is larger than the width, the Haskell's model is a better approximation (*Udias*, 1999).

3.1.3 Stress drop and fault dimensions

Stress drop for an earthquake can be estimated if the seismic moment and the fault dimensions have already been measured for that earthquake. For a circular fault of radius R

$$\Delta\sigma = \frac{7}{16}\frac{M_0}{R^3} \,, \tag{3.6}$$

for a rectangular strike-slip fault with length L and width w

$$\Delta\sigma = \frac{2}{\pi}\frac{M_0}{w^2 L} \,, \tag{3.7}$$

and for a rectangular dip-slip fault

$$\Delta\sigma = \frac{4(\lambda+\mu)}{\pi(\lambda+2\mu)}\frac{M_0}{w^2 L} = \frac{8}{3\pi}\frac{M_0}{w^2 L} \,, \tag{3.8}$$

where the last equation assumes $\lambda = \mu$.

If fault dimensions are known by another method this process is straightforward. For large earthquakes the fault dimensions may be determined from the aftershock area. For smaller earthquakes it is harder to measure the fault dimensions and a small error in fault length or fault radius will result a big error in stress drop (*Lay and Wallace*, 1995; *Stein and Wysession*, 2003). The fault dimensions can also be measured from spectrum of the seismic waves by identifying the corner frequency and then estimating the rupture time of the earthquake. The product of the rupture time and the rupture velocity yields the fault dimension. The rupture velocity and the fault geometry however, should be assumed in this method and again a small uncertainty in the rupture time will result a large uncertainty in the stress drop. In some cases identifying the corner frequency from the spectrum is not easy because of free surface and site effects. For shallow earthquakes, later arriving reflections overlap

14

and give a spectrum different from the source pulse (*Langston*, 1978). Amplification of some frequencies due to site effects can also hide the real corner frequency of the spectrum (*Moya et al.*, 2000).

3.2 Ground motion spectrum

In a simple medium, the acceleration spectrum of horizontally polarized shear waves (*Boore*, 1983) at a distance R from the source is

$$A(\omega) = C M_0 S(\omega, \omega_c) P(\omega, \omega_m) \frac{e^{-\omega R/2Q\beta}}{R} \, , \tag{3.9}$$

where C is a constant, M_0 is the seismic moment, $P(\omega, \omega_m)$ is a high cut filter, and $\frac{e^{-\omega R/2Q\beta}}{R}$ is the path term. The constant C is given by

$$C = \frac{R_{\theta\phi} \cdot FS \cdot PRTITN}{4\pi\rho\beta^3} \, . \tag{3.10}$$

$R_{\theta\phi}$ is the radiation pattern, FS is the free surface amplification, $PRTITN$ accounts for partitioning of energy into two horizontal components, ρ is the density and β is the shear wave velocity. The (acceleration) source spectrum, $S(\omega, \omega_c)$ (*Brune*, 1970) is given by

$$S(\omega, \omega_c) = \frac{\omega^2}{1 + (\omega/\omega_c)^2} \, . \tag{3.11}$$

The high cut filter P accounts for the sharp decrease in acceleration spectrum after a cutoff frequency ω_m and has the following form

$$P(\omega, \omega_m) = [1 + (\omega/\omega_m)^{2s}]^{-1/2} \, , \tag{3.12}$$

where s controls the decay rate at high frequencies. Since both Q and $P(\omega, \omega_m)$ reduce high frequencies, only one of these terms appears.

Seismic moment M_0 and corner frequency f_c– two parameters that control the

spectra of different earthquakes– are related through the following equation

$$f_c = 4.9 \times 10^6 \beta (\Delta\sigma/M_0) \tag{3.13}$$

where f_c is in Hertz, β is in km/sec, $\Delta\sigma$ is in bars and M_0 is in dyne-cm (*Brune*, 1970).

3.3   Stochastic modeling

Ground motions at a site can be obtained by either time domain simulation or estimates of peak motions using random vibration theory for a given spectrum of motion (*Boore*, 2003). For a specified amplitude spectrum it is possible to generate a time series whose spectrum matches the specified one using the stochastic method. *McGuire and Hanks* (1980) and *Hanks and McGuire* (1981) combined the seismological models of radiated spectra with the concept that high frequency ground motion is random in nature. They used Parseval's theorem to predict the rms acceleration (a_{rms}) and then related this rms value to the peak ground acceleration (a_{rms}), using the results from random vibration theory.

3.3.1   Time domain modeling

In the stochastic approach the ground motion is modeled as bandlimited finite duration white Gaussian noise (*Hanks and McGuire*, 1981; *Boore*, 1983). The time domain simulation begins with the generation of a windowed time series of Gaussian noise with zero mean. The variance should be chosen in a way that the average spectral amplitude of the noise will be equal to one. The noise is generated for a duration which is given by the duration of the motion. The spectrum of the noise is then multiplied by the specified spectrum and then will be transformed back into the time domain to give the final time series. *Boore* (2003) mentions that the order of windowing and filtering of white noise is important. If the filtering is done first, the long period part of the motion will be distorted. In practice the simulation is repeated

many times and the average of all simulated spectra's is computed. So the essential inputs of the stochastic method are the spectrum and the duration of the ground motion. A Brune ω^2 source model (equation 3.11) is usually used in simulations. The effects of all other parameters (3.9), are then applied to the simulated spectrum. Defining the duration of motion may require fitting the observed duration to a linear interpolation function.

3.3.2 Random vibration theory

Random vibration theory (RVT) is used to estimate the peak motion (a_{max}, v_{max}, etc) from the spectral shape and duration of motion as a faster alternative to time domain simulation. It provides an estimate of the ratio of peak motion (a_{max}) to rms motion (a_{rms}). *Cartwright and Longuet-Higgins* (1956), (CLH), derived the equations for the ratio of peak to rms motion.

The rms acceleration (a_{rms}) can be obtained from the integral of the squared acceleration spectrum using the Parseval's theorem

$$\int_{-\infty}^{\infty} |a(t)|^2 \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |A(\omega)|^2 \, d\omega, \tag{3.14}$$

where $a(t)$ is the acceleration time history and $A(\omega)$ is its Fourier amplitude spectrum. Assuming that the most significant motion occurs in the shear wave arrival window and that spectral amplitudes will be cut off at $f = f_{max}$ (*Hanks*, 1979) by anelastic attenuation

$$\int_{0}^{T_d} |a(t)|^2 \, dt = \frac{2}{2\pi} \int_{2\pi f_0}^{2\pi f_{max}} |A(\omega)|^2 \, d\omega, \tag{3.15}$$

where $t = 0$ corresponds to the shear wave arrival (R/β) and T_d is the duration of the shear wave signal. The rms acceleration is then

$$a_{rms} \equiv \left[\frac{1}{T_d} \int_0^{T_d} |a(t)|^2 \, dt \right]^{1/2}. \tag{3.16}$$

Assuming that

$$A(\omega) = \Omega_0 (2\pi f_0)^2 \qquad f_0 \leq f \leq f_{max}$$

, and using the approximation that $f_0 = 1/T_d$, the rms acceleration becomes

$$a_{rms} = 2^{1/2}(2\pi)^2 \Omega_0 f_0^3 (f_{max}/f0)^{1/2}, \tag{3.17}$$

where Ω_0 is the spectrum level at low frequencies. The result of (3.17) holds for $f_{max} \gg f_0$. For the Brune scaling $\Delta\sigma = 106\rho R\Omega_0 f_0^3$ (*Hanks and Thatcher*, 1972), the substitution in (3.17) gives

$$a_{rms} = 0.85 \frac{2^{1/2}(2\pi)^2}{106} \frac{\Delta\sigma}{\rho R} \left(\frac{f_{max}}{f_0}\right)^{1/2}. \tag{3.18}$$

The factor 0.85 accounts for the combined effect of free surface amplification, partitioning of horizontal motion, and the radiation pattern (*Hanks*, 1979).

Hanks and McGuire (1981) used an equation based on random vibration theory (*Vanmarcke and Lai*, 1980) to predict the peak acceleration from the rms acceleration

$$\frac{a_{max}}{a_{rms}} = [2 \ln(N)]^{1/2}, \tag{3.19}$$

where N is the number of extrema in the time interval T. Equation (3.19) is based on an assumption that time series is stationary with uncorrelated peaks. If $\tilde{f}$ is the predominant frequency of the motion, in general

$$N = 2\tilde{f}T. \tag{3.20}$$

Equation (3.19) is a valid approximation for large N. For smaller N, CLH equation 6.8 gives an expression for the peak to rms ratio. *Boore* (1983) expanded the integrand of their integral in terms of binomial series and did a term by term integration

$$\frac{E(a_{max})}{a_{rms}} = \sqrt{\frac{\pi}{2}} \sum_{l=1}^{N} (-1)^{l+1} \frac{C_l^N}{\sqrt{l}} \xi^l, \tag{3.21}$$

where $E(a_{max})$ is the expected value of the largest extrema of acceleration and C_l^N

are the binomial coefficients $(= N!/l!(N-l)!)$. ξ is given by

$$\xi = m_2/(m_0 m_4)^{1/2}. \tag{3.22}$$

In fact ξ is a measure of the bandwidth of the spectrum. m's are the moments of the energy density spectrum, where the kth moment is defined as

$$m_k = \frac{1}{\pi} \int_0^\infty \omega^k |A(\omega)|^2 \, d\omega. \tag{3.23}$$

CLH (1956) derived an asymptotic expression for large values of N

$$\frac{E(a_{max})}{a_{rms}} = [2 \ln(N)]^{1/2} + \gamma/[2 \ln(N)]^{1/2}, \tag{3.24}$$

where γ is the Euler's constant $(= 0.5772...)$. Equation (3.19) is the first term of this equation. If we have a_{rms}, we need T and N to use equation (3.24). The usual estimate of T is the inverse of the source corner frequency and N is estimated using (3.20) which in turn needs the predominant frequency $\tilde{f}$. The appropriate frequency when N is the number of extrema,($Boore$, 1983, 2003), is given by

$$\tilde{f}_e = \frac{1}{2\pi} \left(m_4/m_2\right)^{1/2}. \tag{3.25}$$

After choosing the appropriate duration T, the relation between maximum amplitudes and rms amplitudes depends only on moments of ground motion spectrum. $Boore$ (1983) compared the ground motion measures predicted by stochastic time domain modeling and RVT. His results show that the two methods agree well for large N and the RVT results departs from the time domain simulations for small values of N.

3.4 Regression method

3.4.1 General regression for peak filtered motion or Fourier velocity

In the frequency domain observed ground motion can be considered as the multiplication of independent source, site, and path effects. The relationship is additive in the logarithmic sense, so the observed logarithm of ground motion, $A(r, f)$, can be written as

$$A_{ij}(r, f) = E_i(r_{ref}, f) + S_j(f) + D(r_{ij}, f). \tag{3.26}$$

where r is the hypocentral distance, f is the observed frequency, i is the source index, j is the site index, r_{ref} is a reference distance, and E, S, and D are the source excitation, site, and distance (path) functions. The propagation term $D(r, f)$, can be modeled in terms of geometrical spreading $g(r)$ and the frequency dependent anelastic attenuation, $Q(f)$. The term excitation is used here because the regression results do not define any of the source parameters directly. The function $D(r)$ in (3.27) is approximated by piecewise linear segments with a condition of continuity (*Anderson and Lei*, 1994; *Harmsen*, 1997). Using this interpolation function the $D(r)$ is described in terms of the values at N nodes as

$$D(r, f) = \sum_{k=1}^{N} L_k(r) \, D_k. \tag{3.27}$$

where N is the number of nodes and $L_k(r)$ is a linear interpolation function defined as

$$L_k(r) = \begin{cases} \frac{r - r_{k-1}}{r_k - r_{k-1}} & : & if \, r_{k-1} \leq r \leq r_k \ and \ 2 \leq k \leq n \\ \frac{r_{k+1} - r}{r_{k+1} - r_k} & : & if \, r_k \leq r \leq r_{k+1} \ and \ 1 \leq k \leq n-1 \\ 0 & : & otherwise \end{cases}$$

and $D_k \equiv D(r_k, f)$ are node values. The N coefficients are determined by the inversion of each sampling frequency. The nodes are chosen at 10, 20, 30, 40, 50, 75, 90, 105, 120, 135, 150, 175, 200, 250, 300, 400, 500, 600, 700, 800, 900, and 1000 km. The

ground motion regression model is then

$$A_{ij}(r, f) = E_i(r_{ref}, f) + S_j(f) + \sum_{k=1}^{N} L_k(r) \, D_k. \qquad (3.28)$$

Having a set of observations of ground motion at different stations we can construct a matrix form of (3.29). The system of equations can then be solved for the excitation, propagation, and site terms using a linear inversion. A damped least square method with a singular value decomposition (SVD) technique (*Lawson and Hanson*, 1974; *Menke*, 1989) can then be used if we add some constraints to the system of equations. These constraints make the inversion more stable by reducing the number of degrees of freedom. The constraints used here are:

- $D(r_{ref}) = 0$, where r_{ref}=40 km. The reference distance r_{ref} is selected to be in the range of the observed distances. In this case we interpolate within the data set not extrapolate beyond it. It should be large enough to reduce the effect of source depth error on the hypocentral distance and yet not so far that the Moho reflections complicate the motions

- $\sum_j S_j(f) = 0$ for selected data channels.

- $D(r)$ is smooth.

Anderson and Lei (1994) used a linearity constraint on $D(r)$ by requiring the numerical second derivative estimate to be zero

$$D_{l-1} - 2D_l + D_{l+1} = 0$$

. The effectiveness of this constraint can be judged by examining the regression residuals of fit as function of distance.

Given the level of motion $E_j(r_{ref}, f)$ at $r_{ref} = 40$ km, The $D(r, f)$ propagates that motion to the desired distance, r, and the site term adjusts that motion to a particular physical location.

The general regression is done for observed peak filtered velocity and Fourier velocity data.

3.4.2 Implementing general regression

The following linear equation is used for the observed amplitude at a distance $r_k \leq r \leq r_{k+1}$

$$\log A_l(f) = E_i(f) + p\, D_k(f) + (1-p)\, D_{k+1}(f) + S_j(f) \qquad (3.29)$$

for $r_k \leq r \leq r_{k+1}$:

$$p \equiv \frac{r_{k+1} - r_0}{r_{k+1} - r_k}$$

where l is the number of observations replacing ij, $1 \leq l \leq n$, i is the number of earthquakes, $1 \leq i \leq ne$, k is the number of distance terms, $1 \leq k \leq m$, and j is the number of sites, $1 \leq j \leq ns$. The matrix form representation of (3.30) is then

$$A_{(n+m+1)\times 1} = B_{(n+m+1)\times (ne+m+ns)}\, M_{(ne+m+ns)\times 1}, \qquad (3.30)$$

where A_l is the logarithm of the observed processed peak values of the filtered time histories or Fourier spectra. The constraints (section 3.4.1) are applied to the system of equations by adding rows in the matrices. In the expanded matrix form, the rows above the first dotted lines are related to the data, the rows between first and second dotted lines refers to the constraint to the site terms, the rows between second and third dotted lines applied to the smoothing of the attenuation term, and the last row is used to force the reference distance at the proportional functional to be zero.

$$
\begin{pmatrix}
A_1 \\
\vdots \\
A_n \\
\cdots \\
0 \\
\cdots \\
0 \\
\vdots \\
\vdots \\
0 \\
\cdots \\
0
\end{pmatrix}
=
\left(
\begin{array}{cccc|ccccc|cccc}
 & \text{Source terms} & & & & \text{Distance terms} & & & & & \text{Site terms} & & \\
1 & 0 & \cdots & 0 & 0 & p & 1-p & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & p & \cdots & 0 & 0 & 0 & \cdots & 1 \\
\cdots & & & & \cdots & & & & & \cdots & & & \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & u & u & \cdots & u \\
\cdots & & & & \cdots & & & & & \cdots & & & \\
0 & 0 & \cdots & 0 & -2w & w & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & w & -2w & w & \cdots & 0 & \vdots & \vdots & & \vdots \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & w & -2w & 0 & 0 & \cdots & 0 \\
\cdots & & & & \cdots & & & & & \cdots & & & \\
0 & 0 & \cdots & 0 & 0 & w_{ref} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{array}
\right)
$$

$$
\times
\begin{pmatrix}
SRC_1 \\
\vdots \\
SRC_{ne} \\
\cdots \\
D_1 \\
\vdots \\
\vdots \\
D_m \\
\cdots \\
SITE_1 \\
\vdots \\
SITE_{ns}
\end{pmatrix}
$$

3.4.3 Coda normalization method

The coda normalization technique can be used to estimate the source, path and site terms (*Sato and Fehler*, 2000). *Aki* (1980) used this technique to determine the shear wave Q in the crust. For a specific event, this technique divides the amplitude of the direct S Wave by the the amplitude of the S coda (*Frankel et al.*, 1990). The amplitude of the coda is computed for a lapse time equal or greater than twice of the direct S arrival time. The spectral amplitude of the S wave at a specific station is given by

$$A_s(\omega, r) = R_{\theta\phi}I(\omega)S(\omega)G(\omega)r^{-\gamma}e^{-\omega R/2Q\beta}, \tag{3.31}$$

where r is the hypocentral distance, $R_{\theta\phi}$ is the radiation pattern, β is the S wave velocity, $I(\omega)$ is the instrument response, $S(\omega)$ is the source term, and $G(\omega)$ represents the site amplification. γ is the exponent of the geometrical spreading term.

The coda wave amplitude spectrum can be expressed as

$$A_C(\omega, t_c) = I(\omega)S(\omega)G(\omega)C(\omega, t_c), \tag{3.32}$$

where $C(\omega, t_c)$ represents the coda spectral amplitude decay and is called the coda envelope shape (*Frankel et al.*, 1990). It is usually assumed that the site amplification terms are the same for direct and coda (S) waves (*Tsujiura*, 1978). Furthermore, for a given region the coda level is independent of the source receiver distance (*Rautian and Khalturin*, 1978), after some lapse time. This means that we can use the same coda level for a specific earthquake in different stations. Dividing (3.32) by(3.33) then yields

$$A_s(\omega, r)/A_C(\omega, t_c) = R_{\theta\phi}r^{-\gamma}e^{-\omega r/2Q\beta}/C(\omega, t_c) = D(r, \omega) - C(\omega, t_c). \tag{3.33}$$

The division removes the effect of the source and the instrument but not the radiation pattern. This method is useful in cases which there are some unknown instrument responses or when the instrument response is not reliable. For each frequency the normalized amplitude (3.34) can be plotted for different earthquakes and different

stations. The normalized amplitude decays with distance and the decay can be mod-
eled as a function of Q and geometrical spreading. This method needs enough coda
length after twice the direct S arrival time. I will use a regression technique to de-
rive the distance term based on the coda normalization method. This independent
estimate of the distance term will then be compared to the results from the general
regression of direct S waves.

3.4.4   Implementing regression for coda normalization

The equation (3.34) can be written as

$$A_{normalized}(r, f) = \frac{A_s(r, f)}{A_C(f, t_s)}. \tag{3.34}$$

taking the logarithm of both sides gives

$$\log A_{normalized}(r, f) = D(r, f) - C(\omega, t_s), \tag{3.35}$$

where $C(\omega, t_s)$ is the coda envelope that describes the decay of coda with a reference
lapse time t_s. Combining (3.28) and (3.36) gives

$$\log[A_{normalized}(r, f)] = L_K(r) \cdot D_k \cdot + L_{k+1}(r) \cdot D_{k+1} - C(f, t_s). \tag{3.36}$$

If we define $A_i(r) = \log[A_{normalized}(r, f)]$, Equation (3.37) becomes

$$A_i(r) = p \cdot D_j + (1 - p) \cdot D_{(j+1)} - C(f, t_s). \tag{3.37}$$

For $r_k \leq r \leq r_k + 1$

$$A_{(n+m+1)\times 1} = G_{(n+m+1)\times m} \cdot M_{(m)\times 1}, \tag{3.38}$$

where n is the number of observations and m is the number of distance terms. The
smoothness and $D(r_{ref}, f) = 0$ constraints will be added to the system of equations.
The reference distance will be chosen to be 40 km for the same reasons mentioned

25

before in section (3.4.1).

The equation (3.39) can be written as the following matrix form

$$
\begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \\ \cdots \\ 0 \\ \vdots \\ 0 \\ \cdots \\ 0 \end{vmatrix}
=
\begin{vmatrix} 0 & p & 1-p & 0 & \cdots & 0 \\ p & 1-p & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & p & 1-p & \cdots & 0 \\ \hdotsfor{6} \\ -2w & w & 0 & 0 & \cdots & 0 \\ w & -2w & w & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & w & -2w \\ \hdotsfor{6} \\ 0 & 0 & w_{(ref)} & 0 & \cdots & 0 \end{vmatrix}
\times
\begin{vmatrix} D_1 \\ \vdots \\ \vdots \\ D_2 \\ \vdots \\ \vdots \\ D_3 \\ \vdots \\ \vdots \\ \vdots \\ D_m \end{vmatrix} ,
$$

where dotted lines are inserted to separate the observation data from the constraint entries. The constraints will be used more effectively by using the weight parameter w. Applying this technique, yields an initial $D(r, \omega)$ which is an independent, unbiased estimate of the true $D(r, \omega)$.

3.5 Modeling the regression results

The regression is applied to each processed frequency to yield the excitation, path, and site terms. The results then should be interpreted in terms of a theoretical model to estimate the related parameters of each term. An expression for the predicted Fourier velocity spectra for frequency f and distance r is

$$\alpha(r, f) = s(f, \mathbf{M}_W) g(r) e^{-\pi f r / Q(f)\beta} V(f) e^{-\pi f \kappa}, \tag{3.39}$$

where $\alpha(r, f)$ is the Fourier velocity spectra, $s(f, \mathbf{M})$ is the source excitation as a function of moment magnitude, $g(r)$ is the geometrical spreading function, $Q(f)$ is

the frequency dependent quality factor which equals $Q_0(f/1.0)^\eta$, Q_0 is the quality factor at 1.0 Hz, $V(f)$ is a frequency dependent site amplification, and κ controls site dependent attenuation of high frequency. The velocity spectra at a distance of 1 km is defined as

$$s(f, \mathbf{M}) = \frac{CM_0 2\pi f}{1 + (f/f_c)^2)}$$

, where M_0 is the seismic moment, f_c is the source corner frequency, and C is a constant (Boore, 1983).

A comparison of regression parameters of (3.27) with this formula and applying the constraints shows the association:

$$10^E = s(f, \mathbf{M}) g(r_{ref}) e^{-\pi f r_{ref}/Q(f)\beta} \, \overline{V(f)} \, e^{-\pi f \overline{\kappa}} \, , \tag{3.40}$$

$$10^D = \frac{g(r) e^{-\pi f r/Q(f)\beta}}{g(r_{ref}) e^{-\pi f r_{ref}/Q(f)\beta}} \, , \tag{3.41}$$

$$10^{S_i} = \frac{V(f) \, e^{-\pi f \kappa}}{\overline{V(f) \, e^{-\pi f \kappa}}} \, , \tag{3.42}$$

where $\overline{V(f) \, e^{-\pi f \kappa}}$ is the network average of the site effect. In general the geological characteristics of the sites are not known well. To get around this problem an effective κ, κ_{eff}, is defined for the composite effect of network average site amplification by

$$e^{-\pi f \kappa_{eff}} = \overline{V(f) \, e^{-\pi f \kappa}} \, . \tag{3.43}$$

If $V(0) = 1$, both sides of equation (3.44) approach 1.0. $V(f)$ monotonically increases from 1.0 at low frequencies to a fixed high frequency value, and since we have a limited frequency range to fit a logarithmic function, the equation (3.44) may need a constant factor.

By interpreting the regression results based on the theoretical model (3.40) and the corresponding relationships (3.41-3.44), we characterize the excitation term, E, as the network average motion at a reference distance for a given event, The D term

as a function that propagates this motion to the hypocentral distance of interest, and the S as a term that relates a specific site to the network average site response. It should be noted that the regression results are purely empirical and the forward model does not predict absolute motions unless some independent calibration is used.

Each modeling involves the solving of the forward problem assuming some parameters values (e.g. Q and η) using the stochastic method and fitting the model based term ($D(r)$ in this case) to the corresponding observed term. The procedure is repeated many times and the final value of parameters are selected for the case of the best fit. Visual fitting is usually done since in most cases the parameters can be found without too much effort. In some cases however, especially when the number of parameters increases, the visual fitting becomes more difficult and is not preferred. For the modeling part of this study, I use the genetic algorithm (GA) search method to fit the observed to the predicted terms. I also use visual fitting in some cases for comparison purpose. GA is a global search method that can overcome local minima of the objective function in many cases. It can be used for nonlinear inversion as well as linear inversion problems.

3.6   Error analysis

Since the forward modeling using the stochastic method does not provide any uncertainty or error measurement on the estimated parameters such as Q and stress drop, some other methods need to be used for this purpose. I used a bootstrap technique which is a resampling method and is widely used for error estimation when such an estimate is not provided by the inversion or modeling. Resampling techniques can be designed in a way that do not require the usual assumptions about the data probability distribution. More importantly, these techniques can be used to evaluate various statistical properties that can not be determined analytically (*Tichelar and Ruff*, 1989).

In a resampling technique like the bootstarp, the original data set is resampled with replacement to form a large number of data sets and the multiple estimates of the

model give information on model variance. A bootstrap resample is a random selection of n data from n original data and since the resampling is done with replacement, a certain element of the original data set may appear more than once in a new data set. If $\hat{\theta}_i^\star$ is an estimator of the statistic θ calculated for the bootstrap resample i, the bootstrap estimate of standard deviation of θ is

$$\hat{\sigma}_{BOOT} = \left[\frac{1}{L-1} \sum_{i=1}^{L} (\hat{\theta}_i^\star - \overline{\theta})^\mathsf{T} (\hat{\theta}_i^\star - \overline{\theta}) \right]^{1/2} , \tag{3.44}$$

where $\overline{\theta} = \sum_i^{L} \hat{\theta}_i^\star / L$ (*Efron and Tibshirani*, 1994).

3.7 The procedure

I use large data sets of small to moderate crustal earthquakes in the Washington-Oregon, central, and northern California to derive the excitation, propagation, and the site terms for the defined frequency and distance ranges. Geometrical spreading and Q are determined by matching the observed propagation term to the theoretical one and the stress drop is defined in the same way using the excitation term.

- For each study area the data (digital seismograms) are requested from the network. The data preparation includes previewing each trace to discard the clipped, noisy, or otherwise bad waveforms, removing the instrument response to form the velocity time history, and picking the P and S arrival times.

- Each trace is then bandpass filtered at a center frequency, f_c. A highpass Butterworth filter with corner frequency $f_c/\sqrt{2}$, followed by a lowpass filter with corner frequency $\sqrt{2}f_c$. The center frequencies are 0.25, 0.3, 0.5, 1, 2, 3, 4, 6, 8, 10, 12, 14, and 16 Hz.

- The signal duration is defined as the interval within which the integral of filtered velocity squared following the S arrival changes from 5% to 75% of the maximum.

- Peak filtered velocity following the S arrival and the RMS Fourier velocity spectra of the waveform within the duration window between frequencies $f_c/\sqrt{2}$ and $\sqrt{2}f_c$ are computed. The general regression (3.4.1) is performed on these two data sets to yield the excitation, path, and site terms for peak filtered velocity and Fourier velocity at each center frequency.

- RMS signal level as function of time is computed. The regression of coda waves (3.4.3) is performed on this data set.

- Fourier velocity $D(r)$ (derived from the regression) is fitted in terms of $g(r)$ and $Q(f)$ and then the $g(r)$ and $Q(f)$ choice is used to refine the duration $(T(r))$ values by fitting the $D(r)$ for the filtered peak velocity (section 3.5).

- Fitting the excitation term in terms of stress drop using the previously derived $g(r)$ and $Q(f)$.

- Error estimation of derived parameters from the modeling.

I follow the same sequence for each data set and compare the results of different regions. The results and their implications for seismic hazard is discussed at the end.

Chapter 4
Data

Study areas in this study include: Washington-Oregon, Northern California, and Central California regions. Each region has a large data set of small to moderate earthquakes.

4.1 Washington-Oregon

The study area in the Washington-Oregon ranges from 42°N to 49°N in latitude and from -117° to -125° in longitude (Fig. 4.1).
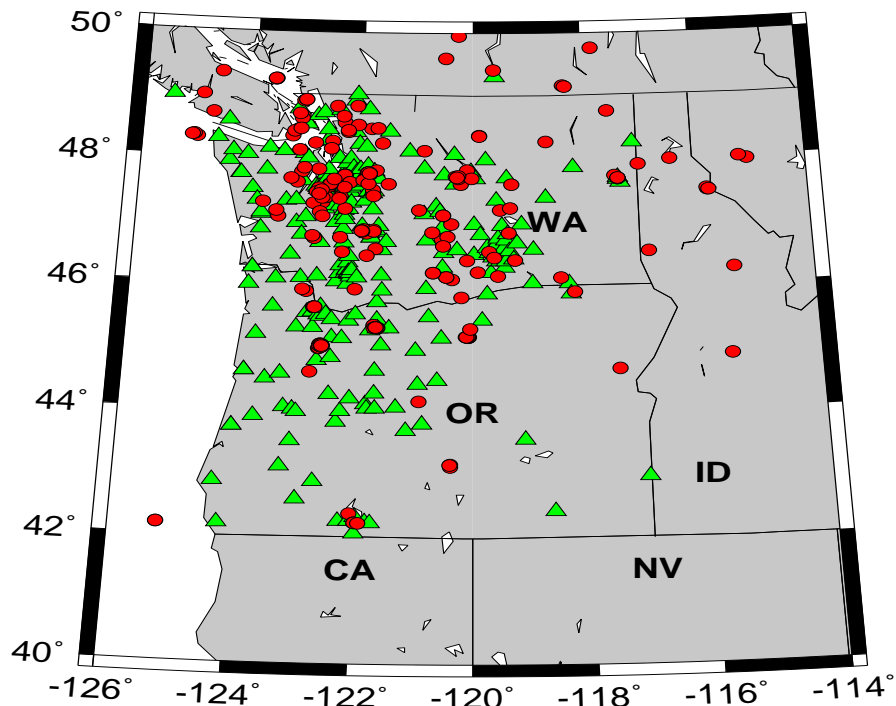
Figure 4.1: Map of earthquakes (circles) and PNSN stations (triangles) in Washington-Oregon.

The data set includes more than 4000 waveforms from 250 earthquakes (M ≥ 3.0)

which occurred during 1993-2003. The data are provided by the Pacific Northwest Seismic Network (PNSN). The PNSN operates more than 100 short period and some broadband stations in the area. PNSN data is used to study earthquake hazard in the Pacific Northwest, volcanic eruptions of Mount St. Helen's, and to determine the location of faults and volcanic magma chambers (www.pnsn.org). Crustal earthquakes are the primary focus of this study but there might be some deeper Benioff zone earthquakes in the data set.

