

# Real-time Information from Seismic Networks

Zur Erlangung des akademischen Grades eines  
DOKTORS DER NATURWISSENSCHAFTEN  
von der Fakultät für Physik des  
Karlsruher Institut für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Geophys. Nina Köhler  
aus Hamburg

Tag der mündlichen Prüfung:

22. Januar 2010

Referent:

Prof. Dr. Friedemann Wenzel

Korreferent:

Prof. Dr. Mustafa Erdik



# Abstract

Earthquake early warning systems provide real-time information about the location and magnitude of an occurring earthquake or the level of imminent ground shaking at a specific user site. By using the fact that modern communication systems can transmit data and information faster than seismic waves propagate, warnings can be issued ahead of the destructive seismic waves. Earthquake early warning systems can provide warning times of a few seconds up to about one minute. The warning time depends on the distances between the earthquake location, the seismic instruments, and the user sites, but also on the algorithms used to rapidly process the incoming ground motion data. Earthquake early warning systems are highly valuable for the mitigation of co-seismic risk.

In this thesis, the performance of a novel methodology for the real-time estimation of seismic source parameters is evaluated that can be used in early warning systems. The method, called *PreSEIS*, has been developed on the example of the mega-city Istanbul by using a set of simulated earthquake scenarios (Böse, 2006). The method combines the advantages of regional and on-site early warning: It uses the ground motion information from a seismic network to estimate earthquake source parameters such as hypocentre location and magnitude. The estimation process starts as soon as the first station of the network is triggered by the seismic waves. The source parameter estimates are continuously updated with ongoing time when more ground motion information becomes available. *PreSEIS* uses artificial neural networks to map the seismic observations onto the likely source parameters. The neural networks must be trained with a set of example earthquakes with known source parameters.

This thesis presents the first application of *PreSEIS* to real earthquake data using two different datasets. The first application uses a dataset of 74 earthquakes from Southern California recorded at stations of the Southern California Seismic Network. The observed earthquake records are combined with empirical relations from Southern California. The second application is based on a dataset of 69 earthquakes from Japan, recorded at stations of the Kyoshin-Net. *PreSEIS* shows robust and stable performances using the real datasets. The source parameter estimates improve significantly with ongoing time. The mean location errors derived from the training processes, defined as the deviation of the estimated hypocentre locations from the true ones, can be reduced to about 13 km within 15 seconds. The according mean magnitude errors can be reduced within the same time span down to  $\pm 0.3$  magnitude units. When the trained neural networks are applied to unknown earthquakes from the same region that were not included into the training dataset, *PreSEIS* is able to estimate the magnitudes of these earthquakes with an accuracy of 90-100%. Finally, *PreSEIS* is used as a benchmark system to evaluate the earthquake early warning per-

---

formance of the *Istanbul Earthquake Rapid Response and Early Warning System*. A potential expansion of the early warning network to regional scale would improve the early warning performance of up to 22%. The application of the original Istanbul early warning system, which is based on the exceedance of amplitude thresholds, to a set of simulated earthquake scenarios reveals possible warning times for Istanbul of up to 10 seconds.

Furthermore, this thesis addresses the simulation of seismic ground motion for specific earthquake scenarios. I use the stochastic simulation technique for finite faults by Beresnev and Atkinson (1997, 1998) to simulate a historic earthquake that hit Istanbul in the year of 1509. The selection of different hypocentre locations and prediction techniques demonstrates the variability of ground motion distribution and thus of potential structural damages which are quantified in terms of predicted building damages for the different scenarios. Depending on the prediction technique, the building damage estimates differ by up to 41%.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Seismic Hazard in Istanbul . . . . .	4
1.2	IERREWS . . . . .	6
<b>2</b>	<b>Simulation of Seismic Ground Motion</b>	<b>9</b>
2.1	The Earthquake Source . . . . .	10
2.2	The Path Effects . . . . .	12
2.3	Local Site Effects . . . . .	13
2.4	Ground Motion Parameters . . . . .	15
2.5	The Simulation Method <i>FINSIM</i> . . . . .	20
2.6	Simulated dataset . . . . .	26
2.7	Optimization of <i>FINSIM</i> parameters . . . . .	29
<b>3</b>	<b>The historic 1509 Istanbul earthquake</b>	<b>35</b>
3.1	Simulation results . . . . .	36
3.2	Building Damage Estimation . . . . .	42
<b>4</b>	<b>Earthquake Early Warning</b>	<b>49</b>
4.1	State-of-the-art . . . . .	49
4.2	PreSEIS . . . . .	55
<b>5</b>	<b>PreSEIS application</b>	<b>65</b>
5.1	Southern California . . . . .	65
5.2	Japan . . . . .	75
5.3	EEW for Istanbul . . . . .	82

---

<b>6</b>	<b>Summary and Outlook</b>	<b>97</b>
<b>A</b>	<b>Simulation of Seismic Ground Motion</b>	<b>103</b>
A.1	Simulated Earthquake Dataset . . . . .	103
A.2	PreSEIS results . . . . .	112
<b>B</b>	<b>Attenuation Relationships</b>	<b>113</b>
B.1	Campbell & Bozorgnia (2008): Global PEER database . . . . .	113
B.2	Özbey et al. (2004): Northwestern Turkey . . . . .	115
B.3	Boore et al. (1997): Western North America . . . . .	117
<b>C</b>	<b>Building Damage Estimation</b>	<b>119</b>
<b>D</b>	<b>The Virtual Seismologist Method</b>	<b>125</b>
<b>E</b>	<b>Earthquake Records from Japan</b>	<b>129</b>
	<b>List of Figures</b>	<b>136</b>
	<b>List of Tables</b>	<b>138</b>
	<b>References</b>	<b>139</b>
	<b>Used Software</b>	<b>147</b>
	<b>Danksagung / Acknowledgments</b>	<b>149</b>

# Chapter 1

## Introduction

Earthquakes regularly cause great numbers of casualties and massive destructions around the world. With increasing urbanization worldwide, earthquakes pose significant threats to lives and objects in seismic active areas. As reported by the Munich Re Group<sup>1</sup>, six of the ten deadliest natural catastrophes in the years between 1980 and 2008 were earthquakes. Among them were the 8 October 2005  $M_w$ 7.6 Pakistan earthquake, the 12 May 2008  $M_w$ 7.9 Sichuan earthquake in China, and the 26 December 2004  $M_w$ 9.0 Sumatra earthquake in south Asia which caused a great tsunami. The ten most costly natural disasters include a number of four earthquakes as well, headed by the 16 January 1995  $M_w$ 6.9 Kobe earthquake in Japan with an overall loss of US\$ 100,000m<sup>1</sup>. Table 1.1 lists the ten earthquakes with the largest number of fatalities between 1980 and 2008, as given by the Munich Re Group. In the year 2008, worldwide losses were caused by a total of 78 earthquakes. The 2008 Sichuan earthquake, for example, was the most expensive catastrophe in macroeconomic terms, with direct losses of at least US\$ 85bn (Munich Re Group, 2009).

In contrast to other natural catastrophes such as hurricanes, tsunamis, or volcanic eruptions, warning times for earthquakes are extremely short. As earthquakes cannot be reliably predicted due to their complex nature, co-seismic earthquake early warning systems can be important tools for reducing the losses caused by a major earthquake. An earthquake early warning (EEW) system provides warnings prior to damaging ground motions using the physical basis that the destructive shear (S-) and surface waves of an occurring earthquake travel slower than the compressional (P-)waves and much slower than electromagnetic signals used to transmit information and warnings. The warning times, defined as the time between the issued warning and the occurrence of strong ground shaking at the user site, range between some seconds and a maximum of about one minute (Wenzel et al., 1999). They depend not only on the distance between earthquake source, seismic network, and user site, but also on the time that is needed to process and transmit data and information.

As first stated by Heaton (1985), an early warning system consists of following components: a

---

<sup>1</sup>Munich Re Group, Geo Risks Research, NatCatSERVICE, January 2009 (<http://www.munichre.com/>)

<i>Date</i>	<i>Region</i>	<i>Fatalities</i>	<i>Overall losses (US\$ )</i>
26 December 2004	South Asia (Tsunami)	220,000	10,000m
8 October 2005	Pakistan	88,000	5,200m
12 May 2008	Sichuan / China	70,000	85,000m
21 June 1990	Gilan Province / Iran	40,000	7,100m
26 December 2003	Bam / Iran	26,200	500m
7 December 1988	Armenia	25,000	14,000m
17 August 1999	Kocaeli (Izmit) / Turkey	17,000	12,000m
26 January 2001	Gujarat / India	15,000	4,500m
19 September 1985	Mexico City / Mexico	9,500	4,000m
30 September 1993	India	7,600	280m

**Table 1.1:** Statistics over the 10 earthquakes with the largest numbers of fatalities between 1980 and 2008<sup>1</sup>. The high number of fatalities for the earthquake in south Asia is mainly due to the tsunami that followed the earthquake.

network of seismic stations for monitoring, a continuous real-time communication for transmitting the recorded data, a central processing facility that infers information from the data, and a broadcasting system that issues the warning to end users. The type of information inferred from the data depends on the requirements of the users. EEW systems can estimate earthquake source parameters such as location and magnitude of the event, but they can also directly estimate the impending ground shaking at a specific user site (Wenzel et al., 1999).

Despite the short warning times the number of possible applications of EEW systems is large. The most common application is the initiation of automatic, pre-defined actions (Harben, 1981), such as:

- shutdown of industrial facilities (e.g. power plants, power distribution centres, factories, refineries)
- shutdown of gas and oil distribution
- stop or slow down trains
- shutdown of computers and disk drivers, initiation of backups
- open fire station doors
- stop elevators and open doors at next floor
- issue audio alarms to alert people

The occurrence of false alarms, however, is an important factor in an early warning system. The unnecessary initiation of actions is costly and might also lead to a loss of credibility for the warning system (Goltz, 2002). A survey undertaken within the frame of the TriNet project in California revealed that many organizations would tolerate only a limited number of false alarms (5 or less) before discontinuing with early warning (Goltz, 2002).

Earthquake early warning is a major component of real-time information systems that provide rapid information for disaster management (Wenzel et al., 2004). The EEW system operates while the earthquake rupture is taking place and issues a warning seconds before strong ground shaking occurs at the sites of interest. Within some time after the earthquake stroke, a rapid response system delivers information about the ground motion distribution in terms of specified ground motion parameters and estimates of damage and human losses. This is crucial information as it enables not only to assess the necessary amount of resources requested for mitigation and to delegate rescue operations, but it also allows for developing priorities in disaster mitigation policy, measuring the efficiency of mitigation actions, and for assessing variations of risk with time (Wenzel et al., 2004).

The necessary ground motion parameters for the rapid response system are peak ground acceleration and velocity, seismic intensity, and spectral parameters at specific periods of motion (see Chapter 2). Their distributions are mapped in shake maps which can be available minutes after the earthquake. The shake maps in combination with databases of buildings and infrastructure and associated vulnerability functions are used to estimate the potential damages. These can then be used to estimate the number of fatalities.

To prepare for a future earthquake in a given urban area, deterministic estimates of ground motions may be done for a specific example scenario. By quantifying the effects of the predicted ground motions on certain urban elements such as infrastructure, buildings, services, and of course human safety, pre-seismic actions like the establishment of building codes or emergency plans, for example, can be done more systematically. It is thus important that the selected scenario earthquake is representative and that additional effects influencing the ground motion characteristics, such as path and local site effects, are defined wisely.

One part of this thesis addresses the topic of stochastic ground motion simulation. The ground motions simulated for a selected scenario earthquake in the Sea of Marmara, northwestern Turkey, will be used to predict likely building damages in the mega-city Istanbul. The simulation technique will also be used to simulate ground motion records at the sites of certain early warning stations in Istanbul that did not record a sufficient amount of real earthquake data necessary for the analysis in this study.

Along with the technical requirements of the seismic early warning stations and real-time communication in an EEW system, the algorithm used for processing the incoming ground motion data is of high importance. Different techniques for the rapid estimation of seismic source parameters or impending ground shaking exist (see Chapter 4). In this study, I will evaluate the performance of a novel approach for the rapid estimation procedure, called *PreSEIS* (Böse, 2006). The methodology has been developed on the example of Istanbul using the stochastic ground motion simulations mentioned above. It is based on artificial neural networks which are used to map the ground motions observed at the stations of a seismic network onto the earthquake source parameters (hypocentre location, magnitude, and extension of the evolving rupture).

In this study, I will apply *PreSEIS* to real earthquake data for the very first time and analyze its performance. Additionally, I will use it as a benchmark system to evaluate the EEW system of Istanbul.

**The outline of this thesis is the following:**

In the next two sections, an introduction to seismic hazard in Istanbul and a description of the early warning and rapid response system in Istanbul are given.

In the first part of Chapter 2, an introduction to the simulation of seismic ground motion is given and common ground motion parameters are defined. The specific simulation technique used in this study is then presented in Section 2.5. Section 2.6 presents the dataset of simulated earthquake scenarios originally established by Böse (2006) that is used in this study in a modified version. The modification, or better optimization, of this data is presented in Section 2.7.

Chapter 3 addresses the above mentioned scenario earthquake: a historic earthquake that occurred close to Istanbul in the year 1509 is simulated to demonstrate the variability of ground motion predictions and to investigate the building damages in the today's city of Istanbul that might arise from a similar earthquake.

The first part of Chapter 4 gives an overview about present earthquake early warning systems and techniques used around the world. The second part summarizes the in this study applied early warning methodology *PreSEIS*.

Chapter 5 addresses the performance tests of *PreSEIS*. First, the method is applied to two real datasets, one from Southern California and one from Japan, to analyze the functionality of *PreSEIS*. Second, *PreSEIS* is used as a benchmark system to evaluate the EEW performance for Istanbul.

Finally, Chapter 6 will summarize the main findings of this thesis and will give an outlook on future tasks.

## 1.1 Seismic Hazard in Istanbul

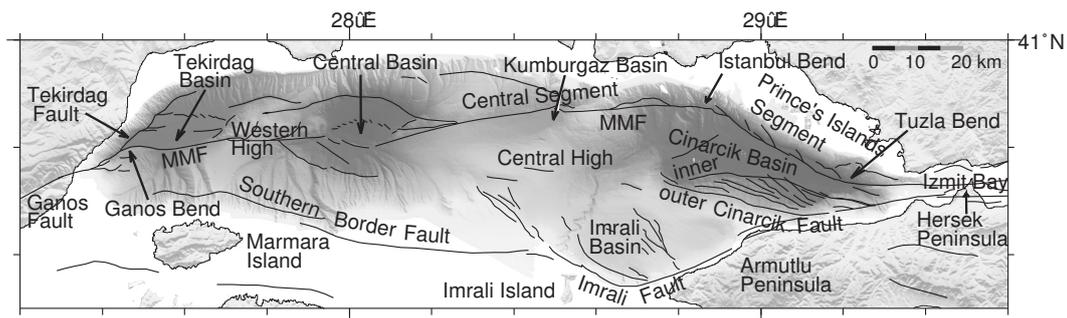
The mega-city Istanbul, located at the Bosphorus which connects the Sea of Marmara with the Black Sea, is the largest city in Turkey. With a population of 12.7 million and a population density of 6.9 inhabitants per km<sup>2</sup> it is also the fifth largest city in the world<sup>2</sup>. The Marmara region, i.e., the region in northwestern Turkey around the Sea of Marmara, is with 67,000 km<sup>2</sup> the smallest but the most densely populated geographical region in Turkey<sup>2</sup>.

The tectonic situation of the Marmara region is dominated by the *North Anatolian Fault* (NAF), starting in the east at a longitude of about 41°E. Until about 30°E, the NAF system is a narrow fault of predominantly dextral strike-slip character that separates the rigid Black Sea and central Anatolia regions (Ambraseys and Jackson, 2000).

---

<sup>2</sup>Wikipedia Encyclopedia, 2009 (<http://www.wikipedia.org/>)





**Figure 1.2:** Map of the northern Marmara Sea basin with bathymetry and fault segments after Armijo et al. (2002) (taken from Hergert (2009)).

the JICA study<sup>3</sup>, it caused the second worst human casualties in the 20th century in Turkey. Seven provinces were affected leading to death of more than 15,000 people. More than 77,000 households were heavily damaged, whereas collapsing buildings accounted for the majority of human losses. The communication system broke down so that no phones (also mobile phones) were usable in the first 48 hours. The initial search and rescue operations were not well organized and hence not effective (JICA<sup>3</sup>).

The devastating 1999  $M_w$ 7.4 Kocaeli and  $M_w$ 7.2 Düzce earthquakes initiated the establishment of the *Istanbul Earthquake Rapid Response and Early Warning System (IERREWS)* in the year 2002 by the Department of Earthquake Engineering of the *Kandilli Observatory and Earthquake Research Institute (KOERI)* of the Bogazici University (Erdik et al., 2003b; Alçik et al., 2009). The purpose of *IERREWS* is to provide near-real time warning for future potentially disastrous earthquakes and to enable rapid shake map and damage assessment after an earthquake (Erdik et al., 2003b). The following section will give a short introduction into the system.

## 1.2 IERREWS

The *Istanbul Earthquake Rapid Response and Early Warning System (IERREWS)* installed and maintained by *KOERI* consists of a dense network of strong motion instruments in the metropolitan area of Istanbul. The seismic stations have a dynamic range of  $\pm 2g$  and record ground acceleration on three channels (Erdik et al., 2003b). As stated by its name, *IERREWS* consists of two parts: a rapid response system and an earthquake early warning (EEW) system. The rapid response part of *IERREWS* consists of 100 strong motion stations which are installed at about 2-3 km interstation distances in dial-up mode for generating rapid response information. Once triggered by an earthquake, each station processes the incoming data and calculates spectral acceleration at pre-defined periods. Additionally, 12 Hz filtered peak ground acceleration and peak ground velocity values are determined. Every 20 s, the parameters are sent in form

<sup>3</sup>The Study on a Disaster Prevention / Mitigation Basic Plan in Istanbul including Seismic Microzonation in the Republic of Turkey (Final report, 2002), by the Japan International Cooperation Agency (JICA) and the Istanbul Metropolitan Municipality (IMM)

of SMS messages to the main data processing centre, located at the Department of Earthquake Engineering at *KOERI*.

To obtain the spectral displacement values at a grid over Istanbul of  $0.01^\circ \times 0.01^\circ$  spacing, the transmitted values are interpolated. For each grid cell, the earthquake demand and instrumental intensity are calculated which are used to estimate building damages and casualties at each grid cell, whereas the casualties are estimated on the basis of the number of collapsed buildings and the degree of damage (more details about the building damage estimation procedure are given in Chapter 3.2 and Appendix C). The distribution of ground shaking and estimated damages and casualties are automatically mapped and transmitted to end users (Governorate of Istanbul, First Army Headquarters, and Istanbul Metropolitan Municipality) within 5 minutes by digital radio modem and GPRS communication (Erdik et al., 2003b).

The EEW part of *IERREWS* consists of 10 additional strong motion stations installed in on-line mode along the northern shoreline of the Sea of Marmara, as close as possible to the Main Marmara Fault. The stations are described in Table 2.3; their locations can be seen in Figure 2.7. The data from the early warning stations are continuously transmitted to the data processing centre by a digital spread spectrum radio modem system (Erdik et al., 2003b).

The implemented early warning method is a simple and robust algorithm which is based on the exceedance of specific threshold time-domain amplitude levels from which three alarm levels are declared:

From the incoming, band-pass filtered acceleration data the peak ground acceleration (*PGA*) is determined and compared with the trigger thresholds. When at any station on any channel *PGA* exceeds a threshold of 0.02g, it is considered a vote. Whenever votes from at least 3 stations within 5 s can be obtained the system will declare alarm level 1. After alarm 1, alarm 2 will be issued whenever the amplitudes exceed 0.05g at any three stations within 5 s. Finally, after alarm 2, the highest alarm level 3 will be declared whenever the amplitudes exceed 0.1g at any three stations within 5 s (Erdik et al., 2003b; Alcik et al., 2009).

As an alternative approach, the cumulative absolute velocity, defined as the time integral of absolute acceleration, can also be used.

The declared alarm levels will be communicated to the recipient facilities which will automatically initiate suitable actions based on the level of alarm. Depending on the earthquake location and the recipient facility the warning time can range up to about 8 s (Erdik et al., 2003b).



## Chapter 2

# Simulation of Seismic Ground Motion

Earthquakes occur on faults which can be characterized as quasiplanar breaks in the earth on which one side moves with respect to the other side. As stated in the *elastic rebound theory* by H. Reid in 1910, crustal stresses resulting from large-scale crustal shearing motions lead to accumulating strain in the immediate vicinity of a fault. When the accumulated strain reaches a threshold, which depends on the properties of the rock and on the fault surface, abrupt frictional sliding occurs. Portions of the accumulated energy are radiated by seismic waves, while other portions of the energy are consumed in heating and fracturing of the rocks (Lay and Wallace, 1995).

An earthquake is characterized by its location, i.e. the geographical latitude and longitude of its epicentre, its depth, and its magnitude. In contrast to the epicentre, the hypocentre includes the depth of the earthquake and represents the location on the fault where the rupture initiates.

The seismic waveform observed at a certain location depends, on the one hand, on the description of the earthquake source. On the other hand, a seismic wave that propagates from the source to the observation site undergoes significant modifications by path and local site effects. If a synthetic waveform is to be modeled, the waveform is characterized as a combination of linear filters representing the source, the path, and the site effects. The waveform  $u(t)$  observed at a point  $\mathbf{x}$  at time  $t$  can be expressed as the convolution of the source  $s(\mathbf{x}, t)$ , the path  $p(\mathbf{x}, t)$  and the site effect  $f(\mathbf{x}, t)$ :

$$u(t) = s(\mathbf{x}, t) * p(\mathbf{x}, t) * f(\mathbf{x}, t). \quad (2.1)$$

As will be shown later, each of these terms depends on the frequency of the seismic wave. It is therefore more appropriate to express equation 2.1 as a function of frequency which is achieved by Fourier transformation. Since convolution in the time domain is equivalent to multiplication in the frequency domain  $u(t)$  can be written as

$$U(\omega) = S(\mathbf{x}, \omega) \cdot P(\mathbf{x}, \omega) \cdot F(\mathbf{x}, \omega), \quad (2.2)$$

where  $\omega$  is the angular frequency  $\omega = 2\pi f$  (e.g. Lay and Wallace, 1995; Stein and Wysession, 2003).

Although the site effects, i.e. the effects appearing in the direct vicinity of the considered site, are part of the propagation effects, it is useful to handle them separately, as it is done in equations 2.1 and 2.2. I will now successively discuss the source and the path and site effects as they will be used in Section 2.5 in the simulation technique.

## 2.1 The Earthquake Source

The observed seismic waveform at a certain site can be represented as a linear combination of the contributions from the different points on the fault: The rupture, starting at the hypocentre location, spreads and causes each point that is passed to slip. This process is expressed by the *representation theorem* which can be written as (Aki and Richards, 1980)

$$u_n(\mathbf{x}, t) = \int_{-\infty}^{+\infty} d\tau \iint_{\Sigma} [u_i(\xi, \tau)] c_{ijpq} v_j \frac{\partial G_{np}(\mathbf{x}, t - \tau; \xi, 0)}{\partial \xi_q} d\Sigma. \quad (2.3)$$

$u_n(\mathbf{x}, t)$  is the  $n$ th component of displacement at point  $\mathbf{x}$  at time  $t$ . The  $i$ th displacement discontinuity across the fault surface  $\Sigma$  is given by  $[u_i(\xi, \tau)] = u_i(\xi, \tau)|_{\Sigma^+} - u_i(\xi, \tau)|_{\Sigma^-}$ , whereas  $\xi$  is a position on  $\Sigma$  and  $\tau$  is the time at which the displacement occurs. The components  $c_{ijpq}$  of a fourth-order tensor describing an isotropic medium are expressed by the Lamé constants  $\lambda$  and  $\mu$ :  $c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$  with the Kronecker function  $\delta_{ij}$  (i.e.,  $\delta_{ij} = 1$  for  $i = j$  and  $\delta_{ij} = 0$  for  $i \neq j$ ).  $v_j$  is the normal vector to  $\Sigma$ . The term  $G_{np}(\mathbf{x}, t; \xi, \tau)$  is the Green's function which describes the  $n$ th component of displacement at location  $\mathbf{x}$  at time  $\tau$ .

Equation 2.3 shows that, if the displacement discontinuity as a function of location and time and the Green's function are known, the total ground motion at the observation site can be determined. The displacement discontinuity implicates a description of the earthquake source, whereas the Green's function represents the propagation of the seismic waves. For large amplitude waves, however, the linear stress-strain relation assumed in equation 2.3 breaks down and becomes non-linear. The representation theorem does not account for these non-linearities (Anderson, 2003).

The size of an earthquake is best described by the *seismic moment*  $M_0$ . According to the representation theorem, the evolution of the seismic moment of an earthquake is as well a function of time. It can be expressed by (Aki and Richards, 1980)

$$M_0(t) = \mu A \bar{u}(t) = \mu L W \bar{u}(t) \quad (2.4)$$

where  $\mu$  is the shear modulus,  $\bar{u}(t)$  is the temporal evolution of the average displacement, and  $A$  is the fault area, i.e. the product of fault length  $L$  and fault width  $W$ . The time which is needed to reach the final average displacement  $\bar{u}(\infty)$  is defined as the *rise time*  $t_r$ . When this final amount of slip is reached the total seismic moment  $M_0$  of the earthquake is given as

$$M_0 = \mu A \bar{u}(\infty). \quad (2.5)$$

Measured values of  $M_0$  range from about  $10^{12}$  dyn-cm for micro-earthquakes up to  $10^{30}$  dyn-cm for strong earthquakes (Aki and Richards, 1980).

Together with the description of the seismic moment  $M_0(t)$  from equation 2.4, the far-field displacement in a homogeneous, isotropic, unbounded medium due to a double couple point source may be given (in spherical coordinates) in the following form (e.g. Lay and Wallace, 1995):

$$\mathbf{u}(\mathbf{x}, t) = \frac{\mathfrak{R}_P^{\theta\phi}}{4\pi\rho\alpha^3 r} \dot{M}_0 \left( t - \frac{r}{\alpha} \right) + \frac{\mathfrak{R}_S^{\theta\phi}}{4\pi\rho\beta^3 r} \dot{M}_0 \left( t - \frac{r}{\beta} \right). \quad (2.6)$$

$\dot{M}_0(t) = \mu A \dot{\bar{u}}(t)$  is the source time function with the time derivative  $\dot{\bar{u}}(t)$  of the average displacement. The distance  $r$  is set to  $|\mathbf{x}|$  and  $\mathfrak{R}_P^{\theta\phi}$  and  $\mathfrak{R}_S^{\theta\phi}$  are the radiation patterns for P- and S-waves given by

$$\begin{aligned} \mathfrak{R}_P^{\theta\phi} &= \sin(2\theta) \cos(\phi) \hat{\mathbf{r}} \\ \mathfrak{R}_S^{\theta\phi} &= \cos(2\theta) \cos(\phi) \hat{\boldsymbol{\theta}} - \cos(\theta) \sin(\phi) \hat{\boldsymbol{\phi}} \end{aligned} \quad (2.7)$$

with  $\hat{\mathbf{r}}$ ,  $\hat{\boldsymbol{\theta}}$ , and  $\hat{\boldsymbol{\phi}}$  being unit vectors pointing into radial and transverse directions, respectively. The description ‘‘far-field’’ refers to distances at least several wavelengths away from the fault. The parameters  $\alpha$  and  $\beta$  in equation 2.6 are the propagation velocities of P- and S-waves, respectively, and are defined as  $\alpha = \sqrt{(\lambda + 2\mu)/\rho}$  and  $\beta = \sqrt{\mu/\rho}$  with  $\rho$  being the density of the medium.

Beresnev and Atkinson (1997) give the modulus of the Fourier-transformed far-field displacement, derived from equation 2.6, as

$$|\mathbf{u}_n^{wave}(\mathbf{x}, \omega)| = \frac{\mathfrak{R}_{wave}^{\theta\phi} M_0}{4\pi\rho c^3 r} \left[ 1 + \left( \frac{\omega}{\omega_c} \right)^2 \right]^{-(n+1)/2} \quad (2.8)$$

where *wave* stands for either P- ( $c = \alpha$ ) or S-waves ( $c = \beta$ ).  $\omega_c$  is the *corner frequency* of the Fourier amplitude spectrum. The parameter  $n$  defines the order of decay of the source spectrum:  $n = 1$  represents the commonly used  $\omega^2$  spectrum after Brune (1970, 1971) where the Fourier amplitudes at high frequencies fall off proportional to  $\omega^{-2}$ . The corner frequency  $\omega_c$  represents the frequency above which this decay takes place.

Following equation 2.8, the displacement spectrum after Brune (1970, 1971) is given as

$$|S_{wave}^d(\mathbf{x}, \omega)| = \frac{\mathfrak{R}_{wave}^{\theta\phi} M_0}{4\pi\rho c^3} \left[ 1 + \left( \frac{\omega}{\omega_c} \right)^2 \right]^{-1}, \quad (2.9)$$

where index  $d$  represents *displacement*.

As we have seen from equation 2.5, the seismic moment  $M_0$  of a point-source dislocation involves the final slip  $\bar{u}(\infty)$ . This slip results in a deformation of  $\bar{u}(\infty)/L$ , with  $L$  being the fault length, which causes a change in stress given as

$$\Delta\sigma = \frac{\mu\bar{u}(\infty)}{L}. \quad (2.10)$$

$\Delta\sigma$  is the *static stress drop*, representing the difference in stress on the fault plane before and after the earthquake.

By inserting equation 2.10 in the definition of the seismic moment (equation 2.5),  $M_0$  can be expressed as

$$M_0 = L^2 \cdot W \cdot \Delta\sigma. \quad (2.11)$$

Assuming  $L \approx W$  gives

$$M_0 \approx L^3 \cdot \Delta\sigma. \quad (2.12)$$

The seismic moment of an earthquake is thus proportional to the stress drop (Beresnev and Atkinson, 1997).

Most earthquakes have typical stress drops between 10 and 100 bars (Lay and Wallace, 1995). This relatively small variability justifies to consider  $\Delta\sigma$  as a constant from which the concept of *self-similarity* for earthquakes follows: The earthquake source processes are scale-invariant; earthquakes of different sizes only differ by a scaling factor.

A constant stress drop results into two conditions: First, the average slip  $\bar{u}(\infty)$  scales with rupture dimension  $L$ , that is  $\bar{u}(\infty) \propto L$ . Second, the seismic moment  $M_0$  scales with rupture duration  $\tau_c$  in the form  $M_0 \propto \tau_c^3$ . The assumption of self-similarity implies that the  $\omega^2$  source model follows the relation

$$M_0 \cdot \omega_c^3 = \text{constant} \quad (2.13)$$

as observed by Aki (1967).

The seismic moment  $M_0$  has been used by Hanks and Kanamori (1979) to define the *moment magnitude*  $M_w$  as

$$M_w = \frac{2}{3} \log M_0 - 10.7 \quad (2.14)$$

with  $M_0$  given in dyn-cm.  $M_w$  is a preferred magnitude scale because it does not saturate for large earthquakes like other magnitude scales.

## 2.2 The Path Effects

Path (or propagation) effects modify the propagating seismic waves on their way from the source to the receiver. The wave amplitudes attenuate with propagation of the seismic waves which is caused by several effects:

The first effect is *scattering* of the seismic waves. It is caused by the wavefield's interaction with small-scale heterogeneities in the traversed medium. The seismic waves are reflected, refracted, and diffracted at interfaces at which physical properties change and conversions from P- to S-waves take place. Scattering is an elastic process which conserves but redistributes the seismic energy and which leads to complex waveforms (Lay and Wallace, 1995).

The second effect is the *geometrical spreading*. It describes the decay of wave amplitudes with

distance  $r$  from the source. Conservation of energy leads to decreasing wave amplitudes as the wavefront expands with increasing distance.

In a homogeneous, spherical earth we must distinguish between near-field, intermediate-field, and far-field waves which relate to seismic waves in the immediate vicinity to the source, far away from the source, and in between these two, respectively. In the near-field, the amplitudes decay with  $1/r^4$ . In the intermediate-field, the decrease is proportional to  $1/r^2$ . The amplitude decay in the far-field is proportional to  $1/r$  (Aki and Richards, 1980).

The third effect is the *inelastic attenuation*. In contrast to scattering and geometrical spreading, which are elastic processes where the energy is conserved, inelastic attenuation is caused by anelasticity - the wave amplitudes attenuate due to the dissipation of energy. The kinetic energy of seismic motion is partially converted to heat by microscopic processes usually referred to as internal friction. Internal friction includes, for example, grain-boundary processes, crystal defects, or thermoelastic processes (Aki and Richards, 1980).

The overall effect of internal friction is summarized by the dimensionless *quality factor*  $Q$ . It is defined by the amount of dissipated energy  $-\Delta E$  per cycle of the oscillating wave:

$$\frac{1}{Q(\omega)} = \frac{1}{2\pi} \frac{-\Delta E}{E}. \quad (2.15)$$

$E$  is the energy stored in the volume (e.g. Aki and Richards, 1980; Lay and Wallace, 1995). The reciprocal definition of  $Q$  implies that large values of  $Q$  represent small attenuation and vice versa.  $Q$  for P-waves is systematically larger than for S-waves and can show typical values from about 10 up to some 1000, depending on the rock type (Lay and Wallace, 1995).

The amplitude attenuation of a seismic wave of angular frequency  $\omega$  with distance is given by

$$A^{att}(\mathbf{x}, \omega) = A_0 \cdot \exp\left(-\frac{\omega}{2} \frac{r(\mathbf{x})}{Q(\omega)c}\right) \quad (2.16)$$

where  $A_0$  is the initial wave amplitude and  $c$  is the seismic wave velocity of either P- or S-wave (Lay and Wallace, 1995). The frequency dependency of  $Q(\omega)$  obeys a power law of the form

$$Q(\omega) = Q_0 \left(\frac{\omega}{2\pi}\right)^\eta. \quad (2.17)$$

With equation 2.17, the inelastic attenuation  $I(\mathbf{x}, \omega)$  can be expressed as

$$I(\mathbf{x}, \omega) = \exp\left(-\frac{\omega}{2} \frac{r(\mathbf{x})}{Q(\omega)c}\right). \quad (2.18)$$

## 2.3 Local Site Effects

*Local site effects* play a fundamental role in earthquake engineering. The local geology (e.g. sedimentary layers or basins) and topography beneath a site modify the traversing seismic waves

and thus the ground motions at the surface. The grade of damage to engineering structures is strongly influenced by local site effects. Soft soils beneath a site form low-velocity layers which trap energy and amplify all frequencies of motion due to the decreased impedance of the layer (Anderson, 2003). The amplifications can reach almost two orders of magnitudes (Kawase, 2003).

Site effects have been observed in a number of earthquakes, for example in Turkey (e.g. Bakir et al., 2002; Özel et al., 2002) or in Mexico (e.g. Singh et al., 1988). Evidence of topographical effects influencing the ground motions were found by Bard and Tucker (1985) or Chávez-García et al. (1996), for example.

An important aspect of site effects is that soft surface layers behave in a non-linear way when they are entered by strong seismic waves. The non-linearity is characterized by a reduction of shear rigidity and thus of shear-wave velocity while the damping factor increases. This results in prolongation of the predominant periods and reduction of the amplification factors (Kawase, 2003). The ground motion simulation technique presented in Section 2.5, however, does not include non-linearities.

For modeling local site effects two separate contributions can be assumed: the *site amplification* and the *high-frequency diminution*.

The site amplification of seismic wave amplitudes assumes a simple model of a single layer over a half-space of infinite, horizontal extension (Kawase, 2003). The amplification is a function of the frequency of motion: For the layer of thickness  $h$  with the shear-wave velocity  $\beta$  the resonant frequencies are given by

$$f_n = \frac{\beta}{4h}(2n + 1), \quad n = 0, 1, 2, \dots \quad (2.19)$$

where the lowest resonant frequency at  $n = 0$  (fundamental frequency) represents maximum amplification (Chen and Scawthorn, 2003).

The high-frequency diminution describes the attenuation of high frequencies of the Fourier amplitude spectrum of ground acceleration. Anderson and Hough (1984) proposed that this attenuation can be approximated by an exponential decay of the form

$$H(\mathbf{x}, \omega) \approx \exp\left(-\pi\kappa\frac{\omega}{2\pi}\right). \quad (2.20)$$

The decay parameter  $\kappa$  is a function of local site condition.

Sites are commonly classified by either the S-wave velocity in the medium beneath the site or the depth of the underlain sediment. Characterizing a site in terms of S-wave velocity can usually be done in two ways: by the so-called *quarter-wavelength* method and by  $V_{S30}$ . The *quarter-wavelength* method uses the average S-wave velocity from the surface to a depth corresponding to a quarter wavelength of the considered frequency (Joyner et al., 1981). The parameter  $V_{S30}$  is defined as the average shear wave velocity over the top 30 m of a site (Campbell, 2003). It is the basis for several building codes such as the widely used *NEHRP* code from the *National Earthquake Hazards Reduction Programme (NEHRP)* (Building Seismic Safety Council, 1997) that

<i>NEHRP class</i>	$V_{S30}$ [m/s]
A: hard rock	> 1500
B: rock	760 - 1500
C: very dense soil and soft rock	360 - 760
D: stiff soil	180 - 360
E: soft clay	< 180
F: soils requiring site-specific evaluations	

**Table 2.1:** *NEHRP* site classifications based on  $V_{S30}$  (Building Seismic Safety Council, 1997).

is used in this work. Table 2.1 describes the *NEHRP* site classes and their respective  $V_{S30}$  ranges.

Boore and Joyner (1997) combined the *quarter-wavelength* method and  $V_{S30}$  to derive frequency-dependent amplification values for certain  $V_{S30}$  sites that correspond to the *NEHRP* site classes. These amplification values are used for the ground motion simulations in this study. Missing amplification values for *NEHRP* class *B* were added by Böse (2006) by non-linear interpolation. The amplification values are listed in Table 2.2.

The total amplification is given by the multiplication of the site amplification  $A(\mathbf{x}, \omega)$  (values from Table 2.2) and the high-frequency diminution  $H(\mathbf{x}, \omega)$ . As seen in Figure 2.1, the amplification at high frequencies is strongly dependent on the choice of the decay parameter  $\kappa$ . The larger  $\kappa$  the stronger is the decay.

## 2.4 Ground Motion Parameters

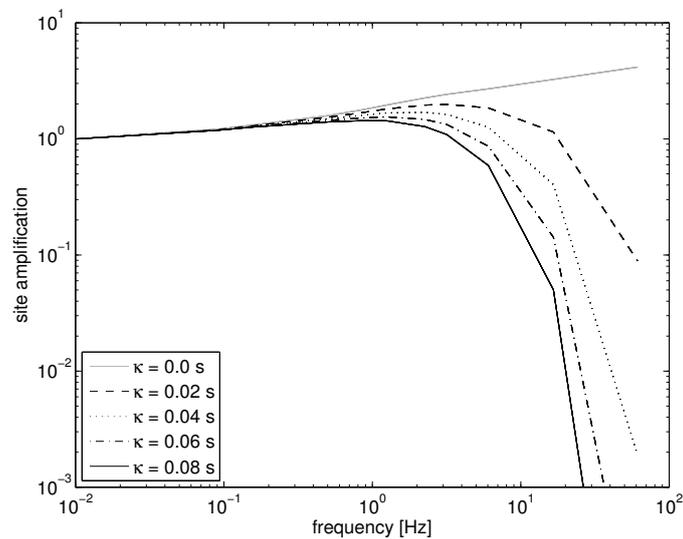
From recorded or simulated ground motion time series at a certain site various parameters can be derived that are commonly used in earthquake engineering. The simplest parameter is the *peak ground acceleration (PGA)*. It is defined as the maximum absolute value of horizontal ground acceleration and thus represents a measure of shaking at the given site. *PGA* can be expressed in terms of the gravitational acceleration  $g$  or in  $\text{m/s}^2$ .

Integrating the acceleration time series once yields velocity, given in  $\text{m/s}$ , and integrating it twice yields displacement, given in  $\text{m}$ . Analogue to the peak ground acceleration, the *peak ground velocity (PGV)* and the *peak ground displacement (PGD)* are defined as the maximum absolute values of the velocity and of the displacement time series, respectively.

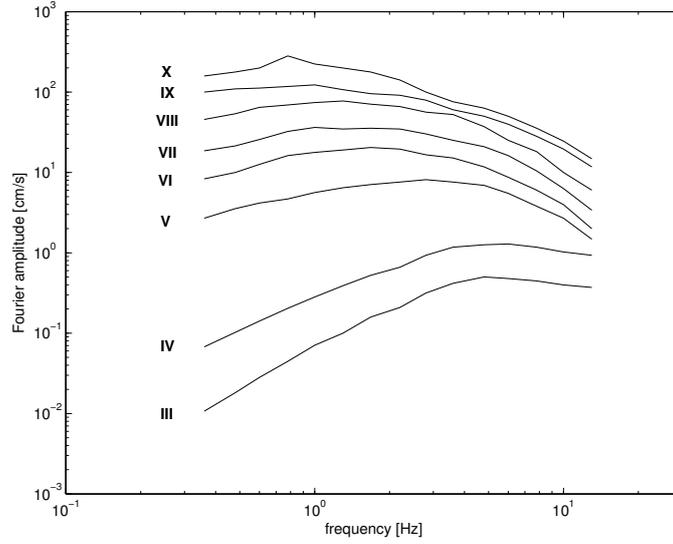
*Seismic intensity I* is a measure of the strength of observed ground motions. It usually refers to human description of the observed damages or experienced ground shaking. The most common intensity scales are the *Modified Mercalli Intensity (MMI)* scale, the *Medvedev-Sponhauer-Karnik (MSK)* scale, the *European Macroseismic Scale (EMS)*, and the *Japan Meteorological Agency (JMA)* scale (Chen and Scawthorn, 2003, and references therein). The *MMI* scale ranges from intensity I (generally unfelt) to XII (total destruction). The *MSK* scale is often assumed to

<i>Frequency</i> [Hz]	<i>Amplification</i> <i>NEHRP A</i> $V_{S30} = 2900 \text{ m/s}$	<i>Frequency</i> [Hz]	<i>Amplification</i> <i>NEHRP B</i> $V_{S30} = 1070 \text{ m/s}$	<i>Amplification</i> <i>NEHRP C</i> $V_{S30} = 520 \text{ m/s}$	<i>Amplification</i> <i>NEHRP D</i> $V_{S30} = 255 \text{ m/s}$
0.01	1.00	0.01	1.00	1.00	1.00
0.10	1.02	0.09	1.03	1.21	1.43
0.20	1.03	0.16	1.06	1.32	1.71
0.30	1.05	0.51	1.21	1.59	2.51
0.50	1.07	0.84	1.34	1.77	2.92
0.90	1.09	1.25	1.49	1.96	3.10
1.25	1.11	2.26	1.80	2.25	3.23
1.80	1.12	3.17	2.01	2.42	3.18
3.00	1.13	6.05	2.39	2.70	3.18
5.30	1.14	16.60	2.93	3.25	3.18
8.00	1.15	61.20	3.75	4.15	3.18
14.00	1.15				

**Table 2.2:** Site amplifications at selected frequencies after Boore and Joyner (1997). The amplification values for *NEHRP B* were determined by Böse (2006).



**Figure 2.1:** Combined effect of site amplification and high-frequency diminution for site class *NEHRP C* ( $V_{S30} = 520 \text{ m/s}$ ) and different decay parameters  $\kappa$  (Boore and Joyner, 1997).



**Figure 2.2:** Mean acceleration spectra for seismic intensities III-X after Sokolov (2002).

be identical with the MMI scale (e.g. Chernov and Sokolov, 1999).

The intensity can also be calculated from ground motion records and is then referred to as *instrumental* or *macroseismic* intensity.

A method to derive instrumental intensity from the Fourier amplitude spectra (FAS) of ground acceleration records is proposed by Chernov and Sokolov (1999) and Sokolov (2002), respectively. Sokolov (2002) gives a revised version of the method and the most recent summary:

From the horizontal components of about 1150 ground motion records the author derived a relationship that allows to calculate the seismic intensity (MMI or MSK scale) directly from spectral amplitudes in the frequency range of 0.4 - 13 Hz.

The relation has the form

$$\log_{10} A_f = 0.49 I - 2.0 \quad (2.21)$$

where  $A_f$  is the Fourier amplitude of the acceleration at a given frequency. The variances  $\sigma$  of  $A_f$  depend on the frequency: Sokolov (2002) assumes that the variances are smallest at the frequencies that are most “representative” for a given intensity. However, for a given intensity not a single representative frequency exist but instead a set of frequencies. Figure 2.2 displays the dependency of the different intensities on the frequencies at which the Fourier amplitudes are selected.

To determine the intensity at a site from a given acceleration record, following assumption can be made: The probability that  $\log_{10} A$  at frequency  $f_j$  will not exceed the mean values  $a_{i,j}$  that refer to the considered intensity  $I = i$  is given as

$$P[x_j \leq a_{i,j}] = 1 - \frac{1}{\sigma_{i,j} \sqrt{2\pi}} \int_{x_{min}}^{x_j} \exp\left(\frac{-(a_{i,j} - x_j)^{1/2}}{2\sigma_{i,j}^2}\right) dx \quad (2.22)$$

with  $x_j = \log_{10}A(f_j)$  and  $x_{min} = a_{i,j} - 5\sigma_{i,j}$ . The probability that the spectral amplitudes at not only one frequency but at all representative frequencies will not exceed the respective values is given as

$$P[x \leq a_i] = \frac{\sum_{j=1}^{n_f} P[x_j \leq a_{i,j}]w_{i,j}}{\sum_{j=1}^{n_f} w_{i,j}} \quad (2.23)$$

where  $n_f$  is the number of considered frequencies and  $w_{i,j} = \sigma_{i,min}^2/\sigma_{i,j}^2$  is the weight. The probability that the intensity  $I$  at the recording site will not exceed  $i$  can then be determined from

$$P[I \leq i] = \prod_i^{i=XII} P[x \leq a_i] \quad (2.24)$$

for intensities III-XII. The desired instrumental intensity value should be determined from either the maximum of the first derivative of the probability function  $P$  or from the intensity value at which  $P = 0.5$  (Sokolov, 2002). In this thesis, I will take the mean value of these two intensities.

