ELECTRIC POWER RESEARCH INSTITUTE

Program on Technology Innovation: Truncation of the Lognormal Distribution and Value of the Standard Deviation for Ground Motion Models in the Central and Eastern United States

1013105

Program on Technology Innovation: Truncation of the Lognormal Distribution and Value of the Standard Deviation for Ground Motion Models in the Central and Eastern United States

1013105

Technical Update, February 2006

EPRI Project Managers
R. Kassawara
L. Sandell

Cosponsor
U.S. Department of Energy
Office of Nuclear Energy
Sciences & Technology
19901 Germantown Road, NE-20
Germantown, MD 20874-1290

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1395 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

DISCLAIMER OF WARRANTIES AND LIMITATION OF LIABILITIES

THIS DOCUMENT WAS PREPARED BY THE ORGANIZATION(S) NAMED BELOW AS AN ACCOUNT OF WORK SPONSORED OR COSPONSORED BY THE ELECTRIC POWER RESEARCH INSTITUTE, INC. (EPRI). NEITHER EPRI, ANY MEMBER OF EPRI, ANY COSPONSOR, THE ORGANIZATION(S) BELOW, NOR ANY PERSON ACTING ON BEHALF OF ANY OF THEM:

(A) MAKES ANY WARRANTY OR REPRESENTATION WHATSOEVER, EXPRESS OR IMPLIED, (I) WITH RESPECT TO THE USE OF ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT, INCLUDING MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, OR (II) THAT SUCH USE DOES NOT INFRINGE ON OR INTERFERE WITH PRIVATELY OWNED RIGHTS, INCLUDING ANY PARTY'S INTELLECTUAL PROPERTY, OR (III) THAT THIS DOCUMENT IS SUITABLE TO ANY PARTICULAR USER'S CIRCUMSTANCE; OR

(B) ASSUMES RESPONSIBILITY FOR ANY DAMAGES OR OTHER LIABILITY WHATSOEVER (INCLUDING ANY CONSEQUENTIAL DAMAGES, EVEN IF EPRI OR ANY EPRI REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES) RESULTING FROM YOUR SELECTION OR USE OF THIS DOCUMENT OR ANY INFORMATION, APPARATUS, METHOD, PROCESS, OR SIMILAR ITEM DISCLOSED IN THIS DOCUMENT.

ORGANIZATION(S) THAT PREPARED THIS DOCUMENT

Norman A. Abrahamson, Inc.

Imperial College London Consultants

This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc.

Copyright © 2006 Electric Power Research Institute, Inc. All rights reserved.

CITATIONS

This document was prepared by

Norman A. Abrahamson, Inc.
152 Dracena Avenue
Piedmont, CA 94611

Principal Investigator
N. Abrahamson

Imperial College London Consultants
47 Prince's Gate
Exhibition Road
London SW7 2QA

Principal Investigator
J. Bommer

This document describes research sponsored by the Electric Power Research Institute (EPRI) and the U.S. Department of Energy (US DOE) under Award No. (DE-FC07-04ID14533). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of Energy.

This publication is a corporate document that should be cited in the literature in the following manner:

Program on Technology Innovation: Truncation of the Lognormal Distribution and Value of the Standard Deviation for Ground Motion Models in the Central and Eastern United States. EPRI, Palo Alto, CA, and US DOE, Washington, DC: 2006. 1013105.

ABSTRACT

In probabilistic seismic hazard analyses, the ground motion variability has a large effect on the computed hazard. The ground motion variability is usually described by a lognormal distribution. It has been common practice in PSHA to truncate the lognormal distribution for ground motion at a specified number of standard deviations (epsilons). Typically, maximum epsilons of 2 to 3 have been used; however, there has not been a strong technical basis for the selection of the maximum epsilon. The computed hazard is reduced as the ground motion distribution is truncated at smaller epsilon values so the choice of a maximum epsilon can have a significant effect on the computed hazard.

One objective of this study was to find a technical basis for the selection of the maximum epsilon value to use based on empirical ground motion data and numerical simulations. This study found that there is no basis for truncating the ground motion distribution at an epsilon value of less than 3 and there are observations of epsilon values greater than 3. We conclude that using an untruncated lognormal ground motion distribution in probabilistic seismic hazard analyses is appropriate for ground motion values that are below the physical limits of the underlying rock or soils. Physical limits are not addressed in this study but are the topic of other ongoing DOE sponsored research.

A second issue that was investigated in this project is the value of the standard deviation for the ground motion variability for the central and eastern United States (CEUS). The value of the standard deviation in the models developed in the EPRI (2004) ground motion study is much larger than recent studies of large data sets of ground motions applicable to the WUS have shown. An evaluation of differences in the standard deviation in the CEUS and WUS based on the variability of the source, path, and site terms indicates that the WUS intra-event standard deviations are generally applicable to the CEUS with some epistemic uncertainty in the effect of focal depth at short distances and that the inter-event standard deviations may be larger in the CEUS than in the WUS based on larger variability in the stress-drops. Alternative models for the total standard deviation (combined intra-event and inter-event) are developed that can be applied to the CEUS. Overall, these new models show a significant reduction in the standard deviation, particularly at short distances. This lower value of the standard deviation will tend to reduce the computed hazard as compared to the EPRI (2004) models.

CONTENTS

1 INTRODUCTION 1-1

2 INFLUENCE OF GROUND MOTION VARIABILITY ON PSHA 2-1
Nature of the ground motion variability 2-1
Sigma and Epsilons 2-7
The Influence of Sigma and Epsilon on Hazard Curves 2-8

3 COMPONENTS OF GROUND MOTION VARIABILITY 3-1
Inter-event and Intra-event variability 3-1

4 MAXIMUM EPSILON VALUES 4-1
Observations of large positive epsilons 4-1
Statistical analysis of large positive epsilons 4-3
Postulated causes of large positive epsilons 4-10
Maximum epsilons from numerical simulations 4-13
Recommended maximum epsilon 4-17

5 ESTIMATION OF SIGMA 5-1
NGA results for WUS 5-1
Distance Dependence 5-5
Sigma from K-NET and KiK-NET data 5-11
Sigma from numerical simulations 5-16

6 SIGMA FOR EUS 6-1
EUS sigma models used in 2004 EPRI EUS ground motion study 6-1
Applicability of sigma from other regions to EUS 6-3
Event-term variability inferred from catalogs 6-3
Data Sources 6-4
Tectonic regions 6-4
Source Variability 6-5
Source-term variability due to limited source parameters 6-8
Path-term variability inferred from simulations (URS) 6-8
Standard Deviation in Ln units 6-15
Site-term variability 6-16
ENA Network Data 6-18
Recommended CEUS sigma 6-20

7 CONCLUSIONS 7-1

8 REFERENCES 8-1

LIST OF FIGURES

Figure 2-1 Ground motions and residuals from the 1994 Northridge earthquake…..…..…. 2-2

Figure 2-2 Residuals of T=0.5 sec spectral accelerations from the equation of Ambraseys et al. (1996) grouped by site classification……..…..…..…..…..……2-2

Figure 2-3 Normal probability plot prepared from the logarithmic residuals of the strong-motion dataset used to derive the equation of Berge-Thierry et al. (2003) for the prediction of PGA……..…..…..…..…..…..…..…..…..…..…….. 2-3

Figure 2-4 Predicted PGA values from an earthquake of M_S 6.4 at rock sites as a function of distance; in the left-hand panel, the numbers indicate the percentiles of the normal distribution and the numbers in paraentheses are the number of logarithmic standard deviations difference from the logarithmic mean…….….. 2-4

Figure 2-5 Total aleatory variability associated with equations for the prediction of PGA, expressed as the ratio of 84th-percentile values to median values…….……2-5

Figure 2-6 PGA residuals from the Italian ground-motion prediction of equation of Sabetta and Pugliese (1996) calculated using all of the available strong-motion records and using the sub-set employed in the regressions……..…….…..…..…..….. 2-6

Figure 2-7 Dependence of total variability of magnitude from the European dataset employed in the regressions by Ambraseys *et al*. 2005).……..…… 2-7

Figure 2-8 Schematic illustration of the interaction between sigma and epsilon in determining a particular level of spectral acceleration, SA.……..…..…….. 2-8

Figure 2-9 Hazard curves derived incorporating and neglecting sigma (Abrahamson, 2000)……..…..…..…..…..…..…..…..…..…..…..…..…..…..…..……2-9

Figure 2-10 Seismic hazards derived for a site in southern California using the ground-motion prediction equation of Boore *et al*. (1997) with different values of sigma (log10 units)……..…..…..…..…..…..…… 2-10

Figure 2-11 Seismic hazard curves derived imposing truncations on the ground-motion distribution at different levels of epsilon (from Bommer et al., 2004)…..…..….. 2-10

Figure 3-1 Inter-event and Intra-event variability……..…….. 3-2

Figure 4-1 Normal probability plots for the PGA residuals of Berge-Thierry et al. (2003) and Lussou et al. (2001) equations (modified from Bommer et al., 2004)…….. 4-2

Figure 4-2 Locations with respect to whole dataset of normalized residuals from all stations available for a single event (Lazio Abruzzo, 07/05/1984, M_S=5.8) plotted against spectral acceleration *(left)*, and distance *(right)*, including site classification (Strasser & Bommer, 2004)……..…..…..…..…..…..…..…..…..…..…….4-3

Figure 4-3 Distribution of selected records in magnitude-distance space…….. 4-5

Figure 4-4 Contribution of individual events to (*left*) complete NGA dataset and (*right*) large positive epsilon selection, at several response periods……..…… 4-7

Figure 4-5 Plots of normalized inter-event residuals (δ_E) vs. normalized intra-event residuals (δ_A) of the NGA dataset, at the 15 response periods considered in this study……..…..…..…...4-9

Figure 4-6 Normalized inter-event for the Chi-Chi, Taiwan sequence…….……4-11

Figure 4-7 Layout of receivers (numbered 1 to 122) used in the simulations……..……..4-14

Figure 4-8 Quantile-quantile plot of the normalized residuals from numerical simulations..4-16

Figure 5-1 Comparison of the standard deviation for PGA from the 1997 models with the prelimianry NGA model from Abrahamson and Silva (2005)…….……..……. 5-2

Figure 5-2 Intra-event residual for PGA for the Abrahamson and Silva (1997) data set….. 5-3

Figure 5-3 Intra-event residual for PGA for the Abrahamson and Silva (2005) data set….. 5-4

Figure 5-4 Inter-event residual for PGA for the Abrahamson and Silva (1997) data set….. 5-4

Figure 5-5 Inter-event residual for PGA for the Abrahamson and Silva (2005) data set….. 5-5

Figure 5-6 Example of range of rupture distances that can occur for short JB distances… 5-6

Figure 5-7 Example of the distance dependence of the standard deviation from the Toro et al (1997) model……..…..……..……. 5-7

Figure 5-8 PGA intra-event residuals of the NGA data set for small magnitudes (M<6) and short distances (R_{JB}<10 km)……..…..…..…..…..……5-8

Figure 5-9 PGA inter-event residuals of the NGA data set for small magnitudes (M<6) and short distances (R_{JB}<10 km)……..…..……5-9

Figure 5-10 PGA total residuals of the Bommer et al. (2004) data set for distances R_{JB}<10 km…….…..…..…..…..…… 5-10

Figure 5-11 PGA total residuals of the Bommer et al. (2004) data set for distances R_{JB}<15 km…….…..…..…..…..…..….. 5-11

Figure 5-12 Distribution of data available for the selection of “well-recorded” shallow events from the K-NET and KiK-NET databases in terms of hypocentral distance…..5-13

Figure 5-13 Influence of sample size on sigma for selected “well-recorded” events from the K-NET and KiK-NET databases……...…..…..…..5-15

Figure 6-1 Comparison of the EPRI (2004) models of the standard deviation for M=5, R_{JB}=5km with the results of the Abrahamson and Silva (2005) model……6-2

Figure 6-2 Comparison of the EPRI (2004) models of the standard deviation for M=6, R_{JB}=10km with the results of the Abrahamson and Silva (2005) model…….. 6-2

Figure 6-3 Comparison of the EPRI (2004) models of the standard deviation for M=7, R_{JB}=30km with the results of the Abrahamson and Silva (2005) model…….. 6-3

Figure 6-4 Map of M>=5 earthquakes from 1978-2004, excluding subduction zones……. 6-5

Figure 6-5 Regression fits to the half-duration from the catalog data…….……6-6

Figure 6-6 Residuals for the half-duration model from the catalog data…… 6-7

Figure 6-7 EPRI Regions for the EUS…….……. 6-9

Figure 6-8 Velocity profiles for the 16 EUS regions…….……..6-10

Figure 6-9 Station locations used for the simulations for evaluations of crustal structure sensitivity……6-12

Figure 6-10 Depth distributions for the EUS and California……...…..……6-13

Figure 6-11 Simulations for the EUS crustal structures……...…..……6-14

Figure 6-12 Simulations for the WUS crustal structures……..……..6-15

Figure 6-13 V_{S30} dependence of the PGA intra-event residuals……...…..…..6-17

Figure 6-14 V_{S30} dependence of the 25 Hz intra-event residuals…….……..…..6-18

Figure 6-15 Stress parameters from ENA network data (from Atkinson, 1993)…….…..…. 6-20

LIST OF TABLES

Table 4-1 Summary of selected records …… 4-6

Table 4-2 Ratio between the event term, δ_E, and the average total residual, $<\delta_T>$, depending on the number of records N_{REG} contributed to the regression……4-8

Table 5-1 Comparisons of standard deviations from Abrahamson and Silva (1997) with the preliminary model from Abrahamson and Silva (2005)……..……5-2

Table 5-2 Variability definitions in empirical GMPE and using the numerical simulations presented in Strasser & Bommer (2005b)……..…… 5-17

Table 6-1 Standard deviation of source half durations and mb-Ms from earthquake catalogs (1978 – 2001) …….…….. 6-8

Table 6-2 Comparison of expected path variability for the EUS and CA based on crustal structures variability and differences in the source depth distributions……..6-15

Table 6-3 Recommended Standard Deviations for the CEUS……6-21

Table 6-4 Recommended Adjustment to the Standard Deviations for the CEUS at short R_{JB}……6-21

1
INTRODUCTION

In probabilistic seismic hazard analyses (PSHA), the ground motion variability has a large effect on the computed hazard (see Chapter 2 for examples). The ground motion variability can be described by a lognormal distribution and it has been common practice in PSHA applications in the Western United States (WUS) to truncate the lognormal distribution for ground motion at a specified number of standard deviations, epsilon. For non-nuclear projects, a maximum of 2 to 3 standard deviations has been commonly used; however, there has not been a sound technical basis for the selection of the truncation level.

This report deals with two aspects of the ground motion variability:

- Sigma: The standard deviation of the natural logarithm of the ground motion
- Epsilon: the number of standard deviations above the median ground motion (epsilon is negative for values below the median)

Two main issues are addressed in the report: (1) can we justify a limit to epsilon, and (2) what is the appropriate value of sigma for the EUS? These two issues are related because the empirical estimates of epsilon are correlated with the value of sigma: if the estimate of sigma does down in a new model, then the epsilon for the same recording will increase.

The first part of the report evaluates empirical and numerically simulated strong motion data to determine if there is a sound technical basis for putting an upper limit on epsilon. Chapter 2 gives background on the effects of the ground motion variability on the computed hazard and discusses the relation between the value of the standard deviation and the number of standard deviations of individual observations. Chapter 3 describes how the ground motion variability is separated into two terms (inter-event and intra-event) in modern attenuation relations. Chapter 4 reviews the residuals from empirical attenuation relations and from numerical simulations and provides recommendations on the maximum number of standard deviations to consider.

The second part of the report evaluates the appropriate value of sigma for use in the EUS. Chapter 5 evaluates the standard deviations from empirical models and numerical simulations. Chapter 6 evaluates how these standard deviations from other regions can be modified to be applicable to the CEUS. Recommended models for the standard deviation for use in the CEUS are provided.

Chapter 7 gives the recommendations of the maximum epsilon as well as models for the value of the sigma for EUS ground motion models.

2
INFLUENCE OF GROUND MOTION VARIABILITY ON PSHA

Nature of the ground motion variability

Ground-motion prediction equations (GMPE) express a given measure of the ground shaking, such as the spectral acceleration at a given response period, as a function of a few variables that characterize the earthquake source, the source-to-site travel path and the effect of the site geology and topography. The models are invariably simple, usually only including four simple variables to model the influence of all of the complex features of the generation, propagation and attenuation of seismic waves that lead to a particular value of the ground-motion measure at a given site. The earthquake source is characterized by the magnitude of the event and, in most cases, the fault rupture mechanism; the latter variable is usually classified into generic groups for different styles of faulting and a coefficient derived for each class. The travel path from the earthquake source to the site is generally characterized only by the length of this path, characterized by a distance metric defined relative either to the hypocentre or the fault rupture. The site itself is generally characterized only by grouping into generic site classes, often based on the average shear-wave velocity over the uppermost 30 m at the site. Several factors that will clearly affect the resulting ground motion at a site are therefore not included in the models: examples include the static and dynamic stress drop, rupture directivity, azimuth of the path, heterogeneties along the path, the deeper geological profile at the site, and the surface relief (topographical effects). As a result of including just a few simple variables in the models, observations of recorded ground motions display very appreciable scatter around the median values predicted by the equations (Figure 2-1). Even if a large number of predictive parameters were used, the high frequency ground motion will still have large variability due to the unpredictability of the precise ground motion that results from the set of predictive parameters.

The data shown in Figure 2-1 are the peak ground accelerations (PGA) from the 1994 Northridge earthquake; as the records are from a single event, the influence of source parameters (magnitude, style-of-faulting) can be ignored. The curve through the data shown in the left-hand panel of the figure is the best fit to the data considering only distance as an explanatory variable, and it can be appreciated that there is considerable scatter in the data. The right-hand panel of the figure displays the residuals, calculated as the difference between the observed PGA values and those predicted by the curve, grouped according to the site classification. The horizontal lines through each group represent the mean values of the residuals for each site class; the differences between these mean values show that on average peak accelerations on soft soil sites are greater than those on stiff soil sites which in turn are greater than those on rock sites, however, the variability of the data within each site class is much larger than the small differences between the mean values of the residuals for each group. Therefore, although the site classification clearly should be included in the model and does result in appreciable (and physically explicable)

differences in the predicted median accelerations, the addition of this explanatory variable has a rather minor impact on the total variability of the data with respect to the model.

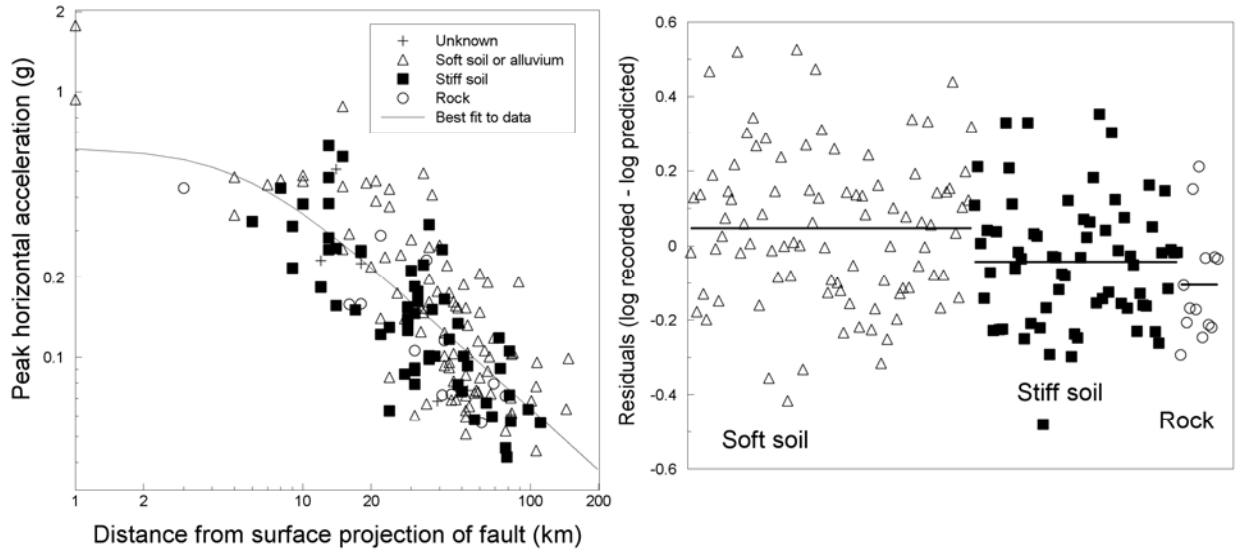

Figure 2-1
Ground motions and residuals from the 1994Northridge earthquake. Left: Best-fit curve to the recorded PGA values; right: residuals grouped by site classification (from Bommer and Boore, 2004).

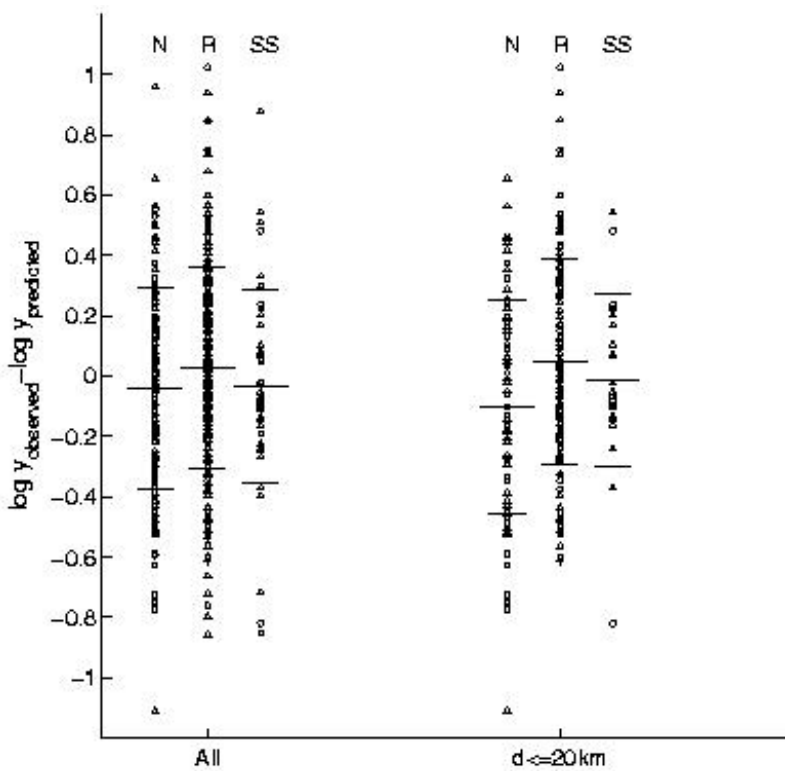

Figure 2-2
Residuals of T=0.5 sec spectral accelerations from the equation of Ambraseys et al. (1996) grouped by site classification: SS – strike-slip, N – Normal, R - reverse (from Bommer et al., 2003).

Figure 2-2 shows the results of adding style-of-faulting as a fourth variable to a ground-motion model (in addition to magnitude, distance and site classification). Here again the differences between the means of each group of residuals reveal an influence of the style-of-faulting, with reverse events producing, on average, higher ground accelerations than strike-slip events which in turn produce slightly higher motions than normal faulting events; the influence of the style-of-faulting is more pronounced when looking only at the residuals for records obtained from relatively short (< 20 km) distances; however, as for the case of site classification illustrated in Figure 2-1, the residuals in Figure 2-2 also display a variability within each style-of-faulting class that is considerably larger than the differences between the mean values of each group of residuals, whence the impact of this additional variable on the scatter is small.

The regressions to obtain empirical ground-motion prediction equations are generally performed on the logarithm of the ground-motion parameter and it is found that the residuals then conform to a normal (Gaussian) distribution. Figure 2-3 shows the residuals from a recently derived European equation displayed on a probability plot in which the distribution of the residuals is compared with that expected for a log-normal distribution; it can be seen that the residuals actually follow the lognormal distribution very closely, right up to values of ±3 standard deviations. For most ground-motion prediction equations, the residuals conform to a lognormal distribution within the range from -2 to +2 standard deviations, but some deviate from the expected distribution beyond these limits (Strasser and Bommer, 2004).

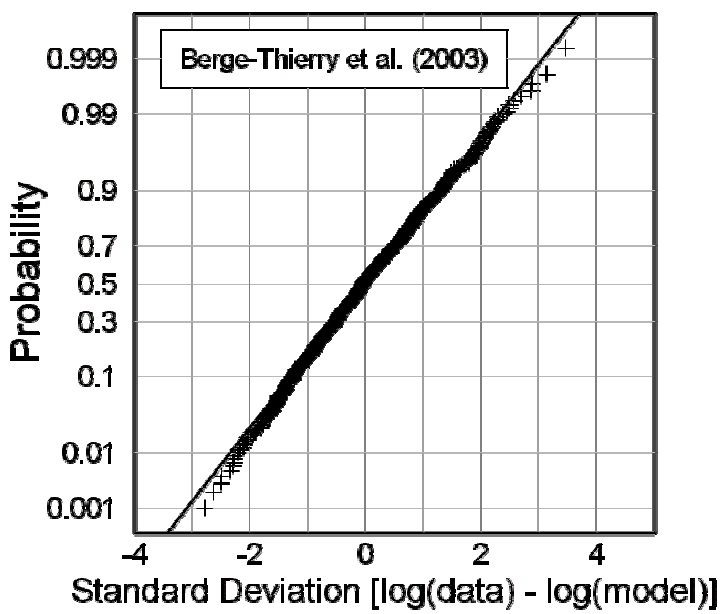

Figure 2-3
Normal probability plot prepared from the logarithmic residuals of the strong-motion dataset used to derive the equation of Berge-Thierry et al. (2003) for the prediction of PGA.

Since the random (aleatory) variability associated with ground-motion prediction equations generally conforms to a lognormal distribution, it can be characterized by the standard deviation, which is an integral and indispensable part of the predictive model. Ground-motion prediction equations are often displayed graphically as curves of the predicted median values, which can be misleading since in effect the equations predict, for a given combination of the explanatory variables (magnitude, distance and site classification) the expected distribution of the ground motion ampltiudes. This is illustrated in Figure 2-4, which displays the predicted distribution of PGA at rock sites for an earthquake of a particular magnitude. The curves are shown on logarithmic (left hand panel) and linear axes (right hand panel) of acceleration, since the latter help convey the extent of the variability. In Figure 2-4, the numbers indicate the percentiles of the normal distribution and the numbers in parentheses are the number of logarithmic standard deviations difference from the logarithmic mean. The values indicate that the median level of PGA at 1 km, which will be exceeded by half of the observations, is 0.54g but the 95-percentile value, which will be exceeded on average once in every 20 realizations of the scenario of an M_s 6.5 at 1 km, is almost three times greater.

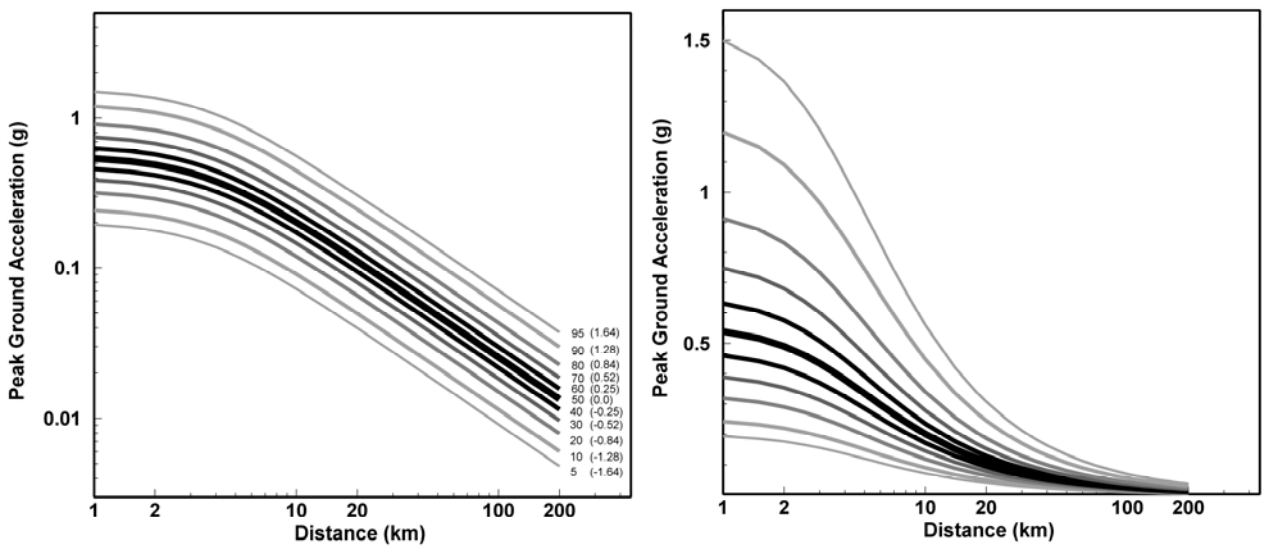

Figure 2-4
Predicted PGA values from an earthquake of M_s 6.5 at rock sites as a function of distance.

The values of the standard deviations associated with ground-motion prediction equations have remained reasonably stable since the first empirical equations were published in the early 1970s. Figure 2-5 shows the aleatory variability scale factor corresponding to one stadard deviation from a large set of ground-motion prediction equations plotted against the date of their publication. Since some equations are expressed in base-10 logarithms and others in natural logarithms, the variability is presented as the ratio of the 84-percentile PGA value to the median PGA value, the former value corresponding to that obtained by adding one logarithmic standard deviation to the logarithmic mean. Although greatly expanded datasets have been used for the more recent equations, and some of them have incorporated large numbers of explanatory variables, there has been no significant reduction in the overall variability associated with the majority of equations.

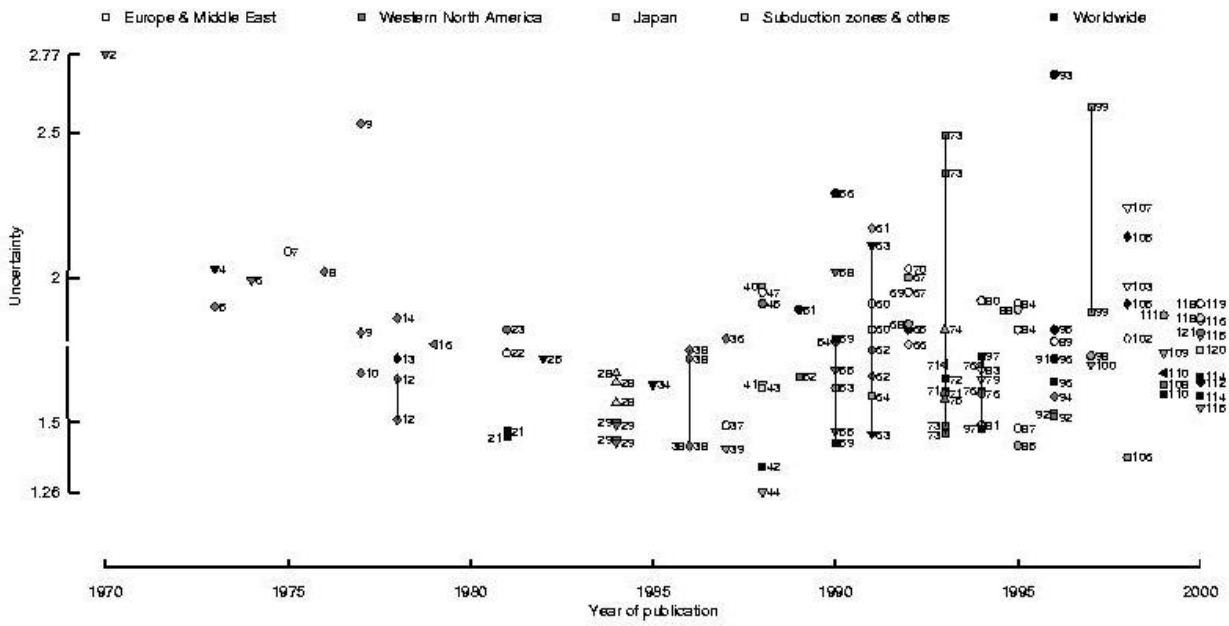

Figure 2-5
Estimates of the total aleatory variability associated with equations for the prediction of PGA. The y-axis, labeled uncertainty, is the ratio of 84-percentile values to median values and thereby represents the scale factor corresponding to 1 standard deviation (from Douglas, 2003).

