

**Time-dependent occurrence rates of large earthquakes in the Dead Sea  
fault zone and applications to probabilistic seismic hazard assessments**

**Dissertation**  
**zur Erlangung des akademischen Grades**  
**"doctor rerum naturalium"**  
**(Dr. rer. nat.)**  
**in der Wissenschaftsdisziplin "Geophysik"**

**eingereicht und verteidigt an der**  
**Mathematisch-Naturwissenschaftlichen Fakultät**  
**der Universität Potsdam**

**vorgelegt von**  
**Amir Hossein Hakimhashemi**

**Potsdam, Oktober 2009**



This work is licensed under a Creative Commons License:  
Attribution - Noncommercial - Share Alike 3.0 Germany  
To view a copy of this license visit  
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Published online at the  
Institutional Repository of the University of Potsdam:  
URL <http://opus.kobv.de/ubp/volltexte/2011/5248/>  
URN <urn:nbn:de:kobv:517-opus-52486>  
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-52486>

## Abstract

The seismicity of the Dead Sea fault zone (DSFZ) during the last two millennia is characterized by a number of damaging and partly devastating earthquakes. These events pose a considerable seismic hazard and seismic risk to Syria, Lebanon, Palestine, Jordan, and Israel. The occurrence rates for large earthquakes along the DSFZ indicate temporal changes in the long-term view.

The aim of this thesis is to investigate, whether the occurrence rates of large earthquakes ( $M_w \geq 6$ ) in different parts of the DSFZ are time-dependent and how. The results are applied to probabilistic seismic hazard assessments (PSHA) in the DSFZ and neighboring areas. Four time-dependent statistical models (distributions), including Weibull, Gamma, Lognormal and Brownian Passage Time (BPT), are applied along with the exponential distribution (Poisson process) as the classical time-independent model. In order to check, if the earthquake occurrence rate follows a unimodal or a multimodal form, a nonparametric bootstrap test of multimodality has been done. A modified method of weighted Maximum Likelihood Estimation (MLE) is applied to estimate the parameters of the models. For the multimodal cases, an Expectation Maximization (EM) method is used in addition to the MLE method. The selection of the best model is done by two methods; the Bayesian Information Criterion (BIC) as well as a modified Kolmogorov-Smirnov goodness-of-fit test. Finally, the confidence intervals of the estimated parameters corresponding to the candidate models are calculated, using the bootstrap confidence sets.

In this thesis, earthquakes with  $M_w \geq 6$  along the DSFZ, with an average width of about 40 km and inside  $29.5^\circ \leq \text{latitude} \leq 37^\circ$ , are considered as the dataset. The completeness of this dataset is calculated since 300 A.D. The DSFZ has been divided into three subzones; the southern, the central and the northern subzone respectively. The central and the northern subzones have been investigated but not the southern subzone, because of the lack of sufficient data.

The results for the central part of the DSFZ show that the earthquake occurrence rate does not significantly pursue a multimodal form. There is also no considerable difference between the time-dependent and time-independent models. Since the time-independent model is easier to interpret, the earthquake occurrence rate in this subzone has been estimated under the exponential distribution assumption (Poisson process) and will be considered as time-independent with the value of  $9.72 * 10^{-3}$  events/year.

The northern part of the DSFZ is a special case, where the last earthquake has occurred in 1872 (about 137 years ago). However, the mean recurrence time of  $M_w \geq 6$  events in this area is about 51 years. Moreover, about 96 percent of the observed earthquake inter-event times (the time between two successive earthquakes) in the dataset from this subzone are smaller than 137 years. Therefore, the

next earthquake in this subzone is overdue. The results for this subzone verify that the earthquake occurrence rate is strongly time-dependent, especially shortly after an earthquake occurrence. A bimodal Weibull-Weibull model has been selected as the best fit for this subzone. The earthquake occurrence rate, corresponding to the selected model, is a smooth function of time and reveals two clusters within the time after an earthquake occurrence. The first cluster begins immediately after an earthquake occurrence, lasts about 80 years, and is explicitly time-dependent. The occurrence rate of this cluster is considerably low immediately after an earthquake occurrence, increases strongly during the following ten years and reaches its maximum of about 0.024 events/year, then decreases over the next 70 years to its minimum of about 0.0145 events/year. The second cluster begins 80 years after an earthquake occurrence and lasts until the next earthquake occurs. The earthquake occurrence rate, corresponding to this cluster, increases extremely slowly, so that it can be considered as an almost constant rate about 0.015 events/year. The results are applied to calculate the time-dependent PSHA in the northern part of the DSFZ and neighboring areas.

# Contents

<b>1</b>	<b>Introduction and Problem Definition</b>	<b>1</b>
1.1	Introduction . . . . .	1
1.2	Problem definition . . . . .	2
1.2.1	Aim of the thesis . . . . .	2
1.2.2	Application to seismic hazard assessment . . . . .	3
1.2.3	Study area . . . . .	5
<b>2</b>	<b>Time-dependent modeling of earthquake occurrence rate, an overview</b>	<b>7</b>
2.1	Physics-based model . . . . .	8
2.2	Statistics-based models . . . . .	8
2.2.1	Short-term time-dependent models, aftershock sequences . . . . .	8
2.2.2	Long-term time-dependent models . . . . .	15
2.3	Conclusions . . . . .	30
<b>3</b>	<b>An adapted approach to time-dependent earthquake occurrence rate modeling</b>	<b>31</b>
3.1	Empirical modeling . . . . .	31
3.2	Time-to-failure modeling . . . . .	32
3.3	Renewal Process . . . . .	32
3.3.1	Earthquake recurrence as a renewal process . . . . .	33
3.4	Statistical distribution of earthquake inter-event times . . . . .	33
3.4.1	Weibull distribution . . . . .	33
3.4.2	Gamma distribution . . . . .	35
3.4.3	Lognormal distribution . . . . .	37
3.4.4	Inverse Gaussian distribution . . . . .	39
3.4.5	Summary . . . . .	41
3.5	Behavior of distributions; multimodality against unimodality . . . . .	41
3.5.1	Multimodality test of a population . . . . .	42
3.6	Modified Weighted Maximum Likelihood Estimation method . . . . .	44
3.6.1	The general approach . . . . .	44
3.6.2	Case of observed censored times . . . . .	45

3.6.3	Weighted MLE with censored times . . . . .	45
3.6.4	Method of solving the MLE problem . . . . .	47
3.7	Selection of the best model . . . . .	51
3.7.1	Methods of information criterion . . . . .	51
3.7.2	Methods of goodness-of-fit test . . . . .	51
3.8	Confidence intervals of estimated parameters . . . . .	52
3.9	Earthquake inter-event time modeling approach . . . . .	53
<b>4</b>	<b>Seismicity in the Dead Sea Fault Zone</b>	<b>54</b>
4.1	Dead Sea Fault Zone and seismicity dataset . . . . .	54
4.1.1	Declustering fore- and aftershocks . . . . .	56
4.1.2	A new method of earthquake catalog completeness check .	58
4.1.3	Final dataset . . . . .	60
4.1.4	Some characteristics of the dataset . . . . .	62
4.2	Temporal cluster behavior in the DSFZ . . . . .	64
<b>5</b>	<b>Analysis of earthquake occurrence rates in the DSFZ</b>	<b>68</b>
5.1	Earthquake occurrence rate in the central part of the DSFZ . . . . .	68
5.2	Earthquake occurrence rate in the northern part of the DSFZ . . . . .	71
5.2.1	The northern part of the DSFZ - a special case . . . . .	71
5.2.2	Multimodality of inter-event times . . . . .	72
5.2.3	Unimodal distributions . . . . .	72
5.2.4	Bimodal (mixed) distributions . . . . .	73
5.2.5	Selection of the best model . . . . .	75
5.2.6	Application of mixed Weibull-Weibull and Lognormal distributions in probabilistic seismic hazard assessments . . . . .	82
<b>6</b>	<b>Summary and conclusions</b>	<b>84</b>
6.1	Summary . . . . .	84
6.2	Conclusions . . . . .	86
	<b>References</b>	<b>90</b>
	<b>Appendix A</b>	<b>98</b>
	<b>Appendix B</b>	<b>104</b>
	<b>Acknowledgments</b>	<b>109</b>
	<b>Curriculum Vitae</b>	<b>110</b>

# List of Figures

1.1	An example for time-dependent and time-independent hazard functions. The red curve shows a time-dependent hazard function of a Brownian Passage Time (BPT) distribution. The blue curve shows a time-independent hazard function of an exponential distribution (Poisson process). . . . .	4
1.2	Location of the DSFZ between the Arabian Plate and the Sinai Block, stretching from the Red Sea in the south to the East Anatolian Fault the north (data from Heidbach & Ben-Avraham 2007; Meghraoui et al. 2003). Red curves represent fault segments of the DSFZ. Arrows show the relative motion of the plates. . . . .	6
3.1	An example of a time-to-failure process, sample evolution of a renewal process $X$ with inter-event times $\tau_1, \tau_2, \dots$ and jump times $J_1, J_2, \dots$ . . . . .	32
3.2	Different Weibull distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function. . . . .	35
3.3	Different Gamma distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function. . . . .	36
3.4	Different Lognormal distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function. . . . .	38
3.5	Different BPT distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function. . . . .	40
3.6	An example of multimodal and unimodal pdfs. The blue curve corresponds to a random variable $T$ , unimodally distributed as $T \sim \text{Weibull}(\beta = 1.8, \eta = 50)$ . The red curve is extracted from a random variable $X$ , multimodally distributed as $X \sim 0.5\text{Weibull}(\beta = 1.1, \eta = 20) + 0.5\text{Weibull}(\beta = 4.0, \eta = 100)$ . . . . .	42
3.7	Different weight functions (equation 3.40) with respect to different shape and scale parameters (location parameter $K = 1$ ); a. $\alpha = 0.5$ , b. $\alpha = 1$ , c. $\alpha = 5$ . . . . .	46



4.1	The Eastern Mediterranean area including the Dead Sea fault zone (modified after Garfunkel 1981). Arrows show relative motions across faults. AF = Arava Fault, AV = Arava Valley, CF = Carmel Fault, DST = Dead Sea Transform, EAFZ = East Anatolian Fault, ESM = Eratosthenes Sea Mount, GAF = Gulf of Aqaba Fault, GF = Gharb Fault, GS = Gulf of Sinai, JV = Jordan Valley, YF = Yamuna Fault. . . . .	55
4.2	Seismicity in the DSFZ, including the subdivision into a southern, a central and a northern zone. The sources of the seismicity data have been described in section 4.1. Red curves show the border of each zone. . . . .	56
4.3	Surface rupture traces of the $M_w \geq 6.0$ earthquakes in the DSFZ since 300 AD. Red lines show the fault segments (data from Heibach & Ben-Avraham 2007; Meghraoui et al. 2003). . . . .	57
4.4	Latitude of the earthquake epicenters vs. earthquake occurrence times. Top: For the entire DSFZ. Bottom: Blue curve for the northern part of the DSFZ and red curve for the central part of the DSFZ. . . . .	66
4.5	Benioff curve for the northern part of the DSFZ. The cumulative seismic moment release shows "active" and "inactive" periods. . .	67
5.1	Results in terms of the achieved significance level (ASL) of the multimodality test with 1000 bootstrap resamplings for the central part of the DSFZ. The inter-event times reveal no specific bimodality. . . . .	68
5.2	The weight function with parameters; $\alpha = 1$ , $p = 6$ and $K = 1$ . . .	69
5.3	Comparison between different distribution functions and the empirical distribution function of earthquake inter-event times for the central part of the DSFZ. . . . .	70
5.4	A schematic clustering behavior of the inter-event times in the northern part of the DSFZ. The crosses present the empirical distribution function, red crosses show the first cluster, and blue crosses demonstrate the second cluster. The red vertical line shows the censored time and its situation comparing with inter-event times. .	72
5.5	The results of the achieved significance level (ASL) of the multimodality test with 1000 bootstrap resamplings for the northern part of the DSFZ. The test does not significantly show, whether unimodality or bimodality is a better model. . . . .	73
5.6	Comparison between different unimodal distribution functions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ. . . . .	74

5.7	Comparison between different bimodal distribution functions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ. WW = mixed Weibull-Weibull, BW = mixed BPT-Weibull, LW = mixed Lognormal-Weibull. . . . .	76
5.8	Comparison between distribution functions of mixed Weibull-Weibull and unimodal Lognormal distributions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ. Results for longer inter-event times. . . . .	77
5.9	The trends of hazard functions of mixed Weibull-Weibull and unimodal Lognormal distributions after the last earthquake in 1872. . . . .	78
5.10	Various hazard functions corresponding to mixed Weibull-Weibull distributions, calculated based on different weight function parameters, within 200 years after the last earthquake in 1872. The red curve is based on the weight function parameters considered in this thesis. The green curve is based on no weights. Changing the weight function parameters does not affect the general shape of the hazard function. . . . .	79
5.11	The trend of the hazard function of mixed Weibull-Weibull distribution after the last earthquake in 1872. Top: The first 80 years between 1872 and 1952. Bottom: Between 1952 and 2152. . . . .	80
5.12	Conditional (red) and unconditional (blue) exceedance probabilities for the northern part of the DSFZ. . . . .	81
B.1	Top: Hazard maps for the northern part of the DSFZ in the year 1873 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). The red polygon shows the boundary of the source zone. For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884. . . . .	104

B.2	Top: Hazard maps for the northern part of the DSFZ in the year 1875 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884. . . . .	105
B.3	Top: Hazard maps for the northern part of the DSFZ in the year 1883 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884. . . . .	106
B.4	Top: Hazard maps for the northern part of the DSFZ in the year 2009 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 2009 - 2100. . . . .	107
B.5	Top: Hazard maps for the northern part of the DSFZ in the year 2100 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 2009 - 2100. . . . .	108

# List of Tables

4.1	Foreshocks and corresponding main shocks along the DSFZ, $M_w \geq 6$ . . . . .	58
4.2	Completeness for different magnitude classes in the DSFZ catalog.	60
4.3	Earthquake dataset of the DSFZ with $M_w \geq 6$ . . . . .	61
4.4	Earthquake dataset of the northern part of the DSFZ with $M_w \geq 6$ .	63
4.5	Earthquake dataset of the central part of the DSFZ with $M_w \geq 6$ . .	64
4.6	Mean value and standard deviation of inter-event times in the DSFZ.	65
5.1	Estimated unimodal distributions for the central part of the DSFZ.	70
5.2	Estimated unimodal distributions for the northern part of the DSFZ.	73
5.3	Estimated bimodal distributions for the northern part of the DSFZ.	75
5.4	The sum of squared errors for the candidate distributions. . . . .	76

# Chapter 1

## Introduction and Problem Definition

### 1.1 Introduction

Large earthquakes belong to the most hazardous natural disasters. Observing only the most recent ten years, four earthquakes with severe consequences have occurred, the 1999 Izmit earthquake in Turkey with the official death toll of 17,127 killed<sup>1</sup> (Barka 1999, Tang 2000, and Marza 2004), the 2003 Bam earthquake in Iran with 26,271 deaths<sup>2</sup>, the 2004 Sumatra-Andaman earthquake and the associated tsunami in the Indian Ocean with a death toll of more than 230,000 people in 11 countries (Lay et al. 2005) and, finally, the 2008 Sichuan earthquake in China that killed at least 68,000 people (Jacobs et al. 2009). Altogether, more than 340,000 people have lost their lives in only these four large earthquakes or their immediate after-effects, as tsunamis, landslides, fires, etc. These figures categorize earthquakes as one of the deadliest natural disasters.

As a matter of fact, any knowledge about the hazard due to earthquakes; i.e. the seismic hazard, can help to reduce the socio-economical losses and to save lives. One of the essential parameters in seismic hazard analyses is the occurrence rate of earthquakes. The occurrence rate describes the number of earthquakes in a time unit. In classic methods of seismic hazard assessments, this rate is considered as constant and extracted from a Poisson process. However, the classic methods cannot properly calculate the rate, when it varies with time. In such cases, time-dependent rate analyses are required.

Recently, various studies (e.g. Ellsworth et al. 1999; Matthews et al. 2002; Parsons 2004; Gonzalez et al. 2006) have shown that the occurrence rate of earthquakes, especially for large events, changes in certain areas with time. In contrast,

---

<sup>1</sup>Some sources suggest the actual figure may have been closer to 40,000 dead (Marza 2004).

<sup>2</sup>Death toll data from BBC News (British Broadcasting Corporation), 2004-03-29, retrieved 2008-10-15. According to Montazeri et al. (2005), over 40,000 deaths.

some studies (see Musson et al. 2005) declare that time-dependency is of possible relevance in short time periods, but in general, for engineering purposes, treating seismicity as time-invariant has usually been considered a useful simplification. Therefore, it is an important issue to test whether the earthquake occurrence rate in a given area is time-dependent or not.

The Dead Sea Fault Zone (DSFZ) is an area, where the occurrence rate for large earthquakes shows an indication of temporal change in the long-term. In this thesis, the earthquake occurrence rate in the DSFZ will be analyzed to discover if it changes significantly with time and how. For that purpose, time-dependent hazard (rate) models will be developed and discussed. The combinations of these time-dependent models with ground motion prediction equations, leading to a seismic hazard assessment will be described later in this chapter.

Similar models have been commonly used in various branches of application of statistical analysis to real-life problems, like: efficiency of treatments in medical sciences, strength of components in manufacturing, financial planning, insurance claims, maintenance of facilities, wind velocity and sunshine intensity in renewable energy, rock blasting in mining, spare part planning for commercial equipment downtime, and many other problems (Murthy 2004). This broad area of usage shows the applicability, effectiveness and flexibility of these models.

In the first chapter of this thesis, the main subject will be declared. The time-dependent methods and models will be reviewed in chapter 2. In chapter 3, a new time-dependent approach will be introduced and discussed in order to model the earthquake inter-event times. The dataset, extracted from DSFZ catalogs, will be presented in chapter 4. The results of application of the approach to the dataset will be illustrated in chapter 5. Finally, a summary of the thesis as well as the outcomes will be summarized and discussed further in chapter 6.

## **1.2 Problem definition**

### **1.2.1 Aim of the thesis**

The aim of this thesis is to analyze the temporal changes of occurrence rate,  $\lambda$ , of large earthquakes in a fault zone or a fault segment.

$\lambda$  is in statistical terminology known as the *hazard* (or *rate*) function. It should be mentioned that the term "hazard" in statistics is different from the term "hazard" in seismic hazard analysis. In statistics, the term "hazard" demonstrates the number of events (here earthquakes above a certain magnitude threshold) in a time unit at a generic time, which is, in fact, the same as the term "occurrence rate" in seismic hazard analysis. The term "hazard" in seismic hazard analysis is always related to the ground motion of a specific site. In this thesis, the term "hazard function"

refers to its meaning in statistical sense. The term "seismic hazard assessment" is used in the sense of its application to probabilistic seismic hazard assessments.

The mathematical interpretation of the hazard function is as:

$$\begin{aligned}
\lambda(t) &= \lim_{\Delta t \rightarrow 0} Pr\{t \leq T \leq t + \Delta t \mid T \geq t\} \\
&= \frac{\lim_{\Delta t \rightarrow 0} Pr\{t \leq T \leq t + \Delta t, T \geq t\}}{Pr\{T \geq t\}} \\
&= \frac{\lim_{\Delta t \rightarrow 0} Pr\{t \leq T \leq t + \Delta t\}}{Pr\{T \geq t\}} = \frac{f(t)}{S(t)} \tag{1.1}
\end{aligned}$$

where  $\lambda(t)$  is the hazard function at a generic time  $t$ ,  $f(t)$  is the *probability density function* (pdf) corresponding to the distribution of the inter-event times (i.e. the time between two consecutive events of interest), and  $S(t)$  is the *survivor* (or *reliability*) function defined as:

$$S(t) = 1 - F(t) \quad \text{with} \quad F(t) = \int_0^t f(x) dx \tag{1.2}$$

where  $F(t)$  is known as the *distribution* or *cumulative distribution function* (cdf). In general, there are two types of hazard functions, namely time-dependent and time-independent. Time-dependent hazard functions vary with time, like hazard functions corresponding to Weibull or Gamma distributions. Time-independent hazard functions do not change with time, like the hazard function of the exponential distribution (Poisson process), which is a constant.

## 1.2.2 Application to seismic hazard assessment

### Classic method, time-independent

A *Probabilistic Seismic Hazard Assessment* (PSHA) for a specific site consists of determining the frequency with which an earthquake ground shaking (e.g. peak or spectral ground acceleration) takes on a defined range of values (e.g.  $> 0.5$  g) with a fixed probability during a given time (e.g. 10% within 50 years). The seismic hazard at a site for earthquake ground shaking  $C$  is defined as the frequency  $\Lambda$  with which a specific value  $c$  is exceeded during time  $t$ . This total frequency  $\Lambda$  is made up of contributions from a number of independent sources  $j$ , where the frequency of exceedance of a specified value  $c$  from each source is calculated as:

$$\Lambda_j(c) = \lambda_j \int_l \int_m S_j(c \mid m \text{ at } l) f_{L|M}(l \mid m) f_M(m) dm dl \tag{1.3}$$

where  $\Lambda_j(c)$  is the frequency with which  $c$  is exceeded at the site of interest from earthquakes at source  $j$ ,  $\lambda_j$  is the occurrence rate of earthquakes of interest at

source  $j$ ,  $m$  is magnitude,  $S_j(c | m \text{ at } l)$  is the probability that  $c$  is exceeded at the site of interest, conditional on an earthquake at source  $j$ , with magnitude  $m$  at distance  $l$ ,  $f_{L|M}(l | m)$  is the probability that an earthquake of given magnitude  $m$  occurs at location  $l$ , and finally  $f_M(m)$  is the probability of occurrence of an earthquake of magnitude  $m$  for source  $j$  (Cornell, 1968; Mc Guire, 2004). Then, the total seismic hazard at the site of interest can be calculated as a summation of hazards on all sources  $j$  as:

$$\Lambda(c) = \sum_j \lambda_j \int_l \int_m S_j(c | m \text{ at } l) f_{L|M}(l | m) f_M(m) dm dl . \quad (1.4)$$

According to the classic methods of PSHA, the parameter  $\lambda_j$  in equations 1.3 and 1.4 (i.e. the hazard function of the earthquake inter-event times) is generally extracted from a Poisson process and is time-independent.

### Toward a time-dependent approach

Time-dependent hazard functions change with time. Suppose that  $t$  is a generic time after an earthquake occurrence, then  $\lambda(t)$  is a non-constant function of  $t$  (see figure 1.1). Considering a time-dependent hazard function, equation 1.4 changes

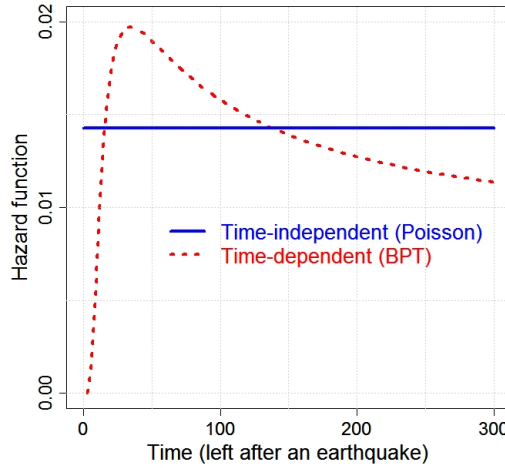


Figure 1.1: An example for time-dependent and time-independent hazard functions. The red curve shows a time-dependent hazard function of a Brownian Passage Time (BPT) distribution. The blue curve shows a time-independent hazard function of an exponential distribution (Poisson process).

to

$$\Lambda(t, c) = \sum_j \lambda_j(t) \int_l \int_m S_j(c | m \text{ at } l) f_{L|M}(l | m) f_M(m) dm dl . \quad (1.5)$$



As mentioned above, the PSHA is generally estimated during some fixed time span in the future. In the case of time-independent PSHA, the seismic hazard in a given time unit (e.g. annual seismic hazard) remains constant during the time (equation 1.4). In contrast, in time-dependent PSHA cases (equation 1.5), the hazard function  $\Lambda(t, c)$  changes with time. In order to calculate time-dependent hazard during a time span  $[t_1, t_2]$ , the  $\lambda_j(t)$  in equation 1.5 can be replaced by:

$$\int_{t_1}^{t_2} \lambda_j(t) dt \quad , \quad (1.6)$$

or for the seismic hazard in time unit by:

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \lambda_j(t) dt \quad . \quad (1.7)$$

According to equations 1.6 and 1.7, considering larger time spans (i.e.  $t_2 - t_1$ ) results in losing details of seismic hazard variations with time. Therefore, the time span has to be considered carefully and as short as possible, especially in those cases, in which the seismic hazard varies vividly with time.

### 1.2.3 Study area

In this thesis, the Dead Sea Fault Zone (DSFZ) in the Middle East has been chosen as the study area. The seismicity of the DSFZ during the last two millennia is characterised by a relatively large number of damaging earthquakes. Earthquakes along the DSFZ pose a considerable seismic hazard to Syria, Lebanon, Palestine, Jordan, and Israel (figure 1.2).

The DSFZ is an active fault zone, forming the boundary between the Arabian plate and the Sinai block (Mahmoud et al. 2005), as a part of the African plate, stretching from the Red Sea in the south to the East Anatolian fault in the north. Based on geological evidence and kinematic considerations, it is generally accepted that the Dead Sea Fault is a continental transform with a total of 105 km accumulated left-lateral displacement (e.g. Garfunkel et al. 1981; Joffe and Garfunkel 1987). The list of historical large earthquakes in the DSFZ provides a very suitable platform in order to apply time-dependent PSHA. Since the seismicity in the DSFZ shows large variations in the past 2000 years, time-dependency of the seismicity rate will be tested for the investigated area. This item will be discussed in detail in chapter 5.

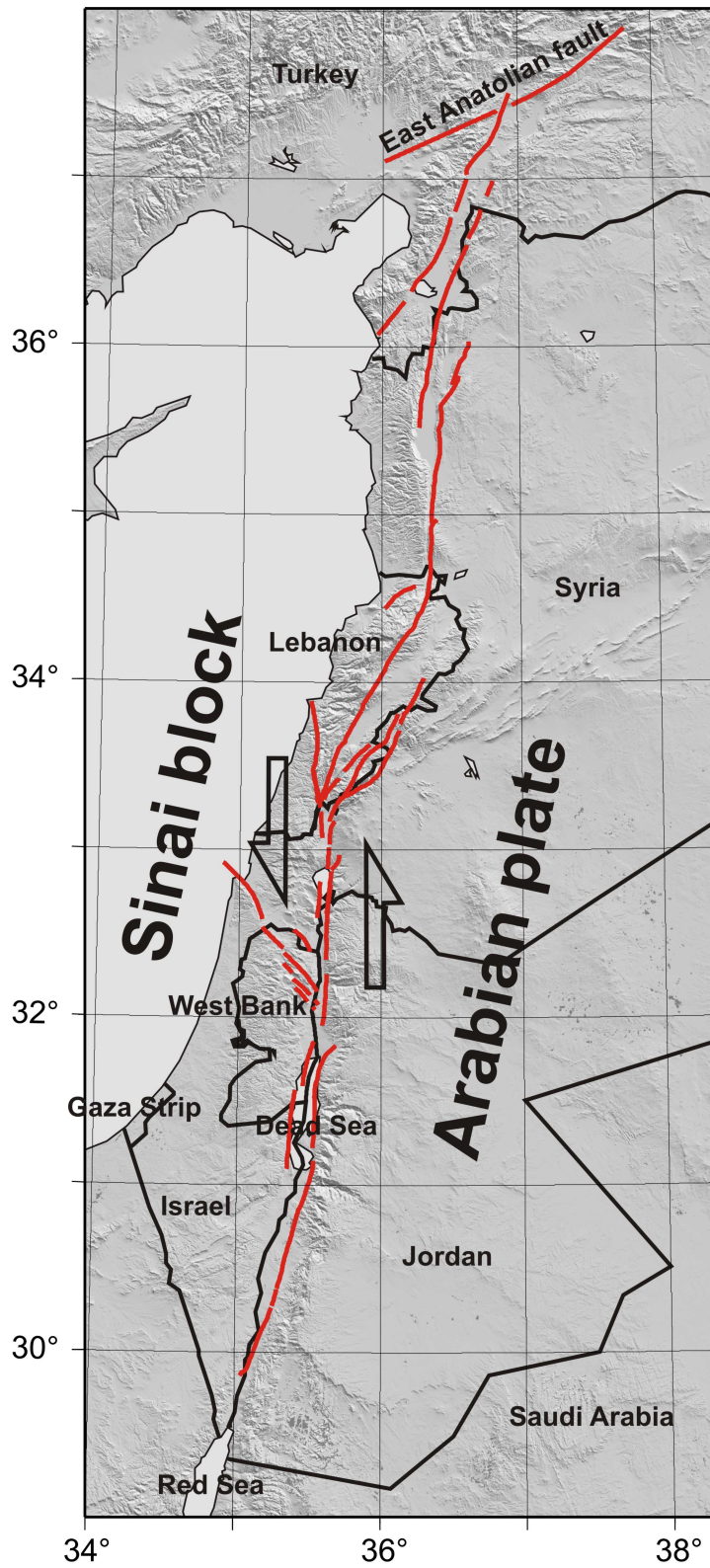


Figure 1.2: Location of the DSFZ between the Arabian Plate and the Sinai Block, stretching from the Red Sea in the south to the East Anatolian Fault the north (data from Heidbach & Ben-Avraham 2007; Meghraoui et al. 2003). Red curves represent fault segments of the DSFZ. Arrows show the relative motion of the plates.

## **Chapter 2**

# **Time-dependent modeling of earthquake occurrence rate, an overview**

Time-dependent modeling of earthquake occurrence rates has recently been developed in many studies (e.g. Ellsworth et al. 1999; Matthews et al. 2002; Parsons 2004; Gonzalez et al. 2006). These approaches have in common that they consider time-dependent earthquake occurrence rates. An overview of these approaches will be given in this chapter.

The time-dependent models are basically divided into two general groups, namely physical and statistical approaches. Physical approaches explain the earthquake rates using the theoretical and experimental aspects of the earthquake processes, considering generally simplified conditions. Although these methods have been more theoretically developed, they deem in practice many simplifications regarding the real conditions, which may cause unexpected and incalculable errors. Statistical approaches, conversely, describe earthquake occurrence rates using empirical earthquake data. These data include earthquake inter-event times, censoring times (the time since the last event), earthquake locations (in the case of spatio-temporal modeling) and earthquake sizes. In general, the earthquake inter-event and censoring times provide much information about earthquake occurrence rates, to be included in statistical methods. In the following an overview of selected time-dependent models will be given.

## 2.1 Physics-based model

### Rate-and-state model

The so-called rate-and-state model (Dieterich 1994, 1995) is commonly applied to model the earthquake recurrence time as a physics-based and forward model. This method is applied to obtain the rate of earthquake activity resulting from the stressing history. Its concept is implemented for faults with rate-and-state-dependent constitutive properties which are derived from laboratory experiments. The goal of this modeling is to find the time at which each source (or area) initiates an earthquake in dependence of the stressing history (Dieterich 1994).

The basic parameters of the rate-and-state models are results of laboratory experiments and field observations. Besides the values and uncertainties of these input parameters, the initial stress state is of key importance, but rarely known. Therefore, for the long-term earthquake nucleation (background seismicity or main shock rate), the determination of the stress level at each source is still a major problem.

The general formulation of this method is specified, giving the time  $t$  at which a particular source nucleates an earthquake from initial conditions and stressing history:

$$t = F[C, \tau(t)] \quad (2.1)$$

where  $C$  represents initial stress conditions and  $\tau(t)$  is some general stressing history.  $C$  can be expressed as the initial stress relative to the stress threshold (Dieterich 1994).

Here, the rate-and-state model will not be discussed further, since the focus in this thesis is on the statistics-based methods. However, it will be included in the discussion of mixed statistics-physics models later in this chapter.

## 2.2 Statistics-based models

### 2.2.1 Short-term time-dependent models, aftershock sequences

The idea of time-dependent earthquake modeling is traditionally related to the modeling of aftershocks, which is performed on short time scales.

#### Omori's law

Omori's law (Omori 1894) is probably the first time-dependent model for describing earthquake rates, where the aftershocks rate triggered by a main shock decays

with the rate

$$\lambda(t) = \frac{K}{t + c} \quad (2.2)$$

where  $t$  is the elapsed time since the main shock occurrence and  $K$  and  $c$  are model parameters.

Utsu (1957, 1961) improved the Omori's law and proposed the power-law form

$$\lambda(t) = \frac{K}{(t + c)^p} \quad (2.3)$$

as the *modified Omori's law*, where  $p$  is the power-law parameter of the relation. Reasenber and Jones (1989) added a magnitude term to the modified Omori's law using the *Gutenberg-Richter* (G-R) law (Gutenberg and Richter 1944),

$$\lambda(t, M) = \frac{10^{a+b(M_m - M)}}{(t + c)^p} \quad (2.4)$$

where  $b$  and  $a$  are the slope and the intercept of the G-R law and  $M_m$  is the magnitude of the main shock. In equation 2.4,  $\lambda(t, M)$  is the aftershock rate at time  $t$  elapsed after the main shock with magnitude  $M$ .

### ETAS model

The Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988) is currently the most applied statistics-based method to model the aftershock rate. This model is based on the typical afterschock decay rate  $\nu(t)$  represented by the modified Omori's law (Utsu 1957, 1961)

$$\nu(t) = K(t + c)^{-p} \quad (2.5)$$

initiated by the main shock at origin time  $t = 0$ .

Typically, the modified Omori's law holds for relatively long periods of the order of some tens of years or more, depending on the background seismicity rate in the neighboring area (Utsu et al., 1995; Ogata and Shimazaki, 1984).

Ogata (1983) proposed a maximum likelihood method to estimate the parameters of the modified Omori's law as:

$$\ln L(\theta) = \sum_{i=1}^N \ln \nu(t_i) - \int_S^T \nu(t) dt, \quad \theta = (K, c, p), \quad (2.6)$$

where  $\ln L(\theta)$  is the corresponding log-likelihood function.

Aftershock activity, however, is not always best predicted by the single modified Omori's law, especially when it includes the secondary aftershock activities

triggered by the large aftershocks, as demonstrated by Guo and Ogata (1997) and Ogata et al. (2003a). Indeed, other complex features of aftershocks can be considered, such as interactively triggered aftershocks, including those among off-fault regions, as discussed by Felzer et al. (2002). Therefore, one can assume that every aftershock can trigger further aftershocks or remote events, and that the occurrence rate at time  $t$  is given by a (weighted) superposition of the modified Omori's relations shifted in time

$$\lambda_\theta(t) = \mu + \sum_{\{j:t_j < t\}} e^{\alpha(M_j - M_c)} \nu(t - t_j) \quad (2.7)$$

where  $\mu$  (shocks/day) represents the rate of the background seismicity, and the summation is taken over every  $j$ -th aftershock occurred before time  $t$  (days) (Ogata and Zhuang, 2006). The weighted size of aftershocks triggered by the  $j$ -th aftershock is made as the exponential function of its magnitude  $M_j$  in accordance with the study by Utsu (1970), where  $M_c$  represents the cut-off magnitude of the fitted data. The coefficient  $\alpha$  measures the efficiency of a shock in generating aftershock activity, relative to its magnitude.

Considering point-process models for the data of occurrence times and locations of earthquakes, where magnitudes are equal to or larger than a certain threshold (cut-off) magnitude  $M_c$ , the occurrence rate  $\lambda(t, x, y|H_t)$  of a space-time point-process is mathematically defined in terms of the occurrence probability of an event at time  $t$  and the location  $(x, y)$  conditional on the past history of occurrences, such that

$$P\{\text{an event in } [t, t + dt] \times [x, x + dx] \times [y, y + dy] | H_t\} = \lambda(t, x, y|H_t) dt dx dy + o(dt dx dy) \quad (2.8)$$

where  $H_t = \{(t_i, x_i, y_i, M_i, S_i); t_i < t\}$  is the history of occurrence times  $\{t_i\}$  up to time  $t$ , with corresponding epicenters  $\{(x_i, y_i)\}$ , magnitudes  $\{M_i\}$ ,  $o(dt dx dy)$  is the function of errors, and  $2 \times 2$ -matrices  $\{S_i\}$  for anisotropic clusters in equations 2.10 - 2.12 (Ogata and Zhuang, 2006). As in the typical space-time extensions of the ETAS model, Ogata (1998) considered the separable form in time and space clustering effects, and compared the following  $g$  functions in equations 2.10 - 2.12 in model

$$\lambda_\theta(t, x, y) = \mu(x, y) + \sum_{\{j:t_j < t\}} \nu(t - t_j) \times g(x - x_j, y - y_j; M_j - M_c), \quad (2.9)$$

where

$$g(x, y; M) = \exp \left[ -\frac{(x, y) S_j(x, y)^t}{2d e^{\alpha(M)}} \right] \quad (2.10)$$

or

$$g(x, y; M) = \frac{e^{\alpha(M)}}{[(x, y)S_j(x, y)^t + d]^q} \quad (2.11)$$

or

$$g(x, y; M) = \left[ \frac{(x, y)S_j(x, y)^t}{e^{\alpha(M)}} + d \right]^{-q} \quad (2.12)$$

where  $S_j$  is an adimensional  $2 \times 2$  positive definite symmetric matrix, and  $(x, y)^t$  indicates the vector transpose. The square brackets have dimension of *degree*<sup>2</sup> where "degree" corresponds to the global distance in latitude. The quadratic form within the brackets indicates that the aftershocks are spatially distributed with ellipsoidal contours. The logarithm of the aftershock area is highly correlated with the main shock's magnitude, leading to the *Utsu-Seki* law, the extension of which is represented by the denominator  $e^{\alpha(M_i - M_c)}$ . The inverse power decay of the aftershock distribution with distance takes the remote triggering phenomenon into account (Ogata and Zhuang, 2006).

