

Alma Mater Studiorum – Università di Bologna

**DOTTORATO DI RICERCA  
IN  
MECCANICA DELLE STRUTTURE**

Ciclo XXI

**Settore scientifico disciplinare di afferenza: ICAR/09 – Tecnica delle Costruzioni**

**ASSESSMENT OF SEISMIC SAFETY: RESPONSE  
SURFACE APPROACH AND ACCELEROGRAM  
SELECTION ISSUES**

**Presentata da: Nicola Buratti**

**Coordinatore Dottorato  
Prof. E. Viola**

**Relatore  
Prof. M. Savoia**

**Correlatori  
Prof. J. J. Bommer  
Dr. P. J. Stafford**

**Esame finale anno 2008**



*Keywords:*

*Accelerogram*

*Response surface*

*Seismic fragility*

*Seismic hazard*

*Seismic risk*



# Index

<b>Introduction</b>	<b>I</b>
<b>1 Seismic risk and seismic hazard</b>	<b>1</b>
1.1 <i>Seismic Risk</i>	1
1.2 <i>Seismic Hazard Analysis</i>	3
1.2.1 Deterministic and probabilistic approaches to seismic hazard analysis	4
1.2.2 Different approaches to probabilistic seismic hazard assessment	6
1.3 <i>Recurrence relationships</i>	7
1.3.1 Modification of the general Gutenberg-Richter model at low magnitudes	8
1.3.2 Modification of the general Gutenberg-Richter model at large magnitudes	10
1.4 <i>Attenuation relationships</i>	11
1.4.1 Variability	13
1.5 <i>Mechanics of the Cornell-McGuire method for probabilistic seismic hazard analysis</i>	16
1.5.1 Integration limits	19
1.5.2 From hazard to probabilities	23
1.5.3 Aleatory and epistemic uncertainties	26
1.5.4 Uniform hazard spectra	27
1.5.5 Disaggregation of probabilistic seismic hazard	28
<b>2 Ground motion selection and scaling issues</b>	<b>31</b>
2.1 <i>Introduction</i>	31
2.2 <i>Ground motion selection</i>	31
2.2.1 Preselecting records	32
2.2.2 Selection based on geophysical parameters	36
2.2.3 Selection based on ground-motion properties and structural response	44
2.3 <i>Scaling</i>	59
2.3.1 Linear amplitude scaling	60
2.3.2 Frequency domain scaling	79
2.3.3 Time domain scaling	80
2.3.4 Number of records required to estimate the median structural response	82
2.4 <i>Bi-directional loading issues</i>	85
2.4.1 Direction of loading	85
2.4.2 Definition of spectral ordinates	89
2.4.3 Ground-motion selection and scaling	91
2.4.4 Vertical component of the ground-motion	97
2.5 <i>Conclusions and possible future work</i>	98
2.5.1 Scenario based selection	98

2.5.2	Selection based on UHS	99
2.5.3	Selection procedures based on non-linear structural response	100
2.5.4	Selection procedures for probabilistic approaches.	100
2.5.5	Bi-directional problems	101
<b>3</b>	<b>Accelerogram selection and scaling procedures for estimating the distribution of drift response</b>	<b>103</b>
3.1	<i>Introduction</i>	103
3.2	<i>Overview of the procedure</i>	104
3.2.1	Definition of the reference distribution	104
3.2.2	Description of the procedure	105
3.3	<i>Case study</i>	110
3.3.1	Ground-motion dataset	111
3.3.2	Structural model and nonlinear analysis	112
3.4	<i>Regression Analyses</i>	113
3.4.1	Spectral accelerations and PGA	114
3.4.2	Interstorey drifts	123
3.4.3	Correlations between $Sa(T_1)$ and $\delta$	129
3.4.4	Record Selection	130
3.4.5	Results	133
3.5	<i>Discussion and conclusions</i>	137
<b>4</b>	<b>Generation of non-stationary artificial accelerograms</b>	<b>141</b>
4.1	<i>Introduction</i>	141
4.2	<i>Model formulation</i>	143
4.3	<i>Computation of the time history</i>	145
4.4	<i>Estimation of the functions defining the spectrogram</i>	146
4.4.1	Bidirectional issues	148
4.4.2	Power amplitude	149
4.4.3	Arias Intensity	152
4.4.4	Duration	157
4.4.5	Central frequency and frequency bandwidth	162
4.4.6	Generation of accelerograms	171
4.5	<i>Artificial accelerograms as input for nonlinear structural analysis</i>	173
4.5.1	Case study structures	175
4.5.2	Definition of the length of plastic hinges	176
4.5.3	Overview of the analyses	179
4.5.4	Results of time-history analyses for the frame structure A	180
4.5.5	Results of time-history analyses for the frame structure B	182
4.5.6	Results of time-history analyses for the frame structure C	185
4.6	<i>Conclusions</i>	187

---

<b>5</b>	<b>Seismic reliability</b>	<b>191</b>
5.1	<i>Introduction</i>	191
5.2	<i>Sources of uncertainty</i>	191
5.3	<i>The reliability problem</i>	193
5.3.1	Limit states	193
5.3.2	Formulation of time invariant reliability problems	193
5.3.3	Statistical and model uncertainties	194
5.3.4	Time variant reliability problems	196
5.4	<i>Probability computation methods</i>	197
5.4.1	First- and second-order methods	198
5.4.2	Simulation methods	200
5.4.3	Response surface	207
5.5	<i>Specialized methods for earthquake engineering</i>	208
5.5.1	Seismic fragility and failure probability	208
5.5.2	Practical reliability methods	209
5.6	<i>Conclusions</i>	216
<b>6</b>	<b>Response Surface Method</b>	<b>217</b>
6.1	<i>Standard Response Surface</i>	217
6.1.1	Least squares method	220
6.1.2	Maximum likelihood estimation	222
6.1.3	Tests on the model	223
6.2	<i>Extension to time dependent problems</i>	224
6.2.1	Random effects and blocks	226
6.2.2	One-way classification	227
6.2.3	Multiple-way classification	230
6.2.4	Tests on the model	233
6.2.5	Prediction of random effects	234
6.3	<i>Design of the simulation plans</i>	236
6.3.1	Some considerations on the variables	236
6.3.2	Properties of a response surface design	237
6.3.3	The central composite design	243
6.3.4	Designs for mixed models	250
<b>7</b>	<b>Assessing seismic fragility of RC frames through response surface method</b>	<b>257</b>
7.1	<i>Seismic fragility</i>	258
7.2	<i>The case study: RC frame structure</i>	261
7.3	<i>Search of the structural capacity</i>	262
7.4	<i>Selection of the random variables</i>	266
7.4.1	Explicit random variables	266

## Index

---

7.4.2	Implicit random variables	266
7.5	<i>RS models and Design of Experiments</i>	268
7.6	<i>Predictions by different models</i>	269
7.6.1	Quadratic models	270
7.6.2	Linear models	274
7.6.3	Constant models	279
7.7	<i>Fragility analysis</i>	281
7.7.1	Quadratic models	282
7.7.2	Linear models	284
7.7.3	Constant models	286
7.7.4	Comparison of the fragility curves	289
7.8	<i>Conclusions</i>	292
<b>References</b>		<b>295</b>
<b>Appendix A</b>		<b>315</b>

# Introduction

Probably, no other branch of engineering has to deal with as much uncertainty as earthquake engineering, e.g. recurrence of earthquakes, intensity of earthquakes, ground motion features, soil effects, topographic effects, structural properties, nonlinear dynamic behaviour of structures, etc. Therefore it is necessary to assess seismic safety of structures using probabilistic procedures able to treat in a systematic and consistent way all the uncertainties involved in the problem. Things are even more complicated because seismic safety assessment procedures require the interaction of experts in different subjects, such as geology, engineering seismology, structural engineers, architects, loss experts, etc.

The main task that structural engineers are required to perform usually consists in assessing the vulnerability of structures. This kind of assessment cannot be performed independently of information regarding the characteristic of seismic action at the sites where the structures under investigation are located. In other words it is extremely important to keep consistency between seismic hazard procedures and seismic fragility assessment procedures.

The present work is divided into two parts, the first one (Chapters 1-4) is mainly focused on the characterization of input in seismic risk analysis, while the second part (Chapters 5-7) is focused on seismic fragility assessment of RC frame structures, and in particular on procedures based on response surfaces.

The concepts of seismic risk, vulnerability and hazard are introduced in Chapter 1. This chapter is mainly focused on presenting seismic hazard assessment procedures, which are the basis for every risk assessment analysis. Among these procedures a particular focus is given to probabilistic seismic hazard assessment procedures which are the *de facto* standard in seismic hazard assessment.

The typical outputs of probabilistic seismic hazard analyses are hazard curves and uniform hazard spectra. These results may not be sufficient, especially if the nonlinear behaviour of structures is to be investigated. In this case, in fact, the seismic action must be described in terms of acceleration time histories, i.e. accelerograms. These latter may be artificially generated or recorded during past earthquakes. This second category of accelerograms is the most widely used in the literature, since artificial accelerograms have been criticized by some researchers. Different methods have been proposed to select recorded accelerograms consistent with results of seismic hazard analyses. These methods often require to modify accelerograms, for example by a linear scaling. A state-of-the-art review of accelerogram selection and scaling procedures is given in Chapter 2.

The objectives of accelerograms selection and modification procedures may be different; the vast majority of those available in the literature are aimed at estimating the mean

value of some structural-response measure. In many applications (e.g. loss assessment) this kind of information does not suffice and information on the probability distribution of the considered -response measure is required. Chapter 3 presents a novel procedure, based on the theory of point approximation of continuous variables, to select recorded accelerograms. The aim of the procedure is estimating, with a low computational effort, the distribution of maximum interstorey-drift demand for reinforced concrete frames.

As previously anticipated, accelerograms may also be artificially generated. It has been shown that the oldest methods proposed in the literature generate accelerograms which are not suitable for nonlinear dynamic analysis of structures. More advanced procedures, recently proposed, are able to generate accelerograms which are free from the issues of the older procedures. Chapter 4, after a state-of-the-art review of the procedures to generate artificial accelerograms, presents a revision of the procedure proposed by Sabetta and Pugliese (1996) in order to generate non stationary artificial accelerograms. The accelerogram generation procedure is based on the approximation of the spectrogram of the accelerograms by means of lognormal functions. The parameters controlling the shape of these latter functions depend on some ground-motion features such as Arias intensity, duration and frequency content. The values of these parameters, required to generate an accelerogram, are estimated, as functions of earthquake magnitude, source-to-site distance and soil stiffness, using ground-motion prediction equations (attenuation relationships). The ground-motion predictive equations originally proposed by Sabetta and Pugliese (1996) have been revisited and updated in the present work, using a more comprehensive database of ground-motions. The accelerograms generated according to the updated Sabetta and Pugliese (1996) procedure have been used as input for nonlinear dynamic analysis of three case study RC frame structures. The results of these analyses suggest that these accelerograms may be a valid alternative to recorded accelerograms.

The second part of the present work is focused on the estimation of seismic fragility of concrete structures. Different methods have been proposed by many researchers with this aim. A review of the procedures which have been proposed in the literature is presented in Chapter 5. The classical random vibration theory, which has produced very important results for linear systems, has some limitations in dealing with strongly nonlinear systems, such as concrete structures. Simulation procedures are much more flexible, i.e. they can be used to address almost every class of problems, but often their high computational cost makes them unpractical. For these reasons many simplified, approximate, methods have been proposed by many researchers. Among these methods, which are described in Chapter 5, those based on response surfaces seem to be particularly promising. Response surfaces are statistical polynomial models which can be used to approximate the limit state surface of the structure under consideration. Data required to calibrate response surface models is collected by performing a series of simulations for some specific values

of the variables controlling the problem. The number of simulations required increases as the number of variables considered increases. This is a strong limitation for seismic reliability problems where the usually the number of variables involved is high. This issue can be addressed by including random factors in the response surface models. These factors allow to take in account implicitly, and therefore maintaining a low computational cost, the effects of some groups of variables. The theory or response surface models is described in Chapter 6, as well as the criteria to plan the simulations to be used to get data for calibrating the models.

Chapter 7 describes a application of the response surface method to the assessment of the seismic fragility of a case study RC frame structure. This case study structure has been used in order to investigate the efficiency of different response surface models as well as of different simulation plans.



# PART 1



# 1 Seismic risk and seismic hazard

## 1.1 Seismic Risk

Earthquake risk reduction is a complex problem involving many people of many vocations, much information, many options, and many decisions and actions.

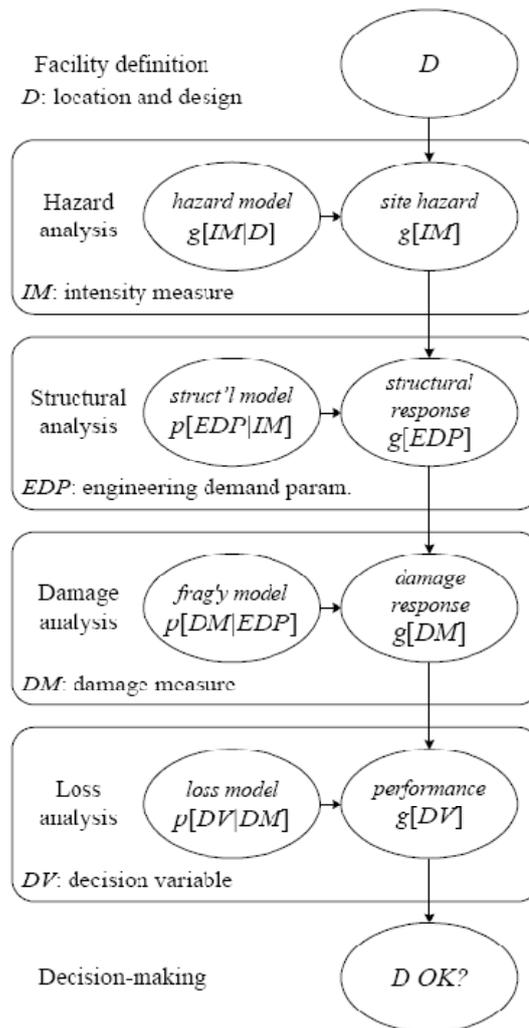
In normal English usage the word risk means exposure to the change of injury or loss. It is noted that the word hazard is almost synonymous with risk, and the two words are used in the risk literature with subtle variations which can be confusing. An authoritative attempt has been made to overcome this difficulty through the publication by the Earthquake Engineering Research Institute's glossary of standard terms for use in this subject (EERI Committee on Seismic Risk, 1984). Thus, the definition of seismic risk is the probability that social or economic consequences of earthquakes will equal or exceed specified values at a site, at several sites, or in an area, during a specified exposure time.

Seismic hazard, on the other hand, is any physical phenomenon (e.g. ground shaking, ground failure) associated with an earthquake that may produce adverse effects on human activities. Thus, hazards may be either purely descriptive terms or quantitatively evaluated, depending on the needs of the situation. In practice, seismic hazard is often evaluated for given probabilities of occurrence. It follows that seismic risk is an outcome of seismic hazard as described by relationships of the form

$$\text{Seismic Risk} = \text{Seismic Hazard} \times \text{Vulnerability} \times \text{Value} \quad (1.1)$$

where vulnerability is the amount of damage induced by a given degree of hazard and expressed as a fraction of the value of the damaged item under consideration. The  $\times$  symbols in Eq. (1.1) do not represent multiplication, in fact, how seismic hazard, vulnerability and value are to be combined depends on how they are expressed. Different ways to do so have been adopted in the numerous procedures proposed in the literature (Dowrick, 2003). To better understand the meaning of Eq. (1.1) it is worth considering an example. One of the most complete risk assessment frameworks recently proposed is the PEER's performance based earthquake engineering methodology (Porter, 2003).

The principal outputs of PEER's approach are system-level performance measures: probabilistic estimates of repair costs, casualties, and loss-of-use duration ("dollars, deaths, and downtime"). The objective of the methodology is to estimate the frequency with which a particular performance metric will exceed various levels for a given design at a given location. These can be used to create probability distributions of the



**Figure 1.1. Schematic representation of the PEER's performance based earthquake engineering procedure (after Porter, 2003).**

performance measures during any planning period of interest. From the frequency and probability distributions can be extracted simple point performance metrics that are meaningful to facility stakeholders, such as an upper-bound economic loss during the owner-investor's planning period. Figure 1.1 illustrates the PEER methodology. As it shows, PEER's PBEE approach involves four stages: hazard analysis, structural analysis, damage analysis, and loss analysis. In the figure, the expression  $p[X|Y]$  refers to the probability density of  $X$  conditioned on knowledge of  $Y$ , and  $g[X|Y]$  refers to the occurrence frequency of  $X$  given  $Y$  (equivalent to the negative first derivative of the frequency with which  $X$  is exceeded, given  $Y$ ). Eq. (1.2) frames the PEER methodology mathematically. Note that Figure 1.1 omits conditioning on  $D$  after the hazard analysis for brevity, but it is nonetheless implicit.

$$g(DV | D) = \iiint P(DV | DM, D) \cdot P(DM | EDP, D) \cdot P(EDP | IM, D) \cdot g(IM | D) dIM dEDP dDM \quad (1.2)$$

Eq. (1.2) can be considered as a specialization of Eq. (1.1): *hazard* maintains the same meaning, *vulnerability* is expressed as the combination of the results of two stages, structural analysis and fragility analysis, and *value* is evaluated by the loss analysis. Within this framework the  $\times$  symbol represents convolution.

Seismic fragility and vulnerability of structures will be described in greater detail in the second part of this dissertation, while this part is focused on seismic hazard assessment and on defining input for structural analysis in terms of accelerograms.

For design or risk assessment purposes the assessment of seismic hazard consists of the following basic steps:

- definition of the nature and location of earthquake sources
- magnitude-frequency relationships for the sources
- attenuation of ground motion with distance from source
- determination of ground motions at the site having the required probability of exceedance.

Because seismic risk and hazard statements are essentially forecasts of future situations, they are inherently uncertain. Seismic hazard assessment attempts to forecast the likely future seismic activity rate and strengths, based on knowledge of the past and present, and significant uncertainties arise partly because the processes involved are not fully understood and partly because relevant data are generally scarce and variable in quality. For reasonable credibility considerable knowledge of both historical seismicity and geology need to be used, together with an appropriate analysis of uncertainties. Where available other geophysical or seismological knowledge, such as crustal strain studies, may also be helpful, particularly in evaluating regional seismic activity patterns.

The present chapter introduces the basis of the procedures that may be used to assess seismic hazard, giving particular relevance to Probabilistic Seismic Hazard Analysis (PSHA). Some of the concepts described here (e.g. hazard curve, attenuation relationship, disaggregation) will be widely used in the following chapters of the present work.

## 1.2 Seismic Hazard Analysis

Seismic hazard could be defined, in the most general sense, as the possibility of potentially destructive earthquake effects occurring at a particular location. With the exception of surface fault rupture and tsunami, all the destructive effects of earthquakes are directly related to the ground shaking induced by the passage of seismic waves.

Textbooks that present guidance on how to assess the hazard of strong ground-motions invariably present the fundamental choice facing the analyst as that between adopting a deterministic or probabilistic approach.

### *1.2.1 Deterministic and probabilistic approaches to seismic hazard analysis*

Probabilistic seismic hazard assessment (PSHA) was introduced by Cornell in 1968 (Cornell, 1968) and has become the most widely used approach to the problem of determining the characteristics of strong ground-motion for engineering design. Some authors, however, have criticized the approach in defence of the deterministic seismic hazard assessment (DSHA), producing a vast debated in the scientific and technical literature. The basis of DSHA (Reiter, 1990; Kramer, 1996; Krinitzsky, 2002) is to develop earthquake scenario, defined by location and magnitude, which could affect the site under consideration. The resulting ground motions at the site, from which the controlling event is determined are then calculated using attenuation relations; in some cases, there may be more than one controlling event to be considered in design.

The mechanics of PSHA are less obvious than those of DSHA. Its essence (Cornell, 1968; Bazzurro and Cornell, 1999; Abrahamson, 2000b; Hanks and Cornell, 2001; Abrahamson, 2006) is to identify all possible earthquakes that could affect a site, including all feasible combinations of magnitude and distance, and to characterise the frequency of occurrence of different size earthquakes through a recurrence relationship. Attenuation equations are then employed to calculate the ground-motion parameters that would results at the site due to each of these earthquakes and hence the rate at which different levels of ground motion occur at the site is calculated. Common to both approaches is the fundamental issue of identifying potential sources of earthquakes. Another common feature is the modelling of the ground motion through the use of attenuation relationships (more correctly called ground-motion prediction equations). The principle difference in the two procedures, resides in those steps of PSHA that are related to characterising the rate at which earthquakes and particular levels of ground motion occur. As Hanks and Cornell (2001) point out the two approaches have far more in common that they do in differences. In fact, the only difference is that a PSHA has units of time and DHSAs does not. In DHSAs the hazard will be defined as the ground motion at the site resulting from the controlling earthquake, whereas in PSHA the hazard is defined as the mean rate of exceedance of some ground-motion amplitude. Once a mean rate of exceedance or probability of exceedance or return period is

selected as the basis for design, the output of PSHA is expressed in terms of a specified ground motion, in the same way as PSHA. Another important difference between the two approaches is related to the treatment of hazard due to different sources of earthquakes. In PSHA, the hazard contributions of different seismogenic sources are combined into a single frequency of exceedance of the considered ground motion parameter; in DSHA, each seismogenic source is considered separately, the design motions corresponding to a single scenario in a single source.

Regarding difference and similarities between the two methods, as Bommer notices (Bommer, 2002), it is often pointed out that probabilities are at least implicitly present in DSHA in so far as the probability of a particular earthquake scenario occurring during the design life of the engineering project is effectively assigned as unity. An alternative interpretation is that within a framework of spatially distributed seismicity, the probability of occurrence of a deterministic scenario is, mathematically, zero, but in general the implied probability of one is a valid interpretation of the scenarios defined in DSHA. As regarding the resulting ground motions, however, the probability depends upon the treatment of the scatter in the strong-motion prediction equation; if the median plus one standard deviation is used, this will correspond to a motion with a 16-percent probability of exceedance, for the particular earthquake specified in the scenario.

There is not a single, universally accepted approach to DSHA. However, a paper by Krinitzsky (Krinitzsky, 2002), has become a sort of standard reference for DSHA. One important difference between DSHA as proposed by Krinitzsky (2002) and DSHA as described by Reiter (1990) and Kramer (1996), is that whereas the latter imply that the ground motions for each scenario should be calculated using median (50-percentile) values from strong-motion scaling relationships, Krinitzsky (2002) proposes the use of the median-plus-one-standard deviation (84-percentile) values.

It can equally be pointed out that any PSHA includes many deterministic elements in so much that the definition of nearly all of the input requires the application of judgements to select from a range of possibilities. This applies in particular to the definition of the geographical limits of the seismic sources zones and the selection of the maximum magnitude. In addition to the various parameter that define the physical model that is the basis for any PSHA, it could also be argued that another parameter, which has pronounced influence on the input to engineering design, is also defined deterministically: the design probability of exceedance. This latter issue, which is of fundamental importance is discussed in Bommer (2002) and Bommer (2006).

### *1.2.2 Different approaches to probabilistic seismic hazard assessment*

The formal beginning of PSHA, can be traced back to the paper by Cornell (Cornell, 1968). Important developments included the development of the software EQRISK by McGuire (McGuire, 1976). A significant difference between EQRISK and the original formulation of Cornell was the inclusion of the influence of the uncertainty or scatter in the strong-motion prediction equation. Two fundamental features of the Cornell-McGuire method are definition of seismogenic zones, as areas or lines, with spatially uniform activity and the assumption of a Poisson process to represent the seismicity, both of which have been challenged by different researchers who have proposed alternatives. Many alternatives to uniformly distributed seismicity within sources defined by polygons have been put forward, such as Bender and Perkins (Bender and Perkins, 1982; Bender and Perkins, 1987) who proposed sources with smoothed boundaries, obtained by defining a standard error on earthquake locations.

There have also been proposals to do away with source zone altogether and use the seismic catalogue itself to represent the possible locations of earthquakes, an approach that may have been used before 1968. Such historic approaches can be non-parametric or parametric. Recent adaptation of these “zone-free” methods include the approach based on spatially smoothed historical seismicity of Frankel et al. (1996) and the kernel method of Woo (1996).

The differences amongst these different approaches to PSHA are not simply academic: Bommer et al. (1998) produced hazard maps for upper-crustal seismicity in El Salvador determined using the Cornell-McGuire approach, two zone-free methods and the kernel method. The four hazard maps, prepared using the same input, showed very significant differences in the resulting spatial distribution of the hazard, and the maximum values of PGA vary amongst the four maps by a factor of more than two.

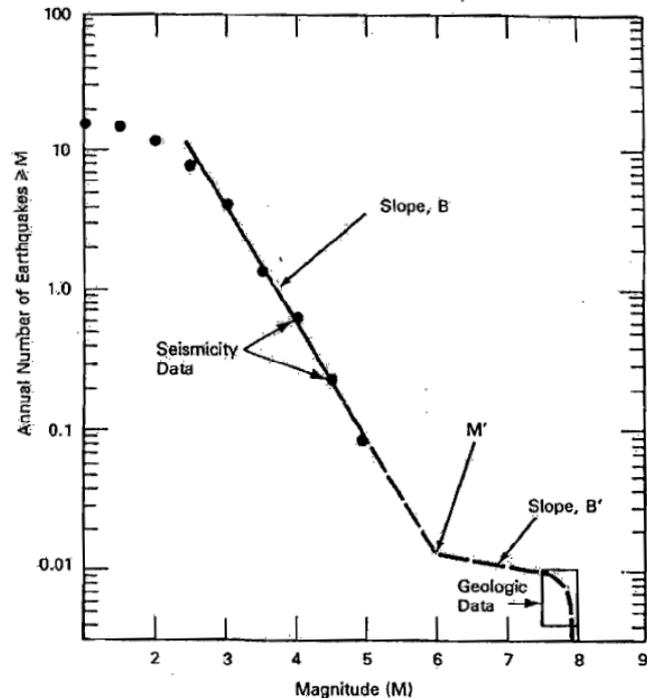


Figure 1.2. Recurrence relationships for earthquakes that fit the characteristic earthquake model (after Reiter, 1990).

### 1.3 Recurrence relationships

The two most widely used models to describe the frequency of earthquakes in seismic sources are the Gutenberg-Richter model (Gutenberg and Richter, 1944) and the characteristic earthquake model (Schwartz and Coppersmith, 1984b). This latter model requires more parameters to be estimated than the first one. There is a lot of evidence to support both models, and a great deal of controversy regarding which of the two is the most appropriate for application in a given situation (Wesnousky, 1994; Kagan, 1996; Wesnousky, 1996). It should be noted from the outset that both of these models are essentially empirical statistical models; there is no generally accepted physical basis to either of them. Additionally, there are other models that may be employed, and that may produce a stronger statistical fit to the data, but they require the specification of more parameters than the Gutenberg-Richter model and consequently their feasibility of their use depends on the dataset size. In the following only the Gutenberg-Richter and the Characteristic model will be briefly described because they are the most widely adopted models. They both share many common attributes. For small to moderate sized events

the two distributions are the same and are based upon the general form of the Gutenberg-Richter relation. The original form of this relation is typically expressed as:

$$\log_{10}(N(m)) = a - b m \quad (1.3)$$

where  $N(m)$  is the number of events of magnitude equal to or greater than occurring throughout the observation period,  $a$  is the number of events with  $m \geq 0$ , and is named the activity parameter, while  $b$  defines the relative frequency of occurrence between events of different sizes. In their initial study, Gutenberg and Richter found values of  $b$  close to 1 over a reasonable range of magnitudes. Since then, Eq. (1.3) has been applied to many other regions of the world, and only small departures from a  $b$  value of 1.0 have been observed.

The basic form of the Gutenberg-Richter relationship given in Eq. (1.3) is usually modified for use in general applications in two fundamental ways. Firstly, the range of magnitudes is restricted so that only events having a magnitude larger than some particular value are considered, and secondly, the form of the relation at large magnitudes is modified. This latter modification constitutes the difference between the general Gutenberg-Richter model and the Characteristic model (see Figure 1.2). These modifications will be shortly described in the following.

### ***1.3.1 Modification of the general Gutenberg-Richter model at low magnitudes***

There are four generic issues to address in regard to this lower bound modification. Firstly, is there some lower physical limit to the size that an earthquake, in the most general sense of the expression, may possess? Secondly, does the power law scaling of earthquake size break down at some level of magnitude? Thirdly, and irrespective of the previous two points, what is the smallest size earthquake that the national seismic network can accurately detect? And finally, what is the magnitude of the smallest size earthquake that is capable of generating ground motions large enough to cause damage to engineered structures?

For the purpose of PSHA the first issue of a theoretical lower limit of the size of magnitudes is of little practical relevance as seismic events at the scales in question are not capable of generating ground motions of interest for the engineering community.

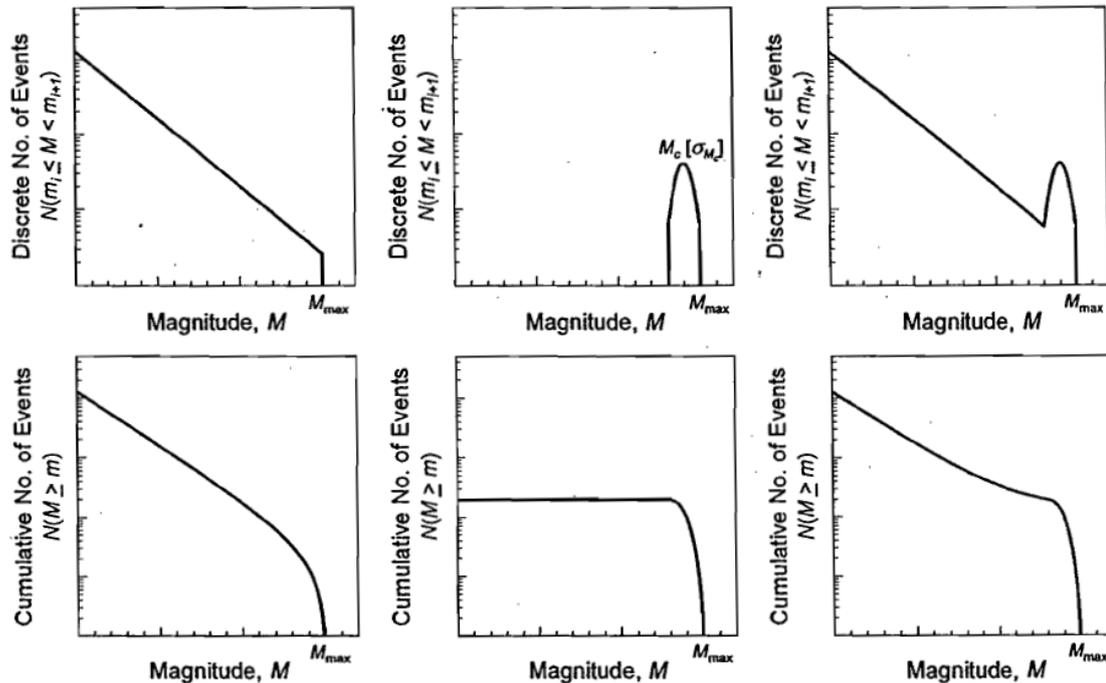


Figure 1.3. Forms of earthquake recurrence relationships in non-cumulative (top row) and cumulative (bottom row) forms: Gutenberg-Richter model (left panels), maximum magnitude model (centre panels), and characteristic earthquake model (right panels). (After Bommer and Stafford, 2007)

The second issue however, is of much greater relevance. Various researchers have given evidence to support that the validity of the power law scaling to very-small magnitudes. However there also exists evidence to the contrary. The general consensus however, is that the Gutenberg-Richter relation is applicable to very low magnitude levels. The issue of the smallest magnitude able to be consistently detected for a seismic network is of direct relevance to PSHA analyses and should be considered when processing earthquake catalogues. Typically the total catalogue can be partitioned into various interval of completeness.

The last of the four issues mentioned above is perhaps the most important one in PSHA framework. When conducting a PSHA, the probabilistic contributions to ground motions exceedances at a particular site are calculated for all magnitudes above a certain minimum size. Therefore, the numerical values assigned to seismic hazard for a site are directly related to the selection of the minimum magnitude.

There are several issues relating to the selection of the minimum magnitude, some of which are discussed by Bender and Campbell (1989). One reason for there not being universally accepted values for this parameter is that the value chosen should be problem specific. Whether or not a certain size event should be considered or not

depends upon the structure for which the assessment is being undertaken. The selection of the most appropriate value must be chosen will be discussed later.

Given the above considerations, no distinction can be made between the Gutenberg-Richter and characteristic models based upon considerations of the lower magnitude limit. The limit that is imposed in determining the parameters of the models normally coincide with the completeness levels of the considered catalogue. The relation is simply truncated at the lowest magnitude considered. Consequently, the probability distributions related to this curtailment need to be normalized to ensure that the total probability theorem is not violated.

### *1.3.2 Modification of the general Gutenberg-Richter model at large magnitudes*

The modification that is made over the domain of large magnitudes represents the point of departure of the Characteristic model from the Gutenberg-Richter models. In fact Schwartz and Coppersmith (1984b) found evidence for recurrence intervals for large magnitudes that significantly exceeded the rates expected from the extrapolation of the Gutenberg-Richter relation to the equivalent magnitudes of the large events. This behaviour is more representative of the magnitude-frequency nature of earthquakes from fault sources. However, there are also many examples where faults do not exhibit this.

One of the arguments presented in support of a characteristic fault model is that the Gutenberg-Richter relation that is commonly observed is the results of a huge number of Characteristic earthquake sources, each relating to varying spatial scales, acting together (Wesnousky, 1994). However, recurrence intervals estimated for large events are usually based upon a very small dataset of a few points at most and inference made with respect to departures from power-scaling may be related with the incompleteness of the earthquake data.

An obvious way to determine whether the characteristic model is more appropriate for modelling future earthquake occurrence than the Gutenberg-Richter relation is to consider the dataset obtained from the seismicity procedures outlined above. This, of course, requires a large enough dataset. The other source of information is from paleoseismic investigations.

Previously, it was implied that the Characteristic model requires more parameters to define the recurrence rates at the upper magnitude ranges. These parameters are usually provided by geologic information rather than from statistics. In the case where geologic

information does not exist it is still possible to employ the Characteristic distribution by making some assumptions based upon the form of characteristic models for faults where the data is available (Schwartz and Coppersmith, 1984b).

Regardless of which model is finally adopted, the probability distributions must be normalized in order to satisfy the total probability theorem. The normalization process for the Characteristic model must be done on a case by case basis as the magnitude, and the distribution for the magnitude, of the annual rate at large magnitudes will vary depending upon the fault source in question. In this case, the Gutenberg-Richter relation, is referred to as doubly bounded Gutenberg-Richter relation (Kramer, 1996).

## 1.4 Attenuation relationships

Numerous options that are available for describing the intensity of ground motions (see Section 2.3.1.1). Now, given a large number of records, one can calculate values for any of these parameters and obtain a robust estimate of the correlation of these values with any other parameter relevant to this suite of records, such as the magnitude of the earthquake from which they came. This type of reasoning is the basis for the development of empirical predictive equations for strong ground-motions. Usually, a relationship is sought between a suite of observed ground-motion parameters and an associated set of independent variables including a measure of the size of the earthquake, a measure of distance from the source to the site, some classification of the style-of-faulting involved and some description of the geological and geotechnical conditions at the recording site.

An empirical ground-motion prediction equation is simply a function of these independent variables that provides an estimate of the expected value of the ground-motion parameter in consideration as well as some measure of the distribution of values about this expected value.

Thus far the development of empirical ground-motion prediction equations has been almost exclusively focussed upon the prediction of peak ground-motions, particularly PGA and, to a far lesser extent, PGV, and ordinates of 5% damped elastic acceleration response spectra (Douglas, 2003; Bommer and Alarcon, 2006). Predictive equations have also been developed for many of the most important ground-motion parameters, but as seismic design actions have historically been derived from PGA or  $S_a(T)$  the demand for such equations is relatively weak. However, the performance of PGA (Wald et al., 1999) and, to a lesser extent,  $S_a(T)$  (Priestley, 2003; Akkar and Özen, 2005 ) for the purposes of predicting structural damage has begun to be questioned. Improvements

in the collaboration between engineering seismologists and structural earthquake engineers has prompted the emergence of research into what really are the key descriptors (such as inelastic spectral ordinates and elastic spectral ordinates for damping ratios other than 5%) of the ground motion that are of importance to structural response and to the assessment of damage in structures.

Regardless of the ground-motion intensity measure in consideration, a ground-motion prediction equation can be represented as a generic function

$$\log(y) = \mu(M, R, \boldsymbol{\theta}) + \varepsilon\sigma_T \quad (1.4)$$

of predictor variables,  $\mu(M, R, \boldsymbol{\theta})$ , where  $M$  is magnitude,  $R$  distance and  $\boldsymbol{\theta}$  is a vector of parameters, and a variance term,  $\varepsilon\sigma_T$ , where  $\sigma_T$  represents the total standard deviation and  $\varepsilon$  is a standard normal variable.

Many developers of ground-motion prediction equations attempt to assign physical significance to the terms in the empirically derived function  $\mu(M, R, \boldsymbol{\theta})$ . In some cases it is possible to derive theoretical equations that may be used as the basis for selecting appropriate functional forms (e.g., Douglas, 2002). Although these theoretical considerations enable us to select appropriate functional forms, once the regression analysis has been conducted the actual values of regression coefficients should not be interpreted as having physical meaning as correlations of varying degrees always exist between the coefficients for different terms of the model.

For most ground-motion measures the values will increase with increasing magnitude and decrease with increasing distance. These two scaling effects form the backbone of prediction equations and many functional forms have been proposed to capture the variation of motions with respect to these two predictors (Douglas, 2003). For modern relationships distinctions are also made between ground motions that come from earthquakes having different styles of faulting, with reverse faulting earthquakes tending to generate larger ground motions than either strike-slip or normal faulting events (Bommer et al., 2003). Historically, account was also taken for site conditions by adding modifying terms similar to those used for the style-of-faulting effects - stiff soil sites have larger motions than rock and soft soil sites have larger motions still. In Europe this use of dummy variables for generic site classes remains the adopted approach in the latest generation of prediction equations (Ambraseys et al., 2005; Akkar and Bommer, 2007a; Akkar and Bommer, 2007b), primarily due to the absence of more detailed site information. However, in the US site response is now modelled using the average shear-wave velocity over the upper 30m, as introduced by Boore et al. (1997).

Furthermore, the influence of non-linear site response, whereby weaker motions tend to be amplified more so than stronger motions due to the increased damping and reduced strength associated with the latter, is also taken into account (Abrahamson and Silva, 1997; Choi and Stewart, 2005).

In addition to the basic scaling of ground motions with magnitude, distance, site conditions, etc., there are additional situations that may result in modified ground motions that are commonly either omitted from developed equations or are later applied as correction factors to the base models. The most common examples include accounting for differences between sites located on the hanging or foot wall of dip-slip fault sources (Abrahamson and Somerville, 1996; Chang et al., 2004), accounting for rupture directivity effects (Somerville et al., 1997a; Abrahamson, 2000a), including models for the velocity pulse associated with directivity effects (Bray and Rodriguez-Marek, 2004), basin effects (Choi et al., 2005) and topographic modifiers (Toshinawa et al., 2004). The most recent predictor variable to be included in prediction equations for peak ground-motions and spectral ordinates is the depth to the top of the rupture (Kagawa et al., 2004; Somerville and Pitarka, 2006). Currently, none of these effects are incorporated into any predictive equations for ground motions in Europe, nor is any account made for non-linearity of site response. Again, this is primarily a result of the lack of well-recorded strong earthquakes in the region.

### 1.4.1 Variability

For any particular ground-motion record the total variance term given in Eq. (1.4) may be partitioned into two components as:

$$\log(y_{ij}) = \mu(M_i, R_{ij}, \boldsymbol{\theta}_{ij}) + \delta_{e,i} + \delta_{a,ij} \quad (1.5)$$

where the terms  $\delta_{e,i}$  and  $\delta_{a,ij}$  represent the inter-event and the intra-event residuals respectively and quantify how far away from the mean estimate of  $\log(y_{ij})$  the motions from the  $i$ -th event and the  $j$ -th recording from the  $i$ -th event are respectively (Abrahamson and Youngs, 1992).

Alternatively, these terms may be expressed in terms of standard normal variates ( $z_{e,i}$  and  $z_{a,ij}$ ) and the standard deviations of the inter-event ( $\tau$ ) and intra-event ( $\sigma$ ) components, i.e.  $\delta_{e,i} = z_{e,i}\tau$  and  $\delta_{a,ij} = z_{a,ij}\sigma$ . The total standard deviation for a predictive equation is obtained from the square root of the sum of the inter-event and intra-event variances,  $\sigma_T^2 = \tau^2 + \sigma^2$ . Later, when discussing PSHA, mention will be made of epsilon,  $\epsilon$ , representing the number of total standard deviations from the

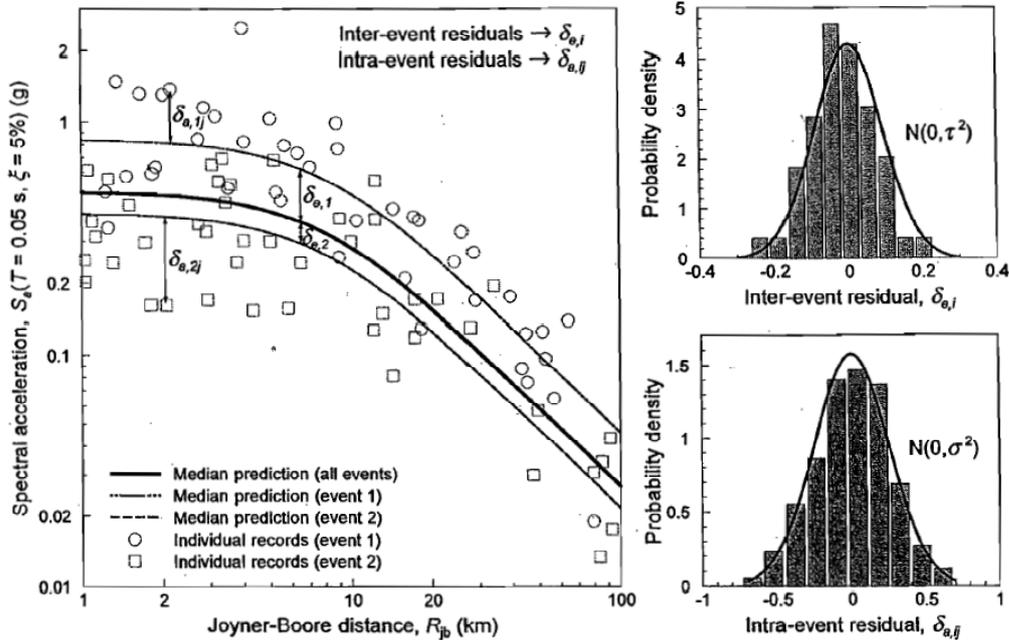


Figure 1.4. Graphical representation of inter-event and intra-event residuals (after Bommer and Stafford, 2007).

median predicted ground-motion. Often ground motion modellers represent the terms  $\delta_{e,i}$  and  $\delta_{a,ij}$ ; by  $\eta_i$  and  $\varepsilon_{ij}$  respectively. Under this convention care must be taken to not confuse the epsilon,  $\varepsilon$ , with the intra-event residual,  $\varepsilon_{ij}$ , term. The two are related via the expression  $\varepsilon = (\eta_i + \varepsilon_{ij})/\sigma_T$ , i.e.  $\varepsilon = (\delta_{e,i} + \delta_{a,ij})/\sigma_T$  using the notation adopted in Eq. (1.5).

When both components of a ground-motion record are used the variance structure of the data is better modelled via the form:

$$\log(y_{ijk}) = \mu(M_i, R_{ij}, \theta_{ijk}) + \delta_{e,i} + \delta_{a,ij} + \delta_{c,ijk} \quad (1.6)$$

where the term  $\delta_{c,ijk}$  represent the inter-component residuals and quantify how far away from the mean estimate of  $\log(y_{ijk})$  the motion of  $k$ -th component from the  $j$ -th recording and the  $i$ -th event is .

Each of these components of variability may be modelled as functions of other parameters such as the magnitude of the earthquake (Youngs et al., 1995), the shear-wave velocity at the site (Abrahamson and Silva, 2008b; Abrahamson and Silva, 2008a), or the amplitude of the ground motion (Campbell, 1997). Exactly how these components are calculated depends upon the regression methodology that is used to derive the equations. However, the most common approach is to adopt random effects

procedures where the correlation between ground motions observed within any particular event is assumed to be the same across events. This concept is shown schematically in Figure 1.4.

Many people think of ground-motion variability as a measure of the lack of fit of a particular predictive equation. However, in most cases it is better to think of a predictive equation as providing an estimate of the distribution of ground motions given a set of predictor variables such as magnitude and distance. From this perspective, the real misfit of the model is related to how well the model's distribution represents the true distribution of ground motions rather than how large are the variance components. People tend not to like large variability, reasoning that this implies that we cannot predict this measure of ground motion with much certainty. However, this perspective is closely related to the paradigm that ground motions are ultimately predictable and that it is only through a result of inadequate modelling and incomplete knowledge that the apparent variability arises. If, on the other hand, one views ground motions as being inherently unpredictable (beyond a certain resolution) then one must view the variability not as a measure of the misfit, but rather as an additional part of the model that describes the range of observable ground motions given an event. Under this latter paradigm there is no reason to like or dislike a particular ground-motion measure simply because predictive equations for this measure have a broad distribution. The only rational basis for judging the importance of a ground-motion measure is to assess the ability of this measure to accurately predict structural response. That said, in most cases, less variability in the ground motion estimate will translate into less variability in the response.

The ground-motion variability plays a very important role in seismic hazard analysis. In many cases the ground-motion variability is the most important component of a ground-motion model as it is the component of the model that defines how likely levels of ground motion are for a particular magnitude-distance scenario. The specification of this distribution of values, and their proper treatment, is central to PSHA yet there are many instances in practice where this element of the model is not treated appropriately. Bommer and Abrahamson (Bommer and Abrahamson, 2006) provide a comprehensive summary of the historical development of PSHA and the important role that the ground-motion variability has played.

## 1.5 Mechanics of the Cornell-McGuire method for probabilistic seismic hazard analysis

While the mathematical formulation of PSHA can seem complex, most of the mechanics behind the framework are very simple.

Formally, basic PSHA may be represented as in Eq. (1.7) (Bazzurro and Cornell, 1999; Abrahamson, 2006; Bommer and Abrahamson, 2006):

$$\lambda_{GM}(gm^*) = \sum_i \left\{ \iiint I[GM > gm^* | m, r, \varepsilon] v_i f_{M,R,E}(m, r, \varepsilon)_i dm dr d\varepsilon \right\} \quad (1.7)$$

where the capital letters represent random variables (GM: a chosen ground-motion parameter (such as PGA, spectral acceleration, Arias intensity, etc.), M: magnitude, R: distance, and E: epsilon) while their lower-case counterparts represent realizations of these random variables. The total rate at which earthquakes occur having a magnitude greater than the minimum considered for source  $i$  is denoted by  $v_i$  (as this term is a constant for each source it may be taken outside the integral in Eq. (1.7)). The joint probability density function of magnitude, distance and epsilon is given by  $f_{M,R,E}(m, r, \varepsilon)_i$  and  $I[GM > gm^* | m, r, \varepsilon]$  is an indicator function equal to one if  $GM > gm^*$  and zero otherwise. Finally,  $\lambda_{GM}(gm^*)$  is the total annual rate at which the target ground-motion value,  $gm^*$ , is exceeded. This is often the way that PSHA is presented in the literature; however it is worth to spend some time to describe the nature of the joint probability density function in magnitude, distance and epsilon appearing in Eq. (1.7).

First of all, it should be observed that because  $\varepsilon$  is statistically independent of  $M$  and  $R$  (although  $\sigma_{\log(IM)}$  might not be so), then the joint probability function may be expressed as:

$$f_{M,R,E}(m, r, \varepsilon) = f_{M,R}(m, r) f_E(\varepsilon) \quad (1.8)$$

where  $f_E(\varepsilon)$  represents the standard Normal distribution (see Section 1.4.1). A further decomposition is almost always made; that being to decompose the joint probability density function in magnitude and distance into the product of the probability density functions corresponding to each variable separately. This methodology is presented in Kramer (Kramer, 1996) and is expressed in Eq. (1.9).

$$f_{M,R,E}(m, r, \varepsilon) = f_{M,R}(m, r) f_E(\varepsilon) = f_M(m) f_R(r) f_E(\varepsilon) \quad (1.9)$$

The probability density function with respect to distance,  $f_R(r)$ , is usually generated by assuming that events are equally likely to occur anywhere with a given source and the probability density function is consequently only related to the geometry of the source. However in reality the decoupling presented in Eq. (1.9) is not possible as modern predictive equations make use of distance measures that do not correspond to point sources. The above decomposition is valid when point source measures of distance are made as in that case one is able to make the assumption that events are equally likely to occur anywhere within the source and that the corresponding source site distance are simply the geometric distance from the site to each point that is considered within the source. It has long been recognised however, that point source distance measures are generally not appropriate for modelling strong ground motions (Bolt and Abrahamson, 2003; Campbell, 2003). If the distance measures used in a PSHA analysis are related to the size of the rupture surface, which in turn is related to the size of the earthquake creating this surface, then the distance from a site to the rupture surface depends upon the hypocentral position as well as the magnitude of the event. The joint probability density function should therefore be written as conditional distribution dependent upon the magnitude of the causative event. The more appropriate form of the joint probability distribution function is therefore given in Eq. (1.10).

$$f_{M,R,E}(m,r,e)_i = f_M(m | \mathbf{x}_{hyp}) f_{\mathbf{x}_{hyp}}(\mathbf{x}_{hyp}) f_R(r | m, \mathbf{x}_{hyp}, \boldsymbol{\theta}_i) f_E(\varepsilon) \quad (1.10)$$

where  $\mathbf{x}_{hyp}$  denotes the hypocentral distance, and the vector  $\boldsymbol{\theta}$  represents a set of parameters that describe the orientation of the rupture surface, such as the strike and dip of the fault, the depth to the bottom of the fault, the segment of the fault the hypocentre is located in, and so on, i.e.  $\boldsymbol{\theta} = \{\theta, \delta, z_{seis}, segment, \dots\}$ . In this case, the probability density function with respect to hypocentral distance,  $f_R(r | m, \mathbf{x}_{hyp}, \boldsymbol{\theta}_i)$ , makes use of the assumption that events are likely to occur anywhere in the source, but the conditional probability density of the actual distance measure used to calculate hazard contributions from the attenuation equations must account for other parameters, including magnitude. This framework for hazard is based on that of Der Kiureghian and Ang (1977). Eq. (1.10) explicitly shows how the hazard computations are dependent upon the nature of the scaling relationships adopted in the analysis. In the following the terms appearing in Eq. (1.10) will be explained in detail.

$f_{\mathbf{x}_{hyp}}(\mathbf{x}_{hyp})$ : the pdf for an event having a hypocentre equal to  $\mathbf{x}_{hyp}$ , where  $\mathbf{x}_{hyp} = (longitude, latitude, depth)$  is any position within source  $i$ . A common assumption that is made, and that was made in Cornell's original presentation of PSHA,

is that hypocentres are equally likely to occur anywhere within a seismic source. This assumption requires the least amount of information regarding the nature of activity for the seismic source.

$f_M(m | \mathbf{x}_{hyp})$  - The conditional pdf of magnitude given the hypocentral position. In many hazard analyses this term is not implicitly considered, instead analyst simply take the previous assumption that earthquakes may occur with equal probability anywhere within a seismic source and also assume that these events may have the full range of magnitudes deemed possible for the source. In this case this term is not conditioned upon the hypocentre position and one simply recovers  $f_M(m)$ , the pdf of magnitude. However, some analysts may wish to address this problem thoroughly and make alternative assumptions using analyses such as those of Somerville et al., 1999 and Mai et al., 2005. For example, it may be assumed that large earthquakes tend to have relatively deep hypocentres and the pdf may be modified accordingly. The pdf of magnitude is often assumed to follow a doubly-bounded exponential distribution for areal sources (Cornell and Vanmarcke, 1969); a modified form of the Gutenberg-Richter equation (Gutenberg and Richter, 1944), and a characteristic distribution for fault sources (Schwartz and Coppersmith, 1984a). However any distribution that relates the relative rates of occurrence of earthquakes of different size is permissible.

$f_R(r | m, \mathbf{x}_{hyp}, \boldsymbol{\theta}_i)$  - The conditional pdf of the distance measure used in the ground-motion prediction equation gives the rupture surface of the earthquake. The rupture surface depends upon the hypocentre, the size of the event and various other parameters encapsulated in  $\boldsymbol{\theta}_i$  including the strike and dip of the fault plane (for fault sources), the depth boundaries of the seismogenic zone, the segment of the fault on which the rupture starts, etc. This term is important as it translates the assumptions regarding the potential locations of earthquakes into measures of distances that are appropriate for use in empirical prediction equations. Note that this term is necessarily different for each distance measure that is considered.

$f_E(\epsilon)$  - The pdf of epsilon. It is important to note that this term is always simply the pdf of the standard normal distribution. For this reason it is not necessary to make this a conditional pdf with respect to anything else. Although standard deviations from ground-motion predictive equations may be dependent upon predictor variables such as magnitude, the pdf of epsilon remains statistically independent of these other variables (Bazzurro and Cornell, 1999).

Given this more complete representation of Eq. (1.10) one must now evaluate the integrals in Eq. (1.7) once they are expressed in terms of the relevant variables. The integrals are not evaluated analytically and all that is required is to discretize the range of

possible parameter values and to determine the contribution to the hazard from each permissible permutation of each parameter.

### *1.5.1 Integration limits*

In the situation defined by Eq. (1.7), the integration must be performed with respect to magnitude, distance, and epsilon. In order to proceed practically, one must define the range of each of these variables for which the integration must be performed. Some of these limits have received considerable attention in the literature, in particular the upper limit on magnitude, while the remaining limits have received relatively little attention.

#### *1.5.1.1 Limits on magnitude*

Usually the magnitude frequency relationship is specified as the double bounded Gutenberg-Richter relationship. The upper bound of which automatically precludes the consideration of any magnitude event greater than this value. The upper limit on the magnitude range is therefore automatically imposed from the fault source model and does not need to be discussed in any more detail here.

Considerable attention has gone into developing procedure for estimating the upper bound on the range of magnitudes that a given source can generate. Conversely, very little attention has been given to the specification of the minimum magnitude, even though this parameter can have a significant influence upon hazard estimates, particularly for short period ordinates of spectral acceleration. The degree to which minimum magnitude affects the calculated seismic hazard depends on many factors, including the level of seismicity, the type of zonation, the maximum magnitude, the variability in ground motion, the period (in the case of response spectra), and the attenuation relationship. Decreasing the minimum magnitude though to be of engineering relevance, causes a drastic increase of the hazard at small mean return periods.

The PSHA procedure must represent the contribution to hazard of all potentially damaging earthquake scenarios. The problem of specifying a minimum magnitude is therefore a problem in defining what a potentially damaging earthquake is. As has been previously mentioned an earthquake scenario consists of, at least, a magnitude-distance pair. This coupling between the magnitude of the event and the distance between the source and the site is critical because the core of the problem is defining ground motion levels at the site of interest.

It can be shown (Stafford, 2006) that reducing the minimum magnitude of engineering relevance will have a more marked effect on increasing the calculated hazard for short

period spectral accelerations (peak ground acceleration included) than long period motions. The degree to which this effect is observed depends upon the relative rates of activity of the seismic source in the vicinity of the site being considered. If sources proximal to the site have high rates of activity then the high frequency components of distal sites is obviously also true. The influence of the distance is minimized, however, by the fact that the high frequency ground motions are severely attenuated in the near surface where lithospheric pressures are low and allow cracks and fissures to exist in the propagating medium.

Cornell and Vanmarcke (1969) presented evidence of the influence that proximal, small magnitude events have on the calculation of hazard. In light of these findings, it is very important to ascertain the destructive potential these small, close events. Bommer et al. (2001), investigated some historical events of small-moderate magnitude that caused significant destructions. The key finding of the Bommer et al. (2001) research was the strong correlation between the destructive potential of these small-moderate events and the focal depth of these events. Shallow events were found to be far more likely to cause destruction than deeper ones. This is a natural consequence of the rate at which near field effects attenuate coupled with the initial strength of these waves. Another finding of the above research was that there appears to be evidence of directivity effects being observed in small-moderate sized events. Traditionally, the small rupture areas associated with small to moderate events implied that these events could be regarded as point sources. Consideration of directivity effect was therefore constrained to cases for magnitudes of 6.5 and above (Somerville et al., 1997a).

It is also worth noting that while the probability of a small events causing significant damage to engineered structures remains low, even taking into account shallow focal depths, site amplification, and forward directivity effects, many of these small events occur. The effect of these numerous events may act in conjunction with each other, i.e. while a single event may not be significant enough to incite large scale damage, it may be significant enough to weaken the existing building stock to the point where a successive event of a similar nature may cause ensuing damage. The cumulative effect of numerous small events should therefore be considered in the selection of the minimum magnitude used in the PSHA.

#### *1.5.1.2 Limits on distance*

The selection of a lower bound to the distance range may seem trivial, i.e. the minimum distance must be zero. However, this may not necessarily be the case. Cornell and Vanmarcke (1969) comment on the influence of this selection with respect to area

sources. In addition, Campbell (1997) uses a measure of distance, that precludes the occurrence of earthquakes in the non-seismogenic near surface. While Campbell (1997) gives a clear and explicit reason for requiring all the events to occur at some distance from the site in question, most other attenuation relationships inadvertently include a parameter that essentially increases the theoretical distance between the source and the site. His parameter that modifies the distance is usually simply described as a regression parameter that enables to better fit to the empirical data. Regardless of its physical meaning, the result is that while the trivial case of a lower bound of zero for the distance variable is commonly adopted, in reality the effective lower bound on the distance is slightly above zero.

The upper bound on distance is rarely mentioned in PSHA. This parameter may have influence upon the results of the hazard analysis if care is not taken in being consistent when defining the seismicity rates for various sources. Again, Cornell and Vanmarcke (1969) address this issue in arriving at the conclusion that small events, close to the site, contribute strongly to the total hazard at a site. There is a coupling between the frequency of events of magnitude greater than some threshold level and the volume of the region considered in the derivation of this value. Obviously, the larger the region considered, the higher the frequency of events exceeding a given magnitude. One must therefore ensure that the activity rates assigned to seismic sources are consistent with the region used in the derivation of these activity rates.

### *1.5.1.3 Limits on epsilon*

The essence of PSHA is to determine the rate at which some ground motion is exceeded. As previously observed the pdf of epsilon is always a standard normal distribution, therefore, in terms of ground motions, this implies that negative values of epsilon characterise events in which the recorded ground motions resulting from an earthquake scenario (magnitude-distance pair) are lower than that which would ordinarily be expected. Therefore, if it is fair to assume that ground motions are lognormally distributed, which many authors suggest it is (Campbell, 1985; Douglas and Smit, 2001; Restrepo-Velez and Bommer, 2003), one can use the limits of the above equation for the normal distribution to infer the lower limit on the value of epsilon to be used in PSHA. The lower bound on epsilon is consequently trivial, and correspond to the conditions where it is certain that a magnitude-distance pair will exceed the target ground motion measure; that being epsilon equal to negative infinity. The upper bound on epsilon on the other hand is far from trivial. Positive values of epsilon represent ground motions that are in excess to the median predicted ground

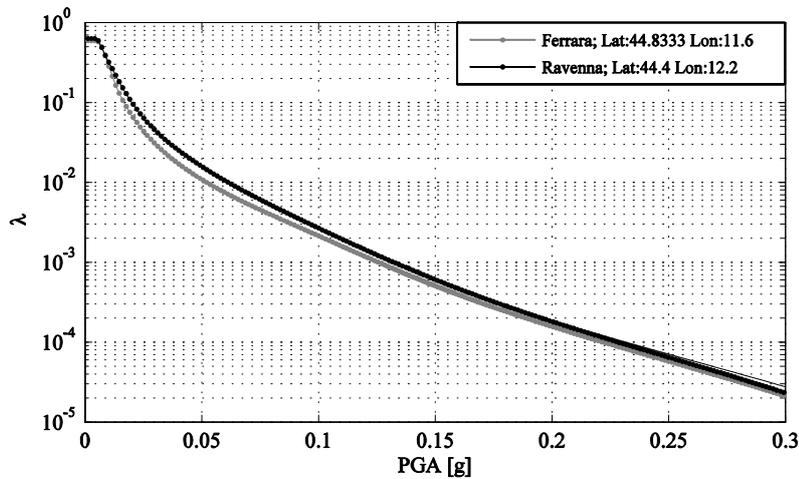


Figure 1.5. Hazard curves for PGA at the sites of Ferrara and Ravenna (Campedel et al., 2008).

motion given a magnitude-distance pair (as well as other relevant parameters). The question that must be addressed when selecting an upper bound on epsilon regards what degree of exceedance on this median will be allowed in the PSHA. This is not a simple question to answer. For a start, we assume that the ground motions associated with a particular magnitude-distance pair are lognormally distributed. This appears to be a very good assumption (Restrepo-Velez and Bommer, 2003) over a wide range of epsilon values ( $\pm 2\epsilon$  at least). However, with any empirical dataset, the number of recorded data points lying in the tails of the distribution is very low, and it is in these regions that only the weakest statistical inference can be made. That said, Restrepo-Velez and Bommer (2003) found that the hypothesis that residuals of ground motion predictive equations are normally distributed could not be rejected when testing using the Anderson-Darling goodness of fit test (Anderson and Darling, 1952). This measure of the goodness of fit is a modification of the Kolmogorov-Smirnov test (Stephens, 1974); modified to better indicate the goodness of fit in the tails of a distribution. Therefore, if the conditions of the Anderson-Darling test are satisfied, one can be confident that the upper bound on epsilon would indeed extend beyond the previously mentioned level of positive two if more data were collected.

The satisfaction of statistical tests is one aspect of the matter, the most important condition that we should seek to meet though relates to the physical limits and the magnitude of ground motions. There must be a physical limit, while conservation of energy is maintained, to the level of ground motion that a particular size event can generate. Exactly what this limit is, is currently an area of active investigation (Bommer et al., 2004; Strasser et al., 2009). In the future, it may be desirable to modify the assumption of lognormality at large values of epsilon. Restrepo-Velez and Bommer

(Restrepo-Velez and Bommer, 2003) propose the upper limit lognormal distribution for this purpose.

### 1.5.2 From hazard to probabilities

The procedure described thus far allows to calculate the rate at which a single target ground-motion is exceeded. If one selects a series of target ground-motion levels and calculate the total rate at which each level is exceeded one may obtain a hazard curve, which is the standard output of a PSHA, i.e. a plot of  $\lambda_{GM}(gm^*)$  against  $gm^*$ . An example of a typical hazard curve is given in Figure 1.5.

Starting from a hazard curve and introducing a model for the occurrence in time of earthquakes it is possible to calculate the probabilities that given ground motion levels are exceeded in a considered time interval. The Poisson process model has been widely used since it was adopted in the initial development of PSHA, but there have been numerous other alternative models proposed.

In order for the occurrence of earthquakes in time to be considered a Poisson process, three conditions must be met (Benjamin and Cornell, 1970):

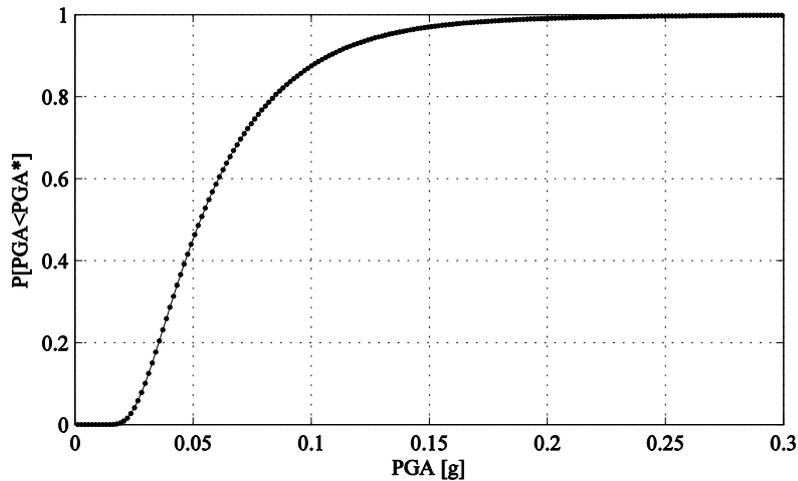
1. Stationarity. The probability of an earthquake event in a short interval of time  $t$  to  $t + \tau$  is approximately  $\lambda\tau$ , for any  $t$ .
2. Non-multiplicity. The probability of two or more earthquakes in a short interval of time is negligible compared to  $\lambda\tau$ .
3. Independence. The number of earthquakes in any interval of time is independent of the number in any other non-overlapping interval of time.

The validity of both the second and third assumptions can be questioned. The condition of non-multiplicity is not met by foreshock or aftershock sequence of earthquakes. Furthermore, the condition of independence does not agree with the earthquake regeneration model of Elastic Rebound Theory. However, the Poisson model, whose probability density function is given in Eq. (1.11)

$$p_x(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!} \quad (1.11)$$

is very simple and easy to implement, requiring only knowledge of the average rate of earthquake occurrence,  $\lambda$ . In Eq. (1.11),  $x$  represents the number of earthquakes and  $t$  time.

The partial violation of the conditions above does not generally pose significant problems to PSHA for the following reasons. The violation of the second condition may



**Figure 1.6. Probability of occurrence of ground motions with different PGA values in 50 years (Campedel et al., 2008).**

be remedied by removing spatially and temporally dependent events from the earthquake catalogues used to estimate the rates of activity for the various seismic sources in the PSHA model, i.e. declustering the seismicity catalogues. The violation of the third condition may present a bigger problem, particularly in the case where evidence of characteristic behaviour is observed, or where the hazard calculation at a site is strongly dependent upon the contribution from a single source (Stafford, 2006). However, Cornell and Winterstein (1988) have pointed out that the sum of non-Poissonian processes may be approximately Poissonian, because the Poisson model is the limit of a sum of point processes (Cornell and Winterstein, 1988; Lomnitz, 1989). Therefore regardless of the temporal dependence of the individual seismic sources within a region, the combination of these sources will yield a regional model of seismic activity that will approach a Poisson model as the number of sources increases. The preponderant use of the Poisson model is therefore justified in the aforementioned case. Adopting the Poisson process as a model for earthquake occurrence, one is readily able to calculate the probability of an event occurring in a given period of time. Defining  $T$  as a random variable representing the time to the first occurrence of an earthquake, the probability that  $T$  exceed some value  $t$  is equal to the probability that no events occur during this interval of time. The probability of no events occurring is found from Eq. (1.11) by setting  $x = 0$ :

$$p_x(x=0) = \frac{(\lambda T)^0 e^{-\lambda T}}{0!} = e^{-\lambda T} \quad (1.12)$$

Obviously, the probability that at least one event occurs during this interval must be the complement of this expression, and is equal to the cumulative distribution function for the occurrence of an earthquake, i.e.  $F_T(T) = 1 - e^{-\lambda T}$ . This latter expression defines an exponential distribution. Figure 1.6 gives an examples of a so calculated cumulative distribution function for the site of Ravenna (Italy) and  $T = 50$  years, calculated by the author (Campedel et al., 2008).

Under the assumption that ground motions may be described by a Poisson distribution over time, the average rate corresponding to the probability,  $P$ , of at least one exceedance within a given time period may be determined using Eq. (1.13):

$$\lambda = \frac{-\log(1-P)}{T} \quad (1.13)$$

For example the 475-year return period used in most seismic design codes throughout the world comes from specifying ground motions having a 10% of being exceeded at least once in any 50 year period (Bommer, 2006). Inserting  $P = 0.1$  and  $T = 50$  years into Eq. (1.13) yields the average annual rate corresponding to this conditions, the reciprocal of which is the return period, that in this case is equal to 475 years. Because Eq. (1.13) is a function of both  $P$  and  $T$  there are an infinitely many combinations of  $P$  and  $T$  that results in a 475-year return period. Once this design criterion is specified, one simply finds the level of ground motion that corresponds to this rate on the hazard curve in order to obtain the design ground-motion.

As previously anticipated, other models have been proposed in the literature, two of the most widely used are the lognormal model and the Brownian passage time model. In order to address the violation of the assumption three of the Poisson process, Nishenko and Buland (1987) have proposed to adopt a lognormal distribution for the inter-arrival times of repeated earthquakes of similar magnitude on individual source. This model has been criticized by Mattheews et al. (2002), who pointed out that for the lognormal model, the hazard rate function, defined as:

$$h(t) = \frac{f_T(t)}{1 - F_T(t)} \quad (1.14)$$

tends to zero for long times. The function in Eq. (1.14) describes the instantaneous failure propensity of the fault under consideration. And according to the currently understanding of the earthquake cycle, the most proper behaviour of the hazard rate function would be that it started at zero (following the occurrence of a major event), and the gradually

increased with time, suggesting that the instantaneous likelihood of failure in turn increases with time. Following a period of increase of the hazard rate function it would then be preferable for the function to approach some asymptotic limit that represents the case in which the stress state around the fault is conducive to failure and that the occurrence of next event can be modelled as a random process. For these reason Matthews et al. (2002) have proposed to adopt the Brownian passage time model, a model derived from a physical basis by adding Brownian perturbation to a steady loading cycle of tectonic stress.

### *1.5.3 Aleatory and epistemic uncertainties*

The PSHA methodology laid out thus far is capable of accounting for all the aleatory variability that exists within the process. However, there is another important component of uncertainty that must also be accounted for – the uncertainty associated with not knowing the applicability of available methods (Bommer, 2003; Abrahamson and Bommer, 2005; Bommer and Abrahamson, 2006). This type of uncertainty is known epistemic uncertainty within the context of PSHA. Aleatory variability and epistemic uncertainty can further be partitioned into modelling and parametric components. These distinctions are not just semantics, each aspect of the overall uncertainty must be treated prudently and each must be approached in a different manner. The logic-tree is the mechanism via which practical application often reveals nuances that require further investigation and many such issues have recently been brought to light as a results of the PEGASOS project (Abrahamson et al., 2002). Aspects such as model selection, model compatibility and the overall sensitivity of PSHA to logic-tree branches for ground motion models have all been addressed (Abrahamson et al., 2002; Bommer et al., 2005; Scherbaum et al., 2005) and have issues associated with how the outputs of the logic-tree are harvested.

The range of ground-motion values corresponding to a given hazard level may vary considerably across fractiles, and as one moves to longer return periods the difference between the mean and median hazard curves may become very large. The first aspect reinforces the importance of taking into account different interpretations of the regional seismo-tectonics as well as different models of approaches to estimating ground motions, while the second aspect demonstrates that one must be clear about how the design ground-motion is to be specified as the results corresponding to the mean hazard and various fractiles may differ considerably.

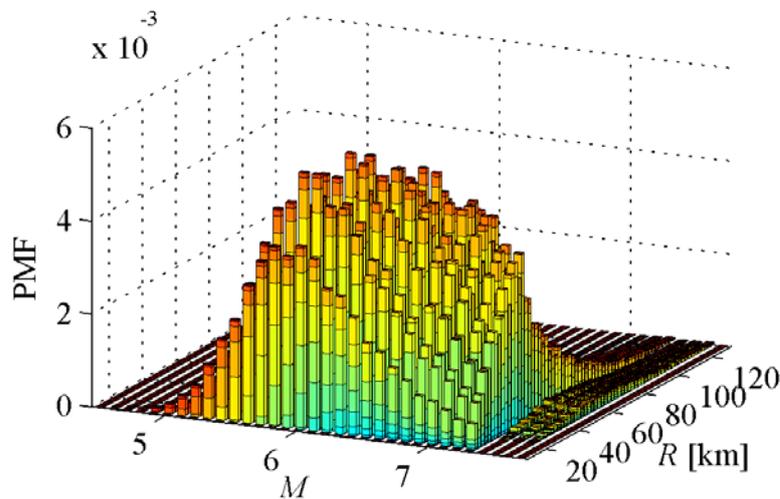


Figure 1.7. Magnitude-distance-epsilon disaggregation for the site of Scalea (CS) (Buratti et al., 2007a).

### 1.5.4 Uniform hazard spectra

The primary output from a PSHA is a suite of hazard curves for response spectral ordinates for different response periods. A design return period is then selected and the response parameter at this return period is determined at each response period and used to construct the elastic response spectrum. A spectrum produced in this way, for which it is known that the return period associated with several response periods is the same, is known as uniform hazard spectrum (UHS) and it is considered an appropriate probabilistic representation of the basic earthquake actions at a particular location. The UHS will often be an envelope of the spectra associated with different sources of seismicity, with short-period ordinates controlled by nearby moderate-magnitude earthquakes and the longer-period part of the spectrum dominated by larger and more distant events. As a consequence, the motion represented by the UHS may not be particularly realistic and this becomes an issue when the motions need to be represented in the form of acceleration time histories. If the only parameter of interest to the engineer is the maximum acceleration that the structure will experience in its fundamental mode of vibration, regardless of the origin of this motion or any other of its features such as duration, then the UHS is a perfectly acceptable format for the representation of the earthquake actions.

### *1.5.5 Disaggregation of probabilistic seismic hazard*

Commonly the results of a Probabilistic Seismic Hazard Analysis (PSHA) are disaggregated in order to show how the resulting hazard at a particular site relates to its constituent components. Usually the integration of the joint probability density function in magnitude distance, and epsilon is consequently approximated in reality by performing a series of simple summations for discrete increments of each variable between the limits of integration. The true value of the integral is recovered in the limit where the increment size of each variable tends to zero. This process of approximation is summarized in Kramer (1996). In its most basic form this disaggregation can be performed with respect to one parameter such as magnitude or distance. In this case the hazard is said to be marginally disaggregated. Typically though, the total hazard is disaggregated with respect to at least two variables; usually magnitude and distance. Because of the numerical procedure adopted to evaluate the integral this is a relatively simple exercise. Only simply stores the contributions to hazard determined for every combination of increment  $s$  between each of the variables before the numerical integration is performed. The contributions to the total hazard for every scenario considered can then be extracted from these stored values.

The representation of the hazard at a site afforded by the disaggregation procedure is very instructive as it takes a step back from the probabilistically framed final outcomes and displays the hazard in terms of the readily understandable parameters. Most people can easily consider the scenario of an earthquake of some magnitude occurring at some distance from them. When the hazard is disaggregated in this manner the result is a suite of magnitude-distance pairs that occur with various probabilities. The identification of the most critical earthquake scenarios is therefore made relatively transparent to the end user of PSHA. There are however some subtleties associated with the disaggregation procedure, such as whether or not the probability mass function, or the probability density function is used to portray the hazard, and how the hazard contributions are assigned to increments of epsilon (McGuire, 1995; Bazzurro, 1998; Bazzurro and Cornell, 1999). An example of a disaggregation histogram, in terms of probability mass function, is given in Figure 1.7, which shows the contributions to the hazard for PGA with a return period of 475 years for the site of Scale (CS) (Buratti et al., 2007a). Different colours represent contributions by different epsilon values.

As well as disaggregating the total hazard with respect to magnitude and distance one should also consider the contributions to the hazard that are made by various values of epsilon. When the hazard disaggregation is performed in terms in terms of the three

primary variables,  $M$ ,  $R$ ,  $e$ , one is able to get a complete picture of where the overall hazard comes from. An additional method that provides even greater clarity is the 4D disaggregation methodology proposed by Bazzurro (Bazzurro, 1998) and Bazzurro and Cornell (Bazzurro and Cornell, 1999) in which the positional contributions are also included by disaggregation the distance, further into latitude and longitude.



## **2 Ground motion selection and scaling issues<sup>1</sup>**

### **2.1 Introduction**

Assessment of seismic demands and their uncertainties necessitates the availability of sets of acceleration time histories that represent the seismic hazard at different return periods, and describe intensity, frequency content, and duration with sufficient comprehensiveness so that central values and measures of dispersion of the demand parameters can be determined with confidence and efficiency. At this time there is no established procedure to select such sets of ground motions. On the other hand, rigorous demand prediction necessitates inelastic time history analysis, which means that records have to be selected (or generated) for the aforementioned purpose. The accepted process is to perform hazard analysis on selected ground-motion parameters, and use the hazard information for record selection and uncertainty propagation. This process implies that the selected ground-motion parameters should be capable of capturing all intensity, frequency content, and duration information that significantly affect the elastic and inelastic response of complex soil-structure systems. No single parameter is ideally suited for this purpose, and, unfortunately, the best choice of parameters depends, sometimes weakly and sometimes strongly, on the structural system and the performance level to be evaluated. This issue is one of the basic challenges of performance-based earthquake engineering (Krawinkler et al., 2003).

### **2.2 Ground motion selection**

Ground motion record selection is considered a critical problem in assessment of structural response based on numerical dynamic analysis. The structural engineer expects the ground motion specialist to select records whose magnitudes, distances, site conditions, and faulting style are representative of the threat to the site. This may be accomplished with the use of disaggregation. In some cases the records are selected to

---

<sup>1</sup> This Chapter has been written with the supervision of Prof. J.J. Bommer and Dr. P.J. Stafford, Department of Civil and Environmental Engineering, Imperial College, London, UK.

provide response spectra that approximate the uniform hazard spectrum or other design response spectrum. The properties (natural period, ductility, etc.) of the structure may or may not be considered in the record selection. The choice of the ground motions may be affected by the interface variable used to measure the intensity of ground motion. This latter is named Intensity Measure (IM) in the PEER Performance Based Earthquake Engineering (PBEE) framework (Porter, 2003). It is extremely important to put the ground motion selection procedures in their context, because different selection criteria must be used depending on the problem under investigation and on the object one likes to achieve.

For example, if a new building is to be designed, the majority of codes prescribe that a suite of ground motions has to be selected such as to be compatible with a Uniform Hazard Spectrum (UHS). These records have to be used to evaluate the mean structural response while the uncertainty on possible ground-motion intensity is already included in the UHS. On the other hand if a probabilistic procedure as the PEER's PBEE (Porter, 2003) is used, the distribution of a structural response parameter (engineering demand parameter) conditional to different levels of ground-motion intensity have to be evaluated. The seismic hazard at the site and the structure are coupled by an intensity measure. The adopted records should be representative of the ground-motion variability for different given values of the intensity measure.

Many different selection procedures have been proposed by different authors in the last years. Following Luco (Luco, 2006) the following objectives can be defined:

1. Evaluation of mean structural response given scenario (magnitude ( $M$ ), distance ( $R$ ), site soil conditions ( $S$ ), fault rupture ( $F$ ), etc.).
2. Evaluation of CDF of structural response given scenario.
3. Evaluation of mean structural response given scenario and IM.
4. Evaluation of distribution of structural response given scenario and IM.
5. Evaluation of mean structural response given UHS
6. Other (e.g. experimental tests).

Tables 1-2 summarize the record selection and scaling methods which will be described in the following.

### *2.2.1 Preselecting records*

Before a record can be considered for selection, it is necessary to check that the properties of the record are such that it is suitable input for dynamic structural analysis. Properties of the record depend on the recording technique of the ground motion and the

**Table 2.1. Summary of the record selection procedure described in this document.**

Procedure	M	R	S	F	$\epsilon$	Dur.	Property matching	Proxy
Chapman 1995								
Shome et al. 1998 and Shome 1999	(x)	(x)						
Bommer et al. 2000	x	x						
Stewart et al. 2001	x	x	x	x	x			
Malhotra 2003	x	x						
Dhakal et al. 2006	x	x					Matching with percentile IDA curves	
Bommer and Acevedo 2004	x	(x)	(x)				UHS	
Baker and Cornell 2005	x	x	(x)		x			
Iervolino and Cornell 2005								
Baker and Cornell 2006b	(x)	(x)	(x)				CMS- $\epsilon$	
Delgado et al. 2006							Acceleration response spectrum	
Watson-Lamprey and Abrahamson 2006b						x	PGA, PGV, Arias intensity (after scaling)	x
Watson-Lamprey and Abrahamson 2006c						x	PGV, $Sa(T_1)$ , $Sa([T_1, T_2])$ (after scaling)	x
Douglas 2006	x	x				x	$S_d(0.1)$ , $S_d(1.0)$ .	
Shantz 2006							Inelastic response surface (after scaling)	x
Trombetti et al. 2006							[PGA, PGV]	
Malhotra 2007	(x)	(x)	(x)			x	[PGA, PGV, PGD]	
ATC-58 Zareian, 2007	x	x						
Geomatrix Consultants 2007	x	x	x				$Sa(T_2) Sa(T_1)$ after scaling.	
ATC-63	x	x	x	x			PGV > 15 cm/sec	
Kircher and Haselton, 2007							PGA > 0.2 g	
Hancock et al. in press	x	x	x				GMPE	
Youngs et al. 2006	x	x	x					
Rathje and Kottke 2007	x	x	x	x			GMPE	
Shome 2007	x	x	x	x			[ $Sa(T_2), Sa(T_3) Sa(T_1)$ ]	
Tothong and Luco 2007								
Zhai and Xie 2007						x	PGA, PGV, PGD, IV, ID, EPA, EPV, demanded yield strength and hysteretic energy	

processing of the data. Most databases provide the processed records only while the raw data is rarely published. All ground motion records contain noise due to imperfections in the recording process (and, for analogue records, the digitizing) and the aim of processing the recorded data is to limit the data to frequency ranges where the signal-to-noise ratio is satisfactory (Boore and Bommer, 2005).

Boore and Bommer (2005) pointed out that the most important processing for all records is the application of low-cut filters to remove the low-frequency parts of the record contaminated by long-period noise. The most important aspect of applying a

**Table 2.2. Summary of the scaling procedures described in this document.**

Procedure	Type of scaling	Reference for scaling
Chapman 1995	linear	PGV
Shome et al. 1998 and Shome 1999	linear	$Sa(T_1)$
Carballo and Cornell 2000	freq-domain	GMPE spectrum
Carballo and Cornell 2000	time-domain	GMPE spectrum
Bommer et al. 2000	linear	UHS
Stewart et al. 2001	linear	UHS. Single period or period range.
Stewart et al. 2001	time-domain	UHS. If a low number of time histories is used.
Malhotra 2003	linear	Part of the UHS
Bommer and Acevedo 2004	linear	Part of the UHS
Baker and Cornell 2005	linear	$Sa(T_1)$
Iervolino and Cornell 2005	linear	$Sa(T_1)$
Baker and Cornell 2006b	linear	$Sa(T_1)$
Delgado et al. 2006		
Watson-Lamprey and Abrahamson 2006b	linear	PGA or PGV or Arias intensity or $Sa([T_1, T_2])$
Watson-Lamprey and Abrahamson 2006c	linear	PGV or $Sa(T_1)$ or $Sa([T_1, T_2])$
Zhai and Xie 2007		
Douglas 2006		
Shantz 2006	linear	Inelastic response surface
Trombetti et al. 2006		
Malhotra 2007	linear	[PGA, PGV, PGD]. Multiple values.
ATC-58 Zareian, 2007	linear	$Sa(T_1)$ . Multiple values (11) defined according to the distribution of $Sa(T_1)$ given by GMPE.
Geomatrix Consultants 2007	linear	$Sa(T_1)$
ATC-63	linear	$Sa(T_1)$
Kircher and Haselton, 2007		
Hancock et al. in press	wavelet	GMPE spectrum (multiple damping ratios).
Youngs et al. 2006	linear	GMPE spectrum (MSE and slope)
Rathje and Kottke 2007	linear	GMPE spectrum. Multiple values.
Shome 2007	linear	$Sa(T_1)$
Tothong and Luco 2007	linear	$Sd_i$ or $IM_{1\&2E}$

low-cut filter is selecting the long-period cut-off, for which a model of the noise is ideally required. The choice is never unambiguous and depends on what is considered to be an acceptable signal-to-noise ratio. From an engineering perspective the most important point is that once the filter frequency is selected, this automatically defines the range of periods over which the data is usable. Akkar and Bommer (2006) suggested that, depending on the type of recording instrument, the magnitude of event and the site class, only spectral displacement ordinates up to 0.65-0.97 times the corner period of the low-cut filter should be trusted. Bazzurro et al. (2004), by studying near-field records, concluded that the more severe is the non-linear behaviour of structures the more important is this issue. These authors also concluded that the value of the high-pass corner frequency,  $f_{HP}$ , has a visible effect on the linear and non-linear response spectrum ratio. Records that are filtered with a larger value of the high-pass corner frequency tend to be more benign if used to compute the non-linear response of oscillators with periods even much lower than the reciprocal of the value of  $f_{HP}$  used. Among the ground-motion parameters considered in their study, only PGD and, to a lesser extent, PGV, were affected by the different aspects of the filtering technique; all the other ground-motion parameters were found to be virtually unaffected by the selection of  $f_{HP}$ . Acausal filters should be preferred to causal filters since they do not distort the phase spectra (Boore and Akkar, 2003). Other characteristics such as the order of the filter are of lesser importance (Bazzurro et al., 2004; Boore and Bommer, 2005).

If ground motions are to be used for bi-directional analysis, the choice of the filter types and parameters are of particular concern: the horizontal components should, be filtered with the same filter parameters (same order and corner periods of the filter); it is also important that the leading zeroes, which should be retained in the filtered record, are of equal length on both components (Boore, 2005). On the other hand, if the vertical component is also included in the analysis, it is not recommended to use the same filter parameters for the vertical as for the horizontal components. Processing all three components with the same filter parameters would lead to loss of low-frequency content of the horizontal components which is valuable to the structural engineer (Boore and Bommer, 2005) for if the same processing is applied to all three components, the filter cut-off will generally be controlled by the vertical component since this will usually have a lower signal-to-noise ratio than the horizontal components, particularly in the long-period range. Similar arguments hold for strongly polarized horizontal components of motion, as may be encountered in near-source recordings, since the stronger

component could be subjected to an unnecessarily severe filter because of the lower signal-to-noise ratio of the fault-parallel component.

Finally, another issue that should be considered is the presence of shifts in the baseline of records; baseline fitting techniques can be used to adjust for reference these shifts in both digitized analog records and digital recordings. Boore and Bommer (2005) pointed out that one advantage of baseline fitting techniques over filters is that the former can enable residual displacements to be recovered although the corrected baseline can rarely be validated by independent observations, especially if displacements are of concern. Bazzurro et al. (2004) pointed out that as far as near-fault ground motions are concerned, the removal or preservation of the residual static displacement has a large impact on structural response and concluded that the ground motions with static residual displacement preserved seem to cause more severe non-linear response of structures with a wide range of natural periods (even much shorter than the typical rise time of the static displacement offset). The authors also expect this effect to be less noticeable on more standard accelerograms recorded where the residual displacements tend to be more limited.

### *2.2.2 Selection based on geophysical parameters*

If scenario-based selection is applied, records are selected which fall in bins around central values of seismic parameters. In most cases, not only one parameter but two or more parameters are used for selection. The three most traditional selection parameters are magnitude ( $M$ ), source-to-site distance ( $R$ ), and site class ( $S$ ). The reason for this is that important characteristics of the record such as frequency content, spectral amplitudes, spectral shape, and duration are correlated to magnitude, distance, and site class. Site classes are commonly based on the shear-wave velocity of the uppermost ground layers. It should be noted that the use of recorded ground motions is based on the implicit assumption that “*provided similar tectonic environments are considered, strong-motion records from one country can be selected and applied in another*” Bommer and Acevedo, 2004.

Reference seismic scenarios can be defined by performing Deterministic Seismic Hazard Analysis (DSHA) or Probabilistic Seismic Hazard Analysis (PSHA) for the site under investigation (See Chapter 2). If the first kind of analysis is performed the design earthquake is fully defined, in terms of magnitude, site-fault distance and nature of surface geology. Usually the 84-percentile ground-motion is considered in this approach (Bommer et al., 2000; Krinitzsky, 2002; Bommer and Acevedo, 2004). If PSHA is adopted, the controlling earthquake scenarios need to be evaluated by disaggregation,

adopting one of the several techniques proposed in literature (Chapman, 1995; McGuire, 1995; Bazzurro and Cornell, 1999). It is worth noting that two slightly different approaches are possible: Bazzurro and Cornell (1999) proposed to disaggregate the hazard conditioned on the spectral acceleration,  $Sa$ , exceeding a given level of interest (i.e.  $Sa \geq s_a^*$ ). McGuire (1995) proposed a slightly different approach that is conditioned on the  $Sa$  equalling the  $Sa$  level of interest (i.e.  $Sa = s_a^*$ ). Depending on the application one of the two approaches can be preferable, for example in assessing structural performance, one is often interested in  $Sa$  equalling a level of interest (10% in 50 years): in this case the latter approach is preferable (Haselton and Deierlein, 2006). Krinitzky (2002) argued that particular care should be taken when disaggregating data from hazard models in which logic trees are used in order to avoid the risk of averaging non-uniform data sets.

In the method proposed by McGuire (1995) a hazard assessment, using all relevant seismic sources, is performed for each spectral response period. Next, a conditional distribution of  $M$ ,  $R$  and  $\epsilon$  (number of logarithmic standard deviations above or below the logarithmic mean from the ground-motion prediction equation used in the analysis) is determined for each source and for each period, where the condition is the matching of the design motion. If one source dominates the hazard for both periods then the distributions will be similar and can be combined to form a composite distribution, the peak of which is used to define the earthquake scenario. If different sources dominate the hazard for different periods then the distributions will be different and a scenario must be defined for each.

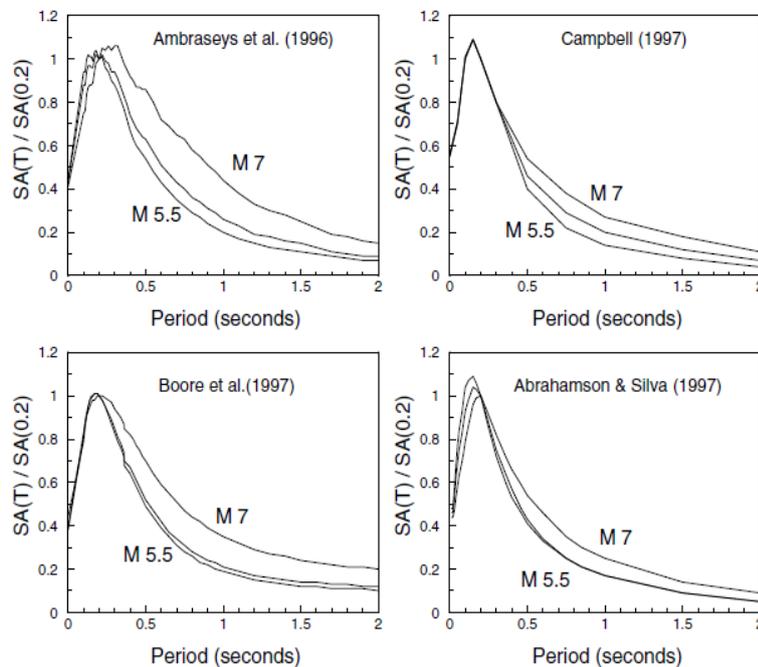
Chapman (1995) defined the hazard-consistent earthquake scenario using the modal event. The underlying probabilistic hazard assessment integrates over all sources, which, the author notes, can sometimes result in the existence of several local maxima, indicating the need to consider multiple design events. The author proposed a procedure to accommodate the difference between the ground motion predicted from the attenuation relationship with the design event and the design ground-motion from the hazard analysis. It is based on counting statistics of a sample of records satisfying the  $M$ - $R$  combination, i.e. a record is chosen which is consistent with the number of standard deviations by which the two values differ. This procedure cannot be applied if only a low number of records compatible with a given  $M$ ,  $R$  scenario can be found, in this case the author suggested to select or synthesize a “*best estimate*” ground motion time-history representative of the modal event, and scale the amplitude of the time-history (Section 2.3.1) such that the PSV response is corresponding to the design expected ground-motion level.

All the aforementioned techniques yield dominant scenarios contributing to the hazard at different parts of the response spectrum, defined by a magnitude, distance and  $\epsilon$ . Once the dominant scenario is defined it is important to decide which parameters should be included in the search and for each parameter how much tolerance should be allowed in the degree of matching between the record and the scenario. The parameter  $\epsilon$ , has not been considered in many ground-motion selection procedures, but recently it was shown (Baker and Cornell, 2005) to be a strong indicator of spectral shape if the spectral acceleration at the natural period of the structure under investigation, is used as intensity measure (Section 2.2.3.3).

Before closing this section it is worth noting that, as Bommer and Acevedo (2004) suggested, if the vertical ground-motion is considered important, its controlling  $M$ - $R$  scenarios may differ from those for the horizontal component.

### 2.2.2.1 Magnitude

There is a strong debate in literature about the importance of magnitude on structural response and hence about the need of taking into account this parameter when selecting records. Many authors demonstrated, by regression analysis, that magnitude has a strong influence on spectral shape and on strong-motion duration. For this reason many



**Figure 2.1. Response spectral shapes (normalised to the ordinate at 0.2 s) for rock sites at 10 km from earthquakes of magnitude 5.5, 6 and 7 using the median values obtained from the following attenuation equations (clockwise from top left ): Ambraseys et al. 1996, Campbell 1997, Abrahamson and Silva 1997 and Boore et al. 1997. After Bommer and Acevedo (2004).**

authors suggested that magnitude should be considered in record selection. Figure 1 shows the dependence of response spectral shapes on magnitude as predicted by different attenuation equations. Bommer et al. (2000) suggested that records should match the reference  $M$ - $R$  pair as closely as possible, stating that it is important that “*the records are from earthquakes with magnitude close to the design value since scaling cannot alter the duration and the spectral shape*”. Stewart et al. (2001) asserted that particular attention should be paid in matching magnitude and suggested that is desirable to use earthquake magnitudes within 0.25 magnitude units of the target magnitude. Bommer and Acevedo (2004) recommended that the search be based on achieving a good match in terms of magnitude and suggested to use earthquakes within 0.2 magnitude units of the target.

Some other authors argued that magnitude and distance matching is not important if records are scaled in amplitude (See Section 2.3), in particular if they are scaled in terms of  $Sa(T_1)$ , i.e. the spectral acceleration at the fundamental frequency of the structure considered. Bazzurro and Cornell (1994) investigated the non-linear response of a simple model of an offshore structure and found that it is not strongly correlated with magnitude, distance and strong-motion duration. Later Bazzurro et al. (1998) stated that “*a careful search of the records that closely match the hypothesized  $M$  and  $R$  scenarios for the site with the aim of computing the structural response of the building is not crucial. Beyond the spectral  $Sa$  level [...]  $M$  and  $R$  play little systematic role in typical non-linear response measures*”. The authors pointed out that  $M$  and  $R$  values may be important only in those cases where near-source directivity effects are expected to be significant. In Shome et al. (1998) and in Shome (1999) it was concluded that provided the records are scaled to match the elastic design spectrum at the fundamental period of the structure considered, then matching the records for the magnitude-distance combination of the design earthquake scenario is not important. However the authors recommend in their conclusions that “*records from roughly the same magnitude*” should be used. Iervolino and Cornell (2005) compared the non-linear response of a suite of model structures to sets of records selected to match a specific magnitude and distance scenario and other sets of records selected arbitrarily. Records in each target set were linearly scaled to the mean spectral acceleration of the set and records in the arbitrary sets were scaled to the mean spectral acceleration of the different target sets. The authors found there is no evidence that suggests that it is necessary to take great care in the selection of records with the respect of these factors. The target sets were defined according to a  $M = 7.0$ ,  $R = 20$  km (closest distance to fault rupture) scenario. Five events with magnitudes ranging from 6.7 to 7.4 were chosen from the PEER ground

motion databank. Six different sets of 10 records (2 for each event) were considered. Almost all the records were in the distance range  $20 \pm 5$  km. As for the arbitrary sets, they contained 10 records too, which had distances spanning from 15 to 50 km and were obtained from events with magnitudes in the range  $6.4 < M < 7.4$ . The method used to infer about importance of matching a  $M$ - $R$  scenario is based on statistical hypothesis testing of the ratios of mean structural response obtained by target and arbitrary sets. It should be noted that the authors adopted small samples and the hypothesis tests are strongly influenced by the high dispersion of the results. Furthermore they supposed the statistical independence of the mean value of structural response obtained by the target and the reference sets while, as the authors stated, the target and the arbitrary sets overlapped each other, i.e. some records were obtained from common events.

Jalayer (2003) investigated structures with very short and very long first-mode periods (a 20-storey structure), and concluded that the structural response is conditionally independent of ground-motion characteristics such as magnitude and source-to-site distance for a given seismic intensity level (e.g.  $Sa(T_1)$ ). This conclusion may justify “*random*” record selection for the site condition and considering the relative position of the site with respect to the major faults around it. The dependence of the response on periods other than  $T_1$  was studied by investigating the (linear) dependence of the residuals of the regression of the structural response (interstorey drift) on the spectral shape factor at a period  $T \neq T_1$ . It was demonstrated, confirming findings of many other authors (e.g. Abrahamson and Silva, 1997), that the shape of the acceleration spectrum depends on the moment magnitude of the ground motion. Furthermore Jalayer found that given  $Sa(T_1)$ , the structural response is positively correlated to spectral shape at a period different than  $T_1$ . For the case of the 20-storey structure, the response was found to be most strongly correlated to spectral shape factors at periods shorter than  $T_1$  and close to the second-mode frequency. For the high-frequency SDOF system, the response was found to be most strongly correlated to shape factors at periods longer than  $T_1$ . For the 20-storey building, the spectral shape factor at  $T_2$  period was negatively correlated to magnitude with a relatively large dispersion around the line fitted to the data. As Jalayer stated the dependence of the response on the shape factor and also the negative correlation between the shape factor and magnitude in the case of a long-period structure like the 20-storey structure may suggest that the response depends on magnitude. Jalayer (2003) found no confirmation of this statement from the results obtained by residual-residual plot for structural displacement and magnitude. These results indicated that given  $Sa(T_1)$ , the displacement-based response for the 20-storey structure is not significantly dependent on magnitude. The author argued that, although

the response might suggest some minor dependence on magnitude for the 20-storey structure, the large variability in the data-points causes the observed dependence not to be (statistically) significant, and that regression analyses that showed dependence of response (given  $T_1$  spectral acceleration) on shape factor and also shape factor on magnitude, do not necessarily imply that there must be a positive (linear) correlation between response (given  $T_1$  spectral acceleration) and magnitude. Jalayer (2003) concluded that there is no evidence in the data, from either the very long or very short period systems he studied, for a dependence of structural response on magnitude. Some inconsistency can be found in the method used by the author to infer the dependence of structural response on magnitude: since a dependence of structural response on spectral shape has been observed and since this latter has been observed to be dependent on magnitude, the only conclusion that can be drawn is that no statistical significance has been found for a linear dependence of the mean structural response (given  $Sa(T_1)$ ) on magnitude.

Krawinkler et al. (2003) and Medina and Krawinkler (2004) investigated the structural response of different frames with the number of stories ranging from 3 to 18 and natural period ranging from 0.3 s to 3.6 s subjected to 80 ground motions recorded during Californian earthquakes of magnitude between 5.8 and 6.9 grouped in 4 different set characterized by different magnitude and distance ranges. It is worth noting that 57 of the aforementioned ground motions were produced by 4 earthquakes. The authors concluded that there appears to be justification to de-emphasize magnitude and distance dependence of seismic demand parameters given  $Sa(T_1)$ . However the authors noted that these conclusions apply only to ground motions outside the near-fault and soft-soil regions. Moreover the results are only limited to magnitude and distance range considered: in fact the median and the dispersion of the response spectra of the accelerograms selected according to the different  $M$  and  $R$  ranges were very similar, i.e. within the range of magnitude and distance covered by the four bins used in this study, the magnitude-distance dependence of the spectral shapes does not have a dominating effect. Dependence of median (or mean) spectra on magnitude and distance has been established through regression analysis in many studies.

As Bommer and Acevedo (2004) pointed out the core of the issue is the degree to which the duration of shaking influences structural demand (Hancock and Bommer, 2007). It is worth noting that many of the aforementioned studies, when making inferences regarding the influence of magnitude matching on structural response adopted non-linear models that did not account for energy-based cyclic deterioration and therefore are not able to give good predictions of the duration effect on structural response.

Furthermore, Hancock and Bommer (2006), pointed out that the influence of duration or effective cycles on the structural response is still debated in the scientific community. Before duration or effective cycles are included as selection criteria one should therefore consider whether duration or effective cycles are expected to affect the structural response. This depends not only on the structure but also on the response parameter which is used to measure the structural response. Cumulative damage parameters often used to assess the damage to reinforced concrete members are, for example, more dependent on the effective number of cycles than the maximum roof displacement of a structure. If it is decided that duration should be considered when selecting records, the choice of definition of duration or effective cycles requires attention. There is a large number of definitions available in the literature for duration (Bommer and Martinez-Pereira, 1999) and effective cycles (Hancock and Bommer, 2006. Bommer et al. 2006) have also shown that the correlation between different duration and effective number of cycles measures is commonly very poor. If a correlation between duration or effective cycles and structural response is expected, the

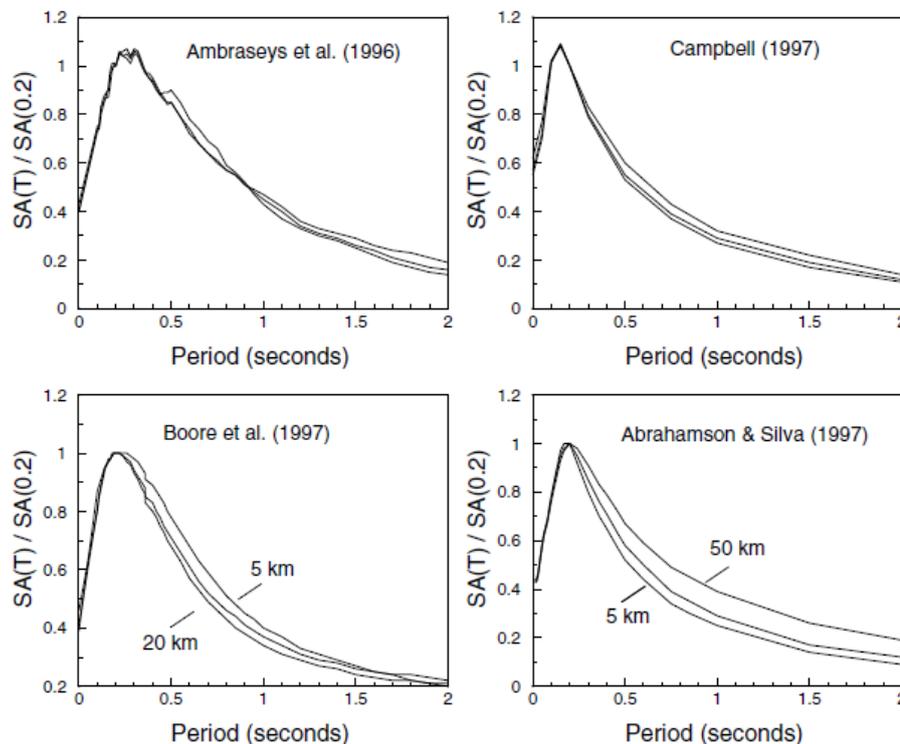


Figure 2.2. Response spectral shapes (normalised to the ordinate at 0.2 s) for rock sites at 5, 20 and 50 km from an earthquake of magnitude 7 using the median values obtained from the following attenuation equations (clockwise from top left): Ambraseys et al. 1996, Campbell 1997, Abrahamson and Silva 1997 and Boore et al. 1997. After Bommer and Acevedo (2004).

choice of an appropriate definition for these parameters is hence important.

#### *2.2.2.2 Distance*

A second parameter that should be considered in record selection is source-to-site distance. As Bommer and Acevedo (2004) suggested by observing attenuation values predicted by different ground-motion prediction equations, the influence of distance on spectral shape is lower if compared to that of magnitude. Figure 2 shows the dependence on distance of response spectral shape for a magnitude 7 earthquake as predicted by different attenuation equations. Furthermore attenuation with distance is almost invariant with period and the ratio of the variation of duration with distance is usually low. Therefore an imperfect correspondence in terms of distance can be fixed by linearly scaling the ground motion record. This draws attention to the issues related to the limits of scaling which are discussed in Section 2.3.1.3.

There are two important exceptions to these conclusions: the first is if records are selected from soft soil sites, since weak, distant motion would not scale linearly for sites closer to the source due to soil non-linearity (Bommer and Acevedo, 2004). The second exception is if near-source rupture directivity effects are to be considered as part of the design scenario. The effect of forward directivity is to produce short-duration motions with high-energy pulses that amplify the spectral ordinates at intermediate or long periods (Somerville et al., 1997b). As Stewart et al. (2001) suggested selected records must be representative of rupture-directivity or fling-step effects.

#### *2.2.2.3 Site classification*

The third parameter that should be included in the search is the site classification, since it strongly influences the ground motion. Soil effects can affect both the amplitude and shape of response spectra. However, specifying a close match for this parameter may not always be feasible since the geotechnical profile has been determined with confidence for a relatively small number of strong-motion recording sites. Furthermore within any site class there can be considerable variation in dynamic response characteristics. Bommer and Acevedo (2004) recommended that if the number of available records matching the magnitude and distance criteria is small, records be considered from sites that are within one site class (e.g., NEHRP or EC8) either side of the classification of the site under consideration. Stewart et al. (2001) suggested that ground motion records should be selected from accelerograph sites having conditions similar to those of the site under consideration. If this is not possible they suggested to scale record by the ratio of soil to rock response spectra for the appropriate magnitude

and distance using an empirical attenuation equation. The authors also suggested that particular care should be taken in selecting records to represent ground motion in basins. Instead of selecting records with a matching site class it is also possible to select ground-motions recorded on rock sites and conduct site response analyses.

#### *2.2.2.4 Other criteria*

The rupture mechanism could also be considered if it is determined as part of the design scenario. There is no definitive evidence for systematic and significant differences between the ground-motions from normal and strike-slip faulting earthquakes, but there is general consensus that reverse-faulting events produce larger amplitudes of motion. Bommer and Acevedo (2004) also suggested that especially when setting up a small suite of real records, the records should not come predominantly from one recording station. A possible exception to this condition would be in the case of the recording station being located very close to the site of interest. Another possible criterion is that any suite of records used in dynamic analyses should not be dominated by accelerograms from a single earthquake event.

### ***2.2.3 Selection based on ground-motion properties and structural response***

#### *2.2.3.1 Selection based on peak values of ground-motion*

Trombetti et al. (2006) proposed a ground-motion selection procedure based on matching both PGA and PGV values obtained by hazard analysis. The main objection that can be raised against this procedure is that the reference PGA and PGV values are calculated by two independent PSHA hence no correlation between the two parameters is considered. This can be an issue if one is interested in estimating the probability of structural failure. Furthermore very few case studies have been reported by the authors: only two SDOF systems with given period. The results considered are the maximum displacement as obtained by incremental dynamic analysis. The authors compared the dispersion in structural response obtained when their selection criterion is used with that obtained using PGA as the intensity measure. Since no information was provided about the period of the SDOF systems considered it is difficult to make any further comment. Nevertheless many authors have shown the inefficiency of PGA as an IM, especially for flexible structures. Finally no information was provided about the possible bias in structural response.

### 2.2.3.2 Selection based on elastic response spectra

Bommer and Ruggeri (2002) and Bommer and Acevedo (2004) performed a review of the guidance in different codes about ground-motion selection and found that guidance given in seismic design codes on how to select appropriate records is usually focused on compatibility with the response spectrum rather than seismological parameters, for the simple reason that the information on seismic source zones and activity rates that underlie zonation maps is not presented and only the uniform hazard spectrum (UHS) is given: “*In current codes, earthquakes are effectively invisible and for this reason the engineer using the code will not easily be able to identify scenario earthquakes*”. Guidance in codes can vary from being purely descriptive to being highly prescriptive. Where specific criteria for selecting records are provided in seismic codes (Bommer and Ruggeri, 2002; Bommer and Acevedo, 2004), they are generally based on the ordinates of the elastic design spectrum although some only specify a match with PGA. Some codes do not specify the relationship between the selected records and the elastic design spectrum, but rather specify that the base shear obtained from dynamic analysis should not be lower than a certain proportion - usually between 0.7 and 0.9 - of that obtained using the equivalent lateral force method, which does not actually help the engineer in making the initial selection of records. Most of the codes that give some guidance on the preparation of suites of acceleration time-series to be used as input to dynamic analyses specify conditions that the records must meet with respect to the ordinates of the elastic design spectrum. Seismic design codes require that it is the average ordinates of the real spectra that have to match the target and not the individual spectra. Bommer and Acevedo (2004) pointed out that this can conceal the issue of the maximum exceedance of the target spectrum by the ordinates from any individual record. Any design spectrum that is derived through the use of PSHA will include the influence of the scatter in the ground-motion prediction equations, which represents the aleatory variability in ground-motion parameters for given combinations of magnitude, distance and site conditions. The strong-motion parameters of the selected records will also display an aleatory variability. Therefore the authors suggested that, if the analysis is performed following the procedure presented in some design codes of selecting three accelerograms, scaling their average ordinates to not fall below the design spectrum, and then using the maximum structural response as the basis for design decisions, the variability in the ground motion is effectively being double counted. A common misconception is to assert that the selected suite of accelerograms should capture the

variability in ground-motion amplitudes, whereas this variability is already fully accounted for in the derivation of the probabilistically-defined response spectrum.

In this document building codes will not be reviewed because most of the conclusions of Bommer and Ruggeri and Bommer and Acevedo are still valid, since as Beyer and Bommer (2007) observed “*guidelines have been copied from codes used in other parts of the world without significant review of the recommendations given therein*”. The only code that will be described in the following is the New Zealand Building, and the drafts produced by the ATC-58 and ATC-62 workgroups will be described.

Different procedures can be used to search for ground motions compatible with the design response spectra. Bommer and Acevedo (2004) proposed to use the matching criteria available in the European Strong Motion Database CD-ROM (Ambraseys et al., 2004). This procedure is based on the root-mean-square difference  $D_{rms}$  between the spectrum of the real record and the reference spectrum:

$$D_{rms} = \frac{1}{N} \sqrt{\sum_{i=1}^N \left( \frac{Sa_0(T_i)}{PGA_0} - \frac{Sa_s(T_i)}{PGA_s} \right)^2} \quad (2.1)$$

where  $N$  is the number of periods at which the spectral shape is specified,  $Sa_0(T_i)$  is the spectral acceleration from the record at period  $T_i$ ,  $Sa_s(T_i)$  is the target spectral acceleration at the same period;  $PGA_0$  and  $PGA_s$  are the peak ground acceleration of the record and the zero-period anchor point of the target spectrum, respectively. The authors sustained that this procedure is superior to matching on the basis of spectrum intensities (area below the response spectrum) in a specified period range, because a good match in that case could easily be obtained with the record having ordinates significantly above the target spectrum at one period and significantly below at another. This procedure limits the maximum deviation of individual peaks or troughs on the spectrum from the target ordinates.

Lately Beyer and Bommer (2007) modified the aforementioned procedure to make allowance for the need of matching the target spectrum over the period range which is of prime interest for the structural analysis. For each record a scale factor  $\alpha$  was determined which minimized the root-mean-square difference  $D_{rms}$  between the scaled geometric mean spectrum,  $Sa_R$ , of the real record and the target spectrum,  $Sa_T$ :

$$D_{rms} = \sqrt{\frac{1}{k-j+1} \sum_{i=j}^k (\alpha Sa_R(T_i) - Sa_T(T_i))^2} \quad (2.2)$$

where  $T_i$  is the  $i$ -th entry of the period vector for which the spectral accelerations are defined;  $j$  and  $k$  define the period interval  $[T_j, T_k]$  for which spectral matching is desired. From all records those with the smallest root-mean-square difference are selected. Another approach to matching which eliminates the influence of the variation of amplitude of the spectral acceleration with period is to normalize the error by the target value at each period:

$$D_{rms} = \sqrt{\frac{1}{k-j+1} \sum_{i=j}^k \left( \frac{\alpha Sa_R(T_i) - Sa_T(T_i)}{Sa_T(T_i)} \right)^2} \quad (2.3)$$

Iervolino et al. (2006a) used a different matching procedure, based on: *i*) the deviation of the average spectrum with respect to the code spectrum ( $\sigma$ ); *ii*) the maximum deviation of a single spectrum within a set with respect to the code spectrum ( $\sigma_{max}$ ); *iii*) records coming from different events within a set; *iv*) small variability of magnitude of events within a set. The maximum deviation of a single spectrum,  $\sigma$ , is defined as:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{Sa_{o,med}(T_i) - Sa_s(T_i)}{Sa_s(T_i)} \right)^2} \quad (2.4)$$

where  $Sa_{o,med}(T_i)$  represents the pseudo-acceleration ordinate related to the mean real spectrum corresponding to the period  $T_i$ , while  $Sa_s(T_i)$  is the value of the spectral ordinate of the code spectrum at the same period and  $N$  is the number of points considered inside the considered range of periods.