Some equations have been published with surprisingly low standard deviations but there is reason to believe that in some cases this does not represent a genuine reduction of the variability but rather has been achieved by selective pruning of the dataset. Figure 2-6 shows the PGA residuals from the Italian strong-motion dataset used by Sabetta and Pugliese (1996), indicating the records that were excluded, without explanation, from the regressions, resulting in an appreciably reduced standard deviation.

Most ground-motion prediction equations include a homoscedastic model of the aleatory variability but a few studies have identified a variation of the sigma with magnitude, with smaller variability at higher magnitudes (Youngs *et al*., 1995; Abrahamson and Silva, 1997). A similar dependence of the variability on magnitude (Figure 2-7) has been found for the recent equations for Europe and the Middle East derived by Ambraseys *et al*. (2005). Campbell (1997) identified a reduction of the scatter with both increasing magnitude and increasing ground-motion amplitude; the latter effect was first identified by Donovan and Bornstein (1978) who attributed the phenomenon in part to the non-linear response of soil sites.

Midorikawa and Ohtake (2004), on the other hand, studied the variance of PGA and PGV in Japanese strong-motion data and found decreasing variance with increasing magnitude and amplitude but also an even stronger decrease with decreasing distance, concluding that the apparent reduction with increasing ground-motion amplitude is the combined effect of the magnitude and distance dependences.

Recently, the NGA project, which is updating the empirical ground motion models for California, have found a lack of a magnitude dependence as compared to the previous models for ground motion in California.

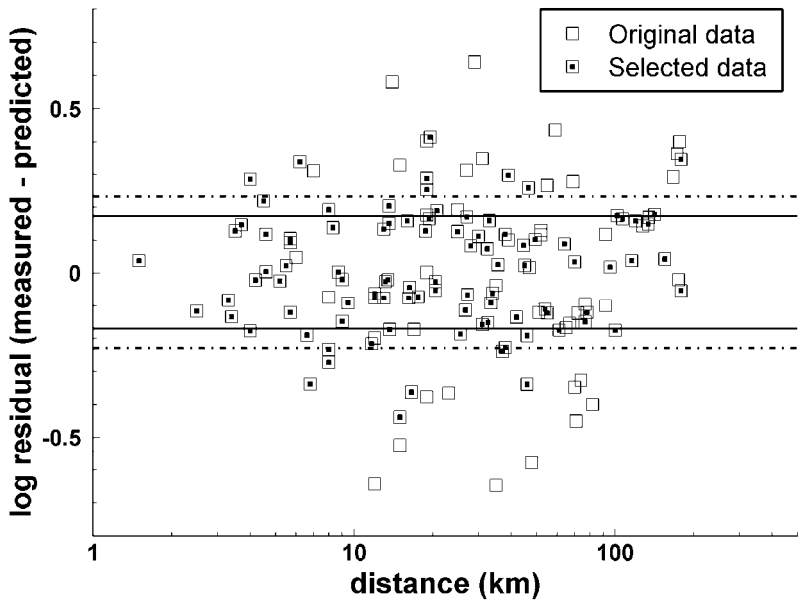

Figure 2-6
PGA residuals from the ground-motion prediction equation of Sabetta and Pugliese (1996) calculated using all of the available Italian strong-motion records and using the sub-set employed in the regressions (squares with dots). The dashed-dotted lines indicate the one standard deviation bounds calculated using the full dataset, whereas the solid lines indicate the values obtained by Sabetta and Pugliese (1996) with the selected sub-set (Bommer and Scherbaum, 2005).

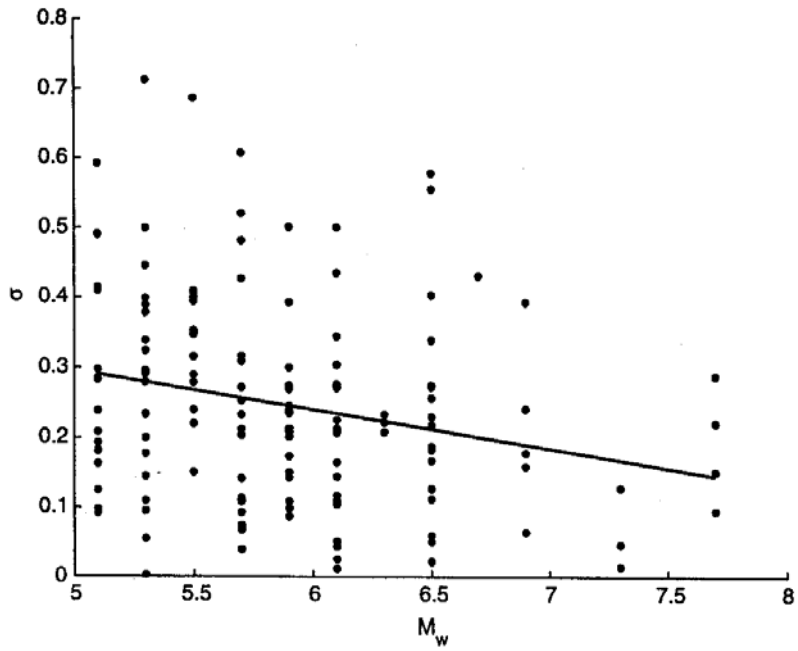

Figure 2-7
Dependence of total variability of magnitude from the European dataset employed in the regressions by Ambraseys et al. (2005).

Sigma and Epsilons

Two terms are used to specify ground-motion amplitudes with respect to the predicted median values from attenuation equations. The first is sigma (σ), which as explained above is the standard deviation of the logarithmic residuals. The second term is epsilon (ε), which is the number of standard deviations above or below the logarithmic mean prediction. A residual, δ, is defined as the difference between an observed value and the predicted value from model; in effect, epsilon can be thought of as a normalized residual, i.e.:

$$\varepsilon = \frac{\delta}{\sigma} \qquad \text{(Equation 2-1)}$$

Figure 2-8 illustrates, schematically, how, for a given ground-motion prediction equation (in terms of median values), the level of ground motion depends on both sigma and epsilon. For an equation with a typical (i.e. large) standard deviation, a relatively high value of ground acceleration will require only a moderate value of epsilon. On the other hand, if the variability associated with the same equation were significantly reduced, then more standard deviations would need to be added to the mean value to reach the same level of acceleration; in other words, for a smaller sigma, a larger epsilon would be required to achieve the same level of acceleration.

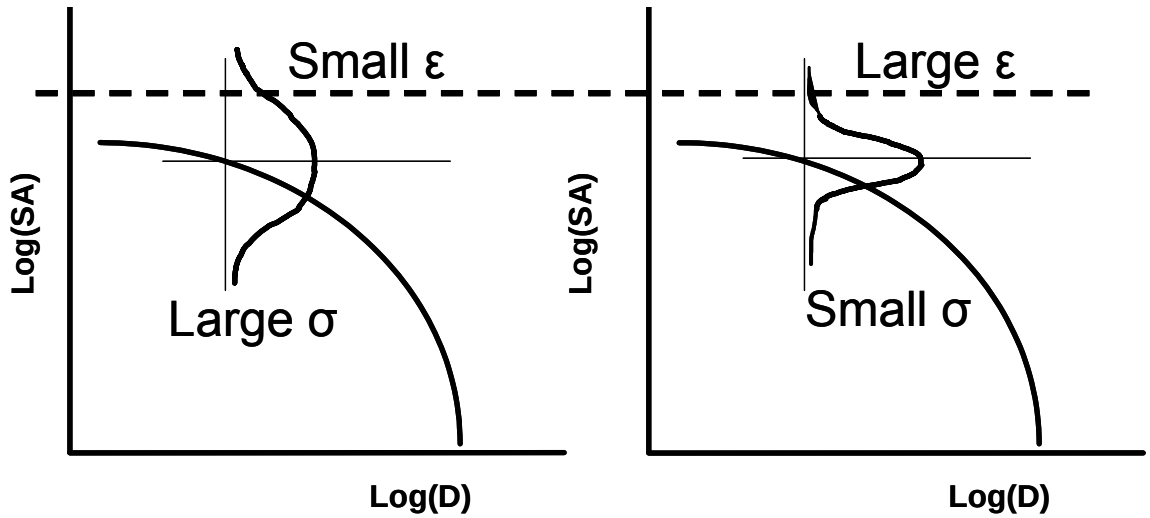

Figure 2-8
Schematic illustration of the interaction between sigma and epsilon in determining a particular level of spectral acceleration, SA.

Since the residuals are described by a probability distribution, the probability of reaching the given level of acceleration will be smaller in the second case as a result of the higher value of epsilon. If $\varepsilon = 1$, the level of acceleration has 16% probability of being exceeded, since this defines the 84-percentile level of motion. If ε is increased to 2, thus defining the 97.7-percentile, the resulting ground-motion amplitude has a probability of only 2.3% of being exceeded.

The Influence of Sigma and Epsilon on Hazard Curves

The value of sigma exerts a strong effect on the outcome of a probabilistic seismic hazard analysis (PSHA). Indeed, if sigma were equal to zero, the hazard would actually saturate at some point (when the ground-motion amplitudes reached the median values from the maximum earthquake magnitude at the closest possible position to the site) and it is only because of this variability in the ground motion that seismic hazard curves continue to move to higher and higher values of acceleration as the annual frequency of exceedance is reduced (Figure 2-9).

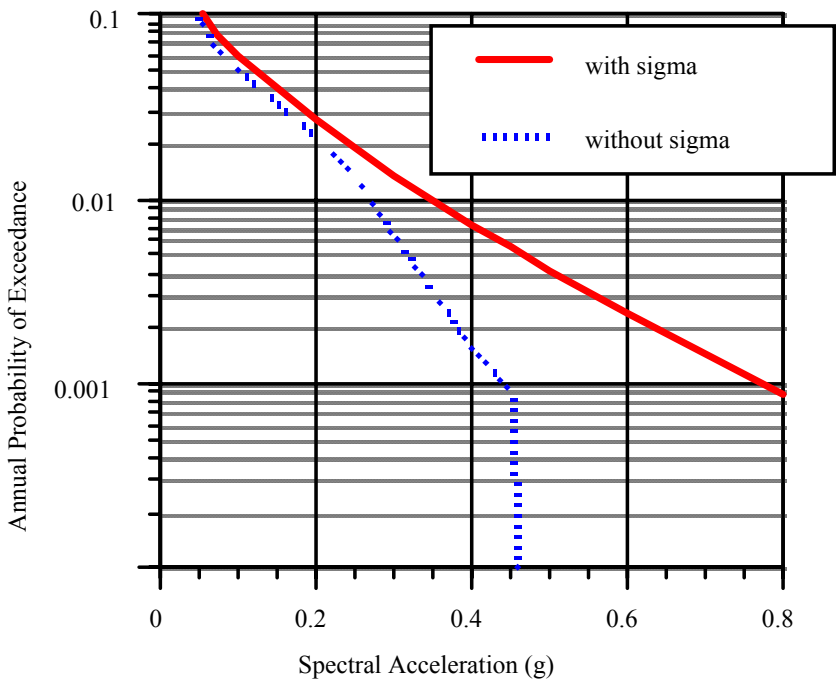

Figure 2-9
Hazard curves derived incorporating and neglecting sigma (Abrahamson, 2000).

A hazard curve derived neglecting sigma cannot be considered to constitute a PSHA. The actual value of sigma, for a given level of seismicity, effectively controls the shape of the seismic hazard curve, as illustrated in Figure 2-10. A common practice in seismic hazard analysis is to include sigma in the calculations but to truncate the distribution at a given number of standard deviations above the median; in other words, sigma is incorporated into the hazard integral but a maximum value is imposed on epsilon. Values of the order of 3 are sometimes assigned for the maximum that epsilon can assume, and some researchers have even proposed truncating the distribution of residuals by imposing a maximum value of 2 on epsilon: e.g. Romeo and Prestininzi (2000) argued that "stronger motions are considered to be unlikely". The issue is precisely that when hazard is calculated for low annual frequencies of exceedance, 'unlikely' events need to be considered since a very low annual exceedance frequency can only be achieved by combining a very infrequent earthquake with a large value of epsilon. Figure 2-11 shows an example of how seismic hazard curves can be affected by truncation at different levels of epsilon.

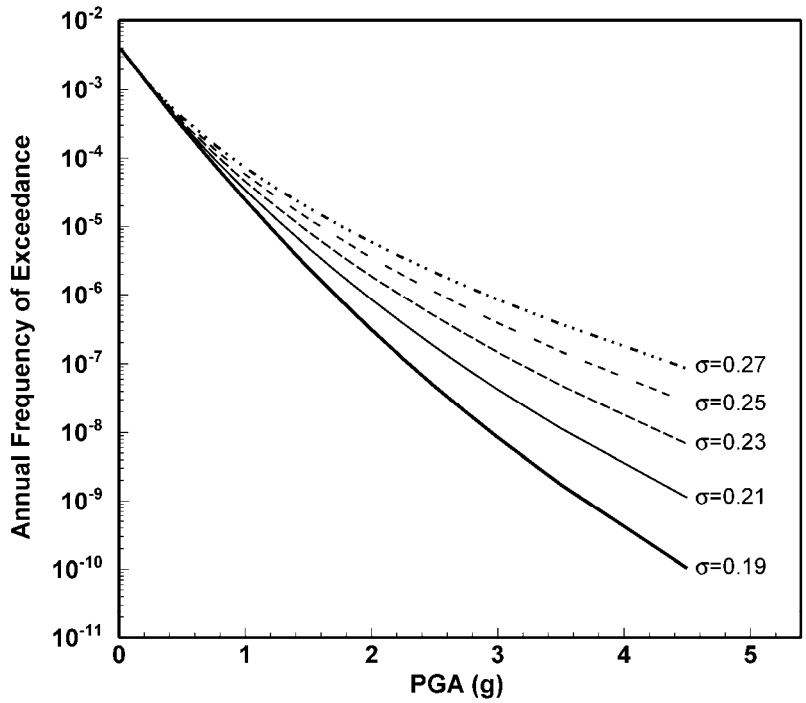

Figure 2-10
Seismic hazards derived for a site in southern California using the ground-motion prediction equation of Boore et al. (1997) with different values of sigma (log10 units).

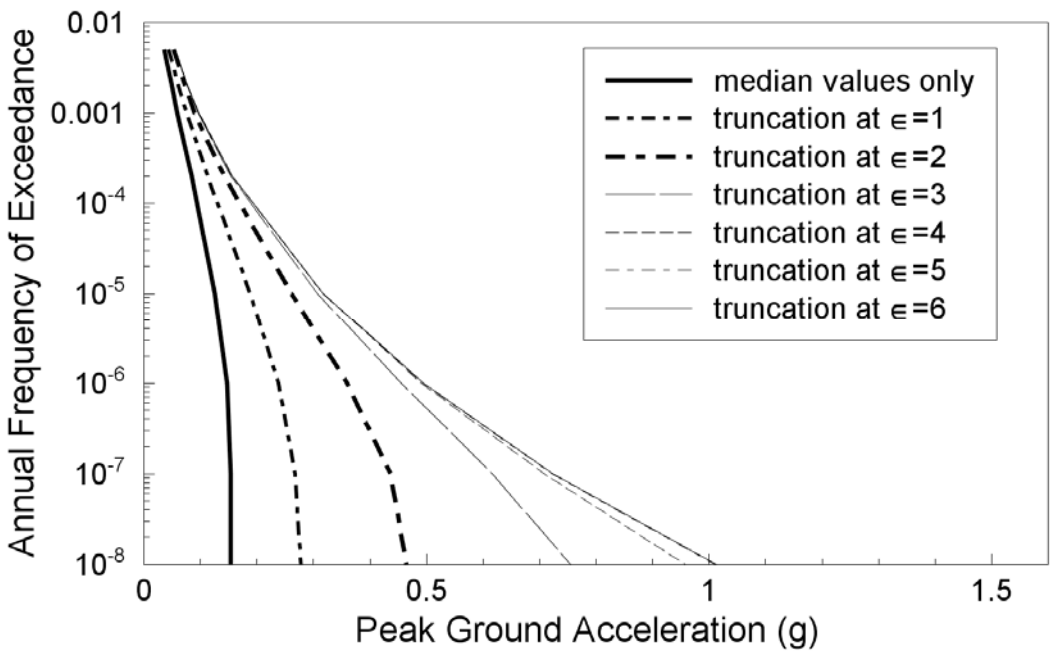

Figure 2-11
Seismic hazard curves derived imposing truncations on the ground-motion distribution at different levels of epsilon (from Bommer et al., 2004).

3
COMPONENTS OF GROUND MOTION VARIABILITY

Inter-event and Intra-event variability

Modern empirical ground motion models account for the correlation of the ground motion from a single event using either the two-step method (Joyner and Boore, 1981) or the random effects method (Abrahamson and Youngs, 1992). The purpose of this section is to provide a brief explanation of these two sources of variability.

The ground motion model has two variability terms as shown in equation (3-1),

$$\ln(Y_{ij}) = f(M_i, F_i, R_{ij}, V_{S30ij}) + \delta_{Aij} + \delta_{Ei} \qquad \text{(Equation 3-1)}$$

where Y_{ii} is the spectral acceleration for the jth recording from the ith earthquake, M_i is the magnitude, F_i is the focal mechanism, R_{ij} is the distance, and V_{S30ij} is the average shear-wave velocity in the top 30 m. The intra-event residual, δ_{Aij}, represents the path and site variability that is observed for a set of ground-motion recordings from a single event. This separation of the variability is shown in Figure 3-1. The dashed lines represent the median curves for the two events. The δ_{Aij} terms represent the variability about the median curves and have a standard deviation of σ_A.

The inter-event residual, δ_{ei}, represents the source variability that is observed for a given magnitude and style-of-faulting. The inter-event variability is also shown in Figure 3-1. These event terms measure the difference between the median curve for a specific earthquake and the median curve for earthquakes with the same magnitude and mechanism. The δ_{Ei} have a standard deviation of σ_E.

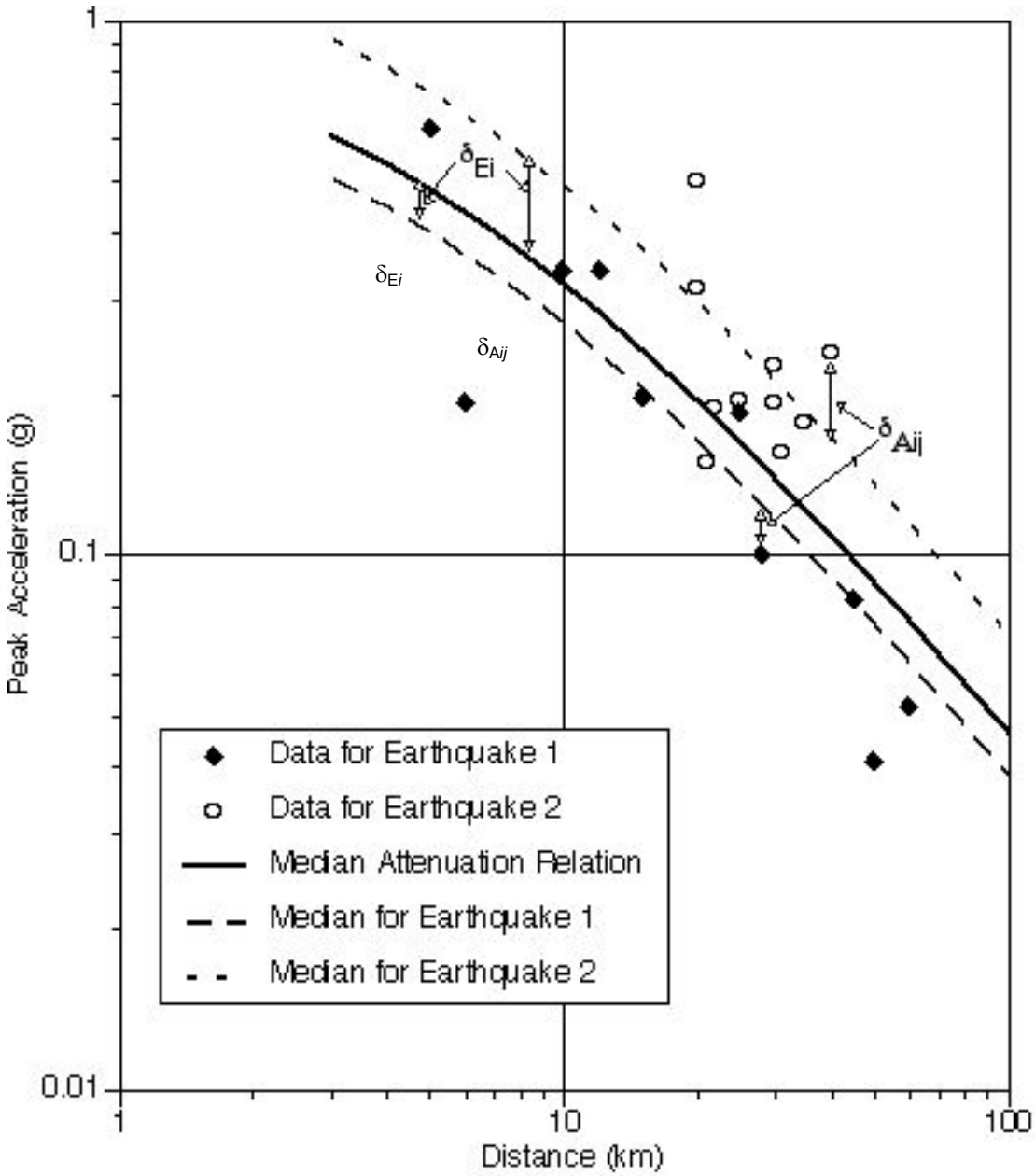

Figure 3-1
Inter-event and Intra-event variability. The differences between the observations and the median for that event are the intra-event variability (δ_{Aij}) and the differences between the median for each earthquake and the median from the attenuation relation are the inter-event variabilities (δ_{Ei}).

4
MAXIMUM EPSILON VALUES

The present chapter summarizes the insights that can be gained both from empirical data and theoretical simulations regarding the level at which the ground-motion distribution can be truncated. Of particular interest is the issue whether records associated with large positive residuals can be considered outliers (*i.e.* anomalous data) or contaminants (*i.e.* data following another distribution than the one under consideration).

Observations of large positive epsilons

Truncation of the ground-motion distribution at a fixed number of standard deviations above the median has been discussed by several authors in the context of hazard analysis. There is, however, a lack of consensus concerning the level of truncation to adopt, suggestions ranging from $+2\sigma$ to $+4\ \sigma$ (e.g. Bernreuter *et al.*, 1989; Anderson & Brune, 1999; Abrahamson, 2000; Romeo & Prestininzi, 2000). These proposals have been based mainly on the representation of the residuals in quantile-quantile or normal probability plots. Such plots display the quantiles of the residuals observed in the dataset against either the theoretical quantiles or the probability associated with a normal distribution; if the points fall along a line, the empirical distribution can be considered normal. If the distribution is truncated, then the probability plot would become vertical at large values on the x-axis. For the datasets used to derive empirical ground motion prediction equations (GMPE), deviations from the best-fitting lognormal distribution are observed above a certain level, generally around $+2\sigma$, as illustrated in Figure 4-1 (after Bommer *et al.*, 2004); however, these plots show that the normal probability plot is becoming flatter, not steeper, indicating that the empirical distribution is broader than expected for log-normal above 2σ. In this figure, the Berge-Thierry et al (2003) model has residuals as large as 3σ and the Lussou et al (2001) data set has observed standard deviations that reach values greater than 4σ.

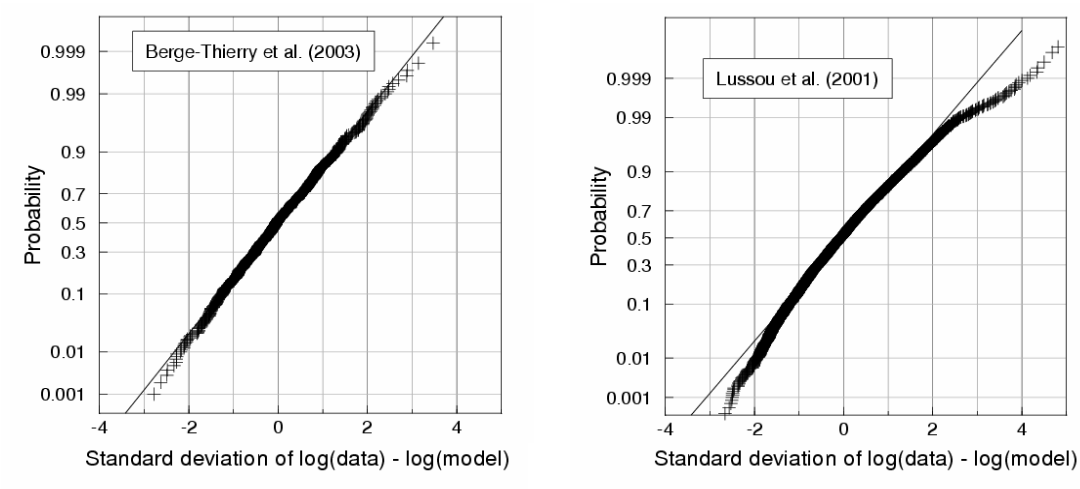

Figure 4-1
Normal probability plots for the PGA residuals of the Berge-Thierry *et al.* (2003) and Lussou *et al.* (2001) equations (modified from Bommer *et al.*, 2004).

In a recent study presented at a workshop on Extreme Ground Motions for the Yucca Mountain project (Strasser & Bommer, 2004), the residuals of a number of strong-motion data sets used for the derivation of predictive equations for horizontal spectral ordinates (Ambraseys *et al.*, 1996; Berge-Thierry *et al.*, 2003; Bommer *et al.*, 2003; Chang *et al.*, 2001; Lussou *et al.*, 2001) have been examined. These equations differ by the extent and provenance of the data, as well as by the number and definitions of the variables used in the regression, and all these equations use a homoscedastic scatter (i.e. σ is independent of the explanatory variables); however, the largest residuals are consistently at least at the 2.5σ to 3σ level. The overall range of values taken by the residuals varies with the frequency considered, the number of variables included in the regression and the definition of horizontal component used, although no systematic trends could be found. As would be expected for well-conditioned datasets, the residuals are not correlated with the explanatory variables. In addition, no pattern could be found with respect to site classification or style-of-faulting when these variables were not included in the regression, which is consistent with the observation that adding these variables does not result in a significant reduction of variance.

In a second step, the 15 highest residuals of the Bommer *et al.* (2003) and Berge-Thierry *et al.* (2003) datasets were analyzed to check whether the lack of distinctive patterns is shared by the subsets constituted by these extreme observations. They found that these subsets showed a lack of correlation with any of the basic explanatory variables used in regression analysis.

It is customary to associate repeated high residuals at a given station with a site-specific response. Similarly, when all records from a single event exhibit high residuals, this is often interpreted as a source characteristic, such as the often cited explanation of “high stress drop” for the 1985 Saguenay, Quebec earthquake (Boore and Atkinson, 1992); however, individual examination of the large-epsilon records and the associated data shows that, in most instances, it is not possible to classify large epsilons unambiguously as source- or site-related. A consistent finding is that ground motions from other stations recorded during the same event as a record

providing a high value of epsilon do not necessarily display high residuals, which points to a high intra-event variability. Figure 4-2 illustrates this in the case of the 1984 Lazio Abruzzo (Italy) earthquake, providing the highest PGA residual for the Bommer *et al.* (2003) equation: there does not seem to be any consistent pattern with respect to distance or site classification, nor any indication that the residuals can be related to a gross source characteristic.

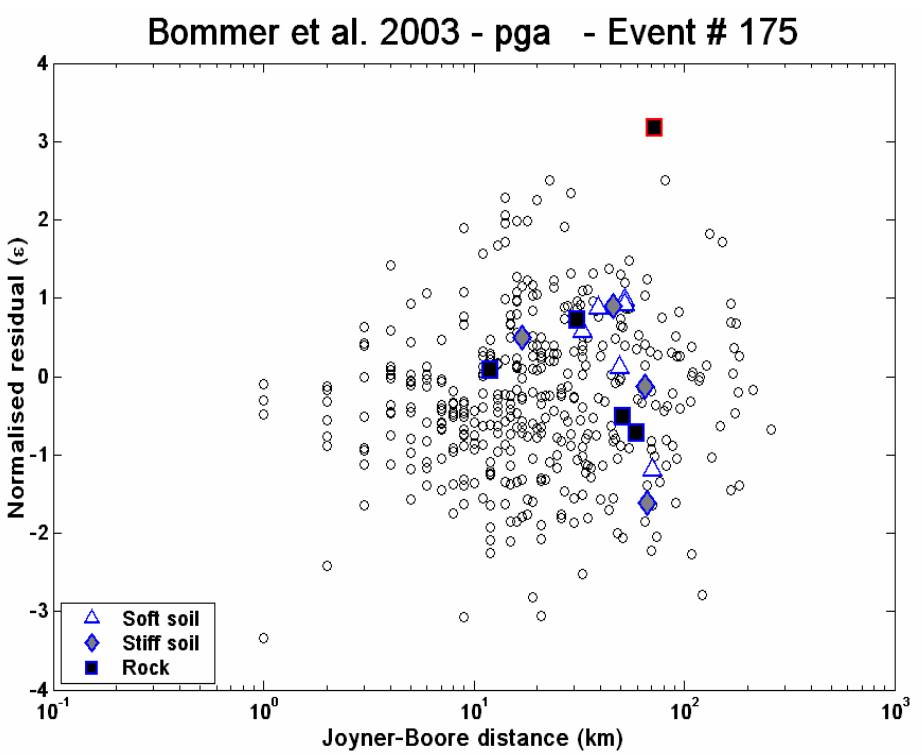

Figure 4-2
Locations with respect to whole dataset of normalized residuals from all stations available for a single event (Lazio Abruzzo, 07/05/1984, M_S=5.8) plotted against distance, including site classification (Strasser & Bommer, 2004). The earthquake that lead to the highest epsilon (red) did not lead to large epsilons for other stations (blue).

Statistical analysis of large positive epsilons

Strasser & Bommer (2005a) examine records associated with large positive epsilons in the strong motion database developed by the Pacific Earthquake Engineering Research (PEER) Center as part of the Next Generation Attenuation (NGA) project. The NGA database includes 3551 records from 173 events and 1455 stations. A subset of this data (2791 records from 102 events and 1150 stations) are used in a regression analysis, which is carried out using the one-stage maximum-likelihood random effects approach described by Abrahamson & Youngs (1992). In the context of this study, "large" epsilons are defined as those with values in excess of 2.5. The selection is based on the total residual, since this is the variable used in seismic hazard calculations to account for ground-motion uncertainty. Records are selected if ε_T exceeds the

threshold level for any of the 15 response periods considered: 0.0s (PGA), 0.02 s, 0.05s, 0.10s, 0.15s, 0.20s, 0.30s, 0.40s, 0.50s, 0.75s, 1.0s, 1.50s, 2.0s, 3.0s and 4.0s.

The choice of the threshold value used in the selection is unavoidably somewhat arbitrary. The reasons for choosing the +2.5 σ_T level are threefold:

All datasets examined in the Strasser and Bommer (2004) study include data points with $\varepsilon_T \geq 2.5$, which therefore constitutes the lowest possible selection value if a comparison between the two studies is to be made.

For a sample of 3,000 points, the theoretical fractile corresponding to the smallest observable frequency is the 99.97th percentile, corresponding to $\varepsilon = 3.4$ (see Strasser & Bommer 2005b, p. 92 for more details), which means that records beyond that level can be expected to be affected by sample size effects. For the meaningful interpretation of large-epsilon records in terms of physical factors, it is desirable to use a selection threshold that is significantly lower than this value.

This level is roughly consistent with threshold levels commonly used in simple statistical tests to detect potential outliers. These tests will be discussed in more detail later in this chapter.

This criterion results in the selection of 136 records, from 24 events. The selection includes recordings from all site classes and all styles-of-faulting, as illustrated in Figure 4-3. Note that only 11 records exceed the $\varepsilon_T = 3.4$ level above which sample size effects are expected to affect the residual values. Most events in the selection contribute only one or two records, with the notable exception of Chi-Chi aftershocks ChiChi-05 and Chi-Chi-06 and the Loma Prieta events, which contribute 87, 10 and 9 records, respectively.