Given the dataset of origin times and space coordinates of earthquakes together with their magnitudes and matrices  $\{(t_i, x_i, y_i, M_i, S_i); M_i \geq M_c, i = 1, \dots, n\}$  during a period  $[0, T]$  and in region A, the log-likelihood function of the parameter  $\theta = (\mu, K, c, \alpha, p, d, q)$  characterizing the spacetime point-process model can be calculated by

$$\ln L(\theta) = \sum_{i=1}^N \ln \lambda_{\theta}(t_i, x_i, y_i) - \int_S^T \int \int_A \lambda_{\theta}(t, x, y) dt dx dy. \quad (2.13)$$

Daley and Vere-Jones (2002) calculated the derivatives of this formula and Ogata (1998) proposed the numerical calculation of it. Helmstetter and Sornette (2003) discuss the predictability in the ETAS model.

ETAS models are widely used to model aftershock sequences. The parameter  $\mu$ , the background earthquake rate or the long-term earthquake rate, is not discussed in this model. However, the model is used in some studies for space-time modeling of main shocks too (see Lombardi & Marzocchi, 2007). According to the structure of the model and its parameters, a complete dataset of aftershocks triggered by a main shock is required in order to estimate the model parameters. These data are not always easily available, especially for the historical earthquake catalogs.

### **Mixed ETAS-rate-and-state model**

Recently, several studies have been developed taking into account the combination of the ETAS model and physical methods; i.e. Coulomb failure stress (see page

25) or rate-and-state model. Console et al. (2006) merged ETAS and rate-and-state models, under the assumption that the temporal behavior of the seismicity, triggered by a Coulomb stress change in an infinite population of faults, is described by the rate-and-state model (Ruina 1983; Dieterich 1986,1992,1994). The background seismicity in this model is considered time-independent (as in the original form of ETAS, Ogata 1988).

Llenos et al. (2009) apply both inverse and forward modeling and develop a combined ETAS-rate-and-state model to explain the variations in seismicity rate triggered by transient aseismic processes such as fluid flow, fault creep or magma intrusion. Two approaches are applied for estimating the time-dependency of the underlying driving mechanisms, namely the stochastic ETAS model and a physical approach based on the rate-and-state model of fault friction. They first identify the parameters related to one another in the two models and then examine their dependence on stressing rate. A particular conflict arises because the rate-and-state model predicts that aftershock productivity scales with stressing rate while the ETAS model assumes that it is time-independent. Llenos et al. (2009) estimate the triggering parameters for 4 earthquake swarms contemporaneous with geodetically observed deformation transients in various tectonic environments. They find that stressing rate transients increase the background seismicity rate without affecting aftershock productivity. Finally, they specify a combined model for seismicity rate variations that will allow future studies to invert seismicity catalogs for variations in aseismic stressing rates.

### EEPAS model

Rhoades and Evison (2004, 2005, 2006) and Rhoades (2007) propose the EEPAS, Every Earthquake a Precursor According to Scale, which is a method of long-range forecasting that uses the previous minor earthquakes in a catalog to forecast the major ones. The EEPAS model has a rate density  $\lambda(t, m, x, y)$  of earthquake occurrence which is defined for any time  $t$ , magnitude  $m$ , and location  $(x, y)$ , where  $m$  exceeds a threshold magnitude  $m_c$  and  $(x, y)$  is a point in a region of surveillance  $R$ . Each earthquake  $(t_i, m_i, x_i, y_i)$  contributes a transient increment  $\lambda_i(t, m, x, y)$  to the future rate density in its vicinity, given by

$$\lambda_i(t, m, x, y) = w_i f_{1i}(t) g_{1i}(m) h_{1i}(x, y) \quad (2.14)$$

where  $w_i$  is a weighting factor that may depend on other earthquakes in the vicinity, and  $f_{1i}$ ,  $g_{1i}$  and  $h_{1i}$  are densities of the probability distributions for time, magnitude, and location, respectively. The magnitude density  $g_{1i}$  is assumed to take the form

$$g_{1i}(m) = \frac{1}{\sigma_M \sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{m - a_M - b_M m_i}{\sigma_M} \right)^2 \right] \quad (2.15)$$



where  $a_M$ ,  $b_M$ , and  $\sigma_M$  are parameters. The time density  $f_{1i}$  is assumed as:

$$f_{1i}(t) = \frac{H(t - t_i)}{(t - t_i)\sigma_T \ln(10)\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\log(t - t_i) - a_T - b_T m_i}{\sigma_T} \right)^2 \right] \quad (2.16)$$

where  $H(s) = 1$  if  $s > 0$  and  $H(s) = 0$  if  $s \leq 0$ , and  $a_T$ ,  $b_T$ , and  $\sigma_T$  are parameters. The location density  $h_{1i}$  is considered as:

$$h_{1i}(x, y) = \frac{1}{2\pi\sigma_A^2 10^{b_A m_i}} \exp \left[ -\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma_A^2 10^{b_A m_i}} \right] \quad (2.17)$$

where  $\sigma_A$  and  $b_A$  are parameters. The total rate density is obtained by summing over all past occurrences, including earthquakes outside  $R$ , which could affect the rate density within  $R$ :

$$\lambda(t, m, x, y) = \mu\lambda_0(t, m, x, y) + \sum_{t_i \geq t_0; m_i \geq m_0} \eta(m_i)\lambda_i(t, m, x, y) \quad (2.18)$$

where  $\mu$  is a parameter,  $\lambda_0$  is a baseline rate density,  $t_0$  is the time of the beginning of the catalog, and  $\eta$  is a normalizing function. The parameter  $\mu$  can be interpreted as the failure-to-predict rate, i.e., the proportion of earthquakes that occur without an appreciable sequence of precursory shocks. The function  $\eta$  needs to be such that the long-run average rate density under the EEPAS model, as a function of magnitude, is the same as that of the catalog. It thus depends on the Gutenberg-Richter  $b$ -value, the parameters  $a_M$ ,  $b_M$  and  $\sigma_M$  and the mean weight  $E(w)$  of earthquakes in the catalog, and can be shown to have the following form:

$$\eta(m) = \frac{b_M(1 - \mu)}{E(w)} \exp \left[ -\beta \left( a_M + (b_M - 1)m + \frac{\sigma_M^2 \beta}{2} \right) \right] \quad (2.19)$$

with  $\beta = b \ln 10$ , where  $b$  is the Gutenberg-Richter  $b$ -value.

The EEPAS model is similar to the ETAS model. Like ETAS, the structure is formally that of a branching process. However, conceptually there is a big difference; the idea of the EEPAS model is not that a small earthquake triggers a larger one but that it provides evidence that a larger earthquake may be in preparation (Evison and Rhoades, 2001). Also, the component distributions by the EEPAS model do not follow power laws in time, location, or earthquake moment. The normal, Lognormal, and bivariate normal distributions adopted for  $g_{1i}$ ,  $f_{1i}$ , and  $h_{1i}$ , respectively, are chosen to be consistent with normally distributed errors in the predictive relations.

In the baseline model adopted in this model, the rate  $\lambda_0$  depends on the Proximity to Past Earthquakes (PPE) in the catalog. The PPE model has most of the features

of the forecasting model proposed by Jackson and Kagan (1999). The PPE model has the form

$$\lambda_0(t, m, x, y) = f_0(t)g_0(m)h_0(x, y) \quad (2.20)$$

where  $f_0$  is the time function,  $g_0$  is the magnitude density and  $h_0$  is the density for location. The time function is given by

$$f_0(t) = \frac{1}{t - t_0}. \quad (2.21)$$

The magnitude density is that implied by the Gutenberg-Richter frequency-magnitude law:

$$g_0(m) = \beta \exp[-\beta(m - m_c)] \quad (2.22)$$

where  $\beta$  is defined as above. Finally, the function  $h_0$  is the sum over all earthquakes from time  $t_0$  up to, but not including, time  $t$  (time-independent) of smoothing Kernels of the form

$$h_{0i}(r_i) = a(m_i - m_c) \frac{1}{\pi} \left( \frac{1}{d^2 + r_i^2} \right) + s \quad (2.23)$$

where  $r_i$  is the distance in km between  $(x, y)$  and the epicenter  $(x_i, y_i)$  of the  $i$ -th earthquake;  $a$  is a normalization constant;  $d$  is a smoothing distance; and  $s$  is a small constant to allow for earthquakes far from past earthquakes. The constant  $a$  is set so that the integral of  $h_0$  over the region of surveillance  $R$  is equal to the number of earthquakes exceeding  $m_c$  within  $R$  over the period to which the model is fitted. The rate density under the PPE model thus diminishes gradually with time between earthquake occurrences and jumps upwards when new earthquakes occur. The function  $h_0$  incorporates the new earthquake locations and  $f_0$  takes into account of the passage of time.

The EEPAS method is generally the summation of the two methods for simultaneously estimations of the aftershock rate as well as the earthquake background rate. The assumption that the rate of aftershock triggering decays based on a Log-normal distribution is not always sufficient. The component  $f_0$  of the long-term earthquake rate covers just time-decreasing types of earthquake rates, which are not always applicable. The role of the censored time (the elapsed time since the last earthquake in the catalog) is not considered in this model. Also the numerous parameters in the model result in more uncertainties in the estimation process, when less data are available.

## 2.2.2 Long-term time-dependent models

### Kernel estimator for hazard function

Faenza et al. (2007) discuss the time-dependent behavior of seismicity with the purpose to quantify its impact on seismic hazard in their study region, the Lower Rhine Embayment, which is one of the regions of highest seismic hazard and risk in Germany (Hinzen and Oemisch, 2001). Faenza et al. (2007) apply the ETAS (Ogata, 1988) to model the short-term clustering of the inter-event times, and via a Monte Carlo technique, the ETAS model is rendered on time scales 50 years of exposure time, which is the standard seismic safety regulations worldwide for normal buildings (structures). For the long-term behavior, Faenza et al. (2007) perform a synthetic test on the impact of possibly cyclic activity on a single major fault. In this case, the impact for varying degrees of knowledge about the past fault activity is quantified.

For the evaluation of the hazard for long-term behavior, a sufficient large number of synthetic catalogs ( $N$ ) of time duration  $t$  is generated. Each catalog can be considered as a representation of seismicity in the time period of interest. Then a set of ground motions is evaluated for each catalog. The probability of non-exceedance of a level  $A^*$  of ground motion at a specific site in the time  $t$  is computed by counting the intervals in which  $A^*$  did not occur

$$P(A^*; t) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N H(A^* - A_{max,i}) \quad (2.24)$$

where  $N$  is the number of catalogs of time duration  $t$ ;  $H$  is the *Heaviside* function and  $A_{max,i}$  is the maximum ground motion value occurred at a site of interest during the  $i$ -th catalog of time duration  $t$ . The complement of  $P(A^*; t)$  is the probability that  $A^*$  is exceeded at least once in the time period  $t$ . An alternative way to estimate the probability of non-exceedance is to consider directly the empirical density function of the maximum acceleration values  $A_{max,i}$ . Each probability of non-exceedance is identical to the corresponding percentile of this distribution (further details in Beauval et al. 2006).

For the short-term behavior, Faenza et al. (2007) first discuss how the coefficient of variation (CV) of the inter-event times can show the behavior of the distribution of the earthquake inter-event times. The CV is expressed as the ratio of the standard deviation to the mean. If the CV is less than 1, the distribution corresponds to a quasi periodic behavior in the temporal domain; if the CV is equal to 1 then the distribution is random in the temporal domain, since the standard deviation and the mean have the same value; and, finally, if the CV is larger than 1 the distribution has a cluster behavior (Cox and Lewis, 1966).

Faenza et al. (2007) discuss two other methods to model the short-term behavior

of the earthquakes. The first method is a non-parametric estimation of the hazard function. The Tanner and Wong (1984) technique is applied and approaches the problem of estimating the hazard function directly by smoothing the empirical rate. The random variables considered are the inter-event time and the censoring time. The consideration of the censoring time, which represents the time elapsed between the most recent event and the end of the catalog, becomes very important in time-dependent analyses. The Kernel estimator of the hazard function  $\lambda(t)$  is

$$\lambda(t) = \sum_{i=1}^n \frac{\delta_i}{n-i+1} K_{\Theta}(t-y_i), \quad (2.25)$$

where  $y_i$  is the  $i$ th ordered random variable of the system,  $\delta_i$  is an indicator associated to  $y_i$ , that is  $\delta_i = 1$  in case of inter-event data,  $\delta_i = 0$  in case of censoring;  $K_{\Theta}$  is the Kernel function depending on a positive smoothing vector  $\Theta$ ; and  $n$  is the total number of data points. The smoothing vector  $\Theta$  is evaluated using a modified likelihood criterion. The details of the algorithm can be found in Tanner and Wong (1984).

Faenza et al. (2007) conclude that neglecting aftershocks lead to an underestimation of the hazard of 8 percent at 90 percent probability of non-exceedance in 50 years. Although the impact is not very high and taking into account that each other ingredient within the chain of the hazard computation has further uncertainties, this result is important since neglecting the time-dependent behavior (aftershock behavior) leads to a systematic underestimation of the estimated values.

The second method is applied with the purpose to examine the effect of the repeating occurrence of main shocks. Faenza et al. (2007) study the behavior of a generic fault that follows the Brownian Passage Time (BPT) distribution (also known as inverse Gaussian distribution). They especially evaluate the influence of the aperiodicity and the elapsed time on the hazard estimation. Knowing the fault and the statistic of the earthquake recurrence times (i.e. the elapsed time and the BPT distribution) the impact on hazard ranges from 5 to 10 percent for the 90 percent probability of non-exceedance in 50 yr. For the two special cases of an ongoing aftershock sequence on the one hand and an elapsed time equal to the mean recurrence time on the other hand, the increase of the hazard value is more than 10 percent at the level of 90 percent probability of non-exceedance. However, after the complete decay of the aftershock activity and for aperiodicity values larger than 0.3, the impact on hazard is negligible for any elapsed time. Faenza et al. (2007) deduce that the implementation of a "long-term" dependence of the seismicity via the BPT distribution does not substantially affect the hazard estimation for the region of the study.

## Proportional hazard function

Faenza et al. (2003) apply a method to model the spatio-temporal occurrence rate of earthquakes through a non-parametric multidimensional hazard function using the proportional hazard model described by Cox (1972) and Kalbfleisch & Prentice (1980). This model deals with two kinds of random variables, namely the inter-event time and the censoring time. At each one of these random variables a vector of covariates is attached, which carries out some information, i.e. spatial/tectonic information in the corresponding subregion.

Following Cox (1972), the proportional hazard model for a time  $t$  since the last event can be written as:

$$\lambda(t, \mathbf{z}) = \lambda_0(t) \exp(\mathbf{z}\beta) \quad (2.26)$$

where  $\mathbf{z}$  is the vector of covariates,  $\beta$  is a vector of coefficients, and  $\lambda_0(t)$  is an arbitrary base-line hazard function. The survivor function,  $S(\cdot)$ , (or a possible survivor function) and probability density function,  $f(\cdot)$ , for this hazard model are

$$\begin{aligned} S(t, \mathbf{z}) &= \exp \left[ - \int_0^t \lambda_0(u) \exp(\mathbf{z}\beta) \, du \right] \\ &= \left( \exp \left[ - \int_0^t \lambda_0(u) \, du \right] \right)^{\exp(\mathbf{z}\beta)} \\ &= S_0(t)^{\exp(\mathbf{z}\beta)} \end{aligned} \quad (2.27)$$

and

$$f(t, \mathbf{z}) = \lambda(t, \mathbf{z})S(t, \mathbf{z}). \quad (2.28)$$

In this model the covariates act multiplicatively on the hazard function and  $\mathbf{z}$  does not depend on time. This means that the form of the base-line  $\lambda_0(\cdot)$  is always the same apart from a multiplication factor. From a physical point of view, the mechanism of earthquake occurrence described by the function  $\lambda_0(\cdot)$  is the same for different areas; only the parameters of the system, i.e.  $\exp(\mathbf{z}\beta)$ , can vary. Kalbfleisch & Prentice (1980) suggest the most likely probabilistic model for the base-line hazard function using the empirical data.

In order to estimate the parameters  $\beta$ , suppose  $N_1$  inter-event times  $t_i$  ( $i = 1, \dots, N_1$ ) with  $t_{(i)}$  as the inter-event times in increasing order and corresponding covariates  $\mathbf{z}_{(1)}, \dots, \mathbf{z}_{(N_1)}$ , as well as  $N_2$  censored times  $\tilde{t}$ , of which  $n_i$  are censored in the  $i$ -th interval  $[t_{(i)}, t_{(i+1)})$ , i.e.  $t_{(i)} \leq \tilde{t}_{i1}, \dots, \tilde{t}_{in_i} < t_{(i+1)}$  ( $i = 0, \dots, N_1$ ), and corresponding covariates  $\mathbf{z}_{i1}, \dots, \mathbf{z}_{in_i}$ . Considering the rank vector  $\mathbf{r}(t_i) =$

$[(1), (2), \dots, (N_1)]$  of the inter-event times, the censored times contain only partial information on the rank vector, because the ordering vector including the censored times is not necessarily the true ordering. The contribution of the censored times in the time interval  $[t_{(i)}, t_{(i+1)})$  to the marginal likelihood is considered as,

$$g(t_{(i)}) = \exp \left[ - \sum_{j=1}^{n_i} \exp(\mathbf{z}_{ij}\beta) \int_0^{t_{(i)}} \lambda_0(u) \, du \right]. \quad (2.29)$$

Then the likelihood function is proportional to

$$\begin{aligned} P(\mathbf{r}; \beta) &= P\{\mathbf{r} = [(1), \dots, (N)]; \beta\} \\ &= \int_0^\infty \int_{t_{(1)}}^\infty \cdots \int_{t_{(N_1-1)}}^\infty \prod_{i=1}^{N_1} f(u_i; \mathbf{z}_i) g(u_i) \, du_{N_1} \cdots du_1 \\ &= \prod_{i=1}^{N_1} \frac{\exp(\mathbf{z}_i\beta)}{\sum_{j \in \Omega'(t_{(i)})} \exp(\mathbf{z}_j\beta)} \end{aligned} \quad (2.30)$$

where  $\Omega'(t_{(i)})$  is the set of labels attached to the inter-event times and censored times with length  $\geq t_{(i)}$ , i.e.  $\Omega'(t_{(i)}) = \{[(j), j_1, \dots, j_{n_j}], j = i, \dots, N_1\}$ .

The case of ties in the dataset has been also discussed. A tie means two or more equal inter-events or censoring times in the dataset. In this case suppose  $d_i$  inter-event times with length  $t_{(i)}$  ( $i = 1, \dots, N_1$ ), then the likelihood function will change to

$$L = \prod_{i=1}^{N_1} \frac{\exp(\mathbf{s}_i\beta)}{\left[ \sum_{j \in \Omega'(t_{(i)})} \exp(\mathbf{z}_j\beta) \right]^{d_i}} \quad (2.31)$$

where  $\mathbf{s}_i$  is the sum of the covariates of the inter-event times with length  $t_{(i)}$ .

The maximum likelihood estimate  $\hat{\beta}$  can be obtained as a solution of the system of equations

$$U_j(\hat{\beta}) = \frac{\partial \ln(L)}{\partial \beta_j} \quad (2.32)$$

where  $U_j$  is the gradient vector. Some mild conditions on the covariates and censoring are required to ensure the asymptotic normality of  $\hat{\beta}$  (Kalbfleisch & Prentice 1980). Then  $\hat{\beta}$  is a normally distributed random variable with the two parameters Mean =  $\beta$  and Variance =  $I(\beta)^{-1}$ , where  $I$  is the Fisher information matrix as:

$$I_{ij} = - \frac{\partial^2 \ln(L)}{\partial \beta_i \partial \beta_j} \quad (2.33)$$

The asymptotic normality of  $\hat{\beta}$  is very important, while it allows very simple and routine tests like to check the relative importance of the factors included in the covariates (Kalbfleisch 1985).

The next step is to estimate the hazard function base-line  $\lambda_0(\cdot)$ . For that reason the survivor function is estimated first. As before,  $t_{(1)}, t_{(2)}, \dots, t_{(N_1)}$  are the ordered inter-event times,  $P_i$  the set of labels associated with inter-event times with length  $t_{(i)}$ , and  $Q_i$  the set of labels associated with the censored times censored in  $[t_{(i)}, t_{(i+1)})$  ( $i = 0, \dots, N_1$ ), where  $t_{(0)} = 0$  and  $t_{(N_1+1)} = \infty$ . The censored times in the interval  $[t_{(i)}, t_{(i+1)})$  are  $\tilde{t}_j$  where  $j$  ranges over  $Q_i$ . Then the proposed likelihood function has the form of

$$L = \prod_{i=0}^{N_1} \left\{ \prod_{j \in P_i} \left[ S_0(t_{(i)})^{\exp(\mathbf{z}_j \beta)} - S_0(t_{(i)}^+)^{\exp(\mathbf{z}_j \beta)} \right] \prod_{k \in Q_i} S_0(\tilde{t}_k^+)^{\exp(\mathbf{z}_k \beta)} \right\} \quad (2.34)$$

where  $S_0(t_{(i)}^+) = \lim_{\delta \rightarrow 0^+} S_0(t_{(i)} + \delta)$  and  $P_0$  is empty. It is clear that  $L$  is maximized by taking  $S_0(\tilde{t}) = S_0(t_{(i)}^+)$  for  $t_{(i)} < \tilde{t} < t_{(i+1)}$  and allowing probability to fall only at the observed inter-event times  $t_{(i)}$ .

The form of the base-line is completely empirical, therefore it has a discrete form. The survivor function at a generic time  $t^*$  since the last event for a discrete process with censored times is

$$S(t^*) = \prod_{j|t_j < t^*} (1 - \lambda_j) \quad (2.35)$$

where  $\lambda_j = P(T = t_j | T \geq t_j)$ . This equation is the discrete form of equation 2.27. The survivor base-line,  $S_0$ , is also measured as a discrete model with hazard contribution  $\lambda_j$  at  $t_{(j)}$  ( $j = 0, \dots, N_1$ ). Then

$$S_0(t_{(i)}) = S_0(t_{(i-1)}^+) = \prod_{j=0}^{i-1} \alpha_j \quad (2.36)$$

where  $\alpha_0 = 1$ . Substituting equations 2.27 and 2.36 in equation 2.34, and rearranging the terms, the equation

$$L = \prod_{i=1}^{N_1} \left\{ \prod_{j \in P_i} \left[ 1 - \alpha_i^{\exp(\mathbf{z}_j \beta)} \right] \prod_{k \in \Omega'(t_{(i)}) - P_i} \alpha_i^{\exp(\mathbf{z}_k \beta)} \right\} \quad (2.37)$$

will be obtained.

The estimation of the survivor function can be carried out by joint estimation of  $\alpha_i$  and  $\beta$  in equation 2.37. More simply, taken  $\beta = \hat{\beta}$  as estimated previously, one

can maximize the logarithm of equation 2.37 with respect to  $\alpha_i$ . In this case, the following equation will be obtained

$$\sum_{j \in P_i} \frac{\exp(\mathbf{z}_j \hat{\beta})}{1 - \hat{\alpha}_i^{\exp(\mathbf{z}_j \hat{\beta})}} = \sum_{k \in \Omega'(t_{(i)})} \exp(\mathbf{z}_k \hat{\beta}). \quad (2.38)$$

If there is no tie, equation 2.38 can be directly solved for  $\hat{\alpha}_i$ , otherwise an iterative solution is required. A suitable initial value for such iteration is

$$\ln(\alpha_{i0}) = \frac{-d_i}{\sum_{k \in \Omega'(t_{(i)})} \exp(\mathbf{z}_k \hat{\beta})}, \quad (2.39)$$

obtained by approximating  $\hat{\alpha}_i^{\exp(\mathbf{z}_j \hat{\beta})} \simeq 1 + \exp(\mathbf{z}_k \hat{\beta}) \ln(\hat{\alpha}_i)$ . The estimated base-line of the survivor function for a generic time  $t^*$  since the last earthquake is therefore

$$\hat{S}_0(t^*) = \prod_{i|t_{(i)} < t^*} \hat{\alpha}_i, \quad (2.40)$$

which is a step function. The equation of the proportional model for a generic time  $t^*$  and covariates  $\mathbf{z}^*$  is

$$\hat{S}(t^*, \mathbf{z}^*) = \prod_{i|t_{(i)} < t^*} \hat{\alpha}_i^{\exp(\mathbf{z}_i \hat{\beta})}. \quad (2.41)$$

In order to check the validity of the modeling, the dataset is divided into two sets, one to use for the modeling (the learning phase) and the other to use for the checking the model (the validation phase). A method of the goodness-of-fit is applied, which can empirically check, how good the model fits the data of the validation phase. Each inter-event time  $t$  of this data set will be transformed through the equation

$$\hat{e}_i = \hat{\Lambda}_0(t_i) \exp(\mathbf{z}_i \hat{\beta}), \quad (2.42)$$

where

$$\hat{\Lambda}_0(t) = \sum_{i|t_{(i)} < t} \left( \sum_{k \in \Omega'(t_{(i)})} \exp(\mathbf{z}_k \hat{\beta}) \right)^{-1}, \quad (2.43)$$

and  $\hat{\beta}$  are estimated by the data of the learning phase. If the model is appropriate, the residuals  $\hat{e}_i$  should be similar to a sample drawn by an exponential distribution with  $\lambda = 1$ . Therefore, a comparison of the cumulative of the residuals  $\hat{e}_i$



with a theoretical exponential curve provides a goodness-of-fit test of the model. The comparison can be statistically checked through a one-sample Kolmogorov-Smirnov test (e.g. Gibbons 1971).

Finally, the authors discuss the application of the model to the Italian seismicity using a set of earthquakes with  $M_w \geq 5.5$  since 1600. They state that their first results obtained from the model explain that large earthquakes ( $M_w \geq 5.5$ ) in Italy tend to cluster in time and space, and the length of the time-cluster may reach a few years, and after this time the distribution of the earthquakes becomes a Poisson distribution. They also did not find any evidence of increasing hazard function. This method has been already used to model the spatio-temporal distribution of the worldwide big earthquakes (see Faenza et al. 2008).

The method belongs to the minimalist models, in which the concept is to use less theory but more information. The model thus requires more accurate data and does not consider any prior form for the distribution of the hazard function.

## **Renewal models**

The renewal models are frequently used to estimate the long-term time-dependent probability of the next large earthquake on specific faults or fault segments, where large shocks occur repeatedly at approximately regular intervals. Based on these types of models, it is assumed that the times between consecutive large earthquakes (inter-event times or recurrence intervals) follow a certain statistical distribution. Since the available records of large earthquakes in a given area is typically scarce (usually including less than ten events), so the empirical statistics are poor and then a theoretical distribution is fitted to the observed inter-event times and used to estimate future earthquake probabilities (Gonzalez et al. 2006).

These types of models can be divided into two groups. The first group; including Gamma, Lognormal, and Weibull distributions, are frequently used for earthquake inter-event time modeling, because they share three properties commonly observed for earthquake inter-event times (Michael, 2005). First, these times must be positive, and these distributions only exist for positive times; second, inter-event times much smaller than the average recurrence interval are rare; and third, the distribution of inter-event times decays slowly for times longer than the average. However, the behavior of Gamma and Weibull distributions, as a generalization of the Exponential distribution, is different from Lognormal distribution, which belongs to the Gaussian based distributions. These distributions and their characteristics will be broadly discussed in the next chapter.

The second group is including distributions derived from two simple physical models of earthquake recurrence which have been proposed as an alternative to those purely empirical approaches (i.e. the first group). They are the Brownian Passage Time model (Kagan and Knopoff, 1987; Ellsworth et al., 1999; Matthews

et al., 2002; WGCEP, 2003), and the Minimalist Model (Vazquez-Prada et al., 2002; Gomez and Pacheco, 2004). The first one represents the tectonic loading of a fault by a variable which evolves by superposition of an increasing linear trend and a Brownian noise term. Then, an earthquake occurs when this variable reaches a given threshold (Matthews et al., 2002). All earthquakes in this model are identical to each other<sup>1</sup>.

The Minimalist Model sketches the plane of a seismic fault, where earthquake ruptures start and propagate according to simplified breaking rules. This model generates earthquakes of various sizes, and only the time between the largest ones (the characteristic earthquakes, that break the whole model fault) are considered for the inter-event time distribution (Gonzalez et al. 2006). The distributions derived from these two models, as well as the Gamma, Lognormal and Weibull, generally represent fairly well the observed distribution of large-earthquake inter-event times (Gomez and Pacheco, 2004). However, they differ significantly in their probability predictions for times much longer than the mean inter-event time of the data. Thus, it seems convenient to take all their different predictions into account.

Gonzalez et al. (2006) use the mentioned models to describe the seven M 6 main shocks in the Parkfield area. They do not consider the censored time, although it is an important variable in the case of time-to-failure modeling. They use a method of moments to estimate the parameters of each continuous parametric model, i.e. Weibull, Gamma, Lognormal and BPT. This method is a simple estimation method, but the estimators corresponding to this method are not necessarily sufficient, because they sometimes fail to take into account all relevant information in the sample (Stigler 1973). The Maximum Likelihood Estimation (MLE) method is mostly proposed in such cases.

In order to find the best model among the different models, the authors focus on the residuals of each distribution. This method is also inconsistent because the number of data and the likelihood of each model are not involved in the model selection process. On the other hand, they do not discuss the goodness-of-fit, which shows, how good the models can basically describe the data.

Hebden and Stein (2009) propose a time-dependent model for hazard assessment for the New Madrid seismic zone and Charleston, South Carolina, and compare the results with the time-independent hazard maps. They apply two Normal distributions, with the same mean recurrence time as for the time-independent model and two different standard deviations, as well as a Lognormal distribution, with the same mean recurrence time, for the inter-event time distribution of large earthquakes. Their results show that the time-dependent model predicts noticeably

---

<sup>1</sup>Here "identical" means that earthquakes follow the same distribution in time. In other word, the earthquake inter-event times are identically distributed

lower hazard for two various 50-year periods, namely 2000-2050 and 2100-2150. They also declare that, if the large earthquake has not occurred by 2200, the hazard predicted in the next 50 years would be higher than predicted by the time-independent model. They state that the dramatic differences between the time-dependent and time-independent hazard maps point out that the hazard estimation in those areas, in which there are rare reliable earthquake data is a very uncertain enterprise. However, they conclude that the lower hazards predicted for the central and eastern part of the United States by the time-dependent models are more plausible.

Ellsworth et al. (1999) discuss the time-dependent time-to-failure models for the earthquake inter-event time modeling. They first introduce different renewal distributions, i.e. Weibull, Gamma, Brownian Passage Time and Lognormal. They declare that at present, it is not possible to discriminate between such candidate models, given the limited and uncertain nature of earthquake recurrence data, although it can be firmly possible to reject a Poisson process or Exponential distribution (Ellsworth, 1995). This declaration is possibly not as trivial as stated. The consideration of more reliable historical data, if possible, can effectively increase the level of reliability of the inter-event time modeling as well as best model choice process. Furthermore, they introduce the Brownian Passage Time distribution as a physically based model and discuss the results of this model using the data from the San Andreas fault in California. After analyzing of 37 series of recurrent earthquakes, with magnitude from -0.7 to 9.2, they suggest a provisional generic value of  $\alpha = 0.5$  (aperiodicity parameter of Brownian Passage Time distribution). For this value of  $\alpha$ , the hazard function exceeds the mean rate for times  $> \mu/2$  and is  $\approx 2/\mu$  for all times  $> \mu$  (where  $\mu$  is the mean inter-event time). Application of this model to the next M6 earthquake on the San Andreas fault at Parkfield, California, suggests that the annual probability of the earthquake is between 1/10 and 1/13. The Brownian Passage Time distribution has been also used in some other studies like, Rundle et al. (2006).

Zöller et al. (2008) discuss a statistical distribution to describe the recurrence of large earthquakes in a specific fault zone. They consider the same class as the Brownian relaxation oscillator proposed by Matthews et al. (2002). Zöller et al. (2008) consider a load state variable that increases over time, reaches a static threshold and relaxes instantaneously back to the ground state. However, considering the Gutenberg-Richter law, they apply a power-law distribution for fluctuations (the load state variables) instead of the Gaussian distribution as assumed in the Brownian relaxation oscillator. Then, they build a renewal loading process and discuss the recurrence time of this process. They show that the coefficient of variation (the standard deviation divided by the mean recurrence times) of recurrence times depends solely on the b-value of the corresponding Gutenberg-Richter law. Zöller et al. (2008) discuss that since the variance of the distribution of the fluctua-

tions is finite and the number of fluctuations is relative large, then the central limit theorem implies that the recurrence times follow a Brownian passage-time (BPT) distribution. They calculate individual recurrence-time distributions for specific fault zones without tuning free parameters. Then, the mean recurrence time can be estimated from geological or paleoseismic data, and the standard deviation is determined from the frequency-size distribution, namely, the Gutenberg-Richter b-value, of an earthquake catalog. Zöller et al. (2008) apply the approach to the Parkfield segment of the San Andreas fault in California and for a long simulation of a numerical fault model. Assuming power-law distributed earthquake magnitudes up to the size of the recurrent Parkfield event (M 6), they find a coefficient of variation that is higher than the value obtained by a direct fit of the BPT distribution to seven large earthquakes. Finally, Zöller et al. (2008) argue that the BPT distribution is a reasonable choice for seismic hazard assessment because it governs not only Brownian motion with drift but also models with power-law statistics for the recurrence of large earthquakes in an asymptotic limit.

### **Clock-change in hazard using stress change models**

Parsons (2005) discusses the variability, parameter sensitivity, and limitations of the time-dependent approach as well as the circumstances under which stress interactions should be added. This effort concentrates solely on the concept of static stress transfer, the theorized lasting change in the stress field resulting from displacements caused by previous fault slip or magmatic intrusions.

The time-dependent probability calculation follows the renewal hypothesis of earthquake regeneration such that earthquake likelihood on a fault is lowest just after the last event. As stress increases, the odds of another earthquake increase. A time-dependent probability calculation sums a probability density function  $f(t)$  as

$$P(t \leq T \leq t + \Delta t) = \int_t^{t+\Delta t} f(t) dt \quad (2.44)$$

where  $f(t)$  can be any distribution, such as Lognormal (Nishenko and Buland, 1987), Weibull (Hagiwara, 1974), Gamma (Utsu, 1984), or Brownian Passage Time (Kagan and Knopoff, 1987; Matthews et al., 2002). These functions distribute around some mean inter-event time ( $\mu$ ), and the width of the distributions represents inherent variability ( $\alpha$ ) on recurrence. For example, a very narrow distribution implies very regular recurrence.

Two commonly applied probability density functions, the Lognormal

$$f(t, \mu, \alpha) = \frac{1}{t\alpha\sqrt{2\pi}} \exp\left(\frac{-(\ln t - \mu)^2}{2\alpha^2}\right) \quad (2.45)$$

and Brownian Passage Time,

$$f(t, \mu, \alpha) = \sqrt{\frac{\mu}{2\pi\alpha^2 t^3}} \exp\left(-\frac{(t - \mu)^2}{2\mu\alpha^2 t}\right) \quad (2.46)$$

have characteristics that qualitatively mimic earthquake renewal. The distributions are asymmetric, with less weight at short recurrence time which, when integrated, translates to very low probability early in the earthquake cycle. They are defined by two parameters, i.e. the mean inter-event time, and the coefficient of variation or aperiodicity that govern their shapes. The distributions differ in their asymptotic behavior; integration of the rate of the Lognormal distribution to very long times asymptotes to zero, whereas the Brownian passage time distribution asymptotes to a fixed value. According to Matthews et al. (2002), this behavior favors the Brownian distribution for hazard calculations. Parsons (2005) explores effects of input parameter variability on time-dependent probability calculation to establish its inherent variation before stress changes are applied.