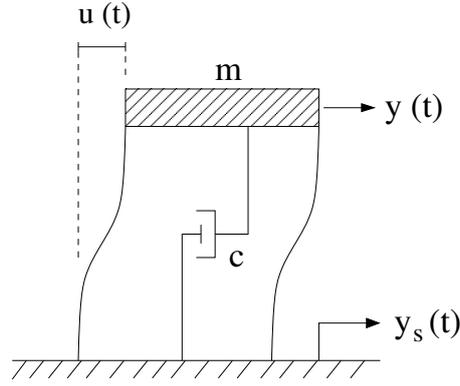
In contrast to the Fourier amplitude spectrum, an analysis of the *response spectrum* of the ground motion allows for analyzing the response of a single-degree-of-freedom (SDOF) system to the seismic excitation. For calculating the response of a building, for example, the system is most simplified modeled by a one-story building with one degree of freedom and with its mass concentrated in a rigid roof (Figure 2.3). The damping of the system is represented by coefficient  $c$ . The response of such a system towards an excitation can be obtained by applying *Duhamel's Integral* (e.g. Paz, 1994):

$$u(t) = \frac{1}{m\omega_D} \int_0^t F_{eff}(\tau) e^{-\xi\omega(t-\tau)} \sin \omega_D(t-\tau) d\tau. \quad (2.25)$$

Equation 2.25 defines the relative displacement  $u(t)$  at time  $t$  of a mass  $m$  caused by the continuous effective force  $F_{eff}(\tau)$ . The integral represents the sum of the differential displacements from time  $\tau = 0$  to time  $t$ . The effective force  $F_{eff}(\tau)$  is defined as the force on an oscillator of mass  $m$  excited by an acceleration function at its base. The damping of the oscillator is described by the damping ratio  $\xi = c/2m\omega$  with  $\omega$  being the natural frequency of the system. The quantities given in equation 2.25 are illustrated in Figure 2.3.

Engineering structures normally have damping ratios well below 1.  $\xi = 0\%$  implies an undamped oscillation, while  $\xi = 20\%$  means critical damping (Paz, 1994). The most commonly found damping ratio of engineering structures is 5%. The maximum response and thus the maximum damage of a building occurs at its fundamental period  $T_0$ , where the ground motion will result into resonance. As a rule of thumb, the fundamental period of a certain building can be estimated from its number of floors ( $N$ ):  $T_0$  is approximately  $N/10$  (Chen and Scawthorn, 2003).

The maximum absolute values of the relative displacement  $u(t)$ , velocity  $\dot{u}(t)$  and acceleration



**Figure 2.3:** Base excitation of a typical SDOF system of mass  $m$  with natural frequency  $\omega$ . The damping ratio is defined as  $\xi = c/2m\omega$ . The effective force is defined as  $F_{eff}(t) = -m\ddot{y}_s(t)$  where  $y_s(t)$  is the absolute displacement of the ground. The time-dependent displacement of mass  $m$  is given by  $u(t)$ , while  $y(t)$  is the absolute displacement of the mass (Paz, 1994).

$\ddot{u}(t)$  define the spectral values  $S_d$ ,  $S_v$ , and  $S_a$ , respectively:

$$\begin{aligned} S_d &= \max(|u(t)|) \\ S_v &= \max(|\dot{u}(t)|) \\ S_a &= \max(|\ddot{u}(t)|) \end{aligned} \quad (2.26)$$

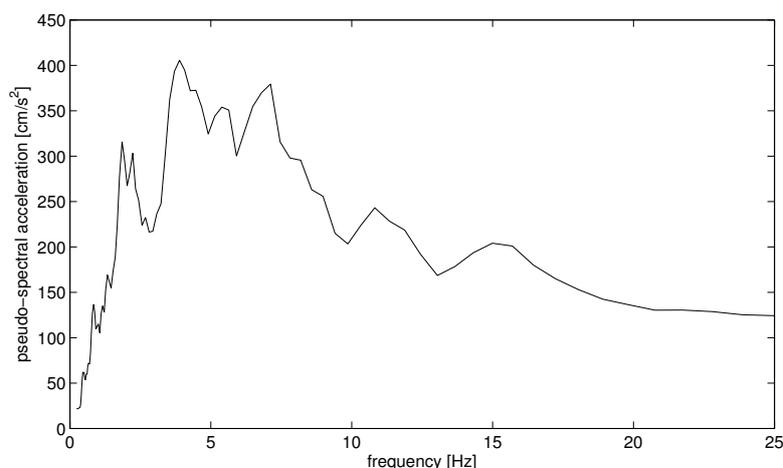
The quantities in equation 2.26 are called *spectral displacement*, *spectral velocity*, and *spectral acceleration*, respectively.  $S_v$  and  $S_a$  can be approximated by the *pseudo-spectral velocity (PSV)* and *pseudo-spectral acceleration (PSA)*:

$$\begin{aligned} PSV &\approx \omega S_d = (2\pi/T) S_d \\ PSA &\approx \omega^2 S_d = (2\pi/T)^2 S_d \end{aligned} \quad (2.27)$$

(e.g. Paz, 1994; Jennings, 2003). Figure 2.4 shows an example of an acceleration response spectrum.

Ground motion parameters are predicted by the application of *attenuation relationships*. These are functions that allow to predict the ground motions at a certain location generated by a specific earthquake. Input parameters into the prediction equations are the magnitude, the source-to-site distance, the local site condition, and usually also information about the fault geometry. Attenuation relationships are based on the extrapolation of observed ground motions which give constraints on the regions and the types of earthquakes that can be predicted by the specific attenuation relations.

The source-to-site distances can be specified differently. For point sources, either the *epicentral* or the *hypocentral distances* are defined. These are, however, poor measures for finite sources. For these, the most common distance measure is the *Joyner-Boore distance*  $r_{jb}$ . It is defined as the closest horizontal distance to the vertical projection of the rupture plane (Joyner and Boore, 1981). Further measures for an extended source are the *rupture distance*  $r_{rup}$  and the *seismic*



**Figure 2.4:** Acceleration response spectrum with 5% damping for a simulated acceleration time series for a  $M_w = 7.3$  earthquake at a source-to-site distance of 42 km and site class *NEHRP C*.

distance  $r_{seis}$ . The former is defined as the closest distance to the rupture plane, while the latter describes the distance to the seismogenic part of the rupture plane (Campbell, 2003, and references given therein).

Examples of attenuation relationships can be found in Ambraseys and Douglas (2003) or Campbell and Bozorgnia (2008) derived from worldwide observed data, in Gülkan and Kalkan (2002), Gülkan and Kalkan (2004) or Özbey et al. (2004) for Turkey, in Ambraseys et al. (1996) for Europe, or in Atkinson and Boore (1995) and Boore et al. (1997) for North America. The relationships from Boore et al. (1997), Campbell and Bozorgnia (2008), and Özbey et al. (2004) will be used in Chapter 2.7 and 3.1 and are defined in Appendix B.1, B.2, and B.3.

## 2.5 The Simulation Method *FINSIM*

For the simulation of seismic ground motion different simulation techniques have been developed and extended by various scientists. The simulations used in this study are generated using *FINSIM*, a method for simulating stochastic acceleration time histories from finite faults, by Beresnev and Atkinson (1997, 1998).

*FINSIM* is based on the stochastic method for simulation of ground motion by Boore (1983, 2003), a commonly used technique assuming the earthquake as a point source. This method has been applied and extended by many scientists, among them Beresnev and Atkinson (1997, 1998). A list of applications of Boore's method for the simulation of earthquakes in different regions of the world can be found in Boore (2003).

Before giving insight into *FINSIM* it is necessary to shortly introduce Boore's stochastic method. It is a simple but effective technique for predicting mean horizontal ground motions from shear waves by considering the earthquake as a point source. It combines functional descriptions of

the ground motion's amplitude spectrum with a random phase spectrum (Boore, 2003). The following portrait of the method is taken from Boore (1983) and Boore (2003):

The stochastic simulation technique originates from the work of Hanks and McGuire (1981), who assumed that high-frequency ground motions can be characterized as band-limited, finite-duration Gaussian white noise with an underlying amplitude spectrum specified by a simple seismological model. Analogue to equation 2.2, Boore's stochastic method describes the spectrum of ground motion as a combination of the effects from the earthquake source ( $S$ ), the path ( $P$ ), and the local site ( $F$ ):

$$Y(\mathbf{x}, \omega) = S(\mathbf{x}, \omega) \cdot P(\mathbf{x}, \omega) \cdot F(\mathbf{x}, \omega). \quad (2.28)$$

The acceleration source spectrum  $S(\mathbf{x}, \omega)$  follows the  $\omega^2$  model by Brune (1970, 1971) and can be expressed with the displacement source spectrum given in equation 2.9:

$$|S_s^a(\mathbf{x}, \omega)| = \omega^2 \frac{F^{surf}}{\sqrt{2}} |S_s^d(\mathbf{x}, \omega)|. \quad (2.29)$$

The index  $a$  stands for *acceleration* and  $s$  for *shear wave*. The parameter  $F^{surf}$  is the amplification due to the free surface and is usually set to 2. The term  $1/\sqrt{2}$  is a reduction factor that accounts for the partitioning of energy into two horizontal components.

As described in equation 2.12 the seismic moment  $M_0$  is linked to the stress drop  $\Delta\sigma$  by

$$M_0 \approx L^3 \cdot \Delta\sigma. \quad (2.30)$$

The path effect

$$P(\mathbf{x}, \omega) = G(\mathbf{x}) I(\mathbf{x}, \omega) \quad (2.31)$$

accounts for the geometrical spreading  $G(\mathbf{x})$  and the inelastic attenuation  $I(\mathbf{x}, \omega)$  as defined in Section 2.2. The path also includes the distance-dependency of the duration  $T_d$  of the seismic signal. The duration is specified by a trilinear model defined by Atkinson and Boore (1995) and will be given later.

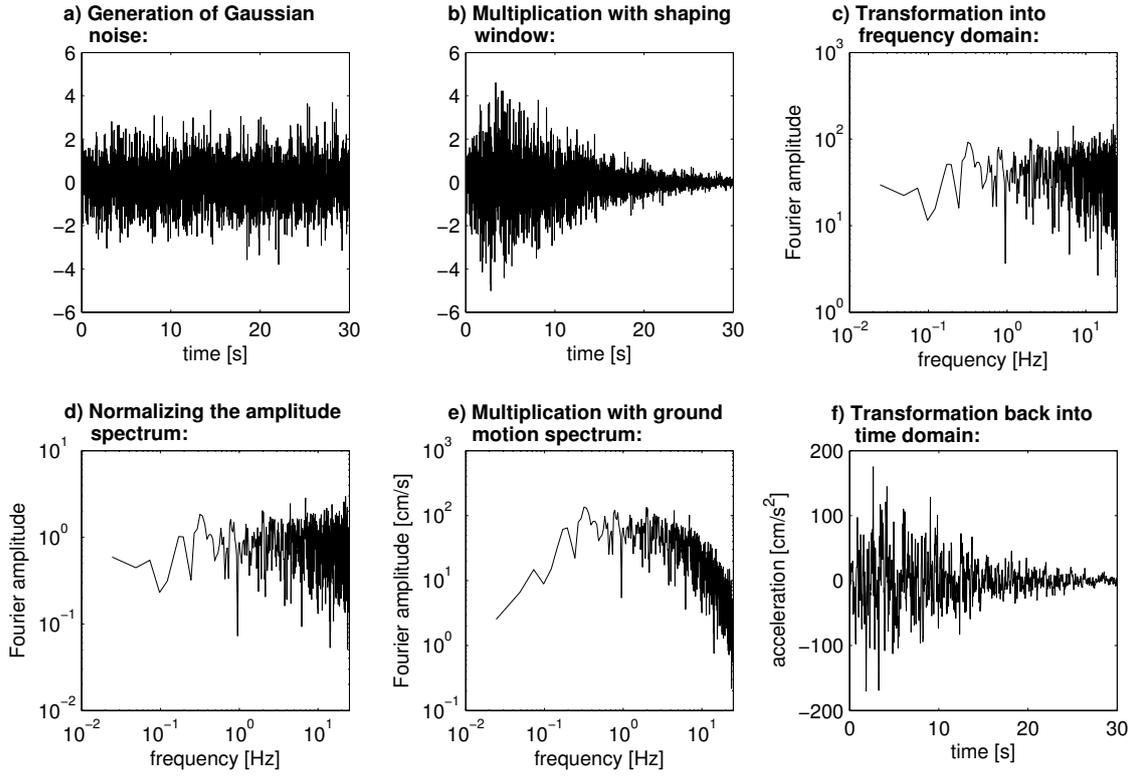
The local site effects are again defined as the product of the site amplification  $A(\mathbf{x}, \omega)$  and the high-frequency diminution  $H(\mathbf{x}, \omega)$ :

$$F(\mathbf{x}, \omega) = A(\mathbf{x}, \omega) \cdot H(\mathbf{x}, \omega). \quad (2.32)$$

The site amplifications are those specified in Table 2.2.

The actual calculation of the ground motion time series follows a simple scheme that is illustrated in Figure 2.5. In the first step, a window of Gaussian white noise with zero mean and unit variance is generated (Figure 2.5a). The length of the window corresponds to the duration of ground motion. This windowed noise is then multiplied with the envelope of the ground motion record to obtain a shape similar to an earthquake record (Figure 2.5b). The applied shaping window is defined after Saragoni and Hart (1974) and follows the functional form

$$w(t; \epsilon; \zeta; t_\zeta) = a(t/t_\zeta)^b \exp(-c(t/t_\zeta)) \quad (2.33)$$



**Figure 2.5:** Simulation of ground motion time series after Boore (2003).

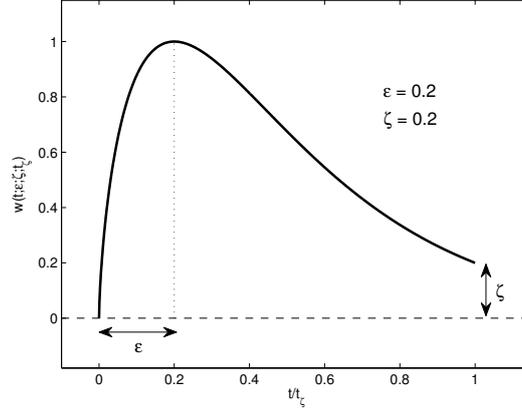
where  $a$ ,  $b$ , and  $c$  determine the position of the peak (Figure 2.6):

$$\begin{aligned}
 a &= (\exp(1)/\epsilon)^b \\
 b &= -(\epsilon \ln \zeta) / [1 + \epsilon (\ln \epsilon - 1)] \\
 c &= b/\epsilon
 \end{aligned} \tag{2.34}$$

Default values used by Böse (2006) are  $\epsilon = \zeta = 0.2$ .

In the third step, the windowed noise is transformed into the frequency domain via Fourier transformation (Figure 2.5c). The amplitude spectrum is then normalized by the square-root of the mean square spectrum (Figure 2.5d) and multiplied by the ground motion spectrum  $Y$  (Figure 2.5e). To receive the final ground motion time series, the resulting spectrum is transformed back into time domain via inverse Fourier transformation (Figure 2.5f). The obtained ground motion represents *mean horizontal* ground motion.

In *FINSIM*, Beresnev and Atkinson (1997, 1998) extended Boore's stochastic method by transferring it to finite faults. The authors follow the example of Hartzell (1978) and Irikura (1983) by subdividing the fault plane into smaller, rectangular subfaults of equal size representing separate point sources. Each point source is characterized by the source spectrum  $S(\mathbf{x}, \omega)$  from equation 2.28. The final seismic signal is obtained by adding up the contributions of all subfault signals.



**Figure 2.6:** Shaping window after Saragoni and Hart (1974) used to give the simulated ground motion time series a shape similar to an earthquake (Figure 2.5b).

After specifying which of the subfault elements contains the hypocentre the rupture spreads radially from this location on the fault. As soon as the rupture reaches the centre of a neighbouring subfault element, this element triggers and starts rupturing from its centre onwards to its edges. The crucial step in this procedure is to link the seismic moment and the corner frequency of each subfault element to its size  $\Delta l$  so that the final sum represents the intended source spectrum.

The relation between the rupture time  $\tau_r$  of each subsource and  $\Delta l$  is given by

$$\tau_r = \frac{\Delta l}{2y\beta} \quad (2.35)$$

The term  $y\beta$  represents the rupture velocity defined as the product of the shear-wave velocity  $\beta$  and a constant  $y$ , usually set to a value of 0.8 (Beresnev and Atkinson, 1998). The factor 2 is included because it is assumed that the rupture starts in the middle of the subfault and slips until it reaches the edges of the subfault, so only accounting for half of the rupture. By defining a parameter  $z$  as

$$z = \tau_r \omega_c \quad (2.36)$$

the corner frequency  $\omega_c$  of the subfault spectrum can be expressed in terms of  $\Delta l$  by

$$\omega_c = K \frac{\beta}{\Delta l} \quad \text{with } K = 2yz. \quad (2.37)$$

The parameter  $z$  is used in the definition of the maximum rate of slip on the fault,  $v_{max}$ :

$$v_{max} = \frac{2yz}{\exp(1)} \frac{\Delta\sigma}{\rho\beta} \quad (2.38)$$

with  $\Delta\sigma$  being the stress drop and  $\rho$  the density of the medium. The parameter  $z$  is controlled by the *FINSIM* input parameter *sfact*:

$$sfact = \frac{z}{1.68}. \quad (2.39)$$

Beresnev and Atkinson (1997, 1998) set  $z = 1.68$  so that  $sfact = 1$ . A variation of  $sfact$  leads to a variation of the maximum slip rate on the fault (equation 2.38).

According to equation 2.30, the link between the seismic moment  $M_{0,sf}$  of the subsource ( $sf$  means *subfault*) and  $\Delta l$  can be approximated by

$$M_{0,sf} \approx \Delta\sigma \Delta l^3. \quad (2.40)$$

The actual target moment  $M_0$  of the total seismic source is obtained by

$$M_0 = N M_{0,sf} = N \Delta\sigma \Delta l^3 \quad (2.41)$$

where  $N = M_0/M_{0,sf}$  is the number of subsources.

The subfault size  $\Delta l$  is implicitly given by defining the number of subfaults along strike and dip of the total fault that must be specified with fault length  $L$  and width  $W$  and with its strike  $\phi$ , dip  $\delta$  and rupture depth  $Z_r$ . Beresnev and Atkinson (1998) state that  $\Delta l$  should be in the range of about 5 to 15 km.

When adding up the contributions of all subsources the time delay between the triggering of the different elements must be taken into account. *FINSIM* implements this by lagging the different contributions by the time required for the rupture to reach the specific element plus the time which is necessary for the seismic waves to travel to the receiver (Beresnev and Atkinson, 1997, 1998).

For the simulation of the path effects  $P(\mathbf{x}, \omega)$  in equation 2.31, *FINSIM* accounts for the geometrical spreading by a trilinear model of the form

$$G(\mathbf{x}) = \begin{cases} r^{pow1}; & r \leq rg1 \\ rg1^{pow1} (r/rg1)^{pow2}; & rg1 < r \leq rg2 \\ rg1^{pow1} (rg2/rg1)^{pow2} (r/rg2)^{pow3}; & r > rg2 \end{cases} \quad (2.42)$$

with the constants  $pow1$ ,  $pow2$ ,  $pow3$ ,  $rg1$ , and  $rg2$  (Table A.1).

The inelastic attenuation  $I(\mathbf{x}, \omega)$  is specified in the form (equations 2.17 and 2.18)

$$I(\mathbf{x}, \omega) = \exp\left(-\frac{\omega}{2} \frac{r(\mathbf{x})}{Q\beta}\right) \quad \text{with} \quad Q = Q_0 \left(\frac{\omega}{2\pi}\right)^\eta. \quad (2.43)$$

As illustrated in Figure 2.5 the duration  $T_d$  of the desired signal must be specified when generating the Gaussian noise. *FINSIM* uses for this the trilinear model given by Atkinson and Boore (1995):

$$T_d = \begin{cases} \tau_r + durmin + b1(r - rmin); & rmin < r \leq rd1 \\ \tau_r + durmin + b1(rd1 - rmin) + b2(r - rd1); & rd1 < r \leq rd2 \\ \tau_r + durmin + b1(rd1 - rmin) + b2(rd2 - rd1) + b3(r - rd2); & r > rd2 \end{cases} \quad (2.44)$$

For distances  $r$  smaller than  $rmin = 10$  km the model defines  $T_d = \tau_r + durmin$ . The constants  $durmin$ ,  $b1$ ,  $b2$ ,  $b3$ ,  $rmin$ ,  $rd1$ , and  $rd2$  are displayed in Table A.1. The envelope function of

the applied shaping window given in equation 2.33 is set to  $w(t; \epsilon; \zeta; t_\zeta) = \zeta$  when  $t = t_\zeta = T_d$  (Boore, 2003).

The local site effects are taken into account by the high-frequency diminution,

$$H(\mathbf{x}, \omega) = \exp\left(-\pi\kappa(\mathbf{x})\frac{\omega}{2\pi}\right), \quad (2.45)$$

and by the frequency-dependent amplification which is defined for each *NEHRP* site classification. *FINSIM* automatically interpolates between the amplification factors given in Table 2.2.

For the application of the earthquake early warning method *PreSEIS* to the simulated earthquake scenarios, *FINSIM* was modified by Böse (2006) in two aspects: First, compressional waves were added to the simulation procedure as both Boore's method and *FINSIM* only simulate shear waves. The existence of P-waves is compulsory, however, because *PreSEIS* starts estimating the earthquake source parameters as soon as the seismic stations are triggered by the P-wave (see Chapter 4.2). The P-wave velocity  $\alpha$  is approximated by  $\alpha \approx \sqrt{3}\beta$ .

Second, a correct time axis was added to the simulated ground motions that account for the travel time differences between the source and the respective seismic stations. For this, the simulated ground motion was embedded into a time series of zero acceleration.

Boore's stochastic method has been specifically designed to simulate intermediate- to high-frequency ground motions that are of main interest for engineering applications. It is therefore highly useful for the applications in this work. The method proved to match successfully ground motions from a number of observed earthquakes despite its simple functional form (Equation 2.28). However, Boore's method suffers from a number of limitations: Due to the point source assumption the simulated ground motions do not include any effects that arise from the rupture propagation such as spatially as well as temporal inhomogeneous radiation of seismic energy, directivity effects due to interfering seismic waves, phase differences, or various components of motion (Boore, 2003). Additionally, the site effects are simplified and do not account for changes in S-wave velocity - non-linearities of soil layers can therefore not be reproduced (Böse, 2006). The extension of Boore's stochastic method to finite faults, on the contrary, is the clear advantage of the *FINSIM* method. By simply adding up the contributions of the stochastic subsources on the fault, Beresnev and Atkinson (1997) obtained a prediction of peak acceleration amplitudes with a precision of more than 15% for most of their simulations. However, more complex and thus more realistic ruptures than a planar fault cannot be realized with *FINSIM*.

Another important finding by Beresnev and Atkinson (1997, 1998) was the fact that the method predicts near-source ground accelerations accurately although near-field terms of the geometrical spreading are explicitly neglected (Beresnev and Atkinson, 1997). Altogether, *FINSIM* can be considered as the much more appropriate technique for simulating ground motion scenarios in the Marmara region for the purpose of this study.

Although simulation techniques drastically simplify the seismic sources and the wave propagation as well as site effects, the dataset generated with *FINSIM* (Section 2.6) is of good quality. Böse (2006) derived attenuation relationships for *PGA*, *PGV*, *PGD*, and *PSA* at 0.3 s, 1.0 s, and 2.0 s from the simulated dataset and compared these attenuations with observed data from

<i>Station No.</i>	<i>Station ID</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>NEHRP class</i>
s1	BOTAS	40.990	27.981	C
s2	BRGAZ	40.885	29.065	B
s3	BUYAD	40.870	29.128	B
s4	FARGE	40.760	29.355	B
s5	HVHRB	40.970	28.845	D
s6	HYBAD	40.876	29.086	B
s7	SINOB	40.999	28.540	C
s8	TUZ01	40.813	29.266	D
s9	YAKUP	40.991	28.668	C
s10	YLVHV	40.695	29.371	C

**Table 2.3:** Coordinates and *NEHRP* site classifications of the early warning stations of *IER-REWS*.

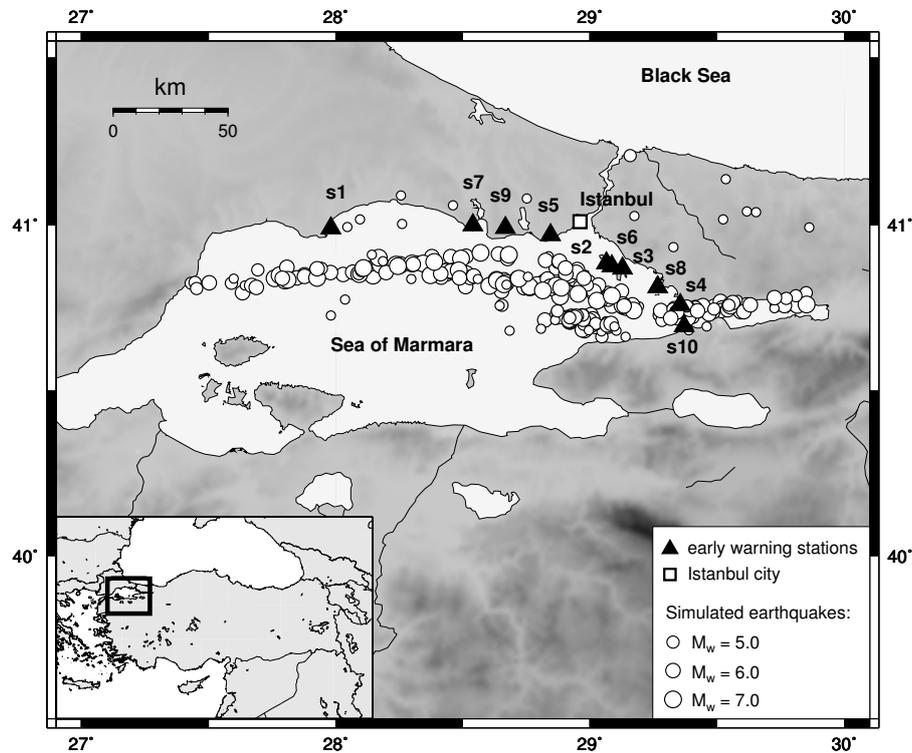
three earthquakes in northwestern Turkey and with attenuation relations from the literature. The simulated data show good agreements up to source-to-site distances of about 30 km. At larger distances, however, the ground motion parameters of the synthetic data attenuate more rapidly. The quality of the synthetic data also depends on the considered frequency of motion: The agreements with the observations and the attenuation relations from the literature are significantly higher for parameters derived from ground motions above 1 Hz (e.g. *PGA* and *PSA* at 0.3 s). For motions of 1 Hz or less the quality of the synthetic data decreases (Böse, 2006).

## 2.6 Simulated dataset

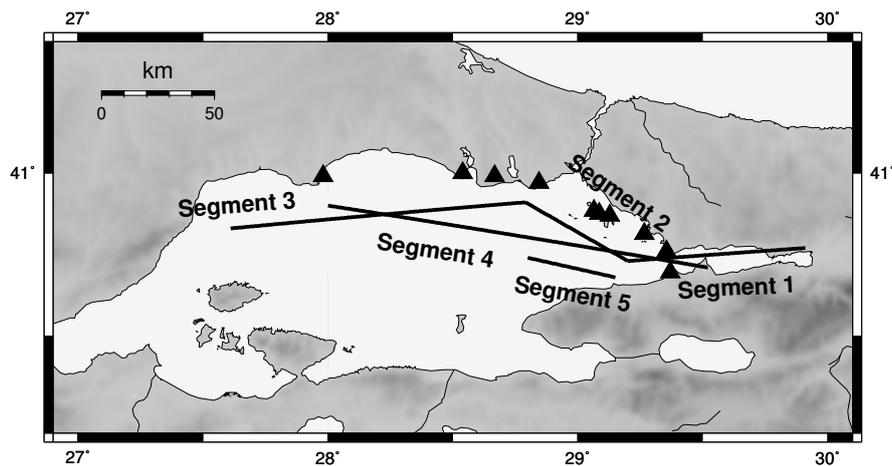
The simulation technique *FINSIM* (Beresnev and Atkinson, 1997, 1998) has been used by Böse (2006) to simulate a suitable dataset for testing the earthquake early warning method *PreSEIS* in the example region Istanbul. Due to the lack of real strong ground motion observations at the early warning stations the dataset was used to train the artificial neural networks (ANNs) and to determine possible warning times for the city of Istanbul.

For this purpose, the dataset had to fulfill several criteria. First, the number of events must be sufficiently large for a stable training process of the neural networks (see Chapter 4.2). Second, the dataset should mimic the spatial distribution of seismicity. This means that not only events located directly on the fault should be included, but also features like clusters and randomly distributed earthquakes further away from the faults to enhance the pattern recognition ability of *PreSEIS* after training of the ANNs. The slip distributions of the simulated earthquakes are generated by randomized weight matrices and show different directions of rupture propagations. The magnitude range includes also medium-sized earthquakes that would not cause damaging ground motions within Istanbul. The *FINSIM* input parameters for generating the dataset are specified in Appendix A.1 in Table A.1.

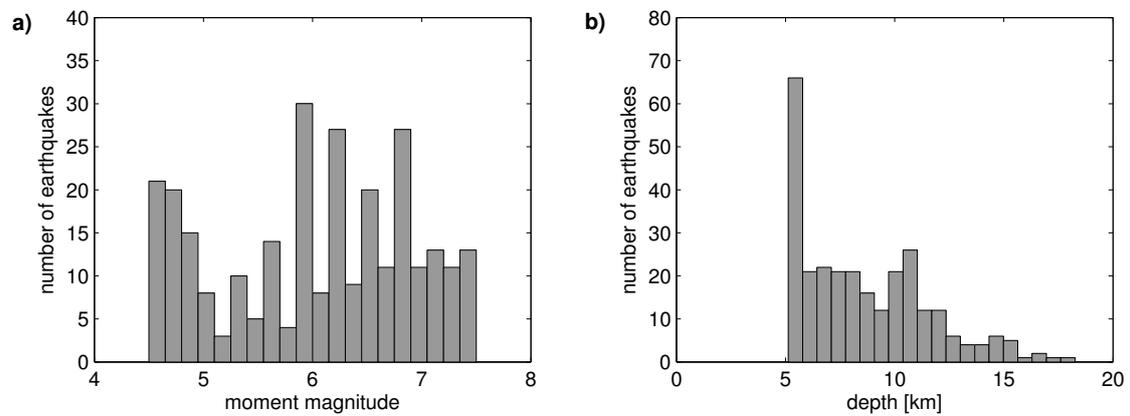
The final dataset consists of 280 simulated earthquake scenarios with moment magnitudes in the



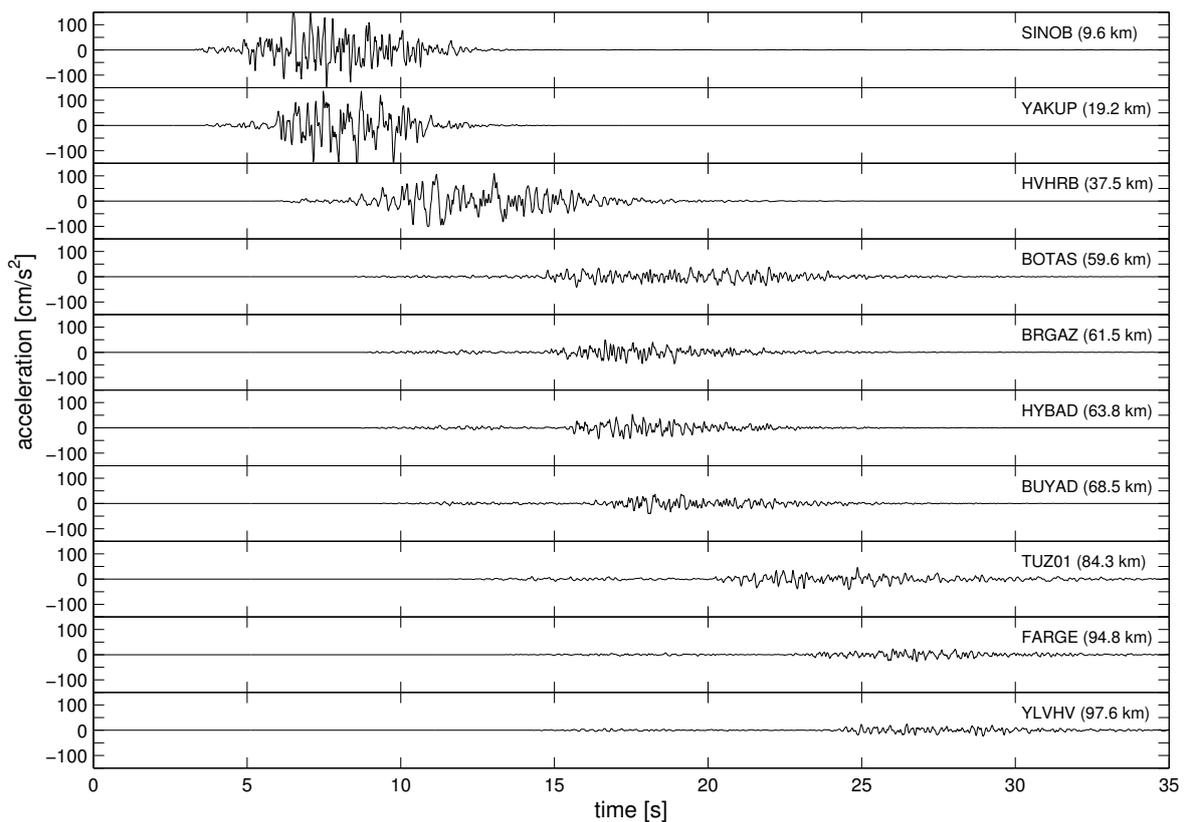
**Figure 2.7:** Distribution of the 280 simulated earthquake scenarios (white circles) in the Marmara region, Turkey. The moment magnitudes range between  $4.5 \leq M_w \leq 7.5$ . The data is simulated for the locations of the 10 stations of *IERREWS* (black triangles).



**Figure 2.8:** Definition of fault segments 1 to 5. A number of 50 earthquake scenarios is simulated along each segment.



**Figure 2.9:** Histograms of moment magnitude (a) and depth (b) of all 280 simulated earthquakes.



**Figure 2.10:** Example of acceleration time series of an  $M_w = 6.2$  earthquake on Segment 3 (Event-no. 106), simulated for the ten early warning stations. The time series are sorted with increasing source-to-site distances.

range of  $4.5 \leq M_w \leq 7.5$ . The spatial distribution of the earthquakes can be seen in Figure 2.7. The acceleration time series were simulated for the locations of the 10 early warning stations of *IERREWS* (Chapter 1.2 and Figure 2.7). The coordinates and site classifications of these stations are given in Table 2.3.

The distribution of the earthquake hypocentres is based on the locations of faults in the Sea of Marmara. Considering both a single through-going and a segmented *Main Marmara Fault*, Böse (2006) defined five fault segments to generate the earthquake scenarios (Figure 2.8): Segment 1 corresponds to the *Izmit Fault*, Segment 2 to the *Cinarcik Fault*, and Segment 3 to the *Western Fault*. Segment 4 characterizes joint ruptures of Segments 1 to 3. Segment 5 accounts for hypocentres located in the southeastern part of the Sea of Marmara. The segments have lengths between 30 and 130 km and are all pure strike-slip faults. Details can be found in Table A.2. Along each of these segments 50 earthquakes are simulated. An additional set of 30 earthquakes, titled as Segment 6, complements the dataset. These are randomly distributed over the whole area and show moment magnitudes of  $4.5 \leq M_w \leq 5.0$ .

The hypocentre locations, moment magnitudes, and rupture lengths of all 280 simulated earthquakes can be found in Tables A.3 to A.8. The distributions of moment magnitudes and depths are illustrated in Figure 2.9. The magnitude histogram shows that the distribution of magnitudes is arbitrary and does not follow the Gutenberg-Richter relation which quantifies the general observation that earthquakes of a given magnitude (e.g.  $M = 6$ ) occur approximately 10 times more frequent than earthquakes of one order of magnitude higher (e.g.  $M = 7$ ).

Figure 2.10 shows an example of the simulated acceleration time series at the ten stations for a  $M_w = 6.2$  earthquake (Event-no. 106 on Segment 3). It can be nicely seen that the onsets of ground motion are lagged with respect to the source-to-site distances. In addition, the longer the travelled distances are, the longer the durations of the records become and the stronger the ground motion amplitudes attenuate.

## 2.7 Optimization of *FINSIM* parameters

The dataset described in the previous section has mainly been established by taking default *FINSIM* input values from the literature (Böse, 2006). As discussed in Section 2.5, the chosen model parameters lead to systematic discrepancies between simulations and observations. The simulations show, for example, a strong attenuation of ground motion parameters at distances greater than about 30 km. Also, many of the simulated acceleration time series show a quite “smooth” shape of the envelope without significant P-wave motions, as can be seen in Figure 2.11c. The characteristics of the P-waves, however, strongly influence the rapid estimation of the earthquake source parameters since they represent the first available information on ground motion at the seismic stations. Many earthquake early warning methods use exclusively the first few seconds of P-wave motion to estimate source parameters or the impending ground motions (Chapter 4).

This section will present an attempt to improve the characteristics of the simulated ground motions with the purpose of improving the *PreSEIS* performance. More characteristic acceleration

Latitude	Longitude	Magnitude	Depth [km]	Date	Region
40.76°N	29.97°E	7.4	19.6	17/08/1999	Izmit
40.77°N	30.10°E	5.8	19.6	13/09/1999	Izmit
40.75°N	29.92°E	5.2	17.7	31/08/1999	Izmit
40.70°N	29.34°E	4.8	12.0	29/09/1999	Yalova

**Table 2.4:** Observed earthquakes selected for the *FINSIM* input parameter optimization.

time series are assumed to significantly influence the pattern recognition ability of the artificial neural networks. For this purpose, I simulated some observed earthquake records from the Marmara region using the same *FINSIM* input parameters as for the synthetic dataset. I then visually compared the observed acceleration records and their simulated counterparts.

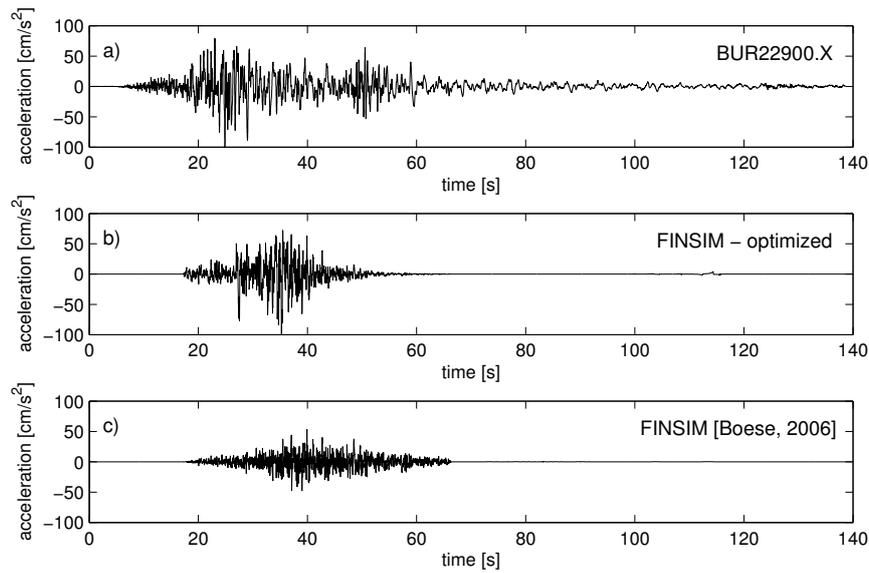
For the optimization of *FINSIM* input parameters, 20 records of four observed earthquakes in the Marmara region were selected from an earthquake catalogue kindly made available by *KO-ERI*. The criteria for selecting the earthquake records were 1) source-to-site distances ranging between 20 and 150 km, 2) magnitudes above 4.5, and 3) records of good quality, i.e. with low noise levels. The four selected events are specified in Table 2.4. Included are records of the 17 August 1999  $M_w$  7.4 Kocaeli earthquake.

The selected records were simulated using *FINSIM* and were then visually compared to the observed records. The visual inspection concentrated mainly on the shape of the envelopes with main focus on the P-wave onsets and on the frequency content. By tuning certain *FINSIM* input parameters such as the stress drop, the parameters for the inelastic attenuation, the high-frequency diminution, or the constants of the applied shaping window, simulation variations could be obtained that better fit the observed earthquake records.

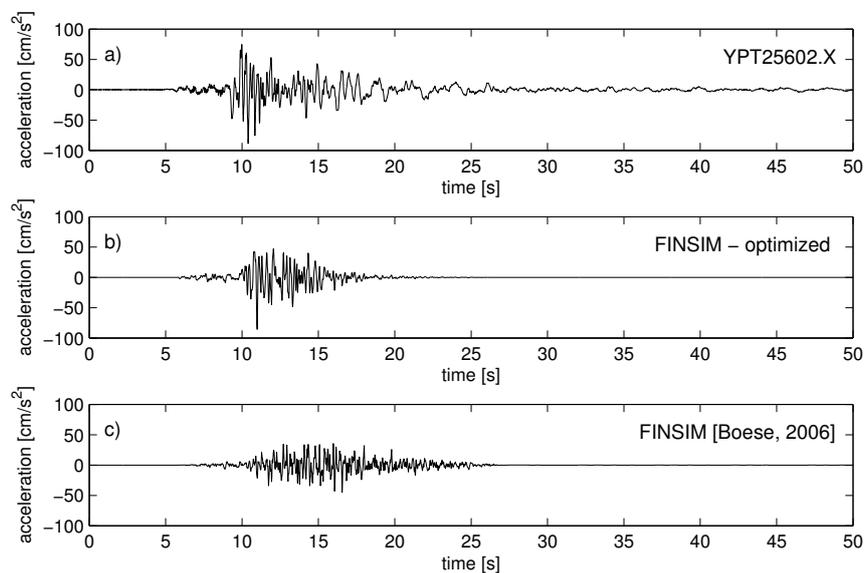
The parameterization that gave the best results consists of following changes: The parameters  $\zeta$  and  $\epsilon$  of the shaping window (Figure 2.6) are set to  $\zeta = 0.01$  and  $\epsilon = 0.005$  (instead of  $\zeta = \epsilon = 0.2$ ). The high-frequency diminution values  $\kappa$  for each *NEHRP* class are increased by a constant of 0.03. This reduces the site amplifications of the high frequencies, as it was shown in Figure 2.1. Additionally, the parameter  $Q_0$  of the crustal attenuation model is changed to  $Q_0 = 70$  (instead of 50).

Figures 2.11 and 2.12 show examples of the improved simulations. Displayed are the records of the north-south component of each the  $M_w$  7.4 Kocaeli earthquake (Figure 2.11a) and the  $M_w$  5.8 earthquake (Figure 2.12a) at epicentral distances of 94.3 km and 28.5 km, respectively. In general, the optimized simulations, compared to the ones obtained from the original parameterization by Böse (2006), fit better in amplitudes and overall shape, although the duration of the records suffers from the increased attenuation.

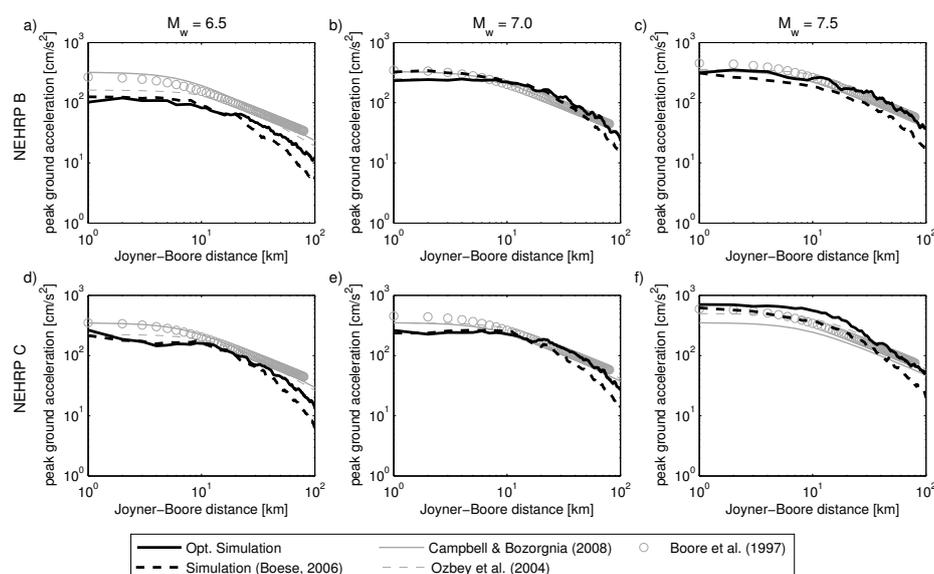
The amplitudes of the P-wave motions increase, while the high frequencies of the original simulations can be slightly reduced which corresponds better to the observations. However, the simulations cannot fully reflect the observed seismograms. *FINSIM* does not account for surface waves, or for reflected, refracted or diffracted waves which make the observed seismograms look



**Figure 2.11:** Observed record (NS-component) of the  $M = 7.4$  Kocaeli earthquake at an epicentral distance of 94.3 km (a). The simulation of this record using the optimized *FINSIM* input parameters is shown in (b), while (c) gives the simulation using the original parameterization after Böse (2006).



**Figure 2.12:** Observed record (NS-component) of the  $M = 5.8$  earthquake at an epicentral distance of 28.5 km (a). The simulation of this record using the optimized *FINSIM* input parameters is shown in (b), while (c) gives the simulation using the original parameterization after Böse (2006).



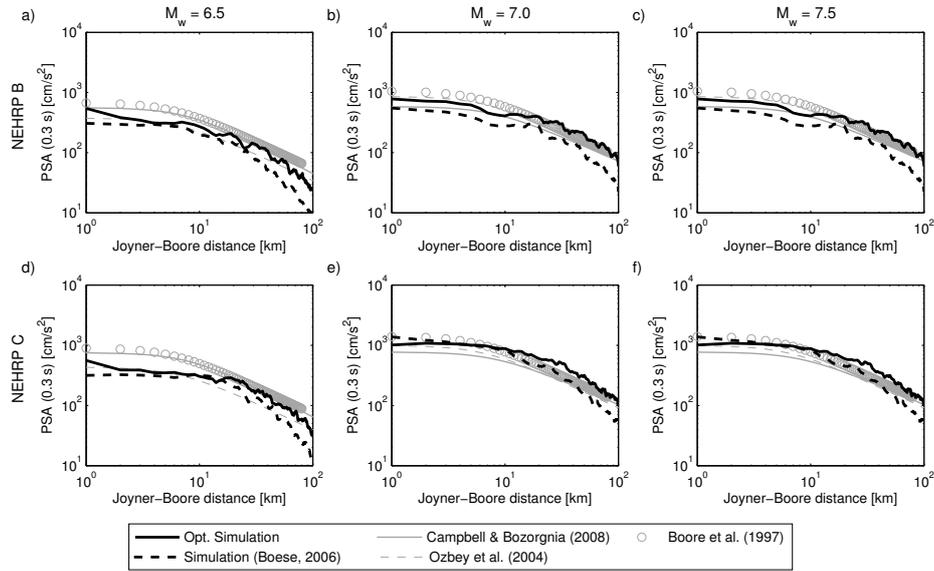
**Figure 2.13:** Attenuation of *PGA* with distance of the optimized simulations (solid bold black lines) and the original simulations (dashed bold black lines) of Böse (2006). The attenuation curves of the simulations are smoothed over 5 km. The attenuation relationships from the literature are displayed in grey symbols.

much more complex. Additionally, the assumed source model might not be appropriate in some cases, as suggested from Figure 2.11. The applied slip distribution may not correctly reflect the real one which shows in general very complex patterns. These points may explain the found differences between observed and simulated waveforms, especially as northwestern Turkey is characterized by a complicated subsurface structure and complex rupture processes.