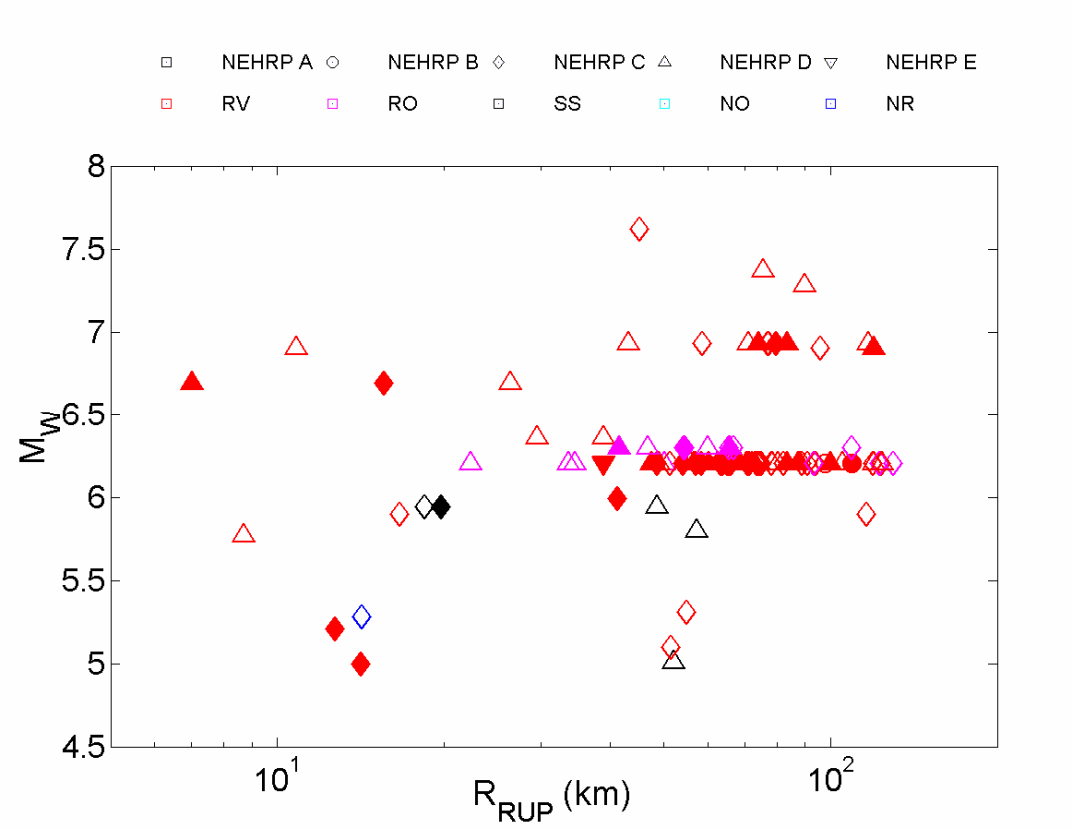

Figure 4-3
Distribution of selected large epsilon records in magnitude-distance space. The shape of the symbols reflect the NEHRP site class, while the color reflects the mechanism associated with the event. Filled symbols correspond to records selected at more than 3 periods.

For the purpose of investigating the factors responsible for these large positive epsilons, the selected records are divided into four categories, according to the number of records (N_{REG}) contributed by the causative event to the regression. Records from the 1999 Chi-Chi, Taiwan sequence are treated separately in view of the overwhelming predominance of these data in the overall dataset. Table 4-1 summarizes the events in each category, along with N_{REG} and the number of records contributed to the selection (N_{SEL}).

Table 4-1
Summary of selected large epsilon records

Category		N_{REG}	Event	N_{SEL}
I	CHI-CHI DATA	399-406	Chi-Chi - 01	1
		3-280	Chi-Chi - 02	1
		71-235	Chi-Chi - 03	4
		66-237	Chi-Chi - 04	1
		79-254	Chi-Chi - 05	87
		82-268	Chi-Chi - 06	10
II	WELL-RECORDED EVENTS ($N_{REG} \geq 45$)	43-47	Coalinga-01	2
		10-112	Whittier Narrows-01	1
		76-77	Loma Prieta	9
		102-154	Northridge-01	3
		2-48	Northridge-06	1
III	INTERMEDIATE EVENTS ($N_{REG} \geq 10$ and $N_{REG} < 45$)	0-16	Imperial Valley-07	1
		7-12	Irpinia-01	1
		7-11	Coalinga-05	1
		20-26	Landers	1
		19-21	Kobe	2
		23-24	Yountville	1
IV	POORLY RECORDED EVENTS ($N_{REG} < 10$)	5-6	Livermore-01	1
		3-6	Westmorland	1
		0-2	Coalinga-07	1
		2-5	Mammoth Lakes-06	3
		1-3	Borah Peak, ID-02	1
		5	Manjil, Iran	1
		9	CA/Baja Border Area	1

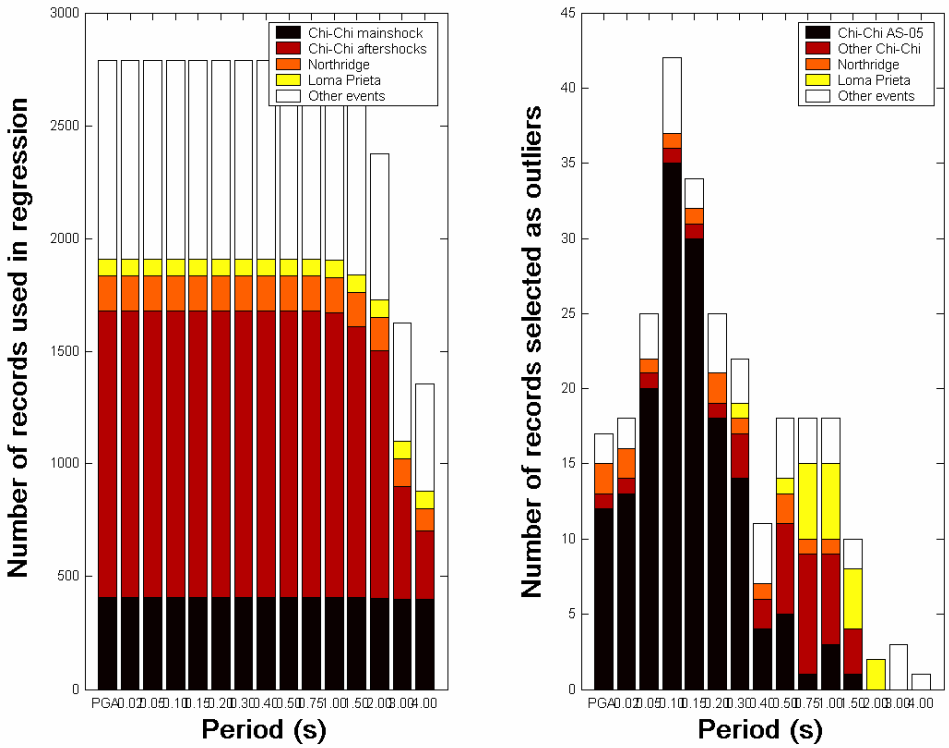

Figure 4-4
Contribution of individual events to (left) complete NGA dataset and (right) large positive epsilon selection, at several response periods.

The need to take into account the number of records contributed by a given event in order to investigate the causative factors arises from the manner the variance components are determined in the random effects model. This model assumes that the inter-event residuals (event terms) are randomly sampled from a normal distribution with zero mean and standard deviation σ_E. As pointed out by Abrahamson & Youngs (1992), this approach leads to a physically meaningful partitioning of the variance into inter-event and intra-event components, all events having the same weight in the determination of the inter-event variability and all records having the same weight in the determination of the intra-event variability.

However, for a strongly unbalanced dataset such as the NGA dataset (see Figure 4-4), this means that the interpretation of the inter-event residuals δ_E (event terms before normalization by σ_E) will be different depending on the number of records contributed to the regression. While δ_E corresponds to the average total residual, $<\delta_T>$, for events that contribute a large number of records, it only corresponds to a fraction of $<\delta_T>$ for events contributing a small number of records, as summarized in Table 4-2.

Table 4-2
Ratio between the event term, δ_E, and the average total residual, $<\delta_T>$, depending on the number of records N_{REG} contributed to the regression. This ratio depends on the relative values of intra- and inter-event variability and is therefore dependent on response period.

T	PGA	0.1s	0.20s	0.30s	0.40s	0.50s	1.0s	2.0s	3.0s	4.0s
σ_{INTRA}	0.51	0.54	0.54	0.55	0.55	0.56	0.6	0.62	0.57	0.58
σ_{INTER}	0.35	0.45	0.39	0.33	0.35	0.35	0.34	0.37	0.38	0.35
N_{REG}	$\delta_E / <\delta_T>$									
1	32%	41%	34%	26%	29%	28%	24%	26%	31%	27%
2	49%	58%	51%	42%	45%	44%	39%	42%	47%	42%
3	59%	68%	61%	52%	55%	54%	49%	52%	57%	52%
4	65%	74%	68%	59%	62%	61%	56%	59%	64%	59%
5	70%	78%	72%	64%	67%	66%	62%	64%	69%	65%
10	82%	87%	84%	78%	80%	80%	76%	78%	82%	78%
20	90%	93%	91%	88%	89%	89%	87%	88%	90%	88%
50	96%	97%	96%	95%	95%	95%	94%	95%	96%	95%
100	98%	99%	98%	97%	98%	98%	97%	97%	98%	97%

Since the reference level for the computation of the intra-event residuals δ_A is defined by $\mu + \delta_E$, this results in events contributing a small number of records being doubly "penalized" in their capacity of contributing residuals to the large-epsilon selection: firstly, a large value of the intra-event residual δ_A is less likely to be observed in a smaller sample, and secondly, δ_E is forced to be closer to the median than it would be if more recordings were available. Combined with the use of a constant ε_T threshold, this results in the preferential selection from events contributing a large amount of data and associated with large positive event terms, as illustrated in Figure 4-5. Conversely, if the number of records contributed to the regression is sufficiently large, a large negative event term may compensate for large positive intra-event residuals.

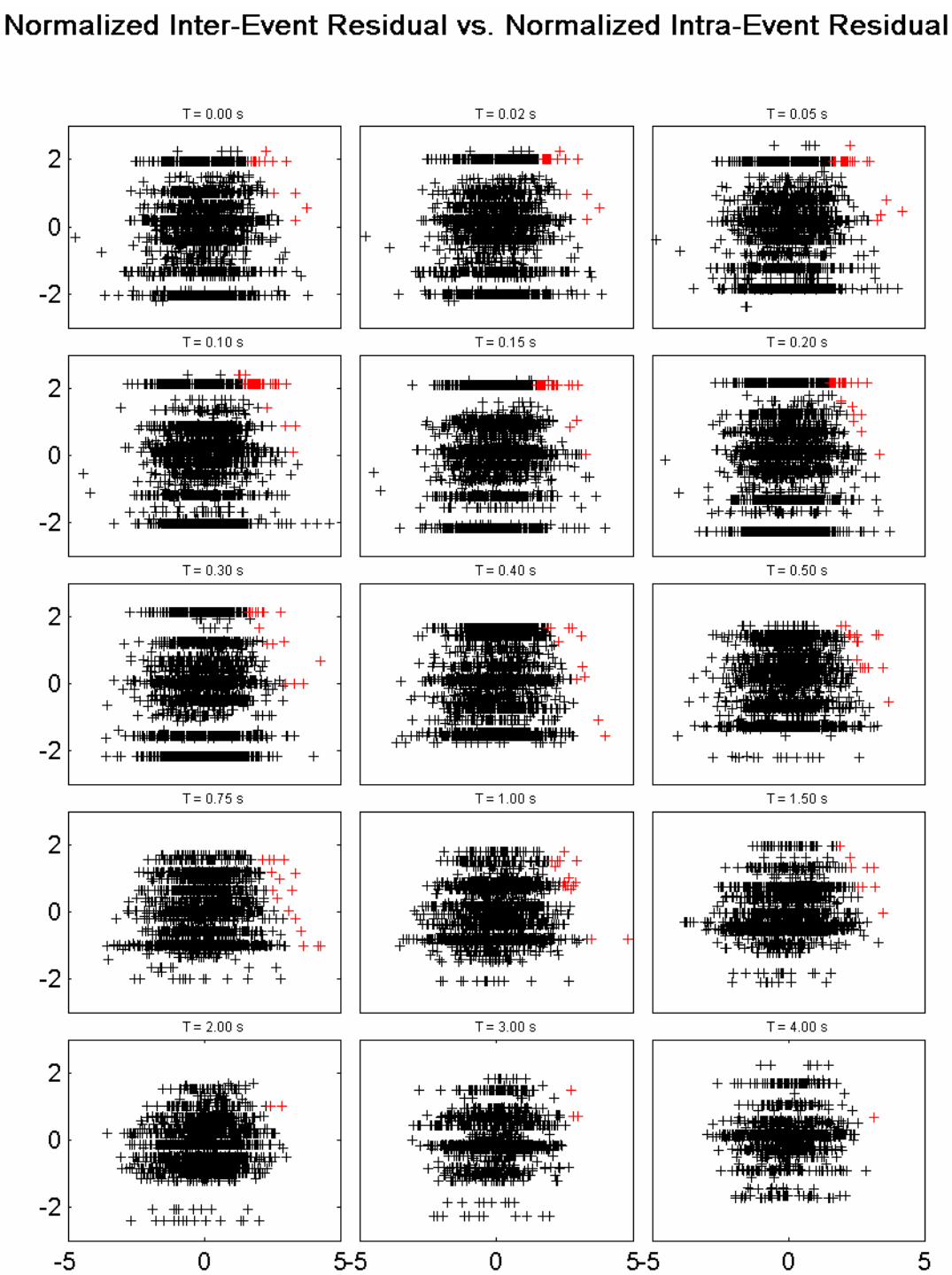

Figure 4-5
Plots of normalized inter-event residuals (σ_E) vs. normalized intra-event residuals (σ_A) of the NGA dataset, at the 15 response periods considered in this study. Records selected because their normalized total residual ε_T exceeds 2.5 are highlighted in red.

Postulated causes of large positive epsilons

This section provides an overview of the interpretations that have been put forward to explain the large positive residuals that correspond to large-epsilon records in the NGA database. Not all records in the selection are discussed, the purpose of this section being to highlight the issues associated with finding physical explanations for large positive epsilons. The reader is referred to Strasser & Bommer (2005a) for a more comprehensive discussion of the physical factors affecting individual records.

Events from the 1999 Chi-Chi, Taiwan, sequence

Data from the mainshock and five larger aftershocks contributes more than 50% of the total dataset, and is therefore not surprising to find records from these events in the large-epsilon selection. What is surprising, however, is the fact that one of the aftershocks (Chi-Chi-05) contributes 87 records, while the mainshock and the other aftershocks only contribute between 1 and 10 records; this is striking because the aftershocks have similar magnitude and contribute comparable numbers of records. Examination of the inter- and intra-event residuals reveals that this pattern is strongly influenced by the value of the inter-event residuals δ_E, shown against response period in Figure 4-6.

Ground-motions from the Chi-Chi mainshock have been noted for their low- amplitudes when compared to predictions from GMPE based on predominantly Californian data (*e.g.* Boore, 2001). The ground-motions recorded during 6 larger aftershocks, however, have been found to agree reasonably well with these predictions (Wang *et al.*, 2004), which would imply that the ground-motions from Chi-Chi-05 are not anomalously high. In fact, they are comparable to the ground-motions observed during the 1984 Morgan Hill event, which has a similar magnitude ($M_W = 6.19$) as Chi-Chi-05 ($M_w = 6.2$) and contributes mainly negative residuals at high frequencies; even the Coyote Lake Dam record, often discussed because of its high PGA value (e.g. Abrahamson & Darragh, 1985; Boore *et al.*, 2004) does not exceed the +1 σ level at high frequencies. The conclusion from this is that the high residuals observed for Chi-Chi-05 are caused not so much by large ground-motions as by a markedly different behavior from the other aftershocks, resulting in an unusually large inter-event variability for this sequence.

This variability is likely to be related to the complex source process involved in this sequence. Based on geological information and waveform inversions, several authors concluded that the aftershock sequence involved the rupture of two conjugate fault planes (e.g. Kao *et al.*, 2002; Chen *et al.*, 2002): the shallowly eastward-dipping Chelungpu Fault which ruptured during the mainshock, and a steeply westward-dipping conjugate plane. Chi-Chi-05 seems to be located at the junction between the two planes, which results in ambiguities about the source mechanism of the event. Uncertainties in the fault geometries associated with these aftershocks have also been noted by Chi & Dreger (2004), who inverted strong-motion data to determine the slip distributions of these events. The same authors also noted that the differences in dip between events resulted in different spatial distribution of the ground-motions, due to radiation pattern and directivity effects.

While this information sheds some light as to what causes residuals from Chi-Chi-05 to be high, it also begs the question of the transportability of variability estimates. The discussion above leads to the conclusion that the large epsilons contributed by Chi-Chi-05 have to be considered as random outliers, i.e. outliers caused by the intrinsic variability of the ground-motions rather than by some particular physical mechanism not included in the regression; however, it also appears that the large inter-event variability associated with the Chi-Chi sequence is caused by a complex source process, which in turn can be related to regional tectonics. Is this level of source complexity representative of the source complexity in stable tectonic region? This issue is addressed in section 6 by comparing source variability from active and stable regions using global earthquake catalogs.

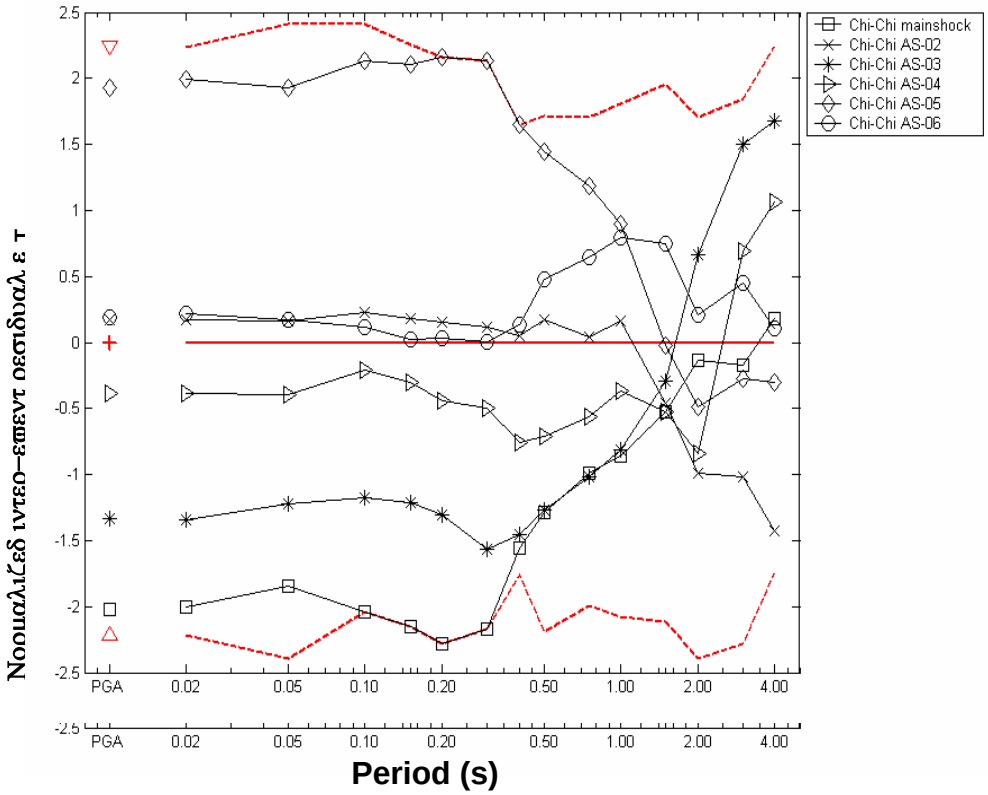

Figure 4-6
Normalized inter-event residuals for the Chi-Chi, Taiwan sequence. The dashed red line give the envelope from all events in the data set.

Well-recorded California events

The records in category II are best-suited for the analysis in terms of causative physical factors, since the number of records contributed by the events is large enough to warrant extensive studies of the ground-motion, but still small enough to allow the selection and analysis of individual records.

This subset includes the largest residual found in this study ($\varepsilon_T = 3.99$), observed at Tarzana during the 1987 Whittier Narrows earthquake, for a response period of 0.30 s. Large residuals

have also been observed for this record at other response periods, and the record clearly departs from the general pattern observed for ground-motions from this event (Campbell, 1988). The preliminary analysis concluded that that a strong site effect was responsible for the large residuals, which was surprising considering the location of the station on a gently sloping hill, the site being classified as rock; however, the matter was not investigated further until the same site again recorded unusually large ground-motions during the 1994 Northridge event. Interpretations of the unusual ground-motions observed at Tarzana include topographic amplification by the hill combined with 3D-effects of the underlying geological structure (Bouchon & Barker, 1996; Spudich *et al.*, 1996), resonant sliding block behavior induced by a small nearby event triggered by the mainshock (Rial, 1996), effects of highly saturated layers at shallow depths (Catchings & Lee, 1996), and effects due to the inclination of incident waves combined with those of local geology and topography (Vahdani & Wikstrom, 2002). The overall conclusion from these studies is that the strong amplification at Tarzana results from a combination of factors which are specific to this particular site. For these types of records, the issue then becomes to determine how likely it is that a similar combination of effects occurs somewhere else, and in particular at the site of interest for the seismic hazard analysis. For the example of Tarzana, the effects listed above are clearly irrelevant to a CEUS hard-rock site located on flat topography.

The 1989 Loma Prieta event contributes 9 records to the selection, all of which are from the San Francisco-Oakland area. The fact that ground motions were higher in the Bay Area compared to locations closer to the source has been related to critical Moho reflections (Somerville & Yoshimura, 1990), combined with forward-directivity effects from a bilateral rupture (Somerville *et al.*, 1994). The selected records constitute only a subset of the Bay Area records, and it is therefore likely that they are affected by additional, more site-specific factors. Explanations include amplification by soft soil sediments (*e.g.* Borcherdt & Glassmoyer, 1994; Field *et al.*, 1994), relative amplification at rock sites located on less competent materials than the Franciscan Complex (Borcherdt & Glassmoyer, 1994), topographic effects (Borcherdt & Glassmoyer, 1994) and influence of ocean-wave motions (Vidale & Bonamassa, 1994). These diagnostics are, however, shared with a number of other records included in the NGA dataset used in the regression that do not exhibit large enough residual values to be included in the selection. For example, the interpretations of the large ground motions recorded at Alameda Naval Air Station as due to the resonance with the soft Bay sediments underlying the stations are identical to those presented for ground motions recorded at nearby Treasure Island (*e.g.* Borcherdt & Glassmoyer, 1994), and studies investigating the response of these stations (*e.g.* Carlisle & Rollins, 1994) use the same reference rock site (Yerba Buena Island). Nevertheless, the Treasure Island record is associated with smaller residuals (ε_T = 1 to 1.5) and therefore was not included in the selection. This example illustrates the difficulty of establishing a robust relation between causative factors and large residual values even in cases where the records are particularly well documented.

Records from events with intermediate sample size

Events in this category can still be considered reasonably well-recorded, but the records are generally not as well documented as those from category-II events. This can be due to two factors: either the records come from a well-documented event, such as Landers or Kobe, but

have received comparatively little attention due to their location far from the source (about 90 and 120 km for the Landers and Kobe records, respectively), or the event as a whole is not particularly well documented. An example of the latter is constituted by the Napa Fire Station record from the 2000 Yountville event. Considering its moderate size (M_w 5.0), this event caused a significant amount of damage in the Napa Valley ($30-50 M damage, 25 injured). Preliminary reports (USGS, 2000;Miranda & Aslani, 2000; EERI, 2000) commented on the unusually high level of the ground-motions recorded at several stations located in the town of Napa, and tentatively attributed them to the combined effects of forward directivity, topography and surface sediments based on a comparison with ground motions recorded on rock at similar distances. There is, however, no information available beyond this circumstantial evidence.

Records from event with small sample size

Records in this category are generally poorly documented. Most of the records selected from events contributing a small number of records to the regression are selected at long response periods, where the dataset is reduced due to the application of criteria pertaining to the quality of the data, and hence more balanced (see Figure 4-4). The Mammoth Lakes-06 event constitutes an exception, contributing 3 records out of 5 to the large-epsilon selection at high frequencies. Examination of the residuals shows that this event is associated with the highest inter-event residual δ_E in the dataset at high frequencies, i.e. the selection of the records is event-driven. A possible explanation for this unusual high-frequency behavior is the volcanic nature of the event. Peppin (1987) discusses several "exotic" phases systematically associated with events occurring in the Long Valley Caldera, which might cause enhanced high-frequency motions. There is some controversy about the nature and location of the scatterers causing these phases, but this debate is unlikely to be relevant to the estimation of ground motions in non-volcanic regions such as the CEUS.

Maximum epsilons from numerical simulations

Strasser & Bommer (2005b) present an analysis in terms of ground-motion vatiability of the results of numerical simulations carried out using the deterministic-stochastic kinematic finite-source model EXWIM developed by Dr Enrico Priolo and co-workers at the Osservatorio di Geofisica Sperimentale (OGS) in Trieste, Italy. This model had previously been used for the estimation of maximum ground motions in the PEGASOS project (Abrahamson *et al.*, 2002).

EXWIM computes synthetic seismograms generated by an extended seismic source and accounts for seismic wave propagation through a vertically heterogeneous anelastic structure. The computational kernel of the approach is the Wavenumber Integration Method (WIM) developed by R. Hermann at St Louis University (Herrmann & Wang, 1985; Herrmann, 1996a; Herrmann, 1996b). The fault rupture is modeled using the kinematic approach developed by Herrero & Bernard (1994), in which the moment density is described by a self-similar random slip distribution with an amplitude decay proportional to k^{-2} at high wavenumbers k, ensuring a high-frequency decay proportional to ω^{-2} in the frequency domain. The deterministic seismograms thus obtained are then subjected to stochastic hybridization in order to model the incoherent nature of ground motions at high frequencies.

As illustrated in Figure 4-7, the simulations were carried out for a dense grid of 122 receivers located at Joyner-Boore distances from 0 to 20 km from the source. All simulations were performed for a $M_w = 6.0$ on a moderately dipping reverse fault (ϕ=25°). A total of nine slip distributions were considered, corresponding to three levels of stress drop (0.09, 0.33 and 2.42 MPa). Nucleation points were assumed to be located on a uniform grid on the fault rupture plane. Thus, events are defined in this study as the combination of a slip distribution and a nucleation point location on the fault plane. The assumptions described above result in a total of 190 x 3 = 570 events for the low stress drop case, 66 x 3 = 198 events for the medium stress drop case, and 14 x 3 = 42 events for the high stress drop case. Finally, the effect of replacing the usual assumption of a spatially uniform rake by a variable rake angle was investigated, effectively doubling the number of events.

An important point to note is that the simulations used in this study were performed for the purpose of assessing the sensitivity of the simulated ground-motions to the various input parameters of the model and were not performed as part of this study. In particular, the calibration in terms of source geometry, crustal model and choice of stress drop levels reflects condition representative of north-eastern Italy, rather than the CEUS; however, while this inevitably influences the absolute values of the simulated ground motions, conclusions regarding ground-motion variability in the near-source region should not be affected. The value of the computed standard deviation will depend on the distributions of the input source parameters. The key conclusion from these simulations is that large epsilon values are generated from source variability and wave propagation through 1-D crustal structures using numerical simulations and thus the large epsilon values cannot discounted as being caused only by unusual site conditions that are not applicable to more typical sites.

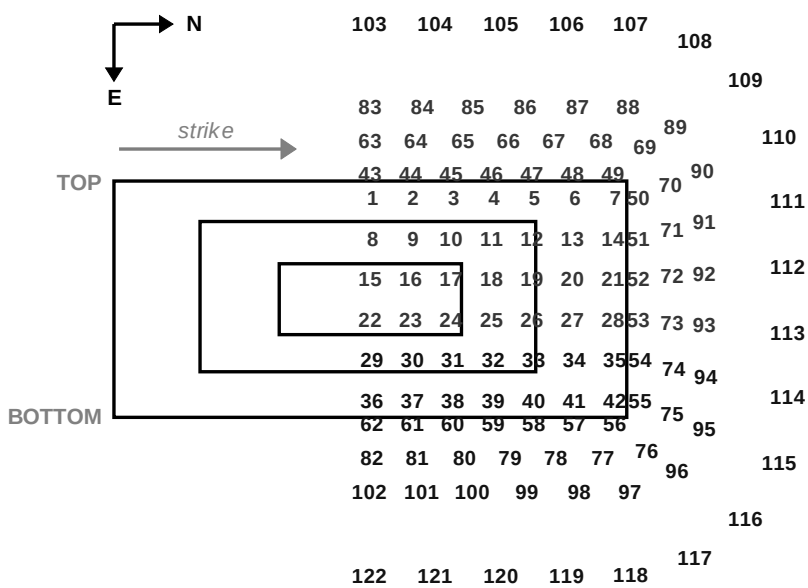

Figure 4-7
Layout of receivers (numbered 1 to 122) used in the simulations. The three rectangles represent

the source rupture areas corresponding to the three levels of stress drop considered.

Different measures of inter- and intra-event variability were investigated, which will be discussed in Chapter 5. A major advantage of the use of numerical simulation results to investigate ground-motion variability is the control over the sample size. In the study discussed here, the number of data points available for a given distance is typically of the order of several thousand data points, i.e. several orders of magnitude greater than the number of points available for a fixed magnitude, distance and site class in the strong-motion data sets used for the derivation of empirical GMPE. The corresponding residuals (defined with respect to the sample mean and normalized by the sample standard deviation) reach the +/- 3 σ or even +/- 4 σ level in a large number of cases.

A good fit to the normal distribution has generally been found in the simulations for motions lying within 2 standard deviations from the sample mean, as illustrated by the quantile-quantile plots shown in Figure 4-8. If the distribution of the data is normal, the plot will be close to linear. Superimposed on the plot is a line joining the first and third quartiles of each distribution, extrapolated out to the ends of the sample. The grey-shaded zones at the top and bottom of the plots correspond to sample quantiles outside the [$-\varepsilon_{TheoMax}$, $\varepsilon_{TheoMax}$] range, where $\varepsilon_{TheoMax}$ is the average theoretical maximum residual expected for a sample of this size. The colored arrows highlight the bounds of the normalized residuals contributed by each stress drop level *(black* = low, *blue* = medium and *red* = high). Beyond the +/- 2 σ range, deviations can be observed, which are of the long-tailed type (*i.e.*, more variability than expected), for receivers located above the fault plane (R_{JB} = 0 km), while receivers located some distance from the fault plane (R_{JB} = 5 and 10 km) show a deviations from log-normal with less variability than expected. The results shown in Figure 4-8 assume uniform weighting of the source parameters; very similar results are obtained when a simple weighting scheme is applied to the source parameters to alleviate biases caused by the uneven contributions of the different stress drop levels.

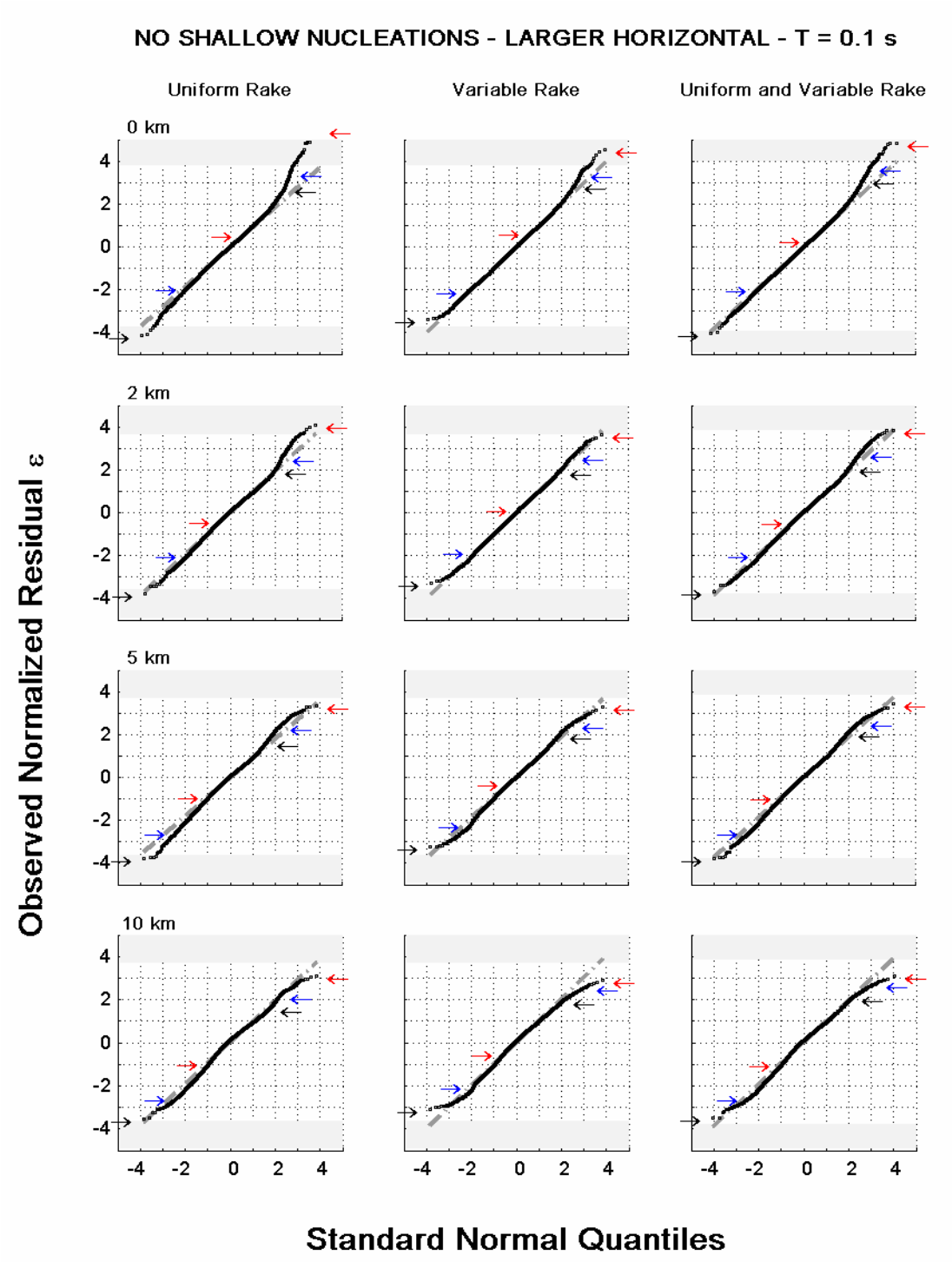

Figure 4-8
Quantile-quantile plot of the normalized residuals from numerical simulations. Results are shown for the larger horizontal component at 0.1 s. (Note: the axes in this figure are switched from those shown in Figure 4-1.)

