Ground Motion Prediction Equations For Time-dependent Seismic Hazard Analysis in Geothermal Areas

Nitin Sharma

April 2, 2013

Ground Motion Prediction Equations For Time-dependent Seismic Hazard Analysis in Geothermal Areas

Doctorate Program In Seismic Risk
XXV Cycle

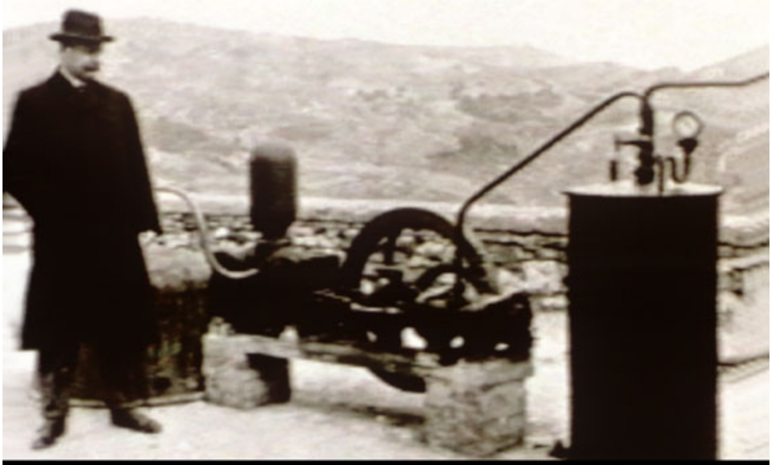
Department of Physics,
University of Naples, Federico II

Tutor: Prof. Aldo Zollo

Co-tutor: Dr. Vincenzo Convertito

Candidate: Dr. Nitin Sharma

Coordinator of XXV cycle in seismic risk program
Prof. Aldo Zollo



First Geothermal Power Plant, 1904, Larderello, Italy

Prince Piero Ginori Conti invented the first geothermal power plant in 1904,at the Larderello dry steam field in Italy.

Slide 50 of 122, © 2000 Geothermal Education Office

Preface and Acknowledgments

This thesis grew out of three years of my work as doctorate student in “Seismic Risk”. The challenges I faced in foreign land during this journey, weather they are scientific or non-scientific, cultural or non-cultural, technical or non-technical, has shaped my character as a growing philosopher in the domain of seismology. My work has dealt with study of seismic hazard analysis in geothermal areas especially in the “Geysers” geothermal area, California. This thesis will be of interest to the beginners and other learners who want to gain an understanding in estimating “Ground Motion Prediction Equations For Time-dependent Seismic Hazard Analysis In Geothermal Areas”.

In some sections, such as in Introduction there is much of the information on “The Geysers”. Similarly, there is an expanded discussion on inversion techniques, estimation of ground motion prediction equation and non-linear mixed effect regression technique. The detailed discussion is also presented on application ground-motion prediction equations in seismic hazard and its time dependent studies. In last chapter the concept of time-dependent ground motion prediction equations is also discussed. I expect that this work will provide basic information to non-technical audiences also. Different readers may get benefit from different sections.

A seismological and statistical background will be helpful to understand and extend the present work. Many colleagues and friends read

all or part of manuscripts at different stages and provided extremely useful advice and comments. In particular, I would like to thank my supervisor Prof. Aldo Zollo whose guidance has result in fruitful thesis. I would like to thank my tutor Dr. Vincenzo Convertito for his guidance, innovative ideas, discussions and encouragement. Last but not the least I would like to thank my family and friends for providing me immense support during this journey.

Dedication

To my family and friends.

Contents

List of Figures

List of Tables

1. Introduction

1.1. Seismicity

Seismicity is defined as intensity, frequency, and spatial distribution of earthquakes in a specific area. It can be at local, regional and global scale.

1.1.1. Types of seismicity

Seismicity is generally divided in two categories which are defined as follows:

a) **Natural seismicity (NS)**: It is associated with occurrence of earthquakes naturally by plate tectonic, Volcanic activity etc. The magnitude range is generally from low to as high as 9.0 M.

b) **Induced seismicity (IS)**: It is referred to typically minor earthquakes and tremors that are caused by human activity that alters the stress field of the Earth's crust. IS events are predominantly of very low magnitudes and, in most cases, human activity is merely the trigger for an earthquake that would have occurred naturally in any case but at a later time. Despite the many similarities between IS and natural earthquakes, there are also notable differences. The most significant being the aspect of human intervention in natural processes.

Generally IS events range from moderately-sized ($4<M<6$), damaging earthquakes with rupture planes of the order of km^2 through rock bursts ($2<M<4$) down to microseismicity and acoustic emissions ($M<2$, rupture planes ≤ 1 m^2) for which humans are insensible and thus can be observed only by sensitive seismic instrumentation. The largest IS event recorded was a M=7.0 gas withdrawal event in Gazli (Uzbekistan) in 1984 Simpson and Leith (1985). The largest mining induced seismic event was a M=5.6 rock burst that occurred in a potash mine near Völkerhausen, Germany in 1989 Knoll and Kuhnt (1990). Significant IS events associated with long-term well injection include the 1966 M=5.2 event at the Rocky Mountain

Figure 1.1.: The damage caused by induced earthquakes, left image is showing Collapse of a block of flats in Welkom in 1976 after a mining-related earthquake of magnitude 5.2 (Images courtesy: of the Council for Geoscience, South Africa (SAWDIS)) and the right image is showing a house damaged in a pair of earthquakes that struck Sparks, Oklahoma, Nov. 6, 2011, in less than 24 hours,which is connected fracking while shale gas extraction(By : Bryan Walsh is a senior writer at TIME).

Arsenal, Denver Evans (1966). Enhanced geothermal system operations caused some notable events arising from short-term stimulation measures, for example the 2003 M=3.7 event at the Cooper basin EGS site, Australia Baisch et al. (2006) and the 2006 M=3.4 event in Basel (Switzerland). Numerous sizable events have been associated with gas withdrawal including the 1996 M=3.8 event near Strachan (Alberta), events up to M 4.2 near Lacq, France Maury et al. (1992) and up to M=3.5 in the Netherlands (e.g., Van Eck et al. (2006)). Enhanced oil and gas recovery events are usually below M=3, but possible IS associated with well operations have local magnitudes as high as 6.5 Nicholson and Wesson (1992). Some of the examples related to damage due to induced earthquakes are shown in figure 1.1.

The previous discussion provided a basis of the study of IS that represents an important topic for nowadays research. That is why study on IS have become an important topic these days. In this thesis we will focus on IS associated with geothermal areas.

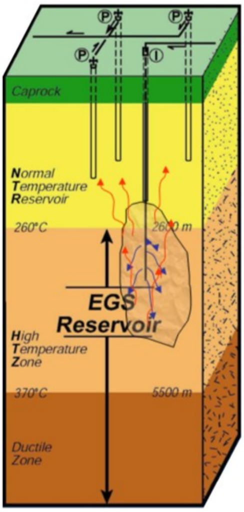


Figure 1.2.: The sketch showing geothermal area. (image: U.S department of energy)

1.2. Enhanced geothermal systems

Enhanced geothermal systems (EGS) have the potential to make a significant contribution to the world energy inventory. These are naturally fractured hydrothermal systems (figure 1.2), easiest sources to extract heat stored in subsurface rocks, but the total resources and its availability tend to be restricted to certain areas. Their development proceeds where conditions are ideal for cost-efficient extraction. These hydrothermal systems sometimes are difficult to locate and also run a high risk of not being commercially feasible, if their geological, physical and chemical characteristics are not favorable. The reasons for developing EGS technology are twofold: (1) to bring uneconomic hydrothermal systems into production by improving their underground conditions (stimulation); and (2) to engineer an underground condition that creates a hydrothermal system, where by injected fluids can be heated by circulation through a hot fractured region at depth and brought to the surface to deliver the captured heat for power conversion and other uses. The second approach expands the available heat resource significantly and reduces the uncertainty of exploitation costs.

However, the process of enhancing permeability and the subsequent extraction of energy may generates small-to-moderate events. The controversial issue of induced seismicity associated with EGS has attained attraction on seismologist worldwide. Although induced seismicity may have few adverse physical effects on operations or on surrounding communities, public concern over the amount and magnitude of seismicity associated with current and future EGS operations still remains. Thus inspired, in this thesis we are discussing about analysis on induced seismicity at the at "The Geysers" geothermal field, California.

1.3. "The Geysers" geothermal area and induced seismicity

The Geysers is a vapor-dominated geothermal field located about 120 km north of San Francisco, California. The area is in coastal ranges and is influenced by the general strike-slip tectonics of Northern California. Oppenheimer (1986) describes the tectonic setting as extensional, with the regional stress field predominating over locally induced stresses, mainly as result of reservoir contraction. The Geyser field is actively exploited since the 1960s, and it is now the most productive geothermal field in the world. It is capable of generating electric power of 1.58GW which is equal to one moderate nuclear power plant. The study of the IS is important also because there is not significant hazard studies in this geothermal area and hazard studies are required to monitor and control the rate of injection as well as seismicity to avoid any damage in surrounding areas.This controversial issue of IS has already caused delays or threatened cancellation of at-least two EGS projects worldwide Majer et al. (2007).

Regular monitoring of induced seismicity since 50 years in this area has result in one of the most comprehensive data sets available. Therefore, The Geysers is a well-suited test case for the study of induced seismicity and the development of related ground-motion prediction models. Many studies have demonstrated that micro-earthquakes at The Geysers are associated with water injection and steam extraction Majer and McEvilly (1979); Eberhart-Phillips and Oppenheimer (1984); Oppenheimer (1986); Enedy et al. (1992); Stark (1992, 2003); Ross et al. (1999); Smith et al. (2003); Majer and Peterson (2005); Majer et al. (2007); Foulger et al. (1997).

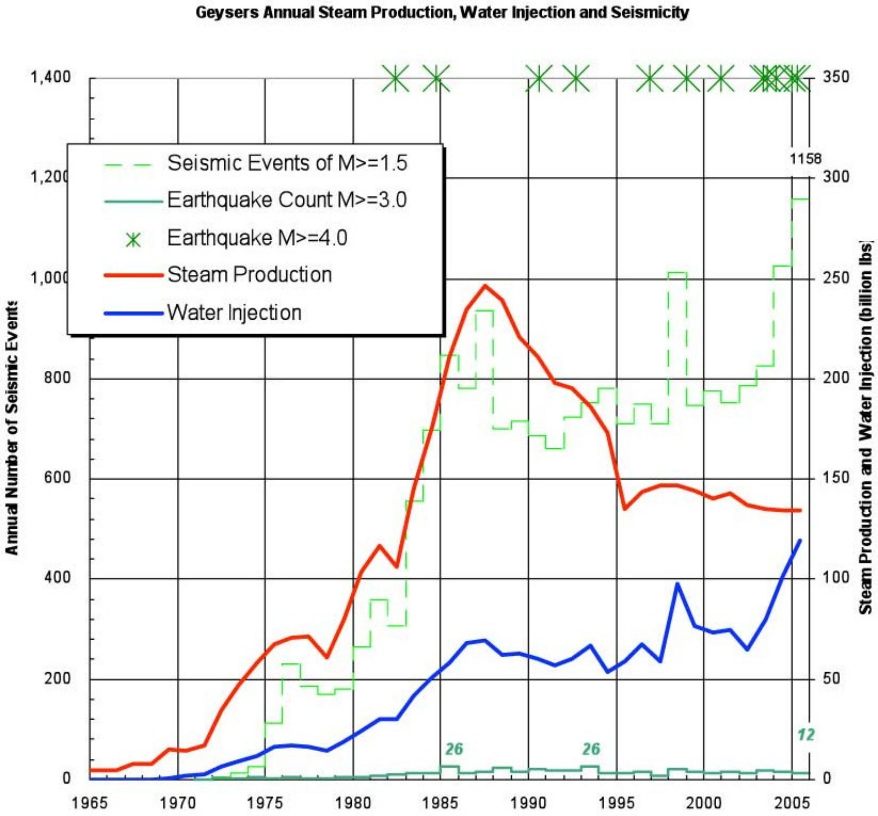


Figure 1.3.: The Geysers annual steam production (red line) plotted with the earthquake activity (M > 1.5 events), and water injection (blue line). Also shown are the M> 3.0 events (green solid line at bottom of plot) and M > 4 events (stars at top of chart), after Majer et al. (2007).

The seismicity since 1965 (roughly the date of significant production at The Geysers) is given in figure 1.3 Majer et al. (2007), which reveals that the seismicity below magnitude 3.0 has increased significantly over the years. The steam production and seismicity trends clearly diverge after additional source of water were used for injection, starting in 1980's. The level of seismicity (magnitude $M>1.2$) has shown positive correlation with steam production. Also, the "injection" chart is scaled such that the injection and seismicity values, at the time of the injection peak in 1998, plot more or less together.

Majer et al. (2007) have noticed that the injection and seismicity plots are now very similar for every year thereafter, as well as being quite similar for all the years previous to 1998. This findings clearly indicates a remarkably strong correlation of seismicity with injection. These data seem to confirm that shallow and deep induced micro earthquakes occurring after the 1980 are correlated to local injection rates, after a certain time lag Stark (1992, 2003); Smith et al. (2003); Enedy et al. (1992). For example, Stark (1992) showed that plumes of micro-earthquakes are clustered around many injection wells, and the seismic activity around each of these wells correlates with its injection rate.

The Geysers zone is subdivided into two seismicity source zones named ZONE1 and ZONE2 (dashed lines in figure 1.4). The separation arguments are supported by Beall et al. (2010); Stark (2003); Beall and Wright (2010). A difference in the seismicity distribution has been also noted by Eberhart-Phillips and Oppenheimer (1984). All the authors observed that the southeastern part of the Geysers reservoir is seismically less active than the northwestern where seismicity extends to greater depth. The differences was basically ascribed to a depth variation in the high-temperature (260-360° C) vapor-dominated reservoir. Specifically, Beall and Wright (2010) identified a net "$M \geq 4.0$ dividing line" which separates the whole area into two different seismic areas.

The northeastern area contains all the earthquakes with magnitude larger than 4.0, whereas the southeastern one is characterized by lower magnitude earthquake. Convertito et al. (2012) have also observed two different zones in the Geysers area by analyzing b-values of the Guttenberg and Richter (1944) relationships for the two areas using Utsu (1992) test.

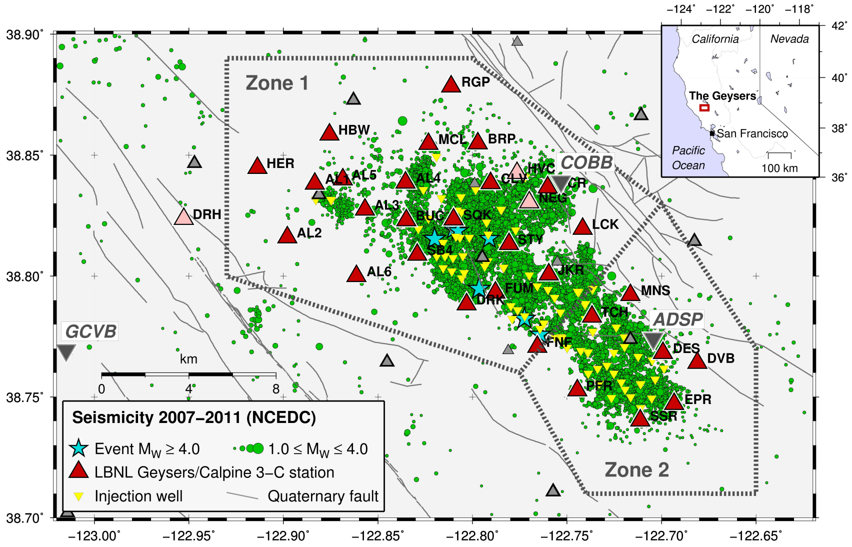


Figure 1.4.: Seismicity map of Geysers geothermal area.

Different temporary and long-term seismic networks have been deployed in The Geysers geothermal field during the last five decades. At present, local seismicity is monitored by the dense Lawrence Berkeley National Laboratory Geysers/Calpine (BG) surface seismic network and with some nearby stations of the Northern California Seismic Network (NCSN). The BG network consists of 32 three-component stations, 29 of which were used for the present study (figure 1.4, red triangles). The BG stations are distributed in an area of about 20x10 km^2, covering the entire geothermal field. Each station was initially equipped with I/O Sensor SM-6 geophones with a natural frequency of 14 Hz. In Fall 2009 these instruments were replaced successively by Oyo GS-11D 4.5 Hz sensors. The BG stations operate in trigger mode, and the waveform segments recorded since the end of July 2007 are made available at the Northern California Earthquake Data center (NCEDC) and the same data set is used in the present studies.

2. Seismic hazard basics

2.1. Seismic hazard

Seismic hazard refers to the study of expected earthquake ground motions at the earth's surface, and its likely effects on existing natural conditions and man-made structures for public safety considerations. The results of such studies are generally published as seismic hazard curves or maps, which identify the relative motion of different areas on a local, regional or global basis.

Thus seismic hazard is used to determine their risks, which are assessed and included in such areas as building codes for standard buildings, designing larger buildings and infrastructure projects, land use planning and determining insurance rates. It should be noticed that seismic hazard and risk are two different terms which are generally confused as one. Seismic risk uses the results of a seismic hazard analysis, and includes both consequence and probability. Seismic risk has been defined for most management purposes, as the potential economic, social and environmental consequences of hazardous events that may occur in a specified period of time.

Seismic Risk (SR) is defined by the degree of loss to a particular event i and the likelihood of Hi, equation 2.1

$$(RS)i = V \times Hi \times E \tag{2.1}$$

where vulnerability (V) is defined by the expected degree of loss ($0 < V < 1$) due to a destructive event, where: 0 means no loss and 1 means total loss and Hi is the hazard for the i^{th} event and E is exposition which signifies the location of the element placed at risk. Seismic risk is a combination of hazard, vulnerability and exposition, see figure 2.1.

High hazard is not always high-risk! A building located in a region of high seismic hazard is at lower risk if it is built to sound seismic engi-

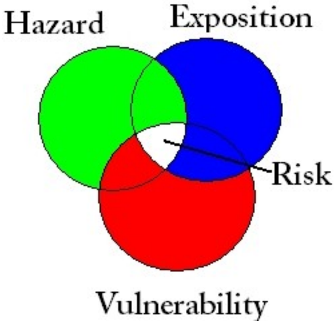


Figure 2.1.: Seismic risk is based on a combination of hazard (H) ,vulnerability (V) and exposition (E).

neering principles. On the other hand, a building located in a region with a history of minor seismicity, in a brick building located on fill subject to liquefaction can be at high or higher risk.

The seismic hazard studies are categorized in two types, deterministic seismic hazard and probabilistic seismic hazard. Before discussing these two types of hazard analysis lets have a look on some basics definitions which are used in hazard analysis mentioned in next section.

2.2. Basic definitions

- **Earthquake:** sudden release of stored strain energy in the earth's crust and propagation energy from the source in the form of shock waves in all directions. The earthquakes can vary from small, moderate to large catastrophic. Earthquakes are generally comment is "An earthquake occurred in Irpinia" which means that Irpinia was the center of damage.

- **Ground motion:** is an earthquake property or characteristic which results in shaking of ground in vertical and horizontal direction. Ground motion can be defined as peak ground displacement, velocity and acceleration.

- **Spectral acceleration:** is approximately what is experienced by a building, as modeled by a particle on a massless vertical rod having the same natural period of vibration as the building.

- **Response spectrum:** is simply a plot of the peak or steady-state response (displacement, velocity or acceleration) of a series of oscillators of varying natural frequency, that are forced into motion by the same base vibration or shock. The resulting plot can then be used to pick off the response of any linear system, given its natural frequency of oscillation.

- **Earthquake damage:** a destructive physical effect on a natural or artificial structures. Examples are the effects of seismic shaking on a building (e.g. broken windows, spalled concrete on columns or broken equipment and piping.)

- **Seismic hazard curve:** a graphical curve depicting the frequency (the number of earthquakes or events per unit time - usually per year) with which selected values of a seismic hazard such as ground motion amplitude are expected to occur (or, more typically, are expected to be exceeded).

- **Damage function:** a relationship between levels of damage and the corresponding levels of shaking. For example, with this function, the damage to a structure for a given ground motion input can be estimated. Damage functions can be derived either empirically or analytically.

- **Loss function:** a relationship between monetary or human loss (for example, the number of causalities) and earthquake damage or levels of ground shaking. Loss may be estimated directly from ground motion amplitudes, either in monetary units or as a fraction of building value, or the levels of damage can be estimated first, and then the loss from the estimated damage can be calculated.

- **Aleatory (or Random) uncertainty:** the probabilistic uncertainty that is inherent in a random phenomenon and cannot be reduced by acquiring additional data or information. In the past, this type of uncertainty has been called randomness or inherent variability.

- **Epistemic (or Knowledge) uncertainty:** that uncertainty that results from lack of knowledge about some model or parameter. This type of uncertainty can be reduced, at least conceptually, by additional data or improved information. In the past this uncertainty has also been called statistical or professional uncertainty, or simply uncertainty.
- **Recurrence interval:** the mean (average) time between occurrences of a given type of earthquake . For example, an earthquake of a specified magnitude on a fault or in a region.
- **Return period:** the mean (average) time between occurrences of a seismic hazard. For example, a certain ground motion at a site, or a certain level of damage or loss.
- **Exposure time:** the mean (average) time for which a structure is exposed to certain seismic hazard.
- **De-aggregation:** statistical decomposition of a hazard to show the relative contribution by magnitude, distance and ground motion deviation.

2.3. Deterministic seismic hazard analysis (DSHA)

Deterministic approach develops scenario for a particular earthquake: an earthquake with specified size, occurring in a specified location - which is assessed based on ground motion at the site of interest. A possible result of deterministic seismic hazard analysis could be: the earthquake hazard at site at X is a peak ground acceleration of 0.5g (50% of the acceleration of gravity) resulting from the occurrence of magnitude 6.5 earthquake on fault Y at the distance of 10 km. Thus this analysis requires 3 basic elements; an earthquake source (fault Y), a controlling earthquake of specified size (magnitude 6.5) and the means of determining hazard, in this case peak ground acceleration (0.5g), at the specified distance (10 km) to the site. Standard deterministic methodology consists of the following steps (figure 2.2):

Step 1 is to identify a source or several possible sources that may affect the site. Configuration of individual sources can by points, lines, areas or

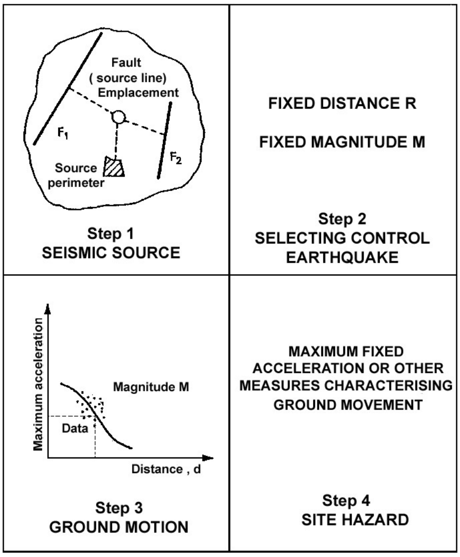


Figure 2.2.: Main steps for deterministic seismic hazard analysis.

volume, depending on the source type and the possibility to define them geologically.

Step 2 is to select the so-called control earthquake. Seismic potential of each seismic source, as described in Step 1, depends on the maximum earthquake that can be generated in the source. This can be the expected earthquake, the maximum credible earthquake or any other type of earthquake. Selection of the specific criterion is one of the most important elements in determining the conservatism level. Responsibility for this control earthquake choice is immense and this constitutes the most vulnerable part in the deterministic analysis. Earthquake magnitude and epicentral intensity are commonly used to define the size (which here can be: duration, magnitude, maximum acceleration, etc..) of the earthquake. In addition to these sizes, there is the requirement of selecting the most appropriate definition of distance, which is the distance (usually the nearest) between the source and location. One of these hypothetical earthquakes will be the control earthquake, an earthquake that generates indices (intensity, peak acceleration, relative speed, predominant period, etc.) which will dominate the effects of other earthquakes. This earthquake will be considered to be the most important in defining the seismic hazard. In this stage several control earthquakes can be used because it is not always clear which event is associated with the largest movement of land at the site of interest. This can happen when a seismic source is as important as other, or when using multiple parameters of the site terrain in defining the seismic hazard (e.g, maximum acceleration, maximum relative velocity, maximum relative displacement, the fundamental period, maximum spectral acceleration, etc.).

Step 3 is the determination of earthquake effect, usually some type of ground-motion (e.g, maximum acceleration, maximum relative velocity, maximum relative displacement, the fundamental period, maximum spectral acceleration, etc.), at the site. Typically this is done by means of an earthquake ground-motion prediction equations which provides estimates of ground motion for an earthquake of a given magnitude at different distances by means of a curve fitted to observed data. If not determined in Step 2, the controlling earthquake is defined based on resulting in the largest ground motion.

Step 4 is the definition of hazard at the site. In the most cases it is the direct output of step 3 . This is usually a simple statement to the effect that the hazard at the site is a specific peak ground acceleration, velocity or other measure that describes earthquake effect.

Expressed in these 4 compact steps, DSHA appears to be very simple procedure and in many respect it is. When applied to the structures for which failure could have catastrophic consequences, such as nuclear power plant and large dams, DSHA provides a straight forward framework for evaluation of worst-case ground motions. However it does not provide any information on the likelihood of occurrence of the controlling earthquake, the likelihood of its occurrence, the level of ground shaking that can be expected during the finite period of time or the effect of uncertainties in the various steps required to compute the resulting ground motion characteristics. Perhaps the most important, DSHA involves subjective decisions, particularly regarding earthquake potential (step 1), that can require the combined expertize and opinions of seismologists, geologists, engineers, risk analysts, economists, social scientists and government officials.

2.4. Probabilistic seismic hazard analysis (PSHA)

Probabilistic approach, proposed by the pioneering work of Cornell (1968), has become a standard method which is widely accepted and used worldwide. The probabilistic approach allows uncertainties in the size, location, and rate of recurrence of earthquake and in the variability of ground motion characteristics with earthquake size and location to be explicitly considered in the evaluation of seismic hazards. It provides the frame-work in which these uncertainties can be identified, qualified, and combined in a rational manner to provide a more complete picture of seismic hazard. It consists of four basic steps (figure 2.3), some of which partially overlap with those of the deterministic approach (Reiter, 1990).

Stage 1 At this stage we define seismic sources. It is generally similar to the step 1 of deterministic analysis except that the sources are explicitly defined as having a uniform seismic potential, i.e, the probability of occurrence of earthquakes or an earthquake by a certain size is the same in the source. Sources may vary from well know fault to entire seismotectonics province.

Stage 2 At this stage we define recurrence seismic characteristics for each source. This step is fundamentally different from step 2 of the deterministic analysis. Instead of checking the control earthquake or maximum earthquake within each source, here each source is characterized by a recurrence relationship or a probability distribution of earthquakes occurrence. A recurrence relationship indicates the probability of occurrence of a given size earthquake, with the epicenter anywhere in the source, within a time-frame, usually one year. Maximum earthquake is selected for each source. Compared with the deterministic procedure, this earthquake is not only the maximum considered earthquake, but the upper limit of the size of earthquakes that will enter the analysis for each source considered.

Stage 3 Now estimate the earthquake effects at site location. In probabilistic analysis a family of attenuation curves for each magnitude is used. Each attenuation curve has its degree of uncertainty with the known data set (curves M1, M2, M3 ...).

Stage 4 Now integrate the entire range of magnitudes and distances for each seismic source to obtain (in a particular location) probabilistic hazard values in the form of cumulative distributions for parameters that describe the movement of land. The effects of all earthquakes of various sizes,

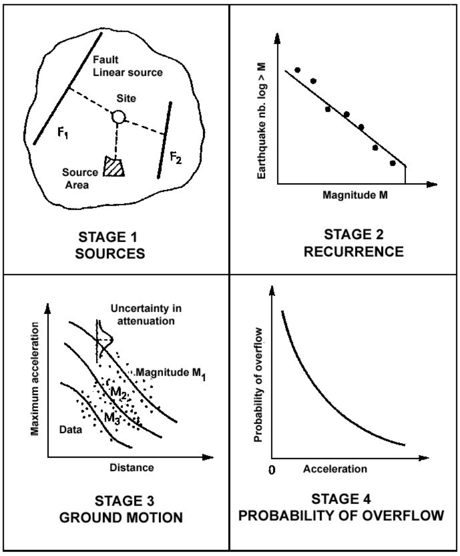


Figure 2.3.: Main stages of the probabilistic seismic hazard analysis (after Reiter (1990)).

produced in different locations and different seismic probabilities are integrated into a single curve, which expresses the probability of exceedance in a specified time period, certain values of parameters describing the seismic motion at the site of interest.

The proper performance of a PSHA requires careful attention to the problems of source characterization and ground motion parameter prediction. A PSHA requires that uncertainties in earthquake location, size, recurrence and ground shaking effects be quantified.

2.4.1. Earthquake source characterization

Characterization of earthquake source requires consideration of the spatial characteristics of the source and of the distribution of earthquakes within that source, of the distribution of the size for each source, and of the distribution of the earthquakes with time. Each of these characteristics involves certain degree of freedom.

Source effect

The geometry of source depends on the type of tectonic activity involved in their formulation. For, example micro-earthquakes generated by drilling process can be considered as *point sources*, well-defined fault planes on which earthquakes can occur at many different locations, can be considered as two dimensional *areal sources*. Areas where mechanism of earthquakes are poorly defined, or where there is extensive faulting to preclude distinction between individual faults, can be treated as three-dimensional *volumetric sources*.

In seismic hazard analysis, *source zones* may be similar to or somewhat different from actual source, depending on the relative geometry of the source and site of interest and also on the quality of information of the source. For example, relatively short fault in figure 2.4a can be modeled as point source since the distance between any point along its length and the site is nearly constant. Similarly the depth of fault plane shown in figure 2.4b is sufficiently small that the variations in hypocentral depth have little influence on hypocenteral distance. In such a case the hazard analysis can be simplified with negligible loss of accuracy by approximating the planar source as a *linear source zone*. If the available data are insufficient to

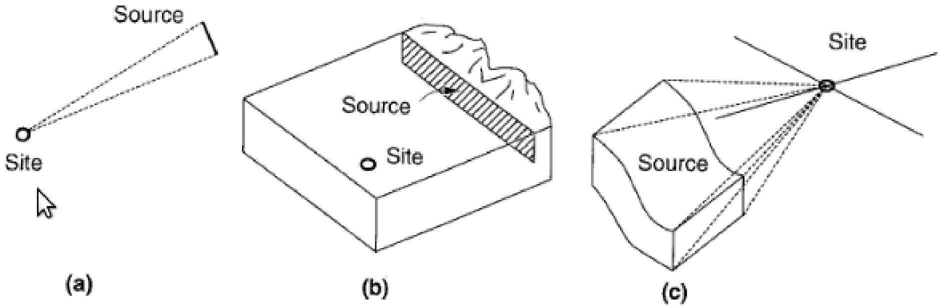


Figure 2.4.: Examples of different source zone geometries: (a) short fault can be modeled as point source; (b) shallow fault can be modeled as linear source; (c) volumetric source zone.

determine the actual geometry of the source accurately, the source can be considered as volumetric source zone shown in figure 2.4c.

For each source zone, uncertainty in earthquake location is characterized by a probability density function of source to site distance. Evaluation of probability density function requires the estimation of the geometry of the source zone and of the distribution of the earthquakes within it.

Magnitude contribution

After identifying and characterizing an earthquake source zone, the seismic hazard analyst's attention is turned toward evaluation of the sizes of earthquakes that the source can be expected to produce. Every source zone is capable to produce maximum earthquake magnitude that rarely exceed; it can be large for some and small for others. In general, a source zone produces earthquakes of different magnitude, with the smaller earthquakes occurring more frequently than larger ones. The distribution of earthquake sizes is described by a *recurrence law.* In PSHA, it is assumed that the recurrence law obtained from the past seismicity is appropriate for the prediction of future seismicity (i.e. seismicity is assumed to be stationary).

Gutenberg-Richter relationship Gutenberg and Richter in 1944, developed the recurrence law after analyzing seismic data from California.

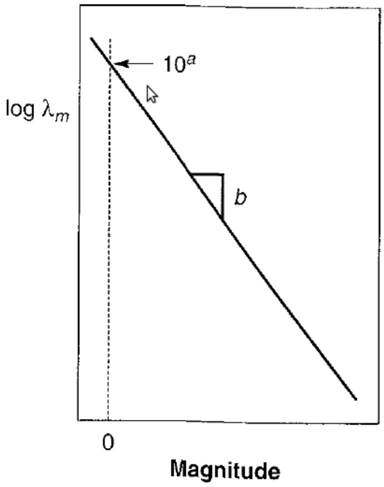


Figure 2.5.: Example of Gutenberg-Richter Law, representing a and b value.

They grouped the events excelling certain magnitude threshold in that period. Then they divided the number of events in each magnitude range by the time length which can be defined a mean annual rate of exceedance, λ_m of an earthquake of magnitude m. The resulting Gutenberg-Richter Law is expressed in equation 2.2

$$\lambda_m = 10^{a-bm} \tag{2.2}$$

where, 10^a is yearly number of earthquakes of magnitude greater than equal to zero. b describes the relative likelihood of small and large events, see figure 2.5. As the b value increases number of larger magnitude earthquake decreases compared to those of smaller magnitude Guttenberg and Richter (1944).

This standard Gutenberg-Richter relation covers an infinite range of magnitudes from $-\infty$ to $+\infty$. However all the earthquakes are not damageable. Moreover to compute PSHA minimum (m_{min}) and maximum (m_{max}) magnitude of earthquake is required. Hence, if the earthquakes smaller than the threshold magnitude (m_{min}) are eliminated the mean annual rate of exceedence can be expressed as McGuire and Arabasz (1990):

$$\lambda_m = \nu exp[-\beta(m - m_{min})] \tag{2.3}$$

The equation 2.3 can be defined as bounded Gutenberg-Richter relationship, where $\nu = exp(\alpha - \beta m_{min})$, $\alpha = aln(10)$ and $\beta = bln(10)$. The probability distribution of magnitude for Gutenberg-Richter relationship (conditioned to magnitude larger than m_{min}) can be expressed as cumulative density function (CDF) (equation 2.4)

$$\begin{aligned} F_M(m) &= P[M < m_{max}|M > m_{min}] = \frac{\lambda_{m_{min}} - \lambda_m}{\lambda_{m_{min}}} \\ &= 1 - exp[-\beta(m - m_{min}] \end{aligned} \tag{2.4}$$

or as probability density function (equation 2.5):

$$f_M = \frac{d}{dm} F_M(m) = \beta exp[-\beta(m - m_{min}] \tag{2.5}$$

On the other side the standard Gutenberg-Richter relation predicts nonzero mean rate of exceedance. If maximum magnitude m_{max}, associated with each zone is known or can be estimated, the mean annual rate of exceedance (McGuire and Arabasz, 1990) can be expressed as;

$$\lambda_m = \nu \frac{exp[-\beta(m - m_{min})] - exp[-\beta(m_{max} - m_{min})]}{1 - exp[-\beta(m_{max} - m_{min})]} \tag{2.6}$$

The equation 2.6 is called bounded recurrence law. The CDF and PDF for the Gutenberg-Richter law with upper and lower bounds can be expressed as:

$$F_M(m) = P[M < m|m_{min} \leq m \leq m_{max}] = \frac{1 - exp[-\beta(m - m_{min})]}{1 - exp[-\beta(m_{max} - m_{min})]} \tag{2.7}$$

$$f_M(m) = \frac{\beta exp[-\beta(m - m_{min})]}{1 - exp[-\beta(m_{max} - m_{min})]} \tag{2.8}$$

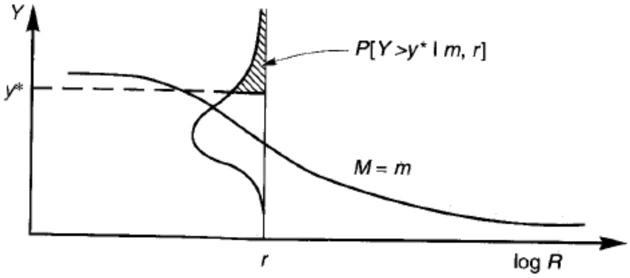


Figure 2.6.: Example showing conditional probability of exceeding a particular value of ground motion parameter for a given magnitude and distance.

2.4.2. Ground-motion prediction equations (GMPE)

This is one of the most essential part of seismic hazard analysis. This section will be discussed in more detail in chapter 4. The GMPEs are functional relationships computed from regression techniques to predict ground motion parameters (PGA, PGV and SA) for different magnitude at different distance ranges. It is generally defined as :

$$Y = Af(m)f(r)f(m,r)f(s)\sigma \tag{2.9}$$

where, A is constant, $f(m)$ is function of magnitude (source effect), $f(r)$ is the function of distance (path effect), $f(m,r)$ is a combined function of magnitude and distance, $f(s)$ is the function of local site effect and σ is the total uncertainty on the predicted ground motion parameter Y. In PSHA the probability that a particular ground motion parameter Y exceeds the $y^*$for an earthquake of given magnitude M and occurring at certain distance R is represented in a figure 2.6

In probabilistic term it is expressed in equation 2.10

$$P[Y > y* |m,r] = 1 - F_y(y^*) \tag{2.10}$$

where, $F_y(Y)$ is the value of CDF of Y at m and r. The value of $F_y(Y)$ depends on the probability distribution used to represent Y. In general, it is assumed that ground motion parameters have log-normal distribution, however the unbounded characteristics of that distribution can attribute a non zero probability to unrealistic values of the ground motion parameter, thus a additional parameter ε is introduced.

2.4.3. Earthquake occurrence model

The distribution of occurrence of earthquakes in given time must be considered to calculate probabilistic seismic hazard analysis. One of the most common and popular model to describe the temporal occurrence of the earthquakes is *Poisson model* which evaluate the probability of occurrence of the event as Poisson process. The Poisson process has following properties:

- The number of occurrence in one time interval are independent of the number that occurs in any other time interval.
- The probability of occurrence during a very short time interval is proportional to the length of the time interval.
- The probability of more than one occurrence during a very short time interval is negligible.

These properties shows that the events of Poisson process occur randomly, with no "memory" of time, size or location of any preceding event. The probability of a random variable n is equal to N, representing the number of particular event during the given interval described by Poisson process is given by equation 2.11

$$P[N = n] = \frac{\mu^n e^{-\mu}}{n!} \tag{2.11}$$

where μ is the expected average number of occurrence of the event. The temporal distribution of earthquakes for PSHA, can be expressed by Poisson model as

$$P[N = n] = \frac{(\lambda t)^n e^{-\lambda t}}{n!} \tag{2.12}$$

where λ is the average rate of occurrence of the event and t is the time period of interest. The probability of occurrence of at least one event in a period of time t is given by equation 2.13

$$\begin{aligned} P[N &\geq 1] = P[N=1] + P[N=2] + \dots \\ &+P[N=\infty] = 1 - P[N=0] = 1 - e^{-\lambda t} \end{aligned} \tag{2.13}$$

Thus Poisson model can be combined with a suitable recurrence law to predict the probability of at least one exceedance in a period of t years by the expression 2.14

$$P[N \geq 1] = 1 - e^{-\lambda t} \tag{2.14}$$

There are several model like Non-homogeneous Poisson models (e.g., Vere-Jones and Ozaki (1982), Time predictable models, slip-predictable models, Markov models etc. Each of the more sophisticated models uses a "pattern" of the earthquake occurrence to reconcile their computed probabilities with the mechanism of the elastic rebound process of the earthquake generations. As a result an extra information is required for each of the models which should be evaluated from historical and instrumental seismicity records that are too sparse for accurate evaluation. But with the time and availability of more data from advanced instruments use of other sophisticated models will increase.

2.4.4. Probabilistic seismic hazard computation

Probabilistic seismic hazard is generally expressed as a curve which indicates the frequency of exceedance of different ground motion parameters. These can also be used to compute the probability of exceeding the selected ground motion parameter in specific period of time. Seismic hazard curve can be obtained for individual source zones and combined to express the aggregate hazard at a particular site.

For a given set of earthquakes, the probability that the ground motion parameter Y will exceed a particular value y^* can be computed using the total probability theorem, that is equation 2.15:

$$P[Y > y^*] = \iiint P[Y > y^* | m, r, \varepsilon] f_M(m) f_R(r) f_\varepsilon(\varepsilon) dm dr d\varepsilon \tag{2.15}$$

where m, r are magnitude and distance that influence Y, $P[Y > y^* | m, r]$ is obtained from GMPE's and $f_M(m)$, $f_R(r)$, and $f_\varepsilon(\varepsilon)$ are probability density functions for the magnitude and distance, respectively. If the number of potential sources in the region are say Ns and each of them has an average rate of threshold magnitude exceedance, $\nu_i[= exp(\alpha_i - \beta_i m_0)]$, the total average exceedance rate for the region will be expressed as

$$\lambda_{y*} = \sum_{i=1}^{N_S} \nu_i \iiint P[Y > y^*|m, r, \varepsilon] f_{Mi}(m) f_{Ri}(r) f_\varepsilon(\varepsilon) dm dr d\varepsilon \tag{2.16}$$

The individual components of the equation 2.16 are for realistic PSHAs, but these complicated integrals cannot be evaluated analytically. Numerical integrations can be performed with different techniques. One approach suggested here is, to divide the possible ranges of magnitude and distances into N_M and N_R segments, respectively. The average rate of exceedance can be estimated by

$$\lambda_{y*} = \sum_{i=1}^{N_S} \sum_{j=1}^{N_M} \sum_{k=1}^{N_R} \nu_i P[Y > y^*|m_j, r_k, \varepsilon] f_{Mi}(m_j) f_{Ri}(r_k) f_\varepsilon(\varepsilon) \triangle m \triangle r \triangle \varepsilon \tag{2.17}$$

where $m_j = m_0 + (j - 0.5)(m_{max} - m_{min})/N_M$,$r_k = r_{min} + (k - 0.5)(r_{max} - r_{min})/N_R$,$\triangle m = (m_{max} - m_{min})/N_m$, and $\triangle r = (r_{max} - r_{min})/N_R$ and $\triangle \varepsilon = \varepsilon / N_s$. This is equivalent to assuming that each source is capable of generation only $N_M$different earthquakes of magnitude m_j at only N_R, different source to site distances, r_k. Equation 2.17is then equivalent to

$$\lambda_{y*} \approx \sum_{i=1}^{N_S} \sum_{j=1}^{N_M} \sum_{k=1}^{N_R} \nu_i P[Y > y^*|m_j, r_k, \varepsilon] P[M = m_j] P[R = r_k] P[\varepsilon] \tag{2.18}$$

The accuracy of the equation increases with N_M, N_R and N_ε. The seismic hazard curve can be easily combined with equation 2.14 to get the probability of exceedance of y^* in the particular time period.

2.5. Deterministic Vs Probabilistic approach

Now the question arises which approach is better and when? Sometimes it is one of the most controversial and difficult question faced by those defining standards for the use of earthquake hazard analysis .While there is a world wide trend among scientists and engineers toward probability criteria, the situation is by no means clear. In many cases the question has been rephrased so that the issue is not "weather" but rather "to what extent" a particular approach should be used.

The deterministic approach provides a clear and track-able method of computing seismic hazard whose assumptions and elements are easily discerned and can be examined by their own merits. It provides engineers and other users with understandable scenarios that can be related to the problem at the hand. However, because it deals with discrete controlling events and has no open or formal way of accounting for uncertainty it tends to disregard the frequency of earthquake occurrence and can lead to wrong assumption that there is no uncertainty. Conclusion based on deterministic seismic hazard analysis can be easily upset by the occurrence of new earthquakes and the development of new hypotheses and understanding. On the other hand, probabilistic seismic hazard analysis is capable of including and integrating a wide range of information and judgment and their associated uncertainties into the flexible framework that also permits the application of societal goals and priorities. Unfortunately its highly integrative nature can obscure those elements which derives the results, its highly quantitative nature can lead to false impressions of accuracy and it open embrace of uncertainty that can make decision-making difficult.

3. Inversion techniques

3.1. What is an inverse problem ?

An inverse problem is a general approach which is used to convert observed measurements into information about a physical object or system that we are interested in. For example, if we have measurements of the Earth's gravity field, then we might ask the question: "given the data that we have available, what can we say about the density distribution of the Earth in that area?" The solution to this problem (i.e. the density distribution that best matches the data) is useful because it generally tells us something about a physical parameter that we cannot directly observe (Menke (1990)). Thus, inverse problems are some of the most important and well-studied problems in science and mathematics.

The forward problem can be conceptually formulated as follows:

Data → Model→Model parameters

The inverse problem is considered the "inverse" to the forward problem which relates the model parameters to the data that we observe:

Model parameters → Model → Data

The transformation from data to model parameters (or vice versa) is a result of the interaction of a physical system with the object that we wish to study. In other words, it relates the physical quantity (i.e. the model parameters) to the observed data.

There are different techniques to solve inverse problem (linear and nonlinear) and we will discuss some of them which are of our interest and within the scope of this thesis.

3.2. Linear regression analysis (LRA)

Linear regression aims to model the relationship between two variables by fitting linear (straight line) equation to the observed data. A linear model can be expressed as

$$y = m_0 + m_1 x + \varepsilon \tag{3.1}$$

where y is dependent variable, x is independent variable, m_0 and m_1 are the model parameters and ε is the total standard error on fit which is estimated with least square method, Draper and Smith (1996). Any linear model can be formulated as

$$d = Gm \tag{3.2}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ 1 & x_n \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \end{bmatrix} \tag{3.3}$$

where d is dependent variable (observed data) matrix, G is matrix of independent variable and m is model parameter matrix. These matrices are solved to explore the values of m_0 and m_1. Hence the model parameters can be computed by expression:

$$m = G^{-1}d \tag{3.4}$$

If G is not a square matrix then equation 3.4 can be expressed as:

$$m = [G^T G]^{-1} G^T d \tag{3.5}$$

We can try to understand more with an example of simple linear model.

$$log(Y) = a + bm + clog(r) \tag{3.6}$$

This equation is a linear model known as ground motion prediction equation (see chapter 2 and 4), where Y is a ground motion parameter (such as peak ground velocity (PGV) or peak ground acceleration (PGA) or spectral acceleration(SA)) that has been recorded by the set of stations for a set of earthquakes and m is magnitude (account for source effect) and r is the source-to-site distance, accounts for geometrical spreading. In order to understand the dependency of Y on m and r, we will solve the equation

Table 3.1.: Linear regression Coefficients with uncertainties analysis for equation 3.6

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	σ_{total}
PGA (m/s^2)	-2.9403 ± 0.0298	1.1558 ± 0.0101	-2.0339 ± 0.0245	0.3856

3.6 and estimate the model parameters *a, b* and *c*. The equation 3.6 can be arranged as matrix:

$$\begin{bmatrix} log(Y_1) \\ log(Y_2) \\ . \\ . \\ log(Y_n) \end{bmatrix} = \begin{bmatrix} 1 & m_1 & log(r_1) \\ 1 & m_2 & log(r_1) \\ . & . & . \\ . & . & . \\ 1 & m_n & log(r_n) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{3.7}$$

This matrix is easy to solve and obtained the model parameters and with any combination of magnitude m and distance r the value of any considered ground motion parameter (say PGA or PGV) can be predicted.

After using the data from 212 earthquakes of The Geysers area (see chapter 4) and solving the matrix shown above, parameters with their respective uncertainties obtained for PGA are listed in table 3.1.

The model is plotted in figure 3.1 and is obtained by setting $m=2.5$ and *r,* distance ranges from 0.4 through 20 km. It can be clearly observed from uniformed residual distribution that model fits quite well to observed data. The residuals here are defined as $logY^{observed\text{-}data}$ - $logY^{predicted\text{-}data}$ The model predicts very high unexpected values at short distances for small-to-moderate magnitude earthquakes.

3.3. Magnitude and distance dependencies

The above discussed model (equation 3.6) depends on the magnitude m and distance r . But we can also check the dependency of the model or data only on the distance or on the magnitude by fixing the effect of one or another. The details are discussed below.

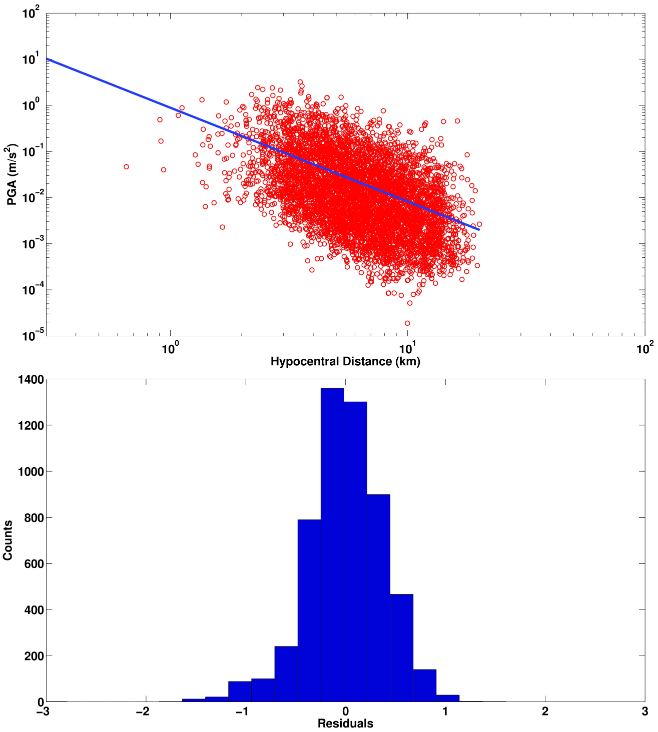


Figure 3.1.: Fitting of model mentioned in equation 3.6 uniformed residual distribution that model fit quite well to observed data, but predicts very high unexpected PGA values at short distances. The model is obtained by setting m=2.5 and hypocentral distance ranges from 0.4 through 20 km.

Table 3.2.: Linear regression coefficients with uncertainties for equation3.9

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	σ_{total}
PGA x r (m/s^2)	-3.7118 ± 0.0270	1.1329 ± 0.0116	0.4440

3.3.1. Effect of magnitude

To understand the dependency of ground motion parameter (PGA) on the magnitude we have to remove the effect of distance from the model, the equation 3.6 can be rearranged as:

$$log(Y) - clog(r) = a + bm \tag{3.8}$$

if c is assumed -1, which corresponds to assuming that direct S-waves carry peak ground acceleration and spherical geometrical spreading is assumed. Thus we can recompute the model parameters a and b and can see the linear dependency of PGA on magnitude. Thus the equation 3.8 can be re-written as

$$log(Y \times r) = a + bm \tag{3.9}$$

the coefficients are mentioned in table 3.2 and the fitting of the model is shown in figure 3.2. It can be clearly observed that when the effect of distance is removed, PGA increases with magnitude. The larger the magnitude (keeping the distance same), the larger will be the peak-ground shaking.

3.3.2. Effect of distance

In order to understand the dependency of ground motion parameter (PGA) on the distance now we have to remove the effect of magnitude. Thus the equation 3.6 can be rearranged as:

$$log(Y) - bm = a + clog(r) \tag{3.10}$$

where b parameter can be assumed as 1 which correspond to assuming that peak ground accelerations are directly dependent on size of earthquake (magnitude), and a and c can be recomputed through linear regression. We can observe the increment or decrement of PGA with distance. The

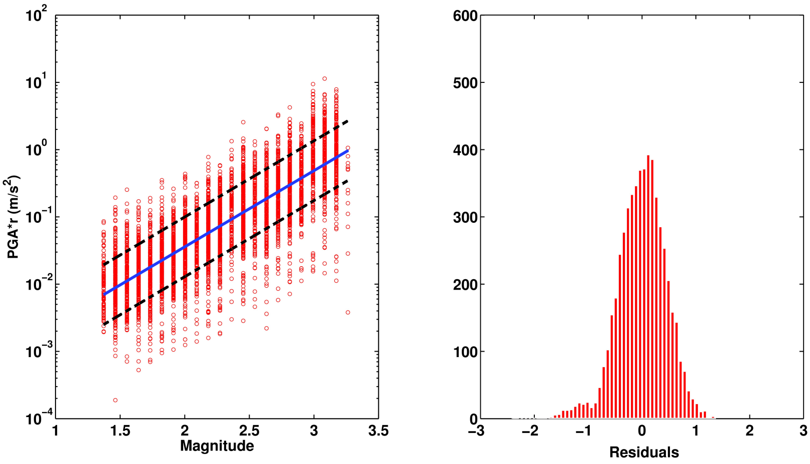


Figure 3.2.: Linear fit of the model mentioned in equation 3.9 showing positive correlation between magnitude and PGA values and uniform residual distribution signifies goodness of the fit.

Table 3.3.: Linear regression coefficients with uncertainties for equation3.10.

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	σ_{total}
PGA - m (m/s^2)	-2.6030 ± 0.1652	-2.0136 ± 0.2004	0.3938

computed coefficients with uncertainty for equation 3.10 are mentioned in table 3.3. The fitting of the model is shown in figure 3.3. It can be observed that the PGA is linearly decreasing with distance. For the set of magnitude value PGA decreases with increase in distance.

As it is observed in the figure 3.1, the model mentioned in equation 3.6 predicts very high values of PGA at short distances, these values are assumed to be unrealistic. This problem can be solve by including a fictitious parameter *(h)* with r (see equation 3.11).

$$log(Y) = a + bm + clog(\sqrt{r^2 + h^2}) \tag{3.11}$$

In many cases it (h parameter) is also regarded as fictitious depth when r is epicentral distance and it is introduced to avoid saturation with distance.

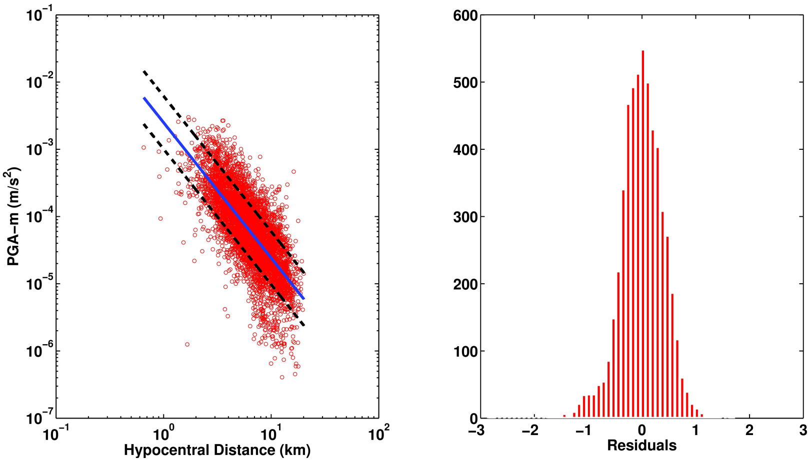


Figure 3.3.: Linear fit of the model mentioned in equation 3.10 showing negative correlation between distance and PGA values and uniform residual distribution signifies goodness of the fit.

However, if r considered as hypo-central distance this parameter cannot be regarded as fictitious depth. In this case it should be considered just as mathematical parameter which is used to avoid unwanted/unrealistic high values of ground motion at smaller distances.

The introduction of the h parameter makes the empirical equation 3.11 a non-linear model, which requires a specific solution different form the matrix inversion. However, this equation can be solved by linearizion, by fixing different values of h and recomputing other 3 coefficients (a, b, and c) and total standard errors. The best h parameter value is decided by minimum standard error and R^2 (defines the goodness of the fitting of model to observed data) statistics too (see table3.4) (Draper and Smith (1996); Emolo et al. (2011); Joyner and Boore (1981)). Figure 3.4 shows that at h=2.0 km total standard error is minimum. The corresponding model together with residual distribution is shown in figure 3.5 (for m= 2.5 and r ranges from 0.01 to 20 km) with blue curve. By comparing with figure 3.1, it is clear that problem of saturation at short distances is solved.

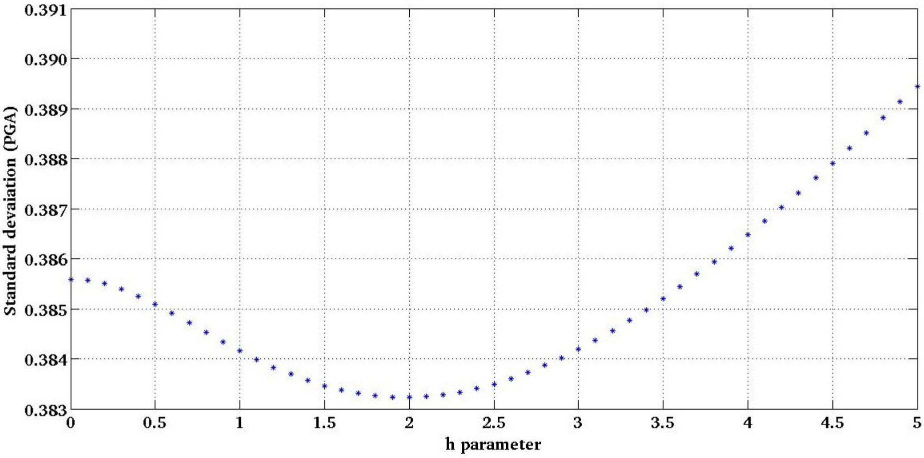


Figure 3.4.: An example illustrating the choice of h parameter by minimizing total standard error, with linear regression analysis for equation 3.11.

Table 3.4.: Linear regression coefficients with uncertainties for equation 3.11

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	h	σ_{total}	R^2
PGA (m/s^2)	-2.6278 ± 0.0319	1.1606 ± 0.0101	-2.3480 ± 0.0280	2.0	0.3832	0.7801

Table 3.5.: Coefficients with uncertainties for equation 3.11, after non linear regression analysis. It is also shown that there is no significant improvement in model fitting as the total standard error and R^2 statistics are same.

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	σ_{total}	R^2
PGA ($\mathrm{m/s^2}$)	-2.6347 ± 0.0659	1.1605 ± 0.0101	-2.3415 ± 0.0622	1.9752 ± 0.2105	0.3832	0.7801

3.4. Non-linear regression analysis (NLRA)

Non-linear regression is the form of regression analysis in which the observed data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. (for further details see Draper and Smith (1996)). For example, equation 3.11 is a non-linear model, if h is considered as a parameter to be computed along with other parameters a, b, and c. Thus now there are 4 model parameters to be estimated instead of 3 parameters computed in linear regression analysis. Table 3.5 lists the computed coefficients with uncertainty.

We can also compare the model fitting of linear regression with non linear regression analysis. It is evident from comparison between table 3.4 and table 3.5, that model fitting has not improved after solving equation 3.11 with non-linear technique (see figure3.5 blue and red curve). The standard deviation and the R^2statistics defines the goodness of the fit of a model to observed data. Thus these values are same in both cases hence the NLRA has not improve the fit of the model to observed data as compared to LRA.

3.5. Non-linear mixed effect regression analysis (NLMRA)

This method is similar to non-linear regression but it has a special advantage that it consider two effect on the non-linear problem i.e fixed effects and random effects (mixed model) that is why it is also regarded as mixed effect regression analysis, Lindstrom and Bates (1990). The mixed models are generally used for repeated measure data. It means that the data

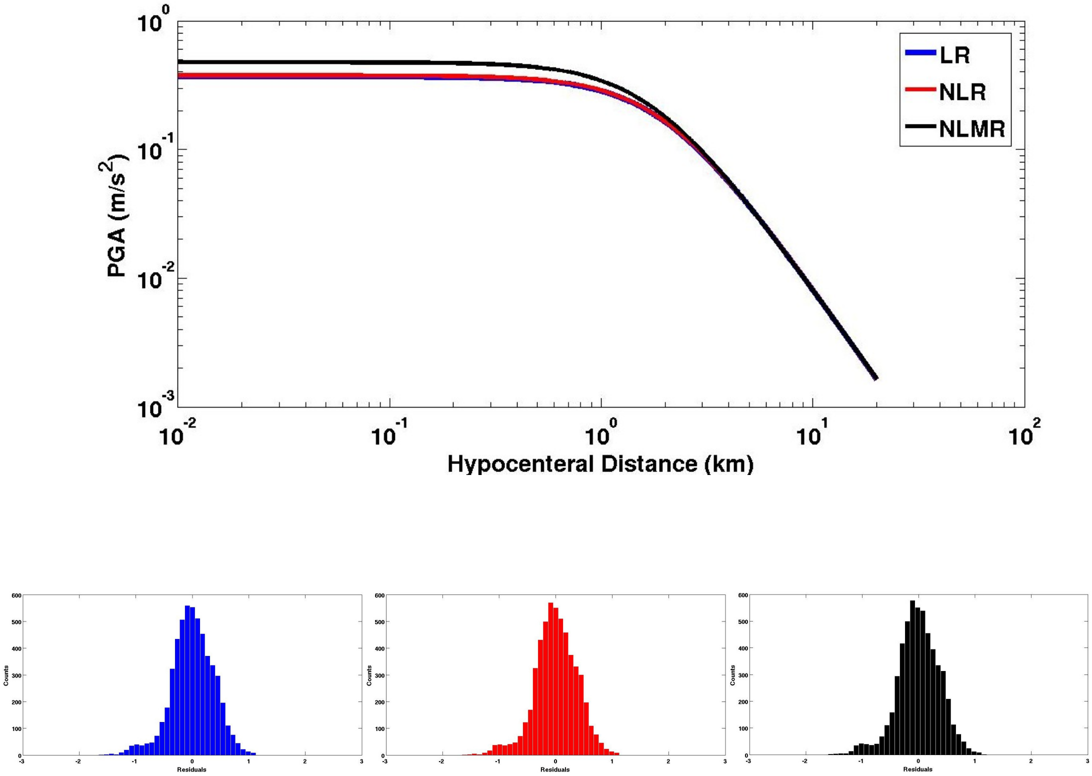


Figure 3.5.: Comparison of model mentioned in equation 3.11, computed by linear regression (LR, blue curve), non-linear regression (NLR, red curve) and non-linear mixed effect regression (NLMR, black curve) for PGA. All the models have same predictions and similar residual distribution shown in lower panels.

is generated by observing a number of individual, which is repeating under different experimental conditions. Here individuals are considered as a random sample from the population of interest. Thus the model based on earthquake data to compute ground motion parameters can also be regarded as mixed model, Abrahamson and Youngs (1992). General form of ground motion prediction model as mixed model can be represented as:

$$logY_{ij} = f(m_i, r_{ij}, \theta) + \eta_i + \varepsilon_{ij} \tag{3.12}$$

where Y_{ij} is ground motion parameter, $f(m,r,\theta)$ is the ground motion prediction equation. m is the magnitude, r is the distance, θ is the vector of model parameters , η_i is the random effect for i^{th} earthquake and ε_{ij} is the error for j^{th} recording from the i^{th} earthquake and it is assumed to be normally distributed. In other words η_i represents inter(between)-event variations and ε_{ij} represent intra(within)-event variations. The inter-events residuals represent the average shift of the observed ground motion from an individual earthquake, from the population median predicted by the ground-motion model. The intra-event residuals represent the misfit between an individual observation at a station from the earthquake-specific median prediction, which is defined as median prediction of the model plus the inter-event term of earthquake (see figure 3.6). Moreover the inter-event and intra-event standard deviations of ground motion model represents earthquake-to-earthquake variability and record-to-record variability, respectively. The inter-event and intra-event residuals are uncorrelated, so the total standard deviation is defined as:

$$\sigma_{total} = \sqrt{\sigma^2_{inter-evet} + \sigma^2_{intra-event}} \tag{3.13}$$

Further the inter-event residuals represents the average source effects and also reflects various factors like stress drop, variation of slip in time and space which are not captured by including magnitude, focal mechanism and source depth. Intra-event residual represents path and site effects (crustal heterogeneity, geological structure and near surface layering) which modify the waveform (in terms of amplitude and frequency) and are not captured by distance metric and site classification on the basis of average shear wave velocity.

Thus we performed the NLMRA to solve equation 3.11 and to compare the results with those of LRA and NLRA (see figure 3.5). The regression

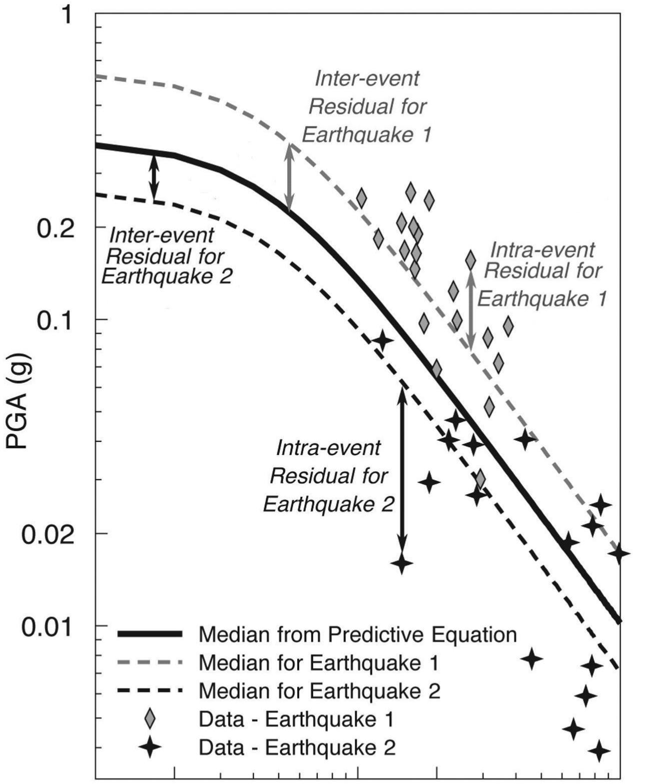


Figure 3.6.: Inter(between)-events and intra(within)-events components of ground motion variability (after Strasser et al. (2009)).

coefficients with uncertainties are listed in table 3.6. The predictions of NLMRA seems a little bit higher at short distances (black curve). However, there is not significant change in standard deviation and R^2 statistics. It was important to perform statistical tests on mean values and standard deviations to check whether their predictions (LRA, NLRA, and NLMRA) are statistically different or not. Thus T-test was done on mean values and F-test on total standard deviations and we found that the null hypothesis was failed to reject that means and standard deviations are equal. Hence the standard deviations and means are statistically equal. Thus all the models have same predictions. But NLMRA has one advantage that it splits total uncertainties into inter-event and intra-event uncertainties which proved useful to study the source effect and site effects on ground-motion recordings separately. Hence from now onwards we will concentrate will NLMRA. Thus we adopt the NLMRA technique to retrieve the coefficients of the GMPEs.

Table 3.6.: Regression coefficients and uncirtainties for equation 3.11 with NLMRA.

			NLMRA					
Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$\sigma_{inter-event}$	$\sigma_{inter-event}$	σ_{total}	R^2
PGA (m/s^2)	-2.6657 ± 0.0717	1.1583 ± 0.0204	-2.3124 ± 0.0552	1.7324 ± 0.1972	0.1420	0.38581	0.3849	0.7862
			NLRA					
PGA (m/s^2)	-2.6347 ± 0.0659	1.1605 ± 0.0101	-2.3415 ± 0.0622	1.9752 ± 0.2105			0.3832	0.7801
			LRA					
PGA (m/s^2)	-2.6278 ± 0.0319	1.1606 ± 0.0101	-2.3480 ± 0.0280	2.0			0.3832	0.7801

4. Predicting ground motion and seismic hazard

4.1. Ground motion

Ground motion is the movement of the earth's surface from earthquakes or explosions. Ground motion is produced by waves that are generated by sudden slip on a fault or sudden pressure at the explosive source and travel through the earth and along its surface. In this chapter we will discuss the important factors which affect the ground motion recordings and their applicability in predicting seismic hazard. The ground motion recordings depend on the source, path, site and the instrument as shown in figure 4.1.

The recorded ground motion U(t) can be expressed as convolution of Source *A(t)*, path *P(t),* site *S(t)* and Instrument *I(t)* mentioned in equation 4.1

$$U(t) = A(t) * P(t) * S(t) * I(t) \tag{4.1}$$

A(t) describes the effect of the size and the nature of the earthquake, *P(t)* describes the effect of the medium on the seismic wave traveling from source to site, *S(t)* describes the effect of upper layer of soil and geology of the recording site, *I(t)* describes the effect of recording instruments.

The seismic waves and the resulting ground motions are divided into two general levels; weak motion, or small amplitude motion from the distant or small earthquakes and strong motion or large amplitude motion from nearby or large earthquake. In seismic hazard analysis, estimates of ground motion from natural earthquakes are defined in the context of strong motion. This kind of motion is of interest to engineers and those concerned with the effects of earthquakes on the society. Various factors which affect ground motions are discussed in next section.

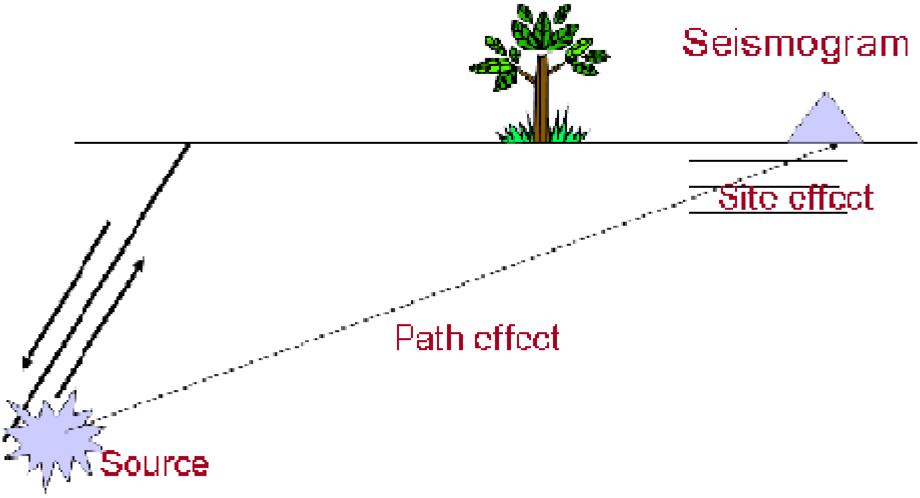


Figure 4.1.: The sketch showing factors affection the ground motion recordings.

4.1.1. Source factor

The earthquake source types effecting ground motion have been already been defined in chapter 2. Apart from source classification some of important factors affecting the ground motion will be discussed in this section.

Magnitude

Magnitude is an empirical measurement of the size of an earthquake which is approximately related to the energy released during the earthquake. The ground shaking depends on the magnitude (energy released) if the distance is constant for the earthquake of different magnitudes. Magnitude can be measured in different scales local magnitude $\mathrm{M_L}$(calculated by using S or surface waves), body wave magnitude $\mathrm{m_b}$(calculated by using P waves), surface wave magnitude $\mathrm{M_s}$ (calculated by using Rayleigh waves), moment magnitude M_w (calculated by using seismic moment). Most of the studies consider M_w accurate because it is related to the rupture area and slip of the fault which strongly affect the ground motion recordings, see Hanks and Kanamori (1979). The relationship between seismic moment and magnitude is reported in equation 4.2

$$Mw = 2/3 * log(SeismicMoment) - 10.73 \tag{4.2}$$

where the units of the moment are in dyne-cm.

Focal mechanism

The focal mechanism of an earthquake describes the inelastic deformation in the source region that generates the seismic waves (Aki and Richards (1980)). In the case of a fault-related event it refers to the orientation of the fault plane that slipped and the slip vector and is also known as a fault-plane solution. Focal mechanisms are derived from a solution of the moment tensor for the earthquake, which itself is estimated by an analysis of observed seismic waveforms. The focal mechanism affects the pattern of "first motions", that is, whether the first arriving P-waves break up or down. For example in case of reverse or thrust faulting where the compressional stress is maximum, highest ground motion appears parallel to the earth's surface and the crust is in state of compression. In case of normal faulting ground motion is considered to be lower and the crust is said to be in state of tension, with strike slip faulting, where maximum and minimum compressional stress are parallel to the earth's surface the ground motion is lower. Examples of high ground motion associated with reverse and thrust faulting include the peak accelerations greater that 1.0g associated with the 1971 San Fernando, California and the 1985 Nahanni, Northwest Territories, Canada earthquakes. The reason of increased ground motion during reverse or thrust faulting may be due to the increase in the strength of the crust (resistance of fault movement) when it is in a state of compression resulting in higher localized stress drop (McGarr (1984)).

Radiation pattern

Radiation Pattern is the azimuthal distribution of amplitude of seismic waves. Each wave has its own type of radian pattern. For example S-wave radiation pattern from a strike slip fault is opposite (rotated by 45° in azimuth) to that of P-wave. The strongest S-wave are generated in the direction which is parallel and perpendicular to fault plane.

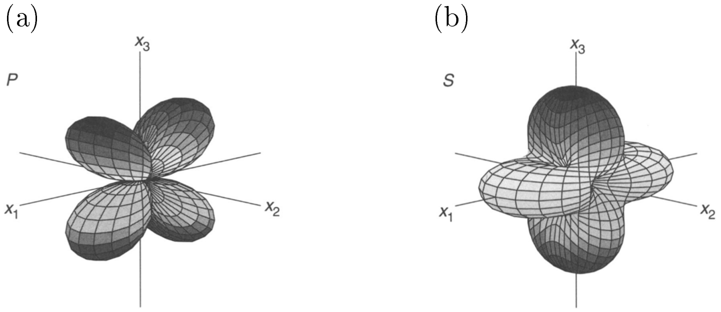


Figure 4.2.: Radiation Pattern P-wave (a) and S-wave(b) for a strike slip fault.

In figure 4.2 it can be observed that SH-wave are strongest in direction where P-waves are weakest for strike-slip fault. The radiation patterns are three dimensional so that fault geometry at depth and the location of interest with respect to that geometry affects ground motion. For example, a site several kilometers away from the earthquake with strike slip focal mechanism would experience higher ground motion, if the fault rupture surface directed toward the site than it would if the fault rupture were perpendicular (Anderson and Luco (1983)). It is not only because the site was closer to fault rupture but also because of different vertical angle to that surface.

Directivity

Earthquake directivity is the focusing of wave energy along the fault in the direction of rupture. When a fault ruptures, unilaterally (with the epicenter at or near one end of the fault break), the radiated waves are stronger in the direction of propagation of rupture. Directivity is caused by the constructive interference (piling up) of the S-waves, due to the rupture propagation.

To understand directivity,see figure 4.3, the fault rupture moves from right to left (i.e. from B to A) from the epicenter. In this case the ground motion is considered as a pulse, radiates outward all the direction. The pulse which started earlier will spread the farthest, while the pulse started latest will spread least. As they have generated at different times, the

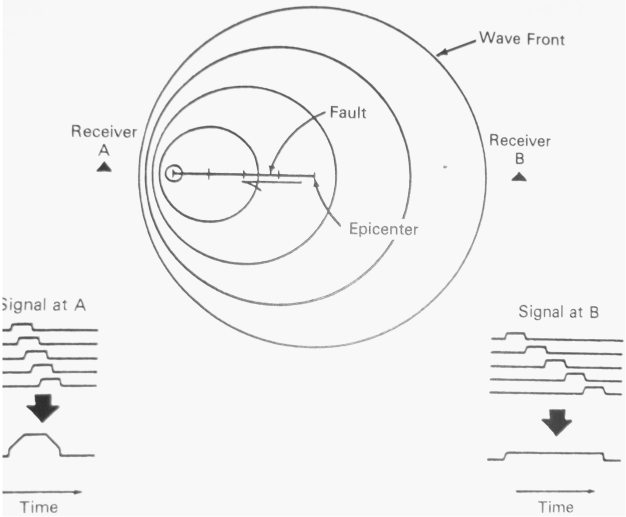


Figure 4.3.: A sketch showing the effect of directivity(after Del Mar Technical Associates 1979).

pulses tend to arrive close to the same time at receiver A (in the direction of rupture propagation) because the pulses that started the latest have the least distance to travel. At receiver B (opposite to the direction of rupture propagation) the pulse started latest have more distance to travel, so that the pulses generated from different parts of rupture tend to arrive spread out over the time. The constructive interference at site A will result in ground motion amplification. Thus in many cases the distance to the fault is not the only consideration for ground motion amplitude but direction is also important.

4.1.2. Path effect factor

The effects of travel path on ground motion are primarily related to attenuation of seismic waves. Attenuation is defined as loss of energy (wave amplitude) with distance. Two main responsible elements of attenuation are discussed below.

Geometrical spreading

It is the result of law of conservation of energy. As the wave front propagate the surface area of the wavefront increases with distance but the energy all over the area must be constant. Thus at larger distances the energy is lower that at short distance. Hence, with the ever increase size of wavefront the energy per unit area falls down. If the earth is considered as homogeneous and isotropic medium, body waves would have spherical wavefront and their amplitude would decrease as $1/R$. Similarly, if earth is considered as uniformly flat layered, surface wave would have cylindrical wave fronts and their amplitudes would decrease $1/\sqrt{R}$, where R is the source-to-site distance.

Absorption (Damping)

Absorption indicates the attenuation due to medium properties, homogeneous or heterogeneous, isotropic or anisotropic. Generally, it is controlled by sliding friction across the cracks, internal friction and grain boundary effects which occur when seismic wave pass through rock (medium). Even temperature play important role, when the rock is hot the absorption is higher. The destructive interference due to scattering of waves in the medium reduce the energy of propagating wave.

Knopoff (1964) defined the term quality factor,Q to quantify the lose of energy of propagating wave in medium. The definition of Q is reported in equation 4.3

$$\frac{1}{Q} = \frac{-\triangle E}{2\Pi E} \tag{4.3}$$

where ΔE is the loss of elastic energy per unit cycle and E is the peak energy in that cycle. Q varies with wave type and also depends on rock properties. The larger the Q value, smaller is the attenuation and vice-versa. The Q also depend on the frequency of the wave, in equation 4.4 it is defined by τ which quantifies the rate of lose of energy with distance,

$$\tau = \frac{\Pi f}{QV} \tag{4.4}$$

where f is frequency and V is seismic wave velocity. Equation 4.4 indicates that the increase in attenuation is directly proportional to frequency.

Nutti (1981) utilized a simple frequency-dependent relationship, expressed in equation 4.5

$$Q = Q_0 f^n \tag{4.5}$$

where Q_0 is Q at 1 Hz. and n varies from 0 to 1. Q, Q_0 and n are functions of medium properties and can vary from material to material (region to region).

4.1.3. Site effect factor

Apart from the source and propagation medium the local site/geological condition plays an important role on the ground motion recordings. Some important factors are discussed below.

Impedance contrast

Impedance contrast is one of the most important reason of increase or decrease in the amplitude of ground motion at the surface. Aki and Richards (1980) point out that impedance is the resistance to particle motion. For horizontally polarized shear wave (SH) it can be defined as the product of density (ρ), the shear wave velocity (β) and the cosine of the angle of incidence (i). (see equation 4.6)

$$Z = \beta \times \rho \times cos(i) \tag{4.6}$$

The angle of incidence, is usually small near earth's surface (according to Snell's law) and its cosine can be assumed equal to one. The particle velocity is inversely proportional to the square root of the impedance. As the seismic wave passes through the region of increasing impedance, the resistance to the motion increases and to preserve energy, the particle velocity and therefore the amplitude of the seismic wave will decrease or vice-versa.

Topographic effect

When the seismic wave encounters sharp change in impedance some of the energy is lost in transmission and other in reflection. And some times these reflected and transmitted waves get trapped with in the bed and produce reverberations. This effect is maximum when reverberating wave are in phase with each other. This is also termed as resonance and is also frequency dependent. In the simplest case the maximum reverberation occurs when the wavelength of the waves are four times the thickness of the layer in which waves are trapped. In other words, for shear waves the frequency which is amplified the most is that which is equal to $\beta/4h$ where the β is shear wave velocity and h is the thickness of the rock bed in which the wave is trapped (Reiter (1990)).

Nonlinear response of soil

Nonlinearity is that phenomenon which is related to the changes in soil properties and therefore changes in soil response as the level of ground motion changes (increase or decrease). Nonlinearity of soil response is also called strain dependence, because the strain the soil goes through during an earthquake increases with the level of stress or ground motion. The strains generated by earthquakes can be large in near field, which can be of the order of 10^{-4} to 10^{-3}, for example see Berrill (1975). These strains reduce the shear modulus (and consequently the shear-wave velocity) and increase the damping ratio (e.g. Seed and Idriss (1969); Hardin and Drnevich (1972)). Such non-linear behavior of soils are capable of explaining the difference between expected high amplifications due to site effects and the lower amplifications recorded in large earthquakes, Seed and Idriss (1969). Thus there are several factors affect the ground motion recordings. Hence it is important to consider all these factor while developing a ground motion prediction equation. In next section we will discuss the case of Geysers geothermal area which is vapor dominated area and induced seismicity is involved. we will see how the above discussed factor (source, medium and site effects) play a key role in ground motion recordings in the Geysers geothermal area.

4.2. Ground motion prediction equations (GMPEs) in "The Geysers" geothermal area.

Some of the common regression techniques have be discussed chapter in 3. Here we will focus on the NLMRA (Lindstrom and Bates (1990) and

Abrahamson and Youngs (1992)) as we know that it includes both fixed effect and random effects and it permits to account for both inter-event and intra-event dependencies on data. We have adopted the technique proposed by Emolo et al. (2011) to introduce a first-order site/station effect correction. First, a model is retrieved, which does not include station effects and represent the reference model. Second, a first-order correction is introduced for station effects based on the Z-test of the residual distribution at each station with respect to reference model. This corrected model is the final model. Details are discussed in next sections.

4.2.1. Data description and processing

To do any kind of analysis the most important part is to create database (here is induced seismicity records). As mentioned in chapter 1, different temporary and long-term seismic networks have been deployed in The Geysers geothermal field during the last five decades. At present, local seismicity is monitored by the dense Lawrence Berkeley National Laboratory Geysers/Calpine (BG) surface seismic network and with some nearby stations of the Northern California Seismic Network (NCSN). The BG network consists of 32 three-component stations, 29 of which were used for the present study (black triangles figure 4.4).

The BG stations are distributed in an area of about 20x10 km^2, covering the entire geothermal field. Each station was equipped with I/O Sensor SM-6 geophones with a natural frequency of 14 Hz. In Fall 2009 these instruments were replaced successively by Oyo GS-11D 4.5 Hz sensors. The BG stations operate in trigger mode, and the waveform segments recorded since the end of July 2007 are made available at the Northern California Earthquake Data center (NCEDC). We retrieved waveforms for the period 1 September 2007 through 15 November 2010, associated the traces with the events from the NCEDC catalog, and prepared a subset of the data for regression analysis. The selected waveforms, correspond to 212 earthquakes (dots in figure 4.4) with a focal depth of less than 5 km. This depth has been selected considering that, the earthquakes observed at depths larger than 5 km are natural earthquakes. The hypocentral distance ranges from 0.5 km to 20 km, and the magnitude range is $1.3 \leq M_W \leq 3.3$. For the largest portion of earthquakes analyzed in this study, the NCEDC catalog provides a duration magnitude $\mathrm{M_D}$ as magnitude mea-

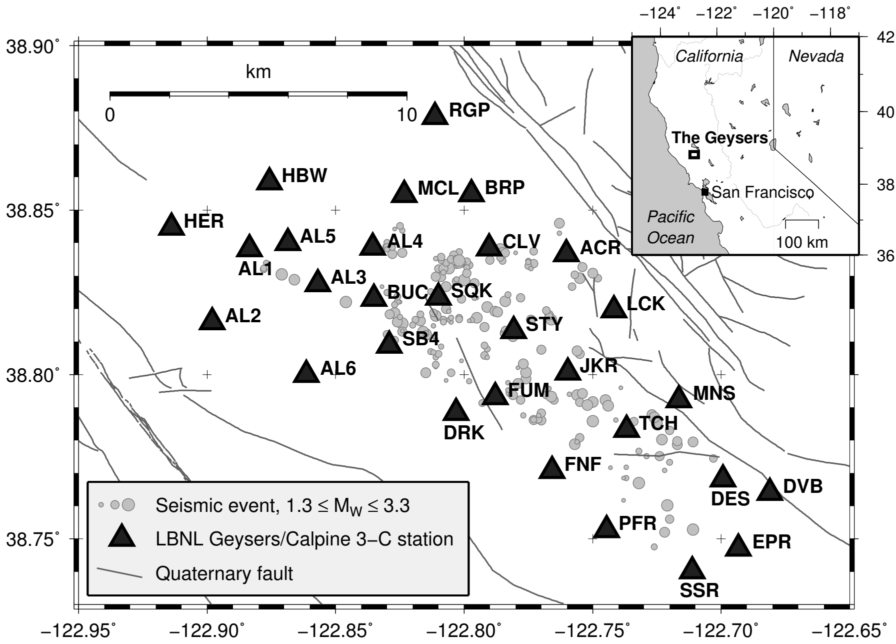


Figure 4.4.: Location of the events (gray circles) and the seismic stations (black triangles) used for regression analysis in the present study.

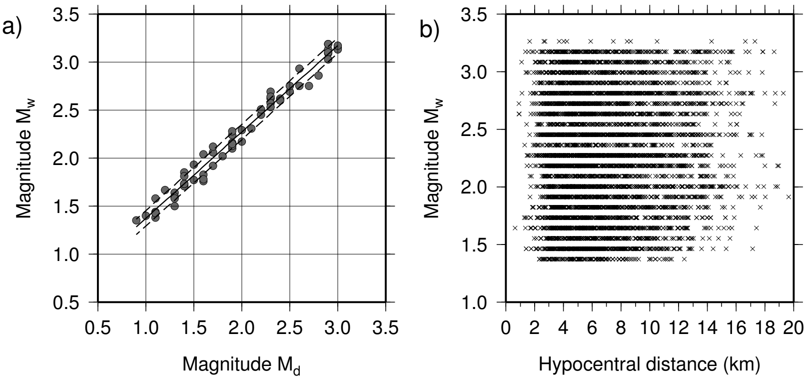


Figure 4.5.: (a) Regression model (solid line) for computing moment magnitude, M_w, from the duration magnitude $\mathrm{M_D}$ with standard deviation of 0.08 are plotted along (dashed lines) and (b) the scatter plot showing the hypocentral distance and magnitude range used for regression analysis.

sure. However, in order to obtain results compatible with other studies and suitable for seismic hazard analysis purposes, we converted $\mathrm{M_D}$ into moment magnitudes $M_w$using a linear relationship retrieved by using M_w data provided by Douglas et al. (2013) is mentioned in equation 4.7 below.

$$M_w = 0.473(\pm 0.035) + 0.900(\pm 0.017)M_D \tag{4.7}$$

The data used and the least-squares fitting with standard deviation (0.08) are shown in figure 4.5a, and figure 4.5b is showing the scatter plot of the hypocentral distance and magnitude used for the present study.

To analyze the best quality data, waveforms with signal-to-noise ratio larger than 10 in the whole analyzed frequency range, which is 0.5 Hz to 35 Hz are selected. Instrument correction to the waveforms within same frequency band is applied. The mean and the trend are also removed. Then zero phase shift and 4 poles butter-worth filter in the frequency band 0.7 Hz and 35 Hz was applied. Most of the processing is done using SAC (Seismic Analysis Code). Further, to measure the correct values of the selected ground-motion parameters we have cut the waveforms in a specific time-window around the event, starting at the origin time and ending at the time corresponding to 98% of total energy contained in waveform which are also tapered with a 0.1 taper width with cosine window. Once the time window has been selected, PGV is measured as the largest value among the two horizontal components. The waveforms are then differentiated and filtered in a frequency band ranging between 0.7 Hz and 35 Hz (to avoid the high frequency noise added while differentiation) to measure PGA and the 5% damped spectral ordinates SA(T) at T = 0.2s, 0.5s and 1s, which, as for PGV, correspond to the largest value among two horizontal components. An example of processed waveforms are shown in figure 4.6.

4.2.2. Regression analysis using NLMRA

After obtaining the peak ground motion parameters, different GMPEs were tested. In particular, the first model (MOD1) has the following formulation:

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) \tag{4.8}$$

where Y represents either PGV expressed in m/s, PGA or SA(T) at T = 0.2s, 0.5s and 1.0s, both expressed in m/s^2, respectively. The model in equation 4.8 accounts for the source effect through the moment magnitude M_w and geometrical spreading through the hypocentral distance R_{hypo} (expressed in km). The h parameter (in km) is introduced to avoid unrealistic high values at short distances (e.g., Joyner and Boore (1981)) (see chapter 3). The inferred coefficients along with the uncertainties are listed in Table 4.1.

We are analyzing data collected in a geothermal area which is continuously perturbed through fluid injection and, as a consequence, can be highly fractured. Thus it is physically reasonable to expect that an-elastic attenuation could also play a key role. Therefore, we introduce an additional linear term depending on the distance, which leads to the model defined by following equation 4.9 (MOD2):

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) + dR_{hypo} \tag{4.9}$$

Using the same dataset as the one used to retrieve MOD1, the NLMRA is performed. The obtained coefficients along and their uncertainties are listed in table 4.2.

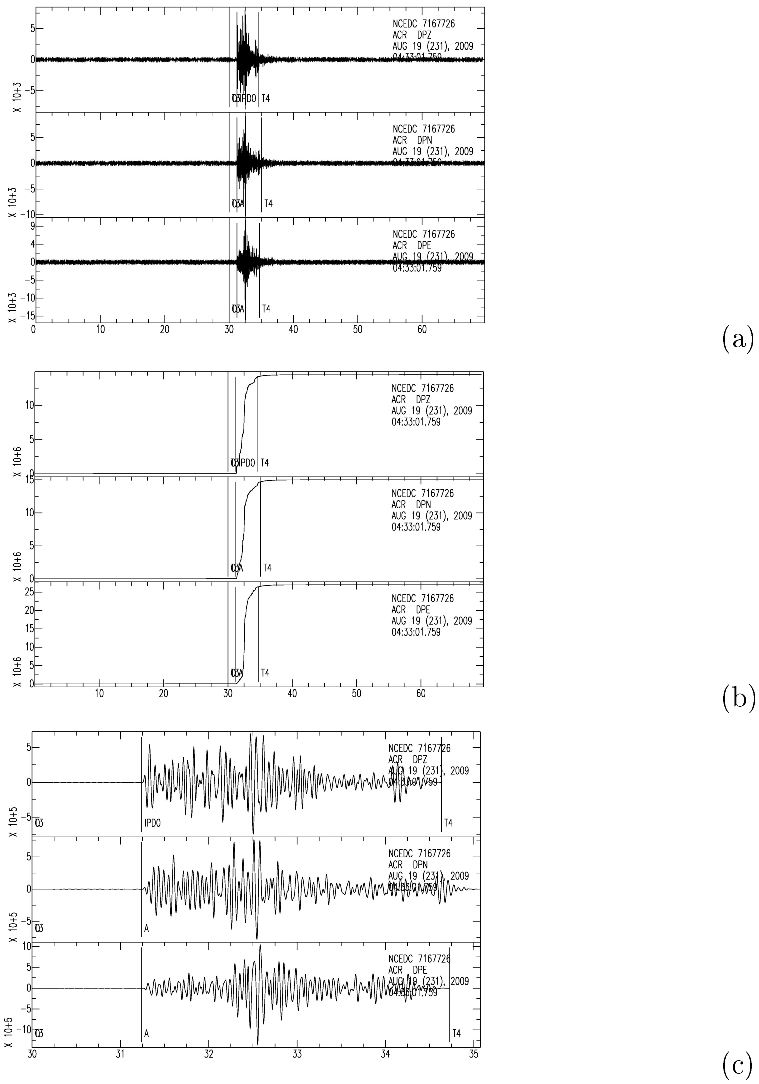


Figure 4.6.: The three component original waveforms (a), curve used to point mark between 0 to 98% of energy contained in waveform (b), cut and instrument corrected filtered (0.7 to 35 Hz) waveforms to measure appropriate peak ground motion parameters (c).