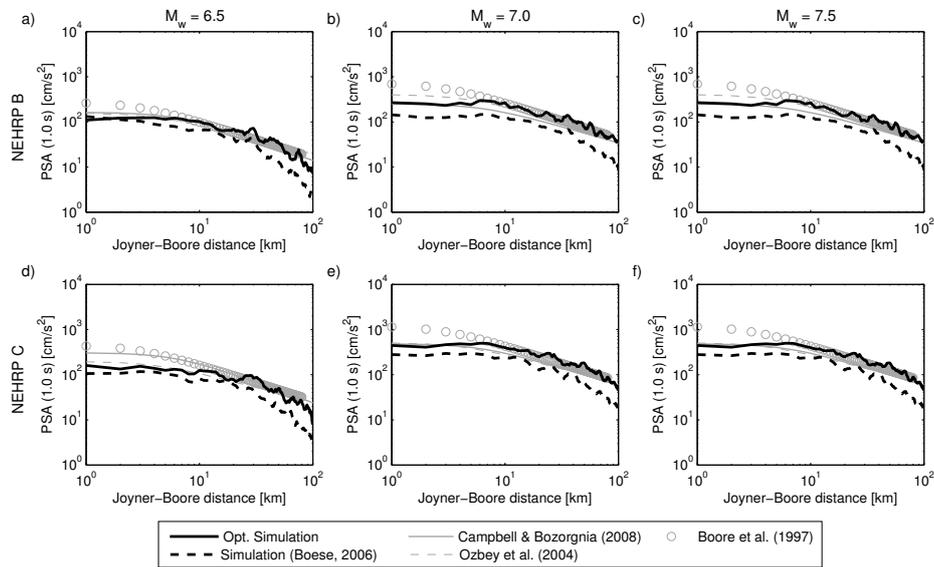
Although the optimization of the simulation parameters is not able to influence the P-wave onsets as strongly as it was intended, and although the variations of input parameters naturally yield different results for each considered earthquake record, the selected variation of the *FINSIM* input parameters represents the best compromise that was realizable by pure visual comparison. I therefore decided to re-calculate the synthetic dataset of the 280 earthquakes using the optimized *FINSIM* parameterization.

The enhancement of the dataset quality is supported by the comparison of attenuation of *PGA* and *PSA* at 0.3 s and 1.0 s with attenuation relationships from the literature (Figures 2.13 to 2.15). Selected were attenuation relations from Campbell and Bozorgnia (2008) for worldwide earthquakes, from Özbey et al. (2004) for northwestern Turkey, and from Boore et al. (1997) for western North America, whereas the one from Özbey et al. (2004) can be considered as the most suitable in this case. The three attenuation relationships are defined in Appendix B.

The attenuation of all three optimized ground motion parameters show a better agreement with the standard attenuation relationships than the original simulations by Böse (2006) which slightly stronger underestimate the ground motions. The attenuation at distances greater than 30 km can



**Figure 2.14:** Attenuation of  $PSA$  at 0.3 s with distance of the optimized simulations (solid bold black lines) and the original simulations (dashed bold black lines) of Böse (2006). The attenuation curves of the simulations are smoothed over 5 km. The attenuation relationships from the literature are displayed in grey symbols.



**Figure 2.15:** Attenuation of  $PSA$  at 1.0 s with distance of the optimized simulations (solid bold black lines) and the original simulations (dashed bold black lines) of Böse (2006). The attenuation curves of the simulations are smoothed over 5 km. The attenuation relationships from the literature are displayed in grey symbols.

also be improved. The optimized simulations show better agreements up to about 80 km. However, both simulations show an underestimation of *PGA* at distances smaller than about 10 km for the smallest considered magnitude  $M_w = 6.5$ . At this magnitude, the *PSA* values are underestimated similarly. However, this mainly holds true for the attenuation relations from Boore et al. (1997) and Campbell and Bozorgnia (2008). These two (and especially the one from Boore et al. (1997)) seem to generally overpredict the ground motions at short distances compared to the attenuation derived from local earthquakes by Özbey et al. (2004). Due to its restriction to earthquakes from the same region than the synthetics the latter seems to be more appropriate.

As a summary, the optimization of the *FINSIM* input parameters significantly improves the quality of the simulated dataset despite the simple procedure of visual comparison. As expected, this improvement manifests itself also in the *PreSEIS* performance: the estimates of the earthquake magnitudes within the first few seconds after the P-wave onsets can be improved by up to 34% compared to the performance using the original data by Böse (2006). This is illustrated in Appendix A.2.

Any results obtained from the synthetic dataset in the following chapters will always refer to the optimized simulations if not stated otherwise.

## Chapter 3

# The historic 1509 Istanbul earthquake

In this chapter, the *FINSIM* technique will be used to simulate a historic earthquake that hit Istanbul in the year of 1509. With a magnitude of about 7.3 the earthquake was one of the largest and most destructive earthquakes that struck Istanbul during the last 500 years. The ground motions in the metropolitan area of Istanbul will be simulated using three different earthquake source models that differ in the location of the hypocentre. Additionally, the ground motions will be predicted by using attenuation relationships. The purpose of this study is the demonstration of the variability of ground motions resulting from different models and prediction techniques. Additionally, the simulated ground motions will be used to predict building damages using the building stock of the today's city of Istanbul to demonstrate what impact the historic earthquake would have had on the modern megacity.

The historic earthquake that is simulated in this study occurred on 10 September 1509, probably around 10 p.m., at 40.9°N and 28.7°E (Ambraseys, 2001). The earthquake had a magnitude of about 7.3, although size estimates differ from study to study: Ambraseys (2001) estimates a moment magnitude of  $M_w = 7.2 \pm 0.3$ . With a rupture length of  $L = 70$  km and an assumed fault width  $W = 10$  km the author estimates a seismic moment of  $M_0 = 7.35 \cdot 10^{19}$  Nm. However, the author allows for an uncertainty in fault size of  $10 \leq W \leq 15$  km and  $40 \leq L \leq 100$  km. Ambraseys and Jackson (2000) estimate a rupture length of 74 km. In both Ambraseys (2002) and Ambraseys and Jackson (2000) the authors estimate a surface wave magnitude of  $M_s = 7.2$ , whereas in Ambraseys (2002) the author additionally assumes that  $M_s = M_w$ . Ambraseys and Finkel (1990) estimate a magnitude of  $M_s = 7.4$ , while Parsons et al. (2000) and Parsons (2004) declare magnitudes of 7.6 and 7.3, respectively.

According to the investigations of Ambraseys (2001) the earthquake caused considerable damage in Istanbul and its direct vicinity on both sides of the Bosphorus, however not at greater distances. There, the damage was only sporadic and not serious, although historic sources prove that the earthquake was felt within a radius of 750 km (Ambraseys and Finkel, 1991). The earthquake killed about 1,500 - 5,000 people in Istanbul which most probably represented between 0.4 and 2% of its population within the city walls; around 10,000 people were injured. The population in the year 1509 is estimated to be around 250,000 people. The contemporary city walls sur-

Fault length:	70.0 km	<b>Fault start:</b>
Fault width:	18.86 km	40.8830°N / 28.8100°E
Fault strike:	265°	<b>Fault end:</b>
Fault dip:	90°	40.8281°N / 28.1829°E
Fault depth:	0.4 km	
No. of subfaults along strike:	7	<b>Hypocentre location:</b> (depth = 9.83 km)
No. of subfaults along dip:	3	Version 1: 40.8556°N / 28.4964°E
Moment magnitude:	7.3	Version 2: 40.8791°N / 28.7652°E
Stress drop:	5.0 MPa	Version 3: 40.8321°N / 28.2277°E

**Table 3.1:** Simulation parameters for the historic 1509 Istanbul earthquake. The remaining *FINSIM* input parameters are set according to Table A.1.

rounded an area of about 17 km<sup>2</sup> with a population density of 14,500 persons per km<sup>2</sup> living in 54,000 households. Most of the ordinary houses were made of wood with probably 1 or 2 stories. Around 1,500 houses were destroyed, corresponding to about 0.5% of the total number of buildings. Assuming this population density and building stock, Ambraseys (2001) estimates seismic intensities in the city of Istanbul ranging between VII and X on the MSK intensity scale. The earthquake also damaged large parts of the city walls and generated sea waves in the Sea of Marmara that flooded large parts of coastal areas (Ambraseys, 2001).

The following section describes how the seismic source of the 1509 earthquake is defined in this study and presents the simulated ground motions.

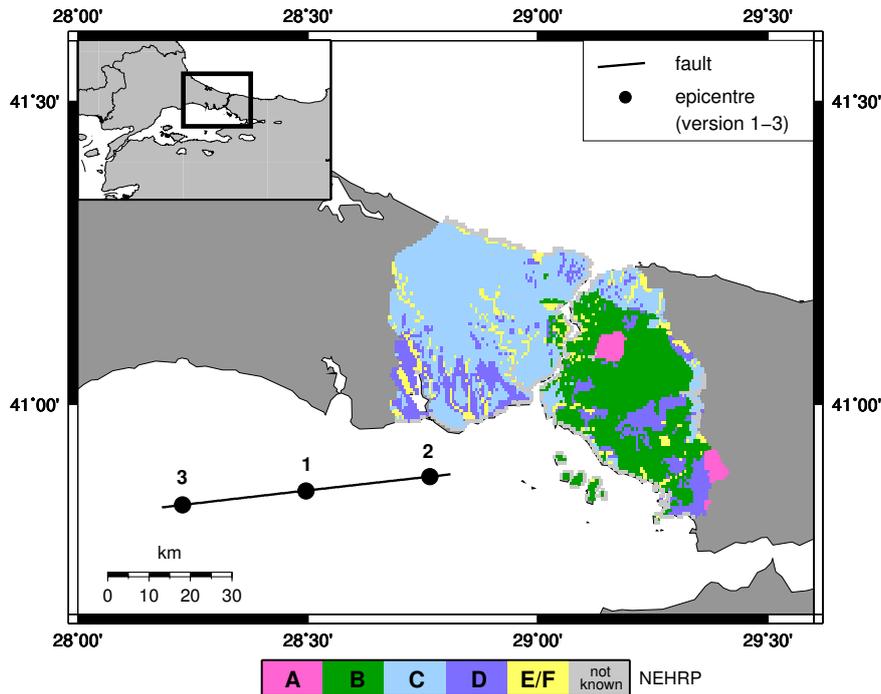
### 3.1 Simulation results

The definition of the fault segment that ruptured during the 1509 earthquake follows the work of Hergert (2009) who investigated what the likely size and location of a future earthquake in the Sea of Marmara would be. I define the fault rupture to be equivalent to Hergert's definition of the Central Segment of the northern Marmara Sea basin, because this segment shows a natural boundary at its eastern end due to the "Istanbul Bend" (see Figure 1.2).

The segment has a rupture length of approximately 70 km which correlates well with the published magnitude of 7.3. The fault is defined as purely strike-slip with a fault strike of 265° and a depth of the upper fault edge of 400 m. The width of the fault is set to almost 19 km (Hergert, 2009, personal communication). Assuming the hypocentre being located in the middle of the fault, its coordinates are given as 40.86°N and 28.50°E at a depth of 9.8 km.

For the *FINSIM* simulation, the fault plane is discretized into 7 subfaults along strike and 3 subfaults along dip, resulting into subfault sizes of 10.0 x 6.29 km. The slip on the subfault elements ranges between 0.0 and 6.6 m and has been generated randomly.

The magnitude in my model is defined as  $M_w = 7.3$  with a static stress drop of 5 MPa. The stress drop is inferred from comparisons of the simulated ground motions with attenuation relationships from Boore et al. (1997), Campbell and Bozorgnia (2008), and Özbey et al. (2004)



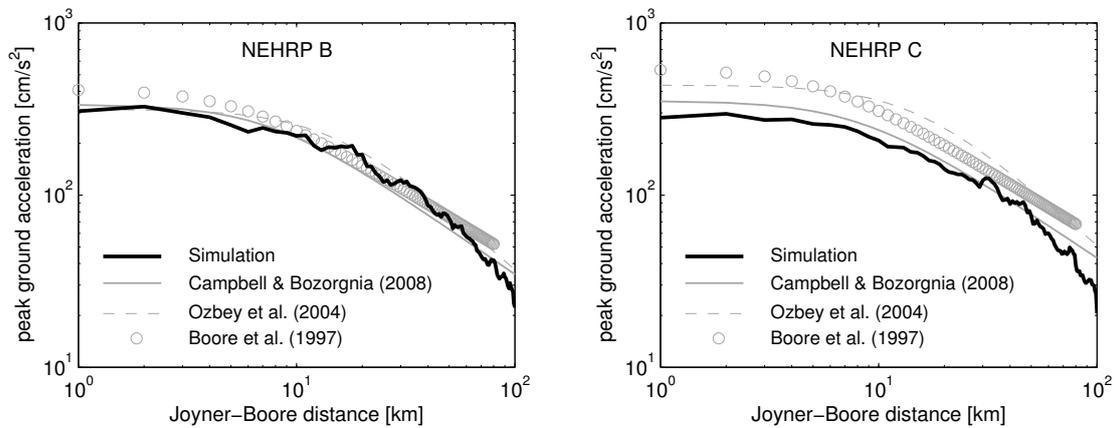
**Figure 3.1:** Location of the ruptured fault and of the epicentre versions 1 to 3. The *NEHRP* site classifications are defined for a dense grid of  $0.005^\circ \times 0.005^\circ$  spacing.

by setting the remaining *FINSIM* input parameters according to the values listed in Table A.1. The model parameters for the 1509 earthquake are summarized in Table 3.1. The selected fault length of 70 km is consistent with the one estimated by Ambraseys (2001).

For the purpose of investigating the influence of directivity effects, three versions of the scenario are calculated: Version 1 assumes the hypocentre being located in the middle of the fault. Version 2 sets the hypocentre to the eastern end of the fault, i.e., closest to Istanbul, while version 3 defines it to be located at the western end of the ruptured fault. The slip distribution on the fault is the same for all three scenarios.

The ground motions in Istanbul for the three scenarios are calculated for a dense grid of 8131 grid elements with  $0.005^\circ \times 0.005^\circ$  spacing. For each grid element the *NEHRP* site classifications (Building Seismic Safety Council, 1997) are defined (e.g. Erdik et al., 2003a). The ruptured fault segment, the hypocentre locations and the site classifications are displayed in Figure 3.1.

To ensure that the source parameters for the simulations are chosen in a reasonable way the simulated ground motions based on the hypocentre location in the middle of the fault are compared with the attenuation relationships of Boore et al. (1997), Campbell and Bozorgnia (2008), and Özbey et al. (2004). Figure 3.2 shows the attenuation of peak ground acceleration over source-to-site distances of up to 100 km for the site classifications *NEHRP B* and *C*. For *NEHRP B* the simulated *PGA* matches the relationships well up to a distance of about 70 km. At greater



**Figure 3.2:** Attenuation of *PGA* with distance for *NEHRP B* and *C* site classifications of the simulated ground motion (bold black curves) and the attenuation relationships from the literature (grey curves). The simulated curves are smoothed over 5 km.

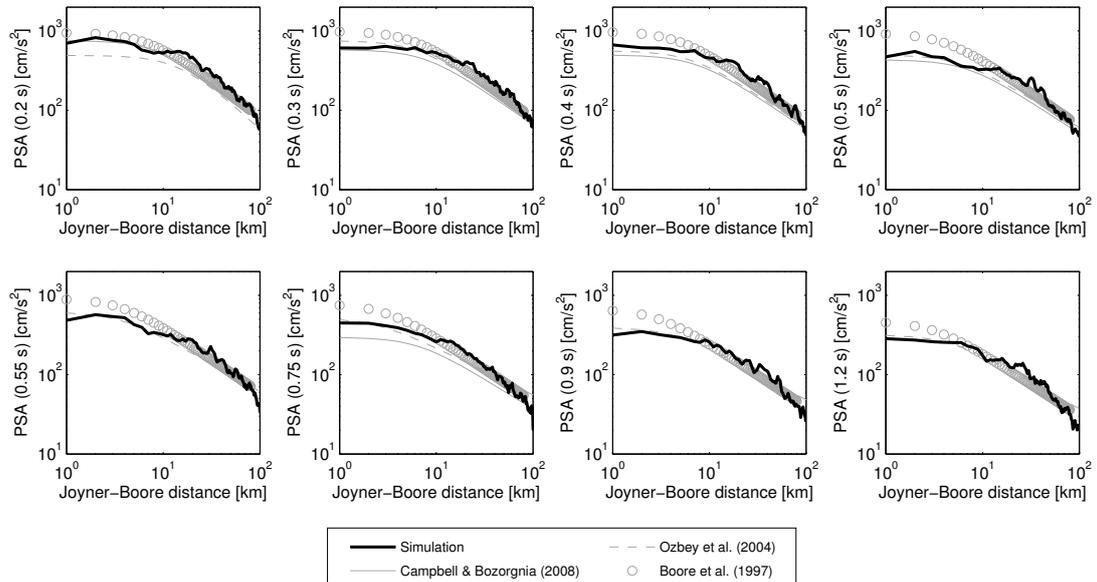
distances the *FINSIM* simulation attenuates stronger. As in Chapter 2.7, the attenuation relationships of Boore et al. (1997) give considerably higher ground motion amplitudes for distances up to 20 or 30 km for both *PGA* and *PSA* for the two site classifications. This agrees with observations summarized in the report of American Red Cross - Bogazici University (2002): Standard attenuation relationships all overpredicted the *PGA* values observed in the 1999 Kocaeli earthquake at distances less than about 20 km.

For *NEHRP C* (Figure 3.2) the peak ground accelerations of the *FINSIM* simulation are consistently lower than the ones predicted by the attenuation relationships.

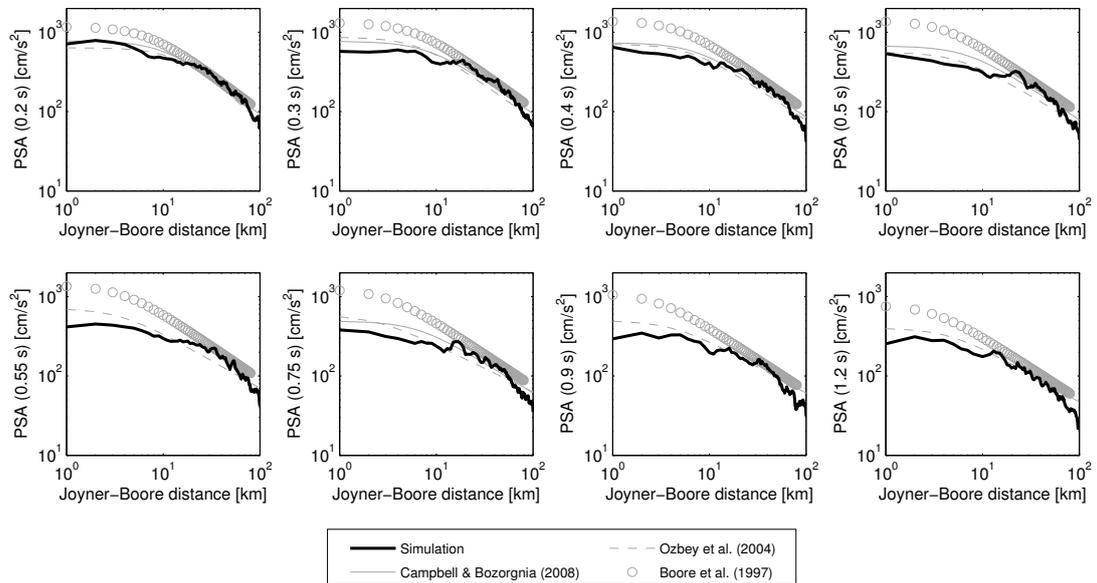
Figure 3.3 and 3.4 compare the attenuation of the pseudo-spectral acceleration at selected frequencies with the predicted values from the attenuation laws. The selected frequencies are those used in the building damage estimation method discussed in the next section. As for *PGA*, the simulations show higher agreements with the attenuation relationships for *NEHRP B* than for *C*, although the difference here is less significant. For *NEHRP B* the *PSA* values of the simulations are slightly overestimated at distances between about 10 and 50 km. In contrast, the simulations underestimate *PSA* for *NEHRP C* at small distances up to about 10 or 15 km.

Nonetheless, the average differences between the simulated ground motions and the attenuation relationships are only minor, proving that the chosen source model for the historic 1509 earthquake is of good quality.

Figure 3.5 shows the resulting distributions of *PGA* and seismic intensity in Istanbul for the three hypocentre versions. In all three cases the highest ground motions occur on the European side with the peak values located in the southwest region. The highest ground motions result from the first version, where the hypocentre is located in the middle of the fault. Here, *PGA* ranges between 23.1 and 456.9 cm/s<sup>2</sup> and the seismic intensity *I* between 4.4 and 9.1. The sec-



**Figure 3.3:** Attenuation of *PSA* at 0.2, 0.3, 0.4, 0.5, 0.55, 0.75, 0.9, and 1.2 s for *NEHRP B* site classification of the simulated ground motion (bold black curves) and the attenuation relationships from the literature (grey curves). The simulated curves are smoothed over 5 km.



**Figure 3.4:** Attenuation of *PSA* at 0.2, 0.3, 0.4, 0.5, 0.55, 0.75, 0.9, and 1.2 s for *NEHRP C* site classification of the simulated ground motion (bold black curves) and the attenuation relationships from the literature (grey curves). The simulated curves are smoothed over 5 km.

ond largest ground motions are achieved by model 3 with the hypocentre located on the western end of the fault. *PGA* ranges here between 25.3 and 445.9 cm/s<sup>2</sup> and *I* between 4.3 and 9.0. The ground motion patterns of versions 1 and 3 look almost similar, although the rupture in version 3 radiates slightly more energy eastwards in direction along the extended fault line. This leads to a minimal increase of the *PGA* values on the Asian side of Istanbul, especially in the coastal area. However, the intensity distributions do not reflect this pattern. They rather show higher intensity values for version 1 on the European part of Istanbul. The intensity shows less attenuation with radial distance around the eastern end of the fault rupture leading to slightly larger intensities in the northernmost part of Istanbul compared to scenario 3.

Version 2, with the hypocentre being located on the eastern end of the fault, results into *PGA* values between 21.1 and 391.2 cm/s<sup>2</sup> and an intensity between 4.2 and 8.6. It does not only exhibit smaller maximum values of ground motion than the other two models, but also shows considerably smaller ground motions over large areas on the Asian part of Istanbul, especially noticeable in the intensity distribution. This is due to the westwards propagating rupture.

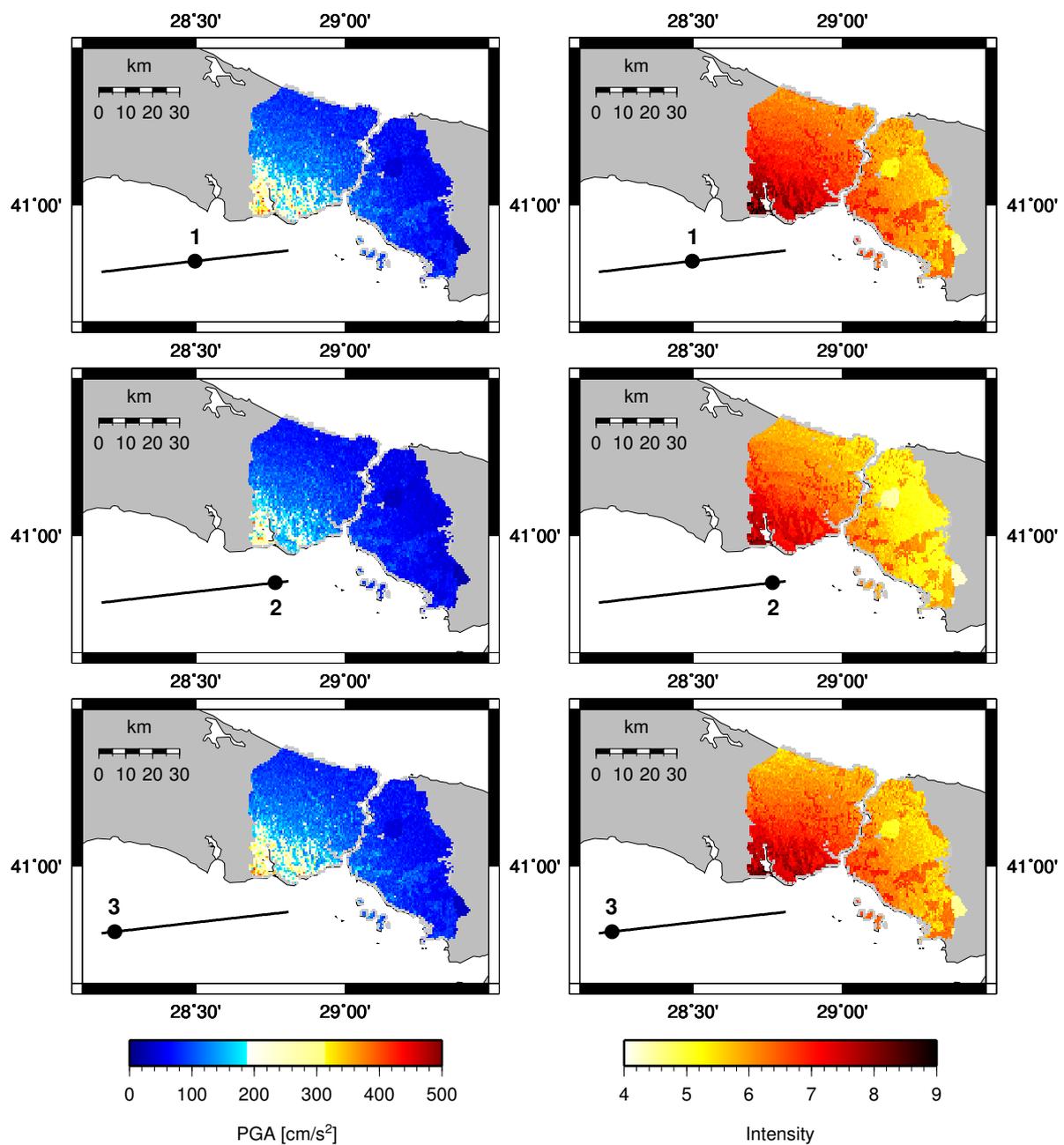
The *PGA* values observed during the 1999 Kocaeli earthquake at near-field stations range between 140.6 and 399 cm/s<sup>2</sup> for distances between 3 and 18 km, depending on the site conditions, whereas the distance is defined as the shortest distance between the observed fault rupture and the station. The maximum value of 399 cm/s<sup>2</sup> was observed at a station at 3 km distance that probably incorporates effects of a shallow thin soil layer (Akkar and Gülkan, 2002). Such a short source-to-site distance is not present in the simulation results for the 1509 earthquake (Figure 3.5). The attenuation curve of the simulated peak ground acceleration (Figure 3.2) shows, however, that *PGA* at 3 km distance is only about 290 cm/s<sup>2</sup> assuming a dense soil site.

Another record of the 1999 Kocaeli earthquake revealed a peak ground acceleration of 376 cm/s<sup>2</sup> at a distance of 18 km (Akkar and Gülkan, 2002). This fits, on average, well to the simulated *PGA* values in the south-west coastal region of Istanbul in Figure 3.5, strongly depending on the site conditions, of course.

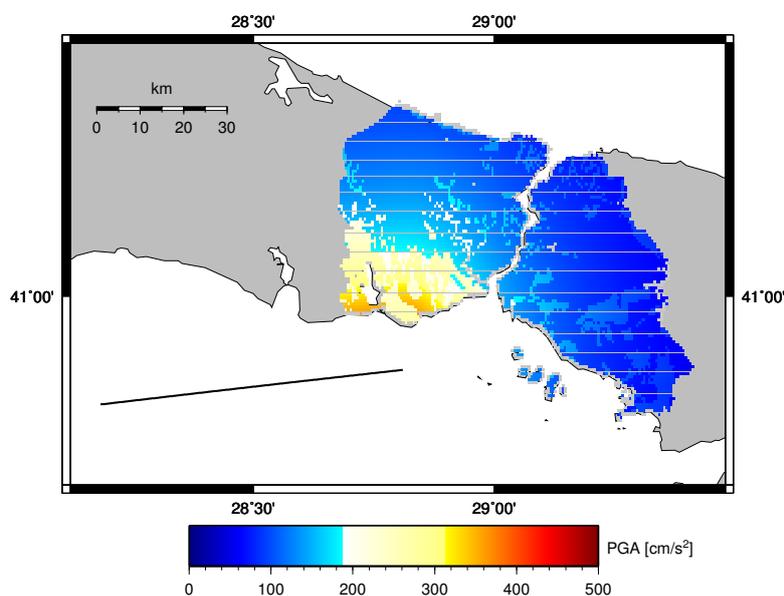
Figure 3.6 shows, for the purpose of comparison, the distribution of peak ground acceleration calculated with the attenuation relationship of Özbey et al. (2004) for northwestern Turkey. The ground motion at each grid cell is calculated using the *Joyner-Boore* distance, i.e., the closest distance from the grid cell to the surface projection of the ruptured fault as it is defined for the 1509 earthquake. *PGA* ranges here between 56.9 and 360.7 cm/s<sup>2</sup>. The minimum *PGA* value is therefore 59% larger than the one obtained from the *FINSIM* simulations. The maximum *PGA* value, in contrast, is 21% smaller than the one obtained from *FINSIM*.

The average peak ground acceleration over all grid elements is 20% larger than the one obtained from *FINSIM*. This is not so much astonishing since Figure 3.2 already indicated that the attenuation relationships predict larger peak ground accelerations, especially at distances up to 10 km and greater 70 or 80 km. The higher average value of *PGA* in Figure 3.6 is hence likely due to higher peak ground accelerations at large distances.

The fact that the *FINSIM* simulations (Figure 3.5) still yield the larger maximum *PGA* values is caused by directivity effects of the propagating rupture. They lead to ground motion distributions that look somehow “speckled”, while the attenuation relationship of Özbey et al. (2004) causes a ground motion pattern that looks “smooth”.



**Figure 3.5:** Distribution of *PGA* (left) and seismic intensity (right) for the simulated 1509 Istanbul earthquake. The ground motions are calculated for all three hypocentre locations.



**Figure 3.6:** Distribution of  $PGA$  calculated with the attenuation relationships of Özbey et al. (2004). Inputs are the *Joyner-Boore* distances from each grid cell to the ruptured fault as it is defined for the 1509 earthquake.

The next section will present building damage estimates calculated from the above described ground motion simulations. The predicted building damages are based on the building stock of today's city of Istanbul.

## 3.2 Building Damage Estimation

The building damage estimates presented in this section are established on the basis of the *KOERI Loss Estimation Methodology (KOERLoss)* developed at *KOERI*. The software performs building damage estimations from earthquakes and is also able to estimate the direct economic losses and casualties that are related to the building damages. The earthquake hazard in the method is calculated both deterministically and probabilistically (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

The building damage estimates by *KOERILoss* are performed in two ways: one method uses the seismic intensity as ground motion input, the other the spectral acceleration. The damages presented in this study are based on the spectral acceleration method.

The fundamental information is the classification of the building stock in Istanbul according to the buildings' vulnerabilities towards earthquakes. The building types depend on the construction type (reinforced concrete frame building, masonry building, reinforced concrete shear wall

building, or precast building), on the number of stories (low-rise (1-4 stories), mid-rise (5-8 stories), or high-rise (8 or more stories)), and on the construction date (before 1979 or after 1980). In Istanbul, low- and mid-rise buildings of reinforced concrete frame structure represent 75% of the total building stock (Erdik et al., 2003a).

The building damage estimations require the establishment of the so-called *spectral displacement demand* of a given structure, determined for a given earthquake. This demand for a certain building type is essentially estimated on the basis of the so-called *capacity diagram*, plotted as spectral acceleration versus spectral displacement. The spectral displacement demand is, on the other hand, obtained by amplifying the spectral displacement by the *spectral displacement amplification factor*. This factor is expressed by the *strength reduction factor*, defined by the developed capacity diagram together with the spectral acceleration at the natural period of the building.

So-called *fragility curves* (cumulative probability of damage to reach or exceed the threshold for a given damage class versus the spectral displacement demand) then allow for calculating the damage probabilities for certain damage classes. These damage classes are defined as *slight*, *moderate*, *extensive*, and *complete damage*. Complete damage refers to the total destruction of the building. The spectral acceleration-based damage estimation procedure is described in more detail in Appendix C. For further information about the loss estimation method please see Erdik et al. (2003a) or the report of American Red Cross - Bogazici University (2002).

The spectral acceleration-based damage estimation method has partially been incorporated by *DELPHI IMM GmbH*<sup>1</sup> in the context of the research project *EDIM*<sup>2</sup>. The building damage estimates discussed in this section are provided by *DELPHI IMM GmbH* by using the spectral acceleration values derived from the simulated ground motions presented in the previous section.

Table 3.2 lists the respective damages for the three 1509 earthquake scenarios as well as for the simulation based on the attenuation relationships of Özbey et al. (2004). The total number of buildings in Istanbul is 737,653. The building damage estimates based on the *FINSIM* simulations result into considerably higher building damages as the simulation based on the attenuation relationships. In the latter case, 17% of all buildings in Istanbul would experience any damage, while the three *FINSIM* scenarios show an average of 27% of damaged buildings which is a bit more than the 1.5-fold of damages.

The distributions of damage within the three *FINSIM* simulations are similar to the ground motion distributions: The versions 1 and 3 yield similar building damage estimates, while version 2 results into less expected damages (almost 24% less). In contrast to the ground motions, however, scenario 3 (hypocentre located at the western end of the fault) shows slightly higher building damage estimates than scenario 1.

<sup>1</sup>DELPHI InformationsMusterManagement GmbH, Potsdam, Germany (<http://www.delphi-imm.de>)

<sup>2</sup>EDIM - An Earthquake Disaster Information system for the Marmara region, Turkey (<http://www.cedim.de/EDIM.php>), funded by the BMBF GEOTECHNOLOGIEN Programme (<http://www.geotechnologien.de>)

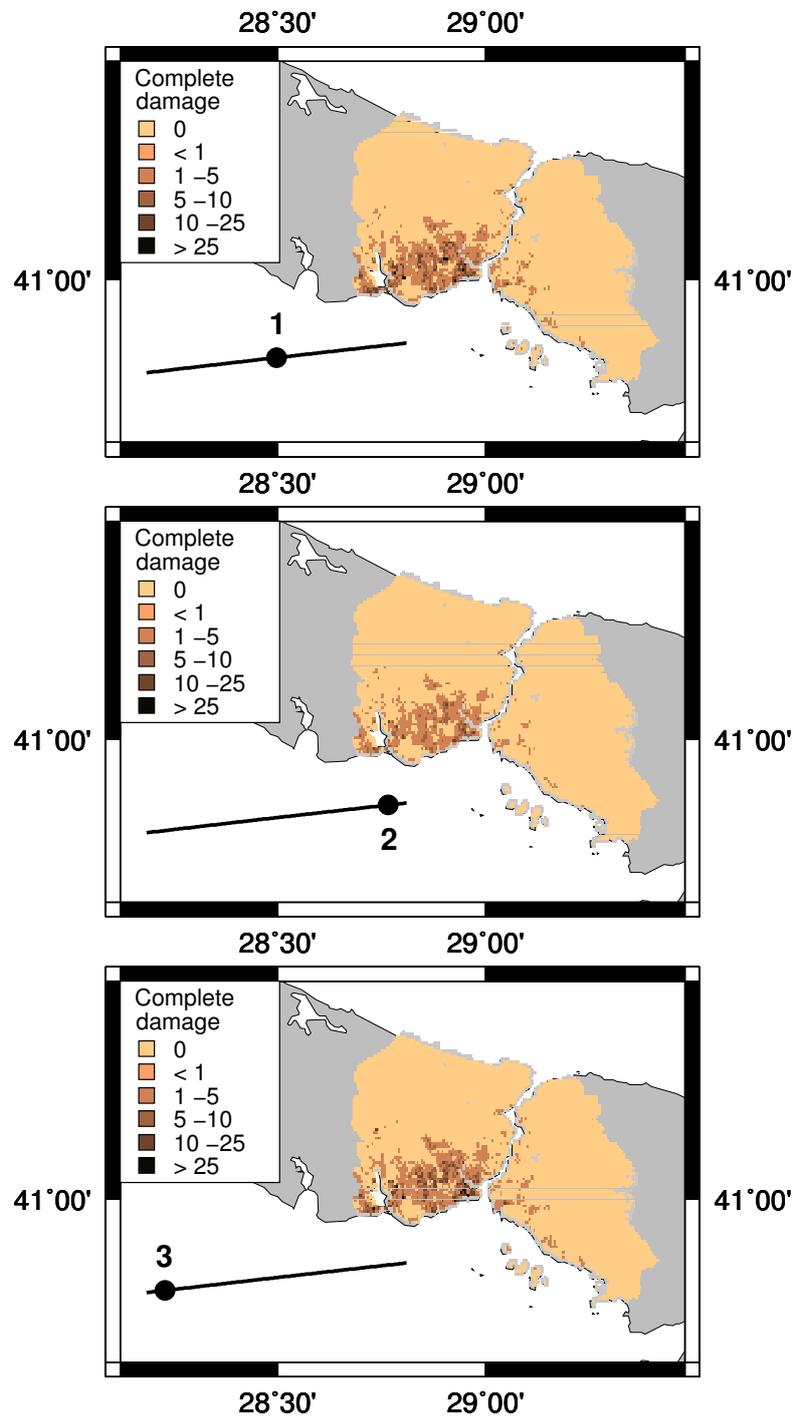
<i>Damage</i>	<i>Scenario 1</i>		<i>Scenario 2</i>		<i>Scenario 3</i>		<i>Özbey et al. (2004)</i>	
	<i>Absolute</i>	<i>%</i>	<i>Absolute</i>	<i>%</i>	<i>Absolute</i>	<i>%</i>	<i>Absolute</i>	<i>%</i>
<i>None</i>	529078.8	71.7	575979.5	78.1	520697.2	70.6	609264.5	82.6
<i>Slight</i>	140113.3	19.0	114874.5	15.6	144400.7	19.6	97418.6	13.2
<i>Moderate</i>	54294.0	7.4	38637.7	5.2	57299.6	7.8	26468.6	3.6
<i>Extensive</i>	10393.9	1.4	6196.1	0.8	11161.3	1.5	3562.4	0.5
<i>Complete</i>	3773.0	0.5	1965.2	0.3	4094.2	0.6	938.9	0.1
<i>Total</i>	208574.2	28.3	161673.5	21.9	216955.8	29.5	128388.6	17.4

**Table 3.2:** Absolute numbers and percentage of estimated building damages from the application of the historic 1509 Istanbul earthquake to today's Istanbul building stock. Scenarios 1 to 3 refer to the three hypocentre locations defined in Table 3.1. The last column gives the damages resulting from the ground motions predicted by the attenuation relationships from Özbey et al. (2004). The damage estimates base on the method of Erdik et al. (2003a) and are provided by *DELPHI IMM GmbH*<sup>1</sup> (Wenzel et al., 2009).

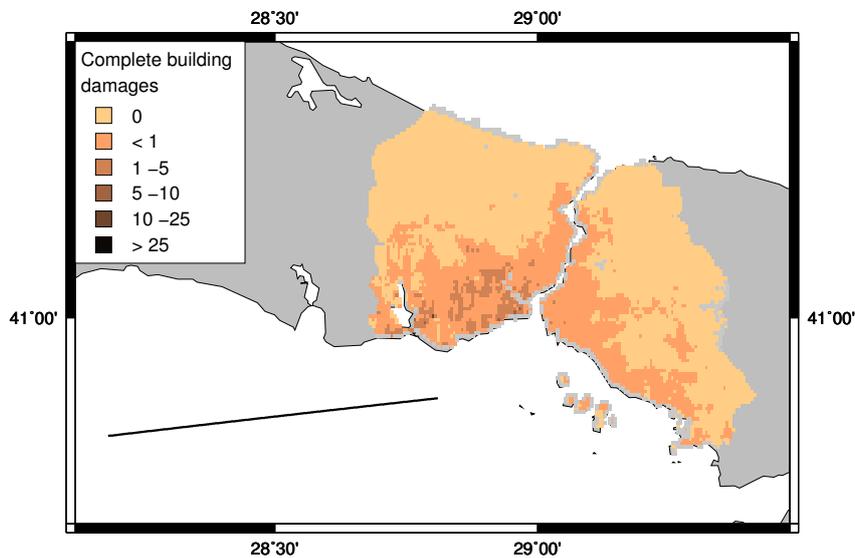
The damage class of completely damaged buildings is the most important in regard to casualties during and after an earthquake. This is, of course, only true when secondary effects like fires, explosions, the emission of contaminative materials, or the occurrence of landslides, for example, are neglected which can also result from less severe ground motions or damages to infrastructure and lifelines. A large number of people is likely to be trapped beneath the debris of destructed houses which is of great importance for rescue operations.

Figure 3.7 shows the distributions of the buildings that are estimated to be completely destructed for the three *FINSIM* simulations of the 1509 earthquake. Most of the damaged buildings are located on the southern part of the European side of Istanbul, with the highest number concentrating in the area around the old town. Again, the damage distributions for scenarios 1 and 3 look almost identical, showing a minimal number of complete damages more for scenario 3, especially along the Bosphorus and on the southern coast of the Asian part. The damage estimates from scenario 2 mainly differ by even less damaged buildings on the Asian part of Istanbul and by a generally decreased number of buildings per grid element on the European side.

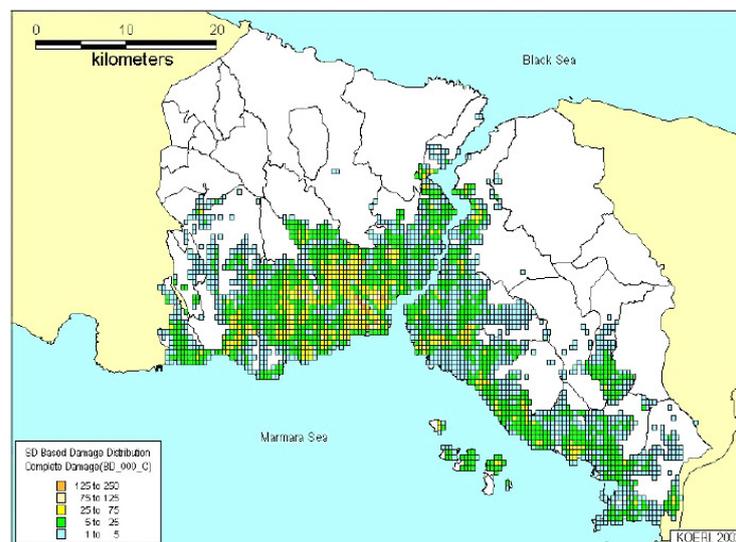
Figure 3.8 presents the according distribution of completely destructed buildings on the basis of the ground motions predicted from the attenuation relationships of Özbey et al. (2004). The distribution reveals a much greater area affected by the damages than obtained from the *FINSIM* simulations, both on the European and Asian part of Istanbul. This can result from the higher ground motions at larger distances. However, as already noted in Table 3.2, the overall number of damages is considerably less; only 0.1% of the buildings experience complete damage. The majority of the affected grid cells only exhibit less than 1 destructed building which is, of course, a purely statistical number. In reality, the damaged buildings (1 or more) would probably occur sporadically at single grid elements rather than being distributed over a coherent, large area. This might then be much more similar to the damage distributions displayed in Figure 3.7.



**Figure 3.7:** Distribution of absolute numbers of expected completely damaged buildings for the simulated 1509 Istanbul earthquake. The damages are calculated for all three hypocentre locations.



**Figure 3.8:** Distribution of absolute numbers of expected completely damaged buildings for the simulated 1509 Istanbul earthquake. The damages are based on the ground motions predicted from the attenuation relationships of Özbey et al. (2004).



**Figure 3.9:** Distribution of absolute numbers of expected completely damaged buildings for the  $M_w = 7.5$  “worst case” scenario earthquake (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

A study by Erdik et al. presents a much more detailed analysis (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a). The authors assume a “worst case” scenario earthquake of moment magnitude 7.5 with a total strike-slip rupture length of 120 km. In contrast to the assumed fault rupture of the historic 1509 earthquake, the rupture of the “worst case” scenario also includes the Istanbul Bend and its eastwards following segment. The average strike of this rupture is therefore roughly parallel to Istanbul’s Sea of Marmara coastline and has an average distance of 20 km to Istanbul. The ground motion is simulated for the same grid and site conditions as for the 1509 scenario in this study, however exclusively determined by taking the average values from several attenuation relationships.

The spectral displacement-based approach for calculating the expected damage yields a total of 34,828 buildings that are damaged beyond repair (complete damage) and 67,395 buildings with extensive damage. Another 195,097 buildings are moderately damaged (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a). The “worst case” scenario thus produces the 7-fold number of complete damages than the average values obtained from the three *FINSIM* simulations of the 1509 earthquake and the 37-fold number than predicted by the laws from Özbey et al. (2004). For the extensive damage class, Erdik et al. (2003a) obtain the 7-fold and the 19-fold damages, respectively, and the 4-fold and the 7-fold damage, respectively, for the damage class “moderate”. The spatial distribution of the completely damaged buildings from Erdik et al. (2003a) is shown in Figure 3.9. Again, the largest numbers of damages occur on the European side around the region of Istanbul’s old town. The overall shape of the distribution is similar to the one in Figure 3.8, although the numbers of destructed buildings per grid cell is considerably higher.

The reason for the much less structural damages estimated for the 1509 earthquake is on the one hand, of course, its smaller magnitude. On the other hand, the directivity effects of the 1509 fault rupture apparently cause the strongest ground motions to concentrate mainly on the European side of Istanbul, whereas the ruptured fault segment assumed by Erdik et al. (2003a) affects both sides of the Bosphorus equally due to its relatively parallel strike to the coastline of the megacity.

As a matter of principle, scenario earthquakes represent only one possible realization of fault rupture and generation of one set of the ground motion amplitudes. We have seen in this chapter that the characteristics of simulated ground motions highly depend on the simulation or prediction technique and on the chosen earthquake source model, leading to a significant degree of variability in seismic ground shaking and damage patterns.



# Chapter 4

## Earthquake Early Warning

As described in Chapter 1, earthquake early warning (EEW) systems provide warnings prior to damaging ground motions (e.g. Allen et al., 2009b) using the physical basis that the destructive S- and surface waves of an occurring earthquake travel slower than the P-waves and much slower than electromagnetic signals used to transmit information and warnings. Extensive research has been undertaken during the last years to develop and improve suitable system configurations and early warning algorithms. I will summarize the recent state-of-the-art.

### 4.1 State-of-the-art of Earthquake Early Warning

Depending on the tectonic settings of a given region and the user demand, different types of EEW systems can be suitable. If the earthquakes originate from a well defined seismic zone, the implementation of a *front detection* system is highly appropriate. In front-detection systems, the seismic sensors are installed as close as possible to the seismic zone so that strong ground motions can be detected as soon as possible to maximize the warning times. The early warning system in Istanbul, as described in Chapter 1, is an example of front-detection. The sensors are installed along the coastline of the Sea of Marmara, as close as possible to the Main Marmara Fault, to send a warning ahead of the seismic waves (Erdik et al., 2003b; Alcik et al., 2009).