The interpretation of the residual values obtained from numerical simulations with respect to those observed in GMPE derived from empirical data is conditional on the validity of the

variability estimates, and more generally on the soundness of the model adopted. Despite the limitations of the variability calculations presented in Strasser & Bommer (2005b), which will be discussed in Chapter 5, it can be concluded that the numerical simulation results presented in that study do not support a truncation of the ground-motion distribution.

Recommended maximum epsilon

As stated previously, there is no consensus regarding the level of truncation to adopt in seismic hazard analysis. This section discusses the recommended values of the maximum epsilon.

The results presented in this Chapter corroborate the view that ground motions follow a log-normal distribution at least up to the 2σ level. This conclusion holds for both empirical and numerically simulated data. There are indications that the ground motion begins to deviate from a lognormal distribution above 2σ, but in many cases, the deviation is leading to broader tails of the distribution, not narrower tails. Suggestion by some authors (e.g. Romeo & Prestininzi, 2000) to truncate the distribution at the level at which the ground motion distribution begins to deviate from lognormal is unconservative.

Simple statistical rules commonly used to detect potential outliers, such as the boxplot rule or Dixons's test, consider thresholds around the 2.5σ level. The boxplot rule labels data points as potential outliers if they lie 1.5 times the interquartile range above the third quartile of the dataset, which for the standard normal distribution corresponds to $\varepsilon = 2.69$. Dixon's test is a recursive test for outlier detection which considers the distribution of the largest normalized residual observed in the sample as a test statistic and compares it to a tabulated critical value at a given significance level, the value tabulated for the 5% significance level being 2.49. Note that all these statistical procedures are based on test statistics assuming the target population to be infinite (*i.e.* a true normal distribution). Barnett & Roberts (1993) [quoted in Barnett & Lewis (1994), p.447] have, however, shown that when the target population is reduced to a finite population and the test statistic is applied to samples drawn from this population, the critical value of the test statistic is non-unique, depending on the finite population used in the first stage. In other words, no discordancy test is available for the detection of potential outliers in a finite population, even in the case where the distributional form of this population is known.

This means that the maximum epsilon cannot be fixed on a purely statistical basis, *i.e.* the exclusion of records based purely on the value of their residual cannot be justified. It is, however, possible to investigate whether these records constitute "deterministic outliers" (see Barnett & Lewis, p. 32) in the sense that they are affected by physical mechanisms that are not relevant to the model under consideration. Earlier in this section, the results of such an analysis applied to the NGA dataset was summarized, and the issues involved in such an approach were highlighted. The strong selection bias implies that individual values of large ε have to be treated cautiously, as they contain information both about the record under consideration and the rest of the dataset.

The example of the Chi-Chi aftershocks shows that large inter-event variability of ground-motions between events with otherwise similar characteristics can cause the inter-event residuals (event terms) to deviate strongly from the median, which in turn affects the values of the total

residuals. While this is a desirable feature in order to estimate the variance components accurately, it complicates the problem of relating large-epsilon values with given physical factors by pushing random residuals, δ_T, out to levels where deterministic outliers would be expected.

Clear-cut cases such as Tarzana, where a large residual can be attributed with a high level of confidence to site-specific effect, are rare; more commonly, interpretations in terms of physical effects can explain why the residuals are positive, but generally fail to capture the totality of the factors, since records exhibiting the same characteristics, but with residuals well below the selection threshold, can be found elsewhere in the dataset, as illustrated by the Loma Prieta records. Therefore, we are not able to robustly associate large values of ε_T with particular physical mechanisms.

We conclude that there is no sound technical basis for selecting a maximum value of epsilon and we do not anticipate any studies in the short term that could be conducted to provide such a technical basis for truncation. Numerical simulations of ground motions are likely the only viable method for evaluating the details of the upper tail of the ground motion distribution for the inter-event term, but current kinematic models require joint probability distributions of the source parameters (e.g. slip-velocity, rupture velocity, slip-distribution, hypocenter location…) that are not currently constrained. Ongoing work to develop constraints on these joint distributions is under study in other projects, but the results from these other studies are several years away. For the intra-event variability, focusing of waves due to complex velocity structure is difficult to rule out and numerical simulations based on 1-D velocity structure models will not address this issue. Therefore, a maximum epsilon for the intra-event variability is unlikely to be determined in the short term.

While at some level, the ground motion will reach the physical limit of transmission in rock or soil, this type of truncation is based on the ground motion level, not on the epsilon value directly. Separately, a study of the maximum ground motion that is possible near the source could be defined. There is an ongoing DOE sponsored project to address the maximum ground motions at Yucca Mountain that could be extended to the CEUS.

Given the lack of a technical basis for defining a maximum epsilon value, we recommend that no truncation of the ground motion distribution be included based on a maximum value of epsilon.

5
ESTIMATION OF SIGMA

NGA results for WUS

The Next Generation Attenuation (NGA) project sponsored by the Pacific Earthquake Engineering Reserach (PEER) Center is developing new attenuation relations for the WUS using a large data set consisting of over 3500 strong motion records from over 173 earthquakes (PEER, 2005). The NGA project includes five developer teams that are each developing updated ground motion models. These teams include the following: Abrahamson and Silva, Boore and Atkinson, Campbell and Bozorgnia, Chiou and Youngs, and Idriss. Preliminary results from the NGA project have indicated that there will be significant revisions in the standard deviation of the attenuation relations. For example, the magnitude dependence of the standard deviations for PGA from Abrahamson and Silva (1997) and Sadigh et al (1997) is compared to that of the Abrahamson and Silva (2005) preliminary NGA model in Figure 5-1. An important change in the models is that the standard deviation in the NGA models is independent of magnitude, whereas, in previous models, the standard deviation decreased with increasing magnitude. The NGA models gives a total standard deviation of about 0.6 natural log units for all magnitudes which is smaller than the standard deviations for the 1997 models for M<6 and larger than standard deviations for the 1997 models for M>6.5. The standard deviations from the preliminary results from Abrahamson and Silva (2005) are listed in Table 5-1. The other developer groups in the NGA project have found similar values for the standard deviation.

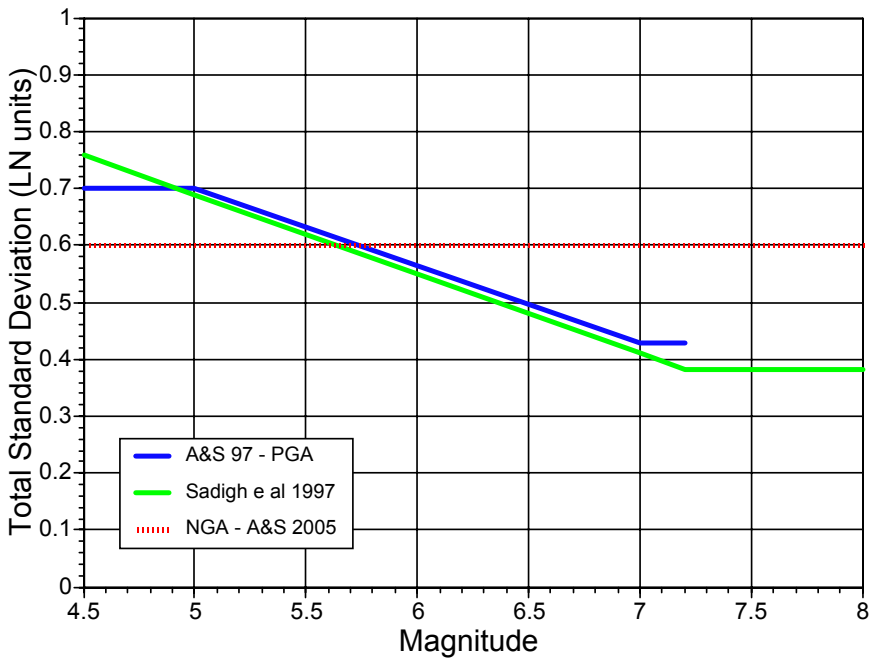

Figure 5-1
Comparison of the standard deviation for PGA from the 1997 models with the preliminary NGA model from Abrahamson and Silva (2005).

Table 5-1
Comparisons of standard deviations from Abrahamson and Silva (1997) with the preliminary model from Abrahamson and Silva (2005).

	Abrahamson and Silva (2005)			Abrahamson and Silva (1997)	
Frequency (Hz)	Intra-event	Inter-event	Total	Total (M=5)	Total (M=7)
PGA	0.51	0.34	0.61	0.70	0.43
25	0.53	0.37	0.64	0.71	0.44
10	0.56	0.41	0.69	0.74	0.47
5	0.55	0.37	0.66	0.77	0.50
2	0.56	0.38	0.68	0.80	0.54
1	0.60	0.36	0.70	0.83	0.59
0.5	0.62	0.41	0.74	0.85	0.61

There are several possible causes for the change in the standard deviation models: the NGA data set has a much larger data set from recent earthquakes, there were corrections to the meta-data for data used in the Abrahamson and Silva (1997) model (particulary corrections to the site classification), and the ground motion parameter was modified to use the geometric mean of the horizontal components after finding the optimal rotational angle (Boore et al, 2005).

To determine the cause for the change in the standard deviation, the standard deviations were computed using the subset of data used by Abrahamson and Silva (1997) but with the updated meta-data, and then with the updated ground motion values. Neither of these changes had a significant effect on the standard deviation, so changes to the meta-data or the definition of the horizontal component are not the main reason for the change in the standard deviation model.

The intra-event residuals for PGA computed using the Abrahamson and Silva (1997) data and the full NGA data set are shown as a function of magnitude in Figures 5-2 and 5-3, respectively. Similarly, the inter-event residuals for PGA computed using the Abrahamson and Silva (1997) data and the full NGA data set are shown as a function of magnitude in Figures 5-4 and 5-5, respectively. The inclusion of the Chi-Chi mainshock (M7.6) causes the increase in the intra-event standard deviation at large magnitudes. The additional earthquakes in the magnitude range of 5-6 have a smaller variability of the event terms resulting in a reduction of the inter-event standard deviation for magnitudes less than 6.0. Therefore, the data from the new earthquakes inthe NGA data set caused the change in the standard deviation for WUS earthquakes, not corrections to the 1997 data sets or the change in the definition of the average horizontal component.

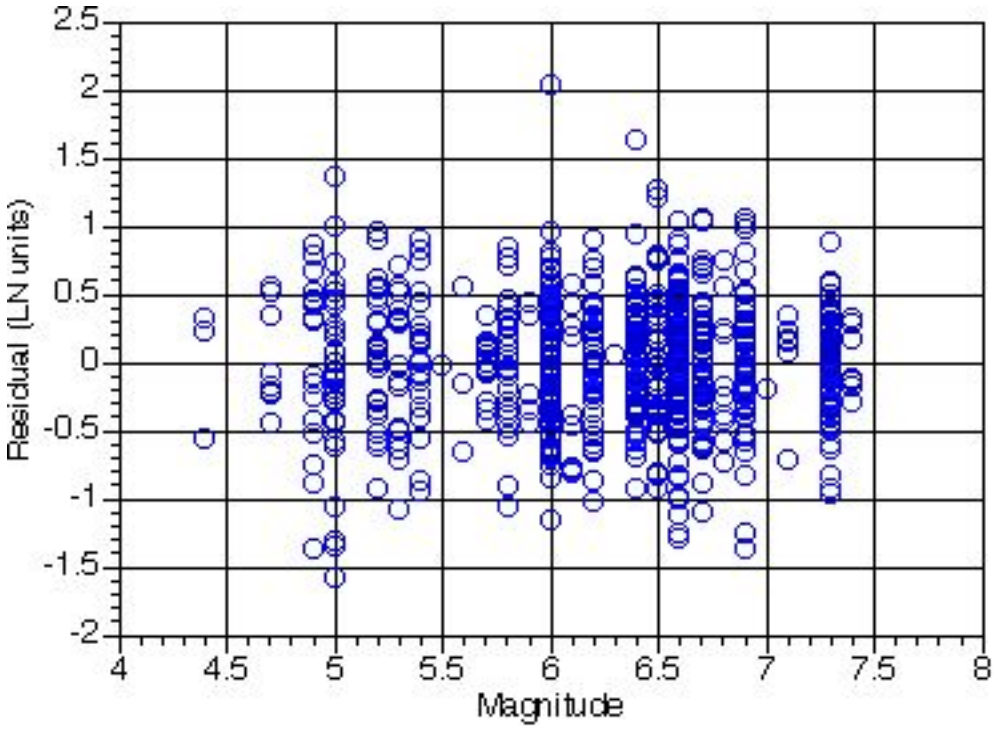

Figure 5-2
Intra-event residual for PGA for the Abrahamson and Silva (1997) data set.

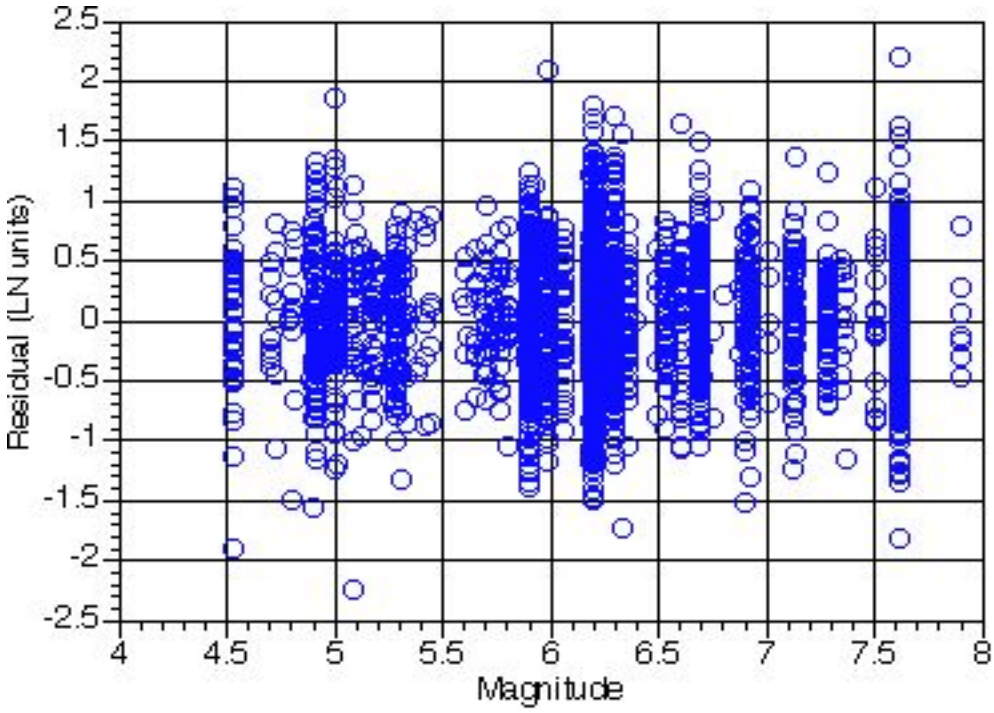

Figure 5-3
Intra-event residual for PGA for the Abrahamson and Silva (2005) data set.

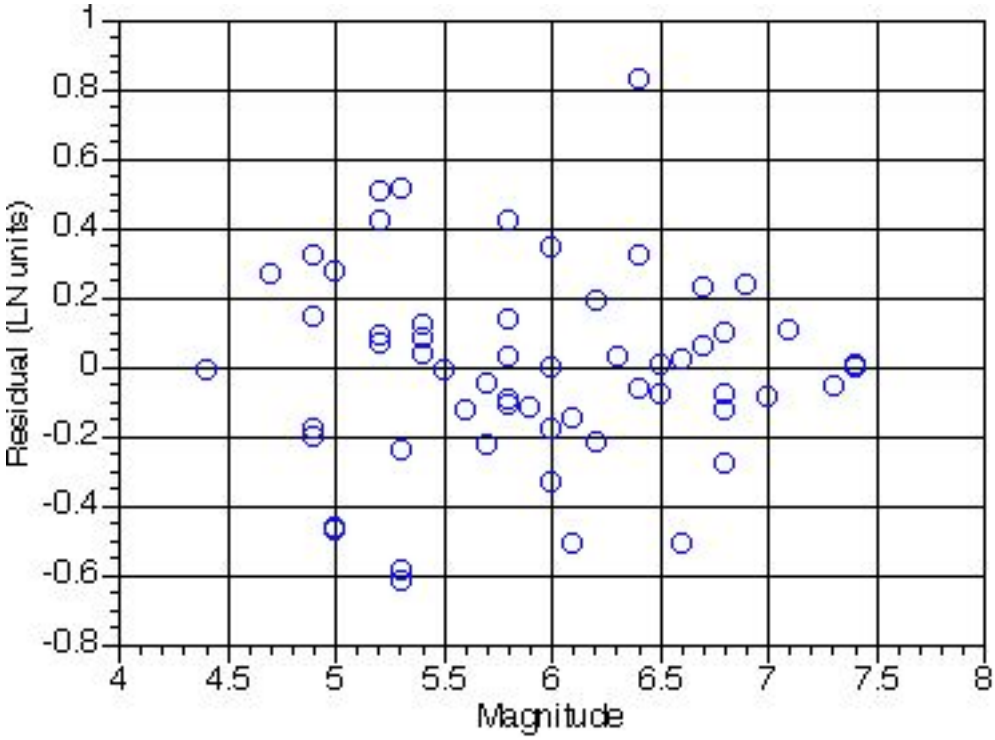

Figure 5-4
Inter-event residual for PGA for the Abrahamson and Silva (1997) data set.

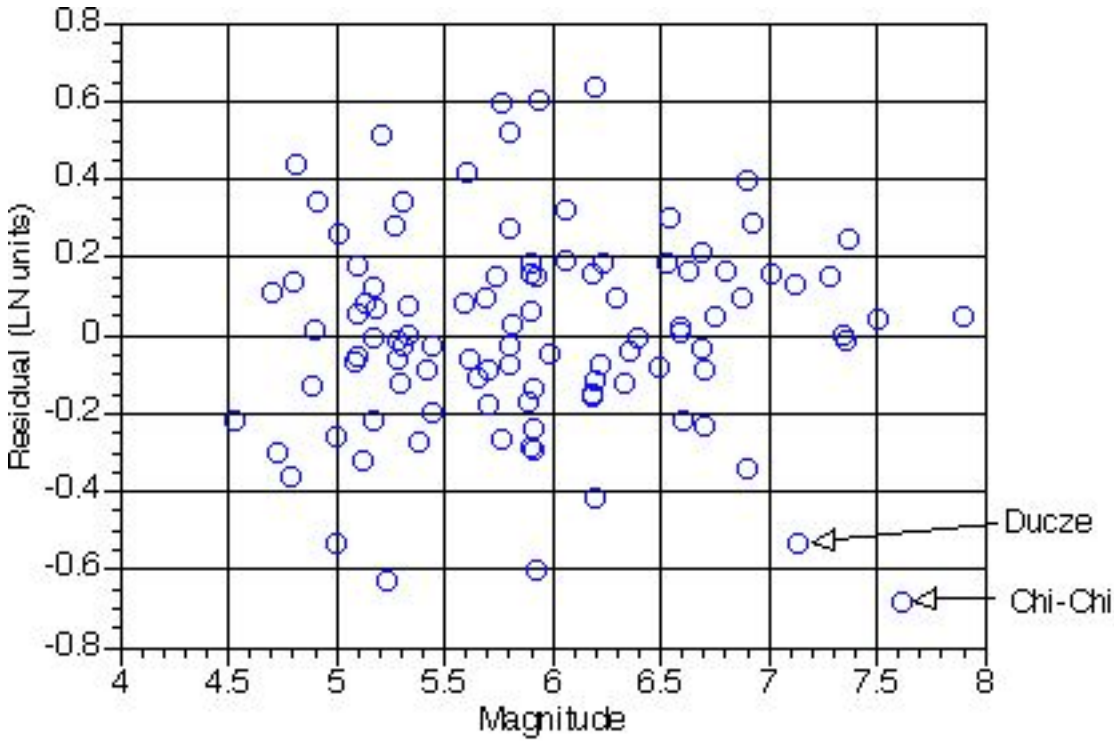

Figure 5-5
Inter-event residual for PGA for the Abrahamson and Silva (2005) data set.

Distance Dependence

The 2004 EPRI ground motion report incorporated a distance dependence to the standard deviation to account for the expected increase in variability at short distances (R_{JB}<25 km) that would result from variability of the source depths that is ignored when using the JB distance. (The Joyner-Boore distance, R_{JB}, is the shortest horizontal distance from the site to the vertical projection of the fault rupture surface) This effect is illustrated in Figure 5-6. Earthquakes with both shallow and deep focal depths would have the same R_{JB}, but would have very different rupture distances (e.g. point source distances). If it is assumed that the ground motion will attenuate as 1/R where R is the point source distance and that there is no systematic differences in the sources as a function of depth (e.g. no stress-drop dependence with depth), then there should be an increase in the standard deviation at short distances due to the use of the R_{JB}. This effect was seen empirically in European strong motion data by Ambraseys and Bommer (1991) in which they found a reduction in the standard deviation as short distances when they replaced the fictitious depth used with the R_{JB} with the actual hypocentral depth. A short-coming of this analysis is that it used ordinary least-squares which ignores correlations in the data from recordings from a single event.

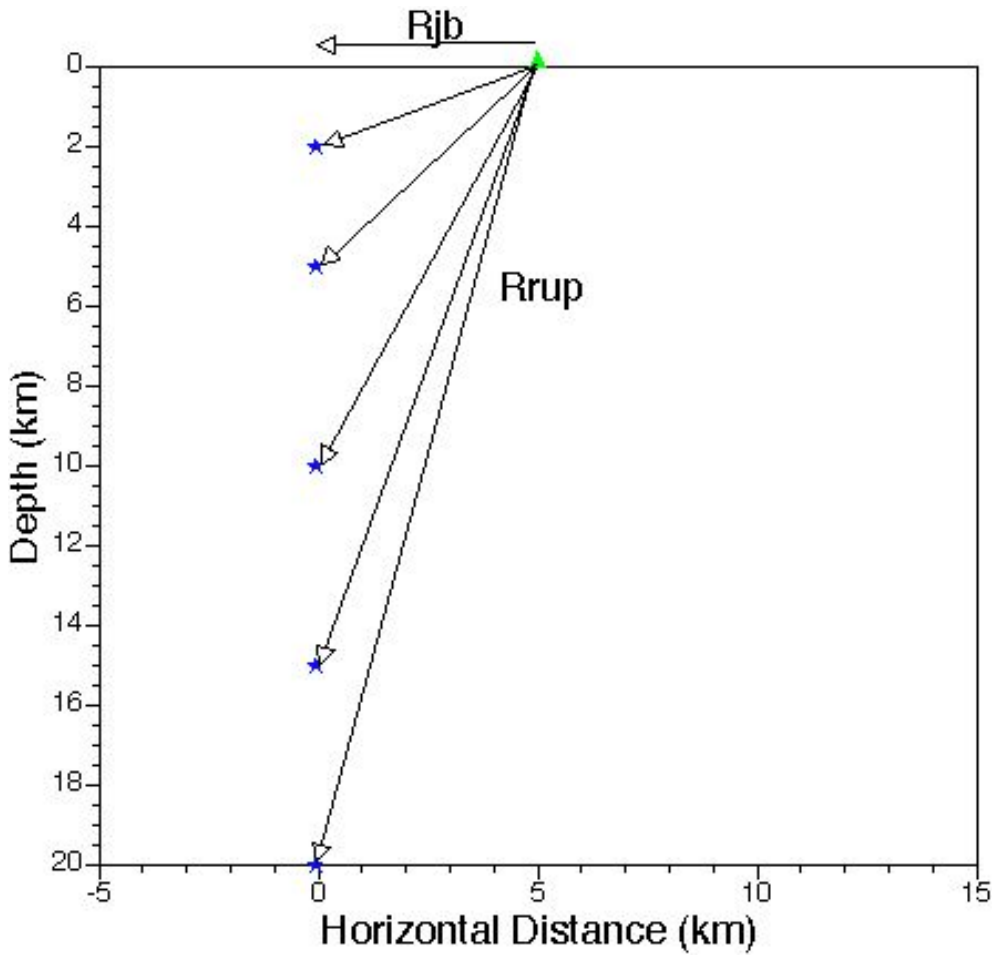

Figure 5-6
Example of range of rupture distances that can occur for short JB distances.

The Toro et al. (1997) EUS ground motion model includes this increase in the standard deviation at short distances as part of the aleatory variability model as shown in Figure 5-7. In this case, the standard deviation increased from 0.62 at distances larger than 20 km to 0.80 at distances shorter than 5 km. This increase is important because the earthquakes at short distances can have a significant effect on the hazard, particularly at low probability levels.

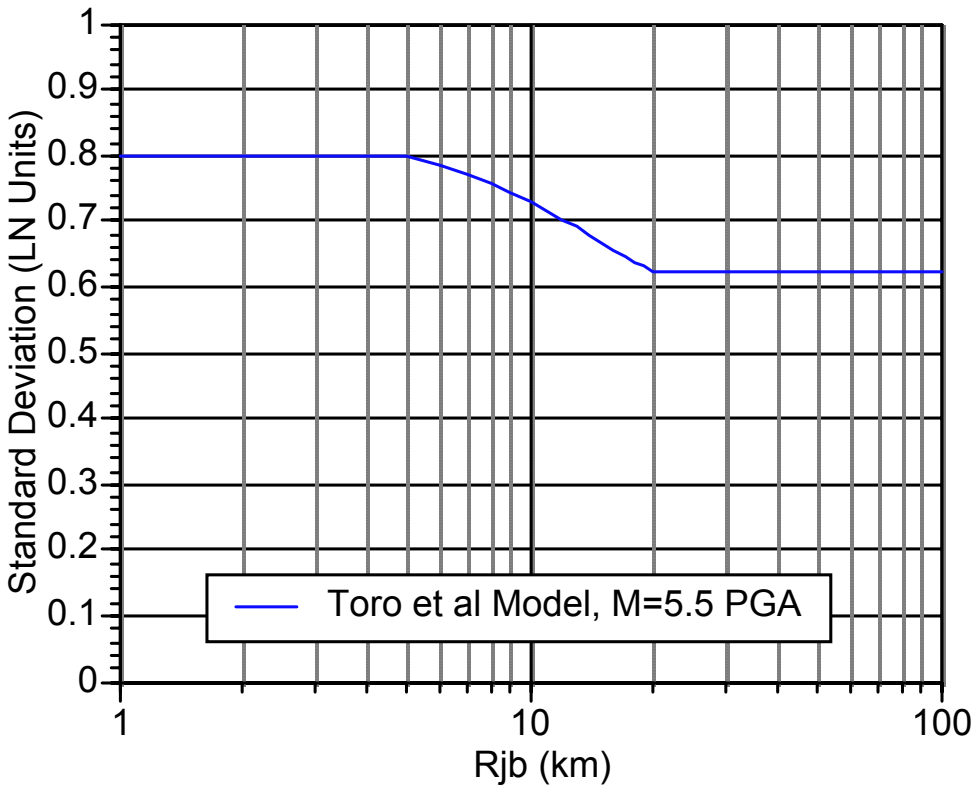

Figure 5-7
Example of the distance dependence of the standard deviation from the Toro et al (1997) model.

There is evidence from the PEER NGA studies that the increase in the standard deviation at short distances may not be appropriate. Using the NGA data set, a regression was conducted using the R_{JB} distance metric. The effect of the source depth variability on the ground motion variability is evaluated by plotting the source depth dependence of PGA residuals from moderate magnitude earthquakes (M<6) for short JB distances (R_{JB}< 10 km). The depth dependence of the intra-event residuals are shown in Figure 5-8. If the standard deviation needs to be increased at short distances for JB distance models, then for small magnitudes and small R_{JB} values, the residuals should show a negative correlation with source depth. For the range of source depths available in the NGA data set, the residuals in Figure 5-8 do not show a negative correlation with source depth, indicating that there is some compensating effect that is occurring.

The PGA event terms from the NGA data set for earthquakes with M<6 are shown in Figure 5-9 as a function of the source depth. This figure shows that there is a correlation of the event term with the depth: earthquakes with shallower depths have smaller event terms. This indicates that the deeper earthquakes are more energetic than the shallower earthquakes. One interpretation is that there is a depth dependence of the median stress-drop with deeper earthquakes having higher stress-drops on average than shallow earthquakes. As a result, the expected distance dependence of the standard deviation is counter-acted by the depth dependence of the stress-drop.

Recently, in a study of southern California earthquakes, Shearer (2005) found that stress-drops increase with depth for shallow depths which is consistent with the results from the NGA data

set; however, other studies have not found stress-drop scaling with depth (Hauksson and Shearer, 2005).

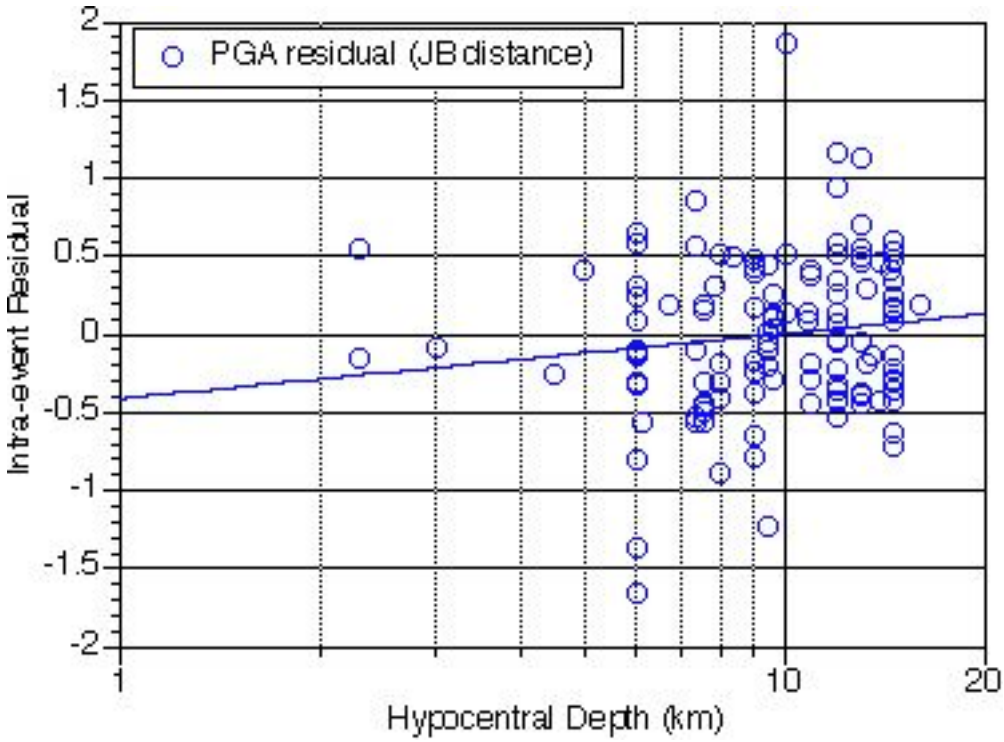

Figure 5-8
PGA intra-event residuals of the NGA data set for small magnitudes (M<6) and short distances (R_{JB}<10 km).