In the next step, he uses the paleoseismic records found on the San Andreas fault in southern California as well as historical catalogs in the same area. Then he constructs a forward approach using Brownian Passage Time distributions with a range of mean inter-event time ( $\mu$ ) and aperiodicity ( $\alpha$ ). In this approach events were repeatedly drawn at random from each distribution in an attempt to match the observed event windows. Distributions that succeeded were tallied. After the consideration of the effects of the unknown elapsed time since the last event, the choice of the distribution as even long paleoseismic catalogs permit a broad array of equally viable earthquake probability forecasts. In some cases the array appears to have a central peak that might provide some basis for expressing a preferred value (Savage, 1991, 1992), but in others not. Time-dependent probability calculations may not be warranted for paleoseismic catalogs with fewer than 10 events. Choice of probability density function and knowledge of elapsed time (censored time) are shown to have potentially broadening effects on the array of probabilities. Moreover, the stress transfer is introduced to the calculations to investigate how large it must be to skew the array of probabilities. The inherent variability and uncertainty from stress change calculations are explored because their influence on earthquake probability scales with their magnitude.

The stress changes are estimated due to the co-seismic slip on the rupture plane and the shear stressing rate (also called tectonic loading) below an a priori defined locking depth of the fault. Technically the solution is achieved by embedding the faults in an elastic half space and using the solution scheme of Okada (1992). For further assessment of the stress change the change of Coulomb Failure Stress (CFS) is calculated for the fault planes of interest (King et al. 1994) using the following definition:

$$\Delta CFS = \Delta\tau + f_c(\Delta\sigma_n + \Delta p) \quad (2.47)$$

where  $\Delta\tau$  is the change in shear stress on the fault (set positive in the direction of fault slip of the future earthquake),  $f_c$  is the coefficient of friction,  $\Delta\sigma_n$  is the change in normal stress acting on the receiver fault (set positive for unclamping), and  $\Delta p$  is the pore pressure change. Positive  $\Delta CFS$  brings the fault closer to failure, indicating a clock-advance with respect to the next earthquake and vice versa, a negative  $\Delta CFS$  departs the fault from failure and delays the next event. Under the renewal model,  $\Delta CFS$  increases with time due to tectonic plate motion only, i.e. simply an increase of shear stress only. Thus, the clock-advance or delay (clock change  $T'$ ) can be estimated by dividing the  $\Delta CFS$  with the shear stressing rate  $\dot{\tau}$ , as:

$$T' = \frac{\Delta CFS}{\dot{\tau}}. \quad (2.48)$$

The probability is accrued from the last earthquake time adjusted by the clock change ( $T_0 + T'$ ) (Working Group on California Earthquake Probabilities, 1990). Alternatively, the earthquake inter-event time  $\mu$  can be adjusted by the clock change as  $\mu = \mu_0 - T'$ .

The choice of whether to change the elapsed time or the inter-event time has a potentially significant effect on the resulting earthquake probability calculation. Probability calculated with a clock change, or elapsed time shift, is most significant at the time of the stress change, and then asymptotes to the maximum probability value with time. A change in recurrence interval has the opposite characteristic. The probability change is smallest at the time of the stress change, and then asymptotes to a permanent offset in probability. These differences are caused because in one instance (recurrence interval change), the mean of the recurrence distribution is changed and thus the maximum calculated probability changes. In the case where the last earthquake time is changed (clock change), the center of the distribution is unaffected; thus, after a long elapsed time, the maximum calculated probability stays the same.

Making a clock change or inter-event time adjustment in probability calculations requires an estimate of the shear stressing rate  $\dot{\tau}$ . A commonly used approach to estimate  $\dot{\tau}$  is to simulate the earthquake fault as a dislocation; the seismogenic part of the fault is locked while the rest of the fault is allowed to slip freely (Parsons, 2002b, 2004).

At this point, Parsons (2005) discusses the effect of time-dependent stress change on the seismicity rate. A transient change in seismicity rate following an earthquake is a long recognized (Omori, 1894; Dutton, 1904), widely observed phenomenon. The physics of the process (Scholz, 1968; Dieterich, 1994; Marcellini, 1997; Kilb et al., 2002; Felzer et al., 2003) and how to work it into earthquake probability calculations (Dieterich and Kilgore, 1996; Matthews et al., 2002; Hardebeck, 2004; J. Gomberg et al., 2005) are still debated. Much like the

overall question whether to incorporate stress transfer into time-dependent earthquake probability calculations, the issue whether to fold in transient seismicity rate changes pits observational evidence against the uncertainty and variability of probability calculations. At present, three methods for transient probability change calculations are in use; two techniques from rate-and-state friction theory (Dieterich and Kilgore, 1996; Toda et al., 1998; Hardebeck, 2004), and one from the Brownian Passage Time model (Matthews et al., 2002; WGCEP, 2003).

The rate-and-state transient effect describes an expected enhanced rate of earthquake nucleation resulting from a stress increase, and can be expressed as a probability. For a stress decrease, the rate of nucleation declines, and eventually recovers. Dieterich (1994) derived a time-dependent seismicity rate  $R(t)$ , after a stress perturbation as

$$R(t) = \frac{r}{[\exp(-\Delta CFS/A\sigma) - 1] \exp(-t/t_a) + 1} \quad (2.49)$$

where  $r$  is the steady state seismicity rate,  $\Delta CFS$  is the change of Coulomb Failure Stress,  $\sigma$  is the normal stress,  $A$  is a fault constitutive constant, and  $t_a$  is an observed aftershock duration, a fault-specific parameter.

The transient change in expected earthquake rate  $R(t)$  after a change of Coulomb Failure Stress can be related to the probability of an earthquake of a given size over the time interval  $\Delta t$  through a nonstationary Poisson process as

$$P(t, \Delta t) = 1 - \exp \left[ - \int_t^{t+\Delta t} R(t) dt \right] = 1 - \exp[-N(t)] \quad (2.50)$$

after Dieterich and Kilgore (1996), where  $N(t)$  is the expected number of earthquakes in the interval  $\Delta t$ . This transient probability change is superimposed on the permanent change (in both probability and seismicity rate) that results from a time shift, or a change in the repeat time as discussed previously. Integrating for  $N(t)$  yields

$$N(t) = r_p \left\{ \Delta t + t_a \ln \left[ \frac{1 + [\exp(-\Delta CFS/A\sigma) - 1] \exp(-\Delta t/t_a)}{\exp(-\Delta CFS/A\sigma)} \right] \right\} \quad (2.51)$$

where  $r_p$  is the expected rate of earthquakes associated with the permanent probability change (Toda et al., 1998). This rate can be determined by again applying a stationary Poisson probability expression as

$$r_p = \left( \frac{-1}{\Delta t} \right) \ln(1 - P_c) \quad (2.52)$$

where  $P_c$  is a conditional probability, and can be calculated using any distribution. There are advantages in using the rate-and-state model, though a number of important assumptions must also be made. Rate-and-state friction includes the friction

parameter via using  $\Delta CFS$ . Under the Dieterich (1994) model, a group of faults or even a single fault is considered an infinite population of earthquake nucleation sites that are near to failure. These conditions are treated as Poissonian (Dieterich and Kilgore, 1996), are independent of the fault rupture history, and are applied on top of the static probability change.

Hardebeck (2004) suggests that superimposing the transient response onto a clock change, or static probability change is double counting, and proposes a method to calculate interaction probabilities that uses time-varying clock change corrections. Integration limits of a probability density function are changed with each time step to simulate a variety of potential earthquake nucleation sites. If a rate-and-state nucleation model (Dieterich, 1994) is implemented, as (Hardebeck, 2004)

$$T_{new} = \frac{A\sigma}{\dot{\tau}} \ln \left[ \left( \exp \left( \frac{T_{old}\dot{\tau}}{A\sigma} \right) - 1 \right) \exp \left( \frac{\Delta CFS}{A\sigma} \right) + 1 \right] \quad (2.53)$$

then a transient response in probability change is obtained. The probability that nucleation will occur between the new limits is calculated using the original density function. Comparison of the two rate-and-state methods shows a similar shape in the evolution of probability with time; a higher peak is obtained with the time-varying clock change method (Hardebeck, 2004) than with the Dieterich and Kilgore (1996) method.

The problem of the application of both statistical (BPT model) and physical (stress change) models simultaneously, which results in a superposition of the role of the stress change in the earthquake time-to-failure modeling, remains anyhow. The use of a forward method to estimate the aperiodicity parameter of the BPT model is not very significant. The goodness-of-fit is not checked for the BPT as far as for the other distributions. The uncertainties of the estimation (or calculation) of the physical parameters and generalization to the real situations are still non-solved.

Gomberg et al. (2005) focus on the combination of both probabilistic and physical models. Using the idea of Stein et al. (1997) and Hardebeck (2004), they develop a general framework based on a simple generalized rate change formulation and apply it to two approaches, namely probabilistic and physical, to show how they relate to one another. They also attempt to show the connection between models of seismicity rate changes applied to populations of independent faults as in background and aftershock seismicity and changes in estimates of the conditional probability of failure of a single fault.

One of the critical issues by using both physical and statistical methods simultaneously is that the statistical methods are generally based on the earthquake data including occurrence times, sizes and locations. These data also hold a stage of the stress change information in the given area. This means that statistical and physical approaches are correlated. This correlation may result in some superimposing of role of the stress change in the earthquake occurrence rate, using both



physical and statistical methods simultaneously. Hardebeck (2004) discusses this issue as well.

Another critical issue is the way the stress change is calculated. Uncertainties in  $\Delta CFS$  models are relative high and model assumptions over simplify the problem.

### **Several other methods**

Kuehn et al. (2008) apply the stress release model proposed by Vere-Jones (1978) and develop a method to consider the stress transfer and stress-triggering between distant faults and seismogenic zones over a multivariate extension of the simple form of the method. The hazard rate function corresponding to the model has an exponential form over the stress scalar (or simply stress), which is a non-decreasing function over stress, but not necessarily over time. The theory behind the model is simple. However, the predefined parameters involved in the model cannot be always easily and correctly guessed or estimated.

Molchan (2005) performs an analysis of the distribution of inter-event times in a stationary point process. His study is motivated by the discovery of unified scaling laws for  $\tau$  (the time between two consecutive events) for the case of seismic events, e.g. modified Omori's law (Utsu 1961) and a new scaling law by Bak et al. (2002) and Corral (2003,2004a,b). He demonstrates that these laws cannot exist simultaneously in a seismogenic area. He shows under some assumptions that if, after rescaling to ensure that statistical expectation of  $\tau$  is equal 1, the inter-event time has a universal distribution  $F$ , then  $F$  must be exponential (the process must be Poissonian). This result is however under the assumption of existence of a universal distribution for  $\tau$ . In the next chapters, it will be shown that the distribution function for  $\tau$  dose not generally follow the exponential distribution, at least in the case of this thesis, and it should be checked for every seismogenic zone, if possible.

Rotondi (1999) applies a multiple-change point problem to identify the periods with various seismicity in a seismic source zone. She defines the problem as to find the number and the time of  $k \geq 1$  change points of the occurrence rate in a sequence of earthquake observations under a generalized Poisson process assumption through a stepwise procedure. To solve this problem, she uses a revisable jump Metropolis-Hastings algorithm under a Markov Chain Monte Carlo (MCMC) method.

La Rocca (2008) assumes a renewal process for the large earthquake recurrence time during long time spans and propose a Bayesian nonparametric estimation of smooth hazard function for the occurrence rate of large earthquakes. This model, like other nonparametric models, requires more reliable data, although assuming a functional form for the Kernel of the nonparametric distribution.

## **2.3 Conclusions**

It can be concluded that most time-dependent models of the earthquake occurrence rate have been focused on aftershock sequences, which are basically short-term analyses of earthquake occurrence rates in time scales not longer than 50 years of exposure time. The methods applied to model the earthquake occurrence rate for long-term are generally based on the renewal process. The scarce information about large earthquakes as well as the broad domain of application of the renewal processes, also in other real-life problems, make the renewal process very applicable in the case of long-term earthquake occurrence rate modeling. In the next chapter, the renewal process will be discussed as well as various commonly used distributions under the assumption of the renewal process.

# Chapter 3

## An adapted approach to time-dependent earthquake occurrence rate modeling

In order to model the earthquake inter-event times, an adapted approach will be introduced in this chapter. Considering the most powerful statistics-based models applied in different forms of time-to-failure modeling, they will be developed for the application to the earthquake inter-event times. Then, the method of parameter estimation will be investigated. Finally, methods for selection of the best model will be discussed.

### 3.1 Empirical modeling

Mathematical models have been used in solving real-life problems from many different disciplines. This requires building a suitable mathematical model. Two different approaches applied to develop mathematical models are:

- *Theory-Based Modeling*, also called *physical-based modeling* or *white-box modeling*, is based on theories (from physical, biological, and social sciences) relevant to the problem.
- *Empirical Modeling*, also called *data-dependent modeling* or *black-box modeling*, is based on the available data forming the basis of model building. It does not require an understanding of the underlying mechanisms involved.

In the empirical approach, an analysis of data is first carried out. Then, the type of mathematical formulation appropriate to describe the data is determined (Murthy et al. 2004).

## 3.2 Time-to-failure modeling

The group of models called *time-to-failure* models is considered to analyze the time between two events, i.e. *failure*, in a process. The aim of this type of models is to find the distribution of the random variable of time between the events.

Consider a stochastic process  $X = (X_t)_{t \in T}$ , with  $T$  as the index set and  $X_t$  as the state of the process at time  $t$ . Let  $X_t$  be the number of failures within time  $t$  since starting time  $t_0$  with  $t \geq t_0$  (here considered as  $t_0 = 0$ ), with  $J_1, J_2, \dots$  as jump times (the time in which an event occurs) and  $\tau_1, \tau_2, \dots$  as inter-event times (or holding times; i.e. the time between two successive event occurrences). This process is illustrated in figure 3.1.

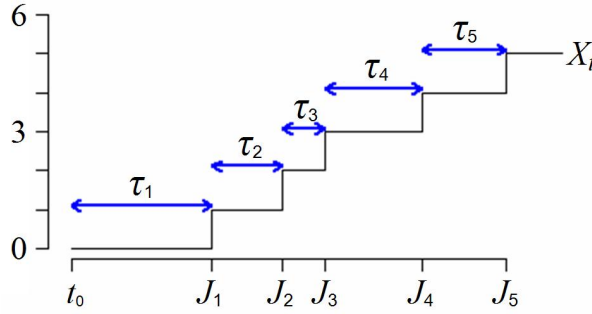


Figure 3.1: An example of a time-to-failure process, sample evolution of a renewal process  $X$  with inter-event times  $\tau_1, \tau_2, \dots$  and jump times  $J_1, J_2, \dots$ .

## 3.3 Renewal Process

A renewal process is a point process characterized by assuming the successive inter-event times  $\tau_1, \tau_2, \dots$  are independently, identically distributed (i.i.d). Let  $J_n = \tau_1 + \tau_2 + \dots + \tau_n$  denote the time up to the  $n$ th event. Then, the stochastic process  $X(t), t \geq 0$ , where  $X(t) = \max \{n : J_n \leq t\}$ , is called a renewal process. This stochastic process has the following properties:

$$\frac{X(t)}{t} \rightarrow \frac{1}{E[\tau_1]}, \text{ as } t \rightarrow \infty, \text{ almost surely.} \quad (3.1)$$

$$\frac{E[X(t)]}{t} \rightarrow \frac{1}{E[\tau_1]}, \text{ as } t \rightarrow \infty \quad (3.2)$$

$$\frac{Var[X(t)]}{t} \rightarrow \frac{Var[\tau_1]}{E[\tau_1]^3}, \text{ as } t \rightarrow \infty \quad (3.3)$$

where  $E$  and  $Var$  are the expectation and the variance of the corresponding random variable (Taylor and Karlin 1998). These properties state that in order to

characterize the renewal process, it is sufficient to estimate the distribution of the corresponding inter-event times  $\tau_1, \tau_2, \dots, \tau_n$ .

If the  $\tau_1, \tau_2, \dots$  are not identically distributed; i.e. they follow different distributions, then the process is not renewal anymore but it can be divided into several subprocesses which are renewal. In such cases, a change-point process will estimate the change-points corresponding to the different subprocesses.

If the  $\tau_1, \tau_2, \dots$  are not independently distributed, then the process is a nonstationary process. In this case, the nonstationary time series can be used in order to model the process.

### **3.3.1 Earthquake recurrence as a renewal process**

The renewal process assumption is generally used in the case of earthquake occurrence rate analysis for strong main shocks (see chapter 2). In this thesis, the process of earthquake recurrence for large earthquakes (i.e. earthquake with moment magnitude  $\geq 6$ ; cf. chapter 4) is assumed as a renewal process, where the inter-event times (i.e. the time between two successive earthquakes) are statistically considered as independent and identically distributed (i.i.d).

## **3.4 Statistical distribution of earthquake inter-event times**

In the case of earthquake inter-event times, the exponential distribution has been applied as the classical time-independent model. The time-dependent distributions, generally applied to earthquake inter-event time modeling, are containing: Weibull (Nishenko 1985), Gamma (Corral 2004a,b; Hainzl et al. 2006), Lognormal (Nishenko & Buland 1987; Michael & Jones 1998) and inverse Gaussian, known as Brownian Passage Time (Ellsworth et al. 1999). These distributions will form the base of earthquake inter-event time modeling in this thesis.

### **3.4.1 Weibull distribution**

The Weibull distribution is a continuous, skewed to the right, probability distribution. It is named after Waloddi Weibull who described it in detail in 1951, although it was first identified by Frechet (1927) and first applied by Rosin & Rammler (1933) to describe the size distribution of particles. The Weibull distribution is often used in the field of life data analysis due to its flexibility; i.e., it can

mimic the behavior of other statistical distributions such as normal and exponential.

The general form of the pdf of the Weibull distribution is:

$$f(t) = \left(\frac{\beta}{\eta}\right) \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right], \quad t \geq 0. \quad (3.4)$$

The parameters  $\beta > 0$  and  $\eta > 0$  are the shape and the scale parameters respectively. Suppose  $T$  is a Weibull distributed random variable, then its mean (expected value) and variance are defined as:

$$E(T) = \eta \Gamma\left(1 + \frac{1}{\beta}\right) \quad (3.5)$$

$$Var(T) = \eta^2 \left[ \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right] \quad (3.6)$$

where

$$\Gamma(y) = \int_0^\infty x^{y-1} \exp(-x) dx \quad (3.7)$$

is the *Gamma function*.

The cdf,  $F(\cdot)$ , and the hazard function,  $h(\cdot)$ , corresponding to the Weibull distribution, are:

$$F(t) = \int_0^t f(x) dx = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^\beta\right], \quad (3.8)$$

$$h(t) = \frac{f(t)}{1 - F(t)} = \left(\frac{\beta}{\eta}\right) \left(\frac{t}{\eta}\right)^{\beta-1}. \quad (3.9)$$

For a special case, when  $\beta = 1$ , Weibull distribution turns into an exponential distribution with  $\lambda = 1/\eta$ . As a matter of fact, the Weibull distribution can be characterized as a generalization of the exponential distribution. The hazard function,  $h(\cdot)$ , of the Weibull distribution is

- decreasing, for  $\beta < 1$ ,
- increasing, for  $\beta > 1$ ,
- constant (exponential distribution), for  $\beta = 1$ .

The Weibull distribution can be applied for a wide range of models, since the corresponding hazard function is flexible and can cover a wide range of hazard function shapes (see figure 3.2).

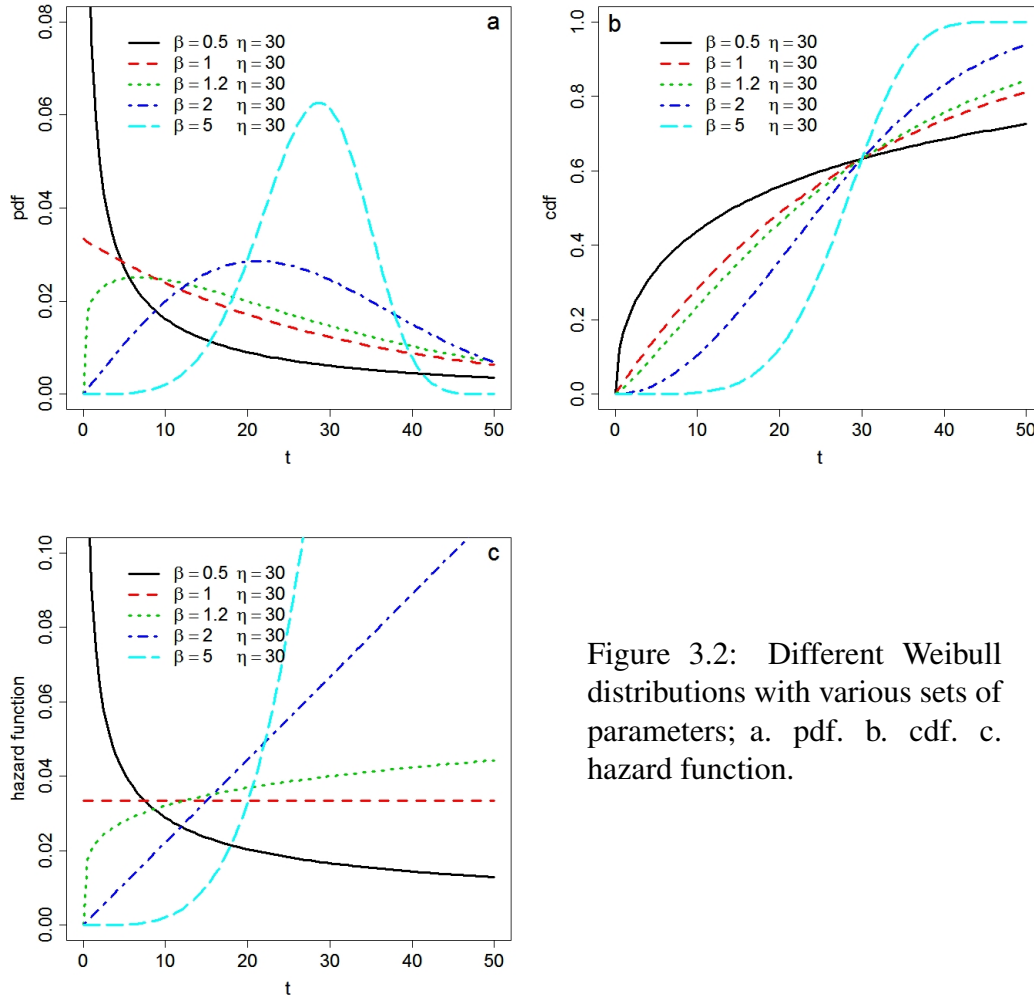


Figure 3.2: Different Weibull distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function.

### 3.4.2 Gamma distribution

The Gamma distribution is a two-parameter, skewed to the right, continuous probability distribution, with scale parameter  $\theta > 0$  and shape parameter  $k > 0$ . If  $k$  is an integer then the distribution represents the sum of  $k$  independent exponentially distributed random variables, each of which has a mean of  $\theta$  (which is equivalent to a rate parameter of  $\theta^{-1}$ ).

The Gamma distribution is frequently used as a probability model for waiting times; for instance, in life testing, the waiting time until death is a random variable that is frequently modeled with a Gamma distribution. Figure 3.3 shows different shapes of the Gamma distribution.

The Gamma distribution has the pdf

$$f(t) = t^{k-1} \frac{\exp\left(-\frac{t}{\theta}\right)}{\theta^k \Gamma(k)}, \quad t \geq 0. \quad (3.10)$$

The corresponding mean and variance of a Gamma distributed random variable

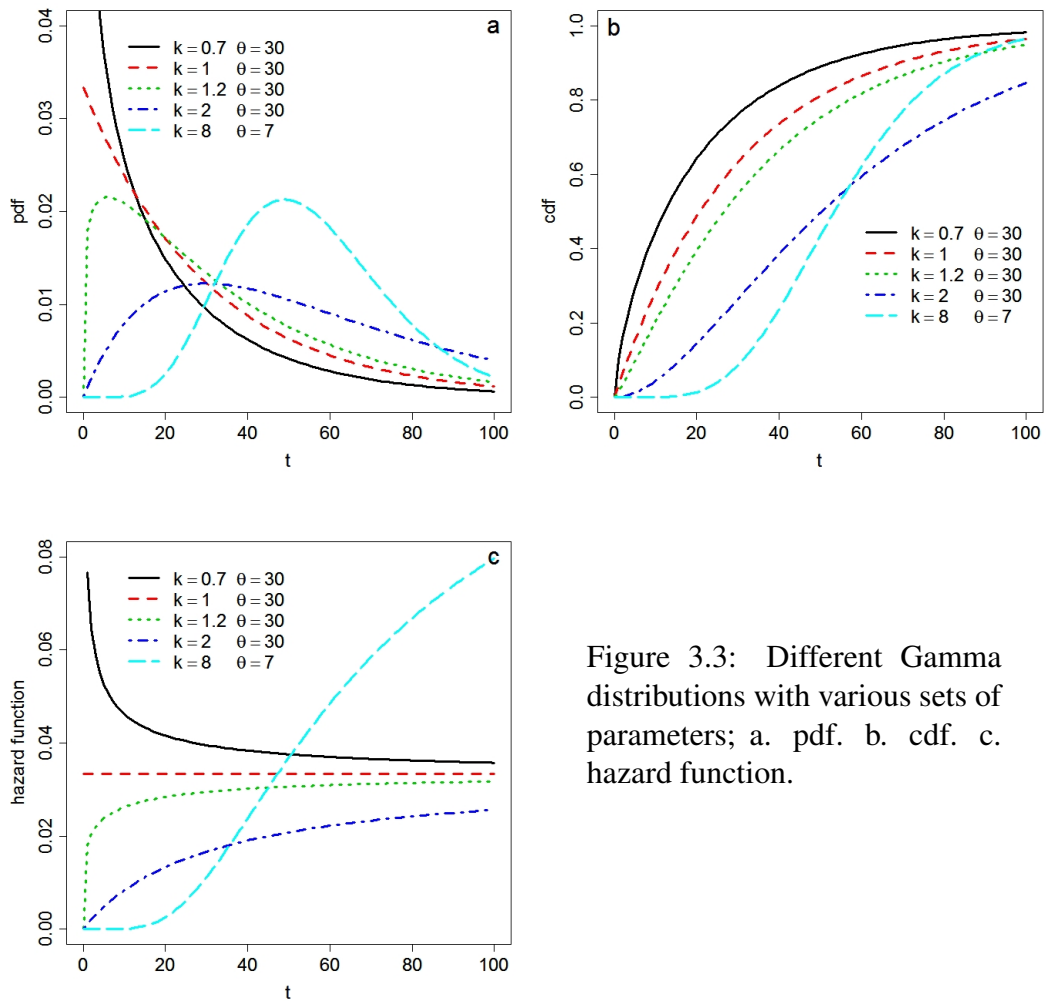


Figure 3.3: Different Gamma distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function.

$T$  are:

$$E(T) = k\theta \quad (3.11)$$

$$Var(T) = k\theta^2. \quad (3.12)$$



The corresponding cdf and hazard function are:

$$F(t) = \frac{\gamma(k, t/\theta)}{\Gamma(k)}, \quad (3.13)$$

$$h(t) = \frac{t^{k-1} \exp(-t/\theta)}{\theta^k (\Gamma(k) - \gamma(k, t/\theta))} = \frac{t^{k-1} \exp(-t/\theta)}{\theta^k \Gamma(k, t/\theta)} \quad (3.14)$$

where

$$\gamma(k, t) = \int_0^t x^{k-1} \exp(-x) dx \quad \text{and} \quad (3.15)$$

$$\Gamma(k, t) = \int_t^\infty x^{k-1} \exp(-x) dx \quad (3.16)$$

are the *lower incomplete Gamma* and *upper incomplete Gamma* functions, respectively.

In a special case, with  $k = 1$ , while  $\Gamma(k = 1) = 1$ , the Gamma distribution turns into an exponential distribution with  $\lambda = 1/\theta$ .

### 3.4.3 Lognormal distribution

The Lognormal distribution is a single-tailed, skewed to the right probability distribution of any random variable whose logarithm is normally distributed (figure 3.4). Given  $X$  as a random variable with a normal distribution, random variable  $Y = \exp(X)$  will be Lognormally distributed; likewise, if  $Y$  is Lognormally distributed, then  $\log(Y)$  will be normally distributed<sup>1</sup>.

The Lognormal distribution has been generally used to model the lives of units whose failure modes are of a fatigue-stress nature. Since this includes most, if not all, mechanical systems, the Lognormal distribution has widespread application. Consequently, the Lognormal distribution is a good companion to the Weibull distribution when modeling these types of units. The Lognormal distribution has the pdf:

$$f(t) = \left( \frac{1}{t\sigma\sqrt{2\pi}} \right) \exp \left[ -\frac{(\ln t - \mu)^2}{2\sigma^2} \right], \quad t \geq 0, \quad (3.17)$$

where the parameters  $\mu$  and  $\sigma > 0$  are the mean and standard deviation of the logarithm of the corresponding random variable<sup>2</sup>. The degree of skewness of the

<sup>1</sup>The base of the logarithmic function does not matter: if  $\log_a(Y)$  is normally distributed, then so is  $\log_b(Y)$ , for any two positive numbers  $a, b \neq 1$ .

<sup>2</sup>Parameter  $\mu$  can be considered as scale parameter, parameter  $\sigma$  as shape parameter.

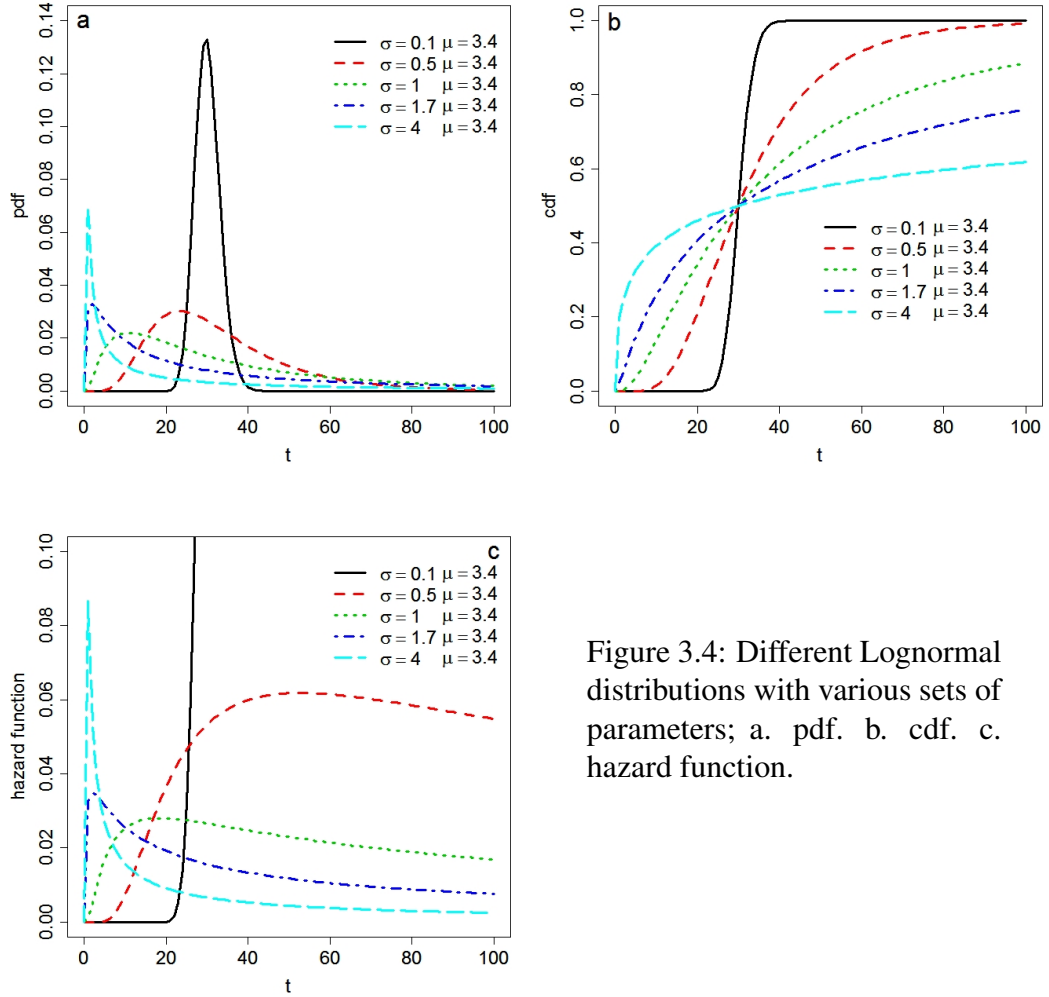


Figure 3.4: Different Lognormal distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function.

Lognormal distribution increases as  $\sigma$  increases for a fixed  $\mu$ , or as  $\mu$  increases for a fixed  $\sigma$  (see figure 3.4).

The mean and the variance of a Lognormally distributed random variable  $T$  are:

$$E(T) = \exp\left(\mu + \frac{\sigma^2}{2}\right), \quad (3.18)$$

$$Var(T) = [\exp(\sigma^2) - 1] \exp(2\mu + \sigma^2). \quad (3.19)$$

The cdf and hazard function corresponding to the Lognormal distribution are:

$$F(t) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\ln(t) - \mu}{\sigma\sqrt{2}} \right) = \Phi \left( \frac{\ln(t) - \mu}{\sigma} \right), \quad (3.20)$$

$$h(t) = \frac{\frac{1}{t\sigma} \phi \left( \frac{\ln(t) - \mu}{\sigma} \right)}{\Phi \left( \frac{\mu - \ln(t)}{\sigma} \right)}, \quad (3.21)$$

where

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dx \quad (3.22)$$

is the *error function*, and the functions  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the pdf and the cdf of the standard normal distribution (a normal distribution with mean = 0 and variance = 1).

### 3.4.4 Inverse Gaussian distribution

In probability theory, the inverse Gaussian distribution (also known as the Wald distribution) is a two-parameter family of continuous, skewed to the right probability distributions in the range  $(0, \infty)$ .

In physics, the inverse Gaussian distribution is known as the Brownian Passage Time (BPT) distribution. This term has its grounds in a relationship with the Brownian motion as a stochastic process. Suppose the stochastic process  $X_t$  given by

$$X_0 = 0, \quad X_t = \nu t + \sigma W_t \quad (3.23)$$

where  $W_t$  is a standard Brownian motion and  $\nu > 0$ . This process is a Brownian motion with drift  $\nu$ . Then the first passage time for a fixed level  $\alpha > 0$  by  $X_t$  is distributed according to an inverse Gaussian as:

$$T_\alpha = \inf\{t > 0 \mid X_t = \alpha\} \sim IG \left( \frac{\alpha}{\nu}, \frac{\alpha^2}{\sigma^2} \right). \quad (3.24)$$

Matthews et al. (2002) discusses a probability model for rupture times on a recurrent earthquake source. By adding Brownian perturbations to steady tectonic loading, a stochastic load-state process is proposed. This load-state process is a Brownian relaxation oscillator. Then, the intervals between events (inter-event times) have a BPT distribution that may serve as a temporal model for time-dependent, long-term earthquake forecasting. This distribution has the following noteworthy properties:

- the probability of immediate re-rupture is zero;
- the hazard function increases steadily from zero at  $t = 0$  to a finite maximum near the mean recurrence time and then decreases asymptotically to a quasi-stationary level, in which the conditional probability of an event becomes time-independent.