Table 4.1.: Regression Coefficients and Relative Uncertainty of Equation 4.8, MOD1

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{Total}	R^2
PGV (m/s)	-4.960 ± 0.074	1.317 ± 0.055	-2.101 ± 0.055	1.850 ± 0.212	0.154	0.344	0.377	0.814
PGA (m/s^2)	-2.666 ± 0.072	1.158 ± 0.020	-2.312 ± 0.055	1.734 ± 0.197	0.142	0.358	0.385	0.786
SA(T=0.2s)(m/s^2)	-3.648 ± 0.087	1.443 ± 0.022	-1.871 ± 0.068	2.393 ± 0.270	0.160	0.356	0.390	0.815
SA(T=0.5s)(m/s^2)	-4.926 ± 0.081	1.552 ± 0.023	-1.688 ± 0.061	2.210 ± 0.275	0.164	0.339	0.377	0.840
SA(T=1.0s)(m/s^2)	-5.341 ± 0.076	1.503 ± 0.022	-1.861 ± 0.056	1.959 ± 0.240	0.157	0.339	0.374	0.840

Table 4.2.: Regression Coefficients and Relative Uncertainty of Equation 4.9, MOD2

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$d \pm \sigma_d$	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{Total}	R^2
PGV (m/s)	-4.585 ± 0.255	1.317 ± 0.021	-2.730 ± 0.390	2.406±0.358	0.027 ± 0.015	0.154	0.344	0.377	0.814
PGA (m/s^2)	-2.307 ± 0.255	1.158 ± 0.020	-2.925 ± 0.376	2.243±0.332	0.026 ± 0.015	0.142	0.358	0.385	0.786
SA(T=0.2s)(m/s^2)	-3.370 ± 0.242	1.442 ± 0.022	-2.309 ± 0.484	2.821±0.492	0.017 ± 0.018	0.160	0.356	0.390	0.815
SA(T=0.5s)(m/s^2)	-4.950 ± 0.327	1.552 ± 0.023	-1.648 ± 0.343	2.158±0.545	0.002 ± 0.014	0.164	0.339	0.377	0.840
SA(T=1.0s)(m/s^2)	-5.158 ± 0.237	1.503 ± 0.022	-2.167 ± 0.364	2.285±0.430	0.013 ± 0.014	0.157	0.339	0.374	0.840

Although we verified through an F-test that the coefficient d in equation 4.9 is statistically significant, looking at the R^2-value and at the total standard deviation, it does not improve the global fitting with respect to MOD1. Hence, in accordance with the Occam's razor or lex parsimoniae (e.g., MacKay (2003)) it is decided to choose the MOD1 rather than MOD2 as the best reference model, as it is described by a less number of parameters.

The qualitative fit of MOD1 as function of hypocentral distance is shown in figure 4.7 for PGV, PGA and SA(T) at T = at 0.2s , 0.5s and 1.0s and for the magnitudes M_w 1.0, 2.0 and 3.0, together with the inter-event and intra-event residuals of the models on entire data set.

4.2.3. Residual analysis and estimation of station effect factor and corrected model

As mentioned in the introduction, we have adopted the technique proposed by Emolo et al. (2011) to introduce a first-order site/station effect correction. It should be noted that the correction must be intended in a broader sense with respect to the standard site-effect because it is not based on the Vs30 values which are not available for the stations used in present study. Moreover, when the Vs30 database obtained from the topography (Wald and Allen (2007)) is investigated (see Data and resources section) The Geysers area falls in C-category soil type of National Earthquake Hazard Reduction Program (NEHRP) classification (Building Seismic Safety Council, 2000 and 2004).

To overcome these problems, we analyzed the residuals distribution at each station obtained using the reference model MOD1 as shown in figures 4.8, 4.9, 4.10, 4.11, 4.12 for PGV, PGA and SA(T) at T = 0.2s, 0.5s and 1.0s, respectively. In particular at each station in these figures, *a* panel shows residual distributions, while *b* panel shows residual as a function of hypocentral distance at each station. The value of slope of the fits is also mentioned in *b* panel. Through Z-test, we tested the null hypothesis of a Gaussian zero-mean distribution at 95% level of confidence. We assume that a deviation from the expected zero-mean value can be reasonably ascribed to a site/station effect which can be corrected by using dummy correction factor. Indeed, due to the assumed definition of the residual ($\log_{10}\mathrm{Y}^{\mathrm{Observed}} - \log_{10}\mathrm{Y}^{\mathrm{Predicted}}$) a positive deviation from the zero-mean

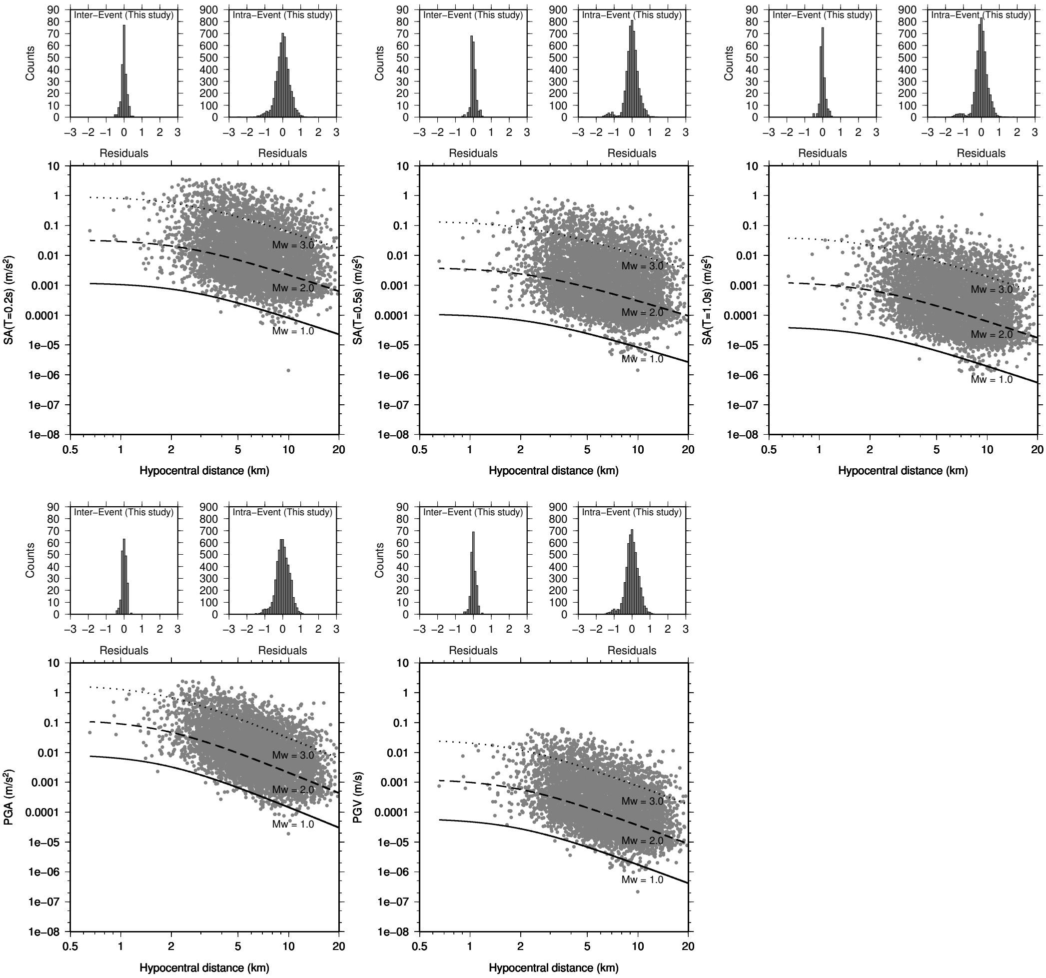


Figure 4.7.: The regression model for magnitude 1.0 (solid line), 2.0 (dashed line) and magnitude 3.0 (dotted line) w.r.t hypocentral distance. The gray dots represent entire data set. Upper panels shows respective inter-event and intra-event residual distribution of respective ground motion parameters PGV, PGA, SA at 0.2s, 0.5s and 1.0s.

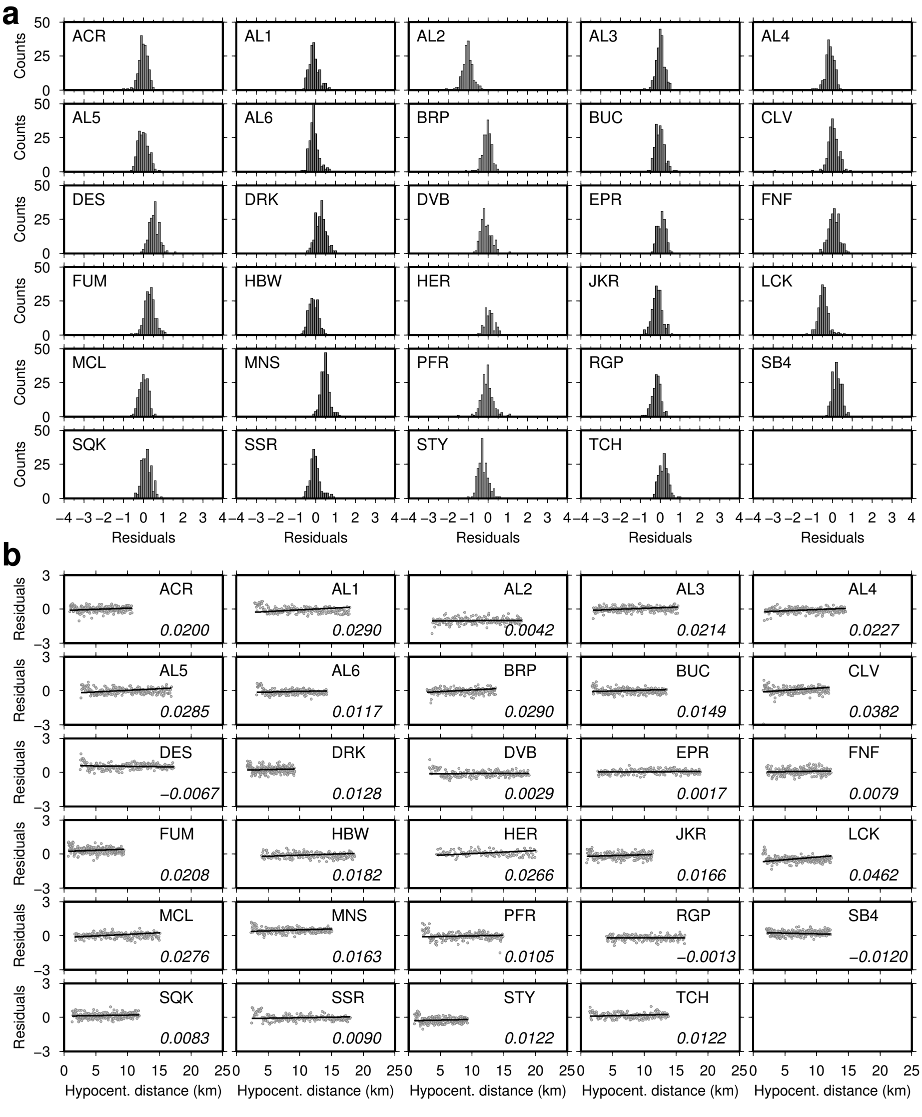


Figure 4.8.: Single station residuals distribution. (a) shows residual histograms and (b) refers to the distribution of residuals as a function of hypocentral distance for model MOD1 (i.e., the reference model), at each station for PGV. The slope of linear fit to residuals at each station is also mentioned.

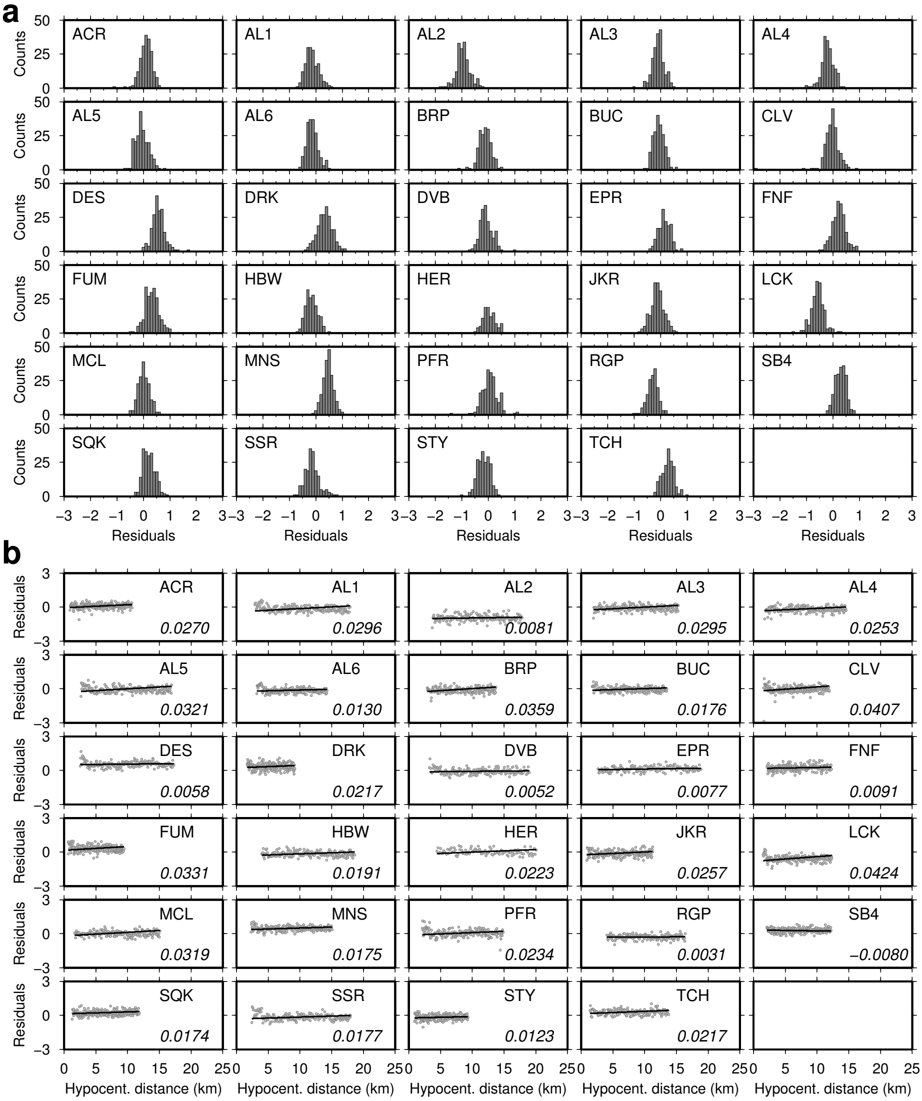


Figure 4.9.: Same as figure 4.8 but for PGA.

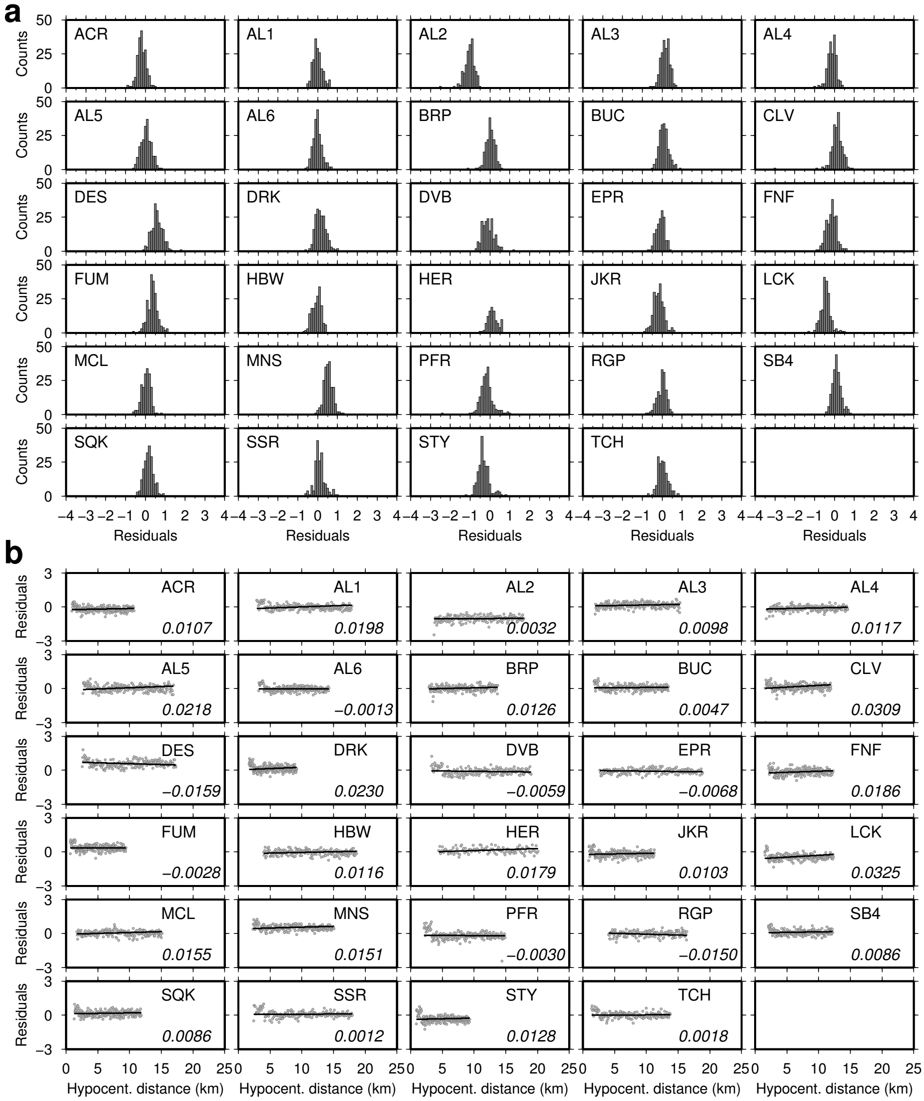


Figure 4.10.: Same as figure 4.8 but for SA(T) at T = 0.2s.

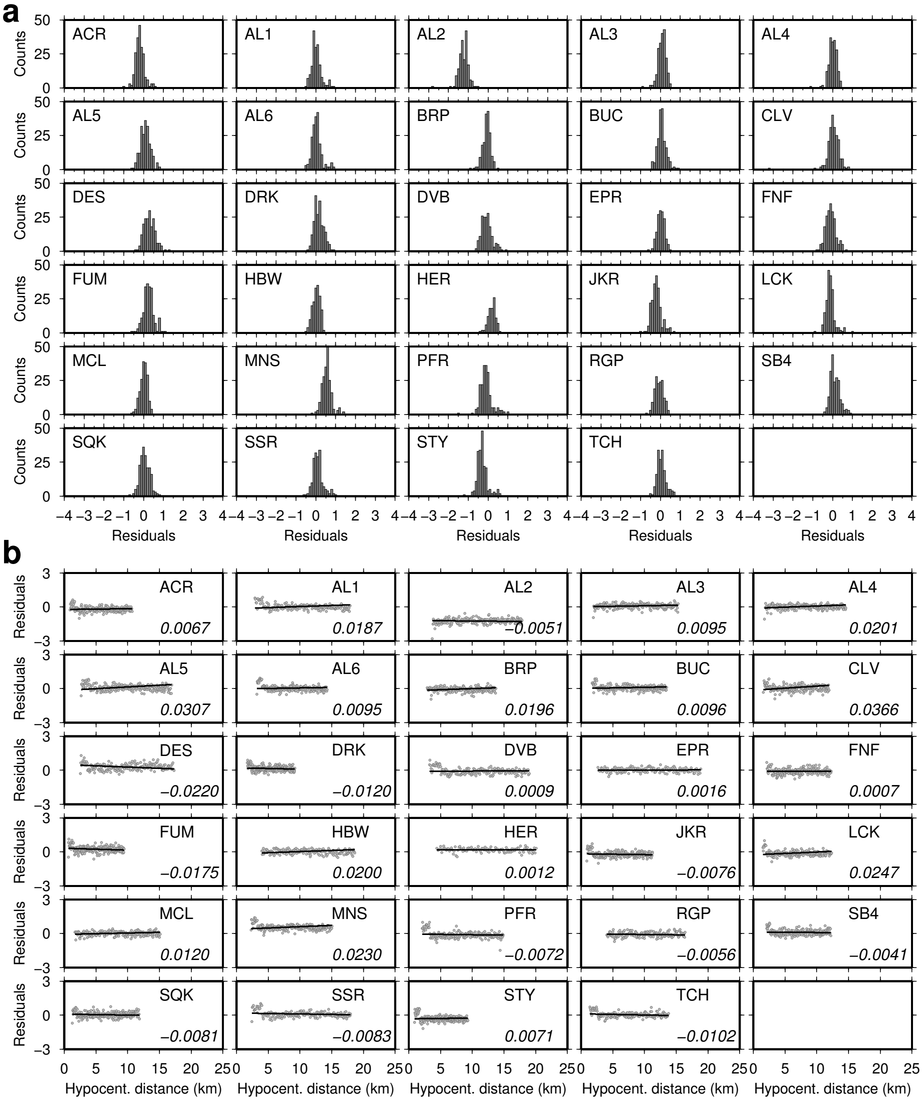


Figure 4.11.: Same as figure 4.8 but for SA(T) at T = 0.5s.

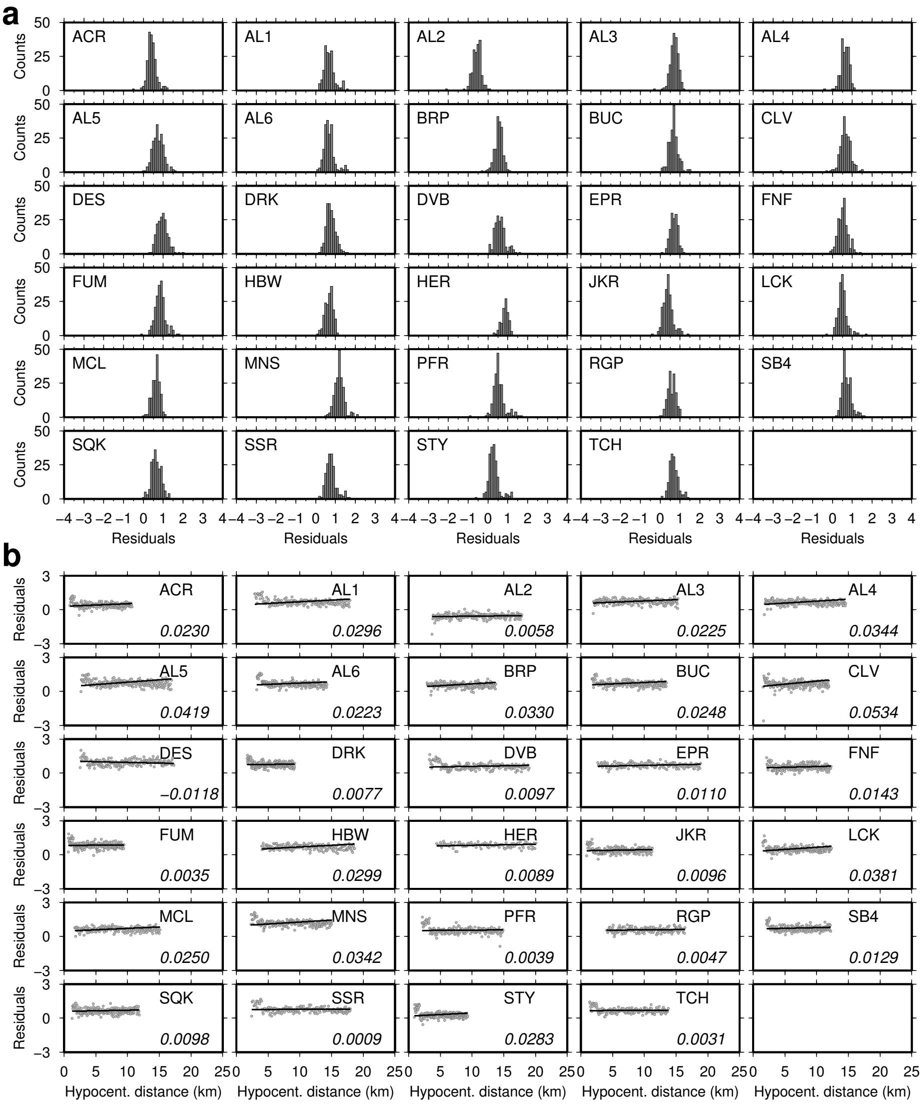


Figure 4.12.: Same as figure 4.8 but for SA(T) at T = 1.0s.

value can be interpreted as an underestimation of the model prediction with respect to the observations, while a negative deviation is to be interpreted as an overestimation of the model prediction. Thus by Z-test in terms of both value and sign, at each station we assigned a dummy variable s whose value is -1,0, or $+1$ depending on the mean residual value negative, zero or positive which allows to recover the observed residual deviation from expected zero mean value. Specifically, we assign a 0 coefficient at all the stations with zero-mean residual (predictions are on average equal to the observed), a $+1$ coefficient at stations having a positive deviation from zero (mean residuals value is greater than zero, that is, predictions underestimate the observations), and a -1 coefficient at stations having a negative deviation from zero (mean residuals value is less than zero, that is, predictions overestimate the observations). The addition of this new parameter in the reference model provides the corrected regression model. The value of s parameter with station code and coordinates, and number of observations ($\mathrm{N_{obs}}$) are listed in table 4.3 for PGV, PGA and SA(T) at T = 0.2s, 0.5s and 1.0s, respectively.

The retrieved s parameter along with a new coefficient e are then used to set up an updated model, that is, the corrected model (hereafter named MOD3), which is formulated as:

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) + es \tag{4.10}$$

The inferred coefficients and their uncertainties are listed in table 4.4.

It can be noted that, after considering the site/station effect, the total standard deviation is reduced and the residuals distribution at each station is improved as their maxima are centered on zero (see figures 4.13, 4.14, 4.15, 4.16, 4.17 for PGV, PGA and SA(T) at T = 0.2s, 0.5s and 1.0s, respectively).

It is also worth noticing that there is a considerable improvement in intra-event standard deviation while there is no significant change in inter-event standard deviation. This could be ascribed to the fact that intra-event residuals are associated with all those factors that are not captured by the distance metric and site classification, Atik et al. (2010) hence confirming the effectiveness of the technique proposed to site/station correction. Furthermore in order to test the effectiveness of the formulation of obtained model, we have also analyzed the distribution of residuals as

Table 4.3.: Site correction parameters for PGV, PGA and SA at 0.2s, 0.5s and 1.0s.

Station	Name	Latitude(°)	Longitude(°)	N_{obs}	S(PGV)	S(PGA)	S(SA@0.2s)	S(SA@0.5s)	S(SA@1.0s)
1	ACR	38.837	-122.76	202	0	1	-1	-1	-1
2	AL1	38.838	-122.883	179	-1	-1	0	1	-1
3	AL2	38.816	-122.898	180	-1	-1	-1	-1	-1
4	AL3	38.828	-122.857	201	1	-1	1	1	1
5	AL4	38.839	-122.836	184	-1	-1	-1	0	-1
6	AL5	38.84	-122.869	208	-1	-1	1	1	1
7	AL6	38.8	-122.861	199	-1	-1	-1	1	0
8	BRP	38.855	-122.797	181	-1	-1	1	-1	-1
9	BUC	38.823	-122.835	197	-1	-1	1	1	1
10	CLV	38.838	-122.79	209	1	-1	1	1	1
11	DES	38.768	-122.699	191	1	1	1	1	1
12	DRK	38.788	-122.803	209	1	1	1	1	1
13	DVB	38.764	-122.681	173	-1	-1	-1	-1	-1
14	EPR	38.747	-122.693	151	1	1	-1	1	1
15	FNF	38.771	-122.766	204	1	1	-1	-1	-1
16	FUM	38.793	-122.788	211	1	1	1	1	1
17	HBW	38.859	-122.876	172	-1	-1	-1	1	1
18	HER	38.845	-122.914	104	1	1	1	1	1
19	JKR	38.801	-122.76	209	-1	-1	-1	-1	-1
20	LCK	38.82	-122.742	191	-1	-1	-1	-1	0
21	MCL	38.855	-122.823	182	1	1	1	1	1
22	MNS	38.792	-122.717	206	1	1	1	1	1
23	PFR	38.753	-122.745	199	-1	1	-1	-1	-1
24	RGP	38.878	-122.811	157	-1	-1	-1	-1	-1
25	SB4	38.809	-122.829	202	1	1	1	1	1
26	SQK	38.823	-122.81	203	1	1	1	1	1
27	SSR	38.74	-122.711	180	-1	-1	1	1	1
28	STY	38.813	-122.781	200	-1	-1	-1	-1	-1
29	TCH	38.783	-122.737	167	1	1	1	1	1

Table 4.4.: Regression coefficients and relative uncertainty of equation 4.10, using MOD 3.

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$e \pm \sigma_e$	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{Total}	R^2
PGV (m/s)	-5.065 ± 0.069	1.320±0.022	-1.966 ± 0.047	1.863±0.191	0.189±0.004	0.164	0.289	0.332	0.857
PGA (m/s^2)	-2.710 ± 0.064	1.165±0.021	-2.244 ± 0.044	1.779±0.158	0.225±0.004	0.151	0.276	0.315	0.859
SA(T=0.2s)(m/s^2)	-3.721 ± 0.082	1.448±0.023	-1.802 ± 0.061	2.629±0.244	0.203±0.004	0.169	0.294	0.339	0.862
SA(T=0.5s)(m/s^2)	-4.833 ± 0.083	1.555±0.024	-1.838 ± 0.061	2.674±0.239	0.182±0.004	0.175	0.290	0.339	0.872
SA(T=1.0s)(m/s^2)	-5.314 ± 0.077	1.506±0.023	-1.918 ± 0.057	2.255±0.216	0.166±0.004	0.168	0.299	0.342	0.867

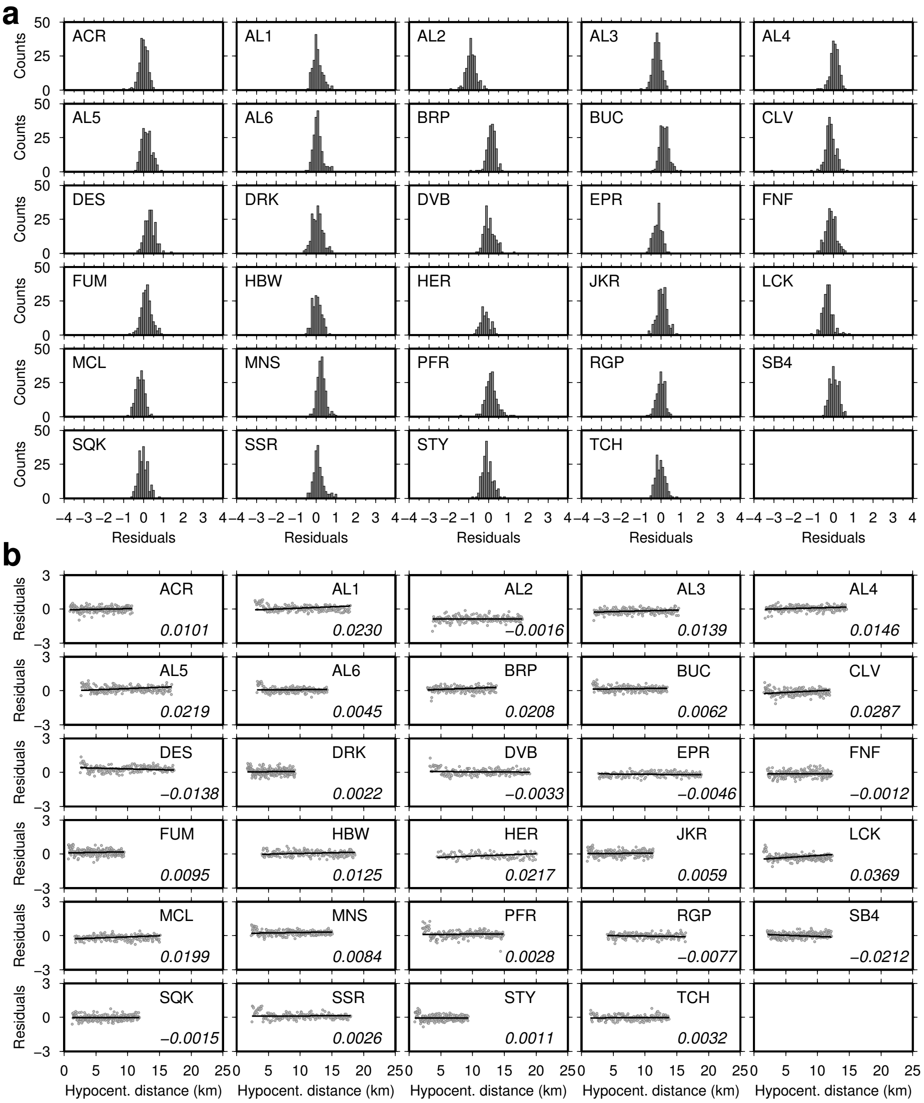


Figure 4.13.: Same as figure 4.8 but for model MOD3(corrected model) PGV. There is an improvement in the model after site correction as now the maxima of residual are centered at zero.

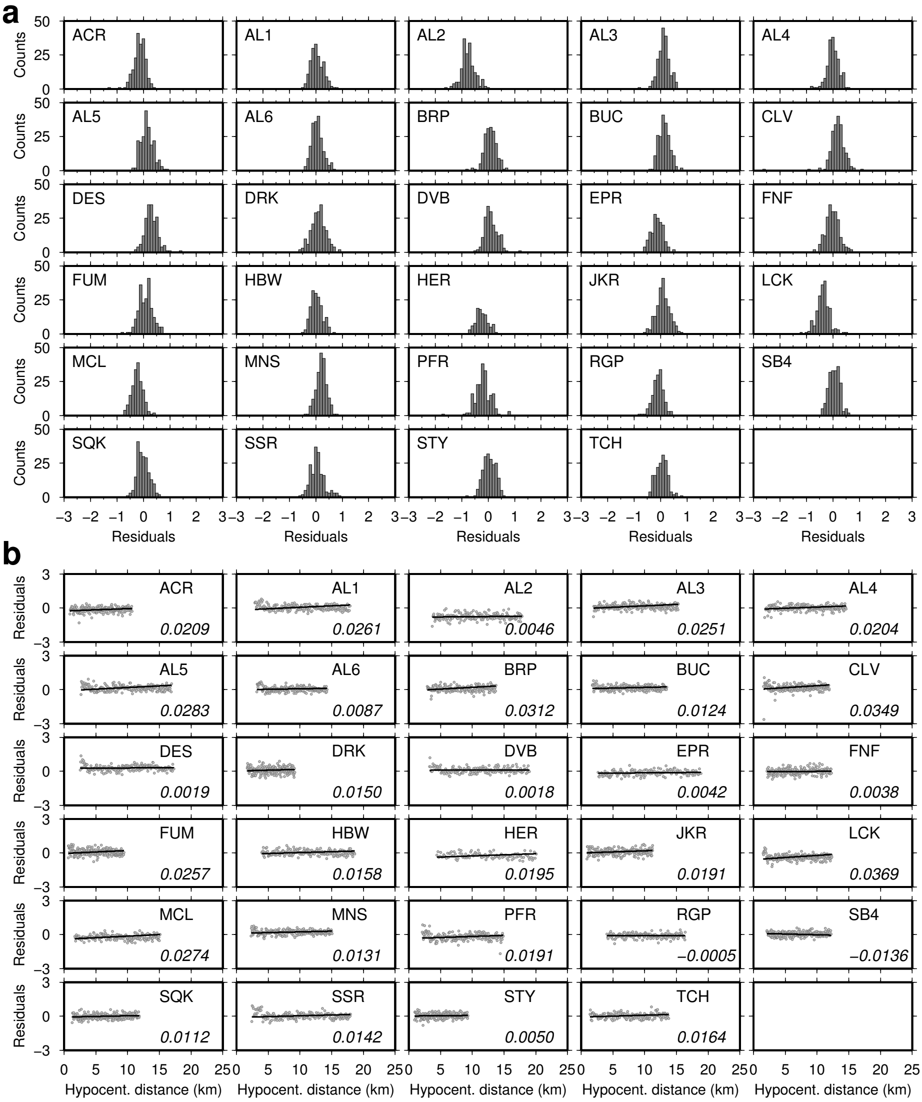


Figure 4.14.: Same as figure 4.9 but for model MOD3(corrected model) PGA. There is an improvement in the model after site correction as now the maxima of residual are centered at zero.

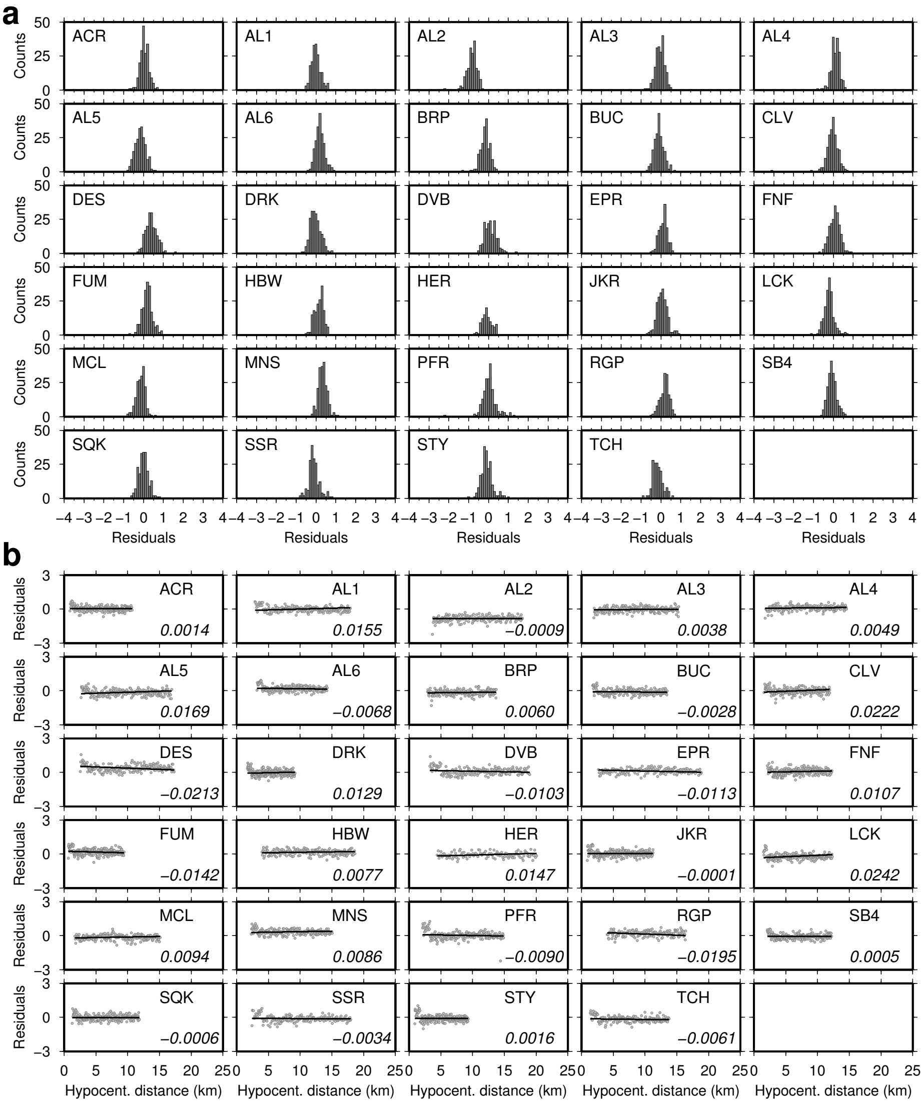


Figure 4.15.: Same as figure 4.10 but for model MOD3(corrected model),SA(T) at T = 0.2s. There is an improvement in the model after site correction as now the maxima of residual are centered at zero.

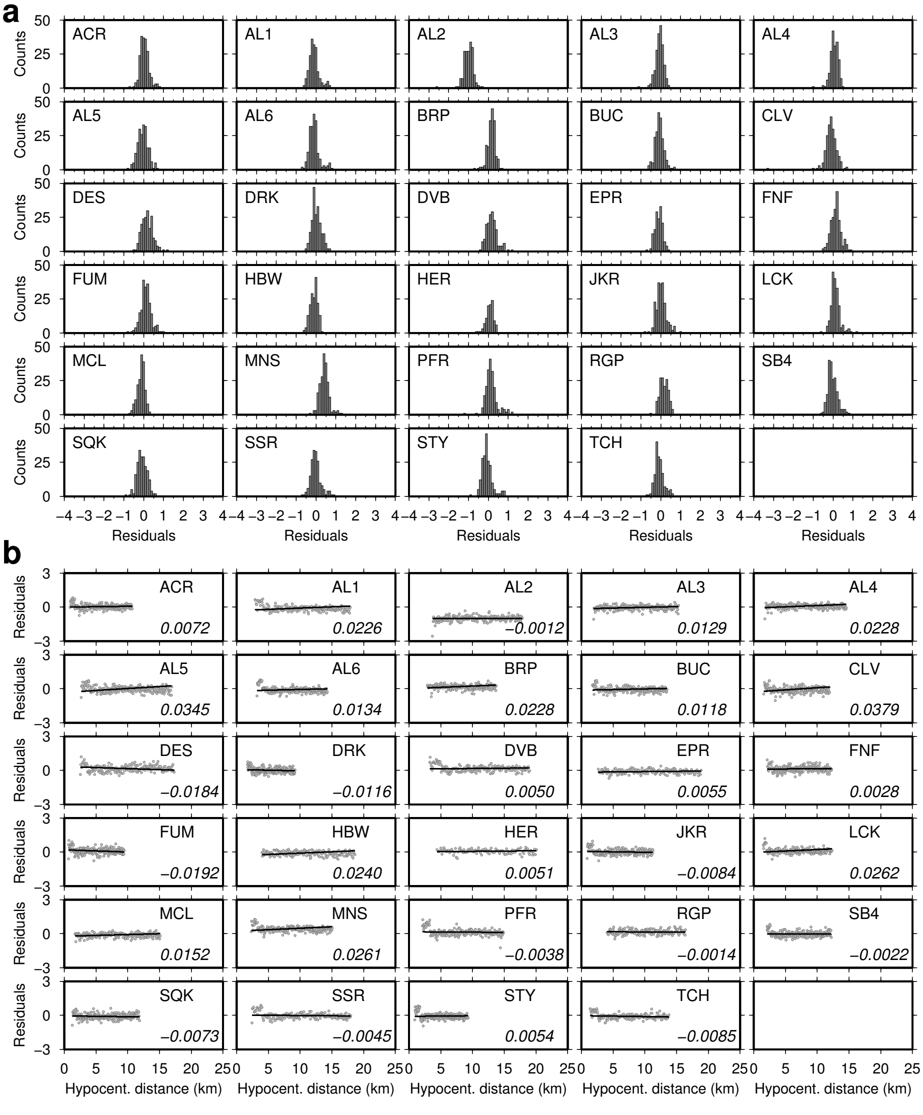


Figure 4.16.: Same as figure 4.11 but for model MOD3(corrected model),SA(T) at T = 0.5s. There is an improvement in the model after site correction as now the maxima of residual are centered at zero.

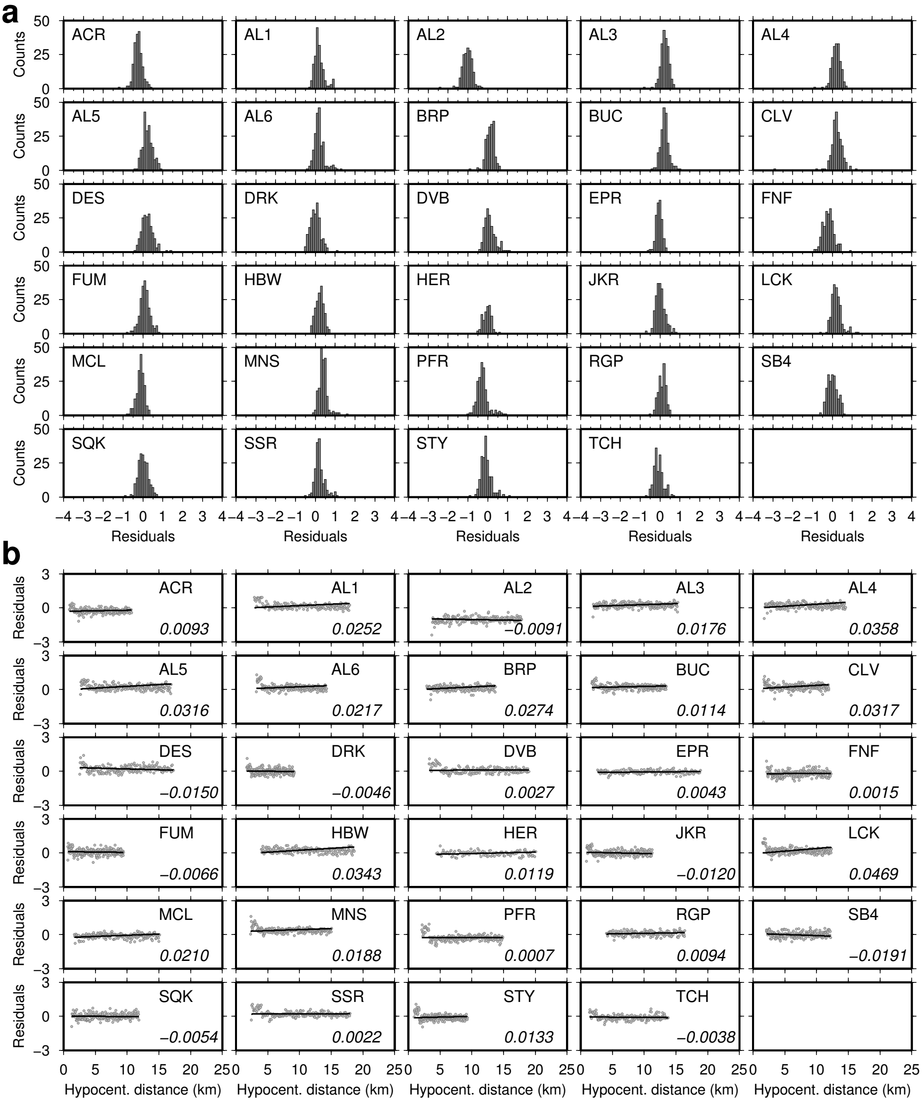


Figure 4.17.: Same as figure 4.12 but for model MOD3(corrected model), SA(T) at T = 1.0s. There is an improvement in the model after site correction as now the maxima of residual are centered at zero.

a function of hypocentral distance and magnitude. No significant trend is observed but as shown in figure 4.18 , for all the ground motion parameters, subset of negative residuals from distance between 6 km through 15 km moved to zero MOD3 is used.

The fit of model MOD3 is plotted in figure 4.19, 4.20 and 4.21 corresponding inter-event and intra-event residuals are also shown. The solid line curves correspond to best model (MOD3) obtained in present study are plotted by setting the site/station correction parameter s = 0, that is, for rock site condition for the magnitude M_w = 1.5, 2.5 and 3.1.

When predictions have to be made at sites other than those used in the present study, we suggest to use MOD3 with s = 0 but adding the contribution of the site-effect using corrective coefficients (e.g. Borcherdt (1944); Cantore et al. (2010)). In the same figures 4.19, 4.20 and 4.21 we have compared the model AB11(Atkinson and Boore (2011)) with Vs30=760m/s for B/C site conditions. As a general consideration, the model AB11 tends to overestimate the observed values for smaller magnitude. But with increase in magnitude there is a good match between both models (MOD3 and AB11).

This final model (MOD3) is applicable with in the range of 20 km of hypocentral distance.

4.2.4. Analysis of station-wise standard deviations

Following Atkinson (2006), we analyzed the standard deviations of the residuals distribution at individual stations before and after site/station correction. The average of these station specific values for all the stations weighted by the number of observations per station, are shown in figure 4.22 (for PGV) and are compared with the standard deviation of the regression obtained from the entire multi-station database (29 stations). The value of weighted mean and total standard deviations are also mentioned in the figure 4.22. The gap between weighted mean standard deviation and total standard deviation is reduced when MOD3 is used. Similar results are observed for PGA, SA(T) at T = 0.2s, 0.5s and 1.0s. (see figures 4.23, 4.24, 4.25, 4.26). The standard deviation at all the stations is lower than that corresponds to the regression model except at station CLV, which is higher than the other stations. This can be due to other effects related to the local geological conditions underneath the station which are not fully

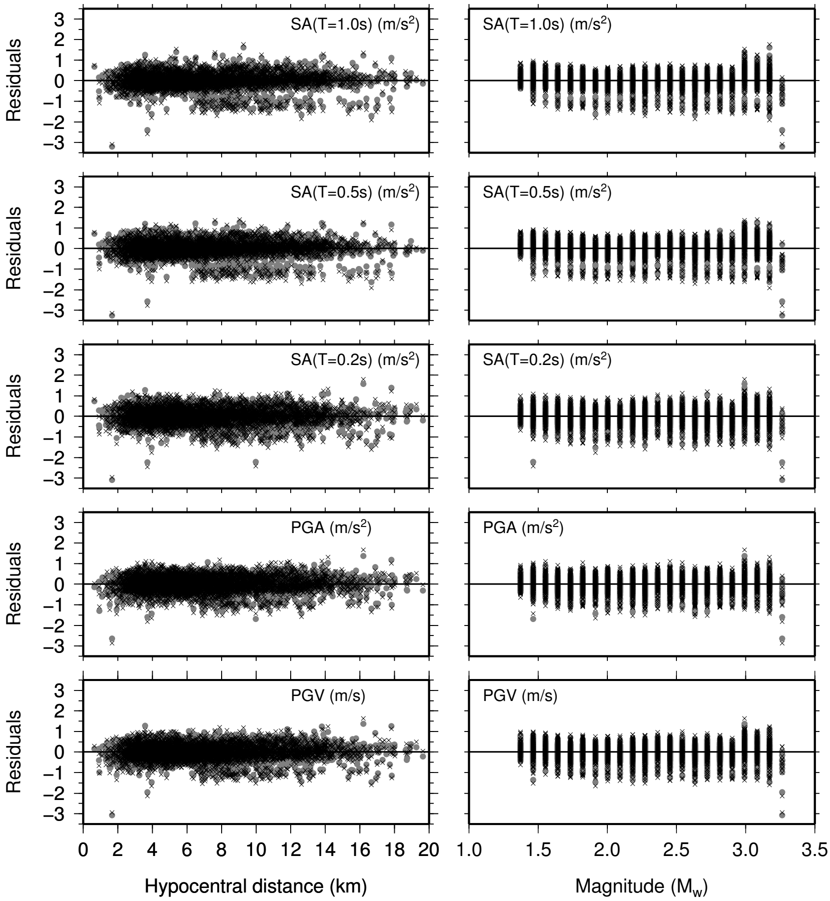


Figure 4.18.: Distribution of residuals as function of hypocentral distance (left) and magnitude (right). Gray dots are showing residuals from MOD3 and black crosses are the residuals from MOD1. It can be clearly observed that the subset of negative data (black crosses) from distance 6 km to 15 km has shifted toward zero which further indicates the improvement in MOD3 after introducing site/station effect correction parameter.

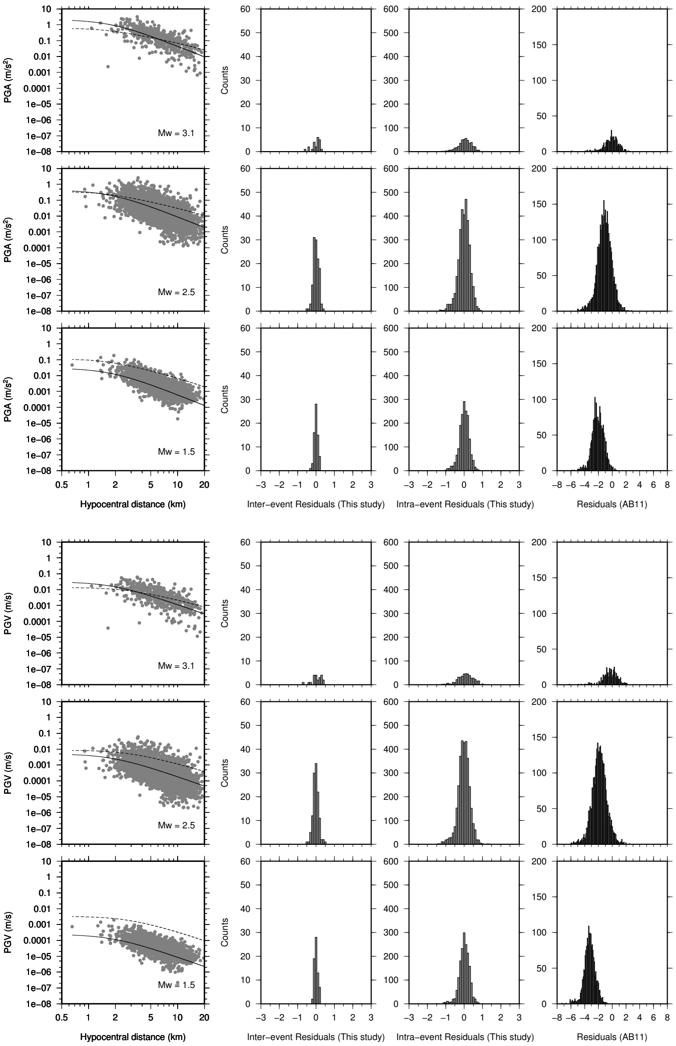


Figure 4.19.: Fitting of model MOD3; solid line, comparison with site/station correction parameter s 0; dashed line, comparison with model AB11; fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for PGV and PGA. For smaller magnitude events, model AB11 overestimates, but with increases in magnitude the predictions of AB11 match our final model (MOD3).

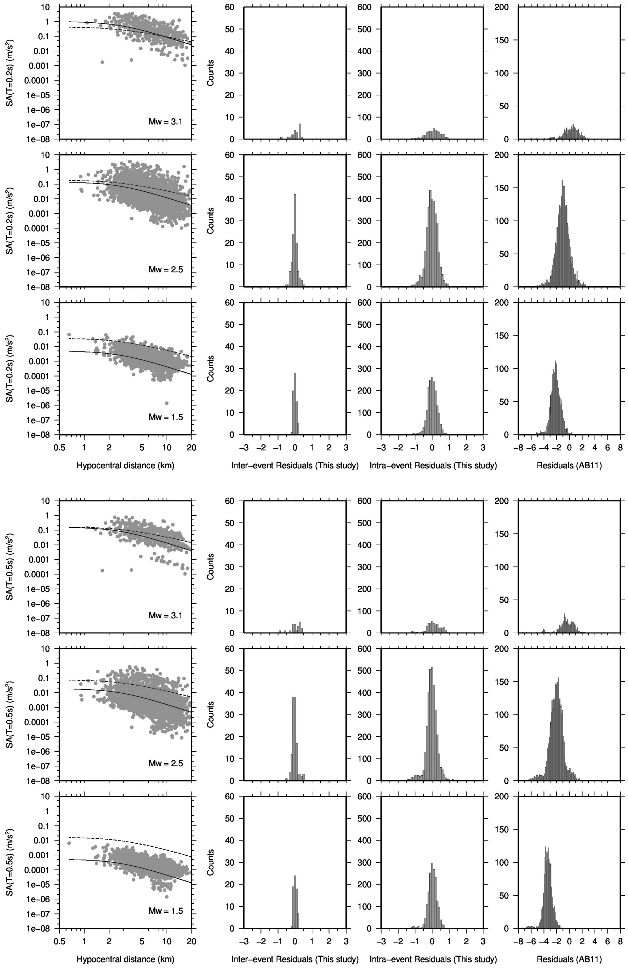


Figure 4.20.: Same as figure 4.19 fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for SA at 0.2s, SA at 0.5s. For smaller magnitude events, model AB11 overestimates, but with increases in magnitude the predictions of AB11 match our final model (MOD3).

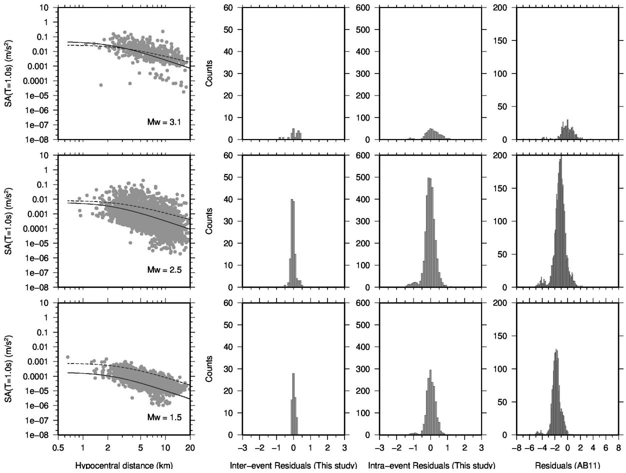


Figure 4.21.: Same as figure 4.19 fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for SA at 1.0s. For smaller magnitude events, model AB11 overestimates, but with increases in magnitude the predictions of AB11 match our final model (MOD3).

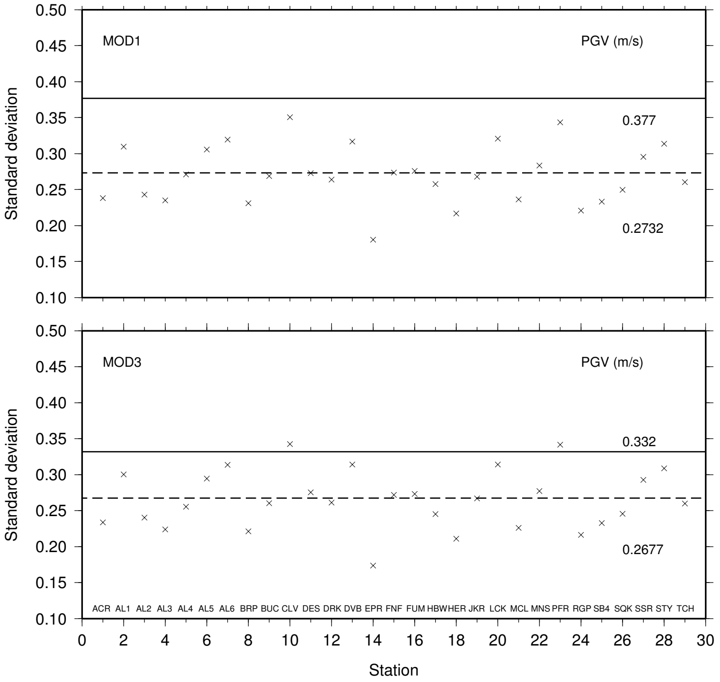


Figure 4.22.: Comparison of standard deviation at each station with standard deviation obtained using complete data for PGV. The dashed line shows weighted mean value of standard deviation over all stations, while the solid line shows the overall standard deviation. Introduction of site correction parameter has improved the model as the gap between weighted mean standard deviation and overall standard deviation has reduced in MOD3 as compared with MOD1. The respective standard deviations are written in the figure.

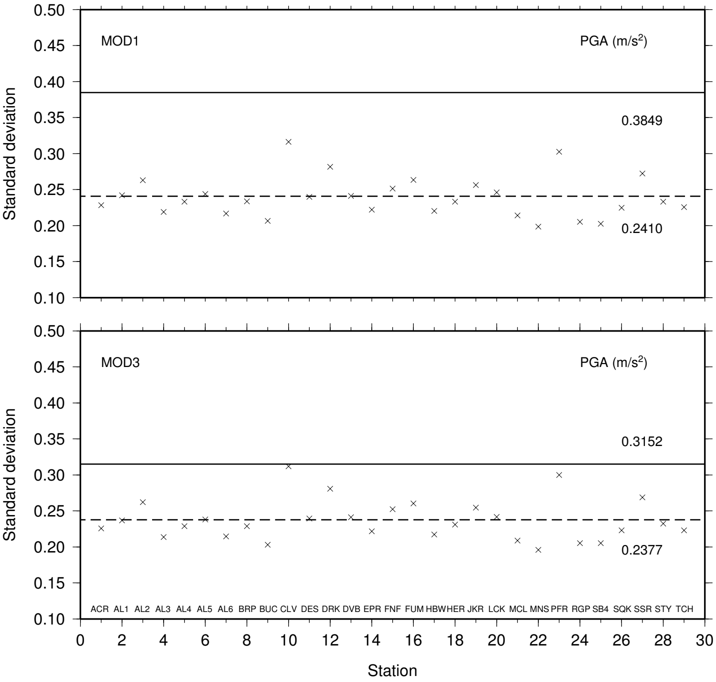


Figure 4.23.: Same as figure 4.22 but for PGA.

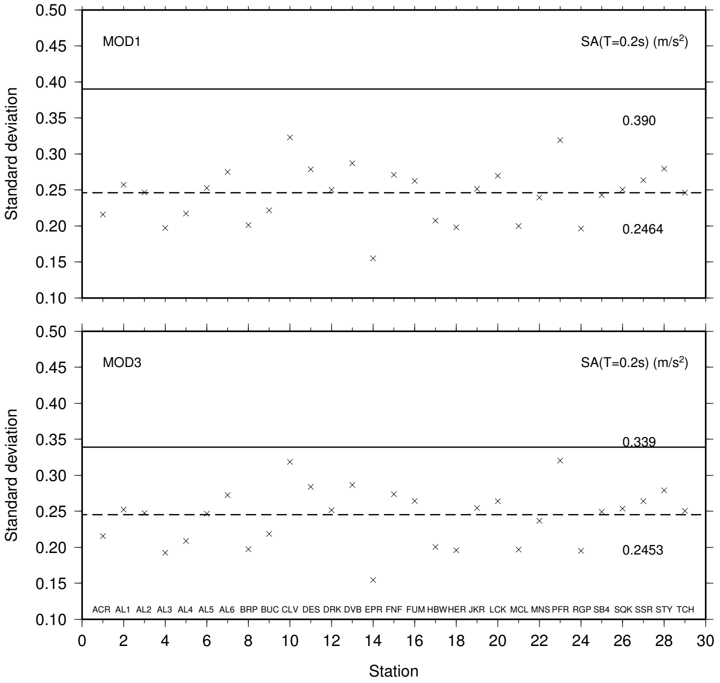


Figure 4.24.: Same as figure 4.22 but for SA (T=0.2s).

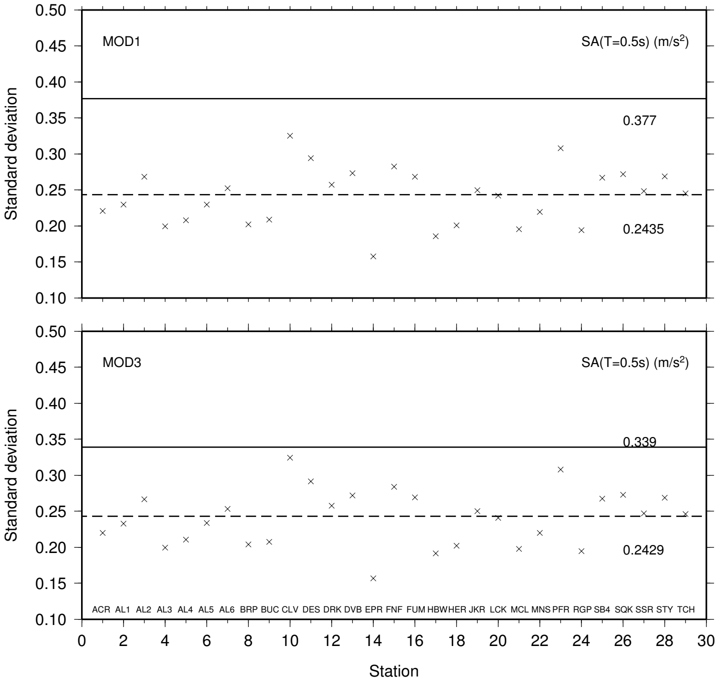


Figure 4.25.: Same as figure 4.22 but for SA (T=0.5s).

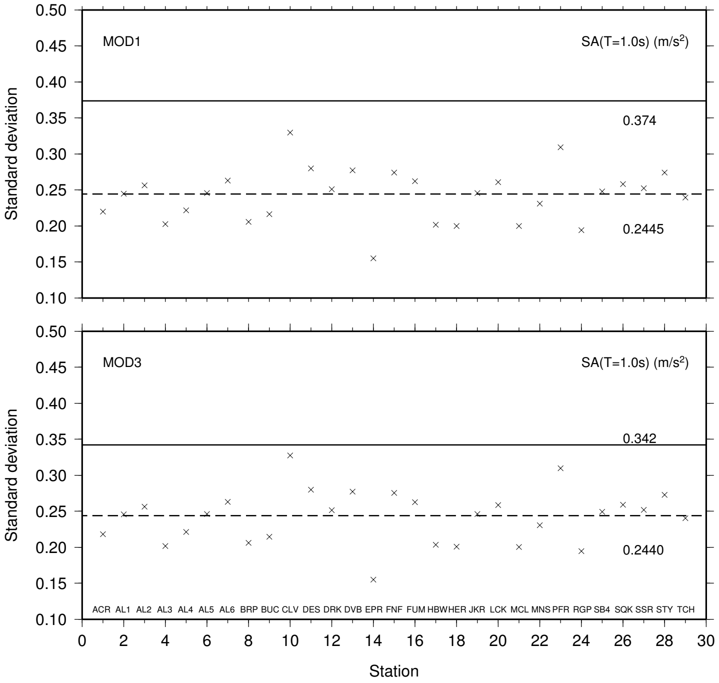


Figure 4.26.: Same as figure 4.22 but for SA (T=1.0s).

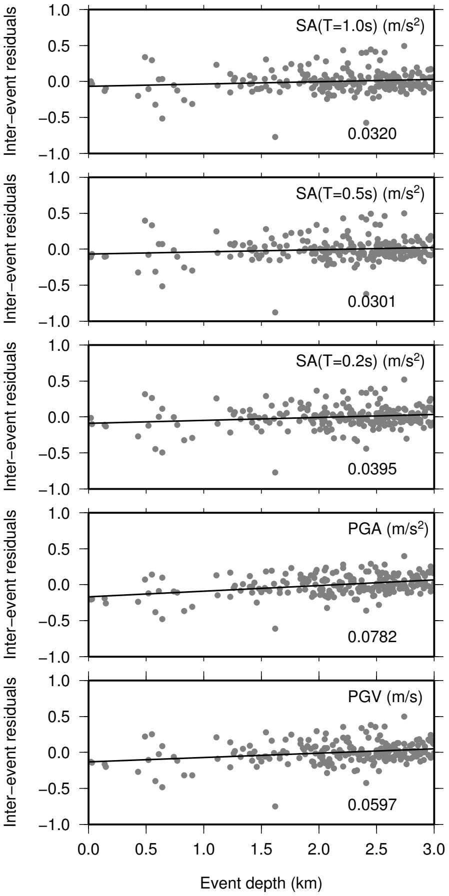


Figure 4.27.: Distribution of inter-event residuals obtained from MOD3 with respect to event-depth the positive trend indicate the increase in stress drop with depth. The same trend is observed for all ground motion parameters. The slope of the fitted line is mentioned in the figure.

accounted by the introduced correction.

4.2.5. Interpretation on inter-event residuals

Finally, using MOD3, we have also analyzed the inter-event residual distribution. Indeed, as reported by Atik et al. (2010), the inter-event residuals account for the influence of source features that are not properly captured by the considered predictor variables (magnitude and distance in the present study). These features are for example the stress drop or the variation of slip in space and time. It could thus be interesting to analyze the dependency of the inter-event residual value as function of depth. As shown in figure 4.27, for all the ground motion parameters we observe only a slight positive trend of inter-event residuals w.r.t. event depth which is reported in each panel. Among other possible explanations, such as a variation in rigidity modulus due to the heterogeneities present in a highly fractured medium, following the interpretation proposed by McGarr (1984) the results could indicate that there is a possible increase in the stress drop with depth.

4.3. Ground motion prediction equation for combined data sets from six geothermal areas.

The work discussed here is the part of the work done in Douglas et al. (2013). In this work various ground-motion datasets of induced and natural seismicity (from Basel, Geysers, Hengill, Roswinkel, Soultz and Voerendaal) are compiled and processed and the moment magnitudes for all events are computed homogeneously. These data are used to show that ground motions from induced and natural earthquakes cannot be statistically distinguished. Empirical GMPEs are derived from these data and it is shown that they have similar characteristics to some recent GMPEs for natural and mining-related seismicity but the standard deviations are higher. Subsequently stochastic models, to account for epistemic uncertainties, are developed based on a single corner frequency and with parameters constrained by the available data. Predicted ground motions from these models are fitted with functional forms to obtain easy-to-use

GMPEs. These are associated with standard deviations derived from the empirical data to characterize the aleatory variability. As an example, it is demonstrated the potential use of these models using data from Campi Flegrei. But here we will confined our discussion related to GMPE's, further details are beyond the scope of the thesis. More details can be found in the article by Douglas et al. (2013). Some details of the Data sets used from above discussed geothermal area are mentioned below.

4.3.1. Basel

The Basel is a 200°C reservoir with depth of 5 km, beneath the city. This EGS project was proposed to provide up to 3MW in electrical production and 20MW thermally. To monitor seismicity in this region a dense network of surface sensors (Swiss Seismological Service, SED; Baden Wüttemberg Seismological Service, LED) and bore-hole sensors (Geo Explorers Ltd.) was deployed. The surface instruments of the SED are either STS-2 broadband seismometers or episensor accelerometers geographically-oriented. Through microzonation studies of Havenith et al. (2007) and site investigations undertaken as part of a probabilistic seismic hazard study for nuclear power facilities in Switzerland (Fah et al. (2009)) the site conditions are well known. Basel lies on a sedimentary basin some several-hundred meters thick. Time-averaged shear-wave velocity down to 30m (VS30) of the sites around Basel tends to be around 400m/s (NEHRP site class C), although rock-site stations at greater distances were also used for determination of magnitudes. The borehole sensors of Geo Explorers Ltd. are short-period geophones with a natural frequency of around 5Hz and a damping coefficient of 0.2. We only use those geophones which are located on the surface for the ground-motion analysis, while the borehole data were included in the magnitude determination. Some instruments were of unknown orientation, so all data are rotated to the direction of maximum amplitude. All events in the Basel dataset are geothermal-induced events, located by the SED. Data were corrected for the amplitude and phase response of the instrument and differentiated to provide acceleration time-series.

4.3.2. Campi Flegrei

Campi Flegrei caldera is a volcanic area, which includes part of the metropolitan area of Naples (southern Italy), one of most densely populated areas in Europe. It is a large depression (with a radius of about 6 km) formed by huge ignimbritic eruptions, the last one having occurred 15,000 years ago (Deino et al. (2004)). This area is also affected, for the past two thousand years at least, by episodes of large uplift and subsidence, as testified by marine ingression levels in Roman and Middle-Age monuments and ruins (Dvorak and Mastrolorenzo (1990)). Started in 1969, unrest episodes has caused maximum uplift of about 3.5m in 15 years (peak rate of about 1m/year) and, after about ten years of stable ground level, restarted in 1982 until 1984. During this final period maximum uplift rates were recorded by leveling networks and tide gauges (De Natale et al. (2006)) and more than 15,000 micro-earthquakes , with magnitudes in the range 0-4.2, occurred (De Natale and Zollo (1986)). Such earthquakes were strongly felt by the population, as they were shallow, with maximum depths of 34 km, but no significant damage was observed (De Natale et al. (1988)). In January 1984, a digital network owned by the University of Wisconsin, was installed at Campi Flegrei, consisting of 13 digital, three-component stations with a 125-250Hz sampling rate (e.g. Aster et al. (1989)). The catalog analyzed here contains the data presented in the aforementioned papers, with the addition of other records that were not analyzed at that time.

4.3.3. Geysers

We analyzed induced seismicity at Geysers between August 2007 and February 2011. For the study region there are waveforms of more than 11,000 located events with magnitudes larger than 1.0 available from the NCEDC. The largest earthquake recorded was the January 4, 2009, M_w 4.3 event. The data was associated with events from the NCEDC earthquake catalog, updated the meta data for all traces, and automatically re-picked the P-wave first arrival times for quality control. Be cause different magnitude types ($\mathrm{M_D}$, ML and M_w) are used in the original catalog, moment magnitude has been recomputed here for all the events with M 1.5 (any scale). To compile a representative subset for this study, we divided the available range of catalog magnitude in bins 0.25 units wide. Within each bin, we selected those ten events with the most validated P-wave picks at BG stations to ensure accurate locations and high-quality waveforms.

4.3.4. Hengill

Ground-motion data used here are recorded close to the Hengill (south-western Iceland) geothermal system by a temporary broadband network installed within the framework of the I-GET FP6 project Jousset and Francois (2006). The temporary network operated from late June until mid-October 2006 and consisted of seven Guralp Systems broadband instruments (CMG3-ESP and CMG40TD) distributed to monitor and explore the Hengill hydrothermal system. In addition, data from the three stations (Lennartz 5s instruments) of the SIL permanent seismic network operated by the Icelandic Meteorological Office (IMO) nearest Hengill (KRO, HEI and SAN) were also collected for this study. Some basic site descriptions are available for the temporary stations, which show variations; however, it is assumed that the stations are all located on rock be cause of the general geology of the Hengill area (shallow volcanic soils overlying lava of various ages). Jousset and Francois (2006) report previous studies suggesting that cooling, mostly due to natural heat loss, and consequential thermal contraction and cracking in the heat source are responsible for the continuous small-magnitude seismicity in the Hengill area. This was deduced by the non-double-couple focal mechanisms with large explosive components, which may be attributable to fluid flow into newly formed cracks. The earthquake catalog of IMO was queried to find those earthquakes with moment magnitudes larger than $M_w 1.0$ that were recorded by one or more of the ten instruments installed in the Hengill region. The acceleration time-histories (derived by time-domain differentiation from the velocity measurements recorded by the broadband sensors) corresponding to these earthquakes were selected.

4.3.5. Roswinkel

The Roswinkel (north-eastern Netherlands) is a natural gas field, situated in a heavily-faulted anticline structure in Triassic sandstones at a depth of around 2.1 km. The field was in production from 1980 to 2005, while seismicity was observed from 1992 to 2006. There were 39 earthquakes in total with a strongest event of magnitude ML = 3.4 and epicentral intensity I_0 = VI. The seismicity that occurred so far has been associated with existing faults on top of the reservoir (Dost and Haak (2007); Van Eck et al. (2006)). The dataset of strong-motion from 27 events recorded at the Roswinkel village obtained with SIG AC-23 sensors are used in this study.

4.3.6. Soultz

The Soultz is a collaborative French-German project and geothermal exploitation began in the late 1980s. The first subterranean circulation of water using the drilled boreholes was achieved in the late 1990s. The boreholes are about 5 km deep and various reservoir stimulations were undertaken in the first decade of this century. The data from three permanent three-component surface stations (FOR, OPS and SRB) installed by Ecole et Observatoire des Sciences de la Terre (EOST) of the University of Strasbourg in 2003 is used here, with lose to the injection wells of the EGS (Charlety et al. (2007)). These stations record amplitudes proportional to ground velocity which is then converted to acceleration by differentiation and application of the calibration factor. The records are of injection experiments conducted in 2003. A high-quality earthquake catalog was provided by EOST. Since the recorded events are all induced in the geothermal reservoir the records are associated with similar hypocentral distances.

4.3.7. Voerendaal

The Roer Valley is an active rift system in the Lower Rhine Embayment in the border area of the Netherlands, Belgium and Germany (Dost and Haak (2007)). Most of the seismicity in the area is situated within the Roer Valley Graben and associated to its main bounding faults: the Peel Boundary Fault to the north-east and the Feldbiss Fault to the south-west. The village of Voerendaal is located on the South Limburg Block, south-west of the Feldbiss Fault. The region around Voerendaal has shown anomalous swarm-like seismicity at relatively shallow depths (around 3 to 8 km). A first swarm was detected in 1985 which lasted for more than a month. Nine events were located, the largest of them was of magnitude ML = 3.0 with maximum epicentral intensity I_0 = IV. On December 20, 2000 a new swarm activity started. This time the swarm lasted for more than a year, with 139 detected events and a strongest event with magnitude ML = 3.9 with intensity I_0 = VI. The dataset used in this study comprises 136 events from the Voerendaal area between April 1999 and August 2009. The waveforms consist of both short-period recordings from the regional seismic network (Willmore Mk III sensors) and accelerometeric recordings from within the Voerendaal village, where since the start of the second swarm three SIG AC-23 strong motion sensors were deployed.

4.4. Regression analysis

Before computation of GMPEs ANOVA test (Douglas et al. 2013) was done to support the idea of mixing all dataset together to generate a common GMPEs for all geothermal areas. These GMPEs allow the scaling of ground motions with respect to magnitude and distance to be studied and compared to those predicted by the stochastic model (Douglas et al. (2013)). In addition, differences between ground motions from different areas are investigated. Thirdly, the predictions are compared to those made by existing GMPEs derived from data from moderate and large earthquakes, specially those by: Ambraseys et al. (2005) (AB05), Bommer et al. (2007) (BM07) and Massa et al. (2008) (MS08), to see whether such models can be extrapolated to the prediction of shaking from small events. The GMPEs are derived using NLMRA (Lindstrom and Bates (1990); Abrahamson and Youngs (1992)), which accounts for between (inter)-event and within (intra)-event variabilities (Atik et al. (2010)) (see chapter 3). Only records with $r_{\text{hypo}} < 50$ km are used. The model selected for regression (Model 1) has a standard functional form (equation 4.11), accounting for first-order effects of magnitude-scaling, near-source saturation, geometrical spreading and anelastic attenuation:

$$ln(Y) = a + bM + cln(\sqrt{r_{hypo}^2 + h^2}) + dr_{hypo} \tag{4.11}$$

where Y is the response variable corresponding to PGA, PGV or PSA at various structural periods (in SI units) and *a, b, c, d* and *h* are regression coefficients. Coefficients obtained from the regression analysis along with their uncertainties and the two principal components of the standard deviation are reported in table 4.5 for PGA, PGV and PSA for three selected periods, for both uncorrected data and corrected data for site effect. The coefficients for periods up to 0.5s are also listed in table 1, 2 of the appendix A.

The magnitude-scaling of the derived GMPEs (coefficient b) closely match the magnitude dependencies reported by Douglas and Jousset (2011) from previous empirical GMPEs derived using data from small (M_w<3) natural and mining-related earthquakes (their figure 1), thereby suggesting that the magnitude-scaling of induced, mining and natural seismicity are comparable. Comparing the coefficient b of the GMPEs for PGV with the coefficients for PGA and PSA shows that PGV for small events is associated with very high frequencies and hence the method of Bommer and Alarcon (2006) to estimate PGV from PSA(0.5s) is not recommended for such earthquakes. The regression coefficients c and d indicate fast decay with distance, which should be attributable to strong an-elastic attenuation (i.e. low Q values). Except at high frequencies, the total standard deviation obtained from the regression on site-corrected data is lower than that corresponding to uncorrected data, which shows that the site effect contributes significantly to ground-motion variability. It should be noted that the high frequency site correction may actually reintroduce some source variability that was effectively hidden by the site attenuation (κ). Consequently we observe an increase in total σ at high frequencies including PGA and PGV, driven by an increase in the between-event variability (τ).

Figure 4.28 shows residual plots for data corrected for site response. As is usual, the within-event residual distributions show larger dispersion compared with the between-event ones. The comparison with other models shows that the residual distributions of BM07 and MS08 are similar to that from our model. On the other hand, the AB05 GMPEs lead to a wide distribution of residuals, which can be attributed to the applicability of this GMPE to larger magnitudes. Indeed, BM07 is based on a dataset where the minimum $M_w$was 3.0, MS08 analyzed data with minimum M_w of 4.0 while AB05 considered strong-motion data relative to earthquakes with $M_w \geq$ 5.0. To investigate what is the largest contribution to the total residual distribution we analyze them as a function of both magnitude and

Table 4.5.: Coefficients for Model 1 (equation 4.11) for selected periods. τ is the between-event, ϕ the within-event and σ the total standard deviations.

Period	$a \pm s_a$	$b \pm s_b$	$c \pm s_c$	$h \pm s_h$	$d \pm s_d$	τ	ϕ	σ
Uncorrected for site- effect								
0.01(PGA)	-5.984 ± 0.423	2.146 ± 0.069	-1.772 ± 0.208	2.511 ± 0.595	-0.023 ± 0.011	0.792	0.829	1.147
0.10	-6.444 ± 0.329	2.376 ± 0.056	-1.410 ± 0.167	1.751 ± 0.704	-0.039 ± 0.009	0.815	0.623	1.026
0.20	-9.513 ± 0.243	2.805 ± 0.058	-0.776 ± 0.121	Constrained to zero	-0.057 ± 0.008	0.800	0.612	1.007
0.50	-11.232 ± 0.480	2.912 ± 0.081	-0.953 ± 0.223	2.772 ± 1.181	-0.022 ± 0.011	0.638	0.816	1.036
PGV	-10.367 ± 0.449	2.018 ± 0.136	-1.124 ± 0.183	2.129 ± 0.895	-0.046 ± 0.010	1.811	0.745	1.958
Corrected for site- effect								
0.01(PGA)	-6.514 ± 0.423	1.995 ± 0.085	-1.468 ± 0.200	2.490 ± 0.688	-0.029 ± 0.010	0.730	1.079	1.303
0.10	-7.991 ± 0.229	2.376 ± 0.063	-0.827 ± 0.106	1.058 ± 1.049	-0.056 ± 0.006	0.589	0.787	0.983
0.20	-10.024 ± 0.219	2.784 ± 0.057	-0.850 ± 0.103	1.080 ± 0.979	-0.041 ± 0.006	0.554	0.658	0.860
0.50	-12.736 ± 0.282	3.056 ± 0.070	-0.675 ± 0.129	1.218 ± 1.434	-0.050 ± 0.008	0.572	0.698	0.903
PGV	-9.999 ± 0.681	1.964 ± 0.122	-0.1405 ± 0.321	2.933 ± 1.088	-0.035 ± 0.016	1.029	1.553	0.1863

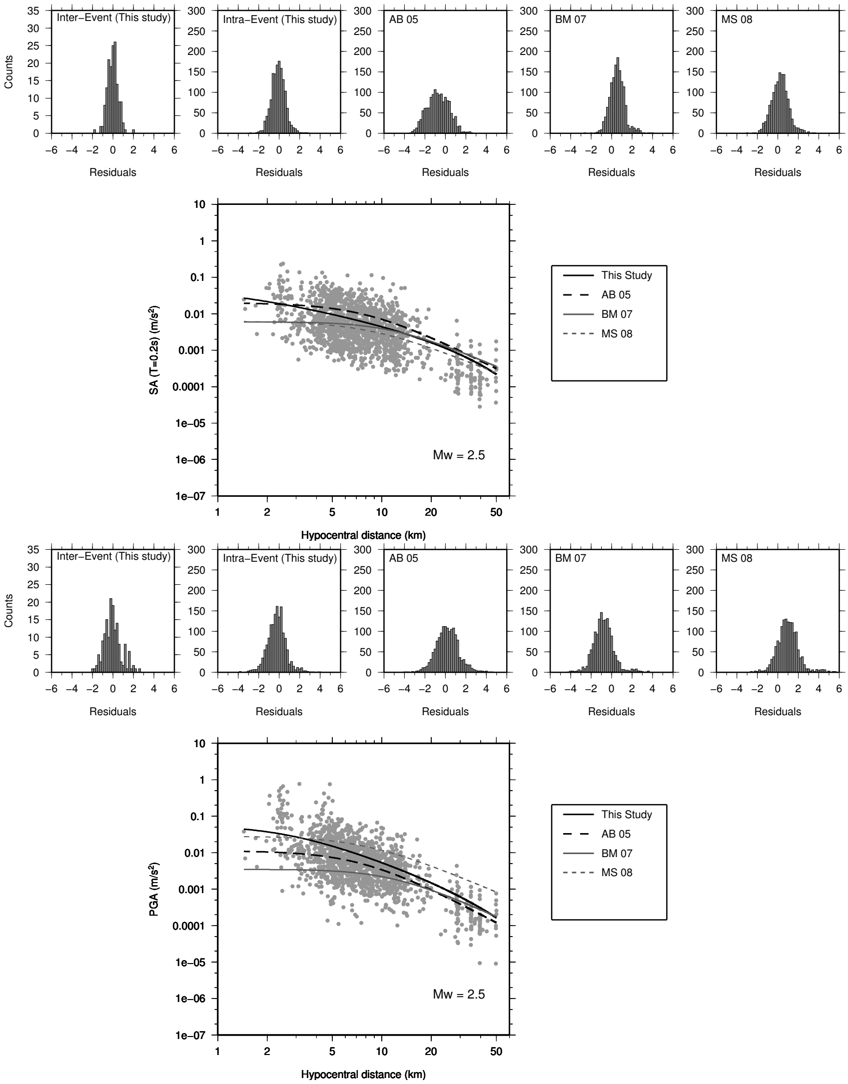


Figure 4.28.: Residual distributions and comparisons between data and predictions (corrected data). The lower panels show the data from earthquakes with $2 < M_W < 3$ and the curves corresponding to M_W 2.5 while the upper panels show the residual distributions using the GMPEs derived here and those of AB05, BM07 and MS08. Moreover, for the GMPEs from the present study the residuals are separated into between-event and within-events components.

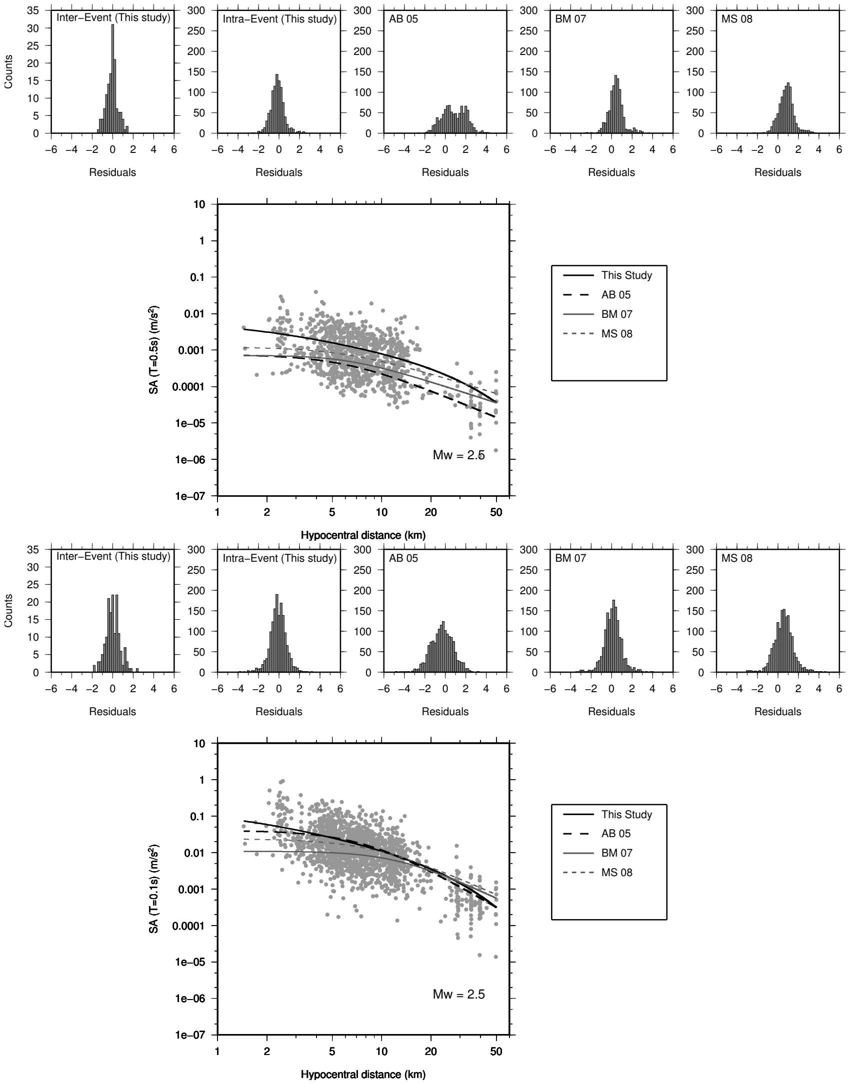


Figure 4.29.: Same as figure 4.28

distance (see figure 4.30, 4.31, 4.32, 4.33). The obtained results which, for brevity are not reported here, show that the largest contribution to the residual dispersion comes from the distance, and, in particular, data recorded at shorter distances feature higher residual values at all structural periods. This indicates that for data collected from earthquakes occurring in these areas it is likely that an-elastic attenuation plays an important role. On the other hand, aside from the results obtained from the AB05 model, the residuals as a function of magnitude are characterized by a quite uniform dispersion, mostly centered on zero.

4.4.1. Effect of focal depth and residual analysis

To analyze the effect of focal depth on the regression models, and thus on the predictions of the ground motion, we implemented two additional models which are compared with the model reported in equation 4.11. From now onwards we refer only to data corrected for site response. First, we selected a model where the regression coefficient h is no longer used in conjunction with r_{hypo}, to discuss the relative importance of this coefficient. Second, we implemented a model in which epicentral distance, r_{epi}, is used instead of r_{hypo}. Models 2 and 3 have the following formulations:

$$ln(Y) = a + bM + cln(r_{hypo}) + dr_{hypo} \tag{4.12}$$

$$ln(Y) = a + bM + cln(\sqrt{r_{epi}^2 + h^2}) + dr_{epi} \tag{4.13}$$

The coefficients and the associated uncertainties for the two models are listed in tables 4.6 and 4.7. The coefficients for periods up to 0.5s are also listed in tables 3, 4 of the appendix A .