Naeim et al. (2004) proposed a genetic algorithms based method to select groups of ground-motion records that in combination are compatible with design response spectra. This method treats any random union of a given number of records (taken from a large database) and the corresponding scaling factors as independent variables and is able to find an optimal solution in terms of distance from a reference response spectrum (the authors used this procedure to select groups of 7 records). The matching function considered by the authors is:

$$\sum_{T=T_0}^{T_n} \left( \sqrt{\frac{\sum_{i=1}^N [S_i Sa_i(T)]^2}{\sum_{i=1}^N S_i^2}} - F_T(T) \right)^2 \quad (2.5)$$

where  $S_i$  is the scaling factor of the record  $i$ ,  $Sa_i(T)$  is the spectral acceleration at period  $T$  of the record  $i$ ,  $F_T$  is the value of the target spectrum at period  $T$  and  $T_0$  and  $T_n$  are respectively the lower and higher period considered. The authors also introduced a penalty function in the optimization problem in order to avoid the solution going below the target spectrum in the period range under consideration.

All the aforementioned procedures were originally proposed to find records compatible with a uniform hazard spectrum, but they can be use with any reference spectrum e.g. with the spectrum predicted by a GMPE or with a conditional mean spectrum (Section 2.2.3.3).

The New Zealand Building Code proposed the following selection and scaling procedure. The stronger component of the ground-motion is considered which is defined as the component with the larger spectral values over the period range  $0.4 - 1.3 T_I$ , where  $T_I$  is the fundamental period of the structure in one specific direction. A record scale factor  $k_I$  is determined which minimizes the error between the target spectrum and the spectrum of the stronger component over the period range of interest. This is done for all records of the suite which will be used for the structural analysis. Finally, a record family scale factor  $k_2$  is determined which ensures that at each period within the period range of interest at least one spectrum of a stronger component scaled with the respective factor  $k_I$  is larger than the target spectrum. The two components of each record are hence scaled with the total factor  $k_I \cdot k_2$ . The stronger component of every record is applied in the direction in which the structural period  $T_I$  was determined. This procedure is repeated for all directions of the structure. Dhakal et al. (2007) investigated the effectiveness of these recommendations and concluded that the intensity measure proposed by the code (based on least squares fit of logarithms of spectral accelerations in the range  $0.4 - 1.3 T_I$ ) is slightly more efficient than the spectral acceleration at the natural period. The authors concluded also that if only three records are used the scale factor  $k_2$  resulted in a overly conservative estimation of structural response.

Malhotra (2003) proposed a different procedure to select strong-motion records for site-specific analysis. Once a site-specific response spectrum is obtained by PSHA the controlling events are identified via disaggregation. Next the duration of the controlling events is estimated using ground-motion prediction equations. The matching with the PSHA spectrum is performed for different ranges of periods and is defined in terms of maximum spectral acceleration  $Sa_{max}$ , control periods  $T_2$  and  $T_3$  as defined by the Newmark and Hall smooth spectrum and strong-motion duration as defined by Trifunac and Brady (1975). Hence for each record in the databank considered the aforementioned

values are calculated by least squares fit between the actual and the smooth spectrum. Different ground motions (e.g. two) are chosen for each controlling event and are linearly scaled to match the UHS in different period ranges

It should be noted that due to the different frequency content of small, local events and large, distant events it might often be the case that the UHS at different response periods is dominated by different earthquake scenarios. Hence, if spectra of real records are scaled to match the UHS over either the entire or a very broad frequency range, the seismic input used for the analysis does not represent a single earthquake scenario. Different warnings have been published in terms of the accuracy of the results that the use of spectrum-compatible accelerograms can offer. Barenberg (1989) suggested that *“a time history derived from a design response spectrum may be dominated by high or low frequency components resulting in a time history that does not represent a realistic earthquake”* and Naeim and Lew (1995) indicated the possibility that spectrum matched records may yield highly conservative results and suggested that *“urge extreme caution in dealing with frequency-domain scaled spectrum compatible time histories in the design of earthquake resistant structures”*. It has become widely accepted in practice that records should not be scaled to the UHS; this limitation applies to all of the codified guidelines cited herein for scaling of records and indeed represents a major shortcoming in current code definitions of seismic actions.

Other procedures have been proposed in literature which do not use the uniform hazard spectrum as reference but the spectra predicted by ground-motion prediction equations. Youngs et al. (2006) worked on the development of a Design Ground Motion Library (DGML), which is an electronic library of selected recorded acceleration time histories defined with the aim of being suitable for use by engineering practitioners for the time history dynamic analysis of various types of facilities in California and other parts of the western United States. The records for this project have been grouped in *M-R-S* bins such that: *i)* the bins do not include combinations of low *M* and/or large *R* that would result in low ground-motion amplitudes for which time history analyses would generally not be performed; *ii)* the bins are wide enough in *M* and *R* that there are significant differences in spectral shape between adjacent bins; *iii)* for near-source bins, the farthest distance of 15 km is selected to capture most records potentially having near-source characteristics. The definition of period ranges appropriate for evaluating response spectral shapes is based on evaluation of typical period ranges of significance for different structure types. The spectral shape of the records in a given *M-R-S* bin and period-range sub-bin is compared to the target median which is determined using ground-motion attenuation equations for the mid-range *M* and *R* of the bin. The records

are first scaled to the level of the smooth median spectra. The scaling criterion is that the spectrum of the record has equal differences above and below the median spectrum over the period range considered. Two measures are used to calculate and evaluate response spectral shapes of records in comparison to the median spectral shape: the first measure is the Mean Squared Error (MSE) of the differences between the median spectrum and the spectrum of the record after scaling. This measure determines the overall “fit” of the spectrum of the record to the target median spectral shape over the period band. The equations for scaling and for determining the MSE of a record are:

$$\text{Scale Factor} = \bar{\varepsilon} = \frac{1}{n} \sum_i \varepsilon(T_i) \quad (2.6)$$

$$MSE = \frac{1}{n} \sum_i (\varepsilon(T_i) - \bar{\varepsilon})^2 \quad (2.7)$$

$$\varepsilon(T_i) = \ln \left[ SA^{\text{target}}(T_i) / SA^{\text{recording}}(T_i) \right] \quad (2.8)$$

where  $n$ , i.e. the number of periods considered, is determined by specifying equally spaced values of  $\ln(T_i)$ . The second measure of the spectral shape of the record relative to the median shape is the “*slope*” of the record spectrum compared to the slope of the median spectrum across the period band. It is determined by a linear regression on the spectral differences with period, between an actual record spectrum and the median spectrum. Spectra with larger slopes (positive or negative) relative to the median spectrum are more skewed relative to the median shape. The equations used to determine the slope of the spectra of the records are given below:

$$\text{Slope} = \frac{\sum_i \left[ \ln(T_i) - \overline{\ln(T)} \right] \times \left[ \varepsilon(T_i) - \bar{\varepsilon} \right]}{\sum_i \left[ \ln(T_i) - \overline{\ln(T)} \right]^2} \quad (2.9)$$

were:

$$\overline{\ln(T)} = \frac{1}{n} \sum_i \ln(T_i) \quad (2.10)$$

The ATC-63 project (Kircher and Haselton, 2007) is working on developing sets of strong- motions (i.e. a Near-Field set and a Far-Field set) appropriate for collapse evaluation of structural systems using Incremental Dynamic Analysis (IDA) methods (Vamvatsikos and Cornell, 2002). The sets have to be generically applicable, i.e.

independent of site, ground motion hazard-level, and structure type (e.g. building fundamental period). Since the set is independent of the building period and site, the set was selected without regard to spectral shape or  $\varepsilon$ . To account for the effects that  $\varepsilon$  has on the collapse fragility, the authors propose to use the post-processing method defined by Haselton (2006). The selection procedure is based on the following criteria: *i*) only records with  $M > 6.5$  are considered; *ii*) ground-motions have to be generated by earthquakes with strike-slip and reverse (thrust) mechanisms; *iii*) recording station located on soft-rock (Site Class C) or stiff soil (Site Class D); *iv*)  $R > 10$  km; *v*) not more than two records per event; *vi*) peak ground acceleration  $> 0.2g$ ; *vii*) peak ground velocity  $> 15$  cm/sec; *viii*) lowest useable frequency  $< 0.25$  Hz; *ix*) recording station in free-field or ground floor of a small building. A scaling procedure to model ground-motion variability is defined: first, individual records (of a given set) are “normalized” by their respective peak ground velocities and factored such that median spectral acceleration of the normalized record set is the same as the median spectral acceleration of the unscaled record set. The second step is to “anchor” the normalized record set to have the target median intensity. Each record in the set is factored by the same amount such that median spectral acceleration ( $S_a$ ) of the record set matches the target at the fundamental period ( $T_f$ ) of the structure.

Douglas (2006) proposed a ground-motion selection procedure based on two steps: in a first phase records are selected according to a seismic hazard scenario, defined in terms of magnitude and distance and focal depth, then, in the second step, the selection procedure proceeds in terms of strong-motion parameters adopting a two-level factorial design, i.e. for each strong-motion parameter accelerograms are selected to fall within two intervals, corresponding to an high and a low value of the parameter considered respectively. The parameters the author considered are two spectral displacements  $Sd(0.1\text{ s})$ ,  $Sd(1.0\text{ s})$  and the ground-motion duration according to Trifunac and Brady (1975).

Finally Delgado et al. (2006) proposed a different ground-motion selection criterion based on cluster analyses. Starting from 20 records selected from a wide range of  $M$  and  $R$ , the authors chose the most homogeneous subset by cluster analysis. Since no reference scenario (spectrum) was defined this technique allows only to indentify a set of similar ground motions in terms of response spectrum, and is probably very sensitive to the initial sample adopted.

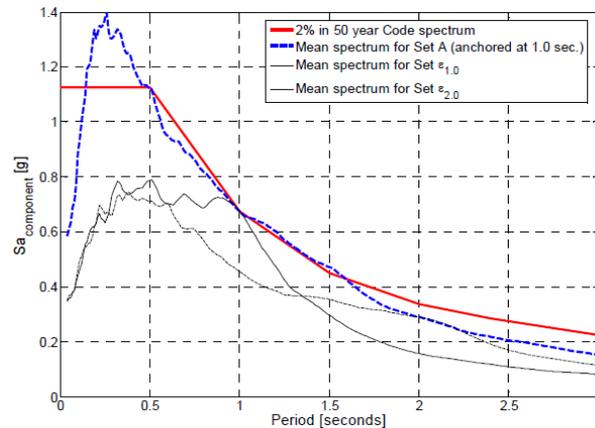


Figure 2.3. Comparison of spectral mean spectral shape of sets of ground-motions selected with: *a*) no regard for  $\epsilon$  (blue line); *b*) accounting for  $\epsilon(1 \text{ s}) = 2$  and *c*) accounting for  $\epsilon(2 \text{ s}) = 2$ . After Haselton and Baker 2006.

### 2.2.3.3 Spectral shape and epsilon

It has long been known that spectral shape has important effects on structural response. This is especially true when higher mode effects are important or when the building is significantly damaged, causing the effective fundamental period to elongate. Baker and Cornell (2005) found that when ground motions are scaled by  $Sa(T_1)$ , the shape of the uniform hazard spectrum is often inappropriate and can lead to extremely conservative predictions of structural responses. This study showed that epsilon, when  $Sa(T_1)$  is used as intensity measure, is an indicator of spectral shape, i.e. is an indicator of the presence of “peaks” or “valleys” in a response spectrum. This can be derived considering the conditional mean response spectrum,  $\mu_{\ln(Sa(T_2))|\ln(Sa(T_1))}$  and its conditional standard deviation. This was found to be especially true for rare ground-motions in coastal California, such as a motion with 2% PE in 50 years (Baker, 2005; Baker and Cornell, 2005; Baker and Cornell, 2006b). This unique spectral shape comes from the fact that  $Sa(T_1)$  is used to define the ground-motion hazard. If the hazard was based on another ground-motion intensity measure such as inelastic spectral displacement (Tothong and Cornell, 2007) or an average spectral-acceleration over a range of periods, this would cause the expected spectral shape to be less peaked for rare ground-motions (Baker and Cornell, 2006b). Thus, there are alternative ways to address the spectral shape issue, either by adjusting the selection and scaling of ground motions or by modifying the definition of ground-motion hazard.

Recent studies have verified that a positive  $\epsilon$  value normally corresponds to a peaked spectral shape. To illustrate this, Figure 2.3 compares the mean spectral shape of three

ground-motion sets (Haselton and Baker, 2006): *a*) a set selected without regard to  $\varepsilon$ , *b*) a set selected to have  $\varepsilon(1\text{ s}) = +2$ , and *c*) a set selected to have  $\varepsilon(2\text{ s}) = +2$ . These records are scaled such that the mean  $Sa(1\text{ s})$  is equal for sets (*a*) and (*b*) and the mean  $Sa(2\text{ s})$  is equal for sets (*a*) and (*c*). This shows that the spectral shapes are distinctly different when the records are selected with or without regard to  $\varepsilon$ . When the records are selected to have positive  $\varepsilon$  values at a specified period, the spectra tend to have a peak at that period. This shape is very different from a standard uniform hazard spectral shape. This makes intuitive sense, because if a ground motion has a much larger than expected spectral acceleration at one period (i.e. high positive  $\varepsilon$ ), then it is unlikely that the spectral accelerations at all other periods are also similarly large.

Furthermore selecting ground motions with proper spectral shape (proper  $\varepsilon$ ) has been shown to significantly increase collapse capacity predictions. This difference in collapse capacity can be explained by comparing the spectral shapes of the aforementioned sets. For example, if the building period is 1.0 second and the ground-motion records are scaled to a common value of  $Sa(1\text{ s})$ , the spectral values of the (*b*) set are smaller for  $Sa(T > 1\text{ s})$  (i.e. the spectral values that are important when the building is damaged and the effective period elongates) and  $Sa(T < 1\text{ s})$  (i.e., the spectral values that are important for higher mode effects).

Baker and Cornell (2006b) studied the effects of various ground-motion properties on the collapse capacity of a seven-storey non-ductile reinforced concrete frame building located in Van Nuys, California with a fundamental period  $T_1$  of 0.8 seconds. They found that the mean collapse capacity increased by a factor of 1.7 when an  $\varepsilon(0.8\text{ s}) = 2.0$  ground-motion set was used in place of a set selected without regard to epsilon (which had mean  $\varepsilon(0.8\text{ s}) = 0.2$ ). Goulet et al. (2006) studied the collapse safety of a modern four-storey reinforced concrete frame building with a period of  $T_1 = 1.0$  second. They compared the collapse capacities for a ground-motion set selected to have a mean  $\varepsilon(1.0\text{ s}) = 1.4$  and a set selected without regard to epsilon (which had a mean  $\varepsilon(1.0\text{ s}) = 0.4$ ). They found that the set selected considering  $\varepsilon$  resulted in a 1.3 to 1.7 times larger mean collapse capacity. Haselton and Baker (2006) used a ductile single-degree-of-freedom oscillator, with a period of  $T_1 = 1.0$  seconds, to demonstrate that a  $\varepsilon(1.0\text{ s}) = 2.0$  ground-motion set resulted in a 1.8 times larger mean collapse capacity as compared to using a ground-motion set selected without regard to epsilon.

For all the aforementioned reasons Backer and Cornell (2005) suggested that epsilon should be considered in the prediction of the annual rate of exceedence of a given hazard when  $Sa(T_1)$  is used as the intensity measure. This can be done either using a vector intensity measure (Section 2.3.1.1.2) or by considering carefully epsilon values

when selecting ground-motions by relaxing the desire of matching distance and magnitude. These considerations about the relationship between epsilon and spectral shape are not appropriate for pulse-type ground motions (Tothong and Cornell, 2007).

#### 2.2.3.4 Procedures based on conditional mean spectrum

Lately Baker and Cornell (2006b) proposed a new ground-motion selection criterion extending their previous studies on epsilon influence on structural response and defining a Conditional Mean Spectrum given  $M$ ,  $R$  and  $\varepsilon$  (CMS- $\varepsilon$ ). As the name suggests a conditional mean spectrum is a spectrum conditional to a given level of  $Sa(T_1)$  taking into account the correlation among spectral accelerations at different periods. The first step required to evaluate the CMS- $\varepsilon$  is performing a PSHA to find the  $Sa(T_1)$  value corresponding to the target probability of exceedance ( $Sa(T_1)^*$ ) at the site under consideration. Disaggregation is then used to find the mean of the  $M$ ,  $R$  and  $\varepsilon$  values ( $\bar{M}$ ,  $\bar{R}$ ,  $\bar{\varepsilon}$ ) that cause occurrence of the  $Sa(T_1)^*$  level.  $\bar{M}$  and  $\bar{R}$  determine the means and standard deviations of response spectral values for all periods via ground-motion prediction models, and  $\bar{\varepsilon}$  the number of standard deviations away from the mean ground-motion at the first-mode period,  $T_1$ . The conditional mean value of the target response spectrum based on  $\bar{M}$ ,  $\bar{R}$  and  $\bar{\varepsilon}$  is computed using the equation:

$$\mu_{\ln Sa(T_2) | \ln Sa(T_1) = Sa(T_1)^*} = \mu_{\ln Sa}(\bar{M}, \bar{R}, T_2) + \sigma_{\ln Sa}(\bar{M}, T_2) \rho_{\ln Sa(T_1), \ln Sa(T_2)} \cdot \bar{\varepsilon}(T_1) \quad (2.11)$$

The terms  $\mu_{\ln Sa}(\bar{M}, \bar{R}, T_2)$  and  $\sigma_{\ln Sa}(\bar{M}, T_2)$  are the marginal mean and standard deviation of  $\ln(Sa(T_2))$ , respectively, and are obtained from a ground-motion prediction equation. The term  $\rho$  is the correlation factor between  $\ln(Sa(T_1))$  and  $\ln(Sa(T_2))$ . This latter can be predicted by regression on empirical observations (Baker, 2005):

$$\rho_{\ln Sa(T_1), \ln Sa(T_2)} = 1 - \cos\left(\frac{\pi}{2} - \left(0.359 + 0.163 I_{I(T_{min} < 0.189)} \ln \frac{T_{min}}{0.189}\right) \ln \frac{T_{max}}{T_{min}}\right) \quad (2.12)$$

where  $T_{min}$  and  $T_{max}$  are the smallest and largest value of  $T_1$  and  $T_2$  respectively and  $I_{I(T_{min} < 0.189)}$  is an indicator function equal to 1 if  $T_{min} < 0.189$  and equal to 0 otherwise. As the authors pointed out using the aforementioned equations the conditional variances of the response spectrum could be easily evaluated, and perhaps it could be used to obtain a criterion to select ground motion to obtain information on the distribution of structural response rather than on its mean value only.

As Baker and Cornell (2006b) pointed out, Eq. (2.11) is, an approximation obtained by substituting the mean values  $\bar{M}$ ,  $\bar{R}$  and  $\bar{\varepsilon}$  for the random values of  $M$ ,  $R$  and  $\varepsilon$

obtained from disaggregation. When the ground-motion hazard is dominated by a single magnitude and distance scenario, the approximation is nearly exact, otherwise  $\bar{M}$ ,  $\bar{R}$  values from disaggregation can correspond to values that do not occur from any fault in the area surrounding the site (Bazzurro and Cornell, 1999), causing some to express concern about using  $\bar{M}$ ,  $\bar{R}$  as target values for record selection. Baker (2005) argued that concern should be lessened with regard to the CMS- $\epsilon$  procedure, in fact  $\bar{M}$  and  $\bar{R}$  are used only to identify a target response spectrum distribution and not to identify target  $M$  and  $R$  values. However Baker (2005) derived an exact formulation for the conditional mean spectrum that takes into account the contributions to hazard given by different faults. Further, when substituting  $\bar{M}$ ,  $\bar{R}$  and  $\bar{\epsilon}$  into Eq. (2.11), one does not necessarily obtain exactly the target  $Sa(T_1)$  value back again. This can be addressed by re-assigning  $\bar{\epsilon}$  to the  $\epsilon$  value that results in a prediction of the  $Sa(T_1)$  target value; the modification will be small and this is consistent with the treatment of  $\epsilon$  by McGuire (1995).

Baker and Cornell (2006b) compared structural response (interstorey drift) produced by records selected according to different criteria: *i*) arbitrarily selection; *ii*)  $M$  and  $R$  matching with no attempt matching  $\epsilon$ ; *iii*) matching  $\epsilon$  with no attempt matching  $M$  and  $R$ ; *iv*) matching the conditional mean response spectrum given  $\bar{M}$ ,  $\bar{R}$  and  $\bar{\epsilon}$ ; and found that records selected according to the first two criteria produce biased results if  $Sa(T_1)$  is used as intensity measure and that the CMS- $\epsilon$  selected records produced lower dispersion in structural response than the other sets of records. This reduces the number of records required to estimate the mean structural response but “*the dispersion is somewhat artificially suppressed [...] this may be problematic for probabilistic drift hazard assessments, where accurate estimates of both the mean and dispersion of response are needed, although no significant errors were observed in the assessments performed here*”.

The authors also suggested that if the structural response parameter of interest is sensitive to spectral-acceleration values at multiple periods, then perhaps this specific peaked spectrum should not be of primary concern. In that case an IM which averages spectral acceleration values over a range of periods might be a better indicator of structural response as suggested by many other authors. In particular Baker and Cornell proposed to adopt the geometric mean of the spectral values in a given range of period and derived specific expressions for the conditional mean spectrum. This latter intensity measure has been recently investigated by Bianchini et al. (in preparation) (Section 2.3.1.1.1).

Recently Shome (2007) proposed a ground-motion selection procedure with the aim of obtaining a set of ground-motion records for non-linear structural dynamic analysis that will result in an accurate estimate of the cumulative distribution function (CDF) and the median of the Engineering Demand Parameter(s) (EDP) of interest for a given structure, earthquake magnitude ( $M$ ), source-to-site distance ( $R$ ), site classification ( $S$ ), style of faulting ( $F$ ), and a given spectrum. The first step of the procedure requires to select records from the  $M$ - $R$ - $S$ - $F$  bin that is consistent with the given scenario, then records are scaled to the spectral acceleration at the first-mode period ( $T_1$ ) of the structure for the given spectrum. A period range such that the non-linear response of the structure is mostly affected by the periods of the ground-motion within the specified range must be selected. The author proposed to adopt the interval  $[T_2, T_3]=[0.5T_1, 2.0T_1]$ . Once the aforementioned range is defined the procedure requires to calculate the conditional joint distribution,  $[Sa(T_2), Sa(T_3)| Sa(T_1)=sa^*]$  for the bin of records. Since the joint distribution of  $[Sa(T_1), Sa(T_2), Sa(T_3)]$  is lognormal, the conditional joint distribution will also be lognormal. This can be done using the correlation coefficients of spectral accelerations at two different periods developed by Baker and Cornell (2006a). Any ground-motion record falls within a region enclosing a percentage of the multivariate normal population  $[Sa(T_2), Sa(T_3)| Sa(T_1) = sa^*]$ . The estimation of the percentages of the population can be obtained by a quadratic form  $Q$  depending on the covariance of the conditional joint distribution. The values of the aforementioned quadratic form can be shown to be distributed as a chi-square distribution with 2 degrees of freedom. The  $Q$ -value of records will indicate if the record falls inside a certain percentage of the lognormal population of  $[Sa(T_2), Sa(T_3)| Sa(T_1)=sa^*]$ . Therefore, the author suggested using  $Q$ -values of the records in the bin as a ranking criterion: the records that have low  $Q$ -values are close to the target spectrum in the specified period range therefore they can be used to estimate the mean EDP for the given spectrum. While a random selection from the bin of records so that the selected records follow the chi-square distribution is proposed to be used to estimate the distribution of EDP for the given spectrum. Some concerns can be raised against this procedure: even if an interval of periods of interest is defined only the spectral values at the boundaries of the interval are used, hence the results are probably very sensible to local fluctuations of the response spectrum. Perhaps a ranking criterion based on matching the conditional mean spectrum over the entire interval could perform better.

### 2.2.3.5 Selection based on non-linear structural response

Dhakal (2006) proposed a procedure to select ground motions for pseudodynamic tests. The method proposed uses results of Incremental Dynamic Analysis to select ground motions. The records are scaled in terms of PGA, although the authors suggested that using  $Sa(T_1)$  could bring better results. In a first stage a suite of ground motions is selected according to a magnitude and distance scenario. Next the obtained records are used to perform an Incremental Dynamic Analysis on the structure that will be tested. Once the 50<sup>th</sup> and 90<sup>th</sup> percentile IDA curves, in terms of displacements, have been evaluated the records with the IDA curves closer to the two aforementioned ones at give value of PGA are selected. This method can have some drawbacks: first of all the percentile curves are evaluated with a small suite of records, hence results can be suite-dependant and as the authors admit PGA is not a good performing IM.

Watson-Lamprey and Abrahamson (2006b) proposed a procedure to select ground motions for slope stability analysis based on using Newmark displacements (Newmark, 1965) as a proxy for more complex non-linear behaviours. The basic observation of the authors is that “*it is difficult to anticipate which time series will give an average response to a non-linear system*” even if spectral shape is considered in the selection together with magnitude distance and site condition. Therefore the authors developed a prediction equation for Newmark displacement as a function of four ground-motion parameters: PGA,  $Sa$ ,  $a_{rms}$  (root mean square acceleration) and ground-motion duration as defined by Trifunac and Brady (1975). The authors showed that Newmark displacement can be modeled with a relatively small variability. Watson-Lamprey and Abrahamson (2006b) concluded that if the accelerograms are selected such that after scaling, the values of the aforementioned time-history parameters lead to a median Newmark displacement similar to that expected for the design event, then that time-history can be expected to give a near average response for a more complicated slope deformation evaluation. This procedure allows ranking of different ground-motions in order to find those best matching the reference values of the search parameters. Some concerns can be raised against the ground-motion records adopted, in fact the authors used all the records in the PEER database with the only exception of records from subduction zones, from the North-eastern California and from non-free-field stations. Post-processing of ground-motion records was not taken in consideration as a selection criterion. Furthermore each record was linearly scaled with factors ranging from 0.5 to 20. It is not clear from the paper how the presence of bias in the predictive equation was checked, in fact even if the authors stated that the model for median displacements is

unbiased for scale factors up to 20, the residuals of the regression are clearly linearly dependant on the scale factor.

Lately, Watson-Lamprey and Abrahamson (2006c) extended their previous procedure and proposed a criterion to select input ground-motions for building assessment. They proposed to predict non-linear structural response using a proxy bilinear SDOF system. The authors developed a prediction equation for the inelastic displacement as a function of the following properties of the ground-motion:  $Sa(2T_1)/Sa(T_1)$ ,  $a_{rms}$ , PGV and ground-motion duration as defined by Trifunac and Brady (1975). The aim of this selection procedure, as described above, is the identification of time-histories whose parameters (i.e.  $Sa(2T_1)/Sa(T_1)$ ,  $a_{rms}$ , PGV) are similar, after scaling, to those giving the design event when substituted in the structural-response predictive equation.

Shantz (2006) proposed a similar procedure for record selection and scaling based on the Normalized Inelastic Displacement Demand, by defining a target displacement surface (defined for a SDOF structure) for peak inelastic displacement over a range of periods and ductilities. Records that satisfy broad magnitude or duration criteria are individually scaled to provide the best match to the target surface. Shantz (2006), by studying the response of a simple bridge, showed that records selected with the criterion he proposed produced a sensible reduction in dispersion of structural response if compared to that produced by record selected according to a  $M, R$  scenario or scaled to the same PGA, without introducing bias.

Cornell, Luco and Tothong (Luco, 2002; Luco and Cornell, 2007; Tothong and Cornell, 2007) proposed an approach to ground-motion selection for Probabilistic Seismic Demand Analysis completely based on the intensity measures. If these latter satisfy some basic properties, which will be described in Section 2.3.1.1.3, the authors stated that the selection can be performed with no regards for magnitude, distance, site soil conditions and fault rupture mechanism. A possible issue in this kind of approach is that if magnitude is neglected in the selection procedure no constrains on ground-motion duration are defined. Moreover since many authors (Iervolino et al., 2006b; Hancock and Bommer, 2007) showed that peak structural displacements are not too much affected by duration, the aforementioned intensity measures themselves are not able to represent duration effects. Finally it is worth noting that all the procedures based on non-linear structural response use simple bilinear SDOF systems as a proxy of more complex structural behaviours. The parameters of these models are usually estimated using non-linear static procedures, therefore the accuracy of the latter is critical for the selection procedures described in this section to work properly. This could be an issue for strongly irregular structures and bi-directional problems.

Zhai and Xie (2007) proposed the concept of the most unfavourable real seismic design ground-motion and selected two groups of input design ground-motions (for different soil conditions): the first group is the most unfavourable real seismic design ground-motions, which Zhai and Xie proposed to use for seismic analysis of very important structures; the second group is the real input design ground-motions with medium damage potential so, according to Zhai and Xie, adequate for seismic analysis of structures located in areas with low or moderate seismicity. The severity of the strong motions was evaluated investigating the response of non-linear SDOF systems. This approach does not allow to include site specific hazard information and to perform structural reliability assessment.

## 2.3 Scaling

Whether records are selected by performing searches in terms of seismological and geophysical parameters, or according to one of the procedure described in the previous section, there will generally be a requirement to ensure that the records conform to some specified level of agreement with the one or more of the ordinates of a reference response spectrum; the latter can be a Uniform Hazard Spectrum, a spectrum obtained by a ground-motion prediction equation or a Conditional Mean Spectrum. The most widely used scaling technique is linear amplitude scaling. According to this technique an accelerogram is simply multiplied by a constant; therefore neither the frequency content nor the duration of the ground motion are modified. Both different definitions of matching and different ground-motion parameters have been proposed in the literature to scale ground motions.

Besides codes, many probabilistic procedures (Porter, 2003) require to evaluate the distribution of a structural response parameter for different given intensities of the ground motion (IM-based procedures). Those approaches are based on the assumption that the chosen IM is a good predictor of the ground-motion damage potential on structures. The variability related to the IM is included in calculations with a fully probabilistic approach performing convolution of hazard and structural fragility. In this case the suites of ground motions have to be representative of the conditional variability of the ground motion *given* a value of the intensity measure (Baker, 2005; Baker and Cornell, 2006b). Most of the probabilistic procedures are based on Incremental Dynamic Analysis (Vamvatsikos and Cornell, 2002) and implicitly assume that the same suite of ground motions can be representative of different intensity levels if properly scaled. More advanced approaches consider different sets of ground-motion

records for different intensity levels (Baker, 2005; Baker and Cornell, 2006b). Finally some procedures try to correct results obtained by scaling the same set of records to different intensity levels by introducing vector intensity measures or by careful post-processing of the results (Haselton and Deierlein, 2006; Baker, 2007c).

Other techniques have the aim of obtaining Spectrum Compatible Records, also known as Spectrum Matched Records. These are artificially generated or edited time-histories of ground acceleration whose response spectral shapes are equal, within a prescribed tolerance, to a predetermined target spectrum. The use of spectrum matched records for the estimation of non-linear structural response has become increasingly widespread in the past 15 years. Its origins arise from the inability of traditional response spectral analyses to be used to estimate maximum responses of non-linear systems, for which a time-integration scheme was deemed more appropriate (Preumont, 1984). The use of spectrum matched accelerograms as surrogates for actual recorded ground motion is attractive for multiple reasons. It is generally accepted that they are able to produce results that present relatively lower dispersion, such that they can more efficiently allow the estimation of seismic demands. This is an important benefit, especially for non-linear analyses that can be highly computer intensive. The iterative generation of Spectrum Matched Records can be done in many ways, but the different schemes can be separated into two major groups: frequency-domain-based, and time-domain-based compatibilization or matching.

Finally, it is worth noting that all the aforementioned scaling techniques do not change the duration on the ground-motion records. Scaling the time axis of an accelerogram can increase or decrease the significant duration, which might be acceptable to compensate for small changes associated with distance, but not to compensate for any mismatch in magnitude because for that it would also be necessary to change the number of cycles of motion. Scaling the time axis of a record changes not only the duration of the motion but also the frequency content of the record over the entire period range. Bommer and Acevedo 2004 strongly recommended that this procedure is to be used with caution.

### ***2.3.1 Linear amplitude scaling***

The most widely used scaling technique is linear amplitude scaling. According to this technique a ground-motion acceleration is simply multiplied by a constant; therefore neither the frequency content or the duration of the ground motion are modified. In Section 2.3.1.1 linear scaling procedures proposed by various authors will be discussed. Some intuitive concerns can be raised about record scaling: namely, that low intensity ground motions have different frequency content than rare or extreme ground motions.

Han and Wen (1994), for example, speculated that “*scaling an earthquake to attain a target damage level of different intensity is questionable since scaling a ground-motion does not account for variations in ground-motion characteristics (e.g., frequency content) which change with intensity*” and Bazzurro et al. (1998) stated that “*a  $M = 5$  record scaled to match the PGA of a  $M = 7$  record will certainly be deficient in the frequency content below 1 Hz*”. Those concerns will be discussed in Section 2.3.1.3.

Recently many authors (Luco, 2002; Giovenale et al., 2004; Luco and Cornell, 2007; Tothong and Cornell, 2007) tried to define a common framework to compare different measures of the intensity of ground motions. Their efforts were focused on probabilistic based procedures. The aim of these approaches is estimating the annual frequencies of exceeding specific limit states. Adopting the PEER framework the latter can be calculated as:

$$\lambda(\text{LS}) = \iint G[\text{LS} | \text{DM}] dG[\text{DM} | \text{IM}] d\lambda(\text{IM}) \quad (2.13)$$

where  $\lambda(\text{LS})$  is the annual frequency of exceeding the limit state LS,  $G[\text{LS} | \text{DM}]$  is the conditional probability of exceeding LS given the damage measure DM,  $G[\text{DM} | \text{IM}]$  is the conditional probability of exceeding DM given the intensity measure IM, and  $\lambda(\text{IM})$ , known as the hazard curve, is the mean annual frequency of occurrence of seismic action with intensities higher than IM at a particular site. The aforementioned authors suggested that the following properties should be considered when assessing the performances of intensity measures:

- efficiency;
- sufficiency;
- hazard computability;
- scaling robustness.

A sufficient intensity measure is defined as one that yields DM conditionally independent, given IM, of earthquake magnitude ( $M$ ), source-to-site distance ( $R$ ) and any other ground-motion characteristics that can affect the structural response (e.g. rupture mechanism, soil type, etc.). Adopting a sufficient IM avoids potentially biased evaluation of  $\lambda(\text{LS})$ , because in this case the function  $G[\text{DM} | \text{IM}]$  is effectively independent from the ground-motion characteristics ( $M$ ,  $R$ , etc.) of the records selected. It also leads to a simpler evaluation, because its use allows the choice of the records to be free from almost any constraints on the values of the ground-motion characteristics. Moreover, adopting a sufficient IM legitimizes the operation of scaling the accelerograms. An efficient IM is one that results in a relatively small variability of DM

given IM. Since  $G[DM | IM]$  in Eq. (2.13) is commonly evaluated by a number of non-linear dynamic analyses the use of an efficient IM leads to a comparatively smaller dispersion in the results. The benefits of this reduction will in general be observed in a narrower confidence interval for the conditional median DM value for a given IM level or, from a different perspective, in a smaller number of non-linear dynamic analyses needed to obtain an acceptable confidence interval. The efficiency and sufficiency of an IM are both criteria that can be quantified via regression analysis with demand measure results (e.g., drift responses) from non-linear dynamic analysis of a structure. The hazard computability can be defined as the effort required in order to determine the hazard curve in terms of the proposed IM. Since the intensity measure (IM) is the link between seismic hazard and structural analysis, the choice of the IM to be used in the non-linear dynamic analysis should keep in consideration the effort required to calculate hazard curves in terms of that variable. Finally the scaling robustness can be defined in terms of the bias induced in structural response when the IM is used to scale ground motions, i.e. the responses for records scaled to different factors but to the same resulting IM level should not show a trend in responses versus scale factors.

### *2.3.1.1 Ground-motion properties (Intensity measures)*

Many authors have investigated ground-motion amplitude scaling. Most of the works that can be found in literature are focused on finding strong motion parameters which are strongly correlated with structural response. Tables 2.3-2.6 summarize the measures of the ground-motion intensity proposed in literature.

#### *2.3.1.1.1 Scalar measures*

Nau and Hall (1984) considered the efficiency of different normalizing factors with the aim of investigating the effectiveness of scaling earthquake response spectra by the peak ground acceleration. Two groups of factors were considered: one based on ground-motion data (integrals of the squared motion, root square, mean square and root mean square motions) and one based on response related quantities (spectrum intensity and amplitude of Fourier spectrum of the ground acceleration). Comparisons were made in terms of elastic and inelastic spectra for a suite of 12 ground-motion records. All the records considered had been recorded in free-field (or in basements of relatively small buildings) and had a PGA greater than 0.15 g. As the authors stated no attempts were made to categorize the records selected, on the contrary “*the rather broad range of characteristics was desired to cover those which might be expected in practice*”.

## Chapter 2 – Ground motion selection and scaling issues

**Table 2.3. Acceleration based intensity measures proposed in literature.**

Measure	Name	Reference
$PGA = \max( a )$	Peak ground acceleration	
$a_{sq} = E_a = \int_0^{t_f} a^2(t) dt = a_{rs}^2$	Squared acceleration	
$P_a = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} a^2(t) dt = \frac{1}{t_{95} - t_{05}} \int_{t_{95}}^{t_{05}} a^2(t) dt = a_{rms}^2$	Earthquake power index	Housner 1975
$I_A = \frac{\cos^{-1} \beta}{\sqrt{1 - \beta^2}} \int_0^{t_f} a^2(t) dt$ ; $\beta$ : damping.	Arias Intensity. Frequency ensemble work.	Arias 1970 and Housner and Jennings 1977
$a_{rms} = \sqrt{P_a}$	Root mean square acceleration	Housner and Jennings 1964
$a_{rs} = \sqrt{E_a}$	Root square acceleration	Housner 1970
$I_c = a_{rms}^{1.5} t_d^{0.5}$ ; $t_d = t_{95} - t_{05}$	Characteristic Intensity	Park et al. 1985
$I_a = a_{\max} t_d^{1/3}$ ; $t_d = t_{95} - t_{05}$		Riddell and Garcia 2001
$\overline{EPA} = \frac{1}{2.5} \frac{1}{0.4} \int_{0.1}^{0.5} Sa(5\%, T) dT$	Effective peak acceleration	FEMA 1994
$Sa(T_1)$	Spectral acceleration at the natural period	Shome et al. Shome et al., 1998
$Sa(T_1 \rightarrow T_\mu) = \frac{1}{T_\mu - T_1} \int_{T_1}^{T_\mu} Sa(5\%, T) dT$		Kurama and Farrow 2003
$T_\mu = T_1 \sqrt{\mu / (\alpha\mu + 1 - \alpha)}$ ; $\mu$ : ductility. $\alpha$ : post yielding stiffness ratio.		
$Sa_{avg}(T_1, \dots, T_n) = \left( \prod_{i=1}^n Sa(T_i) \right)^{1/n}$		Baker and Cornell 2006b
$[Sa(T_1), Sa(T_2)]$		Shome 1999
$[Sa(T_1), \varepsilon]$		Baker and Cornell 2005
$[Sa(T_1), Sa(T_2)/Sa(T_1)]$		Baker and Cornell 2004
$I_{M,D} = \alpha Sa(T_1)^{\beta_1} R_{Sa}^{\beta_2} e^{\beta_3 t_{SM}}$ ; $R_{Sa}$ : $\frac{Sa(\text{"damaged" period})}{Sa(T_1)}$ ;		Mehanny and Deierlein 2000
$t_{SM}$ : significative duration; $\alpha, \beta_1, \beta_2, \beta_3$ : regression coefficients.		
$S_a(T_1) \left[ \frac{S_a(cT_1)}{S_a(T_1)} \right]^\alpha$ , $c = 2, \alpha = \frac{1}{2}$		Cordova et al. 2001
$A_{95} = E_x   E_x = 0.05 E_a$		Sarma and Yang 1987
Where $E_x$ defined as the area bounded by $a(t)^2$ and $A^2$ for a given $A$ . e.g. $E_x = E_a$ for $A = 0$ . $E_x = 0$ for $A = PGA$ .		
$I_{ap} = PGA \cdot Dur_{0.5PGA}^{0.5}$ ; $Dur_{0.5PGA}$ : uniform duration at 0.5PGA		Aptikaev 1982
$P_D = \frac{I_A}{v_0^2}$ ; $v_0$ : number of zero crossings per unit time of the	Potential destructiveness	Araya and Saragoni 1980

Nau and Hall (1984) found that none of the alternative scaling factor based upon recorded ground-motion data provided noteworthy reductions in dispersion compared with that obtained from normalization by the PGA. However the spectrum intensities and mean Fourier amplitude were showed to be potential alternative scaling factors for earthquake spectra, the spectrum intensities being the factor providing lowest dispersion in structural response. The authors also proposed a three-parameter system of spectrum

**Table 2.4. Velocity based intensity measures proposed in literature.**

Measure	Name	Reference
$PGV = v_{\max}$	Peak ground velocity	
$P_v = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} v^2(t) dt = \frac{1}{t_{95} - t_{05}} \int_{t_{95}}^{t_{05}} v^2(t) dt = v_{rms}^2$		
$v_{sq} = E_v = \int_0^{t_f} v^2(t) dt = v_{rs}^2$	Squared velocity	
$v_{rms} = \sqrt{P_v}$	Root mean square acceleration	
$v_{rs} = \sqrt{E_v}$	Root square acceleration	
$I_F = v_{\max} t_d^{0.25}; t_d = t_{95} - t_{05} \text{ (significant duration)}$		Fajfar et al. 1990
$I_v = v_{\max}^{2/3} t_d^{1/3}; t_d = t_{95} - t_{05} \text{ (significant duration)}$		Riddell and Garcia 2001
$\overline{EPV} = Sv(5\%, 1s); Sv: \text{spectral velocity};$	Effective peak velocity	FEMA 1994
$SI_H(\beta, T_1, T_2) = \int_{T_1}^{T_2} Sv(\beta, T) dT;$	Housner's spectral intensity	Housner 1952
$T_1 = 0.1; T_2 = 2.5.$		
$SI_{HC} = SI_H(\beta, 0.1, 1.0);$		Hidalgo and Clough 1974
$MIV = \max(IV)$	Maximum incremental velocity.	Bertero et al. 1976
$IV: \text{area under the accelerogram between two consecutive 0 crossings.}$		
$SI_M = \frac{1}{T} SI_H(T, 2T)$		Matsumura 1992
$SI_{MR} = \frac{1}{T - T_y} SI_H(T, T_h)$		Martínez-Rueda 1998
$T_h: \text{hardening period, i.e. period computed using the tangent stiffness of the hardening branch of the idealized lateral response of the structure.}$		
$SI_{KK} = SI_H(\beta, T_1 - t_1, T_1 + t_2)$		Kappos and Kyriakakis 2000
$T: \text{Natural period of the structure; } t_1 = t_2 = 0.2 T_1$		

intensities, based on pseudovelocity spectrum values for different frequency intervals:  $SI_a$ ,  $SI_v$ ,  $SI_d$ .

Matsumura (1992) investigated the efficiency of four ground-motion parameters, namely: peak ground acceleration, peak ground velocity, mean velocity spectral intensity,  $SI_M$ , between  $T$  and  $2T$  ( $T$  being the natural period of the SDOF system considered in the study) and mean equivalent velocity,  $V_e$ , converted from the input energy,  $E_i$ , between  $T$  and  $2T$ . The conversion was done by  $V_e = \sqrt{2E_i/m}$ . Twelve ground-motion records were used in this study, although the records are reported in the paper the adopted selection criterion is not described by the author who only states that the records “are commonly used in Japan as induced ground acceleration in the analysis of tall buildings”.

The author concluded that PGA and PGV are good measures only in narrow frequency ranges, in the high and in the low frequency ranges respectively. On the other hand  $SI_M$  and  $V_e$  were found to be good measure in a wide range of frequency therefore Matsumura suggested their use as measures of the intensity of the ground motions.

Martinez-Rueda (1998) investigated different parameters based on spectral scales (Nau and Hall criterion,  $SI_{MR}$ ,  $SI_M$ ,  $SI_H$ ,  $SI_{KK}$ ) by studying the response of bilinear SDOF systems with kinematic hardening and non-degrading stiffness with different periods, seismic coefficients and post-yield stiffness ratios.

**Table 2.5. Displacement based intensity measures proposed in literature.**

Measure	Name	Reference
$PGD = d_{\max}$	Peak ground displacement	
$P_d = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} d^2(t) dt = \frac{1}{t_{95} - t_{05}} \int_{t_{05}}^{t_{95}} d^2(t) dt = d_{sq}^2$		
$E_d = \int_0^{t_f} d^2(t) dt = d_{rms}^2$		
$d_{sq} = E_d = \int_0^{t_f} d^2(t) dt$		
$d_{rms} = \sqrt{P_d}$		
$d_{rs} = \sqrt{E_d}$		
$I_d = d_{\max} t_d^{1/3}$ ; $t_d = t_{95} - t_{05}$ (significant duration).		Riddell and Garcia 2001
$Sd(T_1)$		
$IM_{1E}$		Luco and Cornell 2007
$IM_{1I}$		Luco and Cornell 2007
$IM_{1E\&2E}$		Luco and Cornell 2007
$IM_{1I\&2E}$		Luco and Cornell 2007
$IM_{1eff} = \sqrt{\frac{S_d(2T_1, \zeta_1)}{2S_d(T_1, \zeta_1)}} IM_{1E}$		Luco and Cornell 2007
$Sd_i$	Inelastic spectral displacement	Tothong and Cornell 2007

**Table 2.6. Hybrid intensity measures.**

Measure	Name	Reference
$I_z = \frac{\int_0^{Dt} a(t)^2 dt}{PGA \cdot PGV}$		Cosenza and Manfredi 1998
$SI_a = \frac{1}{0.157} SI_H(\beta, 0.028, 0.185) \quad T \in [0.118, 0.500] s$		Nau and Hall 1984
$SI_v = \frac{1}{1.715} SI_H(\beta, 0.285, 2.000) \quad T \in [0.500, 5.000] s$		
$SI_d = \frac{1}{8.333} SI_H(\beta, 0.4.167, 12.500) \quad T \in [5.000, 14.085] s$		

The author used an ensemble of 100 accelerograms, corresponding to the two horizontal components of 50 ground motions. The only selection criteria reported by the author are magnitude greater than 6.0 (six accelerograms with lower magnitudes were included) and PGA greater than 0.1 g. The ground motions used had been recorded at epicentral distances of up to 400 km and produced by earthquakes with magnitudes ranging from 5.4 to 8.1. Soft soil, stiff soil and rock-site recordings were included. The author concluded that there is not a unique spectral intensity that performs better over the entire period range: for structures with medium and high strengths,  $SI_H$  was found to be the most unstable scale and showed poor correlation with displacement ductility demand for short- and long-period structures. Both  $SI_M$  and  $SI_{MR}$  reduced their performance for the intermediate period range. For structures with low strength,  $SI_H$  showed the best performance in the short-period region,  $SI_{MR}$  dominated over the intermediate- and long-period range. The performance of  $SI_H$  were found to deteriorate for increasing periods whereas the opposite occurs in the case in  $SI_M$  and  $SI_{MR}$ . Taking advantage of these observations the author proposed a combined spectral intensity scale that uses, a different spectrum intensity for different values of period, yield seismic coefficient and post yield stiffness. It should be noted that MDOF structures were not investigated in this study.

Shome et al. (1998) observed that the non-linear response from a bin of records chosen from a narrow magnitude and distance interval display wide dispersion and suggested that when the records in each bin are normalized or scaled to the bin median spectral acceleration at the fundamental frequency of the structure, the same median structural response is obtained with reduced variability compared to those of the unscaled sets. Therefore the author suggested that the most efficient way to estimate the non-linear structural response from a given scenario ( $M, R$ ) is to first use an attenuation equation to estimate the median  $Sa(T_1)$  and then to scale record from roughly the same magnitude to this spectral acceleration. Doing so can reduce the number of runs by a factor of 4 (Section 2.3.4) the authors also suggested that apart from the Normalized Hysteretic Energy (NHE) damage measure, scaling records did not appear to bias non-linear response estimates. Shome et al. also found that among several scaling measures considered in their study  $Sa(T_1, 5\%)$  is the one performing best. All their results were obtained for SDOF structures and for a single first mode dominated MDOF structure, hence they require verification to be extended different kinds of MDOF structures.

Later, Shome (1999) investigated the efficiency of scaling with respect to  $Sa(T_1)$  on tall buildings whose response is strongly dependent on higher vibration modes. The author found that the reduction in structural response is lower than in first-mode dominated

structures. Therefore Shome investigated an improved intensity measure composed by weighted average spectral accelerations at different periods, and showed that using this intensity measure a remarkable reduction in the dispersion of structural response can be obtained. The main drawback of this approach is that a structure-specific probabilistic seismic hazard analysis is required. In Shome (1998) and Shome and Cornell (2000) the records used were selected in magnitude-distance bins from stiff soil sites, excluding accelerograms with near-source forward directivity pulses.

Kappos and Kyriakakis (2000) compared the efficiency of different ground-motion parameters. Two different ground-motion datasets were used in this work: 11 records from 11 Greek earthquakes, and 13 records from 8 US earthquakes. Records within each set were almost evenly distributed with respect to rock and alluvia sites. The authors claimed that this “*compilation of records permits consideration of the effect of soil conditions (inevitably in a rough way) as well as of the tectonic regime*”. The first part of the study focused on the effect of scaling on both acceleration and displacement spectra. It was concluded that in the intermediate- and long-period ranges any of the three velocity-related parameters, PGV,  $SI_H$  and  $I_v$  can be used to good effect. The authors investigated also the response of multi-storey frames. Kappos and Kyriakakis suggested that narrow-band spectrum intensities calculated on the basis of elastic and inelastic pseudovelocity spectra as an alternative scaling. In particular the authors found scaling on the basis of the area under the inelastic pseudovelocity spectrum, taken between the initial elastic and the effective inelastic period of the structure to be the most attractive procedure. Their proposals based on the narrow-band elastic SI concept led to COV values between 10% and 40% in the calculated drifts and member ductilities, and a reasonably uniform distribution of scatter along the building height. The case-study considered involves a single structure designed to modern design practice which targets and achieves uniform damage distribution. Bommer and Acevedo (2004) commented that since the Greek records are filtered with a low-frequency cut-off at 1 Hz, therefore the elastic and inelastic displacement spectral ordinates at longer periods, discussed in the paper, should have been neglected. Furthermore Bommer and Acevedo (2004) suggested that differences in the characteristics of the two data sets are attributed to other factors including “*the deeper deposits in some Californian sites (e.g. the Bay area)*”, despite the fact that two of the three records obtained in the San Francisco Bay Area are from rock sites and the third (Parking Garage, Stanford) is not located on Bay mud. Bommer and Acevedo (2004) also suggested the finding that for inelastic spectral responses the COV increases with increasing ductility factors, may simply be due to the large range of magnitudes - and hence durations - in their data sets.

Furthermore Bommer and Acevedo (2004) argued that another explanation for the increase in the COV for the rock data sets than for the alluvial (soil) data sets, which Kappos and Kyriakakis attributed to the different site classifications, could be found in the different ranges of magnitude in the data sets. For the US data, the maximum differences in magnitude amongst the soil records was 1.4 as opposed to 1.8 for the rock data, and for the Greek records the magnitude variation amongst the soil site recordings was just 1.3 compared with 2.1 for the rock site accelerograms.

As was concluded in Shome (1999), spectral shape plays a central role in the response of multi-degree-of-freedom structures, therefore many authors tried to derive specific parameters to address this issue. Mehanny and Deierlein (2000) proposed the following intensity measure to reduce the dispersion in structural response of concrete structures:

$$I_{M,D} = \alpha Sa(T_1, \xi)^{\beta_1} R_{S_a}^{\beta_2} e^{\beta_3 t_{SM}} \quad (2.14)$$

where  $Sa(T_1, \xi)$  is the spectral acceleration at the fundamental frequency of the structure,  $R$  is the ratio of the spectral acceleration at fundamental frequency of the “damaged” structure and  $Sa(T_1, \xi)$ , and  $t_{SM}$  is the strong-motion duration according to Trifunac and Brady (1975). The parameters  $\alpha$ ,  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  were determined by regression analysis. The authors studied the response of two reinforced concrete frames and concluded that the most of the structural dispersion was explained by the spectral shape term ( $R$ ). No noteworthy reduction in structural response was obtained by considering the strong-motion duration. In this work 8 far-field and 8 near-field records were adopted. Moreover three records in the former set were obtained from ground motions produced by the same earthquake. The authors stated that the aforementioned records should be representative of two different scenarios: respectively an high magnitude ( $6.9 < M < 7.4$ ) large distance ( $8.5 < R < 66\text{km}$ ) events and small magnitude ( $6.5 < M < 7.0$ ) short distance ( $1.2 < R < 7.5$ ) events. All the accelerograms were recorded on stiff soil or rock sites. All the records were selected such that  $Sa$  at the fundamental frequency of the structure was as close as possible to the reference target value, i.e. the 2% in 50 years value. As for the near-field records, only ground motions with forward directivity were selected.

Cordova et al. 2001 investigated a simplified version of the aforementioned intensity measure, namely  $S^* = Sa(T_1)(Sa(cT_1)/Sa(T_1))^\alpha$ . The authors showed that this proposed index significantly reduces the record-to-record variability in predicted response obtained from inelastic time history analyses. The same set of ground motions

and of structural models adopted by Mehanny and Deierlein (2000) was used in this study.

Khurama and Farrow (2003) investigated the effectiveness of seven ground-motion scaling methods, based on different IMs, in reducing the scatter in estimated peak lateral displacement demands. This study investigated response of different SDOF and MDOF systems and the influence of soil type. Ground-motion records from very-dense-, stiff- and soft-soil profiles were considered: twenty records were used for each soil category. In addition 20 records chosen from the 40 proposed by the SAC steel project Somerville et al., 1997b were used. Khurama and Farrow considered the following IMs: PGA, EPA,  $A_{95}$ , EPV, MIV (maximum incremental velocity),  $Sa(T_1)$ ,  $Sa(T_0-T_\mu)$ . The main conclusions of this study were:

- for the very-dense, stiff-, and soft-soil ground-motion ensembles at long periods the MIV-based scaling is more effective the PGA based method.
- For the very-dense- and stiff-soil ground-motion ensembles, the method based on  $Sa(T_1)$  is more effective than the method based on PGA, except for short- and very-short-period structures; the effectiveness of the method decreases as the strength ratio increases.
- With the  $Sa(T_0-T_\mu)$  method the scatter is lower that the scatter produced by the  $Sa(T_0)$  method.
- For the very-dense-, stiff- and soft-soil ground-motion sets, the MIV method is more effective than the  $Sa(T_1)$  and  $Sa(T_0-T_\mu)$  methods for a wide range of periods.
- For the soft-soil and near-field ground-motion sets the effectiveness of the  $Sa(T_1)$  and  $Sa(T_0-T_\mu)$  methods with respect to the PGA method is decreased.
- Except for long periods, the  $Sa(T_1)$  and  $Sa(T_0-T_\mu)$  methods are less effective that the PGA based method for the soft-soil set.

This study did not investigate limits of scaling and bias.

Akkar and Ozen (2005) investigated features of PGV for non-linear deformation demands on SDOF systems and for some specific ground-motion features that can be effective in structural response. A total of 60 soil site records with  $2.5 < R < 23$  km and  $5.5 < M < 7.6$  were used. The accelerograms used are low-cut filtered by corner frequencies between 0.1 and 0.2Hz except for the digital 1999 Chi-Chi, Taiwan, earthquake records that are processed by corner frequencies of 0.04 Hz. This information was used in defining the period bound for non-linear response history computations. Ground motions with pulse signals were not considered in this study. The records were grouped, depending on their PGV, in 3 sets of 20 records each. The considered ranges of PGV were  $PGV < 20\text{ cm/sec}$ ,  $20\text{ cm/sec} < PGV < 40\text{ cm/sec}$  and

$40\text{ cm/sec} < \text{PGV} < 60\text{ cm/sec}$ . The non-linear time-history analyses were conducted using non-degrading and stiffness-degrading hysteretic behaviours for constant  $\eta$  and  $R$  levels. The results presented in this study indicate that the PGV intensity measure correlates well with the earthquake magnitude, effective ground-motion duration and frequency content of ground motions. The increase in PGV values is associated with an increase in earthquake magnitude and effective duration of records. The frequency composition of ground motions systematically becomes richer in the long-period range for increasing PGV. The statistical results presented for the mean spectral displacements at different inelastic levels show a consistently increasing trend for ground motions of higher velocity. The mean spectral displacement statistics at different PGV levels also reveal that the spectral region where the inelastic deformations are significantly higher than their elastic counterparts is sensitive to the amplitude of ground velocity. This spectral region shifts towards longer periods with the increase in PGV. The validity of these observations was also tested for PGA and PGV/PGA ratio by computing the period-dependent correlation coefficients at different inelastic deformation levels. The comparisons of correlation coefficients between PGV and these two alternative intensity measures highlighted the superiority of PGV with respect to these two measures. Correlation coefficients computed for spectral acceleration do not resemble a stable behaviour when compared to the general performance of correlation coefficients computed for PGV. The dispersion statistics also reveal a better performance of PGV with respect to  $S_a$  particularly when the short-period SDOF demands are conditioned on PGV rather than on  $S_a$ .

Riddell (2007) investigated the effectiveness of 23 ground-motion intensity indices in terms of correlation with four different structural response variables, i.e. elastic and inelastic deformation demands, input energy and hysteretic energy. Non-linear responses were computed using elastic-plastic, bilinear, and bilinear with stiffness degradation SDOF models. No index showed satisfactory correlation with the non-linear response in all the spectral regions simultaneously. Indeed, acceleration-related indices were found to be the most efficient for rigid systems, velocity-related indices for intermediate-frequency systems, and displacement-related indices for flexible systems; some velocity-related indices also showed and high efficiency in the low-frequency region. In this study 90 ground motions were used, although the author listed all the records adopted no information was provided on the selection criteria adopted.

Bianchini et al. (in preparation) investigated an intensity measure based on the geometric mean of spectral accelerations at different periods and 5% damping:

$$Sa_{avg}(T_1, \dots, T_n) = \left( \prod_{i=1}^n Sa(T_i) \right)^{\frac{1}{n}} \quad (2.15)$$

where  $Sa(T_i)$  is the spectral acceleration at the period  $T_i$  and  $n$  is the total number of periods considered. This intensity measure had been firstly proposed by Baker and Cornell (2006b) who extended an idea of Cordova et al. (2001) and Mehanny (2000). By taking the logarithm of both sides, Eq. (2.15) can be rewritten as:

$$\ln(Sa_{avg}(T_1, \dots, T_n)) = \frac{1}{n} \sum_{i=1}^n \ln(Sa(T_i)) \quad (2.16)$$

This form is particularly convenient for attenuation equations quote the results of regression analyses in terms of the logarithm of spectral accelerations. Therefore a predictive equation for  $\ln(Sa_{avg})$  can be easily developed with an arbitrary set of periods  $T_1; \dots; T_n$  using existing attenuation models. The regression coefficients for  $\ln(Sa_{avg}(T_1; \dots; T_n))$  can be obtained simply taking the mean of the regression coefficients for each  $\ln(Sa(T_i))$ . Thus, PSHA can be performed using  $\ln(Sa_{avg})$  as intensity measure in the same way of any single spectral acceleration value. Concerning the distribution of  $\ln(Sa_{avg})$ , if it is assumed that multiple  $\ln(Sa(T_i))$  values are jointly Gaussian distributed then  $Sa_{avg}$  is also normally distributed for it is a linear function of normal variables. Therefore the parameters of this distribution can be easily derived considering the correlation coefficients derived by Baker (2005).

The authors demonstrated the efficiency of  $Sa_{avg}$  compared to traditional elastic ( $Sa(T_1)$  and peak-ground parameters) and advanced inelastic intensity measures. A huge set of inelastic SDOF and MDOF systems was used to represent the dynamic behaviour of different typologies of buildings. Forty ordinary ground-motions were used in this work. They do not exhibit pulse-type near-fault characteristics and are recorded on stiff soil or soft rock. The source-to-site distance, ranges from 13 to 40 km and the moment magnitude,  $M_w$ , from 6.5 to 6.9. Additional ground-motion features considered defining this set of 40 ordinary ground-motions were: (i) strike-slip, reverse-slip and reverse-oblique fault mechanisms; (iii) aftershocks were not included; (iv) the high-pass corner frequency less or equal than 0.20 Hz. For each recording station, one horizontal component was randomly selected. The authors stated that the use of a single set of GMs is acceptable because it has been shown that inelastic response of systems is not greatly affected by  $M_w$  and  $R$  (except for near fault regions). The size of the set of records was chosen by considering the confidence interval on the estimated response,

i.e. in order to obtain estimates of the median that are within a one-sigma confidence band of 10% as long as the standard deviation of the natural logarithm of the collapse capacities was less than  $0.1/\sqrt{40} = 0.63$ . This set of ground-motions was originally proposed by Medina and Krawinkler 2004.

### **2.3.1.1.2 Vector measures**

Some researchers have recently focused their attention on intensity vectors (Shome, 1999; Baker and Cornell, 2004; Baker, 2005; Baker and Cornell, 2005; Baker, 2007c). The use of a vector intensity measure instead of a scalar one involves many unsolved problems, such as the scaling techniques and the more difficult and unfamiliar hazard calculations (Bazzurro and Cornell, 2002).

The first proposal of a vector measure of the ground-motion intensity can be found in Shome (1999). The author observed that the response of MDOF structure depends on the spectral ordinates at different periods therefore proposed to adopt the following vector measure:  $[Sa(T_1), Sa(T_2)]$ , where  $T_1$  and  $T_2$  are the first and the second natural-period of the structure, respectively. Later Baker and Cornell (2004) proposed to adopt a vector containing  $Sa(T_1)$  and the ratio  $R = Sa(T_1)/Sa(cT_1)$ . Later, the same authors suggested to substitute  $R$  with  $\epsilon$  (see 2.2.3.3) because this latter was shown to be a better indicator of spectral shape than  $R$ . Further research is needed on scaling procedures for vector intensity measures. The state of the practice is to use one of the elements of the vector as reference for scaling and include the second term by careful post-processing of the results (Baker, 2007c).

### **2.3.1.1.3 Measures based on non-linear response**

With the aim of finding highly sufficient and efficient measures Luco and Cornell (2007) and Tothong and Luco (2007) proposed some “*advanced*” structure specific ground-motion intensity measures based on elastic and inelastic structural response. Luco and Cornell (2007) proposed to take in account multi-modal response and inelastic displacements. In this work both ordinary and near-source records were adopted: a first set of time-histories consists of records with closest distances to the rupture surface,  $R$ , between 30 and 46 km and a second of records with  $R$  less than 16 km. The near-source suite was restricted to “forward-directivity” earthquake records, whereas the ordinary suite excludes such earthquake records. The strike-normal component of each of the near-source ground-motion records was considered by the authors (the strike-normal component of each ordinary ground-motion record was considered to maintain consistency). All of the ground-motion records selected also satisfied the following

criteria: *i*) earthquake moment magnitude,  $M$ , greater than or equal to 6.0, *ii*) recorded on “stiff soil” or “very dense soil and soft rock” and *iii*) record processed with a maximum (between two horizontal components) high-pass filter corner frequency less than or equal to 0.25 Hertz. Two other groups of near-source records, selected with less restrictive criteria, were used as a check for the results obtained with the main sets. All of the near-source earthquake records considered were scaled by a factor of two. The ordinary earthquake records were scaled by a factor of eight.

The ground-motion intensity measure denoted  $IM_{1I\&2E}$ , which takes into account second-mode frequency content and inelastic structural behaviour, was demonstrated to be relatively efficient and sufficient with respect to the structural drift response measure. The intensity measure  $IM_{1I\&2E}$  is defined as:

$$IM_{1I\&2E} = \frac{Sd_i(T_1, \beta, d_y)}{Sd(T_1, \beta)} IM_{1E\&2E} \quad (2.17)$$

where  $\beta$  denotes the damping ratio,  $Sd(T_1, \beta)$  the spectral displacement at the natural frequency,  $Sd_i(T_1, \beta, d_y)$  the spectral displacement for an elastic-perfectly-plastic SDOF oscillator with period  $T_1$  and ductility  $d_y$ .  $IM_{1E\&2E}$  is an intensity measure defined as:

$$IM_{1E\&2E} = \sqrt{\left[ PF_1^{[2]} Sd(T_1, \beta) \right]^2 + \left[ PF_2^{[2]} Sd(T_2, \beta) \right]^2} \quad (2.18)$$

where  $PF_1^{[2]}$  and  $PF_2^{[2]}$  are record-dependant participation factors depending on different vibration-modes and on maximum interstorey drift. Further details can be found in Luco 2007. For ordinary records, using these “*advanced*” intensity measures led to the same conclusions obtained using the vector  $IM, \langle S_a, \mathcal{E} \rangle$ . However, using advanced  $IMs$  to evaluate the structural performance for *near-source* pulse-like records is found to be more accurate than using the elastic-based  $IMs$ .

In order for the aforementioned non-linear intensity measures to be applied in a probabilistic earthquake engineering framework, ground-motion attenuation relationships are required. Those latter have been recently developed for  $Sd_i(T_1, \beta, d_y)$ , by Tothong and Cornell (2006). Although their application is probably restricted to special projects because the involved computations must be structure specific (the GMPE depends on structural period and yield displacement).

### 2.3.1.2 Scaling to model the distribution of ground-motions

Normally scaling procedures are adopted to achieve a good degree of compatibility with some kind of reference spectrum. Recently some authors proposed procedures that not only try to match a median spectrum but also try to control the distribution of the ground-motions.

Rathje and Kottke (2007) proposed a “*semi automated*” procedure to select and scale ground-motion records to fit a target response spectrum (given  $M$ ,  $R$ ,  $S$ ,  $F$ ) while controlling the variance. The methodology selects motions for the suite based on matching the spectral shape, and then fits the amplitude and standard deviation of the target by adjusting the scale factors. The method is based on defining the median response spectrum of a set of records as the mean value of the natural logarithm of the spectral accelerations at each period; adopting this formulation the average response spectrum at the period  $i$  can be written as:

$$\begin{aligned}\ln Sa_{avg,i}^{scaled} &= \frac{1}{n} (\ln s_1 + \ln s_2 + \dots + \ln s_n) + \frac{1}{n} (\ln Sa_{1,i} + \ln Sa_{2,i} + \dots + \ln Sa_{n,i}) = \\ &= \ln s_{avd} + \ln Sa_{avg,i}\end{aligned}\quad (2.19)$$

where  $n$  is the number of records considered,  $s_j$  is the scaling factor for the record  $i$  and  $Sa_{j,i}$  is the spectral acceleration of the record  $j$  at the period  $i$ . The first term of the equation controls the amplitude while the second term controls the spectral shape. Since the individual scale factors can be varied without changing  $\ln Sa_{avg,i}^{scaled}$  the authors proposed to use this property to control the standard deviation of the spectra of the set of records; i.e. the average scale factor is used to achieve a good match with the target mean spectrum (in a given period range) while the individual scale factors are chosen in order to match the median  $\pm 1 \sigma$  spectra.

Malhotra (2007) proposed a scaling procedure based on the peak values of acceleration, velocity and displacement. First the median and standard deviation of peak values of ground acceleration, velocity and displacement ( $PGA$ ,  $PGV$  and  $PGD$ ) are estimated from a prediction equation given  $M$ ,  $R$ ,  $S$  and  $F$ . Next, some values of the aforementioned parameters are randomly sampled taking into account their correlation. Records are then chosen which match the median  $\pm 1 \sigma$  values of significant duration at the site (also calculated via prediction equation) and are scaled to match the simulated values of  $PGA$ ,  $PGV$  and  $PGD$ . It is not clear if only one of the three peak values is assumed as reference for scaling or if all are considered together trying to minimize the

mean square error. Furthermore the author did not provide specific recommendations on the number of samples that should be used.

The ATC-58 project (Zareian, 2007) is developing a ground-motion selection procedure with the aim of defining a suite of ground motions that allows an accurate estimate of the cumulative distribution function of a structural response under consideration, given  $M$ ,  $R$  and  $Sa(T_1)$ . The selection procedure is based on defining a  $M$ - $R$  bin and then randomly choosing eleven records from this latter. The chosen records are scaled according to the following procedure: first given  $M$  and  $R$  the median spectra acceleration and its associated dispersion is determined using any ground-motion prediction equation appropriate for the site and the earthquake sources considered. Then eleven values of spectral acceleration are computed using the equation  $Sa_i(T_1) = \theta \exp(\beta \eta_i)$  where  $i = 1$  to 11 and  $\eta_i = \Phi^{-1}(P_i)$ , where  $\Phi^{-1}(\cdot)$  is the inverse standardized normal distribution and  $P_i = (1/11)(i-1) + (1/22)$ . Each individual record is scaled to one of the eleven computed  $Sa_i(T_1)$  values. Some objection can be made to this method: this procedure does not take in account the “record to record” variability because only one records is used for each given  $Sa_i$  value; records are chosen with no regard to the spectral shape; no limits on scaling are defined and finally estimating tails of the structural distribution with 11 records only can produce results very sensitive to the adopted set of ground-motion.

The Geomatrix consultancy group (2007) proposed the following procedure as a part of the Ground-Motion Selection and Modification program. The DGML package (see Section 2.2.3.2) is used to identify records within a user-specified  $M$ - $R$  bin that includes the given  $M$ ,  $R$  and fault mechanism. Next records are scaled to the level of the given first-mode spectral acceleration  $Sa(T_1)$ , so each spectrum matches the target value at the given period. The theoretical conditional mean and conditional mean standard deviation are calculated at a period  $T$  that is well away from  $T_1$  (the authors suggested “*approximately 0.1 T<sub>1</sub>*”). A given number of log-normally distributed points are then randomly sampled based on the theoretical distribution. For each point of the distribution, one scaled record within the data-bin is selected which has  $Sa(T_1)$  closest to that point. Those steps are repeated until the realization that appears to best match the distribution over the full period range is selected. The mean squared error (MSE) of the selected records against the target conditional mean spectrum is calculated, and records are ranked in ascending order of MSE over a specified period range of interest. The authors suggested that groups of 7 records can be used to estimate median EDP, while to estimate the whole probability distribution more records are needed (the authors suggested 28).

### 2.3.1.3 Bias and scaling limits

Many papers on the comparison of different intensity measures only focused on the reduction of dispersion of structural response, but an important point that should always be considered is the degree of which scaling introduces bias in the estimated structural response and therefore what the limits of scaling are.

Concerning this latter point contrasting recommendations can be found in the literature on the topic and as Bommer and Acevedo (2004) suggested, many of the often cited limits to scaling do not have any scientific motivation. The authors reported that Krinitzky and Chang (1977) proposed that if scaling factors of 4 or more needed to be applied to accelerograms, then the records should be rejected, although no justification was given for this assertion. Subsequently Vanmarcke (1979) proposed reduction of the limits on scaling to a factor of 2 for liquefaction analysis, although the limit of 4 was upheld for linear elastic systems. Vanmarcke (1979) based his conclusions on a study of inelastic spectra and of correlations amongst different strong motion parameters, using a dataset of 70 accelerograms. The dataset only included 12 accelerograms that had horizontal peak accelerations of at least 0.2 g and 41 of the accelerograms had PGA values below 0.1 g. Despite the limitations of the data and the analyses underlying the conclusions, the recommendations from these two studies are frequently used as a rule in practice. Malhotra (2003) found that a scaling factor of 5.84 was required for one record used in his study and concluded that this was “*higher than the normally accepted upper limit of 4*”. Bommer and Acevedo (2004) concluded that presumably, the rationale behind imposing limits on scaling is to avoid creating unrealistic ground motions, since this would undermine the inherent value in using real accelerograms in the first place. A more rational approach to investigate the limits of scaling is based on looking for the presence of bias in structural response, in particular on trying to define a dependence between scaling and bias.

It should be noted that it is not the scaling itself that induces bias in structural response, in fact this phenomenon is strictly dependent on the intensity measure used. Quantitative investigations into ground-motion scaling indicated that a suite of ground motions may be safely scaled to the suite’s median spectral acceleration value, at a period  $T_1$ , without biasing the median response of a structure having the same first-mode period  $T_1$  (Bazzurro and Cornell, 1994; Shome et al., 1998; Shome, 1999; Iervolino and Cornell, 2005). But recent work suggests that in some other situations record scaling may induce some bias in structural response (Baker and Cornell, 2005; Baker, 2007b; Luco and Bazzurro, 2007). This bias appears to result from the scaled

ground motions having inappropriate values of spectral shape or the parameter  $\varepsilon$ , which is an indirect measure of spectral shape (Section 2.2.3.3). An important feature of these ground-motion scaling studies is that record selection and scaling approaches are evaluated by studying the response of structures subjected to these motions. If it can be verified that scaled ground motions produce structural responses similar to those from unscaled ground motions having the same intensity, then it can be concluded that the given scaling approach is valid.

In Baker (2007b) and Luco and Bazzurro (2007) it was concluded that inappropriate record scaling can bias estimated structural response; this supports the concern that record scaling might fail to modify all ground-motion properties in an appropriate way. Baker (2007b), through the exploration of conditional mean response spectra, showed that the frequency content of ground-motions does change as the intensity (i.e.,  $Sa(T_1)$ ) changes and concluded that the frequency content is more affected by the variation of  $\varepsilon$  than by the variation of magnitude or distance. The author suggested that if records are selected with the desired spectral shape through a careful record selection scheme (i.e.,  $\varepsilon$ -based or CMS-based selection), they can be scaled without inducing bias. Similar recommendations were suggested by Bommer and Acevedo (2004).

These results may at first glance appear to conflict with the past aforementioned studies that did not detect scaling bias when records were scaled to target  $Sa(T_1)$  values. The reason for the difference can be found in the fact that those studies were considering a specific problem where the mean scale factor among all the records in a suite was approximately one. In those cases, where as many records were scaled up as were scaled down, the median observed maximum interstorey drift ratio was unbiased. This is consistent with the aforementioned results obtained by Baker (2007b) and Luco and Bazzurro (2007), which predict that biases from scaled-up and scaled-down records would offset, resulting in unbiased median response when the average scale factor is approximately one.

Hancock et al. (2008) obtained similar conclusions about the importance of spectral shape in the reduction of bias induced by scaling. The authors proposed a new method to investigate bias. It is based on calculation of the “true” median response by regression analysis, i.e. by fitting an empirical predictive equation for different structural response parameters using data obtained by a huge set of ground motions. As the authors noted “*Although the number of records available for a particular scenario does not change, a great deal of additional constraint on the expected inelastic response is provided by the records from other magnitudes and distance ranges*”. The authors compared sets of ground motions scaled according to different criteria ( $Sa(T_1)$ ),

average spectral acceleration between 0.25s and 1.25s, spectral matching to 5% damped spectrum using wavelets and spectral matching to multiple damping ratios using wavelets) to match a reference spectrum obtained by a ground-motion prediction equation. All the ground motions used were selected according to the criterion proposed by Bommer and Acevedo (2004) i.e. the matching with the reference spectrum was considered. It was concluded that if the accelerograms are selected to match the reference spectrum shape they can be scaled up to factors of 10 without introducing bias. The authors also concluded that since the fatigue and absorbed hysteretic energy damage measures are known to be dependent on ground-motion duration, the lack of bias they observed in these damage measures demonstrates that selecting records to match magnitude is sufficient to prevent scale bias from different ground-motion durations. Hancock et al. (2008) suggested that the bias can be further reduced when the accelerograms are wavelet adjusted to match multiple damping ratios (Section 2.3.3).

### *2.3.2 Frequency domain scaling*

Frequency-Domain techniques for the generation of artificial accelerograms associated to earthquake response spectra have existed for many years. An early proposition is utilized by the computer code SIMQKE (Gasparini and Vanmarcke, 1976); it is based on the relationship between expected response spectral values and the spectral-density function of a random process representation of ground motions. This relationship is derived from analytic Random Vibration Theory (RVT) techniques. A seed record is generated as a simulated realization of this random process, and subsequent iterations are conducted in the frequency domain.

Carballo and Cornell (2000) proposed a procedure for the modification of existing earthquake records in which the target spectrum depended on the response spectrum of the record itself. The modification technique is based on scaling the record's Fourier amplitude spectrum by a smooth "factor function" which is computed by finding the ratio of the predicted response spectrum of the target scenario event to the predicted target of the seed event.

Another widely used Frequency-Domain method is the one coded in the computer program RASCAL (Response Spectra and Acceleration Scaling) (Silva, 1987). It is a semi-empirical procedure, which evolved from random vibration theory based techniques, with the additional (empirical) utilization of the observed Fourier phase spectrum of a real ground motion. The RASCAL procedure modifies the Fourier amplitude spectrum of a seed record in two stages. In the first stage, the factor function applied to the Fourier amplitude spectrum is a smooth one, generated by the ratio of the

RVT-predicted expected smooth response spectrum associated to a Brune-based source spectrum to the target response spectrum. This factor function is then used to modify the Brune Fourier amplitude spectrum. This product is taken as the Fourier amplitude spectrum of the first iteration for a number of predefined iterations. The Brune source spectrum considered is that of the so-called  $\omega$ -squared model. This basic form is then modified to include the radiation pattern, the amplification due to the free surface, high-frequency filtering, and the path attenuation. In the second stage, the factor function is based on the observed response spectrum (computed in the time domain) of the current iteration's time history (which is generated from the current amplitude spectrum and the seed's phase spectrum).

### *2.3.3 Time domain scaling*

Even though the target used as the criterion for generating the artificial accelerogram is based on what looks like a frequency representation (i.e., the target response spectrum), such a representation indicates a (maximum) response that occurred in time. Time-domain schemes take advantage of this fact, working in a backward approach to modify the time-specific event of maximum response (e.g., the maximum observed acceleration of an oscillator). The basic idea of such time-domain modification of earthquake records assumes that the time of the peak response will not be perturbed by adding a small adjustment to the original time history. This way, the observed response is modified to that of the indicated value given by the response spectrum, taking care of the observed difference between it and the accelerogram actual observed maximum response. Sets of “adjustment functions” are then found iteratively, correcting the accelerogram at all the corresponding frequencies of the target, requiring the solution of simultaneous equations to find the amplitudes of the adjustment functions. Two good properties of the procedure (Carballo and Cornell, 2000) are its fast convergence and the ability to conserve for most cases the nonstationary character of the seed time history, provided that the adjustment functions are appropriately selected. Different adjustment functions (wavelets) have been proposed by different authors, a comprehensive review can be found in Hancock 2006 and Hancock et al. 2006.

The use of spectrum compatible records has been common for different non-linear systems. Examples can be found in Preumont (1984), Barenberg (1989) and Naeim and Lew (1995). It should be mentioned that the level of satisfaction of the results obtained by different authors has not been constant. As Carballo and Cornell (2000) suggested, it is important to put the results obtained by these different authors in the context in which they were generated. In contrast to other studies, Carballo and Cornell (2000)

considered the use of spectrum-matched records for obtaining representative demands corresponding to a particular  $M$ - $R$  seismic event. The authors found that for this particular case the use of spectrum-matched accelerograms may yield unconservatively biased estimates of non-linear ductility demands. Carballo and Cornell compared structural response of different SDOF and MDOF structures obtained by unscaled  $M$ - $R$  scenario compatible ground-motions and by compatibilizing the same records to their median response spectrum. The authors found that for the estimation of the ductility demands that correspond to a scenario event, spectrum-matched records will generate results that predict an unconservatively biased estimate of median displacement-based non-linear responses. Some points can be raised here: firstly the authors selected the reference ground-motion according to a scenario defined in terms of magnitude and distance only, recent works (Baker and Cornell, 2005) showed that  $\epsilon$  is an important predictor of spectral shape therefore it should be considered when defining a ground-motion scenario; secondly, the median spectral shape corresponds as well to the median responses of SDOF elastic systems, but this statement is not true for non-linear systems, therefore the bias could be caused by a wrong definition of the reference spectrum.

Carballo and Cornell (2000) results also confirmed what other authors already discussed (Shome, 1999): spectral accelerations other than the one at the structure's fundamental frequency may be important to consider for the efficient prediction of non-linear demands. This implies that spectral shape in general may be important. In this study it was found that a primary contributor to the apparent aforementioned biased estimation relies on the global shape. Local variations (i.e. in the close vicinity of the fundamental elastic frequency) of the spectral shape did not seem to have any major effect on the observed bias.

Hancock et al. (2008) compared structural response of an RC frame building obtained by spectral-matched ground-motion time histories with those obtained by sets of records selected according to different criteria. The authors found that the degree of bias systematically decrease as one applies more constraint on the scaling and matching of accelerograms.

Watson-Lamprey and Abrahamson (2006a) investigated the possible bias caused by the use of spectrum-matched ground motions. In this study the authors compared the response of a SDOF non-linear oscillator (bilinear model) when subjected to two different suites of ground motions, obtained by scaling the same set of records (selected in terms of  $M$ ,  $R$  scenario) with two different procedures: the wavelet based procedure proposed by Hancock et al. (2006). and a linear scaling to match  $Sa(T_1)$ . The reference spectrum considered is the one given by Abrahamson and Silva (1997) GMPE. The

authors found that spectrum-compatible ground motions tend to underestimate the response as compared to scaled ground-motions with troughs at the fundamental period and tend to overestimate the response as compared to scaled ground-motions with peaks at the fundamental period. Their results confirm that spectral shape is a factor that should be taken in account in record selection. Furthermore the authors did not investigate if the set of records adopted is well represented by the database used by Abrahamson and Silva (1997) to calibrate their GMPE, hence the bias they observed could be partially due to a lack of consistence between the two scaling procedures.

### ***2.3.4 Number of records required to estimate the median structural response***

Seismic design codes require that the average ordinates of the real spectra have to match the target and not the individual spectra. Bommer and Acevedo (2004) and Bommer and Ruggeri (2002) observed that the number of ground-motion records that the building codes prescribe can vary from 3 to 7, but all the codes allows to use a minimum number of 3 records. Stewart et al. (2001) proposed that if only three records are used, they should be adjusted with one of the techniques presented in Section 2.3.3, to remove their peaks and troughs so that “*the results of structural analyses are not unduly controlled by the particular time histories that are chosen*”. Bommer and Acevedo (2004) suggested that from a seismological perspective, a preferable approach may be to use at least seven records and then use the average response obtained from the structural analysis.

The number of required records is usually estimated by evaluating confidence intervals on the median structural response, i.e. by evaluating the number of records that have to be used in order to obtain a given margin of error on the estimate of the mean structural response. Given a sample of structural response values and supposing they are distributed according to a normal distribution, the best unbiased linear estimator of the mean of the distribution is the sample mean:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N (X_i) \quad (2.20)$$

where  $N$  is the sample size and  $X_i$  is the  $i$ -th observed value. The Mean Squared Error (MSE) of this estimator can be written as:

$$MSE(\bar{X}) = \left( \frac{\sigma}{\sqrt{N}} \right)^2 \quad (2.21)$$

where  $\sigma$  is the variance of the population. If this latter is known and the population is normal the confidence interval at  $C$  % on the population mean is  $\bar{X} \pm z^* \sigma / \sqrt{N}$  where  $z^*$  is the  $(1-C)/2$  critical value for the standard normal distribution. If the variance of the population is unknown, Eq. (2.21) is not exact and  $\sigma$  has to be replaced by the estimated standard deviation (standard error)  $s$ :

$$s = \left( \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1} \right)^{\frac{1}{2}} \quad (2.22)$$

It can be shown that in this case the distribution of the sample mean is no longer normal, instead the sample mean follows a  $t$  distribution. In this case the exact expression for the confidence interval for the population mean is  $\bar{X} \pm t^* s / \sqrt{N}$  where  $t^*$  is the  $(1-C)/2$  critical value for the  $t$  distribution with  $n-1$  degrees of freedom.

When the population is distributed according to a lognormal distribution the most widely used estimate of the central value is the median, and the estimator is defined as the geometric mean which is the exponential of the average of the natural logarithms of the observed values:

$$\hat{X} = \exp \left( \frac{\sum_{i=1}^N \ln(X_i)}{N} \right) \quad (2.23)$$

and the most widely used estimate for the dispersion is the standard deviation of the natural logarithms of the observed data:

$$s = \left( \frac{\sum_{i=1}^N (\ln(X_i) - \ln(\hat{X}))^2}{N-1} \right)^{\frac{1}{2}} \quad (2.24)$$

For small values the above dispersion measure (e.g. lower than 0.3) is approximately equal to the coefficient of variation. This approximation is adopted in many papers about ground-motion selection and scaling. If this is the case, Eq. (2.21) can be used as an approximation and it can be rearranged to calculate the number of records required to

predict some measure of structural response to a certain accuracy, given the standard deviation of the response,  $\sigma$ . The lognormal distribution and the estimators in Eq. (2.23) and Eq. (2.21) are almost universally used in ground-motion estimation (Shome, 1999). Using the aforementioned approximations Shome (1998) suggested that if one wants to obtain an estimate of the median response within a factor of  $Z$  (e.g.  $\pm 0.1$ ) with a 95% confidence, he must use approximately  $N = 4.0\delta^2/Z^2$  records. The authors suggested that typical values of  $\delta$  for unscaled records can vary from 0.43 to 0.75.

Hancock et al. (2008) compared different structural response parameters values obtained by ground-motions selected according to different criteria and confirmed the findings of Shome and Cornell (1998) and Shome (1999). The standard deviation of the response, and hence the number of records required to predict the response to a given level of confidence, is reduced through scaling the ground-motions to the elastic spectral acceleration at the initial period of the structure when compared to using unscaled accelerograms selected purely on the basis of seismological characteristics. However, this study also showed that further reduction in the number of required records may be achieved by scaling the accelerograms to match the target acceleration spectrum on average between periods of 0.25 and 1.2 seconds. The number of records required to obtain an estimate of peak roof drift to  $\pm 10\%$  accuracy (measured on a log scale) at one standard deviation confidence reduces by a factor of about 4 when records are scaled to the initial period of the structure instead of being unscaled. This reduces by a further factor of about 3 when records are selected and scaled to the average spectral acceleration over a range of periods. This result is particularly useful as it reduces the number of records required to estimate measures of peak response (peak roof and interstorey drifts) to this confidence level from 13 unscaled records to a single record scaled to the average spectral acceleration. What is an even more significant result is that when the selected records are matched using RspMatch2005 (Hancock et al., 2006) to the target spectrum at a single damping ratio, or multiple damping ratios, the number of records required to predict the expected response for other damage measures, including end rotations and the Park and Ang damage index (Park et al., 1985), also decreases significantly from the case where spectra are linearly scaled to the average spectral acceleration to just one or two records at the 10% confidence level.

This confirms the findings of Carballo and Cornell (2000) that the non-linear demands obtained by spectrum-matched records present a significant reduction in dispersion, which allows the use of a substantially smaller suite of accelerograms (of the order of 1/4 the size of an actual recorded ground-motion suite) for estimating the median response. The authors suggested that even though the reduction allowed is large, it is

not large enough to use single spectrum compatible records as some seismic codes call for.

As Hancock et al. (2008) observed the numbers of records required to estimate different local damage indices varies. Low-cycle fatigue and absorbed hysteretic energy have the greatest variability and therefore require more records to predict the inelastic response to a given confidence level than measures of the peak response. The absolute numbers of records required to predict the response to a given confidence level is specific to the structure being considered; however, the trends in the numbers of accelerograms required for the various damage measures and scaling and matching approaches are broadly consistent with previous studies and are therefore expected to be generally representative.

## 2.4 Bi-directional loading issues

Planar structural analysis, including seismic input, has typically been performed with simplistic representations of actual systems. In the dynamic analysis case the effects of the orthogonal horizontal component (out-of-plane) of the ground-motion used (in-plane) are considered to be uncoupled from the in-plane analysis. Typically, although within the same plane, the effects of the vertical component are also considered not important, such that this direction of ground motion is not included in the analysis.

In the following the most critical issues in bi-directional excitation problems will be described. The main reference for this section is the paper by Beyer and Bommer (2007) which investigated many aspects involved in selecting and scaling ground-motion records for bi-directional analysis.

### 2.4.1 Direction of loading

In ground-motion databanks usually records are given with the orientation in which they were recorded. The orientation of the horizontal axes can be rotated by a simple matrix multiplication:

$$\begin{pmatrix} a_{x'}(t) \\ a_{y'}(t) \end{pmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{pmatrix} a_x(t) \\ a_y(t) \end{pmatrix} \quad (2.25)$$

where  $a_x(t)$  and  $a_y(t)$  are the components of the recorded ground-motion in the orientation they were recorded and  $a_{x'}(t)$ ,  $a_{y'}(t)$  are the components of the records

after a rotation by an angle  $\theta$ . Values of response spectra can be very different in different orientations as Figure 2.4 shows.

In some papers (e.g. Marušiač and Fajfar, 2005 and Peruš and Fajfar, 2005) on structural response to bi-directional excitation one of the components is multiplied by  $-1$  before being used on the structure; this transformation does not correspond to any rotation therefore has no physical meaning and should not be used.

Different axes can be defined: studies on response spectrum analysis often refer to the principal components of the ground motion (Penzien and Watabe, 1975). Those latter are defined as a set of axes for which the covariance of the ground motion, considered as a random process, disappears. If the three components of the ground motion are considered, the principal directions are defined by the eigenvectors of the matrix Lopez et al., 2006:

$$\sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ & \sigma_{yy} & \sigma_{yz} \\ & & \sigma_{zz} \end{bmatrix} \quad (2.26)$$

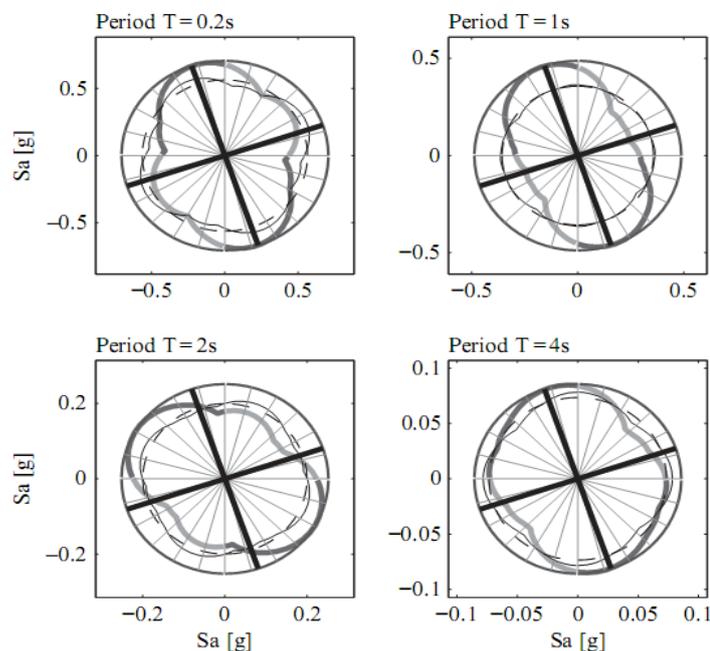
where

$$\sigma_{ij} = \frac{1}{s} \int_0^s a_i(t) a_j(t) dt; i = x, y, z; j = x, y, z \quad (2.27)$$

The orientation of the principal axes is time dependent but Penzien and Watabe 1975 found that the orientation is fairly constant over the strong interval of the ground motion.

It is worth noting that many studies make the approximation that a component of the ground motion is vertical; Lopez et al. 2006 found, by investigating a database containing 97 records, that the average inclination of the pseudo-vertical component is  $11.4^\circ$ , and the modal value of the distribution is between  $5^\circ$  and  $10^\circ$ . The accelerograms in the database were recorded during 25 earthquakes. No events with  $PGA < 0.10$  g were considered. The ground motions were divided in two groups: far-fault ( $R > 15$  km) and near-fault records ( $R < 15$  km). Moment magnitude ranges from 5.4 to 8.2 for the first set and from 6.1 to 7.6 for the second set. The maximum epicentral distance for the far-field records is 187 km. Recording stations were on rock or soil (the authors did not specify the kind of soil).

The authors suggested that a traditional analysis disregarding this inclination may significantly underestimate the critical response (Hernández and López, 2003) for the



**Figure 2.4. Imperial Valley, 1979, Station Delta PEER, 2005: Spectral acceleration as a function of orientation angle  $\theta$  (thick dark and light lines). The outer circle represents the maximum spectral acceleration of one component obtained for all possible orientations of axes. The horizontal and vertical axes correspond to fault-normal and fault-parallel orientation. The thick black radial lines give the orientation of the principal axes with  $\theta_p = 18.8^\circ$ . The solid black line is the geometric mean of the two components as a function of  $\theta$ . The dashed black line represents GMRotD50. After Beyer and Bommer (2007).**

amplification effect of a small inclination of the quasi-vertical component is important in the presence of periods that are close to each other. Those results were obtained by response spectrum analysis only. As far as the acceleration response spectrum is of concern the difference between the vertical and the quasi-vertical component seems to be limited according to Lopez et al. (2006).

Since spectral ordinates vary when the two horizontal components are rotated, the response of the structural system will hence depend on the orientation of the structural axis system relative to the ground motion axes. Therefore, rotating the ground motion will lead to a different structural response. In EC8 it is specified that the seismic action shall “*be applied in both positive and negative directions*”. However, no specifications are made regarding the original orientation. Beyer and Bommer (2007) suggested that it is likely that the components are applied with an arbitrary orientation first and in a second analysis run the polarities of both components are switched. In ASCE 4-98 it is stated that the axes of the ground-motion “*shall, in general, be aligned with the principal axes of the structure*”. However, it is not stated how the ground motion axes

should be orientated. The provisions ASCE 7-05, FEMA 368 and NZS require that the components are applied in the direction that will produce the most adverse effect of the considered parameter. They do not specify, however, how the most critical direction should be established. Normally the direction is found by rotating the ground-motion components at a certain angle interval (e.g.,  $1^\circ$  or  $5^\circ$ ) and analyzing the structure for all orientations of the components. If linear modal analysis is performed the critical directions can be defined according to the procedures proposed by Lopez et al. (2000) and Lopez and Torres (1997).

Recently Rigato and Medina (2007) examined the influence that the angle of incidence of the ground motion has on structural response parameters for a single-storey structure subjected to bi-directional ground motions. Various degrees of inelasticity were considered. The fundamental periods ranged from 0.2 to 2.0 s for both symmetrical and asymmetrical structures. A suite of 39 pairs of horizontal ground-motion records were used. This is a subset of the 80 ground motions used in Krawinkler et al. (2003) and Medina and Krawinkler (2004). The accelerograms are records of ground motions produced by 5 different earthquakes (9 from the Imperial Valley earthquake (1979), 11 from the Loma Prieta earthquake (1989), 14 from the Northridge earthquake (1994), 1 from the San Fernando earthquake, and 4 from the Superstition Hills earthquake (1987)). For a given ground-motion, one of the two horizontal components was classified as either being a major component or a minor component based on its PGA value. The records were applied at various angles of incidence. All major components of the 39 pairs of ground motions were scaled to the same  $S_a(T_1)$  value at the fundamental period of the model, while the minor component was scaled by the same factor used for the major component.

Maximum responses for individual ground motions were found to occur for virtually any angle of incidence (regardless of whether or not a building has a torsional irregularity) and varied with the degree of inelasticity, which implies that inaccurate estimates of structural performance and damage may result if based on ground motions applied at principal orientations alone. The critical angle for a given response parameter varies with fundamental period, model type and the level of inelastic behaviour, and it is difficult to determine a priori like that of an elastic structure. Although an optimal building orientation that minimizes demands for all the parameters considered for a given model cannot be determined explicitly, for a given degree of inelasticity, the average ratio of peak deformation responses based on all angles of incidence to the peak deformation response when the ground-motions are applied at principal building orientations shows stable trends. Generally, these ratios increase with the fundamental

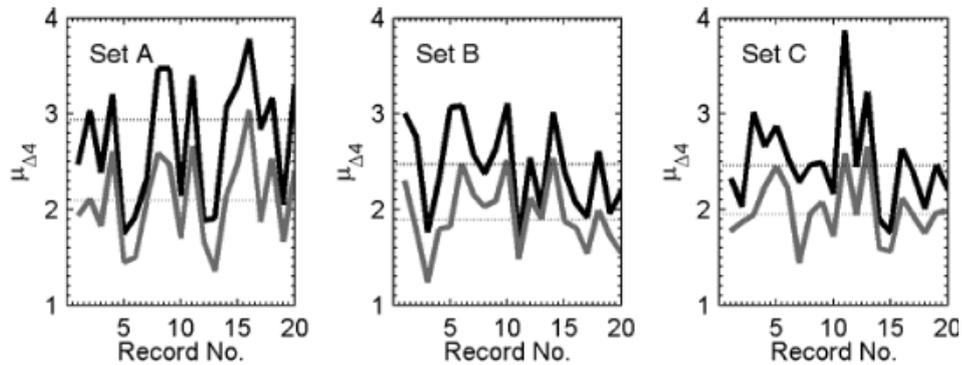


Figure 2.5. Maximum response (black tick line), median response (gray tick line), and their median values (thin lines) as obtained using set of ground-motions selected according to different criteria. After Beyer and Bommer (2007).

period of vibration. These average ratios typically vary between 1.1 and 1.6; however, ratios for individual ground-motions can be as high as 5 for the parameters examined. Beyer and Bommer (2007) studied the structural response of a regular concrete building under bi-directional excitation. When investigating the issue related to direction of the ground-motions the authors considered three different definitions of response (the ductility demand on a structural wall was investigated): the maximum response among all the possible directions of the input; median response from all the possible orientation; random choice for each ground-motion a random direction is chosen. Figure 2.5 shows the results obtained for the different sets of ground-motions. The ratio of these response values varies between 1.05 and 1.68. This great variability in structural response is strictly related to the minimum number of records that should be used to estimate structural response. Beyer and Bommer (2007) concluded that the usual number of records prescribed by codes (i.e. 3 records) it is not sufficient to estimate the maximum median response.

### 2.4.2 Definition of spectral ordinates

A crucial point in bi-directional problems is the definition of the spectral ordinates. Different definitions can be adopted. The following list is adapted from Beyer and Bommer (2006) and Beyer and Bommer (2007):

- x, y: Orientation of the two horizontal components as recorded.
- FN, FP: Components in the fault-normal and fault-perpendicular direction.
- Principal: Components along the principal axes.
- $AM_{xy}$ : Arithmetic mean of the recorded components.

$$Sa_{AMxy}(T) = \frac{Sa_x(T) + Sa_y(T)}{2} \quad (2.28)$$

- $GM_{xy}$ : Geometric mean of the recorded components:

$$Sa_{GMxy}(T) = \sqrt{Sa_x(T) Sa_y(T)} \quad (2.29)$$

- $SRSS_{xy}$ : Square root of the sum of squares of the components x and y:

$$Sa_{SRSSxy}(T) = \sqrt{Sa_x(T)^2 + Sa_y(T)^2} \quad (2.30)$$

- Both: Both horizontal components of a record are considered and treated as two independent realizations of a random process.
- Larger PGA: From the x and y components the one with the larger PGA is chosen and used for all response periods.
- Env: Larger spectral ordinate for the x and y component at each period
- Random: Random choice of one horizontal component from each accelerogram.
- $GMRotD50$ : At each response period, the median value of the geometric mean from all possible orientations of the ground-motion axis system is computed. The orientation corresponding to the median value might vary between different spectral periods (Boore et al., 2006).
- $GMRotI50$ : This ground-motion measure is an approximation of  $GMRotD50$  with a constant axis orientation for all periods, which minimizes the sum of differences between  $GMRotD50$  and  $GMRotI50$  over all considered periods (Boore et al., 2006).
- MaxD: At each period the maximum spectral ordinate from all the possible orientations is determined.
- MaxI: This ground-motion measure is determined following a procedure similar to the one used by Boore et al. (2006) but determining an approximation of MaxD instead of  $GMRotD50$  with a constant axis orientation. The objective function for the angle is slightly different from the one specified by Boore et al. (2006) because it considers the differences between MaxI and MaxD only for periods greater than 0.5 sec.
- Maximum PSA and  $\eta(T, \theta)$ : Maximum pseudo spectral acceleration and ratio  $\eta$  of spectral acceleration for period  $T$  and orientation  $\theta$  to maximum PSA (Hong and Goda, 2007).

Empirical conversion factors between median values of most of the definitions listed and the geometric mean spectrum have been derived by Beyer and Bommer (2006). Watson-Lamprey and Boore (2007) gave adjustment factors that allow converting the

predictions of geometric mean of the ground motion into either the maximum spectral acceleration ( $Sa_{MaxRot}$ ) or the random component of spectral acceleration ( $Sa_{Arb}$ ). The authors provided modification factors for both the mean and the standard deviation of the logarithm of the motions. Moreover the provided conversion factors from  $Sa_{GMRot150}$  to  $Sa_{MaxRot}$  showed that the ratio of  $Sa_{MaxRot}$  over  $Sa_{GMRot150}$  is period dependent, ranging from 1.2 at short periods to 1.35 at long periods. These conversion factors are distance, magnitude, and radiation pattern dependent. The authors suggested that since the dependencies are small for most engineering applications the conversion factors independent of these variables can be used.

Finally Shoja-Taheri and Bolt (1977) defined the concept of Spectral Maximized Record: for each frequency of interest the authors summed in the frequency domain the components of the two horizontal accelerograms orthogonally oriented in the directions  $x$  and  $y$ , i.e. they considered the vector combination of amplitudes and phases for a given azimuthal angle:

$$\begin{aligned} z &= x \cos \theta + y \sin \theta = Z \cos(\omega t + \Phi) \\ x &= A \cos(\omega t + \psi_x) \\ y &= B \cos(\omega t + \psi_y) \end{aligned} \quad (2.31)$$

Eq. (2.31) describes an ellipse therefore it is possible to find the azimuthal angle giving the direction of its major axis:

$$\theta_m = \frac{1}{2} \tan^{-1} \left( \frac{2AB \cos(\psi_x - \psi_y)}{A^2 - B^2} \right) \quad (2.32)$$

and the maximum amplitude  $Z_{max}$ :

$$z_{max}(\omega) = x \cos \theta_m + y \sin \theta_m = Z_{max} \cos(\omega t + \Phi_{max}) \quad (2.33)$$

The authors used this method to generate artificial acceleration time histories compatible with the maximized spectrum.

### 2.4.3 Ground-motion selection and scaling

#### 2.4.3.1 Definition of a target spectrum

The definition of the spectral ordinate of the target spectrum depends on the ground-motion prediction equations which have been used to derive the spectrum. Most GMPEs

use either the geometric mean ( $GM_{xy}$ ) or the envelope spectrum ( $Env_{xy}$ ) of two components with the orientation as recorded (Beyer and Bommer, 2007).

In structural analysis with uni-directional input the spectrum of an arbitrarily selected component and the envelope spectrum of two components have traditionally been widely employed, although some authors pointed out that many studies are not completely consistent (Baker and Cornell, 2006c). When only a single component of the ground motion is needed for the analysis, spectral acceleration is computed only for the selected component at a period equal to the elastic first mode period of the structure, and that is used as the intensity measure. In most cases, no distinction is made between the two components of a ground motion, so using a single component in this case is equivalent to using  $Sa_{Arb}$  as the intensity measure. But frequently the ground-motion hazard analysis has been unwittingly performed with  $Sa_{GM_{xy}}$  to utilize existing attenuation models; this results in an inconsistency when spectral acceleration hazard is coupled with response analysis during performance-based analysis procedures.

Previous studies (Malhotra, 2003; Baker and Cornell, 2006c; Beyer and Bommer, 2007), which have been concerned with bi-directional input for dynamic structural analysis have chosen the geometric mean spectra as the preferred definition of the spectral ordinates of the target spectrum. Beyer and Bommer (2007) suggested that the geometric mean of the spectral ordinates should be used for the following reasons:

- A relatively large set of GMPEs based on this definition is available for use in hazard assessment.
- The measure results in a single spectrum and hence the comparison of target and record spectrum is straightforward.
- The variation of the spectrum with orientation of the ground-motion axes is small.
- Hence, selecting and scaling of records is less sensitive to the orientation of the ground-motion axes than for other definitions.
- The measure is also meaningful in logarithmic space since it corresponds to the arithmetic mean of the logarithmic values. This can be important in structural reliability analysis where logarithmic response parameters are commonly linked to the logarithm of spectral ordinates.
- If the analysis results with bi-directional input are to be compared against results from analysis with uni-directional input, the conversion of the geometric mean spectrum to the spectrum of a single randomly chosen component is fairly straightforward since the median of the two measures are identical while the standard deviation is smaller for the geometric mean than for the single component.

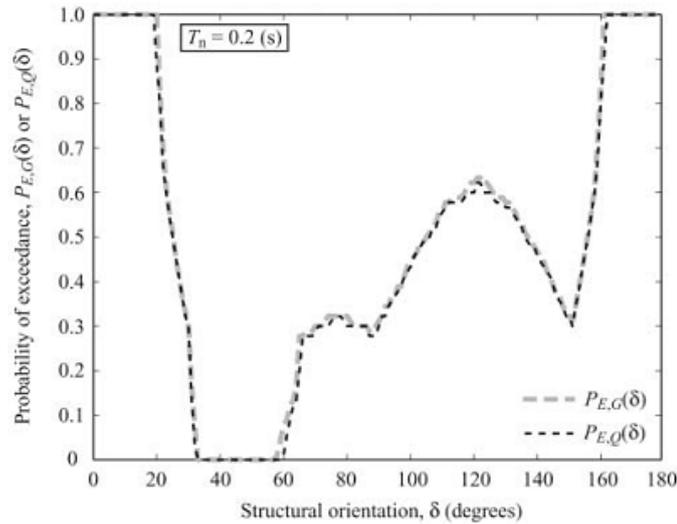


Figure 2.6. Failure probability for a structure designed to resist to  $Sa_{SRSSxy}(\theta)$  or  $Sa_{GMxy}(\theta)$  as a function of the structural orientation. After Hong and Goda (2007).

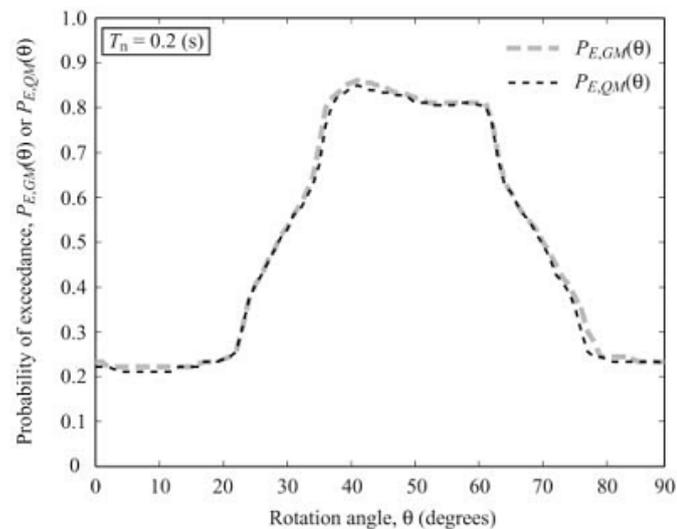


Figure 2.7. Failure probability for a structure of unknown orientation designed to resist to  $Sa_{SRSSxy}(\theta)$  or  $Sa_{GMxy}(\theta)$  as a function of the angle of the recording sensors. After Hong and Goda (2007).

Other definitions such as the envelope spectrum and the SRSS spectrum (which is frequently used in codes) can also be used but it is more difficult to define procedures which are consistent over the process of defining the hazard and selecting and scaling the records. After the geometric mean, the envelope of the two components is probably the most frequently used component definition for GMPEs and hence might be used to define the hazard at a given site.

Hong and Goda (2007) raised some concerns about using the geometric mean or the square root of the sum of squares of the spectral ordinates. The authors showed that it is not possible to infer on the structural failure probability if only  $Sa_{SRSSxy}$  or  $Sa_{GMxy}$  are known: “*if the attenuation relations developed based on  $Sa_{GMxy}(T_n, \theta)$  or  $Sa_{SRSSxy}(T_n, \theta)$  are employed in assessing seismic hazard and in estimating UHS, it is not clear what exactly and actually the return period value entails in the context of codified design and structural safety*”. For example Figure 2.6 shows the failure probability for a structure designed to sustain  $Sa_{SRSSxy}(\theta)$  or  $Sa_{GMxy}(\theta)$  as a function of its orientation  $\delta$  with respect to the ground-motion major axis,  $\theta$  being random and varying uniformly between  $0^\circ$  and  $90^\circ$ . Figure 2.7 shows the failure probability of a structure of unknown orientation  $\delta$  (considered uniformly distributed between  $0$  and  $360^\circ$ ) designed according to  $Sa_{SRSSxy}(\theta)$  or  $Sa_{GMxy}(\theta)$  as a function of the angle  $\theta$  of the recording sensor with respect to the major axis.

As Beyer and Bommer (2007) suggested the geometric mean might be replaced in the future by either *GMRotD50* or *GMRotI50* once new ground-motion prediction equations for these component definitions have been adopted in practice.

#### 2.4.3.2 Ground motion selection

Beyer and Bommer (2007) performed a review of code provisions on ground-motion selection for bi-dimensional analyses. The authors found that most codes do not distinguish between record selection for uni-directional analysis and record selection for bi- or tri-dimensional analysis.

Only the most significant prescriptions are reported here; for a complete description of those provisions see Beyer and Bommer (2007). The use of the same component for both horizontal directions is prohibited by all the guidelines reviewed by the authors. ASCE 4-98 explicitly states that the same component for both horizontal directions must not be used even if the starting time of one component is shifted with respect to the other. ASCE 4-98 requires that geological and seismological settings and local subsurface conditions are appropriate for those of the considered site. It further specifies that the records should be free-field ground motions at the top of the foundation layer and that duration and amplitude of ground-motion parameters such as peak ground acceleration, velocity and displacement shall be representative for the expected ground-motion at the site for that level of hazard. All other codes which were reviewed by Beyer and Bommer (2007) suggested selecting records from earthquake events with geophysical features similar to the design earthquake and in some cases also to the site conditions. The minimum number of records required for structural analysis is three for

all reviewed codes except for ASCE 4-98 which specifies that at least one record should be used unless the structure is sensitive to long-period motion; in this case the minimum required number of records is also three.

ASCE 4-98 states that the components “*shall be statistically independent. (. . .) Two time histories shall be considered statistically independent if the absolute value of the correlation coefficient does not exceed 0.3*”. For structures sensitive to long-period motion “*the input motions in the three orthogonal directions shall, in the frequency range 0.2 to 1.0 Hz, have a correlation coefficient representative of empirical data recorded at sites of similar geotechnical conditions and tectonic environment*”. Beyer and Bommer (2007) argued that these recommendations seem impractical because: *i*) it is not specified whether the correlation coefficient refers to time histories, response spectra, or power spectral density spectra of the components. Moreover, the correlation coefficients of acceleration, velocity, and displacement time histories of the same record are very different. The correlation coefficient further depends on the orientation of the components. *ii*) The authors were unaware of studies which give correlation coefficients as a function of geotechnical and tectonic settings at a site and ASCE 4-98 provides no reference values.

Besides codes very few provision for bi-directional problems have been proposed in the literature. As far as the selection is performed in terms a seismological scenario then there is no difference whether the records are selected for uni- or bi-directional problems. Furthermore if the ordinate of the design spectrum used for the bi-directional analysis is a single component measure, such as for example  $GM_{xy}$ , many of the procedures described in Section 2.2.3.2 can be used. For bi-directional analysis it would also be possible to choose an ordinate definition which retains the two components and to disaggregate the joint ordinate pairs using Vector PSHA 2002. If the ground-motion residual  $\epsilon$  is included in the set of selection parameters it is important that the spectrum of the record matches the definition of the GMPE that is used to determine the ground-motion residual.