4.2   Northern California

The Northern California study area ranges from 37°N to 42°N in latitude and from -119° to -124° in longitude. The data set from Northern California seismic network (NCSN) and Berkeley Digital Seismic Network (BK) includes more than 1900 waveforms from 240 earthquakes (M ≥ 3.0) from 1995 to 2005 (Fig. 4.2).
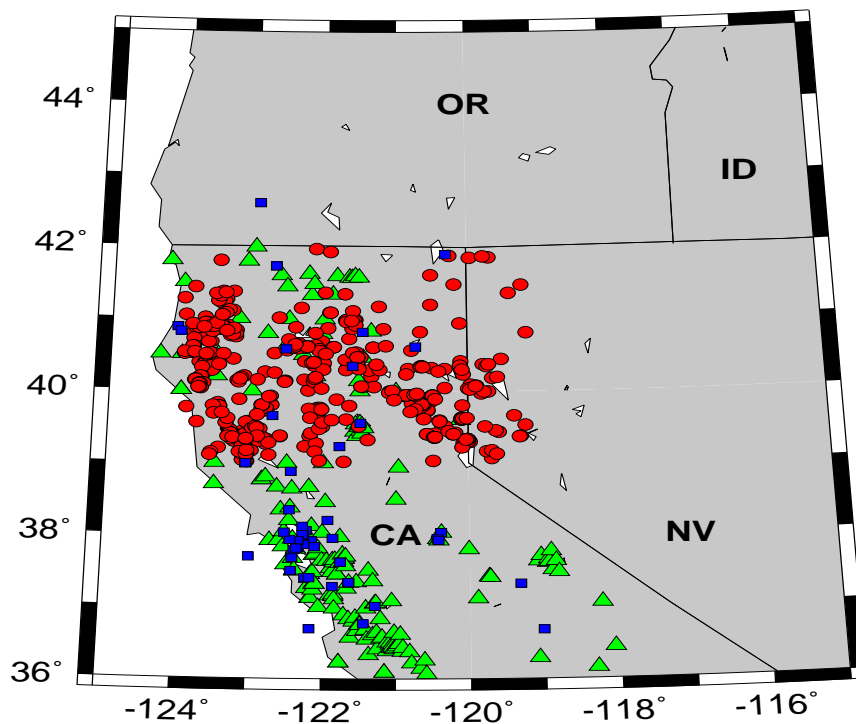
Figure 4.2: Map of earthquakes for this study (circles), NC short period (triangles), and BK broadband (squares) stations in northern California.

Recording channels and their sampling frequencies for the BK and NC networks

are given in Table 4.1.

Sensor	Channels	Rate (samples/sec)	Sampling Mode
Broadband seismometer	BHE, BHN, BHZ	20	Continuous
Broadband seismometer	HHE, HHN, HHZ	80	Continuous
Short-period seismometer	EHE, EHN, EHZ	100	Continuous

Table 4.1: Recording channels of BK and NCSN networks.

4.3   Central California

The map of central California study area is shown in Fig. 4.3. Two distinct clusters of earthquakes can be seen from the map. Since it is interesting to compare the wave propagation and source parameters of the western and eastern parts, I decided to split the data set and study each part separately.
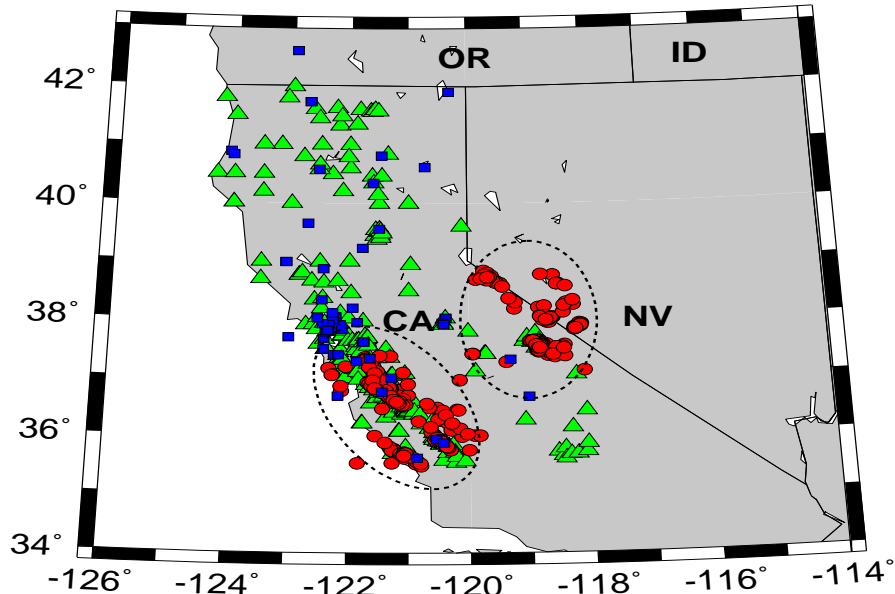
Figure 4.3: Map of earthquakes (this study) and the stations in central California. Ellipses show the earthquakes for western and eastern parts of the the study area.

The western part data set include 2660 waveforms from 230 earthquakes. The eastern part data set consists of 1850 waveforms from 240 earthquakes. the earthquakes of both data sets occurred from 1995 to 2005 and have magnitudes equal to or larger than 3.0.

Chapter 5

Ground motion scaling in

Washington-Oregon

The data set of Washington-Oregon earthquakes (Fig. 5.1) are mostly vertical short period seismograms which cover a distance range of up to 500 km. The obser-
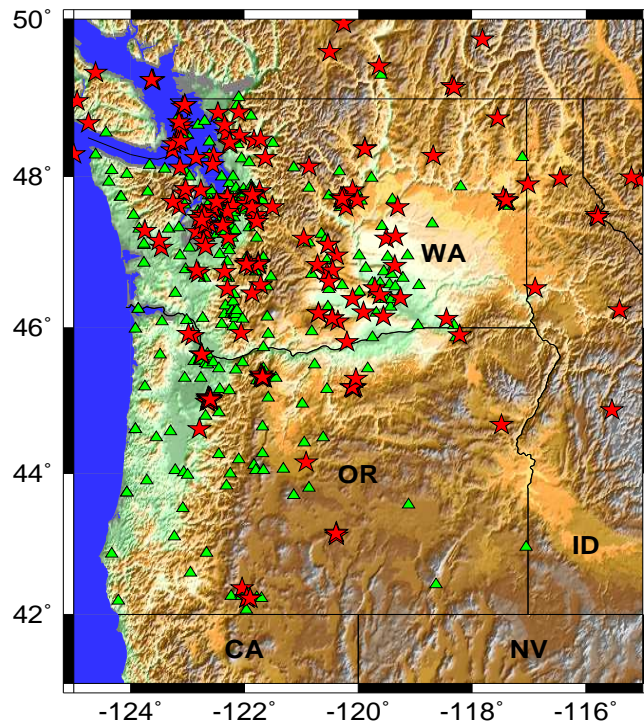
Figure 5.1: Map of stations (triangles) and events (stars) for this study in Pacific Northwest.

vation distances from the events for each station should overlap with that of other stations. This reduces the trade off between the site and the propagation terms in the regression. Since there are only a few broadband horizontals in the data set, I have done the regression and modeling of the whole data set and the vertical but not the horizontal components.

5.1 Data Processing

The coda analysis and regression is first done for the Fourier data. The same procedure is then repeated for the bandpass filtered data. The coda results are used as an independent measurement for comparison purpose.

5.1.1 Coda shape

Figure 5.2 shows the decay shape of the normalized coda at 1.0 Hz as a function of time. All the peak values are normalized to the value of coda at a reference level. A 256 point moving average window is applied to compute the RMS average.
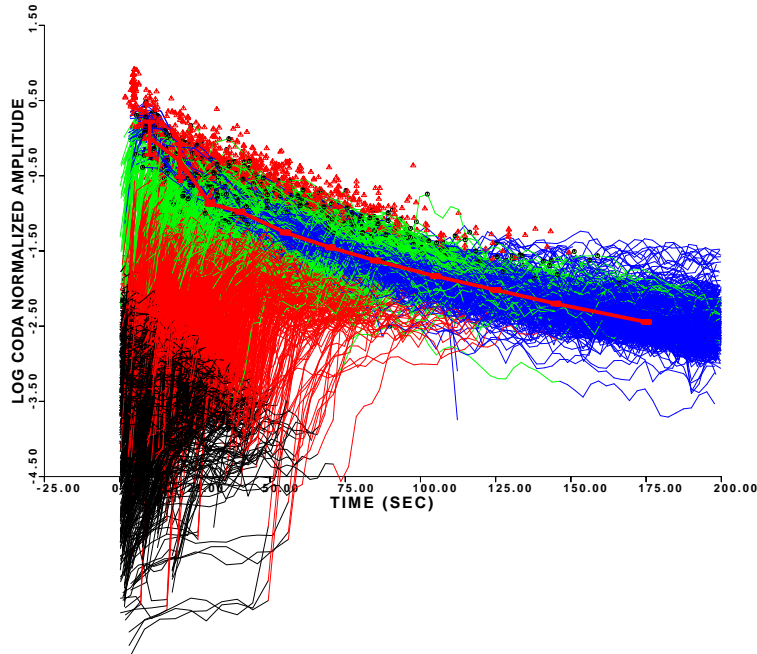
Figure 5.2: Coda decay for 1.0 Hz data in the study area (See text for explanation).

The black lines in the figure show the initial part of the seismogram before the P arrival, red lines between P and S arrivals, green lines the coda between t_s and $2t_s$, and the blue lines are the stable coda after $2t_s$. The coda shape which is empirically determined from data is shown by the heavy red line. The red triangles and the black circles represent the peak amplitude value and the normalized peak amplitudes respectively. These values are used to estimate the propagation $D(r)$ term from the coda.

5.1.2 Regression of Fourier velocity spectra

Figure 5.3 shows the regression results for the propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the Fourier velocity spectra. Top panels show the propagation estimated by coda normalization technique. The mid-panels show the $D(r)$ from the coda and the regression analysis. The $D(r)$ from the two methods are similar at distances less than 100 km. The lower panel show the regression residuals as a function of distance
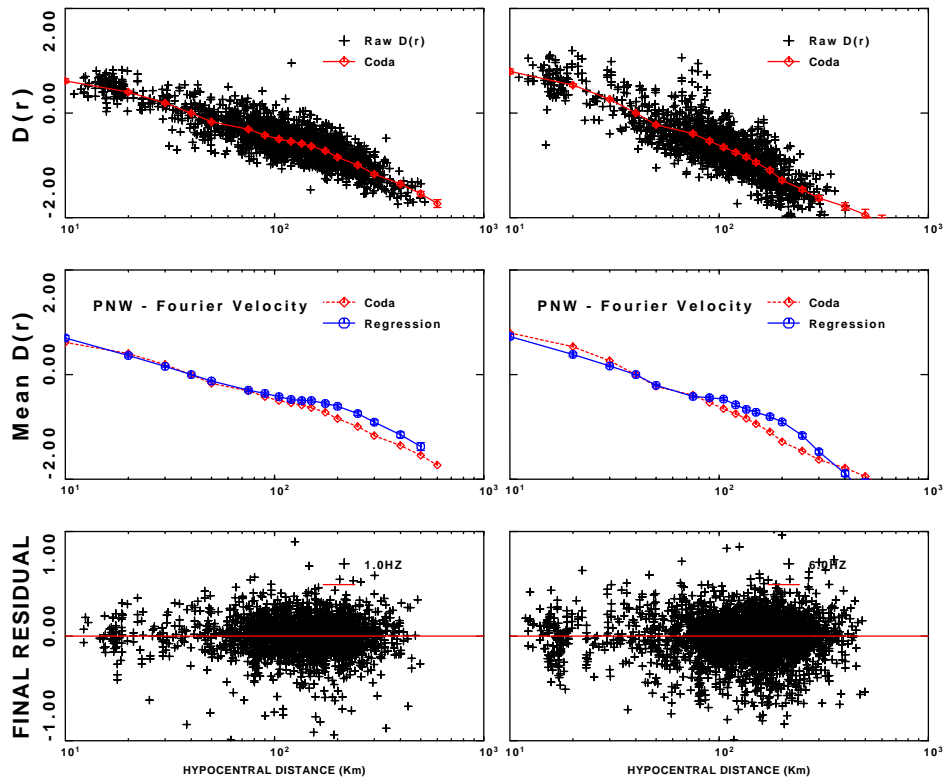
Figure 5.3: Regression results for 3 component Fourier velocity data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions ($D(r)$). Bottom, final residual of the regression analysis.

Figure 5.4 show the 3-component $D(r)$ at different frequencies obtained from the regression of the Fourier velocity data. The curves are corrected for $r^{-1.0}$ to emphasize the departure from $r^{-1.0}$ geometrical spreading. The reference Distance at 40 km is used to normalize the $D(r, f)$ (section 3.4.1). The horizontal dashed line represents the $r^{-1.0}$ trend.

The results for the 0.25, 0.3, and 0.5 Hz show a different trend from other frequencies. These lowest three frequencies may have been affected by the frequency noise, since we use an estimate of duration in the time domain within which the Fourier spectra is computed. The noise in the signal impairs the duration estimation.
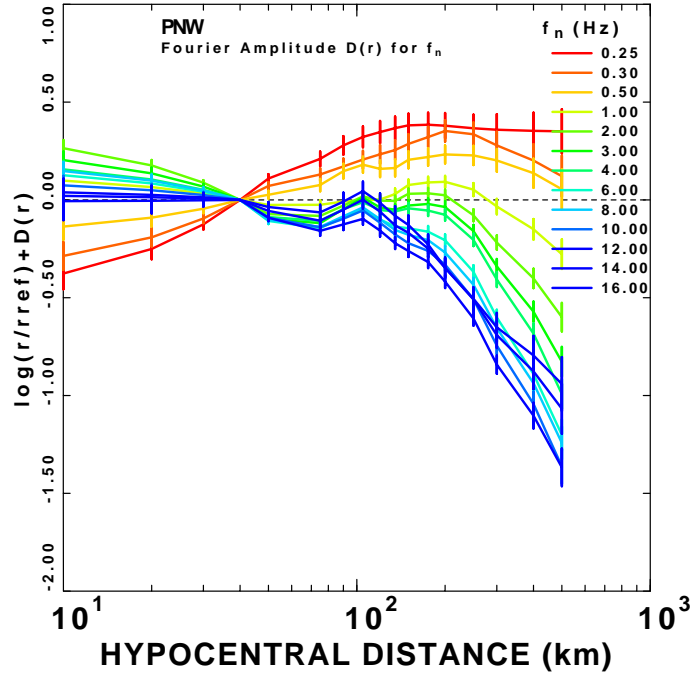
Figure 5.4: The reduced attenuation function $D(r, f)$ obtained from the regression on the 3 component Fourier velocity spectra at frequencies 0.25-16 Hz.

5.1.3   Duration

Boore (1983) used a simple duration estimate to use in RVT predictions of peak motions. *Herrmann* (1985) indicated that real earthquake signals are controlled by source and propagation duration. The RVT predictions will depend on the combined effect of source and distance duration, $T_s + T(r)$, where T_s is the source duration and $T(r)$ is the distance dependent duration which is the wave propagation contribution to the total duration. This term becomes more important in larger distances.

The distance dependent duration is modeled as a piecewise linear function of

distance:

$$T = \sum_{l=1}^{L} T_l N_l(r),$$ (5.1)

where N is the distance nodes. For large earthquakes the source duration may become comparable or larger than the propagation duration. For smaller earthquakes the propagation duration becomes much larger than the source duration at larger distances. The model fits and measured durations are shown in Fig. 5.5 for 0.3, 2.0, 3.0, and 6.0 Hz. The distance dependent duration is estimated using an L_1 norm minimization technique.
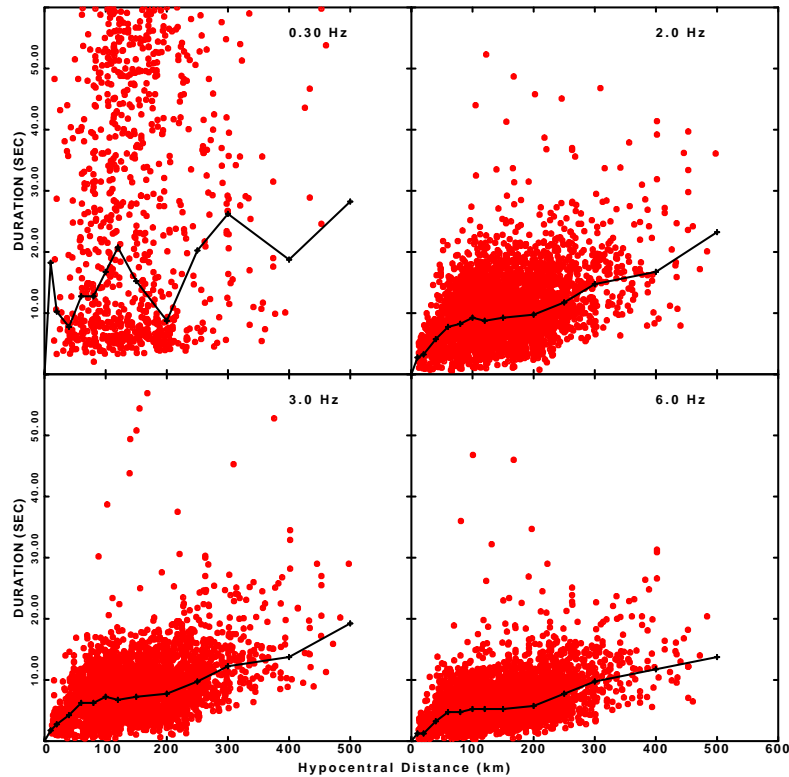
Figure 5.5: Durations at 0.3, 2.0, 3.0, and 6.0 Hz for Pacific Northwest. Red circles indicate individual duration estimates and solid lines represent the duration measured using a median value method.

The individual duration estimates (red circles) show much scatter at 0.3 Hz. The same scatter in results are seen for 0.25 and 0.5 Hz (not shown in here). For use in RVT modeling, we need a simple frequency independent duration as a function of distance. This is done by averaging the duration at each distance over different frequencies and

discarding the outliers. The results for vertical and combined 3 component (All) data is shown in Table 5.1.

Measured durations		
Distance (km)	Z (sec)	All (sec)
0.00	0.0	0.0
10.00	2.20	1.90
40.00	5.45	5.20
80.00	6.45	7.25
100.00	7.90	8.05
150.00	8.10	9.50
200.00	9.15	9.60
250.00	10.30	11.25
300.00	13.50	15.80
400.00	15.50	16.10
500.00	17.40	18.90

Table 5.1: Ground motion durations in Pacific Northwest measured from bandpass filtered data. Z is the vertical and All is the combined 3-component data.

5.1.4   Regression of bandpass filtered velocity

Compared to the Fourier velocity data, the bandpass filtered data are the better data set and the result of the regression is more robust for them. Table 5.2 shows the number of observations for each (3-component) data set.   The number of obser-

Number of observations		
Frequency (Hz)	Peak velocity	Fourier Velocity
0.25	1629	514
0.50	2899	1415
1.00	3661	2412
2.00	3772	3130
4.00	3791	3297
6.00	3795	3206
8.00	3792	3136
10.00	3771	2936
14.00	2915	1779
16.00	2823	1525

Table 5.2: Number of observations in regressions of bandpass filtered and Fourier velocity data in Pacific Northwest.

vations (data points used in regression) for the Fourier data is less than the bandpass filtered data because of the low signal-to-noise ratio. The program will reject the waveforms with a low signal-to-noise ratio. Figure 5.6 shows the regression results

for $D(r)$ at frequencies 1.0 and 6.0 Hz on the bandpass filtered data. Compared to
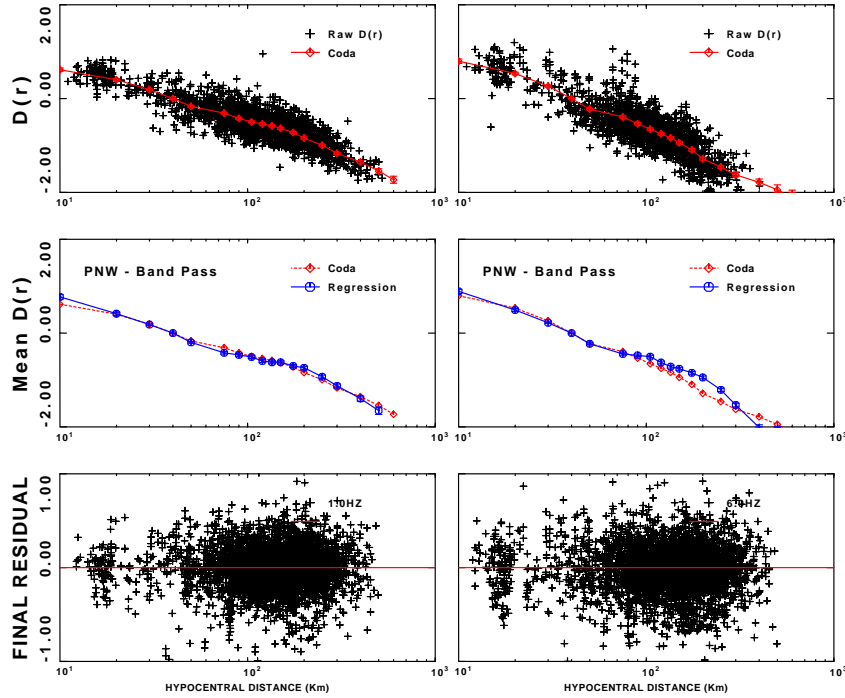
Figure 5.6: Regression results for 3 component bandpass filtered data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions. Bottom, final residual of the regression analysis.

the Fourier data (Fig. 5.3), the regression results on bandpass filtered data show a better agreement with the coda normalization estimate of $D(r)$. The same regression results are shown in Table 5.3 with the errors and number of observations. Table 5.3 shows the regression results for 1.0 and 6.0 Hz. The first and sixth columns are the frequency values. The second and seventh columns are the hypocentral distance. The $D(r, f)$ term is shown in the third and eighth columns and the associated errors estimates are given in the fourth and ninth columns. The fifth and tenth columns are the number of observations. The number of observations is larger in the distance range of 75 to 300 km. The error estimates are higher for the distances with the lower number of observations.

Figure 5.7 shows the 3'-component $D(r)$ at different frequencies obtained from the regression of bandpass filtered data. Like Fig. 5.4 the curves are corrected for r^{-1} geometrical spreading and the reference distance at 40 km is used to normalize

Pacific Northwest bandpass filtered D(r)									
f (Hz)	r (km)	$D(r,f)$	σ	Nobs	f (Hz)	r (km)	$D(r,f)$	σ	Nobs
1.0	10	0.773	0.503E-01	26.10	6.0	10	0.890	0.473E-01	42.67
1.0	20	0.415	0.335E-01	70.54	6.0	20	0.493	0.315E-01	100.67
1.0	40	0.000	0.303E-05	48.95	6.0	40	0.000	0.309E-05	62.60
1.0	75	-0.419	0.262E-01	261.93	6.0	75	-0.444	0.258E-01	273.38
1.0	105	-0.510	0.298E-01	367.75	6.0	105	-0.501	0.290E-01	376.79
1.0	150	-0.622	0.303E-01	351.26	6.0	150	-0.759	0.295E-01	352.66
1.0	200	-0.739	0.318E-01	453.70	6.0	200	-0.944	0.310E-01	462.39
1.0	300	-1.125	0.405E-01	159.05	6.0	300	-1.532	0.402E-01	158.12
1.0	400	-1.396	0.541E-01	57.22	6.0	400	-2.010	0.543E-01	55.696
1.0	500	-1.652	0.799E-01	8.65	6.0	500	-2.497	0.809E-01	8.65

Table 5.3: Attenuation function at 1.0 and 6.0 Hz for 3 component bandpass filtered data in the Pacific Northwest.

$D(r,f)$. The lowest frequency results in this figure show a similar trend with respect to the higher frequencies in the distances more than 100 km, but they are more affected by low frequency noise.
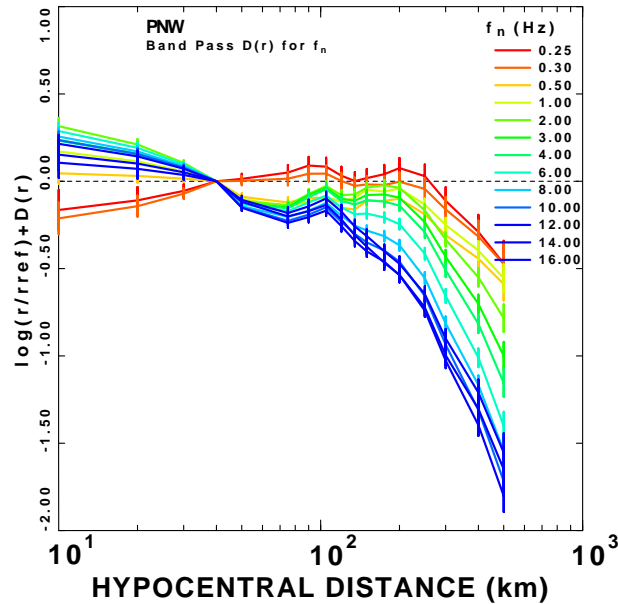
Figure 5.7: The reduced attenuation function $D(r,f)$ obtained from the regression of the 3 component bandpass filtered data at frequencies 0.25-16 Hz.

Modeling the result of the regression depends on whether we use the results of lower frequencies (0.25 -0.5 Hz) or not. In the modeling I have given a lower weight to them but in most cases I did not use them. When contaminated by noise, the

lower frequency observations can highly affect the modeling results and give wrong estimates of the parameters.

5.2 Parameterization and modeling

The parameterization is required to interpret and model the regression results. The interpretation of regression terms are given by Equations 3.41 to 3.43. First, the velocity Fourier spectra $D(r, f)$ is modeled using equation 3.41 by specifying $g(r)$ and $Q(f)$. The specified $Q(f)$ and $g(r)$ are then used with the $T(r, f)$ to make RVT predictions of the bandpass filtered $D(r, f)$. The parameters for Fourier velocity and bandpass filtered $D(r, f)$ should be consistent with each other.

The next step is the parameterization of $E(r, f)$ term for small earthquakes using equation 3.42. This step relies on the previously determined $Q(f)$ and $g(r)$ and network average site effect. In this step an average κ is defined. Finally the observed $E(r, f)$ of larger events with known moment magnitudes will be modeled for absolute scaling of ground motion.

5.2.1 Propagation parameters

The $D(r, f)$ term is a function of geometrical spreading $g(r)$ and attenuation $Q(f)$. The simple forward model for $D(r, f)$ is presented as the multiplication of g and Q. the geometrical spreading $g(r)$ is defined as r^n where r is the hypocentral distance. Attenuation $Q(f)$ is defined as

$$Q(f) = Q_0 (f/f_0)^\eta \, , \tag{5.2}$$

where $f_0 = 1.0$ Hz and η is the frequency dependent exponent that controls the attenuation separation of different frequencies. If η is 0, Q will be constant and the spatial attenuation ($\exp^{-\pi f r / Q(f) \beta}$ term in 3.40) is strongly frequency dependent in which case high frequencies are attenuated more than low frequencies. If η is 1, $Q(f)$ will be strongly frequency dependent and spatial attenuation will be the same at

different frequencies.

5.2.2 Parameterization of Fourier velocity spectra

I used both visual fitting and genetic algorithm (GA) to model the regression results. In most cases the estimates of parameters are very close for the two methods. For some data sets however, the visual fitting has not been so easy and I relied on the results of GA. For each distance and frequency I define the misfit between the observed and modeled $D(r)$ as

$$e = \frac{1}{N} \sqrt{\sum_{j}^{Nd} \sum_{i}^{Nf} w_i (Dt_{ij} - Do_{ij})^2} \tag{5.3}$$

where Dt is the theoretical $D(r)$, Do is the observed $D(r)$, w_i is the weight for different frequencies, and Nd, Nf, and N are the number of distances, frequencies and total observations respectively. The weighting factor is necessary because the low frequency data are highly affected by noise and need to be given a lower weight with respect to the higher frequencies. In some cases I have also given a lower weight to the highest frequencies than to intermediate frequencies.

GA modeling starts with producing an initial population of parameters. The bounds for parameters should be given in the GA input files. For each model (a specific combination of parameters) $D(r, f)$ is first computed using the stochastic method (RVT for peak motions) and then the misfit is calculated using (5.3). The models with lower misfits have more chance to be selected or, in other words, they appear more than other models. This means that for the next generation one model may appear twice but another model may not appear at all. The reproduction of the models is done with the cross over operation. To avoid local minima in the problem the mutation is done which doesn't allow an early convergence. The process is repeated for the next generations of the models. The different parameters used in GA modeling of Fourier velocity $D(r, f)$ are shown in Table 5.4. For a few number of parameters the method converges very fast if the number of models is high in the GA

input. This means that for a large number of models we don't need a large number of generations. Figure 5.8 shows the residuals of the GA fit for the Fourier $D(r, f)$ for

GA modeling of Fourier D(r)		
parameter	lower bound	upper bound
n	-1.5	0.6
Q	100	800
η	0.1	1.0
number of generations = 50		
number of models = 50		

Table 5.4: GA modeling parameters for 3 component Fourier velocity $D(r)$ in the Pacific Northwest.

different generations. In this case for eight parameters (Q, η, and 5 different n) the residuals decrease very rapidly and 20 generation will be enough to get the result.
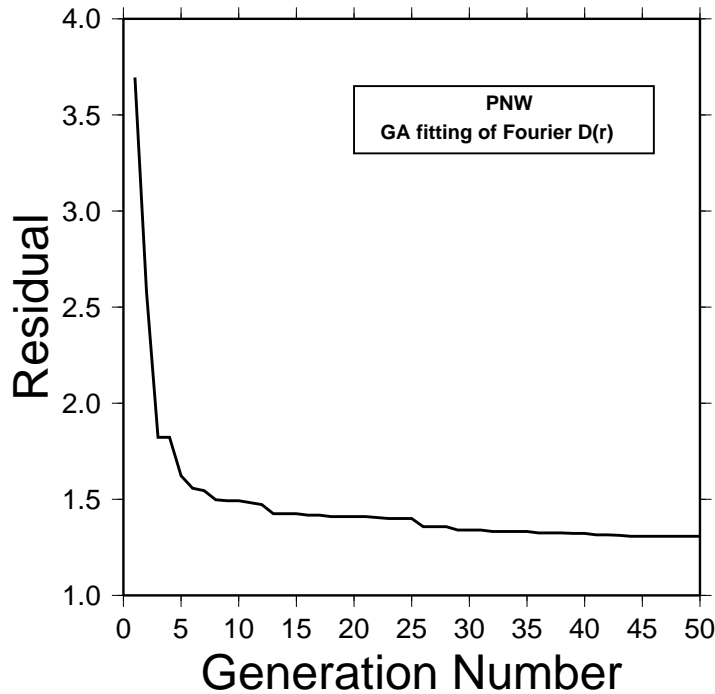
Figure 5.8: GA modeling residuals for 3-component Fourier velocity $D(r)$ in Pacific Northwest.