Concerning PGA, the between- and within-event standard deviations of Models 1 and 2 are equal while the between-event component for T $\leq$ 0.1s for Model 3 is slightly reduced. This could be a consequence of greater scatter in focal depths, which are more poorly defined than the epicenters. Unlike moderate and large earthquakes, which rupture a large proportion of the seismogenic layer, small earthquakes, such as those induced and triggered by EGSs, are associated with rupture of less than a kilometer. Consequently whether this rupture occurs at a depth of, say, 20 km or at 2 km will have a large impact on the shaking at the surface. To test the

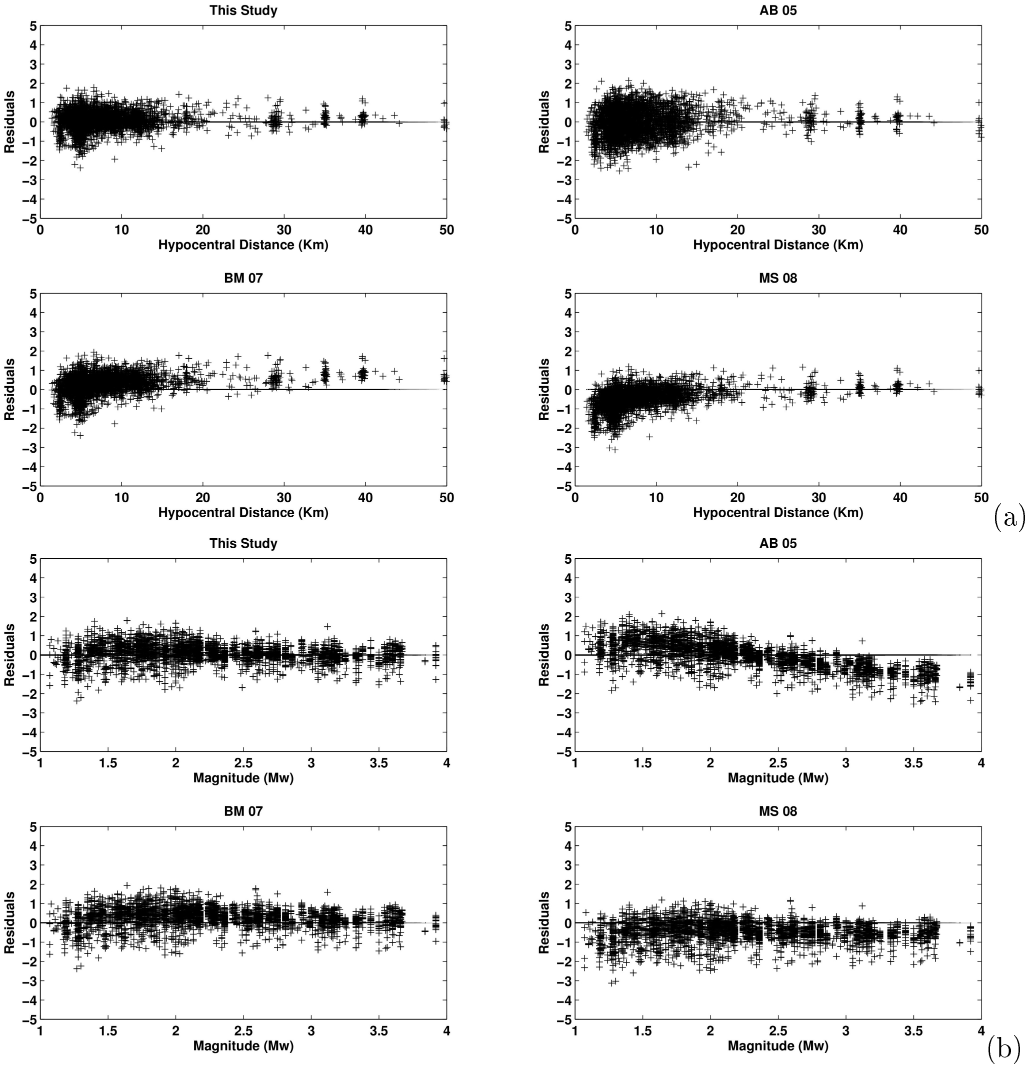


Figure 4.30.: Residual distribution as a function of (a) hypocentral distance and (b) magnitude for Model 1 (corrected for site effect), for PGA showing no significant correlation.

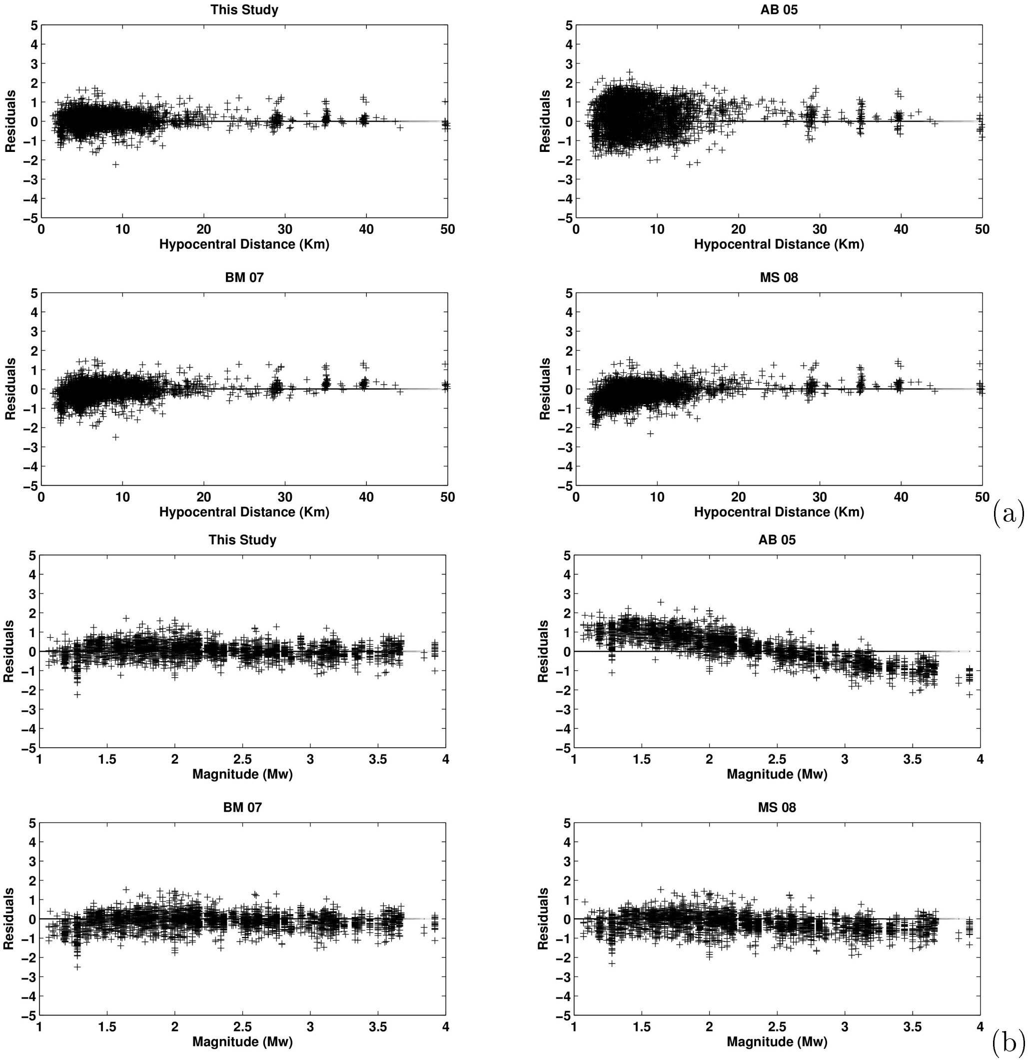


Figure 4.31.: Same as figure 4.30 but for period 0.1s

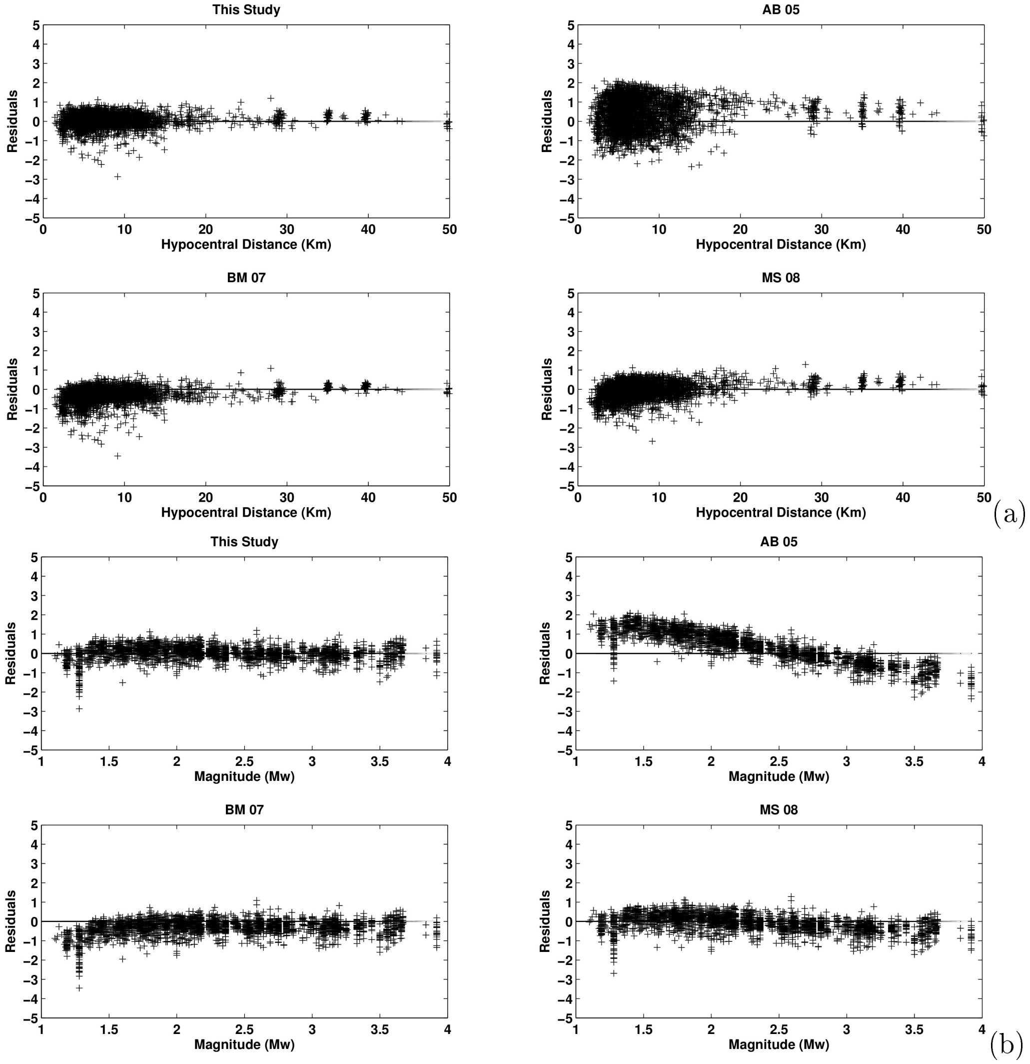


Figure 4.32.: Same as figure 4.30 but for period 0.2s

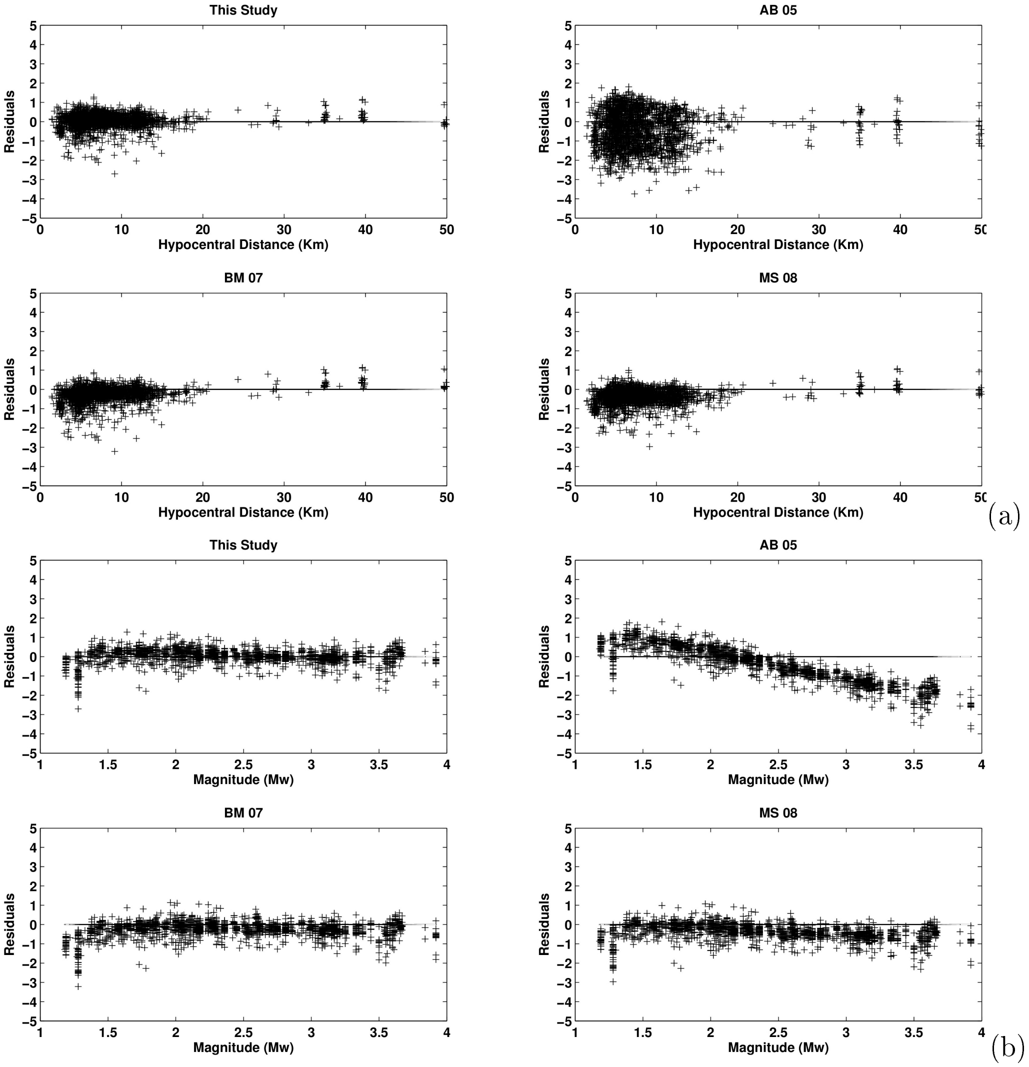


Figure 4.33.: Same as figure 4.30 but for period 0.5s

Table 4.6.: Coefficients for Model 2 (equation 4.12) for selected periods. τ is the between-event, ϕ the within-event and σ the total standard deviations.

Period	$a \pm s_a$	$b \pm s_b$	$c \pm s_c$	$d \pm s_d$	τ	ϕ	σ
0.01(PGA)	-7.198 ± 0.200	1.991 ± 0.085	-1.105 ± 0.062	-0.048 ± 0.005	0.731	1.080	1.304
0.10	-8.079 ± 0.150	2.375 ± 0.063	-0.778 ± 0.050	-0.058 ± 0.004	0.589	0.787	0.983
0.20	-10.118 ± 0.141	2.783 ± 0.057	-0.799 ± 0.048	-0.044 ± 0.004	0.554	0.658	0.860
0.50	-12.829 ± 0.182	3.055 ± 0.070	-0.624 ± 0.060	-0.053 ± 0.005	0.573	0.696	0.902
PGV	-10.846 ± 0.283	1.961 ± 0.121	-0.962 ± 0.088	-0.056 ± 0.007	1.031	1.536	1.850

Table 4.7.: Coefficients for Model 3 (equation 4.13) for selected periods. τ is the between-event, ϕ the within-event and σ the total standard deviations.

Period	$a \pm s_a$	$b \pm s_b$	$c \pm s_c$	$h \pm s_h$	$d \pm s_d$	τ	ϕ	σ
0.01(PGA)	-7.397 ± 0.394	2.117 ± 0.081	-1.282 ± 0.181	3.173 ± 0.469	-0.030 ± 0.010	0.735	1.021	1.258
0.10	-8.824 ± 0.212	2.481 ± 0.062	-0.620 ± 0.095	2.066 ± 0.487	-0.059 ± 0.006	0.592	0.766	0.968
0.20	-10.570 ± 0.255	2.869 ± 0.057	-0.753 ± 0.118	2.675 ± 0.520	-0.040 ± 0.007	0.556	0.666	0.868
0.50	-13.328 ± 0.257	3.104 ± 0.070	-0.497 ± 0.113	2.304 ± 0.761	-0.054 ± 0.007	0.574	0.704	0.908
PGV	-9.672 ± 0.966	2.064 ± 0.123	-1.743 ± 0.432	4.961 ± 0.837	-0.015 ± 0.019	1.029	1.556	1.867

effect of the focal depth on the predicted values from the three models, we considered a set of epicentral distances, and for each distance the actual range of focal depths contained in the data is taken into account. For each term composed of depth, r_{epi} and r_{hypo}, we consider the differences $\ln \mathrm{Y}_{\mathrm{model}i} - \ln \mathrm{Y}_{\mathrm{model}j}$ where i and j correspond to 1, 2 and 3. We note that the differences depend only weakly on structural period and magnitude. Therefore, here we show (in figure 4.34) the results only for PGA and for a representative magnitude $M_w = 2.5$.

As expected, Models 1 and 2 show the same behavior for depths larger than 3 km. The models differ for $r_{epi} < 1$ km and focal depths less than 3 km, where Model 2 provides predictions larger than those by Model 1. This is due to coefficient h in Model 1, which avoids unrealistic Y values at small distances. The comparison between Models 1 and 2 with respect to Model 3 is more important for evaluating effect of the focal depth. Aside from the absolute values, the differences between Models 1 and 3, and those between Models 2 and 3 share the same characteristics. In particular, all the models are similar starting from epicentral distances of 10-15 km. On the other hand, a net difference is observed for $r_{epi} < 5$ km, with a different trend depending on the depth and an inflection point at about 3 km. For depths less than 3 km, the predictions made by Models 1 and 2 are larger than those obtained from Model 3, while the opposite is observed for depths greater than 3 km. Thus depth plays a fundamental role, and r_{hypo} should be more effective than r_{epi}, particularly at short distances. As for natural earthquakes (e.g. Douglas (2007)), it is of interest to investigate the effect of tectonic environment on ground motions from induced events. With this aim, we analyze the residual distributions for each of the six zones. For natural earthquakes , ground motions in different regional stress fields an be significantly different for the same magnitude and source-to-site distance (e.g. McGarr (1984); Bommer et al. (2003); Convertito and Herrero (2004)). For induced seismicity, local stress conditions are mostly driven by field operations, which an reactivate existing faults or generate new ones with mechanisms different to those expected from the regional stress field (Oppenheimer (1986); Li et al. (2011)).

In figure 4.35 we show the PGA residuals as function of r_{hypo}, M_w and depth. It can be noted that dominant contributions come from Geysers and Hengill. As a general comment, for all the models and all the considered

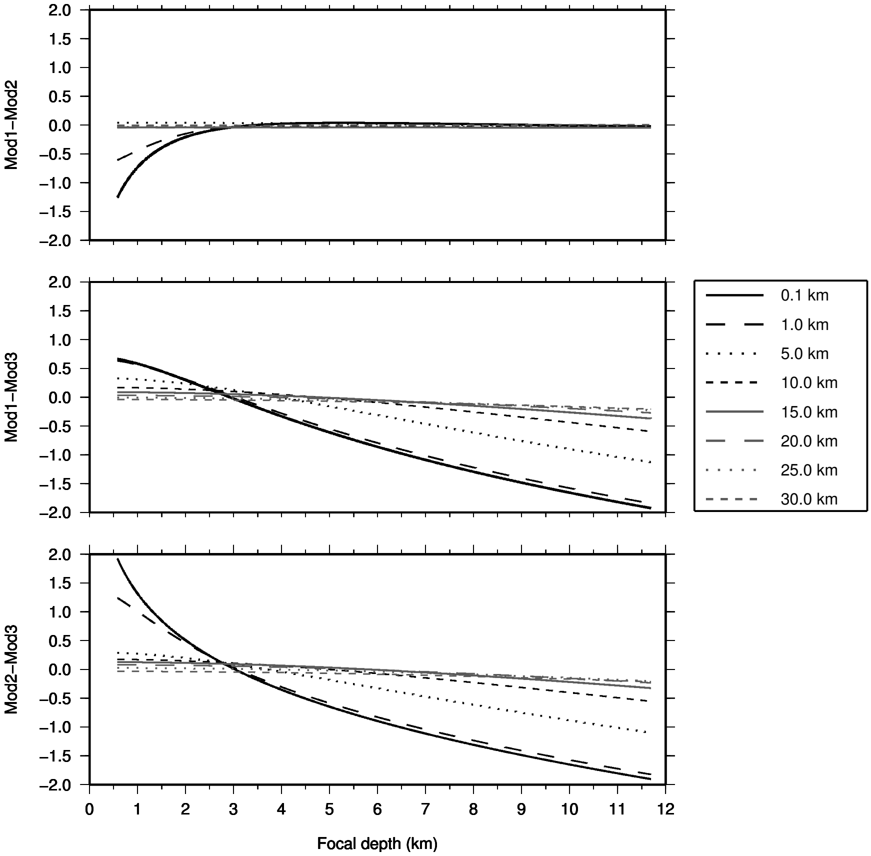


Figure 4.34.: Comparison of predicted PGA of the three models as a function of focal depth.

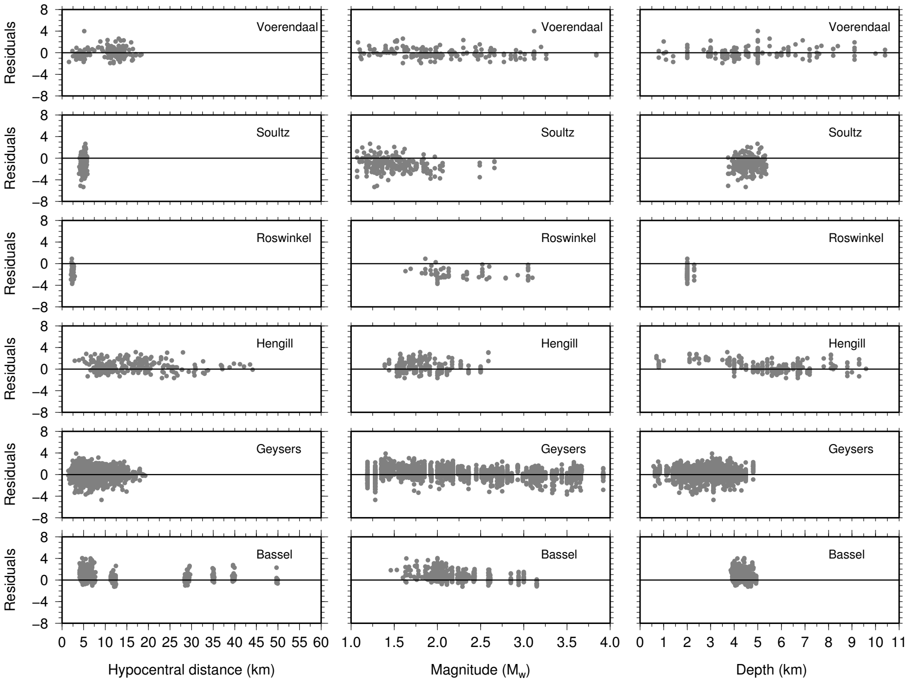


Figure 4.35.: Residual analysis with respect to each area for PGA and model 1.

structural periods we do not observe particular correlations among the residuals and the three variables r_{hypo}, M_w and focal depth.

4.5. Application to time dependent seismic hazard

In chapter 2 we have discussed about seismic hazard analysis and it was discussed that GMPEs are very important for hazard estimation. In this section we will discuss the application of GMPEs in estimation time dependent seismic hazard and as well as monitoring seismic hazard. Convertito et al. (2012) have performed time-dependent seismic hazard analysis for The Geysers geothermal area. The data used for analysis in The Geysers area was between 1 September 2007 and 15 November 2010. It was shown that in the period seismic hazard was not constant with time and space, which is the consequence of filed operations and the variations of both seismicity rate and b-value. The Application to induced seismicity requires some modification to general approach of PSHA. The first reason is the occurrence of earthquakes cannot be constant over the small time-windows because of variations in injection and production rate. Thus the ν parameter as well as the b-value of Gutenberg-Richter relationship vary with time and the hazard integral will be modified from equation 2.16 to :

$$\lambda_{y*} = \sum_{i=1}^{N_S} \iiiint \nu_i(t) P[Y > y^* | m, r, \varepsilon] f_{Mi}(m, b_i(t)) f_{Ri}(r) f_\varepsilon(\varepsilon) dt dm dr d\varepsilon \tag{4.14}$$

where t ranges between $(T, T + \triangle t)$, which corresponds to the time window of interest. The second reason is because of the limited dimension of the seismogenic volume an upper-bound maximum magnitude m_{max}. Thus the upper and truncated truncated probability density function *f(m)* is used. Specifically for The Geysers, probability density function *f(r)* can be assumed uniform (Van Eck et al. (2006)).

4.5.1. Computation of GMPE

The data used here for regression correspond to PGA measured as largest among two horizontal values from 220 earthquakes recorded at 29 LBNL station from September 2007 through November 2010 (see figure 4.36).

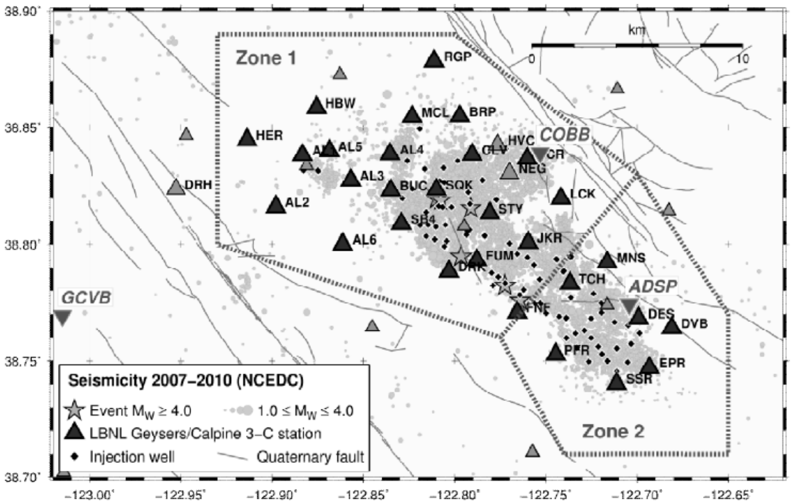


Figure 4.36.: Map of induced seismicity seismicity recorded at The Geysers. Black triangles are the Lawrence Berkeley National Laboratory(LBNL) stations. Additional stations of Northern California Seismic Network (NCSN) are represented by Gray triangles. Two zones are outlined by dashed lines. The inverted triangles are three site chosen for site-specific PSHA which are Anderson Springs (ADSP), Cloverdale(GCVB) and Cobb (COBB).

Table 4.8.: Regression coefficients and related uncertainty of equation 4.15

a	b	c	d	h	e	σ
-2.268 ± 0.356	1.276 ± 0.026	-3.528 ± 0.624	0.053 ± 0.029	3.5	0.218 ± 0.014	*0.324*

The range of magnitude is $1.0 < M_W < 3.5$ and the maximum selected depth is 6.0km while the hypocentral distance is ranging between 0.5 and 20 km. the criteria of selecting waveform was the same as discussed in chapter 4. But this time frequency band was from 1 Hz through 25 Hz. The selected GMPE is formulated as:

$$log(PGA) = a + bM + clog(\sqrt{R^2 + h^2}) + dR + es \tag{4.15}$$

where PGA is in m/s^2.The R is hypocentral distance and M is moment magnitude. h is fictitious depth to avoid distance saturation (Joyner and Boore (1981); Emolo et al. (2011)) and s accounts for site effect. The coefficients and uncertainty are listed in table 4.8.

The best h value is obtained by minimizing the total standard error and maximizing the R^2 statistics. For obtained fictitious depth $h = 3.5$ km minimum standard error is σ =0.324 and R^2 is 0.852. The need to develop GMPE specifically for Geysers is that there are strong chances that already published GMPEs would fail to predict peak-ground-motion for shallow and small-magnitude earthquakes (Bommer and Alarcon (2006)).Further details can be found in Convertito et al. (2012).

4.5.2. Seismic zones

As discussed in chapter 2, it is important to decide potential seismic sources responsible for high ground shaking. In Convertito et al. (2012) it is clearly described that the region considered at The Geysers is divided in to two Zones (see chapter 1 also). The two zones ZONE1 and ZONE2 are outlined by dashed lines in figure 4.36. The zone separation is justified analyzing b-values of the Gutenberg-Richter relationship for two areas using the Utsu (1992) test. Utsu test is done to verify that weather the data used to compute b-value come from different population. Convertito et al. (2012) follow Van Eck et al. (2006) to estimate m_{max} for each source zone by using the technique proposed by Makropoulos and Burton (1983), however it is based on a stationary assumption. After analyzing data sets obtained values are, $m_{max} = 4.5$ for ZONE1 and $m_{max} = 3.8$ for ZONE2. Further details can be found in Convertito et al. (2012).

4.5.3. Computation of PSHA

To compute PSHA the seismicity rate νi and b-values were calculated at several time intervals T_{obs} of 1 month and selected four probabilities of exceedance: 30%, 50%, 70% and 90%. The exposure time chosen was of 2 months and these probabilities corresponds to the return period of about 6 months, 3 months, 2 months, and 1 month respectively. Exposure time was chosen two months to account for at-least two variations in time window period to collect data. To predict PGA, GMPE reported in equation 4.15 was used with *s=0* for rock-site conditions. Time-dependent Gutenberg-Richter relationship was chosen with an upper bound on magnitudes. The time varying truncated version of probability density function *f(m,b(t))*, is formulated as:

$$f_M(m, b(t)) = \frac{\beta(t)exp[-\beta(m - m_{min})]}{1 - exp[-\beta(m_{max} - m_{min})]} \tag{4.16}$$

where $\beta(t) = b(t)ln10$, m_{max} corresponds to maximum magnitude value which is 4.5 in ZONE1 and 3.8 in ZONE2 and minimum m_{min} value of interest is set to 1.2 in both zones, which can regraded as minimum magnitude of completeness. To test the PSHA results observation periods chosen (see Convertito et al. (2012)), were 10 August 2008, 6 February 2009, 4 September 2009, 3 March 2010, 1 June 2010, and 29 September 2010. These specific periods were chosen to monitor the different features of seismicity and b-values and to evaluate their influence on PSHA. The hazard map shown in figure 4.37 describes the probability of exceedance of PGA at 30%, 50%, 70%, and 90%. These map clearly illustrate that hazard is not constant with time. With time varying seismic parameters the largest PGA value changes in both time and space. For example the PGA values are as large as 4.5m/s^{2}with 30% of probability of exceedance in two observation periods (10 August 2008 and 6 February 2009) in ZONE1 while the largest PGA values are expected in ZONE2 during later observation periods (3 March 2010 and 10 June 2010).

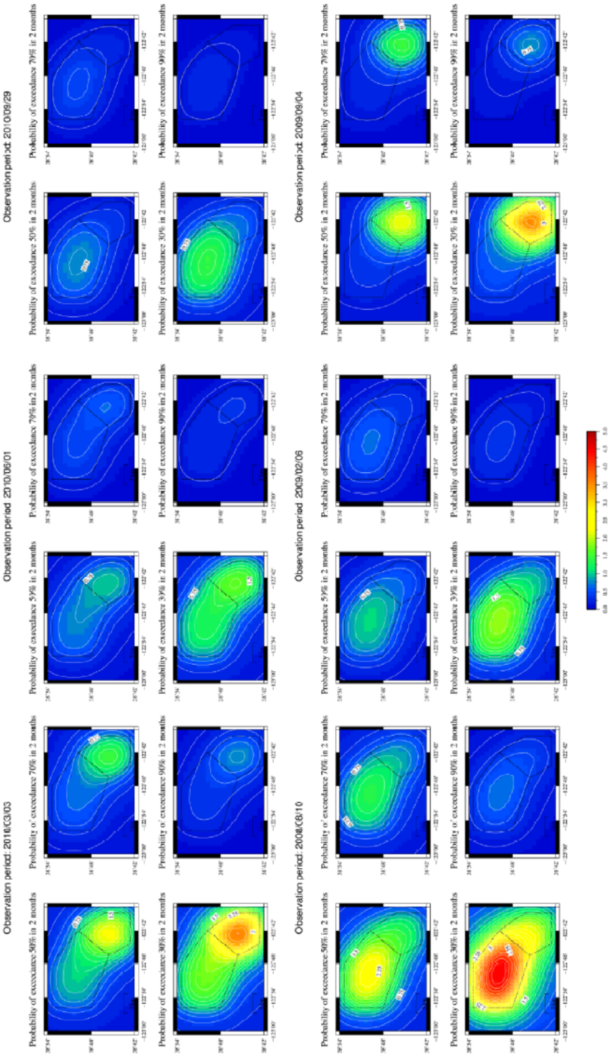


Figure 4.37.: Seismic hazard map reporting the peak-ground motion values having probability of exceedance reported on the top of each map. Each date corresponds to the center time of one-month window centered to that date. Zones are marked with dashed lines. The PGA values are in m/s^2.

4.5.4. Monitoring of Seismic hazard at specific sites

The site-specific seismic hazard analysis is done for 3 sites ADSP, COBB, and GCVB (see figure 4.36)with same probability of exceedance as for the hazard maps discussed in previous section. Authors specified that in the case of hazard associated with low-to-moderate magnitude earthquakes it might be more interesting to consider the medium to highest probabilities instead of considering the lower ones as in the case of standard PSHA because it would be more relevant to know what the actual ground motion level will exceed in near future for monitoring purposes rather than to a assess the values associated with rare events, which contribute more to the lowest probability of exceedance. The results are shown in figure 4.38.

The dashed lines indicate the estimated PGA values resulting from the hazard analysis together with their associated probability of exceedance. The observed PGA values are indicated by the gray squares. The sites experience low-to-moderate shaking for PGA values lower than 1.2m/s^2.The values predicted by hazard analysis are consistent with these observations. It is also observed that the PGA values having 90% of probability of exceedance are systematically exceeded, the PGA values having 30% of probability of exceedance could be used if a cautious value for a more conservative approach is needed.

4.5.5. Concept of traffic light

As discussed in previous section that how time-dependent site-specific hazard analysis can be helpful in monitoring. For example, consider the site ADSP if the effect of induced seismicity or say seismic hazard exceeds a particular fixed threshold then the signal is red which means that the activity should be stopped, if it is just below the threshold then the signal is yellow which means that the activities must me adjusted and if it is reasonably below the threshold then activities can be continued with green signal(see figure 4.39) (Bommer and Alarcon (2006); Convertito et al. (2012)) . This way of monitoring can be used as traffic light controlling the human activities like mining, drilling, injection and extractions of fluid in geothermal areas etc.

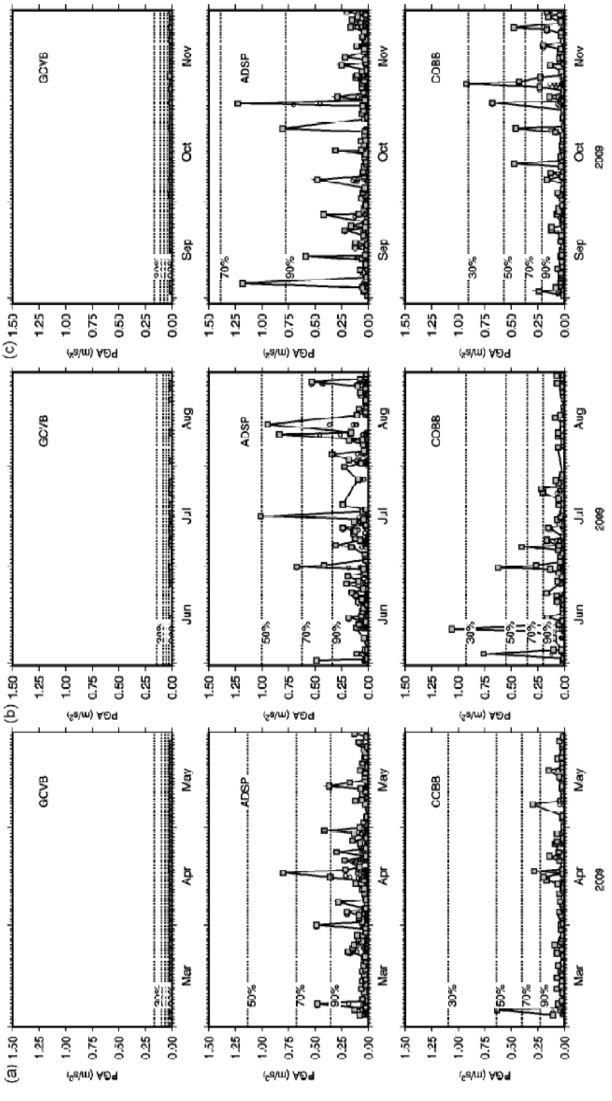


Figure 4.38.: Site-specific seismic hazard analysis for 3 site mentioned in figure 4.36. The dashed lines correspond to the results of the PSHA at the indicated probability of exceedance. Observed PGA values are represented by gray squares. The analysis is performed for three obserevation periods starting 3 March 2009(a), 1 June 2009(b) and 29 September 2009(c).

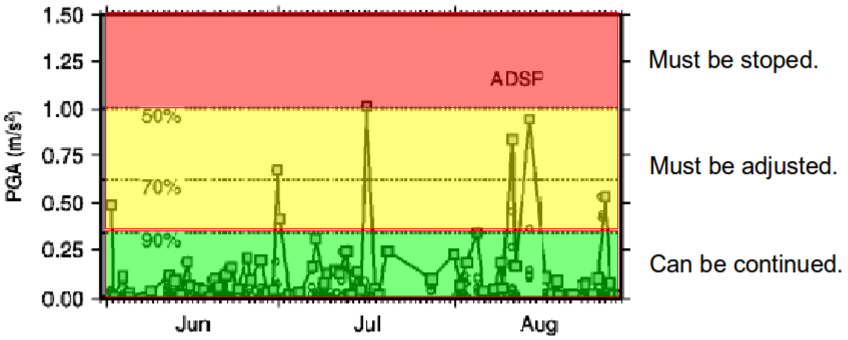


Figure 4.39.: Example showing the concept of traffic light monitoring of seismic hazard for the site ASDP.

5. Time-dependent ground motion prediction equations

In this chapter we discuss about the ongoing work. The concept of time-dependent ground-motion prediction equations inspired from the temporal variations in seismicity due to human intervention in the Geysers geothermal area. Figure 1.3 in chapter 1 is clearly indicating that seismicity is not constant with time. It is changing according to the field operations (Majer et al. (2007)). To understand better, as an example we plot the seismicity maps for period 2007, 2007 to 2008, 2007 to 2009 and 2007 to 2010 of this area, see figures 5.1 and 5.2. We observed that the seismicity is increasing with increase in number of injection wells. The Geysers zone is subdivided into two seismicity source zones named ZONE1 and ZONE2 (dashed lines in figure 1.4). The separation arguments are supported by (Beall et al., 2010; Stark, 2003; Beall and Wright, 2010). A difference in the seismicity distribution has been also noted by (Eberhart-Phillips and Oppenheimer, 1984). Thus the seismicity rate is different in two zones. This time varying seismicity is the effect of time varying stress/strain conditions in upper crust(Majer et al. (2007); Oppenheimer (1986)).

Evolution of new faults and fracture due to drilling of wells introduce more heterogeneity in the medium properties with time. This affects the ground motion recordings as well. (see chapter 4). Hence the time-dependent ground motion prediction equations are important to analyze and interpret the time-dependent changes in ground motion recordings as well as medium properties affecting them. In this chapter we investigate the possibility of analyzing and interpreting the variations in the medium properties due to field operations through the time-dependent ground motion prediction equations. In next sections we will see how time-dependent changes in ground-motion prediction as well as the parameters obtained will be useful to identify the changes in medium properties affecting ground recordings.

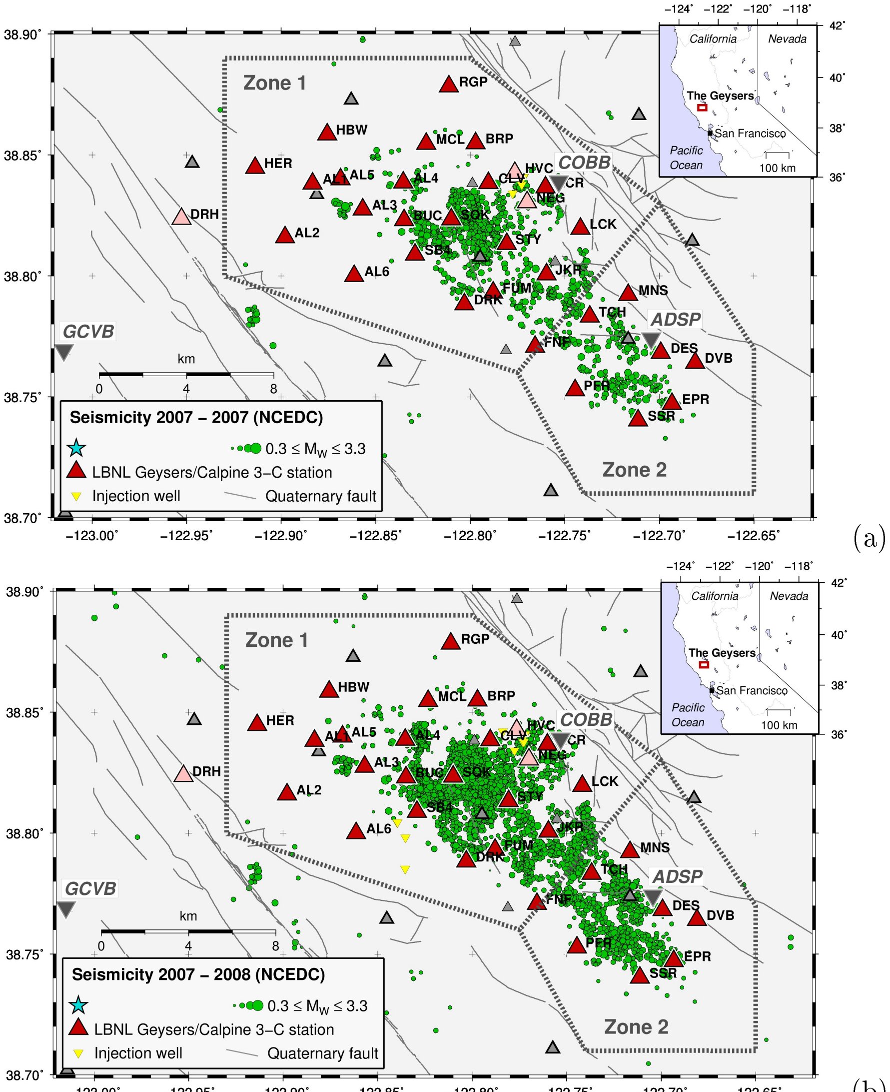


Figure 5.1.: Seismicity map of The Geysers with the time mentioned in the respective figure. It is clearly evident that seismicity is increasing with time.

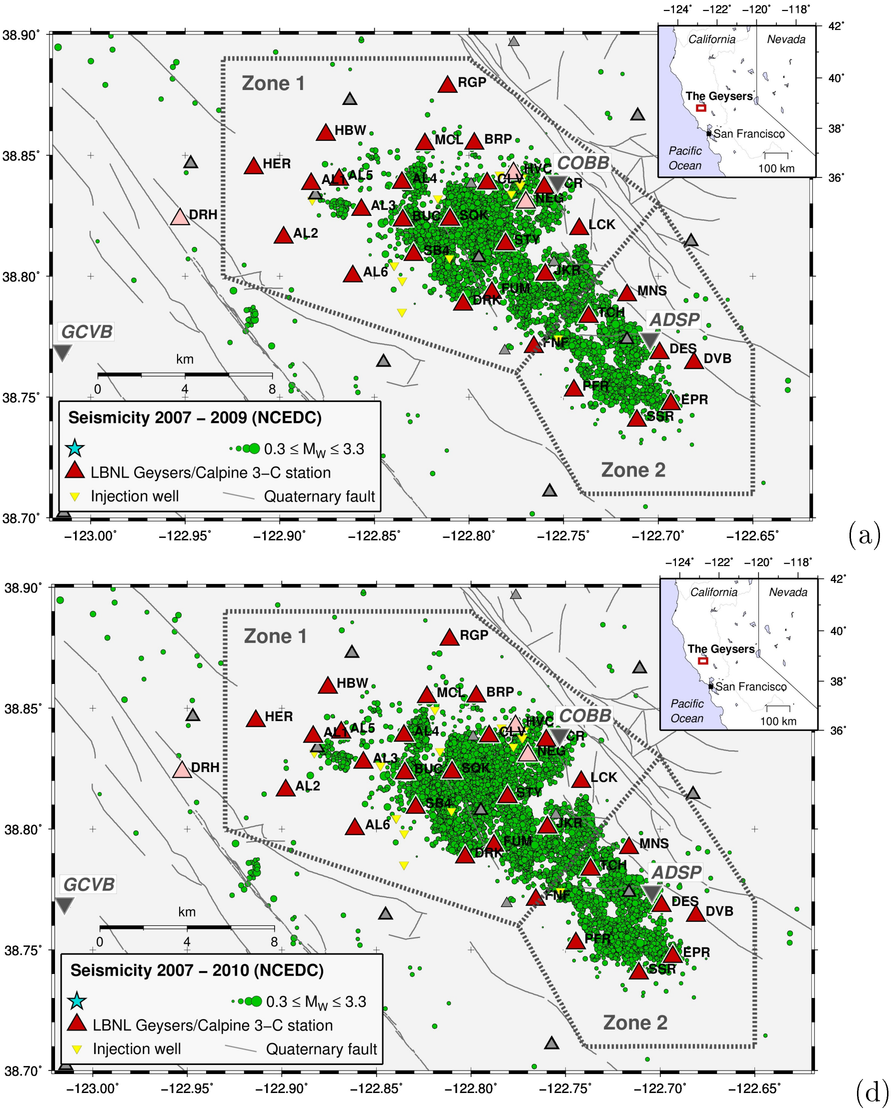


Figure 5.2.: Same as figure 5.1.

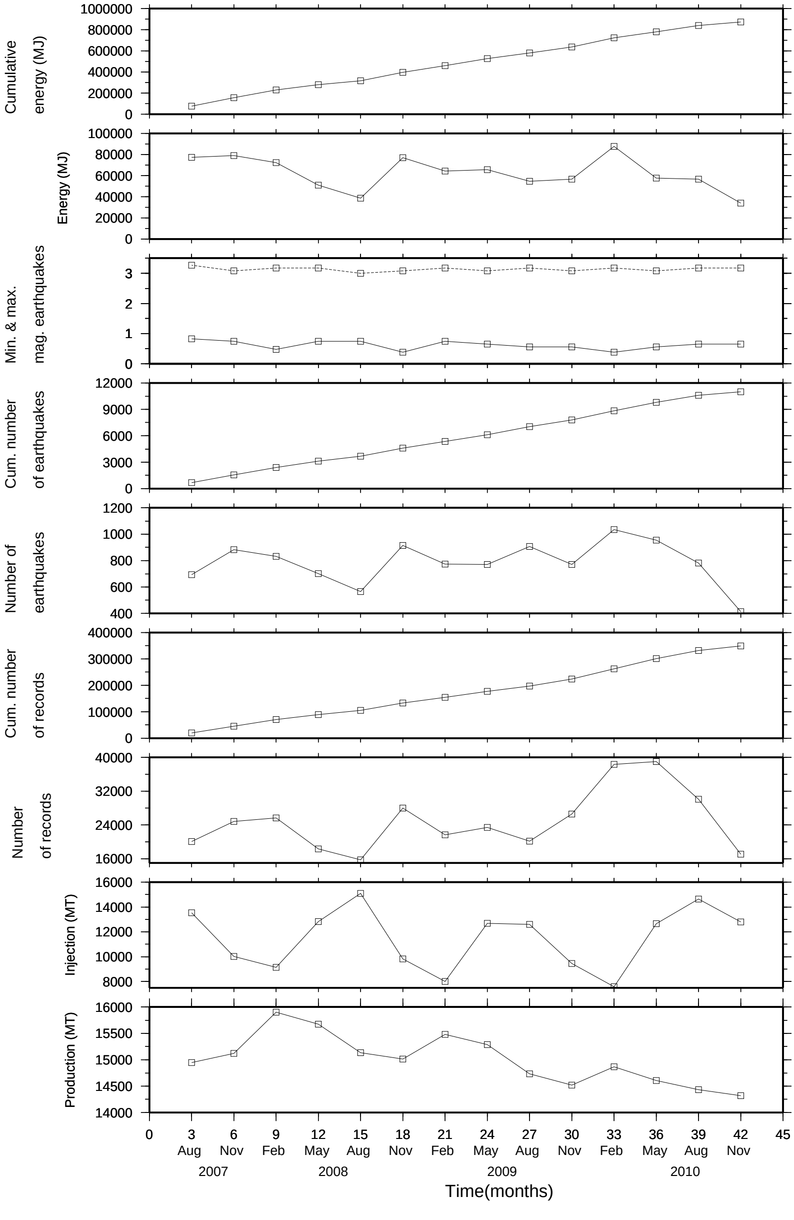


Figure 5.3.: Time-dependent observation of field operation and seismicity.

5.1. Data preparation

The analyzed waveform have been observed in the period 24 July 2007 through 18 November 2010. The waveforms are processed in the same way as mentioned in chapter 4. The magnitude ranges between $0.3 < M_w < 3.3$ and hypocentral distance ranges between $0.1\text{km} < R_{hypo} < 73$ km. Nearly 1 million waveforms from 11,000 events recorded at 29 station are analyzed. In order to infer time-dependent ground motion prediction equations, entire data set is divided in 14 time-windows. The first window contains data of 90 days and every following window contains the data from next 90days in addition to data from all previous windows (e.g. Window1 contains data from first 90 days of catalog, Window2 from 180 days (covering data from last window (Window1) as well), Window3 from 270 days and so no). Thus there is successive addition of data from 90 days in following window. Hence, there are 14 overlapping time-windows and the last window contains data from all 42 months. Time-dependent ground-motion prediction equations are computed using NLMRA for PGV, PGA, SA(T) at T=0.2s, 0.5s, and 1.0s.

We also analyzed the relation between field activities with seismicity in the Geysers geothermal area in each time window. We observed a cyclic negative and positive correlation between fluid injection and number of records (see figure 5.3). The number of records, number of earthquakes and amount of energy released in each time-window increase just after August 2008. Similar trend is observed for cumulative energy as well. But in the same period there is a decrease in fluid injection, thus we think that there is a shift in field activity and occurrence of events. To confirm this observation, we performed correlation analysis between fluid injection and number of events occurred in each window(see figure 5.4). At zero lag we observed a high negative correlation and then negative and positive correlation in cyclic manner. It is also noticed that there is negative correlation during winter months and positive correlation during spring/summer months. It is because the injection rate is changing with the change in seasons, the inject is high during summer and low during winters. The cumulative number of records increase with time and after November 2010 there is the change in trend because of the addition of new stations, same is with the cumulative energy as well. The details are discussed in next section. Magnitude distribution of events with time is

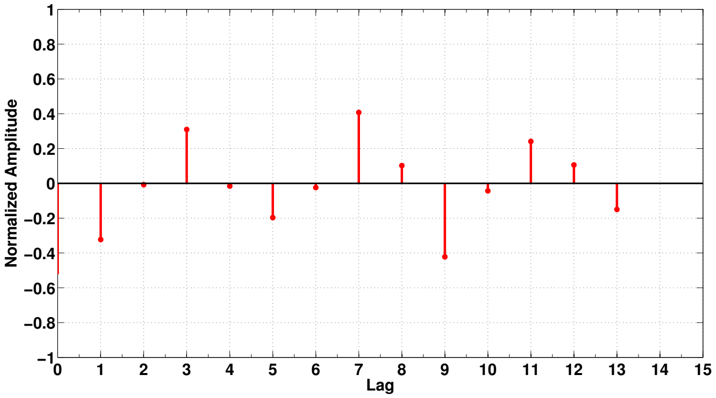


Figure 5.4.: Cross-correlation between fluid injection and induced events in different time windows. A cyclic correlation observed between fluid injection and induced events.

also studied. (see figure 5.5). It can be clearly observed that the number of small magnitude earthquakes (1.5< M_w <1.7) is highest and increases with time which also confirms the similar observation made by Majer et al. (2007) till seismicity of 2005.

5.2. Regression Analysis

The criteria for selecting the model is the same as mentioned in chapter 4. First, we choose the model MOD1(equation 5.1) and then we computed GMPEs for each window. We also considered MOD2 (equation 5.2) to analyze the role of an-elastic attenuation. Second, we analyze residual distribution at each station for each window. Through Z-test on the mean values we estimated site/station correction factor. The final model obtained after site/station correction is MOD3(equation 5.3).

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) \tag{5.1}$$

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) + dR_{hypo} \tag{5.2}$$

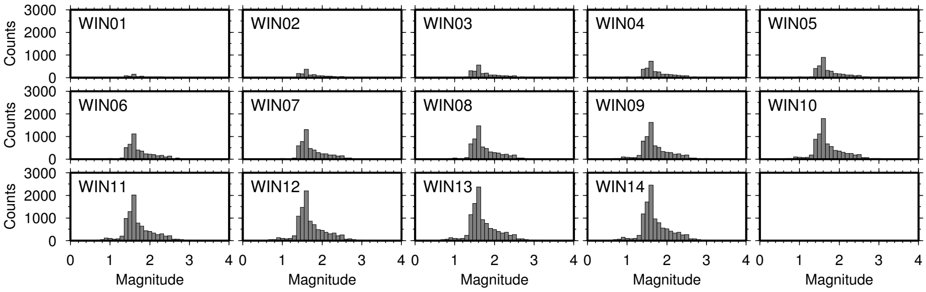


Figure 5.5.: Magnitude distribution in each time-window.

$$Log_{10}(Y) = a + bM_w + cLog_{10}(\sqrt{R_{hypo}^2 + h^2}) + es \tag{5.3}$$

The estimated coefficients for MOD1 at each window are listed in tables 5.1, 5.2, 5.3, 5.4, 5.5. The estimated coefficients for MOD2 at each window are listed in tables 5.6, 5.7, 5.8, 5.9, 5.10.

As discussed before that MOD2 is different from MOD1 because MOD2 considers the effect of an-elastic attenuation, however after comparing the respective standard errors and R^2 statistics, we observed that there is no improvement in the model (see chapter 4), and thus we choose MOD1 as starting model because it is described by less number of parameters.

We introduced site/station effect correction as mentioned in chapter 4 and retrieved s parameter along with a new coefficient e which is then used to set up an updated model, that is, the corrected model MOD3. The coefficients are listed in tables 5.11, 5.12, 5.13, 5.14, 5.15. If we compare the results between MOD1 and MOD3, the standard deviation has reduced for each period and R^2 statistics has increased, which indicate that the goodness of fit of model(MOD3) has improved in each window.

5.3. Time-dependent residual analysis

Inter-event and intra-event residual distributions for MOD1 are shown in figures 5.6 and 5.7. Inter-event and Intra-event residuals shows uniform distribution, however, the spread increases with time and uniformity is maintained for PGA and PGV. But for spectral acceleration SA(T), T=0.2s, 0.5 s and 1.0s inter-event residual shows uniform distribution for first 10 windows (November 2009) but with time (from window11(February 2010) to window14(November 2010)) another peak emerges, which is expected to be correspond to addition large number of small magnitude earthquakes (see figure 5.5) start to evolve from that period . This feature is not evident in PGA and PGV. Intra-event residuals show same behavior at different times for all the periods.

The behavior of inter-event residual distribution is same but there is no significant change after using MOD3. However, the spread of intra-event residuals have reduced significantly for each window thus emphasizing that that station/site effect correction is effective in improving the model fitting (see figures 5.8 and 5.9). It was also observed that there is significant reduction in intra-event uncertainty as it accounts for site effect (see chapter 2 and 4).

5.4. Time-dependent GMPEs and ergodic process

The GMPEs obtained at each time windows are shown in figures 5.10, 5.11, 5.12. The ground-motion prediction equations are shown for $M_w = 2.5$ and black circles represents the observed ground motion parameters for the considered periods. The idea behind analyzing GMPEs with time is to understand if and how ground motion predictions are changing with the addition of new information. To this aim we performed T-test on the mean values and F-test on total standard deviations and in most of the cases we found that null hypothesis stating that the mean prediction are equal with time failed to reject and null hypotheses stating that standard deviations are equal is rejected. Thus standard deviations are not constant with time. But some detailed analysis is still required. As we know that standard deviations of the GMPEs are determined mainly by misfit between observations and predictions at multiple stations for a set of well recorded earthquakes. Thus the standard deviation is related to spatial variability of ground motion.

In seismic hazard generally it is assumed that occurrence of earthquake is an ergodic process (Anderson and Brune (2000)), which is a random process in which the distribution of a random variable in space is the same as the distribution of that same random process at a single point

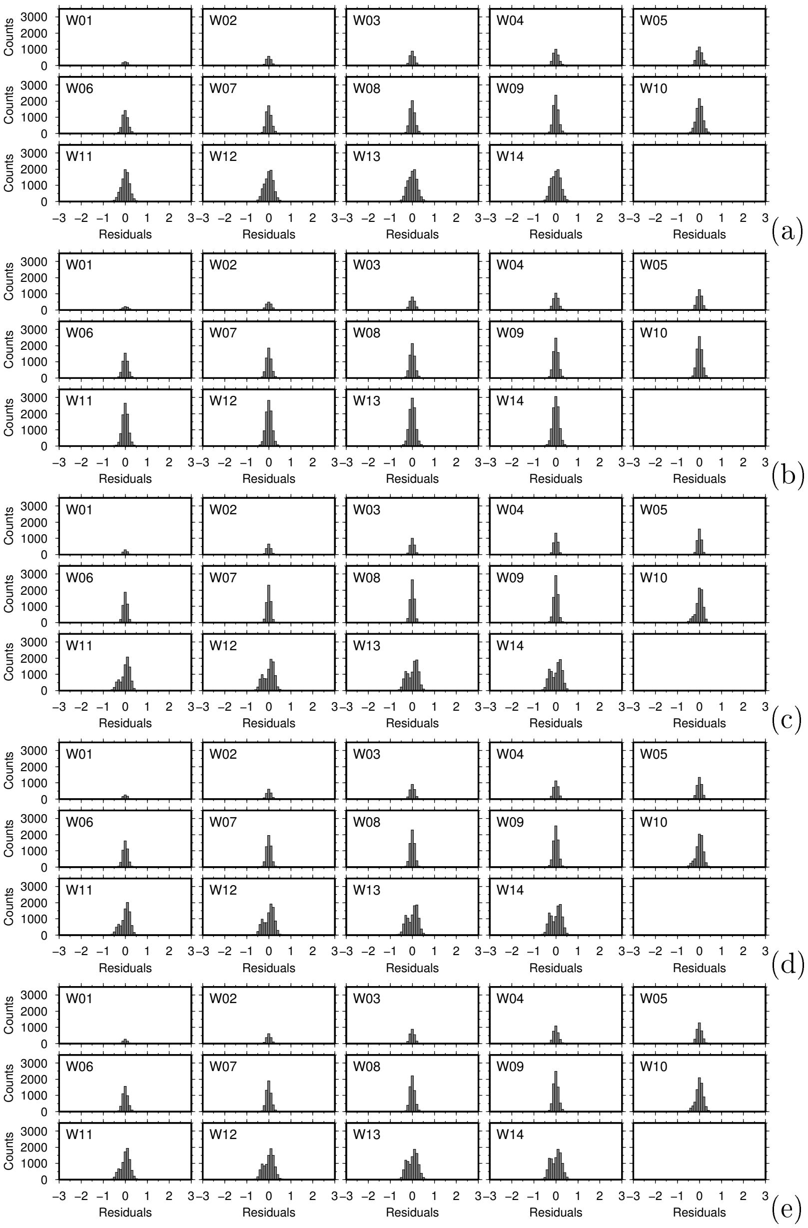


Figure 5.6.: Inter-event residual distribution for each time-window for PGV (a), PGA (b), SA(T=0.2s) (c), SA(T=0.5s) (d), SA(T=1.0s) (e), for MOD1.

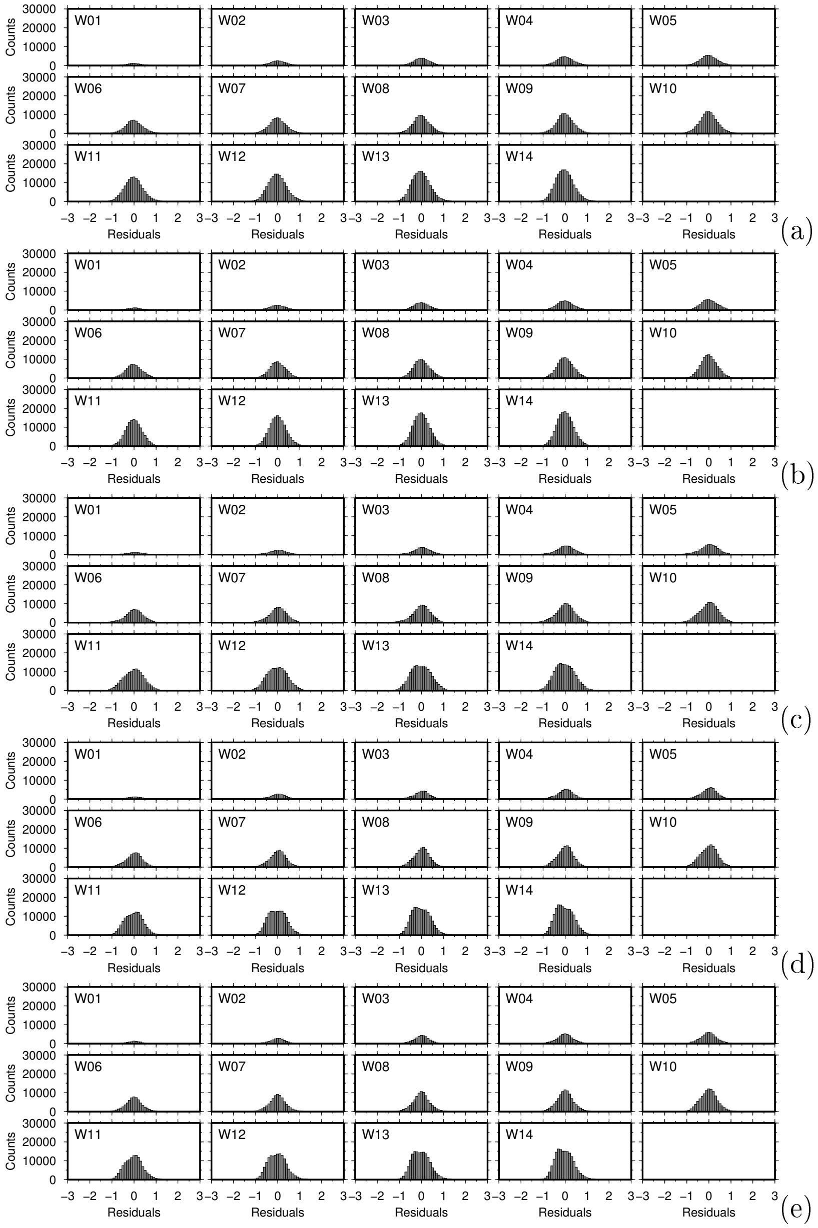


Figure 5.7.: Intra-event residual distribution for each time-window for PGV (a), PGA (b), SA(T=0.2s) (c), SA(T=0.5s) (d), SA(T=1.0s) (e), for MOD1.

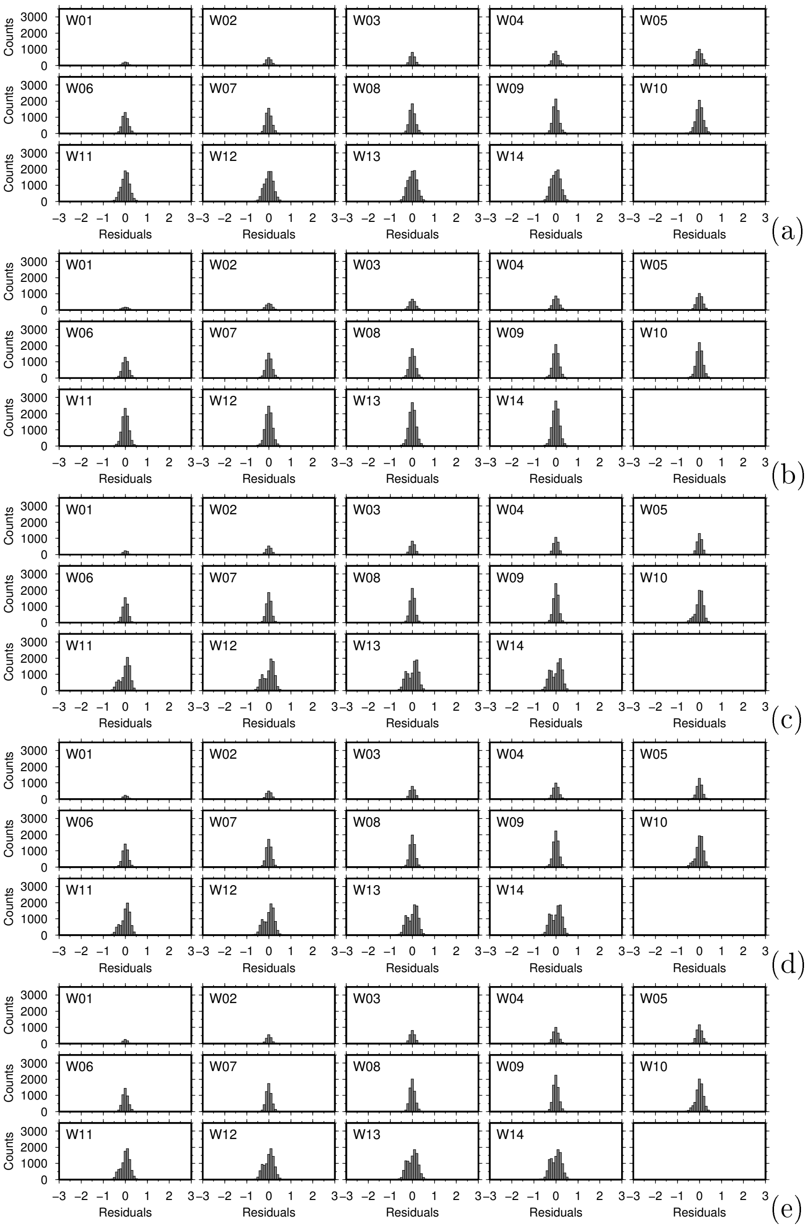


Figure 5.8.: Inter-event residual distribution for each time-window for PGV (a), PGA (b), SA(T=0.2s) (c), SA(T=0.5s) (d), SA(T=1.0s) (e), for MOD3. No significant change is observed after station/site correction factor.

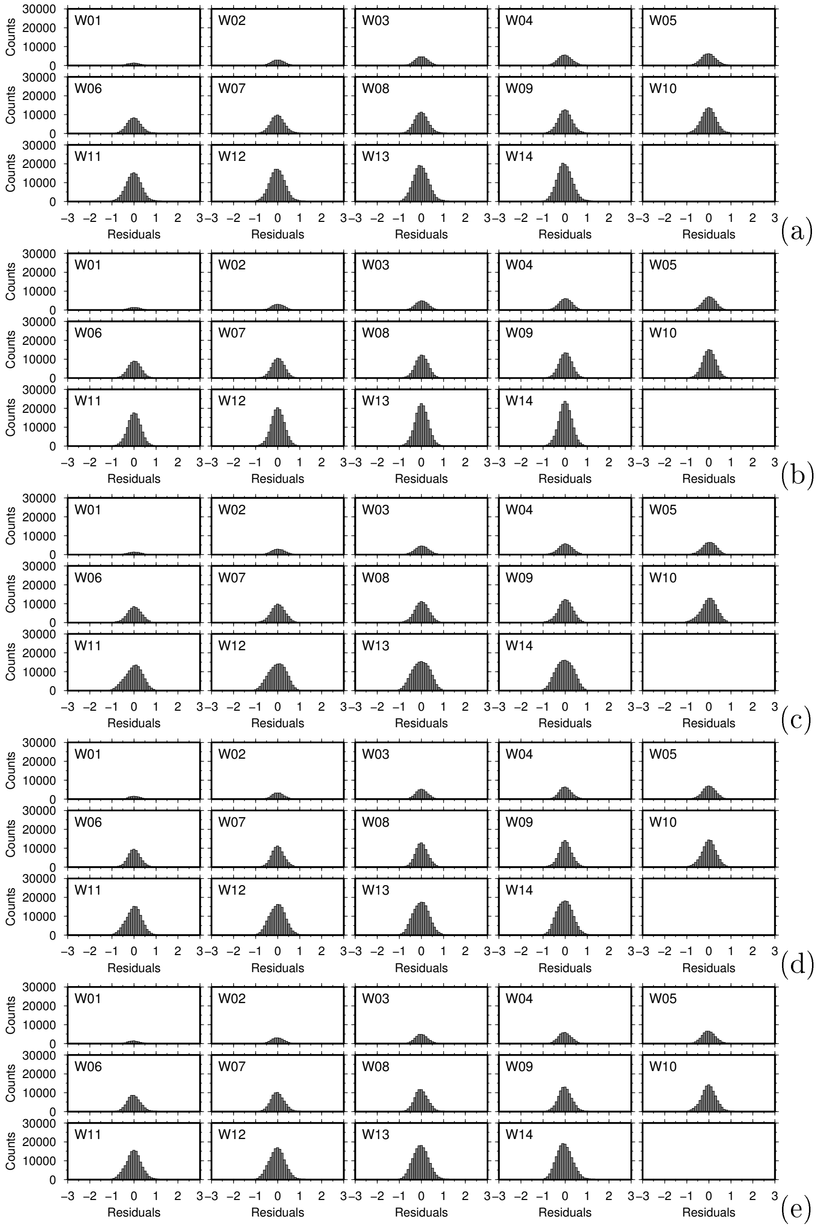


Figure 5.9.: Intra-event residual distribution for each time-window for PGV (a), PGA (b), SA(T=0.2s) (c), SA(T=0.5s) (d), SA(T=1.0s) (e), for MOD3. A significant improvement in fitting is observed as the spread of residuals have reduced.

when sampled as a function of time. Thus ergodic assumption uses standard deviations to describe the temporal distribution of ground motion at the single site over multiple earthquakes (Anderson and Brune (1998, 999a,b)). Ergodic assumption is appropriate for natural seismicity where exposure times are less than the earthquake return period. But this assumption seems inappropriate in case of induced seismicity, because in case of induced seismicity the exposure times are longer than earthquake return period and this can lead to overestimation of ground motion hence the seismic hazard (Anderson and Brune (2000)). Thus ergodic assumption should be avoided when PSHA concerns induced seismicity. Further detailed analysis on this subject will make the results more clear.

5.5. Time-dependent parametric analysis

The parameters obtained for each window for MOD3 are analyzed as a function of time for each period. These coefficients accounts for the dependencies of source and medium which play a major role in controlling ground motions. We found that these coefficients are varying with time (see figures 5.13, 5.14, 5.15, 5.16, 5.17). It is interesting to note that a and h are varying in similar way which is however opposite to b and c parameter variations for all the observed periods (PGA, PGV and SA(T)). But variations in parameter are opposite for high frequencies (PGA) as compared with low frequencies (PGV and SA(T)). For example, *a and h* parameters are increasing with time in case of PGA and decreasing for PGV and SA(T) while *b and c* parameters are decreasing with time for PGA and increasing for SA(T). These time varying changes can be ascribe to the change in source and path effect with the evolution of seismicity. The addition of new events in each time-window is changing the magnitude and the hypocentral range. Thus contributing the variations in ground motion recordings with time. The e parameter accounting for site effect is decreasing with time for all the periods. Moreover, it is also interesting to notice that after August 2009 there is a significant change in inter-event and intra-event uncertainties. The inter-event uncertainty increases in this period. It can be related to the fact that the number of events increase during this period and inter-event accounts for the source effect on ground motion. Moreover it can also be ascribed to aleatory component of uncer-

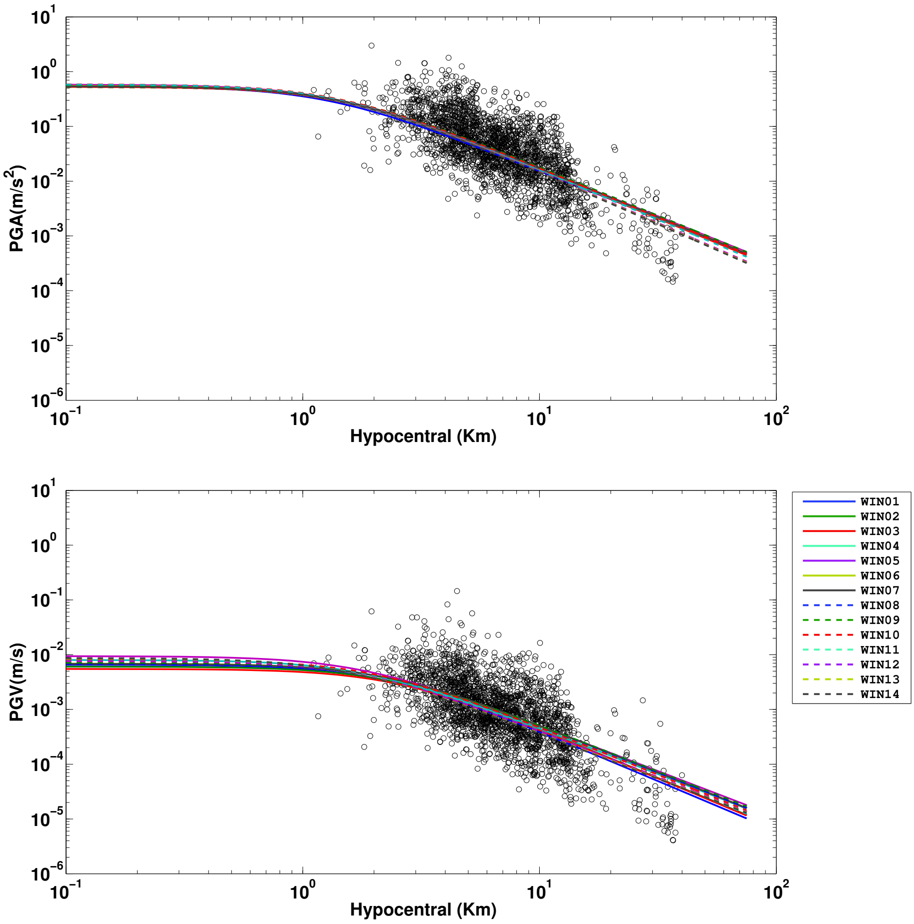


Figure 5.10.: Time-dependent ground motion prediction equations PGA and PGV. They are obtained for event with magnitude M_w=2.5. The data in black circles represent observed ground motion values for the same magnitude.

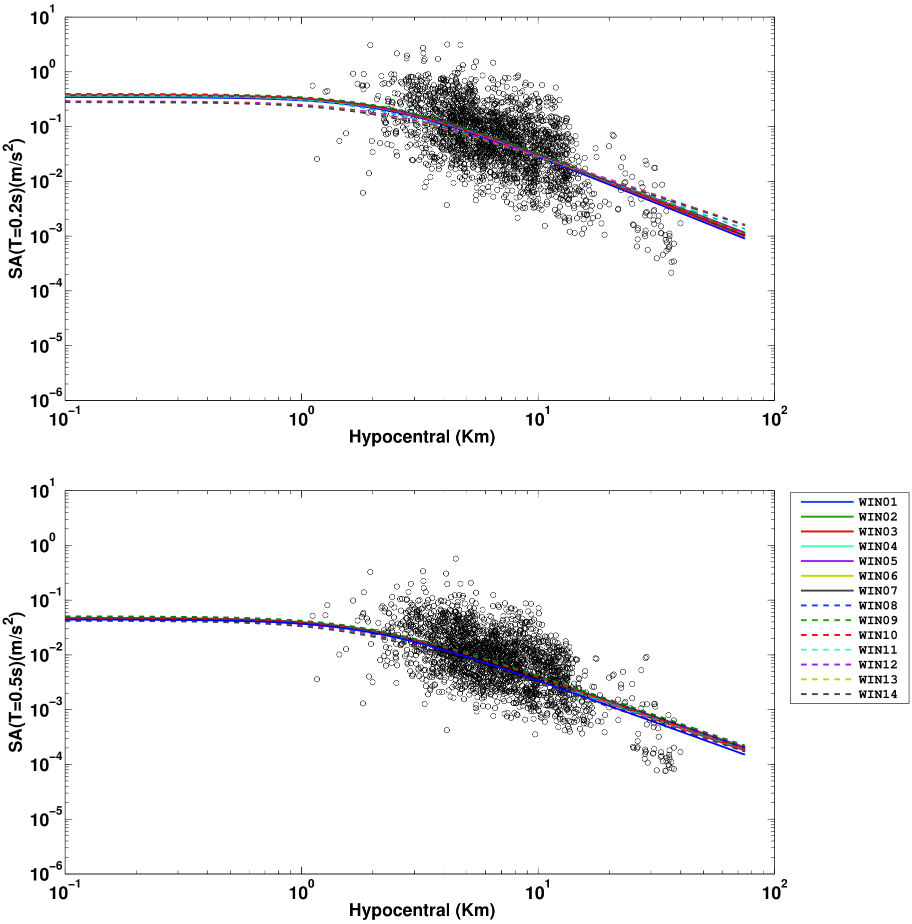


Figure 5.11.: Same as figure 5.10 but for SA(T=0.2s) and SA(T=0.5s).

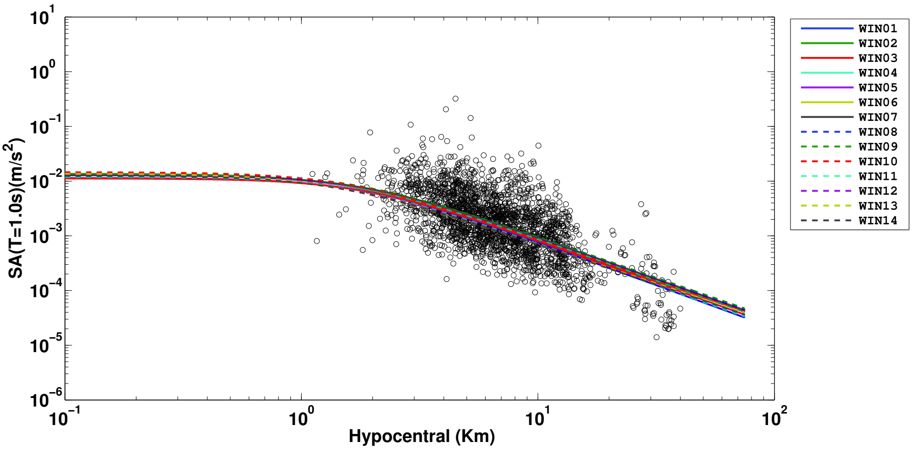


Figure 5.12.: Same as figure 5.10 but for SA(T=1.0s).

tainty which accounts for truly random effects on ground motion observations and the addition of new earthquakes (random distribution of sources) has increased this effect (Anderson and Brune (2000)). On the other hand the intra-event uncertainty accounts for the path and the site effect. Thus a decrease in intra-event uncertainty can be ascribed to epistemic uncertainty component (due to lack of knowledge) which decreases after the addition of new information. (repeatability in the path and site response) (Anderson and Brune (2000)). Thus overall increase in total standard deviation indicates that the effect of inter-event (aleatory) component is more than intra-event (epistemic) component. In conclusion, time-dependent analysis of ground-motion prediction equations has shown interesting and promising results. We found that the ground motion-prediction changes with time because of change in medium properties.

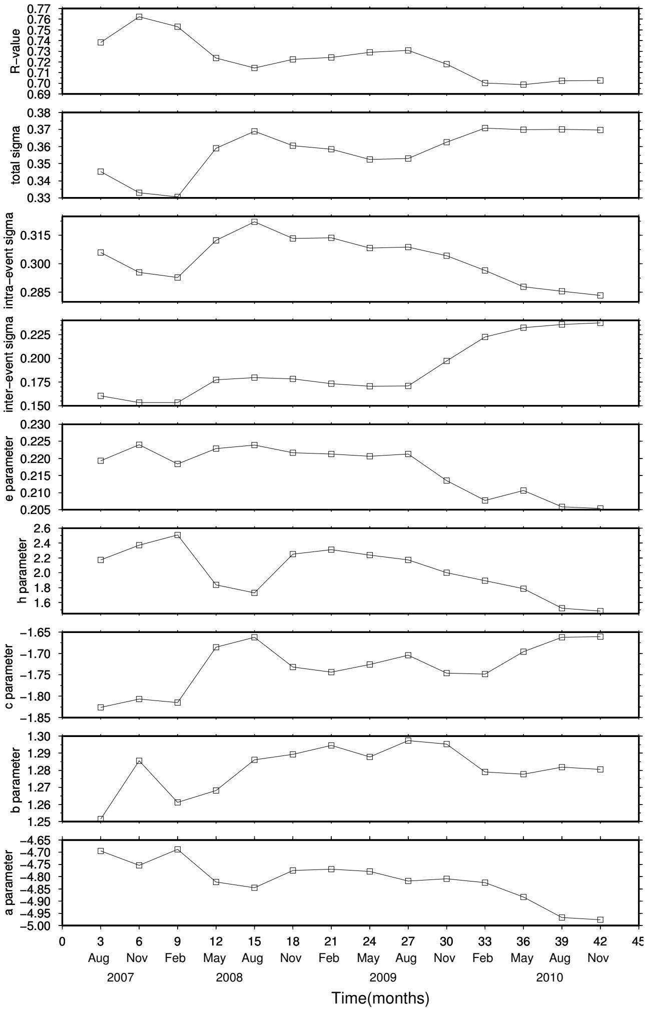


Figure 5.13.: Time-dependent variations in coefficients of PGV, obtained after using MOD3.

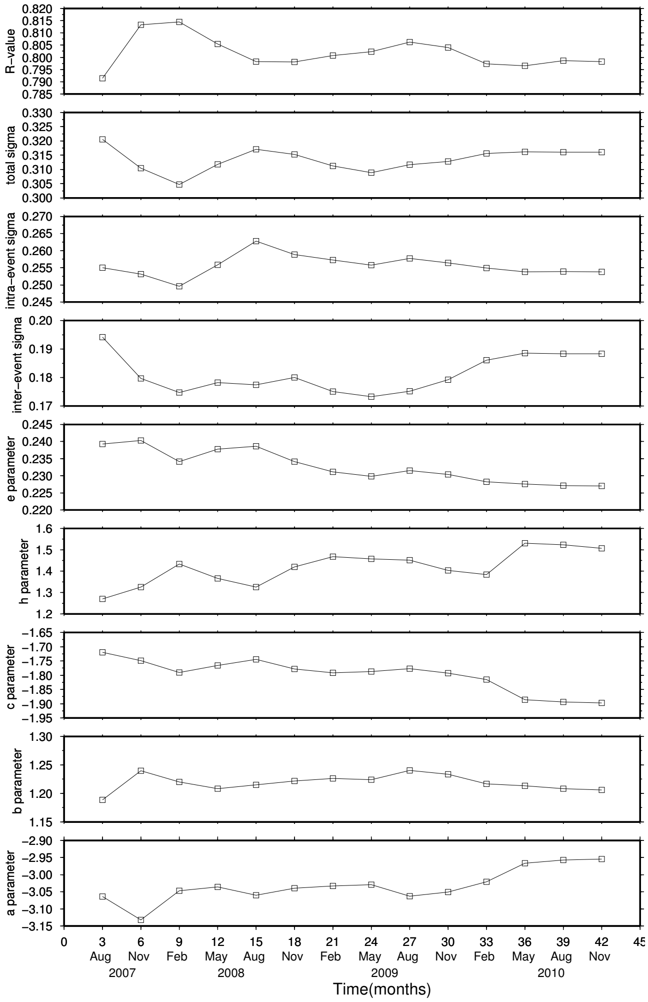


Figure 5.14.: Time-dependent variations in coefficients of PGA, obtained after using MOD3.

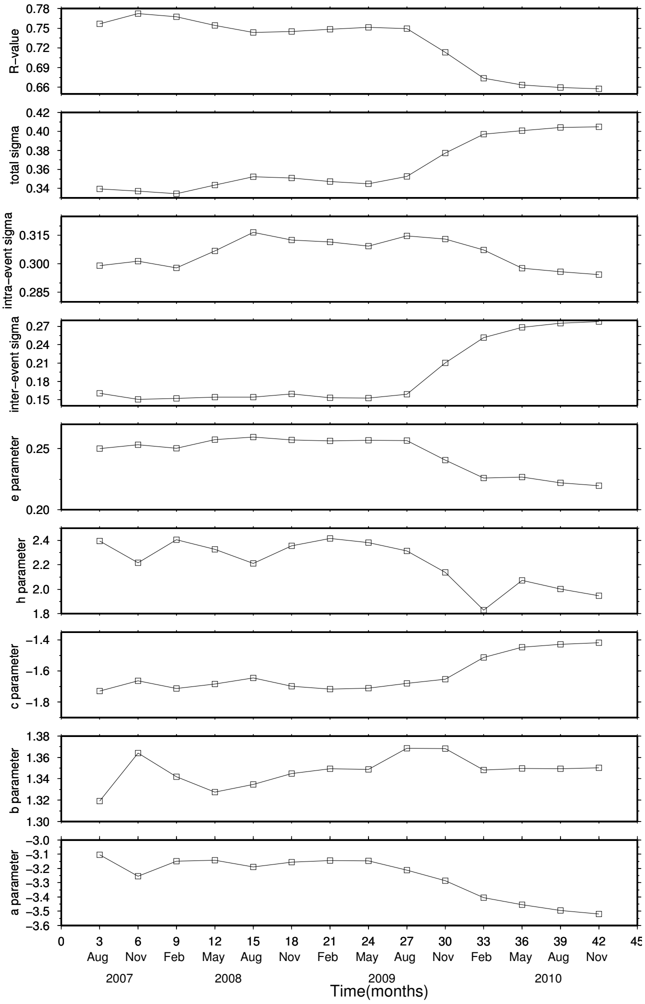


Figure 5.15.: Time-dependent variations in coefficients of SA(T=0.2s), obtained after using MOD3.

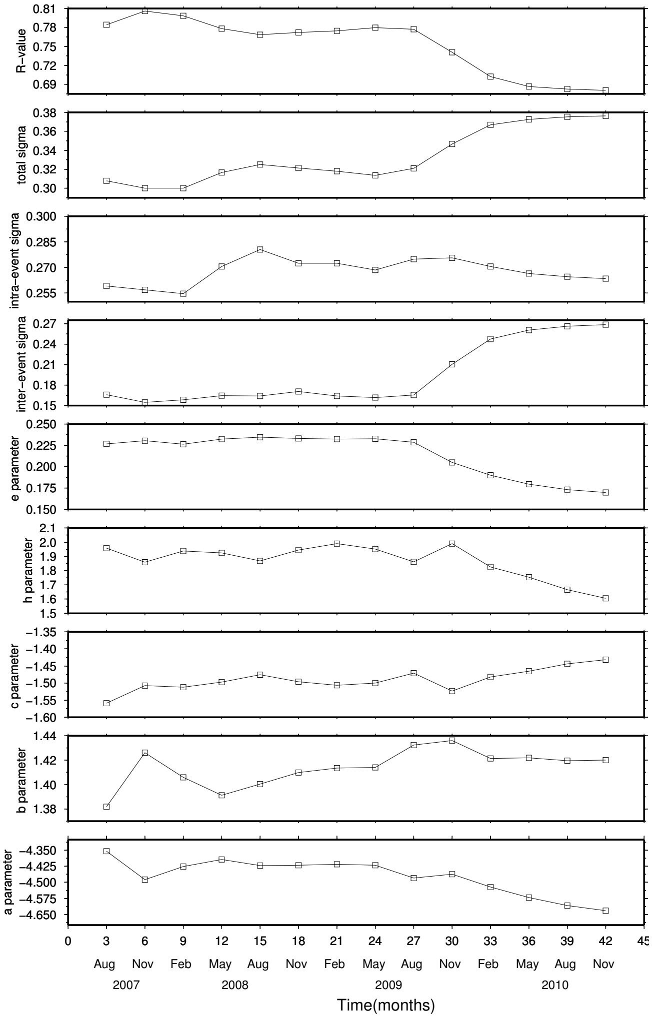


Figure 5.16.: Time-dependent variations in coefficients of SA(T=0.5s), obtained after using MOD3.

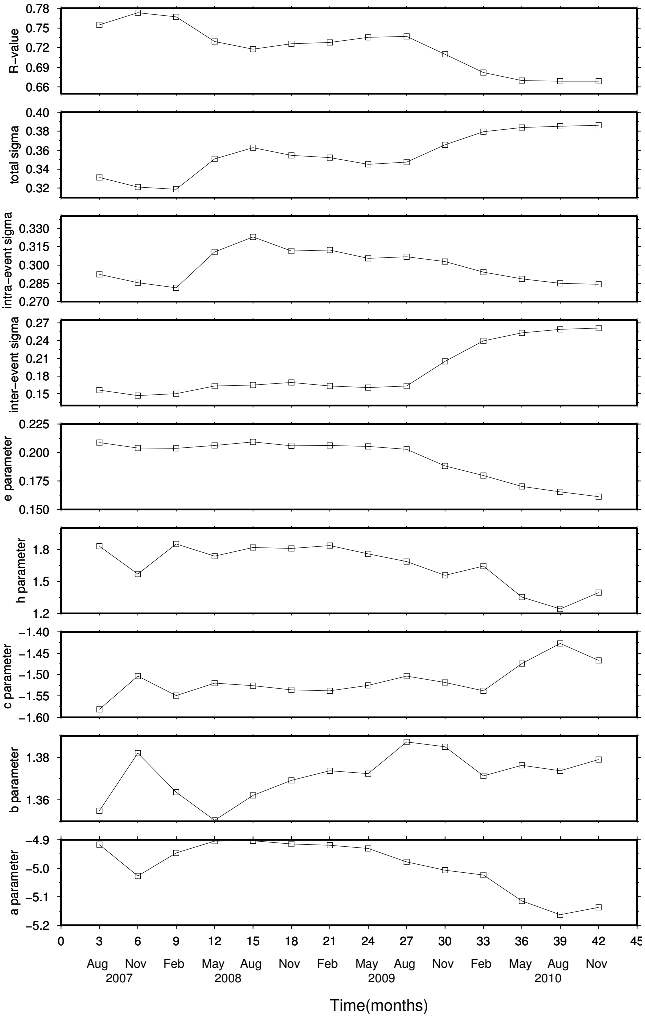


Figure 5.17.: Time-dependent variations in coefficients of SA(T=1.0s), obtained after using MOD3.

Table 5.1.: Estimated coefficients with their uncertainty at each window for PGV using equation 5.1.

