UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Ground motion variability and its effect on the probabilistic seismic hazard analysis

by

Vasily Pavlenko

Submitted in partial fulfilment of the requirements for the degree
Philosophiae Doctor (Physics)
In the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria
August, 2016

Supervisor: Professor Andrzej Kijko

Declaration

I, Vasily Pavlenko declare that this thesis, which I hereby submit for the degree Philosophiae Doctor (Physics) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature:_____

Date: _____

Acknowledgements

I would like to express my sincere appreciation to my supervisor, Professor Andrzej Kijko, for guidance during this study, the provided opportunities, support, and encouraging discussions. I would like to thank my colleagues and co-authors, Ansie Smit from the University of Pretoria Natural Hazard Centre and Dr. Vunganai Midzi from the Council for Geoscience. I am very grateful to Adri van Heerden for stylistic corrections that made the text more readable. I am very greatful to the examiners, Dr. C.J.S. Fourie, Dr. V. Midzi, and Prof. P. Suhadolc for reviewing my research and providing very interesting recommendations. I appreciate the financial support provided by the University of Pretoria and the University of Pretoria Natural Hazard Centre.

Summary

The majority of injuries and casualties during earthquakes occur as a result of partial or complete collapse of buildings. The assessment of possible seismic ground motions for the purposes of earthquake-resistant design can be performed by following the deterministic or probabilistic methodology.

Chapter 1 presents an overview of the current practice in seismic hazard analysis with emphasis on PSHA. At present, the Cornell-McGuire method prevails in PSHA studies. Despite significant development and modifications, this method has several controversial aspects. Absence of an upper bound of the seismic hazard curve is one of the most disputable aspects of the method, as it leads to unrealistic ground motion estimates for very low probabilities of exceedance. This problem stems from using the unbounded log-normal distribution in the modelling of the ground motion variability.

The main objective of the study was to investigate this variability and suggest a more realistic probability distribution which would allow accounting for the finiteness of the ground motion induced by earthquake. Chapter 2 introduces the procedure that is suitable for studying the ground motion variability. Given the data sample, this procedure allows selecting the most plausible probability distribution from a set of candidate models. Chapter 3 demonstrates the application of this procedure to PGA data recorded in Japan. This analysis demonstrated the superiority of the GEVD in the vast majority of considered examples. Estimates of the shape parameter of the GEVD were negative in every considered example, indicating the presence of a finite upper bound of PGA. Therefore, the GEVD provides a model that is more realistic for the scatter of the logarithm of PGA, and the application of this model leads to a bounded seismic hazard curve.

In connection with a revival of interest in seismic intensity as an analogue for physical ground motion parameters, the problem of accounting for anisotropy in the attenuation of MMI is considered in Chapter 4. A set of four equations that could account for this anisotropy was proposed and the applicability of these equations was demonstrated by modelling the isoseismal maps of two well-recorded seismic events that have occurred in South Africa. The results demonstrated that, in general, the new equations were superior to the isotropic attenuation equation, especially as regards to the pronounced anisotropy.

As several different PSHA methods exist, it is important to know how the results of application of these methods corresponded to each other. Chapter 5 presents the comparative study of three major PSHA methods, namely, the Cornell-McGuire method, the Parametric-Historic method, and the method based on Monte Carlo simulations. Two regions in Russia were selected for comparison, and the PGA estimates were compared for return periods of 475 and 2475 years. The results indicated that the choice of a particular method for conducting PSHA has relatively little effect on the hazard estimates when the same seismic source model was used in the calculations. The considered PSHA methods would provide closely related results for areas of moderate seismic activity; however, the difference among the results would apparently increase with increasing seismic activity.

List of abbreviations

AIC	Akaike information criterion
APE	Annual probability of exceedance
CDF	Cumulative distribution function
DSHA	Deterministic seismic hazard analysis
EMS	European macroseismic scale
FMD	Frequency-magnitude distribution
GEVD	Generalised extreme value distribution
GMPE	Ground motion prediction equation
GPD	Generalised Pareto distribution
GR	Gutenberg-Richter
GSHAP	Global Seismic Hazard Assessment Program
IPE	Intensity prediction equation
KS	Kolmogorov-Smirnov
MCE	Maximum credible earthquake
ML	Maximum likelihood
MM	Modified Mercalli
MMI	Modified Mercalli intensity
MSK	Medvedev-Sponheuer-Karnik
PDF	Probability density function
PGA	Peak ground acceleration
PGV	Peak ground velocity
PSHA	Probabilistic seismic hazard analysis

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Introduction and background

Earthquakes count among the most deadly and destructive natural disasters, with the strongest tremors causing untold loss of life and ruinous damage to infrastructure. Such havoc was wrought on 7 December 1988 near Spitak, Armenia, by the largest earthquake (M 7.0) in the region since historical times. The earthquake caused the death of 25000 people, with 50000 injured, and half a million people left homeless (Hadjian, 1993). The scale of the damage was unprecedented, with 90% of the city of Spitak destroyed, whereas 50% of neighbouring Leninakan and 20% of Kirovakan were destroyed (Cisternas et al., 1989). The reconstruction costs amounted to 15 billion roubles (approximately 26 billion US dollars, according to the exchange rate at that time).

Such catastrophic events stimulate attempts to improve the understanding of the mechanism of earthquakes and to develop protective measures to reduce their negative effects. After the destructive earthquake that occurred on 17 January 1995 in Kobe, Japan, the government sponsored the establishment of dense networks of strong-motion stations, and, since 1997, the K-NET and KiK-net networks have been operating continuously. The earthquake data are transmitted to the seismological centre at Tsukuba, and the information is freely available through the website (www.kyoshin.bosai.go.jp).

During the operational period of these networks, a unique volume of strong-motion data has been accumulated. This material has facilitated significant progress in understanding the processes of the radiation and propagation of seismic waves and in assessing the local effects of earthquakes.

Along with the development of engineering seismology, infrastructure design practices have evolved. Earthquake-resistant design is intended to produce structures that can withstand a certain level of shaking without sustaining severe damage. Structures are typically classified by their degree of importance, i.e. infrastructure is considered critical if its failure could potentially amplify the damage and the negative effects of earthquake. Critical structures include, for example, nuclear power plants, nuclear waste repositories, dams, and chemical plants. Some degree of repairable damage would be tolerated for housing units and other non-critical structures; however, the design of critically important structures has to ensure their safe operation after an earthquake.

The input level of shaking for earthquake-resistant design is described by the design ground motion, and seismic hazard analysis aims to quantify the ground motion expected to occur at a particular site. This analysis can be performed by following either of two general approaches

(Reiter, 1990), namely, PSHA and DSHA.

PSHA and DSHA require equal amounts of seismological and geological information, including the history of the seismicity in the region, regional characteristics of the attenuation of seismic waves, information on active seismic faults near the site of interest, and the geological and soil conditions at the site. The fundamental difference, however, is in the process of estimating the seismic hazard. In PSHA, seismic hazard is calculated by taking into account the contributions from all the earthquakes expected to occur in the area surrounding the site during a specified period. DSHA aims to estimate the worst-case scenario ground motion by considering a set of earthquake scenarios and by selecting the one that would accumulate the highest amplitude ground motion at the site. DSHA does not consider the frequency of occurrence of such an event.

Both approaches have strengths and weaknesses, proponents and opponents, and, in practice, both PSHA and DSHA combine deterministic and probabilistic elements at varying degrees (Bommer, 2002). The choice of one approach or the other should be based on the specific goals and objectives of the particular project, and the quality of the available data should be taken into account. The positive features of both approaches are sometimes combined in practical applications. Accordingly, Orozova and Suhadolc (1999) have formulated a hybrid deterministic-probabilistic approach by combining the probabilistic model of earthquake recurrence with the deterministic method of generating synthetic accelerograms. Leyendecker et al. (2000) have applied an innovative and somewhat controversial approach by considering two seismic hazard maps, one of which was obtained by using PSHA for a probability of exceedance of 2% in 50 years, whereas the other was obtained by using DSHA. The hazard estimates from the two maps were combined in such a way that the deterministic estimates were used instead of the probabilistic wherever the probabilistic estimates were higher than the deterministic estimates were. The probabilistic values were used wherever the deterministic estimates turned out to be higher. Therefore, the results of DSHA were used as the upper bound estimates for ground motions.

McGuire (2001) proposed the use of both approaches simultaneously and presented a scheme that described the relative contribution of DSHA and PSHA according to the application. A detailed discussion on this topic, as well as a set of performance criteria for selecting the method of analysis can be found in Klügel (2008).

1.2 Deterministic Seismic Hazard Analysis

Currently, there is no standard for employing DSHA and this analysis can therefore vary in different parts of the world (Gupta, 2002). According to the definition by Kramer (1996), the DSHA incorporates the following steps:

- Identification and characterisation of all earthquake sources capable of producing significant ground motion at the site of interest.
- Estimation of the magnitude of the strongest possible earthquake, also referred to as the MCE, for each seismic source and the selection of a source-to-site distance parameter for each source zone. In most DSHA analyses, the shortest distance between the source zone and the site of interest is selected.

- Estimation of the ground motion at the site for each of the earthquake scenarios.
- Determination of hazard at the site (i.e. the maximum ground motion amplitude of all considered earthquake scenarios).

Several modifications to this procedure have been proposed by Klügel et al. (2006). Recent developments in DSHA include the neo-deterministic approach (NDSHA), a rapidly developing deterministic method that involves waveform modelling techniques and allows for time-dependent earthquake scenarios (Peresan et al., 2011; Zuccolo et al., 2011; Nekrasova et al., 2014).

Traditional DSHA, as defined by Kramer (1996), is usually criticised for being overly conservative. In addition, the complexity and uncertainty associated with the selection of earthquake scenarios for this procedure are also regarded as negative features (Kijko, 2011). Baker (2008) indicated that ground motion variability was an additional problem related to DSHA. Abrahamson (2000) argued that the MCE was not the earthquake scenario that would induce the worst-case ground motion.

1.3 Probabilistic Seismic Hazard Analysis

Similar to DSHA, PSHA is not a single method but a variety of methods that have several similarities and common concepts. In this regard, various techniques were introduced nearly simultaneously, with the milestone in the development of such techniques being the work by Cornell (1968), which was partially based on the ideas suggested by Esteva (McGuire, 2008). Nearly at the same time, Milne and Davenport (1969) developed an alternative technique based purely on seismic event catalogues and historical observations of earthquakes. A significantly more complex, but, at the same time, extremely powerful technique was developed by Molchan et al. (1970). However, despite the significant advantages of this technique, it has remained virtually unknown, probably because of the description in English, which was presented in a single brief paper.

Later, Veneziano et al. (1984) introduced a non-parametric method based solely on the information obtained from a seismic event catalogue. Kijko and Graham (1998; 1999) considered the advantages of parametric and historic methods and derived an alternative parametric-historic method. Another approach to PSHA is represented by the procedures based on the Monte Carlo simulations of long-term seismic catalogues (e.g. Ebel and Kafka, 1999; Musson, 2000; Shumilina et al., 2000; Assatourians and Atkinson, 2013).

The Cornell procedure is straightforward, easy to understand, and easy to implement; such features, together with the computer programs developed by McGuire (1976; 1978) have made the Cornell approach popular. To date, this procedure is the most common PSHA method.

1.4 The Cornell-McGuire method

Budnitz et al. (1997) described the modern Cornell-McGuire method comprehensively, providing detailed recommendations on performing the analysis in relation to the degree of importance of the study object. The Cornell-McGuire method encompasses the following steps:

- Delineation of the seismic zones, that have the potential to produce destructive ground shaking at the site of interest.

- Definition of the spatial, temporal, and magnitude distributions of the seismicity of each seismic source.
- Estimation of the ground motion at the site of interest for all considered earthquake scenarios.
- Calculation of the total hazard curve for the site by combining the contributions from all relevant seismic sources.

1.4.1 Seismic source characterisation

The initial step of the analysis is the detection of every seismic source capable of producing destructive ground shaking at the site of interest. If individual faults cannot be identified because of the short duration of the instrumental records, or because of the low seismic activity of the region, the locations of possible earthquakes are described by area sources.

As regards area sources, usually it is assumed that the earthquake locations are uniformly distributed over the area. Such non-informative probability distribution reflects a lack of knowledge of the characterisation of diffuse seismicity. However, the uniform distribution of epicentres within a source area can be a poor approximation of the actual spatial distribution (e.g. Molina et al., 2001; Kijko, 2011; Spada et al., 2011). Furthermore, the process of delineating the seismic source zones can be difficult and subjective, especially in regions of low seismic activity. The results of the delineation of seismic sources performed by different groups of experts can vary significantly (e.g. McGuire, 1993; Frankel, 1995), with a pronounced influence on the results of seismic hazard assessment.

The difficulties and uncertainties associated with the delineation of source zones have led to the development of alternative techniques. The methods of Milne and Davenport (1969) and Veneziano et al. (1984) are free from subjective decisions on the delineation of seismic sources, as these methods are based directly on earthquake catalogues. However, one significant drawback of these methods is the inability to take into account the stochastic variability of earthquake locations. This drawback can be overcome by introducing various smoothing techniques, as have been proposed, for example, by Frankel (1995), who used a Gaussian function to smooth the estimated activity rates, and by Woo (1996), who used a magnitude-dependent kernel function.

1.4.2 Seismic regime

1.4.2.1 Distribution of earthquake magnitude. The recurrence of earthquakes of different magnitudes is an important characteristic of an active fault or a source zone. The relation between the frequency of occurrence and the magnitude is specified by the FMD. The FMD discovered by Ishimoto and Iida (1939) and Gutenberg and Richter (1944) is the most widely used in seismic hazard studies, although it is not the only one (e.g. Merz and Cornell, 1973; Lomnitz-Adler and Lomnitz, 1979; Taylor et al., 1990; Molchan et al., 1997).

$$\log_{10}[N(M)] = a - bM \tag{1.1}$$

where $N(M)$ is a number of earthquakes with magnitude equal to or greater than M, a is the level of seismic activity and b describes the ratio between the number of small and large earthquakes (for natural seismicity usually $b \approx 1$).

This equation is usually referred to as the GR recurrence law. In its original formulation, the GR law has no upper bound and predicts unlimited earthquake magnitudes. Since magnitude serves as a measure of the elastic energy released during an earthquake, it has become clear that the GR law should not continue indefinitely to higher magnitudes. The physical requirements of finite energy release (Knopoff and Kagan, 1977), or, equivalently, a finite rate of seismic moment release (Anderson, 1979; Molnar, 1979), impose restrictions on the FMD. Therefore, the range of magnitudes for which the linear relation (1.1) holds, is bounded from both sides. If the magnitudes of earthquakes are assumed independent and identically distributed random variables, the bounded GR law is equivalent to assuming an exponential distribution, truncated from the top and shifted to the right. Its CDF is given by:

$$F_M(m) = \frac{1 - e^{-\beta(m-M_c)}}{1 - e^{-\beta(M_{max}-M_c)}}, \quad M_c \leq m \leq M_{max} \tag{1.2}$$

where $\beta = b\ln(10)$, and M_c and M_{max} are the bounds of the range.

The PDF of this distribution has the following form:

$$f_M(m) = \frac{\beta\, e^{-\beta(m-M_c)}}{1 - e^{-\beta(M_{max}-M_c)}}, \quad M_c \leq m \leq M_{max} \tag{1.3}$$

Particular attention has been devoted to modelling the recurrence of large earthquakes, as these events dominate the seismic hazard at large return periods. It was observed that relevant to some large faults, the GR law applies to small and moderate magnitudes, but underestimates the frequency of occurrence of the large earthquakes. These earthquakes are known as characteristic earthquakes, and are described by the characteristic earthquake model (Schwartz and Coppersmith, 1984; Wesnousky, 1994), usually derived from paleoseismic observations.

Although the GR law can be applied to various tectonic regimes and large space volumes, and can describe natural seismicity reasonably well, the GR law cannot be applied in all instances. Bimodal FMDs can be encountered, usually when populations of events with different properties are combined in the same sample (Main, 2000; Wiemer and Wyss, 2000). Instances that are more complicated can be observed in mining areas, where the FMD can be significantly nonlinear (Gibowicz and Kijko, 1994). In addition, nonlinear FMDs have been proposed by various authors to capture possible deviations from linearity. Merz and Cornell (1973) used a function with a quadratic magnitude term for seismic risk calculations, Lomnitz-Adler and Lomnitz (1979) modified the GR law to capture the curvature at the higher magnitude end, and Taylor et al. (1990) proposed a nonlinear model to describe the seismicity observed at Hokkaido Island, Japan. Fig. 1.1 shows the different shapes of the FMDs.

1.4.2.2 Model for earthquake occurrence in time. The homogeneous Poisson process is the basis for modelling earthquake sequence in the majority of PSHA methods, including the Cornell-McGuire method (Cornell, 1968; Cornell and Winterstein, 1988). The homogeneous Poisson

process is a memoryless stochastic process, which means that the occurrence of an event in any time interval is independent of the occurrence of any previous events.

It is worth noting that declustering of the earthquake catalogue is required to obtain a sequence of major events that could be modelled by the Poisson process. Sometimes, the removal of dependent seismic events results in the removal of up to two thirds of the catalogue (Gardner and Knopoff, 1974).

The seismic source becomes less hazardous after the occurrence of a large earthquake compared with sources that have been quiet recently. A long period will be required to accumulate enough strain energy to rupture again; therefore, the time-dependent models accounting for strain accumulation periods could produce more realistic hazard assessments. Various alternatives to the homogeneous Poisson process have been considered.

For instance, Patwardhan et al. (1980) suggested using a semi-Markov model, where the earthquake occurrence depends only on the previous earthquake. Shimazaki and Nakata (1980)

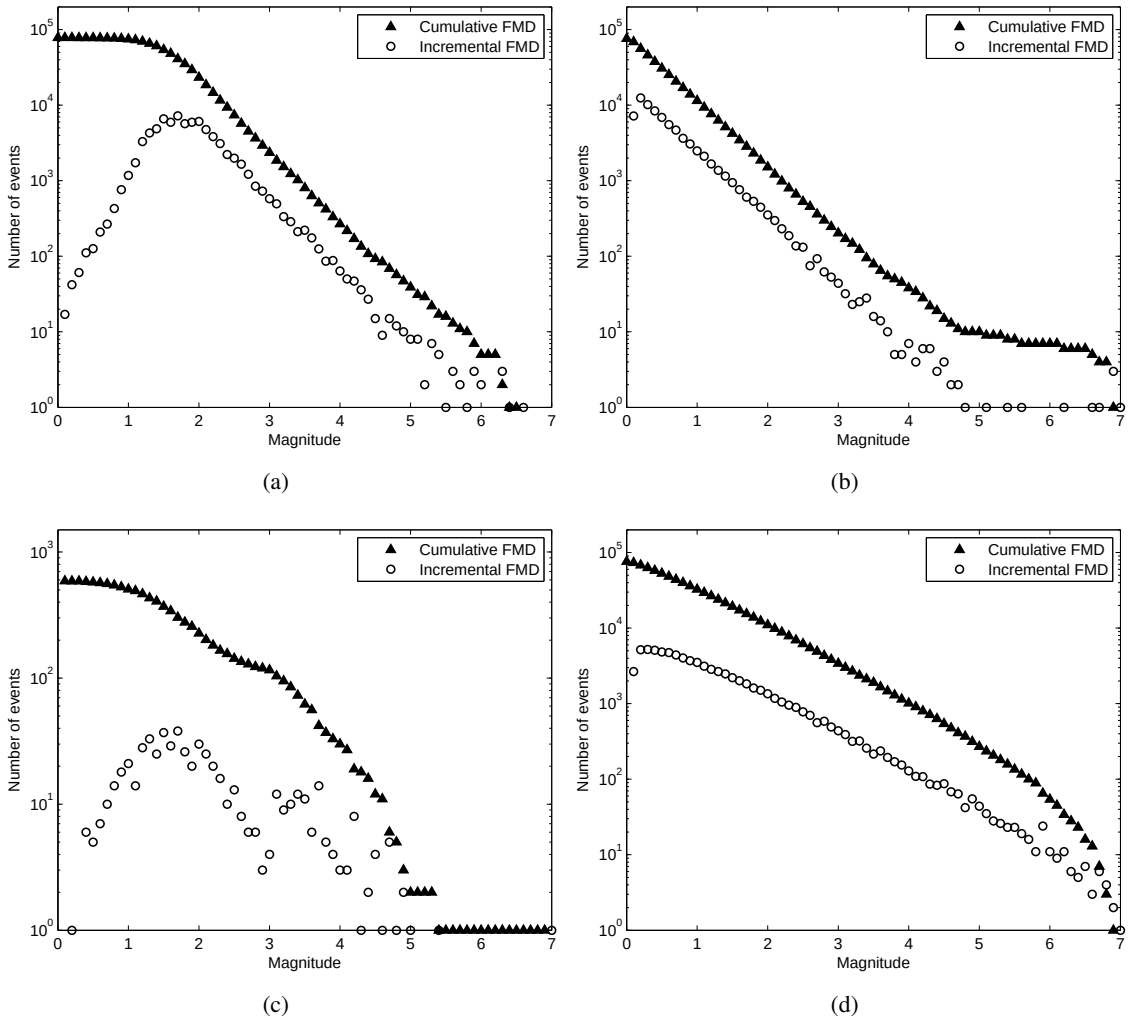


Figure 1.1: Comparison of FMDs. (a) Typical GR FMD. (b) FMD with characteristic earthquake. (c) Bimodal FMD. (d) Nonlinear FMD

suggested two time-dependent models, in which the time-predictable recurrence assumed that an earthquake would occur on a feature when a certain level of stress had accumulated. This regularity allows the prediction of the time of the next event. On the other hand, the slip-predictable recurrence assumes that independently of the accumulated level, the stress drops after each earthquake to some constant level; therefore, the longer the elapsed time since the last earthquake the larger would be the next earthquake.

The self-exciting model (Musmeci and Vere-Jones, 1992) assumes that each event that occurs in a restricted space-time volume increases the conditional occurrence rate within that volume. The self-correcting model (Ogata and Vere-Jones, 1984) assumes that the conditional occurrence rate depends on the current strain at point x_2 and time t_2. As the occurrence of an earthquake at a nearby point x_1 at earlier time t_1 decreases the strain at point x_2, it would generally decrease the conditional occurrence rate at point x_2 and time t_2.

Other time-dependent models focus on the statistical modelling of the inter-occurrence times. Various distributions have been used to fit the observed earthquake inter-occurrence times, among which are the gamma distribution (Udias and Rice, 1975), log-normal distribution (Nishenko and Buland, 1987; Goes, 1996), Weibull distribution (Parvez and Ram, 1997; Yakovlev et al., 2006; Zöller and Hainzl, 2007), the Brownian passage time distribution (Matthews et al., 2002), and the Bayesian combination of the latter three models (Fitzenz and Nyst, 2015).

Comparative studies aiming to measure the differences between the hazard assessments obtained by using the homogeneous Poisson process model and various time-dependent models (Cornell and Winterstein, 1988; Cramer et al., 2000) concluded that with rare exceptions, these differences were not significant enough to provide a basis for discarding the standard approach. On the other hand, recent studies by Boyd (2012) and Iervolino et al. (2014), allowing the inclusion of dependent events into PSHA, have reported an increase of up to 10% and 30%, respectively, of the ground motions at exceedance probability levels of engineering interest, caused by such inclusion of dependent events.

1.4.3 Ground motion at the site of interest

Estimates of ground motion at specific locations are fundamental inputs to seismic hazard analysis. The ground motion is characterised by a particular parameter, usually a horizontal PGA or PGV, spectral acceleration, or seismic intensity on one of the modern scales (MMI, MSK, EMS).

According to modern engineering seismology, the ground motion at the surface is controlled by three groups of factors, namely, the effects of the source (faulting mechanism, magnitude, radiation pattern, and spectral characteristic), the effects of the propagation path (source distance, geometrical spreading, and frequency-dependent inelastic absorption of seismic waves in the medium), and local effects at the site (the influence of surface topography and soil conditions).

There are two general approaches to estimating the ground motions. The first approach is more complex and involves seismogram simulations based on previously determined characteristics of the radiation and propagation of seismic waves (e.g. Costa et al., 1993; Boore, 2003). The second approach is more typical for PSHA and is based on empirical GMPEs, derived by using regression analysis (Boore and Joyner, 1982; Abrahamson and Youngs, 1992; Joyner and

Boore, 1993) of recorded strong ground motions.

The GMPEs should account for factors that contribute to surface ground motion and, at the same time, preserve a reasonable degree of simplicity to remain a useful tool in practical applications. The selection of the most appropriate GMPEs is vital for the proper assessment of seismic hazard. Because of the large number of published GMPEs (Douglas, 2011), it could be difficult to select appropriate models for a particular area. Various approaches have been developed to facilitate this selection, such as the data-driven procedures proposed by Scherbaum et al. (2004; 2009). On the other hand, in many regions of the world, there is a lack of strong-motion data and, therefore, no GMPEs specific to the region exist. In such situations, it is common to adopt the GMPEs developed for a region with similar tectonic properties. The applicability of the adopted GMPEs should be carefully checked by using the available data, as was discussed, for example, by Stafford et al. (2008) and Delavaud et al. (2009).

1.4.4 Hazard estimation

Seismic hazard is quantified by the probability $P(y \geq a_0, T)$ that the ground motion parameter y will exceed the value a_0 at a given site at least once during a specified period T. This probability is found from the mean rate $\lambda(a_0)$ of events that induce ground motions at the site, such that y exceeds a_0:

$$P(y \geq a_0, T) = 1 - e^{-\lambda(a_0)T} \tag{1.4}$$

This equation is based on the assumptions that the rate $\lambda(a_0)$ is constant in time and that the earthquake occurrences are modelled by the Poisson distribution. For a stationary seismic process, $\lambda(a_0)$ could be estimated simply by counting the number of events N_T that induced $y \geq a_0$ at the site during time interval T:

$$\lambda(a_0) = \frac{N_T}{T}$$

The longer the time period T the more accurate would be an estimate obtained from this equation. If the available catalogue of regional seismicity is complete above the lower magnitude of engineering interest, and is supplied with the estimates of ground motion occurring at the site because of each earthquake in the catalogue, and a_k corresponds to the ground motion generated by the k-th event, then the rate of exceedance at the site can be estimated as follows:

$$\lambda(a_0) = \frac{1}{T_0} \sum_k H(a_k - a_0) \tag{1.5}$$

where H is the Heaviside step function and T_0 is the duration of the catalogue.

This estimate would be unbiased only if the available catalogue were representative of all possible earthquakes that could affect the site. This could be achieved if the available catalogue covered an extremely long timespan, ranging from several thousand to hundreds of thousands of years, depending on the seismic activity of the region. Unfortunately, the available earthquake catalogues are generally too short to provide an unbiased estimation of the long-term activity rates.

Moreover, the ground motions are only available for instrumentally recorded earthquakes and only at the sites where seismic stations are deployed. The ground motion estimates are usually obtained by using GMPEs, which allow estimating the median ground motion at the site, and the associated uncertainty as functions of magnitude, source-to-site distance, and other parameters:

$$\ln(y) = f(m, r, \overline{\theta}) + \varepsilon\sigma \tag{1.6}$$

where y is the ground motion parameter of interest, f is the regression function, m is a magnitude, r is a distance measure, $\overline{\theta}$ is a vector of additional explanatory parameters, the term $\varepsilon\sigma$ represents the observed variability of ground motions for the specific earthquake scenario, σ is the standard deviation of the regression model, and ε is usually assumed to be a standard normal random variable.

Using the estimates y and σ, and information regarding the distribution of ε, it is possible to obtain the conditional probability $P(y \geq a_0|m,r)$ that a ground motion level a_0 will be exceeded at a distance r from the source of an earthquake with magnitude m.

Veneziano et al. (1984) proposed a method based on the historical seismicity information, in which the rate $\lambda(a_0)$ is estimated by summation over all the events in a catalogue:

$$\lambda(a_0) = \frac{1}{T_0}\sum_k Q_m P(y \geq a_0|m_k, r_k) \tag{1.7}$$

where Q_m is a correction factor accounting for the underrepresentation of magnitude m events in the catalogue, and m_k and r_k correspond to the k-th event in the catalogue.

Frankel (1995) described a method in which the regular grid is superimposed over the area of study, and the number of events with magnitudes above a threshold m_0 is counted from the seismic catalogue for each cell of the grid. This number of events in a cell is converted from a cumulative value (number of events with $m \geq m_0$) to an incremental value (number of events with $m \in (m_0, m_0 + \Delta m)$, where Δm is a small magnitude increment). The total incremental number of all cells within a certain distance increment r_k from the cite is denoted by N_k, and the annual rate $\lambda(a_0)$ is calculated as follows:

$$\lambda(a_0) = \sum_k \sum_l \frac{N_k}{T_0} 10^{-b(m_l - m_0)} P(y \geq a_0|m_l, r_k) \tag{1.8}$$

where l is the index for magnitude bin, the b-value is taken to be uniform throughout most of the area, and summation over magnitude is carried out to some maximum value.

It is interesting to note that in the original formulation by Cornell (1968), the uncertainty term $\varepsilon\sigma$ of Eq. (1.6) was neglected and that this shortcoming was corrected by Esteva (1969). The importance of this term was recognised and it was considered in later publications (Cornell, 1971; McGuire, 1976). In the modern Cornell-McGuire method, $\lambda(a_0)$ is obtained by using the following equation based on the total probability theorem:

$$\lambda(a_0) = \sum_i \Lambda_i \iint_{M\,R} P(y \geq a_0|m,r) f_{M_i}(m) f_{R_i}(r)\, dr\, dm \tag{1.9}$$

where Λ_i is the activity rate of the i-th seismic source, $f_{M_i}(m)$ and $f_{R_i}(r)$ are the PDFs of magnitude and distance of the i-th source, and summation is performed over all the seismic sources near the site.

By repeating these calculations and using Eq. (1.4), the seismic hazard curve can be constructed for any specific time interval. Seismic hazard at a given site is characterised by the ground motion level that has a specified probability of being exceeded during time interval T. The seismic hazard map can be constructed by applying this procedure to a grid of points. Fig. 1.2 illustrates the elements of the Cornell-McGuire method.

A common goal of seismic hazard analysis is to construct the design response spectrum, which is required for the purposes of structural analysis. In PSHA, this goal is achieved by constructing hazard curves for spectral accelerations at a range of periods, and selecting from each curve an acceleration that has a target probability of exceedance. The obtained spectral amplitudes are plotted versus corresponding periods and the result is termed the uniform hazard spectrum.

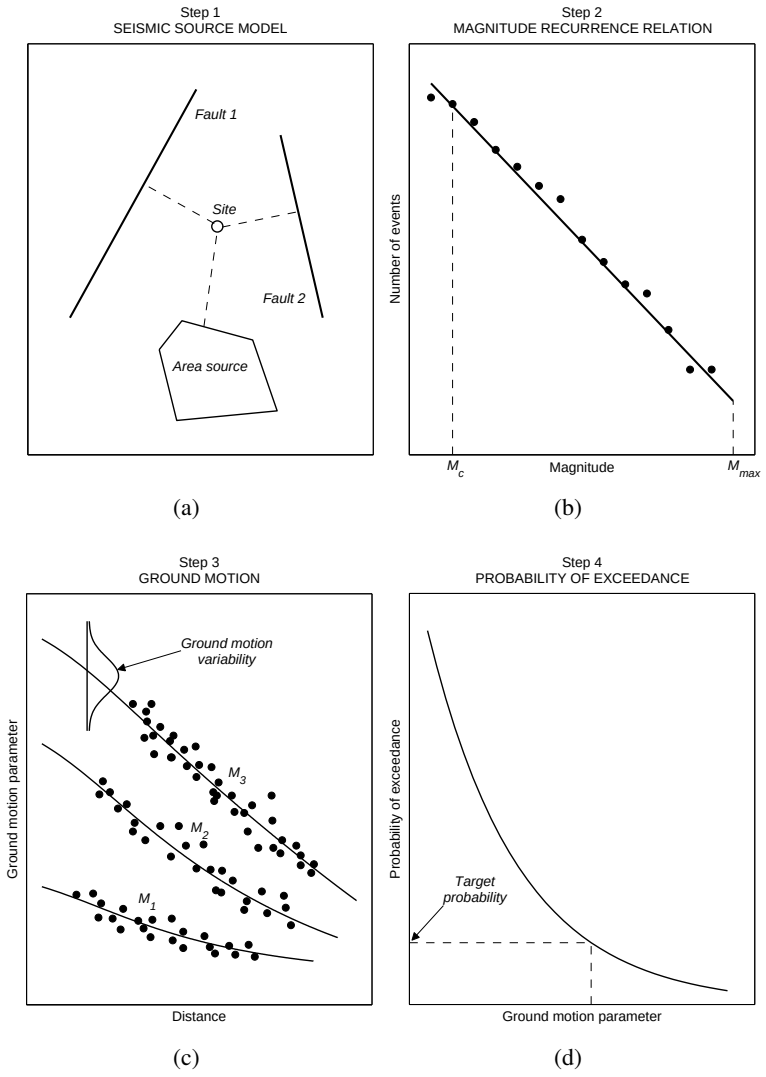


Figure 1.2: Elements of the Cornell-McGuire method

Since the contributions from all the relevant seismic sources are combined into a single hazard curve, the natural question, namely, which earthquake scenario makes the largest pronounced contribution to the result, has no clear answer. This issue is addressed by performing a disaggregation of seismic hazard (e.g. Chapman, 1995; McGuire, 1995; Bazzurro and Cornell, 1999; Romeo and Prestininzi, 2000).

1.4.5 Handling uncertainties

One of the benefits of PSHA is the possibility of taking into account any quantifiable uncertainties. These uncertainties are usually subdivided into two categories that need to be handled differently, namely, epistemic and aleatory uncertainties (Budnitz et al., 1997).

The terms uncertainty and randomness have also been used for epistemic and aleatory uncertainties, respectively. However, these terms are commonly used interchangeably in many seismic hazard studies, and, as a result, they were often confused. The terms epistemic and aleatory uncertainties were introduced to provide unambiguous terminology (Bommer, 2003). Fig. 1.3 illustrates the difference between the two types of uncertainties associated with ground motions.

Epistemic uncertainties are those that arise from a lack of knowledge and the imperfection of the present scientific models. These uncertainties could be reduced through further research and the gathering of more data. Epistemic uncertainties are associated with the use of particular models and are taken into account by using the logic tree formalism, which was introduced to PSHA by Kulkarni et al. (1984), and has become a standard element of PSHA (Reiter, 1990).

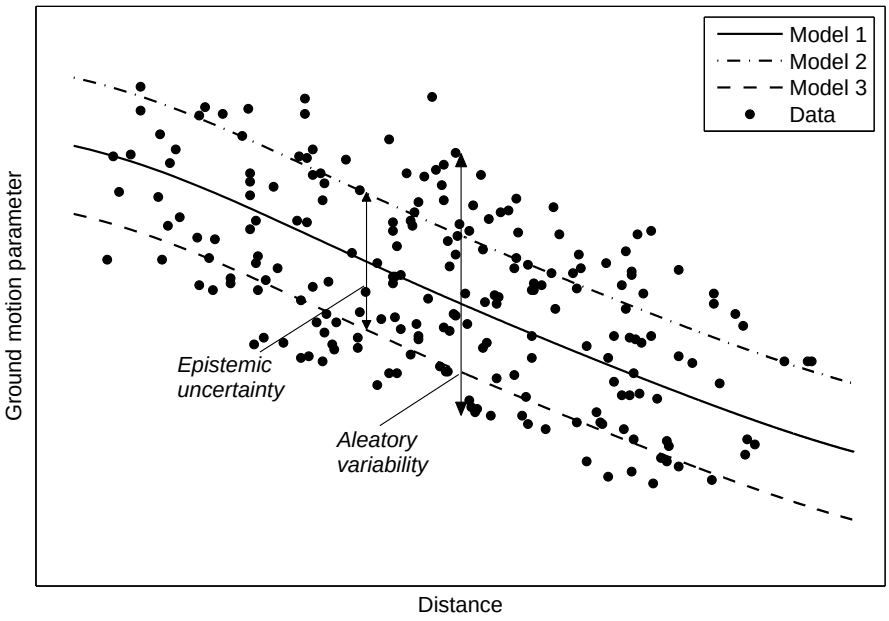


Figure 1.3: Uncertainties associated with ground motions

The logic tree must contain the best estimates of what is known and the potential range of alternatives used to characterise the uncertainty and to reflect the limitations of current knowledge (Bommer, 2012). Each branch of the tree is associated with a normalised weight, which reflects

expert opinion regarding the degree of feasibility of a particular hypothesis (Abrahamson and Bommer, 2005). The hazard calculations are performed for all the possible branches of the logic tree, resulting in a distribution of hazard curves, from which the mean, the median or other fractile can be obtained.