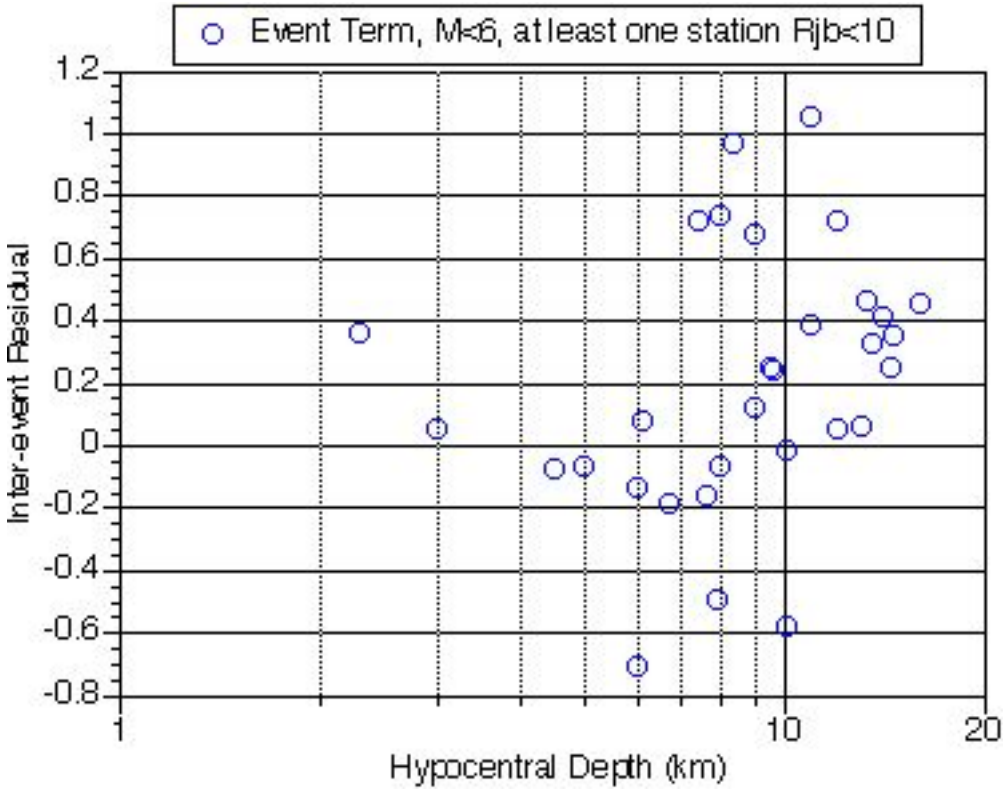

Figure 5-9
PGA inter-event residuals of the NGA data set for small magnitudes (M<6) and short distances (R_{JB}<10 km).

One short-coming of the evaluation of the NGA data set in terms of its relevance to the CEUS is that the depth range of the WUS data set is smaller than the depth range of earthquakes in the CEUS which continue to 30 km depth. The European data used by Bommer et al (2003) has a depth range that is similar CEUS range, but the average depths in the European set are similar to the NGA set. Using the European data set, a similar analysis of the hypocentral depth dependence of the residuals was made. The total residual for PGA is shown in Figures 5-10 and 5-11 for R_{JB} distances less than 10 km and 15 km, respectively. For both plots, the slope in the residuals that is expected if the ground motion attenuated as 1/R is shown. For the R_{JB}<10 km set (Figure 5-10), the computed slope is slightly positive. For the R_{JB}<10 km set (Figure 5-10), the computed slope is -0.30; which has the expected sign. The estimate of this slope is sensitive to the single residual at the shallowest depth. If this point is removed, then the slope is reduced to -0.20. For both subsets of the data, the observed slope with hypocentral depth is much weaker than expected.

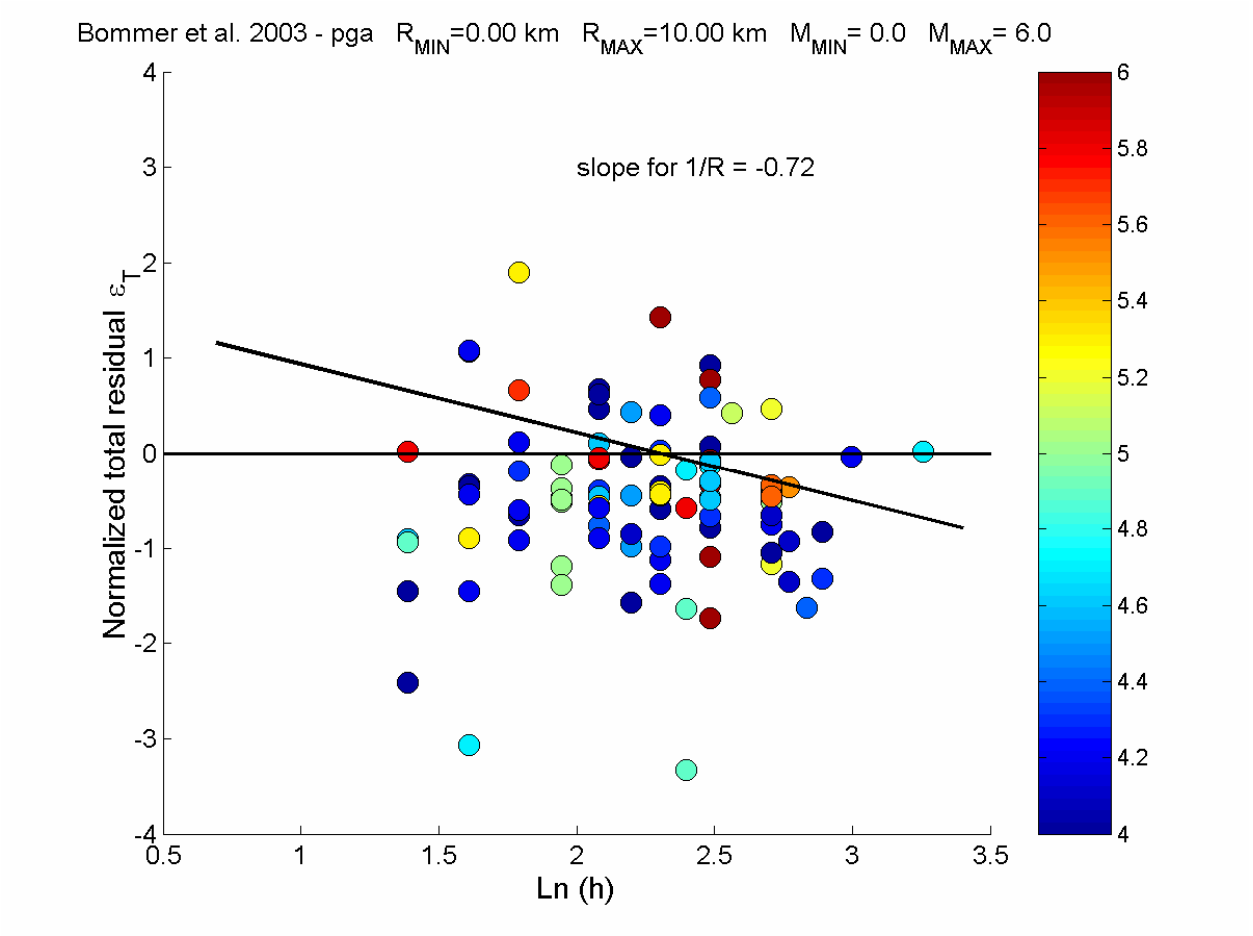

Figure 5-10
PGA total residuals of the Bommer et al. (2004) data set for distances R_{JB}<10 km. The line shows the expected trend of the residuals if the ground motion followed 1/R.

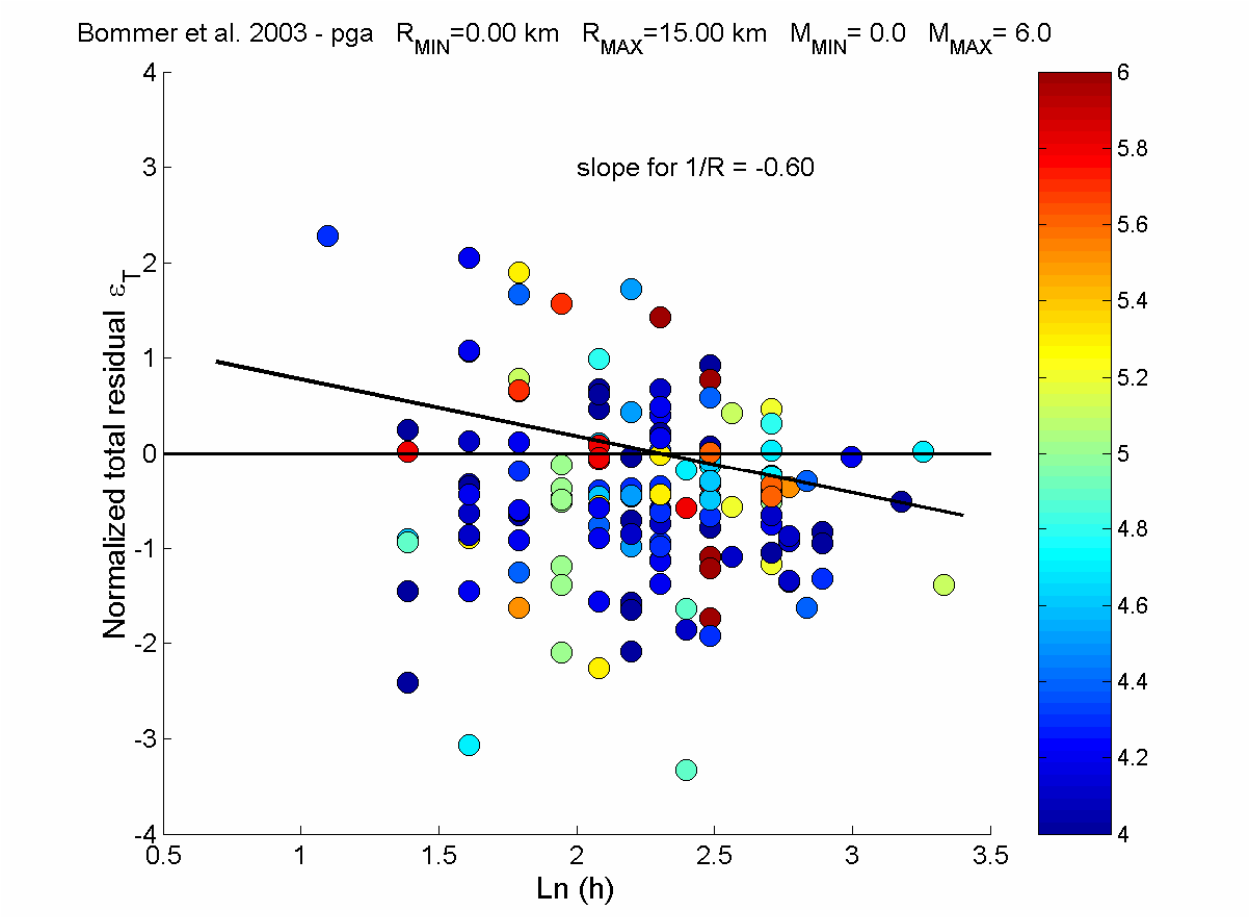

Figure 5-11
PGA total residuals of the Bommer et al. (2004) data set for distances R_{JB}<15 km. The line shows the expected trend of the residuals if the ground motion attenuation followed 1/R.

Sigma from K-NET and KiK-NET data

Strasser & Bommer (2005c) present an analysis of intra-event variability based on data from the recently installed Japanese strong-motion networks Kyoshin-Net (K-NET) and Kiban-Kyoshin-Net (KiK-NET) operated by the National Research Institute for Earth Science and Disaster Prevention (NIED). Due to their fully automated character, these networks are able to record and disseminate large quantities of digital strong-motion data, which represents a potentially very useful resource for the investigation of ground-motion variability associated with small-to-moderate events ($M_w \leq 6.0$) such as those relevant to the current project

This abundance of data could prove particularly useful for the study of intra-event variability, which empirical GMPE suggest constitutes the major component of ground-motion variability. It is, however, unclear how much this conclusion is affected by the fact that the intra-event variability is in most cases computed using records from several earthquakes. To this effect, Strasser & Bommer (2005c) select moderate (M_{JMA}>5.0) shallow (h<30 km) crustal events from the K-NET and KiK-NET databases that can be considered “well-recorded”. The selection

makes use of the search engine provided on the NIED website. To be selected, at least 50 records must be available for the event on either of the networks. This results in the selection of 82 events, 25 of which are from the 2004 Niigata sequence. Somewhat disappointingly, the associated ground-motions tend to be low-amplitude motions from distant stations, as illustrated in Figure 5-12, reproduced from Strasser & Bommer (2005c). The boxplots identify the ranges covered by the data, with the boxes representing the central quartiles. These plots also feature the cut-off distance suggested by Atkinson & Boore (2003) for K-NET data from events with $M_w < 6.0$ to avoid the effects of instrument limitations at high frequencies (thick gray line in the left panel), as well as the approximate lower bound for damaging motions determined by Martinez-Pereira & Bommer (1998, thick gray line in the right panel). This underlines the fact that this abundance of data available from K-NET and KiK-NET is as much a result of the sensitivity of the instruments as of the density of the networks.

The matter is further complicated by the fact that some files associated with a single event in the database actually include arrivals from several events. This feature of the data can be traced back to the fully automated nature of the data treatment and leads to distortions of the spectral shape, which precluded the investigation of spectral response residuals in the Strasser & Bommer (2005c), which instead focused on single-event intra-event variability associated with PGA. No automated methods for the removal of unwanted portions of the records are currently available, and even when the painstaking approach of truncating the records individually is taken (Pousse *et al.*, 2005), the scatter associated with GMPE derived using large sets (several thousand data points) of strong-motion records from K-NET and KiK-NET is of the same order of magnitude or even larger than the scatter associated with GMPE using smaller and less recent datasets.

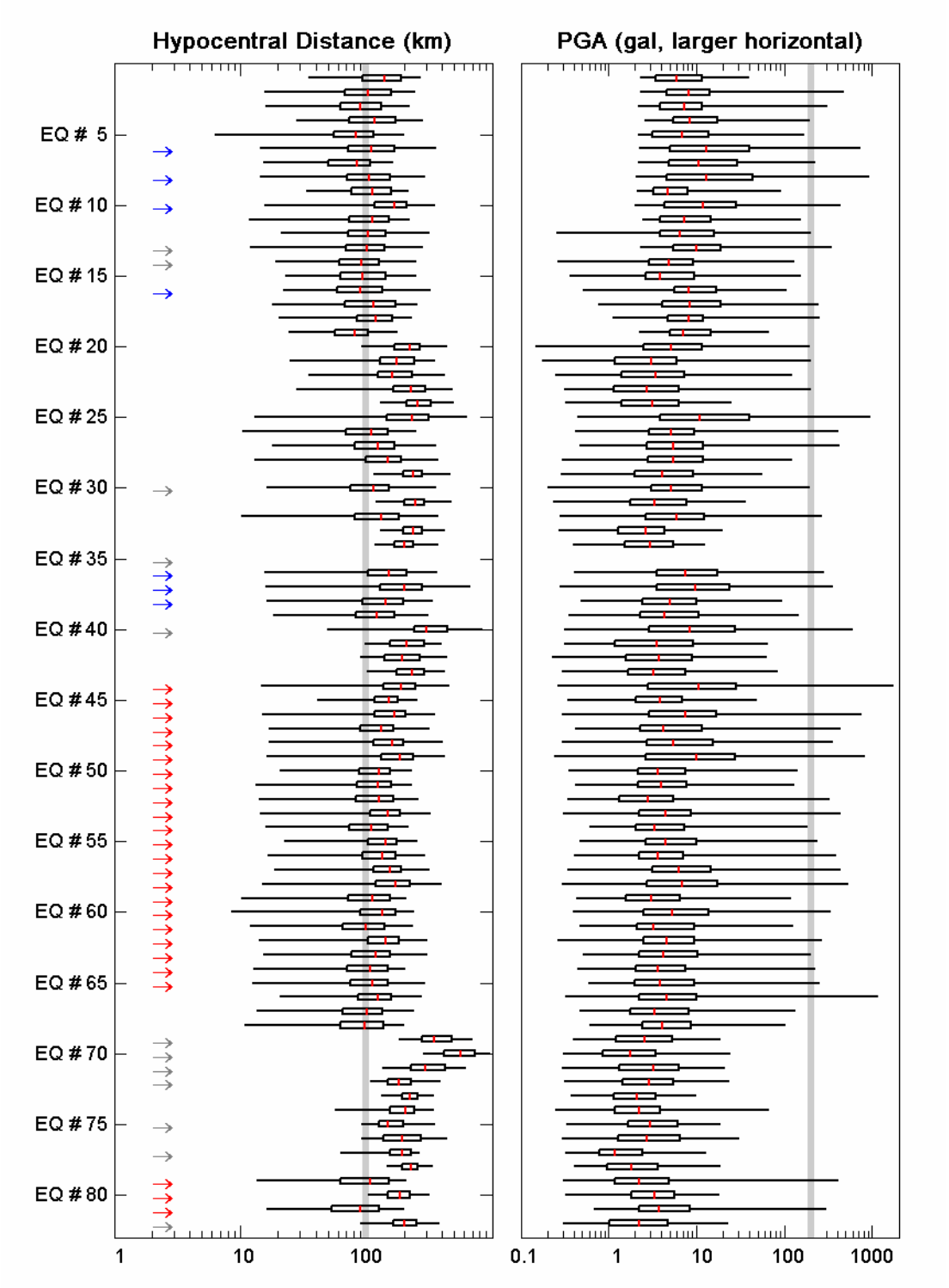

Figure 5-12
Distribution of data available for the selection of 'well-recorded' shallow events from the K-NET and KiK-NET databases in terms of hypocentral distance (left , using the hypocentral location provided by NIED) and PGA (right, larger horizontal component). (Strasser & Bommer, 2005c).

The prevalence of quantity over quantity extends to the description of the site conditions, which is available for a vast majority of the K-NET stations (1031 out of 1034); however, the profiles available for the K-NET sites are rather shallow, with depths limited to 20 m and not exceeding 10 m in about 50% of the cases. Considering that most K-NET stations are located on deep sedimentary sites, this might lead to classification errors when schemes based on parameters requiring a deeper profile (e.g. $V_{S,30}$) are used. In the Strasser & Bommer (2005c) study, the extrapolation was limited to 20 m. Little difference with an extrapolation to 30 m was found, which however gives no indication about the validity of the extrapolation. On the other hand, the 675 KiK-NET stations are associated with deep boreholes, for which reason a good site characterization could be expected, but documentation is readily available for about only 20% of the sites.

The Strasser & Bommer (2005c) study examined intra-event residuals for a number of events which are reasonably well documented in terms of source characteristics. Small-to-moderate (M~6.0) events associated with a complex source process are capable of generating large-amplitude high-frequency motions and can be damaging, as illustrated by the 1997 Kagoshima and 2003 Miyagi events. It is, however, unclear how relevant these observations are to other tectonic frameworks, in particular stable continental regions such as the Central and Eastern United States. Choices made during the regression process, such as the choice of the horizontal component definition, the choice of the distance metric, or the decision whether a site classification is included in the regression or not, have little impact on the intra-event variability or the patterns followed by the residual dataset globally. Similarly, the inclusion of an anelastic attenuation term hardly makes any difference to the results, especially since the range of distances needs to be limited in order to avoid low-quality data. No pattern could be found linking the highest residuals in the dataset to a particular site class. The azimuthal position of the site with respect to the source might be a contributing factor in some cases where the stations in the strike-normal and strike-parallel directions exhibit clearly different residual patterns. Other events, however, exhibit fairly random patterns overall.

In all cases, a reasonably good fit to the log-normal distribution is observed within the +/- 2 σ range. The intra-event variability σ determined from single events ranges from 0.58 to 0.78 ln units and is thus quite large when compared to the values found in empirical predictive equations. While the highest values of σ can probably be related to the propagation of location errors accentuated by the use of data from a single event, values for which no such explanation is available may reach 0.71 ln units. The value of σ is also strongly correlated with the sample size, as illustrated in Figure 5-13. This increase of σ with sample size suggests some systematic error in the dataset, possibly erroneous associations between ground-motion values and source parameters induced by the fully automated nature of the data treatment. Indeed, mismatches can be found between the source parameters listed in the K-NET and KiK-NET data files and those obtained from the F-NET seismometric network also operated by NIED.

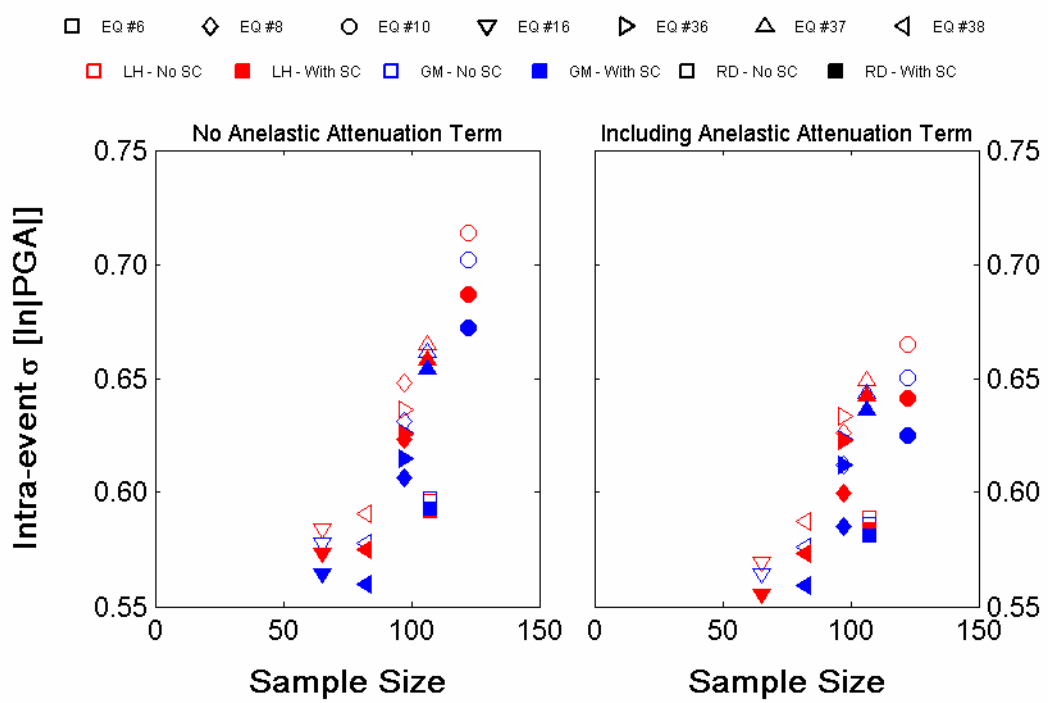

Figure 5-13
Influence of sample size on sigma for selected "well-recorded" events from the K-NET and KiK-NET databases.

The 2004 Niigata sequence provides a unique opportunity of studying the spatial variability of ground-motion residual distributions for events that have a very similar source mechanism, comparable magnitudes and have been recorded on the same network of stations. The recent nature of these earthquakes implied that the documentation available regarding these events was limited at the time the Strasser & Bommer (2005c) study was carried out. In general, no consistency could be found between the intra-event residuals recorded at the same station for a large number of events in the sequence. A few stations exhibit more consistent patterns than others, with residuals tending to have a common sign (positive or negative), but even in those cases exceptions are observed. The highest residuals are reasonably consistent across horizontal component and distance definitions, but there is a distinct lack of correlation across events.

The values of PGA intra-event variability found in the Niigata sequence are consistent with values found from empirical predictive equations, ranging from 0.53 to 0.71 ln units. These values seem to increase with magnitude in the limited range considered (M_{JMA} = 5.0 to 6.8). This is at odds with the findings from some predictive equations using a heteroscedastic model for variability, which find a decrease of intra-event variability with magnitude. A possible explanation is that the apparent dependence of the intra-event variability on magnitude found in the present study is related to a sample size effect: the intra-event variability is also found to increase with sample size, which in turn increases with magnitude due to the combined effects of the magnitude-distance correlation in the available data and the constraints imposed on the data in terms of distance and ground-motion amplitude. Again, this would point to some systematic error in the association of ground-motions and explanatory parameters. This type of explanation

is rendered even more likely in the case of the Niigata sequence by the extremely complex nature of the source process, which included a larger number of aftershocks than expected on average for crustal earthquakes in Japan, and involved rupture on five different fault planes (Hikima & Koketsu, 2005).

To conclude, the data provided by K-NET and KiK-NET is not thought to be of sufficient quality to allow the investigation of the intra-event variability of ground motions associated with moderate (M ~ 6.0) shallow crustal events using data from a single event. Data quality issues affect more particularly the spectral response ordinates. Individual examination of the records may provide some clues as to which particular features of the data are responsible for causing large residual values related to poor data quality, but such a time-consuming approach is totally unpractical for the large sample sizes required in variability computations. This is even more true if one considers that many of the factors found to lead to data quality problems in the present study are unlikely to be relevant to regions other than Japan. In particular, the complex source geometries found in this study for crustal events associated with volcanic activity or subduction tectonics have no equivalent in a stable continental region such as the Central and Eastern United States. Similarly, the strong resonance observed at some response periods is only likely to occur at deep soil sites and need not be considered if rock motions are of interest; however, due to the large degree of overlap between datasets recorded during different events, particularly during the 2004 Niigata sequence, the K-NET/KiK-NET data could still potentially contribute to improved estimates of variability once the data is subjected to a careful screening process such as the one that was applied to the Taiwanese data collected during the 1999 Chi-Chi sequence (Lee *et al.*, 2001).

Sigma from numerical simulations

Strasser & Bommer (2005b) present an analysis in terms of ground-motion variability of the results of numerical simulations carried out using the deterministic-stochastic kinematic finite-source model EXWIM, which has been described in Chapter 4, which also discusses the implications of these simulations in terms of ranges of residual values. This section, on the other hand, focuses more specifically on variability computations.

Simulations such as those presented in this study allow investigating ground-motion variability at short source-to-site distances, where empirical data are generally scarce. A major issue in this type of investigation, however, is to ensure the compatibility between variability estimates estimated from the simulations and the variance components derived in empirical GMPE. No problem exists for the intra-event variability, which is defined as the spatial variability of ground motions for a given set of source parameters; however, apart from distance, the parameters varied in the Strasser & Bommer (2005b) are not routinely included as explanatory variables in empirical GMPE, mainly because of their unpredictability. Therefore, a non-parametric study was carried out, and several measures of inter-event variability defined to investigate the influence of the various source parameters. Table 5-2 summarizes the links (and lack thereof) between the measures of variability defined in Strasser & Bommer (2005b) and those commonly used in empirical GMPE, and whether they include a spatial or temporal component. For the simulations, the temporal component corresponds to variability from repeated runs with different model assumptions. All equivalences assume rock conditions for the empirical equations.

Table 5-2
Variability definitions in empirical GMPE and using the numerical simulations presented in Strasser & Bommer (2005b). The columns labeled 'S' and 'T' indicate whether a spatial or temporal component is included in the definition.

	S	T	Empirical GMPE for rock	Numerical Simulations
TOTAL	●	●	Total Variability σ_{TOTAL}	σ_{SAMPLE} in residual analysis (see Chapter 4)
INTRA	●		Record-to-record variability σ_{INTRA}	Not defined
	●		Site-to-site variability σ_{S}	Intra-event variability σ_{INTRA}
		●	'Residual' record-to-record variability σ_{R}	Not defined
INTER		●	No equivalent (GMPE do not consider azimuth to characterize receiver)	Overall inter-event variability σ_{INTER}
	●	●	Inter-event variability σ_{INTER}	'Pseudo inter-event' variability σ_{PSEUDO}
		●	No equivalent	Inter-event variability at a fixed level of stress drop
		●	No equivalent	Inter-event variability for a fixed slip distribution

The values of σ_{INTRA} calculated range from 0.04 to 1.77 ln units, which corresponds to multiplicative factors on the ground-motion amplitude ranging from about 1 to almost 6. These ranges have been found to depend on the stress drop level, the highest variability being found for medium stress drop level. The largest values of σ_{INTRA} correspond to shallow nucleations and the value of σ_{INTRA} from all of the depths will depend on the relative rate of the shallow events compared to the deeper events. If a smaller value of σ_{INTRA} is found, then the epsilon values from the shallow nucleations will be even larger, providing more evidence that the distributions should not be truncated at a maximum epsilon.

The central part of the data (defined by the second and third quartiles) shows little dependence on the stress drop level, and typically spans the range [0.20-0.90], with median values of about 0.35 to 0.70. These values compare reasonably well with the empirical values, although they tend to be higher. A possible explanation is that the datasets used for the derivation of empirical predictive equation will usually be dominated by data from larger distances than those considered in the Strasser & Bommer (2005b) study, while intra-event variability is likely to be influenced by near-source phenomena such as directivity.

The value of the inter-event variability is strongly dependent on how 'events' are defined in the simulations. Model assumptions may result in biases when they stem from convenience rather than physical considerations. The size of the sample of events over which the variability is computed will also strongly affect the results. Values of σ_{PSEUDO}, the measure of inter-event

variability from the simulations that is suitable for comparison with empirical GMPE, have been found to be significantly higher than the empirical intra-event variability σ_{INTER}. These values range from 0.55 to 1.45 ln units when uniform weighting is applied; this range is slightly reduced to 0.60 to 1.20 ln units when a simple weighting scheme is applied. The purpose of the weighting scheme is to reduce the biases induced by the uneven number of events contributed by different stress drop levels and to modulate the influence of nucleation point locations depending on their location on the fault rupture plane, based on the results of Mai *et al.* (2005). Thus, the values of inter-event variability found in the simulations are comparable to the values of intra-event variability. Again, this could be an effect of the distance range considered, receivers located very close to the source being more sensitive to details of the rupture process. The inclusion of unrealistic and potentially unphysical source parameter combinations is another factor likely to influence the range of values taken by σ_{PSEUDO}. This highlights the need for better constraints on the source parameters used in simulations if the resulting ground-motions are to be used for variability calculations.

In all cases, an overall trend of increased variability with distance can be observed. In particular, receivers located above the fault exhibit much lower values than those located away from the fault; however, there is a considerable amount of dispersion in the values within one distance bin that can be correlated to variations in azimuth; in some cases, they are larger than the differences between adjacent distance bins. The increase of variability with increasing distance in the vicinity of the source is consistent with the finding by Midorikawa & Ohtake (2004) that variability decreases with decreasing distance for shorter distances (<50 km). Both these findings are at odds with results from other empirical GMPE, which show no dependence on distance. The rate of this increase tends to level off with distance, so the discrepancy between empirical results other than the study of Midorikawa & Ohtake (2004) and simulation results might be a consequence of the difference in the distance range considered.

Among the source parameters varied in the Strasser & Bommer (2005b) study, only the nucleation point location significantly influences the ground-motion variability. This is due to the fact that the nucleation point location will effectively control directivity effects. It is possible that the simulation method employed in the study overemphasizes these directivity effects even if unrealistic scenarios such as shallow nucleation followed by down-dip propagation are excluded. On the other hand, changing the stress drop level, slip distribution or replacing the uniform rake assumption with a variable rake assumption, has only a minor impact on the results.

Finally, the response period has been found to have almost no effect on variability. This contradicts the findings from empirical equations, in which the variance typically increases with increasing response period; however, other studies based on stochastic modeling (Atkinson & Boore, 1997a; 1997b) have also found a decrease of variability with response period, so this discrepancy might be a systematic effect. A slight drop-off of variability at longer periods is believed to be related to modeling assumptions, motions at periods longer than 1s being deterministic. A possible explanation for the insensitivity of the results to response period could be the uniformity of the site conditions in the simulations (‘perfect rock’ at all receivers), compared to the variety of site conditions found for empirical data even within a single site class. Thus it would seem that if the trend of increasing variability with increasing response period

observed in empirical data is real, it is not related to purely geometrical factors such as those investigated in the Strasser & Bommer (2005b) study.