Windows	a	b	c	h	σ_a	σ_b	σ_c	σ_h	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.576	1.227	-1.896	2.292	0.061	0.018	0.054	0.188	0.145	0.380	0.406	0.631
2	-4.706	1.259	-1.822	2.116	0.038	0.011	0.034	0.123	0.137	0.374	0.398	0.645
3	-4.677	1.242	-1.809	2.089	0.031	0.009	0.027	0.099	0.140	0.368	0.394	0.641
4	-4.663	1.241	-1.783	2.009	0.029	0.009	0.025	0.094	0.161	0.387	0.419	0.611
5	-4.675	1.259	-1.784	2.035	0.027	0.009	0.024	0.090	0.164	0.396	0.429	0.603
6	-4.679	1.266	-1.802	2.096	0.024	0.008	0.021	0.078	0.164	0.387	0.420	0.613
7	-4.698	1.272	-1.795	2.081	0.022	0.007	0.019	0.073	0.159	0.387	0.419	0.615
8	-4.711	1.265	-1.769	2.008	0.020	0.007	0.017	0.067	0.156	0.382	0.413	0.619
9	-4.748	1.272	-1.745	1.948	0.018	0.006	0.016	0.063	0.157	0.383	0.414	0.619
10	-4.753	1.268	-1.779	1.864	0.018	0.007	0.014	0.056	0.189	0.374	0.419	0.614
11	-4.760	1.253	-1.798	1.776	0.017	0.007	0.013	0.050	0.219	0.364	0.425	0.604
12	-4.785	1.253	-1.807	1.704	0.016	0.007	0.011	0.046	0.231	0.356	0.424	0.602
13	-4.794	1.250	-1.814	1.675	0.015	0.007	0.011	0.043	0.234	0.352	0.423	0.605
14	-4.801	1.249	-1.817	1.658	0.015	0.006	0.010	0.041	0.236	0.350	0.422	0.606

Table 5.2.: Estimated coefficients with their uncertainty at each window for PGA using equation 5.1.

Windows	a	b	c	h	σ_a	σ_b	σ_c	σ_h	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-2.702	1.153	-2.035	2.107	0.056	0.019	0.048	0.156	0.162	0.352	0.388	0.658
2	-2.831	1.195	-1.977	1.888	0.035	0.012	0.030	0.103	0.154	0.352	0.384	0.671
3	-2.806	1.180	-1.968	1.847	0.028	0.010	0.024	0.082	0.152	0.345	0.377	0.671
4	-2.814	1.173	-1.937	1.714	0.024	0.009	0.021	0.075	0.152	0.353	0.384	0.659
5	-2.841	1.181	-1.918	1.651	0.022	0.008	0.019	0.071	0.152	0.358	0.389	0.653
6	-2.835	1.190	-1.941	1.719	0.020	0.007	0.017	0.062	0.155	0.352	0.385	0.659
7	-2.849	1.196	-1.938	1.712	0.018	0.007	0.015	0.057	0.151	0.349	0.380	0.664
8	-2.855	1.193	-1.924	1.676	0.017	0.006	0.014	0.053	0.150	0.347	0.378	0.667
9	-2.880	1.201	-1.906	1.644	0.016	0.006	0.013	0.050	0.151	0.350	0.381	0.665
10	-2.853	1.195	-1.943	1.644	0.015	0.006	0.012	0.046	0.157	0.348	0.381	0.666
11	-2.825	1.180	-1.973	1.610	0.014	0.006	0.011	0.042	0.169	0.344	0.384	0.663
12	-2.821	1.177	-1.989	1.573	0.013	0.005	0.010	0.039	0.172	0.343	0.384	0.662
13	-2.811	1.172	-2.002	1.568	0.013	0.005	0.010	0.037	0.173	0.343	0.384	0.664
14	-2.806	1.171	-2.010	1.565	0.012	0.005	0.009	0.036	0.174	0.342	0.384	0.664

Table 5.3.: Estimated coefficients with their uncertainty at each window for SA(T), at T=0.2s, using equation 5.1.

Windows	a	b	c	h	σ_a	σ_b	σ_c	σ_h	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-2.996	1.291	-1.753	2.403	0.064	0.017	0.058	0.215	0.133	0.394	0.416	0.628
2	-3.143	1.332	-1.676	2.185	0.040	0.011	0.037	0.144	0.129	0.397	0.417	0.637
3	-3.102	1.314	-1.671	2.200	0.033	0.009	0.030	0.117	0.131	0.392	0.414	0.630
4	-3.121	1.303	-1.623	1.981	0.028	0.008	0.025	0.107	0.130	0.404	0.424	0.609
5	-3.159	1.311	-1.590	1.865	0.026	0.008	0.023	0.101	0.130	0.412	0.432	0.599
6	-3.137	1.323	-1.635	2.033	0.024	0.007	0.022	0.089	0.136	0.408	0.430	0.604
7	-3.147	1.328	-1.637	2.029	0.022	0.006	0.020	0.082	0.130	0.407	0.427	0.608
8	-3.155	1.326	-1.620	1.983	0.020	0.006	0.018	0.077	0.129	0.405	0.425	0.609
9	-3.209	1.342	-1.597	1.935	0.019	0.006	0.017	0.073	0.138	0.409	0.432	0.609
10	-3.255	1.344	-1.609	1.845	0.019	0.007	0.015	0.066	0.200	0.397	0.445	0.593
11	-3.310	1.336	-1.596	1.779	0.018	0.008	0.013	0.060	0.249	0.384	0.457	0.572
12	-3.375	1.340	-1.577	1.710	0.017	0.008	0.012	0.055	0.268	0.374	0.460	0.565
13	-3.409	1.340	-1.567	1.683	0.017	0.008	0.011	0.052	0.275	0.369	0.461	0.566
14	-3.428	1.341	-1.563	1.671	0.016	0.008	0.011	0.051	0.279	0.366	0.460	0.565

Table 5.4.: Estimated coefficients with their uncertainty at each window for SA(T), at T=0.5s, using equation 5.1.

Windows	a	b	c	h	σ_a	σ_b	σ_c	σ_h	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.221	1.362	-1.583	2.203	0.055	0.017	0.048	0.199	0.143	0.343	0.371	0.686
2	-4.340	1.401	-1.523	2.116	0.036	0.011	0.031	0.135	0.135	0.342	0.367	0.701
3	-4.313	1.387	-1.508	2.113	0.029	0.009	0.025	0.110	0.141	0.338	0.366	0.693
4	-4.301	1.375	-1.479	1.974	0.026	0.008	0.022	0.103	0.145	0.354	0.383	0.667
5	-4.327	1.384	-1.457	1.893	0.024	0.008	0.021	0.098	0.146	0.363	0.391	0.657
6	-4.328	1.396	-1.480	2.017	0.022	0.007	0.019	0.086	0.151	0.356	0.386	0.663
7	-4.341	1.400	-1.476	2.007	0.020	0.007	0.017	0.080	0.145	0.356	0.384	0.665
8	-4.356	1.400	-1.457	1.948	0.018	0.006	0.016	0.074	0.143	0.353	0.381	0.669
9	-4.412	1.416	-1.431	1.865	0.017	0.006	0.015	0.071	0.149	0.355	0.385	0.671
10	-4.452	1.416	-1.441	1.745	0.017	0.007	0.013	0.063	0.201	0.343	0.398	0.653
11	-4.499	1.404	-1.431	1.628	0.016	0.007	0.011	0.056	0.247	0.331	0.413	0.627
12	-4.552	1.407	-1.421	1.543	0.016	0.007	0.010	0.052	0.264	0.321	0.416	0.619
13	-4.580	1.405	-1.413	1.499	0.015	0.007	0.009	0.049	0.271	0.316	0.416	0.620
14	-4.598	1.406	-1.409	1.471	0.015	0.007	0.009	0.047	0.274	0.314	0.416	0.619

Table 5.5.: Estimated coefficients with their uncertainty at each window for SA(T), at T=1.0s, using equation 5.1.

Windows	a	b	c	h	σ_a	σ_b	σ_c	σ_h	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.742	1.333	-1.681	2.095	0.055	0.017	0.048	0.191	0.142	0.360	0.387	0.667
2	-4.868	1.365	-1.605	1.954	0.035	0.011	0.031	0.128	0.133	0.353	0.377	0.685
3	-4.823	1.347	-1.599	1.985	0.028	0.009	0.025	0.104	0.139	0.347	0.374	0.677
4	-4.786	1.334	-1.578	1.920	0.027	0.009	0.023	0.101	0.152	0.375	0.404	0.637
5	-4.791	1.346	-1.577	1.948	0.026	0.008	0.022	0.098	0.154	0.386	0.416	0.625
6	-4.801	1.355	-1.591	2.007	0.023	0.008	0.020	0.084	0.157	0.375	0.407	0.636
7	-4.822	1.360	-1.581	1.976	0.021	0.007	0.018	0.079	0.152	0.376	0.406	0.637
8	-4.842	1.359	-1.558	1.899	0.019	0.006	0.016	0.072	0.148	0.370	0.399	0.644
9	-4.889	1.373	-1.537	1.841	0.017	0.006	0.015	0.068	0.152	0.370	0.400	0.648
10	-4.908	1.372	-1.571	1.759	0.017	0.007	0.013	0.060	0.199	0.358	0.410	0.635
11	-4.934	1.361	-1.584	1.670	0.017	0.007	0.012	0.053	0.241	0.345	0.421	0.616
12	-4.969	1.363	-1.593	1.612	0.016	0.007	0.010	0.049	0.257	0.336	0.423	0.612
13	-4.987	1.362	-1.597	1.587	0.015	0.007	0.010	0.045	0.263	0.330	0.422	0.614
14	-5.002	1.363	-1.597	1.563	0.015	0.007	0.009	0.044	0.265	0.327	0.421	0.615

Table 5.6.: Estimated coefficients with their uncertainty at each window for PGV using equation 5.2.

Windows	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.716	1.230	-1.663	2.035	-0.010	0.105	0.018	0.153	0.277	0.006	0.145	0.380	0.406	0.632
2	-4.774	1.260	-1.705	1.989	-0.005	0.078	0.011	0.122	0.193	0.005	0.137	0.373	0.398	0.646
3	-4.664	1.242	-1.831	2.112	0.001	0.071	0.009	0.113	0.156	0.005	0.140	0.368	0.394	0.641
4	-4.663	1.241	-1.783	2.009	0.000	0.061	0.009	0.094	0.143	0.004	0.161	0.387	0.419	0.611
5	-4.699	1.259	-1.742	1.990	-0.002	0.059	0.009	0.091	0.141	0.004	0.164	0.396	0.429	0.603
6	-4.693	1.266	-1.778	2.070	-0.001	0.054	0.008	0.084	0.124	0.004	0.164	0.387	0.420	0.613
7	-4.720	1.272	-1.757	2.040	-0.002	0.050	0.007	0.078	0.117	0.003	0.159	0.387	0.419	0.615
8	-4.753	1.265	-1.696	1.927	-0.003	0.044	0.007	0.070	0.110	0.003	0.156	0.382	0.413	0.619
9	-4.760	1.272	-1.724	1.923	-0.001	0.040	0.006	0.062	0.099	0.003	0.157	0.383	0.414	0.619
10	-4.746	1.268	-1.792	1.878	0.001	0.037	0.007	0.058	0.089	0.003	0.189	0.374	0.419	0.614
11	-4.707	1.252	-1.893	1.877	0.005	0.035	0.007	0.054	0.078	0.002	0.219	0.364	0.425	0.603
12	-4.697	1.251	-1.965	1.869	0.008	0.033	0.007	0.050	0.069	0.002	0.230	0.356	0.424	0.601
13	-4.688	1.248	-2.003	1.872	0.009	0.031	0.007	0.048	0.065	0.002	0.234	0.352	0.423	0.604
14	-4.693	1.247	-2.010	1.859	0.009	0.030	0.007	0.046	0.063	0.002	0.237	0.350	0.423	0.605

Table 5.7.: Estimated coefficients with their uncertainty at each window for PGA using equation 5.2.

Windows	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-2.818	1.156	-1.839	1.911	-0.009	0.096	0.019	0.141	0.230	0.006	0.162	0.352	0.388	0.660
2	-2.846	1.195	-1.949	1.861	-0.001	0.071	0.012	0.112	0.157	0.005	0.154	0.352	0.384	0.671
3	-2.730	1.180	-2.103	1.971	0.006	0.064	0.010	0.102	0.125	0.004	0.152	0.345	0.377	0.671
4	-2.738	1.172	-2.075	1.847	0.007	0.052	0.009	0.081	0.109	0.004	0.152	0.353	0.384	0.659
5	-2.738	1.179	-2.106	1.831	0.009	0.049	0.008	0.078	0.102	0.003	0.152	0.358	0.389	0.652
6	-2.733	1.188	-2.125	1.893	0.009	0.046	0.007	0.072	0.091	0.003	0.155	0.352	0.385	0.658
7	-2.770	1.194	-2.081	1.850	0.007	0.041	0.007	0.066	0.086	0.003	0.151	0.349	0.380	0.664
8	-2.799	1.192	-2.027	1.777	0.005	0.038	0.006	0.060	0.081	0.003	0.150	0.347	0.378	0.666
9	-2.810	1.199	-2.033	1.772	0.006	0.034	0.006	0.054	0.075	0.002	0.151	0.350	0.381	0.665
10	-2.792	1.194	-2.055	1.755	0.005	0.032	0.006	0.051	0.069	0.002	0.157	0.348	0.382	0.665
11	-2.732	1.178	-2.143	1.775	0.008	0.030	0.006	0.048	0.062	0.002	0.169	0.344	0.384	0.662
12	-2.713	1.175	-2.186	1.763	0.010	0.028	0.005	0.044	0.057	0.002	0.173	0.343	0.384	0.661
13	-2.695	1.170	-2.212	1.770	0.010	0.027	0.005	0.042	0.054	0.002	0.174	0.342	0.384	0.663
14	-2.697	1.169	-2.208	1.757	0.010	0.026	0.005	0.041	0.053	0.002	0.174	0.342	0.384	0.663

Table 5.8.: Estimated coefficients with their uncertainty at each window for SA(T),at T=0.2s using equation 5.2.

Windows	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-3.196	1.294	-1.421	2.006	-0.014	0.105	0.017	0.155	0.331	0.006	0.132	0.394	0.415	0.629
2	-3.255	1.334	-1.482	1.954	-0.009	0.080	0.011	0.126	0.232	0.005	0.128	0.397	0.417	0.637
3	-3.123	1.315	-1.635	2.162	-0.002	0.077	0.009	0.120	0.187	0.005	0.131	0.392	0.414	0.630
4	-3.140	1.303	-1.588	1.941	-0.002	0.059	0.008	0.092	0.162	0.004	0.130	0.404	0.424	0.610
5	-3.159	1.311	-1.592	1.867	0.000	0.055	0.008	0.088	0.153	0.004	0.130	0.412	0.432	0.599
6	-3.157	1.323	-1.600	1.994	-0.002	0.053	0.007	0.083	0.140	0.004	0.135	0.408	0.429	0.605
7	-3.183	1.328	-1.574	1.957	-0.003	0.048	0.006	0.076	0.132	0.003	0.130	0.407	0.427	0.608
8	-3.215	1.327	-1.514	1.856	-0.005	0.044	0.006	0.070	0.126	0.003	0.129	0.405	0.425	0.610
9	-3.257	1.343	-1.510	1.829	-0.004	0.039	0.006	0.062	0.116	0.003	0.138	0.409	0.432	0.609
10	-3.257	1.344	-1.604	1.839	0.000	0.039	0.007	0.061	0.104	0.003	0.200	0.397	0.445	0.593
11	-3.284	1.335	-1.643	1.833	0.002	0.037	0.008	0.057	0.094	0.003	0.249	0.384	0.458	0.572
12	-3.328	1.339	-1.662	1.812	0.004	0.034	0.008	0.052	0.086	0.002	0.268	0.374	0.460	0.564
13	-3.348	1.339	-1.677	1.815	0.005	0.033	0.008	0.050	0.081	0.002	0.276	0.369	0.461	0.565
14	-3.371	1.340	-1.665	1.795	0.005	0.032	0.008	0.048	0.079	0.002	0.279	0.366	0.461	0.565

Table 5.9.: Estimated coefficients with their uncertainty at each window for SA(T), at T=0.5s using equation 5.2.

Windows	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.372	1.365	-1.331	1.851	-0.011	0.089	0.017	0.132	0.302	0.006	0.142	0.343	0.371	0.687
2	-4.471	1.403	-1.297	1.791	-0.010	0.066	0.011	0.104	0.224	0.005	0.134	0.342	0.367	0.701
3	-4.389	1.387	-1.374	1.934	-0.006	0.061	0.009	0.098	0.185	0.004	0.141	0.338	0.366	0.694
4	-4.377	1.376	-1.345	1.783	-0.006	0.050	0.008	0.079	0.166	0.004	0.145	0.354	0.383	0.668
5	-4.404	1.385	-1.319	1.693	-0.007	0.046	0.008	0.075	0.160	0.003	0.145	0.363	0.391	0.658
6	-4.412	1.397	-1.332	1.809	-0.007	0.044	0.007	0.070	0.144	0.003	0.151	0.356	0.386	0.663
7	-4.442	1.401	-1.297	1.750	-0.008	0.040	0.007	0.064	0.136	0.003	0.145	0.356	0.384	0.666
8	-4.473	1.402	-1.249	1.639	-0.010	0.035	0.006	0.057	0.130	0.003	0.142	0.353	0.381	0.670
9	-4.497	1.418	-1.279	1.638	-0.007	0.032	0.006	0.052	0.117	0.002	0.148	0.355	0.385	0.672
10	-4.497	1.416	-1.359	1.627	-0.004	0.031	0.007	0.050	0.103	0.002	0.201	0.343	0.398	0.653
11	-4.508	1.404	-1.414	1.605	-0.001	0.030	0.007	0.046	0.091	0.002	0.247	0.331	0.413	0.627
12	-4.533	1.406	-1.454	1.591	0.002	0.028	0.007	0.042	0.082	0.002	0.264	0.321	0.416	0.619
13	-4.547	1.404	-1.475	1.588	0.003	0.027	0.007	0.040	0.076	0.002	0.271	0.316	0.416	0.619
14	-4.563	1.405	-1.475	1.567	0.003	0.026	0.007	0.038	0.074	0.002	0.274	0.314	0.416	0.619

Table 5.10.: Estimated coefficients with their uncertainty at each window for SA(T), at T=1.0s using equation 5.2.

Windows	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.791	1.334	-1.598	1.992	-0.004	0.098	0.017	0.144	0.271	0.006	0.142	0.360	0.387	0.668
2	-4.881	1.365	-1.581	1.924	-0.001	0.072	0.011	0.113	0.194	0.005	0.133	0.353	0.377	0.685
3	-4.807	1.347	-1.627	2.019	0.001	0.065	0.009	0.103	0.163	0.004	0.139	0.347	0.374	0.677
4	-4.788	1.334	-1.575	1.916	0.000	0.056	0.009	0.088	0.154	0.004	0.152	0.375	0.404	0.637
5	-4.818	1.346	-1.530	1.887	-0.002	0.054	0.008	0.086	0.153	0.004	0.154	0.386	0.416	0.625
6	-4.815	1.355	-1.567	1.978	-0.001	0.050	0.008	0.079	0.133	0.003	0.157	0.375	0.407	0.636
7	-4.845	1.361	-1.541	1.927	-0.002	0.046	0.007	0.072	0.126	0.003	0.152	0.376	0.406	0.637
8	-4.883	1.359	-1.486	1.805	-0.003	0.040	0.006	0.064	0.118	0.003	0.148	0.370	0.399	0.644
9	-4.892	1.373	-1.532	1.834	0.000	0.037	0.006	0.058	0.106	0.003	0.152	0.370	0.400	0.648
10	-4.885	1.372	-1.611	1.809	0.002	0.035	0.007	0.055	0.094	0.002	0.199	0.358	0.410	0.635
11	-4.872	1.360	-1.695	1.806	0.005	0.033	0.007	0.051	0.082	0.002	0.241	0.345	0.421	0.616
12	-4.874	1.361	-1.763	1.815	0.008	0.031	0.007	0.047	0.073	0.002	0.257	0.336	0.422	0.611
13	-4.875	1.360	-1.800	1.827	0.010	0.030	0.007	0.046	0.068	0.002	0.263	0.330	0.422	0.613
14	-4.884	1.361	-1.810	1.817	0.010	0.029	0.007	0.044	0.066	0.002	0.265	0.327	0.421	0.614

Table 5.11.: Estimated coefficients with their uncertainty at each window for PGV using equation 5.3.

Windows	a	b	c	h	e	σ_a	σ_b	σ_c	σ_h	σ_e	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.695	1.252	-1.826	2.172	0.219	0.052	0.018	0.042	0.154	0.003	0.160	0.306	0.345	0.738
2	-4.753	1.286	-1.806	2.373	0.224	0.035	0.011	0.030	0.106	0.002	0.153	0.296	0.333	0.762
3	-4.688	1.261	-1.815	2.509	0.218	0.030	0.009	0.025	0.088	0.002	0.153	0.293	0.331	0.753
4	-4.821	1.268	-1.685	1.838	0.223	0.025	0.009	0.019	0.079	0.002	0.177	0.312	0.359	0.724
5	-4.846	1.286	-1.662	1.730	0.224	0.023	0.009	0.018	0.075	0.001	0.180	0.322	0.369	0.714
6	-4.775	1.289	-1.732	2.253	0.222	0.022	0.008	0.018	0.069	0.001	0.178	0.313	0.361	0.722
7	-4.769	1.295	-1.744	2.310	0.221	0.021	0.007	0.017	0.065	0.001	0.173	0.314	0.358	0.724
8	-4.779	1.288	-1.725	2.237	0.221	0.019	0.007	0.015	0.059	0.001	0.171	0.308	0.352	0.729
9	-4.818	1.297	-1.704	2.175	0.221	0.017	0.006	0.014	0.056	0.001	0.171	0.309	0.353	0.731
10	-4.809	1.295	-1.746	2.003	0.214	0.016	0.006	0.012	0.048	0.001	0.197	0.304	0.363	0.718
11	-4.824	1.279	-1.748	1.892	0.208	0.015	0.007	0.011	0.043	0.001	0.222	0.297	0.371	0.700
12	-4.883	1.278	-1.696	1.787	0.211	0.015	0.007	0.009	0.040	0.001	0.232	0.288	0.370	0.699
13	-4.967	1.282	-1.662	1.525	0.206	0.013	0.006	0.008	0.038	0.001	0.236	0.286	0.370	0.702
14	-4.976	1.281	-1.660	1.485	0.205	0.013	0.006	0.008	0.036	0.001	0.237	0.283	0.370	0.702

Table 5.12.: Estimated coefficients with their uncertainty at each window for PGA using equation 5.3.

Windows	a	b	c	h	e	σ_a	σ_b	σ_c	σ_h	σ_e	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-3.064	1.189	-1.719	1.270	0.239	0.044	0.020	0.026	0.119	0.003	0.194	0.255	0.321	0.791
2	-3.132	1.240	-1.748	1.326	0.240	0.028	0.013	0.018	0.078	0.002	0.180	0.253	0.310	0.813
3	-3.047	1.220	-1.790	1.434	0.234	0.022	0.010	0.015	0.062	0.001	0.175	0.250	0.305	0.814
4	-3.036	1.209	-1.766	1.367	0.238	0.020	0.009	0.013	0.058	0.001	0.178	0.256	0.312	0.805
5	-3.060	1.215	-1.744	1.326	0.239	0.019	0.009	0.012	0.056	0.001	0.177	0.263	0.317	0.798
6	-3.040	1.222	-1.778	1.420	0.234	0.017	0.008	0.011	0.048	0.001	0.180	0.259	0.315	0.798
7	-3.033	1.226	-1.792	1.468	0.231	0.016	0.007	0.011	0.045	0.001	0.175	0.257	0.311	0.801
8	-3.029	1.224	-1.787	1.457	0.230	0.014	0.006	0.010	0.041	0.001	0.173	0.256	0.309	0.802
9	-3.063	1.241	-1.777	1.451	0.232	0.013	0.006	0.009	0.040	0.001	0.175	0.258	0.312	0.806
10	-3.050	1.234	-1.793	1.404	0.230	0.013	0.006	0.008	0.036	0.001	0.179	0.256	0.313	0.804
11	-3.021	1.217	-1.815	1.385	0.228	0.012	0.006	0.008	0.033	0.001	0.186	0.255	0.316	0.797
12	-2.966	1.214	-1.886	1.530	0.228	0.012	0.005	0.008	0.031	0.001	0.189	0.254	0.316	0.797
13	-2.957	1.209	-1.894	1.523	0.227	0.011	0.005	0.007	0.029	0.001	0.188	0.254	0.316	0.799
14	-2.954	1.206	-1.897	1.507	0.227	0.011	0.005	0.007	0.028	0.001	0.188	0.254	0.316	0.798

Table 5.13.: Estimated coefficients with their uncertainty at each window for SA(T), at T=0.2s using equation 5.3.

Windows	a	b	c	h	e	σ_a	σ_b	σ_c	σ_h	σ_e	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-3.103	1.319	-1.728	2.395	0.250	0.054	0.018	0.045	0.169	0.003	0.160	0.299	0.339	0.757
2	-3.254	1.364	-1.663	2.217	0.253	0.034	0.011	0.029	0.113	0.002	0.150	0.301	0.337	0.772
3	-3.147	1.342	-1.713	2.406	0.250	0.029	0.009	0.025	0.093	0.002	0.152	0.298	0.334	0.767
4	-3.142	1.328	-1.684	2.327	0.257	0.026	0.008	0.022	0.086	0.002	0.154	0.307	0.343	0.754
5	-3.188	1.335	-1.645	2.212	0.260	0.024	0.008	0.020	0.080	0.001	0.154	0.317	0.352	0.743
6	-3.155	1.345	-1.698	2.355	0.257	0.022	0.007	0.019	0.072	0.001	0.159	0.312	0.351	0.745
7	-3.143	1.349	-1.717	2.416	0.257	0.021	0.007	0.018	0.067	0.001	0.153	0.311	0.347	0.749
8	-3.145	1.349	-1.711	2.382	0.257	0.019	0.006	0.016	0.062	0.001	0.153	0.309	0.345	0.751
9	-3.212	1.369	-1.679	2.314	0.257	0.018	0.006	0.015	0.059	0.001	0.159	0.315	0.353	0.749
10	-3.285	1.368	-1.653	2.139	0.241	0.017	0.007	0.013	0.054	0.001	0.210	0.313	0.377	0.713
11	-3.405	1.348	-1.512	1.828	0.226	0.017	0.007	0.011	0.051	0.001	0.252	0.307	0.397	0.674
12	-3.453	1.350	-1.446	2.073	0.227	0.017	0.007	0.011	0.052	0.001	0.268	0.298	0.401	0.664
13	-3.495	1.350	-1.429	2.001	0.222	0.016	0.007	0.010	0.049	0.001	0.275	0.296	0.404	0.660
14	-3.519	1.350	-1.418	1.947	0.220	0.016	0.007	0.009	0.047	0.001	0.278	0.294	0.405	0.658

Table 5.14.: Estimated coefficients with their uncertainty at each window for SA(T), at T=0.5s using equation 5.3.

Windows	a	b	c	h	e	σ_a	σ_b	σ_c	σ_h	σ_e	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.354	1.382	-1.559	1.959	0.227	0.045	0.018	0.034	0.147	0.003	0.166	0.259	0.308	0.784
2	-4.488	1.426	-1.507	1.860	0.231	0.029	0.011	0.022	0.099	0.002	0.155	0.257	0.300	0.806
3	-4.426	1.406	-1.512	1.938	0.226	0.024	0.009	0.018	0.081	0.001	0.159	0.255	0.300	0.798
4	-4.393	1.391	-1.497	1.925	0.232	0.022	0.009	0.017	0.078	0.001	0.165	0.271	0.317	0.778
5	-4.421	1.400	-1.475	1.869	0.235	0.021	0.008	0.016	0.075	0.001	0.164	0.280	0.325	0.768
6	-4.420	1.410	-1.496	1.946	0.233	0.019	0.007	0.014	0.065	0.001	0.170	0.273	0.321	0.772
7	-4.415	1.414	-1.506	1.990	0.232	0.017	0.007	0.013	0.061	0.001	0.164	0.273	0.318	0.775
8	-4.421	1.414	-1.500	1.951	0.233	0.016	0.006	0.012	0.056	0.001	0.162	0.269	0.314	0.779
9	-4.479	1.432	-1.471	1.862	0.229	0.015	0.006	0.011	0.054	0.001	0.166	0.275	0.321	0.777
10	-4.462	1.436	-1.523	1.989	0.205	0.016	0.007	0.011	0.050	0.001	0.210	0.276	0.347	0.741
11	-4.522	1.421	-1.482	1.825	0.190	0.016	0.007	0.010	0.046	0.001	0.248	0.271	0.367	0.702
12	-4.571	1.422	-1.465	1.754	0.180	0.015	0.007	0.009	0.043	0.001	0.260	0.266	0.372	0.687
13	-4.608	1.420	-1.443	1.667	0.173	0.014	0.007	0.008	0.041	0.001	0.266	0.265	0.375	0.683
14	-4.632	1.420	-1.431	1.607	0.170	0.014	0.007	0.008	0.040	0.001	0.268	0.264	0.376	0.681

Table 5.15.: Estimated coefficients with their uncertainty at each window for SA(T), at T=1.0s using equation 5.3.

Windows	a	b	c	h	e	σ_a	σ_b	σ_c	σ_h	σ_e	$\sigma_{inter-event}$	$\sigma_{intra-event}$	σ_{total}	R^2
1	-4.916	1.355	-1.582	1.828	0.209	0.046	0.017	0.036	0.158	0.003	0.156	0.292	0.331	0.755
2	-5.027	1.382	-1.503	1.569	0.204	0.028	0.011	0.022	0.104	0.002	0.147	0.285	0.321	0.773
3	-4.946	1.364	-1.549	1.851	0.204	0.024	0.009	0.019	0.085	0.002	0.150	0.281	0.319	0.767
4	-4.905	1.350	-1.520	1.737	0.206	0.023	0.009	0.018	0.085	0.002	0.163	0.311	0.351	0.729
5	-4.903	1.362	-1.525	1.817	0.210	0.022	0.008	0.018	0.083	0.0010	0.165	0.322	0.363	0.718
6	-4.915	1.369	-1.536	1.807	0.206	0.020	0.008	0.015	0.070	0.001	0.169	0.311	0.354	0.726
7	-4.919	1.374	-1.538	1.836	0.206	0.018	0.007	0.015	0.066	0.001	0.163	0.312	0.352	0.728
8	-4.930	1.372	-1.525	1.757	0.205	0.016	0.006	0.013	0.060	0.001	0.160	0.305	0.345	0.736
9	-4.977	1.387	-1.503	1.684	0.203	0.015	0.006	0.012	0.057	0.001	0.163	0.307	0.347	0.737
10	-5.006	1.385	-1.518	1.557	0.188	0.015	0.007	0.011	0.051	0.001	0.205	0.303	0.366	0.710
11	-5.024	1.371	-1.538	1.643	0.180	0.015	0.007	0.010	0.047	0.001	0.240	0.294	0.380	0.682
12	-5.115	1.376	-1.474	1.353	0.170	0.014	0.007	0.008	0.044	0.001	0.253	0.289	0.384	0.670
13	-5.163	1.374	-1.427	1.242	0.166	0.014	0.007	0.007	0.043	0.001	0.259	0.285	0.385	0.669
14	-5.137	1.379	-1.466	1.393	0.161	0.014	0.007	0.008	0.041	0.001	0.262	0.284	0.386	0.669

6. Conclusion

My thesis deals with detail understanding and analysis of ground-motion prediction equations in geothermal areas with special emphasis on “The Geysers” geothermal area. The interest for this specific area is because it is the largest productive vapor-dominated enhanced geothermal system in the world. No GMPEs were available, specifically for “The Geysers” region before this work. Different techniques are discussed to compute ground motion prediction equations (e.g. linear regression analysis(LRA), non-linear regression analysis(NLRA) and non-linear mixed effect regression analysis(NLMRA)). Although all techniques had same predictions, NLMRA is finally chosen because it has many advantages in comparison with other two (LRA, NLRA) discussed techniques. In fact, it accounts for both fixed effect and random effect dependencies on the the data(Lindstrom and Bates (1990)). Moreover it allows to split the total uncertainty of ground motion into two i.e. inter-event and intra-event uncertainties. Inter-event and intra-event variabilities of ground motion model represents earthquake-to-earthquake variability and record-to-record variability, respectively. Further, the inter-event residuals represents the average source effects and also reflect various factors such as stress drop, variation of slip in time and space which are not captured by including magnitude, focal mechanism and source depth in the GMPEs. Intra-event residual represents path and site effects (crustal heterogeneity, geological structure and near surface layering) which modify the waveform (in terms of amplitude and frequency) and are not captured by distance metric and site classification on the basis of average shear wave velocity (Abrahamson and Youngs (1992)). Thus, the same techniques is applied to predict ground motion in the Geysers, for five different periods, PGV, PGA and SA(T) at T = 0.2s, 0.5s and 1.0s (with S.I. units). The obtained GMPEs are compared with the model proposed by Atkinson and Boore (2011). We found the our GMPEs are more robust when ground motion predictions have to made for induced earthquakes (smaller earthquakes). Two step regression analysis approach is used for computation of GMPEs which is similar to that proposed by Emolo et al. (2011) in the application to natural seismicity recorded in southern Italy. First, a model is retrieved, which does not include station effects and represent the reference model. Second, a first-order correction is introduced for station effects based on the Z-test of the residual distribution at each station with respect to reference model. It is found that if information about local geology is not available, the method of introducing site/station effect correction is effective and shows significant improvement in the model. It is also demonstrated that the two components of standard deviation (i.e., inter-event and intra-event) are related to source effect and local site effects, respectively. The reduction in total standard deviation, after introducing site/station effect correction is actually related to reduction of intra-event component. Inter-event residual residual distribution shows a slightly positive trend with depth, which, as suggested by McGarr (1984) could indicate the possible increase in stress drop with depth. The GMPEs obtained for the Geysers are applicable up to hypocentral distance of 20 km. The model can be used for rock site condition selection site/station correction factor $s = 0$ and $s=+1$ *or* $s=-1$ for which data are available to compute the residual analysis. It is also proposed that for the sites characterized by geology different from the area analyzed in present study, use the final model with $s = 0$ and adding the site-specific effect by using corrective coefficients proposed by Borcherdt (1944), which are also implemented, for example in Shakemap (Wald et al. (1999)).

Work done in Douglas et al. (2013) is also discussed in the thesis. In this work various ground-motion datasets of induced and natural seismicity (from Basel, Geysers, Hengill, Roswinkel, Soultz and Voerendaal) are compiled and processed and the moment magnitudes for all events are computed homogeneously. Ground-motion prediction equations are computed using NLMRA. It is shown that ground motions from induced and natural earthquakes cannot be distinguished. The combined data sets leads to unrealistic high intra-event uncertainty. It is inferred that it is due to the mapping of epistemic uncertainty into aleatory component. However, a reduction in total standard error is observed over mid-to-long period range, but an increase in total standard error is observed in short period range. It is also shown that focal depth plays an important role for shallow earthquakes at shorter distances.

GMPEs are one of the most important ingredients for seismic hazard analysis. An application to time dependent seismic hazard estimation is also discussed in which it is concluded that seismic hazard is not constant with time and space in the Geysers geothermal area. However, the average hazard level remains below the potentially damaging zones. It is strictly affected by industrial operations in the Geysers. The whole region can be divided in to 2 Zones on the basis of different rate of seismicity and b-values. It is demonstrated that site specific hazard analysis can be used for monitoring seismic hazard. If seismic hazard exceeds a particular fixed threshold then the signal for industrial activity can be assumed as red which means that the activity should be stopped, if the predicted hazard is just below the threshold then the signal is yellow which means that the activities must me adjusted and if predicted hazard is reasonably below the threshold then activities can be continued with green signal (Bommer and Alarcon (2006); Convertito et al. (2012)). The proposed technique can be used for monitoring purposes and can be used in the same respect of the traffic light system proposed by Bommer and Alarcon (2006) for controlling human activities such as mining, drilling, injection and extractions of fluid in geothermal areas etc.

Further, a cyclic correlation between seismicity and industrial activities is observed. Thus the concept of time-dependent ground motion prediction equation is introduced. Time-dependent GMPEs are estimated for different periods. It is observed that standard deviations which accounts for spatial uncertainty of ground motion is not constant. In PSHA, the occurrence earthquake phenomenon is regarded as an ergodic process (Anderson and Brune (2000)). The ergodic assumption uses standard deviations to describe the temporal distribution of ground motion at the single site over multiple earthquakes (Anderson and Brune (1998, 999a,b)). The ergodic assumption is appropriate for natural seismicity where exposure times are smaller than the earthquake return period. But it is emphasized that this assumption seems inappropriate in case of induced seismicity, because in case of induced seismicity the exposure times are longer than earthquake return period and this can lead to overestimation of ground motion hence the seismic hazard (Anderson and Brune (2000)). Thus ergodic assumption should be avoided while PSHA with induced seismicity.

Time-dependent parametric analysis shows that a and h are varying in similar way which is however opposite to *b* and *c* parameter variations for all the observed periods (PGA, PGV and SA(T)). But variations in parameter are opposite for high frequencies (PGA) as compared with low frequencies (PGV and SA(T)). For example, *a and h* parameters are increasing with time in case of PGA and decreasing for PGV and SA(T) while *b and c* parameters are decreasing with time for PGA and increasing for SA(T). These time varying changes can be ascribe to the change in source and path effect with the evolution of seismicity. The addition of new events in each time-window changes the magnitude and the hypocentral distance range. Thus contributing the variations in ground motion recordings with time. It is observed that after August 2009 there is a significant change in inter-event and intra-event uncertainties. The inter-event uncertainty increases in this period which can related to the fact that the number of events are increased during this period and inter-event accounts for the source effect on ground motion. Moreover, this observation can be ascribed to aleatory component of uncertainty which accounts for truly random effects on ground motion observations and addition of new earthquakes (random distribution of sources) has increased this effect (Anderson and Brune (2000)). The intra-event uncertainty accounts for the path and the site effect. Thus decrease in intra-event uncertainty is ascribed to epistemic uncertainty component (due to lack of knowledge) which is decreased after addition of new information. (repeatability in the path and site response) (Anderson and Brune (2000)). Thus overall increase in total standard deviation indicates that the effect of inter-event (aleatory) component is larger than intra-event (epistemic) component. Thus overall increase in total standard deviation indicates that the effect of inter-event (aleatory) component is more than intra-event (epistemic) component. The time-dependent analysis of ground-motion prediction equations has shown interesting and promising results. It is found that the ground motion-prediction changes with time because of change in medium properties.

This work represents one of the fewer studies made in EGS till date. This kind of analysis can be extended to other geothermal or mining areas where induced seismicity can be expected. In future, ideas based on this study can be extended or modified to investigate seismicity in new EGS or existing EGS to monitor ground motion levels, which can be dangerous to near by towns.

A. Tables

Table .1: Regression coefficients and related uncertainty for Model 1 (uncorrected for site effect)

Periods	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	σ inter-event	σ intra-event	σ_{total}
0.010	-5.998	2.149	-1.768	2.501	-0.024	0.426	0.069	0.207	0.596	0.011	0.825	0.792	1.143
0.011	-6.015	2.152	-1.760	2.486	-0.025	0.423	0.068	0.206	0.597	0.011	0.820	0.792	1.140
0.011	-6.023	2.154	-1.755	2.476	-0.025	0.422	0.068	0.206	0.598	0.011	0.817	0.792	1.138
0.012	-6.024	2.154	-1.753	2.475	-0.026	0.422	0.068	0.206	0.598	0.011	0.817	0.792	1.138
0.012	-6.020	2.155	-1.754	2.478	-0.026	0.422	0.068	0.206	0.599	0.011	0.817	0.793	1.139
0.013	-6.019	2.155	-1.754	2.475	-0.026	0.422	0.068	0.206	0.599	0.011	0.818	0.794	1.140
0.013	-6.023	2.156	-1.750	2.464	-0.027	0.421	0.068	0.206	0.600	0.011	0.819	0.795	1.141
0.014	-6.022	2.156	-1.750	2.449	-0.027	0.420	0.067	0.205	0.600	0.011	0.808	0.797	1.135
0.014	-6.020	2.156	-1.749	2.441	-0.027	0.419	0.068	0.205	0.601	0.011	0.808	0.798	1.136
0.015	-6.014	2.155	-1.749	2.431	-0.027	0.418	0.068	0.204	0.601	0.011	0.808	0.800	1.137
0.015	-6.002	2.153	-1.752	2.424	-0.027	0.418	0.068	0.204	0.600	0.011	0.809	0.801	1.139
0.016	-5.985	2.151	-1.756	2.419	-0.027	0.419	0.068	0.205	0.600	0.011	0.810	0.804	1.141
0.016	-5.963	2.148	-1.761	2.414	-0.027	0.419	0.068	0.205	0.600	0.011	0.811	0.807	1.144
0.017	-5.937	2.144	-1.768	2.408	-0.027	0.420	0.068	0.206	0.600	0.011	0.812	0.810	1.147
0.018	-5.903	2.140	-1.778	2.407	-0.026	0.422	0.068	0.206	0.599	0.011	0.812	0.814	1.150
0.018	-5.845	2.136	-1.799	2.439	-0.025	0.428	0.068	0.209	0.597	0.011	0.813	0.818	1.154
0.019	-5.769	2.131	-1.828	2.496	-0.024	0.438	0.068	0.214	0.595	0.011	0.815	0.823	1.158
0.020	-5.685	2.126	-1.861	2.558	-0.023	0.449	0.069	0.219	0.593	0.011	0.818	0.827	1.164
0.021	-5.588	2.119	-1.898	2.629	-0.021	0.461	0.069	0.225	0.591	0.011	0.822	0.832	1.169
0.021	-5.483	2.111	-1.938	2.698	-0.019	0.474	0.069	0.230	0.588	0.012	0.825	0.836	1.175
0.022	-5.402	2.104	-1.966	2.728	-0.018	0.480	0.069	0.234	0.585	0.012	0.825	0.841	1.178
0.023	-5.328	2.098	-1.990	2.754	-0.016	0.486	0.070	0.236	0.583	0.012	0.827	0.845	1.182
0.024	-5.251	2.092	-2.016	2.785	-0.015	0.494	0.070	0.240	0.581	0.012	0.828	0.849	1.186
0.025	-5.164	2.086	-2.048	2.820	-0.014	0.502	0.070	0.244	0.579	0.012	0.825	0.855	1.188
0.026	-5.074	2.079	-2.080	2.854	-0.012	0.510	0.070	0.247	0.576	0.012	0.825	0.860	1.192
0.027	-4.977	2.070	-2.111	2.903	-0.011	0.520	0.070	0.252	0.575	0.012	0.826	0.864	1.195
0.028	-4.905	2.062	-2.129	2.931	-0.010	0.526	0.070	0.255	0.575	0.013	0.831	0.867	1.201
0.029	-4.817	2.055	-2.154	2.975	-0.008	0.535	0.070	0.259	0.574	0.013	0.834	0.870	1.205
0.030	-4.746	2.047	-2.170	2.988	-0.008	0.540	0.071	0.261	0.574	0.013	0.837	0.873	1.209
0.031	-4.688	2.039	-2.181	2.975	-0.007	0.539	0.071	0.261	0.572	0.013	0.839	0.877	1.213
0.033	-4.644	2.031	-2.184	2.957	-0.006	0.538	0.071	0.261	0.571	0.013	0.841	0.880	1.217
0.034	-4.629	2.026	-2.171	2.935	-0.007	0.535	0.071	0.259	0.573	0.013	0.847	0.880	1.221
0.035	-4.629	2.023	-2.154	2.909	-0.007	0.531	0.072	0.257	0.575	0.013	0.849	0.880	1.223
0.036	-4.647	2.020	-2.124	2.880	-0.009	0.526	0.072	0.255	0.580	0.013	0.849	0.880	1.223
0.038	-4.632	2.019	-2.111	2.888	-0.009	0.528	0.072	0.256	0.585	0.013	0.851	0.880	1.224
0.039	-4.630	2.018	-2.092	2.881	-0.010	0.527	0.072	0.255	0.589	0.013	0.857	0.879	1.228
0.041	-4.669	2.015	-2.051	2.818	-0.012	0.516	0.072	0.250	0.594	0.012	0.861	0.877	1.229

0.042	-4.720	2.014	-2.006	2.739	-0.015	0.501	0.072	0.244	0.597	0.012	0.863	0.873	1.228
0.044	-4.740	2.013	-1.983	2.681	-0.015	0.491	0.072	0.239	0.596	0.012	0.864	0.870	1.226
0.046	-4.677	2.016	-2.004	2.760	-0.014	0.501	0.073	0.243	0.594	0.012	0.868	0.864	1.225
0.048	-4.620	2.023	-2.026	2.845	-0.013	0.512	0.073	0.248	0.593	0.012	0.868	0.860	1.222
0.049	-4.490	2.033	-2.087	3.009	-0.010	0.537	0.072	0.259	0.591	0.013	0.863	0.857	1.216
0.051	-4.421	2.047	-2.122	3.104	-0.008	0.552	0.071	0.265	0.590	0.013	0.852	0.857	1.208
0.053	-4.353	2.063	-2.158	3.210	-0.007	0.569	0.070	0.273	0.590	0.013	0.838	0.856	1.198
0.055	-4.270	2.078	-2.202	3.327	-0.005	0.587	0.070	0.281	0.589	0.013	0.824	0.854	1.187
0.057	-4.259	2.094	-2.215	3.375	-0.005	0.594	0.068	0.284	0.589	0.013	0.808	0.853	1.174
0.060	-4.301	2.102	-2.190	3.400	-0.006	0.595	0.068	0.284	0.596	0.013	0.793	0.849	1.162
0.062	-4.508	2.122	-2.093	3.275	-0.011	0.568	0.066	0.272	0.604	0.013	0.779	0.844	1.148
0.064	-4.652	2.138	-2.022	3.222	-0.016	0.554	0.065	0.266	0.614	0.013	0.762	0.838	1.132
0.067	-4.787	2.157	-1.958	3.185	-0.020	0.543	0.064	0.261	0.625	0.013	0.746	0.832	1.118
0.070	-4.940	2.173	-1.883	3.125	-0.024	0.527	0.063	0.255	0.636	0.012	0.733	0.825	1.104
0.072	-5.131	2.192	-1.801	2.994	-0.028	0.500	0.062	0.242	0.642	0.012	0.723	0.815	1.089
0.075	-5.324	2.208	-1.722	2.811	-0.031	0.467	0.061	0.228	0.644	0.011	0.710	0.808	1.075
0.078	-5.533	2.225	-1.635	2.554	-0.035	0.426	0.060	0.210	0.647	0.011	0.697	0.803	1.063
0.081	-5.686	2.242	-1.574	2.364	-0.037	0.399	0.060	0.198	0.652	0.010	0.687	0.800	1.055
0.084	-5.832	2.255	-1.520	2.156	-0.039	0.373	0.059	0.186	0.660	0.010	0.676	0.802	1.048
0.087	-6.005	2.276	-1.466	1.913	-0.040	0.346	0.058	0.174	0.674	0.010	0.666	0.803	1.043
0.091	-6.172	2.297	-1.417	1.712	-0.041	0.325	0.057	0.164	0.697	0.009	0.653	0.805	1.036
0.094	-6.289	2.323	-1.398	1.660	-0.041	0.319	0.057	0.161	0.708	0.009	0.643	0.806	1.031
0.098	-6.374	2.350	-1.400	1.683	-0.040	0.322	0.056	0.163	0.708	0.009	0.633	0.809	1.027
0.102	-6.444	2.376	-1.410	1.752	-0.039	0.329	0.056	0.167	0.704	0.009	0.623	0.815	1.026
0.106	-6.571	2.406	-1.399	1.651	-0.039	0.320	0.055	0.163	0.715	0.009	0.617	0.821	1.027
0.110	-6.697	2.428	-1.379	1.543	-0.039	0.311	0.055	0.159	0.735	0.009	0.609	0.826	1.027
0.114	-6.817	2.450	-1.365	1.512	-0.039	0.310	0.055	0.158	0.748	0.009	0.611	0.830	1.031
0.118	-6.943	2.473	-1.351	1.472	-0.038	0.307	0.055	0.157	0.762	0.009	0.614	0.833	1.035
0.123	-7.097	2.499	-1.325	1.451	-0.039	0.306	0.055	0.156	0.781	0.009	0.617	0.834	1.037
0.128	-7.266	2.527	-1.291	1.377	-0.040	0.299	0.055	0.153	0.810	0.009	0.618	0.832	1.036
0.133	-7.425	2.553	-1.263	1.323	-0.040	0.294	0.055	0.150	0.834	0.009	0.616	0.829	1.033
0.138	-7.562	2.577	-1.248	1.303	-0.040	0.292	0.055	0.150	0.847	0.009	0.612	0.829	1.031
0.143	-7.691	2.600	-1.238	1.346	-0.040	0.295	0.055	0.151	0.847	0.009	0.612	0.829	1.031
0.149	-7.826	2.623	-1.228	1.374	-0.040	0.297	0.055	0.152	0.848	0.009	0.612	0.828	1.030
0.154	-8.006	2.647	-1.198	1.323	-0.039	0.293	0.055	0.150	0.877	0.009	0.611	0.828	1.029
0.160	-8.281	2.679	-1.124	1.124	-0.041	0.281	0.056	0.143	0.994	0.009	0.615	0.824	1.028
0.167	-8.623	2.706	-0.994	0.000	-0.048	0.240	0.056	0.122	0.000	0.008	0.605	0.819	1.019
0.173	-8.873	2.734	-0.917	0.000	-0.051	0.241	0.056	0.122	0.000	0.008	0.604	0.816	1.015
0.180	-9.099	2.760	-0.854	0.000	-0.053	0.241	0.056	0.121	0.000	0.008	0.603	0.809	1.009
0.187	-9.290	2.780	-0.798	0.000	-0.057	0.240	0.057	0.121	0.000	0.008	0.605	0.803	1.005
0.194	-9.382	2.787	-0.787	0.000	-0.057	0.241	0.057	0.121	0.000	0.008	0.605	0.801	1.004

0.201	-9.513	2.805	-0.776	0.000	-0.058	0.243	0.058	0.121	0.000	0.008	0.612	0.800	1.007
0.209	-9.636	2.830	-0.780	0.000	-0.056	0.244	0.059	0.121	0.000	0.008	0.620	0.797	1.009
0.217	-9.749	2.842	-0.774	0.000	-0.055	0.248	0.059	0.123	0.000	0.008	0.621	0.793	1.007
0.226	-9.735	2.849	-0.834	0.000	-0.050	0.248	0.060	0.122	0.000	0.008	0.633	0.785	1.008
0.234	-9.773	2.855	-0.867	0.000	-0.045	0.262	0.070	0.123	0.000	0.008	0.766	0.766	1.083
0.244	-9.860	2.865	-0.874	0.000	-0.043	0.262	0.070	0.123	0.000	0.008	0.762	0.760	1.076
0.253	-10.028	2.890	-0.857	0.000	-0.045	0.250	0.061	0.122	0.000	0.008	0.639	0.770	1.001
0.263	-10.156	2.904	-0.849	0.000	-0.045	0.249	0.061	0.121	0.000	0.008	0.640	0.764	0.997
0.273	-10.161	2.910	-0.900	0.815	-0.040	0.280	0.069	0.132	1.485	0.008	0.747	0.745	1.055
0.283	-10.268	2.921	-0.901	0.873	-0.040	0.282	0.069	0.134	1.415	0.008	0.743	0.740	1.048
0.294	-10.393	2.929	-0.879	0.701	-0.041	0.267	0.064	0.129	1.664	0.008	0.662	0.736	0.990
0.306	-10.504	2.935	-0.859	0.836	-0.042	0.277	0.068	0.131	1.501	0.008	0.728	0.722	1.025
0.318	-10.615	2.936	-0.831	0.644	-0.045	0.265	0.066	0.126	1.848	0.008	0.689	0.717	0.994
0.330	-10.741	2.949	-0.817	0.802	-0.045	0.274	0.068	0.129	1.604	0.008	0.714	0.709	1.006
0.343	-10.836	2.954	-0.811	0.878	-0.044	0.276	0.068	0.131	1.522	0.008	0.712	0.705	1.002
0.356	-10.896	2.960	-0.825	1.147	-0.042	0.292	0.069	0.138	1.302	0.008	0.731	0.691	1.006
0.370	-10.929	2.968	-0.857	1.505	-0.039	0.321	0.071	0.152	1.160	0.009	0.751	0.685	1.016
0.384	-10.994	2.965	-0.852	1.643	-0.040	0.334	0.072	0.158	1.153	0.009	0.756	0.682	1.018
0.399	-10.948	2.963	-0.895	2.185	-0.037	0.386	0.075	0.181	1.100	0.010	0.791	0.657	1.028
0.415	-11.080	2.964	-0.873	2.072	-0.037	0.375	0.076	0.176	1.119	0.009	0.790	0.656	1.027
0.431	-11.134	2.960	-0.884	2.154	-0.034	0.384	0.076	0.180	1.110	0.010	0.791	0.654	1.026
0.447	-11.440	2.959	-0.772	1.329	-0.038	0.301	0.067	0.143	1.295	0.008	0.661	0.662	0.936
0.465	-11.270	2.933	-0.853	2.234	-0.033	0.407	0.078	0.191	1.210	0.010	0.795	0.639	1.019
0.483	-11.301	2.931	-0.882	2.384	-0.029	0.426	0.078	0.200	1.195	0.010	0.796	0.638	1.021
0.501	-11.232	2.912	-0.953	2.723	-0.022	0.480	0.081	0.223	1.181	0.011	0.816	0.638	1.036

Table .2: Regression coefficients and related uncertainty for Model 1 (corrected for site effect)

Periods	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	σ inter-event	σ intra-event	σ_{total}
0.010	-6.514	1.995	-1.468	2.490	-0.030	0.423	0.085	0.200	0.688	0.010	1.079	0.730	1.303
0.010	-6.523	2.000	-1.467	2.487	-0.030	0.423	0.085	0.200	0.689	0.010	1.074	0.731	1.299
0.011	-6.543	2.006	-1.463	2.475	-0.030	0.421	0.084	0.199	0.689	0.010	1.066	0.731	1.292
0.011	-6.556	2.012	-1.462	2.467	-0.031	0.419	0.084	0.198	0.689	0.010	1.057	0.731	1.285
0.012	-6.555	2.015	-1.465	2.465	-0.031	0.419	0.083	0.198	0.687	0.010	1.050	0.732	1.280
0.012	-6.550	2.017	-1.471	2.465	-0.031	0.419	0.083	0.199	0.686	0.010	1.044	0.734	1.276
0.013	-6.539	2.018	-1.478	2.466	-0.030	0.420	0.083	0.199	0.684	0.010	1.039	0.736	1.273
0.013	-6.522	2.019	-1.486	2.473	-0.030	0.422	0.083	0.201	0.683	0.010	1.037	0.739	1.273
0.014	-6.495	2.018	-1.499	2.488	-0.030	0.425	0.083	0.202	0.681	0.011	1.036	0.742	1.274
0.014	-6.465	2.017	-1.513	2.503	-0.029	0.429	0.083	0.204	0.679	0.011	1.038	0.745	1.278
0.015	-6.429	2.015	-1.529	2.519	-0.029	0.433	0.083	0.206	0.677	0.011	1.038	0.749	1.280
0.015	-6.393	2.012	-1.542	2.530	-0.028	0.436	0.083	0.208	0.676	0.011	1.040	0.753	1.284
0.016	-6.348	2.008	-1.557	2.544	-0.027	0.440	0.083	0.210	0.675	0.011	1.044	0.758	1.290
0.016	-6.295	2.002	-1.574	2.566	-0.027	0.446	0.084	0.213	0.674	0.011	1.048	0.763	1.297
0.017	-6.240	1.996	-1.590	2.588	-0.026	0.452	0.084	0.216	0.675	0.011	1.054	0.769	1.305
0.018	-6.180	1.988	-1.606	2.608	-0.026	0.459	0.085	0.219	0.676	0.011	1.060	0.776	1.314
0.018	-6.100	1.980	-1.630	2.666	-0.025	0.471	0.085	0.224	0.678	0.011	1.070	0.783	1.326
0.019	-5.990	1.971	-1.664	2.774	-0.024	0.492	0.086	0.234	0.682	0.012	1.080	0.791	1.339
0.020	-5.867	1.961	-1.705	2.887	-0.022	0.514	0.087	0.244	0.686	0.012	1.092	0.799	1.353
0.021	-5.737	1.950	-1.746	3.009	-0.021	0.538	0.088	0.255	0.690	0.013	1.106	0.805	1.368
0.021	-5.596	1.937	-1.789	3.124	-0.019	0.562	0.089	0.266	0.693	0.013	1.116	0.813	1.381
0.022	-5.484	1.927	-1.823	3.194	-0.018	0.579	0.090	0.273	0.694	0.013	1.125	0.819	1.391
0.023	-5.415	1.922	-1.841	3.226	-0.017	0.588	0.090	0.278	0.696	0.013	1.129	0.824	1.398
0.024	-5.340	1.917	-1.864	3.266	-0.017	0.598	0.090	0.282	0.697	0.014	1.123	0.830	1.396
0.025	-5.291	1.913	-1.872	3.276	-0.017	0.603	0.090	0.285	0.700	0.014	1.125	0.835	1.401
0.026	-5.228	1.907	-1.883	3.303	-0.017	0.612	0.090	0.289	0.703	0.014	1.128	0.840	1.407
0.027	-5.190	1.904	-1.884	3.317	-0.017	0.617	0.091	0.291	0.708	0.014	1.134	0.844	1.414
0.028	-5.177	1.901	-1.870	3.313	-0.018	0.619	0.091	0.292	0.716	0.014	1.139	0.848	1.420
0.029	-5.192	1.901	-1.846	3.270	-0.019	0.614	0.092	0.290	0.723	0.014	1.145	0.852	1.428
0.030	-5.183	1.895	-1.826	3.243	-0.020	0.612	0.092	0.289	0.730	0.014	1.153	0.855	1.436
0.031	-5.188	1.894	-1.805	3.209	-0.021	0.607	0.093	0.287	0.736	0.014	1.159	0.858	1.442
0.033	-5.202	1.895	-1.780	3.172	-0.022	0.602	0.093	0.285	0.743	0.014	1.163	0.860	1.446
0.034	-5.245	1.898	-1.744	3.129	-0.024	0.595	0.093	0.282	0.753	0.014	1.166	0.861	1.449
0.035	-5.304	1.904	-1.699	3.093	-0.027	0.589	0.093	0.279	0.768	0.014	1.163	0.861	1.447
0.036	-5.386	1.911	-1.647	3.052	-0.029	0.581	0.093	0.275	0.786	0.014	1.160	0.860	1.444
0.038	-5.435	1.918	-1.609	3.060	-0.032	0.581	0.093	0.276	0.804	0.014	1.157	0.859	1.441
0.039	-5.458	1.926	-1.586	3.097	-0.033	0.585	0.093	0.277	0.817	0.014	1.161	0.854	1.442

0.041	-5.528	1.933	-1.542	3.064	-0.036	0.576	0.093	0.273	0.831	0.013	1.160	0.849	1.437
0.042	-5.610	1.941	-1.495	3.000	-0.038	0.561	0.092	0.266	0.841	0.013	1.149	0.843	1.425
0.044	-5.669	1.953	-1.465	2.994	-0.039	0.557	0.091	0.264	0.852	0.013	1.140	0.838	1.415
0.046	-5.685	1.962	-1.455	3.061	-0.039	0.562	0.090	0.266	0.858	0.013	1.132	0.828	1.403
0.048	-5.735	1.973	-1.433	3.085	-0.040	0.559	0.090	0.265	0.865	0.013	1.122	0.818	1.388
0.049	-5.782	1.986	-1.416	3.117	-0.040	0.558	0.088	0.264	0.870	0.013	1.109	0.808	1.372
0.051	-5.835	2.003	-1.401	3.131	-0.040	0.554	0.087	0.262	0.872	0.013	1.094	0.800	1.355
0.053	-5.892	2.022	-1.384	3.154	-0.040	0.551	0.086	0.261	0.875	0.013	1.074	0.790	1.333
0.055	-5.964	2.038	-1.362	3.144	-0.041	0.540	0.085	0.256	0.874	0.012	1.058	0.778	1.313
0.057	-6.074	2.054	-1.325	3.044	-0.041	0.515	0.083	0.244	0.867	0.012	1.040	0.764	1.290
0.060	-6.251	2.071	-1.257	2.826	-0.043	0.474	0.082	0.225	0.864	0.011	1.023	0.750	1.268
0.062	-6.452	2.090	-1.175	2.550	-0.047	0.428	0.080	0.204	0.870	0.011	1.005	0.737	1.247
0.064	-6.606	2.109	-1.115	2.345	-0.049	0.396	0.079	0.189	0.876	0.010	0.985	0.724	1.222
0.067	-6.711	2.126	-1.080	2.240	-0.051	0.376	0.077	0.180	0.877	0.010	0.961	0.711	1.195
0.070	-6.834	2.145	-1.037	2.139	-0.052	0.358	0.075	0.171	0.887	0.009	0.943	0.697	1.172
0.072	-6.965	2.166	-0.997	2.044	-0.054	0.341	0.074	0.162	0.895	0.009	0.925	0.681	1.148
0.075	-7.099	2.187	-0.958	1.883	-0.055	0.319	0.072	0.152	0.905	0.008	0.906	0.667	1.125
0.078	-7.219	2.207	-0.925	1.711	-0.056	0.298	0.071	0.141	0.916	0.008	0.890	0.652	1.103
0.081	-7.324	2.230	-0.904	1.644	-0.057	0.287	0.070	0.136	0.921	0.008	0.875	0.639	1.083
0.084	-7.433	2.249	-0.883	1.482	-0.057	0.271	0.069	0.127	0.942	0.007	0.858	0.629	1.064
0.087	-7.562	2.271	-0.858	1.234	-0.057	0.251	0.067	0.117	1.000	0.007	0.843	0.619	1.046
0.091	-7.689	2.294	-0.835	1.044	-0.057	0.237	0.066	0.110	1.083	0.007	0.827	0.609	1.027
0.094	-7.813	2.319	-0.813	0.925	-0.058	0.228	0.065	0.105	1.166	0.006	0.813	0.600	1.011
0.098	-7.911	2.348	-0.813	0.943	-0.057	0.226	0.064	0.104	1.139	0.006	0.799	0.593	0.995
0.102	-7.991	2.376	-0.827	1.058	-0.056	0.229	0.063	0.106	1.049	0.006	0.787	0.589	0.983
0.106	-8.075	2.404	-0.845	1.044	-0.054	0.226	0.062	0.106	1.030	0.006	0.775	0.588	0.973
0.110	-8.190	2.426	-0.841	0.809	-0.053	0.215	0.061	0.099	1.186	0.006	0.763	0.584	0.961
0.114	-8.287	2.449	-0.849	0.780	-0.052	0.212	0.061	0.098	1.197	0.006	0.757	0.581	0.954
0.118	-8.375	2.471	-0.862	0.821	-0.049	0.213	0.060	0.099	1.136	0.006	0.752	0.579	0.949
0.123	-8.477	2.494	-0.869	0.868	-0.048	0.213	0.060	0.099	1.081	0.006	0.747	0.575	0.943
0.128	-8.576	2.518	-0.875	0.948	-0.047	0.215	0.060	0.100	1.015	0.006	0.741	0.572	0.936
0.133	-8.658	2.537	-0.887	1.039	-0.045	0.217	0.059	0.102	0.949	0.006	0.731	0.569	0.927
0.138	-8.724	2.554	-0.911	1.159	-0.043	0.221	0.058	0.104	0.876	0.006	0.725	0.565	0.919
0.143	-8.795	2.572	-0.933	1.321	-0.041	0.229	0.058	0.109	0.814	0.006	0.721	0.563	0.915
0.149	-8.853	2.590	-0.964	1.478	-0.039	0.237	0.058	0.113	0.765	0.007	0.716	0.561	0.909
0.154	-8.962	2.610	-0.977	1.494	-0.036	0.237	0.057	0.113	0.752	0.007	0.711	0.560	0.905
0.160	-9.165	2.648	-0.954	1.376	-0.036	0.232	0.058	0.110	0.785	0.006	0.699	0.559	0.895
0.167	-9.372	2.683	-0.916	1.305	-0.037	0.228	0.058	0.108	0.833	0.006	0.690	0.555	0.886
0.173	-9.555	2.716	-0.893	1.183	-0.037	0.223	0.058	0.104	0.881	0.006	0.684	0.553	0.880
0.180	-9.714	2.747	-0.877	1.066	-0.038	0.217	0.057	0.102	0.944	0.006	0.672	0.552	0.870
0.187	-9.871	2.765	-0.843	1.003	-0.040	0.214	0.057	0.100	1.011	0.006	0.668	0.551	0.866

0.194	-9.946	2.766	-0.834	0.930	-0.042	0.210	0.056	0.099	1.063	0.006	0.655	0.552	0.856
0.201	-10.024	2.784	-0.850	1.080	-0.041	0.219	0.057	0.103	0.979	0.006	0.658	0.554	0.860
0.209	-10.193	2.821	-0.842	1.089	-0.042	0.221	0.057	0.104	0.987	0.006	0.660	0.555	0.862
0.217	-10.297	2.843	-0.855	1.117	-0.041	0.226	0.058	0.106	0.986	0.006	0.659	0.557	0.863
0.226	-10.336	2.852	-0.889	1.331	-0.037	0.238	0.058	0.112	0.890	0.007	0.664	0.553	0.864
0.234	-10.481	2.880	-0.892	1.231	-0.036	0.233	0.058	0.110	0.912	0.007	0.653	0.555	0.857
0.244	-10.594	2.896	-0.891	1.156	-0.035	0.231	0.058	0.109	0.943	0.007	0.651	0.557	0.857
0.253	-10.770	2.923	-0.862	0.877	-0.039	0.214	0.052	0.103	1.128	0.006	0.567	0.569	0.803
0.263	-10.921	2.943	-0.843	0.878	-0.040	0.215	0.053	0.103	1.157	0.006	0.570	0.570	0.806
0.273	-10.959	2.959	-0.885	1.369	-0.038	0.247	0.058	0.117	0.908	0.007	0.643	0.558	0.852
0.283	-11.092	2.975	-0.873	1.374	-0.038	0.249	0.059	0.118	0.924	0.007	0.647	0.559	0.855
0.294	-11.220	2.977	-0.845	1.259	-0.041	0.245	0.059	0.115	0.990	0.007	0.645	0.562	0.855
0.306	-11.379	2.989	-0.804	1.107	-0.044	0.238	0.060	0.112	1.099	0.007	0.648	0.566	0.860
0.318	-11.534	2.989	-0.746	0.903	-0.048	0.230	0.060	0.107	1.315	0.007	0.653	0.563	0.862
0.330	-11.665	3.007	-0.736	1.004	-0.049	0.236	0.061	0.110	1.263	0.007	0.657	0.563	0.865
0.343	-11.743	3.012	-0.733	1.072	-0.050	0.239	0.061	0.111	1.230	0.007	0.650	0.563	0.860
0.356	-11.830	3.017	-0.737	1.088	-0.049	0.241	0.061	0.112	1.219	0.007	0.656	0.558	0.861
0.370	-11.953	3.039	-0.741	1.159	-0.049	0.247	0.062	0.114	1.187	0.007	0.663	0.558	0.867
0.384	-12.066	3.047	-0.726	1.194	-0.050	0.251	0.063	0.116	1.204	0.007	0.667	0.559	0.870
0.399	-12.167	3.060	-0.723	1.378	-0.050	0.266	0.064	0.124	1.168	0.007	0.676	0.558	0.876
0.415	-12.295	3.068	-0.706	1.312	-0.051	0.265	0.065	0.123	1.222	0.007	0.679	0.562	0.882
0.431	-12.381	3.063	-0.697	1.280	-0.051	0.264	0.066	0.122	1.252	0.007	0.683	0.563	0.886
0.447	-12.525	3.069	-0.671	1.207	-0.052	0.263	0.067	0.120	1.331	0.007	0.685	0.562	0.886
0.465	-12.608	3.066	-0.666	1.010	-0.052	0.263	0.067	0.121	1.544	0.007	0.685	0.569	0.890
0.483	-12.698	3.070	-0.666	1.073	-0.052	0.268	0.068	0.123	1.504	0.007	0.691	0.568	0.903
0.501	-12.736	3.056	-0.675	1.218	-0.050	0.282	0.070	0.129	1.434	0.008	0.699	0.573	0.903

Table .3: Regression coefficients and related uncertainty for Model 2 (corrected for site effect)

Periods	a	b	c	d	σ_a	σ_b	σ_c	σ_d	σ inter-event	σ intra-event	σ_{total}
0.010	-7.198	1.991	-1.105	-0.048	0.200	0.085	0.062	0.005	1.080	0.731	1.304
0.010	-7.205	1.996	-1.105	-0.048	0.199	0.085	0.063	0.005	1.075	0.732	1.300
0.011	-7.217	2.002	-1.105	-0.048	0.198	0.084	0.063	0.005	1.064	0.732	1.292
0.011	-7.226	2.008	-1.106	-0.048	0.197	0.084	0.063	0.005	1.056	0.732	1.285
0.012	-7.227	2.011	-1.109	-0.048	0.196	0.083	0.063	0.005	1.049	0.733	1.279
0.012	-7.224	2.013	-1.113	-0.048	0.195	0.083	0.063	0.005	1.043	0.735	1.275
0.013	-7.216	2.014	-1.118	-0.048	0.195	0.083	0.063	0.005	1.038	0.737	1.273
0.013	-7.207	2.015	-1.123	-0.048	0.195	0.082	0.063	0.005	1.036	0.740	1.273
0.014	-7.192	2.014	-1.130	-0.048	0.195	0.082	0.063	0.005	1.035	0.743	1.274
0.014	-7.175	2.013	-1.137	-0.048	0.196	0.083	0.064	0.005	1.036	0.746	1.277
0.015	-7.153	2.011	-1.145	-0.048	0.196	0.083	0.064	0.005	1.036	0.750	1.279
0.015	-7.128	2.008	-1.152	-0.047	0.197	0.083	0.064	0.006	1.039	0.754	1.284
0.016	-7.097	2.004	-1.161	-0.047	0.197	0.083	0.065	0.006	1.043	0.759	1.290
0.016	-7.062	1.998	-1.168	-0.047	0.198	0.084	0.065	0.006	1.047	0.764	1.296
0.017	-7.025	1.991	-1.175	-0.047	0.200	0.084	0.066	0.006	1.052	0.770	1.304
0.018	-6.982	1.984	-1.183	-0.047	0.201	0.085	0.066	0.006	1.059	0.777	1.313
0.018	-6.941	1.975	-1.186	-0.047	0.203	0.085	0.067	0.006	1.069	0.784	1.326
0.019	-6.902	1.966	-1.185	-0.047	0.205	0.086	0.067	0.006	1.082	0.792	1.341
0.020	-6.856	1.956	-1.187	-0.047	0.207	0.087	0.068	0.006	1.091	0.800	1.353
0.021	-6.811	1.945	-1.185	-0.048	0.209	0.088	0.069	0.006	1.105	0.807	1.369
0.021	-6.758	1.932	-1.185	-0.048	0.212	0.089	0.069	0.006	1.118	0.814	1.383
0.022	-6.706	1.922	-1.190	-0.048	0.213	0.090	0.070	0.006	1.127	0.820	1.394
0.023	-6.664	1.916	-1.194	-0.049	0.214	0.090	0.070	0.006	1.129	0.826	1.399
0.024	-6.629	1.912	-1.197	-0.049	0.214	0.090	0.071	0.006	1.125	0.831	1.399
0.025	-6.589	1.908	-1.201	-0.049	0.214	0.090	0.071	0.006	1.124	0.837	1.402
0.026	-6.550	1.902	-1.201	-0.049	0.216	0.090	0.072	0.006	1.130	0.842	1.409
0.027	-6.520	1.899	-1.197	-0.050	0.217	0.091	0.072	0.006	1.136	0.846	1.416
0.028	-6.493	1.895	-1.190	-0.050	0.217	0.091	0.072	0.006	1.139	0.850	1.421
0.029	-6.468	1.896	-1.187	-0.051	0.218	0.092	0.073	0.006	1.145	0.854	1.428
0.030	-6.432	1.890	-1.180	-0.051	0.220	0.092	0.073	0.006	1.152	0.857	1.436
0.031	-6.404	1.889	-1.175	-0.051	0.221	0.093	0.073	0.006	1.158	0.859	1.442
0.033	-6.382	1.890	-1.169	-0.052	0.221	0.093	0.073	0.006	1.161	0.862	1.446
0.034	-6.379	1.893	-1.155	-0.053	0.222	0.093	0.073	0.006	1.165	0.862	1.449
0.035	-6.391	1.899	-1.134	-0.054	0.221	0.093	0.073	0.006	1.162	0.862	1.447
0.036	-6.420	1.906	-1.109	-0.055	0.221	0.093	0.073	0.006	1.158	0.862	1.444
0.038	-6.448	1.914	-1.082	-0.057	0.220	0.093	0.073	0.006	1.156	0.860	1.441
0.039	-6.471	1.922	-1.060	-0.059	0.220	0.093	0.073	0.006	1.157	0.856	1.439

0.041	-6.498	1.929	-1.037	-0.060	0.220	0.092	0.072	0.006	1.156	0.850	1.435
0.042	-6.523	1.937	-1.019	-0.061	0.218	0.092	0.072	0.006	1.147	0.844	1.424
0.044	-6.560	1.949	-1.001	-0.062	0.216	0.091	0.071	0.006	1.136	0.839	1.412
0.046	-6.598	1.957	-0.980	-0.062	0.214	0.090	0.071	0.006	1.128	0.829	1.400
0.048	-6.645	1.968	-0.960	-0.063	0.212	0.089	0.070	0.006	1.117	0.819	1.385
0.049	-6.694	1.982	-0.942	-0.063	0.210	0.088	0.069	0.006	1.104	0.810	1.369
0.051	-6.743	1.999	-0.929	-0.063	0.207	0.087	0.068	0.006	1.089	0.801	1.352
0.053	-6.799	2.018	-0.914	-0.063	0.204	0.086	0.067	0.006	1.070	0.791	1.331
0.055	-6.852	2.034	-0.901	-0.063	0.200	0.084	0.066	0.006	1.054	0.779	1.310
0.057	-6.900	2.050	-0.895	-0.062	0.197	0.083	0.065	0.006	1.036	0.766	1.288
0.060	-6.956	2.068	-0.887	-0.061	0.194	0.082	0.064	0.005	1.021	0.751	1.268
0.062	-7.017	2.087	-0.876	-0.061	0.191	0.080	0.063	0.005	1.004	0.738	1.246
0.064	-7.075	2.106	-0.865	-0.062	0.187	0.078	0.062	0.005	0.981	0.725	1.220
0.067	-7.135	2.124	-0.853	-0.062	0.183	0.077	0.060	0.005	0.960	0.711	1.194
0.070	-7.213	2.143	-0.834	-0.063	0.179	0.075	0.059	0.005	0.941	0.697	1.171
0.072	-7.303	2.164	-0.815	-0.063	0.176	0.074	0.058	0.005	0.923	0.681	1.147
0.075	-7.383	2.185	-0.804	-0.063	0.172	0.072	0.057	0.005	0.906	0.667	1.125
0.078	-7.452	2.205	-0.798	-0.063	0.169	0.071	0.056	0.005	0.888	0.652	1.102
0.081	-7.537	2.229	-0.789	-0.063	0.166	0.070	0.054	0.005	0.873	0.639	1.082
0.084	-7.605	2.248	-0.789	-0.062	0.163	0.068	0.054	0.005	0.856	0.630	1.063
0.087	-7.683	2.270	-0.792	-0.060	0.160	0.067	0.053	0.005	0.841	0.619	1.045
0.091	-7.776	2.293	-0.787	-0.059	0.157	0.066	0.052	0.004	0.827	0.609	1.027
0.094	-7.880	2.319	-0.776	-0.060	0.155	0.065	0.051	0.004	0.813	0.600	1.011
0.098	-7.981	2.348	-0.774	-0.059	0.152	0.064	0.050	0.004	0.799	0.593	0.995
0.102	-8.079	2.375	-0.779	-0.058	0.150	0.063	0.050	0.004	0.787	0.589	0.983
0.106	-8.162	2.403	-0.797	-0.056	0.148	0.062	0.050	0.004	0.773	0.588	0.971
0.110	-8.243	2.425	-0.812	-0.055	0.147	0.061	0.050	0.004	0.761	0.585	0.960
0.114	-8.337	2.448	-0.821	-0.053	0.146	0.061	0.049	0.004	0.755	0.582	0.953
0.118	-8.432	2.470	-0.830	-0.051	0.145	0.060	0.049	0.004	0.752	0.579	0.949
0.123	-8.540	2.493	-0.834	-0.050	0.144	0.060	0.049	0.004	0.747	0.575	0.943
0.128	-8.651	2.517	-0.834	-0.049	0.143	0.060	0.049	0.004	0.739	0.572	0.935
0.133	-8.749	2.536	-0.837	-0.048	0.142	0.059	0.048	0.004	0.731	0.569	0.927
0.138	-8.838	2.553	-0.848	-0.046	0.140	0.058	0.048	0.004	0.723	0.565	0.918
0.143	-8.943	2.571	-0.852	-0.045	0.140	0.058	0.048	0.004	0.720	0.563	0.914
0.149	-9.041	2.589	-0.862	-0.044	0.139	0.058	0.048	0.004	0.714	0.561	0.908
0.154	-9.156	2.609	-0.871	-0.042	0.138	0.057	0.048	0.004	0.709	0.561	0.904
0.160	-9.330	2.647	-0.865	-0.040	0.139	0.057	0.048	0.004	0.697	0.559	0.894
0.167	-9.515	2.682	-0.838	-0.041	0.140	0.057	0.048	0.004	0.688	0.556	0.884
0.173	-9.671	2.715	-0.829	-0.041	0.141	0.058	0.048	0.004	0.682	0.553	0.878
0.180	-9.809	2.747	-0.825	-0.041	0.140	0.057	0.048	0.004	0.672	0.552	0.870
0.187	-9.951	2.765	-0.799	-0.043	0.140	0.057	0.048	0.004	0.666	0.552	0.865

0.194	-10.015	2.765	-0.796	-0.044	0.139	0.056	0.048	0.004	0.653	0.552	0.855
0.201	-10.118	2.783	-0.799	-0.044	0.141	0.057	0.049	0.004	0.658	0.554	0.860
0.209	-10.287	2.820	-0.791	-0.044	0.142	0.057	0.049	0.004	0.658	0.555	0.861
0.217	-10.397	2.843	-0.800	-0.043	0.143	0.057	0.049	0.004	0.656	0.558	0.861
0.226	-10.479	2.852	-0.811	-0.041	0.145	0.058	0.049	0.004	0.661	0.554	0.862
0.234	-10.606	2.879	-0.824	-0.039	0.145	0.058	0.049	0.004	0.652	0.556	0.857
0.244	-10.705	2.895	-0.830	-0.039	0.146	0.058	0.050	0.004	0.650	0.558	0.856
0.253	-10.835	2.923	-0.826	-0.041	0.137	0.052	0.051	0.004	0.567	0.569	0.803
0.263	-10.985	2.942	-0.808	-0.042	0.137	0.053	0.051	0.004	0.570	0.570	0.806
0.273	-11.110	2.959	-0.803	-0.042	0.148	0.058	0.051	0.004	0.640	0.559	0.850
0.283	-11.242	2.975	-0.792	-0.043	0.149	0.058	0.051	0.004	0.644	0.560	0.853
0.294	-11.345	2.977	-0.777	-0.044	0.151	0.059	0.052	0.005	0.643	0.562	0.854
0.306	-11.473	2.988	-0.754	-0.047	0.152	0.060	0.053	0.005	0.646	0.566	0.859
0.318	-11.594	2.989	-0.714	-0.050	0.154	0.060	0.053	0.005	0.653	0.563	0.862
0.330	-11.736	3.007	-0.697	-0.051	0.156	0.061	0.053	0.005	0.656	0.563	0.865
0.343	-11.824	3.011	-0.689	-0.052	0.155	0.061	0.053	0.005	0.649	0.563	0.859
0.356	-11.913	3.017	-0.692	-0.051	0.157	0.061	0.054	0.005	0.654	0.558	0.860
0.370	-12.047	3.038	-0.690	-0.051	0.159	0.062	0.054	0.005	0.663	0.558	0.867
0.384	-12.163	3.046	-0.673	-0.053	0.160	0.062	0.054	0.005	0.664	0.559	0.868
0.399	-12.294	3.060	-0.655	-0.054	0.164	0.064	0.055	0.005	0.675	0.558	0.875
0.415	-12.408	3.068	-0.645	-0.054	0.167	0.065	0.055	0.005	0.676	0.562	0.880
0.431	-12.487	3.063	-0.640	-0.054	0.168	0.065	0.056	0.005	0.680	0.564	0.883
0.447	-12.618	3.069	-0.621	-0.055	0.172	0.067	0.056	0.005	0.684	0.562	0.885
0.465	-12.673	3.066	-0.630	-0.054	0.176	0.067	0.059	0.005	0.685	0.569	0.890
0.483	-12.771	3.069	-0.626	-0.054	0.177	0.068	0.059	0.005	0.689	0.569	0.893
0.501	-12.829	3.055	-0.624	-0.053	0.182	0.070	0.060	0.005	0.696	0.573	0.902

Table .4: Regression coefficients and related uncertainty for Model 3 (corrected for site effect)

Periods	a	b	c	h	d	σ_a	σ_b	σ_c	σ_h	σ_d	σ inter-event	σ intra-event	σ_{total}
0.010	-7.397	2.117	-1.283	3.174	-0.030	0.395	0.081	0.181	0.469	0.010	1.021	0.735	1.258
0.010	-7.403	2.122	-1.283	3.178	-0.030	0.395	0.081	0.182	0.470	0.010	1.016	0.736	1.255
0.011	-7.418	2.129	-1.281	3.180	-0.031	0.395	0.081	0.182	0.471	0.010	1.009	0.736	1.249
0.011	-7.431	2.134	-1.280	3.177	-0.031	0.394	0.080	0.182	0.470	0.010	1.001	0.736	1.242
0.012	-7.438	2.138	-1.281	3.166	-0.031	0.392	0.079	0.181	0.469	0.010	0.991	0.737	1.236
0.012	-7.442	2.140	-1.282	3.157	-0.031	0.391	0.079	0.181	0.467	0.010	0.986	0.739	1.232
0.013	-7.441	2.142	-1.285	3.148	-0.031	0.391	0.079	0.181	0.466	0.010	0.983	0.741	1.231
0.013	-7.434	2.143	-1.289	3.143	-0.031	0.391	0.079	0.181	0.465	0.010	0.981	0.744	1.232
0.014	-7.426	2.143	-1.294	3.129	-0.031	0.391	0.079	0.181	0.463	0.010	0.978	0.748	1.231
0.014	-7.409	2.143	-1.303	3.126	-0.031	0.392	0.079	0.182	0.461	0.010	0.980	0.751	1.234
0.015	-7.387	2.141	-1.313	3.123	-0.030	0.393	0.079	0.182	0.460	0.010	0.979	0.755	1.236
0.015	-7.362	2.138	-1.321	3.120	-0.030	0.394	0.079	0.183	0.458	0.010	0.982	0.759	1.241
0.016	-7.337	2.135	-1.329	3.107	-0.029	0.395	0.079	0.183	0.456	0.010	0.983	0.763	1.245
0.016	-7.294	2.130	-1.342	3.117	-0.029	0.399	0.080	0.185	0.457	0.010	0.987	0.769	1.251
0.017	-7.244	2.123	-1.356	3.131	-0.028	0.404	0.080	0.188	0.458	0.010	0.992	0.775	1.259
0.018	-7.192	2.116	-1.369	3.141	-0.028	0.408	0.081	0.190	0.459	0.010	0.998	0.782	1.268
0.018	-7.118	2.108	-1.390	3.181	-0.027	0.418	0.081	0.194	0.463	0.010	1.007	0.789	1.279
0.019	-7.005	2.098	-1.425	3.273	-0.026	0.436	0.082	0.203	0.472	0.011	1.020	0.796	1.294
0.020	-6.891	2.088	-1.461	3.354	-0.025	0.453	0.083	0.210	0.479	0.011	1.030	0.805	1.307
0.021	-6.774	2.075	-1.495	3.437	-0.023	0.470	0.084	0.218	0.486	0.011	1.044	0.811	1.322
0.021	-6.640	2.062	-1.535	3.528	-0.022	0.489	0.085	0.227	0.493	0.012	1.056	0.818	1.336
0.022	-6.535	2.052	-1.566	3.583	-0.021	0.502	0.086	0.233	0.496	0.012	1.064	0.824	1.346
0.023	-6.469	2.046	-1.582	3.612	-0.021	0.510	0.086	0.236	0.499	0.012	1.069	0.830	1.353
0.024	-6.381	2.041	-1.611	3.663	-0.020	0.521	0.086	0.242	0.502	0.012	1.062	0.835	1.352
0.025	-6.321	2.038	-1.625	3.687	-0.019	0.529	0.086	0.245	0.505	0.012	1.064	0.841	1.357
0.026	-6.248	2.031	-1.641	3.724	-0.019	0.538	0.087	0.249	0.509	0.012	1.070	0.846	1.364
0.027	-6.214	2.028	-1.639	3.729	-0.019	0.541	0.087	0.251	0.513	0.013	1.072	0.850	1.369
0.028	-6.174	2.024	-1.637	3.756	-0.020	0.549	0.087	0.254	0.520	0.013	1.081	0.854	1.377
0.029	-6.177	2.024	-1.618	3.734	-0.021	0.547	0.088	0.254	0.525	0.013	1.087	0.858	1.384
0.030	-6.168	2.018	-1.598	3.712	-0.022	0.546	0.088	0.253	0.530	0.013	1.094	0.861	1.392
0.031	-6.167	2.017	-1.580	3.690	-0.023	0.543	0.089	0.252	0.533	0.013	1.100	0.864	1.398
0.033	-6.150	2.018	-1.569	3.695	-0.024	0.546	0.089	0.253	0.540	0.013	1.106	0.866	1.405
0.034	-6.154	2.020	-1.549	3.702	-0.025	0.548	0.090	0.254	0.548	0.013	1.109	0.866	1.408
0.035	-6.199	2.025	-1.509	3.686	-0.027	0.545	0.090	0.253	0.560	0.013	1.110	0.866	1.408
0.036	-6.250	2.031	-1.469	3.685	-0.029	0.544	0.089	0.252	0.575	0.013	1.106	0.866	1.405
0.038	-6.246	2.036	-1.452	3.754	-0.031	0.556	0.089	0.257	0.593	0.013	1.107	0.864	1.404
0.039	-6.201	2.042	-1.458	3.862	-0.031	0.574	0.090	0.265	0.609	0.013	1.113	0.860	1.406

0.041	-6.192	2.047	-1.448	3.937	-0.033	0.584	0.090	0.269	0.624	0.013	1.115	0.854	1.405
0.042	-6.199	2.053	-1.434	3.991	-0.034	0.590	0.089	0.272	0.636	0.013	1.110	0.848	1.397
0.044	-6.214	2.064	-1.422	4.040	-0.034	0.596	0.089	0.274	0.648	0.013	1.102	0.843	1.387
0.046	-6.182	2.070	-1.431	4.142	-0.034	0.608	0.088	0.279	0.656	0.013	1.094	0.833	1.375
0.048	-6.202	2.080	-1.419	4.195	-0.034	0.611	0.087	0.280	0.664	0.013	1.086	0.822	1.362
0.049	-6.236	2.091	-1.405	4.232	-0.034	0.611	0.086	0.280	0.669	0.013	1.077	0.813	1.349
0.051	-6.285	2.107	-1.390	4.240	-0.034	0.605	0.085	0.277	0.671	0.013	1.059	0.804	1.330
0.053	-6.304	2.124	-1.389	4.301	-0.034	0.608	0.084	0.278	0.675	0.013	1.043	0.794	1.311
0.055	-6.351	2.139	-1.377	4.321	-0.034	0.603	0.083	0.276	0.674	0.013	1.027	0.782	1.290
0.057	-6.506	2.155	-1.319	4.192	-0.035	0.570	0.081	0.261	0.666	0.012	1.009	0.768	1.268
0.060	-6.744	2.174	-1.224	3.962	-0.038	0.520	0.080	0.240	0.657	0.012	0.993	0.754	1.247
0.062	-7.001	2.194	-1.117	3.691	-0.043	0.469	0.078	0.217	0.649	0.011	0.971	0.741	1.222
0.064	-7.187	2.213	-1.041	3.503	-0.046	0.432	0.077	0.200	0.642	0.010	0.952	0.728	1.199
0.067	-7.330	2.231	-0.989	3.363	-0.048	0.404	0.075	0.187	0.631	0.010	0.929	0.715	1.172
0.070	-7.505	2.250	-0.922	3.189	-0.051	0.373	0.073	0.173	0.623	0.009	0.909	0.701	1.147
0.072	-7.631	2.270	-0.881	3.128	-0.052	0.357	0.072	0.166	0.622	0.009	0.891	0.685	1.124
0.075	-7.797	2.291	-0.825	2.948	-0.055	0.329	0.070	0.152	0.607	0.008	0.875	0.670	1.102
0.078	-7.941	2.311	-0.781	2.779	-0.056	0.303	0.069	0.140	0.587	0.008	0.858	0.656	1.080
0.081	-8.080	2.335	-0.743	2.648	-0.058	0.284	0.068	0.131	0.573	0.008	0.843	0.643	1.060
0.084	-8.204	2.354	-0.713	2.506	-0.058	0.266	0.067	0.122	0.554	0.007	0.827	0.633	1.041
0.087	-8.367	2.376	-0.671	2.271	-0.059	0.240	0.065	0.109	0.522	0.007	0.813	0.623	1.024
0.091	-8.517	2.400	-0.634	2.095	-0.060	0.222	0.064	0.100	0.500	0.006	0.799	0.613	1.007
0.094	-8.644	2.425	-0.609	2.009	-0.061	0.212	0.063	0.095	0.492	0.006	0.788	0.604	0.992
0.098	-8.747	2.453	-0.605	1.998	-0.061	0.209	0.062	0.093	0.487	0.006	0.776	0.596	0.978
0.102	-8.824	2.481	-0.620	2.066	-0.059	0.212	0.062	0.095	0.487	0.006	0.766	0.592	0.968
0.106	-8.893	2.509	-0.647	2.127	-0.057	0.215	0.061	0.098	0.480	0.006	0.756	0.591	0.960
0.110	-8.997	2.532	-0.649	2.075	-0.056	0.210	0.060	0.095	0.464	0.006	0.745	0.588	0.949
0.114	-9.079	2.555	-0.664	2.116	-0.054	0.212	0.060	0.096	0.460	0.006	0.741	0.585	0.944
0.118	-9.146	2.577	-0.686	2.193	-0.051	0.216	0.060	0.099	0.459	0.006	0.737	0.582	0.939
0.123	-9.230	2.599	-0.701	2.244	-0.049	0.219	0.059	0.100	0.457	0.006	0.733	0.578	0.934
0.128	-9.329	2.623	-0.707	2.261	-0.048	0.219	0.059	0.100	0.454	0.006	0.727	0.575	0.927
0.133	-9.423	2.642	-0.713	2.257	-0.047	0.217	0.058	0.099	0.447	0.006	0.719	0.572	0.919
0.138	-9.496	2.659	-0.733	2.291	-0.045	0.218	0.058	0.100	0.438	0.006	0.712	0.568	0.911
0.143	-9.562	2.676	-0.757	2.382	-0.043	0.224	0.058	0.103	0.439	0.006	0.712	0.565	0.909
0.149	-9.596	2.693	-0.799	2.530	-0.040	0.236	0.057	0.109	0.443	0.006	0.709	0.563	0.905
0.154	-9.696	2.713	-0.815	2.552	-0.037	0.237	0.057	0.110	0.438	0.006	0.706	0.562	0.902
0.160	-9.852	2.744	-0.807	2.554	-0.036	0.239	0.058	0.111	0.446	0.006	0.699	0.561	0.896
0.167	-10.026	2.775	-0.777	2.554	-0.038	0.240	0.058	0.111	0.468	0.006	0.692	0.557	0.889
0.173	-10.155	2.802	-0.773	2.585	-0.037	0.244	0.058	0.112	0.476	0.007	0.689	0.555	0.884
0.180	-10.337	2.836	-0.747	2.469	-0.038	0.234	0.057	0.107	0.470	0.006	0.675	0.554	0.873
0.187	-10.476	2.853	-0.720	2.464	-0.040	0.234	0.057	0.108	0.491	0.006	0.668	0.554	0.868

0.194	-10.514	2.853	-0.729	2.563	-0.041	0.242	0.057	0.112	0.506	0.007	0.660	0.554	0.861
0.201	-10.570	2.870	-0.753	2.676	-0.040	0.255	0.057	0.118	0.520	0.007	0.666	0.556	0.868
0.209	-10.731	2.904	-0.746	2.685	-0.040	0.259	0.058	0.120	0.532	0.007	0.670	0.557	0.871
0.217	-10.860	2.926	-0.744	2.618	-0.040	0.255	0.058	0.118	0.526	0.007	0.668	0.559	0.871
0.226	-10.870	2.933	-0.789	2.780	-0.037	0.271	0.059	0.125	0.530	0.007	0.675	0.555	0.874
0.234	-11.018	2.960	-0.788	2.710	-0.035	0.266	0.059	0.123	0.519	0.007	0.666	0.557	0.868
0.244	-11.202	2.976	-0.751	2.488	-0.037	0.248	0.059	0.114	0.504	0.007	0.659	0.560	0.865
0.253	-11.430	3.000	-0.689	2.222	-0.042	0.221	0.053	0.104	0.494	0.006	0.570	0.572	0.808
0.263	-11.568	3.019	-0.674	2.244	-0.043	0.224	0.053	0.105	0.514	0.006	0.573	0.574	0.811
0.273	-11.554	3.032	-0.740	2.594	-0.039	0.262	0.059	0.121	0.544	0.007	0.656	0.560	0.863
0.283	-11.679	3.047	-0.729	2.605	-0.040	0.265	0.060	0.122	0.560	0.007	0.661	0.562	0.867
0.294	-11.790	3.047	-0.705	2.598	-0.042	0.268	0.061	0.123	0.585	0.007	0.662	0.564	0.869
0.306	-11.966	3.057	-0.654	2.480	-0.046	0.259	0.061	0.119	0.605	0.007	0.665	0.568	0.874
0.318	-12.166	3.055	-0.570	2.184	-0.052	0.235	0.061	0.106	0.616	0.007	0.665	0.565	0.873
0.330	-12.284	3.072	-0.565	2.248	-0.052	0.242	0.062	0.109	0.641	0.007	0.669	0.565	0.875
0.343	-12.350	3.076	-0.569	2.307	-0.053	0.246	0.062	0.112	0.656	0.007	0.661	0.565	0.869
0.356	-12.440	3.081	-0.570	2.304	-0.052	0.247	0.062	0.112	0.655	0.007	0.667	0.560	0.870
0.370	-12.575	3.103	-0.568	2.293	-0.052	0.248	0.063	0.112	0.657	0.007	0.674	0.560	0.876
0.384	-12.704	3.109	-0.542	2.196	-0.054	0.242	0.063	0.109	0.664	0.007	0.676	0.561	0.878
0.399	-12.812	3.119	-0.534	2.182	-0.054	0.245	0.065	0.110	0.684	0.007	0.684	0.559	0.884
0.415	-12.925	3.126	-0.523	2.158	-0.055	0.246	0.066	0.110	0.698	0.007	0.687	0.563	0.889
0.431	-12.993	3.121	-0.522	2.182	-0.054	0.250	0.066	0.112	0.711	0.007	0.690	0.565	0.891
0.447	-13.122	3.123	-0.497	2.121	-0.056	0.247	0.067	0.109	0.733	0.007	0.691	0.564	0.892
0.465	-13.227	3.117	-0.478	1.861	-0.056	0.236	0.068	0.103	0.716	0.007	0.692	0.569	0.896
0.483	-13.307	3.119	-0.481	1.913	-0.056	0.241	0.069	0.106	0.731	0.007	0.698	0.569	0.901
0.501	-13.328	3.104	-0.497	2.034	-0.054	0.257	0.070	0.113	0.761	0.007	0.704	0.574	0.908

B. Publications

Deciphering shallow mantle stratification through information dimension

D.S. Ramesh, P. Appala Raju, Nitin Sharma[†], and S. Das Sharma
NATIONAL GEOPHYSICAL RESEARCH INSTITUTE (COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH), HYDERABAD 500007, INDIA

ABSTRACT

Applying the concept of cluster entropy, we demonstrate that the primary P-to-s (Ps) conversions originating from the well-recognized seismological boundaries (Moho, 410 km, and 660 km depth interfaces) in Earth's interior and multiple reflections (reverberations) between the surface and the Moho—the prominent crust-mantle boundary of Earth—can be unambiguously identified. Their attendant information dimension can act as a discriminant. Documentation of shallow mantle stratification at depths ~150–300 km (L/X layers) using Ps signals is scarce and is sometimes attributed to their weak registration in the seismograms together with interference by reverberations from shallow structure. The slopes of best fits to mean cluster information dimension (as function of epicentral distance) of Ps signals from target boundaries such as the Moho, 410 km, and L/X discontinuities are consistently gentle with a slight negative character. In contrast, those related to multiples show a steep positive behavior. This opposite nature can potentially discriminate between reverberations and direct converted waves, even when they tend to arrive in overlapping time windows. Our approach thus enables unequivocal identification of shallow mantle layering from Ps data recorded at diverse tectonic provenances of wide antiquity.