The process of constructing the logic tree and assigning weights to the branches is not a trivial matter, and, if done incorrectly, leads to underestimation of the seismic hazard (Abrahamson, 2000). Bommer et al. (2005) provided detailed analysis of the issues related to this process by using a logic tree for capturing epistemic uncertainty related to the selection of appropriate GMPEs. Fig. 1.4 demonstrates a simple example of a logic tree.

Aleatory uncertainties are those associated with inherent stochastic randomness of natural processes. According to a definition given by Budnitz et al. (1997), these uncertainties are irreducible and cannot be known in detail, although they are susceptible to analysis.

One of the most important aleatory uncertainties in seismic hazard studies is the ground motion variability associated with GMPEs (e.g. Strasser et al., 2008a; 2009). This variability is reflected in the range of possible values of the ground motion parameter for specific earthquake scenarios and is handled by integrating over the distribution of ground motion amplitudes. As indicated by Bommer and Abrahamson (2006), in earlier PSHA studies either the ground motion variability was often completely neglected or its effect was artificially reduced, which resulted in substantial underestimation of the seismic hazard.

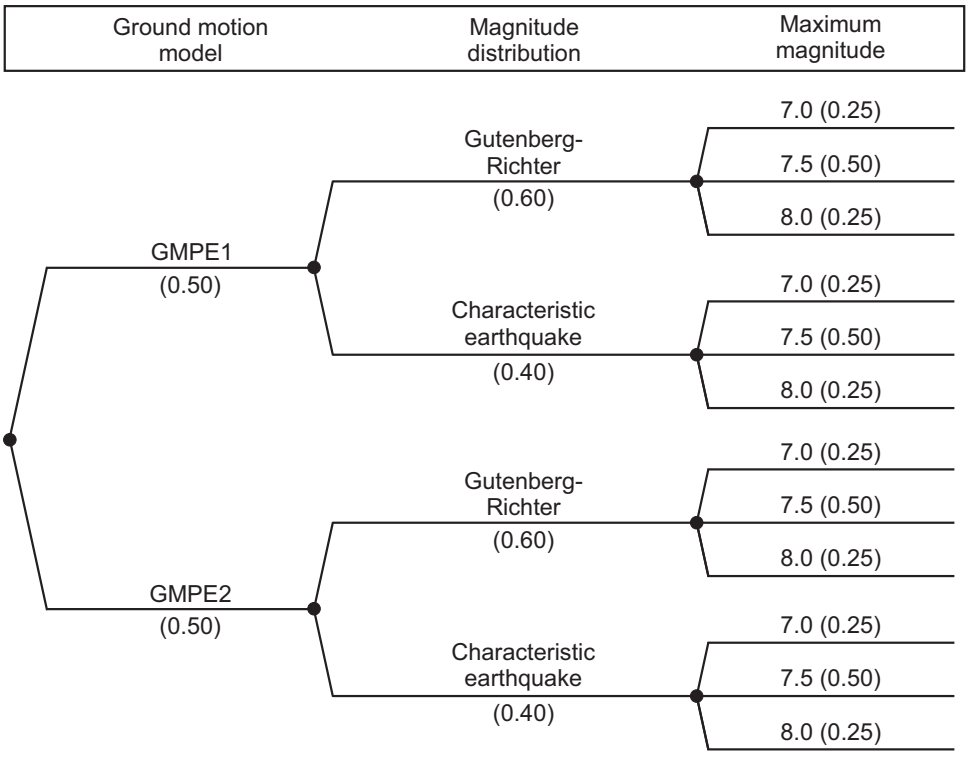


Figure 1.4: An example of a logic tree

1.5 Parameters of seismic regime

In the classic Cornell-McGuire method, each seismic source is characterised by four parameters:

- magnitude of completeness M_c
- annual rate of seismic events with magnitudes above M_c, Λ
- parameter b of the GR recurrence relation
- magnitude of the largest possible earthquake M_{max}.

The estimation of these parameters is associated with various difficulties, which are discussed in the following sections.

1.5.1 Magnitude of completeness M_c

The magnitude of completeness of seismic data M_c is defined as the lowest magnitude at which a seismic network is able to detect earthquakes reliably and completely. A fraction of seismic events with magnitudes below M_c is absent from the catalogue.

This parameter identifies the lower end of the linear segment in the FMD, and, therefore, accurate knowledge of M_c is essential for reliable estimation of the b-value and the activity rate Λ. Overestimating M_c reduces the amount of data, whereas, underestimating it could potentially lead to incorrect results of further analyses.

It is well known that M_c changes with time in most catalogues, usually decreasing because the number of seismic stations increases, in the process, improving the detection capabilities of the local networks. The spatial differences of M_c, which can be observed in catalogues collected from areas with different densities of seismic networks, should also be taken into account in large-scale seismic hazard studies. Typical observations of the seismicity of a source zone are shown in Fig. 1.5.

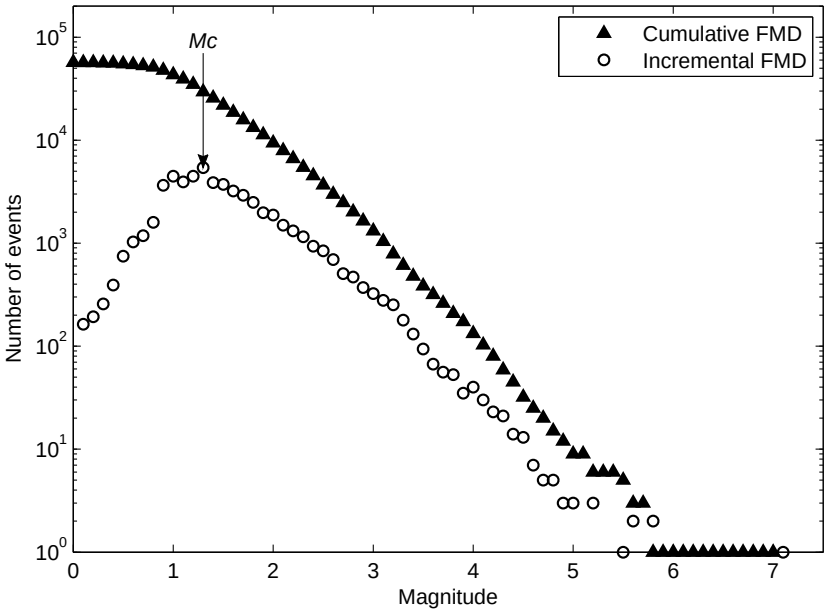


Figure 1.5: Example of the FMD of a source zone, the point that typically corresponds to M_c is also shown

The majority of existing methods for assessing M_c could be subdivided into two categories:

- Methods based on the information from seismic catalogues.
- Methods that investigate the detection capabilities by studying the signal-to-noise ratios at each station of the network (e.g. Ringdal, 1975; Gomberg, 1991; Kværna et al., 2002).

The methods of the first category prevail in seismic hazard studies; therefore, the present material will focus on these methods, as applied to instrumental catalogues. The completeness analysis of historical data is discussed, for example, by Albarello et al. (2001) and Stucci et al. (2004).

One of the first methods for estimating M_c was proposed by Stepp (1972), based on the assumption that the sequence of earthquakes follows a stationary Poisson process. In this instance, the mean rate of earthquake occurrence Λ is stable for the complete part of the catalogue, and its standard deviation is inversely proportional to the time interval $\sigma_\Lambda = \sqrt{\Lambda/T}$. A plot of $\sigma_\Lambda(T)$ for each specific magnitude allows detecting points where the graphs start to depart from $1/\sqrt{T}$ behaviour, indicating the end of the period of complete reporting.

Rydelek and Sacks (1989) developed a method to estimate M_c based on the assumptions that earthquakes follow a Poisson distribution and that detection capabilities increase at night relative to daytime (because of lower cultural activity and wind noise). The significant bias in the day-to-night detection of events in the given magnitude interval can be tested by using the Poisson assumption, and the catalogue is classified incomplete at and below this interval if the bias confirmed.

Wiemer and Wyss (2000) proposed two methods based on the assumption of self-similarity. The first method requires the detection of the point of maximum curvature of the frequency-magnitude curve. The second method is based on a comparison between the observed and the synthetic FMDs by using the goodness-of-fit statistic, calculated from the absolute difference between the numbers of events in the same magnitude bins of the two FMDs.

Cao and Gao (2002) proposed that M_c be estimated by using the stabilisation of the b-value as a function of the lower cut-off magnitude M_{co}. Assuming that the b-value increases for $M_{co} < M_c$ and $M_{co} \gg M_c$, and remains almost constant for $M_{co} \geq M_c$, the M_c can be estimated as the magnitude for which the increment between two successive values of b is sufficiently small (e.g. smaller than 0.03).

Kagan (2003) derived a method for estimating M_c based on fitting the Pareto distribution to the observed seismic moment data and using the KS test. Marsan (2003) proposed estimating M_c by assessing the likelihood of completeness, which is defined as the probability that the GR law, fitted to the data above the lower cut-off magnitude, would match the number of events in the magnitude bin that precedes the cut-off value. The M_c is selected so that two conditions must be met, namely, drop of the b-value for smaller magnitudes and drop of the likelihood of completeness at M_c.

Woessner and Wiemer (2005) developed a method based on the modelling of the entire magnitude range, in which the data below M_c is modelled by the normal distribution and the data above M_c is modelled by the GR law. The parameters of the model are estimated by the ML method and M_c is defined as the magnitude that maximises the log-likelihood function.

Amorèse (2007) employed the change-point test to detect the points in the incremental FMD,

where the significant and stable changes of the slope occur. Several change points can be found by this procedure, and the Wilcoxon-Mann-Whitney non-parametric test is applied to each such change point to test the null hypothesis that there is no change in the sequence at this point, and M_c is selected as the point for which the probability of making an error when rejecting the null hypothesis is the smallest.

Several techniques have been developed for mapping the spatially varying M_c, including the non-parametric method of Schorlemmer and Woesner (2008), the Bayesian method of Mignan et al. (2011), and the multiscale technique of Vorobieva et al. (2013).

The problem of estimating M_c has attracted much attention, as knowledge of this value is essential for many studies that involve statistical analysis of seismic catalogues. A comprehensive review and analysis of existing estimation methods for M_c can be found in Mignan and Woessner (2012). Despite the considerable progress that has been achieved, there is still no agreement on the applicability of the assumption of self-similarity in a range of small magnitudes (Rydelek and Sacks, 2003; Wiemer and Wyss, 2003).

1.5.2 Annual rate of seismic activity Λ and the b-value

In some earthquake sequences, temporal declines of the b-value have been observed before the main event (e.g. Gibowicz, 1973; Smith, 1981; Nuannin et al., 2005), which is in agreement with the hypothesis that the b-value is inversely related to the level of stress in the crust (Scholz, 1968; Amitrano, 2003; Schorlemmer et al., 2005). In addition, the temporal declines of seismic activity have been recognised as possible precursors of large earthquakes (Wyss and Habermann, 1988; Wyss et al., 1999; Zöller et al., 2002). Therefore, these parameters not only have practical importance for seismic hazard studies but also they can be useful for earthquake prediction.

The annual activity rate Λ is not usually discussed, as its ML estimator is simply the number of events per year:

$$\hat{\Lambda} = N/T \tag{1.10}$$

where T is a time span of the seismic catalogue.

The main concern, therefore, is the reliable estimation of the b-value. The initial studies on this topic were based on the unbounded GR law, and considered the probability distribution of unlimited earthquake magnitudes, for which the PDF has the following form:

$$f_M(m) = \beta e^{-\beta(m-M_c)}, \quad m \geq M_c \tag{1.11}$$

Utsu (1965) derived an estimator for b by applying the method of moments, whereas Aki (1965) obtained the same result by applying the ML method:

$$\hat{b} = \frac{\log_{10} e}{\bar{M} - M_c} \tag{1.12}$$

where $\bar{M}$ is the mean magnitude of the sample.

In addition, Aki (1965) provided a close analytical form for the standard deviation associated

with this estimator:

$$\sigma_{\hat{b}} = \frac{\hat{b}}{\sqrt{N}} \tag{1.13}$$

Shi and Bolt (1982) derived a new formula for the uncertainty of the b-value:

$$\sigma_{\hat{b}} = \frac{\hat{b}^2}{\log_{10}(e)} \sqrt{\frac{\sum_{i=1}^{N}(M_i - \bar{M})^2}{N(N-1)}} \tag{1.14}$$

Unfortunately, the estimator in Eq. (1.12) was derived without taking into account the presence of the finite upper bound magnitude $M_{\max}$. In addition, to be applicable, this formula requires a seismic catalogue that is complete above the threshold M_c. A typical seismic catalogue can be subdivided into two parts, namely, the pre-instrumental part that contains only the largest seismic events, and a relatively short instrumental part, which can be subdivided into a number of sub-catalogues, each with its own level of completeness M_c.

Subsequently, various generalisations and modifications of (1.12) have been proposed, aimed at solving the aforementioned issues, as well as other problems. Page (1968) considered a magnitude range with the upper bound $M_{\max}$ assumed as known, and obtained a recursive relation for estimating the b-value, which tends to (1.12) as the magnitude range $(M_c, M_{\max})$ increases.

Weichert (1980) generalised the ML estimators of the b-value and the activity rate Λ to allow utilising unequal observational periods and to account for the magnitude being a discrete and not a continuous variable, known with some accuracy, usually up to one decimal place. In addition, Bender (1983) discussed the influence of grouping magnitudes into classes and stressed the importance of the correction for bias, which is especially significant if magnitudes were recovered from historic intensity data, for which the grouping interval can be quite large. The corrected estimator can be expressed as:

$$\hat{b} = \frac{\log_{10} e}{\bar{M} - (M_c - \Delta M/2)} \tag{1.15}$$

where ΔM is the grouping interval of the catalogue.

The uncertainty of reported magnitudes is another factor that has significant influence on the estimation of seismic parameters. Tinti and Mulargia (1985) introduced the terms apparent magnitude $\tilde{M}$ and true magnitude M. According to this terminology, the reported apparent $\tilde{M}$ is the true magnitude, perturbed by a random error ε:

$$\tilde{M} = M + \varepsilon$$

Considering the normally distributed ε, Tinti and Mulargia (1985) concluded that the presence of magnitude uncertainty generally led to an overestimation of Λ, whereas the b-value remained unaffected. Various researchers suggested the b-value estimation technique, based on the least-squares fitting of a straight line to the frequency-magnitude plot (e.g. Guttorp, 1987; Nuannin et al., 2005).

In most parts of the world, instrumental seismic catalogues are relatively short; therefore,

employing the pre-instrumental catalogues would significantly expand the volume of available information. Kijko and Sellevoll (1989) expanded the ML estimators of seismicity parameters in the instance of mixed data, composed of pre-instrumental macroseismic observations and recent complete instrumental catalogues. In a subsequent study, Kijko and Sellevoll (1992) considered the effect of magnitude uncertainties on the results of estimation for instances of uniformly and normally distributed uncertainties.

A review of the most common estimation methods for the assessment of the b-value was provided by Marzocchi and Sandri (2003). These authors pointed out that the problem of the experimental estimation of b still had no universally recognised solution in the seismological community and therefore remained unsolved.

Castellaro et al. (2006) discussed in detail the influence of magnitude conversions on the estimates of seismic parameters. Converting magnitudes from one scale to another is a conventional step of pre-processing the data in seismic hazard studies, since a homogenised earthquake catalogue is required. In their study, Castellaro et al. (2006) have highlighted the superiority of the generalised orthogonal regression in deriving the relations for converting magnitudes. While application of the standard linear regression for this purpose leads to significantly biased assessment of Λ and b, the unbiased estimates can be obtained by converting the magnitudes by employing the generalised orthogonal regression.

Among the most recent developments, is the simple and general ML procedure introduced by Kijko and Smit (2012) for estimating b and Λ from multiple catalogues with different levels of completeness. A comprehensive study on the performance of the most frequently used b-value estimators has been conducted by Bengoubou-Valérius and Gibert(2013). These authors performed a series of statistical tests based on the Monte Carlo simulations and bootstrap analysis, and provided a number of recommendations regarding the b-value assessment. Summarising these recommendations, it can be concluded that the least-squares estimator should be discarded and, in general, the ML estimator should be preferred.

1.5.3 Magnitude of the largest possible earthquake M_{max}

In PSHA, $M_{\max}$ usually refers to the magnitude of the largest earthquake a seismic source is capable of producing. Alternatively, some authors suggest a formal smoothing of the FMD truncated in the range of large magnitudes (Main and Burton, 1984; Kagan, 1991), preserving the possibility of larger magnitudes. However, Pisarenko et al. (2010) put forward the definition of $M_{\max}$ as the maximum for a given future time interval. The following material is restricted to the definition traditionally used in PSHA. The estimation of $M_{\max}$ is crucial for seismic hazard analysis because large earthquakes produce the most severe ground motions and dominate the hazard at large return periods. The available methods for estimating $M_{\max}$ are classified as either deterministic or probabilistic.

Deterministic methods. The most frequently used methods of this category are those based on empirical relationships between the magnitude and the physical dimensions of the ruptured area (Kanamori and Anderson, 1975). Fault rupture length has been extensively used for estimating the earthquake magnitude, and numerous studies aimed at investigating the correlation between

magnitude and fault rupture length (e.g. Tocher, 1958; Mark, 1977; Bonilla et al., 1984; Wells and Coppersmith, 1994) have demonstrated the general nature of these relationships. As seismic moment is proportional to fault rupture area, it appears that magnitude would be more fundamentally related to fault rupture area than fault rupture length. The relationships in terms of fault rupture area (e.g. Geller, 1976; Wyss, 1979; Singh et al., 1980; Wells and Coppersmith, 1994) demonstrate smaller variations compared with the relations of fault rupture length. However, any attempt to specify the rupture scenario for a future earthquake is associated with considerable uncertainties that inevitably propagate to $M_{\max}$ estimates.

Alternatively, $M_{\max}$ can be inferred from the fault slip rate (e.g. Anderson and Luco, 1983; Youngs and Coppersmith, 1985) or from the seismic moment release rate (Smith, 1976). Therefore, assuming the applicability of the GR law with an upper bound $M_{\max}$, Smith (1976) derived the following formula:

$$\hat{M}_{\max} = \frac{\log_{10}[d(d-b)^{-1}T\dot{M}_0] - c}{d} \tag{1.16}$$

where c and d are constants of empirical relationship $\log_{10}(M_0) = dM_w + c$ between seismic moment M_0 and moment magnitude M_w, b is the slope of the GR law, $\dot{M}_0$ is a moment release rate, and T is the time interval.

In addition, Frohlich (1998), assuming that the GR law holds, proposed a simple estimator:

$$\hat{M}_{\max} = M_{\mathrm{c}} + \frac{1}{b}\log_{10} N \tag{1.17}$$

where N denotes the number of events in the catalogue with magnitudes above M_{c}.

Ward (1997) developed a technique for simulating the rupture processes on fault segments and estimating $M_{\max}$ resulting from interactions occurring between those segments.

Additionally, $M_{\max}$ is commonly estimated by extrapolating the FMD derived from the available seismic data to some large recurrence interval, for example, 500 or 1000 years. When the available data are insufficient to establish the FMD, $M_{\max}$ is sometimes obtained by adding an increment to the maximum observed magnitude $M_{\max}^{\mathrm{obs}}$:

$$\hat{M}_{\max} = M_{\max}^{\mathrm{obs}} + \delta \tag{1.18}$$

where δ varies from 0.0, which corresponds to the lower bound estimate of $M_{\max}$, and up to 1.0 (e.g. McGuire, 1993; Wheeler, 2009).

It has to be taken into account that the results of the deterministic estimation of $M_{\max}$ are generally associated with significant uncertainties that could be as high as one unit on the magnitude scale (Kijko, 2004).

Probabilistic methods. Various probabilistic methods have been developed over the years. Among the first, were those based on the extreme value theory. A study by Epstein and Lomnitz (1966), proved that the type I extreme value distribution for the largest magnitudes could be derived directly from the assumptions usually made in PSHA that the sequence of main seismic events is a

Poisson process and that the GR law is satisfied. Subsequently, various researchers have sought an estimation of $M_{\max}$ and corresponding recurrence intervals by fitting an extreme value distribution to a subset of maxima of observed magnitudes, usually annual maxima.

For instance, Kárník and Schenková (1974) and Schenková and Schenk (1975) applied the type I extreme value distribution to fit the maxima of magnitudes, while Yegulalp and Kuo (1974) and Schenková and Kárník (1978) applied the type III distribution, which assumes a finite upper bound of a variable. The latter obtained a better fit to the data compared with that of the type I extreme value distribution.

Subsequently, Kijko and Sellevoll (1981), using the modified GR law proposed by Lomnitz-Adler and Lomnitz (1979), derived a triple exponential distribution for the maxima of earthquake magnitudes. A statistical test based on the χ^2 statistic demonstrated that the triple exponential distribution approximations of the observations were superior to both the type I and the type III extreme value distributions.

Although the methods based on fitting a distribution to the maxima of earthquake magnitudes have the advantage of no small magnitude data being required, the main drawback of these methods is that the available data are generally incomplete, even for the maxima of magnitudes, in the pre-instrumental period and even in the recent periods of instrumental recordings. In general, a catalogue subdivided into a number of equal time intervals will demonstrate gaps in seismicity. To address this issue, Kijko and Dessokey (1987) proposed an improved technique that allows the use of maxima extracted from unequal time intervals.

In addition, the methods developed for estimating the endpoints of distribution functions were applied in the field of seismology. In general, the largest observation $M_{\max}^{\text{obs}}$ is a crucial parameter for these methods. An estimator of $M_{\max}$ can be written in a generic form:

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + \Delta \tag{1.19}$$

where Δ is a positive correction factor.

Let $M_1 \leq M_2 \leq \ldots \leq M_n$ be a sample of n earthquake magnitudes sorted in ascending order, $M_n = M_{\max}^{\text{obs}}$ and M_{n-1} is the second largest observation. Then the estimator derived by Robson and Whitlock (1964) takes the form:

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + (M_{\max}^{\text{obs}} - M_{n-1}) \tag{1.20}$$

Considering the above estimator, Cooke (1979) proposed that if the value of parameter ν (which determines the decay of the right tail of distribution) were known, an improved estimator, with a smaller mean squared error than (1.20), could be achieved:

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + (2\nu)^{-1}(M_{\max}^{\text{obs}} - M_{n-1}) \tag{1.21}$$

where ν is an exponent of Gnedenko's (1943) condition:

$$\lim_{M \to 0_-} \frac{1 - F_M(cM + M_{\max})}{1 - F_M(M + M_{\max})} = c^{1/\nu}$$

with c constant.

Another estimator proposed by Cooke (1980) can be applied when an extremely limited amount of information is available. This requires only a few largest observations and, with the modification discussed by Kijko and Singh (2011), can be written as:

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + \frac{1}{k}\left(M_{\max}^{\text{obs}} - \frac{1}{k-1}\sum_{i=2}^{k} M_{n-i+1}\right) \tag{1.22}$$

where k is the number of largest observations.

Furthermore, the methods based on generic equations for endpoints of distribution functions are widely applied. When the available information is sufficient for establishing the distribution of magnitude, $M_{\max}$ can be estimated by using the equation proposed by Tate (1959). Assuming that the expected value of $M_{\max}$ were equal to the largest observed value $E(M_{\max}) = M_{\max}^{\text{obs}}$, this equation could be written as:

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + \frac{1}{n f_M(M_{\max}^{\text{obs}})} \tag{1.23}$$

Pisarenko et al. (1996) derived unbiased and Bayesian estimators based on this equation. Another extensively used generic equation was derived by Cooke (1979):

$$\hat{M}_{\max} = M_{\max}^{\text{obs}} + \int_{M_c}^{M_{\max}} [F_M(m)]^n dm \tag{1.24}$$

Using Eq. (1.24), Kijko and Sellevoll (1989) derived an approximate equation that is valid for large samples and can be solved by an iterative scheme. Kijko and Graham (1998) considered its Bayesian analogue that accounts for moderate deviations from the assumed GR law, whereas Kijko and Singh (2011) achieved an exact solution for Eq. (1.24).

When no specific parametric function $F_M(m)$ is assumed, Eq. (1.24) can be applied in combination with some non-parametric technique for approximating the magnitude distribution; for example, with the kernel smoothing method, as demonstrated by Kijko et al. (2001), or with approximation based on order statistics, as discussed by Kijko and Singh (2011). This approach is especially useful where the empirical magnitude distribution is nonlinear, multi-modal, or if the characteristic event is observed.

An extremely powerful estimation procedure, which allows incorporating all the available information from various sources into the analysis, is the Bayesian estimation technique formulated by Cornell (1994) and improved by Kijko (2012).

Finally, $M_{\max}$ can be estimated by fitting a specified distribution $F_M(m)$ to the observed FMD, with the use of the standard least-squares regression, by minimising the absolute value of the misfit, or by using the adaptive regression procedure described by Kijko (1994).

Each of the described methods has advantages and limitations, which are discussed in detail by Wheeler (2009) and Kijko and Singh (2011). The choice of the most appropriate procedure should be based on the assumptions regarding the analysed seismicity and should take into account the quality and quantity of the available information.

1.6 Controversies of the Cornell-McGuire method

Despite significant development in the Cornell-McGuire method, a degree of controversy remains, that requires attention and further research. Several aspects of the Cornell-McGuire method are subject to criticism, among which are the treatment of uncertainties (e.g. Klügel, 2008; 2011), and the use of logic tree formalism (e.g. Krinitzsky, 1995; 2003; Castaños and Lomnitz, 2002). In addition, some authors have raised question on the correctness of the mathematical foundations of the method (e.g. Wang and Zhou, 2007; Wang, 2009).

The upper bound of a hazard curve has been indicated as an absent piece of PSHA (Bommer, 2002), and such absence has led to extremely inconsistent hazard assessments for very low probabilities of exceedance (Stepp et al., 2001; Corradini, 2003; Stamatakos, 2004; Klügel, 2005). The main factor contributing to seismic hazard assessments at low probabilities is ground motion variability. The log-normal distribution commonly used to model ground motion variability is unbounded, and, as a result, unrealistic high values of ground motion parameters are encountered when very low probabilities are considered in the analysis.

Recently, a number of devastating earthquakes have occurred in unexpected locations (Ellsworth, 2012), in particular the 2011 Tohoku earthquake. Such events have stimulated debate on how adequate the available seismic hazard maps and the methods applied for their preparation were (e.g. Stein et al., 2011; 2012; Hanks et al., 2012; Stirling, 2012), emphasising the necessity of testing the seismic hazard assessments with observations.

The results of the GSHAP (1992-1999) have been subjected to systematic tests. Kossobokov and Nekrasova (2012) compared the shaking predicted by the GSHAP map (Giardini et al., 1999) with the shaking observed during strong earthquakes that occurred in 2002-2009. Wyss et al. (2012) have performed a similar comparison for the number of casualties. The authors of these studies concluded that both tested quantities have been severely underestimated by the GSHAP, and that the methods applied by this program therefore require reassessment and modifications.

Chapter 2

Effect of alternative distributions of ground motion variability on results of probabilistic seismic hazard analysis

(Published in Natural Hazards, September 2015)

2.1 Abstract

PSHA is a regularly applied practice that precedes the construction of important engineering structures. The Cornell-McGuire procedure is the most frequently applied method of PSHA. This paper examines the fundamental assumption of the Cornell-McGuire procedure for PSHA, namely, the log-normal distribution of the residuals of the ground motion parameters. Although the assumption of log-normality is standard, it has not been rigorously tested. Moreover, the application of the unbounded log-normal distribution for the calculation of the hazard curves results in non-zero probabilities of the exceedance of physically unrealistic amplitudes of ground motion parameters. In this study, the distribution of the residuals of the logarithm of PGA is investigated, using the database of the Strong-motion Seismograph Networks of Japan and the GMPE of Zhao and co-authors. The distribution of residuals is modelled by a number of probability distributions, and the one parametric law that approximates the distribution most precisely is chosen by the statistical criteria. The results of the analysis show that the most accurate approximation is achieved with the GEVD for a central part of a distribution and the GPD for its upper tail. The effect of replacing a log-normal distribution in the main equation of the Cornell-McGuire method is demonstrated by the calculation of hazard curves for a simple hypothetical example. These hazard curves differ significantly from one another, especially at low annual exceedance probabilities.

2.2 Introduction

PSHA is a complicated and crucial problem of modern seismology as it is related to the effects of strong earthquakes and their consequences for the inhabitants. PSHA is applied to estimate the possible amplitudes of destructive seismic ground motion and to provide the design loads for the construction of critical structures such as dams and power plants. The main goal of such analysis is to minimise the negative effect of future strong earthquakes. Although there are several methods of PSHA (Cornell, 1968; Shumilina et al., 2000; Kijko, 2008), the most frequently used method is the Cornell-McGuire procedure (Cornell, 1968). The theoretical foundations, formulated by C.A. Cornell and L.Esteva (McGuire, 2008), were supplemented by the computer programs developed

by R.K. McGuire (1976; 1978), which led to a method of PSHA known as the Cornell-McGuire procedure.

Ground motion variability is an important component of this method (Bender, 1984; Bommer and Abrahamson, 2006). This component was introduced in the Cornell-McGuire procedure to account for the effect of the scatter of the amplitude of seismic ground motion at a site (Cornell, 1971) and is included in the main equation of this procedure. The common assumption is that the ground motion variability can be modelled by a random variable with a log-normal distribution (Joyner and Boore, 1981). This implies that the residuals of ground motion parameters are log-normally distributed about the predicted value or, equivalently, that the residuals of the logarithms of these parameters are normally distributed. However, this hypothesis has not been reliably tested. Moreover, the assumption of log-normally distributed residuals has become a standard, and as a result usually is not tested but is accepted as a given.

The evidence for a log-normal distribution was confirmed by the KS test at the 90 per cent confidence limit (Campbell, 1981). Nevertheless, although the hypothesis was not rejected by the KS test, it does not imply that the hypothesis is true. The KS test does perform well in a central part of a distribution, however, it is widely known that the test demonstrates poor sensitivity to deviations from the hypothesised distribution that occur in the tails.

The log-normal assumption is criticised in (Raschke, 2013), where author notes that the natural distribution for residuals of maxima, such as PGA, is the GEVD. The theory of extreme values is widely applied in the analysis of natural disasters in general and in the analysis of seismic hazard in particular. Pisarenko and Rodkin (2010) provides the results of the application of the extreme value theory for various aspects of the analysis of natural disasters.

In general, a PSHA is applied to estimate ground motions with an APE down to 10^{-4}, a typical annual exceedance probability value designated for nuclear power plant design. However, in a PSHA performed for the Yucca Mountain nuclear waste repository, probabilistic hazard curves were extrapolated to an annual exceedance probability of 10^{-8}. The peak characteristics of ground motion corresponding to an annual exceedance probability of 10^{-7} were as high as 20 g for PGA and up to 1800 cm/s for PGV (Corradini, 2003; Stamatakos, 2004). This instance revealed a controversy in a fundamental assumption of the modern Cornell-McGuire method. As pointed out, for example, in Abrahamson (2000), at these low annual probabilities, the hazard estimates are controlled by the tail of the distribution of the ground motion residuals. Since log-normal distribution is unbounded, extrapolation of a hazard curve leads to the unlimited increase of the amplitudes of expected ground motions, with the decreasing of the APE.

On the other hand, some recent studies of the results of the GSHAP (Giardini et al., 1999) revealed discrepancies between the observed seismicity and that predicted by the resulting maps of this program (e.g. Kossobokov and Nekrasova, 2011; 2012; Wyss et al., 2012). The authors of these studies concluded that the common methods of PSHA are inadequate and need to be revised and probably modified.

One probable source of the revealed inadequacy is the assumption of the log-normal distribution of residuals of the ground motion parameter (e.g. PGA). An upper tail of the distribution of the ground motion residuals controls the behaviour of hazard curves at long return periods. There-

fore, an accurate modelling of this distribution, especially at an upper tail region, is a significant problem.

The current study is methodological in nature and its main purpose is to introduce a suitable method of studying the ground motion variability. In this study, an analysis of the distribution of the residuals of the logarithm of PGA is performed in order to select a parametric law that describes this distribution most accurately. Data obtained from the Japanese Strong-motion Seismograph Networks were used in the study. The Japanese database was chosen mainly because of a dense net of strong-motion stations that allows obtaining enough observations. A GMPE of Zhao et al. (2006) was used for the calculation of the forecast values of PGA. Statistical criteria show that the best approximation for the distribution of residuals of the logarithm of PGA is achieved with the GEVD. The GPD is used to capture the behaviour of an upper tail more accurately.

2.3 Methods

The method for studying the distribution of residuals is based on the sequential application of the KS test and the AIC, and a quantile-quantile plot. The distribution of the residuals of the natural logarithm of PGA is modelled by a number of parametric distributions. The residual is defined as

$$\varepsilon = \ln(\mathrm{PGA}_{\mathrm{observed}}) - \ln(\mathrm{PGA}_{\mathrm{predicted}}) \tag{2.1}$$

where $\mathrm{PGA}_{\mathrm{observed}}$ is the observed value and $\mathrm{PGA}_{\mathrm{predicted}}$ is the value calculated by using an appropriate GMPE.

The typical GMPEs allow the calculation of median values of the ground motion parameters by using their dependence on the magnitude, source to site distance, local soil conditions at a site, source mechanism, and others. Such equations often have an empirical nature and are developed based on vast databases of observed values of ground motion parameters (Boore and Joyner, 1982). The selection of the most appropriate GMPE is not a trivial task and some guidance and criteria for choosing the most appropriate GMPE for the application in a PSHA for a particular site can be found in Scherbaum et al. (2009) and Arroyo et al. (2014). A comprehensive list of GMPEs developed during the period 1964-2010 is presented in Douglas (2011).

In this study, data recorded by the Japanese Strong-motion Seismograph Networks were used. The GMPE of Zhao et al. (2006) was used for the calculation of the forecast values of PGA. This GMPE was developed for the calculation of the ground motion parameters of subduction zone earthquakes, it allows calculating a geometrical mean of the horizontal components of PGA, or 5% damped acceleration response spectrum.

In this study, the logistic distribution, the Student's t-distribution and the GEVD were considered as alternatives to the normal distribution. Following a standard notation, where μ is a location parameter and σ is a scale parameter, PDFs of these distributions can be written as follows:
The PDF of Student's t-distribution is defined as

$$f_{\mu,\sigma,n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}\sigma}\left(1+\frac{1}{n}\frac{(x-\mu)^2}{\sigma^2}\right)^{-\frac{n+1}{2}}, \quad x \in (-\infty;+\infty)$$

where Γ is an Euler's gamma function, n is a number of degrees of freedom.