If duration is of concern, Beyer and Bommer (2007) stated that defining the duration or effective cycles for in-plane motion is not straightforward. Malhotra (2003) suggested to take simply the arithmetic mean of the durations of each component. Another common way to count cycles of in-plane motion is to consider the SRSS sum of the component time histories. Beyer and Bommer (2007) concluded that neither of the presented definitions for duration or effective cycles of in-plane motion seems to have a robust physical basis.

### 2.4.3.3 *Scaling*

When scaling records to a target design spectrum, it is important that the definition of the spectral ordinates, in terms of treatment of the two horizontal components, is consistent between the design spectrum and the spectrum of the scaled record. However, the underlying definition of the horizontal component of code spectra is not always clear.

Beyer and Bommer (2007) concluded that two options seem possible to match the geometric mean target spectrum:

- Scaling of the geometric mean spectrum of the record to match the target spectrum. The record was scaled to minimize the error between record and target spectrum.
- Scaling of the components individually so that not only the geometric mean but also each component matches the target spectrum as well as possible. This is done by first scaling the geometric mean spectrum of the record to match the target spectrum; this scaling factor is called  $\alpha$ . In a second step one component is scaled by a factor  $\beta$  while the other is scaled by the factor  $1/\beta$ . In this way the geometric mean spectra of the two components remains unaltered.

Carballo and Cornell (2000) considered a simple application to a bi-dimensional problem by considering three sets of ground-motions: the first, adopted as reference, contained ground-motions selected according to a  $M, R$  scenario. It should be noted that only four records were used due to limited availability of computational resources. All the components of the ground motion were scaled by a factor defined such that the spectral acceleration of the larger component of the ground-motion corresponded to some reference values. The second set contained records which have been spectrally matched in order to make each of their components compatible with the median spectra of the respective components of the reference set. Finally the third set was made of ground motions for which both components had been edited in order to achieve compatibility with the response spectrum predicted by the ground-motion predictive equation proposed by Abrahamson and Silva (1997). This procedure introduces an intensity ratio of the two components of 1. All vertical components of the ground motion were compatibilized to the corresponding vertical-motion attenuation relationship; the ratio of horizontal-to-vertical intensity, for the unit horizontal intensity ratio case, was directly taken from there as well. The stronger components (as indicated by the spectral acceleration of interest) of the ground-motion input were applied in one structural direction and the weaker components in the other direction. As for the second suite of records, the authors observed a little reduction in the dispersion of structural response compared to results of the reference set. The authors also observed an

unconservative bias ranging from 10% to 35%. As for the third set, obviously unconservative structural responses were observed. The same way one would not use the same accelerogram of one component of ground-motion simultaneously in both directions, one should not use the same level of spectral target for both directions. The authors suggested the shapes to be considered should be generated by attenuation relationship constructed with different sets of ground-motion records (e.g., fault-parallel and fault-normal). As the authors stated, the sample size adopted is extremely small and the effect of different orientations of ground-motion were neglected therefore the results obtained are hardly extendable. Furthermore, some objections can be raised about the consistency of the procedure: each component of the records of the third set was spectrally matched to the spectrum predicted by the GMPE proposed by Abrahamson and Silva (1997) which is the geometric mean of the spectra of the two components.

#### *2.4.4 Vertical component of the ground-motion*

Typically, although within the same plane, the effects of the vertical component are considered not important, such that this direction of the ground motion is not included in the analysis. Many authors showed that the high-frequency spectral acceleration of the vertical component of the ground motion can be stronger than the horizontal ones in the near-field. Bozorgnia and Campbell (2004) found that the observed and predicted V/H spectra are strong functions of natural period, source-to-site distance, and local site conditions; and a relatively weak function of earthquake magnitude and faulting mechanism. The behaviour of the V/H spectra with distance is different for firm soil than for stiffer soil and rock deposits, at firm soil sites approaching a factor of 1.8 at short periods, close distances, and large magnitudes.

Based on the observed and predicted behaviour of the V/H spectra, Bozorgnia and Campbell (2004) proposed a tentative simplified model for estimating V/H spectral ordinates for practical engineering applications (see Figure 2.8). A reasonable agreement between the simplified and predicted V/H spectral ordinates was found over a wide range of seismological parameters. A simple procedure for generating a preliminary vertical design spectrum was also proposed.

Similar results were obtained by Lopez et al. (2006). The authors calculated the response spectra in the principal directions for 97 ground-motion records (see Figure 2.9). The pseudo-vertical to horizontal ratio can be greater than 1 for short distances and high frequencies.

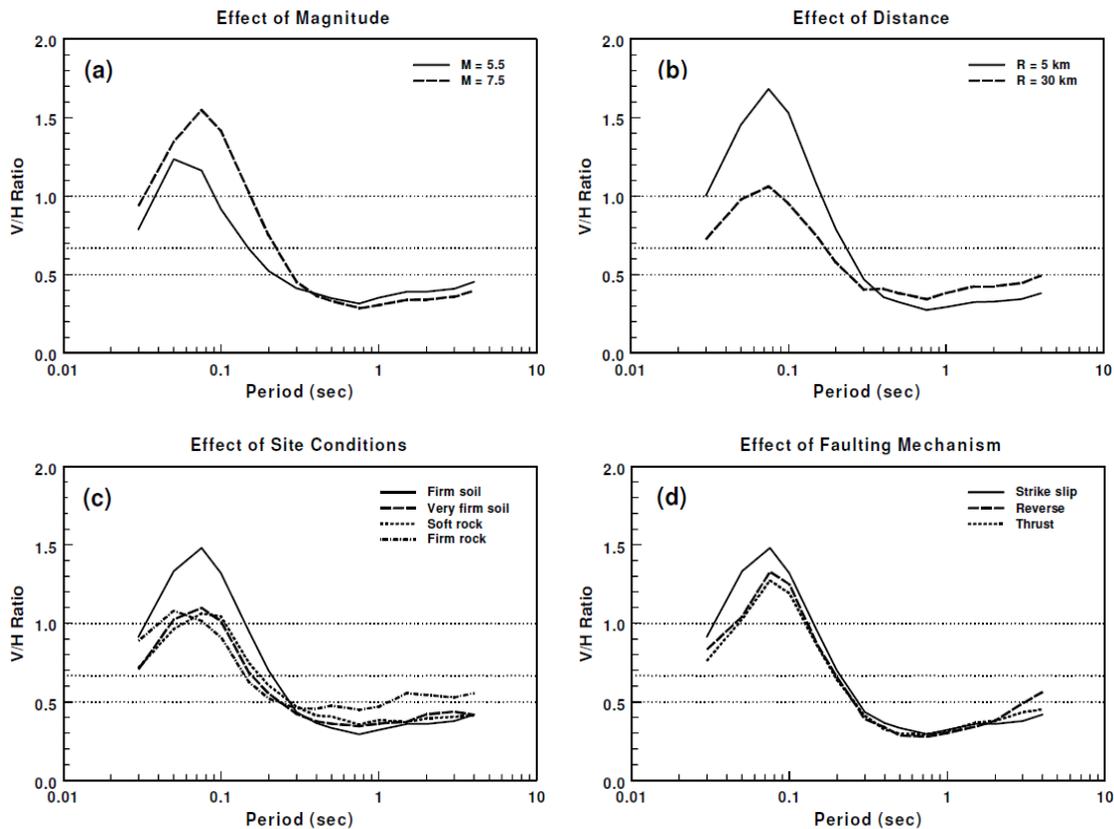


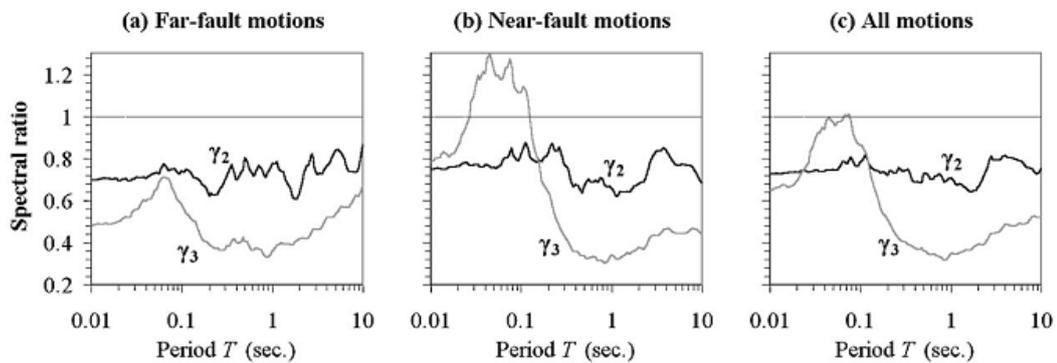
Figure 2.8. V/H ratio as predicted by the ground-motion prediction model by Bozorgnia and Campbell. Effects of magnitude, distance, site conditions and faulting mechanism. After Bozorgnia and Campbell (2004).

## 2.5 Conclusions and possible future work

### 2.5.1 Scenario based selection

Disaggregation has to be used to define reference scenarios in terms of magnitude, site-fault distance, epsilon etc. Disaggregation results can be not easy to manage if a single scenario is not strongly predominant. In this case multiple scenarios have to be considered. Particular care should be taken if logic trees are adopted to combine different ground-motion prediction equations and seismic sources.

*Magnitude.* It has been shown by many authors that magnitude has a strong influence on spectral shape and duration. This parameter should therefore be considered in ground-motion selection. Some researchers stated that magnitude matching is not important if records are scaled to spectral acceleration. This statement has recently been shown to be incorrect, in fact if spectral shape is not considered in selection structural response can be biased.



**Figure 2.9.** Ratios of the acceleration response spectra in the principal directions of ground-motions.  $\gamma_2$  is the ratio of the spectral component in the principal direction 2 and the spectral component in the principal direction 1.  $\gamma_3$  is the ratio of the spectral component in the principal direction 3 and the spectral component in the principal direction 1. After Lopez et al. 2006.

*Distance.* A lack of matching in terms of distance seems to be of less importance and it can be compensated by linear scaling. This procedure is strictly depends on which parameter is used to define the intensity of the strong motion and therefore to define the reference for scaling. If multiple spectral values are used, high scaling factors can be allowed, on the other hand if a more “local” parameter such as PGA, PGV,  $Sa(T_1)$  is used it can give biased results. The literature on this topic is limited because even if many authors studied the correlation of structural response with different ground-motion parameters few of them investigated the possible presence of bias in structural response. In fact the aim of these works was mainly trying to find the best descriptor of the intensity ground-motion effects on building. Furthermore in the majority of these studies very simple structures and small ground-motion sets were used.

*Epsilon.* Recently it has been shown that if the intensity of the ground-motion is measured via  $Sa(T_1)$ ,  $\epsilon$  is a strong indicator of spectral shape. Therefore if this intensity measure is used spectral shape can be included in the selection by using  $\epsilon$ .

### 2.5.2 Selection based on UHS

Most of the current building codes prescribe to select sets of ground-motions such as their “mean” response spectrum matches the UHS given by the codes. This is probably the state of the practice in many engineering companies. The aim of these procedures is trying to define a set of records giving a structural response with the same level of probability of occurrence of the UHS. Some critical aspects can be highlighted and could be investigated:

- *Definition of the reference spectrum.* In particular many codes do not contain information on which definition of spectral is used in their spectra (geometric mean, random component etc...). Furthermore, given the non-linear behaviour of

the structures, does a set of records matching a given UHS give a mean structural response with the desired probability of occurrence?

- *Components of the ground-motion and direction.* Many codes do not give instructions on which of the two horizontal components should be used and usually ground-motions are used according to the recording-axes: no rotation is considered. This is particularly important since structural failure probability can be strongly dependant on ground-motion orientation if structures are designed according to definitions of the response spectrum not dependant on orientation (e.g. geometric mean).
- *Definition of matching.* Which period range should be considered? Which tolerance should be allowed? Can the median structural response be affected by the dispersion of the ground-motions around their mean? And therefore, should this dispersion be controlled?
- *Number of ground-motions.* Many codes allow to use 3 records only and to consider the maximum structural response. Many authors suggested that this practice should be abandoned.
- *Limits on scaling.* Many codes do not give any instructions about scaling.
- *Records from the same event.* No prescription are usually given about the number of records obtained by the same event that can be used. Does this affect structural response?
- *Magnitude and distance.* Many codes do not prescribe limits on magnitude and distance. How does this affect structural response?

### ***2.5.3 Selection procedures based on non-linear structural response***

Some authors proposed to adopt simple non linear SDOF systems as proxies of more complex structures and to select ground motions with characteristics similar to those of the records giving a desired value of the proxy response. The main issue in these methods is related to the definition a simple system able to mimic the dynamic behaviour of more complex structures.

### ***2.5.4 Selection procedures for probabilistic approaches.***

Recently Performance Based Earthquake Engineering procedures received great attention by many researchers. The aim of these procedures is to evaluate the probability of reaching some given limit states. To do so structural response must be combined with seismic hazard in a fully probabilistic framework. The best state of the practice is to use an Intensity Measure as interface variable, i.e. expressing both hazard and structural response in terms of this variable. Therefore the selection and scaling of ground-motions is strictly related to this interface variable. The most attractive possibility would be selecting ground motions in terms of the IM only, for this to be done the IM should have some properties as: efficiency, sufficiency, hazard computability, scaling robustness. For

many years  $Sa(T_1)$  has been used as IM and it has been stated that ground-motion could be select with little regard for magnitude and distance, recently it has been shown that doing so can produce biased results. At the state of the practice no IM seem to have the required properties to allow for a “IM” based selection. The current state of the best practice is to use  $Sa(T_1)$  as IM and select ground motions according to a Conditional Mean Spectrum.

Finally, in fully probabilistic procedures the variability of structural response is as important as its mean value, because the probability distribution of the structural response has to be estimated. Some procedures have been proposed to define set of ground-motions with this aim but much research is still required in this field.

### 2.5.5 *Bi-directional problems*

- *Definition of the reference spectrum and of scaling procedures.* Definition of the reference spectrum is a crucial aspect in bidirectional problems. The most widely adopted definition is the geometric mean of the two horizontal components. This definition does not account for all the possible orientations of the ground motion. Therefore criteria based on GMRotD50 or GMRotI50 could produce more stable results.
- *Effects of the vertical component of the motion:* usually the vertical component of the motion is not included in structural analyses. Recently it has been shown that the high frequency content of this component can be particularly important. No criteria have been proposed in the literature about the ground-motion selection criteria in this case. This is particularly important if disaggregation is used, because the controlling scenarios could be different for the vertical and the horizontal ground motion.



# 3 Accelerogram selection and scaling procedures for estimating the distribution of drift response<sup>2</sup>

## 3.1 Introduction

For the majority of applications in which the effects of earthquakes on structures are considered, seismic actions are represented in the form of pseudo-acceleration or displacement response spectra. There are, however, many situations for which the specification of structural actions via a response spectrum is deemed insufficient. In such cases, the structural response for a given earthquake scenario is estimated by subjecting the structure to acceleration time-histories that are compatible with the scenario in question. These time-domain analyses are far more computationally expensive than response-spectrum-based alternatives but with ongoing improvements in the power of desktop computers commonly encountered in design offices, this issue is becoming less of a constraint. Recent work, such as that of Hancock et al. (2008), has demonstrated that robust estimates of the median structural response of a typical MDOF structure may be obtained with the use of very few accelerograms if their spectra initially have an appropriate shape and the records are then scaled and adjusted with wavelets to ensure that they have particular characteristics. The study of Hancock et al. (2008), is just one recent example of the numerous approaches that have been proposed for selecting, scaling and modifying real accelerograms in order to obtain robust estimates of the median structural response. Approaches based upon linearly scaling or wavelet adjusting real accelerograms tend to be favoured now over earlier procedures that were proposed for synthesizing spectrum-compatible accelerograms (Gasparini and Vanmarcke, 1976; Boore, 2003) as these earlier approaches result in accelerograms with unrealistic energy contents (Naeim and Lew, 1995; Bommer and Acevedo, 2004). Whether or not accelerograms are obtained through scaling or adjustment of real

---

<sup>2</sup> This Chapter has been written with the supervision of Prof. J.J. Bommer and Dr. P.J. Stafford, Department of Civil and Environmental Engineering, Imperial College, London, UK.

records or via synthetic means, the objective once the suite has been obtained is generally to determine a central estimate, such as the mean or median, of the structural response that may then be used for design purposes. As a result, the research that has been conducted thus far has principally focused upon how to estimate the median response most efficiently as well as on how to identify the characteristics of records that lead to biased estimates of this response measure. There are numerous applications for which knowledge of a central estimate of the response is not sufficient and for which one requires an estimate of the full distribution of the structural response. Generally, these applications are related to the assessment of existing structures whereas the approaches focusing on estimating the median response are primarily geared towards the specification of loading for structural design, in fact the great majority of building codes require to estimate the median structural response through an ensemble of ground motions (although, ultimately, knowledge of the distribution of response may have implications for the specification of design requirements in codes). For example, in earthquake loss assessment one must not only consider the potential damage associated with the expected response, but also the damage due to the full range of possible responses that may be experienced under a particular scenario. Likewise, when undertaking structural assessments one should consider the utility of alternative options for retrofit or demolition on the basis of costs estimated from consideration of all possible levels of future damage that might be experienced by the structure in question. Intuitively, one would expect that more time-history analyses are required in order to estimate the characteristics of the full distribution of a particular response measure than are required to obtain an estimate of the median response. The purpose of this study is to outline an efficient approach via which the full distribution of drift response may be approximated using estimates of the first two moments of the distribution. While the numbers of records that are required under this approach are greater than that required to estimate the median alone, it is shown that the numbers of records are not as high as one might imagine and that it is plausible to run these analyses in a design-office environment.

## **3.2 Overview of the procedure**

### *3.2.1 Definition of the reference distribution*

The objective of the research presented herein is to identify methods via which the distribution of drift response may be estimated for a given seismological scenario. In order to evaluate the performance of any approach one must first establish a basis for

comparison. In the present case the true distribution of drift response is unknown and we must therefore define some reference distribution. To this end the approach adopted by Hancock et al. (2008), is implemented whereby a structure is analyzed under the action of a very large number of unscaled natural accelerograms.

The drift responses that are obtained during these analyses are then compiled and empirical models that relate the calculated drift values to common seismological parameters are derived via regression analysis. An empirical model for estimating spectral accelerations over a broad range of periods is also derived. Given these models one may take an earthquake scenario, defined during a seismic hazard analysis, and obtain the distribution of spectral acceleration or drift values that are associated with this scenario. Provided that the dataset used for the development of the empirical models is large and that the seismological scenario being considered is not beyond or near the magnitude-distance range of applicability of the models (Bommer et al., 2007) it is reasonable to assume that the distributions thus obtained will be good approximations to the true, but unknown, distributions. Although this assumption cannot be validated, most empirically-based methods within engineering seismology and earthquake engineering are founded upon similar assumptions.

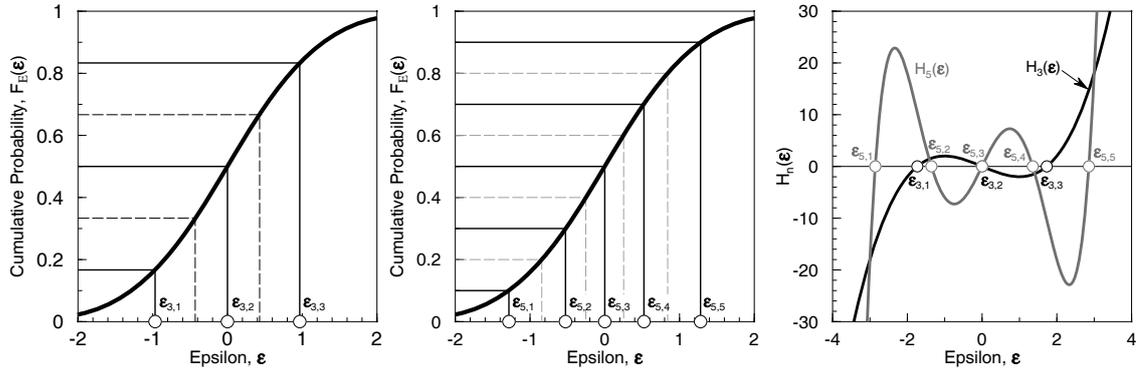
A fundamental assumption of the regression analyses that have been conducted and are presented herein is that the ordinates of spectral acceleration and the inter-storey and roof drifts are lognormally distributed. This assumption has been used by many researchers (Bazzurro et al., 1998; Shome et al., 1998; Cornell et al., 2002; Baker and Cornell, 2006b; Stoica et al., 2007) and is well supported by the distributions of residuals that are observed following the regression analyses (see 0). This is of particular importance in the present study as by assuming a lognormal distribution during the regression analysis we are also making an implicit assumption regarding the nature of the reference distribution. Making the assumption that the drift values are lognormally distributed is also very convenient as it means that the distribution of drift may be fully described by the first and second moments of the drift only, i.e., if one can obtain robust estimates of the mean and standard deviation of the logarithmic drift then the complete distribution is known (Shome et al., 1998; Cornell et al., 2002).

### ***3.2.2 Description of the procedure***

The large number of time-history analyses that are required for deriving the empirical relationships, and hence the reference distributions, may also be used to infer relationships among the ground-motion parameters and measures of structural response. For example, if one plots the calculated inter-storey drifts against the spectral

accelerations at the fundamental period of the structure one is able to observe very strong dependencies. Such dependencies are useful for identifying functional relationships among different parameters such as spectral acceleration and roof drift, and the existence of such trends implies that a robust model for the median drift may be obtained from the median spectral acceleration. For the purposes of the present study, however, the distributions of spectral acceleration and drift values about these median values are more important. In particular if it is possible to demonstrate that a higher-than-average, or positive epsilon (Shome et al., 1998; Cornell et al., 2002) spectral acceleration also leads to a higher-than-average drift in a systematic manner (this is not obvious, given the non-linear behaviour of structures under earthquake loading) then we may be well placed to relate the distribution of drift values to the distribution of spectral acceleration values.

The above reasoning forms the basis of the approach taken in this study. First a large number of time-history analyses are conducted and the results of these analyses are used to derive empirical models for inter-storey and roof drift. These models are then analysed in order to look for correlations that exist among the residuals of the models for the drifts and the residuals of the model for spectral acceleration at the initial fundamental period of the structure. If strong correlations exist then it is possible to relate particular levels of spectral acceleration to particular levels of drift response. Hence, the rationale is that if one linearly scales a record to a particular level above or below the median spectral acceleration for a given scenario, then one may anticipate that the drift response will also be above or below the median drift response to a similar degree, as measured in units of standard deviation (assuming a positive correlation in this case). If such inferences are possible then all that remains is to identify the optimal way of mapping levels of spectral acceleration into levels of drift and hence recovering the distribution of drift. Two different approaches are taken for mapping the distribution of logarithmic spectral acceleration into the distribution of logarithmic drift. It must be noted that although two approaches are investigated herein, there are many other alternative approaches for representing a continuous distribution by a discrete approximation (Cornell et al., 2002; Baker and Cornell, 2006b). The first approach that is taken is to subdivide the realm of possible logarithmic spectral acceleration values into ranges that have an equal probability of occurring. This is achieved by subdividing the range of cumulative probabilities into equal intervals and then using the inverse cumulative distribution function to obtain the logarithmic spectral acceleration values that correspond to the centres of these probability intervals. The logarithmic spectral acceleration values are identified by their corresponding epsilon values, i.e., the number



**Figure 3.1.** Three (a) and five (b) point approximations of a normal probability distribution accordingly to the equal-probability rule. Hermite polynomials (c), and their zeros used for three and five point approximation procedure proposed by 1983.

of standard deviations above the median value. This approach is demonstrated in the left and centre panels of Figure 3.1, where three- and five-point approximations to the standard normal distribution are shown. Note that although the ordinate is subdivided into equal intervals the resulting epsilon values on the abscissa are not equally spaced for four-point approximations and higher. This method of approximating a continuous distribution is rather common due to its simplicity (Keefer and Bodily, 1983), but it is known to have some deficiencies. For example, Miller and Rice (1983) demonstrate that while the mean value of both the cumulative and discrete distributions will be the same, the even moments of the discrete distribution should be less than those of the continuous distribution. The second approach adopted is more sophisticated and makes use of an approximation based upon Gauss-Hermite quadrature.

This latter approach is based on the results of Miller and Rice (1983) who have demonstrated that the moments of a continuous distribution  $f_E$  may be exactly replicated by a discrete distribution consisting of epsilon values (nodes) and associated probabilities (weights) that correspond to the nodes and weights used in Gauss-Hermite quadrature. This approach approximates the integral of the product of a function  $g(\epsilon)$  and a weighting function  $w(\epsilon)$  by evaluating  $g(\epsilon)$  at several values  $\epsilon_i$  of  $\epsilon$ , and computing a weighted sum of the results:

$$\int_a^b g(\epsilon)w(\epsilon)d\epsilon = \sum_{i=1}^n w_i g(\epsilon_i) \quad (3.1)$$

where  $w_i$  is the  $i$ -th weight. To establish the correspondence between the numerical integration formula and a discrete approximation of a probability distribution, Miller and Rice (1983) associated the distribution with the weighting function,  $w(\epsilon)$ , and the

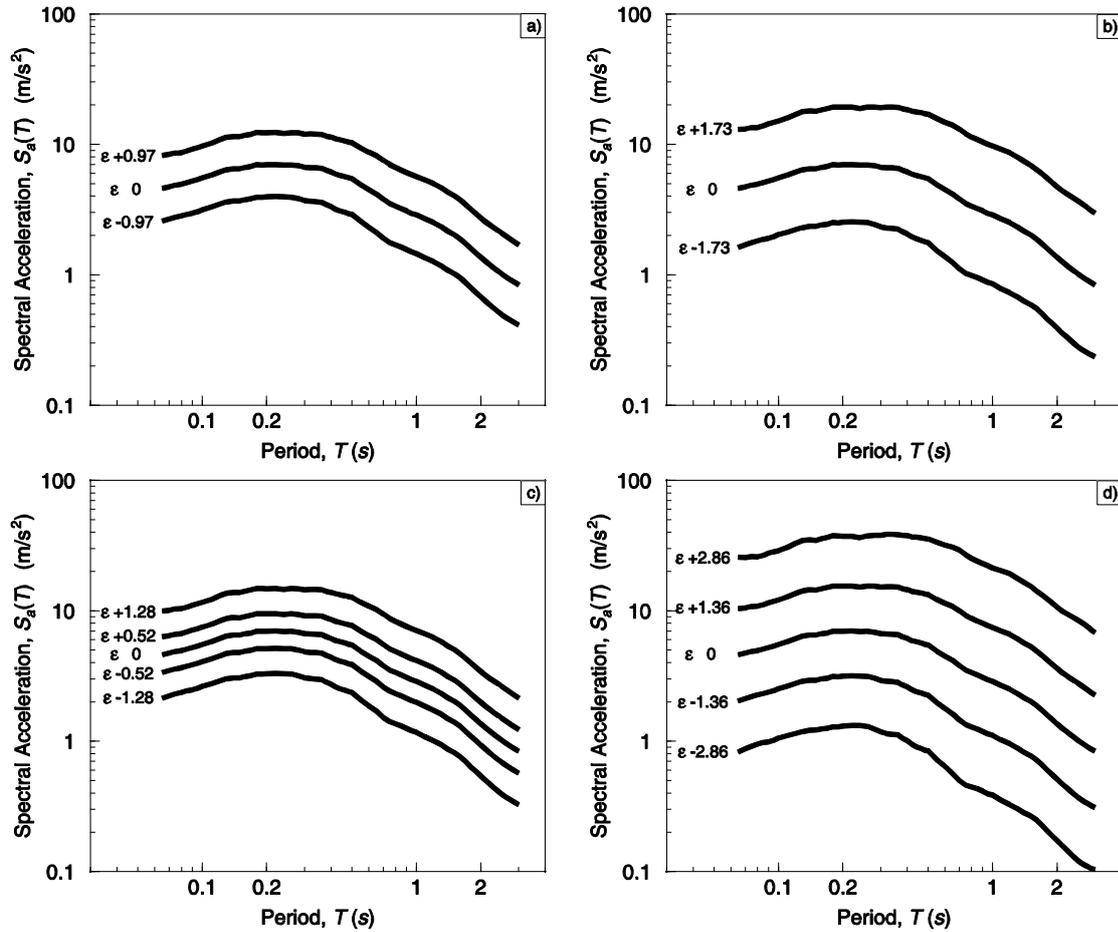
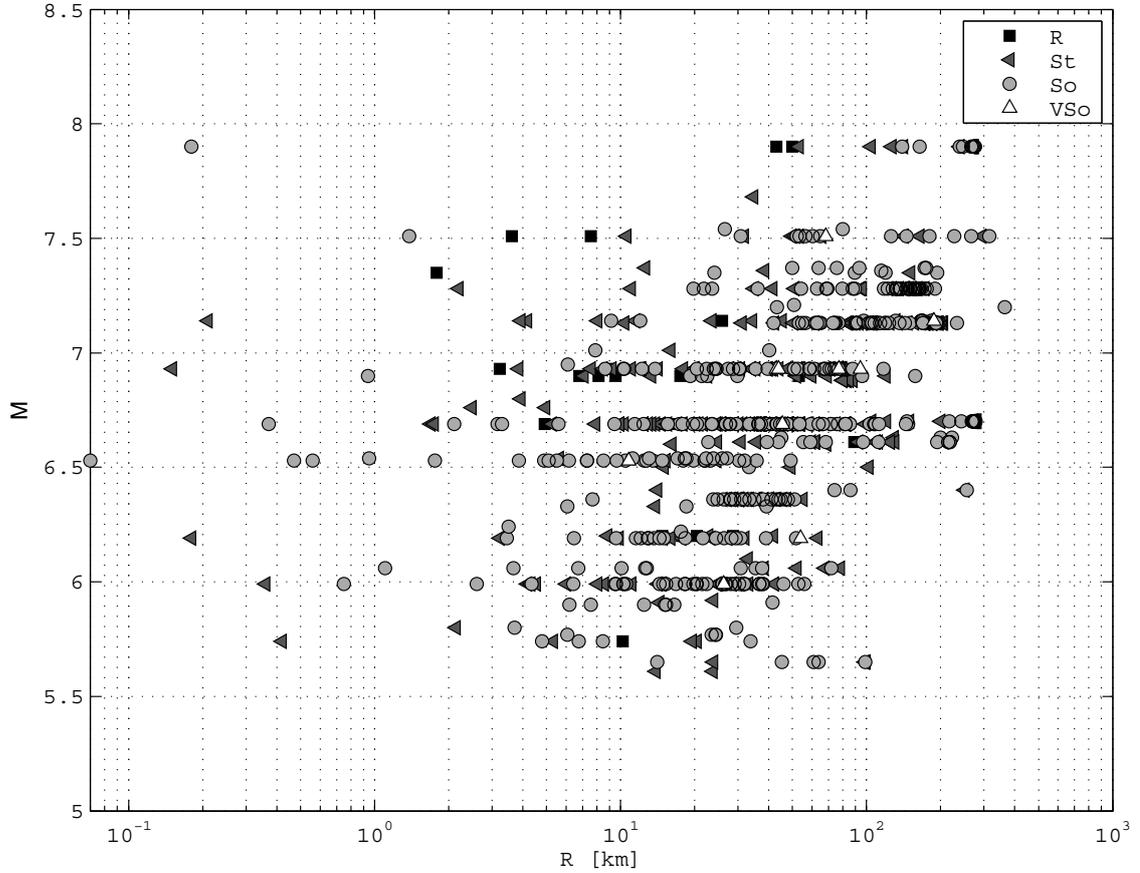


Figure 3.2. Reference spectra defined according to a) equal-probability three point approximation, b) Gauss-Hermite quadrature three point approximation, c) equal-probability five point approximation and d) Gauss-Hermite quadrature five point approximation.

probabilities,  $p_i$ , with the weights  $w_i$ . Furthermore they approximated  $g(\varepsilon)$  by a polynomial, and chose  $\varepsilon_i$  and  $p_i$  (or  $w_i$ ) to provide an adequate approximation for each term of the polynomial. Thus, one must find a set of values  $\varepsilon_i$  and probabilities  $p_i$  such that:

$$\langle \varepsilon^k \rangle = \int_{-\infty}^{+\infty} \varepsilon^k f_E d\varepsilon = \sum_{i=1}^N p_i \varepsilon_i^k \quad k = 0, 1, 2, \dots \quad (3.2)$$

where  $\langle \cdot \rangle$  stands for expected value. It is possible to demonstrate that the  $\varepsilon_i$  values to be used correspond to the roots of Hermite polynomials (Miller and Rice, 1983). The  $n$ -th order Hermite polynomial may be defined as



**Figure 3.3. Moment magnitude – distance distribution of the dataset used.**

$$H_n(\varepsilon) = (-1)^n e^{\varepsilon^2/2} \frac{d^n}{d\varepsilon^n} e^{-\varepsilon^2/2} \quad (3.3)$$

which leads to the following expressions for the third- and fifth-order polynomials

$$H_3(\varepsilon) = \varepsilon^3 - 3\varepsilon \quad (3.4)$$

$$H_5(\varepsilon) = \varepsilon^5 - 10\varepsilon^3 + 15\varepsilon \quad (3.5)$$

These polynomials along with their roots are plotted in Figure 3.1. The roots of these polynomials correspond to the epsilon values that should be used to define target levels for scaling the accelerograms. The results that are obtained from the structural analyses conducted using records scaled according to these epsilon values are then combined with the corresponding weights to obtain estimates of the moments of the distribution. The moments of the logarithmic drift may be expressed as:

$$\langle (\log_{10} \delta)^k \rangle = \sum_{i=1}^n p_i (\log_{10} \delta_i)^k \quad \text{for } k = 0, 1, 2, \dots \quad (3.6)$$

Where  $\langle (\log_{10} \delta)^k \rangle$  is the  $k$ -th moment of  $\log_{10}(\delta)$ , i.e., the mean and variance of  $\log_{10}\delta$  correspond to  $k = 1$  and  $k = 2$  respectively, and  $p_i$  are weights which are given by Miller and Rice (1983).

From a statistical point of view this second approach should perform better than the simpler first approach as the moments obtained via Eq. (3.6) using the discrete nodes and weights are identical to the moments of the equivalent continuous distribution (Spiegel and Liu, 1999). Figure 3.1 presents examples corresponding to three- and five-point approximations to normal distributions and these approximations are those that we opt to use for the remainder of the analysis. However, it should be noted that there is no particular reason for having chosen to use three- and five-point approximations.

Whether or not superior performance may be achieved using alternative orders for the polynomials remains to be tested. Once the epsilon values have been defined, these may be used in conjunction with the empirical model for spectral acceleration in order to define either three or five target spectra that correspond to a particular seismological scenario. Examples of three and five target spectra obtained according to the two aforementioned criteria are given in Figure 3.2. Note that in this study the common assumption that the spectral ordinates are fully correlated is adopted, i.e., for a given target spectrum the ordinates at every period are the same number of standard deviations away from the median spectral level. This is a widely adopted assumption, the first attempt to consider correlation among spectral values are being made by Baker et al. (2006b) and Jayaram et al. (2008) (.).

The scaled accelerograms whose spectra best match these multiple target spectra, as quantified by the root-mean-square difference between the logarithmic target and candidate spectra over a range of periods, are then used as inputs into time-history analyses and the drift responses of the structure under these records are obtained. The statistics of the drift values are then compared to the statistics of the reference distribution defined by the empirical models derived from the time-history analyses conducted on the unscaled accelerograms.

### 3.3 Case study

The principle of the procedure is very simple but it relies heavily on the assumption that the characteristics of ground-motions that lead to high spectral acceleration values also

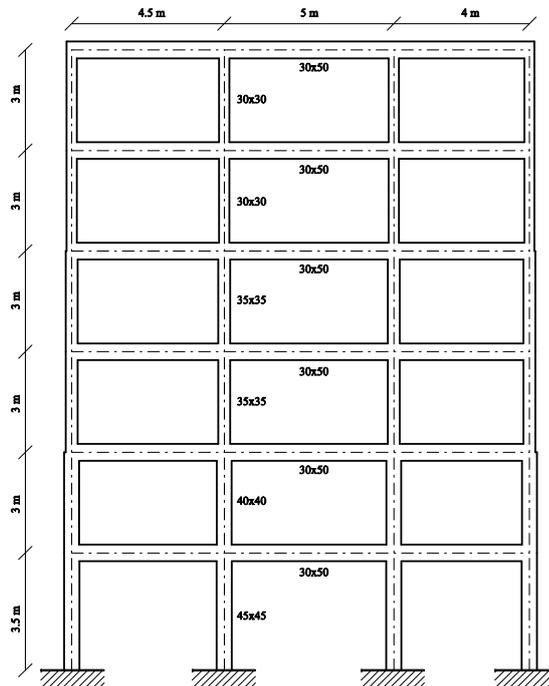


Figure 3.4. Geometry of the case study frame structure considered.

lead to high drift values. In order to test this assumption an example is necessarily required. The particular application of the proposed procedure is outlined in the following sections.

### 3.3.1 Ground-motion dataset

Dataset Accelerograms were selected from a subset of the records in the Next Generation of Attenuation (NGA) project database (Power et al., 2006; Power et al., 2008) . A total of 1666 accelerograms were used for the analysis. This subset was defined by excluding all records from the Chi-Chi earthquake sequence, to prevent its 1813 accelerograms dominating the results, as well as any records with only one horizontal component and records for which appropriate metadata were not available (moment magnitude,  $M_w$ , Joyner Boore distance,  $R_{JB}$ , and shear wave velocity over the topmost 30 m,  $V_{S30}$ ). All of the selected records have a maximum usable period of at least 3 seconds (Boore and Bommer, 2005), i.e. records where either component has a high-pass filter frequency greater than 0.33 Hz (less than 3s period) have been excluded. The site classification is based on the average shear-wave velocity over the upper 30 m ( $V_{S30}$ ). The magnitude-distance distribution of the dataset is given in Figure 3.3. For a more straightforward understanding of the data distribution in Figure 3.3 the record are grouped into three categories according to their  $V_{S30}$  value. The grouping follows Boore

and Joyner (1982) and Ambraseys et al. (1997) as follows: sites with  $V_{s30}$  higher than 760 m/s are classified as rock (R), sites with  $V_{s30}$  between 360 m/s and 760 m/s are classified as stiff soil (St); sites with a  $V_{s30}$  between 360 m/s and 180 m/s are classified as soft soil (So) and those with  $V_{s30}$  lower than 180 m/s as very soft (VSo). The style-of-faulting was not taken into account because tests on regression results showed it had little significance. A detailed list of the ground-motion database used is provided in Appendix A.

### 3.3.2 Structural model and nonlinear analysis

A six-storey, three-span reinforced-concrete frame building is considered in this study. The structure has an initial fundamental period of 0.93 seconds (after the application of gravity loads). The time-domain analyses take into account both geometric nonlinearity and material inelasticity. Structural members are modelled using force-based fiber elements (Spacone et al., 1996b; Spacone et al., 1996c) which have the advantage, over displacement-based approaches, of satisfying equilibrium in the nonlinear range, alleviating the need for mesh refinement. The confined concrete constitutive behaviour is modelled using the model proposed by Mander (Mander et al., 1989), the unconfined concrete constitutive behaviour is modelled using the model developed by Saenz Pinto (Ceb-Fip, 1993) and the steel constitutive behaviour is modelled using the Menegotto–Pinto’s model (Menegotto and Pinto, 1973). A confinement factor of 1.2 was considered for the confined concrete. All nonlinear time-history analyses were conducted using the software OpenSees (McKenna et al., 2000). Figure 3.5 gives push-over curves for the columns at the different floors of the structure (the curves were obtained individually for each column). Gravity and live loads were applied before performing the push-over procedure.

**Table 3.1. Reinforcements in the members of the case study structure.**

Member	As [cm <sup>2</sup> ]	As' [cm <sup>2</sup> ]
Columns		
I Storey	6.03	6.03
II Storey	6.03	6.03
III Storey	6.03	6.03
IV Storey	6.03	6.03
V Storey	4.02	4.02
VI Storey	4.02	4.02
Beams		
All	8.17	3.08

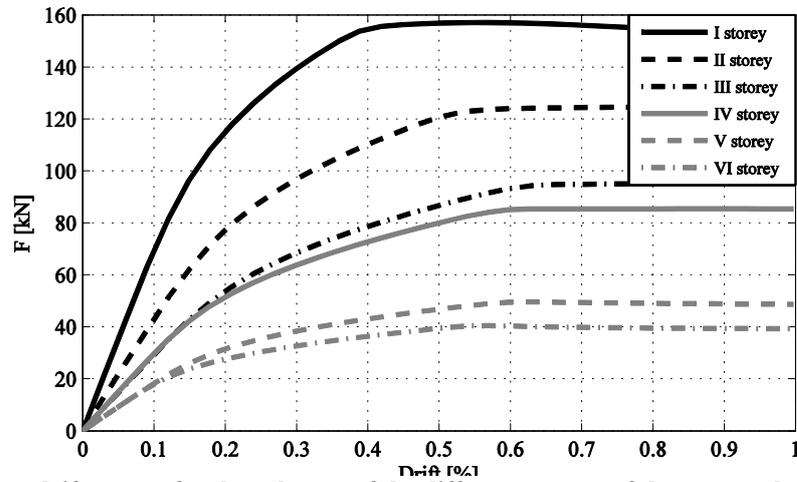


Figure 3.5. Force-drift curves for the columns of the different storeys of the case study structure.

### 3.4 Regression Analyses

In this study, regression analysis is used to derive predictive equations for spectral acceleration, for interstorey drifts and for roof drift. According to the results of some preliminary statistical tests performed the same statistical model was used for spectral acceleration and interstorey drifts:

$$\log_{10}(Y) = c_1 + c_2 M_w + c_3 M_w^2 + (c_4 + c_5 M_w) \log_{10} \left( \sqrt{R_{JB}^2 + c_6^2} \right) + c_7 \log_{10} (V_{S_{30}}) \quad (3.7)$$

Here  $\log_{10}(Y)$  is the base ten logarithm of the parameter to be regressed (i.e. spectral acceleration at different periods and interstorey drifts);  $M_w$  is moment magnitude (Hanks and Kanamori, 1979);  $R_{JB}$  is the closest distance to the surface projection of the fault rupture, as proposed by Joiner and Boore (1981);  $V_{S_{30}}$  is the shear-wave-velocity over topmost 30 m.

The functional form adopted in this study is similar to that previously used by Hancock et al. (1981) but with two key differences being the inclusion of the average shear-wave velocity over the uppermost 30 m at the site as an additional predictor variable and the inclusion of a magnitude dependant geometric spreading. The functional form adopted for this study also omits the previously included dummy variables for style-of-faulting. The style of faulting is not included as the regression analyses did not lead to statistically significant coefficients for the dummy variables that we used to account for style-of-faulting. It is well known that style-of-faulting influences the amplitude of ground motions, but one might argue that this effect influences spectral shape. While in

some models the shape will be modified slightly by the style-of-faulting factors it is certainly not a significant effect and it is not included in all of the models. For example, the Abrahamson and Silva (2008a) NGA model, that can be considered the most comprehensive model available has a constant factor for style-of-faulting for normal events across all periods from PGA to 10 seconds. Although their factors for reverse events do vary with period, the effect is very small and hardly influences spectral shape. Furthermore another key difference may be found in the adopted structure for the variance. The total standard deviation,  $\sigma_T$ , of each empirical model may be decomposed into three independent components: the inter-event standard deviation,  $\sigma_E$ ; the intra-event standard deviation,  $\sigma_A$ ; and the inter-component standard deviation,  $\sigma_C$ , which must be considered when both horizontal components of a ground-motion recording are used in the regression analysis, as is done in this study (Boore et al., 1997; Douglas, 2003; Boore and Bommer, 2005; Baker and Cornell, 2006c). As these three components of variability are independent, the total standard deviation may be represented as:

$$\sigma_T = \sqrt{\sigma_E^2 + \sigma_A^2 + \sigma_C^2} \quad (3.1)$$

The coefficients and the variance components of the models were obtained using the nonlinear mixed effects procedure of Lindstrom and Bates (1990) as implemented in the statistical computing package R (Pinheiro and Bates, 1995; R Development Core Team, 2008).

### ***3.4.1 Spectral accelerations and PGA***

**In order to define the reference acceleration spectra, regression analysis was used to develop ground-motion prediction equations (attenuation relationships) for acceleration spectrum values at 66 periods and for peak ground acceleration. The period values were defined accordingly to Brady et al. (1973). Figure 3.6 and Figure 3.7 give an example of attenuation of PGA over magnitude and distance.**

Table 3.2 gives the so obtained regression coefficients as well as the standard deviations.

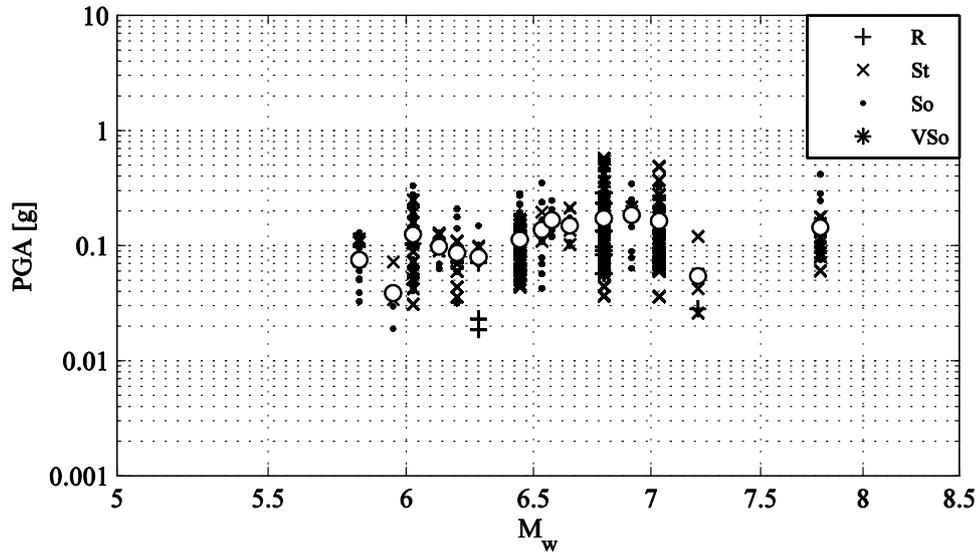


Figure 3.6. Attenuation of PGA over magnitude for ground motions recorded at distances in the interval [18.3 km, 41.43 km]. Circles indicate mean values of PGA for different magnitude bins.

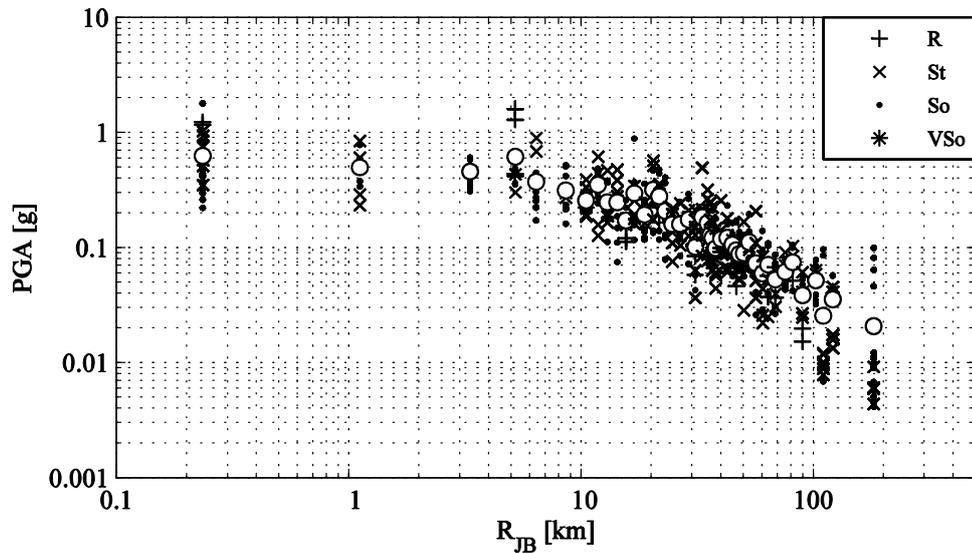


Figure 3.7. Attenuation of PGA over distance for ground motions generated by earthquakes with magnitude in the interval [6.5, 6.7]. Circles indicate mean values of PGA for different distance bins.

**Table 3.2. 5% damping spectral values: estimates of the regression parameters and of the standard deviations of Equation (3.7). Predicted spectral accelerations are in  $m/s^2$ .**

$T$ [s]	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$	$\sigma_E$	$\sigma_C$	$\sigma_A$	$\sigma_T$
PGA	-2.13	1.41	-0.11	-2.31	0.19	8.90	-0.27	0.11	0.20	0.07	0.24
0.05	-1.82	1.40	-0.12	-2.71	0.23	9.86	-0.19	0.13	0.21	0.07	0.25
0.055	-1.47	1.32	-0.11	-2.74	0.23	10.46	-0.19	0.12	0.21	0.06	0.25
0.06	-1.30	1.29	-0.11	-2.77	0.23	10.91	-0.18	0.13	0.21	0.06	0.25
0.065	-1.48	1.20	-0.09	-2.38	0.17	10.89	-0.14	0.14	0.21	0.06	0.26
0.07	-1.23	1.13	-0.09	-2.33	0.16	11.22	-0.15	0.13	0.21	0.06	0.26
0.075	-0.70	1.00	-0.08	-2.43	0.18	11.72	-0.14	0.13	0.21	0.06	0.25
0.08	-0.43	0.99	-0.08	-2.63	0.20	12.50	-0.14	0.12	0.21	0.06	0.25
0.085	-0.28	0.96	-0.08	-2.65	0.20	12.66	-0.14	0.12	0.21	0.06	0.25
0.09	0.03	0.88	-0.07	-2.65	0.20	12.85	-0.15	0.13	0.21	0.06	0.25
0.095	0.19	0.86	-0.07	-2.71	0.21	13.43	-0.16	0.13	0.21	0.06	0.25
0.1	0.23	0.88	-0.08	-2.74	0.21	13.69	-0.17	0.12	0.21	0.06	0.25
0.11	0.07	0.97	-0.08	-2.88	0.23	14.33	-0.17	0.12	0.21	0.06	0.25
0.12	0.47	0.85	-0.07	-2.82	0.22	14.72	-0.18	0.12	0.21	0.06	0.26
0.13	0.31	0.98	-0.09	-3.05	0.25	15.70	-0.20	0.13	0.21	0.07	0.26
0.14	0.55	0.96	-0.09	-3.18	0.26	16.26	-0.20	0.12	0.21	0.07	0.26
0.15	0.45	0.98	-0.09	-3.19	0.27	16.01	-0.19	0.12	0.21	0.07	0.25
0.16	0.21	1.05	-0.09	-3.17	0.27	15.95	-0.19	0.12	0.21	0.07	0.26
0.17	-0.11	1.13	-0.10	-3.10	0.26	15.83	-0.20	0.12	0.21	0.07	0.26
0.18	-0.68	1.28	-0.11	-3.05	0.26	15.76	-0.20	0.12	0.21	0.07	0.26
0.19	-0.39	1.14	-0.10	-2.87	0.23	15.25	-0.20	0.12	0.22	0.07	0.26
0.2	-0.39	1.12	-0.09	-2.81	0.23	15.04	-0.21	0.11	0.22	0.07	0.25
0.22	-0.37	1.15	-0.10	-2.87	0.24	14.39	-0.25	0.11	0.22	0.07	0.25
0.24	0.09	1.00	-0.08	-2.76	0.23	13.39	-0.28	0.11	0.21	0.08	0.25
0.26	0.03	0.93	-0.08	-2.44	0.19	11.79	-0.29	0.11	0.22	0.08	0.26
0.28	-0.05	0.91	-0.07	-2.21	0.17	10.38	-0.32	0.12	0.22	0.08	0.26
0.3	-0.37	0.98	-0.08	-2.08	0.15	9.33	-0.35	0.11	0.22	0.08	0.26
0.32	-1.25	1.25	-0.10	-2.10	0.16	8.64	-0.37	0.12	0.22	0.08	0.27
0.34	-1.70	1.37	-0.10	-2.07	0.16	8.09	-0.38	0.12	0.22	0.08	0.27
0.36	-2.27	1.57	-0.12	-2.22	0.18	8.24	-0.39	0.12	0.23	0.08	0.27
0.38	-2.84	1.73	-0.13	-2.25	0.19	7.85	-0.39	0.12	0.23	0.09	0.27
0.4	-3.04	1.74	-0.13	-2.07	0.17	7.47	-0.38	0.13	0.23	0.08	0.28
0.42	-3.42	1.82	-0.14	-1.98	0.16	7.24	-0.38	0.13	0.23	0.09	0.28
0.44	-3.68	1.90	-0.14	-1.99	0.16	7.29	-0.39	0.14	0.23	0.09	0.28
0.46	-3.80	1.94	-0.14	-1.99	0.16	7.19	-0.41	0.14	0.23	0.09	0.28
0.48	-3.89	2.01	-0.15	-2.07	0.17	7.21	-0.43	0.14	0.23	0.09	0.28
0.5	-3.97	2.06	-0.16	-2.16	0.18	7.16	-0.45	0.14	0.23	0.09	0.28
0.55	-3.91	2.01	-0.15	-2.00	0.17	6.16	-0.48	0.14	0.24	0.09	0.29

Chapter 3 – Accelerogram selection and scaling procedures for estimating the distribution of drift response

---

$T$ [s]	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$	$\sigma_E$	$\sigma_C$	$\sigma_A$	$\sigma_T$
0.6	-4.34	2.11	-0.16	-1.88	0.16	5.08	-0.51	0.16	0.24	0.09	0.30
0.65	-4.09	2.04	-0.15	-1.89	0.16	4.50	-0.54	0.16	0.24	0.09	0.30
0.7	-3.44	1.83	-0.14	-1.83	0.16	3.79	-0.56	0.17	0.24	0.09	0.31
0.75	-3.27	1.76	-0.13	-1.80	0.16	3.33	-0.57	0.17	0.24	0.09	0.31
0.8	-3.93	1.94	-0.14	-1.85	0.16	3.25	-0.56	0.16	0.24	0.09	0.31
0.85	-4.59	2.15	-0.16	-1.98	0.18	3.37	-0.55	0.16	0.24	0.10	0.31
0.9	-5.07	2.31	-0.17	-2.11	0.20	3.64	-0.56	0.16	0.24	0.10	0.31
0.95	-5.63	2.49	-0.19	-2.21	0.22	3.88	-0.56	0.16	0.24	0.10	0.31
1	-6.07	2.64	-0.20	-2.27	0.22	4.22	-0.57	0.16	0.24	0.10	0.30
1.1	-7.57	3.09	-0.23	-2.21	0.22	4.07	-0.61	0.16	0.24	0.10	0.31
1.2	-7.65	3.05	-0.23	-2.03	0.19	3.71	-0.62	0.16	0.25	0.10	0.31
1.3	-7.72	3.04	-0.22	-1.93	0.18	3.52	-0.64	0.16	0.25	0.10	0.31
1.4	-7.76	3.01	-0.22	-1.86	0.17	3.57	-0.63	0.16	0.24	0.10	0.31
1.5	-7.52	2.92	-0.21	-1.88	0.17	3.69	-0.65	0.16	0.24	0.10	0.31
1.6	-7.85	3.01	-0.22	-1.90	0.17	3.59	-0.66	0.16	0.24	0.10	0.31
1.7	-8.72	3.25	-0.23	-1.89	0.17	3.26	-0.67	0.16	0.24	0.10	0.31
1.8	-8.91	3.30	-0.24	-1.88	0.17	2.94	-0.68	0.16	0.25	0.11	0.31
1.9	-9.12	3.34	-0.24	-1.83	0.17	2.65	-0.69	0.16	0.24	0.11	0.31
2	-9.20	3.34	-0.24	-1.87	0.18	2.46	-0.69	0.16	0.25	0.11	0.31
2.1	-9.03	3.28	-0.23	-1.93	0.19	2.41	-0.69	0.16	0.25	0.11	0.31
2.2	-8.79	3.21	-0.23	-1.99	0.19	2.44	-0.69	0.16	0.25	0.11	0.32
2.3	-8.90	3.24	-0.23	-1.98	0.19	2.42	-0.70	0.17	0.25	0.11	0.32
2.4	-9.08	3.26	-0.23	-1.88	0.18	2.39	-0.70	0.17	0.25	0.11	0.32
2.5	-9.21	3.26	-0.23	-1.79	0.16	2.39	-0.70	0.17	0.25	0.11	0.32
2.6	-9.23	3.23	-0.22	-1.73	0.16	2.41	-0.71	0.17	0.25	0.11	0.32
2.7	-9.25	3.20	-0.22	-1.65	0.14	2.41	-0.71	0.17	0.25	0.11	0.32
2.8	-9.33	3.19	-0.22	-1.61	0.14	2.42	-0.72	0.17	0.25	0.11	0.32
2.9	-9.51	3.22	-0.22	-1.61	0.14	2.47	-0.71	0.17	0.25	0.11	0.32
3.0	-9.69	3.25	-0.22	-1.62	0.14	2.57	-0.70	0.16	0.25	0.11	0.32

Table 3.2 It should be noted that, the standard deviations vary with period by cover the range (in base 10 logarithm) [0.24, 0.32] with generally increasing values with increasing period. These values are entirely consistent with other empirical ground-motion models for spectral acceleration (e.g. Ambraseys et al., 2005; Akkar and Bommer, 2007b; Abrahamson and Silva, 2008a).

Using the coefficients given in Table 3.2 together with Eq. (3.7) one can evaluate response spectra for different scenarios in terms of magnitude, distance, and shear wave velocity. As an example Figure 3.8 gives the median acceleration response spectra at 5% damping predicted for sites with  $V_{S30} = 1000$  m/s at distances of 10, 50, 100 and 200 km for a  $M = 6.5$  event. Figure 3.9 and Figure 3.10 give the predicted attenuation over magnitude and distance of spectral acceleration at 0.5 s for sites with  $V_{S30} = 1000$  m/s.

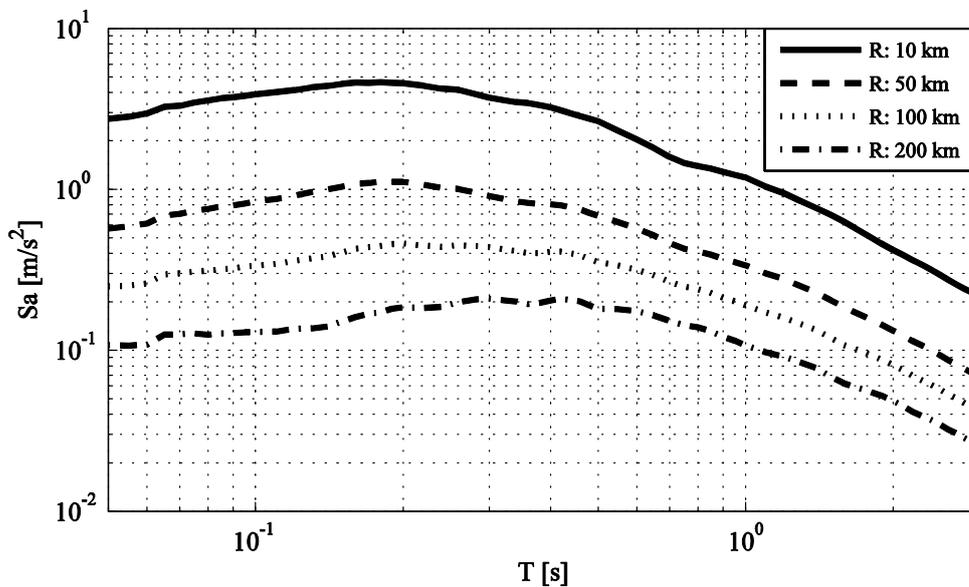


Figure 3.8. Acceleration response spectra (5% damping ratio) predicted by the regressions developed in the present work ( $M_w=6.5$ ;  $V_{S30}=1000$  m/s).

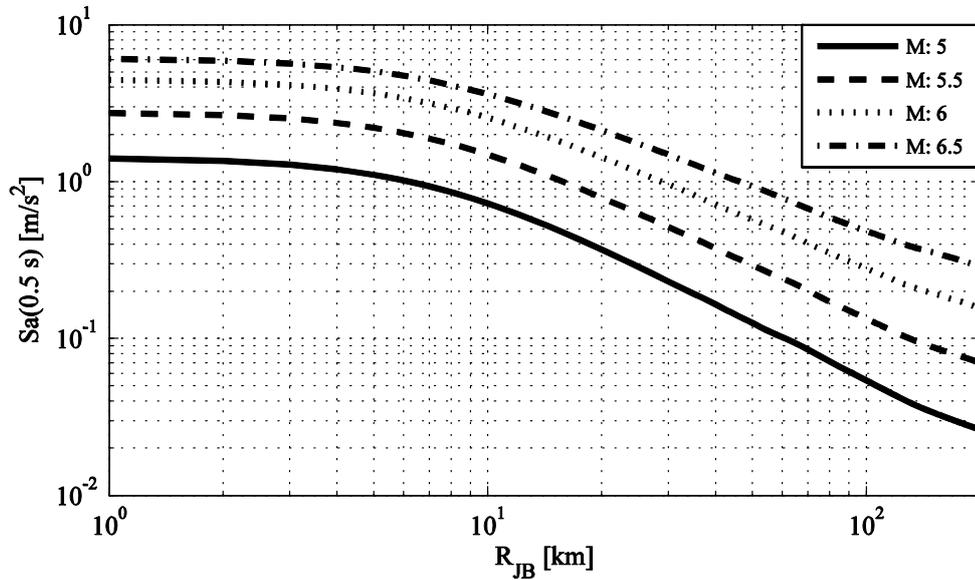


Figure 3.9. Attenuation of spectral acceleration (5% damping ratio) over distance predicted by the regressions developed in the present work ( $V_{s30} = 1000$  m/s).

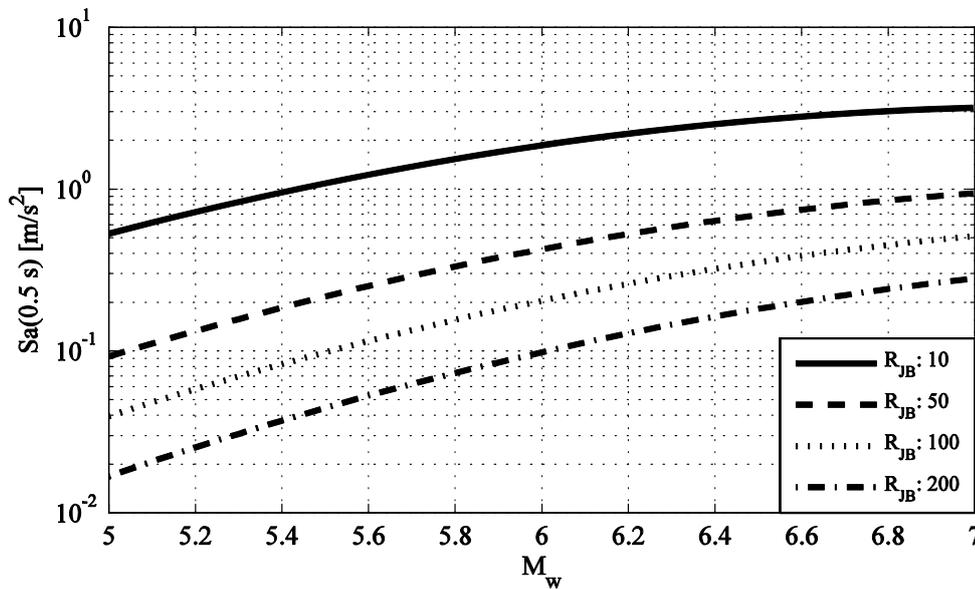


Figure 3.10. Attenuation of spectral acceleration (5% damping ratio) over magnitude predicted by the regressions developed in the present work.

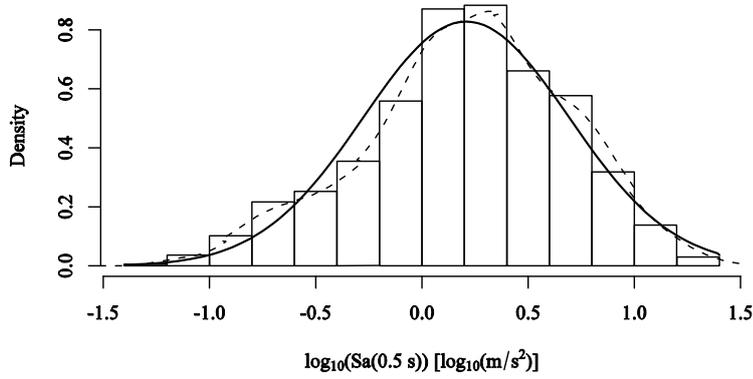


Figure 3.11. Histogram of  $\log_{10}(\text{Sa}(0.5 \text{ s}))$ , the solid black line is a normal distribution and the dashed line is obtained through a kernel smoothing procedure.

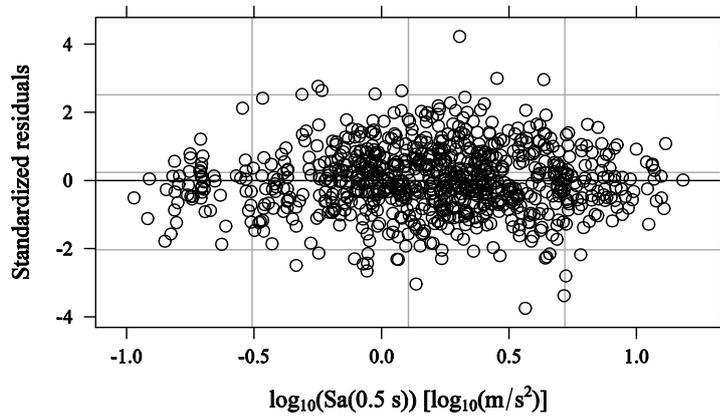


Figure 3.12. Dependency of standardized residuals on fitted values of  $\text{Sa}(0.5 \text{ s})$ .

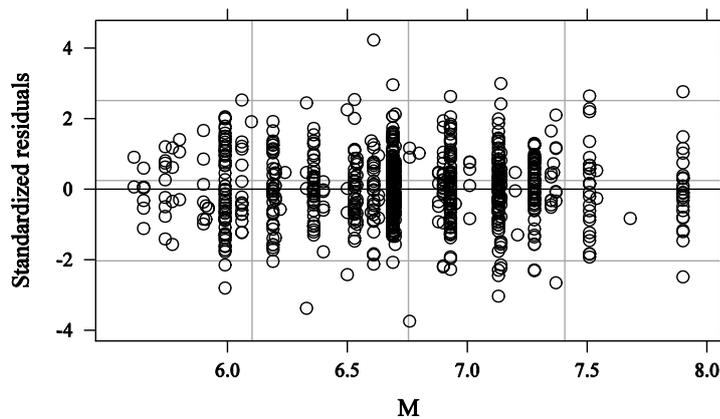


Figure 3.13. Dependency of standardized residuals on magnitude for  $\text{Sa}(0.5 \text{ s})$ .

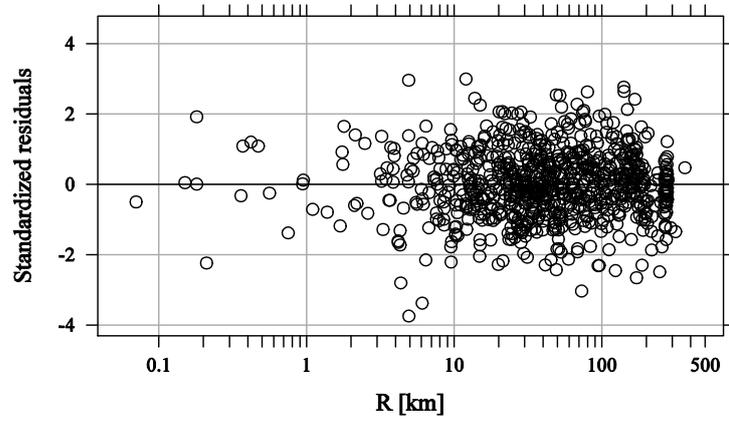


Figure 3.14. Dependency of standardized residuals on distance for  $S_a(0.5\text{ s})$ .

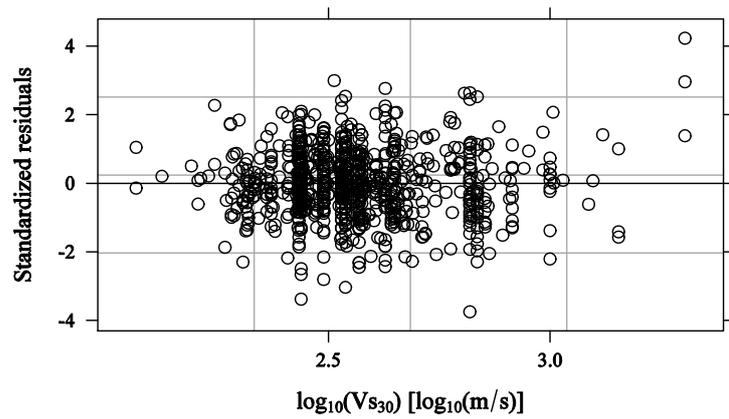


Figure 3.15. Dependency of standardized residuals on base 10 logarithm of shear wave velocity over the upper 30 meters for  $S_a(0.5\text{ s})$ .

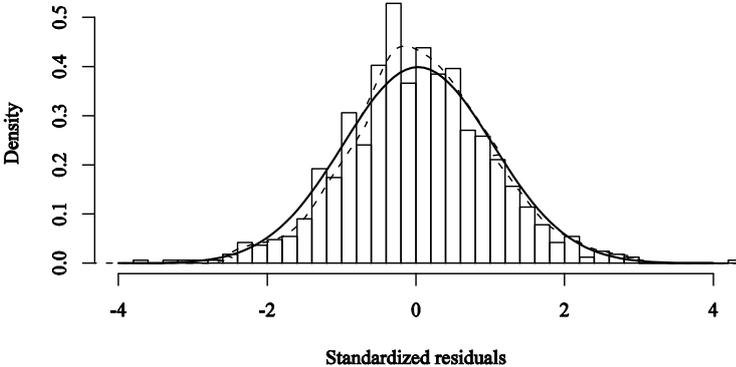


Figure 3.16. Histogram of the standardized residuals for Sa(0.5 s). The solid black line is a standard normal distribution and the dashed line is obtained through a kernel smoothing procedure.

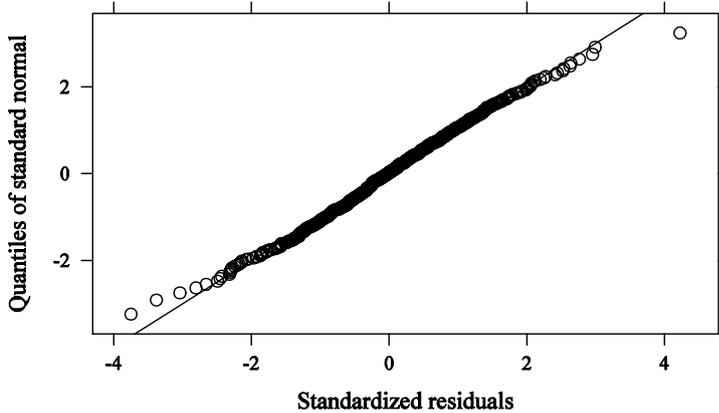


Figure 3.17. Normal quantile-quantile plot for the standardized residuals for Sa(0.5 s).

When performing the regressions the goodness of fit was checked for all the parameters considered, and in particular the hypothesis of normal distribution of  $\log_{10}(Y)$  was verified (Y being the generic variable considered). As an example some diagnostic plots for the regressions performed for Sa(0.5 s) are reported in Figure 3.11-Figure 3.17. Figure 3.11 shows a histogram which allows to check the normality of the logarithm of the data. The solid line represents a reference normal distribution and the dashed line was obtained using a kernel smoothing procedure on the data.

A plot of the standardized residual versus the fitted values is depicted in Figure 3.12. No dependency of the standardized residual on the fitted values can be observed, the average residual is 0 and its distribution can be assumed constant with good approximation. Similar plots investigating the presence of trends of the residuals versus distance, magnitude and shear wave velocity are given in Figure 3.13-Figure 3.15. These plots confirm the goodness of the regression. Finally Figure 3.16 and Figure 3.17 allow to check the normality of the residuals, in fact Figure 3.16 reports an histogram of the

**Table 3.3. Estimates of regression coefficients and standard deviations (see Equation (3.7)) for the interstorey drift ratios and for the roof displacement. Predicted interstorey drifts are not dimensional (drift ratios) and roof displacement is in m.**

Drif	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$	$\sigma_E$	$\sigma_C$	$\sigma_A$	$\sigma_T$
I Storey	-6.64	1.73	-0.12	-1.34	0.09	2.62	-0.57	0.17	0.09	0.24	0.31
II Storey	-7.12	1.98	-0.14	-1.53	0.12	3.12	-0.58	0.17	0.09	0.24	0.31
III Storey	-7.46	2.11	-0.15	-1.64	0.13	3.34	-0.57	0.17	0.09	0.24	0.31
IV Storey	-6.93	1.95	-0.14	-1.72	0.15	3.77	-0.55	0.16	0.08	0.23	0.29
V Storey	-7.37	2.21	-0.16	-1.82	0.15	2.89	-0.62	0.15	0.10	0.26	0.32
VI Storey	-7.07	2.02	-0.14	-1.80	0.16	3.38	-0.5	0.13	0.08	0.23	0.28
Roof	-6.36	2.16	-0.15	-1.65	0.14	3.05	-0.60	0.17	0.09	0.25	0.32

**Table 3.4. p-values for the statistical tests for normality conducted upon the considered drift measures. A-D: Anderson-Darling; K-S: Kolmogorov-Smirnov; L: Lilliefors; and S-W: Shapiro-Wilks tests. Bold test indicates values less than 0.05.**

Drift	Inter-event residuals				Intra-event residuals			
	A-D	K-S	L	S-W	A-D	K-S	L	S-W
I St.	0.899	0.974	0.787	0.703	0.298	0.222	0.107	0.192
II St.	0.704	0.928	0.741	0.570	0.543	0.400	0.479	0.114
III St.	0.592	0.903	0.890	0.573	0.329	0.374	0.287	0.150
IV St.	0.650	0.851	0.805	0.624	0.080	0.281	<b>0.027</b>	0.131
V St.	0.539	0.778	0.754	0.456	<b>0.001</b>	<b>0.035</b>	<b>0.001</b>	<b>0.000</b>
VI St.	0.293	0.696	0.786	0.237	<b>0.000</b>	0.093	<b>0.031</b>	<b>0.000</b>
Roof	0.675	0.922	0.926	0.568	0.088	0.312	0.085	0.132

residuals, it can be observed that their distribution (dotted line) is pretty close to a normal distribution (solid line) and that their mean value is 0. Figure 3.17 gives a quantile-quantile plot which allows to check the same assumptions, apart from little deviations in very high and low quantiles the normality hypothesis can be accepted. It should be noted that other attenuation relationship for spectral accelerations available in the literature could have been used to estimate spectral acceleration distribution. In the present work it was chosen to develop specific relationships in order to keep the procedure as much internally consistent as possible in order to reduce uncontrollable source of error.

### 3.4.2 Interstorey drifts

As previously stated regression analysis was used to develop predictive equations for the demand measures under investigation. The predictive equations were then used to define the reference probability distributions of the demand measures. Goodness of fit tests showed that the same attenuation formula used for spectral acceleration could be used for interstorey and roof drifts as well. Examples of attenuation of the second floor

interstorey drift,  $\delta_2$ , over distance and magnitude are depicted in Figure 3.18 and Figure 3.19, which if compared with Figure 3.6 and Figure 3.7 respectively, show many similarities.

Diagnostic plots for second floor interstorey drift are reported in Figures 1.23-1.29. Their meaning is similar to those concerning the regression on  $S_a(0.5\text{ s})$ . It is worthily noted that Figure 3.22 supports the hypothesis of normal distribution of the base ten logarithm of the drift. Since attenuation relationships for drift ratios are not common in the literature (being Hancock et al. (2008) and Watson-Lamprey (2006) the only other examples), the normality of the logarithm of the drift ratios was checked using four different statistical tests. The so obtained results are given in Table 3.4. It can be observed that that the intra-event residuals for the interstorey drifts of the V storey are not lognormally distributed (at 5% significance); conflicting test results are obtained also for the VI storey with some tests indicating not to reject the hypothesis that the residuals are lognormally distributed. However, for all other storeys and for the roof drift all of the tests allow to retain the assumption that the logarithmic drift values are lognormally distributed.

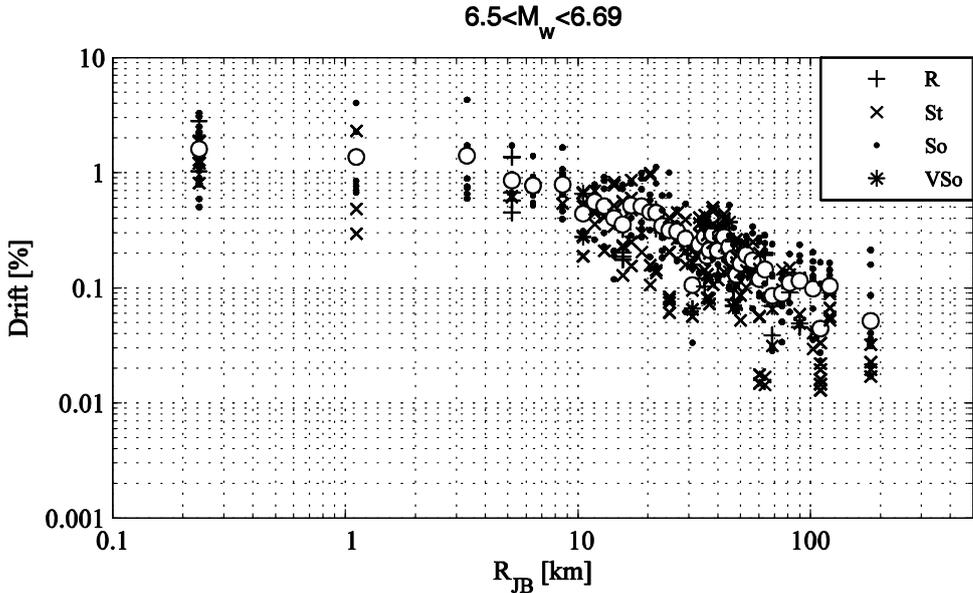


Figure 3.18. Second-floor interstorey drift: attenuation over distance. Data are grouped in  $R_{JB}$  bins which mean values are depicted as circles. Interstorey drift shows an attenuation with a shape similar to that of PGA and spectral acceleration.

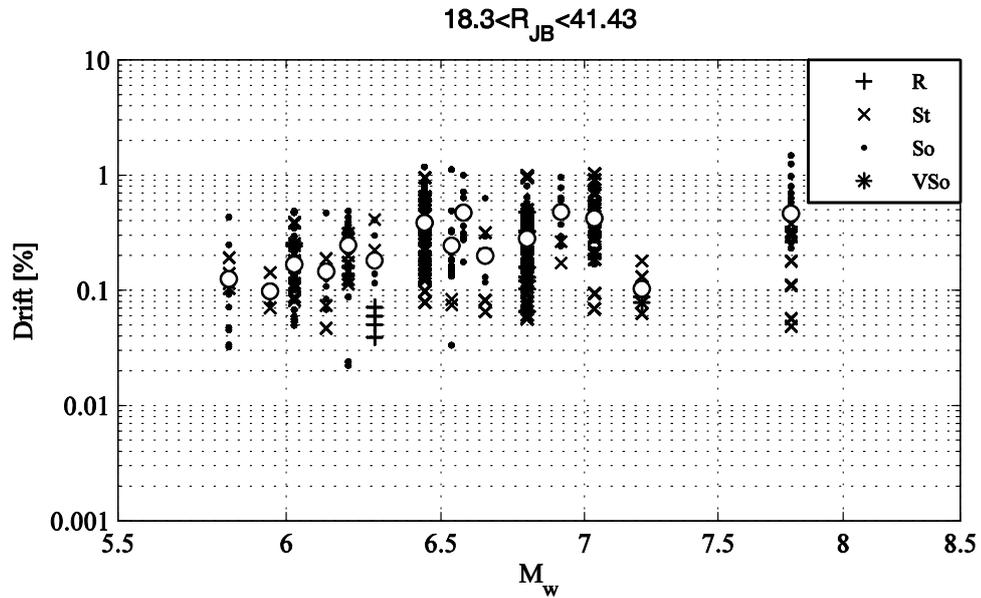


Figure 3.19 Second-floor interstorey drift: attenuation over magnitude. Data are grouped in  $M$  bins which mean values are depicted as circles. Interstorey drift shows an attenuation with a shape similar to that of PGA and spectral acceleration.

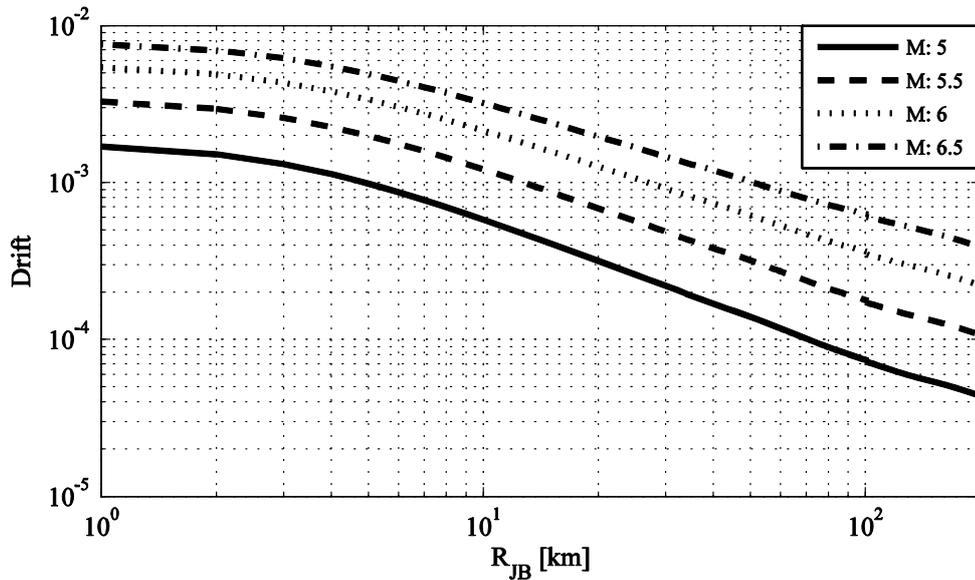


Figure 3.20. Predicted attenuation over distance of the II storey drift. A  $V_{S30}$  value equal to 1000 m/s is considered.

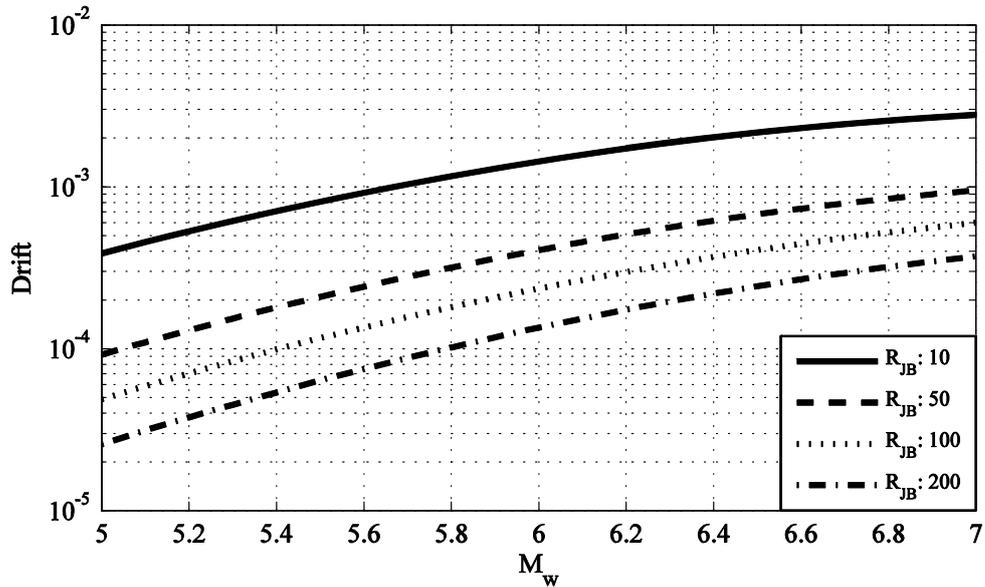


Figure 3.21. Predicted attenuation over magnitude of the II storey drift. A  $V_{s30}$  value equal to 1000 m/s is considered.

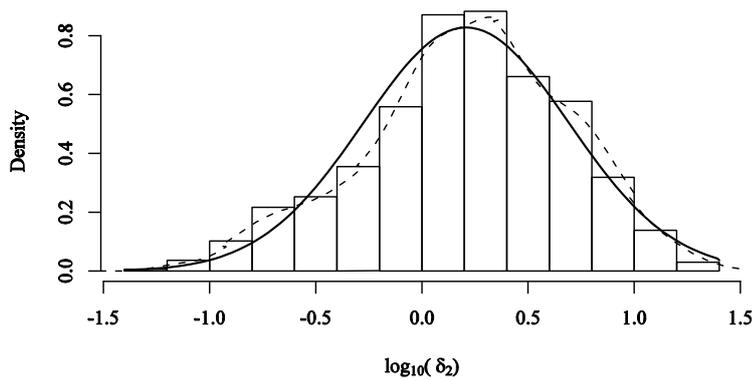


Figure 3.22. Histogram of  $\log_{10}(\text{drift}_2)$ , the solid black line is a normal distribution with and the dashed line is obtained through a kernel smoothing procedure.

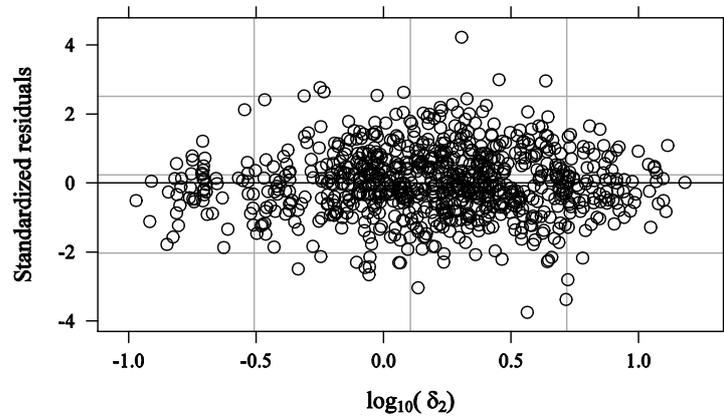


Figure 3.23. Dependency of standardized residuals on fitted values.

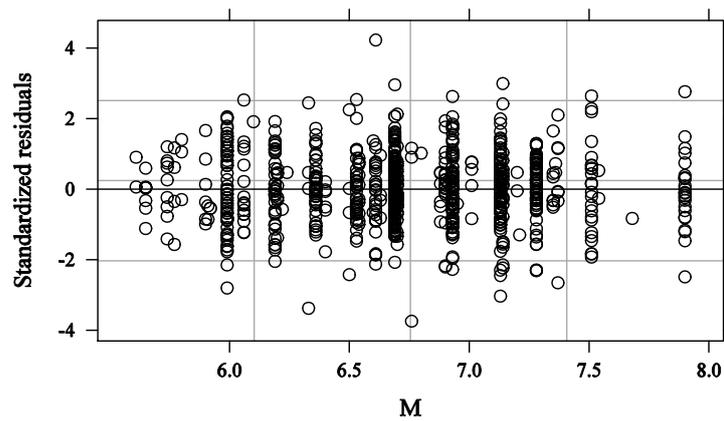


Figure 3.24. Dependency of standardized residuals upon magnitude.

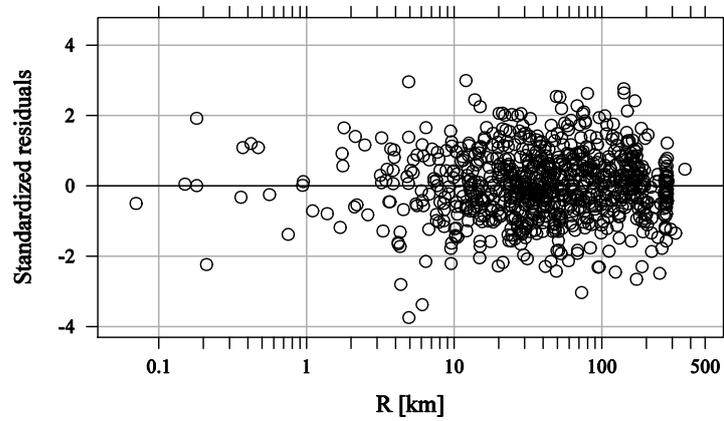


Figure 3.25. Dependency of standardized residuals on distance.

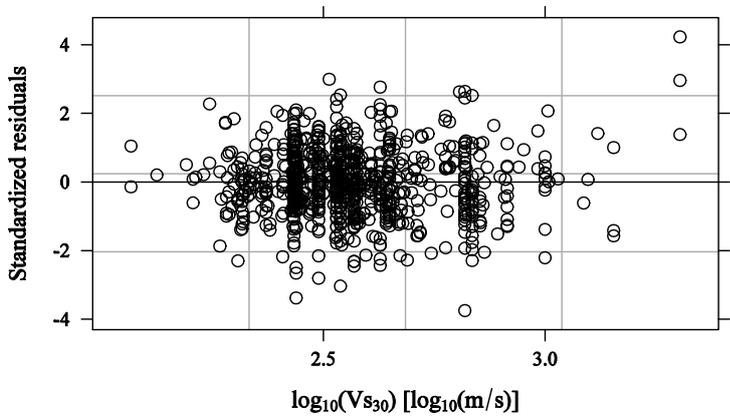


Figure 3.26. Dependency of standardized residuals on log<sub>10</sub>(vs30)

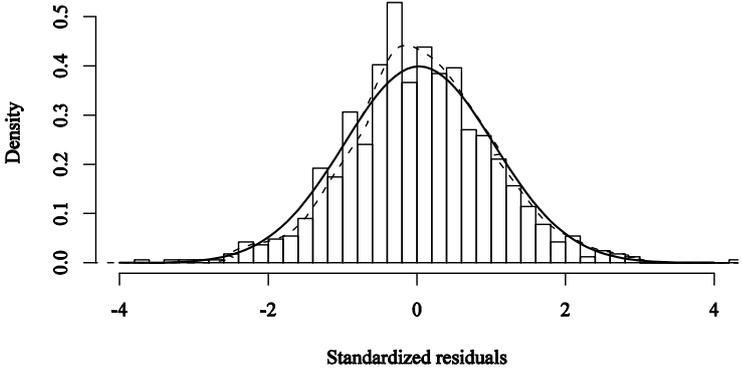


Figure 3.27. Histogram of the standardized residuals for second storey drift. The solid black line is a standard normal distribution and the dashed line is obtained through a kernel smoothing procedure.

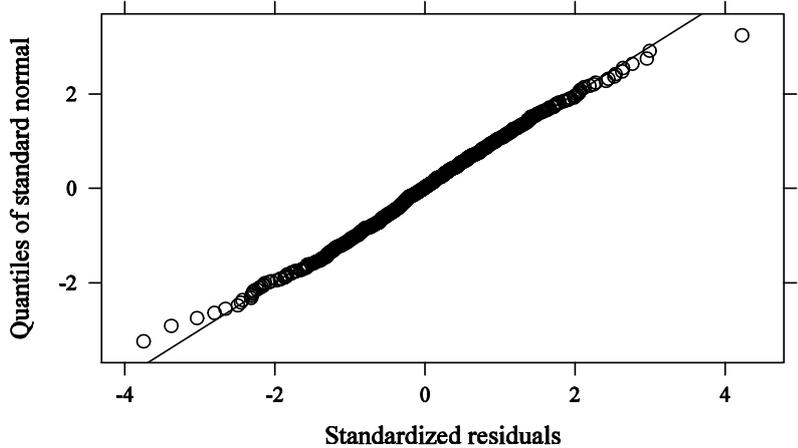


Figure 3.28. Quantile-quantile plot for the standardized residuals for second storey drift regression.

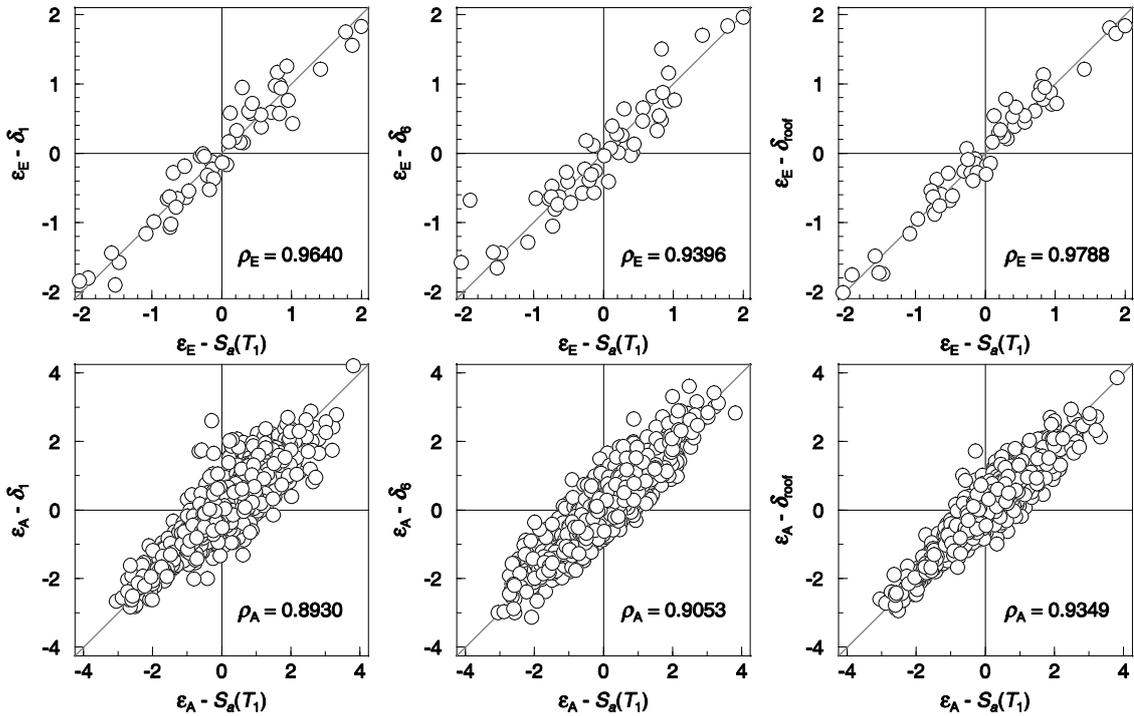


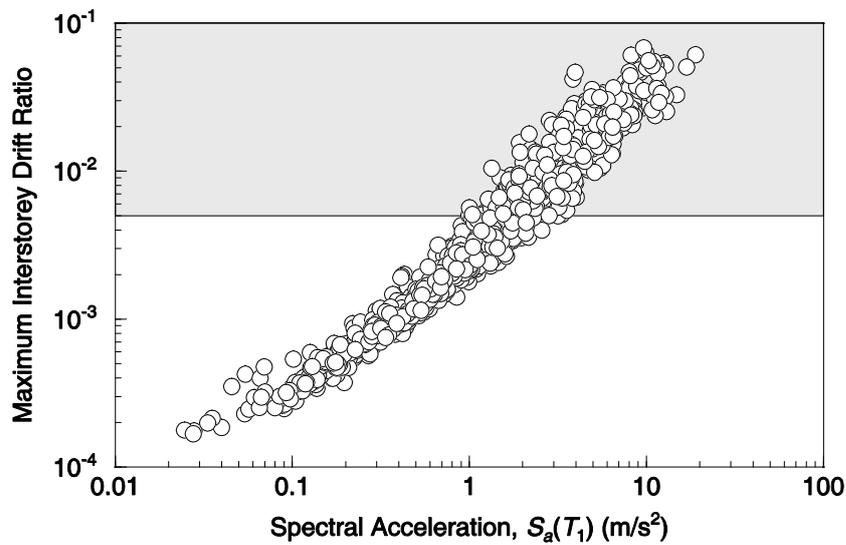
Figure 3.29. Correlation of regression residuals for first storey drift (left column), sixth storey drift (centre column) and roof drift (right column) with spectral acceleration at natural period of the structure considered

### 3.4.3 Correlations between $Sa(T_1)$ and $\delta$

Following the development of the empirical models, the residuals were examined in order to identify the strength of any correlations that exists. For the purpose of determining the correlations it is important to make the distinction between the inter-event and intra-event components of the total residuals. The correlations between all of the considered drift measures (the six inter-storey drifts and the roof drift) and the spectral acceleration at the initial fundamental period of the structure are very strong. The strength of these correlations may be appreciated by inspection of Table 3.5 where all of the calculated correlations are presented, as well as in Figure 3.29 where examples of

Table 3.5. Correlation of regression residuals with spectral acceleration at natural period of the structure.

	Correlation with $Sa(T_1)$						
	I storey drift, $\delta_1$	II storey drift, $\delta_2$	III storey drift, $\delta_3$	IV storey drift, $\delta_4$	V storey drift, $\delta_5$	VI storey drift, $\delta_6$	Roof drift, $\delta_{roof}$
Inter-event	0.9640	0.9709	0.9742	0.9814	0.9615	0.9396	0.9788
Intra-event	0.8930	0.9126	0.9240	0.9357	0.9250	0.9053	0.9349
Total	0.9177	0.9327	0.9416	0.9515	0.9387	0.9224	0.9502



**Figure 3.30. Relationship between maximum inter-storey drift and the spectral acceleration at the fundamental period of the structure,  $S_a(T_1)$ , found from the 1666 time-domain analyses conducted in this study using unscaled natural records. The gray shaded area indicates drift demands causing non linear behaviour of the structure-**

the normalized residuals are shown. The weakest correlation determined from the analyses of the unscaled records is between the spectral acceleration and the first storey drift, while the strongest correlations are found between the spectral acceleration and the fourth storey drift.

The strength of these correlations strongly suggests that if one were to scale an accelerogram to a level of, say, one epsilon, then one should expect that the drift values that are obtained following a time-history analysis conducted with this scaled record will also be very close to a one epsilon (i.e., 84th-percentile) level in terms of drift.

It could be argued that the strength of the correlation depends on the extent to which the structure behaves nonlinearly during the analyses. Figure 3.30 shows the relationships between maximum interstorey drift ratios and  $S_a(T_1)$  obtained by the analyses performed on the 1666 reference ground motions. Comparing Figure 3.30 with Figure 3.5 one can conclude that many ground motions bring the structure in the non linear range.

### **3.4.4 Record Selection**

Once some reference acceleration response spectra have been defined by applying the theory of point approximations of continuous variables on the response spectra distributions predicted by attenuation relationships, accelerograms are to be selected according to those reference response spectra. This section describes the criteria adopted to select accelerograms consistent with the reference spectra.

#### *3.4.4.1 Reference scenario*

Prior to describing the approach to record selection it necessary to define the seismological scenario that was considered for this case study. The scenario corresponds to an event with a moment magnitude of  $M_w$  7.0 located at a Joyner-Boore distance of  $R_{JB} = 10$  km from a site with an average shear-wave velocity over the uppermost 30 m of  $V_{S30} = 300$  m/s. This scenario is purely hypothetical and was not selected with a view to representing any particular situation. However, the scenario was also not defined arbitrarily. The primary reason for working with this scenario is that it represents a severe loading case that should ensure that the structure behaves in a nonlinear manner. This combination of magnitude, distance and site class is also one for which a reasonable number of records exist, which thus provides one with more options when implementing the adopted selection procedure outlined in what follows. The final key reason is associated with the confidence that it may be placed in the empirical models that have been derived for the drift responses. Given that there are a good number of records that have seismological characteristics similar to those of the specified scenario and that the scenario is well within the range of parameter values used to derive the empirical models, it is reasonable to be confident that the model will perform robustly and will act as a good surrogate for the unknown true distribution of drift. Hence, one can be confident that any biases that one observes in the final results are predominantly due to actual biases rather than being due to a poor definition of the basis for comparison.

#### *3.4.4.2 Spectral matching procedure*

Numerous researchers have demonstrated, either explicitly (e.g., Hancock et al., 2008; Baker and Cornell, 2006b), or implicitly (Luco and Bazzurro, 2007), that an effective way to prevent biased estimates of structural response when using scaled records is to ensure that careful attention is paid to the spectral shapes of the accelerograms. That is, it is important to ensure that when accelerograms are scaled their spectra match a target spectrum at multiple periods so that one does not inadvertently select records that contain significant peaks or troughs at particular periods. As the focus in this study is upon estimating the distribution of drift response, selection on the basis of spectral shape should take priority over other commonly imposed constraints such as restrictions on the magnitude of the event from which the original accelerogram came. This prioritization is justified as Hancock (2006) and Hancock and Bommer (2006) have demonstrated that peak responses like drift are not particularly sensitive to the duration of shaking (which is correlated to magnitude). That said, when selecting records for the

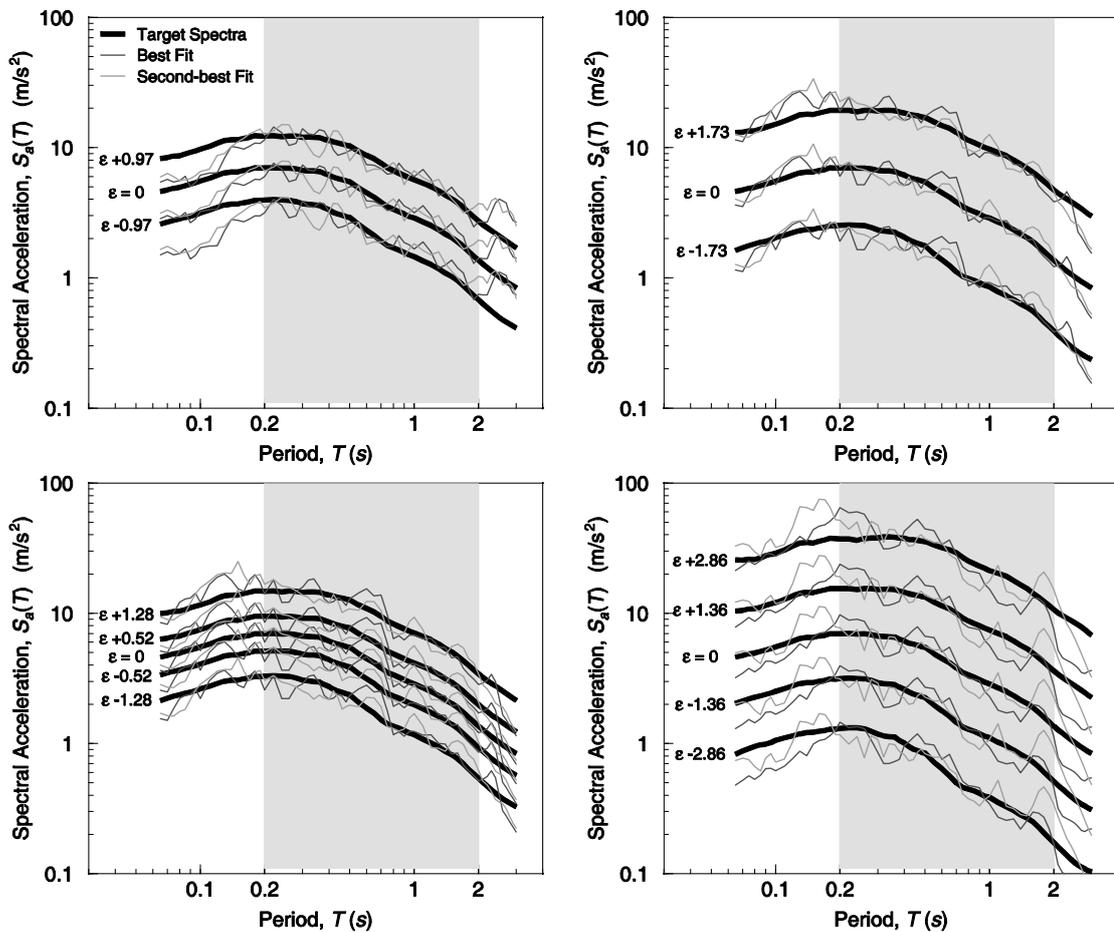


Figure 3.31. Examples of the selection procedure: reference spectra and best matching accelerogram spectra.

purpose of estimating the structural response due to a particular seismological scenario it is, in principle, preferable to select records that are at least broadly consistent with this scenario and for this reason an initial screening was made to limit the selection of accelerograms to those coming from events having magnitudes within  $\pm 0.2$  units of the specified scenario. Following this initial screening, records were selected on the basis of their ability to fit, when scaled, the target spectra over a period band ranging from roughly half the initial fundamental period of the structure to roughly double the initial fundamental period of the structure, i.e., over the range  $[0.2, 2.0]$  seconds. The boundaries of the period range are defined following consideration in the framework of direct displacement-based design (Priestley et al., 2007).

The goodness-of-fit is quantified by the root-mean-square difference ( $D_{rms}$ ) between the logarithmic ordinates of the candidate spectrum,  $S_{aR}$ , and the logarithmic ordinates of the target spectrum,  $S_{aT}$ , in the interval  $[T_a, T_b]$ :

$$D_{rms}(\alpha) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\log_{10}(\alpha Sa_R(T_i)) - \log_{10}(Sa_T(T_i)))^2} \quad (3.2)$$

where  $\alpha$  is a scaling factor and  $n$  is the number of periods considered. For each candidate record the scaling factors that minimize the  $D_{rms}$  are calculated for each of the different target spectral levels. The records that are finally selected are those with the smallest  $D_{rms}$  (Beyer and Bommer, 2007).

#### 3.4.4.3 Selected suites of ground motions

The selection procedure outlined above is approached in two ways. In the first approach, each of the three or five target spectral levels (prescribed either via the equal-probability, EP, approach or the Gauss-Hermite, GH, approach) are considered independently and the records whose scaled spectra best fit each of these levels are selected. In this first approach the selection is performed without replacement so that once a record has been selected it is no longer a candidate for selection at any of the other target levels. In the second approach, all target levels are considered simultaneously and the selected records are those whose spectra are able to be scaled (by three or five different factors as the case may be) in order to provide a good match to all target levels. Initially only the first approach was considered but it was observed that the standard deviation of the empirical model for spectral acceleration is fairly constant across periods with the implication (when specifying target ordinates using the same epsilon over all periods) that the shapes of the target spectra are all very similar. Hence, if a record has a spectrum that matches one level it will also have a reasonable fit to the other levels and one may therefore work with a reduced set of accelerograms. Note that this does not imply a reduction in the number of structural analyses, just a smaller number of accelerograms that must then be scaled multiple times and that the constraint of non-replacement following selection becomes redundant. Figure 3.31 demonstrates the results of this second selection procedure for the four different methods of specifying target spectral levels considered in this study.

Note that in each case the target spectrum corresponding to  $\epsilon = 0$  is the same yet different accelerograms are chosen due to their ability to match all of the target levels simultaneously.

### 3.4.5 Results

**Table 3.6. Maximum drift ratio demands (see Figure 3.5 for capacity curves) for the record selection based upon single records scaled to multiple reference spectra. GH: Gauss-Hermite, EQ: Equal Probability.**

Method	Storey					
	I	II	III	IV	V	VI
GH 3 levels	0.038	0.037	0.037	0.029	0.064	0.048
GH 5 levels	0.073	0.083	0.06	0.032	0.25	0.064
EP 3 levels	0.021	0.021	0.022	0.015	0.05	0.03
EP 5 levels	0.033	0.032	0.035	0.022	0.058	0.038

**Table 3.7. Maximum drift ratio demands (see Figure 3.5 for capacity curves) for the record selection based upon a different record scaled to each reference spectrum. GH: Gauss-Hermite, EQ: Equal Probability.**

Method	Storey					
	I	II	III	IV	V	VI
GH 3 levels	0.038	0.037	0.037	0.029	0.061	0.048
GH 5 levels	0.056	0.067	0.055	0.032	0.081	0.064
EP 3 levels	0.01	0.01	0.012	0.011	0.032	0.022
EP 5 levels	0.033	0.03	0.026	0.018	0.051	0.025

For both the EP and GH approaches the procedure based upon 3-point and 5-point approximations was evaluated for all inter-storey drifts and the roof drift. Previously, in Table 3.5 it was observed that the weakest correlation determined from the analyses of the unscaled records was between the spectral acceleration and the first-storey drift (although these parameters were still strongly correlated) while the strongest correlations were found to exist between the spectral acceleration and the fourth-storey drift followed very closely by the roof drift.

The underlying assumptions of the methodology proposed herein would suggest that the worst performance of the method should be observed for the first-storey drift while the best performance should be observed for the fourth-storey and roof drifts. For this reason the results of the analyses conducted on the scaled records for the first-storey drift in Figure 3.32 and the roof drift in Figure 3.33 are presented. For all of the other measures of drift, for which the results are not presented herein, the performance of the approaches exhibited similar trends and the results shown in Figure 3.32 and Figure 3.33 may be regarded as being indicative of the general performance of the approaches for this structure.

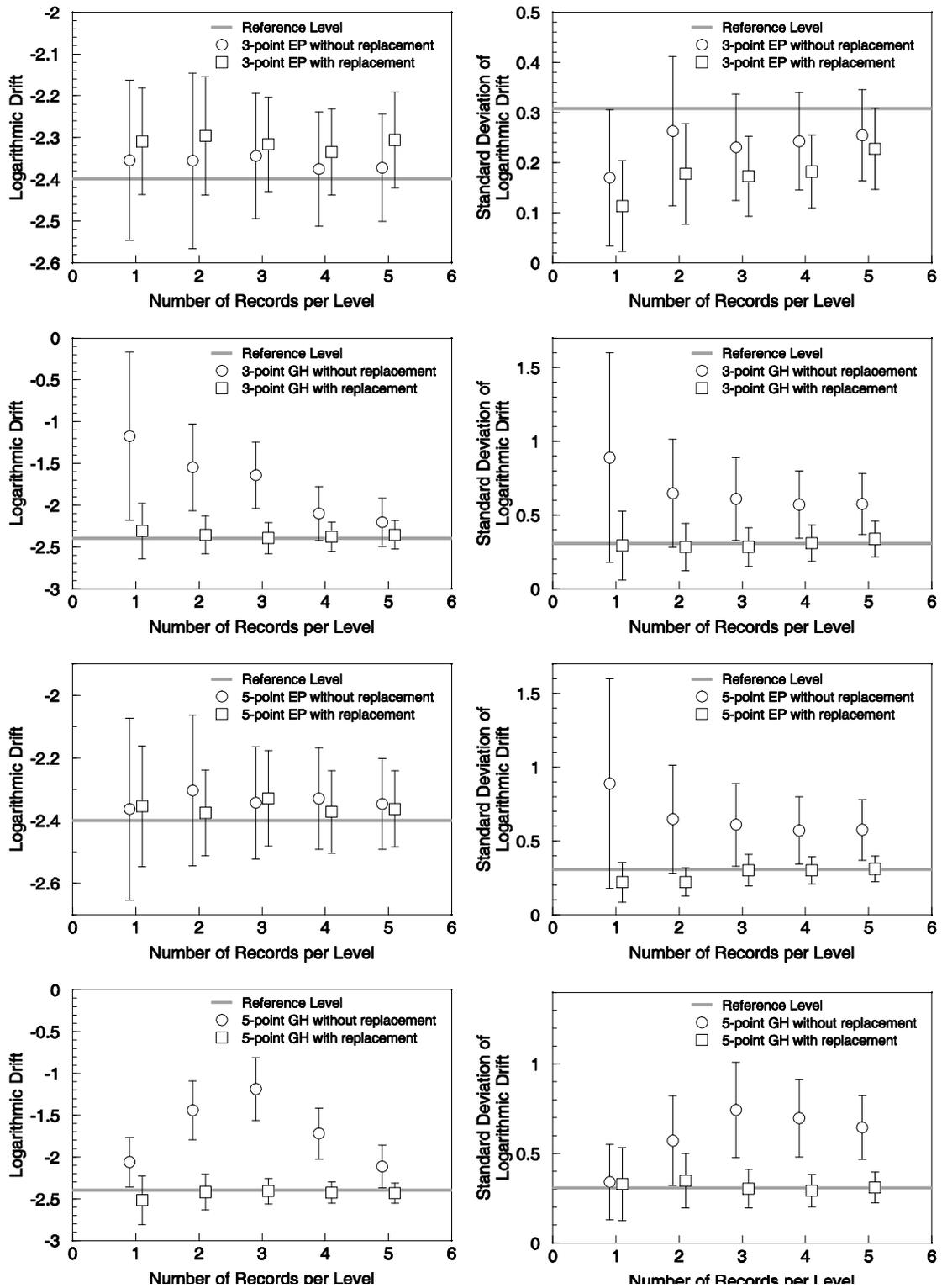


Figure 3.32. Predictions, obtained by the different procedures considered, of mean value (left column) and standard deviation (right column) of the logarithm of first storey drift.