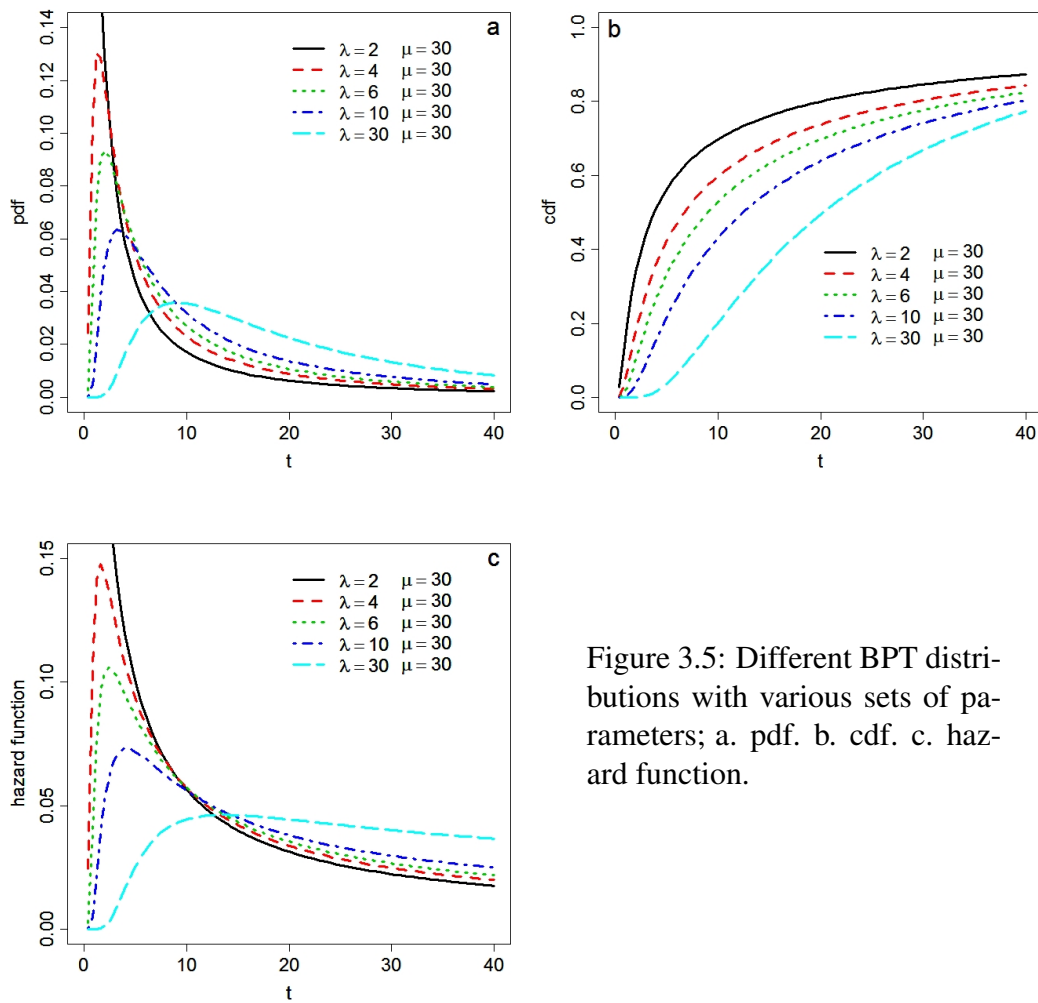


Figure 3.5: Different BPT distributions with various sets of parameters; a. pdf. b. cdf. c. hazard function.

The pdf of the BPT distribution can be written as:

$$f(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left[-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right] \quad (3.25)$$

where the parameters  $\mu$  and  $\lambda$  are the scale and the shape parameters, respectively. The degree of skewness of the BPT distribution increases as the shape parameter  $\lambda$  increases for a fixed scale parameter  $\mu$ , or as  $\mu$  increases for a fixed  $\lambda$ . Figure 3.5 shows different shapes of the BPT distribution.

The mean and the variance of a BPT distributed random variable  $T$  are:

$$E(x) = \mu \quad , \quad Var(T) = \frac{\mu^3}{\lambda}. \quad (3.26)$$

The cdf corresponding to the BPT distribution can be described as:

$$F(t) = \Phi \left( \sqrt{\frac{\lambda}{t}} \left[ \frac{t}{\mu} - 1 \right] \right) + \exp \left( \frac{2\lambda}{\mu} \right) \Phi \left( -\sqrt{\frac{\lambda}{t}} \left[ \frac{t}{\mu} + 1 \right] \right), \quad (3.27)$$

where  $\Phi(\cdot)$  is the cdf of the standard normal distribution.

### 3.4.5 Summary

All the above mentioned distributions; i.e. Weibull, Gamma, Lognormal and BPT, are from the family of exponential distributions and skewed to the right. They have been used in different cases of time-to-failure modeling (Hogg & Craig 1978, Chikara & Folks 1989, Limpert et al. 2001, and Murthy et al. 2004) and will be the base of the statistical modeling in this thesis.

## 3.5 Behavior of distributions; multimodality against unimodality

In several cases of statistical modeling, the data cannot be sufficiently described by a single unimodal distribution. For example, in the case of clustered data, the multimodal distributions can generally fit the data better than unimodal distributions. Therefore, it is important to find out whether the data are multimodally or unimodally distributed. Figure 3.6 shows the a unimodal and a multimodal pdf.

A multimodal (or mixed) pdf has the following general form

$$f(t) = \sum_{j=1}^g \pi_j f_j(t), \quad \sum_{j=1}^g \pi_j = 1 \quad (3.28)$$

where  $g$  is the number of modes (i.e. the number of unimodal pdfs involved in the multimodal pdf  $f(\cdot)$ ),  $f_j(\cdot)$  is the  $j$ -th unimodal pdf involved in the multimodal pdf  $f(\cdot)$ , and  $\pi_j$  is the percental contribution of the  $j$ -th pdf  $f_j(\cdot)$  to the multimodal pdf. If  $g > 1$ , then  $f(\cdot)$  is a multimodal pdf.

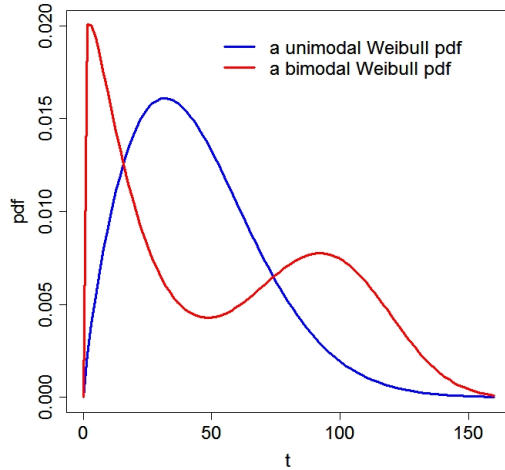


Figure 3.6: An example of multimodal and unimodal pdfs. The blue curve corresponds to a random variable  $T$ , unimodally distributed as  $T \sim \text{Weibull}(\beta = 1.8, \eta = 50)$ . The red curve is extracted from a random variable  $X$ , multimodally distributed as  $X \sim 0.5\text{Weibull}(\beta = 1.1, \eta = 20) + 0.5\text{Weibull}(\beta = 4.0, \eta = 100)$ .

### 3.5.1 Multimodality test of a population

Efron & Tibshirani (1993) propose a bootstrap method to test the multimodality of a population. In this method, the histogram of the data is monitored. However, by just considering the histogram, it will be difficult to confirm if the data are unimodal or multimodal, because the histogram is not smooth. In order to obtain a smoother estimate, a *Gaussian Kernel estimate* has been proposed. Denoting the observed (calculated) inter-event times by  $t_1, t_2, \dots, t_n$ , a Gaussian Kernel density can be estimated by

$$\hat{f}(t; h) = \frac{1}{nh} \sum_{i=1}^n \phi\left(\frac{t - t_i}{h}\right) \quad (3.29)$$

where  $\phi(t)$  is the standard normal density  $(1/\sqrt{2\pi}) \exp(-t^2/2)$ . The parameter  $h$  is called the *band width* and determines the amount of *smoothing* that is applied to the data. Larger values of  $h$  produce a smoother density estimate. One can think of equation 3.29 as adding up  $n$  little Gaussian density curves centered at each point  $t_i$ , each having standard deviation  $h$ . Clearly, any opinion drawn from the data depends strongly on the chosen value of  $h$ . Then, in order to approach the problem in terms of hypothesis testing, there is a natural way to choose  $h$ , since as  $h$  increases, the number of modes in a Gaussian Kernel density estimate is non-decreasing. Now consider testing

$$H_0 : \text{number of modes} = 1 \quad (3.30)$$

versus number of modes  $> 1$ . Since the number of modes decreases as  $h$  increases, there is a smallest value of  $h$  such that  $\hat{f}(t; h)$  has one mode. Call this  $\hat{h}_1$ . It seems reasonable to use  $\hat{f}(t; \hat{h}_1)$  as the estimated null distribution for the test of  $H_0$ . In a sense, it is the density estimate "closest" to the data that is consistent with  $H_0$ . "Closest" means here that it uses the least amount of smoothing (smallest value of  $h$ ) among all estimates with one mode.

There is one small adjustment which has to be made to  $\hat{f}(\cdot; \hat{h}_1)$ . Equation 3.29 artificially increases the variance of the estimate (see Efron & Tibshirani, 1993), so it should be rescaled to have variance equal to the sample variance. Denote the rescaled estimate by  $\hat{g}(\cdot; \hat{h}_1)$ .

Now, the test statistic can be selected. A natural choice is  $\hat{h}_1$ , the smallest band width producing a density estimate with one mode. A large value of  $\hat{h}_1$  indicates that a large amount of smoothing must be applied to create an estimate with one mode and is therefore evidence against  $H_0$ .

Putting all of this together, the bootstrap hypothesis test for  $H_0$  (number of modes = 1) is based on the achieved significance level

$$\text{ASL}_{\text{boot}} = \text{Prob}_{\hat{g}(\cdot; \hat{h}_1)} \left\{ \hat{h}_1^* > \hat{h}_1 \right\} \quad (3.31)$$

where  $\hat{h}_1$  is fixed at its observed value, the bootstrap sample  $t_1^*, t_2^*, \dots, t_n^*$  is drawn from  $\hat{g}(\cdot; \hat{h}_1)$  and  $\hat{h}_1^*$  is the smallest value of  $h$  producing a density estimate with one mode from the bootstrap data  $t_1^*, t_2^*, \dots, t_n^*$ .

The  $y_1^*, y_2^*, \dots, y_n^*$  are sampled from a smooth estimate of the data, which is called *smooth bootstrap*, by replacement from  $t_1, t_2, \dots, t_n$ . Then,  $t_1^*, t_2^*, \dots, t_n^*$  are set as:

$$t_i^* = \bar{y}^* + (1 + \hat{h}_1^2 / \hat{\sigma}^2)^{-1/2} (y_i^* - \bar{y}^* + \hat{h}_1 \epsilon_i); \quad i = 1, 2, \dots, n \quad (3.32)$$

where  $\bar{y}^*$  is the mean of  $y_1^*, y_2^*, \dots, y_n^*$ ,  $\hat{\sigma}^2$  is the plug estimate of variance of the data and  $\epsilon_i$  are standard normal random variables. The factor  $(1 + \hat{h}_1^2 / \hat{\sigma}^2)^{-1/2}$  scales the estimate so that its variance is approximately  $\hat{\sigma}^2$ .

A summary of the steps of the algorithm, applied in this thesis, can be described as following:

1. Draw  $B$  bootstrap samples of size  $n$  from  $\hat{g}(\cdot; \hat{h}_1)$ .
2. For each bootstrap sample compute  $\hat{h}_1^*$  the smallest band width that produces a density estimate with one mode. Denote the  $B$  values of  $\hat{h}_1^*$  by  $\hat{h}_1^*(1), \dots, \hat{h}_1^*(B)$ .
3. Approximate  $\text{ASL}_{\text{boot}}$  by
$$\widehat{\text{ASL}}_{\text{boot}} = \# \left\{ \hat{h}_1^*(b) \geq \hat{h}_1 \right\} / B.$$

## 3.6 Modified Weighted Maximum Likelihood Estimation method

The maximum likelihood estimation (MLE) method is a commonly used statistical method, applied to fit a statistical model to desired data, and provides estimates for the model's parameters. The MLE was recommended, analyzed and vastly popularized by Fisher between 1912 and 1922, although it had been used earlier by Gauss, Laplace, Thiele, and F. Y. Edgeworth (Aldrich 1997).

The method of MLE can be applied to many well-known estimation problems in statistics. For example, suppose the heights of Iranians are studied. A sample of some number of Iranians, but not the entire population, is available, and their heights are recorded. Further, the heights are assumed to be normally distributed with some unknown mean and variance. The sample mean is then the maximum likelihood estimator of the population mean, and the sample variance is a close approximation to the maximum likelihood estimator of the population variance. For a fixed set of data and an underlying probability model, the MLE picks the values of the model parameters that make the data "more likely" than any other values of the parameters would make them. The MLE gives a unique and easy estimation in the case of the normal distribution and many other problems. If a uniform prior distribution is assumed for the parameters, the maximum likelihood estimate coincides with the most probable values thereof.

### 3.6.1 The general approach

Suppose  $T$  is a continuous random variable with pdf  $f_{\Theta}(t)$ , where  $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$  is the set of  $k$  unknown parameters which need to be estimated. Suppose a set of  $N$  independent observations,  $t_1, t_2, \dots, t_N$ , of the random variable  $T$ . Then, the likelihood function is given by the following product:

$$L(t_1, t_2, \dots, t_N | \Theta) = \prod_{i=1}^N f_{\Theta}(t_i). \quad (3.33)$$

The logarithmic likelihood function is given by:

$$\Lambda(t_1, t_2, \dots, t_N | \Theta) = \ln [L(t_1, t_2, \dots, t_N | \Theta)] = \sum_{i=1}^N \ln [f_{\Theta}(t_i)]. \quad (3.34)$$

The maximum likelihood estimators of  $\theta_1, \theta_2, \dots, \theta_k$  are obtained by maximizing  $L$  or  $\Lambda$ .

By maximizing  $\Lambda$ , which is much easier to work with than  $L$ , the ML-estimators



of  $\theta_1, \theta_2, \dots, \theta_k$  are the simultaneous solutions of  $k$  equations such that:

$$\frac{\partial(\Lambda)}{\partial(\theta_j)} = 0, \quad j = 1, 2, \dots, k. \quad (3.35)$$

### 3.6.2 Case of observed censored times

In the case of time-to-failure modeling, the time since the last event is called the censored time. In order to use this information for the MLE method, the following modified MLE approach is proposed in this thesis:

$$L = \prod_{i=1}^N f_{\Theta}(t_i) * S_{\Theta}(t_c) \quad (3.36)$$

$$\Lambda = \ln L = \sum_{i=1}^N \ln [f_{\Theta}(t_i)] + \ln [S_{\Theta}(t_c)] \quad (3.37)$$

where  $t_c$  is the censored time, and  $S_{\Theta}(\cdot)$  is the *survivor function* or *reliability function* of the random variable  $T$  with  $S(t) = 1 - F(t) = 1 - \int_{-\infty}^t f(x) dx = \int_t^{\infty} f(x) dx$ , where  $F(\cdot)$  is the corresponding cdf.

### 3.6.3 Weighted MLE with censored times

In some cases of time-to-failure modeling, for example in cases of data with different levels of uncertainties, the roles of different observed inter-event times in the likelihood function are not equivalent. In these cases, the ML-function can be discriminated using a method of weighting.

In this thesis, where historical earthquake data are used, the use of weights in the MLE is forced for two reasons. The first is the intrinsic uncertainty of historical earthquake data. Uncertainties in the earthquake parameters, especially for the older ones, are neither recognizable, nor calculable. The second reason is the fact that earthquake rates probably change with time. Therefore, the estimation process requires higher weights for earthquakes occurring in the recent past.

Applying weights to the relations of the MLE, they change to

$$L = \prod_{i=1}^N f_{\Theta}(t_i)^{w_i} * S_{\Theta}(t_c)^{w_c} \quad (3.38)$$

$$\Lambda = \ln L = \sum_{i=1}^N w_i \ln [f_{\Theta}(t_i)] + w_c \ln [S_{\Theta}(t_c)] \quad (3.39)$$

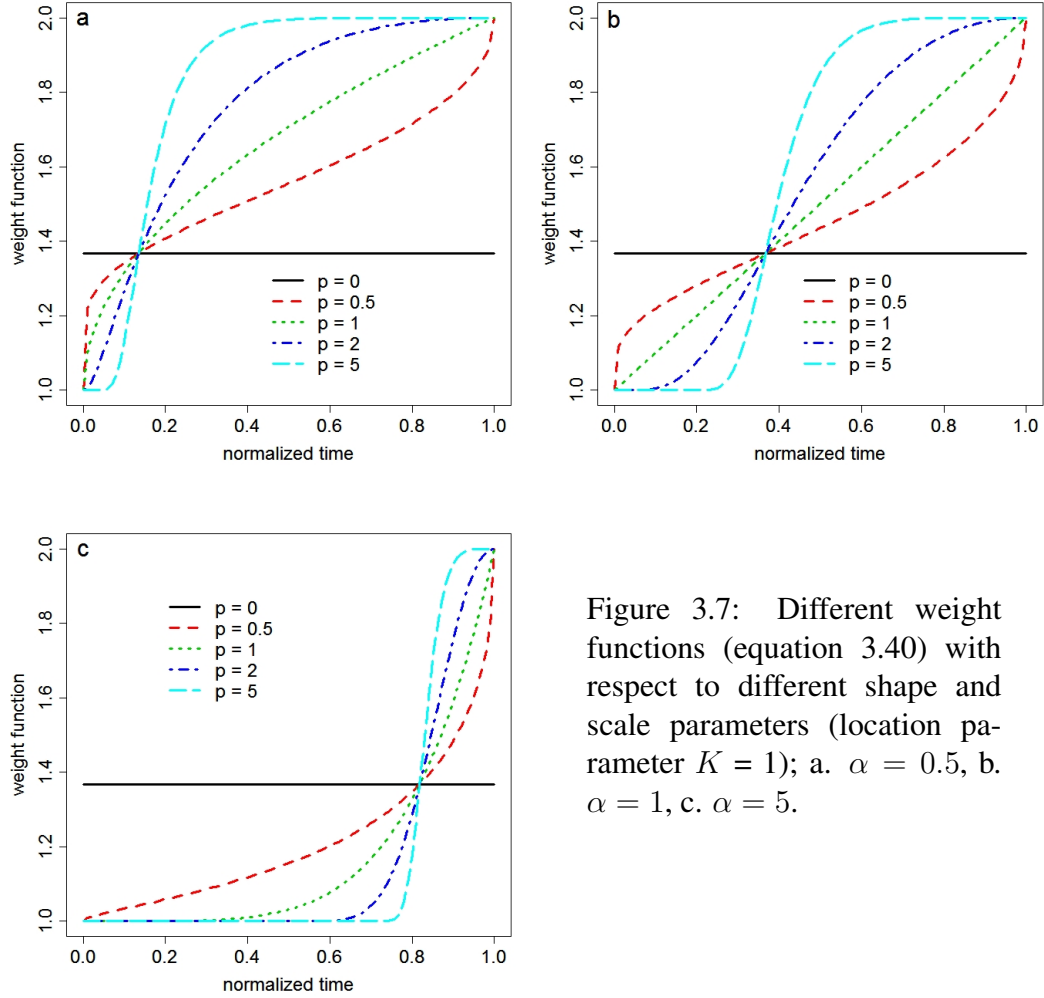


Figure 3.7: Different weight functions (equation 3.40) with respect to different shape and scale parameters (location parameter  $K = 1$ ); a.  $\alpha = 0.5$ , b.  $\alpha = 1$ , c.  $\alpha = 5$ .

where  $W = \{w_1, w_2, \dots, w_n\}$  is the set of corresponding weights assigned to the inter-event times  $t_1, t_2, \dots, t_n$ , and  $w_c$  is the weight assigned to the censored time  $t_c$ .

In this thesis, the following function is suggested to calculate the weights,

$$\omega(x) = \exp(-|\alpha \ln(x)|^p) + K, \quad 0 \leq x \leq 1 \quad (3.40)$$

where  $\alpha \geq 0$ ,  $p \geq 0$  and  $K \geq 0$  are the scale, shape and location parameters respectively. The corresponding weights  $w_i$  can be defined as:

$$w_i = \omega\left(\frac{ot_i}{\max_i(ot_i)}\right) = \exp\left(-\left|\alpha \ln\left(\frac{ot_i}{\max_i(ot_i)}\right)\right|^p\right) + K \quad (3.41)$$

where  $ot_i$  is the occurrence time of the  $i$ -th event. The reason for considering such a function is that, the function in equation 3.40 covers different possible shapes of weight functions and is not decreasing with time (figure 3.7). The parameters of the weight function in equation 3.40 can be imported in the model as known parameters, or estimated via the MLE process like other parameters in the likelihood function.

### 3.6.4 Method of solving the MLE problem

#### Case of unimodality

Attaching the censored time to the likelihood function makes the maximizing problem of the MLE method more complex. In this thesis, the numerical method of *Nonlinear Conjugate Gradients with Newton-Raphson and Fletcher-Reeves* (Shewchuk 1994) is applied in order to solve the problem of maximization of the likelihood function. Given a function  $f$ , a starting value  $x$ , a maximum number of conjugate gradients iterations  $i_{max}$ , a conjugate gradients error tolerance  $\epsilon_{CG} < 1$ , a maximum number of Newton-Raphson iterations  $j_{max}$ , and a Newton-Raphson error tolerance  $\epsilon_{NR} < 1$ ; the algorithm is defined as:

Algorithm 3.1: Nonlinear conjugate gradients with Newton-Raphson and Fletcher-Reeves steps.

```

 $i \leftarrow 0$ 
 $k \leftarrow 0$ 
 $r \leftarrow -f'(x)$ 
 $d \leftarrow r$ 
 $\delta_{new} \leftarrow r^T r$ 
 $\delta_0 \leftarrow \delta_{new}$ 

while  $i < i_{max}$  and  $\delta_{new} > \epsilon_{CG}^2 \delta_0$  do
   $i \leftarrow 0$ 
   $\delta_d \leftarrow d^T d$ 
  do
     $\alpha \leftarrow -\frac{[f'(x)]^T d}{d^T f''(x) d}$ 
     $x \leftarrow x + \alpha d$ 
     $j \leftarrow j + 1$ 
  while  $j < j_{max}$  and  $\alpha^2 \delta_d > \epsilon_{NR}^2$ 
   $r \leftarrow -f'(x)$ 
   $\delta_{old} \leftarrow \delta_{new}$ 
   $\delta_{new} \leftarrow r^T r$ 
   $\beta \leftarrow \frac{\delta_{new}}{\delta_{old}}$ 
   $d \leftarrow r + \beta d$ 
   $k \leftarrow k + 1$ 
  If  $k = n$  or  $r^T d \leq 0$ 
     $d \leftarrow r$ 
     $k \leftarrow 0$ 
   $i \leftarrow i + 1$ 

```

where  $T$  means "transpose". This algorithm terminates when the maximum number of iterations  $i_{max}$  has been exceeded, or when  $\|r_{(i)}\| \leq \epsilon_{CG} \|r_{(0)}\|$ . Each Newton-Raphson iteration adds  $\alpha d$  to  $x$ ; the iterations are terminated when each update  $\alpha d$  falls below a given tolerance ( $\|\alpha d\| \leq \epsilon_{NR}$ ), or when the number of sub-iterations exceeds  $j_{max}$ . A fast inexact line search can be accomplished by using a small  $j_{max}$  and/or by approximating the Hessian  $f''(x)$  with its diagonal. In Appendix A the log-likelihood functions of five distributions and their first and second derivatives are presented in order to determine  $f'(x)$  and the Hessian  $f''(x)$ . Nonlinear conjugate gradients method is restarted (by setting  $d \leftarrow r$ ) whenever a search direction is computed that is not a descent direction. It is also restarted once every  $n$  iterations, to improve convergence for small  $n$ .

The calculation of  $\alpha$  may result in a divide-by-zero error. This may occur because the starting point  $x_{(0)}$  is not sufficiently close to the desired optimal point, or because  $f$  is not twice continuously differentiable. In the former case, the solution is to choose a better starting point or a more sophisticated line search.

### Case of multimodality

The likelihood function in case of multimodal distributions is more complex. Simultaneous estimates of the distribution parameters  $\theta_1, \dots, \theta_k$  in equation 3.34, and the proportion parameters  $\pi_1, \dots, \pi_g$  in equation 3.28, is too complex, if possible at all. Therefore, a method of *Expectation Maximization* (EM) is applied (Dempster et al. 1977, McLachlan & Krishnan 1996, Cherkassky & Mulier 1998, McLachlan & Peel 2000, Tsai et al. 2001, Dellaert 2002, Borman 2004, Molenberghs & Verbeke 2005).

EM is an effective iterative procedure to compute the ML-estimate in the presence of missing or hidden data as well as to estimate some unknown parameters for a given dataset. EM algorithms are used in statistics for finding maximum likelihood estimates of parameters in probabilistic models, where the model depends on unobserved latent variables. EM is an iterative method, which alternates between performing an expectation step (E-step), which computes an expectation of the log-likelihood with respect to the current estimate of the distribution for the latent variables, and a maximization step (M-step), which computes the parameters which maximize the expected log-likelihood found on the E-step. These parameters are then used to determine the distribution of the latent variables in the next E-step. In the E-step, the missing data, or multimodality parameters, are estimated given the observed data and current estimate of the model parameters. This is achieved using a conditional expectation. In the M-step, the likelihood function is maximized under the assumption that the missing data, or the multimodality

parameters, are known. Suppose a mixed pdf as:

$$f(t | \psi, \Xi) = \sum_{j=1}^g \pi_j f_j(t; \Theta_j), \quad \sum_{j=1}^g \pi_j = 1 \quad (3.42)$$

where  $\Theta_j$  are the sets of parameters of  $f_j(\cdot; \Theta_j)$ ,  $j = 1, \dots, g$ ,  $\Xi = \bigcup_{j=1}^g \Theta_j$  is the set of distribution parameters, and  $\psi = (\pi_1, \dots, \pi_{g-1})^T$  is corresponding to  $g - 1$  mixing proportions. The corresponding log-likelihood function can be written as:

$$\begin{aligned} \Lambda(t_1, \dots, t_N | \Theta, \Xi) &= \sum_{i=1}^N \ln [f(t_i | \psi, \Xi)] \\ &= \sum_{i=1}^N \ln \left[ \sum_{j=1}^g \pi_j f_j(t_i; \Theta_j) \right]. \end{aligned} \quad (3.43)$$

Then the EM algorithm can be defined as follows:

**Algorithm 3.2: Expectation Maximization (EM) algorithm.**

1. Observed random sample obtained from mixture density (equation 3.42)

$$\begin{aligned} y &= (u_1^T, \dots, u_N^T)^T \\ u_i &= (t_i, f(t_i | \psi, \Xi)) = (t_i, \sum_{j=1}^g \pi_j f_j(t_i; \Theta_j)). \end{aligned}$$

2. Introduce  $z = (z_1^T, \dots, z_N^T)^T$  as the unobservable or missing data, where  $z_i$  are  $g$ -dimensional vectors of zero-one indicator variables.
3. Define  $z_{ij} = (z_i)_j$ ,  $z_{ij}$  is equal to one, if  $u_i$  arose from  $j$ -th component.
4. **E-step.** Calculate the current conditional expectation of  $Z_{ij}$ , given the observed data  $y$

$$\begin{aligned} E(Z_{ij} | y, \psi^{(k)}, \Xi) &= P \{ Z_{ij} = 1 | y, \psi^{(k)}, \Xi \} = z_{ij}^{(k)} \\ z_{ij}^{(k)} &= \frac{\pi_j^{(k)} f_j(w_i; \Theta_j)}{f(w_i | \psi^{(k)}, \Xi)} = \frac{\pi_j^{(k)} f_j(w_i; \Theta_j)}{\sum_{l=1}^g \pi_l^{(k)} f_l(w_i; \Theta_l)} \end{aligned}$$

5. **M-step**

$$\pi_j^{(k+1)} = \frac{\sum_{i=1}^N z_{ij}^{(k)}}{n}, \quad j = 1, \dots, g.$$

The convergence of the EM algorithm is discussed in detail by McLachlan & Krishnan (1996).

Considering censored time (equations 3.36 and 3.37), as well as the weights (equations 3.38 and 3.39) in the MLE process, the likelihood function and the log-likelihood function in equation 3.43 will change to

$$L = \prod_{i=1}^N f(t_i | \psi, \Xi)^{w_i} * S(t_c | \psi, \Xi)^{w_c}, \quad (3.44)$$

$$\Lambda = \sum_{i=1}^N w_i \ln \left[ \sum_{j=1}^g \pi_j f_j(t_i; \Theta_j) \right] + w_c \ln \left[ \sum_{j=1}^g \pi_j S_j(t_c; \Theta_j) \right]. \quad (3.45)$$

Then the EM algorithm will change to algorithm 3.3.

**Algorithm 3.3:** EM algorithm for weighted MLE with censored time.

1. Observed random sample obtained from mixture density (equation 3.42)

$$\begin{aligned} y &= (u_1^T, \dots, u_N^T, u_c^T)^T \\ u_i &= (t_i, f(t_i | \psi, \Xi)) = (t_i, \sum_{j=1}^g \pi_j f_j(t_i; \Theta_j)) \quad , \quad i = 1, \dots, N \\ u_c &= (t_c, S(t_c | \psi, \Xi)) = (t_c, \sum_{j=1}^g \pi_j S_j(t_c; \Theta_j)). \end{aligned}$$

2. Introduce  $z = (z_1^T, \dots, z_N^T, z_c^T)^T$  as the unobservable or missing data, where  $z_1, \dots, z_g, z_c$  are  $g$ -dimensional vectors of zero-one indicator variables.
3. Define  $z_{ij} = (z_i)_j$ ,  $z_{ij}$  is equal to one, if  $u_i$  arose from  $j$ -th component, and  $z_{cj} = (z_c)_j$ ,  $z_{cj}$  is equal to one, if  $u_c$  arose from  $j$ -th component.
4. **E-step.** Calculate the current conditional expectation of  $Z_{ij}$  and  $Z_{cj}$  given the observed data  $y$

$$E(Z_{ij} | y, \psi^{(k)}, \Xi) = P \{ Z_{ij} = 1 | y, \psi^{(k)}, \Xi \} = z_{ij}^{(k)} \quad i \in \{1, \dots, N, c\}$$

$$\begin{aligned} z_{ij}^{(k)} &= \frac{\pi_j^{(k)} f_j(u_i; \Theta_j)}{f(u_i | \psi^{(k)}, \Xi)} = \frac{\pi_j^{(k)} f_j(u_i; \Theta_j)}{\sum_{l=1}^g \pi_l^{(k)} f_l(u_i; \Theta_l)} \quad i \in \{1, \dots, N\} \\ z_{cj}^{(k)} &= \frac{\pi_j^{(k)} S_j(u_c; \Theta_j)}{S(u_c | \psi^{(k)}, \Xi)} = \frac{\pi_j^{(k)} S_j(u_c; \Theta_j)}{\sum_{l=1}^g \pi_l^{(k)} S_l(u_c; \Theta_l)} \end{aligned}$$

5. **M-step**

$$\pi_j^{(k+1)} = \frac{\sum_{i=1}^N z_{ij}^{(k)} + z_{cj}^{(k)}}{n+1} \quad , \quad j = 1, \dots, g.$$

## 3.7 Selection of the best model

In this thesis, two methods have been applied in order to choose the best distribution. These methods include a method of *information criterion* and a *goodness-of-fit test*.

### 3.7.1 Methods of information criterion

These methods are based on the comparison among recalculated (or adjusted) measures of the maximum log-likelihoods calculated using a MLE method. The most commonly applied information criteria in statistical analyses are the *Akaike Information Criterion* (AIC, Akaike 1974) and the *Bayesian Information Criterion* (BIC, Volinsky & Raftery 1999).

The AIC chooses the best model based on the minimum value of

$$\text{AIC} = 2k - 2 \ln(L) = 2k - 2\Lambda \quad (3.46)$$

where  $k$  is the number of parameters and  $L$  and  $\Lambda$  are the maximum likelihood and the maximum log-likelihood of ML-estimates with respect to a model (distribution). Parameter  $N$  (i.e. the number of the observed data) does not play any role in the AIC.

The BIC chooses the best model based on the minimum value of the equation

$$\text{BIC} = k \ln(N) - 2 \ln(L) = k \ln(N) - 2\Lambda. \quad (3.47)$$

The BIC criterion results in consistent outcomes (see Csiszár & Shields 1999). In this thesis the BIC is used.

### 3.7.2 Methods of goodness-of-fit test

In order to find out, whether the models chosen by the information criteria can sufficiently explain the data, the goodness-of-fit tests will be applied. There are different goodness-of-fit methods divided into two general categories.

The first category includes the tests which concentrate on the average of the data. An example from this category is the *Anderson-Darling* test (AD-test, Anderson & Darling 1952). The AD-test hypotheses are defined as:

$$\begin{aligned} H_0 &= \text{The data follow a specified distribution} \\ H_a &= \text{The data do not follow the specified distribution} \end{aligned}$$

The AD-test statistic is defined as:

$$A_N^2 = -N - S \quad (3.48)$$

$$S = \sum_{i=1}^N \frac{2i-1}{N} [\ln(F(t_i)) + \ln(1 - F(t_{N+1-i}))] \quad (3.49)$$

where  $F$  is the cdf corresponding to the specified distribution,  $t_i$  are the *ordered* data, and  $N$  is the size of the sample.

The second category includes the tests which concentrate on the details of the data. The *Kolmogorov-Smirnov* test (KS-test, Eadie et al. 1971) is in this category. The test hypotheses of the KS-test are the same as for the AD-test. The KS-test statistic is defined as:

$$D_N = \sup_t |F_N(t) - F(t)| \quad (3.50)$$

where  $F_n(\cdot)$  is the empirical cdf for  $N$  i.i.d observations  $t_i$ , defined as:

$$F_N(t) = \frac{1}{N} \sum_{i=1}^N I_{t_i \leq t}. \quad (3.51)$$

$I_{t_i \leq t}$  is the *indicator function*<sup>3</sup>.  $F(\cdot)$  is the theoretical cdf of the distribution being tested, which must be a continuous distribution (i.e., no discrete distributions such as binomial or Poisson).

By the Glivenko-Cantelli theorem (Shorak & Wellner 1986), if the sample comes from the distribution  $F(\cdot)$ , then  $D_N$  converges to zero almost surely. If  $F(\cdot)$  is continuous, then, under the null hypothesis,  $\sqrt{N}D_N$  converges to the *Kolmogorov distribution*. The cdf of the Kolmogorov distribution is given as (Blackman 1956)

$$\Pr(T \leq t) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 t^2} = \frac{\sqrt{2\pi}}{t} \sum_{i=1}^{\infty} e^{-(2i-1)^2 \pi^2 / 8t^2}. \quad (3.52)$$

In this thesis, a modified method of the KS-test has been developed in order to test the goodness-of-fit of different candidate distributions. According to the weighted MLE with censored time, the KS-test has been adjusted such that the empirical cdf in equation 3.51 will change to

$$F_N(t) = \frac{\sum_{\{i:t_i \leq t_c\}} w_i I_{t_i \leq t}}{\sum_{i=1}^N w_i + w_c} I_{t \leq t_c} + \frac{\sum_{\{i:t_i \leq t_c\}} w_i + \sum_{\{i:t_i > t_c\}} w_i I_{t_i \leq t}}{\sum_{i=1}^N w_i} I_{t > t_c} \quad (3.53)$$

where  $t_i$ ,  $t_c$ ,  $w_i$  and  $w_c$  are defined as above.

### 3.8 Confidence intervals of estimated parameters

A classic method to calculate the confidence intervals of model parameters is to use the *Fisher information matrix*, i.e. the Hessian matrix of the log-likelihood

---

<sup>3</sup> $I_{t_i \leq t} = 1$  if  $t_i \leq t$ , and  $I_{t_i \leq t} = 0$  if  $t_i > t$



function. However, the confidence intervals calculated by the Fisher information matrix are not always reliable, especially in cases in which the proportion of the number of observations (data size) divided by the number of parameters is not large enough. In such cases, nonparametric methods, like *Bootstrap confidence intervals*, have been suggested.

In this thesis, the *Bootstrap-t* method (Efron, 1982) is used in order to calculate the confidence intervals. This method is based on a given studentized "pivot"  $\mathfrak{R}_n = (\hat{\theta}_n - \theta)/\hat{\sigma}_n$ , where  $\hat{\theta}_n$  is an estimator of  $\theta$ ; i.e. the designated parameter, and  $\hat{\sigma}_n^2$  is a variance estimator for  $\hat{\theta}_n$ . If the distribution  $G_n$  of  $\mathfrak{R}_n$  is unknown, then it can be estimated by the bootstrap estimator  $G_{BOOT}$  defined by

$$G_{BOOT}(x) = P_*\{\mathfrak{R}_n^* \leq x\} \quad (3.54)$$

where  $\mathfrak{R}_n^* = (\hat{\theta}_n^* - \hat{\theta}_n)/\hat{\sigma}_n^*$ , and  $\hat{\theta}_n^*$  and  $\hat{\sigma}_n^*$  are bootstrap analogs of  $\hat{\theta}_n$  and  $\hat{\sigma}_n$ , respectively. The resulting confidence interval for  $\theta$  is

$$C_n = \left[ \hat{\theta}_n - \hat{\sigma}_n G_{BOOT}^{-1} \left( 1 - \frac{\alpha}{2} \right), \hat{\theta}_n - \hat{\sigma}_n G_{BOOT}^{-1} \left( \frac{\alpha}{2} \right) \right], \quad (3.55)$$

which has been called the bootstrap-t confidence interval for  $\theta$  with level  $1 - \alpha$ , where  $P\{\theta \in C_n\} \geq 1 - \alpha$ ,  $0 < \alpha < 1$ .