The PDF of logistic distribution is defined as

$$f_{\mu,\sigma}(x) = \frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\sigma\left(1+\exp\left(-\frac{x-\mu}{\sigma}\right)\right)^2}, \quad x \in (-\infty;+\infty)$$

The PDF of the GEVD is defined as

$$f_{\mu,\sigma,\xi}(x) = \frac{1}{\sigma}\begin{cases} \left(1+\xi\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}-1}\exp\left\{-\left(1+\xi\frac{(x-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right\}, & 1+\xi\frac{(x-\mu)}{\sigma} \geq 0, \xi \neq 0 \\ \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)-\frac{x-\mu}{\sigma}\right\}, & x \in (-\infty;+\infty), \xi = 0 \end{cases}$$

where ξ is a shape parameter.

Statistical analysis was performed in the following order:

- Estimation of the distribution parameters by the ML method.
- Testing the hypothesis that a sample belongs to the current distribution by the KS test at 0.05 significance level.
- Calculation of the AIC for hypotheses that were accepted by the KS test.

The application of the KS test (Massey, 1951) for one sample allows rejecting the distributions that do not fit the empirical data. The test statistic of this test with the Bol'shev's amendment (Bol'shev and Smirnov, 1965) is calculated by using the formula

$$S_k = \frac{6nD_n + 1}{6\sqrt{n}} \tag{2.2}$$

where $D_n = \max(D_n^+, D_n^-)$, $D_n^+ = \max\left\{\frac{i}{n} - F(x_i,\theta)\right\}$, $D_n^- = \max\left\{F(x_i,\theta) - \frac{i-1}{n}\right\}$; n is a sample size, $x_1,\ldots,x_n$ - elements of a sample, sorted in ascending order, $F(x,\theta)$ is a CDF of a parametric model that undergoes the test.

An attractive feature of this test is that the distribution of its test statistic itself does not depend on the underlying CDF being tested. However, in composite hypotheses testing, when the parameters of the probability distribution are estimated on the analysed sample, the KS test loses this feature. In such instances the conditional distribution of a test statistic depends on a number of factors (such as form of $F(x,\theta)$, number of estimated parameters, method of parameter estimation etc). Lemeshko and Lemeshko (2009) presents the updated results (tables of percentage points and models of the distributions of statistics) for nonparametric goodness of fit tests in testing composite hypotheses in case of using ML estimations. The KS test rejects hypotheses for which the maximum deviation of the theoretical CDF from the empirical CDF exceeds a critical value at a given significance level.

However, the KS test alone does not allow unambiguous conclusion about which parametric model approximates the empirical distribution most accurately. Such a conclusion can be made based on the calculation of the AIC (Akaike, 1974) for hypotheses that were accepted by the KS

test. The criterion is defined as

$$AIC = -2\ln(L) + 2k \tag{2.3}$$

where L is a maximised likelihood function, k is a number of parameters of the probability distribution model. The parametric distribution for which the value of criterion is minimal is considered the best approximation among the considered alternatives for the empirical distribution.

The quantile-quantile plot allows comparing the quantiles of empirical and theoretical distributions. The conception of such a plot has emerged from the observation that for important classes of distributions, the quantiles are linearly related to the corresponding quantiles of a standard example from this class (Beirlant et al., 2004). Linearity in a graph can be easily checked by the eye and can further be quantified by means of a correlation coefficient.

2.4 Results and discussion

The results of the statistical analysis show that the best approximation of the distribution of residuals is achieved with the GEVD. It is important to note that a similar conclusion was reached in Dupuis and Flemming (2006) from theoretical considerations. In Dupuis and Flemming (2006) the regression analysis was performed using both the GEVD and the normal distribution as a model for the distribution of residuals, it was demonstrated that a better fit to the data and in turn more accurate acceleration estimates are obtained with the use of the GEVD. A similar conclusion in regards of the distribution of the ground motion residuals was also reached in Raschke (2013).

Figure 2.1 demonstrates the histogram of residuals together with the fitted PDFs. Corresponding values of the AIC are presented in Table 2.1.
The CDF of the GEVD is defined as

$$H_{\xi,\mu,\sigma}(x) = \begin{cases} \exp\left\{-\left(1+\xi\frac{(x-\mu)}{\sigma}\right)^{-1/\xi}\right\}, & \xi \neq 0 \\ \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, & \xi = 0 \end{cases}$$

This is generalised form, also known as the Jenkinson-von Mises representation, which combines three types of extreme value distributions. When $\xi = 0$ it is equivalent to the Gumbel distribution (type I), when $\xi > 0$ it is equivalent to the Fréchet distribution (type II) and when $\xi < 0$ it is equivalent to the Weibull distribution (type III) (Embrechts et al., 1997).

Model	AIC
GEVD	531.098
Normal	532.056
Student's t	534.054
Logistic	537.948

Table 2.1: Values of the AIC for considered distributions

Figure 2.2 demonstrates the quantile-quantile plot for the quantiles of the GEVD. It can be seen from the plot, that an upper tail of the distribution of the residuals slightly deviates from the

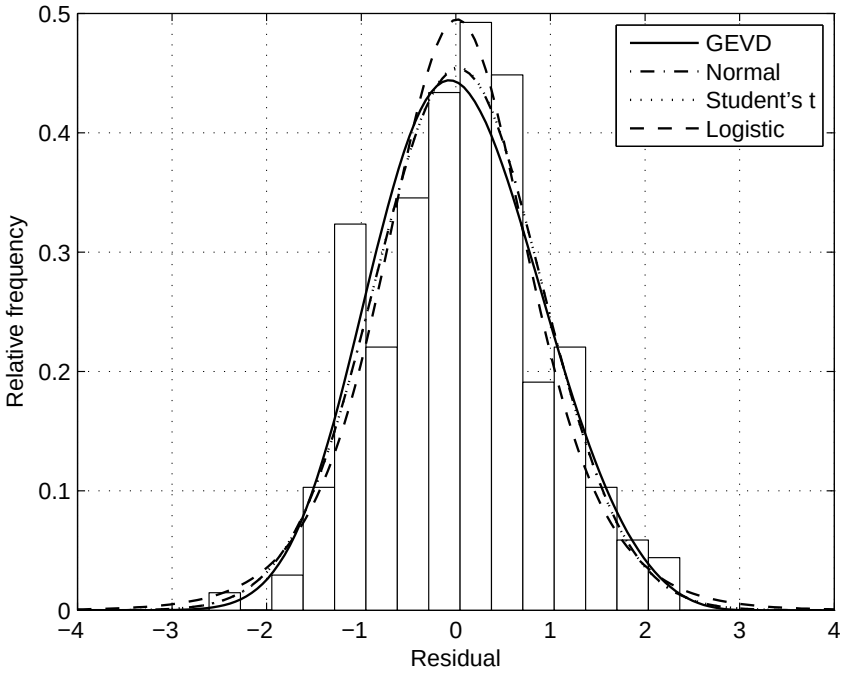


Figure 2.1: Histogram of residuals and the fitted PDFs

GEVD. Accurate modelling of that tail is an important problem because it defines the hazard curve at long return periods.

Therefore, the "peaks over threshold" method was applied to fit an upper tail of the distribution of residuals more precisely. This method is based on fitting the GPD to values that exceed a reasonably large threshold (Embrechts et al., 1997). The GPD arises as a limiting distribution of the excesses for a sufficiently large threshold value and is often used for modelling the tails of empirical distributions. The CDF of the GPD is defined by the following function

$$G_{\xi,\nu,\beta}(x) = \begin{cases} 1 - \left(1 + \xi \frac{(x-\nu)}{\beta}\right)^{-1/\xi}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{x-\nu}{\beta}\right), & \xi = 0 \end{cases}$$

Similar to the GEVD, the GPD is also characterised by three parameters, namely location ν, scale β and the shape parameter ξ. When the GPD is used as a model for a tail of some other distribution its parameter ν defines the threshold from which a tail region of that distribution begins. When $\xi = 0$ the GPD is equivalent to the exponential distribution, when $\xi > 0$ the GPD has a heavy tail, when $\xi < 0$ the GPD has a finite upper bound defined as $x_F = \nu - \frac{\beta}{\xi}$. Three possible types of tail of the PDF of the GPD, with different values of the shape parameter are shown in Fig. 2.3.

There are several methods for the estimation of the shape parameter. Well-known methods, such as the Hill estimator (1975) and the Pickands estimator (1975) are both based on the asymptotic properties and require a significant number of observations. Applicability of these methods to the real observations is doubtful (Pictet et al., 1998). In this study, therefore, a shape parameter was estimated by using the ML method.

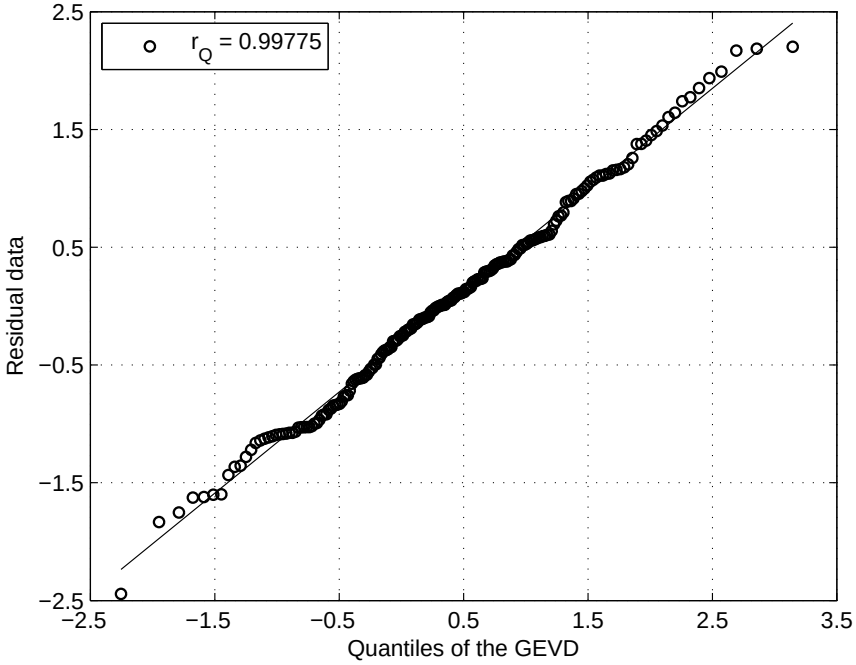


Figure 2.2: Quantile-quantile plot of sample data vs quantiles of the GEVD

The robust estimation of the shape parameter $\hat{\xi}$ requires an optimal choice of the threshold value $\hat{v}$. If too high a value of $\hat{v}$ is chosen, too few exceedances and, consequently, high variance estimators will be the result. When $\hat{v}$ is too small, the estimators become biased. The procedure for the optimal determination of the threshold value is proposed in Embrechts et al. (1997). This procedure utilises the linearity of the mean excess function for the GPD, which is defined as

$$e(\hat{v}) = E(X - \hat{v} | X > \hat{v}) = \frac{\hat{\beta} + \hat{\xi}\hat{v}}{1 - \hat{\xi}} \tag{2.4}$$

This procedure suggests selecting the threshold value $\hat{v}$ as a starting point of a linear segment of the mean excess graph. Such a graph for sample data is presented in Fig. 2.4.

By varying the threshold value and observing changes in the estimates of the rest of the parameters of the GPD, the optimal threshold value can be determined. Evidence of such a choice is the stabilization of the estimates of the scale and shape parameters. Once again, a quantile-quantile plot is used as a tool for comparing the data and the model. An estimator of the quantile of the GPD can be written as

$$\hat{x}_p = \hat{v} + \frac{\hat{\beta}}{\hat{\xi}} \left[\left(\frac{n}{N_{\hat{v}}} (1 - p) \right)^{-\hat{\xi}} - 1 \right] \tag{2.5}$$

where n and $N_{\hat{v}}$ are sample size and number of exceedances, respectively, and $\hat{v}$, $\hat{\beta}$ and $\hat{\xi}$ are estimates of the GPD parameters.

Quantile-quantile plots of a tail of residual data versus quantiles of the GPD, with the values of the shape parameter estimated by the ML method for the GEVD and the GPD are shown in Figs. 2.5 and 2.6. The estimates of a shape parameter differ for the instances where the GEVD is

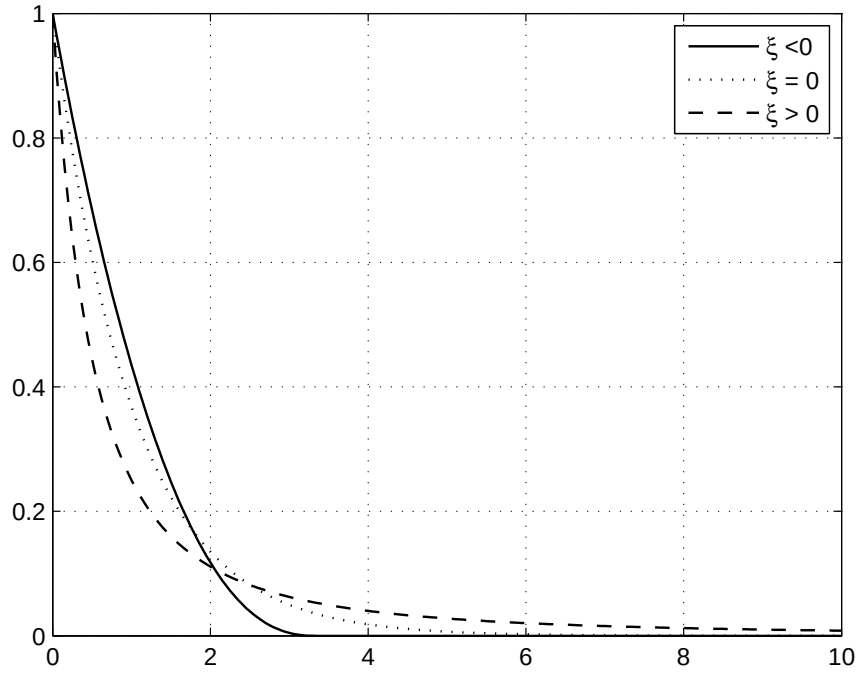


Figure 2.3: Tails of PDF of the GPD with different values of parameter ξ

used as a model for a full range of residuals and where the GPD is used additionally to fit an upper tail more accurately. The estimates are $\hat{\xi} = -0.245$ for the first instance and $\hat{\xi} = -0.359$ for the second.

Therefore, the distribution of residuals is represented by a hybrid distribution model that consists of the GEVD in a central region and the GPD in a region of an upper tail.

A similar analysis was performed during this study by using the GMPEs of Atkinson and Boore (2003) and Kanno et al. (2006) to check how generally applicable these results are. The results obtained with these GMPEs are very close to those presented in this study.

2.5 Implication for Probabilistic Seismic Hazard Analysis

For a demonstration of the effect of replacing a normal distribution, hazard curves were calculated in the following manner:

1. Using an unbounded normal distribution.
2. Using a normal distribution, truncated at a specified level of ground motion.
3. Using the GEVD.
4. Using a hybrid distribution model that consists of the GEVD for a central region and the GPD for an upper tail.

A brief review of the Cornell-McGuire PSHA procedure could be helpful for understanding the following material.

To begin with, recap of a PDF of log-normal distribution could be useful. If the distribution of a random variable is log-normal, its PDF has the following form

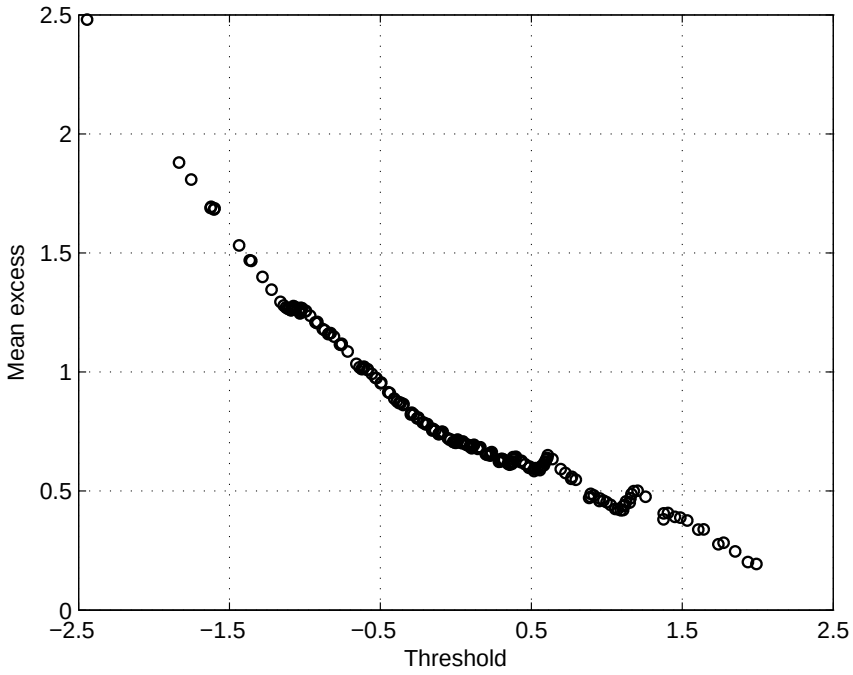


Figure 2.4: Mean excess graph of sample data

$$f_{\mu,\sigma}(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, \quad x > 0$$

The transformation $Y = \ln X$ leads to a normally distributed random variable with a location μ and a scale σ parameters. The values of these parameters are estimated by using an appropriate GMPE. Given an earthquake with magnitude m, the probability can be calculated that ground motion at distance r from the source will exceed a particular level a_0 by the following equation

$$P(y \geq \ln(a_0)|m,r) = \frac{1}{\sqrt{2\pi}\sigma} \int_{a_0}^{\infty} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \tag{2.6}$$

This equation can be conveniently expressed in terms of the standard normal distribution

$$P(y \geq \ln(a_0)|m,r) = 1 - \Phi(z) \tag{2.7}$$

where $z = \frac{\ln(a_0)-\mu}{\sigma}$ is a standardised normal random variable and $\Phi(z)$ is the standard normal CDF.

Next, consider a site surrounded by N seismic sources. Each seismic source is characterised by magnitude M_i, distance to site R_i and annual activity rate ν_i. The parameters of future seismic events are yet unknown, therefore M_i and R_i are random variables with corresponding PDFs $f_{M_i}(m)$ and $f_{R_i}(r)$. The total annual rate of exceedance of a particular level of ground motion a_0 can be calculated as follows

$$\lambda(y \geq \ln(a_0)) = \sum_{i=1}^{N} \nu_i \iint P(y \geq \ln(a_0)|m,r) f_{M_i}(m) f_{R_i}(r)\, dr dm \tag{2.8}$$

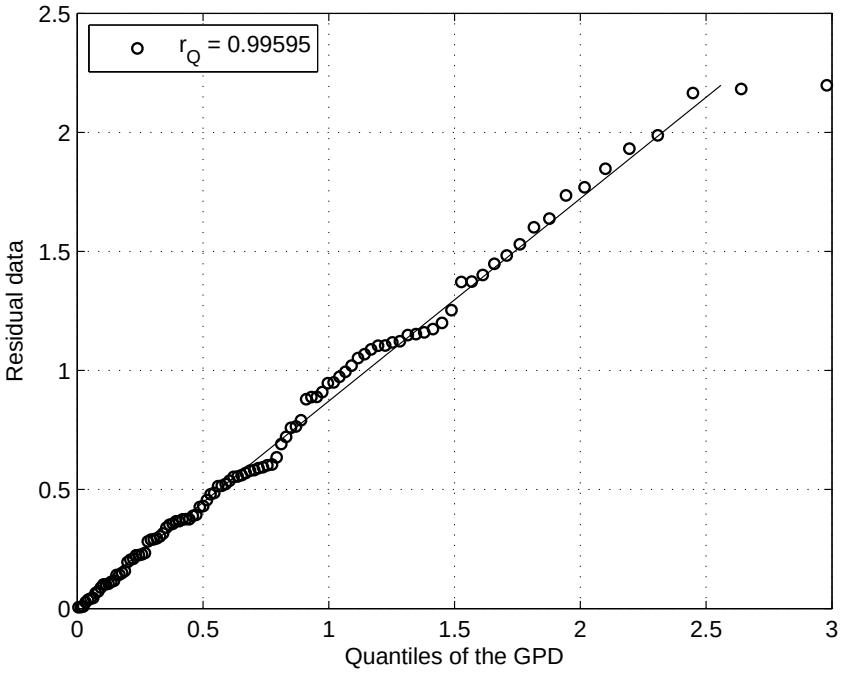


Figure 2.5: Quantile-quantile plot of the tail fraction of residuals, ξ is estimated for the GEVD

From an assumption that the sequence of major seismic events can be modelled by the Poisson distribution it follows that the probability for a particular level of ground motion a_0 to be exceeded at least once during the time interval T can be calculated (Anderson and Brune, 1999) as follows

$$P(y \geq \ln(a_0), T) = 1 - \exp(-\lambda(y \geq \ln(a_0)) \times T) \tag{2.9}$$

The equation 2.9 with $T = 1$ year defines the seismic hazard curve, the main result of the PSHA. For small values of the annual rate of exceedance ($\lambda(y \geq \ln(a_0)) \ll 1$), equation 2.9 can be approximated as

$$P(y \geq \ln(a_0), T = 1) = 1 - \exp(-\lambda(y \geq \ln(a_0))) \cong \lambda(y \geq \ln(a_0)) \tag{2.10}$$

As emphasised in Wang (2011), $T = 1$ year is neglected on the right side of 2.10, thus both sides of this equation contain a dimensionless quantity, i.e. the APE.

As can be seen from equation 2.8, the ground motion variability is explicitly incorporated in the calculation of the seismic hazard. It is, namely used in a calculation of the conditional exceedance probability of a ground motion of a particular level a_0. The normal distribution is unbounded, therefore the further a hazard curve extrapolated, the higher the level of ground motion is expected to be exceeded.

The necessity of an upper bound of the ground motion, as well as the difficulties related to its determination are summarised in Bommer et al. (2004). Strasser et al. (2004) proposed the truncation of the distribution of residuals at a level of three standard deviations above the median as a measure to prevent the effect of unbounded normal distribution. Given a normal distribution,

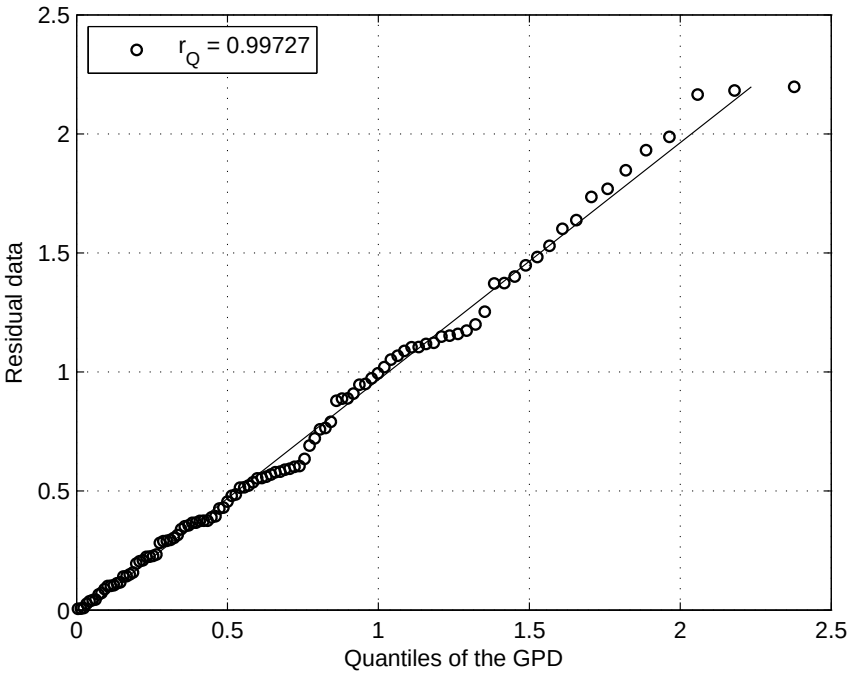


Figure 2.6: Quantile-quantile plot of the tail fraction of residuals, ξ is estimated for the GPD

truncated at a value a_T, the PDF has to be renormalised to satisfy the fundamental properties of PDF. Then, the probability that an earthquake with magnitude m will produce ground motion at distance r from the source that exceeds a particular level a_0 can be expressed as

$$P(y \geq \ln(a_0)|m,r) = \begin{cases} 1 - \frac{\Phi(z)}{\Phi(z_T)}, & y \leq a_T \\ 0, & y > a_T \end{cases} \tag{2.11}$$

where $z_T = \frac{\ln(a_T)-\mu}{\sigma}$.

After the replacement of a normal distribution by the GEVD, the same probability can be expressed as

$$P(y \geq \ln(a_0)|m,r) = 1 - H_\xi(z) \tag{2.12}$$

where H_ξ is a standardised CDF of the GEVD.

And after the replacement of a normal distribution by a hybrid model, this probability can be written as

$$P(y \geq \ln(a_0)|m,r) = \begin{cases} 1 - (1-p)\frac{H_\xi(z)}{H_\xi(z_\nu)}, & y \leq \mu + \nu \\ p(1 - G_\xi(z)), & y > \mu + \nu \end{cases} \tag{2.13}$$

where $z_\nu = \frac{\ln(a_\nu)-\mu}{\sigma}$, $a_\nu = \exp(\mu + \nu)$, $G_\xi(z)$ is a standardised CDF of the GPD, p is a fraction of the residual values that fall in a tail region.

For the purpose of demonstration, a simple hypothetical example was considered. This example is similar to an example used in Baker (2008) and assumes there are two seismic sources

that may affect the site. Both sources are subduction slab sources. The first source is capable of producing an earthquake of magnitude $m_1 = 5.5$ every 100 years ($\nu_1 = 0.01$) and is located at a depth of $d_1 = 30$ km and a distance of $r_1 = 140$ km from the site. The second source is capable of producing an earthquake of magnitude $m_2 = 6.5$ every 500 years ($\nu_2 = 0.002$) and is located at a depth of $d_2 = 30$ km and a distance of $r_2 = 200$ km from the site. The soil conditions at a site are characterised as medium soil ($V_{S30} = 250$ $\frac{m}{s}$). For the given combinations of parameters, a GMPE of Zhao et al. (2006) gives ln(PGA) values of $\mu_1 = 1.8404$ $\frac{cm}{s^2}$, $\mu_2 = 2.0233$ $\frac{cm}{s^2}$ and a standard deviation $\sigma = 0.6840$, which is a constant in this GMPE for seismic events generated by sources of identical type. With the defined earthquake scenarios, equation 2.8 simplifies to the following

$$\lambda(y \geq \ln(a_0)) = \nu_1 \times P(y \geq \ln(a_0)|m_1, r_1) + \nu_2 \times P(y \geq \ln(a_0)|m_2, r_2) \tag{2.14}$$

By repeating these calculations for a range of values of PGA, a total hazard curve can be constructed. Hazard curves calculated by using the above-mentioned distributions are represented in Fig. 2.7. As can be seen, the hazard curve calculated by using the GEVD displays the highest ground motion estimates, almost down to an annual exceedance probability of 10^{-6} where it crosses with the hazard curve calculated by using an unbounded normal distribution.

The hazard curve calculated by using a hybrid distribution model is very close to the curve calculated by using a truncated normal distribution, down to an annual exceedance probability of 10^{-5}, after which it estimates higher ground motions and the difference gradually increases.

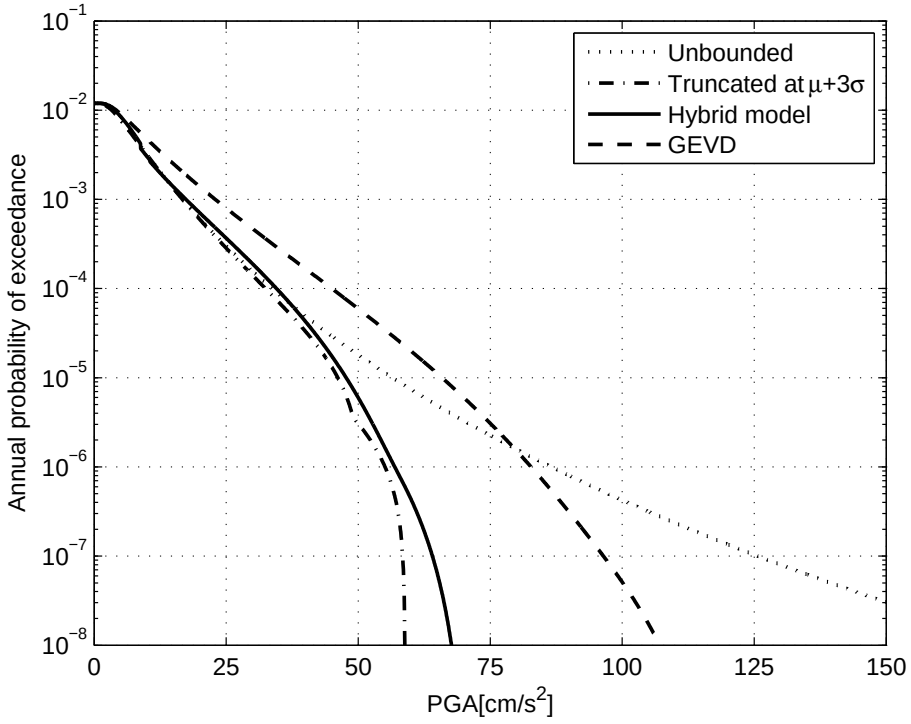


Figure 2.7: Hazard curves calculated by using different parametric distributions

As can be seen from Fig. 2.7, the hazard curves calculated by using the GEVD and a hybrid distribution model depend strongly on the shape parameter ξ. Although the method applied for statistical analysis in this study is satisfactory, the estimations of the shape parameter can only be called preliminary. These estimations were obtained based on the strong ground motion records from a particular region and, therefore, are valid only for this particular region, which, in this instance is Japan. Such analysis should be performed for a multiple number of datasets of the recordings of seismic ground motions that were induced by earthquakes of various types and magnitudes, and were recorded worldwide, for a possible generalisation of these results.

2.6 Conclusion

In this study, the distribution of the residuals of $\ln(\text{PGA})$ was modelled by a number of probability distribution laws, using the database of the Strong-motion Seismograph Networks of Japan and a GMPE of Zhao et al. (2006). The results of the analysis indicate that the best approximation for the distribution of residuals was obtained with the GEVD. This result is consistent with the conclusions of Dupuis and Flemming (2006) and Raschke (2013). The ”peaks over threshold” method was applied in an attempt to model an upper tail of the distribution of residuals more precisely. Thus, the resulting distribution of residuals is a hybrid model that consists of the GEVD in a central region and the GPD in a region of an upper tail. Similar analysis was performed during this study by using GMPEs of Atkinson and Boore (2003) and Kanno et al. (2006), which demonstrated analogous regularities.

The estimations of the shape parameter of the GEVD and the GPD resulted in negative values, indicating that the distribution of residuals has a finite upper bound. Consequently, a maximum value of PGA can be associated with an earthquake scenario involved in the PSHA. This approach is preferred to the truncation procedures proposed in Strasser et al. (2004), because a maximum value of PGA, unlike the truncation of a distribution, has a clear physical meaning.

Hazard curves were calculated for a simple hypothetical example to demonstrate the effect of the replacement of the normal distribution. Hazard curves were calculated by using the GEVD and a hybrid distribution model, which differ from each other and from the curves calculated by using the normal distribution. This difference is particularly evident at low annual exceedance probabilities.

Acknowledgements I would like to express my appreciation to Professor Andrzej Kijko for formulating the problem, guidance during this study, the review of this paper and the valuable comments that helped to significantly improve its quality. I would like to thank two anonymous reviewers for their valuable comments and suggestions that allowed to make this paper even better. I am grateful to the K-NET and KiK-net Strong-Motion Seismograph Networks of Japan for seismic ground motion records (available online at http://www.kyoshin.bosai.go.jp).

Chapter 3

Estimation of the upper bound of seismic hazard curve by using the generalized extreme value distribution

3.1 Abstract

The problem considered in this study is that of unrealistic ground motion estimates, which arise in the Cornell-McGuire method when the seismic hazard curve is calculated for extremely low APEs. This problem stems from using the normal distribution in the modelling of the variability of the logarithm of ground motion parameters. In this study, the database of the strong-motion seismograph networks of Japan was used to examine the distribution of the logarithm of PGA. The normal distribution and the GEVD models were considered in the analysis, with the preferred model being selected based on the statistical criteria. The results of the analysis demonstrated the superiority of the GEVD in the vast majority of considered examples. The estimates of the shape parameter of the GEVD were negative in every considered example, indicating the presence of a finite upper bound of PGA. Therefore, the GEVD provides a model that is more realistic for the scatter of the logarithm of PGA, and the application of this model leads to a bounded seismic hazard curve.

3.2 Introduction

PSHA is an important field of modern seismology, related to the effects of strong earthquakes and their consequences. The main purpose of PSHA is to estimate the design ground motion that can be utilised in earthquake structural engineering to produce a structure that can withstand a certain level of shaking without severe damage, and thus reduce the negative effects of strong earthquakes, i.e. casualties and damage to infrastructure. Several PSHA methods exist (Cornell, 1968; Milne and Davenport, 1969; Molchan et al., 1970; Veneziano et al., 1984; Frankel, 1995; Kijko and Graham, 1998; 1999; Ebel and Kafka, 1999; Shumilina et al., 2000); however, the most widely applied method is the Cornell-McGuire procedure (Cornell, 1968; 1971; McGuire, 1976; 1978). This method was formulated by C.A.Cornell and L.Esteva (Bommer and Abrahamson, 2006; McGuire, 2008) and was supplemented with computer programs developed by R.K.McGuire.

Although extensive further development has taken place since the Cornell-McGuire method was originally published, some controversial aspects remain in its mathematical formulation. One of these is the limitlessness of the seismic hazard curves at very low APE. Pertinent examples of such a flaw in modern PSHA include seismic hazard studies that were performed for the Yucca Mountain nuclear waste repository in the USA (Stepp et al., 2001) and the PEGASOS project in Switzerland (Abrahamson et al., 2002). Both studies considered the extremely low APE (down to 10^{-8} and 10^{-7}, respectively) and provided unrealistically high values of ground motion parameters (Stepp et al., 2001; Corradini, 2003; Stamatakos, 2004; Klügel, 2005).

For very low APE, the hazard estimates are controlled by the tail of the distribution of the ground motion residuals (Anderson and Brune, 1999; Abrahamson, 2000; Wang, 2011). In current practice, the distribution of the residuals of the ground motion parameters is assumed to correspond to a log-normal distribution and, therefore, the distribution of logarithmic residuals is modelled by a normal (Gaussian) distribution. Since normal distribution is unlimited, the estimations of the ground motion parameters at very low APE are unlimited as well.

The determination of the possible upper bounds of seismic ground motions has been discussed in engineering seismology for a long time. The history of the development of this subject is discussed in Bommer et al. (2004) and Strasser and Bommer (2009). The values that were suggested as possible limits of ground motion parameters, such as PGA and PGV, have been gradually increasing because of the accumulation of strong motion records and the consequent increase of the observed maximum values of these parameters.

The Bayesian procedure for estimating the maximum value of a parameter of ground motion that would occur at a given site within the specified time interval was proposed by Pisarenko and Lyubushin (1997; 1999) and was applied by Lyubushin et al. (2002). This procedure utilises the catalogue of seismic events and the GMPE; it is applicable for any ground motion parameter that may be estimated by using the GMPE (e.g., PGA, components of acceleration response spectra, seismic intensity).