Figure 5.9 shows the results of modeled $D(r, f)$ using the GA method. The modeled (black) and observed (colored) curves are plotted together. The $D(r, f)$ values are corrected for an $r^{-1.0}$ trend to show the departure from simple $r^{-1.0}$ spreading. The horizontal dashed line in the upper plot represent $r^{-1.0}$ trend. The lower plot
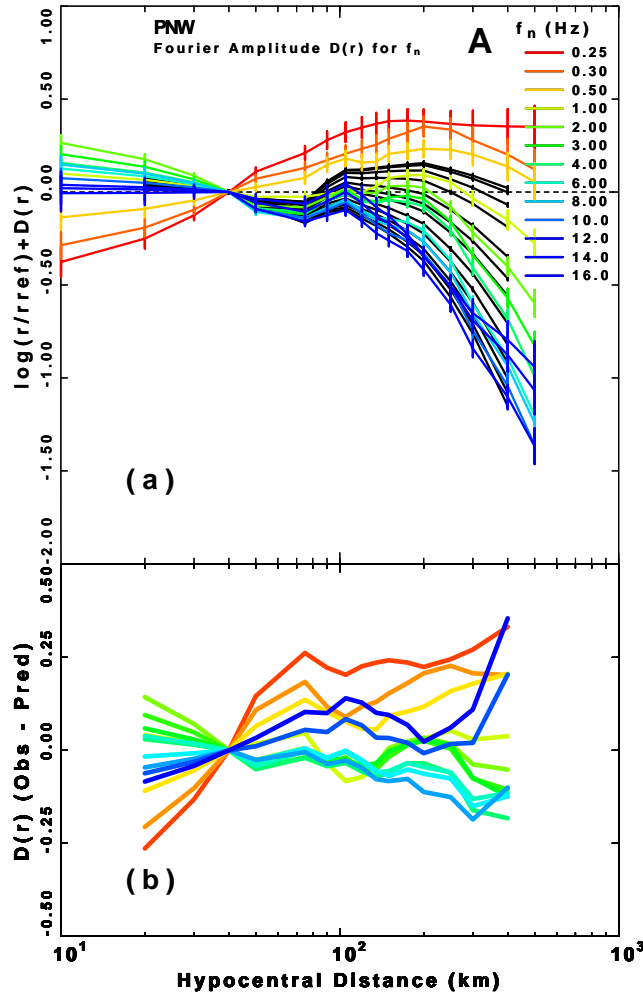
Figure 5.9: (a) observed (colored) and modeled (black) 3-component Fourier velocity attenuation function for 0.25-16 Hz in the Pacific Northwest (b) The residuals of the model fit to the observed $D(r, f)$.

shows the residuals of the model fit using GA. The residuals for the 0.25, 0.3, and 0.Hz are high. The model fit is good for all other frequencies . The final propagation parameters are $Q_0 = 280$, $\eta = 0.55$, and

$$
g(r) = \begin{cases}
r^{-1.0} & r < 40 km \\
r^{-1.2} & 40 < r < 70\ km \\
r^{0.3} & 70 < r < 100\ km \\
r^{-0.7} & 100 < r < 200\ km \\
r^{-0.9} & r > 200\ km
\end{cases}
$$

As mentioned before in the previous sections the parameterization using visual fitting or GA modeling does not provide error measurements. I have used a bootstrap method throughout this study to estimate the error bounds of the parameters. The bootstrap method can be applied in several ways. A more reliable approach is to randomly resample the original data set, apply the regression and then model the regression results. For my large data sets this approach takes a long time since every regression could take from half to one hour. Then for each regression result the GA modeling is run (with 50 models and generations for example) to get a set of parameters. In my case each GA run took 5-8 hours. Computing a bootstrap with 200 runs of GA the required running time will be very long unless a cluster is used. Instead, I took another approach and started the procedure from the results of the regressions. I randomly resampled the regression results within the error bounds of the results and used GA for the modeling of the observations. This latter approach takes a much shorter time. The data can be resampled within 1σ or 2σ but resampling within 2σ may overestimate the errors of the estimated parameters. For the whole study I did the resampling within 1σ of the regression results for both $D(r, f)$ and $E(r, f)$ terms. Table 5.5 shows the bootstrap mean and error estimates of the propagation parameters. the parameters of the best model are also shown in the second column for comparison.

Bootstrap and best model results-Fourier $D(r)$			
parameter	best model	bootstrap mean	bootstrap σ
Q_0	280	311	38
η	0.55	0.57	0.06
n_1	-1.0	-1.01	0.09
n_2	-1.2	-1.09	0.08
n_3	0.3	0.185	0.07
n_4	-0.7	-0.95	0.09
n_5	-0.9	-0.75	0.06
number of GA runs = 100			
No. of models in each run = 50			

Table 5.5: Bootstrap error estimate of modeled propagation parameters for Fourier $D(r, f)$ in the Pacific Northwest.

Note that the mean of each parameter is not the same as that of the best model. Different ns are the power of r (in geometrical spreading function, $g(r)$) at five differ-

ent distances. Based on the results I peak the best model value for each parameter and the σ from the bootstrap as the related error. I used 1σ as the error so $Q_0 = 280 \pm 38$ and $\eta = 0.55 \pm 0.06$.

5.2.3   Parameterization of bandpass filtered velocity

The same procedure has been used for parameterization of bandpass filtered $D(r, f)$. In this part the previously determined Q, η, and n (from the Fourier domain) are used with the ground motion durations for fitting the time domain $D(r, f)$ using RVT. During the modeling the parameter values and durations may need to be changed for the best fit. The final values of parameters may be different for the Fourier and time domain since the $D(r, f)$ values are different for the two domains. For this study the modeling of the time domain (bandpass filtered) $D(r, f)$ are the most important because they are obtained from the regression of the better and the more reliable data set. But in any case the parameterization of the Fourier domain is done first. For the best model $Q_0 = 280$, $\eta = 0.55$ and

$$
g(r) = \begin{cases}
r^{-1.0} & r < 40\,km \\
r^{-1.3} & 40 < r < 70\ km \\
r^{0.5} & 70 < r < 100\ km \\
r^{-0.6} & 100 < r < 200\ km \\
r^{-0.9} & r > 200\ km
\end{cases}
$$

The geometrical spreading is different in some distance ranges from the Fourier domain results. For distance range of 70 to 100 km there are strong reflections from Moho and a positive exponent is needed for the r. In both Fourier and time domains the visual and GA modeling results are very close but I have only shown the GA results.

Figure 5.10 shows the modeling of the 3-component bandpass filtered $D(r, f)$. I did not model the 0.25, 0.3, and 0.5 Hz frequencies because the residuals at these frequencies were high (like Fourier domain) and even with giving lower weights they

47
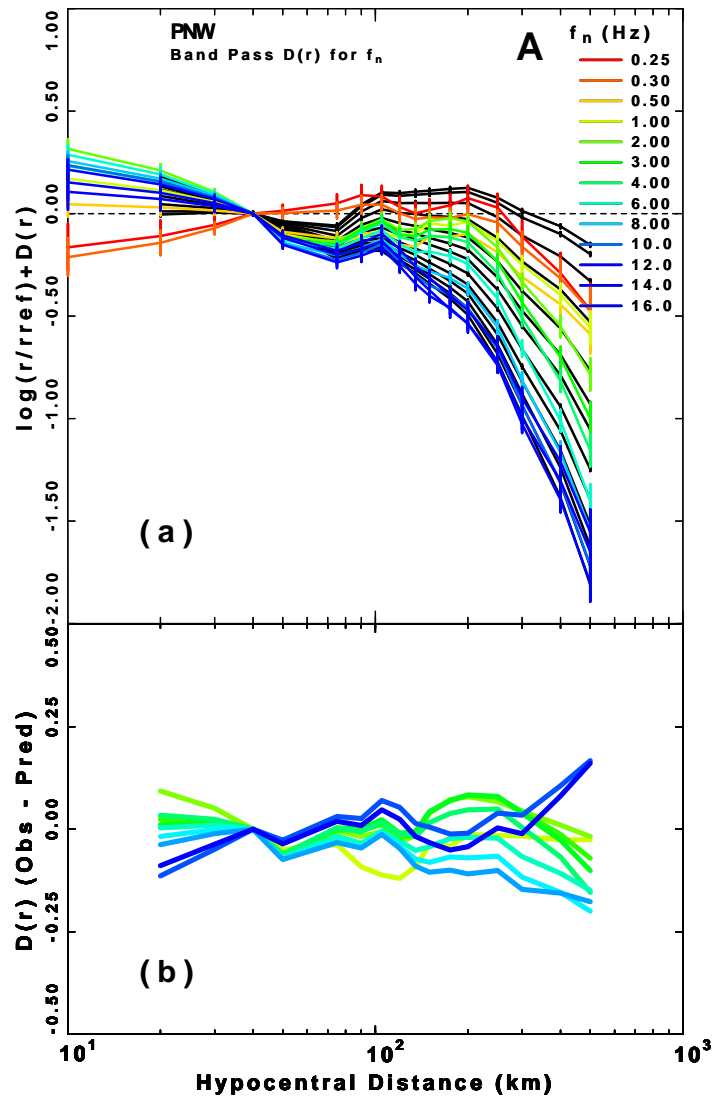
Figure 5.10: (a) observed (colored) and modeled (black) 3-component bandpass filtered velocity attenuation function for 0.25-16 Hz in the Pacific Northwest (b) The residuals of the model fit to the observed $D(r, f)$ for 1.0-16 Hz.

affected the results for other frequencies. The plots of modeled $D(r, f)$ for those lowest frequencies in Figure 5.10 are based on the results of higher frequencies. The residuals are low for distances up to 300 km for frequencies 1.0-16.0 Hz.

Table 5.6 shows the bootstrap mean and standard deviations of the modeled propagation parameters of bandpass $D(r, f)$. The last column in the Table 5.6 shows the best model values and related errors. I have chosen the best model values not the bootstrap means because the best model values are obtained from the observed not

Bootstrap and best model results-All bandpass $D(r)$				
parameter	best model	bootstrap mean	bootstrap σ	selected value
Q_0	280	296	34	280 ± 34
η	0.55	0.54	0.05	0.55 ± 0.05
n_1	-1.0	-1.01	0.09	-1.0 ± 0.09
n_2	-1.3	-1.19	0.07	-1.3 ± 0.07
n_3	0.5	0.55	0.09	0.4 ± 0.09
n_4	-0.6	-0.85	0.08	-0.6 ± 0.08
n_5	-0.9	-1.01	0.07	-0.9 ± 0.07
number of GA runs for bootstrap = 100				
No. of models in each run = 50				

Table 5.6: Bootstrap error estimate of modeled propagation parameters for 3 component bandpass $D(r, f)$ in the Pacific Northwest.

the perturbed data.

There is a strong trade off between Q_0 and η and between $Q(f)$ and $g(r)$ so none of the determined parameters are unique. Figure 5.11 shows the trade off between Q_0 and η for a fixed $r^{-1.0}$ geometrical spreading. The misfits (Eq. 5.3) in the Figure 5.11
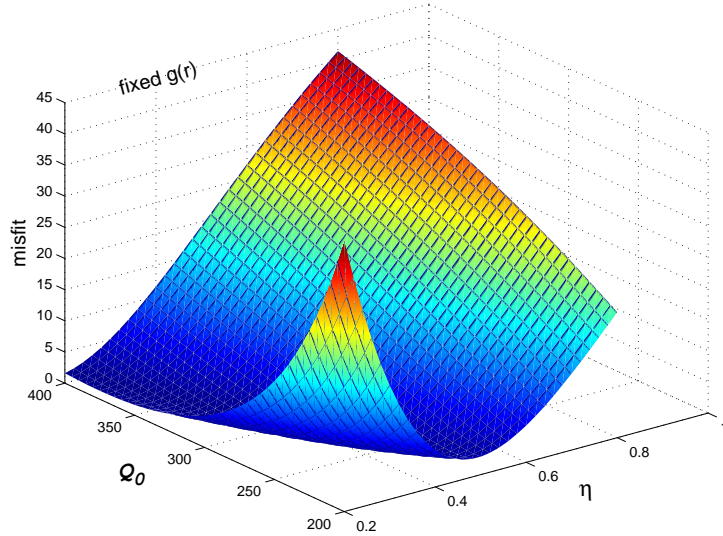
Figure 5.11: Trade off between Q_0 and η. A fixed $r^{-1.0}$ geometrical spreading is used in the parameterization.

are obtained by 1270 runs of GA (1270 different combination of Q_0 and η) for modeling the observed $D(r, f)$ using RVT. The $r^{-1.0}$ geometrical spreading is used to show the trade off between Q_0 and η. The use of a fixed geometrical spreading will decrease the degrees of freedom and hence the trade off to some extent. Numerical analysis

of crustal wave propagation (*Wang and Herrmann*, 1980) and empirical studies of ground motion attenuation (*Campbell*, 1991) however, show that a varying exponent of geometrical spreading in different distances is well justified.

Since more than 90% of the PNW data are vertical short periods, I have only used 3-component (all waveforms) and vertical component analysis. The horizontal data did not form a good data set for a regression analysis. Figure 5.12 shows the regression results for vertical component data at 1.0 and 6.0 Hz. In the middle
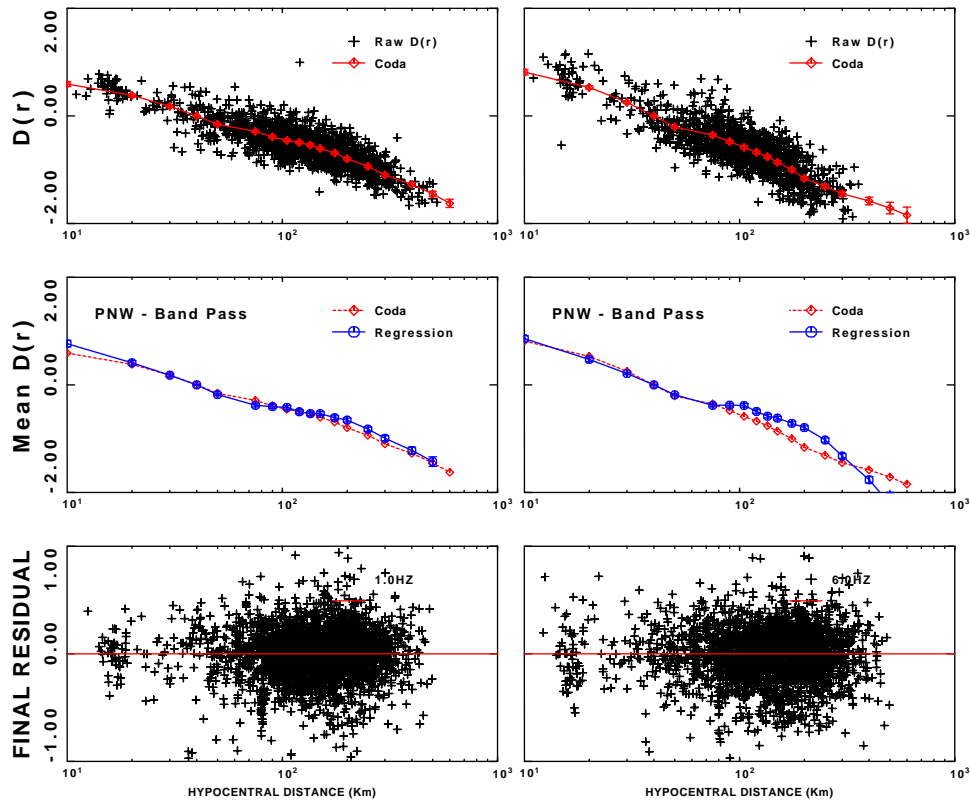
Figure 5.12: Regression results for vertical bandpass filtered data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions ($D(r)$). Bottom, final residuals of the regression analysis.

panel of the figure the obtained $D(r, f)$ by coda normalization method and regression are compared. Usually the agreement are good between two method for frequencies 1-10 Hz. The coda are more affected by noise and I only model the results of the regression. But the coda estimate of $D(r, f)$ provide a good estimate for comparison purpose. Figure 5.13 shows the observed and modeled vertical component bandpass

filtered $D(r, f)$. The lower panel shows the residual of fit. As in Figure 5.10 I modeled the frequencies 1-16 Hz and residuals shown are for that frequency range. For the-Z
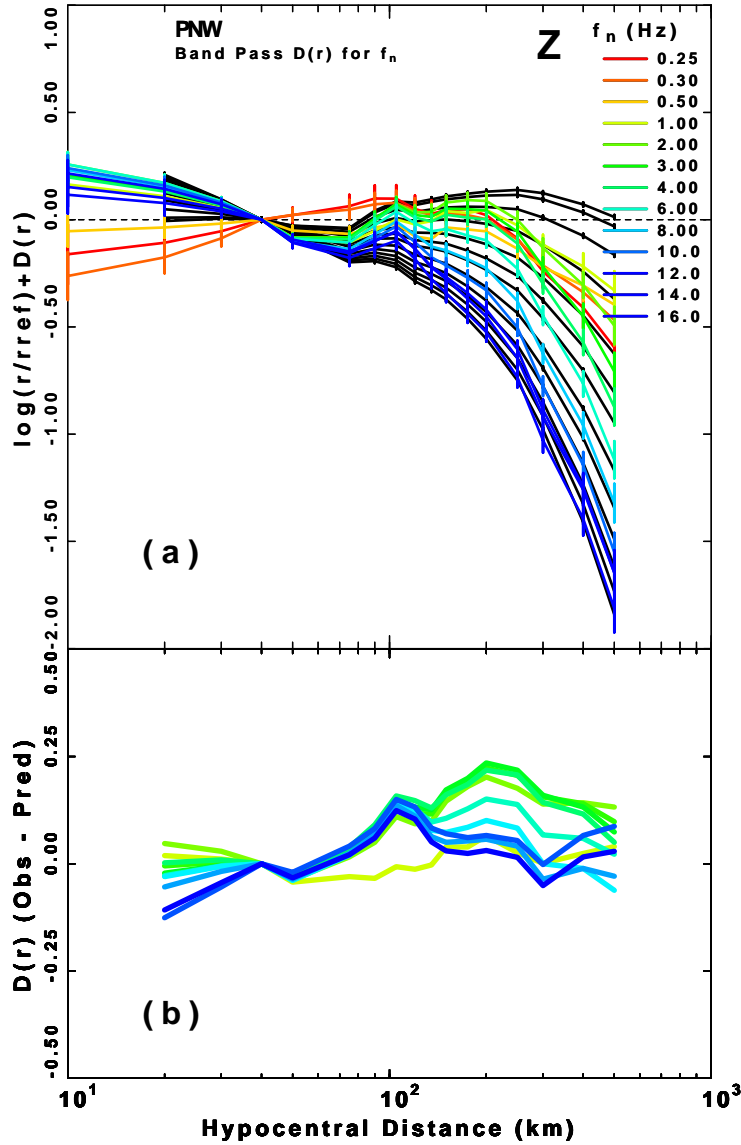
Figure 5.13: (a) Observed (colored curves) and modeled (black curves) vertical component bandpass filtered velocity attenuation function for 0.25-16.0 Hz in the Pacific Northwest. (b) Residuals of the model fit to the observed $D(r, f)$ for 1-16 Hz.

component modeling the residuals are high for distances larger than 200 km at 2-6 Hz. The GA modeling results can change by changing weights for different frequencies, but for the Z component, not all the frequencies could fit well regardless of weighting scheme. The parameterization result for the bandpass filtered Z-component is a little

different from the 3-component results. The differences can be attributed to the different measured durations (Table 5.1) and different number of observations for each data set which results in different regression results. The best model for the Z-component yields $Q_0 = 290$, $\eta = 0.50$, and

$$
g(r) = \begin{cases}
r^{-1.0} & r < 40 km \\
r^{-1.2} & 40 < r < 70 \ km \\
r^{0.2} & 70 < r < 100 \ km \\
r^{-0.7} & 100 < r < 200 \ km \\
r^{-0.8} & r > 200 \ km
\end{cases}
$$

The error analysis results and the final selected values for each parameter are shown in Table 5.8. The errors for the Z-component are somewhat larger than all (3-

Bootstrap and best model results-Z bandpass $D(r)$				
parameter	best model	bootstrap mean	bootstrap σ	selected value
Q_0	290	310	41	290 ± 41
η	0.50	0.48	0.06	0.50 ± 0.06
n_1	-1.0	-1.18	0.09	-1.0 ± 0.09
n_2	-1.2	-1.01	0.11	-1.2 ± 0.11
n_3	0.2	0.32	0.095	0.2 ± 0.095
n_4	-0.7	-0.82	0.08	-0.7 ± 0.08
n_5	-0.8	-0.98	0.09	-0.8 ± 0.09
number of GA runs for bootstrap = 100				
No. of models in each run = 50				

Table 5.7: Bootstrap error estimate of modeled propagation parameters for vertical component bandpass $D(r, f)$ in the Pacific Northwest.

component) data. The 3-component data includes some horizontal seismograms and therefore is a more complete data set. The distance and frequency coverage for 3-component data set is better than that of the Z-component. So the regression result for the 3-component data might be a more robust solution with smaller error. In all study area I will use the results of the 3-component data as the final solution. For example I will use a value of 280 ± 34 for Q_0 in the Pacific Northwest based on the modeling results of 3 component bandpass filtered $D(r, f)$.

Once the propagation parameters are determined the excitation at the reference

distance of 40 km will be modeled using the values of Q_0, η, and n.

5.2.4 Parameterization of bandpass filtered Excitation

The excitation at 40 km obtained from the regression will be modeled using Equation 3.41. The most important parameter in the excitation modeling is the stress drop. The value of κ is also needed for this parameterization. I used an ω^{-2} Brune source model and previously determined attenuation function (section 5.2.3) with RVT to produce theoretical excitations. Usually κ is first defined by fitting some small events and stress drop is obtained by fitting the observed excitations of a number of events with known moment magnitude. Figure 5.14 shows the observed and a set of theoretical excitations for the 3-component bandpass filtered data. The black curves show the observed excitations from the regression and the red curves are the theoretical ones for moment magnitudes 3.0-6.0.
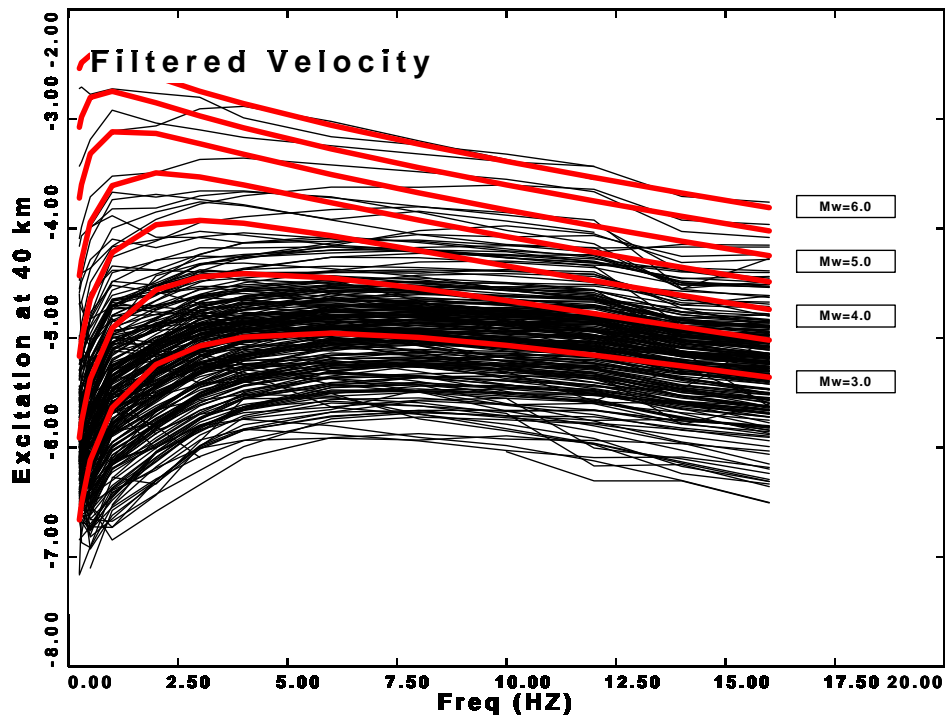
Figure 5.14: Excitation at 40 km for 3-component bandpass filtered data in the Pacific Northwest. Black curves represent the excitation of different events obtained from the regression. Red curves are the theoretical curves for events with moment magnitudes 3.0 to 6.0.

The red curves in the Figure 5.14 are computed assuming a stress drop and a κ values and they were not fitted to the observed ones. The results of actual modeling is shown in Figure 5.15. In this figure the frequency axis is also logarithmic. The small boxes in the right side are the moment magnitudes of theoretical curves. The boxes at the left side are the moment magnitude of the selected events with their observed excitation are shown by dotted blue curves. I have used 6 events with
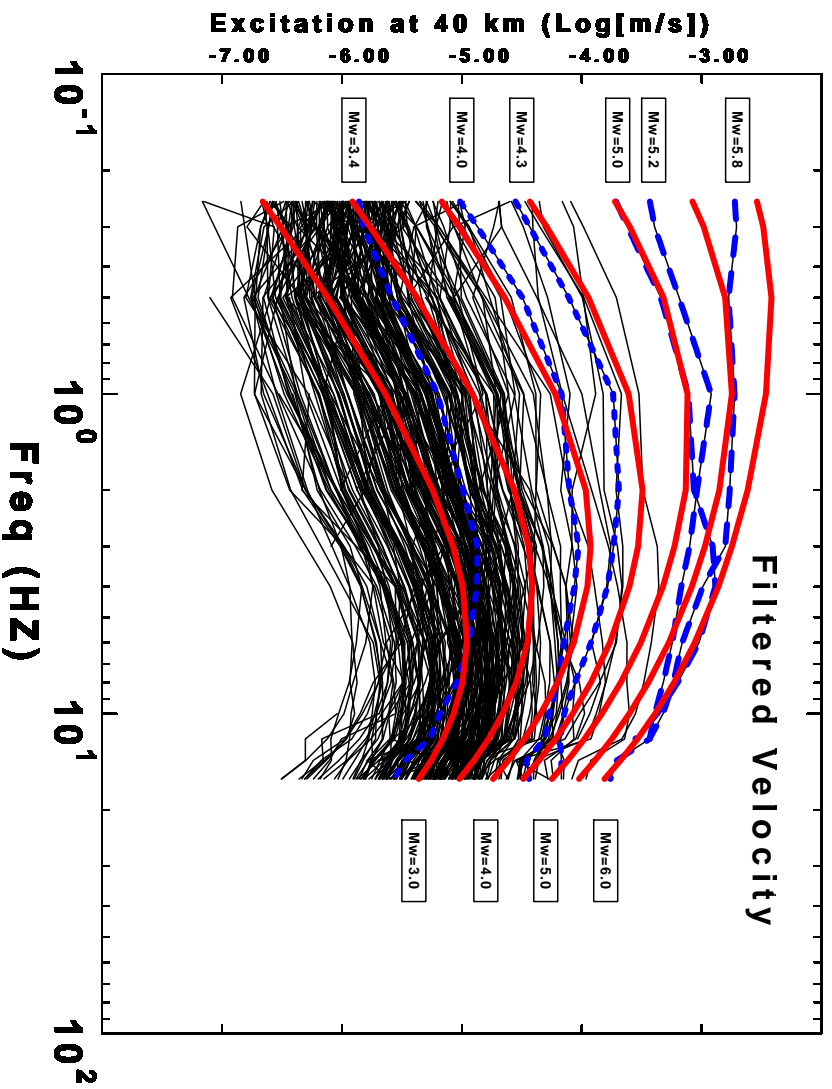
Figure 5.15: Excitation at 40 km for 3-component bandpass filtered data in the Pacific Northwest. Black curves are observed excitations. Red curves are the theoretical curves for events with moment magnitudes 3.0 to 6.0. Blue dotted curves are the observed excitations of selected events with known moment magnitudes.

known moment magnitudes of 3.4, 4.0, 4.3, 5.0, and 5.8 and modeled their excitation simultaneously to obtain the values of $\Delta\sigma$ and κ. The best models gives $\Delta\sigma = 30$ bars and $\kappa_{eff} = 0.045$ sec.

The obtained parameters are based on the assumption that stress drop is independent of the source size for a specific region. So regardless of the number of events that

are modeled at one time and what magnitude they have, the result will be a constant stress drop. Table 5.8 shows the result of the error analysis for the modeling of the Fig. 5.15. The bootstrap mean is very close to the best model value. The errors from

Bootstrap and best model results- bandpass $E(40, f)$				
parameter	best model	bootstrap mean	bootstrap σ	selected value
$\Delta\sigma$ [bars]	30	27	9	30 ± 9
κ_{eff} [sec]	0.045	0.042	0.009	0.045 ± 0.009
number of GA runs for bootstrap = 100				
No. of models in each run = 50				

Table 5.8: Bootstrap error estimate of modeled excitation at 40 km for 3-component bandpass $E(r, f)$ in the Pacific Northwest.

the bootstrap are also small. The residuals of the fit for most of the events are also in an acceptable range (Fig. 5.16). For the six selected events with known moment
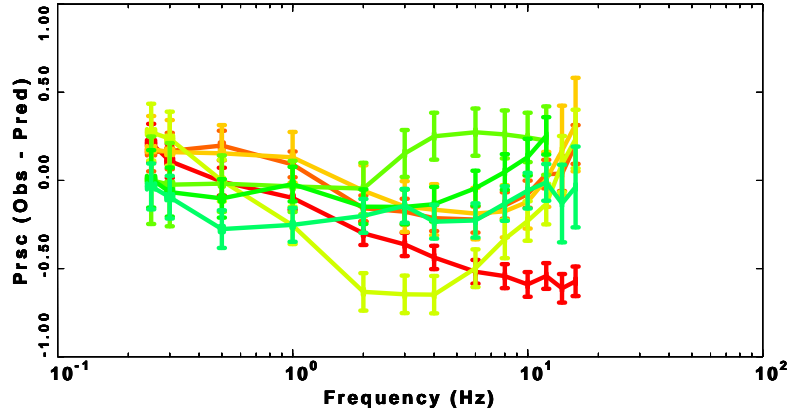
Figure 5.16: Residuals of the fit for the 3-component bandpass filtered excitation at 40 km for the 6 events of the Fig. 5.15.

magnitude (blue curves Fig. 5.15) the modeling residuals converged very fast. Figure 5.17 shows the residuals of the fit for different generations. In general Modeling the excitation is harder than the propagation, $D(r)$, term. The assumption of a constant stress drop and using Brune's point source model are partially responsible for the problems in fitting the excitation term. In this case the six selected events could fit well and they covered a range of magnitudes good enough for modeling. In some cases using the 2 corner Brune's model may improve the fitting of the model to the observed. For the PNW however, it was not necessary to do this.
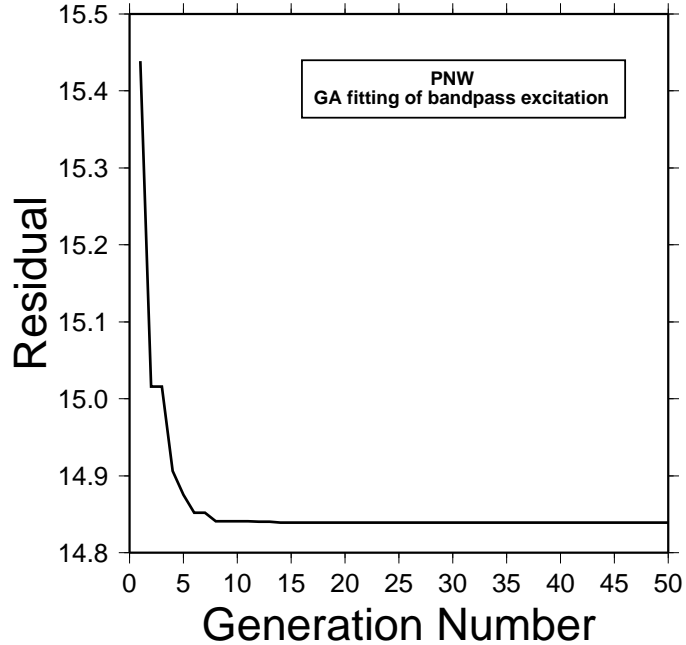
Figure 5.17: Residuals of each generation for 3-component bandpass filtered excitation at 40 km for the 6 events of the Fig. 5.15.