Another typical front-detection system is the Seismic Alert System (SAS) for Mexico City. It has been established after the devastating Michoacán earthquake in 1985 which caused serious damage to the city. The system is designed to detect earthquakes that are generated in the Guerrero Gap, a subduction zone about 300 to 450 km away from Mexico City. Despite these large distances the earthquakes can be highly damaging because of large amplification of seismic waves due to soft, water-saturated clay soils on which large parts of the mega-city are built (Suárez et al., 2009). The seismic waves radiated from large earthquakes in this zone take approximately 60 to 80 s to reach Mexico City resulting in large warning times of at least 40 s (Espinosa-Aranda et al., 1995, 2009; Suárez et al., 2009).

The SAS was the first system that emitted warnings to the public. Until end of April 2009, the

SAS issued 66 alerts of which 53 were issued as preventive alert, given when the estimated magnitude for the earthquake is between 5 and 6. The remaining 13 alerts were emitted as public alerts, given when the estimated magnitude is 6 or larger. The preventive alerts are only sent to registered users of the SAS. The public alert is sent to all users, transmitted as a sound signal by television and radio stations (Suárez et al., 2009).

Another example of front-detection is the EEW system for the Romanian capital Bucharest. The source region of potentially damaging earthquakes is defined to an about 40 km x 80 km x 110 km large volume in the Vrancea zone. Three strong motion sensors have been installed in the epicentral area. For estimating the peak ground acceleration in Bucharest, amplitude scaling relationships using filtered *PGA* are applied. Average hypocentral distances of 160 km provide warning times of about 25 s (Böse et al., 2007; Wenzel et al., 1999).

The task of an EEW system is to provide reliable warnings within the shortest amount of time. However, high reliability requires more time, while a fast system has less information available, resulting in less accuracy. The two main approaches in EEW, *on-site* and *regional* early warning, reflect this conflict. On-site early warning uses the observed ground motion at a single site to predict the upcoming strong shaking at the same site. Regional early warning, on the contrary, uses a network of seismic sensors to locate the earthquake and to estimate its magnitude and the ground shaking level at other sites. On-site early warning is fast but less reliable. Regional early warning, in contrast, is more reliable but also requires more time since it must be waited until the seismic waves arrive at several stations. Regional early warning is thus not useful in regions where the distances to the user sites are only short (Kanamori, 2005).

The first real-time EEW system in operation was the *Urgent Earthquake Detection and Alarm System (UrEDAS)* in Japan, a typical on-site system that started operation in 1984. It is a P-wave detection and alarm system for the Japanese high-speed trains. It calculates parameters such as back azimuth, predominant frequency, and vertical to horizontal ratio from the first 3 s following the P-wave arrival to estimate the location and magnitude of an earthquake. The processing is done at each station so that time consuming data transmission to a central processing facility is not necessary. A second part, Compact UrEDAS, has later been installed to estimate the expected damage potential of an occurring earthquake already 1 s after the P-wave arrival (Nakamura and Saita, 2007).

The second EEW system in Japan uses both on-site and regional approaches. The system is operated by the Japan Meteorological Agency (JMA) and aims on providing the maximum expected seismic intensity and the earliest S-wave arrival time in each subprefectural area before the strong motion arrives. The system uses a total of 1000 seismic stations of approximately 20 km spacing throughout Japan to locate the hypocentre and estimate the magnitude, which are then used to predict the seismic intensity (Kamigaichi et al., 2009). The system became fully operational in 2007.

JMA uses several methods for locating the earthquake hypocentre, depending on the number of triggered stations. For one triggered station, the slope of the P-wave onset is converted into epicentral distance by using an inverse relationship between the slope and the distance found by Odaka et al. (2003). The azimuth is estimated by fitting an ellipsoid to the P-phase particle

motion. In addition, a network approach is used. For one or two triggered stations, a “territory” region to each station is defined. The geometrical centroid of the territory is used as the initial epicentre estimate. With an additional triggered station, the territory is divided into subsections and the epicentre estimate is updated. The hypocentre depth is fixed to 10 km. For three to five triggered stations, a number of possible hypocentre locations is defined on a grid, whereas the location which fits best to the observed P-wave arrival time differences is considered as the hypocentre location (Kamigaichi, 2004). To increase the accuracy, the procedure includes information on the not-yet-triggered stations surrounding the epicentre region, as introduced by Horiuchi et al. (2005).

The magnitude is estimated by using a scaling relationship between the magnitude and the maximum displacement amplitude in combination with the hypocentre location. The estimate is updated each second using the latest hypocentre location estimate, whereas two different scaling relationships are used, one for the P-waves and one for the S-waves (Kamigaichi, 2004). The prediction of seismic intensity is accomplished via the estimation of peak ground velocity using hypocentral distance, focal depth, magnitude, and amplification of the seismic wave amplitudes at a given site (Hoshiha et al., 2008).

JMA additionally included the concept of on-site early warning. Whenever the observed amplitude of ground acceleration exceeds a threshold of  $100 \text{ cm/s}^2$  an initial warning is issued (Kamigaichi et al., 2009).

Since the system became fully operational in 2007, 11 public alarms have been issued, whereas three alarms overestimated the seismic intensities. Another two earthquake alarms were missed (Kamigaichi et al., 2009). The JMA warnings are issued to two categories of users, online limited users and general users. The online limited users include, for example, railway companies, industries, and other organizations that obtain the EEW information online to initiate automatic actions. The general users receive the warnings via television, radio, mobile phones, or internet, for example, but also by sound signals (Hoshiha et al., 2008).

Many EEW methods use parameters derived exclusively from the first few seconds of P-wave data to estimate the magnitude of an earthquake or to predict the upcoming strong ground motion directly. One of these parameters is the *predominant period*  $\tau_p$ , calculated as

$$\tau_{p,i} = 2\pi \sqrt{\frac{\alpha X_{i-1} + x_i^2}{\alpha D_{i-1} + (dx/dt)_i^2}}, \quad (4.1)$$

where  $i$  indicates the time step.  $x_i$  is the observed ground velocity,  $X$  is the smoothed ground velocity squared,  $D$  is the smoothed velocity derivative squared, and  $\alpha$  is a smoothing constant (e.g. Allen and Kanamori, 2003). The in EEW used relation of  $\tau_p$  to magnitude is based on the general observation that small earthquakes radiate relatively high-frequency energy, while the seismic waves of large-magnitude earthquakes exhibit lower frequencies.

A second P-wave parameter relating the frequency content of the P-wave to magnitude is  $\tau_c$  (Kanamori, 2005).  $\tau_c$  is the effective period of the P-wave over a fixed time interval, usually set to 3 s. The third commonly used P-wave parameter, however not related to the signal’s frequency content, is the peak ground motion of the first few seconds of the P-wave. The peak displacement

$P_d$  was found to be the most robust one (Allen et al., 2009b).

The advantage of P-wave parameters extracted from the first 3 to 5 s of P-wave data is that estimates of magnitude or ground shaking are quickly available. However, the approach is controversial; is it indeed possible to estimate the size of an earthquake from the first few seconds of rupture? The so-called cascade model for fault rupture assumes that there is no difference between the beginnings of large and small earthquakes. The model considers the fault as being divided into single patches. A large earthquake occurs when a small earthquake triggers a cascade of increasingly larger slip events. The model thus implicates that earthquakes are not deterministic (e.g. Ellsworth and Beroza, 1995).

Nonetheless, different studies showed that the predominant period of the first few seconds lead to successful magnitude estimates. Olson and Allen (2005), for example, state that when the maximum predominant period from a time series over a certain time window,  $\tau_p^{max}$ , is plotted against the magnitude on a log-linear scale, a linear scaling relation emerges. Their analysis concentrated on a large set of earthquakes from Japan, Taiwan, California, and Alaska. The observed scatter in their log-linear plot is related, according to the authors, to measurement errors, station effects, and path effects. Olson and Allen (2005) propose that the final magnitude of an earthquake is partially controlled by the initiation process within the first few seconds of the rupture, and partially by the physical state of the surrounding fault plane.

Rydelek and Horiuchi (2006), on the contrary, repeated the analysis done by Olson and Allen (2005) by using a set of earthquakes from Japan. They find no evidence that the final earthquake magnitude can be estimated before the rupture has completed. Wolfe (2006) demonstrated on the basis of numerical simulations that the parameter  $\tau_p$  is a non-linear function of spectral amplitude and period that gives greater weight to higher amplitudes and higher frequencies. The author states that averaging the predominant period over several stations does therefore not provide an estimate of the average frequency content of an earthquake. Wolfe (2006) also suggests that the parameter  $\tau_p^{max}$  includes errors arising from the influence of both the amplitude spectrum and the phase spectrum.

Yamada and Ide (2008) found that the values of  $\tau_p^{max}$  have upper and lower limits. For larger earthquakes,  $\tau_p^{max}$  has an upper limit that depends on the length of the investigated time window. For smaller earthquakes, the authors found a lower limit that is proportional to the sampling interval of the data. For intermediate-sized earthquakes, Yamada and Ide (2008) found that the values of  $\tau_p^{max}$  are close to their typical source durations and can exhibit a large variety due to the complexity of the source process. Yamada and Ide (2008) conclude that the linear relationship used by Olson and Allen (2005) may be useful as long as it is kept in mind that this relationship not necessarily reflects physical properties of the initial earthquake rupture. They state that the dependency of  $\tau_p^{max}$  on the final magnitude is not sufficient to conclude whether earthquake rupture is deterministic or not.

The above listed P-wave parameters are used by various EEW methods. One of them is the Earthquake Alarm Systems (ElarmS) methodology, originally developed for southern California (e.g. Allen and Kanamori, 2003; Allen et al., 2009a; Wurman et al., 2007). ElarmS uses the maximum predominant period  $\tau_p^{max}$  in combination with the peak displacement  $P_d$  of the first 1 to 4 s

of the P-wave to estimate the earthquake magnitude. It also provides alert maps (Wurman et al., 2007) which are updated every second by predicting the distribution of ground shaking using attenuation relationships and site corrections (Allen et al., 2009b). ElarmS has been tested using data from California, Taiwan, Japan, and the Pacific Northwest of the United States (e.g. Brown et al., 2009; Lockman and Allen, 2007; Olson and Allen, 2005). ElarmS is one of three algorithms currently tested in real-time as part of the *California Integrated Seismic Network (CISN)* early warning project.

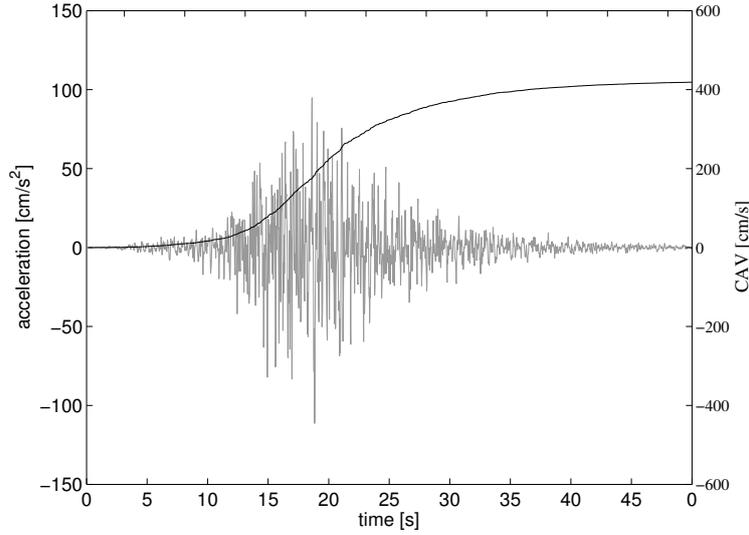
The EEW system of Taiwan combines P-wave parameters with a Virtual Subnetwork (VSN) method (Hsiao et al., 2009). The latter has been in operation since 2001. It is a regional early warning method using a nationwide network of strong motion stations installed by the Taiwan Central Weather Bureau. The VSN method estimates the earthquake magnitude and can provide warning times of more than 20 s to cities at distances greater than 145 km from the source (Wu and Teng, 2002). Due to the geometry of network and source region, the VSN system has a blind zone with a radius of 70 km around the epicentre of an earthquake in which warnings cannot be issued before the onset of strong ground motion. This motivated the establishment of a second EEW component that uses P-wave information. It operates parallel to the VSN method. When the average  $P_d$  value from the five nearest station exceeds 0.1 cm,  $\tau_p^{max}$  and  $\tau_c$  are calculated from the first 3 s of data. The magnitude is estimated using empirical relationships between  $M$  and  $\tau_p^{max}$  and  $\tau_c$ . When a magnitude estimate of both the VSN and the P-wave parameters is larger than 6.0, the shake map will be calculated. Warnings from the system are currently not available to the public (Hsiao et al., 2009).

The second EEW algorithm currently tested in real-time in California is an on-site approach based on  $\tau_c$  and high-pass filtered  $P_d$ . The algorithm is based on a single sensor observation and calculates the P-wave parameters from the initial 3 s to estimate moment magnitude and  $PGV$  (Böse et al., 2009).

The third implemented algorithm at CISN is the Virtual Seismologist (VS) method, a Bayesian approach to earthquake early warning. It uses peak acceleration, velocity, and displacement at every second to estimate the most probable location and magnitude of an occurring earthquake (Cua, 2005; Cua and Heaton, 2007; Cua et al., 2009). Baye's theorem allows the VS to use prior information such as health status of the seismic network, Gutenberg-Richter relationship or known fault locations to assist in resolving trade-offs in magnitude and location that cannot be resolved in the initial stages of earthquake rupture due to scarce observations (Cua and Heaton, 2007). More details about the VS method can be found in Appendix D of this thesis.

Continuous research on EEW is also going on in Italy (e.g. Zollo et al., 2009), where a probabilistic evolutionary approach for earthquake hypocentre determination is currently tested in real-time (Satriano et al., 2008; Weber et al., 2007). The tests are performed using the Irpinia seismic network in southern Italy. Methods for the estimation of magnitude include the usage of the squared velocity integral over the first few seconds of P- and S-waves (Festa et al., 2008) and the peak amplitudes (Lancieri and Zollo, 2008; Zollo et al., 2006).

A novel approach for EEW is based on the development of low-cost sensors. The Self-organizing Seismic Early Warning Information Network (SOSEWIN) consists of single low-cost sensing units that can form dense decentralized, self-organizing wireless mesh networks. Each sensing unit undertakes its own data processing, analysis, and communication of data and early warning



**Figure 4.1:** Cumulative absolute velocity (black curve) of a simulated acceleration record (grey).

messages. The data processing for the EEW application currently includes the determination of peak acceleration, velocity, and displacement, cumulative absolute velocity, Arias intensity, and predominant period. SOSEWIN is currently running in a testbed in Istanbul (Fleming et al., 2009).

The algorithm *PreSEIS* (Pre-SEISmic shaking) discussed in this thesis is a neural network-based approach to earthquake early warning developed by Böse (2006). The method estimates hypocentre location, magnitude, and the final expansion of the evolving seismic rupture by using the ground motion information from a network of seismic stations. The hypocentre is estimated using the P-wave arrival time differences from the different stations by including the information of not-yet-triggered stations.

*PreSEIS* takes advantage of both regional and on-site early warning. As soon as the P-wave reaches the first station, the method starts estimating the seismic source parameters. The estimates are updated at regular time steps with ongoing time when more ground motion information becomes available. The magnitude and the evolving rupture are estimated using the *cumulative absolute velocity* (CAV) together with the hypocentre estimates at each time step (Böse, 2006; Böse et al., 2008).

CAV is defined as the integral of absolute acceleration over a certain time window, usually set to the duration of the record:

$$CAV \equiv \int_0^t |a(t)| dt \quad (4.2)$$

Figure 4.1 illustrates the cumulative absolute velocity of one of the simulated acceleration time series used in this study.

Benjamin and Associates (1988) found that the cumulative absolute velocity is one out of ten

investigated ground motion parameters that predicts damage with highest reliability. *CAV* has the advantage that it reflects both the amplitude and the duration of motion.

PreSEIS makes use of artificial neural networks which map the observations from the seismic stations onto the source parameters. It has been developed on the example of Istanbul, using the simulated ground motion data described in Chapter 2. Already 0.5 s after the P-waves arrive at the first seismic station, the magnitude estimates show errors of less than  $\pm 0.8$  magnitude units. This can be decreased to  $\pm 0.3$  units after 15.0 s. The mean location errors are 10 km after 0.5 s and 5 km after 15.0 s (Böse, 2006; Böse et al., 2008).

At present, PreSEIS is a new approach under development, rather than a final EEW algorithm. The next section will describe the methodology in more detail.

## 4.2 The PreSEIS Methodology

The following description of the *PreSEIS* methodology is based on Böse (2006) and Böse et al. (2008):

As mentioned before, *PreSEIS* makes use of artificial neural networks (ANNs) to map the ground motion observations from each seismic station onto the earthquake source parameters. Used are so-called *Two-Layer Feed Forward* (TLFF) neural networks. Each neural network consist of simple processing units, called *neurons*. A TLFF network is built of three layers of neurons, an *input layer*, a *hidden layer*, and an *output layer* (Figure 4.2a). Each neuron is connected to all neurons in the subsequent layer.

In the case of *PreSEIS*, the input layer is represented by the input parameters that are derived from the ground motion time series of each seismic station. The number of neurons in the input layer thus depends on the total number  $I$  of stations  $i$  in the seismic network ( $i = 1, \dots, I$ ). Since this layer does not perform any processing of data it is not numbered.

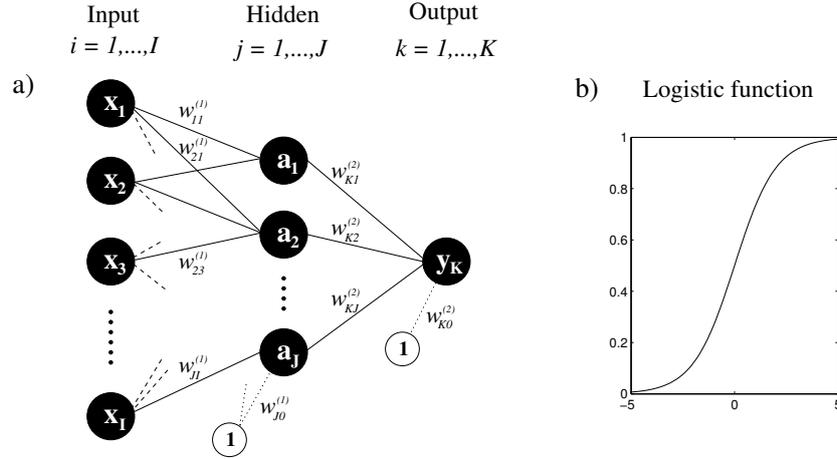
The middle layer that does not represent the final output of the network is called the hidden layer with  $j = 1, \dots, J$  neurons. The number of neurons in the hidden layer controls the performance of the TLFF network and must be determined empirically. This will be explained later.

The output layer represents the final outputs, e.g. the earthquake source parameters that shall be estimated. The number of neurons is thus equivalent to the number of estimated source parameters.

A *feed-forward* neural network strictly operates in direction from the input to the output layer without any recursion.

The importance of each connection between the different neurons is controlled by the *weights*  $w_{ji}$ . The weights of the connections from the input to the hidden layer are marked with the superscript (1), while (2) represents the connections between the hidden and the output layer. The sum of the weights  $w_{ji}$  and inputs  $x_i$  is calculated for each neuron, defined as the input  $a_j$  for this neuron:

$$a_j = \sum_{i=1}^I w_{ji}^{(1)} x_i + w_{j0}^{(1)} = \sum_{i=0}^I w_{ji}^{(1)} x_i, \quad (4.3)$$



**Figure 4.2:** (a) Sketch of a Two-Layer Feed-Forward neural network. The neurons are arranged in the input, the hidden, and the output layer. The connections between the neurons of the different layers are controlled by the weights  $w_{ji}$ . (b) Logistic activation function defined in equation 4.5.

whereas  $x_0 = 1$ . The parameter  $w_{j0}^{(1)}$  is called *threshold* and determines whether the neuron will take the activated value  $z_j$ . This activation is realized by applying an *activation function*  $g(\cdot)$ :

$$z_j = g(a_j) = g\left(\sum_{i=0}^I w_{ji}^{(1)} x_i\right) \quad (4.4)$$

A common choice for an activation function is the *logistic* (or *sigmoid*) function, defined as

$$g(x) \equiv \frac{1}{1 + \exp(-x)}. \quad (4.5)$$

The logistic function is displayed in Figure 4.2b.

The proper determination of the weights  $w_{ji}$  in the training process of the TLFF network (and thus the “learning” of the network) can be achieved by the application of various learning algorithms. The training is accomplished by feeding a dataset with known input and output values into the TLFF network. The squared errors between the correct output and the network output obtained for a certain combination of weights are minimized using the so-called Levenberg optimization method (Levenberg, 1944). *PreSEIS* allows for exchanging the learning technique with a different algorithm.

The complete training is performed by several iterations. The obtained errors between the correct and the network outputs are fed back through the network and the weights of the following iteration are adapted according to this information. By this way, the errors of each iteration can be reduced. In *PreSEIS*, the initial weights of the first iteration are set randomly.

A fundamental point is the determination of the numbers of necessary iterations. If the training is stopped too early the network will not have achieved its optimum accuracy. On the contrary, if

the training continues for too many iterations the network will become overfitted to the training data and will lose its generalization capability. To determine the most optimum stopping point, an independent *validation* subset of the training dataset is determined. Böse (2006) realized this by taking off 10% of the training data. As soon as the error of this independent validation subset increases, the training of the TLFF network is terminated.

A similar procedure is performed to determine the optimum number of neurons in the hidden layer: After training, the neural network is fed with a *test* set of data which follow the same statistical patterns as the training data but which are “unknown” to the network, i.e. the data have not been included in the training dataset. The optimum number of neurons in the hidden layer will be the one where the network performs well for *both* the training and the test data.

For the architecture of *PreSEIS* and the given number of available ground motion simulations Böse (2006) determined  $J = 6$  as the best number of neurons in the hidden layer.

*PreSEIS* performs three different source parameter estimations from the seismic ground motion time series of an earthquake. The first one is the estimation of the earthquake hypocentre location, expressed by the geographical latitude  $\Phi$  and longitude  $\Theta$  and by the source depth  $Z$ . The second is the estimation of the moment magnitude  $M$  and the third is the prediction of the expansion of the earthquake rupture. The latter is represented by the start and end coordinates of the rupture,  $\Phi_{\text{start}}$ ,  $\Theta_{\text{start}}$ ,  $\Phi_{\text{end}}$ , and  $\Theta_{\text{end}}$ .

The source parameter estimates start as soon as the P-wave arrives at the first station in the seismic network. With proceeding time the ground motion time series become longer and are available from an increasing number of stations. To account for the time-dependent rupture process *PreSEIS* updates the source parameter estimates at regular time intervals. These time intervals are set to 0.5 s, but can be changed to different values.

The current stage of *PreSEIS* requires that all seismic stations are in proper operational mode at all times and do not fail.

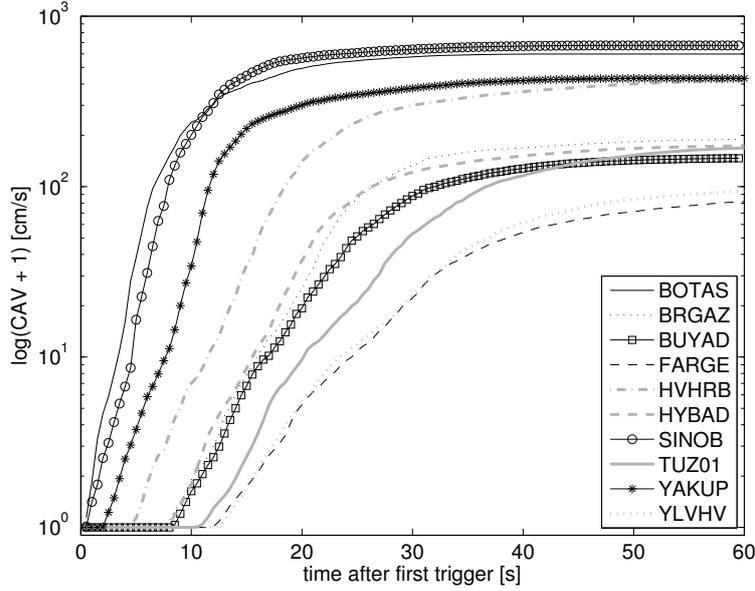
*PreSEIS* consists of three TLFF networks at each time step: one for estimating the hypocentre location, one for estimating the magnitude, and one for predicting the evolving rupture expansion. The input parameters to the TLFF networks responsible for estimating the hypocentre location are the arrival time differences

$$\Delta t_i = t_i - t_{i_0} \quad (4.6)$$

between the P-wave arrival  $t_i$  at station  $i$  and the arrival at the first station,  $t_{i_0}$ . The initial arrival time  $t_{i_0}$  is assigned to  $t_{i_0} = 1$ .

At time  $t^n$ , where  $n$  is the index of the current time step, the P-wave will only have arrived at a subset  $i$  of stations. The information on the not-yet-triggered stations contributes to the confinement of the possible range of solutions since certain earthquake locations can be ruled out. The time difference between the P-wave arrival at station  $i$  and the initial P-wave arrival at station  $i_0$  at time  $t^n$  is thus given as

$$\Delta \tau_i^n \equiv \begin{cases} t^n - \tau_{i_0}; & \tau_i > t^n \\ \tau_i - \tau_{i_0}; & \tau_i \leq t^n \end{cases} \quad (4.7)$$



**Figure 4.3:** Logarithm of  $CAV + 1$  calculated from each of the ten EEW stations (Table 2.3) for a simulated example earthquake. The time axis starts with the time at which the P-wave arrives at the first station.

where  $n = 1, \dots, N$  and  $\tau_{i_0} \leq \tau_i$  for  $i = 1, \dots, I$ . The upper relation gives the minimum time difference for a not-yet-triggered station at time  $t^n$ .

The second parameter used for the magnitude and rupture estimates relies on the cumulative absolute velocity  $CAV$  (Equation 4.2). Using the simulated ground motion time series, Böse (2006) calculated  $CAV$  from the 0.05-12.0 Hz band-pass filtered, combined-horizontal acceleration time series:

$$\log(CAV_i^n + 1) \equiv \begin{cases} 0; & \tau_i > t^n \\ \log\left(\sum_{t=\tau_i}^{t^n} |a_i(t)| dt + 1\right); & \tau_i \leq t^n \end{cases} \quad (4.8)$$

where  $|a_i(t)|$  is the absolute amplitude of ground acceleration given in  $\text{cm/s}^2$  observed at sensor  $i$  at time  $t$ . The quantity  $dt$  is the sampling interval of the ground acceleration time series. The logarithm accounts for the large range of  $CAV$  values from different stations; since the logarithm is not defined at zero, a constant of 1 is added to each  $CAV$  value. An example of the  $\log(CAV + 1)$  inputs is given in Figure 4.3. The parameter is determined at each time step (here 0.5 s) at each seismic station.

With Equations 4.3 and 4.4, the TLFF networks for the *PreSEIS* source parameter estimates at each time step  $n$  can now be defined. The first network used for estimating the hypocentre location is given as

$$\hat{\Phi}^n = \sum_{j=0}^J w_{1j}^{\text{loc}(2),n} g\left(\sum_{i=0}^I w_{ji}^{\text{loc}(1),n} \Delta\tau_i^n\right), \quad (4.9)$$

$$\hat{\Theta}^n = \sum_{j=0}^J w_{2j}^{\text{loc}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{loc}(1),n} \Delta\tau_i^n \right), \quad (4.10)$$

$$\hat{Z}^n = \sum_{j=0}^J w_{3j}^{\text{loc}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{loc}(1),n} \Delta\tau_i^n \right). \quad (4.11)$$

The network used for predicting the magnitude is quantified by

$$\begin{aligned} \hat{M}^n = & \sum_{j=0}^J w_{1j}^{\text{mag}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{mag}(1),n} \log(\text{CAV}_i^n + 1) \right. \\ & \left. + w_{j(I+1)}^{\text{mag}(1),n} \hat{\Phi}^n + w_{j(I+2)}^{\text{mag}(1),n} \hat{\Theta}^n + w_{j(I+3)}^{\text{mag}(1),n} \hat{Z}^n \right). \end{aligned} \quad (4.12)$$

Along with the derived *CAV* values, the location estimates from the same time step are used as additional input for the magnitude estimation. Figure 4.4 illustrates the TLFF networks used to estimate hypocentre location and magnitude at time step  $n$ . Both networks have 6 neurons in the hidden layer.

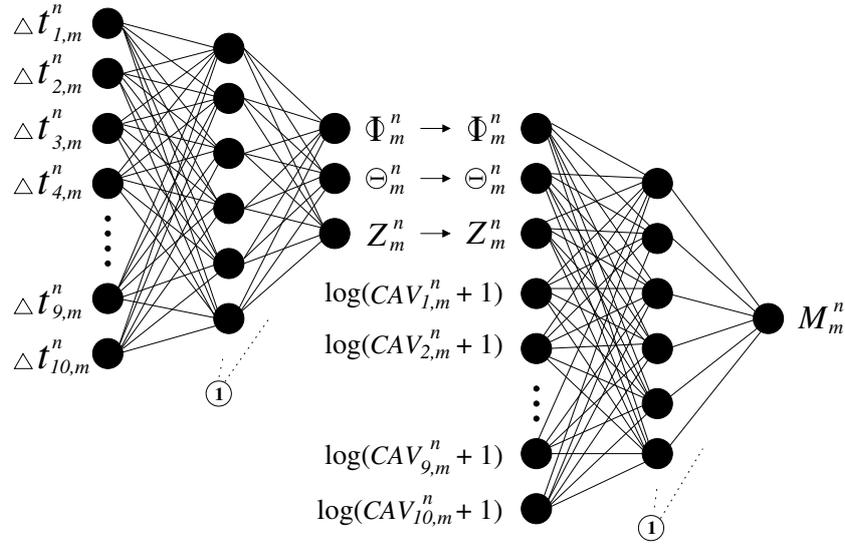
Finally, the third network for estimating the expansion of the evolving seismic rupture can be described. It incorporates both the hypocentre estimates and the magnitude estimate of the same time step:

$$\begin{aligned} \hat{\Phi}_{\text{start}}^n = & \sum_{j=0}^J w_{1j}^{\text{rup}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{rup}(1),n} \log(\text{CAV}_i^n + 1) \right. \\ & \left. + w_{j(I+1)}^{\text{rup}(1),n} \hat{\Phi}^n + w_{j(I+2)}^{\text{rup}(1),n} \hat{\Theta}^n + w_{j(I+3)}^{\text{rup}(1),n} \hat{Z}^n + w_{j(I+4)}^{\text{rup}(1),n} \hat{M}^n \right), \end{aligned} \quad (4.13)$$

$$\begin{aligned} \hat{\Theta}_{\text{start}}^n = & \sum_{j=0}^J w_{2j}^{\text{rup}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{rup}(1),n} \log(\text{CAV}_i^n + 1) \right. \\ & \left. + w_{j(I+1)}^{\text{rup}(1),n} \hat{\Phi}^n + w_{j(I+2)}^{\text{rup}(1),n} \hat{\Theta}^n + w_{j(I+3)}^{\text{rup}(1),n} \hat{Z}^n + w_{j(I+4)}^{\text{rup}(1),n} \hat{M}^n \right), \end{aligned} \quad (4.14)$$

$$\begin{aligned} \hat{\Phi}_{\text{end}}^n = & \sum_{j=0}^J w_{3j}^{\text{rup}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{rup}(1),n} \log(\text{CAV}_i^n + 1) \right. \\ & \left. + w_{j(I+1)}^{\text{rup}(1),n} \hat{\Phi}^n + w_{j(I+2)}^{\text{rup}(1),n} \hat{\Theta}^n + w_{j(I+3)}^{\text{rup}(1),n} \hat{Z}^n + w_{j(I+4)}^{\text{rup}(1),n} \hat{M}^n \right), \end{aligned} \quad (4.15)$$

$$\begin{aligned} \hat{\Theta}_{\text{end}}^n = & \sum_{j=0}^J w_{4j}^{\text{rup}(2),n} g \left( \sum_{i=0}^I w_{ji}^{\text{rup}(1),n} \log(\text{CAV}_i^n + 1) \right. \\ & \left. + w_{j(I+1)}^{\text{rup}(1),n} \hat{\Phi}^n + w_{j(I+2)}^{\text{rup}(1),n} \hat{\Theta}^n + w_{j(I+3)}^{\text{rup}(1),n} \hat{Z}^n + w_{j(I+4)}^{\text{rup}(1),n} \hat{M}^n \right). \end{aligned} \quad (4.16)$$



**Figure 4.4:** Sketch of the TLFF neural networks used for estimating the hypocentre location (latitude  $\Phi$ , longitude  $\Theta$ , depth  $Z$ ) and the magnitude ( $M$ ) at time step  $n$  for earthquake  $m$ . The input parameters are derived from stations 1 to 10. The estimation of location and magnitude are performed almost simultaneously.

The outputs of each TLFF network are linear combinations of the input parameters and the weights  $w_{ji}$ . As mentioned above, the weights are determined during the training of the neural networks with a training dataset where the inputs and the outputs are known.

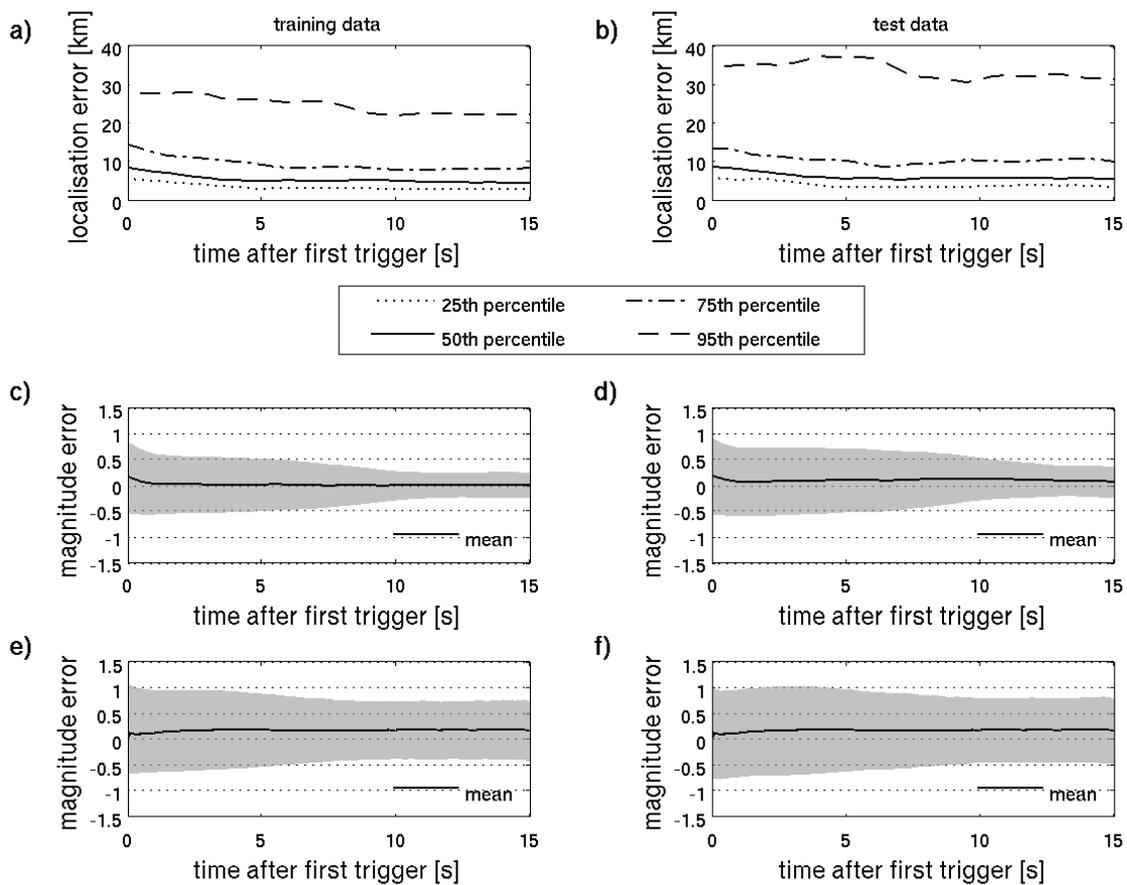
When building neural networks it is important to adjust the size of the networks to the number of training examples. In order to set up a robust ANN model, each free parameter in the network requires at least 10 training examples (Bishop, 1995). The number of weights  $W$  in a TLFF network is given by

$$W = (I + 1) \cdot J + (J + 1) \cdot K \quad (4.17)$$

where  $I$  is the number of input parameters,  $J$  is the number of neurons in the hidden layer, and  $K$  is the number of network outputs. In the case of *PreSEIS*,  $I$  is either 10, 13, or 14,  $J$  is 6, and  $K$  is either 3, 1, or 4. Using Equation 4.17, this leads to 87 up to 118 weight parameters, meaning that the TLFF networks require at least 870 to 1180 training examples. However, the simulated database of scenario earthquakes includes only 280 examples and is thus insufficient.

Böse (2006) solved this problem by artificially enlarging the database: *PreSEIS* adds five sets of events with erroneously picked P-wave onsets showing random delays of up to 1 s. In this way, the number of training examples is increased to 1680 patterns.

Figure 4.5 illustrates the mean location and magnitude errors from the application of *PreSEIS* to the 280 simulated earthquake scenarios (Chapter 2). The TLFF networks are trained with 80% of the scenarios which were selected randomly. The trained networks are then applied to the test dataset, i.e. the remaining 20% of the dataset. Each subfigure displays the temporal development of the source parameter estimates which are updated every 0.5 s. The time axes start at the time



**Figure 4.5:** Absolute errors of hypocentre locations (a+b) and mean magnitude errors (c-f) from the application of *PreSEIS* to the 280 simulated earthquake scenarios. The training dataset consists of 80% of the scenarios, while the remaining 20% define the test subset. The magnitude errors in (c) and (d) result from the training with the parameterized ground motion observations, while (e) and (f) result from usage of the original acceleration times series. The grey-shaded areas give the standard deviation of the magnitude errors. All curves are smoothed over 1.5 seconds.

where the P-waves of each earthquake scenario reach the first seismic station.

The location errors (Figure 4.5a and b) are defined as the differences between the *PreSEIS* estimates and the true hypocentre locations. They are classified by their 25th, 50th, 75th, and 95th percentile, whereas the 50th percentile represents the median. The location errors decrease with ongoing time as additional stations trigger.

Figure 4.5c and d display the mean magnitude errors and the standard deviations (grey-shaded areas) of all 280 earthquake scenarios derived from the training and test datasets. It can be clearly seen that the errors again decrease with ongoing time. These magnitude errors are obtained from feeding the neural networks with the parameterized ground motion (*CAV* values), as previously described.

For the purpose of comparison, I repeated the analysis using the identical training and test subsets. However, the neural networks were this time not trained with parameterized ground motion, but with the observed one. This means I fed the original acceleration time series of 0.02 s sampling rate into the neural networks without taking absolute values or calculating *CAV*, and let the ANNs map this original ground motion onto the source parameters. The source parameter estimates are thus updated each 0.02 s.

Naturally, the location errors are not influenced by this procedure since they are based on the P-wave arrival times which did not change. I do not show them in Figure 4.5. The magnitude estimates are influenced, however, as can be seen in Figure 4.5e and f. The standard deviations are considerably higher and do not show such a strong decrease with ongoing time as the ones before (Figure 4.5c and d). This stresses the importance of parameterizing observed ground motions for earthquake early warning purposes.

The results shown in Figure 4.5 are smoothed over 1.5 s using a moving average procedure to eliminate possible outliers. This is a legal and commonly used procedure with the outputs of TLFF networks (Böse, 2006).

In addition to the above described source parameter estimates, *PreSEIS* is able to predict the expected level of ground shaking at a specified user site at each time step. This is done on the basis of empirical attenuation relationships. Böse (2006) determined the seismic intensity  $I_{FAS}$  of each ground motion simulation using the Fourier amplitude spectrum (FAS) method by Sokolov (2002), as described in Chapter 2.4. An empirical attenuation law has then been established from the simulated database for the Marmara region. The seismic intensity at a given site for a given earthquake with magnitude  $M$  can be estimated from

$$\hat{I}_{FAS}^n(\lambda, \theta) = \exp(1.2655 + 0.2089\hat{M}^n - 0.2451 \log(\hat{d}_{rup}^n + 2.1502\hat{M}^n) - 0.0011\hat{d}_{rup}^n) \quad (4.18)$$

where  $\hat{d}_{rup}^n$  is the distance between the sites  $(\lambda, \theta)$  and  $(\hat{\phi}, \hat{\varphi})$  at each time step  $n$  (Böse et al., 2008):

$$\hat{\phi}^n = \hat{\Phi}_{start}^n + \frac{(\lambda - \hat{\Phi}_{start}^n)(\hat{\Phi}_{end}^n - \hat{\Phi}_{start}^n) + (\theta - \hat{\Theta}_{start}^n)(\hat{\Theta}_{end}^n - \hat{\Theta}_{start}^n)}{(\hat{\Phi}_{start}^n - \hat{\Phi}_{end}^n)^2 + (\hat{\Theta}_{start}^n - \hat{\Theta}_{end}^n)^2} \cdot (\hat{\Phi}_{end}^n - \hat{\Phi}_{start}^n) \quad (4.19)$$

and

$$\hat{\varphi}^n = \hat{\Theta}_{\text{start}}^n + \frac{(\lambda - \hat{\Phi}_{\text{start}}^n)(\hat{\Phi}_{\text{end}}^n - \hat{\Phi}_{\text{start}}^n) + (\theta - \hat{\Theta}_{\text{start}}^n)(\hat{\Theta}_{\text{end}}^n - \hat{\Theta}_{\text{start}}^n)}{(\hat{\Phi}_{\text{start}}^n - \hat{\Phi}_{\text{end}}^n)^2 + (\hat{\Theta}_{\text{start}}^n - \hat{\Theta}_{\text{end}}^n)^2} \cdot (\hat{\Theta}_{\text{end}}^n - \hat{\Theta}_{\text{start}}^n) \quad (4.20)$$

The introduction of a specific intensity threshold allows to define specific warning levels. If the estimated intensity is smaller than the defined threshold, the earthquake is probably non-damaging. If the estimated intensity is equal or greater than the threshold, an alarm is necessary.

Böse (2006) defined following alarm levels:

A *correct* alarm means that the earthquake was correctly identified as being damaging. A *missed* alarm means that no alarm was issued although the earthquake is damaging. *Correct all-clear* means that an event was correctly identified as non-damaging, and a *false* alarm means an alarm that was issued although the earthquake is non-damaging.

Böse (2006) set the intensity threshold to 5.5. An ideal performance of *PreSEIS* will only issue correct alarms and correct all-clear.



# Chapter 5

## Application of PreSEIS

So far, the *PreSEIS* methodology described in the previous chapter has been applied successfully to the synthetic data simulated for the Marmara region in Turkey (Böse, 2006; Böse et al., 2008). The study showed a robust performance of the algorithm and demonstrated a clear and fast convergence of the source parameter estimates toward correct solutions. However, the application to synthetic data is of limited meaning; aspects such as simulated and thus simplified local site effects or added noise, for example, can only reflect the real effects up to a certain extent.

In the following section, *PreSEIS* will be applied to a dataset from Southern California. The study represents the first performance test of *PreSEIS* using real earthquake data which are, in this case, combined with empirical relations from Southern California by Cua (2005). A summary of the study has been published in Köhler et al. (2009).

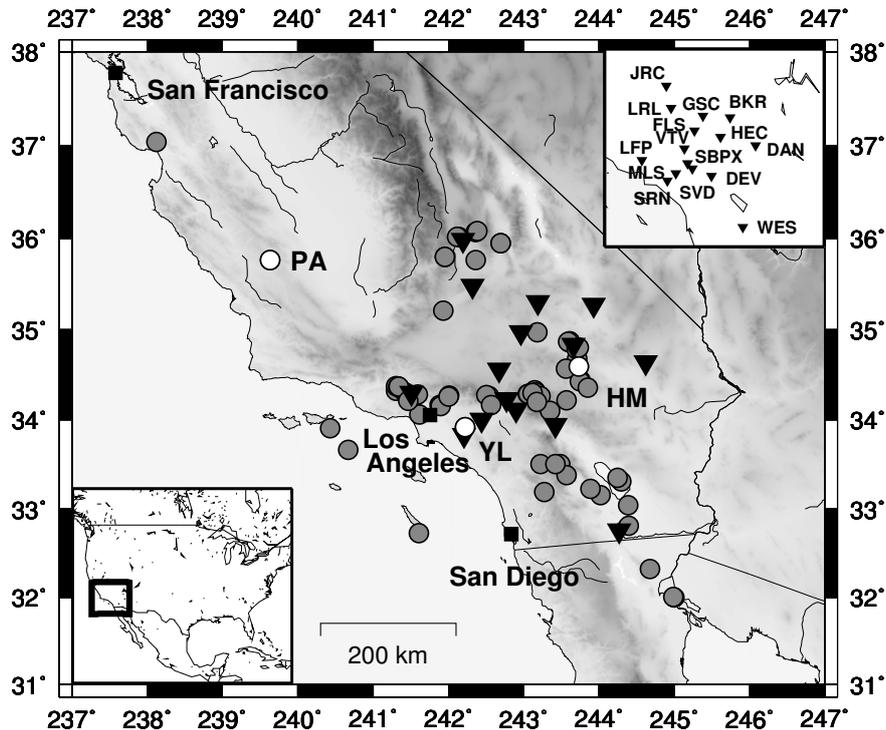
In Section 5.2, the method will be applied to a set of earthquakes from Japan recorded at stations of the Japanese K-NET. The dataset consists exclusively of observed ground motion records.

In Section 5.3, *PreSEIS* will be used as a benchmark system for evaluating earthquake early warning in Istanbul using the synthetic ground motion data simulated for the sites of the ten EEW stations from *IERREWS*. The dataset will also be used to analyze the performance of the amplitude threshold system of *IERREWS*.

### 5.1 Rapid Source Parameter Estimates of Southern California Earthquakes

The purpose of this *PreSEIS* performance test is to analyze the method's functionality in terms of two aspects: first, the capability of *PreSEIS* to handle real earthquake data and second, the operational suitability of *PreSEIS*. Furthermore, questions that remain open before a possible future implementation of the method need to be defined.