LITHOSPHERE; v. 2; no. 6; p. 462–471. doi: 10.1130/L117.1

INTRODUCTION

A stratified Earth model receives support from the presence of prominent seismic velocity discontinuities at average depths of (1) 33 km (termed as the Moho), (2) 410 km (410 km), and (3) 660 km (660 km) detected globally, in addition to the discovery of a shallow mantle boundary at ~200 km, first recognized and reported by Lehmann (1959), which is restricted to certain regions. Since then, several researchers have documented the presence of shallow mantle layers in diverse tectonic settings using a variety of seismic waves at varied depths (Revenaugh and Jordan, 1991; Li et al., 2002; Gung et al., 2003; Deuss and Woodhouse, 2004). Based on their depth disposition, these boundaries are termed the Lehmann (L, 180–250 km) and X- discontinuities (250–330 km). With the advent of P-to-s (Ps) receiver function technique (Vinnik, 1977), owing to its superior vertical resolution imaging power, unambiguous detection of a stratified mantle by way of well-documented presence of the 410 km and 660 km discontinuities almost everywhere became possible. For the same reason, delineation of shallow mantle layering seemed promising. In spite of their better sensitivity to vertical layering, observations related to the possible presence of the shallow depth boundaries ~150–300 km in the Ps receiver function data generally remained scarce for several reasons. There is a perception that weaker registrations of Ps conversions from these shallow mantle depth layers, compared to those related to the well-documented 410 km and 660 km discontinuities, often render the former relatively ambiguous (e.g., see Grunewald et al., 2001). On the contrary, the amplitudes of Ps conversions related to the Lehmann discontinuity (~250 km; PLs) are the highest beneath the northwestern half of the Colorado Plateau–Rio Grande Rift–Great Plains Seismic Transect (LA RISTRA) (Wilson et al., 2005) and eastern segment of the Massachusetts Broadband Seismometer Experiment (MOMA) profile (Li et al., 2002). Notwithstanding these data, importantly, primary Ps converted waves from the 150–300 km depth interfaces and reverberations (multiples) due to shallow structure can arrive in overlapping time windows. Several studies can be cited in this context: (1) Li et al. (2002) reported complex scattering originating from shallow structure in the western segment of the MOMA profile that interfered with PLs-like arrivals; (2) the nature of arrivals corresponding to 250–300 km depth beneath the Tanzania craton remains inconclusive (Owens et al., 2000); and (3) the problem of unambiguous detection of PLs arising due to interference by multiples from Moho and sediments was reported by Landes et al. (2006). Traditional methods based on moveout trends (Yuan et al., 1997) remained unhelpful in such situations. Such a dilemma was also expressed recently while dealing with a negative phase in study of the lithosphere-asthenosphere boundary in North America from receiver functions (Abt et al., 2010). Details of the factors that render detection of depth boundaries at ~150–300 km in the Ps data relatively ambiguous and contentious can be found in Sheehan et al. (2000). Hence, given the proximity of crustal reverberations to possible direct converted phases, such as the PLs in our study, this signal cannot be unambiguously attributed to a direct phase.

The opposing trends of the moveout curves for Ps and the multiples based on slowness criteria (Yuan et al., 1997) together with some innovative data segregation (Ramesh et al., 2005), or with S-to-p (Sp) data as a guide (Farra and Vinnik, 2000; Ramesh et al., 2010), would reduce the ambiguity associated with identification of Ps conversions from shallow mantle boundaries. Though the opposite moveout of primary conversions and multiples can potentially discriminate between them, its success becomes partial when the receiver functions are restricted to a narrow range of slowness values, and also due to possibly small moveout values of PLs-like phase arrival times with ray parameter (Sheehan et al., 2000). Therefore, techniques other than those based on slowness criteria alone deserve attention. One fundamental reason for the limited success

---

[†] Present address: Department of Physics, Laboratorio di Ricerca in Sismologia Sperimentale e Computazionale, University of Naples "Federico II," Via Cintia, 26, 80126, Monte S. Angelo, Napoli, Italy.

of the moveout-based discrimination of seismic phases could be that we are still operating in the measurement space itself. Instead, it may be interesting to view data in the generic space where the inherent differences related to the origin of various seismological phases could be preserved better. One such approach involves application of principles of information theory to our observations.

Use of information theory in general, particularly employing methods based on measures of entropy to solve geophysical problems, is still in its infancy and has gained momentum in recent years. For example, various measures of entropy were successfully applied to investigate dynamical complexity of the Earth's system (Earth's magnetosphere, to be precise) recently by Balasis et al. (2009). Their results provide conclusive evidence for a significant decrease in complexity and accession of persistency in disturbance storm time (D_{st}) series as a magnetic storm approaches. This finding, according to them, can be used as a diagnostic tool for detection of magnetospheric injury (global instability). To our knowledge, use of measures of entropy in seismology is yet to gain prominence. The present study to discern seismic phases of varied origin by applying a measure of cluster entropy is perhaps a pioneering attempt to understand Earth's stratification.

Documentation of shallow mantle layering assumes significance both in the context of mantle stratification and candidate mineral transformations at these depth regimes (Deuss and Woodhouse, 2004; Ganguly and Frost, 2006). In this work, based on the concept of cluster entropy (Carbone et al., 2004; Carbone, 2007; Carbone and Stanley, 2007; Balasis et al., 2009; see also Shannon, 1948), we determine the cluster information dimension of primary Ps conversions from various depth discontinuities (e.g., Moho:Pms; Lehmann:PLs; 410 km:P410s) and the crustal reverberation phases (e.g., Pps and Pss; see Fig. 1). We further demonstrate that the cluster information dimension can indeed discriminate between direct converted waves and multiply reflected phases registered in Ps receiver functions.

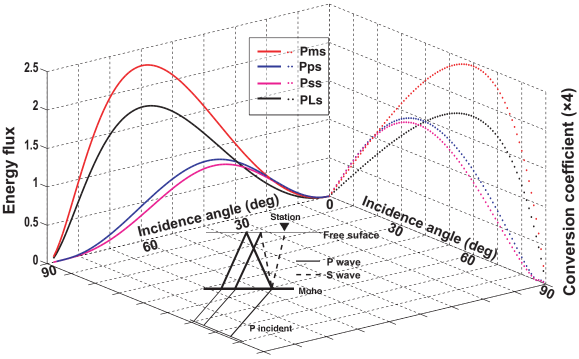


Figure 1. Energy flux (solid curves) and conversion coefficients (stippled curves) of some prominent seismic phases recorded in a Ps receiver function. Conversion coefficients (displacement amplitude) and energy fluxes associated with depth converted phases Pms and PLs together with crustal multiples Pps and Pss are shown as a function of incidence angle. A schematic ray diagram depicting Pms, Pps, and Pss phases is also shown. Calculations for P410s and P660s phases were also performed but are not shown here. Computations were carried out using parameters from the PREM standard earth model. Average parameter values until 1000 km depth below the target conversion boundary (e.g., Moho:Pms; 220 km:PLs; 400 km:P410s, etc.) and above it (to 0 km) were used to arrive at the corresponding P and S velocities and densities across the boundary of interest in our computations. All computations were done using MATLAB 7.6.

MOTIVATION, METHODS, AND DATA

Background

The reflection and transmission coefficients together with the energy flux associated with converted phases such as Pms, PLs, and crustal multiples (viz. Pps and Pss) as a function of incidence angle were computed and are presented in Figure 1 based on approaches suggested by Aki and Richards (1980) and verified using Young and Braile (1976). The most significant observation is that the conversion coefficients (displacement amplitudes) and the associated energy flux values of Ps conversions (e.g., Pms and PLs phases), and the reverberating phases (Pps and Pss) follow distinctly different distributions (Fig. 1). Also, the varied converted and multiple phases show internal consistency in consonance with their nature. Based on the typical ray geometries of converted and multiple phases and their attendant distinct energy flux distributions shown above (see Fig. 1), it is reasonable to anticipate that their respective average amount of information (entropy) could be different. This possible difference in entropy between the direct conversions and reverberation phases would result in divergent information dimensions of these seismic arrivals. Hence, some measure of attendant information dimension can perhaps potentially discriminate between primary converted waves and multiply reflected phases that might arrive even in overlapping time windows.

Method to Estimate Cluster Information Dimension (D_{Cluster})

In order to establish tangible criteria for the detection of shallow mantle layers and to have better insight into the nature of varied seismic phases registered in Ps receiver function data, we calculated cluster entropy of a time series adapting the approaches mentioned in Carbone et al. (2004), Carbone (2007), Carbone and Stanley (2007), and Balasis et al. (2009).

In information theory, any data string "X" is termed as a source, and the data points constitute the source alphabet. Prior to sampling of this source, there is a certain amount of uncertainty associated with the outcome, which turns into some information on sampling. Thus, concepts of uncertainty and information are related. Cluster entropy is defined as "S" = $-\Sigma\, P_i \times \log P_i$, where P_i is the probability of occurrence of *i*th event. The quantity $\log P_i$ is the amount of information one gains when an event with probability of occurrence P_i happens. Hence, the average amount of information or the expected value of information for one trial of experiment is the quantity "S." When all events are equally likely, then $P_i = 1/N$, where N is the number of options. Hence, the average information will be $\log P_i = \log N$. To calculate the cluster entropy "S" of a given data string, the samples are first grouped into histograms of stipulated bin width, and the area under the histogram is normalized to yield the probability distribution. The scaling exponent (factor) that describes the scaling of the computed entropy with bin width is termed the cluster information dimension. Therefore, cluster information dimension (D_{Cluster}) is calculated using $D_{\text{Cluster}} = \Sigma\, P_i(r) \times \log P_i(r)/\Sigma\, P_i(r) \times \log r_i$, where r_i denotes the bin width of the histogram.

The important steps used to calculate the cluster information dimension as applicable to our study in a practical recipe form are: (1) imparting stationarity to the time series by applying a moving average to the data (Fig. 2), and (2) obtaining the differential amplitudes between the time series and its moving average (Fig. 2), to result in a stochastic sequence, which is followed by (3) construction of frequency distributions of this amplitude sequence in order to compute its cluster entropy using the corresponding probability distribution functions associated with each class. (4) Later, from the ensemble of entropy of all the classes, the cluster information dimension of the seismic phase under consideration is estimated

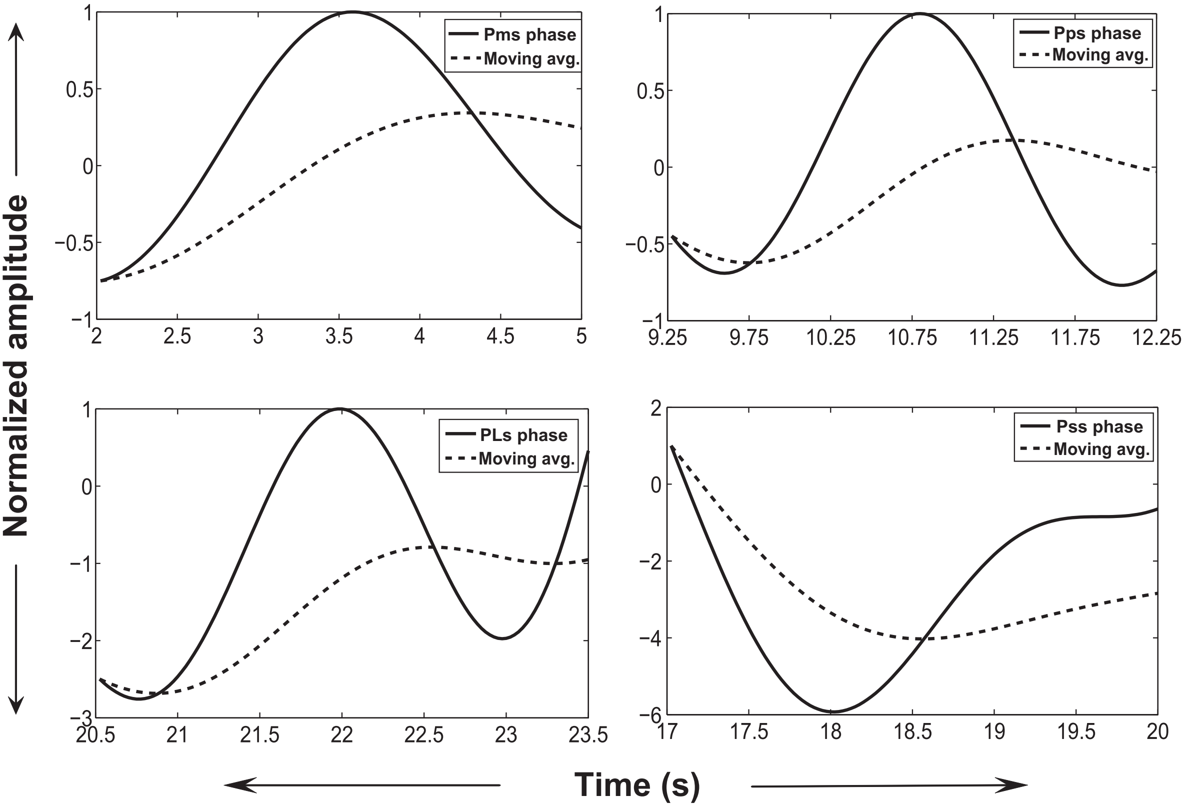


Figure 2. Representative seismic phases and moving average curves. Examples are given of seismic phases (solid curves) of interest along with moving averages (stippled curves) showing the approach to impart stationarity to the recorded time series.

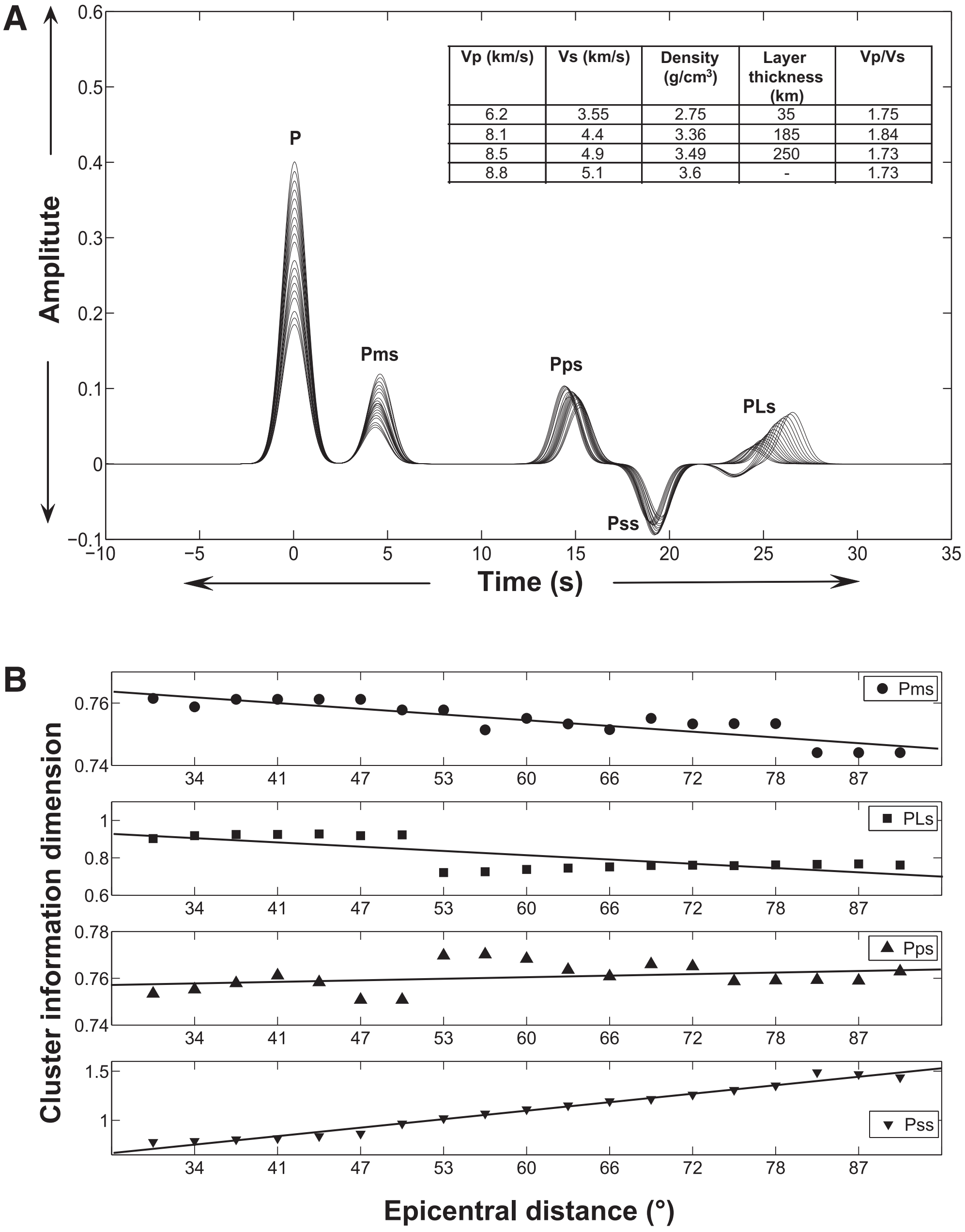
according to the previous formula for D_{Cluster}, and (5) finally, the entire process (1–4) is repeated on each receiver function to enable us to compute D_{Cluster} of various seismic phases (e.g., Pms, PLs, P410s, Pps, Pss, etc.) for the total epicentral distance range of the data.

Based on this foundation, we first demonstrate the efficacy of our technique on synthetic receiver function data generated using a model with the "Lehmann" boundary incorporated (see Fig. 3A). The synthetic data were obtained adopting standard methods (e.g., Haskell, 1962; Langston, 1977, 1979). The D_{Cluster} values as a function of epicentral distance for various seismic phases (Pms, PLs, Pps, and Pss) along with their best fits are shown in Figure 3B. It is interesting to note that the best-fit slopes of primary converted phases (Pms and PLs) show a negative trend, in contrast to the positive slopes exhibited by the multiples (Pps and Pss phases). This opposing slope-character distilled from the synthetic seismogram data could form the fundamental basis to discriminate between depth converted phases and multiply reflected seismic phases.

Application to the Observed Data

With this premise, a few hundreds of broad band data from six seismic stations distributed mainly in India and North America, and published by us earlier (Ramesh et al., 2002, 2010), were reanalyzed. Two of these stations (HYB—Hyderabad; CUD—Cudappah) are located in southern India, two (FFC—Flin Flon; FRB—Frobisher Bay) are situated in Canada, while the remaining (HRV—Harvard; PAS—Pasadena) are located in the United States. We followed the methods described by Vinnik (1977) to construct and treat the Ps receiver functions. The number of data, their quality, and processing details together with the published main results for these stations from India and North America can be found respectively in Ramesh et al. (2010) and Ramesh et al. (2002). Following Vinnik (1977), to obtain a Ps receiver function, we first rotated the original ZNE recordings to ZRT using back-azimuth information. Subsequently, the Z and R components were further rotated into local ray coordinate system L (P energy) and Q (SV energy) components involving the angle of incidence for better isolation of the P and SV energy from the incident wave field. This was followed by deconvolution of the L component from the Q component to obtain P-to-s (Ps) receiver functions.

The Ps distance moveout-corrected (Yuan et al., 1997) receiver function stack sections for six stations are presented in Figure 4. Apart from the well-recognized Pms, Pps, and Pss arrivals, positive polarity energy corresponding to possible PLs(?) arrivals can be observed in the data at ~20 s consistently across a large distance range (slowness) at three seismic stations, namely, PAS, CUD, and HYB (Figs. 4B–4D). This additional feature (PLs), though observed at these stations in the aforementioned studies, was not discussed earlier in view of its arrival proximity to the multiple phases (Pps and Pss), and its often inclined nature akin



Vp (km/s)	Vs (km/s)	Density (g/cm^3)	Layer thickness (km)	Vp/Vs
6.2	3.55	2.75	35	1.75
8.1	4.4	3.36	185	1.84
8.5	4.9	3.49	250	1.73
8.8	5.1	3.6	-	1.73

Figure 3. Application of the method to synthetic data. (A) Synthetic receiver functions generated using the model presented in the inset. Synthetics are in the distances between 29° and 90° corresponding to slowness range 0.08–0.042 s/km in steps of 0.002 s/km. The receiver function phases are Pms (Ps conversion from the Moho boundary at 35 km depth) and PLs (Ps conversion from "Lehmann" boundary placed at 220 km depth), and the crustal multiple phases are Pps and Pss, for which ray paths are also shown in Figure 1. Note that the synthetic seismograms are presented in the ZRT system, as opposed to the LQT used for our real data. (B) Cluster information dimension for 19 synthetic receiver functions shown in Figure 3A as a function of epicentral distance for various seismic phases (Pms, PLs, Pps, and Pss).

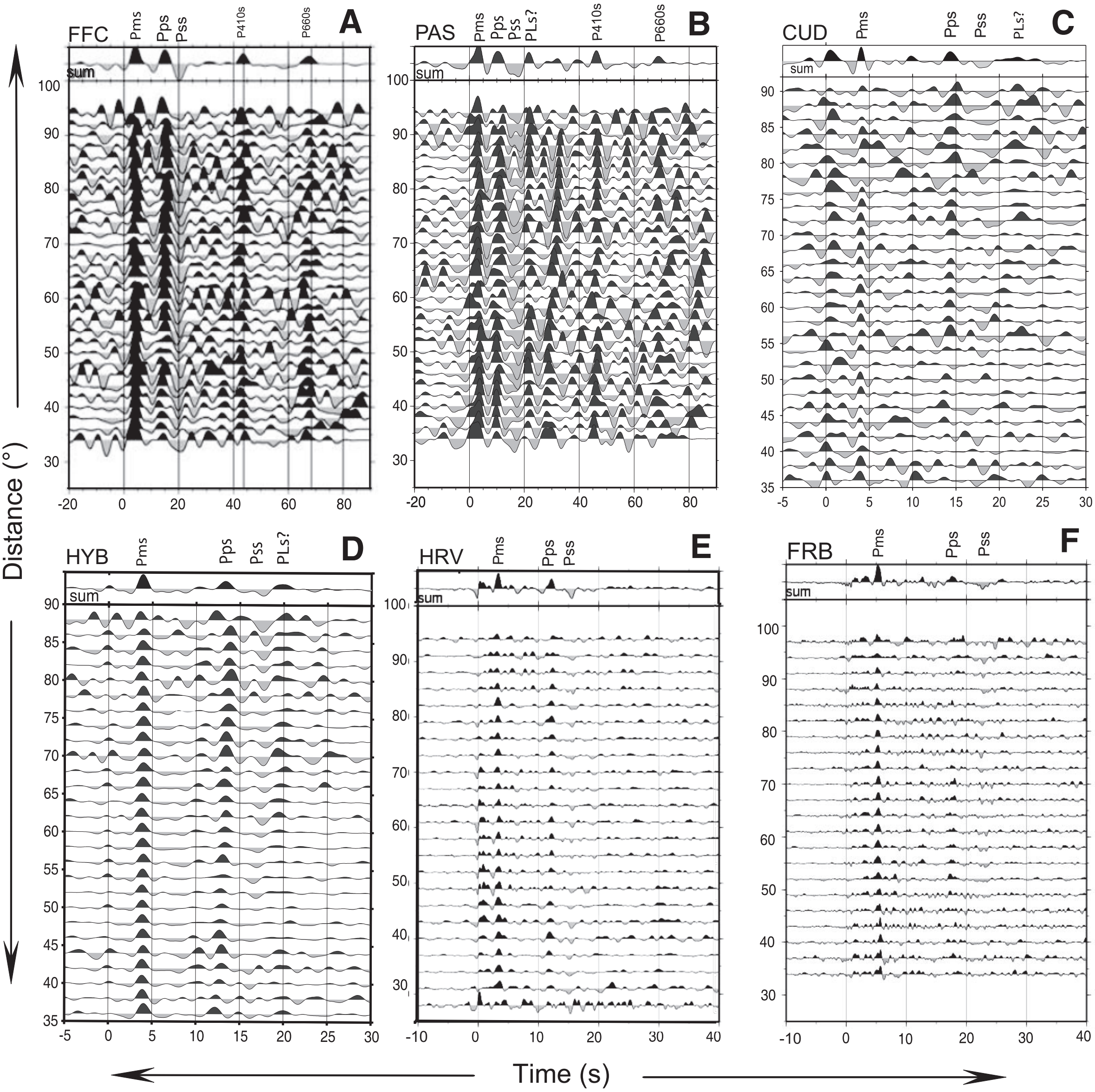


Figure 4. Ps moveout-corrected P receiver function distance stack sections at all the six stations (see text for station locations). In the top graph, the corresponding sum trace is shown. The Moho, Pms arrival, possible Lehmann-related arrival (220 km, PLs), conversions from the 410 km (P410s) and 660 km (P660s) boundaries, in addition to the crustal multiples, Pps and Pss phases, are marked in the sum traces.

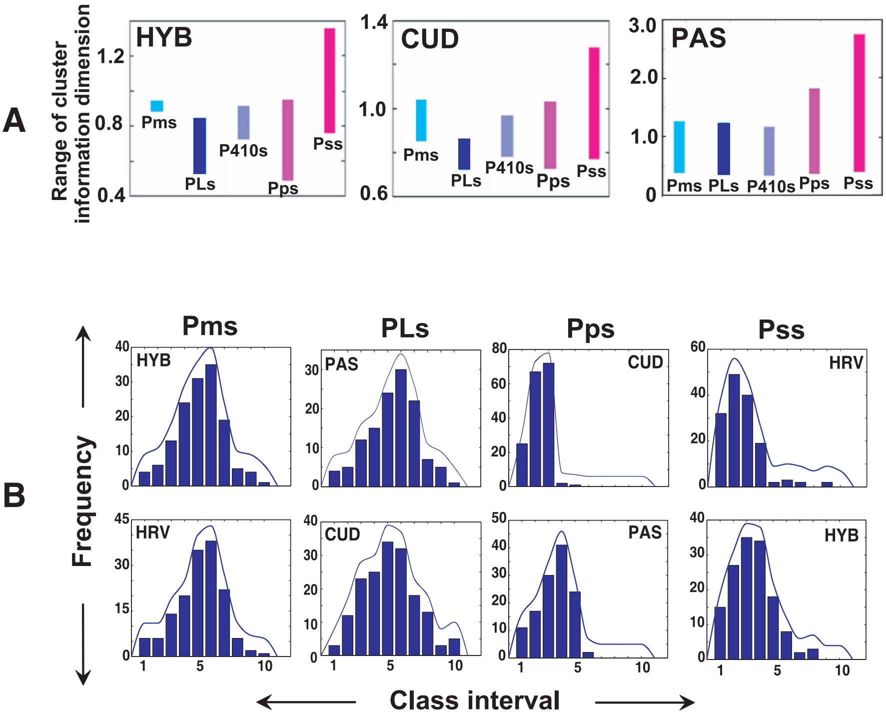


Figure 5. Examples of spread in cluster information dimension ($D_{Cluster}$) and histogram plots of $D_{Cluster}$ associated with various seismic phases. (A) The range of $D_{Cluster}$ for converted phases Pms, PLs, and P410s and crustal multiple phases Pps and Pss computed for stations HYB, CUD, and PAS. Primary conversions generally show smaller spreads compared to the large scatter shown by multiples. (B) Cluster information dimension histograms (bars) and corresponding probability distribution function (*pdf*) curves show the distinct character of the direct converted phases and multiple arrivals. Note mutual consistency in *pdf* patterns of primary converted waves (between Pms and PLs) and the multiply reflected waves (i.e., Pps and Pss). This can be prescribed as one diagnostic to discriminate conversions from multiples based on $D_{Cluster}$ alone.

to the multiples (see Figs. 4B–4D) in the Ps moveout-corrected sections. These factors render identification of possible PLs phase as a primary conversion ambiguous. For these reasons, we then preferred to err on the side of caution, especially in the absence of robust primary and multiple phase discriminants, in addition to the lack of additional constraints or information in the form of Sp receiver functions, which are now available at some locations to unambiguously identify these signals (see, for example, Ramesh et al., 2010).

Therefore, our motivation to apply the concept of cluster entropy (Carbone, 2007; Balasis et al., 2009; see also Shannon, 1948) to study the nature of energetic arrivals close to the identified Pps and Pss arrivals, preferably within a 20–30 s delay time window, present in the Ps moveout-corrected stack sections (e.g., Fig. 4) of our previously cited publications seems reasonable. Thus, we computed the cluster entropy and the corresponding cluster information dimension for the well-recognized direct converted phases Pms and P410s, and multiples Pps and Pss, in addition to the ambiguous PLs arrivals.

In computation of the cluster entropy, ~15% of data did not yield a cluster using the attendant moving averages of the corresponding seismic phases. Temporal variations in the ambient noise beneath a seismic station could perhaps be one contributing factor to this. This minor limitation can however be overcome with availability of moderate amounts of data.

RESULTS AND DISCUSSION

The computed cluster information dimension of the seismic arrivals designated as Pms, PLs, and P410s, in general, share a narrow restricted range, as opposed to the wide scatter exhibited by the multiples Pps and Pss (Fig. 5A). This indicates that the range of uncertainty associated with the ambiguous PLs phase reflects more commonality with the recognized converted waves (Pms and P410s) compared to the well-established reverberated arrivals (Pps and Pss). This could perhaps serve as one diagnostic to discriminate between direct conversions and the multiples.

We constructed the probability distribution functions (*pdf*) of the cluster information dimension corresponding to the seismic arrivals Pms, PLs, P410s, Pps, and Pss at various seismic stations. Samples of such *pdf*s at representative stations presented in Figure 5B illustrate the distinct distribution character of the converted waves and the multiples. This observation is valid in most of the cases. Such contrasting behavior of the candidate phases is already documented in their respective energy flux and conversion coefficient distributions (Fig. 1). This observation reinforces the fact that reverberation phases seem to follow a slightly skewed distribution, while primary conversions from target boundaries within Earth are invariably associated with more symmetric distributions (Fig. 5B). This could serve as another diagnostic to discriminate shallow mantle primary conversions (like the PLs phase) from the multiple arrivals.

To further distill our diagnostics and observations in order to arrive at a more robust measure of the discriminant, the cluster information dimension ($D_{Cluster}$) of Pms, PLs, P410s, Pps, and Pss seismic phases, as a function of epicentral distance, was studied in two ways. The first deals with variation of $D_{Cluster}$ corresponding to each receiver function with epicentral distance (Fig. 6), which replicates the behavior already documented by the synthetic data presented in Figure 3B.

In the other, the Ps receiver functions obtained over a wide distance range (30°–90°; 1° ≈ 111 km) at each station were grouped into uniform-sized distance bins. The choice of the distance bin width at each station

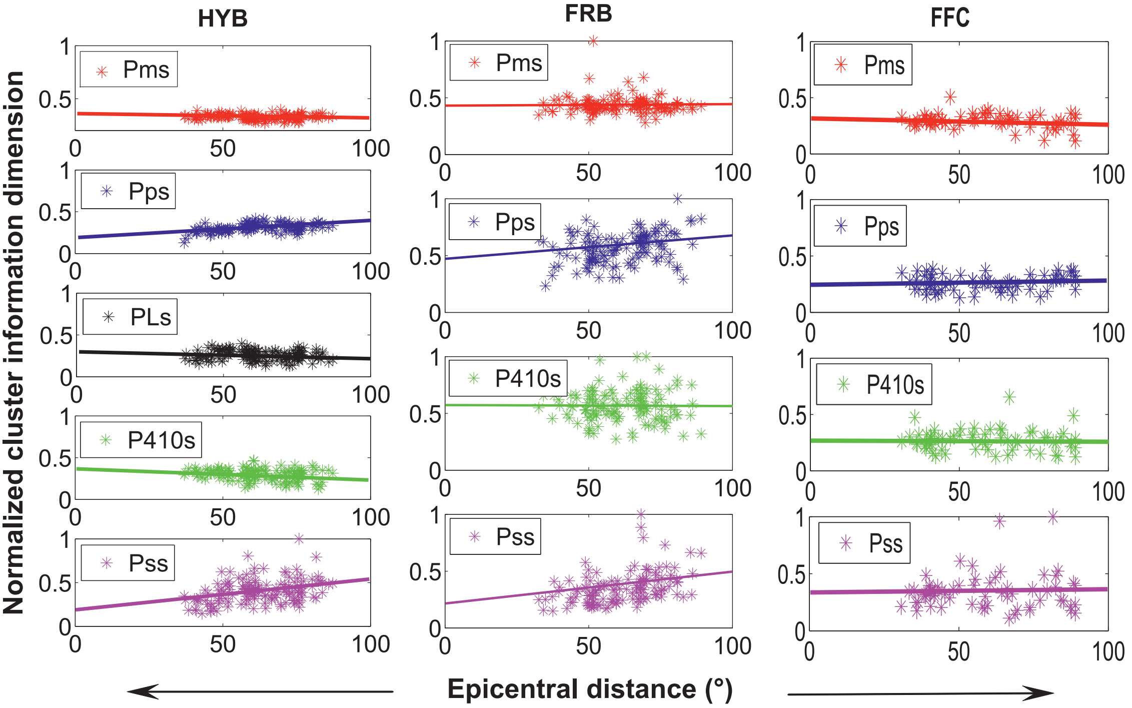


Figure 6. $D_{Cluster}$ for each receiver function at a few sample stations. Cluster information dimensions for various seismic phases from individual receiver functions retain the same trends as exhibited by synthetic data results (Fig. 3B).

varied in order to maintain a comparable number of receiver functions per distance bin. The mean value of the cluster information dimension associated with the stated direct and multiple phases corresponding to each distance range (bin width) were then computed. A weighted least square fit corresponding to each seismic phase mean information dimension at a few stations sampling diverse tectonic regimes of varied ages is shown in Figure 7. Also, individual mean information dimension values for phases Pms and Pps are shown. Importantly, all the primary converted phases (Pms, P410s) show nearly flat to marginally negative best-fit slopes, while those arising from the multiples (Pps and Pss) are characterized contrastingly by relatively steep positive slopes. Significantly, the weighted best fit for the ambiguous PLs phase shares similar attitude and attributes as the confirmed direct converted waves.

From this, we find that the observed opposing slope-character is indeed an important diagnostic to discriminate between primary converted waves and multiply reflected phases. Our results from individual (Fig. 6) and binned receiver function data (Fig. 7) imply that the nature of observed opposing slopes for direct and reverberated phases would be preserved even when some epicentral distances remain unrepresented or under-represented. We would therefore like to add that a large amount of data is not a stringent requirement in our approach to extract the diagnostics, because even moderate amount of data (Fig. 3B) seems to yield the desired results without influencing our final conclusions significantly.

Another important issue that arises is related to the choice of bin width and its influence on the observed slopes. This is addressed by focusing on the PLs observations recorded at stations PAS, HYB, and CUD, where the distance bin widths at each station were varied in order to maintain a comparable number of receiver functions per distance bin. Figure 8 demonstrates that the nature of the best-fit slope to the PLs phase at these stations is still preserved as in individual station plots (Fig. 7), directly suggesting that the choice of bin width has no significant role in determination of $D_{Cluster}$ and on our final results.

Plots similar to these (Fig. 7), combining cluster information dimension results from all the six stations corresponding to each seismic phase, are presented in Figure 9. Considering their location on distant continents of the globe with different age provinces of tectonic diversity, the consistent adherence to gentle-negative slopes by the direct phases (Pms, PLs, and P410s) and to steep positive slopes by the multiples (Pps and Pss) in the mean cluster information dimension versus average epicentral distance bin plots at all locations (Fig. 9) is certainly remarkable. This therefore constitutes a robust measure for discriminating between the competing seismological phases of interest.

CONCLUSIONS AND IMPLICATIONS

This is the first attempt to discriminate between various seismological phases, such as the primary converted waves arising from several depth boundaries in Earth's interior, and the multiply reflected wave types using the concept of cluster entropy and related information dimension. After demonstrating that the distributions of amplitude displacements (conversion coefficients) and corresponding energy fluxes of the direct converted waves and reverberation phases are different, we exploit these dissimilar characteristics to compute their cluster information dimension. Three distinctive diagnostics have emerged to discriminate between primary

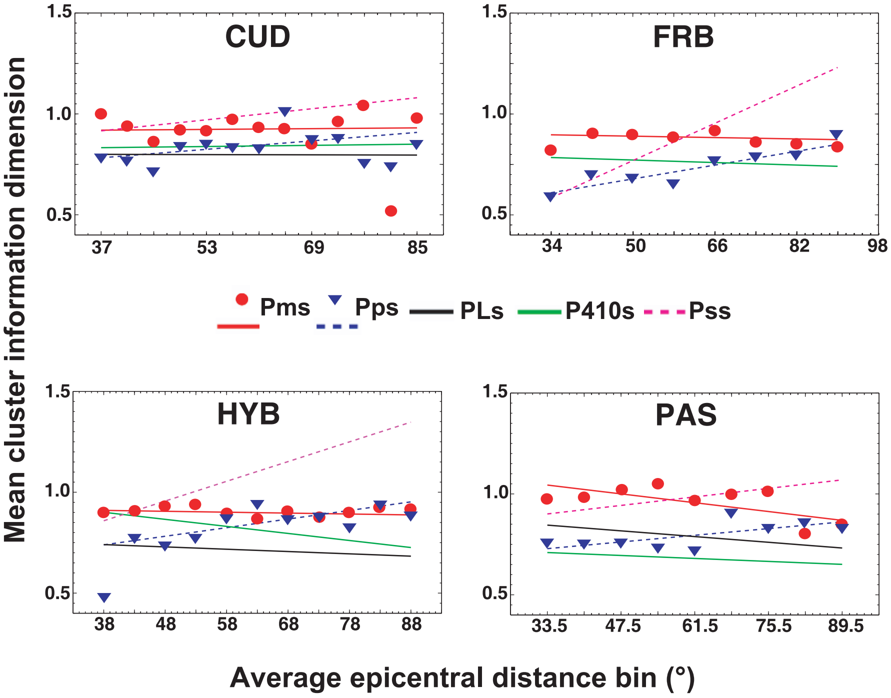


Figure 7. Mean $D_{Cluster}$ of prominent Ps receiver function phases recorded at various stations. A weighted least square fit to variation of mean cluster information dimension with epicentral distance bin corresponding to each seismic phase is presented at a few stations of the study. Solid line fits correspond to primary converted waves, while the dashed lines are best fits for multiples. The direct converted phases (Pms and P410s) and the ambiguous PLs share consistent gentle negative best-fit slopes, while in contrast the multiples (Pps and Pss) follow a decisive positive steep slope. Therefore, the best-fit slopes of $D_{Cluster}$ of the seismic phases can distinguish between the direct converted phases and the multiples.

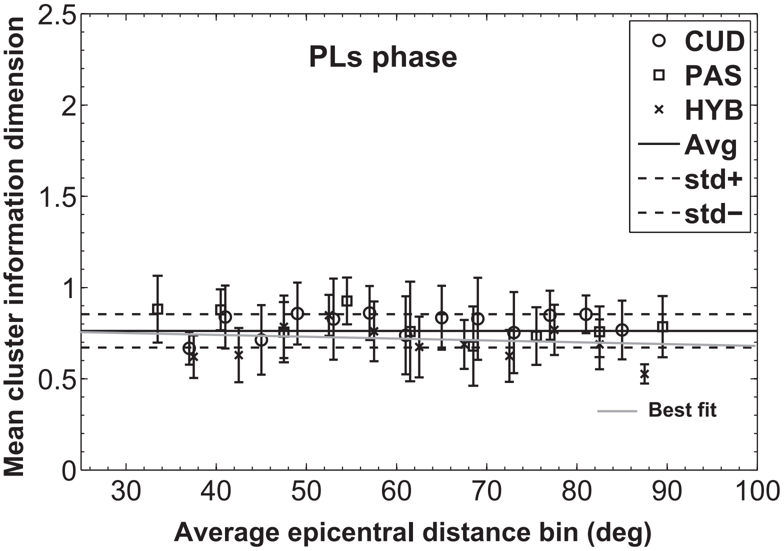


Figure 8. $D_{Cluster}$ values of the PLs phase recorded at stations PAS, HYB, and CUD. Error bars associated with each $D_{Cluster}$ value for the PLs phase recorded at individual stations are shown. The average (Avg) and corresponding standard deviation (std ±) lines are for data recorded at all three stations. The best fit (gray line) for the entire data set (all stations combined) is same as that shown for the PLs phases in Figure 9 and is superposed from it.

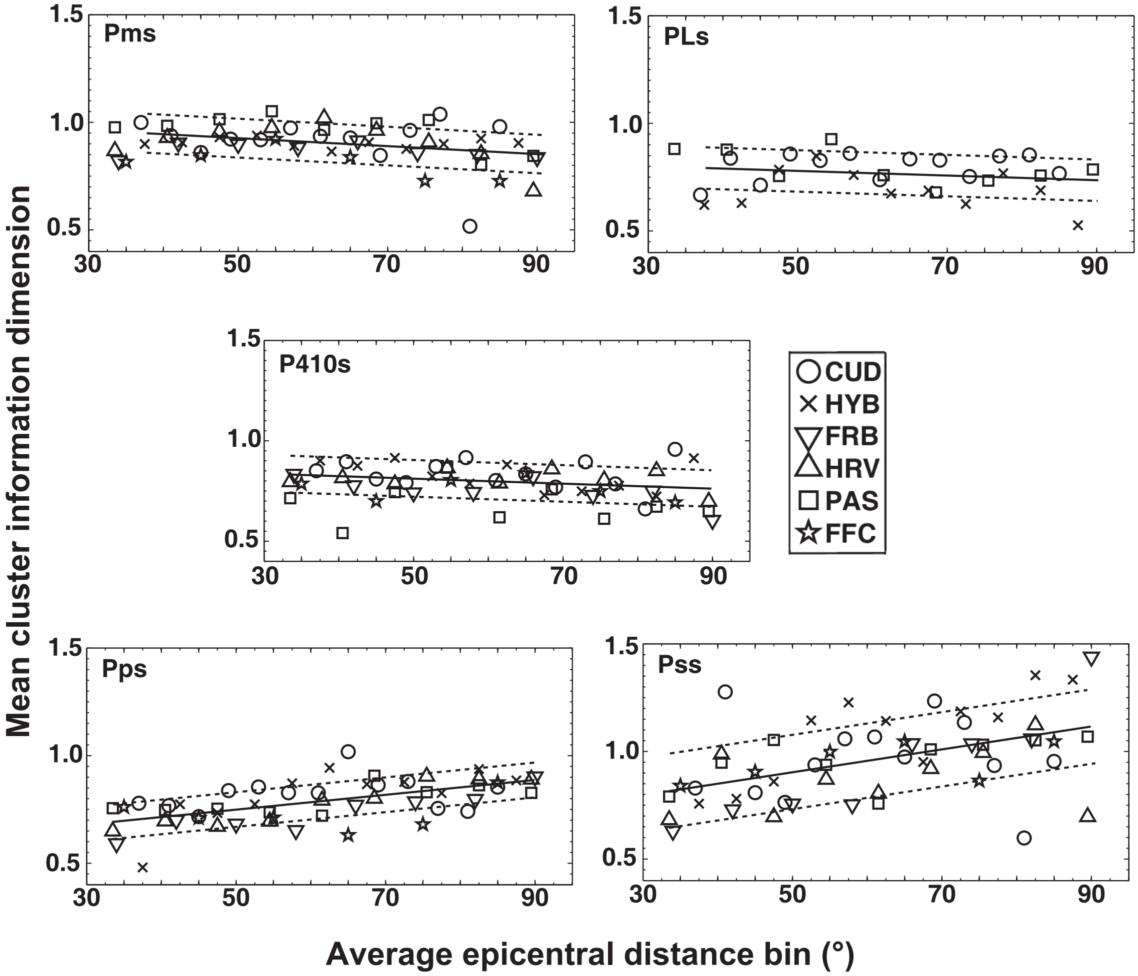


Figure 9. Same as in Figure 7, showing D_{Cluster} results from all the six stations. Note that only the corresponding phases recorded at these stations are shown. Solid line is the simple least square fit line for the entire data set, while the dashed lines are the error bounds of the obtained least square fit. The remarkable consistency in the respective slopes of various seismic phases revealed in Figure 7 is preserved at all the stations in spite of obvious differences in their tectonic settings and provenance antiquity. Thus, primary and multiple phases may be discriminated by their contrasting slopes in the mean cluster information dimension plots as a function of epicentral distance.

conversions and multiples upon application of Shannon's theory to the data. These are (1) dichotomy in their D_{Cluster} range of uncertainties, (2) distinct D_{Cluster} distribution character of corresponding *pdf*s, which are more symmetric for conversions and skewed for reverberations (Fig. 5B), and, finally, (3) consistently diverse mean D_{Cluster} best-fit slopes (Figs. 6, 7, and 9) of the primary converted waves and the multiply reflected wave types registered globally. We therefore successfully address the relatively long-standing issue of existence and unambiguous detection of shallow mantle layering (~150–300 km depth) from Ps receiver function data by applying concepts based on information theory. The reliability of this new approach is demonstrated by the cluster entropy-delineated PLs phase as confirmed by the Sp receiver function results of Ramesh et al. (2010) at stations HYB and CUD. The confirmed presence of velocity layering beneath the Indian stations, demonstrated by us in this study, at depths akin to the Lehmann boundary, along with the results presented in Ramesh et al. (2010), has at least one important implication. When viewed together with the latest reports on the nature of lithospheric layering in the North American craton, and its attendant implications (Yuan and Romanowicz, 2010), our findings clearly suggest that the depth to the lithosphere-asthenosphere boundary beneath the cratonic regions of SE India definitely exceeds 200 km.

Our new approach may provide a platform to test the validity of the existing phase transition hypotheses on the origin of shallow mantle layers by way of comparison with cluster entropy associated with candidate mineral phase reactions from mineral physics experiments. This has wide implications for geoscientists.

Scope for Future Work

In the near future, as a follow-up to the current study, we wish to attempt the following:

(1) Based on the pierce-point distribution of the receiver functions, the receiver functions would be segregated according to their sampling variable geology in a region. This could enable us test the degree of correlation (e.g., Arianos and Carbone, 2009) of the signals and their origin, leading to better geologic/geophysical explanation of the results. For example, it would be interesting to test our method on data across some of the confirmed suture zones where two distinct geologic provinces are now in welded contact.

(2) Similarly, a correlation analysis of earthquakes recorded by several stations sampling the same geology but originating from different tectonic environments (e.g., convergence, rift-related, transform faulting, and plate interior regions) will be taken up. Such a study of events from diverse tectonic settings will enable us to characterize the similarities and dissimilarities between various source regions on the globe based on cluster entropy.

ACKNOWLEDGMENTS

We are grateful to Pravin K. Gupta (Indian Institute of Technology, Roorkee, India) and Vinod K. Gaur (Indian Institute of Astrophysics, Bangalore, India) for clearing up concepts related to information theory and for their critical review of the manuscript. Rainer Kind (GeoForschungsZentrum, Potsdam, Germany) is thanked for critical reading of the manuscript. Suggestions by all above have improved the quality of this presentation. Thanks are due to the two anonymous reviewers and senior editor of the journal, Raymond M. Russo, who provided clear perception for improvement of our original submission. Help from our colleagues, R.K. Chadha for Indian data and P. Mandal for help with with the synthetics, is acknowledged. This work received support from the Council of Scientific & Industrial Research (New Delhi) in the form of the National Geophysical Research Institute Supra Institutional Program (SIP) and in-house main laboratory project MLP-6509-28 (to Das Sharma).

REFERENCES CITED

Abt, D.L., Fischer, K.M., French, S.W., Ford, H.A., Yuan, H., and Romanowicz, B., 2010, North American lithospheric discontinuity structure imaged by Ps and Sp receiver functions: Journal of Geophysical Research, v. 115, p. B09301, doi: 10.1029/2009JB006914.

Aki, K., and Richards, P.G., 1980, Quantitative Seismology Theory and Methods: Volume 1: New York, W.H. Freeman and Company, 557 p.

Arianos, S., and Carbone, A., 2009, Cross-correlation of long-range correlated series: Journal of Statistical Mechanics, doi: 10.1088/1742-5468/2009/03/P03037.

Balasis, G., Daglis, I.A., Papadimitriou, C., Kalimeri, M., Anastasiadis, A., and Eftaxias, K., 2009, Investigating dynamical complexity in the magnetosphere using various entropy measures: Journal of Geophysical Research, v. 114, p. A00D06, doi: 10.1029/2008JA014035.

Carbone, A., 2007, Algorithm to estimate the Hurst exponent of high-dimensional fractals: Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, v. 76, p. 056703, doi: 10.1103/PhysRevE.76.056703.

Carbone, A., and Stanley, H.E., 2007, Scaling properties and entropy of long-range correlated time series: Physica A, v. 384, p. 21–24, doi: 10.1016/j.physa.2007.04.105.

Carbone, A., Castelli, G., and Stanley, H.E., 2004, Analysis of clusters formed by the moving average of a long-range correlated time series: Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, v. 69, p. 026105, doi: 10.1103/PhysRevE.69.026106.

Deuss, A., and Woodhouse, J.H., 2004, The nature of the Lehmann discontinuity from its seismological Clapeyron slopes: Earth and Planetary Science Letters, v. 225, p. 295–304, doi: 10.1016/j.epsl.2004.06.021.

Farra, V., and Vinnik, L.P., 2000, Upper mantle stratification by P and S receiver functions: Geophysical Journal International, v. 141, p. 699–712, doi: 10.1046/j.1365-246x.2000.00118.x.

Ganguly, J., and Frost, D.J., 2006, Stability of anhydrous phase B: Experimental studies and implications for phase relations in subducting slab and the X discontinuity in the mantle: Journal of Geophysical Research, v. 111, p. B06203, doi: 10.1029/2005JB003910.

Grunewald, S., Weber, M., and Kind, R., 2001, The upper mantle under Central Europe: Indications for the Eifel plume: Geophysical Journal International, v. 147, p. 590–601, doi: 10.1046/j.1365-246x.2001.01553.x.

Gung, Y., Panning, M., and Romanowicz, B., 2003, Global anisotropy and the thickness of continents: Nature, v. 422, p. 707–711, doi: 10.1038/nature01559.

Haskell, N.A., 1962, Crustal reflection of plane P and SV waves: Journal of Geophysical Research, v. 67, p. 4751–4768, doi: 10.1029/JZ067i012p04751.

Landes, M., Ritter, J.R.R., O'Reilly, B.M., Readman, P.W., and Do, V.C., 2006, A N-S receiver function profile across the Variscides and Caledonides in SW Ireland: Geophysical Journal International, v. 166, p. 814–824, doi: 10.1111/j.1365-246X.2006.03052.x.

Langston, C.A., 1977, The effect of planar dipping structure on source and receiver responses for constant ray parameter: Bulletin of the Seismological Society of America, v. 67, p. 1029–1050.

Langston, C.A., 1979, Structure under Mount Rainier, Washington, inferred from teleseismic body waves: Journal of Geophysical Research, v. 84, p. 4749–4762, doi: 10.1029/JB084iB09p04749.

Lehmann, I., 1959, Velocities of longitudinal waves in the upper part of the Earth's mantle: Annales de Geophysique, v. 15, p. 93–118.

Li, A., Fischer, K.M., van der Lee, S., and Wysession, M.E., 2002, Crust and upper mantle discontinuity structure beneath eastern North America: Journal of Geophysical Research, v. 107, doi: 10.1029/2001JB000190.

Owens, T.J., Nyblade, A.A., Gurrola, H., and Langston, C.A., 2000, Mantle transition zone structure beneath Tanzania, East Africa: Geophysical Research Letters, v. 27, p. 827–830, doi: 10.1029/1999GL005429.

Ramesh, D.S., Kind, R., and Yuan, X., 2002, Receiver function analysis of the North American crust and upper mantle: Geophysical Journal International, v. 150, p. 91–108, doi: 10.1046/j.1365-246X.2002.01697.x.

Ramesh, D.S., Kumar, M.R., Uma Devi, E., Solomon Raju, P., and Yuan, X., 2005, Moho geometry and upper mantle images of northeast India: Geophysical Research Letters, v. 32, p. L14301, doi: 10.1029/2005GL022789.

Ramesh, D.S., Bianchi, M., and Das Sharma, S., 2010, Images of possible fossil collision structures beneath the Eastern Ghats Belt, India, from P and S receiver functions: Lithosphere, v. 2, p. 84–92, doi: 10.1130/L70.1.

Revenaugh, J., and Jordan, T.H., 1991, Mantle layering from ScS reverberations: 3. The upper mantle: Journal of Geophysical Research, v. 96, p. 19,781–19,810, doi: 10.1029/91JB01487.

Shannon, C.E., 1948, A mathematical theory of communication: Bell Systems Technical Journal, v. 27, p. 379–423, p. 623–656.

Sheehan, A.F., Shearer, P.M., Gilbert, H.J., and Dueker, K.G., 2000, Seismic migration processing of P-SV converted phases for mantle discontinuity structure beneath the Snake River Plain, western United States: Journal of Geophysical Research, v. 105, p. 19,055–19,065, doi: 10.1029/2000JB900112.

Vinnik, L.P., 1977, Detection of waves converted from P to SV in the mantle: Physics of the Earth and Planetary Interiors, v. 15, p. 39–45, doi: 10.1016/0031-9201(77)90008-5.

Wilson, D., Aster, R., Ni, J., Grand, S., West, M., Gao, W., Baldridge, W.S., and Semken, S., 2005, Imaging the seismic structure of the crust and upper mantle beneath the Great Plains, Rio Grande Rift, and Colorado Plateau using receiver functions: Journal of Geophysical Research, v. 110, p. B05306, doi: 10.1029/2004JB003492.

Young, G.B., and Braile, L.W., 1976, A computer program for the application of Zoeppritz's amplitude equations and Knott's energy equations: Bulletin of the Seismological Society of America, v. 66, p. 1881–1885.

Yuan, H., and Romanowicz, B., 2010, Lithospheric layering in the North American craton: Nature, v. 466, p. 1063–1068, doi: 10.1038/nature09332.

Yuan, X., Ni, J., Kind, R., Mechie, J., and Sandvol, E., 1997, Lithospheric and upper mantle structure of southern Tibet from a seismological passive source experiment: Journal of Geophysical Research, v. 102, p. 27,491–27,500, doi: 10.1029/97JB02379.

MANUSCRIPT RECEIVED 23 JUNE 2010
REVISED MANUSCRIPT RECEIVED 25 AUGUST 2010
MANUSCRIPT ACCEPTED 9 SEPTEMBER 2010

Printed in the USA

Bulletin of the Seismological Society of America, Vol. 102, No. 6, pp. 2563–2573, December 2012, doi: 10.1785/0120120036

From Induced Seismicity to Direct Time-Dependent Seismic Hazard

by Vincenzo Convertito, Nils Maercklin,* Nitin Sharma, and Aldo Zollo

Abstract The growing installation of industrial facilities for subsurface exploration worldwide requires continuous refinements in understanding both the mechanisms by which seismicity is induced by field operations and the related seismic hazard. Particularly in proximity of densely populated areas, induced low-to-moderate magnitude seismicity characterized by high-frequency content can be clearly felt by the surrounding inhabitants and, in some cases, may produce damage. In this respect we propose a technique for time-dependent probabilistic seismic-hazard analysis to be used in geothermal fields as a monitoring tool for the effects of on-going field operations. The technique integrates the observed features of the seismicity induced by fluid injection and extraction with a local ground-motion prediction equation. The result of the analysis is the time-evolving probability of exceedance of peak ground acceleration (PGA), which can be compared with selected critical values to manage field operations.

To evaluate the reliability of the proposed technique, we applied it to data collected in The Geysers geothermal field in northern California between 1 September 2007 and 15 November 2010. We show that the period considered the seismic hazard at The Geysers was variable in time and space, which is a consequence of the field operations and the variation of both seismicity rate and *b*-value. We conclude that, for the exposure period taken into account (i.e., two months), as a conservative limit, PGA values corresponding to the lowest probability of exceedance (e.g., 30%) must not be exceeded to ensure safe field operations. We suggest testing the proposed technique at other geothermal areas or in regions where seismicity is induced, for example, by hydrocarbon exploitation or carbon dioxide storage.

Introduction

Subsurface exploration aimed at producing energy exploiting the internal heat of the Earth is attracting large attention in many countries. Although new energies are beneficial, field operations such as high-pressure fluid pumping and hydraulic stimulation of reservoirs in geothermal fields induce seismicity. Thus, it is mandatory both to gain understanding of the mechanisms by which earthquakes are induced, and to study the effects of the induced seismicity in terms of seismic hazard. These are the main factors for mitigating seismic risk for the exposed community (Giardini, 2009).

In the framework of Enhanced Geothermal Systems (EGSs) and induced-seismicity analysis, the present paper focuses on The Geysers geothermal field, which is a vapor-dominated geothermal field located in northern California. The main steam reservoir has a temperature of about 235° C and underlies an impermeable caprock with its base 1.1–3.3 km below the surface. Commercial exploitation of the field began in 1960, and seismicity became more frequent in the area and increased with increasing field development (e.g., Eberhart-Phillips and Oppenheimer, 1984; Majer *et al.*, 2007). The observed induced seismicity is concentrated within the upper 6 km of the crust, in the reservoir below producing wells and near injection wells (Eberhart-Phillips and Oppenheimer, 1984). Induced seismicity has been monitored since the mid-1970s, and its temporal and spatial distribution has been analyzed to understand the causative mechanisms (e.g., Allis, 1982; Eberhart-Phillips and Oppenheimer, 1984; Oppenheimer, 1986). Seven earthquakes with $M \geq 4$ have been recorded in the period from April 2007 through October 2010.

Probabilistic seismic-hazard analysis (PSHA) in a time-dependent approach can be used as a tool to monitor the ongoing effects of the induced seismicity and can help guide the field operations for minimizing seismic risk. In this respect, in past years several studies have been performed which used either standard approaches with slight modifications to hazard analysis or proposed new techniques. For example, Van Eck *et al.* (2006) applied the standard Cornell (1968) approach to study hazard in the Netherlands related to

---

*Also at Dipartimento di Scienze Fisiche, Università degli studi di Napoli "Federico II", Napoli, Complesso Universitario di Monte S. Angelo, Via Cintia, 80126 Napoli, Italy.

seismicity induced by exploitation of a gas field, considering both peak ground acceleration (PGA) and peak ground velocity (PGV) as ground-motion parameters. As a modification to the standard approach, in order to account for the duration of gas production in a given field, they considered shorter return periods compared with the standard 475 years.

A new approach to tackle the problem of controlling the threat from induced seismicity has been proposed by Bommer *et al.* (2006). The authors correlated thresholds of tolerable ground motion with PGV and proposed a "traffic light" system whose response depends on the frequency of occurrence of the induced earthquakes. Based on the analysis of real-time recorded seismicity levels in terms of frequency of earthquake occurrence and the recorded PGV values, the system can issue three different colors: green, amber, and red. These colors alert the operators about the level of induced seismicity in terms of frequency of occurrence and ground-motion values, and can help to decide whether field operations can continue (green), must be adjusted (amber), or must be stopped (red).

Analyzing data from the Basel 2006 earthquake sequence, Bachmann *et al.* (2011) introduced a probability-based monitoring approach as an alternative to the "traffic light" system. The authors assumed that seismic sequences triggered by fluid injection can be treated as a sequence of clustered earthquakes. As a consequence, they modeled them using the Reasenberg and Jones (1989) model and the Epidemic-Type Aftershock Sequence model (Ogata, 1988). The models were first tested and their forecasting performance evaluated in a retrospective approach. The forecast was then translated into seismic hazard in terms of probabilities of exceedance of ground-motion intensity level.

In the present work, a time-dependent PSHA was performed by using a modified version of the approach proposed by Convertito and Zollo (2011). The technique is aimed at integrating the observations of the seismicity parameters and the ground motion prediction equations (GMPEs) to estimate the probability of exceedance of PGA that represents one of the inputs for risk analysis. The other inputs are the vulnerability that provides the probability of exceedance of a given damage level, and the exposure representing a qualitative and quantitative estimate of the elements exposed to the risk. Specifically, data were collected during different time windows of the field operations (T_{obs}), and PGA values having selected return periods were estimated by solving the hazard integral (Cornell, 1968). In order to test whether the proposed approach is able to predict what should be observed if the field operations could remain on the same level as that in the current T_{obs}, the estimated PGA values having selected probabilities of exceedance were compared with the observed values. As noted by Van Eck *et al.* (2006), the exposure periods and the return periods to be considered must be different from those used in standard PSHA because of the duration of the catalog in general and of the duration of EGS's operations.

We implemented the procedure proposed in this paper using data recorded by the Lawrence Berkeley National Laboratory Geysers/Calpine (LBNL) seismic network and the Northern California Seismic Network. The whole dataset is contained in the Northern California Earthquake Data Center (NCEDC) catalog which covers the period 2 October 1969 to present, although data from before 1975 may not be reliable. However, here we considered only the period 1 September 2007 through 15 November 2010, due to the aim of the present analysis, and because of the limited availability of the waveforms before that period. Moreover, in order to be confident of accounting only for induced seismicity, we selected only earthquakes having a depth of less than 4 km.

Methodology Overview

For any selected site, PSHA furnishes a hazard curve that represents the probability of exceedance of the ground-motion parameter, A, in a given time interval; as for example, during the design life of a building, a bridge, or other infrastructure. The computation of the hazard curve requires evaluation of the hazard integral (Cornell, 1968; Bazzurro and Cornell, 1999), which provides the mean annual rate of exceedance for the ith selected earthquake source zone and for a range of possible magnitudes and distances. The hazard integral is given by the equation:

$$E_i(A > A_o) = \alpha_i \int_M \int_R \int_\varepsilon I[A > A_o | m, r, \epsilon] \times f(m) f(r) f(\epsilon) dm dr d\epsilon, \qquad (1)$$

where I is an indicator function that equals one, if A is larger than A_o for a given distance r, a given magnitude m, and a given ε. Both distance and magnitude range between a minimum and a maximum value of interest, that is $R_{\mathrm{min}} \le r \le R_{\mathrm{max}}$, and $M_{\mathrm{min}} \le m \le M_{\mathrm{max}}$. The ε variable represents the residual variability of the A parameter with respect to the median value predicted by the selected GMPE (e.g., Bazzurro and Cornell, 1999; Convertito *et al.*, 2009). The GMPE, in its standard formulation, expresses the variation of the strong ground-motion parameter as a function of the source-to-site distance and the magnitude of an earthquake. The probability density function (PDF) for m, $f(m)$, depends on the adopted earthquake-recurrence model, which can be, for example, the standard Gutenberg and Richter (1944) (GR). The PDF for the distance r, $f(r)$, depends on the site location and source geometry. Given the magnitude and the distance, ε is defined as the number of logarithmic standard deviations by which the logarithm of the ground motion deviates from the median. Thus, via its associated PDF, $f(\varepsilon)$, ε accounts for the residual variability of the ground-motion parameter for which the hazard is estimated (e.g., Bazzurro and Cornell, 1999). Finally, α_i represents the mean annual rate of occurrence of the earthquakes within each identified source zone and is estimated from the seismic catalogs. In the classic approach to PSHA, a homogeneous Poissonian recurrence model is assumed where α_i is constant in time. Assuming that the event

$A > A_o$ is a selective process, for a given site, equation (1) allows the computation of the probability of exceedance P in a time interval t as

$$P(A > A_0, t) = 1 - e^{-\sum_{i=1}^{N} E_i(A>A_0)}, \qquad (2)$$

where the sum is taken over all of the sources that contribute to the hazard. Doing the analysis for a set of sites in an area of interest, and setting the exposure time (~50 years for civil structures) and the probability of exceedance, a hazard map can be obtained (Reiter, 1990).

The application to induced seismicity requires some modifications to the classical approach to PSHA. First, because of the variations of injection and production rate, the earthquake occurrence may not be stationary over small time windows. As a consequence, it is required that the α parameter, as well as the b-value of the GR relationship, varies with time, and the hazard integral is modified as:

$$E_i(A > A_o) = \int_T \int_M \int_R \int_\varepsilon \alpha_i(t) I[A > A_o|m, r, \varepsilon] \times f[m, b_i(t)] f(r) f(\varepsilon) dt dm dr d\varepsilon, \qquad (3)$$

where t ranges between $(T, T + \Delta t)$, which corresponds to the time window of interest. Second, because of the limited dimension of the seismogenic volume an upper-bound maximum magnitude $M_{\max}$ must be selected. Thus, an upper and lower truncated formulation of the $f(m)$ is used instead of the classic unbounded one in which the b-value varies with time. Moreover, the assumption of $f(r)$ being a uniform distribution based on the extension of the geothermal field may be not strictly valid. However, while recognizing that more sophisticated approaches, such as that proposed by Lasocki (2005), do exist for application to The Geysers, because of the level of recorded seismicity and the presence of several injection wells, we assume that $f(r)$ can be reasonably assumed as uniform. A similar assumption has been utilized by Van Eck *et al.* (2006).

To compute the conditional probability in equation (1), we used the GMPE proposed by Convertito *et al.* (2011) specifically for the area of interest. The data used for regression correspond to PGA measured as the largest between the two horizontal components from 220 earthquakes recorded at 29 LBNL stations from September 2007 through November 2010. The analyzed magnitude range is $1.0 < M_w < 3.5$, whereas the selected maximum depth value is 6 km, which corresponds to hypocentral distances ranging between 0.5 and 20 km. These values provide also the range of validity of the GMPE. For consistency with the analysis presented in Predicted Versus Observed Ground-Motion Values for Monitoring Purposes, we have not included in the dataset the PGA values recorded during 2009 at stations COBB, ADSP, and GCVB (see Fig. 1) which have been selected for site-specific seismic-hazard analysis.

Based on a pre-processing of the waveforms we selected only the best quality data for developing the GMPE. Specifically, these are earthquakes with at least 20 P picks and only those waveforms with a signal-to-noise ratio larger than

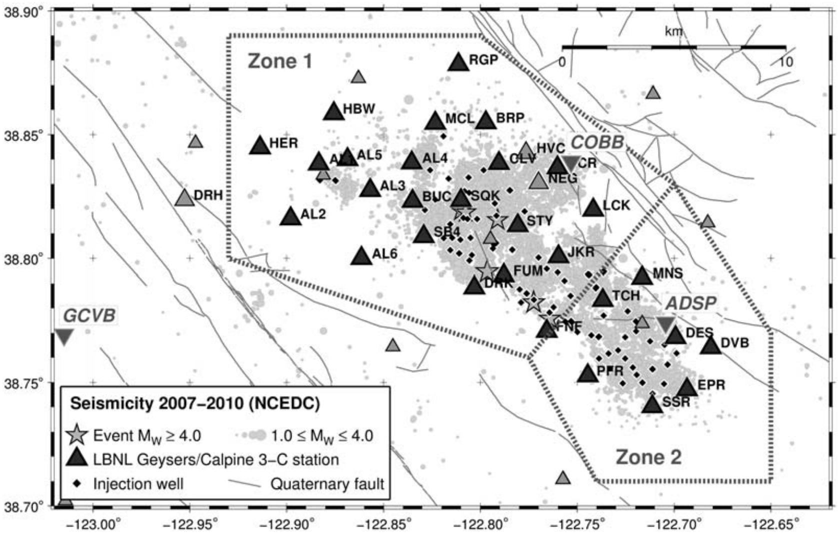


Figure 1. Map of induced seismicity recorded at The Geysers. Black triangles indicate the seismic stations of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network used in this study, and gray triangles are additional stations from the Northern California Seismic Network (NCSN) in the region. Dashed lines outline the two seismic-source zones used for the PSHA. The inverted triangles mark the three locations of site-specific PSHA, which are Anderson Springs (ADSP), Cloverdale (GCVB), and Cobb (COBB).

10. Moreover, to compute correct physical units, we applied the appropriate instrument-response correction to the waveforms within the frequency band ranging between 1 and 25 Hz. The selected formulation of the GMPE is defined as

$$\log \mathrm{PGA} = a + bM + c \log \sqrt{R^2 + h^2} + dR + es, \quad (4)$$

where PGA corresponds to peak ground acceleration in $\mathrm{m/sec}^2$. The distance metric R is the hypocentral distance in km and **M** is the moment magnitude. Additionally, the h parameter is introduced to avoid unrealistic high values at short distances (e.g., Joyner and Boore, 1981; Emolo *et al.*, 2011) and s accounts for site effect. The coefficients and their uncertainties are listed in Table 1. We obtained the best h value by minimizing the total standard error and maximizing the R^2 statistic that measures how well the regression curve approximates the real data points (Draper and Smith, 1996). The obtained fictitious depth $h = 3.5$ km yields the minimum standard error $\sigma = 0.324$ and a maximum R^2 of 0.852. The need of developing a specific GMPE arises from the fact that the probability is high that published GMPEs fail to predict peak-ground-motion values from shallow-focus, small-magnitude earthquakes, as noted by Bommer *et al.* (2006). Moreover, to date no other GMPEs published for The Geysers geothermal field exist. Further, the use of a specific GMPE allows for an accounting of the possible differences in static stress-drop conditions and anelastic attenuation between earthquakes induced within geothermal fields and natural earthquakes occurring elsewhere.

Seismic-Hazard Analysis

Seismic Source Identification

Based on the earthquakes' locations, magnitude distribution, and the locations of injection wells, we divided the region considered at The Geysers into two source zones named ZONE1 (Z1) and ZONE2 (Z2) that are outlined by the dashed lines shown in Figure 1. Our separation is supported by arguments by Stark (2003), Beall and Wright (2010), and Beall *et al.* (2010) for The Geysers field. A difference in the seismicity distribution has been also noted by Eberhart-Phillips and Oppenheimer (1984). All the authors noted that the southeastern part of The Geysers reservoir is less active seismically than the northwestern part where seismicity extends to greater depth. The difference was basically ascribed to a depth variation in the high-temperature (at 260–360° C) vapor-dominated reservoir. Specifically, Beall and Wright (2010) identified a net $M \geq 4$ dividing line which separates the whole area into two different seismic areas. The northwestern area contains all the earthquakes with magnitude larger than 4.0, whereas the southeastern one is characterized by lower magnitude earthquakes.

Moreover, the separation is justified from a statistical point of view by analyzing the b-values of the GR relationship for the two areas using the Utsu (1992) test. The Utsu test allows verification that data used for estimating the b-values come from different populations, and hence are associated with two different seismic sources. Using data for the entire duration of the analyzed catalog from 2007 through 2010, we obtained a probability of the order of 1.0E-6.

Maximum-Magnitude Estimation

In studies dealing with induced seismicity, one of the most debated issues concerns the selection of the maximum magnitude value $M_{\max}$ that could be induced by field operations. As an example, Shapiro *et al.* (2007) proposed a technique for estimating $M_{\max}$ from an analysis of injection duration, the strength of the injection source, and rock properties such as hydraulic diffusivity. Recently, Shapiro *et al.* (2010) also proposed the seismogenic index to quantify the seismotectonic state at an injection location. The seismogenic index depends only on the tectonic features and is independent of injection time or other injection characteristics, whose value correlates with the probability of a significant magnitude event. However, particularly for the period of interest analyzed in the present paper, detailed data about injection rate and extraction rate are not freely available. As a consequence, we follow Van Eck *et al.* (2006) and estimate $M_{\max}$ for each source zone by using the technique of Makropoulos and Burton (1983), although it is based on a stationary assumption. The technique assumes that the total energy that may be accumulated and released in a seismogenic volume is fairly constant in the considered time window. This hypothesis can be considered valid if the whole duration of the analyzed dataset is taken into account. Then, when the cumulative energy is plotted as function of time (Fig. 2), the distance between the two parallel lines enveloping the released energy correlates with the upper limit of the energy that would be observed in the region, if the accumulated energy during the time was released by a single earthquake. From the analyzed dataset we obtain $M_{\max}$ values of 4.5 for Z1 and 3.8 for Z2, respectively, as shown in Figure 2. These estimated values are coherent with the observations reported in Table 2 that confirm that the magnitude M 4.5 has never been exceeded in the analyzed period.

Table 1
Regression Coefficients and Relative Uncertainty of Equation (4).

a	b	c	d	h	e	σ
−2.268 ± 0.356	1.276 ± 0.026	−3.528 ± 0.624	0.053 ± 0.029	3.5	0.218 ± 0.014	0.324

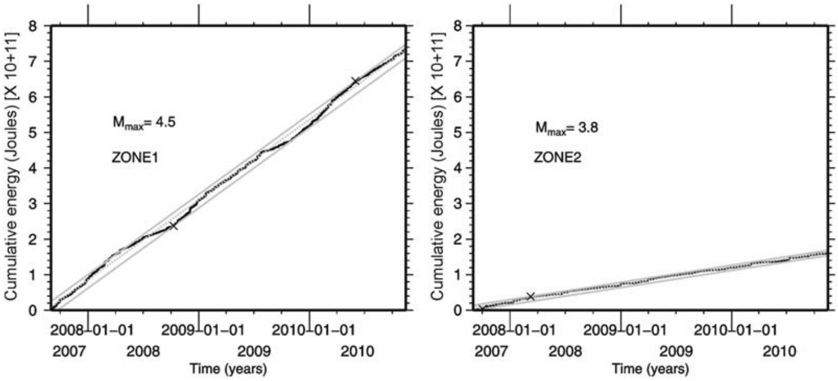


Figure 2. Results of the maximum-magnitude estimate obtained using the technique proposed by Makropoulos and Burton (1983). The left panel refers to zone Z1 and the right panel refers to zone Z2. In each panel the black line corresponds to the cumulative energy in the respective zone, the dashed-gray line corresponds to the average trend, and continuous-gray lines correspond to the upper and lower limit of the cumulative energy. The two crosses indicate the maximum and minimum cumulative energy, respectively, whose difference provides the estimated maximum magnitude.

Time-Dependent Seismicity Parameters Estimation

Once the seismic sources are identified and the expected $M_{\max}$ values are estimated, seismicity rate and the b-value need to be calculated as function of time for the time-dependent PSHA. To this aim, we first estimated the minimum magnitude of completeness M_c of the available catalog by using a technique similar to the Goodness-of-Fit technique proposed by Wiemer and Wyss (2000). Basically, the technique employs a comparison between the observed and the theoretical $f(m)$ PDF with the aim of minimizing the root mean square for a set of trial values for M_c, while the uncertainty is obtained by using a Monte Carlo approach. The results are shown in Figure 3b for Z1 and Figure 4b for Z2. Second, we estimated the b-values by using the Aki (1965) technique, while their uncertainties are obtained according to the formula of Shi and Bolt (1982). The obtained b-values are shown as gray dots in Figure 3c for Z1 and in Figure 4c for Z2. The observed different temporal behavior of the two analyzed parameters in Z1 and Z2 further supports the hypothesis of two different zones from the seismogenic point of view. The gray lines in Figures 3d and 4d indicate the weekly seismicity rates that provide a detailed picture of the activity rate. However, for the purposes of performing a stable PSHA, a monthly observation time is considered. A one month time window permits a statistically significant data sample for computing both the b-value and seismicity rate. The same time window has been selected by Eberhart-Phillips and Oppenheimer (1984) and Majer and Peterson (2007) as it enables capture of the main features of the temporal seismicity evolution, such as the difference between winter and summer months. Thus, assuming a non-homogenous Poisson model for earthquake occurrence, we integrated the seismicity rate $\alpha(t)$ in each time interval of one month length, and the resulting α values are plotted as black lines in Figure 3d for Z1 and Figure 4d for Z2. As a general consideration, it is evident that the monthly seismicity rate level in Z1 is on average three times that in Z2. Particularly for Z1, some interesting insights can be gained from the analysis of the plots. At a large timescale, three main peaks in the seismicity rate can be observed corresponding to March 2008, January 2009, and November 2009, while the rate is quite constant between January 2010 and July 2010. In each year, a seasonal variation of the seismicity rate can be noted,

Table 2

List of the Earthquakes with $M > 4.0$ within the Study Region*

Date (yyyy/mm/dd)	Time (hh:mm:ss)	Latitude (°)	Longitude (°)	Depth (km)	M
2007/04/24	21:08:29	38.795	−122.797	2.5	4.4
2007/12/01	20:50:12	38.815	−122.791	3.04	4.0
2008/02/24	05:32:10	38.819	−122.810	2.99	4.0
2008/05/30	04:48:30	38.779	−122.768	1.00	4.3
2008/05/30	04:48:36	38.776	−122.764	1.90	4.1
2009/01/04	17:27:10	38.782	−122.772	4.68	4.3
2010/07/15	15:31:43	38.819	−122.808	2.93	4.0

*As listed in the NCEDC earthquake catalog.

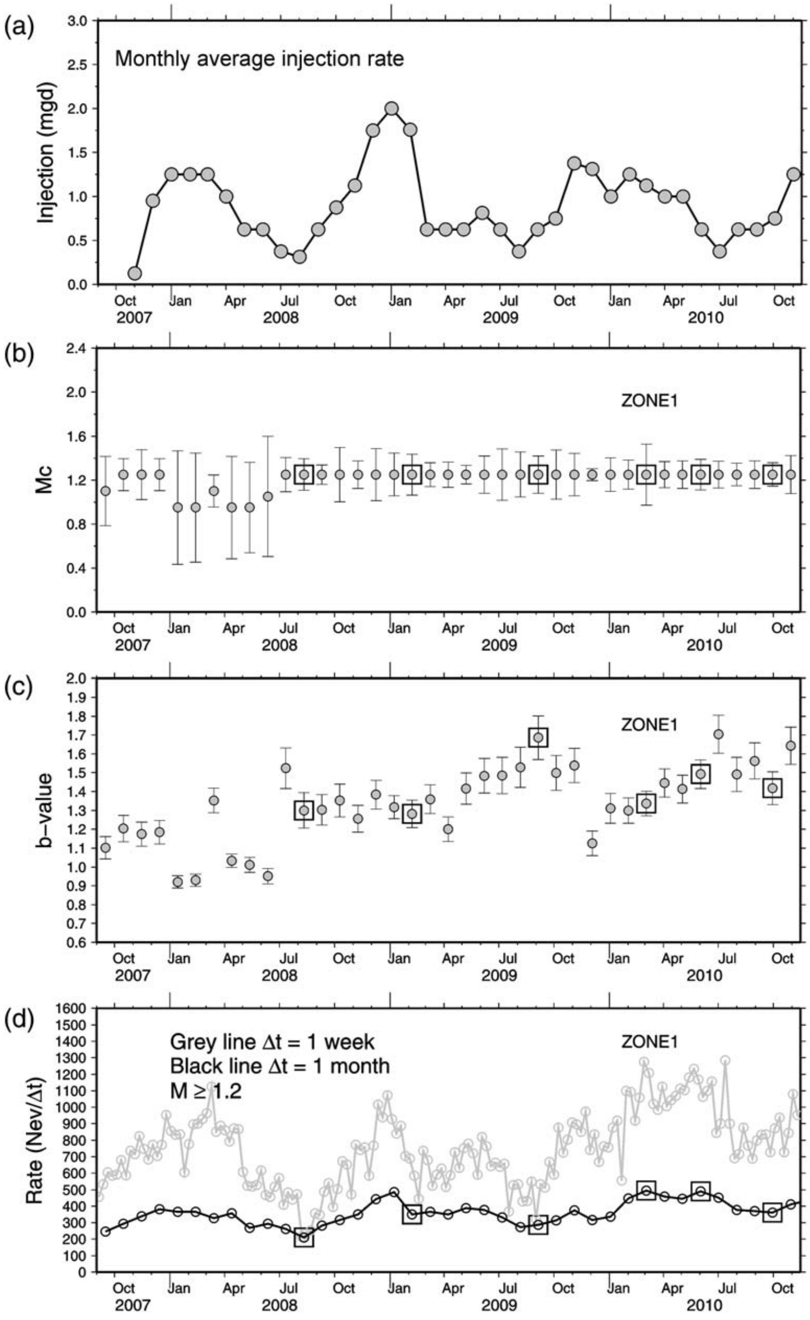


Figure 3. Analysis of induced seismicity recorded in the period 1 September 2007 through 15 November 2010 for Z1. Panel 3a shows the monthly average injection rate in millions of gallons per day (mgd) for the entire Geysers field. Panel 3b shows the minimum magnitude of completeness M_c as function of time, including uncertainties. Panel 3c shows the b-values and the uncertainties as functions of time using a one month time window. Panel 3d shows the weekly seismicity rate (gray lines) and the monthly seismicity rate (black lines). The squares identify the dates of the selected observation periods during which seismic-hazard analysis was performed.

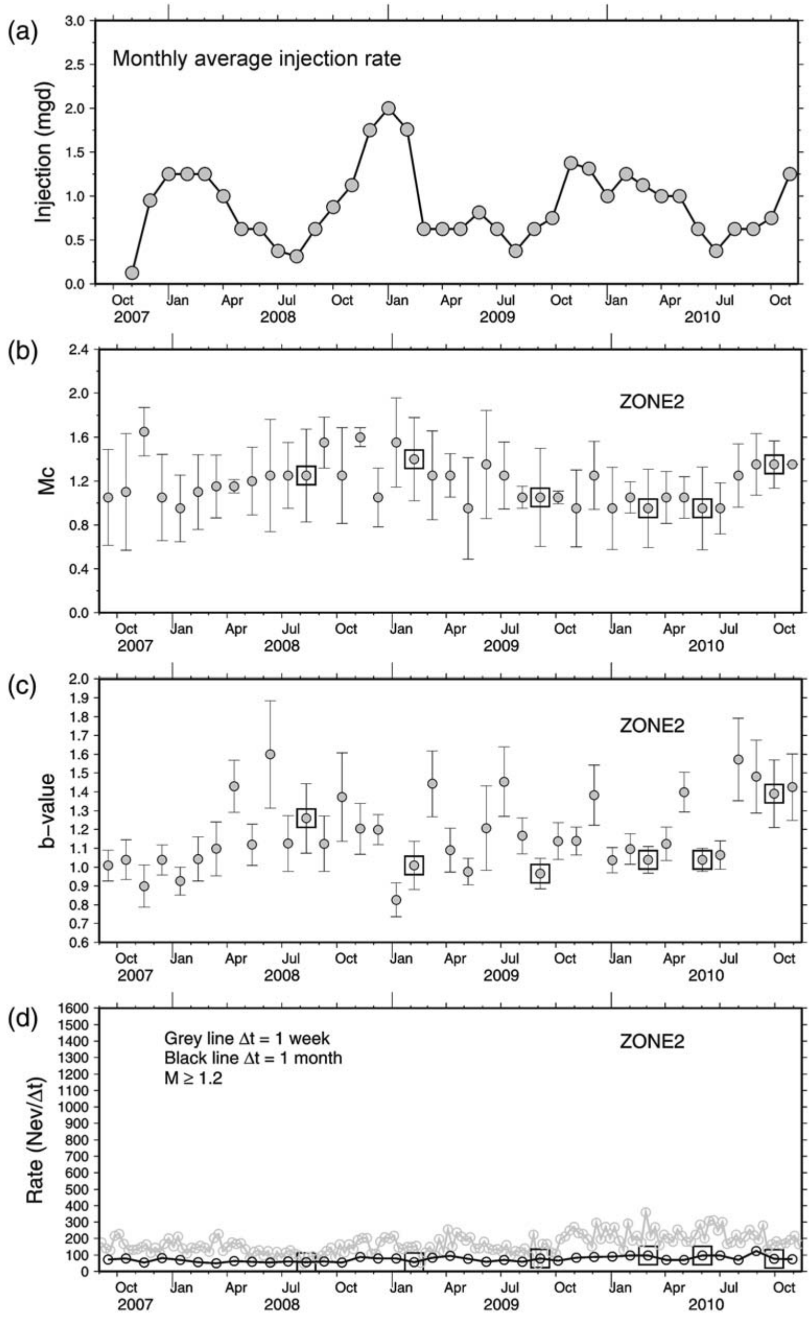


Figure 4. Analysis of induced seismicity recorded in the period 1 September 2007 through 15 November 2010 for Z2. Panel 4a shows the monthly average injection rate in millions of gallons per day (mgd) for the entire Geysers field. Panel 4b shows the minimum magnitude of completeness M_c as function of time, including uncertainties. Panel 4c shows the b-values and the uncertainties as functions of time using a one month time window. Panel 4d shows the weekly seismicity rate (gray lines) and the monthly seismicity rate (black lines). The squares identify the dates of the selected observation periods during which seismic-hazard analysis was performed.

which is related to changes in the amount of water injection throughout the year as shown in Figures 3a and 4a. A similar observation has been made by Majer and Peterson (2007) at The Geysers for the period from 2000 to mid-2006. However, these observations must be accompanied with the analysis of M_c, which shows several fluctuations, particularly for Z2. This means that the catalog has variable minimum-magnitude of completeness, which reflects several effects, such as seismic-network malfunctioning or improvements. Looking at the M_c value variation as function of time in our present application, we choose to use only events with magnitude larger than 1.2 for subsequent analyses.

As a further consideration, representing a key aspect for PSHA, the variations of the seismicity rate and M_c also lead to a variation of the b-values. Figures 3c and 4c compare the temporal variation of the b-values and their uncertainties with the seismicity rate for the two zones Z1 and Z2, respectively. While the estimated b-values for Z2 show a large scatter, distinct trends can be seen for Z1, corresponding to the following periods: September 2007 to January 2008, July 2008 to April 2009, May 2009 to November 2009, and January 2009 to June 2010. Because the M_c values are stable within these periods, these observed trends in the b-values can be considered as real features of the induced seismicity. Thus, it is expected that the probabilities of occurrence of larger magnitude earthquakes relative to lower magnitude events change in time and are different within the two zones. These temporal and spatial variations can affect the results of any PSHA and have to be included in such analyses.