To further investigate the validity of a constant drop in the PNW region, I have done another modeling in which I tried to separately fit the excitation terms for some the events (with known moment magnitude) in the data set. In the modeling each event has a different stress drop and κ. Among the events I selected 25 events with computed excitation terms. The moment magnitude of those events have been computed by the staff of the College of Oceanic and Atmospheric Sciences, Oregon State University (http://quakes.oce.orst.edu/moment tensor). The moment tensor solutions are available for events from 1994 to 1998. The events which I have used have moment magnitudes from 3.4 to 5.8. The summarized results of modeling of the individual events are shown in the table 5.9. For 2 events from the total of 25

Stress drop and κ-PNW				
parameter	min value	max value	mean	standard deviation
$\Delta\sigma$ [bars]	10	91	31	24
κ [sec]	0.0071	0.09	0.048	0.03

Table 5.9: Mean and standard deviation of stress drop and κ for 23 events with known moment magnitudes in the Pacific Northwest. Brune's source model is used for all events.

the computed stress drops were very high (270 and 315 bars). The excitation terms had errors larger than other similar events and smaller number of observations. Since more than 90% of the events had stress drops less than 100 bars, I excluded those two events from the computations. A more reliable estimate would be possible if there were more events with known moment magnitudes. The results of Table 5.9 shows that the mean of the stress drop and κ for the 23 events are very close to the result of the simultaneous modeling using only 6 events. The standard deviations of both parameters however, are very high for the 23 events, almost comparable to the parameter's values.

Figure 5.18 shows the result of this modeling using the GA algorithm. The left
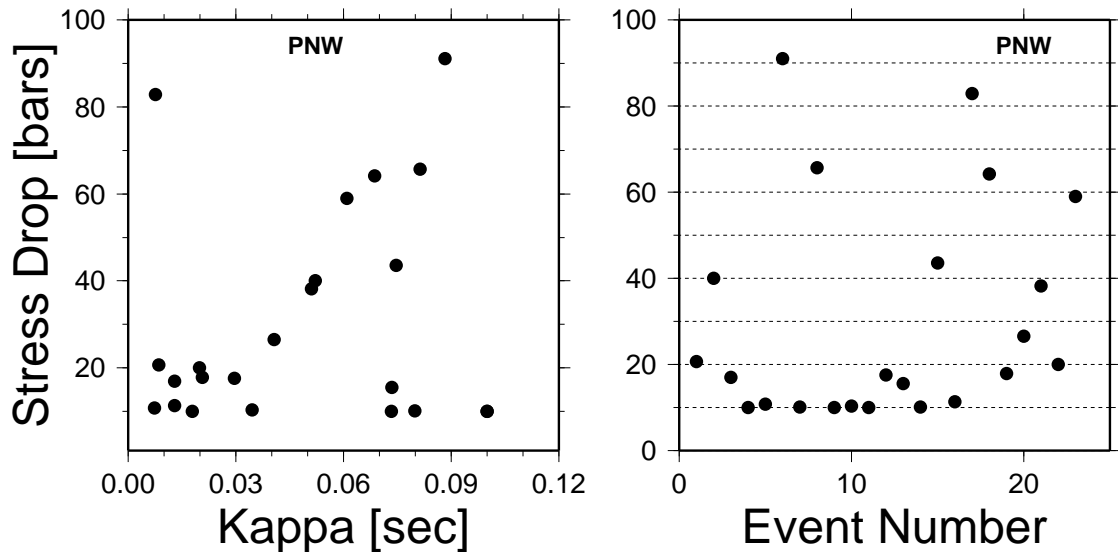
Figure 5.18: (left) Stress drop versus κ for 23 events with known moment magnitudes in the Pacific Northwest. (right) Stress drop for each event. For each event the excitation terms has been modeled separately.

figure shows stress drop versus κ. There is no correlation between the two parameters. The right figure shows the stress drop for each event which ranges from 10 to 90 bars. It seems reasonable to pick the mean κ and use it as a fixed parameter to estimate stress drop from the excitation terms in the Pacific Northwest (as done in modeling of Fig. 5.15). Using the simultaneous modeling of 6 events and individual modeling of 23 events I will select a constant stress drop of 30 bars for the Pacific Northwest region. Conclusion of a constant stress drop should be based on the results of studying

more events. Statistically, 23 events are not enough to conclude a constant or non constant stress drop. But some other studies in the Cascadia and Pacific Northwest also reported a similar average value for the stress drop (*Atkinson*, 1995; *Dewberry and Crosson*, 1995).

5.2.5 Site terms

The effect of site is very important on ground motions. The parameter κ which appears in 3.41 and 3.43 (excitation and site) is used to account for the decay of high frequency motions. I don't model the site terms in this study. The parameters κ however, determined in the previous section with the modeling of the excitation term.

Figure 5.19 shows the site terms from the regression results on vertical and 3-component bandpass filtered velocity.
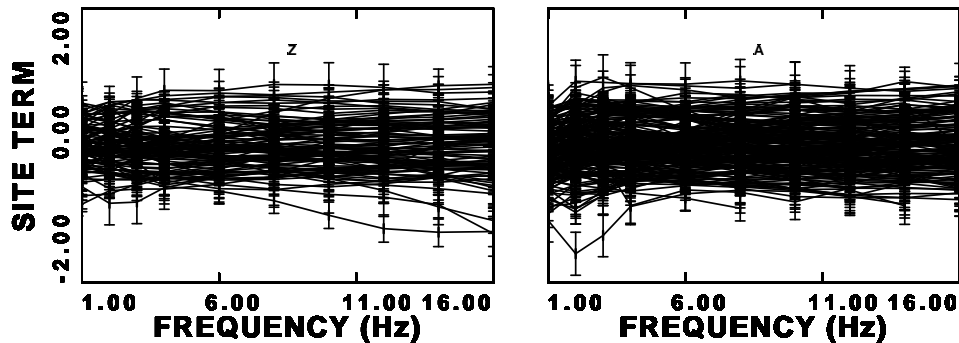
Figure 5.19: Inverted vertical (left) and 3-component (right) site terms for the PNSN network sites.

Figure 5.20 shows the individual horizontal to vertical site term ratios and their average. Correlating the site terms and H/V ratios with the geological conditions needs a detailed knowledge of the shallow geology at each site.

5.2.6 Attenuation and stress drop comparisons

Atkinson (1995) and *Dewberry and Crosson* (1995) reported an average value of 30 and 32 bars for stress drop in the Cascadia and Pacific Northwest regions. *Atkinson*

Empirical H/Z Ratio
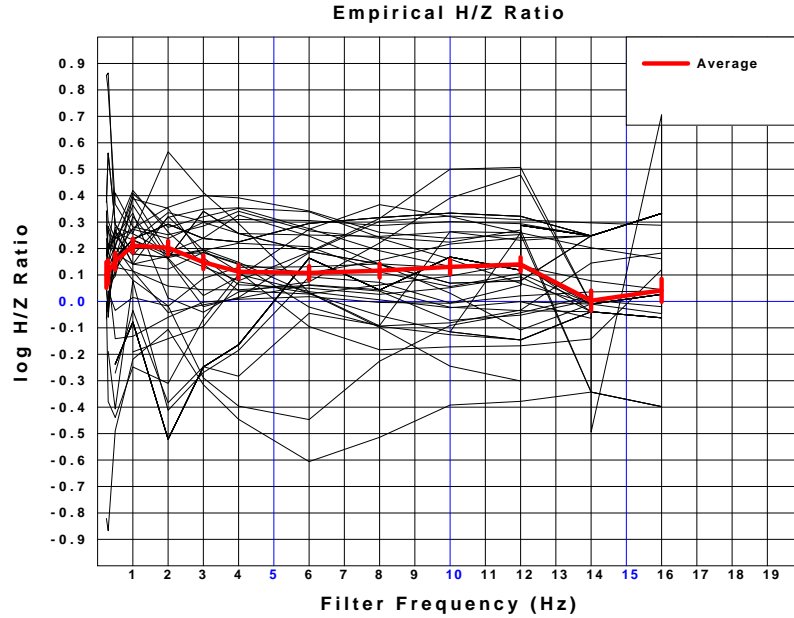
Figure 5.20: Observed and mean log H/Z ratio from the station terms of Fig. 5.19. Thin lines represent individual station ratios. The thick read line represent the men ratio.

(1995) used over 1000 vertical-component Fourier-spectra in the Cascadia region of southwestern British Columbia and northwestern Washington. A part of her study area overlaps with the northern part of this study. *Dewberry and Crosson* (1995) used more than 8000 coda amplitude measurements of 78 local earthquakes in the northern Washington (northern part of this study). The average stress drop of this study (30 bars) is the same as their results.

Atkinson (1995) reported 3 different attenuation functions using a simple r^{-1} geometrical spreading. She divided the data set into shallow crustal ($h < 10$ km) and deep crustal and plate ($h > 20$) events. Her results (using r^{-1}) are compared to the results of my study in the table 5.10. My result for intrinsic attenuation ($280f^{0.55}$)

Atkinson 1995			This study
whole data	shallow data	plate data	whole data
$380f^{0.39}$	$174f^{0.58}$	$263f^{0.49}$	$280f^{0.55}$

Table 5.10: Intrinsic attenuation function comparison between this study and Atkinson 1995.

is close to Atkinson's result for the plate (deeper crustal and in slab) data. My data

59

set does not have any earthquakes from British Columbia and Atkinson's data set does not include any event from southern Washington and Oregon. The propagation paths for the two studies therefore are different. I did not divide my data set into shallow and deep data instead I did a regression for the whole data set. My data have some deep events too but most of the events are crustal earthquakes and the goal of this study is to analyze those events. And finally I did not use an r^{-1} geometrical spreading. The regression results for the propagation term $(D(r, f))$ required a more complicated form of geometrical spreading (5.2.3).

Chapter 6

Ground motion scaling in northern California

The data set of northern California earthquakes (Fig. 6.1) are vertical short period and 3-component broad band seismograms which cover a distance range of up to 350 km. The data set I used for this study area was from the NCSN and BK networks.
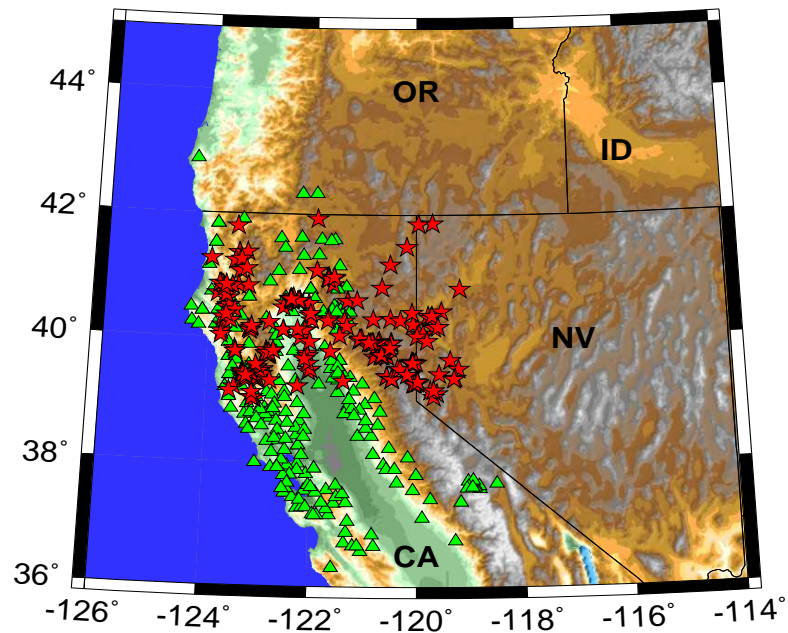
Figure 6.1: Map of stations (triangles) and events (stars) for this study in northern California.

Most of the seismograms were cut after the $S-$wave arrival. I used the EVTFAST option to get the data from the networks. Using this option the requested data will not have enough coda length in time because they are used for the location purpose. So the coda analysis results are not good for this data set. Beside this problem the data quality for most of the stations and earthquakes were not good. So I couldn't do a complete analysis (as for PNW) of the data set. In this chapter I will show the

results of those parts of data processing which I think are reliable but they are still affected by the low quality of the data. The 3-component bandpass filtered data were the best among the other and I will show the results for these part of the processed data.

6.1   Regression of bandpass filtered data

Figure 6.2 shows the regression results for propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the 3-component bandpass filtered data. Top panels show the propagation term estimated by coda normalization technique. The mid-panels show the $D(r)$ from the coda and the regression analysis.
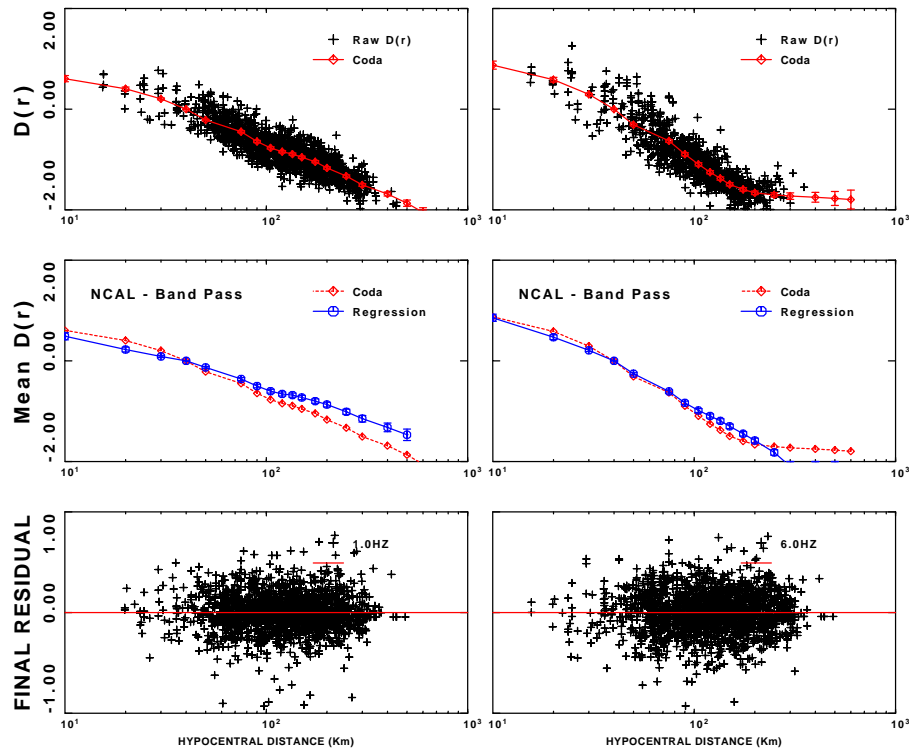
Figure 6.2: Regression results for 3-component bandpass filtered data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions ($D(r)$). Bottom, final residual of the regression analysis.

Figure 6.3 shows the same results for 2 and 10 Hz. The agreement between coda and regression estimates are better for the 6 and 10 Hz. Since the data points for the
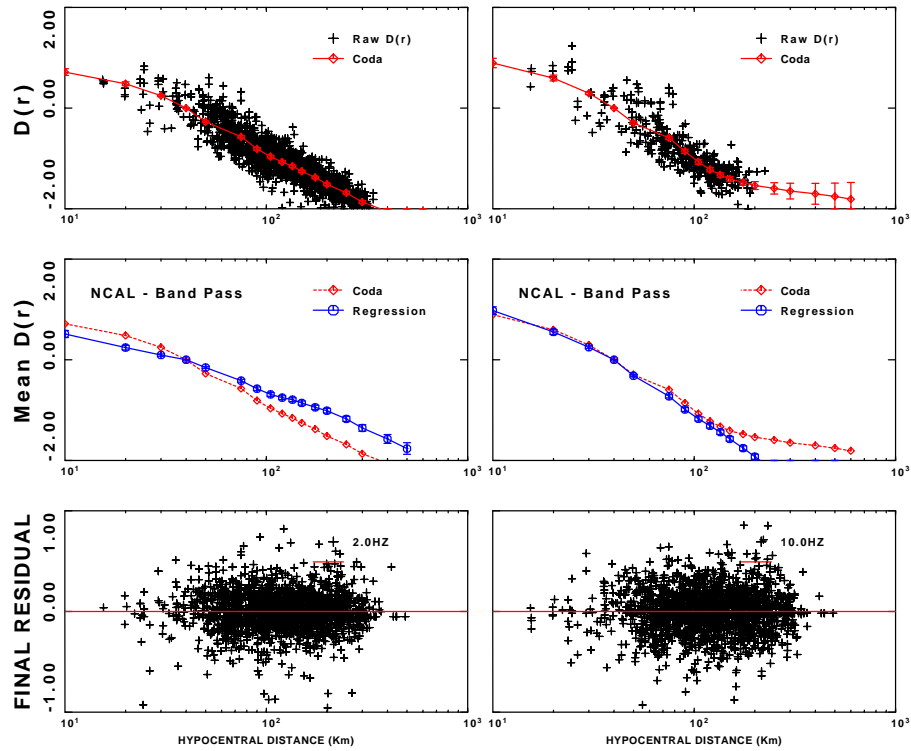
Figure 6.3: Regression results for 3-component bandpass filtered data at 2.0 and 10.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions ($D(r)$). Bottom, final residual of the regression analysis.

coda method are less than the regression method, I can not rely on the coda method results. Beside the data quality, most of the seismograms did not have enough coda length. The figures show a rapid decrease of amplitude for the frequencies 6.0 and 10.0 Hz.

The regression results for the Fourier velocity are not reliable because the total number of observations are much less than those for the bandpass filtered data. This was the case for the PNW data too, but for the northern California even the bandpass filtered data do not have a high number of observations. So the distance and frequency coverage of the data for this study area is not as good as the Pacific Northwest.

Table 6.1 compares the number of observations for the 3-component bandpass filtered data and the Fourier velocity data. Based on the results of the Pacific Northwest, northern California, and central California (next chapter), the minimum number

63

of observation required for a stable solution is between 2000-2500. For the smaller number of observations in this kind of study, the results of the regression should be used carefully. Figure 6.4 shows the 3-component $D(r)$ obtained from the regression

Number of observations		
Frequency (Hz)	Peak velocity	Fourier Velocity
0.25	627	123
0.50	1409	571
1.00	1751	1045
2.00	1798	1457
4.00	1793	1586
6.00	1781	1532
8.00	1767	1426
10.00	1731	1247
14.00	1338	742
16.00	1212	614

Table 6.1: Number of observations in regressions of bandpass filtered and Fourier velocity data in northern California.

of bandpass data at different frequencies. The curves are corrected for the r^{-1} geo-
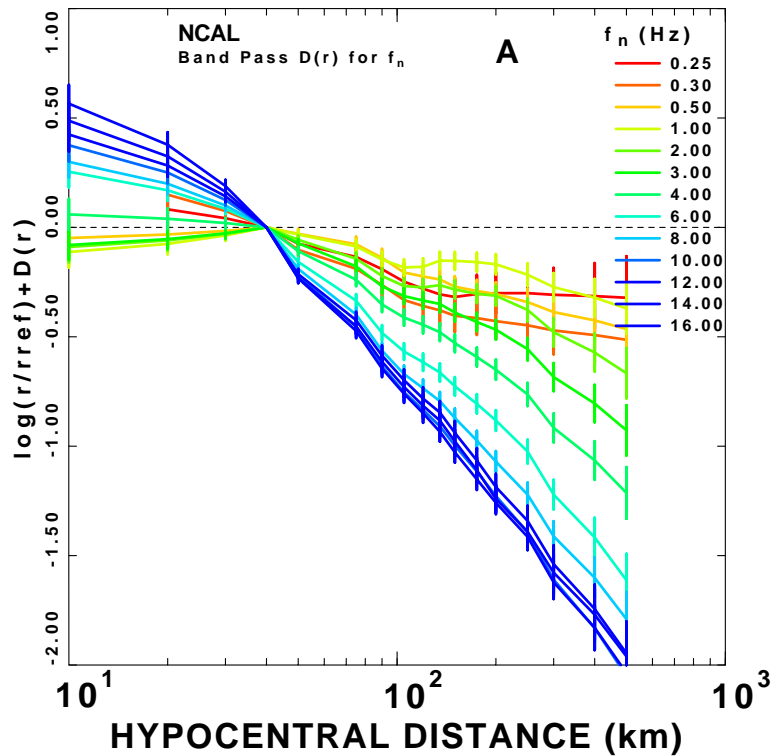
Figure 6.4: The reduced attenuation function $D(r, f)$ obtained from the regression of the 3 component bandpass filtered data at frequencies 0.25-16 Hz.

metrical spreading. The attenuation is very strong for the higher frequencies. The lower frequency results are affected by the noise and therefore are not reliable. The results of the regression on the Z-component bandpass filtered data are shown in the Figure 6.5 for 1.0 to 16.0 Hz
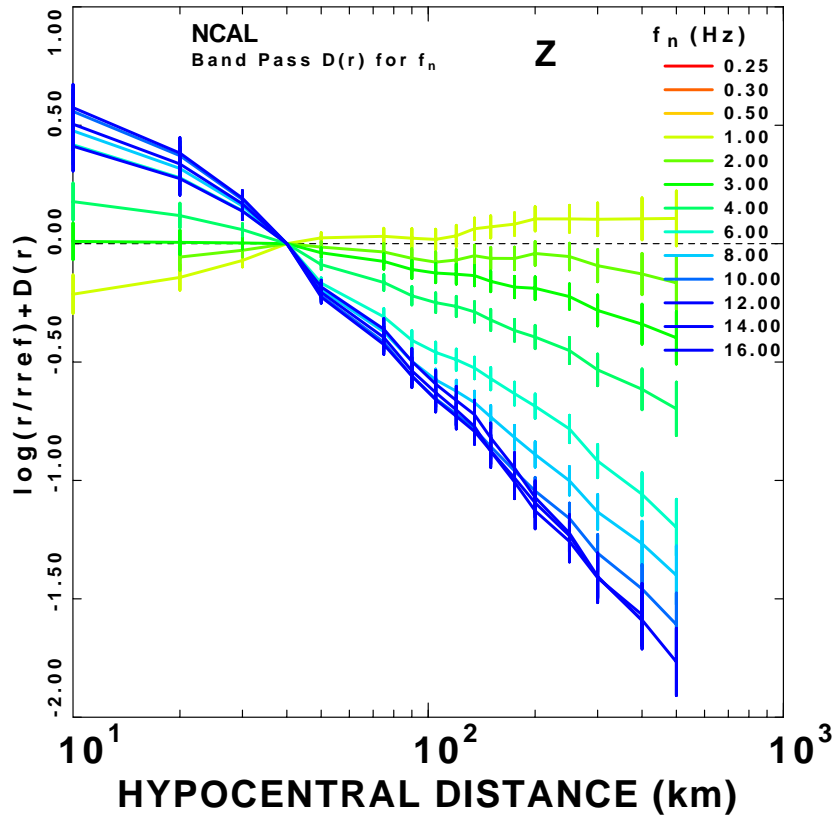
Figure 6.5: The reduced attenuation function $D(r, f)$ obtained from the regression of the Z-component bandpass filtered data at frequencies 1-16 Hz.

The results for 3- and Z-component data show that the study area of northern California is highly attenuating. The effect of the reflections from Moho are not strong. In fact for some frequencies it can not be seen in the figures. The rapid decrease in the amplitude is more clear for higher frequencies (5-16 Hz) for both 3- and Z-component data.

Two different patterns of attenuation can be seen from the derived $D(r, f)$ function. The modeling results for the $D(r, f)$ function confirmed the validity of two different types of attenuation for two different group of frequencies.

65

6.2 Parameterization of bandpass filtered velocity

I tried to model the $D(r, f)$ in both the time and Fourier domain and for different components. Because of the low quality of the data I could only model the 3-component bandpass $D(r, f)$ in a reliable way. The modeling of Z-component gives similar results but with higher residuals. For the 3-component bandpass $D(r, f)$, one attenuation function alone can not fit the whole frequency range (Fig. 6.6). The col-
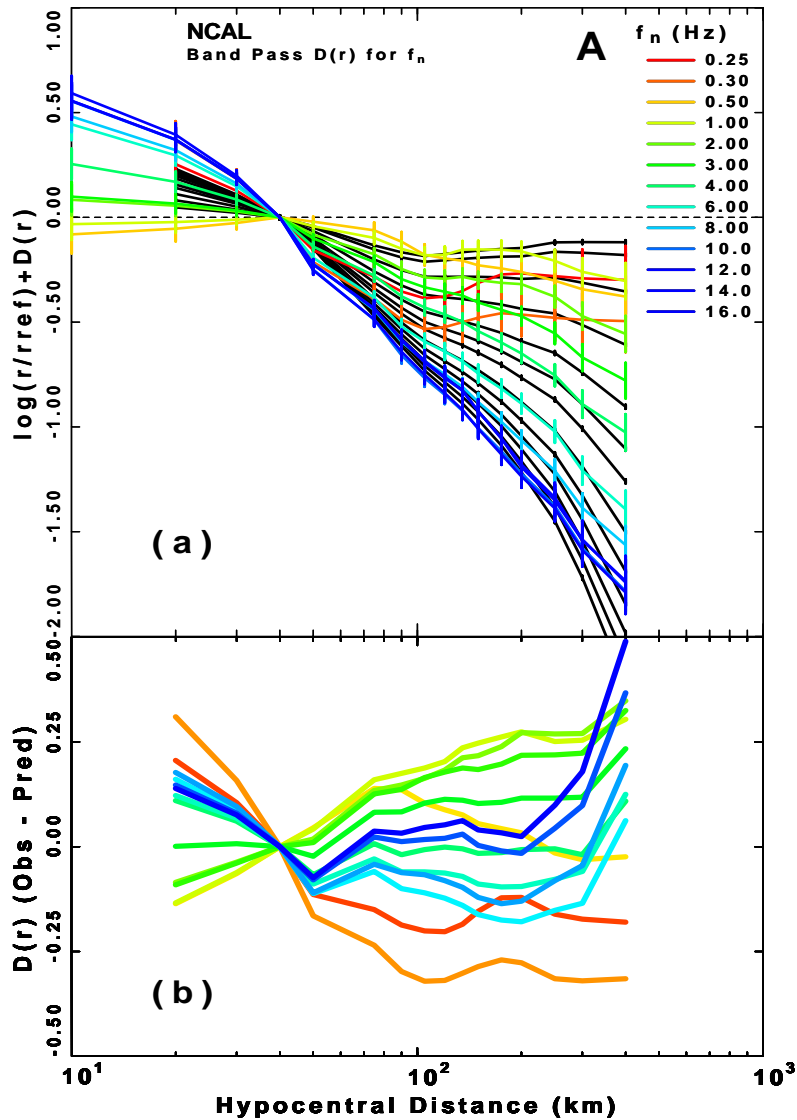
Figure 6.6: (a) observed (colored) and modeled (black) 3-component bandpass velocity attenuation function for 0.25-16 Hz in the northern California (b) The residuals of the model fit to the observed $D(r, f)$.

ored and black curves represent the observed and the modeled $D(r, f)$ respectively. The residuals of the fit for each frequency in the bottom panel. The residuals for the frequencies 1.0-4.0 Hz are very high. The residuals for the frequencies 6.0-16.0 Hz are small. The lowest frequency range 0.25-0.50 have been given a low weight and their residuals are not considered to be important. Many different runs of GA and different models of visual fitting gave the same results: One frequency range fits well but the other have high residuals. The total residuals from GA were also very high. To get around this problem (assuming that 3-component bandpass $D(r, f)$ from the regression are acceptable), I modeled the $D(r, f)$ using two sets of different attenuation functions for the two frequency ranges 1-4 Hz and 6-16 Hz. Table 6.2 shows the modeling results of the 3-component bandpass $D(r, f)$ using two different sets of parameters

Modeling bandpass $D(r, f)$- northern California		
Parameter	Frequency (1-4 Hz)	Frequency (6-16 Hz)
Q_0	210 ± 40	210 ± 40
η	0.55 ± 0.10	0.50 ± 0.10
$g(r)$	$r^{-1.0} \quad r < 40\,km$ $r^{-0.8} \quad 40 < r < 70\,km$ $r^{-0.6} \quad 70 < r < 100\,km$ $r^{-0.5} \quad r > 100\,km$	$r^{-1.3} \quad r < 40km$ $r^{-1.5} \quad 40 < r < 70\,km$ $r^{-1.2} \quad 70 < r < 100\,km$ $r^{-0.8} \quad r > 100\,km$

Table 6.2: Intrinsic attenuation and geometrical spreading parameters derived by modeling 3-component bandpass $D(r, f)$ in the northern California.

The best models for the two frequency ranges have been obtained using GA modeling and were tested by visual fitting. The errors of Q_0 and η are computed using the bootstrap method. The geometrical spreading for the frequency range 6-16 Hz are much stronger. The Q_0 values are less than the Pacific Northwest. The residual for both frequency groups are low after this modeling.

Boatwright et al. (2003) analyzed peak ground velocity and peak ground acceleration from moderate and large earthquakes in northern California. Their results show that for $r >$100 km, a simple r^{-n} geometrical spreading can not fit the attenuation of peak motions. They instead used a combination of power law (r^{-n}) and an exponential function as $\exp(-\eta r)$. My results are similar to their results in having higher

attenuation due to geometrical spreading. Their results however, show higher attenuation at distances larger than 100 km while I obtained strong effect of geometrical spreading at smaller distances. I have also different results for two frequency groups. Their study area includes some parts of the central California but I have strictly studied the northernmost part. Those differences may be the reason in getting different results. Their values for Q_0 and η are close to my results and more importantly both studies show unusual strong attenuation in northern California.