Kijko and Graham (1999) have proposed two approaches for estimating the maximum value of PGA at a site of interest. The first approach suggests the straightforward estimation of PGA by using the appropriate GMPE, under assumption that the strongest possible earthquake (i.e., an earthquake with magnitude $\hat{M}_{\max}$, where $\hat{M}_{\max}$ is estimated by using one of the methods such as developed by Pisarenko et al., 1996; Kijko, 2004; Kijko and Singh, 2011) occurred very close to the site (e.g., at a distance of 10 km). The second approach utilises the distribution of logarithm of PGA at a site, derived by Kijko and Graham (1999). The derived distribution of logarithm of PGA at a site is of the same type as the distribution of earthquake magnitude, obtained under the assumption of applicability of the GR (Gutenberg and Richter, 1944) recurrence law. The similarity of two distributions allows estimation of the maximum possible PGA at a site by modifying the techniques developed for assessing the upper limit of earthquake magnitude.

Some studies (e.g., Strasser et al., 2008b) have sought a solution to the problem by truncation of the distribution of ground motion residuals. However, such a procedure has no clear physical meaning; moreover, to a certain extent it is arbitrary.

An example of such a study is that of Romeo and Prestininzi (2000), who proposed a truncation at two standard deviations above the median of the distribution. On the other hand, Strasser et al. (2004) indicated that the truncation should be performed at least at a level three standard deviations above the median. Conversely, McGuire (1976) suggested that the distribution of residuals should be truncated at a level six standard deviations above the median value, or truncation should be performed in such a way that the ground motion amplitude at a site could not be greater than that value at the epicenter (McGuire, 1977).

Bommer and Abrahamson (2006) have indicated that truncation at a level above three standard deviations has little effect on the hazard curves in the range of the return periods that are generally used in engineering design. If the seismic activity of the region is not very high, this effect would remain small, even for a return period of 10^4 years.

An interesting discussion on the influence of truncating the distribution of ground motion residuals on the results of PSHA can be found in Wu et al. (2011). These authors (Wu et al., 2011) concluded that the number of standard deviations depends on the range of APE under consideration and it should gradually increase with the decrease of the APE.

Hypothetical hazard curves, calculated by using the unbounded normal distribution, as well as the normal distribution, truncated at levels of three, four, and five standard deviations above the median value, are presented in Fig. 3.1

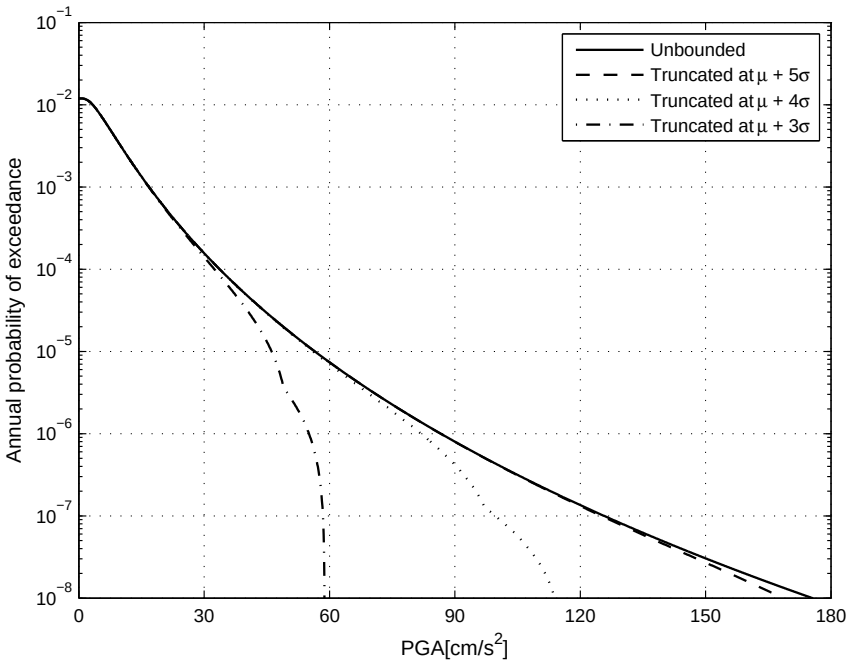


Figure 3.1: Effect of truncation of the distribution of residuals on the hazard estimates

However, truncation of the distribution of the ground motion residuals potentially leads to an exclusion of certain strong motions that have low probability of occurrence from the scope of analysis. Moreover, it is exactly those particular strong motions that define the behavior of the seismic hazard curve at a long return period. Therefore, from this perspective, the truncation of ground motion variability seems doubtful.

On the other hand, some studies have focused on investigating the distribution of ground motion variability. Accordingly, Lavallée and Archuleta (2005) have investigated the distribution of the absolute values of PGA obtained from the strong motion records of the 1999 Chi-Chi earthquake and have suggested to model it by the Lévy distribution. Dupuis and Flemming (2006) have indicated that the distribution of residuals of PGA should theoretically correspond to the GEVD. Dupuis and Flemming (2006) have performed regression analysis under assumption of the log-normal distribution of residuals, as well as under assumption of the residuals being distributed according to the GEVD. The results demonstrated that a superior fit to the data and, in turn, more accurate acceleration estimates are obtained under the latter assumption. Similar considerations were expressed by Raschke (2013), who criticised the log-normal assumption and noted that the GEVD is the natural distribution for residuals of maxima such as PGA.

Huyse et al. (2010) applied the peaks over threshold method to analyse both the raw PGA data and the logarithmic residuals of PGA, and concluded that the GPD, with the negative shape parameter, provided a model that was more accurate for the tail fractions of both the studied datasets. Similar results were obtained by Pavlenko (2015), who found that the GEVD was a more appropriate model for logarithmic residuals of PGA.

In the present study, the analysis of the statistical properties of $\ln(\text{PGA})$ is continued. The analysis is based on the data of the strong-motion seismograph networks of Japan (K-NET, KiK-net). Statistical criteria are used to compare the performance of the normal distribution and the GEVD models. The results indicate the superior performance of the GEVD in the vast majority of considered examples.

3.3 The Cornell-McGuire procedure

In PSHA, the seismic hazard is characterised by the probability $P(y \geq a_0, T) = P(a_0, T)$ that the ground motion parameter y will exceed the value a_0 at a given site at least once during a specified period of time T. It is usually assumed in PSHA studies that the sequence of major seismic events can be modelled by the Poisson process (e.g. Anderson and Brune, 1999). This assumption allows the calculation of $P(a_0, T)$:

$$P(a_0, T) = 1 - e^{-\lambda(a_0)T} \tag{3.1}$$

where $\lambda(a_0) = \lambda(y \geq a_0)$ is the mean annual rate of exceedance of ground motions level a_0 at the site.

For $T = 1$ year and for $(\lambda(a_0) \ll 1)$, eq. (3.1) can be approximated as:

$$P(a_0, T = 1) = 1 - e^{-\lambda(a_0)} \cong \lambda(a_0) \tag{3.2}$$

Equation (3.2) is an approximation of the APE and $T = 1$ year is neglected on the right hand side of (3.2); therefore, both sides of this equation contain a dimensionless quantity (Wang, 2011). For a single seismic source, $\lambda(a_0)$ can be calculated as:

$$\lambda(a_0) = \nu P(y \geq a_0) \tag{3.3}$$

where ν is the annual rate of occurrence of earthquakes with magnitude greater than or equal to m_0, which is the lower threshold of magnitude of earthquakes capable of producing ground motions with $y \geq a_0$ at a site. The probability of exceedance $P(y \geq a_0)$, given the occurrence of an earthquake, can be calculated by using the total probability theorem:

$$P(y \geq a_0) = \iint P(y \geq a_0 | m, r) f_M(m) f_R(r)\, dm\, dr \tag{3.4}$$

where $f_M(m)$ and $f_R(r)$ denote the PDFs of magnitude and source to site distance, respectively, and $P(y \geq a_0 | m, r)$ is the conditional probability that an earthquake of magnitude m would cause ground motion $y \geq a_0$ at distance r from the source.

The generalisation of Eq. (3.3) for an instance of N seismic sources is straightforward, as the total annual rate of exceedance is the sum of the rates of individual seismic sources:

$$\lambda(a_0) = \sum_{i=1}^{N} \lambda_i(a_0) = \sum_{i=1}^{N} \nu_i P(y \geq a_0) \tag{3.5}$$

where subscript i indicates the i-th seismic source.

The substitution of Eq. (3.4) into Eq. (3.5) yields:

$$\lambda(a_0) = \sum_{i=1}^{N} \nu_i \iint P(y \geq a_0 | m, r) f_{M_i}(m) f_{R_i}(r)\, dm\, dr \tag{3.6}$$

The variability of the ground motion was recognised as an important element of seismic hazard calculations (Bender, 1984), and integration over the distribution of possible values of y was included in the calculations (Cornell, 1971; McGuire, 1976).

For a given combination of m and r, the GMPE allows estimation of the mean value of $\ln(y)$ and its standard deviation. The scatter of observed values of $\ln(y)$ around the mean is taken into account by considering the conditional probability distribution of $\ln(y)$ (Fig. 3.2). The standard assumption in current practice is that for the given m and r, y has a log-normal distribution (e.g. Abrahamson, 1988), or, equivalently, that $\ln(y)$ is normally distributed. Consequently, the conditional probability of exceedance $P(y \geq a_0 | m, r)$ is

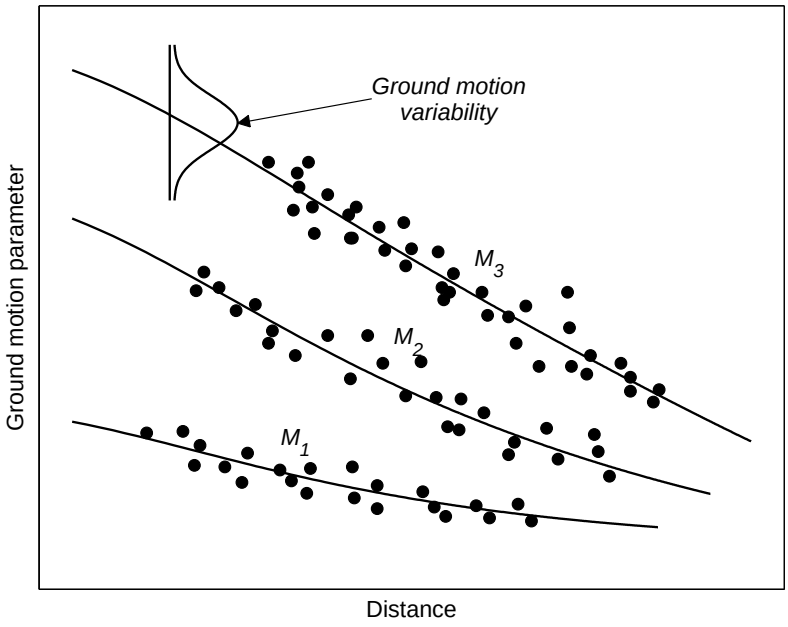


Figure 3.2: Aleatory variability of ground motion

calculated by using the normal distribution:

$$P(y \geq a_0|m,r) = \frac{1}{\sqrt{2\pi}\sigma}\int_{a_0}^{\infty} e^{-\frac{(\ln(u)-\mu)^2}{2\sigma^2}}\,du = 1-\Phi(z) \tag{3.7}$$

where $z = [\ln(a_0)-\mu]/\sigma$ is the standardised normal random variable and $\Phi(z)$ is the standard normal CDF.

As the normal distribution has unbounded support, any value of $\ln(y)$ receives a non-zero probability of being exceeded. Clearly, this is not appropriate, as the amount of energy released during an earthquake is finite, and, therefore, the ground motion should be bounded. As a result, truncating the distribution at some level above the median has become the standard practice. However, this practice has shortcomings, as was discussed above. A model that is more appropriate for the distribution of $\ln(y)$ would account for the finiteness of the ground motion induced by an earthquake of magnitude m at distance r.

3.4 Generalised extreme value distribution

The extreme value theory is devoted to the statistical analysis of rare events, and is used widely in various fields of knowledge, such as hydrology, meteorology, structural engineering, and earth sciences. This theory has a broad range of applications in the analyses of natural disasters (Pisarenko and Rodkin, 2010; 2014), and it is used in analysing the largest possible earthquakes (e.g. Epstein and Lomnitz, 1966; Kijko and Sellevoll, 1981).

The general result in the extreme value theory, the Fisher-Tippett-Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943), postulates that a properly normalised maximum from a sample $\{X_n, n \geq 1\}$ of independent identically distributed random variables, with distribution function F, can only converge in distribution to one of the three possible limiting distributions. Specifically, assume a sequence of constants $a_n > 0$, and $b_n \in \mathbb{R}$ $(n \geq 1)$, such that a normalised sample maximum has a non-degenerate limiting distribution:

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \tag{3.8}$$

Then, G must be one of the following three extreme value distributions:

$$\begin{array}{lll} \textit{Gumbel (type I):} & \Lambda(x) = \exp(-e^{-x}), & x \in \mathbb{R} \\ \textit{Fréchet (type II):} & \Phi_\alpha(x) = \exp(-x^{-\alpha}), & x > 0,\ \alpha > 0 \\ \textit{Weibull (type III):} & \Psi_\alpha(x) = \exp(-(-x)^{\alpha}), & x \leq 0,\ \alpha > 0 \end{array}$$

Three extreme value distributions can be combined into a single generalised form by introducing the shape parameter ξ so, that:

$$\xi = \begin{cases} 0 & \text{corresponds to } \Lambda(x) \\ \alpha^{-1} > 0 & \text{corresponds to } \Phi_\alpha(x) \\ -\alpha^{-1} < 0 & \text{corresponds to } \Psi_\alpha(x) \end{cases}$$

The following form is the GEVD, also called the Jenkinson-von Mises representation:

$$G_\xi(z) = \begin{cases} \exp(-(1+\xi z)^{-1/\xi}), & 1+\xi z > 0,\ \xi \neq 0 \\ \exp(-e^{-z}), & z \in \mathbb{R},\ \xi = 0 \end{cases} \tag{3.9}$$

where $z = (x-\mu)/\sigma$, and μ, σ and ξ are the location, scale, and shape parameters, respectively.

The distribution function F belongs to the domain of attraction of G_ξ if (3.8) holds with $G = G_\xi$. The Gumbel domain of attraction $D(G_0)$ includes a large variety of distributions, of which the tails can be differ significantly; ranging from moderately heavy, such as the log-normal distribution to very light, such as the normal distribution. The Fréchet domain of attraction $D(G_{\xi_+})$ consists of the heavy-tailed distributions, of which the right tail behaves like a power law. Such distributions include the Pareto, Cauchy, Student's-t, and the Fréchet distributions. The Weibull domain of attraction $D(G_{\xi_-})$ includes distributions with finite right endpoints (Fig. 3.3), for example, the uniform and the beta distributions.

Suppose a is the horizontal acceleration induced by an earthquake with magnitude m at distance r from the source. Let $a_{\max}$ denote the upper limit of a, then the possible values of $\ln(a)$ would be bounded by $(-\infty, \ln(a_{\max})]$, and the distribution of $\ln(a)$ would belong to the Weibull domain of attraction. As PGA is a maximum of a, it would be reasonable to expect that the distribution of $\ln(\text{PGA})$ would converge to the Weibull extreme value distribution.

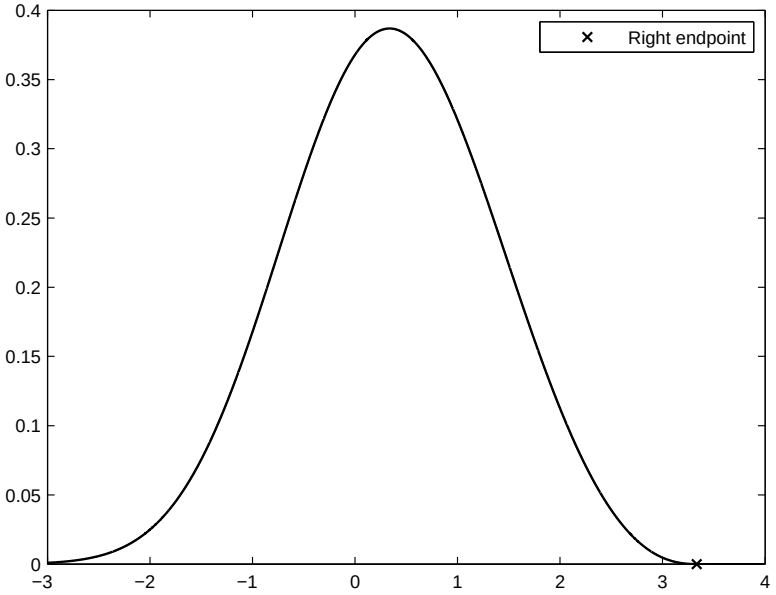


Figure 3.3: PDF of a distribution with the support bounded on the right

3.5 Applied procedure and data

The same procedure applied in Pavlenko (2015) was used in this study to examine the probability distribution of ln(PGA), with the normal distribution and the GEVD being considered as potential models in the current analysis. The data of the strong-motion seismograph networks of Japan (www.kyoshin.bosai.go.jp) were used in the study. The stations located on very dense soil (National Earthquake Hazards Reduction Program [NEHRP] class C) were selected, and the records of events with focal depths from 10 to 20 km and hypocentral distances from 46 to 54 km were used in the analysis. Some extension of the distance range is inevitable to gain enough data for statistical analysis. It is believed that the data obtained are representative of the random scatter in ln(PGA) for the given m and r. The data were grouped into bins according to magnitude, and the empirical distributions of ln(PGA) were modelled for the magnitude bins that contained the bulk of the data.

Various parameters have been used in ground motion studies to describe the level of horizontal acceleration, such as the square root of the sum of squares of two horizontal components (e.g. Kanno et al., 2006), and the geometric mean of the two components (e.g. Zhao et al., 2006). In this study, these two parameters are used for horizontal PGA, the first one is denoted PGA_{SR} (square root) and the second is denoted PGA_{GM} (geometric mean).

The PDFs of the normal distribution and the GEVD are given by:

$$\phi(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-z^2/2),\ z \in \mathbb{R} \tag{3.10}$$

$$g(z) = \begin{cases} \exp\left(-(1+\xi z)^{-1/\xi}\right)(1+\xi z)^{-1/\xi-1}, & 1+\xi z > 0,\ \xi \neq 0 \\ \exp\left(-e^{-z}-z\right), & z \in \mathbb{R},\ \xi = 0 \end{cases} \tag{3.11}$$

The parameters of the normal distribution were estimated by the ML method. The basic principle of this method is that the parameter values that maximise the likelihood of obtaining the given sample in a series of experiments should be taken as the most plausible estimates. In practice, this is achieved by maximisation of the likelihood function L, or its natural logarithm ℓ, called the log-likelihood function:

$$\ell = \ln(L) = \sum_{i=1}^{n} \ln[f(x_i|\theta)] \tag{3.12}$$

The details of the method for normal distribution are well known and therefore require no explanation. The estimators for parameters μ and σ are the sample mean and the sample standard deviation. The estimation of the parameters of the GEVD has its nuances; the support of G_ξ depends on the unknown values of the parameters. Furthermore, the applicability of the estimation methods depends on the value of ξ. Various methods exist for estimating ξ, which could be applied in different circumstances. For instance, the well-known Hill estimator (1975) is applicable only for positive values of ξ; the ML estimator is valid for $\xi > -0.5$; the probability-weighted moment estimator (Hosking et al., 1985) is valid for $\xi < 1$; and the Pickands estimator (1975) and the moment estimator proposed by Dekkers et al. (1989) can be applied in the general instance ($\xi \in \mathbb{R}$). Detailed reviews of these estimation techniques can be found in Embrechts et al. (1997), Beirlant et al. (2004), and de Haan and Ferreira (2006).

In the present study, the condition $\xi > -0.5$ was fulfilled for the analysed data and therefore the ML method could be applied. The log-likelihood function of the sample $\{X_n, n \geq 1\}$ of GEVD random variables for the instance $\xi \neq 0$ is given by:

$$\ell = -n\ln(\sigma) - \left(\frac{1}{\xi}+1\right)\sum_{i=1}^{n}\ln(1+\xi z_i) - \sum_{i=1}^{n}(1+\xi z_i)^{-1/\xi} \tag{3.13}$$

where $z_i = (X_i - \mu)/\sigma$.
Differentiating (3.13) with respect to μ, σ, and ξ, yields the following likelihood equations:

$$\begin{cases} \sum_{i=1}^{n} a_i b_i = 0 \\ \sum_{i=1}^{n} z_i a_i b_i - n = 0 \\ \sum_{i=1}^{n} z_i a_i b_i + \frac{1}{\xi} \sum_{i=1}^{n} \ln(a_i)(b_i - \xi) = 0 \end{cases} \tag{3.14}$$

where $a_i = (1 + \xi z_i)^{-1}$, $b_i = 1 + \xi - (1 + \xi z_i)^{-1/\xi}$.

These equations have no explicit solution and therefore should be solved numerically. Iterative numerical procedures were proposed for this purpose by among others Prescott and Walden (1980) and Hosking (1985). In the instance $\xi > -0.5$, the ML method provides consistent, efficient, and asymptotically normal estimators. The covariance matrix of vector $\hat{\theta} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ can be obtained from the inverse of the Fisher information matrix:

$$I_{ij} = -\left(\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} \right) \Bigg|_{\theta = \hat{\theta}} \tag{3.15}$$

The KS test (Kolmogorov, 1933) was used to test for significant deviations between the theoretical and the empirical distributions. The goodness of the fit of the models was compared by using the AIC (Akaike, 1974):

$$\mathrm{AIC} = -2\ln(L) + 2k \tag{3.16}$$

where L is the maximum value of the likelihood function, and k is a number of parameters of probability distribution. The AIC allows estimating the information loss as a result of using the particular model. When a set of candidate models is considered, with AIC_i denoting the AIC value of the i-th model, and $\mathrm{AIC}_{\min}$ denoting the minimum of those values, then, for the i-th model, the relative likelihood can be calculated as follows:

$$\tilde{L} = \exp[(\mathrm{AIC}_{\min} - \mathrm{AIC}_i)/2] \tag{3.17}$$

This quantity measures the relative probability of the i-th model to minimise the information loss.

3.6 Results and discussion

The histograms of $\ln(\mathrm{PGA}_{SR})$ and $\ln(\mathrm{PGA}_{GM})$, with fitted normal distribution, and the GEVD are shown in Figs. 3.4 and 3.5, and the relative likelihoods are listed in Table 3.1. The majority of histograms have a similar shape, with a slightly elongated right tail. There is good agreement between the results obtained for $\ln(\mathrm{PGA}_{SR})$ and $\ln(\mathrm{PGA}_{GM})$, and, relevant to both parameters, the normal distribution performed better than the GEVD only in one instance out of twelve. Although there is no obvious trend, all the estimates of ξ were negative, which confirms the convergence of data to the bounded Weibull extreme value distribution. Similar results were obtained by Huyse et al. (2010) and Pavlenko (2015).

The applicability of the GEVD for peak ground motion parameters is supported by the extreme value theory. The GEVD is a flexible distribution that can assume a variety of shapes and its shape parameter ξ governs the decay of the tail of distribution. In instance $\xi < 0$, the support of the GEVD is bounded on the right. The inverse distribution function of the GEVD is:

$$Q(p) = \begin{cases} \mu + \sigma\{[-\ln(p)]^{-\xi} - 1\}/\xi & \xi \neq 0 \\ \mu - \sigma\{\ln[-\ln(p)]\} & \xi = 0 \end{cases} \tag{3.18}$$

In instance $\xi < 0$, the right endpoint of the support is given by:

$$x^F = Q(1) = \mu - \frac{\sigma}{\xi} \tag{3.19}$$

The estimate of x^F can be obtained by substituting the estimates of the parameters into Eq. (3.19). Thereby, the GEVD allows accounting for the finiteness of the seismic ground motion, and provides a rational way of estimating the maximum value of PGA for a specified earthquake scenario. This is a viable alternative to the common practice of using the truncated normal distribution (e.g. Strasser et al.,2008b) for modelling the scatter of the logarithm of peak ground motion parameters.

	PGA_{SR}		PGA_{GM}	
M	$\tilde{L}(G_\xi)$	$\tilde{L}(\Phi)$	$\tilde{L}(G_\xi)$	$\tilde{L}(\Phi)$
3.4	1.0	0.497	1.0	0.313
3.5	1.0	0.133	1.0	0.092
3.6	1.0	0.344	1.0	0.113
3.7	1.0	0.032	1.0	0.013
3.8	1.0	0.173	1.0	0.084
3.9	1.0	0.479	1.0	0.394
4.0	0.016	1.0	0.028	1.0
4.1	1.0	0.312	1.0	0.321
4.2	1.0	0.847	1.0	0.872
4.3	1.0	0.029	1.0	0.034
4.4	1.0	0.402	1.0	0.414
4.5	1.0	0.758	1.0	0.699

Table 3.1: Relative likelihoods of the GEVD (G_ξ) and the normal distribution (Φ)

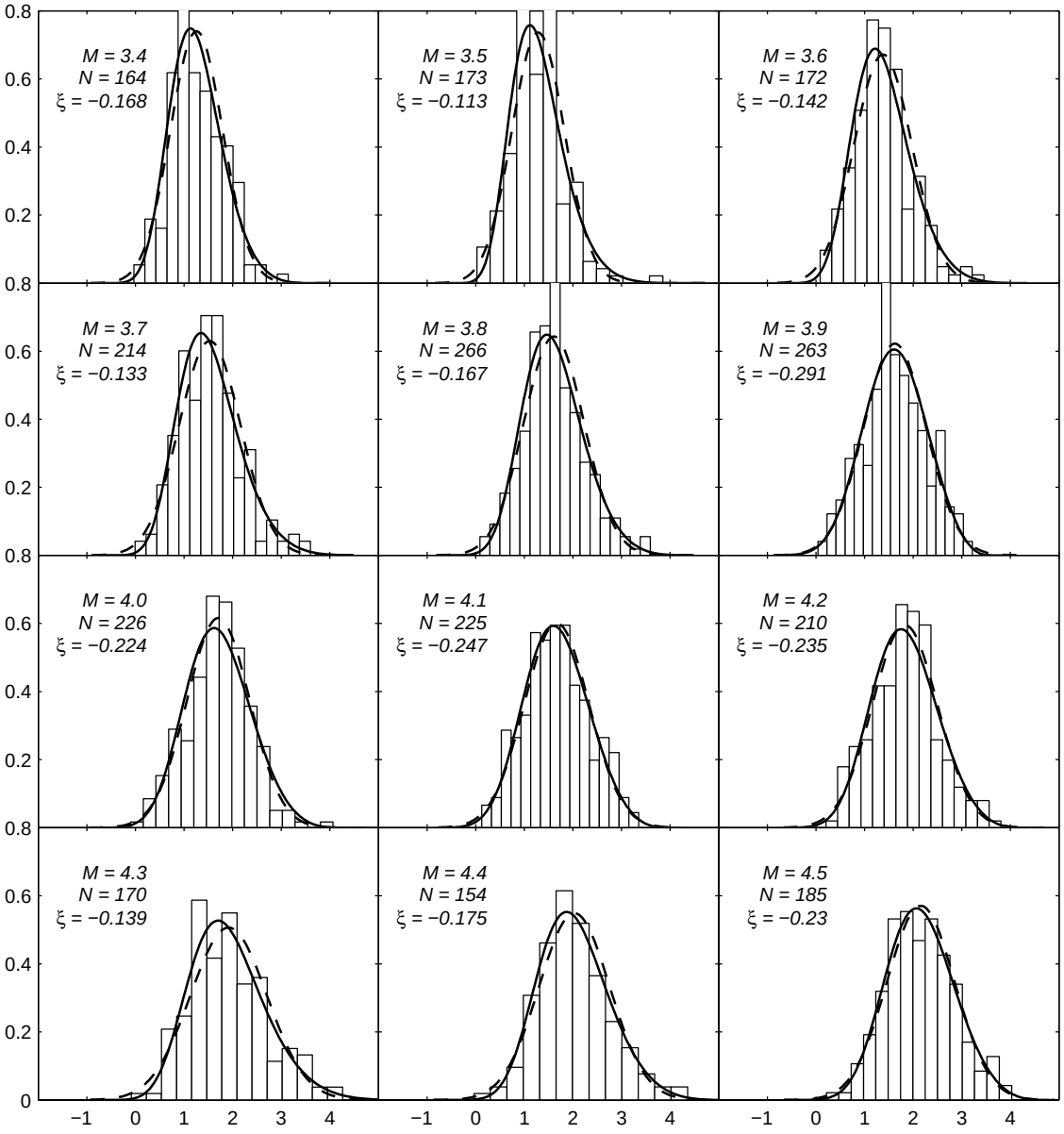


Figure 3.4: Sample histograms of $\ln(\mathrm{PGA}_{SR})$. The PDFs of the normal distribution and the GEVD are shown by the broken and the solid lines, respectively. Magnitude (M), sample size (N) and estimated value of ξ are shown for each histogram

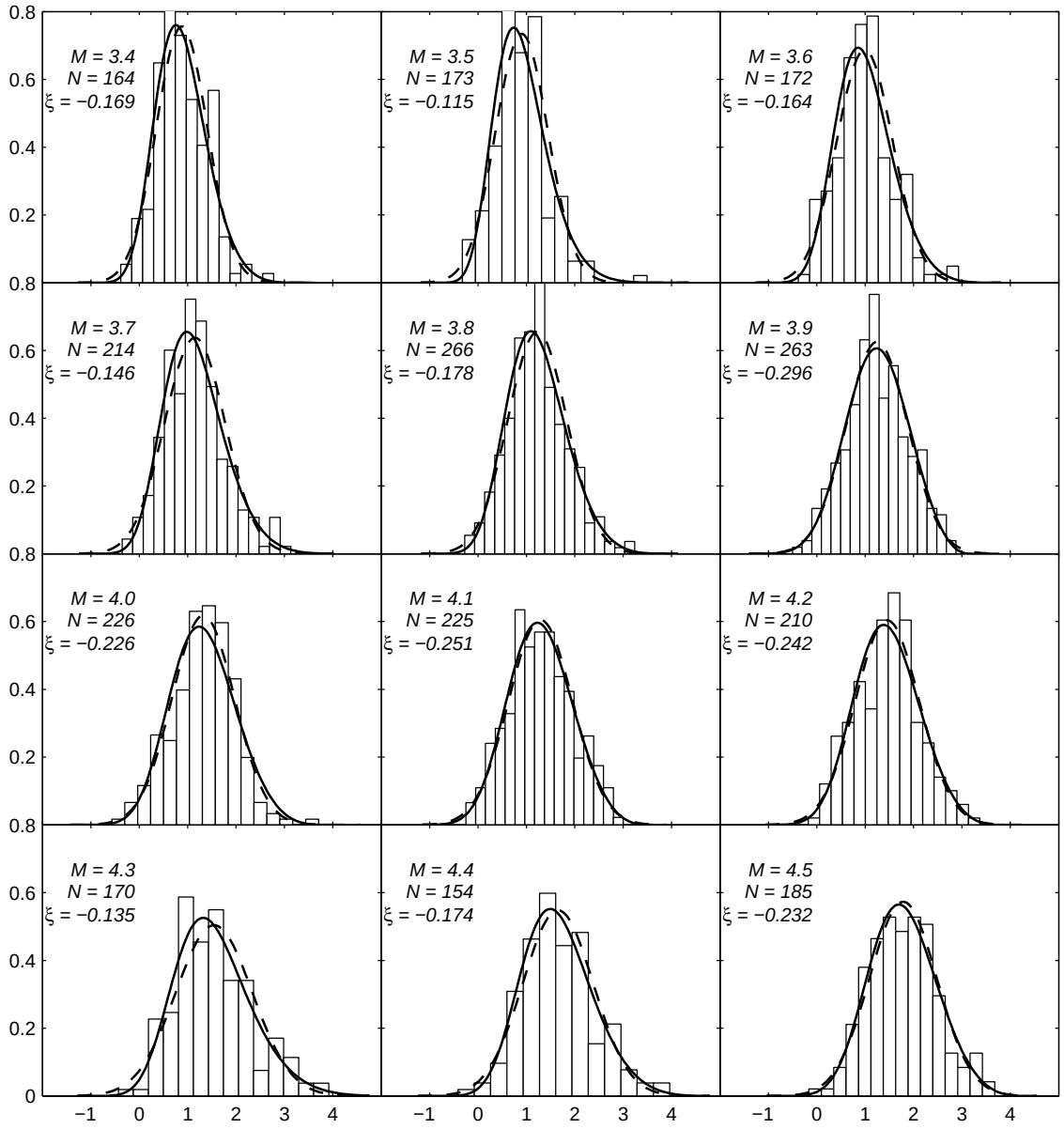


Figure 3.5: The same as in Fig. 3.4 for $\ln(\mathrm{PGA}_{GM})$

3.7 Conclusion

In this study, the distribution of ln(PGA) was investigated by using the data of the strong-motion seismograph networks of Japan. The normal distribution and the GEVD were used for modelling the empirical distribution of ln(PGA). Two definitions of horizontal PGA were used, namely, the square root of the sum of squares of two horizontal components and the geometric mean of the two horizontal components. Similar results were obtained for both definitions, the GEVD provided a better fit than the normal distribution in eleven out of twelve instances. The estimated values of the shape parameter of the GEVD were negative in every instance, indicating that the support of the distribution is bounded on the right. Therefore, the GEVD provides a more realistic model for the scatter of the ln(PGA), which allows accounting for the finiteness of the ground motion induced by a specified earthquake scenario. The maximum value of PGA can be estimated directly from the parameters of the GEVD. Correct modelling of the ground motion parameters is important for realistic seismic hazard assessment and the studies on the statistical properties of these parameters should therefore be continued.

Acknowledgements I would like to thank Professor Andrzej Kijko for the review of the text, and the valuable suggestions that helped to improve its quality. I am grateful to the K-NET and KiK-net strong-motion seismograph networks of Japan for seismic ground motion records (http://www.kyoshin.bosai.go.jp).

Chapter 4

On Anisotropic Attenuation Law of Modified Mercalli Intensity

4.1 Abstract

Seismic intensity is frequently used to describe the effects of earthquakes. In many instances, isoseismal maps demonstrate an elongated elliptical shape, indicating the influence of anisotropy in the attenuation of the seismic intensity. In this purely methodological study, a set of equations that could account for this anisotropy is proposed. These equations were validated by modelling the isoseismal maps of two well-recorded seismic events that occurred in South Africa. The ability of the new equations to model the observed data is compared with that of the classic isotropic attenuation law. The results of the analysis indicate that, in general, the new equations provide a superior fit to the observed data, especially as regards pronounced anisotropy.

4.2 Introduction

An intensity scale is defined as the distance-dependent value that describes the effects of a seismic event on humans, structures, and on the surface of the Earth at a particular site. The observed effects become less noticeable as the distance from the hypocentre of the seismic event increases. PGA, the maximum measured (or expected) acceleration amplitude of an earthquake, is an alternative measure often used to describe seismic effects.

Despite the concept of seismic intensity measurements being at least 80 years old, it remains a subject of extensive research worldwide. The MM intensity scale and the MM IPEs are the focus of continuous research in USA (Dengler and Dewey, 1998; Bakun, 2000; Dewey et al., 2000; Kaka and Atkinson, 2004; Bakun, 2006), Canada (Jalpa and Atkinson, 2012), and Central Asia (Bindi et al., 2014).

The IPEs form integral part of the ShakeMap code (Atkinson and Kaka, 2007), and are used, e.g., in Chile (Barrientos et al., 2004) and Japan (Bakun, 2005), and are considered the standard in Europe (EMS-98, 1998; Bakun and Scotti, 2006). In addition, these equations are used widely to determine the location and magnitude of historical earthquakes (Bakun and Wentworth, 1997; Bakun, 2005, 2006; Bakun and Scotti, 2006; Park and Hong, 2016).

