Extension of the Parametric-Historic Procedure for Probabilistic Seismic Hazard Analysis

by

Petrus Johannes Vermeulen

University of Pretoria

Submitted in partial fulfilment of the degree

Master of Science in Engineering Geology

Department of Geology

Faculty of Natural and Agricultural Sciences

University of Pretoria

March 2014

Supervisor: Professor Andrzej Kijko

Co-supervisor: Professor Jan Louis van Rooy

© University of Pretoria

I, Petrus Johannes Vermeulen, hereby declare that this thesis is my own original work, which I submit for the degree Master of Science in Engineering Geology at the University of Pretoria. I also declare that this work has not previously been submitted by me for a qualification at any tertiary institution. I have, to the best of my knowledge, duly acknowledged all other people's work that I have used.

SIGNATURE:_____

DATE:_____

© University of Pretoria

Dedicated, without hesitation or afterthought, to my parents:

Jannie and Melinda Vermeulen

© University of Pretoria

Abstract:

Some validations of and extensions to the Parametric-Historic procedure as developed in a two-paper sequel by Andrzej Kijko and Gerhard Graham in 1998 and 1999 are presented.

The source-free distribution they derive is validated through a somewhat more rigorous mathematical derivation. The approach is also extended in two ways. The first extension involves direct application of the estimators, which they used to determine maximum regional magnitude, now to directly determine the maximum possible peak ground acceleration at a specific site. The second extension is a generalization of the distribution of peak ground acceleration in a semi-closed form solution to incorporate a wider range of ground motion prediction equations. The first extension is straightforward and is simultaneously developed and validated by example application to actual ground motion data. The second extension is derived and then applied to a specific ground motion prediction equation as example to illustrate its performance.

© University of Pretoria

Acknowledgements

I would like to thank my supervisors, Professors Andrzej Kijko and Louis van Rooy, not only for their leading but also for their kind assistance and support (both technically and morally) throughout the course of my research project. I also thank my other more senior colleagues Ms Ansie Smit from the Natural Hazard Centre and Mr Yazeed Van Wyk and Mr Matthys Dippenaar from the Department of Geology who also provided me with much guidance. I thank Professor S. Lasocki from the Polish Academy of Science for providing us with high quality contemporary data. I thank Savo Smocilac, Bronwyn Strydom, Jacq Crous and Andrea Kreuiter for their advice (informal but important). I am much indebted to the Natural Hazard Centre for financial support. Last but not least, special thanks to my family for reasons that would require more pages than this work is long if I were to name them; however, these include material, moral, and informal technical support.

© University of Pretoria

Table of Contents

© University of Pretoria

© University of Pretoria

List of Figures

© University of Pretoria

© University of Pretoria

List of Tables

© University of Pretoria

List of abbreviations

cdf – cumulative distribution function

GMPE – ground motion prediction equation

G-R law – Gutenberg-Richter law

K-S estimator – Kijko Sellevoll estimator of maximum magnitude

K-S-B estimator – Kijko Sellevoll Bayesian estimator of maximum magnitude

KG – reference to the two-paper sequel in which Kijko and Graham (1998 and 1999) developed their Parametric-Historic approach

log-PGA – logarithm of peak ground acceleration

m – magnitude

m_{max},M_{max} – maximum possible magnitude

m_{min} – minimum magnitude in range under consideration

NGA equations – new generation attenuation equations

PGA – peak ground acceleration

PGV – peak ground velocity

PGD – peak ground displacement

SA – spectral acceleration

PGA_{max} – maximum possible peak ground acceleration

PGA_{min} – minimum peak ground acceleration in range under consideration

pdf – probability density function

PSHA – probabilistic seismic hazard analysis

SD – spectral displacement

© University of Pretoria

SHA – seismic hazard analysis

SV – spectral velocity

T-P estimator – Tate-Pisarenko estimator of maximum magnitude

TP statistic – statistic based on “log moments” used to identify deviation from an exponential distribution

© University of Pretoria

Definitions

Crossover value – value in the tail of a distribution from which it starts to deviate significantly from the assumed model.

Extreme value distribution – distribution of the maxima of groups of a certain number of events. For a seismic catalogue data this may be considered as equivalent to the distribution of the largest magnitudes during time intervals of the same size.

Frequency of exceedance – expected frequency with which a certain value is expected to be exceeded, typically per annum.

Ground motion prediction equation – equation that predicts the central trend of a ground motion parameter. Ground motion prediction equations are generalizations of what is called attenuation relations that predict, in addition to magnitude and distance scaling, the ground motion parameter as a function of all source, path and site effects that can be sufficiently characterized.

Hard bound/hard cutoff – sharp truncation to a pdf.

Parametric-Historic approach – a parametric, data-driven probabilistic seismic hazard analysis, as opposed to deductive procedures that require subjective input. (See McGuire, 1993)

Point estimator – estimator giving an exact value (typically central trend) as opposed to e.g. confidence intervals.

Poisson process – stochastic process for which each event is independent of all previous events.

Probabilistic seismic hazard analysis – defined by Kijko (2011) as “quantification of the probability that a specified level of ground motion will be exceeded at least once at a site or in a region during a specified exposure time”.

Seismic Hazard – defined by Kijko (2011) as “any physical phenomena associated with an earthquake... and their effects on land, man-made structure, and socio-economic systems that have the potential to produce a loss. It is also used without regard to a loss to indicate the probable level of ground shaking occurring at a given point within a certain period of time.”

Soft bound/ soft cutoff – point on a pdf where probability starts to decrease more rapidly.

© University of Pretoria

1. Introduction

The classic paper by Cornell (1968), based on his doctoral thesis (McGuire 2007) marked the beginning of what is known today as probabilistic seismic hazard analysis. Esteva (e.g. 1963; in McGuire, 2007) developed a similar methodology almost simultaneously, but independently. In years to come Cornell and Esteva would interact and exchange ideas to fully develop the formalisms. (McGuire 2007)

The approach they developed is what is described by McGuire (1993) as a *deductive* approach, because it relies heavily on interpretation; the whole hazard analysis can be carried out without data. Later on Veneziano et al. (1984; in McGuire 1993) developed a method that requires only data and makes use of the empirical distribution of the data. It is therefore completely non-parametric. McGuire (1993) uses the term *historic* approach for this latter non-parametric, data-driven approach.

The advantage of the deductive approach is that it accommodates the inclusion of possible seismic gaps, migration of seismicity, and other hypotheses or interpreted effects. The advantage of the historic approach is that it has no need of models or parameters and has no need of interpretative input. The disadvantage of the historic approach is that it cannot make predictions about values that have low probabilities of occurring; that is, it is not suitable for values with longer return periods than the time span of the catalogue. (McGuire, 1993)

Kijko and Graham (1998 and 1999; hereafter referred to as KG1998 and KG1999, respectively, or KG for the two-paper sequel) derived a very elaborate approach which in McGuire's (1993) scheme would be classified as a parametric-historic approach (hereafter "the Parametric-Historic procedure" will refer specifically to the procedure developed in KG). KG1998 also notes that a drawback of the historic approach is that it cannot accommodate catalogues with different levels of completeness, but that one of the largest drawbacks of the deductive approach is that it requires proper delineation of sources, which in many cases is not possible (Southern Africa is an example of such a case).

The aim of KG is to develop an approach with the advantages of both the deductive and the historic approaches, eliminating their weaknesses. KG1998 deals to a large extent only with the development of an estimate of the maximum possible magnitude from historical data, which is an essential parameter both in their Parametric-Historic procedure and in seismology in general.

© University of Pretoria

KG1999 develops the rest of the details of this original approach. It draws from the work of Kijko and Sellevoll (1989 and 1992) the ability to utilise catalogues with different levels of completeness and with considerable uncertainty; it uses maximum likelihood estimates to estimate parameters from catalogue data; it derives a source-free distribution of peak ground acceleration under certain assumptions of ground motion attenuation; and it derives a method to estimate PGA_{max} from m_{max} while taking the uncertainty the GMPE into account. It turns out that the distribution of the logarithm of PGA is of the same form as that of magnitude. It is important to note that the parametric nature of KG's whole approach allows Bayesian incorporation of *a priori* information.

© University of Pretoria

2. Aim of this project

In Kijko and Graham's (1999) development of their *Parametric-Historic* approach, they stated and justified briefly that the cdf of the distribution of log-PGA data is a truncated exponential distribution:

$$P[\ln(PGA) \leq y] = \begin{cases} 0, & y < \ln(PGA_{\min}) \\ \dfrac{1-\exp[-\gamma(y-\ln(PGA_{\min})]}{1-\exp[-\gamma(\ln(PGA_{\max})-\ln(PGA_{\min})]}, & \ln(PGA_{\min}) \leq y \leq \ln(PGA_{\max}) \\ 1, & y > \ln(PGA_{\max}) \end{cases} \quad (2.1)$$

where γ is the mean value in the case of an untruncated exponential distribution.

The aims of this research project are:

1) To validate the distribution model stated in equation (2.1)
2) To extend the Parametric-Historic approach by direct application of the maximum estimators discussed in Kijko and Singh (2011).
3) To extend the Parametric-Historic approach to incorporate GMPE's of a more general form in an analytical or semi-analytical representation of the cdf of log-PGA.

© University of Pretoria

3. Background

The goal of most SHAs is to make statements about, or to quantify the strong ground motion to be expected in future at a site of interest (Budnitz *et al.*, 1997). This information is then typically used for risk assessments and engineering design purposes. It is common practice to express ground motion in terms of PGV, PGA, SA, SV, SD, etc. (Bommer and Abrahamson, 2006; Gupta, 2002; Campbell, 2003). Currently methodology relies upon two conceptual approaches (Kijko, 2011; Gupta, 2002; Kramer, 1996): deterministic seismic hazard analysis (DSHA), which uses deterministic, physically *certain* principles to analyse seismic hazard and usually focuses on a specific (often worst case) scenario; and probabilistic seismic hazard analysis (PSHA), which determines the hazard by probabilistic means and provides return periods of a whole spectrum of values of strong ground motion. The results of PSHA is thus not certain, but rather is expressed in terms of the probability of a certain value of ground motion being exceeded (commonly referred to in seismology as the *frequency of exceedance* of that value) or the return period ground motion exceeding that value. Contrary to past popular views, PSHA and DSHA are not dogmatically mutually exclusive, but rather constitute the ends of a continuous spectrum (Bommer, 2002; Kijko, 2011; McGuire, 2001). Paraphrasing McGuire, deterministic (physical) concepts and data can – or should – be used to benchmark probabilistic results (probabilistic results should be consistent with deterministic results). Determinism may also be incorporated into PSHA as far as feasible by use of the *Bayesian formalism* (Kijko and Sellevoll, 1989; Kijko and Sellevoll and 1992).

3.1. Basic theoretical aspects pertaining to earthquake physics

One of the mechanical models in fault mechanics that is very pervasive in the literature (eg. Kawamur *et al.*, 2012; Erickson *et al.*, 2011; Ben-Zion, 2008; Carlson *et al.*, 1994) is known as the Burridge-Knopoff slider block model. It is used as starting point here; it is not the only widely used model, but is sufficient to introduce the basic concepts. This model involves an array of blocks connected to a surface on one side, by leaf springs and to each other by harmonic springs, and scratches (i.e. contact involving friction) a surface that is opposite the surface the blocks are connected to (Carlson *et al.*, 1994). The one surface is then moved relative to the other, and the behaviour of this model as a simulated fault model is then analysed. Figure 3.1 is a depiction of a one dimensional slider-block model. Parameters and behaviour of this model investigated by seismologists range from mere energy release, propagation pattern (Figure 3.2.), rupture velocity, to the characteristics of slip pulses within the fault (Erickson *et al.*, 2011; Carlson *et al.*, 1994). Workers such as Carlson *et al.* (1994) have shown that energy release of this model behaves as predicted by

© University of Pretoria

the Gutenberg-Richter frequency-magnitude relation. The Gutenberg-Richter relation for earthquake magnitude implies a Pareto (or power law) distribution of earthquake energy release per event, which, in turn is consistent with the theory of self-organized criticality in many chaotic systems (Sornette, 2006; Sornette and Sornette, 1989; Bak and Tang, 1989; Bak, 1999).). Erickson *et al.* (2011) have shown that the model possess nonlinear, chaotic behaviour.

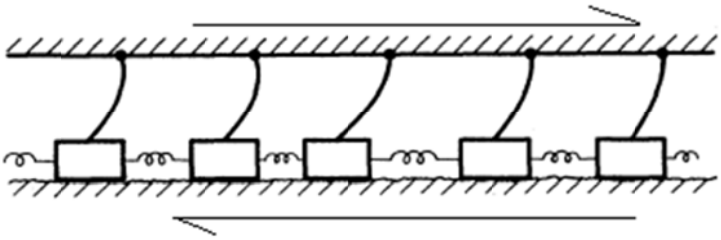

Figure 3.1. One dimensional version of Burridge-Knopoff slider-block model. (Adapted from Carlson *et al.*, 1994)

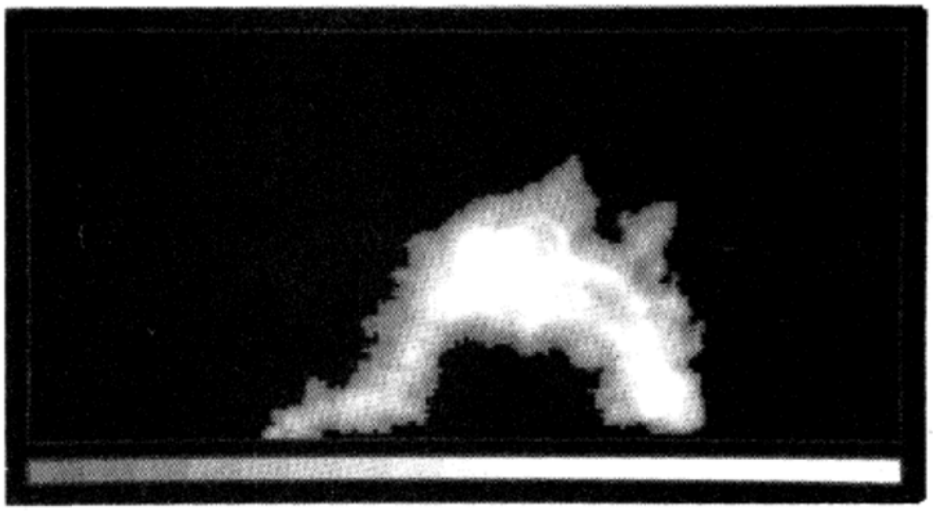

Figure 3.2. Slip distribution of a slider-block model simulation (Carlson *et al.*, 1994)

Another aspect of earthquake physics which is of interest here is the propagation of seismic waves, and for a hazard analysis specifically wave propagation to a site of interest. It is not hard to comprehend how complex travel paths of waves may become in heterogeneous geological media. Seismologists are not yet at a point where complex geological models can be used to accurately determine ray paths for earthquake-seismological purposes. Ground motion prediction equations only capture the central tendencies of wave attenuation, energy dissipation, and group wave effects (Bommer and Abrahamson, 2006). Some pervasive geological ray path effect has in fact been introduced into the newest ground motion prediction equations such as the scattering effects of the

© University of Pretoria

Moho discontinuity (e.g. Atkinson and Boore, 2006). Figure 3.3 (Atkinson and Boore, 2006) illustrates this effect – note the discontinuous slope in the curves (each solid line represents the GMPE at a different earthquake magnitude) – this is representative of the scattering effects of the Moho discontinuity.

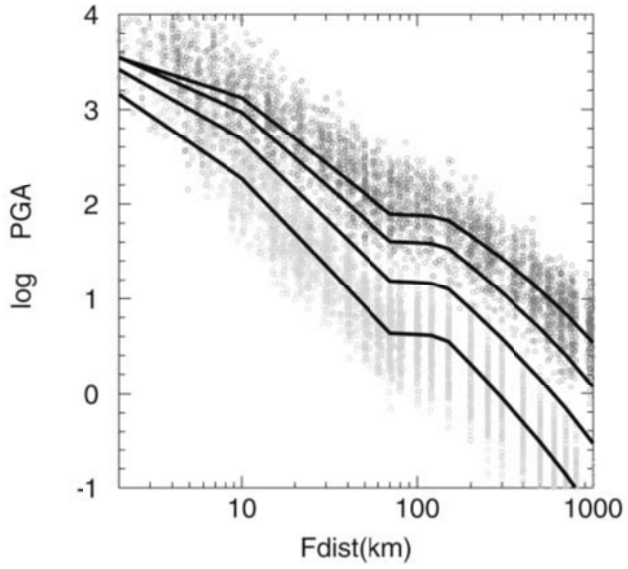

Figure 3.3 Illustration of GMPE curves with discontinuous slopes (grey dots represent data and black curves are GMPE's fitted to the data through regression analysis) (Atkinson and Boore, 2006).

3.2. Probabilistic Seismic Hazard Analysis

Kijko (2011) defines *seismic hazard* as "any physical phenomena associated with an earthquake... and their effects on land, man-made structure, and socio-economic systems that have the potential to produce a loss. It is also used without regard to a loss to indicate the probable level of ground shaking occurring at a given point within a certain period of time." Probabilistic seismic hazard analysis Kijko (2011) defines as "quantification of the probability that a specified level of ground motion will be exceeded at least once at a site or in a region during a specified exposure time". Cornell (1968) introduced what forms the basic formalism of today's PSHA in a paper titled *Engineering Seismic Risk Analysis*. Development and refinement followed in series of articles by numerous experts, causing PSHA to develop into an applied field of science (Bommer and Abrahamson, 2006; McGuire, 2008). In current practice the basic method, known as the Cornell-

© University of Pretoria

Mcguire procedure (Kijko, 2011), includes four steps (Reiter, 1990 in Kramer, 1996; Reiter, 1990 in Kijko, 2011):

1) Identify sources (faults or areas) and assign a spatial probability distribution to each source (in practice a uniform distribution is usually assigned).
2) Specify the temporal distribution of earthquake occurrence (referred to as a recurrence law) for each of these sources. This is usually done by assuming earthquakes are generated according to a Poisson process with recurrence period λ (a Poisson process is just a process for which no event in time depends on previous events – a memoryless process). The most common recurrence law is the Gutenberg-Richter relationship, or modifications thereof (see e.g. Kijko, 2011; Gupta, 2002; Utsu, 1999; Kramer, 1996). According to the Gutenberg-Richter relation the logarithm of the frequency of exceedance, λ_m, of a given magnitude m is linearly related to magnitude, m:

$$\log(\lambda_m) = a - bm \tag{3.1}$$

where a and b are constants dependent on the source.
3) Determine the ground motion prediction equation (GMPE) applicable to the area along with the uncertainty on the GMPE (Kijko, 2011). Traditionally a GMPE incorporates the decrease in ground motion as seismic waves travel from the source to the site of interest. Recently scientists and engineers started to incorporate many factors that have an effect on the ground motion that an earthquake of a given magnitude causes at a site (Gupta, 2008).
4) The previous steps' information is incorporated by probabilistic formalisms into a single cumulative probability distribution of ground motion for a specified site.

Delineating seismic sources and assigning a spatial probability distribution to them is usually a very uncertain exercise. Some researchers (Wald and Heaton, 1994; Kagan and Knopoff, 1980; Kagan, 1981; Kagan, 2006; Ben-Zion and Sammis, 2003; McGuire, 2001) have done extensive research on determination of spatial probability distributions of earthquake sources, but common practice is to delineate a seismogenic zone (2 dimensional feature) or a fault (1 dimensional feature) and assume a uniform distribution of earthquakes over it (Kramer, 1996). The details of these spatial distributions are very complex indeed (Kagan, 2006; Ben-Zion

© University of Pretoria

and Sammis, 2003). McGuire (2001) describes a recursive technique whereby only the main faults or zones are investigated in detail.

The Gutenber-Richter frequency-magnitude relationship can also be written as

$$P[M \leq m] = F(m) = \begin{cases} 0, & m < m_0 \\ 1 - \exp[-\beta(m - m_0)], & m_0 \leq m \leq m_{max} \\ 1, & m > m_{max} \end{cases} \quad (3.2)$$

where m_0 is the lowest magnitude value of interest and β =2.303b (see Kramer, 1996 for more details on the derivation of equation (3.2)). If plotted on a graph with a logarithmically scaled ordinate, the result is a line with slope $-\beta$. For small to intermediate magnitudes relation (3.2) suffices, but it is commonly accepted that arbitrarily large earthquakes are not possible (eg. Kijko and Graham, 1998; Gupta, 2002; Kramer, 1996). It has also been verified empirically that some large earthquakes with very low recurrence frequencies possibly do not follow equation (3.2) (Pisarenko and Sornette, 2004). If it is assumed that relation (3.2) stops abruptly at some magnitude (referred to a sharp cutoff), the relation changes to

$$P[M \leq m] = F(m) = \begin{cases} 0, & m < m_0 \\ \dfrac{1 - \exp[-\beta(m - m_0)]}{1 - \exp[-\beta(m_{max} - m_0)]}, & m_0 \leq m \leq m_{max} \\ 1, & m > m_{max} \end{cases} \quad (3.3)$$

where m_{max} is the upper bound of magnitude values an earthquake may take on; it is merely rescaled by the factor $(1 - \exp[-\beta(m_{max} - m_0)])^{-1}$ so that the value of the cdf is 1 at m_{max}.

Another common modification of the G-R law is the so-called characteristic earthquake model (Youngs and Coppersmith, 1985). This model assumes the exponential distribution (the distribution in eq. (3.2)) over most of the range of magnitudes, but an anomalously high recurrence rate for an interval of the largest earthquakes (these are the *characteristic* earthquakes). Figure 3.4 shows a theoretical characteristic earthquake distribution, and Figure 3.5 some real life examples of data suspected to follow a characteristic earthquake distribution from Youngs and Coppersmith (1985).

© University of Pretoria

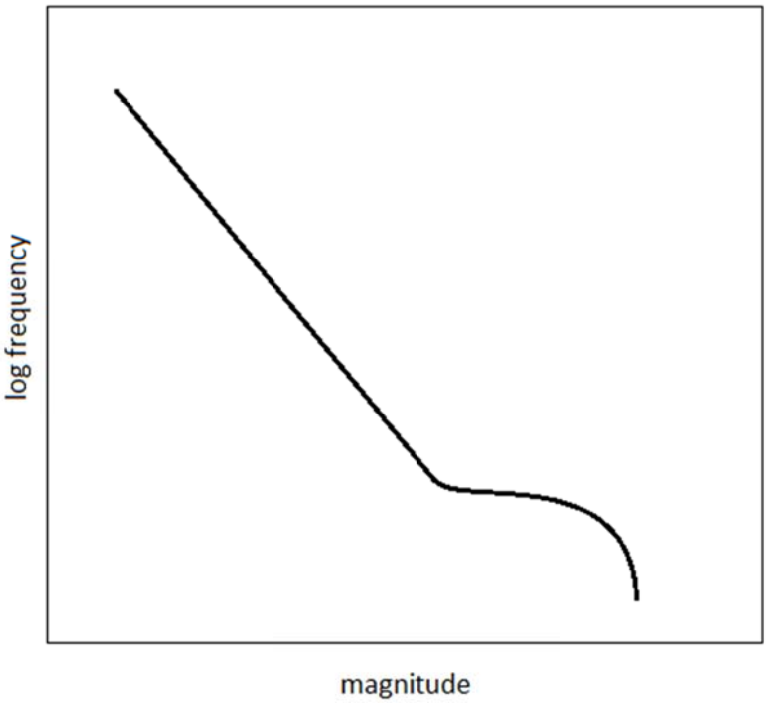

Figure 3.4 Theoretical shape of characteristic earthquake model (Adapted from Kramer, 1996)

© University of Pretoria

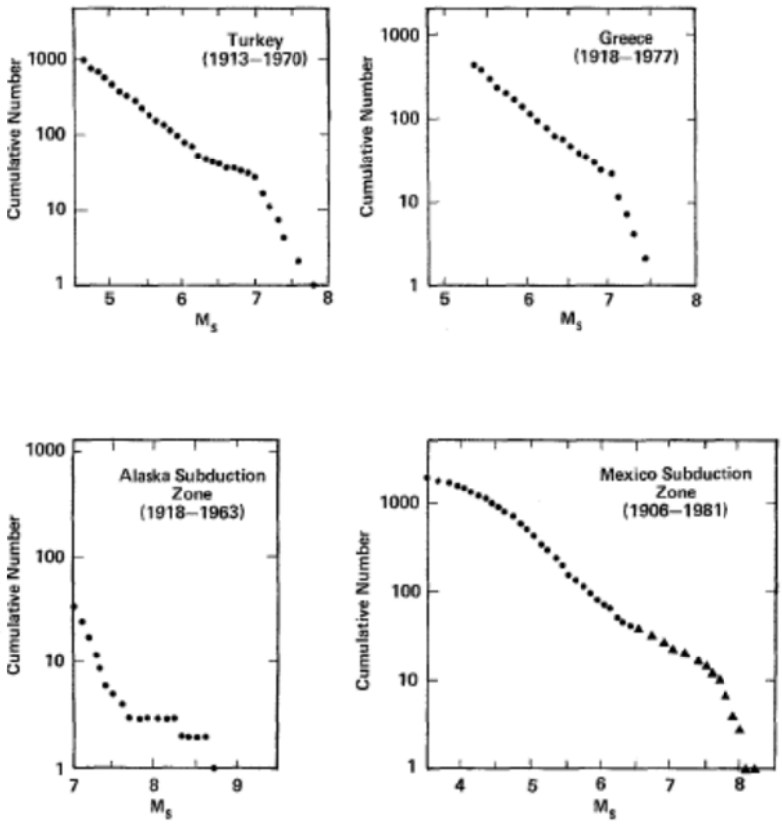

Figure 3.5 Magnitude data suspected to follow a characteristic earthquake distribution (Youngs and Coppersmith, 1985)

In general the assumption that events follow a Poisson process in time (i.e. a *memory-less* process, as discussed previously) may not strictly be correct. It is a simplified assumption, but applies more or less when foreshocks and aftershocks are not considered (Kramer, 1996), and may even be a good approximation in most cases for the catalogue as a whole, as discussed later in this section.

Ground motion prediction equations will be discussed in detail in section 3.2; suffices it to say that theoretical models have been developed based on simplified assumptions such as a spherically symmetrical seismic source process and homogeneous propagation media that do not capture the full complexity of the problem (Hong and Goda, 2007). Because of the extreme complexity of theoretical considerations beyond the simple model just mentioned, empiricism has become the norm (see e.g. Power *et al.* 2008; Douglas, 2011; Campbell, 2003).

Reiter's step 4 is the combination of the data from the previous steps to determine the probability of exceeding a ground motion parameter at least once. The results are usually presented as a *hazard curve*, which is the complement of the cdf, also known as a *survivor curve*. Figure 3.6 (Kijko, 2011) shows an example of a hazard curve. The result is for a specific site, but

© University of Pretoria

analyses may be done for a grid of points to produce a *seismic hazard map* (Figure 3.7, Giardini *et al.*, 1999) at a given probability of exceedance, or recurrence period.

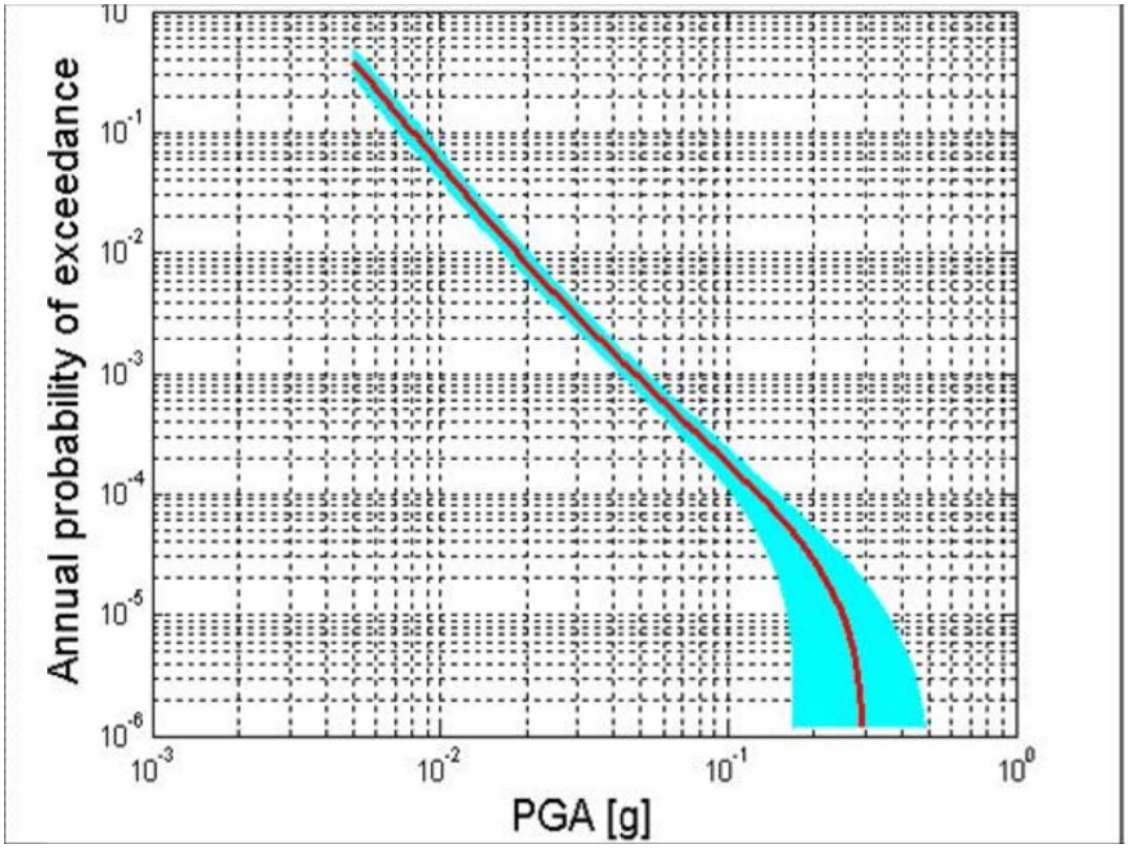

Figure 3.6 Hazard curve: probability of exceedance of levels of PGA (Kijko, 2011). Blue shading indicates a confidence interval.

© University of Pretoria

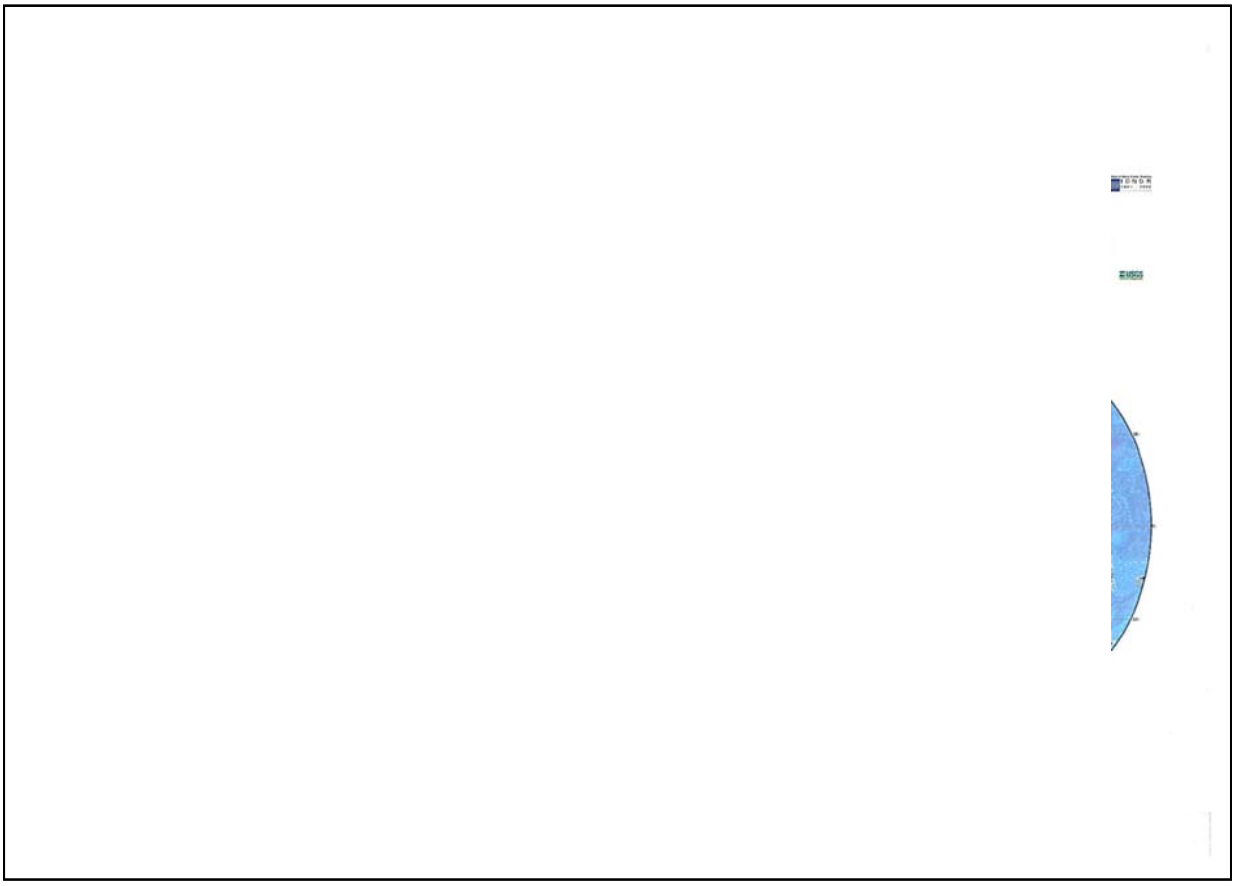

Figure 3.7 Global seismic hazard map of PGA expected at 10% probability of exceedance in 50 years (Giardini *et al.*, 1999).