6.3 Parameterization of bandpass filtered excitation

The results of the excitation modeling for 3-component data set is shown in Figure 6.7. The red curves computed using the best model results by simultaneous
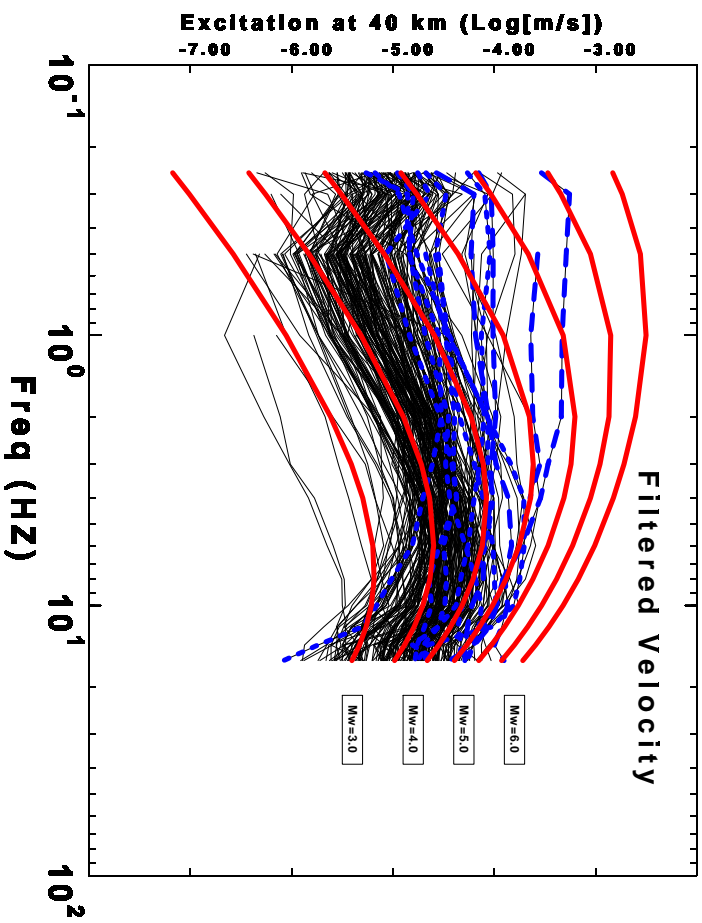
Figure 6.7: Excitation at 40 km for 3-component bandpass filtered data in the northern California. Black curves are observed excitations. Red curves are the theoretical curves for events with moment magnitudes 3.0 to 6.0. Blue dotted curves are the observed excitations of 10 selected events with known moment magnitudes from 3.8 to 5.5.

fitting of 10 events with known moment magnitudes from 3.8 to 5.5. Using the previously determined attenuation function the best model yields $\Delta\sigma = 90 \pm 30$ bars and $\kappa = 0.04 \pm 0.01$ sec. The computed error for the stress drop is high. The residuals of the fit are also larger than those of the Pacific Northwest (Fig. 6.8). In summary the
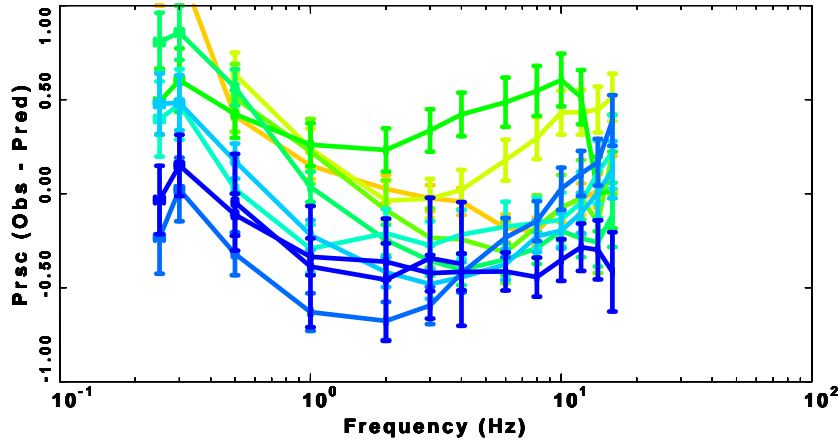
Figure 6.8: Residuals of the fit for the 3-component bandpass filtered excitation at 40 km for the 10 events of the Fig. 6.7.

northern California results show a strong effect of geometrical spreading and lower Q compared to the Pacific Northwest. I did not compute the stress drop for each event. I just modeled the excitation term assuming a constant stress drop. The computed stress drop (90 $\pm$ 30 bars) is in the range of average stress drop of 70 to 100 bars for California (*Atkinson*, 1995).

Chapter 7

Ground motion scaling in central California

The study area of central California is shown in Figure 7.1. There are two distinct areas of seismicity in central California. For this reason I have divided the data set into central east and central west parts. Studying the differences in regional wave propagation are important and interesting. Two ellipses in the figure show the data
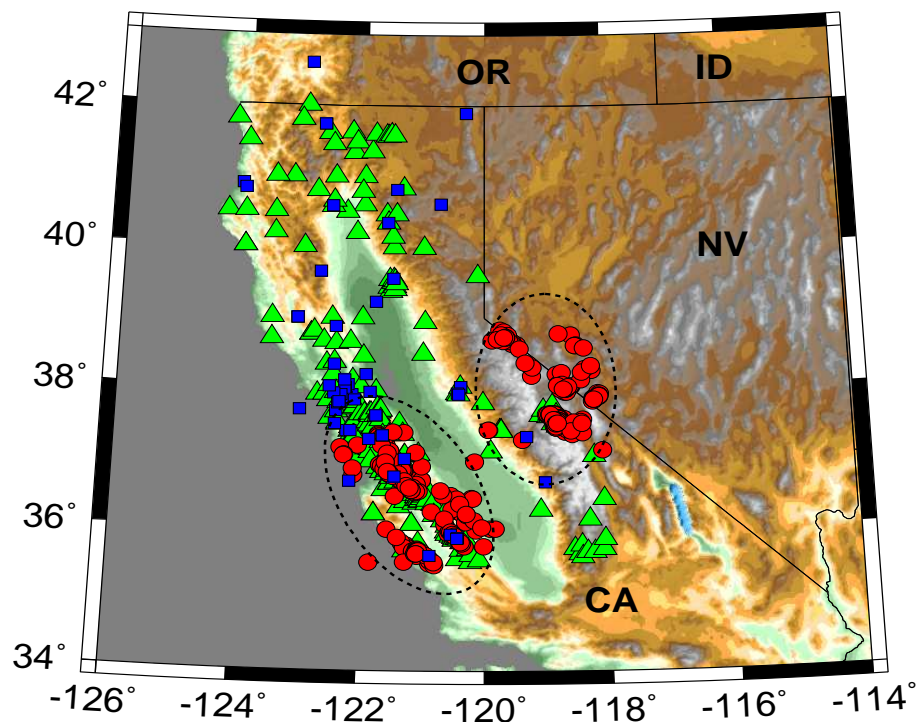
Figure 7.1: Map of stations and events (circles) for this study in central California. Triangles represent short period vertical stations and squares represent 3-component broad band stations.

for each part. The Bay area is located in the central west part of this study. I have done the regression analysis and parameterization for each part.

7.1 Ground motion scaling in central east California

The good quality data set for the eastern part included 1850 seismogram from 240 earthquakes. The data set has both vertical short period and 3-component broadband seismograms. The analysis has been done for the Fourier velocity and bandpass filtered data.

7.1.1 Regression and modeling of Fourier velocity data

Figure 7.2 shows the regression results for the propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the Fourier velocity data. Top panels show the propagation
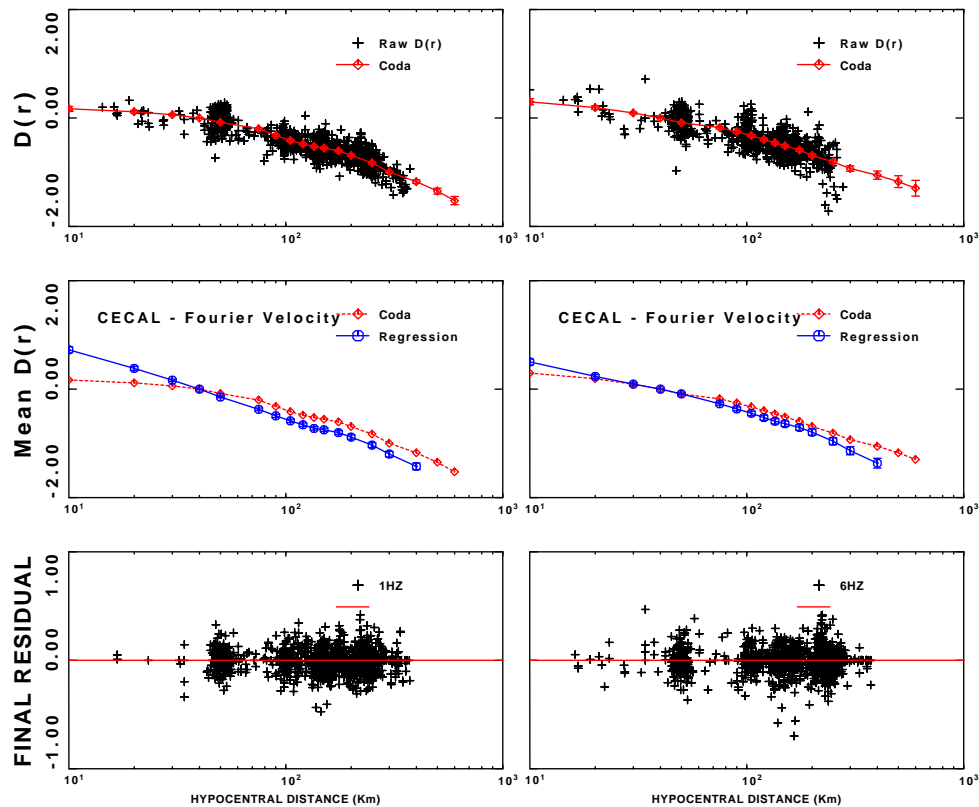
Figure 7.2: Regression results for 3 component Fourier velocity data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions, $D(r)$. Bottom, final residual of the regression analysis.

estimated by coda normalization technique. The mid-panels show the $D(r)$ from the coda and the regression analysis. The lower panels are the final residuals of the regression analysis. The two estimates of $D(r)$ agree in the distances with enough data points. Figure 7.3 shows the 3-component $D(r)$ at different frequencies obtained from the regression of the Fourier velocity data. The attenuation pattern is very dif-
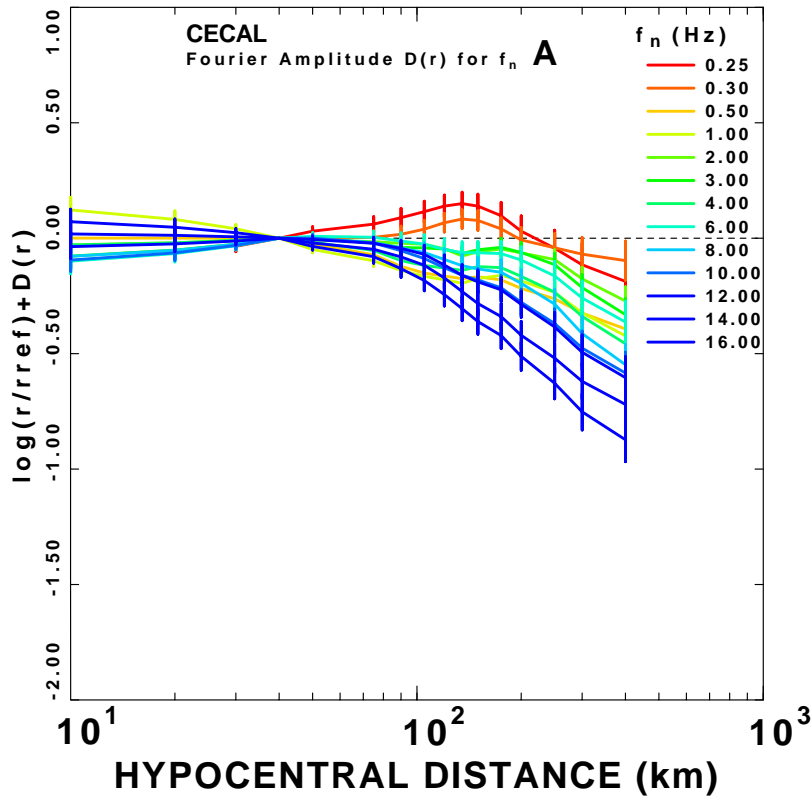
Figure 7.3: The reduced attenuation function $D(r, f)$ obtained from the regression on the 3 component Fourier velocity spectra at frequencies 0.25-16 Hz.

ferent from northern California. The central east attenuation doesn't need a strong geometrical spreading for modeling $D(r, f)$.

GA modeling and visual fitting give very close results for the parameters. Table 7.1 shows the parameter bounds I have used for both central east and central west California in modeling the observations of $D(r, f)$. In general, modeling the $D(r)$ in central east California is easier because it has a smoother shape compared to the Pacific Northwest and northern California. It means a simpler geometrical spreading is needed even though I have used the same distance nodes in all regres-

GA modeling of Fourier $D(r)$		
parameter	lower bound	upper bound
n	-1.5	0.6
Q	100	800
η	0.1	1.0
number of generations = 50		
number of models = 50		

Table 7.1: GA modeling parameters for 3 component Fourier velocity $D(r)$ in the central California.

sions. I have first used a simple r^{-1} for all distances but it required a Q_0 value of 380 for the best fit. This value of Q_0 seems to be high for the California so I decided to use a more complex form of geometrical spreading for the central east California. But this form of $g(r)$ is still simpler than those of the Pacific Northwest and central California. The modeling results using a different geometrical spreading yields $Q_0=$ 290, $\eta = 0.59$, and

$$g(r) = \begin{cases} r^{-1.0} & r < 40 \text{ km} \\ r^{-0.7} & 40 < r < 100 \text{ km} \\ r^{-0.4} & r > 100 \text{ km} \end{cases}$$

The results of error analysis for this modeling is shown in the Table 7.2. I have used 3 different $g(r)$ at 3 distances. The bootstrap mean for Q_0 is considerably higher the best model value. The results of the bootstrap depends on the original data and the model we use in our analysis. I will use the results of bandpass filtered velocity since it is the better data set.

Bootstrap and best model results-Fourier $D(r)$			
parameter	best model	bootstrap mean	bootstrap σ
Q_0	290	360	41
η	0.59	0.67	0.08
n_1	-1.0	-1.01	0.09
n_2	-0.8	-1.09	0.08
n_3	-0.5	-0.98	0.07
number of GA runs = 100			
No. of models in each run = 50			

Table 7.2: Bootstrap error estimate of modeled propagation parameters for Fourier $D(r, f)$ in the central east California.

Figure 7.4 shows the observed and modeled 3-component Fourier velocity $D(r, f)$.

The residual of 1.0 Hz is high for this modeling. Frequencies 0.25, and 0.3 Hz have
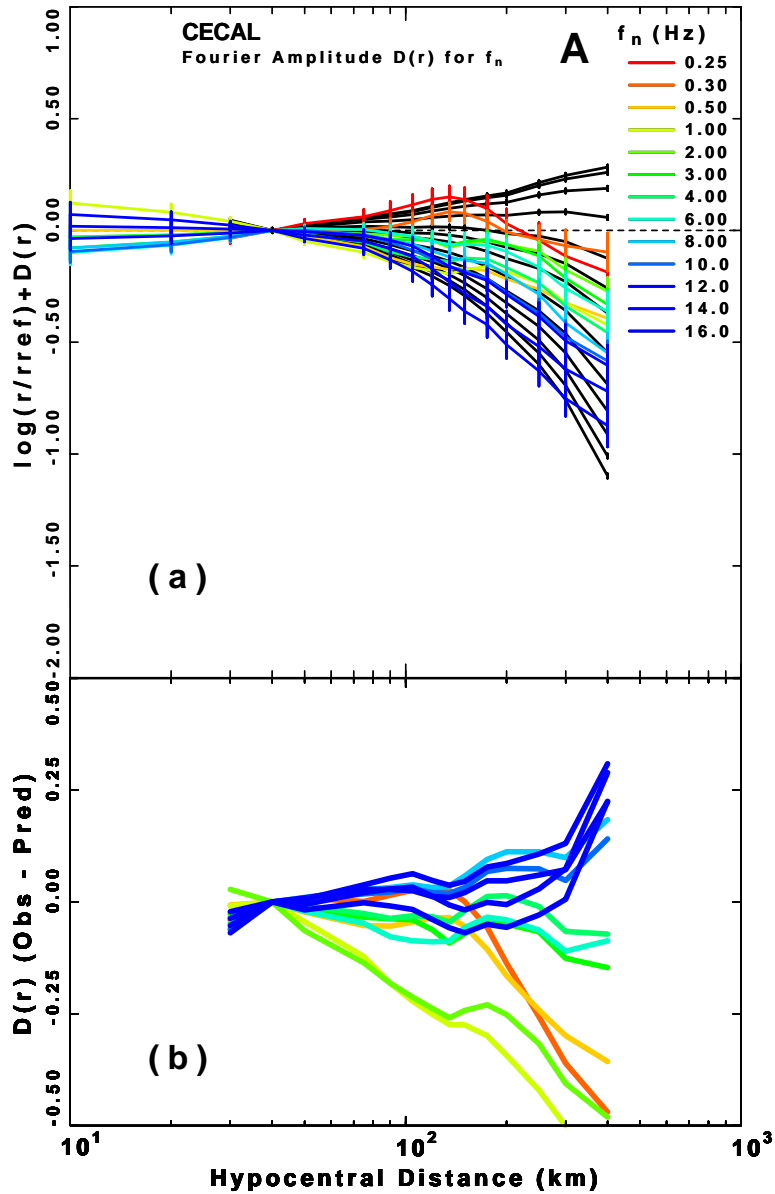
Figure 7.4: (a) Observed (colored) and modeled (black) 3-component Fourier velocity attenuation function for 0.25-16 Hz in the central east California (b) The residuals of the model fit to the observed $D(r, f)$.

also high residuals at distances larger than 200 km. For the modeling of the bandpass data I have given a small weight to the lower frequencies to get the results of the best model based more on higher frequencies. The results of vertical and horizontal component for the Fourier domain are not good enough for modeling. The combined 3-component data set however, can be modeled in both domains.

7.1.2 Regression and modeling of bandpass filtered data

Figure 7.5 shows the regression results for propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the bandpass filtered data. In general the coda and regression
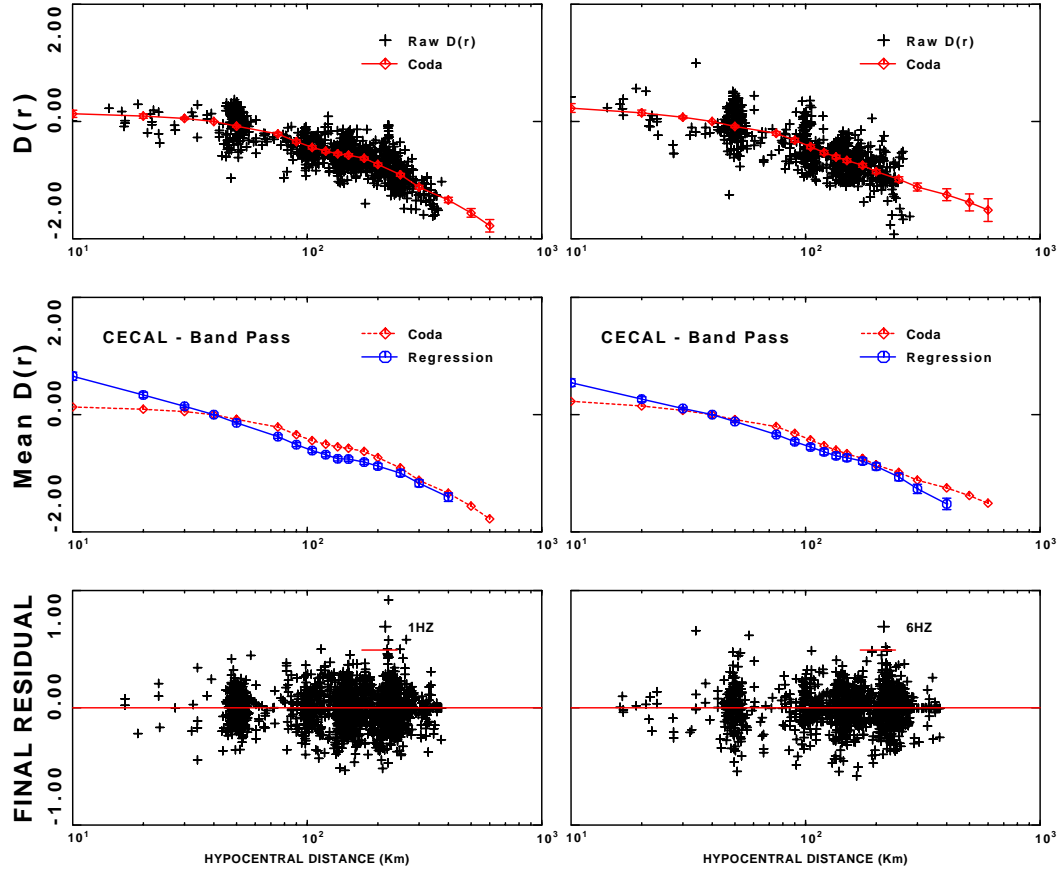
Figure 7.5: Regression results for 3 component bandpass filtered data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions, $D(r)$. Bottom, final residual of the regression analysis.

estimates of $D(r)$ in here agree better than the Fourier velocity data. I got the same results for other frequencies which are not shown in here. The regression derived $D(r)$ for both domains show the same features (except for lower frequencies) with a little difference in the amplitude. This is true for the case when the Fourier data are not affected by noise. The vertical and short period data are not very similar in different domains.

Figure 7.6 show the 3-component $D(r)$ at different frequencies obtained from

the regression of the Fourier velocity data. Comparing this figure with Figure 7.3
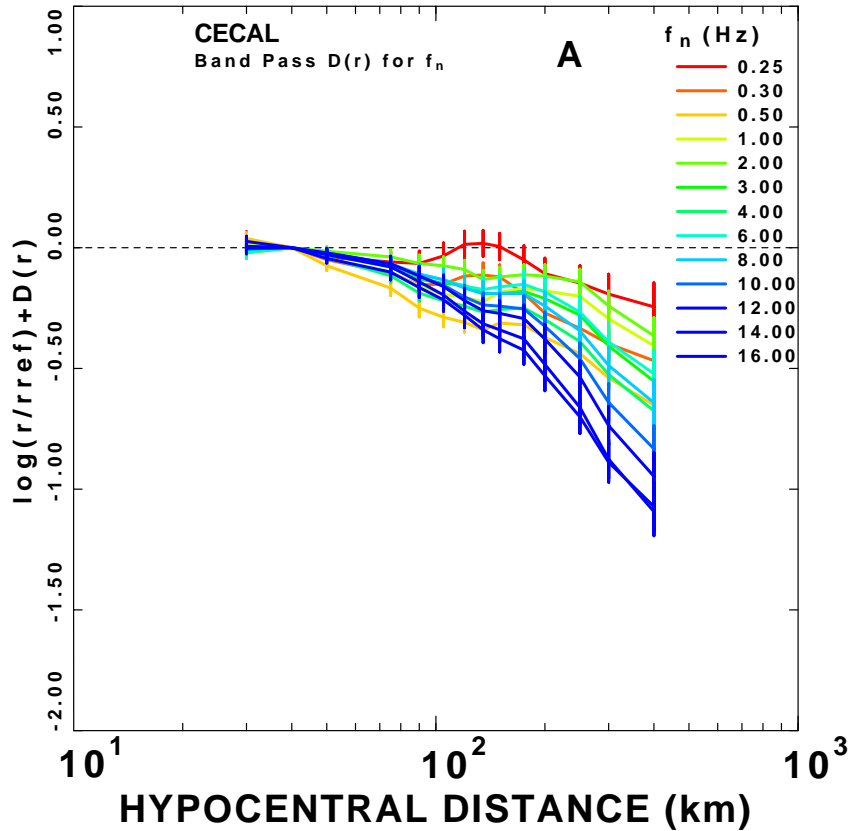
Figure 7.6: The reduced attenuation function $D(r, f)$ obtained from the regression on the 3 component bandpass filtered data at frequencies 0.25-16 Hz.

(Fourier data) shows that for distances higher than 40 km, the attenuation functions are similar. The only difference is that for Fourier data the higher frequencies are more apart which means higher values of η is needed for modeling.

Figure 7.7 shows regression derived horizontal and vertical component $D(r)$ at different frequencies. The $D(r)$ curves of the horizontal component for higher frequencies are separated from the lower frequencies. The modeling of this curves is not easy since it needs a frequency dependent η and maybe different geometrical spreading. It is likely that the data set is not good enough for a regression on a single component like horizontal or vertical. In fact the number of observations (data points) is not so high for central east California and because of this problem I decided to model only the 3-component data in both domains to get more robust results.

Figure 7.8 shows the measured and model fit durations at frequencies 0.3, 2.0, 3.0
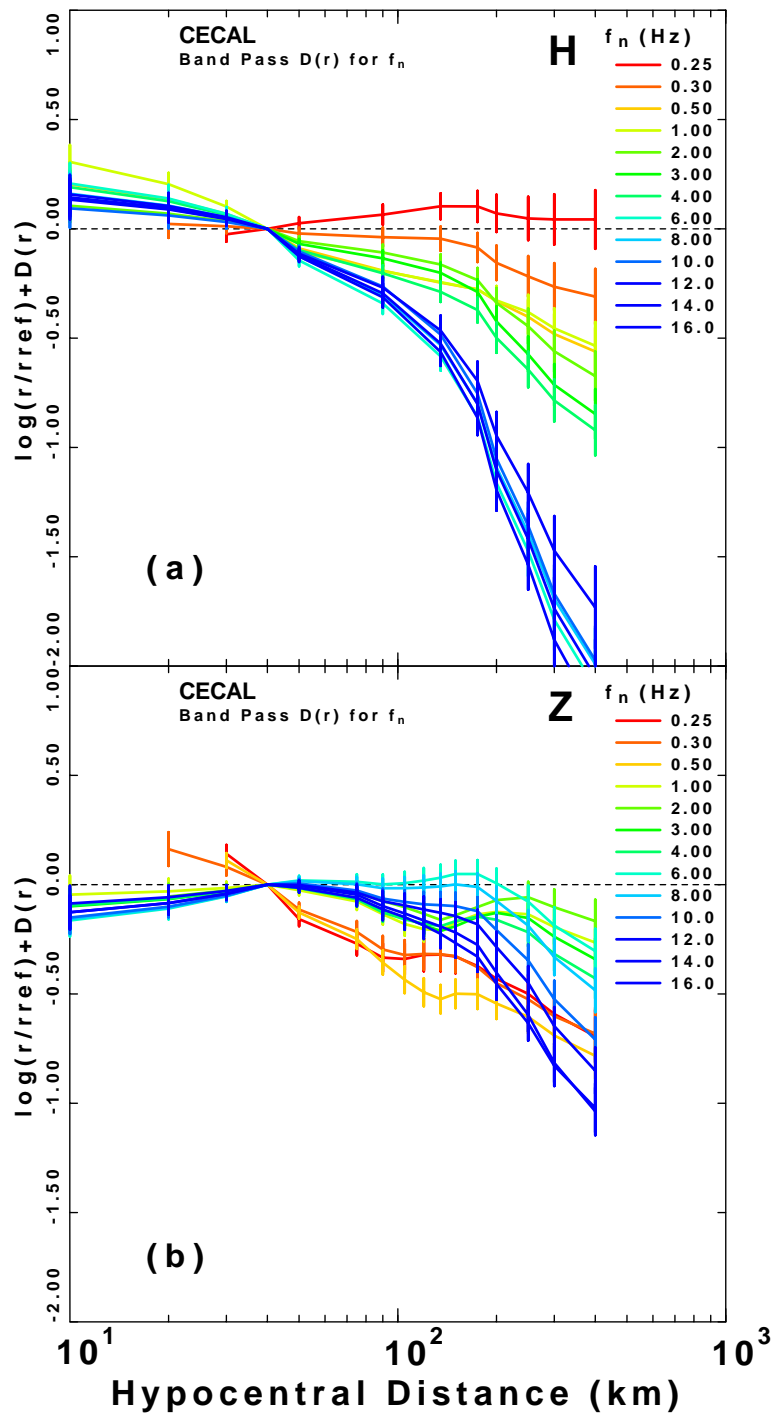
Figure 7.7: The reduced attenuation function $D(r, f)$ obtained from the regression on the horizontal (top) and vertical (bottom) component bandpass filtered data at frequencies 0.25-16 Hz.

and 6.0 Hz. The duration estimates are used in the modeling of bandpass filtered data. Except for the lowest frequencies there is not much scatter in the duration
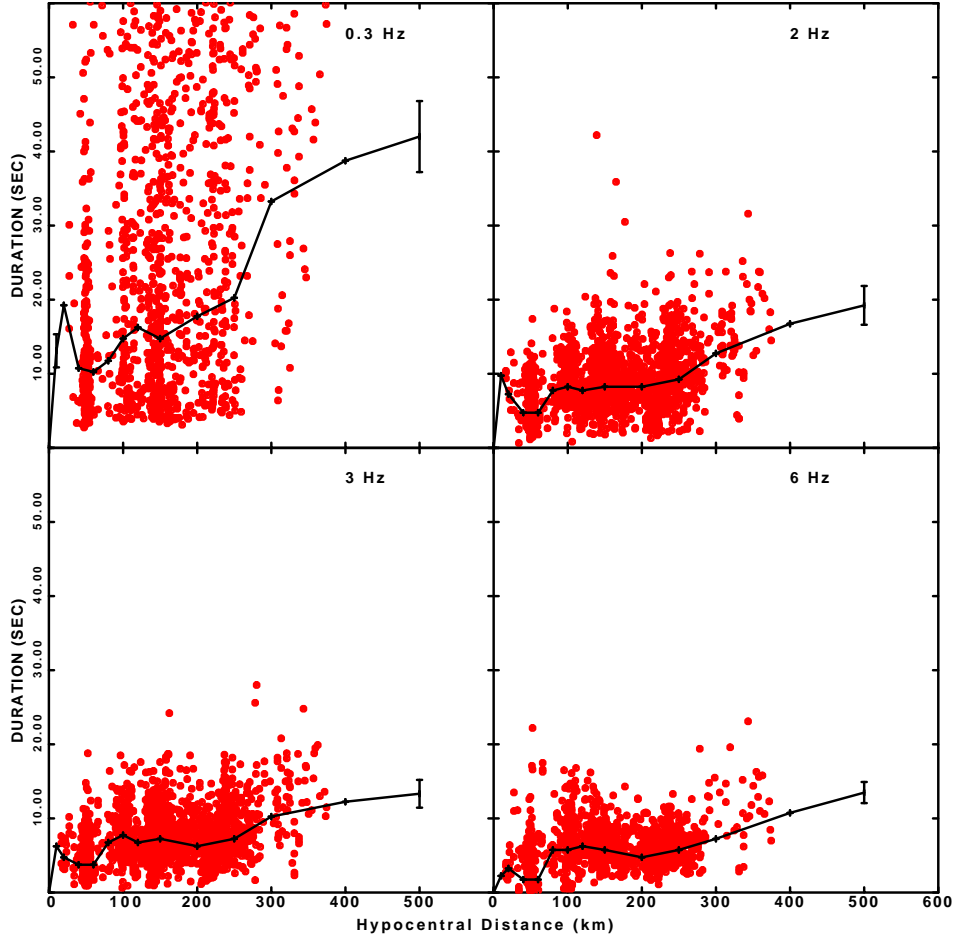
Figure 7.8: Durations at 0.3, 2.0, 3.0, and 6.0 Hz for central east California. Red circles indicate individual duration estimates and solid lines represent the duration measured using a median value method.

data.