The introduction of the Prompt Assessment of Global Earthquakes for Response (PAGER) system (Wald et al., 2008; Earle et al., 2009) has contributed significantly to the application and improved understanding of the MM intensity scale by the public, the media, and disaster management authorities. Worden et al. (2010) and Allen et al. (2012) have provided excellent reviews on the application of the intensity scale in ShakeMaps in active crustal regions.

Using the MM intensity-based prediction equations posses several significant advantages compared with using the GMPEs based on the PGA. One of these advantages is related to the effect of local conditions on the amplitude of the surface ground motion. Accounting for the site effect could be a difficult task with

respect to the PGA (e.g., Chiou et al., 2008); however, this effect is taken into account automatically in regional IPEs.

Another advantage of using MM IPE is observed in the application of the vulnerability curves, often expressed in terms of MM intensities, especially in the insurance industry. Employing MM intensity-based damage curves allows for the calculation of seismic risk, without the additional conversion of the PGA to the MM intensity or vice versa. If such conversions are required in a risk assessment procedure, various conversion equations could be used. These include the equations by Neumann (1954), Gutenberg and Richter (1956), Ambraseys (1974), Trifunac and Brady (1975), Murphy and O'Brien (1977), Saragoni et al. (1982), Wald et al. (1999), and Worden et al. (2012).

However, similar to the GMPE problem, no systematic database of damage to infrastructure exists and, therefore, no region-characteristic vulnerability curves relevant to seismicity in South Africa exist. Moreover, it would not be possible to construct such curves in the near future. Since the required calibration data for South African conditions are non-existent, the above conversion equations cannot be applied with any level of accuracy. Consequently, the ATC-13 (1985) damage curves, or similar, are used for South African seismic risk assessment. Alternative vulnerability curves, e.g., Risk-UE (Mouroux et al., 2004) were excluded because of significant uncertainties relevant to the predicted damages (G. Trendafiloski, personal communication 2011).

Currently, IPEs are often applied that do not account for anisotropy of the attenuation of seismic intensity. However, in many instances, the anisotropic representation of the medium facilitates significantly superior agreement with the results of the observations than the isotropic representation does. Therefore, there is a clear need to develop appropriate attenuation equations. The current study is purely methodological and its main purpose is to suggest a suitable method to account for the anisotropy in IPEs. Accordingly, a suitable set of equations is proposed, which is validated by modelling the isoseismal maps of two well-recorded seismic events.

4.3 Applied models

The dependence of amplitude A of an elastic wave on distance r from the emitter can be expressed as:

$$A(r) \sim A_0 e^{-\alpha r} r^{-\beta} \tag{4.1}$$

where A_0 is the amplitude at the emitter, α is the frequency-dependent coefficient of the attenuation, $r^{-\beta}$ is the geometrical spreading factor, and β is a distance-dependent coefficient (Boore, 2003).

It can be shown that this type of dependence leads to an attenuation equation of seismic intensity of the following form (Peruzza, 1996):

$$I_0 - I = -a_1 - a_2 \ln(r) - a_3 r \tag{4.2}$$

where I_0 is the maximum (focal) MM intensity at the epicentre, I is the MM intensity at the site of interest, a_1, a_2, and a_3 are parameters, and r is the epicentral/hypocentral distance (km).

The numerical values of parameters a_1, a_2, and a_3 vary depending on the region of investigation and they are usually estimated from MM intensity distribution maps by using the standard least squares regression (Carnahan et al., 1969).

The empirical relation between earthquake magnitude m and the MM intensity at the epicentre I_0, is given by (Richter, 1958):

$$I_0 = \frac{3}{2} m - 1 \tag{4.3}$$

The IPE (4.2) does not take into account that the ground-motion amplitude variation is a function of the direction of the rupture propagation, focal mechanism, and medium anisotropy. All these factors contribute to the ground-motion amplitude variation as a function of azimuth and, in contrast with intensity, their effect on ground motion acceleration is well understood (e.g. Yamazaki and Türker, 1992; Somerville et al., 1997; Miyake et al., 2001; Spudich and Chiou, 2008).

Seismic anisotropy is a characteristic of the medium that reflects the horizontal component of the dependence of the velocity of propagation and the attenuation of the seismic waves on the azimuth. The effect of anisotropy can be introduced into IPE in several ways.

Accordingly, Papazachos (1992), in examining the principal parameters of the attenuation of the upper crust in Greece, suggested an anisotropic model for the radiation of seismic intensity at the source. The new model appeared to be significantly more reliable compared with the previously proposed isotropic models.

An approach proposed by Teramo et al. (1995) is based on the suitable mechanical modelling of the macroseismic fields and allows for the quantitative characterisation of their anisotropy. The resulting attenuation model was compared with previously developed models by analysing the macroseismic data from a set of earthquakes that occurred in two regions of Italy. The comparison demonstrated the greater adaptability of the new model to the observed data.

Termini et al. (2005) applied the same modelling technique and obtained an anisotropic attenuation law with the same structure as that of Teramo et al. (1995). The reliability of the proposed attenuation law was confirmed by analysing local and regional macroseismic data.

Zonno et al. (2009) utilised the Bayesian method to estimate the probability distribution of macroseismic intensity at a particular site and obtained two probabilistic models for forecasting the intensity levels in Italy. The applicability of the models was demonstrated by forecasting the scenario produced by the Colfiorito earthquake of 1997.

Sørensen et al. (2010) derived new IPEs, based on a previously developed model (Sørensen et al., 2009), by introducing a regional correction factor that accounts for the anisotropic intensity distribution characteristic for the Vrancea region of Romania.

The IPEs suggested in this study are modifications of the IPE (4.2) to account for anisotropy of the attenuation of seismic intensity. It is anticipated that in the horizontal plane variation of characteristics of the medium with an azimuth has a period of 180°, that is, in opposite directions these characteristics would be similar. Moreover, there is a point where the differences between the characteristics on the two orthogonal directions reach maximum values. In view of these considerations, the following set of equations is proposed, suitable for accounting for the effect of anisotropy:

$$I_0 - I = -a_1 - a_2 \ln(r) - a_3 r|\cos(\Phi)| \tag{4.4}$$

$$I_0 - I = -a_1 - a_2 \ln\{r[a_3 + a_4|\cos(\Phi)|]\} - r[a_3 + a_4|\cos(\Phi)|] \tag{4.5}$$

$$I_0 - I = -a_1 - a_2 \ln(r) - a_3 r\cos(2\Phi) \tag{4.6}$$

$$I_0 - I = -a_1 - a_2 \ln\{r[a_3 + a_4\cos(2\Phi)]\} - r[a_3 + a_4\cos(2\Phi)] \tag{4.7}$$

where $a_1, \ldots, a_k$ are the parameters to be determined, r is the epicentral distance to the apparent epicentre, I is the MM intensity at distance r, I_0 is the maximum MM intensity at the epicentre, $\Phi = (Az - \phi)$, Az is the azimuth from the apparent epicentre to a point, and ϕ is the angle between the dominant axis of the MM intensity distribution and the northern direction.

The apparent epicentre is defined as the centre point of the MM intensity distribution, whereas the epicentre refers to the point on the surface of the Earth directly above the hypocentre. Various factors lead

to the epicentre and the apparent epicentre not always coinciding as, e.g., in the instance of the August 5, 2014 Orkney (South Africa) seismic event. Likewise, Sørensen et al. (2010) have indicated a systematic shift between the epicentre and location of the maximum intensities of the earthquakes of the Vrancea region in Romania.

Equations (4.4) and (4.6) are linear with respect to parameters a_1, a_2, and a_3; therefore the standard linear least square procedure could be applied. The solution to the system of three equations can be written in matrix form, as follows:

$$(A^{-1}A)a = A^{-1}y \tag{4.8}$$

where A is the following $(n \times 3)$ matrix:

$$A = \begin{pmatrix} -1 & -\ln(r_1) & -r_1F_1 \\ \vdots & \vdots & \vdots \\ -1 & -\ln(r_n) & -r_nF_n \end{pmatrix} \tag{4.9}$$

with

$$F_i = \begin{cases} \cos[2(Az_i - \phi)] \\ |\cos(Az_i - \phi)| \end{cases} \tag{4.10}$$

Vectors y and a are column vectors defined as $y = (I_0 - I_1, \ldots, I_0 - I_n)^T$, and $a = (a_1, \ldots, a_k)^T$. Subscript i $(= 1, 2, \ldots, n)$ represents the individual intensity data points, with a total of n such data points available. The epicentral distance r_i is defined as the distance between the apparent epicentre and point i. The MM intensity observed at point i is denoted by I_i and Az_i is the azimuth from the apparent epicentre to point i. The matrix operator T denotes the transpose of the respective matrix or vector.

Equations (4.5) and (4.7) are not linear with respect to parameters a_1, a_2, a_3, and a_4; therefore, estimating these values represents a more sophisticated problem. For these equations, the parameters were estimated by minimising the sum of the squared residuals, with the application of the derivative-free Nelder-Mead simplex procedure (Nelder and Mead, 1965). The starting point for the simplex algorithm was obtained as the solution of a linearised system similar to (4.8), with matrix A_{L} given by:

$$A_{\mathrm{L}} = \begin{pmatrix} -1 & -\ln(r_1) & -r_1 & -r_1F_1 \\ \vdots & \vdots & \vdots & \vdots \\ -1 & -\ln(r_n) & -r_n & -r_nF_n \end{pmatrix} \tag{4.11}$$

The variance-covariance matrix for the vector of the unknown parameters a, for each of the four defined models (4.4) to (4.7), is given by:

$$D(\hat{a}) = s^2 \left(A^T A\right)^{-1} \tag{4.12}$$

where

$$s^2 = \frac{(\hat{y} - y)^T (\hat{y} - y)}{n - k} \tag{4.13}$$

As regards equations (4.4) and (4.6), A is the $(n \times 3)$ matrix defined by equation (4.9). As regards

equations (4.5) and (4.7), A is the $(n \times 4)$ matrix of partial derivatives evaluated at the solution point $\hat{a}$:

$$A = \begin{pmatrix} -1 & -\ln[r_1(a_3+a_4F_1)] & \frac{-a_2}{a_3+a_4F_1} - r_1 & \left(\frac{-a_2}{a_3+a_4F_1} - r_1\right)F_1 \\ \vdots & \vdots & \vdots & \vdots \\ -1 & -\ln[r_n(a_3+a_4F_n)] & \frac{-a_2}{a_3+a_4F_n} - r_n & \left(\frac{-a_2}{a_3+a_4F_n} - r_n\right)F_n \end{pmatrix}\Bigg|_{(a=\hat{a})} \tag{4.14}$$

Vector $\hat{y}$ denotes the estimated values of $I_0 - I_i$. By definition, the sample standard deviation (standard error) of each unknown parameter is therefore given by $\sigma_{a_j} \equiv \sqrt{D_{jj}}$ for $j = 1, \ldots, k$ and D_{jj} is the diagonal element of matrix D.

The results of applying models (4.2),(4.4),(4.5),(4.6), and (4.7) were compared based on the AIC (Akaike, 1971), defined as:

$$\text{AIC} = -2\ln(L) + 2k \tag{4.15}$$

where L is a maximised likelihood function, and k is the number of parameters to be estimated.

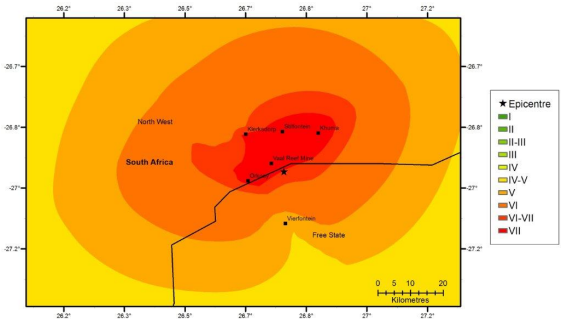
The calculated values of the AIC were normalised. Consequently, AIC = 1 corresponds to the model providing the worst approximation for the observed data, whereas the model for which AIC was minimal, provided the best approximation among the considered alternatives.

4.4 Case studies

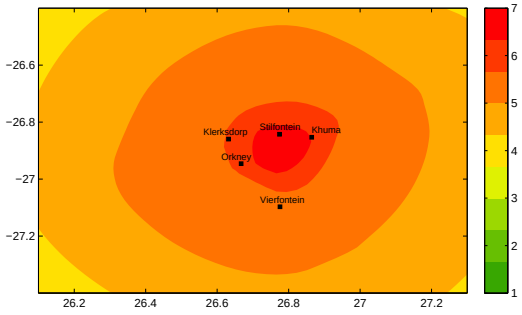
Two well-recorded seismic events in South Africa were considered to demonstrate the performance of the proposed set of models. The first event is an earthquake of M_{L} 5.5 that occurred in the Orkney area of the North West Province of South Africa on August 5, 2014. The earthquake shaking was felt widely in southern Africa and as far from the epicentre as Cape Town, Maputo (Mozambique), and Gaborone (Botswana). One person was killed and more than 600 houses were damaged. The coordinates of the epicentre were estimated as 26.942°S, 26.818°E, and the focal depth was estimated as 4.7 km. The communities located in the epicentral area experienced a maximum MM intensity of VII (Midzi et al., 2015).

The second event is an earthquake of M_{L} 6.3 that occurred on September 29, 1969 in the Ceres-Tulbagh region of the Cape Province of South Africa. This event caused several casualties and considerable damage to properties in and around the towns of Ceres, Tulbagh, and Wolseley, with reports of significant damage covering an area of approximately 2000 km^2. The strong shock caused panic in Cape Town, located 100 km from the source region, while reliable reports were received from locations as far as Johannesburg (1150 km away). The epicentral coordinates were estimated as 32.9°S, 19.7°E (Green and Bloch, 1971). The reported estimates of the focal depth vary from 4 to 33 km; however, a recent estimate of 15 km was given by Scherbaum and Krüger (2014). This seismic event remains the most destructive earthquake ever experienced in South Africa.

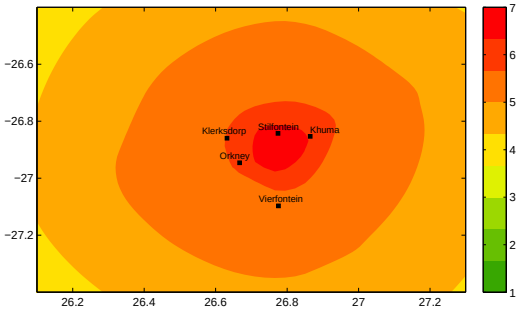
Figs. 4.1 and 4.2 demonstrate the actual and modelled isoseismal maps of the Orkney and Ceres-Tulbagh earthquakes, respectively. The estimated parameters of the IPEs, corresponding standard errors, and the normalised values of the AIC are listed in Tables 4.1 and 4.2. IPE (4.5) provides the best approximation for the isoseismal map of the Orkney seismic event (Table 4.1), whereas that of the Ceres-Tulbagh seismic event is provided by IPE (4.6) (Table 4.2).



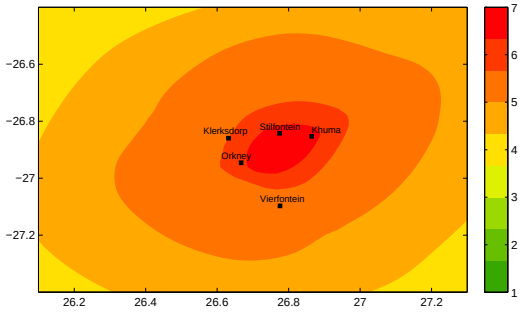
(a) Actual isoseismal map



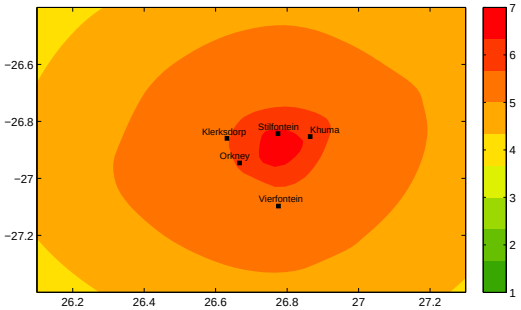
(b) Result of IPE (4.2)



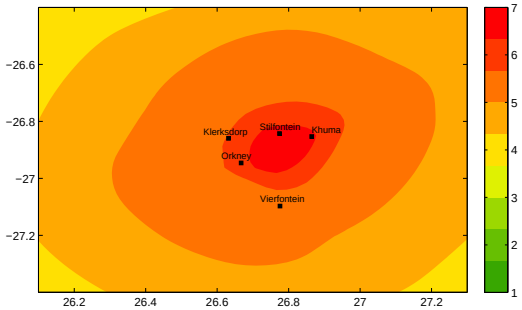
(c) Result of IPE (4.4)



(d) Result of IPE (4.5)



(e) Result of IPE (4.6)

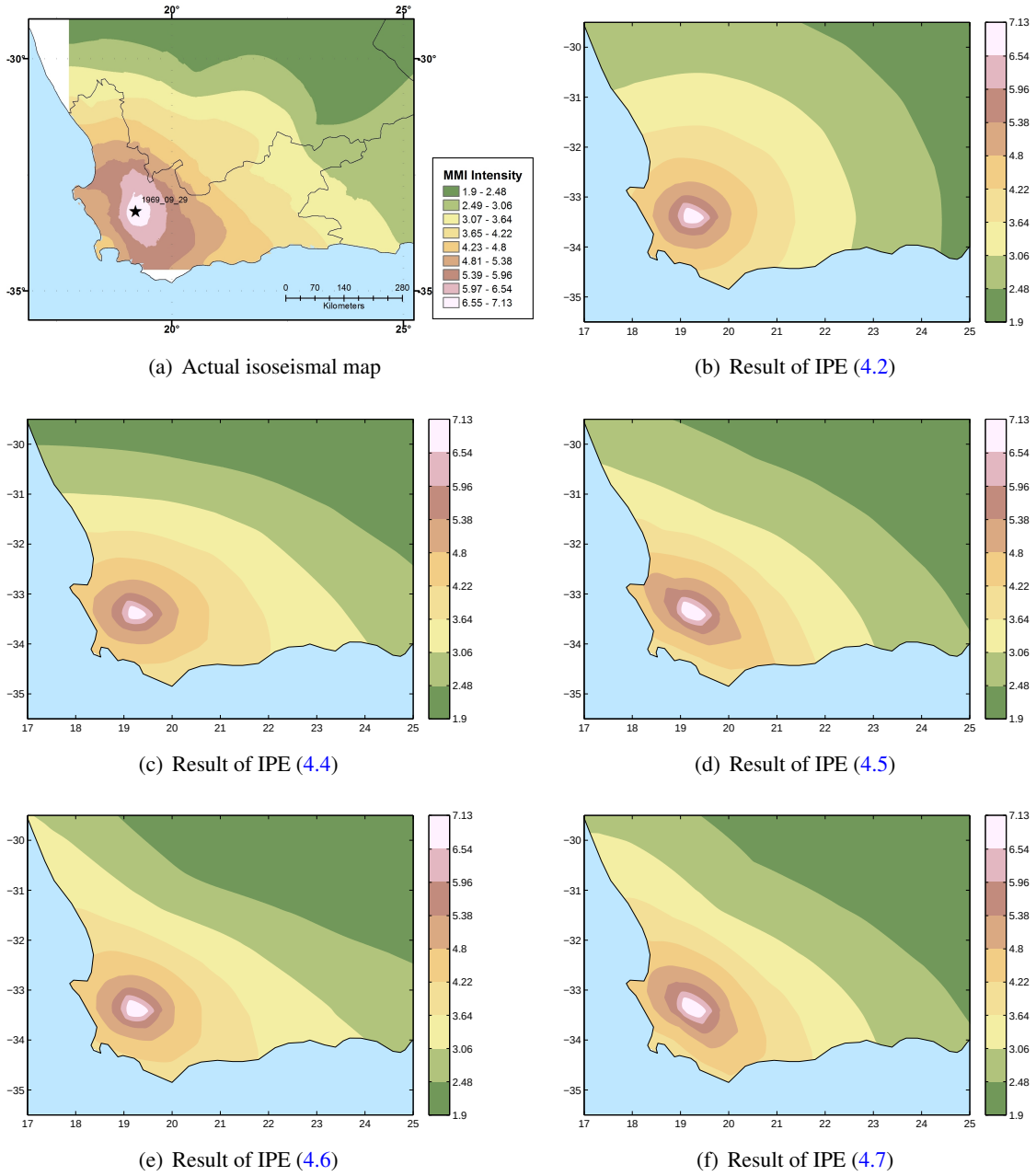


(f) Result of IPE (4.7)

Figure 4.1: Actual and modelled isoseismal maps of the Orkney earthquake of August 5, 2014

	$a_1 \pm \sigma_{a_1}$	$a_2 \pm \sigma_{a_2}$	$a_3 \pm \sigma_{a_3}$	$a_4 \pm \sigma_{a_4}$	AIC
IPE (4.2)	2.0116 ± 0.2841	-1.0767 ± 0.0691	0.0011 ± 0.0005	-	0.980
IPE (4.4)	1.9280 ± 0.2457	-1.0538 ± 0.0552	0.0018 ± 0.0006	-	0.966
IPE (4.5)	-5.2010 ± 0.1942	-1.1402 ± 0.0719	0.0012 ± 0.0004	0.0006 ± 0.0003	0.953
IPE (4.6)	1.5801 ± 0.2250	-0.9498 ± 0.0448	0.0002 ± 0.0002	-	1.000
IPE (4.7)	-5.2418 ± 0.2285	-1.1101 ± 0.0717	0.0014 ± 0.0005	0.0002 ± 0.0001	0.977

Table 4.1: Parameters, standard errors, and the normalised values of the AIC estimated from the macroseismic data of the Orkney earthquake of August 5, 2014



(a) Actual isoseismal map

(b) Result of IPE (4.2)

(c) Result of IPE (4.4)

(d) Result of IPE (4.5)

(e) Result of IPE (4.6)

(f) Result of IPE (4.7)

Figure 4.2: Actual and modelled isoseismal maps of the Ceres-Tulbagh earthquake of September 29, 1969

	$a_1 \pm \sigma_{a_1}$	$a_2 \pm \sigma_{a_2}$	$a_3 \pm \sigma_{a_3}$	$a_4 \pm \sigma_{a_4}$	AIC
IPE (4.2)	4.6170 ± 0.7929	-1.6038 ± 0.1793	0.0005 ± 0.0006	-	1.000
IPE (4.4)	5.6311 ± 0.5710	-1.8512 ± 0.1242	0.0023 ± 0.0006	-	0.888
IPE (4.5)	-7.6241 ± 1.2954	-1.5226 ± 0.1579	0.0003 ± 0.0003	0.0003 ± 0.0003	0.896
IPE (4.6)	4.2693 ± 0.3729	-1.5231 ± 0.0690	0.0016 ± 0.0003	-	0.871
IPE (4.7)	-7.5821 ± 1.2190	-1.5396 ± 0.1564	0.0005 ± 0.0005	0.0001 ± 0.0002	0.879

Table 4.2: Parameters, standard errors, and the normalised values of the AIC estimated from the macroseismic data of the Ceres-Tulbagh earthquake of September 29, 1969

4.5 Conclusion

In this methodological study, a set of models was proposed that is suitable for describing the anisotropy in the attenuation of seismic intensity. The applicability of the proposed models was demonstrated by modelling the isoseismal maps of two well-recorded seismic events that occurred in South Africa on September 29, 1969 and August 5, 2014. By calculating the AIC, the results of the application of the proposed models were compared with the results obtained by using the isotropic model.

A surprising finding was that the performance of the isotropic IPE (4.2) was superior to that of the IPE (4.6) when the isoseismal map of the Orkney seismic event was modelled (Table 4.1). This result could probably be attributed to the relatively weak anisotropy of the observed macroseismic fields. However, relevant to the more pronounced anisotropy observed for the Ceres-Tulbagh seismic event, a new set of IPEs facilitated a considerable improvement compared with the results of the isotropic IPE (4.2) (Table 4.2). In general, the proposed set of IPEs provided a superior fit to the observations. This result agrees with the conclusions reached by Papazachos (1992) and Teramo et al. (1995).

The findings of this study indicate that the best approximations for the two considered seismic events were obtained by different models. This could be interpreted as evidence that no single universal anisotropic IPE would be applicable equally to every seismic event.

Acknowledgements We are grateful to the editor of the Journal of Seismology, Dr. Mariano García Fernández, for help and suggestions regarding available literature on anisotropy of attenuation of seismic intensity. We appreciate helpful comments and suggestions by Dr. Päivi Mäntyniemi from the Institute of Seismology, University of Helsinki.

Chapter 5

Comparative study of three probabilistic methods for seismic hazard analysis: Case studies of Sochi and Kamchatka

5.1 Abstract

This study examines the effect of the procedures used in three different PSHA methods for estimating the rates of exceedance of ground motion. To evaluate the effect of these procedures, the Cornell-McGuire and Parametric-Historic methods, and the method based on Monte Carlo simulations are employed, and the seismic source model, based on spatially smoothed seismicity, is used in the calculations. Two regions in Russia were selected for comparison, and seismic hazard maps were prepared for return periods of 475 and 2475 years. The results indicate that the choice of a particular method for conducting PSHA has relatively little effect on the hazard estimates. The Cornell-McGuire method yielded the highest estimates, with the two other methods producing slightly lower estimates. The variation among the results based on the three methods appeared to be virtually independent of the return period. The variation in the results for the Sochi region was within 6%, and that for the Kamchatka region was within 10%. Accordingly, the considered PSHA methods would provide closely related results for areas of moderate seismic activity; however, the difference among the results would apparently increase with increasing seismic activity.

5.2 Introduction

The PSHA methodology allows estimation of the probability that various ground motion levels would be exceeded at a particular site during a specified time interval. Such analysis should precede the construction of infrastructure facilities in seismically active regions. At present, the Cornell-McGuire method (Cornell, 1968; McGuire, 1976; 1978) is applied most frequently for PSHA. This method incorporates information on seismic source zones (in the form of active faults or areal sources), FMDs (e.g. the GR relation), and GMPEs to estimate seismic hazard at a particular site. In this method, considerable attention is paid to the problem of accounting for various uncertainties (Budnitz et al., 1997), by using probability distributions (aleatory variability) and logic trees (epistemic uncertainties).

The initial step of the Cornell-McGuire method requires delineating the seismic sources, which are characterised by uniform spatial distribution of seismicity and homogeneous seismic parameters. Over time, it has become clear that the uniform distribution in many instances does not reflect the actual spatial distribution of epicentres (e.g. Wiemer et al., 2009; Spada et al., 2011). Moreover, the process of defining the source zones can be difficult and subjective, potentially leading to significant differences in the resulting source geometries prepared by different groups of experts (e.g. McGuire, 1993; Frankel, 1995; Budnitz et al., 1997). In addition, the estimation of the seismic parameters in areas of relatively low seismicity presents a substantial problem.

Such difficulties have stimulated the development of alternative methods that do not require the defini-

tion of the seismic source zones. These methods include, e.g. the techniques of Milne and Davenport (1969) and Veneziano et al. (1984) that are based entirely on the information from the seismic event catalogues, the methods of Frankel (1995) and Woo (1996) that use the spatial smoothing of seismicity, and the method of Kijko and Graham (1998; 1999) that combines the strong features of the previous techniques. In addition, there are PSHA procedures based on the Monte Carlo simulations (e.g. Ebel and Kafka, 1999; Musson, 2000; Shumilina et al., 2000; Assatourians and Atkinson, 2013).

In view of these different PSHA methods, the question is how the hazard estimates resulting from these different methods corresponded. In large countries, such as Russia, where seismogenic provinces differ significantly, different groups of experts often use different methods to analyse the seismic hazard in their regions. In such instances, it is important to know in what way the results of these analyses corresponded to each other.

Several studies have been devoted to this question, such as those by Molina et al. (2001), Beauval et al. (2006), Hong et al. (2006), and Goda et al. (2013). These studies were primarily focused on investigating the influences of different seismicity models on the estimated seismic hazard.

Molina et al. (2001) and Beauval et al. (2006) have compared the hazard estimates obtained with the conventional zoning approach - the basis of the Cornell-McGuire method - with those obtained by using the Kernel Smoothing method of Woo (1996). Using synthetic earthquake catalogues, Hong et al. (2006) conducted a comparison of seismic hazard estimates based on the Cornell-McGuire method, the Davenport-Milne method, and the Epicentral Cell method (an extension of the latter). Goda et al. (2013) used synthetic earthquake data to evaluate the effects of different smoothing approaches by employing the Cornell-McGuire, the Kernel Smoothing, and the Epicentral Cell methods.

The findings of these studies have demonstrated, among other things, that the assumption of a homogeneous activity rate within a seismic source zone is a poor representation of the true activity rate (Molina et al., 2001). In addition, it was shown that the hazard estimates based on the Cornell-McGuire method are generally higher than are those based on other methods.

In addition to the differences of the applied seismic source models, the existing PSHA methods use different procedures to estimate the rates of exceedance of ground motion. In contrast with previous comparative studies, the current work is focused on investigating the effect of the procedures used in the different PSHA methods to estimate the exceedance rates. Based on the same seismic source model, three major PSHA methodologies are compared, namely, the conventional Cornell-McGuire method, the Parametric-Historic method, and the method based on Monte Carlo simulations.

5.3 Materials and Methods

5.3.1 Earthquake catalogue

The main earthquake catalogue used in this study represents the entire territory of Eurasia, and the timespan is from ancient times to the end of 2011. These data were generously provided by Dr Nina Medvedeva from the Schmidt Institute of Physics of the Earth of the Russian Academy of Sciences (http://ifz.ru). All dependent seismic events were removed during the preparation of the catalogue. The catalogue lists the date, epicentral coordinates, focal depth, and the surface-wave magnitude M_s of each event, and, for some events, the focal intensity and the azimuth of rupture propagation are available.

Two areas in Russia have been selected for investigation in this study, namely, the city of Sochi on the Black Sea coast and surrounding area, and the Kamchatka Peninsula in far eastern Russia. Both these areas are characterised by high seismic activity, with the seismicity of Sochi and surroundings being characterised by crustal earthquakes and that of Kamchatka by subduction earthquakes. Moreover, the characteristics of the radiation and propagation of seismic waves in these two areas differ substantially (Pavlenko, 2011).

More recent (2012-2016) seismic data for the selected areas were obtained from the United States Geological Survey (USGS) website (http://earthquake.usgs.gov/earthquakes/search/). The USGS data mostly provide the body-wave magnitude m_b. The space-time windowing algorithm of Gardner and Knopoff (1974) was applied to these data to remove clusters and dependent seismic events. To obtain a homogeneous catalogue, magnitudes were converted to the moment magnitude scale M_w. Conversions were performed, based on the regional relations, if available; otherwise, global relations were applied. Therefore, small values of m_b in the regional catalogue for the Sochi area were converted by using the relation of Gasperini et al. (2013), which is based on the European-Mediterranean dataset. Conversions of both m_b and M_s for the Kamchatka Peninsula were performed by using the regional relations obtained by Gusev (1991). These relations were presented as a discrete set of points, therefore simple polynomial interpolation was applied to obtain continuous relations. Outside of the applicability ranges of regional relations, global relations were used. The adopted conversion scheme is shown in Table 5.1.

The resulting catalogue for the Sochi area contains 3958 earthquakes, with M_w from 3.0 to 8.1; whereas, that for the Kamchatka Peninsula contains 10389 earthquakes, with M_w from 3.6 to 9.0. The epicentres of these earthquakes are shown in Figs. 5.1 and 5.2. By using the cumulative plots of the number of seismic events as a function of time, the regional catalogues have been divided into sections, i.e. those containing pre-instrumental historical records and those containing complete instrumental earthquake data. For each instrumental sub-catalogue, the magnitude of completeness M_c was estimated based on the procedure proposed by Amorèse (2007).

Sochi		
$m_b < 5.0$:	$M_w = \exp(1.19 + 0.16 m_b) - 2.26$	(Gasperini et al., 2013)
$m_b \geq 5.0, H \geq 70 km$:	$M_w = 0.165 m_b^2 - 0.372 m_b + 2.816$	(Tsampas et al., 2016)
$H < 70 km$:	$M_w = 1.64 m_b - 3.18$	(Das et al., 2011)
$H < 70 km, M_s \in [3.0, 6.1]$:	$M_w = 0.67 M_s + 2.12$	(Das et al., 2011)
$M_s \in [6.2, 8.4]$:	$M_w = 1.06 M_s - 0.38$	(Das et al., 2011)
$H \geq 70 km, M_s \in [3.3, 4.3]$:	$M_w = 0.67 M_s + 2.33$	(Das et al., 2011)
$M_s \in [4.4, 7.7]$:	$M_w = -0.006 M_s^2 + 0.850 M_s + 1.540$	(Tsampas et al., 2016)
Kamchatka		
$m_b < 4.5$:	$M_w = \exp(-0.60 + 0.34 m_b) + 2.15$	(Gasperini et al., 2013)
$m_b \in [4.5, 6.3]$:	Interpolation	(Gusev, 1991)
$M_s < 3.9, H \geq 70 km$:	$M_w = 0.67 M_s + 2.33$	(Das et al., 2011)
$H < 70 km$:	$M_w = 0.67 M_s + 2.12$	(Das et al., 2011)
$M_s \geq 3.9$:	Interpolation	(Gusev, 1991)

Table 5.1: Applied magnitude conversions

5.3.2 Seismic source model

The seismic source model used in this study resembles the approach proposed by Frankel (1995) for mapping seismic hazard in the United States (e.g. Frankel et al., 2002; Petersen et al., 2014). The area of study is covered by a regular grid of points forming a square of cells. Each cell is treated as a point seismic source and earthquakes are assumed to occur as a stationary Poisson process with constant rate Λ.