This emphasizes the need of using results from several simulation methods to assess the contribution of modeling uncertainty to variability computations from numerical simulations. Further research needs include simulations to include the magnitude-dependence of ground-motions (which could not be investigated in the Strasser & Bommer study) as well as improvements in the calibration in order to restrict source parameter combinations to physically possible ones, so that simulation results can be used with confidence to constrain the tails of the ground-motion distribution. The realism of simulation results could also be increased by including a site-specific response.

6
SIGMA FOR EUS

EUS sigma models used in 2004 EPRI EUS ground motion study

The 2004 EPRI EUS ground motion study (EPRI, 2004) considered four alternative models for the aleatory variability. These four models are described below.

The first model is based on the empirical model of Abrahamson and Silva (1997) for WUS earthquakes with modifications for EUS conditions. This model assumes that the standard deviation from well recorded WUS strong ground motion data sets are applicable to stable continental regions with the exception that the standard deviation of the stress-drop in the CEUS was assumed to be larger than for the WUS. This increase in stress-drop variability in the CEUS was based on the assumption that the 1988 Saguenay earthquake was approximately a $+2\sigma$ stress-drop earthquake. The model was also modified to account for an increase in the standard deviation at short distances that occurs from the use of the JB distance metric if the ground motion is assumed to scale as 1/R (this distance modification was discussed in more detail earlier).

The second model is based on the Toro et al (1997) ground motion prediction equation which is based on the total aleatory variability estimated for the point source stochastic model (e.g. numerical simulation). Following Abrahamson et. al. (1990), there are two parts to the variability computed for numerical simulations: modeling variability, and parametric variability. The modeling variability is empirically derived and is computed from the misfit between observed and predicted ground motions when the event-specific model parameters (such as stress-drop) are optimized for each earthquake. The parametric variability is computed from suites of forward simulations that sample the distribution of event-specific parameters that were optimized in the estimation of the modeling variability. Silva (1996) updated his estimate of the modeling variability based on comparisons of the model predictions using a larger number of earthquakes. The updated modeling variability is larger than the modeling variability used in the Toro et al. (1997) model. For example, at PGA, the standard deviation of the modeling variability was increased from 0.32 to 0.48 natural log units. In the EPRI (2004) model, standard deviation given in Toro et al (1997) was increased to account for this larger modeling variability term.

The third model is based on the USNRC (2002) model. The USNRC model was revised to account for an increase in the standard deviation at short distance (R<25 km) to account for the use of JB distance. The fourth model is based on Silva (2002).

The resulting standard deviations are shown in Figures 6-1, 6-2, and 6-3 for three cases: M=5, R_{JB}=5 km; M=6, R_{JB}=10 km; M7, R_{JB}=30 km. For the M5, R5 case, the standard deviations from the EPRI models are, on average, about 0.2 natural log units greater than the standard deviations from the NGA model. For the M6, R10 km case, the differences are about 0.1 natural log units, and for the M7 R30 case, the models are similar.

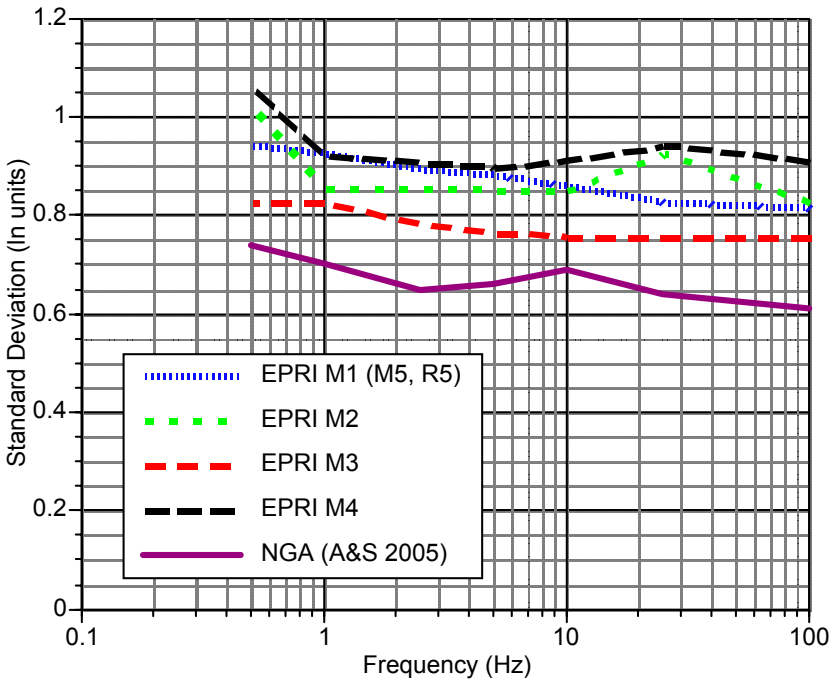

Figure 6-1
Comparison of the EPRI (2004) models of the standard deviation for M=5, R_{JB}=5km with the results of the preliminary NGA model of Abrahamson and Silva (2005).

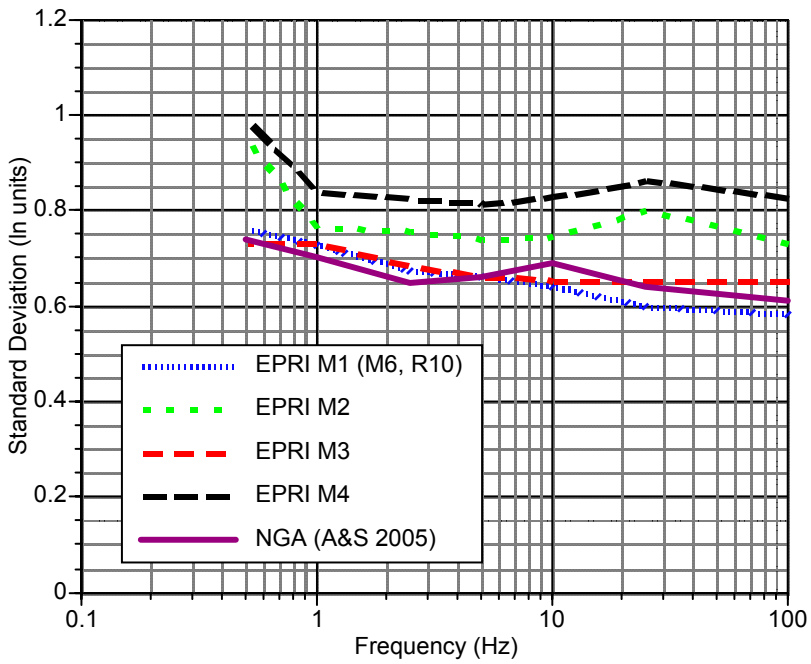

Figure 6-2
Comparison of the EPRI (2004) models of the standard deviation for M=6, R_{JB}=10km with the results of the preliminary NGA model of Abrahamson and Silva (2005).

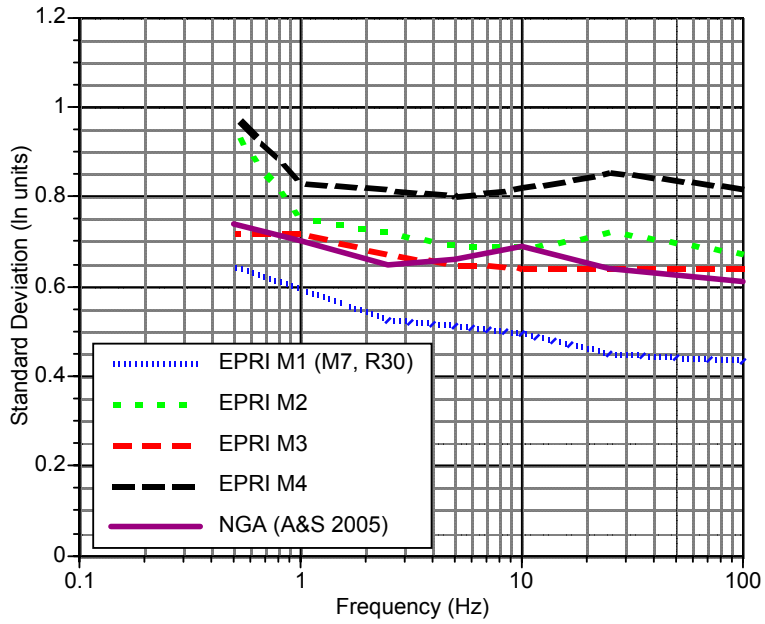

Figure 6-3
Comparison of the EPRI (2004) models of the standard deviation for M=7, R_{JB}=30km with the results of the preliminary NGA model of Abrahamson and Silva (2005).

The EPRI (2004) standard deviations at short distances and small magnitudes are much larger than found for other regions with large data sets. An important question is why should the aleatory variability for the EUS be so much larger than for other regions with more data. We emphasize that the issue here is the value of the aleatory variability, not the epistemic uncertainty. The differences in the epistemic uncertainty for EUS vs. WUS is captured through the use of the alternative models for the median ground motion and the alternative models for the aleatory variability. Below, we address the applicability of the aleatory variability from other regions to the ground motions in the EUS.

Applicability of sigma from other regions to EUS

As noted above, for small magnitudes and short distances, the aleatory variability in the EPRI (2004) model is much larger than for the NGA models for the WUS. With the sparse strong motion data set for the EUS, we need to consider if this increased standard deviation is warranted. That is, is there a reason to expect that the inter-event variability and/or intra-event variability for EUS earthquakes is significantly different from that of WUS earthquakes?

Event-term variability inferred from catalogs

Ideally, we would have strong ground motions from a large set of earthquakes in the EUS that could be used to compare with the inter-event variability from the WUS data sets. As a proxy for a large set of earthquakes with recorded strong ground motions, earthquake catalogs are used. Simple measures of the strength of the source that are available from earthquake catalogs are used. We then compare the variability of these simple source parameters for earthquakes in active regions and earthquakes in stable regions. If the standard deviation of the simple source

parameter is similar between the two regions, then we assume that the inter-event standard deviation from the WUS would be applicable to the EUS.

Data Sources

Hypocentral information (origin time, latitude, longitude, focal depth), Ms, and mb estimates were compiled from the Internatinoal Seismologucal Center (ISC) online bulletin database. The data search was restricted to events with ISC magnitudes (of any type) greater than or equal to 5.0 and focal depths less than 35 km.

Estimates of the half duration of slip on the source and the centroid time measured from the origin time were compiled from the Havard online catalog. The half duration is a measure of how long the source is slipping. A shorter duration would lead to larger ground motions since the energy is packed in to a shorter time. The ISC catalog is available for events between 1978 and 2002. The USGS PDE is used for earthquakes occuring after 2002.

Tectonic regions

The data from the earthquake catalogs were associcated with the following tectonic regions: shallow active regions, subduction zones. oceanic regions, and stable regions. As an initial classification, the earthquakes were classified by the 1995 Flinn-Engdahl regionalization scheme (Young et al , 1996). The earthquakes are shown grouped by geographic region in Figure 6-4. By their nature, stable regions have few earthquakes. The East Afrian rift is a zone of thin crust with active spreading that may not be representative of earthquakes ocurring along the New Madrid zone since New Madrid is a failed rift that is now in compression. In addition, the cluster of events from NE China (just west of Korea) may represent active regions and an earthquakes near Indonesia may be a subduction event. Therefore, we considered three subsets for the stable regions: one that includes all the events that are yellow in Figure 6-4; one that excludes the East African Rift from stable regions, and one that excludes the East African Rift, NE China, and the possible subduction events near Indonesia.

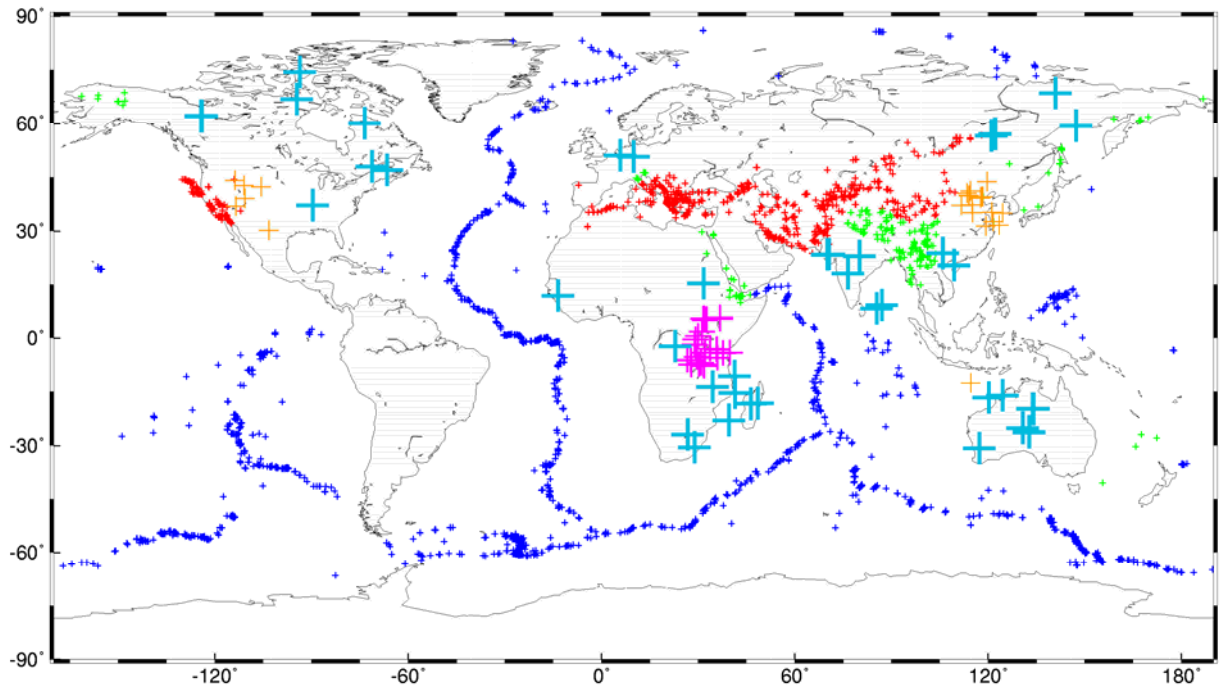

Figure 6-4
Map of M>=5 earthquakes from 1978-2004, excluding subduction zones. The colors indicate the Flinn-Engdahl region: Red=active regions, blue=oceanic, yellow=stable regions, and green=regions classified as stable but considered to be active or oceanic.

Source Variability

Our objective is to compare the source variability for stable continental regions and active regions. This is done by comparing the standard deviation of the source half durations for earthquake soruces in these two regions. The source half duration results were fit to the following model:

$$\ln(HD(\sec)) = c_1 + c_2 M_S \qquad \text{(Equation 6-1)}$$

The coefficients c_1 and c_2 are estimated using ordinary least-squares. The fits to the data from active regions and the three subsets of stable regions are shown in Figure 6-5.

The standard deviation of the fits are listed in Table 6-1. This comparison shows that there is no increase in the standard deviation for stable regions as compared to active regions. This suggests that the gross source properties for earthquakes in stable regions are no more variable than those in active regions.

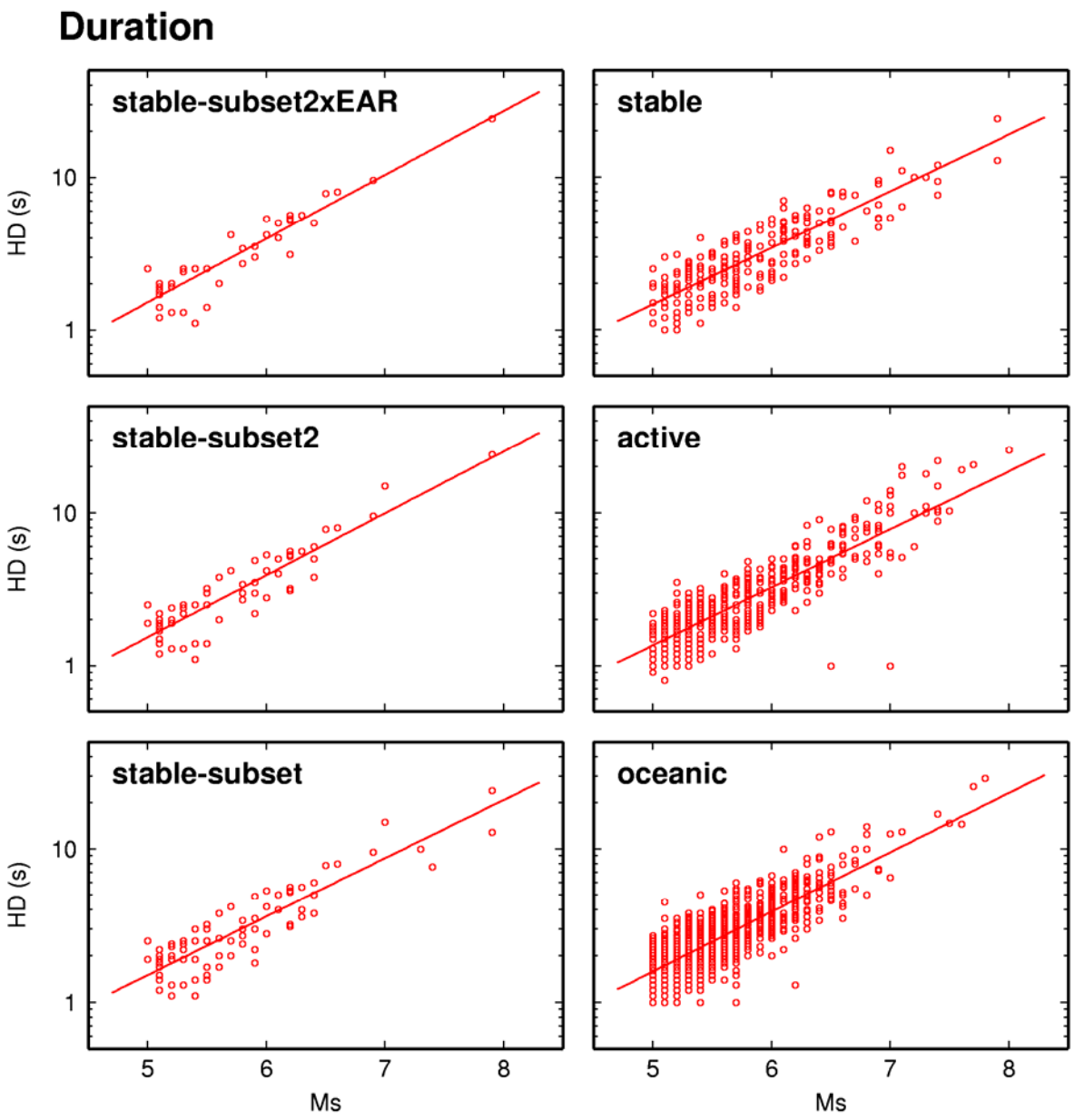

Figure 6-5
Regression fits to the source half-durations from the catalog data. The top right frame has all events that fall into Stable Regions (yellow and green in Figure 6-4). The left hand side has stable subset 1, 2 and 3 in the bottom, middle, and top frames, respectively.

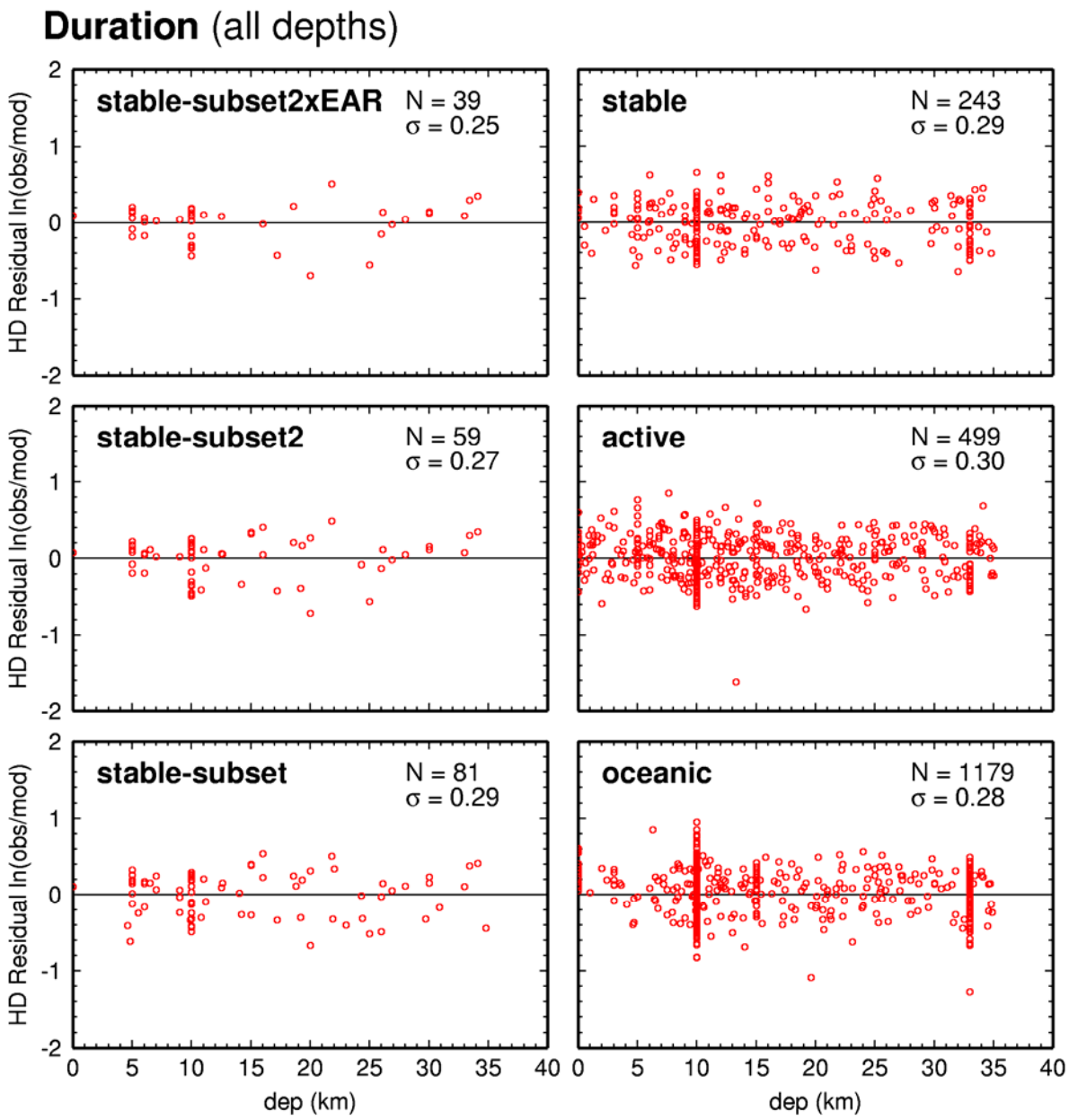

Figure 6-6
Residuals for the source half-duration model from the catalog data. The top right frame has all events that fall into Stable Regions (yellow and green in Figure 6-4). The left hand side has stable subset 1, 2 and 3 in the bottom, middle, and top frames, respectively.

Table 6-1
Standard deviation of source half durations from earthquake catalogs (1978 – 2001)

	Source Half-Durations	
Region	N	Sigma (Ln units)
Active	499	0.30 ± 0.01
Stable subset1	81	0.29 ± 0.02
Stable subset 2	59	0.27 ± 0.02
Stable subset 3	39	0.25 ± 0.03

Source-term variability due to limited source parameters

The WUS attenuation relations include more source parameters than the CEUS models such as style-of-faulting, and hanging wall/footwall terms. These additional parameters have a negligible effect on the total standard deviation. Repeating the NGA regresssions without these additional model parameters increases the total standard deviation by less than 0.01 natural log units. This change is smaller than the accuracy of the estimate of the standard deviation and therefore is not considered further.

Path-term variability inferred from simulations (URS)

Another source of possible differences in the standard deviations between the WUS and CEUS is the variability of the crustal structure for these two regions. To evaluate this possible difference, high frequency simulations were performed using the URS semi-stochastic method (Graves and Pitarka, 2004) to compare the variability of ground motions from stable and active regions due to the crustal structure variability in these two regions. The goal of the simulation is to determine if there is an expected difference in the variability of the ground motion due to known differences in the range of crustal structures for active and stable regions. For each region, simulations were conducted for a suite of crustal structures for a M=5.5 earthquake. A small magnitude was used in the simulations since smaller magnitude events will be more sensitive to the variability in hypocentral depth than larger magnitude earthquakes.

For the CEUS, 16 regions developed by EPRI (1993) are used to sample the range of crustal models applicable to the CEUS (Figure 6-7). The 1-D crustal models for these 16 regions are shown in Figure 6-8. For each model, a surface shear-wave velocity of 2830 m/s is used.

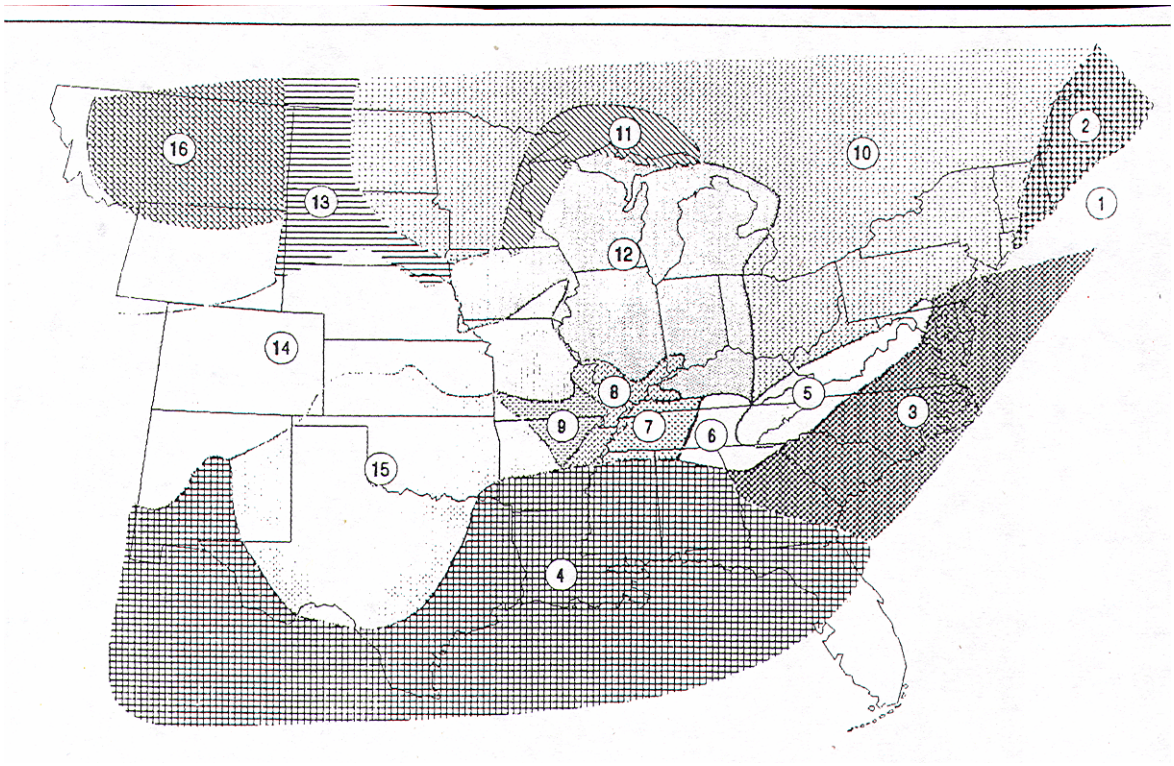

Figure 6-7
Sixteen Regions for the CEUS used in the EPRI (1993) ground motion study.

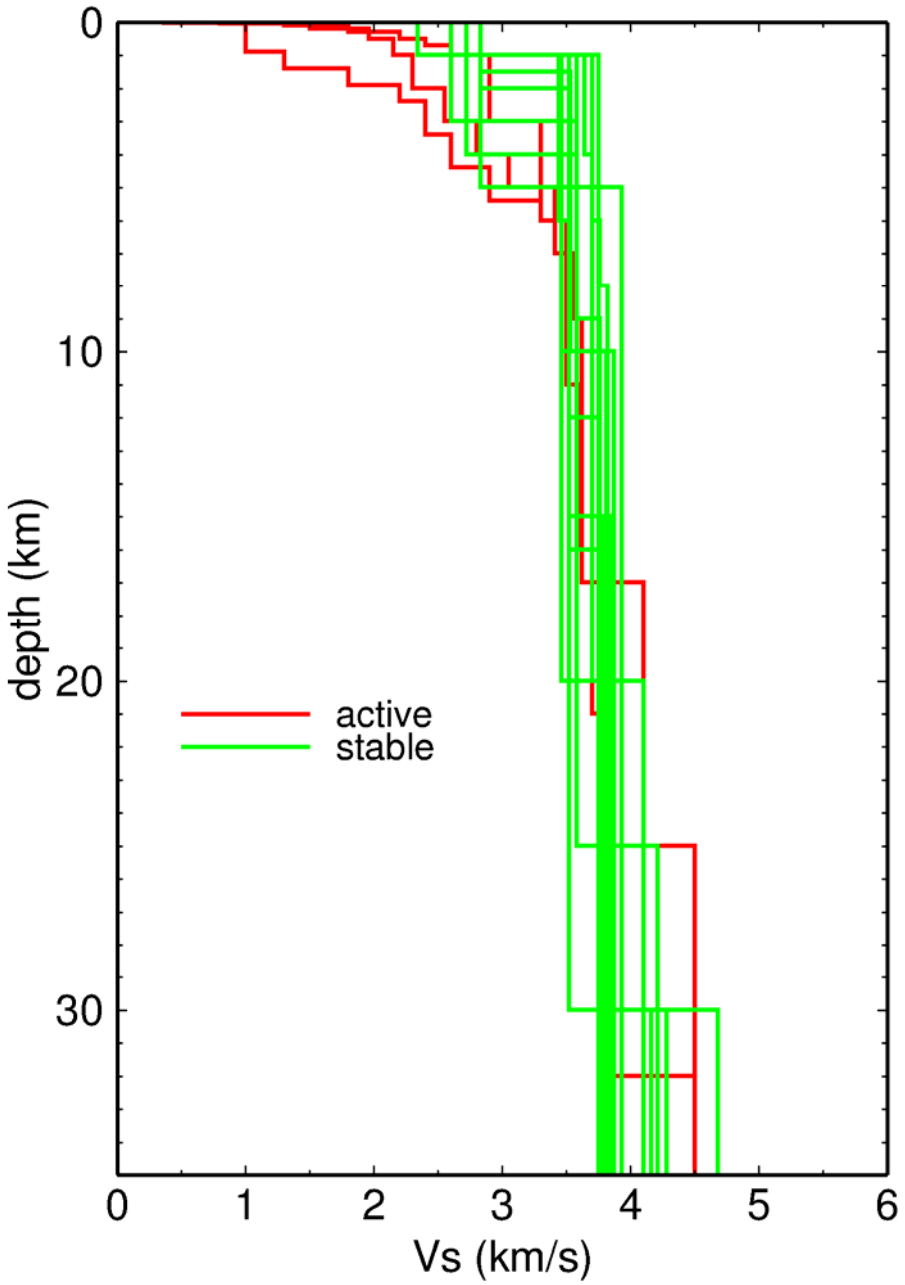

Figure 6-8
Velocity profiles for the 16 EUS regions shown in Figure 6-7 (Green) and for the three CA models (red).

For the WUS, 3 crustal models were considered: a southern California rock model, a southern California basin model, and a northern California model. For the active region, all of the models (including the basin model) have a surface shear-wave velocity of 760 m/s. These models are shown by the red curves in Figure 6-8.

For both the stable and active regions, the simulations were conducted at a suite of 40 stations located at distances up to 50 km. The station loations are shown in Figure 6-9. Seven source depths were sampled: 5, 10, 15, 20, 25, 30, and 35 km. For each depth and site, six different realizations were run using a different random seed and a different source mechanism. This lead to a total of 1680 3-component accelerograms for each crustal model.

A weighted regression analysis was conducted to fit the simulated ground motions to a simple attenuation curve. The weights were applied to appriximate the average depth distributions for each region. For the stable regions, the depth distribution from EPRI (1993) is used and for the active region, the distribution of focal depths in California from 1970-2005 is used. These two depth distributions are shown in Figure 6-10. The depth distribution for the CEUS model extends to larger depths which could lead to greater variability for CEUS ground motions than for the WUS.