Using the measured durations the modeling results for the 3-component bandpass filtered data in central east California yields,$Q_0 = 280$, $\eta = 0.50$, and

$$
g(r) = \begin{cases}
r^{-1.0} & r < 40 \text{ km} \\
r^{-0.8} & 40 < r < 100 \text{ km} \\
r^{-0.5} & r > 100 \text{ km}
\end{cases}
$$

The parameters are very close to those of the 3-component Fourier velocity data. Figure 7.9 shows the modeling results of the 3-component bandpass filtered $D(r, f)$. Except for the 3 lowest frequencies, the residuals are in an acceptable range. The

78
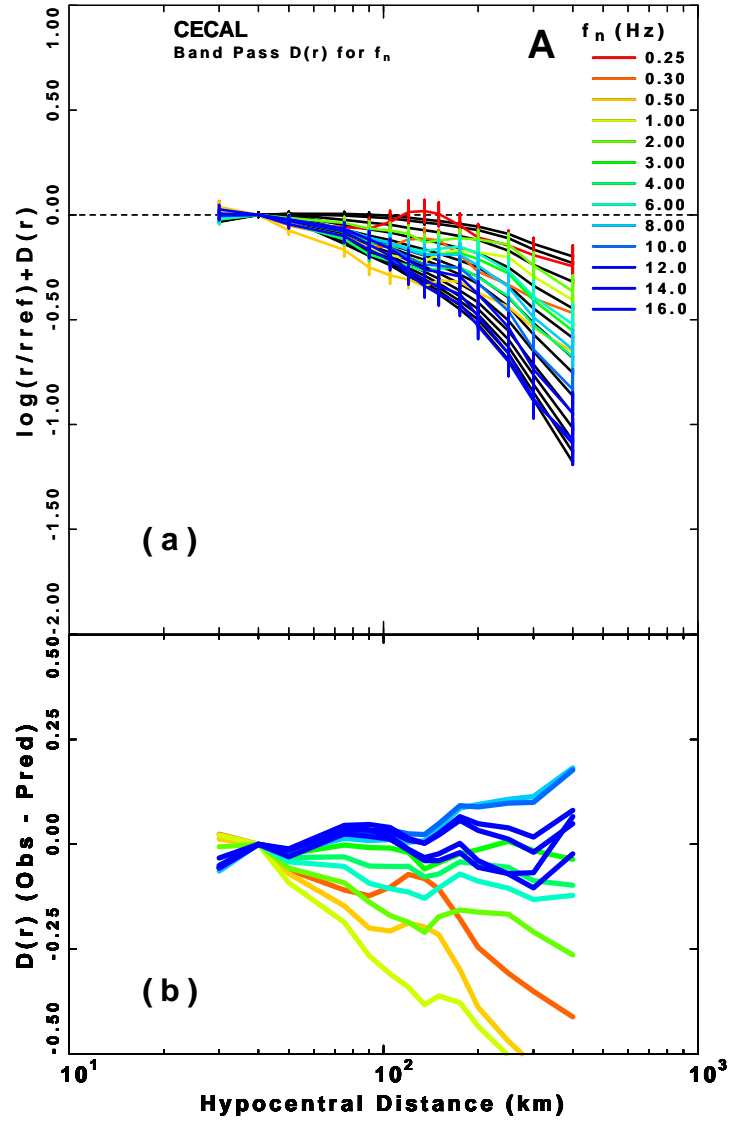
Figure 7.9: (a) Observed (colored) and modeled (black) 3-component bandpass filtered velocity attenuation function for 0.25-16 Hz in the central east California (b) The residuals of the model fit to the observed $D(r, f)$ for 1.0-16 Hz.

estimated errors for Q_0 and η are 36 and 0.09 respectively. The estimated error for the geometrical spreading power is 0.095 at all distances. So $Q_0 = 280 \pm 36$ and $\eta = 0.50 \pm 0.09$ are the final selected values for the central east California.

Using the modeled attenuation parameters I have modeled the 3-component excitation term in this region. Figure 7.10 shows the observed and modeled excitation terms. A simultaneous modeling has been done for 11 events with known moment magnitudes from 3.6 to 5.5 using a constant stress drop model. The parameters of

Excitation at 40 km (Log[m/s])

-6.00 -5.00 -4.00 -3.00 -2.00

Filtered Velocity

Mw=3.6 Mw=4.2 Mw=4.5 Mw=5.2 Mw=5.5

Mw=3.0 Mw=4.0 Mw=5.0 Mw=6.0

10^{-1} 10^{0} 10^{1} 10^{2}

Freq (HZ)
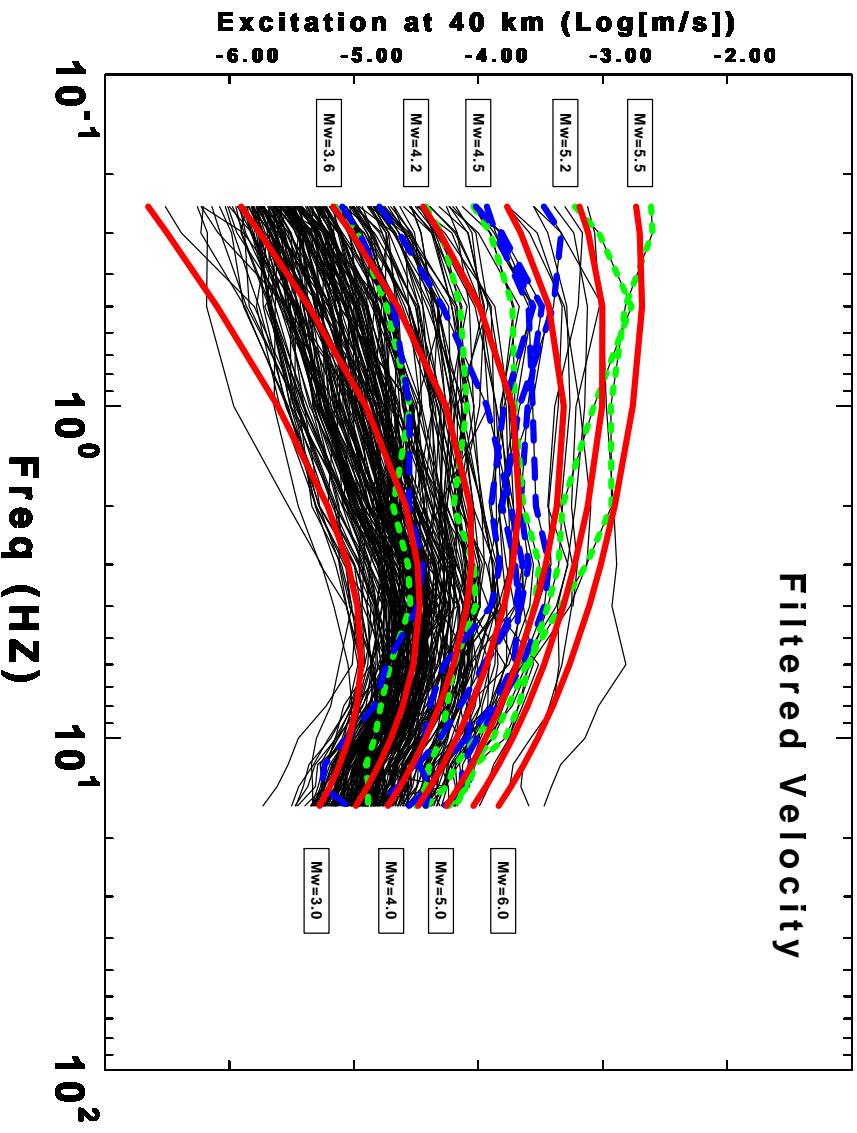
Figure 7.10: Excitation at 40 km for 3-component bandpass filtered data in the central east California. Black curves are observed excitations. Red curves are the theoretical curves for events with moment magnitudes 3.0 to 6.0. Blue and green dotted curves are the observed excitations of 11 selected events with known moment magnitudes from 3.6 to 5.5. The small boxes in the left show the moment magnitude of green dotted curves.

the best model are $\Delta\sigma = 15 \pm 7$ bars and $\kappa = 0.028 \pm 0.008$ sec. The residuals of the fit are shown in Figure 7.11. The residuals are high in the lower frequencies. A stress drop of 15 bars seems to be very low for the California. Again a constant stress drop assumption may not be true for the events of this region and using only a few events may not be the best way to estimate the stress drop. To compare different estimates of stress drop, I have done several parameterizations using more events. In the data set 43 events had known seismic moment and moment magnitude. I have used Berkeley regional moment tensor solutions for all moment magnitudes of the events in north-ern and central California (http://seismo.berkeley.edu/ dreger/mtindex.html). For another estimation of stress drop I used all 43 events and modeled all of them with a
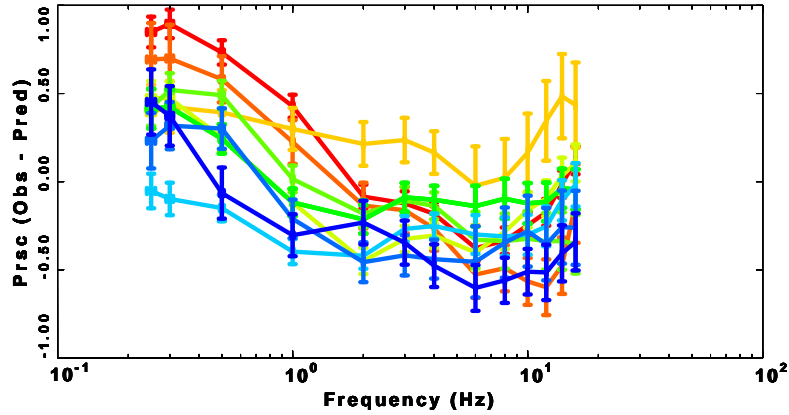
Figure 7.11: Residuals of the fit for the 3-component bandpass filtered excitation at 40 km for the 11 events of the Fig. 7.10.

constant stress drop. The results of this modeling give $\Delta\sigma = 24$ bars and $\kappa = 0.034$ sec. In the next step I modeled the excitation terms for each event. The result of this modeling is shown in Figure 7.12 in which κ is plotted versus the stress drops for 43 events. The results of the modeling in Figure 7.12 show that the individual
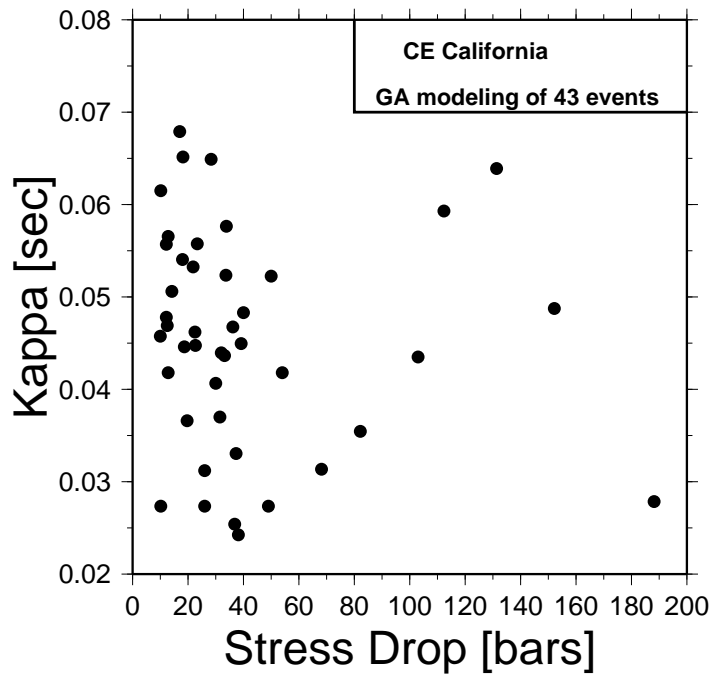
Figure 7.12: κ versus stress drop for 43 events with known moment magnitudes in the central east California.

stress drops vary from 10 to 190 bars with most of them lie between 10 and 60 bars. the means and standard deviations are 44 and 32 bars for stress drop and 0.045 and

0.012 sec for κ. The standard deviation of stress drops is high and comparable to the mean. Although there is no strong correlation between $\Delta\sigma$ and κ in Figure 7.12, in fact these two parameters trades off with each other (*Boore et al.*, 1992b). I also tried to use a fix κ and repeated the modeling using $\kappa = 0.045$ for all events. this value is the mean value of 43 different κs (Fig. 7.12). Figure 7.13 shows stress drop versus moment magnitude for two different case of variable and fixed κ. In both cases
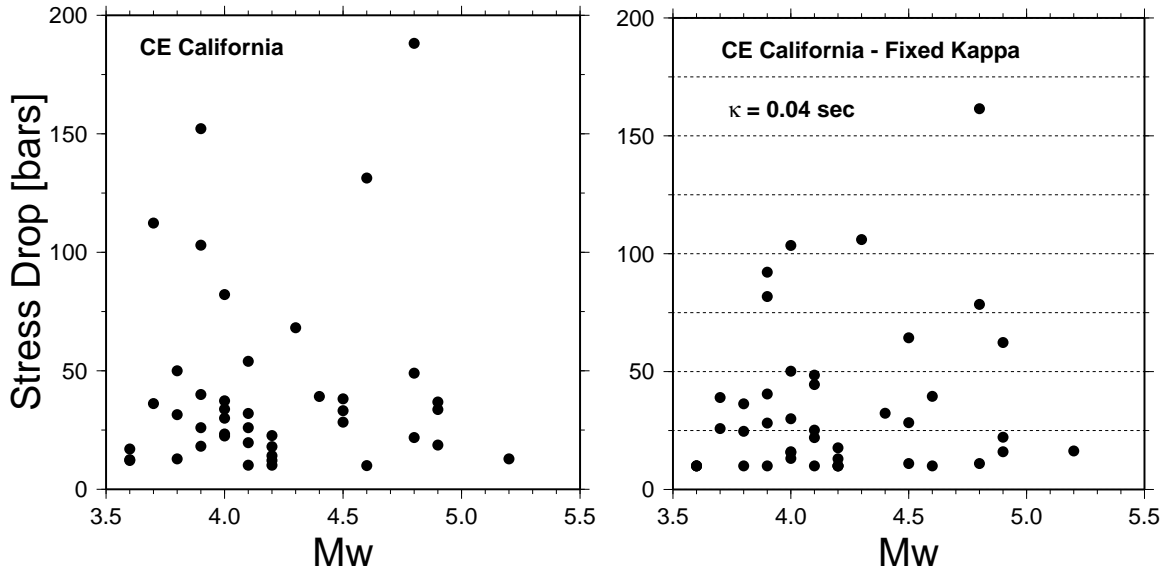
Figure 7.13: (left) Stress drop versus moment magnitude for 43 events in the central east California. Each event has a different κ (Fig 7.12). (right) Same as left but a fixed value of 0.04 for κ was used in modeling.

the majority of the events have stress drops between 10 and 60 bars. The mean and standard deviation of stress drop for the case of fixed κ are 36 and 26 bars. Since the mean of stress drops for two cases are not significantly different, using a constant κ in the modeling seems to be reasonable.

As the final attempt in modeling the excitation terms I tried to compute the errors of stress drop for each event. The error and the mean were computed from 100 runs of the GA code for each single event. The means of 100 runs are not necessarily the same as the best model result for each event. The average stress drop of all events from this computation is 48 bars. Figure 7.14 shows the individual error computation results.

I have summarized the results of this part in the Table 7.3. The first row of
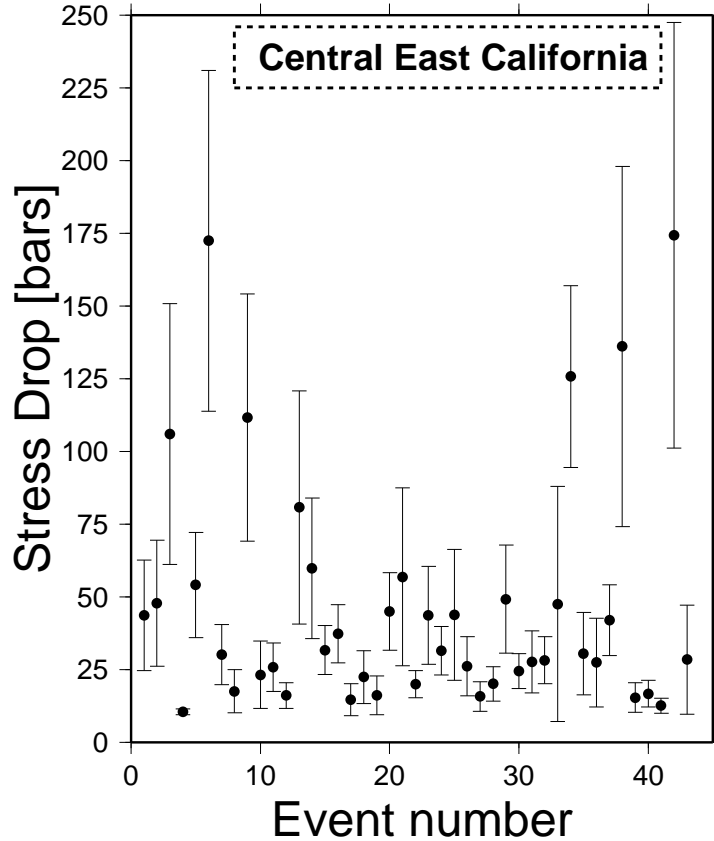
Figure 7.14: Estimated errors of stress drops for 43 events in central east California. Each individual event has been modeled using a Brune point source model. A bootstrap analysis has been done for each event. The stress drop values are the mean values of each bootstrap. The average of all stress drops is 48 bars

the results show the estimated parameters obtained from modeling 11 events using a constant stress drop (Fig. 7.10). second row show the same results for modeling all 43 events with known moment magnitude using a constant stress drop model. The results of the third row were obtained by modeling each event separately in which each event had different stress drop and κ. The results of stress drop and κ for this case are shown as the mean and standard deviation of all computed values. Individual results are shown in Fig 7.13. The bootstrap results for this case is shown in Figure 7.14 which are actually the errors of each computed stress drop. The fourth row is the same as third row except that I used a fixed κ. The individual results are shown in the right panel of Figure 7.13. I haven't done a bootstrap for this part (fixed κ) because the stress drop results are not significantly different from the case of the variable κ.

Based on the results of different parameterizations I selected an average value of

Bootstrap and best model results for $\Delta\sigma$ and κ				
No. of events used	best model result		bootstrap	
	$\Delta\sigma$ (bars)	κ (sec)	$\Delta\sigma$ (mean and std)	κ (mean and std)
11 (const. $\Delta\sigma$)	15	0.028	19 ± 7	0.036 ± 0.008
43 (const. $\Delta\sigma$)	24	0.034	30 ± 12	0.029 ± 0.009
43 (non-const. $\Delta\sigma$)	Ave: 44 ± 32	Ave: 0.045 ± 0.012	-	-
43 (non-const. $\Delta\sigma$)	Ave: 36 ± 26	0.045 (fixed)	-	-

Table 7.3: Best model and error analysis results of stress drop and κ for 43 events with known moment magnitudes in the central east California. Brune's source model is used for all events.

45 ± 20 bars for stress drop in my study area of central California.

7.2 Ground motion scaling in central west California

The left cluster of earthquakes in Figure 7.1 shows the earthquakes that I have used to study the ground motion scaling in the central west California. The data set includes 2660 good quality waveforms. The number of horizontal components in this Region is greater than for the central east California. The central west data set is better in general; It also covers smaller distance ranges and has more observations compared to the central east data set.

7.2.1 Regression and modeling of Fourier velocity data

Figure 7.15 shows the regression results for propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the Fourier velocity data. Compared to the results of the same frequency for the central east (Fig. 7.2), The coda and regression results agree well in all distances. There are also more data points in the distances less than 40 km. The data cover a distance range up to 300 km.

I have modeled the Fourier velocity $D(r)$ to obtain the propagation parameters. The same parameter bounds of other regions have been used in parameterizations and the same procedure has been repeated in here by starting from the 3-component Fourier velocity data. The best model parameters for $D(r)$ are $Q_0 = 250$, $\eta = 0.30$,

Figure 7.15: Regression results for 3 component Fourier velocity data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions, $D(r)$. Bottom, final residual of the regression analysis.

and

$$
g(r) = \begin{cases}
r^{-1.3} & r < 40 \text{ km} \\
r^{-1.1} & 40 < r < 60 \text{ km} \\
r^{0.0} & 60 < r < 90 \text{ km} \\
r^{-0.6} & 90 < r < 150 \text{ km} \\
r^{-0.9} & r > 150 \text{ km}
\end{cases}
$$

The estimated errors of Q_0 and η are 26 and 0.07 respectively. Figure 7.16 shows the observed and modeled 3-component Fourier velocity $D(r, f)$. Except for 0.25 and 0.3 Hz, the residuals are low for distances up to 300 km. These two low frequencies had a very low weight in the modeling. The results show a strong geometrical spreading for distances larger than 150 km. The Q_0 values obtained for central California in this study are larger than some values reported by other studies for different parts

Figure 7.16: (a) Observed (color) and modeled (black) 3-component Fourier velocity attenuation function for 0.25-16 Hz in the central west California (b) The residuals of the model fit to the observed $D(r, f)$.

of California (*Raoof et al.*, 1999; *Benz et al.*, 1997). As mentioned before, Q_0 trades off with η and my η values are smaller. *Raoof et al.* (1999) obtained $Q(f) = 180 f^{0.45}$ for southern California using the regression of 3-component velocity data. *Benz et al.* (1997) obtained $Q(f) = 187 f^{0.56}$ for southern and central California by studying L_g attenuation in the continental US.

7.2.2 Regression and modeling of bandpass filtered data

Figure 7.17 shows the regression results for propagation term, $D(r)$, at frequencies 1.0 and 6.0 Hz on the bandpass filtered velocity data. There is almost a perfect
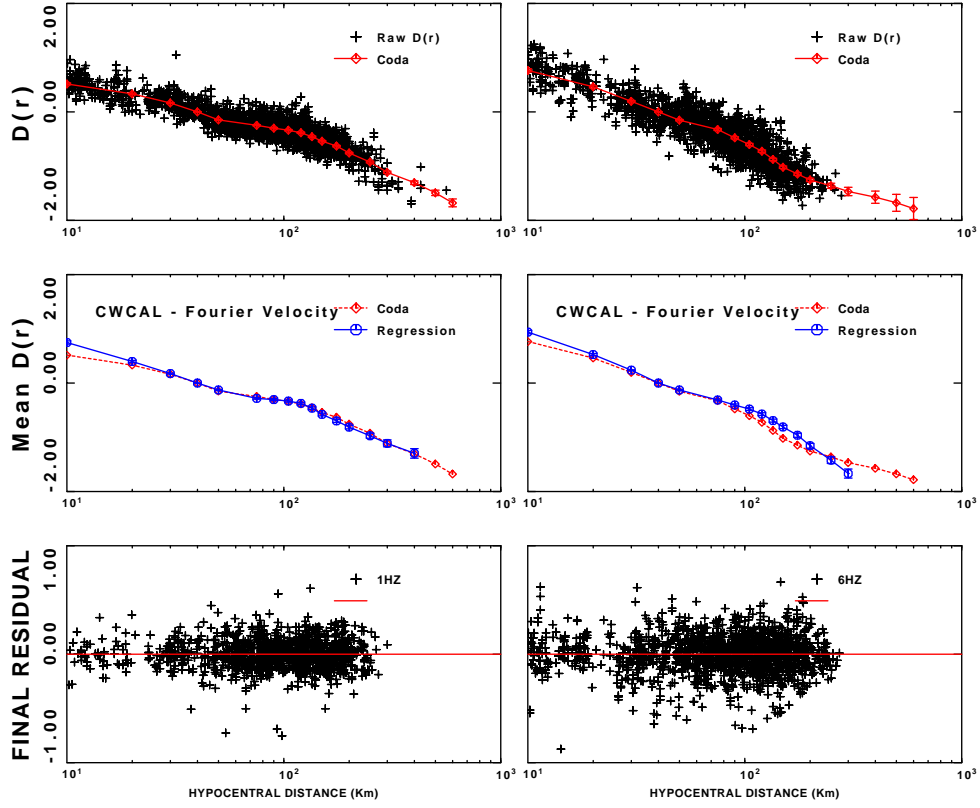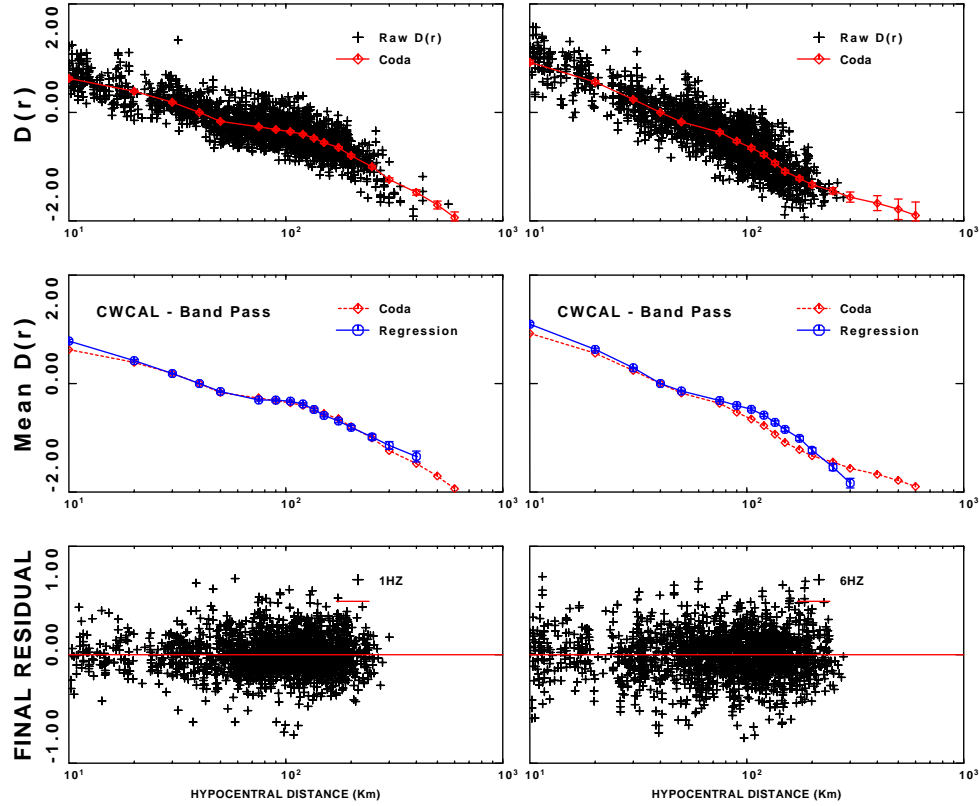

Figure 7.17: Regression results for 3 component bandpass filtered velocity data at 1.0 and 6.0 Hz. Top, coda estimate of $D(r)$ using coda normalization method. Middle, coda and regression propagation functions, $D(r)$. Bottom, final residual of the regression analysis.

agreement between the coda normalization and the regression results. This match can be seen up to 300 km. The curves have been extrapolated beyond that distance since there are no data points for distances more than 300 km.

The same parameters used for the modeling the Fourier velocity $D(r)$, can match the band pass filtered $D(r)$ too, except that the $g(r)$ terms are a little different between the two domains. . I have used the same Q_0 and η of Fourier $D(r)$ and obtained the geometrical spreading term from modeling. I have done this because otherwise the resultant Q_0 values would be high (around 300) and η would be small

(0.25 - 0.28) for the best model. Although the combined $Q_0 f^\eta$ would be an acceptable value, I kept the Q_0 to be constant (= 250) because the value of $Q_0 = 300$ seems to be high for the California region. Figure 7.18 shows the observed and modeled 3-component bandpass filtered velocity $D(r, f)$. Table 7.4 shows the best model and
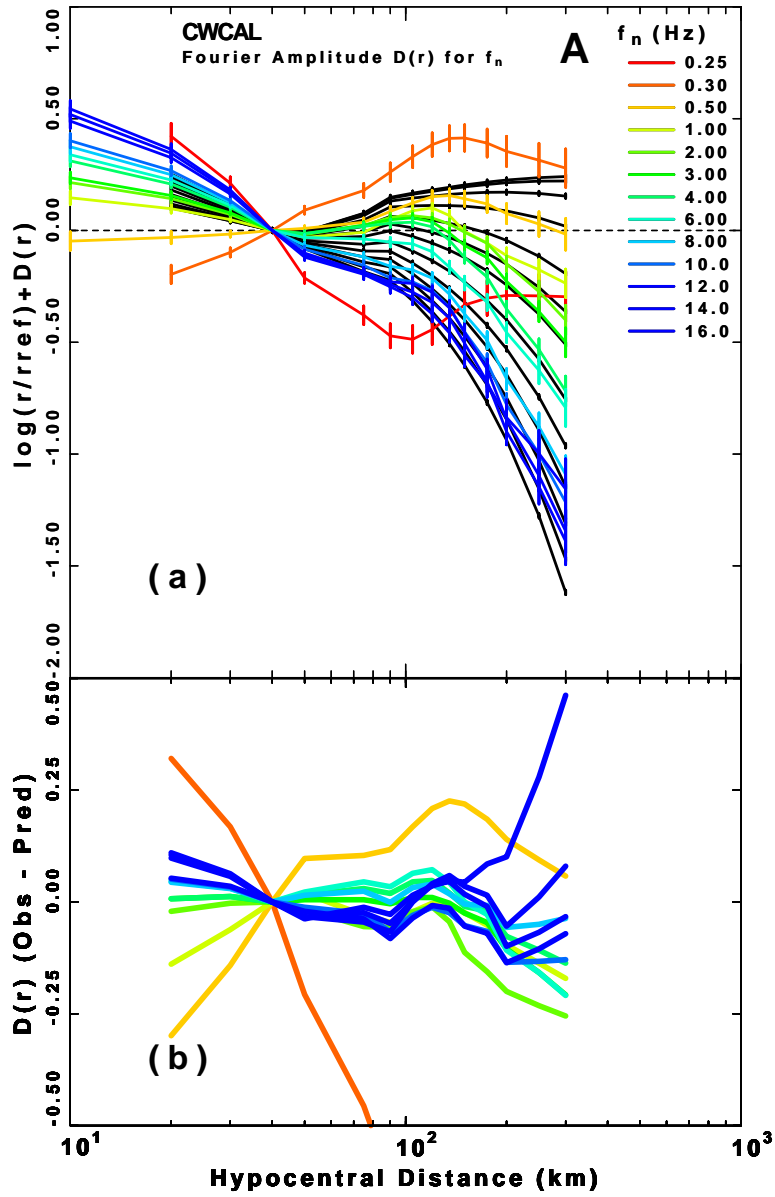

Figure 7.18: (a) Observed (colored) and modeled (black) 3-component bandpass filtered velocity attenuation function for 0.25-16 Hz in the central west California (b) The residuals of the model fit to the observed $D(r, f)$.

error analysis results of the propagation parameters for bandpass filtered $D(r)$. The

Bootstrap and best model results-All bandpass $D(r)$				
parameter	best model	bootstrap mean	bootstrap σ	selected value
Q_0	250	276	19	260 ± 19
η	0.30	0.32	0.05	0.35 ± 0.05
n_1	-1.3	-1.3	0.11	-1.3 ± 0.11
n_2	-1.0	-1.1	0.09	-1.0 ± 0.09
n_3	0.1	0.2	0.02	0.1 ± 0.02
n_4	-0.5	-0.56	0.06	-0.5 ± 0.06
n_5	-0.95	-1.01	0.08	-1.01 ± 0.08
number of GA runs for bootstrap = 100				
No. of models in each run = 50				

Table 7.4: Bootstrap error estimate of modeled propagation parameters for 3 component bandpass $D(r, f)$ in the central wets California.

ns are the geometrical spreading exponents at different distances. The final selected values are $Q(f) = 250^{+19}_{-19} f^{0.30 \pm 0.05}$ and

$$g(r) = \begin{cases} r^{(-1.3 \pm 0.11)} & r < 40 \text{ km} \\ r^{(-1.0 \pm 0.09)} & 40 < r < 60 \text{ km} \\ r^{(0.1 \pm 0.02)} & 60 < r < 90 \text{ km} \\ r^{(-0.5 \pm 0.06)} & 90 < r < 150 \text{ km} \\ r^{(-0.95 \pm 0.08)} & r > 150 \text{ km} \end{cases}$$