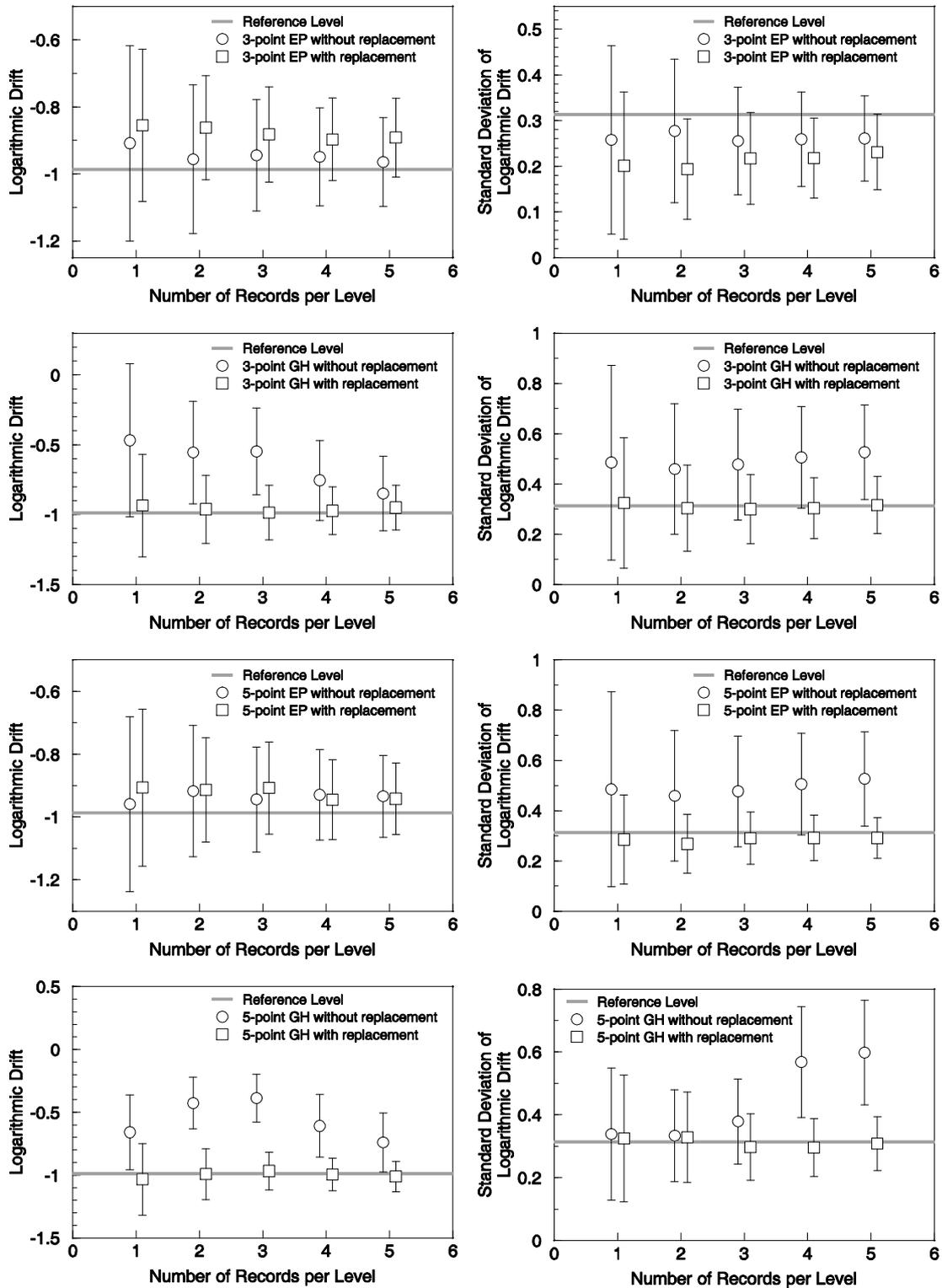


Figure 3.33. Predictions, obtained by the different procedures considered, of mean value (left column) and standard deviation (right column) of the logarithm of roof drift.

In both Figure 3.32 and Figure 3.33, the estimates of the first two moments of drift ratios are shown in addition to the 95% confidence interval about this estimate. It is very clear that the performance of the Gauss-Hermite approach is significantly superior to that of the Equal-Probability approach, although it is important to note that different scales have been used on the ordinates of these figures. In the vast majority of cases it may be appreciated that the 95% confidence interval includes mean values and standard deviations of drift ratio distributions estimated through attenuation relationships, and assumed to represent the reference drift distributions. This implies that in most cases it is not possible to state that the estimates based upon the scaled analyses are significantly different (in a statistical sense) from the reference levels shown in these figures by the heavy gray lines.

Finally Table 3.6 and Table 3.7 give the maximum interstorey drifts observed during the analyses performed with the ground motions selected according to the different procedures investigated. Comparing those tables with Figure 3.5 one can observe that the structure has a nonlinear behaviour during the analyses.

### **3.5 Discussion and conclusions**

The results presented in Figure 3.32 and Figure 3.33 seem to be very promising. When the target spectral levels are defined according to the roots of Hermite polynomials and individual records are scaled to match all of these different targets, it may be appreciated that both the 3-point and 5-point approximations work extremely well. This statement holds for both the estimates of the means and of the standard deviations of the logarithmic drifts. Good approximations to the full distribution of drift may therefore be obtained by using a single record with a spectral shape that is well matched to the specified scenario. This approximation is generally improved, but only very slightly, when a second record is used. It also appears that the use of a 5-point approximation does not provide significant advantages over a 3-point approximation which means that as few as six (two records, each scaled to three target levels) carefully designed time-history analyses may be all that is required in order to obtain a reasonable estimate of the full distribution of drift response. Of course, while the 95% confidence intervals for these estimates in Figure 3.32 and Figure 3.33 were presented no statement was made regarding how many records would be required to achieve a particular precision in these estimates of the first and second moments. However, it can be observed from these figures that for the preferred GH approach, and with replacement allowed during the record selection, there are no significant trends, which means that an estimate of the

number of records that are required to achieve a certain level of precision can be obtained from consideration of the formulae for determining the standard errors of the estimates of the mean and standard deviation.

The approach using a single record scaled to match all target levels of the response spectrum performs very well, which is perhaps not unexpected given that the shape of the target spectra are always very similar, as discussed previously. When different records are selected to match each of the target levels separately, and the rule not allowing replacement of the selected records is invoked, the method is somewhat less efficient because of the still rather limited coverage of the global strong-motion databank. Given that the total number of available records in any magnitude distance-site class bin is generally small, if a separate record is required for each target level of the spectrum, it is almost inevitable that for the third or fifth record, as the case may be, the match to the target spectral shape will not be as good as for the first record selected. Therefore, for as long as the number of candidate records remains rather limited it is advantageous to scale individual records to multiple target levels.

In this study the records were selected by first finding the record that provided the best match to the median target spectrum before progressively moving to target spectra more distant from the median.

This approach makes some intuitive sense when the objective is to estimate the median drift response but it may be that for the purposes of modelling the full distribution greater priority could be placed upon ensuring better fits to the target levels away from the median.

This study has only been concerned with the estimation of the distribution of drift response and the proposed methodology in this case has proven to be effective for the single structure and earthquake scenario considered herein. Whether or not this approach would yield similar quality results when applied to alternative damage measures such as fatigue damage or absorbed hysteretic energy remains to be seen but this will most likely be related to the strength of the correlation between spectral acceleration and these parameters. For these damage measures it is likely that more importance would have to be placed upon ensuring that the duration and frequency contents of the selected records were consistent with the specified seismological scenario (Hancock, 2006; Hancock and Bommer, 2006; Hancock and Bommer, 2007).

The recent study of Hancock et al. (2008) suggests that the approach presented in this study might work even more effectively if in addition to linearly scaling the accelerograms so that their spectra matched the target spectra over a broad period range, these records were then adjusted with wavelets so that this match was further improved.

Furthermore, it would seem that the chances that the proposed approach would work when applied to estimate the distributions of fatigue damage or absorbed hysteretic energy would be greatly enhanced if the wavelet adjustments were made to not only multiple target spectra but also to spectra constructed for different damping ratios. Hancock et al. (2008) have demonstrated that this approach further ensures that the duration of the selected records are consistent with the prescribed seismological scenario.

This work represents a first attempt at investigating the problem of how many records are required in order to obtain a robust estimate of the full distribution of a response parameter. The study has been conducted for a single six-storey reinforced concrete frame and for a single seismological scenario and the performance of the proposed approaches should be considered in this light. Clearly, further work is required for different earthquake scenarios and different structures in order to qualify and generalize the findings of this study as well as to refine the overall procedure.

However, the results that have been presented herein are promising and suggest that it is conceivable for engineers to obtain stable estimates of the distribution of structural response parameters from a relatively small number of time-history analyses.



# 4 Generation of non-stationary artificial accelerograms

## 4.1 Introduction

In the last ten years an increasing effort to install strong-motion recording networks has been done worldwide. In spite of this initiative, the availability of natural records is sometime limited due to the absence of nearby recording stations or because the site is in a low or moderate seismicity region. For this reason, the generation of synthetic accelerograms is important to provide engineers with reliable input excitation for sites with no strong motion data. Thus, simulated accelerograms must have realistic energy duration and frequency content, representing the physical conditions of natural recorded acceleration time histories. Moreover, providing information about the aleatory variability of ground motion predictions is necessary for probabilistic seismic hazard assessment (Thrainsson and Kiremidjian, 2002; Boore, 2003). In addition, the increase of computational power and the development of numerical methods enable the use of acceleration time histories in the study of linear and nonlinear dynamic behaviour of structures.

Many engineering methods have been proposed for simulating ground motions. A first attempt to simulate accelerograms was made by filtering and windowing a Gaussian noise or by using Autoregressive Moving Average models (Saragoni and Hart, 1974; Nau et al., 1982). The time histories are in general disconnected from any specific geophysical parameter. Other approaches provide accelerograms with response spectra matching a target response spectrum (Gasparini and Vanmarcke, 1976; Kaul, 1978). The core of the problem is that the matching procedure generates too many cycles of strong motion. The artificial accelerograms have consequently an unreasonably high energy content and Naiem and Lew (1995) demonstrated the inefficiency of using such artificial record for non-linear analysis.

A second category of methods for simulating acceleration time histories relies on a more physical approach, for which the ground motion is modelled by convolving the source, path and site effect (Aki and Richards, 1980). Some important efforts have been

devoted to the modelling of the source process. For example, point source summation (Hartzell, 1978; Irikura, 1983) and stochastic subevent summation (Schneider et al., 1993; Zeng et al., 1994; Beresnev and Atkinson, 1997) have been developed to simulate the source complexity. However, when used, empirical Green's function may poorly represent the variability of actual records in different source or site conditions. Other authors have concentrated their effort on accurate modelling of wave propagation in 3D structures based on finite-difference modelling (Kristek and Moczo, 2003). But these methods still demand extremely high computational resources and very good knowledge of the media.

To overcome these difficulties, a stochastic method that combines seismological models of the spectral amplitude of ground motion with the engineering notation that high frequency motions are basically random has been developed (Hanks, 1979; Hanks and McGuire, 1981; Boore, 1983; Boore, 2003). High frequency ground motions are modelled as band-limited Gaussian noise in which the radiated energy is distributed over a specified duration. The application of this method requires the spectral shape as a function of earthquake size (Boore, 1983; Atkinson and Somerville, 1994), which implies the knowledge of parameters characterizing the source process (e.g. corner frequency) and the wave propagation. However the variability of the motion is only taken into account by the random generation of the phase, and the method assumes stationarity of the frequency content with time.

The purpose of this chapter is to develop a stochastic model based on the Sabetta and Pugliese (1996) approach to simulate artificial accelerograms. The accelerogram simulations are derived from a spectrogram which is approximated by some strong motion indicators: the strong motion duration, the arias intensity, the central frequency and the frequency bandwidth of the signal. The values of these indicators are obtained through the use of ground motion prediction equations (attenuation relationships). The model allows to simulate the signal non stationarity, as well as the time dependence of the instantaneous power. In addition, this approach seems to be promising in modelling the ground motion natural variability.

The original Sabetta and Pugliese (1996) procedure was developed using a limited database of ground-motions recorded in Italy. Recently Pousse et al. (Pousse, 2005; Pousse et al., 2006) proposed an enhanced version of the procedure and derived the required predictive equations using ground-motions from the K-NET database (Kinoshita, 1998). Neither of these two papers investigated the performances of the so obtained ground-motions when used for structural analysis. In the present work, a modified version of the Sabetta and Pugliese (1996) procedure is developed using

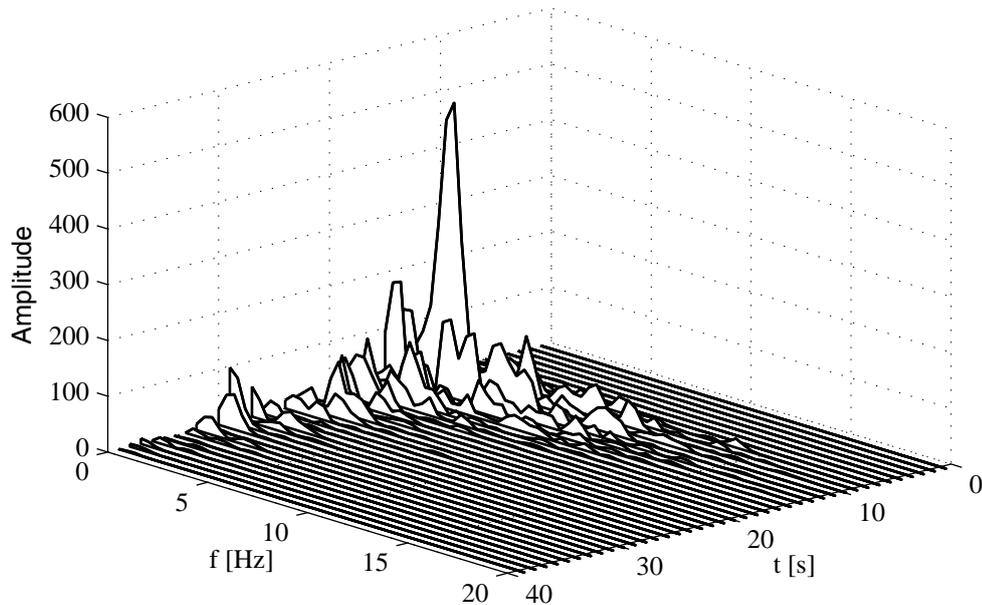
ground-motions from a subset of the NGA database and the performances of the so obtained accelerograms for non-linear structural analysis are assessed.

## 4.2 Model formulation

The simulation of non-stationary ground-motions is performed through an empirical model, where time dependence and frequency content of the signal are represented through a spectrogram,  $PS(t,f)$ , named physical spectrum after Mark (1970), and defined as

$$PS(t, f) = \left| \int_{-\infty}^{+\infty} w(t-u) x(u) \exp(-i2\pi fu) du \right|^2 \quad (4.1)$$

where  $t$  is time,  $f$  is frequency,  $x(u)$  is the ground-motion acceleration at time  $u$ , and  $w(t-u)$  is a running Hamming window. Figure 4.1 shows a spectrogram calculated for one of the accelerograms considered in the present work.



**Figure 4.1.** Example of spectrogram for one of the accelerograms in the database. The spectrogram can be considered as a time dependent power density spectrum.

Integration of equation (4.1) over time and frequency yields the Arias intensity,  $IA$ , of the accelerogram (Arias, 1970). This parameter is proportional to the total energy of the ground motion and is calculated as the integral of the squared acceleration.

The spectrogram can be factorized as a series of power spectral densities, calculated at different times  $PS_i(f)$ , which, according to the Sabetta and Pugliese's formulation (Sabetta and Pugliese, 1996) can be fitted with a lognormal function defined through tree

parameters derived from the theory of spectral moments (Vanmarcke, 1980; Lai, 1982). The above mentioned parameters are the average total power  $Pa$ , corresponding to the area under  $PS_i$ ; the central frequency,  $Fc$ , giving a measure of where the spectral density is concentrated along the frequency axis, and the frequency bandwidth,  $Fb$ , corresponding to the dispersion of the spectral density around the central frequency. Therefore, the factorized spectrogram can be written as

$$PS(t, f) = PS_i(f) Pa(t) \quad (4.2)$$

where both  $PS_i(f)$  and  $Pa(t)$ , are fitted through lognormal functions. Considering time dependence, the definition of the  $j$ -th spectral moment at time  $t$  becomes (Lai, 1982)

$$M_j(t) = \int_0^{+\infty} f^j PS(t, f) df \quad j = 0, 1, 2 \quad (4.3)$$

where  $f$  is frequency. The afore mentioned parameters used to approximate the spectrogram are defined as:

$$Pa(t) = M_0(t) \quad (4.4)$$

$$Fc(t) = \frac{M_1(t)}{M_0(t)} \quad (4.5)$$

$$Fb(t) = \left( \frac{M_2(t)}{M_0(t)} - \left( \frac{M_1(t)}{M_0(t)} \right)^2 \right)^{1/2} \quad (4.6)$$

$Pa(t)$ , the instantaneous average power, is the time envelope function describing the amplitude variation of the ground motion. Its integral in the time domain is equal to the integral of  $PS$  in the time frequency plane and therefore corresponds to the Arias intensity  $IA$ :

$$IA = \int a(t)^2 dt = \iint PS(t, f) dt df = \int Pa(t) dt \quad (4.7)$$

Central frequency,  $Fc(t)$  and frequency bandwidth,  $Fb(t)$ , represent the non-stationarity of the frequency content and correspond respectively to the centroid of  $PS$  and to the radius of gyration of  $PS$  with respect to  $Fc$  on the frequency plane.

With the above-defined parameters, it is possible to derive a lognormal function approximating the  $PS$  (Sabetta and Pugliese, 1996):

$$PS_{approx}(t, f) = \frac{Pa(t)}{\sqrt{2\pi}\delta f} \exp\left(-\frac{1}{2}\left(\frac{\log(f) - \log(\beta(t))}{\delta}\right)^2\right) \quad (4.8)$$

where  $\beta(t)$  and  $\delta$  are derived from  $Fc(t)$  and  $Fb(t)$  in the following way:

$$\log(\beta(t)) = \log(Fc(t)) - \delta^2/2 \quad (4.9)$$

$$\delta = \sqrt{\log(1 + Fb^2(t)/Fc^2(t))} \quad (4.10)$$

$Fc(t)$  and  $Fb(t)$  are derived through empirical ground motion prediction equations (see Section 4.4.5), while  $Pa(t)$  is supposed to follow a lognormal distribution

$$Pa(t) = \frac{IA}{t\sqrt{2\pi}\mu} \exp\left(-\frac{1}{2}\left(\frac{\log(t) - \sigma}{\mu}\right)^2\right) \quad (4.11)$$

with mean  $\mu$  and standard deviation  $\sigma$  empirically determined and depending on strong motion duration (see Section 4.4.5). As Eq. (4.11) shows, the integral of  $Pa(t)$  in the interval  $[0, \infty]$  is scaled to the Arias intensity,  $IA$ , in order to recover the predicted energy in the simulated accelerograms. The reference values of the Arias intensity are calculated through an empirical attenuation relationship (see 4.4.3).

### 4.3 Computation of the time history

To generate accelerograms Sabetta and Pugliese (1996) assumed that the ground motion at time  $t$  results from the contribution of random uncorrelated phases. In the frequency range of interest for engineers, this assumption is physically reasonable because of the complexity of the source and propagation effects (Boore, 2003).

The accelerograms are simulated by calculating the inverse Fourier transform  $\alpha(t, \mu)$  of  $\sqrt{PS(t, f)} \exp(i\varphi(f))$ , where  $\varphi$  is a random phase:

$$\alpha(t, \mu) = \frac{1}{2\pi} \int \sqrt{PS(t, f)} \exp(i2\pi f\mu + i\varphi(f)) df \quad (4.12)$$

Using the definition of  $PS$  in Eq. (4.1) and taking  $\mu = t$ , Eq. (4.12) yields the following inversion formula:

$$a(t) = \frac{1}{2\pi} \int \sqrt{PS(t, f)} \exp(i2\pi ft + i\varphi(f)) df \quad (4.13)$$

The simulated accelerogram corresponds to the sum, of a Fourier series terms with amplitude  $A(t, f) = \sqrt{PS(t, f)}$  and uniformly distributed random phase  $\phi$ . Since  $PS(t, f)$  depends on the central frequency  $Fc(t)$  and on frequency bandwidth  $Fb(t)$  the time histories are non-stationary.

## 4.4 Estimation of the functions defining the spectrogram

The database originally used by Sabetta and Pugliese (1996) contained 190 horizontal components and 95 vertical components of strong motions recorded from 17 Italian earthquakes. The distribution of recordings with respect to magnitude, distance, and site classification is given in Figure 4.2. Acceleration time histories were baseline and instrument corrected with a frequency domain routine. The cut-off frequencies for the high-pass filtering ranged from 0.2 to 0.7 Hz with typical values of 0.3 Hz; the cut-off frequencies for the low-pass filtering ranged from 20 to 35 Hz with typical values of 25 Hz. The magnitude scale adopted by Sabetta and Pugliese corresponded to the surface-wave magnitude ( $M_s$ ) when both local magnitude ( $M_L$ ) and  $M_s$  were greater than or equal to 5.5 and corresponded to  $M_L$  when magnitude was lower than this value. Moment magnitude (Hanks and Kanamori, 1979), which unlike other scales does not saturate at high magnitudes, was not used because reliable estimates of seismic moments were not available for some of the smaller earthquakes used in the regression analysis.

In the present work, the ground motion prediction equations developed by Sabetta and Pugliese have been revisited using a new and more comprehensive database. Accelerograms were selected from a subset of the records in the Next Generation of Attenuation (NGA) project database (Power et al., 2006; Power et al., 2008). A total of 1504 accelerograms were used for the analysis. This subset was defined by excluding all records from the Chi-Chi earthquake sequence, all pulse-like records, all records with only one horizontal component and records for which appropriate metadata were not available (moment magnitude,  $M_w$ , Joyner-Boore distance  $R_{JB}$  and shear wave velocity over the upper 30 meters,  $V_{S30}$ ). Pulse like ground-motions were identified using the procedure proposed by Baker (2007d). All the selected records have a maximum usable period of at least 3 seconds (Boore and Bommer, 2005). The Mw 7.6 Chi-Chi, Taiwan, earthquake was recorded in 1999 by a very large number of strong motion instruments very close to the rupture surface of the event. Consequently, this event may potentially increase the number of near source ground motion records. However, it is thought that the ground motions from this event are lower than those that might be expected from future earthquake of similar magnitude. Because this event

contributes such a large number of near field records, the inclusion of these records would have a significant impact upon the nature of the near source scaling of ground motion measures. However, the impact of including this event in a regression analysis is to cause the resulting equations to underestimate ground motions in this important near field region. For this reason Campbell and Bozorgnia (Campbell and Bozorgnia, 2003a; Campbell and Bozorgnia, 2003b; Campbell and Bozorgnia, 2004) refrained from including records from the Chi-Chi event. Speculated reasons for the lower than expected ground motions from the Chi-Chi event include large amounts of surface rupturing, large amounts of slip on the fault plane and long rise times over the fault (Campbell and Bozorgnia, 2003b). Currently these effects are not taken into account in empirical regression equations.

On the other hand it is noteworthy, that the coefficients of the predictive equations are obtained using mixed effect models, are able to take into account systematic differences between events. These models are employed in the present work and there may be ample justification for including the Chi-Chi events, but it should be noted that the structure of the mixed model adopted to develop predictive equations is not capable to account for systematic events which are not simply additive in terms of the regressed parameter. A complete list of the ground motions used is provided in Appendix A.

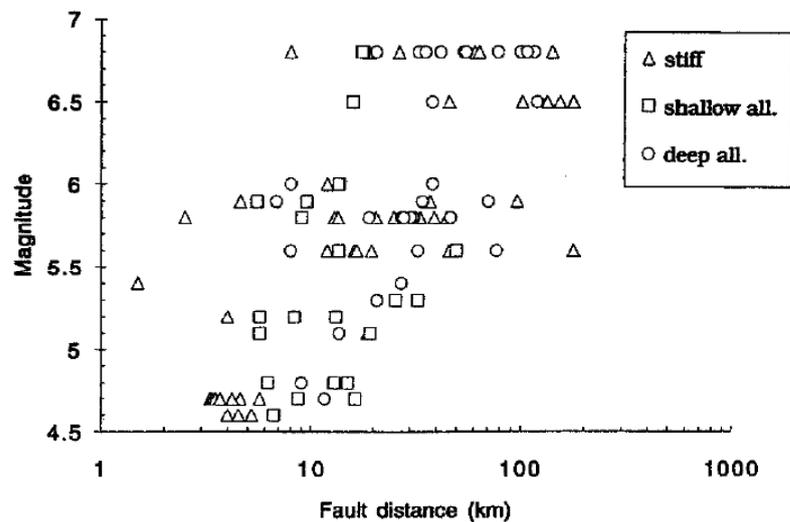


Figure 4.2. Magnitude-distance distribution of the ground-motion database originally used by Sabetta and Pugliese (1996) (after Sabetta and Pugliese, 1996).

#### 4.4.1 Bidirectional issues

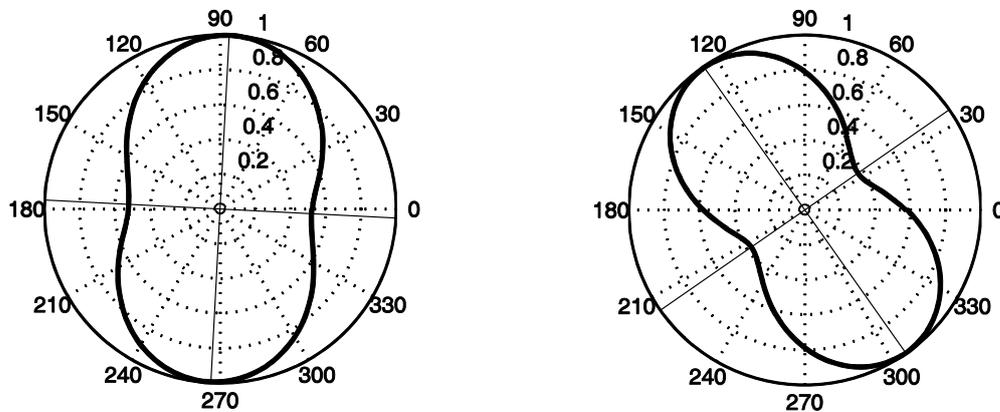


Figure 4.3. Normalized Arias intensity values for recording 1093 (left panel) and 1099 (right panel), calculated for different directions. Directions 0° and 90° correspond to the as recorded components.

One of the problems that one must face when developing ground-motion prediction equations is the definition of the component of the ground motion to be used. In fact, excluding the vertical component, each recording contains two orthogonal acceleration time histories. The orientation of these components with respect to the rupture direction is usually random, depending on the sensor orientation. Different approaches have been proposed in the literature for dealing with this issue. A briefly summary has been given in Section 2.4. Further details can be found in Douglas (2003), Baker and Cornell (2006c), Boore et al. (2006), Watson-Lamprey and Boore (2007), Beyer and Bommer (2006) and Beyer and Bommer (2007). One of the most widely adopted approaches is to use the geometric mean, of parameter under investigation, of the two components. This approach can have some drawbacks, in fact, if on one hand it gives a more stable prediction of the mean response, on the other hand it can provide a reduced estimate of the ground-motion variability. This issue can be overcome using the procedure proposed by Joyner and Boore (Joyner and Boore, 1993; Joyner and Boore, 1994) which was adopted in the previous chapter when fitting attenuation relationships (See 3.4). Sabetta and Pugliese (1996) adopted a different approach. For the Arias intensity,  $IA$  (Arias, 1970), the horizontal component of the motion having the larger value of  $IA$  was used. This approach gives higher values than the previous one but predicts a lower variability since only one component per ground motion is used.

Even using the Joyner and Boore's procedure (Joyner and Boore, 1993; Joyner and Boore, 1994) one can argue that using the two components of the ground-motion as recorded one can underestimate the actual variability of the ground motion. As an example, consider the Arias intensity as measure of the ground motion intensity. Using the simple rotation formula

$$\begin{pmatrix} a_{x'}(t) \\ a_{y'}(t) \end{pmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{pmatrix} a_x(t) \\ a_y(t) \end{pmatrix} \quad (4.14)$$

where  $a_x(t)$  and  $a_y(t)$  are the recorded accelerograms and  $a_{x'}(t)$  and  $a_{y'}(t)$  are the acceleration time histories after a rotation  $\theta$  of the axes, one can calculate the  $IA$  values at different angles. Figure 4.3 shows the dependence of  $IA$  values on the angle  $\theta$  for two of the recordings in the considered database. The values of Arias intensity have been normalized to their respective maxima. As figure Figure 4.3b suggests, the maximum and the minimum values of  $IA$  may not correspond to the orientation of the sensors ( $0^\circ$  and  $90^\circ$  in the figure). Therefore in the case depicted in Figure 4.3a considering the as recorded components to develop an attenuation relationship will give a good prediction of the variability of  $IA$ , on the other hand considering the as recorded components depicted in Figure 4.3b gives lower estimates of variability of  $IA$ . Similar considerations can be made for different measures of the ground-motion intensity, such as PGA and spectral accelerations. These aspects may have consequences on structural reliability assessment because the orientation of a structure is usually fully random with respect to the direction of the possible ground-motions (see Section 2.4). Therefore it is very important to have sound estimates of both mean value and variance of the considered ground motion intensity measures.

To overcome these difficulties, in the present work, for each couple of accelerograms in the database, Arias intensity was calculated for all the possible rotation angles (with  $1^\circ$  steps) and the accelerograms in the direction of the maximum and minimum values have been used in order to obtain a good estimate of the ground motion variability. The considered durations (see Section 4.4.4) are those of the so obtained components.

#### 4.4.2 Power amplitude

According to Sabetta and Pugliese (1996), power amplitude,  $Pa(t)$ , is approximated through a lognormal function the parameters of which depend on strong motion duration  $DV$ . Furthermore the unit area under the lognormal function is scaled to the Arias intensity,  $IA$ :

$$Pa(t) = IA \cdot \frac{1}{t\sqrt{2\pi\mu}} \exp\left(-\frac{1}{2}\left(\frac{\log(t) - \sigma}{\mu}\right)^2\right) \quad (4.15)$$

where the mean value,  $\mu$ , and the standard deviation,  $\sigma$ , of the lognormal function are defined by Sabetta and Pugliese as:

$$\mu = \log(T2) + \sigma \quad \sigma = \log(T3/T2)/2.5 \quad (4.16)$$

with:

$$\begin{aligned} T1 &= R/7 \\ T2 &= T1 + 0.5 \cdot DV \\ T3 &= T1 + 2.5 \cdot DV \\ Tot.duration &= 1.3 \cdot T3 \end{aligned} \quad (4.17)$$

where  $T1$  corresponds to the time delay in seconds between S waves, with propagation velocity  $V_s$ , and P waves, with propagation velocity  $V_p$ , and is calculated by dividing the distance  $R$  (in kilometres) by the factor  $V_p \cdot V_s / (V_p - V_s)$  assumed to be equal to 7 km/sec. The choices of  $T2$ ,  $T3$  and  $\sigma$  were derived in order to have *i*) a modal value of the lognormal function in Eq. (4.15) at time  $t = T2$ , correlated to the distance; *ii*) a standard deviation proportional to the strong-motion duration  $DV$ ; *iii*) an area equal to the Arias intensity  $IA$ ; and *iv*) a total duration 30% greater than the value of  $T3$ , corresponding to the modal value plus  $2 \cdot DV$ . To specify the duration of the ground motion,  $DV$ , Sabetta and Pugliese adopted the definition given by Vanmarcke and Lai (1980). In the present work, the definition of duration given by Trifunac and Brady (1975) was adopted. According to this definition, duration is related to the release of ground motion energy and it is defined as the time during which the 90% of the Arias intensity is released. Figure 4.4 explains the adopted definition of duration showing a Husid plot (Husid, 1969) for one of the considered accelerograms. This plot shows the increase of Arias intensity over time, in fact it shows the values of the function

$$H(t) = \frac{\int_0^t a^2(t) dt}{\int_0^{t_f} a^2(t) dt} = \frac{\int_0^t a^2(t) dt}{IA} \quad (4.18)$$

where  $t_f$  is the total accelerogram duration. Of course, the values given by (4.18) are in the interval  $[0,1]$ . Accordingly to the definition adopted, the duration is written as

$$DV = t_{0.05-0.95} = t|_{H=0.95} - t|_{H=0.05} \quad (4.19)$$

where  $t | H = x$  indicates time at which Eq. (4.18) gives a value equal to  $x$ . Since, in the present work a different definition of ground-motion duration and a different database have been used the coefficients in Eq. (4.17) have been modified.

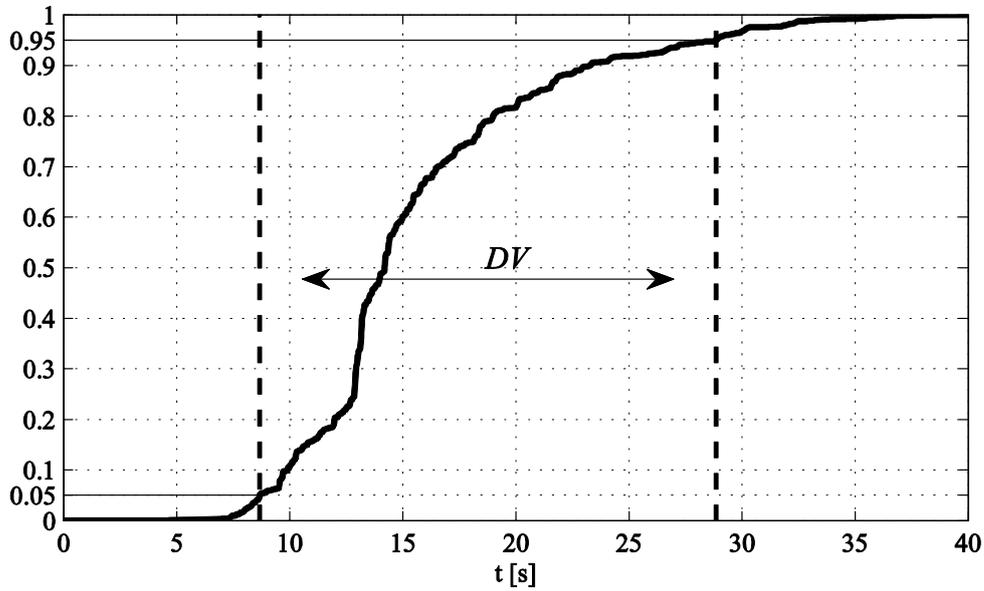


Figure 4.4. Husid plot for one of the ground motions contained in the database considered in the present work.

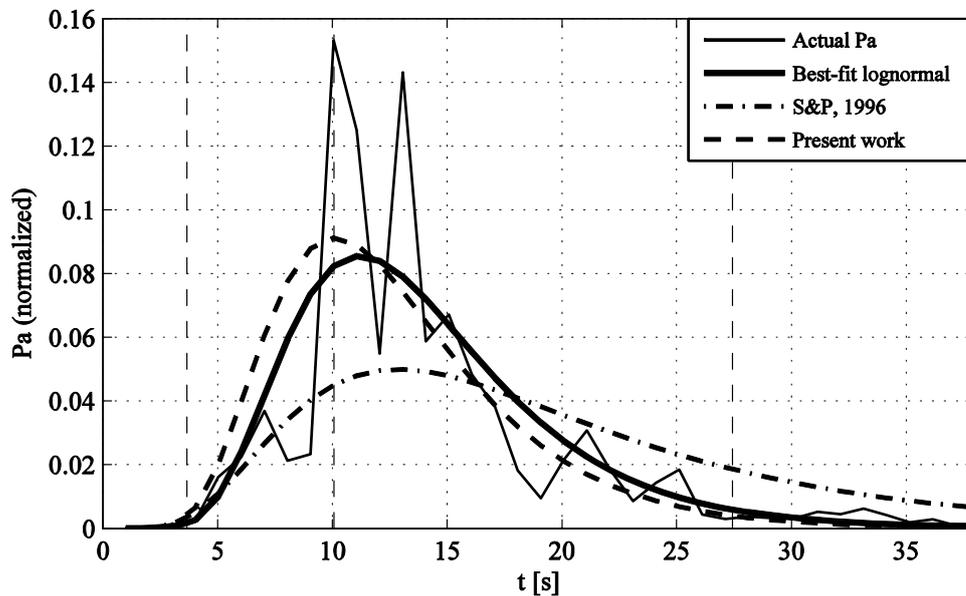


Figure 4.5. Comparison power amplitude,  $Pa(t)$ , defined according to different criteria: actual accelerogram  $Pa$  (thin solid line), best fit lognormal (solid thick line), original Sabetta and Pugliese model (dash dot line) and model developed in the present study (dashed line). Curves are normalized to unit area.

In order to do so, Power amplitude,  $Pa(t)$ , was calculated for each accelerogram considered using Eq. (4.4) and fitted using a lognormal function. Then the values of the parameters in Eq. (4.17) were modified in order to obtain the best average fit over the whole ground-motion database. The so obtained relationships are given in Eq. (4.20)

$$\begin{aligned}
 T1 &= R/7 \\
 T2 &= T1 + 0.35 \cdot DV \\
 T3 &= T1 + 1.3 \cdot DV \\
 Tot.duration &= 1.3 \cdot T3
 \end{aligned}
 \tag{4.20}$$

where the symbols have the same meaning of those in Eq. (4.17). Figure 4.5 shows a comparison of the original and the modified model against the best fitting lognormal function for the power amplitude calculated for one of the considered accelerograms.

### 4.4.3 Arias Intensity

As previously stated the lognormal function approximating power amplitude is scaled in order to obtain an area equal to Arias intensity. The reference value for this latter parameter is estimated through a ground motion attenuation relationship. In the present work the following expression has been used in order to express attenuation of Arias intensity

$$\begin{aligned}
 \log_{10}(IA) &= c_1 + c_2 M + c_3 (M - 6)^2 \\
 &+ (c_4 + c_5 M) \log_{10} \left( \sqrt{R_{JB}^2 + c_6^2} \right) + c_7 \log_{10}(V_{S_{30}})
 \end{aligned}
 \tag{4.21}$$

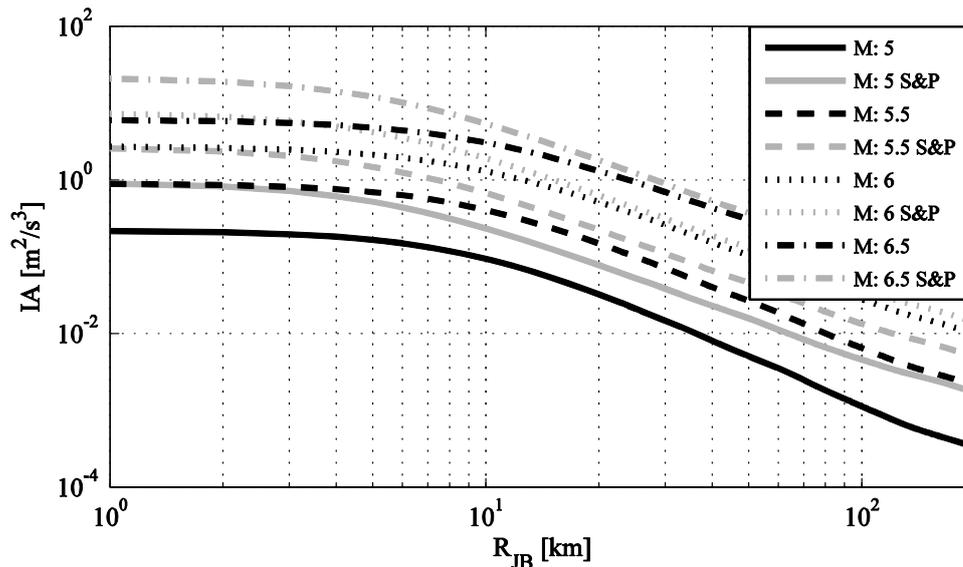
where  $IA$  is Arias intensity,  $M$  is moment magnitude,  $R_{JB}$  is Joyner and Boore distance and  $V_{S_{30}}$  is shear wave velocity over the last 30 meters. As done in Section 3.4 the total standard deviation,  $\sigma_T$ , of each empirical model is decomposed into three independent components: the inter-event standard deviation,  $\sigma_E$ ; the intra-event standard deviation,  $\sigma_A$ ; and the inter-component standard deviation,  $\sigma_C$ , which must be considered when both horizontal components of a ground-motion recording are used in the regression analysis, as done in this study. It is worth recalling that the two horizontal components used in the present work are those having the maximum and the minimum Arias intensity. The estimates of the regression parameters and of the variance are listed in Table 4.1. Figure 4.6 and Figure 4.7 compare the median values of Arias intensity as predicted by Eq. (4.21) for a rock site, with those predicted by the attenuation models proposed by Sabetta and Pugliese (1996).

**Table 4.1. Predictions of regression coefficients and standard deviations of Eq. (4.21) for Arias Intensity.**

$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$	$\sigma_E$	$\sigma_C$	$\sigma_A$
1.42	0.54	-0.27	-3.68	0.29	9.36	-0.88	0.23	0.10	0.35

It is worth recalling that the definitions of magnitude and distance adopted in the present work are different from those adopted by Sabetta and Pugliese (1996), therefore the results of the comparisons should be taken with care. Though, it can be concluded that the Sabetta and Pugliese's attenuation relationships predict larger Arias intensities over the whole magnitude-range considered. This can be explained considering that those regressions were performed on the larger Arias intensity among the two recorded components, while in the present work, the models have been fitted using two components of the ground motion.

As done in the previous chapter the goodness of fit of the regression was tested using different diagnostic plots (Pinheiro and Bates, 2000) which are given in Figures 4.8-4.14. Figure 4.8 depicts an histogram of the fitted data and allows to check the normality assumption. Figures from Figure 4.9 to Figure 4.12 give plots of the standardized residuals against the fitted values and against the independent variables of the regression model. From these plots one can conclude that the residuals are independent of the fitted value and on the independent variables of the regression model. Finally using Figure 4.13 and Figure 4.14 it is possible to check the normality of the standardized residuals.

**Figure 4.6. Attenuation of Arias intensity over distance, as predicted by Eq. (4.21).**

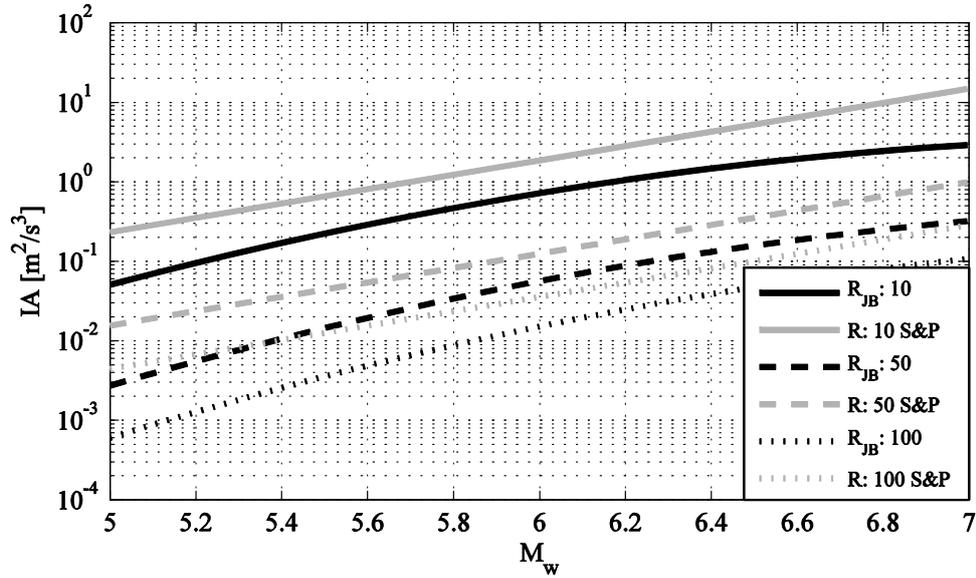


Figure 4.7. Attenuation of Arias intensity over magnitude as predicted by Eq. (4.21).

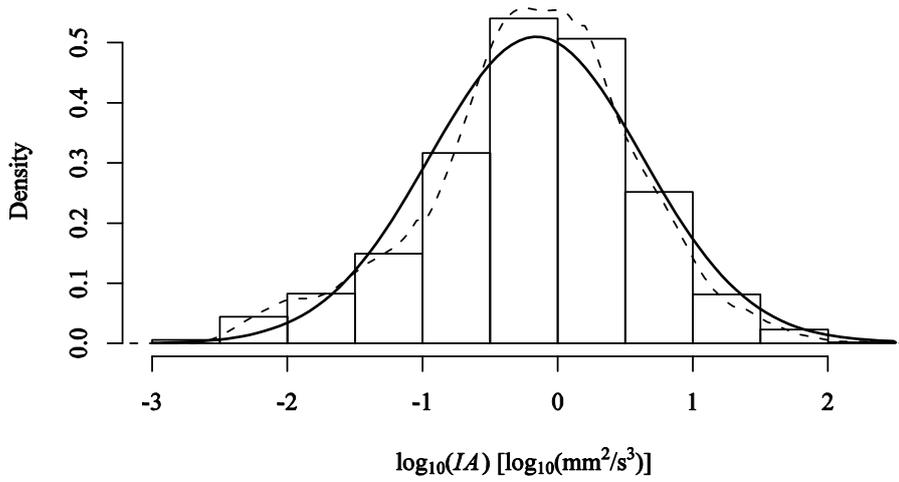


Figure 4.8. Histogram of  $\log_{10}(IA)$  used as input data for regression. Solid line represents a normal distribution and dotted line is obtained using a kernel smoothing procedure.

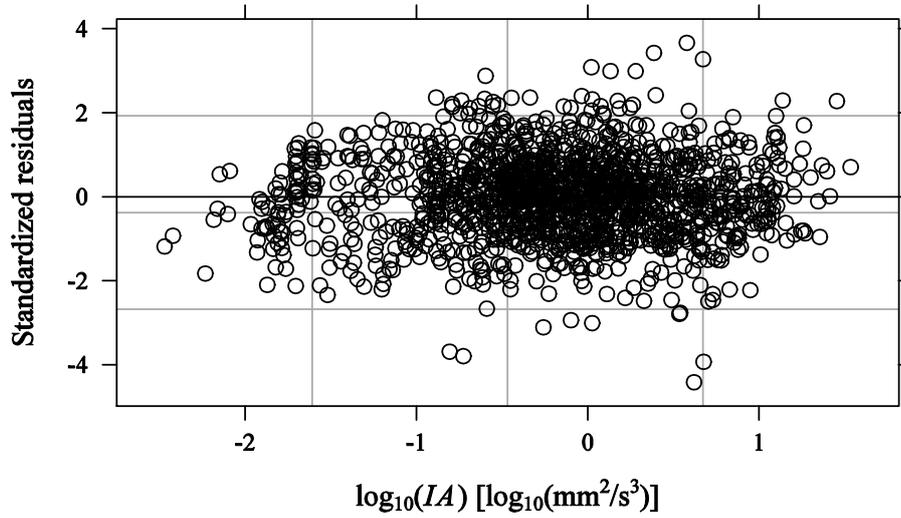


Figure 4.9. Standardized residuals of IA regression versus fitted values.

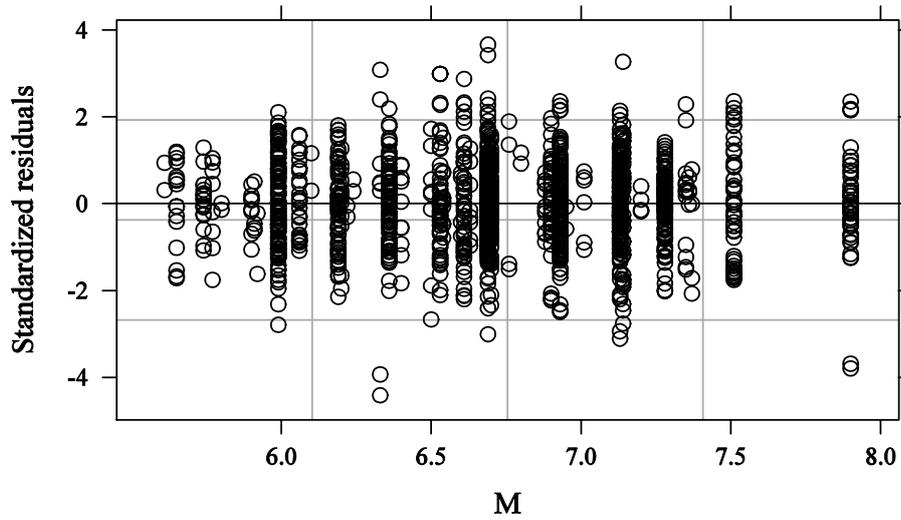


Figure 4.10. Standardized residuals of IA regression versus moment magnitude.

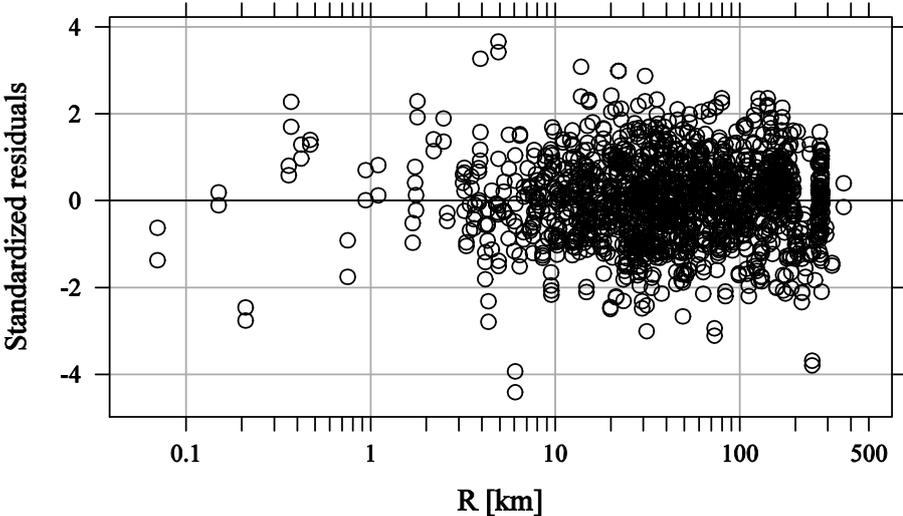


Figure 4.11. Standardized residuals of IA regression versus distance.

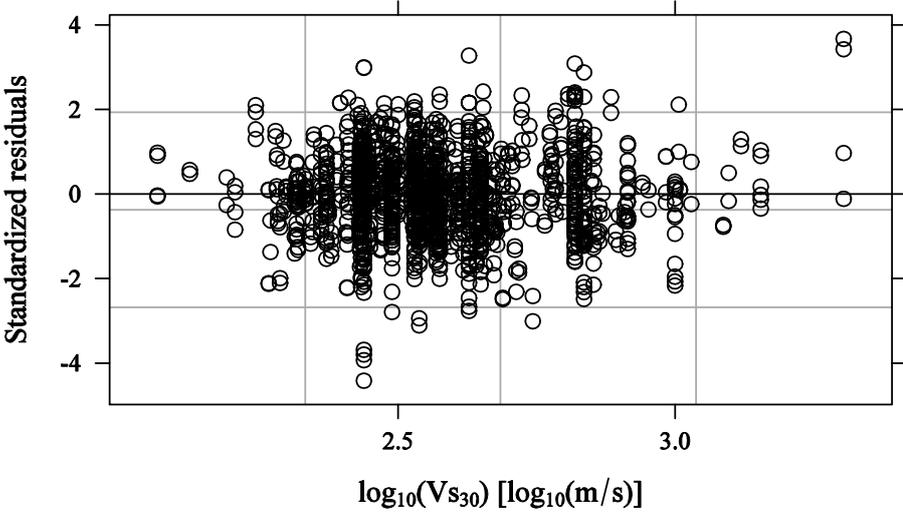


Figure 4.12. Standardized residuals IA regression versus  $\log_{10}(V_{s30})$ .

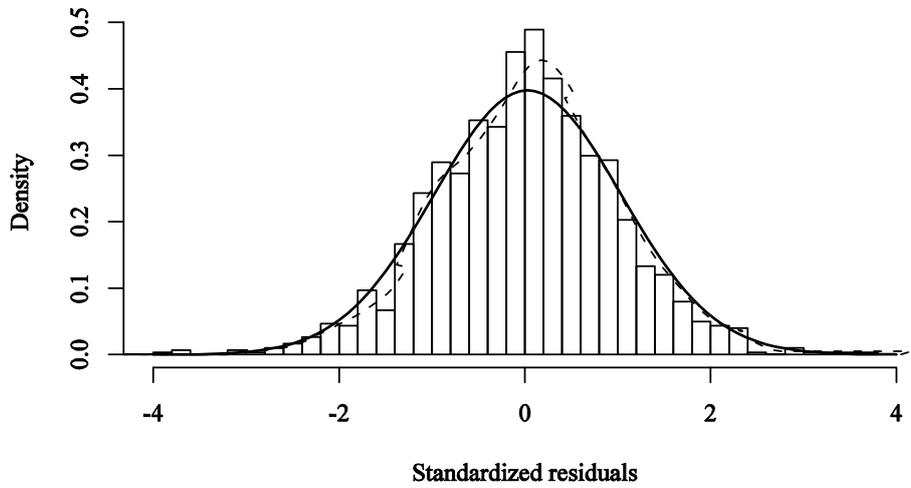


Figure 4.13. Histograms of standardized residuals of IA regression. Solid line shows a standard normal distribution and the dashed line is obtained using a kernel smoothing procedure.

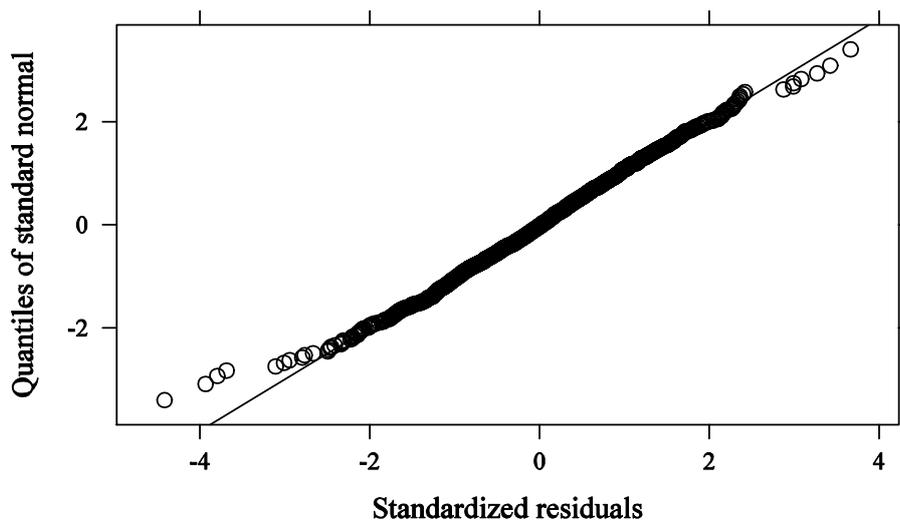


Figure 4.14. Normal quantile-quantile plot for standardized residuals of IA regression.

#### 4.4.4 Duration

Another parameter used to approximate the spectrogram is ground-motion duration. Therefore in order to generate an accelerogram one needs to estimate of its value using an attenuation relationship.

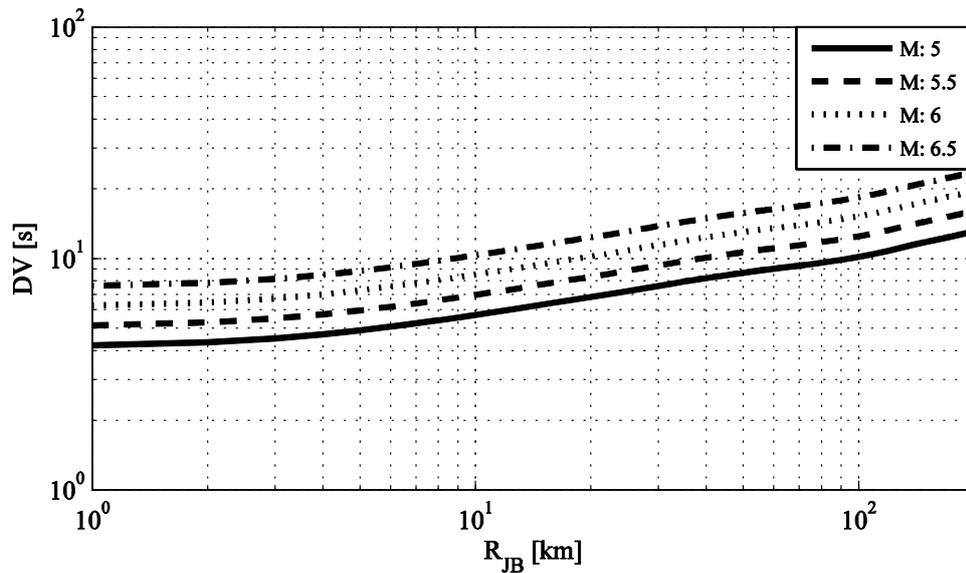
The expression adopted to develop a ground-motion prediction equation for duration is given in Eq. (4.22)

$$\log_{10} DV = c_1 + c_2 M + c_3 \log_{10} \left( \sqrt{R_{JB}^2 + c_4^2} \right) + c_5 \log_{10} (Vs_{30}) \quad (4.22)$$

where the symbols have the same meaning as in Eq. (4.21) and  $c_1, \dots, c_5$  are regression parameters to be determined. Comparing Eq. (4.21) with Eq. (4.22) it is evident that the latter has a simpler form. This simpler form was defined accordingly to the results of significance tests performed on the regression parameters of a model with the form given in Eq. (4.21). Regression results are reported in Table 4.2. Figure 4.15 and Figure 4.16 show the attenuation of duration predicted by Eq. (4.22). Since Sabetta and Pugliese (1996) used a different definition of duration no comparison is possible. Diagnostic plot for the regression are given in Figures 4.17-4.23. These plots confirm the assumption of normality of the base ten logarithm of duration and show that the standardized residuals are independent of the fitted values and of the independent variables of the model

**Table 4.2. Estimates of regression parameters and standard deviations of Eq. (4.22).**

$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$\sigma_E$	$\sigma_C$	$\sigma_A$
0.41	0.17	0.28	3.36	-0.30	0.13	0.04	0.15



**Figure 4.15. Attenuation of duration over distance as predicted by regression.**

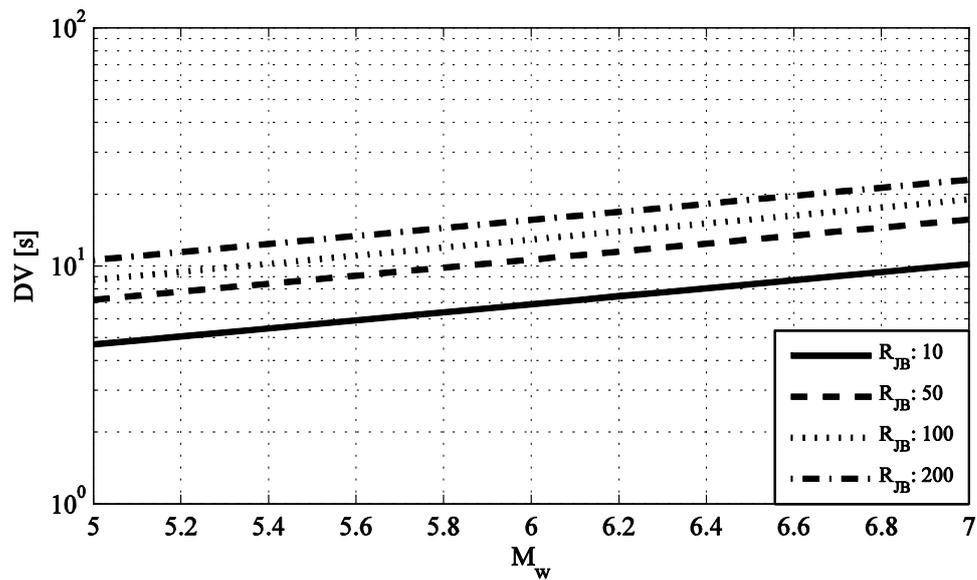


Figure 4.16. Attenuation of duration over magnitude as predicted by regression.

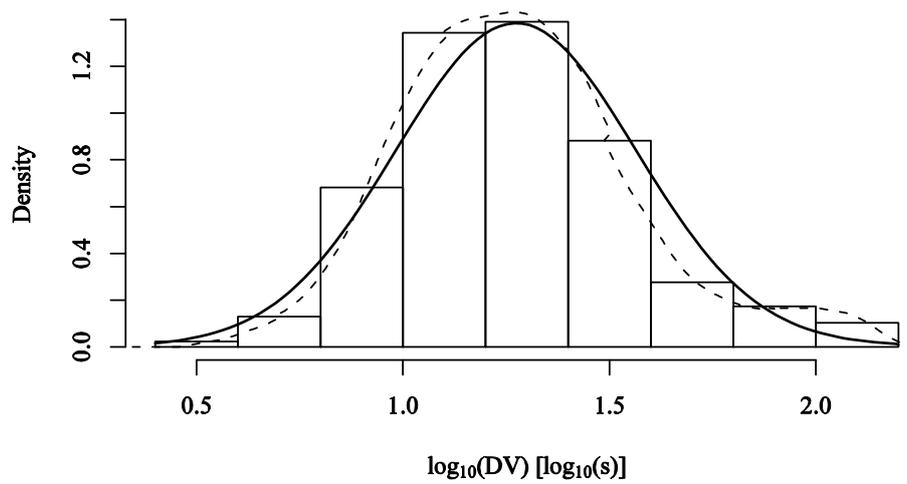


Figure 4.17. Histograms of  $\log_{10}(DV)$  used as input data for regression. Solid line shows a normal distribution and the dotted line is obtained using a kernel smoothing procedure.

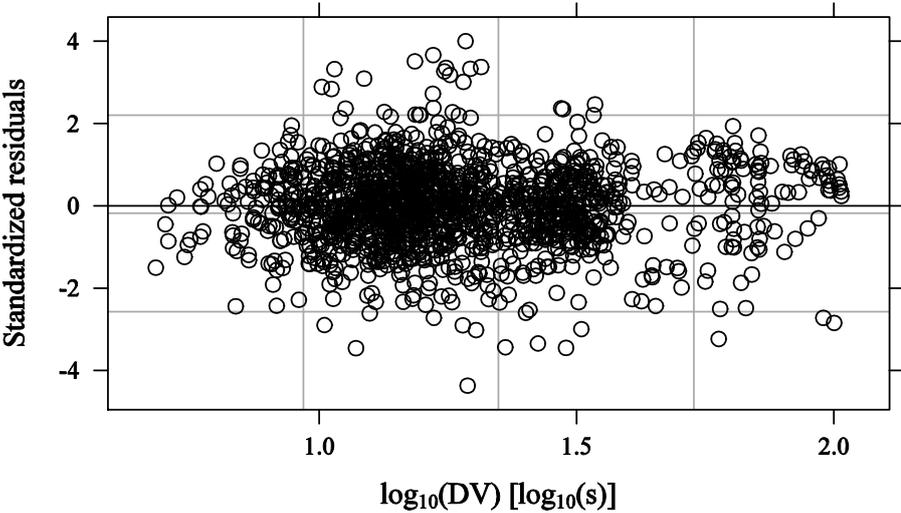


Figure 4.18. Standardized residuals versus fitted values.

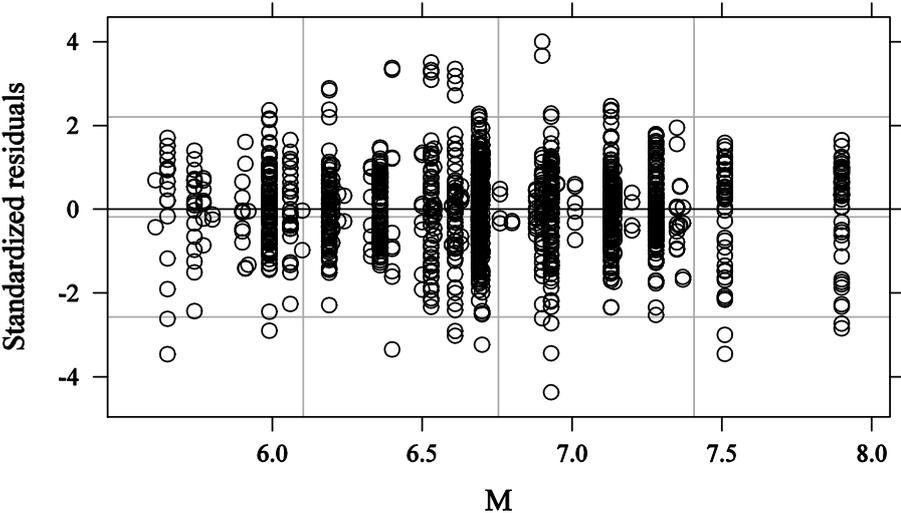


Figure 4.19. Standardized residual against magnitude for duration regression.

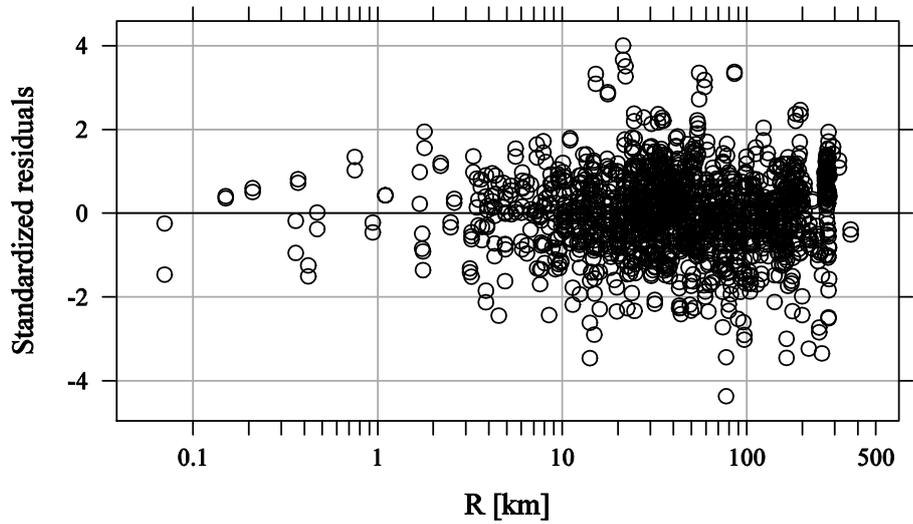


Figure 4.20. Standardized residuals of duration regression versus distance.

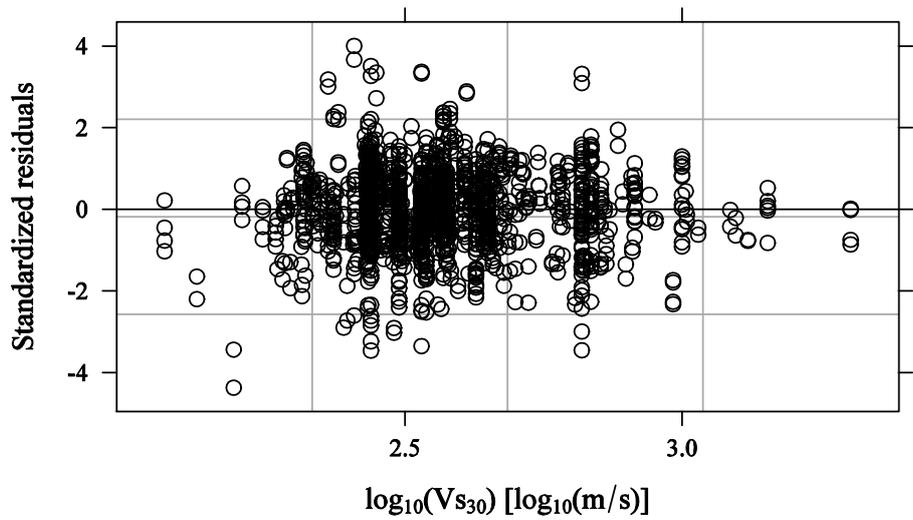


Figure 4.21. Standardized residuals of duration regression versus  $\log_{10}(V_{s30})$ .

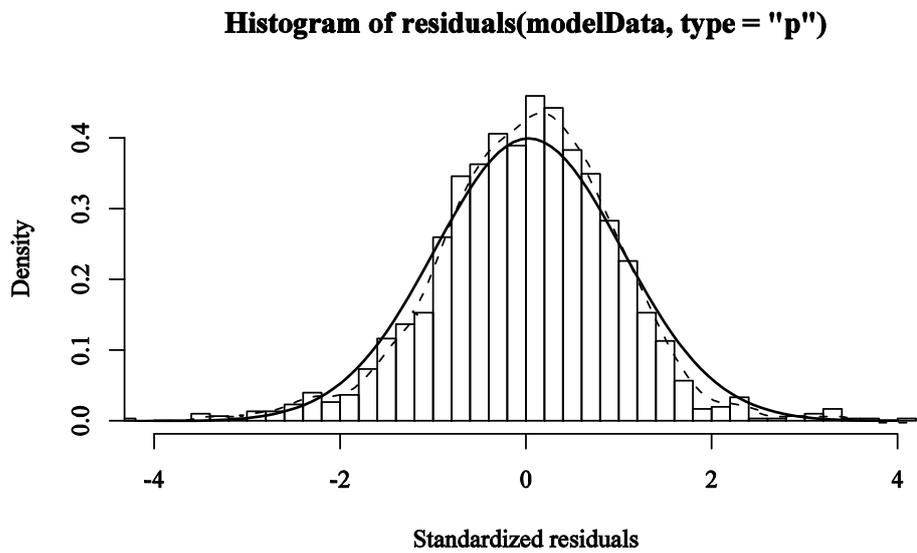


Figure 4.22. Histogram of standardized residuals of duration regression. Solid line is a standard normal distribution and the dashed line is obtained using a kernel smoothing procedure.

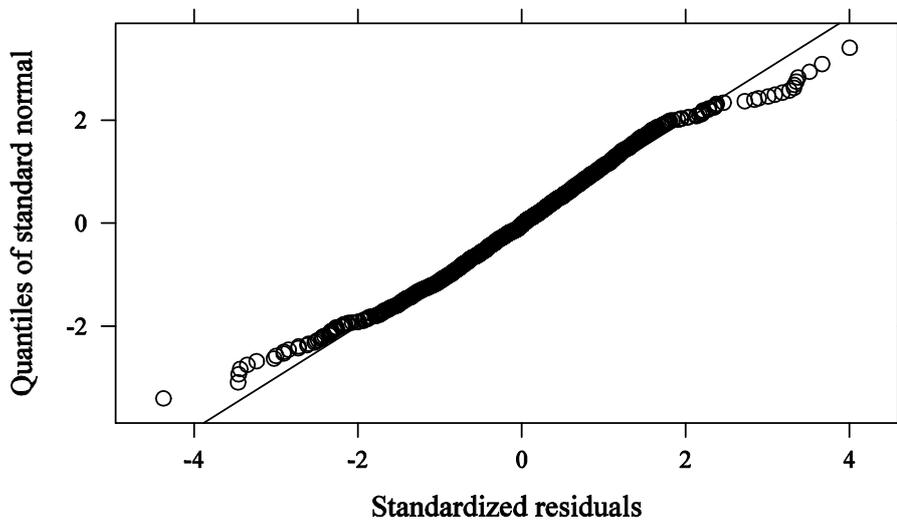


Figure 4.23. Normal quantile-quantile plot for duration regression standardized residuals.

#### 4.4.5 Central frequency and frequency bandwidth

As far as the predictive equation for central frequency ( $F_c$ ) is concerned, the exponential decay originally proposed by Sabetta et al. (1986) has been maintained but

the predictive model has been modified with respect to that proposed by Sabetta and Pugliese (1996) which was

$$\log(Fc) = 3.4 - 0.35 \log(t) - 0.218M - 0.15S \quad (4.23)$$

where  $t$  is time and  $S$  is a dummy variable used to take into account soil conditions has been modified as follows

$$\log(Fc) = c_1 + c_2M + (c_3 + c_4M + c_5 \log_{10}(Vs_{30})) \log(t + 1.25) + c_6 \log_{10}(Vs_{30}) \quad (4.24)$$

The structure of the variance was partitioned into three components similarly to what was done for  $IA$  and  $DV$ . The definitions of intra-event and of inter-event variances was maintained but the definition of the variance,  $\sigma_C$ , related to the two components of the ground motion was modified. In fact in this case  $\log(Fc)$  depends on time,  $t$ , and therefore there is more than one  $Fc$  value for each component. Conceptually the meaning of  $\sigma_C$  is unchanged but it must be calculated in a different way. Estimates of regression parameters and standard deviations of the model are given in Table 4.3. In Figure 4.24 the  $Fc$  values predicted by the present model are compared with those predict by the Sabetta and Pugliese's model and with some data points used to fit the models. Figure 4.25 compares the prediction, for different magnitudes, of the present model with those of the Sabetta and Pugliese's model (Sabetta and Pugliese, 1996). Diagnostic plots for the regression on  $Fc$  are given in Figures 4.26-4.30.

**Table 4.3. Estimates of the regression parameters and of the standard deviations of the regression model for  $Fc$ .**

$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$\sigma_C$	$\sigma_E$	$\sigma_A$
2.60	0.20	-1.46	-0.04	0.44	-0.47	0.36	0.30	0.43

The second parameter defining the frequency content of the ground motion is frequency bandwidth  $Fb$ . Following Sabetta and Pugliese (1996) the predictive equation has been defined in terms of the ratio  $Fb/Fc$ . The adopted functional model is

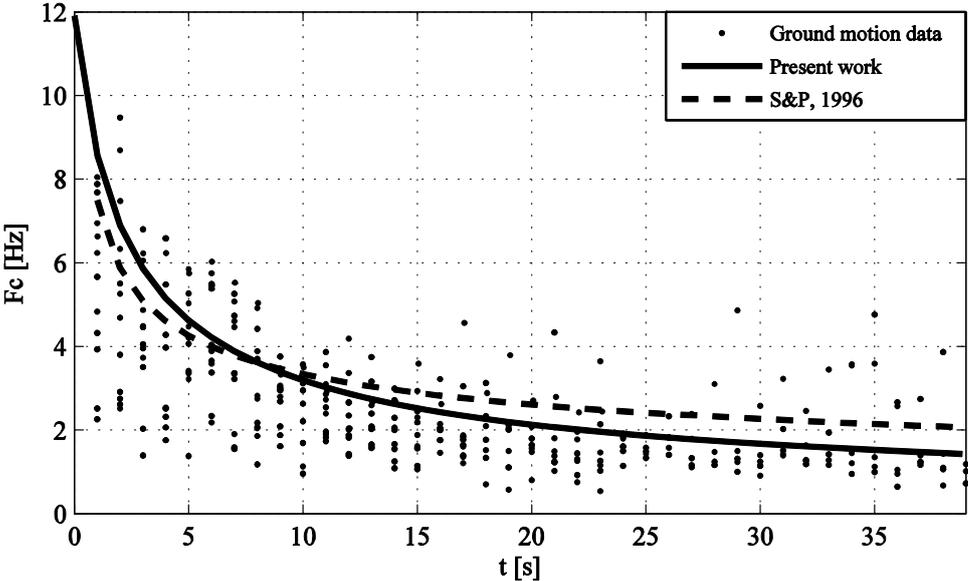
$$\log\left(\frac{Fb}{Fc}\right) = c_1 + c_2M + c_3 \log_{10}(Vs_{30}) + c_4t \quad (4.25)$$

In contrast to the predictive equation proposed by Sabetta and Pugliese (1996) time dependence was included in the predictive equation for  $Fb/Fc$ . Estimates of regression parameters and standard deviations of the model in Eq. (4.25) are given in Table 4.4. Figure 4.32 compares the prediction of the ratio  $Fb/Fc$  predicted by the present model

(see Eq. (4.25)) and by the Sabetta and Pugliese’s model with some of the data points used in the present work. Figure 4.31 depicts predictions for different magnitudes. Diagnostic plots for the regression on  $Fb/Fc$  are given in Figures 4.33-4.37.

**Table 4.4.** Estimates of the regression parameters and of the standard deviations of the regression model for  $Fb/Fc$ .

$c_1$	$c_2$	$c_3$	$c_4$	$\sigma_C$	$\sigma_E$	$\sigma_A$
0.42	0.0026	-1.46	0.0038	0.20	0.12	0.32



**Figure 4.24.** Dependence of  $F_c$  on time for ground motions 905, 908, 768, 913, 873, 767, 821, 860, 914 (see Appendix A): actual data (dots), prediction my the model in Eq. (4.25) and prediction by the Sabetta and Pugliese (1996) model.

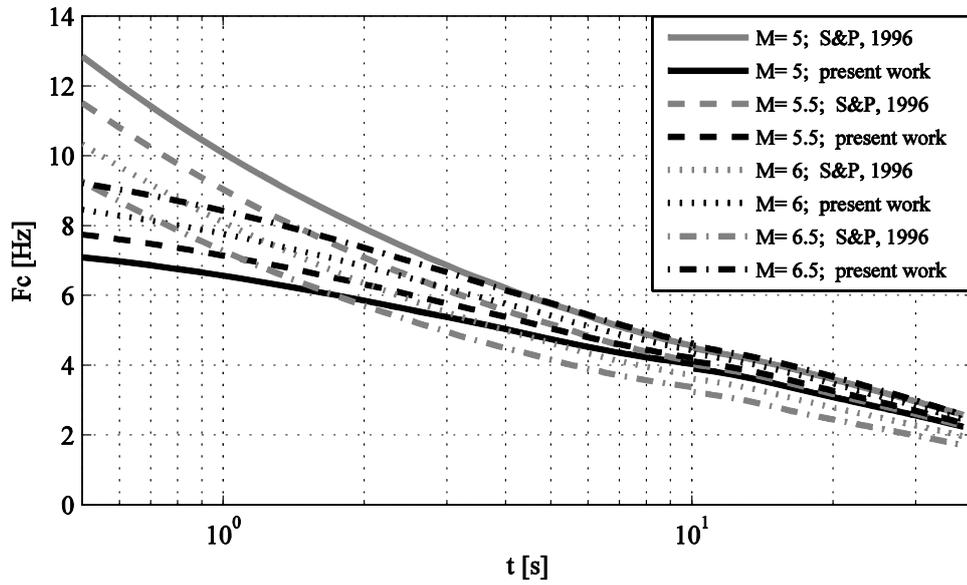


Figure 4.25. Predictions of  $F_c$  for different magnitudes as given by the model in Eq. (4.25) and by the Sabetta and Pugliese (1996) model.

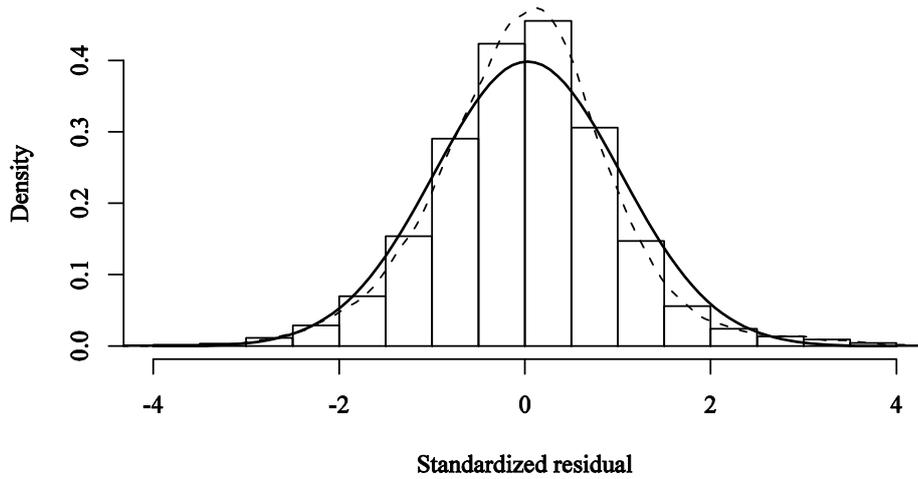


Figure 4.26. Histogram of standardized residuals of  $F_c$  regression. Solid line is a standard normal distribution and dashed line is obtained using a kernel smoothing procedure.

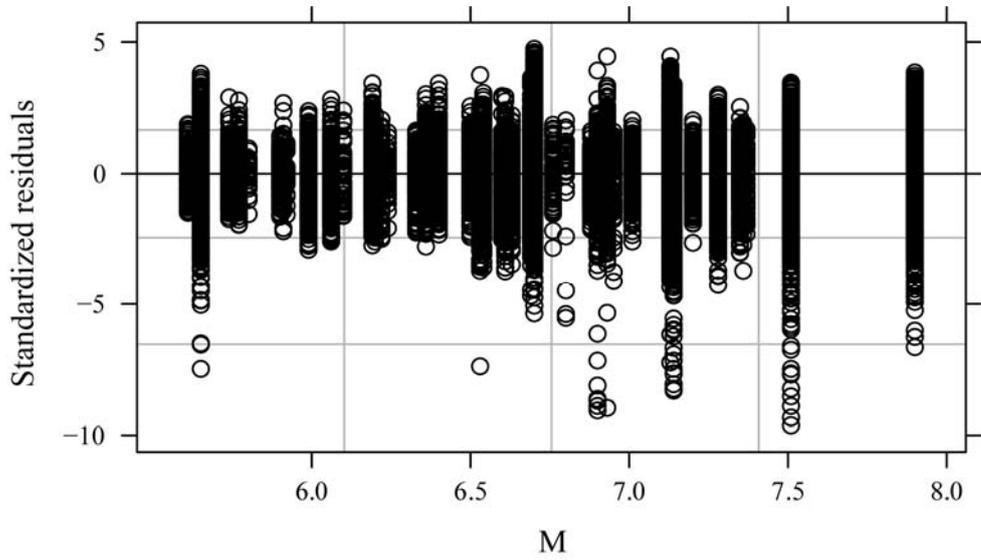


Figure 4.27. Standardized residuals of  $F_c$  regression versus magnitude.

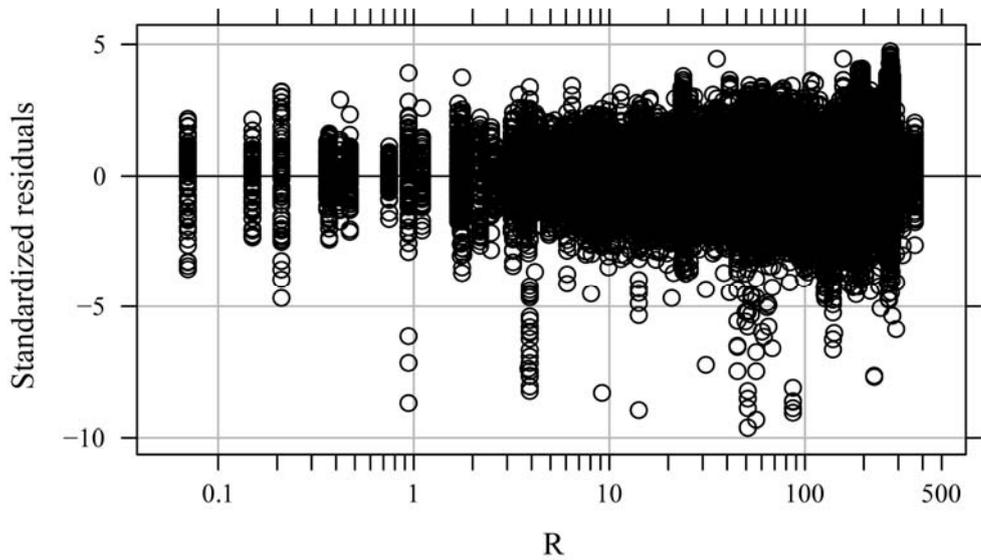


Figure 4.28. Standardized residuals of  $F_c$  regression versus distance (log scale).

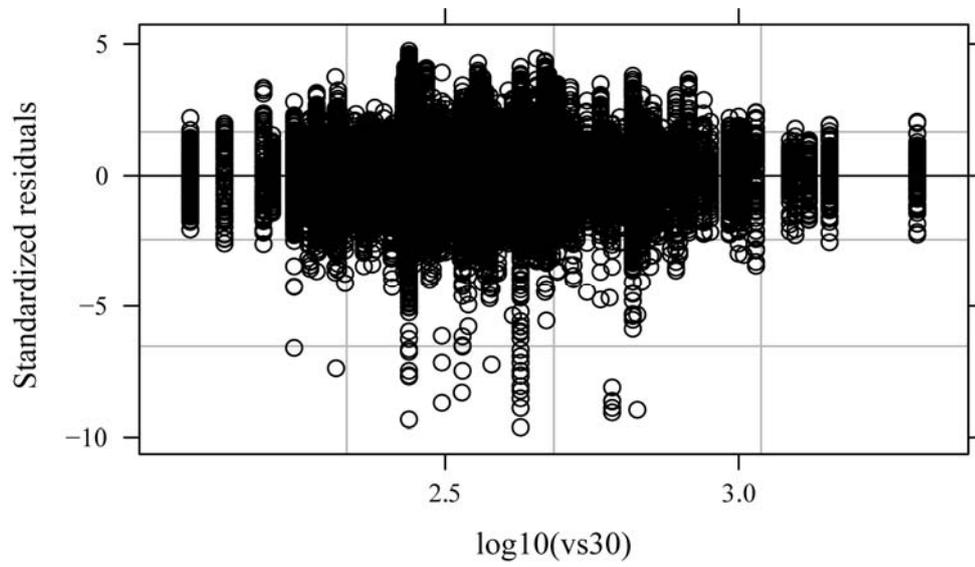


Figure 4.29. Standardized residuals of  $F_c$  regression versus  $\log_{10}(Vs30)$ .

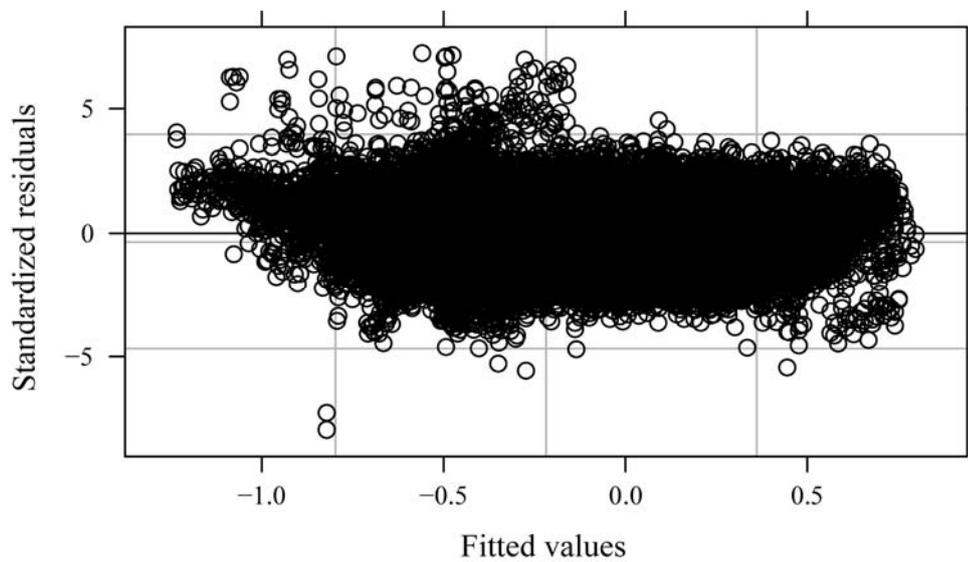


Figure 4.30. Standardized residuals of  $F_c$  regression versus fitted values.

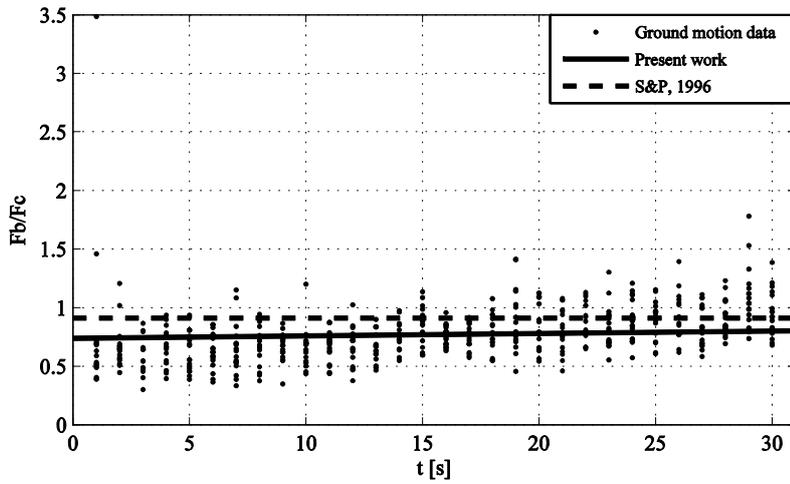


Figure 4.31. Dependence of  $F_b/F_c$  on time as predicted by different models: actual data (dots), prediction by the model in Eq. (4.25) and by the model proposed by Sabetta and Pugliese (1996). Note that the latter model neglected time dependence.

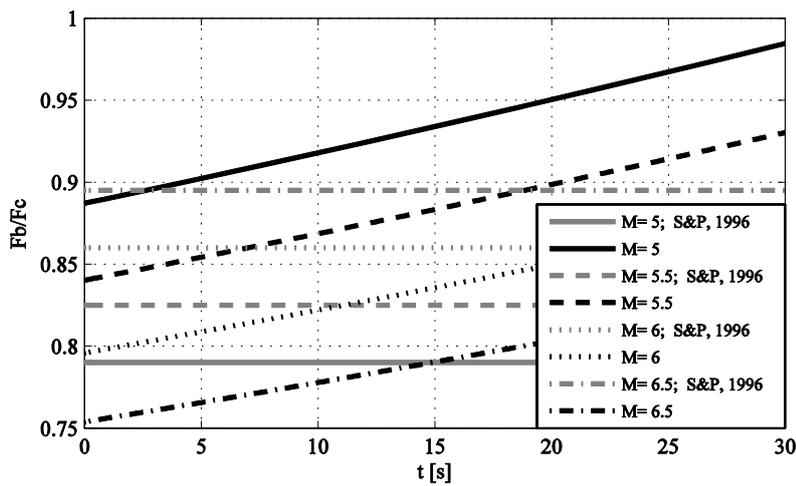


Figure 4.32. Predictions of  $F_b/F_c$  for different magnitudes given by the model in Eq. (4.25).

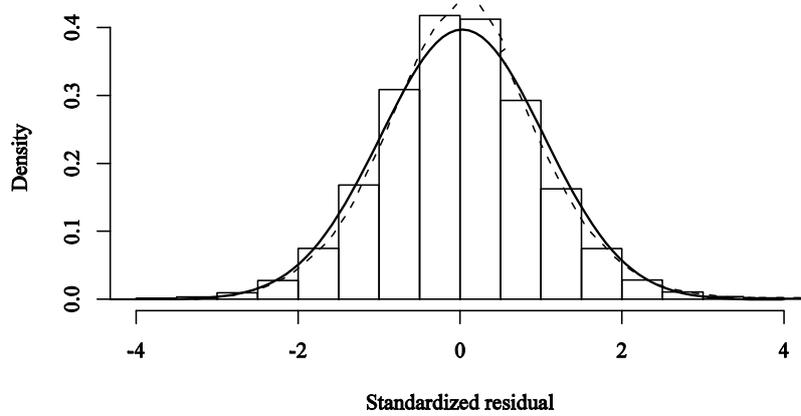


Figure 4.33. Histogram of standardized residuals of Fb/Fc regression. The solid line is a standard normal distribution and the dashed line is obtained using a kernel smoothing procedure.

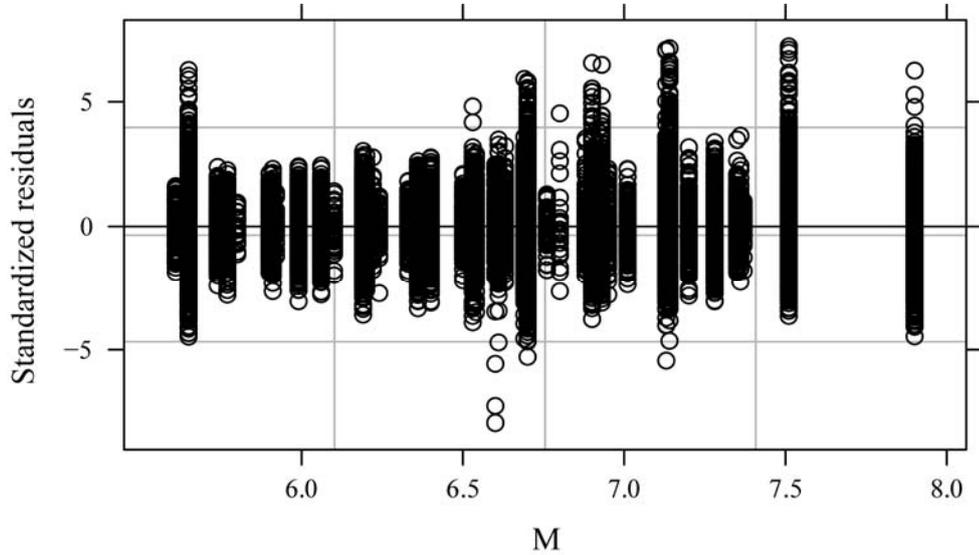


Figure 4.34. Standardized residuals of Fb/Fc regression versus magnitude.

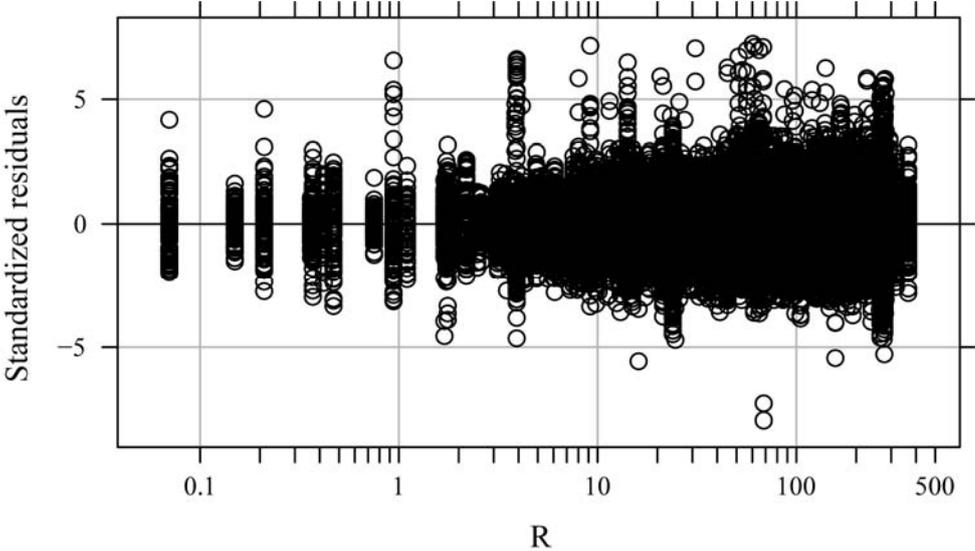


Figure 4.35. Standardized residuals of Fb/Fc regression versus distance (log scale).

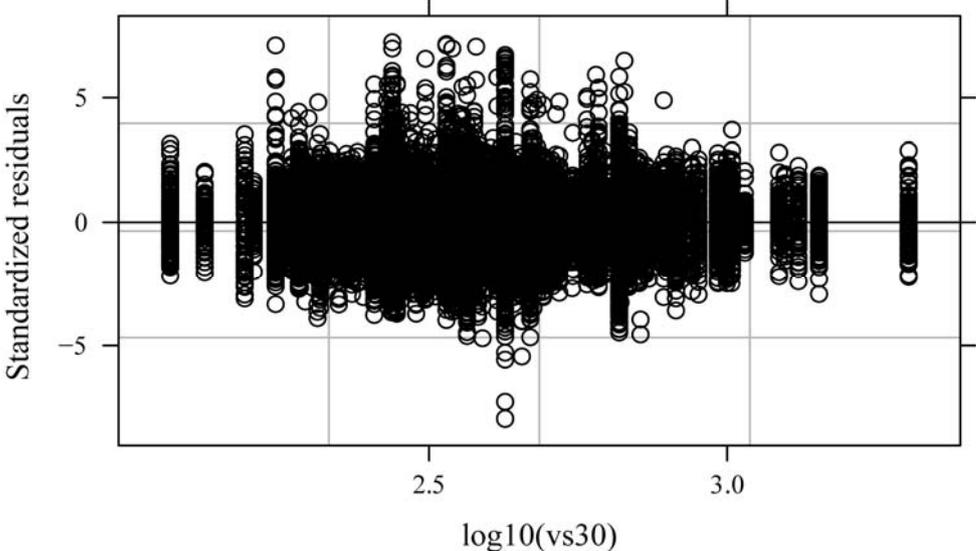


Figure 4.36. Standardized residuals of Fb/Fc regression versus log<sub>10</sub>(Vs30).

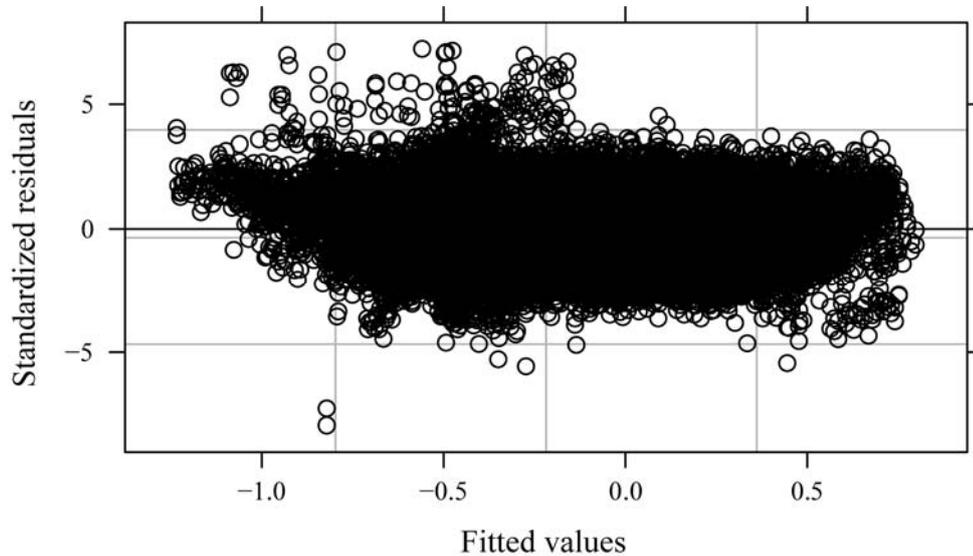


Figure 4.37. Standardized residuals of  $Fb/Fc$  regression versus fitted values.

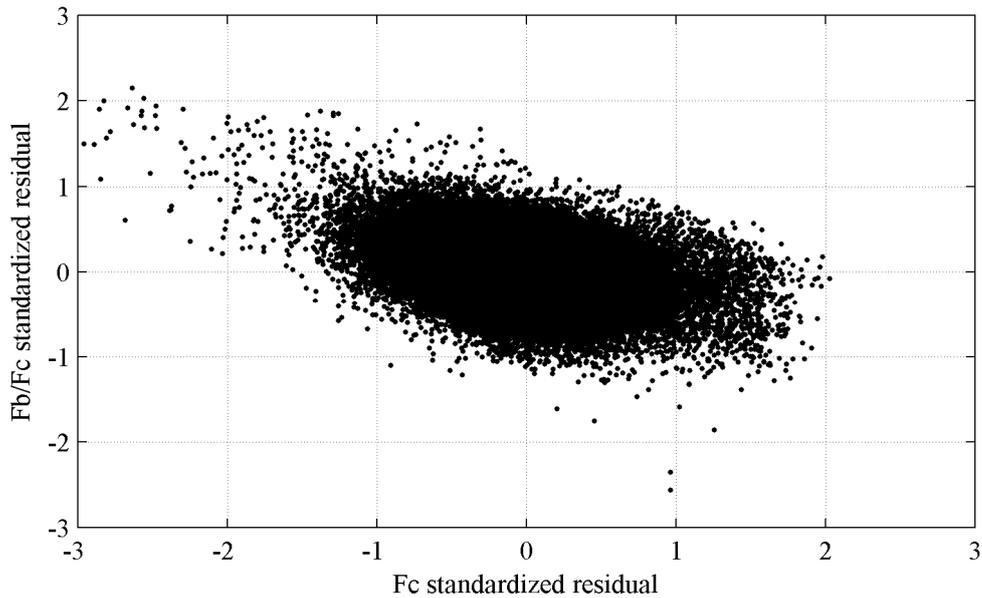
#### 4.4.6 Generation of accelerograms

Using the predictive equations developed in the previous sections it is possible to generate artificial accelerograms once a scenario is defined in terms of magnitude, distance and soil conditions.

One of the objectives of the present work was developing artificial accelerograms suitable for estimating not only the mean value of structural response (e.g. drift) but also its variance. In fact many procedures for generating artificial accelerograms have been developed in order to obtain estimates of mean structural response and therefore produce ground-motions with an unnatural low variability.

The ground-motion variability plays a very important role in seismic hazard analysis. In many cases the ground-motion variability is the most important component of a ground-motion model. For example, the specification of ground motion distribution values is central to PSHA (Cornell, 1968; Bommer and Abrahamson, 2006; Bommer and Abrahamson, 2007a; Bommer and Abrahamson, 2007b; Strasser et al., 2009).

Often, the of ground-motion variability is considered as a measure of the lack of fit of a particular predictive equation. However, in most cases it is better to think of a predictive equation as providing an estimate of the distribution of ground motions, given a set of predictor variables such as magnitude and distance. From this perspective, the real misfit of the model is related to how well the model's distribution represents the true distribution of ground motions rather than how large are the variance components.



**Figure 4.38. Standardized residuals of  $Fb/Fc$  regression versus standardized residuals of  $Fc$  regression.**

With this perspective, using the predictive equations developed in the previous sections, it is possible to obtain estimates of the distributions of  $IA$ ,  $DV$ ,  $Fc$  and  $Fb/Fc$  as a function of magnitude, distance and soil stiffness. According to the present approach, the logarithm of each one of the four parameters defining the physical spectrum is a normal variable. Its mean value can be calculated using Eqs. (4.21), (4.22), (4.24), (4.25) with estimated regression parameter values provided in Tables 4.1, 4.2, 4.3, 4.4. Those tables give the estimated standard deviations of the models as well.

Therefore, when an artificial accelerogram is generated values  $IA$ ,  $DV$ ,  $Fc$  and  $Fb/Fc$  can be randomly sampled according to their distributions and used in Eq. (4.8) and Eq. (4.11) to obtain an approximated spectrogram. Since four random variables are involved in the simulation procedure, before proceeding further it is mandatory to check for the presence of correlation among them. This can be done plotting the residuals one against another. The only correlation found was the one between the residuals of the regression of  $Fc$  and of  $Fb/Fc$ . As Figure 4.38 shows there is a negative correlation the value of which is equal to -0.396.

As an example average acceleration response spectra of 8 accelerograms generated using the proposed procedure are compared in Figure 4.39 with acceleration response spectra predicted by the attenuation relationships developed in Section 3.4.1. Figure 4.39 shows a good match between the two kinds of spectra. It should be noted that the artificial accelerograms are not scaled in order to enhance spectral compatibility as done in the procedure proposed by Pousse et al. (Pousse, 2005; Pousse et al., 2006).

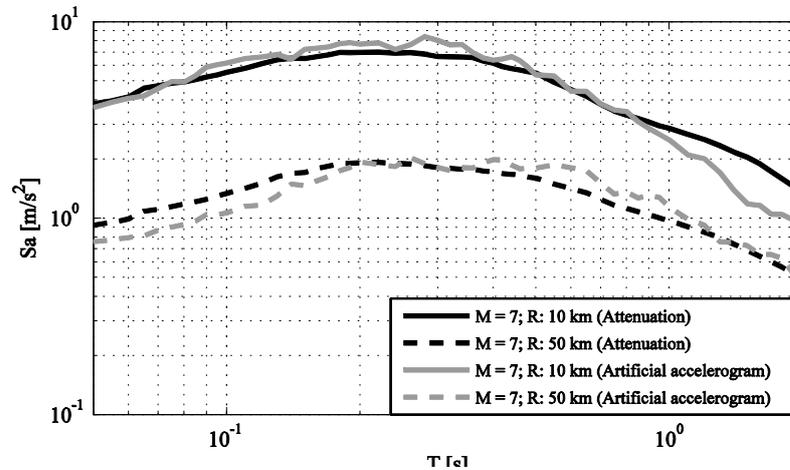


Figure 4.39. Comparisons among acceleration response spectra predicted by attenuation relationships developed in Chapter 3.4.1 and average spectra of eight accelerograms generated according to the modified Sabetta and Pugliese (1996) method.

## 4.5 Artificial accelerograms as input for nonlinear structural analysis

The comparisons given in Figure 4.39 suggest that the acceleration response spectra of the accelerograms generated according to the present procedure are in good agreement with the spectra given by ground-motion predictive equations. This does not give any information on the suitability of the accelerograms as input for non-linear dynamic analysis. It should be noted that the common concern that artificial ground-motions have unreasonably high energy content (Naeim and Lew, 1995) does not hold for the ground motions generated according to this procedure, since the energy content of the accelerograms (measured by the Arias intensity) is defined in a physically consistent way.

In order to investigate this important aspect a series of nonlinear dynamic analyses were performed on three case study structures, in order to compare the results given by artificial accelerograms (generated for different scenarios in terms of magnitude, distance and soil conditions) with those given by recorded accelerograms and considered as reference. Results given by the two different kinds of accelerograms were compared in terms of maximum interstorey-drift-ratio demand because this is one of the most widely used measures of structural demand.

In order to obtain the reference distributions of interstorey drift, the same procedure adopted in Chapter 3 was used. These distributions are obtained by fitting attenuation

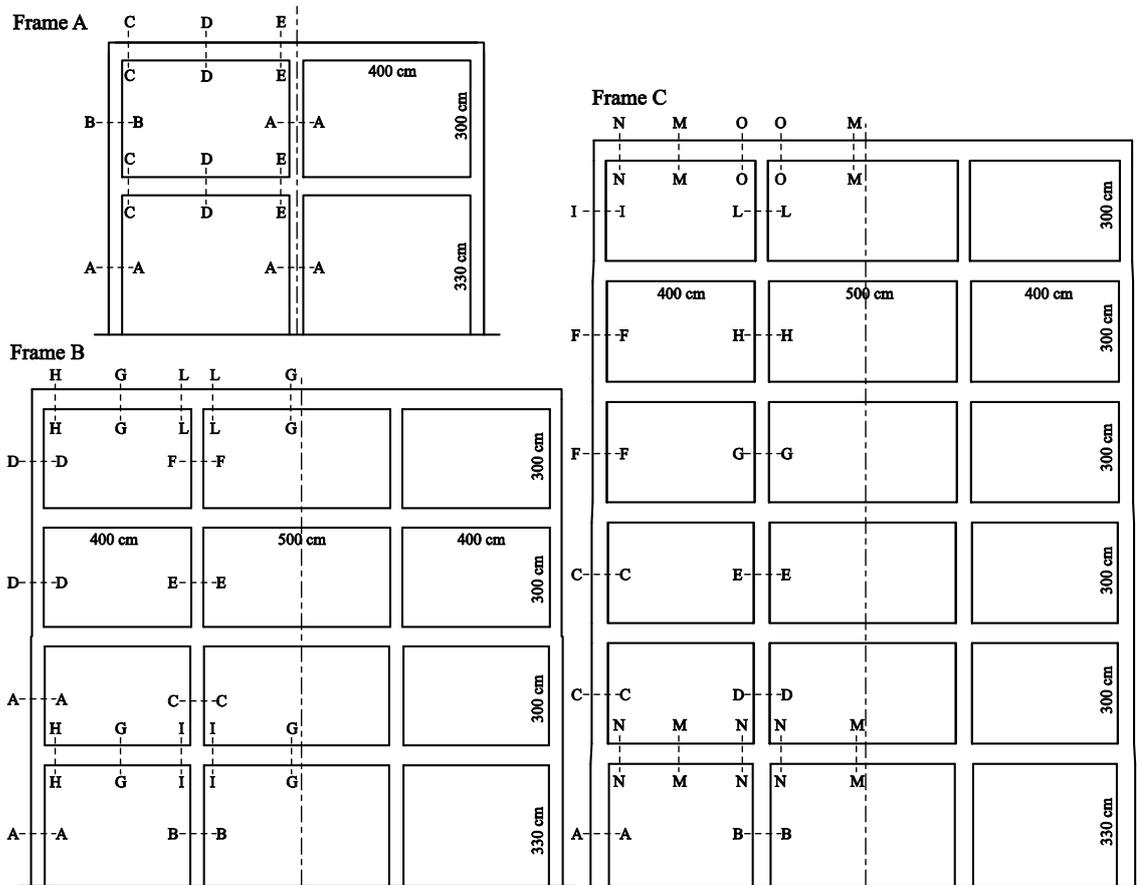


Figure 4.40. Frame structures considered in the present study. Natural periods in the cracked range: frame A = 0.46 s; frame B = 0.69 s; frames C = 0.84 s. Cross section details are given in Table 4.5.

relationships on the maximum drifts observed during nonlinear dynamic analyses performed using all the recorded accelerograms in the considered database.

Predictive equations (one for the maximum drift of each floor) used have the form:

$$\log_{10}(Y) = c_1 + c_2 M_w + c_3 M_w^2 + (c_4 + c_5 M_w) \log_{10} \left( \sqrt{R_{JB}^2 + c_6^2} \right) + c_7 \log_{10}(V_{S30}) \quad (4.26)$$

where  $\log_{10}(Y)$  is the base ten logarithm of the parameter to be regressed;  $M_w$  is moment magnitude;  $R_{JB}$  is the closest distance to the surface projection of the fault rupture, as proposed by Joiner and Boore (1981);  $V_{S30}$  is the shear-wave-velocity over upper 30 m. The total standard deviation,  $\sigma_T$ , of each empirical model may be decomposed into three independent components: the inter-event standard deviation,  $\sigma_E$ ; the intra-event standard deviation,  $\sigma_A$ ; and the inter-component standard deviation,  $\sigma_C$ , which must be considered when both horizontal components of a ground-motion recording are used in

**Table 4.5. Cross section details for the RS frames considered in the present study. All lengths are in cm.**

Section	Frame A				Frame B				Frame C			
	b	h	As	As'	b	h	As	As'	b	h	As	As'
A-A	30	30	4f16	4f16	35	35	3f20	3f20	40	45	2f24	2f24
B-B	30	30	4f20	4f20	35	35	4f24	4f24	40	45	5f24	5f24
C-C	30	50	2f14	2f14+2f18	35	35	3f24	3f24	40	40	2f24	2f24
D-D	30	50	4f14	2f14	30	30	4f20	4f20	40	40	5f24	5f24
E-E	30	50	2f14	2f14+4f18	30	30	3f24	3f24	40	40	4f24	4f24
F-F					30	30	3f20	3f20	35	35	3f24	3f24
G-G					30	50	2f14	2f14	35	35	5f24	5f24
H-H					30	50	2f14	2f14+2f20	35	35	3f24	3f24
I-I					30	50	2f14	2f14+3f20	30	30	3f24	3f24
L-L					30	50	2f14	2f14+2f20	30	30	2f24	2f24
M-M									30	50	2f14	2f14
N-N									30	50	2f14	2f14+4f20
O-O									30	50	2f14	2f14+3f18

the regression analysis, as is done in this study (Boore et al., 1997; Douglas, 2003; Boore and Bommer, 2005; Baker and Cornell, 2006c). As done in Section 3.4.2 these predictive equation were used to define reference drift distributions as a function of magnitude, distance and soil conditions. Details on the regressions and goodness of fit tests have been omitted.