The procedure by which all contributing probability factors are combined, also known as the *theorem of total probability* (McGuire, 2008) states that if the probability of an event A varies with the outcome x_i of some process X, then the actual probability $P[A]$ of A happening is calculated by the sum

$$P[A] = \sum_{i=1}^{n} (P[A|X = x_i] \times P[X = x_i]) \tag{3.4}$$

where $P[A|X = x_i]$ is the probability that A happens if the outcome of X is x_i, and $P[X = x_i]$ is the probability that the outcome of the process X is x_i. If the collection of all possible outcomes of X is an interval on the real line, the continuous (integral) version is used

© University of Pretoria

$$P[A] = \int_{x_{min}}^{x_{max}} P[A|x] f(x) dx \tag{3.5}$$

where $f(x)$ is the differential probability that x is the outcome of X. For more details on this topic see Soong (2004), especially Chapters 2 and 3, or any other book on the mathematical fundamentals of probability theory.

This formalism is applied to PGA by the integral

$$P[pga > a] = \int_{m_{min}}^{m_{max}} \int_{r_{min}}^{r_{max}} P[\,pga > a|m,r] f_{m,r}(m,r) dm dr \tag{3.6}$$

where $f_m(m)$ is the differential probability (or probability density function - pdf) of magnitude, and $f_r(r)$ is the differential probability of distance. Although the integrand in equation (3.6) may vary with measure of distance (e.g. epicentral distance, hypocentral distance, closest distance to ruptured surface), the general form holds for all of these. If one needs to incorporate different source zones with different rates of seismicity λ, the total probability is calculated as:

$$P[pga > a] = \sum_{i=1}^{n} \frac{\lambda_i}{\sum_{j=1}^{n}(\lambda_j)} \int_{m_{min}}^{m_{max}} \int_{r_{min}}^{r_{max}} P[\,pga > a|m,r] f_{m,r}(m,r) dm dr \tag{3.7}$$

Uncertainty in the calculation of the total probability

A word on the philosophy of uncertainty of primary importance in earthquake seismology is in order. Uncertainties are divided into those of truly stochastic nature (at least in the current paradigm and model frame of reference), called *aleatory* uncertainty; and those that are caused by scientists' lack of knowledge, called *epistemic* uncertainty (Strasser *et al.*, 2009). Aleatory uncertainty cannot be improved by including more data. An interesting example of this would be the b value in the G-R recurrence law. Many scientists consider that it varies with time (e.g. Kijko and Graham, 1998; Cao and Gao, 2002; Latchman *et al.*, 2008), which is a stochastic process and belongs

© University of Pretoria

to the category of aleatory variability (uncertainty). As an example of epistemic uncertainty, consider Figure 3.5. The slope (which is characterised by the b-value) clearly varies with magnitude, but if insufficient data is available one cannot distinguish subtle variations in the slope. The variation in the b-value for different values is there, but cannot be specified due to a lack of knowledge. This lack of knowledge may be decreased by the accumulation of more data. The aleatory uncertainty is incorporated into (3.7) by integration over the uncertainty in the same way as the other variables. Equation (3.7) is usually assumed to represent the central trend. Kijko and Sellevoll (1989 and 1992) give a detailed approach for the incorporation of Gaussian uncertainties in the magnitude distribution, as discussed in Section 5.1. Strasser *et al.* (2009) characterise in detail the incorporation of aleatory uncertainties in ground motion, but is in this work it will be assumed that the uncertainties in log-PGA data is normally distributed (or, less strictly, that it follows a symmetrical distribution).

In essence, epistemic uncertainties are also incorporated by the use of theorem of total probability by considering different possible models and assigning a probability weighting to each model. The *logic tree formalism* which is used to accomplish this incorporation of epistemic uncertainties is a popular method in the practical application of PSHA. This formalism assigns probability weights (relating to plausibility based on expert judgement) to given starting point assumptions relating to the seismic source distribution. Alternative GMPE's are given weights considering each of the alternative sets of assumptions relating to the seismic sources. The weights of the alternative GMPE's when considering one set of source assumptions to be true should add up to one. All the different cases of GMPE and source assumptions are considered individually, and weights are assigned to different models of the response of the local ground material for each scenario such that the weights of the respective ground response models add up to one. The process is repeated for alternative models of aleatory uncertainty. The result of this exercise may be summarised in a figure called a *logic tree*. Figure 3.8 depicts a hypothetical logic tree. (Kramer, 2006; Gupta, 2008; Kijko, 2011)

© University of Pretoria

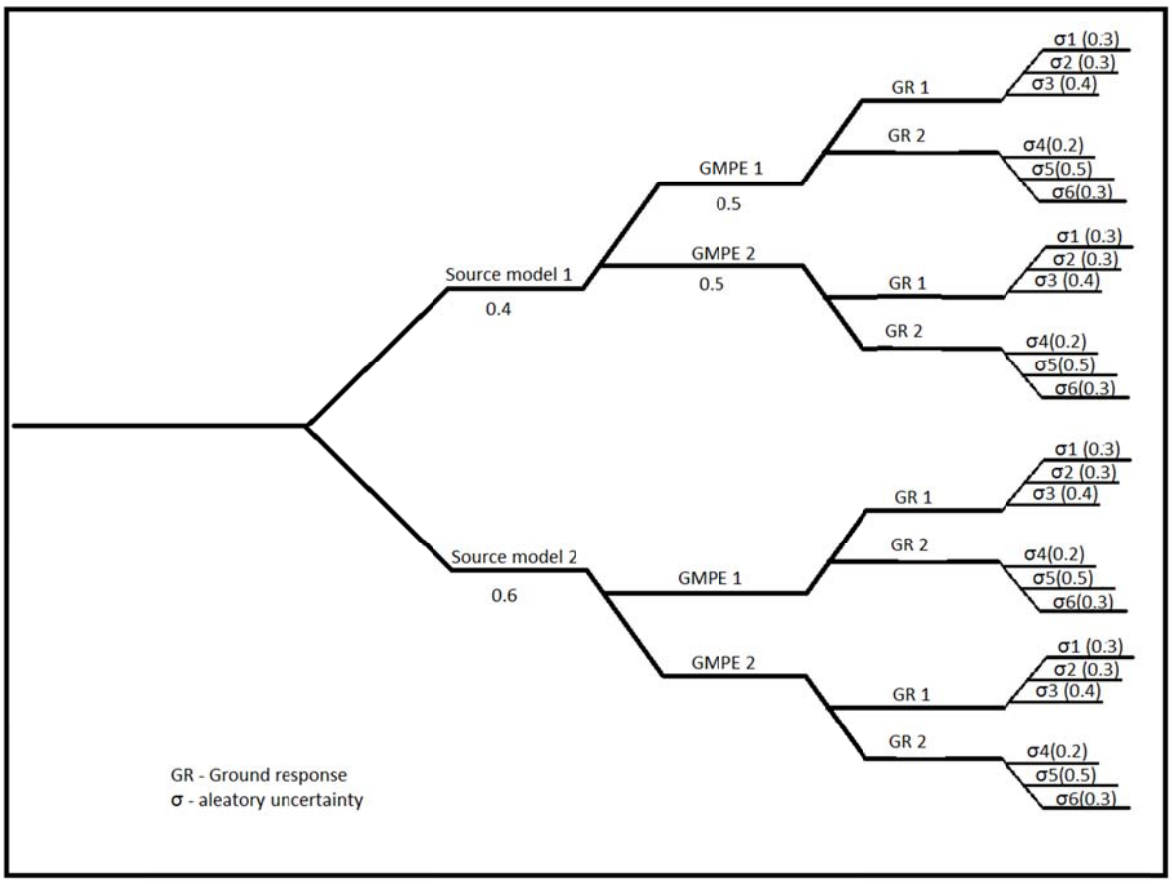

Figure 3.8 Hypothetical Logic tree

A short note on the applicability of the Poisson process to seismic hazard

The Poisson process may be shown to maximize information entropy when only the time rate (rate of seismicity in the case of PSHA) is specified (Kapur and Kesavan, 1992). Cornell and Winterstein (1988) investigated how conservative the assumption of a Poisson process is. Specifically, they investigated the performance of the Poisson process compared to semi-Markov processes where the following dependencies apply respectively and/or in combinations:

(1) the marginal probability of the time lapse from the present to the next event depends on the time lapse since the previous event to the present
(2) the marginal probability of the time lapse between the previous event and the next event depends on the magnitude of the previous earthquake
(3) the marginal probability of the magnitude of the next event depends on the time lapse since the previous event to the next event
(4) the marginal probability of the magnitude of the next event depends on the magnitude of the previous event

Cornell and Winterstein (1988) concluded that for single or a small amount of seismogenic entities (features such as faults or fault segments that display the dependencies in points 1-4 above) the Poisson process models the hazard conservatively for time lapses between successive events that

© University of Pretoria

are not too long relative to the expected time between successive events (i.e. when no large seismic gaps occur). They also concluded that, for an increasing amount of seismogenic entities, the Poisson process approximation becomes increasingly conservative.

The Pareto distribution as a statistical model for seismic energy release

It is shown here that seismic energy release follows a Pareto (also referred to as a power law) distribution under the assumption of the Gutenberg Richter relation.

Seismic energy E is related to magnitude by a relateion of the form

$$m = d\,\log(E) + c \tag{3.8}$$

where c and d are constants. Let $e \coloneqq 10^{\frac{m-c}{d}}$. Throughout the rest of this section it should be understood that only $m\epsilon[m_{min}, m_{max}]$ is considered here; it is implicitly assumed to avoid cluttered equations and relations. Substituting equation (3.8) in equation (3.2) as *the random variable* in equation (3.2) (that is, m in 3.8 corresponds to m in equation (3.2)), we get the following sequence of implications

$$\begin{aligned}
& P[d\,\log(E) + c \le m] = 1 - \exp[-\beta(m - m_0)] \\
\therefore \quad & P[E \le 10^{\frac{m-c}{d}}] = 1 - \exp[-\beta(m - m_0)] \\
\therefore \quad & P[E \le 10^{\frac{m-c}{d}}] = 1 - \exp[-\beta(m - m_0)] \\
\therefore \quad & P[E \le e] = 1 - \exp[-\beta(d\log(e) + c - m_0)] \\
& \qquad = 1 - \exp[-\beta(d\log(e) + c - m_0)] \\
& \qquad = 1 - ke^{-\frac{\beta d}{\ln(10)}}
\end{aligned} \tag{3.9}$$

where $k = \exp[-\beta(c - m_0)]$. The result $P[E \le e] = 1 - ke^{-\frac{\beta d}{\ln(10)}}$ is a Pareto distribution. Given the expression $m_0 = d \cdot \log(x_{min}) + c$ it is easily verified that the result in the set of implications 3.9 may be written in the form

© University of Pretoria

$$\mathrm{P}[E \leq e] = \left(\frac{e}{e_{min}}\right)^{-\frac{\beta d}{\ln(10)}} \tag{3.10}$$

where, e_{min} is the level of completeness in terms of energy. The truncated version of equation (3.10), where energy is considered to have a maximum possible value it may take on, may be expressed as:

$$\mathrm{P}[E \leq e] = \frac{(e_{min})^{-\frac{\beta d}{\ln(10)}} - (e)^{-\frac{\beta d}{\ln(10)}}}{(e_{min})^{-\frac{\beta d}{\ln(10)}} - (e_{max})^{-\frac{\beta d}{\ln(10)}}} \tag{3.11}$$

Ground motion

At a given site an earthquake manifests in ground motion, which is the focus of most of engineering seismology and the goal of most seismic hazard analyses. Energy propagates as elastic waves from the source to a site under consideration where ground shaking is experienced. In a simplified continuum regime the value of a measure of ground motion is determined by two parameters: (1) the earthquake magnitude, related to the amount of energy released at the source; and (2) the distance from the source, as energy spreads out over a wider surface as the wave front moves further away from its source. The energy is not only reduced by this spreading out over a greater surface area with distance; energy is also dissipated, or absorbed, as heat energy by the medium (the medium is not completely *elastic*). All these effects are quantified by a *ground motion prediction equation* (GMPE) is typically to be of the form (Cornel, 1969; Boore and Joyner, 1982; Kijko and Graham, 1999 etc.):

$$\ln(Y) = c_1 + c_2 M + \phi(r) + \varepsilon \tag{3.12}$$

where Y is a ground motion parameter such as PGA, PGV or PGD, r denotes distance from source to site, $\phi(r)$ is some function of r, ε denotes the uncertainty (or residual) and the c_i's are positive constants which are to be determined empirically. $\phi(r)$ is often assumed to take the form

© University of Pretoria

$-c_3 r - c_4 \ln(r)$ and the residual term ε to follow a normal distribution (Kjko and Graham, 1998; Campbell, 2002; Lay and Wallace, 1995). Such a functional form is typical for a spherically symmetrical explosion or implosion in a homogeneous propagation medium (Lay and Wallace, 1995).

Because of the deviation of different focal mechanisms from a spherically symmetry, equation (3.8) does not hold exactly – or, in the least, the constants will vary with the azimuth relative to the fault strike. In addition, the radiation pattern of compressional waves and shear waves also differ (Convertito and Herrero., 2006). The so-called double couple mechanism gives rise to a four leaf pattern where, if the fault trace is put on one axis, a leaf falls in each quadrant for P waves, and a leaf falls on each axis for S-waves. In addition, the double couple mechanism gives rise to a compressional P wave in two opposite quadrants, and a dilational wave in the remaining two opposite quadrants (Lay and Wallace, 1995) (see Figure 3.9).

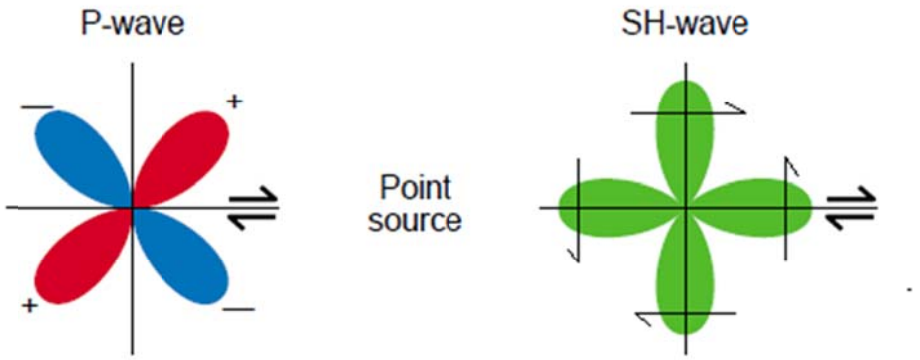

Figure 3.9 Radiation pattern of an infinitesimally small double couple mechanism (adapted from Robinson *et al.*, 2001)

Effects of ground motion on engineering structures are complex; structures respond to different frequencies by resonating to or damping the motions of these different frequencies to different degrees. Housner (1941), Biot (1932), and Biot *et al.* (1943) did pioneering work on these effects. Seed *et al.* (1976) and Trifuncac and Lee (Trifunac and Lee, 1985a and 1985b, In Trifunac, 1991; Trifunac 1976a and 1976b; Trifunac and Lee 1989; Trifunac, 1991) did extensive research on the response of local soil material on different frequency spectra. The ground motion parameter PGA fundamentally relates to the force a seismic motion exerts on a unit mass in contact with ground surface, but because of the complexity of oscillatory behaviour of engineering structures PGA does by no means adequately describe the effects on engineering structures, so PGV and PGD are commonly used measures as well (e.g. Bommer and Abrahamson, 2006; Gupta, 2002; Campbell,

© University of Pretoria

2003; Trifunac, 1991). Furthermore the SD, SV, and/or SA (which are displacement, velocity, and acceleration amplitudes of different frequencies, respectively) are often reported in seismic records to account for effects that Trifuncac and Lee (eg. Trifunac, 1991) addressed concerning the frequency response of different soil material.

Authors such as Sommerville (2003) and Hong and Goda (2007) emphasise that much of the variability in ground motion is due to fault orientation relative to a specific location as well as the fault mechanism. Bommer *et al.* (2004) also point out that all the complicated details of rupture propagation affect the nature of the radiation pattern. Fault movement does not take place simultaneously across a whole fault surface; movement only proceeds along a *rupture front*, where failure through irregularities and coalesced discontinuities takes place (Bommer*et al.*, 2004; Aki, 1984; Ben-Zion, 2008). Local areas (referred to as asperities by Aki, 1984) of very high rupture velocity create an effect much like a sonic boom that concentrates a large amount of energy a single pulse. Burridge (1973) concluded that it is possible for fault rupture to proceed at super-shear wave velocity. Such rupture, or even rupture slightly below shear wave velocity, causes energy to be 'bunched up' in a single pulse in the direction of rupture propagation (Sommerville, 2003). This leads to what is referred to as rupture *directivity effects*, which is usually perceived in the near field in the direction in line with the fault and in the direction which rupture takes place (Sommerville, 2003). Note that earthquake magnitude measures do not take such energy concentrating effects into account. Effects such as the fault mechanism, rupture pattern and directivity effects may, to a large extent be the cause of intra-event variability in light of works such as those of Sommerville (2003), Convertito and Herrero (2004), Strasser and Bommer (2009) and Spagnuolo *et al.* (2012). Figure 3.10 illustrates directivity effects; note how different rupture propagation directions distort the pattern of the point source model in Figure 3.9. The most extreme motion caused by directivity effects only tend to play a role at sites located close to the source, which is largely due to nonlinear dissipation of energy with large movements in soil (Beresnev and Wen, 1996).

© University of Pretoria

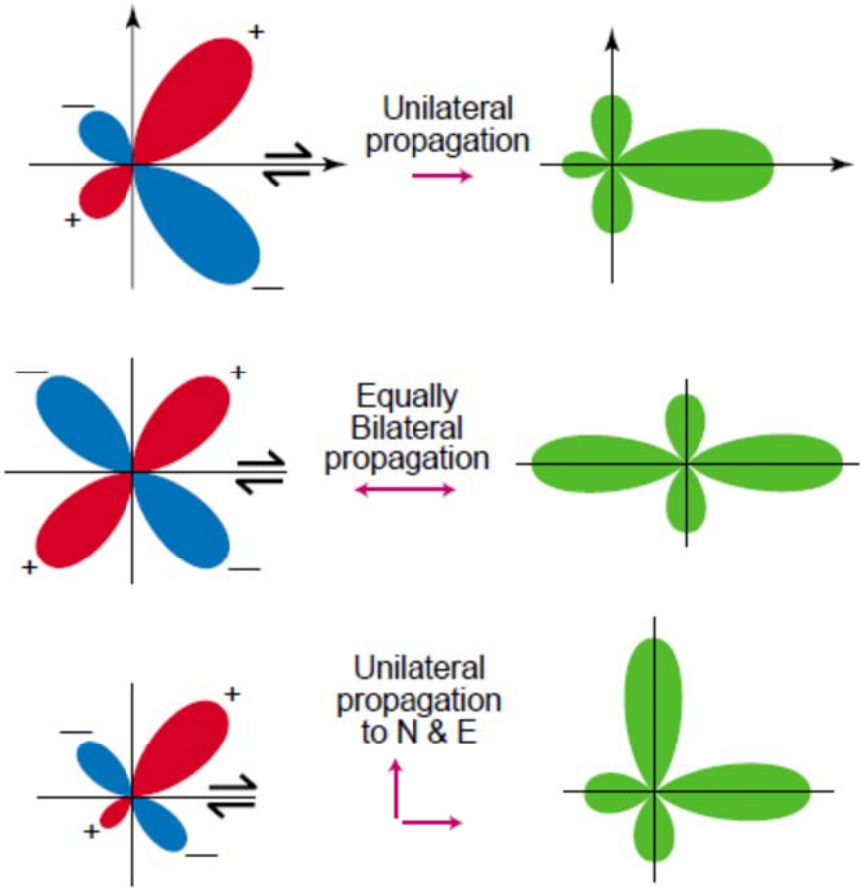

Figure 3.10 Directivity effects. Note how the pattern distorts compared to the point source pattern in Figure 3.9 (Robinson *et al.*, 2001)

It was recognised as early Anderson and Luco (1983) that the focal mechanism (dip angle and rake of rupture process) also affects peak ground motion attenuation (or scaling). However, it was only in 2004 (Convertito and Herrero, 2004) that a method was put forth to take the dip and rake angles into account theoretically. The method is only stated very generally and exact results are presented as graphs for the generic mechanisms (normal, strikeslip, and thrust) suggesting that closed form solutions are not even available for these. The general method, however, is simple: it is a weighting factor as the radiation pattern averaged over all possible strikes (in the sense of *a priori* information), divided by the average of all possible (in the sense of minimal information) focal mechanisms, radiation patterns, and distances. This is not included in GMPE's, but rather the predicted value of ground motion resulting from a GMPE not accounting for these effects is considered to be the average, and the value predicted by a particular GMPE is multiplied by the weighting factor. The authors note that the general concept of fault style is not taken into account

© University of Pretoria

by this, and that the overall tectonic stress regime may still have an influence on the peak ground motion parameters (Convertito and Herrero, 2006).

As in optics, curved interfaces between media with different wave propagation velocities can scatter and focus waves both by reflection and refraction – phenomena known as caustics. Rial (1984) proposed that this might be applicable in seismology, and its effects are well recognized at present, although usually not stated as simply as Rial (1984) did by using ray theoretical concepts. At present continuum mechanical methods seem to be preferred over ray theoretical methods (compare Chen *et al.*, 2012; Tsaur and Chang, 2008; Tsaur and Chang, 2009; Boore, 1973). These effects are the well-known topographical effects that are caused by basin boundaries, valleys and ridges that are commonly taken into account in seismic hazard analyses and earthquake engineering.

Because fault mechanisms are often more complex than a spherically symmetric explosion source, and because travel path and ground response effects are not yet well understood, ground motion prediction equations have been extended by empirical means. The Pacific Earthquake Engineering Research Center (PEER) undertook an ambitious and eventually prolific project in which groups of researchers developed empirical GMPE's from a large database of strong motion records (specifically for shallow earthquakes in active tectonic regions of North America and areas with similar tectonic regimes). The equations that have been developed cover PGA as well as SA (peak acceleration at different spectral intervals). In spite of the fact that development of these GMPE's were empirically driven, they were not entirely neglectful of current advances in earthquake physics: regression analyses on data took place with parameters and functional forms reflecting knowledge of earthquake physics, yet keeping the equations simple enough to ensure practical applicability. Factors they took into consideration to a greater or lesser extent were the special consideration of earthquakes at close distances, attenuation at close and far distances, directivity effects, wave trapping effects of the hanging wall (top wedge of a dipping fault), the style of faulting, depth to the fault if it does not intercept ground surface, rupture area, site effects, and basin trapping resonance effects. (Power *et al.*, 2008)

© University of Pretoria

4. Literature review part 1: boundedness of peak ground acceleration

4.1. Boundedness of peak ground acceleration in the traditional PSHA Procedure

The problem of assessing PGA_{max} is something of a controversial subject at present in engineering seismology and has been identified as the *missing piece* in SHA (Bommer, 2002). In a comprehensive article, Strasser and Bommer (2009) record what an enigma the question has caused during a period of a few decades while more data was accumulating and the formalisms of seismic hazard analysis were being developed. As the maximum observed value of PGA increased, so did the estimated upper bound PGA_{max}. In the course of the past few decades the prior estimates were exceeded in due course (but at least the estimates seemed to keep ahead; Bommer *et al.,* 2004). A solution to this enigma is not just of theoretical scientific interest, but also necessary for design values of facilities such as nuclear power plants, nuclear storage facilities and other structures that are hazardous when damaged (Strasser and Bommer, 2009). At Yucca Mountain Nuclear Storage Facility researchers attempted in effect to bypass the problem in the Cornell McGuire procedure by extrapolating the observed distribution to an arbitrary but extremely low annual return frequency. This, however, led to values considered physically impossible (Andrews *et al.*, 2007) (values as high as 20g is mentioned by Corrandi, 2003). This spurred a new wave of research on the subject, with much emphasis on physical, deterministic bounds (Hanks *et al.*, 2006; Bommer *et al.*, 2004; Andrews *et al.*, 2007; Strasser *et al.*, 2009). Andrews *et al.* (2007), for instance, have shown that a given fault can only release a certain amount of energy determined by the current stress state, and that the medium through which waves propagate can only support limited amplitudes of spectral ground motion determined by the strength of the medium material. They identified and investigated three controlling physical factors contributing to PGA_{max} at a specific site, *viz*. the maximum seismic energy and the radiation that can emanate from a given source, the effects of the source-to-site path, and the maximum ground motion the geological material at the site can sustain given their shear strength.

Despite the certainty that PGA_{max} exists, the great scatter of data makes determination and verification of PGA_{max} daunting and controversial. In fact, Strasser and Bommer (2009) consider that it is mainly this scatter that the Cornell-Mcguire procedure has failed to bind. Strasser *et al.* (2009) devote a paper to the discussion of the uncertainty in the ground motion prediction equations. Uncertainty is divided into an epistemic component and an aleatory component (Atkinson, 2011). The aleatory component, in turn, is divided into one component of inter-event uncertainty and another of intra-event uncertainty (Strasser and Bommer, 2009). The inter-event uncertainty is

© University of Pretoria

caused by a variation from one event to another and is present even at a single location. Intra-event uncertainty is caused by the variation from one site to another and is present even for a single earthquake. Inter-event uncertainty is interpreted to be due to the details of the rupture process not reflected by the magnitude or other measures in use, and the intra-event uncertainty due to differences in the travel path and site specific conditions.

Much of the intra-event variability is caused by the damping effects of the paths seismic waves travel along to the site (Kramer, 1996), as well as both the damping and resonation effects caused by the unconsolidated geological material at the site (Ambraseys, 1970). A layer of loose soil over hard bedrock constitutes a large impedance contrast in acoustic terms, which allows for resonance of the soil layer (energy effectively becomes trapped within the soil layer). If soil is contained within a basin-shaped structure, this energy-trapping effect is even more pronounced (Stewart *et al.*, 2002).

Site specific conditions that have an effect on the amount of ground motion are due to impedance difference (reflected by the difference in stiffness) of the bedrock and the overlying soil, the geometry of the soil deposit-bedrock interface (Ambraseys, 1970; Steward *et al.* 2002), and the surface topography (Rial, 1984).

A sedimentary layer of low shear wave velocity is analogous to an oscillator fixed at one end; it can be modelled by an elastic continuum of finite thickness that is fixed at the bottom and free at the top (Gazetas, 1982). It has harmonic frequencies at which it resonates and thus amplifies motions of specific frequency content. This effect is well known in seismology and is known as *seismic site effect.* At very high amplitudes of ground motion soils start to behave as a visco-elastic continuum, so that energy is dissipated and the amplification effect is reduced (Beresnev and Wen, 1996). Ambraseys (1970) explains that soil responds in an elastic fashion if the soil is stiff and/or at small SA amplitudes are small. However, for looser soils and/or larger SA the soil starts to dissipate energy and amplification is less pronounced than expected. Figure 4.1 shows this in a qualitative way (stiffness, as here referred to, is related to the undrained shear strength of the soil divided by the overburden pressure). Concerning the effects of topography, Rial (1984) explains that concave bedrock topography acts in analogous way to optical lenses, focussing rays onto so-called caustic surfaces. These caustic surfaces intersect ground surface to form patterns such as those illustrated in Figure 4.2 (Figure 4.2 displays the results of simulations done by Rial, 1984). In the extreme case, the intersection with the caustic surface with the ground is theoretically a point (analogous to the focal point of an optical lens). Steward *et al.* (2002) explains that, from a ray theoretical perspective, a layer of soil over bedrock in effect traps the rays inside the layer of soil because of the much lower

© University of Pretoria

impedance of the soil layer. If the bedrock forms a basin this effect is much more pronounced, because wave energy becomes trapped in the soil in the basin structure.

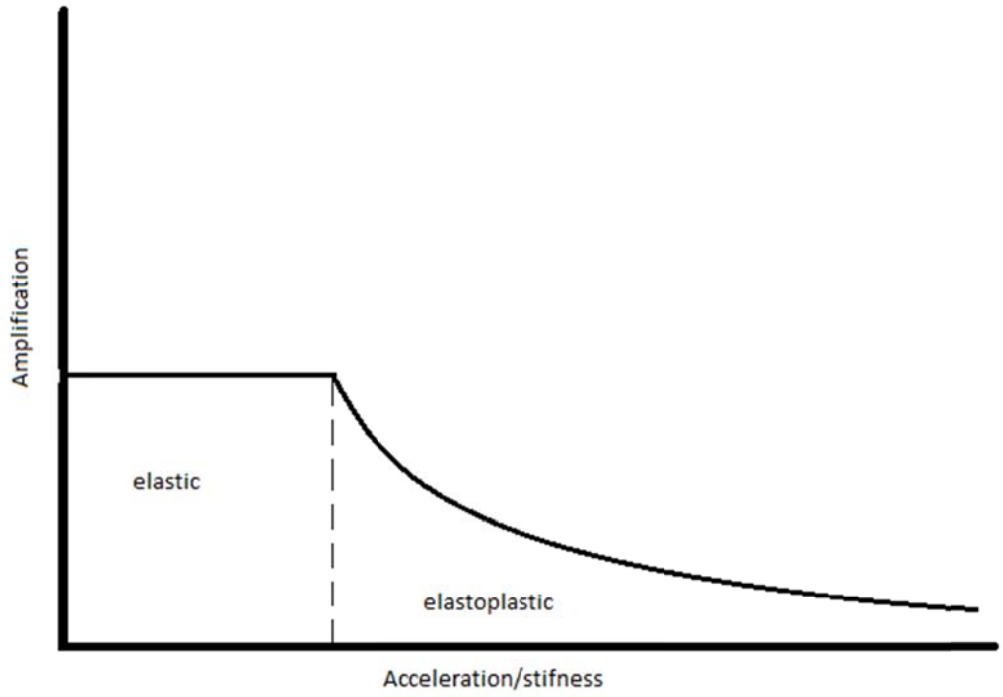

Figure 4.1 Nonelastic, energy dissipative effect of soil on amplification of seismic energy. (Adapted from Ambraseys, 1970)

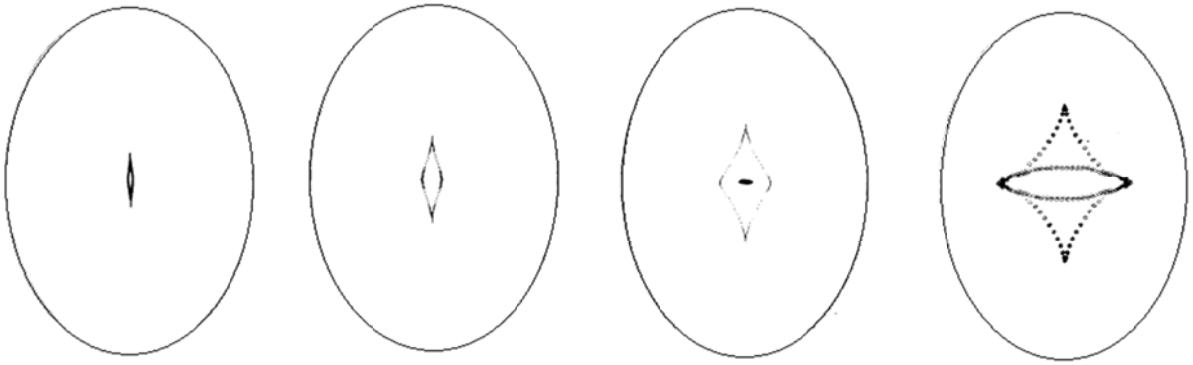

Figure 4.2 Intersection of caustic surfaces with ground surface of a sediment basin. In this simulation the source is located directly beneath the basin, and the depth of the source increases from left to right. (adapted from Rial, 1984)

© University of Pretoria

4.2. Boundedness of magnitude

The common assumption of the boundedness of magnitude itself implies that PGA has to be bounded. According to Brune (1970) the maximum near-field acceleration in a frequency interval $[0, \omega]$ may be estimated by

$$\ddot{u}(t) = \frac{1}{\pi}\frac{\sigma}{\mu}\beta\omega \cdot \left(\frac{\sin(\omega t)}{\omega t}\right) \tag{4.1}$$

where σ is the effective shear stress across a fault plane, μ is the shear modulus of the rock in which the fault is situated, β is shear wave velocity, and t is time. Consider the worst case scenario where the stress drop, $\Delta\sigma$, is equal to the effective shear stress on the fault plane. Kanamori and Anderson (1975) give the relation

$$\log(M_0) = \frac{3}{2}\log(A) + \log\left(16\Delta\sigma/7\pi^{3/2}\right) \tag{4.2}$$

where M_0 is the moment magnitude and A is the fault rupture area. Kanaromi and Hanks (1979) defined the seismic moment magnitude as a linear function of $\log(M_0)$. Given that the rupture area must be finite and the assumption that $\sigma = \Delta\sigma$, equations (4.1) and (4.2) makes it clear that finite magnitude necessarily gives rise to finite PGA, at least for finite frequencies.