For the central west California data set I feel more confident about the results of the combined horizontal and vertical (3-component) data set. This was true for other regions of my study area too because 3-component data set provide a better distance and frequency coverage than a single component data alone. To see the effect of the attenuation I have also done the regression and modeling for vertical and horizontal components in the central west data set. The best model for the bandpass filtered vertical component data gives $Q_0 = 260$, $\eta = 0.30$ and

$$g(r) = \begin{cases} r^{-1.3} & r < 40 \text{ km} \\ r^{-1.2} & 40 < r < 60 \text{ km} \\ r^{0.0} & 60 < r < 90 \text{ km} \\ r^{-0.5} & 90 < r < 150 \text{ km} \\ r^{-0.6} & r > 150 \text{ km} \end{cases}$$

The geometrical spreading terms are a little different from the 3-component results. The differences arises because the $D(r)$ terms for different components are not the same. Figure 7.19 shows the results of modeling the bandpass $D(r)$ for vertical component data. The horizontal component was modeled with the same values of Q_0
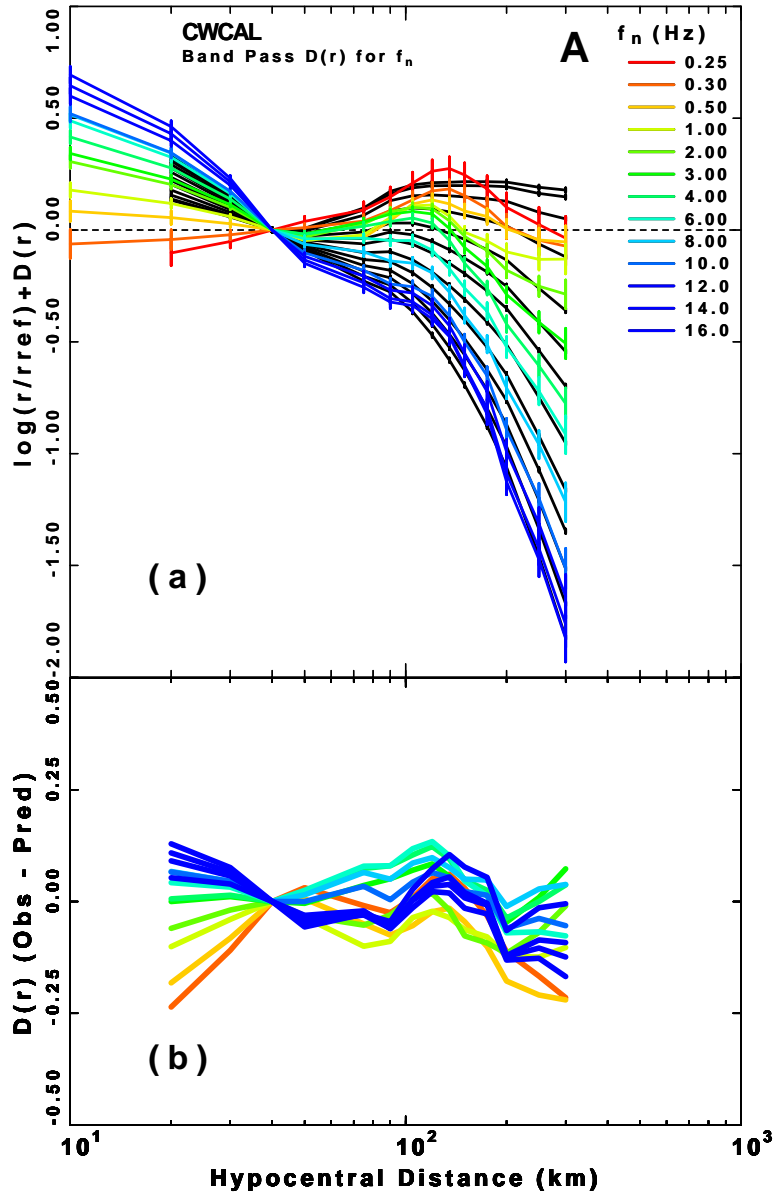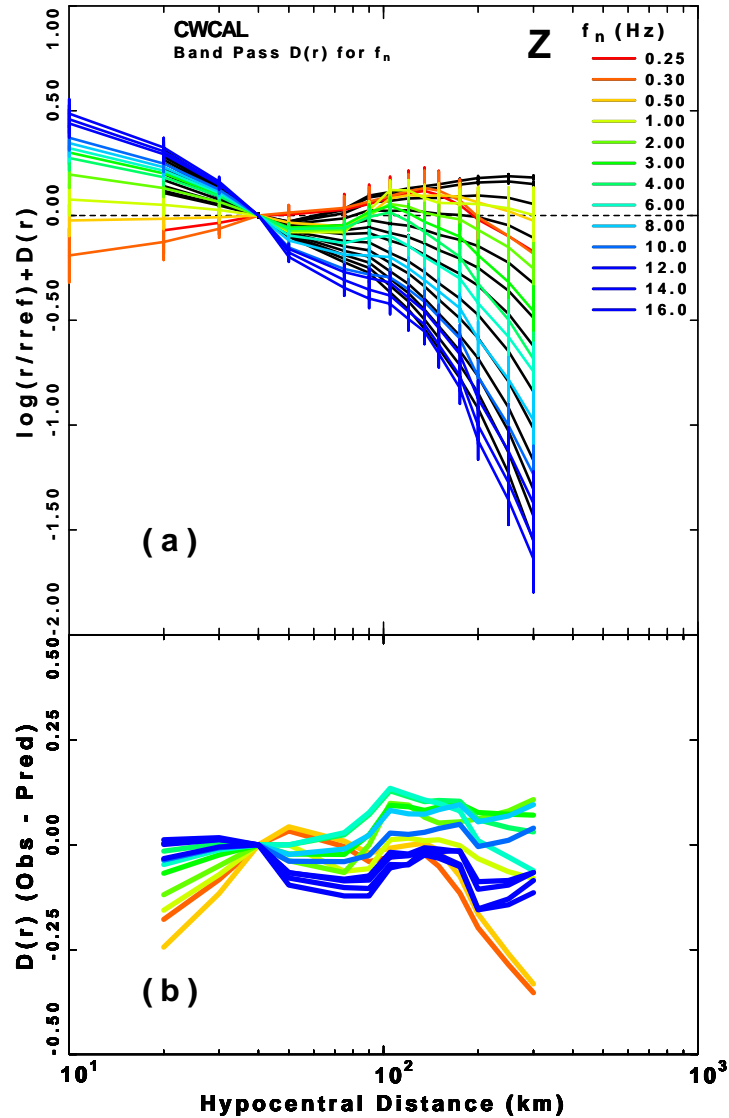

Figure 7.19: (a) Observed (colored) and modeled (black) vertical component bandpass filtered velocity attenuation function for 0.25-16 Hz in the central west California (b) The residuals of the model fit to the observed $D(r, f)$.

and η (250 and 0.30). Figure 7.20 shows the observed and modeled $D(r)$ for horizontal component data. The residuals are valid up to 300 km because the observations exist only up to 300 km as can be seen from the bottom panel of Fig. 7.17. The

observed $D(r)$ curves have been extrapolated for distances larger than 300 km in central west California and their modeling for these larger distances is not so realistic. The residuals for the distances up to 300 km are small for vertical, horizontal, and 3-component data. Using the propagation parameters I have modeled the excitation
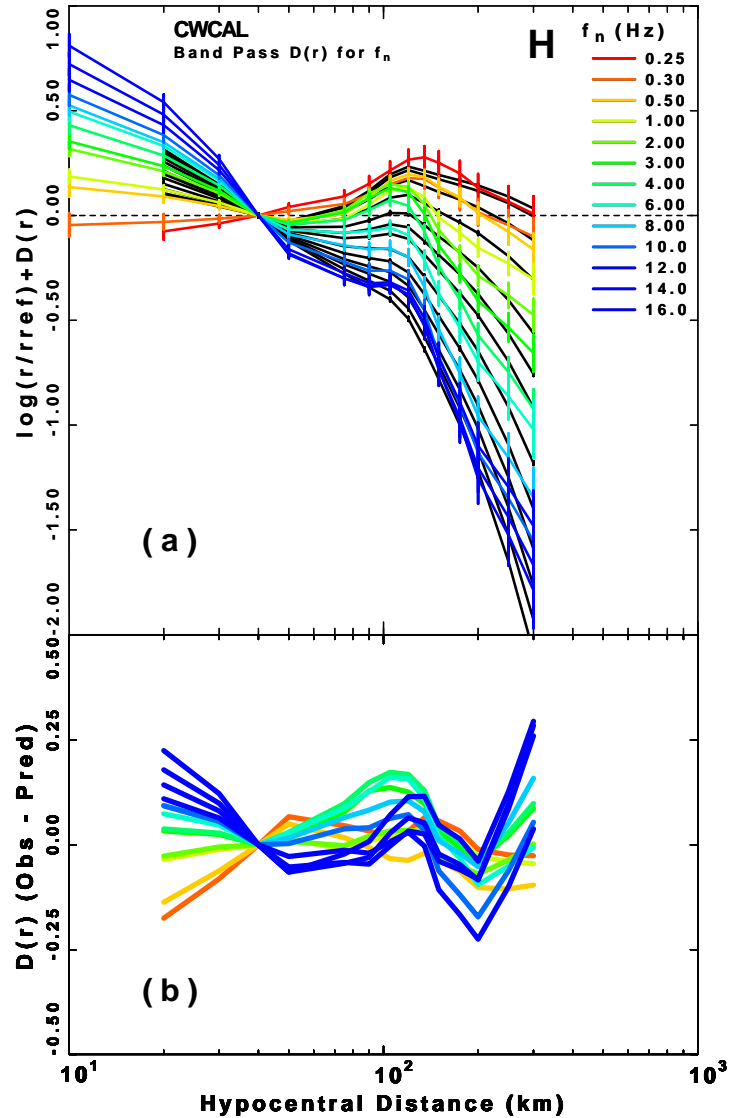
Figure 7.20: (a) Observed (color) and modeled (black) horizontal component bandpass filtered velocity attenuation function for 0.25-16 Hz in the central west California (b) The residuals of the model fit to the observed $D(r, f)$.

terms for 3-component bandpass filtered data. As usual I started with modeling a few events with known moment magnitudes using a constant stress drop and κ for all. For this step I used 15 events with known moment magnitudes from 3.5 to 6.5.

The residuals of the fit are shown in Figure 7.22 for some of the modeled excitations.
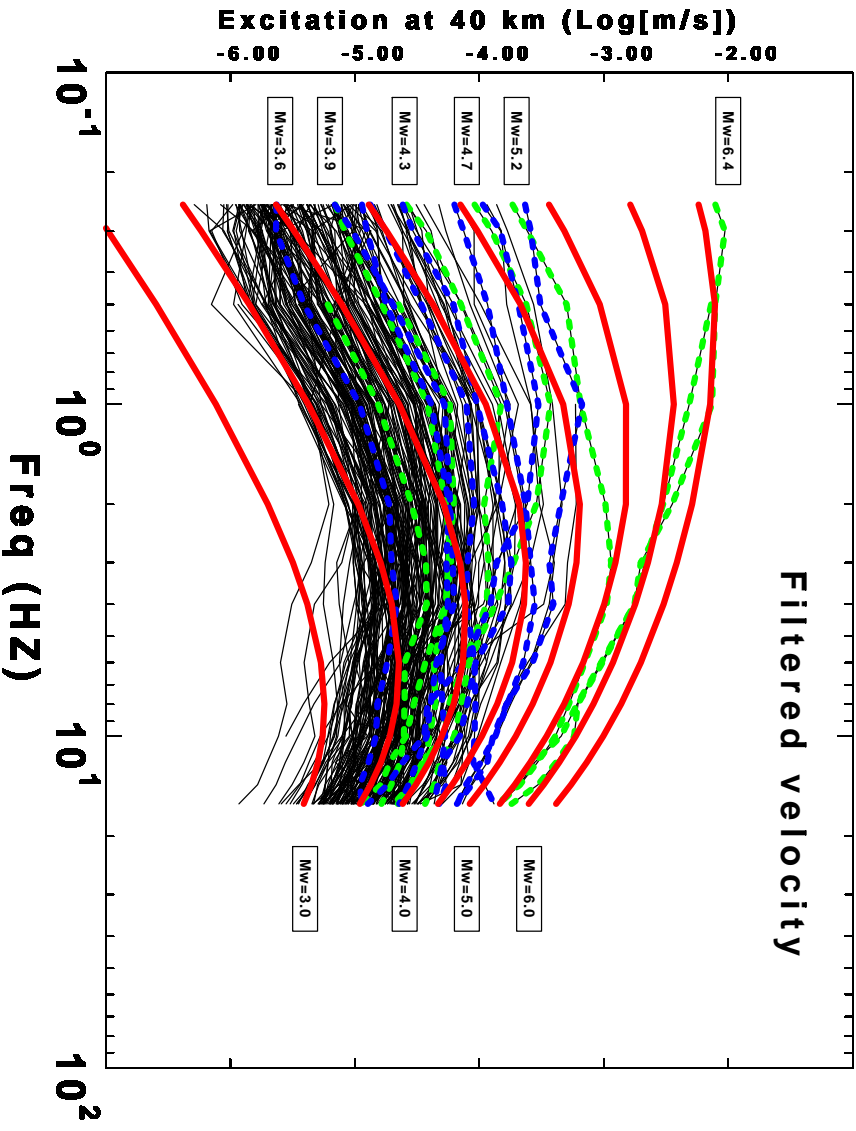
Figure 7.21: Excitation at 40 km for 3-component bandpass filtered data in the central west California. Black curves are observed excitations. Red curves are the theoretical curves for events with moment magnitudes 3.0 to 6.5. Blue and green dotted curves are the observed excitations of 15 selected events with known moment magnitudes from 3.5 to 6.5. The small boxes in the left show the moment magnitude of green dotted curves.

Most of the residuals are large at frequencies less than 1.0 Hz. For some events the higher frequencies also have large residuals. Using the previously determined Q_0 and η, the best model gives $\Delta\sigma = 110$ bars and $\kappa = 0.035$ sec. The high residuals imply that the model is not appropriate at least for the lower frequencies. The obtained stress drop is considerably larger than the average estimated value for the central east California (45-50 bars) in this study. I have done the same computations as the central east California to obtain different estimates of stress drop. First I modeled the excitation terms for 49 events with known moment magnitudes. Using a constant stress drop for all events the best model gives $\Delta\sigma = 124$ bars and $\kappa = 0.026$ sec. In
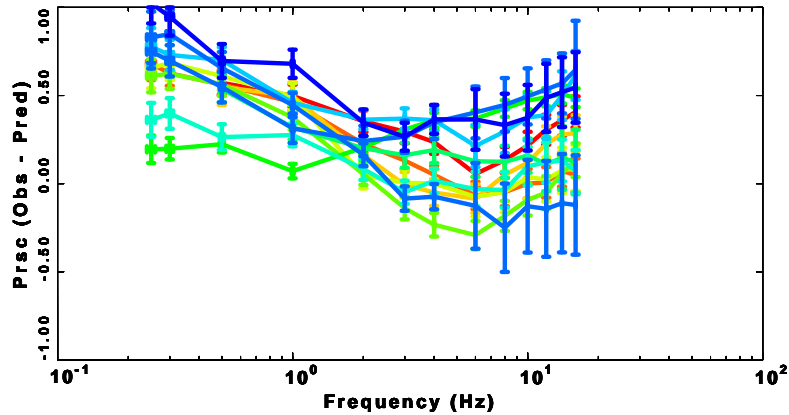
Figure 7.22: Residuals of the fit for the 3-component bandpass filtered excitation at 40 km for the 15 events of the Fig. 7.21.

the next step I have done the same modeling while letting each event has a different stress drop and κ. The results are shown in Figure 7.23 in which κ for each event is plotted versus the stress drop.
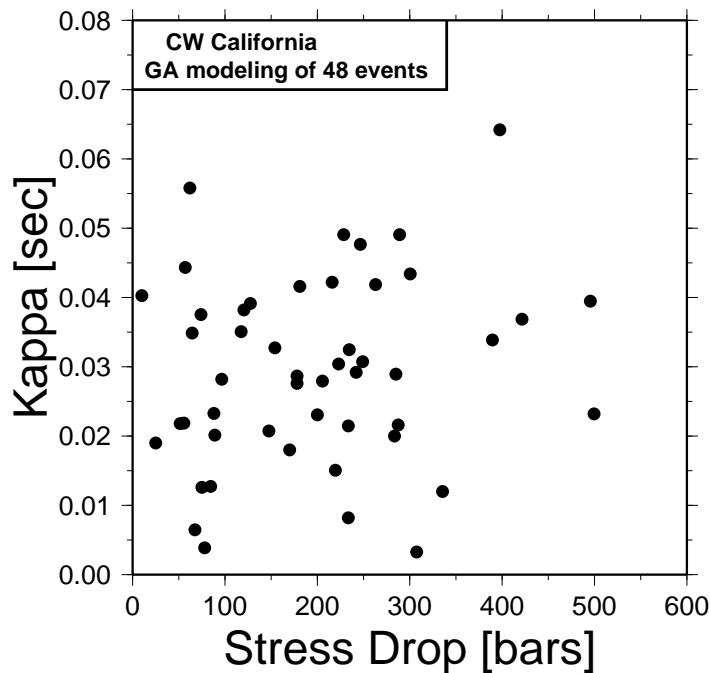
Figure 7.23: κ versus stress drop for 49 events with known moment magnitudes in the central west California.

There is no correlation between the two parameters in this plot. The stress drops range from 10 to 500 bars. The mean and standard deviation are 190 and 115 bars for stress drop and 0.029 and 0.013 sec for κ. The purpose of this part is to obtain

93

the mean κ and repeat the modeling while κ is fixed. I did the modeling with a fixed value of 0.03 sec for κ to obtain the stress drops of the same 49 events. Stress drops for two different cases of variable and fixed κ are plotted against moment magnitude in Figure 7.24. For the fixed κ case the standard deviation and mean of the stress
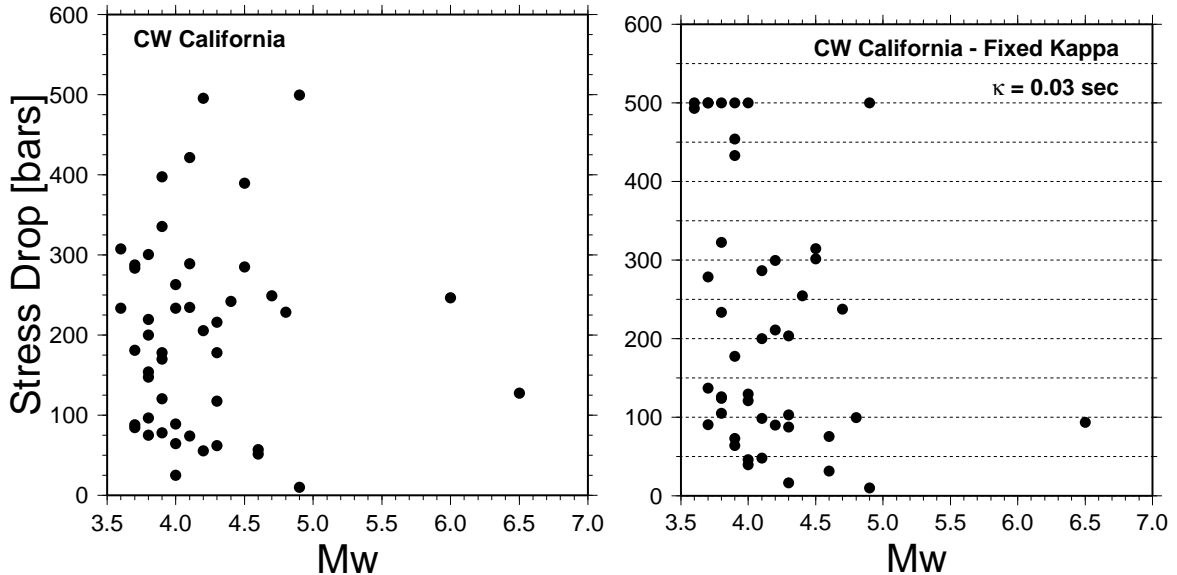
Figure 7.24: (left) Stress drop versus moment magnitude for 49 events in the central west California. Each event has a different κ (Fig. 7.23). (right) Same as left but a fixed value of 0.03 for κ was used in the modeling.

drop are 215 and 150 respectively. Compared to the results of the variable κ (left panel) the fixed κ model shows more scatter in the stress drop results. Using a fixed κ a few events have stress drops equal to or greater than 500 bars. This results in a larger mean and standard deviation. The final modeling was done to obtain the error of each estimated stress drop. The error and the mean were computed from 100 simulations for each single event. The means of 100 runs are not necessarily the same as the best model result for each event. The average stress drop and standard deviation for events from this computation are 202 and 111 bars respectively. The individual error estimates are shown in Figure 7.25 for 39 events. From the 49 events of Figure 7.24, 10 events had high residuals in modeling and therefore were excluded in error estimation of individual stress drops. In fact stress drops higher than 300 bars are high and may not be not realistic. If I exclude events with stress drops higher than 400 bars (which also have high residuals in modeling), the average and mean

94
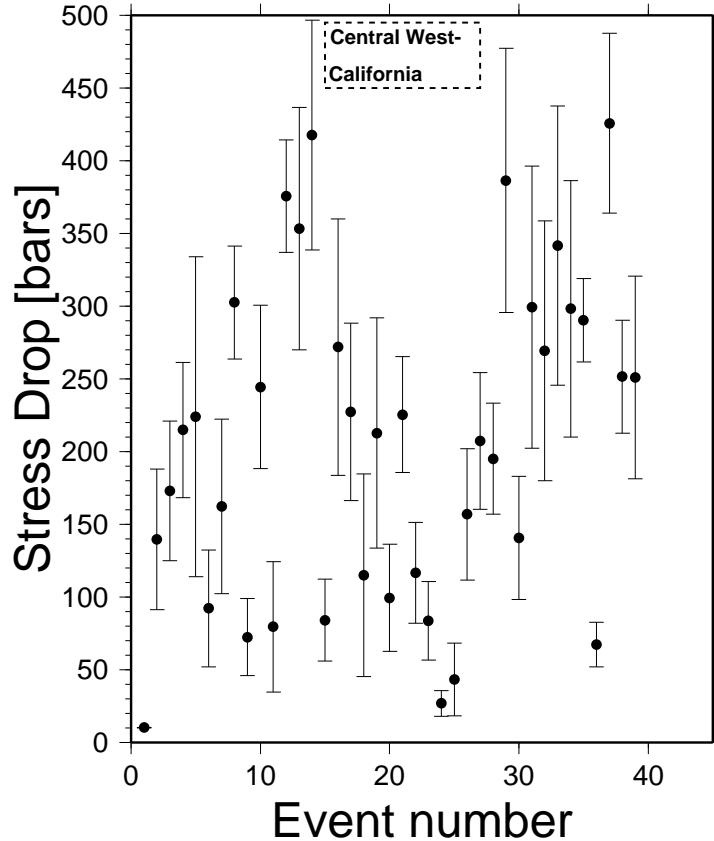
Figure 7.25: Estimated errors of stress drops for 39 events in central west California. Each individual event has been modeled using a Brune point source model. A bootstrap analysis has been done for each event. The stress drop values are the mean values of each bootstrap. The average of all stress drops is 202 bars.

stress drop for the rest of the events will be 142 and 86 bars. This result is closer to the result of the modeling all events with a constant stress drop and κ.

The results of this part are summarized in Table 7.5. The first row of the results show the estimated parameters obtained from modeling of 15 events using a constant stress drop (Fig. 7.21). Second row shows the same results obtained from modeling of 49 events using a constant stress drop model. The results of the third row were obtained by modeling each event separately in which every event have different stress drop and κ. The results of stress drop and κ for this case are shown as the mean and standard deviation of all computed values. Individual results are shown in Figure 7.24. The bootstrap results for this case are shown in Figure 7.25. The fourth row is the same as first row except that I used a fixed κ. The individual results are shown in

the left panel of Figure 7.24

Bootstrap and best model results for $\Delta\sigma$ and κ				
No. of events used	best model result		bootstrap	
	$\Delta\sigma$ (bars)	κ (sec)	$\Delta\sigma$ (mean and std)	κ (mean and std)
15 (const. $\Delta\sigma$)	110	0.035	160 ± 46	0.039 ± 0.01
49 (const. $\Delta\sigma$)	124	0.026	130 ± 34	0.036 ± 0.009
49 (non-const. $\Delta\sigma$)	Ave: 190 ± 115	Ave: 0.029 ± 0.013	-	-
49 (non-const. $\Delta\sigma$)	Ave: 215 ± 150	0.030 (fixed)	-	-

Table 7.5: Best model and error analysis results of stress drop and κ for 49 events with known moment magnitudes in the central west California. Brune's source model is used for all events.

The results of the different modelings show that the stress drop is higher than 100 bars. Based on the results the final selected average values are $\Delta\sigma = 125 \pm 35$ bars and $\kappa = 0.035 \pm 0.01$ sec.

7.3   Statistical tests on regression results

It is always important to do some statistical tests on the regression or inversion results. A least square technique has been used in the regression of observed ground motions throughout this study. Solving an inversion problem by least squares assumes a Gaussian distribution of errors. Outliers in the data can cause the error distributions to deviate from a Gaussian one. Figure 7.26 shows histogram plots of regression residuals for central west California data at 6 different frequencies. The histograms are very close to the normal in appearance and means are very close to zero. I have done the test of normality in two ways. In the first method I used Chi-Square goodness of fit test. This test is applied to binned data in which the data range is divided into k subintervals. The number of data points in each subinterval is counted and the empirical histogram is then compared with the theoretical PDF (Gaussian in here). The formula for the Chi-Square statistic is

$$\chi^2 = \sum_{i=1}^{k} \frac{(e_i - o_i)^2}{e_i} \sim \chi^2_{k-1-\text{nep}}$$

Figure 7.26: Histograms of 3-component regression residuals in central west California at 0.25, 0.3, 1.0, 2.0, 10.0, and 16.0 Hz.

where e_i is the expected number of data points in cell i, o_i is the actual number of data points in cell i, and $k - 1 - \text{nep}$ is the Chi-Square degrees of freedom where nep is the number of estimated parameters. This test is sensitive to the choice of bins. Usually $k > 5$ is selected but with a large number of data points different choices of k should be tested. I have used $k = 10$ in the computations. The number of estimated parameters are two since we estimate mean and standard deviation. So for the choice of k=10, we will have $k - 1 - \text{nep}$=7. For a test confidence level, α, we will find the value of $\chi^2_{1-\alpha,k-1-\text{nep}}$ from the table. For our choice of $k = 10$ and $\alpha = 0.05$, $\chi^2_{0.95,7} = 14.07$. If the result of the test from the above formula is smaller than the Chi-Square table value, we can assume that the distribution of our data set is normal. The Chi-Square test on the regression residuals showed than none of the residuals for different frequencies have normal distribution. I tried different number of subintervals since the Chi-Square test is very sensitive to this number. For some choices of k, the Chi-Square test results were close to the table value (14.07). The results of this test on the 3-component regression residuals in the central west California are shown in

Table 7.6. The mean and skewness for each frequency are also shown in the table.

Statistical test results on regression residuals				
Freq. (Hz)	No. Obs.	Mean	Skewness	χ^2 value
0.25	1031	8.0970E-04	0.1670	115
0.30	1179	-2.110E-05	-0.343	60
0.50	1662	1.3300E-04	-0.949	71
1.00	2104	6.8500E-04	-1.368	108
2.00	2236	-1.470E-05	-0.552	38
3.00	1374	-5.900E-06	-0.341	41
4.00	2404	-2.120E-05	-0.230	49
6.00	1926	-1.700E-05	-0.179	41
8.00	1932	-1.700E-05	-0.172	46
10.00	1931	-4.800E-06	-0.123	51
12.00	1819	-7.600E-06	-0.113	51
14.00	1809	-1.600E-05	-0.069	32
16.00	1780	-1.300E-05	0.1270	31

Table 7.6: Mean, skewness, and Chi-Square test results for the regression residuals of bandpass filtered data in central west California. The number of subintervals for the test is 10.

The χ^2 value (k=10), for all frequencies is larger than the $\chi^2_{0.95,7}$ value (14.07) from the table. It means that the residuals do not have a normal distribution for this choice of k. The number of observations (second column) are large for all frequencies. The usual choices of k may not be good for large number of samples. The histograms show that residual distributions may not be far from the normal. I repeated the test with randomly selecting 100 samples from the residuals. For 1000 tests of randomly selected samples the mean and standard deviation of χ^2 values were 13.5 and 7.8 for 0.25 Hz frequency. This variability of results for different values of k indicates that the normality test should not be judged by Chi-Square test alone and some other tests should be used as well.

In the second approach I used quantile-quantile (Q-Q) plots (*Jain*, 1991) which is a way to visually test if a data set has a specific distribution. This plot has been shown in Figure 7.27 for the regression residuals at frequencies 0.25, 0.30, 1.0, 2.0, 10.0, and 16.0 Hz. In this plot the actual residuals of the inversion are plotted against those expected for a Gaussian distribution. The quantile of the residuals (vertical axis) are just the ordered residuals. If the distribution of residuals is Gaussian, the Q-Q plot should look like a straight line. The Q-Q plots show that at some
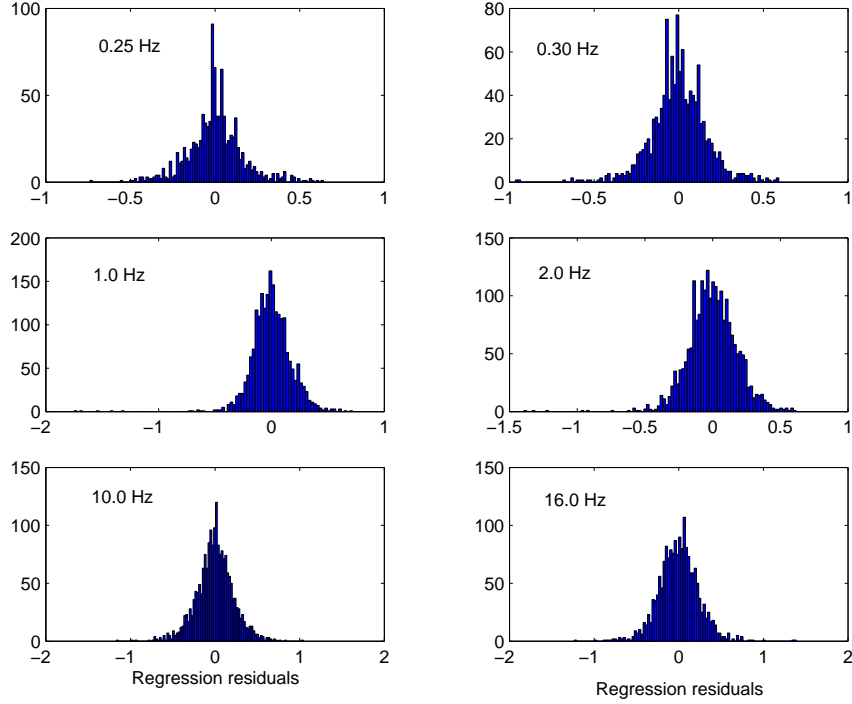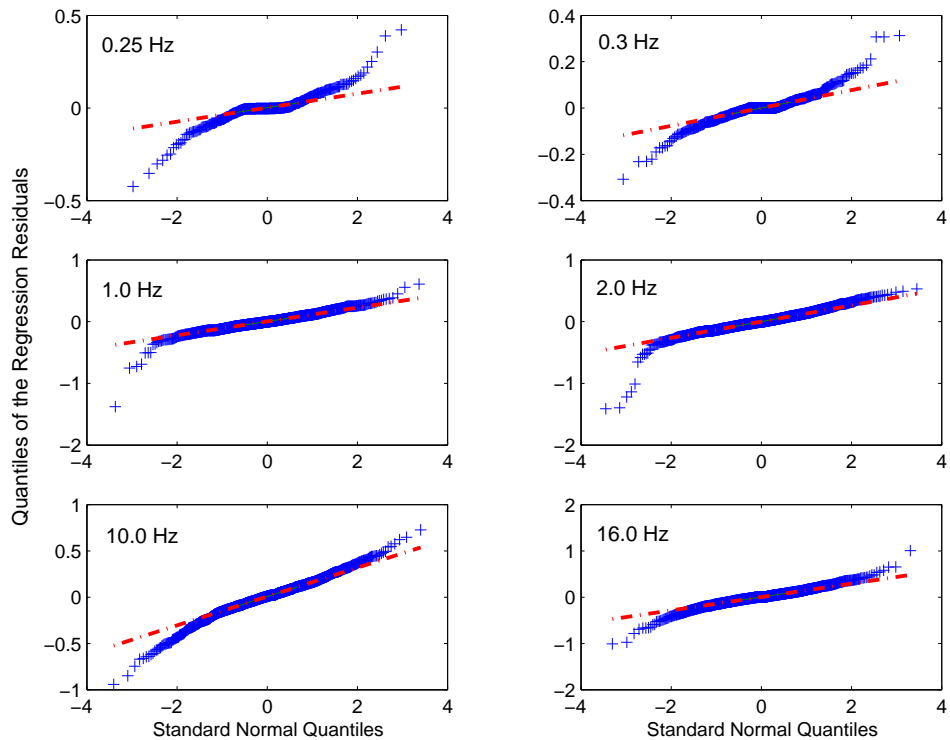
Figure 7.27: Q-Q plots for 3-component regression residuals in central west California at 0.25, 0.3, 1.0, 2.0, 10.0, and 16.0 Hz. The dashed red line represents the trend for a Gaussian distribution.

frequencies only a few data points do not lie on the straight line (Gaussian trend). So we can conclude that for those frequencies (larger than 1.0 Hz), the assumption of Gaussian distribution is valid. At the same time it is possible to look at the regression results, remove some outliers and repeat the regression. Throughout this study I have repeated the regression while I have excluded some of the outliers from the observations. The regression results also show which data (stations and events) produce unusual results. Those data could be eliminated in the processing when the procedure is repeated.