#### 4.5.1 Case study structures

Three reinforced concrete frames were used as case study structures. In the following they will be referred as Building A, Building B and Building C. Details on the geometry and on the reinforcing bars are given in Figure 4.40 and Table 4.5.

The time-domain analyses take into account both geometric nonlinearity and material inelasticity. Structural members are modelled using the force-based elements with plastic hinges proposed by Scott and Fenves (2006). These elements were chosen in order to avoid the localization problems that elements with distributed inelasticity can experience when sections have strain-softening behaviour (Spacone et al., 1996a; Spacone et al., 1996b; Spacone et al., 1996c; Coleman and Spacone, 2001; Scott and Fenves, 2006). Both displacement-based and force-based finite elements are affected by this issue. The displacement-based approach causes localization of response over a single finite element (De Borst et al., 1994; Bazant and Planas, 1998). The length of the element undergoing softening response controls the structural response, thus leading to non-objectivity because the structural response depends on the choice of the characteristic length in the finite element discretization (Scott and Fenves, 2006). On

the other hand, with force-based beam-column elements deformations localize at a single integration point rather than across an entire element, making the characteristic length equal to the integration weight associated with the section undergoing strain softening. This leads to a loss of objectivity because the response changes as a function of the number of element integration points rather than as a function of the element length (Scott and Fenves, 2006). In the present study, this problem has been addressed by using the integration method proposed by Scott and Fenves (2006), which confines non-linear constitutive behaviour to plastic hinge regions of a specified length maintaining objectivity. Non-linear behaviour of sections in the plastic hinge regions is modelled using fibre discretization. The confined concrete constitutive behaviour is modelled using the model proposed by Mander (Mander et al., 1989), the unconfined concrete constitutive behaviour is modelled using the model developed by Saenz (Ceb-Fip, 1993) and the steel constitutive behaviour is modelled using the Menegotto and Pinto's model (Menegotto and Pinto, 1973).

All nonlinear time-history analyses were run using the software OpenSees (McKenna et al., 2000).

#### ***4.5.2 Definition of the length of plastic hinges***

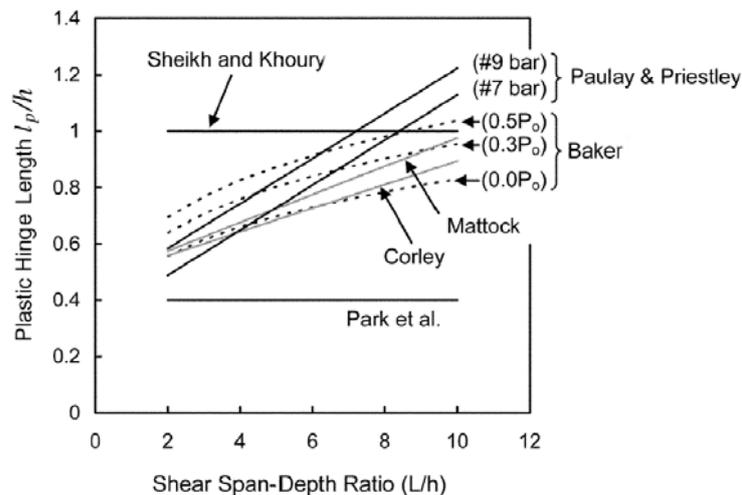
The adopted finite element beam-column model requires to define the length of plastic hinges (Scott et al., 2008). This parameter is very important, because it has a strong influence on analysis results (Inel and Ozmen, 2006). Different methods have been proposed in the literature in order to evaluate this length.

Indeed, the most accurate way of defining this parameter is using results of experimental tests of RC column under cycling loading and trying to reproduce the experimental force-displacement behaviour with the numerical model. According to this approach the plastic hinge length is defined together with the hysteretic behaviour of the hinge (Ibarra and Krawinkler, 2005; Ibarra et al., 2005). Of course, this approach is not always practicable since results of experimental tests under cycling-loading for columns similar to those of the columns of the structure that is to be modelled may not be available.

A simpler approach, which is widely used in the literature (Scott et al., 2008), is to determine the length of plastic hinge through empirical equations calibrated using data from experimental test. Many researchers have proposed equations with this aim (e.g. Baker, 1956; Mattock, 1964; Corley, 1966; Mattock, 1967; Priestley and Park, 1987; Paulay and Priestley, 1992; Sheikh and Houry, 1993; Mendis, 2001; Panagiotakos and Fardis, 2001; Priestley et al., 2007; Bae and Bayrak, 2008). The length of a plastic hinge depends on many factors: *i*) level of axial load; *ii*) moment gradient; *iii*) level of shear

stress in the plastic hinge region; *iv*) mechanical properties of longitudinal and transverse reinforcement; *v*) concrete strength; and *vi*) level of confinement and its effectiveness in the potential hinge region (Bae and Bayrak, 2008). Though, some of the models proposed in the literature do not consider the dependence on some of the aforementioned parameters (e.g. axial load). A very comprehensive review of the available models can be found in Bae and Bayrak (Bae and Bayrak, 2008).

The dependency of the ratios  $l_p/h$ , where  $l_p$  is plastic hinge length and  $h$  is cross section height, on shear span-depth ratio,  $L/h$ , predicted by different models available in the literature is given in Figure 4.41. This Figure shows that there is a wide scatter in the length predicted by the different models. Furthermore, using empirically relationships determined directly from experimental data to define the length of plastic hinges in finite element models can be questionable. In fact, the plastic hinge lengths predicted by empirical models are usually defined on the basis of physical considerations. Considerations which may not be consistent with the definition of plastic hinge length used in the numerical model. For this reason this approach was not used in the present work.



**Figure 4.41. Plastic hinge lengths versus shear-span ratio, as predicted by different empirical models available in the literature (after Bae and Bayrak, 2008).**

Another possible way to define plastic hinge length is that described by Scott and Fenves (2006) who developed a procedure that Coleman and Spacone (2001) had proposed to address localization issues in force-based beam-column elements. According to this procedure the plastic hinge length is determined imposing that the numerical model maintains a constant energy release after strain-softening initiates.

Even if this definition has a theoretical basis, it is not possible to relate plastic hinge length with experimental data.

For the aforementioned reasons and for lack of experimental data, in the present work a different method has been used in order to define the plastic hinges length. According to the adopted procedure the plastic hinge length is calculated in such a way that the finite element model gives, for a cantilever column, an ultimate displacement equal to that predicted by the Fardis's model (Panagiotakos and Fardis, 2001). This model predicts the ultimate displacement of concrete columns as a function of some geometric parameters, of materials characteristic and, confinement ration and axial load level. Fardis's model, which is also included in Eurocode 8, was chosen because it has been fitted using a very comprehensive database of experimental data (Panagiotakos and Fardis, 2001). According to the proposed procedure the plastic hinge length is determined, using an iterative algorithm, such a way that the ultimate displacement, i.e. the displacement corresponding to the attainment of the ultimate deformation in one of the cross sections, given by the numerical model corresponds to that predicted by the Fardis's model. Ultimate deformations for steel and concrete are defined according to the criteria given in Priestley et al. (2007).

Figure 4.42 shows, for a 30x30 cm<sup>2</sup> cross section with a reinforcing-steel-ratio equal to 0.03, the dependency of the plastic hinge length, obtained according to the adopted procedure, on axial load ratio,  $v = N/(f_c \cdot b \cdot h)$ , and on concrete compressive strength,  $f_c$ . The change in the slope of the curves at  $v = 0.05$  is a consequence of the change of failure type, in fact for  $v \leq 0.05$  the ultimate strain is attained by the reinforcing steel and for  $v > 0.05$  by the confined concrete. Increasing the reinforcing steel ratio may eliminate the steel-side rupture. Figure 4.43 shows dependency of plastic hinge length on axial load ratio, for  $f_c = 30$  MPa, reinforcing-steel-ratio equal to 0.05, and different cross section sizes.

This procedure has been adopted to define the length of plastic hinges for the beam-column finite-elements used to model the case study structures. It is worth noting that  $l_p$  depends on the axial load level. Axial loads produced by gravity loads have been used in order to calculate plastic hinge lengths for the case study structures.

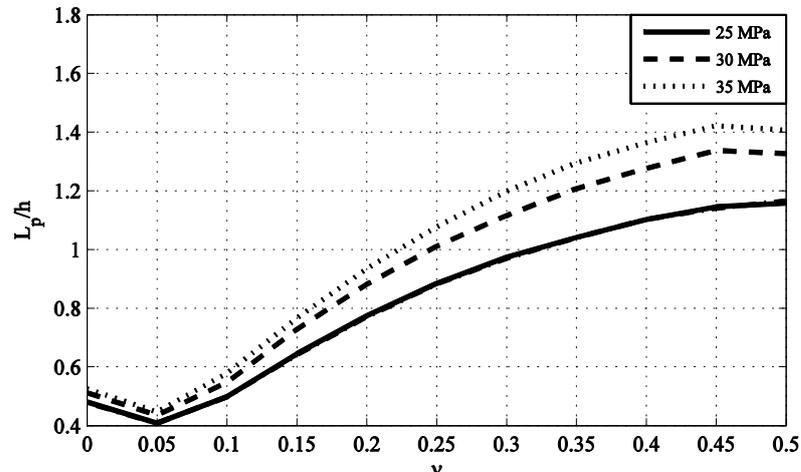


Figure 4.42. Plastic hinge length versus axial load ratio for three different concrete strengths, reinforcing-steel-ratio equal to  $\rho = 0.03$ , and  $30 \times 30 \text{ cm}^2$  cross section.

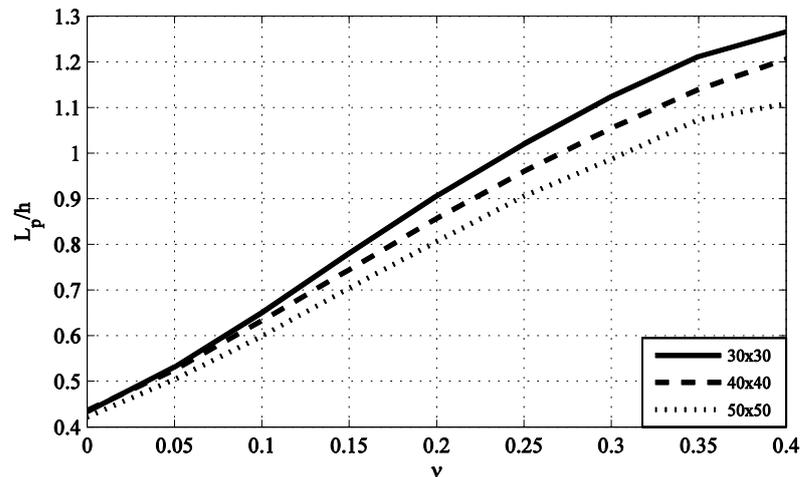


Figure 4.43. Plastic hinge length versus axial load ratio for three different cross sections, reinforcing-steel-ratio  $\rho = 0.05$ , and concrete strength  $f_c = 30 \text{ MPa}$ .

### 4.5.3 Overview of the analyses

In this section, the drifts obtained performing time-history analyses with unscaled artificial accelerograms generated according to different magnitude, distance and soil stiffness scenarios will be discussed for the case study building. The following scenarios were considered: *i*)  $M = 6.5$ ,  $R = 20 \text{ km}$ ,  $V_{S30} = 300 \text{ m/s}$ ; *ii*)  $M = 6.5$ ,  $R = 100 \text{ km}$ ,  $V_{S30} = 300$ ; *iii*)  $M = 7.0$ ,  $R = 10 \text{ km}$ ,  $V_{S30} = 300$ ; *iv*)  $M = 7.0$ ,  $R = 20 \text{ km}$ ,  $V_{S30} = 300$  and *v*)  $M = 7.0$ ,  $R = 50 \text{ km}$ ,  $V_{S30} = 300 \text{ m/s}$ . Furthermore, two different sets of accelerograms were generated for each scenario, one set (A) considering the variability of  $IA$  and  $DV$  only and the other set (B) considering the variability of all the parameters ( $IA$ ,  $DV$ ,  $F_c$  and

*Fb*) defining the approximated spectrogram. 100 accelerograms were generated for each set.

#### ***4.5.4 Results of time-history analyses for the frame structure A***

Figures 4.44-4.48 show the results, in terms of distribution of maximum drift at different storeys, obtained for the sets A and B and the five scenarios considered. The reference drift distribution was obtained, as described in Section 4.5, by fitting attenuation relationships on the results of time history analyses performed using all the recorded accelerograms in the considered database. These reference values are depicted by gray lines; the solid line indicates the reference median response, the dashed lines indicate  $\pm\sigma$  intervals, and the dash-dot lines indicate  $\pm 2\sigma$  intervals. As done in Chapter 3 a lognormal distribution was assumed for maximum drift. This choice has been verified performing different statistical tests, as done in Section 3.4.2, which are not reported here.

Artificial ground motions of both sets give median values of drift in good agreement with the reference values, but they give a variability of drift response which is less accurate. This loss of accuracy in predicting response variability increases as the non-linear demand on the structure increases. In fact, comparing Figure 4.45 and Figure 4.46 one can see that for the  $M = 6.5$ ,  $R = 100$  km scenario, which brings the structure to very low inelastic excursions, the drift predictions are very good in terms of both mean value and standard deviation. In particular, for this scenario, the set B, generated taking into account the variability of all the parameters defining the spectrogram gives very good predictions. On the other hand for the  $M = 7.0$   $R = 10$  km scenario, which brings the structure to higher inelastic excursions there is a loss of accuracy especially as far as drift variability is concerned.

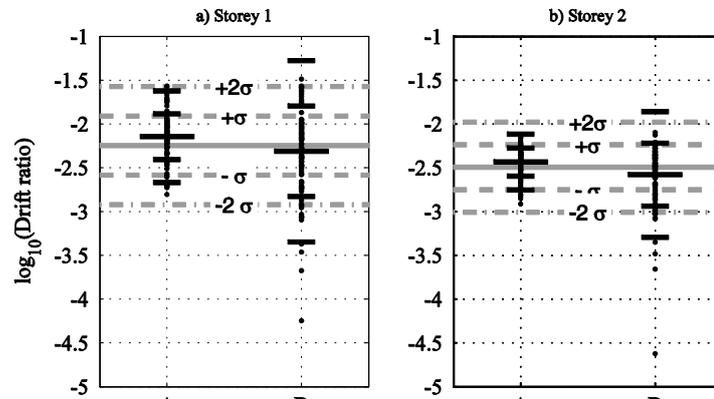


Figure 4.44. Frame A:  $M = 6.5$ ,  $R = 20$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

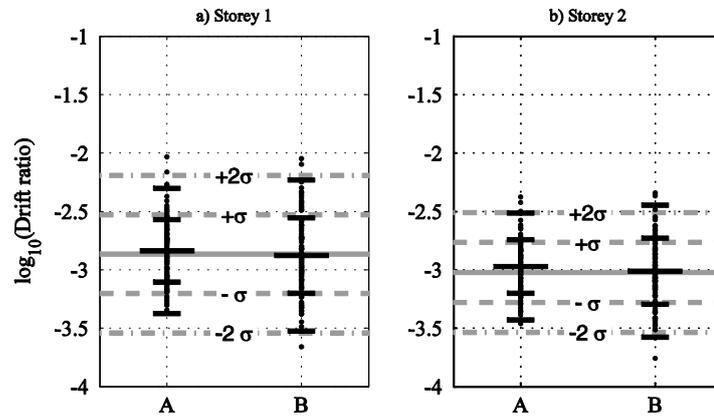


Figure 4.45. Frame A:  $M = 6.5$ ,  $R = 100$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

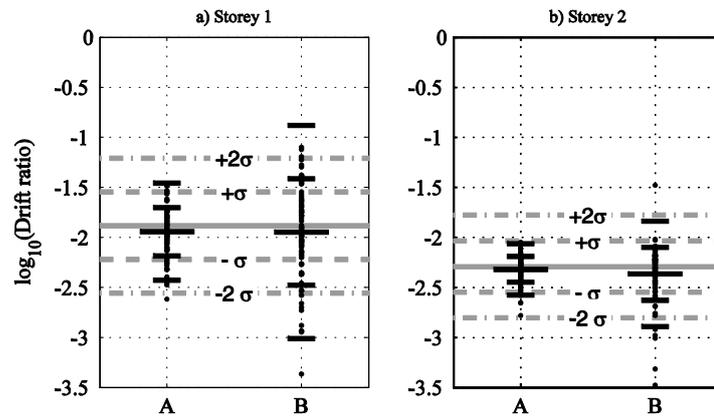


Figure 4.46. Frame A:  $M = 7.0$ ,  $R = 10$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

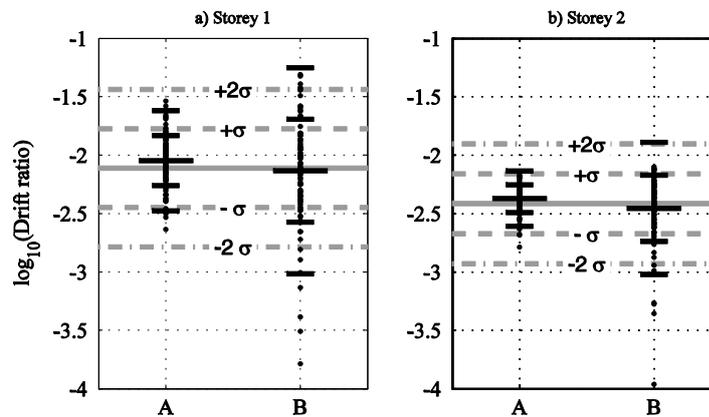


Figure 4.47 Frame A:  $M = 7.0$ ,  $R = 20$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

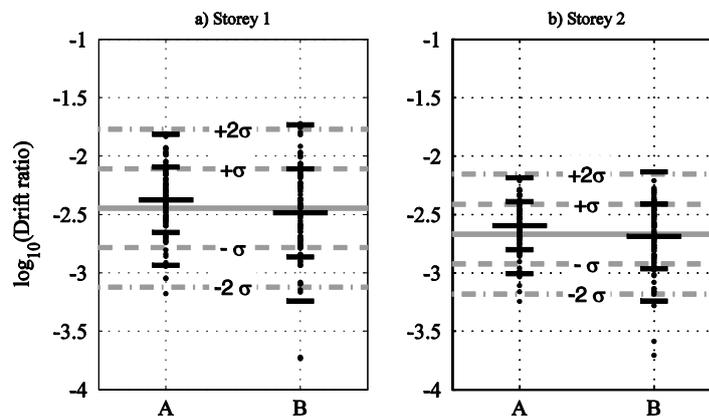


Figure 4.48 Frame A:  $M = 7.0$ ,  $R = 50$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

### 4.5.5 Results of time-history analyses for the frame structure B

Figures 4.49-4.53 give drift values obtained for the RC frame B. Similarly to what was obtained for the frame A, the artificial ground motions give predictions of mean drift values in good agreement with the reference values given by the reference attenuation. As previously observed, predicting response variability is more problematic especially when the structure has strong inelastic excursions. The variability predicted by artificial ground motions of set A is usually lower than the reference one while variability predicted by set B is usually higher. Furthermore the goodness of prediction, in terms of variability, decreases as the inelastic demand on the structure increases (compare Figure 4.50 with Figure 4.51).

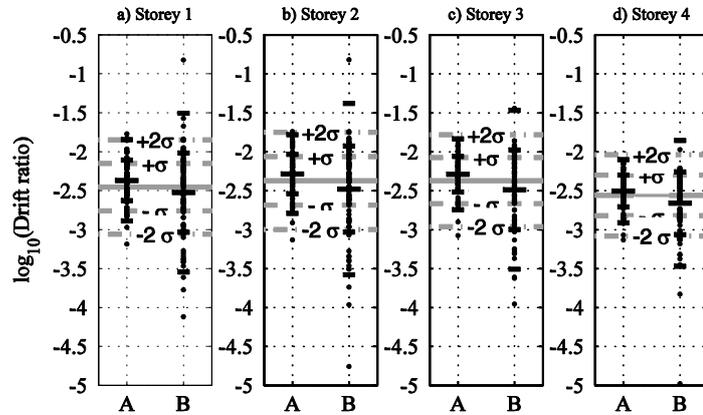


Figure 4.49 Frame B:  $M = 6.5$ ,  $R = 20$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

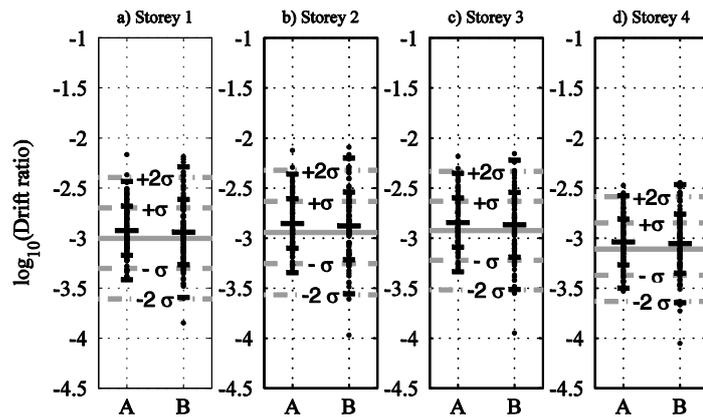


Figure 4.50 Frame A:  $M = 6.5$ ,  $R = 100$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

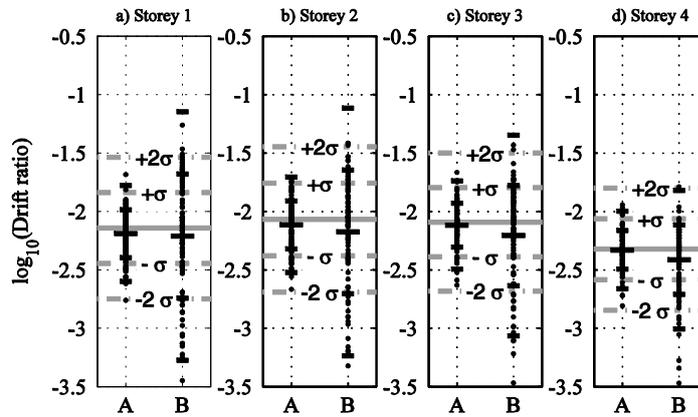


Figure 4.51 Frame B:  $M = 7.0$ ,  $R = 10$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

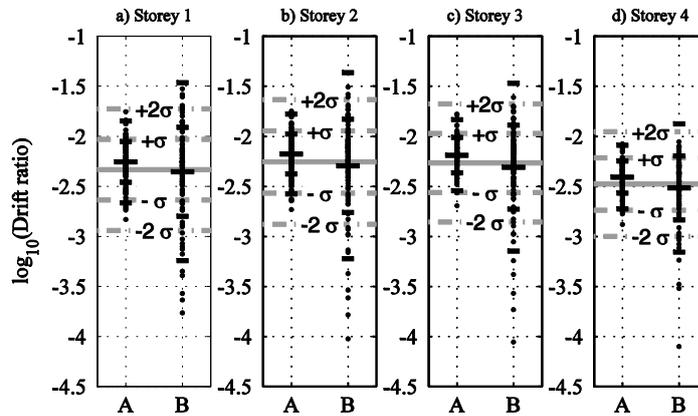


Figure 4.52 Frame B:  $M = 7.0$ ,  $R = 20$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

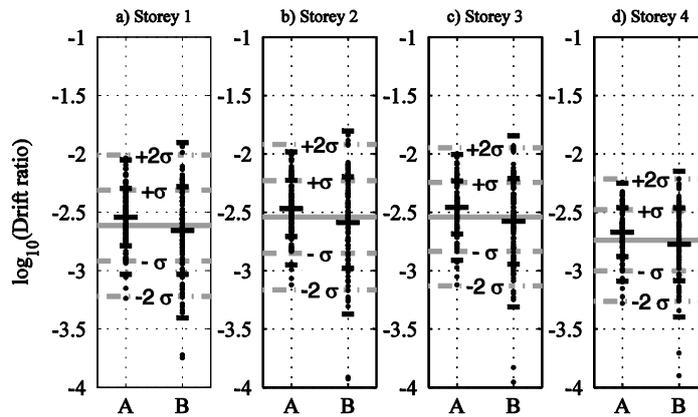


Figure 4.53. Frame B:  $M = 7.0$ ,  $R = 50$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

### 4.5.6 Results of time-history analyses for the frame structure C

Finally, results in terms of maximum drift demand for structure C are given in Figures 4.54-4.58. These results confirm the conclusions that were drawn for buildings A and B. That is, the artificial ground motions generated following the procedure proposed give good results in terms of mean drift response but give less accurate results in terms of drift variability especially when high inelastic excursions occur.

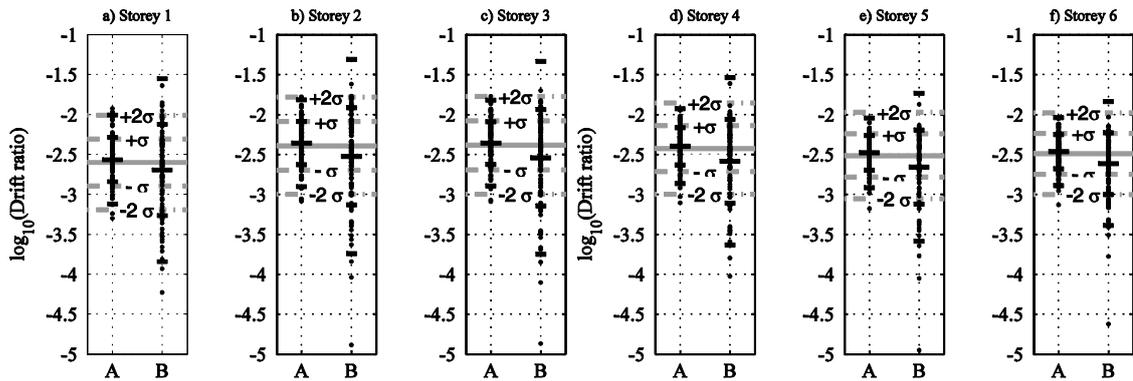


Figure 4.54 Frame C:  $M = 6.5$ ,  $R = 20$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

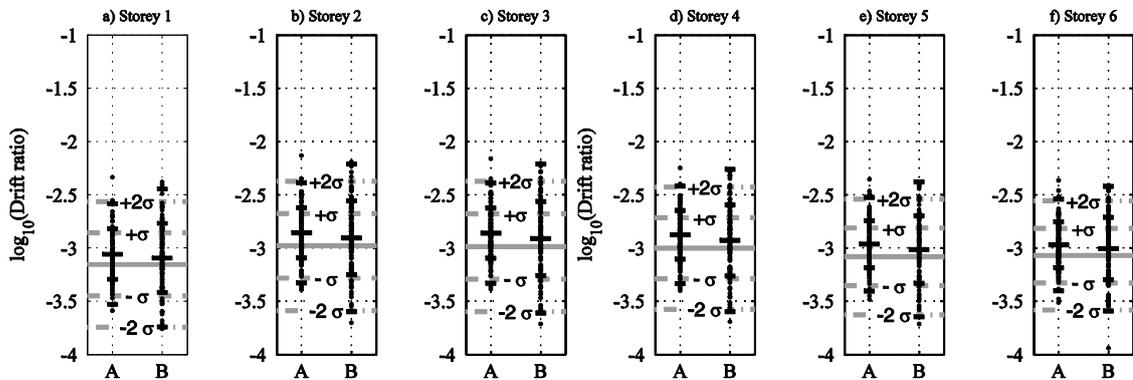


Figure 4.55 Frame C:  $M = 6.5$ ,  $R = 100$  km,  $V_{s30} = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

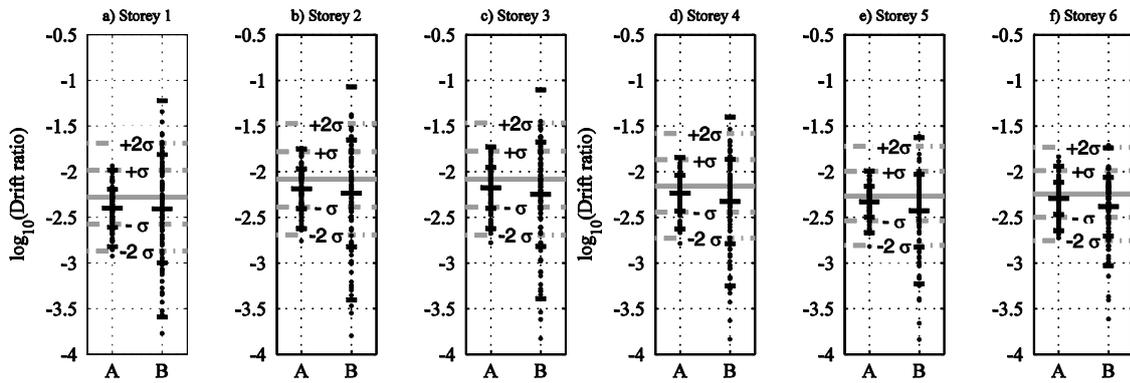


Figure 4.56 Frame C:  $M = 7.0$ ,  $R = 10$  km,  $Vs30 = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

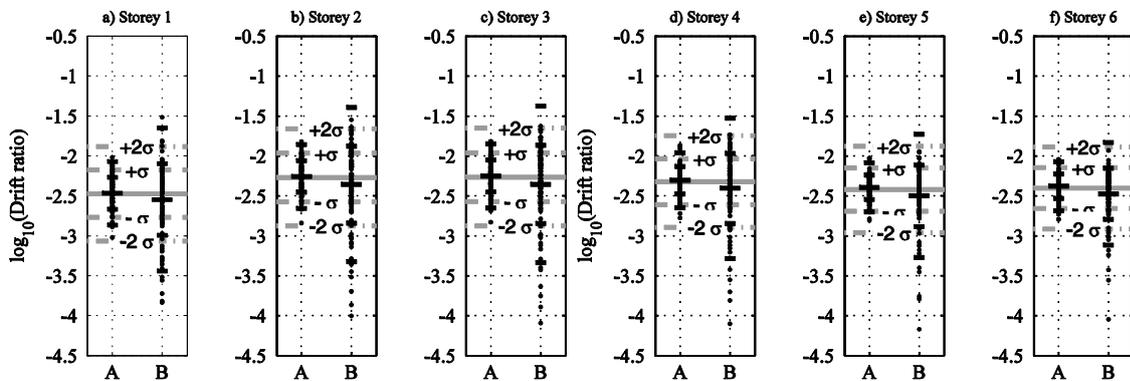


Figure 4.57 Frame C:  $M = 7.0$ ,  $R = 20$  km,  $Vs30 = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

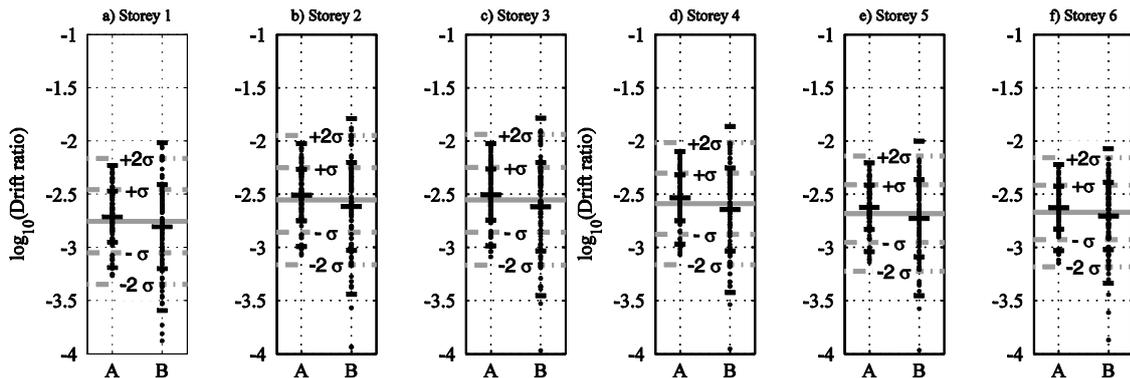


Figure 4.58. Frame C:  $M = 7.0$ ,  $R = 50$  km,  $Vs30 = 300$  m/s scenario. Gray lines indicate mean,  $\pm 1$  and  $\pm 2$  standard deviation values of the reference drift distribution. Dots represent results of analyses, their mean,  $\pm 1$  and  $\pm 2$  standard deviation values are indicated with horizontal segments.

## 4.6 Conclusions

The procedure for generating artificial accelerograms originally proposed by Sabetta and Pugliese (1996) has been revisited and enhanced by *i*) proposing new forms for the ground-motion prediction equations involved in the model, *ii*) using a bigger and more comprehensive database of accelerograms, *iii*) using a more consistent definition of magnitude, *iv*) proposing a new criterion for taking into account the directions of recording sensors and giving a better description of ground motion variability.

The attenuation relationships originally developed by Sabetta and Pugliese had very simple functional form and, most importantly, had a very simple representation of the variance structure of the data. In fact the total variance of the model was not partitioned into intra- and inter-event components. This representation of the variance is nowadays used in almost any attenuation relationship because it provides a better representation of the actual data correlation. The functional form adopted in the present work are consistent with those recently proposed by Stafford, 2006 when developing an attenuation relationship for Arias intensity in New Zealand.

Furthermore a bigger accelerogram database was used when performing regression analysis: Sabetta and Pugliese considered 190 accelerograms while in the present work 1504 accelerograms were considered. Though, it is worth noting that the original database contained only Italian earthquakes while in the present work a subset (See Section 4.4) of the NGA database was used which contains events recorded all around the world, even though it is particularly US focused. Recent studies (Stafford et al., 2008) performed on the NGA attenuation relationships suggest that these latter can be adopted for the Europe and Middle East as well.

In contrast to what was done by Sabetta and Pugliese (1996), in the present work, moment magnitude (Hanks and Kanamori, 1979) has been used to measure earthquake intensity. It is recognized that this measure is preferable to other magnitude measures because it does not saturate at high values.

Finally, in the present work, particular care has been taken to correctly represent the variability of ground motion. In particular two horizontal components of ground-shaking were used, defined in such a way that one corresponds to the direction with the maximum Arias intensity and the other to the direction with the minimum Arias intensity.

Accelerograms generated according to the proposed procedure, for different scenarios in terms of magnitude, distance and soil conditions, were used as input for non-linear dynamic analysis of three case study RC buildings. The so obtained results, in terms of

maximum-interstorey-drift-demand distribution, were compared with reference distributions obtained using the procedure described in Chapter 3. That is fitting attenuation relationships on the maximum drift values obtained from dynamic analyses performed using all the recorded accelerograms contained in the database used to develop attenuation relationships for Arias intensity, duration, central frequency and frequency bandwidth.

Maximum-drift-demand distributions obtained from artificial accelerograms are generally in good agreement in terms of mean value, while the agreement in terms of variability is lower. It is worth noting that both artificial and recorded accelerograms were not scaled. Most probably introducing a linear scaling procedure could enhance results, especially in terms of mean structural response.

## PART 2



# 5 Seismic reliability

## 5.1 Introduction

Perhaps no other discipline within engineering has to deal with as much uncertainty as the field of earthquake engineering (Der Kiureghian, 1996). The randomness in the occurrence of earthquakes in time and space, the vast uncertainty in predicting intensities of resulting ground motions, and the inability to accurately assess capacities of structures under cyclic loading all compel to make use of probabilistic methods in order to consistently account for the underlying uncertainties and make quantitative assessments of structural safety. While the need for probabilistic methods in earthquake engineering has long been recognized, their use in practice has been limited because of the analytical and computational difficulties that these methods impose on the design and evaluation process.

The final objective in earthquake engineering often is decision on design specifications. This usually requires risk analysis, involving the assessment of probabilities as well as costs associated with each design alternative and the corresponding consequences, should structural failure occur. A simpler, more practical framework is provided by probabilistic design codes, which implicitly account for the underlying uncertainties and optimize (in an approximate manner) the expected utility derived from each design.

## 5.2 Sources of uncertainty

Der Kiureghian (1996) provides a list of the dominant types of uncertainty in structural engineering: (a) inherent randomness, which arises from intrinsic variability in materials and in environmental effects, such as loads and support movements; (b) statistical uncertainty, which arises in the course of estimating parameters of probability distributions from observational samples of limited size; and (c) model uncertainty, which arises from the imperfection of mathematical models used to describe complex physical phenomena, such as models describing loads and capacities of soils or structures. Whereas the uncertainty due to inherent randomness is irreducible, statistical and model uncertainties can be reduced, the former by collection of additional samples and the latter by use of more refined models.

More recently Der Kiureghian and Ditlevsen (2009) further detailed the above list and expanded some of the points given above, suggesting the following list of sources of uncertainty: (a) Uncertainty inherent in the basic random variables, such as the uncertainty inherent in material property constants and load values, which can be directly measured; (b) uncertain model error resulting from selection of the form of the probabilistic sub-model used to describe the distribution of basic variables; (c) uncertain modelling errors resulting from selection of the physical sub-models used to describe the derived variables; (d) statistical uncertainty in the estimation of the parameters of the probabilistic sub-model; (e) statistical uncertainty in the estimation of the parameters of the physical sub-models; (f) uncertain errors involved in measuring of observations, including errors involved in indirect measurement; (g) uncertainty modelled by the random variables corresponding to the derived variables, which may include, in addition to all the above uncertainties, uncertain errors resulting from computational errors, numerical approximations or truncations.

Given these numerous sources of uncertainties, it is particularly important to threat and propagated them correctly. Many approaches have been proposed in the literature, the reader is referred to Ditlevsen (1982), Der Kiureghian (1989), Der Kiureghian (1996), Ditlevsen and Madssen (1996), Gardoni et al. (2002), and Der Kiureghian and Ditlevsen (2009).

While there can be many sources of uncertainty, it is convenient to categorize the character of uncertainties as either aleatory or epistemic. An aleatory uncertainty is one that is presumed to be intrinsic randomness of a phenomenon. On the other hand, an uncertainty is one that is presumed as being caused by lack of knowledge. As Der Kiureghian and Ditlevsen (2009) suggest, the above distinction may raise the philosophical question whether there is any aleatory uncertainty at all. This distinction does not make sense outside the model universe: from a semantic point of view, all uncertainties are the same as lack of knowledge. However, it is convenient within a probabilistic model to introduce the categorization of uncertainties into aleatory and epistemic. The advantage of separating the uncertainties is that this makes clear which uncertainties can be reduced and which uncertainties are less prone to reduction (at least in the near-term, i.e. before major advances occur in scientific knowledge). Therefore the categorization of uncertainties is a choice that must be made by the model builder, and generally depends on the context and application.

## 5.3 The reliability problem

### 5.3.1 *Limit states*

The concept of limit state related to a specific requirement is defined as a state of the structure including its loads at which the structure is just on the point of not satisfying the requirement. Usually the requirement will be interpreted and formulated within a mathematical model for the geometric and mechanical properties of the structure and for the actions on the structure. Let  $x_1, x_2, \dots, x_n$  be those variable that independently contribute to that part of the mathematical model that concerns geometry, strength properties and actions. To each choice of values corresponds a uniquely defined structure with uniquely defined loads. This structure with its loads is a pure mathematical object that does or does not satisfy a given limit state requirement.

Some reliability problems can be formulated in terms of a finite number of variables, but there are relevant reliability problems where the model formulation most conveniently requires use of an infinity of variables (functions). In particular this is the case when random temporal and spatial property variations of actions and resistances are relevant. A given limit-state requirement divides the domain of definition of the model in two sets, the safe set and the failure set, in which the requirement is satisfied and not satisfied, respectively. The boundary of the safe set is called the limit state.

Limit states can be of different categories. The principal ones are collapse limit states (ultimate limit states) and serviceability limit states. A collapse limit state usually represents a situation where the structure is just at the point of losing its integrity, that is, to pass into an irreversible state that may have a catastrophic nature and from which the structure only recovers by repair or reconstruction. A serviceability limit-state corresponds to the limit between an acceptable and a not acceptable state under normal use. Such a state is with respect to direct damage of the structure often reversible in the sense that the structure by unloading passes unchanged back to the safe set. However, passages of a serviceability limit-state can also cause permanent damages of the structure such as formation of cracks or other visible defects. Generally these damages will not raise a reliability problem of the collapse limit-state category provided the structure is subject to general running maintenance.

### 5.3.2 *Formulation of time invariant reliability problems*

The essence of the structural reliability problem is the probability integral

$$P_F = \int_F f(\mathbf{x}) d\mathbf{x} \quad (5.1)$$

where  $P_F$  denotes the failure probability,  $f(\mathbf{x})$  denotes the PDF of a vector of random variables  $\mathbf{x}$  that represent time-invariant uncertain quantities influencing the state of the structure under consideration and  $F$  denotes a subset of the outcome space where failure occurs. By failure, usually the exceedance of a prescribed serviceability or safety limit is implied. For mathematical analysis, it is necessary to describe the failure domain  $F$  in an analytical form. Usually this is done in terms of a performance function, i.e.  $F = \{\mathbf{x} : g(\mathbf{x}) \leq 0\}$ , where  $g(\mathbf{x})$  is known as the limit-state function. The boundary of  $F$  is defined by  $g(\mathbf{x}) = 0$  and is known as the limit-state surface. Obviously the safe set is defined by  $g(\mathbf{x}) > 0$ .

A measure of interest, related to Eq. (5.1), is the generalized “reliability index”, defined by

$$\beta = \Phi^{-1}(1 - P_F) \quad (5.2)$$

where  $\Phi(\cdot)$  denotes the inverse of the standard normal cumulative probability. When the failure domain is linear and the variables are standard normal,  $\beta$  can be interpreted as the minimum distance from the origin (the mean point) to the failure domain. This interpretation has been found useful in second-moment reliability analysis and in developing probabilistic codes.

### 5.3.3 Statistical and model uncertainties

The formulation presented in Eq. (5.1) only accounts for the inherent randomness represented by the random variables  $\mathbf{x}$ . To explicitly account for statistical and model uncertainties, it is convenient to write the PDF of  $\mathbf{x}$  as  $f(\mathbf{x}|\boldsymbol{\theta})$  and the limit state functions as  $g(\mathbf{x}, \boldsymbol{\theta})$ , where  $\boldsymbol{\theta}$  is a vector collecting all the uncertain parameters present in the distribution function and the mathematical models describing the limit-states. The probability of failure and the reliability index are now dependent on  $\boldsymbol{\theta}$  and hence are uncertain. There are different positions in the literature as to how these uncertainties should be treated. One view, supported by Madsen et al. (1986) and Ditlevsen (1982), is to not differentiate between the uncertainties in  $\boldsymbol{\theta}$  from those in  $\mathbf{x}$ . The result is the expected probability of failure obtained from the total probability formulation

$$E(P_F(\boldsymbol{\theta})) = \int_F f(\mathbf{x}|\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\mathbf{x}d\boldsymbol{\theta} \quad (5.3)$$

where  $f(\boldsymbol{\theta})$  is the distribution of  $\boldsymbol{\theta}$ , which may be obtained, for example, from regression or from Bayesian updating. This integral has the same form as in Eq. (5.1), provided that  $\mathbf{x}$  and  $\boldsymbol{\theta}$  are combined to form a single vector. Hence methods used for the solution of Eq. (5.1) can also be used to evaluate Eq. (5.3).

Another approach for the analysis of uncertainties, supported by Der Kiureghian (Der Kiureghian, 1989; Der Kiureghian, 1996; Der Kiureghian and Ditlevsen, 2009) recognizes the fundamental difference between the inherent randomness represented by  $\mathbf{x}$ , which are irreducible, and the statistical and model uncertainties represented by  $\boldsymbol{\theta}$ , which are reducible. In this case, failure probability,  $P_F$ , and reliability index,  $\beta$ , are treated as random variables with probability distributions that reflect the effect of statistical and model uncertainties on their estimated values. The cumulative distribution function of  $\beta$ , for example is given by

$$P(\beta \leq \beta_0) = \int_{\beta(\boldsymbol{\theta}) - \beta_0 \geq 0} f(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (5.4)$$

and its PDF can be seen as the sensitivity of the above probability with respect to the parameter  $\beta_0$ . The integral in Eq. (5.4) is also of the type shown in equation (5.1), except that the limit state function here involves the conditional reliability index  $\beta(\boldsymbol{\theta})$ .

In applications, one is often interested in a simple measure to describe the effect of statistical and model uncertainties on computed  $P_F$  and  $\beta$ . One measure of this kind is an interval on  $P_F$  or  $\beta$  with a specified probability or confidence. The probability distribution described above can be used to generate such intervals. Another convenient measure is the variance of  $P_F(\boldsymbol{\theta})$  or  $\beta(\boldsymbol{\theta})$  with  $\boldsymbol{\theta}$  regarded as random variables. These variances can be obtained from the distribution described above, however, it is simpler to use a first order approximation, which for  $\beta$  leads to

$$\text{Var}(\beta(\boldsymbol{\theta})) \approx \nabla_{\boldsymbol{\theta}}\beta \cdot \mathbf{C}_{\boldsymbol{\theta}\boldsymbol{\theta}} \cdot \nabla_{\boldsymbol{\theta}}\beta^T \quad (5.5)$$

where  $\nabla_{\boldsymbol{\theta}}\beta$  is the gradient vector of  $\beta$  at the mean of  $\boldsymbol{\theta}$  and  $\mathbf{C}_{\boldsymbol{\theta}\boldsymbol{\theta}}$  is the covariance matrix of  $\boldsymbol{\theta}$ . This variance together with an assumed distribution can be used to obtain an approximate confidence interval on  $\beta$ . A similar equation for the variance of  $P_F(\boldsymbol{\theta})$  can be written, which is, however less accurate, because typically  $P_F$  has a strong non-linear dependence on  $\boldsymbol{\theta}$  (Der Kiureghian, 1996). For this reason Der Kiureghian (1996) suggests to use the confidence interval of  $\beta$  together with Eq (5.2) to obtain a confidence interval on  $P_F$ .

### 5.3.4 Time variant reliability problems

A reliability problem is said to be time-variant when the limit-state function depends on time,  $t$ . One important case is when some of the uncertain variables are stochastic in nature, as in  $g(\mathbf{x}, \mathbf{y}(t))$ , where  $\mathbf{y}(t)$  denotes a vector of stochastic processes. For example,  $\mathbf{x}$  may denote uncertain mass, stiffness or damping properties of a structure, which are usually time-invariant, and  $\mathbf{y}(t)$  may denote ground acceleration processes at the support points of the structure. For this class of reliability problems, the failure event constitutes the out-crossing of the vector process  $\mathbf{y}(t)$  through the limit-state surface  $g(\mathbf{x}, \mathbf{y}) = 0$ . Usually it is necessary to solve this problem by conditioning on  $\mathbf{x}$ , i.e.

$$P_F = \int P\left(\min_{0 \leq t \leq T} (g(\mathbf{x}, \mathbf{y}(t))) \leq 0 \mid \mathbf{x}\right) f(\mathbf{x}) d\mathbf{x} \quad (5.6)$$

where  $T$  denotes the structure lifetime. The conditioned failure probability for given  $\mathbf{x}$  is solved by the methods of stochastic process theory and the integral is evaluated by one of the methods described later on. The nested reliability approach described in the following section is particularly suitable for this purpose, provided certain smoothness conditions on the out-crossing problem are satisfied.

The exact solution of the out-crossing probability unfortunately is not available even for the special case of a linear limit-state function and stationary Gaussian processes. For small failure probabilities, however, the following upper bound provides a good approximation:

$$P\left(\min_{0 \leq t \leq T} (g(\mathbf{x}, \mathbf{y}(t))) \leq 0 \mid \mathbf{x}\right) \leq P(g(\mathbf{x}, \mathbf{y}(0))) + \int_0^T v(\mathbf{x}, \mathbf{t}) dt \quad (5.7)$$

where  $v(\mathbf{x}, \mathbf{t})$  denotes the mean rate of out-crossing of  $\mathbf{y}(t)$  for given  $\mathbf{x}$ . The first term on the right-hand side is a time invariant reliability problem of the type defined in Eq. (5.1). The main difficulty is in computing the mean out-crossing rate and considerable effort in recent years has been devoted to finding exact or approximate solutions to this problem. For a linear limit-state surface, the out-crossing problem can be converted to a scalar process up-crossing, a problem for which a well known solution by Rice is available. For a nonlinear limit-state surface, a generalization of the Rice formula is available that requires integration over the surface. Based on this formula, solutions of the out-crossing rate for special cases of the limit-state surface (e.g., polyhedral or ellipsoid surfaces) have been derived. Asymptotic solutions for certain classes of

processes have also been derived. These solutions are rather involved and require a considerable amount of computation. Another idea has been to liberalize the limit-state surface at an appropriate point (either the point of maximum mean out-crossing rate or the point of maximum likelihood) and convert the problem into an up-crossing rate problem. A comprehensive review of the afore mentioned classical random approaches is provided by Pinto et al. (2004), who point out that limitations of these approaches are severe, especially in view of aspects peculiar to earthquake engineering, i.e. non-linear behaviour and the presence of uncertainties in the structural model.

Another class of time-variant reliability problems is defined by a limit-state function which is a direct function of time, i.e. of the form  $g(\mathbf{x}, t) = 0$ . For example, an uncertain structure subjected to a time-varying deterministic load belongs to this class. In some cases it is possible to represent stochastic processes in terms of random variables and deterministic functions of time. In that case, a problem defined by a limit-state function of the form  $g(\mathbf{x}, \mathbf{y}) = 0$  can be converted to one of the form  $g(\mathbf{x}, t) = 0$ , where  $\mathbf{x}$  now includes the random variables defining. The probability of failure for this class of problems is described by the integral in Eq. (5.1), where the failure domain now is  $D = \{\min_{0 \leq t \leq T} (g(\mathbf{x}, t)) \leq 0\}$ . Hence, the methods available for solving time-invariant reliability problems are also applicable to this class of problems. However, caution should be exercised since the limit-state surface in this case is not continuously differentiable and can be strongly non-linear (Der Kiureghian, 1996). The above formulations can be extended to space-variant reliability problems, such as those involving random fields of material properties or loads, or those where the limit state function is explicitly dependent on a spatial coordinate. Furthermore, one can formulate problems where the limit-state function is dependent on both time and space coordinates.

## 5.4 Probability computation methods

Researchers devoted great deal of effort to developing efficient algorithms for computing probability integrals of the type in Eq. (5.1). A straightforward integration, by analytical or numerical means, usually is not possible because of the arbitrary nature of the integration domain and the typically high dimension of the problem. Often the size of  $\mathbf{x}$  is large, and consequently indirect approaches for evaluating the integral are essential. A review of the most widely used of these methods is presented in the present section. Several of the methods described below require transformation of the random variables in the original space,  $\mathbf{x} \in \mathbb{R}^d$ , where  $d$  represents the dimension of the

problem, into the standard normal space,  $\mathbf{u} \in \mathbb{R}^d$ , where each component of the vector  $\mathbf{u}$  is associated with an independent central unit Gaussian standard distributions. The transformation, which is nonlinear for non Gaussian random variables, is expressed as  $\mathbf{u} = \mathbf{u}(\mathbf{x})$ , where  $\mathbf{u}$  has the standard normal density. These transformations,  $\mathbf{u} = T_{xu}(\mathbf{x})$  and  $\mathbf{x} = T_{ux}(\mathbf{u})$ , are established by applying for example the Rosenblatt's or the Nataf's transformation (Ditlevsen and Madsen, 1996; Pinto et al., 2004).

There is a vast literature of different procedures on how to estimate the probability of failure given the performance function. However, the performance function,  $g(\mathbf{x})$ , can only be specified explicitly for rather simple and particular cases. In most cases of practical interest however, the function  $g(\mathbf{x})$  will not be known explicitly in terms on an analytically tractable expression. In cases where a deterministic finite-element-method (FEM) analysis code will be used to compute the structural response, the performance function will be known only point wise, i.e. the performance  $g_i = g(\mathbf{x}^{(i)})$  can be computed for each vector  $\mathbf{x}^{(i)}$ . Using the transformations mentioned above, the performance function  $g_u$  defined in the standard normal space, can be determined as follows:

$$g_u(\mathbf{u}) = g(\mathbf{x}) = g(T_{ux}(\mathbf{u})) \quad (5.8)$$

In other words, the evaluation of the performance at a single point  $\mathbf{u}^{(i)}$  in the standard normal space requires a transformation into the original space, a complete run of the FEM model and the computation of the performance form the response. The computational cost of evaluating the failure probability is governed by the number of structural analyses that have to be carried out. Hence, in view of feasibility and efficiency, the basic problem in structural reliability using deterministic FEM code should be specified as follows (Schuëller et al., 2004): determine the probability of failure within a specified confidence interval such that the number of required structural analyses is small.

### 5.4.1 First- and second-order methods

The first- and second-order reliability methods (FORM and SORM) take advantage of the fact that the point  $\mathbf{u}^* = \min(|\mathbf{u}|, g(\mathbf{u}) = 0)$ , where  $g(\mathbf{u})$  indicates the transform of the limit-state function in the  $\mathbf{u}$  space, is a point located on the limit state surface with minimum distance from the origin. Therefore this point has the highest probability density among all the failure points in the standard normal space. This point is known as the *design point* or the *most likely failure point*. It is evident that probability densities in the standard normal space are rotationally symmetric and decay exponentially with the

square of distance in both radial and tangential directions from  $\mathbf{u}^*$ . It follows that the major contribution to the probability integral in Eq. (5.1) comes from the neighbourhood of  $\mathbf{u}^*$ , provided the surface is not strongly nonlinear and there is only one significant design point. These conditions, for example, are satisfied for most structural component reliability problems (Der Kiureghian, 1996). Based on this, the limit-state surface in the neighbourhood of the design point is approximated by a first- or second-order surface for which the solution of the probability integral is available. Specifically, in FORM, the limit-state surface is replaced by the tangent hyper-plane at  $\mathbf{u}^*$  and the first order approximation of the failure probability is given by

$$P_F \approx \Phi(-\beta) \quad (5.9)$$

where  $\beta = \boldsymbol{\alpha} \mathbf{u}^*$ , the first-order reliability index, is the distance from the origin and  $\boldsymbol{\alpha}$  denotes the unit normal vector at the design point directed toward the failure set. In SORM, the limit-state surface is fitted with a second-order surface (usually a paraboloid) at  $\mathbf{u}^*$ , and the second-order approximation of the failure domain is given in terms of  $\beta$  and the principal curvatures,  $\kappa_i$ , of the paraboloid. The exact expression of this probability is given in terms of a single-fold integral, which must be evaluated numerically.

An important step in FORM/SORM is the determination of the design point. This requires the solution of a constrained optimization problem. Most algorithms construct a converging sequence of points according to the rule  $u_{k+1} = u_k + \lambda_k d_k$ ,  $k = 0, 1, \dots$ , where  $d_k$  denotes the search direction vector at step  $k$  and  $\lambda_k$  denotes the step size. In most efficient algorithms  $d_k$  is expressed in terms of the gradient of the constrain function,  $\nabla G$ .

When the limit-state surface in the standard normal space is strongly nonlinear, which may arise from strong nonlinearity of  $g(\mathbf{x})$  or from strong non-normality of the basic variables  $\mathbf{x}$ , multiple design points on the surface may occur with significant neighbourhood densities. For such cases, an improved approximation may be obtained by fitting first- or second-order approximations at all significant local design points followed by a series system analysis, provided certain convexity conditions are satisfied. Finding these local design points, however, is a challenge by itself since it is difficult to prevent convergence of the optimization algorithm to the global design point.

The major advantage of FORM/SORM is that the probability integral is computed with a relatively small number of computations of the limit-state function and its gradient.

Typically (Schuëller et al., 2004), for FORM the number of computations of each  $g$  and  $\nabla g$  is of order 10-100, and for SORM depending on the method used, extra computations of  $g$  of order  $10^d$  or  $d^2$  are required to determine the principal curvature where  $d$  denotes the number of random variables. These numbers are independent of the magnitude of the failure probability. These methods have the additional advantage of providing the coordinates of the most likely failure point and, in SOMR, the shape of the limit-state surface in the neighbourhood of the design point. Furthermore FORM readily provides the sensitivities of the failure probability, which are important in many applications.

As for the disadvantages, FORM provides a point estimate, subject to linearization errors, without confidence. Moreover, it requires the evaluation of the design point, which becomes difficult in high dimensions for non-linear limit state functions in the standard normal space. The efforts to compute the design point grows proportionally with the dimension  $d$ . SORM requires in addition to the design point  $n-1$  main curvature which cannot be obtained in a feasible manner for high dimensions  $d$ . The procedure implies that the domain close to the design point is the important domain which is not the case for high dimensions.

## 5.4.2 Simulation methods

### 5.4.2.1 Monte Carlo simulation

Among the procedures developed for the estimation of  $P_F$ , a prominent position is held by simulation methods. These techniques are based on traditional Monte Carlo sampling. If an indicator function  $I_f(\mathbf{x})$  is introduced which assumes a value equal to 1 if  $\mathbf{x}$  is in the failure domain and equal to 0 otherwise, then the integral in Eq. (5.1) can be rewritten as:

$$P_f = \int_{F \cup S} I_f(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = E[I_f(\mathbf{x})] \quad (5.10)$$

which shows that failure probability is the expected value of the indicator function according to the probability density function  $f_{\mathbf{x}}(\mathbf{x})$ . Therefore, in Monte Carlo simulation an estimator of the form

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N I_F(\mathbf{x}^{(i)}) \quad (5.11)$$

is used, where the samples  $\mathbf{x}^{(i)}$  in Eq. (5.11) are independently identically distributed (i.i.d) according to  $f(\mathbf{x})$ . It can be shown that the estimator in Eq. (5.11) is unbiased

(Pinto et al., 2004). The greatest advantage of Monte Carlo is its generality; in fact it can be applied to almost any class of problems. Furthermore Monte Carlo simulation the convergence rate, which in the mean square sense is measured by the coefficient of variation  $\delta_{MC} = \sqrt{\text{Var}[\hat{P}_F]} / P_F = \sqrt{(1 - P_F) / NP_F}$  of  $\hat{P}_F$ , is independent of the dimensionality of the random vector  $\mathbf{x}$ . Its main disadvantage is its inefficiency in estimating small failure probabilities  $P_F$  due to the large number (proportional to  $1/P_F$ ) of samples or equivalently system analyses needed to achieve an acceptable level of accuracy. Many variance reducing techniques have been proposed in the literature, e.g. importance sampling, directional sampling, importance sampling using design points, adaptive importance sampling, subset simulation, etc. It should be noted that some of these methods, directional simulation for example, are usually limited by the dimensionality of the problem. A brief overview of the most important and efficient (even in high dimensions) among these methods will be given in the following.

#### 5.4.2.2 Importance sampling

Importance sampling has been one of the most prevalent approaches in the context of simulation based methods for the estimation of structural reliability. The underlying concept is to draw samples of the vector of random parameters  $\mathbf{x}$  from a distribution  $f(\mathbf{x})$  which is concentrated in the important region of the random parameter space, that is the failure domain  $F$ . This requires re-expressing the probability of failure  $P_F$  as follows

$$P_F = \int I_F(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} = \int \frac{I_F(\mathbf{x}) f(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) d\mathbf{x} = E_h \left[ \frac{I_F(\mathbf{x}) f(\mathbf{x})}{h(\mathbf{x})} \right] \quad (5.12)$$

and using an estimator of the form

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N \frac{I_F(\mathbf{x}^{(i)}) f(\mathbf{x}^{(i)})}{h(\mathbf{x}^{(i)})} \quad (5.13)$$

where  $\{\mathbf{x}^{(i)}\}_{i=1}^N$  are distributed according to  $h(\mathbf{x})$ . The variance of  $\hat{P}_F$  when  $\{\mathbf{x}^{(i)}\}_{i=1}^N$  are i.i.d. is given by:

$$\text{Var}[\hat{P}_F] = \frac{1}{N} \left( \int \frac{I_F(\mathbf{x}) f^2(\mathbf{x})}{h^2(\mathbf{x})} f(\mathbf{x}) d\mathbf{x} - P_F^2 \right) \quad (5.14)$$

it is straightforward to verify that the latter becomes zero when

$$h(\mathbf{x}) = h_{opt}(\mathbf{x}) = \frac{I_F(\mathbf{x}) f(\mathbf{x})}{P_F} \quad (5.15)$$

which represents the optimal choice for the importance sampling density which is practically infeasible since it requires the knowledge of the failure probability a priori. Several techniques have been developed in order to approximate the optimal sampling density or to construct a different one which, exhibits a decreased variance of the estimator in Eq. (5.13). At present, the most prevalent approaches are those based on kernel density estimators or on design points.

Importance sampling is more robust and accurate than FORM or SORM, but not competitive to methods like sampling and subset simulation, because it is generally impossible to sample according to the optimal sampling density. The approach has difficulties to deal with multiple failure domains if they are not well separated. Used together with a surrogate response surface, the robustness and accuracy of importance sampling is lost, because such global simple analytical approximation cannot be determined efficiently in a robust and accurate manner for high dimensions.

#### 5.4.2.3 Importance sampling using design points

The design point – which was defined previously, is also the point of maximum likelihood (Der Kiureghian, 1996; Pinto et al., 2004). This is the best choice unless additional information on the true limit state surface is available. However, it should be noted that the contrary to FORM and SORM, the estimates obtained by importance sampling are not sensitive to the exact position of the design point, which therefore does not necessarily need to be determined with great accuracy. For a convex failure domain the centre of the sampling density might be shifted into the failure region or into the safe domain for a convex safe region.

#### 5.4.2.4 Importance sampling using kernel density estimators

As demonstrated in the previous sections, the design points and their neighbourhood do not always represent the most important region of the failure domain, especially in high dimensional spaces. Furthermore, the computational cost associated with their determination can be quite high which adversely affects the efficiency of the method.

Therefore, it has been proposed to use points lying inside the failure domain  $F$  in order to construct kernel density estimators of the optimal importance sampling density as described in the sequel. In the earliest versions of this method, the points in  $F$  were generated using standard rejection sampling from the original pdf (Ang et al., 1992). This of course is extremely inefficient in cases where PF is small since it requires the simulation of an exuberant number of points in order to obtain sufficient samples in the failure domain (Schuëller et al., 2004; Schuëller and Pradlwarter, 2007). The aforementioned problem was addressed in (Au and Beck, 1999) where a Markov Chain metropolis algorithm was used. Markov Chain techniques provide an ideal tool when one needs to sample from a distribution which is known up to a constant as is the case for the optimal importance sampling density (Eq. (5.15)). Points  $\mathbf{x}^{(i)}$  which are (asymptotically) distributed according to  $h_{opt}$  can be obtained as intermediate states of an irreducible Markov Chain whose unique stationary distribution is  $h_{opt}$ . The initial point  $\mathbf{x}^{(0)} \in F$  is selected either by rejection sampling or using engineering judgement (and this does not present serious difficulties in most systems of practical interest). The sequence of points  $\mathbf{x}^{(i)}$  obtained, will lie in  $F$ . Subsequently, a kernel sampling density estimator is constructed

$$k(\mathbf{x}) = \frac{1}{M} \sum_{i=1}^M \frac{1}{(w\lambda_i)^d} K\left(\frac{\mathbf{x} - \mathbf{x}^{(i)}}{w\lambda_i}\right) \quad (5.16)$$

using  $M$  points  $\mathbf{x}^{(i)}$  obtained by the Markov Chain. In the equation above,  $w$  is the window width,  $\lambda_i$  is the local bandwidth factor and  $K$  is the kernel pdf which is most commonly selected as the normal pdf:

$$K(\mathbf{x}) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi}} e^{-x_i/2} \quad (5.17)$$

The density in Eq. (5.16) is used as the importance sampling density in order to estimate PF based on Eq. (5.13).

The total cost of this approach, in terms of number of systems analyses needed equals the number of points generated by the Markov Chain in order to construct the kernel density estimate ( $M$  in Eq. (5.16)) and the number of points sampled from  $k(\mathbf{x})$  in the estimation of the failure probability ( $N$  in Eq. (5.13)). The quality of the estimate of the optimal sampling pdf provided by Eq. (5.16) depends on the particular parameters ( $w$ ,  $\lambda_i$ ,  $K$ ) but most importantly on the probabilistic characteristics of the points  $\mathbf{x}^{(i)}$  used.

Further details on the issues regarding this procedure can be found in Schuëller et al. (Schuëller et al., 2004).

#### 5.4.2.5 Subset simulation

Recently, a novel procedure called subset simulation was proposed by Au and Beck (2001) for structural reliability problems. It was further developed and investigated in the context of seismic risk dynamic analysis (Au and Beck, 2003). The basic concept on which it is based first appeared in the context of statistical physics under the name umbrella sampling. Subset simulation overcomes the inefficiency of direct Monte Carlo in estimating small probabilities, by expressing  $P_F$  as a product of larger, conditional probabilities. This is achieved by defining a decreasing sequence of events (subsets)  $\{F_i\}_{i=1}^m$  such that  $F_m = F$  and  $F_1 \supset F_2 \supset \dots \supset F_m = F$ . Due to the latter property,  $\bigcap_{i=1}^k F_i = F_k \forall k \leq m$ . As a result, the probability of failure  $P_F$  can be written as:

$$P_F = P(F_m) = P(F_1) \prod_{i=1}^{m-1} P(F_{i+1} | F_i) \quad (5.18)$$

It is obvious that by an appropriate selection of  $\{F_i\}_{i=1}^{m-1}$  the probabilities  $P(F_1)$  and  $P(F_{i+1} | F_i) \geq 1$  can be made sufficiently large so that their estimation can be performed efficiently by direct Monte Carlo estimators. Hence, the original problem is broken up into a series of  $m$  intermediate problems, for each one of which a solution can be obtained with a small number of simulations. Distinct estimators  $\hat{P}_1$  and  $\hat{P}_{i+1} \forall i \geq 1$  appearing in Eq. (5.18). A key ingredient in the overall efficiency of the method is the use of Markov Chains which allow the generation of the samples needed for  $\hat{P}_{i+1}$  by using the samples simulated in the previous step for calculating  $\hat{P}_i$ . For that purpose a Metropolis-Hastings algorithm can be applied similar to the discussed in the previous section. Unfortunately, the gain in efficiency that is achieved with the use of Markov Chains, comes at the expense of the independence of the samples generated which leads to biased estimated PF. The samples produced this way will be correlated which results in a decelerated convergence of the estimators  $\hat{P}_{i+1}$  compared to the i.i.d. case. Despite the dependence introduced in the samples generated, their use is justified as long as the respective estimators converge when the number of samples  $N$  goes to infinity.

The coefficient s of variation  $\delta_{i+1}$  of the estimators  $\hat{P}_{i+1}$  are given by

$$\delta_{i+1} = \sqrt{\frac{1 - P_{i+1}}{NP_{i+1}}} \sqrt{1 + \gamma_{i+1}} \quad (5.19)$$

where  $P_{i+1} = P(F_{i+1} | F_i)$ . The factor  $\gamma_{i+1} \geq 0$  depends on the correlation between the samples used for the estimation of  $P_{i+1}$  and is equal to zero when those are independent. The coefficient of variation of  $P_F$  is bounded from above

$$\delta^2 \leq \sum_{i,j=1}^m \delta_i \delta_j \quad (5.20)$$

If, for example  $P_F = 10^{-4}$  and  $m = 4$ ,  $P_i = 0.1$  and  $\gamma_i = \gamma$  then  $\delta$  will always be less than  $(12/N)\sqrt{1+\gamma}$ . In contrast, when direct Monte Carlo simulation is used, the coefficient of variation decays as  $(50/\sqrt{N})$  and therefore it will be larger than  $\delta$  unless the correlation between the samples, as measured by the factor  $\gamma$ , is considerably high. Subset simulation has a wide range of applicability. It performs well irrespectively of the geometry and number of the failure domains. It is also applicable to non-Gaussian distributed random variables. It retains the basic advantage of Monte Carlo whose performance is unaffected by the dimension of the random parameter vector.

#### 5.4.2.6 Line Sampling

Line sampling evolved from the need to treat high dimensional reliability problems with an implicitly available performance function  $f(\mathbf{x})$  obtained from deterministic FEM-analyses.

As an introduction to the procedure, consider the widely used importance sampling procedure, which can be written as in Eq. (5.13) repeated here for ease of lecture:

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N \frac{I_F(\mathbf{x}^{(i)}) f(\mathbf{x}^{(i)})}{h(\mathbf{x}^{(i)})} \quad (5.21)$$

In the standard normal space  $f(\mathbf{x})$  is standard normal distributed. Considering the estimator in Eq. (5.21), it is obvious that its variance, and hence its efficiency, depends only on two factors. First on the ratio  $N_F/N = (\sum_{i=1}^N I_F(\mathbf{x}^{(i)}))/N$ , and on the variance of the ratios  $R^{(i)} = f(\mathbf{x}^{(i)})/h(\mathbf{x}^{(i)})$ . In the ideal case of optimal importance sampling the first ratio is one and the second ratios are constant and equal to  $P_F$ . As already mentioned such sampling procedure is not realizable. However, something quite close to optimal can be obtained using line sampling if an important direction,  $\alpha$ , can be computed which points toward the failure domain nearest to the origin. Hence, the important direction assumes the role which the design point plays in importance sampling using design point and in FOMR/SORM. It is not required that the vector  $\alpha$

points exactly to the design point, nor are any assumptions made regarding the shape of the limit function.

Without loss of generality, it is possible to assume that  $x_1$  points in the direction of  $\boldsymbol{\alpha}$ . This can be always be ensured by a suitable rotation. If the sampling in direction  $x_1$  is assumed to be independent from sampling in all remaining directions, the estimate in Eq. (5.21) assumes the form:

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N I_F(\mathbf{x}^{(i)}) \frac{\exp\left(-\left(x_1^{(i)}\right)^2/2\right)}{h\left(\mathbf{x}_1^{(i)}\right)\sqrt{2\pi}} \prod_{k=2}^d \frac{\exp\left(-\left(x_k^{(i)}\right)^2/2\right)}{\sqrt{2\pi}} \frac{1}{h\left(x_2, \dots, x_d\right)} \quad (5.22)$$

The variation in the last term of the product will vanish if the sampling density  $h(x_2, \dots, x_d)$  normal to the direction  $\boldsymbol{\alpha}$  assumes the  $(d-1)$ -dimensional standard normal distribution. The optimal sampling normal to the important direction should follow a standard normal distribution, which corresponds exactly to direct Monte Carlo simulation in the subspace normal to  $\boldsymbol{\alpha}$ . Hence, under the condition that direct Monte Carlo will be used for the subspace  $x^\perp$ , the estimator in Eq. (5.22) reduces to the simple form:

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N I_F(\mathbf{x}^{(i)}) \frac{\exp\left(-\left(x_1^{(i)}\right)^2/2\right)}{h\left(x_1^{(i)}\right)\sqrt{2\pi}} \quad (5.23)$$

Clearly, it can never be efficient to perform importance sampling for a single standard normal distributed variable, because it approximates just the integral of the standard normal density function in the failure domain which can be computed quasi exactly using the cumulative standard normal distribution function  $\Phi(\cdot)$ . As a consequence, the above estimate for the failure probability is most efficiently approximated by

$$\hat{P}_F = \frac{1}{N} \sum_{i=1}^N P_F^{(i)} \quad (5.24)$$

with the conditional failure probabilities

$$P_F^{(i)} = \int_{-\infty}^{+\infty} I_F(\mathbf{x}^{(i)}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_1^{(i)})^2}{2}\right) = \Phi(\beta_1^{(i)}) - \Phi(\beta_x^{(i)}) \quad (5.25)$$

This procedure can be generalized for the case where  $x_1$  does not point in the direction  $\alpha$  (Schuëller et al., 2004). The independent estimates  $P_F^{(i)}$  for the failure probability allow to compute an unbiased estimate for the failure probability and to establish the variance of this estimate

$$\text{var}(P_F) = \frac{1}{N(N-1)} \sum_{i=1}^N (P_F^{(i)} - \hat{P}_F)^2 \quad (5.26)$$

which is the basis for deriving further confidence intervals. Line sampling can also be applied to an arbitrary number of directions. For multiple failure regions, each point  $\mathbf{x}$  of the  $d$ -dimensional standard space is associated in disjunct unique manner to one of the given direction.

The performances of line sampling can be enhanced by using a stepwise algorithm (Koutsourelakis et al., 2004). It utilizes the basic idea of subset simulation that was illustrated in the previous section. Specifically, the original problem of estimating  $P_F$  using Eq. (5.24) is split into a series of intermediate problems. The information gathered at every step is utilized in subsequent steps in order to obtain accurate solutions using a small number of samples.

Line sampling is capable to take advantage of simple flat limit states in standard normal space and samples in the most important domain, without assuming a linear or quadratic limit state surface or requiring a design point computation. Line sampling can be applied for any irregular limit state and can deal with any arbitrary number of failure domains. The advantages of line sampling become most pronounced when  $d \rightarrow \infty$ , but is also competitive in low dimensions. Furthermore, the application of the stepwise procedure proposed in Koutsourelakis et al. (Koutsourelakis et al., 2003) can further reduce the computational cost.

### 5.4.3 Response surface

An alternative approach from computing probabilities of the type in Eq. (5.1) is to replace the integration boundary by an approximating response surface and then perform the integration by an appropriate means without having to engage the actual limit-state function. This approach is particularly useful when the limit state function is algorithmic in form and its gradient is difficult to compute. To construct the response

surface one typically computes  $g(\mathbf{x})$  at a number of points and then fits a polynomial surface to the points by the least squares method. One important consideration in using the response surface methodology is the selection of the experiment points, which can be done following the theory of design of experiments. It is important to note that the number of simulations needed to calibrate the model grows as the number of variables in the model increase. Therefore the model is not computationally efficient in very high dimensions. However, this issue can be addressed using response surfaces with random effects (Faravelli, 1989) which allow to take into account implicitly the effects of some of variables involved in the considered problem. Doing so allows to greatly increase the computational efficiency but on the other hand reduces accuracy. Further details on response surface models can be found in Chapter 6 and Chapter 7.

## 5.5 Specialized methods for earthquake engineering

### 5.5.1 Seismic fragility and failure probability

The term fragility function, or simply fragility stands for the probability of exceeding a give state of structural performance (i.e. a limit state), as a function of one parameter describing the intensity of the ground motion, most typically the PGA or the  $S_a$  (see Section 2.3.1.1). In a structural engineering sense, such limit states for specific structural components and systems may be either strength or deformation-related. In a broader socioeconomic context, the LS may be related to repair costs that are in excess of a desired amount, opportunity losses, or morbidity/mortality. Limit state identification requires a thorough understanding of the behaviour of the safety-related systems within the plant and the role of structural components and systems in ensuring acceptable behaviour of such systems (Wen et al., 2004). With the limit state (LS) identified, its probability

$$P(LS) = \sum P(LS | IM = im) P(IM = im) \quad (5.27)$$

where  $IM$  is a random variable (or vector) describing the intensity of the demand (ground motion) on the system,  $P(LS | IM = im)$  is the conditional limit state probability given that  $IM = im$ , and the summation is taken over all the possible values of  $IM$ . The conditional probability  $P(LS | IM = im) = F_{LS}(im)$  is the fragility. The probability  $P(IM = im)$  defines the hazard (in earthquake engineering, the seismic hazard is defined by the cumulative distribution function  $P(IM > im)$ ), which can be determined for according to the procedures described in Chapter 1.

The fragility of a component or system defines the conditional probability of its attaining a performance limit state, which may range from loss of function to incipient collapse, given the occurrence of a particular operational or environmental demand. Furthermore, fragility provides a probabilistic measure of safety margin with respect to design-basis or other events.

Most commonly the states of interest in the non-linear range of behaviour therefore evaluation of fragilities normally requires some form of simulation, i.e. performing a set of non-linear analyses to obtain the statistics of the maximum response, on set for each ground-motion intensity. Given the statistics of the response and a suitable probabilistic model for it (the most frequently used one being the lognormal distribution), and knowing the CDF of the corresponding capacity, the fragility is evaluated as for the fundamental time-invariant reliability problem described in:

$$F_{LS}(im) = \int_0^{\infty} f_D(\alpha | IM = im) F_C(\alpha) d\alpha \quad (5.28)$$

where  $F_C$ , the CDF of capacity, and  $f_D$ , the pdf of demand, are both expressed in terms of the same structural response parameter  $\alpha$  (e.g. interstorey drift, roof drift, hysteretic energy).

It is noted that although only the intensity of the ground motion appears as a variable in  $F_{LS}(im)$ , this latter is also a function of all the other characteristics of the ground motion, notably the frequency content and the duration, from which the response is significantly affected. Hence, the fragility is not exclusively a property of the structure, but also of the features of the expected ground motion, and it will be in general different in different tectonic environments. The influence of the detailed features of the motion depends to a significant extent on the choice of the intensity parameter.

### ***5.5.2 Practical reliability methods***

So called practical reliability and risk methods specialized for seismic problems have been developing since the mid-seventies, mainly for nuclear and other important installations, but it is only recently that these approaches are finding a theoretical stability and wider application also for more ordinary types of construction. Practicality has obviously a cost in terms of loss of accuracy and generality with respect to the rigorous formulations: the nature and the consequences of these losses need to be well understood in order for these methods to be used properly. They rest on two assumptions, both referring to the time dimensions of the problem (Pinto et al., 2004): the temporal sequence of the seismic event and the dynamic response of the structure

during an event. The first one consists in choosing an appropriate interval of time, normally one year, and considering that if failure is to occur within any such interval, it will do so under the seismic event of largest intensity. Since the yearly maximum event is obviously random, the implication is that what is needed is a statistical basis and a probabilistic model for the CDF of the yearly maxima. If afterwards one wants to pass from the annual CFF to the same quantity referred to a longer period of time, the common practice is to advocate independence among annual events and to introduce the annual rate of occurrences of events, which assuming the process as Poissonian, allows the CDF of the maxima relative to any desired period of time to be obtained. The second assumption is that in a structure subjected to ground acceleration, failure occurs when the response attains its maximum. In more refined versions the dependence of the capacity of some of the failure models on the demand is accounted for, so that the failure event is defined as  $C(t, D(t)) - D(t) \leq 0$  and it may not occur when  $D(t) = D_{max}$ . This assumption is quite consequential, since it allows the collection, through a number of dynamic analyses, of the statistics of the maximum responses so as to obtain a vector of correlated random variables, to be subsequently compared with the corresponding random capacities, as for the case of a time-invariant problem. It remains to note that the definition of maximum demand is obvious in the case of SDOF structures or of a structure whose response is dominated by a single mode, in the latter case  $D_{max}$  coinciding with the maximum modal response. In a MDOF structure with no single dominant mode, the demands on its various members attain generally their maxima at different instants of time. The statistics of the individual maxima can still be collected, as well as their correlation: in general, however, a larger number of dynamic analyses are needed to obtain stable estimates of the statistical parameters. The components of the vector  $D$  are obviously not simultaneous: this recognised fact, that tends to reduce the correlation between them, cannot be accounted for in this category of approaches.

#### *5.5.2.1 Determination of seismic fragility*

In the literature it is possible to find a high number of practical procedures and methods proposed for defining seismic fragility or to directly obtain failure probability.

The fragility functions can be generated by using: (i) observation of damages caused by earthquakes, (ii) experimental data (usually at the component level), (iii) detailed analytical models, (iv) simplified analytical models, and (v) design information and engineering judgement.

Post-earthquake survey-based procedures rely theoretically on the most reliable data source — damage data from earthquakes. In a number of such procedures (e.g. Gulkam and Sozen, 1999; Yucemen et al., 2004), some basic structural information (e.g., number of stories, structural system), material properties (e.g., in-situ concrete strength), apparent structural deficiencies (e.g., vertical and plan irregularities), and building site location were collected with damage data through a post-earthquake survey. This information was utilized to arrive at a rating score or index in which the numerical value usually determines whether the building is safe or unsafe, with respect to the traditional goal of assuring life safety. Other procedures (e.g. Shinozuka et al., 2000; Rossetto and Elnashai, 2003) utilized the collected information for developing seismic assessment tools in the form of fragility curves.

These observation-based or empirical procedures are highly specific to a particular seismo-tectonic, geotechnical, and built environment (Rossetto and Elnashai, 2003). Application of these procedures to regions other than those for which they were developed often does not yield satisfactory damage estimates when seismic characteristics and building infrastructure differ. Consequently, they have found limited use only in highly seismic regions, where they have been used to rank seismic vulnerability of buildings. In contrast, more recent seismic vulnerability and risk assessment procedures require multiple performance evaluations within a performance or consequence-based framework, which may only be feasible with simulation-based analytical procedures. Before the resources for intensive analytical simulations required for the derivation of fragility curves were available, expert opinion had been the substitute for analytical simulations. The ATC relied on expert opinion with limited observational data from the 1971 San Fernando earthquake when preparing the ATC-13 report (ATC, 1985), which is one of the first applications of fragility modelling to civil infrastructure subjected to earthquake demand. The reliability of the fragilities in ATC-13, which were identified in terms of damage state probability matrices, is questionable in that the fragilities are subjective and the associated degree of conservatism is unknown (Rossetto and Elnashai, 2003).

The well-known loss estimation software package, HAZUS, developed under the sponsorship of FEMA (FEMA, 2003), is also based on expert opinion to a considerable degree (Celik, 2007). HAZUS incorporates fragilities for 36 categories of building and 4 damage states, where the fragilities are modeled by lognormal distributions with the distribution parameters based primarily on expert opinion. The loss estimation methodology classifies the buildings in terms of building type on the basis of their height and structural system and seismic design level on the basis of the seismic

standard used in their design, the seismic zones in which they are built, their design vintage, and their use (Kircher et al., 1997a). Based on this classification, building capacity is represented by a non-linear static pushover curve in the form of base shear versus roof displacement, and building response to an input scenario earthquake, considering the local site conditions, is determined with the capacity spectrum method. The building response is then entered into the associated built-in fragility curves defined at the thresholds of four discrete damage states (slight, moderate, extensive, and complete), defined separately for the structural system and for drift- and acceleration-sensitive non-structural components, to perform the loss estimation calculations given the occupancy class of the building (e.g., residential, commercial) (Kircher et al., 1997b). Whitman et al. (1997) observed that the losses estimated using HAZUS should be viewed with caution since they may be off by a factor of two, or even more in the CEUS. But the most significant limitation is that, HAZUS does not provide for the analysis or propagation of uncertainty (Celik, 2007).

More recent approaches have relied on analytical simulations. The relation between structural response and earthquake ground motion intensity, which is the basic ingredient for deriving the fragility curves, is established through analytical simulations with varying comprehensiveness. Differences also exist in characterization of earthquake hazard, structural damage, performance limits, etc.

Hwang and Jaw (1990) proposed a procedure to calculate fragility curves taking into account uncertainties in both ground-motion and structure. The uncertainty in each parameter defining the earthquake-structure system is characterized by several representative values that are selected considering the uncertainty range of the parameter and its use in engineering practices. Samples of structures and earthquake motions are constructed from the combination of these representative values, and then the latin hypercube sampling technique is used to construct the samples of earthquake-structure system. For each sample, the nonlinear seismic analysis is performed to produce response data, which are then statistically analyzed. Five limit states representing various degrees of structural damage are defined and the statistics of the structural capacity corresponding to each limit state can be established. The fragility curve is generated by evaluating the limit state probabilities at different levels of peak ground acceleration.

Singhal and Kiremidjian (1996) developed fragility curves for low-, mid-, and high-rise RC frames that were designed using seismic provisions. The uncertainty in structural capacity and demand was taken into account through Monte Carlo simulations. Stochastically generated frame models randomly paired with simulated ground motion

records were used in the non-linear time history analyses. Structural demand versus seismic intensity relationships were determined from so-called stripe analyses (i.e. incremental dynamic analyses with ground motions scaled to different intensity levels (Vamvatsikos and Cornell, 2002)). The structural demand at each seismic intensity level was assessed using ground motions scaled to that particular intensity level and was represented by a lognormal probability density function. The lognormal model of demand was then utilized to compute fragility estimates (for the performance limits considered) at that particular level. Finally, fragility curves were represented by lognormal cumulative distribution functions that were fit to individual fragility estimates, computed at several seismic intensity levels. Singhal and Kiremidjian (1998) later presented a Bayesian method for updating the fragility curves that they developed earlier for low-rise RC frames and estimating confidence bounds on those fragility curves, by using the observed building damage data from the 1994 Northridge earthquake.

Mosalam et al. (1997) developed fragility curves for RC frames with and without masonry infill walls. Single-degree-of-freedom (SDOF) models were employed in the analysis. These models were obtained from the adaptive nonlinear static pushover analyses of the frame models, which were generated using Monte Carlo simulations to take into account the uncertainty in structural material properties. The structural responses of these SDOF models to each ground motion (i.e., each model was paired with each ground motion rather than randomly as in Singhal and Kiremidjian (1996)) were used to determine the fragility estimates (for the performance limits considered) for that particular ground motion.

Shinozuka et al. (2000) developed both empirical and analytical fragility curves for bridges. The empirical fragility curves utilized the observed bridge damage data from the 1995 Kobe earthquake. In contrast, the analytical fragility curves utilized such data that were simulated from the nonlinear time history analyses of stochastically generated models of two bridges, taking into account the uncertainty in structural material properties. Both fragility curves were represented by lognormal distribution functions with the distribution parameters estimated using the maximum likelihood method. Confidence intervals for the distribution parameters were also provided.

Porter et al. (2001) proposed an assembly-based vulnerability framework for assessing the seismic vulnerability of buildings on a building-specific basis. The proposed approach differs from the above analytical procedures in that “a vulnerability function” that relates the seismic losses to the seismic intensity was developed for a particular building and the damage to individual assemblies was determined for this purpose. The

seismic losses were assessed using stripe analyses. The structural response to each scaled ground motion was entered into assembly fragility curves, and the associated damage to each structural and non-structural element in the building and to its contents was determined as outlined in the study. The total damage was then expressed in terms of the sum of repair and loss-of-use costs as a fraction of replacement cost. After performing a regression analysis on the generated data, the seismic vulnerability function was obtained for a particular building. The application of the proposed framework to a steel moment frame building revealed that substantial uncertainty exists in the vulnerability function derived for the building.

Cornell et al. (2002) developed a probabilistic framework for seismic design and assessment of structures in a demand and capacity format, addressing the uncertainties in hazard, structural, damage, and loss analyses. Structural-demand versus seismic-intensity relationships were determined from a so-called cloud analysis (i.e. nonlinear dynamic analyses using accelerograms not scaled to the same intensity levels). The structural demand was assessed using a suite of ground motions and the median structural demand was represented by a log-linear function of seismic intensity. The structural demand was assumed to be distributed lognormally about the median with constant logarithmic standard deviation. This framework provided the probabilistic basis for the design recommendations that resulted from the SAC project. The procedure proposed by Cornell et al. (2002) has been extended by Lupoi et al. (2002) in order to be applied to reinforced concrete frame structures.

Franchin et al. and Schotanus et al. (Franchin et al., 2003b; Franchin et al., 2003a; Franchin et al., 2004; Schotanus et al., 2004) developed fragility curves using a response surface with random block effects approach. This procedure allows to take into account uncertainties in both ground-motion and structure and can be used in conjunction with state-of-the-art FEM models. Structural capacity is approximated by a polynomial response surface as a function of the uncertain structural parameters. Uncertainty in ground-motion or in spatial variability of material strength is taken into account implicitly introducing random factors in the response surface model. Data required to calibrate the model is gathered performing an ensemble of non-linear incremental dynamic analysis planned according to the theory of the Design of Experiments. This method has also been investigated by the author (Buratti et al., 2006b; Buratti et al., 2006c; Buratti et al., 2006a; Buratti et al., 2007b; Buratti et al., 2007c). Further details and a case-study application can be found in Chapter 6 and Chapter 7 respectively.

Erberik and Elnashai (2004) used the same methodology as in Singhal and Kiremidjian (1996) for deriving fragility curves for mid-rise flat-slab RC buildings with masonry infill walls. However, they paired each stochastically generated building model with each ground motion record considered in the study rather than randomly matching the models with the ground motions. Performance limits for which the fragility curves were developed were identified from nonlinear static pushover analysis of the building.

Rossetto and Elnashai (2005) developed fragility curves for low-rise RC frames with masonry infill walls that were designed according to the seismic design code in place in Italy in 1982. Structural demand versus seismic intensity relationships were determined using the same methodology as in Erberik and Elnashai (2004) but the capacity spectrum method with adaptive nonlinear static pushover analysis was employed. A response surface equation was fit to the demand versus intensity data. Fragility curves were then developed using a larger dataset at refined seismic intensity levels, which was generated through a re-sampling process from the response surface equation. Confidence bounds were also identified on the fragility curves.

Kwon and Elnashai (2006) developed fragility curves for low-rise GLD RC frames but the problematic reinforcing details associated with such frames, such as the inadequate joint shear capacity and the insufficient positive beam bar anchorage, were not considered. The finite element model of the three-story GLD RC frame was validated using experimental data from the shake table tests (Bracci et al., 1995a; Bracci et al., 1995b). The fragility curve derivation methodology followed that in Erberik and Elnashai (2004) with full combination of the randomly generated material strength parameters in the generated frame models. The analysis of structural demand statistics indicated that the effect of material uncertainty is negligible with respect to that of ground motion uncertainty. Furthermore, the comparison of fragility curves that were developed using different sets of ground motions revealed a dependency on the choice of the ensemble as in Mosalam et al. (1997).

In a recent study, Ramamoorthy et al. (2006) also developed fragility curves for low-rise RC frames. The structural demand was assessed using a cloud analysis based on nonlinear time history analysis and the median demand was represented by a bilinear function rather than a linear function as in Cornell et al. (2002) with the regression parameters estimated from a Bayesian methodology presented in the study. More recently (2008) extended the procedure adopted in the aforementioned study in order to provide confidence bounds for the fragility curves.

Lupoi et al. (Lupoi et al., 2006) proposed a method for computing fragility function with the aim of taking in account different aspect involved in the reliability problem: that is,

the possibility of multiple failure modes to occur and their reciprocal interaction, the uncertainty in structural capacity, and the influence on dynamic response of the variability of system parameters. The method has some important similarities with the one presented by Gardoni et al. (Gardoni et al., 2002; Gardoni et al., 2003). In fact, they both share the idea of obtaining the system fragility for a structure having multiple modes of failure, by solving a reliability problem that involves structural capacities at the component level and the corresponding, earthquake-induced demands. The main feature of the method in Gardoni et al. (Gardoni et al., 2002; Gardoni et al., 2003) consists of using Bayesian techniques to correct for the systematic and random errors in the capacity and demand models, while in the method in Lupoi et al. (2006) it is assumed that these corrections are already implemented into the models.

## 5.6 Conclusions

The present chapter presented an introduction of the structural reliability problem, with particular focus on seismic reliability of structures. Different formulations and solution procedures have been presented and discussed.

Most simulation procedures are computationally too expensive in order to be applied to seismic reliability problems, and are mainly used as reference for validating practical procedures. Therefore, in the great majority of seismic reliability problems, practical methods adopted. These procedures, aimed at calculating seismic fragility, are often based on strong assumptions and simplifications. For example many approaches use very simplified structural models (e.g. SDOF systems), take into account only uncertainty in ground-motion, not considering uncertainty in structural parameters. Among the methods described above, the most promising seem to be those based on response surfaces. In fact response surface approach can be applied to different structural systems, it can take into account multiple failure modes, it can be used together with state-of-the-art non-linear models, and has good computational efficiency, especially when random factors are used to describe implicitly the effects of some groups of variables involved in the problem.

For these reasons the remaining chapters of the present work are dedicated to assessment of seismic fragility via response surface method. The theoretical basis of the method is described in Chapter 6 and a case study is presented in Chapter 8. In particular, in the case study, some critical aspects of response surface models are investigated.

## 6 Response Surface Method

The Response Surface (RS) method is based on the definition of a statistical model expressing a structural response parameter as a function of a set of variables, called *factors* in the statistical language. The RS is typically based on a polynomial function, and has been used in many research fields (Box and Draper, 1987; Searle et al., 1992; Khuri and Cornell, 1996). This chapter describes the theory of standard response surfaces as well as the theory of response surfaces with random factor. This chapter also introduces the theory of design of experiments which gives criteria to define the experiments needed to obtain data for response surface calibration.

### 6.1 Standard Response Surface

The response variable is the measured quantity the value of which is assumed to be dependent upon the levels of the factors. The true value of the response corresponding to any particular combination of the factor levels and in the absence of experimental error of any kind is denoted by  $\eta$ . The dependence of  $\eta$  on the factors can be written as:

$$\eta = \phi(x_1, x_2, \dots, x_k) \quad (6.1)$$

where  $\phi$  is an unknown function, called the true response function, and it is assumed to be continuous. Considering the response function for a single factor it is possible to represent it locally to any required degree of approximation with a Taylor series expansion about some arbitrary point  $x_{10}$ , i.e.,

$$\eta = \phi(x_{10}) + (x_1 - x_{10})\phi'(x_{10}) + \frac{1}{2}(x_1 - x_{10})^2\phi''(x_{10}) + \dots \quad (6.2)$$

where  $\phi'$  and  $\phi''$  are, respectively, the first and second derivatives of  $\phi$  with respect to  $x_1$  and evaluated at  $x_{10}$ . The expansion reduced to a polynomial of the form

$$\eta = \phi(x_1) = \beta_0 + \beta_1 x_1 + \beta_{11} x_1^2 + \dots \quad (6.3)$$

where the coefficients  $\beta_0$ ,  $\beta_1$  and  $\beta_{11}$  are parameters which depend on  $x_{10}$  and the derivatives of the response function at  $x_{10}$ . Therefore if  $k$  factors are considered Eq. (6.2) can be rewritten as

$$\eta = \varphi(\mathbf{x}) + \nabla\varphi(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_{10}) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_{10})^T \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_{10}) + \dots \quad (6.4)$$

where  $\mathbf{x}$  is a vector containing the levels of the factors,  $\mathbf{x}_0$  is the centre point of the Taylor expansion,  $\nabla\varphi$  is the gradient of  $f$  and  $\mathbf{H}$  is the hessian matrix. Eq. (6.4) can be rewritten as a polynomial equation in the factor levels, in a form similar to that of Eq. (6.3):

$$\eta = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \sum_{j=1}^k \beta_{ij} x_i x_j \quad (6.5)$$

A statistical model of this kind is called fixed effect model because the effects are related to a finite set of levels of the factors. The model in Eq. (6.3) is non-linear as far as the  $x_i$  variables are concerned but is linear with regard to  $\beta$ . The latter are called regression coefficients or parameters. The coefficients  $\beta_i$  are the values of the first-order partial derivatives and are referred to as first-order effects. The coefficients  $\beta_{ij}$  are defined as the values of the second order partial derivatives and are called second-order effects. The structural form of  $\eta$  is usually unknown and therefore an approximating form is sought using a polynomial or some type of empirical model equation. Furthermore, as far as experimental data is concerned, the real value of the response is not known and is substituted by the expected values of the response parameter  $E(y)$ . The model in Eq. (6.5) can therefore be rewritten as (Searle et al., 1992; Khuri and Cornell, 1996):

$$E(Y) = \eta = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} \quad (6.6)$$

where  $\mathbf{f}(\mathbf{x})$  and  $\boldsymbol{\beta}$  are defined as

$$\mathbf{f}(\mathbf{x}) = \left[ 1 \quad x_1 \quad \dots \quad x_k \quad x_1^2 \quad \dots \quad x_k^2 \quad x_1 \cdot x_2 \quad \dots \quad x_1 \cdot x_k \quad \dots \quad x_{k-1} \cdot x_k \right]^T \quad (6.7)$$

and

$$\boldsymbol{\beta} = \left[ \beta_0 \quad \beta_1 \quad \dots \quad \beta_k \quad \beta_{11} \quad \dots \quad \beta_{kk} \quad \beta_{12} \quad \dots \quad \beta_{1k} \quad \dots \quad \beta_{k-1,k} \right]^T \quad (6.8)$$

The deviation of the  $i$ -th observation  $y_i$  (the result of a numerical simulation in the present study) from the expected value  $E(y_i)$  can be written as:

$$\varepsilon_i = y_i - E(y_i) = y_i - \mathbf{f}(\mathbf{x}_i)^T \boldsymbol{\beta} \quad (i = 1, \dots, N) \quad (6.9)$$

Therefore, considering  $n$  observations, the model in Eq. (6.9) can be expressed in matrix notation, as Khuri and Cornell, 1996:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (6.10)$$

where  $\mathbf{y}$  is a vector collecting the observed response values

$$\mathbf{y} = [y_1 \dots y_n]^T \quad (6.11)$$

$\mathbf{X}$  is a  $n \times p$  matrix the  $i$ -th row of which is the vector of monomials  $\mathbf{f}(\mathbf{x})^T$  calculated at the values assumed by the variables  $x$  at the  $i$ -th trial

$$\mathbf{X} = \begin{bmatrix} \mathbf{f}(\mathbf{x}_1)^T \\ \vdots \\ \mathbf{f}(\mathbf{x}_n)^T \end{bmatrix} \quad (6.12)$$

and  $\boldsymbol{\varepsilon}$  is the vector containing the random errors with respect to  $E(y_i)$ :

$$\boldsymbol{\varepsilon} = [\varepsilon_1 \dots \varepsilon_n]^T \quad (6.13)$$

As an example, if 2 factors and 6 experiments are considered, the model in Eq. (6.10) can be written as:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{2,1} & x_{1,1}^2 & x_{2,1}^2 & x_{1,1}x_{2,1} \\ 1 & x_{1,2} & x_{2,2} & x_{1,2}^2 & x_{2,2}^2 & x_{1,2}x_{2,2} \\ 1 & x_{1,3} & x_{2,3} & x_{1,3}^2 & x_{2,3}^2 & x_{1,3}x_{2,3} \\ 1 & x_{1,4} & x_{2,4} & x_{1,4}^2 & x_{2,4}^2 & x_{1,4}x_{2,4} \\ 1 & x_{1,5} & x_{2,5} & x_{1,5}^2 & x_{2,5}^2 & x_{1,5}x_{2,5} \\ 1 & x_{1,6} & x_{2,6} & x_{1,6}^2 & x_{2,6}^2 & x_{1,6}x_{2,6} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_{11} \\ \beta_{22} \\ \beta_{12} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix} \quad (6.14)$$

where  $x_{i,j}$  is the  $j$ -th level of the  $i$ -th factor. Matrix  $\mathbf{X}$  is named design matrix.

### 6.1.1 Least squares method

Given a model in the form of Eq. (6.10) the regression coefficient values can be estimated using the least squares method: one of the simplest in statistics. The following assumptions are to be made about the errors  $\boldsymbol{\varepsilon}$ :

1. Random errors have zero mean and common variance,  $\sigma^2$ .
2. Random errors  $\varepsilon_i$  are mutually independent in the statistical sense.

For tests of significance ( $t$ - and  $F$ -statistics), and confidence interval estimation, an additional assumption must be satisfied:

3. Random errors  $\varepsilon_i$  are normally distributed.

The method of least squares selects as estimates for the unknown parameters in Eq. (6.10), those values,  $b_0, b_1, \dots, b_k$ , which minimize the quantity:

$$R(\boldsymbol{\beta}) = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (6.15)$$

Setting to zero the derivatives of (6.15) with respect to  $\boldsymbol{\beta}$  gives:

$$\mathbf{X}^T \mathbf{X}\boldsymbol{\beta} - \mathbf{X}^T \mathbf{y} = 0 \quad (6.16)$$

from which the estimates of the least squares estimates of the elements of  $\boldsymbol{\beta}$  can be obtained:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (6.17)$$

It is worthily noted that this method does not require the value of the variance of the error to be known. The statistical properties of the estimator  $\mathbf{b}$  derive from the assumptions concerning the elements of  $\boldsymbol{\varepsilon}$ . The expectation vector of  $\mathbf{b}$  is

$$\begin{aligned} E(\mathbf{b}) &= E\left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}\right] = E\left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})\right] = \\ &= \boldsymbol{\beta} + E(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \boldsymbol{\varepsilon} = \boldsymbol{\beta} \end{aligned} \quad (6.18)$$

therefore  $\mathbf{b}$  is an unbiased estimator of  $\boldsymbol{\beta}$ . The variance-covariance matrix of the vector of the estimates is

$$\mathbf{C}_{bb} = \text{var}(\mathbf{b}) = \text{var}\left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}\right] = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{var}(\mathbf{y}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \quad (6.19)$$

and, since the covariance matrix of the vector  $\mathbf{y}$  can be written (accordingly to assumption 1) as

$$\mathbf{C}_{YY} = \text{var}(\mathbf{y}) = E[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T] = E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T] = \text{var}(\boldsymbol{\varepsilon}) = \mathbf{C}_{\varepsilon\varepsilon} = \sigma_\varepsilon^2 \mathbf{I}_n, \quad (6.20)$$

the variance-covariance matrix of  $\mathbf{b}$  can be rewritten as

$$\mathbf{C}_{bb} = \text{var}(\mathbf{b}) = (\mathbf{X}^T \mathbf{X})^{-1} \sigma_\varepsilon^2 \quad (6.21)$$

Along the main diagonal of the matrix  $\mathbf{C}_{bb}$ , the  $ii$ -th element, is the variance of  $b_i$ . The  $ij$ -th element of  $\mathbf{C}_{bb}$  is the covariance between the elements  $b_i$  and  $b_j$  of  $\mathbf{b}$ . If the errors are jointly normally distributed, then  $\mathbf{b}$  is distributed as a normal multivariate:

$$\mathbf{b} \sim N\left[\boldsymbol{\beta}, (\mathbf{X}^T \mathbf{X})^{-1} \sigma_\varepsilon^2\right] \quad (6.22)$$

Another important property of the estimator  $\mathbf{b}$  is that it produces the minimum variance estimates of the elements of  $\boldsymbol{\beta}$ , therefore it is called best linear estimator of  $\boldsymbol{\beta}$ .

One of the purposes in obtaining a fitted model is to use the model for predicting response values at points throughout the experimental region. Let  $\mathbf{x}$  denote a  $p \times 1$  vector the elements of which correspond to the elements of a row of the matrix  $\mathbf{X}$  in Eq. (6.10). An expression for the predicted value of the response, at any point  $\mathbf{x}$  in the experimental region is:

$$\hat{y}(\mathbf{x}) = \mathbf{x}'\mathbf{b} \quad (6.23)$$

A measure of the precision of the prediction  $\hat{y}(\mathbf{x})$ , defined as the variance of  $\hat{y}(\mathbf{x})$ , is expressed as

$$\begin{aligned} \text{var}(\hat{y}(\mathbf{x})) &= \text{var}\left[\mathbf{f}(\mathbf{x})^T \mathbf{b}\right] = \mathbf{f}(\mathbf{x})^T \text{var}(\mathbf{b}) \mathbf{f}(\mathbf{x}) = \\ &= \mathbf{f}(\mathbf{x})^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{f}(\mathbf{x}) \sigma_\varepsilon^2 \end{aligned} \quad (6.24)$$

It is worth noting that the variance of the prediction depends on  $\mathbf{x}$ , i.e. it is not constant throughout the experimental region.

In Eq. (6.19) for the variance-covariance matrix of  $\mathbf{b}$ , as well as in Eq. (6.24) for the variance of  $\hat{y}(\mathbf{x})$ , the variance of errors,  $\sigma_\varepsilon^2$ , was assumed known. This assumption is seldom true and usually an estimate  $\hat{\sigma}_\varepsilon^2$  is needed. The estimate is obtained from the

analysis of the data values. For the general case where the fitted model contains  $p$  parameters and the total number of observations is  $n$  ( $n > p$ ), the estimate,  $s^2$ , is computed from

$$\begin{aligned} s^2 &= \frac{1}{n-p} \sum_{i=1}^n \varepsilon_i^2 = \frac{1}{n-p} (\mathbf{y} - \mathbf{Xb})^T (\mathbf{y} - \mathbf{Xb}) = \\ &= \frac{1}{n-p} S(\mathbf{b}) = \frac{1}{n-p} SSE \end{aligned} \quad (6.25)$$

where SSE is the sum of squared residuals. The divisor  $n - p$  is the degrees of freedom of the estimator  $s^2$ , which is an unbiased estimator of  $\sigma_\varepsilon^2$ .

### 6.1.2 Maximum likelihood estimation

Let  $\mathbf{x}$  be a vector of random variables with joint probability density function  $f(\mathbf{x})$ , the parameters of which are denoted by  $\boldsymbol{\theta}$ ;  $f(\mathbf{x})$  can be considered from two different perspectives. In fact, on one hand  $f(\mathbf{x})$  can be considered as a probability density function and therefore  $\boldsymbol{\theta}$  is known, i.e.  $f(\mathbf{x}) = f(\mathbf{x} | \boldsymbol{\theta})$ . On the other hand the vector  $\mathbf{x}$  may be considered as containing known data and  $\boldsymbol{\theta}$  may be considered unknown. In this case  $f(\mathbf{x})$  is a function of this latter only, i.e.  $f(\mathbf{x}) = f(\boldsymbol{\theta} | \mathbf{x})$ , it is called likelihood and indicated as  $L(\boldsymbol{\theta} | \mathbf{x})$ . From the mathematical point of view the two representations refer to the same object  $f(\mathbf{x} | \boldsymbol{\theta}) \equiv L(\boldsymbol{\theta} | \mathbf{x})$  but they are used for the sake of clarity. The likelihood function is the base of the maximum likelihood estimation method, which takes as an estimator of  $\boldsymbol{\theta}$  the vector  $\tilde{\boldsymbol{\theta}}$  which maximizes the likelihood function. When response surfaces are concerned, the unknown parameters are the regression parameters  $\boldsymbol{\beta}$  and the variance of the error term  $\sigma_\varepsilon^2$ .

Assuming, that  $\boldsymbol{\varepsilon}$  is normally distributed it follows that

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{C}_{\varepsilon\varepsilon}) \rightarrow \mathbf{y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{C}_{YY}) \quad (6.26)$$

and therefore the likelihood function has the following form

$$L(\boldsymbol{\beta}, \sigma_\varepsilon^2 | \mathbf{y}) = \frac{\exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{Z}\boldsymbol{\beta})^T \mathbf{C}_{YY}^{-1} (\mathbf{y} - \mathbf{Z}\boldsymbol{\beta})\right]}{(2\pi)^{\frac{1}{2}N} |\mathbf{C}_{YY}|^{\frac{1}{2}}} \quad (6.27)$$

If  $L(\boldsymbol{\theta} | \mathbf{x})$  is sufficiently regular, as in Eq. (6.27),  $\tilde{\boldsymbol{\theta}}$  can be calculated by setting to zero the partial derivatives of the likelihood function with respect to the unknown parameters

$$\begin{aligned}\frac{\partial L(\boldsymbol{\beta}, \sigma_e^2 | \mathbf{y})}{\partial \boldsymbol{\beta}} &= 0 \\ \frac{\partial L(\boldsymbol{\beta}, \sigma_e^2 | \mathbf{y})}{\partial \sigma_e^2} &= 0 \\ \sigma_e^2 &> 0\end{aligned}\tag{6.28}$$

Maximizing the likelihood function  $L$  (or  $\log(L)$ ) is, in general, a difficult numerical problem. Setting to zero the first derivatives of  $\log(L)$  gives a non-linear system of equations. Moreover, the likelihood function contains the inverse of an  $n \times n$  matrix, which is computationally expensive especially if the number of observations is large. In addition, the maximization problem is a constrained problem. In the present study, the EM (Expectation-Maximization) algorithm has been used, alternating calculation of the conditional expected values and maximization of simplified likelihoods Searle et al., 1992.

### 6.1.3 Tests on the model

The usual test of significance of the fitted regression equation is a test of the null hypothesis  $H_0$ : “all values of  $\beta_i$  ( $i \neq 0$ ) are zero”. The alternative hypothesis is  $H_a$ : “at least one value of  $\beta_i$   $i \neq 0$  is not zero”. Assuming normality of the errors, the test of  $H_0$  involves first calculating the value of the  $F$ -statistics:

$$F = \frac{SSR/(p-1)}{SSE/(N-p)}\tag{6.29}$$

where the sum of squares due to regression (SSR) is

$$SSR = \sum_{u=1}^N (\hat{y}(\mathbf{x}_u) - \bar{y})^2\tag{6.30}$$

where  $\bar{y}$  is the overall average, and the sum of squares accounted for by the fitted model (SSE) is

$$SSE = \sum_{u=1}^N (y_u - \hat{y}(\mathbf{x}_u))^2\tag{6.31}$$

If the null hypothesis is true the  $F$ -statistics in Eq. (6.29) follows a  $F$  distribution with  $p-1$  and  $N-p$  degrees of freedom in the numerator and in the denominator,

respectively. Therefore the test is performed by comparing the value given by Eq. (6.29) with the upper  $100\alpha$  percent point of the  $F$  distribution with  $p-1$  and  $N-p$  degrees of freedom. If this value is exceeded then the null hypothesis is rejected at the  $\alpha$  level of significance.

### 6.1.3.1 Tests concerning the individual terms of the model

The  $F$ -statistics in Eq. (6.29) is a test of the hypothesis that all of the parameters (excluding  $\beta_0$ ) in the proposed model are 0. In general, tests of hypotheses concerning the individual parameters in the proposed model are performed by comparing the parameter estimates in the fitted model to their respective estimated standard errors. Denoting the least squares estimate of  $\beta_i$  with  $b_i$  and the estimated standard error of  $b_i$  by “est. s.e. ( $b_i$ )”, then a test of the null hypothesis  $H_0: \beta_i = 0$ , is performed by calculating the value of the following test statistic

$$t = \frac{b_i}{\text{est. s.e.}(b_i)} \quad (6.32)$$

and comparing the so obtained value with a table value from the  $t$ -table. The choice of this latter value depends upon the alternative hypothesis,  $H_a$ , the level of significance  $\alpha$ , and the degrees of freedom for  $t$ . If the alternative hypothesis is  $H_a: \beta_i \neq 0$  the test is called two-sided test, and the table value is  $t_{\alpha/2}$ , i.e. the upper  $100(\alpha/2)$  percentage point of the  $t$ -distribution. If, on the other hand, the alternative hypothesis is  $H_a: \beta_i < 0$  or  $H_a: \beta_i > 0$ , the test is a one-sided test, and the table value is  $t_\alpha$ . Finally it should be noted that when the model contains more than one unknown parameter, say  $\beta_i$  and  $\beta_j$  (in addition to  $\beta_0$ ) and the estimates of the coefficients of the terms are not uncorrelated, then the tests of  $H_0: \beta_i = 0$  and  $H_0: \beta_j = 0$  are not independent. Consequently when the null hypothesis is expressed as  $H_0: \beta_i = 0$  versus  $H_a: \beta_i \neq 0$ , what is actually being tested is the hypothesis that the term  $\beta_i X_i$  does not explain any additional amount of variation in the response values above that which is explained by the other terms in the model.

## 6.2 Extension to time dependent problems

The standard formulation of RS cannot be used for applications in earthquake engineering because not all the random variables on which the structural behaviour depends can be expressed in explicit form as reported in Eq. (6.10). In fact, the evaluation of the seismic fragility requires to take the variability of the seismic action

and mechanical parameters over the structure into account. As for the first aspect, the earthquake ground-motion is a non-stationary process, with amplitude and frequency content variable in time, therefore a fully probabilistic description of it would require an enormous amount of variables. Moreover, the material properties can be very inhomogeneous among different structural members, due to building construction phases or degradation. Mathematical description of this aspect, would also require many random variables. The response surface method is a good alternative to Monte Carlo methods (in terms of reduced computational effort) only if the number of variables is low (6-8 variables), because the number of simulations required to calibrate the RS (called *the model*) strongly increases with the number of variables (Franchin et al., 2003a; Schotanus et al., 2004). In order to reduce the number of random variables in the response surface, they are divided in two groups: explicit  $\mathbf{x}_E$  and implicit  $\mathbf{x}_I$  variables (Veneziano et al., 1983; Faravelli, 1989; Casciati and Faravelli, 1991). The effects of the first ones are accounted for explicitly, while the effects of the latter are considered implicitly through random factors. The formulation of the model given in the previous sections needs to be extended, in order to take into account the differences between fixed and random factors. This type of model is named *mixed model*. In the following description of the model the notation used by Khuri and Cornell, 1996 and Searle et al., 1992 has been employed.