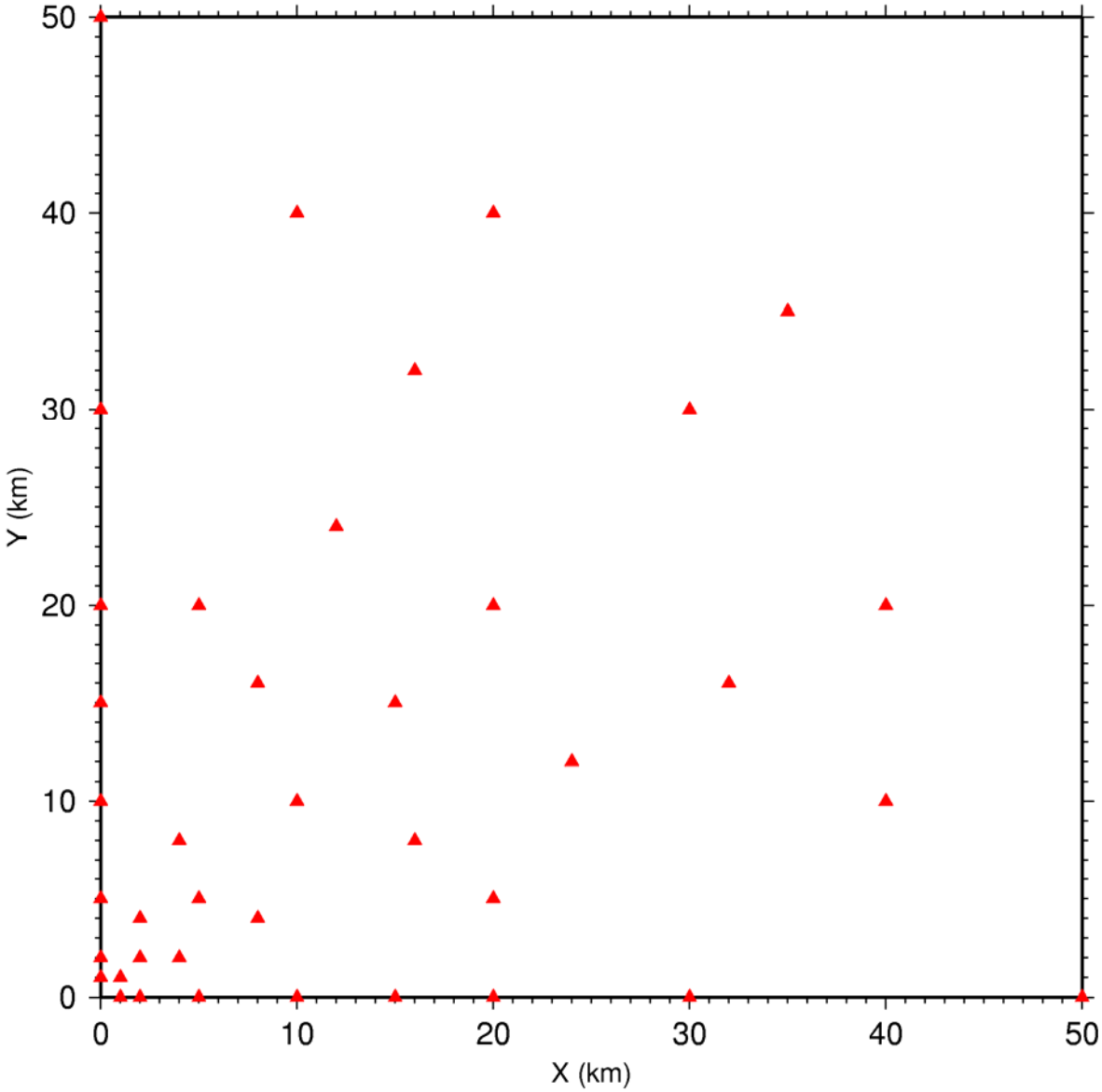

Figure 6-9
Station locations used for the simulations for evaluations of variability due to crustal structure.

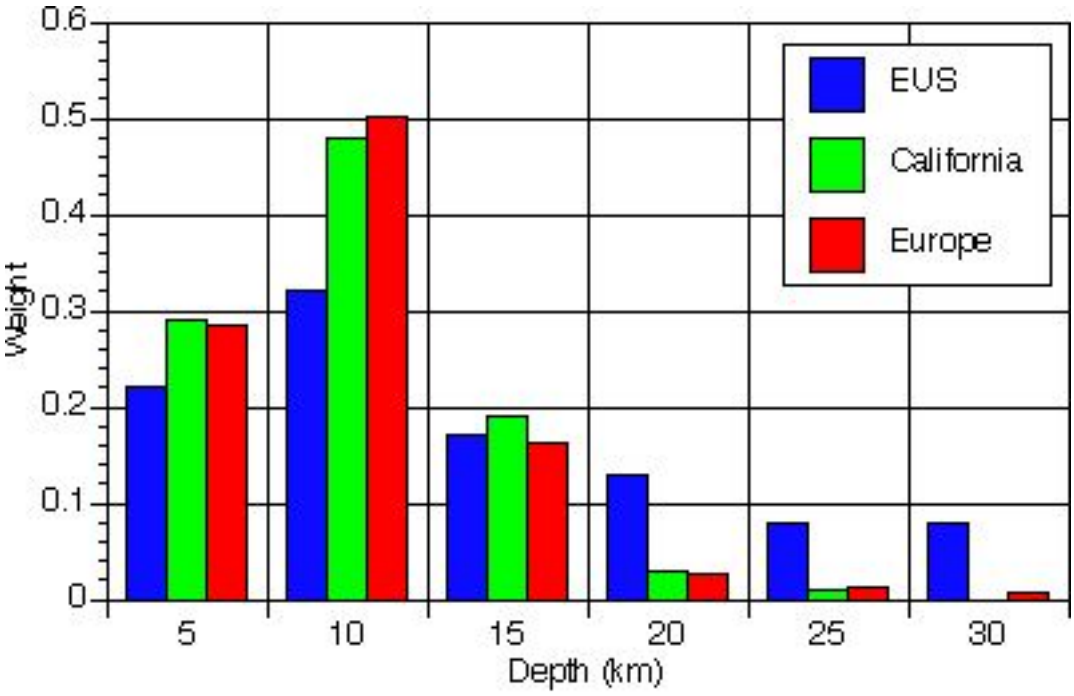

Figure 6-10
Depth distributions for the CEUS and California. The distribution from depths from European strong motion data sets is similar to that in California.

The simulated 10 Hz response spectral values are shown in Figure 6-11 and 6-12 for the CEUS and CA simulations, respectively. These simulations were fit to simple attenuation relations. The resulting standard deviations are shown in Table 6-2. This table shows that there is no expected increase in the standard deviation due to crustal structure differences between the CEUS and the WUS.

Intra-event variability could also be caused by variabiliy in Q which we have not considered here because Q has a small effect at distances of less than 100 km which dominate the WUS strong motion data used for developing the variability .

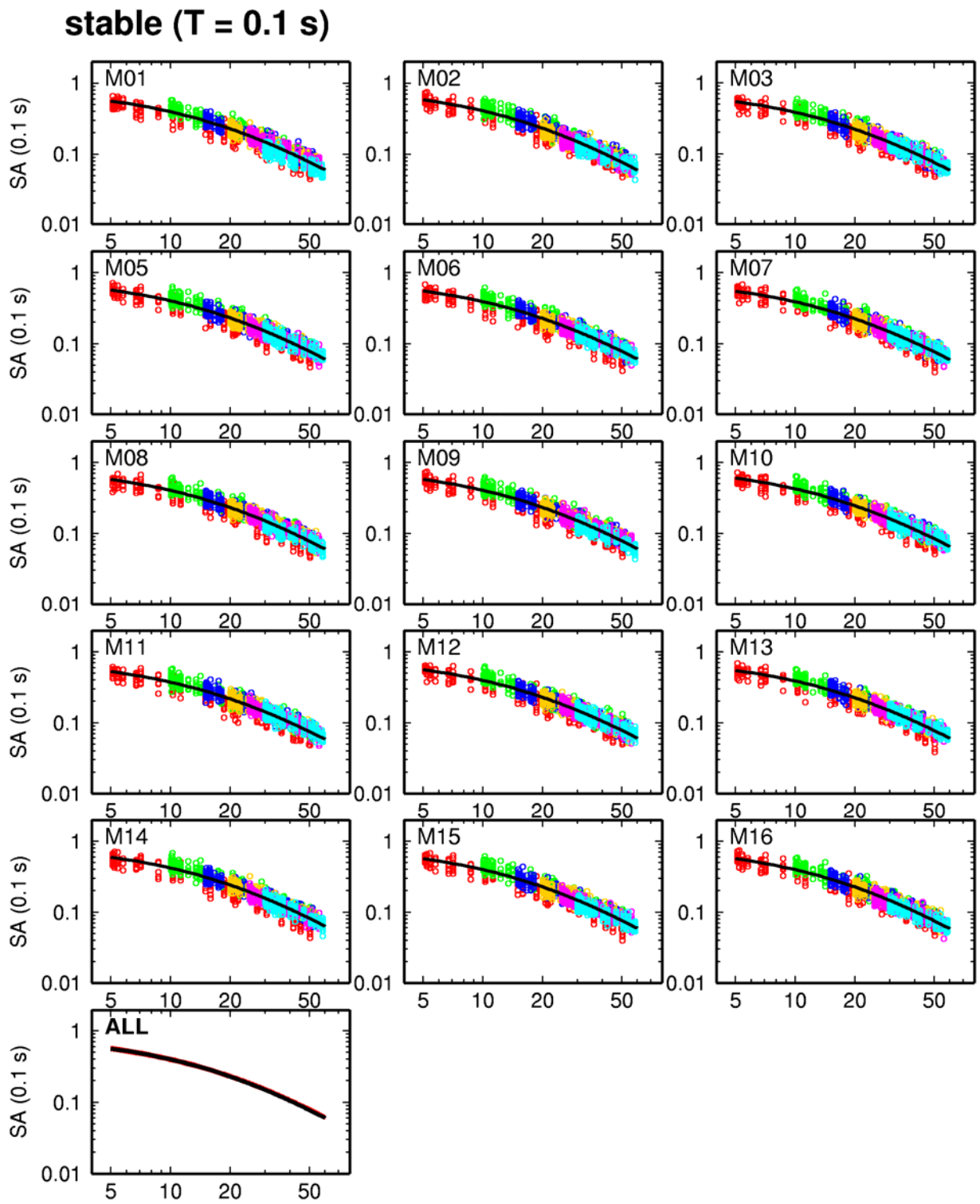

Figure 6-11
Simulations for the CEUS crustal structures. The colors are for different hypocentral depths.

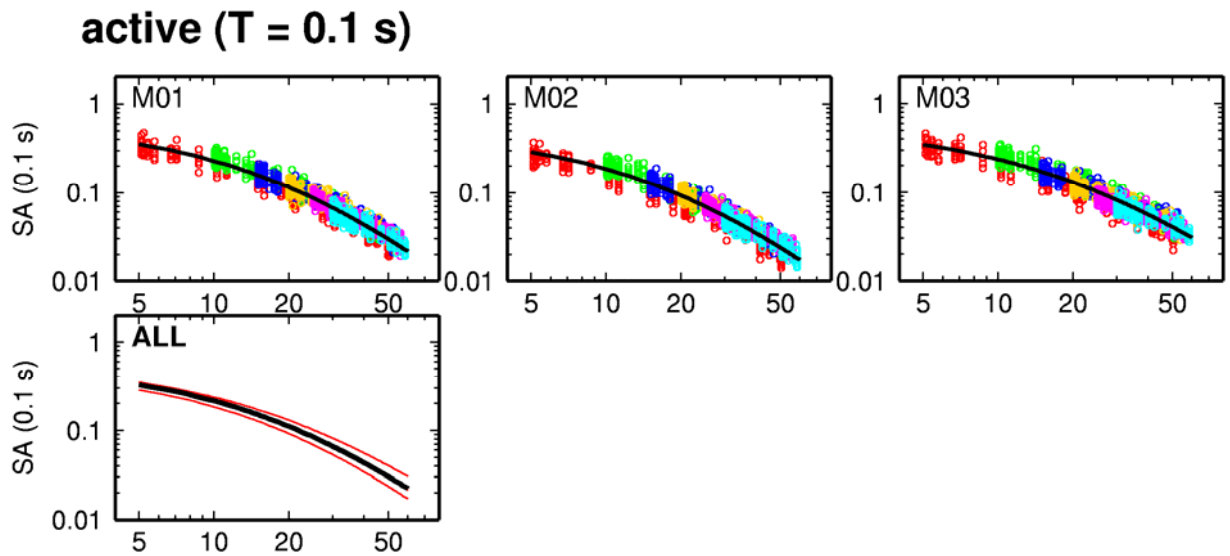

Figure 6-12
Simulations for the three WUS crustal structures.

Table 6-2
Comparison of expected path variability for the CEUS and CA based on crustal structures variability and differences in the source depth distributions

Region	Standard Deviation in Ln units						
	PGA	**25 Hz**	**10 Hz**	**5 Hz**	**2 Hz**	**1 Hz**	**0.5 Hz**
Active	0.20	0.21	0.23	0.23	0.21	0.33	0.33
Stable	0.16	0.17	0.17	0.21	0.20	0.30	0.28

Site-term variability

The NGA models use V_{s30} as a site paremeter. For the CEUS, the hazard is typically computed using attenuation relations for hard-rock site conditions. For example, the EPRI (2004) models are for a V_{S30}=2800 m/s. Since we are concerned with hard-rock site conditions, we need to check that the standard deviation from the NGA model is applicable to hard-rock conditions. The residuals from the NGA model are plotted as a function of the V_{S30} in Figures 6-13 and 6-14 for peak acceleration and 25 Hz spectral acceleration, respectively. There is no systematic increase in variability as a function of the V_{S30}. The intra-event standard deviation for all of the data is 0.49 and the intra-event standard deviation for V_{S30}>750m/s is 0.46. This indicates that there may be a small reduction in the site-term variability for the hard-rock sites.

At high frequencies (> 10 Hz), the intra-event variability is slightly larger for the hard-rock sites than for rock or soil sites. For example, at 25 Hz, the standard deviation for V_{S30}>950 m/s is 0.55 compared to the total data set value of 0.51. This additional variability results from the additional high frequency content that is in the spectrum. If there is no energy to resonate, then the response spectral values will be similar to the PGA and the variability will be equal to the PGA variability. As you move to lower frequencies, then there is energy to resonate and the variability of the response spectral values increases as compared to the PGA variability. This variability is approximately constant with frequency for frequencies with significant engery to resonate, however, the variability stars to increase at frequencies less than 1 Hz. We can see this effect in the WUS data set: in the WUS data, the 25 Hz spectral acceleration is very close to the PGA since most sites attenuate the high frequencies, but if we consider hard rock sites in the WUS data set (e.g. V_{S30}>950 m/s), the standard deviation is 0.55 which is similar to the 5-10Hz standard deviation of 0.55-0.56 for the full data set. For the CUES, the ground motion has larger high frequency content than the WUS. To accommodate the difference in the high frequency content between the CEUS and the WUS, the WUS standard deviations at high frequencies should be increased to the level of the standard deviation in the frequency range with significant energy. That is, we are assuming that there is adequate high frequency energy in the EUS data that will lead to the increase in variability seen for moderate frequency response spectral values in the WUS.

Atkinson (2005) compared the total standard deviation of recordings from a single site with the standard deviation from multiple sites. Atkinson found that the single station standard deviation was 10% smaller than the standard deviation for the total data set. The standard CEUS ground motion models are for a V_{S30}=2800 m/s. For this hard-rock condition, the lateral variability of the crustal structure in the CEUS may be less than in the WUS. The 10% reduction in the standard deviation found for a single site can be considered to be an upper bound of the potenital reduction for a more laterally homogeneous crust.

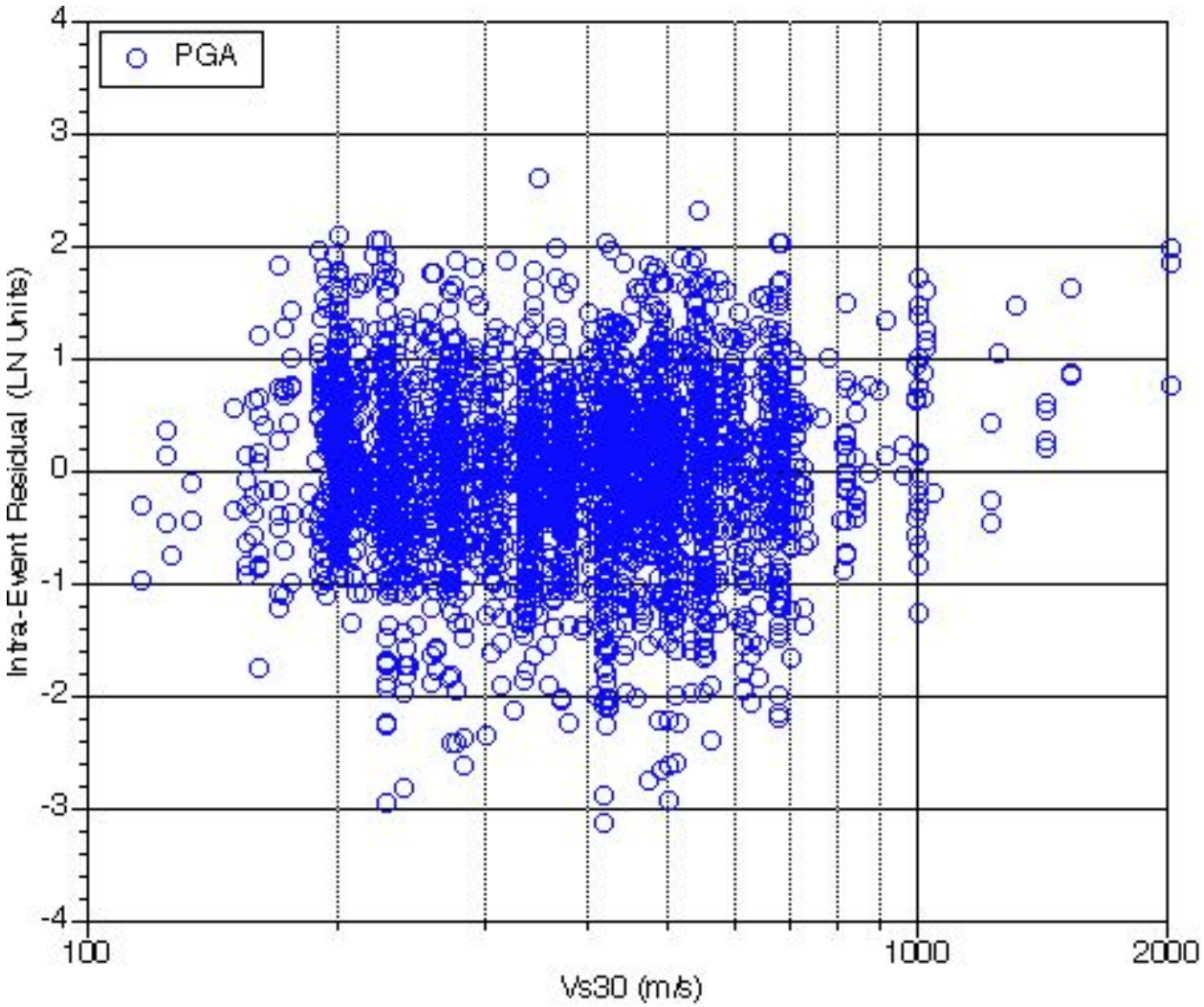

Figure 6-13
V_{S30} dependence of the PGA intra-event residuals.

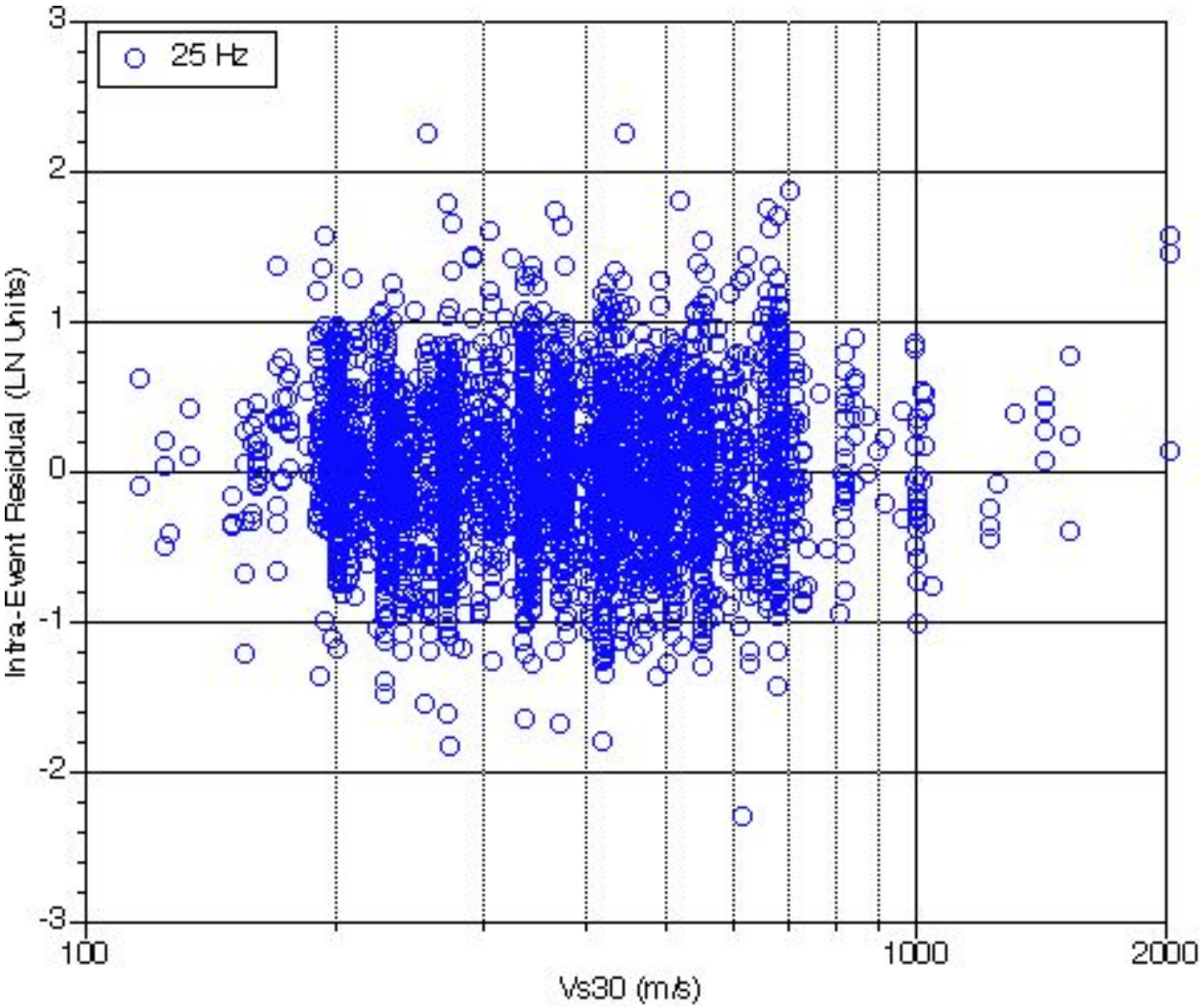

Figure 6-14
V_{S30} dependence of the 25 Hz intra-event residuals.

ENA Network Data

Based on teleseismic source effects discussed above, the inter-event ground motion variability in the CEUS is similar to that in the WUS. Regional network ground motion data can also be used to evaluate the inter-event variability for the CEUS.

Atkinson and Boore (2005) evaluated ENA network data and determined stress-parameters and high-frequency amplitude levels for ENA earthquakes with moment magnitudes >=4. The stress-parameters from the Atkinson and Boore (2005) study are shown in Figure 6-15. Assuming no magnitude dependence of the stress-parameter, the standard deviation of the stress-parameters from Atkinson and Boore is $\sigma_{\ln\Delta\sigma}$=0.60 natural log units.

The standard deviation of the high frequency response spectral values can be computed from the standard deviaton of the stress-parameter. By computing the high frequency response spectral values for a range of stress-parameters, keeping all other parameters constant, the scaling of the

high frequency ground motion with stress-parameter can be computed. Based on forward modeling of the stress-drop dependence, for spectral frequencies above the corner frequency of the earthquake, Abrahamson and Becker (1997) relate the response spectral value to the point source stress-parameter

$$\ln(Sa) = c_1 + 0.74 \ln(\Delta\sigma) \qquad \text{(Equation 6-2)}$$

where Sa is the spectral acceleration at 5% damping. The standard deviation of the ground motion due to stress-parameter variability is then given by 0.74 $\sigma_{\ln\Delta\sigma}$. Therefore, the stress-parameter standard deviation from Atkinson and Boore (2005) corresponds to an inter-event standard deviation of 0.44 natural log units (0.60 x 0.74). This value is similar to the inter-event standard deviation from the WUS data of about 0.4 (Table 5-1).

Atkinson and Boore (2005) also includes a model of the high frequency level from the ENA network data. The standard deviation of the high frequency level using a quadratic magnitude scaling is 0.39 natural log units. The high frequency level can be related to the stress-parameter using the point source spectrum. In the single corner frequency model, the source spectrum is given by

$$S(f) = \frac{c_1 M_o f^2}{1 + \left(\frac{f}{f_c}\right)^2} \qquad \text{(Equation 6-3)}$$

and the corner frequency is given by

$$f_c = c_2 \left(\frac{\Delta\sigma}{M_o}\right)^{1/3} \qquad \text{(Equation 6-4)}$$

At high frequencies, the source spectrum is approximately

$$S_{HF} = c_1 c_2^2 M_o^{1/3} \Delta\sigma^{2/3} \qquad \text{(Equation 6-5)}$$

The relation between the $\ln(\sigma_{HF})$ and the $\ln(\Delta\sigma)$ is given by

$$\ln(\Delta\sigma) = c_3 + 1.5 \ln(\sigma_{HF}) \qquad \text{(Equation 6-6)}$$

Combining this scaling with the scaling of spectral acceleration with stress-drop, the high frequency response spectral standard deviation is given by 1.11 σ_{HF}.(0.74 x 1.5) Therefore, this model leads to an inter-event standard deviation of 0.43 natural log units (1.11 x 0.39). This value is consistent with the standard deviation determined from the stress-parameters. This agreement is expected since the stress-parameters were derived from the hogh frequency levels.

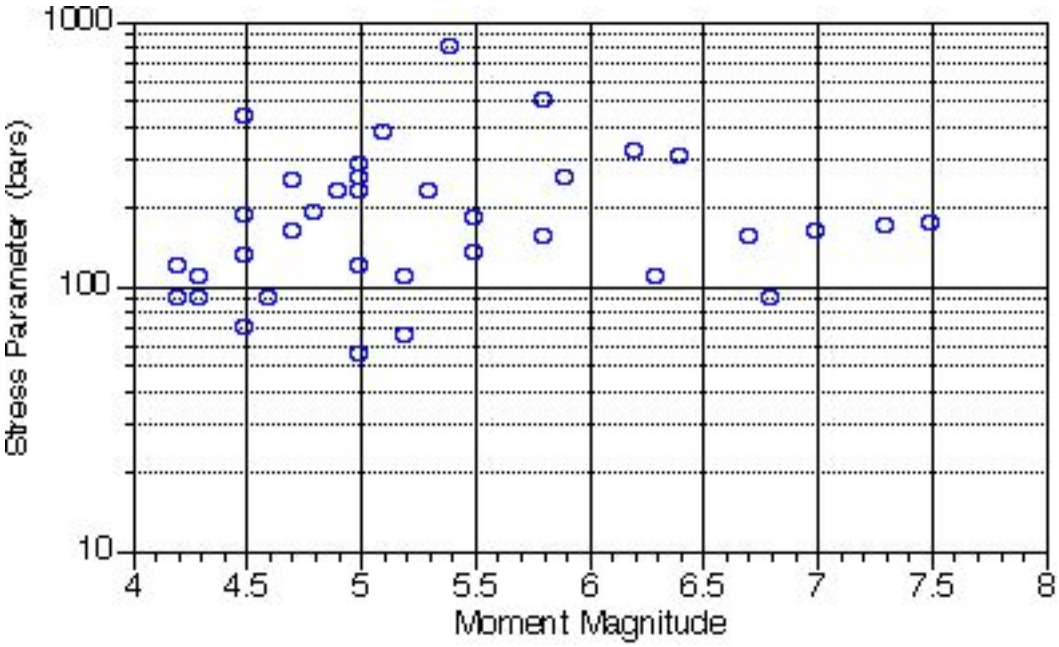

Figure 6-15
Stress parameters from ENA network data (from Atkinson and Boore, 2005)

Recommended CEUS sigma

Based on the observations that the inter-event variability in the CEUS is similar to the inter-event variability in the WUS we develop a suite of alternative aleatory variability models for the CEUS.

For the intra-event variability, the standard deviation from the WUS data base is considered to be applicable to the CEUS with the following modification: increase the standard deviation for high frequencies to account for the additional high frequency energy in the EUS spectra. A model with a reduced intra-event variability that could be caused by a more homogeneous crustal structure (at the V_S=2800 m/s depth) is also considered. In this model, the variability is 0.03 natural log units smaller, which is about 1/2 of the reduction Atkinson (2005) found for a single station. The reduction in the standard deviation for PGA at hard-rock sites determined from the WUS data set is similar (0.03 ln units). While a reduction for a single station is well founded, a reduction for the CEUS rock is more speculative. Therefore, a weight of 0.3 is given to the reduction alternative and a weight of 0.7 is given to the no reduction alternative.

For the inter-event variability, only one model is considered since the ENA data showed inter-event variability consistent with the WUS models. The high frequency ENA data gave slightly larger variability (0.43 as compared to 0.40). To account for this larger inter-event variability, the WUS inter-event standard deviations are increased by 0.03 units.

This leads to two models for the standard deviation at distances greater than 20 km. The recommended values for these two alternatives listed in Table 6-3.

Three alternative models are considered for the additional variability at short distances. In the first model, no additional variability is added. This is based on the lack of a trend in the

residuals with hypocentral depth for R_{JB}<10 km. In the second model, the variability is increased based on a slope of –0.2. (This is from the European data evaluated for R_{JB}<15 km excluding the one recording discussed earlier.) In the third model, the variability is increased based on a slope of –0.4. This value is selected as a value between the computed –0.2 and a value of –0.7 for 1/R scaling. There is no empirical evidence to support this large slope, but it is included to account for the uncertainty in the slopes. The no increase model is preferred since it is consistent with the empirical data at distances less than 10 km. A weight of 0.6 is given to this model. The small increase model is given a weight of 0.3 and the larger increase model is given a weight of 0.1.

The additional standard deviation is model by:

$$\sigma_2 = \begin{cases} a_1 & for R_{JB} \le 10km \\ a_1\left(1 - \dfrac{\ln(R_{JB}) - \ln(10)}{\ln(20) - \ln(10)}\right) & for 10km < R_{JB} < 20km \\ 0 & for R_{JB} \ge 20km \end{cases} \qquad \text{(Equation 6-7)}$$

The total standard deviation is computed by combining the additional standard deviation due to the use of the JB distance (σ_2) with the combined standard deviation from the intra-event and intra-event and inter-event terms for distances greater than 20 km (σ_1):

$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2} \qquad \text{(Equation 6-8)}$$

where

$$\sigma_{Total} = \sqrt{\sigma_{\text{int}\,er}^2 + \sigma_{\text{int}\,ra}^2} \qquad \text{(Equation 6-9)}$$

The standard deviations developed in this study do not include the correction for the effects of using point source distances rather than using the JB distance.

The standard deviation models recommended are based on empirical studies. The EPRI (2004) models also used numerical simulations to estimate the standard deviation. In these simulation studies, the variability was divided into modeling variability and parametric variability. The modeling variability was estimated from comparisons of simulated and recorded ground motions. The parametric variability was computed from the variability of simulated ground motions in which the event parameters that were optimized in the modeling variability exercise (such as stress-drop) were then randomized for future earthquakes. If these sources of variability are independent, then this approach should give a valid estimate of the variability. An issue is if the variability using this modeling approach is too large because of some correlations between the modeling variability and the parametric variability.

Table 6-3
Recommended Standard Deviation, σ_1, for the CEUS. Values are in Ln units.

	Model 1A WUS Inter-event WUS Intra-event wt = 0.7			Model 1B WUS Inter-event WUS Intra-event Reduced for homogeneous crustal structure wt = 0.3		
Frequency (Hz)	Intra-event	Inter-Event	Total	Intra-event	Inter-event	Total
PGA	0.51	0.37	0.63	0.48	0.37	0.61
25	0.56	0.43	0.71	0.53	0.43	0.68
10	0.56	0.43	0.71	0.53	0.43	0.68
5	0.56	0.43	0.71	0.53	0.43	0.68
2	0.56	0.43	0.71	0.53	0.43	0.68
1	0.60	0.43	0.74	0.57	0.43	0.71
0.5	0.62	0.43	0.75	0.59	0.43	0.73

Table 6-4
Recommended value of coefficient a1 (in Equation 6-7) for the CEUS.