7.4   Ground motion scaling comparisons

Studying regional variation of wave propagation parameters in the Pacific Northwest and northern California was one of the objectives of this study. The results of

99

regressions and modeling show that in fact there are variations from one region to an-
other region. Figure 7.28 shows the regression derived $D(r)$ for the 4 different regions
of this study. Among the different regions, northern California shows a strong effect
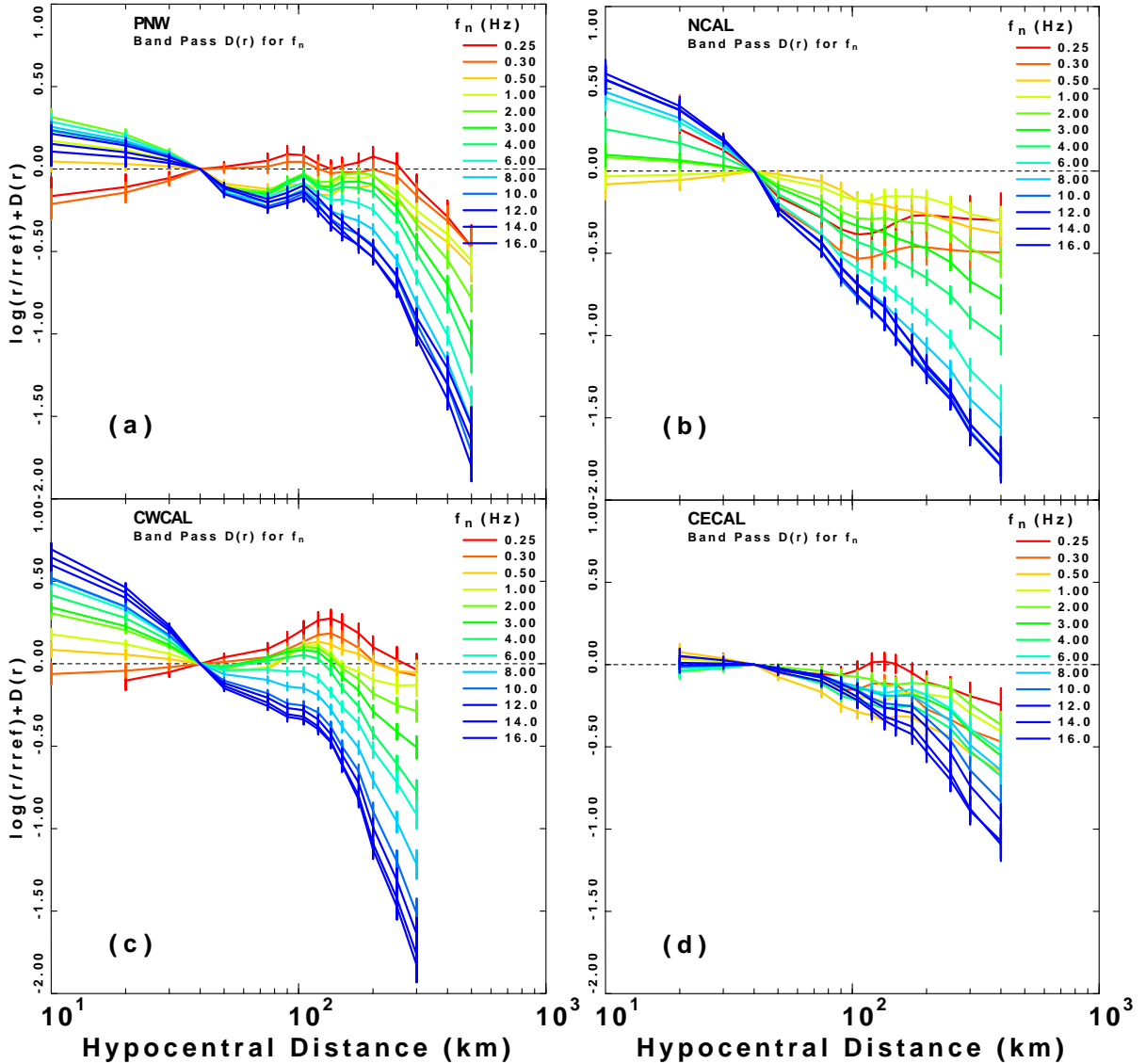
Figure 7.28: Regression derived $D(r)$ for 3-component bandpass filtered data in (a)
Pacific Northwest, (b) northern California, (c) central west California, and (d) central
east California.

of geometrical spreading. In fact modeling the $D(r)$ term for northern California
requires using a frequency dependent geometrical spreading and η. The central east
California shows a smoother form of $D(r)$ compared to other regions and a simpler
form of geometrical spreading can be used to model the observations in this region.

The modeling results for $D(r)$ are shown in Table 7.7. The mean values of Q_0 varies from 210 to 280 between the different regions. The values for η change from 0.30 in the central west California to 0.55 in the Pacific Northwest. The range of both η and Q_0 values indicate that the variation of the propagation parameters is important.

Wave propagation parameters		
Pacific Northwest	**C. East California**	**C. West California**
$Q(f) = 280^{+34}_{-34}f^{0.55\pm0.05}$	$Q(f) = 280^{+36}_{-36}f^{0.50\pm0.09}$	$Q(f) = 250^{+19}_{-19}f^{0.30\pm0.05}$
$g(r) = \begin{cases} r^{-1.0} & r < 40\,\text{km} \\ r^{-1.2} & 40 < r < 70\,\text{km} \\ r^{0.5} & 70 < r < 100\,\text{km} \\ r^{-0.7} & 100 < r < 200\,\text{km} \\ r^{-0.9} & r > 200\,\text{km} \end{cases}$	$\begin{matrix} r^{-1.0} & r < 40\,\text{km} \\ r^{-0.8} & 40 < r < 100\,\text{km} \\ r^{-0.5} & r > 100\,\text{km} \end{matrix}$	$\begin{matrix} r^{-1.3} & r < 40\,\text{km} \\ r^{-1.0} & 40 < r < 60\,\text{km} \\ r^{0.1} & 60 < r < 90\,\text{km} \\ r^{-0.5} & 90 < r < 150\,\text{km} \\ r^{-0.90} & r > 150\,\text{km} \end{matrix}$
Northern California		
$f = 1 - 13$ Hz	$f = 1 - 4$ Hz	$f = 6 - 13$ Hz
$Q(f) = 210^{+40}_{-40}f^{0.50\pm0.1}$	$\begin{matrix} r^{-1.0} & r < 40\,\text{km} \\ r^{-0.8} & 40 < r < 70\,\text{km} \\ r^{-0.6} & 70 < r < 100\,\text{km} \\ r^{-0.5} & r > 100\,\text{km} \end{matrix}$	$\begin{matrix} r^{-1.3} & r < 40\text{km} \\ r^{-1.5} & 40 < r < 70\,\text{km} \\ r^{-1.2} & 70 < r < 100\,\text{km} \\ r^{-0.8} & r > 100\,\text{km} \end{matrix}$

Table 7.7: Summary of wave propagation parameters for 4 different regions of this study.

The source scaling parameters for all the regions are shown in Table 7.8. The average values of stress drop change from 30 bars in the Pacific Northwest to 125 bars in the central California. The obtained values for the central east California are much smaller than those of central west. It should be emphasized that for each region, the source scaling results are tied to the corresponding propagation model of the region (table 7.7).

Source scaling parameters				
Parameter	**Pacific Northwest**	**N. California**	**CE. California**	**CW. California**
$\Delta\sigma$ [bars]	30 ± 9	90 ± 30	45 ± 15	125 ± 35
κ [sec]	0.05 ± 0.009	0.04 ± 0.01	0.036 ± 0.008	0.035 ± 0.01

Table 7.8: Summary of source scaling parameters for 4 different regions of this study.

Chapter 8

Discussion and Conclusions

I have studied regional ground motion scaling in the Pacific Northwest, northern, and central California. The data set for each region consisted of vertical short-period and 3-component broadband seismograms. I obtained the results for the frequency range of 0.25-16 Hz and distances 20-500 km. In this study I have processed the data in the frequency and time domain. For the time domain the data are bandpass filtered within narrow frequency bands to obtain the peak motions.

I used coda normalization and the regression methods to obtain excitation, propagation, and site terms from the amplitudes of the observed strong motions. The derived propagation and excitation terms are modeled to obtain the attenuation and source parameters. The coda normalization results are used for comparison with the regression results. The two methods agreed well when there were sufficient number of observations for both methods but in general the regression technique provided better results.

The $D(r)$ term shows a smoother form in the central east California compared to other regions. For northern California, $D(r)$ shows a rapid decrease of amplitude with distance especially for frequencies higher than 4 Hz. The $D(r)$ term for the Pacific Northwest and central west California show a typical trend of attenuation with some effects of supercritical reflections from the Moho.

I modeled the observations of $D(r)$ using visual and genetic algorithm (GA) methods. The visual method is faster and can be used if the number of parameters are limited. It is also good to obtain a rough estimate of parameters using only a few runs. For the larger number of parameters the visual method becomes harder to use unless we keep some of the parameters constant. GA is good for these cases and searches the parameter space in a better way.

By modeling the $D(r, f)$ I found $Q(f) = 280\ f^{0.55}$ for the Pacific Northwest. The residuals of the fit are small for this model. The geometrical spreading for this region consists of a 5 segment function. Strong supercritical reflections requires $r^{0.5}$ parameterization in the distance range of 70-100 km for the best fit. The best model for the central east California gives $Q(f) = 280\ f^{0.50}$ which is very close to $Q(f)$ for the Pacific Northwest. The geometrical spreading is much simpler in this region. The geometrical spreading exponents are -1.0, -0.8, and -0.5 for distances less than 40 km, between 40 and 100 km, and larger than 100 km respectively. The central west California best model has $Q(f) = 250\ f^{0.30}$ and a 5 segment geometrical spreading function.

I could not model the $D(r, f)$ for the northern California as I did for the other regions. The specific shape of this term for northern California required using a frequency dependent η and frequency dependent geometrical spreading. For the frequency range of 1-4 Hz, $Q(f) = 210\ f^{0.55}$ and for 6-13 Hz, $Q(f) = 210\ f^{0.55}$. The change in η from one frequency range to another is not that large but the differences in geometrical spreading are profound. The 6-13 Hz geometrical spreading are much stronger (Table 7.7).

The Q values obtained in this study are in the range of the values obtained for western north America by some previous studies (*Baqer and Mitchell*, 1998). The Q_0 values of this study however, are a little different from some other studies of wave attenuation (*Raoof et al.*, 1999; *Boatwright et al.*, 2003). These differences are expected since different frequency ranges and different methods of Q measurement will give different results. Specific regions analyzed may also be a reason for these differences (*Shi et al.*, 1996). *Boatwright et al.* (2003) studied the data from northern and some parts of central California. I have done separate studies for northern and central California.

Since in most studies Q_0 and η are determined together, we have to consider the trade-off between these two parameters. This trade-off precludes obtaining a unique solution of either of parameters. So in comparing different studies, the combined effect of Q_0 and η should be considered although the effect of η for higher frequencies is

stronger. Among the four regions studied here, the rapid amplitude decrease of higher frequencies in northern California is noticeable. The results for three other regions doesn't show such a strong effect of geometrical spreading. The Q_0 value is also smaller (210) in northern California. *Boatwright et al.* (2003) analyzed peak ground acceleration and velocity data from northern California and some parts of central California and observed strong geometrical spreading effect in distances larger than 100 km.

The trade-off between $g(r)$ and $Q(f)$ is also important and should be considered. Some studies assume $g(r) = r^{-1.0}$ for the all distances and solve for $Q(f)$. For some applications (seismic hazard) the combined effect of $g(r)$ and Q is important and the trade-off does not pose a problem. Many empirical attenuation relationships have been developed to use in seismic hazard studies. Those relationships express the ground motion amplitude as a function of magnitude and distance. The decay form of amplitude is the combined effect of the Q and $g(r)$.

I have modeled the excitation terms using the previously determined $D(r)$. I have used a Brune point source model with RVT to obtain the stress drop parameter and κ. I modeled the excitation terms in two ways: using a few events with a constant stress drop and using all events with known moment magnitudes while each event has a different stress drop parameter. I have also repeated the modeling with a fixed value of κ for the second method. But in each case I have selected an average value for the parameters. The values of κ vary from 0.03 to 0.05 between different regions. An average value of 30 bars is obtained for the Pacific Northwest. The model fit is generally good and residuals are not large. The northern California average value is 90 bars. The best estimates for central east and central west California are 45 and 125 bars respectively. For both regions of central California the residuals are high for the frequencies less than 1.0 Hz. Using the three different stress drops, the average value for the whole study area of California is 70 bars. The individual stress drops for central east and central west California show much scatter ranging from 10 to a few hundred bars. In making the averages, I excluded a few events in the central west California with very high stress drops and high residuals of modeling. Because

of large spread in stress drops of central west California, using an average value may not be justified and a constant stress drop value should not be used.

The uncertainties in the results are also important. The errors obtained for the stress drops are high for some of the events. The errors on constant stress drops are smaller. Since estimating the stress drops is not easy in general and they also depend on the predefined $Q(f)$ values, the modeling results for $Q(f)$ are more robust.

The values of Q_0 obtained in this study ranges from 210 to 280. The average stress drops change from 30 bars in the Pacific Northwest to 125 in central west California. These ranges of parameters indicate that the ground motion parameters of a specific region may not be used for another region. For example using the stress drops of central California may overestimate the ground motion level predictions in the Pacific Northwest.

References

Aki, K. (1967), Scaling law of seismic spectrum, *J. Geophys. Res.*, *72*, 1217–1231.

Aki, K. (1969), Analysis of the seismic coda of local earthquakes as scattered waves, *J. Geophys. Res.*, *74*, 615–631.

Aki, K. (1980), Attenuation of shear waves in the lithosphere, *J. Geophys. Res.*, *85*, 6496–6504.

Aki, K., and B. Chouet (1975), Origin of coda waves: Source, attenuation, and scattering effects, *J. Geophys. Res.*, *80*, 3322–3342.

Anderson, J. G., and Y. Lei (1994), Nonparametric description of peak acceleration as a function of magnitude, distance, and site in Guerro, Mexico, *Bull. Seism. Soc. Am.*, *84*, 1003–1017.

Andrews, D. J. (1982), Separation of source and propagation spectra of seven Malmoth Lakes aftershocks, *Tech. Rep. 82-591*, U.S. Geol. Surv. Open File Rep.

Andrews, D. J. (1986), Objective determination of source parameters and similarity of earthquakes of different size, in *Earthquake Source Mechanics*, edited by S. Das, J. Boatwright, and C. Scholz, pp. 259–267, American Geophysical Union.

Atkinson, G. M. (1993), source spectra for earthquakes in eastern north america, *Bull. Seism. Soc. Am.*, *83*, 1778–1798.

Atkinson, G. M. (1995), Attenuation of source parameters of earthquakes in the Cascadia region, *Bull. Seism. Soc. Am.*, *85*, 1327–1342.

Atkinson, G. M., and D. M. Boore (1997), Stochastic point-source modeling of ground motions in the Cascadia region, *Seismol. Res. Lett.*, *68*, 74–85.

Atkinson, G. M., and D. M. Boore (2003), Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions, *Bull. Seism. Soc. Am.*, *93*, 1703–1729.

Atkinson, G. M., and W. Silva (1997), An empirical study of earthquake source spectra for California earthquakes, *Bull. Seism. Soc. Am.*, *87*, 97–113.

Atkinson, G. M., and W. Silva (2000), Stochastic modeling of California ground motions, *Bull. Seism. Soc. Am.*, *90*, 255–274.

Baqer, S., and B. J. Mitchell (1998), Regional variation of L_g coda Q in continental United States and its relation to crustal structure evolution, *Pure Appl. Geophys.*, *153*, 613–638.

Benz, H., A. Frankel, and D. M. Boore (1997), Regional L_g attenuation for the continental United States, *Bull. Seism. Soc. Am.*, *87*, 606–619.

Boatwright, J., and G. Choy (1992), Acceleration source spectra anticipated for large earthquakes in eastern north america, *Bull. Seism. Soc. Am.*, *82*, 660–682.

Boatwright, J., J. B. Fletcher, and T. E. Fumal (1991), A general inversion scheme for source, site, and propagation characteristics using multiply recorded sets of moderate-sized earthquakes, *Bull. Seism. Soc. Am.*, *81*, 1754–1782.

Boatwright, J., H. Bundock, J. Luetgert, L. Seekins, L. Gee, and P. Lombard (2003), The dependence of pga and pgv on distance and magnitude inferred from northern California shake map data, *Bull. Seism. Soc. Am.*, *93*, 2043–2055.

Boore, D. M. (1983), Stochastic simulation of high frequency ground motion based on seismological models of radiated spectra, *Bull. Seism. Soc. Am.*, *73*, 1865–1894.

Boore, D. M. (2003), Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.*, *160*, 636–676.

Boore, D. M., W. B. Joyner, and T. E. Fumal (1992a), Estimation of response spectra and peak acceleration from western North America earthquakes: An interim report, *Tech. Rep. 93-509*, U.S. Geol. Surv. Open File Rep.

Boore, D. M., W. B. Joyner, and L. Wennerberg (1992b), Fitting the stochastic ω^{-2} source model to observed response spectra in western north america: trade-offs between $\Delta\sigma$ and κ, *Bull. Seism. Soc. Am.*, *82*, 1956–1963.

Borcherdt, R. D. (1970), Effects of local geology on ground motion near San Fransisco Bay, *Bull. Seism. Soc. Am.*, *60*, 29–61.

Borcherdt, R. D., G. Galssmoyer, A. D. Kiureghian, and E. Granswick (1989), Results and data from seismologic and geologic studies following earthquakes of December 7, 1988 near Spitak, Armenia, S.S.R., *Tech. Rep. 89-163A*, U.S. Geol. Surv. Open File Rep.

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

Campbell, W. B. (1981), Near source attenuation of peak horizontal acceleration, *Bull. Seism. Soc. Am.*, *71*, 2039–2070.

Campbell, W. B. (1991), An empirical analysis of peak horizontal acceleration for the Loma Prieta, California, earthquake of 18 October 1989, *Bull. Seism. Soc. Am.*, *71*, 1838–1858.

Campbell, W. B., and Y. Bozorgnia (1994), Near source attenuation of peak horizontal attenuation from worldwide accelerograms recorded from 1957 t0 1993, in *Fifth U.S. National Conference on Earthquake Engineering*, vol. 1, pp. 283–292, Earthquake Engineering Institute, Berkeley, California.

Cartwright, D. E., and M. S. Longuet-Higgins (1956), The statistical distribution of the maxima of a random function, *Proc. Roy. Soc. London, Ser*, pp. 212–223.

Chin, B. H., and K. Aki (1991), Simultaneous study of the source, path, and site effects on strong ground motion during the 1989 Loma Prieta earthquake: a preliminary result on pervasive non-linear site effects, *Bull. Seism. Soc. Am.*, *81*, 1859–1854.

Choy, G. L., and J. Boatwright (1990), Source characteristics of the Loma Prieta California, earthquake of october 18, 1989 from global digital seismic data, *Geophys. Res. Lett.*, *17*, 1183–1186.

Choy, G. L., and S. H. Kirby (1994), Apparent stress, fault maturity, and seismic hazard for normal-fault earthquakes at subduction zones, *Geophys. J. Int.*, *159*, 991–1012.

Cornell, C. A. (1968), Engineering seismic risk analysis, *Bull. Seism. Soc. Am.*, *58*, 1583–1606.

Dewberry, S. R., and R. S. Crosson (1995), Source scaling and moment estimation for the Pacific Northwest seismograph network using s-Coda amplitudes, *Bull. Seism. Soc. Am.*, *85*, 1309–1326.

Efron, B., and R. J. Tibshirani (1994), *An Introduction to the Bootstrap*, Chapman and Hall.

Field, E. H., and K. H. Jacob (1995), A comparison and test of various site-response estimation techniques including three that are not reference-site dependent, *Bull. Seism. Soc. Am.*, *85*, 1127–1143.

Frankel, A., A. McGarr, J. Bicknell, J. Mori, L. Seeber, and E. Cranswick (1990), Attenuation of high-frequency shear waves in the crust: Measurements from New York state, South Africa and Southern California, *J. Geophys. Res.*, *95*, 17,441–17,454.

Gregor, N. J., W. J. Silva, I. G. Wong, and R. R. Young (2002), Ground-motion relationships for Cascadia subduction zone megathrust earthquakes based on a stochastic finite-fault model, *Bull. Seism. Soc. Am.*, *92*, 1923–1932.

Haddon, R. A. W. (1996), Earthquake source spectra in eastern North America, *Bull. Seism. Soc. Am.*, *86*, 1300–1313.

Hanks, T. C. (1979), b values and $\omega^{-\gamma}$ seismic source models: Implications for tectonic stress variations along active crustal fault zones and the estimation of high-frequency strong ground motion, *J. Geophys. Res.*, *84*, 2235–2242.

Hanks, T. C., and R. K. McGuire (1981), The character of high frequency ground motion, *Bull. Seism. Soc. Am.*, *71*, 2071–2095.

Hanks, T. C., and W. R. Thatcher (1972), A graphical representation of seismic source parameters, *J. Geophys. Res.*, *77*, 4393–4405.

Harmsen, S. C. (1997), Estimating the diminution of shear-wave amplitude with distance: Application to the Los Angeles, California, urban area, *Bull. Seism. Soc. Am.*, *87*, 888–903.

Haskell, N. A. (1964), Total energy and energy spectral density of elastic wave radiation from propagating faults, *Bull. Seism. Soc. Am.*, *54*, 1811–1841.

Herrmann, R. B. (1985), An extension of random vibration theory estimates of strong ground motion, *Bull. Seism. Soc. Am.*, *75*, 1447–1453.

Herrmann, R. B., and L. Malagnini (2005), Interpretation of high frequency ground motion from regional seismic network observations, to be published in *Bull.Seiem.Soc.Am.*.

Hough, S. E., R. D. Brocherdt, P. A. Friberg, R. Busby, E. Field, and K. H. Jacob (1990), The role of sediment-induced amplification in the collapse of the Nimitz freeway during the October 17, 1989 Loma Prieta earthquake, *Nature*, *344*, 853–855.

Jain, R. (1991), *The Art of Computer System Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*, Wiley.

Jeon, Y. S. (2003), High frequency earthquake ground motion scaling in Utah, Master's thesis, Saint Louis University, Saint Louis, Missouri.

Joyner, W. B., and D. Boore (1998), Measurement, characterization, and prediction of strong ground motion, in *Earthquake Engineering and Soil Dynamics II - Recent Advances in Ground-Motion Evaluation, Geotechnical Special Publication*, vol. 20, pp. 43–102, ASCE.

Kanamori, H., and D. L. Anderson (1975), Theoretical basis of some empirical relations in seismology, *Bull. Seism. Soc. Am.*, *65*, 1073–1095.

Kato, K., K. Aki, and M. Takemura (1995), Site amplification from coda waves: Validation and application to *s*-wave site response, *Bull. Seism. Soc. Am.*, *85*, 467–477.

King, J. L., and B. E. Tucker (1984), Observed variations of earthquake motion across a sediment filled vally, *Bull. Seism. Soc. Am.*, *74*, 137–151.

Kramer, S. (1996), *Geotechnical Earthquake Engineering*, Prentice Hall, Upper Saddle River, NJ.

Langston, C. A. (1978), Moments, corner frequencies, and the free surface, *J. Geophys. Res.*, *83*, 3422–3426.

Langston, C. A. (1979), Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *jgr*, *84*, 4749–4762.

Lawson, C. L., and R. J. Hanson (1974), *Solving least square problems*, Prentice-Hall Inc., Englewood Cliffs, NJ.

Lay, T., and T. C. Wallace (1995), *Modern Global Seismology*, Academic Press.

Lermo, J., S. Fransisco, and J.Chavez-Garcia (1993), Site effect evaluation using spectral ratios with only one station, *Bull. Seism. Soc. Am.*, *83*, 1574–1594.

Malagnini, L., R. B. herrmann, and M. D. Bona (2000a), Ground motion scaling in Apennines (Italy), *Bull. Seism. Soc. Am.*, *90*, 1062–1081.

Malagnini, L., R. B. herrmann, and K. Koch (2000b), Ground motion scaling in Central Europe, *Bull. Seism. Soc. Am.*, *90*, 1052–1061.

Malagnini, L., K. Mayeda, A. Akinchi, and P. L. Bragato (2004), Estimating absolute site effects, *Bull. Seism. Soc. Am.*, *94*, 1343–1352.

Mayeda, K., S. Koyanagi, and K. Aki (1991), Site amplification from S-wave coda in the Long Vally caldera region, California, *Bull. Seism. Soc. Am.*, *81*, 2194–2213.

McGuire, R. K., and T. C. Hanks (1980), Rms accelerations and spectral amplitudes of strong ground motion during the San Fernando, California, earthquake, *Bull. Seism. Soc. Am.*, *70*, 1907–1919.

Menke, W. (1989), *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press.

Moya, A., J. Aguirre, and K. Irikura (2000), Inversion of source parameters and site effects from strong motion records using genetic algorithms, *Bull. Seism. Soc. Am.*, *90*, 977–992.

Ortega, R. (2000), High frequency ground motion in Central Mexico: Site, excitation and attenuation, Ph.D. thesis, Saint Louis University, Saint Louis, Missouri.

Phillips, W. S., and K. Aki (1986), Site amplification of coda waves from local earthquakes in central California, *Bull. Seism. Soc. Am.*, *76*, 627–648.

Raoof, M., R. B. Herrmann, and L. Mlagnini (1999), Attenuation and excitation of three-component ground-motion in southern California, *Bull. Seism. Soc. Am.*, *90*, 888–902.

Rautian, T. G., and V. I. Khalturin (1978), The use of the coda for determination of the earthquake source spectrum, *Bull. Seism. Soc. Am.*, *68*, 923–948.

Reiter, L. (1991), *Earthquake Hazard Analysis: Issues and Insights*, Columbia University Press.

Sato, H., and M. Fehler (2000), *Seismic wave propagation and scattering in the heterogeneous earth*, Springer Verlag, New York.

Shi, J., W. Y. Kim, and P. G. Richards (1996), Variability of crustal attenuation in the northern united States from L_g waves, *J. Geophys. Res.*, *101*, 25,231–25,242.

Silva, W., and C. Stark (1992), Source, path, and site ground motion model for the 1989 M 6.9 Loma Prieta earthquake, *Tech. rep.*, CDMG.

Silva, W., R. Darragh, and I. Wong (1990), Engineering characterization of earthquake strong ground motions with applications to the Pacific Northwest, in *Proceedings at the Third NEHRP Workshop on Earthquake Hazards in the Pouget Sound /Portland Region*, edited by W. Hays, U.S. Geol. Surv. Open File Rep.

Singh, S. K., J. lemro, T. Dominguez, J. M. Espinosa, E. Mena, and R. Quass (1988), The Mexico earthquake of September 19. 1985- a study of amplification of seismic waves in the valley of Mexico with respect to a hill zone site, *Earthquake Spectra*, *4*, 653–673.

Stein, S., and M. Wysession (2003), *An Introduction to Seismology, Earthquakes, and Earth Structure*, Blackwell.

Stiedl, J. H., A. G. Tumarkin, and R. L. Archuleta (1996), What is a reference site, *Bull. Seism. Soc. Am.*, *86*, 1733–1748.

Tichelar, B. W., and L. J. Ruff (1989), How good are our best models? Jacknifing, bootstraping, and earthquake depth, *Eos Trans. AGU, 70*.

Toro, G. R., N. A. Abrahamson, and J. F. Schneider (1995), Engineering of strong ground motions from earthquakes in the central and the western United States, *Earthquake Spectra, 11*.

Tsujiura, M. (1978), Spectral analysis of the coda waves from local earthquakes, *Bull. Earthquake Rse. Inst. Univ. Tokyo, 53*, 1–48.

Udias, A. (1999), *Principles of Seismology*, Campbridge University Press.

Vanmarcke, E. H., and S. P. Lai (1980), Strong-motion duration and rms amplitude of earthquakes records, *Bull. Seism. Soc. Am.*, *70*, 1293–1307.

Wang, C. Y., and R. B. Herrmann (1980), A numerical study of $P-$, $SV-$, and $SH-$wave generation in a plane layered medium, *Bull. Seism. Soc. Am.*, *70*, 1015–1036.

Vita Auctoris

Ali Fatehi was born in Baneh, a small town northwest of Iran, in September 1965. He finished high school in his home town and then attended Tehran University in 1985 where he did his undergraduate studies in physics and graduated in February 1991. After two years of military service he was admitted to the graduate studies program of Institute of Geophysics, Tehran University in 1995 and obtained his master degree in October 1997.

He joined the International Institute of Earthquake Engineering and Seismology (IIEES) in February 1998. With his colleagues at IIEES he was working on developing seismic hazard maps and at the same time he was involved in building the Iranian National Seismic Network.

During some communications with Charles Ammon about seismological problems, he was encouraged to continue his graduate studies at the Department of Earth and Atmospheric Sciences of Saint Louis University, where he started his PhD study in August 2002.