The model including fixed and random factors can be written as:

$$E(y(\mathbf{x}_1, \mathbf{x}_2) | \delta) = \mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta} + \delta \quad (6.33)$$

introducing the error as previously done

$$\varepsilon = y - E(y | \delta) = y - (\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta} + \delta) \quad (6.34)$$

it is possible to write:

$$y(\mathbf{x}_1, \mathbf{x}_2) = \mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta} + \delta + \varepsilon \quad (6.35)$$

where  $\mathbf{f}(\mathbf{x}_1)^T \boldsymbol{\beta}$  represents the fixed effects, which are treated explicitly, and  $\delta$  is the random factor describing implicitly the effects of the variables  $\mathbf{x}_I$ , on the response. In this type of model the distribution of the random factor  $\delta$  has to be determined together with the regression parameters. The random factor is assumed to be a random variable with null mean value with no lack of generality, in fact the mean effect of the random factor is included in the constant term of the fixed effects portion of the model. A

second, and stronger hypothesis, is that  $\delta$  has a Gaussian distribution with variance  $\sigma_{\delta}^2$  which becomes the only additional parameter to be estimated. Statistically the random effects can be thought as being related to an infinite set of levels of which only a random sample is present the data. In this kind of problems there are two different sources of variability, one related to the random factors and one related to the error  $\varepsilon$ . These terms are called variance components and their sum is the total variability of the observed quantity.

### 6.2.1 Random effects and blocks

The partition of the experiments into blocks (a block is simply a group of experiments) is a procedure originally developed for conditions when they could not be performed under homogeneous conditions. These non homogeneous conditions can be uncontrollable, for example if materials produced by different factories are to be tested together. As an alternative, experiments can be deliberately performed under non homogeneous conditions in order to evaluate their effects on the response. This is what it is done in the present work: experiment blocking is used to include the uncertainty on the possible ground motions into the model.

When the experiments are blocked the assumption of independence of the errors must be relaxed. Furthermore in the following it is assumed that a random factor can influence the response only additively. Partitioning  $n$  total experiments into  $b$  block of  $n_i$  elements each:

$$\sum_{i=1}^b n_i = n \quad (6.36)$$

the response surface mixed model can be written as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \quad (6.37)$$

where  $\boldsymbol{\delta}$  is the vector of the random effects and  $\mathbf{Z}$  is a  $n \times b$  matrix which assigns each random effect realization to its corresponding block. As an example if 3 blocks with 2 experiments each are considered the matrix  $\mathbf{Z}$  becomes

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.38)$$

In the most general case matrix  $\mathbf{Z}$  can be written as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{1}_{n_1} & \mathbf{0}_{n_1} & \cdots & \mathbf{0}_{n_1} \\ \mathbf{0}_{n_2} & \mathbf{1}_{n_2} & \cdots & \mathbf{0}_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{n_b} & \mathbf{0}_{n_b} & \cdots & \mathbf{1}_{n_b} \end{bmatrix} \quad (6.39)$$

where  $\mathbf{1}_{n_i}$  and  $\mathbf{0}_{n_i}$  represent  $1 \times n_i$  vectors containing ones and zeros respectively. In the following sections details will be given for models for which the blocking is controlled by one or more random effects.

### 6.2.2 One-way classification

Recalling the general form of the model given in Eq. (6.37) which was obtained under the assumptions

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} \quad (6.40)$$

$$E(\mathbf{y} | \boldsymbol{\delta}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} \quad (6.41)$$

$$\boldsymbol{\varepsilon} = \mathbf{y} - E(\mathbf{y} | \boldsymbol{\delta}) \quad (6.42)$$

and assuming that the random factor is normally distributed,  $\boldsymbol{\delta} \sim N(\mathbf{0}, \sigma_{\delta}^2 \mathbf{I}_b)$ , independently from  $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma_{\varepsilon}^2 \mathbf{I}_n)$ , i.e. assuming that  $\text{cov}(\boldsymbol{\delta}, \boldsymbol{\varepsilon}) = \mathbf{0}$ , the following form of the covariance matrix of  $\mathbf{y}$  is obtained:

$$\begin{aligned} \mathbf{C}_{YY} &= E\left[(\mathbf{y} - E(\mathbf{y}))(\mathbf{y} - E(\mathbf{y}))^T\right] = E\left[(\mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon})(\mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon})^T\right] = \\ &= \sigma_{\delta}^2 \mathbf{Z}\mathbf{Z}^T + \sigma_{\varepsilon}^2 \mathbf{I}_n = \sigma_{\varepsilon}^2 \mathbf{A} \end{aligned} \quad (6.43)$$

where  $\mathbf{A}$  is a block diagonal matrix,  $\mathbf{A} = \text{diag}[\mathbf{A}_1 \dots \mathbf{A}_b]$ , with

$$\mathbf{A}_i = \gamma \mathbf{J}_{n_i} + \mathbf{I}_{n_i}, \quad (6.44)$$

where  $\mathbf{J}_{n_i}$  is a  $n_i \times n_i$  matrix of ones and  $\gamma$  is the ratio of the variances,  $\gamma = \sigma_{\delta}^2 / \sigma_{\varepsilon}^2$ . This representation of the variance structure is named Hartley-Rao representation. As an example if 6 experiments are partitioned into 3 blocks of size 2 Eq. (6.43) becomes

$$\begin{aligned} \mathbf{C}_{YY} &= \sigma_{\delta}^2 \mathbf{Z}\mathbf{Z}^T + \sigma_{\varepsilon}^2 \mathbf{I}_n = \\ &= \sigma_{\delta}^2 \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} + \sigma_{\varepsilon}^2 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \\ &= \sigma_{\varepsilon}^2 \begin{bmatrix} 1+\gamma & \gamma & 0 & 0 & 0 & 0 \\ \gamma & 1+\gamma & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1+\gamma & \gamma & 0 & 0 \\ 0 & 0 & \gamma & 1+\gamma & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1+\gamma & \gamma \\ 0 & 0 & 0 & 0 & \gamma & 1+\gamma \end{bmatrix} = \\ &= \bar{\sigma}_{\varepsilon}^2 \begin{bmatrix} 1 & \rho & 0 & 0 & 0 & 0 \\ \rho & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & \rho & 0 & 0 \\ 0 & 0 & \rho & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & 1 \end{bmatrix} \end{aligned} \quad (6.45)$$

where  $\bar{\sigma}_{\varepsilon}^2 = (1+\gamma)\sigma_{\varepsilon}^2$  and  $\rho = \gamma / (1+\gamma)$ . From Eq. (6.45) it is evident how blocking of experiments changes the variance structure and introduces a correlation among the blocks. This correlation is constant after the assumptions made on the distributions of  $\delta$  and  $\varepsilon$ .

### 6.2.2.1 Parameter estimation

If the variance ratio  $\gamma$  is known, values of the unknown parameters of the model can easily be obtained via generalized least squares. In fact in this case the likelihood function can be written as:

$$L(\boldsymbol{\beta}, \sigma_\varepsilon^2 | \mathbf{y}, \gamma) = L(\boldsymbol{\beta}, \sigma_\varepsilon^2 | \mathbf{y}, \sigma_\delta^2) = \frac{\exp\left[-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{C}_{YY}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right]}{(2\pi)^{\frac{1}{2}N} |\mathbf{C}_{YY}|^{\frac{1}{2}}} \quad (6.46)$$

Eq. (6.46) clearly shows that the maximization of  $L$  corresponds to the maximization of the argument of the exponential function. This lead to the following generalized least squares estimate of the regression parameters:

$$\mathbf{b} = (\mathbf{X}^T \mathbf{C}_{YY} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{C}_{YY} \mathbf{y} \quad (6.47)$$

Eq. (6.47) corresponds to Eq. (6.17) when errors are not correlated, i.e. when  $\mathbf{A} = \mathbf{I}$ . Normally the variance ratio  $\gamma$  is not known and has to be estimated, in this case least square method can no longer be used, contrariwise maximum likelihood method can still be applied and the likelihood function can be written as follows:

$$L(\boldsymbol{\beta}, \sigma_\varepsilon^2, \gamma | \mathbf{y}) = L(\boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\delta^2 | \mathbf{y}) \quad (6.48)$$

Once again Eq. (6.48) can be maximized by setting equal to zero its partial derivatives

$$\begin{aligned} \frac{\partial L(\boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\delta^2 | \mathbf{y})}{\partial \boldsymbol{\beta}} &= 0 \\ \frac{\partial L(\boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\delta^2 | \mathbf{y})}{\partial \sigma_\varepsilon^2} &= 0 \\ \frac{\partial L(\boldsymbol{\beta}, \sigma_\varepsilon^2, \sigma_\delta^2 | \mathbf{y})}{\partial \sigma_\delta^2} &= 0 \end{aligned} \quad (6.49)$$

under the following constrains:

$$\begin{aligned}\sigma_\varepsilon^2 &> 0 \\ \sigma_\delta^2 &\geq 0\end{aligned}\tag{6.50}$$

Once the estimates of the regression parameters  $\mathbf{b}$  and of the variances  $\hat{\sigma}_\varepsilon^2$ ,  $\hat{\sigma}_\delta^2$  have been calculated the covariance matrix of  $\mathbf{b}$  can be estimated as

$$\text{var}(\mathbf{b}) = (\mathbf{X}^T \hat{\mathbf{A}}^{-1} \mathbf{X})^{-1} \hat{\sigma}_\varepsilon^2\tag{6.51}$$

and the predicted response and its variance as

$$\hat{y}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \mathbf{b}\tag{6.52}$$

$$\text{var}(\hat{y}) = \mathbf{f}(\mathbf{x})^T (\mathbf{X}^T \hat{\mathbf{A}}^{-1} \mathbf{X})^{-1} \mathbf{f}(\mathbf{x}) \hat{\sigma}_\varepsilon^2\tag{6.53}$$

### 6.2.3 Multiple-way classification

Random effects can be used to describe more than one source of variability. In the present work, for example, they are used to take in account not only the variability of the ground motions but also the variability of the concrete strength among different members of structures.

In this case, under assumptions similar to Eq. (6.40) -(6.42), the generic observation can be expressed as

$$y = \mathbf{f}^T \boldsymbol{\beta} + \sum_{i=1}^r \delta_i + \varepsilon\tag{6.54}$$

where  $r$  is the number of random factors. The results of  $n$  experiments can be collected into a vector and the model can be written in matrix form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}\tag{6.55}$$

Eq. (6.55) is formally similar to Eq. (6.37) but  $\boldsymbol{\delta}$  does not collect anymore the levels (the number of which is denoted by  $q_i$  in the following) of a single factor only, but contains the levels of all the factors  $\delta_i$  considered. Therefore  $\boldsymbol{\delta}$  can be thought as being made up of  $r$  different  $\delta_i$  vectors, each containing the  $q_i$  levels of a single random factor  $\delta_i$ , i.e.,

$$\boldsymbol{\delta} = [\boldsymbol{\delta}_1 \dots \boldsymbol{\delta}_r]^T\tag{6.56}$$

Matrix  $\mathbf{Z}$  is build by combining  $r$  matrices of size  $n \times q_i$  which relate each factor to the proper experiments

$$\mathbf{Z}\boldsymbol{\delta} = [\mathbf{Z}_1 \dots \mathbf{Z}_r] \begin{bmatrix} \boldsymbol{\delta}_1 \\ \vdots \\ \boldsymbol{\delta}_r \end{bmatrix} = \sum_{i=1}^r \mathbf{Z}_i \boldsymbol{\delta}_i \quad (6.57)$$

where  $\boldsymbol{\delta}_i$  contains the  $q_i$  levels of the factor  $i$ . As an example if 4 experiments and 2 random factor with 2 levels each are considered the matrix  $\mathbf{Z}$  and the vector  $\boldsymbol{\delta}$  can be written as

$$\mathbf{Z}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \mathbf{Z}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{Z} = \begin{bmatrix} 1 & 0 & | & 1 & 0 \\ 1 & 0 & | & 0 & 1 \\ 0 & 1 & | & 1 & 0 \\ 0 & 1 & | & 0 & 1 \end{bmatrix} \quad (6.58)$$

$$\boldsymbol{\delta} = [\boldsymbol{\delta}_{1,1} \quad \boldsymbol{\delta}_{1,2} \quad \boldsymbol{\delta}_{2,1} \quad \boldsymbol{\delta}_{2,2}]^T \quad (6.59)$$

where  $\boldsymbol{\delta}_{i,j}$  denotes the level  $j$  of the factor  $i$ .

As done for the simpler models described in the previous sections it assumed that  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ , that  $\boldsymbol{\varepsilon}$  and  $\boldsymbol{\delta}_i$  are mutually independent and with constant variance

$$\begin{aligned} E(\boldsymbol{\delta}_i) &= \mathbf{0} \\ \text{var}(\boldsymbol{\varepsilon}) &= \sigma_{\boldsymbol{\varepsilon}}^2 \mathbf{I}_n \\ \text{var}(\boldsymbol{\delta}_i) &= \sigma_{\boldsymbol{\delta}_i}^2 \mathbf{I}_{q_i} \\ \text{cov}(\boldsymbol{\delta}_i, \boldsymbol{\varepsilon}) &= \mathbf{0} \\ \text{cov}(\boldsymbol{\delta}_i, \boldsymbol{\delta}_j) &= \mathbf{0} \end{aligned} \quad (6.60)$$

From Eq. (6.60) it is possible to derive the following variance covariance matrices for  $\boldsymbol{\delta}$

$$\mathbf{C}_{\boldsymbol{\delta}\boldsymbol{\delta}} = \text{var}(\boldsymbol{\delta}) = \begin{bmatrix} \sigma_{\boldsymbol{\delta}_1}^2 \mathbf{I}_{q_1} & & & \\ & \sigma_{\boldsymbol{\delta}_2}^2 \mathbf{I}_{q_2} & & \\ & & \ddots & \\ & & & \sigma_{\boldsymbol{\delta}_r}^2 \mathbf{I}_{q_r} \end{bmatrix} \quad (6.61)$$

and for  $\mathbf{Y}$ :

$$\mathbf{C}_{YY} = \text{var}(\mathbf{y}) = \mathbf{Z}\mathbf{C}_{\delta\delta}\mathbf{Z}^T + \sigma_\varepsilon^2\mathbf{I}_n = \sum_{i=1}^r \sigma_{\delta_i}^2 \mathbf{Z}_i \mathbf{Z}_i^T + \sigma_\varepsilon^2 \mathbf{I}_n \quad (6.62)$$

which for the example for which Eq. (6.58) and Eq. (6.59) were derived becomes

$$\mathbf{C}_{YY} = \begin{bmatrix} \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & \sigma_{\delta_1}^2 & \sigma_{\delta_2}^2 & 0 \\ \sigma_{\delta_1}^2 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & 0 & \sigma_{\delta_2}^2 \\ \sigma_{\delta_2}^2 & 0 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 & \sigma_{\delta_1}^2 \\ 0 & \sigma_{\delta_2}^2 & \sigma_{\delta_1}^2 & \sigma_{\delta_1}^2 + \sigma_{\delta_2}^2 + \sigma_\varepsilon^2 \end{bmatrix} \quad (6.63)$$

Comparing Eq. (6.63) and Eq. (6.45) one can clearly notice that the correlation is now present not only inside the blocks but also among different blocks.

In order to derive a more synthetic form for the equations of the model it is convenient to include  $\varepsilon$  into the vector  $\boldsymbol{\delta}$  by setting:

$$\boldsymbol{\delta}_0 = \boldsymbol{\varepsilon} \quad q_0 = n \quad \mathbf{Z}_0 = \mathbf{I}_n \quad (6.64)$$

From Eq. (6.64) it follows that Eq. (6.55) can be rewritten as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \sum_{i=0}^r \mathbf{Z}_i \boldsymbol{\delta}_i \quad (6.65)$$

and Eq. (6.62) as

$$\mathbf{C}_{YY} = \sum_{i=0}^r \mathbf{Z}_i \mathbf{Z}_i^T \sigma_{\delta_i}^2 \quad (6.66)$$

Also in this case it is possible to adopt the Hartley-Rao formulation which is similar to that shown in the previous section:

$$\begin{aligned} \mathbf{C}_{YY} &= \sigma_\varepsilon^2 \mathbf{A} \\ \mathbf{A} &= \sum_{i=1}^r \gamma_i \mathbf{Z}_i \mathbf{Z}_i^T + \mathbf{I}_n \\ \gamma_i &= \frac{\sigma_{\delta_i}^2}{\sigma_\varepsilon^2} \end{aligned} \quad (6.67)$$

Regression parameter values and variances can be estimated using the maximum likelihood method which, in this case, is formally similar to what was described in section 6.2.3. The variance matrix of  $\mathbf{b}$ , the predicted response  $\hat{y}(\mathbf{x})$  and its variance can be calculated using expressions similar to those in Eq. (6.51), (6.52) and (6.53) the only difference being the matrix  $\mathbf{A}$  which has to be derived from Eq. (6.67) instead of Eq. (6.44).

## 6.2.4 Tests on the model

### 6.2.4.1 Tests concerning fixed effects

Tests concerning the elements of  $\boldsymbol{\beta}$  in the model of Eq. (6.55) are carried out using approximate  $t$  statistics. For example, a test statistics for testing the null hypothesis  $H_0 : \beta_i = 0$  versus  $H_a : \beta_i \neq 0$  is given by

$$t_i = \frac{b_i}{s_i} \quad (6.68)$$

where  $b_i$  is the  $i$ -th element of  $\mathbf{b}$ , the estimator of  $\boldsymbol{\beta}$ , and  $s_i$  is the square root of the corresponding diagonal element of  $\mathbf{C}_{bb}$ . Under  $H_0$ ,  $t_i$  has approximately a  $t$ -distribution. Further details can be found in Khuri and Cornell, 1996 and Khuri, 1992.

### 6.2.4.2 Tests concerning the block effect

A test concerning the block effect can be performed considering the Type III sum of squares for  $\boldsymbol{\delta}$  in the model of Eq. (6.55). The following hypothesis concerning the random block effect is considered:  $H_0 : \sigma_{\delta}^2 = 0$  versus  $H_a : \sigma_{\delta}^2 \neq 0$ . The Type III sum of squares,  $R(\boldsymbol{\delta} | \boldsymbol{\beta})$ , is equal to Searle et al., 1992

$$R(\boldsymbol{\delta} | \boldsymbol{\beta}) = \mathbf{y}^T \left( \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \right) \mathbf{y} \quad (6.69)$$

where  $\mathbf{W} = [\mathbf{X} : \mathbf{Z}]$ . The expected value of  $R(\boldsymbol{\delta} | \boldsymbol{\beta})$  is given by Searle et al., 1992

$$E(R(\boldsymbol{\delta} | \boldsymbol{\beta})) = \sigma_{\delta}^2 \text{tr} \left( \mathbf{Z}^T \left( \mathbf{I}_N - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X} \right) \mathbf{Z} \right) + (b-1) \sigma_{\varepsilon}^2 \quad (6.70)$$

it follows that a test statistics for testing the null hypothesis is given by the ratio

$$F = \frac{R(\boldsymbol{\delta} | \boldsymbol{\beta})}{(b-1)MS_E} \quad (6.71)$$

where  $MS_E$  is the residual mean square defined as

$$MS_E = \frac{\mathbf{y}^T \left( \mathbf{I}_N - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \right) \mathbf{y}}{n-p-b} \quad (6.72)$$

where  $b$  is the number of blocks. Under the null hypothesis the statistic in Eq. (6.71) has an  $F$ -distribution with  $b-1$  and  $n-p-b$  degrees of freedom. Further details can be found in Khuri and Cornell, 1996 and Khuri, 1992.

### 6.2.5 Prediction of random effects

In this section some brief details will be given on the problem of predicting the unobservable realized values,  $\tilde{\boldsymbol{\delta}}$ , of random effects that are part of a mixed model. The generic element of  $\tilde{\boldsymbol{\delta}}$ ,  $\tilde{\delta}_i$  can be considered as the realized value, in a block of experiments, of a random variable. The vector  $\tilde{\boldsymbol{\delta}}$  is called prediction and not estimate because estimation concerns parameters only.

The general problem can be stated as follows: suppose  $\boldsymbol{\Lambda}$  and  $\mathbf{Y}$  are jointly distributed vectors of random variables, with those in  $\mathbf{Y}$  being observable but those in  $\boldsymbol{\Lambda}$  not being observable. The problem is to predict  $\boldsymbol{\Lambda}$  from some realized observed value of  $\mathbf{Y}$ . Usually  $\mathbf{Y}$  contains more elements than  $\boldsymbol{\Lambda}$  and this latter is often scalar.

A criterion that can be used to find a prediction is to look for a predictor that minimizes the mean square error. When  $f(\boldsymbol{\delta}, \mathbf{y})$  is the joint density function of the random variables  $\boldsymbol{\Lambda}$  and  $\mathbf{Y}$  the mean square error of prediction is

$$E(\tilde{\boldsymbol{\delta}} - \boldsymbol{\delta})^T \mathbf{S}(\tilde{\boldsymbol{\delta}} - \boldsymbol{\delta}) = \iint (\tilde{\boldsymbol{\delta}} - \boldsymbol{\delta})^T \mathbf{S}(\tilde{\boldsymbol{\delta}} - \boldsymbol{\delta}) f(\boldsymbol{\delta}, \mathbf{Y}) d\mathbf{y} d\mathbf{u} \quad (6.73)$$

where  $\mathbf{S}$  is any positive symmetric matrix. From minimizing Eq. (6.73) it is possible to derive the best predictor (best because it is derived using a minimum least square criterion) of  $\tilde{\boldsymbol{\delta}}$  which can be expressed as:

$$\tilde{\boldsymbol{\delta}} = E(\boldsymbol{\delta} | \mathbf{y}) \quad (6.74)$$

this results holds for all probability density functions  $f(\boldsymbol{\delta}, \mathbf{y})$  and does not depend on the positive symmetric matrix  $\mathbf{S}$ . Furthermore the best predictor is not biased in fact

$$E_{\mathbf{y}}(\tilde{\boldsymbol{\delta}}) = E_{\mathbf{y}}[E_{\delta|\mathbf{y}}(\boldsymbol{\delta}|\mathbf{y})] = E(\boldsymbol{\delta}) \quad (6.75)$$

Noteworthy  $\tilde{\boldsymbol{\delta}} = E(\boldsymbol{\delta}|\mathbf{y})$  is a random variable, being a function of  $\mathbf{y}$  and unknown parameters. Thus the problem of estimating the best predictor  $\tilde{\boldsymbol{\delta}}$  remains, and demands some knowledge of the joint density  $f(\boldsymbol{\delta}, \mathbf{y})$ . When this is normal,

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{y} \end{bmatrix} = N\left(\begin{bmatrix} \boldsymbol{\mu}_{\delta} \\ \boldsymbol{\mu}_Y \end{bmatrix}, \begin{bmatrix} \mathbf{C}_{\delta\delta} & \mathbf{C}_{Y\delta} \\ \mathbf{C}'_{Y\delta} & \mathbf{C}_{YY} \end{bmatrix}\right) \quad (6.76)$$

which when combined with Eq. (6.80) gives

$$\tilde{\boldsymbol{\delta}} = \boldsymbol{\mu}_{\delta} + \mathbf{C}_{Y\delta}\mathbf{C}_{YY}^{-1}(\mathbf{y} - \boldsymbol{\mu}_Y) \quad (6.77)$$

The estimation problem is clearly visible in these results. The predictor is given in Eq. (6.74) but it and its succeeding properties cannot be estimated without having values for, or estimating, the four parameters  $\boldsymbol{\mu}_{\delta}$ ,  $\boldsymbol{\mu}_Y$ ,  $\mathbf{C}_{Y\delta}$  and  $\mathbf{C}_{YY}$ .

The best predictor is not necessarily linear in  $\mathbf{y}$ . Supposing attention is now confined to predictors that are linear, they must have the form

$$\tilde{\boldsymbol{\delta}} = \mathbf{a} + \mathbf{b}\mathbf{y} \quad (6.78)$$

where  $\mathbf{a}$  is a vector and  $\mathbf{b}$  is a matrix. Minimizing Eq. (6.73), in order to obtain the best linear predictor, leads to

$$\tilde{\boldsymbol{\delta}} = \boldsymbol{\mu}_{\delta} + \mathbf{C}_{Y\delta}\mathbf{C}_{YY}^{-1}(\mathbf{y} - \boldsymbol{\mu}_Y) \quad (6.79)$$

where the parameters are defined as in(6.76) but without the assumption of normality.

An immediate observation on Eq. (6.79) is that it is identical to Eq. (6.77). This shows that the best linear predictor, derivation of which demands no knowledge of the form of  $f(\boldsymbol{\delta}, \mathbf{Y})$ , is identical to the best predictor under normality. Therefore this predictor is called best linear unbiased predictor (BLUP). Recalling the assumptions made previously

$$\begin{aligned} \mathbf{C}_{\delta\delta} &= \text{var}(\boldsymbol{\delta}) \\ \mathbf{C}_{YY} &= \text{var}(\mathbf{Y}) = \mathbf{Z}\mathbf{C}_{\delta\delta}\mathbf{Z}^T + \sigma_{\epsilon}^2\mathbf{I} \\ \mathbf{C}_{Y\delta} &= \text{cov}(\mathbf{Y}, \boldsymbol{\delta}) = \mathbf{C}_{\delta\delta}\mathbf{Z}^T \end{aligned} \quad (6.80)$$

and that  $\boldsymbol{\mu}_\delta = E(\boldsymbol{\delta}) = \mathbf{0}$  and,  $\boldsymbol{\mu}_Y = \mathbf{X}\boldsymbol{\beta}$  it follows that a prediction linear in  $\mathbf{y}$  and unbiased can be expressed as:

$$BLUP(\boldsymbol{\delta}) = \tilde{\boldsymbol{\delta}} = \mathbf{C}_{\delta\delta} \mathbf{Z}^T \mathbf{C}_{YY}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \quad (6.81)$$

## 6.3 Design of the simulation plans

### 6.3.1 Some considerations on the variables

In planning the program of experiments, the experimenter is faced with

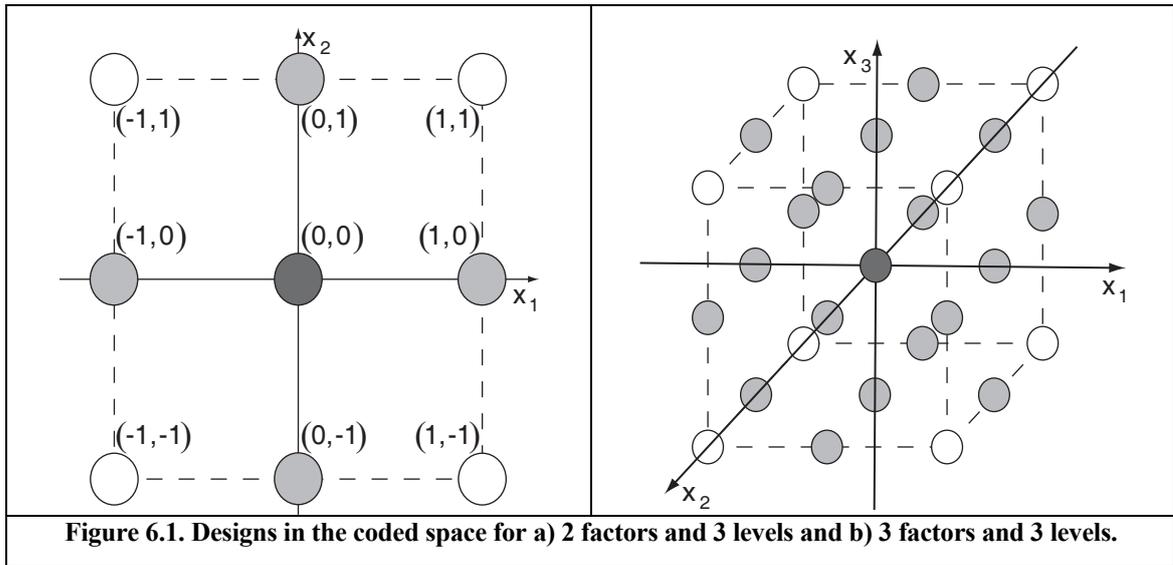
1. choosing the factors or input variables to be used in the experiment;
2. selecting the range of values and the number of levels of each factor in order to adequately measure the effects of the factors on the response.

Upon answering these questions the experimental region is defined.

The used of coded variables in place of the input variables facilitates the construction of the experimental designs. Coding removes the units of measurements of the input variables and consequently distance measured along the axes of the coded variables in a  $k$ -dimensional space are standardized. A convenient coding formula for defining the coded variables is

$$x_i = \frac{2X_i - (X_{iL} + X_{iH})}{X_{iH} - X_{iL}} \quad (6.82)$$

where  $X_{iL}$  and  $X_{iH}$  are the low and high levels of the factor  $X_i$ , respectively. If only two levels are considered for a factor, Eq. (6.82) gives the notation  $\pm 1$ . The region defined by those two levels is a cuboidal region in a  $k$ -dimensional space. Geometrically, the cuboidal region has  $2^k$  vertices where each vertex is defined by the coordinate  $\pm 1$  settings in  $x_1, x_2, \dots, x_k$ . When a factor has three levels and the mid level is in the middle between the lower and upper levels, the coding formula in Eq. (6.82) produces the coded levels  $x_i = -1, 0, +1$  associated with the low, middle and high values of  $X_i$ , respectively. When all factors have three levels, again the region in the coded variables is a  $k$ -dimensional cuboidal region. However, the number of factor combinations is now  $3^k$  and  $2^k$  of the combinations are the vertices of the  $k$ -dimensional cuboidal region, the remaining  $3^k - 2^k$  combinations define the centroids of all the lower dimensional boundaries of the  $k$ -dimensional cube along with the centroid,  $\mathbf{0}$ , of the cuboidal region. Figure 6.1 gives examples of designs with 3 levels for 2 and 3 factors respectively.



There are several advantages to using coded variables rather than the original input variables when fitting polynomial models. Two of the most obvious advantages are:

1. computational ease and increased accuracy in estimating the model coefficients;
2. enhanced interpretability of the coefficient estimates.

Both the advantages stem from the fact that the  $\mathbf{X}^T \mathbf{X}$  matrix in the coded variables is usually of a simpler form than the  $\mathbf{X}^T \mathbf{X}$  matrix in the original variables. The simpler form of the  $\mathbf{X}^T \mathbf{X}$  matrix, the easier it is to invert  $\mathbf{X}^T \mathbf{X}$ , and as a result, the greater will be the computational accuracy of the model parameter estimates and the reduction in computing time. Finally since the coding transformation is a one-to-one transformation, any linear polynomial equation in the values of  $x_i$  is expressible as (and equivalent to) a polynomial equation of the same degree in the values of  $X_i$ .

### 6.3.2 Properties of a response surface design

Prior to performing the actual experimentation, quite often it is necessary to decide whether the main emphasis is measuring the effects of the input variables or whether one is really more interested in modelling the response surface for prediction purposes. In the present section designs in which the primary emphasis is in modelling and exploring response surface over the experimental region are concerned.

Box and Wilson, 1951 listed some properties that a response surface design to be used when fitting a polynomial model should have:

1. the design should generate a satisfactory distribution of information throughout the region of interest;

2. the design should ensure that the fitted value at  $\mathbf{x}$ ,  $\hat{Y}(\mathbf{x})$ , be as close as possible to the true value at  $\mathbf{x}$ ,  $\eta(\mathbf{x})$ ;
3. the design should have good detectability of model lack of fit;
4. the design should allow experiments to be performed in blocks;
5. the design should allow designs of increasing order to be built up sequentially;
6. the design should provide an internal estimate of the error variance;
7. the design should require a minimum number of experimental points;
8. the design should ensure simplicity of calculation of the model parameter estimates.

This list of properties will serve as a reference in next sections. In addition to the properties listed above, there are times when the design is required to possess the property of orthogonally and/or the property of rotatability. An orthogonal design is one in which the terms in the fitted model are uncorrelated with one another and thus the parameter estimates are uncorrelated. In his case, the variance of the predicted response at any point  $\mathbf{x}$  in the experimental region, is expressible as a weighted sum of the variances of the parameter estimates in the model. A first order design (i.e. a design suited for fitting first order models), for example, is orthogonal if and only if the corresponding  $\mathbf{X}^T\mathbf{X}$  matrix is diagonal. If second order designs are concerned, it is not possible to obtain a diagonal  $\mathbf{X}^T\mathbf{X}$  matrix as a consequence of the second order terms, in this case different conditions must be satisfied which will be discussed in the following.

With a rotatable design, on the other hand, the variance of  $\hat{Y}(\mathbf{x})$ , which is known to depend on the location of the point  $\mathbf{x}$ , is a function only of the distance from the point  $\mathbf{x}$  to the centre of the design. Thus, with a rotatable design, the prediction variance,  $\text{var}[\hat{Y}(\mathbf{x})]$ , is the same at all points that are equidistant from the design centre. Consequently, in the space of the input variables, surfaces of constant prediction variance form concentric hyperspheres. One of the desirable features of rotatability is that the quality of prediction, as measured by the magnitude of  $\text{var}[\hat{Y}(\mathbf{x})]$ , is invariant to any rotation of the coordinate axes in the space of the input variables.

To better understand the rotatability and the orthogonality properties it is worth to concentrate on certain parameters of the distribution of design points called design moments. Supposing that the model to be fitted is expressed in matrix notation in the form of Eq. (6.10) and supposing that this model is of order  $d$  and is a function of  $k$  input variables (which are considered to have been coded), then a design moment of order  $\delta$  ( $\delta = 0, 1, \dots, 2d$ ), denoted by  $[1^{\delta_1} 2^{\delta_2} \dots k^{\delta_k}]$ , is equal to

$$[1^{\delta_1} 2^{\delta_2} \dots k^{\delta_k}] = \frac{1}{n} \sum_{i=1}^n x_{i1}^{\delta_1} \cdot x_{i2}^{\delta_2} \cdot \dots \cdot x_{ik}^{\delta_k} \quad (6.83)$$

where  $n$  is the total number of observations and  $\delta_1, \delta_2, \dots, \delta_k$  are non-negative integers such as  $\sum_{i=1}^k \delta_i = \delta$ . The design moments are the elements of the matrix  $\frac{1}{n} \mathbf{X}^T \mathbf{X}$  which is called the moment matrix. For example, if a first order model in the  $k$  coded variables is considered and

$$x_{ui} = \frac{X_{ui} - \bar{X}_i}{s_{x_i}}, \quad s_{x_i} = \left( \sum_{u=1}^n \frac{(X_{ui} - \bar{X}_i)^2}{n} \right)^{1/2} \quad \begin{matrix} u = 1, 2, \dots, n \\ i = 1, 2, \dots, k \end{matrix} \quad (6.84)$$

where the denominator,  $s_{x_i}$ , is a measure of the spread of the design points in the direction of the  $X_i$  axis, then the following moment matrix is obtained

$$\frac{1}{n} \mathbf{X}^T \mathbf{X} = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 & \dots & x_k \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{matrix} & \begin{bmatrix} 1 & [1] & [2] & [3] & \dots & [k] \\ [11] & [12] & [13] & \dots & [1k] \\ & [22] & [23] & \dots & [2k] \\ & & [33] & \dots & [3k] \\ & sim & & \ddots & \vdots \\ & & & & [kk] \end{bmatrix} \end{matrix} \quad (6.85)$$

with

$$[i] = \frac{1}{N} \sum_{u=1}^N x_{ui} \quad [ii] = \frac{1}{N} \sum_{u=1}^N x_{ui}^2 \quad [ij] = \frac{1}{N} \sum_{u=1}^N x_{ui} x_{uj} \quad i, j = 1, 2, \dots, k \quad (6.86)$$

The moment  $[i]$  is the average of the  $x_{ui}$  values over the  $n$  observations and is called moment of the first order. The moments  $[ii]$  and  $[ij]$  have similar meaning and are called pure second-order moment and mixed second-order moment, respectively. However, by the coding convention in Eq. (6.84) it follows that  $\sum_{u=1}^N x_{ui} = 0$  and  $\sum_{u=1}^N x_{ui}^2 = n$  and consequently  $[i] = 0$  and  $[ii] = 1$ . Thus the moment matrix in Eq. (6.85) is simplified to

$$\frac{1}{n} \mathbf{X}^T \mathbf{X} = \begin{matrix} & x_1 & x_2 & x_3 & \cdots & x_k \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & [12] & [13] & \cdots & [1k] \\ & & 1 & [23] & \cdots & [2k] \\ & & & 1 & \cdots & [3k] \\ & & & & \ddots & \vdots \\ & & & & & 1 \end{bmatrix} \end{matrix} \quad (6.87)$$

When the fitted model is of the second order, the moment matrix contains moments up to order four. With two variables, for example, the coding convention in Eq. (6.84) produces the moment matrix

$$\frac{1}{n} \mathbf{X}^T \mathbf{X} = \begin{matrix} & x_1 & x_2 & x_1^2 & x_2^2 & x_1 x_2 \\ \begin{matrix} x_1 \\ x_2 \\ x_1^2 \\ x_2^2 \\ x_1 x_2 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & [12] \\ 0 & 1 & [12] & [111] & [122] & [112] \\ & & 1 & [112] & [222] & [122] \\ & & & [1111] & [1122] & [1112] \\ & & & & [2222] & [1222] \\ & & & & & [1122] \end{bmatrix} \end{matrix} \quad (6.88)$$

where

$$\begin{aligned} [111] &= \frac{1}{n} \sum_{i=1}^N x_{i1}^3 & [122] &= \frac{1}{n} \sum_{i=1}^N x_{i1} x_{i2}^2 \\ [1111] &= \frac{1}{n} \sum_{i=1}^N x_{i1}^4 & [1122] &= \frac{1}{n} \sum_{i=1}^N x_{i1}^2 x_{i2}^2 \end{aligned} \quad (6.89)$$

Both orthogonality and rotatability conditions for second-order design can be defined in terms of the moment matrix.

An orthogonal-second order designs can be obtained if the variables in the model defined by Eq. (6.10) are expressed in terms of orthogonal polynomials as shown below. Let  $P_m(x_{ui})$  be the orthogonal polynomial of degree  $m$  ( $m \geq 0$ ) for the  $i$ -th input variable  $x_i$  ( $i = 1, 2, \dots, k$ ) then,

$$P_m(x_{ui}) = x_{ui}^m + \alpha_{m-1,m} x_{ui}^{m-1} + \dots + \alpha_{1m} x_{ui} + \alpha_{0m} \quad (6.90)$$

where the values of the  $\alpha$ s are chosen such as

$$\sum_{u=1}^N P_m(x_{ui}) P_{m-j}(x_{ui}) = 0 \quad P_0(x_{ui}) = 1 \quad j = 1, 2, \dots, m \quad (6.91)$$

The original second-order model can now be expressed in terms of these orthogonal polynomials as

$$\mathbf{Y} = (\mathbf{XP})(\mathbf{P}^{-1}\boldsymbol{\beta}) + \boldsymbol{\varepsilon} = \dot{\mathbf{X}}\dot{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \quad (6.92)$$

where  $\mathbf{P}$  is the non-singular matrix transforming the terms  $x_{ui}^m$  into  $P_m(x_{ui})$ ,  $\dot{\mathbf{X}} = \mathbf{XP}$  and  $\dot{\boldsymbol{\beta}} = \mathbf{P}^{-1}\boldsymbol{\beta}$ . Under the scaling convention given in Eq. (6.84) it is obtained that

$$P_1(x_{ui}) = x_{ui} \quad P_2(x_{ui}) = x_{ui}^2 - [iii]x_{ui} - 1 \quad i = 1, 2, \dots, k \quad (6.93)$$

and the response surface model (see Eq. (6.10)) can be written, in an explicit form, as

$$Y_u = \left( \beta_0 + \sum_{i=1}^k \beta_{ii} \right) + \sum_{i=1}^k (\beta_i + [iii]\beta_u) P_1(x_{ui}) + \sum_{i=1}^k \beta_{ii} P_2(x_{ui}) + \sum_{i=1}^{k-1} \sum_{\substack{j=2 \\ j>i}}^k \beta_{ij} P_1(x_{ui}) P_1(x_{uj}) + \boldsymbol{\varepsilon}_u \quad u = 1, 2, \dots, N \quad (6.94)$$

Since  $\sum_{u=1}^N P_m(x_{ui}) = 0$  for  $m = 1, 2$  and  $i = 1, 2, \dots, k$  and  $\sum_{u=1}^N P_1(x_{ui}) P_2(x_{ui}) = 0$  for  $i = 1, 2, \dots, k$  then the moment matrix  $\frac{1}{N} \dot{\mathbf{X}}^T \dot{\mathbf{X}}$  is diagonal if

$$\begin{aligned} [ii] &= [ijj] = [iij] = 0 \text{ per } i < j \\ [iij] &= 1 \text{ per } i < j \\ [iijn] &= [ijjn] = [ijn] = 0 \text{ per } i < j < n \\ [iij] &= [ijj] = 0 \text{ per } i < j \\ [ijns] &= [ijn] = 0 \text{ per } i < j < n < s \end{aligned} \quad (6.95)$$

A simple characterization of rotatability can, as well, be given in terms of the elements of the  $\mathbf{X}^T \mathbf{X}$  matrix, or the moment matrix. In fact a necessary and sufficient condition for a design to be rotatable is that the moment of order  $\delta$  be of the form

$$[1^{\delta_1} 2^{\delta_2} \dots k^{\delta_k}] = \begin{cases} 0 & \text{if any } \delta_i \text{ is odd} \\ \lambda_d \prod_{i=1}^k (\delta_i)! & \text{if all } \delta_i \text{ are even} \\ 2^{\delta/2} \prod_{i=1}^k \left( \frac{\delta_i}{2} \right)! & \end{cases} \quad (6.96)$$



It has been shown by Box and Hunter (1975) that with any rotatable second-order design the variance of the predicted response at any point  $\mathbf{x}$  in the experimental region is given by

$$\text{Var}(\hat{y}(\mathbf{x})) = A \left[ 2(k+2)\lambda_4^2 + 2\lambda_4(\lambda_4 - 1)(k+2)\rho^2 + ((k+1)\lambda_4 - (k-1))\rho^4 \right] \quad (6.99)$$

where  $\rho^2 = \mathbf{x}'\mathbf{x}$  and  $A = [2N\lambda_4((k+2)\lambda_4 - k)]^{-1}$ . For the particular case of an orthogonal second-order rotatable design, i.e., when  $\lambda_4 = 1$ , Eq. (6.99) becomes

$$\text{Var}(\hat{y}(\mathbf{x})) = \frac{\sigma^2(k+2+\rho^4)}{2N} \quad (6.100)$$

from which it follows that the variance of  $\hat{y}(\mathbf{x})$  at the design centre, i.e. when  $\rho = 0$ , is

$$\text{Var}(\hat{y}(\mathbf{0})) = \sigma^2 \lambda_4 \left( N \left( \lambda_4 - \frac{k}{k+2} \right) \right)^{-1} \quad (6.101)$$

It can be seen from Eq. (6.101) that  $\text{Var}(\hat{y}(\mathbf{0}))$  is a decreasing function of  $\lambda_4$  for  $\lambda_4 > k/(k+2)$ . As  $\lambda_4$  approaches or exceeds unity, the precision at the center of the design increases. Finally it is worth noting that when  $\lambda_4$  approaches  $k/(k+2)$ , the values of  $A$  in Eq. (6.99) becomes infinitely large, which renders the rotatable second-order design useless. This occurs then all the design points of a second-order design are equidistant from the design centre.

### 6.3.3 The central composite design

Box and Wilson ( ) introduced the class of central composite designs (CCD) for problems involving response surfaces. A central composite design consist of:

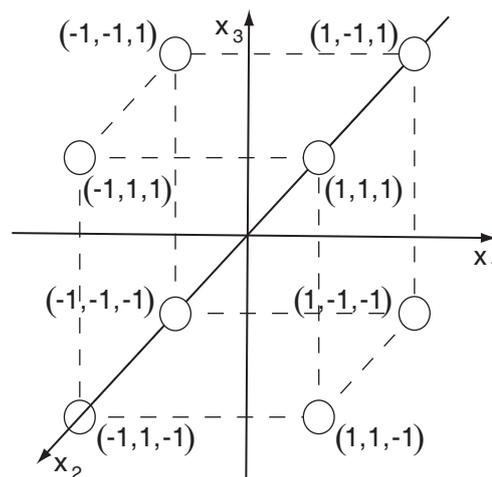
- a complete (or fraction of a)  $2^k$  factorial design, where the factor levels are coded to the usual -1, +2 values. This is called factorial protion of hte design;
- $n_0$  centre points;
- two axial points on the axix of each design variable at a distance of  $\alpha$  from the design centre. This is called axial portion of the design.

A  $2^k$  factorial design consists of all the  $2^k$  points with levels  $(x_1, x_2, \dots, x_k) = (\pm 1, \pm 1, \dots, \pm 1)$  where every possible combination of + and - signs is selected in turn. To better understand the features of the design it is convenient to list the runs in standard order. This standard order is obtained by writing alternate - and + signs in the column corresponding to the first factor, alternate pairs --, ++ in the second

column, alternate fours ----, ++++ in the third column and so on. Table 6.1 shows an example of runs set out in standard order.

**Table 6.1.  $2^3$  factorial design sorted in standard order:  $X_1, X_2, X_3$  are actual factor levels,  $x_1, x_2, x_3$  are coded levels and  $y$  is the observed response.**

Actual values			Coded values			$y$
$X_1$	$X_2$	$X_3$	$x_1$	$x_2$	$x_3$	
400	15	35	-1	-1	-1	2.83
460	15	35	1	-1	-1	3.56
400	25	35	-1	1	-1	2.23
460	25	35	1	1	-1	3.06
400	15	45	-1	-1	1	2.47
460	15	45	1	-1	1	3.30
400	25	45	-1	1	1	1.95
460	25	45	1	1	1	2.56



**Figure 6.2  $2^3$  Factorial design.**

Assuming that experiments defined according to a complete factorial design are run the parameters of the model can be estimated either by using Eq. (6.17) or by a different procedure, described in the following, which allows to better understand the features of factorial designs. First of all some definitions must be given. The main effect of a given variable is the average difference in the level of response as one moves from the low to the high level of that variable. As an example, if the data in Table 6.1 is considered, the main effect, indicated by 1, of the variable  $x_1$  is

$$1 \leftarrow \frac{1}{4}(3.56 + 3.06 + 3.30 + 2.56) - \frac{1}{4}(2.83 + 2.23 + 2.47 + 1.95) = 0.75 \quad (6.102)$$

where the symbol  $\leftarrow$  is read as “is estimated by”. A valuable property of the factorial design is that it makes possible not only the calculation of main effects (average effects) but also of interaction effects between variables as well. Two variables, say  $x_1$  and  $x_3$ , are said to interact, in their effect on the response, if the effect of 1 is different at the two different levels of 3. In the previous example if the first four runs, in which  $x_3$  is at its lower level, the main effect of 1 is

$$(1|x_3 = -1) \leftarrow \frac{1}{2}(3.56 + 3.06) - \frac{1}{2}(2.83 + 2.23) = 0.78 \quad (6.103)$$

For the last four runs, with  $x_3$  at its upper level, the main effect of 1 is

$$(1|x_3 = 1) \leftarrow \frac{1}{2}(3.30 + 2.56) - \frac{1}{2}(2.47 + 1.95) = 0.72 \quad (6.104)$$

The interaction between variable 1 and 3 is defined as half the difference between the main effect of 1 at the upper level of  $x_3$  and the main effect of 1 at the lower level of  $x_3$ . This interaction is denoted by the symbol 13, so that

$$13 \leftarrow \frac{1}{2}[(1|x_3 = 1) - (1|x_3 = -1)] \leftarrow \frac{1}{2}[0.72 - 0.78] = -0.03 \quad (6.105)$$

It is worthily noticed that interchanging the roles of variables 1 and 3 does not change the value of the interaction. Performing analogous calculations the other two-factor interactions 12 and 23 can be obtained. Those two-factor interactions may be different at different levels of the variable not involved in their definition. For example 12 may be different at different levels of the variable 3:

$$(12|x_3 = -1) \leftarrow \frac{1}{2}[(3.06 - 2.23) - (3.56 - 2.83)] = 0.05 \quad (6.106)$$

$$(12|x_3 = 1) \leftarrow \frac{1}{2}[(2.56 - 1.95) - (3.30 - 2.47)] = -0.1$$

Half the difference between these quantities is called the 123 interaction

$$123 \leftarrow \frac{1}{2}[(12|x_3 = 1) - (12|x_3 = -1)] \leftarrow -0.08 \quad (6.107)$$

**Table 6.2. Table of sign for the design given in Table 6.1.**

	<b>I</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>12</b>	<b>13</b>	<b>23</b>	<b>123</b>	<b>y</b>
	+	-	-	-	+	+	+	-	2.83
	+	+	-	-	-	-	+	+	3.56
	+	-	+	-	-	+	-	+	2.23
	+	+	+	-	+	-	-	-	3.06
	+	-	-	+	+	-	-	+	2.47
	+	+	-	+	-	+	-	-	3.30
	+	-	+	+	-	-	+	-	1.95
	+	+	+	+	+	+	+	+	2.56
Divisor	8	4	4	4	4	4	4	4	

There is an easy systematic way of making the previous calculations using the columns of signs in

Table 6.2. Considering, for example, the calculation of the main effect 1 as in Eq. (6.102), it could be written as

$$1 \leftarrow \frac{1}{4}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8) \quad (6.108)$$

In other words, the main effect 1 would be obtained by multiplying the column of data  $y$  by the column of signs in the column labelled **1** and dividing by the divisor 4 indicated here. The indicated here. The divisors are the number of + signs in the corresponding columns. The two- and three- The two- and three-factor interactions can be obtained performing similar calculations. A table such as A table such as

Table 6.2 is easily constructed for any  $2^k$  level factorial design as follows. The first column is a column of ones of length  $2^k$ . The following  $k$  columns are + and – signs for design written down in standard order. The  $2^k - k - 1$  columns **12**, **13**, ..., **123**, ..., **123...k**, are then obtained by multiplying signs, row by row, in the way indicated by the headings.

A factorial design does not suffice in order to estimate all the parameters of a complete polynomial response surface of second degree. For this reason, in CCD, centre points and axial points are added to the factorial design. Figure 6.3. shows the points added to the  $2^k$  factorial design.

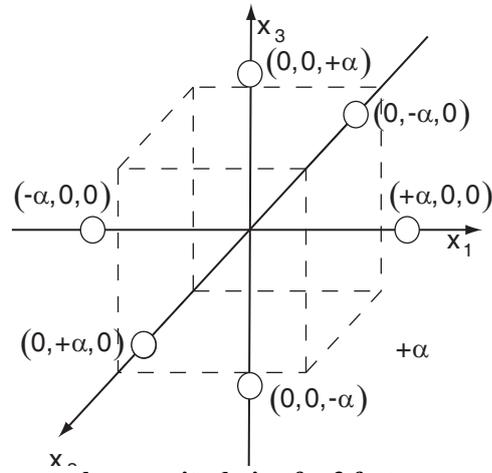


Figure 6.3. Axial portion of a central composite design for 3 factors.

The total number of experiments to run is therefore  $N = 2^k + 2k + n_0$ . The values of  $\alpha$  and  $n_0$  are chosen in order to give the CCD some useful properties. The criteria to do that will be described in the following. As an example a design matrix in 2 variables with  $n_0 = 1$  and  $\alpha = \sqrt{2}$  has the form

$$X = \begin{matrix} & x_1 & x_2 \\ \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \\ \sqrt{2} & 0 \\ -\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & -\sqrt{2} \\ 0 & 0 \end{bmatrix} \end{matrix}$$

Adopting the scaling convention given by Eq. (6.84) it can be easily verified that the odd moments up to order 4 are 0 for a CCD, i.e.

$$\begin{aligned}
 [i] &= 0 & i = 1, 2, \dots, k \\
 [iii] &= 0 & i = 1, 2, \dots, k \\
 [ij] &= 0 & i, j = 1, 2, \dots, k \quad i \neq j \\
 [ijj] &= 0 & i, j = 1, 2, \dots, k \quad i \neq j \\
 [ijk] &= 0 & i \neq j \neq k \\
 [iiij] &= 0 & i, j = 1, 2, \dots, k \quad i \neq j \\
 [iijk] &= 0 & i \neq j \neq k
 \end{aligned} \tag{6.109}$$

The even moments,  $[ii]$ ,  $[iiii]$  and  $[iijj]$  ( $i, j = 1, 2, \dots, k \quad i \neq j$ ) are non-zero with  $[ii] = 1$  by the scaling convention. The latter two moments are influenced by the choice of the number of centre points  $n_0$  and by the value  $\alpha$  of the axial points setting.

A CCD is rotatable if  $[iiii] = 3[iijj]$  for  $i, j = 1, 2, \dots, k; i \neq j$ . If  $g$  is a scale factor chosen such as that  $[ii] = 1$ , then  $g = \left( N / (F + 2\alpha^2) \right)^{1/2}$ , where  $F$  is the number of points in the factorial portion of the design, and  $N = 2^k + 2k + n_0$  is the total number of points in the CCD. It follows that the condition of rotatability for a CCD can be written as

$$Fg^4 + 2\alpha^4 g^4 = 3Fg^4 \tag{6.110}$$

or equivalently

$$\alpha = F^{1/4} \tag{6.111}$$

As for orthogonality, a CCD can have this property if the mixed fourth-order moment  $[iijj]$ , or equivalently  $\lambda_4$  is equal to unity. In terms of the scale factor  $g$ ,  $[iijj]$  is written as

$$[iijj] = \frac{Fg^4}{N} = \frac{FN}{(F + 2\alpha^2)^2} \tag{6.112}$$

In order to have  $[iijj] = 1$  it must be

$$(F + 2\alpha^2)^2 = FN \tag{6.113}$$

Solving Eq. (6.113) brings to the conclusion that a CCD can be made orthogonal by choosing the value of the axial setting,  $\alpha$ , to be equal to

**Table 6.3. Partial table of signs for a 2<sup>6</sup> factorial design.**

Experiment	1	2	3	4	5	6	123456	Block
1	-	-	-	-	-	-	+	II
2	+	-	-	-	-	-	-	I
3	-	+	-	-	-	-	-	I
4	+	+	-	-	-	-	+	II
5	-	-	+	-	-	-	-	I
6	+	-	+	-	-	-	+	II
7	-	+	+	-	-	-	+	II
8	+	+	+	-	-	-	-	I
9	-	-	-	+	-	-	-	I
10	+	-	-	+	-	-	+	II
11	-	+	-	+	-	-	+	II
12	+	+	-	+	-	-	-	I
13	-	-	+	+	-	-	+	II
14	+	-	+	+	-	-	-	I
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
63	-	+	+	+	+	+	-	I
64	+	+	+	+	+	+	+	II

$$\alpha = \left( \frac{(FN)^{1/2} - F}{2} \right)^{1/2} \tag{6.114}$$

If it is desired for the CCD to be orthogonal as well as rotatable, it is possible to choose  $\alpha$  and  $n_0$  to achieve both properties. To do so, it suffices to replace  $\alpha^2$  in Eq. (6.113) by  $\sqrt{F}$  in order to respect the rotatability condition, the brings the following equation

$$(F + 2\sqrt{F})^2 = F(F + 2k + n_0) \tag{6.115}$$

which if solved for  $n_0$  gives

$$n_0 \approx 4\sqrt{F} + 4 - 2k \tag{6.116}$$

where  $n_0$  is equal to the integer closest to the expression on the right-hand side.

### 6.3.4 Designs for mixed models

In the previous sections the central composite design has been introduced. In its standard formulation this kind of design does allow to take into account fixed effect only. For this reason it need to be modified in order to be used with mixed models. In fact the central composite design must be partitioned in blocks to which different levels of the random factors can be associated. This section describes some criteria that can be used in order to block central composite design maintaining properties such as orthogonality and rotatability. First blocking for cases with one random factor only will be discussed (one-way classification) then the model will be extended to case with more random factors (multiple-way classification).

#### 6.3.4.1 One way blocking

First of all blocking criteria for the factorial part of the CCD will be illustrated. To do so an example design with  $2^6 = 64$  experiments will be used. A partial table of signs of this example design is given in Table 6.3. Considering the partial table of signs, which shows the design in standard order and supposing that a partitioning into two blocks is required one could allocate runs for which the product column **123456** has a minus sign to block I, and runs for which **123456** has a plus sign to block II. In this case according to Box and Draper, 1987 it is said that the block effect (**B**) is completely confused with the 123456 interaction, i.e.,

$$\mathbf{B} = \mathbf{123456} \quad (6.117)$$

Thus it is not possible to estimate (at least with the same accuracy) the 123456 interaction. However, this very high-order interaction will be negligible in most practical cases. It is worth noting that the 2 level factorial design has an important orthogonal property whereby a sequence of signs corresponding to a particular effect is orthogonal to every other such sequence. The important implication of this, is that altering the apparent 123456 effect by superimposing on it the difference between blocks does not change the estimate on any of the other effects.

Suppose now that the same design is to be arranged into four blocks each containing 16 run. This can be done by confounding two high-order interactions with block contrast. For example the contrasts 123456 and 23456 could be associated with blocks. Thus the blocking generators for the contrast would be

$$\mathbf{B}_1 = \mathbf{123456} \quad \mathbf{B}_2 = \mathbf{23456} \quad (6.118)$$

and the runs would then be allocated to the four blocks according to the following scheme:

$$\begin{array}{rcc}
 & & \mathbf{B}_2 = 23456 \\
 & & \quad - \quad + \\
 \mathbf{B}_1 = 123456 & - & I \quad II \\
 & + & III \quad IV
 \end{array} \tag{6.119}$$

That is, runs would be allocated to the four blocks *I*, *II*, *III* and *IV* as the sign associated with the columns 123456 and 23456 took the values (- -), (- +), (+ -), (+ +). Unfortunately this arrangement presents serious issues. In fact, there are 3 independent contrasts among the four blocks of 16 runs each. If  $\mathbf{B}_1$  and  $\mathbf{B}_2$  are associated with two of these contrasts, the third must be the interaction  $\mathbf{B}_1 \times \mathbf{B}_2 = \mathbf{B}_1 \mathbf{B}_2$ . However, then the latter interaction will be confounded with the interaction between **123456** and **23456** which is

$$\begin{array}{rcccccccccccccccc}
 \mathbf{123456} & + & - & - & + & - & + & + & - & - & + & + & - & . & . & . & - & + \\
 \mathbf{23456} & - & - & + & + & + & + & - & - & + & + & - & - & . & . & . & + & + \\
 \mathbf{123456} \times \mathbf{23456} & - & + & - & + & - & + & - & + & - & + & - & + & . & . & . & - & +
 \end{array}$$

that is, the interaction between is the main effect **1**. Therefore the main effect of the first variable is confounded with block differences. In order to make better blocking it is first mandatory to understand how to calculate the interaction between complex coefficients. First, it should be noted that if one takes any set of signs for any effect and multiplies it by the signs of the same effect one obtains a row of +’s which are denoted by the identity **I**

$$\begin{array}{rcccccccccccccccc}
 \mathbf{1} & - & + & - & + & - & + & - & + & - & + & - & + & . & . & . & - & + \\
 \mathbf{1} & - & + & - & + & - & + & - & + & - & + & - & + & . & . & . & - & + \\
 \mathbf{1} \times \mathbf{1} = \mathbf{1}^1 = \mathbf{I} & + & + & + & + & + & + & + & + & + & + & + & + & . & . & . & + & +
 \end{array}$$

Thus the multiplication sign is used to imply the multiplication of the signs in corresponding positions in two rows. It is possible to write

$$\mathbf{1} \times \mathbf{1} = \mathbf{1}^2 = \mathbf{I} \quad \mathbf{2} \times \mathbf{2} = \mathbf{2}^2 = \mathbf{I} \quad \mathbf{12} \times \mathbf{12} = \mathbf{1}^2 \mathbf{2}^2 = \mathbf{I}^2 = \mathbf{I} \tag{6.120}$$

Furthermore multiplication of any contrast by the identity **I** gives the contrast unchanged.

$$\mathbf{1} \times \mathbf{I} = \mathbf{1} \quad \mathbf{23} \times \mathbf{I} = \mathbf{23} \quad \mathbf{12345} \times \mathbf{I} = \mathbf{12345} \tag{6.121}$$

therefore applying this rule in the case of the design above it is possible to obtain

$$\mathbf{B}_1\mathbf{B}_2 = 123456 \times 23456 = 12^2 3^2 4^2 5^2 6^2 = 1 \times \mathbf{I}^5 = \mathbf{1} \quad (6.122)$$

indicating that  $B_1B_2$  and 1 are confounded.

An arrangement in four blocks of 16 runs so that all the interaction confounded with block are of the highest possible order may be obtained by using as generators two four-factor interactions in which only two symbols overlap. For example a possible choice is

$$\mathbf{B}_1 = 1234 \quad \mathbf{B}_2 = 3456 \quad (6.123)$$

which gives

$$\mathbf{B}_1\mathbf{B}_2 = 1256 \quad (6.124)$$

Once the block generators are defined it is possible to allocate the runs to the four blocks corresponding to the sign combinations  $(\mathbf{B}_1, \mathbf{B}_2) = (-, -), (-, +), (+, -), (+, +)$  as follows

$\mathbf{B}_1 = 1234$	+	-	-	+	-	+	+	-	-	+	+	-	.	.	.	-	+
$\mathbf{B}_2 = 3456$	+	+	+	+	-	-	-	-	-	-	-	-	.	.	.	+	+
Block	IV	II	II	IV	I	III	III	I	I	III	III	I	.	.	.	II	IV

Under the common assumption that the blocks contribute only additive effects, the main effects, two-factors interactions, and three-factor interactions will all remain unconfounded with any block effect.

Finally it is considered the case in which the example design is to be partitioned into eight blocks. If some three-factor interactions can be confounded the design can be split choosing the generators

$$\mathbf{B}_1 = 1234 \quad \mathbf{B}_2 = 3456 \quad \mathbf{B}_3 = 123 \quad (6.125)$$

which give the following confounding pattern for the seven contrasts among the eight blocks:

$$\begin{aligned}
 \mathbf{B}_1 &= 1234 \\
 \mathbf{B}_2 &= 3456 \\
 \mathbf{B}_1\mathbf{B}_2 &= 1256 \\
 \mathbf{B}_3 &= 136 \\
 \mathbf{B}_1\mathbf{B}_3 &= 246 \\
 \mathbf{B}_2\mathbf{B}_3 &= 145 \\
 \mathbf{B}_1\mathbf{B}_2\mathbf{B}_3 &= 235
 \end{aligned} \tag{6.126}$$

Once the factorial portion of the CCD has been split into block the centre points and the axial portion must be assigned to blocks as well. An operative criterion can be derived from the requirements to have orthogonality of the blocks, which means that the fixed effects can be estimated independently from the random effects (i.e. least squares method can be used). Adopting for the response surface the following expression

$$y_u = \beta_0 + \sum_{i=1}^k \beta_i x_{ui} + \sum_{i=1}^k \beta_{ii} x_{ui}^2 + \sum_{i=1}^k \sum_{\substack{j=1 \\ j>i}}^k \beta_{ij} x_{ui} x_{uj} + \sum_{l=1}^b \delta_l z_{ul} + \epsilon_u \tag{6.127}$$

where  $\delta_l$  denotes the effect of the  $l$ -th block ( $l=1, 2, \dots, b$ ),  $z_{ul}$  is a dummy variable taking the values 1, if the  $u$ -th trial is carried out in the  $l$ -th block and 0 otherwise and  $\epsilon_u$  is the random error. Alternatively the model can be written as

$$y_u = \beta'_0 + \sum_{i=1}^k \beta_i x_{ui} + \sum_{i=1}^k \beta_{ii} x_{ui}^2 + \sum_{i=1}^k \sum_{\substack{j=1 \\ j>i}}^k \beta_{ij} x_{ui} x_{uj} + \sum_{l=1}^b \delta_l (z_{ul} - \bar{z}_l) + \epsilon_u \tag{6.128}$$

where  $\bar{z}_l = (1/N) \sum_{u=1}^N z_{ul}$  and  $\beta'_0 = \beta_0 + \sum_{l=1}^b \delta_l \bar{z}_l$ . The formulation in Eq. (6.128) allows to derive the following conditions for orthogonal blocking

$$\sum_{u=1}^N x_{ui} (z_{ul} - \bar{z}_l) = 0 \quad i = 1, 2, \dots, k \quad l = 1, 2, \dots, b \tag{6.129}$$

$$\sum_{u=1}^N x_{ui} x_{uj} (z_{ul} - \bar{z}_l) = 0 \quad i, j = 1, 2, \dots, k \quad i \neq j \quad l = 1, 2, \dots, b \tag{6.130}$$

$$\sum_{u=1}^N x_{ui}^2 (z_{ul} - \bar{z}_l) = 0 \quad i, j = 1, 2, \dots, k \quad l = 1, 2, \dots, b \tag{6.131}$$

If it is required for the design to be also rotatable it is to be assumed that the first-order moments as well as the mixed second-order moments of the design are zero

$$\sum_{u=1}^N x_{ui} = 0 \quad i = 1, 2, \dots, k \quad (6.132)$$

$$\sum_{u=1}^N x_{ui} x_{uj} = 0 \quad i, j = 1, 2, \dots, k \quad i \neq j \quad (6.133)$$

Combining Eqs. (6.129)-(6.133) and considering the features of the  $z$  variables, it is possible to derive the following conditions on the  $x_{ui}$  settings for orthogonal blocking

$$\sum_{u^{(l)}} x_{ui} = 0 \quad i = 1, 2, \dots, k \quad l = 1, 2, \dots, b \quad (6.134)$$

$$\sum_{u^{(l)}} x_{ui} x_{uj} = 0 \quad i, j = 1, 2, \dots, k \quad i \neq j \quad l = 1, 2, \dots, b \quad (6.135)$$

$$\frac{\sum_{u^{(l)}} x_{ui}^2}{\sum_{u=1}^n x_{ui}^2} = \frac{n_l}{N} \quad i = 1, 2, \dots, k \quad l = 1, 2, \dots, b \quad (6.136)$$

where  $\sum_{u^{(l)}}$  denotes summation extended only over those values of  $u$  in the  $l$ -th block; and  $n_l$  is the number of runs in the  $l$ -th block. Equations (6.134)-(6.136) can be summarized as follows:

1. Conditions in Eq. (6.134) and Eq. (6.136) imply that the column arrays associated with  $x_1, x_2, \dots, x_k$  are orthogonal and sum to zero within each block. Hence each block must consist of a first-order orthogonal design.
2. Condition in Eq. (6.136) implies that the fraction of the total sum of squares for variable  $x_i$  in each block must be equal to the fraction of the total number of runs allotted to that block.

The CCD can be made to block orthogonally. In fact, each of the factorial and axial portion of the design forms a first-order orthogonal design. These portions provide a basis for a first division of the CCD into two blocks. The number of centre point replications will have to be determined to satisfy condition given in Eq. (6.136). The composition of these two blocks can be described as follows:

- Block 1. A factorial portion consisting of  $F = 2^k$  points, in addition to  $n_{OF}$  centre point replications.
- Block 2. An axial portion consisting of  $2k$  points plus  $n_{OA}$  centre point replications. From Eq. (6.136) follows that the axial setting must have the value

$$\alpha = \sqrt{\frac{F(2k + n_{OA})}{2(F + n_{OF})}} \quad (6.137)$$

Furthermore if such a design is also required to be rotatable, the condition of which is  $\alpha = F^{1/4}$ , then from Eq. (6.136) follows

$$\sqrt{F} = \frac{2(F + n_{OF})}{2k + n_{OA}} \quad (6.138)$$

or equivalently

$$2F - \sqrt{F}(2k + n_{OA}) + 2n_{OF} = 0 \quad (6.139)$$

It is worth noting from Eq. (6.139) that for some values of  $k$ , it is not always possible to find a rotatable CCD that blocks orthogonally. A necessary condition for the satisfaction of Eq. (6.139) is

$$(2k - n_{OA})^2 - 16n_{OF} \geq 0 \quad (6.140)$$

For those values of  $k$  for which Eq. (6.140) is not satisfied, it is possible to achieve

**Table 6.4. Orthogonal blocking arrangements for rotatable or near rotatable central composite designs.**  $F$ : runs in the factorial portion of the design,  $b_F$  blocks into which the factorial portion of the design is partitioned,  $n_{0l}$  number of centre point added to each block,  $n_l$  number of runs in each block,  $\alpha_{orth,blocks}$ :  $\alpha$  value to achieve orthogonality,  $\alpha_{rot}$ :  $\alpha$  value to achieve rotatability and  $n$ : total number of runs.

k	2	3	4	5	6
Factorial blocks					
$F$	4	8	16	32	64
$b_F$	1	2	2	4	8
$n_{0l}$	3	2	2	2	9
$n_l$	7	6	10	10	1
Axial block					
$n_{0l}$	3	2	2	4	6
$n_l$	7	8	10	14	18
$\alpha_{orth,blocks}$	1.4142	1.6330	2.0	2.3664	2.8284
$\alpha_{rot}$	1.4142	1.6818	2.0	2.3784	2.8284
$n$	14	20	30	54	90

orthogonal blocking and near rotatability for a CCD. For example, when  $k = 3$ ,  $F = 8$ ,  $n_{OF} = 4$  and  $n_{OA} = 2$  Eq. (6.137) gives  $\alpha = 1.633 \approx 8^{1/4} = 1.68$  which is close to the rotatable setting. A list of orthogonal blocking arrangements for rotatable or near-rotatable central composite designs is given in Table 6.4.

A further partitioning of the axial portion of the design is not compatible with the condition 1 listed above. On the other hand, the factorial portion of the design can be partitioned into more than one block until the resulting designs are of order greater or equal to 3 and the number of repetitions of the centre point is the same for each block.

# 7 Assessing seismic fragility of RC frames through response surface method

In this chapter, the response surface (RS) method with random block effects is used to approximate the structural capacity of an existing reinforced-concrete (RC) frame structure. The spectral acceleration, at the natural frequency, causing the failure of the structure is the parameter adopted to represent the structural capacity. In the following sections, the numerical model adopted to perform the non-linear time-history analyses, the criteria adopted to define the explicit and the implicit variables and the method adopted to generate the input ground-motions will be described.

The procedure is illustrated with reference to the seismic fragility assessment of a RC frame structure. RS models with random block effects have been used to approximate the dependence of the structural capacity from the random parameters considered in the analysis. Accordingly, some variables are considered explicitly in the definition of the RS (concrete strength, steel yielding stress, live loads), whereas the uncertainties related to the seismic action and the variability of the mechanical properties over the structure are considered in an implicit way.

As for the ground-motion, simulated acceleration time-histories are used, according to the method proposed by Sabetta et al. and Pugliese et al. (Sabetta and Pugliese, 1996; Buratti et al., 2006b). Artificial accelerograms are preferred here over recorded time-histories because as many ground-motions as required by the probabilistic model for reliability analysis can be generated, according to a given seismological scenario. In fact, since the Monte Carlo simulation method has been used to compute the reference values of the fragility, a very large number of statistically independent numerical simulations (and therefore of ground-motions) is needed. In Buratti et al. (Buratti et al., 2007a), it has been shown that the ground-motion generation procedure used in the present work gives acceleration time histories that: (i) do not introduce bias in the structural response and (ii) have a degree of variability similar to that of recorded ground-motions.

As for the concrete strength, non destructive testing on existing old RC frame structures often shows that the strength can be very variable, not only with respect to the value prescribed by design, but also among the various parts of the structure, for instance from one floor to another or even among the columns of the same floor. Therefore, the variation of the concrete strength has been described by a correlated random field.

RS models based on polynomial functions of different degree and obtained with different simulation-plans have been employed and compared in terms of computational efficiency and accuracy. The reference values for the structural capacity and the failure probabilities have been computed by MC simulations, obtained by fully random generation of ground-motions and mechanical/loading characteristics of the structure.

Quadratic polynomial models have shown a high sensitivity to the type of design used and require a higher number of simulations than linear models in order to give accurate results. On the contrary, these latter models have shown less sensitivity to the selected simulation plan.

## 7.1 Seismic fragility

In a seismic reliability framework, the seismic fragility curve is defined as the probability of failure of a structure conditional to the ground-motion intensity.

The structural failure is attained when the limit state function, defined as the difference between structural capacity and demand, is less than or equal to zero:

$$g = \min_t [C(\mathbf{x}, t) - D(\mathbf{x}, t)] \leq 0 \quad (7.1)$$

where both capacity,  $C$ , and demand,  $D$ , depend on the set  $\mathbf{x}$  of random variables adopted to describe the variability of some properties of the structure (e.g., related to the strength of structural elements), and on the time  $t$ .

In the definition of the limit state function in Eq. (7.1), the explicit dependence on time is eliminated, because the minimum value over the entire ground-motion duration is taken (Veneziano et al., 1983; Casciati and Faravelli, 1991). If the structural behaviour is non-linear, Monte Carlo (MC) simulation methods are typically used to solve Eq. (7.1) (Pinto, 2001). As well known, these methods require a great computational effort, especially when values of the fragility curve corresponding to low probability levels must be evaluated. Computationally less demanding procedures such as FORM and SORM may have some limitations when Eq. (7.1) is not analytical or when a suitable

probabilistic model is not available for all the random variables involved in the problem (e.g. to realistically describe the seismic action) (Pinto, 2001).

In the present study, different RS models, together with different criteria for the design of simulation-plans are used to approximate the limit state function, in order to reduce the number of simulations required to perform the fragility analysis.

First of all, an intensity measure (Shome et al., 1998; Luco and Cornell, 2007) must be chosen to express the structural capacity  $C$  and demand  $D$  adopted in the limit state function. This intensity measure shall be the same used to express the seismic hazard for the site where the structure is located. Some different proposals can be found in the literature (Shome et al., 1998; Luco and Cornell, 2007). In the present study, according to Veneziano et al. (Veneziano et al., 1983), Casciati et al. (Casciati and Faravelli, 1991) and Shome et. al (Shome et al., 1998), both structural capacity and demand are written as a function of the spectral acceleration, corresponding to the first natural frequency of the structure,  $S_a$ . Eq. (7.1) is then rewritten as:

$$g = S_{a,C}(\mathbf{x}) - S_{a,D} \quad (7.2)$$

where  $S_{a,C}$  represents the spectral acceleration causing the structural failure and  $S_{a,D}$  the spectral acceleration the structure is subjected to. Following this approach,  $S_{a,D}$  is independent from the structural capacity, because the fragility is conditioned on this parameter. The first step of the reliability analysis is to determine  $S_{a,C}$  as a function of the random variables  $\mathbf{x}$  modelling the variability of the mechanical properties and the loads for the structure. Then, the seismic fragility curve is calculated by evaluating the structural failure probability for different values of  $S_{a,D}$ , i.e.  $P_f = P(g < 0 \mid S_{a,D})$ . The evaluation of the dependence of structural capacity  $S_{a,C}$  on the random variables  $\mathbf{x}$  involved in the problem is the most computationally expensive step due to the non-linear structural behaviour and the high number of random variables. In the present study, the RS method has been used to approximate the unknown dependence of  $S_{a,C}$  on  $\mathbf{x}$ . The natural frequency of the structure, adopted to select the value of spectral acceleration, is calculated using the mean values of the random variables involved in the problem (Franchin et al., 2003b; Franchin et al., 2003a). The 5% damping-ratio spectrum has been used.

It is important to note that the effectiveness of response surface approximation depends on the efficiency of the intensity measure adopted to describe seismic action. Spectral acceleration is by far the most widely adopted parameter but some authors have recently demonstrated that other measures may give better performances as far as non-linear

response of concrete structures is concerned. Besides, scaling to spectral acceleration can lead to results affected by systematic errors (see Section 2.3).

Furthermore considering uncertainty of parameters defining structural behaviour makes spectral acceleration even less effective because the natural frequency of the structure under consideration depends on these mentioned parameters. One of the most promising intensity measure to be used in RS applications is the one proposed by Baker and Cornell (2006b), which is defined as the geometric mean of spectral acceleration values at a set of periods:

$$Sa_{avg}(T_1, \dots, T_n) = \left( \prod_{i=1}^n Sa(T_i) \right)^{1/n} \quad (7.3)$$

where  $T_1, \dots, T_n$  are  $n$  periods of interest. Eq. (7.3) can also be expressed as arithmetic mean of the values of logarithmic spectral accelerations

$$\log(Sa_{avg}(T_1, \dots, T_n)) = \frac{1}{n} \sum_{i=1}^n \log(Sa(T_i)) \quad (7.4)$$

This formulation is particularly convenient, because ground motion prediction equations (attenuation relationships) can be easily obtained for  $\log(Sa_{avg}(T_1, \dots, T_n))$ , with an arbitrary set of periods,  $T_1, \dots, T_n$ , using existing models. In fact the mean and variance of  $\log(Sa_{avg}(T_1, \dots, T_n))$  are given by (Baker and Cornell, 2006b):

$$E(\log(Sa_{avg}(T_1, \dots, T_n))) = \frac{1}{n} \sum_{i=1}^n E(\log(Sa(T_i))) \quad (7.5)$$

$$\text{Var}(\log(Sa_{avg}(T_1, \dots, T_n))) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \rho_{\log(Sa_{T_i}), \log(Sa_{T_j})} \sigma_{\log(Sa_{T_i})} \sigma_{\log(Sa_{T_j})} \quad (7.6)$$

where  $E(\log(Sa(T_i)))$  and  $\sigma_{\log(Sa_{T_i})}$  are the conditional mean and standard deviation of  $\log(Sa(T_i))$ , available from popular ground motion prediction models (e.g. Ambraseys et al., 2005; Abrahamson and Silva, 2008b). The term  $\rho_{\log(Sa_{T_i}), \log(Sa_{T_j})}$  describes the correlation between  $\log(Sa(T_i))$  and  $\log(Sa(T_j))$  and can be evaluated via regression analysis of empirical observations (Baker and Cornell, 2006a; Baker, 2007a; Baker and Jayaram, 2008). Eq. (7.5) and Eq. (7.6) are the conditional logarithmic mean and variance, given magnitude, distance, etc. as with standard ground motion prediction models for spectral acceleration. If  $\log(Sa(T_i))$  are assumed to be jointly Gaussian (Stewart et al., 2001; Bazzurro and Cornell, 2002; Jayaram and Baker, 2008), their sum

is also Gaussian and Eq. (7.5) and Eq. (7.6) completely defines the distribution of  $\log(Sa_{avg}(T_1, \dots, T_n))$ . Therefore probabilistic seismic hazard analysis and disaggregation can then be performed using this intensity measure using the same procedures used for any single spectral acceleration value.

In the following, even if the described approach may have some advantages in terms of accuracy, since  $Sa_{avg}$  is a much more efficient intensity measure than  $Sa(T_1)$ , for the sake of simplicity a more standard intensity measure, such as spectral acceleration at the natural frequency, has been used.

## 7.2 The case study: RC frame structure

The geometry of the case study RC frame structure, the cross-sections and steel reinforcements of beams and columns are reported in Figure 7.1 and Table 7.1. The structure has two bays, 500 cm and 550 cm wide respectively, and three 300 cm high storeys. In addition to beam and column self-weights, the dead load  $q_D = 33 \text{ kN/m}$  is prescribed along all the beams. The live load  $q_L$  acting on the first floor beams is 1.5 times greater than on the two upper floors and is defined according to a lognormal distribution. All the above parameters have been defined on the basis of engineering judgement with reference to the construction practice in Italy, for this kind of building, during the 90s.

Non-linear dynamic analyses have been performed using the Finite Element software OpenSees (McKenna, 1997; McKenna et al., 2000; Scott et al., 2008). Beams and columns have been modelled as finite elements with distributed inelasticity, based on the flexibility formulation (Spacone et al., 1996b; Spacone et al., 1996c). The cross sections have been discretized into fibres and 5 and 7 control sections have been used for beams and columns, respectively (McKenna et al., 2000). Non-linear constitutive behaviours have been defined for concrete and steel. Each concrete section has been subdivided in unconfined and confined zones: Saenz's (Ceb-Fip, 1993) and Mander's (Mander et al., 1989) models have been used for concrete constitutive laws, respectively. The constitutive behaviour of reinforcing steel has been modelled according to a bilinear model.

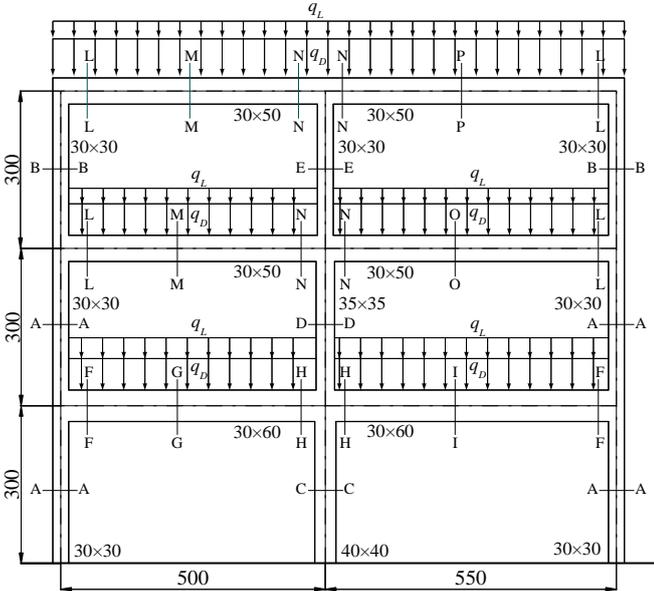


Figure 7.1. Live loads, geometry and cross sections for the case study frame structure considered in the present chapter. See also Table 7.1 for cross section details.

### 7.3 Search of the structural capacity

The seismic capacity of the structure is defined as the value of the spectral acceleration  $S_{a,C}$  causing the failure of the structure. Therefore, a limit state condition for the structure must be defined.

Different measures can be adopted to define the capacity of the structure to be compared with the seismic demand, e.g. maximum shear, interstorey drift, chord rotation, etc. In this study, only an ultimate limit state condition is defined: structural failure is attained when, in a section of the RC frame, the maximum strain in the confined concrete reaches a prescribed ultimate value,  $\epsilon_{cu}$  (Kowalsky, 2000; Priestley et al., 2007). The ultimate strain capacity  $\epsilon_{cu}$  is evaluated using the Mander’s model (Mander et al., 1989; Priestley et al., 2007) for confined concrete, in order to take into account the beneficial effect of closely spaced steel stirrups in the critical regions of the frame (end regions of columns and beams), where plastic deformations can be very high.

The structural capacity of the structure,  $S_{a,C}$ , is calculated by scaling the considered ground-motion up to the attainment of the structural failure. In each simulation, in order to obtain the required scaling factor, an incremental dynamic analysis (Vamvatsikos, 2002; Vamvatsikos and Cornell, 2002; Vamvatsikos and Cornell, 2004) is performed: according to this method, the structure is subjected to an accelerogram, linearly scaled (in

amplitude) to increasing levels of intensity until the limit state condition is exceeded. Then, the iterative Brent's method (Press et al., 2002) is used to obtain, up to the desired accuracy, the scaling factor causing the structural collapse and, therefore,  $S_{a,C}$  which is given by the product of the scaling factor and the spectral acceleration at the natural frequency of the structure of the unscaled accelerogram. The Brent's method combines root bracketing, bisection and inverse quadratic interpolation. This algorithm could be extended (Vamvatsikos and Cornell, 2002) in order to take into account more than one performance level.

The aforementioned procedure has been automated developing a software (in the Matlab framework) that executes all the incremental analyses required by a given experimental design. In particular the software performs the following tasks:

- 1) definition of the experimental plans (CCDs and factorial designs) with and without blocking (see Section 6.3);
- 2) search of structural capacity for each experiment;
  - a) modification of the structural model according to the values of explicit and implicit variables;
  - b) search for the initial interval to be used as input for Brent's algorithm;

**Table 7.1. Cross section dimension and reinforcing bars for the case study structure.**

Section	Width cm	Height cm	As	As'	Stirrups
A-A	30	30	2 $\phi$ 20	2 $\phi$ 20	$\phi$ 10@20
B-B	30	30	3 $\phi$ 20	3 $\phi$ 20	$\phi$ 10@20
C-C	40	40	2 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
D-D	35	35	2 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
E-E	30	30	2 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
F-F	30	60	2 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
G-G	30	60	4 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
H-H	30	60	2 $\phi$ 18+ 1 $\phi$ 20	2 $\phi$ 18+ 4 $\phi$ 20	$\phi$ 10@20
I-I	30	60	2 $\phi$ 18+ 2 $\phi$ 20	2 $\phi$ 18	$\phi$ 10@20
L-L	30	50	2 $\phi$ 18	3 $\phi$ 18	$\phi$ 10@20
M-M	30	50	4 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
N-N	30	50	2 $\phi$ 18+ 3 $\phi$ 20	2 $\phi$ 18+ 4 $\phi$ 20	$\phi$ 10@20
O-O	30	50	5 $\phi$ 18	2 $\phi$ 18	$\phi$ 10@20
P-P	30	50	2 $\phi$ 18+ 3 $\phi$ 20	2 $\phi$ 18	$\phi$ 10@20

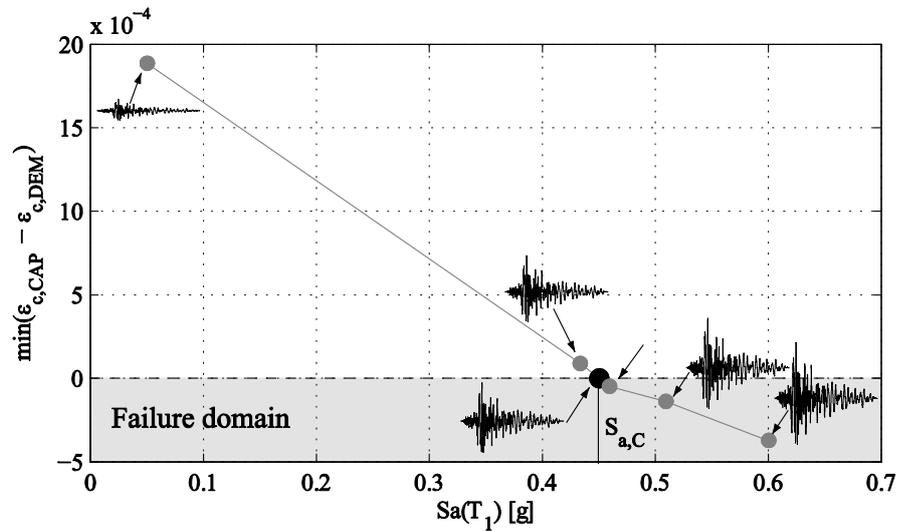
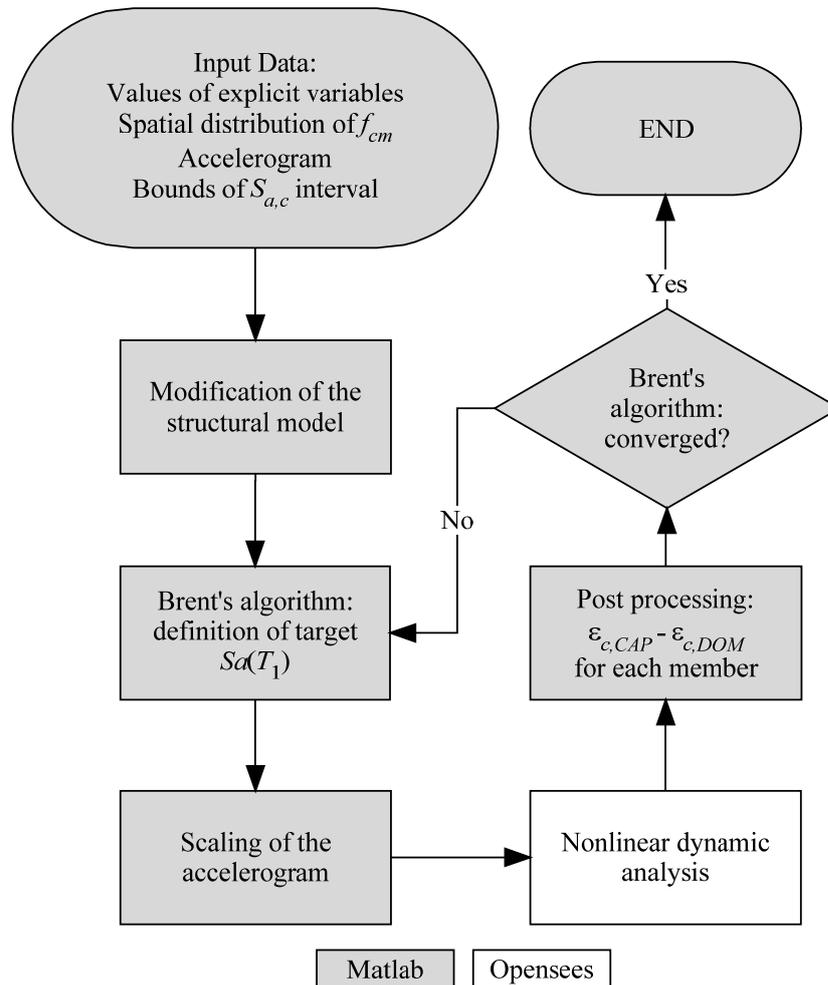


Figure 7.2. Example of application of Brent's algorithm to search structural capacity in terms of spectral acceleration at natural period of the structure. Both structural capacity,  $\epsilon_{c,CAP}$ , and demand,  $\epsilon_{c,DEM}$  are expressed in terms of maximum concrete compressive strain.

- c) search of structural capacity using Brent's algorithm;
- 3) calibration of response surfaces;
- 4) calculation of fragility curves using Monte Carlo simulation.

It is worth noting that during step 2, the Matlab software interacts with Opensees which is used to perform the nonlinear dynamic analyses. At step 2b, the software performs a sequence of nonlinear dynamic analyses at increasing levels of spectral acceleration until failure is attained. The highest spectral acceleration value not bringing the structure to collapse and the smallest spectral acceleration value bringing the structure to collapse are used to define the starting interval for Brent's algorithm (step 2c). A block diagram of the procedure used to search structural capacity at step 2c is given in Figure 7.3. At the end of each analysis, results are automatically post-processed by the software and the attainment of the considered limit state is verified. To this purpose the maximum compressive deformations recorded for each section of each structural member, during the analysis are compared with their corresponding capacity values. It should be observed that capacity values may be different for each structural member as a consequence of the considered spatial variability of concrete mean strength (See Section 7.4.2).



**Figure 7.3. Search of structural capacity: block diagram of the structure of the software developed to automate the procedure.**

As well known, the spectral acceleration is not a completely *efficient* intensity measure of the ground-motion, because it is only fully correlated with the structural response (Luco and Cornell, 2007). Hence, a number of ground-motions must be used in order to obtain an estimate of the capacity  $S_{a,C}$  (and consequently of the fragility), which is independent from the input ground-motions used to perform the simulations. In the present work, artificial acceleration time histories, compatible with the dominant seismic hazard scenario at the considered site, have been generated using the procedure proposed by Sabetta and Pugliese (Sabetta et al., 1986; Sabetta and Pugliese, 1996) described in detail in Chapter 4.

**Table 7.2. Distributions of the explicit random variables considered in the present study.**

Variable	Distribution	Mean value	COV
$q_L$	LN	4.31 $kN/m$	0.8
$f_y$	LN	515.0 $N/mm^2$	0.1
$f_{cm}$	LN	33.0 $N/mm^2$	0.2
$\Delta_\epsilon$	LN	1.0	0.2

## 7.4 Selection of the random variables

The criteria adopted to define the random variables involved in the analyses, ground-motion time-histories and to calibrate the RS are described here.

### 7.4.1 Explicit random variables

Four explicit random variables, related with the mechanical properties of the RC structure and applied loads, are considered in the analysis. Their distributions are summarized in Table 7.2. The live load acting on the structure is defined by the random variable  $q_L$ . Steel yielding stress and mean concrete compressive strength are denoted by  $f_y$  and  $f_{cm}$ , respectively. Moreover,  $\Delta_\epsilon$  represents the uncertainty in the definition of the ultimate confined concrete strain. In other words, the ultimate strain is defined as  $\epsilon_{cu} \cdot \Delta_\epsilon$ , where  $\epsilon_{cu}$  is a function of the concrete compressive strength  $f_c$  (according to the Mander's model) and  $\Delta_\epsilon$  represents the uncertainty on that value. This criterion is used in order to introduce variability in the concrete ultimate strain while maintaining correlation between concrete compressive strength  $f_c$  and ultimate concrete strain  $\epsilon_{cu}$ .

### 7.4.2 Implicit random variables

The concrete compressive strength is considered non uniform among different structural members. The actual value,  $f_c$ , of the strength of a beam/column is given as  $f_c = f_{cm} \cdot \mu_c$ , i.e., the product between the mean value  $f_{cm}$ , valid for the entire structure and taken as an explicit random variable, and a local fluctuation  $\mu_c$ . Both random variables have lognormal distributions. The local strength fluctuation  $\mu_c$  is modeled using a random field with discrete support (one different value for each beam or column) with partial (prescribed) correlation among the columns of the same floor and no correlation among those of different floors. This model should provide a good representation, in a statistical sense, of the concrete strength distribution in existing RC structures, due to the sequence of building stages. The mean value of the random variable  $\mu_c$  is 1. Its Coefficient of Variation (COV) is calculated assuming COV = 0.3 for  $f_c$  and COV = 0.2

for  $f_{cm}$  (Ceb-Fip, 1993). The random variables describing the local fluctuation  $\mu_c$  among the frame elements are taken as implicit variables, and their effects on the structural capacity are considered through the random factor  $\delta_{\mu_c}$ . A different sample of the random field (i.e. a different strength distribution over the structure) is adopted in each block related to  $\delta_{\mu_c}$ .

A second random factor  $\delta_{sis}$  is used to represent the uncertainty on the seismic action, i.e., a different time-history is adopted in each block related to  $\delta_{sis}$ . In order to perform a consistent sensitivity analysis of results, the selected accelerograms must constitute a homogeneous population, fully representative of the possible variability of the ground-motion occurring in a given site. Two kinds of accelerograms are typically used in for non-linear dynamic analyses: artificial or recorded accelerograms (Bommer and Acevedo, 2004; Buratti et al., 2007a). In the first case, an as wide as required population can be generated, but artificial accelerograms may be quite different from natural accelerograms (Bommer and Acevedo, 2004). On the opposite side, the selection of recorded accelerograms may not give a homogeneous population, because of the incompleteness of ground-motion databases.

In the present work, the method proposed by Sabetta and Pugliese (1996), and described in detail in Chapter 4, has been adopted. It allows to generate artificial ground-motions starting from magnitude, epicentral distance, and soil type of the site considered. This model adopts four strong-motion indicators to generate the artificial time-histories. These indicators are the ground-motion duration, the Arias intensity, the central frequency ( $F_c$ ) and the frequency bandwidth ( $F_b$ ) of the signal. Their values are estimated through empirical predictive equations (attenuation relationships). The simulation of a non-stationary strong ground-motion is achieved through an empirical model, where time dependence and frequency content of the signal are represented through the signal spectrogram.

Recent studies performed by the present author (Buratti et al., 2007a) have shown that, for the case study structure, artificial time-histories generated by this method are characterized by a degree of variability comparable with that of recorded ground-motions, and that they do not introduce bias in the structural response. The adopted ground motions have been generated assuming rock soil.

**Table 7.3. Number of repetitions of the simulation designs adopted for the different RS models.**

Design	Response surface models			
	$Q_{fix}$	Q	L	C
RND	1	-	-	-
CCD	-	2-9	2-9	2-9
F1	-	-	2-9	2-9
F2	-	-	2-9	2-9

## 7.5 RS models and Design of Experiments

Response surface models and experimental designs have been described in great detail in Chapter 6, but for ease of reading it is worth repeating the general expression of the polynomial response model which has been used in the present work. The adopted model can be written as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_{seis} \boldsymbol{\delta}_{seis} + \mathbf{Z}_{\mu_c} \boldsymbol{\delta}_{\mu_c} + \boldsymbol{\varepsilon} \quad (7.7)$$

where  $\mathbf{Y}$  is a column vector collecting the results of the simulations; the design matrix,  $\mathbf{X}$  contains, the values given to the explicit variables in each simulation;  $\boldsymbol{\beta}$  is a vector of unknown regression parameters;  $\boldsymbol{\delta}_{seis}$  and  $\boldsymbol{\delta}_{\mu_c}$  are vectors containing the unknown values assumed by the two random factors considered (see Section 7.4); and  $\boldsymbol{\varepsilon}$  is a vector containing realizations of the random error term. The unknown values assumed by the random factors are associated to their corresponding blocks of simulations by the Boolean matrices  $\mathbf{Z}_{seis}$  and  $\mathbf{Z}_{\mu_c}$ .

With reference to the model reported in Eq. (7.7), a simulation plan must be selected to obtain the data for the estimation of the RS parameters giving a good prediction of the true response in the region of interest (Box and Draper, 1987). In the present work, different RS models and simulation plans have been adopted, with the objective of obtaining a good compromise between reliability of results and computational saving. Previous works by the present authors (Buratti et al., 2006c) suggested that RS models simpler than quadratic may be, in some cases, sufficiently accurate in calculating fragility curves, and require a computational cost significantly smaller, because a reduced number of numerical simulations is needed to calibrate the model.

The models and the simulation designs investigated in the present study are summarized in Table 3. The following polynomial models have been adopted for the RS:

- $Q_{fix}$ : Quadratic model for the mean response and no random block.
- Q: Quadratic model for the mean response and two random block effects

- L: Linear model for the mean response and two random block effects.
- C: Constant model for the mean response and two random block effects.

Moreover, the following simulation plans have been adopted:

- RND: fully random design. It is used to build a reference model. In each simulation, a new random sample of implicit and explicit variables is generated. 1100 random samples have been used to calibrate the model (i.e., 1100 simulations).
- CCD: central composite design with  $2^k$  factorial points at  $\pm 1\sigma$  and centre points at  $\alpha = \pm 2\sigma$  for the explicit random variables  $\mathbf{x}_E$ , 3 blocks for  $\delta_{sis}$  (2 from factorial portion, 1 from axial portion) and 1 block for  $\delta_{\mu_c}$ . 2 centre point replicates are added to each  $\delta_{sis}$  block. The total number of simulations is  $N = 30$ .
- F1:  $2^k$  factorial design with axial points at  $\pm 1\sigma$  for  $\mathbf{x}_E$ , 2 blocks for  $\delta_{sis}$  and 2 blocks for  $\delta_{\mu_c}$ . The total number of simulations is  $N = 16$ .
- F2:  $2^k$  factorial design with axial points at  $\pm 2\sigma$  for  $\mathbf{x}_E$ , 2 blocks for  $\delta_{sis}$  and 2 blocks for  $\delta_{\mu_c}$ . The total number of simulations is  $N = 16$ .

For the last three designs, the simulations have been repeated a number of times in order to increase the total number of blocks, i.e. to increase the number of different time-histories and random strength distributions used to calibrate the model.

In the following, the models will be named with the notation: Model-Design-Repetitions. For instance Q-CCD-6 is a quadratic polynomial model, calibrated using 6 repetitions of the basic CCD described before. Therefore, the total number of blocks is 18 for  $\delta_{sis}$  and 6 for  $\delta_{\mu_c}$  and the total number of simulations is  $N = 180$ . Since in each block a different random sample of the implicit variables is used (i.e. different accelerograms and concrete strength distributions are used), different results are obtained when different design repetitions are performed, even if the values of explicit random variables are the same. Different calibrations of the same model and design starting from different ground-motions and concrete strength distributions will be distinguished by appending a letter to the already described notation, e.g., Q-CCD-6a, Q-CCD-6b.

The numerical values reported in Table 3 indicate the number of repetitions of the basic designs used to calibrate the different models.

## 7.6 Predictions by different models

Once the values of the random variables have been selected and the numerical simulations have been performed, the RS models are calibrated following the procedure

described in Section 6.2. It is convenient to use the RS to approximate  $\ln(S_{a,C})$  instead of  $S_{a,C}$ , so avoiding the predictions of negative values of the structural capacity which would be meaningless.

First of all, the RSs calibrated using different models are compared. Figures 1.4-1.17 show some sample sections of the mean response surface obtained by different RS models setting in turns three of the explicit variables to their mean values, and considering only the variability of the response parameter  $\ln(S_{a,C})$  with respect to the fourth, free, explicit variable. Some reference capacity values to be used for comparison have been obtained via Monte Carlo simulation. For each reference value, one time-history, one strength distribution (implicit variables) and one value of the free explicit variable are randomly sampled. This set of results, which are plotted as crosses ( $\times$ ), will be called *reduced Monte Carlo* and used as the reference for comparison with polynomial RS models.

### 7.6.1 Quadratic models

Figures 7.4-7.7 depict some sections of the Q-CCD-9 quadratic response surface models. The thick line indicates the mean surface given by  $Q_{fix}$ -RND (the reference model); dash-dot lines indicate intervals at 68.2% and 95.4% confidence level. The thin lines represent the mean response surfaces given by different Q-CCD-9 models and the dots some simulation results obtained from different repetitions of CCD used to calibrate the RS. Crosses indicate the results of the reduced Monte Carlo simulation. The regression parameters defining the RSs have been obtained from regression together with the variances of  $\varepsilon$ ,  $\delta_{sis}$  and  $\delta_{\mu_c}$ . On the right side of the figures, the structural capacity as predicted by the models for different values of the free explicit variable, is represented as a normal variable with mean value equal to that predicted by the RS and variance equal to  $\text{Var}[\varepsilon]$  for the  $Q_{fix}$ -RND model (no random block effects) and  $\text{Var}[\varepsilon + \delta_{sis} + \delta_{\mu_c}]$  for the models with random effects. In order to predict the structural fragility with accuracy, a model giving good estimates of both the mean value and the variance of the RS is required. It is worth noting how logarithmic structural capacity depends on the explicit variables considered: it increases almost linearly with concrete compressive strength (Figure 7.4); it increases with  $\Delta_\varepsilon$  (Figure 7.5), but with a lower increase rate than in the previous case and; it is almost independent of  $f_y$  (Figure 7.6); and it decreases almost linearly with  $q_L$ .

According to Figures 7.4-7.7, all Q-CCD-9 models give good predictions of the structural capacity at points close to the centre of the design, but the accuracy decreases

if RS are compared with the reference models at a certain distance from this point. It is important to note that all the terms of the polynomial models have been used but, in some cases (see Figure 7.4 and Figure 7.7) this leads to a lack of accuracy as the distance from the centre of design increases.

Sections of the RSs for different Q-CCD-5 and Q-CCD-2 models are depicted in Figure 7.8 and Figure 7.9 and compared with the *reduced-MC* model. It is clear that, the low number of repetitions of the base designs may give completely wrong predictions of both mean response surface and variance even at the centre of the design  $\bar{x}_E$ . Furthermore these latter models are also very sensitive to the samples of implicit variables used; in fact the RSs corresponding to different repetitions of the designs are very spread.

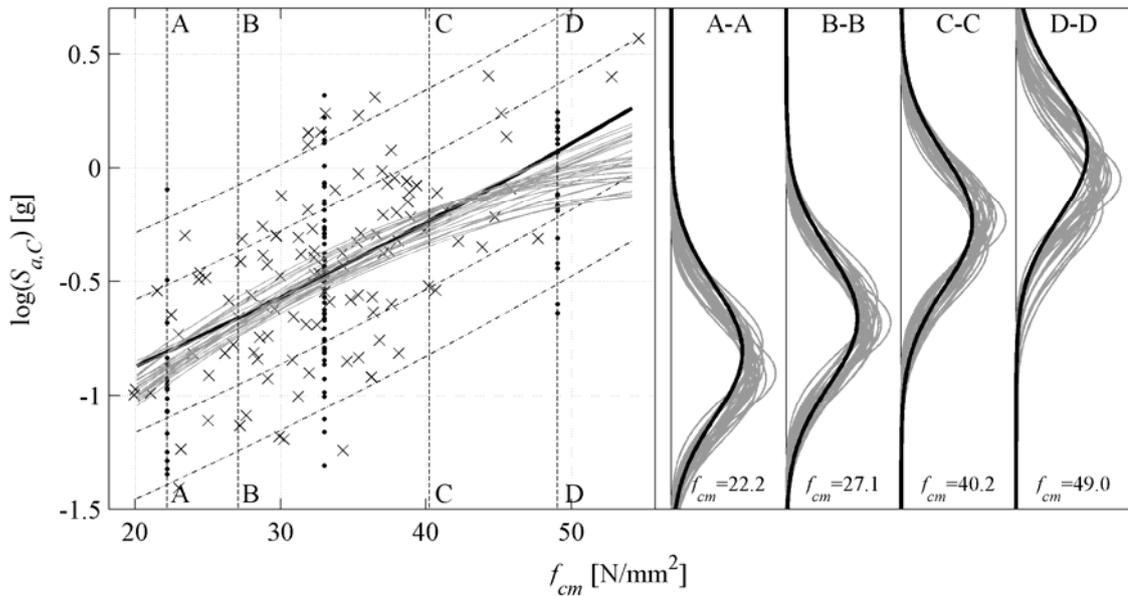


Figure 7.4. Quadratic models: sections (with  $q_L, f_y, \Delta_e$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

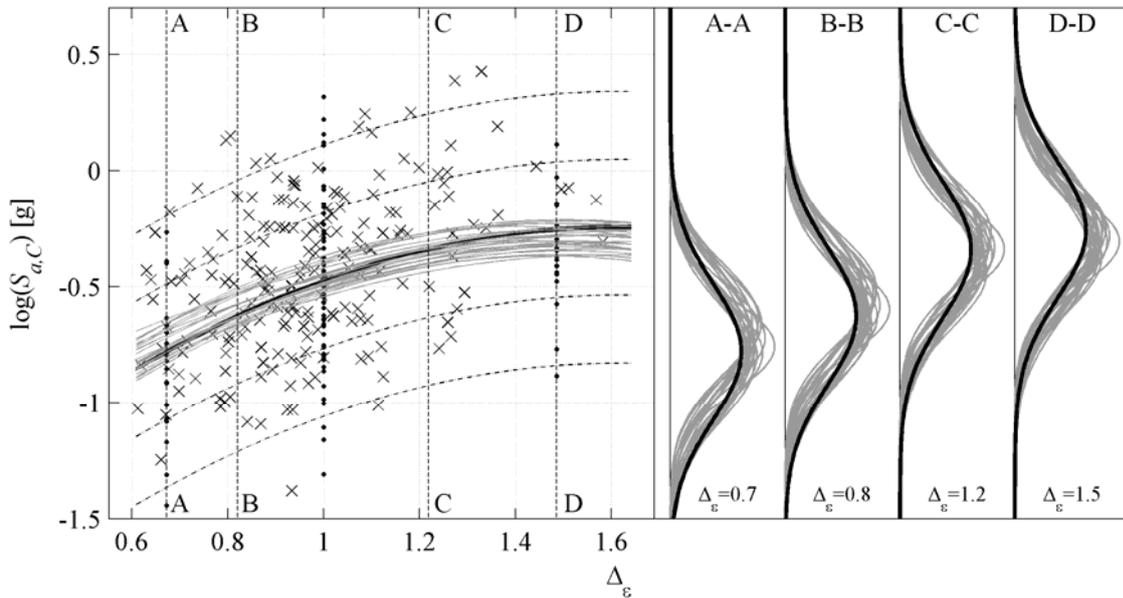


Figure 7.5. Quadratic models: sections (with  $f_{cm}$ ,  $q_L$ ,  $f_y$ , equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

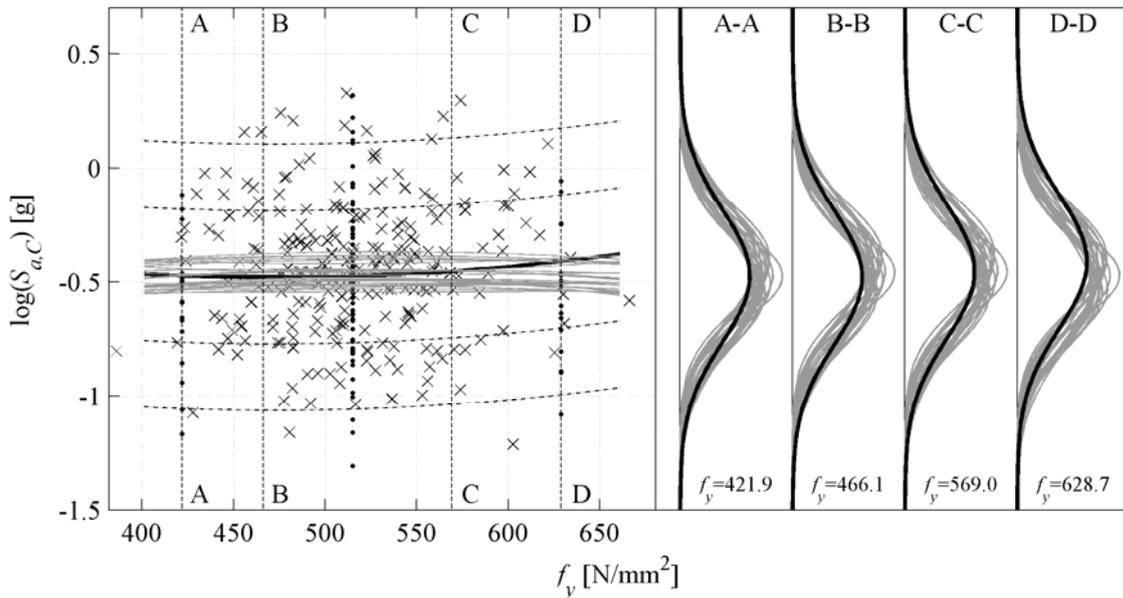


Figure 7.6. Quadratic models: sections (with  $f_{cm}$ ,  $\Delta_\epsilon$ ,  $q_L$ , equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

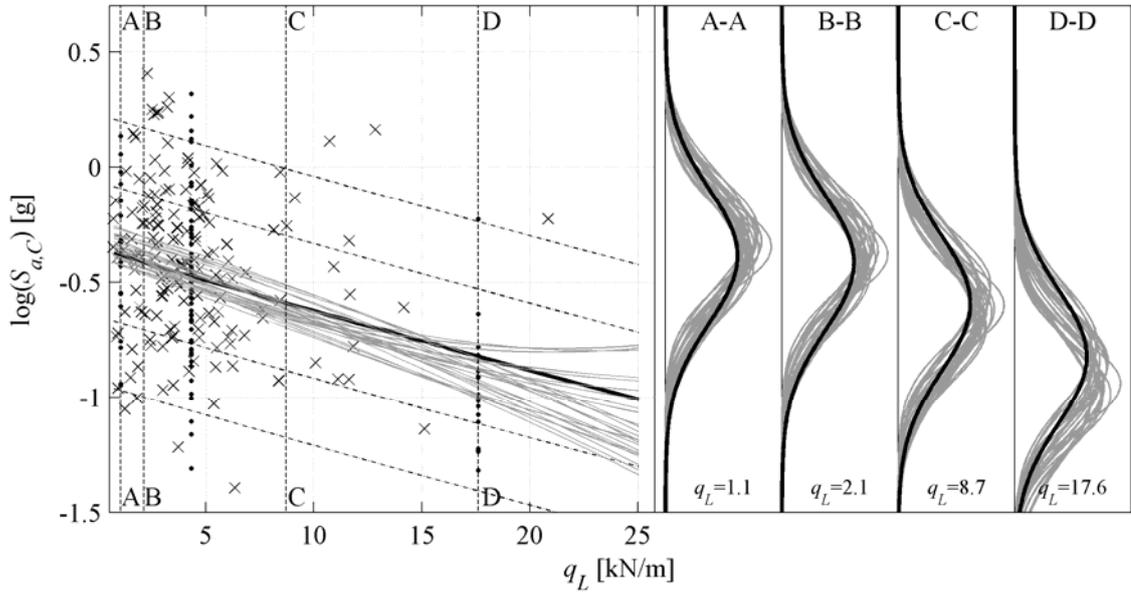


Figure 7.7. Quadratic models: sections (with  $f_{cm}$ ,  $\Delta_{\epsilon}$ ,  $f_y$ , equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

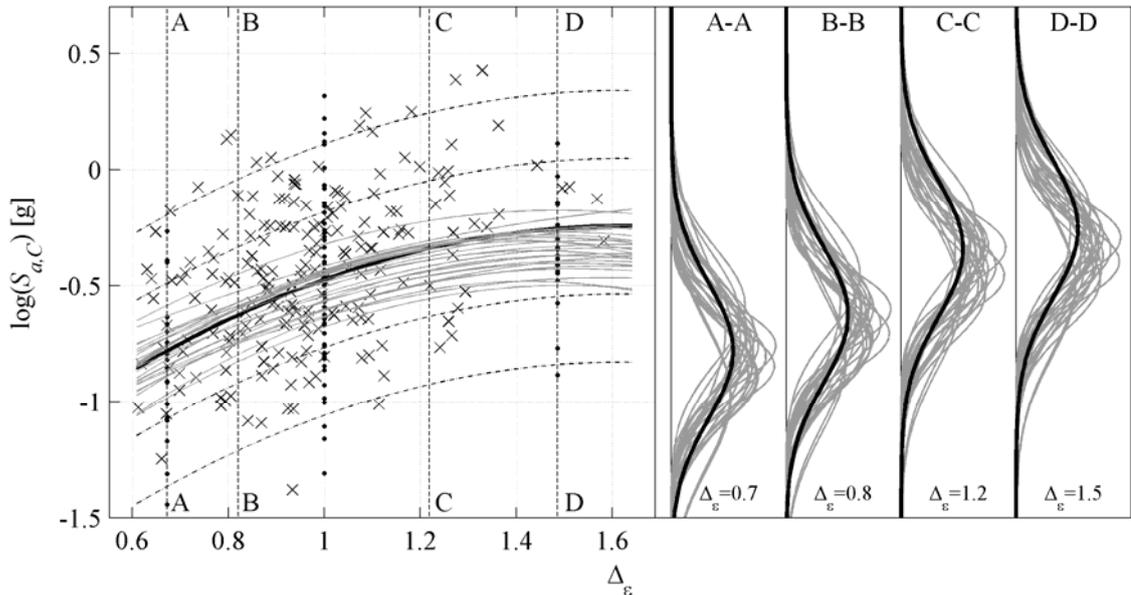


Figure 7.8. Quadratic models: sections (with  $f_{cm}$ ,  $q_L$ ,  $f_y$ , equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-5 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

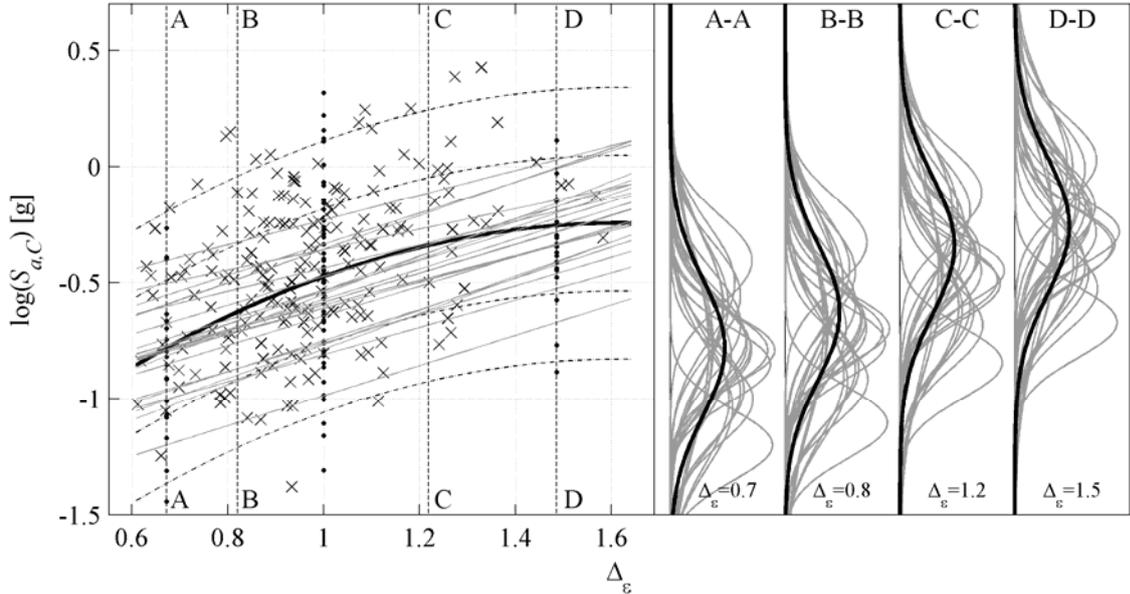


Figure 7.9. Quadratic models: sections (with  $q_L, f_y, \Delta_\epsilon$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and Q-CCD-2 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

7.6.2 Linear models

Figures 7.10-7.13 show some sections of the response surfaces obtained from the L-F1-9 models (i.e., linear models with 2-level factorial designs and 9 repetitions, total number of simulations  $N = 144$ ). Different L-F1-9 models give practically the same result, because the number of blocks is sufficiently high to describe well the variability of the implicit variables. Nevertheless, a (systematic) error is present for both mean surface and variance. In fact all these models underestimate both the mean structural capacity and its variance. In Figure 7.14, the sections of L-F1-2 model confirms the lower sensitivity of linear models to the number of blocks used, in fact the curves are less spread than those for Q-CCD-2 models, which are depicted in Figure 7.9.