The FMD describes the relation between the frequency of occurrence and the magnitude of earthquakes. The most common FMD in seismic hazard studies is the GR relation (Gutenberg and Richter, 1944):

$$\log_{10} N(M) = a - bM \tag{5.1}$$

where $N(M)$ is the number of seismic events with magnitude equal to or greater than M, a is a measure of the level of seismic activity, and b is the slope of the recurrence curve. If the magnitudes of the seismic events

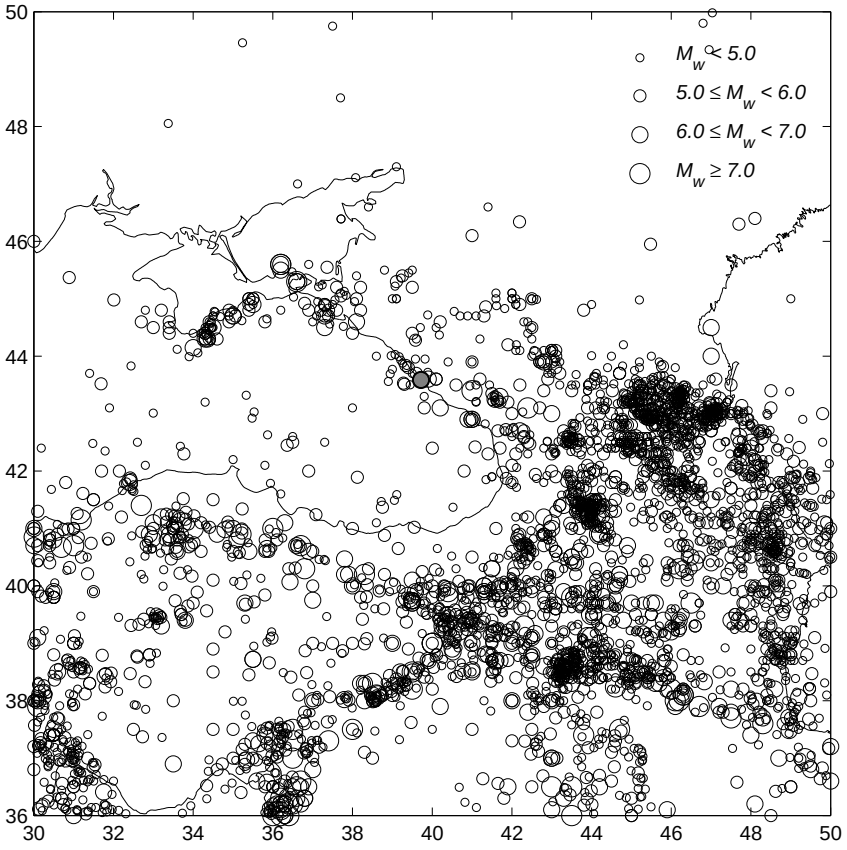


Figure 5.1: Location of epicentres in area surrounding Sochi

are assumed to be independent identically distributed random variables, and the magnitude range is bounded from the top, then, the distribution of magnitude has the form of a truncated exponential distribution, with the following CDF:

$$F_M(m) = \frac{1-\exp\left[-\beta(m-M_c)\right]}{1-\exp\left[-\beta(M_{max}-M_c)\right]}, \quad M_c \le m \le M_{max} \tag{5.2}$$

where $\beta = b\ln(10)$, and M_{max} is a magnitude of the strongest possible earthquake in the area. The uncertainty of parameter β can be handled by applying the compound (Bayesian) distribution (DeGroot, 1970; Hamada et al., 2008; Klugman et al., 2008). In general, if the random variable M has a CDF $F_M(m,\theta)$, with vector of parameters θ, and $f_\theta(\theta)$ denotes the PDF of θ, the compound CDF would be calculated as follows:

$$F_M(m) = \int_{\Omega_\theta} F_M(m,\theta)\, f_\theta(\theta)\, d\theta \tag{5.3}$$

Assuming that the variation of β can be modelled by a gamma distribution, the following compound CDF can be obtained (Campbell, 1982; Kijko and Graham, 1998; Kijko et al., 2016):

$$F_M(m) = C_\beta \left[1 - \left(\frac{p}{p+m-M_c}\right)^q\right], \quad M_c \le m \le M_{max} \tag{5.4}$$

where $p = \frac{\bar{\beta}}{\sigma_{\bar{\beta}}^2}$, $q = \left(\frac{\bar{\beta}}{\sigma_\beta}\right)^2$, $\bar{\beta}$ is the mean value of β and σ_β is its standard deviation, and C_β is a normalising

constant given by:

$$C_\beta = \left[1 - \left(\frac{p}{p + M_{\max} - M_c}\right)^q\right]^{-1} \tag{5.5}$$

The corresponding compound PDF is expressed as:

$$f_M(m) = \bar{\beta} C_\beta \left(\frac{p}{p + m - M_c}\right)^{q+1}, \quad M_c \le m \le M_{\max} \tag{5.6}$$

Thereby, the set of required parameters for each seismic source consists of Λ, β, and $M_{\max}$. These parameters were estimated in the following manner. First, $M_{\max}$ was estimated based on a few largest observed magnitudes (Kijko and Singh, 2011):

$$\hat{M}_{\max} = M_{\max}^{obs} + \frac{1}{n_0}\left(M_{\max}^{obs} - \frac{1}{n_0 - 1}\sum_{i=2}^{n_0} M_{n-i+1}\right) \tag{5.7}$$

where $M_{\max}^{obs}$ is the largest observed earthquake magnitude, $n_0 = 10$, and M_{n-i+1} is the $(n-i)$-th largest observation.

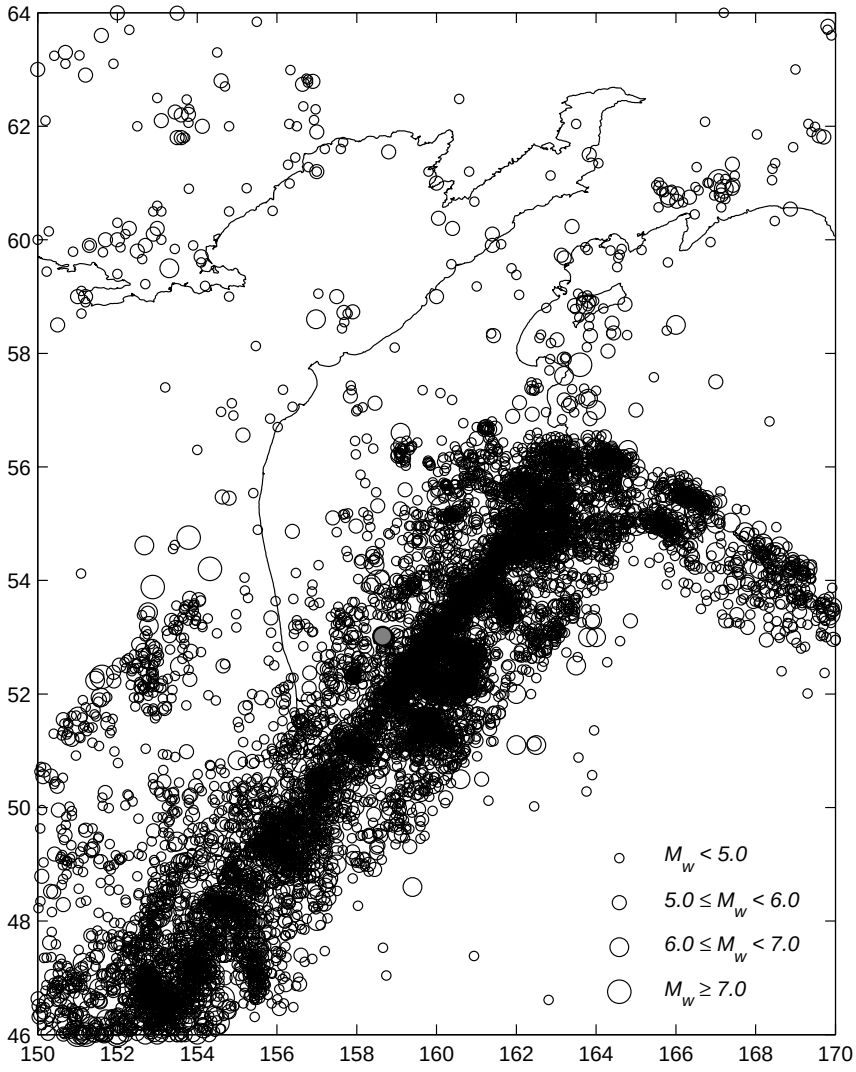


Figure 5.2: Location of epicentres at the Kamchatka Peninsula

Second, Λ and β were estimated by maximising the joint likelihood function, as described by Kijko and Sellevoll (1989; 1992). This procedure allows using the information from the whole seismic catalogue (i.e. pre-instrumental data and complete data), and accounting for uncertainties associated with the FMD and magnitude determination.

5.3.3 Hazard calculations

The seismic hazard at a particular site is characterised by the ground motion that has a specified probability to be exceeded at least once during the specified period of time. The assumption that the occurrence of earthquakes conforms to a stationary Poisson process allows the calculation of the probability that ground motion parameter y would exceed the value a_0 at the site at least once during time interval T:

$$P[y \geq a_0, T] = 1 - e^{-\lambda(a_0)T} \tag{5.8}$$

where $\lambda(a_0)$ is the annual rate of exceedance of ground motion level a_0 at the site.

The estimation procedure for the exceedance rate $\lambda(a_0)$ depends on the applied PSHA methodology. For the Cornell-McGuire method, this value is estimated by using the following equation, based on the total probability theorem:

$$\lambda(a_0) = \sum_{i}^{N} \Lambda_i \int_R \int_M P[y \geq a_0 | m, r] f_{M_i}(m) f_{R_i}(r)\, dr\, dm \tag{5.9}$$

where the summation is taken over all seismic sources capable of inducing significant ground motion at the site, $f_{R_i}(r)$ is the PDF of distance R, $f_{M_i}(m)$ is the PDF of magnitude M given in eq. (5.6), $M_{\min}$ is the smallest magnitude considered in the analysis (in this study, $M_{\min} = 4.0$), the conditional probability of exceedance $P[y \geq a_0 | m, r]$ reflects the inherent variability of ground motion y for given magnitude m and distance r, usually calculated by using the normal (Gaussian) distribution.

In this study, as the seismic sources are modelled as a regular grid of point sources, eq. (5.9) reduces to the following:

$$\lambda(a_0) = \sum_{i}^{N} \Lambda_i \int_{M_{\min}}^{M^i_{\max}} P[y \geq a_0 | m, r] f_{M_i}(m) dm \tag{5.10}$$

By increasing the value a_0 and repeating these calculations, the seismic hazard curve is constructed.

The Parametric-Historic method is oriented more empirically. Only the sources that induce ground motions y in excess of a fixed lower threshold $a_{\min}$ are taken into account. The magnitude range is sub-divided into small intervals Δm and for each source the cumulative rate of exceedance is calculated by summing the incremental rates. The total annual rate of exceedance $\lambda(a_0)$ is calculated by summation over all contributing sources:

$$\lambda(a_0) = P[y \geq a_0] \sum_{i}^{N_s} \sum_{j}^{N_m} \Lambda_i (a \geq a_{\min}) \int_{m_j - \Delta m/2}^{m_j + \Delta m/2} f_{M_i}(m)\, dm \tag{5.11}$$

where N_s is the number of seismic sources that induce ground motions y in excess of the fixed lower threshold $a_{\min}$, N_m is the number of intervals Δm between $M_{\min}$ and $M^i_{\max}$, and $P[y \geq a_0]$ is the probability of exceedance, estimated from the empirical distribution of y at a site of interest.

The third considered PSHA procedure is based on a synthetic catalogue of seismic events. The synthetic catalogue is generated by the Monte Carlo simulation technique, based on the estimated seismic

parameters Λ, β, and $M_{\max}$. The duration of the synthetic catalogue T_c should be sufficient to allow for reliable hazard estimation. In this study, two probability levels are considered, namely, 10% and 2% probabilities of exceedance in 50 years, which correspond to return periods of approximately 475 and 2475 years, respectively. In this study, the duration of the synthetic catalogues was equal to a hundred times the return period.

The estimation of the annual rate of exceedance $\lambda(a_0)$ from the synthetic catalogue is straightforward. The ground motion parameter y is calculated for each event of the synthetic catalogue and a cumulative histogram of y is calculated. Subsequently, the cumulative histogram is normalised by the timespan of the synthetic catalogue T_c, and the required annual rate of exceedance $\lambda(a_0)$ is obtained from the normalised histogram.

5.3.4 Selection of the GMPEs

Estimates of the expected ground motion at the site of interest are fundamental factors in seismic hazard analysis. The ground motion is characterised by a particular parameter, usually a horizontal PGA, PGV, or spectral acceleration. GMPEs are employed to estimate ground motion parameters for use in both the deterministic and the probabilistic seismic hazard analyses. These equations allow estimation of the median values of the ground motion parameters, based on the earthquake magnitude, source to site distance, local soil conditions, fault mechanism, and other parameters. The GMPEs are usually empirical equations obtained by means of regression analysis (e.g. Joyner and Boore, 1993). A large variety of GMPEs for different parts of the world have been developed over the years (Douglas, 2011).

As regards the regions selected for this study, the main characteristics of the radiation and propagation of seismic waves at the North Caucasus region and in the vicinity of Sochi have been studied by Pavlenko (2008; 2009; 2016) and Pavlenko and Pavlenko (2016), but a regional GMPE has not been developed yet. Despite continuous studies on strong ground motions at Kamchatka (Gusev et al., 1997; Petukhin et al., 1999; Chubarova et al., 2010), a reliable GMPE for this region has not been established yet. In such instances, a common practice is to adopt the GMPEs developed for other regions with similar tectonic properties (e.g. Stafford et al., 2008; Delavaud et al., 2009). Since the purpose of the current study is to compare the different methods of PSHA rather than to assess the seismic hazard itself, the adoption of particular GMPEs should not affect the results radically. Therefore, in this study, a set of GMPEs recommended by the Global Earthquake Model (GEM, http://www.globalquakemodel.org) project is used. These GMPEs are listed in Table 5.2.

Reference	Scaling parameters	Sochi	Kamchatka	Weight
Akkar and Bommer (2007)	M_w, R_{JB}	✓	✓	0.2
Boore and Atkinson (2008)	M_w, R_{JB}	✓	✓	0.2
Campbell and Bozorgnia (2008)	M_w, R_{rup}	✓	✓	0.2
Cauzzi and Faccioli (2008)	M_w, R_{hyp}	✓	✓	0.2
Chiou and Youngs (2008)	M_w, R_{rup}	✓	✓	0.2
Youngs et al. (1997)	M_w, R_{rup}	–	✓	0.34
Atkinson and Boore (2003)	M_w, R_{rup}	–	✓	0.33
Kanno et al. (2006)	M_w, R_{rup}	–	✓	0.33

Table 5.2: The GMPEs recommended by the GEM

The GMPEs of Akkar and Bommer (2007) and Cauzzi and Faccioli (2008) have been developed for implementation in the European region. Akkar and Bommer (2007) used strong motion data from Europe and the Middle East for their study, while Cauzzi and Faccioli (2008) compiled the database by including the strong motion records from Japan, Iran, California, Turkey, Iceland, and Italy. The GMPEs of Boore

and Atkinson (2008), Campbell and Bozorgnia (2008), and Chiou and Youngs (2008) have been developed as contributions to the Next Generation Attenuation project of the Pacific Earthquake Engineering Research Center (peer.berkeley.edu), and are considered globally applicable to shallow crustal earthquakes in active tectonic regions.

Youngs et al. (1997) and Atkinson and Boore (2003) have developed globally applicable GMPEs for subduction zone earthquakes, based on the global strong motion databases, whereas, the GMPE of Kanno et al. (2006) has been developed by using the data from Japan, California, and Turkey. Youngs et al. (1997) and Atkinson and Boore (2003) categorised earthquakes as interface events (shallow-angle thrust events that occur on the interface between the subducting and the overriding plates, usually not deeper than 50 km) and intraslab events (events that occur within the subducting oceanic plate and which are typically high-angle normal faulting events). Kanno et al. (2006) distinguished between shallow and deep earthquakes.

The seismic hazard in Sochi and the surrounding region is partially attributable to the proximity of the area to the Caucasus Mountains, a part of the Iran-Caucasus-Anatolia seismic region, characterised by high seismic activity. Furthermore, numerous strong and moderate earthquakes have been reported in the Sochi area and in other parts of the Black Sea coastal area, as well as in the Black Sea itself.

In comparison with the seismicity of the Sochi region, that of the Kamchatka Peninsula represents a greater challenge. The seismicity of the Kamchatka Peninsula is characterised by subduction earthquakes on the south eastern coast, where seismicity is dominated by the events occurring at the Kuril-Kamchatka Trench. At the north end of the peninsula, seismicity is characterised by less frequent crustal earthquakes. To the west of the peninsula, in the Sea of Okhotsk, several large deep-focus earthquakes have occurred in recent times, including the great M_w 8.3 earthquake of May 24, 2013 (Chebrova et al., 2015).

Along the south eastern coast, in area that extends deep into the peninsula, the ground motions could be caused by either the subduction earthquakes occurring on the dipping Pacific Plate beneath the peninsula, or by the crustal earthquakes. This area was modelled as a transition zone of mixed seismicity, where both types of earthquakes contribute to strong ground motions. The structure of the Kamchatka subduction zone was explored in detail by Gorbatov et al. (1997). The ground motions in the transition zone were estimated as the weighted average of the outputs of two GMPEs, one developed for crustal seismicity, and the second developed for subduction earthquakes:

$$\begin{cases} y = p_c y_c + p_s y_s \\ \sigma = \sqrt{p_c^2 \sigma_c^2 + p_s^2 \sigma_s^2} \end{cases} \tag{5.12}$$

where p represent normalised weights that reflect the relative probability for ground motion to be induced by a subduction or crustal event, y is the median ground motion value, and σ is the corresponding standard deviation, subscripts c and s mean crustal and subduction.

Epistemic uncertainty was handled by using the logic tree formalism (e.g. Bommer et al., 2005), and a set of alternative hazard curves was calculated by using the GMPEs, from which the mean hazard curve was selected to characterise seismic hazard. The weights of the set of GMPEs are listed in Table 5.2.

5.4 Results and discussion

Figures 5.3 and 5.4 show the seismic hazard maps for PGA for the two considered areas. In these figures, the hazard maps for exceedance probability of 10% in 50 years (return period of 475 years) are shown in the upper row; whereas, in the lower row, the maps are shown for exceedance probability of 2% in 50 years (return period of 2475 years). In the computation of seismic hazard, the grid size was set to 0.1°for the Sochi area and to 0.2°for the Kamchatka area.

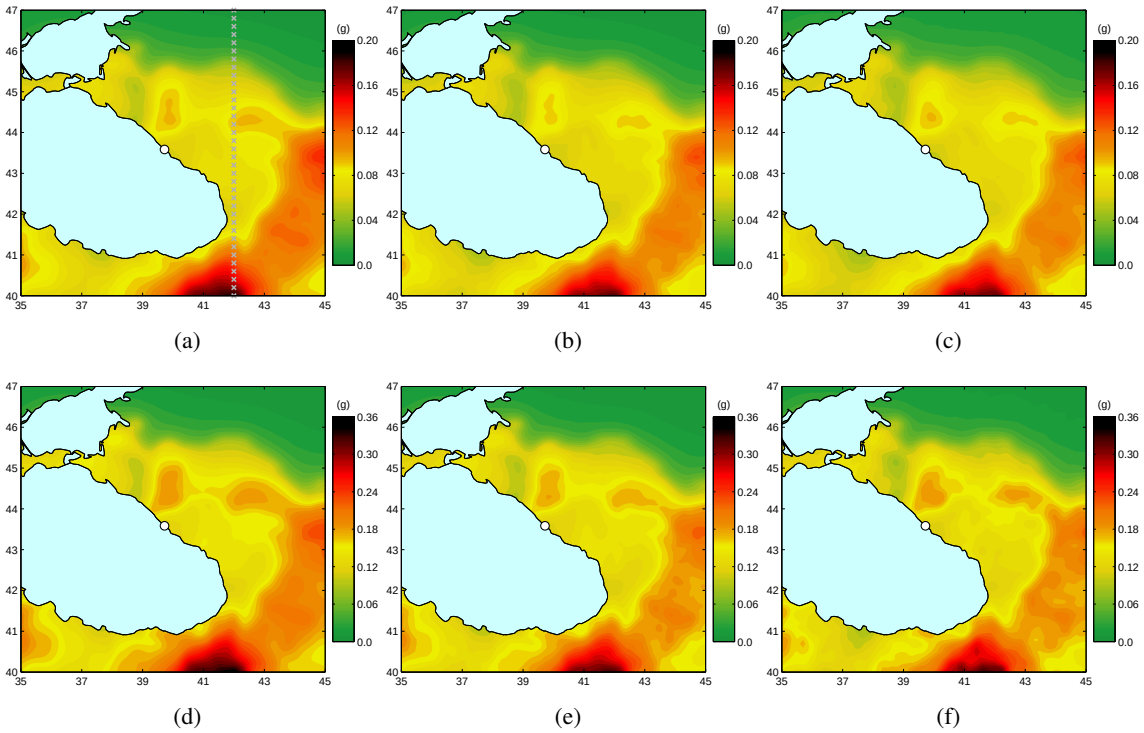


Figure 5.3: Comparison of seismic hazard maps for Sochi and surrounding area. Left to right: the Cornell-McGuire method, the Parametric-Historic method, and the method based on the Monte Carlo technique. Upper row: T_R = 475 years, lower row: T_R = 2475 years. White circle shows location of Sochi. Profile 1 is shown by grey crosses in Fig. 5.3(a)

The shape of the hazard contours reflects the observed regional seismicity, with the higher hazard being concentrated near the epicentres of the major seismic events. When the return period is increased from 475 years to 2475 years, the seismic hazard estimates increase by a factor of nearly 1.9, while the shape of the contours remains unchanged.

It could be interesting to subject these maps to a test to indicate which of the methods provides the most realistic result. Objective testing of the PSHA results is a significant problem, which has been discussed by several authors (e.g. McGuire, 1979; Ward, 1995; Ordaz and Reyes, 1999; Beauval et al., 2008; Stirling and Gerstenberger, 2010; Stein et al., 2011; Kossobokov and Nekrasova, 2012; Stirling, 2012; Wyss et al., 2012; Stein et al., 2012; Mezcua et al., 2013; Stein et al., 2015), with various approaches being proposed.

There are two main categories of tests, of which the first relates to testing the modelled rate of exceedance against the observed number of ground motion exceedances (e.g. Ordaz and Reyes, 1999; Stirling and Gerstenberger, 2010; Mezcua et al., 2013). The second category of tests relies on comparison between the modelled and the observed levels of ground motion (e.g. Ward, 1995; Miyazawa and Mori, 2009; Kossobokov and Nekrasova, 2012).

However, the question of the adequacy of the PSHA methods is beyond the scope of the current study. Moreover, the testing of seismic hazard estimates is a relatively new and debatable aspect of PSHA (e.g. Stein et al., 2011; Hanks et al., 2012; Stirling, 2012), and no consensus has been reached on how such testing should be performed. Consequently, the current analysis is restricted to a quantitative comparison of the obtained PSHA maps.

The hazard estimates were compared at the sites along two profiles, passing through the highest, moderate, and the lowest hazard areas of the maps, as shown in Figs. 5.3(a) and 5.4(a). Figures 5.5 and 5.6 and show the levels of PGA along these profiles.

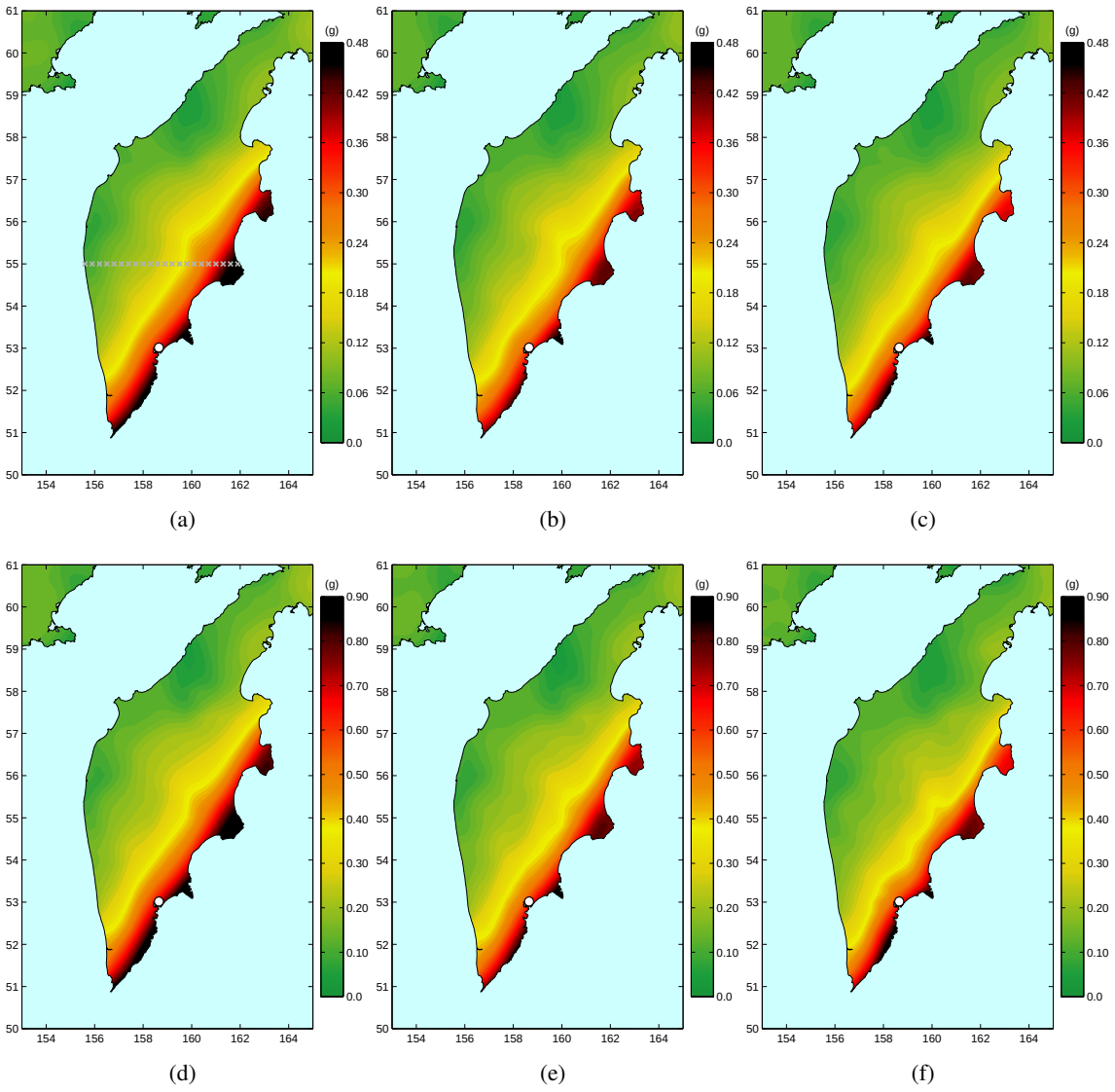


Figure 5.4: Comparison of seismic hazard maps for Kamchatka. Left to right: the Cornell-McGuire method, the Parametric-Historic method, and the method based on the Monte Carlo technique. Upper row: T_R = 475 years, lower row: T_R = 2475 years. White circle shows location of Petropavlovsk-Kamchatskiy. Profile 2 is shown by grey crosses in Fig. 5.4(a)

The trends observed along both profiles are similar, namely, the Cornell-McGuire method yields the highest hazard estimates; whereas, those obtained with the Parametric-Historic method and the method based on the Monte Carlo technique are slightly lower. The ratios of the PGA estimates along the two profiles were calculated for more explicit comparison (Table 5.3).

The relative difference between the PGA estimates along profile 1 slightly increases with the increasing return period. On average, the relative difference between the PGA estimates based on the Parametric-Historic and the Cornell-McGuire methods is about 5% for both return periods; whereas, that of the method based on the Monte Carlo technique on average 6% below the estimates of the Cornell-McGuire method for both return periods. Along profile 2, the relative difference between the PGA estimates decreases slightly as the hazard level decreases (Table 5.3), and the ratios are similar for both return periods. The PGA estimates of the Parametric-Historic method are on average 8% below the estimates of the Cornell-McGuire method; whereas, those of the method based on the Monte Carlo technique are on average 10% below the estimates

of the Cornell-McGuire method. Judging by the averaged values of the ratios, the variation among the results based on the three methods is relatively low, but has a systematic nature.

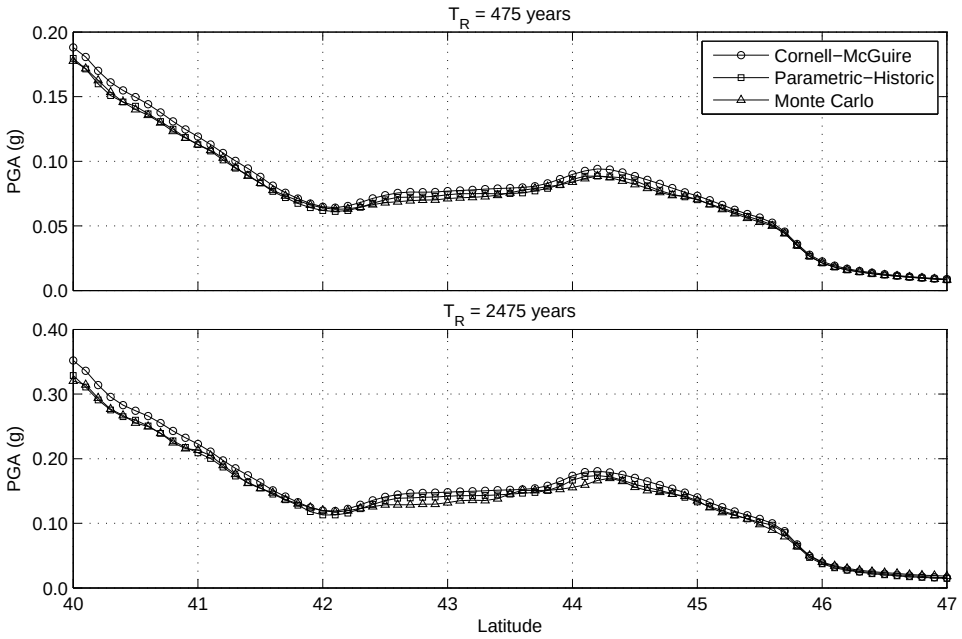


Figure 5.5: Hazard levels along profile 1

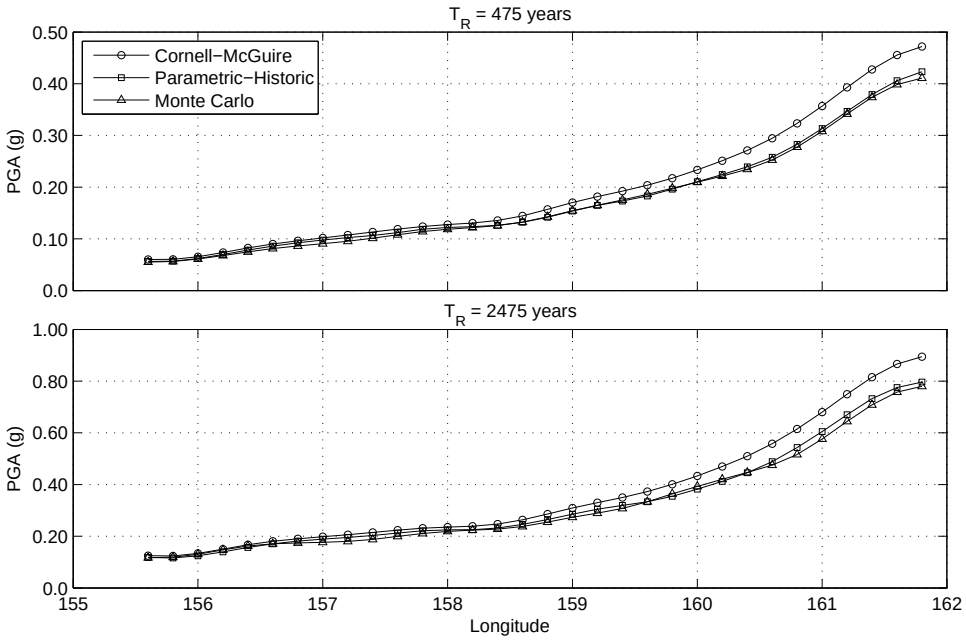


Figure 5.6: Hazard levels along profile 2

	Profile 1				Profile 2			
T_R = 475 years	$A_{\max}(g)$	$R_{\max}$	$< R >$	$R_{\min}$	$A_{\max}(g)$	$R_{\max}$	$< R >$	$R_{\min}$
Cornell-McGuire	0.188				0.472			
Parametric-Historic	0.180	0.955	0.952	0.935	0.423	0.896	0.919	0.940
Monte Carlo	0.178	0.944	0.943	0.935	0.411	0.871	0.900	0.919
T_R = 2475 years								
Cornell-McGuire	0.352				0.894			
Parametric-Historic	0.329	0.934	0.948	0.921	0.797	0.891	0.919	0.943
Monte Carlo	0.320	0.910	0.937	0.920	0.780	0.872	0.902	0.924

$A_{\max}$ - maximum PGA estimate, $R_{\max} = (A_{\max}/A^{CM}_{\max})$, $< R >=< (A/A^{CM}) >$, $R_{\min} = (A_{\min}/A^{CM}_{\min})$
A^{CM} - estimates of the Cornell-McGuire method, A - estimates of the two other methods

Table 5.3: Comparison of seismic hazard estimates along two profiles

5.5 Conclusion

In this study, seismic hazard estimates obtained by using three different PSHA methods were compared. For comparison, the seismic hazard maps were prepared for two regions of Russia, and the PGA estimates were compared for return periods of 475 and 2475 years. The results indicated that the choice of a particular method for conducting PSHA has relatively little effect on the hazard estimates when the same seismic source model was used in the calculations. The comparison indicated that the Cornell-McGuire method systematically yielded the highest estimates of PGA, whereas the Parametric-Historic method and the method based on the Monte Carlo technique produced similar results, which were slightly below that of the Cornell-McGuire method. The analysis for the two regions considered indicated that the relative difference between the results of the three methods was systematic, remaining virtually unchanged when the return period increased from 475 to 2475 years. For the Sochi region, characterised by high seismic activity, this difference was within 6%; whereas, for the Kamchatka region, where seismic activity is very high, the difference was up to 10%. These results suggest that for regions of moderate seismic activity, all three methods would provide closely related seismic hazard estimates. However, difference among the results would apparently become more pronounced for regions characterised by high seismic activity.

Acknowledgements We are grateful to Dr. Nina Medvedeva from the Schmidt Institute of Physics of the Earth of the Russian Academy of Sciences for provided earthquake catalogue, and to the USGS for their data.

Chapter 6

Concluding remarks

In this thesis several important methodological problems of the modern PSHA were considered.

In Chapter 1, an overview of the current practice in seismic hazard analysis was presented, and several disputable aspects of the most widely applied PSHA method were identified.

In Chapter 2, the distribution of the residuals of the logarithm of PGA was considered and the procedure that is suitable for studying this distribution was proposed. The proposed procedure does not require a priori assumptions and allows selecting the most plausible probability distribution based on the available data. It was shown that the hazard curves calculated by using different probability distributions can differ significantly from one another, especially at very low annual probabilities of exceedance. In addition, the results of the analysis indicated that the best approximation for the distribution of residuals was obtained with the GEVD.

In Chapter 3, a similar analysis was performed by using a large dataset of PGA recorded in Japan. This analysis demonstrated that the variability of the logarithm of PGA is more precisely modelled by the GEVD with negative shape parameter. This distribution assumes a finite upper bound of a variable and, therefore, allows to account for finiteness of the ground motion induced by specific earthquake scenario. Therefore, the approach based on using the GEVD eliminates the need for truncating the ground motion variability. It provides a rational way of estimating the maximum value of PGA induced by a specific earthquake scenario. This approach is preferred to the truncation of a distribution, because a maximum value of PGA, unlike the truncation of a distribution, has a clear physical meaning.

In Chapter 4, a set of equations was proposed that allows accounting for anisotropy in the attenuation of seismic intensity. The results of the application of the proposed equations were compared with the results obtained by using the isotropic attenuation law. The results of this study demonstrated that anisotropic attenuation equations allow improving agreement with the results of the observations, especially when modelling significantly anisotropic isoseismal maps.

In Chapter 5, seismic hazard estimates obtained by using three different PSHA methods were compared. Seismic hazard maps were prepared for two regions of Russia and PGA estimates were compared for return periods of 475 and 2475 years. The analysis for the two regions considered indicated that the relative difference between the results of the three methods was systematic, remaining virtually unchanged when the return period increased from 475 to 2475 years. The results suggested that for regions of moderate seismic activity, all three methods would provide closely related seismic hazard estimates. However, difference among the results would apparently become more pronounced for regions characterised by high seismic activity.

The results of this thesis contribute to obtaining more realistic models for PSHA.

List of References

Abrahamson, N. A. (1988). Statistical properties of peak ground accelerations recorded by the SMART 1 array. *Bull. Seismol. Soc. Am. 78*(1), 26–41.

Abrahamson, N. A. (2000). State of the Practice of Seismic Hazard Evaluation. In *Proceedings of GeoEng 2000, International Conference on Geotechnical & Geological Engineering*, Volume 1, Melbourne, Australia, pp. 659–685.