Time-Dependent Hazard Analysis

Time-dependent PSHA proposed in the present paper utilizes a modified version of the technique proposed by Convertito and Zollo (2011) that was originally developed for analyzing so-called syn-crisis seismicity before an impending volcanic eruption. In the present application, we calculate the seismicity rate α and the b-values at several time intervals T_{obs} of one month, and select four probabilities of exceedance: 30%, 50%, 70%, and 90%. Taking an exposure time of two months, these probabilities correspond to return periods of about six months, three months, two months and one month, respectively. We chose the value of two months for the exposure time to account for at least two variations in the time-window period used to collect data.

We applied the GMPE reported in equation (4) to predict PGA, assuming that $s = 0$ that corresponds to a rock-site condition. Concerning the GR relationship and the related magnitude PDF, an upper bound on the magnitude values must be imposed. Here, we chose a time-varying truncated version of $f[m, b(t)]$, which is formulated as

$$f[m, b(t)] = \frac{\beta(t) e^{-\beta(t)(m - M_{\min})}}{1 - e^{-\beta(t)(M_{\max} - M_{\min})}}. \quad (5)$$

In equation (5), $\beta(t) = b(t) \ln 10$, $M_{\max}$ corresponds to the maximum magnitude value in each zone, that is, 4.5 for Z1 and 3.8 for Z2, and the minimum magnitude of interest $M_{\min}$ is set to 1.2 for both zones, which corresponds to the selected minimum magnitude of completeness.

To test the effect of the time variations of the input seismic parameters on the PSHA results, we computed hazard maps for different observation periods. The observation periods are 10 August 2008, 6 February 2009, 4 September 2009, 3 March 2010, 1 June 2010, and 29 September 2010 that are indicated by the squares in Figures 3b, c, d, and 4 for Z1 and Z2, respectively. Specifically, because we have used a monthly representation for both the b-value and the seismicity rate $\alpha(t)$, the indicated dates represent the central value of the corresponding one month time period. We selected these specific periods to be able to monitor the different features of seismicity rate and b-values and to evaluate their influence on PSHA. The corresponding hazard maps showing the PGA values having 30%, 50%, 70%, and 90% of exceedance probabilities are shown in Figure 5. These maps illustrate that the hazard is not constant with time. As an effect of the time variation of the seismic parameters, the largest PGA value changes both in time and space. For example, for the first two observation periods (10 August 2008 and 6 February 2009) the zone with higher hazard is Z1, where PGA values as large as 4.5 m/s^2 are estimated as values having 30% of probability of exceedance. On the other hand, largest PGA values are expected in Z2 during the later observation periods (3 March 2010 and 10 June 2010). Because Z1 is characterized by a larger $M_{\max}$ value than Z2, the larger PGA values in Z2 during the later periods can be mainly ascribed to the differences in the b-values.

Predicted Versus Observed Ground-Motion Values for Monitoring Purposes

As an additional analysis, we evaluated the use of PSHA as a monitoring tool to help to reduce seismic risk during the field operations. Benefiting from the availability of an independent dataset of PGA values for the analyzed period, we could follow a simple strategy to check the reliability of the PSHA results. We compared the estimated PGA values with those PGA values actually recorded in the area after the observation period considered for PSHA. In particular, we did a site-specific seismic-hazard analysis at three sites named ADSP, COBB, and GCVB (see Fig. 1), considering the same probabilities of exceedance as for the hazard maps computed in the Time-Dependent Hazard Analysis. For each site and each of the selected T_{obs}, we considered the successive three months and PGA values. In particular, for the two sites ADSP and COBB, data have been retrieved from the USGS (Data and Resources), which, for the time period of interest, contains PGA values for 2009. For the GCVB site, the NCEDC database (Data and Resources) has been accessed and PGA values were measured from the waveforms with the same procedure adopted for preparing the dataset

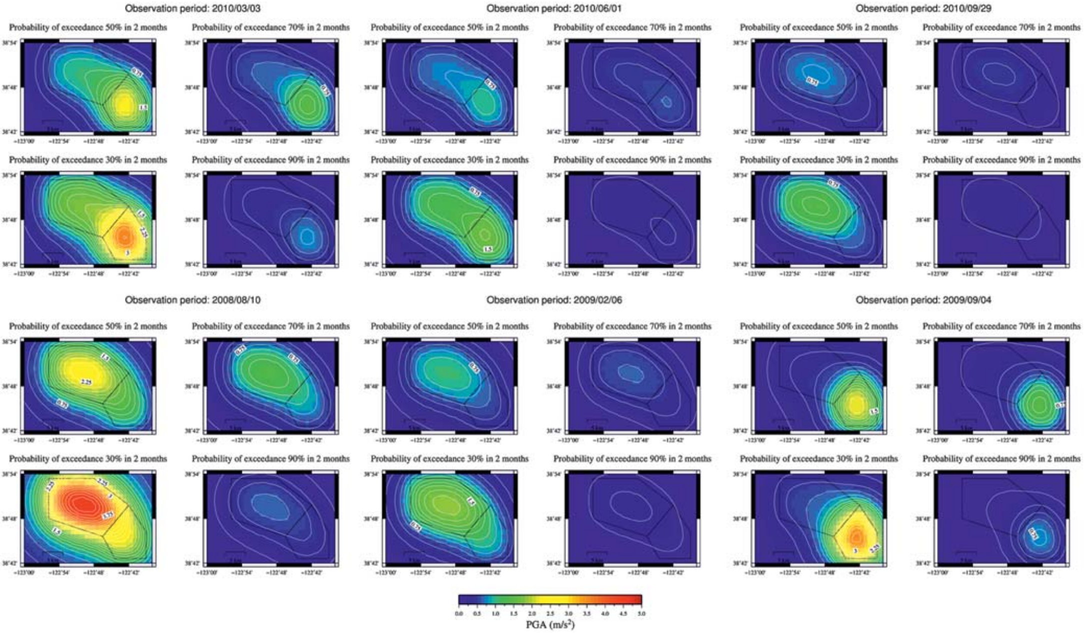


Figure 5. Seismic-hazard maps relative to the dates indicated by the squares in Figures 3 and 4. The reported PGA values are expressed in m/s^2 and represent the peak-ground-motion values having the probability of exceedance reported on the top of each map. Each date corresponds to the central time of a one month window centered on that date. The dashed lines outline the two seismic zones shown in Figure 1.

used for retrieving the GMPE. Thus, the selected observation periods start at March, June, and September 2009, respectively. Before discussing the results it is important to note that in the case of hazard associated with low-to-moderate magnitude earthquakes it might be more interesting to consider the medium-to-highest probabilities instead of considering the lower ones as in the case of the standard PSHA. This is because for monitoring purposes it could be more relevant to know what actual ground-motion level will be exceeded in the near future, rather than to assess the values associated with rare events, which contribute more to the lowest probability of exceedance.

The results of the site-specific PSHA in the monitoring context are shown in Figure 6. The estimated PGA values resulting from the hazard analysis together with their associated probability of exceedance are indicated by the dashed horizontal lines. The observed PGA values are indicated by the gray squares. Based on the recorded PGA values, which are lower than 1.2 m/s^2, all the sites experienced light-to-moderate shaking, and the values predicted by the hazard analysis are consistent with these observations. While the PGA values having 90% of probability of exceedance are systematically exceeded, the PGA values having a 30% of probability of exceedance could be used if a cautious value for a more conservative approach is needed.

Summary and Conclusions

In this paper we suggested and applied a procedure for time-dependent probabilistic seismic-hazard analysis in the context of induced seismicity in geothermal reservoirs. The main aim consists in verifying that hazard levels produced by induced seismicity stay below critical values during all phases of industrial operations. The analyzed dataset is made of waveforms and a parametric catalog of earthquakes recorded at The Geysers geothermal field in northern California between 1 September 2007 and 15 November 2010.

Here, we used a modified version of the approach proposed by Convertito and Zollo (2011), which is based on the following main points that could be used as guidelines for a real monitoring application:

- identify the seismic source zones at The Geysers, considering both well locations and recorded earthquakes (see Fig. 1);
- estimate the expected maximum magnitude in each zone, that is 4.5 in Z1 (northwest Geysers) and 3.8 in Z2;
- compute the minimum magnitude of completeness M_c of the network and monitor its time variation (M_c set to 1.2 for the entire analyzed period); and
- compute the b-values and seismicity rates as functions of time (Figs. 3 and 4);

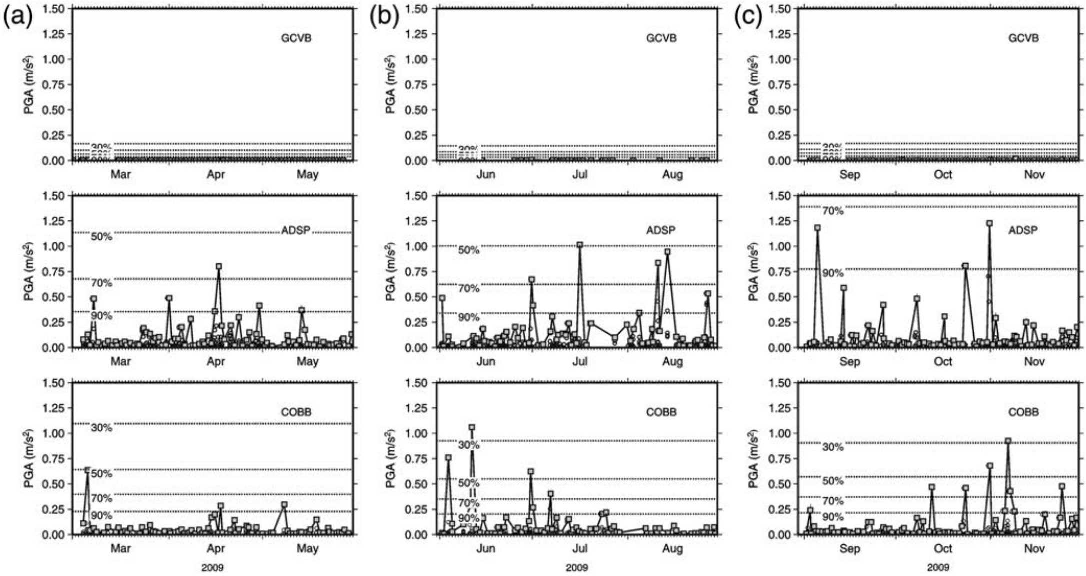


Figure 6. Site-specific seismic-hazard analysis for the three sites indicated in Figure 1. The analysis has been performed for three observation periods (a) starting 3 March 2009, (b) 1 June 2009, (c) and 29 September 2009, respectively, and whose duration is three months. The dashed horizontal lines correspond to the result of the PSHA at the indicated probability of exceedance. Gray squares represent the PGA values observed at the specific station.

- select proper critical ground-motion thresholds related to potential damage or perceived shaking; and
- recalibrate field operations if predicted ground-motion values exceed the thresholds.

From all these parameters and their time variation we obtained the time-dependent seismic hazard at four probabilities of exceedance (30%, 50%, 70%, and 90%), using PGA as the ground-motion parameter of interest. The main result is that, based on the current value of the input parameters, seismic hazard at The Geysers is not constant in time and space. However, with reference to the values used in ShakeMap® (Wald *et al.*, 1999), the average hazard level remains below potentially damaging values.

In addition, the availability of waveforms allowed us to investigate the possibility of using the time-dependent PSHA results for monitoring purposes. To this aim, we performed site-specific PSHA at a set of observation times that provided PGA values, having given probabilities of exceedance, to be used as predictions. We then compared these predicted PGA values with the PGA values observed in the period after the respective observation time. The obtained results show that PGA values corresponding to the lowest probability of exceedance (e.g., 30%) can be used as a conservative limit, in accordance with the fact that lower probabilities are associated more with rare events. On the other hand, the observations confirm that the predicted PGA values corresponding to higher probabilities of exceedance are actually exceeded.

A larger waveform database, together with the analysis of other ground-motion parameters, such as the PGV or the spectral ordinates at several structural periods, and possibly a longer study period seems to be necessary to justify and demonstrate the capabilities of time-dependent probabilistic seismic-hazard analysis as a monitoring tool during industrial operations. However, the results obtained in the present study are encouraging, and we suggest testing our hazard analysis and monitoring approach at other geothermal areas and, for example, also at gas field or carbon dioxide storage sites.

Data and Resources

Waveforms and parametric data have been retrieved from the Northern California Earthquake Data Center (NCEDC; network BG and station NC.GCVB; http://www.ncedc.org, last accessed January 2012). The PGA values used for the analysis, presented in Predicted Versus Observed Ground-Motion Values for Monitoring Purposes have been retrieved from USGS (ftp://ehzftp.wr.usgs.gov/luetgert/calpine/sm_sum.txt, last accessed April 2012). Figures 3a and 4a have been reproduced from an original figure by Calpine Corporation downloaded at http://www.geysers.com/docs/20110818_Hartline_NW_Geysers_EGS_FINAL_Template.pdf (last accessed January 2012). All figures have been generated with the Generic Mapping Tools (Wessel and Smith, 1991).

Acknowledgments

This work has been supported financially within the 7th Research Program of the European Union through the project Geothermal Engineering Integrating Mitigation of Induced Seismicity in Reservoirs (GEISER; ENERGY: 2009.2.4.1; Contract Number 241321). We wish to thank Associate Editor Arthur McGarr and two anonymous reviewers whose suggestions improved the quality of the manuscript.

References

Aki, K. (1965). Maximum likelihood estimate of b in the formula $\log N = a - bM$ and its confidence limits, *Bull. Earthquake Res. Inst./Tokyo Daigaku Jishin Kenkyusho Iho* **43,** 237–239.

Allis, R. G. (1982). Mechanisms of induced seismicity at The Geysers geothermal reservoir, California, *Geophys. Res. Lett.* **9,** 629–632.

Bachmann, C. E., S. Wiemer, J. Woessner, and S. Hainzl (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for Enhanced Geothermal Systems, *Geophys. J. Int.* **186,** no. 2, 793–807, doi: 10.1111/j.1365-246X.2011.05068.x.

Bazzurro, P., and C. A. Cornell (1999). Disaggregation of seismic hazard, *Bull. Seismol. Soc. Am.* **89,** 501–520.

Beall, J. J., and M. C. Wright (2010). Southern extent of The Geysers high temperature reservoir based on seismic and geochemical evidence, *Trans. Geotherm. Resour. Counc.* **34,** 53–56.

Beall, J. J., M. C. Wright, A. S. Pingol, and P. Atkinson (2010). Effect of high injection rate on seismicity in The Geysers, *Trans. Geotherm. Resour. Counc.* **34,** 47–52.

Bommer, J. J., S. Oates, J. M. Cepeda, C. Lindholm, J. Bird, R. Torres, G. Marroquin, and J. Rivas (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal project, *Eng. Geol.* **83,** 287–306, doi: 10.1016/j.enggeo.2005.11.002.

Convertito, V., and A. Zollo (2011). Assessment of pre-crisis and syn-crisis seismic hazard at Campi Flegrei and Mt. Vesuvius volcanoes, Campania, southern Italy, *Bull. Volcanol.* **73,** no. 6, 767–783, doi: 10.1007/s00445-011-0455-2.

Convertito, V., I. Iervolino, and A. Herrero (2009). Importance of mapping design earthquakes: Insights for the Southern Apennines, Italy, *Bull. Seismol. Soc. Am.* **99,** no. 5, 2979–2991, doi: 10.1785/0120080272.

Convertito, V., N. Sharma, N. Maercklin, A. Emolo, and A. Zollo (2011). Seismic hazard analysis as a controlling technique of induced seismicity in geothermal systems (Abstract S41C-2194), *Fall Meeting, American Geophysical Union*, San Francisco, California, 5–9 December S41C-2194.

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

Draper, N. R., and H. Smith (1996). *Applied Regression Analysis*, Third Ed., John Wiley and Sons, Inc., New York, 407 pp.

Eberhart-Phillips, D., and D. H. Oppenheimer (1984). Induced seismicity in The Geysers Geothermal Area, California, *J. Geophys. Res.* **89,** 1191–1207.

Emolo, A., V. Convertito, and L. Cantore (2011). Ground-motion predictive equations for low-magnitude earthquakes in the Campania–Lucania area, Southern Italy, *J. Geophys. Eng.* **8,** no. 1, 46–60, doi: 10.1088/1742-2132/8/1/007.

Giardini, D. (2009). Geothermal quake risks must be faced, *Nature* **462,** 848–849, doi: 10.1038/462848a.

Gutenberg, B., and C. R. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34,** 185–188.

Joyner, W. B., and D. M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, *Bull. Seismol. Soc. Am.* **71,** 2011–2038.

Lasocki, S. (2005). Probabilistic analysis of seismic hazard posed by mining induced events, in Y. Potvin and M. Hudyma (Editors), *Proc. 6th Int. Symp. On Rockburst in Mines "Controlling Seismic Risk", Australian Center for Geomechanics (ACG)*, Perth, 151–156.

Majer, E. L., and J. E. Peterson (2007). The impact of injection on seismicity at The Geysers, California Geothermal Field, *Int. J. Rock Mech. Min. Sci.* **44,** 1079–1090, doi: 10.1016/j.ijrmms.2007.07.023.

Majer, E. L., R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma (2007). Induced seismicity associated with enhanced geothermal systems, *Geothermics* **36,** 185–222.

Makropoulos, K. C., and W. Burton (1983). Seismic risk of circum-Pacific earthquakes I. Strain energy release, *Pure Appl. Geophys.* **121,** no. 2, 247–267, doi: 10.1007/BF02590137.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point-processes, *J. Am. Stat. Assoc.* **83,** 9–27.

Oppenheimer, D. C. (1986). Extensional tectonics at the Geysers Geothermal Area, California, *J. Geophys. Res.* **91,** 11,463–11,476.

Reasenberg, P. A., and L. M. Jones (1989). Earthquake hazard after a mainshock in California, *Science* **243,** no. 4895, 1173–1176, doi: 10.1126/science.243.4895.1173.

Reiter, L. (1990). *Earthquake Hazard Analysis*, Columbia University Press, New York, 254 pp.

Shapiro, S. A., C. Dinske, and J. Kummerow (2007). Probability of a given-magnitude earthquake induced by fluid injection, *Geophys. Res. Lett.* **34,** no. L22314, 5, doi: 10.1029/2007GL031615.

Shapiro, S. A., C. Dinske, and C. Langenbruch (2010). Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulation, *The Leading Edge* **29,** no. 3, 304–309, doi: 10.1190/1.3353727.

Shi, Y., and B. A. Bolt (1982). The standard error of the magnitude-frequency b value, *Bull. Seismol. Soc. Am.* **72,** 1677–1678.

Stark, M. (2003). Seismic evidence for a long-lived enhanced geothermal system (EGS) in the northern Geysers reservoir, *Trans. Geotherm. Resour. Counc.* **27,** 727–731.

Utsu, T. (1992). On seismicity, in *Report of cooperative research of the institute of statistical mathematics*, Vol. **34**, Mathematical Seismology VII, Annals of the Institute of Statistical Mathematics, 139–157.

Van Eck, T., F. Goutbeek, H. Haak, and B. Dost (2006). Seismic hazard due small-magnitude, shallow-source, induced earthquakes in The Netherlands, *Eng. Geol.* **87,** 105–121, doi: 10.1016/j.enggeo.2006.06.005.

Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet "ShakeMaps": Rapid generation of peak ground motion and intensity maps for earthquakes in Southern California, *Earthquake Spectra* **15,** 537–555, doi: 10.1193/1.1586057.

Wessel, P., and W. H. F. Smith (1991). Free software helps map and display data, *Eos Trans. AGU* **72,** 445–446.

Wiemer, S., and M. Wyss (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the Western United States, and Japan, *Bull. Seismol. Soc. Am.* **90,** 859–869.

Istituto Nazionale di Geofisica e Vulcanologia
Osservatorio Vesuviano
Via Diocleziano 328
80124 Napoli, Italy
(V.C.)

AMRA S.c.a.r.l.
Analysis and Monitoring of Environmental Risk
Via Nuova Agnano 11
80125 Napoli, Italy
(N.M.)

Dipartimento di Scienze Fisiche
Università degli Studi di Napoli "Federico II", Napoli
Complesso Universitario di Monte S. Angelo
Via Cintia
80126 Napoli, Italy
(N.S., A.Z.)

Manuscript received 1 February 2012

Bulletin of the Seismological Society of America, Vol. 103, No. 1, pp. 117–130, February 2013, doi: 10.1785/0120120138

Ground-Motion Prediction Equations for The Geysers Geothermal Area based on Induced Seismicity Records

by Nitin Sharma, Vincenzo Convertito, Nils Maercklin,* and Aldo Zollo

Abstract The Geysers geothermal field in Northern California, which has been actively exploited since the 1960s, is the world's largest geothermal field. The continuous injection of fluids and the consequent stress perturbations induce seismicity that is clearly felt in the surrounding communities. In order to evaluate seismic hazard due to induced seismicity and the effects of seismicity rate level on the population and buildings in the area, reliable ground-motion prediction equations (GMPEs) must be developed. This paper introduces the first GMPEs specific for The Geysers area in terms of peak ground velocity (PGV), peak ground acceleration (PGA), and 5% damped spectral acceleration SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s. The adopted non-linear mixed-effect regression technique to derive the GMPE includes both fixed and random effects, and it permits to account for both inter-event and intra-event dependencies in the data. Site-specific effects are also estimated from the data and are corrected in the final ground-motion model. We used data from earthquakes recorded at 29 stations of the Berkeley-Geysers network during the period September 2007 through November 2010. The magnitude range is $1.3 \leq M_w \leq 3.3$, whereas the hypocentral distances range between 0.5 km and 20 km. The comparison of our new GMPE for The Geysers with a standard model derived in a different tectonic context shows that our model is more robust when predictions have to be made for induced earthquakes in this geothermal area.

Introduction

Enhanced geothermal systems (EGSs) provide an increasingly important contribution to the world energy inventory. However, one of the most controversial issues associated with EGSs is induced seismicity, which has been the cause of delays and threatened cancellation of at least two EGS projects worldwide (Majer *et al.*, 2007). Although induced seismicity may have light-to-moderate adverse physical effects on operations or on the surrounding communities, public concern about the rate and magnitude of the seismicity associated with current and future EGS operations still remains.

The present paper is focused on The Geysers, a vapor-dominated geothermal field located about 120 km north of San Francisco, California. The field has been actively exploited since the 1960s, and it is now the most productive geothermal field in the world. Regular monitoring of induced seismicity in the field began already in the 1960s and has been continuously improved since then, resulting in one of the most comprehensive data sets available. Therefore, The Geysers is a well-suited test case for the study of induced seismicity and the development of related ground-motion prediction models. Many studies have demonstrated that microearthquakes at The Geysers are associated with water injection and steam extraction (Majer and McEvilly, 1979; Eberhart-Phillips and Oppenheimer, 1984; Oppenheimer, 1986; Enedy *et al.*, 1992; Stark, 1992, 2003; Foulger *et al.*, 1997; Ross *et al.*, 1999; Smith *et al.*, 2000; Majer and Peterson, 2005; Majer *et al.*, 2007).

Currently, there are no specific ground-motion prediction equations (GMPEs) available for The Geysers field, although they play a key role in seismic-hazard analysis. Time-dependent study of GMPEs and seismic hazard helps in monitoring the effects of the seismicity rate levels on inhabitants living in surrounding areas and on structures (Convertito *et al.*, 2012). The aim of the present study is the estimation of suitable GMPEs for The Geysers geothermal area. In particular, the GMPEs are retrieved for peak ground velocity (PGV), peak ground acceleration (PGA), and 5% damped spectral acceleration [SA(T)] at $T = 0.2$ s, 0.5 s, and 1.0 s, as these ground-motion parameters are readily available after each earthquake, can be used to estimate potential damage, and are well-correlated with levels of human sensitivity (Bommer *et al.*, 2006). This issue is particularly important in all the applications devoted to real-time

---

*Also at Dipartimento di Scienze Fisiche, Unità di Recerca in Sismologia Sperimentale e Computazionale, Università degli Studi di Napoli "Federico II", Complesso Universitario di Monte S. Angelo, via Cintia, 80126 Napoli, Italy.

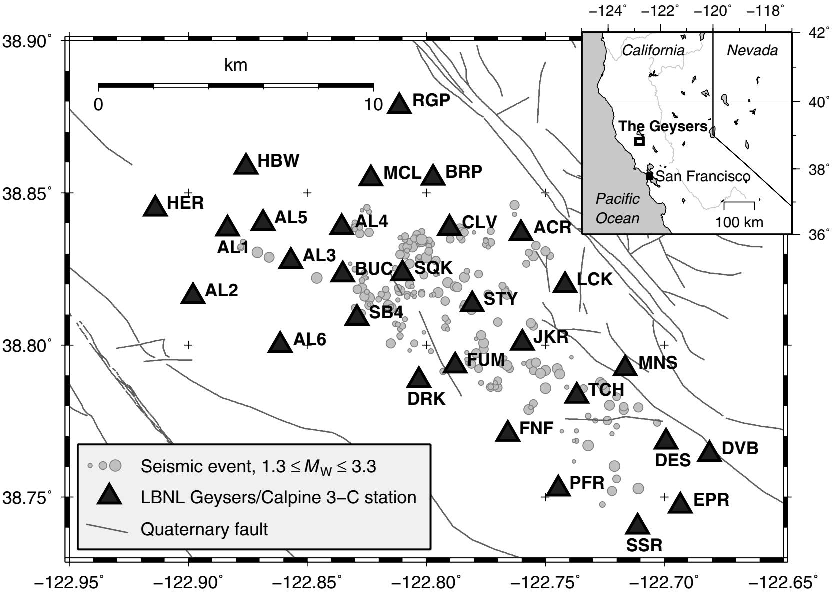


Figure 1. Location of the events and seismic stations used for regression analysis in the present study. Gray circles, event locations; black triangles, seismic stations; LBNL, Lawrence Berkeley National Laboratory.

monitoring and seismic risk reduction in geothermal areas (Bommer *et al.*, 2006; Bachmann *et al.*, 2011; Convertito and Zollo, 2011; Convertito *et al.*, 2012).

In order to account for site-specific effects, and for the model to be also applicable to the sites where there is lack of geologic information such as the time-averaged shear-wave velocity in the upper 30 m (V_{S30}), we have implemented a two-step regression analysis similar to that proposed by Emolo *et al.* (2011) in the application to natural microseismicity recorded in southern Italy. In particular, in the first step a model is retrieved, which does not explicitly include possible station effects and represents a reference model. In the second step, a first-order correction is introduced for station effects based on the analysis of the residuals distribution at each station with respect to the reference model. The residuals are defined here as the difference between the common logarithm of the observed peak ground-motion parameter value and common logarithm of the predicted one. The Z-test is used to test the null hypothesis of a zero-mean residuals distribution. Stations for which a statistically significant deviation from zero is observed are assumed to be affected by a station effect (Emolo *et al.*, 2011). Thus,

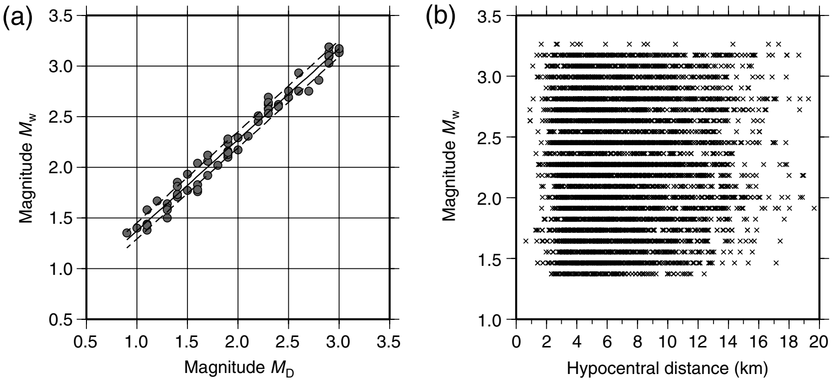


Figure 2. (a) Regression model for computing moment magnitude, M_w, from the duration magnitude M_D with standard deviation of 0.08; solid line, regression model; dashed lines, regression model with standard deviation; and (b) scatter plot showing hypocentral distance and magnitude range used for regression analysis.

based on the Z-test results, we assign a corrective coefficient to be applied at each station, which would allow reduction of the deviation of the residual distribution from the expected zero-mean value. Specifically, we assign a 0 coefficient at all the stations with zero-mean residual (predictions are on average equal to the observed), a +1 coefficient at stations having a positive deviation from zero (mean residuals value is greater than zero, that is, predictions underestimate the observations), and a −1 coefficient at stations having a negative deviation from zero (mean residuals value is less than zero, that is, predictions overestimate the observations). The addition of this new parameter in the reference model provides the corrected regression model.

In the present application, we tested different models, which included different dependencies on magnitude and distance, representing the predictor variables that account for source and propagation effects such as geometrical and anelastic attenuation (as the zone is highly fractured and fluid dominated). Finally, we chose the best model on the basis of the total standard deviation and R^2 statistic, which measure how well the regression curve approximates the real data points (Draper and Smith, 1996). To test the reliability of the obtained results, we have also compared our final model with the model proposed by Atkinson and Boore (2011; herein after, AB11), which is the successor to the Boore and Atkinson (2008) model. The comparisons are quite good, but some difference is observed for lower magnitude events and the comparison between both models improves with increase in magnitude.

Data Description

Different temporary and long-term seismic networks have been deployed in The Geysers geothermal field during the last five decades. At present, local seismicity is monitored by the dense Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine (BG) surface seismic network and with some nearby stations of the Northern California Seismic Network. The BG network consists of 32 three-component stations, 29 of which were used for the present study (black triangles in Fig. 1). The BG stations were distributed in an area of about 20 × 10 km^2, covering the entire geothermal field. Each station was equipped with I/O Sensor SM-6 geophones with a natural frequency of 14 Hz. In the fall of 2009 these instruments were replaced successively by Oyo GS-11D 4.5 Hz sensors. The BG stations operate in trigger mode, and the waveform segments recorded since the end of July 2007 are made available at the Northern California Earthquake Data Center (NCEDC). We retrieved waveforms for the period 1 September 2007 through 15 November 2010, associated the traces with the events from the NCEDC catalog, and prepared a subset of the data for regression analysis. The selected waveforms correspond to 212 earthquakes (dots in Fig. 1) with a focal depth of less than 5 km. This depth has been selected considering that the earthquakes observed at depths greater than 5 km are natural earthquakes. The hypocentral distance ranges from 0.5 km to 20 km, and the magnitude range is $1.3 \le M_\mathrm{w} \le 3.3$. For the largest portion of earthquakes analyzed in this study, the NCEDC catalog provides a duration magnitude M_D as magnitude measure. However, in order to obtain results compatible with other studies and suitable for seismic-hazard analysis purposes, we converted M_D into moment magnitudes M_w using a linear relationship retrieved by using M_w data provided by J. Douglas *et al.*, unpublished manuscript, 2012, mentioned in equation (1) as follows:

$$M_\mathrm{w} = 0.473(\pm 0.035) + 0.900(\pm 0.017) M_\mathrm{D}. \qquad (1)$$

The data used and the least-squares fitting with standard deviation (0.08) are shown in Figure 2a, and Figure 2b shows

Table 1

Regression Coefficients and Relative Uncertainty of Equation (2), Using MOD1

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$\sigma_{\text{Inter-event}}$	$\sigma_{\text{Intra-event}}$	σ_{Total}	R^2
PGV (m/s)	−4.960 ± 0.074	1.317 ± 0.022	−2.101 ± 0.055	1.850 ± 0.212	0.154	0.344	0.377	0.814
PGA (m/s^2)	−2.666 ± 0.072	1.158 ± 0.020	−2.312 ± 0.055	1.734 ± 0.197	0.142	0.358	0.385	0.786
SA($T = 0.2$ s) (m/s^2)	−3.648 ± 0.087	1.443 ± 0.022	−1.871 ± 0.068	2.393 ± 0.270	0.160	0.356	0.390	0.815
SA($T = 0.5$ s) (m/s^2)	−4.926 ± 0.081	1.552 ± 0.023	−1.688 ± 0.061	2.210 ± 0.275	0.164	0.339	0.377	0.840
SA($T = 1.0$ s) (m/s^2)	−5.341 ± 0.076	1.503 ± 0.022	−1.861 ± 0.056	1.959 ± 0.240	0.157	0.339	0.374	0.840

Table 2

Regression Coefficients and Relative Uncertainty of Equation (3), Using MOD2

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$d \pm \sigma_\mathrm{d}$	$\sigma_{\text{Inter-event}}$	$\sigma_{\text{Intra-event}}$	σ_{Total}	R^2
PGV (m/s)	−4.585 ± 0.255	1.317 ± 0.021	−2.730 ± 0.390	2.406 ± 0.358	0.027 ± 0.015	0.154	0.344	0.377	0.814
PGA (m/s^2)	−2.307 ± 0.242	1.158 ± 0.020	−2.925 ± 0.376	2.243 ± 0.332	0.026 ± 0.015	0.142	0.358	0.385	0.786
SA($T = 0.2$ s) (m/s^2)	−3.370 ± 0.327	1.442 ± 0.022	−2.309 ± 0.484	2.821 ± 0.492	0.017 ± 0.018	0.160	0.356	0.390	0.815
SA($T = 0.5$ s) (m/s^2)	−4.950 ± 0.222	1.552 ± 0.023	−1.648 ± 0.343	2.158 ± 0.545	−0.002 ± 0.014	0.164	0.339	0.377	0.840
SA($T = 1.0$ s) (m/s^2)	−5.158 ± 0.237	1.503 ± 0.022	−2.167 ± 0.364	2.285 ± 0.430	0.013 ± 0.014	0.157	0.339	0.374	0.840

the scatter plot of the hypocentral distance and magnitude used for the present study.

To analyze the best quality data, we selected only waveforms with signal-to-noise ratio greater than 10 in the whole analyzed frequency range, 0.5–35 Hz. We applied the instrument correction to the waveforms within the same frequency band. The mean and the trend were also removed. Then zero phase shift and a four-pole Butterworth filter in the frequency band 0.7–35 Hz was applied. Further, to measure the correct values of the selected ground-motion parameters we cut the waveforms in a specific time window around the event, starting at the origin time and ending at the time corresponding to 98% of total energy contained in waveform, which were also tapered with a 0.1 taper width with a cosine window.

Once the time window was selected, PGV was measured as the largest value among the two horizontal components. The waveforms were then differentiated and filtered in a frequency band ranging between 0.7 Hz and 35 Hz (to avoid the high frequency noise added while differentiating) to measure PGA and the 5% damped spectral ordinates SA(T) at $T = 0.2$ s, 0.5 s, and 1 s, which, as for PGV, correspond to the largest value among two horizontal components.

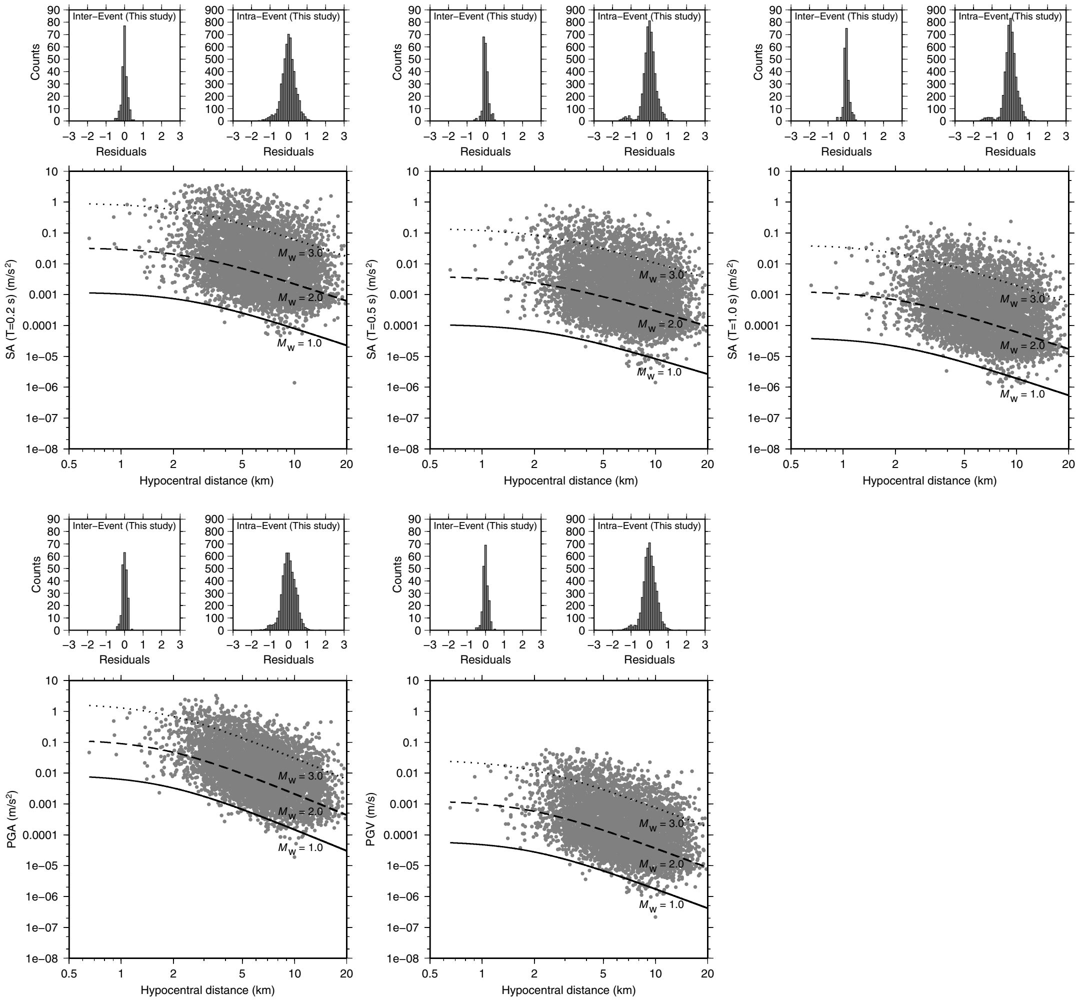


Figure 3. Regression model for magnitudes 1.0, 2.0, and 3.0 with respect to hypocentral distance; solid line, magnitude 1.0; dashed line, magnitude 2.0; dotted line, magnitude 3.0; gray dots, entire data set. Smaller pairs of panels show inter- and intra-event residual distributions for each of five ground-motion parameters (PGV, PGA, and SA at 0.2 s, 0.5 s, and 1.0 s) depicted in the larger panels.

Regression Analysis

In the present study, we have used non-linear mixed-effect regression analysis (NLMRA) to obtain the GMPEs. The technique is described in detail in papers by Lindstrom and Bates (1990) and Abrahamson and Youngs (1992), and has the advantage to consider inter-event and intra-event effects (e.g., Abrahamson and Silva, 1997). The predictor variables are moment magnitude M_w, and hypocentral distance, R_{hypo}, whereas the response variables correspond to PGV, PGA, and SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s. We have tested two models that differ in the term accounting for anelastic attenuation. In particular, the first model (hereafter named MOD1) has the following formulation:

$$\log_{10} Y = a + bM_w + c \log_{10}[\sqrt{(R_{hypo}^2 + h^2)}], \quad (2)$$

for which Y represents either PGV expressed in meters per second, PGA, or SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s, both expressed in meters per second squared, respectively. The model in equation (2) accounts for the source effect through the moment magnitude M_w and geometrical spreading through the hypocentral distance R_{hypo} (expressed in kilometers). The h parameter (in kilometers) is introduced to avoid unrealistically high values at short distances (e.g., Joyner and Boore, 1981). The inferred coefficients along with the uncertainties are listed in Table 1. We are analyzing data collected in a geothermal area, which is continuously perturbed through fluid injection and, as a consequence, can be highly fractured.

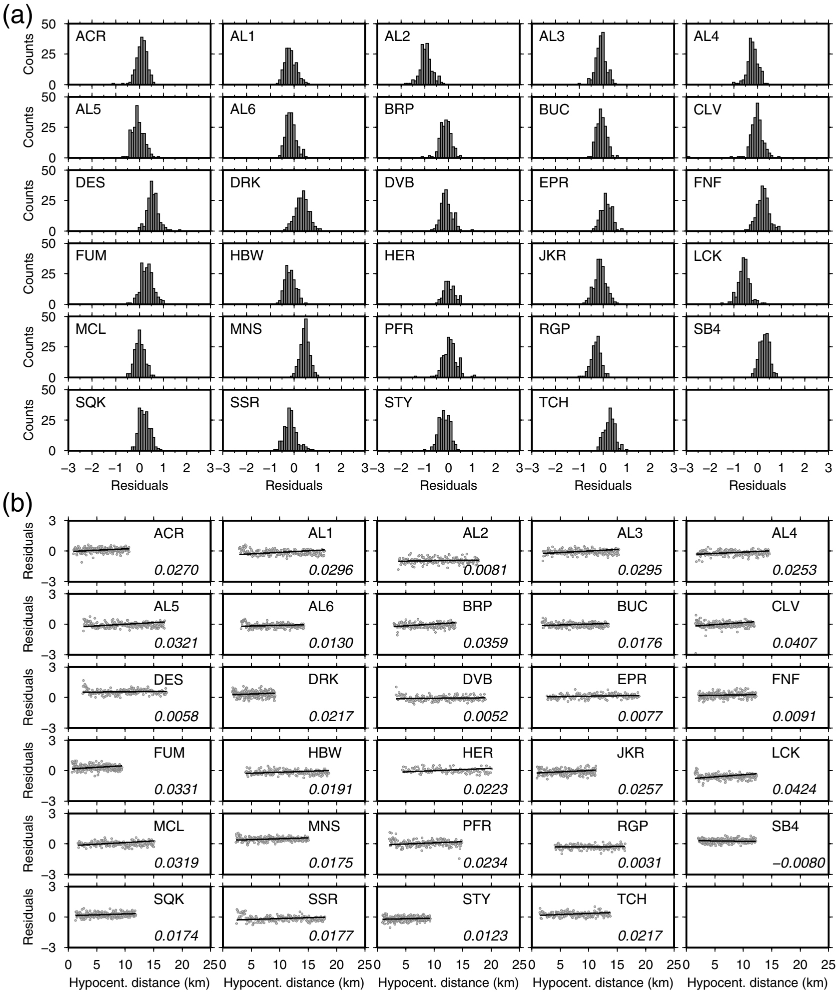


Figure 4. Single-station residuals distribution. (a) Residual histograms; (b) residuals distribution as a function of hypocentral distance for model MOD1 (reference model) at each station for PGA only, including slope of linear fit to residuals at each station.

It is physically reasonable to expect that anelastic attenuation could also play a key role. Therefore, we introduce an additional linear term depending on the distance, which leads to the model defined by the following equation (3) (hereafter named MOD2):

$$\log_{10} Y = a + bM_{\mathrm{w}} + c \log_{10}[\sqrt{(R^2_{\mathrm{hypo}} + h^2)}] + dR_{\mathrm{hypo}}. \tag{3}$$

Using the same data set as the one used to retrieve MOD1, the NLMRA is performed. The obtained coefficients and their uncertainties are listed in Table 2. Although we verified through an F-test that the coefficient d in equation (3) is statistically significant; however, looking at the R^2-value and the total standard deviation, it does not improve global fitting with respect to MOD1. Hence, in accordance with Occam's Razor, or *lex parsimoniae* (e.g., MacKay, 2003), we decided to choose MOD1 rather than MOD2 as the best reference model, as it is described by fewer parameters. The qualitative fit of MOD1 as function of hypocentral distance is shown in Figure 3 for PGV, PGA, and SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s and for the magnitudes M_{w} 1.0, 2.0, and 3.0, together with the inter-event and intra-event residuals of the models for the entire data set.

Station Effect and Corrected Model

As mentioned in the Introduction, we have adopted the technique proposed by Emolo *et al.* (2011) to introduce a first-order site/station-effect correction. It should be noted

Table 3

Site Correction Parameters for PGV, PGA, and SA at 0.2 s, 0.5 s, and 1.0 s

Station	Name	Latitude (°)	Longitude (°)	Nobs	S (PGV)	S (PGA)	S (SA at 0.2 s)	S (SA at 0.5 s)	S (SA at 1.0 s)
1	ACR	38.837	−122.760	202	0	1	−1	−1	−1
2	AL1	38.838	−122.883	179	−1	−1	0	1	−1
3	AL2	38.816	−122.898	180	−1	−1	−1	−1	−1
4	AL3	38.828	−122.857	201	1	−1	1	1	1
5	AL4	38.839	−122.836	184	−1	−1	−1	0	−1
6	AL5	38.840	−122.869	208	−1	−1	1	1	1
7	AL6	38.800	−122.861	199	−1	−1	−1	1	0
8	BRP	38.855	−122.797	181	−1	−1	1	−1	−1
9	BUC	38.823	−122.835	197	−1	−1	1	1	1
10	CLV	38.838	−122.790	209	1	−1	1	1	1
11	DES	38.768	−122.699	191	1	1	1	1	1
12	DRK	38.788	−122.803	209	1	1	1	1	1
13	DVB	38.764	−122.681	173	−1	−1	−1	−1	−1
14	EPR	38.747	−122.693	151	1	1	−1	1	1
15	FNF	38.771	−122.766	204	1	1	−1	−1	−1
16	FUM	38.793	−122.788	211	1	1	1	1	1
17	HBW	38.859	−122.876	172	−1	−1	−1	1	1
18	HER	38.845	−122.914	104	1	1	1	1	1
19	JKR	38.801	−122.760	209	−1	−1	−1	−1	−1
20	LCK	38.820	−122.742	191	−1	−1	−1	−1	0
21	MCL	38.855	−122.823	182	1	1	1	1	1
22	MNS	38.792	−122.717	206	1	1	1	1	1
23	PFR	38.753	−122.745	199	−1	1	−1	−1	−1
24	RGP	38.878	−122.811	157	−1	−1	−1	−1	−1
25	SB4	38.809	−122.829	202	1	1	1	1	1
26	SQK	38.823	−122.810	203	1	1	1	1	1
27	SSR	38.740	−122.711	180	−1	−1	1	1	1
28	STY	38.813	−122.781	200	−1	−1	−1	−1	−1
29	TCH	38.783	−122.737	167	1	1	1	1	1

Table 4

Regression Coefficients and Relative Uncertainty of Equation (4), Using MOD3

Parameter	$a \pm \sigma_a$	$b \pm \sigma_b$	$c \pm \sigma_c$	$h \pm \sigma_h$	$e \pm \sigma_e$	$\sigma_{\mathrm{Inter\text{-}event}}$	$\sigma_{\mathrm{Intra\text{-}event}}$	σ_{Total}	R^2
PGV (m/s)	-5.065 ± 0.069	1.320 ± 0.022	-1.966 ± 0.047	1.863 ± 0.191	0.189 ± 0.004	0.164	0.289	0.332	0.857
PGA (m/s^2)	-2.710 ± 0.064	1.165 ± 0.021	-2.244 ± 0.044	1.779 ± 0.158	0.225 ± 0.004	0.151	0.276	0.315	0.859
SA($T = 0.2$ s) (m/s^2)	-3.721 ± 0.082	1.448 ± 0.023	-1.802 ± 0.061	2.629 ± 0.244	0.203 ± 0.004	0.169	0.294	0.339	0.862
SA($T = 0.5$ s) (m/s^2)	-4.833 ± 0.083	1.555 ± 0.024	-1.838 ± 0.061	2.674 ± 0.239	0.182 ± 0.004	0.175	0.290	0.339	0.872
SA($T = 1.0$ s) (m/s^2)	-5.314 ± 0.077	1.506 ± 0.023	-1.918 ± 0.057	2.255 ± 0.216	0.166 ± 0.004	0.168	0.299	0.342	0.867

that the correction must be intended in a broader sense with respect to the standard site effect because it is not based on the V_{S30} values, which are not available for the stations used in the present study. Moreover, when the V_{S30} database obtained from the topography (Wald and Allen, 2007) is investigated (see Data and Resources) The Geysers area falls in the C-category soil type of the National Earthquake Hazard Reduction Program (NEHRP) classification (Building Seismic Safety Council, 2000, 2004). To overcome these problems, we analyzed the residuals distribution at each station obtained using the reference model MOD1 as shown in Figure 4 for PGA only for the sake of brevity. In particular at each station, Figure 4a shows residual distributions, whereas Figure 4b shows residual as a function of hypocentral distance at each station. The value of the slope of the fits is also mentioned in Figure 4b. As explained before through the Z-test, we tested the null hypothesis of a Gaussian zero-mean distribution at the 95% level of confidence. We assume that a deviation from the expected zero-mean value can be reasonably ascribed to a site/station effect that can be corrected by using a dummy correction factor. Indeed, due to the assumed definition of the residual ($\log_{10} Y^{\mathrm{Observed}} - \log_{10} Y^{\mathrm{Predicted}}$) a positive deviation from the zero-mean value can be interpreted as an underestimation of the model prediction with respect to the observations, whereas a negative deviation is to be interpreted as an overestimation of the model prediction. Thus, by Z-test in terms of both value and sign, at each station we assigned a dummy variable s,

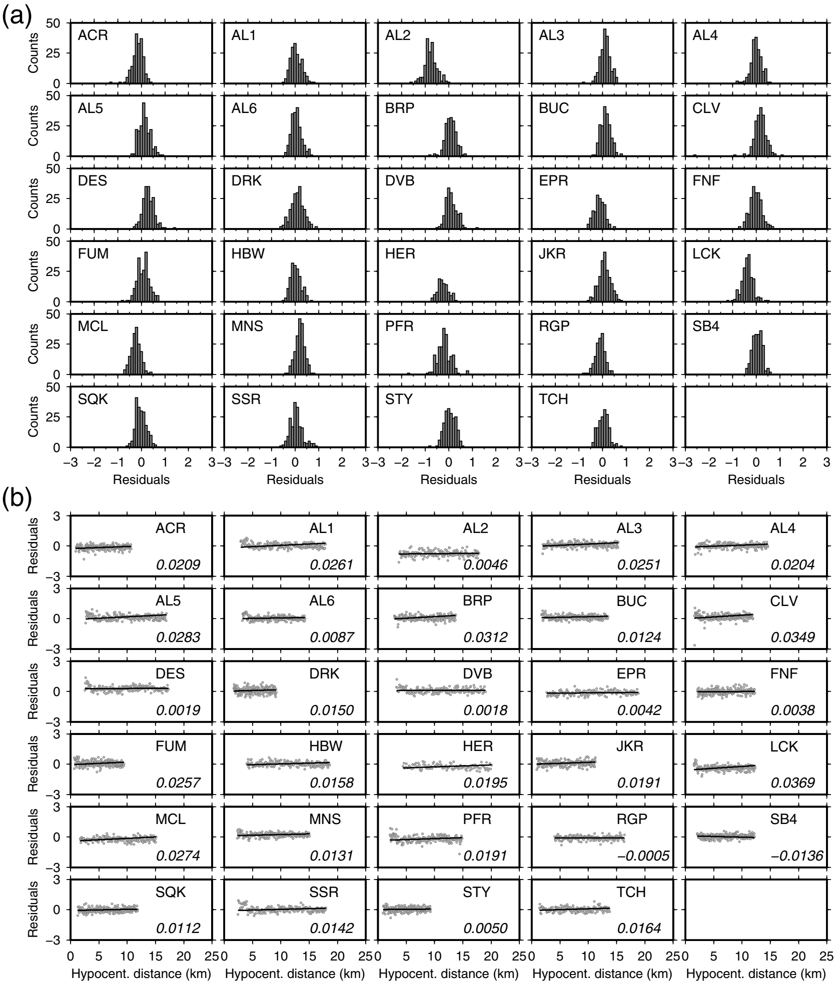


Figure 5. Same as for Figure 4, but for model MOD3 (corrected model). There is an improvement in the model after site corrections as now the maxima of residuals are centered at zero.

the value for which is −1, 0, or +1, depending on the mean residual value negative, zero, or positive, which allows for recovery of the observed residual deviation from the expected zero-mean value. The value of the s parameter with station code and coordinates, and the number of observations (Nobs) are listed in Table 3 for PGV, PGA, and SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s, respectively. The retrieved s parameter, along with a new coefficient e, is then used to set up an updated model, that is, the corrected model (hereafter named MOD3), which is formulated as

$$\log_{10} Y = a + bM_{\mathrm{w}} + c \log_{10}[\sqrt{(R_{\mathrm{hypo}}^2 + h^2)}] + es. \quad (4)$$

The inferred coefficients and their uncertainties are listed in Table 4. It can be noted that, after considering the site/station effect, the total standard deviation is reduced and the residuals distribution at each station is improved as their maxima are centered on zero (see Fig. 5). It is also worth noticing that there is a considerable improvement in intra-event standard deviation, whereas there is no significant change in inter-event standard deviation. This could be ascribed to the fact that intra-event residuals are associated with all those factors that are not captured by the distance metric and site classification (Atik *et al.*, 2010), hence confirming the effectiveness of the technique proposed for site/station correction. As an example, in Figure 5 we show the residuals histograms for PGA at each station and the residuals as a function of hypocentral distance. The slope of the fit line is also indicated in Figure 5b. An improvement in the results compared with those shown in Figure 4 can be clearly observed. Similar

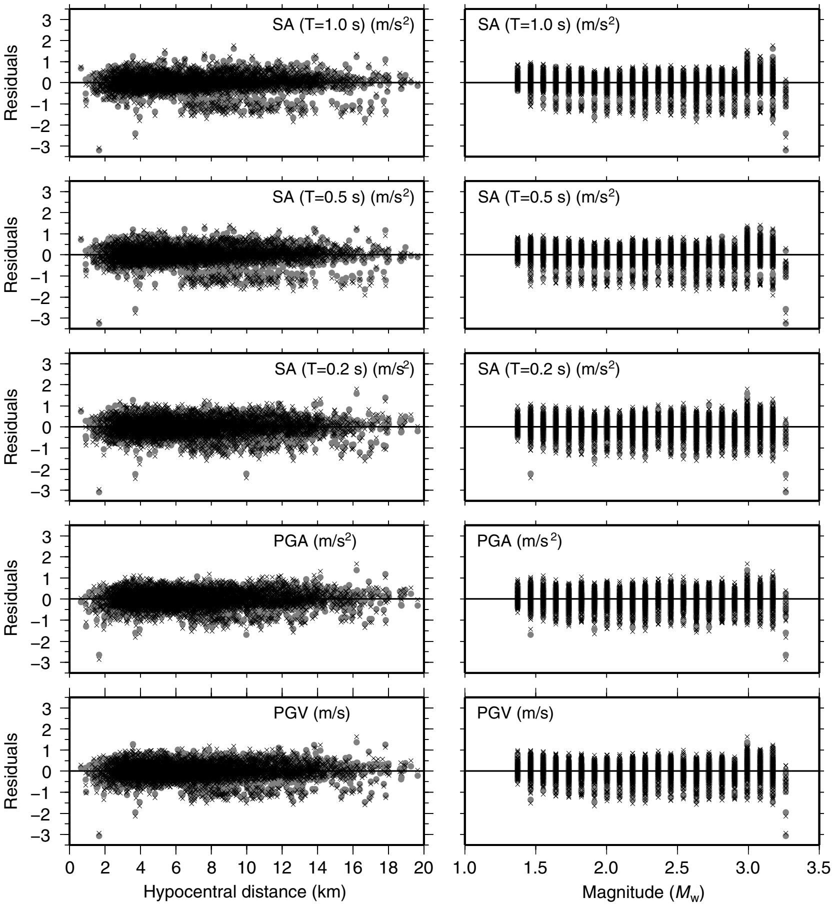


Figure 6. Distribution of residuals as function of hypocentral distance and magnitude; gray dots, show MOD3 residuals; black crosses, MOD1 residuals; black crosses, the subset of negative data at distances from 6 km to 15 km has shifted toward zero, which further indicates the improvement in MOD3 after introducing site/station effect correction parameter.

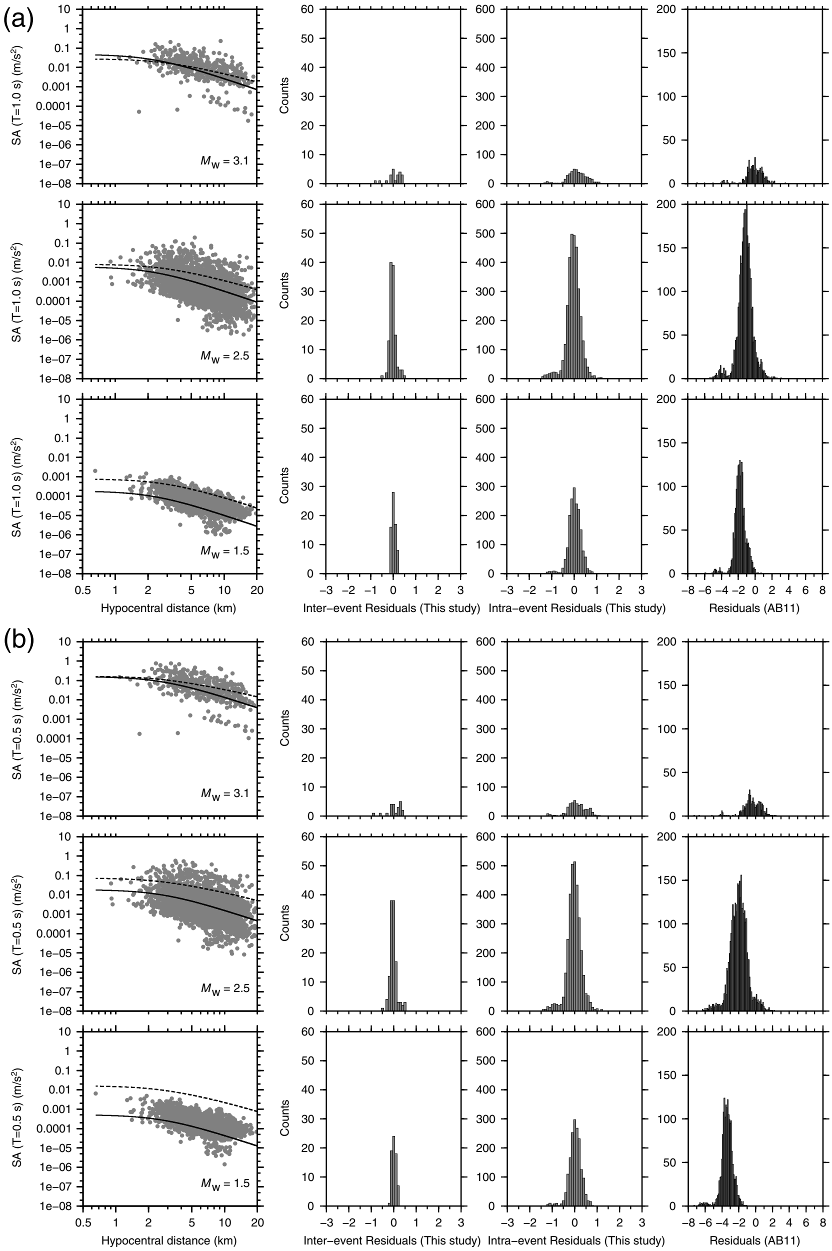


Figure 7. Fitting of model MOD3; solid line, comparison with site/station correction parameter $s = 0$; dashed line, comparison with model AB11; (a) fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for SA at 1.0 s; (b) fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for SA at 0.5 s; (c) fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for SA at 0.2 s; (d) fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for PGV; and (e) fitting for magnitudes $M_w \geq 3.0$, $2.0 < M_w < 3.0$, and $M_w \leq 2.0$, with corresponding inter-event and intra-event residuals for PGA. For smaller magnitude events, model AB11 overestimates, but with increases in magnitude the predictions of AB11 match our final model (MOD3). (*Continued*)

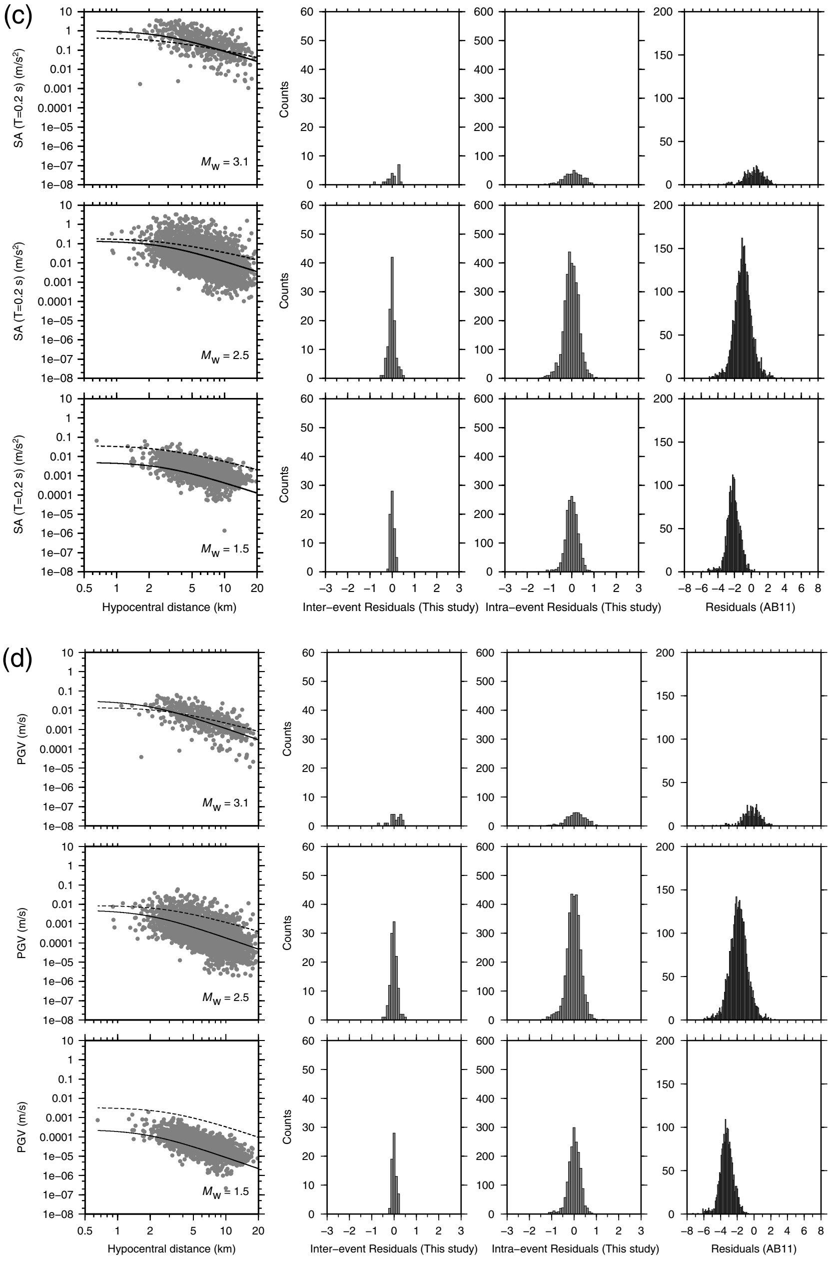


Figure 7. Continued.

improvement is observed for other ground-motion parameters also. Furthermore, in order to test the effectiveness of the formulation of the obtained model, we have also analyzed the distribution of residuals as a function of hypocentral distance and magnitude. No significant trend is observed but as shown in Figure 6, for all the ground-motion parameters, subset of negative residuals from distance between 6 km through 15 km moved to zero when MOD3 is used.

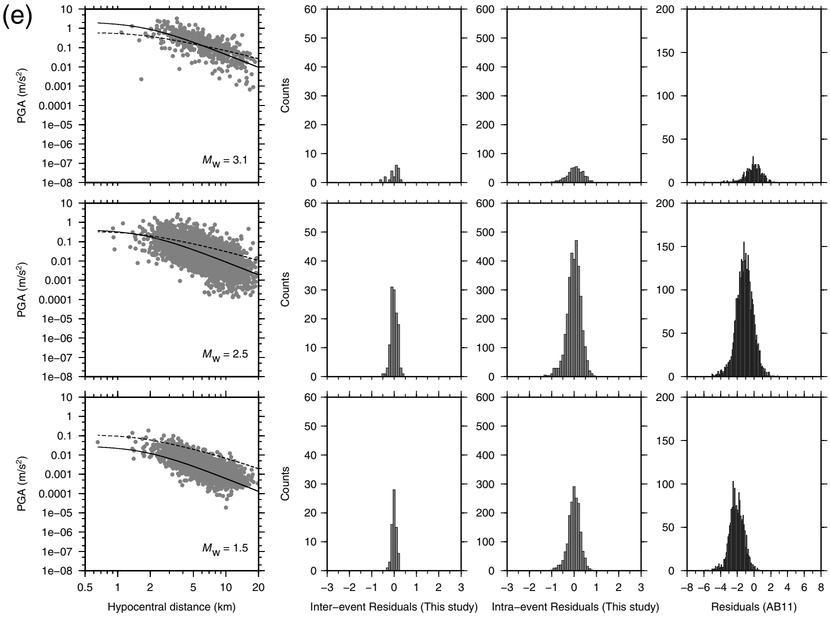


Figure 7. Continued.

The fit of model MOD3 is plotted in Figure 7a,b, and corresponding inter-event and intra-event residuals are also shown. The solid-line curves, which correspond to the best model (MOD3) obtained in the present study, are plotted by setting the site/station correction parameter $s = 0$, that is, for rock site condition for the magnitude $M_w = 1.5$, 2.5, and 3.1. When predictions have to be made at sites other than those used in the present study, we suggest using MOD3 with $s = 0$, but adding the contribution of the site effect using corrective coefficients (e.g., Borcherdt, 1994; Cantore *et al.*, 2010). In the same figure, we have compared the model AB11 with $V_{S30} = 760$ m/s for B/C site conditions which referred as dense soil to soft rock according to NEHRP site classification. As the general consideration, model AB11 tends to overestimate the observed values for smaller magnitudes. But with increase in magnitude there is a good match between both models (MOD3 and AB11). Our final model (MOD3) is applicable up to 20 km of hypocentral distance.

Following Atkinson (2006), we analyzed the standard deviations of the residuals distribution at individual stations before and after site/station correction. The average of these station specific values for all the stations, weighted by the number of observations per station, is shown in Figure 8, and is compared with the standard deviation of the regression obtained from the entire multi-station database (29 stations). The values of weighted mean and total standard deviations

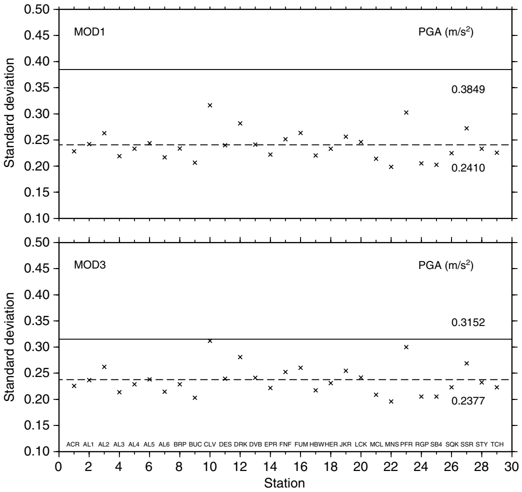


Figure 8. Comparison of standard deviation at each station with standard deviation obtained using complete data for PGA only; dashed line, weighted mean value of standard deviation over all stations; solid line, overall standard deviation. The introduction of the site correction parameter has improved the model as the gap between weighted mean standard deviation and overall standard deviation is reduced in MOD3 compared with MOD1. The respective standard deviations are written in the figure.

are also mentioned in Figure 8. The gap between weighted mean standard deviation and total standard deviation is reduced when MOD3 is used. Similar results are observed for PGV, SA(T) at $T = 0.2$ s, 0.5 s, and 1.0 s, but for the sake of brevity, we show results for PGA only. The standard deviation at all the stations is lower than that corresponding to the regression model, except at station CLV, which is higher than the other stations. This can be due to other effects related to the local geological conditions beneath the station, which are not fully accounted for by the introduced correction.

Finally, using MOD3, we have also analyzed the inter-event residual distribution. Indeed, as reported by Atik *et al.* (2010), the inter-event residuals account for the influence of source features that are not properly captured by the considered predictor variables (magnitude and distance in the present study). These features are, for example, the stress drop or the variation of slip in space and time. It could thus be interesting to analyze the dependency of the inter-event residual value as a function of depth. As shown in Figure 9, for all the ground-motion parameters we observed, only a slight positive trend of inter-event residuals with respect to event depth which is reported in each panel. Among other possible explanations, such as a variation in rigidity modulus due to the heterogeneities present in a highly fractured medium, following the interpretation proposed by McGarr (1984), the results could indicate that there is a possible decrease in stress with depth.

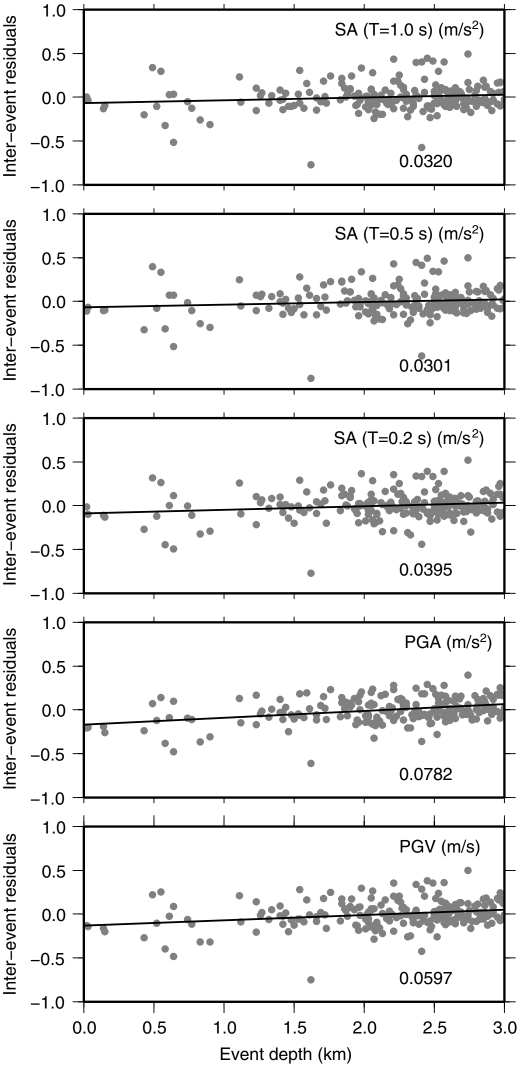


Figure 9. Distribution of inter-event residuals obtained from MOD3 with respect to event depth. The positive trend indicates the increase in stress drop with depth. The same trend is observed for all ground-motion parameters. The slope of the fitted line is stated in the figure.

Conclusions

In the present work we have performed an NLMRA, which allows one to separate the contributions of the inter-event and intra-event error to the total standard error (e.g., Abrahamson and Silva, 1997; Atik *et al.*, 2010). The intra-event dependencies on GMPEs are obtained for PGV, PGA, and SA(T) at $T = 0.2$ s, 0.3 s, and 1.0 s. The comparison of our GMPE models with the model proposed by Atkinson and Boore (2011) shows that our model is more robust when ground-motion predictions have to be made for induced earthquakes (smaller events) in The Geysers geothermal field. Moreover, if information about local geology is not available, our approach of introducing site/station effect correction has shown a significant improvement in the models as well. We have also demonstrated how the two components of the standard deviation (i.e., inter-event and intra-event) are related to source effect and local site effects, respectively. The reduction in total standard deviation after introduction of site/station effect correction is actually associated with a reduction of the intra-event component. We observe that the inter-event residuals distribution shows a slightly positive trend with depth, which, as suggested by McGarr (1984), could indicate a possible decrease in stress with depth (McGarr, 1984). The results obtained are promising and we recommend the use of the GMPE MOD3 for modeling of ground motion at The Geysers, which is applicable up to a hypocentral distance of 20 km. The model can be used for rock-site conditions selecting $s = 0$ for the site/station correction and $s = +1$ or -1 for which data are available to compute the residual distribution analysis. However, for sites characterized by geology different from the area analyzed in the present study, we suggest predicting ground-motion values by using MOD3 with $s = 0$ and adding the site-specific effect by using the corrective coefficients proposed by

Borcherdt (1994), which are also implemented, for example, in ShakeMap (Wald *et al.*, 1999).

Data and Resources

Waveforms and parametric data of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network have been retrieved from the Northern California Earthquake Data Center (NCEDC; network code BG; http://www.ncedc.org, last accessed January 2012). Figures have been generated with the Generic Mapping Tools (Wessel and Smith, 1991). The information about V_{S30} has been retrieved from http://earthquake.usgs.gov/hazards/apps/vs30/ (last accessed August 2012).

Acknowledgments

We wish to thank all those people who supported this work directly or indirectly. We are thankful to Ben Edwards who helped in converting duration magnitude to moment magnitude. We are also thankful to Carlo Cauzzi and another anonymous reviewer from whom suggestions have potentially improved this paper. The work has been supported financially by the 7th Research Program of the European Union through the Geothermal Engineering Integration Mitigation of Induced Seismicity in Reservoirs project.

References

Abrahamson, N. A., and W. J. Silva (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismol. Res. Lett.* **68,** 94–127.

Abrahamson, N. A., and R. R. Youngs (1992). A stable algorithm for regression analysis using the random effect model, *Bull. Seismol. Soc. Am.* **82,** 505–510.

Atik, A. L., N. A. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81,** 794–801.

Atkinson, M. G. (2006). Single-station sigma, *Bull. Seismol. Soc. Am.* **96,** 446–455.

Atkinson, G. M., and D. M. Boore (2011). Modification to existing ground-motion prediction equations in light of new data, *Bull. Seismol. Soc. Am.* **101,** 1121–1135.

Bachmann, C. E., S. Wiemer, J. Woessner, and S. Hainzl (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: Introducing a probability-based monitoring approach for enhanced geothermal systems, *Geophys. J. Int.* **186,** no. 2, 793–807, doi: 10.1111/j.1365-246X.2011.05068.x.

Bommer, J. J., S. J. Oates, M. Cepeda, C. Lindholm, J. Bird, R. Torres, G. Marroqøin, and J. Rivas (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal project, *Eng. Geol.* **83,** 287–306.

Boore, D. M., and G. M. Atkinson (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthquake Spectra* **24,** 99–138.

Borcherdt, R. D. (1994). Estimates of site-dependent response spectra for design (methodology and justification), *Earthquake Spectra* **10,** no. 4, 617–653.

Building Seismic Safety Council (2000). *National Earthquake Hazards Reduction Program (NEHRP) Part 1: Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, 2000 Edition (FEMA 368)*, Building Seismic Safety Council for the Federal Emergency Management Agency, Washington, D.C.

Building Seismic Safety Council (2004). *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, 2003 Edition (FEMA 450)*, Building Seismic Safety Council, National Institute of Building Sciences, Washington, D.C.

Cantore, L., V. Convertito, and A. Zollo (2010). Development of a site-conditions map for the Campania–Lucania region (southern Apennines, Italy), *Ann. Geophys.* **53,** no. 4, 27–37, doi: 10.4401/ag-4648.

Convertito, V., and A. Zollo (2011). Assessment of pre-crisis and syn-crisis seismic hazard at Campi Flegrei and Mt. Vesuvius volcanoes, Campania, southern Italy, *Bull. Volcanol.* **73,** no. 6, 767–783, doi: 10.1007/s00445-011-0455-2.

Convertito, V., N. Maercklin, N. Sharma, and A. Zollo (2012). From induced seismicity to direct time-dependent seismic hazard, *Bull. Seismol. Soc. Am.* **102,** no. 6, 2563–2573, doi: 10.1785/0120120036.

Draper, N. R., and H. Smith (1996). *Applied Regression Analysis*, Third Ed., John Wiley and Sons, Inc., New York, 407 pp.

Eberhart-Phillips, D., and D. H. Oppenheimer (1984). Induced seismicity in The Geysers geothermal area, California, *J. Geophys. Res.* **89,** 1191–1207.

Emolo, A., V. Convertito, and L. Cantore (2011). Ground-motion predictive equations for low-magnitude earthquakes in the Campania–Lucania area, southern Italy, *J. Geophys. Eng.* **8,** no. 1, 46–60, doi: 10.1088/1742-2132/8/1/007.

Enedy, S. L., K. L. Enedy, and J. Maney (1992). Reservoir response to injection in the southeast geysers, *Geoth. Resour. Counc. Spec. Rep.* **17,** 211–219.

Foulger, G. R., C. C. Grant, B. R. Julian, and A. Ross (1997). Industrially induced changes in earth structures at The Geysers geothermal area, California, *Geophys. Res. Lett.* **24,** 135–137.

Joyner, W. B., and D. M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake, *Bull. Seismol. Soc. Am.* **71,** 2011–2038.

Lindstrom, M. J., and D. M. Bates (1990). Nonlinear mixed-effects models for repeated measures data, *Biometrics* **46,** 673–687.

MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, Cambridge.

Majer, E. L., and T. V. McEvilly (1979). Seismological investigation at The Geysers geothermal field, *Geophysics* **44,** 246–269.

Majer, E. L., and J. E. Peterson (2005). Application of micro-earthquake monitoring for evaluating and managing the effects of fluid injection at naturally fractured EGS sites, *Geotherm. Resour. Counc. Trans.* **29,** 103–107.

Majer, E. L., R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma (2007). Induced seismicity associated with enhanced geothermal systems, *Geothermics* **36,** 185–222.

McGarr, A. (1984). Scaling of ground motion parameters, state of stress, and focal depth, *J. Geophys. Res.* **89,** 6969–6979.

Oppenheimer, D. C. (1986). Extensional tectonics at The Geysers geothermal area, California, *J. Geophys. Res.* **91,** 11,463–11,476.

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

Smith, J. L. B., J. J. Beall, and M. A. Stark (2000). Induced seismicity in the SE Geysers field, *Geotherm. Resour. Counc. Trans.* **24,** 24–27.

Stark, M. A. (1992). Micro-earthquakes a tool to track injected water in The Geysers reservoir, *Geotherm. Res. Counc. Spec. Rep.* **17,** 111–117.

Stark, M. A. (2003). Seismic evidence for a long-lived enhanced geothermal system (EGS) in the northern Geysers reservoir, *Geotherm. Resour. Counc. Trans.* **24,** 24–27.

Wald, D. J., and T. I. Allen (2007). Topographic slope as a proxy for seismic site conditions and amplification, *Bull. Seismol. Soc. Am.* **97,** 1379–1395.

Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet ShakeMaps: Rapid generation of

instrumental ground motion and intensity maps for earthquakes in southern California, *Earthquake Spectra* **15,** 537–555.
Wessel, P., and W. H. F. Smith (1991). Free software helps map and display data, *Eos Trans. AGU* **72,** 445–446.

Dipartimento di Scienze Fisiche, Unità di Recerca in Sismologia Sperimentale e Computazionale
Università degli Studi di Napoli “Federico II”
Complesso Universitario di Monte S. Angelo, via Cintia
80126 Napoli, Italy
(N.S., A.Z.)

Istituto Nazionale di Geofisica e Vulcanologia
Osservatorio Vesuviano
via Diocleziano 328
8014 Napoli, Italy
(V.C.)

AMRA S.c.a.r.l.
Analysis and Monitoring of Environmental Risk
Via Nuova Agnano 11
80125 Napoli, Italy
(N.M.)

Manuscript received 10 April 2012

Bulletin of the Seismological Society of America, Vol. 103, No. 3, pp. –, June 2013, doi: 10.1785/0120120197

Ⓔ

Predicting Ground Motion from Induced Earthquakes in Geothermal Areas

by John Douglas, Benjamin Edwards, Vincenzo Convertito, Nitin Sharma, Anna Tramelli,
Dirk Kraaijpoel, Banu Mena Cabrera, Nils Maercklin, and Claudia Troise

Abstract Induced seismicity from anthropogenic sources can be a significant nuisance to a local population and in extreme cases lead to damage to vulnerable structures. One type of induced seismicity of particular recent concern, which, in some cases, can limit development of a potentially important clean energy source, is that associated with geothermal power production. A key requirement for the accurate assessment of seismic hazard (and risk) is a ground-motion prediction equation (GMPE) that predicts the level of earthquake shaking (in terms of, for example, peak ground acceleration) of an earthquake of a certain magnitude at a particular distance. Few such models currently exist in regard to geothermal-related seismicity, and consequently the evaluation of seismic hazard in the vicinity of geothermal power plants is associated with high uncertainty.

Various ground-motion datasets of induced and natural seismicity (from Basel, Geysers, Hengill, Roswinkel, Soultz, and Voerendaal) were compiled and processed, and moment magnitudes for all events were recomputed homogeneously. These data are used to show that ground motions from induced and natural earthquakes cannot be statistically distinguished. Empirical GMPEs are derived from these data; and, although they have similar characteristics to recent GMPEs for natural and mining-related seismicity, the standard deviations are higher. To account for epistemic uncertainties, stochastic models subsequently are developed based on a single corner frequency and with parameters constrained by the available data. Predicted ground motions from these models are fitted with functional forms to obtain easy-to-use GMPEs. These are associated with standard deviations derived from the empirical data to characterize aleatory variability. As an example, we demonstrate the potential use of these models using data from Campi Flegrei.

Online Material: To be provided by editor.

1

Introduction

There is growing interest worldwide in using geothermal energy for electrical power production. In Europe, for example, there are at least 40 geothermal projects for electricity generation (most of them very small scale) in various stages of development, according to the International Geothermal Association (see Data and Resources). In areas with sufficient heat but inadequate fluid conductivity, enhanced geothermal systems (EGSs) are being initiated. In an EGS, the permeability of a geothermal reservoir is enhanced using hydraulic stimulation. This procedure induces microearthquakes by design but may also trigger larger events due to existing (tectonic) stresses. These larger events may be felt and provoke alarm in the local population or, in extreme cases, cause damage. In the case of the Deep Heat Mining project (Basel, Switzerland), a main event of relatively moderate magnitude (M_L 3.4, M_w 3.2) was triggered along with thousands of smaller shocks. The increase in seismicity was a cause for concern to the local population and led to a project shutdown and insurance claims amounting to more than $9 million (Giardini, 2009). The review article by Majer *et al.* (2007) discusses various cases of seismicity triggered and induced by geothermal power production.

The evaluation of seismic hazard requires ground-motion models linking event parameters (e.g., magnitude and location) to site parameters such as peak ground acceleration (PGA; Convertito *et al.*, 2012). These models are generally in the form of ground-motion prediction equations (GMPEs) although simulations (e.g., Douglas and Aochi, 2008) could

be envisioned. GMPEs for induced earthquakes should provide robust predictions for small ($1 \lesssim M_{\mathrm{w}} \lesssim 5$), shallow
2 (focal depths, $h \lesssim 5$ km) earthquakes at close source-to-site distances (hypocentral distances, $r_{\mathrm{hyp}} \lesssim 20$ km), a distance–magnitude–depth range that is poorly covered by existing ground-motion models. Douglas (2011) summarizes almost 300 empirical GMPEs for the prediction of PGA and nearly 200 empirical GMPEs for elastic-response spectral ordinates. In addition, Douglas (2011) lists dozens of GMPEs developed via approaches other than regression analysis on recorded ground-motion data. However, in that study, no GMPEs derived for earthquakes induced by geothermal activity were identified. There are a handful of models available that have used data from nonnatural earthquakes, such as mining-induced tremors (McGarr *et al.*, 1981; McGarr and Fletcher, 2005) and nuclear explosions (Hays, 1980), but their relevance to geothermally induced tremors is not clear despite the fact that mining-related GMPEs often cover the magnitude–distance–depth range of relevance to geothermal projects.

Bommer *et al.* (2006) developed a GMPE for peak ground velocity (PGV) for use in the area surrounding the Berlín (El Salvador) geothermal power project. The authors did this by adjusting a previously published GMPE, derived from accelerograms of natural seismicity in Europe and the Middle East, so that it better fit ground-motion data from seismic swarms with shallow focal depths in this area; records associated with geothermal power production were not available at the time. This relied on the assumption that shaking from these swarm events and those associated with geothermal power production are similar. After the installation of a monitoring network during and following the period of hydraulic stimulation, ground-motion data from local earthquakes (some of which were induced by the stimulation) were used to update the derived GMPE because the original model was found to overestimate ground motions in the geothermal field. Such an approach could be followed in seismically active areas for which local ground-motion data exist (such as El Salvador), but this adjustment procedure would not be possible for EGS projects in regions with little or no history of earthquakes or ground-motion monitoring. Various previous studies have demonstrated that earthquake shaking from small earthquakes ($M_{\mathrm{w}} \lesssim 6$) shows greater interregion variability than ground motions from larger events (Douglas, 2007; Chiou *et al*, 2010). Therefore, it is probable that GMPEs for induced seismicity in one (host) region would need adjustment for application in another (target) region.

As part of its Seventh Framework Programme (FP7), the European Commission funded a collaborative international research project, the Geothermal Engineering Integrating Mitigation of Induced Seismicity in Reservoirs (GEISER), to study various aspects of induced seismicity and how they modify the local seismic hazard. The current article is the result of an effort within the GEISER project to develop consensus ground-motion models for the prediction of PGA, PGV, and elastic-response spectral ordinates associated with induced earthquakes in geothermal areas. EGS project operators need to be able to estimate future seismic hazard before reservoir stimulations start; before stimulation, it is likely that ground-motion data would be limited and with a poor magnitude–distance distribution from which to derive site-specific GMPEs. Consequently, we sought generic GMPEs for use in future EGS projects that can be made more site specific once observations become available from local networks.

In this study, available ground-motion data associated with various locations and seismicity types are presented and analyzed to investigate the potential dependence of earthquake shaking on the type of seismicity (natural, geothermal-related events, or induced by gas extraction) and to derive empirical GMPEs based on these data. Stochastic models then are developed that account for epistemic uncertainty in the prediction of median ground motions. These models are used to simulate PGA, PGV, and response spectral accelerations that are then regressed to produce easy-to-use GMPEs. To assign estimates of the aleatory variability to be used with these GMPEs, the empirical data are analyzed to present models of the between-event (interevent) and within-event (intraevent) components of this variability for application within seismic hazard assessments for EGS sites. A possible approach to assign weights to these models is then presented for an example site.

Data Selection and Processing

For this article, we assembled data from surface instruments in Basel (Switzerland), Campi Flegrei (Italy), Geysers (United States), Hengill (Iceland), Roswinkel and Voerendaal (the Netherlands), and Soultz-sous-Forêts (France). Data from Campi Flegrei are only used in the independent validation of the GMPEs developed here. Some of these sets contain records from natural earthquakes (some of which are in areas of natural geothermal activity; i.e., Campi Flegrei and Hengill), geothermal-related events, and shocks induced 3
by gas extraction. Charlety *et al.* (2007) noted that the smallest geothermal event felt by the local population near Soultz-sous-Forêts was M_{L} 1.4 (a similar threshold holds for tremors felt in the Netherlands). We therefore aim to provide robust predictions down to such magnitudes. In the following sections, more details of these datasets are given; and, following that, methods for the computation of moment magnitudes and the correction for local site responses are presented.

All instrument-corrected data were assessed for their quality through a combination of visual inspection and analysis of the signal-to-noise ratio (SNR). For each record, the frequency range for which the SNR was greater than three was assessed. This criterion led to a rapid drop-off in the number of records available for analysis, from 5 Hz (0.2 s) downward. By 2 Hz (0.5 s), less than half the data can be used, and by 1 Hz (1 s) almost all are unusable. Based on this observation, we do not seek to use spectral accelerations beyond 0.5 s in our analysis, nor do we use data from

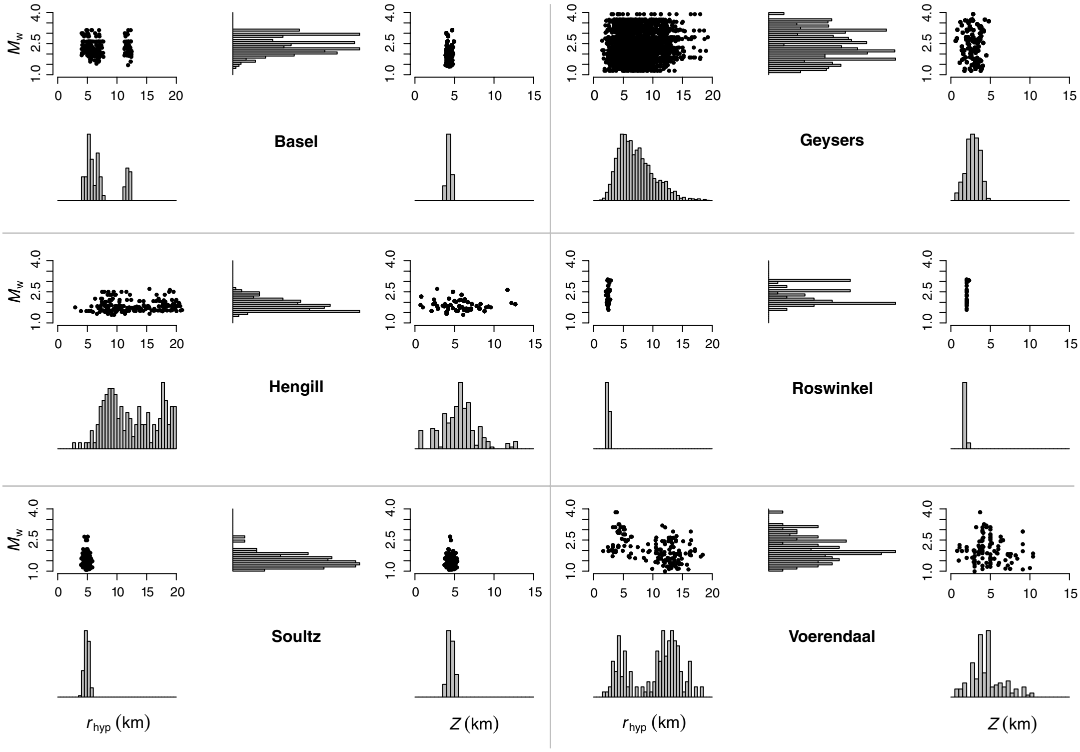


Figure 1. Magnitude–hypocentral distance–location distributions and magnitude–focal depth–location distributions of records used in this study. Only the records up to $r_{hyp} = 20$ km are plotted because there are few records beyond this distance range, with the exceptions of Basel and Hengill.

earthquakes below M_w 1. An analysis of the influence of high-cut filtering on the observations shows that, contrary to some expectations, response spectral accelerations from accelerograms are not very sensitive to high-frequency noise, confirming the conclusions of Douglas and Boore (2011). However, some records required a high cut-off frequency of less than 10 Hz, which means that the PGA (and other high-frequency parameters) are likely to be significantly reduced. Therefore, we excluded all records requiring a high cut-off frequency of less than 10 Hz. As an initial test, the residuals between the observed PGAs and those predicted by the GMPE of Bommer *et al.* (2007) were computed. This equation was chosen because it used data down to M_w 3 and hence covered part of the magnitude range of interest. The residual plot showed that a small number of records, predominantly from Geysers, had very low PGAs (more than 100 times smaller) relative to the predicted value. The inclusion of these data could hamper any analysis performed and were therefore removed. (Because Geysers contributes thousands of records to the analysis, this removal is not a significant loss). The reason for these very small amplitudes is not known but is likely due to instrument malfunction or misassociation of the record to a particular earthquake.