### 3.9 Earthquake inter-event time modeling approach

Considering the methods mentioned in this chapter, the following approach will be adapted for inter-event time modeling:

#### 1. Multimodality test of inter-event times

#### 2. Parameter estimation

For five different distributions: Exponential (Poisson process), Weibull, BPT, Gamma and Lognormal, as well as for bimodal mixtures of these distributions in the case of multimodal inter-event time distributions

- Using the modified MLE method for the unimodal cases
- Using the Expectation Maximization and the modified MLE method for the multimodal cases

#### 3. The choice of the "best" model using

- Bayesian Information Criterion
- Kolmogorov-Smirnov goodness-of-fit tests

#### 4. Calculation of the confidence intervals

# Chapter 4

## Seismicity in the Dead Sea Fault Zone

### 4.1 Dead Sea Fault Zone and seismicity dataset

The Dead Sea Fault Zone (DSFZ, figure 4.1) is a major left-lateral strike-slip fault zone (e.g. Garfunkel et al. 1981) that accommodates the relative motion between Africa and Arabia (Tapponnier 1977, Courtillot et al. 1987, Le Pichon & Gaulier 1988, Salamon 1996), connecting a region of extension in the Red Sea with the Taurus collision zone in the north (Klinger et al. 2000). Large earthquakes, with  $M \sim 6 - 7.3$ , are known to have occurred repeatedly along the DSFZ throughout the historical period (Grünthal & Wahlström 2009; Grünthal et al. 2009c).

According to the distribution of the earthquakes and associated surface displacements along the DSFZ (Khair et al. 1997; Janssen et al. 2007), the study area along the DSFZ has been divided into three subareas, a southern, a central and a northern area (figure 4.2). About 105 km of left lateral displacement has been observed along the southern and central part of the DSFZ with  $29.5^\circ \leq \text{latitude} \leq 33.2^\circ$ . This part of the fault was additionally subdivided into the southern, with  $29.5^\circ \leq \text{latitude} \leq 31^\circ$ , and the central, with  $31^\circ \leq \text{latitude} \leq 33.2^\circ$ , subzones considering the seismicity distribution. About 60 km of displacement have been estimated for the northern part of the DSFZ (Khair et al. 1997). This part is considered as the northern subzone within  $33.2^\circ \leq \text{latitude} \leq 37^\circ$ . The time-dependent approach has not been applied for the southern part of the DSFZ, because only two main shocks with  $M_w \geq 6$  have occurred in this area since 300 AD, which are not sufficient for a rational statistical analysis.

The DSFZ earthquake catalog is obtained from available catalogs and special studies for the Middle East area, according to the procedure described by Grünthal et al. (2009a), Grünthal & Wahlström (2009) and Grünthal et al. (2009c). The

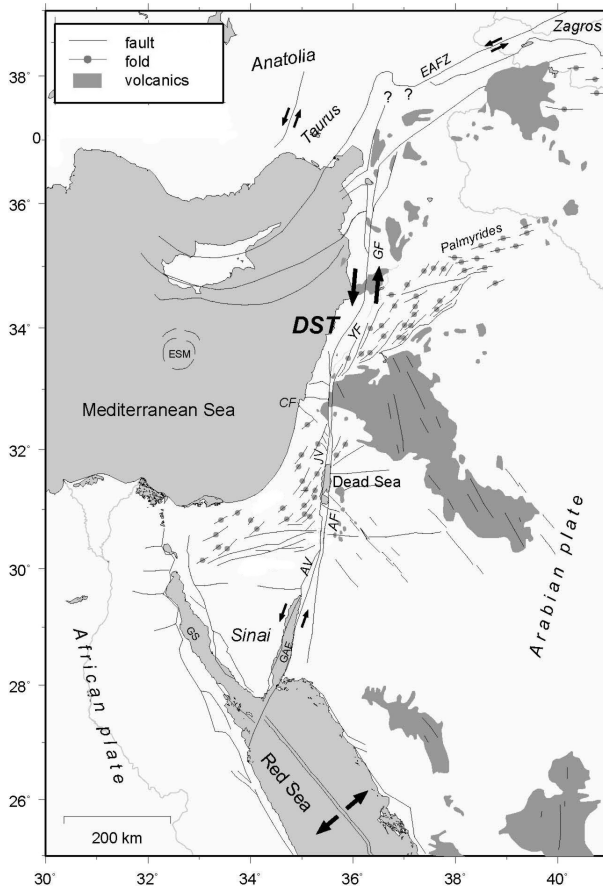


Figure 4.1: The Eastern Mediterranean area including the Dead Sea fault zone (modified after Garfunkel 1981). Arrows show relative motions across faults. AF = Arava Fault, AV = Arava Valley, CF = Carmel Fault, DST = Dead Sea Transform, EAFZ = East Anatolian Fault, ESM = Eratosthenes Sea Mount, GAF = Gulf of Aqaba Fault, GF = Gharb Fault, GS = Gulf of Sinai, JV = Jordan Valley, YF = Yamuna Fault.

dataset used in this thesis is excerpt from the EMEC catalog (Euro-Mediterranean Earthquake Catalog; Grünthal & Wahlström 2009) which is under construction at GFZ in a similar way to and as an extension of CENEC (Grünthal et al. 2009a). All known earthquakes with  $M_w \geq 6$  along the DSFZ, defined with an average width of about 40 km (figures 4.2 and 4.3), are obtained from the EMEC catalog and are here called the "DSFZ catalog". This dataset is declustered, i.e. the fore- and aftershocks are removed from the dataset, using a method from Grünthal (1985) described in details by Burkhard & Grünthal (2009) and Grünthal et al. (2009b). The completeness period of this dataset has been proved using a new method to start at about 300 AD. This new completeness check, as well as the method of declustering, will be described in the following sections.

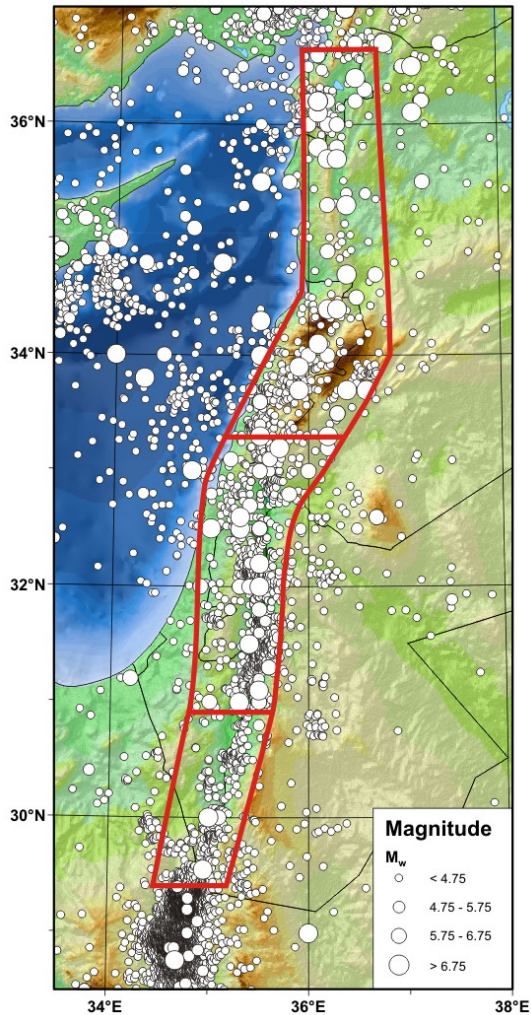


Figure 4.2: Seismicity in the DSFZ, including the subdivision into a southern, a central and a northern zone. The sources of the seismicity data have been described in section 4.1. Red curves show the border of each zone.

### 4.1.1 Declustering fore- and aftershocks

The declustering is done to identify and to separate the fore- and aftershocks from the main shocks in a catalog. Several approaches have been applied in order to decluster earthquake catalogs (Utsu 1969; Gardner and Knopoff, 1974; Grünthal, 1985; Reasenber 1985; Youngs et al., 1987). Zhuang et al. (2002) summarize that most of these methods remove the earthquakes in a space-time window around a large event (the main shock); the differences among the methods relate to the choice of window sizes and some other details (e.g., Utsu 1969; Gardner and Knopoff 1974; Grünthal, 1985; Kellis-Borok and Kossobokov 1986). In general, larger magnitudes of the main shocks result in larger space-time windows.

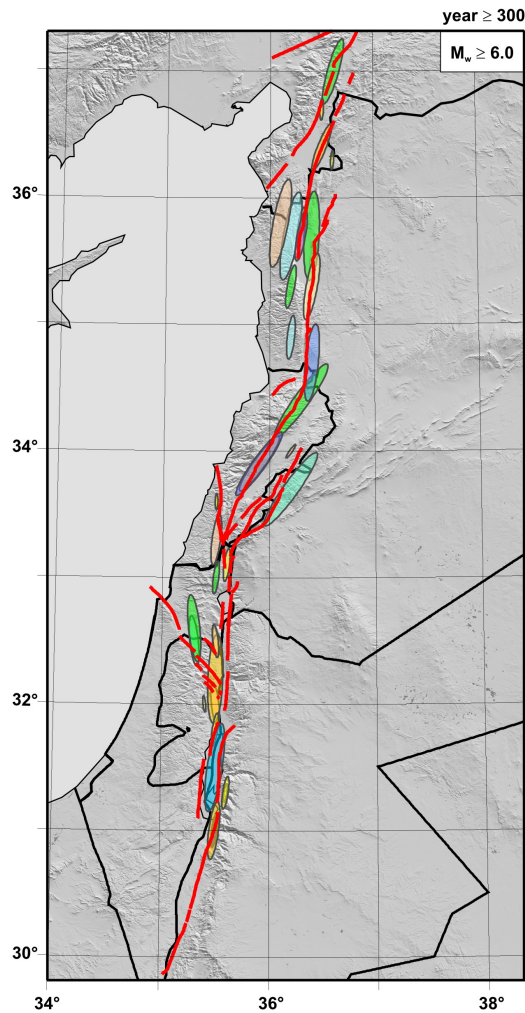


Figure 4.3: Surface rupture traces of the  $M_w \geq 6.0$  earthquakes in the DSFZ since 300 AD. Red lines show the fault segments (data from Heidbach & Ben-Avraham 2007; Meghraoui et al. 2003).

Zhuang et al. (2002) explain that an alternative to the window-based declustering methods is the link method, based on a space-time distance between events (e.g., Reasenber 1985; Frohlich and Davis 1990; Davis and Frohlich 1991).

In reality, these approaches do not produce considerably different results in cases where the earthquake catalogs (or datasets) include only large events bounded to a fault zone. The earthquake dataset used in this thesis has been declustered using an approach by Grünthal (1985) extended to larger magnitudes by Burkhard & Grünthal (2009)<sup>1</sup>. Based on this approach, the time windows are calculated for

<sup>1</sup>Burkhard & Grünthal (2009) describe the advantages of this method and compare it with several other methods.

foreshocks as:

$$dT_f(M_w) = \begin{cases} \exp(-4.77 + \sqrt{0.62 + 17.32M_w}) & \text{if } M_w < 7.8 \\ \exp(6.44 + 0.055M_w) & \text{otherwise} \end{cases}, \quad (4.1)$$

and for aftershocks as:

$$dT_a(M_w) = \begin{cases} \exp(-3.95 + \sqrt{0.62 + 17.32M_w}) & \text{if } M_w < 6.6 \\ \exp(6.44 + 0.055M_w) & \text{otherwise} \end{cases}. \quad (4.2)$$

The distance window is calculated as:

$$dR(M_w) = \exp\left(1.77 + \sqrt{0.037 + 1.02M_w}\right). \quad (4.3)$$

For a given main shock all smaller magnitude events within  $dR$ , occurring within the fore- or aftershock time windows, are considered dependent events and removed. Applying this approach to the dataset marks two foreshocks and no aftershocks. Table 4.1 shows these foreshocks, as well as the corresponding main shocks.

Table 4.1: Foreshocks and corresponding main shocks along the DSFZ,  $M_w \geq 6$ .

year	month	day	latitude	longitude	Mw	kind of event
362	5	24	31.3	35.6	6.4	foreshock
363	5	19	31.5	35.5	7.2	main shock
1759	10	30	33.1	35.6	6.5	foreshock
1759	11	25	33.7	36.2	7.3	main shock

### 4.1.2 A new method of earthquake catalog completeness check

The aim of checking the completeness of a catalog is to find out, since what time all events in a given magnitude class (e.g.  $M_w \geq 6$ ) are included in the catalog. The commonly used approaches to check the completeness of catalogs are a method by Stepp (1972), as well as the visual method<sup>2</sup>. Both methods assume that the earthquake inter-event times follow a Poisson distribution. This assumption is not true if the earthquake occurrence rate is time-dependent. Another method of completeness check, introduced by Rotondi & Garavaglia (2002), has been applied based on a change point process. This method also assumes that the earthquake inter-event times follow a Poisson process and is therefore not suitable for

<sup>2</sup>The visual method is based on monitoring the trend of the cumulative number of earthquakes for different magnitude classes. The completeness starts at a point where the trend has a fixed slope until the end of the catalog.

cases with time-dependent earthquake occurrence rates. Furthermore, although this method may statistically seem more reliable, it is neither simple to apply nor applicable for those cases in which there is no plain change of slope between complete and incomplete datasets.

In this thesis, a new method for investigating data completeness has been developed and applied to the DSFZ catalog. In this method, a deviation criterion is considered to compare the compatibility of the earthquake inter-event times in different time spans. The complete data start, where the deviation criterion remains stable and does not significantly change with time. Since no prior distribution has been considered for the earthquake inter-event times in this new method, the only parameter to be monitored is the standard deviation of the earthquake inter-event times for earthquakes above a magnitude threshold (here  $M_w \geq 6$ ). Therefore, it is also applicable for the case with time-dependent earthquake occurrence rates. Consider the set of different time spans as:

$$\Theta = \{T_i | i = 2, \dots, n\} \quad (4.4)$$

where  $T_i = T - \{t_{n-j} | j = 0, \dots, n - i\}$  ( $i = 2, \dots, n$ ),  $T$  is the set of all earthquake inter-event times and  $t_i$  is the  $i$ -th ranked inter-event time based on occurrence times. Now define

$$SD(i) = \left( \frac{\sum_{t \in T_i} (t - \bar{T}_i)^2}{N(T_i) - 1} \right)^{1/2}, \quad \bar{T}_i = \frac{\sum_{t \in T_i} t}{N(T_i)}, \quad i = 2, \dots, n \quad (4.5)$$

as the unbiased estimation of the standard deviation of the set  $T_i$ , where  $N(T_i)$  is the number of inter-event times in the set  $T_i$ .

In the next step, the difference between two successive standard deviations will be calculated as

$$DSD(i) = SD(i + 1) - SD(i), \quad i = 2, \dots, n. \quad (4.6)$$

Finally, the  $DSDs$  will be monitored against occurrence times. The completeness of the catalog can then be defined, when the  $DSDs$  begin to show a uniform trend. Based on this method, the completeness of the DSFZ catalog has been estimated. Table 4.2 shows completeness times for different magnitude classes in the DSFZ catalog. "Completeness time" in table 4.2 represents since what time (year) the data is complete. It should be mentioned that the "number of events" in this table refers to the number of events in the main catalog of the DSFZ. According to the completeness times and number of earthquakes in each class in table 4.2, earthquakes with  $M_w \geq 6$  have been considered as large earthquakes in the DSFZ catalog.

Table 4.2: Completeness for different magnitude classes in the DSFZ catalog.

magnitude class	completeness time	number of events in complete part
$M_w \geq 5.0$	1850	26
$M_w \geq 5.5$	1500	29
$M_w \geq 6.0$	300	78
$M_w \geq 6.5$	150	49
$M_w \geq 7.0$	0	25

### 4.1.3 Final dataset

According to the aim of this thesis, the long-term behavior of the seismicity of large earthquakes will be analyzed in order to investigate whether this behavior is time-dependent or not. Therefore, the dataset considered in this thesis should be as long as possible and should contain sufficient number of events. This dataset can be achieved considering the dataset of earthquakes with  $M_w \geq 6$  which can be considered complete since 300 AD. The final dataset of this thesis includes 50 earthquakes with  $M_w \geq 6$  (table 4.3) since 300 AD, where the northern zone and the central zone contain 31 and 17 earthquakes, respectively (tables 4.4-4.5). The dataset is declustered. The southern zone contains only two earthquakes, which are not enough for a rational statistical analysis. Therefore, the statistical approach has been applied to the northern and the central subzones.

The first column in tables 4.3-4.5, "year", shows the occurrence year of earthquakes in the dataset. The second and third columns, "latitude" and "longitude", present the coordinates of the earthquake epicenter. The fourth column, " $M_w$ ", presents the *moment magnitude* of the corresponding earthquake. The moment magnitude (denoted as  $M_w$ ), defined by Hanks and Kanamori (1979), is based on the seismic moment of the earthquake, which is equal to the rigidity of the Earth multiplied by the average co-seismic slip on the fault and the size of the rupture area<sup>3</sup>.  $M_w$  is the preferred measure of earthquake sizes in the seismological community concerning to PSHA. It is a physically based measure and is used in most modern ground motion prediction equations. Consequently, the catalog of the DSFZ is prepared, following the concept of harmonized  $M_w$  (Grünthal & Wahlström, 2003; Grünthal et al., 2009a).

The fifth column, "inter-event time", represents the time difference between every two successive earthquakes in the dataset. The sixth column, "rupture length", gives the estimated subsurface rupture length of earthquakes in the dataset, calcu-

<sup>3</sup>"Glossary of Terms on Earthquake Maps". USGS. [http:// earthquake.usgs.gov/ eqcenter/ glossary.php](http://earthquake.usgs.gov/eqcenter/glossary.php) . Retrieved 2009-03-21.



Table 4.3: Earthquake dataset of the DSFZ with  $M_w \geq 6$ .

year	latitude	longitude	Mw	inter-event time (yr)	surface rupture length (km)
341	36.2	36.1	6.1	-	15.2
363	31.5	35.5	7.2	22.4	75.1
457	36.1	36.1	6.3	94.3	21.4
494	35.8	36.3	6.5	36.3	26.9
500	36.2	36.1	7.0	6.0	59.7
526	36.2	36.1	6.7	26.4	37.8
528	36.2	36.1	6.8	2.5	42.4
551	33.9	35.9	7.1	22.6	67.0
565	36.0	36.2	6.1	13.5	15.2
588	36.2	36.2	6.5	23.8	29.3
634	32.0	35.5	6.7	45.2	40.0
658	32.5	35.5	6.4	24.5	24.0
713	36.0	36.0	6.7	54.8	37.8
749	32.2	35.5	7.2	35.8	75.1
756	32.0	35.5	6.1	7.1	15.6
835	36.0	36.0	6.0	78.8	13.5
847	34.4	36.3	7.3	12.9	84.1
854	32.5	35.3	6.5	6.1	123.7
860	35.7	36.4	7.2	6.0	47.5
972	36.1	36.1	6.0	112.0	13.5
972	31.0	35.5	6.5	0.0	26.9
1033	32.5	35.5	6.5	61.9	26.9
1047	31.0	35.5	6.5	13.1	26.9
1060	32.2	35.5	6.1	13.0	15.6
1063	34.4	36.3	6.8	3.6	42.4
1068	29.6	35.0	7.0	4.6	59.7
1068	32.6	35.3	7.0	0.2	59.7
1091	36.1	36.1	6.5	23.3	29.3
1150	31.8	35.5	6.5	58.3	29.3
1157	35.3	36.4	7.0	7.6	59.7
1170	34.7	36.4	7.1	12.9	67.0
1202	33.9	35.9	7.0	31.9	59.7
1212	30.0	35.0	6.9	9.9	47.5
1259	32.5	35.5	6.5	46.9	26.9
1269	37.0	36.6	7.0	10.1	59.7

Table 4.3: Continued.

year	latitude	longitude	M <sub>w</sub>	inter-event time (yr)	rupture length (km)
1287	33.0	35.5	6.5	18.0	26.9
1404	35.7	36.2	7.2	116.9	75.1
1407	35.7	36.3	6.9	3.2	47.5
1408	35.8	36.1	7.2	1.7	75.1
1458	31.0	35.5	6.9	49.9	53.3
1656	34.9	36.2	6.7	197.2	37.8
1726	36.3	36.6	6.1	70.2	17.0
1738	36.7	36.5	6.2	12.4	19.1
1759	33.7	36.2	7.3	21.2	84.1
1796	35.3	36.2	6.7	36.4	37.8
1802	34.0	36.2	6.1	5.7	15.2
1834	31.3	35.6	6.1	32.4	17.0
1837	33.3	35.5	6.9	2.6	47.5
1872	36.4	36.5	6.9	35.3	47.5
1927	32.0	35.4	6.1	55.3	15.2

lated through

$$\log(RLD) = a + bM_w \quad (4.7)$$

where  $RLD$  is the subsurface rupture length (km),  $a = -2.57$  and  $b = 0.62$  are parameters, and  $M_w$  is the moment magnitude (Wells & Coppersmith 1994).

#### 4.1.4 Some characteristics of the dataset

A simple test for spatial clustering of earthquakes is to plot epicenter latitudes against occurrence times (figures 4.4a & 4.4b), which shows no significant regularity for the dataset at hand. A very simple correlation test, i.e. a linear regression test, also indicates no significant correlation between earthquake locations (i.e. latitudes) and earthquake occurrence times in the entire DSFZ, or in the central and northern parts of the DSFZ separately. The mean value of the earthquake inter-event times and their standard deviation are shown in table 4.6 for different parts of the DSFZ. According to table 4.6, the ratio of the standard deviation to the mean value for each region is greater than one. This topic will be discussed in the next section in this chapter.

Table 4.4: Earthquake dataset of the northern part of the DSFZ with  $M_w \geq 6$ .

year	latitude	longitude	$M_w$	inter-event time (yr)	surface rupture length (km)
341	36.2	36.1	6.1	-	15.2
457	36.1	36.1	6.3	116.7	21.4
494	35.8	36.3	6.5	36.3	26.9
500	36.2	36.1	7.0	6.0	59.7
526	36.2	36.1	6.7	26.4	37.8
528	36.2	36.1	6.8	2.5	42.4
551	33.9	35.9	7.1	22.6	67.0
565	36.0	36.2	6.1	13.5	15.2
588	36.2	36.2	6.5	23.8	29.3
713	36.0	36.0	6.7	124.4	37.8
835	36.0	36.0	6.0	121.8	13.5
847	34.4	36.3	7.3	12.9	84.1
860	35.7	36.4	7.2	12.1	47.5
972	36.1	36.1	6.0	112.0	13.5
1063	34.4	36.3	6.8	91.6	42.4
1091	36.1	36.1	6.5	28.1	29.3
1157	35.3	36.4	7.0	65.9	59.7
1170	34.7	36.4	7.1	12.9	67.0
1202	33.9	35.9	7.0	31.9	59.7
1269	37.0	36.6	7.0	66.9	59.7
1404	35.7	36.2	7.2	134.8	75.1
1407	35.7	36.3	6.9	3.2	47.5
1408	35.8	36.1	7.2	1.7	75.1
1656	34.9	36.2	6.7	247.1	37.8
1726	36.3	36.6	6.1	70.2	17.0
1738	36.7	36.5	6.2	12.4	19.1
1759	33.7	36.2	7.3	21.2	84.1
1796	35.3	36.2	6.7	36.4	37.8
1802	34.0	36.2	6.1	5.7	15.2
1837	33.3	35.5	6.9	35.0	47.5
1872	36.4	36.5	6.9	35.3	47.5

Table 4.5: Earthquake dataset of the central part of the DSFZ with  $M_w \geq 6$ .

year	latitude	longitude	Mw	inter-event time (yr)	surface rupture length (km)
363	31.5	35.5	7.2	-	75.1
634	32.0	35.5	6.7	270.6	40.0
658	32.5	35.5	6.4	24.5	24.0
749	32.2	35.5	7.2	90.6	75.1
756	32.0	35.5	6.1	7.1	15.6
854	32.5	35.3	6.5	97.8	123.7
972	31.0	35.5	6.5	118.0	26.9
1033	32.5	35.5	6.5	61.9	26.9
1047	31.0	35.5	6.5	13.1	26.9
1060	32.2	35.5	6.1	13.0	15.6
1068	32.6	35.3	7.0	8.4	59.7
1150	31.8	35.5	6.5	81.6	29.3
1259	32.5	35.5	6.5	109.2	26.9
1287	33.0	35.5	6.5	28.0	26.9
1458	31.0	35.5	6.9	171.6	53.3
1834	31.3	35.6	6.1	375.5	17.0
1927	32.0	35.4	6.1	93.1	15.2

## 4.2 Temporal cluster behavior in the DSFZ

The *aperiodicities*<sup>4</sup> of earthquake inter-event times in the DSFZ and its corresponding subzones; i.e. the northern and the central part, are larger than one. This indicates a *clustering behavior* of the earthquake occurrence rate of main shocks in the DSFZ<sup>5</sup> (see Mucciarelli 2007). The clustering behavior indicates that the earthquake occurrence may change in time.

Another indication to a clustering behavior appears when considering the *Benioff curve*. The Benioff curve shows how the cumulative *seismic moment release*, cumulative  $M_0$ , increases with time. Following Hanks and Kanamori (1979), the seismic moment release due to an earthquake can be calculated, as:

$$M_0 = 10^{\frac{3(M_w + 6.033)}{2}}. \quad (4.8)$$

<sup>4</sup>The aperiodicity is a parameter, which represents the ratio of the standard deviation to the mean value of a given dataset.

<sup>5</sup>This type of clustering should not be confused with the clustering in terms of main shocks, foreshocks and aftershocks. Neglecting fore- and aftershocks, this type of clustering is due to only main shocks. One can consider it as *long-term clustering*.

Table 4.6: Mean value and standard deviation of inter-event times in the DSFZ.

region	mean value of inter-event times (yr)	standard deviation of inter-event times (yr)
entire DSFZ	32	37
central part of the DSFZ	98	102
northern part of the DSFZ	51	56

Figure 4.5 shows the Benioff curve for the northern part of the DSFZ. According to figure 4.5, two different processes for cumulative  $M_0$ ; i.e. cumulative energy release as a measure of seismic activity, can be identified. The first process takes place, when cumulative  $M_0$  dramatically increases within a relatively short period. This period can be called an "active period". When the cumulative  $M_0$  exceeds a certain higher "level of seismic moment release", the second process begins. The second process occupies a relatively longer period where the amount of seismic moment release increases relatively slowly. This period can be called an "inactive period". As soon as the second process exceeds a "time threshold", the first process will be activated again. It seems, from figure 4.5, that the current situation in the DSFZ (at least in the northern part) is more compatible with the second process; i.e. the area is currently experiencing an inactive period.

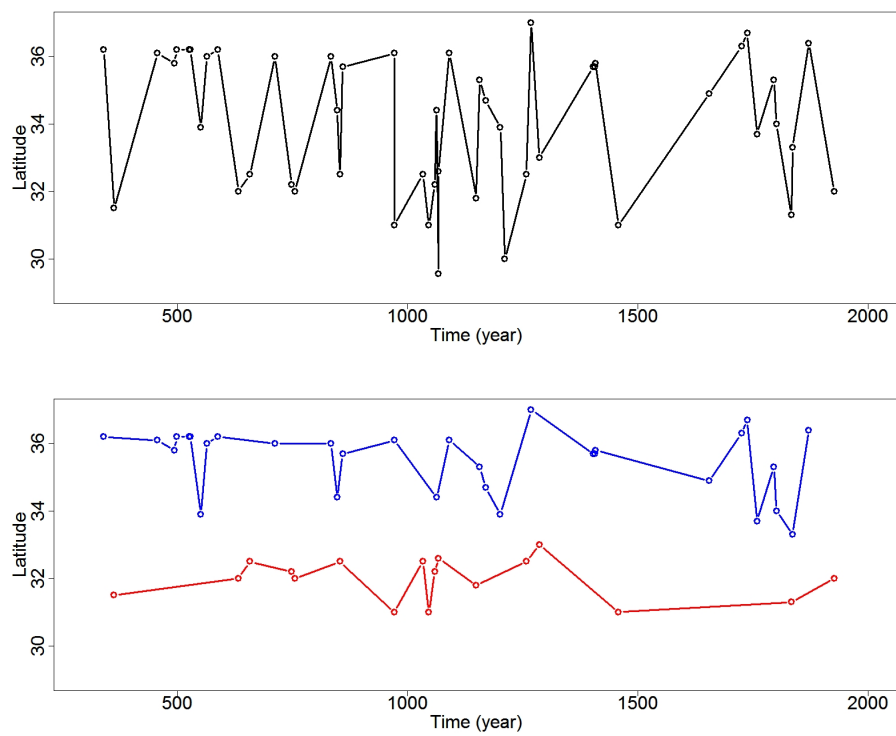


Figure 4.4: Latitude of the earthquake epicenters vs. earthquake occurrence times. Top: For the entire DSFZ. Bottom: Blue curve for the northern part of the DSFZ and red curve for the central part of the DSFZ.

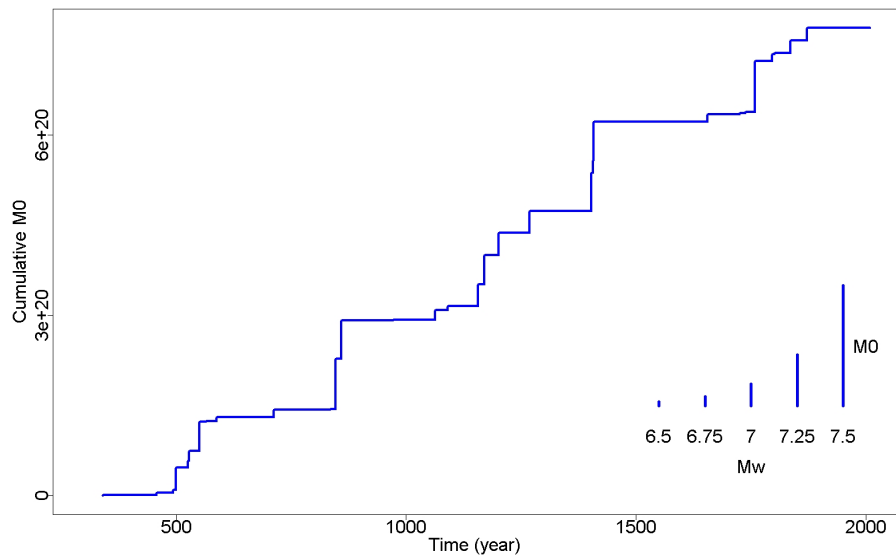


Figure 4.5: Benioff curve for the northern part of the DSFZ. The cumulative seismic moment release shows "active" and "inactive" periods.

# Chapter 5

## Analysis of earthquake occurrence rates in the DSFZ

In this chapter the approach discussed in chapter 3 will be applied to the data; i.e. the earthquake inter-event times, extracted in chapter 4. The method has been applied for two cases, namely the central and the northern part of the DSFZ. For each case the output will be discussed separately.

### 5.1 Earthquake occurrence rate in the central part of the DSFZ

According to the approach of chapter 3, the first step is to test the multimodality of the earthquake inter-event times. In order to test the multimodality, a bootstrap method has been applied. The output of the test indicates no bimodality of the inter-event times for the central part of the DSFZ (see figure 5.1). Considering

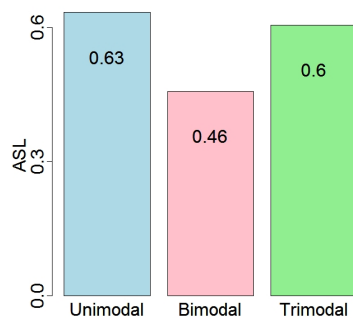


Figure 5.1: Results in terms of the achieved significance level (ASL) of the multimodality test with 1000 bootstrap resamplings for the central part of the DSFZ. The inter-event times reveal no specific bimodality.

figure 5.1, the earthquake occurrence rate could be also considered trimodal for



the central part of the DSFZ, based on its ASL. However the number of model parameters required for a trimodal distribution is eight (two proportion parameters plus 3\*2 modal parameters, two parameters for each modal). The size of the earthquake inter-event times dataset is 16 for the central part of the DSFZ (17 earthquakes in total), which is not sufficient to make a significant parameter estimation for the trimodal distribution. Since the ASL of the unimodality in figure 5.1 is larger than the ASL of the trimodality, the earthquake occurrence rate in the central part of the DSFZ will be considered as unimodal.

Following the modified weighted MLE method discussed in chapter 3, the parameters of the candidate distributions have been estimated. The five candidate distributions are: Weibull, Gamma, Lognormal and BPT as time-dependent distributions against the exponential (based on the Poisson process) as the time-independent distribution. The weight function parameters (in equation 3.40) are considered as  $\alpha = 1$ ,  $p = 6$  and  $K = 1$  (figure 5.2). These parameter values

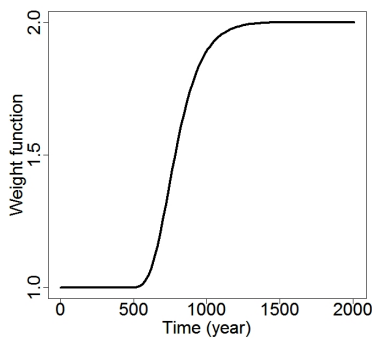


Figure 5.2: The weight function with parameters;  $\alpha = 1$ ,  $p = 6$  and  $K = 1$ .

form a weight function, which is strongly discriminative between two time spans, before 500 AD and after 1200 AD. In between these two time spans, the weight function is strictly increasing. This form of the weight function is in agreement with the reliability of the earthquake data in the catalog<sup>1</sup>.

Table 5.1 presents the estimation results for the five candidate distributions. In order to select the best model, the values of the BIC (Bayesian information criterion) as well as the KS-test (Kolmogorov-Smirnov test) p-value have been estimated. Based on these two criteria, the Weibull distribution, which is the best time-dependent model, is not significantly better than the exponential distribution, which is the time-independent model. Therefore, the annual earthquake occurrence rate for the central part of the DSFZ can be considered as time-independent. Figure 5.3 shows the cdfs of the five different distributions comparing with the

<sup>1</sup>The earthquake data in the main catalog are more reliable for the recent time; i.e. after 1000 AD.

Table 5.1: Estimated unimodal distributions for the central part of the DSFZ.

Distribution	Parameters		ML	BIC	KS-test
	scale	shape			
Exponential	102.88	-	-89.4	181.7	0.92
Weibull	85.32	0.88	-89.8	185.3	0.97
Gamma	98.88	1.04	-89.4	184.5	0.92
Lognormal	4.09	1.19	-89.8	185.2	0.92
BPT	97.76	38.51	-90.1	185.8	0.56

empirical distribution (see equation 3.53) of the inter-event times in the central part of the DSFZ. According to the estimated exponential distribution, the annual

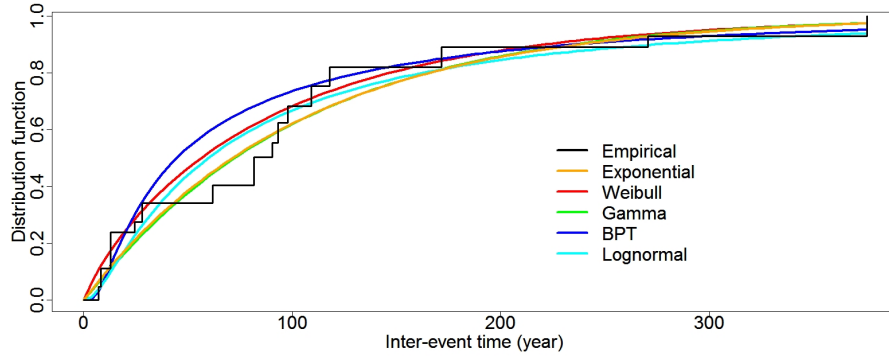


Figure 5.3: Comparison between different distribution functions and the empirical distribution function of earthquake inter-event times for the central part of the DSFZ.

earthquake occurrence rate for  $M_w \geq 6$  is  $\hat{\lambda} = 1/102.88 = 0.00972$  in the central part of the DSFZ. This rate results in a mean recurrence time of  $1/\hat{\lambda} = 102.88$  years. Following the bootstrap confidence method discussed in chapter 3, the confidence interval of the estimated parameter  $\hat{\lambda}$  has been calculated (equations 3.54 and 3.55). Suppose the distribution of  $\mathfrak{R}_n^*$  as  $H_{BOOT}$  (a nonparametric bootstrap estimator of  $G_{BOOT}$ ). Then equation 3.55 will change to:

$$\mathbf{C}_n = \left[ \hat{\theta}_n - \hat{\sigma}_n H_{BOOT}^{-1} \left( 1 - \frac{\alpha}{2} \right), \hat{\theta}_n + \hat{\sigma}_n H_{BOOT}^{-1} \left( \frac{\alpha}{2} \right) \right]. \quad (5.1)$$

Using 1000 bootstrap iterations to estimate  $H_{BOOT}^{-1}(\cdot)$ , plus 1000 bootstrap iterations to estimate  $\hat{\sigma}_n^*$  in each iteration (altogether  $1000 * 1000 = 1,000,000$  bootstrap iterations), the confidence interval for a 80% significance level has been

calculated as:  $C_n(\hat{\lambda}) = [0.00600, 0.01298]$  and  $C_n(1/\hat{\lambda}) = [77.04, 166.67]$ . The confidence interval for a 95% level of significance is:  $C_n(\hat{\lambda}) = [0.00388, 0.01612]$  and  $C_n(1/\hat{\lambda}) = [62.03, 257.73]$ .