Abrahamson, N. A., P. Birkhauser, M. Koller, D. Mayer-Rosa, P. Smit, C. Sprecher, S. Tinic, and R. Graf (2002). PEGASOS - A Comprehensive Probabilistic Seismic Hazard Assessment for Nuclear Power Plants in Switzerland. In *Proceedings of the 12th European Conference on Earthquake Engineering*, London, UK. Paper No. 633.

Abrahamson, N. A. and J. J. Bommer (2005). Probability and Uncertainty in Seismic Hazard Analysis. *Earthquake Spectra 21*(2), 603–607. doi: 10.1193/1.1899158.

Abrahamson, N. A. and R. R. Youngs (1992). A stable algorithm for regression analyses using the random effects model. *Bull. Seismol. Soc. Am. 82*(1), 505–510.

Akaike, H. (1971). On entropy maximization principle. In P. R. Krishnaiah (Ed.), *Applications of Statistics*, Amsterdam, pp. 27–41. North-Holland.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr. 19*(6), 716–723. doi: 10.1109/TAC.1974.1100705.

Aki, K. (1965). Maximum Likelihood Estimate of b in the Formula $\log N = a - bM$ and its Confidence Limits. *Bull. Earthq. Res. Inst., Univ. Tokyo 43*, 237–239.

Akkar, S. and J. J. Bommer (2007). Prediction of elastic displacement response spectra in Europe and the Middle East. *Earthq. Eng. Struct. Dyn. 36*, 1275–1301. doi: 10.1002/eqe.679.

Albarello, D., R. Camassi, and A. Rebez (2001). Detection of Space and Time Heterogeneity in the Completeness of a Seismic Catalog by a Statistical Approach: An Application to the Italian Area. *Bull. Seismol. Soc. Am. 91*(6), 1694–1703. doi: 10.1785/0120000058.

Allen, T. I., D. J. Wald, and C. B. Worden (2012). Intensity attenuation for active crustal regions. *J. Seismol. 16*(3), 409–433. doi: 10.1007/s10950-012-9278-7.

Ambraseys, N. (1974). The correlation of intensity with ground motions. In *Proceedings of the 14th Conference European Seismological Commission*, Volume 1, Trieste, pp. 335–341.

Amitrano, D. (2003). Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the b value. *J. Geophys. Res. 108*(B1). doi: 10.1029/2001JB000680.

Amorèse, D. (2007). Applying a Change-point Detection Method on Frequency-Magnitude Distributions. *Bull. Seismol. Soc. Am. 97*(5), 1742–1749. doi: 10.1785/0120060181.

Anderson, J. G. (1979). Estimating the seismicity from geological structure for seismic-risk studies. *Bull. Seismol. Soc. Am. 69*(1), 135–158.

Anderson, J. G. and J. N. Brune (1999). Probabilistic Seismic Hazard Analysis without the Ergodic Assumption. *Seismol. Res. Lett. 70*(1), 19–28. doi: 10.1785/gssrl.70.1.19.

Anderson, J. G. and J. E. Luco (1983). Consequences of slip rate constraints on earthquake occurrence relations. *Bull. Seismol. Soc. Am. 73*(2), 471–496.

Applied Technology Council (1985). *ATC-13: Earthquake Damage Evaluation Data for California*. Redwood City, California.

Arroyo, D., M. Ordaz, and R. Rueda (2014). On the Selection of Ground-Motion Prediction Equations for Probabilistic Seismic-Hazard Analysis. *Bull. Seismol. Soc. Am. 104*(4), 1860–1875. doi: 10.1785/0120130264.

Assatourians, K. and G. M. Atkinson (2013). EqHaz: An Open-Source Probabilistic Seismic-Hazard Code Based on the Monte Carlo Simulation Approach. *Seismol. Res. Lett. 84*(3), 516–524. doi: 10.1785/0220120102.

Atkinson, G. M. and D. M. Boore (2003). Empirical Ground-Motion Relations for Subduction-Zone Earthquakes and Their Application to Cascadia and Other Regions. *Bull. Seismol. Soc. Am. 93*(4), 1703–1729. doi: 10.1785/0120020156.

Atkinson, G. M. and S. I. Kaka (2007). Relationships between Felt Intensity and Instrumental Ground Motion in the Central United States and California. *Bull. Seismol. Soc. Am. 97*(2), 497–510. doi: 10.1785/0120060154.

Baker, J. W. (2008). An Introduction to Probabilistic Seismic Hazard Analysis (PSHA) version 1.3. Available at http://web.stanford.edu.

Bakun, W. H. (2000). Seismicity of California's North Coast. *Bull. Seismol. Soc. Am. 90*(4), 797–812. doi: 10.1785/0119990138.

Bakun, W. H. (2005). Magnitude and location of historical earthquakes in Japan and implications for the 1855 Ansei Edo earthquake. *J. Geophys. Res. 110*(B02304). doi: 10.1029/2004JB003329.

Bakun, W. H. (2006). MMI Attenuation and Historical Earthquakes in the Basin and Range Province of Western North America. *Bull. Seismol. Soc. Am. 96*(6), 2206–2220. doi: 10.1785/0120060045.

Bakun, W. H. and O. Scotti (2006). Regional intensity attenuation models for France and the estimation of magnitude and location of historical earthquakes. *Geophys. J. Int. 164*(3), 596–610. doi: 10.1111/j.1365-246X.2005.02808.x.

Bakun, W. H. and C. M. Wentworth (1997). Estimating earthquake location and magnitude from seismic intensity data. *Bull. Seismol. Soc. Am. 87*(6), 1502–1521.

Barrientos, S., E. Vera, P. Alvarado, and T. Monfret (2004). Crustal seismicity in central Chile. *J. S. Am. Earth Sci. 16*(8), 759–768. doi: 10.1016/j.jsames.2003.12.001.

Bazzurro, P. and C. A. Cornell (1999). Disaggregation of Seismic Hazard. *Bull. Seismol. Soc. Am. 89*(2), 501–520.

Beauval, C., P.-Y. Bard, S. Hainzl, and P. Guéguen (2008). Can strong-motion observations be used to constrain probabilistic seismic-hazard estimates? *Bull. Seismol. Soc. Am. 98*(2), 509–520. doi: 10.1785/0120070006.

Beauval, C., O. Scotti, and F. Bonilla (2006). The role of seismicity models in probabilistic seismic hazard estimation: comparison of a zoning and a smoothing approach. *Geophys. J. Int. 165*(2), 584–595. doi: 10.1111/j.1365-246X.2006.02945.x.

Beirlant, J., Y. Goedebeur, T. Teugels, J. Segerus, D. De Waal, and C. Ferro (2004). *Statistics of Extremes: Theory and Applications*. Chichester: John Wiley & Sons.

Bender, B. (1983). Maximum likelihood estimation of *b*-values for magnitude grouped data. *Bull. Seismol. Soc. Am. 73*(3), 831–851.

Bender, B. (1984). Incorporating acceleration variability into seismic hazard analysis. *Bull. Seismol. Soc. Am. 74*(4), 1451–1462.

Bengoubou-Valérius, M. and D. Gibert (2013). Bootstrap determination of the reliability of *b*-values: an assessment of statistical estimators with synthetic magnitude series. *Nat. Hazards 65*(1), 443–459. doi: 10.1007/s11069-012-0376-1.

Bindi, D., M. Massa, L. Luzi, G. Ameri, F. Pacor, R. Puglia, and P. Augliera (2014). Pan-European ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods up to 3.0 s using the RESORCE dataset. *Bull. Earthquake Eng. 12*(1), 391–430. doi: 10.1007/s10518-013-9525-5.

Bol'shev, L. N. and N. V. Smirnov (1965). *Tables of Mathematical Statistics*. Moscow: Fizmatigiz. Translated into English.

Bommer, J. J. (2002). Deterministic vs. probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy. *J. Earthquake Eng. 6*(Special Issue 1), 43–73. doi: 10.1080/13632460209350432.

Bommer, J. J. (2003). Uncertainty about the uncertainty in seismic hazard analysis. *Eng. Geol. 70*(1-2), 165–168. doi: 10.1016/S0013-7952(02)00278-8.

Bommer, J. J. (2012). Challenges of Building Logic Trees for Probabilistic Seismic Hazard Analysis. *Earthquake Spectra 28*(4), 1723–1735. doi: 10.1193/1.4000079.

Bommer, J. J. and N. A. Abrahamson (2006). Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates? *Bull. Seismol. Soc. Am. 96*(6), 1967–1977. doi: 10.1785/0120060043.

Bommer, J. J., N. A. Abrahamson, F. O. Strasser, A. Pecker, P.-Y. Bard, H. Bungum, F. Cotton, D. Fäh, F. Sabetta, F. Scherbaum, and J. Studer (2004). The Challenge of Defining Upper Bounds on Earthquake Ground Motions. *Seismol. Res. Lett. 75*(1), 82–95. doi: 10.1785/gssrl.75.1.82.

Bommer, J. J., F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta, and N. A. Abrahamson (2005). On the Use of Logic Trees for Ground-Motion Prediction Equations in Seismic-Hazard Analysis. *Bull. Seismol. Soc. Am. 95*(2), 377–389. doi: 10.1785/0120040073.

Bonilla, M. G., R. K. Mark, and J. J. Lienkaemper (1984). Statistical relations among earthquake magnitude, surface rupture length, and surface fault displacement. *Bull. Seismol. Soc. Am. 74*(6), 2379–2411.

Boore, D. M. (2003). Simulation of Ground Motion Using the Stochastic Method. *Pure Appl. Geophys. 160*(3), 635–676. doi: 10.1007/PL00012553.

Boore, D. M. and G. M. Atkinson (2008). Ground-Motion Prediction Equations for the Average Horizontal Component of PGA, PGV, and 5%-Damped PSA at Spectral Periods between 0.01 s and 10.0 s. *Earthquake Spectra 24*(1), 99–138. doi: 10.1193/1.2830434.

Boore, D. M. and W. B. Joyner (1982). The empirical prediction of ground motion. *Bull. Seismol. Soc. Am. 72*(6), S43–S60.

Boyd, O. S. (2012). Including Foreshocks and Aftershocks in Time Independent Probabilistic Seismic Hazard Analyses. *Bull. Seismol. Soc. Am. 102*(3), 909–917. doi: 10.1785/0120110008.

Budnitz, R. J., G. Apostolakis, D. M. Boore, L. S. Cluff, K. J. Coppersmith, C. A. Cornell, and P. A. Morris (1997). Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts. U.S. Nuclear Regulatory Commission Report NUREG/CR-6372.

Campbell, K. W. (1981). Near source attenuation of peak horizontal acceleration. *Bull. Seismol. Soc. Am. 71*(6), 2039–2070.

Campbell, K. W. (1982). Bayesian analysis of extreme earthquake occurrences. Part I. Probabilistic hazard model. *Bull. Seismol. Soc. Am. 72*(5), 1689–1705.

Campbell, K. W. and Y. Bozorgnia (2008). NGA Ground Motion Model for the Geometric Mean Horizontal Component of PGA, PGV, PGD and 5% Damped Linear Elastic Response Spectra for Periods Ranging from 0.01 to 10 s. *Earthquake Spectra 24*(1), 139–171. doi: 10.1193/1.2857546.

Cao, A. M. and S. S. Gao (2002). Temporal variation of seismic *b*-values beneath northeastern Japan island arc. *Geophys. Res. Lett. 29*(9), 48–1–48–3. doi: 10.1029/2001GL013775.

Carnahan, B., H. A. Luther, and J. O. Wilkes (1969). *Applied Numerical Methods*. New York: John Wiley and Sons.

Castaños, H. and C. Lomnitz (2002). PSHA: is it science? *Eng. Geol. 66*(3-4), 315–317. doi: 10.1016/S0013-7952(02)00039-X.

Castellaro, S., F. Mulargia, and Y. Y. Kagan (2006). Regression problems for magnitudes. *Geophys. J. Int. 165*(3), 913–930. doi: 10.1111/j.1365-246X.2006.02955.x.

Cauzzi, C. and E. Faccioli (2008). Broadband (0.05 to 20 s) prediction of displacement response spectra based on worldwide digital records. *J. Seismol. 12*, 453–475. doi: 10.1007/s10950-008-9098-y.

Chapman, M. C. (1995). A probabilistic approach to ground-motion selection for engineering design. *Bull. Seismol. Soc. Am. 85*(3), 937–942.

Chebrova, A. Y., V. N. Chebrov, A. A. Gusev, A. V. Lander, E. M. Guseva, S. V. Mityushkina, and A. A. Raevskaya (2015). The impacts of the M_W 8.3 Sea of Okhotsk earthquake of May 24, 2013 in Kamchatka and worldwide. *J. Volcanol. Seismol. 9*(4), 223–241. doi: 10.1134/S074204631504003X.

Chiou, B., R. Darragh, N. Gregor, and W. Silva (2008). NGA Project Strong-Motion Database. *Earthquake Spectra 24*(1), 23–44. doi: 10.1193/1.2894831.

Chiou, B. S.-J. and R. R. Youngs (2008). An NGA Model for the Average Horizontal Component of Peak Ground Motion and Response Spectra. *Earthquake Spectra 24*(1), 173–215. doi: 10.1193/1.2894832.

Chubarova, O. S., A. A. Gusev, and V. N. Chebrov (2010). The ground motion excited by the Olyutorskii earthquake of April 20, 2006 and by its aftershocks based on digital recordings. *J. Volcanol. Seismol. 4*(2), 126–138. doi: 10.1134/S0742046310020065.

Cisternas, A., H. Philip, J. C. Bousquet, M. Cara, A. Deschamps, L. Dorbath, C. Dorbath, H. Haessler, E. Jimenez, A. Nercessian, L. Rivera, B. Romanowicz, A. Gvishiani, N. V. Shebalin, I. Aptekman, S. Arefiev, B. A. Borisov, A. Gorshkov, V. Graizer, A. Lander, K. Pletnev, A. I. Rogozhin, and R. Tatevossian (1989). The Spitak (Armenia) earthquake of 7 December 1988: field observations, seismology and tectonics. *Nature 339*, 675–679. doi: 10.1038/339675a0.

Cooke, P. (1979). Statistical inference for bounds of random variables. *Biometrika 66*(2), 367–374. doi: 10.1093/biomet/66.2.367. Available at http://www.jstor.org/stable/2335672.

Cooke, P. (1980). Optimal linear estimation of bounds of random variables. *Biometrika 67*(1), 257–258. doi: 10.1093/biomet/67.1.257.

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bull. Seismol. Soc. Am. 58*(5), 1583–1606.

Cornell, C. A. (1971). Probabilistic analysis of damage to structures under seismic loads. In D. A. Howells, I. P. Haigh, and C. Taylor (Eds.), *Dynamic waves in civil engineering*, New York, pp. 473–488. Society for Earthquake and Civil Engineering: John Wiley and Sons.

Cornell, C. A. (1994). Statistical analysis of maximum magnitudes. In A. C. Johnston, K. J. Coppersmith, L. R. Kanter, and C. A. Cornell (Eds.), *The Earthquakes of Stable Continental Regions*, Volume 1. Assessment of Large Earthquake Potential, pp. 5.1–5.27. Palo Alto: Electric Power Research Institute.

Cornell, C. A. and S. R. Winterstein (1988). Temporal and magnitude dependence in earthquake recurrence models. *Bull. Seismol. Soc. Am. 78*(4), 1522–1537.

Corradini, M. L. (2003). Letter from chairman of the U.S. Nuclear Waste Technical Review Board to the Director of the Office of Civilian Radioactive Waste Management. Available at http://www.nwtrb.gov/corr/mlc010.pdf.

Costa, G., G. F. Panza, P. Suhadolc, and F. Vaccari (1993). Zoning of the Italian territory in terms of expected peak ground acceleration derived from complete synthetic seismograms. *J. Appl. Geophys. 30*(1-2), 149–160. doi: 10.1016/0926-9851(93)90023-R.

Cramer, C. H., M. D. Petersen, T. Cao, T. R. Toppozada, and M. Reichle (2000). A Time-Dependent Probabilistic Seismic-Hazard Model for California. *Bull. Seismol. Soc. Am. 90*(1), 1–21. doi: 10.1785/0119980087.

Das, R., H. R. Wason, and M. L. Sharma (2011). Global regression relations for conversion of surface wave and body wave magnitudes to moment magnitude. *Nat. Hazards 59*(2), 801–810. doi: 10.1007/s11069-011-9796-6.

de Haan, L. and A. Ferreira (2006). *Extreme Value Theory: An Introduction*. New York: Springer Science and Business Media, LLC.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.

Dekkers, A. L. M., J. H. J. Einmahl, and L. de Haan (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Stat. 17*(4), 1833–1855. Available at https://projecteuclid.org.

Delavaud, E., F. Scherbaum, N. Kuehn, and R. C. (2009). Information-Theoretic Selection of Ground-Motion Prediction Equations for Seismic Hazard Analysis: An Applicability Study Using Californian Data. *Bull. Seismol. Soc. Am. 99*(6), 3248–3263. doi: 10.1785/0120090055.

Dengler, L. A. and J. W. Dewey (1998). An intensity survey of households affected by the Northridge, California, earthquake of 17 January 1994. *Bull. Seismol. Soc. Am. 88*(2), 441–462.

Dewey, J. W., D. J. Wald, and L. A. Dengler (2000). Relating Conventional USGS Modified-Mercalli Intensities to Intensities Assigned with Data Collected via the Internet. *Seismol. Res. Lett. 71*(2), 264. doi: 10.1785/gssrl.71.2.184.

Douglas, J. (2011). Ground-motion prediction equations 1964-2010. Final Rept. RP-59356-FR, Bureau de Recherches Géologiques et Minières (BRGM), Orléans, France, 444 pp. Available at http://peer.berkeley.edu.

Dupuis, D. J. and J. M. Flemming (2006). Modelling peak accelerations from earthquakes. *Earthq. Eng. Struct. Dyn. 35*(8), 969–987. doi: 10.1002/eqe.565.

Earle, P. S., D. J. Wald, K. S. Jaiswal, T. I. Allen, M. G. Hearne, K. D. Marano, A. J. Hotovec, and J. M. Fee (2009). *Prompt Assessment of Global Earthquakes for Response (PAGER): A system for rapidly determining the impact of earthquakes worldwide*. U.S. Geological Survey. Open-File Report 2009-1131.

Ebel, J. E. and A. L. Kafka (1999). A Monte Carlo Approach to Seismic Hazard Analysis. *Bull. Seismol. Soc. Am. 89*(4), 854–866.

Ellsworth, W. L. (2012). Heavy Tails and Earthquake Probabilities. *Seismol. Res. Lett. 83*(3), 483–484. doi: 10.1785/gssrl.83.3.483.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997). *Modelling Extremal Events for Insurance and Finance*. Berlin: Springer-Verlag.

Epstein, B. and C. Lomnitz (1966). A Model for Occurrence of Large Earthquakes. *Nature 211*, 954–956. doi: 10.1038/211954b0.

Esteva, L. (1969). Seismic risk and seismic design decisions. In *Proceedings of MIT Seminar on the Earthquake Resistant Design of Nuclear Reactors*, Cambridge, Massachusetts, pp. 142–182.

European Macroseismic Scale 1998 (EMS-98) (1998). Luxembourg: Centre Européen de Géodynamique et de Séismologie. Available at: http://www.franceseisme.fr/EMS98_Original_english.pdf.

Fisher, R. A. and L. H. C. Tippett (1928). Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Math. Proc. Camb. Phil. Soc. 24*(2), 180–190. doi: 10.1017/S0305004100015681.

Fitzenz, D. D. and M. Nyst (2015). Building Time-Dependent Earthquake Recurrence Models for Probabilistic Risk Computations. *Bull. Seismol. Soc. Am. 105*(1), 120–133. doi: 10.1785/0120140055.

Frankel, A. (1995). Mapping Seismic Hazard in the Central and Eastern United States. *Seismol. Res. Lett. 66*(4), 8–21. doi: 10.1785/gssrl.66.4.8.

Frankel, A. D., M. D. Petersen, C. S. Mueller, K. M. Haller, R. L. Wheeler, E. V. Leyendecker, R. L. Wesson, S. C. Harmsen, C. H. Cramer, D. M. Perkins, and K. S. Rukstales (2002). *Documentation for the 2002 Update of the National Seismic Hazard Maps*. U.S. Geological Survey. Open-File Report 02-420.

Frohlich, C. (1998). Does Maximum Earthquake Size Depend on Focal Depth? *Bull. Seismol. Soc. Am. 88*(2), 329–336.

Gardner, J. K. and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? *Bull. Seismol. Soc. Am. 64*(5), 1363–1367.

Gasperini, P., B. Lolli, and G. Vannucci (2013). Body-Wave Magnitude m_b Is a Good Proxy of Moment Magnitude M_w for Small Earthquakes ($m_b < 4.5 - 5.0$). *Seismol. Res. Lett. 84*(6), 932–937. doi: 10.1785/0220130105.

Geller, R. J. (1976). Scaling relations for earthquake source parameters and magnitudes. *Bull. Seismol. Soc. Am. 66*(5), 1501–1523.

Giardini, D., G. Grünthal, K. M. Shedlock, and P. Zhang (1999). The GSHAP Global Seismic Hazard Map. *Ann. Geofisc. 42*(6), 1225–1228. doi: 10.4401/ag-3784.

Gibowicz, S. J. (1973). Variation of the frequency-magnitude relation during earthquake sequences in New Zealand. *Bull. Seismol. Soc. Am. 63*(2), 517–528.

Gibowicz, S. J. and A. Kijko (1994). *An Introduction to Mining Seismology*. San Diego: Academic Press.

Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire (On the limiting distribution of the maximum term of random series). *Ann. Math. 44*(3), 423–453. doi: 10.2307/1968974. Available at http://www.jstor.org/stable/1968974.

Goda, K., W. Aspinall, and C. A. Taylor (2013). Seismic Hazard Analysis for the U.K.: Sensitivity to Spatial Seismicity Modelling and Ground Motion Prediction Equations. *Seismol. Res. Lett. 84*(1), 112–129. doi: 10.1785/0220120064.

Goes, S. D. B. (1996). Irregular recurrence of large earthquakes: An analysis of historic and paleoseismic catalogs. *J. Geophys. Res. 101*(B3), 5739–5749. doi: 10.1029/95JB03044.

Gomberg, J. (1991). Seismicity and Detection/Location Threshold in the Southern Great Basin Seismic Network. *J. Geophys. Res. 96*(B10), 16401–16414. doi: 10.1029/91JB01593.

Gorbatov, A., V. Kostoglodov, G. Suárez, and E. Gordeev (1997). Seismicity and structure of the Kamchatka subduction zone. *J. Geophys. Res. 102*(B8), 17883–17898. doi: 10.1029/96JB03491.

Green, R. W. E. and S. Bloch (1971). The Ceres, South Africa, earthquake of September 29, 1969: I. Report on some aftershocks. *Bull. Seismol. Soc. Am. 61*(4), 851–859.

Gupta, I. D. (2002). The state of the art in seismic hazard analysis. *ISET Journal of Earthquake Technology 39*(4), 311–346.

Gusev, A. A. (1991). Intermagnitude relationships and asperity statistics. *Pure Appl. Geophys. 136*(4), 515–527. doi: 10.1007/BF00878585.

Gusev, A. A., E. I. Gordeev, E. M. Guseva, A. G. Petukhin, and V. N. Chebrov (1997). The First Version of the $A_{\max}(M_W, R)$ Relationship for Kamchatka. *Pure Appl. Geophys. 149*(2), 299–312. doi: 10.1007/s000240050027.

Gutenberg, B. and C. F. Richter (1944). Frequency of earthquakes in California. *Bull. Seismol. Soc. Am. 34*(4), 185–188.

Gutenberg, B. and C. F. Richter (1956). Earthquake magnitude, intensity, energy, and acceleration: (Second paper). *Bull. Seismol. Soc. Am. 46*(2), 105–145.

Guttorp, P. (1987). On least-squares estimation of *b* values. *Bull. Seismol. Soc. Am. 77*(6), 2115–2124.

Hadjian, A. H. (1993). The Spitak, Armenia earthquake of 7 December 1988 - why so much destruction. *Soil Dynam. Earthquake Eng. 12*(1), 1–24. doi: 10.1016/0267-7261(93)90053-T.

Hamada, M. S., A. G. Wilson, C. S. Reese, and H. F. Martz (2008). *Bayesian Reliability*. New York: Springer.

Hanks, T. C., G. C. Beroza, and S. Toda (2012). Have Recent Earthquakes Exposed Flaws in or Misunderstandings of Probabilistic Seismic Hazard Analysis? *Seismol. Res. Lett. 83*(5), 759–564. doi: 10.1785/0220120043.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Stat. 3*(5), 1163–1174. Available at https://projecteuclid.org.

Hong, H. P., K. Goda, and A. G. Davenport (2006). Seismic hazard analysis: a comparative study. *Can. J. Civ. Eng. 33*(9), 1156–1171. doi: 10.1139/l06-062.

Hosking, J. R. M. (1985). Algorithm AS 215: Maximum likelihood estimation of the parameters of the generalized extreme value distribution. *Applied Statistics 34*(3), 301–310. doi: 10.2307/2347483. Available at http://www.jstor.org/stable/2347483.

Hosking, J. R. M., J. R. Wallis, and E. Wood (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics 27*(3), 251–261. doi: 10.2307/1269706. Available at http://www.jstor.org/stable/1269706.

Huyse, L., R. Chen, and J. A. Stamatakos (2010). Application of Generalized Pareto Distribution to Constrain Uncertainty in Peak Ground Accelerations. *Bull. Seismol. Soc. Am. 100*(1), 87–101. doi: 10.1785/0120080265.

Iervolino, I., M. Giorgio, and B. Polidoro (2014). Sequence-Based Probabilistic Seismic Hazard Analysis. *Bull. Seismol. Soc. Am. 104*(2), 1006–1012. doi: 10.1785/0120130207.

Ishimoto, M. and K. Iida (1939). Observations of earthquakes registered with the microseismograph constructed recently. *Bull. Earthq. Res. Inst., Univ. Tokyo 17*, 443–478.

Jalpa, D. P. and G. M. Atkinson (2012). Scenario ShakeMaps for Ottawa, Canada. *Bull. Seismol. Soc. Am. 102*(2), 650–660. doi: 10.1785/0120100302.

Joyner, W. B. and D. M. Boore (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley, california, earthquake. *Bull. Seismol. Soc. Am. 71*(6), 2011–2038.

Joyner, W. B. and D. M. Boore (1993). Methods of regression analysis of strong motion data. *Bull. Seismol. Soc. Am. 83*(2), 469–487.

Kagan, Y. Y. (1991). Seismic moment distribution. *Geophys. J. Int. 106*(1), 123–134. doi: 10.1111/j.1365-246X.1991.tb04606.x.

Kagan, Y. Y. (2003). Accuracy of modern global earthquake catalogs. *Phys. Earth Planet. In. 135*(2-3), 173–209. doi: 10.1016/S0031-9201(02)00214-5.

Kaka, S. I. and G. M. Atkinson (2004). Relationships between Instrumental Ground-Motion Parameters and Modified Mercalli Intensity in Eastern North America. *Bull. Seismol. Soc. Am. 94*(5), 1728–1736. doi: 10.1785/012003228.

Kanamori, H. and D. L. Anderson (1975). Theoretical basis of some empirical relations in seismology. *Bull. Seismol. Soc. Am. 65*(5), 1073–1095.

Kanno, T., A. Narita, N. Morikawa, H. Fujiwara, and Y. Fukushima (2006). A New Attenuation Relation for Strong Ground Motion in Japan Based on Recorded Data. *Bull. Seismol. Soc. Am. 96*(3), 879–897. doi: 10.1785/0120050138.

Kárník, V. and Z. Schenková (1974). Application of the theory of largest values to earthquake occurrence in the balkan region. *Studia Geophysica et Geodaetica 18*(2), 134–139. doi: 10.1007/bf01614976.

Kijko, A. (1994). Seismological Outliers: L_1 or Adaptive L_p Norm Application. *Bull. Seismol. Soc. Am. 84*(2), 473–477.

Kijko, A. (2004). Estimation of the Maximum Earthquake Magnitude, $m_{\max}$. *Pure Appl. Geophys. 161*(8), 1655–1681. doi: 10.1007/s00024-004-2531-4.

Kijko, A. (2008). Data Driven Probabilistic Seismic Hazard Assessment Procedure for Regions with Uncertain Seismogenic Zones. In E. S. Husebye (Ed.), *Earthquake Monitoring and Seismic Hazard Mitigation in Balkan Countries*, pp. 237–251. Springer Netherlands. doi: 10.1007/978-1-4020-6815-7_16.

Kijko, A. (2011). Seismic Hazard. In H. K. Gupta (Ed.), *Encyclopedia of Solid Earth Geophysics*, Volume 1, pp. 1107–1121. Springer Netherlands. doi: 10.1007/978-90-481-8702-7_10.

Kijko, A. (2012). On Bayesian procedure for maximum earthquake magnitude estimation. *Res. Geophys. 2*(1), 46–51. doi: 10.4081/rg.2012.e7.

Kijko, A. and M. M. Dessokey (1987). Application of the extreme magnitude distributions to incomplete earthquake files. *Bull. Seismol. Soc. Am. 77*(4), 1429–1436.

Kijko, A. and G. Graham (1998). Parametric-historic Procedure for Probabilistic Seismic Hazard Analysis. Part I: Estimation of Maximum Regional Magnitude $m_{\max}$. *Pure Appl. Geophys. 152*(3), 413–442. doi: 10.1007/s000240050161.

Kijko, A. and G. Graham (1999). "Parametric-historic" Procedure for Probabilistic Seismic Hazard Analysis. Part II: Assessment of Seismic Hazard at Specified Site. *Pure Appl. Geophys. 154*(1), 1–22. doi: 10.1007/s000240050218.

Kijko, A., S. Lasocki, and G. Graham (2001). Non-parametric Seismic Hazard in Mines. *Pure Appl. Geophys. 158*(9), 1655–1675. doi: 10.1007/PL00001238.

Kijko, A. and M. A. Sellevoll (1981). Triple exponential distribution, a modified model for the occurrence of large earthquakes. *Bull. Seismol. Soc. Am. 71*(6), 2097–2101.

Kijko, A. and M. A. Sellevoll (1989). Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes. *Bull. Seismol. Soc. Am. 79*(3), 645–654.

Kijko, A. and M. A. Sellevoll (1992). Estimation of earthquake hazard parameters from incomplete data files. Part II. Incorporation of magnitude heterogeneity. *Bull. Seismol. Soc. Am. 82*(1), 120–134.

Kijko, A. and M. Singh (2011). Statistical Tools for Maximum Possible Earthquake Magnitude Estimation. *Acta Geophys. Pol. 59*(4), 674–700. doi: 10.2478/s11600-011-0012-6.

Kijko, A. and A. Smit (2012). Extension of the Aki-Utsu *b*-Value Estimator for Incomplete Catalogs. *Bull. Seismol. Soc. Am. 102*(3), 1283–1287. doi: 10.1785/0120110226.

Kijko, A., A. Smit, and M. A. Sellevoll (2016). Estimation of Earthquake Hazard Parameters from Incomplete Data Files. Part III. Incorporation of Uncertainty of Earthquake-Occurrence Model. *Bull. Seismol. Soc. Am. 106*(3). doi: 10.1785/0120150252.

Klügel, J.-U. (2005). Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants. *Eng. Geol. 78*(3-4), 285–307. doi: 10.1016/j.enggeo.2005.01.007.

Klügel, J.-U. (2008). Seismic Hazard Analysis - Quo vadis? *Earth Sci. Rev. 88*(1), 1–32. doi: 10.1016/j.earscirev.2008.01.003.

Klügel, J.-U. (2011). Uncertainty Analysis and Expert Judgment in Seismic Hazard Analysis. *Pure Appl. Geophys. 168*(1), 27–53. doi: 10.1007/s00024-010-0155-4.

Klügel, J.-U., L. Mualchin, and G. F. Panza (2006). A scenario-based procedure for seismic risk analysis. *Eng. Geol. 88*(1-2), 1–22. doi: 10.1016/j.enggeo.2006.07.006.

Klugman, S. A., H. H. Panjer, and G. E. Willmot (2008). *Loss Models: From Data to Decisions*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Knopoff, L. and Y. Y. Kagan (1977). Analysis of the Extremes as Applied to Earthquake Problems. *J. Geophys. Res. 82*, 5647–5657. doi: 10.1029/JB082i036p05647.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione (On the empirical determination of a distribution law). *Giorn. Ist. Ital. Attuar. 4*(1), 83–91.

Kossobokov, V. G. and A. K. Nekrasova (2011). Global seismic hazard assessment program (GSHAP) maps are misleading. *Probl. Eng. Seismol. 38*(1), 65–76. (in Russian).

Kossobokov, V. G. and A. K. Nekrasova (2012). Global Seismic Hazard Assessment Program Maps are Erroneous. *Seism. Instrum. 48*(2), 162–170. doi: 10.3103/S0747923912020065.

Kramer, S. L. (1996). *Geotechnical Earthquake Engineering*, Chapter Seismic Hazard Analysis, pp. 106–142. International Series in Civil Engineering and Engineering Mechanics. Prentice-Hall.

Krinitzsky, E. L. (1995). Deterministic versus probabilistic seismic hazard analysis for critical structures. *Eng. Geol. 40*(1-2), 1–7. doi: 10.1016/0013-7952(95)00031-3.

Krinitzsky, E. L. (2003). How to combine deterministic and probabilistic methods for assessing earthquake hazards. *Eng. Geol. 70*(1-2), 157–163. doi: 10.1016/S0013-7952(02)00269-7.

Kulkarni, R. B., R. R. Youngs, and K. J. Coppersmith (1984). Assessment of confidence intervals for results of seismic hazard analysis. In *Proceedings of the 8th World Conference on Earthquake Engineering*, Volume 1, San Francisco, USA, pp. 263–270.

Kværna, T., F. Ringdal, J. Schweitzer, and L. Taylor (2002). Optimized Seismic Threshold Monitoring - Part 1: Regional Processing. *Pure Appl. Geophys. 159*(5), 969–987. doi: 10.1007/s00024-002-8668-0.

Lavallée, D. and R. J. Archuleta (2005). Coupling of the random properties of the source and the ground motion for the 1999 Chi Chi earthquake. *Geophys. Res. Lett. 32*(L08311). doi: 10.1029/2004GL022202.

Lemeshko, B. and S. Lemeshko (2009). Distribution models for nonparametric tests for fit in verifying complicated hypotheses and maximum-likelihood estimators. Part 1. *Meas. Tech. 52*(6), 555–565. doi: 10.1007/s11018-009-9330-3.

Leyendecker, E. V., R. J. Hunt, A. D. Frankel, and K. S. Rukstales (2000). Development of Maximum Considered Earthquake Ground Motion Maps. *Earthquake Spectra 16*(1), 21–40. doi: 10.1193/1.1586081.

Lomnitz-Adler, J. and C. Lomnitz (1979). A modified form of the Gutenberg-Richter magnitude-frequency relation. *Bull. Seismol. Soc. Am. 69*(4), 1209–1214.

Lyubushin, A. A., T. M. Tsapanos, V. F. Pisarenko, and G. C. Koravos (2002). Seismic Hazard for Selected Sites in Greece: A Bayesian Estimate of Seismic Peak Ground Acceleration. *Nat. Hazards 25*(1), 83–98. doi: 10.1023/a:1013342918801.

Main, I. (2000). Apparent Breaks in Scaling in the Earthquake Cumulative Frequency-Magnitude Distribution: Fact or Artifact? *Bull. Seismol. Soc. Am. 90*(1), 86–97. doi: 10.1785/0119990086.

Main, I. G. and P. W. Burton (1984). Information theory and the earthquake frequency-magnitude distribution. *Bull. Seismol. Soc. Am. 74*(4), 1409–1426.