The distribution of the final selected data in terms of magnitude, hypocentral distance, focal depth, and location is shown in Figure 1. Although records from overlapping magnitude and distance ranges from various sites are included, some data come from distance ranges that are not well covered by data from other sites (e.g., Roswinkel). Because induced seismicity and local monitoring networks are closely localized in space, available records often come from similar distances, and magnitude–distance plots show banding. This lack of overlap in the available data impedes statistical analyses. In total, 3968 records (963 from Basel, 2328 from Geysers, 231 from Hengill, 61 from Roswinkel, 223 from Soultz, and 162 from Voerendaal) from 535 earthquakes and 119 stations are used to develop the GMPEs. Fifty-five records from 22 earthquakes and 13 stations from Campi Flegrei are used in an example application of these GMPEs. The focal mechanisms of the majority of these events are not known; therefore, we assume strike-slip faulting when comparing the observations to previously published GMPEs.

Basel

The Basel EGS project was proposed to provide up to 3 MW in electrical production and 20 MW in thermal production. A 200°C reservoir was to be created at a depth of 5 km beneath the city. A dense network of surface sensors (Swiss Seismological Service, SED; Baden Wüttemberg Seismological Service, LED) and borehole sensors (Geo Explorers Ltd.) was deployed to monitor seismicity. The SED surface instruments are either STS-2 broadband seismometers or EpiSensor accelerometers, geographically oriented. Site conditions are well known through the microzonation studies of Havenith *et al.* (2007) and site investigations undertaken as part of a probabilistic seismic-hazard study for nuclear power facilities in Switzerland (Fäh *et al.*, 2009). Basel lies on a sedimentary basin some several-hundred meters thick. Time-averaged shear-wave velocity down to 30 m (V_{S30}) of the sites around Basel tends to be around 400 m/s (National Earthquake Reductions Program [NEHRP] site class C), although rock-site stations at greater distances were also used for determination of magnitudes. The borehole sensors from Geo Explorers Ltd. are short-period geophones with a natural frequency of around 5 Hz and a damping coefficient of 0.2. Although they are located at various depths, we only use those on the surface for the ground-motion analysis (the borehole data were included in the magnitude determination). Some instruments were of unknown orientation, so all data were rotated to the direction of maximum amplitude. All events in the Basel dataset are geothermally induced events located by the SED. Data were corrected for the amplitude and phase response of the instrument and differentiated to provide acceleration-time series.

Campi Flegrei

Campi Flegrei caldera is a volcanic area that includes part of the metropolitan area of Naples (southern Italy), one of most densely populated areas in Europe. It is a large depression (with a radius of about 6 km) formed by huge ignimbritic eruptions, the last one having occurred 15,000 years ago (Deino *et al.*, 2004). For at least the past 2000 years, the area has also been affected by episodes of large uplift and subsidence, as shown by marine ingression levels in Roman and Middle Age monuments and ruins (Dvorak and Mastrolorenzo, 1990). The latest episodes of unrest, causing maximum uplift of about 3.5 m in 15 years (peak rate of about 1 m/yr), started in 1969 and, after about 10 years of stable ground level, restarted in 1982 and continued 4 until 1984. During this final period, maximum uplift rates were recorded by leveling networks and tide gauges (De Natale *et al.*, 2006), and more than 15,000 microearthquakes occurred with magnitudes from 0 to 4.2 (De Natale and Zollo, 1986). As these earthquakes were generally shallow with maximum depths of 3–4 km, they were strongly felt by the local population but did not produce significant damage (De Natale *et al.*, 1988). In January 1984, a digital network owned by the University of Wisconsin was installed at Campi Flegrei, consisting of 13 digital, three-component stations with a 125–250 Hz sampling rate (e.g., Aster *et al.*, 1989). The catalog analyzed here contains the data presented in the aforementioned papers, with the addition of other records that were not analyzed at that time.

Geysers

Geysers is a vapor-dominated geothermal field located in northern California. The main steam reservoir has a temperature of about 235°C and underlies an impermeable caprock with its base 1.1–3.3 km below the surface. Commercial exploitation of the field began in 1960. Since then, seismicity has become more frequent in the area and has increased with further field development (e.g., Majer *et al.*, 2007). The induced seismicity is concentrated within the upper 4 km of the crust, in the reservoir below production wells, and near injection wells.

Different temporary and long-term seismic networks have been deployed in the area during the last five decades. At present, local seismicity is monitored by the dense Berkeley–Geysers (BG) surface seismic network and some nearby stations of the Northern California Seismic Network (NCSN). The BG network consists of 29, three-component stations distributed over an area of about 20 × 10 km^2, covering 5 the entire geothermal field. Initially, each BG station was equipped with I/O Sensor SM-6 geophones with natural frequencies of 14 Hz. Toward the end of 2009, these instruments were replaced by OYO Geospace GS-11D 4.5 Hz 6 sensors. The BG stations operate in trigger mode, and the waveform segments recorded since the end of July 2007 are made available by the Northern California Earthquake Data Center (NCEDC). No information about V_{S30} of the recording sites is currently available.

We analyzed induced seismicity at Geysers between August 2007 and February 2011. For the study region, there are waveforms of more than 11,000 located events with magnitudes larger than 1.0 available from the NCEDC. The largest earthquake recorded was the 4 January 2009 M_w 4.3 event. We associated all data with events from the NCEDC earthquake catalog, updated the metadata for all traces, and automatically repicked the *P*-wave first-arrival times for quality control. Because different magnitude types (M_D, M_L, and M_w) are used in the original catalog, moment magnitude has been recomputed here for all the events with $M \geq 1.5$ (any scale). To compile a representative subset for this study, we divided the available range of catalog magnitude into bins 0.25 units wide. Within each bin, we selected those 10 events with the most validated *P*-wave picks at BG stations to ensure accurate locations and high-quality waveforms.

Hengill

Ground-motion data recorded close to the Hengill (southwestern Iceland) geothermal system by a temporary broadband network installed within the framework of the I-GET FP6 project (Jousset and François, 2006) were used 7 here. The temporary network operated from late June 2006 until mid-October 2006 and was composed of seven Güralp Systems broadband instruments (CMG-3ESP and CMG-40TD) distributed to monitor and explore the Hengill hydrothermal system. In addition, data from the three stations 8 (Lennartz LE3-5s instruments) of the South Icelandic 9 Lowland (SIL) permanent seismic network, operated by the Icelandic Meteorological Office (IMO), nearest Hengill (KRO, HEI, and SAN) were also collected for this study. Some basic site descriptions are available for the temporary stations, which show variations; however, it is assumed that the stations are all located on rock because of the general geology of the Hengill area (shallow volcanic soils overlying lava of various ages). Jousset and François (2006) report previous studies suggesting that cooling, mostly due to natural heat loss, and consequential thermal contraction and cracking in the heat source are responsible for the continuous small-magnitude seismicity in the Hengill area. This was deduced by the non-double-couple focal mechanisms with large explosive components, which may be attributable to fluid flow into newly formed cracks. The IMO earthquake catalog was queried to find those earthquakes with moment magnitudes larger than M_w 1.0 that were recorded by one or more of the ten instruments installed in the Hengill region. The acceleration time histories (derived by time-domain differentiation from the velocity measurements recorded by the broadband sensors) corresponding to these earthquakes were selected.

Roswinkel

The Roswinkel (northeastern Netherlands) natural gas field is situated in a heavily faulted anticline structure in Triassic sandstones at a depth of around 2.1 km. The field was in production from 1980 to 2005, while seismicity was observed from 1992 to 2006, with 39 earthquakes in total and a strongest event of magnitude M_L 3.4 and epicentral intensity of $I_0 = VI$. The seismicity that occurred so far has been associated with existing faults on top of the reservoir (van Eck *et al.*, 2006; Dost and Haak, 2007). The dataset used in this study contains 27 events with strong-motion recordings from the Roswinkel village obtained using SIG AC-23 sensors. 10

Soultz

The Soultz geothermal exploitation began in the late 1980s as a collaborative French–German project. The first subterranean circulation of water using the drilled boreholes was achieved in the late 1990s. The boreholes were then deepened to about 5 km and various reservoir stimulations were undertaken in the first decade of this century. Electricity has been produced since June 2008. The data used here come from three permanent three-component surface stations (FOR, OPS, and SRB) installed by École et Observatoire des Sciences de la Terre (EOST) of the University of Strasbourg in 2003 close to the injection wells of the EGS (Charlety *et al.*, 2007). These stations record amplitudes that are proportional to ground velocity, which is then converted to acceleration by differentiation and application of the calibration factor. The records are of injection experiments conducted in 2003. A high-quality earthquake catalog was provided by EOST. Because the recorded events are all induced in the geothermal reservoir, the records are associated with similar hypocentral distances.

Voerendaal

The Roer Valley Rift System is an active rift system in the Lower Rhine embayment in the border area of the Netherlands, Belgium, and Germany (Dost and Haak, 2007). Most of the seismicity in the area is situated within the Roer Valley graben, and this is associated to the main bounding faults: the Peel boundary fault to the northeast and the Feldbiss fault to the southwest. The village of Voerendaal is located on the South Limburg block, southwest of the Feldbiss fault. The region around Voerendaal has shown anomalous swarmlike seismicity at relatively shallow depths (around 3–8 km). A first swarm was detected in 1985 and lasted for more than a month. Nine events were located in 1985, the largest of which had a magnitude of M_L 3.0 and a maximum epicentral intensity of $I_0 = IV$. After 15 years, a new swarm of events started in the same area on 20 December 2000. This time the swarm lasted for more than a year, with 139 detected events and a strongest event with magnitude M_L 3.9 and intensity of $I_0 = VI$. The dataset used in this study is composed of 136 events from the Voerendaal area between April 1999 and August 2009. The waveforms consist of both short-period recordings from the regional seismic network (Willmore MkIII sensors) and accelerometric recordings from within the Voerendaal village, where three SIG AC-23 strong-motion sensors have been deployed since the start of the second swarm.

Computation of Moment Magnitudes

One of the main difficulties in analyzing data from small earthquakes in different areas is the lack of a mutually consistent magnitude scale. The magnitudes of such small shocks are generally not computed by international agencies, such as the International Seismological Centre (ISC), or by teleseismic scales such as surface-wave, body, or moment magnitude. Consequently, local or duration magnitudes are the only measures of the size of small events usually available. These magnitude scales are notorious for being network dependent, and therefore it is difficult to know if, for example, an M_L 2 earthquake in one area is truly the same size as an M_L 2 earthquake in another. In light of this, using the technique presented by Edwards *et al.* (2010) we have calculated, using an automatic procedure, moment magnitudes (M_w) for the vast majority of earthquakes considered here. 11

The method is based on the far-field spectral model of Brune (1970, 1971), and it was shown previously to provide magnitudes consistent within ±0.1 units of moment tensor solutions of $M > 3$ events in Switzerland. The method is extended to lower magnitudes, relying only on a sufficient bandwidth (at least one decade, $\log_{10}$ unit) of the spectrum being visible above the noise. In their comparison of Swiss 12 moment tensor solutions, Edwards *et al.* (2010) showed that the site effect on M_w in Switzerland was negligible because the majority of recordings were from hard-rock sites. However, in the case of limited distance coverage of records analyzed herein, the instruments may be located entirely on sediments. We therefore allowed for site effects in the case of the Basel data, using the rock reference of Poggi *et al.* (2011); and, because we wanted to predict ground motions at the surface, we excluded borehole data.

The Geysers data are also located on sediments, which considerably amplify ground motions. In the case of unknown reference-velocity profile and amplification, there is a coupled trade-off between magnitude and site effects. However, for several events, moment magnitudes have been independently determined by Berkeley Seismological Laboratory (BSL). The moment tensor analysis used for their determination is insensitive to site effects due to the long-period signals that were analyzed. To constrain our joint inversion for site amplification and unknown M_w, we fixed the magnitudes of events with known $M_w > 3.5$. Given sufficient recordings across all stations from these events, the resulting inversion for unknown M_w and site effects is decoupled, extending the determination of M_w to low magnitudes. For the Hengill and Soultz data, no reference (M_w or site amplification) exists, so we assume negligible site effects due to the rock-site classifications.

As a check of the calculated M_w values and to determine magnitude-scaling relations between published and recalculated magnitudes, we compared all moment magnitudes to agency magnitudes when possible. Detailed analysis and interpretation of the M_w determination is presented by B. Edwards and J. Douglas (2012, unpublished manuscript). As few independent M_w estimates exist, it is difficult to objectively assess our calculated magnitudes. Nevertheless, comparison with magnitudes from BSL, where values were not fixed ($M_w < 3.5$), showed agreement to within 0.15 units down to M_w 3.2. Furthermore, the scaling of M_w determined for this study was found to be consistent with the M_L:M_w scaling relations of Grünthal *et al.* (2009) and Goertz-Allmann *et al.* (2011), down to M_w 1.

For some of the poorly recorded (generally the smallest) earthquakes, moment magnitudes could not be calculated with this approach. In the case of Basel, existing formulas from Goertz-Allmann *et al.* (2011) were available to estimate M_w from M_L for such events. For the events from other locations, region-specific magnitude-conversion formulas were used (Table 1). These formulas were derived using linear least-squares regression, assuming that the catalog magnitudes are definitive and all errors in M_w are equal. Although the catalog magnitudes are not, in fact, definitive, the majority were provided without error estimates, limiting the scope for analysis. These conversions introduce uncertainties into the analysis, especially when the standard deviations of the formulas are large (i.e., for Hengill and Voerendaal), but they have the benefit of increasing the number of records available for analysis.

Table 1
Magnitude-Conversion Formulas Used to Estimate M_w for Those Earthquakes for which M_w Could Not Be Calculated

Site	Equation	Standard Deviation
Basel	$M_w = 0.594M_L + 0.985$ for $M_L < 2$	0.159 for $M_L < 2$
	$M_w = 1.327 + 0.253M_L + 0.085M_L^2$ for $2 \leq M_L \leq 4$	0.134 for $2 \leq M_L < 4$
	$M_w = M_L - 0.3$ for $M_L > 4$ (Goertz-Allmann *et al.*, 2011)	0.175 for $M_L > 4$
Geysers	$M_w = 0.90M_D + 0.47$	0.08
Hengill	$M_w = 0.546M_{w,IMO} + 1.072$	0.28
Roswinkel	$M_w = 0.578M_L + 1.168$	0.10
Soultz	$M_w = 0.614M_L + 0.433$	0.19
Voerendaal	$M_w = 0.641M_L + 1.018$	0.33

Correction for Site Response

GMPEs are usually corrected for local site response and attenuation conditions relative to a given reference before application in local or regional hazard studies (e.g., Delavaud *et al.*, 2012). Knowing the precise reference site condition of a GMPE is therefore important because not having this information can lead to the introduction of large epistemic uncertainties in probabilistic seismic-hazard analysis. In the case of most GMPEs, the reference condition is rather loosely defined (often in terms of a V_{S30} value or an NEHRP site class).

To address the problem of the starkly different site conditions across the different datasets in our study, all waveforms were corrected for a site-specific amplification response in addition to site attenuation, κ (Anderson and Hough, 1984). The velocity profile used for the reference condition was the generic rock profile determined by Poggi *et al.* (2011) for Switzerland, while the corresponding reference attenuation was $\kappa = 0.016$ s (Edwards *et al.*, 2011). Site-specific estimates of amplification and κ were obtained as a side product of the magnitude-determination procedure (Edwards *et al.*, 2008; Edwards and Rietbrock, 2009), which introduced a degree of compatibility between amplification, attenuation, and magnitudes. The Fourier response (amplification and attenuation relative to the reference condition) was removed from each waveform in the frequency domain, and the corrected spectrum was then returned to the time domain preserving the phase, thereby providing site-corrected accelerograms. As a result of this process, all data were cor- 13 rected to a commonly known site condition. To use a GMPE developed with these data for a specific site condition, appropriate amplification and attenuation corrections needed to be applied.

Empirical Analysis

In this section we analyze the ground-motion data to investigate possible differences in shaking from the locations considered here. In agreement with current practice, the geometric mean of the two horizontal components of pseudospectral acceleration (PSA) for 5% damping is considered. Of particular interest in relation to ground-motion variability is the similarity of time histories observed at a particular geothermal site. Because the recordings are of events in a relatively small source region and the waves follow similar paths to closely spaced receivers, this should provide an insight into single-station, single-source sigma (Atkinson, 2006; Rodriguez-Marek *et al.*, 2011). This issue is discussed in
14 a subsequent section. The two types of data analysis considered here (analysis of variance [ANOVA] and regression) are complementary because ANOVA does not impose a functional form and relies on many records from overlapping magnitude–distance ranges from different locations, whereas regression analysis requires a functional form (or forms) to be assumed but data from different regions do not need to overlap.

Analysis of Variance

Douglas (2004a,b, 2007) uses one-way ANOVA to test the differences in observed ground motions for a given magnitude and distance in two or more regions. In this technique, ln(PSA)s computed from ground-motion records in different regions are binned into small magnitude and distance intervals, and the mean ln(PSA)s for each magnitude–distance range are computed for each region along with their standard deviations. For bins with sufficient records, ANOVA allows the computation of the significance level of the difference between average ground motions in the different areas. The advantage of this approach over techniques involving the derivation and comparison of GMPEs for the different regions is that no assumptions need to be made on the functional form of the GMPEs. Furthermore, assessing the significance of differences between predictions from GMPEs is not straightforward (Douglas, 2007). The disadvantage of this technique is that it requires dense overlapping datasets for the magnitude–distance ranges of interest because there must be at least two (and preferably many more) records per region per bin, which often limits its application to small earthquakes at moderate distances.

Because of the short (and accurately determined) hypocentral distances, it was decided to use a fine grid in terms of
15 r_{hyp} and a coarser grid in terms of magnitude. Therefore, the data space was gridded into 0.5 $M_{\mathrm{w}} \times 2$ km bins, within which it was assumed that the ground motions were similar. Bins with more than five records from a single location were identified to highlight those magnitude–distance ranges in which sufficient records for robust ANOVA existed. The average site-corrected spectra for each location and the magnitude–distance bins with data from more than one location are displayed in Figure 2.

This figure indicates that very-near-source ground motions (within 7 km) vary significantly in the different locations across the whole period range of interest. At greater distances, the differences between most of the spectra become insignificant. This apparent difference in ground motions between locations could be attributed to differences in stress drop because the differences are more apparent at short periods and the spectra converge at longer periods. Reasons for these regional differences in stress drop could be related to variations in average focal depths and elastic/mechanical properties of the media. There does not seem to be any correlation between the type of seismicity and ground-motion amplitudes because, for instance, the Soultz spectra are generally much higher than average for a certain magnitude–distance range whereas the Basel spectra are generally lower than average, and both are examples of induced seismicity. Similarly Voerendaal and Hengill present significantly different average spectra, and these datasets are both examples of natural seismicity.

Regression Analysis

As a complementary technique to ANOVA, in this section we develop GMPEs by regression analyses of the available ground-motion data. These GMPEs allow the scaling of ground motions with respect to magnitude and distance to be studied and compared to those predicted by the stochastic model developed in the following section, Development of Generic Stochastic Models. In addition, differences between ground motions from different areas are investigated. Finally, the predictions are compared to those made by existing GMPEs derived from data from moderate and large earthquakes, specifically those by Ambraseys *et al.* (2005; herein referred to as AB05), Bommer *et al.* (2007; referred to as BM07), and Massa *et al.* (2008; referred to as MS08), to see if such models can be extrapolated to the prediction of shaking from small events.

We derive a specific GMPE using nonlinear mixed-effect regression (Lindstrom and Bates, 1990; Abrahamson and Youngs, 1992), which accounts for between-event and within-event variabilities (Al Atik *et al.*, 2010). Only records with $r_{\mathrm{hyp}} < 50$ km are used. The model selected for regression (model 1) has a standard functional form (equation 1) accounting for first-order effects of magnitude scaling, near-source saturation, geometrical spreading, and anelastic attenuation:

$$\ln Y = a + bM + c \ln \sqrt{r_{\mathrm{hyp}}^2 + h^2} + d r_{\mathrm{hyp}}, \qquad (1)$$

where Y is the response variable corresponding to PGA, PGV, or PSA at various structural periods (in SI units) and a, b, c, d, and h are regression coefficients. Coefficients obtained from the regression analysis, along with their uncertainties and the two principal components of the standard deviation, are reported in Table 2 for PGA, PGV, and PSA for three selected periods. (See Ⓔ Table S2 of the electronic supplement to this 16
article for coefficients for periods up to 0.5 s.)

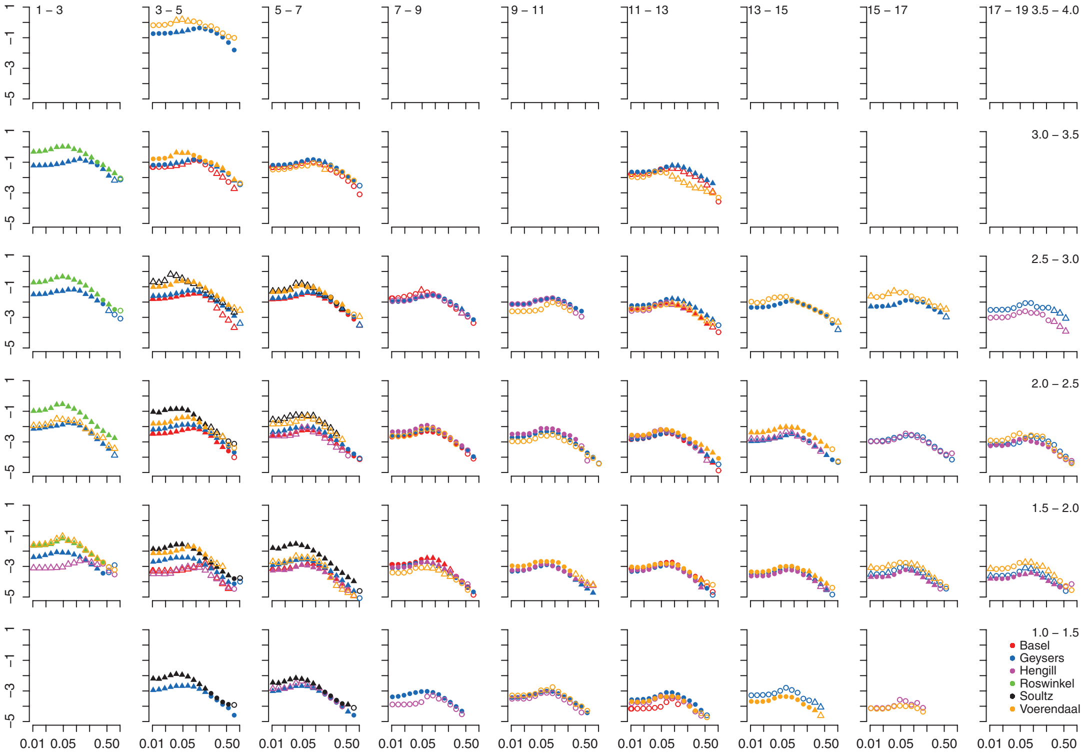


Figure 2. Average site-corrected response spectra observed in each location in M_w-r_{hyp} bins of 0.5 M_w × 2 km. A triangle means a significant difference was found at the 5% level between PSAs at the different locations, and a circle implies no significant difference was found. Filled symbols indicate more than five records from a location are in the M_w-r_{hyp} bin, and unfilled indicates five records or less. 34

Table 2

36 Coefficients for Model 1 for Selected Periods in which τ is the Between Event, ϕ the Within Event, and σ the Total Standard Deviations

	Period	$a \pm s_a$	$b \pm s_b$	$c \pm s_c$	$h \pm s_h$	$d \pm s_d$	τ	ϕ	σ
Uncorrected for site effects	PGA 0.01	−5.984 ± 0.427	2.146 ± 0.069	−1.772 ± 0.208	2.511 ± 0.595	−0.023 ± 0.011	0.792	0.829	1.147
	0.10	−6.444 ± 0.329	2.376 ± 0.056	−1.410 ± 0.167	1.751 ± 0.704	−0.039 ± 0.009	0.815	0.623	1.026
	0.20	−9.513 ± 0.243	2.805 ± 0.058	−0.776 ± 0.121	Constrained to zero	−0.057 ± 0.008	0.800	0.612	1.007
	0.50	−11.232 ± 0.480	2.912 ± 0.081	−0.953 ± 0.223	2.722 ± 1.181	−0.022 ± 0.011	0.638	0.816	1.036
	PGV	−10.367 ± 0.449	2.018 ± 0.136	−1.124 ± 0.183	2.129 ± 0.895	−0.046 ± 0.010	1.811	0.745	1.958
Corrected for site effects	PGA 0.01	−6.514 ± 0.423	1.995 ± 0.085	−1.468 ± 0.200	2.490 ± 0.688	−0.029 ± 0.010	0.730	1.079	1.303
	0.10	−7.991 ± 0.229	2.376 ± 0.063	−0.827 ± 0.106	1.058 ± 1.049	−0.056 ± 0.006	0.589	0.787	0.983
	0.20	−10.024 ± 0.219	2.784 ± 0.057	−0.850 ± 0.103	1.080 ± 0.979	−0.041 ± 0.006	0.554	0.658	0.860
	0.50	−12.736 ± 0.282	3.056 ± 0.070	−0.675 ± 0.129	1.218 ± 1.434	−0.050 ± 0.008	0.572	0.698	0.903
	PGV	−9.999 ± 0.681	1.964 ± 0.122	−1.405 ± 0.321	2.933 ± 1.088	−0.035 ± 0.016	1.029	1.553	1.863

The magnitude scaling of the derived GMPEs (coefficient b) closely matches the magnitude dependencies reported by Douglas and Jousset (2011) from previous empirical GMPEs derived using data from small ($M_w \lesssim 3$) natural and mining-related earthquakes (their fig. 1), thereby suggesting that the magnitude scaling of induced, mining, and natural seismicity are comparable. Comparing the coefficient b of the GMPEs for PGV with the coefficients for PGA and PSA shows that the PGV for small events is associated with very high frequencies, and hence the method of Bommer and Alarcón (2006) to estimate PGV from PSA (0.5 s) is not recommended for such earthquakes. The regression coefficients c and d indicate fast decay with distance, which could be attributable to strong anelastic attenuation (i.e., low Q values). Except at high frequencies, the total standard deviation obtained from the regression on site-corrected data is lower than that corresponding to uncorrected data, which shows that the site effect contributes significantly to ground-motion variability. It should be noted that the high-frequency site correction may actually reintroduce some source variability that was effectively hidden by the site attenuation (κ). Consequently, we observe an increase in total σ at high frequencies, including PGA and PGV, driven by an increase in the between-event variability (τ).

Figure 3 shows residual plots for data corrected for site response. As is usual, the within event residual distributions show larger dispersion compared with the between-event distributions. The comparison with other models shows that the residual distributions of BM07 and MS08 are similar to that of our model. On the other hand, the AB05 GMPEs lead to a wide distribution of residuals, which can be attributed to the applicability of this GMPE to larger magnitudes. Indeed, BM07 is based on a dataset in which the minimum M_w was 3.0, MS08 analyzed data with minimum M_w of 4.0, and AB05 considered strong-motion data relative to earthquakes with $M_w \geq 5.0$.

To investigate what consitutes the largest contribution to the total residuals' distribution, we analyze them as a function of both magnitude and distance. The obtained results, which for brevity are not reported here, show that the largest contribution to the residual dispersion comes from the distance; and, in particular, data recorded at shorter distances feature higher residual values at all structural periods. This indicates that, for data collected from earthquakes occurring in these areas, it is likely that anelastic attenuation plays an important role. On the other hand, aside from the results obtained from the AB05 model, the residuals as a function of magnitude are characterized by a quite uniform dispersion, mostly centered on zero.

To analyze the effect of focal depth on the regression models, and thus on the predictions of the ground motion, we implemented two additional models to be compared with the model reported in equation (1). From now on we refer only to data corrected for site response. First, we selected a model in which the regression coefficient h is no longer used in conjunction with r_{hyp} so as to discuss the relative importance of this coefficient. Second, we implemented a model in which epicentral distance, r_{epi}, is used instead of r_{hyp}. Models 2 and 3 have the following formulations:

$$\ln Y = a + bM + c \ln r_{\mathrm{hyp}} + d r_{\mathrm{hyp}}, \tag{2}$$

and

$$\ln Y = a + bM + c \ln \sqrt{r_{\mathrm{epi}}^2 + h^2} + d r_{\mathrm{epi}}. \tag{3}$$

The coefficients and the associated uncertainties for the two models are listed in Table 3. Concerning the PGA, the between- and within-event standard deviations of models 1 and 2 are equal, while the between-event component for $T \leq 0.1$ s for model 3 is slightly reduced. This could be a consequence of greater scatter in focal depths, which are more poorly defined than the epicenters.

Unlike moderate and large earthquakes, which rupture a large proportion of the seismogenic layer, small earthquakes, such as those induced and triggered by EGSs, are associated with ruptures of less than a kilometer. Consequently, whether this rupture occurs at a depth of, say, 20 km or at 2 km will have a large impact on the shaking at the surface. To test the effect of the focal depth on the predicted values from the three models, we considered a set of epicentral distances, and for each distance the actual range of focal depths contained in the data is taken into account. For each term composed of depth, r_{epi} and r_{hyp}, we considered the differences $\ln Y_{\mathrm{model}\ i} - \ln Y_{\mathrm{model}\ j}$ where i and j correspond to 1, 2, and 3. We note that the differences depend only weakly on structural period and magnitude. Therefore, we show (in Fig. 4) results only for the PGA and for a representative magnitude, M_w 2.5. As expected, models 1 and 2 show the same behavior for depths larger than 3 km. The models differ for $r_{\mathrm{epi}} <$ 1 km and focal depths less than 3 km, for which model 2 provides predictions larger than those of model 1. This is due to coefficient h in model 1, which avoids unrealistic Y values at small distances. The comparison between models 1 and 2 with respect to model 3 is more important for evaluating the effect of the focal depth. Aside from the absolute values, the differences between models 1 and 3 and those between models 2 and 3 share the same characteristics. In particular, all the models are similar starting from epicentral distances of 10–15 km. On the other hand, a net difference is observed for $r_{\mathrm{epi}} < 5$ km, with a different trend depending on the depth and an inflection point at about 3 km. For depths less than 3 km, the predictions made by models 1 and 2 are larger than those obtained from model 3, while the opposite is observed for depths greater than 3 km. Thus, depth plays a fundamental role, and r_{hyp} should be more effective than r_{epi}, particularly at short distances.

As for natural earthquakes (e.g., Douglas, 2007), it is of interest to investigate the effect of tectonic environment on ground motions from induced events. With this aim, we analyze the residual distributions for each of the six zones. For

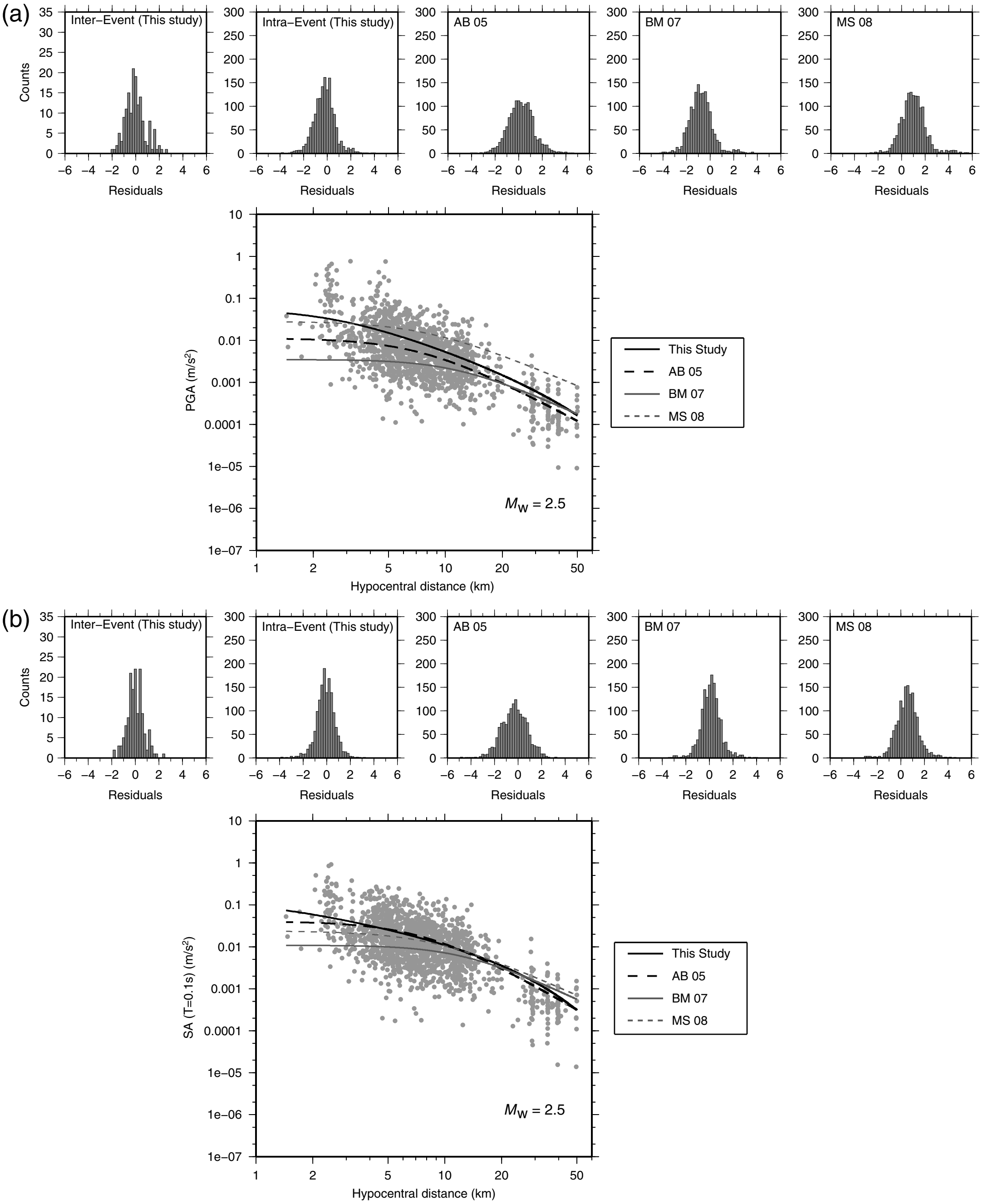


Figure 3. Residual distributions and comparisons between data and predictions (corrected data) include (a) PGA, (b) PSA (0.10 s), (c) PSA (0.20 s), and (d) PSA (0.50 s). The lower panels show the data from earthquakes with $2 \leq M_w \leq 3$ and curves corresponding to M_w 2.5, while the upper panels show the residual distributions using the GMPEs derived here and those of Ambraseys *et al.* (2005), Bommer *et al.* (2007), and Massa *et al.* (2008). Moreover, for the GMPEs from the present study, the residuals are separated into between-event and within-event components. (*Continued*)

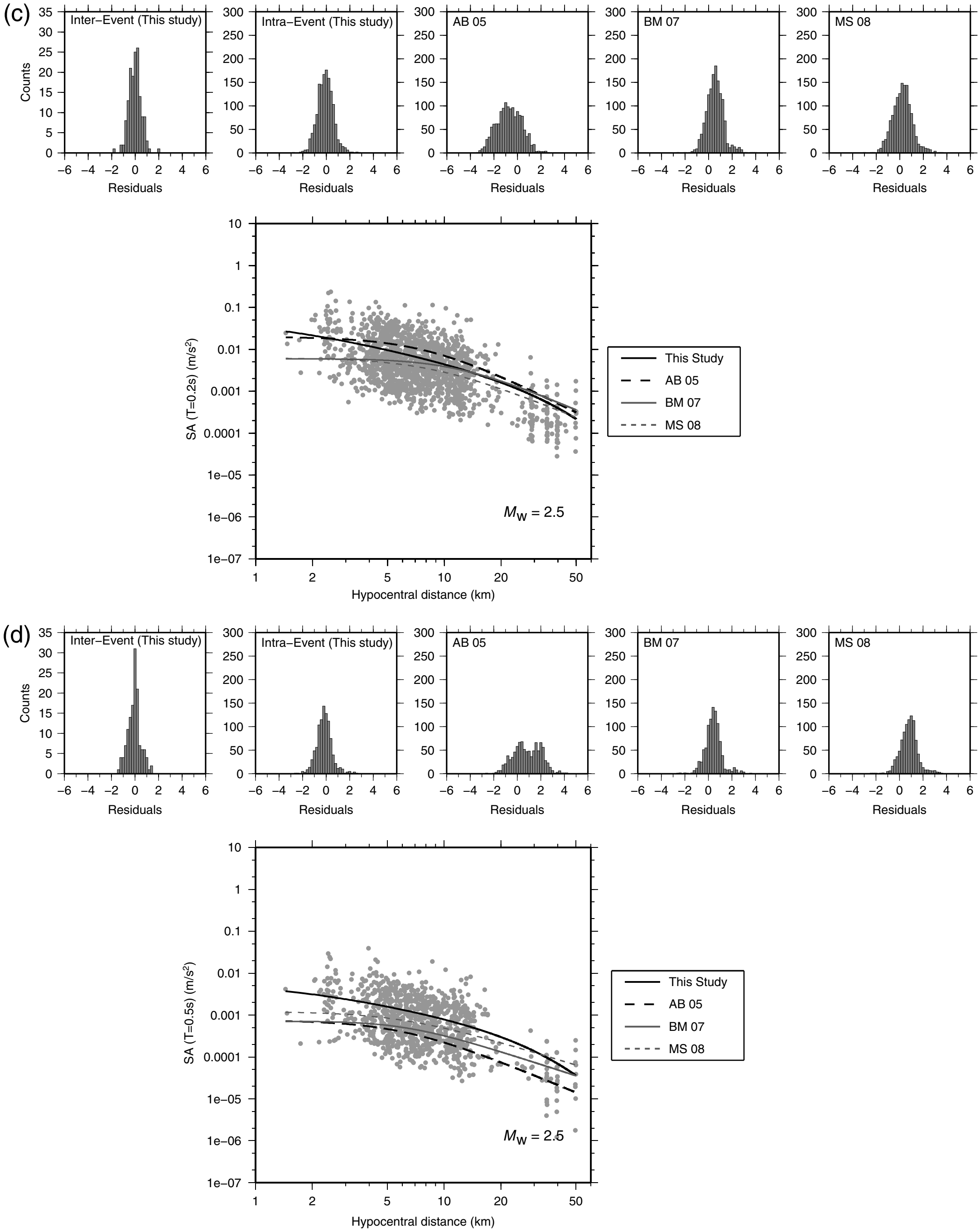


Figure 3. Continued.

Table 3

Coefficients for Models 2 (Equation 2) and 3 (Equation 3)

	Period	$a \pm s_a$	$b \pm s_b$	$c \pm s_c$	$h \pm s_h$	$d \pm s_d$	τ	ϕ	σ
Model 2	PGA 0.01	−7.198 ± 0.200	1.991 ± 0.085	−1.105 ± 0.062		−0.048 ± 0.005	0.731	1.080	1.304
	0.10	−8.079 ± 0.150	2.375 ± 0.063	−0.778 ± 0.050		−0.058 ± 0.004	0.589	0.787	0.983
	0.20	−10.118 ± 0.141	2.783 ± 0.057	−0.799 ± 0.048		−0.044 ± 0.004	0.554	0.658	0.860
	0.50	−12.829 ± 0.182	3.055 ± 0.070	−0.624 ± 0.060		−0.053 ± 0.005	0.573	0.696	0.902
	PGV	−10.846 ± 0.283	1.961 ± 0.121	−0.962 ± 0.088		−0.056 ± 0.007	1.031	1.536	1.850
Model 3	PGA 0.01	−7.397 ± 0.394	2.117 ± 0.081	−1.282 ± 0.181	3.173 ± 0.469	−0.030 ± 0.010	0.735	1.021	1.258
	0.10	−8.824 ± 0.212	2.481 ± 0.062	−0.620 ± 0.095	2.066 ± 0.487	−0.059 ± 0.006	0.592	0.766	0.968
	0.20	−10.570 ± 0.255	2.869 ± 0.057	−0.753 ± 0.118	2.675 ± 0.520	−0.040 ± 0.007	0.556	0.666	0.868
	0.50	−13.328 ± 0.257	3.104 ± 0.070	−0.497 ± 0.113	2.304 ± 0.761	−0.054 ± 0.007	0.574	0.704	0.908
	PGV	−9.672 ± 0.966	2.064 ± 0.123	−1.743 ± 0.432	4.961 ± 0.837	−0.015 ± 0.019	1.029	1.556	1.867

natural earthquakes, ground motions in different regional stress fields can be significantly different for the same magnitude and source-to-site distance (e.g., McGarr, 1984; Bommer *et al.* 2003; Convertito and Herrero, 2004). For induced seismicity, local stress conditions are mostly driven by field operations, which can reactivate existing faults or generate new ones with mechanisms different to those expected from the regional stress field (Oppenheimer, 1986; Li *et al.*, 2011). In Figure 5 we show the PGA residuals as a function of r_{hyp}, M_{w}, and depth. It can be noted that dominant contributions come from Geysers and Hengill. As a general comment, for all the models and all the considered structural periods, we do not observe particular correlations among the residuals and the three variables, M_{w}, r_{hyp}, and focal depth.

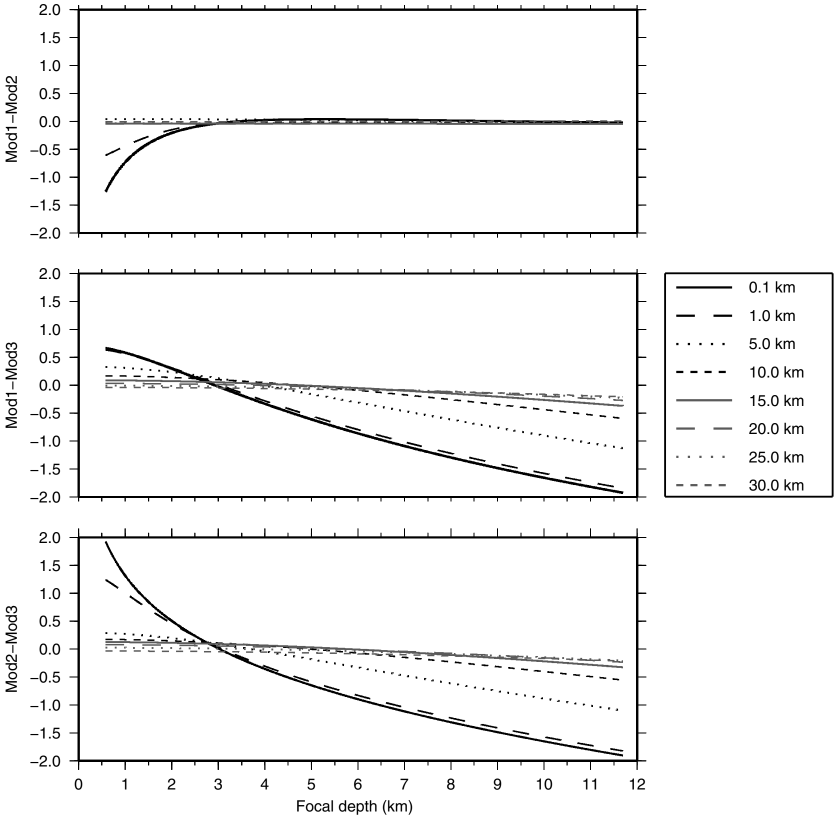


Figure 4. A comparison of predicted PGA of the three models as a function of focal depth.

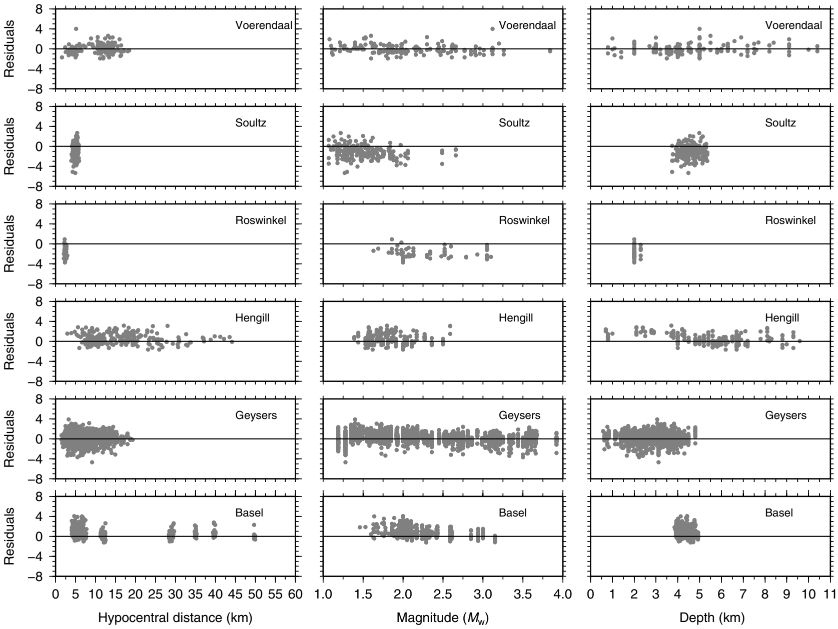


Figure 5. Residual analysis with respect to each area for PGA and model 1.

Development of Generic Stochastic Models

Because the GMPEs developed in the previous section may lead to erroneous predictions for $M_w > 3$ due to limited data, the purpose of this section is to develop stochastic models that are applicable close ($r_{hyp} < 50$ km) to shallow earth-
17 quakes of $1 \le M_w \le 5$. As discussed earlier in this article, this is outside the magnitude–distance range of the vast majority of published GMPEs. The empirical ground-motion models derived herein combine data from different sites and hence include within their aleatory variability (σ) large site-to-site variation (this is discussed in detail in a later
18 section). In keeping with state-of-the-art seismic-hazard assessments, it is preferable to explicitly separate epistemic uncertainty, which can be reduced through the collection of additional data, from true aleatory variability, which is randomness intrinsic to the model. The development of many (site-specific) stochastic models for geographical zones where geothermal exploitation is ongoing or likely in the future allows this separation to be made. The way in which these models could be used in EGS projects and
19 with what values of σ is discussed in subsequent sections.

Q and κ values were first obtained for each region and station, respectively, following the spectral fitting method detailed in Edwards *et al.* (2011). In the case of Geysers, Hengill, and Basel, sufficient data were available to define regional Q values: 199, 657, and 1575 (based on a velocity of $\beta = 3500$ m/s), respectively. However, in the case of Basel, we used $Q = 1200$, based on a study of a much larger Swiss dataset by Edwards *et al.* (2011). For cases in which Q was not determined (due to records from a narrow distance range), we assumed low attenuation ($Q = 1200$), such that the majority of attenuation was assigned to the station κ value. Station κ values were estimated based on these Q models such that the average path attenuation, t^*, is given by $t^* = \kappa + r_{hyp}/(Q\beta)$.

For selected datasets, we compared κ values computed using the methods of Anderson and Hough (1984) and Edwards *et al.* (2011) and found negligible differences. Stress (drop) parameters (Fig. 6) for the datasets were computed based on the definition assumed by the Stochastic-Method SIMulation (SMSIM) software as $\Delta\sigma = M_0[f_c/(0.4906\beta)]^3$ (Boore, 2003), with the seismic moment M_0, and the event–common source corner frequency f_c, obtained through a

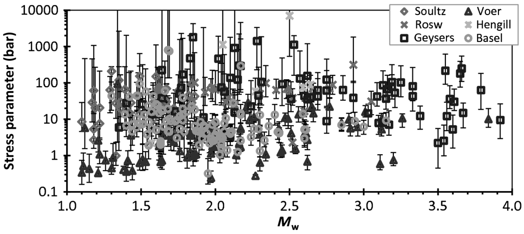


Figure 6. Observed stress-parameter values from the different datasets. Error bars indicate the uncertainty based on the range of possible f_c values within ±5% of the minimum misfit.

Table 4
A Summary of Stochastic Model Parameters

Parameter	Values	
Source type	(Brune 1970, 1971)	
Stress parameter	1, 10, 100 bar	
Q	200, 600, 1800	
Geometrical decay	$1/R$	
κ	0.005, 0.02, 0.04, 0.06 s	
Amplification function	Frequency	Amplification
	0.10	1.01
	0.20	1.03
	0.40	1.10
	0.80	1.31
	1.60	1.62
	3.20	1.75
	6.42	1.78
	12.0	1.79
	100	1.82
V_{S30}	1100 m/s	
Duration	$1/f_c + 0.05R_{\text{hyp}}$	
Density (ρ), reference velocity (β)	2800 kg/m^3, 3500 m/s	
Partition factor, radiation, free surface	0.71, 0.55, 2	

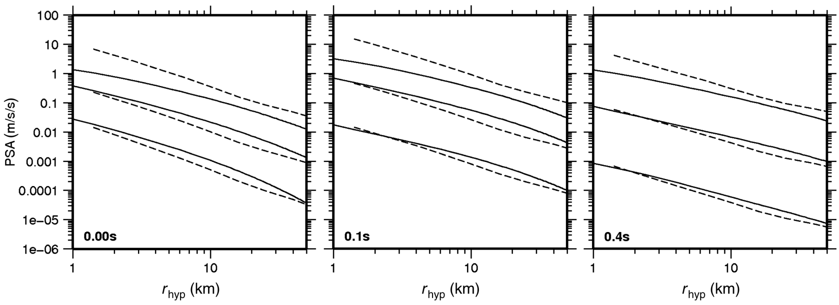
second spectral inversion of log–log spectra with fixed Q. There appears to be regional variation of the stress parameter, but this may also be due to the influence of trade-offs in the inversion and data limitations (such as available spectral bandwidth). To address such trade-offs, we include an uncertainty based on the range of possible f_c values within ±5% of the minimum misfit, although this is based on the assumption of known attenuation.

To capture the range of possible median ground motions for different stress-drop and attenuation scenarios, we simulated $3 \times 3 \times 4 = 36$ different stochastic models. The models were made using a combination of Q (200, 600, and 1800), Brune (1970, 1971) ω^2 stress parameter ($\Delta\sigma$, 1, 10, and 100 bar), and κ (0.005, 0.02, 0.04, and 0.06 s). The choice of these values was designed to cover the range of observed average values. In fact, the 36 models could be reduced to 12 if we consider that, for the magnitude range of interest and moderate κ and Q, the stress parameter is relatively insignificant (Douglas and Jousset, 2011). A known rock reference is imposed through the specification of the amplification and corresponding site-attenuation model of Poggi *et al.* (2011) and Edwards *et al.* (2011). The duration model is the theoretical model presented by Herrmann (1985), while geometrical spreading is assumed to follow $1/r$ decay. The stochastic model parameters are listed in Table 4.

While these models do not explicitly represent the region-specific data, we later describe a method for the production 20 of a mixture model to best describe a particular dataset and region. In the same way, it will be possible to produce a best-estimate model for a new region (given an expected stress parameter $\Delta\sigma$, κ, and Q value), which can then be dynamically updated when data become available. In these simulations we do not aim to address the issue of ground-motion variability, rather we are looking to cover epistemic uncertainty of potential median ground-motion models. Indeed, because of the complex interaction of stochastic-model parameters, it is difficult to justify the use of simulated ground motions for analyzing such variability (Rietbrock *et al.*, 2013). Instead, we provide measures of variability based on empirical analysis of data (see below). 21

A comparison of the generic model with $Q = 600$, $\kappa = 0.02$ s, and stress parameter 10 bar with a model for Switzerland (Edwards and Fäh, 2013) based on weak-motion data is shown in Figure 7. The comparison with the Swiss model shows they are quite similar, although the Swiss model has a magnitude-dependent stress parameter, which leads to a higher PSA for M_w 4.5 events. Furthermore, the Swiss model shows stronger geometrical decay in the first 20 km, then less decay at greater distances. 22

To make the developed stochastic model easier to use for hazard assessments, median ground motions (PGA, PSA, and PGV) for various magnitudes and distances are predicted using the stochastic model and random-vibration theory implemented in SMSIM (Boore, 2005) to which functions are fitted



35 **Figure 7.** A comparison of simulated PGA and PSA (gray lines) at 0.1 and 0.4 s to the model of Edwards and Fäh (2013) for the Swiss foreland for M_w 1.5, 3, and 4.5. For the comparison, the generic model with $Q = 600$ is used along with $\kappa = 0.02$ s and a $\Delta\sigma$ of 10 bar.

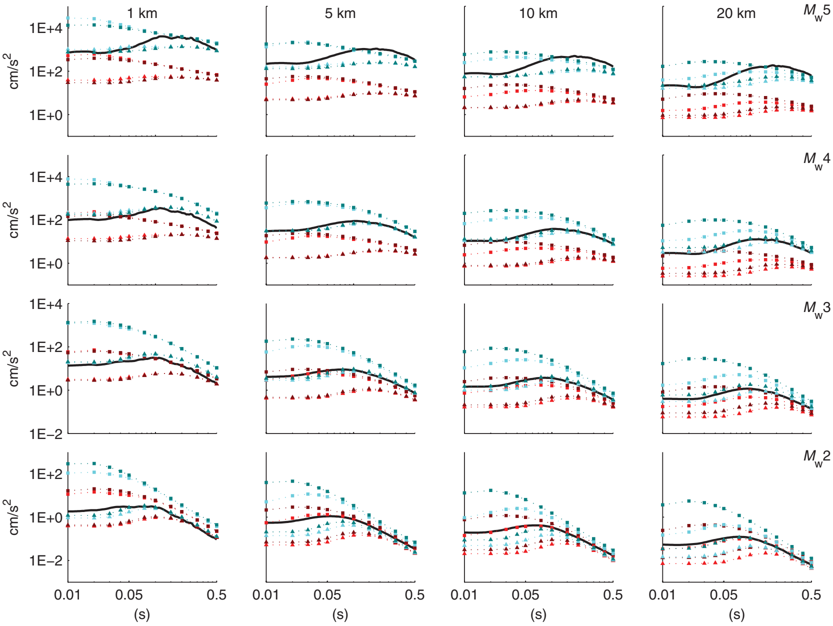


Figure 8. Trellis plot showing the predicted spectra (from the GMPEs derived from the simulated motions) for the six end-member models at M_w 2, 3, 4, and 5 and $r_\mathrm{hyp} = 1$, 5, 10, and 20 km. The models are red (1 bar) and blue (100 bar); light shade ($Q = 200$) and dark shade ($Q = 1800$); square ($\kappa = 0.005$ s) and triangle ($\kappa = 0.06$ s). Also shown is the empirical model 1 (black line).

by regression analysis. This is the same approach adopted by, for example, Atkinson and Boore (2006). To focus on induced events, simulations were performed at distances of $r_\mathrm{hyp} = 1$ to 50 km in 1 km intervals and magnitudes 1–5 in intervals of 0.25 for PGA, PGV, and PSA at 18 periods (0.01, 0.02, 0.03, 0.04, 0.05, 0.075, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.75, 1, 1.5, 2, 4 s, and 10 s). To capture the behavior of the predicted ground motions, various functional forms were tested, and it was found that the following simple functional form was sufficient:
23

$$\ln y = b_1 + b_2(M_\mathrm{w} - 3) + b_3(M_\mathrm{w} - 3)^2 + b_4(M_\mathrm{w} - 3)^3 + b_5 \ln(r_\mathrm{hyp} + b_h) + b_6(r_\mathrm{hyp} + b_h). \quad (4)$$

The standard deviations from using standard least-squares regression to fit this functional form to the simulated PSAs are all smaller (much smaller for $T > 0.1$ s) than 0.3, in terms of natural logarithms, thus showing that the functional
24 form is adequate. The coefficients of the 36 fitted models are listed in Ⓔ Table S1 of the electronic supplement.

Rather than present individual models for different $\Delta\sigma$, Q, and κ, an attempt was made to derive a single model using a functional form based on the theoretical dependency of the Fourier amplitude spectrum on these three parameters. However, the regression equations obtained by this approach were associated with large standard deviations, and predictions from the metamodel did not closely match the simulations. Therefore, we do not recommend these equations, and the coefficients are not given here. Atkinson and Boore (2006, 2007) propose an approximate procedure to adjust $\Delta\sigma$ in their GMPEs, derived using the stochastic model without listing coefficients for many different values of the stress parameter. These equations were tested to adjust the models for 10 bar up and down, but it was found that the predictions from the adjusted models for larger events ($M_\mathrm{w} \geq 4$) and $T > 0.1$ s did not match those from the individual GMPEs for $\Delta\sigma = 1$ bar and 100 bar. Therefore, the adjustments of Atkinson and Boore (2006, 2007) are not recommended for the models presented here.

Comparisons of the different generic models are shown in Figure 8. The predictions from the stochastic models with different Q diverge even close to the source, $r_\mathrm{hyp} < 10$ km. However, the effect of $\Delta\sigma$ is limited below M_w 3 and for 25
higher κ. This effect is initially surprising; however, as shown by Douglas and Jousset (2011), for example, high-frequency attenuation (Q and κ) means that changing $\Delta\sigma$ has a limited impact because it changes the plateau of the source spectrum for frequencies higher than the corner frequency, but these frequencies are then highly attenuated by the path and site. As the corner frequency decreases (magnitude increases), the impact of the high-frequency attenuation becomes less important and the models with different $\Delta\sigma$ diverge. For small events, κ is indeed controlling the

different values of PSA, particularly at high frequencies. However, for larger events (e.g., M_w 5), the stress parameter has a strong influence. Therefore, although it is not critical for fitting small events, the choice of stress parameter may have a strong impact on hazard if magnitudes approach 5. Empirical model 1 (black curve) is roughly in the middle of the generic stochastic models for $M_w \leq 4$, for which there are sufficient records to constrain the regression. It is, however, at the higher end of predictions for M_w 5, for which the regression is unconstrained by data and the assumption of the same (linear) magnitude scaling used for smaller events breaks down (Douglas and Jousset, 2011).

Because of the large epistemic uncertainty in the prediction of median ground motions from geothermal areas, the approach followed here is to propose a set of possible stochastic models that we suggest should be used as a basis of logic trees for seismic-hazard assessments associated with geothermal projects. At the beginning of a project when no ground-motion data are available from a site, the various models derived here could be assigned equal weight within the logic tree or weighted dependent on previous estimates of Q, $\Delta\sigma$, and κ in the region. Once seismograms become available from the local network installed as part of the project, the weights can be revised by comparing these data to the different models and assessing the likelihood of each model to be the correct one for the site. A simple example of this type of approach is presented in the next section.

Derivation of Weights for Stochastic Models

As an example of the possible application of the GMPEs derived above, we use the independent dataset from Campi Flegrei. These data are measurements of natural seismicity within a geothermally active zone rather than of induced events, but they are selected because they have many of the characteristics that EGS site records would have, such as small magnitudes, shallow focal depths, and short hypocentral distances. Only 55 records were analyzed from this site so the situation corresponds to either prestimulation monitoring of background seismicity or early on in the stimulation process. We analyzed 14 seismic events recorded by various networks during the bradiseismic crises of 1982–1984 and 8 events recorded during the smaller crisis of 2006. We selected earthquakes having a clear S wave with respect to the background noise recorded by at least three stations. The analyzed earthquakes have depths between 1 and 4 km below sea level, and almost all are located close to the center of the Campi Flegrei caldera. Spectra of ground displacement of a 2.56 s time window containing the S-wave first arrival (starting 0.3 s before the S-wave arrival time) were fitted by using a theoretical ω^2 model corrected for both $Q = 125$ and $\kappa = 0.015$ s (De Natale *et al.*, 1987) to estimate low-frequency levels, Ω_0, and corner frequencies, f_c. The observed spectra were also corrected for the site functions found by Tramelli *et al.* (2010).

Seismic moments were obtained from the spectra using this relationship (Aki and Richards, 2002):

$$M_0 = \frac{4\pi\rho V_S^3 r_{\text{hyp}}\Omega_0}{F R_\theta}, \qquad (5)$$

with $\rho = 2000$ kg/m^3, $V_S = 1700$ m/s, F is the free-surface correction, and the average S-wave radiation-pattern coefficient, R_θ, is 0.3 (De Natale *et al.*, 1987). The stress drop, $\Delta\sigma$, was calculated using the Brune (1970, 1971) model, $\Delta\sigma = 0.44 M_0/r^3$, where the source radius r is given by $0.37 V_S/f_c$. The stress drops calculated for the earthquakes of the 1982–1984 crises are close to 5 bar, as previously found by De Natale *et al.* (1987), while the values found for the 2006 earthquakes are a little higher, as shown in Table 5.

Based on the values of Q (125) and κ (0.015 s) estimated by De Natale *et al.* (1987) and the value of $\Delta\sigma$ (5 bar) found here, it could be argued to give highest weight within a seismic-hazard assessment for this area to the GMPEs for $Q = 200$, $\kappa = 0.02$ s, and $\Delta\sigma = 1$ bar and 10 bar, which are the nearest 26
available stochastic models of the 36 derived here. However, given the large epistemic uncertainties on these estimates of Q, κ, and $\Delta\sigma$ for Campi Flegrei, a more observational-based approach may be preferred due to the limited data.

The magnitudes of the available data only covers the range of $0.4 \leq M_w \leq 2.1$, and over this range the effect of $\Delta\sigma$ is limited. Consequently, we only consider a single stress parameter (1 bar). We also exclude the models using $\kappa = 0.005$ s, which is unrealistic for Campi Flegrei sites. This means that the 36 potential GMPEs are reduced to $3 \times 3 = 9$ models, which include $Q = 200$, 600, and 1800 and $\kappa = 0.02$, 0.04, and 0.06 s. Figure 9 compares the predicted and observed PSA (0.1 s) and PSA (0.5 s) for data within $0.5 M_w$ units of the mean magnitude of the analyzed data (M_w 1.4), which shows the large inherent variability in observed ground-motion data from small earthquakes and the difficulty in preferring certain models over others. Nevertheless, the lower group of stochastic models (corresponding to $\kappa = 0.06$ s) provides a better fit than the other GMPEs. Consequently, slightly higher weights for these GMPEs could be appropriate, but the other models cannot be excluded from the logic tree because of the high-epistemic uncertainty when conducting seismic-hazard assessments using limited observations (such as this case). Figure 9 also illustrates that empirical model 1 and the stochastic models predict similar ground motions for this magnitude and range of distances. In practice, ground-motion data from a local monitoring network installed as part of an EGS project would allow continual updating of the weights assigned to each of the considered GMPEs as the epistemic uncertainty in the median ground motions decreases as more data are recorded.

Aleatory Variability

One of the most active areas of engineering seismology research in the past decade is in the understanding and characterization of aleatory variability of earthquake

Table 5

Corner Frequencies, Stress Drops, Moments, Magnitudes, and Locations of the Analyzed Earthquakes

Event	f_c (Hz)	$\Delta\sigma$ (bar)	M_0 (N·m)	M_w	Northing (m)	Easting (m)	Depth (m)
	18.07	6.02	3.91×10^{10}	1.0	4519785	426892	1850
241	9.20	1.83	3.64×10^{11}	1.6	4520998	425892	2450
253	8.53	5.05	1.56×10^{12}	2.1	4521374	427104	2950
2843	15.20	4.48	4.49×10^{11}	1.7	4521168	429309	3950
2845	12.67	10.90	6.49×10^{11}	1.8	4521177	428311	2650
2855	9.60	5.40	7.03×10^{11}	1.8	4521368	427709	1950
2884	10.13	1.58	3.78×10^{10}	1.0	4519280	427379	1250
2910	11.04	8.05	9.94×10^{11}	1.9	4516482	427858	2950
2958	15.27	7.26	2.10×10^{11}	1.5	4520695	426493	2050
2977	10.17	2.53	4.68×10^{10}	1.1	4519377	426986	1250
2981	14.20	0.70	5.18×10^{9}	0.4	4519284	426985	1250
2983	12.00	1.96	2.03×10^{10}	0.8	4519376	427085	1150
2988	10.13	0.45	2.74×10^{10}	0.9	4520213	424788	1950
2992	8.42	0.99	1.03×10^{11}	1.3	4519777	427693	1750
6001	16.80	5.41	3.95×10^{10}	1.0	4519673	426947	1600
6008	20.00	24.60	3.36×10^{11}	1.6	4519770	428382	1950
6011	19.67	9.26	1.28×10^{11}	1.3	4519922	427976	1570
6012	17.40	15.08	2.33×10^{11}	1.5	4519625	428141	1800
6013	19.80	6.84	8.62×10^{10}	1.2	4520220	427838	1600
6019	14.47	51.72	5.13×10^{11}	1.7	4521625	427922	1190
6020	17.95	50.58	1.84×10^{11}	1.5	4521403	427906	920
6023	14.20	14.99	5.40×10^{11}	1.8	4519936	428496	1760

The first 14 rows indicate analyzed earthquakes from 1982 to 1984; the last 8 rows show those from 2006.

shaking, commonly summarized by the standard deviation (σ) of the logarithm of a ground-motion parameter such as PGA (e.g., Strasser *et al.*, 2009, and the references therein). It is important that the σ used within a probabilistic seismic-hazard assessment correctly captures the true variability in the ground motion and that it is not unrealistically small or large.

One cause of overestimated σ is that the independent parameters (metadata), such as magnitude and distance, are inaccurate. This would lead to a mapping of these uncertainties into the computed σ. This has been proposed as a possible reason for the common observation that ground motions from small earthquakes seem to be more variable than those from large earthquakes. This is because metadata of small earthquakes are likely to be poorer than those of large earthquakes because national and international seismic networks have difficulty locating and quantifying the size of small shocks (e.g., Youngs *et al.*, 1995). The analyses conducted here should provide estimates of σ largely free of contributions from inaccurate metadata because the records used all come from locations well covered by high-quality local seismic networks. As an example, Jousset *et al.* (2011) relocalized local earthquakes as part of a tomographic study for the Hengill area. The locations reported by IMO (used here) and Jousset *et al.* (2011) are generally within
27 0.5 km (horizontally) and within 1 km (vertically) of one another. Tremors within geothermal reservoirs are probably even more accurately located because of dense local arrays. In addition, magnitudes were all carefully recomputed for this study using a consistent method.

A comparison between the aleatory variabilities of various recent GMPEs and those for the empirical model 1 is shown in Figure 10. One clear difference between the variabilities of recent GMPEs for moderate and large earthquakes and model 1 and the mining-related GMPE of McGarr and Fletcher (2005) is the strong period dependency of these two models with a peak in σ for very short periods and a rapid decrease as period increases. As shown in Figure 8,

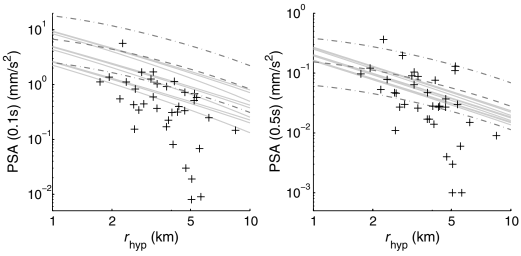


Figure 9. Comparisons of predicted ground motions by nine stochastic GMPEs and empirical model 1 (corrected for site effects) and observed PSA (0.1 s) and PSA (0.5 s) for data from Campi Flegrei. Median PSAs are plotted for the stochastic (solid light gray lines) and empirical (dashed dark gray lines) models, and the 16th and 84th percentiles (dash-dotted dark gray lines) for the empirical GMPEs.

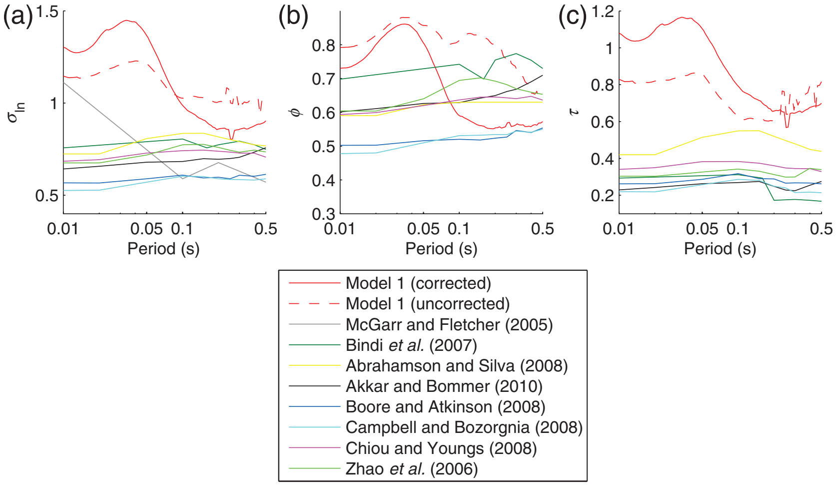


Figure 10. A comparison of aleatory variabilities of various recent GMPEs for an M_w 3 earthquake and a rock site ($V_{S30} = 1000$ m/s) using natural logarithms: (a) total σ; (b) within-event ϕ; and (c) between-event τ. A comparison of aleatory variabilities (left, total σ; middle, within-event ϕ; and right, between-event τ) of various recent GMPEs for an M_w 3 earthquake and a rock site ($V_{S30} = 1000$ m/s) using natural logarithms. Six commonly used GMPEs for active crustal regions (Zhao *et al.*, 2006; Abrahamson and Silva, 2008; Boore and Atkinson, 2008; Campbell and Bozorgnia, 2008; Chiou and Youngs, 2008; Akkar and Bommer, 2010) are considered, along with one model for natural seismicity covering a similar magnitude-distance range to the data used here (Bindi *et al.*, 2007), one model for mining-induced seismicity (McGarr and Fletcher, 2005; the authors of this model only report the total σ of their model), and the sigmas for model 1 (uncorrected and corrected for site response).

there is much variation in predictions from the derived stochastic models at very short periods, which is being mapped in empirical models into the aleatory variability rather than being considered as epistemic uncertainty in the values of $\Delta\sigma$, κ, and Q. In addition, the sigmas of these two GMPEs are much higher than those of the other considered GMPEs. These observations can be related to much higher (more than twice for PGA) between-event variabilities (τ). The within-event variabilities (ϕ) from all models are comparable. This is despite the finding of Douglas and Jousset (2011) that variations in near-surface attenuation (characterized by κ) lead to larger variations in high-frequency ground motions from small earthquakes than they do in large earthquakes due to the interaction between source-corner frequency and attenuation. Such an effect would affect ϕ, which could be an explanation for slightly higher values at short periods for model 1. Although the reduction in ϕ from using the site-corrected data is evident, this site correction leads to higher τ (due to the reduction of the local attenuation, κ, to that consistent with a hard-rock site). Because κ acts as a low-pass filter, this reduction recovers a previously unseen source variability.

Are the much higher τ values from the data analyzed here realistic? Or should a different model for τ be developed or adopted, such as one of the Next Generation Attenuation (NGA) models (Abrahamson *et al.*, 2008)? One reason for much larger short-period τ is greater variability in $\Delta\sigma$; another is variations in focal depth, which would have a stronger effect for small shocks than for large events. Any variation among records from a single station of induced earthquakes of similar magnitudes in a well-constrained zone is almost entirely due to differences in the source (e.g., mechanism and stress drop). Many records from Soultz and Basel fall into this category. These data allow estimates of this component of τ to be made. The τ computed here is based on regression analysis of data from six different zones, and future EGS sites will likely only be affected by a single source, which should be less variable than multiple sources. Therefore, zone-specific τ have been computed for Soultz and Basel, which are well-defined induced seismicity sources with sufficient data for robust statistics. These are computed from the between-event residuals with respect to model 1 for these two zones. These zone-specific τ_{ZS} values are compared to the overall τ in Figure 11. τ_{Soultz} and τ_{Basel} are much lower than the original τ and are also similar to τ of previously published GMPEs (Fig. 10). This implies that the high τ values obtained by regressing on data from six zones is due in large part to the aleatory variability in the earthquake sources between zones that should not be accounted for in site-specific hazard assessments for EGS. However, if this component of aleatory variability is removed from the hazard assessment, then it requires that the epistemic uncertainty in the assessment of median ground motion for the considered EGS be correctly accounted for (see Derivation of Weights for Stochastic Models). 28

By studying many records from California, Atkinson (2006) concludes that in the situation of a single station recording earthquakes on a single fault, the associated

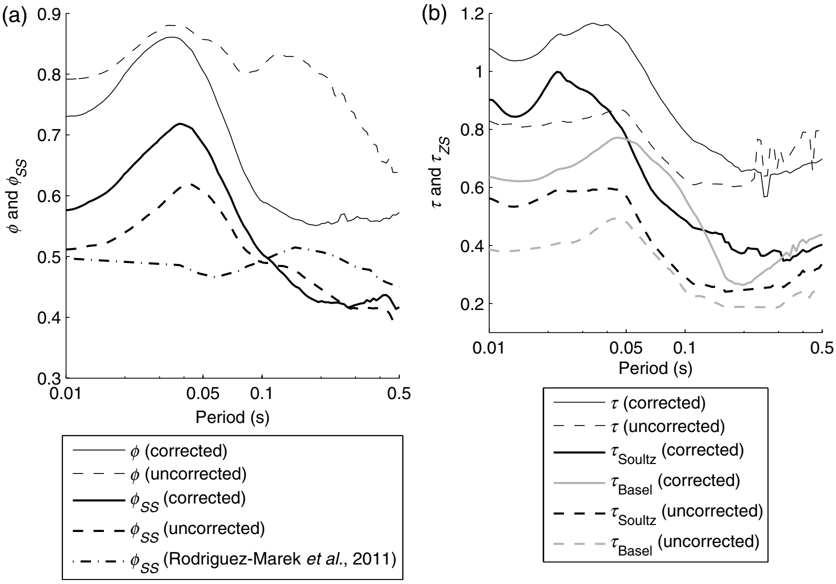


Figure 11. A comparison of (a) overall and single-station within-event variabilities (ϕ and ϕ_{SS}, respectively), and (b) overall and zone-specific between-event variabilities (τ and τ_{ZS}, respectively), for model 1 (uncorrected and corrected for site response).

variability of the ground motions (characterized in terms of standard deviation) is about 60% of its value for many sites recording earthquakes on various faults. Geothermal power projects offer an ideal situation in which single-station, single-source adjustments of σ could and should be made. In areas of low and moderate seismicity (the case for most current EGS projects; e.g., Soultz) the ground motions of induced seismicity close to an EGS are likely to dominate those from natural seismicity for high probabilities of exceedance (short return periods) and hence the geothermal reservoir could be considered as a single source. For regulatory and production purposes, it is likely that a dense monitoring network will be installed close to the EGS. Consequently, it would be possible to characterize the local site conditions and estimate the site-correction factors for locations affected by the induced seismicity. Therefore, the aleatory variability will be lower without a consequential increase in the epistemic uncertainty.

Single-station ϕ, (also called $\phi_{SS,S}$ by Rodriguez-Marek *et al.*, 2011) was computed using the approach of Atkinson (2006) and model 1 (both data uncorrected and corrected for site response) for all 62 stations recording 10 or more earthquakes in all considered zones. As previously found by Atkinson (2006) and Rodriguez-Marek *et al.* (2011), $\phi_{SS,S}$ varies considerably from one station to the next. The mean single-station ϕ (ϕ_{SS}) is plotted on Figure 11 alongside the estimate of this variability by Rodriguez-Marek *et al.* (2011) for Japanese surface stations and the ϕ found for model 1. As expected, the removal of variability coming from mixing sites leads to a significant drop in ϕ and values of ϕ_{SS} similar to those reported by Rodriguez-Marek *et al.* (2011). Stations recording the induced seismicity at Soultz and Basel (that are close to the ideal single-station, single-fault situation) do not show lower values of ϕ_{SS}. Using these values of ϕ_{SS} for site-specific hazard assessments requires that the epistemic uncertainty in the estimation of the median-site correction for considered locations be accounted for (see Derivation of Weights for Stochastic Models). Ⓔ Table S3 of the electronic supplement provides the estimates of ϕ_{SS}, τ_{Soultz}, and τ_{Basel} derived here.

Conclusions

This article has investigated ground motions generated by induced earthquakes and those associated with EGSs in particular. We sought to answer the question of whether ground motions from induced earthquakes are significantly different than those from natural earthquakes using various statistical techniques. We developed stochastic models and subsequently GMPEs to estimate earthquake shaking in terms of PGA, PSA, and PGV that are valid from M_{w} 1 to 5. We also developed a homogenized database of PSA and corresponding metadata from several sites (Basel, Geysers, Hengill, Roswinkel, Soultz, and Voerendaal). To account for varying site conditions, a correction for site-specific amplification and attenuation was applied to the data. We showed that this resulted in a reduction in σ over the mid- to long-period range but an increase in σ in the short-period range (i.e.,

<0.1 s). This is because site attenuation effectively hides a significant proportion of source variability for microearthquakes. As we brought the sites to a relatively hard-rock condition ($\kappa = 0.016$ s), this source variability was reintroduced, as seen in the increase in τ values. Clearly, subsequent correction to a suitable surface condition will tend to reduce this uncertainty in high-frequency spectral ordinates as high damping is typically applied in the upper layers. Nevertheless, hard-rock predictions are usually sought in order to predict bedrock-referenced ground motions for hazard assessments. This high τ is, therefore, an important feature to consider. A further observation was that source effects varied considerably from region to region. Combining the regions to produce a common τ leads to unrealistically high values. We infer that the high, combined τ is due to the mapping of epistemic uncertainty into the aleatory component. Strong variation in mean stress drop was observed across the study regions, which is a likely source of regional variability. However, in hazard assessments, this should be treated as epistemic uncertainty, which may be reduced in the case of sufficient observations. Restricting the computation of τ to the region-specific case, we instead observe uncertainty in ground-motion prediction due to variations in stress drop, which may not be predicted. Our uncertainty model is thus constructed using a combination of sources. Epistemic uncertainty is covered by implementing various prediction models, which may be weighted or eliminated according to expert judgment or observations. Aleatory variability is covered in terms of single-site, single-region σ, which is comprised of ϕ_{SS} and single-region τ.

Given the significant site-to-site variability in geothermal events highlighted in this study, the application of a unique empirical GMPE, as developed as part of this study, would lead to bias, in addition to overestimation of ground-motion variability at a specific site. Effectively, the site-to-site variability (ϕ) includes a significant component of epistemic uncertainty. Neglecting this component in any hazard analysis would lead not only to bias, but also to the possibility of the residual misfit lying outside the predicted uncertainty. Nevertheless, to a first approximation, in light of no other seismological information, the use of the empirical GMPEs with total σ ensures that the possible range of ground motions is covered.

On the other hand, given a database of site-specific earthquake recordings, or the potential to update our knowledge (reduce epistemic uncertainty) with time, we are presented with the opportunity to better represent the observed ground motions, possibly based on extrapolation of physical parameters. To this end, a logic-tree approach is commonly taken in order to account for epistemic uncertainty in probabilistic seismic-hazard assessment, with weights assigned related to the belief of the analyst that, for example, a given ground-motion model is the correct one. Given the range of the 36 stochastic ground-motion models presented in this study, we can foresee that the weighting of such models is (at least partially) determined through either residual or likelihood analyses (Scherbaum *et al.*, 2009; Kale and Akkar, 2013). Alternatively, spectral parameters (such as $\Delta\sigma$, Q, and κ) could be determined from available data to better select relevant models. In this case, the site-to-site component of variability should be appropriately reduced (Rodriguez-Marek *et al.*, 2011) because this is accounted for through different stochastic models to cover the epistemic uncertainty.

For applications of the stochastic GMPEs, we recommend at first to consider the 36 different models to account for epistemic uncertainty in the median with between-event variability τ equal to the average of the values of τ_{ZS} derived for Soultz and Basel and within-event variability ϕ equal to the values of ϕ_{SS} computed here. Also we recommend that data recorded by local seismic networks (which will often be installed as part of an EGS project) be used to subsequently winnow and weight the stochastic GMPEs and potentially adjust the associated models for τ and ϕ, which are likely to be site specific. However, the adjustment of these variabilities requires a considerable number of records for stable estimates to be made, as does the assessment of weights for the GMPEs.

The GMPEs developed here could be used within probabilistic seismic-hazard assessment accounting for induced seismicity as performed, for example, by van Eck *et al.* (2006) for induced seismicity related to gas extraction in the Netherlands and by Convertito *et al.* (2012) for geothermally induced seismicity at Geysers. In addition to being applicable to ground motions associated with induced seismicity from geothermal power production, the GMPEs presented here may also be applicable for hazard assessments of geological carbon dioxide storage projects, which involve the injection of high-pressure fluids into geological structures, but their value for this situation has not yet been evaluated.

Data and Resources

The data from Basel were provided by Geo Explorers Ltd, the Swiss Seismological Service, and the Landeserdbebendienst Baden-Württemberg. Campi Flegrei data were preprocessed in the framework of the coordination project, "Integrated Seismic Methods Applied to the Investigations of the Active Volcano Structure: An Application to the Campi Flegrei Caldera," launched by the Dipartimento della Protezione Civile of the Istituto Nazionale di Geofisica e Vulcanologia (INGV) during the 2000–2004 National Framework **30** Program, coordinated by the National Group of Volcanology (GNV) of INGV (see Capuano *et al.*, 2006). Waveforms from the Lawrence Berkeley National Laboratory Geysers/Calpine seismic network (BG) and the related earthquake catalog have been retrieved from the publicly accessible website of the Northern California Earthquake Data Center (NCEDC, www.ncedc.org, last accessed May 2012). The data from Hengill were recorded by a temporary network installed within the framework of the I-GET FP6 project (Jousset and François, 2006), which were provided by Philippe Jousset. Data from Voerendaal and Roswinkel were recorded by the Seismology Division of the Koninklijk Nederlands Meteorologisch Instituut. The data from Soultz were recorded by a permanent network installed by École et Observatoire des Sciences de la Terre (EOST) of the University of Strasbourg and were provided by Michel Frogneux (EOST). Information on current geothermal projects was obtained from the website of the International Geothermal Association (www.geothermal-energy.org, last accessed December 2012).

Acknowledgments

This study was partially funded by the Geothermal Engineering Integrating Mitigation of Induced Seismicity in Reservoirs (GEISER) project under contract 241321 of the European Commission Research Seventh
31 Framework Programme (FP7). We thank Dave Boore for help conducting the analysis of the effect of noise on response spectra. We thank Philippe Jousset for providing the data from Hengill, Michel Frogneux for the data from Soultz, and the Lawrence Berkeley National Laboratory Geysers/Calpine seismic network and the Northern California Earthquake Data Center for the data from The Geysers. Xavier Rachez, Georgia Cua, Giuseppe De Natale, Bernard Dost, and Falko Bethmann helped with various aspects of this study. Finally, we thank S. Lasocki and an anonymous reviewer for their constructive comments on a previous version of this article.

References

Abrahamson, N., and W. Silva (2008). Summary of the Abrahamson & Silva NGA ground-motion relations. *Earthq. Spectra* **24,** no. 1, 67–97, doi: 10.1193/1.2924360.

Abrahamson, N., G. Atkinson, D. Boore, Y. Bozorgnia, K. Campbell, B. Chiou, I. M. Idriss, W. Silva, and R. Youngs (2008). Comparisons of the NGA ground-motion relations, *Earthq. Spectra* **24,** no. 1, 45–66, doi: 10.1193/1.2924363.

Abrahamson, N. A., and R. R. Youngs (1992). A stable algorithm for regression analyses using the random effects model, *Bull. Seismol. Soc. Am.* **82,** no. 1, 505–510.

Aki, K., and P. G. Richards (2002). *Quantitative Seismology*, University Science Books, Sausalito, California, 700 pp.

Akkar, S., and J. J. Bommer (2010). Empirical equations for the prediction of PGA, PGV and spectral accelerations in Europe, the Mediterranean region and the Middle East, *Seismol. Res. Lett.* **81,** no. 2, 195–206.

Al Atik, L., N. Abrahamson, J. J. Bommer, F. Scherbaum, F. Cotton, and N. Kuehn (2010). The variability of ground-motion prediction models and its components, *Seismol. Res. Lett.* **81,** no. 5, 794–801, doi: 10.1785/gssrl.81.5.794.

Ambraseys, N. N., J. Douglas, S. K. Sarma, and P. M. Smit (2005). Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from Europe and the Middle East: Horizontal peak ground acceleration and spectral acceleration, *Bull. Earthq. Eng.* **3,** no. 1, 1–53, doi: 10.1007/s10518-005-0183-0.

Anderson, J. G., and S. E. Hough (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bull. Seismol. Soc. Am.* **74,** no. 5, 1969–1993.

Aster, R., R. P. Meyer, G. De Natale, A. Zollo, M. Martini, E. Del Pezzo, G. Iannaccone, and R. Scarpa (1989). Seismic investigation of the Campi Flegrei: A synthesis and summary of results, in *Volcanic Seismology*, Volcanic Series III, K. Aki, P. Gasparini, and R. Scarpa (Editors), Springer-Verlag, Amsterdam, 462–483.

Atkinson, G. M (2006). Single-station sigma, *Bull. Seismol. Soc. Am.* **96,** no. 2, 446–455, doi: 10.1785/0120050137.

Atkinson, G. M., and D. M. Boore (2006). Earthquake ground-motion prediction equations for eastern North America, *Bull. Seismol. Soc. Am.* **96,** no. 6, 2181–2205, doi: 10.1785/0120050245.

Atkinson, G. M., and D. M. Boore (2007) Erratum to "Earthquake ground-motion prediction equations for eastern North America", *Bull. Seismol. Soc. Am.* **97,** no. 3, 1032, doi: 10.1785/0120070023.

Bindi, D., S. Parolai, H. Grosser, C. Milkereit, and E. Durukal (2007). Empirical ground-motion prediction equations for northwestern Turkey using the aftershocks of the 1999 Kocaeli earthquake, *Geophys. Res. Lett.* **34,** L08305, doi: 10.1029/2007GL029222.

Bommer, J. J., and J. E. Alarcón (2006). The prediction and use of peak ground velocity, *J. Earthq. Eng.* **10,** no. 1, 1–31.

Bommer, J. J., J. Douglas, and F. O. Strasser (2003). Style-of-faulting in ground-motion prediction equations, *Bull. Earthq. Eng.* **1,** no. 2, 71–203.

Bommer, J. J., S. Oates, J. M. Cepeda, C. Lindholm, J. Bird, R. Torres, G. Marroqun, and J. Rivas (2006). Control of hazard due to seismicity induced by a hot fractured rock geothermal project, *Eng. Geol.* **83,** 287–306, doi: 10.1016/j.enggeo.2005.11.002.

Bommer, J. J., P. J. Stafford, J. E. Alarcón, and S. Akkar (2007). The influence of magnitude range on empirical ground-motion prediction, *Bull. Seismol. Soc. Am.* **97,** no. 6, 2152–2170, doi: 10.1785/0120070081.

Boore, D. M. (2003). Simulation of ground motion using the stochastic method, *Pure Appl. Geophys.* **160,** no. 3–4, 635–676, doi: 10.1007/PL00012553.

Boore, D. M. (2005). SMSIM—Fortran programs for simulating ground motions from earthquakes: Version 2.3—A revision of OFR 96-80-A, *U.S. Geol. Surv. Open-File Rept. 00-509*, modified version, describing the program as of 15 August 2005 (Version 2.30).

Boore, D. M., and G. M. Atkinson (2008). Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01 s and 10.0 s, *Earthq. Spectra* **24,** no. 1, 99–138, doi: 10.1193/1.2830434.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes, *J. Geophys. Res.* **75,** no. 26, 4997–5009.

Brune, J. N. (1971) Correction, *J. Geophys. Res.* **76,** no. 20, 5002.

Campbell, K. W., and Y. Bozorgnia (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 s to 10 s, *Earthq. Spectra* **24,** no. 1, 139–171, doi: 10.1193/1.2857546.

Capuano, P., G. Russo, T. Vanorio, R. Prevete, E. Auger, M. Bonagura, G. Caielli, V. Convertito, N. Damiano, L. D'Auria, A. Emolo, L. Lovisa, and M. Moretti (2006). 1984 Campi Flegrei seismic waveforms compilation, in *Geophysical Exploration of the Campi Flegrei (Southern Italy) Caldera Interiors: Data, Methods and Results*, A. Zollo, P. Capuano, and M. Corciulo (Editors), Doppia Voce, Naples, Italy, 15–24.

Charlety, J., N. Cuenot, L. Dorbath, C. Dorbath, H. Haessler, and M. Frogneux (2007). Large earthquakes during hydraulic stimulations at the geothermal site of Soultz-sous-Forets, *Int. J. Rock Mech. Min. Sci.* **44,** no. 8, 1091–1105, doi: 10.1016/j.ijrmms.2007.06.003.

Chiou, B., R. Youngs, N. Abrahamson, and K. Addo (2010). Ground-motion attenuation model for small-to-moderate shallow crustal earthquakes in California and its implications on regionalization of ground-motion prediction models, *Earthq. Spectra* **26,** no. 4, 907–926, doi: 10.1193/1.3479930.

Chiou, B. S. J., and R. R. Youngs (2008). An NGA model for the average horizontal component of peak ground motion and response spectra, *Earthq. Spectra* **24,** no. 1, 173–215, doi: 10.1193/1.2894832.

Convertito, V., and A. Herrero (2004). Influence of focal mechanism in probabilistic seismic hazard analysis, *Bull. Seismol. Soc. Am.* **94,** no. 6, 2124–2136, doi: 10.1785/0120040036.

Convertito, V., N. Maercklin, N. Sharma, and A. Zollo (2012). From induced seismicity to direct time-dependent seismic hazard, *Bull. Seismol. Soc. Am.* **102,** no. 6, 2563–2573, doi: 10.1785/0120120036.

De Natale, G., and A. Zollo (1986). Statistical analysis and clustering features of the Phlegraean Fields earthquake sequence, May '83–May '84, *Bull. Seismol. Soc. Am.* **76,** no. 801–814.

De Natale, G., G. Iannaccone, M. Martini, and A. Zollo (1987). Seismic sources and attenuation properties at Campi Flegrei volcanic area, *Pure Appl. Geophys.* **125,** no. 6, 883–917.

De Natale, G., E. Faccioli, and A. Zollo (1988). Scaling of peak ground motions from digital recordings of small earthquakes at Campi Flegrei, southern Italy, *Pure Appl. Geophys.* **126,** no. 1, 37–53.

De Natale, G., C. Troise, F. Pingue, G. Mastrolorenzo, L. Pappalardo, M. Battaglia, and E. Boschi (2006). The Campi Flegrei caldera: Unrest mechanisms and hazards, in *Mechanisms of Activity and Unrest at Large Calderas*, Vol. 269, C. Troise, G. De Natale, and C. R. J. Kilburn (Editors), Geological Society of London, 25–45.

Deino, A. L., G. Orsi, M. Piochi, and S. de Vita (2004). The age of the Neapolitan Yellow Tuff caldera-forming eruption (Campi Flegrei caldera—Italy) assessed by 40Ar/39Ar dating method, *J. Volcanol. Geoth. Res.* **133,** no. 1–4, 157–170.

Delavaud, E., F. Cotton, S. Akkar, F. Scherbaum, L. Danciu, C. Beauval, S. Drouet, J. Douglas, R. Basili, M. A. Sandikkaya, M. Segou, E. Faccioli, and N. Theodoulidis (2012). Toward a ground-motion logic tree for probabilistic seismic hazard assessments in Europe, *J. Seismol.* **16,** no. 3, 451–473, doi: 10.1007/s10950-012-9281-z.