The conditional exceedance probability, i.e. the conditional probability of an earthquake occurrence within time  $t$ , given the censored time  $t_c$  ( $t \geq t_c$ ), is calculated as:

$$Pr \{T \leq t \mid T \geq t_c\} = \frac{Pr \{t_c \leq T \leq t\}}{Pr \{T \geq t_c\}} = \frac{F(t) - F(t_c)}{1 - F(t_c)} \quad (5.2)$$

where  $T$  is the random variable of the earthquake inter-event times,  $Pr(\cdot)$  is the corresponding probability function, and  $F(\cdot)$  is the cdf regarding to  $T$ . Here, the random variable  $T$  is exponentially distributed. Therefore,

$$Pr \{T \leq t \mid T \geq t_c\} = \frac{F(t) - F(t_c)}{1 - F(t_c)} = 1 - e^{-\lambda(t-t_c)} = F(t - t_c). \quad (5.3)$$

The censored time  $t_c$  in this area is  $\sim 81$  years, since the last earthquake occurred in 1927. By plugging  $t_c = 81$  into equation 5.3, the conditional exceedance probability of  $M_w \geq 6$  earthquakes in the central part of the DSFZ within the next 30 years; i.e.  $t - t_c = 30$ , is calculated as

$$Pr \{T \leq t \mid T \geq t_c\} = Pr \{T \leq 30 + t_c \mid T \geq t_c\} = F(30) \approx 0.25. \quad (5.4)$$

Equation 5.4 explains that the conditional probability of an earthquake occurrence does not depend on the censored time  $t_c$ , but just the time span between  $t$  and  $t_c$ . This shows the time-independency of the exponential distribution.

## 5.2 Earthquake occurrence rate in the northern part of the DSFZ

The dataset used in the northern part of the DSFZ ( $33.2^\circ \leq \text{latitude} \leq 37^\circ$ ) includes 31 historical earthquakes (30 earthquake inter-event times) with  $M_w \geq 6$ .

### 5.2.1 The northern part of the DSFZ - a special case

According to the dataset, the last  $M_w \geq 6$  earthquake in this area occurred in 1872. Therefore, the time since the last earthquake (i.e. the censored time) is  $\sim 137$  years. The average earthquake inter-event time in this area is  $\sim 51$  years, which is much smaller than 137 years. Another fact is that more than 96% of the observed inter-event times in this area are shorter than the censored time.

Furthermore, the earthquake inter-event times in the northern part of the DSFZ reveal a clustering property. The inter-event time dataset can be divided into two different parts, one with inter-event times  $< 37$  years, another with inter-event times  $> 65$  years (figure 5.4). This clustering property is also an indication of multimodality of the earthquake inter-event times in the northern part of the DSFZ.

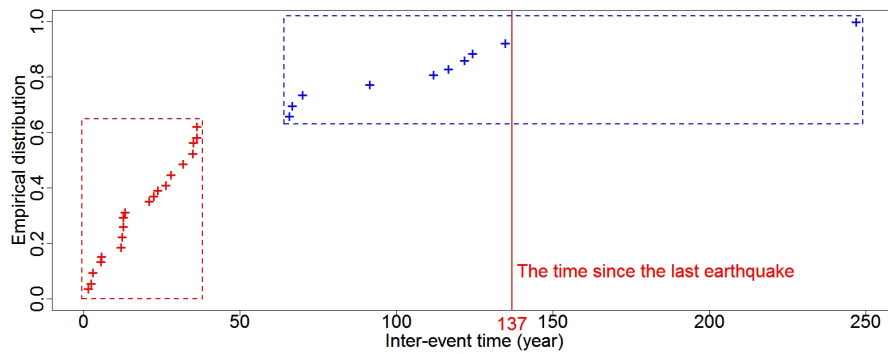


Figure 5.4: A schematic clustering behavior of the inter-event times in the northern part of the DSFZ. The crosses present the empirical distribution function, red crosses show the first cluster, and blue crosses demonstrate the second cluster. The red vertical line shows the censored time and its situation comparing with inter-event times.

These two facts together make the case of the northern part of the DSFZ special. It may be considered "a case with an *overdue earthquake*".

## 5.2.2 Multimodality of inter-event times

The first step of the estimation process is to check if the distribution of the earthquake inter-event times in this area is multimodal. Here again, the bootstrap method has been applied to test the multimodality of the distribution of earthquake inter-event times. The results of the test indicate no significant difference between unimodality and bimodality (figure 5.5). Since the multimodality test cannot significantly distinguish between unimodality and bimodality, the distribution of the earthquake inter-event times will be estimated under both assumptions; i.e. unimodality and bimodality.

## 5.2.3 Unimodal distributions

As is shown in table 5.2, based on the BIC and the KS-test p-value, the best unimodal distribution for the northern part of the DSFZ is the Lognormal distribution.

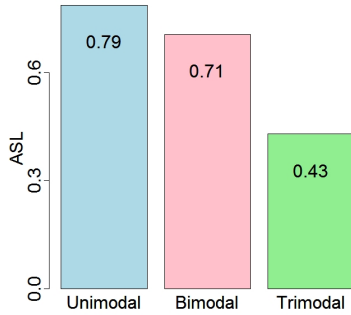


Figure 5.5: The results of the achieved significance level (ASL) of the multimodality test with 1000 bootstrap resamplings for the northern part of the DSFZ. The test does not significantly show, whether unimodality or bimodality is a better model.

Figure 5.6 demonstrates the cdf of different unimodal distributions against the empirical distribution function (extracted from the earthquake inter-event times in the northern part of the DSFZ).

Table 5.2: Estimated unimodal distributions for the northern part of the DSFZ.

Distribution	Parameters		ML	BIC	KS-test
	scale	shape			
Exponential	58.44	-	-151.1	305.6	0.44
Weibull	45.46	0.93	-151.5	310.0	0.94
Gamma	65.46	0.90	-150.9	308.7	0.58
Lognormal	3.43	1.30	-151.2	309.4	0.99
BPT	51.04	15.73	-153.7	314.3	0.37

## 5.2.4 Bimodal (mixed) distributions

Three bimodal distributions have been tested for the earthquake inter-event time dataset for the northern part of the DSFZ. The first model is a mixed Lognormal-Weibull distribution with the pdf:

$$f(t) = \rho \left[ \left( \frac{1}{t\sigma\sqrt{2\pi}} \right) \exp \left( -\frac{(\ln t - \mu)^2}{2\sigma^2} \right) \right] + (1 - \rho) \left[ \left( \frac{\beta}{\eta} \right) \left( \frac{t}{\eta} \right)^{\beta-1} \exp \left( -\left( \frac{t}{\eta} \right)^\beta \right) \right], \quad t \geq 0, \quad (5.5)$$

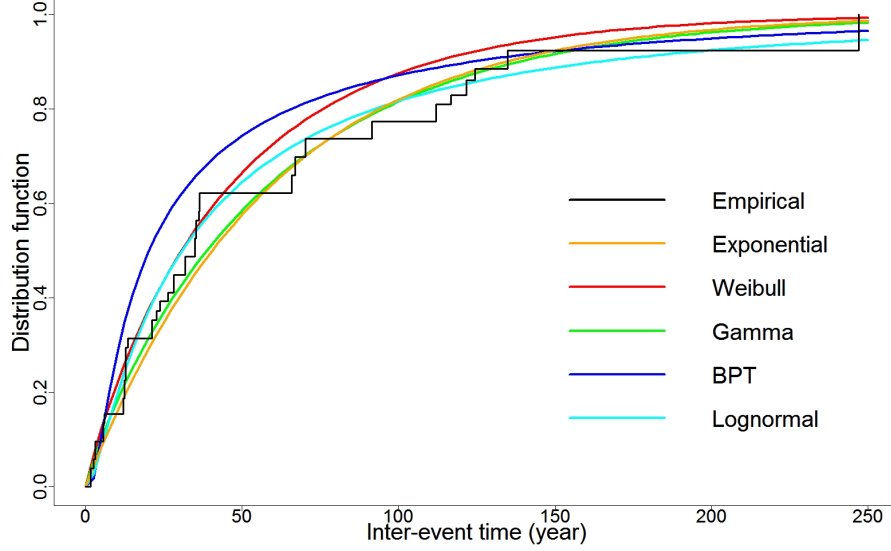


Figure 5.6: Comparison between different unimodal distribution functions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ.

where  $\rho \in [0, 1]$  is the proportion parameter.

The second model is a BPT-Weibull distribution with the pdf:

$$f(t) = \rho \left[ \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left(-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right) \right] + (1-\rho) \left[ \left(\frac{\beta}{\eta}\right) \left(\frac{t}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{t}{\eta}\right)^\beta\right) \right], \quad t \geq 0, \quad (5.6)$$

and the third model is a Weibull-Weibull distribution with the pdf:

$$f(t) = \rho \left[ \left(\frac{\beta_1}{\eta_1}\right) \left(\frac{t}{\eta_1}\right)^{\beta_1-1} \exp\left(-\left(\frac{t}{\eta_1}\right)^{\beta_1}\right) \right] + (1-\rho) \left[ \left(\frac{\beta_2}{\eta_2}\right) \left(\frac{t}{\eta_2}\right)^{\beta_2-1} \exp\left(-\left(\frac{t}{\eta_2}\right)^{\beta_2}\right) \right], \quad t \geq 0. \quad (5.7)$$

The three basic unimodal distributions used as basis for these three mixed distributions are the Weibull, BPT and Lognormal distributions. The reasons to choose these three mixed distributions are:

- the Lognormal and Weibull unimodal distributions fit the data well.

- the BPT distribution has a physical basis (Matthews et al., 2002).
- the statistical hazard functions of these three distributions are compatible with the behavior of the first cluster of the desired mixed distribution (the hazard function increases fast shortly after an event occurrence, seeks a maximum, then decreases).
- the hazard function of the Weibull distribution is compatible with the second part of the desired mixed distribution, as it can be constructed to increase with time.

Following the methods of EM described in chapter 3, the parameters of the three mixed distributions have been estimated. The parameters of the weight function are considered again as:  $\alpha = 1$ ,  $p = 6$  and  $K = 1$ . The results are shown in table 5.3. In table 5.3, BW, LW and WW denote the BPT-Weibull, the Lognormal-Weibull, and the Weibull-Weibull mixed distributions, respectively. Figure 5.7 shows a comparison between the empirical distribution function and the bimodal distribution functions. According to table 5.3, the best bimodal distribution, based on the BIC and KS-test p-value, is the mixed Weibull-Weibull distribution.

Table 5.3: Estimated bimodal distributions for the northern part of the DSFZ.

Distri- bution	Parameters					ML	BIC	KS-test
	$\rho$	scale 1	shape 1	scale 2	shape 2			
BW	0.51	17.18	14.28	112.52	1.50	-151.3	316.3	0.88
LW	0.29	2.52	1.14	72.03	1.06	-150.4	314.5	0.87
WW	0.28	17.44	1.33	73.63	1.06	-150.4	314.4	0.97

### 5.2.5 Selection of the best model

Considering tables 5.2 and 5.3, two distributions have been selected as candidate models in the northern part of the DSFZ:

- the mixed Weibull-Weibull distribution, which is the best bimodal distribution considering the BIC and KS-test p-value,
- the unimodal Lognormal distribution, because it is the best unimodal distribution, considering the BIC and the KS-test p-value.

These models are compared, focusing only on the longer inter-event times, i.e. inter-event times  $> 60$  years, in figure 5.8. It seems that the mixed Weibull-Weibull distribution can fit this part of inter-event times better than the unimodal

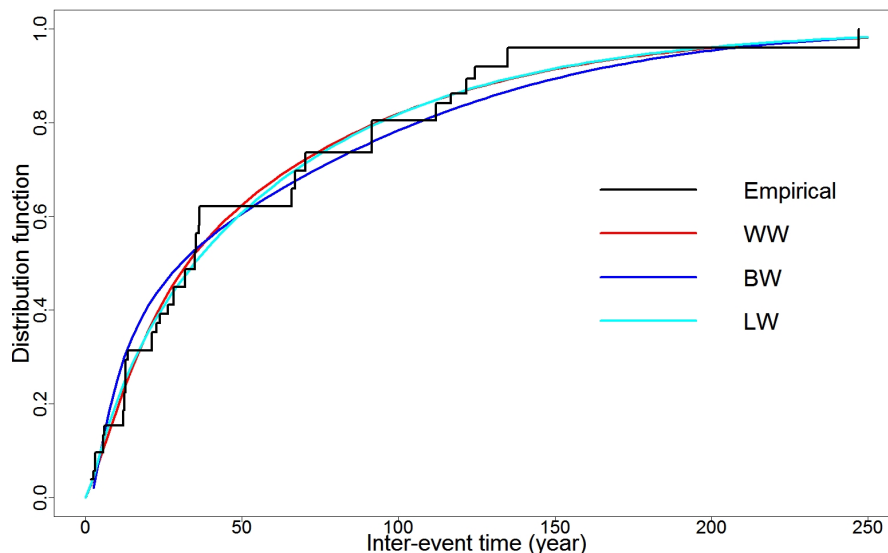


Figure 5.7: Comparison between different bimodal distribution functions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ. WW = mixed Weibull-Weibull, BW = mixed BPT-Weibull, LW = mixed Lognormal-Weibull.

Lognormal distribution.

The sum of squared errors ( $SSE = \sum [estimate - observed]^2$ ) has been calculated for the two candidate distributions against the empirical distribution for the inter-event times  $> 60$  years (table 5.4). Following table 5.4, although both distri-

Table 5.4: The sum of squared errors for the candidate distributions.

Distribution	ML	BIC	KS-test	$SSE_{t>60}$
Mixed WW	-150.4	314.4	0.97	0.007
Lognormal	-151.2	309.4	0.99	0.013

butions perform well in likelihood and KS-test, the smaller  $SSE_{t>60}$  obtained for the mixed Weibull-Weibull distribution confirms that the best fit for the inter-event times  $> 60$  years is the mixed Weibull-Weibull distribution.

This statement plays an important role in the process of the best model selection, as the time since the last  $M_w \geq 6$  earthquake (i.e. censored time) is much larger than the average earthquake inter-event time, as has already been discussed. Another important argument to select the mixed Weibull-Weibull distribution for



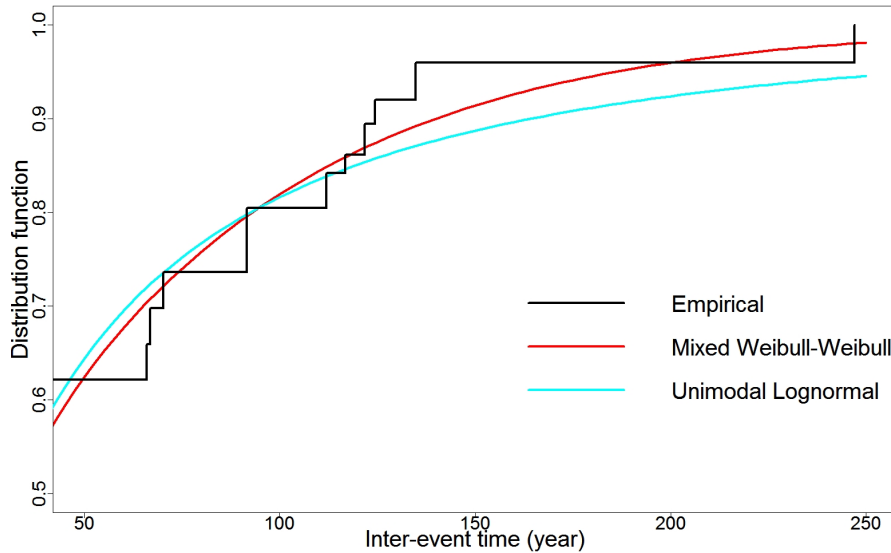


Figure 5.8: Comparison between distribution functions of mixed Weibull-Weibull and unimodal Lognormal distributions and the empirical distribution function of earthquake inter-event times for the northern part of the DSFZ. Results for longer inter-event times.

the earthquake inter-event times in the northern part of the DSFZ appears, considering the trends of the corresponding statistical hazard functions (figure 5.9). Considering figure 5.9, the hazard function of the mixed Weibull-Weibull distribution increases immediately after an earthquake occurrence to a maximum value ( $\sim 0.024$ ), then decreases to a minimum value ( $\sim 0.015$ ) and after that increases very slowly with time. In contrast, the hazard function of the unimodal Lognormal distribution increases after an earthquake occurrence to a maximum value ( $\sim 0.026$ ), then decays with the time. Among these two completely different trends, an increasing (the mixed Weibull-Weibull) and a decreasing (the unimodal Lognormal) hazard function for the long-term, the increasing hazard function for the long-term is in agreement with the long-term stress increase. Therefore, the mixed Weibull-Weibull distribution is selected to model the inter-event times of  $M_w \geq 6$  earthquakes in the northern part of the DSFZ.

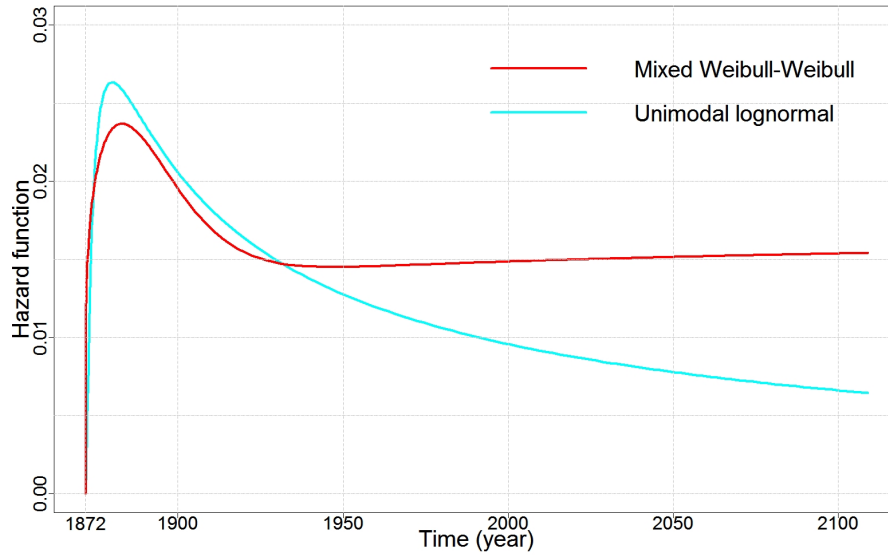


Figure 5.9: The trends of hazard functions of mixed Weibull-Weibull and unimodal Lognormal distributions after the last earthquake in 1872.

### The statistical hazard function according to the mixed Weibull-Weibull distribution

The selected mixed Weibull-Weibull is a time-dependent distribution. This distribution also describes how the earthquake inter-event times in the northern part of the DSFZ can be divided into two clusters. The first cluster begins shortly after an earthquake occurrence and lasts about 80 years, and the second one begins about 80 years after an earthquake occurrence and lasts until the next earthquake occurrence. It should be mentioned that the general shape of this hazard function does not change by changing the parameters of the weight function (see figure 5.10). The hazard function (statistical) of the mixed Weibull-Weibull distribution is also time-dependent. Equations 5.8 and 5.9 explain the hazard function of this distribution:

$$h(t) = \frac{f(t)}{1 - F(t)} \quad (5.8)$$

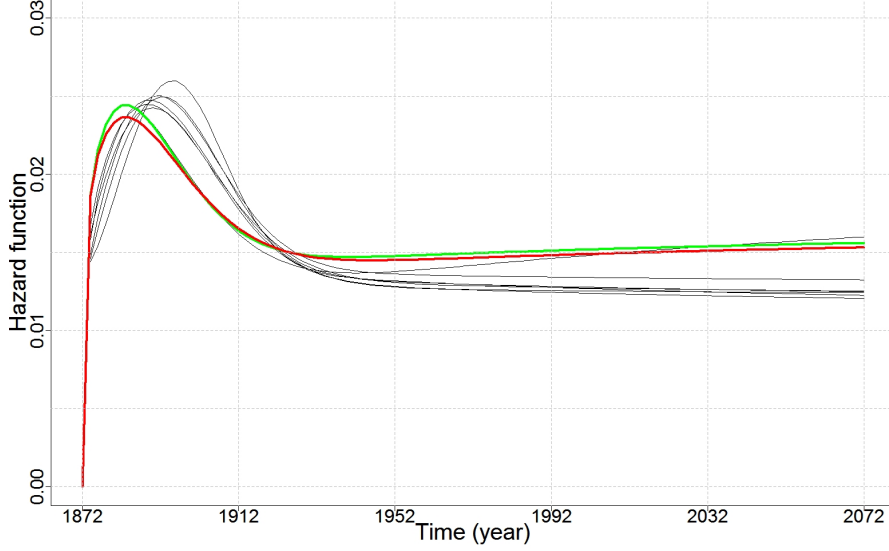


Figure 5.10: Various hazard functions corresponding to mixed Weibull-Weibull distributions, calculated based on different weight function parameters, within 200 years after the last earthquake in 1872. The red curve is based on the weight function parameters considered in this thesis. The green curve is based on no weights. Changing the weight function parameters does not affect the general shape of the hazard function.

with

$$f(t) = \rho \left( \frac{\beta_1}{\eta_1} \right) \left( \frac{t}{\eta_1} \right)^{\beta_1-1} \exp \left( - \left[ \frac{t}{\eta_1} \right]^{\beta_1} \right) + (1 - \rho) \left( \frac{\beta_2}{\eta_2} \right) \left( \frac{t}{\eta_2} \right)^{\beta_2-1} \exp \left( - \left[ \frac{t}{\eta_2} \right]^{\beta_2} \right) \quad (5.9)$$

and,

$$F(t) = \int_0^t f(x) dx \quad (5.10)$$

where  $(\rho, \eta_1, \beta_1, \eta_2, \beta_2) \approx (0.28, 17.44, 1.33, 73.63, 1.06)$  is the vector of the estimated parameters. Figure 5.11 represents the shape of the hazard function for different time spans after the last  $M_w \geq 6$  earthquake in 1872 in this area.

According to figure 5.11, the hazard increases immediately after an earthquake occurrence. After about ten years, it reaches the maximum value of  $\sim 0.024$ . Then

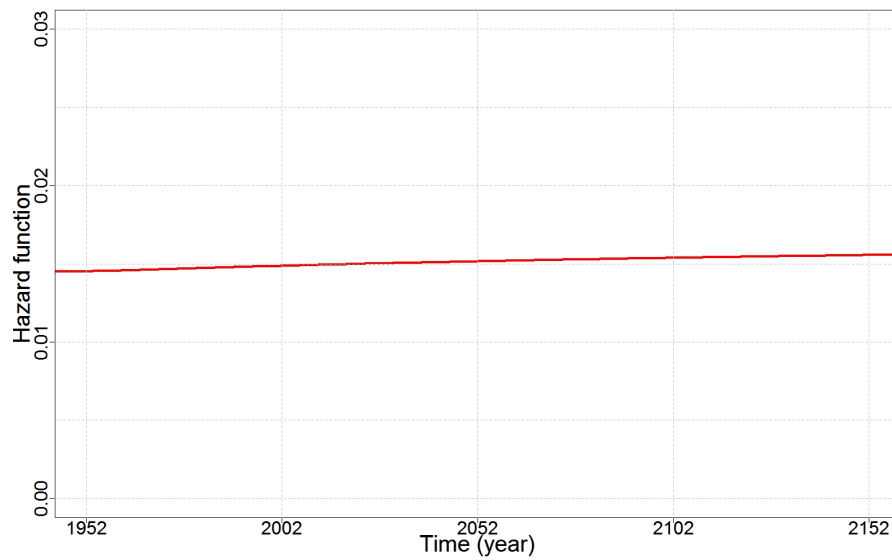
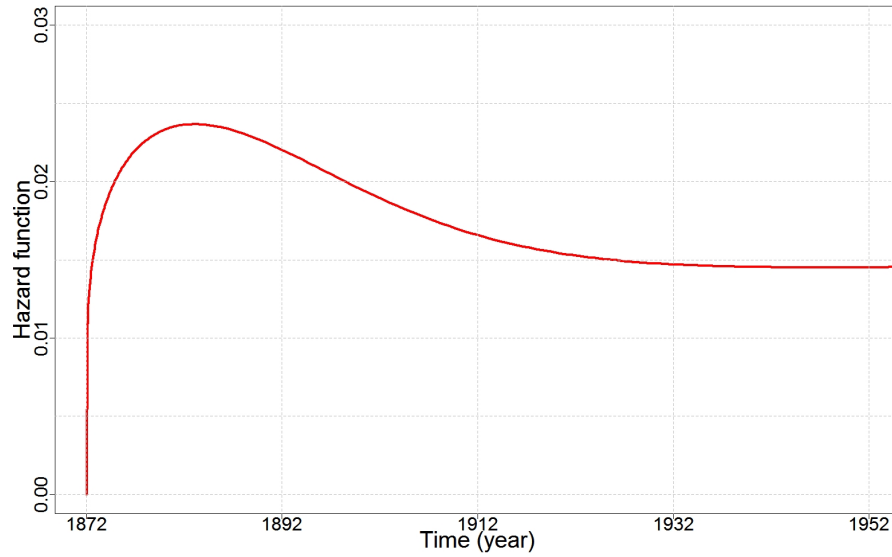


Figure 5.11: The trend of the hazard function of mixed Weibull-Weibull distribution after the last earthquake in 1872. Top: The first 80 years between 1872 and 1952. Bottom: Between 1952 and 2152.

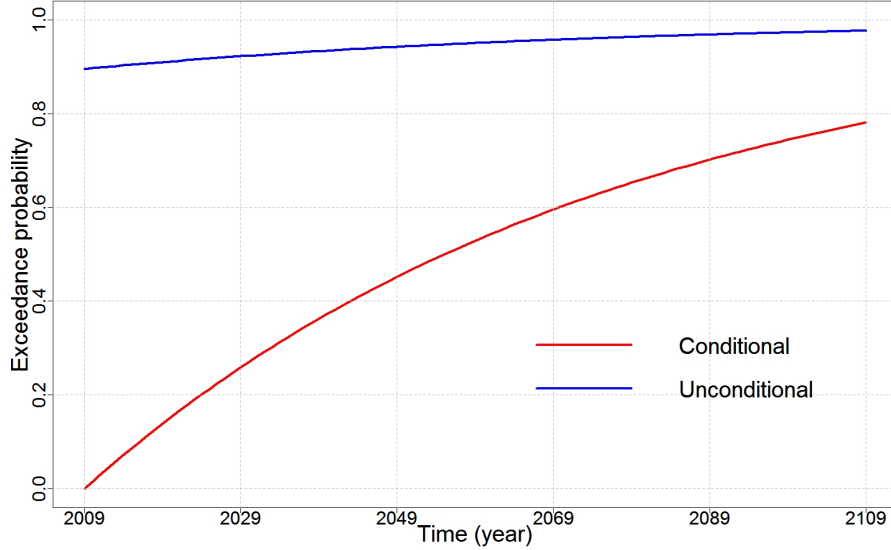


Figure 5.12: Conditional (red) and unconditional (blue) exceedance probabilities for the northern part of the DSFZ.

the hazard decreases within the next 70 years and reaches the minimum value of  $\sim 0.0145$ . At this point, it starts again to increase very slowly and reaches the amount of  $\sim 0.0149$  in 2009. Starting from now, the hazard function remains almost constant at about 0.015 ( $\sim 0.0154$  in 2100 and  $\sim 0.0157$  in 2200). This behavior of the hazard function confirms the long-term cluster-property of the large earthquakes inter-event times in the northern part of the DSFZ.

### Conditional exceedance probability

According to equation 5.2, the conditional exceedance probability within time  $t$  in the northern part of the DSFZ can be calculated as:

$$Pr \{T \leq t \mid T \geq t_c\} = \frac{F(t) - F(t_c)}{1 - F(t_c)} \quad (5.11)$$

where the function  $F(\cdot)$  is the same cdf as in equation 5.10.

By plugging  $t_c = 137$  into equation 5.11, the conditional exceedance probability has been calculated for the north part of the DSFZ. Figure 5.12 presents the conditional and unconditional exceedance probability. The conditional exceedance probability within the next 30 years is  $\sim 0.36$  in the northern part of the DSFZ.

## 5.2.6 Application of mixed Weibull-Weibull and Lognormal distributions in probabilistic seismic hazard assessments

The time-dependent occurrence rates corresponding to the mixed Weibull-Weibull; i.e. the selected model for the northern part of the DSFZ, and the Lognormal distributions, have been applied to a PSHA for the DSFZ and neighboring areas. The PSHA based on the Lognormal distribution is considered only in order to be compared with the PSHA based on the mixed Weibull-Weibull distribution. The details of the hazard calculation and the comparison with the classical time-independent approach are presented by Grünthal et al. (2009c). The hazard maps are calculated for 10% exceedance probability in 50 years. The results for the northern part of the DSFZ are shown and discussed here for different time periods within 1872 - 2100, since this part of the DSFZ requires a time-dependent approach. The qualitative differences discussed in this chapter for the two statistical distributions are confirmed in the hazard maps.

The obtained mean hazard maps are presented in Appendix B. The hazard maps are calculated for five different years; 1873, 1875, 1883, 2009 and 2100 after the last large earthquake in 1872 in the considered source zone (see Appendix B, figure B.1). The time-dependent portion of the hazard maps have been prepared considering earthquakes with magnitude  $M_w \geq 6$ . In Appendix B, the presented hazard maps are shown based on the mixed Weibull-Weibull distribution, the unimodal Lognormal distribution, and the classical time-independent model; i.e. exponential distribution based on the Poisson process. The hazard maps based on the exponential distribution do not change with time. More details about the applied PSHA approach can be found in the scientific technical report by Grünthal et al. 2009c.

The hazard maps for the northern part of the DSFZ in the year 1873 (one year after the last large earthquake) show that the seismic hazard based on the mixed Weibull-Weibull distribution is larger than the seismic hazard calculated by the Lognormal model and the exponential (time-independent under Poisson process) distribution. In the year 1875, the seismic hazard based on the mixed Weibull-Weibull distribution is between the seismic hazard based on the Lognormal and based on the exponential distribution. In the year 1883 (about 11 years after the last large event), the seismic hazard based on the mixed Weibull-Weibull is much larger than the seismic hazard based on the exponential distribution, but still less than the seismic hazard based on the Lognormal distribution. In the year 2009 (current time, about 137 years after the last large earthquake), the seismic hazard based on the mixed Weibull-Weibull distribution is slightly less than the seismic hazard based on the exponential distribution, while the seismic hazard based on the Lognormal distribution is much lower. In the year 2100, the seismic hazard based on the mixed Weibull-Weibull is almost equal to the seismic hazard based

on the exponential distribution; whereas the seismic hazard based on the Lognormal distribution is extremely low. Finally, it can be concluded that hazard maps for the northern part of the DSFZ show considerable differences in seismic hazard calculated based on the two time-dependent distributions; i.e. the mixed Weibull-Weibull and the Lognormal distribution, and the classical time-independent distribution; i.e. the exponential distribution, within the first cluster; i.e. 80 years after an earthquake occurrence. However, the seismic hazard calculated based on the mixed Weibull-Weibull and the exponential distribution are only slightly different during the second cluster; i.e. more than 80 years after an earthquake, where the seismic hazard calculated based on the Lognormal distribution continuously decreases with time.

It should be noticed that the time-dependent seismic hazard depends strongly on the applied time-dependent statistical model (distribution). The time-dependent approach developed in this thesis can be applied to every fault zone or seismic region. However, the choice of the best statistical distribution of the earthquake inter-event times can change from region to region. In other words, there is still no universal statistical distribution to model the earthquake inter-event times for large earthquakes in every fault zone or seismic region. This problem is caused by the lack of knowledge about the earthquake recurrence processes. The modified weighted maximum likelihood estimation method developed in this thesis is considered to use as much information as possible. However, the problem of the quantification of *epistemic* uncertainties (Mc Guire 2004) coupled to the earthquake information remains open. The epistemic uncertainties include the two cases of uncertainties mentioned above; i.e. the lack of knowledge about earthquake recurrence processes as well as the lack of earthquake information for large earthquakes.

# Chapter 6

## Summary and conclusions

### 6.1 Summary

In this thesis, the behavior of the occurrence rates of large earthquakes in the DSFZ has been analyzed using the historical earthquake catalog. The main challenge has been to investigate whether the occurrence rate of large earthquakes in this area is time-dependent and how. Therefore, a set of historical  $M_w \geq 6$  earthquakes has been considered as the main dataset of large earthquakes in the DSFZ. The study area contains the earthquakes within an average width of about 40 km along the DSFZ within  $29.5^\circ \leq \text{latitude} \leq 37^\circ$ . The completeness of this dataset has been determined to begin at 300 AD.

The dataset has been divided into three subareas including the southern, the central and the northern zones. The earthquake behavior has been studied for the central and the northern parts of the DSFZ, but not the southern part. Since only two large earthquakes have occurred here since 300 AD, which are not sufficient for a reliable statistical analysis. In order to apply the time-to-failure modeling concept to analyze the occurrence rate of earthquakes, the earthquake inter-event times have been calculated for the central and northern parts. Then the problem is reduced to estimate the best fit of the earthquake inter-event times in both areas.

A statistical approach has been developed and applied to fit the inter-event times. The approach has been developed based on the assumption that earthquake inter-event times are i.i.d (independently identically distributed). This assumption leads to a renewal process assumption for the inter-event times.

In the next step, five candidate models have been selected: exponential, Weibull, Gamma, Lognormal, and inverse Gaussian (BPT) distributions. Here, the time-dependency describes whether the hazard function corresponding to a specific distribution is a function of time (time-dependent) or not (time-independent).

A multimodality test has been applied to each inter-event time dataset; i.e. the



central and northern part of DSFZ, to check whether the inter-event times are unimodally or multimodally distributed. The five distributions have been directly used for the case of unimodal distributions, while the four time-dependent distributions have been considered as basic models involved in the structure of multimodal distributions, where multimodality is applicable.

A modified maximum likelihood method has been developed to estimate the parameters of the different distributions. Two important terms have been added to the classic method of MLE:

- **The censored time**

Not only the inter-event times but also the censored time (i.e. the time since the last event) have been considered in this method. In general, the censored time is important in time-to-failure modeling, since it includes the most recent information related to the current situation of a given time-to-failure process. In the case of earthquake inter-event time modeling, where the lack of reliable information is a common problem, the censored time is of major importance.

- **Weighting information**

Two reasons lead to the use of a weighted MLE. The first is the intrinsic uncertainty of historical earthquake data. Uncertainties in the earthquake parameters, especially for the older events, are neither recognizable, nor calculable. The second reason is the fact that earthquake rates can change with time<sup>1</sup>. Therefore, the estimation process requires larger weights for earthquakes occurring in the recent time.

The best distribution has been selected according to two criteria, a BIC and a KS-test p-value. Finally, the confidence intervals have been calculated by a bootstrap confidence interval method for each parameter corresponding to the selected models.

### **Results for the central part of DSFZ**

The distribution of the earthquake inter-event times in the central part of the DSFZ is considered as unimodal. The selected model in this area is a unimodal exponential distribution (Poissonian process). The hazard function corresponding to this distribution is time-independent.