In order to evaluate the functionality of *PreSEIS*, a set of Southern California earthquakes recorded



**Figure 5.1:** Map of Southern California showing the locations of the 15 SCSN station sites (black inverted triangles) and the epicentral locations of the 74 training earthquakes (grey circles). The white circles represent the locations of the 1999  $M_w$  7.1 Hector Mine (HM), 2002  $M_w$  4.8 Yorba Linda (YL), and 2004  $M_w$  6.0 Parkfield (PA) earthquakes. The black squares indicate the locations of large cities within the study area.

at the *Southern California Seismic Network (SCSN)* is used. The data is a subset of the data used by Cua (2005) to establish the *Virtual Seismologist (VS)* method. The data subset consists of 70 earthquakes plus an additional set of four events that occurred in the same source region. The earthquakes have source depths ranging between 0.01 and 31.39 km and moment magnitudes in the range  $2.0 \leq M_w \leq 7.3$ . In this study, I limit the number of recording stations to 15. However, if enough training examples are available, the possible number of stations that can be included in *PreSEIS* is not limited.

Figure 5.1 shows a map of Southern California with the locations of the 74 earthquakes. The earthquake located furthest northwest is a  $M_w = 7.0$  event which I decided to include into the catalogue for investigating how the location estimates will be influenced if the seismicity pattern shows statistical outliers. For the purpose of comparison, the event will later be removed from the catalogue and the source parameter estimation process will be repeated.

The records of the 74 earthquakes are used as training data for the neural networks. The source parameter estimation performance is then tested on three example events from the catalogue, the 3 September 2002  $M_w$  4.8 Yorba Linda earthquake, located at  $33.92^\circ\text{N}$  and  $-177.78^\circ\text{W}$  at a depth

of 12.9 km (Hauksson et al., 2002), the 16 October 1999  $M_w$ 7.1 Hector Mine mainshock, located at 34.59°N and 116.27°W at a depth of 5 km (Graizer et al., 2002), and the 8 September 2004  $M_w$ 6.0 Parkfield earthquake, located by SCSN at 35.82°N and 120.37°W at a depth of 7.9 km (Figure 5.1). These events represent characteristic settings: the Hector Mine and Yorba Linda earthquakes are located within the network; the Hector Mine earthquake is located near one station but with some distance to the next stations; the Parkfield earthquake is located remotely.

The earthquake records were downloaded from the *Southern California Earthquake Data Center*<sup>1</sup>. When possible, the 100 samples per second, high gain, broadband (HH) channel was taken. A baseline correction was applied and the data were corrected for the instrument gain to obtain ground motion velocity. The velocity records were differentiated to obtain ground motion acceleration. When the HH channel was clipped, the 100 samples per second, low gain accelerometer channel was downloaded instead. A baseline correction was applied and the data were again corrected for the instrument gain to obtain ground motion acceleration. The acceleration data were integrated to get ground motion velocity (Cua, 2005; Cua and Heaton, 2007).

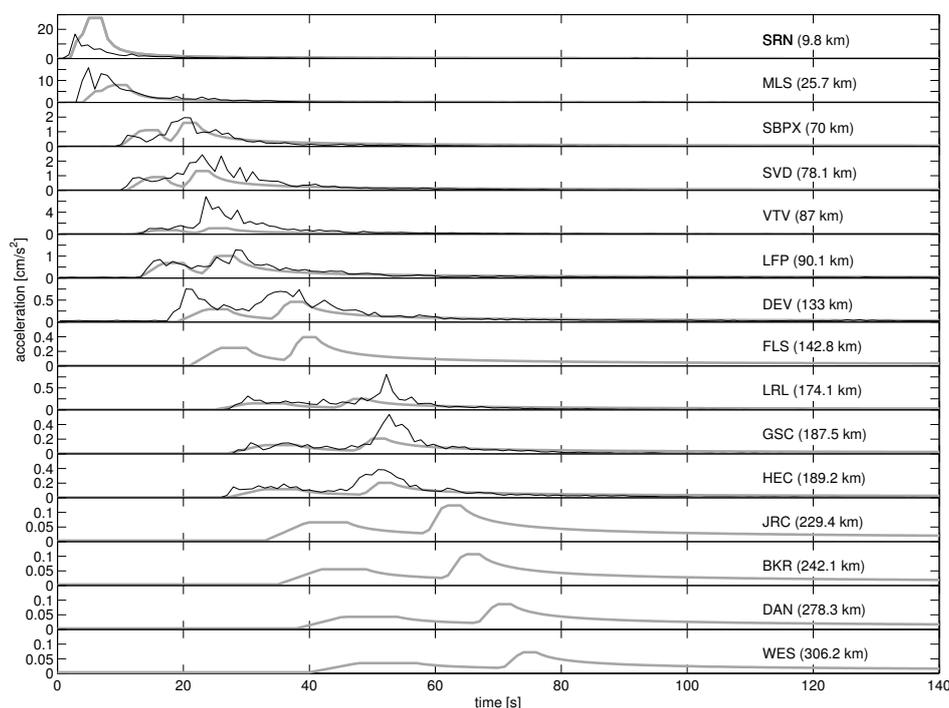
Due to missing records or poor signal-to-noise ratios, only 40% of the records from the 15 selected stations were available. Since *PreSEIS* requires a complete dataset (see Chapter 4.2) the missing records are replaced by synthetic envelopes, predicted by applying the envelope attenuation relationships established by Cua (2005). These relationships allow to calculate the ground motion envelope as a function of time, given the magnitude and epicentral distance of the earthquake. The relationships were inferred from observed envelopes by parameterizing them as a function of P- and S-wave envelopes and ambient noise at the station. The relationships predict envelopes for peak vertical and root mean square of the peak horizontal acceleration, velocity, and displacement data for both rock and soil site classifications. Details on the envelope attenuation relationships are given in Appendix D.

For the prediction of missing envelopes in this study, the P- and S-wave onset times at the 15 stations were determined assuming constant seismic velocities of  $\alpha = 6.8$  km/s and  $\beta = 3.9$  km/s. Because of the possibility to predict missing records, the dataset from the VS method (Cua, 2005; Cua and Heaton, 2007) is highly suitable for the application of *PreSEIS*.

Figure 5.2 shows, as an example, the observed and predicted ground motion envelopes of vertical acceleration from the Yorba Linda earthquake. The records are sorted according to their source-to-site distances. At the stations where no observed record was available, only the predicted envelopes are displayed. Except for stations MLS and DEV, the predicted P-wave envelopes fit the observed ones reasonably well. For the closest source-to-site distance (station SRN), the synthetic S-wave envelope is highly overpredicted. On the contrary, the greater the source-to-site distances become, the more underpredicted become the synthetic S-wave envelopes. Nonetheless, the envelope attenuation relationships represent the best possibility to synthesize the missing earthquake records, whereas the envelope coefficients for predicting acceleration envelopes give better results than the ones for velocity (Cua, personal communication). In the following analyses, the predicted envelopes will only be used when no observations are available.

---

<sup>1</sup><http://www.data.scec.org/>



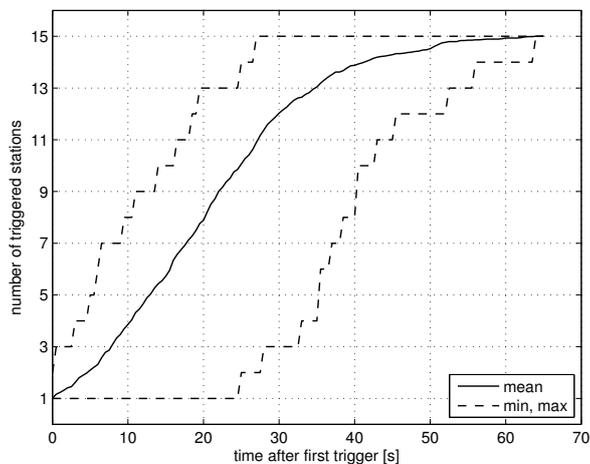
**Figure 5.2:** Observed (black curves) and predicted (grey curves) ground motion envelopes of vertical acceleration from the  $M_w$ 4.8 Yorba Linda earthquake at the 15 SCSN stations. The records start at the earthquake origin time. The epicentral distances are given in brackets.

### 5.1.1 PreSEIS Source Parameter Estimates

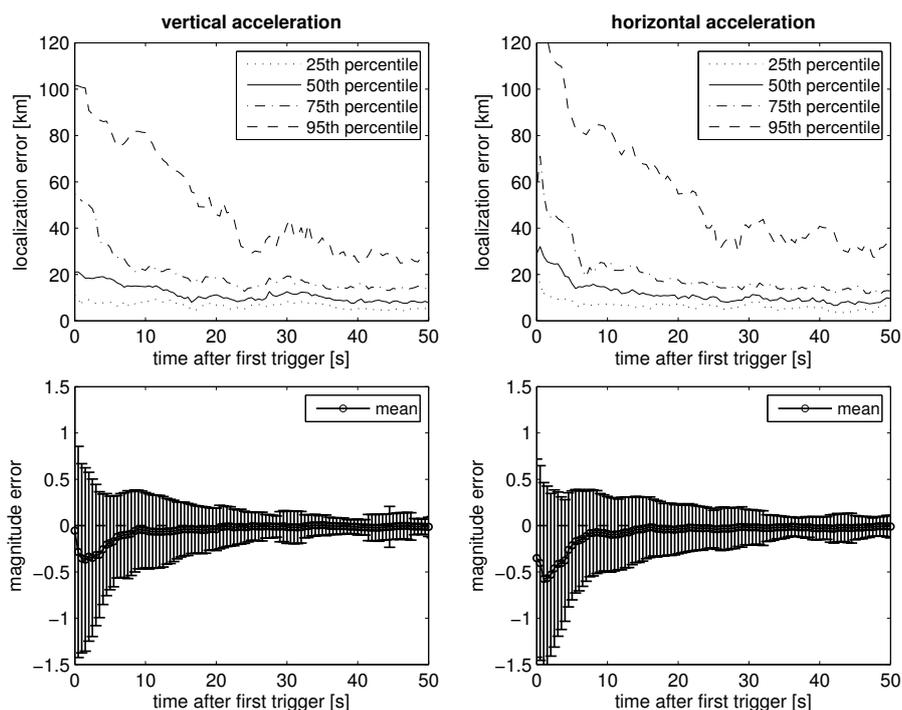
The neural networks are trained with the P-wave onset times of the 74 earthquakes at all 15 stations and with the cumulative sum of their envelopes. The training and the source parameter estimations of *PreSEIS* are performed for both vertical and horizontal acceleration and velocity. The predicted envelopes represent a combined horizontal component. For the observed envelopes I use the NS-components; comparisons between the NS- and EW-components gave no significant differences. Additionally, the presented results are obtained from the use of acceleration data. Using velocity data instead did not change the overall results.

As it can be seen from Figure 5.1, the epicentral distances between the earthquakes and the stations vary widely. This is reflected in the average time which is needed until additional ground motion information becomes available at subsequent triggered stations. Figure 5.3 demonstrates that, after the first station has been triggered, ground motion information from a second station is, on average, available after 5.5 s. Information from a third station is, on average, available after 8.5 s. To reach all 15 stations, the P-waves need an average time of 64.5 s.

Figure 5.4 displays the absolute errors in hypocentre location (top) and mean magnitude errors



**Figure 5.3:** Mean temporal distribution of the number of triggered stations for all 74 earthquakes (solid line) and their minima and maxima (dashed lines).

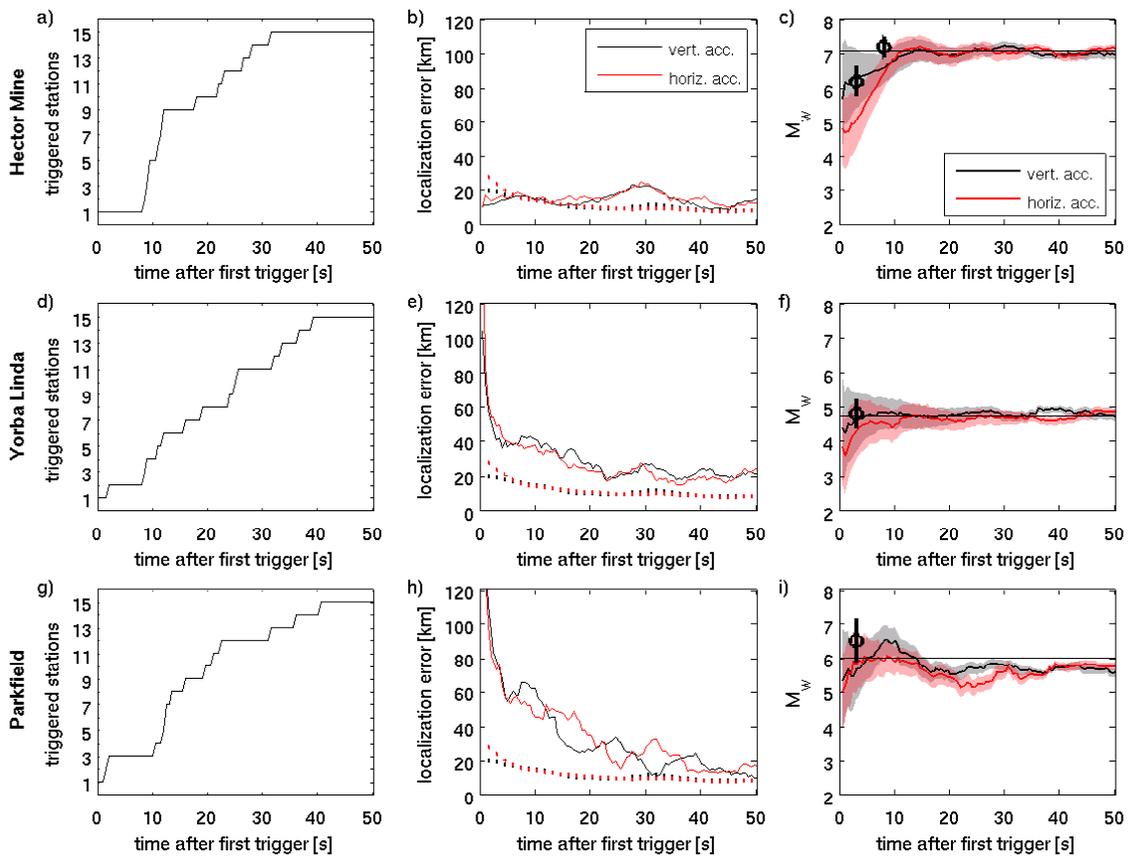


**Figure 5.4:** Absolute errors of hypocentre locations (top) and mean magnitude errors (bottom) from the training process with all 74 earthquakes using vertical (left) and horizontal (right) acceleration data. The localization errors are specified by the 25th, 50th, 75th, and 95th percentiles of the error distributions. The magnitude errors are given by their means (circles) and their standard deviations (error bars). The time axes start with the P-wave arrivals at the first respective seismic station.

(bottom) with proceeding time after the initial P-wave onsets that result from training the neural networks with all 74 earthquakes. The localization errors are defined as the differences between the true (*SCSN*) and the estimated hypocentre locations, thus including the source depths. The errors are classified as their 25th, 50th, 75th, and 95th percentiles, with the 50th percentile marking the median. The results presented on the left hand side of Figure 5.4 are obtained by using vertical acceleration data, while the results on the right hand side are based on the horizontal components.

The localization errors of both components clearly decrease with ongoing time, with median errors of 21.0 km (vertical components) and 29.0 km (horizontal components) at 0.5 s after the initial P-wave arrival. At 8.5 s, i.e., at the time at which data from a third station is available, the errors decrease to 14.9 km and 15.8 km, respectively. After about 15 s, the location errors reach a roughly constant level with a median error of about 9 km. The magnitude errors (Figure 5.4 bottom) also decrease with ongoing time. At 0.5 s after the initial P-wave arrival, standard deviations of  $\pm 1.55$  (vertical components) and  $\pm 1.07$  magnitude units (horizontal components) are obtained. At 8.5 s, the standard deviations are reduced to  $\pm 0.45$  and  $\pm 0.46$  magnitude units, respectively. They continuously decrease down to  $\pm 0.08$  and  $\pm 0.09$  magnitude units, respectively, until the end of the investigated time span.

Figure 5.5 shows the temporal distribution of triggered stations and the *PreSEIS* source parameter estimates for the Hector Mine (top row), Yorba Linda (middle row), and Parkfield earthquake (bottom row). Although the  $M_w 7.1$  Hector Mine earthquake occurred close to station HEC, it takes 9.0 s until the P-wave arrives at the second-nearest station (Figure 5.5a). Another seven stations trigger within the following 3.5 s. After 32.0 s, the P-wave has arrived at all 15 stations. The localization errors of vertical (black curves) and horizontal (red curves) acceleration data (Figure 5.5b) show a roughly constant level with proceeding time. The initial location errors at 0.5 s are 11.2 and 10.7 km for the vertical and horizontal components, respectively. At 9.0 s (corresponding to two triggered stations) the errors have values of 16.2 and 16.4 km, respectively. Averaging the location errors over the analyzed time span of 50 s yields mean errors of 14.6 and 16.0 km which is 1.0 km larger than the median training error from vertical acceleration data and 1.1 km larger than the median training error from horizontal acceleration data. The initial moment magnitude estimates for the Hector Mine earthquake (Figure 5.5c) at 0.5 s are  $M_w = 5.7 \pm 1.6$  for the vertical components and  $M_w = 4.8 \pm 1.1$  for the horizontal components and are therefore 1.4 and 2.3 magnitude units smaller than the true *SCSN* magnitude. With triggering of the second station (after 9.0 s), the estimates improve to  $M_w = 6.6 \pm 0.4$  and  $M_w = 6.8 \pm 0.5$ . Although the initial magnitude estimate using the horizontal components is almost 1 magnitude unit smaller than the one obtained from vertical data, the estimates reach the true magnitude level slightly earlier, after almost 11.0 s. The magnitude estimates from the vertical components reach the true magnitude level after 13.5 s. For comparison, the *Virtual Seismologist (VS)* magnitude estimates for this event at 3 and 8 s are indicated as black error bars and have values of  $M_w(3s) = 6.2 \pm 0.5$  and  $M_w(8s) = 7.2 \pm 0.3$  (Cua, 2005). The *VS* estimate at 3 s agrees well with the *PreSEIS* estimate obtained from vertical component data. However, the *VS* estimate at 8 s indicates that the method predicts the true magnitude level for this event faster than *PreSEIS*. It must be noted, however, that the *VS* uses observations from a larger set of



**Figure 5.5:** Number of triggered stations (left) and *PreSEIS* source parameter estimates for the  $M_w$ 7.1 Hector Mine (top),  $M_w$ 4.8 Yorba Linda (middle), and  $M_w$ 6.0 Parkfield earthquake (bottom). Middle column: Absolute errors of hypocentre locations (solid curves) with mean errors (dotted curves) obtained from the training process in Figure 5.4 (50th percentile). Right column: Moment magnitude estimates (solid curves) with standard deviations from the training process (shaded areas). The black curves represent the results obtained from vertical components, whereas the red curves represent the results obtained from horizontal components. All curves are smoothed over 6.5 s.

seismic stations than the 15 selected stations used in this study.

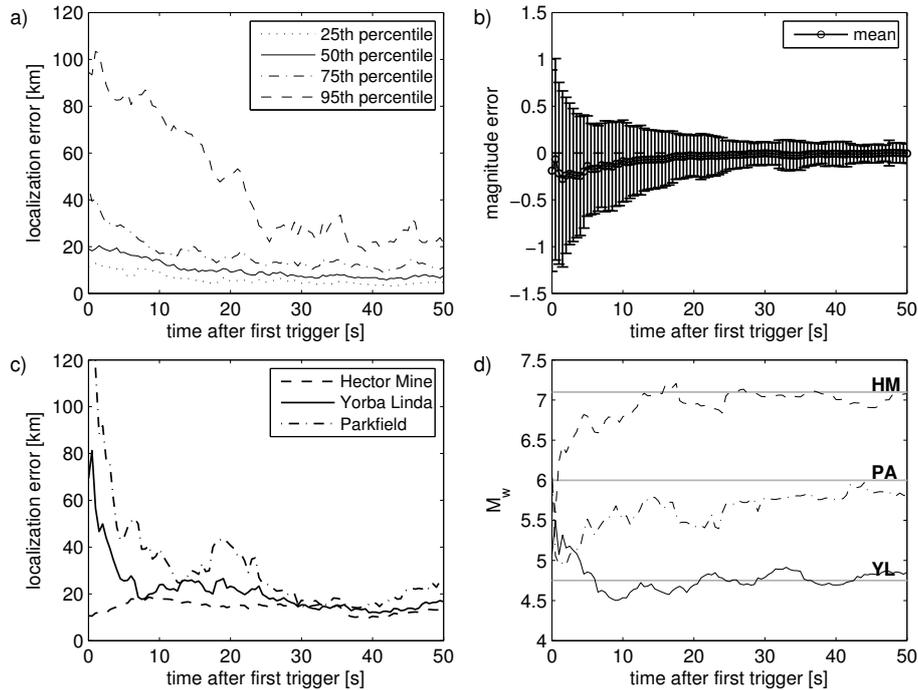
The epicentre of the  $M_w$ 4.8 Yorba Linda earthquake is closely located to two of the seismic stations, whereas the second station triggers only 1.8 s after the first one (Figure 5.5d). The third and fourth stations trigger within the subsequent 7.7 s. It takes 39.5 s until ground motion information is available from all 15 stations. The initial localization errors (Figure 5.5e) at 0.5 s from the two components are high, showing values of 104.1 km (vertical components) and 158.4 km (horizontal components). With the contribution from the second station after 1.8 s, the errors can be reduced to 51.2 and 54.4 km, respectively. After about 4 s, both location errors decrease to about 40 km. With ongoing time the errors can be continuously reduced. The initial magnitude estimates (Figure 5.5f) yield values of  $M_w = 4.4 \pm 1.6$  (vertical components) and  $M_w = 3.9 \pm 1.1$  (horizontal components). At the initial time step the magnitudes are thus underestimated by 0.4 and 0.9 magnitude units, respectively. The correct magnitude level is reached after 5.0 s for the vertical components and 10.0 s for the horizontal components. The VS estimate by Cua (2005) gives a magnitude of  $M_w(3s) = 4.8 \pm 0.4$  at 3 s, again being in good agreement with the PreSEIS results from the vertical component acceleration data.

For the  $M_w$ 6.0 Parkfield earthquake (Figure 5.5g), the first three stations trigger within 2.5 s. It takes another 8.5 s until the P-wave reaches the fourth station. All 15 stations are triggered after 41 s. As for the Yorba Linda earthquake, the initial localization errors are large (Figure 5.5h), showing values of 189.5 km (vertical components) and 176.1 km (horizontal components). Within 5-10 s, the errors are reduced to about 60 km. The location errors continuously decrease with ongoing time. The initial magnitude estimates (Figure 5.5i) at 0.5 s are  $M_w = 5.3 \pm 1.6$  (vertical components) and  $M_w = 5.0 \pm 1.1$  (horizontal components) and thus underestimate the size of the earthquake by 0.7 and 1.0 magnitude units, respectively. After 2.5 s, the estimates improve to  $M_w = 5.4 \pm 0.9$  and  $M_w = 5.9 \pm 0.9$ , respectively. After reaching the correct magnitude level, the magnitude is overestimated for a couple of seconds by using the vertical components. The VS estimate by Cua (2005), which is  $M_w(3s) = 6.5 \pm 0.7$  at 3 s, suggests a similar trend.

Figure 5.6 illustrates how the previously presented results change if the  $M_w$ 7.0 event, which is located far outside of the network (Figure 5.1), is removed from the training catalogue. In the following, I will only concentrate on the results obtained from vertical acceleration data.

Figure 5.6a shows that the location errors decrease by a minimal amount compared to the complete catalogue before (Figure 5.4). The initial estimate of the 75th percentile is reduced by 9 km. The average value of the 50th percentile over all 50 s is 10.1 km and thus 3.5 km smaller than before. The magnitude errors show an initial standard deviation of  $\pm 1.1$  magnitude units which is half a magnitude unit less than for the complete catalogue. The variations of the standard deviations at later time steps (between 30 and 50 s) also decreased. Figure 5.6c and d present the source parameter estimates for the Hector Mine, Yorba Linda, and Parkfield earthquakes. For the Hector Mine event, the localization error and the magnitude estimates do not show significant changes. However, the initial magnitude estimate is smaller than in Figure 5.5 and it takes about 2 s longer until the true magnitude level of 7.1 is reached.

For the Yorba Linda earthquake, the initial localization error is with 70 km about 35 km smaller



**Figure 5.6:** Absolute errors of hypocentre locations (a) and mean magnitude errors (b) from the training process removing the  $M_w 7.0$  event from the training catalogue (vertical acceleration data). (c) and (d) show the location errors and magnitude estimates for the  $M_w 7.1$  Hector Mine (HM, dashed curves),  $M_w 4.8$  Yorba Linda (YL, solid curves), and  $M_w 6.0$  Parkfield (PA, dash-dotted curves) earthquake. The curves in (c) and (d) are smoothed over 6.5 s.

than in the performance before. After about 8 s, the error decreases to an average error around 20 km. The initial magnitude estimate gives  $M_w = 5.1 \pm 1.1$  and thus overestimates the earthquake size by 0.3 magnitude units.

The initial localization error for the Parkfield earthquake is again larger than 120 km, but can be reduced slightly faster than in Figure 5.5. The initial magnitude estimate gives  $M_w = 5.9 \pm 1.1$  which is 0.6 magnitude units larger than the estimate obtained from the complete catalogue and only 0.1 magnitude units smaller than the true magnitude. However, the magnitude estimate decreases significantly within the following seconds and reaches the true magnitude level only after more than 40 s. Removing the  $M_w 7.0$  earthquake from the training data does therefore not influence the overall source parameter estimates significantly, and no positive influence can be noticed for the magnitude estimation of the Parkfield earthquake which is also located further outside of the network. The enclosure of statistical outliers in the training data thus widens the generalization capability of the neural networks.

In addition to the presented analyses, I repeated the *PreSEIS* performance tests by replacing all observed ground motion envelopes with predicted ones so that the database exclusively consisted

of synthetic data. The source parameter estimates obtained from this purely synthetic dataset did not differ significantly from the ones obtained from the combination of observed and synthetic data, neither in the size of errors nor in their temporal development. This stresses on the one hand the great flexibility of the neural networks and, on the other hand, the high quality of the predicted ground motion envelopes. In order to give the predicted ground motion envelopes a more realistic image, I additionally added a constant Gaussian noise signal of 1 Hz to all predicted envelopes and repeated the training process. Again, this did not influence the source parameter estimates significantly.

### 5.1.2 Discussion

The *PreSEIS* performance using the data from Southern California clearly shows two different results: The first one is the large size of initial location errors for the Yorba Linda and Parkfield earthquakes. However, these errors should be considered relative to the overall expansion of the study area. The Yorba Linda earthquake is closely located between two stations, but of course, not the absolute travel times to these stations are considered for the *PreSEIS* estimates but only the arrival time differences. The two clusters of earthquakes in the training database from which similar arrival time differences could arise are located northwest and southeast of the Yorba Linda earthquake in average distances of about 100 and 150 km. This fits well to the observed initial location errors of 104 and 158 km, depending on the used component. For the Parkfield earthquake, the training database does not contain any events that are located in the closer vicinity. The variability of possible earthquake locations that satisfy the observed arrival time differences within the first one or two time steps is probably too high to obtain more accurate location estimates for these two earthquakes. This hypothesis is supported by the fact that, once the P-waves arrive at two to three stations, the location errors for both earthquakes drastically decrease.

Second, the magnitude estimates are of remarkably good quality despite the large location errors, considering the fact that the location estimates contribute to the magnitude estimation. The magnitude estimates clearly converge toward their correct solutions with proceeding time and increasing number of triggered stations. This stresses the great generalization capability of ANNs. Once ground motion information is available from two stations, the magnitude estimates are of 90.0% (Parkfield earthquake), 94% (Hector Mine earthquake), and 100% (Yorba Linda earthquake) accuracy, respectively. The time until the magnitudes are estimated correctly is strongly dependent on the station density around the epicentre.

The overall source parameter estimates do not significantly depend on the used component of ground motion. Nonetheless, the magnitude estimates for the three example earthquakes show better initial magnitude estimates from the vertical acceleration data, whereas the magnitude estimates obtained from the horizontal data reach the true values faster than the ones from the vertical data.

The standard deviations of the source parameter estimates, derived from the training process with all 74 events, show highest values within the first few seconds and clearly decrease with progressing time. This demonstrates the inverse relationship between the reliability of estimates and remaining warning time in an EEW system. It also stresses the importance of updating the

source parameter estimates with ongoing time as it is performed in *PreSEIS*.

The major objective of this study was to investigate the functionality of *PreSEIS* in terms of 1) its capability to handle real data, and 2) its operational suitability.

1) The application of *PreSEIS* to real data cases combined with synthetics leads to a stable and robust performance. The quality of outputs depends neither on the number of missing observations that had to be replaced by predictions nor on added noise. Indeed, repeating the test using exclusively synthetic data did not change the quality of results. The major remaining problem of the application of *PreSEIS* to real data will thus be the existence of non-functional stations. However, I am confident that suitable training epochs including various configurations of failing stations can reduce the impact of missing observations or even completely eliminate this effect.

2) This study confirms that ANNs can indeed be suitable and attractive tools for EEW applications. The stability of results proves the tolerance of ANNs toward various types of input data, which is of great advantage in transferring the method to different regions. The study also shows that the time necessary to obtain reliable source parameter estimates can be reduced by more appropriate geometrical settings and shorter interstation distances. A condition for applying the method to a new region would be the availability of a suitable training dataset, which should be as extensive and various as possible to allow for a vast amount of *a priori* information.

In summary, I conclude that the major remaining challenge for an implementation of the *PreSEIS* methodology is the incorporation of non-functional early warning stations and, in some study areas, the availability of suitable training examples. However, these tasks are potentially resolvable and can be addressed in future work.

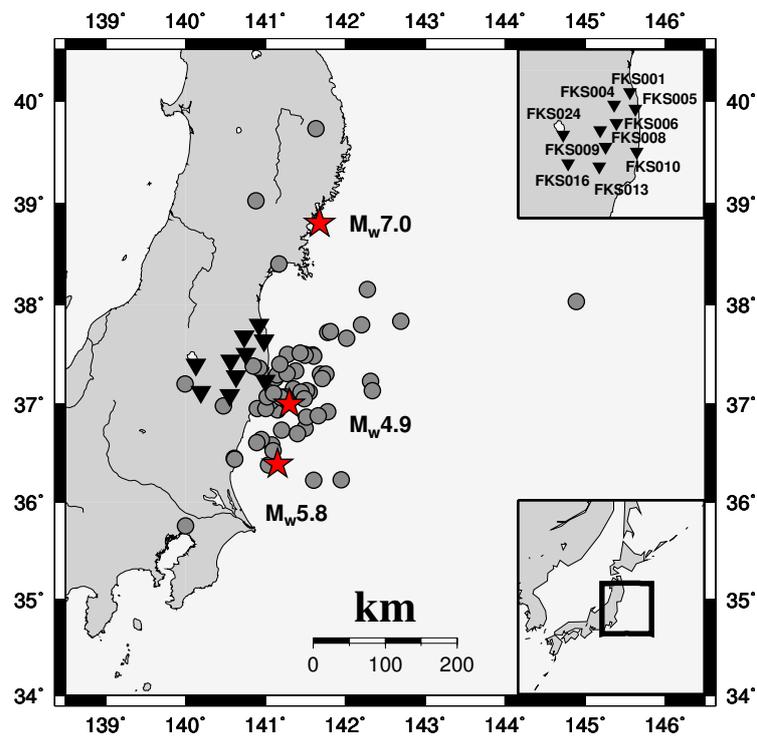
## 5.2 Rapid Source Parameter Estimates of Japanese Earthquakes

This section presents the second *PreSEIS* performance test using real data cases. The dataset consists of 69 earthquakes from Japan in the magnitude range of  $3.8 \leq M_w \leq 7.2$  recorded at ten stations of the *Kyoshin Net (K-NET)*<sup>2</sup>. K-NET consists of 1000 strong motion stations deployed all over Japan with an average interstation distance of about 25 km. The stations record ground acceleration in the range of  $\pm 2000$  Gal (Kinoshita, 2003). The advantage of K-NET data is the vast amount of strong motion records that are freely available. In contrast to the previously described *PreSEIS* application to earthquakes from Southern California, this study will use exclusively observed earthquake records without any additional predictions.

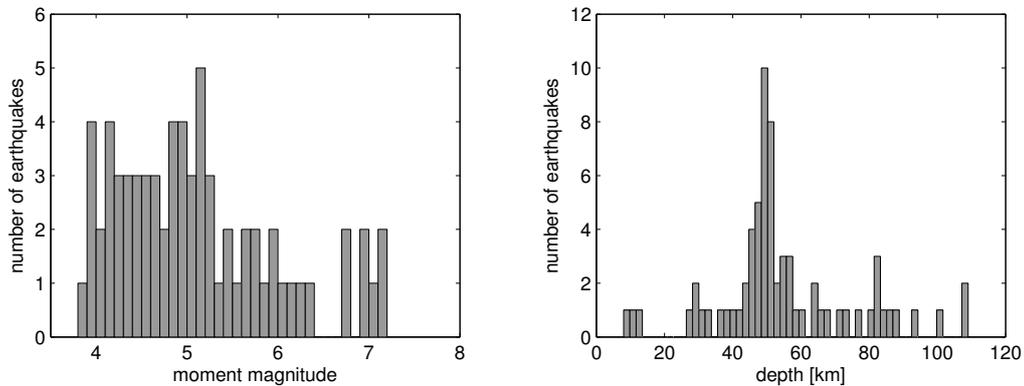
Figure 5.7 displays the locations of the epicentres and the ten seismic stations. The earthquakes were chosen from a large catalogue of earthquakes recorded at all K-NET stations in the years between 1997 and 2009. The subset used in this study represents the largest number of events that have been recorded at the same subset of K-NET stations. This satisfies the requirement of

---

<sup>2</sup><http://www.k-net.bosai.go.jp/>



**Figure 5.7:** Map of Japan with the ten K-NET station sites (black inverted triangles) and epicentral locations of the 69 earthquakes (grey circles). The three example earthquakes used for testing are marked by red stars.



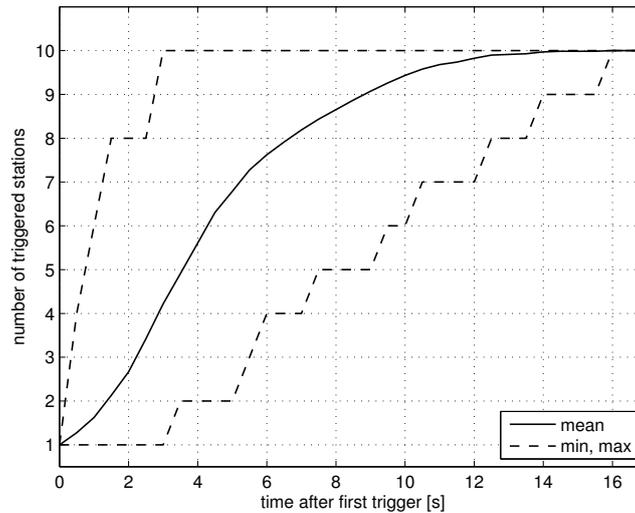
**Figure 5.8:** Distribution of moment magnitudes (left) and depths (right) of all 69 earthquakes from Japan used in this study.

the current *PreSEIS* stage to use complete datasets where no stations failed.

Figure 5.8 shows the moment magnitude and depth distributions of all 69 earthquakes. The earthquake depths have values between 8 and 109 km, whereas the majority shows depths around 48-50 km.

For the application of *PreSEIS*, I removed the mean of the 100 samples per second acceleration time series and applied a 0.5-12 Hz band-pass filter. As for the synthetic data from the Marmara region, the *PreSEIS* magnitude estimates base on the cumulative absolute velocity, *CAV*, calculated for each of the records. The neural networks are trained with 66 of the 69 earthquakes. The *PreSEIS* performance is then tested with the remaining three earthquakes that are not included into the training dataset. The first event is a  $M_w$ 7.0 earthquake located at 141.68°E and 38.81°N which occurred at a depth of 71 km on 26 May 2003. The second event is a  $M_w$ 5.8 earthquake located at 141.15°E and 36.39°N at a depth of 49 km. It took place on 4 April 2004. The third earthquake used for testing the *PreSEIS* performance is a  $M_w$ 4.9 event located at 141.30°E and 37.00°N at a depth of 50 km which occurred on 3 January 2009. The three events were selected arbitrarily in a way that three different magnitude levels were represented. The locations of the three example earthquakes are indicated in Figure 5.7 by red stars. The unfiltered ground acceleration records of these three earthquakes at the ten K-NET stations are presented in Appendix E.

The dense station configuration of the ten selected K-NET stations is reflected in the P-wave arrival time differences between the single stations. Figure 5.9 displays the temporal distribution of triggered stations for the training catalogue. On average, the P-waves arrive at a second station 1.4 s after the arrival at the first station. Ground motion information from three stations are available, on average, 2.2 s after the initial P-wave arrival. To trigger all ten stations the P-waves of the training earthquakes need an average time of about 16 s. In the best case, three stations



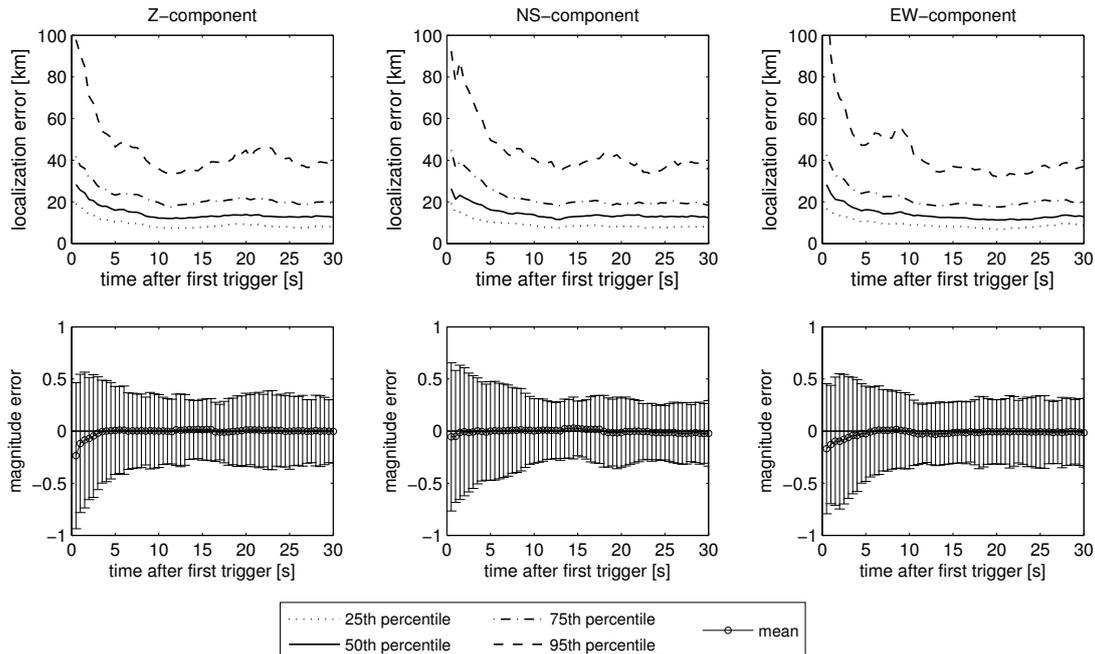
**Figure 5.9:** Mean temporal distribution of the number of triggered stations for all 66 training earthquakes (solid line) and their minima and maxima (dashed lines).

are triggered within 0.3 s, whereas in the most unfavourable case it takes 5.5 s to trigger three stations.

Figure 5.10 shows the hypocentre location errors (top) and mean magnitude errors (bottom) obtained from training the neural networks with the 66 training earthquakes. The training has been accomplished separately for all three ground motion components. The different components show similar results. At the initial time step (0.5 s) the mean location errors (50th percentile) are 28.3 km, 26.3 km, and 28.3 km for the vertical component, the NS-component, and the EW-component, respectively. The location errors are again defined as the deviation of the estimated hypocentre locations from the true ones. After 3 s, corresponding to three triggered stations for most earthquakes, the location errors decrease to 18.7 km, 19.8 km, and 18.5 km, respectively. After 15 s, when the P-waves have already arrived at the majority of stations, the mean location errors can be reduced to 12.9 km, 13.0 km, and 12.5 km, respectively.

The magnitude errors decrease significantly during the first 15 s as well. The three components show initial magnitude errors (after 0.5 s) of  $\pm 0.70$  (vertical component),  $\pm 0.71$  (NS-component), and  $\pm 0.62$  (EW-component) magnitude units. After 3 s, the errors decrease to  $\pm 0.54$ ,  $\pm 0.54$ , and  $\pm 0.59$  magnitude units, respectively. After 15 s, the errors can be reduced to  $\pm 0.29$ ,  $\pm 0.26$ , and  $\pm 0.29$  units, respectively. The magnitude errors do not show further decrease after 15 s; the CAV values obtained from the later arriving phases at the last triggered stations do thus not influence the magnitude estimates anymore.

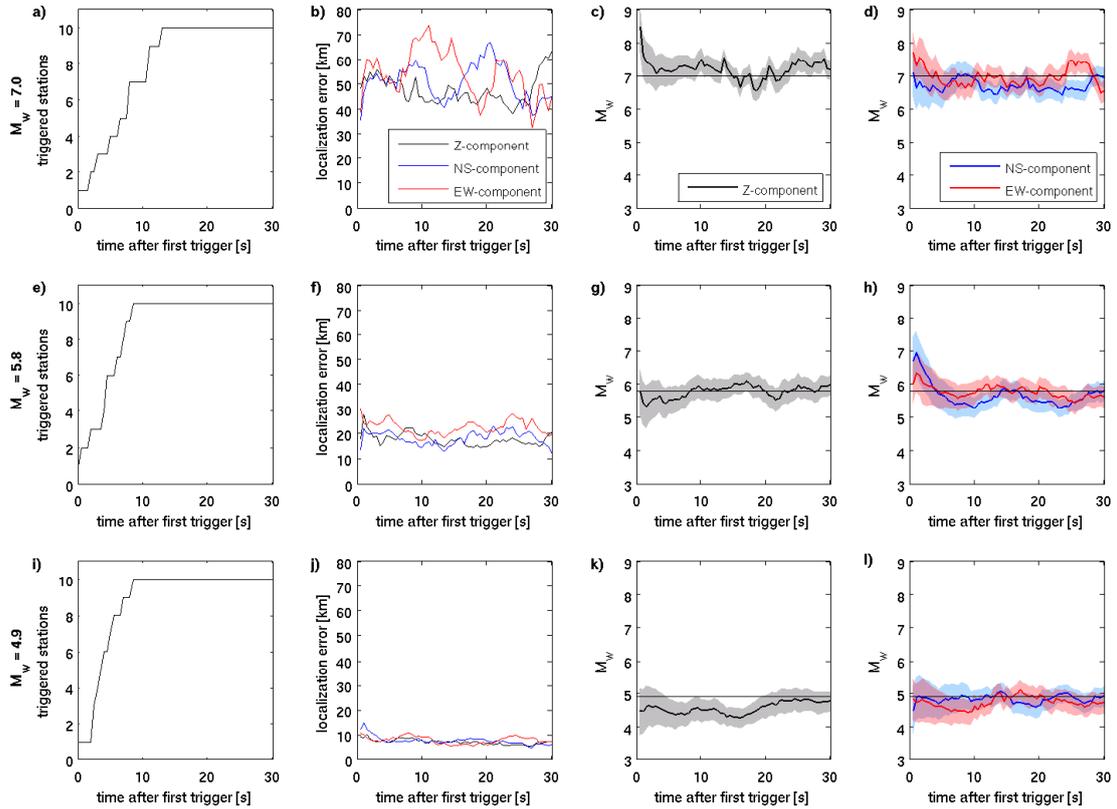
Figure 5.11 shows the location and magnitude estimates obtained from the application of the trained neural networks to the three unknown example earthquakes, again for all three components of ground motion. For the  $M_w 7.0$  earthquake the P-wave arrives at the second nearest



**Figure 5.10:** Absolute errors of hypocentre locations (top) and mean magnitude errors (bottom) obtained from the training process with the 66 earthquakes for the vertical components (left), NS-components (middle), and EW-components (right). The curves are smoothed over 4.5 s.

station only 2.5 s later than at the first station. Three stations are reached after 3.5 s. All ten stations are triggered within 13.5 s (Figure 5.11a). The location errors for all three components (Figure 5.11b) have average values of 47.60 km (Z-component), 47.77 km (NS-component) and 50.84 km (EW-component) over the analyzed time span of 30 s. The curves are fluctuating and do not show a consistent decreasing or increasing trend. However, the initial estimates from all three components at 0.5 s are about 10-12 km smaller than the average location errors over all 30 s.

The initial magnitude estimate obtained from the vertical components (Figure 5.11c) gives a moment magnitude of  $M_w = 8.48 \pm 0.70$  and therefore overestimates the magnitude by 1.48 units. After 4 s, when ground motion information is available from 3 stations, the magnitude estimate decreases to  $M_w = 7.20 \pm 0.48$  which is only 0.2 magnitude units larger than the true magnitude. For the remaining time period, the magnitude estimate slightly fluctuates but shows an average overestimation of magnitude of about 0.1-0.2 magnitude units. The horizontal components give two different magnitude estimates (Figure 5.11d): The initial estimate using the NS-component,  $M_w = 7.09 \pm 0.71$ , fits the true magnitude very well. The remaining time period, however, reveals a slight underestimation of magnitude by about 0.3-0.4 magnitude units. At 4 s, for example, a value of  $M_w = 6.63 \pm 0.47$  is predicted. The estimates obtained from the EW-components, in contrast, also overestimate the magnitude at the initial time step similar to the vertical components. The initial estimate gives  $M_w = 7.72 \pm 0.62$ . After 4 s, this value decreases to  $M_w = 7.13 \pm 0.54$ .



**Figure 5.11:** Number of triggered stations (first column) and *PreSEIS* source parameter estimates for the  $M_w$  7.0 (top),  $M_w$  5.8 (middle), and  $M_w$  4.9 earthquake (bottom). The localization errors (second column) are defined as the differences between the true and the estimated hypocentre locations. The source parameter estimates are separately obtained from the vertical component data (black curves), NS-components (blue curves), and EW-components (red curves). The third and fourth columns display the mean magnitude estimates (solid curves) for the three components, whereas the standard deviations (shaded areas) are obtained from the training process. All curves are smoothed over 4.5 s.

The later time steps show a slight fluctuation around the true magnitude level.

For the  $M_w$ 5.8 earthquake, the second station triggers only 1 s after the first one, and the third station triggers 2.5 s after the first one (Figure 5.11e). The P-wave reaches all ten stations within 9 s. The location errors from the three components (Figure 5.11f) have average values of 18.02 km (Z-component), 18.66 km (NS-component), and 20.16 km (EW-component) over all 30 s. The errors only show a minimal decrease within the first 3 s and can rather be considered as more or less constant during the investigated time period. The location can thus be well confined from the arrival time differences at the first three stations. The magnitude estimates from the vertical component data (Figure 5.11g) are highly accurate already from the first time step on. At 0.5 s, the magnitude is estimated as  $M_w = 5.78 \pm 0.70$  which is only 0.02 units smaller than the true magnitude. Although the following 8 s show a minimal underestimation, the general trend over all 30 s reveals a high consistency of estimates with the true magnitude. Similar results can be obtained from the EW-components (Figure 5.11h), whereas the initial estimate at 0.5 s gives a magnitude of  $M_w = 6.04 \pm 0.62$  and thus overestimates the magnitude by 0.24 magnitude units. The NS-components show a higher initial estimate of  $M_w = 6.72 \pm 0.71$  but decreases within the first 7 s to an average magnitude level of about 5.6.