Deterministic methods used to estimate an upper limit on magnitude (m_{max}) often make use of empirical relationships with rupture parameters such as length, area, etc. (Gupta, 2002). One of these is particularly elegant in that it uses slip rate, which relates to the rate of seismic energy release. This relationship, derived by Anderson (1979; in Youngs and Coppersmith, 1985), also serves as an argument of the boundedness of magnitude from a theoretical perspective, so its derivation may be considered instructive; an adaption of this argument given in Gupta (2002) is followed here.

Let $\lambda(M)$ be the average rate of occurrence of earthquakes exceeding magnitude M. $d\lambda(M)$ may be viewed as the average recurrence rate of earthquakes within a small magnitude increment containing M. Let M_0 denote moment magnitude (seismic moment can be seen as shear strain energy release, and moment magnitude is related to the logarithm of seismic moment as in equation (3.8)). The average frequency with which shear strain energy is released per unit time by events of magnitude larger than M is obtained by integrating over the rate of occurrence of events, which for this purpose is not $\lambda(M)$, but $\lambda(M_{\max}) - \lambda(M)$, the reason being that we are interested in the rate of *exceedance* involving the *complementary* cdf.

© University of Pretoria

$$\dot{M}_0 = \int_{\lambda(-\infty)}^{\lambda(M_{\max})} M_0(M) d\{\lambda(M_{\max}) - \lambda(M)\} = -\int_{-\infty}^{M_{\max}} M_0(M) \left(-\frac{d\lambda(M)}{dM}\right) dM \qquad (4.3)$$

where $-\frac{d\lambda(M)}{dM}$. Constitutes the density of events in a small magnitude increment. The reason for the minus sign follows from the fact that we used $\lambda(M_{\max}) - \lambda(M)$ as integrator instead of $\lambda(M)$. If the use of $\lambda(M_{\max}) - \lambda(M)$ instead of $\lambda(M)$ still seems strange, think of the fact that $\lambda(M_{\max}) < \lambda(-\infty)$ in the integral bounds. (For the reader who finds the use of integrals of the form $\int_a^b g(x)dF(x)$ in the following paragraphs unfamiliar, Appendix A provides an explanation in terms of the Riemann-Stieltjies definition of the integral). Substituting the relationship

$$\log(M_0) = l + kM \qquad (4.4)$$

along with the Gutenberg-Richter law (equation (3.1)) we obtain

$$\dot{M}_0 = \int_{-\infty}^{M_{\max}} bln(10) \cdot 10^{a+l+(k-b)M} dM \qquad (4.5)$$

It is assumed that the constant k exceeds the constant b which Knopoff and Kagan (1977) note is usually the case. In this case the integral is does not converge as M_{max} approaches infinity, emphasizing the need of a finite bound (if it so happens that b exceeds k there is still the straightforward constraint that the finite geometrical dimensions and finite strength of rock places on the strain energy that may be accumulated on a single fault).

From this consideration a theoretical estimate of M_{max} may be obtained. Evaluating the integral in equation (3.5) we obtain

$$\dot{M}_0 = \frac{b}{(k-b)} \cdot 10^{a+l+(k-b)M_{max}} \qquad (4.6)$$

Noting that the recurrence period T_{Mmax} of M_{max} is the inverse of the recurrence rate (or frequency) $10^{a-bMmax}$ in accordance with equation (3.1), we solve for M_{max}

$$M_{max} = \frac{\log\left(\frac{(k-b)T_{\mathrm{Mmax}}\dot{M}_0}{b}\right) - l}{k} \qquad (4.7)$$

© University of Pretoria

$\dot{M}_0$ may be determined by (Gupta, 2002)

$$\dot{M}_0 = \mu A \dot{u} \tag{4.8}$$

where μ is the shear modulus of the rock, A is the fault area (or the seismogenic, seismically active part of it), and $\dot{u}$ is the slip rate along the fault plane. T_{Mmax}, Gupta (2002) notes, is typically estimated from paleoseismic investigations.

Other deterministic, or physical estimates of m_{max} are based on empirically determined relationships between magnitude and fault parameters such as its length or area, and fault displacement (Gupta, 2002). It is easy to comprehend the validity of such relations from equation (4.8) and its related relation without the time derivative $\mathrm{M}_0 = \mu A u$, in which the fault dimensions, displacement and slip rate feature. Thus the finiteness of these parameters guarantees the finitude of magnitude.

Probabilistic estimators of m_{max} provide good prospects, but in their comprehensive review of these estimators, Kijko and Singh (2011) note that very little work has been done in this respect. In the paper just mentioned the authors discuss twelve methods to estimate m_{max}. In the following section this compilation and review by Kijko and Singh (2011) of methods to estimate the maximum possible earthquake magnitude from historical records is discussed. These methods they discuss assume a truncated distribution, or hard cutoff, as opposed to methods assuming a tapering, or soft cutoff. Other methods, some of which do consider a distribution with a soft cutoff, are discussed in section 4.2.2.

4.2.1. Twelve Methods for estimation of maximum possible magnitude

Kijko and Singh (2011) published a review of twelve methods (hereafter TM) to estimate m_{max} from historical data, all of which assume that

$$m_{max} = m_{max}^{obs} + \Delta \tag{4.9}$$

where m_{max}^{obs} is the maximum observed magnitude amongst the historical data and Δ is a positive correction term. Most of the procedures discussed in TM are based on an adaption of Cooke's (1979) method to the maximum earthquake magnitude problem, which they refer to as *the generic equation*. It assumes that the largest observed magnitude is close to m_{max}; otherwise stated, it

© University of Pretoria

requires that the largest observed magnitude would have been exceeded had the distribution been unbounded. It does assume the frequency-magnitude distribution is known, but does not specify it, so magnitude size distributions other than the Gutenberg-Richter distribution may also be used. The derivation of the generic equation follows:

Suppose magnitude is distributed according to some cumulative distribution function (cdf) $F(m)$. Consider a sequence of n random observations $(m_1, m_2, m_3, \dots, m_n)$ rank ordered in increasing order. The cdf of the largest magnitude, m_n is $[F(m)]^n$. The expected value of m_n is

$$E[m_n] = \int_{m_{min}}^{m_{max}} m \, d[F(m)]^n \tag{4.10}$$

Integrating by parts we obtain

$$E[m_n] = m_{max} - \int_{m_{min}}^{m_{max}} [F(m)]^n dm \tag{4.11}$$

Now, given an actual catalogue with n entries, the best estimate for $E[m_n]$ is just the maximum observed magnitude, m_{max}^{obs}. Substituting this and rearranging we obtain the relation

$$m_{max} \approx m_{max}^{obs} + \int_{m_{min}}^{m_{max}} [F(m)]^n \, dm \tag{4.12}$$

which is *the generic equation*. In the context of equation (4.9) we have

$$\Delta = \int_{m_{min}}^{m_{max}} [F(m)]^n \, dm \tag{4.13}$$

A brief discussion of the methods presented in the TM follows.

The first method in the TM is due to Gibowicz and Kijko (1994), which they call the Tate-Pisarenko (T-P) estimator. The expected value of the CDF at m_n, that is identified as

$$E[F(m_n)] = \frac{n}{n+1} \tag{4.14}$$

Taking $F(m_{\max}^{obs})$ as the best estimate of $E[F(m_n)]$, one obtains the estimator

$$F(m_{max}^{obs} | m_{max}) = \frac{n}{n+1} \tag{4.15}$$

© University of Pretoria

($F(m_{max}^{obs})$ is replaced by $F(m_{max}^{obs}|m_{max})$ just to make clear the interest in m_{max} as a parameter). The root of equation 4.13 is then an estimate for m_{max}. As the first method they assume the G-R law. The method is, of course, open to different magnitude distributions in the same way as the generic equation.

In the second and third methods, called the Kijko-Sellevoll (K-S) procedure in the TM, the generic equation is merely iterated, assuming it converges to a fixed point:

$$\hat{m}_{max,current} \approx m_{max}^{obs} + \int_{m_{min}}^{\hat{m}_{max,prior}} d[F(m)]^n \tag{4.16}$$

where $\hat{m}_{max,current}$ is the current estimate of m_{max}, and $\hat{m}_{max,prior}$ the previous. Although this is a very simple procedure, Kijko and Graham (1998) and Lasocki and Urban (2011) have found that it performs very well, although the latter authors recommend only a single iteration. The inherent variance of such estimator is approximately (Kijko and Graham, 1998):

$$Var(\hat{m}_{max}) = \Delta^2 \tag{4.17}$$

i.e. the square of the integral term of equation 4.15. Each iteration requires integration, but an approximation is given by

$$\Delta = \frac{E_1(n_2) - E_1(n_1)}{\beta \exp(-n_2)} + m_{min}\exp(-n) \tag{4.18}$$

where $n_1 = n/\{1 - \exp[-\beta(m_{max} - m_{min})]\}$, $n_2 = n_1\{1 - \exp[-\beta(m_{max} - m_{min})]\}$, and $E_1(z) = \int_z^{\infty} \frac{\exp(\tau)}{\tau} d\tau$. Using this approximation constitutes the second method. The third uses an exact solution which is only valid for integer values of n in equation (4.18):

© University of Pretoria

$$\Delta = \frac{m_{max} - m_{min} + \frac{1}{\beta}\sum_{i=1}^{n} \frac{(-1)^i}{i}\binom{n}{i}(1 - \exp[-i\beta(m_{max} - m_{min})])}{(1 - \exp[-\beta(m_{max} - m_{min})])^n} \tag{4.19}$$

The fourth and fifth methods in the TM are Bayesian adaptations of T-P and the K-S estimators where the b value in the Gutenberg-Richter relation is assumed to vary. For the K-S procedure this results in a probability distribution of which the cdf is given by

$$F(m) = \begin{cases} 0, & x \geq m_{max} \\ C_\beta \left[1 - \left(\frac{p}{p + m - m_{min}}\right)^q\right], & m_{min} \leq x \leq m_{max} \\ 1, & x \geq m_{max} \end{cases} \tag{4.20}$$

where

$$C_\beta = \left[1 - \left(\frac{p}{p + m_{max} - m_{min}}\right)^q\right]^{-1} \tag{4.21}$$

and p and q are related to the variance of β:

$$p = \frac{\bar{\beta}}{var(\beta)} \tag{4.22}$$

$$q = \left(\frac{\bar{\beta}}{\sqrt{var(\beta)}}\right)^2 \tag{4.23}$$

where $\bar{\beta}$ is the mean value of the varying parameter β. For the T-P estimate the root of the equation using this Bayesian version of the G-R law is

© University of Pretoria

$$m_{max} = m_{max}^{obs} + \frac{1}{n\bar{\beta}C_{\beta}}\left(\frac{p}{p + m_{max}^{obs} - m_{min}}\right)^{-(q+1)} \tag{4.24}$$

For the Bayesian variation of the G-R law in the K-S procedure no exact solution has been found, but an approximation is

$$\Delta = \frac{\delta^{\frac{1}{q}}\exp[nr^q/(1-r^q)]}{\bar{\beta}}\left[\Gamma\left(-\frac{1}{q};\delta r^q\right) - \Gamma\left(-\frac{1}{q};\delta\right)\right] \tag{4.25}$$

where $r = p/(p + m_{max} - m_{min})$, $c_1 = exp\left[-n(1-C_{\beta})\right]$, $\delta = nC_{\beta}$, and $\Gamma(.;.)$ is the complementary incomplete gamma function.

The remaining eight procedures are all non-parametric procedures, not assuming any distribution.

The sixth method may be seen as the non-parametric counterpart of the Kijko-Sellevoll procedure developed by Kijko *et al.* (2001): instead of assuming a distribution, it uses a Gaussian kernel approximation to the magnitude distribution; that is, it fits a sum of Gaussian functions to the distribution. Accordingly, Δ takes the form

$$\Delta = \int_{m_{min}}^{m_{max}} \left[\frac{\sum_{i=1}^{n}\left[\mathrm{Erf}\left(\frac{m-m_i}{h}\right) - \mathrm{Erf}\left(\frac{m_{min}-m_i}{h}\right)\right]}{\sum_{i=1}^{n}\left[\mathrm{Erf}\left(\frac{m_{max}-m_i}{h}\right) - \mathrm{Erf}\left(\frac{m_{min}-m_i}{h}\right)\right]}\right]^{n} dm \tag{4.26}$$

where m_i is the magnitude of the ith event in the catalogue, $Erf(.)$ is the cumulative distribution of the Gaussian pdf, and h is a smoothing factor. With substitution of equation (4.26) generic equation (4.12) is again solved by iteration.

© University of Pretoria

The seventh method is due to Cooke (1979). Here the empirical distribution function $\hat{F}(m)$ is used as the cumulative distribution function for the magnitude in the correction term Δ. For any sample of size n, the empirical distribution function is defined as

$$\hat{F}(m) = \begin{cases} = 0, & m < m_1 \\ = \dfrac{i}{n}, & m_i \leq m \leq m_{i+1} \\ = 1, & m > m_n \end{cases} \tag{4.27}$$

where m_i is the i[th] largest magnitude in the catalogue. (Note how the values are used in order of ascending magnitude. Statistical methods which reorders observations in such ways are also known as order statistics). Thus, delta becomes

$$\Delta = \int_{m_{min}}^{m_{max}} \left[\hat{F}(m)\right]^n dm = \sum_{i}^{n} \left[\frac{i}{n}\right]^n (m_{i+1} - m_i) \tag{4.28}$$

No iteration is required: note that m_{max} does not appear in the sum in equation (4.28). The sum is simplified by noting some applicable manipulations and the approximation $(1 - \frac{1}{n})^n \approx e$:

$$\sum_{i}^{n} \left[\frac{i}{n}\right]^n (m_{i+1} - m_i) \approx (1 - e^1) \sum_{i=0}^{n-1} e^i m_{n-i} \tag{4.29}$$

The eighth method is based on a statement by Gnedenko (1943; in Kijko and Singh, 2011) which suggests that for a broad class of cumulative distributions which are linear close to their upper end point, an estimate of the upper end point may take the form

© University of Pretoria

$$\hat{m}_{max} = \sum_{i=1}^{n_0} a_i m_{n-i+1} \tag{4.30}$$

Where $\{m_{n-n_0}, m_{n-n_0+1}, \dots, m_n\}$ are the n_0 largest order statistics. This is a useful form of an estimate if only the very largest events were recorded. Cooke (1980) uses the coefficients $a_1 = 1 + 1/n_0$, $a_2 = a_2 = \cdots = a_{n_0-1=0}$, and $a_{n_0} = -1/n_0$. Kijko and Singh (2011) suggest $a_2 = a_2 = \cdots = a_{n_0} = -\frac{1}{n_0(n_0-1)}$.

The ninth method is due to Robson and Whitlock (1964). Kijko and Singh (2011) note that this procedure is asymptotically mean unbiased, but has a large mean squared error. This estimate is given by

$$\hat{m}_{max} = m_{max}^{obs} + \left(m_{max}^{obs} - m_{n-1}\right) \tag{4.31}$$

In this case $\Delta = (m_{max}^{obs} - m_{n-1})$.

The tenth method is an improvement of the eighth, and is due to Cooke (1979). The mean squared error of Robertson and Whitlock's estimate is multiplied by a factor that is a function to the exponent in the asymptotic shape of the tail. For the Gutenberg-Richter distribution

$$\Delta = 0.5\left(m_{max}^{obs} - m_{n-1}\right) \tag{4.32}$$

The last two procedures in TM are the least squares regression and a least absolute value regression (referred to in the TM as L_2-norm regression and L_1-norm regression for least squares and least absolute value regression, respectively), used to fit an analytic function to the cdf. It is noted in TM that least squares regression is not recommended for data that is not reliable; nor should it be used if residuals are not known to be normally distributed. Least absolute values regression does not

© University of Pretoria

possess these drawbacks. Regression does, however, require a predetermined shape of the distribution.

4.2.2. Other approaches to estimating or characterizing m_{max}

Kagan and Schoenberg (2001) and Pisarenko and co-workers (Pisarenko *et al.*, 2003; Pisarenko and Sornette 2004; Pisarenko and Sornette, 2006; Pisarenko *et al*., 2008a) did extensive work on the previously mentioned soft cutoff models of m_{max}. They assume that at some value, called the corner value by Kagan (2002) and crossover magnityde by Pisarenko *et al*. (2003), the probability decreases much more rapidly than the G-R law predicts. Figure 4.3 illustrates the concept of a corner or crossover magnitude. It should be made clear that soft cutoff values are not the maximum possible values.

Kagan and Schoenberg (2001) introduce a model that they call a tapered Pareto law for the probability distribution of seismic moment. The cdf that characterises this law is given by

$$F(mom) = 1 - \left(\frac{a}{mom}\right)^{\beta} \exp\left(\frac{a - mom}{\theta}\right), \qquad mom \geq a \tag{4.33}$$

β is the traditional G-R parameter as it appears in equation (3.2), a is the level of completeness under the transformation in equation (3.8), and θ is the corner moment. Note that in this case the corner magnitude is not associated with a discontinuity in the slope, but marks the point above which considerable deviation from the traditional G-R law commences (Kagan and Schoenberg, 2001).

Kagan and Schoenberg (2001) provide the relation

$$mom \simeq 10^{\frac{3}{2}m+6} \tag{4.34}$$

© University of Pretoria

between seismic moment mom and moment magnitude m. So by substitution of relation (4.34) into equation (4.33) one obtains the corresponding frequency relation for moment magnitude.

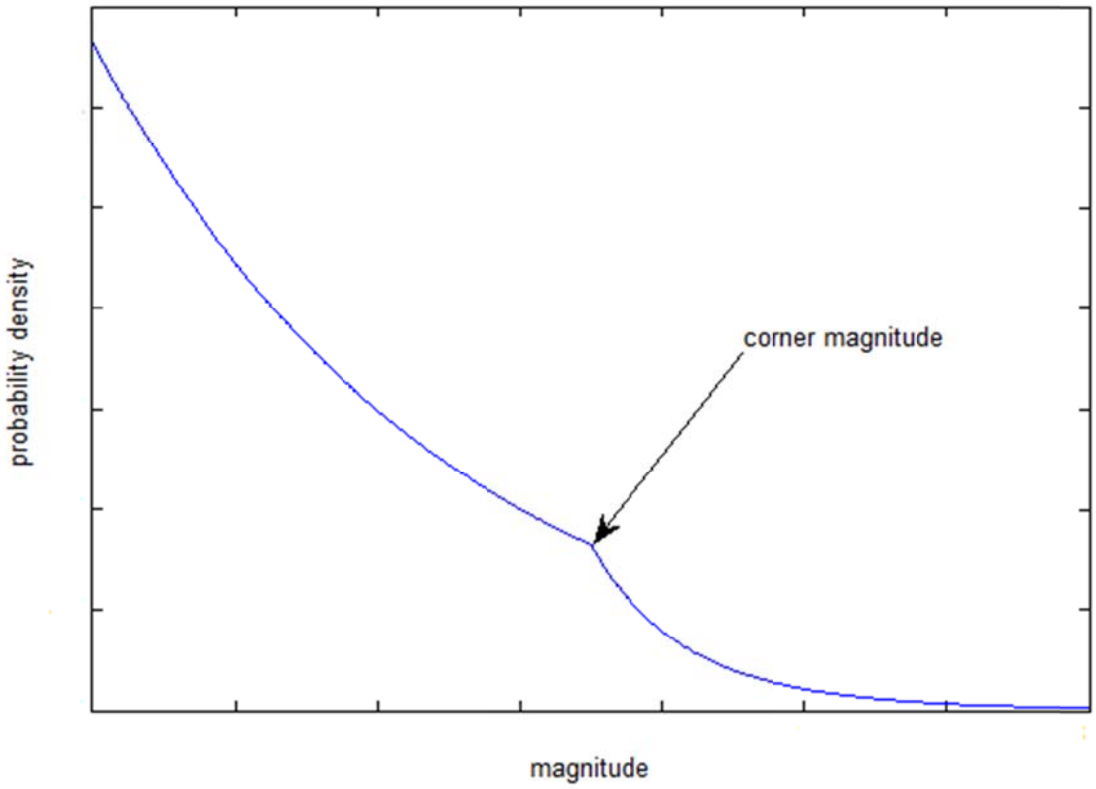

Figure 4.3 Illustration of the concept of a corner magnitude. (Take care to note that the probability density is NOT log scaled in this figure)

Pisarenko and Sornette (2004) constructed a statistical test for detection deviation from a given parametric model and a method for detection of the crossover value to a different distribution from the proposed distribution. To obtain the statistic the observations are first rank ordered to obtain $y_1 \geq y_2 \geq \cdots \geq y_{i-1} \geq y_i \geq y_{i+1} \geq \cdots \geq u$, where u is some threshold value. The rank ordered are transformed to a uniform distribution through the cdf function $F(.)$ to obtain $x_i \coloneqq F(y_i)$, under the hypothesis that the observations $\{y_i\}$ follow the distribution $F(.)$. The mean and standard deviation of x_i are given, respectively, by (Hajek and Sidak, 1967 in Pisarenko and Sornette, 2004):

© University of Pretoria

$$E[x_i] = \frac{i}{N+1} \tag{4.35}$$

$$stD[x_i] = \frac{\sqrt{i(N-i+1)(N+2)}}{N+1} \tag{4.36}$$

The set is standardized to the set $\{\rho_i\}$ by deducting the expected value and dividing by the standard deviation:

$$\rho_i = \frac{x_i - E[x_i]}{Std[x_i]} \tag{4.37}$$

According to the authors the statistic

$$\hat{\varepsilon}_N = \Gamma\left(\frac{N}{2}; \frac{\sum_{i=1}^{N} \rho^2}{2}\right) \tag{4.38}$$

is a dimensionless statistic that would give the probability of exceeding $\sum_{i=1}^{N} \rho^2$ in N observations under the assumption that $\{\rho_i\}$ is normally distributed. This is then plotted on a graph against the threshold value u, analogous to how the Hill-plot for extreme values is done.

The statistic $\hat{\varepsilon}_N$ is used to test if the hypothesis that the assumed distribution is followed throughout, or if it is to be rejected from some value and higher. $\hat{\varepsilon}_N$ is not used to place exactly the cross-over value from which deviation takes place, but rather to detect the existence thereof. If there is some value from which the assumed model is rejected, Pisarenko and Sornette (2004) propose the use of a maximum likelihood estimate for some continuous distribution where the function crosses over to a different, tapering distribution. The authors propose two different possible distributions to append at the crossover value: a Pareto distribution and an exponential

© University of Pretoria

distribution. In case studies on data from different tectonic regimes the authors found that the parameter estimates of the function above the truncation point indicate some continuation beyond the crossover, whereas in other cases it approaches a step function as is the case with a truncated distribution. They also conclude that whether the appended distribution is exponential or Pareto is immaterial for the estimation of the crossover value if the extreme tail is not densely populated.

Pisarenko *et al.* (2003) and Pisarenko and Sornette (2006) developed, as a continuation on the rationale in (Pisarenko and Sornette, 2004), a nonparametric statistic to detect deviation of earthquake energy recurrence from a Pareto distribution (nonparametric in the sense that does not require a parametric form of the function where it deviates from a Pareto distribution). This statistic, which they call the TP statistic (not to be confused with the T-P estimate in Section 4.2.1), is a linear combination of the first two 'log moments' that is equal to zero for a power law distribution:

$$TP = (E_1)^2 - 0.5E_2 = 0 \tag{4.39}$$

where

$$E_1 = \int_u^\infty \log\left(\frac{x}{u}\right) dF(x) \text{ and } E_2 = \int_u^\infty \left(\log\left(\frac{x}{u}\right)\right)^2 dF(x) \tag{4.40}$$

and u is the lower threshold value. In empirical form it is given by

$$TP = \left[\left(\frac{1}{n}\right)\sum_{k=1}^{n} \log\left(\frac{x_k}{u}\right)\right]^2 - 0.5\left(\frac{1}{n}\right)\sum_{k=1}^{n}\left[\log\left(\frac{x_k}{u}\right)\right]^2 \tag{4.41}$$

The authors give an estimate of the standard deviation:

© University of Pretoria

$$STD[TP] = n^{-0.5}\left[2E_1 \log\left(\frac{x_k}{u}\right) - 0.5\left(\log(x_k/u)\right)^2\right] \quad (4.42)$$

where E_1 is estimated from the sample according to:

$$E_1 = \left(\frac{1}{n}\right)\sum_{k=1}^{n} \log\left(\frac{x_k}{u}\right) \quad (4.43)$$

The estimate is plotted against the threshold value and compared to confidence bounds determined by Monte Carlo simulation.

Pisarenko *et al.* (2008a) have developed a method that makes use of extreme value theory, dividing a catalogue into time windows and works with the collection of maximum observed magnitude values from all the time windows. The approach they discuss uses the Generalized Extreme Value (GEV) distribution which describes the distribution of maxima as the extent of the time windows approach infinity, if such a distribution exists. To estimate m_{max} for a given length of time intervals, quantile estimates on the GEV distribution is used.

Another approach which they develop (Pisarenko *et al*., 2008b) also propose the use of the asymptotic (i.e. "limiting") distribution of the actual number of events (or, more accurately, the expected number of events) for large events, and not the asymptotic distribution for maxima in time windows. The asymptotic distribution of large events, without the use of time windows, is termed the Generalized Pareto Distribution (GPD). The GPD itself behaves asymptotically the same as the Pareto distribution (Pisarenko and Sornette, 2004). In Pisarenko *et al*. (2008b) quantile estimates are for time windows using what they call the *Lomnitz Formula* from the estimated GPD parameters. The reason for their use of extreme value theory is, for instance, the fact that the algorithms for the removal of fore- and aftershocks may not be efficient enough. They also claim that all attempts of point estimates of m_{max} in the past failed to produce satisfactory accuracy because of the scatter involved in all of these estimates. Raschke (2011) noted that the use of time windows or so-called block estimates which are required for the use of the GEV distribution make estimates relatively ineffective when the largest values in a block is not close enough to the m_{max}.

© University of Pretoria

Raschke (2011) developed an estimate based on fiducial intervals. The fiducial distribution was introduced by Pisarenko (1991) but was found inadequate by Kijko (2004). For his method, Raschke improved Pisarenko's form of the distribution. The fiducial distribution of Pisarenko (1991), (in Rashcke, 2011), is

$$P(m_{max} < z) = 1 - F^n\left(m_{max}^{obs}|z\right) \tag{4.44}$$

where $F(.|z)$ is the conditional cdf, given $m_{max} = z$, and the assumed form – in this case the truncated exponential distribution of the Gutenberg-Richter law – of the cdf. Raschke scaled this distribution:

$$P(m_{max} < z) = \frac{1 - \left[\frac{F(m_{max}^{obs}|\infty)}{F(z|\infty)}\right]^n}{1 - F^n(m_{max}^{obs}|\infty)} \tag{4.45}$$

on $[m_{max}^{obs}, \infty)$. Here $F(.|\infty)$ denotes the untruncated exponential distribution. Raschke then continues to use this distribution for obtaining an estimate of m_{max}, rather than just using it for confidence intervals. He does this by computing the expected value of the pdf corresponding to (4.45), and he notes that this estimate is asymptotically unbiased. It should be noted that this is not a point estimate, but yields a quantile-based estimate.

5. Literature review part II: the Parametric-Historic Procedure

5.1. The main aspects of the Parametric-Historic procedure for Probabilistic Seismic Hazard Analysis (Kijko and Graham, 1998 and 1999)

Kijko and Graham (1999) state that the Parametric-Historic approach addresses the problems of subjective judgement (especially relating to definition of seismic sources) and incompleteness and/or the different levels of completeness (i.e. magnitude above which the catalogue is complete) as well as inaccuracy of different parts of seismic catalogues. The problem of

© University of Pretoria

overly subjective judgement and deduction is eliminated by directly using data to derive parameters of theoretical distribution models used in the analysis. The problem of incompleteness, different levels of completeness and inaccuracies are dealt with by in-depth procedures developed in Kijko and Sellevoll (1989 and 1992), developed specifically for this purpose. Figure (4.4) schematically illustrates the condition of completeness and uncertainty of a typical catalogue.

The historic part of the catalogue generally contains only very large events, so it is incorporated by considering these events to be the largest events in a time window. These events can then be assumed to follow an extreme value distribution derived in Kijko and Sellevolll (1989 and 1992):

$$\mathrm{F}_m^{max}(m|m_0, m_{max}, t) = \frac{\exp\{-\lambda_0 t[1 - F_m(m|m_0, m_{max})]\} - \exp(-\lambda_0 t)}{1 - \exp(-\lambda_0 t)}, \qquad (5.1)$$

$$m_{min} \leq m \leq m_{max}$$

where λ_0 is the mean activity rate of earthquakes larger than m_0, and m_0 is the lower limit of the extreme part of the entire catalogue. Note that $\exp(-\lambda_0 t)$ becomes negligibly small for large t, as is the case in most practical situations according to KG1999, so it is ignored.