Dost, B., and H. W. Haak (2007). Natural and induced seismicity, in *Geology of the Netherlands*, T. E. Wong and J. de Jager (Editors), Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands, 223–239.

Douglas, J. (2004a). An investigation of analysis of variance as a tool for exploring regional differences in strong ground motions, *J. Seismol.* **8,** no. 4, 485–496.

32 Douglas, J. (2004b). Use of analysis of variance for the investigation of regional dependence of strong ground motions, in *Proc. of the Thirteenth World Conf. on Earthquake Engineering*, Paper Number 29.

Douglas, J. (2007). On the regional dependence of earthquake response spectra, *ISET J. Earthq. Technol.* **44,** no. 1, 71–99.

Douglas, J. (2011). Ground-motion prediction equations 1964–2010, *Final Rept.*, BRGM/RP-59356-FR, Orléans, France, http://www.brgm.fr/publication/rapportpublic.jsp (last accessed, December 2012)..

Douglas, J., and H. Aochi (2008). A survey of techniques for predicting earthquake ground motions for engineering purposes, *Surv. Geophys.* **29,** no. 3, 187–220, doi: 10.1007/s10712-008-9046-y.

Douglas, J., and D. M. Boore (2011). High-frequency filtering of strong-motion records, *Bull. Earthq. Eng.* **9,** no. 2, 395–409, doi: 10.1007/s10518-010-9208-4.

Douglas, J., and P. Jousset (2011). Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes, *Seismol. Res. Lett.* **82,** no. 4, 504–508, doi: 10.1785/gssrl.82.4.504.

Dvorak, J. J., and G. Mastrolorenzo (1990). History of vertical movement movement in Pozzuoli Bay, southern Italy: The result of regional extension related to evolution of the Tyrrhenian Sea and of local volcanic activity, *Special Pap. Geol. Soc. Am.* **263,** 47 pp.

Edwards, B., and D. Fäh (2013). A stochastic ground-motion model for Switzerland, *Bull. Seismol. Soc. Am.* **103,** no. 1, doi: 10.1785/0120110331

Edwards, B., and A. Rietbrock (2009). A comparative study on attenuation and source-scaling relations in the Kanto, Tokai, and Chubu regions of Japan, using data from Hi-Net and KiK-Net, *Bull. Seismol. Soc. Am.* **99,** no. 4, 2435–2460.

Edwards, B., A. Rietbrock, J. J. Bommer, and B. Baptie (2008). The acquisition of source, path, and site effects from microearthquake recording using *Q* tomography: Application to the United Kingdom, *Bull. Seismol. Soc. Am.* **98,** no. 4, 1915–1935, doi: 10.1785/0120070127.

Edwards, B., B. Allmann, D. Fäh, and J. Clinton (2010). Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude, *Geophys. J. Int.* **183,** 407–420, doi: 10.1111/j.1365-246X.2010.04743.x.

Edwards, B., D. Fäh, and D. Giardini (2011). Attenuation of seismic shear wave energy in Switzerland, *Geophys. J. Int.* **185,** no. 2, 967–984, doi: 10.1111/j.1365-246X.2011.04987.x.

Fäh, D., S. Fritsche, V. Poggi, G. Gassner-Stamm, P. Kästli, J. Burjanek, P. Zweifel, S. Barman, J. Clinton, L. Keller, P. Renault, and S. Heuberger (2009). Determination of site information for seismic stations in Switzerland, *Tech. Rept. SED/PRP/R/004/20090831*, Swiss Seismological Service, for the swiss nuclear Pegasos Refinement Project.

Giardini, D. (2009). Geothermal quake risks must be faced, *Nature* **462,** 848–849, doi: 10.1038/462848a.

Goertz-Allmann, B. P., B. Edwards, F. Bethmann, N. Deichmann, J. Clinton, D. Fäh, and D. Giardini (2011). A new empirical magnitude scaling relation for Switzerland, *Bull. Seismol. Soc. Am.* **101,** no. 6, 3088–3095, doi: 10.1785/0120100291.

Grünthal, G., R. Wahlström, and D. Stromeyer (2009). The unified catalogue oearthquakes in central, northern, and northwestern Europe (CENEC)—Updated and expandedto the last millennium, *J. Seismol.* **13,** 517–541.

Havenith, H. B., D. Fäh, U. Polom, and A. Roullé (2007). *S*-wave velocity measurements applied to the seismic microzonation of Basel, Upper Rhine graben, *Geophys. J. Int.* **170,** no. 346–358, doi: 10.1111/j.1365-246X.2007.03422.x.

Hays, W. W. (1980). Procedures for estimating earthquake ground motions, *U.S. Geol. Surv. Profess.l Pap. 1114*, 77 pp.

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

Jousset, P., and B. François (2006). Set-up of a seismological broadband network at Hengill geothermal field (Iceland), *Final Rept.*, BRGM/RP-54971-FR, Orléans, France, 54 pp.

Jousset, P., C. Haberland, K. Bauer, and K. Arnason (2011). Hengill geothermal volcanic complex (Iceland) characterized by integrated geophysical observations, *Geothermics* **40,** no. 1, 1–24, doi: 10.1016/j.geothermics.2010.12.008.

Kale, O., and S. Akkar (2013). A new procedure for selecting and ranking 33 ground-motion prediction equations (GMPEs): The Euclidean-distance based ranking (EDR) method, *Bull. Seismol. Soc. Am.* **103,** no. 2A, 1069–1084.

Li, J., H. S. Kuleli, H. Zhang, and M. N. Toksöz (2011) Focal mechanism determination of induced microearthquakes in an oil field using full waveforms from shallow and deep seismic networks, *Geophysics* **76,** no. 6, WC87–WC101, doi: 10.1190/geo2011-0030.1.

Lindstrom, M. J., and D. M. Bates (1990). Nonlinear mixed effects models for repeated measures data, *Biometrics* **46,** no. 3, 673–687, doi: 10.2307/2532087.

Majer, E. L., R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith, and H. Asanuma (2007). Induced seismicity associated with Enhanced Geothermal Systems, *Geothermics* **36,** 185–222, doi: 10.1016/j.geothermics.2007.03.003.

Massa, M., P. Morasca, L. Moratto, S. Marzorati, G. Costa, and D. Spallarossa (2008). Empirical ground-motion prediction equations for northern Italy using weak- and strong-motion amplitudes, frequency content, and duration parameters, *Bull. Seismol. Soc. Am.* **98,** no. 3, 1319–1342, doi: 10.1785/0120070164.

McGarr, A. (1984). Scaling of ground motion parameters, state of stress, and focal depth, *J. Geophys. Res.* **89,** no. B8, 6969–6979.

McGarr, A., and J. B. Fletcher (2005). Development of ground-motion prediction equations relevant to shallow mining-induced seismicity in the Trail Mountain area, Emery County, Utah, *Bull. Seismol. Soc. Am.* **95,** no. 1, 31–47, doi: 10.1785/0120040046.

McGarr, A., R. W. E. Green, and S. M. Spottiswoode (1981). Strong ground motion of mine tremors: Some implications for near-source ground motion parameters, *Bull. Seismol. Soc. Am.* **71,** no. 1, 295–319.

Oppenheimer, D. H. (1986). Extensional tectonics at the Geysers geothermal area, California, *J. Geophys. Res.* **91,** no. Bll 11,463–11,476.

Poggi, V., B. Edwards, and D. Fäh (2011). Derivation of a reference shear-wave velocity model from empirical site amplification, *Bull. Seismol. Soc. Am.* **101,** no. 1, 258–274, doi: 10.1785/0120100060.

Rietbrock, A., F. Strasser, and B. Edwards (2013). A stochastic earthquake ground-motion prediction model for the United Kingdom, *Bull. Seismol. Soc. Am.* **103,** no. 1, 57–77.

Rodriguez-Marek, A., and F. Bonilla (2011). Analysis of single-station standard deviation using the KiK-net data, *Bull. Seismol. Soc. Am.* **101,** no. 3, 1242–1258, doi: 10.1785/0120100252.

Scherbaum, F., E. Delavaud, and C. Riggelsen (2009). Model selection in seismic hazard analysis: An information-theoretic perspective, *Bull. Seismol. Soc. Am.* **99,** no. 6, 3234–3247, doi: 10.1785/0120080347.

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, insights, and challenges, *Seismol. Res. Lett.* **80,** no. 1, 40–56, doi: 10.1785/gssrl.80.1.40.

Tramelli, A., D. Galluzzo, E. Del Pezzo, and M. A. Di Vito (2010). A detailed study of the site effects in the volcanic area of Campi Flegrei using empirical approaches, *Geophys. J. Int.* **182,** 1073–1086, doi: 10.1111/j.1365-246X.2010.04675.x.

van Eck, T., F. Goutbeek, H. Haak, and B. Dost (2006). Seismic hazard due to small-magnitude, shallow-source, induced earthquakes in The Netherlands, *Eng. Geol.* **87,** no. 1–2, 105–121, doi: 10.1016/j.enggeo.2006.06.005.

Youngs, R. R., N. Abrahamson, F. I. Makdisi, and K. Sadigh (1995). Magnitude-dependent variance of peak ground acceleration, *Bull. Seismol. Soc. Am.* **85,** no. 4, 1161–1176.

Zhao, J. X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H. K. Thio, P. G. Somerville, Y. Fukushima, and Y. Fukushima (2006). Attenuation relations of strong ground motion in Japan using site classification based on predominant period, *Bull. Seismol. Soc. Am.* **96,** no. 3, 898–913, doi: 10.1785/0120050122.

Risks and Prevention Division/Seismic and Volcanic Risks Unit (DRP/RSV)
Bureau de Recherches Géologiques et Minières (BRGM)
3 avenue C. Guillemin, BP 36009
45060 Orléans Cedex 2, France
(J.D.)

Swiss Seismological Service
Eidgenössische Technische Hochschule (ETH) Zürich
Sonneggstrasse 5
8092 Zurich, Switzerland
(B.E., B.M.C.)

Istituto Nazionale di Geofisica e Vulcanologia
Osservatorio Vesuviano
via Diocleziano 328
80124 Napoli, Italy
(V.C., A.T., C.T.)

Dipartimento di Scienze Fisiche
Unità di Ricerca in Sismologia Sperimentale e Computazionale
Università degli Studi di Napoli Federico II
Complesso Universitario Monte S. Angelo, via Cintia
80126 Napoli, Italy
(N.S.)

Seismology Division
Koninklijk Nederlands Meteorologisch Instituut (KNMI)
PO Box 201
NL-3730 AE De Bilt, Netherlands
(D.K.)

AMRA S.c. a r.l., Analysis and Monitoring of Environmental Risk
Via Nuova Agnano 11
80125 Napoli, Italy
(N.M.)

Manuscript received 8 June 2012

Queries

1. AU: Our deepest apologies for the earlier difficulties with the copyedit of your article. Your article has been re-edited, which has resulted in more queries for your attention. Thank you for your patience with the additional requests.
2. AU: Would you be amenable to changing the current wording to " …..provide robust predictions for small, shallow earthquakes at close source-to-site distances ($1 \lesssim M_w \lesssim 5$; focal depth $h \lesssim 5$ km; and hypocentral distances $r_{hyp} \lesssim 20$ km)"? (It would smooth the flow of the sentence and might be simpler for the reader to see the parameters together.)
3. AU: Which change would be more consistent with your intended meaning? (1) From the use of "i.e.," it appears that Campi Flegrei and Hengill have the only instruments in areas of natural geothermic activity, in which case the text could be shortened to "(Campi Flegrei and Hengill are in areas of natural geothermal activity)." (2) Alternatively, if other sites from your list also are in areas of natural geothermic activity, then we should change "i.e." to "e.g." to indicate that Campi Flegrei and Hengill are only examples.
4. AU: Would the following change be correct to both clarify and shorten the text? Change "started in 1969 and, after about 10 years of stable ground level, restarted in 1982 and continued until 1984" to "were from 1969 to 1972 and 1982 to 1984, with about 10 years of stable ground level in the interrim."
5. AU: I have changed the dimensions to "20×10 km^2"; please let us know if you prefer "20 km × 10 km."
6. AU: BSSA uses the company name followed by model details; is OYO Geospace the company that produced the sensors described here?
7. AU: Please provide the full name for this project; I-GET FP6 would be included in parentheses after the full name.
8. AU: In keeping with BSSA listing of make and model numbers for instruments, I have tentatively changed "Lennartz 5s" to "Lennartz LE3D-5s" instruments; is this correct?
9. AU: Please verify if "South Icelandic Lowland" is the currently used full name for the SIL permanent seismic network. (Please provide correction if needed.)
10. AU: The closest make/model I located is the GeoSIG AC-23 accelerometer. (1) Would changing "SIG" to "GeoSIG" be correct for the instruments you used? (2) Is this the same model used in the section on Voerendaal as well? (If so, it will be changed there as well.)
11. AU: Instead of the current sentence beginning with "In light of …..," would this replacement be correct? "In light of this, we have used an automatic procedure to apply the technique presented by Edwards *et al.* (2010) to calculate moment magnitudes (M_w) for the vast majority of earthquakes considered here." If not, please provide alternative phrasing to explain the calculations.
12. AU: Both the original manuscript and the first proof you received use the phrasing "In their comparison with Swiss moment-tensor solution ….. ." The editor's query "Do you mean "of? In other words, do you mean comparing of Swiss moment-tensor solutions by themselves or compared to something else?" was asking whether "with" should be changed to "of" but did not actually change the manuscript. However, in your corrections, you indicate, "We do not agree with this change as it changes our intent. Please change 'comparison with Swiss' to 'comparison of Swiss'." The first part of your correction seems to indicate that we should not change "with" to "of," but then your specific instruction asks for that change to be made. Please indicate which word ("with" or "of") correctly represents the work done by Edwards *et al.* 2010.
13. AU: My apologies for what may be a minor point, but in the sentence beginning with "The Fourier response," would the previously edited text of "preserving the phrase, and providing site-corrected accelerograms" better approximate your original meaning if it is changed to "preserving the phrase, thereby providing site-corrected accelerograms"?
14. AU: In the electronic issues, BSSA generally provides hyperlinks to cross-referenced sections in an article. Please provide the complete section title to which you refer when indicating "a subsequent section"; that title will be added to the text here using the format, "This issue is discussed in a subsequent section, [title here]."
15. AU: Would it be technically correct to change "use a fine grid in terms of r_{hyp} and a coarser grid in terms of magnitude" to "use a fine grid for r_{hyp} and a coarser grid for magnitude"?
16. AU: Please note that BSSA requires tables to be called out (cross-referenced) in sequence, but Table S2 is called out before Table S1. Please verify there will not be a problem with the logical presentation of content if Table S1 and Table S2 are reordered in the electronic supplement and their respective citations corrected in the article itself.
17. AU: BSSA avoids using "above" and "below" to refer to locations within an article; please provide alternative wording if you prefer something other than "earlier in this article."
18. AU: Please provide the title of the section so that it may be included with a hyperlink in the electronic edition. (Do you mean the "Aleatory Variability" section?)
19. AU: Do you also wish to name the sections here so that hyperlinks may be added in the electronic edition?
20. AU: Is there a specific section that should be named here for adding a hyperlink for the electronic edition?

21. AU: Please provide alternative wording (e.g., section title if applicable) that may be substituted for “see below” so as to better comply with BSSA style.
22. AU: Grammatically we needed to avoid ending a sentence with “beyond”; is “at greater distances” consistent with your intention here?
23. AU: Please verify that the left side of equation (4) is correct as “ln y” instead of “ln Y” (as used in equations 1–3).
24. AU: Our apologies that the query in the previous proof was not sufficiently clear. The editor was requesting revised wording for the sentence, “The standard deviations from fitting, using standard least squares regression, this functional form to the simulated PSAs are all smaller (much smaller for $T > 0{:}1$ s) than 0.3, in terms of natural logarithms, showing that the functional form is adequate.” Please review the currently edited sentence for accuracy and to be certain your intended meaning has been retained.
25. AU: In reviewing the initial edit to the sentence beginning with “The predictions from .....,” I was concerned that the edit might have prompted you to change the meaning of your original sentence when you made your correction. I have proceeded with the correction you specified but would like to offer an alternative. Would you prefer us to substitute the following sentence for the current one? “Even close to the source ($r_{\mathrm{hyp}} < 10$ km), the predictions from the stochastic models with different Q diverge as the distance increases.”
26. AU: Please review the edit that added $\Delta\sigma$ to “for $Q = 200$, $\kappa = 0.02$ s, 1 bar and 10 bar”; is this consistent with your intended meaning?
27. AU: Would it be correctly phrased if just the parentheses (but not their content) were removed so that the text reads “within 0.5 km horizontally and within 1 km vertically”?
28. AU: Please verify that the section you wish to cross-reference is Derivation of Weights for Stochastic Models.
29. AU: Please verify that the section you wish to cross-reference is Derivation of Weights for Stochastic Models.
30. AU: I substituted Dipartimento della Protezione Civile of the Istituto Nazionale di Geofisica e Vulcanologia (INGV) for the Italian Department of Civil Protection because BSSA encourages use of the native language for names of organizations. Is this the agency you intended to list?
31. AU: EC needed to be defined in the Acknowledgments as per BSSA style, but the word “Research” was not included in the description of this project in the text itself, which suggested that the “EC-Research” may be a subgroup of the European Commission. Please provide any needed changes to the proper name in the text of the article and Acknowledgments so that it is in the correct form and consistently used in both places.
32. AU: Please provide (1) the name of the city and country and (2) the days and month(s) in which the meeting was held.
33. AU: Kale and Akkar (2013): This article is now published in BSSA. So the volume number and page range have been inserted. Please confirm whether this is okay.
34. AU: Please note that for consistency, the bin grids must use the same units as in the description in the text, so I have added M_{w} after the 0.5.
35. AU: Please provide revised figures for Figures 7 and 8 to correct the following style inconsistencies: (1) BSSA uses scientific notation of the type 1×10^{-6} (which is also what is used in Figure 9). Please correct the scientific notation in Figures 7 and 8 in the axis labels to match BSSA style. (2) Also in Figure 7, please change the x-axis labels to r_{hyp} to match earlier figures (but must use roman “r” instead of italic), and use parentheses rather than square brackets for the units. (3) Also in Figure 8, please provide full labels (rather than just units) to better match other figures in article and to be consistent with BSSA style.
36. AU: For both Tables 2 and 3: Please review the modifications to column 2 that are intended to better match BSSA style.

Elsevier Editorial System(tm) for
Geothermics
Manuscript Draft

Manuscript Number:

Title: Analysis of induced seismicity at The Geysers geothermal area, California

Article Type: Special Issue: Induced Seismicity

Keywords: Induced seismicity, Source and medium properties, Geothermal areas, The Geysers

Corresponding Author: Dr. Antonio Emolo, Ph.D.

Corresponding Author's Institution: University Federico II in Naples

First Author: Antonio Emolo, Ph.D.

Order of Authors: Antonio Emolo, Ph.D.; Ortensia Amoroso; Antonella Orefice; Emanuela Matrullo; Nitin Sharma; Vincenzo Convertito; Nils Maercklin; Aldo Zollo

Abstract: Understanding induced seismicity in a geothermal system can help to characterize extent and conditions of the reservoir, and to assess the potential seismic hazard associated with field operations. Here we summarize analyses of induced seismicity at The Geysers geothermal field in Northern California by analyzing data recorded from July 2007 through October 2011. We compiled a dataset of more than 15000 events with magnitudes between 1.0 and 4.5 and re-picked first-arrival times. Using the highest-quality picks, we determined a minimum 1-D velocity model, which is also used as an initial model for 3-D seismic tomography. The well-constrained 1-D P-velocity model shows a sharp velocity increase at 1-2 km depth (from 3 to 5 km/s) and then a gradient-like trend down to about 5 km depth, where velocities reach values of 6-7 km/s. The station corrections show coherent, relatively high, positive travel-time delays in the NW zone, thus indicating a strong lateral variation of the P-wave velocities, which is further corroborated by the 3-D tomographic model. After probabilistic, non-linear relocation of the events, the seismicity appears spatially diffused in a 15x10 km2 area elongated in NW-SE direction, with hypocentral depths between 0 and 6 km and generally deeper in the NW sector. The average vP/vS ratio is about 1.67, consistent with previous estimates for the same time period. To infer the present stress regime, we computed focal mechanisms from P-wave first-arrival polarities. The inferred fault-plane solutions show a dominant strike-slip and normal faulting mechanisms, with P and T axes coherently oriented as expected for the regional stress field. We also determined the seismic moments, source radii and stress drops using a multi-step, iterative inversion procedure of P-wave displacement spectra, assuming a four-parameters spectral model and a constant-Q attenuation factor. We observe a self-similar scaling of the computed source parameters in the whole investigated magnitude range, with a nearly

constant stress-drop of 20 and 120 MPa depending on the use of the Brune (1970) or Madariaga (1976) source model, respectively.

Suggested Reviewers: Roland Gritto
Roland.Gritto@arrayinfotech.com

Volker Oye
volker@norsar.no

Catherine Dorbath
catherine.dorbath@eost.u-strasbg.fr

Cover Letter

Università degli Studi "Federico II"
Dipartimento di Scienze Fisiche
Complesso Universitario di Monte S.Angelo – Edificio 6
Via Cintia
80126 – Napoli – Italy

Naples, February 5, 2013

Dear Editor,

I would like to submit the manuscript "*Analysis of induced seismicity at The Geysers geothermal area, California*" by Antonio Emolo, Ortensia Amoroso, Antonella Orefice, Emanuela Matrullo, Nitin Sharma, Vincenzo Convertito, Nils Maercklin, and Aldo Zollo, for publication in the special issue about *Induced Seismicity* of the journal *Geothermics.*

This is an original research paper which has neither previously, nor simultaneously, in whole or in part been submitted anywhere else. All authors have read and approved the final version submitted.

In this manuscript we present results about the propagation medium properties and earthquake source inferred from the analysis of the microseismicity at The Geysers geothermal field in California.

Three potential independent reviewers who have excellent expertise in the studied field of this paper are:

Roland Gritto, Roland.Gritto@arrayinfotech.com

Volker Oye, volker@norsar.no

Catherine Dorbat, catherine.dorbath@eost.u-strasbg.fr

Best regards,

Antonio Emolo

Antonio Emolo

Highlights of the paper ‘*Analysis of induced seismicity at The Geysers geothermal area, California*’, by A. Emolo et al., submitted for publication in *Geothermics*.

We analyzed about 4 years of microseismicity data collected at The Geysers geothermal field (California) to infer information about medium properties and earthquake sources.

We infer a 1-D P-wave velocity model to be used as initial model for the 3-D tomography. The 1-D model presents a sharp velocity increase at 1-2 km depth (from 3 to 5 km/s) and then a gradient-like trend down to about 5 km depth, where velocities reach values of 6-7 km/s. The 3-D P-wave tomographic model shows a strong lateral variation. Vp/Vs ratio is larger (about 2) in the north-western sector of the inverstigated area, than in the south-eastern part. After a non-linear relocation of events, the seismicity appears spatially diffused in a 15x10 km^2 area elongated in NW-SE direction, with hypocentral depths between 0 and 6 km and generally deeper in the NW sector. We observe a self-similar scaling of source parameters as corner-frequency, stress-drop, source radii, and seismic moment, with a nearly constant stress-drop of 20 and 120 MPa depending on the use of the Brune (1970) or Madariaga (1976) source model, respectively.

Analysis of induced seismicity

at The Geysers geothermal area, California

***Antonio Emolo*[1], *Ortensia Amoroso*[2], *Antonella Orefice*[1], *Emanuela Matrullo*[3,4], *Nitin Sharma*[1],**
***Vincenzo Convertito*[2], *Nils Maercklin*[3,5], *Aldo Zollo*[1]**

[1] Dipartimento di Fisica, Università degli Studi Federico II, Napoli, Italy

[2]Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Napoli, Italy

[3]AMRA scarl, Napoli, Italy

[4]Laboratoire de Géologie, Ecole Normale Supérieure – CNRS, Paris, France

[5]TGS Geophysical Company Ltd, Bedford, United Kingdom

Author for correspondence: Antonio Emolo
Dipartimento di Fisica
Università degli Studi Federico II
Complesso Universitario di Monte S. Angelo – Edificio 6
via Cintia
80126 – Napoli – Italy
Phone: +39 081 676822
E-mail: antonio.emolo@unina.it

Abstract

Understanding induced seismicity in a geothermal system can help to characterize extent and conditions of the reservoir, and to assess the potential seismic hazard associated with field operations. Here we summarize analyses of induced seismicity at The Geysers geothermal field in Northern California by analyzing data recorded from July 2007 through October 2011. We compiled a dataset of more than 15000 events with magnitudes between 1.0 and 4.5 and re-picked first-arrival times. Using the highest-quality picks, we determined a minimum 1-D velocity model, which is also used as an initial model for 3-D seismic tomography. The well-constrained 1-D P-velocity model shows a sharp velocity increase at 1-2 km depth (from 3 to 5 km/s) and then a gradient-like trend down to about 5 km depth, where velocities reach values of 6-7 km/s. The station corrections show coherent, relatively high, positive travel-time delays in the NW zone, thus indicating a strong lateral variation of the P-wave velocities, which is further corroborated by the 3-D tomographic model. After probabilistic, non-linear relocation of the events, the seismicity appears spatially diffused in a 15x10 km^2 area elongated in NW-SE direction, with hypocentral depths between 0 and 6 km and generally deeper in the NW sector. The average v_P/v_S ratio is about 1.67, consistent with previous estimates for the same time period. To infer the present stress regime, we computed focal mechanisms from P-wave first-arrival polarities. The inferred fault-plane solutions show a dominant strike-slip and normal faulting mechanisms, with P and T axes coherently oriented as expected for the regional stress field. We also determined the seismic moments, source radii and stress drops using a multi-step, iterative inversion procedure of P-wave displacement spectra, assuming a four-parameters spectral model and a constant-Q attenuation factor. We observe a self-similar scaling of the computed source parameters in the whole investigated magnitude range, with a nearly constant stress-drop of 20 and 120 MPa depending on the use of the Brune (1970) or Madariaga (1976) source model, respectively.

Keywords: *Induced seismicity, Source and medium properties, Geothermal areas, The Geysers.*

1 Introduction

1.1 Enhanced geothermal systems and induced seismicity

Geothermal systems have the potential to provide a significant contribution to the world energy demand. These are naturally or artificially fractured hydrothermal systems, providing accessible sources to extract heat stored in subsurface rocks, which may be used directly for heating purposes or to generate electricity. The total resources and their availability tend to be restricted to certain areas, and the field development proceeds only where conditions are ideal for cost-efficient extraction. Such hydrothermal systems sometimes are difficult to locate and also run a high risk of not being commercially feasible, if their geological, physical and chemical characteristics are not favorable. In that case the underground heat source requires enhancement (e.g., Nielson et al., 2001).

The reason for developing Enhanced Geothermal Systems (EGS) is two-fold: (1) to bring uneconomic hydrothermal systems into production by improving their underground conditions (hydraulic fracturing, reservoir stimulation); and (2) to engineer an underground condition that creates a new fracture network system, where injected fluids are heated by circulation through hot fractured rock at depth. The hot fluid or steam is then brought to the surface to deliver the captured heat for power conversion and other uses. The second approach expands the available heat resources significantly and reduces the uncertainty of exploitation costs.

However, fluid injection, steam extraction, and reservoir stimulation in EGS lead to induced seismicity. Although only in rare cases induced events may be large enough to pose a hazard to population or installed facilities, on the other hand the microseismicity provides valuable information on the extent and properties of the reservoir (e.g., Eberhart-Phillips and Oppenheimer, 1984; Stark, 2003). Therefore, microseismic monitoring is important, both for mitigation of the unwanted effects of industrial operations and for the continuous assessment of reservoir conditions. The controversial issue of induced seismicity in EGS has attracted seismological research. Although induced seismicity has in fact few adverse physical effects on operations or on surrounding communities, public concern over the amount and magnitude of seismicity associated with current and future EGS operations still remains.

To better understand induced seismicity and the reservoirs in which it occurs, we applied a set of different seismological analyses to a high-quality seismic dataset collected at The Geysers geothermal field, located about 120 km north of San Francisco, California.

1.2 Seismicity, data and equipment of The Geysers, California

The Geysers is a vapor-dominated field with the top of the main steam reservoir at about 1-3 km below the surface. The field has been actively exploited since the 1960's and is now the most productive geothermal field in the world (Bertani, 2012). With increasing field development also the seismicity increased in the area (e.g., Majer et al., 2007, Majer and Peterson, 2007). As reported for example by Stark (2003), the dominant mechanism of induced seismicity generation at The Geysers EGS is the temperature contrast between the injected water and the hotter rock fracture surfaces.

The Geysers area can be divided into two distinct seismicity source zones which are outlined by the dashed lines shown in Figure 1a and here named ZONE1 and ZONE2. The separation arguments are supported by Stark (2003), Beall and Wright (2010), and Beall et al. (2010). Some differences in the seismicity distribution were also noted by Eberhart-Phillips and Oppenheimer (1984). All the authors observed that the southeastern part of the Geysers reservoir is seismically less active than the northwestern where seismicity extends to greater depth. The differences were basically ascribed to a depth variation in the high-temperature (260–360° C) vapor-dominated reservoir. Specifically, Beall and Wright (2010) identified a net "$M \geq 4.0$ *dividing line*" which separates the whole area into two seismically different zones. In fact, the northwestern area (ZONE1) contains all the earthquakes having magnitude larger than 4.0, whereas the southeastern one (ZONE2) is characterized by lower magnitude events. Further supporting arguments were more recently brought by Convertito et al. (2012) who also distinguished two different zones in the Geysers area from the analysis of the *b*-values of the Gutenberg-Richter relationship based on the Utsu (1992) test.

Different temporary and long-term seismic networks have been deployed in The Geysers geothermal field during the last five decades to monitor the seismic activity. At present, local seismicity is monitored by the dense Lawrence Berkeley National Laboratory Geysers/Calpine (BG) surface seismic network and by some nearby stations of the Northern California Seismic Network (NCSN). The BG network consists of 32 three-component stations, 29 of which were used for the present study (red triangles in Figure 1a). The BG stations are distributed over an area of about 20x10 km^2, covering the entire geothermal field. Each station was equipped with I/O Sensor SM-6 geophones with a natural frequency of 14 Hz. In Fall 2009 these instruments were replaced by Oyo GS-11D 4.5 Hz sensors.

For the present studies we compiled a waveform database of 15476 events recorded at BG stations between 2007/07/24 and 2011/10/30, and associated to earthquakes in the catalog of the Northern California Earthquake Data Center (NCEDC). Figure 1b shows the magnitude-frequency distribution of events at The Geysers found in the NCEDC catalog (orange) versus the distribution of waveforms present in our database (blue), indicating that our analyses include all the events with magnitude larger than 1.2 in the studied period which was identified as the average minimum-magnitude of completeness by Convertito et al. (2012). Travel times as a function of the epicentral distance are shown in panel c of Figure 1 to illustrate the regional extent and number of arrival-time picks in our database.

1.3 1-D starting P-wave velocity model

To better constrain the earthquakes hypocentral locations and determine a reliable reference model for the 3-D seismic tomography, we preliminarily searched for the optimal 1-D P-wave velocity model for the studied area by inverting the first arrival times using the VELEST software (Kissling et al., 1994).

The inversion procedure was initially limited to a reduced number of data obtained from the whole data-base selecting only those events having at least five P arrival-time readings, an azimuthal gap smaller than 200°, a maximum location error (for both horizontal and vertical directions) of 2 km, and maximum Root Mean Square (RMS) travel-time residual of 0.038 s. The final data-set is thus composed of 14483 first P arrival-time readings, corresponding to 565 very well located seismic events. Time picking accuracy was estimated to be in the range 0.02-0.2 s.

As starting models for the inversion procedure we selected different 1-D velocity models available in the literature for The Geyser area. These models (left panel in Figure 2) present a very broad range for the P-wave velocities, especially in the upper few first kilometers. Moreover, the number and the depths of the interfaces characterizing each model, strongly differ from each other probably reflecting different assumptions, tools, and data used in each study.

For each starting velocity model, the convergence of the inversion procedure to a stable solution was obtained after 15-20 iterations and the retrieved final models (right panel in Figure 2) are all characterized by RMS values not larger than 0.0136 s. In order to account for the information carried by all the considered final models, an average velocity model (dashed black curve in the left panel of Figure 2) has been used as starting model for a further inversion step whose solution will represents our *best 1-D velocity model* (dashed black curve in the right panel of Figure 2). For the best model we obtained a final RMS value of the travel time residuals of 0.0135 s.

1.4 V_P/V_S ***ratio, earthquake relocation, and focal mechanisms***

An optimal V_P/V_S ratio has been estimated in order to better constrain the earthquake relocation. To this aim, we selected a total of 1320 events recorded at least at 15 stations and with a minimum magnitude of 2. Moreover, we manually picked the first P- and S-wave arrival times obtaining in this way a high-quality data-set. A weighting factor was assigned to the P- and S-phase readings according to the estimated uncertainties. Specifically, we identified five different uncertainties intervals (lower than 0.02s, 0.02-0.02s, 0.04-0.08s, 0.08-0.20s, and larger than 0.20s) at which we assigned a decreasing weighting factor. The final data-set consists of 31414 P- and 9752 S-wave arrival times, respectively.

A first evaluation of an average V_P/V_S ratio has been performed by analyzing the modified Wadati diagram obtained by plotting the difference between the S-wave arrival-times (T_{Si}-T_{Sj}) as a function of the P-wave arrival-time differences (T_{Pi}-T_{Pj}) for each couple (*i*, *j*) of stations (Chatalain, 1978). Assuming a homogeneous half-space, it can be easily shown that the travel-time differences should follow a straight line whose slope is equal to the V_P/V_S ratio of the half-space:

$$T_{Pi} - T_{Pj} = \frac{\upsilon_P}{\upsilon_S}\left(T_{Si} - T_{Sj}\right) \qquad (1)$$

By fitting the time-difference data, we estimated a V_P/V_S ratio of about 1.670 with a root mean square error of 0.002.

Seismic events have been then relocated in the best 1-D velocity model (Figure 3). The seismicity appears spatially diffused in a 15x10 km^2 area elongated in the NW-SE direction, with earthquakes depths ranging between 0 and 6 km. As pointed out in previous studies about The Geyser geothermal field (e.g., Eberhart-Phillips and Oppenheimer, 1984; Stark, 2003), we also find that events occurring in the NW sector (ZONE1) are, on average, deeper than in the SE area (ZONE2) according the variation of the high-temperature zone (Stark, 2003).

Moreover, we computed the focal mechanisms for the largest seismic events in our data-set ($M > 3$) using the P-wave first-arrival polarities and the FPFIT algorithm (Reasenberg and Oppenheimer, 1985). In accordance with previous studies (e.g., Eberhart-Phillips and Oppenheimer, 1984; Mossop, 1998), the fault-plane solutions found here show a dominant strike-slip and normal faulting mechanisms, with P and T axes coherently oriented with respect to the expected regional tectonic regime in Northern California, associated with the relative motion between the North American and Pacific plates.

2 3-D P- and S-wave velocity models

The 3-D P- and S-wave velocity models are obtained from the tomographic inversion of the first P- and S-wave arrival times described in the previous section. The inversion is performed by using an improved technique based on the accurate finite-difference travel time computation and a simultaneous inversion of both velocity models and earthquake locations (Latorre et al., 2004; Amoroso et al., 2012). The available stations and events distribution allowed us to investigate a volume of $30 \times 25 \times 6$ km^3 which embeds the whole geothermal area. The model is parameterized by a nodal representation, described by a tridimensional grid in which the spacing is equal to $3 \times 3 \times 1$ km^3. The best 1-D velocity model and the retrieved Vp/Vs ration described in the previous section were used as starting model for the 3-D seismic tomographic inversion. The misfit function defined as the sum of the squared time delay, is *a posteriori* analyzed and it was found that the convergence is reached after 10 iterations. The final misfit value is equal to 0.1s with a reduction of about 50%.

Figure 5 shows the P-wave velocity model at a set of horizontal slices corresponding to different depths (0, 1, 1.5 and 2 km). The images indicate the presence of a strong variation of the seismic velocity along the NW-SE direction further confirming the presence of a high temperature reservoir whose top separates deep, injection-induce the seismicity from shallower seismic activity (Stark, 2003).

By using the retrieved velocity models, we deduced V_P/V_S images. Figure 6 shows a map view of the V_P/V_S ratio at the same depths used to represent the P-wave velocity model shown in Figure 5. The V_P/V_S maps indicate a relevant anomaly in the NW area (ZONE1) which extends down to 3-4 km depth and is characterized by values larger than 2. This anomaly can be associated with a dense fractured/porous partially fluid-saturated rock volume (Ito et al., 1979)

3 Spectral analyses

Following the study by Boatwright (1980), the displacement spectrum $U(\omega)$ can be described through the relation

$$U(\omega) = C_S \frac{\Omega_0}{1+\left(\frac{\omega}{\omega_c}\right)^{\gamma}} \exp\left(-\frac{\pi t^*}{2}\right) \qquad (2)$$

where Ω_o is the low-frequency spectral level ($\omega << \omega_c$), that is related to seismic moment M_0, ω_c is the angular corner-frequency, that is related to source radius r, and γ is the high-frequency spectral fall-off parameter, that in the present sutdy is assumed to vary in the range 1-3. The selection of source model which can differ from the standard ω^{-2} model may be justified by the fact that the triggering mechanism of the induced seismicty, that for The Geysers is ascribed to a temperature contrast can modify the spectral fall-off. The constant C_S accounts for direct P-wave amplification due to a vertically varying velocity structure (Aki and Richards, 1980) while the exponential $\exp\left(-\frac{\pi t^*}{2}\right)$ accounts for the anelastic body-wave attenuation along the travel path, being t^* the attenuation parameter which depends on the travel-time T and the quality factor Q ($t^*=T/Q$).

3.1 Anelastic attenuation

Induced seismicity is generally characterized by low-to-moderate magnitude earthquakes. Particularly for the relatively low magnitude events ($M_L<1.5$) we expect that the corner frequency is above the available frequency range that in the present analysis is 6-20 Hz. Thus, for those earthquakes the only parameter that can be estimated are t^* and the seismic moment M_0. To this aim, we use a technique which infers the intercept and slope of the logarithm of the displacement spectrum for frequencies smaller than the expected corner-frequency. Indeed, from equation (2), given a source i and a station j, the displacement spectrum at frequencies smaller than the corner-frequency on a log-log scale can be written as

$$\log \Omega_{ij}(\omega < \omega_c) = \log\left(\Omega_o^i\right) - \frac{\omega t_{ij}^*}{2}$$

where Ω_0 is the low-frequency spectral level which is proportional to seismic moment M_0.

In a first approximation, we assumed the attenuation parameter t^* independent on the frequency. Given a set of M microearthquakes and N recording stations, one has to solve the NxM system of linear equations (3) to determine the M values of Ω_0 and NxM values of t^*. Due to the expected degree of sparseness of the matrix, we adopted a trial-and-error method where the value of Ω_0 for a given earthquake is determined by exploring a range of values around an initial guess inferred from the local magnitude which for the dataset analyzed in the present study is reported in Northern California Earthquake Data Center (NCEDC) catalog.

For this analysis we inverted the P-wave displacement spectra of about 700 earthquakes and applied the non-linear Levenberg-Marquardt least-square algorithm (Marquardt, 1963), implemented in the software package GNUPLOT (Janert, 2009), for curve fitting and parameters estimation. We estimated an average t^* value of (0.029 ± 0.012) s.

Moreover, we analyzed the station residuals of P-travel time and the P-attenuation characteristic time t^*, comparing them to the single-station V_P/V_S ratio (Figure 7) as function of the distance along the profile NW-SE oriented shown in Figure 1. Specifically, the V_P/V_S ratio at each single station has been obtained as the average value of the difference between S- and P-wave arrival times divided by the P-wave travel-time for all the events recorded at that station (e.g., Wadati, 1933).

The three quantities show relatively high values in the ZONE1 as compared to ZONE2 confirming once again the differences between the NW and SE areas already pointed out from earthquake locations, P- and V_P/V_S seismic tomography images.

3.2 Source parameter and scaling relationship

We use a frequency-domain parametric approach to estimate the seismic moment M_0, the corner-frequency f_c, and the high-frequency spectral fall-off parameter γ from P-wave micro-earthquakes displacement spectra. For the spectral analysis, we have chosen magnitude-dependent frequency band limits, by extending it at higher frequencies for the analysis of the smallest events. To account for noise content, the signal-to-noise ratio (S/N) retrieved in the considered frequency-bands, is used as a weighting factor in the inversion procedure. The recorded signals have been preliminarily corrected for the overall instrument response curves, considering for each station its specific seismic sensor type and data-logger transfer function.

By setting in the spectral model (2) the average t^* value obtained for smallest earthquakes ($<t^*> = 0.029$ s), the P-wave displacement spectra are inverted to obtain the refined estimations for the source parameters (M_o, f_c) for about 2400 earthquakes with magnitude ranging between 0.7 and 4.5.

In Figure 8 the corner-frequency, static stress-drop, and source radius are plotted as a function of the seismic moment (panels a, b, and c, respectively). In particular, the source radius is estimated from the corner frequency assuming the Madariaga (1976) model. The seismic moment ranges between $2.8 \cdot 10^{11}$ and $2.3 \cdot 10^{16}$ Nm, while the corner frequency varies in the interval [0.7-143.6] Hz. The seismic moment and circular fault radius are then used to estimate the static stress-drop (Keilis-Borok, 1959). We observe a self similarity scaling-law with constant Madariaga’s stress drop of about 64 MPa, which corresponds to a Brune's stress drop of about 12 MPa.

In Figure 9a the spatial variation of stress-drop, moving from ZONE1 to ZONE 2 along the NW-SE direction, is shown. It is clear that the highest values of stress-drop are observed in the ZONE1.

On the other hand, Figure 9b shows the variation of the static stress-drop with depth for the two different zones. Each point in plot has been obtained by averaging the data grouped in 0.5km depth bins. We observe that the static stress-drop increases with depth in both zones, remaining higher in ZONE1. This feature was also noted by Sharma et al. (2013) by analyzing peak-ground motion parameters in the same area and for the same period and can be a supporting argument for the observed larger magnitude event occurrence in the ZONE1.

4 Discussion and Conclusions

Our studies on subsurface properties and induced seismicity at The Geysers confirm and refine the subdivision of the geothermal field into two distinct zones (ZONE1 and ZONE2) exhibiting significantly different medium and seismic source properties. The difference is corroborated by the results obtained from the 3-D seismic tomography and the V_P/Vs ratio images, from the analysis at each station of the P-wave travel time residuals, V_P/Vs ratio, and the anelastic attenuation. The V_P/V_S values retrieved from the 3D analysis are larger than 2.0, particularly for the northwestern sector, which is a clear indication of a fractured, fluid-saturated rock crustal volume.

A constant stress-drop scaling of source parameters is observed over about four orders of magnitude of variation of seismic moment. The median value of Madariaga static stress drop is relatively high, about 64 MPa, which corresponds to Brune's stress drop of about 12 Mpa. This confirms that stress-drop of induced micro-fractures in geothermal areas are relatively high and comparable to the values observed for natural earthquakes recorded in tectonically active areas. We could also demonstrate a clear spatial variation of stress-drop moving from ZONE1 to ZONE2 along the NW-SE direction with the higher values observed in the latter area. This can explain the difference in the seismogenic potential among the two zone and in particular the observation proposed by Beall and Wright (2010) concerning the occurrence of the larger event (M≥4) in the ZONE1.

We interpret the relatively high stress-drop values as associated with the dominant fluid fault-lubrication effect which strongly reduce the friction level during microfractures.

We observe a difference in the depths distribution of the induced earthquakes occurring in the two zones. In particular, ZONE1 is characterized by a peak of the seismicity at depths ranging between 2 and 3 km while in ZONE2 the peak is at about 2 km. This feature is correlated with a known difference in the depth of the reservoir which is deeper in the ZONE1 with respect to ZONE2 (e.g., Beall and Wright, 2010). The additional interesting result obtained in the present study concerns the inferred variation of the static stress-drop as function of the depth. In particular, we observe a quite linear increasing of the static stress-drop up to 3.0 for both the two areas. However, deeper part of the ZONE1 features larger values (up to 250MPa) with respect to ZONE2 and, as consequence, could originate larger events.

Fault-plane solutions for a selected number of earthquakes having magnitude larger than 3, show a dominant strike-slip and normal faulting mechanisms, characterized by P and T axes coherently oriented as expected for the regional stress field.

Acknowledgements

This work has been financially supported within the 7th Research Program of European Union through the project GEISER “Geothermal Engineering Integration Mitigation of Induced Seismicity in Reservoirs”. Waveforms and parametric data of the Lawrence Berkeley National Laboratory Geysers/Calpine seismic network have been retrieved from the Northern California Earthquake Data Center (NCEDC; network code BG; http://www.ncedc.org). Figures have been generated with the Generic Mapping Tools (GMT; Wessel and Smith, 1998).

References

Aki K. and P. G. Richards (1980). Quantitative Seismology, Theory and Methods (2 volumes). W. H. Freeman, San Francisco, 932 pp.

Amoroso O., A. Zollo and J. Virieux J. (2012). 3D seismic imaging of an active, normal fault zone in southern Apennines (Italy): clues on fluid driven microearthquake fracturing. Abstract S31C-07 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.

Beall J. J. and M. C. Wright (2010). Southern extent of The Geysers high temperature reservoir based on seismic and geochemical evidence. In: GRC Transactions, Geothermal Resources Council. Vol. 34. Geothermal Resources Council, Geothermal Resources Council, Davis, California, pp. 1199-1202.

Beall J. J., M. C. Wright, A. S. Pingol and P. Atkinson (2010). Effect of high rate injection on seismicity in The Geysers. In: GRC Transactions, Geothermal Resources Council. Vol. 34. Geothermal Resources Council, Davis, California, pp. 1203-1208.

Bertani R. (2012). Geothermal power generation in the world 2005-2010 update report. Geothermics **41**, 1-29, doi:10.1016/j.geothermics.2011.10.001.

Boatwright J. (1980). A spectral theory for circular seismic sources; simple estimates of source dimension, dynamic stress drop, and radiated seismic energy. Bull. Seism. Soc. Am. **70**, 1–27.

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

Chatalain J. L. (1978). Etude fine the la sismicite en zone de collision continental à l’aide d’un réseau de stations portables: la region Hindu-Kush-Pamir. Thèse de 3 éme cicle, Univ. de Grenoble, Grenoble, France.

Convertito V., N. Maercklin, N. Sharma and A. Zollo (2012). From induced seismicity to direct time-dependent seismic hazard. Bull. Seism. Soc. Am. **102**, 2563-2573, doi:10.1785/0120120036.

Eberhart-Phillips, D. (1986). Three-dimensional velocity structure in the northern California coast ranges from inversion of local earthquake arrival times. Bull. Seism. Soc. Am.**76**, 1025-10521.

Eberhart-Phillips D. and D. H. Oppenheimer (1984). Induced seismicity in The Geysers geothermal area, California. J. Geophys. Res. **89**, 1191-1207, doi:10.1029/JB089iB02p01191.

Ito H., J. De Vilbiss and A. Nur (1979). Compressional and shear waves in saturated rock during water-stream transition. J. Geophys. Res. **84**, 4731-4736.

Janert K. P. (2009). Gnuplot in action: understanding data with graphs. Manning Publications Co., 396 pp.

Keilis-Borok V. I. (1959). On the estimation of the displacement in an earthquake source and of source dimensions. Ann. Geophis. **12**, 205-214.

Kissling E., W. L. Ellsworth, D. Eberhart-Phillips and U. Kradolfer (1994). Initial reference model in local earthquake tomography. J. Geophys. Res. **99**, 19635-19646.

Latorre D., J. Virieux, T. Monfret, V. Monteiller, T. Vanorio, J.-L. Got and H. Lyon-Caen (2004). A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set. Geophys. J. Int. **159**, 1013–1031.

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

Majer E. L., R. Baria, M. Stark, S. Oates, J. Bommer, B. Smith and H. Asanuma (2007). Induced seismicity associated with enhanced geothermal systems. Geothermics **36**, 185-222, doi:10.1016/j.geothermics.2007.03.003.

Majer E. L. and J.E. Peterson (2007). The impact of injection on seismicity at The Geysers, California geothermal field. International Journal of Rock Mechanics and Mining Sciences **44**, 1079-1090, doi:10.1016/j.ijrmms.2007.07.023.

Marquardt D. (1963). An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics **11**, 431–441, doi:10.1137/0111030.

Mossop A. (1998). Unpublished work on mechanisms of Geysers injection-induced seismicity, available at Stanford University c/o Dr. Paul Segall.

Nielson D. L., S. Garg, B. Koenig, A. Truesdell, M. A. Walters, M. Stark, W. T. Box and J. J. Beall (2001). Concept for an enhanced geothermal reservoir at The Geysers. Geothermal Resources Council Transactions, Volume 25, Pages 191-194.

Reasenberg P. and D. Oppenheimer (1985). FPFIT, FPPLOT and FPPAGE: Fortran computer programs for calculating and displaying earthquake fault plane solutions. U.S. Geol. Surv., Open File Report, 85-739.

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

Sharma N., V. Convertito, N. Maercklin and A. Zollo (2013). Ground-Motion Prediction Equations for The Geysers geothermal area based on induced seismicity records. Bull Seism. Soc. Am. **3**, doi: 10.1785/0120120138.

Stark M. (2003). Seismic evidence for a long-lived enhanced geothermal system (EGS) in the northern geysers reservoir. In: GRC Transactions, Geothermal Resources Council. No. 27. Geothermal Resources Council, Davis, California, pp. 727-732.

Utsu T. (1992). On seismicity. In: Report of cooperative research of the Institute of Statistical Mathematics, Vol. 34, Mathematical Seismology VII, Annals of the Institute of Statistical Mathematics, 139–157.

Wadati K. (1933). On the travel time of earthquake waves, Part II. Geophys. Mag. **7**, 101-111.

Warren, D. H. (1981). Seismic-refraction measurements of crustal structure near Santa Rosa and Ukiah, California, U.S. Geol. Surv. Prof. Pap. 1141, 167-181.

Wessel P. and W. H. F. Smith (1998). New, improved version of Generic Mapping Tools released. EOS Transactions, American Geophysical Union **79**(47), 579, doi:10.1029/98EO00426.

Figures Captions

Figure 1. (a) Induced seismicity map at The Geysers (green circles). The red triangles indicate the seismic stations of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network used in this study. Yellow triangles correspond to the injection wells. The dashed lines outline the two seismic source zones (ZONE1 and ZONE2). Black line roughly oriented NW-SE identifies the profile adopted for figure 7. (b) Magnitude-frequency distribution of events at The Geysers as found from the NCEDC catalog (orange) *versus* the distribution of waveforms in ourdatabase (blue), indicating that our analyses include all events with magnitude larger than 1.2 in the studied period. (c) Travel-times *versus* epicentral distances to illustrate the regional extent of data and number of arrival time picks in our database.

Figure 2. 1-D P-wave velocity models available from the literature for the investigated area and used as starting models for the inversion procedure (a) and final velocity models retrieved using the VELEST code (b) are shown in different colors. Dashed line in panel (a) corresponds to the average final model used to infer the *best 1-D velocity model* (dashed line in panel b). RMS value in seconds are also provided for each model in the panel b). Model 1: Ross et al. (1999); model 2: Eberhart-Phillips and Oppenheimer (1984); model 3: Eberhart-Phillips (1986); model 4: Calnet model, available in Eberhart-Phillips and Oppenheimer (1984); model 5: Warren, D. H. (1981).

Figure 3. Map view of the relocated epicenters (top, red dots) and E-W vertical section of the seismicity and histogram of the events as function of depth (bottom). Triangles indicate the seismic stations of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network used in this study.

Figure 4. Focal mechanisms for earthquake with magnitude larger than 3, obtained using the first P-wave polarities. Different colors correspond to the earthquakes depths according to color-scale shown. Triangles indicate the seismic stations of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network used in this study.

Figure 5. Map view of P-wave velocity model at 0, 1, 1.5 and 2 km depth from 3-D travel-time tomography. Regions not covered by ray-paths are in grey. The seismic stations of the Lawrence Berkeley National Laboratory (LBNL) Geysers/Calpine seismic network, are shown at the 0 km depth layer as open triangles while black dots in all panels represent the earthquakes locations.

Figure 6. Map view showing the V_P/V_S ratio inferred from the 3-D tomography, at 0, 1, 1.5 and 2 km depth. Convention are similar to Figure 5.

Figure 7. Spatial variation along a NW-SE profile (see Figure 1) of the single station V_P/V_S (top), and station residuals for P-wave travel-times (middle) and P-wave attenuation parameter t* (bottom). Dashed lines outline the boundary between the two seismic source zones (ZONE 1 on the left and ZONE2 on the right). Staions codes are reported in all panels.

Figure 8. Corner-frequency, static stress-drop (according to the Madariaga, 1976, source model) and source radius are reported as a function of the seismic moment in the left, middle, and right panels, respectively. Error bars are also shown in grey. The red line in the middle panel corresponds to the median value of the Madariaga' static stress-drop (64 MPa). In each panel, dashed lines outline the constant stress-drop values expressed in MPa.

Figure 9: (a). Spatial distribution of the static stress-drop values for the analyzed events. Circles corresponds to the earthquake epicenters and their color is proportional to the stress-drop according to the color-scale in the figure. (b). Static stress-drop *versus* depth for the analyzed earthquakes, distinct for ZONE1 (black line) and ZONE2 (grey line). (c). Distribution of hypocentral depths for the ZONE1 (black line) and ZONE2 (grey line).

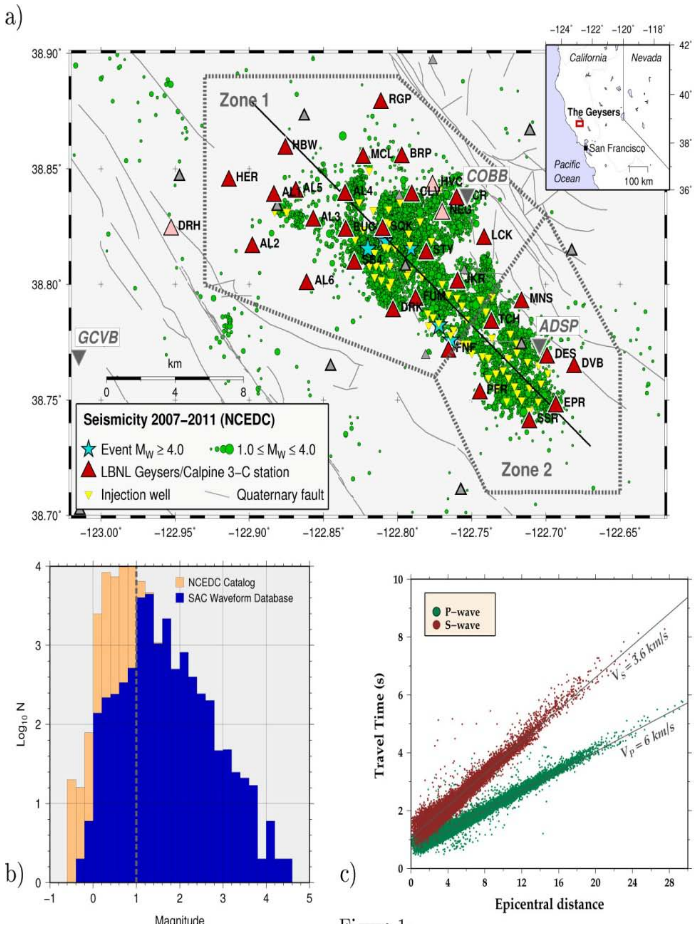

a)
Zone 1
Zone 2
COBB
ADSP
GCVB
RGP
HBW
MCL
BRP
HER
AL5
AL4
AL3
AL2
AL6
CLV
HVC
NEG
DRH
BUC
SQK
STY
SB4
LCK
JKR
FUM
DRK
MNS
TCH
FNF
DES
DVB
PFR
EPR
SSR
km
0
4
8
Seismicity 2007–2011 (NCEDC)
Event $M_W \geq 4.0$
$1.0 \leq M_W \leq 4.0$
LBNL Geysers/Calpine 3-C station
Injection well
Quaternary fault
38.90°
38.85°
38.80°
38.75°
38.70°
−123.00°
−122.95°
−122.90°
−122.85°
−122.80°
−122.75°
−122.70°
−122.65°
California
Nevada
The Geysers
San Francisco
Pacific Ocean
100 km
−124°
−122°
−120°
−118°
42°
40°
38°
36°
b)
NCEDC Catalog
SAC Waveform Database
Log_{10} N
Magnitude
−1
0
1
2
3
4
5
c)
P−wave
S−wave
$V_S = 3.6$ km/s
$V_P = 6$ km/s
Travel Time (s)
Epicentral distance
0
4
8
12
16
20
24
28
2
6
10

Figure 1

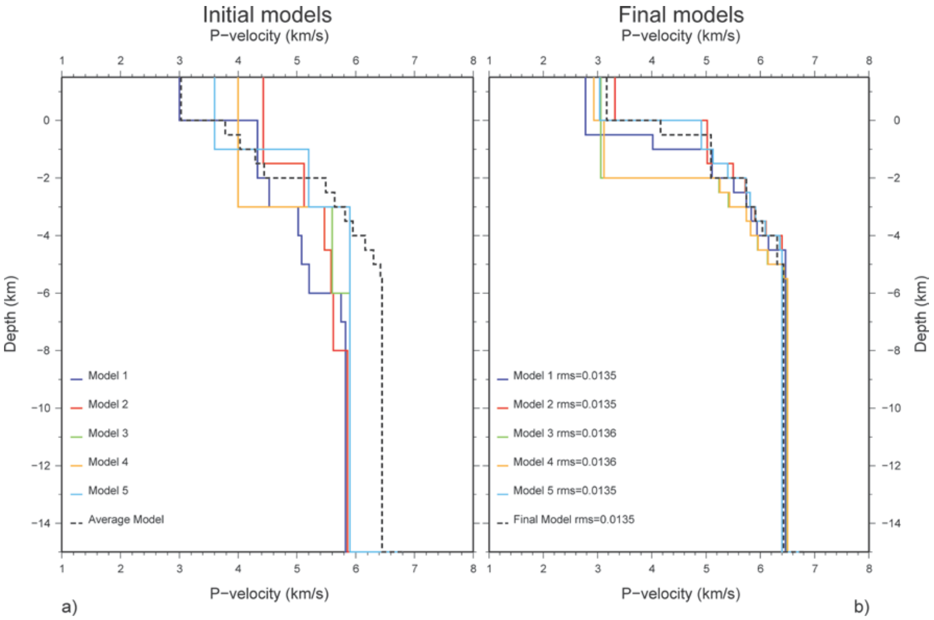

Initial models
P−velocity (km/s)
Final models
P−velocity (km/s)
1
2
3
4
5
6
7
8
0
−2
−4
−6
−8
−10
−12
−14
Depth (km)
Model 1
Model 2
Model 3
Model 4
Model 5
Average Model
Model 1 rms=0.0135
Model 2 rms=0.0135
Model 3 rms=0.0136
Model 4 rms=0.0136
Model 5 rms=0.0135
Final Model rms=0.0135
a)
b)

Figure 2

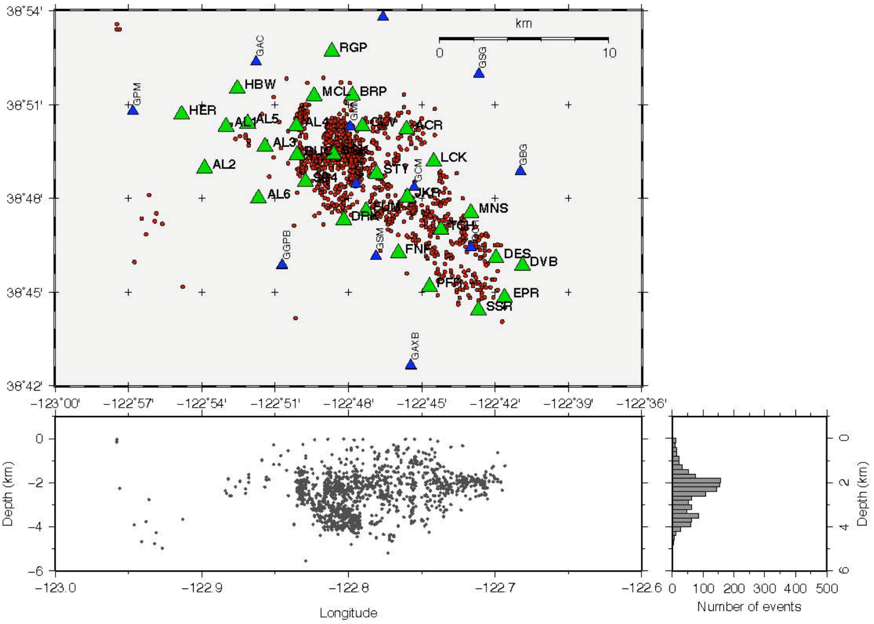

38°54'
38°51'
38°48'
38°45'
38°42'
−123°00'
−122°57'
−122°54'
−122°51'
−122°48'
−122°45'
−122°42'
−122°39'
−122°36'
km
0
10
RGP
GAC
GSG
HBW
MCL
BRP
GPM
HER
AL1
AL5
AL4
ACR
AL3
LCK
GBG
AL2
STY
GCM
AL6
JKR
DRK
MNS
GGPB
GSM
FNF
DES
DVB
PFR
EPR
SSR
GAXB
Depth (km)
0
−2
−4
−6
−123.0
−122.9
−122.8
−122.7
−122.6
Longitude
0
100
200
300
400
500
Number of events
0
2
4
6
Depth (km)

Figure 3

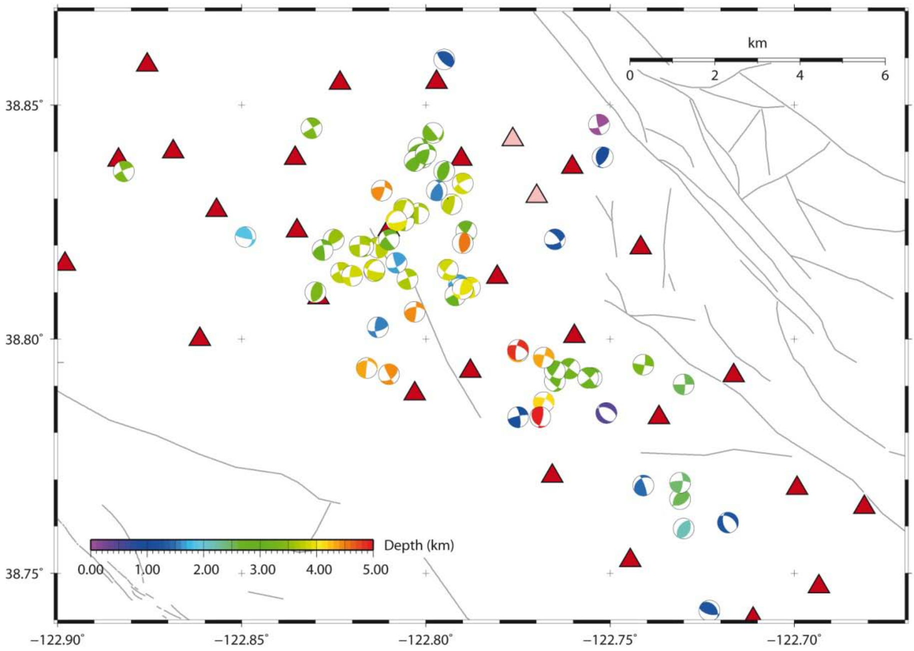

km
0
2
4
6
38.85°
38.80°
38.75°
Depth (km)
0.00
1.00
2.00
3.00
4.00
5.00
−122.90°
−122.85°
−122.80°
−122.75°
−122.70°

Figure 4

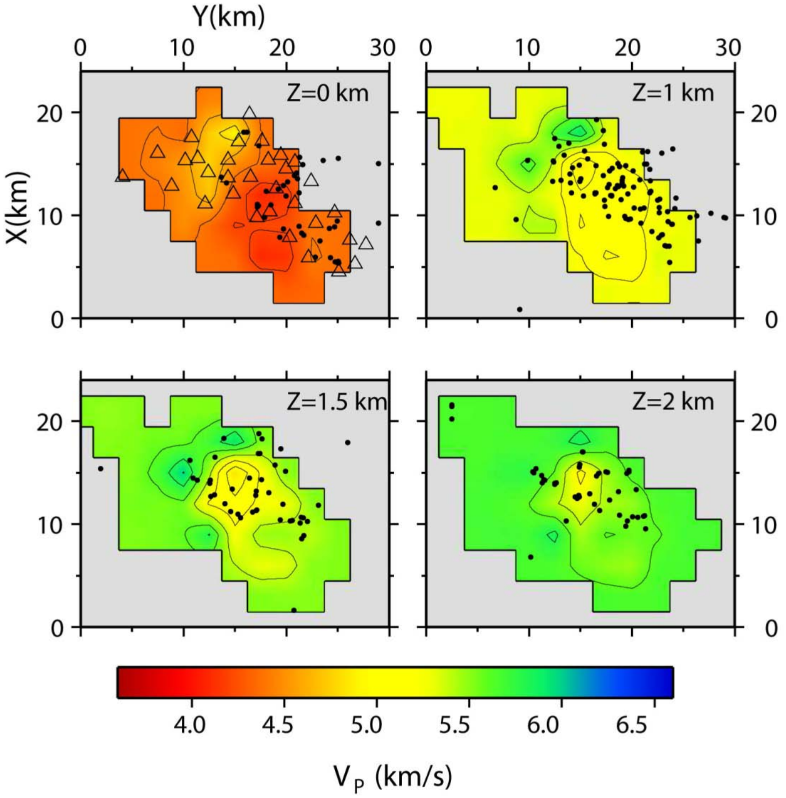


Figure 5

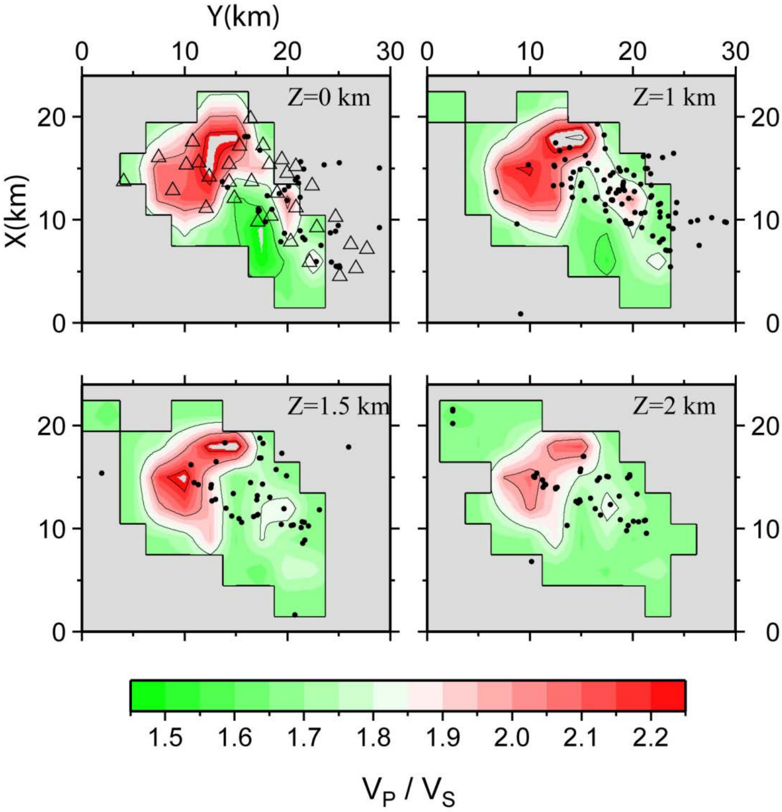
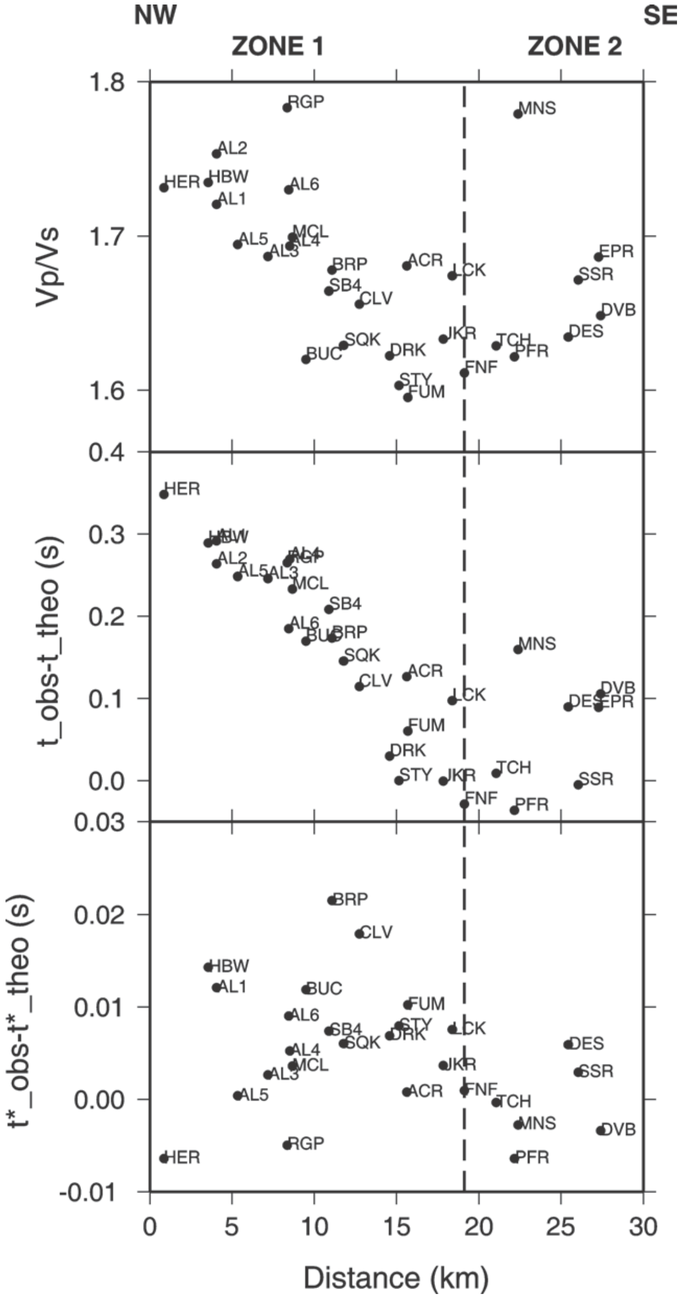

Y(km)
X(km)
0
10
20
30
Z=0 km
Z=1 km
Z=1.5 km
Z=2 km
1.5
1.6
1.7
1.8
1.9
2.0
2.1
2.2
V_P / V_S

Figure 6

581



582

583 Figure 7

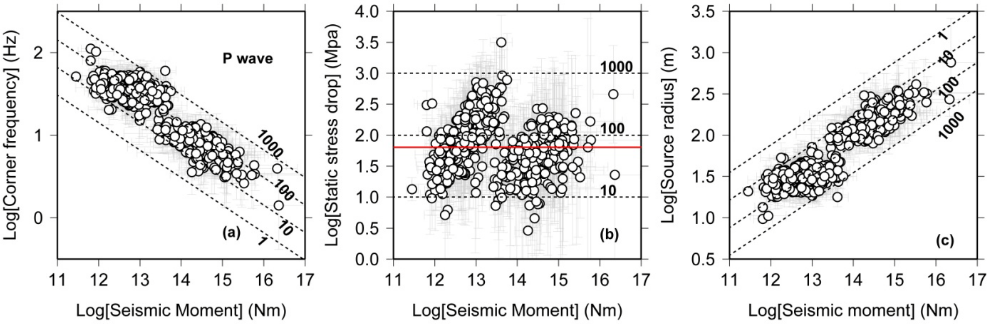

P wave
Log[Corner frequency] (Hz)
Log[Seismic Moment] (Nm)
1000
100
10
1
(a)
Log[Static stress drop] (Mpa)
Log[Seismic Moment] (Nm)
1000
100
10
(b)
Log[Source radius] (m)
Log[Seismic moment] (Nm)
1
10
100
1000
(c)

Figure 8

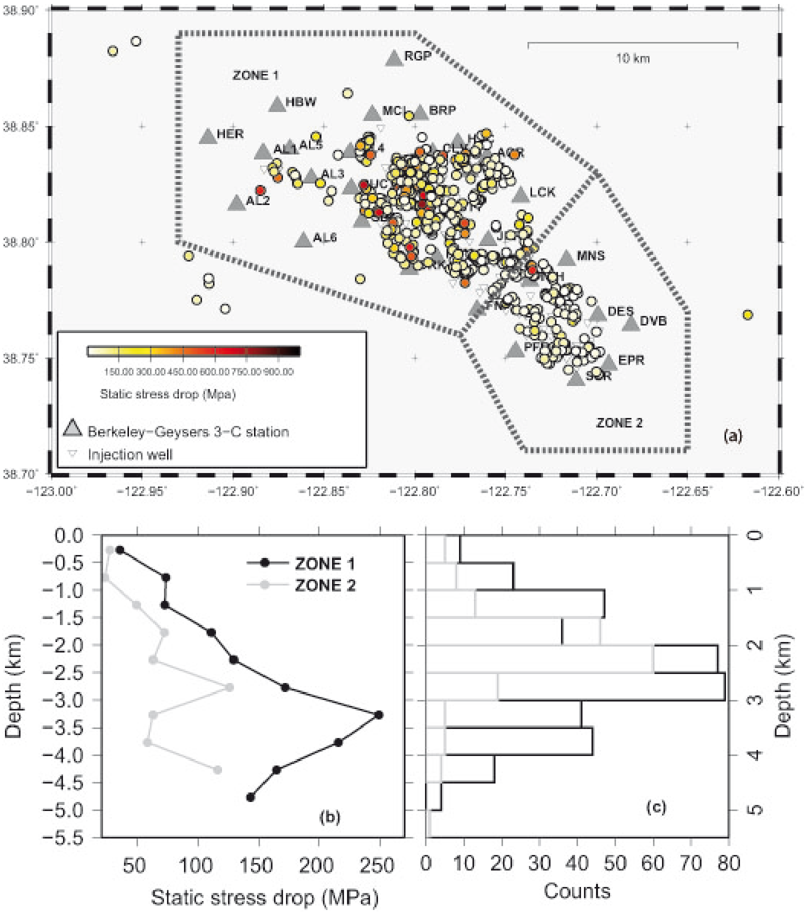

38.90°
38.85°
38.80°
38.75°
38.70°
−123.00°
−122.95°
−122.90°
−122.85°
−122.80°
−122.75°
−122.70°
−122.65°
−122.60°
ZONE 1
ZONE 2
10 km
RGP
HBW
MCL
BRP
HER
AL1
AL5
AL4
AL3
AL2
AL6
LCK
ACR
MNS
DES
DVB
EPR
SQR
PFR
JC
(a)
150.00 300.00 450.00 600.00 750.00 900.00
Static stress drop (Mpa)
Berkeley−Geysers 3−C station
Injection well
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
−4.0
−4.5
−5.0
−5.5
Depth (km)
50
100
150
200
250
Static stress drop (MPa)
ZONE 1
ZONE 2
(b)
0
10
20
30
40
50
60
70
80
Counts
0
1
2
3
4
5
Depth (km)
(c)

Figure 9

Bibliography

Abrahamson, N. A. and Silva, W. J. (1997). Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismological Research Letters*, 68(1):94–127.

Abrahamson, N. A. and Youngs, R. R. (1992). A stable algorithm for regression analysis using the ramdom effect model. *Bulletin of the Seismological Society of America*, 82(1):505–510.

Aki, K. and Richards, P. G. (1980). Quantitative seismology: Theory and methods. *San Francisco California: W H Freeman*, 1 and 2.

Ambraseys, N. N., Douglas, J., Sarma, S. K., and Smit, P. M. (2005). Equations for the estimation of strong ground motions from shallow crustal earthquakes using data from europe and the middle east: Horizontal peak ground acceleration and spectral acceleration. *Bulletin of Earthquake Engineering*, 3(1):1–53.

Anderson, J. G. and Brune, J. N. (1998). Non-ergodic probabilistic seismic hazard analysis. *Seismological Research Letters*, 69(2):71.

Anderson, J. G. and Brune, J. N. (1999a). Methodology for using precarious rocks in nevada to test seismic hazard models,. *Bulletin of the Seismological Society of America*, 89:456–467.

Anderson, J. G. and Brune, J. N. (1999b). Probablistic seismic hazard analysis without ergodic assumption. *Seismological Research Letters*, 70:19–28.

Anderson, J. G. and Brune, J. N. (2000). Probablistic seismic hazard analysis: improving consistency with precarious rock observation by removing the ergodic assumption. *12th World Conference on Earthquake Engineering*, Auckland, New Zealand:1548.

Anderson, J. G. and Luco, J. E. (1983). Consequences of slip rate constraints on earthquake occurrence relation. *Bulletin of the Seismological Society of America*, 73:471–496.

Aster, R., Meyer, R. P., De Natale, G., Zollo, A., Martini, M., Del Pezzo, E., Iannaccone, G., and Scarpa, R. (1989). Seismic investigation of the campi flegrei: A synthesis and summary of results. *in Volcanic Seismology*, K. Aki, P. Gasparini, and R. Scarpa (Editors), Volcanic series III, Springer-Verlag.

Atik, A. L., Abrahamson, N. A., Bommer, J. J., Scherbaum, F., Cotton, F., and Kuehn, N. (2010). The variability of ground-motion prediction models and its components. *Seismological Research Letters*, 81(5):794–801.

Atkinson, G. M. and Boore, D. M. (2011). Modification to existing ground-motion prediction equation equations in light of new data. *Bulletin of the Seismological Society of America*, 101:1121–1135.

Atkinson, M. (2006). Single-station sigma. *Bulletin of the Seismological Society of America*, 96(2):446–455.

Bachmann, C. E., Wiemer, S., Woessner, J., and Hainzl, S. (2011). Statistical analysis of the induced basel 2006 earthquake sequence: introducing a probability-based monitoring approach for enhanced geothermal systems. *Geophysical Journal International*, 186(2):793–807.

Baisch, S., Weidler, R., Voros, R., Wyborn, D., and Graaf, L. d. (2006). Induced seismicity during the stimulation of a geothermal hfr reservoir in the cooper basin, australia. *Bulletin of the Seismological Society of America*, 96(6):2242–2256.

Beall, J. J. and Wright, M. C. (2010). Southern extent of the geysers high temperature reservoir based on seismic and geochemical evidence, geothermal resources council transactions. *Geothermal Resources Council Transactions*, 34:53–56.

Beall, J. J., Wright, M. C., S, P. A., and Atkinson, P. (2010). Effect of high injection rate on seismicity in the geysers. *Geothermal resources council transactions*, 34:47–52.

Berrill, J. B. (1975). A study of high-frequency strong ground motion from the san fernando earthquake. *Ph.D. thesis California Institute of Technology, Pasadena, California.*

Bommer, J. J. and Alarcon, J. E. (2006). The prediction and use of peak ground velocity. *Journal of Earthquake Engineering*, 10(1):1–31.

Bommer, J. J., Douglas, J., and Strasser, O. F. (2003). Style-of-faulting in ground-motion prediction equations. *Bulletin of Earthquake Engineering*, 1(2):71–123.

Bommer, J. J., Stafford, P. J., Alarcon, J. E., and Akkar, S. (2007). The influence of magnitude range on empirical ground-motion prediction. *Bulletin of the Seismological Society of America*, 97(6):2152–2170.

Borcherdt, R. D. (1944). Estimates of site-dependent response spectra for design (methodology and justification). *Earthquake Spectra*, 10(4):617–653.

BuildingSeismicSafetyCouncil (2000). *NEHRP recommended provisions for seismic regulations for new buildings and other structures.*

BuildingSeismicSafetyCouncil (2004). *NEHRP Recommended Provisions for seismic regulations for new buildings and other structures*, volume Building Seismic Safety Council, National Institute of Building Sciences. Washington D.C.

Cantore, L., Convertito, V., and Zollo, A. (2010). Development of a site-conditions map for the campania-lucania region (southern apennines, italy). *Annals of Geophysics*, 53(4):ag–4648.

Charlety, J., Cuenot, N., Dorbath, L., Dorbath, C., Haessler, H., and Frogneux, M. (2007). Large earthquakes during hydraulic stimulations at the geothermal site of soultz-sous-forets. *International Journal of Rock Mechanics and Mining Sciences*, 44(8):1091–1105.

Convertito, V. and Herrero, A. (2004). Influence of focal mechanism in probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America*, 94(6):2124–2136.

Convertito, V., Maerklin, N., Sharma, N., and Zollo, A. (2012). From induced seismicity to direct time-dependent seismic hazard. *Bulletin of the Seismological Society of America*, 102:2563–2573.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58:1583–1606.

De Natale, G., Faccioli, E., and Zollo, A. (1988). Scaling of peak ground motions from digital recordings of small earthquakes at campi flegrei, southern italy. *Pure and Applied Geophysics*, 126(1):37–53.

De Natale, G., Troise, C., Pingue, F., Mastrolorenzo, G., Pappalardo, L., Battaglia, M., and Boschi, E. (2006). The campi flegrei caldera: Unrest mechanisms and hazards. *in Mechanisms of Activity and Unrest at Large Calderas*, C. Troise, G. De Natale, and C. R. J. Kilburn (Editors), Geologocal Society London 269:25–45.

De Natale, G. and Zollo, A. (1986). Statistical analysis and clustering features of the phlegraean fields earthquake sequence, may'83-may'84. *Bulletin of the Seismological Society of America*, 76:801–814.

Deino, A. L., Orsi, G., Piochi, M., and S, V. D. (2004). The age of the neapolitan yellow tuff caldera-forming eruption (campi flegrei caldera-italy) assessed by 40ar/39ar dating method. *Journal of Volcanology and Geothermal Research*, 133(1-4):157–170.

Dost, B. and Haak, H. W. (2007). Natural and induced seismicity, in wong,t. e., and j. de jager (editors). Geology of the Netherlands Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands:223–239.

Douglas, J. (2007). On the regional dependence of earthquake response spectra. *ISET, Journal of Earthquake Technology*, 71(1):71–99.

Douglas, J., Edwards, B., Convertito, V., Sharma, N., Tramelli, N., Kraaijpoel, D., Cabrera, B. M., Maerklin, N., and Troise, C. (2013). Predicting ground motion from induced earthquakes in geothermal areas. *Bulletin of the Seismological Society of America*, 103(3):xx–xx.

Douglas, J. and Jousset, P. (2011). Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes. *Seismological Research Letters*, 82(4):504–508.

Draper, N. R. and Smith, H. (1996). Applied regression analysis. Third Edition.

Dvorak, J. J. and Mastrolorenzo, G. (1990). History of vertical movement in pozzuoli bay, southern italy: The result of regional extension related to evolution of the tyrrhenian sea and of local volcanic activity. *Special paper, Geological Society of America*, 263:47.

Eberhart-Phillips, D. and Oppenheimer, D. H. (1984). Induced seismicity in the geysers geothermal area, california. *Journal of geophysical research*, 89(2):1191–1207.

Emolo, A., Convertito, V., and Cantore, L. (2010). Ground-motion predictive equations for low-magnitude earthquakes in the campania–lucania area, southern italy. *Journal of Geophysics and Engineering*, 8(1):46–60.

Emolo, A., Convertito, V., and Cantore, L. (2011). Ground motion predictive equations for low-magnitude earthquakes in the campania lucania area, southern italy. *Journal of Geophysics and Engineering*, 8(1):46–60.

Enedy, S. L., Enedy, K. L., and Maney, J. (1992). Reservoir response to injection in the southeast geysers. *Geothermal Resources Council, Special Report*, 17:211–219.

Evans, D. M. (1966). The denver area earthquakes and the rocky mountain arsenal disposal well. *The Mountain Geologist*, 3(1):23–26.

Fah, D., Fritsche, S., Poggi, V., Gassner-Stamm, G., Kastli, P., Burjanek, J., Zweifel, P., Barman, S., Clinton, J., Keller, L., Renault, P., and Heuberger, S. (2009). Determination of site information for seismic stations in switzerland. *Tech. Rep. SED/PRP/R/004/20090831*, Swiss Seismological Service, for the swiss nuclear Pegasos Refinement Project.

Foulger, G. R., Grant, C. C., Julian, B. R., and Ross, A. (1992). Industrially induced changes in earth structures at the geysers geothermal area, california. *Geophysical Research Letters*, 24:135–137.

Foulger, G. R., Grant, C. C., Ross, A., and Julian, B. R. (1997). Industrially induced changes in earth structure at the geysers geothermal area, california. *Geophysical Research Letters*, 24(2):135–137.

Guttenberg, B. and Richter, C. F. (1944). Frequency of earthquakes in california. *Bulletin of the Seismological Society of America*, 34(4):1985–1988.

Hanks, T. C. and Kanamori, H. (1979). Moment magnitude scale. *Journal of Geophysical Research*, 84(B5):2348–2350.

Hardin, B. O. and Drnevich, V. P. (1972). Shear modulus and damping in soils: Design equations and curves. *Journal of the Soil Mechanics and Foundations Division, ASCE*, 98(SM7):667–692.

Havenith, H. B., Fah, D., Polom, U., and Roulle, A. (2007). S-wave velocity measurements applied to the seismic microzonation of basel, upper rhine graben. *Geophysical Journal International*, 170:346–358.

Jousset, P. and Francois, B. (2006). Set-up of a seismological broadband network at hengill geothermal field (iceland). *Final report RP-54971-FR*, BRGM, OrÃ©ans, France,(18 ill):54.

Joyner, W. B. and Boore, D. M. (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley, california, earthquake. *Bulletin of the Seismological Society of America*, 71:2011–2038.

Knoll, P. and Kuhnt, W. (1990). Seismological and technical investigations of the mechanics of rockbursts. *Rockbursts and Seismicity in Mines*, pages 129–138.

Knopoff, L. (1964). Q. *Reviews of Geophysics*, 2:625–660.

Li, J., Kuleli, H. S., Zhang, H., and Toksoz, M. N. (2011). Focal mechanism determination of induced microearthquakes in an oil field using full waveforms from shallow and deep seismic networks. *Geophysics*, 76(6):WC87–WC101.

Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, pages 673–687.

MacKay, D. J. C. (2003). Information theory, inference and learning algorithms. *Cambridge University Press*, (ISBN0521642981).

Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J. J., Smith, B., and Asanuma, H. (2007). Induced seismicity associated with enhanced geothermal systems. *Geothermics*, 36(3):185–222.

Majer, E. L. and McEvilly, T. V. (1979). Seismological investigation at the geysers geothermal field. *Geophysiscs*, 44:246–269.

Majer, E. L. and Peterson, J. E. (2005). Application of micro-earthquake monitoring for evaluating and managing the effects of fluid injection at naturally fractured egs sites. *Geothermal Resources Council Transactions*, 29:103–107.

Makropoulos, K. C. and Burton, W. (1983). Seismic risk of circum-pacific earthquakes i. strain energy release. *Pure and Applied Geophysics*, 121:247–266.

Massa, M., Morasca, P., Moratto, L., Marzorati, S., Costa, G., and Spallarossa, D. (2008). Empirical ground-motion prediction equations for northern italy using weak- and strong-motion amplitudes, frequency content, and duration parameters. *Bulletin of the Seismological Society of America*, 98(3):1319–1342.

Maury, R. M. V., Grasso, R. J., and Wittlinger, G. (1992). Monitoring of subsidence and induced seismicity in the lacq gas field (france): The consequences on gas production and field operation. *Engineering Geology*, 2(3):123–135.

McGarr, A. (1984). Scaling of ground motion parameters, state of stress, and focal depth. *Journal of Geophysical Research*, 89(B8):6969–6979.

McGuire, R. K. and Arabasz, W. J. (1990). An introduction to seismic hazard analysis. *Geothechnical and Environmental Geophysics*, 1:333–353.

Menke, W. (1990). Geophysical data analysis: Discrete inverse theory. *International Geophysics Series*, 45.

Nicholson, C. and Wesson, R. L. (1992). Triggered earthquakes and deep well activities. *Pure and Applied Geophysics*, 139:561–578.

Nutti, O. W. (1981). Similarities and differences between western and eastern united states earthquakes, and their consequences for earthquake engineering. *Earthquakes and Earthquake Engineering: The Eastern United States*, In J.E. Bearvers, ed.:25–51.

Oppenheimer, D. C. (1986). Extensional tectonics at the geysers geothermal area, california. *Journal of Geophysical Research*, 91(11):11,463–11,476.

Reiter, L. (1990). Earthquake hazard analysis-issues and insights. *Columbia university press New York*, page 254.

Ross, A., Foulger, G. R., and Julian, B. R. (1999). Source processes of industrially-induced earthquakes at the geysers geothermal area, california. *Geophysics*, 64(6):1877–1889.

Seed, H. B. and Idriss, I. M. (1969). Influence of soil conditions on ground motions during earthquakes. *Journal of the Soil Mechanics and Foundations Division, ASCE*, 95(SMI):99–37.

Simpson, D. W. and Leith, W. (1985). The 1976 and 1984 gazli, ussr, earthquakes- were they induced? *Bulletin of the Seismological Society of America*, 75(5):1465–1468.

Smith, J. L. B., Beall, J. J., and A, S. M. (2003). Induced seismicity in the se geysers field. *Geothermal Resources Council Transactions*, 24:24–27.

Stark, M. A. (1992). Micro-earthquakes a tool to track injected water in the geysers reservoir. *Geothermal Resources Council, Special Report*, 17:111–117.

Stark, M. A. (2003). Seismic evidence for a long-lived enhanced geothermal systems (egs) in the northern geysers reservoir. *Geothermal resources council transactions*, 24:24–27.

Strasser, F. O., Abrahamson, N. A., and Bommer, J. J. (2009). Sigma: Issues, insight and challenges. *Seismological Research Letters*, 80(1):40–54.

Utsu, T. (1992). On seismicity, in: Report of cooperative research of the institute of statistical mathematics. *Mathematical Seismology VII, Annals of the Institute of Statistical Mathematics*, 34:139–157.

Van Eck, T., Goutbe, F., Haak, H., and Dost, B. (2006). Seismic hazard due to small-magnitude, shallow-source,induced earthquakes in the netherlands. *Engineering Geology*, 87:105–121.

Vere-Jones, D. and Ozaki, T. (1982). Some examples of statistical estimation applied to eathquake data.i.cyclic poisson and self-exciting models. *Annals of the Institute of statistics and mathematics*, 34(Part B):189–207.

Wald, D. J. and Allen, T. I. (2007). Topographic slope as a proxy for seismic site conditions and amplification. *Bulletin of the Seismological Society of America*, 97(5):1379–1395.

Wald, D. J., Quitoriano, V., Heaton, T. H., Kanamori, H., Scrivner, C. W., and Worden, C. B. (1999). rinet shakemaps: Rapid generation of instrumental ground motion and intensity maps for earthquakes in southern california. *Earthquake Spectra*, 15:537–555.