Figure 7.15 shows cross sections of L-CCD-9 model. It should be noted that the use of CCD design instead of F1 design reduces the systematic error in the mean surface model.

Finally sections of L-F2-9 models are given in Figure 7.16 and Figure 7.17. These models are clearly the least accurate in predicting structural capacity, since they systematically predict lower that reference values.

It is worth noting that the black dots in Figures 7.10-7.13 represents results from CCDs, because variable values in factorial designs are not intersected by planes passing from the origin of the space and parallel to one of the axes. There point are plotted in the

afore mentioned figures because they give indications on the actual structural capacity distribution,

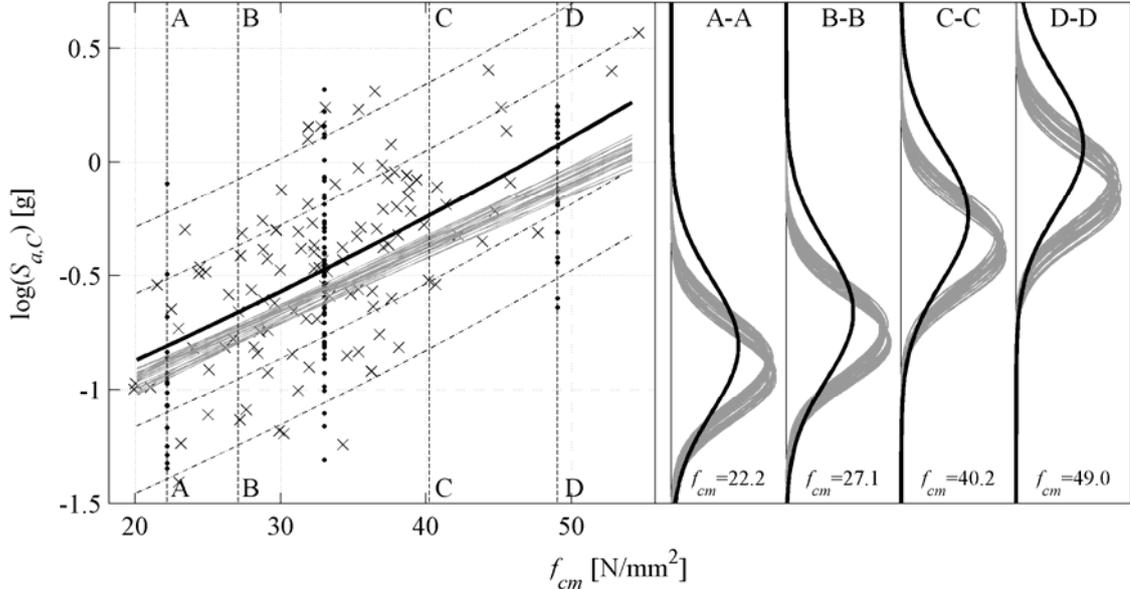


Figure 7.10. Linear models: sections (with  $\Delta_e$ ,  $q_L$ ,  $f_y$ , equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F1-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

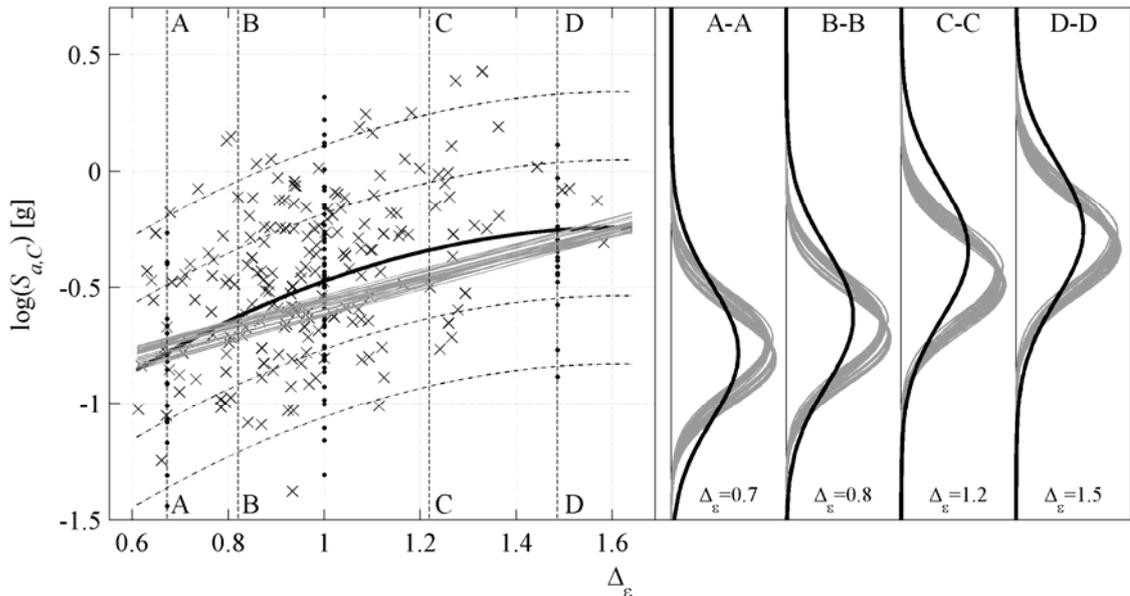


Figure 7.11. Linear models: sections (with  $f_{cm}$ ,  $f_y$ ,  $q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F1-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

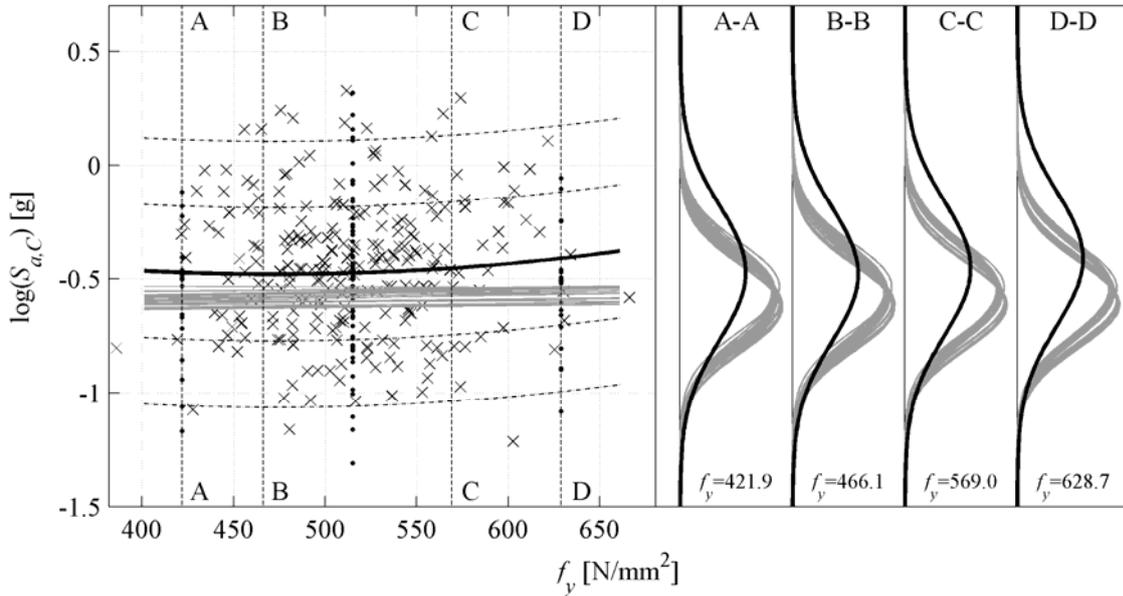


Figure 7.12. Linear models: sections (with  $f_{cm}$ ,  $\Delta_e$ ,  $q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F1-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

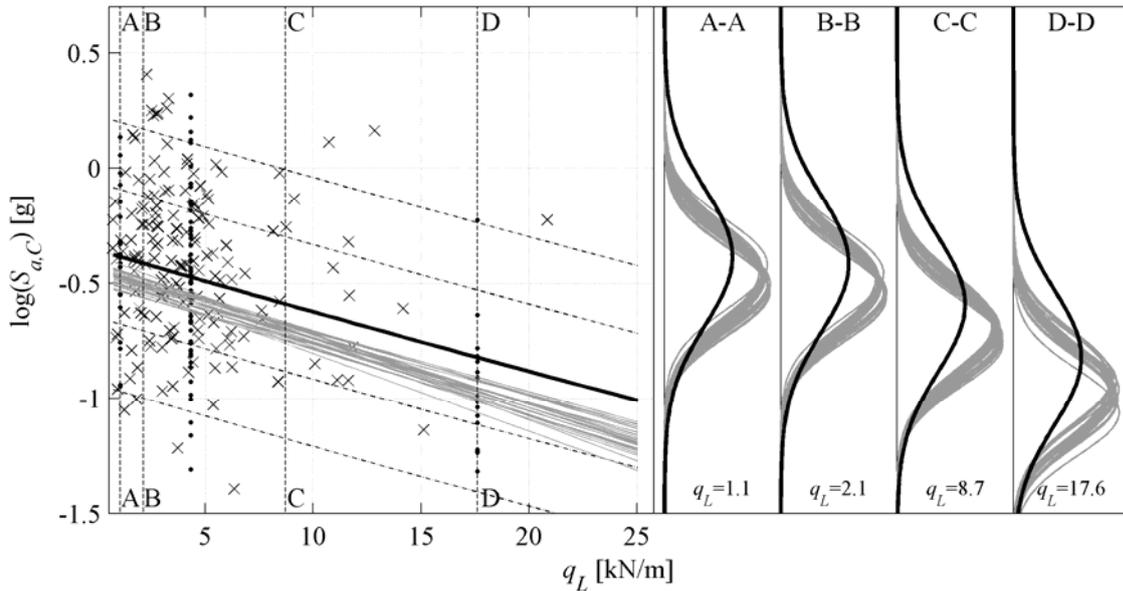


Figure 7.13. Linear models: sections (with  $f_{cm}$ ,  $\Delta_e$ ,  $f_y$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F1-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

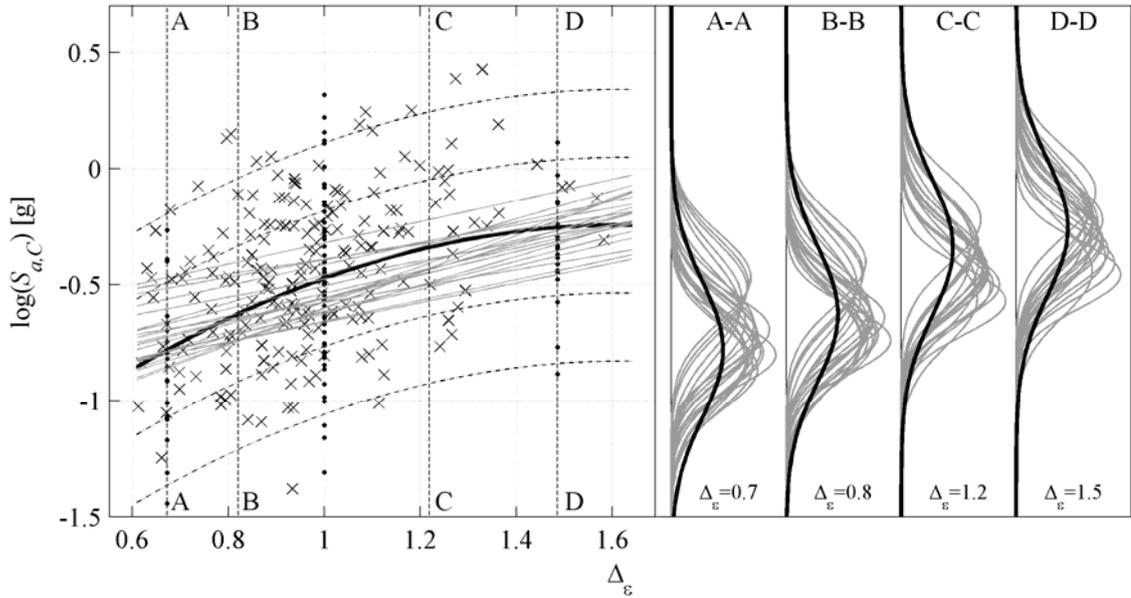


Figure 7.14. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F1-2 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

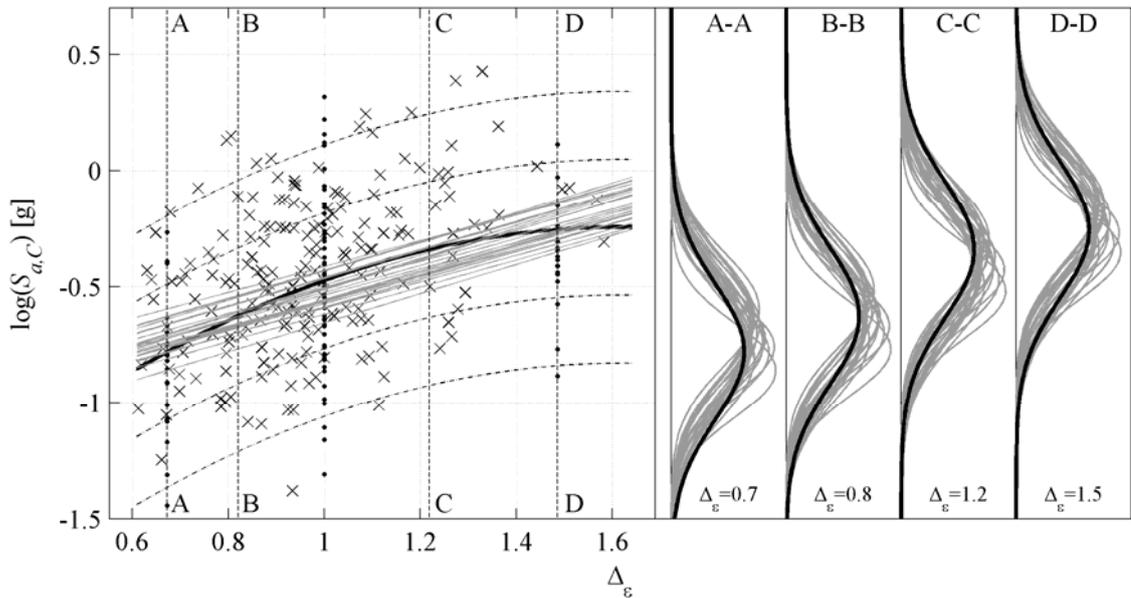


Figure 7.15. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

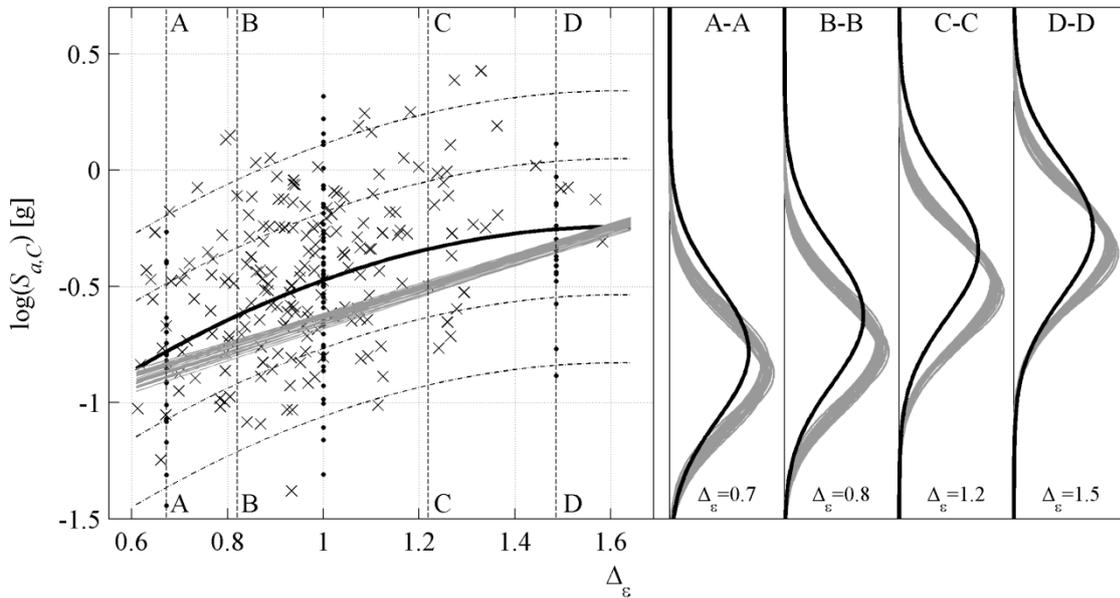


Figure 7.16. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F2-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

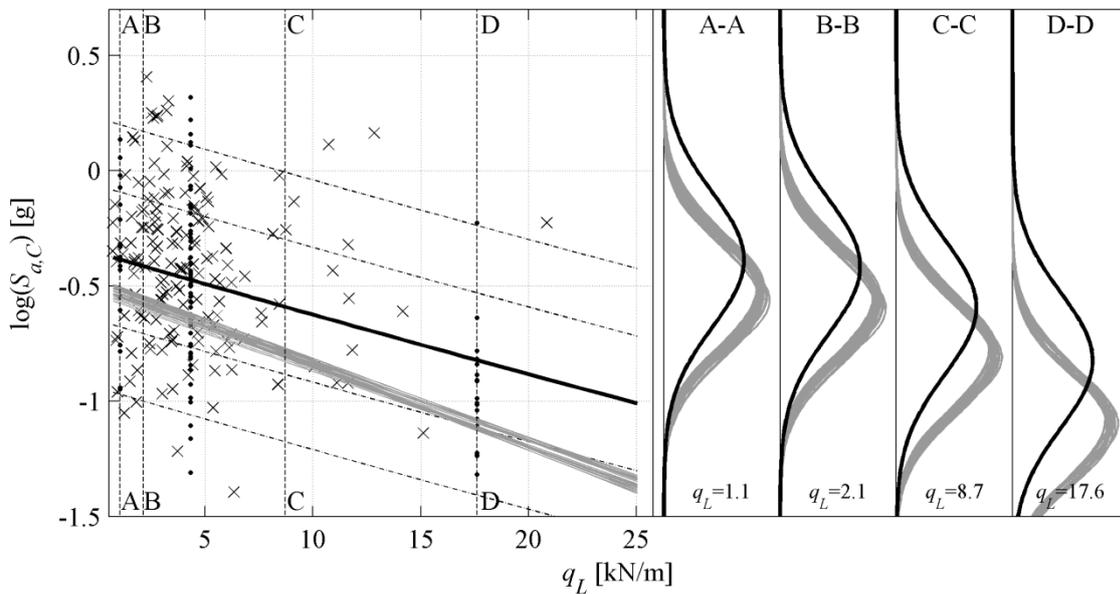


Figure 7.17. Linear models: sections (with  $f_{cm}, \Delta_\epsilon, f_y$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and L-F2-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

### 7.6.3 Constant models

Constant models have also been fitted using data from CCD, F1 and F2 designs. Figures 7.18-7.20 depict sections of constant models fitted using data from different designs (C-CCD-9, C-F1-9 and C-F2-9). Of course these models consider only the average structural capacity and do not take into account the explicit dependence on the explicit variables.

The idea that may suggest to use these models is that the variability in structural response induced by different accelerograms is so high that an accurate estimate of the dependency of structural capacity on the explicit variables is not required. As it is evident from Figures 7.18-7.20, C-CCD-9 can still provide a reasonable approximation of structural capacity distribution whereas the models fitted using first order designs, C-F1-9 and C-F2-9, are far less accurate. It is important to note that constant models are advantageous only if they can be fitted using simplified designs because most of the computational cost lies in performing the non linear incremental analyses. This will be further discussed in the following when fragility values obtained by different response surface models will be compared.

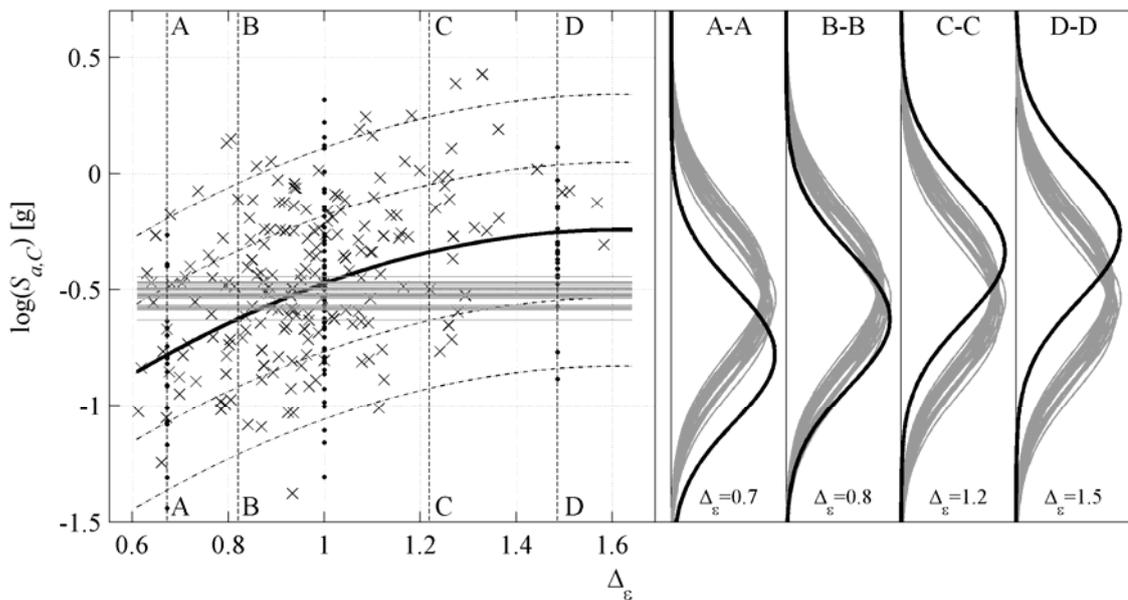


Figure 7.18. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and C-CCD-9 (thin gray lines) models.  $\times\times\times$ : capacity values from *reduced-MC* simulation and  $\bullet\bullet\bullet$  from the CCD used for RS calibration.

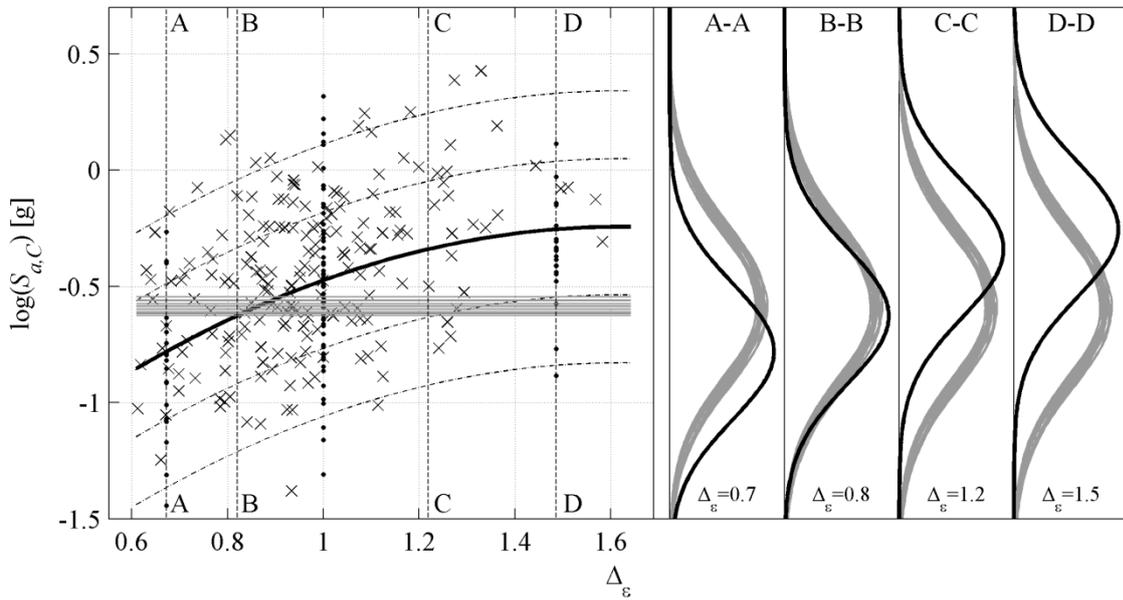


Figure 7.19. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and C-F1-9 (thin gray lines) models. xxx: capacity values from *reduced-MC* simulation and ●● from the CCD used for RS calibration.

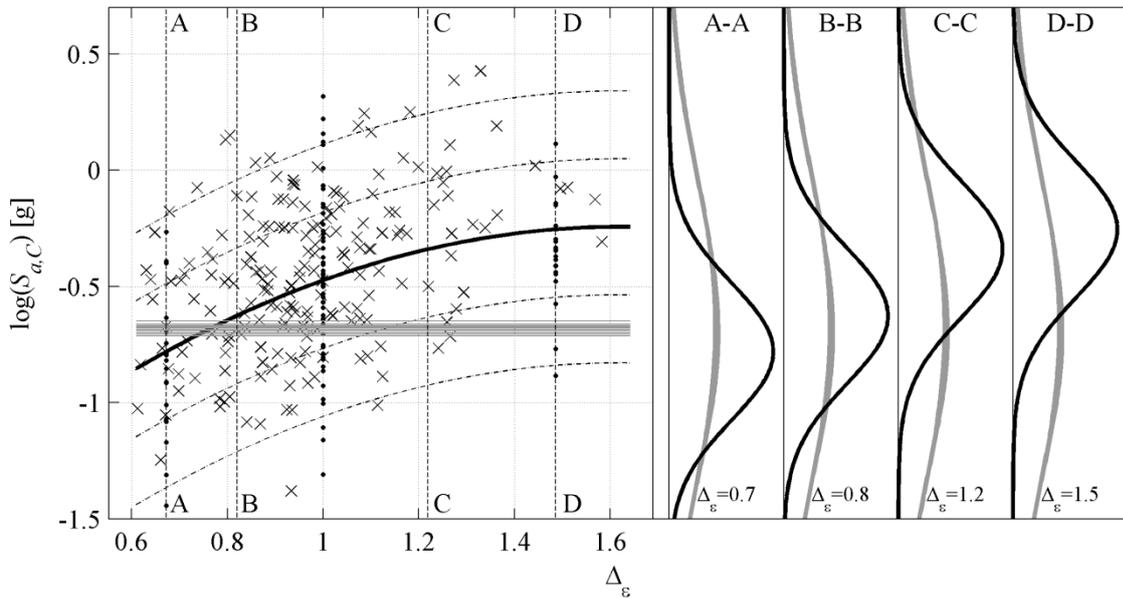


Figure 7.20. Linear models: sections (with  $f_{cm}, f_y, q_L$  equal to their mean values) of the mean response surfaces predicted by  $Q_{fix}$ -RND (thick black lines) and C-F2-9 (thin gray lines) models. xxx: capacity values from *reduced-MC* simulation and ●● from the CCD used for RS calibration.

## 7.7 Fragility analysis

Using the RS models to express the structural capacity, the approximated limit state function (Eq. (7.2)) for the structure can be expressed in analytical form as:

$$\begin{aligned}
 g(\mathbf{x}_E, \boldsymbol{\beta}, \varepsilon, \delta_{sis}, \delta_{\mu_c} | S_{a,D}) &= \ln(S_{a,C}) - \ln(S_{a,D}) = \\
 &= \mathbf{f}(\mathbf{x}_E)^T \boldsymbol{\beta} + \delta_{sis} + \delta_{\mu_c} + \varepsilon - \ln(S_{a,D})
 \end{aligned}
 \tag{7.8}$$

Hence,  $g$  in Eq. (7.8) is a function of the explicit random variables  $\mathbf{x}_E$ , distributed according to the values reported in Table 2, and of  $\varepsilon$ ,  $\delta_{sis}$ ,  $\delta_{\mu_c}$ , representing the error in the RS and the random effect of implicit variables, normally distributed (see Section 6.2). In order to evaluate the conditional probability of the structural failure ( $P_f$ ) for different levels of structural demand (i.e., the seismic fragility curve), the Monte Carlo simulation method can be applied to Eq. (7.8) with a minimum computational effort.

In order to compare the fragility curves obtained by different RSs, some reference values of  $P_f$  for different values of  $S_{a,D}$  have been evaluated using a full Monte Carlo approach. For each value of  $S_{a,D}$ , the corresponding  $P_f$  is calculated by performing a sequence of non-linear dynamic analyses (simulations), with independent random sampling of both explicit and implicit variables. These results are indicated with circles in the fragility graphs reported in the following and are used as the reference solution for the comparison between fragility curves obtained by different RS models. The 68% confidence intervals for the values of the fragilities estimated by MC simulations have been depicted with a thick vertical interval.

Confidence intervals can also be calculated for the fragility curves derived from response surfaces, starting from the variance-covariance matrix of the estimates of the regression parameters. In fact this matrix can be thought as a measure of the confidence of the estimate of the mean surface. Using a Monte Carlo simulation approach, it is therefore possible to compute confidence intervals for fragility curves. As an example mean,  $\pm\sigma$  and  $\pm 2\sigma$  fragility curves for  $Q_{fix}$ -RND are given in Figure 7.21. The spread of these curves can be considered a measure of the goodness of fit of the RS model. In the Q-RND model, the goodness of fit is generally very high and decreases as the distance from the centre of the design increases (See Chapter 6).

In the following, fragility curves obtained by different RS models will be compared and in order to do not make the figures too complicated the confidence intervals will be omitted and plotted for the Monte Carlo reference solution only. As for the fragility curves obtained from RS models, only the mean curves will be plotted.

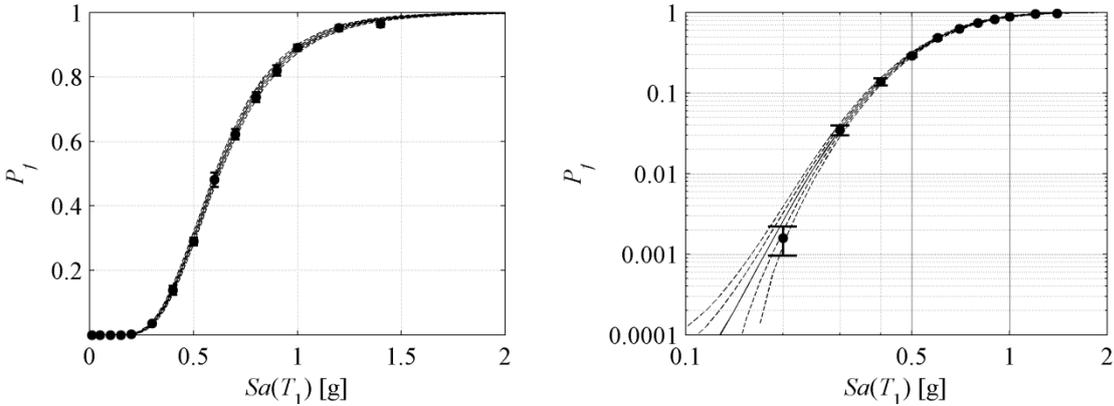


Figure 7.21. Mean and  $\pm 1$  and  $\pm 2$  standard deviation fragility curves for  $Q_{fix}$ -RND model. Dots indicate reference values, and their 68% confidence intervals, obtained by full MC simulation.

### 7.7.1 Quadratic models

Figure 7.22 shows the fragility curves obtained using different quadratic Q-CCD-9 models (gray lines) together with their envelope (black lines). These models give fragility curves in good agreement with the results of MC simulation. Nevertheless there is a lack of accuracy as far as the low fractiles are concerned. This portion on the fragility curves is usually the most important in structural reliability problems because it is related to low  $S_{a,D}$  values. In fact, when the seismic risk has to be evaluated, the fragility curves must be combined with the seismic hazard curve of the considered site which is higher for low spectral accelerations (Pinto et al., 2004).

The error of the different Q-CCD-9 models in the prediction of the value of  $S_{a,D}$  corresponding to  $P_f = 10^{-3}$  is between 5% and 30%. This circumstance can be expected because the tails of the fragility curve are related with the tails of the probability density functions of the explicit random variables collected in  $\mathbf{x}_E$  and, as Figures 7.4-7.7 clearly show, the error of Q-CCD-9 models on the prediction of the mean surface increases with the distance from the mean value of the explicit random variables,  $\mathbf{x}_E$ .

The envelopes and the fragility curves obtained using different Q-CCD-5 and Q-CCD-2 are reported in Figure 7.23 and Figure 7.24, respectively. In this case, i.e. adopting few replications of CCD, the fragility curves are very scattered and usually not in agreement with the reference results. Figures 7.8-7.9 clearly show that this error depends on the low accuracy in the prediction of both the mean value and the variance of the structural capacity.

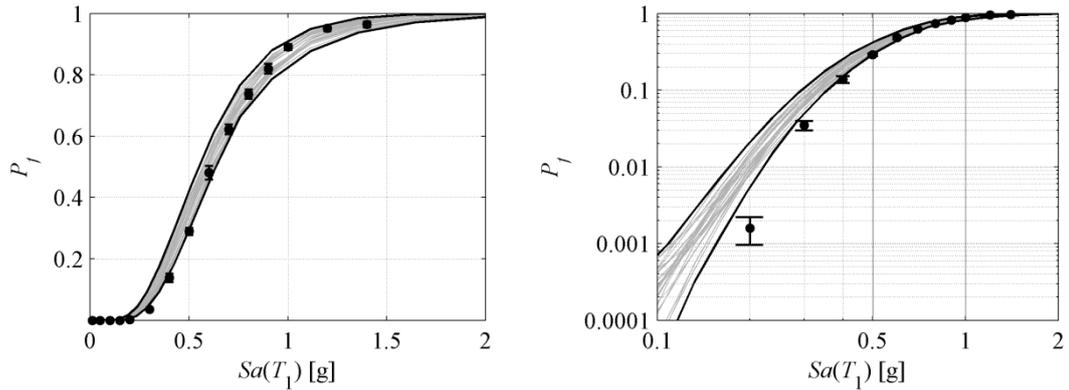


Figure 7.22. Q-CCD-9 models: fragility curves (gray lines) and their envelopes (black lines) obtained, in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

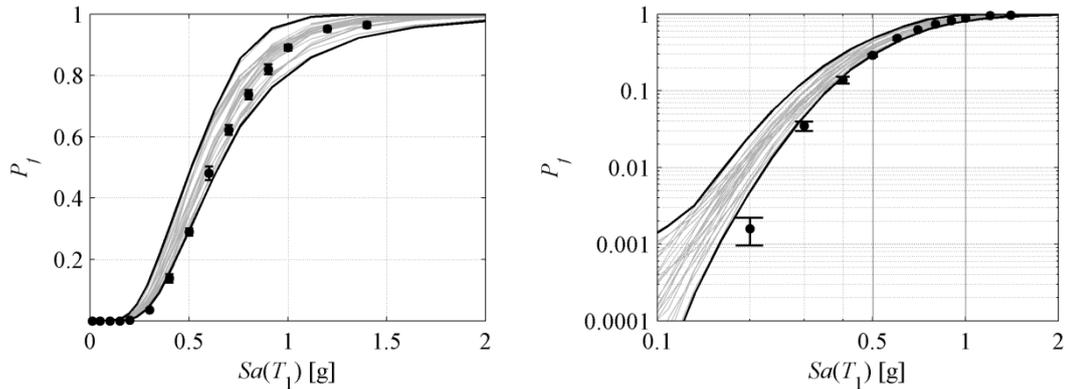


Figure 7.23. Q-CCD-5: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

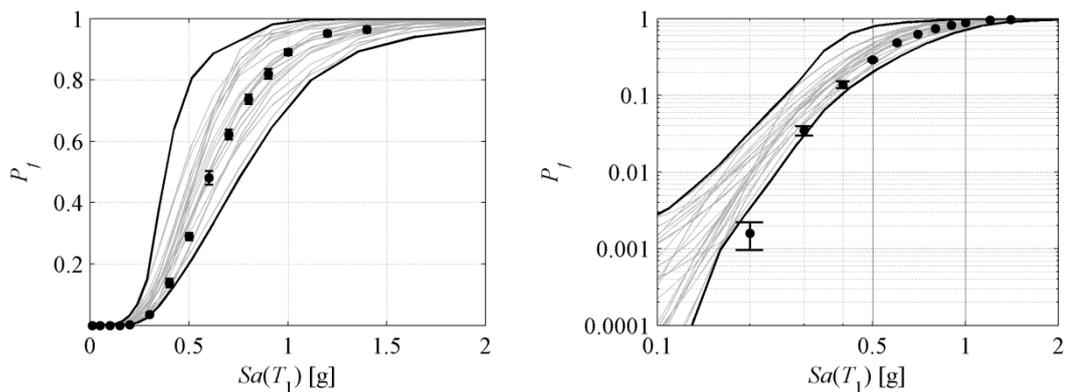


Figure 7.24. Q-CCD-2: fragility curves (gray lines) and their envelopes (black lines) obtained by Q-CCD-2 models, in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

### 7.7.2 Linear models

Figure 7.25 and Figure 7.26 show the fragility curves obtained by adopting L-CCD-9 and L-CCD-2 models, respectively. In the first case, the fragility curves are in good agreement with MC simulation for both tails and median. In the second case, the dispersion of the curves is much higher and therefore the sensitivity to the samples of implicit variables used is higher.

Figure 7.27 and Figure 7.28 show the fragility curves obtained through L-F1-9 and L-F1-2 models. The accuracy of fragility curves derived from L-F1-9 is good for very low spectral acceleration values but is poor for medium/high spectral acceleration values. Reducing the number of repetitions of the design (L-F1-2), increases the dispersion of the curves. Comparison of Figure 7.24 to Figure 7.28 confirms the higher sensitivity of Q-CCD models to the number of repetitions of the design.

Finally Figure 7.29 and Figure 7.30 depict the fragility curves given by L-F2-9 and L-F2-2 models. According to the conclusions drawn for Figure 7.19 and Figure 7.20, these models give poor approximations of structural capacity dependence on explicit variables and therefore, fragility curves with low accuracy.

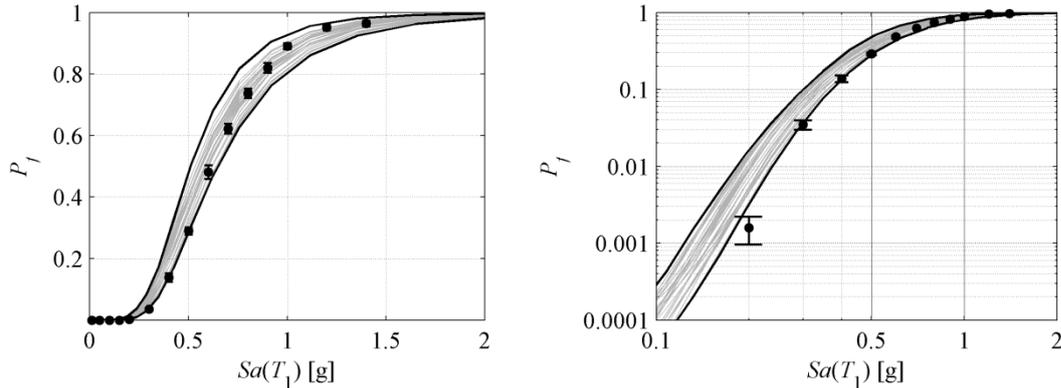


Figure 7.25. L-CCD-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

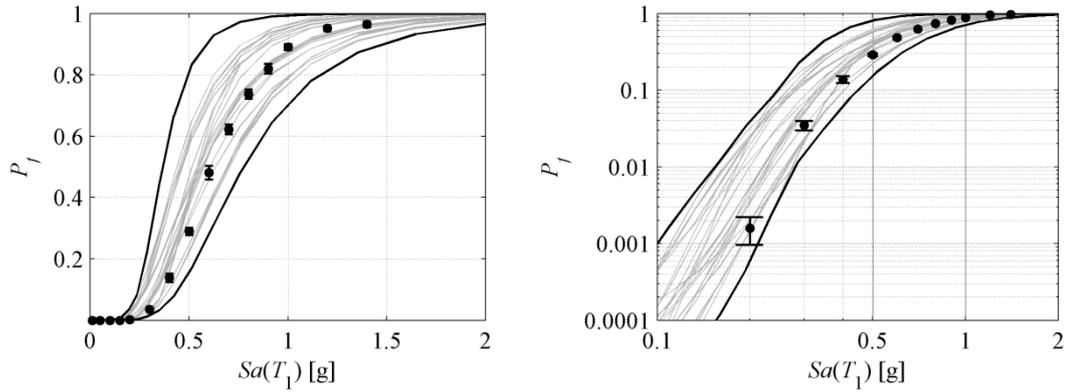


Figure 7.26. L-CCD-2: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

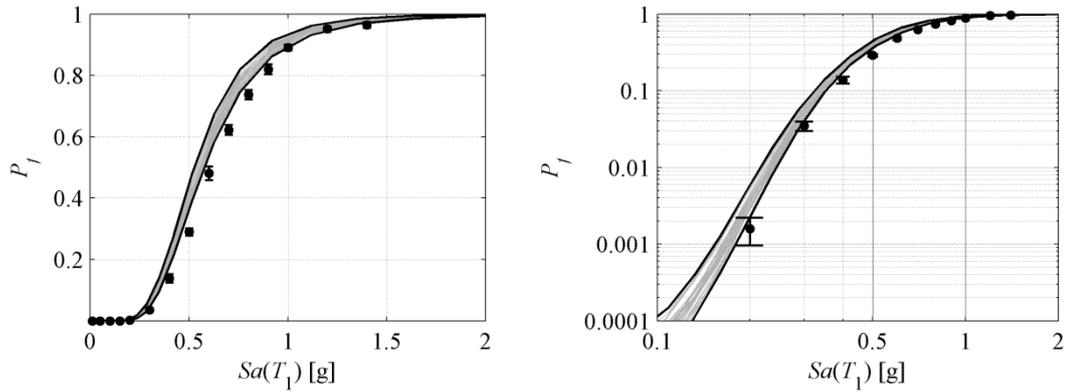


Figure 7.27. L-F1-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

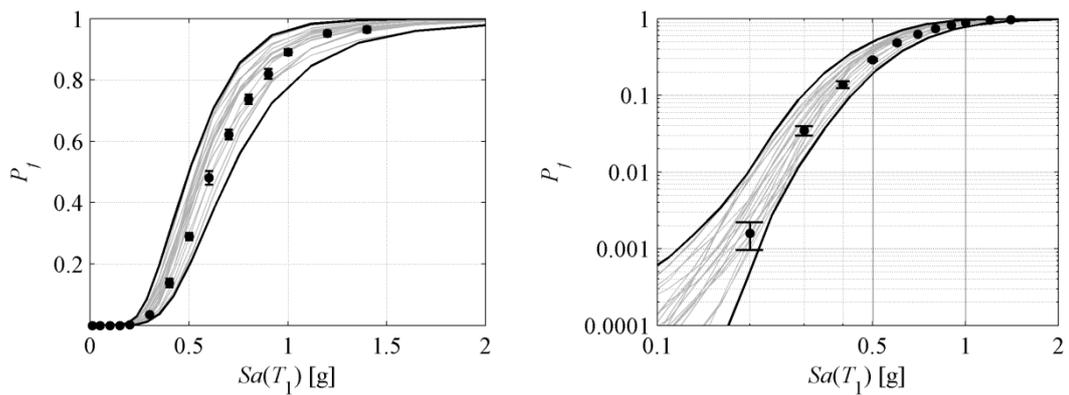


Figure 7.28. L-F1-2: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

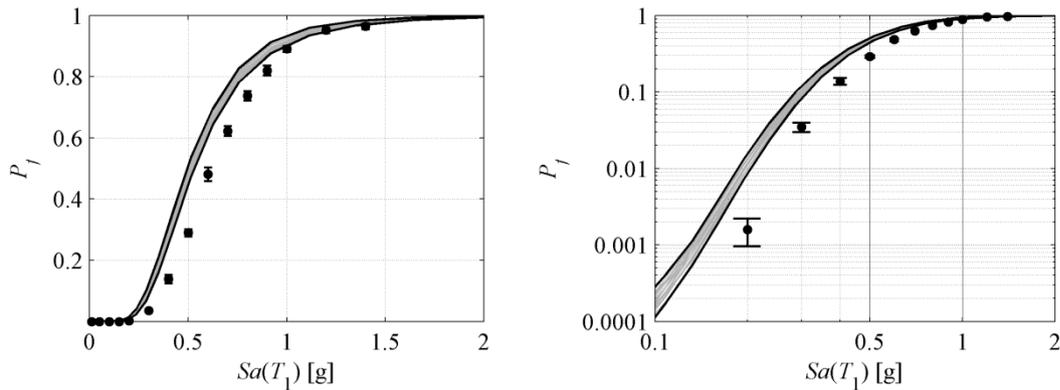


Figure 7.29. L-F2-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

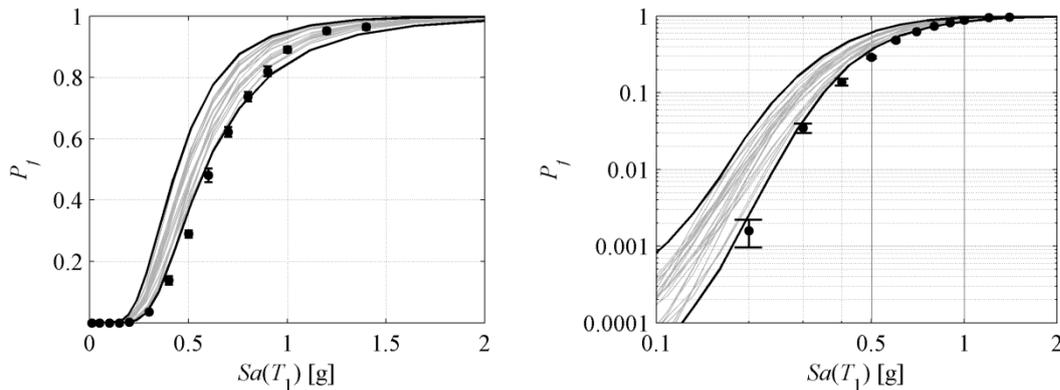


Figure 7.30. L-F2-2: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

### 7.7.3 Constant models

Figure 7.31 and Figure 7.32 depict the fragility curves obtained using C-CCD-9 and C-CCD-2 models, respectively. C-CCD-9 models give fragility curves in good agreement with MC simulation results, with low sensitivity to the number of repetitions of the design. The sensitivity of C-CCD-2 models is much higher.

Figure 7.33 and Figure 7.34 show the results obtained using C-F1-9 and C-F1-2 models. As previously concluded for L-F1 models (Figure 7.27 and Figure 7.28) these models give fragility curves with good accuracy for low spectral acceleration values but with very low accuracy for medium/high spectral acceleration values.

Finally Figure 7.35 shows fragility curves obtained from C-F2-9 models. These curves have an extremely low accuracy for all spectral acceleration values, because C-F2-9 models give very poor prediction of structural capacity (see Figure 7.20).

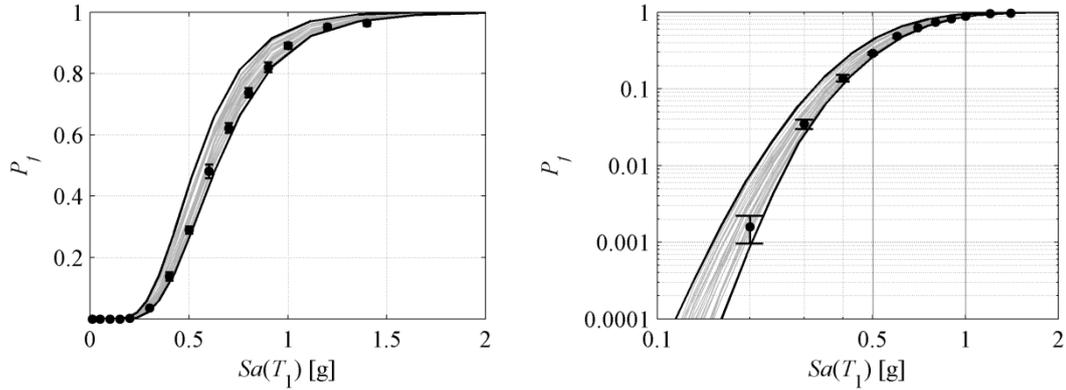


Figure 7.31. C-CCD-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

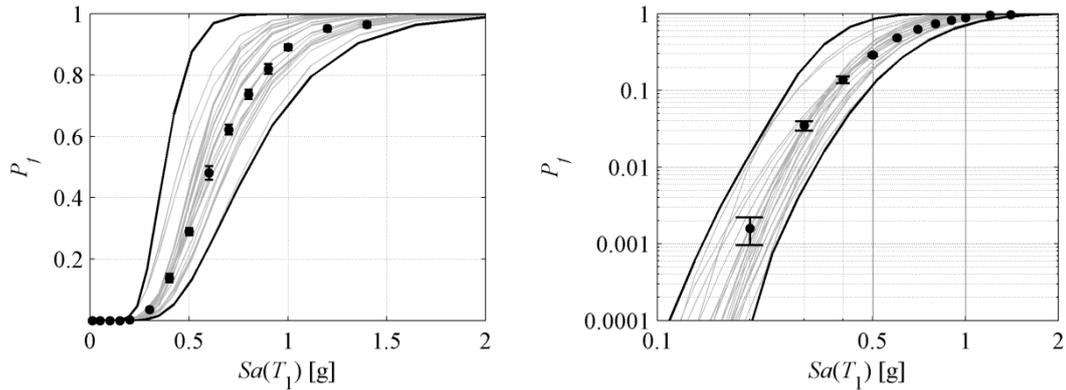


Figure 7.32. C-CCD-2: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

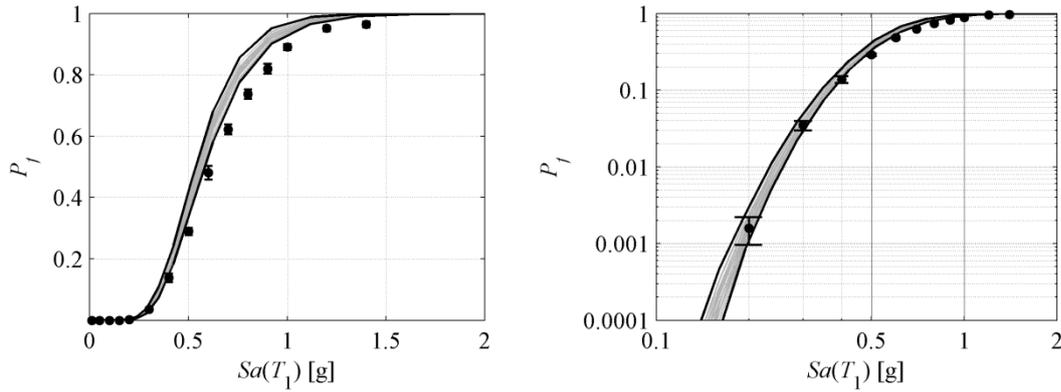


Figure 7.33. C-F1-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

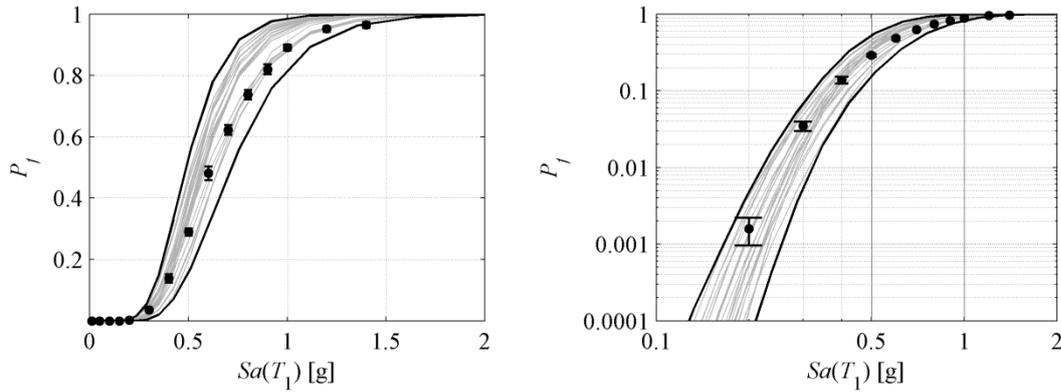


Figure 7.34 . C-F1-2: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

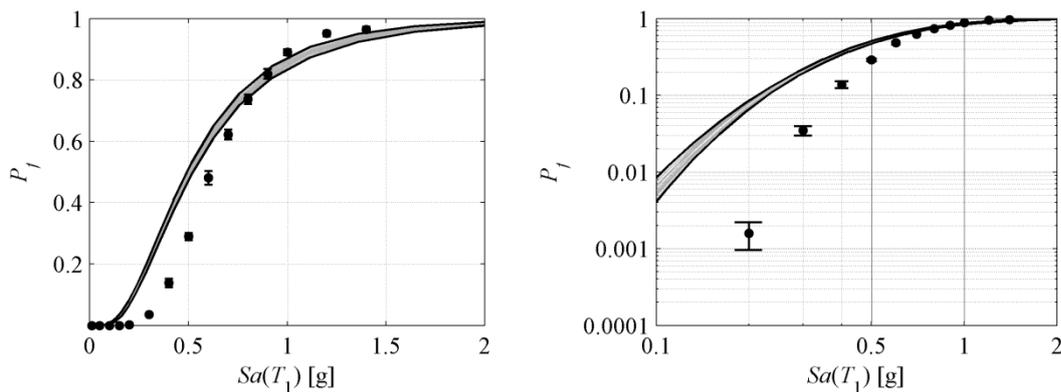


Figure 7.35. C-F2-9: fragility curves (gray lines) and their envelopes (black lines), in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation with  $\pm\sigma$  confidence intervals.

### 7.7.4 Comparison of the fragility curves

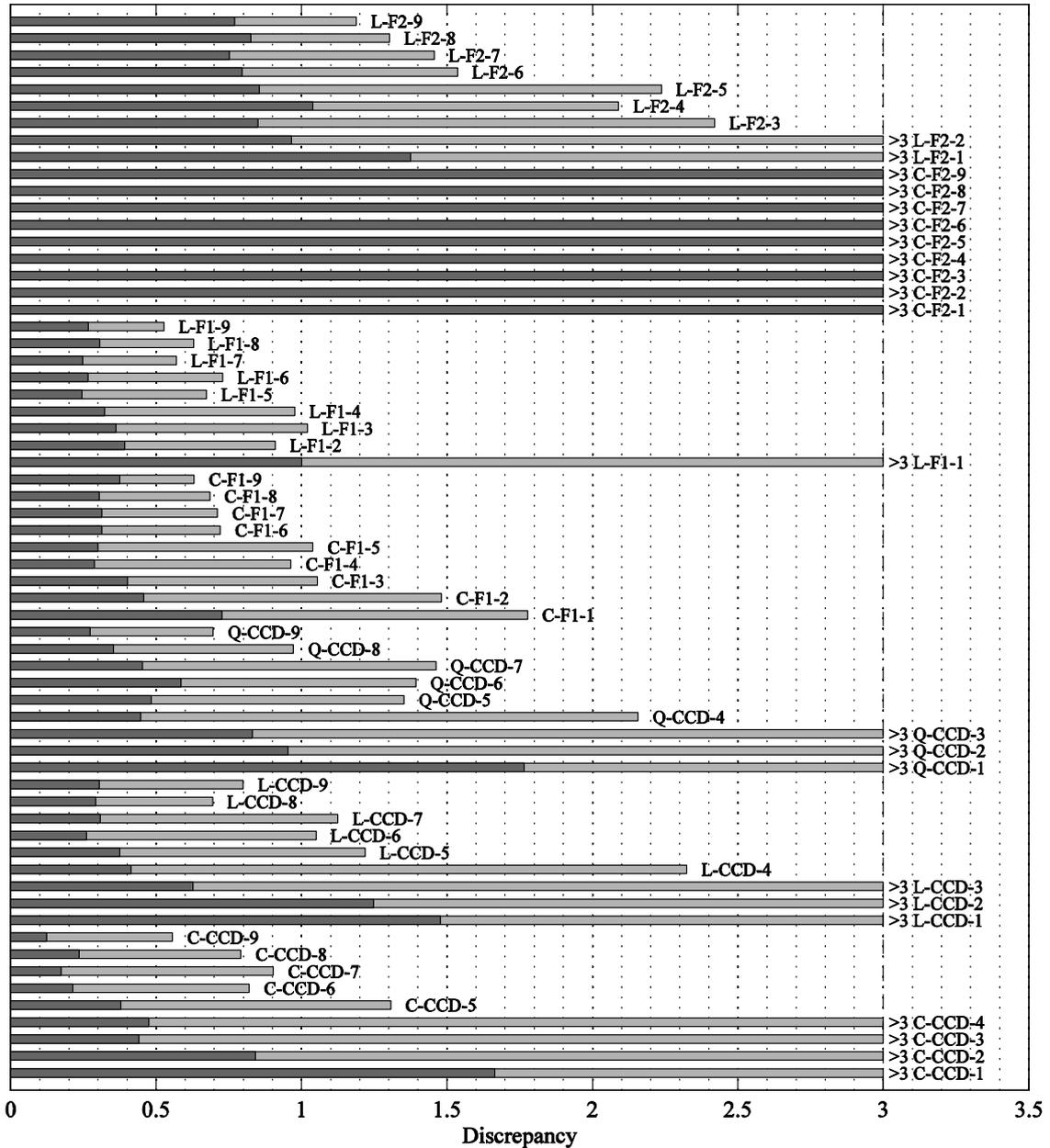


Figure 7.36. Discrepancy Index, DI, (see Eq. (7.9)) for the complete response surface models. Light gray bars indicate DI discrepancy and dark gray bars indicate mean DI.

In order to compare the fragility curves obtained through the different RS models considered, a measure of their discrepancy with respect to the fragility values obtained by MC simulation has been defined. Considering that MC failure probabilities have been calculated at  $n_{MC}$  spectral acceleration values. The following definition of Discrepancy Index, DI, is adopted:

$$DI = \sum_i^{n_{MC}} \left( \frac{(P_{RS}(Sa_i) - P_{MC}(Sa_i))^2}{P_{MC}(Sa_i)(1 - P_{MC}(Sa_i))} \right) \quad (7.9)$$

where  $P_{RS}(Sa_i)$  and  $P_{MC}(Sa_i)$  indicate the probability of failure at spectral acceleration  $Sa_i$  given by MC and RS models, respectively. The definition in Eq. (7.9) is a measure of the distance between some selected points of two distributions, defined as weighted sum of squares. It has been derived from the Anderson's test (Anderson and Darling, 1952; Stephens, 1974): a statistical test of whether there is evidence that a given sample of data did not arise from a given probability distribution, which has good sensitivity at the tails and at the median of distributions.

Figure 7.36 shows the values of DI in Eq. (7.9) for all the RS models considered. Light gray bars show the maximum value of discrepancy among all the considered RS models with the same order and design and dark gray bars the average value. The difference between these two bars is a measure of the sensitivity of the results to the specific samples of the implicit variables (i.e. accelerograms and concrete strength distributions) that are used to calibrate the model. In fact, when the number of repetitions of the designs increases, and therefore more realizations of implicit variables are used, the spread of the fragility curves reduces (compare, for example, Figure 7.23 with Figure 7.24) and therefore the difference between the dark- and the light-gray bars in Figure 7.36 reduces as well.

According to the results given in Figure 7.29 and Figure 7.30 the models with the highest DI (i.e. lowest accuracy) in Figure 7.36 are those fitted using the F2 designs. In fact, even increasing the repetitions of designs, and therefore the number of ground-motions considered, accuracy does not increase. Figure 7.36 suggests that for the other models, more than 4 repetitions of the design seem to be required to achieve a good accuracy. Furthermore, the accuracy of L-F1 models is stable from 5 to 9 repetitions and of C-F1 models from 6-9 repetitions. Similar conclusions can be drawn for C-CCD models, but in this case the maximum DI is slightly higher than that of C-F1 models. Q-CCD models are more sensitive than L-F1 models to the number of repetitions of the design and therefore even for 7 repetitions of the design the maximum discrepancy is still high. Counter-intuitively, according to Figure 7.36, the model with the highest average accuracy (lowest DI) is C-CCD-9. This model does not consider any explicit dependence of structural capacity on the explicit variables.

This result may be explained considering that only complete RS models have been considered, i.e. significance test have not been performed on the terms of the polynomial models. Consequently low-significance terms can decrease the accuracy of the results especially so far as low fractiles are concerned. In fact, these points are usually dependant

on response surface values far from the centre of the design. For example, considering Figure 7.6 one can observe that the dependence of structural capacity on steel yield stress,  $f_y$ , is very low. For this reason, a polynomial model with a quadratic term for  $f_y$  will give a very poor approximation especially for point far from the centre of the design.

In order to address this issue and better compare the designs, for each design models with significant terms only have been fitted. These models, referred to as most significant models in the following, have been defined as those having only the terms which are significant according to significance tests (see Section 6.2.4) with a confidence level of 5%. It should be noted that the number of significant terms may be different for each design and even for different repetitions of the same design. In fact, when the number of repetitions increases and therefore more information is provided, higher-order terms begin to become more significant. For CCD design, the highest-degree-model is complete quadratic polynomial. For F1 and F2 designs, pure quadratic terms are excluded and therefore only first order and interaction terms are considered. Figure 7.37 shows the DI for the most significant models. Of course the models in Figure 7.37 are less than those in Figure 7.36 because when considering the most significant models there is only one model for each design. Therefore the first letter of the codes previously used to refer to models (e.g. Q-CCD-9) is not anymore needed (e.g. CCD-9). Figure 7.37 confirms some of the conclusions drawn commenting Figure 7.36. Models fitted using F2 designs give the worst performances: even increasing the number of repetitions of the design their average discrepancy is high. Fragility curves calculated from CCD-9 models have the highest accuracy, and are depicted in Figure 7.38. It should be noted that the most significant CCD-9 model has a lower DI than the C-CCD-9 model (see Figure 7.36). The fragility curves calculated from CCD designs are more sensitive to the number of design repetitions than those calculated from F1 designs.

These results highlight the importance of performing significance tests on response surface models, because non-significant terms can reduce the accuracy of the models to a degree proportional to the distance from the centre of the design of experiments.

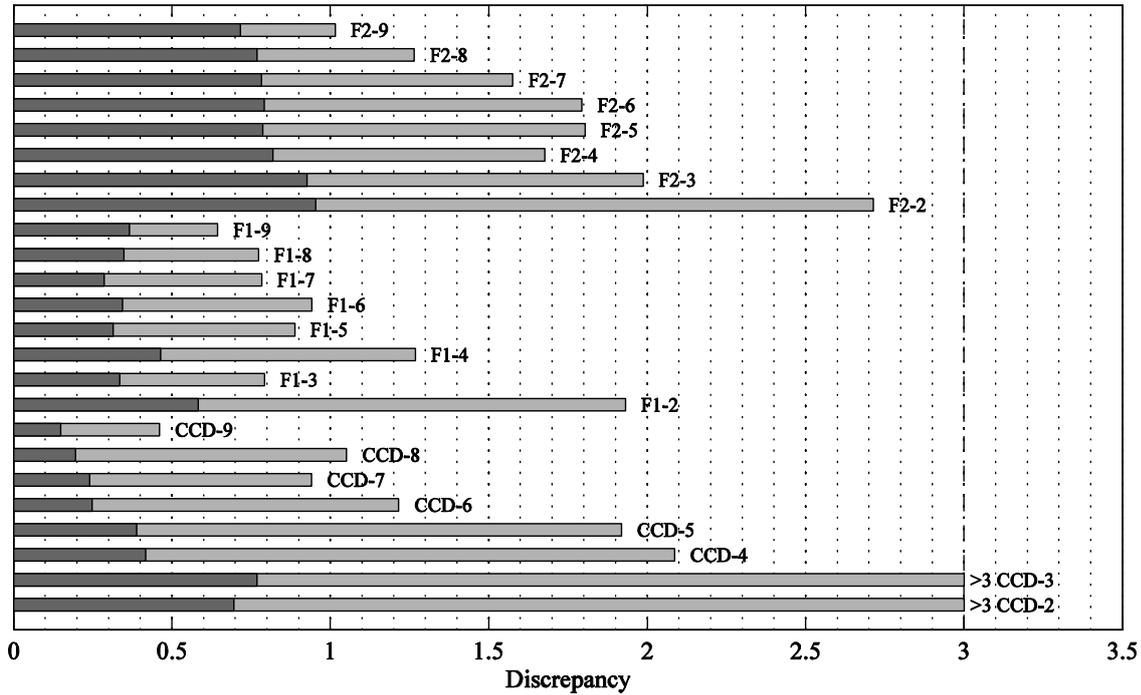


Figure 7.37. Discrepancy Index, DI, (see Eq. (7.9)) for the complete response surface models. Light gray bars indicate DI discrepancy and dark gray bars indicate mean DI

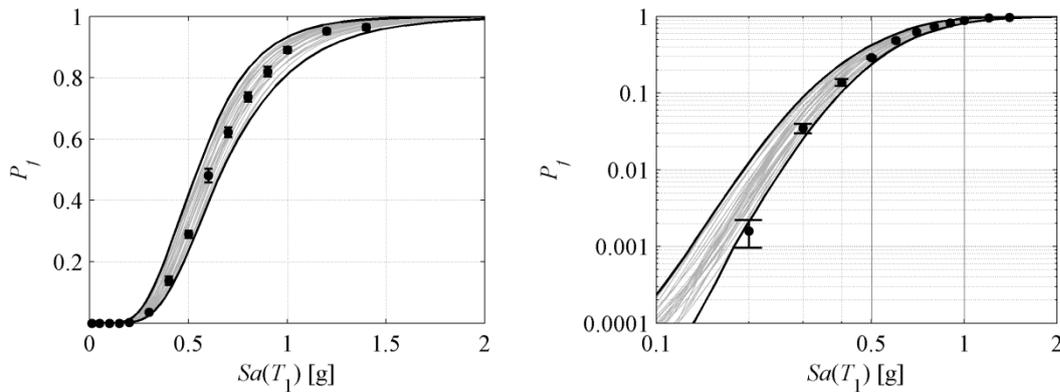


Figure 7.38. Fragility curves (gray lines) and their envelopes (black lines) obtained by CCD-9 most significant model, in linear scale (left panel) and log-scale (right panel). ●●●: reference values obtained by full MC simulation.

## 7.8 Conclusions

In the present study, a RS-based method has been developed to evaluate the seismic fragility curve of a RC frame structure. The variability of the material properties and live loads have been taken into account explicitly in the adopted RS model, whereas the

variability of ground-motions and concrete strength among individual members (beams/columns) of the structure have been taken into account implicitly via random factors. In order to obtain a statistically homogeneous sample, artificial ground-motions, generated by the method proposed by Sabetta and Pugliese Sabetta and Pugliese, 1996, have been adopted.

Different combinations of RS models and simulation plans have been adopted in order to find a good compromise between the accuracy of the results and the computational effort. As far as the RS models are concerned, quadratic and linear RS have been adopted. As for the design of the simulation plans, CCD and  $2^k$  factorial designs have been investigated. The results obtained through these models have been compared with those given by full MC simulations.

It is shown that quadratic polynomial models are very sensitive to the design considered. Moreover, from the present study, a minimum number of 9 different ground-motion time-histories is suggested in order to include in the model the ground-motion variability with sufficient accuracy. Reduced designs can produce completely wrong results. Particular care must be used if, as in reliability analysis, low fractiles of the fragility curve must be evaluated. It is shown that linear models are slightly less accurate in approximating the dependence of the structural capacity on the explicit variables, but are less sensitive to the design adopted and can be, in some cases, preferable with respect to quadratic models.



# References

- Abrahamson, N. A. (2000a). "Effects of rupture directivity on probabilistic seismic hazard analysis." 6th International conference on seismic zonation, Palm Springs, CA.
- Abrahamson, N. A. (2000b). "State of the practice of seismic hazard evaluation." GeoEng 2000, Melbourne, Australia.
- Abrahamson, N. A. (2006). "Seismic hazard assessment: problems with current practice and future developments." First European conference on earthquake engineering and seismology, Geneva, Switzerland.
- Abrahamson, N. A., Birkhauser, M., Koller, D., 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, UK.
- Abrahamson, N. A., and Bommer, J. J. (2005). "Probability and Uncertainty in Seismic Hazard Analysis." *Earthquake Spectra*, 21(2), 603-607.
- Abrahamson, N. A., and Silva, W. J. (1997). "Empirical Response Spectral Attenuation Relations for Shallow Crustal Earthquakes." *Seismological Research Letters*, 68(1), 94-127.
- Abrahamson, N. A., and Silva, W. J. (2008a). "Abrahamson & Silva NGA ground motion relations for the geometric mean horizontal component of peak and spectral ground motion parameters." Report, Final report prepared for the Pacific Earthquake Engineering Research Center, February 2008.
- Abrahamson, N. A., and Silva, W. J. (2008b). "Summary of the Abrahamson & Silva NGA Ground-Motion Relations." *Earthquake Spectra*, 24(1), 67-97.
- Abrahamson, N. A., and Somerville, P. (1996). "Effects of the hanging wall and footwall on ground motions recorded during the Northridge earthquake." *Bulletin of the Seismological Society of America*, 86(1), S93-S99.
- Abrahamson, N. A., and Youngs, R. R. (1992). "A stable algorithm for regression analyses using the random effects model." *Bulletin of the Seismological Society of America*, 82(1), 505-510.
- Aki, K., and Richards, P. G. (1980). *Quantitative Seismology: Theory and Methods. Vol. 1.*, W. H. Freeman and Co., New York.
- Akkar, S., and Bommer, J. J. (2006). "Influence of long-period filter cut-off on elastic spectral displacements." *Earthquake Engineering & Structural Dynamics*, 35(9), 1145-1165.
- Akkar, S., and Bommer, J. J. (2007a). "Empirical prediction equations for peak ground velocity derived from strong-motion records from Europe and Middle East." *Bulletin of the Seismological Society of America*, 97(1), 511-530.
- Akkar, S., and Bommer, J. J. (2007b). "Prediction of elastic displacement response spectra in Europe and at the Middle East." *Earthquake Engineering & Structural Dynamics*, 36, 1275-1301.
- Akkar, S., and Özen, Ö. (2005). "Effect of peak ground velocity on deformation demands for SDOF systems." *Earthquake Engineering & Structural Dynamics*, 34(13), 1551-1571.
- Ambraseys, N. N., Douglas, J., Rinaldis, D., Berge-Thierry, C., Suhadolc, P., Costa, G., Sigbjörsson, R., and Smit, P. (2004). "Dissemination of European strong-motion

- data, Vol. 2." CD-ROM Collection, Engineering and Physical Sciences Research Council, United Kingdom.
- Ambraseys, N. N., Douglas, J., Sarma, S. K., and Smit, P. M. (2005). "Equations for the Estimation of Strong Ground Motions from Shallow Crustal Earthquakes Using Data from Europe and the Middle East: Horizontal Peak Ground Acceleration and Spectral Acceleration." *Bulletin of Earthquake Engineering*, 3, 1-53.
- Ambraseys, N. N., Simpson, K. A., and Bommer, J. J. (1996). "The prediction of horizontal response spectra in Europe." *Earthquake Engineering & Structural Dynamics*, 25, 371-400.
- Anderson, T. W., and Darling, D. A. (1952). "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes." *Annals of Mathematical Statistics*, 23(2), 193-212.
- Ang, G. L., Ang, A. H.-S., and Tang, W. H. (1992). "Optimal importance sampling density estimator." *Journal of Engineering Mechanics*, 118(6), 1146-1163.
- Aptikaev, F. F. (1982). "On the correlations of MM Intensity with parameters of ground shaking." 7th European Conference on Earthquake Engineering, Athens, Greece, 95-97.
- Araya, R., and Saragoni, G. R. (1980). "Capacidad de los movimientos sismicos de producir daño estructural." Report, Division of Structural Engineering, Department of Engineering, University of Chile, Santiago, Chile.
- Arias, A. (1970). "A measure of Earthquake Intensity." Seismic Design for Nuclear Power Plants, R. J. Hansen, ed., Massachusetts Institute of Technology Press, Cambridge, Massachusetts 438-483.
- ATC. (1985). "Earthquake damage evaluation data for California." Report, Applied Technology Council, Redwood City.
- Atkinson, G. M., and Somerville, P. G. (1994). "Calibration of time history simulation methods." *Bulletin of the Seismological Society of America*, 84(2), 400-414.
- Au, S. K., and Beck, J. L. (1999). "A new adaptive importance sampling scheme for reliability calculations." *Structural Safety*, 21, 135-158.
- Au, S. K., and Beck, J. L. (2001). "Estimation of small failure probabilities in high dimensions by subset simulation." *Probabilistic Engineering Mechanics*, 16, 263-277.
- Au, S. K., and Beck, J. L. (2003). "Subset Simulation and its Application to Seismic Risk Based on Dynamic Analysis." *Journal of Engineering Mechanics*, 129(8), 901-917.
- Bae, S., and Bayrak, O. (2008). "Plastic Hinge Length of Reinforced Concrete Columns." *ACI Structural Journal*, 105(3), 290.
- Baker, A. L. L. (1956). *Ultimate load theory applied to the design of reinforced concrete frames*, Concrete Publications Ltd., London, UK.
- Baker, J. W. (2005). "Vector-valued ground motion intensity measures for probabilistic seismic demand analysis," University of Stanford, Stanford, CA, U.S.
- Baker, J. W. (2007a). "Correlation of ground motion intensity parameters used for predicting structural and geotechnical response." 10th International Conference on Application of Statistic and Probability in Civil Engineering (ICASP10), Tokyo, JP.
- Baker, J. W. (2007b). "Measuring bias in structural response caused by ground motion scaling, in Proceedings." 8th Pacific Conference on Earthquake Engineering, Nanyang Technological University, Singapore.
- Baker, J. W. (2007c). "Probabilistic structural response assessment using vector-valued intensity measures." *Earthquake Engineering & Structural Dynamics*, 36(13), 1861-1883.

- Baker, J. W. (2007d). "Quantitative Classification of Near-Fault Ground Motions." *Bulletin of the Seismological Society of America*, 97(5), 1486-1501.
- Baker, J. W., and Cornell, C. A. (2004). "Choice of a Vector Ground Motion Intensity Measures for Seismic Demand Hazard Analysis." 13th World Conference on Earthquake Engineering, Vancouver, Canada.
- Baker, J. W., and Cornell, C. A. (2005). "A vector-valued ground motion intensity measure consisting of spectral acceleration and epsilon." *Earthquake Engineering & Structural Dynamics*, 34(10), 1193 - 1217.
- Baker, J. W., and Cornell, C. A. (2006a). "Correlation of response spectral values for multicomponent ground motions." *Bulletin of the Seismological Society of America*, 96(1), 215 - 27.
- Baker, J. W., and Cornell, C. A. (2006b). "Spectral shape, epsilon and record selection." *Earthquake Engineering & Structural Dynamics*, 35(9), 1077-1095.
- Baker, J. W., and Cornell, C. A. (2006c). "Which Spectral Acceleration Are You Using?" *Earthquake Spectra*, 22(2), 293-312.
- Baker, J. W., and Jayaram, N. (2008). "Correlation of Spectral Acceleration Values from NGA Ground Motion Models." *Earthquake Spectra*, 24(1), 299-317.
- Barenberg, M. E. (1989). "Inelastic Response of a Spectrum-Compatible Artificial Accelerogram." *Earthquake Spectra*, 5(3), 477-493.
- Bazant, Z. P., and Planas, J. (1998). *Fracture and size effect in concrete and other quasi brittle materials*, CRC Press, Boca Raton, Florida.
- Bazzurro, P. (1998). "Probabilistic Seismic Demand Analysis," Stanford University, Stanford.
- Bazzurro, P., and Cornell, C. A. (1994). "Seismic hazard analysis of nonlinear structures. I: Methodology." *Journal of Structural Engineering*, 120(11), 3320-3344.
- Bazzurro, P., and Cornell, C. A. (1999). "Disaggregation of seismic hazard." *Bulletin of the Seismological Society of America*, 89(2), 501-520.
- Bazzurro, P., and Cornell, C. A. (2002). "Vector-valued probabilistic seismic hazard analysis (VPSHA)." 7th U.S. National Conference on Earthquake Engineering, Boston, Massachusetts.
- Bazzurro, P., Cornell, C. A., Shome, N., and Carballo, J. E. (1998). "Three proposals for characterizing MDOF nonlinear seismic response." *Journal of Structural Engineering*, 124(11), 1281-1289.
- Bazzurro, P., Sjöberg, B., Luco, N., Silva, W. J., and Darragh, R. (2004). "Effects of Strong Motion Processing Procedures on Time Histories, Elastic and Inelastic Spectra." Workshop on Strong Motion Record Processing (COSMOS).
- Bender, B., and Campbell, K. W. (1989). "Selection of minimum magnitude for use in seismic hazard analysis." *Bulletin of the Seismological Society of America*, 79(1), 199-204.
- Bender, B., and Perkins, D. M. (1982). "SEISRISK II, A computer program for seismic hazard estimation." Report, US Geological Survey.
- Bender, B., and Perkins, D. M. (1987). "SEISRISK III, A computer program for seismic hazard estimation." *US Geological Survey Bulletin*, 1772, 1-20.
- Benjamin, J. R., and Cornell, C. A. (1970). *Probability, statistics, and decision for civil engineers*, McGraw-Hill, New York.
- Beresnev, I. A., and Atkinson, G. M. (1997). "Modeling finite fault radiation from the  $\omega^2$  spectrum." *Bulletin of the Seismological Society of America*, 87, 67-84.
- Bertero, V. V., Herrera, R. A., and Mahin, S. A. (1976). "Establishment of design earthquakes - Evaluation of present methods." International Symposium on Earthquake Structural Engineering, St. Luis.

- Beyer, K., and Bommer, J. J. (2006). "Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion." *Bulletin of the Seismological Society of America*, 96(4 A), 1512 - 1522.
- Beyer, K., and Bommer, J. J. (2007). "Selection and Scaling of Real Accelerograms for Bi-Directional Loading: A Review of Current Practice and Code Provisions." *Journal of Earthquake Engineering*, 11(Supplement 1), 13-45.
- Bianchini, M., Cornell, C. A., and Baker, J. W. (in preparation). "An earthquake intensity measure based on an average of spectral accelerations."
- Bolt, B. A., and Abrahamson, N. A. (2003). "Estimation of strong motions." International handbook of earthquake engineering seismology, W. H. K. Lee, H. Kanamori, P. C. Jennings, C. Kisslinger, and C. London, eds., Academic Press, 81B: 983-1001.
- Bommer, J. J. (2002). "Deterministic vs. probabilistic seismic hazard assessment: An exaggerated and obstructive dichotomy." *Journal of Earthquake Engineering*, 6(Sa), 43-73.
- Bommer, J. J. (2003). "Uncertainty about the uncertainty in seismic hazard analysis." *Engineering Geology*, 70(1-2), 165-168.
- Bommer, J. J. (2006). "Re-thinking seismic hazard mapping and design return periods." First european conference on earthquake engineering and seismology, Geneva, Switzerland.
- Bommer, J. J., and Abrahamson, N. A. (2006). "Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates?" *Bulletin of the Seismological Society of America*, 96, 1967-1977.
- Bommer, J. J., and Abrahamson, N. A. (2007a). "Reply to 'Comment on 'Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates?' by Julian J. Bommer and Norman A. Abrahamson' by Jens-Uwe Klugel." *Bulletin of the Seismological Society of America*, 97(6), 2208-2211.
- Bommer, J. J., and Abrahamson, N. A. (2007b). "Reply to 'Comment on 'Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates?' by Julian J. Bommer and Norman A. Abrahamson' by Zhenming Wang and Mai Zhou." *Bulletin of the Seismological Society of America*, 97(6), 2215-2217.
- Bommer, J. J., Abrahamson, N. A., Strasser, F. O., Pecker, A., Bard, P. Y., Bungum, H., Cotton, F., Fah, 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., and Acevedo, A. B. (2004). "The use of real earthquake accelerograms as input to dynamic analysis." *Journal of Earthquake Engineering*, 8(Special Issue 1), 43-91.
- Bommer, J. J., Douglas, J., and Strasser, F. O. (2003). "Style-of-Faulting in Ground-Motion Prediction Equations." *Bulletin of Earthquake Engineering*, 1(2), 171-203.
- Bommer, J. J., Georgallides, G., and Tromms, I. J. (2001). "Is there a near-field for small-to-moderate magnitude earthquakes." *Journal of Earthquake Engineering*, 5(3), 395-423.
- Bommer, J. J., Hancock, J., and Alarcón, J. E. (2006). "Correlations between duration and number of effective cycles of earthquake ground motion." *Soil Dynamics and Earthquake Engineering*, 26(1), 1-13.
- Bommer, J. J., and Martinez-Pereira, A. (1999). "The effective duration of earthquake strong motion." *Journal of Earthquake Engineering*, 3, 127-172.

- Bommer, J. J., McQueen, C., Salazar, W., Scott, S., and Woo, G. (1998). "A case study of the spatial distribution of seismic hazard (El Salvador)." *Natural Hazards*, 18, 145-166.
- Bommer, J. J., and Ruggeri, C. (2002). "The Specification of Acceleration Time-histories in Seismic Design Codes." *European Earthquake Engineering*, 16(1), 3-17.
- Bommer, J. J., Scherbaum, F., Bungum, H., Cotton, F., Sabetta, F., and Abrahamson, N. A. (2005). "On the use of logic trees for ground-motion prediction equations in seismic-hazard analysis." *Bulletin of the Seismological Society of America*, 95(2), 377-389.
- Bommer, J. J., Scott, S. G., and Sarma, S. K. (2000). "Hazard-consistent earthquake scenarios." *Soil Dynamics and Earthquake Engineering*, 19(4), 219-231.
- Bommer, J. J., and Stafford, P. J. (2007). "Engineering Seismology. Course support material. Autumn term 2007.", Imperial College, London.
- Bommer, J. J., Stafford, P. J., Alarcón, J. E., and Akkar, S. (2007). "The influence of magnitude range on empirical ground-motion prediction." *Bulletin of the Seismological Society of America*, 97(6), 2152-2170.
- Boore, D. M. (1983). "Stochastic simulation of high-frequency ground motions based on seismological model of the radiated spectra." *Bulletin of the Seismological Society of America*, 73, 1865-1894.
- Boore, D. M. (2003). "Simulation of ground motion using the stochastic method." *Pure and applied geophysics*, 160, 635-676.
- Boore, D. M. (2005). "On Pads and Filters: Processing Strong-Motion Data." *Bulletin of the Seismological Society of America*, 95(2), 745-750.
- Boore, D. M., and Akkar, S. (2003). "Effect of causal and acausal filters on elastic and inelastic response spectra." *Earthquake Engineering & Structural Dynamics*, 32(11), 1729-1748.
- Boore, D. M., and Bommer, J. J. (2005). "Processing of strong-motion accelerograms: needs, options and consequences." *Soil Dynamics and Earthquake Engineering*, 25(2), 93-115.
- Boore, D. M., and Joyner, W. B. (1982). "The empirical prediction of ground motion." *Bulletin of the Seismological Society of America*, 72(6B), S43-60.
- Boore, D. M., Joyner, W. B., and Fumal, T. E. (1997). "Equations for estimating horizontal response spectra and peak ground acceleration from western North America earthquakes: a summary of recent work." *Seismological Research Letters*, 68(1), 128-153.
- Boore, D. M., Watson-Lamprey, J. A., and Abrahamson, N. A. (2006). "Orientation-Independent Measures of Ground Motion." *Bulletin of the Seismological Society of America*, 96(4A), 1502-1511.
- Box, G. E. P., and Draper, N. R. (1987). *Empirical model-building and response surfaces*, John Wiley & Sons, New York.
- Box, G. E. P., and Wilson, K. B. (1951). "On the experimental attainment of optimum conditions." *Journal of the royal statistical society*, B13, 1-45.
- Bozorgnia, Y., and Campbell, K. W. (2004). "The vertical-to-horizontal response spectral ratio and tentative procedures for developing simplified v/h and vertical design spectra." *Journal of Earthquake Engineering*, 8(2), 175-207.
- Bracci, J. M., Reinhorn, A. M., and Mander, J. B. (1995a). "Seismic resistance of reinforced concrete frame structures designed for gravity loads: performance of structural system." *ACI Structural Journal*, 92(5), 597-609.

- Bracci, J. M., Reinhorn, A. M., and Mander, J. B. (1995b). "Seismic retrofit of reinforced concrete buildings designed for gravity loads: performance of structural model." *ACI Structural Journal*, 92(6), 711-723.
- Brady, A. G., Trifunac, M. D., and Hudson, D. E. (1973). "Analyses of strong motion earthquake accelerograms - response spectra." Report, Earthquake Engineering Research Laboratory - California Institute of Technology.
- Bray, J. D., and Rodriguez-Marek, A. (2004). "Characterization of forward-directivity ground motions in the near-fault region." *Soil Dynamics and Earthquake Engineering*, 24(11), 815-828.
- Buratti, N., Ferracuti, B., and Savoia, M. (2006a). "Reliability of r.c. structures against seismic action: response surface approach." XVI Convegno italiano di meccanica computazionale (GIMC), Bologna, ITALY.
- Buratti, N., Ferracuti, B., and Savoia, M. (2006b). "Seismic fragility of existing RC structures by response surface method." First European Conference on Earthquake Engineering and Seismology, Geneva.
- Buratti, N., Ferracuti, B., and Savoia, M. (2006c). "Seismic risk assessment of R/C structures through response surface method." III European Conference on Computational Mechanics, Lisbon, Portugal.
- Buratti, N., Ferracuti, B., and Savoia, M. (2007a). "Recorded and simulated ground motion time histories for seismic fragility analysis of RC structures." XII Conference of the Associazione Nazionale Italiana di Ingegneria Sismica, Pisa.
- Buratti, N., Ferracuti, B., and Savoia, M. (2007b). "A response surface approach with random block effects for seismic fragility assessment of r.c. structures." 10th International conference on application of statistics and probability in civil engineering (ICASP), Tokyo, JP.
- Buratti, N., Ferracuti, B., and Savoia, M. (2007c). "Seismic fragility of RC frame structures using response surface with random block effects." XVIII Convegno associazione italiana di meccanica teoria ed applicata (AIMETA), Brescia, Italy.
- Campbell, K. W. (1985). "Strong motion attenuation relations: a ten-year perspective." *Earthquake Spectra*, 1(4), 759-804.
- Campbell, K. W. (1997). "Empirical near-source attenuation relationships for horizontal and vertical components of peak ground acceleration, peak ground velocity, and pseudo-absolute acceleration response spectra." *Seismological Research Letters*, 68(1), 154-179.
- Campbell, K. W. (2003). "Strong-motion attenuation relations." International Handbook of Earthquake Engineering and Seismology, W. H. K. Lee, H. Kanamori, P. C. Jennings, and C. Kisslinger, eds., Academic Press, London, UK, 1003-1012.
- Campbell, K. W., and Bozorgnia, Y. (2003a). "Updated near-source ground-motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra." *Bulletin of the Seismological Society of America*, 93(1), 314-331.
- Campbell, K. W., and Bozorgnia, Y. (2003b). "Updated near-source ground-motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra (vol 93, pg 314, 2003)." *Bulletin of the Seismological Society of America*, 93(4), 1872-1872.
- Campbell, K. W., and Bozorgnia, Y. (2004). "Updated near-source ground-motion (attenuation) relations for the horizontal and vertical components of peak ground acceleration and acceleration response spectra (vol 93, pg 314, 2003)." *Bulletin of the Seismological Society of America*, 94(6), 2417.
- Campedel, M., Cozzani, V., Buratti, N., Ferracuti, B., and Savoia, M. (2008). "Quantitative risk assessment of accidents induced by seismic events in

- industrial sites." 3rd International Conference on Safety and Environment in Process Industry, Rome, Italy.
- Carballo, J. E., and Cornell, C. A. (2000). "Probabilistic seismic demand analysis: spectrum matching and design." Report: *RMS-41*, Department of Civil and Environmental Engineering, Stanford University.
- Casciati, F., and Faravelli, L. (1991). *Fragility analysis of complex structural systems*, John Wiley & Sons, New York.
- Ceb-Fip. (1993). *Model Code 1990*, Redwood Books, Trowbridge.
- Celik, O. C. (2007). "Probabilistic Assessment of Non-Ductile Reinforced Concrete Frames Susceptible to Mid-America Ground Motions," Georgia Institute of Technology.
- Chang, T.-Y., Cotton, F., Tsai, Y.-B., and Angelier, J. (2004). "Quantification of hanging-wall effects on ground motion: Some insights from the 1999 Chi-Chi earthquake." *Bulletin of the Seismological Society of America*, 94(6), 2186-2197.
- Chapman, M. C. (1995). "A probabilistic approach to ground-motion selection for engineering design." *Bulletin of the Seismological Society of America*, 85(3), 937-942.
- Choi, Y. J., and Stewart, J. P. (2005). "Nonlinear site amplification as function of 30 m shear wave velocity." *Earthquake Spectra*, 21(1), 1-30.
- Choi, Y. J., Stewart, J. P., and Graves, R. W. (2005). "Empirical model for basin effects accords for basin depth and source location." *Bulletin of the Seismological Society of America*, 95(4), 1412-1427.
- Coleman, J., and Spacone, E. (2001). "Localization Issues in Force-Based Frame Elements." *Journal of Structural Engineering*, 127(11), 1257-1265.
- Cordova, P. P., Deierlein, G. G., Mehanny, S. S. F., and Cornell, C. A. (2001). "Development of a two-parameter seismic intensity measure and probabilistic assessment procedure." 2nd U.S.-Japan Workshop on Performance-Based Earthquake Engineering Methodology for Reinforced Concrete Building Structures, Sapporo, Hokkaido.
- Corley, W. G. (1966). "Rotational capacity of reinforced concrete beams." *Journal of the Structural Division*, 92(ST5), 121-146.
- Cornell, C. A. (1968). "Engineering seismic risk analysis." *Bulletin of the Seismological Society of America*, 58(5), 1583-1606.
- Cornell, C. A., Jalayer, F., Hamburger, R. O., and Foutch, D. A. (2002). "Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines." *Journal of Structural Engineering*, 128(4), 526 - 533.
- Cornell, C. A., and Vanmarcke, E. H. (1969). "The major influences on seismic risk." Fourth world conference on earthquake engineering, Santiago, Chile.
- Cornell, J. A., and Winterstein, S. R. (1988). "Temporal and magnitude dependence in earthquake recurrence models." *Bulletin of the Seismological Society of America*, 78(4), 1522-1537.
- Cosenza, E., and Manfredi, G. (1998). "A seismic design method including damage effect." 11th European Conference on Earthquake Engineering, Paris, France.
- De Borst, R., Feenstra, P. H., Pamin, J., and Sluys, L. J. (1994). "Some current issues in computational mechanics of concrete structures." Computer Modelling of Concrete Structures, H. Mang, N. Bićanić, and R. De Borst, eds., Swansea, UK.
- Delgado, R., Marques, M., Monteiro, R., Delgado, P., Romao, X., and Costa, A. (2006). "Setting up real or artificial earthquake records for dynamic analysis." 1st European Conference on Earthquake Engineering and Seismology (ECEES), Geneva, Switzerland.

- Der Kiureghian, A. (1989). "Measures of structural safety under imperfect states of knowledge." *Journal of Structural Engineering*, 115(5), 1119-1140.
- Der Kiureghian, A. (1996). "Structural reliability methods for seismic safety assessment: A review." *Engineering Structures*, 18(6), 412 - 424.
- Der Kiureghian, A., and Ditlevsen, O. (2009). "Aleatory or epistemic? Does it matter." *Structural Safety*, 31, 105-112.
- Dhakal, R. P., Mander, J. B., and Mashiko, N. (2006). "Identification of critical ground motions for seismic performance assessment of structures." *Earthquake Engineering & Structural Dynamics*, 35(8), 989-1008.
- Dhakal, R. P., Singh, S., and Mander, J. B. (2007). "Effectiveness of earthquake selection and scaling method in New Zealand." *Bulletin of the New Zealand Society for Earthquake Engineering*, 40(3), 160-171.
- Ditlevsen, O. (1982). "Model uncertainty in structural reliability." *Structural Safety*, 1(1), 73-86.
- Ditlevsen, O., and Madsen, H. O. (1996). *Structural reliability methods*, J. Wiley & Sons, New York.
- Douglas, J. (2002). "Note on scaling of peak ground acceleration and peak ground velocity with magnitude." *Geophysical Journal International*, 148(2), 336-339.
- 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(1-2), 43-104.
- Douglas, J. (2006). "Strong-motion records selection for structural testing." 1st European Conference on Earthquake Engineering and Seismology (ECEES), Geneva, Switzerland.
- Douglas, J., and Smit, P. M. (2001). "How accurate can strong ground motion attenuations be?" *Bulletin of the Seismological Society of America*, 91(6), 1917-1923.
- Dowrick, D. J. (2003). *Earthquake Risk Reduction*, John Wiley and Sons, Chichester, West Sussex, England.
- EERI Committee on Seismic Risk. (1984). "Glossary of terms for probabilistic seismic-risk and hazard analysis." *Earthquake Spectra*, 1(1), 33-40.
- Erberik, M. A., and Elnashai, A. S. (2004). "Fragility analysis of flat-slab structures." *Engineering Structures*, 26(7), 937-948.
- Fajfar, P., Vidic, T., and Fischinger, M. (1990). "A measure of earthquake motion capacity to damage medium period structures." *Soil Dynamics and Earthquake Engineering*, 9(5), 236-242.
- Faravelli, L. (1989). "Response-surface approach for reliability analysis." *Journal of Engineering Mechanics*, 115(12), 2763 - 2781.
- FEMA, F. E. M. A. (1994). *NEHRP Recommended provisions for the development of seismic regulations for new buildings*, Washington, D.C.
- FEMA, F. E. M. A. (2003). "Multi-hazard loss estimation methodology, earthquake model, HAZUS-MH MR1 technical manual." Report, Washington, D.C.
- Franchin, P., Lupoi, A., Pinto, P. E., and Schotanus, M. I. J. (2003a). "Response surface for seismic fragility analysis of RC structures." *Applications of Statistics and Probability in Civil Engineering*, San Francisco.
- Franchin, P., Lupoi, A., Pinto, P. E., and Schotanus, M. I. J. (2003b). "Seismic Fragility of Reinforced Concrete Structures Using a Response Surface Approach." *Journal of Earthquake Engineering*, 7(Special Issue 1), 45-77.
- Franchin, P., Lupoi, A., Pinto, P. E., and Schotanus, M. I. J. (2004). "Seismic Fragility Analysis of RC Structures, Use of Response Surface for a Realistic Application." XIII World Conference on Earthquake Engineering, Vancouver.

- Frankel, A., Mueller, C., Barnhard, T., Perkins, D., Leyendecker, E. V., Dickman, N., Hanson, S., and Hopper, M. (1996). "National Seismic-Hazard Maps: Documentation June 1996." Report, U.S. Geological Survey.
- Gardoni, P., Der Kiureghian, A., and Mosalam, K. M. (2002). "Probabilistic Capacity Models and Fragility Estimates for Reinforced Concrete Columns based on Experimental Observations." *Journal of Engineering Mechanics*, 128(10), 1024-1038.
- Gardoni, P., Mosalam, K. M., and Der Kiureghian, A. (2003). "Probabilistic seismic demand models and fragility estimates of RC bridges." *Journal of Earthquake Engineering*, 7(Special issue 1), 79-106.
- Gasparini, D., and Vanmarcke, E. H. (1976). "SIMQKE, Simulated Earthquake Motions Compatible With Prescribed Response Spectra." Report, MIT.
- Geomatrix. (2007). "Modified Application of DGML to Match the Distribution about the Conditional Mean Spectrum for Nonlinear Structural Analysis." GSM workgroup: synopses of methods adopted.
- Giovenale, P., Cornell, C. A., and Esteva, L. (2004). "Comparing the adequacy of alternative ground motion intensity measures for the estimation of structural responses." *Earthquake Engineering & Structural Dynamics*, 33(8), 951-979.
- Goulet, C., Haselton, C. B., Mitrani-Reiser, J., Stewart, J. P., Taciroglu, E., and Deierlein, G. G. (2006). "Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - Part I: ground motion selection and structural collapse simulation." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, U.S.
- Gulkam, P., and Sozen, M. (1999). "Procedure for determining seismic vulnerability of building structures." *ACI Structural Journal*, 96(3), 336-342.
- Gutenberg, B., and Richter, C. F. (1944). "Frequency of earthquakes in California." *Bulletin of the Seismological Society of America*, 34, 185-188.
- Han, S. W., and Wen, Y. K. (1994). "Method of reliability-based calibration of seismic structural parameters." Report, University of Illinois and Urbana-Champaign, Champaign, IL.
- Hancock, J. (2006). "The influence of duration and the selection and scaling of accelerograms in engineering design and assessment," Imperial College, University of London, London, GB.
- Hancock, J., and Bommer, J. J. (2006). "A State-of-Knowledge Review of the Influence of Strong-Motion Duration on Structural Damage." *Earthquake Spectra*, 22(3), 827-845.
- Hancock, J., and Bommer, J. J. (2007). "Using spectral matched records to explore the influence of strong-motion duration on inelastic structural response." *Soil Dynamics and Earthquake Engineering*, 27(4), 291-299.
- Hancock, J., Bommer, J. J., and Stafford, P. (in press). "Numbers of Scaled and Matched Accelerograms Required for Inelastic Dynamic Analyses." *Earthquake Engineering and Structural Dynamics*.
- Hancock, J., Bommer, J. J., and Stafford, P. J. (2008). "Numbers of Scaled and Matched Accelerograms Required for Inelastic Dynamic Analyses." *Earthquake Engineering & Structural Dynamics*, 37(14), 1585-1702.
- Hancock, J., Watson-Lamprey, J. A., Abrahamson, N. A., Bommer, J. J., Markatis, A., McCoy, E., and Mendis, R. (2006). "An improved method of matching response spectra of recorded earthquake ground motion using wavelets." *Journal of Earthquake Engineering*, 10(SPEC ISS 1), 67 - 89.
- Hanks, T. C. (1979). "b values and  $\omega^{-\gamma}$  seismic source model: implication for tectonic stress variations along active crustal fault zones and the estimation of high-

- frequency strong ground motions." *Journal of Geophysical Research*, 84, 2235-2242
- Hanks, T. C., and Cornell, C. A. (2001). "Probabilistic seismic hazard analysis: a beginner's guide." *Earthquake Spectra*, Submitted.
- Hanks, T. C., and Kanamori, H. (1979). "A moment magnitude scale." *Journal of Geophysical Research*, 77(23), 4393-4405.
- Hanks, T. C., and McGuire, R. K. (1981). "The character of high-frequency strong ground motion." *Bulletin of the Seismological Society of America*, 71, 2071-1095.
- Hartzell, S. H. (1978). "Earthquakes aftershocks as Green's Functions." *Geophysical Research Letters*, 5, 1-4.
- Haselton, C. B., and Baker, J. W. (2006). "Ground motion intensity measures for collapse capacity prediction: choice of optimal spectral period and effect of spectral shape." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, U.S.
- Haselton, C. B., and Deierlein, G. G. (2006). "Assessing Seismic Collapse Safety of Modern Reinforced Concrete Moment Frame Buildings." Report, Stanford University, Stanford, CA.
- Hernández, J. J., and López, O. A. (2003). "Evaluation of combination rules for peak response calculation in three-component seismic analysis." *Earthquake Engineering & Structural Dynamics*, 32(10), 1585-1602.
- Hidalgo, P., and Clough, R. W. (1974). "Earthquake simulator study of a reinforced concrete frame." Report, EERC, University of California, Berkeley, CA.
- Hong, H. P., and Goda, K. (2007). "Orientation-Dependent Ground-Motion Measure for Seismic-Hazard Assessment." *Bulletin of the Seismological Society of America*, 97(5), 1525-1538.
- Housner, G. W. "Spectrum Intensities of Strong-Motion Earthquakes." *Symposium on Earthquake and Blast Effects on Structures*, Los Angeles, CA, 20-36.
- Housner, G. W. (1970). "Strong Ground Motion." *Earthquake Engineering*, R. L. Wiegell, ed., Prentice-Hall Inc., Englewood Cliffs, N. J.
- Housner, G. W. "Measures of Severity of Ground Shaking." *U.S. Conference on Earthquake Engineering*, Earthquake Engineering Research Institute, 25-33.
- Housner, G. W., and Jennings, P. C. (1964). "Generation of artificial earthquakes." *Journal of the Engineering Mechanics Division*, 90, Proceedings paper 3806.
- Housner, G. W., and Jennings, P. C. (1977). "The Capacity of Extreme Earthquake Motions to Damage Structures." *Structural and Geotechnical Mechanics*, W. J. Hall, ed., Prentice-Hall Inc., Englewood Cliffs, N. J., 102-116.
- Husid, L. R. (1969). "Características de terremotos - Analisis general." *Revista del IDIEM, Santiago del Chile*, 8, 21-42.
- Hwang, H. H. M., and Jaw, J.-W. (1990). "Probabilistic Damage Analysis of Structures." *Journal of Structural Engineering*, 116(7), 1992-2007.
- Ibarra, L. F., and Krawinkler, H. (2005). "Global collapse of frame structures under seismic excitations." Report, Department of Civil and Environmental Engineering, Stanford University, Stanford, CA.
- Ibarra, L. F., Medina, R. A., and Krawinkler, H. (2005). "Hysteretic models that incorporate strength and stiffness deterioration." *Earthquake Engineering & Structural Dynamics*, 34(12), 1489-1511.
- Iervolino, I., and Cornell, C. A. (2005). "Record selection for nonlinear seismic analysis of structures." *Earthquake Spectra*, 21(3), 685 - 713.

- Iervolino, I., Maddaloni, G., and Cosenza, E. (2006a). "Unscaled records sets compliant with eurocode 8." 1st European Conference on Earthquake Engineering and Seismology, Geneva, Switzerland.
- Iervolino, I., Manfredi, G., and Cosenza, E. (2006b). "Ground motion duration effects on nonlinear seismic response." *Earthquake Engineering & Structural Dynamics*, 35(1), 21-38.
- Inel, M., and Ozmen, H. B. (2006). "Effects of plastic hinge properties in nonlinear analysis of reinforced concrete buildings." *Engineering Structures*, 28, 1494-1502.
- Irikura, K. (1983). "Semi-empirical estimation of strong ground motions during large earthquakes." *Bulletin of the Disaster Prevention Research Institute (Kyoto University)*, 33, 63-104.
- Jalayer, F. (2003). "Direct probabilistic seismic analysis: implementing non-linear dynamic assessments," Stanford University, Stanford.
- Jayaram, N., and Baker, J. W. (2008). "Statistical tests for the joint distribution of spectral acceleration values." *Bulletin of the Seismological Society of America*, 98(5), 2231-2243.
- Joyner, W. B., and Boore, D. M. (1981). "Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley, California, earthquake." *Bulletin of the Seismological Society of America*, 71(6), 2011-2038.
- Joyner, W. B., and Boore, D. M. (1993). "Methods for regression analysis of strong-motion data." *Bulletin of the Seismological Society of America*, 83(2), 469-487.
- Joyner, W. B., and Boore, D. M. (1994). "ERRATA: Methods for regression analysis of strong-motion data." *Bulletin of the Seismological Society of America*, 84(3), 955-956.
- Kagan, Y. Y. (1996). "Comment on "The Gutenberg-Richter or characteristic earthquake distribution, which is it?" by Steven G. Wesnousky." *Bulletin of the Seismological Society of America*, 86(1A), 274-285.
- Kagawa, T., Irikura, K., and Somerville, P. G. (2004). "Differences in ground motion and fault rupture process between the surface and buried earthquakes." *Earth Planets and Space*, 56(1), 3-14.
- Kappos, A. J., and Kyriakakis, P. (2000). "A re-evaluation of scaling techniques for natural records." *Soil Dynamics and Earthquake Engineering*, 20(1-4), 111-123.
- Kaul, M. K. (1978). "Spectrum-consistent time-history generation." *Journal of the Engineering Mechanics Division*, 104(ME4), 781-788.
- Keefer, D. L., and Bodily, S. E. (1983). "Three-point approximations for continuous random variables." *Management Science*, 29(5), 595-609.
- Khuri, A. I. (1992). "Response surface models with random block effects." *Technometrics*, 34(1), 26-37.
- Khuri, A. I., and Cornell, J. A. (1996). *Response surfaces: designs and analyses*, Marcel Dekker, New York.
- Kinoshita, S. (1998). "Kyoshin Net (K-NET)." *Seismological Research Letters*, 69(4), 309-332.
- Kircher, C. A., and Haselton, C. B. (2007). "ATC-63 Method." GSM workgroup: synopses of methods adopted.
- Kircher, C. A., Nassar, A. A., Kustu, O., and Holmes, W. T. (1997a). "Development of building damage functions for earthquake loss estimation." *Earthquake Spectra*, 13(4), 663-682.
- Kircher, C. A., Reitherman, R. K., and Alavi, B. (1997b). "Estimation of earthquake losses to buildings." *Earthquake Spectra*, 13(4), 703-720.

- Koutsourelakis, P. S., Pradlwarter, H. J., and Schuëller, G. I. (2003). "Reliability of structures in high dimensions." *Proceedings in Applied Mathematics and Mechanics*, 3(1), 495-496.
- Koutsourelakis, P. S., Pradlwarter, H. J., and Schuëller, G. I. (2004). "Reliability of structures in high dimensions. Part I. Algorithms and applications." *Probabilistic Engineering Mechanics*, 19.
- Kowalsky, M. J. (2000). "Deformation Limit States for Circular Reinforced Concrete Bridge Columns." *Journal of Structural Engineering*, 126(8), 869-878.
- Kramer, S. L. (1996). *Geotechnical earthquake engineering*, Prentice-Hall.
- Krawinkler, H., Medina, R. A., and Alavi, B. (2003). "Seismic drift and ductility demands and their dependence on ground motions." *Engineering Structures*, 25(5), 637-653.
- Krinitzsky, E. L. (2002). "How to obtain earthquake ground motions for engineering design." *Engineering Geology*, 65(1), 1-16.
- Krinitzsky, E. L., and Chang, F. K. (1977). "Specifying peak motion for design earthquakes." Report, US Army Corps of Engineers, Vicksburg, Mississippi.
- Kristek, J., and Moczo, P. (2003). "Seismic-wave propagation in viscoelastic media with material discontinuities: a 3D fourth-order staggered-grid finite-difference modeling." *Bulletin of the Seismological Society of America*, 93, 2273-2280.
- Kurama, Y. C., and Farrow, K. T. (2003). "Ground motion scaling methods for different site conditions and structure characteristics." *Earthquake Engineering & Structural Dynamics*, 32(15), 2425-2450.
- Kwon, O. S., and Elnashai, A. S. (2006). "The effect of material and ground motion uncertainty on the seismic vulnerability of RC structures." *Engineering Structures*, 28(2), 289-303.
- Lai, S.-S. P. (1982). "Statistical characterization of strong ground motions using power spectral density function." *Bulletin of the Seismological Society of America*, 72(1), 259-274.
- Lindstrom, M. J., and Bates, D. M. (1990). "Nonlinear mixed effects model for repeated measures data." *Biometrics*, 46(3), 673-687.
- Lomnitz, C. (1989). "Comment on 'Temporal and magnitude dependence in earthquake recurrence models' by C. A. Cornell and S. R. Winterstein." *Bulletin of the Seismological Society of America*, 79(5), 1662-1662.
- Lopez, O. A., Chopra, A. K., and Hernandez, J. J. (2000). "Critical response of structures to multicomponent earthquake excitation." *Earthquake Engineering & Structural Dynamics*, 29(12), 1759-1778.
- Lopez, O. A., Hernandez, R. B., and Fernandez, A. (2006). "Response spectra for multicomponent structural analysis." *Earthquake Spectra*, 22(1), 85-113.
- Lopez, O. A., and Torres, R. (1997). "Critical angle of seismic incidence and the maximum structural response." *Earthquake Engineering & Structural Dynamics*, 26(9), 881-894.
- Luco, N. (2002). "Probabilistic seismic demand analysis, SMRF connection fractures, and near-source effects," Department of Civil and Environmental Engineering, Stanford University, CA, USA.
- Luco, N. (2006). "Categorization of Ground Motion Selection and Modification Methods (Presentation)." 2006 COSMOS Annual Meeting Technical Session.
- Luco, N., and Bazzurro, P. (2007). "Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses?" *Earthquake Engineering & Structural Dynamics*, 36(13), 1813-1835.