© University of Pretoria

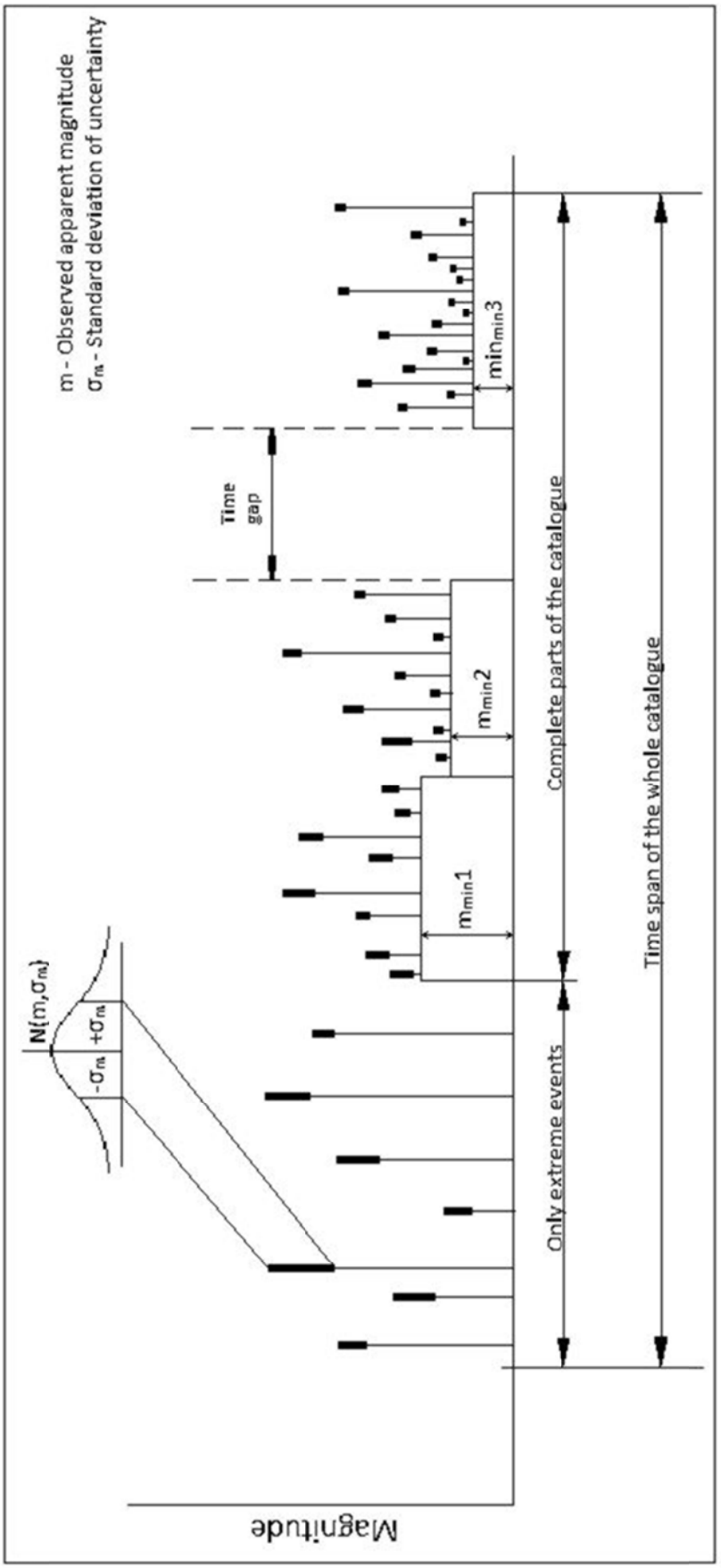

Figure 5.1 Schematic of a typical seismic catalogue. The first part contains only the largest events, with large uncertainty. The second part contains instrumental data of which the level of completeness ($m_{min}1$, $m_{min}2$, $m_{mon}3$) typically becomes lower (better) and standard deviation decreases with time because of the increase in sophistication of the instruments, networks and methods. Time gaps where no data was recorded may occur (such as during the world wars). (After Kijko and Sellevoll, 1992)

© University of Pretoria

For an assumed normal distribution for magnitude uncertainty with standard deviation σ_m, the convolution the assumed distribution of magnitude and the normal distribution yields the following functions for the pdf and cdf of magnitude, respectively (Gibowicz and Kijko, 1994)

$$f_m(m|m_{min}, m_{max}, \sigma_m) = f_m(m|m_{min}, m_{max}) \left\{ \frac{e^{\chi^2}}{2} \left[\mathrm{erf}\left(\frac{m_{max} - m}{\sqrt{2}\sigma_m} + 1 \right) + \mathrm{erf}\left(\frac{m - m_{min}}{\sqrt{2}\sigma_m} - \chi \right) \right] \right\} \tag{5.2}$$

$$F_m(m|m_{min}, m_{max}, \sigma_m)$$

$$= F_m(m|m_{min}, m_{max}) \left\{ \mathrm{A}(m_{min}) \left[\mathrm{erf}\left(\frac{m_{max} - m}{\sqrt{2}\sigma_m} \right) + 1 \right] + \left[\mathrm{A}(m_{max}) \left(\frac{m_{max} - m}{\sqrt{2}\sigma_m} \right) - 1 \right] \right. \tag{5.3}$$

$$\left. - 2 \frac{f_m(m|m_{min}, m_{max}, \sigma_m)}{f_m(m|m_{min}, m_{max})} A(m) \right\} / 2[A_1 - A(m)]$$

where $A(m) = \exp(-\beta m)$, $\mathrm{erf}(.)$ is the cdf of the normal distribution (or the error function), and $\chi = \beta\sigma_m/\sqrt{2}$. The cdf has to be renormalized to include only values above m_{min} (KG1999):

$$\tilde{F}_m(m|m_{max}, \sigma_m) = \frac{F_m(m|m_{min}, m_{max}, \sigma_m) - F_m(m_{min}|m_{min}, m_{max}, \sigma_m)}{1 - F_m(m_{min}|m_{min}, m_{max}, \sigma_m)} \tag{5.4}$$

According to Tinti and Mulgaria (1985a and 1985b, in Kijko and Graham, 1999) the mean activity rate has to be replaced by the apparent mean activity rate $\tilde{\lambda}(m) = \lambda(m)\exp(\chi^2)$. $\tilde{f}_m(m|m_{max}, \sigma_m)$ is obtained by differentiating equation (4.3). To account for magnitude uncertainty in the extreme value distribution (5.1) the $\tilde{F}_m(m|m_{max}, \sigma_m)$ is thus substituted instead of $F_m(m|m_0, m_{max})$. The resulting pdf is given by (KG1999):

$$\tilde{f}_m^{max}(m|m_0, m_{max}, t)$$

$$= \frac{\tilde{\lambda}_0 t \tilde{f}_m(m|m_0, m_{max}, t, \sigma_m) \exp\{-\tilde{\lambda}_0 t[1 - \tilde{F}_m(m|m_0, m_{max}, t, \sigma_m)]\}}{1 - \exp(-\lambda_0 t)}, \tag{5.5}$$

$$m_{min} \leq m \leq m_{max}$$

The unknown pair of parameters (β, λ_0) is estimated by maximising the likelihood function

$$L_0(\beta, \lambda_0) = K_0 \cdot \prod_{j=1}^{n_0} \tilde{f}_m^{max}(m_{0j}|m_0, m_{max}, t_{0j}, \sigma_{m0j}) \tag{5.6}$$

© University of Pretoria

where m_{0j} is the apparent magnitude of the strongest earthquake during the time interval t_{0j}, σ_{m0j} is the standard deviation on m_{0j}, and n_0 is the number of entries in the extreme part of the catalogue. K_0 is a normalization constant. Thus the whole of the extreme part of the catalogue may be utilized to determine the parameters (β, λ_0) on the extreme part of the catalogue where data is sparse because of the low activity rate.

Now for the i[th] complete subcatalogue define $l_{i1}(\beta)$ and $l_{i2}(\lambda_i, \beta)$ as follows:

$$l_{i1}(\beta) = K_\beta \cdot \prod_{j=1}^{n_i} \tilde{f}_m(m_{ij}|m_{min}^i, m_{max}, \sigma_{mij}) \tag{5.7}$$

where m_{ij} is the j[th] entry in the i[th] subcatalogue, m_{min}^i is the level of completeness of the i[th] subcatalogue, and σ_{mij} is the standard deviation on m_{ij}. K_β is a constant.

$$l_{i2}(\lambda_i, \beta) = K_\lambda \cdot \left(\tilde{\lambda}_i t_i\right)^{n_i} \exp(-\tilde{\lambda}_i t_i) \tag{5.8}$$

where

$$\tilde{\lambda}_i = \lambda_i \exp(\chi^2) \tag{5.9}$$

and

$$\lambda_i = \lambda(1 - F(m_{min}^i|m_{min}, m_{max})) \tag{5.10}$$

and K_λ is again some constant. For the i[th] complete subcatalogue the joint likelihood of the pair (β, λ) can be expressed a:

© University of Pretoria

$$L_i(\beta,\lambda) = l_{i1}(\beta)\, l_{i2}(\lambda_i,\beta) \tag{5.11}$$

Equations (5.6) and (5.11) define the joint likelihood function of the whole catalogue:

$$L(\beta,\lambda) = \prod_{i=0}^{n_s} L_i(\beta,\lambda) \tag{5.12}$$

where n_s is the total number of subcatalogues in the combined complete parts of the catalogue. An estimate of (β,λ) is obtained by maximizing the likelihood function (5.12). Note that, because of the incorporation of historical catalogues, neither the estimator

$$\beta = \frac{1}{\bar{m} - m_{min}} \tag{5.13}$$

given by Aki (1965), nor the estimator

$$\frac{1}{\beta} = \bar{m} - m_{min} + \frac{(m_{max} - m_{min})\exp[-\beta(m - m_{min})]}{1 - \exp[-\beta(m_{max} - m_{min})]} \tag{5.14}$$

given by Page (1968) can be used in this case.

After arriving at the MLE for (β,λ), the authors justify the use of the doubly truncated exponential distribution for log-transformed PGA data (discussed further in sections 4.4 and 5.2). Also assuming the occurrence of ground motion at a site as a Poisson process, an analogous approach is followed for direct application to log-transformed PGA data. They do not, however, explicitly include uncertainty in this case. The resulting equations for the MLE of the parameters $\left(\gamma,\lambda_{\ln(pga)}\right)$, analogous to (β,λ) for magnitude distribution, they give as

$$\begin{cases} \dfrac{1}{\lambda_{\ln(pga)}} = \dfrac{\langle t\rangle A_1 - \langle tA\rangle}{A_2 - A_1} - \left\langle \dfrac{t\exp(-\lambda_{\ln(pga)}t)}{1 - \exp(-\lambda_{\ln(pga)}t)} \right\rangle \\ \dfrac{1}{\gamma} = \langle x\rangle - \dfrac{B_2 - B_1}{A_2 - A_1} + \lambda_{\ln(pga)}\left[\dfrac{(\langle t\rangle A_2 - \langle tA\rangle)(B_2 - B_1)}{(A_2 - A_1)^2} - \dfrac{\langle t\rangle B_2 - \langle tB\rangle}{A_2 - A_1}\right] \end{cases} \tag{5.15}$$

© University of Pretoria

where $A_1 = \exp(-\gamma x_{min}),\ A_2 = \exp(-\gamma x_{max}), B_1 = x_{min}A_1,\ B_2 = x_{max}A_2,$ $\langle \frac{t\exp(-\lambda_{\ln(pga)}t)}{1-\exp(-\lambda_{\ln(pga)}t)} \rangle = \frac{1}{n}\sum_{i=1}^{n} \frac{t\exp(-\lambda_{\ln(pga)}t_i)}{1-\exp(-\lambda_{\ln(pga)}t_i)}, \langle t \rangle = \sum_{i=1}^{n}\left(\frac{t_i}{n}\right), \langle x \rangle = \sum_{i=1}^{n}\left(\frac{x_i}{n}\right), \langle tA \rangle = \sum_{i=1}^{n}\left(\frac{t_i A(x_i)}{n}\right),$ $\langle tB \rangle = \sum_{i=1}^{n}\left(\frac{t_i B(x_i)}{n}\right), A(x) = \exp(-\gamma x),$ and $B(x) = xA(x)$.

In an example application KG1999 estimates the maximum possible PGA by substituting m_{max} and the closest possible distance into the GMPE and calculating a large upper confidence limit in the uncertainty distribution of the GMPE. They do suggest the use of the maximum estimators developed in KG1998 directly on log-transformed PGA data (not necessarily single station data, but also data derived through GMPEs) – one of the aims of this research project is to develop this suggestion more fully as an extension to the Parametric-Historic procedure.

5.2. Theoretical Distribution of peak ground acceleration: The Pareto distribution

Kijko and Graham (1999) very briefly derived a site specific distribution of peak ground acceleration and noted that it is in fact source-free. The distribution they give is in terms of the logarithm of PGA is

$$\Pr[\ln(pga) \le x] = \begin{cases} 0, & x < \ln(PGA_{min}) \\ \dfrac{\exp[-\gamma \ln(PGA_{min})] - \exp[-\gamma x]}{\exp[-\gamma \ln(PGA_{min})] - \exp[-\gamma \ln(PGA_{max})]}, & \ln(PGA_{min}) \le x \le \ln(PGA_{max}) \\ 1, & x > \ln(PGA_{max}) \end{cases} \quad (5.16)$$

which, when transformed to give the cdf of PGA, results in a truncated Pareto distribution:

$$\Pr[PGA < x] = \begin{cases} 0, & x < \ln(PGA_{min}) \\ \dfrac{1 - \left(\frac{x}{PGA_{min}}\right)^{-\gamma}}{1 - \left(\frac{PGA_{max}}{PGA_{min}}\right)^{-\gamma}}, & \ln(PGA_{min}) \le x \le \ln(PGA_{max}) \\ 1, & x > \ln(PGA_{max}) \end{cases} \quad (5.17)$$

© University of Pretoria

where $\gamma = \frac{\beta}{c_2}$, and c_2 corresponds to the constant in equation (3.11).

The Pareto distribution is a distribution in which the frequency of observation of a value is related to some negative power of that value (Newman, 2005):

$$\mathrm{f_x} = Cx^{-b} \tag{5.18}$$

Here, and in the rest of the section, it will be taken for granted that $x \geq x_{min}$, for some minimum value x_{min} that x may take on. Many natural disasters also follow a Pareto distribution (in some cases considered to be a product of Self Organized Criticality; Bak, 1996), thus it is often encountered in hazard analysis and the reinsurance industry. Newman (2005) gives several examples of phenomena that follow (at least in part) Pareto distributions: populations of cities, moon crater diameters, intensity of solar flares, and intensity of wars - just to name a few. In extreme value theory, many distributions are considered to approximately follow a Pareto distribution for large values (i.e. in the upper tail part) (Caserta and De Vries, 2003). Such distributions are often referred to as heavy tailed distributions, and the Pareto distribution is the classic example of a heavy tailed distribution (Resnick, 2007). The exponent b in equation (5.18) is the value that determines how 'heavy' the tail of the distribution is, and is technically referred to as the tail index.

Heavy tailed distributions derive their name from the fact that their tails largely affect the location of the mean of the distribution (the mean may be infinite if the tail is too 'heavy'). The technical definition implies that a heavy tailed distribution has the property that very large values are almost equally likely (Sigman, 1999). This property causes a very counter-intuitive stochastic behaviour and is one of the reasons why risk may be misjudged on the basis of one's intuitive judgement (Naylor *et al.*, 2009). For this reason the Pareto distribution is much better characterised by a *recurrence period* of events larger than a certain size and by the *mean excess* over a certain threshold value, rather than by a single central trend. The mean excess of a distribution *f(x)* over a given threshold is merely the mean (or expected value) of the part of the distribution that is larger than the threshold value. The precise definition of the mean excess function (also known as the mean residual lifetime, depending on the context) is (Smith, 2003; Nieboer, 2011):

© University of Pretoria

$$M(t) = E[f(x|x > t)] = \int_{x_{min}}^{\infty} xf(x|x > t)d\,x \tag{5.19}$$

where *M(t)* denotes the mean excess over threshold *t*. The integral's upper limit is at infinity, but in many cases distribution is truncated at some value.

Another property which characterises the Pareto distribution is what is known as self-similarity (Nieboer, 2011) or the property of being scale free (Newman, 2005). This means that the scale (or units) by which one measures does not affect the shape of the distribution (it is only rescaled by a multiplicative constant) (Newman, 2005). Mathematically it is stated as

$$f(cx) = h(c)f(x) \tag{5.20}$$

Related to the property in equation (5.20) is a convenient formulation of the *conditional* Pareto distribution $p(x|x > t)$:

$$p(x|x > t) = k(t)p\left(\frac{x}{t}\right) \tag{5.21}$$

where $k(t) = \frac{s(t)}{h(1/t)}$, and $s(t)$ is a scaling factor rescaling the distribution to satisfy the criterion $\int_t^{\infty} p(x|x > t)dx = 1$. To be specific, a Pareto distribution with any minimum cutoff value x_{min} is given by

$$p(x) = \frac{(b-1)}{x_{min}}\left(\frac{x}{x_{min}}\right)^{-b} \tag{5.22}$$

and the conditional Pareto distribution by

© University of Pretoria

$$p(x|x>t) = \frac{{x_{min}}^{-b-1}}{t} \times \frac{(\alpha-1)}{x_{min}} \left(\frac{\left(\frac{x}{t}\right)}{x_{min}} \right)^{-b} = \frac{(b-1)}{t} \left(\frac{x}{t}\right)^{-b} \tag{5.23}$$

Substituting (5.23) in (5.19) and evaluating the improper integral one obtains

$$M(t) = \frac{1-b}{2-b} t \tag{5.24}$$

This implies that *M(t) is linear in t*. Moreover, coming back to the property of self-similarity (equation (5.20)), the slope depends only on b, no matter what scale or units one measures in. Two very useful heuristics (rules of thumb) follow from this (Nieboer, 2011): first, a plot of the empirical mean excess function $\widehat{M}(t)$ of Pareto distributed data against the threshold value, t, theoretically follows a linear trend; second, if one *aggregates* data by summing together groups of data where each group contains the same amount of data points, the resulting empirical mean excess function looks similar to that of the original data – specifically, the linear trend followed by the aggregated data has the same slope. These properties for $\widehat{M}(t)$ only strictly hold when the mean is finite, that is when $b < 1$.

Now, because equations (5.22) and (5.23) are essentially the same, we see from (5.24) that a Pareto distributed population has mean

$$E[p(x)] = \frac{1-b}{2-b} x_{min} \tag{5.25}$$

which is the mean (or mean excess) over the chosen threshold x_{min} .The variance of the pareto distribution is (Johnson and Kotz, 1994)

© University of Pretoria

$$V[p(x)] = \frac{(b-1)^2}{(b-2)^2(b-3)} {x_{min}}^2 \tag{5.26}$$

and the mean becomes infinite when $b < 1$, and the variance when $b < 2$. The cdf of the Pareto distribution is

$$\mathcal{P}(x) = 1 - \left(\frac{x}{x_{min}}\right)^{-(b-1)} \tag{5.27}$$

Note that the truncated Pareto distribution (compare with equations (3.11) and (5.22)) differs in its functional form only by a scale parameter, which is $\frac{1}{\mathcal{P}(x_{max})}$. Applying this to ground motion, the cdf of PGA is given by

$$\Pr[PGA < x] = \left[\frac{\mathcal{P}(x)}{\mathcal{P}(x_{max})}\right], \qquad x \leq x_{max} \tag{5.28}$$

© University of Pretoria

6. Theoretical validation of use of the Pareto distribution to model the distribution of peak ground acceleration and transformation to an exponential distribution

In this chapter the Pareto distribution is validated as a model for the distribution of PGA data (and other peak ground motion parameters following an assumed attenuation law). As a synopsis, this validation draws together the Gutenberg-Richter law and the ground motion prediction equation in equation (3.12), resulting in a Pareto distribution. After the validation, the logarithmic transformation to an exponential distribution is introduced. The logarithmic transformation results in an exponential distribution, identical in form to what the Gutenberg-Richter law gives rise to. This transformation therefore has the advantage that statistical methods familiar in seismology, as applied to magnitude data under the assumption of the G-R recurrence law, may be applied to log-transformed PGA data. Finally, parameters of the exponential distribution and estimation of these parameters is discussed.

6.1. Combination of the Gutenberg-Richter law and the ground motion prediction equation

It will be taken for granted here that defined cdf's are 0 when $m \leq m_{min}$ and 1 when $m \geq m_{max}$. As mentioned in Section 3.2, the Gutenberg-Richter law gives rise to the exponential distribution of magnitude with cdf:

$$F(m) = 1 - e^{-\beta(m-m_{min})} \tag{6.1}$$

If a truncation point, that is m_{max}, is imposed, then equation (6.1) is normalised as

$$F(m) = \frac{1 - e^{-\beta(m-m_{min})}}{1 - e^{-\beta(m_{max}-m_{min})}} \tag{6.2}$$

© University of Pretoria

Combining ground motion prediction equation (3.12) and the truncated G-R relation in equation (3.3), but considering the error term ε in (3.12) to be zero for the present, one obtains a cumulative distribution for the expected PGA:

$$\begin{aligned}[\ln(PGA) < x|r] &= P[c_1 + c_2 M + \phi(r) < x|r] \\ &= P\left[M < \frac{x - c_1 - \phi(r)}{c_2}|r\right] \\ &= \frac{1 - \exp\left[-\beta\left(\frac{x-c_1-\phi(r)}{c_2} - m_{min}\right)\right]}{1 - \exp[-\beta(m_{max} - m_{max})]}\end{aligned} \tag{6.3}$$

If ε is a symmetrical distribution (such as the normal distribution) we are considering in effect the mean (and median) by assuming $\varepsilon = 0$. Integration over all possible source distances gives

$$\begin{aligned}&P[\ln(PGA) < x] \\ &= \frac{\int_{max\{r_{min},\phi^{-1}(x-c_1-c_2 m_{max})\}}^{min\{r_{max},\phi^{-1}(x-c_1-c_2 m_{min})\}} f_R(r)dr - \exp\left[-\beta\left(\frac{x-c_1-\ln(a(x))/\left(\frac{\beta}{c_2}\right)}{c_2} - m_{min}\right)\right]}{1 - \exp[-\beta(m_{max} - m_{min})]}\end{aligned} \tag{6.4}$$

where

$$a(x) = \int_{max\{r_{min},\phi^{-1}(x-c_1-c_2 m_{max})\}}^{min\{r_{max},\phi^{-1}(x-c_1-c_2 m_{min})\}} e^{\frac{\beta\phi(r)}{c_2}} f_R(r)dr \tag{6.5}$$

$f_R(r)$ being the pdf of the source being at distance r.

© University of Pretoria

Given the expressions

$$m_{min} = \frac{\ln(PGA_{\min}) - c_1 - \phi(r_{max})}{c_2} \tag{6.6}$$

$$m_{max} = \frac{\ln(PGA_{\max}) - c_1 - \phi(r_{min})}{c_2} \tag{6.7}$$

One obtains

$$P[\ln(PGA) < x] = \frac{1 - \exp\{-\gamma[x - (\ln(PGA_{min}) - (\phi(r_{max}) - \ln(a)/\gamma))]\}}{1 - \exp\left[-\gamma\left(\ln(PGA_{\max}) - (\phi(r_{min}) - \ln(a)/\gamma) - (\ln(PGA_{\min}) - (\phi(r_{max}) - \ln(a)/\gamma))\right)\right]} \tag{6.8}$$

where $\gamma = \frac{\beta}{c_2}$.

If it is assumed that

$$\phi(r_{min}) - \phi(r_{max}) \ll c_2(m_{max} - m_{\min}) \tag{6.9}$$

it is justified to approximate

$$a(x) = \int_{\max\{r_{min}, \phi^{-1}(x - c_1 - c_2 m_{max})\}}^{\min\{r_{max}, \phi^{-1}(x - c_1 - c_2 m_{min})\}} e^{\gamma\phi(r)} f_R(r) dr \approx \int_{r_{min}}^{r_{max}} e^{\gamma\phi(r)} f_R(r) dr \tag{6.10}$$

© University of Pretoria

$$= e^{\gamma\phi(r_{max})} - \int_{r_{min}}^{r_{max}} \gamma\phi'(r)e^{\gamma\phi(r)}F_R(r)dr$$

Now

$$\left|\int_{r_{min}}^{r_{max}} \gamma\phi'(r)e^{\gamma\phi(r)}F_R(r)dr\right| < \left|\int_{r_{min}}^{r_{max}} \gamma\phi'(r)e^{\gamma\phi(r)}dr\right| = |e^{\gamma\phi(r_{min})} - e^{\gamma\phi(r_{max})}| \tag{6.11}$$

$$\leq \gamma|\phi(r_{min}) - \phi(r_{max})|$$

because ϕ is negative on $[r_{max}, r_{min}]$, and γ is a Lipschitz constant for the function $e^{\gamma\phi}$ for $\phi < 0$. The final term in the inequalities (6.11) is negligible compared to the range of values PGA takes on (see inequality (6.9)), so it is ignored. Furthermore, because of (6.9), the approximation

$$\int_{max\{r_{min},\phi^{-1}(x-c_1-c_2m_{max})\}}^{min\{r_{max},\phi^{-1}(x-c_1-c_2m_{min})\}} f_R(r)dr \approx \int_{r_{min}}^{r_{max}} f_R(r)dr = 1 \tag{6.12}$$

may be made. If $\phi(r_{min}) - \phi(r_{max})$ is approximated to be zero in the denominator as well (because of its negligible effect compared to magnitude), the result is the exponential distribution:

$$P[\ln(PGA) < x] = \frac{1 - \exp\{-\gamma[x - \ln(PGA_{min})]\}}{1 - \exp[-\gamma(\ln(PGA_{\max}) - \ln(PGA_{\min}))]} \tag{6.13}$$

© University of Pretoria

This result was originally derived in a very briefly by Kijko and Graham (1999), and will form the basis of the proposed extension to the Parametric-Historic procedure in Chapter 5. It also follows that the PGA is distributed according to

$$P[PGA < y] = \frac{(PGA_{min})^{-\gamma} - (y)^{-\gamma}}{(PGA_{min})^{-\gamma} - (PGA_{max})^{-\gamma}} \tag{6.14}$$

This is the truncated Pareto distribution as given by Johnson *et al.* (1994).

6.2. Transformation to an exponential distribution

As was seen in equation (6.13), the logarithm of PGA is readily described as an exponential distribution under the stated assumptions. Practical methods developed by authors such as Aki (1965), Page (1968) and Kijko and Graham (1998 and 1999) are applicable to the exponential distribution, so the application of these methods to ground motion data makes the logarithmic transformation attractive in the sense of its familiarity. Pisarenko and Rodkin (2013) also state "*The treatment of heavy-tailed data is often facilitated by using logarithms of original values. Switching to logarithms (which can be done only when the original numerical values are positive) ensures almost always that all the statistical moments exist, and hence the Law of Large Numbers and the Central Limit Theorem are applicable to the sums of logarithms."* The common assumption that the residual term ε in equation (3.12) for the logarithm of PGA follows a normal distribution also favours the use of such transformation. Furthermore, the exponential distribution is in a sense "better behaved" in that it does not possess the heavy tailed properties that give rise to rather counter-intuitive behaviour of observations from the Pareto distribution and the problem of nonexistence of the mean and variance in some cases is eliminated. Johnson *et al.* (1996) mention that such a transformation is common for handling power-law distributed data, although they caution that an understanding of the process giving rise to a Pareto distribution cannot be explained by characterizations in terms of the exponential distribution. It is best, they say, to analyse the raw data to determine the fitness of the Pareto model *per se*. It may be noted, however, that goodness of fit statistics of the Kolmagorov-Smirnov type are independent of data transformations, a fact that will be used later on. The transformation also does not affect maximum likelihood estimates of parameters.

© University of Pretoria

6.3. Parameters and their estimation

Equation 5.16 has 3 parameters that have to be estimated – that is γ, $\ln(PGA_{min})$ and $\ln(PGA_{max})$ – if we mean to interpret it as a doubly truncated exponential distribution.

PGA_c is the threshold of completeness, i.e. the lower limit above which *all* PGA values were recorded during the recording time. The data below the threshold is incomplete due to insensitivity of instrumentation, in the case of catalogues consisting of data from different stations and the cumulative effects of different cut-offs where values are considered to be too low to be of interest (Woessner and Wiemer, 2005). It will be assumed that $PGA_c = PGA_{min}$, where PGA_{min} is the value from which equation (6.14) holds. Thus, if PGA_{min} is placed below this level the shape of the distribution is not correctly represented by the data, if above, useful data is lost. Wiemer and Woessner (2005) and Mignan *et al.* (2011) give a comprehensive list of techniques to estimate the level of completeness of magnitude catalogues. Because of equation (6.14), estimates on a log-log histogram or survivor curve of PGA data of the point where "data loss curvature" starts may be used to estimate $\ln(PGA_{min})$ (e.g. Woessner and Wiemer, 2005; Amorèse, 2007). Figures 6.1 and 6.2 provide illustrations (these figures were done graphically, merely to illustrate the concept). Two methods of note that do not depend on the assumption of equation (6.14) are: a method that identifies the threshold where signal-to-noise-ratio variations can be detected diurnally (Rydelek and Sacks, 1989); and the intricate yet robust method of Schrolemer and Woessner (2008) that computes detection probabilities as a function of distance and magnitude. Techniques that are based on analysis of station specific data and probability of detection may be particularly useful for the use of PGA data.

© University of Pretoria

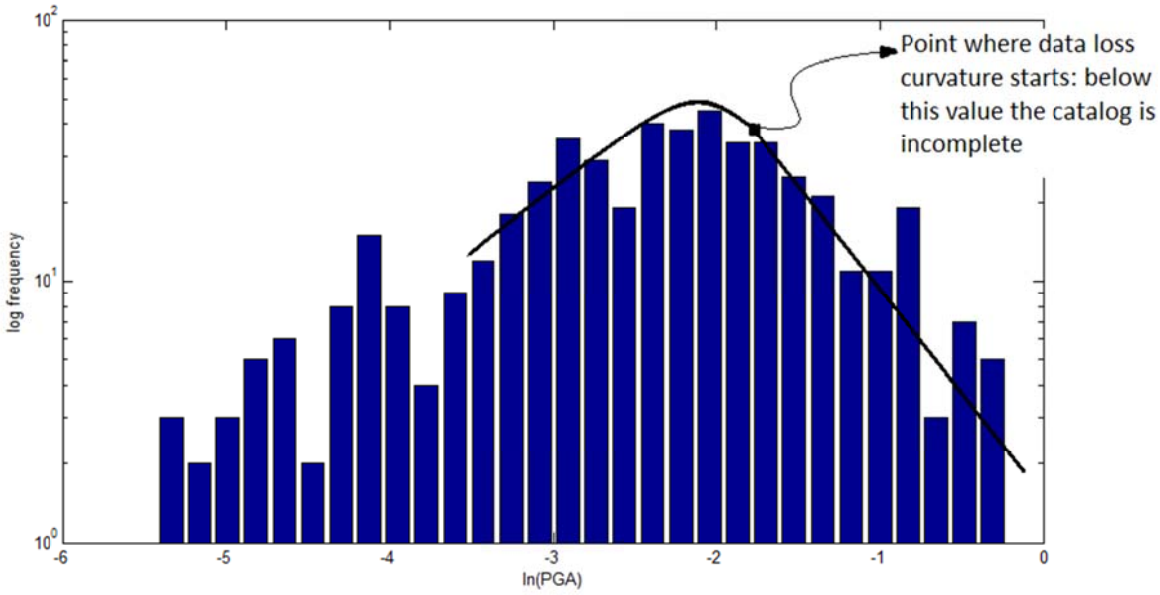

Figure 6.1 Estimation of level of completeness of data with similar site effect from the PEER NGA database from a log-log histogram.

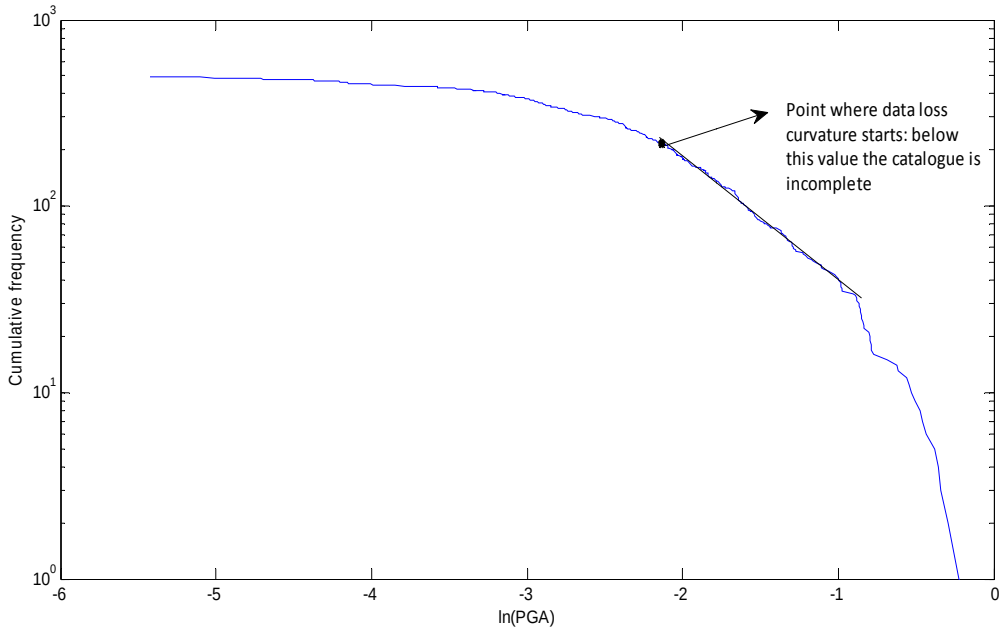

Figure 6.2 Estimation of level of completeness of data with similar site effect from the PEER NGA database (Pacific Earthquake Engineering Research Center, 2005) from a survivor curve (deviations from linearity for high values are not due to incompleteness, but possibly rather due to boundedness, or statistical noise in the tail – the tail is always very scantly populated – see Naylor *et al.*, 2009; Newman, 2005).