Finally, for the  $M_w$ 4.9 earthquake the second and third stations trigger simultaneously at 3 s. The P-wave reaches all ten stations after 9 s (Figure 5.11i). The location errors for this example event are the smallest with average errors of 6.82 km (Z-component), 7.45 km (NS-component), and 7.64 km (EW-component) over all 30 s (Figure 5.11j). Again, the arrival time differences from the first three stations are sufficient to estimate the hypocentre location with the smallest error. The magnitude estimates obtained from the vertical components (Figure 5.11k) show a slight underestimation of magnitude. At 0.5 s, the magnitude is predicted as  $M_w = 4.46 \pm 0.70$ . With three triggered stations at 3 s, the estimate slightly increases to  $M_w = 4.60 \pm 0.54$ . A better fit with the true magnitude is obtained after about 22 s. The magnitude estimates from the EW-components show a similar trend (Figure 5.11l). At 0.5 s the magnitude is predicted as  $M_w = 4.83 \pm 0.62$  and at 3 s as  $M_w = 4.73 \pm 0.59$ . The NS-components show an initial estimate of  $M_w = 4.49 \pm 0.71$ . Already at 3 s, this estimate increases to  $M_w = 4.84 \pm 0.54$  which is only 0.06 magnitude units smaller, on average, than the true magnitude.

### 5.2.1 Discussion

The application of *PreSEIS* to the real earthquake data from Japan proves the high quality of the neural network methodology. The mean location and magnitude errors decrease significantly during the first 10-15 s and show smallest average location errors between 12.5 and 13.0 km. The smallest mean magnitude errors obtained from the training have standard deviations of  $\pm 0.26$  and  $\pm 0.29$  magnitude units, depending on the used ground motion component.

The application of the trained neural networks to three unknown example earthquakes results in location and magnitude errors of high quality. The comparatively high average location errors of 47.60 to 50.84 km for the  $M_w$ 7.0 earthquake result from the fact that no training examples

are available in the immediate vicinity of this example event. The magnitude is slightly overestimated for this earthquake using vertical ground acceleration. In general, the magnitude estimates of the three example events highly agree with the true magnitude levels already from the beginning of the investigated time periods. The magnitude estimates obtained from the EW-components are slightly more similar to the ones obtained from the vertical components than to the ones from the NS-components. The location errors for the two earthquakes that are located within the main cluster of training events are small and have average errors of 18.02 to 20.16 km and 6.82 to 7.64 km, respectively, depending on the component of ground motion.

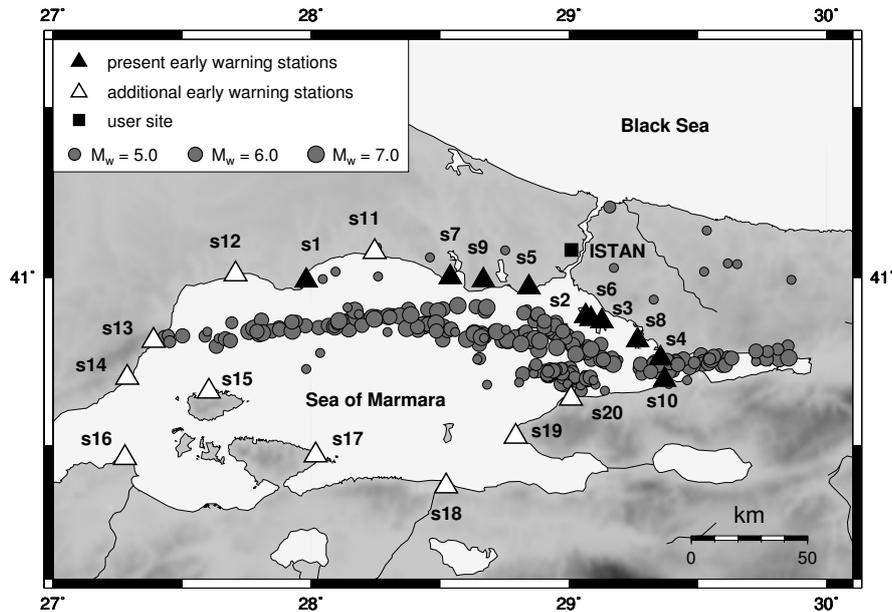
I expect that the parallel integration of all three ground motion components into *PreSEIS* would result in the most optimum performance. This would require some modifications of the *PreSEIS* code which are not addressed in this study but which would be feasible.

The presented study proves that *PreSEIS* can handle datasets consisting exclusively of observed earthquake records as successful as synthetic data. This is an important finding since an implementation of an earthquake early warning algorithm naturally implies the application to pure observations. This second application to real data confirms again the suitability of artificial neural networks for earthquake early warning applications.

### 5.3 EEW for Istanbul

As a consequence of its immediate vicinity to the Main Marmara Fault, the Turkish mega-city Istanbul is exposed to high seismic risk. Earthquake early warning (EEW) is thus an important task for the safety of millions of people living in and around Istanbul. The study presented in this section will evaluate the performance of the EEW system integrated in the *Istanbul Earthquake Rapid Response and Early Warning System (IERREWS)* (Erdik et al., 2003b) which is described in Chapter 1.2. The present EEW network consists of 10 strong motion stations installed along the northern shoreline of the Sea of Marmara, between Istanbul and the Main Marmara Fault (Figure 5.12, Table 2.3). It was recently discussed to expand the network to regional scale by installing an additional set of 10 seismic stations. These stations would be installed around the Sea of Marmara, as can be seen in Figure 5.12.

To evaluate the contribution of the additional 10 stations to early warning, I will make use of *PreSEIS* as a benchmark system; in the previous sections the method has successfully proven its functional suitability. As *PreSEIS* has the advantage of using the full incoming seismic signals it is highly valuable for investigating the performance differences between the current and the expanded network. Furthermore, I will also investigate the possible contribution of the additional EEW stations to the alarming system that is implemented within *IERREWS* (Alcik et al., 2009; Erdik et al., 2003b; see Chapter 1.2). Due to the lack of strong motion observations at the EEW stations, the investigations are based on the synthetic dataset simulated with *FINSIM*, which is presented in Chapter 2.6. It consists of 280 earthquake scenarios (Figure 5.12) with moment magnitudes between 4.5 and 7.5. For the purpose of this study, I also simulated the according acceleration time series of all 280 events for the sites of the 10 additionally planned stations,

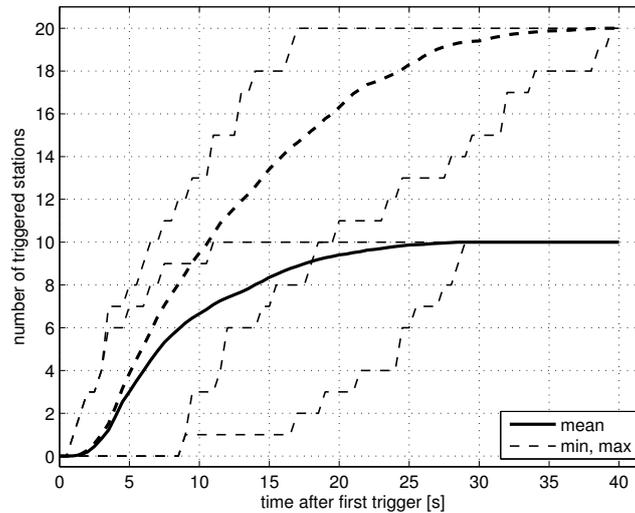


**Figure 5.12:** Map of the Marmara region, Turkey. The black triangles indicate the positions of the current 10 EEW stations (Erdik et al., 2003b). The white triangles represent the 10 additionally planned stations. The grey circles show the distribution of the 280 simulated earthquakes. ISTAN defines a user site within the city of Istanbul.

using the same *FINSIM* parameterization as for the first 10 stations.

Figure 5.13 displays the average number of stations that are triggered by the P-waves of all 280 earthquakes with ongoing time, starting at the earthquake occurrence times. On average, the time between the P-wave arrival at the first station and the time at which three stations are triggered can be reduced in the expanded network by 1 s, from 5 s for the case of 10 stations to 4 s for 20 stations. Even in the most unfavourable case, the expanded network configuration leads to 9 s until three stations are triggered, while the present network of 10 stations needs a maximum time of 18 s. This leads to an increase of the theoretically possible warning times for all 280 events, as displayed in Figure 5.14.

The warning time is defined as the time difference between the initial P-wave detection at the first station and the S-wave arrival at a selected user site ISTAN in Istanbul (Figure 5.12). For the present network of 10 stations the theoretical warning times range between -0.6 s and 32.3 s with an average warning time of 12.6 s (Figure 5.14). A negative warning time means that the earthquake is located in a way that the S-wave would arrive at site ISTAN before any station is triggered by the P-wave. This might happen for one of the events located on the northern or northeastern side of ISTAN. The expanded network of 20 stations yields warning times between -0.6 s and 37.7 s with an average warning time of 13.1 (Figure 5.14). Although the average warning time only increases by 0.5 s, it can be clearly seen in Figure 5.14 that the warning times in the case of the expanded network are distributed more equally than for the present network of



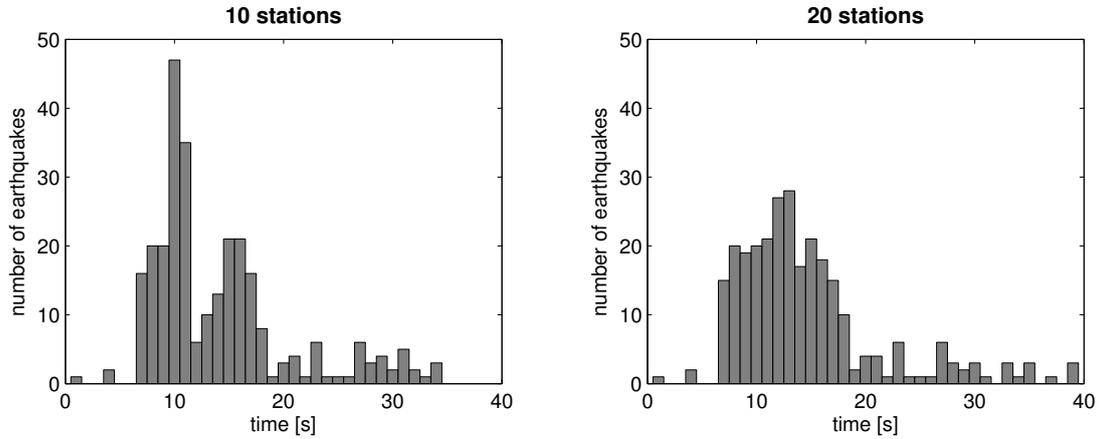
**Figure 5.13:** Average number of triggered stations with progressing time for all 280 earthquakes for the present network of 10 stations (bold solid curve) and for the expanded network of 20 stations (bold dashed curve). The dashed curves in normal font represent the according minima and maxima.

10 stations.

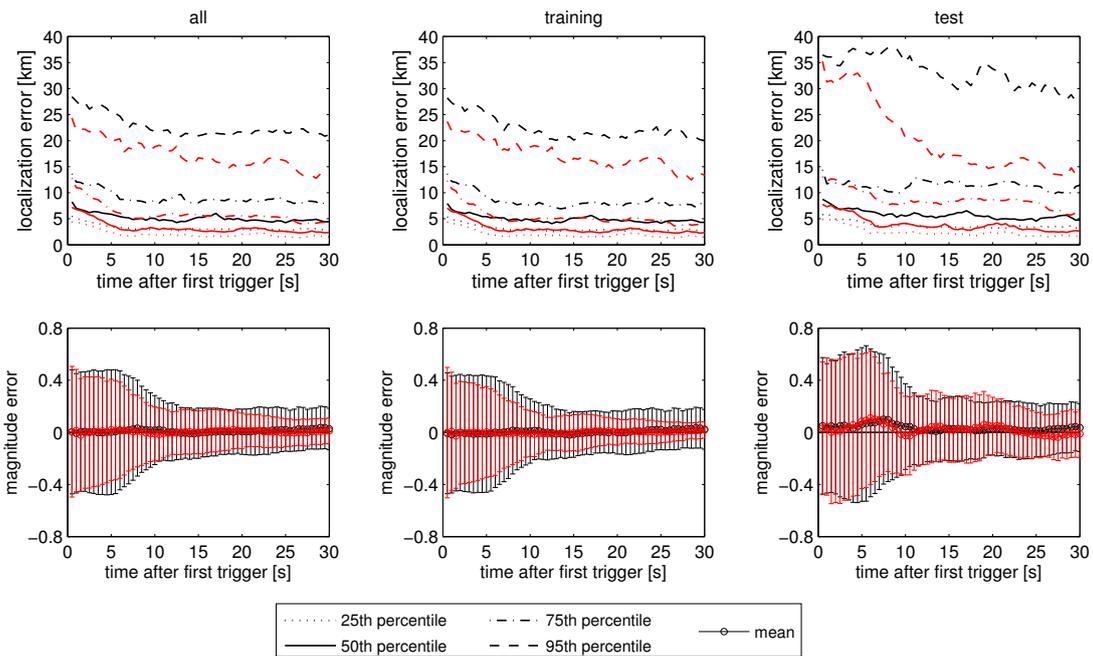
As mentioned above, *PreSEIS* is used as a benchmark method to evaluate the EEW performance differences between the two network configurations. Figure 5.15 shows the location and magnitude errors with ongoing time for all 280 earthquake scenarios (left) as well as for the training (middle) and test (right) processes separately. The training dataset consists of 224 randomly chosen events, while the remaining 56 events build the test dataset which is not included into the training process. The training and test subsets are identical for both network configurations. The location errors are again split into their 25th, 50th, 75th, and 95th percentiles with the 50th percentile marking the median.

As it can be seen at first glance, the expanded network of 20 stations leads to consistently smaller source parameter estimation errors than the present network of 10 stations. While the initial errors at 0.5 s are not significantly different, the subsequent time steps show a clear decrease of errors for 20 stations. After about 3-4 s, the median location errors become 2-3 km smaller. The magnitude errors are, after 2 s, already 8% smaller for all 280 events and 12 % smaller for the training subset. After 5 s, these magnitude errors are already 19% and 22% smaller, respectively. The magnitude errors of the test subset, however, are not significantly influenced by the network expansion.

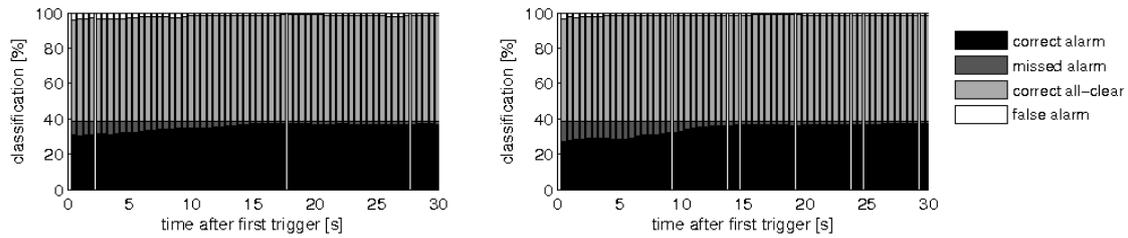
The general improvement of the *PreSEIS* source parameter estimates is naturally caused by the fact that, with 20 stations, the P-waves of the earthquakes can be detected earlier and ground motion information from subsequently triggered stations are not only available from more stations but also at earlier time steps. Thus, the expansion of the EEW network to regional scale leads to



**Figure 5.14:** Warning times of all 280 earthquake scenarios in the case of the original (left) and the expanded network (right).



**Figure 5.15:** Absolute hypocentre location errors (top) and mean magnitude errors (bottom) with progressing time for the network configurations of 10 stations (black curves) and 20 stations (red curves). The left column shows the combined errors for all 280 earthquake scenarios. The middle column presents the errors obtained from training the neural networks with 224 randomly chosen events. The right column displays the results obtained from the application of the trained neural networks to the remaining 56 events.



**Figure 5.16:** *PreSEIS* alert classifications for all 280 earthquakes based on the present EEW network of 10 stations (left) and on the planned expanded network of 20 stations (right).

more reliable source parameter estimates when the EEW method *PreSEIS* is applied.

As described in Chapter 4.2, *PreSEIS* additionally uses an empirical attenuation law derived from the simulated dataset to estimate the seismic intensity  $I$  at a pre-defined site such as ISTAN (Böse, 2006). By defining an *alert threshold* at  $I = 5.5$ , each event can be classified as potentially damaging ( $I \geq 5.5$ ) or non-damaging ( $I < 5.5$ ) at the user site (Böse, 2006). The intensity at site ISTAN of each event is estimated at each time step; whenever this value is equal or larger than 5.5 an alert is issued. The alerts are classified as following:

A *correct alert* is defined as the correct recognition of an earthquake causing  $I \geq 5.5$ . A *missed alert* means that an earthquake which would indeed exceed the intensity threshold has not triggered an alert. A *correct all-clear* is issued when an earthquake has correctly been identified as non-damaging, and a *false alert* describes a wrongly issued alert for a non-damaging event (Böse, 2006).

Figure 5.16 shows the *PreSEIS* alert classifications for all 280 earthquake scenarios for the present network of 10 stations and the expanded one of 20 stations. On average, the present EEW network of 10 stations produces 35.3% correct alerts, 3.3% missed alerts, 59.5% correct all-clear, and 1.9% false alerts. The expanded network of 20 stations produces 34.2% correct alerts, 4.4% missed alerts 59.9% correct all-clear, and 1.6% false alerts, on average. Similar to the magnitude estimates of the test dataset (Figure 5.15), the alert classification using the expanded network does not differ significantly from the one obtained from the original network. This stresses, on the one hand, the great advantage of *PreSEIS* to perform equally well for different input data configurations. On the other hand, they indicate that the present EEW network for Istanbul already presents a highly efficient network configuration for the purpose of early warning.

### 5.3.1 Amplitude Threshold-Based Alarms

As described in Chapter 1.2, the early warning algorithm implemented in *IERREWS* is based on the exceedance of specific amplitude thresholds (Erdik et al., 2003b; Alcik et al., 2009). From the incoming acceleration data, peak ground acceleration is determined and compared to the threshold values. Whenever *PGA* exceeds the thresholds, an alarm is issued (See Chapter 1.2). I will

apply this algorithm to the simulated dataset using both the original and the expanded network configurations. The algorithm I applied is thus the following:

When at any station on any channel *PGA* exceeds a threshold of 0.02g, it is considered a vote. Whenever votes from at least 3 stations within 5 s can be obtained the alarm 1 will be declared. After alarm 1, alarm 2 will be issued whenever the amplitudes exceed 0.05g at any three stations within 5 s. After alarm 2, alarm 3 will be declared whenever the amplitudes exceed 0.1g at any three stations within 5 s (Erdik et al., 2003b).

The performance of such a threshold-based EEW system is highly dependent on the selected amplitude threshold values. If the thresholds are chosen too small, the number of false alarms is likely to increase since also non-damaging earthquakes might trigger an alarm. If the thresholds are chosen too high, alarms are likely to be missed. The in *IERREWS* defined threshold values were chosen based on expert judgment but have not been quantitatively related to the ground shaking in Istanbul expected from the earthquakes triggering one of the alarm levels. This has been investigated by Oth et al.<sup>3</sup>. The authors evaluated the EEW network in terms of the amplitude thresholds and the location of the seismic stations. They found out that the current system is highly sensitive to rather small ground motion amplitudes. By classifying the earthquake scenarios into warning classes depending on the peak ground shaking they will cause at a user site in Istanbul, and by systematically changing the threshold values of the three alarm levels, the early warning performance of the network could be optimized. The threshold values that lead to the best performance (largest number of correctly classified events with best available warning times) are 0.03g for alarm 1, 0.12g for alarm 2, and 0.16g for alarm 3. Additionally, Oth et al.<sup>3</sup> simulated an expansion of the current network by six more onshore stations and three ocean-bottom seismometers. A micro-genetic algorithm is used to find the best potential station locations around Istanbul in addition to the current system that yield the best early warning performance. The study shows that there is no significant improvement in the early warning performance if more than 10 stations are used. In addition, the analysis indicates that the current locations of the 10 early warning stations are already almost optimum (Oth et al.<sup>3</sup>).

To compare the early warning performances in this study I will also apply the optimized threshold values to the current system of 10 stations.

Table 5.1 summarizes the number of events that issue the respective alarms for the original and the planned expanded network as well as for the original network using the optimized thresholds. The expanded network of 20 stations yields an increase of all three alarms. It declares 7% more alarms 1, 6% more alarms 2, and 4% more alarms 3 than the network of 10 stations and on average 0.9 s faster. This leads to an average increase of warning times by 2.2 s, whereas the warning time is defined as the time between the triggering of the alarm and the S-wave arrival at ISTAN. When applying the optimized amplitude thresholds to the 10 stations, the number of issued alarms decreases by 9% for alarm 1 and by 54% for both alarms 2 and 3. The time necessary to declare the alarms, on the contrary, increases, leading to considerably smaller warning

---

<sup>3</sup>Oth, A., Böse, M., Wenzel, F., Köhler, N., and M. Erdik: Optimizing seismic networks for earthquake early warning - the case of Istanbul (Turkey). *Submitted to J. Geophys. Res.*, 2010

	<i>10 stations</i>	<i>20 stations</i>	<i>10 stations opt. thresholds</i>
<b>Alarm 1</b>	209 in $\bar{t} = 9.7$ s	227 in $\bar{t} = 8.7$ s	190 in $\bar{t} = 10.1$ s
<b>Alarm 2</b>	169 in $\bar{t} = 10.9$ s	186 in $\bar{t} = 9.8$ s	78 in $\bar{t} = 12.6$ s
<b>Alarm 3</b>	104 in $\bar{t} = 11.8$ s	115 in $\bar{t} = 11.4$ s	48 in $\bar{t} = 14.2$ s
<b>Average warning time</b>	4.9 s after alarm 1 3.8 s after alarm 2 2.3 s after alarm 3	7.3 s after alarm 1 6.2 s after alarm 2 4.2 s after alarm 3	4.5 s after alarm 1 1.8 s after alarm 2 0.3 s after alarm 3

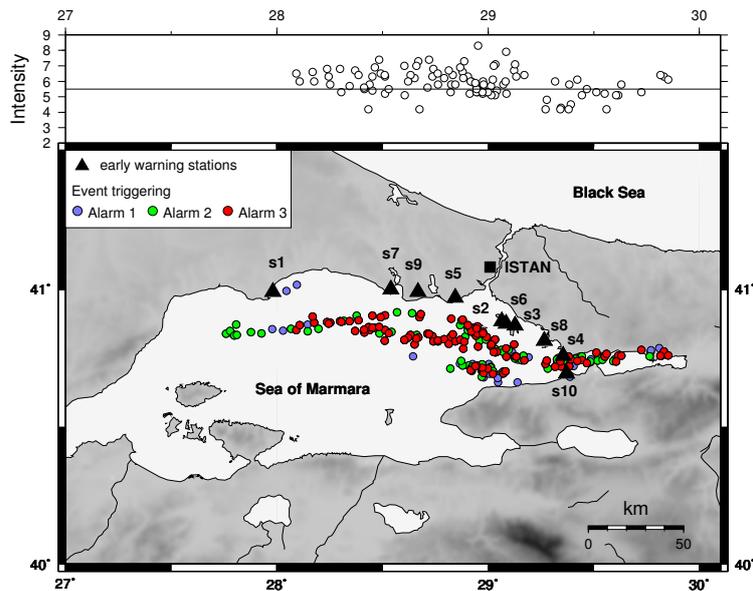
**Table 5.1:** Number of earthquake scenarios causing alarm 1, 2, and 3, respectively, together with the average time  $\bar{t}$  that is needed to issue the alarms with respect to the earthquake occurrence times. The left and the middle column give the alarms obtained from the present network configuration of 10 stations and the expanded one of 20 stations. The right column displays the results obtained from the optimized trigger thresholds. The average warning times are defined as the time differences between the triggering of the alarms and the S-wave arrivals at ISTAN.

times, especially for alarm level 2 and 3.

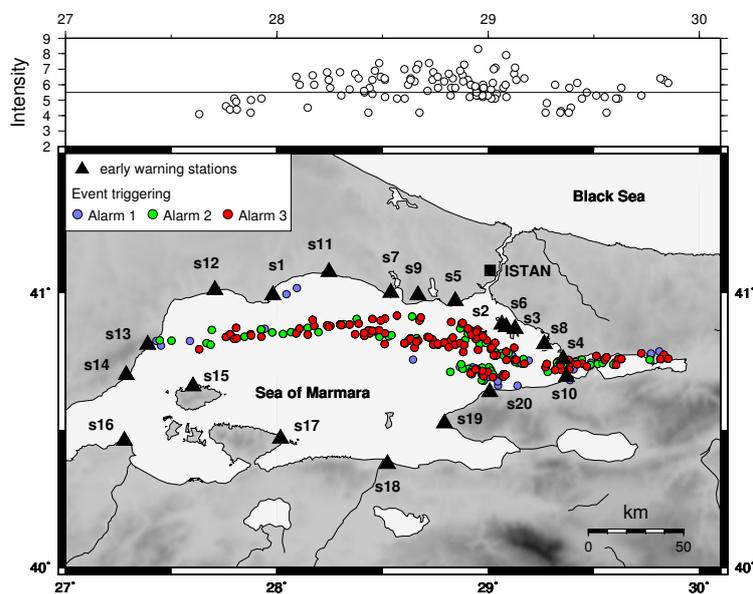
Figures 5.17, 5.18, and 5.19 show the distributions of the events that cause the alarms described in Table 5.1, whereas not all events triggered all three alarm levels. The seismic intensity at site ISTAN of these earthquakes is plotted above the distribution maps. The majority of the additional alarms in the expanded network (5.18) is caused by earthquakes located further west in the Sea of Marmara which show systematically lower intensities at ISTAN. The earthquake distribution derived from the optimized trigger thresholds (Figure 5.19) naturally reflects that the remaining alarm-triggering events are stronger and cause higher ground shakings at ISTAN compared to Figure 5.17, due to the fact that the optimized thresholds are set to higher values.

The histograms in Figure 5.20 show that, even with 20 stations, the majority of votes leading to the respective alarms occur at stations s1-s10 which are the stations from the original network. The only additional stations contributing noteworthy to the declaration of alarms are stations s11-s13 and s20, whereas station s20 has the greatest impact due to its proximity to the simulated earthquake locations. The distribution of votes at the 10 stations is independent from the applied threshold values (Figure 5.21), except for the fact that the higher, optimized thresholds result in less votes, of course.

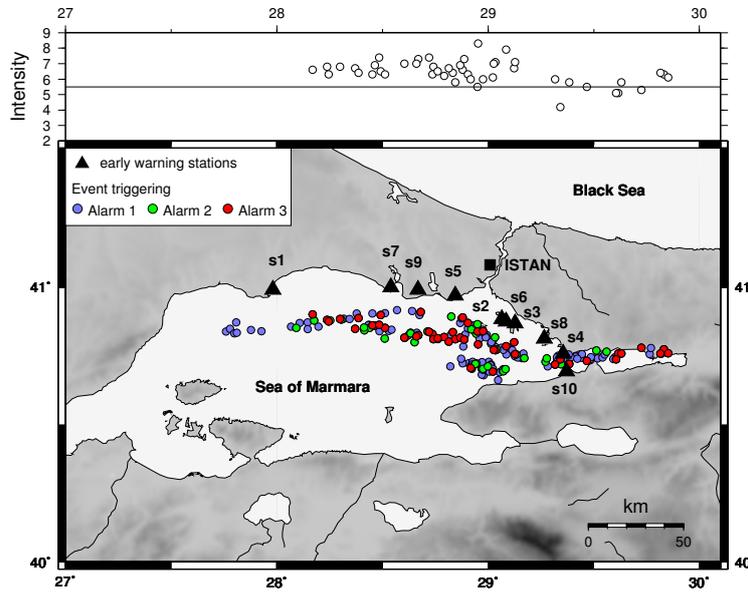
Now where we have analyzed the distribution of alarms it is necessary to investigate if the declared alarms are correct in terms of potentially damaging ground shaking at ISTAN. Figure 5.18 has already indicated that, if the thresholds are set to equal values at all stations, an expansion of the EEW network to regional scale would not make a significant contribution to earthquake early warning for Istanbul. Although the expanded network triggers more alarms, most of these alarms occur for events located further away from the city of Istanbul. They do not cause seismic



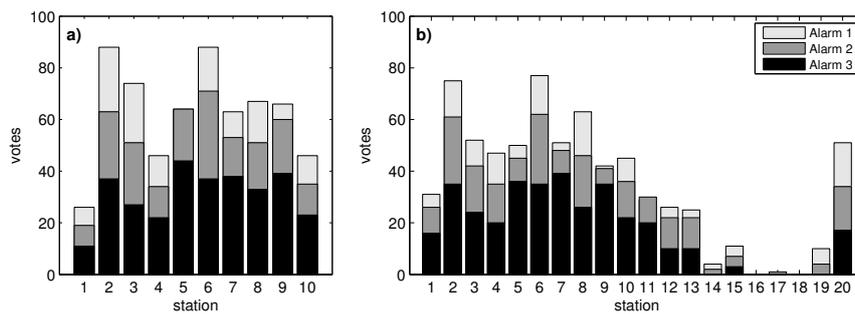
**Figure 5.17:** Distribution of earthquake scenarios that trigger alarm 1 (blue circles), both alarm 1 and alarm 2 (green circles), and all 3 alarms (red circles) in the case of the present EEW network of 10 stations. The seismic intensities at site ISTAN caused by the respective events are plotted above the map. The solid line indicates the previously defined intensity level of  $I = 5.5$ .



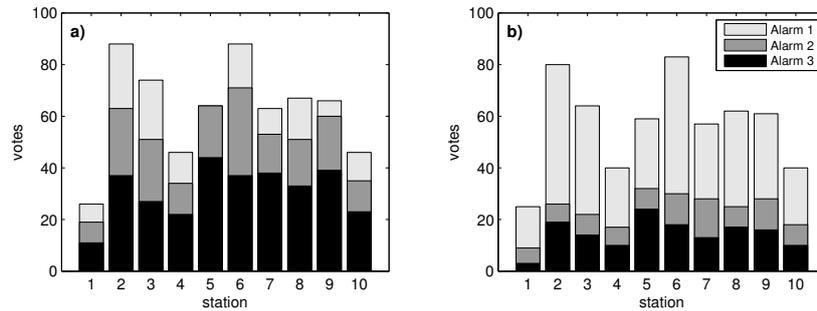
**Figure 5.18:** Distribution of earthquake scenarios that trigger alarm 1 (blue circles), both alarm 1 and alarm 2 (green circles), and all 3 alarms (red circles) in the case of the expanded EEW network of 20 stations. The seismic intensities at site ISTAN caused by the respective events are plotted above the map. The solid line indicates the previously defined intensity level of  $I = 5.5$ .



**Figure 5.19:** Distribution of earthquake scenarios that trigger alarm 1 (blue circles), both alarm 1 and alarm 2 (green circles), and all 3 alarms (red circles) in the case of the present EEW network of 10 stations with the optimized amplitude thresholds. The seismic intensities at site ISTAN caused by the respective events are plotted above the map. The solid line indicates the previously defined intensity level of  $I = 5.5$ .



**Figure 5.20:** Distribution of votes leading to the three alarms at each station of the current network (a) and of the expanded network (b).



**Figure 5.21:** Distribution of votes leading to the three alarms at each station of the current network after the application of the current amplitude thresholds (a) and the optimized ones (b).

intensities at ISTAN that are above our threshold of 5.5 and can therefore not be considered as potentially damaging. The additionally issued alarms in the case of 20 stations are thus likely false alarms.

Table 5.2 confirms this hypothesis. Listed are the alarm classifications for the 280 earthquake scenarios with respect to the user site ISTAN. The final alarm is considered as alarm level 3. The network of 20 stations is not capable to increase the number of correct alarms, although it neither misses more alarms than with only 10 stations. It produces ten more false alarms, however, due to the earthquakes located far west in the Sea of Marmara.

The application of the optimized alarm thresholds significantly reduces the number of false alarms and predicts correctly the damage potential of an increased number of events. The higher threshold values, however, also increase the number of missed alarms. This might be simply due to the fact that the simulated dataset used in this study is considerably larger than the one used by Oth et al., which consists of 180 earthquake scenarios. A larger number of moderate-sized earthquakes apparently exist in my database that cause amplitudes slightly below the optimized threshold values, although they still have damage potential. On the other hand, the chosen intensity level of 5.5 (Böse, 2006) considerably influences the statistics. If the intensity level would only be minimal higher, a number of earthquakes would not be classified as damaging anymore and the number of missed alarms would thus be reduced. Finally, the fact that the alarm classification of Oth et al. is based on peak ground acceleration at the user site instead of intensity is another reason (maybe the most important reason) why the optimized threshold values do not improve my alarm classification in all aspects.

Finally, the warning times for the events that lead to an alarm declaration need to be considered. Figure 5.22 illustrates the warning times resulting from the original 10 stations and the expanded network of 20 stations. Plotted are the warning times obtained from the declaration of all three alarms. The dotted lines indicate the damaging events that generate  $I \geq 5.5$  at ISTAN. As discussed before, both network configurations lead to a considerable number of false alarms. It is obvious that the warning times for the majority of events do not change significantly when

	<i>10 stations</i>	<i>20 stations</i>	<i>10 stations (opt. thresholds)</i>
<i>Correct alarms</i>	71	71	43
<i>Missed alarms</i>	2	2	30
<i>Correct all-clear</i>	174	164	202
<i>False alarms</i>	33	43	5

**Table 5.2:** Alarm classification of all 280 earthquake scenarios on the basis of alarm 3 of the threshold system.

using 20 stations. For the actual correct alarms, the maximum warning times are in the range of about 10 s which is in good correlation to the 8 s stated by Erdik et al. (2003b). Only few earthquakes show negative warning times.

For the events of numbers 100 to 150, however, the warning times increase drastically up to almost 30 s using the expanded network. These are the events simulated around Segment 3 (Figure 2.8), which is the westernmost segment. Most of these events are located too far away from Istanbul to be damaging anymore; they represent false alarms.

Figure 5.23 shows the same comparison but this time for the network of 10 stations using the original and the optimized threshold values. Figure 5.23a is identical to Figure 5.22a, only with a smaller range on the time axis.

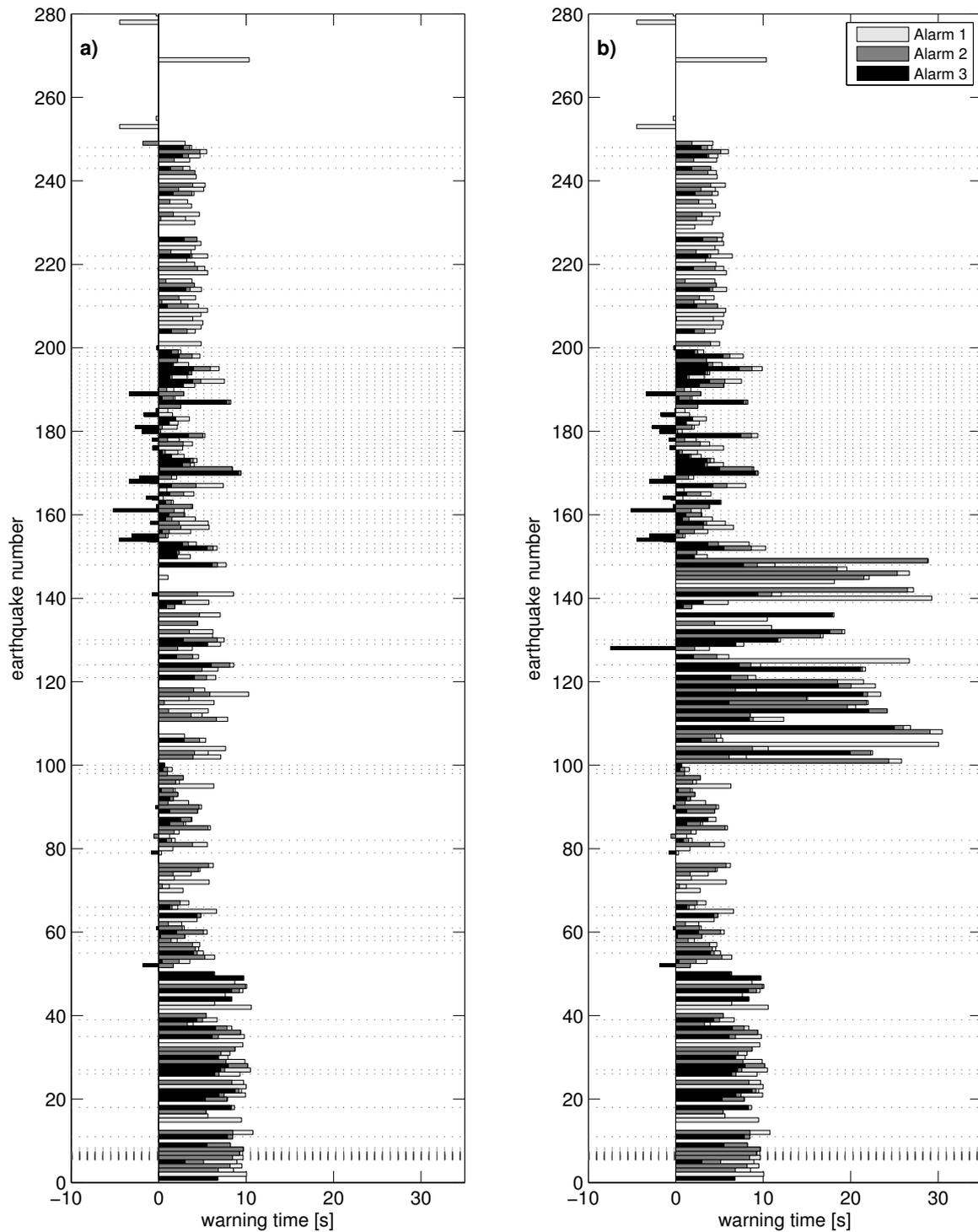
As previously discussed, the optimized threshold values trigger much less votes for alarms 2 and 3 compared to the original thresholds. This is the reason why Table 5.2 reveals more missed alarms than for the original thresholds, since Table 5.2 defines the final alarm as alarm level 3. In addition, a larger number of negative warning times occur for alarm levels 2 and 3 for events on Segment 4 (event-no. 150-200).

The overall warning times available after alarm level 1 do not depend noteworthy on the original or the optimized thresholds, since the thresholds for alarm 1 only differ by 0.01g.

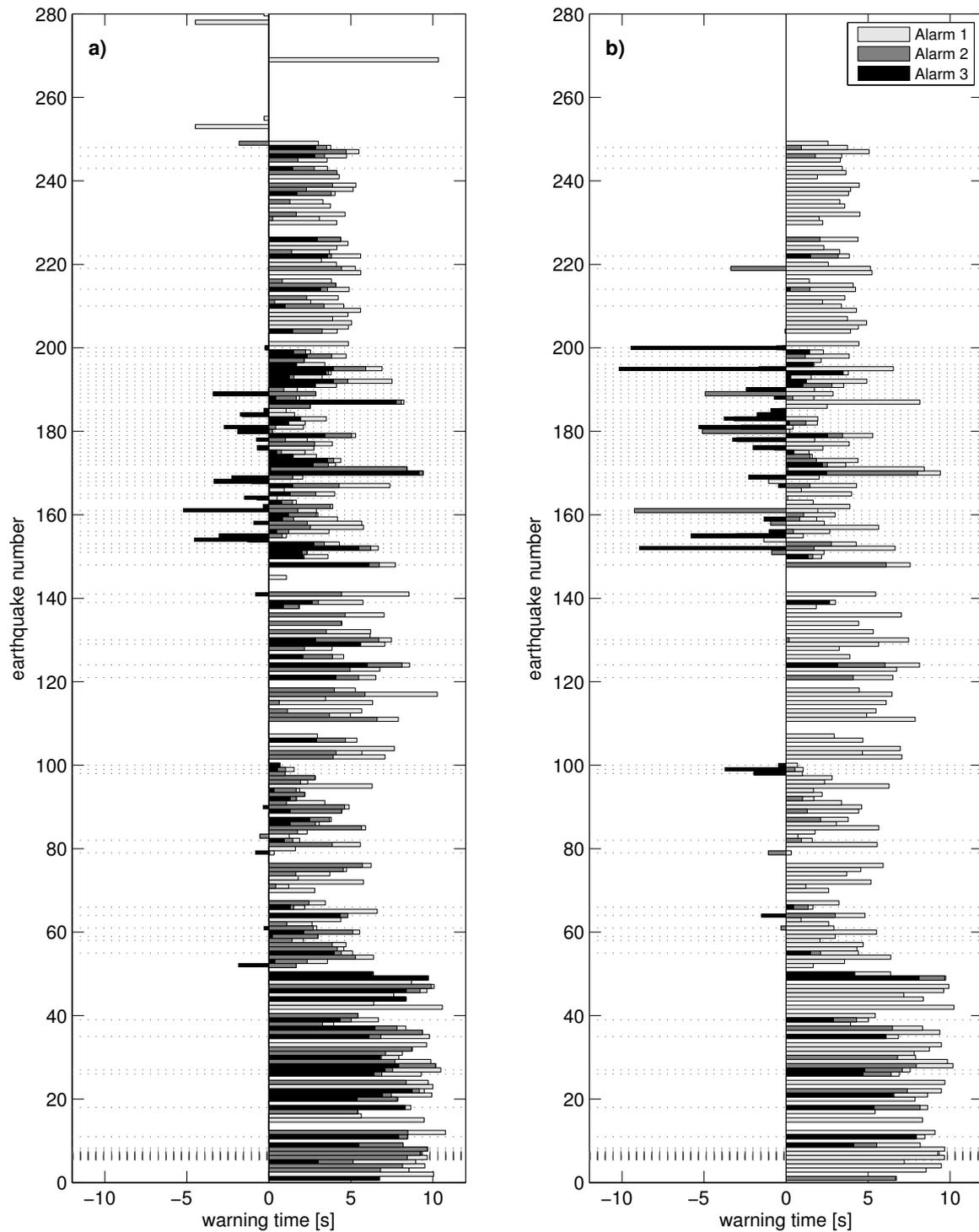
### 5.3.2 Summary

The presented performance evaluation of the earthquake early warning network in Istanbul reveals a high quality performance of the current network consisting of 10 early warning stations. When using an early warning algorithm like *PreSEIS* that integrates the information from all stations at all time steps, an expansion of the network to the Marmara region could even improve the source parameters estimates by up to 22%.

When using the threshold based alarm system under the assumption that the threshold values would be identical at all stations, an expansion of the network to the Marmara region would not contribute significantly to the reliability of early warning for Istanbul. This was expected since earthquakes occurring further away from Istanbul would still trigger alarms even though they are too far away to be threatening. To avoid the false alarms it would be necessary to adapt the threshold values at the additional stations. Another possibility would be to use the additional stations exclusively to issue warnings to cities surrounding the northwestern and southern Sea of Marmara.



**Figure 5.22:** Available warning times for all 280 earthquake scenarios using the *IERREWS* threshold system with 10 stations (a) and with 20 stations (b). The warning times are defined for all three alarm levels. The dotted lines indicate the earthquake scenarios that cause  $I \geq 5.5$  at ISTAN.



**Figure 5.23:** Available warning times for all 280 earthquake scenarios using the original network of 10 stations with the original *IERREWS* thresholds (a) and the optimized thresholds (b). The dotted lines indicate the earthquake scenarios that cause  $I \geq 5.5$  at ISTAN.

It must be kept in mind, however, that this study only represents a theoretical performance evaluation. Modifying the characteristics of the earthquake database, the alarm thresholds and the intensity threshold used to classify events as damaging and non-damaging will result in different early warning performances. The comparison with the application of the optimized alarm thresholds obtained by Oth et al. is also of limited meaning since their classification whether an earthquake would be damaging or not is based on the exceedance of *PGA* at the user site, whereas mine is based on the intensity of ground motion.



# Chapter 6

## Summary and Outlook

Earthquake early warning (EEW) systems are of great value for mitigating co-seismic risk. The physical basis for EEW is the difference in propagation speed of the destructive shear and surface waves of an occurring earthquake and electromagnetic signals used to transmit information on ground motion and warnings. Depending on the distance of a strong earthquake from a user site such as an urban area or industrial facility, warnings may be issued a few seconds up to about one minute before the strong ground shaking takes place at this site. Despite these short warning times the number of possible applications of EEW systems is large. The most common application is the initiation of automatic actions such as the shutdown of industrial facilities, gas distribution, or computers, or the stoppage or braking of high speed trains, for example.

Earthquake early warning is a main component of real-time information systems. Besides the warning shortly before the occurrence of strong ground shaking, real-time information systems provide information about the ground motion distribution and estimates of structural damages and human losses within some time after the earthquake stroke. This is crucial information not only for rescue operations but also for future improvement of disaster mitigation strategies.

One contribution to earthquake preparedness in a hazardous region is the deterministic estimation of ground motion distribution for a specific example earthquake. This earthquake scenario must be as representative as possible so that likely effects of ground motion on infrastructure, buildings, or services can be quantified. The information inferred from such a scenario earthquake can assist in the establishment of emergency plans or building codes.

One part of this thesis deals with the simulation of seismic ground motion for specific example earthquakes. The final ground motion can be characterized as a combination of linear filters representing the earthquake source, the path effects, and the local site effects that modify the traversing seismic waves. Chapter 2 gives an introduction into this. The ground motion simulations in this thesis are accomplished with *FINSIM*, a stochastic simulation technique for finite faults (Beresnev and Atkinson, 1997, 1998). *FINSIM* is an extension of the stochastic point source method by Boore (1983, 2003) and is described in Chapter 2.5. It is a simple and robust approach that combines functional descriptions of the ground motion's amplitude spectrum with a random phase spectrum. The final waveform at a specific site obtained for a specific earth-

quake is represented by the sum of contributions from all subfault elements on the defined fault. I use a modified version of *FINSIM* established by Böse (2006) who included the simulation of compressional waves as the original method only accounts for shear waves.

In Chapter 3 the *FINSIM* technique is used to simulate a historic earthquake that hit the Turkish city Istanbul in the year of 1509. With a magnitude of about 7.3 it was one of the largest and most destructive earthquakes that struck Istanbul during the last 500 years. The mega-city Istanbul is a major focus in this thesis due to its extremely high seismic risk. It is closely located to the western continuation of the North Anatolian Fault in the Sea of Marmara whose segments partially show distances of 10 to 20 km to the mega-city. According to Parsons (2004), a 35-70% probability exists that a  $M \geq 7$  earthquake will affect Istanbul in the years between 2004 and 2034. The occurrence of such a large earthquake on one of the fault segments in the Sea of Marmara will likely cause immense damages and large numbers of fatalities. The 1999  $M_w$  7.4 Kocaeli (Izmit) earthquake, for example, had disastrous impact on Istanbul; more than 15,000 people were killed and more than 77,000 households were heavily damaged.