- Luco, N., and Cornell, C. A. (2007). "Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions." *Earthquake Spectra*, 23(2), 257-392.
- Lupoi, G., Franchin, P., Lupoi, A., and Pinto, P. E. (2006). "Seismic Fragility Analysis of Structural Systems." *Journal of Engineering Mechanics*, 132(4), 385-395.
- Lupoi, G., Lupoi, A., and Pinto, P. E. (2002). "Seismic risk assessment of RC structures with the "2000 SAC/FEMA" method." *Journal of Earthquake Engineering*, 6(4), 499-512.
- Madsen, H. O., Krenk, S., and Lind, N. C. (1986). *Methods of structural safety*, Prentice-Hall, Englewood Cliffs.
- Mai, P. M., Spudich, P., and Boatwright, J. (2005). "Hypocenter Locations in Finite-Source Rupture Models." *Bulletin of the Seismological Society of America*, 95(3), 965-980.
- Malhotra, P. K. (2003). "Strong-Motion Records for Site-Specific Analysis." *Earthquake Spectra*, 19(3), 557 - 578.
- Malhotra, P. K. (2007). "Simulating Site-Specific Ground Motions for Future Earthquakes." GSM workgroup: synopses of methods adopted.
- Mander, J. B., Priestley, M. J. N., and Park, R. (1989). "Theoretical stress-strain model for confined concrete." *Journal of Structural Engineering*, 114(8), 1804-1826.
- Mark, W. D. (1970). "Spectral analysis of the convolution and filtering of nonstationary stochastic processes." *Journal of Sound and Vibration*, 11, 19-63.
- Martínez-Rueda, J. E. (1998). "Scaling procedure for natural accelerograms based on a system of spectrum intensity scales." *Earthquake Spectra*, 14(1), 135-152.
- Marušić, D., and Fajfar, P. (2005). "On the inelastic seismic response of asymmetric buildings under bi-axial excitation." *Earthquake Engineering & Structural Dynamics*, 34(8), 943-963.
- Matsumura, K. (1992). "On the intensity measure of strong motions related to structural failures." 10th World Conference on Earthquake Engineering, Madrid, 375-380.
- Matthews, M. V., Ellsworth, W. L., and Reasenber, P. A. (2002). "A Brownian model for recurrent earthquakes." *Bulletin of the Seismological Society of America*, 46(6), 1271-1282.
- Mattock, A. H. (1964). "Rotational capacity of hinging regions in reinforced concrete beams." Flexural mechanics of reinforced concrete, American Concrete Institute (ACI), Farmington Hills, MI.
- Mattock, A. H. (1967). "Discussion of "Rotational capacity of hinging regions in reinforced concrete beams"." *Journal of the Structural Division*, 93(ST2), 519-522.
- McGuire, R. K. (1976). "FORTRAN computer program for seismic risk analysis." Report, US Geological Survey Open-File Report 76-67.
- McGuire, R. K. (1995). "Probabilistic Seismic Hazard Analysis and Design Earthquakes: Closing the Loop." *Bulletin of the Seismological Society of America*, 85(5), 1275-1284.
- McKenna, F., Fenves, G. L., and Scott, M. H. (2000). "Open system for earthquake engineering simulation (OpenSees)." Univ. of California, Berkeley, CA.
- McKenna, F. T. (1997). "Object-Oriented Finite Element Programming: Frameworks for Analysis, Algorithms and Parallel Computing," University of California, Berkeley.
- Medina, R., and Krawinkler, H. (2004). "Seismic Demands for Nondeteriorating Frame Structures and Their Dependence on Ground Motions." Report: *PEER Report 2003/15*, Pacific Earthquake Engineering Research Center.

- Mehanny, S. S. S., and Deierlein, G. G. (2000). "Modelling of assessment of seismic performance of composite frames with reinforced concrete columns and steel beams." Report, Department of Civil and Environmental Engineering, Stanford University, Sanford.
- Mendis, P. (2001). "Plastic hinge lengths of normal and high-strength concrete in flexure." *Advances in structural engineering*, 4(4), 189-195.
- Menegotto, M., and Pinto, P. E. (1973). "Method for analysis for cyclically loaded R.C. plane frames including changes in geometry and non-elastic behavior of elements under combined normal force and bending." Symposium on the resistance and ultimate deformability of structures acted on by well defined repeated loads, Zurich, Switzerland, 15-22.
- Miller, A. C., and Rice, T. R. (1983). "Discrete approximations for probability distributions." *Management Science*, 29(3), 352-362.
- Mosalam, K. M., Ayala, G., White, R. W., and Roth, C. (1997). "Seismic fragility of LRC frames with and without masonry infill walls." *Journal of Earthquake Engineering*, 1(4), 693-720.
- Naeim, F., Alimoradi, A., and Pezeshk, S. (2004). "Selection and scaling of ground motion time histories for structural design using genetic algorithms." *Earthquake Spectra*, 20(2), 413 - 426.
- Naeim, F., and Lew, M. (1995). "On the Use of Design Spectrum Compatible Time Histories." *Earthquake Spectra*, 11(1), 111-127.
- Nau, J. M., and Hall, W. J. (1984). "Scaling methods for earthquake response spectra." *Journal of Structural Engineering*, 110(7), 1533-1548.
- Nau, R. F., Oliver, R. M., and Pister, K. S. (1982). "Simulating and analyzing artificial nonstationary earthquake ground motions." *Bulletin of the Seismological Society of America*, 72(615-636).
- Newmark, N. M. (1965). "Effects of earthquakes on dams and embankments." *Geotechnique*, 15(2), 139-160.
- Nishenko, S. P., and Bulad, R. (1987). "A generic recurrence interval distribution for earthquake forecasting." *Bulletin of the Seismological Society of America*, 77(4), 1382-1399.
- Panagiotakos, T. B., and Fardis, M. N. (2001). "Deformations of Reinforced Concrete Members at Yielding and Ultimate." *ACI Structural Journal*, 98(2), 135.
- Park, Y.-J., Ang, A. H.-S., and Wen, Y. K. (1985). "Seismic damage analysis of reinforced concrete buildings." *Journal of Structural Engineering*, 111(4), 740-757.
- Paulay, T., and Priestley, M. J. N. (1992). *Seismic design of reinforced concrete and masonry buildings*, John Wiley and Sons, New York.
- PEER. (2005). "Next Generation of Attenuation (NGA) database. PEER strong-motion database (<http://peer.berkeley.edu/nga/index.html>)." Pacific Earthquake Engineering Research Centre, Berkeley.
- Penzien, J., and Watabe, M. (1975). "Characteristics of 3-dimensional earthquake ground motions." *Earthquake Engineering & Structural Dynamics*, 3, 365-373.
- Peruš, I., and Fajfar, P. (2005). "On the inelastic torsional response of single-storey structures under bi-axial excitation." *Earthquake Engineering & Structural Dynamics*, 34(8), 931-941.
- Pinheiro, J. C., and Bates, D. M. (1995). "Model building for nonlinear mixed-effects models." Report, Department of Biostatistics, University of Wisconsin.
- Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-effects Models in S and S-PLUS*, Springer, New York.

- Pinto, P. E. (2001). "Reliability methods in earthquake engineering." *Progress in Structural Engineering and Materials*, 3(1), 76-85.
- Pinto, P. E., Giannini, R., and Franchin, P. (2004). *Seismic reliability analysis of structures*, IUSS Press, Pavia.
- Porter, K. A. (2003). "An Overview of PEER's Performance-Based Earthquake Engineering Methodology." 9th International Conference on Applications of Statistics and Probability in Civil Engineering, San Francisco, CA, U.S.
- Porter, K. A., Kiremidjian, A. S., and S., L. J. (2001). "Assembly-based vulnerability of buildings and its use in performance evaluation." *Earthquake Spectra*, 17(2), 291-312.
- Pousse, G. (2005). "Analyse des données accélérométriques de K-Net et Kik-Net: implications pour la prédiction du mouvement sismique - accélérogrammes et spectres de réponse - et la prise en compte des effets de site non-linéaire," Université de Grenoble, Grenoble, FR.
- Pousse, G., Bonilla, L. F., Cotton, F., and Margerin, L. (2006). "Nonstationary stochastic simulation of strong ground motion time histories including natural variability: Application to the K-net Japanese database." *Bulletin of the Seismological Society of America*, 96(6), 2103-2117.
- Power, M., Chiou, B., Abrahamson, N. A., Bozorgnia, Y., Shantz, T. J., and Roblee, C. (2008). "An Overview of the NGA Project." *Earthquake Spectra*, 24(1), 3-21.
- Power, M., Chiou, B., Abrahamson, N. A., and Roblee, C. (2006). "The 'Next Generation of Ground Motion Attenuation Models' (NGA) project: an overview." Proceedings of the 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, Paper No. 2022.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and P., F. B. (2002). *Numerical recipes in C++: the art of scientific computing*, Cambridge University Press, Cambridge.
- Preumont, A. (1984). "The generation of spectrum compatible accelerograms for the design of nuclear power plants." *Earthquake Engineering & Structural Dynamics*, 12(4), 481-497.
- Priestley, M. J. N. (2003). *Myths and fallacies in earthquake engineering - revisited: The Mallet Milne Lecture*, IUSS Press, Pavia, Italy.
- Priestley, M. J. N., Calvi, G. M., and Kowalsky, M. J. (2007). *Displacement-based seismic design of structures*, IUSS Press, Pavia, Italy.
- Priestley, M. J. N., and Park, R. (1987). "Strength and ductility of concrete bridge columns under seismic loading." *ACI Structural Journal*, 84(1), 61-76.
- R Development Core Team. (2008). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Ramamoorthy, S. K., Gardoni, P., and Bracci, J. M. (2006). "Probabilistic Demand Models and Fragility Curves for Reinforced Concrete Frames." *Journal of Structural Engineering*, 132(10), 1563-1572.
- Ramamoorthy, S. K., Gardoni, P., and Bracci, J. M. (2008). "Seismic Fragility and Confidence Bounds for Gravity Load Designed Reinforced Concrete Frames of Varying Height." *Journal of Structural Engineering*, 134(4), 639-650.
- Rathje, E. M., and Kottke, A. R. (2007). "Procedures for selection and scaling of earthquake motions for dynamic response analysis." 1st U.S.-Italy Seismic Bridge Workshop, Pavia, Italy.
- Reiter, L. (1990). *Earthquake hazard analysis: issues and insights*, Columbia University Press.
- Restrepo-Velez, L., and Bommer, J. J. (2003). "An exploration of the nature of the scatter in ground motion prediction equations and the implications for seismic

- hazard assessment." *Journal of Earthquake Engineering*, 7(Special Issue 1), 171-199.
- Riddell, R. (2007). "On ground motion intensity indices." *Earthquake Spectra*, 23(1), 147-173.
- Riddell, R., and Garcia, J. E. (2001). "Hysteretic energy spectrum and damage control." *Earthquake Engineering & Structural Dynamics*, 30(12), 1791-1816.
- Rigato, A. B., and Medina, R. A. (2007). "Influence of angle of incidence on seismic demands for inelastic single-storey structures subjected to bi-directional ground motions." *Engineering Structures*, 29(10), 2593-2601.
- Rossetto, T., and Elnashai, A. S. (2003). "Derivation of vulnerability functions for European-type RC structures based on observational data." *Engineering Structures*, 25(10), 1241-1263.
- Rossetto, T., and Elnashai, A. S. (2005). "A new analytical procedure for the derivation of displacement-based vulnerability curves for populations of RC structures." *Engineering Structures*, 27(3), 397-409.
- Sabetta, F., Masiani, R., and Giuffrè, A. (1986). "Frequency nonstationarity in Italian strong motion accelerograms." 8th European Conference on Earthquake Engineering, Lisbon, Portugal, 25-32.
- Sabetta, F., and Pugliese, A. (1996). "Estimation of response spectra and simulation of nonstationary earthquake ground motions." *Bulletin of the Seismological Society of America*, 86(2), 337-352.
- Saragoni, G. R., and Hart, G. C. (1974). "Simulation of artificial earthquakes." *Earthquake Engineering Soil Dynamics*, 2, 249-267.
- Sarma, S. K., and Yang, K. S. (1987). "An evaluation of strong motion records and a new parameter  $A_{95}$ ." *Earthquake Engineering & Structural Dynamics*, 15, 119-132.
- Scherbaum, F., Bommer, J. J., Bungum, H., Cotton, F., and Abrahamson, N. A. (2005). "Composite ground-motion models and logic trees: methodology, sensitivities, and uncertainties." *Bulletin of the Seismological Society of America*, 95(5), 1575-1593.
- Schneider, J. F., Silva, W. J., and Stark, C. (1993). "Ground motion model for the 1989 M 6.9 Loma Prieta earthquake including effect of source, path and site." *Earthquake Spectra*, 9, 251-287.
- Schotanus, M. I. J., Franchin, P., Lupoi, A., and Pinto, P. E. (2004). "Seismic fragility analysis of 3D structures." *Structural Safety*, 26(4), 421 - 441.
- Schuëller, G. I., and Pradlwarter, H. J. (2007). "Benchmark study on reliability estimation in higher dimensions of structural systems - An overview." *Structural Safety*, 29(3), 167-182.
- Schuëller, G. I., Pradlwarter, H. J., and Koutsourelakis, P. S. (2004). "A critical appraisal of reliability estimation procedures for high dimensions." *Probabilistic Engineering Mechanics*, 19(4), 463-474.
- Schwartz, D. P., and Coppersmith, K. (1984a). "Fault behavior and characteristic earthquakes - examples from the Wasatch and San-Andreas fault zones." *Journal of Geophysical Research*, 89(NB7), 5681-5698.
- Schwartz, D. P., and Coppersmith, K. J. (1984b). "Fault behavior and characteristic earthquakes - examples from the Wasatch and San-Andreas fault zones." *Journal of Geophysical Research*, 89(NB7), 5681-5698.
- Scott, M. H., and Fenves, G. L. (2006). "Plastic hinge integration methods for force-based beam-column elements." *Journal of Structural Engineering*, 132(2).

- Scott, M. H., Fenves, G. L., McKenna, F., and Filippou, F. C. (2008). "Software patterns for nonlinear beam-column models." *Journal of Structural Engineering*, 134(4).
- Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance components*, John Wiley & Sons, New York.
- Shantz, T. J. (2006). "Selection and scaling of earthquake records for nonlinear dynamic analysis of first mode dominate bridge structures." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, U.S.
- Sheikh, S. A., and Houry, S. S. (1993). "Confined concrete columns with stubs." *ACI Structural Journal*, 90(4), 414-431.
- Shinozuka, M., Feng, M. Q., Lee, J., and Naganuma, T. (2000). "Statistical Analysis of Fragility Curves." *Journal of Engineering Mechanics*, 126(12), 1224-1231.
- Shoja-Taheri, J., and Bolt, B. A. (1977). "A generalized strong-motion accelerogram based on spectral maximization from two horizontal components." *Bulletin of the Seismological Society of America*, 67(3), 863-876.
- Shome, N. (1999). "Probabilistic Seismic Demand Analysis of Nonlinear Structures," Department of Civil and Environmental Engineering, Stanford University, CA, USA.
- Shome, N. (2007). "Sa(T1) scaling for a given spectrum." GSMW workgroup: synopses of methods adopted.
- Shome, N., and Cornell, C. A. (2000). "Structural seismic demand analysis: Consideration of collapse." ASCE Speciality Conference on Probabilistic Mechanics and Structural Reliability, Notre Dame, Indiana, U.S.
- Shome, N., Cornell, C. A., Bazzurro, P., and Carballo, J. E. (1998). "Earthquakes, records, and nonlinear responses." *Earthquake Spectra*, 14(3), 469 - 500.
- Silva, W. J. (1987). "WES Rascal code for synthesizing Earthquake Ground Motions." Report, US Army Corps of Engineers State-of-the-Art for Assessing Earthquake Hazards in the United States.
- Singhal, A., and Kiremidjian, A. S. (1996). "Method for Probabilistic Evaluation of Seismic Structural Damage." *Journal of Structural Engineering*, 122(12), 1459-1467.
- Singhal, A., and Kiremidjian, A. S. (1998). "Bayesian Updating of Fragilities with Application to RC Frames." *Journal of Structural Engineering*, 124(8), 922-929.
- Somerville, P., Irikura, K., Graves, R. W., Sawada, S., Wald, D. J., Abrahamson, N. A., Iwasaki, Y., Kagawa, N., Smith, N., and Kowada, A. (1999). "Characterizing earthquake slip models for the prediction of strong ground motion." *Seismological Research Letters*, 70(1), 59-80.
- Somerville, P., and Pitarka, A. (2006). "Differences in earthquake source and ground motion characteristics between surface and buried faulting earthquakes." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA.
- Somerville, P., Smith, N., Graves, R. W., and Abrahamson, N. A. (1997a). "Modification of empirical strong ground motion attenuation relations to include the amplitude and duration effects of rupture directivity." *Seismological Research Letters*, 68(1), 199-222.
- Somerville, P. G., Smith, N., Punyamurthula, S., and Sun, J. (1997b). "Development of ground motion time histories for phase 2 of the FEMA/SAC steel project." Report: SAC/BD-97/04, SAC Joint Venture, Sacramento, CA, USA.
- Spacone, E., Ciampi, V., and Filippou, F. C. (1996a). "Mixed formulation of nonlinear beam finite element." *Computers and Structures*, 58(1), 71-83.

- Spacone, E., Filippou, F. C., and Taucer, F. F. (1996b). "Fibre beam-column model for non-linear analysis of R/C frames: part I. Formulation." *Earthquake Engineering & Structural Dynamics*, 25(7), 711-725.
- Spacone, E., Filippou, F. C., and Taucer, F. F. (1996c). "Fibre beam-column model for non-linear analysis of R/C frames: part II. Applications." *Earthquake Engineering & Structural Dynamics*, 25(7), 727-742.
- Spiegel, M. R., and Liu, J. (1999). *Mathematical handbook of formulas and tables*, McGraw-Hill, New York.
- Stafford, P. J. (2006). "Engineering Seismological Studies and Seismic Design Criteria for the Buller Region, South Island, New Zealand," University of Canterbury, Christchurch, New Zealand.
- Stafford, P. J., Strasser, F., and Bommer, J. J. (2008). "An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region." *Bulletin of Earthquake Engineering*, 6(2), 149-177.
- Stephens, M. A. (1974). "EDF Statistics for Goodness of Fit and Some Comparisons." *Journal of the American Statistical Association*, 69(347), 730-737.
- Stewart, J. S., Chiou, S. J., Bray, J. D., Graves, R. W., Sommerville, P. G., and Abrahamson, N. A. (2001). "Ground Motion Evaluation Procedures for Performance-Based Design." Report: 2001/09, PEER.
- Stoica, M., Medina, R. A., and McCuen, R. H. (2007). "Improved probabilistic quantification of drift demands for seismic evaluation." *Structural Safety*, 29(2), 132-145.
- Strasser, F. O., Abrahamson, N. A., and Bommer, J. J. (2009). "Sigma: Issues, Insights, and Challenges." *Seismological Research Letters*, 80, 40-56.
- Thrainsson, H., and Kiremidjian, A. S. (2002). "Simulation of digital earthquake accelerograms using the inverse discrete Fourier transform." *Earthquake Engineering & Structural Dynamics*, 31, 2023-2048.
- Toshinawa, T., Hisada, Y., Kono, K., Shibayama, A., Honkawa, Y., and Ono, H. (2004). "Topographic site response at a quaternary terrace in Hachioji, Japan, observed in strong motion and microtremors." 13th World Conference on Earthquake Engineering, Vancouver, B.C., Canada.
- Tothong, P., and Cornell, C. A. (2006). "An Empirical Ground-Motion Attenuation Relation for Inelastic Spectral Displacement." *Bulletin of the Seismological Society of America*, 96(6), 2146-2164.
- Tothong, P., and Cornell, C. A. (2007). "Probabilistic Seismic Demand Analysis Using Advanced Ground Motion Intensity Measures, Attenuation Relationships, and Near-Fault Effects." Report: 2006/11, PEER.
- Tothong, P., and Luco, N. (2007). "Probabilistic seismic demand analysis using advanced ground motion intensity measures." *Earthquake Engineering & Structural Dynamics*, 36(13), 1837-1860.
- Trifunac, M. D., and Brady, A. G. (1975). "A study on the duration of strong duration earthquake ground motion." *Bulletin of the Seismological Society of America*, 65(3), 585-626.
- Trombetti, T., Gasparini, G., Silvestri, S., and Malavolta, D. (2006). "A methodology for determination of efficient earthquake bins for performance based seismic design." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, USA.
- Vamvatsikos, D. (2002). "Seismic performance, capacity and reliability of structures as seen through incremental dynamic analysis," Stanford University, Stanford.
- Vamvatsikos, D., and Cornell, C. A. (2002). "Incremental dynamic analysis." *Earthquake Engineering & Structural Dynamics*, 31(3), 491 - 514.

- Vamvatsikos, D., and Cornell, C. A. (2004). "Applied incremental dynamic analysis." *Earthquake Spectra*, 20(2), 523 - 553.
- Vanmarcke, E. H. (1979). "Representation of earthquake ground motion: scaled accelerograms and equivalent response spectra." Report, US Army Corps of Engineers, Vicksburg, Mississippi.
- Vanmarcke, E. H. (1980). "Parameters of the spectral density function, their significance in the time and frequency domain." *MIT Civil Engineering and Design*, 60(1).
- Vanmarcke, E. H., and Lai, S.-S. P. (1980). "Strong-motion duration and RMS amplitude of earthquake records." *Bulletin of the Seismological Society of America*, 70(4), 1293-1307.
- Veneziano, D., Casciati, F., and Faravelli, L. (1983). "Method of seismic fragility for complicated systems." 2nd CNSI Specialist meeting on probabilistic methods in seismic risk assessment for NPP, Lawrence Livermore Laboratory, CA.
- Watson-Lamprey, J. A. (2006). "Evaluation of GSM Methods Results." Presentation at the GSM meeting on 17th November 2006.
- Watson-Lamprey, J. A., and Abrahamson, N. A. (2006a). "Bias caused by use of spectrum compatible motions." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, U.S.
- Watson-Lamprey, J. A., and Abrahamson, N. A. (2006b). "Selection of ground motion time series and limits on scaling." *Soil Dynamics and Earthquake Engineering*, 26(5), 477 - 482.
- Watson-Lamprey, J. A., and Abrahamson, N. A. (2006c). "Selection of time series for analyses of response of buildings." 1st European Conference on Earthquake Engineering and Seismology (ECEES), Geneva, Switzerland.
- Watson-Lamprey, J. A., and Boore, D. M. (2007). "Beyond Sa\_GMRotI: Conversion to Sa\_Arb, Sa\_SN, and Sa\_MaxRot." *Bulletin of the Seismological Society of America*, 97(5), 1511-1524.
- Wen, Y. K., Ellingwood, B. R., and Bracci, J. M. (2004). "Vulnerability Function Framework for Consequence-based Engineering." Report, Middle America Earthquake Center.
- Wesnousky, S. G. (1994). "The Gutenberg-Richter or characteristic earthquake distribution, which is it?" *Bulletin of the Seismological Society of America*, 84(6), 1940-1959.
- Wesnousky, S. G. (1996). "Reply to Yan Kagan's comment on 'The Gutenberg-Richter or characteristic earthquake distribution, which is it?'" *Bulletin of the Seismological Society of America*, 86(1A), 286-291.
- Whitman, R. V., Anagnor, T., Kircher, C. A., Lagorio, H. J., Lawson, R. S., and Schneider, P. (1997). "Development of a national earthquake loss estimation methodology." *Earthquake Spectra*, 13(4), 643-661.
- Woo, G. (1996). "Kernel estimation methods for seismic hazard area source modelling." *Bulletin of the Seismological Society of America*, 86, 353-362.
- Youngs, R. R., Abrahamson, N. A., Makdisi, F. I., and Sadigh, K. (1995). "Magnitude-dependent variance of peak ground acceleration." *Bulletin of the Seismological Society of America*, 85(4), 1161-1176.
- Youngs, R. R., Power, M. S., and Chin, C.-C. (2006). "Design ground motion library." 8th U.S. National Conference on Earthquake Engineering, San Francisco, CA, U.S.
- Yucemen, M. S., Ozcebe, G., and Pay, A. C. (2004). "Prediction of potential damage due to severe earthquakes." *Structural Safety*, 26(3), 349-366.

## References

---

- Zareian, F. (2007). "ATC-58 Scenario-Based Assessments Method." GSM workgroup: synopses of methods adopted.
- Zeng, Y., Anderson, J. G., and Yu, G. (1994). "A composite source model for computing realistic strong motions." *Geophysical Research Letters*, 21, 725-728.
- Zhai, C. H., and Xie, L. L. (2007). "A new approach of selecting real input ground motions for seismic design: The most unfavourable real seismic design ground motions." *Earthquake Engineering & Structural Dynamics*, 36(8), 1009-1027.

# Appendix A

In the present work a subset of the NGA database (Power et al., 2006; Power et al., 2008) was used to develop the ground-motion prediction equations given in Chapter 3 and Chapter 4. Table A.1 gives details of this subset. Criteria used to identify to select ground motion recordings have been given in Section 3.3.1 and Section 4.4. The databases used in Chapter 3 and Chapter 4 are almost completely overlapping, the only difference being the exclusion of pulse like ground motions from the subset used in Chapter 4. Record Sequence Number (RSN) is a unique ID number used in the NGA database ((Power et al., 2006; Power et al., 2008). Pulse classification has been performed according to the criterion proposed by Baker (2008). Each ground-motion recording in Table A.1 contains 2 horizontal components and 1 vertical component. Horizontal components only have been used in the present work.

**Table A.1. Ground motion database considered in the present work.**

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$ [km]	$V_{S30}$ m/s	Pulse 0 = No 1 = Yes
6	'Imperial Valley-02'	1940	'El Centro Array #9'	6.95	6.09	213.4	0
12	'Kern County'	1952	'LA - Hollywood Stor FF'	7.36	114.62	316.5	0
15	'Kern County'	1952	'Taft Lincoln School'	7.36	38.42	385.4	0
28	'Parkfield'	1966	'Cholame - Shandon Array #12'	6.19	17.64	408.9	0
30	'Parkfield'	1966	'Cholame - Shandon Array #5'	6.19	9.58	289.6	0
31	'Parkfield'	1966	'Cholame - Shandon Array #8'	6.19	12.9	256.8	0
32	'Parkfield'	1966	'San Luis Obispo'	6.19	63.34	712.8	0
33	'Parkfield'	1966	'Temblor pre-1969'	6.19	15.96	527.9	0
36	'Borrego Mtn'	1968	'El Centro Array #9'	6.63	45.12	213.4	0
37	'Borrego Mtn'	1968	'LA - Hollywood Stor FF'	6.63	222.42	316.5	0
38	'Borrego Mtn'	1968	'LB - Terminal Island'	6.63	199.84	229.8	0
40	'Borrego Mtn'	1968	'San Onofre - So Cal Edison'	6.63	129.11	442.9	0
51	'San Fernando'	1971	'2516 Via Tejon PV'	6.61	55.2	280.6	0
53	'San Fernando'	1971	'Bakersfield - Harvey Aud'	6.61	111.88	271.4	0
54	'San Fernando'	1971	'Borrego Springs Fire Sta'	6.61	214.32	271.4	0
55	'San Fernando'	1971	'Buena Vista - Taft'	6.61	111.37	438.3	0
58	'San Fernando'	1971	'Cedar Springs Pumphouse'	6.61	92.25	477.2	0
59	'San Fernando'	1971	'Cedar Springs, Allen Ranch'	6.61	89.37	813.5	0
60	'San Fernando'	1971	'Cholame - Shandon Array #2'	6.61	217.54	184.8	0
61	'San Fernando'	1971	'Cholame - Shandon Array #8'	6.61	218.17	256.8	0
62	'San Fernando'	1971	'Colton - So Cal Edison'	6.61	96.81	302	0
64	'San Fernando'	1971	'Fort Tejon'	6.61	59.52	394.2	0
65	'San Fernando'	1971	'Gormon - Oso Pump Plant'	6.61	43.95	308.4	0
67	'San Fernando'	1971	'Isabella Dam (Aux Abut)'	6.61	130	684.9	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
68	'San Fernando'	1971	'LA - Hollywood Stor FF'	6.61	22.77	316.5	0
69	'San Fernando'	1971	'LB - Terminal Island'	6.61	58.99	229.8	0
74	'San Fernando'	1971	'Maricopa Array #1'	6.61	193.25	271.4	0
75	'San Fernando'	1971	'Maricopa Array #2'	6.61	108.56	438.3	0
76	'San Fernando'	1971	'Maricopa Array #3'	6.61	109.01	438.3	0
77	'San Fernando'	1971	'Pacoima Dam (upper left abut)'	6.61	0	2016.1	0
81	'San Fernando'	1971	'Pearblossom Pump'	6.61	35.54	529.1	0
84	'San Fernando'	1971	'San Diego Gas & Electric'	6.61	205.77	370.8	0
86	'San Fernando'	1971	'San Onofre - So Cal Edison'	6.61	124.79	442.9	0
87	'San Fernando'	1971	'Santa Anita Dam'	6.61	30.7	684.9	0
88	'San Fernando'	1971	'Santa Felita Dam (Outlet)'	6.61	24.69	389	0
89	'San Fernando'	1971	'Tehachapi Pump'	6.61	61.75	669.5	0
90	'San Fernando'	1971	'UCSB - Fluid Mech Lab'	6.61	124.38	370.8	0
92	'San Fernando'	1971	'Wheeler Ridge - Ground'	6.61	68.38	347.7	0
93	'San Fernando'	1971	'Whittier Narrows Dam'	6.61	39.45	298.7	0
94	'San Fernando'	1971	'Wrightwood - 6074 Park Dr'	6.61	61.64	486	0
95	'Managua, Nicaragua-01'	1972	'Managua, ESSO'	6.24	3.51	288.8	0
121	'Friuli, Italy-01'	1976	'Barcis'	6.5	49.13	424.8	0
122	'Friuli, Italy-01'	1976	'Codroipo'	6.5	33.32	274.5	0
124	'Friuli, Italy-01'	1976	'Feltre'	6.5	102.05	659.6	0
125	'Friuli, Italy-01'	1976	'Tolmezzo'	6.5	14.97	424.8	0
126	'Gazli, USSR'	1976	'Karakyr'	6.8	3.92	659.6	0
131	'Friuli, Italy-02'	1976	'Codroipo'	5.91	41.37	274.5	0
133	'Friuli, Italy-02'	1976	'San Rocco'	5.91	14.37	659.6	0
135	'Santa Barbara'	1978	'Cachuma Dam Toe'	5.92	23.75	438.3	0
136	'Santa Barbara'	1978	'Santa Barbara Courthouse'	5.92	0	515	0
137	'Tabas, Iran'	1978	'Bajestan'	7.35	119.77	338.6	0
138	'Tabas, Iran'	1978	'Boshrooyeh'	7.35	24.07	338.6	0
139	'Tabas, Iran'	1978	'Dayhook'	7.35	0	659.6	0
140	'Tabas, Iran'	1978	'Ferdows'	7.35	89.76	274.5	0
141	'Tabas, Iran'	1978	'Kashmar'	7.35	193.91	274.5	0
142	'Tabas, Iran'	1978	'Sedeh'	7.35	150.33	424.8	0
143	'Tabas, Iran'	1978	'Tabas'	7.35	1.79	766.8	0
145	'Coyote Lake'	1979	'Coyote Lake Dam (SW Abut)'	5.74	5.3	597.1	0
146	'Coyote Lake'	1979	'Gilroy Array #1'	5.74	10.21	1428	0
147	'Coyote Lake'	1979	'Gilroy Array #2'	5.74	8.47	270.8	0
148	'Coyote Lake'	1979	'Gilroy Array #3'	5.74	6.75	349.9	0
149	'Coyote Lake'	1979	'Gilroy Array #4'	5.74	4.79	221.8	0
150	'Coyote Lake'	1979	'Gilroy Array #6'	5.74	0.42	663.3	0
151	'Coyote Lake'	1979	'Halls Valley'	5.74	33.69	281.6	0
153	'Coyote Lake'	1979	'SJB Overpass, Bent 5 g.l.'	5.74	20.44	370.8	0
154	'Coyote Lake'	1979	'San Juan Bautista, 24 Polk St'	5.74	19.46	370.8	0
158	'Imperial Valley-06'	1979	'Aeropuerto Mexicali'	6.53	0	274.5	0
159	'Imperial Valley-06'	1979	'Agrarias'	6.53	0	274.5	0
160	'Imperial Valley-06'	1979	'Bonds Corner'	6.53	0.47	223	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$ [km]	$V_{S30}$ m/s	Pulse 0 = No 1 = Yes
161	'Imperial Valley-06'	1979	'Brawley Airport'	6.53	8.54	208.7	0
162	'Imperial Valley-06'	1979	'Calexico Fire Station'	6.53	10.45	231.2	0
163	'Imperial Valley-06'	1979	'Calipatria Fire Station'	6.53	23.17	205.8	0
164	'Imperial Valley-06'	1979	'Cerro Prieto'	6.53	15.19	659.6	0
165	'Imperial Valley-06'	1979	'Chihuahua'	6.53	7.29	274.5	0
166	'Imperial Valley-06'	1979	'Coachella Canal #4'	6.53	49.1	345.4	0
167	'Imperial Valley-06'	1979	'Compuertas'	6.53	13.52	274.5	0
169	'Imperial Valley-06'	1979	'Delta'	6.53	22.03	274.5	0
170	'Imperial Valley-06'	1979	'EC County Center FF'	6.53	7.31	192.1	0
171	'Imperial Valley-06'	1979	'EC Meloland Overpass FF'	6.53	0.07	186.2	0
172	'Imperial Valley-06'	1979	'El Centro Array #1'	6.53	19.76	237.3	0
173	'Imperial Valley-06'	1979	'El Centro Array #10'	6.53	6.17	202.9	1
174	'Imperial Valley-06'	1979	'El Centro Array #11'	6.53	12.45	196.3	0
175	'Imperial Valley-06'	1979	'El Centro Array #12'	6.53	17.94	196.9	0
176	'Imperial Valley-06'	1979	'El Centro Array #13'	6.53	21.98	249.9	0
178	'Imperial Valley-06'	1979	'El Centro Array #3'	6.53	10.79	162.9	1
179	'Imperial Valley-06'	1979	'El Centro Array #4'	6.53	4.9	208.9	0
180	'Imperial Valley-06'	1979	'El Centro Array #5'	6.53	1.76	205.6	0
181	'Imperial Valley-06'	1979	'El Centro Array #6'	6.53	0	203.2	1
182	'Imperial Valley-06'	1979	'El Centro Array #7'	6.53	0.56	210.5	1
183	'Imperial Valley-06'	1979	'El Centro Array #8'	6.53	3.86	206.1	0
184	'Imperial Valley-06'	1979	'El Centro Differential Array'	6.53	5.09	202.3	1
185	'Imperial Valley-06'	1979	'Holtville Post Office'	6.53	5.51	202.9	1
186	'Imperial Valley-06'	1979	'Niland Fire Station'	6.53	35.64	207.5	0
187	'Imperial Valley-06'	1979	'Parachute Test Site'	6.53	12.69	348.7	0
188	'Imperial Valley-06'	1979	'Plaster City'	6.53	30.33	345.4	0
189	'Imperial Valley-06'	1979	'SAHOP Casa Flores'	6.53	9.64	338.6	0
190	'Imperial Valley-06'	1979	'Superstition Mtn Camera'	6.53	24.61	362.4	0
191	'Imperial Valley-06'	1979	'Victoria'	6.53	31.92	274.5	0
192	'Imperial Valley-06'	1979	'Westmorland Fire Sta'	6.53	14.75	193.7	0
230	'Mammoth Lakes-01'	1980	'Convict Creek'	6.06	1.1	338.5	0
231	'Mammoth Lakes-01'	1980	'Long Valley Dam (Upr L Abut)'	6.06	12.56	345.4	0
265	'Victoria, Mexico'	1980	'Cerro Prieto'	6.33	13.8	659.6	0
266	'Victoria, Mexico'	1980	'Chihuahua'	6.33	18.53	274.5	0
268	'Victoria, Mexico'	1980	'SAHOP Casa Flores'	6.33	39.1	338.6	0
269	'Victoria, Mexico'	1980	'Victoria Hospital Sotano'	6.33	6.07	274.5	0
283	'Irpinia, Italy-01'	1980	'Arienzo'	6.9	52.93	1000	0
284	'Irpinia, Italy-01'	1980	'Auletta'	6.9	9.52	1000	0
285	'Irpinia, Italy-01'	1980	'Bagnoli Irpinio'	6.9	8.14	1000	0
286	'Irpinia, Italy-01'	1980	'Bisaccia'	6.9	17.51	1000	0
288	'Irpinia, Italy-01'	1980	'Brienza'	6.9	22.54	500	0
289	'Irpinia, Italy-01'	1980	'Calitri'	6.9	13.34	600	0
290	'Irpinia, Italy-01'	1980	'Mercato San Severino'	6.9	29.79	350	0
292	'Irpinia, Italy-01'	1980	'Sturno'	6.9	6.78	1000	1

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse 0 = No 1 = Yes
					[km]	m/s	
293	'Irpinia, Italy-01'	1980	'Torre Del Greco'	6.9	59.63	659.6	0
294	'Irpinia, Italy-01'	1980	'Tricarico'	6.9	51.74	460	0
295	'Irpinia, Italy-02'	1980	'Auletta'	6.2	28.69	1000	0
296	'Irpinia, Italy-02'	1980	'Bagnoli Irpinio'	6.2	17.79	1000	0
297	'Irpinia, Italy-02'	1980	'Bisaccia'	6.2	14.73	1000	0
299	'Irpinia, Italy-02'	1980	'Brienza'	6.2	41.73	500	0
300	'Irpinia, Italy-02'	1980	'Calitri'	6.2	8.81	600	0
302	'Irpinia, Italy-02'	1980	'Rionero In Vulture'	6.2	22.68	530	0
303	'Irpinia, Italy-02'	1980	'Sturno'	6.2	20.38	1000	0
314	'Westmorland'	1981	'Brawley Airport'	5.9	15.28	208.7	0
315	'Westmorland'	1981	'Niland Fire Station'	5.9	15.16	207.5	0
316	'Westmorland'	1981	'Parachute Test Site'	5.9	16.54	348.7	1
317	'Westmorland'	1981	'Salton Sea Wildlife Refuge'	5.9	7.57	191.1	0
319	'Westmorland'	1981	'Westmorland Fire Sta'	5.9	6.18	193.7	1
322	'Coalinga-01'	1983	'Cantua Creek School'	6.36	23.78	271.4	0
323	'Coalinga-01'	1983	'Parkfield - Cholame 12W'	6.36	55.05	408.9	0
324	'Coalinga-01'	1983	'Parkfield - Cholame 1E'	6.36	42.76	338.5	0
326	'Coalinga-01'	1983	'Parkfield - Cholame 2WA'	6.36	43.83	184.8	0
327	'Coalinga-01'	1983	'Parkfield - Cholame 3E'	6.36	40.01	376.1	0
328	'Coalinga-01'	1983	'Parkfield - Cholame 3W'	6.36	44.82	338.5	0
329	'Coalinga-01'	1983	'Parkfield - Cholame 4AW'	6.36	46.73	338.5	0
330	'Coalinga-01'	1983	'Parkfield - Cholame 4W'	6.36	45.49	438.3	0
331	'Coalinga-01'	1983	'Parkfield - Cholame 5W'	6.36	47.88	289.6	0
333	'Coalinga-01'	1983	'Parkfield - Cholame 8W'	6.36	50.98	256.8	0
334	'Coalinga-01'	1983	'Parkfield - Fault Zone 1'	6.36	41.04	338.5	0
335	'Coalinga-01'	1983	'Parkfield - Fault Zone 10'	6.36	30.34	438.3	0
336	'Coalinga-01'	1983	'Parkfield - Fault Zone 11'	6.36	27.1	376.1	0
337	'Coalinga-01'	1983	'Parkfield - Fault Zone 12'	6.36	27.96	338.5	0
338	'Coalinga-01'	1983	'Parkfield - Fault Zone 14'	6.36	28.11	338.5	0
339	'Coalinga-01'	1983	'Parkfield - Fault Zone 15'	6.36	28	376.1	0
340	'Coalinga-01'	1983	'Parkfield - Fault Zone 16'	6.36	26.2	338.5	0
341	'Coalinga-01'	1983	'Parkfield - Fault Zone 2'	6.36	37.92	338.5	0
342	'Coalinga-01'	1983	'Parkfield - Fault Zone 3'	6.36	36.14	370.8	0
343	'Coalinga-01'	1983	'Parkfield - Fault Zone 4'	6.36	33.42	338.5	0
344	'Coalinga-01'	1983	'Parkfield - Fault Zone 6'	6.36	31.64	438.3	0
345	'Coalinga-01'	1983	'Parkfield - Fault Zone 7'	6.36	29.91	370.8	0
346	'Coalinga-01'	1983	'Parkfield - Fault Zone 8'	6.36	28.58	376.1	0
347	'Coalinga-01'	1983	'Parkfield - Fault Zone 9'	6.36	30.43	438.3	0
348	'Coalinga-01'	1983	'Parkfield - Gold Hill 1W'	6.36	35.04	338.5	0
349	'Coalinga-01'	1983	'Parkfield - Gold Hill 2E'	6.36	31.85	338.5	0
350	'Coalinga-01'	1983	'Parkfield - Gold Hill 2W'	6.36	35.93	376.1	0
351	'Coalinga-01'	1983	'Parkfield - Gold Hill 3E'	6.36	28.72	370.8	0
352	'Coalinga-01'	1983	'Parkfield - Gold Hill 3W'	6.36	38.1	438.3	0
353	'Coalinga-01'	1983	'Parkfield - Gold Hill 4W'	6.36	40.13	438.3	0
354	'Coalinga-01'	1983	'Parkfield - Gold Hill 5W'	6.36	42.72	438.3	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
355	'Coalinga-01'	1983	'Parkfield - Gold Hill 6W'	6.36	47.04	438.3	0
356	'Coalinga-01'	1983	'Parkfield - Stone Corral 2E'	6.36	35.29	376.1	0
357	'Coalinga-01'	1983	'Parkfield - Stone Corral 3E'	6.36	32.81	376.1	0
358	'Coalinga-01'	1983	'Parkfield - Stone Corral 4E'	6.36	30.3	376.1	0
359	'Coalinga-01'	1983	'Parkfield - Vineyard Cany 1E'	6.36	24.83	338.5	0
362	'Coalinga-01'	1983	'Parkfield - Vineyard Cany 2W'	6.36	29.01	338.5	0
363	'Coalinga-01'	1983	'Parkfield - Vineyard Cany 3W'	6.36	30.91	297.2	0
364	'Coalinga-01'	1983	'Parkfield - Vineyard Cany 4W'	6.36	33.28	376.1	0
366	'Coalinga-01'	1983	'Parkfield - Vineyard Cany 6W'	6.36	39.94	438.3	0
367	'Coalinga-01'	1983	'Pleasant Valley P.P. - bldg'	6.36	7.69	257.4	0
368	'Coalinga-01'	1983	'Pleasant Valley P.P. - yard'	6.36	7.69	257.4	0
369	'Coalinga-01'	1983	'Slack Canyon'	6.36	25.98	684.9	0
436	'Borah Peak, ID-01'	1983	'CPP-601'	6.88	82.6	424.8	0
437	'Borah Peak, ID-01'	1983	'CPP-610'	6.88	83	424.8	0
438	'Borah Peak, ID-01'	1983	'PBF (second bsmt)'	6.88	87.69	659.6	0
439	'Borah Peak, ID-01'	1983	'TAN-719'	6.88	84.8	424.8	0
			'TRA-642 ETR Reactor				
440	'Borah Peak, ID-01'	1983	Bldg(Bsmt)'	6.88	79.59	659.6	0
			'TRA-670 ATR Reactor				
441	'Borah Peak, ID-01'	1983	Bldg(Bsmt)'	6.88	80	659.6	0
446	'Morgan Hill'	1984	'APEEL 1E - Hayward'	6.19	51.68	220	0
447	'Morgan Hill'	1984	'Agnews State Hospital'	6.19	24.48	239.7	0
448	'Morgan Hill'	1984	'Anderson Dam (Downstream)'	6.19	3.22	488.8	0
449	'Morgan Hill'	1984	'Capitola'	6.19	39.08	288.6	0
450	'Morgan Hill'	1984	'Corralitos'	6.19	23.23	462.2	0
451	'Morgan Hill'	1984	'Coyote Lake Dam (SW Abut)'	6.19	0.18	597.1	1
452	'Morgan Hill'	1984	'Foster City - APEEL 1'	6.19	53.89	116.4	0
454	'Morgan Hill'	1984	'Gilroy - Gavilan Coll.'	6.19	14.83	729.7	0
455	'Morgan Hill'	1984	'Gilroy Array #1'	6.19	14.9	1428	0
456	'Morgan Hill'	1984	'Gilroy Array #2'	6.19	13.68	270.8	0
457	'Morgan Hill'	1984	'Gilroy Array #3'	6.19	13.01	349.9	0
458	'Morgan Hill'	1984	'Gilroy Array #4'	6.19	11.53	221.8	0
459	'Morgan Hill'	1984	'Gilroy Array #6'	6.19	9.85	663.3	0
460	'Morgan Hill'	1984	'Gilroy Array #7'	6.19	12.06	333.9	0
461	'Morgan Hill'	1984	'Halls Valley'	6.19	3.45	281.6	0
462	'Morgan Hill'	1984	'Hollister City Hall'	6.19	30.76	198.8	0
463	'Morgan Hill'	1984	'Hollister Diff Array #1'	6.19	26.42	215.5	0
464	'Morgan Hill'	1984	'Hollister Diff Array #3'	6.19	26.42	215.5	0
465	'Morgan Hill'	1984	'Hollister Diff Array #4'	6.19	26.42	215.5	0
466	'Morgan Hill'	1984	'Hollister Diff Array #5'	6.19	26.42	215.5	0
467	'Morgan Hill'	1984	'Hollister Diff. Array'	6.19	26.42	215.5	0
470	'Morgan Hill'	1984	'San Juan Bautista, 24 Polk St'	6.19	27.15	370.8	0
471	'Morgan Hill'	1984	'San Justo Dam (L Abut)'	6.19	31.88	622.9	0
472	'Morgan Hill'	1984	'San Justo Dam (R Abut)'	6.19	31.88	622.9	0
480	'Lazio-Abruzzo, Italy'	1984	'Pontecorvo'	5.8	29.58	338.6	0

Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse 0 = No 1 = Yes
					[km]	m/s	
495	'Nahanni, Canada'	1985	'Site 1'	6.76	2.48	659.6	0
496	'Nahanni, Canada'	1985	'Site 2'	6.76	0	659.6	1
497	'Nahanni, Canada'	1985	'Site 3'	6.76	4.93	659.6	0
511	'N. Palm Springs'	1986	'Anza - Red Mountain'	6.06	38.22	684.9	0
512	'N. Palm Springs'	1986	'Anza - Tule Canyon'	6.06	51.91	684.9	0
514	'N. Palm Springs'	1986	'Cabazon'	6.06	6.74	345.4	0
518	'N. Palm Springs'	1986	'Fun Valley'	6.06	12.79	345.4	0
520	'N. Palm Springs'	1986	'Hesperia'	6.06	71.7	345.4	0
522	'N. Palm Springs'	1986	'Indio'	6.06	35.34	207.5	0
527	'N. Palm Springs'	1986	'Morongo Valley'	6.06	3.67	345.4	0
529	'N. Palm Springs'	1986	'North Palm Springs'	6.06	0	345.4	0
530	'N. Palm Springs'	1986	'Palm Springs Airport'	6.06	10.08	207.5	0
531	'N. Palm Springs'	1986	'Puerta La Cruz'	6.06	67.38	370.8	0
532	'N. Palm Springs'	1986	'Rancho Cucamonga - FF'	6.06	77.98	390.2	0
535	'N. Palm Springs'	1986	'San Jacinto - Valley Cemetary'	6.06	30.71	338.5	0
538	'N. Palm Springs'	1986	'Sunnymead'	6.06	37.66	271.4	0
540	'N. Palm Springs'	1986	'Whitewater Trout Farm'	6.06	0	345.4	0
543	'Chalfant Valley-01'	1986	'Benton'	5.77	24.25	271.4	0
544	'Chalfant Valley-01'	1986	'Bishop - LADWP South St'	5.77	23.38	271.4	0
546	'Chalfant Valley-01'	1986	'Lake Crowley - Shehorn Res.'	5.77	24.37	338.5	0
547	'Chalfant Valley-01'	1986	'Zack Brothers Ranch'	5.77	6.07	271.4	0
548	'Chalfant Valley-02'	1986	'Benton'	6.19	21.55	271.4	0
549	'Chalfant Valley-02'	1986	'Bishop - LADWP South St'	6.19	14.38	271.4	0
550	'Chalfant Valley-02'	1986	'Bishop - Paradise Lodge'	6.19	14.97	345.4	0
551	'Chalfant Valley-02'	1986	'Convict Creek'	6.19	29.35	338.5	0
553	'Chalfant Valley-02'	1986	'Long Valley Dam (Downst)'	6.19	18.3	345.4	0
554	'Chalfant Valley-02'	1986	'Long Valley Dam (L Abut)'	6.19	18.3	345.4	0
556	'Chalfant Valley-02'	1986	'McGee Creek - Surface'	6.19	28.2	359.2	0
558	'Chalfant Valley-02'	1986	'Zack Brothers Ranch'	6.19	6.44	271.4	0
568	'San Salvador'	1986	'Geotech Investig Center'	5.8	2.14	545	1
569	'San Salvador'	1986	'National Geographical Inst'	5.8	3.71	350	1
586	'New Zealand-02'	1987	'Maraenui Primary School'	6.6	68.74	424.8	0
587	'New Zealand-02'	1987	'Matahina Dam'	6.6	16.09	424.8	0
591	'Whittier Narrows-01'	1987	'Anaheim - W Ball Rd'	5.99	25.81	234.9	0
592	'Whittier Narrows-01'	1987	'Arcadia - Campus Dr'	5.99	4.53	367.5	0
594	'Whittier Narrows-01'	1987	'Baldwin Park - N Holly'	5.99	4.34	308.6	0
595	'Whittier Narrows-01'	1987	'Bell Gardens - Jaboneria'	5.99	10.31	308.6	0
597	'Whittier Narrows-01'	1987	'Beverly Hills - 14145 Mulhol'	5.99	29.09	355.8	0
602	'Whittier Narrows-01'	1987	'Burbank - N Buena Vista'	5.99	20.37	271.4	0
603	'Whittier Narrows-01'	1987	'Calabasas - N Las Virg'	5.99	52.76	338.5	0
604	'Whittier Narrows-01'	1987	'Canoga Park - Topanga Can'	5.99	45.98	267.5	0
607	'Whittier Narrows-01'	1987	'Carson - Catskill Ave'	5.99	29.85	361.2	0
608	'Whittier Narrows-01'	1987	'Carson - Water St'	5.99	26.3	160.6	0
611	'Whittier Narrows-01'	1987	'Compton - Castlegate St'	5.99	18.32	308.6	0
613	'Whittier Narrows-01'	1987	'Covina - W Badillo'	5.99	9.49	271.4	0

Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
614	'Whittier Narrows-01'	1987	'Downey - Birchdale'	5.99	14.9	245.1	0
615	'Whittier Narrows-01'	1987	'Downey - Co Maint Bldg'	5.99	14.95	271.9	0
616	'Whittier Narrows-01'	1987	'El Monte - Fairview Av'	5.99	0.75	308.6	0
618	'Whittier Narrows-01'	1987	'Fountain Valley - Euclid'	5.99	36.69	270.2	0
619	'Whittier Narrows-01'	1987	'Garvey Res. - Control Bldg'	5.99	0.36	468.2	0
620	'Whittier Narrows-01'	1987	'Glendale - Las Palmas'	5.99	14.68	446	0
622	'Whittier Narrows-01'	1987	'Hacienda Heights - Colima'	5.99	9.6	337	0
624	'Whittier Narrows-01'	1987	'Huntington Beach - Lake St'	5.99	42.16	370.8	0
626	'Whittier Narrows-01'	1987	'LA - 116th St School'	5.99	18.23	301	0
627	'Whittier Narrows-01'	1987	'LA - Baldwin Hills'	5.99	21.51	297.1	0
628	'Whittier Narrows-01'	1987	'LA - Centinela St'	5.99	28	234.9	0
630	'Whittier Narrows-01'	1987	'LA - Century City CC South'	5.99	25.95	278	0
632	'Whittier Narrows-01'	1987	'LA - Cypress Ave'	5.99	8.56	446	0
633	'Whittier Narrows-01'	1987	'LA - E Vernon Ave'	5.99	10.5	308.6	0
634	'Whittier Narrows-01'	1987	'LA - Fletcher Dr'	5.99	11.07	446	0
637	'Whittier Narrows-01'	1987	'LA - N Figueroa St'	5.99	6	405.2	0
638	'Whittier Narrows-01'	1987	'LA - N Westmoreland'	5.99	15.34	315.1	0
641	'Whittier Narrows-01'	1987	'LA - Saturn St'	5.99	20.35	308.7	0
642	'Whittier Narrows-01'	1987	'LA - W 70th St'	5.99	16.77	294.2	0
645	'Whittier Narrows-01'	1987	'LB - Orange Ave'	5.99	19.8	270.2	0
646	'Whittier Narrows-01'	1987	'LB - Rancho Los Cerritos'	5.99	24.61	405.2	0
647	'Whittier Narrows-01'	1987	'LB - Recreation Park'	5.99	30.4	370.8	0
649	'Whittier Narrows-01'	1987	'La Habra - Briarcliff'	5.99	14.17	361.2	0
650	'Whittier Narrows-01'	1987	'La Puente - Rimgrove Av'	5.99	10.24	308.6	0
652	'Whittier Narrows-01'	1987	'Lakewood - Del Amo Blvd'	5.99	22.4	234.9	0
664	'Whittier Narrows-01'	1987	'N Hollywood - Coldwater Can'	5.99	28.37	446	0
666	'Whittier Narrows-01'	1987	'Newhall - W Pico Canyon Rd.'	5.99	55.93	285.9	0
667	'Whittier Narrows-01'	1987	'Northridge - 17645 Saticoy St'	5.99	38.04	280.9	0
668	'Whittier Narrows-01'	1987	'Norwalk - Imp Hwy, S Grnd'	5.99	14.37	270.2	0
672	'Whittier Narrows-01'	1987	'Pacoima Kagel Canyon USC'	5.99	31.8	271.4	0
673	'Whittier Narrows-01'	1987	'Panorama City - Roscoe'	5.99	32.13	271.4	0
675	'Whittier Narrows-01'	1987	'Pasadena - CIT Athenaeum'	5.99	4.18	415.1	0
677	'Whittier Narrows-01'	1987	'Pasadena - CIT Calif Blvd'	5.99	4.3	370.8	0
678	'Whittier Narrows-01'	1987	'Pasadena - CIT Indust. Rel'	5.99	4.3	370.8	0
683	'Whittier Narrows-01'	1987	'Pasadena - Old House Rd'	5.99	8.03	455.4	0
694	'Whittier Narrows-01'	1987	'Studio City - Coldwater Can'	5.99	26.91	294.2	0
695	'Whittier Narrows-01'	1987	'Sun Valley - Roscoe Blvd'	5.99	30.33	308.6	0
696	'Whittier Narrows-01'	1987	'Sun Valley - Sunland'	5.99	26.71	271.4	0
697	'Whittier Narrows-01'	1987	'Sunland - Mt Gleason Ave'	5.99	24.82	446	0
699	'Whittier Narrows-01'	1987	'Sylmar - Sayre St'	5.99	36.64	338.5	0
701	'Whittier Narrows-01'	1987	'Terminal Island - S Seaside'	5.99	37.67	229.8	0
702	'Whittier Narrows-01'	1987	'Torrance - W 226th St'	5.99	33.07	370.8	0
705	'Whittier Narrows-01'	1987	'West Covina - S Orange Ave'	5.99	6.42	308.6	0
706	'Whittier Narrows-01'	1987	'Whittier Narrows Dam upstream'	5.99	2.6	298.7	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
718	'Superstition Hills-01'	1987	'Wildlife Liquef. Array'	6.22	17.59	207.5	0
719	'Superstition Hills-02'	1987	'Brawley Airport'	6.54	17.03	208.7	0
720	'Superstition Hills-02'	1987	'Calipatria Fire Station'	6.54	27	205.8	0
721	'Superstition Hills-02'	1987	'El Centro Imp. Co. Cent'	6.54	18.2	192.1	0
722	'Superstition Hills-02'	1987	'Kornbloom Road (temp)'	6.54	18.48	207.5	1
723	'Superstition Hills-02'	1987	'Parachute Test Site'	6.54	0.95	348.7	0
724	'Superstition Hills-02'	1987	'Plaster City'	6.54	22.24	345.4	0
725	'Superstition Hills-02'	1987	'Poe Road (temp)'	6.54	11.16	207.5	0
726	'Superstition Hills-02'	1987	'Salton Sea Wildlife Refuge'	6.54	25.88	191.1	0
727	'Superstition Hills-02'	1987	'Superstition Mtn Camera'	6.54	5.61	362.4	0
728	'Superstition Hills-02'	1987	'Westmorland Fire Sta'	6.54	13.03	193.7	0
729	'Superstition Hills-02'	1987	'Wildlife Liquef. Array'	6.54	23.85	207.5	0
731	'Loma Prieta'	1989	'APEEL 10 - Skyline'	6.93	41.71	391.9	0
732	'Loma Prieta'	1989	'APEEL 2 - Redwood City'	6.93	43.06	133.1	0
733	'Loma Prieta'	1989	'APEEL 2E Hayward Muir Sch'	6.93	52.53	271.1	0
734	'Loma Prieta'	1989	'APEEL 3E Hayward CSUH'	6.93	52.39	517.1	0
735	'Loma Prieta'	1989	'APEEL 7 - Pulgas'	6.93	41.68	415.3	0
			'APEEL 9 - Crystal Springs				
736	'Loma Prieta'	1989	Res'	6.93	40.85	449.6	0
737	'Loma Prieta'	1989	'Agnews State Hospital'	6.93	24.27	239.7	0
			'Alameda Naval Air Stn				
738	'Loma Prieta'	1989	Hanger'	6.93	70.9	190	1
739	'Loma Prieta'	1989	'Anderson Dam (Downstream)'	6.93	19.9	488.8	0
740	'Loma Prieta'	1989	'Anderson Dam (L Abut)'	6.93	19.9	488.8	0
741	'Loma Prieta'	1989	'BRAN'	6.93	3.85	376.1	0
742	'Loma Prieta'	1989	'Bear Valley #1, Fire Station'	6.93	61.15	338.5	0
			'Bear Valley #10, Webb				
743	'Loma Prieta'	1989	Residence'	6.93	66.89	304.1	0
			'Bear Valley #12, Williams				
744	'Loma Prieta'	1989	Ranch'	6.93	50.71	331.2	0
			'Bear Valley #14, Upper Butts				
745	'Loma Prieta'	1989	Rn'	6.93	71.28	376.1	0
746	'Loma Prieta'	1989	'Bear Valley #5, Callens Ranch'	6.93	53.46	391	0
747	'Loma Prieta'	1989	'Bear Valley #7, Pinnacles'	6.93	68.22	597.1	0
748	'Loma Prieta'	1989	'Belmont - Envirotech'	6.93	43.94	627.6	0
749	'Loma Prieta'	1989	'Berkeley - Strawberry Canyon'	6.93	78.32	477.7	0
750	'Loma Prieta'	1989	'Berkeley LBL'	6.93	79.16	597.1	0
751	'Loma Prieta'	1989	'Calaveras Reservoir'	6.93	35.28	513.7	0
752	'Loma Prieta'	1989	'Capitola'	6.93	8.65	288.6	0
753	'Loma Prieta'	1989	'Corralitos'	6.93	0.15	462.2	0
754	'Loma Prieta'	1989	'Coyote Lake Dam (Downst)'	6.93	20.44	295	0
755	'Loma Prieta'	1989	'Coyote Lake Dam (SW Abut)'	6.93	19.97	597.1	0
756	'Loma Prieta'	1989	'Dublin - Fire Station'	6.93	58.68	271.4	0
			'Dumbarton Bridge West End				
757	'Loma Prieta'	1989	FF'	6.93	35.31	274.5	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
758	'Loma Prieta'	1989	'Emeryville - 6363 Christie'	6.93	76.87	198.7	0
759	'Loma Prieta'	1989	'Foster City - APEEL 1'	6.93	43.77	116.4	0
761	'Loma Prieta'	1989	'Fremont - Emerson Court'	6.93	39.66	284.8	0
762	'Loma Prieta'	1989	'Fremont - Mission San Jose'	6.93	39.32	367.6	0
763	'Loma Prieta'	1989	'Gilroy - Gavilan Coll.'	6.93	9.19	729.7	0
764	'Loma Prieta'	1989	'Gilroy - Historic Bldg.'	6.93	10.27	338.5	1
765	'Loma Prieta'	1989	'Gilroy Array #1'	6.93	8.84	1428	0
766	'Loma Prieta'	1989	'Gilroy Array #2'	6.93	10.38	270.8	0
767	'Loma Prieta'	1989	'Gilroy Array #3'	6.93	12.23	349.9	1
768	'Loma Prieta'	1989	'Gilroy Array #4'	6.93	13.81	221.8	0
769	'Loma Prieta'	1989	'Gilroy Array #6'	6.93	17.92	663.3	0
770	'Loma Prieta'	1989	'Gilroy Array #7'	6.93	22.36	333.9	0
771	'Loma Prieta'	1989	'Golden Gate Bridge'	6.93	79.71	641.6	0
772	'Loma Prieta'	1989	'Halls Valley'	6.93	30.25	281.6	0
773	'Loma Prieta'	1989	'Hayward - BART Sta'	6.93	54.01	370.8	0
775	'Loma Prieta'	1989	'Hollister - SAGO Vault'	6.93	29.54	684.9	0
776	'Loma Prieta'	1989	'Hollister - South & Pine'	6.93	27.67	370.8	0
777	'Loma Prieta'	1989	'Hollister City Hall'	6.93	27.33	198.8	0
778	'Loma Prieta'	1989	'Hollister Diff. Array'	6.93	24.52	215.5	0
779	'Loma Prieta'	1989	'LGPC'	6.93	0	477.7	1
780	'Loma Prieta'	1989	'Larkspur Ferry Terminal (FF)' 'Lower Crystal Springs Dam'	6.93	94.56	169.7	0
781	'Loma Prieta'	1989	dwnt'	6.93	48.24	712.8	0
782	'Loma Prieta'	1989	'Monterey City Hall'	6.93	39.69	684.9	0
783	'Loma Prieta'	1989	'Oakland - Outer Harbor Wharf'	6.93	74.16	248.6	0
784	'Loma Prieta'	1989	'Oakland - Title & Trust'	6.93	72.09	306.3	1
785	'Loma Prieta'	1989	'Olema - Point Reyes Station'	6.93	117.02	338.5	0
786	'Loma Prieta'	1989	'Palo Alto - 1900 Embarc.'	6.93	30.56	209.9	0
787	'Loma Prieta'	1989	'Palo Alto - SLAC Lab'	6.93	30.62	425.3	0
788	'Loma Prieta'	1989	'Piedmont Jr High'	6.93	72.9	895.4	0
789	'Loma Prieta'	1989	'Point Bonita'	6.93	83.37	1315.9	0
790	'Loma Prieta'	1989	'Richmond City Hall'	6.93	87.78	259.9	0
791	'Loma Prieta'	1989	'SAGO South - Surface'	6.93	33.94	684.9	0
792	'Loma Prieta'	1989	'SF - 1295 Shafter'	6.93	68.05	338.5	0
793	'Loma Prieta'	1989	'SF - Cliff House'	6.93	78.58	712.8	0
794	'Loma Prieta'	1989	'SF - Diamond Heights'	6.93	71.23	582.9	0
795	'Loma Prieta'	1989	'SF - Pacific Heights'	6.93	75.96	1249.9	0
796	'Loma Prieta'	1989	'SF - Presidio'	6.93	77.34	594.5	0
797	'Loma Prieta'	1989	'SF - Rincon Hill'	6.93	74.04	873.1	0
798	'Loma Prieta'	1989	'SF - Telegraph Hill'	6.93	76.4	712.8	0
799	'Loma Prieta'	1989	'SF Intern. Airport'	6.93	58.52	190.1	0
800	'Loma Prieta'	1989	'Salinas - John & Work'	6.93	28.66	271.4	0
801	'Loma Prieta'	1989	'San Jose - Santa Teresa Hills'	6.93	14.18	671.8	0
802	'Loma Prieta'	1989	'Saratoga - Aloha Ave'	6.93	7.58	370.8	0
803	'Loma Prieta'	1989	'Saratoga - W Valley Coll.'	6.93	8.48	370.8	1

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
804	'Loma Prieta'	1989	'So. San Francisco, Sierra Pt.'	6.93	63.03	1020.6	0
806	'Loma Prieta'	1989	'Sunnyvale - Colton Ave.'	6.93	23.92	267.7	0
807	'Loma Prieta'	1989	'Sunol - Forest Fire Station'	6.93	47.41	400.6	0
808	'Loma Prieta'	1989	'Treasure Island'	6.93	77.32	155.1	0
809	'Loma Prieta'	1989	'UCSC'	6.93	12.15	714	0
810	'Loma Prieta'	1989	'UCSC Lick Observatory'	6.93	12.04	714	0
811	'Loma Prieta'	1989	'WAHO'	6.93	11.03	376.1	0
812	'Loma Prieta'	1989	'Woodside'	6.93	33.87	454	0
813	'Loma Prieta'	1989	'Yerba Buena Island'	6.93	75.07	659.8	0
814	'Griva, Greece'	1990	'Edessa (bsmt)'	6.1	32.84	424.8	0
821	'Erzican, Turkey'	1992	'Erzincan'	6.69	0	274.5	1
825	'Cape Mendocino'	1992	'Cape Mendocino'	7.01	0	513.7	1
826	'Cape Mendocino'	1992	'Eureka - Myrtle & West'	7.01	40.23	338.5	0
827	'Cape Mendocino'	1992	'Fortuna - Fortuna Blvd'	7.01	15.97	457.1	0
828	'Cape Mendocino'	1992	'Petrolia'	7.01	0	712.8	1
829	'Cape Mendocino'	1992	'Rio Dell Overpass - FF'	7.01	7.88	311.8	0
832	'Landers'	1992	'Amboy'	7.28	69.21	271.4	0
833	'Landers'	1992	'Anaheim - W Ball Rd'	7.28	144.9	234.9	0
834	'Landers'	1992	'Arcadia - Arcadia Av'	7.28	137.25	308.6	0
835	'Landers'	1992	'Arcadia - Campus Dr'	7.28	135.22	367.5	0
836	'Landers'	1992	'Baker Fire Station'	7.28	87.94	271.4	0
837	'Landers'	1992	'Baldwin Park - N Holly'	7.28	131.92	308.6	0
838	'Landers'	1992	'Barstow'	7.28	34.86	370.8	0
839	'Landers'	1992	'Bell Gardens - Jaboneria'	7.28	154.26	308.6	0
840	'Landers'	1992	'Big Tujunga, Angeles Nat F'	7.28	144.13	446	0
841	'Landers'	1992	'Boron Fire Station'	7.28	89.69	345.4	0
842	'Landers'	1992	'Brea - S Flower Av'	7.28	137.44	308.6	0
843	'Landers'	1992	'Buena Park - La Palma'	7.28	150.09	308.6	0
844	'Landers'	1992	'Burbank - N Buena Vista'	7.28	157.94	271.4	0
845	'Landers'	1992	'Calabasas - N Las Virg'	7.28	190.05	338.5	0
846	'Landers'	1992	'Chatsworth - Devonshire'	7.28	172.45	376.1	0
847	'Landers'	1992	'Compton - Castlegate St'	7.28	161.23	308.6	0
848	'Landers'	1992	'Coolwater'	7.28	19.74	271.4	0
849	'Landers'	1992	'Covina - W Badillo'	7.28	128.06	271.4	0
850	'Landers'	1992	'Desert Hot Springs'	7.28	21.78	345.4	0
851	'Landers'	1992	'Downey - Co Maint Bldg'	7.28	157.46	271.9	0
852	'Landers'	1992	'Duarte - Mel Canyon Rd.'	7.28	126.33	446	0
853	'Landers'	1992	'El Monte - Fairview Av'	7.28	135.88	308.6	0
854	'Landers'	1992	'Featherly Park - Maint'	7.28	121.8	308.6	0
855	'Landers'	1992	'Fort Irwin'	7.28	62.98	345.4	0
856	'Landers'	1992	'Fountain Valley - Euclid'	7.28	146.89	270.2	0
857	'Landers'	1992	'Glendale - Las Palmas'	7.28	148.07	446	0
858	'Landers'	1992	'Glendora - N Oakbank'	7.28	122.61	446	0
859	'Landers'	1992	'Hacienda Heights - Colima'	7.28	136.29	337	0
860	'Landers'	1992	'Hemet Fire Station'	7.28	68.66	338.5	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
861	'Landers'	1992	'Huntington Bch - Waikiki'	7.28	156	234.9	0
862	'Landers'	1992	'Indio - Coachella Canal'	7.28	54.25	345.4	0
863	'Landers'	1992	'Inglewood - Union Oil'	7.28	167.27	316	0
864	'Landers'	1992	'Joshua Tree'	7.28	11.03	379.3	0
865	'Landers'	1992	'LA - 116th St School'	7.28	164.36	301	0
866	'Landers'	1992	'LA - E Vernon Ave'	7.28	157.69	308.6	0
867	'Landers'	1992	'LA - Fletcher Dr'	7.28	153.04	446	0
870	'Landers'	1992	'LA - Obregon Park'	7.28	151.7	349.4	0
871	'Landers'	1992	'LA - S Grand Ave'	7.28	161.56	308.6	0
872	'Landers'	1992	'LA - W 15th St'	7.28	160.99	405.2	0
873	'Landers'	1992	'LA - W 70th St'	7.28	163.96	294.2	0
874	'Landers'	1992	'LB - Orange Ave'	7.28	160.85	270.2	0
876	'Landers'	1992	'La Habra - Briarcliff'	7.28	143.12	361.2	0
877	'Landers'	1992	'La Puente - Ringrove Av'	7.28	132.08	308.6	0
878	'Landers'	1992	'Lakewood - Del Amo Blvd'	7.28	157.41	234.9	0
879	'Landers'	1992	'Lucerne'	7.28	2.19	684.9	0
883	'Landers'	1992	'Northridge - 17645 Saticoy St'	7.28	172.32	280.9	0
884	'Landers'	1992	'Palm Springs Airport'	7.28	36.15	207.5	0
885	'Landers'	1992	'Pomona - 4th & Locust FF'	7.28	117.5	229.8	0
886	'Landers'	1992	'Puerta La Cruz'	7.28	94.48	370.8	0
887	'Landers'	1992	'Riverside Airport'	7.28	96	370.8	0
888	'Landers'	1992	'San Bernardino - E & Hospitality'	7.28	79.76	271.4	0
889	'Landers'	1992	'San Gabriel - E Grand Ave'	7.28	141.92	401.4	0
890	'Landers'	1992	'Santa Fe Springs - E.Joslin'	7.28	150.1	308.6	0
891	'Landers'	1992	'Silent Valley - Poppet Flat'	7.28	50.85	684.9	0
892	'Landers'	1992	'Sun Valley - Roscoe Blvd'	7.28	163.54	308.6	0
893	'Landers'	1992	'Sun Valley - Sunland'	7.28	158.25	271.4	0
894	'Landers'	1992	'Sunland - Mt Gleason Ave'	7.28	151.53	446	0
895	'Landers'	1992	'Tarzana - Cedar Hill'	7.28	175.65	257.2	0
896	'Landers'	1992	'Tustin - E Sycamore'	7.28	136.72	234.9	0
897	'Landers'	1992	'Twentynine Palms'	7.28	41.43	684.9	0
898	'Landers'	1992	'Villa Park - Serrano Ave'	7.28	132.94	308.6	0
899	'Landers'	1992	'West Covina - S Orange Ave'	7.28	132.32	308.6	0
900	'Landers'	1992	'Yermo Fire Station'	7.28	23.62	353.6	0
942	'Northridge-01'	1994	'Alhambra - Fremont School'	6.69	35.66	550	0
943	'Northridge-01'	1994	'Anacapa Island'	6.69	65.84	821.7	0
944	'Northridge-01'	1994	'Anaheim - W Ball Rd'	6.69	66.33	234.9	0
945	'Northridge-01'	1994	'Anaverde Valley - City R'	6.69	37.67	446	0
946	'Northridge-01'	1994	'Antelope Buttes'	6.69	46.65	821.7	0
947	'Northridge-01'	1994	'Arcadia - Arcadia Av'	6.69	39.41	308.6	0
948	'Northridge-01'	1994	'Arcadia - Campus Dr'	6.69	41.11	367.5	0
949	'Northridge-01'	1994	'Arleta - Nordhoff Fire Sta'	6.69	3.3	297.7	0
950	'Northridge-01'	1994	'Baldwin Park - N Holly'	6.69	47.72	308.6	0
951	'Northridge-01'	1994	'Bell Gardens - Jaboneria'	6.69	41.27	308.6	0

Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
952	'Northridge-01'	1994	'Beverly Hills - 12520 Mulhol'	6.69	12.39	545.7	0
953	'Northridge-01'	1994	'Beverly Hills - 14145 Mulhol'	6.69	9.44	355.8	0
954	'Northridge-01'	1994	'Big Tujunga, Angeles Nat F'	6.69	19.1	446	0
955	'Northridge-01'	1994	'Brea - S Flower Av'	6.69	64.63	308.6	0
956	'Northridge-01'	1994	'Buena Park - La Palma'	6.69	59.35	308.6	0
957	'Northridge-01'	1994	'Burbank - Howard Rd.'	6.69	15.87	821.7	0
958	'Northridge-01'	1994	'Camarillo'	6.69	34.78	234.9	0
959	'Northridge-01'	1994	'Canoga Park - Topanga Can'	6.69	0	267.5	0
			'Canyon Country - W Lost				
960	'Northridge-01'	1994	Cany'	6.69	11.39	308.6	0
961	'Northridge-01'	1994	'Carson - Catskill Ave'	6.69	46.05	361.2	0
962	'Northridge-01'	1994	'Carson - Water St'	6.69	45.44	160.6	0
963	'Northridge-01'	1994	'Castaic - Old Ridge Route'	6.69	20.1	450.3	0
964	'Northridge-01'	1994	'Compton - Castlegate St'	6.69	42.96	308.6	0
965	'Northridge-01'	1994	'Covina - S Grand Ave'	6.69	57.29	405.2	0
966	'Northridge-01'	1994	'Covina - W Badillo'	6.69	53.21	271.4	0
967	'Northridge-01'	1994	'Downey - Birchdale'	6.69	45.68	245.1	0
968	'Northridge-01'	1994	'Downey - Co Maint Bldg'	6.69	43.2	271.9	0
969	'Northridge-01'	1994	'Duarte - Mel Canyon Rd.'	6.69	48.37	446	0
970	'Northridge-01'	1994	'El Monte - Fairview Av'	6.69	44.51	308.6	0
971	'Northridge-01'	1994	'Elizabeth Lake'	6.69	36.2	234.9	0
972	'Northridge-01'	1994	'Featherly Park - Maint'	6.69	82.01	308.6	0
973	'Northridge-01'	1994	'Garden Grove - Santa Rita'	6.69	63.73	234.9	0
974	'Northridge-01'	1994	'Glendale - Las Palmas'	6.69	21.64	446	0
977	'Northridge-01'	1994	'Hemet - Ryan Airfield'	6.69	144.62	338.5	0
978	'Northridge-01'	1994	'Hollywood - Willoughby Ave'	6.69	17.82	234.9	0
979	'Northridge-01'	1994	'Huntington Bch - Waikiki'	6.69	66.43	234.9	0
980	'Northridge-01'	1994	'Huntington Beach - Lake St'	6.69	74.7	370.8	0
981	'Northridge-01'	1994	'Inglewood - Union Oil'	6.69	37.18	316	0
982	'Northridge-01'	1994	'Jensen Filter Plant'	6.69	0	373.1	0
983	'Northridge-01'	1994	'Jensen Filter Plant Generator'	6.69	0	525.8	0
984	'Northridge-01'	1994	'LA - 116th St School'	6.69	36.39	301	0
985	'Northridge-01'	1994	'LA - Baldwin Hills'	6.69	23.51	297.1	0
986	'Northridge-01'	1994	'LA - Brentwood VA Hospital'	6.69	12.92	416.6	0
987	'Northridge-01'	1994	'LA - Centinela St'	6.69	20.36	234.9	0
988	'Northridge-01'	1994	'LA - Century City CC North'	6.69	15.54	278	0
989	'Northridge-01'	1994	'LA - Chalon Rd'	6.69	9.87	740.1	0
990	'Northridge-01'	1994	'LA - City Terrace'	6.69	35.03	365.2	0
991	'Northridge-01'	1994	'LA - Cypress Ave'	6.69	28.98	446	0
992	'Northridge-01'	1994	'LA - E Vernon Ave'	6.69	33.33	308.6	0
993	'Northridge-01'	1994	'LA - Fletcher Dr'	6.69	25.66	446	0
994	'Northridge-01'	1994	'LA - Griffith Park Observatory'	6.69	21.2	1015.9	0
995	'Northridge-01'	1994	'LA - Hollywood Stor FF'	6.69	19.73	316.5	0
996	'Northridge-01'	1994	'LA - N Faring Rd'	6.69	12.42	255	0
997	'Northridge-01'	1994	'LA - N Figueroa St'	6.69	30.19	405.2	0

Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
998	'Northridge-01'	1994	'LA - N Westmoreland'	6.69	23.4	315.1	0
1000	'Northridge-01'	1994	'LA - Pico & Sentous'	6.69	27.82	270.2	0
1001	'Northridge-01'	1994	'LA - S Grand Ave'	6.69	29.52	308.6	0
1002	'Northridge-01'	1994	'LA - S. Vermont Ave'	6.69	27.89	270.2	0
1003	'Northridge-01'	1994	'LA - Saturn St'	6.69	21.17	308.7	0
1004	'Northridge-01'	1994	'LA - Sepulveda VA Hospital'	6.69	0	380.1	0
1005	'Northridge-01'	1994	'LA - Temple & Hope'	6.69	28.82	376.1	0
1006	'Northridge-01'	1994	'LA - UCLA Grounds'	6.69	13.8	398.4	0
1007	'Northridge-01'	1994	'LA - Univ. Hospital'	6.69	32.39	376.1	0
1008	'Northridge-01'	1994	'LA - W 15th St'	6.69	25.6	405.2	0
			'LA - Wadsworth VA Hospital				
1009	'Northridge-01'	1994	North'	6.69	14.55	392.2	0
			'LA - Wadsworth VA Hospital				
1010	'Northridge-01'	1994	South'	6.69	14.55	413.8	0
1011	'Northridge-01'	1994	'LA - Wonderland Ave'	6.69	15.11	1222.5	0
1012	'Northridge-01'	1994	'LA 00'	6.69	9.87	706.2	0
1013	'Northridge-01'	1994	'LA Dam'	6.69	0	629	1
1014	'Northridge-01'	1994	'LB - City Hall'	6.69	53.94	381.2	0
1015	'Northridge-01'	1994	'LB - Rancho Los Cerritos'	6.69	47.79	405.2	0
1016	'Northridge-01'	1994	'La Crescenta - New York'	6.69	17.81	446	0
1017	'Northridge-01'	1994	'La Habra - Briarcliff'	6.69	58.32	361.2	0
1019	'Northridge-01'	1994	'Lake Hughes #1'	6.69	35.46	425.3	0
1020	'Northridge-01'	1994	'Lake Hughes #12A'	6.69	20.77	602.1	0
1021	'Northridge-01'	1994	'Lake Hughes #4 - Camp Mend'	6.69	31.27	821.7	0
			'Lake Hughes #4B - Camp				
1022	'Northridge-01'	1994	Mend'	6.69	31.3	554	0
1023	'Northridge-01'	1994	'Lake Hughes #9'	6.69	24.86	670.8	0
1024	'Northridge-01'	1994	'Lakewood - Del Amo Blvd'	6.69	53.57	234.9	0
1025	'Northridge-01'	1994	'Lancaster - Fox Airfield Grnd'	6.69	51.88	271.4	0
1026	'Northridge-01'	1994	'Lawndale - Osage Ave'	6.69	34.31	361.2	0
1027	'Northridge-01'	1994	'Leona Valley #1'	6.69	36.86	684.9	0
1028	'Northridge-01'	1994	'Leona Valley #2'	6.69	36.9	446	0
1029	'Northridge-01'	1994	'Leona Valley #3'	6.69	37	684.9	0
1030	'Northridge-01'	1994	'Leona Valley #4'	6.69	37.23	446	0
1031	'Northridge-01'	1994	'Leona Valley #5 - Ritter'	6.69	37.47	446	0
1032	'Northridge-01'	1994	'Leona Valley #6'	6.69	37.7	327.4	0
1033	'Northridge-01'	1994	'Littlerock - Brainard Can'	6.69	46.31	821.7	0
1034	'Northridge-01'	1994	'Malibu - Point Dume Sch'	6.69	26.77	349.5	0
1035	'Northridge-01'	1994	'Manhattan Beach - Manhattan'	6.69	33.56	405.2	0
1036	'Northridge-01'	1994	'Mojave - Hwys 14 & 58'	6.69	85.28	308.6	0
1037	'Northridge-01'	1994	'Mojave - Oak Creek Canyon'	6.69	75.64	338.5	0
1038	'Northridge-01'	1994	'Montebello - Bluff Rd.'	6.69	43.22	405.2	0
1039	'Northridge-01'	1994	'Moorpark - Fire Sta'	6.69	16.92	405.2	0
1040	'Northridge-01'	1994	'Mt Baldy - Elementary Sch'	6.69	71.3	338.5	0
1041	'Northridge-01'	1994	'Mt Wilson - CIT Seis Sta'	6.69	35.53	821.7	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
1042	'Northridge-01'	1994	'N Hollywood - Coldwater Can'	6.69	7.89	446	0
1043	'Northridge-01'	1994	'Neenach - Sacatara Ck'	6.69	51.61	308.6	0
1044	'Northridge-01'	1994	'Newhall - Fire Sta'	6.69	3.16	269.1	0
1045	'Northridge-01'	1994	'Newhall - W Pico Canyon Rd.'	6.69	2.11	285.9	1
1046	'Northridge-01'	1994	'Newport Bch - Irvine Ave. F.S'	6.69	83.01	405.2	0
1047	'Northridge-01'	1994	'Newport Bch - Newp & Coast'	6.69	82.03	370.8	0
1048	'Northridge-01'	1994	'Northridge - 17645 Saticoy St'	6.69	0	280.9	0
1049	'Northridge-01'	1994	'Pacific Palisades - Sunset'	6.69	13.34	446	0
1050	'Northridge-01'	1994	'Pacoima Dam (downstr)'	6.69	4.92	2016.1	0
1051	'Northridge-01'	1994	'Pacoima Dam (upper left)'	6.69	4.92	2016.1	0
1052	'Northridge-01'	1994	'Pacoima Kagel Canyon'	6.69	5.26	508.1	0
1053	'Northridge-01'	1994	'Palmdale - Hwy 14 & Palmdale'	6.69	41.37	551.6	0
1055	'Northridge-01'	1994	'Pasadena - N Sierra Madre'	6.69	35.77	455.4	0
1056	'Northridge-01'	1994	'Phelan - Wilson Ranch'	6.69	85.75	308.6	0
1057	'Northridge-01'	1994	'Playa Del Rey - Saran'	6.69	24.42	405.2	0
1058	'Northridge-01'	1994	'Point Mugu - Laguna Peak'	6.69	36.61	376.1	0
1059	'Northridge-01'	1994	'Port Hueneme - Naval Lab.'	6.69	47.58	271.4	0
1060	'Northridge-01'	1994	'Rancho Cucamonga - Deer Can'	6.69	79.83	821.7	0
1061	'Northridge-01'	1994	'Rancho Palos Verdes - Hawth' Rancho Palos Verdes -	6.69	48.02	477.7	0
1062	'Northridge-01'	1994	'Luconia'	6.69	50.47	508.8	0
1063	'Northridge-01'	1994	'Rinaldi Receiving Sta'	6.69	0	282.3	1
1064	'Northridge-01'	1994	'Riverside Airport'	6.69	98.83	370.8	0
1065	'Northridge-01'	1994	'Rolling Hills Est-Rancho Vista'	6.69	44.89	376.1	0
1066	'Northridge-01'	1994	'Rosamond - Airport'	6.69	64.6	345.4	0
1067	'Northridge-01'	1994	'San Bernardino - CSUSB Gr' San Bernardino - Co Service	6.69	103.04	345.4	0
1068	'Northridge-01'	1994	'Bldg - Freefield' San Bernardino - E & Hospitality'	6.69	107.59	271.4	0
1069	'Northridge-01'	1994	'Hospitality'	6.69	108.18	271.4	0
1070	'Northridge-01'	1994	'San Gabriel - E Grand Ave'	6.69	38.86	401.4	0
1071	'Northridge-01'	1994	'San Jacinto - CDF Fire Sta'	6.69	147.47	271.4	0
1072	'Northridge-01'	1994	'San Marino - SW Academy'	6.69	34.55	379.4	0
1073	'Northridge-01'	1994	'San Pedro - Palos Verdes'	6.69	53.24	376.1	0
1074	'Northridge-01'	1994	'Sandberg - Bald Mtn'	6.69	41.26	821.7	0
1075	'Northridge-01'	1994	'Santa Barbara - UCSB Goleta'	6.69	107.2	338.5	0
1076	'Northridge-01'	1994	'Santa Fe Springs - E.Joslin'	6.69	48.06	308.6	0
1077	'Northridge-01'	1994	'Santa Monica City Hall'	6.69	17.28	336.2	0
1078	'Northridge-01'	1994	'Santa Susana Ground'	6.69	1.69	715.1	0
1079	'Northridge-01'	1994	'Seal Beach - Office Bldg'	6.69	61.46	370.8	0
1082	'Northridge-01'	1994	'Sun Valley - Roscoe Blvd'	6.69	5.59	308.6	0
1083	'Northridge-01'	1994	'Sunland - Mt Gleason Ave'	6.69	12.38	446	0
1086	'Northridge-01'	1994	'Sylmar - Olive View Med FF'	6.69	1.74	440.5	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
1087	'Northridge-01'	1994	'Tarzana - Cedar Hill A'	6.69	0.37	257.2	0
1088	'Northridge-01'	1994	'Terminal Island - S Seaside'	6.69	53.43	229.8	0
1089	'Northridge-01'	1994	'Topanga - Fire Sta'	6.69	10.31	376.1	0
1090	'Northridge-01'	1994	'Tustin - E Sycamore'	6.69	81.66	234.9	0
1092	'Northridge-01'	1994	'Ventura - Harbor & California'	6.69	54.28	271.4	0
1093	'Northridge-01'	1994	'Villa Park - Serrano Ave'	6.69	76.38	308.6	0
1094	'Northridge-01'	1994	'West Covina - S Orange Ave'	6.69	51.46	308.6	0
1095	'Northridge-01'	1994	'Whittier - S. Alta Dr'	6.69	48.36	376.1	0
1096	'Northridge-01'	1994	'Wrightwood - Jackson Flat'	6.69	64.46	821.7	0
1097	'Northridge-01'	1994	'Wrightwood - Nielson Ranch'	6.69	81.54	345.4	0
1098	'Northridge-01'	1994	'Wrightwood - Swarthout'	6.69	71.51	338.5	0
1099	'Double Springs'	1994	'Woodfords'	5.9	12.48	345.4	0
1103	'Kobe, Japan'	1995	'FUK'	6.9	158.08	256	0
1105	'Kobe, Japan'	1995	'HIK'	6.9	95.72	256	0
1106	'Kobe, Japan'	1995	'KJMA'	6.9	0.94	312	0
1107	'Kobe, Japan'	1995	'Kakogawa'	6.9	22.5	312	0
1109	'Kobe, Japan'	1995	'MZH'	6.9	69.04	609	0
1111	'Kobe, Japan'	1995	'Nishi-Akashi'	6.9	7.08	609	0
1112	'Kobe, Japan'	1995	'OKA'	6.9	86.94	609	0
1113	'Kobe, Japan'	1995	'OSAJ'	6.9	21.35	256	0
1116	'Kobe, Japan'	1995	'Shin-Osaka'	6.9	19.14	256	0
1117	'Kobe, Japan'	1995	'TOT'	6.9	119.64	609	0
1126	'Kozani, Greece-01'	1995	'Kozani'	6.4	14.13	659.6	0
1127	'Kozani, Greece-01'	1995	'Larisa'	6.4	74.06	338.6	0
1136	'Dinar, Turkey'	1995	'Balikesir'	6.4	255.44	338.6	0
1140	'Dinar, Turkey'	1995	'Denizli'	6.4	85.8	338.6	0
1141	'Dinar, Turkey'	1995	'Dinar'	6.4	0	219.8	0
1143	'Dinar, Turkey'	1995	'Izmir Trigger #2'	6.4	250.31	659.6	0
1144	'Gulf of Aqaba'	1995	'Eilat'	7.2	43.29	354.9	0
1145	'Gulf of Aqaba'	1995	'Hadera'	7.2	365.14	277	0
1147	'Kocaeli, Turkey'	1999	'Ambarli'	7.51	68.09	175	0
1148	'Kocaeli, Turkey'	1999	'Arcelik'	7.51	10.56	523	1
1149	'Kocaeli, Turkey'	1999	'Atakoy'	7.51	56.49	274.5	0
1151	'Kocaeli, Turkey'	1999	'Balikesir'	7.51	180.24	338.6	0
1152	'Kocaeli, Turkey'	1999	'Bornova'	7.51	315.9	274.5	0
1153	'Kocaeli, Turkey'	1999	'Botas'	7.51	126.04	274.5	0
1154	'Kocaeli, Turkey'	1999	'Bursa Sivil'	7.51	65.53	659.6	0
1155	'Kocaeli, Turkey'	1999	'Bursa Tofas'	7.51	60.43	274.5	0
1156	'Kocaeli, Turkey'	1999	'Canakkale'	7.51	266.17	274.5	0
1157	'Kocaeli, Turkey'	1999	'Cekmece'	7.51	64.95	346	0
1159	'Kocaeli, Turkey'	1999	'Eregli'	7.51	141.37	659.6	0
1160	'Kocaeli, Turkey'	1999	'Fatih'	7.51	53.34	338.6	0
1161	'Kocaeli, Turkey'	1999	'Gebze'	7.51	7.57	792	0
1162	'Kocaeli, Turkey'	1999	'Goynuk'	7.51	31.74	424.8	0
1164	'Kocaeli, Turkey'	1999	'Istanbul'	7.51	49.66	424.8	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse 0 = No 1 = Yes
					[km]	m/s	
1165	'Kocaeli, Turkey'	1999	'Izmit'	7.51	3.62	811	0
1166	'Kocaeli, Turkey'	1999	'Iznik'	7.51	30.74	274.5	0
1167	'Kocaeli, Turkey'	1999	'Kutahya'	7.51	145.06	274.5	0
1168	'Kocaeli, Turkey'	1999	'Manisa'	7.51	293.35	659.6	0
1169	'Kocaeli, Turkey'	1999	'Maslak'	7.51	52.96	659.6	0
1170	'Kocaeli, Turkey'	1999	'Mecidiyekoy'	7.51	51.17	424.8	0
1172	'Kocaeli, Turkey'	1999	'Tekirdag'	7.51	164.18	659.6	0
1175	'Kocaeli, Turkey'	1999	'Usak'	7.51	226.72	274.5	0
1176	'Kocaeli, Turkey'	1999	'Yarimca'	7.51	1.38	297	1
1177	'Kocaeli, Turkey'	1999	'Zeytinburnu'	7.51	51.98	274.5	0
1599	'Duzce, Turkey'	1999	'Ambarli'	7.14	187.99	175	0
1600	'Duzce, Turkey'	1999	'Arcelik'	7.14	131.17	523	0
1601	'Duzce, Turkey'	1999	'Aslan R.'	7.14	130.8	274.5	0
1602	'Duzce, Turkey'	1999	'Bolu'	7.14	12.02	326	1
1603	'Duzce, Turkey'	1999	'Bursa Tofas'	7.14	166.07	274.5	0
1604	'Duzce, Turkey'	1999	'Cekmece'	7.14	182.78	346	0
1605	'Duzce, Turkey'	1999	'Duzce'	7.14	0	276	1
1606	'Duzce, Turkey'	1999	'Fatih'	7.14	167.29	338.6	0
1608	'Duzce, Turkey'	1999	'Hava Alani'	7.14	177.28	424.8	0
1609	'Duzce, Turkey'	1999	'Kocamustafapaba Tomb'	7.14	168.46	338.6	0
1610	'Duzce, Turkey'	1999	'Kutahya'	7.14	168.26	274.5	0
1611	'Duzce, Turkey'	1999	'Lamont 1058'	7.14	0.21	424.8	0
1612	'Duzce, Turkey'	1999	'Lamont 1059'	7.14	4.17	424.8	0
1613	'Duzce, Turkey'	1999	'Lamont 1060'	7.14	25.78	782	0
1614	'Duzce, Turkey'	1999	'Lamont 1061'	7.14	11.46	481	0
1615	'Duzce, Turkey'	1999	'Lamont 1062'	7.14	9.15	338	0
1616	'Duzce, Turkey'	1999	'Lamont 362'	7.14	23.41	517	0
1617	'Duzce, Turkey'	1999	'Lamont 375'	7.14	3.93	424.8	0
1618	'Duzce, Turkey'	1999	'Lamont 531'	7.14	8.03	659.6	0
1619	'Duzce, Turkey'	1999	'Mudurnu'	7.14	34.3	659.6	0
1620	'Duzce, Turkey'	1999	'Sakarya'	7.14	45.16	471	0
1621	'Duzce, Turkey'	1999	'Yarimca'	7.14	97.51	297	0
1626	'Sitka, Alaska'	1972	'Sitka Observatory'	7.68	34.61	659.6	0
1627	'Caldiran, Turkey'	1976	'Maku'	7.21	50.78	274.5	0
1628	'St Elias, Alaska'	1979	'Icy Bay'	7.54	26.46	274.5	0
1629	'St Elias, Alaska'	1979	'Yakutat'	7.54	80	274.5	0
1633	'Manjil, Iran'	1990	'Abbar'	7.37	12.56	724	0
1634	'Manjil, Iran'	1990	'Abhar'	7.37	75.58	274.5	0
1636	'Manjil, Iran'	1990	'Qazvin'	7.37	49.97	274.5	0
1637	'Manjil, Iran'	1990	'Rudsar'	7.37	63.96	274.5	0
1638	'Manjil, Iran'	1990	'Tehran - Building & Housing'	7.37	174.55	274.5	0
1639	'Manjil, Iran'	1990	'Tehran - Sarif University'	7.37	171.75	274.5	0
1640	'Manjil, Iran'	1990	'Tonekabun'	7.37	93.3	274.5	0
1643	'Sierra Madre'	1991	'LA - City Terrace'	5.61	23.67	365.2	0
1646	'Sierra Madre'	1991	'Pasadena - USGS/NSMP'	5.61	13.91	370.8	0

Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$ [km]	$V_{S30}$ m/s	Pulse 0 = No 1 = Yes
			Office'				
1740	'Little Skull Mtn,NV'	1992	'Station #1-Lathrop Wells'	5.65	14.12	274.5	0
1741	'Little Skull Mtn,NV'	1992	'Station #2-NTS Control Pt. 1'	5.65	23.83	659.6	0
1742	'Little Skull Mtn,NV'	1992	'Station #3-Beaty'	5.65	45.15	338.6	0
1743	'Little Skull Mtn,NV'	1992	'Station #4-Pahrump 2'	5.65	61.04	274.5	0
1744	'Little Skull Mtn,NV'	1992	'Station #5-Pahrump 1'	5.65	63.82	274.5	0
1745	'Little Skull Mtn,NV'	1992	'Station #6-Las Vegas Calico Basin'	5.65	99.44	659.6	0
1746	'Little Skull Mtn,NV'	1992	'Station #7-Las Vegas Ann Road'	5.65	98.85	274.5	0
1747	'Little Skull Mtn,NV'	1992	'Station #8-Death Valley Scotties Castle'	5.65	98.09	659.6	0
			'12440 Imperial Hwy, North Grn'				
1759	'Hector Mine'	1999	Grn'	7.13	176.59	274.5	0
1760	'Hector Mine'	1999	'Alhambra - LA Co PW HQ FF'	7.13	174.9	370.8	0
1761	'Hector Mine'	1999	'Altadena - Eaton Canyon'	7.13	166.11	370.8	0
1762	'Hector Mine'	1999	'Amboy'	7.13	41.82	271.4	0
1763	'Hector Mine'	1999	'Anza - Pinyon Flat'	7.13	89.98	724.9	0
1764	'Hector Mine'	1999	'Anza - Tripp Flats Training'	7.13	102.4	684.9	0
1765	'Hector Mine'	1999	'Arleta - Nordhoff Fire Sta'	7.13	193.8	297.7	0
1766	'Hector Mine'	1999	'Baker Fire Station'	7.13	64.08	271.4	0
1767	'Hector Mine'	1999	'Banning - Twin Pines Road'	7.13	83.43	684.9	0
1768	'Hector Mine'	1999	'Barstow'	7.13	61.2	370.8	0
1769	'Hector Mine'	1999	'Beverly Hills Pac Bell Bsmt'	7.13	195.87	294.2	0
1770	'Hector Mine'	1999	'Big Bear Lake - Fire Station'	7.13	61.85	338.5	0
1771	'Hector Mine'	1999	'Bombay Beach Fire Station'	7.13	120.69	257	0
1772	'Hector Mine'	1999	'Burbank Airport'	7.13	187.73	271.4	0
1773	'Hector Mine'	1999	'Cabazon'	7.13	76.89	345.4	0
1775	'Hector Mine'	1999	'Castaic - Old Ridge Route'	7.13	205.16	450.3	0
1776	'Hector Mine'	1999	'Desert Hot Springs'	7.13	56.4	345.4	0
			'Devore - Devore Water Company'				
1777	'Hector Mine'	1999	Company'	7.13	106.17	370.8	0
1778	'Hector Mine'	1999	'Downey - Co Maint Bldg'	7.13	184.17	271.9	0
1779	'Hector Mine'	1999	'El Centro Array #10'	7.13	186.86	202.9	0
1780	'Hector Mine'	1999	'Featherly Park - Maint'	7.13	149.43	308.6	0
1781	'Hector Mine'	1999	'Fillmore Pac Bell'	7.13	232.72	271.4	0
1782	'Hector Mine'	1999	'Forest Falls Post Office'	7.13	74.92	345.4	0
1783	'Hector Mine'	1999	'Fort Irwin'	7.13	65.04	345.4	0
1784	'Hector Mine'	1999	'Frink'	7.13	118.5	345.4	0
1785	'Hector Mine'	1999	'Fun Valley'	7.13	54.68	345.4	0
1786	'Hector Mine'	1999	'Heart Bar State Park'	7.13	61.21	684.9	0
1787	'Hector Mine'	1999	'Hector'	7.13	10.35	684.9	0
1788	'Hector Mine'	1999	'Hemet Fire Station'	7.13	104.55	338.5	0
1789	'Hector Mine'	1999	'Hesperia - 4th & Palm'	7.13	89.87	345.4	0
1790	'Hector Mine'	1999	'Huntington Beach - Lake St'	7.13	184.03	370.8	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse 0 = No 1 = Yes
					[km]	m/s	
1791	'Hector Mine'	1999	'Indio - Coachella Canal' 'Indio - Riverside Co Fair	7.13	73.55	345.4	0
1792	'Hector Mine'	1999	Grnds'	7.13	74	207.5	0
1793	'Hector Mine'	1999	'Jensen Filter Plant Generator'	7.13	196.78	525.8	0
1794	'Hector Mine'	1999	'Joshua Tree'	7.13	31.06	379.3	0
1795	'Hector Mine'	1999	'Joshua Tree N.M. - Keys View'	7.13	50.42	684.9	0
1796	'Hector Mine'	1999	'LA - 116th St School'	7.13	191.45	301	0
1797	'Hector Mine'	1999	'LA - City Terrace'	7.13	177.94	365.2	0
1798	'Hector Mine'	1999	'LA - MLK Hospital Grounds'	7.13	190.3	270.2	0
1799	'Hector Mine'	1999	'LA - Obregon Park'	7.13	179.29	349.4	0
1800	'Hector Mine'	1999	'LA - Pico & Sentous'	7.13	186.79	270.2	0
1801	'Hector Mine'	1999	'LA - Temple & Hope'	7.13	183.96	376.1	0
1802	'Hector Mine'	1999	'LAX Fire Station'	7.13	203.21	370.8	0
1803	'Hector Mine'	1999	'LB - City Hall'	7.13	194.67	381.2	0
1804	'Hector Mine'	1999	'La Canada - Wald Residence'	7.13	172.58	370.8	0
1805	'Hector Mine'	1999	'Lake Hughes #1'	7.13	184.96	425.3	0
1806	'Hector Mine'	1999	'Leona Valley - Fire Station #1'	7.13	172.05	345.4	0
1807	'Hector Mine'	1999	'Little Rock Post Office' 'Los Angeles - Acosta Residence'	7.13	146.51	442	0
1808	'Hector Mine'	1999	Residence'	7.13	172.49	376.1	0
1809	'Hector Mine'	1999	'Lytle Creek Fire Station'	7.13	112.26	345.4	0
1810	'Hector Mine'	1999	'Mecca - CVWD Yard'	7.13	91.96	345.4	0
1811	'Hector Mine'	1999	'Mentone Fire Station #9'	7.13	91.15	271.4	0
1812	'Hector Mine'	1999	'Mill Creek Ranger Station'	7.13	84.87	370.8	0
1813	'Hector Mine'	1999	'Morongo Valley'	7.13	53.17	345.4	0
1814	'Hector Mine'	1999	'Newhall - Fire Sta'	7.13	198.13	269.1	0
1815	'Hector Mine'	1999	'Newport Bch - Irvine Ave. F.S' 'North Palm Springs Fire Sta #36'	7.13	178.25	405.2	0
1816	'Hector Mine'	1999	#36'	7.13	61.84	345.4	0
1817	'Hector Mine'	1999	'North Shore - Durmid'	7.13	110.72	338.5	0
1818	'Hector Mine'	1999	'Pacoima Kagel Canyon'	7.13	186.31	508.1	0
1819	'Hector Mine'	1999	'Palmdale Fire Station' 'Pasadena - Fair Oaks & Walnut'	7.13	156.46	452.9	0
1820	'Hector Mine'	1999	Walnut'	7.13	171.89	430.7	0
1821	'Hector Mine'	1999	'Pomona - 4th & Locust FF'	7.13	143.36	229.8	0
1822	'Hector Mine'	1999	'Riverside Airport'	7.13	123.79	370.8	0
1823	'Hector Mine'	1999	'Salton City' 'San Bernardino - Del Rosa Wk Sta'	7.13	123.25	324.5	0
1824	'Hector Mine'	1999	Sta'	7.13	96.91	684.9	0
1825	'Hector Mine'	1999	'San Bernardino - E & Hospitality'	7.13	105.2	271.4	0
1826	'Hector Mine'	1999	'San Bernardino - Fire Sta. #10'	7.13	103.62	271.4	0
1827	'Hector Mine'	1999	'San Bernardino - Fire Sta. #4'	7.13	101.71	271.4	0
1828	'Hector Mine'	1999	'San Bernardino - Fire Sta. #9'	7.13	108.02	271.4	0
1829	'Hector Mine'	1999	'San Bernardino - Mont. Mem	7.13	104.95	271.4	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse
					[km]	m/s	0 = No 1 = Yes
			Pk'				
			'San Bernardino - N Verdemont				
1830	'Hector Mine'	1999	Sch'	7.13	104.67	370.8	0
1831	'Hector Mine'	1999	'San Jacinto - Soboba'	7.13	92.71	370.8	0
			'Seven Oaks Dam Project				
1832	'Hector Mine'	1999	Office'	7.13	87.2	659.6	0
1833	'Hector Mine'	1999	'Snow Creek'	7.13	72.88	345.4	0
			'Sylmar - County Hospital				
1834	'Hector Mine'	1999	Grounds'	7.13	191.59	338.5	0
1835	'Hector Mine'	1999	'Temecula - 6th & Mercedes'	7.13	134.11	370.8	0
1836	'Hector Mine'	1999	'Twentynine Palms'	7.13	42.06	684.9	0
1837	'Hector Mine'	1999	'Valyermo Forest Fire Station'	7.13	135.77	345.4	0
1838	'Hector Mine'	1999	'Whitewater Trout Farm'	7.13	62.91	345.4	0
1839	'Hector Mine'	1999	'Whittier - Scott & Whittier'	7.13	169.59	338.5	0
			'Whittier Narrows Dam				
1840	'Hector Mine'	1999	downstream'	7.13	169.83	298.7	0
1841	'Hector Mine'	1999	'Wrightwood - Nielson Ranch'	7.13	113.45	345.4	0
1842	'Hector Mine'	1999	'Wrightwood Post Office'	7.13	118.71	338.6	0
			'Anchorage - Aho Res				
2059	Alaska'	2002	(Basement)'	6.7	270.65	274.5	0
			'Anchorage - DOI Off. of				
2060	Alaska'	2002	Aircraft'	6.7	272.87	279.4	0
			'Anchorage - Dowl Eng				
2061	Alaska'	2002	Warehouse'	6.7	271.02	360	0
			'Anchorage - K2-02'				
2062	Alaska'	2002	'Anchorage - K2-02'	6.7	265.48	366	0
			'Anchorage - K2-03'				
2063	Alaska'	2002	'Anchorage - K2-03'	6.7	264.91	474	0
			'Anchorage - K2-04'				
2064	Alaska'	2002	'Anchorage - K2-04'	6.7	273.69	279.4	0
			'Anchorage - K2-05'				
2065	Alaska'	2002	'Anchorage - K2-05'	6.7	269.64	284	0
			'Anchorage - K2-06'				
2066	Alaska'	2002	'Anchorage - K2-06'	6.7	269.22	491	0
			'Anchorage - K2-07'				
2067	Alaska'	2002	'Anchorage - K2-07'	6.7	275.34	270	0
			'Anchorage - K2-08'				
2068	Alaska'	2002	'Anchorage - K2-08'	6.7	272.21	274	0
			'Anchorage - K2-09'				
2069	Alaska'	2002	'Anchorage - K2-09'	6.7	268.72	582	0
			'Anchorage - K2-10'				
2070	Alaska'	2002	'Anchorage - K2-10'	6.7	277.35	269	0
			'Anchorage - K2-11'				
2071	Alaska'	2002	'Anchorage - K2-11'	6.7	274.42	467.9	0
			'Anchorage - K2-12'				
2072	'Nenana Mountain,	2002	'Anchorage - K2-12'	6.7	272.52	514	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$	$V_{S30}$	Pulse 0 = No 1 = Yes
				[km]	m/s		
Alaska'							
2073	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-13'	6.7	278.07	354	0
2074	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-14'	6.7	275.58	467.9	0
2075	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-15'	6.7	279.32	412	0
2076	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-16'	6.7	277.15	424.8	0
2077	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-17'	6.7	242.85	274.5	0
2078	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-18'	6.7	216.47	274.5	0
2079	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-19'	6.7	271.9	191.3	0
2080	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-20'	6.7	276.74	279.4	0
2081	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-21'	6.7	275.24	279.4	0
2082	'Nenana Mountain, Alaska'	2002	'Anchorage - K2-22'	6.7	280.41	467.9	0
2083	'Nenana Mountain, Alaska'	2002	'Anchorage - NOAA Weather Fac.'	6.7	275.47	274.5	0
2084	'Nenana Mountain, Alaska'	2002	'Anchorage - New Fire Station #1'	6.7	267.24	274.5	0
2085	'Nenana Mountain, Alaska'	2002	'Anchorage - New Fire Station #7'	6.7	276.49	274.5	0
2086	'Nenana Mountain, Alaska'	2002	'Anchorage - Police Headquarters'	6.7	270.21	467.9	0
2087	'Nenana Mountain, Alaska'	2002	'Anchorage - State Fish & Game'	6.7	273.46	360	0
2088	'Nenana Mountain, Alaska'	2002	'Anchorage International Airport'	6.7	272.9	274.5	0
2089	'Nenana Mountain, Alaska'	2002	'Fairbanks - Ester Fire Station'	6.7	146.32	274.5	0
2090	'Nenana Mountain, Alaska'	2002	'Fairbanks - Geophysic. Obs, CIGO'	6.7	148.82	424.8	0
2091	'Nenana Mountain, Alaska'	2002	'TAPS Pump Station #07'	6.7	199.27	659.6	0
2092	'Nenana Mountain, Alaska'	2002	'TAPS Pump Station #08'	6.7	119.84	424.8	0
2093	'Nenana Mountain, Alaska'	2002	'TAPS Pump Station #09'	6.7	104.73	382.5	0
2094	'Nenana Mountain, Alaska'	2002	'Valdez - Valdez City Hall'	6.7	276.74	338.6	0

## Appendix A

RNS	Earthquake Name	Year	Station Name	$M_w$	$R_{JB}$ [km]	$V_{S30}$ m/s	Pulse 0 = No 1 = Yes
2095	'Denali, Alaska'	2002	'Anchorage - DOI Off. of Aircraft'	7.9	272.51	279.4	0
2096	'Denali, Alaska'	2002	'Anchorage - Dowl Eng Warehouse'	7.9	270.27	360	0
2097	'Denali, Alaska'	2002	'Anchorage - K2-02'	7.9	264.47	366	0
2098	'Denali, Alaska'	2002	'Anchorage - K2-03'	7.9	263.55	474	0
2099	'Denali, Alaska'	2002	'Anchorage - K2-04'	7.9	273.56	279.4	0
2100	'Denali, Alaska'	2002	'Anchorage - K2-05'	7.9	269.09	284	0
2101	'Denali, Alaska'	2002	'Anchorage - K2-06'	7.9	268.17	491	0
2102	'Denali, Alaska'	2002	'Anchorage - NOAA Weather Fac.'	7.9	275.13	274.5	0
2103	'Denali, Alaska'	2002	'Anchorage - New Fire Station #1'	7.9	266.61	274.5	0
2104	'Denali, Alaska'	2002	'Anchorage - New Fire Station #7'	7.9	275.91	274.5	0
2105	'Denali, Alaska'	2002	'Anchorage - Police Headquarters'	7.9	269.02	467.9	0
2106	'Denali, Alaska'	2002	'Anchorage - State Fish & Game'	7.9	272.64	360	0
2107	'Denali, Alaska'	2002	'Carlo (temp)'	7.9	49.94	963.9	0
2108	'Denali, Alaska'	2002	'Eagle River - AK Geologic Mat'	7.9	246.25	274.5	0
2109	'Denali, Alaska'	2002	'Fairbanks - Ester Fire Station'	7.9	139.27	274.5	0
2110	'Denali, Alaska'	2002	'Fairbanks - Geophysic. Obs, CIGO'	7.9	140.71	424.8	0
2111	'Denali, Alaska'	2002	'R109 (temp)'	7.9	42.99	963.9	0
2112	'Denali, Alaska'	2002	'TAPS Pump Station #08'	7.9	104.16	424.8	0
2113	'Denali, Alaska'	2002	'TAPS Pump Station #09'	7.9	53.01	382.5	0
2114	'Denali, Alaska'	2002	'TAPS Pump Station #10'	7.9	0.18	329.4	1
2115	'Denali, Alaska'	2002	'TAPS Pump Station #11'	7.9	126.4	376.1	0
2116	'Denali, Alaska'	2002	'TAPS Pump Station #12'	7.9	164.66	338.6	0
2117	'Denali, Alaska'	2002	'Valdez - Valdez City Hall'	7.9	239.28	338.6	0
2118	'Denali, Alaska'	2002	'Valdez - Valdez Dock Company'	7.9	239.52	659.6	0
3548	'Loma Prieta'	1989	'Los Gatos - Lexington Dam'	6.93	3.22	1070.3	0
3549	'Northridge-01'	1994	'Monte Nido Fire Station'	6.69	15.46	659.6	0
3550	'Northridge-01'	1994	'Loma Linda; VA Hospital, North Freefield'	6.69	112.32	274.5	0
3551	'Northridge-01'	1994	'Loma Linda; VA Hospital, South Freefield'	6.69	112.31	274.5	0