© University of Pretoria

PGA_{max} is understood to be a physical limit as discussed in Chapter 2. $\ln(PGA_{max})$ may be estimated with methods discussed in Sections 4.2.1 and 4.2.2, again from the log-transformed data in the cases where the exponential distribution is applicable.

The parameter γ is equivalent to the β parameter of the G-R distribution (equation 5.2). Following KG, the *maximum likelihood method*, of which the result is referred to as the *maximum likelihood estimate* (MLE) of the parameter, will be used to estimate this parameter. It is determined by maximizing the value of the *likelihood function*, which is proportional to the joint probability of observing all the individual observed values from the assumed distribution. The likelihood function, expressed as a function of the parameter to be estimated is

$$\mathcal{L}(\gamma|\boldsymbol{x}) = f(\boldsymbol{x}|\gamma) = \prod_{i=1}^{n} f(x_i|\gamma) \tag{6.15}$$

where $\boldsymbol{x}$ (boldface) denotes the vector containing all the observed values and x_i the ith observed value (note that in our case $\boldsymbol{x}$ and x_i would refer to the log-transformed data). The MLE of β is the value that maximises likelihood function $\mathcal{L}(.)$. Page (1968) has shown that the MLE for γ uniquely satisfies

$$\frac{1}{\gamma} = \bar{x} - x_{min} + \frac{(x_{max} - x_{min})\exp[-\gamma(x_{max} - x_{min})]}{1 - \exp[-\gamma(x_{max} - x_{min})]} \tag{6.16}$$

where $\bar{x}$ is the sample mean, and, in the case under consideration, $x_{min} = \ln(PGA_{min})$, and $x_{max} = \ln(PGA_{max})$. Although the solution cannot be obtained explicitly, it can be estimated numerically by fixed point iteration.

Some of the estimators of x_{max} require that the value of γ should be specified, and in equation (6.16) x_{max} has to be specified. As recommended by Kijko and Graham (1998) a fixed point iteration is done on the pair $< x_{max}, \gamma >$ to overcome this difficulty, and the estimator

© University of Pretoria

$$\gamma = \bar{x}^{-1} \tag{6.17}$$

for the untruncated exponential distribution or the estimator given by Aki (1965) for a distribution only truncated from below

$$\gamma = \frac{1}{\bar{x} - x_{min}} \tag{6.18}$$

is used to obtain an initial approximation. The initial approximation is used to estimate $\ln(PGA_{max})$ with the K-S or K-S-B estimators discussed in Section 4.2.1. The estimate of γ is refined by equation 4.22, and $\ln(PGA_{max})$ again estimated with the new value of γ. The iteration is continued until a satisfactory close approximation is obtained.

© University of Pretoria

7. Direct estimation of PGA_{max} by application of the methods previously used to estimate m_{max}

In this chapter an extension to the Parametric-Historic procedure by direct application of maximum estimators discussed in Section 4.2.1 to PGA data is introduced. This extension follows naturally, so the bulk of this chapter is devoted to an example application to single-station PGA data. This application is both validative and illustrative.

7.1. Estimating PGA_{max}

Pareto distributed data is often analysed in log-transformed form (Johnson *et al.*, 1994); the log-transformed data is exponentially distributed and standard methods for exponential distributions apply. This is also why, as discussed in the previously, the Pareto distribution of earthquake energy translates to the exponential distribution of earthquake magnitude (see Section 3.1). In the same way Pareto distributed PGA data implies that log-PGA is exponentially distributed. Advantage is taken of some decades' work on the statistical and probabilistic properties of the Gutenberg-Richter relation, which translates to that of an exponential distribution, which was shown to also to be applicable to log-PGA data. Specifically, methods discussed in Section 4.2 allows one to estimate PGA_{max} in terms of $\ln(PGA_{max})$. Pisarenko and Lyubushin (1997) published a paper that used ideas along this line, but, according the author's knowledge, no one has further worked on the idea till present. It is essential to draw attention to the openness of the idea: methods (of which there are many) used to estimate the upper limit of magnitude may also be used to estimate the upper limit of PGA from log-transformed data. Only one assumption is made: the ground motion data under consideration is Pareto distributed by and large, so as safeguard against using it inappropriately one has to test if the data conforms to the Pareto distribution. If data is not distributed according to a Pareto distribution, one of the non-parametric methods in Section 4.2.1 may be used, preferably on log-transformed data.

The methodology proposed above is conceptually straightforward. The rest of this section is devoted to the specific case where a Pareto distribution is assumed, but note that the condition (6.9) is actually not met in the case study in subsequent sections, for which the Pareto distribution was found to hold in section 6.1.

© University of Pretoria

7.2. Remarks on when it is justified to assume a Pareto distribution and variations thereof as a distribution model of peak ground acceleration

The K-S procedure is expected to be effective in cases where ground motion clearly follows a Pareto distribution. As noted previously, the existence of such cases is both well established and explainable. One also has to keep in mind that it is often a good idea to 'let the data speak for itself', as many scientists have often admonished, even though something might not be readily explainable. But in cases where data clearly deviates from a Pareto distribution, yet shows close resemblance, the K-S-B estimator may be a good choice as an estimator. It was developed for a stochastically varying b-value in the Gutenberg-Richter law (equation (3.1)), which (b-value) many scientists consider to have a direct physical connotation; but note that it also determines the slope of the graph at a specific point (in fact, it determines all the derivatives of the relation). One of the main results used in extreme value theory is that extreme values follow the so-called Generalized Pareto Distribution (see also Section 4.2.2) (Embrechts *et al*., 1997; De Haan and Ferreira, 2006; Pisarenko *et al*., 2008b).

Deviation from the Pareto distribution this may well be approximated by a variation in the b-value with a distribution with a Gamma distribution. In a rigorous sense, a function may be expressed a mixture (i.e. a weighted sum or integral) of exponentials if it is *completely monotonic* (Miller and Samko, 2001), that is, if it possesses derivatives of all orders and satisfies

$$(-1)^n f^{(n)}(x) \geq 0 \tag{7.1}$$

The only considerable limitation of completely monotonic functions is that they cannot have inflection points, which is evident from the fact mentioned by Miller and Samko (2001) that, firstly, a completely monotonic function is either identically zero or never zero, and secondly that $f^{(2n)}(x)$ and $-f^{(2n-1)}(x)$ are also completely monotonic.

If ground motion data clearly does not even approximately follow a Pareto distribution such as when it has several modes are observed (Kijko *et al*., 2001) (which might well be the case, especially in mine induced seismicity; Gibowicz and Kijko, 1994; Kijko *et al.*, 2001), then the non-parametric estimation with Gaussian kernels is recommended. If data is sparse so that no good

© University of Pretoria

estimates of the distribution shape and/or parameters can be made, the R-W procedure may be considered because of the fact that, although it has large scatter, it is mean unbiased (e.g. Kijko and Singh, 2011).

7.3. Data

As an illustration of the above formalism, it is applied to catalogues of PGA values records in the region of the Żelazny Most slimes dam in the Legnica-Głogów Copper District, Poland. The records of six stations, named Grodowiec, Guzice, Komorniki, Trzebcz, ZM2WP, and ZM8WP here for convenience, were used. The longest of these catalogues span a 10 year period (2001-2011) and are all located in a relatively small area – no two stations are more than 6 km apart, though the stations were not installed simultaneously. Figure 7.1 shows the locations of the stations relative to the slimes dam, along with the seismogenic mining region.

© University of Pretoria

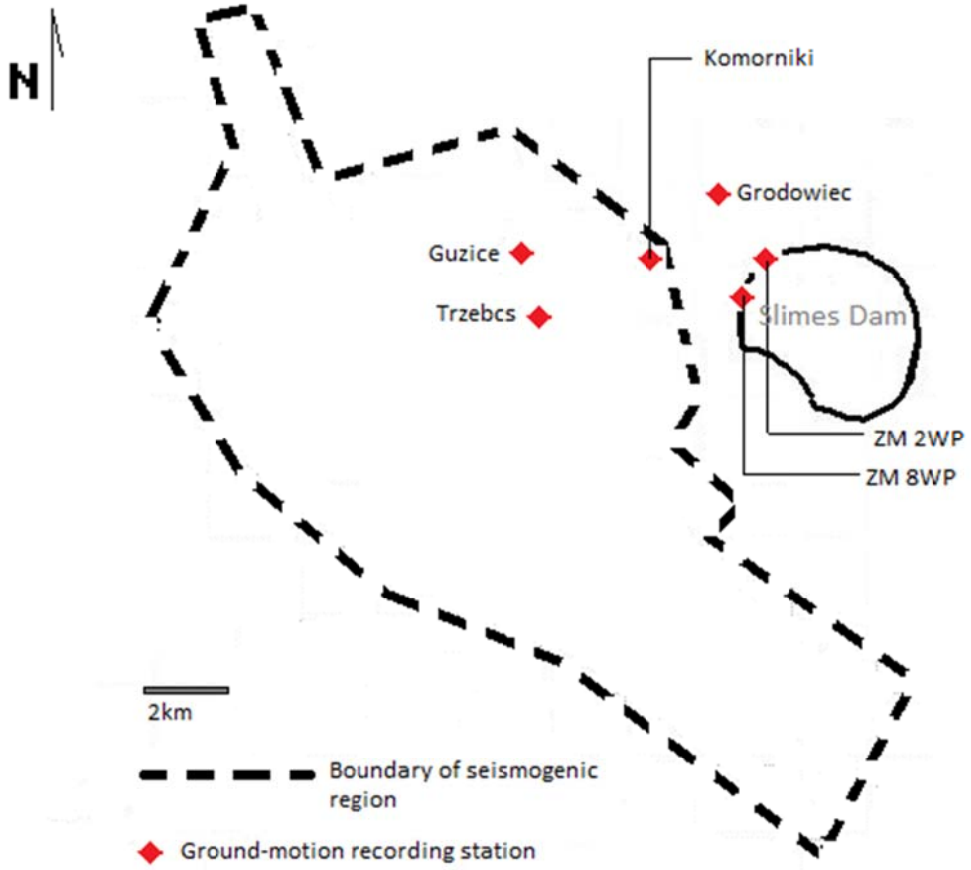

Figure 7.1 Locations of recording stations relative to slimes dam (derived from Lasocki, 2005; and Orlecka-Sikora *et al.*, 2012).

In this district, copper is extracted from very hard, competent rock which is prone to violent brittle failure manifesting in rock bursts and seismicity, and the safety and stability of the large Żelazny Most slimes dam is a concern (Orlecka-Sikora *et al.*, 2012). This spurred detailed investigations and careful monitoring of the seismicity in the area (e.g. Orlecka-Sikora *et al.*, 2012; Lasocki, 2005). Figure 7.2 shows histograms of the PGA data from each station. The deviation from the power law trend in small values are considered to be due to incompleteness, which is typical also in magnitude catalogues, and arises because of insensitivity of instrumentation and subsequent increasing loss of data below a certain threshold, called the level of completeness (Woesner and Wiemer, 2005). For purposes of this work the level of completeness was simply estimated graphically.

7.4. Estimation of parameters and goodness of fit

© University of Pretoria

Only the *total* PGA was used for the analysis of the data (not vertical, horizontal, etc. components). For each station's data a suitable minimum cutoff vale was determined visually by inspecting histograms of the data (see Section 6.3). The log-transformed data was used to calculate ln(PGA_{max}) with utilization of a software package called HA2 (Hazard Area 2) version 2.05 developed by Kijko (2006), using an option implementing the so-called Kijko-Sellevoll-Bayes procedure (see equation (4.20)). The software uses an iterative procedure whereby estimates of γ and $\ln(PGA_{max})$ of equation 6.13 are subsequently refined in by iteration in the way discussed in Section 5.1. The transformed data was scaled and shifted by a constant to fit in an interval required by the software, so the actual transformation is of the form $a\ln(PGA) + b$, where a and b are constants (the reason for this transformation is that the software is designed to handle earthquake catalogues, and some subroutines assume typical values taken on by magnitude). Note that, in the case of such transformation, if the data is Pareto distributed with tail index γ, the parameter in the resulting exponential distribution is γ/a .

© University of Pretoria

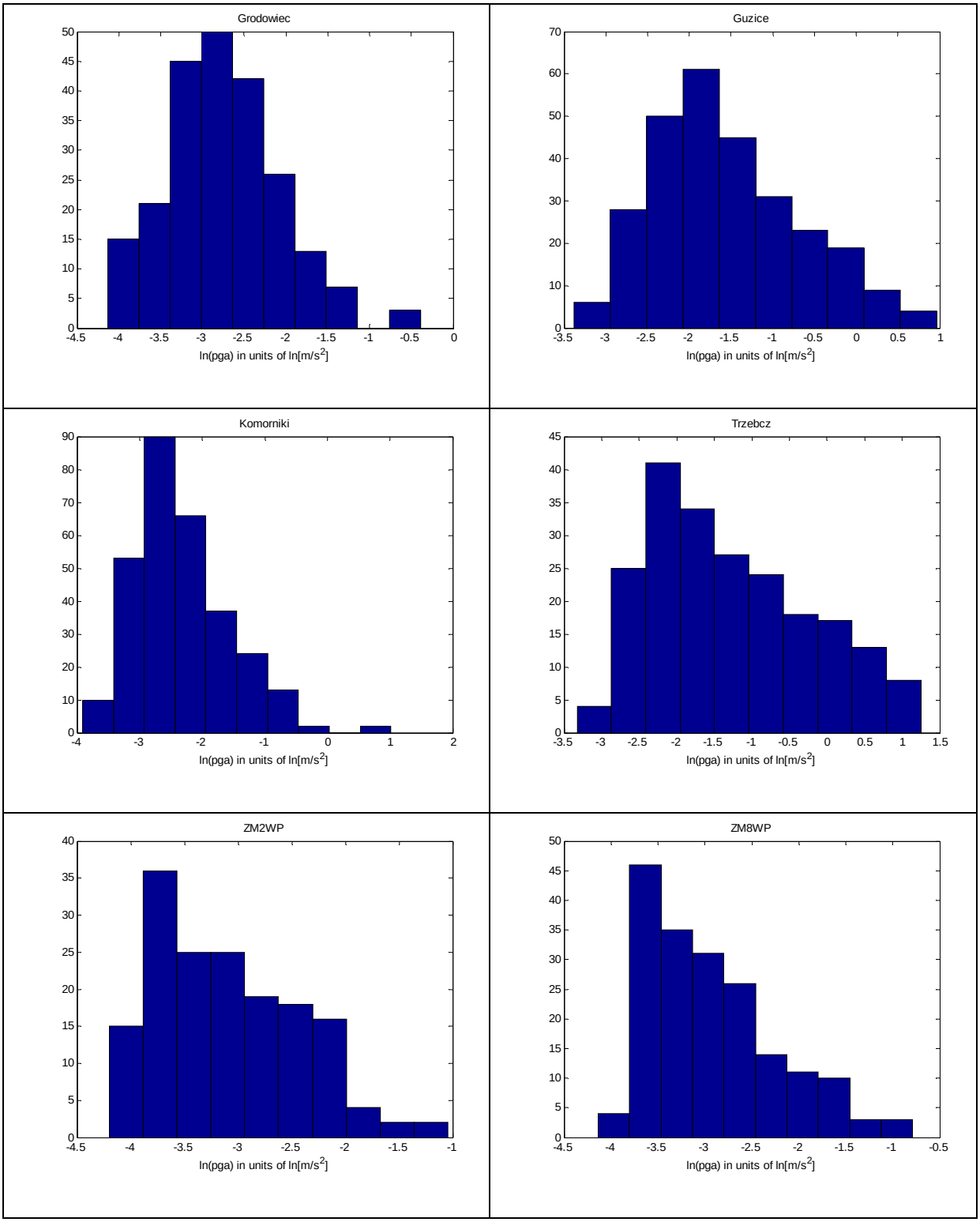

Figure 7.2 Histograms of PGA values from each catalogue

The Cramer-von Mises Goodness of Fit statistic was calculated for the fitted exponential distribution obtained, as recommended by Stephens (1974). Stephens recommends the Cramer-von Mises statistic for the case where the value of the exponent (γ/a in this case) was estimated from

© University of Pretoria

the sample. Table 7.2 gives the resulting figures for the goodness of fit test (adapted confidence levels were obtained from Stephens, 1974). It should be noted that the Cramer-von Mises statistic itself is not affected by the transformation of the data, so the figures indicate the fitness of the Pareto model to the original data. The constants *a* and *b* used in the transformation are shown in Table 7.1.

Table 7.1 Constants for logarithmic transformation a*ln*(PGA)+b

	Grodowiec	Guzice	Komorniki	Trzebcz	ZM2WP	ZM8WP
a	1.5369	1.296	1.0445	1.1267	1.4779	1.4162
b	7.1014	5.2479	5.4386	5.0943	8.0432	7.6152

Table 7.2 presents all the resulting values for the evaluation or application excercise with the transformed data, and Table 7.3 presents the resulting PGA_{max} values and distribution tail index after reversing the transformation. Note that the tail index is the power of the probability density function, not that of the cumulative distribution; so it is obtained it by multiplying the beta value by the transformation constant and adding one. Figures 7.3 (a) through (f) show survivor plots of the complete parts of each station's catalogue along with that of the fitted exponential-gamma model. As an aside, the exponential-gamma model gives rise to distribution of the class of Pareto distribution (Johnson *et al.*, 2004).

© University of Pretoria

Table 7.2 Resulting values for application of exponential-gamma model and direct estimation of PGA_{max} to mining data

	Grodowiec	Guzice	Komorniki	Trzebcz	ZM2WP	ZM8WP
γ/a	0.92	0.69	1.17	0.46	0.69	0.73
$\sigma(\gamma/a)$ †	0.1	0.07	0.09	0	0.09	0.08
$a ln(PGA_{max})+b$	6.79	6.61	6.9	6.59	6.65	6.65
W^2 *	0.2920	0.0898	0.2072	0.0604	0.1468	0.0859
Level of confidence	0.01<a<0.025	a>0.15	0.05<a<0.1	a>0.15	a>0.15	a>0.15

*Cramer-von Mises statistic

Note: $a\ln(PGA_{min})+b=2.5$ in each case

† Note that this is specifically the standard deviation used in the mixed exponential distribution

Table 7.3 Resulting figures after reversing the transformation

	Grodowiec	Guzice	Komorniki	Trzebcz	ZM2WP	ZM8WP
PGA_{max}	0.82	2.9	4.1	3.8	0.39	0.51
$\gamma + 1$	2.4	1.89	2.2	1.52	2.0	2.0

© University of Pretoria

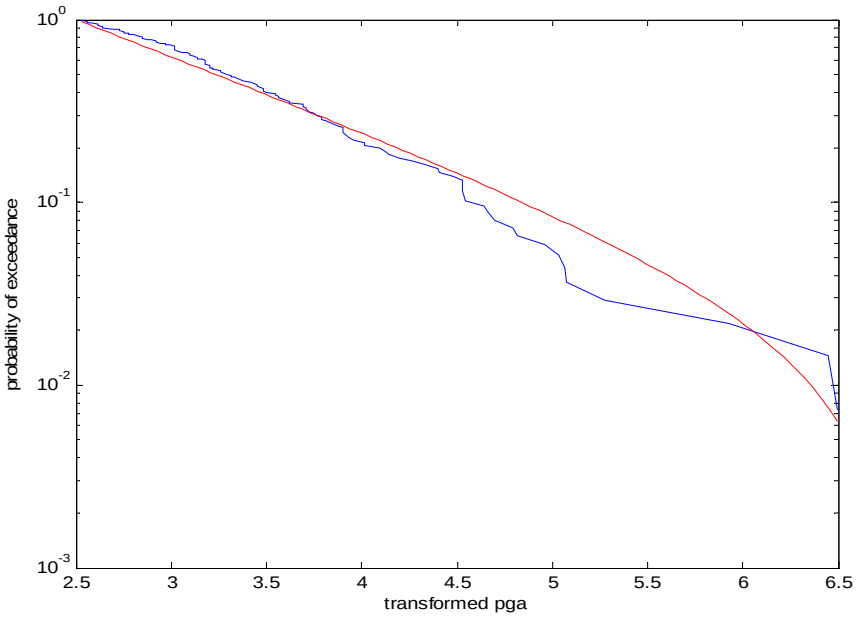

Figure 7.3 (a) Survivor plot for complete part of data from Grodowiec. Red line indicates the survivor function fitted model, the blue line that of the actual data.

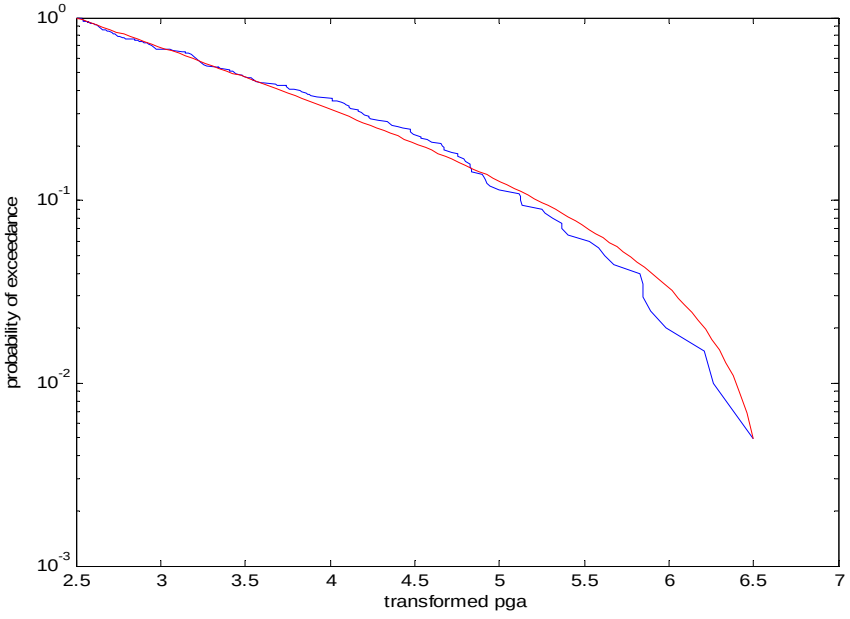

Figure 7.3 (b) Survivor plot for complete part of data from Guzice. Red line indicates the survivor function fitted model, the blue line that of the actual data.

© University of Pretoria

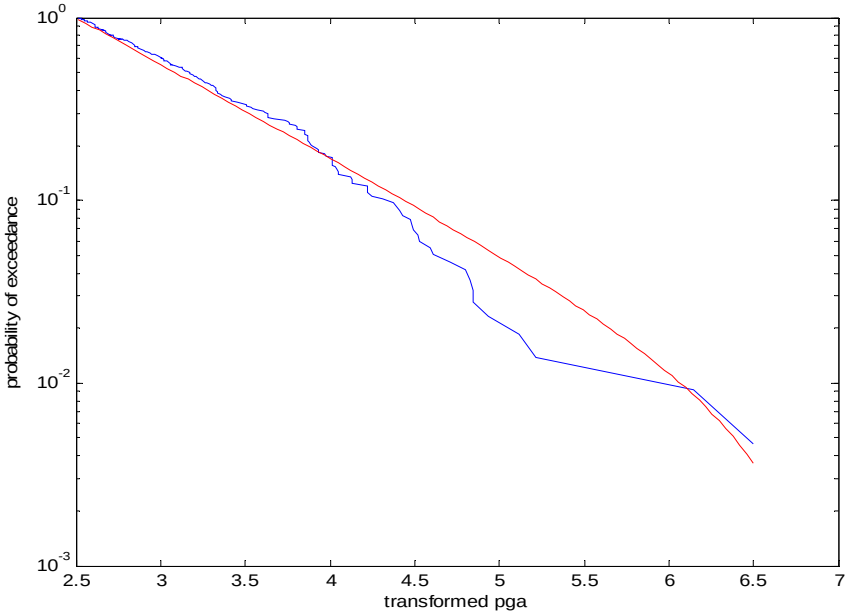

Figure 7.3 (c) Survivor plot for complete part of data from Komorniki. Red line indicates the survivor function fitted model, the blue line that of the actual data.

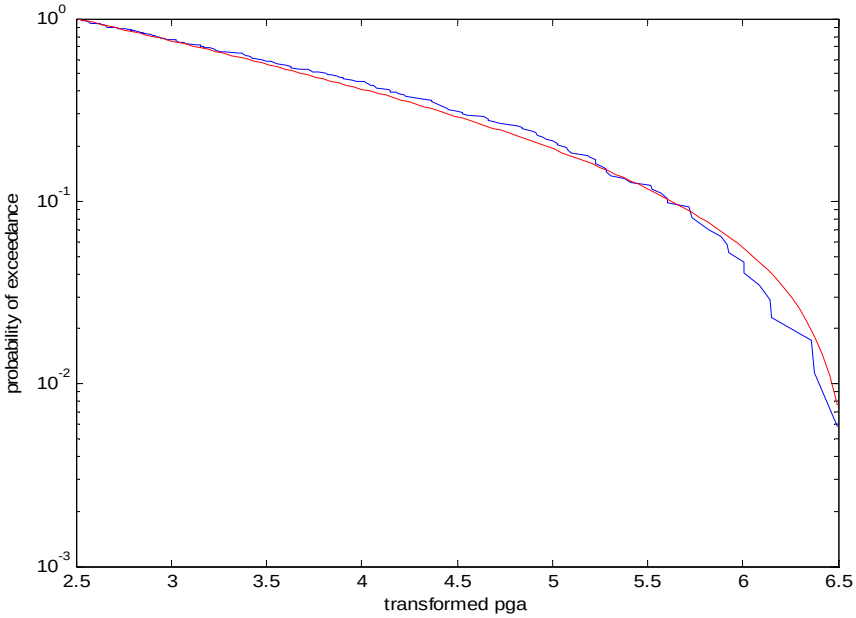

Figure 7.3 (d) Survivor plot for complete part of data from Trzebcz. Red line indicates the survivor function fitted model, the blue line that of the actual data.

© University of Pretoria

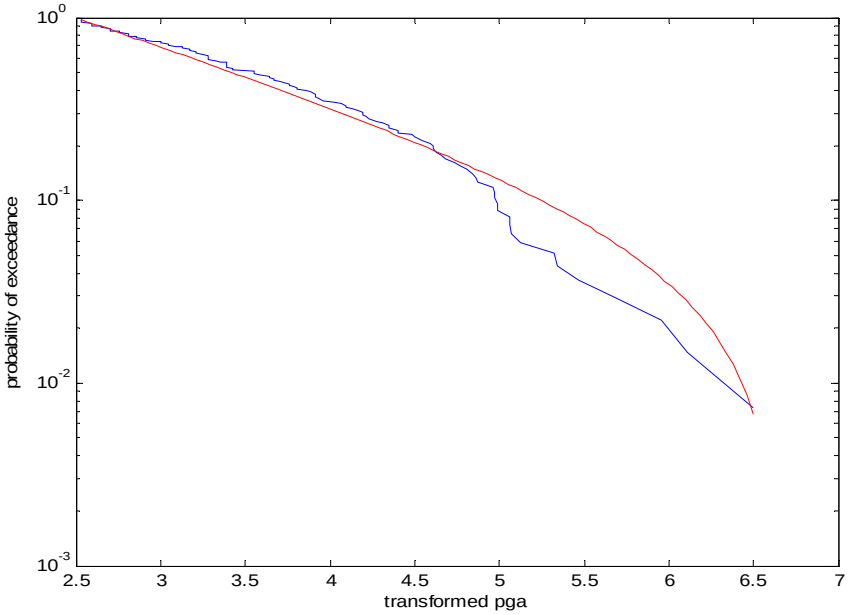

Figure 7.3 (e) Survivor plot for complete part of data from ZM2WP. Red line indicates the survivor function fitted model, the blue line that of the actual data.

© University of Pretoria

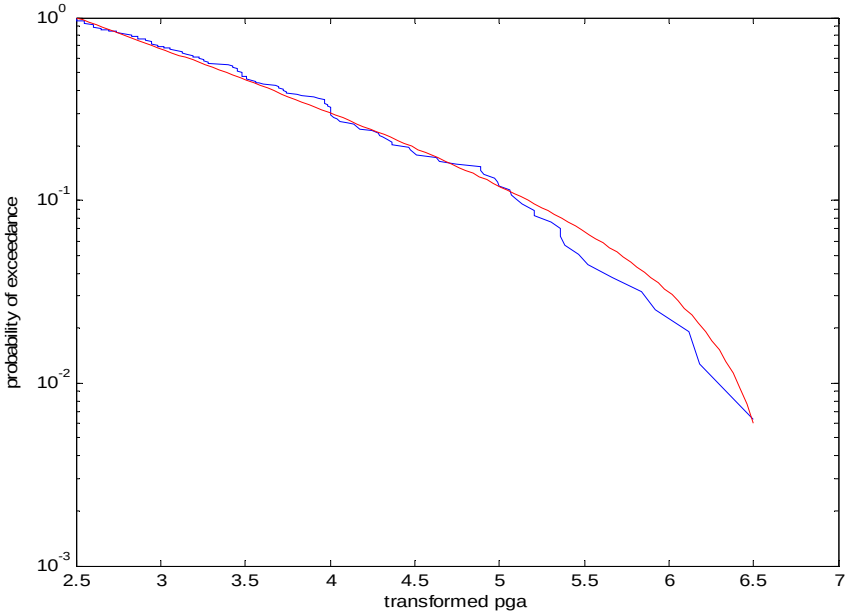

Figure 7.3 (f) Survivor plot for complete part of data from ZM8WP. Red line indicates the survivor function fitted model, the blue line that of the actual data.

© University of Pretoria

7.5. Test for possible deviation from Pareto distribution using the TP statistic

To test for possible deviation from the Pareto distribution for large values, the TP statistic discussed in Section 4.2.2 is used. The estimate is plotted against the threshold value. The estimate was applied to each station's data, the results of which are shown in Figures 7.4 (a) through (f). Monte Carlo confidence bounds at 99% confidence are also displayed. These bounds seem to indicate a bias, which is indeed so because of the boundedness of the samples (that is, because random samples were draw from an upper truncated distribution). Figure 7.5 (a) through (f) show plots of the TP statistic of data with confidence bounds determined for exponential-gamma mixture. Note that, in both Figures 7.4 and 7.5, the parameters used are those of the fitted exponential-gamma mixture (K-S-B method, section 4.2.2), but the resulting values do not differ much from those obtained by the K-S estimator (K-S method, 4.2.2).