### **Results for the northern part of the DSFZ**

The distribution of the earthquake inter-event times in the northern part of the DSFZ is considered as bimodal. The selected model in this area is a bimodal

---

<sup>1</sup>This is obvious, at least in the case of this thesis, in the northern part of the DSFZ.

Weibull-Weibull mixed distribution. The hazard function corresponding to this distribution is time-dependent.

## 6.2 Conclusions

### Efficiency of the approach

The previous approaches for long-term time-dependent modeling of earthquake occurrence rates (adapted for main shocks) are not sufficient. They consider either very simple unimodal statistical distributions or nonparametric methods. The first case; i.e. unimodal statistical distributions, cannot fit the data in the case of clustered data. The second group; i.e. nonparametric methods, is more data oriented and needs more reliable data, which are, often, very scarce in the case of large earthquakes. Many of these approaches do not use significant statistical methods for parameter estimation and model selections. Hence, a significant approach to analyze the earthquake occurrence rate is still missing.

In this thesis, a statistical approach has been developed in order to analyze the earthquake occurrence rate for large earthquakes. This approach has been shown to model the data efficiently. This approach

- uses all available temporal information of large historical earthquakes by considering both inter-event times and censoring time.
- uses a new method to determine completeness times for a set of large earthquakes.
- produces more reliable results, since it considers a weighting system where more reliable data are assigned higher level of importance in the estimation process.
- investigates whether the occurrence rates of large earthquakes are significantly time-dependent or if a time-independent model is sufficient.
- checks if the data are multimodally distributed; i.e. temporally clustered in the long-term, or unimodally distributed; i.e. no long-term temporal clustering in the data.
- considers various forms of distributions, respectively hazard functions, which makes it more flexible to fit the entire dataset. All these distributions have been commonly used in other scientific research areas.
- uses significant statistical methods (a modified MLE and an EM) to estimate the model parameters.

- selects the best model using two methods, a Bayesian Information Criterion which selects the best model considering the average, and a Kolmogorov-Smirnov goodness-of-fit test (a modified case in this thesis) selects the best model which selects the best model considering the details.
- can significantly fit different clusters of the data, in the case of clustering data.

### **Central part of the DSFZ**

The earthquake inter-event times in the central part of the DSFZ show no clustering property. The bootstrap multimodality test confirms this statement.

The earthquake occurrence rate in this area is assumed to be time-independent, since the time-dependent models are not significantly better than the time-independent model (exponential distribution under a Poissonian process). Although the most significant time-dependent model (the Weibull distribution in this case) is slightly better, taking into account the KS-test (with p-value 0.97 against 0.92 for the exponential distribution), the time-independent model is chosen since it is more simple and has only one parameter (against two parameters for the Weibull distribution). The following can be concluded from the results of the approach in this area:

- There is no significant difference between time-dependent and time-independent models, neither in KS-test nor in BIC.
- The earthquake occurrence rate can be considered as time-independent in the central part of DSFZ.

### **Northern part of the DSFZ**

In the northern part of the DSFZ, the inter-event times show a clustering property. The bootstrap multimodality test does not reject the clustering property in this area. On the other hand, the results of applying the approach to the inter-event times in this area show a very explicit time-dependency for the earthquake occurrence rate. Therefore, both multimodal and unimodal models have been applied to the dataset. A mixed Weibull-Weibull model has been selected as the best model in this area. This model produces a very high KS-test p-value (more than 0.97) and its BIC is relative low. Its statistical hazard function increases in the long-term. Therefore, the model is in agreement with the long-term stress increase.

The hazard function corresponding to the mixed Weibull-Weibull distribution reveals two clusters for the earthquake occurrence rate within the time after an earthquake occurrence. The first cluster begins immediately after an earthquake occurrence and is explicitly time-dependent. This cluster lasts about 80 years. The second cluster begins 80 years after an earthquake occurrence and lasts until the next

earthquake occurs. The hazard function corresponding to this cluster increases very slowly with time, such that it can be considered as a time-independent exponential distribution with rate  $\lambda \approx 0.015$ .

The following can be concluded for this area:

- The earthquake occurrence rate is significantly time-dependent.
- The mixed Weibull-Weibull distribution explains the data better than other candidate distributions, because the two different components of the model can fit the two different clusters of the inter-event times better than unimodal models.
- The rate of earthquake occurrence is considerably lower immediately after an earthquake, increases strongly during the following ten years and reaches its maximum of about 0.024, then decreases over the next 70 years to its minimum of about 0.0145.
- More than 80 years after an earthquake, the earthquake occurrence rate increases extremely slowly, such as it can be considered as an almost constant rate about 0.015.
- The classical Poissonian method (time-independent) is not sufficient to explain the earthquake inter-event times shortly after an earthquake (up to 80 years). However, for longer time spans, especially for an overdue earthquake, it would be sufficient.
- The hazard maps for the northern part of the DSFZ show considerable differences between the time-dependent and the classical (time-independent) seismic hazard within the first cluster; i.e. 80 years after an earthquake occurrence. However, the time-dependent and the classical (time-independent) seismic hazard are only slightly different during the second cluster; i.e. more than 80 years after an earthquake.

### **Future studies**

The time-dependent seismic hazard depends strongly on the applied time-dependent statistical model (distribution). In this thesis, a time-dependent approach has been developed. This approach can be applied to every seismic region in the world. However, the obtained results; i.e. the selected distributions, will change in different seismic zones depending on the earthquake datasets connected to these regions. Up to now no universal statistical distribution exists that can model the earthquake inter-event times for large earthquakes in every fault zone or seismic region. Whereas it is not known if such a model exists, it is a fact that the problem

of lack of knowledge about the earthquake recurrence processes is still unsolved. A modified weighted maximum likelihood estimation method has been developed in this thesis. Although this method helps to consider more information about earthquake inter-event times, the problem of the quantification of epistemic uncertainties coupled to the earthquake information remains open. The epistemic uncertainties include the two cases of uncertainties mentioned above; i.e. the lack of knowledge about earthquake recurrence processes as well as the lack of information about inter-event times for large earthquakes. These two problems can be considered as subjects of future studies in the DSFZ and other seismic regions.

## References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6): 716-723, doi:10.1109/TAC.1974.1100705.
- Aldrich, J., 1997. R. A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science* 12 (3): 162-176, doi:10.1214/ss/1030037906.
- Anderson, T. W., Darling, D. A., 1952. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathematical Statistics* 23: 193-212, doi:10.1214/aoms/1177729437.
- Bak, P., Christensen, K., Danon, L., Scanlon, T., 2002. Unified Scaling Law for Earthquakes. *Physical Review Letters* 88 (17), doi:10.1103/PhysRevLett.88.178501.
- Barka, A., 1999. The 17 August 1999 Izmit Earthquake. *Science*, 285: 1858-1859.
- Beauval, C., Hainzl, S., Scherbaum, F., 2006. Probabilistic seismic hazard estimation in low-seismicity regions considering non-Poissonian seismic occurrence. *Geophysical Journal International*, 164: 543-550, doi:10.1111/j.1365-246X.2006.02863.x.
- Blackman, J., 1956. An Extension of the Kolmogorov Distribution. *Annals of Mathematical Statistics* 27 (2): 513-520.
- Borman, S., 2004. The expectation maximization algorithm - a short tutorial. Available at [http://www.seanborman.com/publications/EM\\_algorithm.pdf](http://www.seanborman.com/publications/EM_algorithm.pdf).
- Burkhard, M., Grünthal, G., 2009. Seismic source zone characterization for the seismic hazard assessment project PEGASOS by the Expert Group 2 (EG1b). *Swiss Journal of Geosciences* 102: 149-188, doi: 10.1007/s00015-009-1307-3.
- Cherkassky, V. S., Mulier, F., 1998. *Learning from Data: Concepts, Theory, and Methods*. John Wiley & Sons, Inc., New York, NY.
- Chikara, R. S., Folks, J. L., 1989. *The Inverse Gaussian Distribution*. Marcel Dekker, New York.
- Console, R., Murru, M., Catalli, F., 2006. Physical and stochastic models of earthquake clustering. *Tectonophysics* 417: 141-153.
- Cornell, C. A., 1968. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 58: 1583-1606.
- Corral, A., 2003. Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Physical Review E* 68, 035102.
- Corral, A., 2004a. Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes. *Physical Review Letter* 92 (10): 1-4.
- Corral, A., 2004b. Universal local versus unified global scaling laws in the statistics of seismicity. *Physica A* 340: 590-597.
- Courtillot, V., Armijo, R., Tapponnier, P., 1987. The Sinai Triple Junction Revisited. *Tectonophysics* 141: 181-190.
- Cox, D. R., 1972. *Regression Models and Life Tables (with Discussion)*. *Journal of the Royal Statistical Society, Series B* 34: 187-220.
- Cox, D. R., Lewis, P. A. W., 1966. *The Statistical Analysis of Series of Events*. Methuen's Monographs on Applied Probability and Statistics. John Wiley, London. Viii + 285 pp.
- Csiszár, I., Shields, P., 1999. Consistency of the BIC order estimator. *Electronic Research Announcements of the American Mathematical Society* 5: 123-127.
- Daley D. J., Vere-Jones D., 2002. *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods, Second Edition*. Springer.
- Dellaert F., 2002. *The Expectation Maximization Algorithm*. College of Computing, Georgia Institute of Technology, Technical Report number GIT-GVU-02-20.

- Dempster, A. P., Laird, N. M., Rubin, D. B., 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)* 39 (1): 1-38.
- Dieterich, J. H., 1986. A model for the nucleation of earthquake slip. In: Das, S., Boatwright, J., Scholz, C.H. (Eds.). *Earthquake Source Mechanics, Geophysical Monograph 37, Maurice Ewing Series, 6, American Geophysics Union, Washington, DC: 37-47.*
- Dieterich, J. H., 1992. Earthquake nucleation on faults with rate-and-state-dependent strength. *Tectonophysics* 211: 115-134.
- Dieterich, J. H., 1994. A constitutive law for rate of earthquake production and its application to earthquake clustering. *Journal of Geophysical Research* 99: 2601-2618.
- Dieterich, J. H., 1995. Earthquake simulations with time-dependent nucleation and long-range interactions. *Nonlinear Processes in Geophysics* 2: 109-120.
- Dieterich, J.H., Kilgore, B.D., 1996. Imaging surface contacts: power law contact distributions and contact stresses in quartz, calcite, glass and acrylic plastic. *Tectonophysics* 256: 219-239.
- Dutton, C. E., 1904. *Earthquakes in the Light of the New Seismology*. New York: G. P. Putnam's Sons, 314 pp.
- Eadie, W. T., Dryard, D., James, F. E., Roos, M., Sadoulet, B., 1971. *Statistical Methods in Experimental Physics*. North Holland, Netherlands.
- Efron, B., Tibshirani, R., 1993. *An introduction to the bootstrap*. New York: Chapman and Hall.
- Efron, B., 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia.
- Ellsworth, W. L., Matthews, M. V., Nadeau, R. M., Nishenko, S. P., Reasenber, P. A., Simpson, R. W., 1999. A physically-based earthquake recurrence model for estimation of long-term earthquake probabilities. *United States Geological Survey Open-File Report 99-552*, 22 p.
- Ellsworth, W. L., 1995. Characteristic earthquakes and long-term earthquake forecasts: implications of central California seismicity. In F. Y. Cheng and M. S. Sheu (Editors), *Urban Disaster Mitigation: The Role of Science and Technology*, Elsevier Science Ltd., Oxford.
- Evison, F. F., D. A. Rhoades, 2001. Model of long-term seismogenesis. *Annali di Geofisica* 44: 81-93.
- Faenza, L., Marzocchi, W., Boschi, E., 2003. A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the Italian catalogue. *Geophysical Journal International* 155 (2): 521-531.
- Faenza, L., Hainzl, S., Scherbaum, F., Beauval, C., 2007. Statistical analysis of time-dependent earthquake occurrence and its impact on hazard in the low seismicity region Lower Rhine Embayment. *Geophysical Journal International* 171 (2): 797-806, doi:10.1111/j.1365-246X.2007.03564.x.
- Faenza, L., Marzocchi, W., Serretti, P., Boschi, E., 2008. On the spatio-temporal distribution of M 7.0+ worldwide seismicity with a non-parametric statistics. *Tectonophysics* 449: 97-104.
- Felzer, K. R., Becker, T. W., Abercrombie, R. E., Ekström, G., Rice, J. R., 2002. Triggering of the 1999 MW 7.1 Hector Mine earthquake by aftershocks of the 1992 MW 7.3 Landers earthquake. *Journal of Geophysical Research (Solid Earth)* 107 (B9), doi:10.1029/2001JB000911.
- Felzer, K. R., Abercrombie, R. E., Ekström, G., 2003. Secondary Aftershocks and Their Importance for Aftershock Forecasting. *Bulletin of the Seismological Society of America* 93 (4): 1433-1448.
- Fréchet, M., 1927. Sur la loi de probabilité de l'cart maximum. *Annales de la Société Polonaise de Mathématique, Cracovie* 6: 93-116.
- Garfunkel, Z., Zak, I., Freund, R., 1981. Active Faulting in the Dead Sea Rift. *Tectonophysics* 80: 1-26.

- Garfunkel, Z., 1981. Internal structure of the Dead Sea leaky transform (rift) in relation to plate kinematics. *Tectonophysics* 80: 81-108.
- Gibbons, J. D., 1971. *Nonparametric Statistical Inference*. McGraw-Hill Book Company, New York.
- Gomberg, J., Belardinelli, M. E., Cocco, M., Reasenber, P. A., 2005. Time-dependent earthquake probabilities. *Journal of Geophysical Research* 110, doi:10.1029/2004JB003404.
- Gómez, J.B., Pacheco, A.F., 2004. The Minimalist Model of characteristic earthquakes as a useful tool for description of the recurrence of large earthquakes. *Bulletin of the Seismological Society of America* 94: 1960-1967.
- González, A., Gómez, J. B., Pacheco, A. F., 2006. Updating seismic hazard at Parkfield. *Journal of Seismology* 10: 131-135, doi:10.1007/s10950-005-9005-8.
- Grünthal, G., 1985. The up-dated earthquake catalogue for the German Democratic Republic and adjacent areas - statistical data characteristics and conclusions for hazard assessment. 3rd International Symposium on the Analysis of Sismicity and Seismic Risk, Liblice/Czechoslovakia, 17-22 June 1985 (Proceeding Vol. I, 19-25).
- Grünthal, G., Wahlström, R., 2003. An  $M_w$  based earthquake catalogue for central, northern and northwestern Europe using a hierarchy of magnitude conversions. *Journal of Seismology* 7 (4): 507-531.
- Grünthal, G., Wahlström, R., 2009. A harmonized seismicity data base for the Euro-Mediterranean region. *Cahiers du Centre Europeen de Godynamique et de Sismologie*, Proceedings of the 27th ECGS Workshop "Seismicity Patterns in the Euro-Med Region", Luxembourg (in press).
- Grünthal, G., Wahlström, R., Stromeyer, D., 2009a. The unified catalogue of earthquakes in central, northern, and northwestern Europe (CENEC) - updated and expanded to the last millennium. *Journal of Seismology*, doi:10.1007/s10950-008-9144-9.
- Grünthal, G., Bosse, C., Stromeyer, D., 2009b. Die neue Generation der probabilistischen seismischen Gefährdungseinschätzung der Bundesrepublik Deutschland, Version 2007 mit Anwendung für die Erdbeben-Lastfälle der DIN 19700:2004-07 "Stauanlagen". Scientific Technical Report STR09/07, GFZ German Research Centre for Geosciences.
- Grünthal, G., Hakimhashemi, A. H., Schelle, H., Bosse, C., Wahlström, R., 2009c. Study of the long-term temporal behaviour of the seismicity of the Dead Sea Fault Zone and its implication for time-dependent seismic hazard assessments. Scientific Technical Report STR09/09, GFZ German Research Centre for Geosciences.
- Guo Z., Ogata Y., 1997. Statistical relations between the parameters of aftershocks in time, space and magnitude. *Journal of Geophysical Research* 102: 2857-2873.
- Gutenberg, R., Richter, C. F., 1944. Frequency of earthquakes in California. *Bulletin of the Seismological Society of America* 34: 185-188.
- Hagiwara, Y., 1974. Probability of earthquake occurrence as obtained from a Weibull distribution analysis of crustal strain. *Tectonophysics* 23: 323-318.
- Hainzl, S., Scherbaum, F., Beauval, C., 2006. Estimating background activity based on interevent-time distribution. *Bulletin of the Seismological Society of America* 96: 313-320.
- Hanks, T.C., Kanamori, H., 1979. A moment magnitude scale. *Journal of Geophysical Research* 84 (B5): 2348-2350.
- Hardebeck, J. L., 2004. Stress Triggering and Earthquake Probability Estimates. *Journal of Geophysical Research* 109, B04310, doi:10.1029/2003JB002437.
- Hebden J. S., Stein, S., 2009. Time-dependent Seismic Hazard Maps for the New Madrid Seismic Zone and Charleston, South Carolina, Areas. *Seismological Research Letters* 80 (1): 12-20, doi:10.1785/gssrl.80.1.12.



- Heidbach, O., Ben-Avraham, Z., 2007. Stress evolution and seismic hazard of the Dead Sea fault system. *Earth and Planetary Science Letters* 257: 299-312.
- Helmstetter, A., Sornette, D., 2003. Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research (Solid Earth)* 108 (B10), 2457, doi:10.1029/2003JB002409 01.
- Hinzen, K., Oemisch, M., 2001. Location and magnitude from seismic intensity data of recent and historic earthquakes in the northern Rhine area, Central Europe. *Bulletin of the Seismological Society of America* 91: 40-56.
- Hogg, R., Craig, A., 1978. *Introduction to mathematical statistics*. Macmillan Publishing Co., Inc., New York, 4th edition.
- Jackson, D. D., Kagan, Y. Y., 1999. Testable earthquake forecasts for 1999. *Seismological Research Letters* 70: 393-403.
- Jacobs, A., Wong, E., Yuanxi, H., 2009. China Reports Student Toll for Quake. *New York Times*. <http://www.nytimes.com/2009/05/08/world/asia/08china.html>. Retrieved 2009-05-14.
- Janssen, C., Hoffmann-Rothe, A., Bohnhoff, M., Wetzell, H. -U., Matar, A., Khatib, M., 2007. Different styles of faulting deformation along the Dead Sea Transform and possible consequences for the recurrence of major earthquakes. *Journal of Geodynamics* 44: 6689.
- Joffe, S., Garfunkel, Z., 1987. Plate kinematics of the circum Red Sea a re-evaluation. *Tectonophysics* 141 (1-3): 5-22.
- Reasenberg, P. A., Jones, L. M., 1989. Earthquake hazard after a mainshock in California. *Science* 243: 1173-1176.
- Kagan, Y. Y., Knopoff, L., 1987. Statistical short-term earthquake prediction. *Science* 236: 1563-1567.
- Khair, K., Tsokas, G. N., Sawaf, T., 1997. Crustal structure of the northern Levant region: multiple source Werner deconvolution estimates for Bouguer gravity anomalies. *Geophysical Journal International* 128: 605-616.
- Kalbfleisch, J. D., Prentice, R. L., 1980. *The statistical analysis of failure time data*. New York: John Wiley & Sons.
- Kalbfleisch, J. D., 1985. *Probability and statistical inference*. 2nd edition, Springer-Verlag, New York.
- Kilb, D., Gombert, J., Bodin, P., 2002. Aftershock triggering by complete Coulomb stress changes. *Journal of Geophysical Research* 107(B4), 2060, doi:10.1029/2001JB000202.
- King, G. C. P., Stein, R. S., Lin, J., 1994. Static Stress Changes and the Triggering of Earthquakes. *Bulletin of the Seismological Society of America* 84: 935-953.
- Klinger, Y., Avouac, J. P., Abou Karaki, N., Dorbath, L., Bourles, D., Reyss, J. L. 2000a. Slip rate on the Dead Sea transform fault in northern Araba valley (Jordan). *Geophysical Journal International* 142: 755-768.
- Kuehn, N. M., Hainzl, S., Scherbaum, F., 2008. Non-Poissonian earthquake occurrence in coupled stress release models and its effect on seismic hazard. *Geophysical Journal International* 174 (2): 649-658.
- Lay, T., Kanamori, H., Ammon, C., Nettles, M., Ward, S., Aster, R., Beck, S., Bilek, S., Brudzinski, M., Butler, R., DeShon, H., Ekström, G., Satake, K., Sipkin, S., 2005. The Great Sumatra-Andaman Earthquake of December 26, 2004. *Science* 308: 1127-1133, doi:10.1126/science.1112250.
- Le Pichon, X., Gaulier, J. M., 1988. The rotation of Arabia and the Levant fault system. *Tectonophysics* 46: 369-406.
- Limpert, E., Stahel, W. A., Abbt, M., 2001. Lognormal distributions across the sciences: keys and clues. *Bioscience* 51 (5): 341-352.

- Llenos, A. L., McGuire, J. J., Ogata, Y., 2009. Modeling Seismic Swarms Triggered by Aseismic Transients. *Earth and Planetary Science Letters* 281: 59-69.
- Lombardi, A. M., Marzocchi, W., 2007. Evidence of clustering and non-stationarity in the time distribution of large worldwide earthquakes. *Journal of Geophysical Research* 112, B02303.
- Mahmoud, S., Reilinger, R., McClusky, S., Vernant, P., Tealeb, A., 2005. GPS evidence for northward motion of the Sinai Block: Implications for E. Mediterranean tectonics. *Earth and Planetary Science Letters* 238 (1-2): 217-224.
- Marcellini, A., 1997. Physical model of aftershock temporal behaviour. *Tectonophysics* 277 (1-3): 137-146. doi:10.1016/S0040-1951(97)00082-6.
- Marza, V. I., 2004. On the death toll of the 1999 Izmit (Turkey) major earthquake. ESC General Assembly Papers, Potsdam: European Seismological Commission.
- Matthews, M. V., Ellsworth, W. L., Reasenberg, P. A., 2002. A Brownian model for recurrent earthquakes. *Bulletin of the Seismological Society of America* 92: 2233-2250.
- McGuire, R. K., 2004. Seismic hazard and risk analysis. Earthquake Engineering Research Institute, MNO-10, 240 p.
- McLachlan, G. J., Krishnan, T., 1996. *The EM Algorithm and Extensions*. New York, Wiley.
- McLachlan, G. J., Peel, D. A., 2000. *Finite Mixture Models*. New York, Wiley.
- Meghraoui, M., Gomez, F., Sbeinati, R., Van der Woerd, J., Mouty, M., Darkal, A. N., Radwan, Y., Layyous, I., Al Najjar, H., Darawcheh, R., Hijazi, F., Al-Ghazzi, R., Barazangi, M., 2003. Evidence for 830 years of seismic quiescence from palaeoseismology, archaeoseismology, and historical seismicity along the Dead Sea fault in Syria. *Earth and Planetary Science Letters* 210: 35-52.
- Michael, A. J., Jones, L. M., 1998. Seismicity alert probabilities at Parkfield, California, revisited. *Bulletin of the Seismological Society of America* 88: 117-130.
- Michael, A. J., 2005. Viscoelasticity, postseismic slip, fault interactions, and the recurrence of large earthquakes. *Bulletin of the Seismological Society of America* 95 (5): 1594-1603, doi:10.1785/0120030208.
- Molchan, G. M., 2005. Interevent time distribution of seismicity: A theoretical approach. *Pure and Applied Geophysics* 162: 1135-1150.
- Molenberghs, G., Verbeke, G., 2005. *Models for Repeated Discrete Data*. New York, Springer-Verlag.
- Montazeri, A., Baradaran, H., Omidvari, S., Azin, S. A., Ebadi, M., Garmaroudi, G., Harirchi, A. M., Shariati, M., 2005. Psychological distress among Bam earthquake survivors in Iran: a population-based study. *BMC Public Health* 5 (4), doi:10.1186/1471-2458-5-4. PMID 15644145.
- Mucciarelli M., 2007. A test for checking earthquake aperiodicity estimates from small samples. *Natural Hazards and Earth System Sciences* 7: 399-404.
- Murthy D. N. P., Xie, M., Jiang, R., 2004. *Weibull Models*. New York, Wiley.
- Musson, R. M. W., Toro, G. R., Coppersmith, K. J., Bommer, J. J., Deichmann, N., Bungum, H., Cotton, F., Scherbaum, F., Abrahamson, N. A., 2005. Evaluating hazard results for Switzerland and how not to do it: A discussion of "Problems in the application of the SSHAC probability method for assessing earthquake hazards at Swiss nuclear power plants" by J-U Kigel. *Engineering Geology* 82 (1): 43-55.
- Nishenko, S. P., Buland, R., 1987. A generic recurrence interval distribution for earthquake forecasting. *Bulletin of the Seismological Society of America* 77: 1382-1399.
- Nishenko, S. P., 1985. Seismic potential for large and great interplate earthquakes along the Chilean and southern Peruvian margins of South America: a quantitative reappraisal. *Journal of Geophysical Research* 90: 3589-3615.

- Ogata, Y., Shimazaki, K., 1984. Transition from aftershock to normal activity. *Bulletin of the Seismological Society of America* 74 (5): 1757-1765.
- Ogata, Y., Jones, L. M., Toda, S., 2003. When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California. *Journal of Geophysical Research* 108 (B6), 2318, doi:10.1029/2002JB002009 (ESE1-12).
- Ogata, Y., Zhuang, J., 2006. Space-time ETAS models and an improved extension. *Tectonophysics* 413: 13-23.
- Ogata, Y., 1983. Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure. *Journal of Physics of the Earth* 31: 115-124.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association* 83 (401): 9-27.
- Ogata, Y., 1998. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* 50 (2): 379-402.
- Okada, Y., 1992. Internal deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America* 82: 1018-1040.
- Omori, F., 1894. On the aftershock of earthquakes. *Journal of the College of Sciences of the Imperial University of Tokyo* 7: 111-200.
- Parsons, T., 2002b. Post-1906 stress recovery of the San Andreas fault system from 3-D finite element analysis. *Journal of Geophysical Research* 107, 2162, doi:10.1029/2001JB001051.
- Parsons, T., 2004. Recalculated probability of M 7 earthquakes beneath the Sea of Marmara, Turkey. *Journal of Geophysical Research* 109 (B5), B05304, doi:10.1029/2003JB002667.
- Parsons, T., 2005. Significance of stress transfer in time-dependent earthquake probability calculations. *Journal of Geophysical Research* 110, B05S02.
- Rhoades, D. A., Evison, F. F., 2004. Long-range earthquake forecasting with every earthquake a precursor according to scale. *Pure and Applied Geophysics* 161: 47-71.
- Rhoades, D. A., Evison, F. F., 2005. Test of the EEPAS forecasting model on the Japan earthquake catalogue. *Pure and Applied Geophysics* 162: 1271-1290.
- Rhoades, D. A., Evison, F. F., 2006. The EEPAS forecasting model and the probability of moderate-to-large earthquakes in the Kanto region, central Japan. *Tectonophysics* 417: 119-130.
- Rhoades, D. A., 2007. Application of the EEPAS model to forecasting earthquakes of moderate magnitude in southern California. *Seismological Research Letters* 78 (1): 110-115.
- La Rocca, L., 2008. Bayesian Non-parametric Estimation of Smooth Hazard Rates for Seismic Hazard Assessment. *Scandinavian Journal of Statistics* 35 (3): 524-539.
- Rosin, P., Rammler, E., 1933. The Laws Governing the Fineness of Powdered Coal. *Journal of the Institute of Fuel* 7: 29-36.
- Rotondi, R., 1999. Statistical analysis of temporal variations of seismicity level in some Italian regions. *Natural Hazards* 19 (2-3): 139-150.
- Rotondi, R., Garavaglia, E., 2002. Statistical analysis of the completeness of a seismic catalogue. *Natural Hazards* 25 (3): 245-258.
- Ruina, A. L., 1983. Slip instability and state variable friction laws. *Journal of Geophysical Research* 88 (B12): 10359-10370.
- Rundle, J. B., Rundle, P. B., Donnellan, A., Li, P., Klein, W., Morein, G., Turcotte, D. L., and Grant, L., 2006. Stress transfer in earthquakes, hazard estimation and ensemble forecasting: Inferences from numerical simulations. *Tectonophysics* 413: 109-125.
- Salamon, A., Hofstetter, A., Garfunkel, Z., Ron, H., 1996. Seismicity of the eastern Mediterranean region: Perspective from the Sinai subplate. *Tectonophysics* 263: 293-305.
- Savage, J. C., 1991. Criticism of some forecasts of the national earthquake prediction evaluation council. *Bulletin of the Seismological Society of America* 81: 862-881.

- Savage, J. C., 1992. The uncertainty in earthquake conditional probabilities. *Geophysical Research Letters* 19: 709-712.
- Scholz, C. H., 1968. Experimental study of the fracturing process in brittle rock. *Journal of Geophysical Research* 73: 1447-1454.
- Shewchuk, J. R., 1994. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*. Edition 1-1/4. School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213.
- Shorak, G. R., Wellner, J. A., 1986. *Empirical Processes with Applications to Statistics*. Wiley.
- Stein, R. S., Barka, A., Dieterich, J., 1997. Progressive failure on the north anatolian fault since 1939 by earthquake stress triggering. *Geophysical Journal International* 128: 594-604.
- Stepp, J. C., 1972. Analysis of the completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard. *Proceedings of the International Conference on Microzonation for Safer Construction: Research and application*, Seattle, Washington, U.S.A., 64, 1189-1207.
- Stigler, S., 1973. Simon Newcomb, Percy Daniell, and the history of robust estimation 1885-1920. *Journal of the American Statistical Association* 68: 872-879.
- Tang, A. K., 2000. Izmit (Kocaeli), Turkey, earthquake of August 17, 1999 including Duzce Earthquake of November 12, 1999 Lifeline Performance. *Technical Council of Lifeline Earthquake Engineering Monograph No. 17*. Reston, VA : American Society of Civil Engineers, 2000, p. 259.
- Tanner, M. A., Wong, W. H., 1984. Data-based nonparametric estimation of the hazard function with applications to model diagnostics and exploratory analysis. *Journal of the American Statistical Association* 79: 123-456.
- Tapponnier, P., 1977. Evolution tectonique du système alpin en Méditerranée, poinçonnement et écrasement rigide-plastique. *Bulletin de la Societe Geologique de France* 3: 437-460.
- Taylor, H. M., Karlin, S., 1998. *An Introduction To Stochastic Modeling*. 3rd edition, Academic Press London 1998.
- Toda, S., Stein, R. S., Reasenberg, P. A., Dietrich, J. H., 1998. Stress transferred by the  $M_w = 6.5$  Kobe, Japan shock: effect on aftershocks and future earthquake probabilities. *Journal of Geophysical Research* 103: 24543- 24565.
- Tsai, A., Zhang, J., Willsky, A. S., 2001. Expectation-maximization algorithms for image processing using multiscale models and mean-field theory, with applications to laser radar range profiling and segmentation. *Optical Engineering* 40 (7): 1287-1301.
- Utsu, T., 1957. Magnitude of earthquakes and occurrence of their aftershocks. *Journal of the Seismological Society of Japan* 10: 35-45.
- Utsu, T., 1961. A statistical study on the occurrence of aftershocks. *Geophysical Magazine* 30: 521-605.
- Utsu, T., 1970. Aftershocks and earthquake statistics (II) - Further investigation of aftershocks and other waergquake sequences based on a new classification of earthquake sequences. *Journal of the Faculty of Science, Hokkaido Imperial University, Ser.VII, 3*: 197-266.
- Utsu, T., 1984, Estimation of parameters for recurrence models of Earthquakes. *Bulletin of the Earthquake Research Institute, University of Tokyo* 59: 53-66.
- Utsu, T., Ogata, Y., Matsuúra, R. S., 1995. The Centenary of the Omori formula for a decay law of aftershock activity. *Journal of Physics of the Earth* 43: 1-133.
- Vázquez-Prada, M., González, A., Gómez, J. B., Pacheco, A. F., 2002. A minimalist model of characteristic earthquakes. *Nonlinear Processes in Geophysics* 9: 513-519.
- Vere-Jones, D., 1978. Earthquake prediction - a statistician's view. *Journal of the Physics of the Earth* 25: 129- 146.

- Volinsky, C. T., Raftery, A., E., 1999. Bayesian information criterion for censored survival models. Technical Report 349, Department of Statistics, University of Washington (<http://www.stat.washington.edu/tech.reports/tr349.ps>).
- Weibull, W., 1951. A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, transactions of the ASME 18: 293-297.
- Wells, D. L., Coppersmith, K. L., 1994. New empirical relationships among magnitude, rupture width, rupture area, and surface displacement. *Bulletin of the Seismological Society of America* 84: 974-1002.
- Working Group on California Earthquake Probabilities, 1990, Probabilities of large earthquakes in the San Francisco Bay region, California, U.S. Geological Survey Circular 1053, 61p.
- Working Group on California Earthquake Probabilities. 2003. Earthquake Probabilities in the San Francisco Bay Region: 2002 to 2031 [R]. USGS Open-File Rept 03-214, 1 235.
- Zhuang, J., Ogata, Y., Vere-Jones, D., 2002. Stochastic Declustering of Space-Time Earthquake Occurrences. *Journal of the American Statistical Association* 97: 369-380.
- Zöller, G., Hainzl, S., Holschneider, M., 2008. Recurrent Large Earthquakes in a Fault Region: What Can Be Inferred from Small and Intermediate Events? *Bulletin of the Seismological Society of America* 98 (6): 2641-2651, doi: 10.1785/0120080146.