Mark, R. K. (1977). Application of linear statistical models of earthquake magnitude versus fault length in estimating maximum expectable earthquakes. *Geology 5*(8), 464–466. doi: 10.1130/0091-7613(1977)5.

Marsan, D. (2003). Triggering of seismicity at short timescales following Californian earthquakes. *J. Geophys. Res. 108*(B5). doi: 10.1029/2002JB001946.

Marzocchi, W. and L. Sandri (2003). A review and new insights on the estimation of the *b*-value and its uncertainty. *Ann. Geophys. 46*(6), 1271–1282. doi: 10.4401/ag-3472.

Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *J. Am. Stat. Assoc. 46*(253), 68–78. doi: 10.2307/2280095. Available at http://www.jstor.org/stable/2280095.

Matthews, M. V., W. L. Ellsworth, and P. A. Reasenberg (2002). A Brownian Model for Recurrent Earthquakes. *Bull. Seismol. Soc. Am. 92*(6), 2233–2250. doi: 10.1785/0120010267.

McGuire, R. K. (1976). *Fortran program for seismic risk analysis*. U.S. Geological Survey. Open-File Report 76-67.

McGuire, R. K. (1977). Effects of uncertainty in seismicity on estimates of seismic hazard for the east coast of the United States. *Bull. Seismol. Soc. Am. 67*(3), 827–848.

McGuire, R. K. (1978). *FRISK: computer program for seismic risk analysis using faults as earthquake sources*. U.S. Geological Survey. Open-File Report 78-1007.

McGuire, R. K. (1979). Adequacy of simple probability models for calculating felt-shaking hazard, using the Chinese earthquake catalog. *Bull. Seismol. Soc. Am. 69*(3), 877–892.

McGuire, R. K. (1993). Computations of seismic hazard. *Ann. Geofisc. 36*(3-4), 181–200. doi: 10.4401/ag-4263.

McGuire, R. K. (1995). Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bull. Seismol. Soc. Am. 85*(5), 1275–1284.

McGuire, R. K. (2001). Deterministic vs. probabilistic earthquake hazards and risks. *Soil Dynam. Earthquake Eng. 21*(5), 377–384. doi: 10.1016/s0267-7261(01)00019-7.

McGuire, R. K. (2008). Probabilistic seismic hazard analysis: Early history. *Earthq. Eng. Struct. Dyn. 37*(3), 329–338. doi: 10.1002/eqe.765.

Merz, H. A. and C. A. Cornell (1973). Seismic risk analysis based on a quadratic magnitude-frequency law. *Bull. Seismol. Soc. Am. 63*(6), 1999–2006.

Mezcua, J., J. Rueda, and R. M. García Blanco (2013). Observed and Calculated Intensities as a Test of a Probabilistic Seismic-Hazard Analysis of Spain. *Seismol. Res. Lett. 84*(5), 772–780. doi: 10.1785/0220130020.

Midzi, V., B. Zulu, B. Manzunzu, T. Mulabisana, T. Pule, T. Myendeki, and W. Gubela (2015). Macroseismic survey of the ML5.5, 2014 Orkney earthquake. *J. Seismol. 19*(3), 741–751. doi: 10.1007/s10950-015-9491-2.

Mignan, A., M. J. Werner, S. Wiemer, C.-C. Chen, and Y.-M. Wu (2011). Bayesian Estimation of the Spatially Varying Completeness Magnitude of Earthquake Catalogs. *Bull. Seismol. Soc. Am. 101*(3), 1371–1385. doi: 10.1785/0120100223.

Mignan, A. and J. Woessner (2012). Estimating the magnitude of completeness for earthquake catalogs. *Community Online Resource for Statistical Seismicity Analysis*. doi: 10.5078/corssa-00180805. Available at http://www.corsa.org.

Milne, W. G. and A. G. Davenport (1969). Distribution of earthquake risk in Canada. *Bull. Seismol. Soc. Am. 59*(2), 729–754.

Miyake, H., T. Iwata, and K. Irikura (2001). Estimation of rupture propagation direction and strong motion generation from azimuth and distance dependence of source amplitude spectra. *Geophys. Res. Lett. 28*(14), 2727–2730. doi: 10.1029/2000GL011669.

Miyazawa, M. and J. Mori (2009). Test of seismic hazard map from 500 years of recorded intensity data in Japan. *Bull. Seismol. Soc. Am. 99*(6), 3140–3149. doi: 10.1785/0120080262.

Molchan, G. M., V. I. Keilis-Borok, and G. V. Vilkovich (1970). Seismicity and Principal Seismic Effects. *Geophys. J. Roy. Astron. Soc. 21*(3), 323–335. doi: 10.1111/j.1365-246X.1970.tb01795.x.

Molchan, G. M., T. Kronrod, and G. F. Panza (1997). Multi-Scale Seismicity Model for Seismic Risk. *Bull. Seismol. Soc. Am. 87*(5), 1220–1229.

Molina, S., C. D. Lindholm, and H. Bungum (2001). Probabilistic seismic hazard analysis: zoning free versus zoning methodology. *Bollettino di Geofisica Teorica ed Applicata 42*(1-2), 19–39.

Molnar, P. (1979). Earthquake recurrence intervals and plate tectonics. *Bull. Seismol. Soc. Am. 69*(1), 115–133.

Mouroux, P., M. Bertrand, M. Bour, B. L. Brun, S. Depinois, P. Masure, and Risk-UE team (2004). The European Risk-UE Project: an advanced approach to earthquake risk scenarios. In *Proceedings of 13th World Conference on Earthquake Engineering*, Vancouver, B.C., Canada. Paper No. 3329.

Murphy, J. R. and L. J. O'Brien (1977). The correlation of peak ground acceleration amplitude with seismic intensity and other physical parameters. *Bull. Seismol. Soc. Am. 67*(3), 877–915.

Musmeci, F. and D. Vere-Jones (1992). A space-time clustering model for historical earthquakes. *Ann. Inst. Stat. Math. 44*(1), 1–11. doi: 10.1007/BF00048666.

Musson, R. M. W. (2000). The use of Monte Carlo simulations for seismic hazard assessment in the U.K. *Ann. Geofisc. 43*(1), 1–9. doi: 10.4401/ag-3617.

Nekrasova, A., V. Kossobokov, A. Peresan, and A. Magrin (2014). The comparison of the NDSHA, PSHA seismic hazard maps and real seismicity for the Italian territory. *Nat. Hazards 70*(1), 629–641. doi: 10.1007/s11069-013-0832-6.

Nelder, J. and R. Mead (1965). A simplex method for function minimization. *Comput. J. 7*(4), 308–313. doi: 10.1093/comjnl/7.4.308.

Neumann, F. (1954). *Earthquake Intensity and Related Ground Motion*. Seattle: Univ. of Washington Press.

Nishenko, S. P. and R. Buland (1987). A generic recurrence interval distribution for earthquake forecasting. *Bull. Seismol. Soc. Am. 77*(4), 1382–1399.

Nuannin, P., O. Kulhanek, and L. Persson (2005). Spatial and temporal *b* value anomalies preceding the devastating off coast of NW Sumatra earthquake of December 26, 2004. *Geophys. Res. Lett. 32*(L11307). doi: 10.1029/2005GL022679.

Ogata, Y. and D. Vere-Jones (1984). Inference for earthquake models: A self-correcting model. *Stochastic Processes and their Applications 17*(2), 337–347. doi: 10.1016/0304-4149(84)90009-7.

Ordaz, M. and C. Reyes (1999). Earthquake hazard in Mexico City: Observations versus computations. *Bull. Seismol. Soc. Am. 89*(5), 1379–1383.

Orozova, I. M. and P. Suhadolc (1999). A deterministic-probabilistic approach for seismic hazard assessment. *Tectonophysics 312*(2-4), 191–202. doi: 10.1016/s0040-1951(99)00162-6.

Page, R. (1968). Aftershocks and microaftershocks of the great Alaska earthquake of 1964. *Bull. Seismol. Soc. Am. 58*(3), 1131–1168.

Papazachos, C. B. (1992). Anisotropic Radiation Modelling of Macroseismic Intensities for Estimation of the Attenuation Structure of the Upper Crust in Greece. *Pure Appl. Geophys. 138*(3), 445–469. doi: 10.1007/BF00876882.

Park, S. and T.-K. Hong (2016). Joint Determination of Event Epicenter and Magnitude from Seismic Intensities. *Bull. Seismol. Soc. Am. 106*(2), 499–511. doi: 10.1785/0120150158.

Parvez, I. A. and A. Ram (1997). Probabilistic Assessment of Earthquake Hazards in the North-East Indian Peninsula and Hindukush Regions. *Pure Appl. Geophys. 149*(4), 731–746. doi: 10.1007/s000240050049.

Patwardhan, A. S., R. B. Kulkarni, and D. Tocher (1980). A semi-Markov model for characterizing recurrence of great earthquakes. *Bull. Seismol. Soc. Am. 70*(1), 323–347.

Pavlenko, O. V. (2008). Characteristics of seismic wave attenuation in the crust and upper mantle of the Northern Caucasus. *Izv., Phys. Solid Earth 44*(6), 487–494. doi: 10.1134/S1069351308060049.

Pavlenko, O. V. (2009). The study of the radiation characteristics and propagation of seismic waves in the North Caucasus by modeling the accelerograms of the recorded earthquakes. *Izv., Phys. Solid Earth 45*(10), 874–884. doi: 10.1134/S106935130910005X.

Pavlenko, O. V. (2011). Regional differences in the characteristics of seismic wave emission and propagation in Kamchatka and the Northern Caucasus. *Doklady Earth Sciences 438*(2), 846–852. doi: 10.1134/S1028334X11060225.

Pavlenko, O. V. (2016). The Q-factor estimates for the crust and upper mantle in the vicinity of Sochi and Anapa (North Caucasus). *Izv., Phys. Solid Earth 52*(3), 353–363. doi: 10.1134/S1069351316030101.

Pavlenko, V. A. (2015). Effect of alternative distributions of ground motion variability on results of probabilistic seismic hazard analysis. *Nat. Hazards 78*(3), 1917–1930. doi: 10.1007/s11069-015-1810-y.

Pavlenko, V. A. and O. V. Pavlenko (2016). The seismic wave absorption in the crust and upper mantle in the vicinity of the Kislovodsk seismic station. *Izv., Phys. Solid Earth 52*(4), 492–502. doi: 10.1134/S1069351316030113.

Peresan, A., E. Zuccolo, F. Vaccari, A. Gorshkov, and G. F. Panza (2011). Neo-Deterministic Seismic Hazard and Pattern Recognition Techniques: Time-Dependent Scenarios for North-Eastern Italy. *Pure Appl. Geophys. 168*(3), 583–607. doi: 10.1007/s00024-010-0166-1.

Peruzza, L. (1996). Attenuating intensities. *Ann. Geofisc. XXXIX*(5), 1079–1093. doi: 10.4401/ag-4037.

Petersen, M. D., M. P. Moschetti, P. M. Powers, C. S. Mueller, K. M. Haller, A. D. Frankel, Y. Zeng, S. Rezaeian, S. C. Harmsen, O. S. Boyd, N. Field, R. Chen, K. S. Rukstales, N. Luco, R. L. Wheeler, R. A. Williams, and A. H. Olsen (2014). *Documentation for the 2014 Update of the United States National Seismic Hazard Maps*. U.S. Geological Survey. Open-File Report 2014-1091, doi: 10.3133/ofr20141091.

Petukhin, A. G., A. A. Gusev, E. M. Guseva, E. I. Gordeev, and V. N. Chebrov (1999). Preliminary Model for Scaling of Fourier Spectra of Strong Ground Motion Recorded on Kamchatka. *Pure Appl. Geophys. 156*(3), 445–468. doi: 10.1007/s000240050307.

Pickands, J. (1975). Statistical inference using extreme order statistics. *Ann. Stat. 3*(1), 119–131. Available at http://www.jstor.org/stable/2958083.

Pictet, O. V., M. M. Dacorogna, and U. A. Muller (1998). Hill, bootstrap and jackknife estimators for heavy tails. In R. J. Adler, R. E. Feldman, and M. S. Taqqu (Eds.), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, pp. 283–310. Birkhäuser.

Pisarenko, V. F. and A. A. Lyubushin (1997). Statistical estimation of maximum peak ground acceleration at a given point of a seismic region. *J. Seismol. 1*(4), 395–405. doi: 10.1023/a:1009795503733.

Pisarenko, V. F. and A. A. Lyubushin (1999). A Bayesian Approach to Seismic Hazard Estimation: Maximum Values of Magnitudes and Peak Ground Accelerations. *Earthquake Res. China (English Edition) 13*(1), 45–57. doi: 10.1023/a:1013342918801.

Pisarenko, V. F., A. A. Lyubushin, V. B. Lysenko, and T. V. Golubeva (1996). Statistical Estimation of Seismic Hazard Parameters: Maximum Possible Magnitude and Related Parameters. *Bull. Seismol. Soc. Am. 86*(3), 691–700.

Pisarenko, V. F. and M. V. Rodkin (2010). *Heavy-Tailed Distributions in Disaster Analysis*. Heidelberg: Springer.

Pisarenko, V. F. and M. V. Rodkin (2014). *Statistical Analysis of Natural Disasters and Related Losses*. Cham: Springer.

Pisarenko, V. F., D. Sornette, and M. V. Rodkin (2010). Distribution of maximum earthquake magnitudes in future time intervals: application to the seismicity of Japan (1923-2007). *Earth Planets Space 62*(7), 567–578. doi: 10.5047/eps.2010.06.003.

Prescott, P. and A. T. Walden (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika 67*(3), 723–724. doi: 10.1093/biomet/67.3.723.

Raschke, M. (2013). Statistical modeling of ground motion relations for seismic hazard analysis. *J. Seismol. 17*(4), 1157–1182. doi: 10.1007/s10950-013-9386-z.

Reiter, L. (1990). *Earthquake Hazard Analysis. Issues and Insights*. New York: Columbia University Press.

Richter, C. F. (1958). *Elementary Seismology*. San Francisco: W. H. Freeman and Company.

Ringdal, F. (1975). On the estimation of seismic detection thresholds. *Bull. Seismol. Soc. Am. 65*(6), 1631–1642.

Robson, D. S. and J. H. Whitlock (1964). Estimation of a truncation point. *Biometrika 51*(1-2), 33–39. Available at http://www.jstor.org/stable/2334193.

Romeo, R. and A. Prestininzi (2000). Probabilistic versus deterministic hazard analysis: An integrated approach for siting problems. *Soil Dynam. Earthquake Eng. 20*(1-4), 75–84. doi: 10.1016/S0267-7261(00)00039-7.

Rydelek, P. A. and I. S. Sacks (1989). Testing the completeness of earthquake catalogues and the hypothesis of self-similarity. *Nature 337*, 251–253. doi: 10.1038/337251a0.

Rydelek, P. A. and I. S. Sacks (2003). Comment on "Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan," by Stefan Wiemer and Max Wyss. *Bull. Seismol. Soc. Am. 93*(4), 1862–1867. doi: 10.1785/0120020035.

Saragoni, G. R., J. Crempien, and R. Araya (1982). Characteristics of Chilean seismic strong motions. *Revista del I.D.I.E.M., Universidad de Chile, Chile 21*, 67–87. In Spanish.

Schenková, Z. and V. Kárník (1978). The third asymptotic distribution of largest magnitudes in the Balkan earthquake provinces. *Pure Appl. Geophys. 116*(6), 1314–1325. doi: 10.1007/bf00874690.

Schenková, Z. and V. Schenk (1975). Return periods of earthquakes and trends of seismic activity. *Pure Appl. Geophys. 113*(1), 683–693. doi: 10.1007/bf01592952.

Scherbaum, F., F. Cotton, and P. Smit (2004). On the Use of Response Spectral-Reference Data for the Selection and Ranking of Ground-Motion Models for Seismic-Hazard Analysis in Regions of Moderate Seismicity: The Case of Rock Motion. *Bull. Seismol. Soc. Am. 94*(6), 2164–2185. doi: 10.1785/0120030147.

Scherbaum, F., E. Delavaud, and C. Riggelsen (2009). Model Selection in Seismic Hazard Analysis: An Information-Theoretic Perspective. *Bull. Seismol. Soc. Am. 99*(6), 3234–3247. doi: 10.1785/0120080347.

Scherbaum, F. and F. Krüger (2014). The 29 September 1969, Ceres, South Africa, Earthquake: Full Waveform Moment Tensor Inversion for Point Source and Kinematic Source Parameters. *Bull. Seismol. Soc. Am. 104*(1), 576–581. doi: 10.1785/0120130209.

Scholz, C. H. (1968). The frequency-magnitude relation of microfracturing rock and its relation to earthquakes. *Bull. Seismol. Soc. Am. 58*(1), 399–415.

Schorlemmer, D., S. Wiemer, and M. Wyss (2005). Variations in earthquake-size distribution across different stress regimes. *Nature 437*, 539–542. doi: 10.1038/nature04094.

Schorlemmer, D. and J. Woessner (2008). Probability of Detecting an Earthquake. *Bull. Seismol. Soc. Am. 98*(5), 2103–2217. doi: 10.1785/0120070105.

Schwartz, D. P. and K. J. Coppersmith (1984). Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones. *J. Geophys. Res. 89*(B7), 5681–5698. doi: 10.1029/JB089iB07p05681.

Shi, Y. and B. A. Bolt (1982). The standard error of the magnitude-frequency *b* value. *Bull. Seismol. Soc. Am. 72*(5), 1677–1687.

Shimazaki, K. and T. Nakata (1980). Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett. 7*(4), 279–282. doi: 10.1029/GL007i004p00279.

Shumilina, L. S., A. A. Gusev, and V. M. Pavlov (2000). An improved technique for determination of seismic hazard. *J. Earthq. Predict. Res. 8*, 104–110.

Singh, S. K., E. Bazan, and L. Esteva (1980). Expected earthquake magnitude from a fault. *Bull. Seismol. Soc. Am. 70*(3), 903–914.

Smith, S. W. (1976). Determination of maximum earthquake magnitude. *Geophys. Res. Lett. 3*(6), 351–354. doi: 10.1029/GL003i006p00351.

Smith, W. D. (1981). The *b*-value as an earthquake precursor. *Nature 289*, 136–139. doi: 10.1038/289136a0.

Somerville, P. G., N. F. Smith, R. W. Graves, and N. A. Abrahamson (1997). Modification of Empirical Strong Ground Motion Attenuation Relations to Include the Amplitude and Duration Effects of Rupture Directivity. *Seismol. Res. Lett. 68*(1), 199–222. doi: 10.1785/gssrl.68.1.199.

Sørensen, M. B., D. Stromeyer, and G. Grünthal (2009). Attenuation of Macroseismic Intensity: A New Relation for the Marmara Sea Region, Northwest Turkey. *Bull. Seismol. Soc. Am. 99*(2A), 538–553. doi: 10.1785/0120080299.

Sørensen, M. B., D. Stromeyer, and G. Grünthal (2010). A macroseismic intensity prediction equation for intermediate depth earthquakes in Vrancea region, Romania. *Soil Dynam. Earthquake Eng. 30*(11), 1268–1278. doi: 10.1016/j.soildyn.2010.05.009.

Spada, M., S. Wiemer, and E. Kissling (2011). Quantifying a Potential Bias in Probabilistic Seismic Hazard Assessment: Seismotectonic Zonation with Fractal Properties. *Bull. Seismol. Soc. Am. 101*(6), 2694–2711. doi: 10.1785/0120110006.

Spudich, P. and B. S.-J. Chiou (2008). Directivity in NGA Earthquake Ground Motions: Analysis Using Isochrone Theory. *Earthquake Spectra 24*(1), 279–298. doi: 10.1193/1.2928225.

Stafford, P. J., F. O. Strasser, and J. J. Bommer (2008). An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region. *Bull. Earthquake Eng. 6*(2), 149–177. doi: 10.1007/s10518-007-9053-2.

Stamatakos, J. A. (2004). Review by the Office of Nuclear Material Safety and Safeguards of the U.S. Department of Energy's responses to key technical issue agreements SDS.2.01 and SDS.2.02 for a potential Geologic Repository at Yucca Mountain, Nevada, Project No: WM-011. U.S. Nuclear Regulatory Commission Contract NRC-02-02-012. Available at http://pbadupws.nrc.gov/docs/ml0428/ml042870242.pdf.

Stein, S., R. Geller, and M. Liu (2011). Bad Assumptions or Bad Luck: Why Earthquake Hazard Maps Need Objective Testing. *Seismol. Res. Lett. 82*(5), 623–626. doi: 10.1785/gssrl.82.5.623.

Stein, S., R. Geller, and M. Liu (2012). Why earthquake hazard maps often fail and what to do about it. *Tectonophysics 562-563*, 1–25. doi: 10.1016/j.tecto.2012.06.047.

Stein, S., B. D. Spencer, and E. M. Brooks (2015). Metrics for Assessing Earthquake-Hazard Map Performance. *Bull. Seismol. Soc. Am. 105*(4), 2160–2173. doi: 10.1785/0120140164.

Stepp, J. C. (1972). Analysis of completeness of the earthquake sample in the puget sound area and its effects on statistical estimates of earthquake hazard. In *Proceedings of the International Conference on Microzonation for Safer Construction Research and Application*, Seattle, USA, pp. 897–909.

Stepp, J. C., I. Wong, J. Whitney, R. Quittemeyer, N. Abrahamson, G. Toro, R. Youngs, K. Coppersmith, J. Savy, T. Sullivan, and Yucca Mountain PSHA Project Members (2001). Probabilistic Seismic Hazard Analyses for Ground Motions and Fault Displacements at Yucca Mountain, Nevada. *Earthquake Spectra 17*(1), 113–151. doi: 10.1193/1.1586169.

Stirling, M. (2012). Earthquake Hazard Maps and Objective Testing: The Hazard Mapper's Point of View. *Seismol. Res. Lett. 83*(2), 231–232. doi: 10.1785/gssrl.83.2.231.

Stirling, M. and M. Gerstenberger (2010). Ground motion-based testing of seismic hazard models in New Zealand. *Bull. Seismol. Soc. Am. 100*(4), 1407–1414. doi: 10.1785/0120090336.

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, Insights, and Challenges. *Seismol. Res. Lett. 80*(1), 40–56. doi: 10.1785/gssrl.80.1.40.

Strasser, F. O. and J. J. Bommer (2009). Strong Ground Motions - Have We Seen the Worst? *Bull. Seismol. Soc. Am. 99*(5), 2613–2637. doi: 10.1785/0120080300.

Strasser, F. O., J. J. Bommer, and N. A. Abrahamson (2004). The need for upper bounds on seismic ground motion. In *Proceedings of 13th World Conference on Earthquake Engineering*, Vancouver, B.C., Canada. Paper No. 3361.

Strasser, F. O., J. J. Bommer, and N. A. Abrahamson (2008a). Estimating ground-motion variability: issues, insights & challenges. In *Proceedings of 14th World Conference on Earthquake Engineering*, Beijing, China.

Strasser, F. O., J. J. Bommer, and N. A. Abrahamson (2008b). Truncation of the distribution of ground-motion residuals. *J. Seismol. 12*(1), 79–105. doi: 10.1007/s10950-007-9073-z.

Stucchi, M., P. Albini, C. Mirto, and A. Rebez (2004). Assessing the completeness of italian historical earthquake. *47*(2-3), 659–673. doi: 10.4401/ag-3330.

Tate, R. F. (1959). Unbiased Estimation Functions of Location and Scale Parameters. *Ann. Math. Stat. 30*(2), 341–366. Available at http://www.jstor.org/stable/2237086.

Taylor, D. W. A., J. A. Snoke, I. S. Sacks, and T. Takanami (1990). Nonlinear frequency-magnitude relationships for the Hokkaido corner, Japan. *Bull. Seismol. Soc. Am. 80*(2), 340–353.

Teramo, A., E. Stillitani, and A. Bottari (1995). On an anisotropic attenuation law of the macroseismic intensity. *Nat. Hazards 11*(3), 203–221. doi: 10.1007/BF00613407.

Termini, D., A. Teramo, C. Bottari, and T. Tuvè (2005). An Anisotropic Attenuation Law of Macroseismic Intensity Performed on Virtual Intensity Distribution of Seismogenic Zones. *Pure Appl. Geophys. 162*(4), 707–714. doi: 10.1007/s00024-004-2634-y.

Tinti, S. and F. Mulargia (1985). Effects of magnitude uncertainties on estimating the parameters in the Gutenberg-Richter frequency-magnitude law. *Bull. Seismol. Soc. Am. 75*(6), 1681–1697.

Tocher, D. (1958). Earthquake energy and ground breakage. *Bull. Seismol. Soc. Am. 48*(2), 147–153.

Trifunac, M. D. and A. G. Brady (1975). On correlation of seismoscope response with earthquake magnitude and Modified Mercalli Intensity. *Bull. Seismol. Soc. Am. 65*(2), 307–321.

Tsampas, A. D., E. M. Scordilis, C. B. Papazachos, and G. F. Karakaisis (2016). Global-Magnitude Scaling Relations for Intermediate-Depth and Deep-Focus Earthquakes. *Bull. Seismol. Soc. Am. 106*(2), 418–434. doi: 10.1785/0120150201.

Udias, A. and J. Rice (1975). Statistical analysis of microearthquake activity near San Andreas geophysical observatory, Hollister, California. *Bull. Seismol. Soc. Am. 65*(4), 809–827.

Utsu, T. (1965). A method for determining the value of b in the formula $\log N = a - bM$ showing the magnitude-frequency relation for earthquakes. *Geophys. Bull. Hokkaido Univ. 13*, 99–103. in Japanese with English summary.

Veneziano, D., C. A. Cornell, and T. O'hara (1984). *Historical Method for Seismic Hazard Analysis*. Electric Power Research Institute. Report, NP-3438, Palo Alto.

Vorobieva, I., C. Narteau, P. Shebalin, F. Beauducel, A. Nercessian, V. Clouard, and M.-P. Bouin (2013). Multiscale Mapping of Completeness Magnitude of Earthquake Catalogs. *Bull. Seismol. Soc. Am. 103*(4), 2188–2202. doi: 10.1785/0120120132.

Wald, D. J., P. S. Earle, T. I. Allen, K. Jaiswal, K. Porter, and M. Hearne (2008). Development of the U.S. Geological Survey's PAGER system (Prompt Assessment of Global Earthquakes for Response). In *Proceedings of 14th World Conference on Earthquake Engineering*, Beijing, China, pp. 27–41.

Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999). TriNet "ShakeMaps": Rapid Generation of Peak Ground Motion and Intensity Maps for Earthquakes in Southern California. *Earthquake Spectra 15*(3), 537–556. doi: 10.1007/10.1193/1.1586057.

Wang, Z. (2009). Comment on "Sigma: Issues, Insights, and Challenges" by F. O. Strasser, N. A. Abrahamson, and J. J. Bommer. *Seismol. Res. Lett. 80*(3), 491–493. doi: 10.1785/gssrl.80.3.491.

Wang, Z. (2011). Seismic Hazard Assessment: Issues and Alternatives. *Pure Appl. Geophys. 168*(1-2), 11–25. doi: 10.1007/s00024-010-0148-3.

Wang, Z. and M. Zhou (2007). Comment on "Why Do Modern Probabilistic Seismic-Hazard Analyses Often Lead to Increased Hazard Estimates?" by Julian J. Bommer and Norman A. Abrahamson. *Bull. Seismol. Soc. Am. 97*(6), 2212–2214. doi: 10.1785/0120070004.

Ward, S. N. (1995). Area-based tests of long-term seismic hazard predictions. *Bull. Seismol. Soc. Am. 85*(5), 1285–1298.

Ward, S. N. (1997). More on $M_{\max}$. *Bull. Seismol. Soc. Am. 87*(5), 1199–1208.

Weichert, D. H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes. *Bull. Seismol. Soc. Am. 70*(4), 1337–1346.

Wells, D. L. and K. J. Coppersmith (1994). New Empirical Relationships among Magnitude, Rupture Length, Rupture Width, Rupture Area, and Surface Displacement. *Bull. Seismol. Soc. Am. 84*(4), 974–1002.

Wesnousky, S. G. (1994). The Gutenberg-Richter or characteristic earthquake distribution, which is it? *Bull. Seismol. Soc. Am. 84*(6), 1940–1959.

Wheeler, R. L. (2009). *Methods of Mmax Estimation East of the Rocky Mountains*. U.S. Geological Survey. Open-File Report 2009-1018.

Wiemer, S., M. García-Fernández, and J.-P. Burg (2009). Development of a seismic source model for probabilistic seismic hazard assessment of nuclear power plant sites in Switzerland: the view from PEGASOS Expert Group 4 (EG1d). *Swiss J. Geosci. 102*(1), 189–209. doi: 10.1007/s00015-009-1311-7.

Wiemer, S. and M. Wyss (2000). Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan. *Bull. Seismol. Soc. Am. 90*(4), 859–869. doi: 10.1785/0119990114.

Wiemer, S. and M. Wyss (2003). Reply to "Comment on 'Minimum Magnitude of Completeness in Earthquake Catalogs: Examples from Alaska, the Western United States, and Japan,' by Stefan Wiemer and Max Wyss," by Paul A. Rydelek and I. S. Sacks. *Bull. Seismol. Soc. Am. 93*(4), 1868–1871. doi: 10.1785/0120020103.

Woessner, J. and S. Wiemer (2005). Assessing the Quality of Earthquake Catalogues: Estimating the Magnitude of Completeness and Its Uncertainty. *Bull. Seismol. Soc. Am. 95*(2), 684–698. doi: 10.1785/0120040007.

Woo, G. (1996). Kernel Estimation Methods for Seismic Hazard Area Source Modeling. *Bull. Seismol. Soc. Am. 86*(2), 353–362.

Worden, C. B., M. C. Gerstenberger, D. A. Rhoades, and D. J. Wald (2012). Probabilistic Relationships between Ground-Motion Parameters and Modified Mercalli Intensity in California. *Bull. Seismol. Soc. Am. 102*(1), 204–221. doi: 10.1785/0120110156.

Worden, C. B., D. J. Wald, T. I. Allen, K. Lin, D. Garcia, and G. Cua (2010). A Revised Ground-Motion and Intensity Interpolation Scheme for ShakeMap. *Bull. Seismol. Soc. Am. 100*(6), 3083–3096. doi: 10.1785/0120100101.

Wu, J., G. Mengtan, C. Kun, and H. Bei (2011). Discussion on the influence of truncation of ground motion residual distribution on probabilistic seismic hazard assessment. *Earthquake Eng. Eng. Vib. 10*(3), 379–392. doi: 10.1007/s11803-011-0074-0.

Wyss, M. (1979). Estimating maximum expectable magnitude of earthquakes from fault dimensions. *Geology 7*(7), 336–340. doi: 10.1130/0091-7613(1979)7.

Wyss, M. and R. E. Habermann (1988). Precursory Quiescence Before the August 1982 Stone Canyon, San Andreas Fault, Earthquakes. *Pure Appl. Geophys. 126*(2), 333–356. doi: 10.1007/bf00879002.

Wyss, M., A. Hasegawa, S. Wiemer, and N. Umino (1999). Quantitative mapping of precursory seismic quiescence before the 1989, *M* 7.1 off-Sanriku earthquake, Japan. *Ann. Geofisc. 42*(5), 851–869. doi: 10.4401/ag-3765.

Wyss, M., A. Nekrasova, and V. Kossobokov (2012). Errors in expected human losses due to incorrect seismic hazard estimates. *Nat. Hazards 62*(3), 927–935. doi: 10.1007/s11069-012-0125-5.

Yakovlev, G., D. L. Turcotte, J. B. Rundle, and P. B. Rundle (2006). Simulation-Based Distributions of Earthquake Recurrence Times on the San Andreas Fault System. *Bull. Seismol. Soc. Am. 96*(6), 1995–2007. doi: 10.1785/0120050183.

Yamazaki, F. and T. Türker (1992). Spatial variation study on earthquake ground motion observed by the Chiba array. In *Earthquake Engineering, Tenth World Conference*, Rotterdam, pp. 651–656. Balkema.

Yegulalp, T. M. and J. T. Kuo (1974). Statistical prediction of the occurrence of maximum magnitude earthquakes. *Bull. Seismol. Soc. Am. 64*(2), 393–414.

Youngs, R. R., S.-J. Chiou, W. J. Silva, and J. R. Humphrey (1997). Strong Ground Motion Attenuation Relationships for Subduction Zone Earthquakes. *Seismol. Res. Lett. 68*(1), 58–73. doi: 10.1785/gssrl.68.1.58.

Youngs, R. R. and K. J. Coppersmith (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bull. Seismol. Soc. Am. 75*(4), 939–964.

Zhao, J. X., J. Zhang, A. Asano, Y. Ohno, T. Oouchi, T. Takahashi, H. Ogawa, K. Irikura, H. K. Thio, P. G. Somerville, Y. Fukushima, and Y. Fukushima (2006). Attenuation Relations of Strong Ground Motion in Japan Using Site Classification Based on Predominant Period. *Bull. Seismol. Soc. Am. 96*(3), 898–913. doi: 10.1785/0120050122.

Zöller, G. and S. Hainzl (2007). Recurrence Time Distributions of Large Earthquakes in a Stochastic Model for Coupled Fault Systems: The Role of Fault Interaction. *Bull. Seismol. Soc. Am. 97*(5), 1679–1687. doi: 10.1785/0120060262.

Zöller, G., S. Hainzl, J. Kurths, and J. Zschau (2002). A Systematic Test on Precursory Seismic Quiescence in Armenia. *Nat. Hazards 26*(3), 245–263. doi: 10.1023/a:1015685006180.

Zonno, G., R. Rotondi, and C. Brambilla (2009). Mining Macroseismic Fields to Estimate the Probability Distribution of the Intensity at Site. *Bull. Seismol. Soc. Am. 99*(5), 2876–2892. doi: 10.1785/0120090042.

Zuccolo, E., F. Vaccari, A. Peresan, and G. F. Panza (2011). Neo-Deterministic and Probabilistic Seismic Hazard Assessments: a Comparison over the Italian Territory. *Pure Appl. Geophys. 168*(1), 69–83. doi: 10.1007/s00024-010-0151-8.

Appendix A

Thesis Outputs

A.1 Publications

- Pavlenko, V. A., 2015. Effect of alternative distributions of ground motion variability on results of probabilistic seismic hazard analysis. Nat. Hazards 78(3):1917-1930. doi: 10.1007/s11069-015-1810-y
- Pavlenko, V. A., 2015. Uncertainty in the estimates of peak ground acceleration in seismic hazard analysis. Izv., Phys. Solid Earth 51(6):878-884. doi: 10.1134/s1069351315060099
- Pavlenko, V., 2016. Estimation of the upper bound of seismic hazard curve by using the generalized extreme value distribution. Manuscript submitted.
- Pavlenko, V., Kijko, A., Smit, A., Midzi, V., 2016. On Anisotropic Attenuation Law of Modified Mercalli Intensity. Manuscript submitted.
- Pavlenko, V., Kijko, A., 2016. Comparative study of three probabilistic methods for seismic hazard analysis: case studies of Sochi and Kamchatka. Manuscript submitted.

A.2 Conference outputs

- Pavlenko, V., 2014. Uncertainty in the estimates of peak ground acceleration in seismic hazard analysis. *Moscow Science Week*. Poster presentation (in Russian).
- Pavlenko, V., 2014. Data-driven approach to ground motion variability and its effect on the calculation of hazard curves. *3rd Annual International Conference on Geological & Earth Sciences*. Poster presentation.
- Pavlenko, V., 2015. Effect of alternative distributions of ground motion variability on results of Probabilistic Seismic Hazard Analysis. *26th General Assembly of International Union of Geodesy and Geophysics*. Poster presentation.