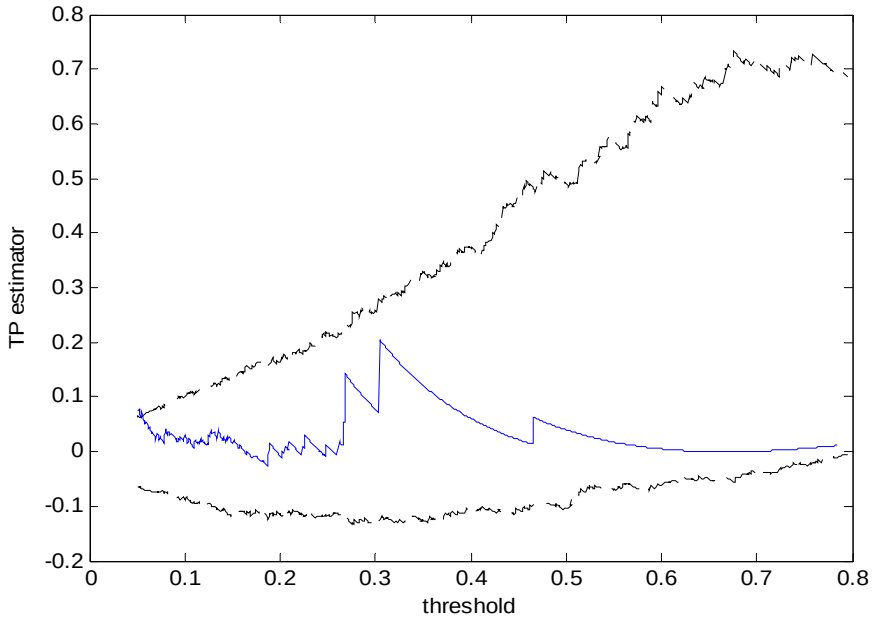

Figure 7.4 (a) TP statistic for Grodowiec; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

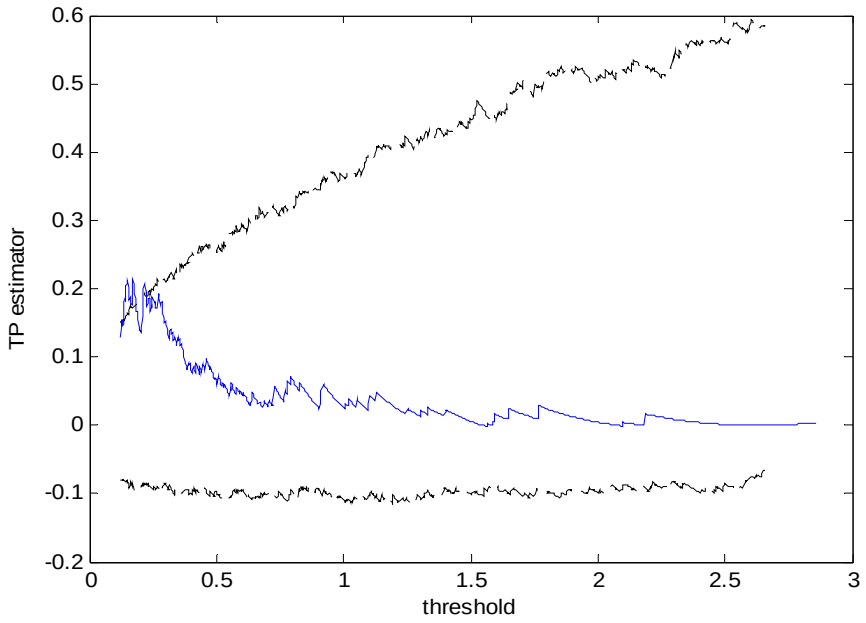

Figure 7.4 (b) TP statistic for Guzice; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

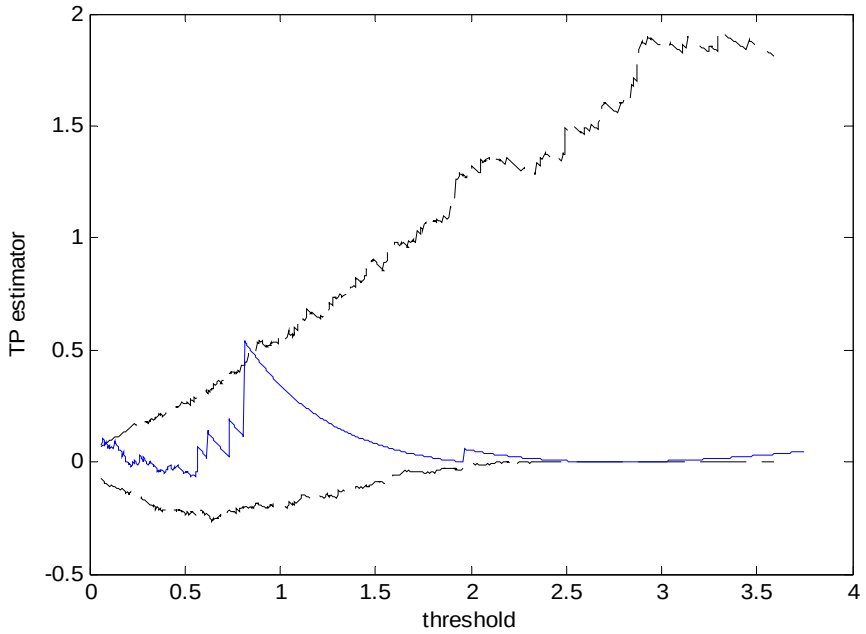

Figure 7.4 (c) TP statistic for Komorniki; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

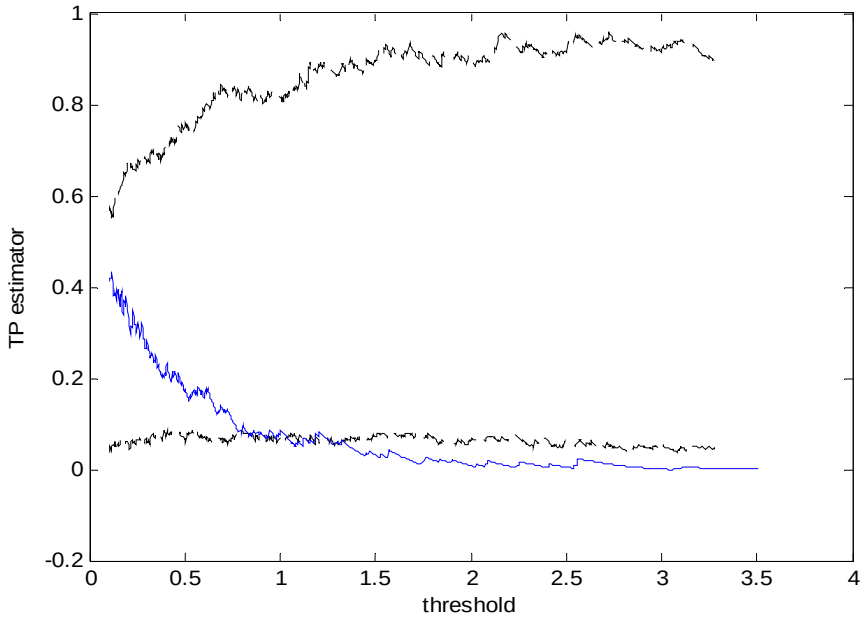

Figure 7.4 (d) TP statistic for Trzebcz; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

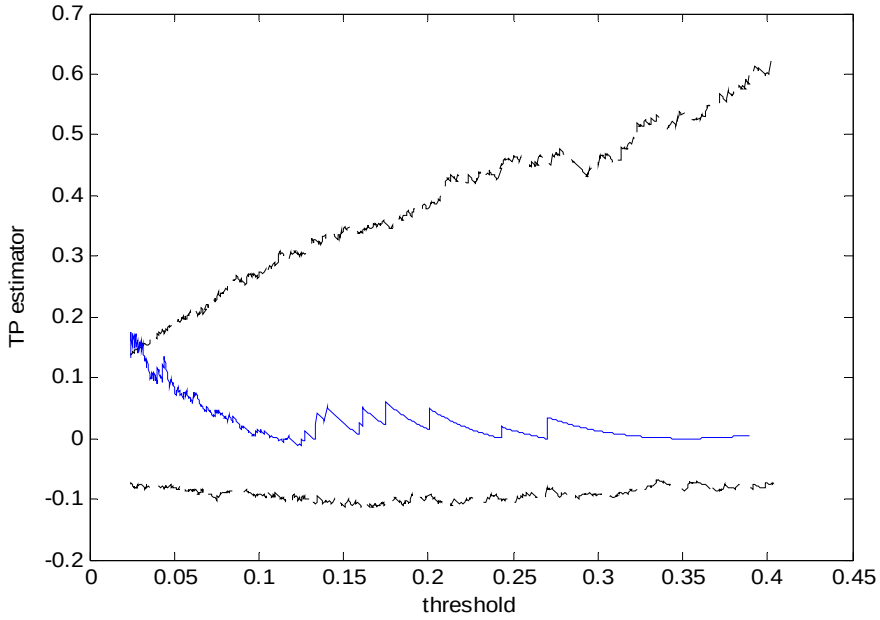

Figure 7.4 (e) TP statistic for ZM2WP; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

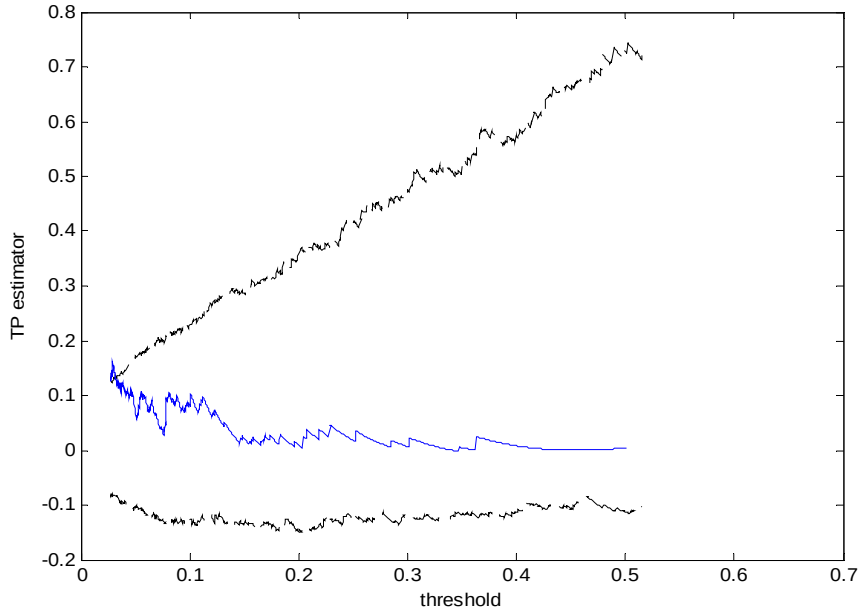

Figure 7.4 (f) TP statistic for ZM8WP; confidence bounds generated from an exponential distribution. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

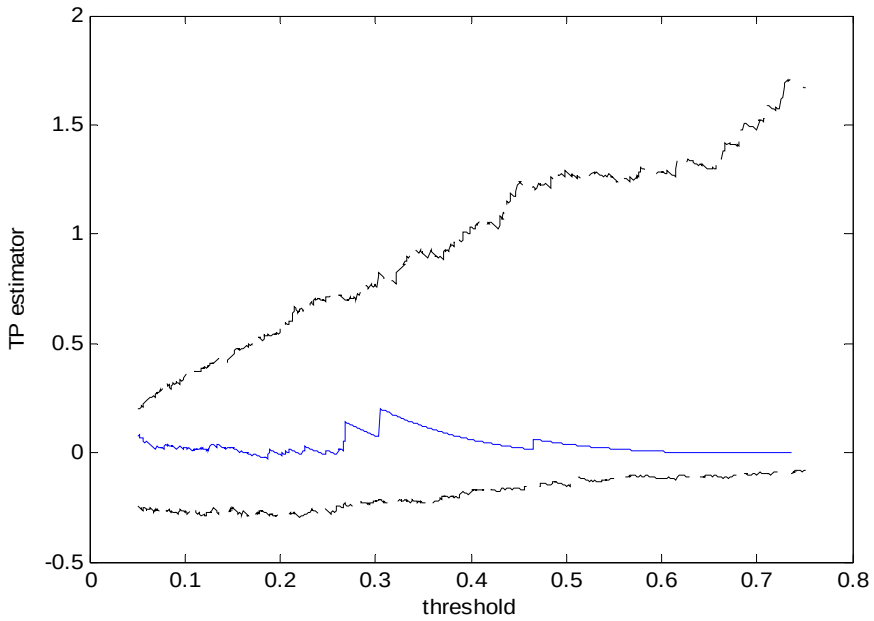

Figure 7.5 (a) TP statistic for Grodowiec; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

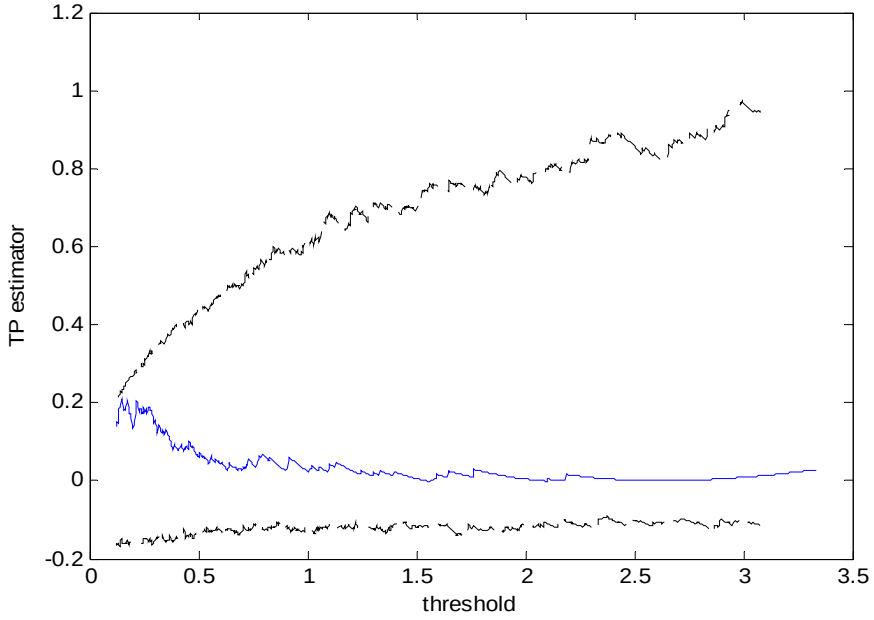

Figure 7.5 (b) TP statistic for Guzice; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

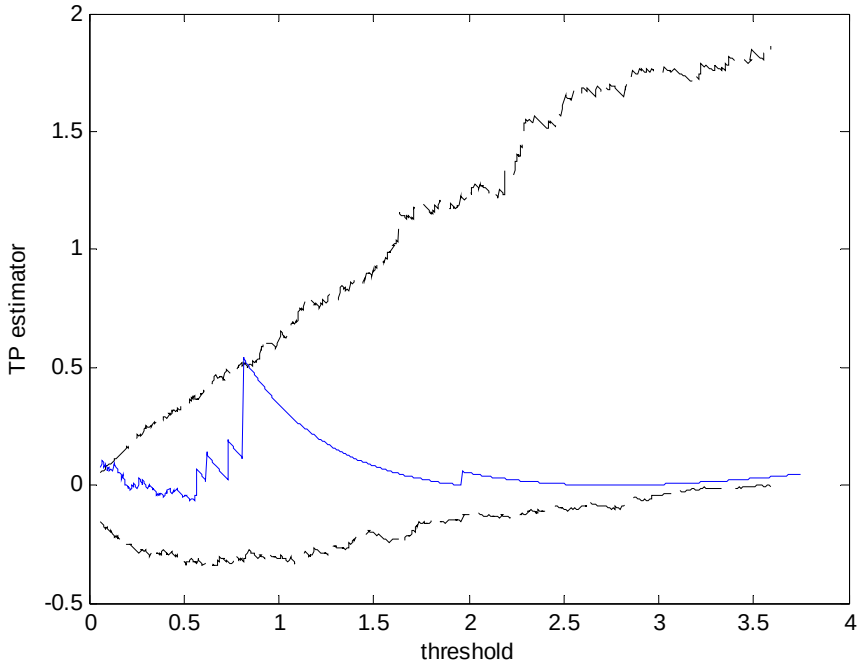

Figure 7.5 (c) TP statistic for Komorniki; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

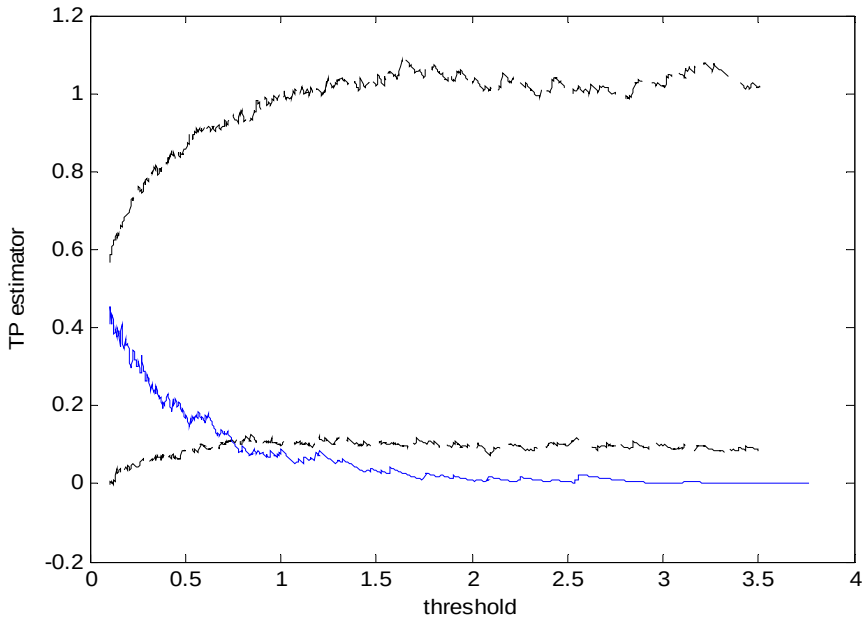

Figure 7.5 (d) TP statistic for Trebcz; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

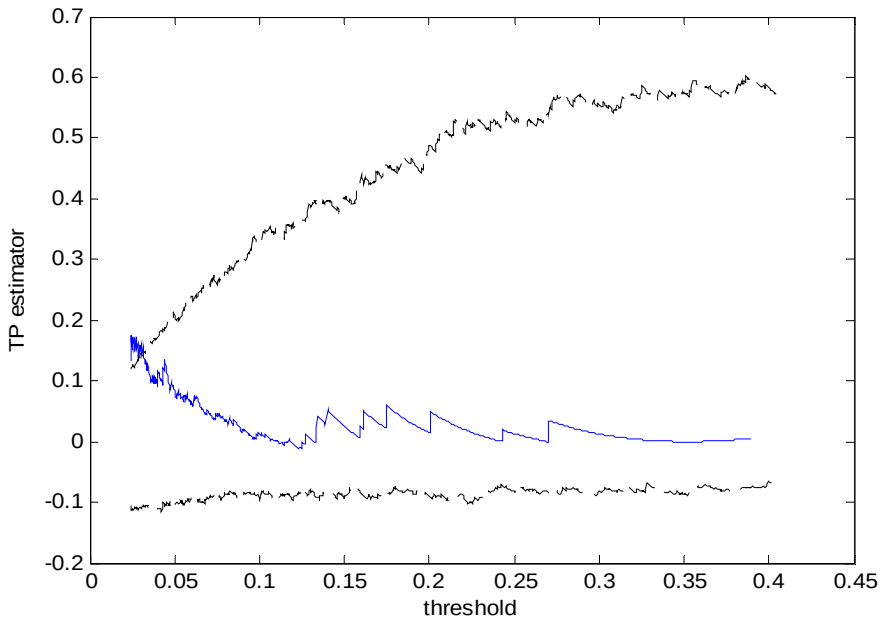

Figure 7.5 (e) TP statistic for ZM2WP; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

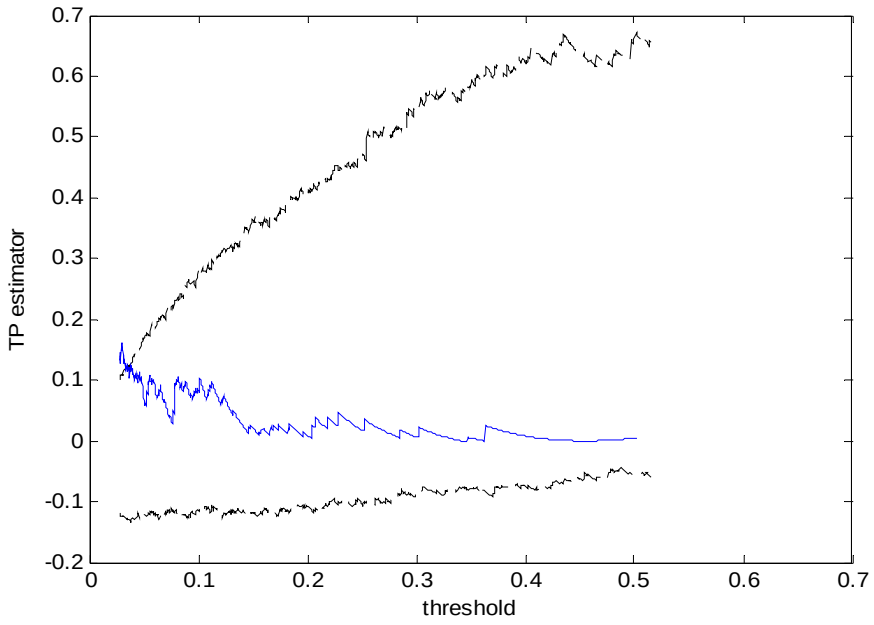

Figure 7.5 (f) TP statistic for ZM8WP; confidence bounds generated from an exponential-gamma mixture. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

7.6. Discussion

The procedure followed in this chapter, applied to an example case, served to develop the method of estimation of PGA_{max} directly from PGA records. The illustration used only one of the possible procedures discussed in section 4.2, but any of the procedures may be applied (depending on the data). Using Stephens' (1974) adapted confidence intervals of the Cramer-von Mises goodness of fit statistic the procedure seems to indicate good overall fit.

The T-P statistic serves as an indicator of deviation from a Pareto distribution in the tail end of the distribution. For all stations but Komorniki and Trzebcz do the values of the T-P statistic fall within both the 99% confidence bounds of both the exponential and exponential-gamma distributions. The data for Komorniki clearly shows an improvement from the exponential to the exponential-gamma distribution; the statistic for Trzebcz falls further outside the confidence bounds of the exponential-gamma than the exponential distribution. The author's suspicion is that the problem with this specific station might be the estimated maximum value: note the extreme bias of the confidence bounds towards positive values, whereas an infinite distribution fluctuates about zero. The statistic for Trzebcz station itself is close to zero, which suggests that its maximum value may be higher than the estimated value used. To test the plausibility of this argument the data was plotted with confidence bounds for the exponential distribution with a higher maximum (7.00 as opposed to 6.52) – see Figure 7.6. Indeed it seems a plausible argument – the statistic is much closer

© University of Pretoria

to falling within the 99% confidence bound with a higher maximum. As an aside, such use of the T-P statistic may be used as a guide to whether data is sufficient to use the generic gquation in Section 4.2.1.

As a concluding remark to this chapter, it should be noted that, at present, single station observation catalogues are usually not large enough for application as directly as was done in this chapter; the amount of single station observations in this case study is exceptionally large. Instead, PGA data may be obtained from magnitude data through the use of a GMPE. Methods are also available to remove variability (e.g. methods used by Morikawa *et al.,* 2008) , so ground motion from a region may be aggregated and the method applied exactly as described in this chapter with variations from Section 4.2.

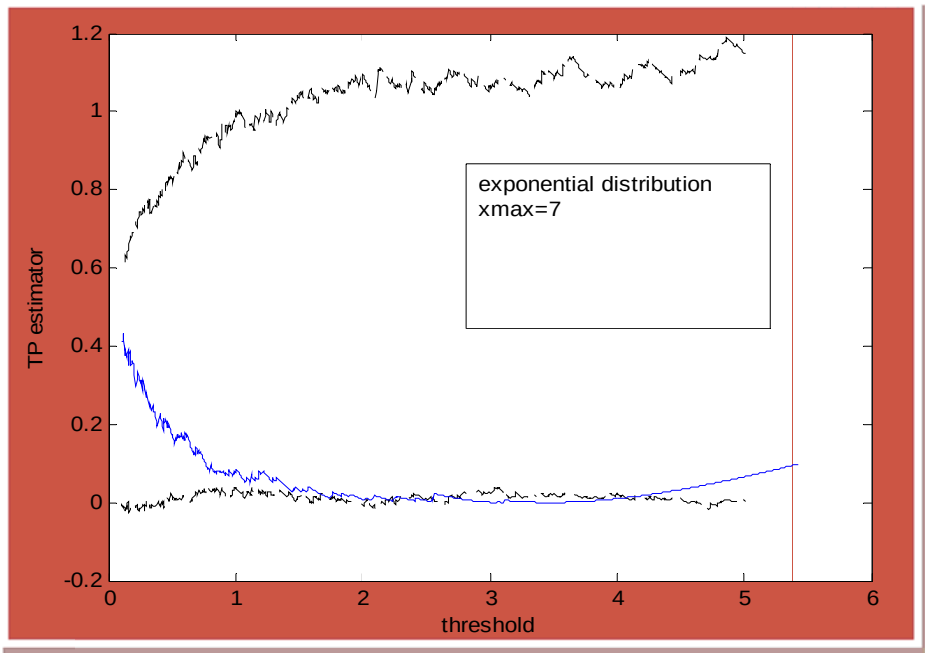

Figure 7.6 TP statistic for Trebcz; confidence bounds generated from and exponential-gamma mixture with x_{max} at a higher level than the estimated value. Blue line indicates the statistic value of the data, dotted lines the 99% conficence bounds.

© University of Pretoria

8. Incorporation of nonlinear terms of Ground motion prediction equations: theoretical results for future development

8.1. Motivation

Tsang *et al.* (2011) states that the problem with the seismogenically source-free methods is that they are based on the simplest form of the GMPE's (equation 3.11). There is thus the need to incorporate more complex terms into the parametric-historic, data-driven ground motion distribution. It is only natural to attempt to develop distributions for more complex GMPE's to refine this estimate.

8.2. Two methods that may be used to incorporate more complex functional forms of ground motion prediction equations

In this section two methods that may be used to incorporate more complex functional forms of GMPE's are proposed. The first, which will be called the *re-substitution method*, is based on a simple approximation of magnitude in terms of a common parameter (or function) that is dependent on distance and PGA. The second method, which will be called the *inverse Taylor method*, only differs in that it uses a more advanced numerical method developed by Itsikov *et al.* (2012) to compute the Taylor expansion of the inverse of a function.

A GMPE of the following general form is assumed:

$$\ln(pga) = q_1 + q_2 m + \varphi_r + \varphi_{m,r}(m) \tag{8.1}$$

($\varphi_{m,r}(m)$ also (possibly) depends on *r*). From the comprehensive overview by Douglas (2011) this functional form seems to incorporate most of the forms of GMPEs.

Now let

© University of Pretoria

$$x = x(\ln(pga), r) \coloneqq \ln(pga) - q_1 - \phi_r \tag{8.2}$$

and

$$q = q_2 \tag{8.3}$$

So we also have

$$x = x(m, r) = qm + \phi_{m,r}(m) \tag{8.4}$$

The idea is to obtain as close as possible to an explicit solution of the form

$$\Omega(x) \coloneqq m(x) \tag{8.5}$$

in order to obtain the distribution of PGA:

$$\mathrm{P}[\mathrm{PGA} \geq x] = \begin{cases} 0, & x < \ln(PGA_{min}) \\ \dfrac{1 - exp[-\beta(\Omega(x) - \ln(PGA_{min}))]}{1 - \exp[-\beta(\ln(PGA_{max}) - \ln(PGA_{min}))]}, & \ln(PGA_{min}) \leq x \leq \ln(PGA_{max}) \\ 1, & x > \ln(PGA_{max}) \end{cases} \tag{8.6}$$

Iterative re-substitution Method:

In many cases it is justified to assume (see the list of GMPES in Douglas, 2011) that

© University of Pretoria

$$qm \gg \phi_{m,r}(m) \tag{8.7}$$

It is proposed that the sequence $\Omega_n^{I.R.\cdot}(x)$ defined in the following way provides consecutively better approximations to $\Omega(x)$ (see Appendix B for a detailed motivation):

$$\begin{aligned}
\Omega_0^{I.R.\cdot}(x) &\approx \frac{x}{q} \\
\Omega_1^{I.R.\cdot}(x) &\approx \frac{x - \phi_{m,r}\left(\frac{x}{q}\right)}{q} \\
\Omega_3^{I.R.\cdot}(x) &\approx \frac{x - \phi_{m,r}\left(x - \phi_{m,r}\left(\frac{x}{q}\right)\right)}{q}
\end{aligned} \tag{8.8}$$

■ ■ ■

Inverse Taylor Method:

Note that

$$\Omega(x) = \boldsymbol{inv}\left[qm + \phi_{m,r}\right](x) \tag{8.9}$$

© University of Pretoria

where $\boldsymbol{inv}[.]$ denotes the inverse function (or, more precisely, inverse *functional*). The Taylor expansion to an arbitrary order *n*, with $n \geq 2$, is

$$\Omega(x) \approx T_n\left[\text{inv}\left(qm + \phi_{m,r}\right)\right](x) = m_0 + \sum_{k=1}^{n} \frac{(x - x_0)^k}{k!} \text{M}^{(n)}(x_0) \tag{8.10}$$

where T_n denotes the Taylor series expanded to n terms. The derivatives $\text{M}^{(n)}(x_0)$ are solved via an algorithmic method developed by Itsikov *et al*. (2012) and is explained in Appendix B.

8.3. An example of the performance of the iterative re-substitution and inverse Taylor methods

As an example case to evaluate the performance of these two methods they are applied to Atkinson and Boore's (2006) equation for the Eastern North America:

$$LogPGA = c_1 + c_2 m + c_3 m^2 + (c_4 + c_5 m) f_1 + (c_6 + c_7 m) f_2 + (c_8 + c_9 m) f_0 + c_{10} + S \tag{8.11}$$

where $f_0 = \max\left(\log\left(\frac{R_0}{R}\right), 0\right)$; $f_1 = \min(\log(R), \log(R_1))$; $f_2 = \max\left(\log\left(\frac{R}{R_2}\right), 0\right)$, and $R_0 = 10; R_{10} = 70;\ R_2 = 140$

The constants c_i are: $c_1 = 0.907$; $c_2 = 0.983$; $c_3 = -0.066$; $c_4 = -2.7$; $c_5 = 0.159$; $c_6 = -2.8; c_7 = 0.212; c_8 = -0.301; c_9 = -0.0653; c_{10} = 0.000448$.

© University of Pretoria

Expand relation on intervals $[R_{min}, R_0], (R_0, R_1], (R_1, R_2], (R_2, R_{max}]$ and arrange the resulting equations in the format

$$logPGA = c_1 + dm + \mathcal{J}_m(m) + \mathcal{J}_{m,r}(m,r) + \mathcal{J}_r(r) + S + K \tag{8.12}$$

where $\mathcal{J}_m(m)$ denotes a term depending on m alone, $\mathcal{J}_{m,r}(m,r)$ denotes terms depending on both m and r, and $\mathcal{J}_r(r)$ the combination of all terms depending on r alone. K denotes all the rest of the terms constant in r and m. The resulting expressions are as follows:

On $R_{min} \leq R \leq R_0$

$$logPGA = c_1 + (c_2 + c_9\log(R_0))m + c_3m^2 + (c_5 - c_9)m\log(R) + (c_4 - c_8)\log(R) + c_{10}R + S \tag{8.13.a}$$

On $R_0 < R \leq R_1$

$$logPGA = c_1 + c_2m + c_3m^2 + c_5m\log(\mathrm{R}) + c_4\log(R) + c_{10}R + S \tag{8.13.b}$$

On $R_1 < R \leq R_2$

$$logPGA = c_1 + (c_2 + c_5\log(R_1))m + c_3m^2 + c_{10}R + S + c_s\log(R_1) \tag{8.14.c}$$

© University of Pretoria

On $R_2 < R \leq R_{max}$

$$logPGA = c_1 + \big(c_2 + c_5 \log(R_1) - c_7 \log(R_2)\big)m + c_3 m^2 + c_7 m \log(R) + c_6 \log(R) + c_{10} R + S + c_4 \log(R_1) - c_6 \log(R_2) \quad (8.14.d)$$

Now let

$$\phi_{m,r}(m,r) = \varphi_m(m) + \varphi_{m,r}(m,r) \quad (8.15)$$

for use in the iterative re-substitution method (see equation (8.8)).

To test the accuracy of the method a set of PGA values, say $\{PGA_i\}$ were computed that correspond to a range of values of m, say $\{m_i\}$, for different values of r. Because it is then known which PGA corresponds to which value of m at a given value of r, the probability of exceedance of these values of $\{PGA_i\}$ is known to be that corresponding values $\{m_i\}$. In this way it is possible to test the accuracy of the probability values given by the proposed values by comparison to the exact values. Figure 8.1 displays the graph resulting from approximating $\Omega(x)$with $\Omega_0^{I.R.}(x)$, Figure 8.2 the result of approximating with $\Omega_1^{I.R.}(x)$, and Figure 8.3 that of $\Omega_1^{I.R.}(x)$. Figure 8.4 displays the result of the application of the inverse Taylor method with Taylor series expanded to 2 terms, Figure 8.5 to 3 terms, and Figure 8.6 expansion to 5 terms. Note that the different blue lines correspond to different values of r.