	Model 2A Weight=0.6	Model 2B Weight=0.3	Model 2C Weight=0.1
PGA	0.00	0.12	0.23
25	0.00	0.12	0.23
10	0.00	0.12	0.23
5	0.00	0.12	0.23
2	0.00	0.12	0.23
1	0.00	0.12	0.23
0.5	0.00	0.12	0.23

7
CONCLUSIONS

There are two key issues for the ground motion variability: should the ground motion distribution be truncated at a maximum number of standard deviations, and what value of the standard deviation that should be used for the CEUS.

Although it has been common practice in PSHA in the WUS to truncate the ground motion distribution at a maximum epsilon of about 3, we found no sound technical basis for a truncation at this epsilon level. Residuals from empirical attenuation relations based on large data sets have epsilon values ranging up to 3 or 4. The ground motions with largest epsilon values were evaluated in detail and no systematic physical feature associated with large epsilon values was found that could be used to exclude these large epsilon values from the empirical data. Since large epsilon values are also observed in ground motions from kinematic numerical simulations that use simplified wave propagation (e.g. 1-D crustal models), we have no evidence that epsilon values greater than 3 are not possible. Their rate of occurrence depends on the joint probability distribution of the parameters used in the numerical simulations. Limits on epsilon from numerical simulations may be appropriate if there are correlations of the source parameters that exclude certain combinations that lead to large epsilon values. Currently, these joint distributions are too poorly constrained to permit a reasonably supported conclusion. It is important to note that epsilon is a relative measure of the variability since it is a normalized number of standard deviations. If the value of the standard deviation is reduced, then the value of epsilon will increase.

While at some level, the ground motion will reach the physical limit of the rock or soil, this type of truncation is based on the ground motion level, not on the epsilon value directly. Separately, a study of the maximum ground motion that is possible near the source could be defined. Given the lack of a technical basis for defining a maximum epsilon value, we recommend that no truncation of the ground motion distribution be included based on a maximum value of epsilon for probabilistic seismic hazard analysis. This conclusion supports the approach used in the recent seismic hazard analyses conducted by EPRI (2004) that did not include a truncation of the ground motion distribution.

The value of the standard deviation used in the EPRI (2004) ground motion study is larger than recent studies of large data sets of ground motions applicable to the WUS. Possible causes for an increase in the standard deviation from the WUS to the CEUS were evaluated in terms of the source, path, and site contributions to the variability. These comparisons indicate that the intra-event standard deviation from the WUS is applicable to the CEUS and the inter-event variability may be larger in the CEUS based on the variability of stress-parameter estimates. To capture the epistemic uncertainty in the standard deviation, two alternative models of the intra-event standard deviations are developed. The total standard deviation is given by combining the inter-event and intra-event standard deviations. These two models with their recommended weights are given in Table 6-3. These models are for the standard deviation based on JB distance

models. The additional variability considered by EPRI (2004) to account for the effect of using point sources rather than extended sources is not included in this study.

The standard deviations discussed above represent the standard deviation at JB distances greater than 20 km. For distances less than 20km, there is a key issue regarding the need for an increase in the standard deviation due to the use of the JB distance metric which is independent of depth. The EPRI (2004) ground motion models were based on the JB distance metric and a significant increase in the standard deviation for JB distances less than 20 km was included in all of the sigma models developed in that study. The empirical ground motion data evaluated in this study do not support a large increase in the standard deviation at short distances, but some increase may be justified. Three alternative models of the additional contribution to the standard deviation at short distances are developed. The parameters for these three models and the recommended weights are given in Table 6-4. Note that most of the weight is given to the model with zero increase.

Combining the two models of the standard deviation (Table 6-3) with the three models of the increase in the standard deviation at distances less than 20 km (Table 6-4) leads to a total of 6 standard deviation models. The weights for the two sets of models are multiplied to give the total weight for each of the 6 models.

The resulting models of the standard deviation are, on average, lower than the standard deviation models developed in the EPRI (2004) study. There are three key reasons for this decrease. First, the inter-event variability of the ENA stress-parameters given by Atkinson and Boore (2005) is less than used by EPRI (2004). Second, the increase a short distances included in the EPRI models to account for the use of the JB distance (assuming that the ground motion attenuates with 1/R and that there is no depth dependence of stress-drop) has been reduced based on evaluations of residuals from recent ground motion studies. Finally, the intra-event variability is modeled based on empirical data from the WUS rather than using modeling methods.

Since the computed hazard increases with an increase in the standard deviation, the hazard results computed using the EPRI (2004) standard deviations will be conservative compared to hazard computed using the standard deviations recommended in this report if all other aspects of the source and ground motion models held fixed.

8
REFERENCES

Abrahamson, N.A. (2000). State of the practice of seismic hazard evaluation. *Proceedings of GeoEng 2000*, Melbourne, 19-24 November, vol. **1**, 659-685.

Abrahamson, N.A. & R.R. Darragh (1985). The Morgan Hill earthquake of April 24, 1984 – The 1.29g acceleration at Coyote Lake Dam: Due to directivity, a double event, or both? *Earthquake Spectra* **1**(3), 445-455.

Abrahamson, N. A, P. G. Somerville, and C. A. Cornell (1990). Uncertainty in numerical strong motion predictions, Proc. 4th US Nat. Conf. Earth. Engin., 1, 407-416.

Abrahamson, N.A. & R.R. Youngs (1992). A stable algorithm for regression analysis using the random effects model. *Bulletin of the Seismological Society of America* **88**(1), 505-510.

Abrahamson, N. A. and A. Becker (1997). Ground motion data package, vol 1, Yucca Mountain PSHA project, February 1997.

Abrahamson, N.A., Birkhauser P., Koller, M., Mayer-Rosa, D. Smit, P., Sprecher, C., Tinic, S. and Graf, R. (2002). PEGASOS – a comprehensive probabilistic seismic hazard assessment for nuclear power plants in Switzerland. *12th European Conference on Earthquake Engineering*, London, Paper No. 633.

Abrahamson, N. A. and W. J. Silva (2005). Preliminary results of the A&S 2005 attenuation relation, presented at the USGS workshop on Attenuation Relations to be used in the USGS National Seismic Hazard Maps, Menlo Park, CA, October 24, 2005.

Ambraseys, N.N., K.A. Simpson & J.J. Bommer (1996). Prediction of horizontal response spectra in Europe. *Earthquake Engineering & Structural Dynamics* **25**(4), 371-400.

Ambraseys, N.N., J. Douglas, S.K. Sarma & 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. *Bulletin of Earthquake Engineering* **3**(1), 1-53.

Anderson, J.G. & J.N. Brune (1999). Probabilistic seismic hazard assessment without the ergodic assumption. Seismological Research Letters 70(1), 19-28.

Atkinson, G.M. & D.M. Boore (1992). Source spectra for the 1988 Saguenay, Quebic earthquakes, Bull. Seism Soc. Am., **82**, 683-719.

Atkinson, G. M. (1993). Earthquake source spectra in Eastern North America, Bull. Seism. Soc. Am., 83, 1778-1798.

Atkinson, G.M. & D.M. Boore (1997a). Some comparisons between recent ground-motion relations. *Seismological Research Letters* **68**(1), 41-57.

Atkinson, G.M. & D.M. Boore (1997b). Stochastic point-source modeling of ground motions in the Cascadia region. *Seismological Research Letters* **68**(1), 74-85.

Atkinson, G.M. & D.M. Boore (2003). Empirical ground-motion relations for subduction-zone earthquakes and their application to Cascadia and other regions. *Bulletin of the Seismological Society of America* **93**(4), 1703-1729.

Atkinson, G. M (2005). Single station sigma, Bull. Seism. Soc. Am., in Press.

Atkinson, G. M. and D. M. Boore (2005). Earthquake ground-motion prediction equations for eastern North America, submitted to Bull. Seism. Soc. Am., December 4, 2005.

Berge-Thierry, C., F. Cotton, O. Scotti, D.-A. Griot-Pommera & Y. Fukushima (2003). New empirical response spectral attenuation laws for moderate European earthquakes. *Journal of Earthquake Engineering* **7**(2), 193-222.

Bernreuter, D.L., J.B. Savy, R.W. Mensing & J.C. Chen (1989). Seismic hazard characterization of 69 nuclear power plant sites East of the Rocky Mountains. *Report NUREG/CR-5250*, US Nuclear Regulatory Commission.

Bommer, J.J., Abrahamson, N.A., Strasser, F.O., Pecker, A., Bard, P-Y., Bungum, H., Cotton, F., Fäh, D., Sabetta, F., Scherbaum, F. and Studer, J. (2004). The challenge of defining upper bounds on earthquake ground motions. *Seismological Research Letters* **75**(1), 82-95.

Bommer, J.J., J. Douglas & F.O. Strasser (2003). Style-of-faulting in ground-motion prediction equations. *Bulletin of Earthquake Engineering* **1**(2), 171-203.

Bommer, J.J. & D.M. Boore (2004). Engineering Seismology. In*: Encyclopedia of Geology*, Academic Press, vol. 1, pp.499-514.

Bommer, J.J. & F. Scherbaum (2005). Capturing and limiting ground-motion uncertainty in seismic hazard assessment. In: *Future Directions in Strong Motion Instrumentation*, Gulkan, P. & J.G. Anderson (Eds.), Springer, 315 pp.

Boore, D.M. (2001). Comparisons of ground motions from the 1999 Chi-Chi earthquake with empirical predictions largely based on data from California. *Bulletin of the Seismological Society of America*, **91**(5), 1212-1217.

Boore, D.M., V.M. Graizer, J.C. Tinsley & A.F. Shakal (2004). A study of possible ground-motion amplification at the Coyote Lake Dam, California. *Bulletin of the Seismological Society of America*, **94**(4), 1327-1342.

Borcherdt, R.D. & G. Glassmoyer (1994). Influences of local geology on strong and weak ground motions recorded in the San Francisco Bay region and their implications for site-specific building-code provisions, in Borcherdt, R.D. (ed.), The Loma Prieta, California, Earthquake of October 17, 1989 – Strong Ground Motion. *United States Geological Survey Professional Paper 1551-A*, 77-108.

Bouchon, M. & J.S. Barker (1996). Seismic response of a hill: The example of Tarzana, California, *Bulletin of the Seismological Society of America*, **86**(1), 66-72.

Campbell, K.W. (1988). The Whittier Narrows, California earthquake of October 1, 1987 – Preliminary analysis of peak horizontal acceleration. *Earthquake Spectra* **4**(1), 115-137.

Campbell, K.W. (1997). Empirical near-source attenuation of horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra. *Seismological Research Letters* **68**(1), 154-179.

Carlisle, H. & K.M. Rollins (1994). Ground-response studies at the Alameda Naval Air Station, in Borcherdt, R.D. (ed.), The Loma Prieta, California, Earthquake of October 17, 1989 – Strong Ground Motion. *United States Geological Survey Professional Paper 1551-A*, 123-143.

Catchings, R.D. & W.H.K. Lee (1996). Shallow velocity structure and Poisson's ratio at the Tarzana, California, strong-motion accelerometer site. *Bulletin of the Seismological Society of America*, **86**(6), 1704-1713.

Chang, T.-Y., F. Cotton & J. Angelier (2001). Seismic attenuation and peak ground acceleration in Taiwan. *Bulletin of the Seismological Society of America* **91**(5), 1229-1246.

Chen, K.-C., B.-S. Huang, J.-H. Wang & H.-Y. Yen (2002). Conjugate thrust faulting associated with the 1999 Chi-Chi, Taiwan, earthquake sequence. *Geophysical Research Letters* **29**(8), Art. No. 1277.

Chi, W.-C. & D. Dreger (2004). Crustal deformation in Taiwan: Results from finite source inversions of six $M_w > 5.8$ Chi-Chi aftershocks. *Journal of Geophysical Research* **109**(7), Art. No. B07305.

Donovan, N.C. & A.E. Bornstein (1978). Uncertainties in seismic risk analysis. *Journal of the Geotechnical Engineering Division*, ASCE, **104**, 869-887.

Douglas, J. (2003). Earthquake ground motion estimation using strong-motion records: a review of equations for the estimation of peak ground acceleration and response spectral ordinates. *Earth-Science Reviews* **61**, 43-104.

EERI (2000). The Napa earthquake of September 3, 2000. *EERI Special Earthquake Report*, Earthquake Engineering Research Institute, November 2000.

EPRI, Dominion Energy, Entergy Nuclear, and Exelon Generation Company (2004). Report CEUS Ground Motion Project Final Report, 1009684.

http://www.eeri.org/lfe/pdf/usa_napa_eeri_preliminary_report.pdf

Field, E.H., S.E. Hough, K.H. Jacob & P.A. Friberg (1994). Site response in Oakland, California, near the failed section of the Nimitz Freeway, in Borcherdt, R.D. (ed.), The Loma Prieta, California, Earthquake of October 17, 1989 – Strong Ground Motion. *United States Geological Survey Professional Paper 1551-A*, 169-180.

Hauksson, E. and P. Shearer (2005). Southern California hypocenter relocations with waveform cross-correlation, Part 1: Results using double-difference methods, Bull. Seism. Soc. Am., **95** (3), 896-903.

Herrero, A. & P. Bernard (1994). A kinematic self-similar rupture process for earthquakes. *Bulletin of the Seismological Society of America* **84**(4), 1216-1228.

Herrmann, R.B. & C.Y. Wang (1985). A comparison of synthetic seismograms. *Bulletin of the Seismological Society of America* **75**(1), 41-56.

Herrmann, R.B. (1996a). Computers program in seismology. An overview of synthetic seismogram computation. Version 3.0. Department of Earth and Atmospheric Sciences, Saint Louis University, Saint Louis, Missouri.

Herrmann, R. B. (1996b). Computers program in seismology. Volume VI: wavenumber integration. Version 3.0. Department of Earth and Atmospheric Sciences, Saint Louis University, Saint Louis, Missouri.

Hikima, K. & K. Koketsu (2005). Rupture processes of the 2004 Chuetsu (mid-Niigata prefecture) earthquake, Japan: a series of events in a complex fault system. *Geophysical Research Letters* **32**(18), Art. No. L18303.

Kao, H., Y.-H. Liu, W.-T. Liang & W.-P. Chen (2002). Source parameters of regional earthquakes in Taiwan:1999-2000 including the Chi-Chi earthquake sequence. *TAO* **13**(3), 279-298.

Lee, W.H.K., T.C. Shin, K.W. Kuo, K.C. Chen & C.F. Wu (2001). CWB free-field strong-motion data from the 21 September, Chi-Chi, Taiwan, earthquake. *Bulletin of the Seismological Society of America* **91**(5), 1370-1376.

Lussou, P., P.Y. Bard & F. Cotton (2001). Seismic design regulation codes: Contribution of K-Net data to site effect evaluation. *Journal of Earthquake Engineering* **5**(1), 13-33.

Mai, P.M., P. Spudich & J. Boatwright (2005). Hypocenter locations in finite source rupture models. *Bulletin of the Seismological Society of America* **95**(3), 965-980.

Martinez-Pereira, A. & J.J. Bommer (1998). What is the near-field? In *Seismic Practice into the Next Century*, ed. Booth, E., Balkema, 245-252.

Midorikawa, S. & Y. Ohtake (2004). Variance of peak ground acceleration and velocity in attenuation relationships. *Proceedings of the 13th World Conference on Earthquake Engineering*, Vancouver, Paper no. 325.

Miranda, E. & H. Aslani (2000). Brief Report on the September 3, 2000 Yountville/Napa California Earthquake. *Report prepared for the Pacific Earthquake Engineering Center.* http://nisee.berkeley.edu/yountville/

Molas, G.L. & F. Yamazaki (1995). Attenuation of earthquake ground motion in Japa including deep-focus events. *Bulletin of the Seismological Society of America* **85**(5), 1343-1358.

PEER (2005). NGA Strong Motion Database. http://peer.Berkeley.edu/NGA.

Peppin, W.A. (1987). Exotic seismic phases recorded near Mammoth Lakes and their use in the delineation of shallow-crustal (magma) anomalies. *Pure and Applied Geophysics* **125**(6), 1008-1023.

Pousse, G., C. Berge-Thierry, L.F. Bonilla & P.-Y. Bard (2005). Eurocode 8 design response spectra evaluation using the K-NET Japanese database. *Journal of Earthquake Engineering* **9**(4), 547-574.

Rial, J. A. (1996). The anomalous seismic response of the ground at the Tarzana Hill site during the Northridge 1994 southern California earthquake: a resonant, sliding block?, *Bulletin of the Seismological Society of America*, **86**(6), 1714-1723.

Romeo, R. & A. Prestininzi (2000). Probabilistic versus deterministic hazard analysis: an integrated approach for siting problems. *Soil Dynamics & Earthquake Engineering* **20**(1-4), 75-84.

Sabetta, F. & A. Pugliese (1996). Estimation of response spectra and simulation of nonstationary earthquake ground motions. *Bulletin of the Seismological Society of America* **86**(2), 337-352.

Shearer, P. (2005). Gutenberg Lecture: Seismic Data Crunching for Fun and Profit, Eos Trans, AUG, 86(52), Fall Meeting Suppl., Abstract S14A-01.

Somerville, P.G., N.F. Smith & R.G. Graves (1994). Effect of critical reflections from the Moho on the attenuation of strong ground motion, in Borcherdt, R.D. (ed.), The Loma Prieta, California, Earthquake of October 17, 1989 – Strong Ground Motion. *United States Geological Survey Professional Paper 1551-A*, 67-76.

Somerville, P. & J. Yoshimura (1990). The influence of critical Moho reflections on strong ground motions recorded in San Francisco and Oakland during the 1989 Loma Prieta earthquake. *Geophysical Research Letters* **17**(8), 1203-206.

Spudich, P., M. Hellweg & W.H.K. Lee (1996). Directional topographic site response at Tarzana observed in aftershocks of the 1994 Northridge, California, earthquake: Implications for mainshock motions, *Bulletin of the Seismological Society of America*, **86**(1B), S193-S208.

Strasser, F.O. & J.J. Bommer (2004). Outliers in strong-motion datasets *Proceedings of Workshop on Extreme Ground Motions at Yucca Mountain*, Menlo Park, California, August 2004.

Strasser, F.O. & J.J. Bommer (2005a). Large ground-motion residuals in the NGA/PEER database. *Research Report*, Department of Civil and Environmental Engineering, Imperial College London.

Strasser, F.O. & J.J. Bommer (2005b). Preliminary simulations with EXWIM: Calibration and assessment of ground-motion variability in the near-source region. *Research Report*, Department of Civil and Environmental Engineering, Imperial College London.

Strasser, F.O. & J.J. Bommer (2005c). Analysis of intra-event ground-motion residuals from K-NET and KiK-NET data. *Research Report*, Department of Civil and Environmental Engineering, Imperial College London.

Toro, G.R., N.A. Abrahamson, and J.F. Schneider (1997). *Model of Strong Ground Motions from Earthquakes in Central and Eastern North America: Best Estimates and Uncertainties*, Seism. Res. Let., 68, 58-73.

USGS (2000). The September 3, 2000 Yountville Earthquake, U.S. Geological Survey, Earthquake Hazards Program online Report (last modified 7/12/2000). http://quake.wr.usgs.gov /recent/reports/napa/index.html

Vahdani, S. & S. Wikstrom (2002). Response of the Tarzana strong motion site during the 1994 Northridge earthquake. *Soil Dynamics and Earthquake Engineering* **22**, 837-848.

Vidale, J.E. & O. Bonamassa (1994). Influence of near-surface geology on the direction of ground motion above a frequency of 1 Hz, in Borcherdt, R.D. (ed.), The Loma Prieta, California, Earthquake of October 17, 1989 – Strong Ground Motion. *United States Geological Survey Professional Paper 1551-A*, 61-66.

Wang, G.-Q., D.M. Boore, H. Igel & X.-Y. Zhou (2004). Comparisons of ground motions from five aftershocks of the 1999 Chi-Chi, Taiwan, earthquake with empirical predictions largely based on data from California. *Bulletin of the Seismological Society of America*, **94**(6), 2198-2212.

Youngs, R.R., N. Abrahamson, F.I. Makdisi, & K. Singh (1995). Magnitude-dependent variance of peak ground acceleration. *Bulletin of the Seismological Society of America* **85**(4), 1161-1176.

A
COMMENTS AND RESPONSES TO NRC REQUESTS FOR ADDITIONAL INFORMATION (RAIs)

The NRC submitted 13 RAIs related to the G1.3 Task on June 1, 2006. The industry responses and resolutions for 12 of the 13 RAIs are documented below.

NRC Comment/Question:

3.1 Provide hazard curves for the twenty-eight CEUS reactor sties with the revised sigma model so that NRC staff can assess the overall impact.

Industry Response/Resolution:

3.1 This will be addressed in a separate report.

NRC Comment/Question:

3.2 The report indicates that there are some epsilon >3 records in San Francisco, CA from the Loma Prieta earthquake in 1989. However, using the latst NGA attenuation relationships, Boore and Atkinson (2995), Abrahamson & Silva (20050, and Campbell and Bozorgnia (2005), no instances of epsilon > 3 were found, see table attached. Please explain the difference

Industry Response/Resolution:

3.2 The epsilon value depends on both the median and the standard deviation (see page 2-7, sigmas and epsilons). If the standard deviation increases, as occurred in the NGA models, then the epsilon decreases. If the median value changes, then the epsilon also changes. This is why the epsilon is not a robust parameter. The change in the epsilon values for the data from the Loma Prieta earthquake is not a surprise. Similar changes can be expected each time the ground motion models are revised.

NRC Comment/Question:

3.3 The report postulates that there are three causes for changes in NGA's standard deviation model: a) a larger data set, b) corrections of meta-data, and c) use of the geometric mean

of the horizontal components after finding the optimal rotational angle. Two of the three possible causes, b) and c) are eliminated after further comparison with the earlier data used in the Abrahamson and Silva (1997) model, while added data from the new earthquakes is determined to be the cause. Did the other NGA modeling teams all reach the same result? Please provide further explanation regarding the other modeling teams and possible causes for changes in their standard deviations.

Industry Response/Resolution:

3.3 N. Abrahamson carried out this evaluation as part of the NGA project and made these results available to the other models. We are not aware of similar evaluations of the cause for the changes in the sigma being made by the other NGA teams.

NRC Comment/Question:

3.4 The report indicates that the meta-data change has no significant impact on the sigma. However, the 1997 data seem to differ significantly from the 2005 data in the inter-event part, when comparing Figures 5-4 and 5-5. Please explain this difference (and also at regional and teleseismic distances).

Industry Response/Resolution:

3.4 As noted, the value of the inter-event residuals for the 1997 and 2005 data sets are significantly different. This change occurs because the median model is different or the two models. The models have different distance scaling, magnitude scaling, style-of-faulting and site effects. The change in the median model is not due to the change in the meta-data, rather it is due to the increased data set size and the change in the functional form of the model.

NRC Comment/Question:

3.5 Please explain why variation in half-duration time derived from Harvard Moment Tensor catalog is a reliable indicator of source variation for different regions. Two strong aftershocks, which occurred on 9/22/1999 00:14 UT and 9/20/1999 21:46 UT, respectively, from Chi-Chi earthquake have the same moment magnitude (Mw=6.4) and the same half-durations (4.0 seconds) in the Havard Catalog, but two events produced substantially different PGA and SA (5Hz) for stations within 70 km distance.

Industry Response/Resolution:

3.5 The objective of the catalog comparison was to compare the variability of source properties for earthquakes from different tectonic regions. The half-duration provides a overall measure of the dynamic stress-drop. Teleseismic data are for moderate to long periods. We agree that the half duration may not be a good measure of the high-

frequencies generated by the earthquake, but it is the best that is available from teleseismic catalogs. We are using teleseismic catalogs to supplement the sparse ground motion data from ENA earthquakes above magnitude 5.

The approach used is to compare the variability of available source properties from teleseismic data for different tectonic regions. The half-duration is an available parameter. If there were other source properties available in the teleseismic data set that were more representative of the high frequency radiation, we would use them.

NRC Comment/Question:

3.6 The author stated that the path effects should give the same variability in the CEUD and WUS because synthetic seismograms using various crustal models for the CEUS show the same variability as seismograms for the WUS. Please explain how much plane-layered models can capture the variability of path effects, where there is likely to be lateral variations of crustal structure.

Industry Response/Resolution:

3.6 We used ranges of plane layered models applicable to the CEUS to compare to the range of plane layered models applicable to the WUS. This does not imply that the lateral variations in the crustal structure are not important. Rather, we have assumed that the lateral variations in the crustal structure in the CEUS and WUS are similar over distances of less than 100 km. The modeling is used to determine if we should expect larger variability in the CUES as compared to the WUS due to the range of 1-D crustal structures that are applicable to the CUES as compared to the range of 1-D crustal structure applicable to the WUS.

NRC Comment/Question:

3.7 Please explain the USNRC 2002 ground motion model referred on page 6-1 in section 6.

Industry Response/Resolution:

3.7 This it the USNRC (2002) report referenced in the EPRI (2004) ground motion report. It is "Guidence for Performing Probabilistic Seismic Hazard Analysis for a Nuclear Plant Site: Example Application to the Southeastern United States", NUREG/CR-6607

NRC Comment/Question:

3.8 The report indicates that the inclusion of the Chi-Chi mainshock causes the increase in the intra-event standard deviation at large magnitudes and the inclusion of additional earthquakes in the magnitude range of 5-6 have a similar variability of the event terms resulting in a reduction of the inter-event standard deviation of magnitudes less than 6.0.

Since the WUS ground motion functions mean to be used to predict the ground motions in California and other west US earthquakes, is it appropriate to introduce data from other regions which will obviously bias the ground motion prediction in this specific geographic area? Are the sigmas developed using additional data from other regions still valid for the area?

Industry Response/Resolution:

3.8 The ground motion data from the WUS is not used to predict the median ground motion from a specific magnitude and distance. We are only using the variability of the ground motion from the WUS to predict the variability of the ground motion in the EUS. While use of the WUS data to predict the median for EUS earthquakes would obviously introduce a bias, it is not at all obvious that the use of the WUS variability will introduce a bias. The estimates of the variability appear to be similar around the world for similar levels of model parameterization. One of the main objectives of the report was to check for reasons to expect that the variability in the EUS is significantly different from the WUS.

NRC Comment/Question:

3.9 On page 6-18, the report states that the standard deviation of the stress-drop from the paper submitted to BSSA by Atkinson and Boore is 0.6 natural log units (median value 140 bar). However, this quote of standard deviation was not found in Atkinson and Boore's 2005 submittal. Please explain.

Industry Response/Resolution:

3.9 The value of the standard deviation was computed using the stress-drops listed in Table 4 of the Atkinson and Boore (2006) paper, assuming a log-normal distribution.

NRC Comment/Question:

3.10 Atkinson and Boore provided a direct estimate of total sigma (aleatory with inter and intra event terms) contained in their 2004 report to USGS. In that report the sigma was calculated from the residuals of the observations with respect to the predictions of their 2004 attenuation relationships. Their 2005 submittal to BSSA used the same dataset as their 2004 ground motion functions which appear to be the same as the 2005 ground motion function. The sigma reported in 2004 is 0.3 at 0.2-0.5 Hz and 0.4 at 5-10 Hz. It gives a direct estimate of aleatory sigma of 0.92 ln unit at 5 Hz. This sigma is significantly higher than 0.7 unit recommended by NGA. Please explain.

Industry Response/Resolution:

3.10 The large standard deviation given in the Atkinson and Boore 2004 report for 5-10Hz was computed using both horizontal components and vertical components converted to

horizontal components. As a result, the standard deviation is biased to larger values since it includes the additional variability due to conversions from the vertical to the horizontal, particularly at high frequencies. In addition, much of the network data is for distances greater than 200 km which can have larger variabilities. Therefore, the large value of the standard deviation of 0.92 is not considered to be applicable to horizontal ground motion at distances less than 200 km.

NRC Comment/Question:

3.11 Please justify using the sigma derived from stress drop as the inter-event sigma when using attenuation relationships derived with a different median stress drop from the one found by Atkinson and Boore (2005).

Industry Response/Resolution:

3.11 In the stochastic point source model, the source terms are the moment and the stress-drop. The magnitude is defined by the moment, so the only source term remaining is the stress-drop. The variability in the stress-drop for a given moment will lead to different values of the ground motions for all distances. Therefore, the stress-drop is like an event term. If an attenuation relation was derived with event-terms based on point source stochastic models, then the event terms would be simply function of the stress-drop. The scaling of the ground motion (event-term) with stress drop was addressed in the report (see equation 6-2). The variability of the ground motion due to stress-drop variability does not depend on the median stress-drop, only on the variability of the stress-drop. Therefore, it is applicable to ground motion models that used different median stress-drops.

NRC Comment/Question:

3.12 The report recommends a inter-event sigma of 0.43, which is substantially less than the inter-event sigma value of 0.53 derived from the observed standard deviation of stress-drop reported by Atkinson and Boore (2005). Please explain this discrepancy.

Industry Response/Resolution:

3.12 The value of 0.43 is consistent with the data in Atkinson and Boore (2006). The stress-drop variability is 0.6 natural log units (see question 3.9). The ground motion variability is computed by multiplying the stress-drop variability by a factor of 0.74 (see page 6-19). This results in an inter-event variability of 0.44 (e.g. 0.74 x 0.60=0.44). The standard deviation of the high frequency level for the ENA network data is 0.39, which corresponds to a ground motion sigma of 0.43 (see page 6-19, 0.39 x 1.5 x 0.74 = 0.43). So, the standard deviation of the ground motion sigma based the high frequency level is consistent with the ground motion sigma based on the stress drop variability. These are both consistent with the inter-event standard deviation of 0.40 from the WUS data.

We do not understand how the value of 0.53, cited in the RAI, was derived from the stress-drops reported in Atkinson & Boore (2005).

NRC Comment/Question:

3.13 The 2004 EPRI ground motion report incorporates a distance dependence to the standard deviation to account for the use of the Joyner-Boore distance, RJB. However, PEER NGA studies indicate that the increase in the standard deviation at short distances may not be appropriate. The report states that this may be due to the deeper earthquakes being more energetic than the shallower earthquakes and postulates that the deeper earthquakes have higher stress-drops on average than shallower earthquakes, "As a result, the expected distance-dependence of the standard deviation is counter acted by the depth dependence of stress-drop". However, one of the four EPRI (2004) models for aleatory variability assumes a larger stress-drop variability for the CEUS compared to the WUS. The report also mention on page 5-7 that "other studies have not found stress-drop scaling with depth".

Considering 1) the lack of distance-dependence of the stress-drop, 2) the large variability of stress-drop for the CEUS, 3) the sparseness of the CEUS data, and 4) the statement on page 5-5 of the report:

"If it is assumed that the ground motion will attenuate as 1/R where R is the point source distance and that there is no systematic differences in the sources as a function of depth (e.g. not stress-drop dependence with depth), then there should be an increase in the standard deviation at short distances due to the use of RJB"

Please justify the decision to eliminate the additional sigma to account for the use of RJB for events at distances less than 20 km.

Industry Response/Resolution:

3.13 The basis for the decision to eliminate the additional sigma to account for the use of RJB for events at distances less than 20 km is given in the report. (see pages 5-5 to 5-11) This basis is summarized below, but the report gives a more complete discussion.

There is no empirical evidence of an increase in the standard deviation at short distances in the WUS or European data due to the use of RJB. Therefore, we consider it reasonable to assume that there will not be an increase for sites in the CEUS.

This is, if the ground motion attenuated as 1/R and the stress-drop was independent of depth for WUS or European data, then we should see an increase in the standard deviation at short distances. Since we don't see this, we conclude that these conditions do not hold in the WUS or Europe. We have no basis for assuming that these conditions would hold in the CEUS and but not hold in the WUS or Europe.

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI)

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, California, and Charlotte, North Carolina, was established in 1973 as an independent, nonprofit center for public interest energy and environmental research. EPRI brings together members, participants, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of electricity generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents nearly 15% of EPRI's total research, development, and demonstration program.

Together…Shaping the Future of Electricity

© 2006 Electric Power Research Institute (EPRI), Inc. All rights reserved. Electric Power Research Institute and EPRI are registered service marks of the Electric Power Research Institute, Inc.

Printed on recycled paper in the United States of America

1013105

ELECTRIC POWER RESEARCH INSTITUTE
3420 Hillview Avenue, Palo Alto, California 94304-1395 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com