The purpose of the in Chapter 3 presented simulation of the historic 1509 Istanbul earthquake is the demonstration of the variability of ground motions resulting from different earthquake models and prediction techniques. Compared are the distributions of peak ground acceleration and seismic intensity in the metropolitan area of the today's city of Istanbul simulated for three different hypocentre locations of the historic earthquake. The distribution of peak ground acceleration is additionally predicted using attenuation relationships and compared with the *FINSIM* simulations. The results show significant differences in ground motion distribution depending on the hypocentre location and prediction technique. Different are not only the amplitudes of ground motion but also their spatial distribution. Maximum values of peak ground acceleration up to  $457 \text{ cm/s}^2$  and of intensities up to 9.1 can be obtained for the European part of Istanbul. The assumption of a finite fault leads to directivity effects which strongly influence the resulting ground motions. To demonstrate the impact of this ground motion variability on the resulting damages the building damages in Istanbul arising from the different scenarios of the 1509 earthquake were predicted and compared.

The building damage estimates are established on the basis of the *KOERI Loss Estimation Methodology (KOERILoss)* developed at the *Kandilli Observatory and Earthquake Research Institute (KOERI)* at the Bogazici University in Istanbul. The fundamental information for the damage predictions is the classification of the building stock in Istanbul in terms of the buildings' vulnerabilities toward earthquakes. The establishment of the *spectral displacement demand* of a given structure, determined for a given earthquake, in combination with *fragility curves* allow for calculating the damage probabilities for certain damage classes (Appendix C). Chapter 3.2 presents the predicted building damages for the 1509 earthquake scenarios. More than 4000 buildings are predicted to be completely destroyed, mostly located on the European part of Istanbul around the old town. The study stresses that scenario earthquakes can represent only one possible realization of fault rupture and distribution of ground motion and damages.

The task of an earthquake early warning system is the provision of reliable warnings within the shortest amount of time. The *warning time*, defined as the time between the declared alarm

and the arrival of destructive S- and surface waves at the user site, depends not only on the distances between the source, seismic network, and user site, but also on the time that is necessary for processing the observed ground motion data. With a suitable seismic network, an earthquake can be located and its magnitude can be estimated reliably. This approach is called *regional* early warning. *On-site* early warning, in contrast, uses the observed ground motion at a single site to predict the impending strong shaking at the same site. This approach is less reliable but faster than regional early warning since the latter requires the arrivals of seismic waves at several stations. Both techniques are applied in various EEW systems and methodologies, either in operation or under development. Chapter 4 gives an overview on the recent state of the art.

This thesis focusses on a novel approach to earthquake early warning. The methodology *PreSEIS* (Pre-SEISmic shaking) developed by Böse (2006) is a combination of regional and on-site early warning. It estimates the earthquake hypocentre location, the magnitude, and the final expansion of the fault rupture on the basis of ground motion information available from a seismic network. The hypocentre location is estimated from the P-wave arrival time differences at the respective stations, whereas information on not-yet-triggered stations is used to confine the possible source locations. The magnitude and the rupture expansion are estimated from the *cumulative absolute velocity*, defined as the integral of absolute acceleration over a certain time window. As soon as the P-wave reaches the first seismic station the source parameter estimation process starts. The estimates are continuously updated at regular time steps when more ground motion information becomes available.

*PreSEIS* makes use of artificial neural networks which map the ground motion observations onto likely earthquake source parameters. *PreSEIS* uses *two-layer feed forward* neural networks that consist of three layers of processing units (neurons) - one input, one hidden, and one output layer. The neural networks have to be trained with a dataset of example earthquakes with known source parameters. After successful training the networks can be applied to unknown data that follow the statistical patterns from the training earthquakes. For a robust performance of the neural networks, the training dataset must be sufficiently large and the numbers of neurons in each neural network must be adapted to the number of input and output parameters. The current version of *PreSEIS* requires that all stations in the seismic network are fully operational and do not fail. Chapter 4.2 gives a detailed description of the *PreSEIS* methodology.

The *PreSEIS* methodology has been developed on the example of Istanbul. After the devastating 1999 Kocaeli and Düzce earthquakes, *KOERI* installed the *Istanbul Rapid Response and Early Warning System (IERREWS)*. It consists of a dense network of strong motion instruments installed in the metropolitan area of Istanbul. The rapid response part consists of 100 stations that are used to provide rapid shake maps and damage and loss estimates. The early warning part consists of ten strong motion stations installed along the northern shoreline of the Sea of Marmara, between Istanbul and the likely source region of future strong earthquakes. The alarm system is based on the exceedance of ground motion amplitude thresholds at the stations (Erdik et al., 2003b). Chapter 1.2 describes the system in more detail.

Due to the lack of strong ground motion observations at the early warning stations, Böse (2006) used *FINSIM* to simulate a large set of 280 earthquake scenarios in the Sea of Marmara for the application of the early warning method *PreSEIS*. The ground motions are simulated for the sites

of the ten *IERREWS* early warning stations. Chapter 2.6 presents the dataset. The dataset has mainly been established by taking default *FINSIM* input parameters leading to some systematic discrepancies between simulations and observations. To improve the suitability of the simulated dataset for EEW purposes, I simulated a number of observed earthquake records with *FINSIM* and optimized the input parameters by visually comparing the simulations with the observations. The visual inspection concentrated mainly on the shapes of the envelopes and the P-wave onsets. I then re-calculated the synthetic dataset of the 280 earthquakes with the optimized *FINSIM* input parameters. The optimized ground motion records lead to an improved *PreSEIS* performance with up to 34% more accurate source parameter estimates. The optimization of the simulated dataset is described in Chapter 2.7.

The early warning method *PreSEIS* has successfully been applied to the simulated dataset (Böse, 2006). However, the application to synthetic data is of limited meaning. The full functionality of *PreSEIS* can only be revealed using real earthquake data. Chapter 5.1 and 5.2 present the first applications of *PreSEIS* to real earthquake observations. In Chapter 5.1, the method is applied to a set of 74 earthquakes from Southern California, recorded at 15 stations of the Southern California Seismic Network. The dataset does not consist exclusively of real data; missing observations are replaced with predictions. In this study, the *PreSEIS* magnitude estimates base not on the cumulative absolute velocity but on the cumulative envelopes of ground motion. Envelope attenuation relationships established by Cua (2005) allow to predict the missing observations. The neural networks are trained with all 74 earthquakes; the performance is then demonstrated by three example events representing different characteristic settings. The errors of the source parameter estimates clearly decrease with ongoing time. Once ground motion information is available from two stations the magnitude estimates have accuracies between 90 and 100%. The study shows a stable and robust *PreSEIS* performance. The neural networks prove to be suitable and attractive tools for EEW applications. They are tolerant toward various types of input data which is of great advantage in transferring the method to different regions.

In Chapter 5.2, *PreSEIS* is applied to a set of 69 earthquakes from Japan recorded at ten stations of the Kyoshin-Net. As for the synthetic data the magnitude estimates are based on the cumulative absolute velocity. The neural networks are trained with 66 of the 69 earthquakes. The source parameter estimates are then tested on the remaining three earthquakes which are unknown to the neural networks. The dataset used in this study consists exclusively of real observations. *PreSEIS* shows again a stable performance. With ongoing time the average location errors can be reduced to about 13 km, depending on the ground motion component. The mean magnitude errors can be reduced to about  $\pm 0.3$  magnitude units. The application to the three unknown events results in location and magnitude estimates of high quality. The largest location errors of about 50 km are obtained for an event where no training examples were available in its direct vicinity. The location errors, however, also include the source depths of the earthquakes. The average location errors obtained for the other two example earthquakes are around 19 km and 7 km. The magnitudes of the example earthquakes can be estimated accurately; *PreSEIS* is able to estimate the overall magnitude level of the events within the first few seconds of available ground motion data.

The performance evaluations of *PreSEIS* show that the method is highly suitable for early warn-

ing applications. In Chapter 5.3 it is thus used as a benchmark system to evaluate the performance of the present early warning system in Istanbul. Since it has recently been discussed to expand the early warning system to the Marmara region by installing an additional set of 10 stations around the Sea of Marmara, I compare the theoretical performances of both the present and the expanded network. For this, I simulated the ground motion from the 280 earthquake scenarios for the sites of the additionally planned stations and applied *PreSEIS* to both station configurations. By using the expanded network, the source parameter estimates can be improved by up to 22%. By calculating the seismic intensity at a user site in Istanbul for each of the 280 earthquake scenarios, an alarm classification can be established, whereas damaging earthquakes are classified as ones where the intensity at the user site is equal or larger than 5.5. By estimating the intensities at each time step, *PreEIS* can be used to issue alarms. The performance shows that the alarm classifications do not differ significantly for the present and the expanded early warning network.

In the second part of Chapter 5.3, the simulated earthquake scenarios are used to test the threshold-based early warning algorithm implemented in Istanbul. The results clearly show that the present station configuration together with the present amplitude thresholds shows the best performance. An expansion of the network would increase the number of false alarms due to the fact that earthquakes located in the western part of the Sea of Marmara would trigger alarms although the events are too far away to be damaging. This presumes, however, that the thresholds at the additional stations would be set to identical values than at the original ten stations. For the present EEW system I obtain warning times for Istanbul of up to 10 s.

## Outlook

Despite the successful performance tests of *PreSEIS*, the method still suffers from the limitation to a fully operational seismic network. In reality, seismic stations will fail from time to time and may not be able to contribute in the case of an occurring earthquake. It is necessary to integrate information on non-functional stations into *PreSEIS*. One approach might be to train the neural networks for all possible combinations of failing stations. This would lead to time consuming training processes, but might be effective.

The other main problem that needs to be addressed before a possible implementation of the method is the availability of suitable training data. Combining observations with synthetics might be a possible solution, as this thesis demonstrates. The combination of strong and weak motion data might be another option to enlarge the training datasets. As soon as *PreSEIS* can handle incomplete datasets, however, the problem of finding suitable datasets will be reduced.

The third step necessary to include would be a reliable, automatic P-wave detection. Since other, already implemented, early warning methods also use automatic P-wave detection and arrival time differences between the different seismic stations to locate an earthquake I consider this problem as fairly resolvable.

Furthermore, the application of *PreSEIS* to the earthquake data from Japan (Chapter 5.2) sug-

gests that a parallel incorporation of all three ground motion channels would lead to the most optimum performance. This would require modifications of the *PreSEIS* code that have not yet been addressed in this work.

# Appendix A

## Simulation of Seismic Ground Motion

### A.1 Simulated Earthquake Dataset

Table A.1 displays the *FINSIM* input parameters (Chapter 2.5) that are used to generate the 280 earthquake scenarios described in Chapter 2.6.

Table A.2 describes the fault geometries of Segments 1 to 5 from Chapter 2.6 (Figure 2.8). Along each of these segments a number of 50 earthquake scenarios are simulated (Figure 2.7).

Tables A.3 to A.8 describe the hypocentre locations, moment magnitudes, and rupture lengths of all 280 simulated earthquake scenarios (Figures 2.7, 2.9, and 2.10), whereas Segment 6 corresponds to a number of 30 events distributed randomly over the region.

<b>Source</b>		
Fault length:	$L$	see Table A.3 - A.8
Fault width:	$W$	1.0 - 16.4 km
Fault strike:	$\phi$	265° - 300°
Fault dip:	$\delta$	90°
Fault depth:	$Z_r$	5.0 km
Moment magnitude:	$M_w$	4.5 - 7.5
High-freq. radiation strength:	$sfact$	0.9 - 1.3
Stress drop:	$\Delta\sigma$	$6 \cdot 10^6 - 13 \cdot 10^6 \text{ N/m}^2$
Density of medium:	$\rho$	3000 kg/m <sup>3</sup>
P-wave velocity:	$\alpha$	5716 m/s
S-wave velocity:	$\beta$	3300 m/s
<b>Path</b>		
Trilinear model for geometrical spreading:	$rg1$	2 km
	$rg2$	400 km
	$pow1$	-1.5
	$pow2$	-0.8
	$pow3$	-0.7
Inelastic attenuation:	$Q_0$	50 (Böse, 2006) 70 (this work)
	$\eta$	1.09
Duration model:	$rmin$	10 km
	$rd1$	70 km
	$rd2$	130 km
	$durmin$ (B,C,D)	2.0 s, 2.2 s, 2.4 s
	$b1$ (B,C,D)	0.25, 0.3, 0.4
	$b2$	0.1
	$b3$	0.04
<b>Site</b>		
High-frequency diminution:	$\kappa$ (B,C,D)	0.035, 0.04, 0.045 (Böse, 2006) 0.065, 0.07, 0.075 (this work)
Frequency-dependent amplification:	<i>NEHRP B, C, D</i>	see Table 2.2

**Table A.1:** *FINSIM* input parameters for the 280 simulated earthquakes according to Böse (2006). The parameters  $Q_0$  and  $\kappa$  have been modified in this work (Chapter 2.7).

	<i>SEGMENT [°N / °E]</i>				
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
<i>Fault start</i>	40.77 / 29.91	40.73 / 29.20	40.91 / 28.79	40.71 / 29.52	40.68 / 29.15
<i>Fault end</i>	40.73 / 29.20	40.91 / 28.79	40.83 / 27.61	40.90 / 28.00	40.74 / 28.80
<i>Strike</i>	265°	300°	265°	280°	282°
<i>Dip</i>	90	90	90	90	90
<i>Length [km]</i>	60	40	100	130	30
<i>Depth [km]</i>	5	5	5	5	5

**Table A.2:** Description of Segment 1 to 5 (Figure 2.8) after Böse (2006).

No.	Latitude [ $^{\circ}$ N]	Longitude [ $^{\circ}$ E]	Depth [km]	$M_w$	Rupture length [km]
1	40.77	29.51	10.48	6.9	43.79
2	40.72	29.40	5.44	5.1	3.26
3	40.75	29.52	6.20	6.2	16.75
4	40.74	29.52	7.34	6.0	11.98
5	40.74	29.44	7.12	6.5	24.21
6	40.74	29.54	5.78	6.0	12.11
7	40.75	29.60	9.14	6.4	22.95
8	40.74	29.48	10.13	6.1	14.79
9	40.76	29.62	7.64	6.8	40.30
10	40.76	29.76	5.46	5.1	3.57
11	40.73	29.47	13.43	6.9	46.32
12	40.74	29.59	7.04	6.4	22.17
13	40.77	29.79	9.61	6.0	11.55
14	40.76	29.73	5.75	5.0	1.00
15	40.75	29.33	5.63	5.6	7.34
16	40.77	29.77	6.61	6.0	12.93
17	40.74	29.33	9.82	6.7	32.66
18	40.72	29.38	6.91	7.0	48.47
19	40.78	29.83	8.01	5.6	6.60
20	40.76	29.55	7.88	7.0	49.16
21	40.74	29.61	12.31	6.7	33.43
22	40.73	29.27	9.85	5.8	9.13
23	40.76	29.36	5.12	4.5	1.00
24	40.75	29.42	10.53	6.1	13.77
25	40.77	29.58	5.25	4.7	1.00
26	40.78	29.83	10.39	6.9	42.11
27	40.76	29.82	6.97	7.0	52.31
28	40.72	29.34	6.78	6.2	16.14
29	40.74	29.33	6.00	5.9	11.08
30	40.77	29.56	8.87	6.3	19.69
31	40.76	29.76	9.66	6.3	18.05
32	40.76	29.44	6.67	6.1	13.92
33	40.75	29.49	8.28	5.7	8.08
34	40.80	29.85	9.62	5.7	8.17
35	40.76	29.63	13.52	6.9	47.49
36	40.71	29.28	7.47	5.3	4.22
37	40.74	29.28	10.56	6.3	17.86
38	40.75	29.41	10.39	6.9	42.08
39	40.76	29.85	11.18	7.1	57.56
40	40.74	29.36	10.28	6.8	40.09
41	40.75	29.73	5.75	4.6	1.00
42	40.75	29.35	5.25	4.5	1.00
43	40.79	29.81	6.96	5.7	8.03
44	40.75	29.39	10.36	6.9	41.57
45	40.78	29.77	7.90	5.9	10.12
46	40.76	29.38	9.85	5.9	10.20
47	40.75	29.33	8.53	5.3	4.40
48	40.76	29.55	5.64	5.2	3.92
49	40.74	29.34	10.30	6.0	12.52
50	40.78	29.72	10.56	6.9	45.20

**Table A.3:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the simulated earthquake scenarios along Segment 1 after Böse (2006).

<i>No.</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>Depth [km]</i>	<i>M<sub>w</sub></i>	<i>Rupture length [km]</i>
51	40.80	29.07	5.25	4.5	1.00
52	40.86	28.91	6.74	6.2	15.29
53	40.83	28.99	6.28	6.3	19.15
54	40.74	29.16	6.62	5.4	5.17
55	40.77	29.03	8.79	6.8	36.23
56	40.78	29.10	10.43	6.2	16.89
57	40.78	29.04	10.93	6.2	16.16
58	40.83	28.96	8.50	6.2	15.61
59	40.80	29.01	6.76	6.2	15.93
60	40.76	29.09	6.13	6.1	14.68
61	40.83	28.94	11.49	6.3	19.99
62	40.83	28.89	7.97	5.9	10.76
63	40.81	29.02	8.40	5.2	4.04
64	40.80	29.12	11.72	6.6	27.54
65	40.75	29.19	5.50	4.9	1.00
66	40.84	28.95	10.75	6.3	19.27
67	40.83	28.96	7.25	5.9	11.00
68	40.88	28.87	5.25	4.5	1.00
69	40.91	28.84	5.83	4.6	1.00
70	40.86	28.93	7.13	5.4	5.02
71	40.87	28.87	5.82	6.1	13.57
72	40.77	29.11	5.50	4.8	1.00
73	40.86	28.90	5.75	4.9	1.00
74	40.82	28.99	5.65	5.7	7.80
75	40.78	29.06	6.44	6.5	25.16
76	40.75	29.13	7.84	5.5	5.82
77	40.83	28.98	5.25	4.6	1.00
78	40.82	29.03	5.25	4.5	1.00
79	40.89	28.83	11.49	6.3	19.95
80	40.83	28.94	6.65	5.4	5.38
81	40.77	29.07	5.63	5.7	7.50
82	40.85	28.92	10.95	6.4	20.78
83	40.89	28.87	7.32	5.5	6.08
84	40.86	28.91	10.31	5.9	11.24
85	40.76	29.17	8.23	6.0	13.00
86	40.85	28.97	8.11	5.4	5.49
87	40.77	29.04	10.97	6.4	20.97
88	40.81	29.02	5.50	4.6	1.00
89	40.82	29.03	6.20	6.2	16.50
90	40.80	29.01	6.84	6.0	11.47
91	40.81	28.96	9.57	5.8	8.94
92	40.87	28.95	10.65	6.1	14.52
93	40.86	28.92	10.01	5.9	11.00
94	40.86	28.95	10.27	6.2	15.77
95	40.75	29.14	8.88	5.4	5.46
96	40.85	28.90	10.49	6.0	12.19
97	40.82	28.99	9.57	5.9	11.40
98	40.89	28.88	11.77	6.6	27.95
99	40.84	28.98	10.17	6.8	38.21
100	40.87	28.90	7.38	6.7	31.53

**Table A.4:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the simulated earthquake scenarios along Segment 2 after Böse (2006).

<i>No.</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>Depth [km]</i>	<i>M<sub>w</sub></i>	<i>Rupture length [km]</i>
101	40.86	27.69	5.96	5.9	10.09
102	40.89	28.32	10.23	6.8	39.26
103	40.84	27.76	8.10	7.1	57.92
104	40.87	28.15	10.62	6.3	18.27
105	40.82	27.43	8.70	5.6	6.97
106	40.91	28.51	5.87	6.2	15.36
107	40.86	28.44	11.99	6.6	30.15
108	40.83	27.45	8.59	5.9	9.93
109	40.80	27.63	11.14	6.4	22.37
110	40.82	27.64	5.12	4.6	1.00
111	40.85	28.15	14.48	7.1	60.69
112	40.90	28.38	8.51	6.2	15.69
113	40.85	27.78	8.05	7.1	55.79
114	40.86	27.93	6.64	6.1	13.47
115	40.83	27.80	10.85	7.0	50.83
116	40.85	28.03	8.66	6.2	17.37
117	40.87	27.81	7.55	6.8	37.09
118	40.86	28.07	10.27	6.8	39.96
119	40.85	27.88	11.24	6.4	23.20
120	40.84	27.85	10.54	6.3	17.68
121	40.88	28.25	14.73	7.2	64.48
122	40.85	27.94	7.19	5.1	3.20
123	40.84	27.88	10.48	6.9	43.85
124	40.88	28.30	15.25	7.2	72.58
125	40.82	27.59	8.52	5.4	4.89
126	40.90	28.45	5.93	6.3	17.85
127	40.87	28.14	5.25	4.7	1.00
128	40.92	28.57	11.04	7.0	54.57
129	40.88	28.34	7.23	7.2	68.95
130	40.85	28.09	7.29	7.3	73.53
131	40.86	27.98	12.19	6.7	32.20
132	40.84	27.81	11.12	7.1	56.24
133	40.85	28.09	7.53	5.6	7.42
134	40.91	28.64	10.39	6.2	16.62
135	40.86	28.10	8.36	5.7	8.50
136	40.84	27.93	14.45	7.1	60.19
137	40.86	28.38	5.83	4.5	1.00
138	40.90	28.68	9.04	6.4	21.79
139	40.90	28.49	14.20	7.1	56.61
140	40.81	27.45	6.45	5.9	10.17
141	40.87	28.11	11.31	7.1	60.42
142	40.82	27.67	6.95	6.3	20.11
143	40.84	27.60	5.25	4.5	1.00
144	40.87	27.94	8.78	5.4	5.15
145	40.84	27.79	9.88	6.7	33.61
146	40.81	27.69	9.66	6.0	11.86
147	40.87	27.93	11.31	6.5	23.82
148	40.88	28.18	15.45	7.1	59.55
149	40.83	27.50	7.12	6.5	24.15
150	40.91	28.68	15.61	7.3	78.46

**Table A.5:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the simulated earthquake scenarios along Segment 3 after Böse (2006).

<i>No.</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>Depth [km]</i>	<i>M<sub>w</sub></i>	<i>Rupture length [km]</i>
151	40.74	29.17	12.12	7.3	79.64
152	40.88	28.24	7.34	7.3	77.06
153	40.85	28.44	6.46	6.5	26.02
154	40.83	28.82	10.94	7.0	52.54
155	40.83	28.74	7.93	7.0	51.10
156	40.85	28.51	14.51	7.1	61.07
157	40.86	28.20	10.57	6.9	45.40
158	40.84	28.41	10.54	6.9	44.79
159	40.81	28.74	10.55	6.9	44.98
160	40.80	28.65	6.15	6.6	29.11
161	40.84	28.63	8.36	7.2	69.67
162	40.82	28.61	6.69	6.8	36.71
163	40.86	28.43	10.29	6.8	40.24
164	40.81	28.76	14.40	7.1	59.61
165	40.83	28.62	13.18	6.7	33.86
166	40.83	28.66	10.60	6.9	45.96
167	40.85	28.37	16.88	7.5	101.86
168	40.79	28.88	12.73	7.4	96.50
169	40.82	28.79	7.99	7.0	53.44
170	40.72	29.32	6.49	6.6	27.34
171	40.88	28.30	10.38	6.9	41.96
172	40.76	29.13	8.19	7.1	61.99
173	40.76	29.13	14.46	6.9	47.38
174	40.81	28.51	7.76	6.9	44.52
175	40.77	29.03	8.77	7.4	90.92
176	40.82	28.60	12.21	7.3	82.05
177	40.84	28.53	12.94	6.7	31.70
178	40.84	28.83	7.48	6.7	34.86
179	40.88	28.24	16.44	7.4	93.29
180	40.83	28.63	11.57	7.2	66.46
181	40.81	28.84	12.54	7.4	91.04
182	40.82	28.87	13.24	7.5	111.75
183	40.83	28.67	15.57	7.3	77.90
184	40.82	28.66	12.63	7.4	93.53
185	40.81	28.89	12.41	7.4	87.54
186	40.85	28.54	9.99	6.7	35.23
187	40.74	29.42	12.18	6.7	32.01
188	40.81	28.76	7.89	7.0	49.52
189	40.84	28.71	7.62	6.8	39.49
190	40.80	28.81	15.31	7.1	57.69
191	40.86	28.46	14.07	7.0	54.86
192	40.86	28.45	7.80	6.9	46.01
193	40.84	28.72	14.00	7.0	53.86
194	40.79	29.09	18.27	7.5	103.00
195	40.90	28.17	16.13	7.4	87.63
196	40.86	28.48	16.98	7.5	103.76
197	40.84	28.49	10.46	6.9	43.23
198	40.85	28.41	8.73	6.7	35.03
199	40.79	28.95	14.29	7.4	88.12
200	40.89	28.39	12.73	7.4	96.27

**Table A.6:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the simulated earthquake scenarios along Segment 4 after Böse (2006).

<i>No.</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>Depth [km]</i>	<i>M<sub>w</sub></i>	<i>Rupture length [km]</i>
201	40.73	29.00	5.78	6.0	11.93
202	40.71	28.87	5.75	4.9	1.00
203	40.72	29.00	5.75	4.5	1.00
204	40.70	28.98	11.83	6.6	28.58
205	40.69	29.05	9.87	6.0	13.12
206	40.72	29.01	5.80	5.6	6.54
207	40.68	28.98	10.11	6.1	14.68
208	40.70	29.02	5.75	5.0	1.00
209	40.66	29.05	5.60	5.6	6.63
210	40.70	28.98	11.93	6.4	23.13
211	40.71	28.82	6.94	6.3	19.78
212	40.72	28.91	7.92	5.9	10.26
213	40.70	29.04	5.75	4.8	1.00
214	40.71	28.92	8.33	6.6	27.01
215	40.72	29.01	9.07	6.4	22.08
216	40.70	28.97	9.61	6.0	11.56
217	40.72	28.96	5.64	5.2	3.94
218	40.70	29.07	7.23	5.4	5.53
219	40.70	29.07	12.31	6.5	23.45
220	40.73	28.93	9.40	5.9	10.42
221	40.73	28.87	9.00	5.5	5.87
222	40.69	29.02	11.84	6.6	28.64
223	40.73	28.97	6.58	6.0	12.37
224	40.70	28.97	10.18	5.9	10.63
225	40.68	29.05	5.25	4.9	1.00
226	40.71	29.00	6.27	6.3	19.00
227	40.70	29.07	5.75	4.9	1.00
228	40.71	29.06	5.75	4.6	1.00
229	40.70	29.02	5.25	4.9	1.00
230	40.69	29.00	6.88	5.6	7.23
231	40.72	28.88	5.62	6.0	11.61
232	40.68	28.97	6.07	6.0	12.91
233	40.70	29.02	5.75	4.9	1.00
234	40.70	28.98	8.97	5.6	6.49
235	40.72	28.92	7.14	6.5	24.95
236	40.68	29.10	5.50	4.5	1.00
237	40.72	28.97	6.30	6.3	19.88
238	40.73	28.97	9.94	5.8	9.55
239	40.70	29.02	6.68	5.5	5.64
240	40.71	28.86	5.75	4.9	1.00
241	40.70	28.92	7.32	5.5	6.06
242	40.71	28.92	10.07	5.8	9.48
243	40.72	28.95	11.54	6.5	25.88
244	40.68	29.00	5.75	4.7	1.00
245	40.71	28.92	9.29	6.5	24.92
246	40.72	28.94	11.29	6.5	23.68
247	40.70	29.07	9.20	6.5	23.71
248	40.70	29.08	11.42	6.5	24.77
249	40.74	28.87	7.24	5.9	10.89
250	40.72	28.92	5.50	5.0	1.00

**Table A.7:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the simulated earthquake scenarios along Segment 5 after Böse (2006).

<i>No.</i>	<i>Latitude [°N]</i>	<i>Longitude [°E]</i>	<i>Depth [km]</i>	<i>M<sub>w</sub></i>	<i>Rupture length [km]</i>
251	41.14	29.53	5.75	4.7	1.00
252	41.04	29.62	5.83	4.5	1.00
253	40.99	28.05	5.75	4.8	1.00
254	40.68	28.68	5.75	4.7	1.00
255	40.66	29.14	5.50	4.8	1.00
256	40.84	28.76	5.83	4.7	1.00
257	40.76	28.65	5.25	5.0	1.00
258	40.73	27.98	5.25	4.7	1.00
259	40.89	28.19	5.25	4.9	1.00
260	41.02	29.52	5.75	4.7	1.00
261	41.00	28.26	5.50	4.8	1.00
262	40.99	29.86	5.62	4.5	1.00
263	41.04	29.65	5.25	4.9	1.00
264	40.69	29.46	5.25	4.7	1.00
265	41.21	29.16	5.50	5.0	1.00
266	40.84	28.30	5.75	4.7	1.00
267	40.78	28.04	5.50	4.7	1.00
268	40.69	29.10	5.50	4.6	1.00
269	40.68	29.39	5.50	4.6	1.00
270	40.69	28.81	5.25	4.9	1.00
271	40.82	28.67	5.25	4.7	1.00
272	40.91	28.14	5.83	4.8	1.00
273	41.06	28.46	5.50	4.6	1.00
274	40.84	28.56	5.25	4.9	1.00
275	41.09	28.26	5.25	4.7	1.00
276	40.93	29.33	5.25	4.7	1.00
277	41.03	29.17	5.50	4.9	1.00
278	40.76	28.64	5.50	4.9	1.00
279	41.08	28.75	5.25	4.9	1.00
280	41.02	28.10	5.50	4.5	1.00

**Table A.8:** Locations, moment magnitudes  $M_w$ , and rupture lengths of the 30, randomly distributed earthquake scenarios. They are referred to as Segment 6 after Böse (2006).

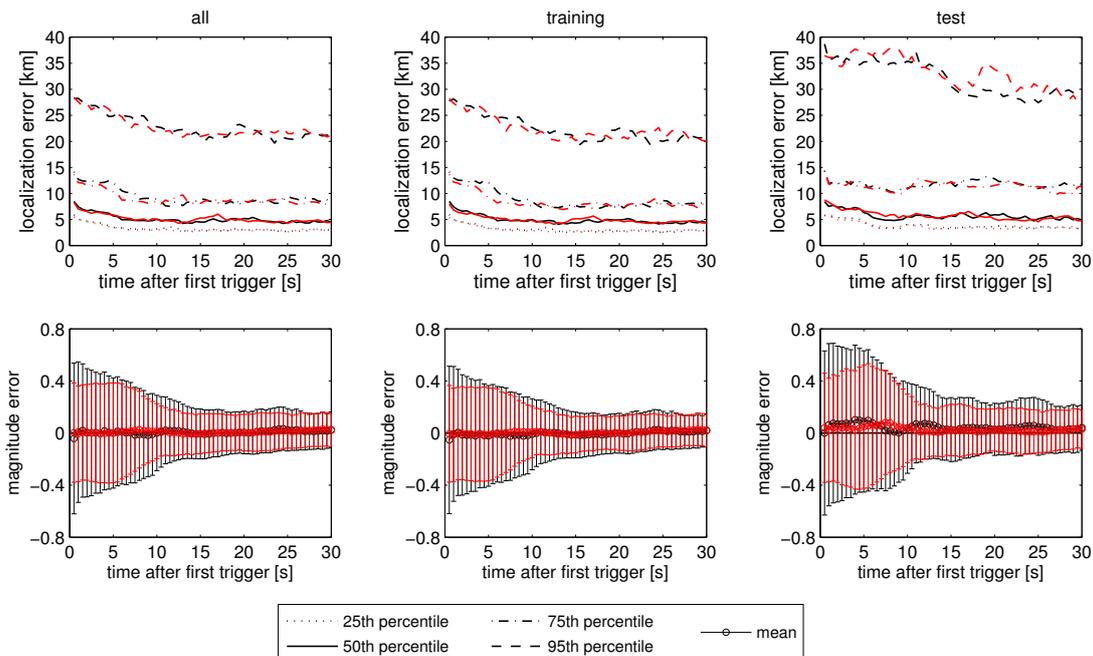
## A.2 PreSEIS Results: Optimized *FINSIM* Simulations

*PreSEIS* has been applied to the optimized simulated earthquake dataset derived in Chapter 2.7. Figure A.1 shows the errors of the localization and magnitude estimates of the *PreSEIS* performance (red curves) compared to the results obtained using the original dataset by Böse (2006) (black curves).

The left column presents the errors with ongoing time obtained from the complete dataset, i.e. all 280 earthquakes. The middle column shows the results from the training process with 80% of the events in the catalogue. The right column gives the results obtained from the application of the trained system to the remaining (unknown) 20% of the data.

The magnitude errors (bottom) decrease significantly when using the optimized parameterization. Already at the first time step of 0.5 seconds, the magnitude errors are 34% smaller. The localization errors (top) are not influenced because they are determined from the P-wave onset times only. These are not changed by the parameter optimization.

Details to the *PreSEIS* methodology can be found in Chapter 4.2.



**Figure A.1:** Results from the application of *PreSEIS* to both the original simulated dataset (black curves) as well as to the optimized simulations (red curves). The top row shows the absolute localization errors with proceeding time after the first station trigger, while the bottom row presents the according mean magnitude errors.

# Appendix B

## Attenuation Relationships

### B.1 Campbell & Bozorgnia (2008): Global PEER database

Campbell and Bozorgnia (2008) give an empirical ground motion model for the geometric mean of horizontal peak ground acceleration (*PGA*), velocity (*PGV*), and displacement (*PGD*) as well as for 5% damped linear elastic response spectra for periods ranging between 0.01 and 10 s. The ground motion model is derived from a set of crustal earthquakes from tectonically active regimes worldwide, which is a subset of the PEER (Pacific Earthquake Engineering Research Center) strong motion database. The ground motion model is given in terms of moment magnitude, style of faulting, rupture depth, hanging wall geometry, site response, basin response, and inter-event and intra-event variability.

The ground motion model is valid for earthquakes with moment magnitudes  $M_w$  between 4.0 and 7.5 and for distances ranging from 0 to 200 km.

The attenuation relation is given as

$$\ln Y = f_{mag} + f_{dis} + f_{flt} + f_{hng} + f_{site} + f_{sed} \quad (\text{B.1})$$

where  $Y$  is the respective ground motion parameter in g.

The magnitude term is defined by

$$f_{mag} = \begin{cases} c_0 + c_1 M_w; & M_w \leq 5.5 \\ c_0 + c_1 M_w + c_2 (M_w - 5.5); & 5.5 < M_w \leq 6.5 \\ c_0 + c_1 M_w + c_2 (M_w - 5.5) + c_3 (M_w - 6.5); & M_w > 6.5 \end{cases} \quad (\text{B.2})$$

The distance term is given by

$$f_{dis} = (c_4 + c_5 M_w) \ln \left( \sqrt{R_{RUP}^2 + c_6^2} \right) \quad (\text{B.3})$$

where  $R_{RUP}$  is the closest distance to the co-seismic rupture plane in km.

The style-of-faulting term is expressed as

$$f_{flt} = c_7 F_{RV} f_{flt,Z} + c_8 F_{NM} \quad (B.4)$$

where  $F_{RV}$  represents reverse and reverse-oblique faulting and  $F_{NM}$  normal and normal-oblique faulting:

$$F_{RV} = \begin{cases} 1; & 30^\circ < \lambda < 150^\circ \\ 0; & \text{else} \end{cases} \quad (B.5)$$

$$F_{RV} = \begin{cases} 1; & -150^\circ < \lambda < -30^\circ \\ 0; & \text{else} \end{cases} \quad (B.6)$$

$\lambda$  is the rake angle.

$f_{flt,Z}$  is defined as

$$f_{flt,Z} = \begin{cases} Z_{TOR}; & Z_{TOR} < 1; \\ 1; & Z_{TOR} \geq 1 \end{cases} \quad (B.7)$$

with  $Z_{TOR}$  being the depth to the top of the co-seismic rupture plane in km.

The hanging-wall term is given by

$$f_{hng} = c_9 f_{hng,R} f_{hng,M} f_{hng,Z} f_{hng,\delta} \quad (B.8)$$

with

$$f_{hng,R} = \begin{cases} 1; & r_{jb} = 0 \\ (\max(R_{RUP}, \sqrt{r_{jb}^2 + 1}) - r_{jb}) / \max(R_{RUP}, \sqrt{r_{jb}^2 + 1}); & r_{jb} > 0, Z_{TOR} < 1 \\ (R_{RUP} - r_{jb}) / R_{RUP}; & r_{jb} > 0, Z_{TOR} \geq 1 \end{cases} \quad (B.9)$$

$$f_{hng,M} = \begin{cases} 0; & M_w \leq 6.0 \\ 2(M_w - 6.0); & 6.0 < M_w < 6.5 \\ 1; & M_w \geq 6.5 \end{cases} \quad (B.10)$$

$$f_{hng,Z} = \begin{cases} 0; & Z_{TOR} \geq 20 \\ (20 - Z_{TOR}) / 20; & 0 \leq Z_{TOR} < 20 \end{cases} \quad (B.11)$$

$$f_{hng,\delta} = \begin{cases} 1; & \delta \leq 70 \\ (90 - \delta) / 20; & \delta > 70 \end{cases} \quad (B.12)$$

$r_{jb}$  is the Joyner-Boore distance in km and  $\delta$  is the dip of the rupture plane.

The shallow site response term is expressed as

$$f_{site} = \begin{cases} c_{10} \ln\left(\frac{V_{S30}}{k_1}\right) + k_2 \left\{ \ln\left(A_{1100} + c\left(\frac{V_{S30}}{k_1}\right)^n\right) - \ln(A_{1100} + c) \right\}; & V_{S30} < k_1 \\ (c_{10} + k_2 n) \ln\left(\frac{V_{S30}}{k_1}\right); & k_1 \leq V_{S30} < 1100 \\ (c_{10} + k_2 n) \ln\left(\frac{1100}{k_1}\right); & V_{S30} \geq 1100 \end{cases} \quad (B.13)$$

<i>Period [s]</i>	$c_0$	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$	$c_7$
PGA	-1.715	0.500	-0.530	-0.262	-2.118	0.170	5.60	0.280
0.20	-0.486	0.500	-0.446	-0.398	-2.220	0.170	7.60	0.280
0.30	-1.171	0.500	-0.294	-0.511	-2.095	0.170	6.04	0.280
0.40	-1.466	0.500	-0.186	-0.592	-2.066	0.170	5.30	0.280
0.50	-2.569	0.656	-0.304	-0.536	-2.041	0.170	4.73	0.280
0.75	-4.844	0.972	-0.578	-0.406	-2.000	0.170	4.00	0.280

<i>Period [s]</i>	$c_8$	$c_9$	$c_{10}$	$c_{11}$	$c_{12}$	$k_1$	$k_2$	$k_3$
PGA	-0.120	0.490	1.058	0.040	0.610	865.0	-1.186	1.839
0.20	-0.012	0.490	2.194	0.040	0.610	748.0	-2.188	1.856
0.30	0.000	0.490	2.460	0.040	0.750	587.0	-2.518	1.865
0.40	0.000	0.490	2.587	0.040	0.850	503.0	-2.657	1.874
0.50	0.000	0.490	2.544	0.040	0.883	457.0	-2.669	1.883
0.75	0.000	0.490	2.133	0.077	1.000	410.0	-2.401	1.906

**Table B.1:** Model coefficients of equation B.1 for estimating the geometric mean of horizontal peak ground acceleration (*PGA*) and 5% damped spectral acceleration (both in g) for selected natural periods after Campbell and Bozorgnia (2008).

$V_{S30}$  is the averaged S-wave velocity in the top 30 m of the site profile in m/s (see Chapter 2.3).  $A_{1100}$  is the median estimate of PGA on a reference rock outcrop ( $V_{S30} = 1100$  m/s) in g.

The basin response term is given by

$$f_{sed} = \begin{cases} c_{11}(Z_{2.5} - 1); & Z_{2.5} < 1 \\ 0; & 1 \leq Z_{2.5} \leq 3 \\ c_{12}k_3e^{-0.75}(1 - e^{-0.25(Z_{2.5}-3)}); & Z_{2.5} > 3 \end{cases} \quad (\text{B.14})$$

where  $Z_{2.5}$  is the depth in km to the 2.5 km/s shear-wave velocity horizon, typically referred to as basin or sediment depth.

The model coefficients for the natural periods used to estimate the building damages in Chapter 3,  $c_0$  to  $c_{12}$  and  $k_1$  to  $k_3$ , are listed in Table B.1. However, not all periods necessary for the building damage estimates are available in this study. The remaining coefficients  $c$  and  $n$  are set to  $c = 1.88$  and  $n = 1.18$  for all periods.

## B.2 Özbey et al. (2004): Northwestern Turkey

Özbey et al. (2004) give empirical attenuation relationships for the geometric mean of horizontal peak ground acceleration (*PGA*) and 5% damped spectral acceleration ( $S_a$ ) for earthquakes in northwestern Turkey. The equations give ground motion in terms of moment magnitude,

<i>Period [s]</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>h</i>	<i>e</i>	<i>f</i>
PGA	3.287	0.503	-0.079	-1.1177	14.82	0.141	0.331
0.20	3.518	0.494	-0.094	-1.1162	14.87	0.113	0.285
0.30	3.040	0.549	-0.095	-0.8762	6.54	0.062	0.320
0.40	2.825	0.593	-0.112	-0.8089	6.48	0.102	0.394
0.50	2.685	0.653	-0.171	-0.7302	5.58	0.051	0.385
0.55	2.581	0.685	-0.177	-0.6928	3.56	0.061	0.393
0.75	2.247	0.750	-0.170	-0.5946	2.34	0.054	0.396
0.90	2.272	0.791	-0.172	-0.6630	4.21	0.102	0.416
1.20	2.267	0.874	-0.267	-0.6910	4.49	0.103	0.397

**Table B.2:** Attenuation model coefficients of equation B.15 for estimating the geometric mean of horizontal peak ground acceleration (*PGA*) and 5% damped spectral acceleration (both in  $\text{cm/s}^2$ ) for selected natural periods after Özbey et al. (2004).

*Joyner-Boore* distance, and site conditions for normal and strike-slip earthquakes. The attenuation relationships are valid for earthquakes with moment magnitudes  $\geq 5.0$ .

The empirical attenuation model is given as

$$\log Y = a + b (M_w - 6) + c (M_w - 6)^2 + d \log \sqrt{r_{jb}^2 + h^2} + e G_1 + f G_2 \quad (\text{B.15})$$

where  $Y$  is the ground motion parameter (*PGA* or  $S_a$ ) in  $\text{cm/s}^2$ ,  $M_w$  is the moment magnitude, and  $r_{jb}$  is the *Joyner-Boore* distance in km.

$a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$ ,  $f$ , and  $h$  are the attenuation model coefficients. Depending on the site conditions, the coefficients  $G_1$  and  $G_2$  are set to

Site class A + B:  $G_1 = 0, G_2 = 0$

Site class C:  $G_1 = 1, G_2 = 0$

Site class D:  $G_1 = 0, G_2 = 1$ .

The respective site classes are defined as following average S-wave velocities over the top 30 m ( $V_{S30}$ ):

Site class A:  $> 750$  m/s

Site class B: 360 - 750 m/s

Site class C: 180 - 360 m/s

Site class D:  $< 180$  m/s

The attenuation model coefficients for calculating *PGA* and  $S_a$  at the natural periods used to estimate the building damages in Chapter 3 can be found in Table B.2.

### B.3 Boore et al. (1997): Western North America

Boore et al. (1997) give equations to estimate random horizontal-component peak acceleration and 5% damped pseudo-acceleration response spectra in terms of the natural logarithm of the ground motion parameter for shallow earthquakes in western North America. The equations give ground motion in terms of moment magnitude  $M_w$ , *Joyner-Boore* distance, and site conditions for strike-slip, reverse-slip, or unspecified faulting mechanisms. For the latter it is assumed that reverse slip earthquakes have positive rake angles and the absolute value of the rake for left-lateral slip is  $< 90^\circ$  and that strike-slip events are those with a rake angle within  $30^\circ$  of horizontal. The relationships are defined for earthquakes in the magnitude range  $5.5 \leq M_w \leq 7.5$  and for distances not larger than 80 km.

The ground motion equation is given as

$$\ln Y = b_1 + b_2 (M_w - 6) + b_3 (M_w - 6)^2 + b_5 \ln r + b_V \ln \frac{V_S}{V_A} \quad (\text{B.16})$$

where

$$r = \sqrt{r_{jb}^2 + h^2} \quad (\text{B.17})$$

and

$$b_1 = \begin{cases} b_{1SS} & \text{for strike-slip earthquakes;} \\ b_{1RS} & \text{for reverse-slip earthquakes;} \\ b_{1ALL} & \text{if mechanism is not specified.} \end{cases} \quad (\text{B.18})$$

$Y$  is the ground motion parameter (peak horizontal acceleration or pseudo-acceleration response) in g,  $M_w$  is the moment magnitude,  $r_{jb}$  is the *Joyner-Boore* distance in km, and  $V_{S30}$  is the average S-wave velocity over the top 30 m in m/s.

$b_{1SS}$ ,  $b_{1RS}$ ,  $b_{1ALL}$ ,  $b_2$ ,  $b_3$ ,  $b_5$ ,  $h$ ,  $b_V$ , and  $V_A$  are the coefficients determined in the regression.  $V_{S30}$  is recommended as:

NEHRP site class B:	1070 m/s
NEHRP site class C:	520 m/s
NEHRP site class D:	250 m/s
Rock:	620 m/s
Soil:	310 m/s

The regression coefficients for calculating the peak acceleration (*PGA*) and the response spectra at the natural periods used to estimate the building damages in Chapter 3 can be found in Table B.3.

<i>Period [s]</i>	$b_{ISS}$	$b_{IRS}$	$b_{IALL}$	$b_2$	$b_3$	$b_5$	$b_V$	$V_A$	$h$
PGA	-0.313	-0.117	-0.242	0.527	0.000	-0.778	-0.371	1396	5.57
0.20	0.999	1.170	1.089	0.711	-0.207	-0.924	-0.292	2118	7.02
0.30	0.598	0.803	0.700	0.769	-0.161	-0.893	-0.401	2133	5.94
0.40	0.212	0.423	0.311	0.831	-0.120	-0.867	-0.487	1954	4.91
0.50	-0.122	0.087	-0.025	0.884	-0.090	-0.846	-0.553	1782	4.13
0.55	-0.268	-0.063	-0.176	0.907	-0.078	-0.837	-0.579	1710	3.82
0.75	-0.737	-0.562	-0.661	0.979	-0.046	-0.813	-0.653	1507	3.07
0.90	-0.993	-0.848	-0.933	1.018	-0.035	-0.802	-0.685	1432	2.89
1.20	-1.345	-1.265	-1.315	1.064	-0.032	-0.794	-0.710	1400	3.14

**Table B.3:** Regression coefficients of equation B.16 for estimating peak horizontal acceleration (*PGA*) and 5% damped pseudo-acceleration response spectra (both in *g*) for selected natural periods after Boore et al. (1997).