© University of Pretoria

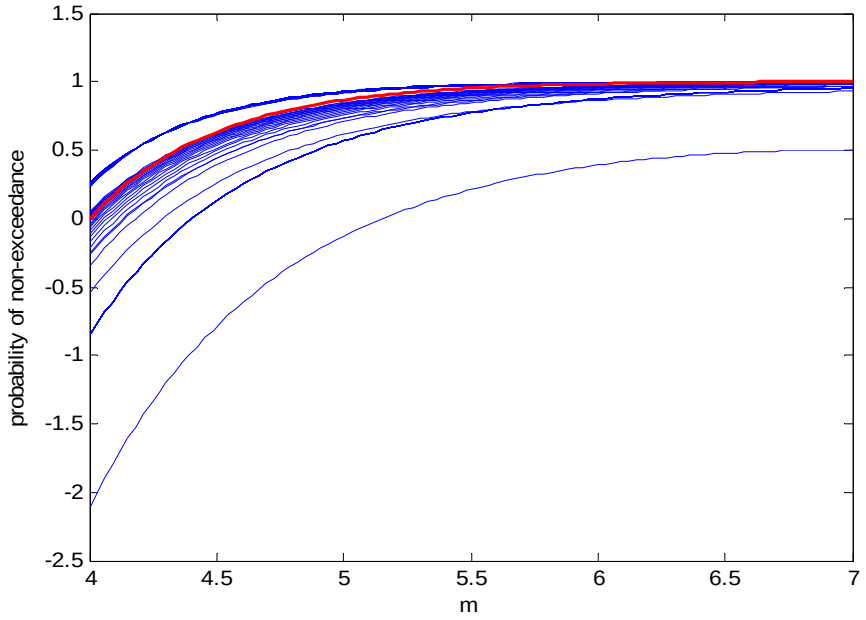

Figure 8.1 Iterative re-substitution method with no re-substitutions used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

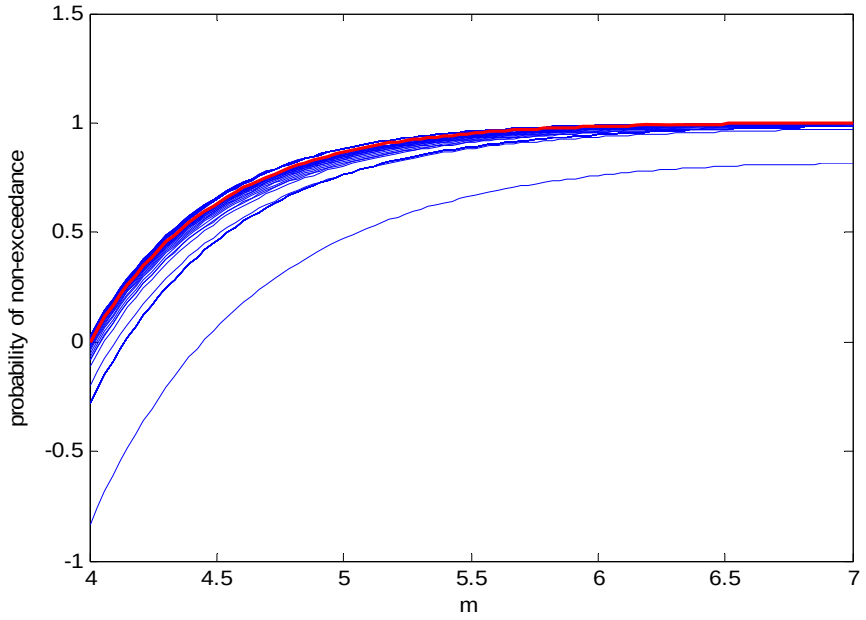

Figure 8.2 Iterative re-substitution method with one re-substitution used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

© University of Pretoria

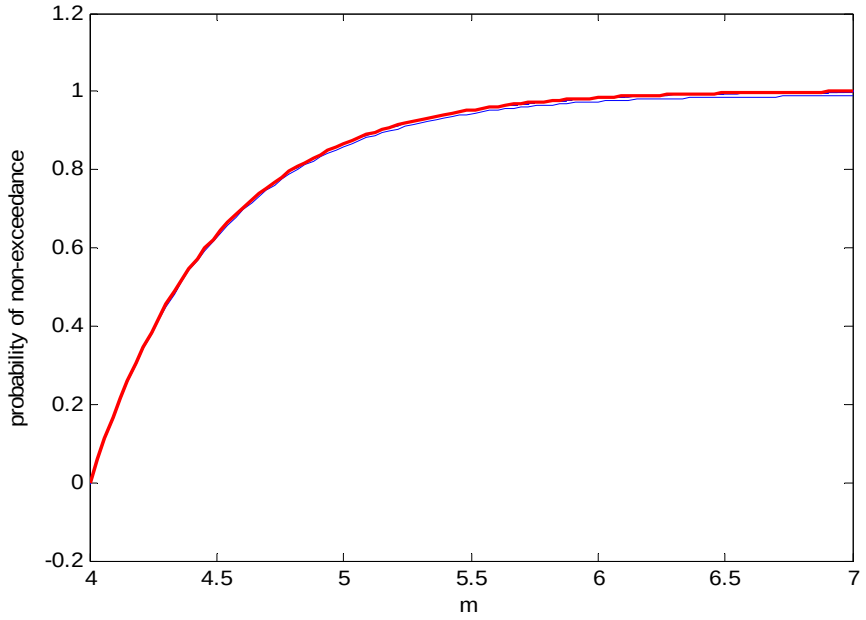

Figure 8.3 Iterative re-substitution method with ten re-substitutions used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

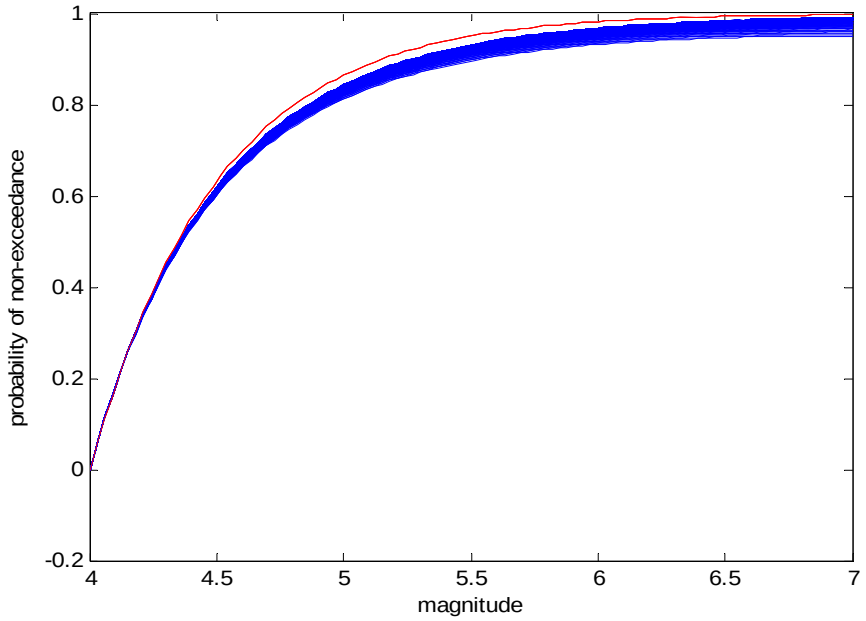

Figure 8.4 Inverse Taylor method expanded to 5 terms used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

© University of Pretoria

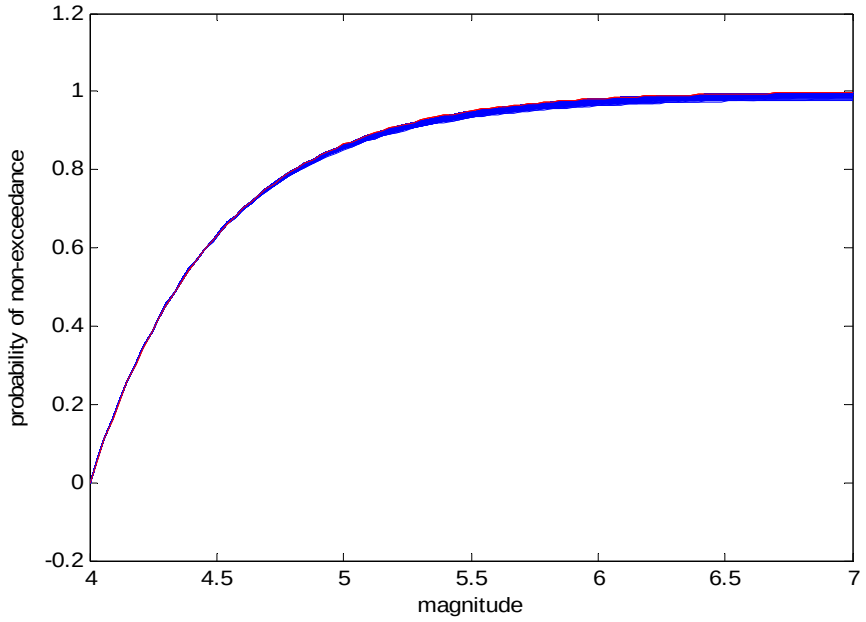

Figure 8.5 Inverse Taylor method expanded to 5 terms used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

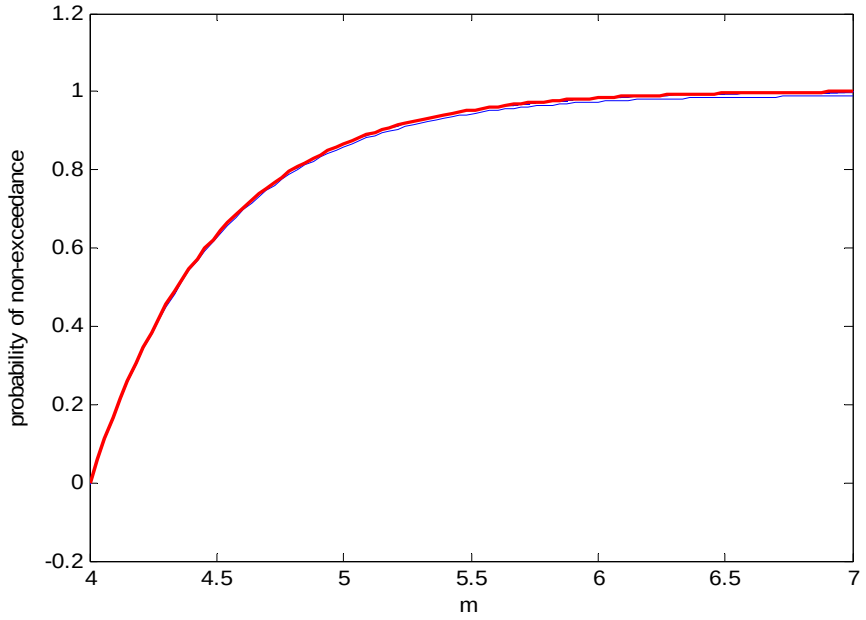

Figure 8.6 Inverse Taylor method expanded to 5 terms used to calculate probabilities of exceedance. Red line denotes actual values, blue lines approximations at constant r.

8.4. Discussion of the outcomes of the two approximations

© University of Pretoria

The iterative re-substitution method is heuristically closer to a semi-closed form or analytical method than the Inverse Taylor method because of its simplicity (although not necessarily technically so). The Inverse Taylor method is robust and the general idea is straightforward, but computationally complex with a large amount of iterative details (see Appendix B). It tends to be, heuristically, more numerical in nature (referring to the detailed calculation of the derivatives if the derivatives of the inverse function), whereas the iterative re-substitution method can more easily be comprehended as an analytical approximation. The advantage of the first method is therefore that it is easy to comprehend in its details and to implement, and is computationally less complex, whereas the second is guaranteed to yield converging under the assumption that $\Omega(x)$ is well defined. (For analytic properties of the Taylor expansion of the inverse function implemented in the second method, consult the paper by the authors of the method: Itsikov *et al.* (2012)).

© University of Pretoria

9. Final discussion:

KG1999 briefly derives an exponential distribution of the log-PGA values, which is equivalent to a Pareto distribution for PGA values. The mathematical derivation of the distribution is validated in Section 6.1 as a very good approximation under the assumption that the range of log-PGA values for a fixed magnitude and varying distance is negligible compared to the range of values log-PGA takes on for a fixed distance and varying magnitude with distance and magnitude within their respective ranges of interest. This mathematical derivation will therefore hold for cases where data from a limited range of distances have produced the earthquakes with PGA values of interest. This makes it valid in cases where large values of PGA are caused by, say, the closest earthquakes as is often the case. The effect of the error when the range of PGA values for a fixed magnitude and varying distance is similar to that compared to the range of values PGA takes on for a fixed distance and varying magnitude with distance is (theoretically) more pronounced; thus makes the nature of the error comprehensible. Figure 9.1 shows the distribution of the logarithm of PGA for a site located some distance along strike from a fault modelled as a linear source (take note that the lower part and the upper part deviates significantly). This is, however, a theoretical model and does not eliminate the applicability of the Pareto distribution to any real life data, but does serve as a caution not to use the Pareto distribution without verification of the applicability (and, in general, as caution against what Tsang *et al.* (2011) refer to as "black box syndrome"). The TP statistic used in Chapter 7 to test for possible deviation from the Pareto distribution is designed to be sensitive to tail artefacts as seen in Figure 9.1.

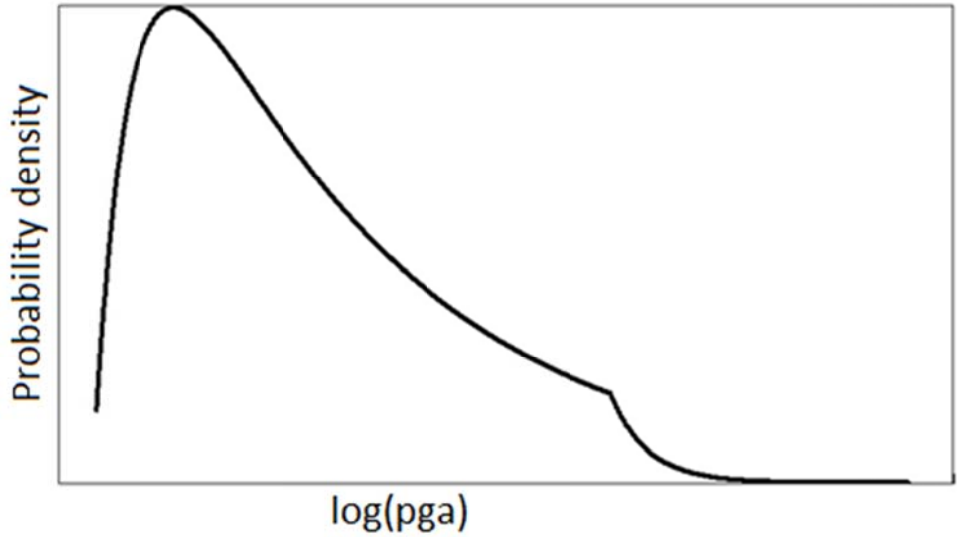

Figure 9.1 Shape of the distribution of log-PGA where the site is located some distance from a fault directly along strike. The range of PGA values for a fixed magnitude and varying distance is similar to that compared to the range of values PGA takes on, which causes large artefacts (here in the extreme) in the upper and lower tails.

© University of Pretoria

As Kijko (2012) noted, and Bommer and co-workers continually emphasise (e.g. Bommer, 2002; Bommer *et al.*, 2004; Strasser and Bommer, 2009), one of the main problems in probabilistic seismic hazard analysis is the determination of an upper bound on peak ground acceleration. PSHA provides a means of predicting recurrence of ground motions over any range that is considered plausible, but, especially because of the large uncertainty involved, is unable to estimate an upperbound on PGA from integration over all of the influencing factors such as magnitude, distance, etc. The method proposed in Chapter 7 is a data-driven approach based directly on the PGA data for a specific site. It was shown in Chapter 6 that the distribution of the log-transformed data is (at least roughly) of the same type as that of magnitude.

As an extension of the KG's Parametric-Historic method, the methods discussed in Section 4.2 for estimation of m_{max} may be applied to the PGA data (in log-transformed form where applicable for methods developed for application to magnitude, directly for methods developed for seismic moment) to estimate PGA_{max}. The methods from TM (see Section 4.2.1) accommodates both parametric and non-parametric estimates from the data. Note that the generic equation (4.12) is open to different parameterizations, even though the case of an exponential distribution has been investigated in detail in KG1998.

In the specific application in Chapter 7 a Pareto distribution of PGA data (equivalently an exponential distribution of the logarithmically transformed data) was assumed. For the sake of clarity it should be emphasised again that the data was analysed in log-transformed form (although both the Bayesian maximum likelihood estimates of parameters and Kolmogorov-Smirnov type statistics are invariant under this transformation). K-S-B method described in TM was used to estimate PGA_{max}. The flexibility to incorporate many different cases using the K-S-B method, which was used in the application, is evident from the combined flexibility of the Gamma distribution (used as a weight distribution for the exponential mixture in the method) and the flexibility of exponential mixtures. The flexibility of exponential mixtures does not, however, rule out the possibility of existence of other types of distributions (the theoretical distribution in Figure 9.1., for example, cannot be expressed as a mixture of exponential distributions due to the inflection in the lower part and the discontinuity of the first derivative in the upper part). In such cases the non-parametric methods of Section 4.2 may be applied. It should be noted that catalogues with few entries are unlikely to capture the deviation in the tail and, due to the fast decrease in probability above crossover value, and likely to result in an underestimation of the maximum.

Cases where deviations in the upper tail occur are discussed extensively by Kagan and Schoenberg (2001), Kagan (2002), Pisarenko *et al.*, (2003), Pisarenko and Sornette (2004), and

© University of Pretoria

Pisarenko and Sornette (2006). In Section 7.5 the T-P statistic of Pisarenko and co-workers (Pisarenko *et al.*, 2003; Pisarenko and Sornette, 2006) was used to test for such deviations from a Pareto distribution. For 4 out of the 6 stations' data one cannot reject the hypothesis of the PGA data following the fitted Pareto distribution within a 99% confidence bound for large values (stations Komorniki and Trzebcz were the ones raising concern). When the actual exponential-gamma mixture was used to determine the 99% confidence bounds, Komorniki still raises concern, but doesn't fall that far outside of the bounds. Trzebcz still falls far outside the confidence bounds where the exponential-gamma mixture is used. Noting that, theoretically, the T-P statistic is always zero – and that the T-P statistic for Trzebcz falls toward zero from not too large values, it is natural to suspect that the PGA_{max} is in fact underestimated for the PGA data of Trzebcz station. This hypothesis was tested *ad-hoc* by using confidence intervals for an exponential distribution with a higher upper bound (Figure 7.6). The T-P statistic falls better within these 99% confidence bounds. (As an aside, the ad-hoc test may be developed into a heuristic for the validity of our PGA_{max} estimates.) The fact that equation 4.10 (the generic equation) is amsymptotically unbiased (Kijko and Graham, 1998), so that more data always increases the accuracy, leads to the conclusion that the data for Trzebcz station was most probably insufficient.

In Chapter 8 a second extension is made to the parametric-historic procedure in the form of two numerical estimators developed to incorporate the nonlinear magnitude terms in equation (7.1). The incorporation of Atkinson and Boore's (2006) equation was considered as an example. The effect that the incorporation of non-linear terms might have on the distribution is left for future research. These distributions are, unfortunately, not source-free as the distribution derived in KG1999, but may be made so by including any solely distant dependant terms into the function dependent on both distance and magnitude, but exact conditions of convergence of iterative re-substitution method is not known and more uncertain if the approximation stated conditions in the derivation are not met.

© University of Pretoria

10. Summary

Kijko and Graham (1998 and 1999) developed a procedure they call a Parametric-Historic procedure for PSHA after Mcguire's (1993) classification of PSHA procedures as deductive, historic, and possibly parametric-historic. Kijko and Graham (1998) developed a robust method to estimate the maximum regional magnitude of an area from seismic catalogues. Kijko and Graham (1999) further developed their parametric-historic procedure where they included the description of maximum likelihood estimators of parameters from catalogues, ways to deal with inaccurate and/or incomplete catalogues, and, finally, a methodology for site-specific analysis using the catalogue data. In Kijko and Graham (1999) an equation of the distribution of the distribution of log-PGA was also briefly derived and turned out to be, independently of the source distribution, of the same form as the distribution of magnitude, that is, an exponential distribution; equivalently, PGA data follows a Pareto distribution. Other methods that preceded or followed from Kijko and Grahams' (1998) work on a point estimator of m_{max} are described in Section 4.2.1. Methods mostly concerned with a soft cutoff are discussed in Section 4.2.2.

In Chapter 6 the applicability of exponential distribution to the logarithm of PGA was validated under the assumption that the range of log-PGA values for a fixed magnitude and varying distance is negligible compared to the range of values log-PGA takes on for a fixed distance and varying magnitude with distance and magnitude within their respective ranges of interest. A theoretically justified use of the exponentially distribution would be cases where only the closest sources are considered.

In Chapter 7 a simple extension to the Parametric-Historic procedure is introduced. This involves direct application of the maximum estimators discussed in Section 4.2 directly to log-transformed PGA data. It is applied to a specific instance of data from Legnica-Głogów Copper District, Poland. The data is hypothesised to follow a Pareto distribution. The level of completeness is estimated by identifying the point on the lower range of a log-log survivor curve where it starts to deviate from linearity. The parameters $\ln(PGA_{max})$ and γ are determined from the log-transformed PGA data by the use of maximum likelihood estimators to estimate γ, and the generic equation 4.10 applied to a gamma mixture of exponential distributions (an exponential distribution where γ is allowed to vary according to a gamma distribution) is used to estimate $\ln(PGA_{max})$. A fixed point iteration is carried out, iteratively refining the pair $< \gamma, \ln(PGA_{max}) >$ by the aforementioned procedures. The Cramer-von Mises Goodness of Fit statistic was applied to individual station's data from the Legnica-Głogów Copper District and the estimated models, of which the level of confidence figures are reported in Table 7.3. Since the Cramer-von Mises Goodness of Fit statistic is

© University of Pretoria

representative of the central trend for the whole set of data, the T-P statistic introduced by Pisarenko *et al.* (2003) and Pisarenko and Sornette (2006) was employed to test for possible deviation especially in the upper tail of the distribution. The statistic is biased for a truncated distribution, but Monte-Carlo confidence bounds can still be plotted. Only for two stations was there deviation at 99% confidence level. One of these falls better into the confidence bounds if γ is allowed to vary (confidence bounds are plotted for an exponential-gamma mixture). For the other it does not, but the statistic falls to zero as the unlimited Pareto distribution does theoretically. The suspicion is that $\ln(pga_{max})$ was underestimated because of an insufficient number of observations.

In Chapter 8 another extension of the Parametric-Historic procedure: two methods of incorporating a more general form of GMPE into the distribution of PGA in a semi-closed form, one is named the re-substitution method, the other the inverse Taylor method. In both methods magnitude is solved for, which turns out not to be possible in a fully closed form explicitly. The same method of substitution may then be used that Kijko and Graham (1999) used to obtain a distribution of PGA, given the distance. In the iterative re-substitution method an approximation of magnitude (*m*) in terms of PGA and distance (*r*) is substituted after an arbitrary amount of re-substitutions of an implicit solution for *m* into itself. The inverse Taylor method involves the use of Taylor series of an inverse function, involving iterative procedures developed by Itsikov *et al*. (2012) for which each of the derivatives may be solved exactly. The iterative re-substitution method and the inverse Taylor method were applied to Atkinson and Boore's (2006) GMPE for Eastern North America. Iterative re-substitution can refine the accuracy indefinitely in this case. The inverse Taylor method is guaranteed to converge for an invertible (and 'well behaved') function, but is computationally more intensive than the iterative re-substitution method.

© University of Pretoria

11. Conclusion

- Under typical assumptions in PSHA, the doubly truncated exponential probability distribution is applicable to log-transformed peak ground motion – as suggested by Kijko and Graham (1999) – is valid under mild additional assumptions.
- For the investigated case using data from the Legnica-Głogów Copper District, the above-mentioned theoretical form turns out to be plausible even outside the additional theoretical restrictions.
- If the doubly truncated exponential distribution of log-transformed peak ground motion data is plausible, statistics developed for estimating m_{max} may also be used to estimate $\ln(pga_{max})$.
- As a corollary, hard cutoff models and soft cutoff models may be used in a complimentary way, as was done in section 7.5. This is opposed to the idea that hard cutoff estimations and soft cutoff estimations belong to mutually exclusive schools of thought.
- The iterative re-substitution and inverse Taylor methods, introduced in Chapter 8, may be used to incorporate more general GMPE's into a distribution for peak ground motion data than the doubly truncated exponential distribution, albeit in a semi-closed form.

© University of Pretoria

12. Future research

The Pareto distribution as a model of PGA at a given site has only been validated for a relatively limited case. It has been observed in Chapters 5 and 6, and previously by other authors (e.g. Pisarenko *et al.*, 1996; Newman, 2005), that ground motion at a specific site seems to follow a Pareto distribution. Further verification of the extent of this hypothesis is still necessary, both empirically as larger single station data becomes available, and theoretically in terms of physical factors or mathematical derivations.

The method that was proposed for estimation of PGA_{max} is straightforward given the variety of methods discussed in Section 4.2.1, but applicability of the methods gives more reason for further development on the generic equation, both further work on the particular functional forms already developed to some extent (these were not exhaustively developed), and development of closed forms and approximations for functional forms not yet considered. For example: (1) it would be beneficial to have some quantification of the amount of data necessary to make a reliable estimate of the maximum value in the case of an exponential or exponential-gamma distribution (as in the case in Section 7.5 and 7.6 of station Trzebcz for which it was suspected that the number of observations are not enough); (2) the ad-hoc test used in section 7.6 may be developed as a qualitative heuristic to determine if the number of observations is enough to make a reliable estimate of the maximum value; (3) if deviations from the exponential distribution does occur at some crossover value it is expedient to know what is the effect of a subtle crossover value on the estimate of the maximum value; (4) determine what other distributions could be used to model the distribution ground motion data, and develop an estimator along the same lines as KG1998; (5) determine the applicability and limitations of exponential mixtures as models for log-transformed PGA.

The extension of the parametric-historic procedure by incorporation of more general GMPEs in Chapter 8 was not developed exhaustively: (1) rigorous conditions for the convergence of the iterative re-substitution method needs to be determined; (2) some quantitative or qualitative measure for the amount of iterative substitutions in the iterative re-substitution and the number of terms in the expansion in the inverse Taylor method to obtain satisfactory convergence may be useful.

Thus far hard cutoffs (or point estimators) and soft cutoff-based statistics were developed as two competing schools of thought. In Sections 7.4 through 7.6 methods from both

© University of Pretoria

schools were used in a complementary way. More detailed investigation into the trade-offs and complements between the two schools should be made.

© University of Pretoria

13. Bibliography

Aki, K. (1984). Asperities, barriers, characteristic earthquakes and strong motion prediction. *Journal of Geophysical Research*, 89(B7): 5867-5872.

Ambraseys, N. N. (1970). Factors controlling the earthquake response of foundation materials. *Proceedings of the Third European Symposium on Earthquake Engineering*: 309-317

Amorèse, D. (2007). Applying a change-point detection method on frequency-magnitude distributions. *Bulletin of the Seismological Society of America*, 97(5): 1742-1749.

Andrews, D. J. (1976). Rupture velocity of plane strain shear cracks. *Journal of Geophysical Research*, 81(32): 5679-5687.

Andrews, D. J., Hanks, T. C., Whitney, J. W. (2007). Physical limits on ground motion at Yucca Mountain. *Bulletin of the Seismological Society of America*, 97(6): 1771-1792.

Anderson, J. G. (1979). Estimating the seismicity from geological structure for seismic-risk studies. *Bulletin of the Seismological Society of America*, 69(1): 135-158.

Anderson, J. G., Luco, J. E. (1983). Parametric study of near-field ground motions for oblique-slip and dip-slip dislocation models. *Bulletin of the Seismological Society of America*, 73(1): 45-57.

Arnold, B. C. (2008). Pareto and generalized pareto distributions. *In Modeling Income Distributions and Lorenz Curves*: 119-145. Springer, New York.

Atkinson, G. M., Boore, D. M. (2006). Earthquake ground-motion prediction equations for eastern North America. *Bulletin of the Seismological Society of America*: 96(6): 2181-2205.

Atkinson, G. M. (2011). An empirical perspective on uncertainty in earthquake ground motion prediction 1 This paper is one of a selection of papers in this Special Issue in honour of Professor Davenport. *Canadian Journal of Civil Engineering*, 38(9): 1002-1015.

Bak, P., & Tang, C. (1989). Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research: Solid Earth*, 94(B11): 15635-15637.

Bak, P. (1997). *How nature works*. Oxford: Oxford university press.

Bartle, R. G. (1964). *The elements of real analysis*. New York: John Wiley and Sons.

© University of Pretoria

Ben-Zion, Y., Sammis, C. G. (2003). Characterization of fault zones. *Pure and Applied Geophysics*: 160(3-4), 677-715.

Ben-Zion, Y. (2008). Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Reviews of Geophysics*, 46(4): RG4006.

Beresnev, I. A., & Wen, K. L. (1996). Nonlinear soil response—a reality?. Bulletin of the Seismological Society of America, 86(6), 1964-1978.

Biot, M. A. (1932). Transient oscillations in elastic systems (Doctoral dissertation, California Institute of Technology).

Biot, M. A. (1943). Analytical and experimental methods in engineering seismology. *Transactions of the American Society of Civil Engineers*, 108(1): 365-385.

Bommer, J.J. (2002). Deterministic vs. probabilistic seismic hazard assessment: an exaggerated and obstructive dichotomy. *Journal of Earthquake Engineering*, 6(special issue no. 1): 43-73.

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

Bommer, J. J., & Abrahamson, N. A. (2006). Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bulletin of the Seismological Society of America*, 96(6): 1967-1977.

Boore, D. M. (1973). The effect of simple topography on seismic waves: implications for the accelerations recorded at Pacoima Dam, San Fernando Valley, California. *Bulletin of the Seismological Society of America*, 63(5): 1603-1609.

Boore, D. M., Joyner, W. B. (1982). The empirical prediction of ground motion. *Bulletin of the Seismological Society of America*, 72(6B): S43-S60.

Brune, J. N. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *Journal of geophysical research*, 75(26): 4997-5009.

Brune, J. N. (1999). Precarious rocks along the Mojave section of the San Andreas Fault, California: Constraints on ground motion from great earthquakes. *Seismological Research Letters*, *70*(1), 29-33

© University of Pretoria

Budnitz, R. J., Apostolakis, G., Boore, D.M., Cluff, L.M., Coppersmith, K.J., Cornell, C.A., Morris, P.A. (1997). *Recommendations for probabilistic seismic hazard analysis: guidance on uncertainty and use of experts* (Vol. 1). Washington, DC: US Nuclear Regulatory Commission.

Burridge, R. (1973). Admissible speeds for plane-strain self-similar shear cracks with friction but lacking cohesion. *Geophysical Journal of the Royal Astronomical Society*, 35(4): 439-455.

Caserta, S., & Vries, C. D. (2003). Extreme value theory and statistics for heavy tail data. In: Field, P. (ed.) *Modern Risk Management - A History*: 169-178. London: Risk Books

Campbell, K. W. (2003), Engineering models of strong ground motion. In: Chen, W., Shawthorn, C. (eds.) *Earthquake Engineering Handbook*. pp. 205-280. Boca Raton, Florida, USA: CRC Press

Campbell, K. W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America. *Bulletin of the Seismological Society of America*, 93(3): 1012-1033.

Cao, A., Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical research letters*, 29(9): 48-1 – 48-3

Carlson, J. M., Langer, J. S., Shaw, B. E. (1994). Dynamics of earthquake faults. Reviews of Modern Physics, 66(2): 657-670.

Chen, J. T., Lee, J. W., Shyu, W. S. (2012). SH-wave scattering by a semi-elliptical hill using a null-field boundary integral equation method and a hybrid method. *Geophysical Journal International*, 188(1): 177-194.

Convertito, V., & Herrero, A. (2004). Influence of focal mechanism in probabilistic seismic hazard analysis. *Bulletin of the Seismological Society of America*, 94(6): 2124-2136.

Convertito, V., & Herrero, A. (2006). Reply to "Comment on 'Influence of Focal Mechanism in Probabilistic Seismic Hazard Analysis' by Vincenzo Convertito and André Herrero," by FO Strasser, V. Montaldo, J. Douglas, and JJ Bommer.*Bulletin of the Seismological Society of America*, 96(2): 754-756.

Cooke, P. (1979), Statistical inference for bounds for random variables, *Biometrika,* 66(2): 367-374

Cooke, P. (1980), Optimal linear estimation of bounds of random variables, *Biometrika,* 67(1): 257-258.