## Appendix A

The modified weighted log-likelihood function and its first and second partial derivatives for:

1 - the exponential distribution

$$\begin{aligned}\Lambda &= \sum_{i=1}^N w_i [\ln(\lambda) - \lambda t_i] - w_c \lambda t_c \\ \frac{\partial \Lambda}{\partial \lambda} &= \sum_{i=1}^N w_i \left[ \frac{1}{\lambda} - t_i \right] - w_c t_c \\ \frac{\partial^2 \Lambda}{\partial \lambda^2} &= \sum_{i=1}^N -\frac{w_i}{\lambda^2}\end{aligned}$$

2 - the Weibull distribution

$$\begin{aligned}\Lambda &= \sum_{i=1}^N w_i \left[ \ln(\beta) - \ln(\eta) + (\beta - 1) [\ln(t_i) - \ln(\eta)] - \left(\frac{t_i}{\eta}\right)^\beta \right] \\ &\quad - w_c \left(\frac{t_c}{\eta}\right)^\beta \\ \frac{\partial \Lambda}{\partial \beta} &= \sum_{i=1}^N w_i \left[ \frac{1}{\beta} + \ln(t_i) - \ln(\eta) - [\ln(t_i) - \ln(\eta)] \left(\frac{t_i}{\eta}\right)^\beta \right] \\ &\quad - w_c [\ln(t_c) - \ln(\eta)] \left(\frac{t_c}{\eta}\right)^\beta \\ \frac{\partial \Lambda}{\partial \eta} &= \sum_{i=1}^N w_i \left[ \frac{\beta}{\eta} \left( \left(\frac{t_i}{\eta}\right)^\beta - 1 \right) \right] + w_c \left(\frac{\beta}{\eta}\right) \left(\frac{t_c}{\eta}\right)^\beta \\ \frac{\partial^2 \Lambda}{\partial \beta^2} &= \sum_{i=1}^N w_i \left[ -\frac{1}{\beta^2} - [\ln(t_i) - \ln(\eta)]^2 \left(\frac{t_i}{\eta}\right)^\beta \right] \\ &\quad - w_c [\ln(t_c) - \ln(\eta)]^2 \left(\frac{t_c}{\eta}\right)^\beta \\ \frac{\partial^2 \Lambda}{\partial \eta^2} &= \sum_{i=1}^N w_i \left[ \frac{\beta}{\eta^2} \left( 1 - (1 + \beta) \left(\frac{t_i}{\eta}\right)^\beta \right) \right] + w_c \left(\frac{\beta}{\eta^2}\right) \left(\frac{t_c}{\eta}\right)^\beta (1 + \beta)\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \Lambda}{\partial \eta \partial \beta} &= \sum_{i=1}^N w_i \left[ \frac{1}{\eta} \left( \left( \frac{t_i}{\eta} \right)^\beta (1 + \beta [\ln(t_i) - \ln(\eta)]) - 1 \right) \right] \\ &\quad + w_c \left( \frac{1}{\eta} \right) \left( \frac{t_c}{\eta} \right)^\beta (1 + \beta [\ln(t_c) - \ln(\eta)]) \\ \frac{\partial^2 \Lambda}{\partial \beta \partial \eta} &= \frac{\partial^2 \Lambda}{\partial \eta \partial \beta}\end{aligned}$$

### 3 - the Gamma distribution

$$\begin{aligned}\Lambda &= \sum_{i=1}^N w_i \left[ (k-1) \ln(t_i) - \frac{t_i}{\theta} - k \ln(\theta) - \ln(\Gamma(k)) \right] \\ &\quad + w_c \left[ \ln \left( \Gamma \left( k, \frac{t_c}{\theta} \right) \right) - \ln(\Gamma(k)) \right]\end{aligned}$$

$$\frac{\partial \Lambda}{\partial k} = \sum_{i=1}^N w_i \left[ \ln(t_i) - \ln(\theta) - \frac{\Gamma_k(k)}{\Gamma(k)} \right] + w_c \left[ \frac{\Gamma_k \left( k, \frac{t_c}{\theta} \right)}{\Gamma \left( k, \frac{t_c}{\theta} \right)} - \frac{\Gamma_k(k)}{\Gamma(k)} \right]$$

$$\frac{\partial \Lambda}{\partial \theta} = \sum_{i=1}^N w_i \left( \frac{t_i}{\theta^2} - \frac{k}{\theta} \right) + w_c \left[ \frac{\frac{1}{\theta} \left( \frac{t_c}{\theta} \right)^k \exp \left( -\frac{t_c}{\theta} \right)}{\Gamma \left( k, \frac{t_c}{\theta} \right)} \right]$$

$$\begin{aligned}\frac{\partial^2 \Lambda}{\partial k^2} &= \left( \sum_{i=1}^N w_i + w_c \right) \left[ -\frac{\Gamma_{kk}(k) \Gamma(k) - \Gamma_k^2(k)}{\Gamma^2(k)} \right] \\ &\quad + w_c \left[ \frac{\Gamma_{kk} \left( k, \frac{t_c}{\theta} \right) \Gamma \left( k, \frac{t_c}{\theta} \right) - \Gamma_k^2 \left( k, \frac{t_c}{\theta} \right)}{\Gamma^2 \left( k, \frac{t_c}{\theta} \right)} \right]\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \Lambda}{\partial \theta^2} &= \sum_{i=1}^N w_i \left( -\frac{2t_i}{\theta^3} + \frac{k}{\theta^2} \right) \\ &\quad + w_c \left( \frac{1}{\theta^2} \right) \left( \frac{t_c}{\theta} \right)^k \exp \left( -\frac{t_c}{\theta} \right) \left[ \frac{\left( \frac{t_c}{\theta} - k - 1 \right) \Gamma \left( k, \frac{t_c}{\theta} \right) - \left( \frac{t_c}{\theta} \right)^k \exp \left( -\frac{t_c}{\theta} \right)}{\Gamma^2 \left( k, \frac{t_c}{\theta} \right)} \right]\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \Lambda}{\partial \theta \partial k} &= -\sum_{i=1}^N \frac{w_i}{\theta} \\ &\quad + w_c \left( \frac{1}{\theta} \right) \left( \frac{t_c}{\theta} \right)^k \exp \left( -\frac{t_c}{\theta} \right) \left[ \frac{[\ln(t_c) - \ln(\theta)] \Gamma \left( k, \frac{t_c}{\theta} \right) - \Gamma_k \left( k, \frac{t_c}{\theta} \right)}{\Gamma^2 \left( k, \frac{t_c}{\theta} \right)} \right]\end{aligned}$$

$\Gamma_k$  and  $\Gamma_{kk}$  are the first and the second derivatives of  $\Gamma$ .

$$\begin{aligned}\Gamma_k(k) &= \frac{d}{dk}\Gamma(k) = \Gamma(k)\psi_0(k) \\ \Gamma_{kk}(k) &= \frac{d}{dk}\Gamma_k(k) = \Gamma(k)\psi_1(k) + \frac{\Gamma_k^2(k)}{\Gamma(k)} \\ \Gamma_k(k, t) &= \frac{d}{dk}\Gamma(k, t) = \int_t^\infty x^{k-1} \ln(x) \exp(-x) dx \\ \Gamma_{kk}(k, t) &= \frac{d}{dk}\Gamma_k(k, t) = \int_t^\infty x^{k-1} \ln^2(x) \exp(-x) dx\end{aligned}$$

where  $\psi_0$  and  $\psi_1$  are *Digamma* and *Trigamma* functions respectively.

$$\begin{aligned}\psi_0(k) &= \frac{d}{dk} \log(\Gamma(k)) = \int_0^\infty \left( \frac{\exp(-x)}{x} - \frac{\exp(-kx)}{1 - \exp(-x)} \right) dx \\ \psi_1(k) &= \frac{d}{dk} \psi_0(k) = - \int_0^1 \frac{x^{k-1} \ln(x)}{1 - x} dx\end{aligned}$$

#### 4 - the inverse Gaussian (BPT) distribution

$$\begin{aligned}\Lambda &= \sum_{i=1}^N w_i \left( \frac{1}{2} [\ln(\lambda) - \ln(2\pi t_i^3)] - \frac{\lambda(t_i - \mu)^2}{2\mu^2 t_i} \right) + w_c \ln(S(t_c)) \\ \frac{\partial \Lambda}{\partial \lambda} &= \sum_{i=1}^N w_i \left( \frac{1}{2\lambda} - \frac{(t_i - \mu)^2}{2\mu^2 t_i} \right) + w_c \left[ \frac{S_\lambda(t_c)}{S(t_c)} \right] \\ \frac{\partial \Lambda}{\partial \mu} &= \sum_{i=1}^N w_i \left( \frac{\lambda t_i}{\mu^3} - \frac{\lambda}{\mu^2} \right) + w_c \left[ \frac{S_\mu(t_c)}{S(t_c)} \right] \\ \frac{\partial^2 \Lambda}{\partial \lambda^2} &= \sum_{i=1}^N -\frac{w_i}{2\lambda^2} + w_c \left[ \frac{S_{\lambda\lambda}(t_c)S(t_c) - S_\lambda^2(t_c)}{S^2(t_c)} \right] \\ \frac{\partial^2 \Lambda}{\partial \mu^2} &= \sum_{i=1}^N w_i \left( -\frac{3\lambda t_i}{\mu^4} + \frac{2\lambda}{\mu^3} \right) + w_c \left[ \frac{S_{\mu\mu}(t_c)S(t_c) - S_\mu^2(t_c)}{S^2(t_c)} \right] \\ \frac{\partial^2 \Lambda}{\partial \lambda \partial \mu} &= \sum_{i=1}^N w_i \left( \frac{t_i}{\mu^3} - \frac{1}{\mu^2} \right) + w_c \left[ \frac{S_{\lambda\mu}(t_c)S(t_c) - S_\lambda(t_c)S_\mu(t_c)}{S^2(t_c)} \right]\end{aligned}$$



where  $S(\cdot)$  is the survivor function of the BPT distribution,  $S_\lambda$ ,  $S_\mu$ ,  $S_{\lambda\lambda}$ ,  $S_{\mu\mu}$  and  $S_{\lambda\mu}$  are the first and second derivatives of  $S$  over  $\lambda$  and  $\mu$ , with:

$$\begin{aligned}
S(t_c) &= \frac{1}{2} \left[ 1 - \operatorname{erf} \left( \sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} - 1 \right] \right) \right] - \frac{e^{\frac{2\lambda}{\mu}}}{2} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
S_\lambda(t_c) &= -\frac{1}{\sqrt{2\pi\lambda t_c}} \left( \frac{t_c}{\mu} - 1 \right) \exp \left( -\frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \\
&\quad - \frac{e^{\frac{2\lambda}{\mu}}}{\mu} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
&\quad + \frac{1}{\sqrt{2\pi\lambda t_c}} \left( \frac{t_c}{\mu} + 1 \right) \exp \left( \frac{2\lambda}{\mu} - \frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right) \\
S_\mu(t_c) &= \frac{1}{\mu^2} \sqrt{\frac{\lambda t_c}{2\pi}} \exp \left( -\frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) + \frac{\lambda e^{\frac{2\lambda}{\mu}}}{\mu^2} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
&\quad - \frac{1}{\mu^2} \sqrt{\frac{\lambda t_c}{2\pi}} \exp \left( \frac{2\lambda}{\mu} - \frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right) \\
S_{\lambda\lambda}(t_c) &= \frac{1}{2\sqrt{2\pi\lambda t_c}} \left( \frac{t_c}{\mu} - 1 \right) \left( \frac{1}{\lambda} + \frac{1}{t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \exp \left( -\frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \\
&\quad - \frac{2e^{\frac{2\lambda}{\mu}}}{\mu^2} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
&\quad + \frac{1}{\sqrt{2\pi\lambda t_c}} \left( \frac{t_c}{\mu} + 1 \right) \left( \frac{3}{\mu} - \frac{1}{2\lambda} - \frac{1}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right) \exp \left( \frac{2\lambda}{\mu} - \frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right) \\
S_{\mu\mu}(t_c) &= \frac{1}{\mu^3} \sqrt{\frac{2\lambda t_c}{\pi}} \left( \frac{\lambda}{2\mu} \left[ \frac{t_c}{\mu} - 1 \right] - 1 \right) \exp \left( -\frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \\
&\quad - \frac{2\lambda(\lambda + \mu)e^{\frac{2\lambda}{\mu}}}{\mu^4} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
&\quad + \frac{1}{\mu^3} \sqrt{\frac{2\lambda t_c}{\pi}} \left( \frac{2\lambda}{\mu} + 1 - \frac{\lambda}{2\mu} \left[ \frac{t_c}{\mu} + 1 \right] \right) \exp \left( \frac{2\lambda}{\mu} - \frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right)
\end{aligned}$$

$$\begin{aligned}
S_{\lambda\mu}(t_c) &= \frac{1}{2\mu^2} \sqrt{\frac{\lambda}{2\pi t_c}} \left( \frac{t_c}{\lambda} - \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \exp \left( -\frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} - 1 \right]^2 \right) \\
&\quad + \frac{(2\lambda + \mu)e^{\frac{2\lambda}{\mu}}}{\mu^3} \left[ 1 + \operatorname{erf} \left( -\sqrt{\frac{\lambda}{2t_c}} \left[ \frac{t_c}{\mu} + 1 \right] \right) \right] \\
&\quad - \frac{1}{\mu^2} \sqrt{\frac{\lambda}{2\pi t_c}} \left( \frac{3t_c}{\mu} + 1 + \frac{t_c}{2\lambda} - \frac{1}{2} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right) \exp \left( \frac{2\lambda}{\mu} - \frac{\lambda}{2t_c} \left[ \frac{t_c}{\mu} + 1 \right]^2 \right)
\end{aligned}$$

## 5 - the Lognormal distribution

$$\begin{aligned}
\Lambda &= \sum_{i=1}^N w_i \left( -\ln(t_i) - \ln(\sigma) - \frac{1}{2} \ln(2\pi) - \frac{(\ln(t_i) - \mu)^2}{2\sigma^2} \right) + w_c \ln(S(t_c)) \\
\frac{\partial \Lambda}{\partial \mu} &= \sum_{i=1}^N w_i \left( \frac{\ln(t_i) - \mu}{\sigma^2} \right) + w_c \left[ \frac{S_\mu(t_c)}{S(t_c)} \right] \\
\frac{\partial \Lambda}{\partial \sigma} &= \sum_{i=1}^N w_i \left( \frac{(\ln(t_i) - \mu)^2}{\sigma^3} - \frac{1}{\sigma} \right) + w_c \left[ \frac{S_\sigma(t_c)}{S(t_c)} \right] \\
\frac{\partial^2 \Lambda}{\partial \mu^2} &= \sum_{i=1}^N -\frac{w_i}{\sigma^2} + w_c \left[ \frac{S_{\mu\mu}(t_c)S(t_c) - S_\mu^2(t_c)}{S^2(t_c)} \right] \\
\frac{\partial^2 \Lambda}{\partial \sigma^2} &= \sum_{i=1}^N w_i \left( \frac{1}{\sigma^2} - \frac{3(\ln(t_i) - \mu)^2}{\sigma^4} \right) + w_c \left[ \frac{S_{\sigma\sigma}(t_c)S(t_c) - S_\sigma^2(t_c)}{S^2(t_c)} \right] \\
\frac{\partial^2 \Lambda}{\partial \mu \partial \sigma} &= \sum_{i=1}^N \frac{-2w_i(\ln(t_i) - \mu)}{\sigma^3} + w_c \left[ \frac{S_{\mu\sigma}(t_c)S(t_c) - S_\sigma(t_c)S_\mu(t_c)}{S^2(t_c)} \right]
\end{aligned}$$

where  $S(\cdot)$  is the survivor function of the Lognormal distribution,  $S_\sigma$ ,  $S_\mu$ ,  $S_{\sigma\sigma}$ ,  $S_{\mu\mu}$  and  $S_{\mu\sigma}$  are the first and second derivatives of  $S$  over  $\sigma$  and  $\mu$ , with:

$$\begin{aligned}
S(t_c) &= \frac{1}{2} \left[ 1 - \operatorname{erf} \left( \frac{\ln(t_c) - \mu}{\sigma\sqrt{2}} \right) \right] \\
S_\mu(t_c) &= \frac{1}{\sigma\sqrt{2\pi}} \exp \left( \frac{-(\ln(t_c) - \mu)^2}{2\sigma^2} \right)
\end{aligned}$$

$$\begin{aligned}
S_{\sigma}(t_c) &= \frac{\ln(t_c) - \mu}{\sigma^2 \sqrt{2\pi}} \exp\left(\frac{-(\ln(t_c) - \mu)^2}{2\sigma^2}\right) \\
S_{\mu\mu}(t_c) &= \frac{\ln(t_c) - \mu}{\sigma^3 \sqrt{2\pi}} \exp\left(\frac{-(\ln(t_c) - \mu)^2}{2\sigma^2}\right) \\
S_{\sigma\sigma}(t_c) &= \frac{\ln(t_c) - \mu}{\sigma^3 \sqrt{2\pi}} \left[ \frac{(\ln(t_c) - \mu)^2}{\sigma^2} - 2 \right] \exp\left(\frac{-(\ln(t_c) - \mu)^2}{2\sigma^2}\right) \\
S_{\mu\sigma}(t_c) &= \frac{1}{\sigma^2 \sqrt{2\pi}} \left[ \frac{(\ln(t_c) - \mu)^2}{\sigma^2} - 1 \right] \exp\left(\frac{-(\ln(t_c) - \mu)^2}{2\sigma^2}\right)
\end{aligned}$$

## Appendix B

Time-dependent and time-independent hazard maps for the northern part of the DSFZ as discussed in section 5.2.6. For more details, see Grünthal et al. 2009c.

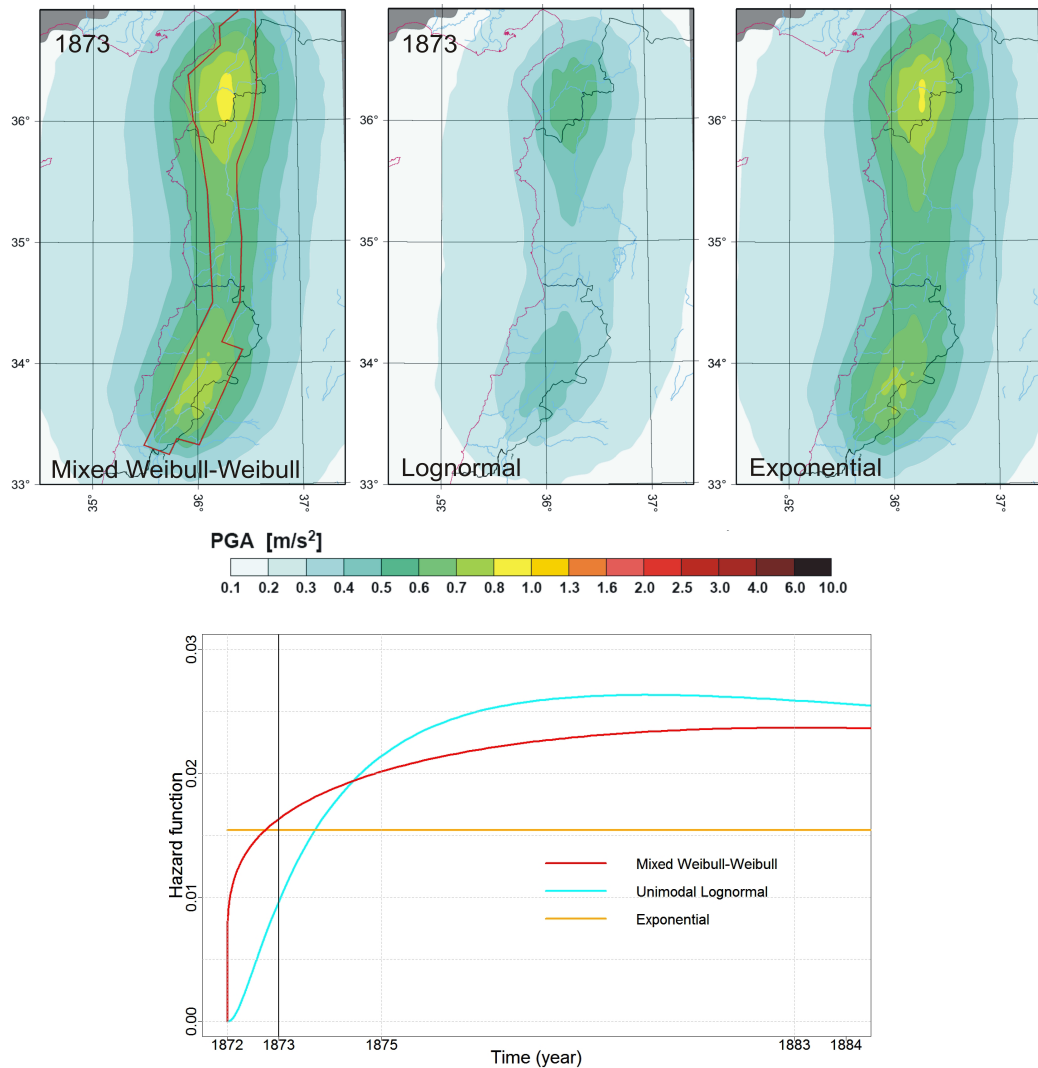


Figure B.1: Top: Hazard maps for the northern part of the DSFZ in the year 1873 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). The red polygon shows the boundary of the source zone. For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884.

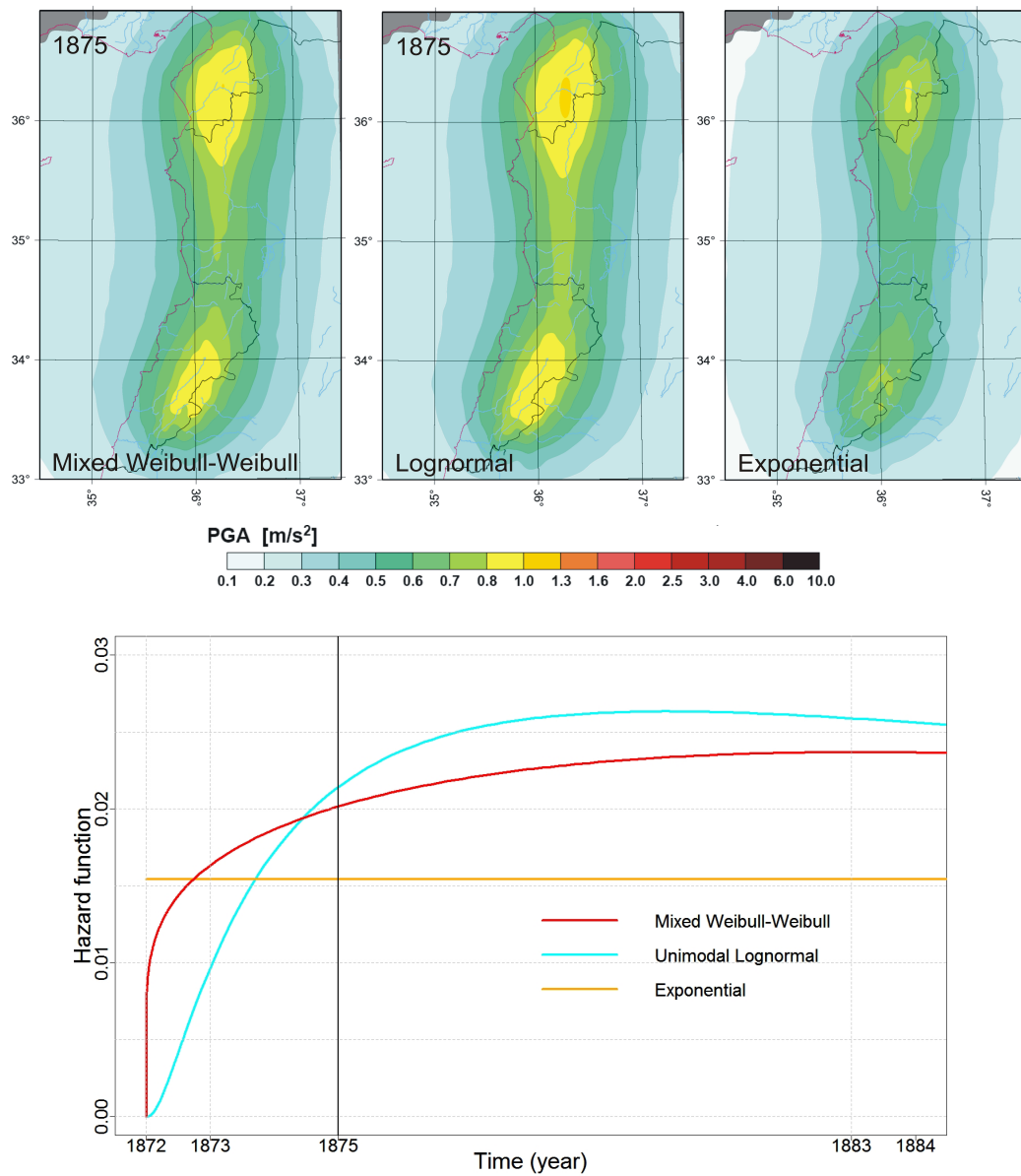


Figure B.2: Top: Hazard maps for the northern part of the DSFZ in the year 1875 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884.

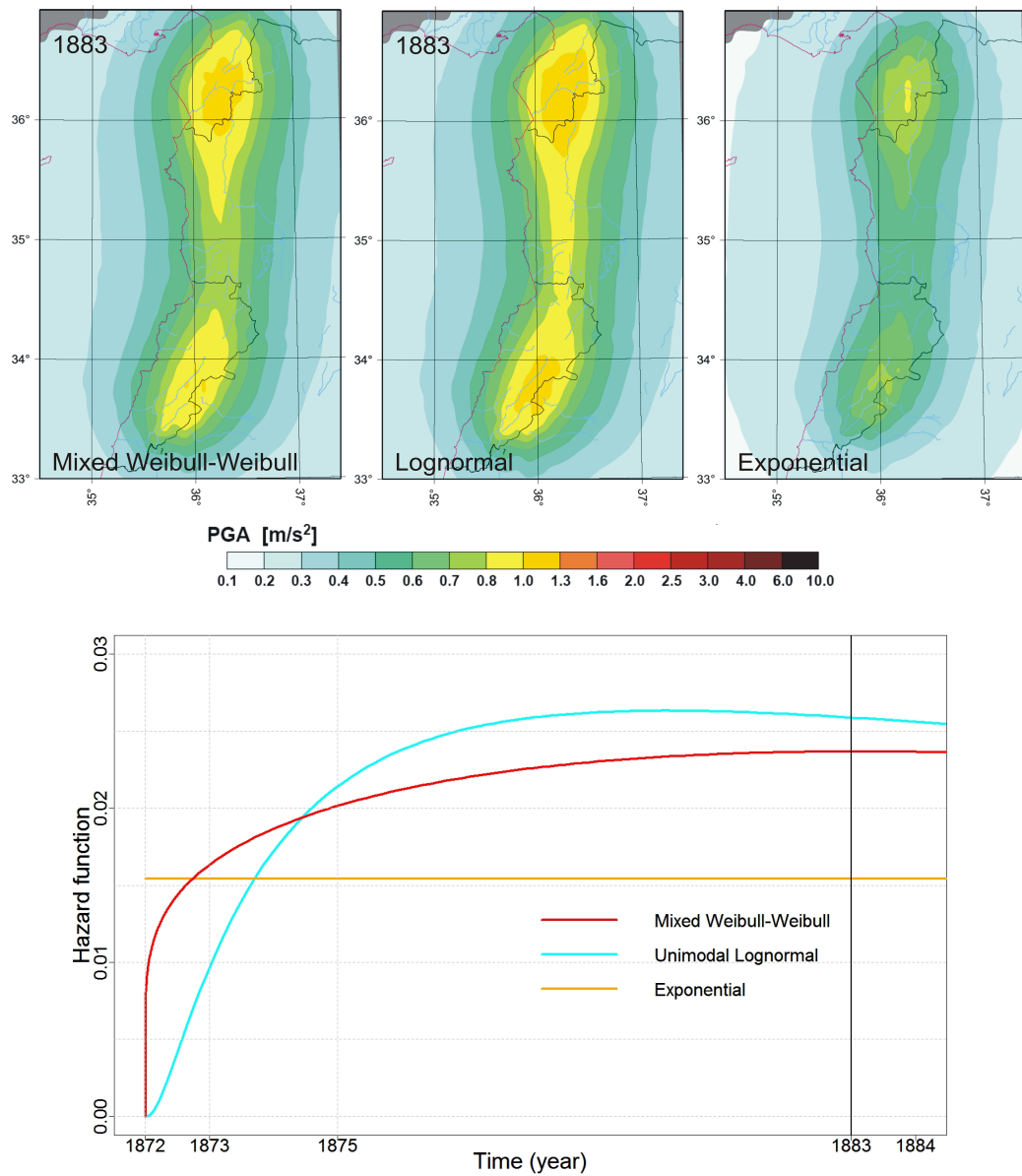


Figure B.3: Top: Hazard maps for the northern part of the DSFZ in the year 1883 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 1872 - 1884.

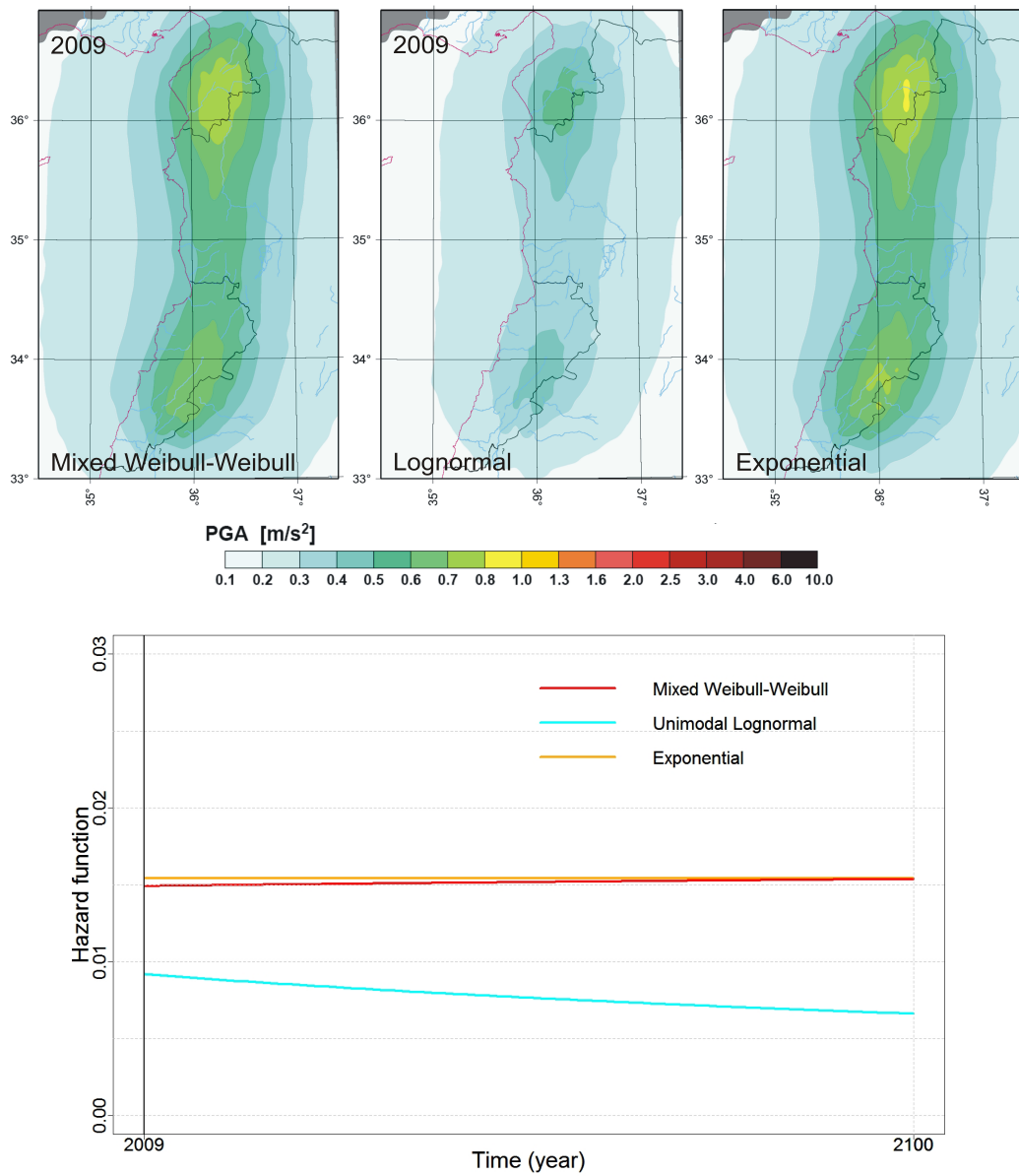


Figure B.4: Top: Hazard maps for the northern part of the DSFZ in the year 2009 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 2009 - 2100.

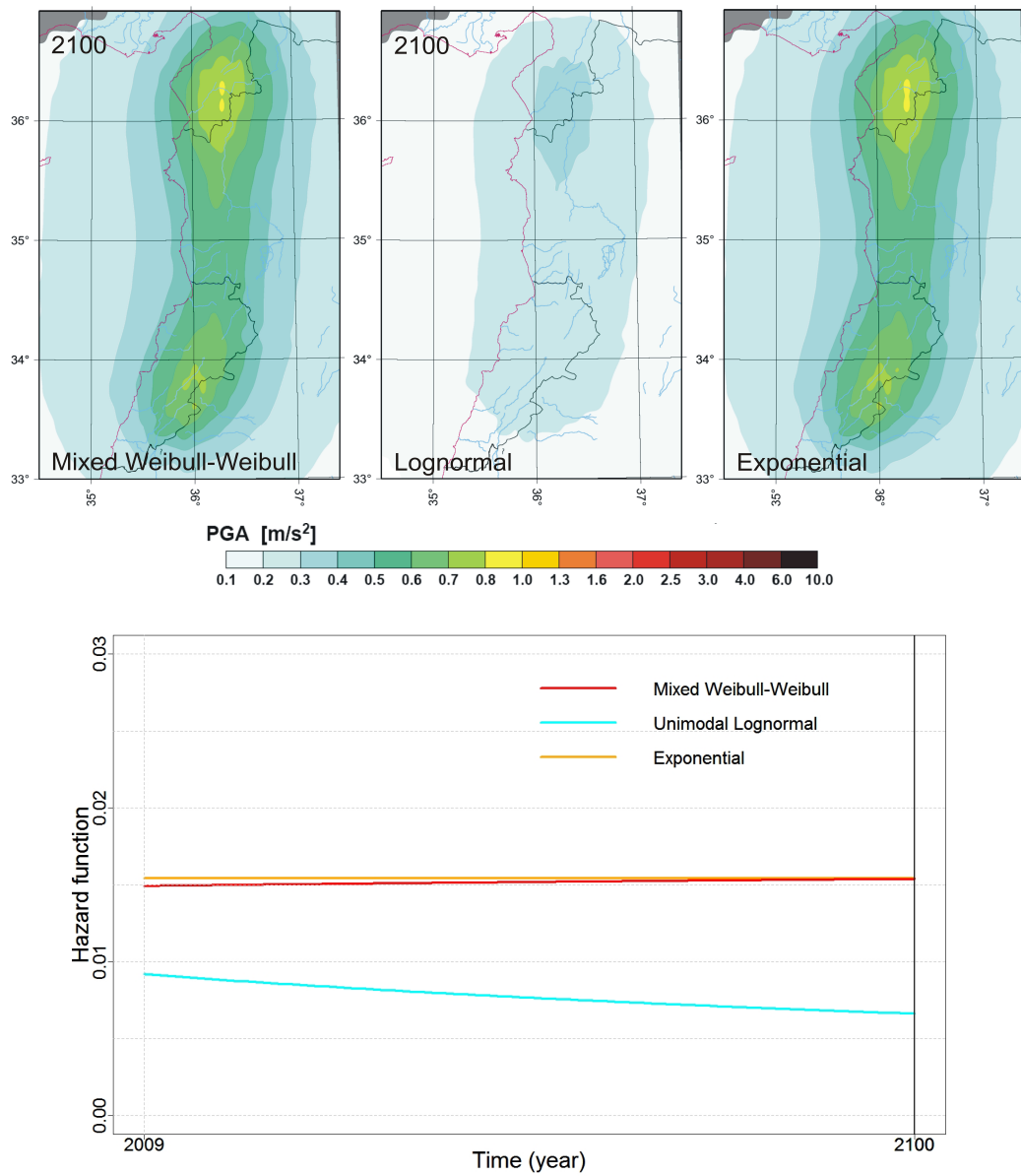


Figure B.5: Top: Hazard maps for the northern part of the DSFZ in the year 2100 for 10% exceedance probability in 50 years considering occurrence rates of large events obtained by the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution (classical time-independent model under the Poisson process). For more details, see Grünthal et al. 2009c. Bottom: Statistical hazard functions (occurrence rates) for the mixed Weibull-Weibull, the unimodal Lognormal, and the exponential distribution within 2009 - 2100.



## Acknowledgments

My special gratefulness goes to Dr. Gottfried Grünthal, head of Hazard Assessment and Stress Field section at GFZ, who helped, supported and guided me during my PhD. He had always enough patience to answer my questions about seismological issues and treated my more mathematical mind to a multidisciplinary direction. I thank Dr. Holger Schelle, who helped me to learn more about physics of earthquakes. I also thank Dr. Jean-Pierre Stockis and Dr. Gerald Kroisandt, who helped me very effectively in mathematical and statistical parts of my thesis. I thank my supervisor at the University of Potsdam and head of Geophysical Deep Sounding section at GFZ, Prof. Dr. Michael Weber, who supported and helped me in GFZ during my PhD. Special thanks go to Dr. Christine Böckmann, who kindly accepted my participation in her University course and helped me to obtain some teaching experience. Throughout discussions with Dr. Sebastian Hainzl, I could improve my understanding of main subjects in my thesis. I would also like to thank my other colleagues from Hazard Assessment and Stress Field section at GFZ for their kind helps and supports. Finally, I thank my parents in Iran for their constant support.

*This study was part of the DEad Sea Integrated Research project, DESIRE, founded by the Deutsche Forschungsgemeinschaft (German Research Foundation). I appreciate this support and the support from the German Research Centre for Geosciences, GFZ.*

## Curriculum Vitae

First Name: Amir Hossein  
Last Name: Hakimhashemi  
Birth Date: 11 February 1973  
Place of Birth: Tehran, Iran

### Education

1978 - 1986 Basic and Secondary School in different schools in Tehran, Iran  
1986 - 1990 Saeedi High School, Tehran, Iran  
High School degree in Mathematics and Physics  
1990 - 1996 Isfahan University of Technology, Isfahan, Iran  
Bachelor degree in Applied Mathematics  
Specialization in Operations Research  
1996 - 1999 Institute for Research in Planning and Development, Teheran, Iran  
Master degree in Socio-economic Systems Engineering  
Specialization in Transportation Planning  
2003-2006 Technical University of Kaiserslautern, Germany  
Master degree in Mathematics  
Specialization in Optimization and Statistics  
2006-2009 University of Potsdam, Germany  
PhD study in Geophysics

### Career

2006-present Researcher at Department of Physics of the Earth  
GeoForschungsZentrum Potsdam (Potsdam, Germany)  
1999-2003 Expert at the Iran Road and Maintenance Organization, Tehran, Iran

**Tutoring Experience** Assistance in the lecture course "Mathematics II for Geoecologists and Geoscientists"  
Department of Mathematics  
University of Potsdam (Potsdam, Germany)

Potsdam den 27 October 2009

/Amir Hossein Hakimhashemi/