© University of Pretoria

Corradini, M. L. (2003). *Letter from the chairman of the U.S. Nuclear Waste Technical Review Board to the Director of the Office of Civilian Radioactive Waste Management.* [Online] Available from: http://www.nwtrb.gov/corr/mlc010.pdf (last accessed May 2013).

Cornell, C. A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, 58(5): 1583-1606.

Cornell, C. A., & Winterstein, S. R. (1988). Temporal and magnitude dependence in earthquake recurrence models. *Bulletin of the Seismological Society of America*, 78(4): 1522-1537.

Das, S. (2003). Spontaneous complex earthquake rupture propagation. *Pure and Applied Geophysics*, 160(3-4): 579-602.

De Haan, L. L., Ferreira, A. (2006). Extreme value theory. New York: Springer.

Douglas, J. (2011). *Ground motion prediction equations 1964-2010*, BRGM/RP-59356-FR. 444 pp. 9 illustrations

Embrechts, P. (1997). *Modelling extremal events: for insurance and finance* (Vol. 33). New York: Springer.

Erickson, B. A., Birnir, B., & Lavallée, D. (2011). Periodicity, chaos and localization in a Burridge–Knopoff model of an earthquake with rate-and-state friction. *Geophysical Journal International*. 187(1): 178-198.

Gazetas, G. (1982). Vibrational characteristics of soil deposits with variable wave velocity. *International Journal for Numerical and Analytical Methods in Geomechanics*, 6(1): 1-20.

Giardini, D., Grünthal, G., Shedlock, K. M., & Zhang, P. (1999). The GSHAP global seismic hazard map. *Annals of Geophysics*, 42(6).

Gibowicz, S. J., Kijko, A. (1994). *An introduction to mining seismology*. San Diego: Academic Press.

Gnedenko, B. (1943), Sur la distribution limite du terme maximum d'unesérie aléatoire, *Annnals of Mathematics*, 44(3): 423-453 (in French)

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

Hájek, J., Šidák, Z., & Sen, P. K. (1967). *Theory of rank tests* (p. 297). New York: Academic press.

© University of Pretoria

Hanks, T. C., & Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth,* 84(B5): 2348-2350.

Hanks, T. C., N. A. Abrahamson, M. Board, D. M. Boore, J. N. Brune, C. A. Cornell (2006). In: Gulkan, P., Anderson, J. G. (eds.) *Observed ground motions, extreme ground motions, and physical limits to ground motions, in Future Directions in Strong Motion Instrumentation*: 55–59. New York: Springer.

Hong, H. P. Goda, K. (2007). Orientation-dependent ground-motion measure for seismic-hazard assessment. *Bulletin of the Seismological Society of America*, 97(5): 1525-1538.

Housner, G. W. (1941). An investigation of the effects of earthquakes on buildings (Doctoral dissertation, California Institute of Technology).

Irikura, K., & Kurahashi, S. (2012). Strong ground motions during the 2011 Pacific coast off Tohoku, Japan Earthquake. In *International Symposium on Engineering Lessons Learned from the Giant Earthquake*.

Itskov, M., Dargazany, R., & Hörnes, K. (2012). Taylor expansion of the inverse function with application to the Langevin function. *Mathematics and Mechanics of Solids*: 17(7), 693-701.

Johnson, K., & Kotz, S., Balakrishnan, N. (1994). Continuous Univariate Distributions. Vol 1 (2nd edition) John Wiley and Sons Inc. 756 pp.

Kagan, Y. Y., & Knopoff, L. (1980). Spatial distribution of earthquakes: the two-point correlation function. Geophysical Journal of the Royal Astronomical Society, 62(2), 303-320.

Kagan, Y. Y. (1981). Spatial distribution of earthquakes: the four-point moment function. Geophysical Journal International, 67(3), 719-733.

Kagan, Y. Y., & Schoenberg, F. (2001). Estimation of the upper cutoff parameter for the tapered Pareto distribution. *Journal of Applied Probability*, 38: 158-175.

Kagan, Y. Y. (2002). Seismic moment distribution revisited: I. Statistical results. Geophysical Journal International, 148(3), 520-541.

Kagan, Y. Y. (2007). Earthquake spatial distribution: the correlation dimension. *Geophysical Journal International*, 168(3): 1175-1194.

© University of Pretoria

Kanamori, H., & Anderson, D.L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, 65(5): 1073-1095.

Kapur, J. N., & Kesavan, H. K. (1992). *Entropy optimization principles with applications*. San Diego: Academic Press.

Kawamura, H., Hatano, T., Kato, N., Biswas, S., & Chakrabarti, B. K. (2012). Statistical physics of fracture, friction, and earthquakes. Reviews of Modern Physics, 84(2), 839.

Kijko, A., & Sellevolll, M. A. (1989). Estimation of earthquake hazard parameters from incomplete data files. Part I. Utilization of extreme and complete catalogs with different threshold magnitudes. *Bulletin of the Seismological Society of America*, 79(3): 645-654.

Kijko, A., & Sellevolll, M. A. (1992). Estimation of earthquake hazard parameters from incomplete data files. Part II. Incorporation of magnitude heterogeneity. *Bulletin of the Seismological Society of America*, 82(1): 120-134.

Kijko A, Graham G (1998) "Parametric-historic" Procedure for Probabilistic Seismic Hazard Analysis Part I: Estimation of Maximum Regional Magnitude mmax. *Pure and Applied Geophysics,* 152:413–442

Kijko, A., Graham, G. (1999). "Parametric-historic" Procedure for Probabilistic Seismic Hazard Analysis Part II: Assessment of Seismic Hazard at Specified Site. *Pure and Applied geophysics*, 154(1): 1-22.

Kijko, A., Lasocki, S., & Graham, G. (2001). Non-parametric seismic hazard in mines. *Pure and Applied Geophysics*, 158(9-10): 1655-1675.

Kijko, A. (2011). Seismic Hazard. In Gupta, H. (2011) (Ed.). *Encyclopedia of Solid Earth Geophysics.* ISBN: 978-90-481-8707-0. Berlin: Springer, 2011., 1.Chicago Klügel, J. U. (2007). Error inflation in probabilistic seismic hazard analysis.*Engineering Geology*, *90*(3), 186-192.

Kijko A. (2004) Estimation of the maximum earthquake magnitude mmax. *Pure and Applied Geophysics*, 161:1655–1681

Kijko, A. (2006). *HA2 version 2.05*, Matlab software package: Pretoria, University of Pretoria. Available from the author on request.

© University of Pretoria

Kijko, A., & Singh, M. (2011). Statistical tools for maximum possible earthquake magnitude estimation. *Acta Geophysica*, 59(4): 674-700.

Knopoff, L., &Kagan, Y. (1977). Analysis of the theory of extremes as applied to earthquake problems. *Journal of Geophysical Research*, 82(36): 5647-5657.

Kramer, S. L. (1996). Geotechnical earthquake engineering. *Prentice-Hall Civil Engineering and Engineering Mechanics Series,* Upper Saddle River, New Jersey: Prentice Hall

Klügel, Jens-Uwe. *Seismic Hazard Analysis—Quo vadis*?. *Earth-Science Reviews* 88.1 (2008): 1-32.

Latchman, J. L., Morgan, F. D. O., & Aspinall, W. P. (2008). Temporal changes in the cumulative piecewise gradient of a variant of the Gutenberg–Richter relationship, and the imminence of extreme events. *Earth-Science Reviews*, 87(3): 94-112.

Lasocki, S. (2005). Probabilistic analysis of seismic hazard posed by mining induced events. In Proc. 6th Int. Symp. on Rockburst in Mines "Controlling Seismic Risk". ACG, Perth: 151-156.

Lasocki, S., & Urban, P. (2011). Bias, variance and computational properties of Kijko's estimators of the upper limit of magnitude distribution, Mmax. *Acta Geophysica*, 59(4): 659-673.

Lay, T., & Wallace, T. C. (1995). *Modern global seismology* (Vol. 58). San Diego: Academic Press.

McGruire, R. K., (1993). Computations of seismic hazard, *Annali di Geofisica* 36(3-4):181-200

McGuire, R. K. (2001). Deterministic vs. probabilistic earthquake hazards and risks. *Soil Dynamics and Earthquake Engineering*, 21(5): 377-384.

McGuire, R. K. (2008). Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering & Structural Dynamics*, 37(3): 329-338.

Miller, K. S., & Samko, S. G. (2001). Completely monotonic functions. *Integral Transforms and Special Functions*, 12(4): 389-402.

Mignan, A., Werner, M. J., Wiemer, S., Chen, C. C., & Wu, Y. M. (2011). Bayesian estimation of the spatially varying completeness magnitude of earthquake catalogs. *Bulletin of the Seismological Society of America*, *101*(3): 1371-1385.

© University of Pretoria

Morikawa, N., Kanno, T., Narita, A., Fujiwara, H., Okumura, T., Fukushima, Y., & Guerpinar, A. (2008). Strong motion uncertainty determined from observed records by dense network in Japan. *Journal of seismology*, 12(4): 529-546.

Naylor, M., Greenhough, J., McCloskey, J., Bell, A. F., & Main, I. G. (2009). Statistical evaluation of characteristic earthquakes in the frequency-magnitude distributions of Sumatra and other subduction zone regions. *Geophysical Research Letters*, 36(20).

Newman, Mark EJ. Power laws, Pareto distributions and Zipf'slaw. *Contemporary physics* 46(5): 323-351.

Nieboer, D. (2011). Heuristics of heavy-tailed distributions and the Obesity index (M.Sc. Thesis, Delft University of Technology).

Orleca-Sikora, B., Cesca, S., Lasocki, S., Lizurek, G., Wiejacz, P., Rudziński Ł., Urban, P., Kozłowska, M. (2012). Complex source mechanisms of mining-induced seismic events and their exceptional surface impacts: the case from Rudna copper-ore mine in Poland. *Journal of Geophysical Research*, submitted.

Page, R. (1968). Aftershocks and microaftershocks of the great Alaska earthquake of 1964. *Bulletin of the Seismological Society of America*, 58(3): 1131-1168.

Pacific Earthquake Engineering Research Center (2005 edition; first published 1999).*PEER NGA Database Flatfile*. [Online] Available from: http://peer.berkeley.edu/nga/flatfile.html [Last accessed: 2013-09-02]

Pisarenko VF (1991). Statistical evaluation of maximum possible earthquake. *Physics of the Solid Earth* 27:757–763

Pisarenko, V.F., A.A. Lyubushin, V.B. Lysenko, and T.V. Golubieva (1996). Statistical estimation of seismic hazard parameters: Maximum possible magnitude and related parameters, *Bulletin of Seismological Society of America.* 86(3): 691-700

Pisarenko, V. F., & Lyubushin, A. A. (1997). Statistical estimation of maximum peak ground acceleration at a given point of a seismic region. *Journal of Seismology*, 1(4): 395-405

Pisarenko, V., Sornette, D., Rodkin, M. (2003). Deviations of the distributions of seismic energies from the Gutenberg-Richter law. *arXiv preprint physics/0312020*.

© University of Pretoria

Pisarenko, V. F., Sornette, D. (2004). Statistical detection and characterization of a deviation from the Gutenberg-Richter distribution above magnitude 8. *pure and applied geophysics*, *161*(4), 839-864.

Pisarenko, V., Sornette, D. (2006). New statistic for financial return distributions: Power-law or exponential?. *Physical and Statistical Mechanics and its Applications*, 366: 387-400.

Pisarenko, V.F., A. Sornette, D. Sornette and M.V. Rodkin (2008a). New Approach to the Characterization of Mmax and of the Tail of the Distribution of Earthquake Magnitudes, Pureand Applied Geophysics 165, 847-888.

Pisarenko, V. F., Sornette, A., Sornette, D., & Rodkin, M. V. (2008b). Characterization of the Tail of the Distribution of Earthquake Magnitudes by combining the GEV and GPD descriptions of Extreme Value Theory. arXiv preprint arXiv:0805.1635.

Pisarenko, V., Rodkin, M., (2014). *Statistical analysis of natural disasters and related losses*. (Uncorrected Proof) Springer, 10.1007/978-3-319-01454-8

Power, M., Chiou, B., Abrahamson, N., Bozorgnia, Y., Shantz, T., Roblee, C. (2008). An overview of the NGA project. *Earthquake Spectra*, 24(1), 3-21.

Raschke, M. (2012). Inference for the truncated exponential distribution. *Stochastic Environmental Research and Risk Assessment*, 26(1): 127-138.

Reiter, L. (1990). Earthquake hazard analysis: issues and insights. Columbia University Press.

Resnick, S. (2007). Heavy-tail phenomena: probabilistic and statistical modeling. New York: Springer

Rial, J. A. (1984). Caustics and focusing produced by sedimentary basins: applications of catastrophe theory to earthquake seismology. *Geophysical journal of the Royal Astronomical Society*, *79*(3), 923-938.

Robinson, D. P., Henry, C., Das, S., Woodhouse, J. H. (2001).Simultaneous rupture along two conjugate planes of the Wharton Basin earthquake. *Science*, 292(5519): 1145-1148. (Supplementary material)

Robson, D.S., and J.H. Whitlock (1964), Estimation of a truncation point, *Biometrika,* 51(1-2): 33-39.

© University of Pretoria

Rydelek, P.A., Sacks, I.S. (1989). Testing the completeness of earthquake catalogues and the hypothesis of self-similarity. *Nature*, 337(6204): 251-253

Schorlemmer, D., Woessner, J. (2008). Probability of detecting an earthquake. *Bulletin of the Seismological Society of America*, 98(5): 2103-2117.

Seed, H. B., Ugas, C., & Lysmer, J. (1976). Site-dependent spectra for earthquake-resistant design. *Bulletin of the Seismological Society of America*, 66(1): 221-243.

Sigman, K. (1999). Appendix: A primer on heavy-tailed distributions. *Queueing Systems*, 33(1): 261-275.

Smith, R. L. (2003). Statistics of extremes, with applications in environment, insurance and finance. Extreme values in finance, telecommunications and the environment, 1-78.

Somerville, P. G. (2003). Magnitude scaling of the near fault rupture directivity pulse. *Physics of the earth and planetary interiors*, 137(1): 201-212.

Soong, T. T. (2004). *Fundamentals of probability and statistics for engineers*. West Sussex, UK: John Wiley and Sons

Sornette, A., Sornette, D. (1989). Self-organized criticality and earthquakes. *Europhysics Letters*, 9(3): 197.

Sornette, D., Knopoff, L., Kagan, Y. Y., Vanneste, C. (1996). Rank-ordering statistics of extreme events: Application to the distribution of large earthquakes.*Journal of Geophysical Research: Solid Earth,* 101(B6): 13,883-13,893.

Sornette, D. (2006). Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools (Springer Series in Synergetics). Germany: Springer

Spagnuolo, E., Herrero, A., Cultrera, G. (2012). The effect of directivity in a PSHA framework. *Geophysical Journal International*, 191(2): 616-626.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. Journal of the American Statistical Association, 69(347): 730-737.

© University of Pretoria

Stepp, J. C., Wong, I., Whitney, J., Quittmeyer, R., Abrahamson, N., Toro, G., Youngs, R., Coppersmith, K., Savy, J., Sullivan, T. (2001). Probabilistic seismic hazard analyses for ground motions and fault displacement at Yucca Mountain, Nevada. *Earthquake Spectra,* 17(1): 113-151.

Stewart, J. P., Chiou, S. J., Bray, J. D., Graves, R. W., Somerville, P. G., & Abrahamson, N. A. (2002). Ground motion evaluation procedures for performance-based design. *Soil dynamics and earthquake engineering*, *22*(9): 765-772.

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

Strasser, F. O., Abrahamson, N. A., Bommer, J. J. (2009). Sigma: Issues, insights, and challenges. *Seismological Research Letters*, 80(1): 40-56.

Strasser, F. O., Bommer, J. J. (2009). Review: Strong Ground Motions—Have We Seen the Worst?. *Bulletin of the Seismological Society of America*, 99(5), 2613-2637.

Tinti, S., Mulargia, F. (1985a). Effects of magnitude uncertainties on estimating the parameters in the Gutenberg-Richter frequency-magnitude law. *Bulletin of the Seismological Society of America*, 75(6): 1681-1697.

Tinti, S., Mulargia, F. (1985b). Application of the extreme value approaches to the apparent magnitude distribution of the earthquakes. *Pure and Applied Geophysics*, 123(2): 199-220.

Trifunac M.D., Lee V.F., (1985a). *Preliminary empirical model for scaling Fourier amplitude spectra of strong ground acceleration in terms of earthquake magnitude source to station distance, site intensity and recording site conditions* (Report No. CE 85-03), Deptartment of Civil Engineering , University of Southern California, Los Angeles, California.

Trifunac M.D., Lee V.F., (1985b). *Preliminary empirical model for scaling pseudo relative velocity spectra of strong earthquake acceleration in terms of magnitude, distance, site intensity and recording site conditions* (Report No. CE 85-04), Deptartment of Civil Engineering , University of Southern California, Los Angeles, California.

Trifunac, M. D. (1976a). Preliminary empirical model for scaling Fourier amplitude spectra of strong ground acceleration in terms of earthquake magnitude, source-to-station distance, and recording site conditions. *Bulletin of the Seismological Society of America*, 66(4): 1343-1373.

© University of Pretoria

Trifunac, M. D. (1976b). Preliminary empirical model for scaling Fourier amplitude spectra of strong ground acceleration in terms of earthquake magnitude, source-to-station distance, and recording site conditions. *Bulletin of the Seismological Society of America*, 66(4): 1343-1373.

Trifunac, M. D., & Lee, V. W. (1989). Empirical models for scaling Fourier amplitude spectra of strong ground acceleration in terms of earthquake magnitude source to station distance, site intensity and recording site conditions. *Soil Dynamics and Earthquake Engineering*, 8(3): 110-125.

Trifunac, M. D. (1991) Empirical scaling of Fourier spectrum amplitudes of recorded strong earthquake accelerations in terms of Modified Mercalli Intensity, local soil conditions and depth of sediments. *Soil Dynamics and Earthquake Engineering* 10.1 (1991): 65-72.

Tsang, H. H., Yaghmaei-Sabegh, S., Anbazhagan, P., & Sheikh, M. N. (2011). A checking method for probabilistic seismic-hazard assessment: case studies on three cities. *Natural hazards*, 58(1): 67-84.

Tsaur, D. H., Chang, K. H. (2009). Scattering and focusing of SH waves by a convex circular-arc topography. *Geophysical Journal International*, 177(1): 222-234.

Tsaur, D. H., Chang, K. H. (2008). An analytical approach for the scattering of SH waves by a symmetrical V-shaped canyon: shallow case. *Geophysical Journal International*, 174(1): 255-264.

Utsu, T. (1999). Representation and Analysis of the Earthquake Size Distribution: A Historical Review and Some New Approaches. *pure and applied geophysics*, 155(2-4): 509-535.

Veneziano, D., C. A. Cornell, O'Hara, T. (1984). *Historical method of seismic hazard analysis*. Electric Power Research Institute: Palo Alto, California.

Wald, D. J., Heaton, T. H. (1994). Spatial and temporal distribution of slip for the 1992 Landers, California, earthquake. *Bulletin of the Seismological Society of America*, 84(3): 668-691.

Woessner, J., Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, 95(2): 684-698.

Youngs, R. R., Coppersmith, K. J. (1985). Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates. *Bulletin of the Seismological society of America*, 75(4): 939-964.

© University of Pretoria

Appendix A: The Riemann-Stieltjes integral

This appendix has been added for the convenience of the reader not familiar with integrals of the form

$$\int_a^b g(x)dF(x) \qquad (A.1)$$

This explanation is adapted from Bartle (1964) to be brief, informal and intuitive.

An informal analytical definition of the Riemann-Steltjies Integral: Divide the interval (a,b) into sub-intervals and denote the collection of sub-intervals by P, and denote the endpoints of the intervals by x_i. We construct what we call a Riemann-Stieltjes sum:

$$S(P,g,f) = \sum_{k=2}^{n} g(t_{i-1})\{f(x_i) - f(x_{i-1})\} \qquad (A.2)$$

where t_j is some arbitrary point in the interval (x_{i-1}, x_i). Figure A1 illustrates this idea of *partitioning* graphically.

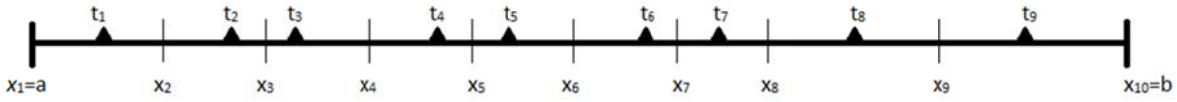

Figure A1: Illustration of partitioning in a Riemann-Stieltjes sum

If we refine P consecutively, and $S(P,g,f)$ converges, the value to which it converges is what we call the Riemann-Stieltjes integral and is denoted by

$$\int_a^b g(x)dF(x)$$

The idea is by no means far removed from the usual Riemann integral; in fact, it is exactly equivalent to rescaling the abscissa (horizontal axis) to $F(x)$ and evaluating the resulting integral. The following characteristic allows one to evaluate the Riemann-Stieltjes integral (at least for *well-behaved* functions):

$$\int_a^b g(x)dF(x) = \int_a^b g(x)\left[\frac{dF(x)}{dx}\right]dx \qquad (A.3)$$

© University of Pretoria

Appendix B: Derivation of the iterative re-substitution method and the inverse Taylor method

A GMPE of the following general form is assumed:

$$\ln(pga) = q_1 + q_2 m + \varphi_r + \varphi_{m,r}(m) \tag{B.1}$$

Now let

$$x = x(\ln(pga), r) \coloneqq \ln(pga) - q_1 - \phi_r \tag{B.2}$$

and

$$q = q_2 \tag{B.3}$$

So we also have

$$x = x(m, r) = qm + \phi_{m,r}(m) \tag{B.4}$$

The idea is to obtain an explicit solution

$$\Omega(x) \coloneqq m(x) \tag{B.5}$$

© University of Pretoria

Development of the iterative re-substitution Method:

In many cases it is justified to assume (see the list of GMPES in Douglas, 2011) that

$$qm \gg \phi_{m,r}(m) \tag{B.6}$$

It follows that

$$x \approx qm \tag{B.7}$$

From equation (B.4) we have

$$m = \frac{x - \phi_{m,r}(m)}{q} \tag{B.8}$$

Then, rearranging and substituting approximation (B.7) in (B.8) we have

$$m \approx \frac{x - \phi_{m,r}\left(\frac{x}{q}\right)}{q} \tag{B.9}$$

which is a better approximation than (B.7) alone, under the assumption that the second derivative of φ is not too large (that is so that the first derivative does not increase too rapidly; a motivation is given in the appendix to this chapter). Substitution of (B.9) into (3.3) (which is the cdf form of the Gutenberg-Richter relation) yields

© University of Pretoria

$$F_x(x) = \frac{1 - e^{-\beta\left(\frac{x-\phi_{m,r}\left(\frac{x}{q}\right)}{q} - m_{min}\right)}}{1 - e^{-\beta(m_{max} - m_{min})}} \tag{B.10}$$

Further improvement may be made by re-substitution of equation (B.8) into itself, and afterward substituting $x \approx \frac{x}{q}$. The result of 3 such re-substitutions results in the expression

$$F(x) = \frac{1 - exp\left[-\beta\left(\frac{x-\phi_{m,r}\left(\frac{x-\phi_{m,r}\left(\frac{x-\phi_{m,r}\left(\frac{x}{q}\right)}{q}\right)}{q}\right)}{q} - m_{\min}\right)\right]}{1 - \exp[-\beta(m_{\max} - m_{\min})]} \tag{B.11}$$

In equation (B.11) we have approximated $\Omega(x) \approx \frac{x-\phi_{m,r}\left(\frac{x-\phi_{m,r}\left(\frac{x}{q}\right)}{q}\right)}{q}$.

In this way we obtain a sequence of approximations which is expected to converge to $\Omega(x)$:

$$\Omega_0^{I.R.}(x) \approx \frac{x}{q}$$

$$\Omega_1^{I.R.}(x) \approx \frac{x - \phi_{m,r}\left(\frac{x}{q}\right)}{q} \tag{B.12}$$

$$\Omega_3^{I.R.}(x) \approx \frac{x - \phi_{m,r}\left(x - \phi_{m,r}\left(\frac{x}{q}\right)\right)}{q}$$

■ ■ ■

Development of the Inverse Taylor Method:

© University of Pretoria

Let

$$\Omega(x) = \boldsymbol{inv}[qm + \phi_{m,r}](x) \tag{B.13}$$

Where $\boldsymbol{inv}[.]$ denotes the map of a function to its invese.The Taylor expansion to an arbitrary order *n*, with $n \geq 2$, is

$$\Omega(x) \approx T_n[\text{inv}(qm + \phi_{m,r})](x) = m_0 + \frac{(x - x_0)}{1!}\text{M}'(x_0) + \sum_{k=2}^{n} \frac{(x - x_0)^k}{k!}\text{M}^{(n)}(x_0) \tag{B.14}$$

where T_n denotes the Taylor series expanded to n terms.

We have that

$$\Omega'(x_0) = \frac{1}{\frac{d}{dm}\left(qm + \phi_{m,r}(m)\right)\Big|_{x=x_0}} = \frac{1}{q + \frac{d}{dm}\phi_{m,r}(m_0)} \tag{B.15}$$

Now one still has to solve $\Omega^{(n)}$ for $n > 1$, which we would like to express explicitly if possible. Itsikov *et al.* (2012) derived the following algorithmic formulae:

Denote the inverse of $\Omega(.)$ by $\psi(.)$ (so in our case $\psi(m) = qm + \phi_{m,r}(m)$). Then

$$\Omega^{(n)} = -\frac{n!}{(\psi')^n}\sum_{j=1}^{n-1}\frac{\Omega^{(j)}}{j!}P_{j,n}, n = 2,3, \ldots \tag{B.16}$$

© University of Pretoria

Where

$$P_{j,k} = \begin{cases} (\psi')^j, & j = k \\ \dfrac{1}{(k-j)\psi'} \displaystyle\sum_{l=1}^{k-j} (lj - k + j + l) \frac{\psi^{(l+1)}}{(l+1)!} P_{j,k-l} & k \geq j+1 \end{cases} \tag{B.17}$$

which renders a solution to equation (B.14).

© University of Pretoria

Appendix C: Analytical motivation for validity of iterative re-substitution method

The following is an analytical motivation (in some cases more intuitive than mathematically rigorous) of the proposal that the approximation $\Omega_1^{I.R.}(x)$ is better than $\Omega_0^{I.R.}(x)$ (see equation (8.8)) when the second derivative of φ is sufficiently small, but may be invalid if the condition is not met. Convergence of subsequent terms of the sequence $\left(\Omega_n^{I.R.}(x)\right)$ by induction is not proven, but the improvement made by the first two terms provides a basis for such an argument.

Assume d is positive and that there exists constant m_{zero} such that $\varphi(m_{zero}) = 0$. The error due to the approximation is

$$m \approx \frac{x - \varphi\left(\frac{x}{d}\right)}{d} \tag{C.1}$$

$$\varepsilon_1 = \left| m - \frac{x - \varphi\left(\frac{x}{d}\right)}{d} \right| = \left| \frac{x - \varphi(m)}{d} - \frac{x - \varphi\left(\frac{x}{d}\right)}{d} \right| = \frac{\left|\varphi\left(\frac{x}{d}\right) - \varphi(m)\right|}{d} \tag{C.2}$$

Equivalently

$$\varepsilon_1 = \frac{\left|\varphi\left(\frac{x}{d}\right) - \varphi\left(\frac{x-\varphi(m)}{d}\right)\right|}{d} \tag{C.3}$$

The error due to $m \approx \frac{x}{d}$ alone is

$$\varepsilon_2 = \left| m - \frac{x}{d} \right| = \left| \frac{x - \varphi(m)}{d} - \frac{x}{d} \right| = \frac{|\varphi(m)|}{d} \tag{C.4}$$

So $\varepsilon_1 < \varepsilon_2$ exactly when $\left|\varphi\left(\frac{x}{d}\right) - \varphi(m)\right| < |\varphi(m)|$.

In short: Assuming $|\varphi''|$ is not very large, we have that

1) For most values of m we have $\varepsilon_1 < \varepsilon_2$.
2) For very small values it might be the case that $\varepsilon_1 > \varepsilon_2$, but by this time we know that ε_1 is negligible.

Figure 1 illustrates the need of the assumption that $|\varphi''|$ should not be very large.

© University of Pretoria

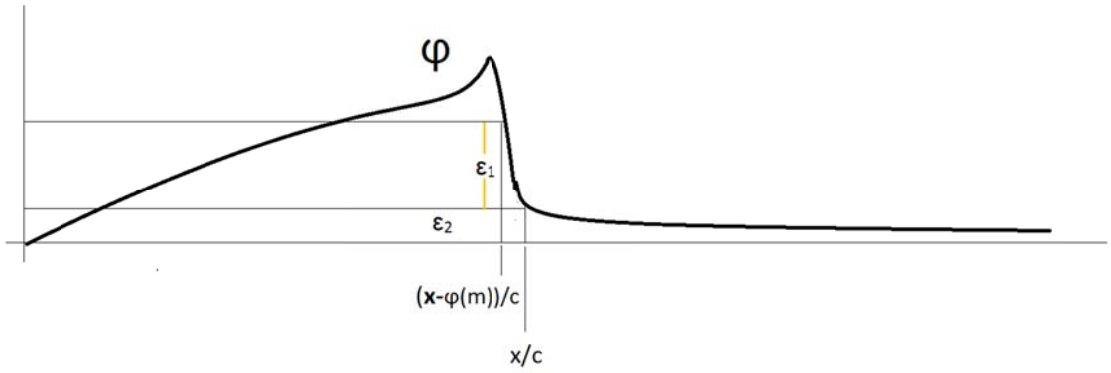

Figure A8.1. Illustrative motivation of the need for $|\varphi''|$ to not be very large

More elaborately: By the mean value theorem there exist constants a and b in the intervals $\left[\min(m,\frac{x}{d}),\max(m,\frac{x}{d})\right]$ and $[\min(m,m_{zero}),\max(m,m_{zero})]$, respectively, such that

$$|\varphi'(a)|\left|\frac{x}{d}-m\right| = \left|\varphi\left(\frac{x}{d}\right)-\varphi(m)\right| \qquad (C.5)$$

$$|\varphi'(b)||m_{zero}-m| = |\varphi(m)| \qquad (C.6)$$

Note, in relation (C.4), that

$$\left|\frac{x}{d}-m\right| = \frac{|\varphi|}{d} \qquad (C.7)$$

by the definition of x.

Now assume that

$$\left|\frac{x}{d}-m\right| = \frac{|\varphi|}{d} \ll |m_{zero}-m| \qquad (C.8)$$

which, because of (C.6), is equivalent to assumption that m is not very close to m_{zero}. Then, to have

$$\varepsilon_1 > \varepsilon_2 \qquad (C.9)$$

(which is the opposite of what we want) we need

$$\frac{\left|\varphi\left(\frac{x}{d}\right)-\varphi\left(\frac{x-\varphi(m)}{d}\right)\right|}{d} > \frac{|\varphi(m)|}{d} \qquad (C.10)$$

which, dividing both sides by *d* and substituting equation (C.6) gives

© University of Pretoria

$$\left|\varphi\left(\frac{x}{d}\right) - \varphi(m)\right| > |\varphi(m)| \qquad (C.11)$$

which leads to

$$|\varphi'(a)|\left|\frac{x}{d} - m\right| > |\varphi'(b)||m_{zero} - m| \qquad (C.12)$$

and, substituting equation (C.6) again, shows that we need

$$|\varphi'(a)|\frac{|\varphi|}{d} > |\varphi'(b)||m_{zero} - m| \qquad (C.13)$$

which requires that $|\varphi'(a)| \gg |\varphi'(b)|$ because of relations (C.8) and (C.13). This, of course, requires that $|\varphi''|$ should be large. Turning the argument around, this shows that if $|\varphi''|$ is not large, then

$$\varepsilon_1 \lesssim \varepsilon_2 \qquad (C.14)$$

When $\frac{|\varphi|}{d} \gtrsim |m_{zero} - m|$, it is reasonable to assume that ε_1 will be negligible, since $dm >> \varphi_m(m)$ implies that $\frac{|\varphi|}{d}$ is very small compared to m_{zero} in a region close to to m_{zero}, given that it does not "blow up" too quickly – this we assure again by requiring that $|\varphi''|$ is not too large.

© University of Pretoria