# Appendix C

## Building Damage Estimation

The estimation of building damages due to earthquakes requires information about the building stock in the study area. The *KOERILoss* method (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a) uses a vulnerability-based building classification that incorporates three different categories:

1) Construction type:

- $I = 1$ : Reinforced concrete frame building
- $I = 2$ : Masonry building
- $I = 3$ : Reinforced concrete shear wall building
- $I = 4$ : Precast building

2) Number of stories (including basement):

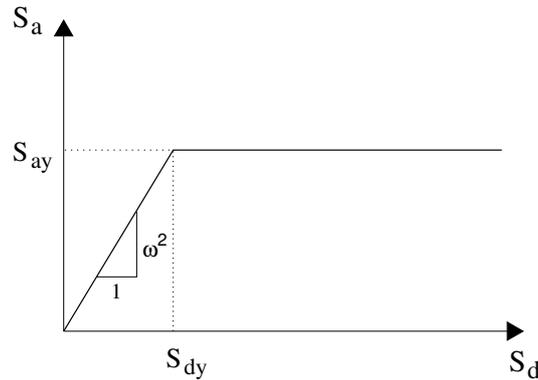
- $J = 1$ : Low-rise (1-4 stories)
- $J = 2$ : Mid-rise (5-8 stories)
- $J = 3$ : High-rise (8 and more stories)

3) Construction date:

- $K = 1$ : Pre-1980 (1980 included)
- $K = 2$ : Post-1980

In Istanbul, low- and mid-rise buildings of reinforced concrete frame structure represent 75% of the total building stock (Erdik et al., 2003a).

As stated in Chapter 3, the spectral acceleration-based method for estimating damages requires the establishment of the so-called *spectral displacement demand* of a given structure, determined for a given earthquake. In *KOERILoss*, the estimation of the spectral displacement demand fol-



**Figure C.1:** Idealized capacity diagram (spectral acceleration versus spectral displacement) representing elastic perfectly-plastic behaviour, after Erdik et al. (American Red Cross - Bogazici University, 2002).

lows the HAZUS<sup>1</sup> damage estimation methodology. In HAZUS, the building is idealized as an equivalent non-linear, single-degree-of-freedom system subjected to a slowly increasing lateral loading pattern (Chen and Scawthorn, 2003).

The structural capacity of the building is represented by the *capacity spectrum* (capacity diagram), plotted as spectral acceleration versus spectral displacement. Figure C.1 presents a typical, idealized capacity diagram given by Erdik et al. in the report of American Red Cross - Bogazici University (2002). The diagram represents an elastic perfectly-plastic behaviour with the initial slope corresponding to the square of the natural frequency.

The capacity diagram of a given structure can be estimated by its *yield spectral acceleration*  $S_{ay}$  (Figure C.1) which is given as

$$S_{ay} = C_s \gamma \lambda \frac{g}{\alpha_1}. \quad (\text{C.1})$$

$C_s$  is the approximate value of the estimated design lateral strength factor assumed to be valid for the given date of construction. The quantity  $\gamma$  represents the ratio of the yield strength to the design strength and  $\lambda$  is defined as the ratio of the ultimate strength to the yield strength. The parameter  $g$  is the gravitational acceleration and  $\alpha_1$  refers to the mass ratio. The parameters  $C_s$ ,  $\gamma$ ,  $\lambda$ , and  $\alpha_1$  depend on the building category and are defined in Table C.1 (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

The elastic spectral acceleration  $S_{ae}$  associated with the natural period  $T$  of the building (obtained from the simulated ground accelerations) allows for calculating the *strength reduction factor*  $R_y$  by using the above obtained yield spectral acceleration:

$$R_y = \frac{S_{ae}(T)}{S_{ay}}. \quad (\text{C.2})$$

<sup>1</sup>HAZUS (1999): Earthquake Loss Estimation Methodology, Technical Manual, RMS, NIBS, and FEMA

<i>I</i>	<i>J</i>	<i>T</i> [s]	$\alpha_1$	$\gamma$	$\lambda$	$C_s$		$C_2$	
						<i>K</i> = 1	<i>K</i> = 2	<i>K</i> = 1	<i>K</i> = 2
1	1	0.40	0.80	1.30	2.00	0.06	0.08	1.2	1.1
1	2	0.75	0.80	1.15	2.00	0.06	0.08	1.1	1.0
1	3	1.20	0.75	1.00	2.00	0.04	0.06	1.0	1.0
2	1	0.30	0.75	1.00	2.00	0.06	0.06	1.2	1.2
2	2	0.50	0.75	1.00	2.00	0.06	0.06	1.2	1.2
3 / 4	1	0.30	0.75	1.50	2.50	0.06	0.08	1.1	1.1
3 / 4	2	0.55	0.75	1.30	2.50	0.06	0.08	1.0	1.0
3 / 4	3	0.90	0.65	1.15	2.50	0.06	0.08	1.0	1.0

**Table C.1:** Building capacity parameters after Erdik et al. (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a)

$R_y$  is used to calculate the *spectral displacement amplification factor*  $C_1$ , defined as

$$C_1 = \frac{1 + (R_y - 1) T_s / T}{R_y} \quad (C.3)$$

$T_s$  refers to the transition period from the constant acceleration region to the constant velocity region of the acceleration spectrum (Figure C.1). For  $T \geq T_s$ , the amplification factor  $C_1$  is set to 1.

Together with the elastic spectral displacement  $S_{de}(T)$ , the spectral displacement amplification factor  $C_1$  can finally be used to calculate the required inelastic spectral displacement demand  $S_{di}(T, R_y)$ . It is obtained from

$$S_{di}(T, R_y) = C_1 \cdot C_2 \cdot S_{de}(T) \quad (C.4)$$

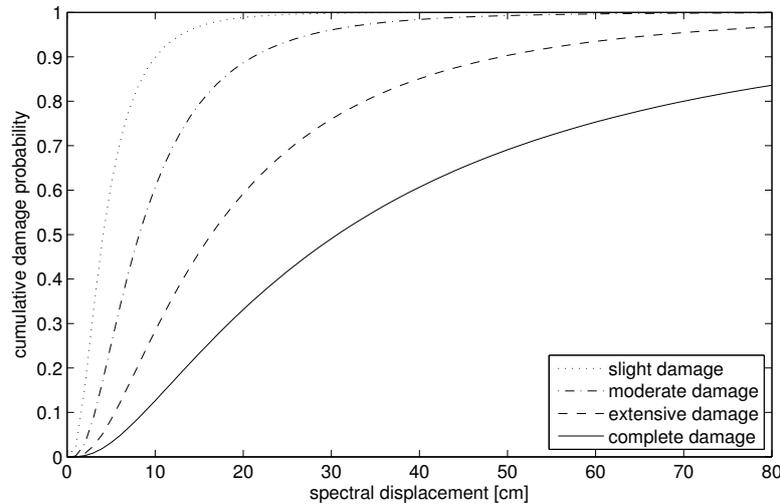
whereas the elastic spectral displacement is related to the elastic spectral acceleration  $S_{ae}(T)$  in the form

$$S_{de}(T) = \left( \frac{T}{2\pi} \right)^2 \cdot S_{ae}(T). \quad (C.5)$$

The coefficient  $C_2$  in equation C.4 is also specified in Table C.1 (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

The inelastic spectral displacement demand  $S_{di}$  is now entered into displacement-based fragility curves, where the horizontal axis represents the spectral displacement demand. The vertical axis gives the cumulative probability of structural damage reaching or exceeding the threshold of a given damage level. Each fragility curve is represented by a distribution function  $\Phi$ :

$$P[D \geq ds | S_{di}] = \Phi \left[ \frac{1}{\beta_{ds}} \ln \left( \frac{S_{di}}{S_{d,ds}} \right) \right]. \quad (C.6)$$



**Figure C.2:** Spectral displacement-based vulnerability curves for reinforced concrete frame, mid-rise, pre-1980 buildings for the different damage levels after the *KOERILoss* method by Erdik et al. (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

$D$  refers to the damage. The quantity  $S_{d,ds}$  is the median spectral displacement corresponding to the threshold of the reached damage level,  $ds$ , classified as *slight*, *moderate*, *extensive*, and *complete* damage.  $\beta_{ds}$  is the standard deviation of the natural logarithm of the spectral displacement corresponding to the considered damage level. The function  $\Phi$  is the cumulative standard normal distribution function.

The parameter  $S_{d,ds}$  in equation C.6 is expressed as

$$S_{d,ds} = \alpha_2 \cdot D_{ds} \cdot H. \quad (C.7)$$

$D_{ds}$  refers to the median story drift ratio estimated for the considered damage level.  $H$  is the height of the building and  $\alpha_2$  is the modal parameter (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).

The coefficients  $H$ ,  $\alpha_2$ ,  $D_{ds}$ ,  $S_{d,ds}$ , and  $\beta_{ds}$  for the fragility curves are given in Table C.2.

Figure C.2 presents an example of the spectral displacement-based vulnerability curves for reinforced concrete frame ( $I = 1$ ), mid-rise ( $J = 2$ ), pre-1980 ( $K = 1$ ) buildings. The vertical axis gives the cumulative probability to reach or exceed the damage levels slight, moderate, extensive, and complete.

				$K = 1$											
				<i>Slight (s)</i>			<i>Moderate (m)</i>			<i>Extensive (e)</i>			<i>Complete (c)</i>		
<i>I</i>	<i>J</i>	<i>H</i> [m]	$\alpha_2$	$D_s$ %	$S_{d,s}$ [cm]	$\beta_s$	$D_m$ %	$S_{d,m}$ [cm]	$\beta_m$	$D_e$ %	$S_{d,e}$ [cm]	$\beta_e$	$D_c$ %	$S_{d,c}$ [cm]	$\beta_c$
1	1	6	.75	.40	1.80	.95	.80	3.60	.91	2.00	9.00	.85	4.00	18.00	.97
1	2	15	.75	.35	3.94	.70	.80	9.00	.74	1.60	18.00	.86	3.00	33.75	.98
1	3	30	.65	.30	6.75	.70	.80	18.00	.81	1.20	27.00	.89	2.50	56.25	.98
2	1	6	.75	.30	1.35	.99	.70	3.15	1.05	1.60	7.20	1.10	3.20	14.40	1.08
2	2	10	.75	.35	2.63	.70	.80	6.00	.74	1.60	12.00	.86	3.00	22.50	.98
3/4	1	6	.75	.40	1.80	.95	.80	3.60	.91	2.00	9.00	.85	4.00	18.00	.97
3/4	2	15	.75	.30	3.38	.70	.60	6.75	.74	1.20	13.50	.86	2.67	30.04	.98
3/4	3	30	.60	.30	6.75	.70	.80	18.00	.81	1.20	27.00	.89	2.50	56.25	.98

				$K = 2$											
				<i>Slight (s)</i>			<i>Moderate (m)</i>			<i>Extensive (e)</i>			<i>Complete (c)</i>		
<i>I</i>	<i>J</i>	<i>H</i> [m]	$\alpha_2$	$D_s$ %	$S_{d,s}$ [cm]	$\beta_s$	$D_m$ %	$S_{d,m}$ [cm]	$\beta_m$	$D_e$ %	$S_{d,e}$ [cm]	$\beta_e$	$D_c$ %	$S_{d,c}$ [cm]	$\beta_c$
1	1	6	.75	.50	2.25	.89	1.00	4.50	.90	2.50	11.25	.90	5.00	22.50	.89
1	2	15	.75	.40	4.50	.70	1.00	11.25	.70	2.00	22.50	.70	4.00	45.00	.89
1	3	30	.65	.40	9.00	.66	1.00	22.50	.66	1.60	36.00	.76	3.20	72.00	.91
2	1	6	.75	.40	1.80	.99	0.90	4.05	1.05	2.00	9.00	1.10	4.00	18.00	1.08
2	2	10	.75	.40	3.00	.70	1.00	7.50	.70	2.00	15.00	.70	4.00	30.00	.89
3/4	1	6	.75	.50	2.25	.89	1.00	4.50	.90	2.50	11.25	.90	5.00	22.50	.89
3/4	2	15	.75	.40	4.50	.70	0.70	7.88	.70	1.60	18.00	.70	3.33	37.46	.89
3/4	3	30	.60	.40	9.00	.66	1.00	22.50	.66	1.60	36.00	.76	3.20	72.00	.91

**Table C.2:** Spectral displacement-based fragility curve data after the *KOERILoss* method by Erdik et al. (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a).



# Appendix D

## The Virtual Seismologist Method

The *Virtual Seismologist (VS)* method is a Bayesian approach to earthquake early warning. It uses ground motion amplitudes, pre-defined prior information, and ground motion envelope attenuation relationships to estimate the most probable location and magnitude of an occurring earthquake (Cua, 2005; Cua and Heaton, 2007; Cua et al., 2009).

Bayes' theorem considers the most probable earthquake source estimates at any given time step as a combination of prior information and a likelihood function. Prior information can include hazard maps, fault traces, the health status of the network, or the Gutenberg-Richter relationship, for example. The likelihood function accounts for the information derived from an ongoing earthquake. The estimates of the earthquake source parameters are updated at every second (Cua, 2005; Cua and Heaton, 2007; Cua et al., 2009).

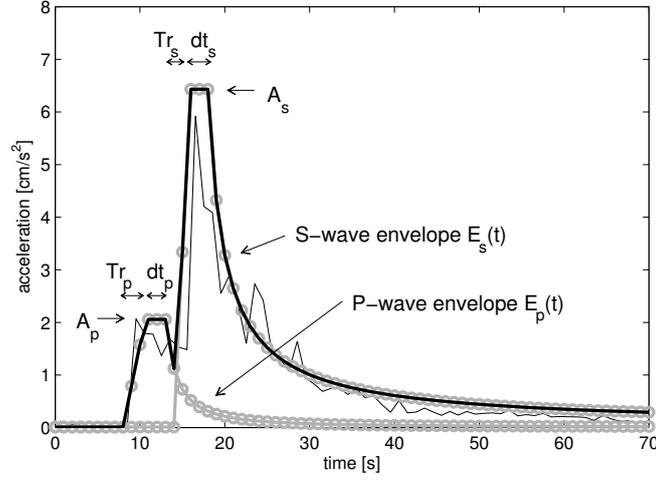
The likelihood function in the *VS* can be described as a set of relationships which map the information inferred from an ongoing earthquake into the source parameter estimates. The information inferred from the observations include the arrival times and the ground motion envelope amplitudes. The ground motion envelope is defined as the maximum absolute value of ground motion on a given channel over a one-second time window. Cua et al. (2009) summarize the required relationships as following:

The first relationship discriminates between the amplitudes of P- and S-waves ( $PS$ ) and accounts for the fact that P-waves show larger amplitudes on the vertical channels, while S-waves have larger amplitudes on the horizontal channels:

$$PS = 0.4 \log_{10}(ZA) + 0.55 \log_{10}(ZV) - 0.46 \log_{10}(HA) - 0.55 \log_{10}(HV) \quad (D.1)$$

$ZA$  and  $ZV$  represent the envelope values of vertical acceleration ( $A$ ) and velocity ( $V$ ).  $HA$  and  $HV$  are the envelope values of root mean square horizontal acceleration and velocity, respectively. If  $PS > -0.1$ , then a P-wave is considered, otherwise an S-wave.

The second relationship defines the ground motion ratio between the vertical acceleration ( $ZA$ ) and vertical displacement ( $ZD$ ) envelope values which correlates with magnitude (Cua, 2005;



**Figure D.1:** Parameterization of ground motion envelope after Cua (2005). It is defined as a combination of P- and S-wave envelope, each characterized by a rise time  $t_r$ , a duration  $dt$ , and an amplitude  $A$ . The decay of the envelopes is determined by the two decay parameters  $\tau$  and  $\gamma$ .

Cua and Heaton, 2007):

$$\begin{aligned}
 ZAD &= 0.36 \log_{10}(ZA) - 0.93 \log_{10}(ZD) \\
 M_{ZAD} &= -1.63 \cdot ZAD + 8.94, & \text{if P-wave, with } \sigma_{M_{ZAD}} &= 0.45 \\
 M_{ZAD} &= -1.46 \cdot ZAD + 8.05, & \text{if S-wave, with } \sigma_{M_{ZAD}} &= 0.41
 \end{aligned} \tag{D.2}$$

$ZAD$  is inverse proportional to the size of the event.  $M_{ZAD}$  is the single-station magnitude estimate and can be calculated as soon as 2 s of P-wave amplitude data are available.

Thirdly, the envelope attenuation relationships (or ground motion models) are defined. Cua (2005) developed a parameterization that decomposes an observed ground motion envelope into P-wave, S-wave, and ambient noise envelopes (Figure D.1). The parameterization has been applied to a set of about 30,000 observed ground motion envelope records of vertical and horizontal acceleration, velocity, and filtered displacement. The records originate from 70 southern California earthquakes in the magnitude range  $2 < M < 8$  which were recorded at stations of the *Southern California Seismic Network (SCSN)* within source-to-site distances of up to 200 km. The developed envelope attenuation relationships can be used to predict the expected ground motion envelope at a certain site as a function of time, given the magnitude and epicentral distance of the earthquake.

The observed ground motion envelope is parameterized as

$$E_{obs}(t) = \sqrt{E_P^2(t) + E_S^2(t) + E_{ambient}^2} + \epsilon \tag{D.3}$$

whereas  $E_P$  is the P-wave envelope and  $E_S$  the envelope of the S-wave and of later arriving phases.  $E_{ambient}$  is the ambient noise at the station and  $\epsilon$  is defined as the difference between the

predicted and observed envelopes.  $E_{ambient}$  is modeled as a constant, while  $E_P$  and  $E_S$  are each described by a rise time ( $t_{rise_P}$ ,  $t_{rise_S}$ ), a constant amplitude ( $A_P$ ,  $A_S$ ), a duration ( $\Delta t_P$ ,  $\Delta t_S$ ), and two decay parameters ( $\tau_P$ ,  $\tau_S$ ;  $\gamma_P$ ,  $\gamma_S$ ) (Figure D.1):

$$E_{i,j}(t) = \begin{cases} 0; & t < T_i \\ \frac{A_{i,j}}{t_{rise_{i,j}}}(t - T_i); & T_i \leq t < T_i + t_{rise_{i,j}} \\ A_{i,j}; & T_i + t_{rise_{i,j}} \leq t < T_i + t_{rise_{i,j}} + \Delta t_{i,j} \\ A_{i,j} \frac{1}{(t - T_i - t_{rise_{i,j}} - \Delta t_{i,j} + \tau_{i,j})^{\gamma_{i,j}}}; & t \geq T_i + t_{rise_{i,j}} + \Delta t_{i,j} \end{cases} \quad (D.4)$$

with  $i$  standing for either P- or S-wave,  $j$  being the  $ZV$ ,  $HA$ ,  $HV$ , and  $HD$  channels, and  $T_i$  being the P- and S-wave arrival times. Altogether, the description of the observed ground motion envelopes is based on 11 envelope parameters (Cua, 2005; Cua and Heaton, 2007).

In order to find the best 11 envelope parameters that fit each observed envelope in the dataset Cua (2005) applied the *Neighbourhood Algorithm (NA)* developed by Sambridge (1999a,b). The *NA* is a direct search method for nonlinear inversion which separates the one nonlinear problem into many small nonlinear inverse problems.

The application of the *NA* resulted into following envelope attenuation relationships for peak vertical and the root mean square of the peak horizontal acceleration, velocity, and displacement amplitudes for both rock and soil sites:

$$\log_{10} \bar{Y}(M, R) = a \cdot M + b \cdot R(M) + d \cdot \log_{10} R(M) + e \quad (D.5)$$

with

$$R(M) = R + C(M)$$

and

$$C(M) = c_1 \cdot \arctan(M - 5) \cdot \exp(c_2 \cdot (M - 5)).$$

$\log_{10} \bar{Y}(M, R)$  represents the specific envelope parameter.  $M$  is the earthquake magnitude and  $R$  is either the epicentral distance in km for events with  $M < 5$  or the *Joyner-Boore* distance for  $M > 5$ . The coefficients  $a$ ,  $b$ ,  $c_1$ ,  $c_2$ ,  $d$ , and  $e$  are defined for the maximum P- and S-wave amplitudes for vertical and horizontal acceleration, velocity, and displacement on both rock and soil sites (Cua et al., 2009). Tables with the coefficients are given in Cua (2005) and Cua and Heaton (2007). The final envelope is obtained by combining all 11 envelope parameters and by adding a constant noise parameter to the envelope.

Finally, the last relationships define the multiple-station magnitude and location estimates:

$$L(M, \text{lat}, \text{lon}) = \sum_{k=1}^n \sum_i^{P,S} L(M, \text{lat}, \text{lon})_{ik}$$

$$L(M, \text{lat}, \text{lon}) = \frac{(ZAD_{ik} - \bar{Z}_i(M))^2}{2\sigma_{ZAD_k}^2} + \sum_{j=1}^4 \frac{(\log_{10} Y_{obs,ijk} - \log_{10} Y_{ijk}(M, \text{lat}, \text{lon}))^2}{2\sigma_{ijk}^2} \quad (D.6)$$

$$\begin{aligned} \bar{Z}(M)_{kP} &= -0.62 \cdot M + 5.50, & \text{with } \sigma_{Z(M)P} &= 0.28 \\ \bar{Z}(M)_{kS} &= -0.69 \cdot M + 5.52, & \text{with } \sigma_{Z(M)S} &= 0.25 \end{aligned}$$

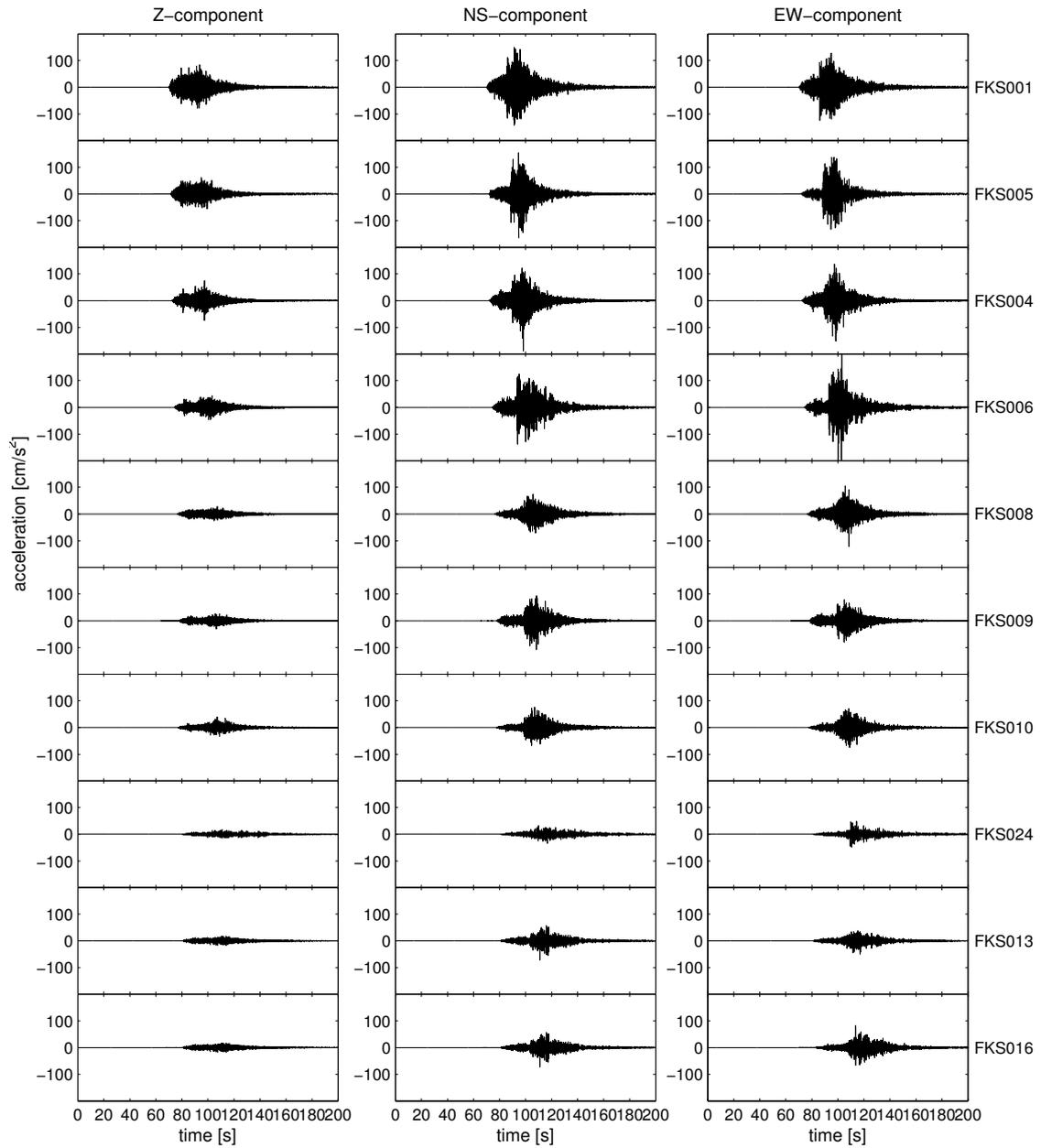
for  $i = \text{P- or S-wave}$ ,  $j = 1, \dots, 4$  for  $ZV$ ,  $HA$ ,  $HV$ , and  $HD$  channels, and  $k = 1, \dots, n$  where  $n$  is the number of stations.  $ZAD$  and  $\log_{10} Y_{ijk}$  are those described in equations D.2 and D.5, respectively. The parameters  $\log_{10} Y_{obs,ijk}$  are available observed envelope amplitudes. The equations for  $\bar{Z}(M)_{kP}$  and  $\bar{Z}(M)_{kS}$  are the inverses of the single-station magnitude relationships  $M_{ZAD}$  in equation D.2.

The magnitude and location coordinates that minimize equation D.6 correspond to the most probable magnitude and location estimates (Cua et al., 2009). The *Virtual Seismologist* method is currently tested in real-time as part of the *California Integrated Seismic Network* early warning project (Cua et al., 2009). Please see Cua (2005), Cua and Heaton (2007), and Cua et al. (2009) for further details about the *Virtual Seismologist* method.

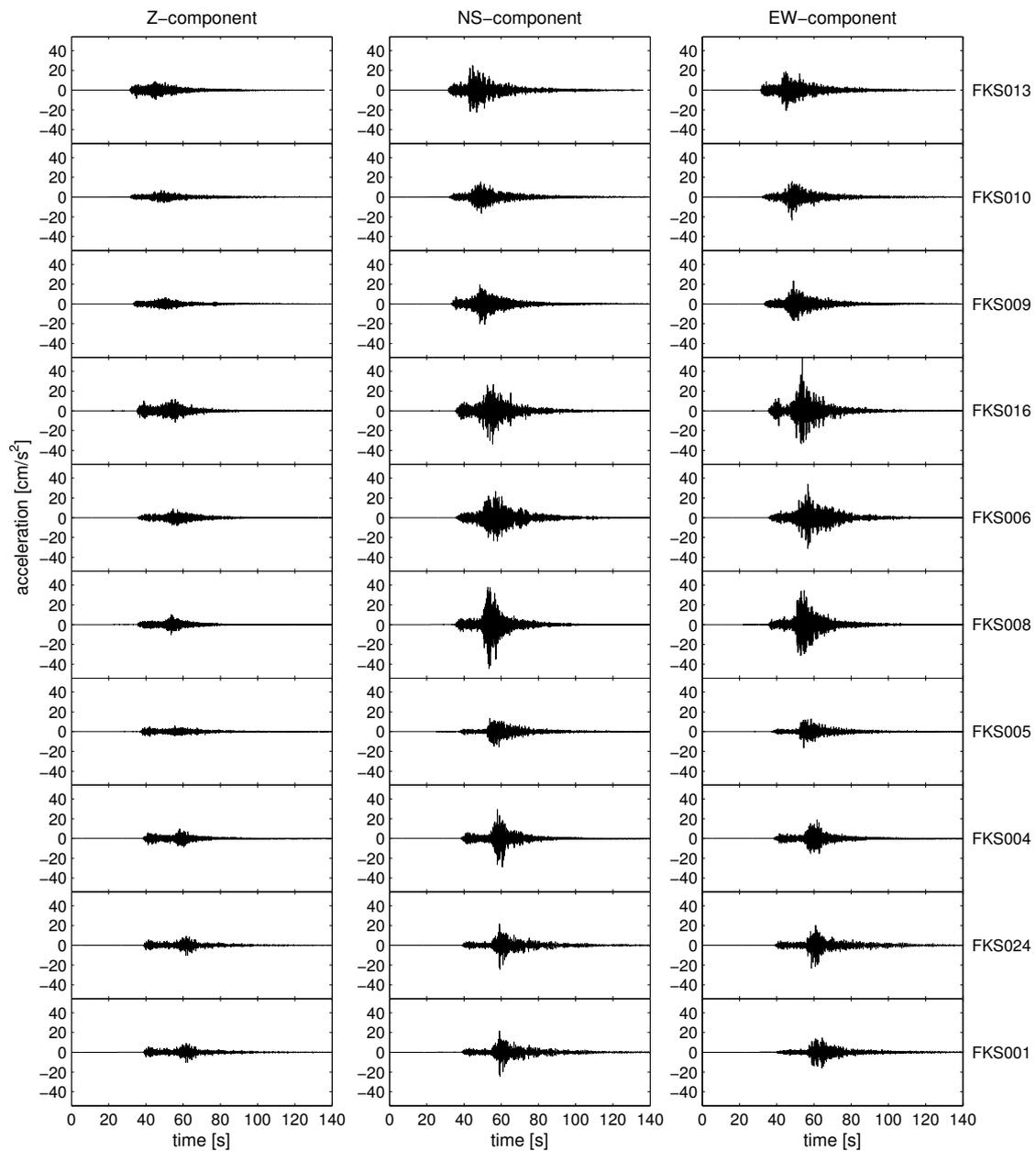
# Appendix E

## Earthquake Records from Japan

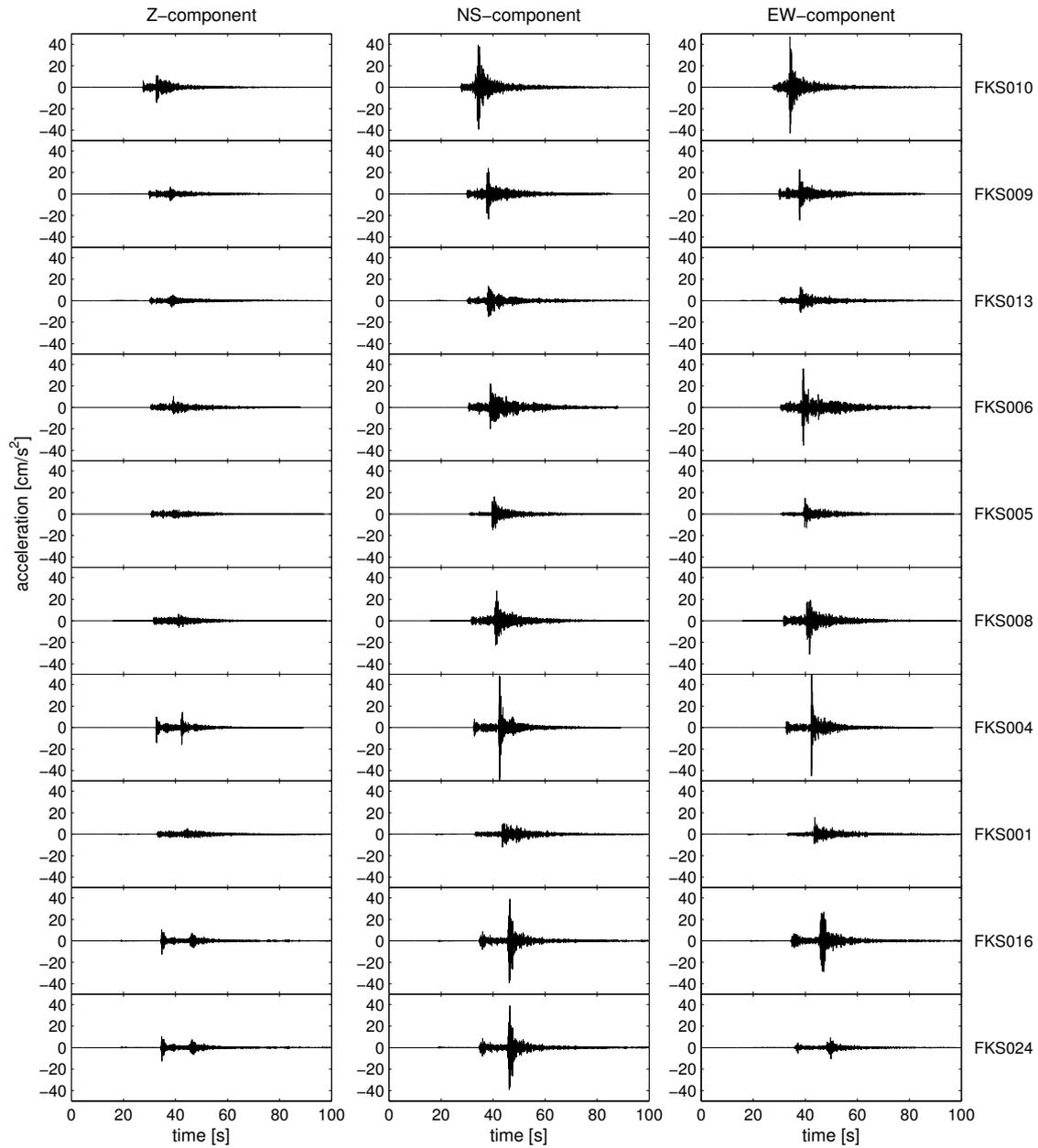
Figures E.1 to E.3 display the acceleration records from the three example earthquakes used in Chapter 5.2 to test the *PreSEIS* performance. Displayed are all three components of ground motion.



**Figure E.1:** Unfiltered ground acceleration records from the  $M_w 7.0$  earthquake located at  $141.68^\circ\text{E}$  and  $38.81^\circ\text{N}$  at a depth of 71 km. The records are sorted by P-wave arrival times at the ten K-NET stations.



**Figure E.2:** Unfiltered ground acceleration records from the  $M_w$ 5.8 earthquake located at  $141.15^\circ\text{E}$  and  $36.39^\circ\text{N}$  at a depth of 49 km. The records are sorted by P-wave arrival times at the ten K-NET stations.



**Figure E.3:** Unfiltered ground acceleration records from the  $M_w 4.9$  earthquake located at  $141.30^\circ\text{E}$  and  $37.00^\circ\text{N}$  at a depth of 50 km. The records are sorted by P-wave arrival times at the ten K-NET stations.

# List of Figures

1.1	Epicentres of major earthquakes ( $M_S \geq 6.8$ ) in the Marmara region since the year 1500 (after Ambraseys (2002)) . . . . .	5
1.2	Map of the northern Marmara Sea basin with bathymetry and fault segments after Armijo et al. (2002) . . . . .	6
2.1	Combined effect of site amplification and high-frequency diminution for site class NEHRP C ( $V_{S30} = 520$ m/s) and different decay parameters $\kappa$ . . . . .	16
2.2	Mean acceleration spectra for seismic intensities III-X after Sokolov (2002) . . . . .	17
2.3	Base excitation of a typical SDOF system of mass $m$ with natural frequency $\omega$ . . . . .	19
2.4	Acceleration response spectrum with 5% damping for a simulated acceleration time series for a $M_w = 7.3$ earthquake at a source-to-site distance of 42 km and site class NEHRP C . . . . .	20
2.5	Stochastic simulation of ground motion time series . . . . .	22
2.6	Shaping window after Saragoni and Hart (1974) . . . . .	23
2.7	Distribution of the 280 simulated earthquakes . . . . .	27
2.8	Definition of fault segments 1 to 5 . . . . .	27
2.9	Moment magnitude and depth histograms of the simulated dataset . . . . .	28
2.10	Example of acceleration time series of an $M_w = 6.2$ earthquake on Segment 3 . . . . .	28
2.11	Observed record (NS-component) of the $M = 7.4$ Kocaeli earthquake compared to the optimized and original <i>FINSIM</i> parameterization . . . . .	31
2.12	Observed record (NS-component) of the $M = 5.8$ earthquake compared to the optimized and original <i>FINSIM</i> parameterization . . . . .	31
2.13	Attenuation of PGA with distance of the optimized and the original simulations, compared to attenuation relationships from the literature . . . . .	32
2.14	Attenuation of PSA at 0.3 s with distance of the optimized and the original simulations, compared to attenuation relationships from the literature . . . . .	33

2.15	Attenuation of PSA at 1.0 s with distance of the optimized and the original simulations, compared to attenuation relationships from the literature . . . . .	33
3.1	Location of the ruptured fault and of the epicentre versions 1 to 3. The <i>NEHRP</i> site classifications are defined for a dense grid of $0.005^\circ \times 0.005^\circ$ spacing. . . . .	37
3.2	Attenuation of PGA with distance for <i>NEHRP</i> B and C site classifications of the simulated ground motion and the attenuation relationships from the literature . . .	38
3.3	Attenuation of PSA at 0.2, 0.3, 0.4, 0.5, 0.55, 0.75, 0.9, and 1.2 s for <i>NEHRP</i> B site classification of the simulated ground motion and the attenuation relationships from the literature . . . . .	39
3.4	Attenuation of PSA at 0.2, 0.3, 0.4, 0.5, 0.55, 0.75, 0.9, and 1.2 s for <i>NEHRP</i> C site classification of the simulated ground motion and the attenuation relationships from the literature . . . . .	39
3.5	Distribution of PGA and seismic intensity for the simulated 1509 Istanbul earthquake. The ground motions are calculated for all three hypocentre locations . . .	41
3.6	Distribution of PGA calculated with the attenuation relationships of Özbey et al. (2004) . . . . .	42
3.7	Distribution of absolute numbers of expected completely damaged buildings for the simulated 1509 Istanbul earthquake. . . . .	45
3.8	Distribution of absolute numbers of expected completely damaged buildings for the simulated 1509 Istanbul earthquake. The damages are based on the ground motions predicted from the attenuation relationships of Özbey et al. (2004). . . . .	46
3.9	Distribution of absolute numbers of expected completely damaged buildings for the $M_w = 7.5$ “worst case” scenario earthquake (American Red Cross - Bogazici University, 2002; Erdik et al., 2003a) . . . . .	46
4.1	Cumulative absolute velocity of a simulated acceleration record . . . . .	54
4.2	(a) Sketch of a Two-Layer Feed-Forward neural network. The neurons are arranged in the input, the hidden, and the output layer. (b) The logistic activation function. . . . .	56
4.3	Logarithm of $CAV + 1$ calculated from each of the ten EEW stations for a simulated example earthquake . . . . .	58
4.4	Sketch of the TLFF neural networks used for estimating the hypocentre location and the magnitude at time step $n$ for earthquake $m$ . . . . .	60
4.5	Absolute errors of hypocentre locations and mean magnitude errors from the application of PreSEIS to the 280 simulated earthquake scenarios . . . . .	61

5.1	Map of Southern California showing the locations of the 15 SCSN station sites and the epicentral locations of the 74 training earthquakes . . . . .	66
5.2	Observed and predicted ground motion envelopes of vertical acceleration from the $M_w$ 4.8 Yorba Linda earthquake at the 15 SCSN stations . . . . .	68
5.3	Mean temporal distribution of the number of triggered stations for all 74 earthquakes and their minima and maxima . . . . .	69
5.5	Number of triggered stations and PreSEIS source parameter estimates for the $M_w$ 7.1 Hector Mine, $M_w$ 4.8 Yorba Linda, and $M_w$ 6.0 Parkfield earthquake . . . . .	71
5.7	Map of Japan with the ten K-NET station sites and epicentral locations of the 69 earthquakes . . . . .	76
5.8	Distribution of moment magnitudes and depths of all 69 earthquakes from Japan used in this study . . . . .	77
5.9	Mean temporal distribution of the number of triggered stations for all 66 training earthquakes and their minima and maxima . . . . .	78
5.10	Absolute errors of hypocentre locations and mean magnitude errors obtained from the training process with the 66 earthquakes for the vertical components, NS-components, and EW-components . . . . .	79
5.11	Number of triggered stations and <i>PreSEIS</i> source parameter estimates for the $M_w$ 7.0, $M_w$ 5.8, and $M_w$ 4.9 earthquake . . . . .	80
5.12	Map of the Marmara region, Turkey, with the locations of the present EEW stations and the additionally planned stations . . . . .	83
5.13	Average number of triggered stations with progressing time for all 280 earthquakes for the present network of 10 stations and for the expanded network of 20 stations . . . . .	84
5.14	Warning times of all 280 earthquake scenarios in the case of the original and the expanded network . . . . .	85
5.15	Absolute hypocentre location errors and mean magnitude errors with progressing time for the network configurations of 10 and 20 stations . . . . .	85
5.16	PreSEIS alert classifications for all 280 earthquakes based on the present EEW network of 10 stations and on the planned expanded network of 20 stations . . . . .	86
5.17	Distribution of earthquake scenarios triggering the alarms together with their seismic intensities at ISTAN for the present network of 10 stations . . . . .	89
5.18	Distribution of earthquake scenarios triggering the alarms together with their seismic intensities at ISTAN for the expanded network of 20 stations . . . . .	89
5.19	Distribution of earthquake scenarios triggering the alarms together with their seismic intensities at ISTAN for the present network of 10 stations with the optimized amplitude thresholds . . . . .	90

5.20	Distribution of votes leading to the three alarms at each station of the current network and of the expanded network . . . . .	90
5.21	Distribution of votes leading to the three alarms at each station of the current network after the application of the current amplitude thresholds and the optimized ones . . . . .	91
5.22	Available warning times for all 280 earthquake scenarios using the IERREWS threshold system with 10 stations (a) and with 20 stations (b) . . . . .	93
5.23	Available warning times for all 280 earthquake scenarios using the original network of 10 stations with the original IERREWS thresholds (a) and the optimized thresholds (b) . . . . .	94
A.1	Results from the application of PreSEIS to both the original as well as the optimized simulations . . . . .	112
C.1	Idealized capacity diagram (spectral acceleration versus spectral displacement) representing elastic perfectly-plastic behaviour after Erdik et al. . . . .	120
C.2	Spectral displacement-based vulnerability curves for reinforced concrete frame, mid-rise, pre-1980 buildings for the different damage levels . . . . .	122
D.1	Parameterization of ground motion envelope after Cua (2005) . . . . .	126
E.1	Unfiltered ground acceleration records from the $M_w$ 7.0 earthquake located at 141.68°E and 38.81°N at a depth of 71 km . . . . .	130
E.2	Unfiltered ground acceleration records from the $M_w$ 5.8 earthquake located at 141.15°E and 36.39°N at a depth of 49 km . . . . .	131
E.3	Unfiltered ground acceleration records from the $M_w$ 4.9 earthquake located at 141.30°E and 37.00°N at a depth of 50 km . . . . .	132

# List of Tables

1.1	Statistics over the ten earthquakes with the largest numbers of fatalities between 1980 and 2008. . . . .	2
2.1	NEHRP site classifications based on $V_{S30}$ . . . . .	15
2.2	Site amplifications at selected frequencies after Boore and Joyner (1997) . . . . .	16
2.3	Coordinates and NEHRP site classifications of the early warning stations of IER-REWS . . . . .	26
2.4	Observed earthquakes selected for the <i>FINSIM</i> optimization . . . . .	30
3.1	Simulation parameters for the historic 1509 Istanbul earthquake. . . . .	36
3.2	Absolute numbers and percentage of estimated building damages for the application of the historic 1509 Istanbul earthquake to today's Istanbul building stock . . . . .	44
5.1	Distribution of alarms using the original network of 10 stations, the expanded network of 20 stations, and the optimized amplitude thresholds . . . . .	88
5.2	Alarm classification of all 280 earthquake scenarios on the basis of alarm 3 of the threshold system . . . . .	92
A.1	FINSIM input parameters for the 280 simulated earthquakes . . . . .	104
A.2	Description of Segment 1 to 5 . . . . .	105
A.3	Description of the simulated earthquakes on Segment 1 . . . . .	106
A.4	Description of the simulated earthquakes on Segment 2 . . . . .	107
A.5	Description of the simulated earthquakes on Segment 3 . . . . .	108
A.6	Description of the simulated earthquakes on Segment 4 . . . . .	109
A.7	Description of the simulated earthquakes on Segment 5 . . . . .	110

---

A.8	Description of the simulated earthquakes on Segment 6 . . . . .	111
B.1	Model coefficients of equation B.1 for estimating the geometric mean of horizontal peak ground acceleration ( <i>PGA</i> ) and 5% damped spectral acceleration (both in g) for selected natural periods after Campbell and Bozorgnia (2008). . . . .	115
B.2	Attenuation model coefficients of equation B.15 for estimating the geometric mean of horizontal peak ground acceleration ( <i>PGA</i> ) and 5% damped spectral acceleration (both in $\text{cm/s}^2$ ) for selected natural periods after Özbey et al. (2004). . . . .	116
B.3	Regression coefficients of equation B.16 for estimating peak horizontal acceleration ( <i>PGA</i> ) and 5% damped pseudo-acceleration response spectra (both in g) for selected natural periods after Boore et al. (1997). . . . .	118
C.1	Building capacity parameters after Erdik et al. . . . .	121
C.2	Spectral displacement-based fragility curve data after the KOERILoss method by Erdik et al. . . . .	123

# Bibliography

- Aki, K. (1967). Scaling law of seismic spectrum. *J. Geophys. Int.*, 72:1217–1231.
- Aki, K. and Richards, P. (1980). *Quantitative Seismologie, Theory and Methods*. W.H. Freeman and Company, San Francisco, ISBN: 0-7167-1058-7.
- Akkar, S. and Gülkan, P. (2002). A critical examination of near-field accelerograms from the Sea of Marmara region earthquakes. *Bull. Seism. Soc. Am.*, 92(1):428–447.
- Alcik, H., Ozel, O., Apaydin, N., and Erdik, M. (2009). A study on warning algorithms for Istanbul earthquake early warning system. *Geophys. Res. Lett.*, 36:L00B05.
- Allen, R., Brown, H., Hellweg, M., Khainovski, O., Lombard, P., and Neuhauser, D. (2009a). Real-time earthquake detection and hazard assessment by ElarmS across California. *Geophys. Res. Lett.*, 36:L00B08.
- Allen, R., Gasparini, P., Kamigaichi, O., and Böse, M. (2009b). The status of earthquake early warning around the world: an introductory overview. *Seismol. Res. Lett.*, 80(5):682–693.
- Ambraseys, N. (2001). The earthquake of 1509 in the Sea of Marmara, Turkey, revisited. *Bull. Seism. Soc. Am.*, 91(6):1397–1416.
- Ambraseys, N. (2002). The seismic activity of the Marmara Sea Region over the last 2000 years. *Bull. Seism. Soc. Am.*, 92(1):1–18.
- Ambraseys, N. and Douglas, J. (2003). Near-field horizontal and vertical earthquake ground motions. *Soil Dyn. Earthq. Eng.*, 23:1–18.
- Ambraseys, N. and Finkel, C. (1990). The Marmara Sea earthquake of 1509. *Terra Nova*, 2(2):167–174.
- Ambraseys, N. and Finkel, C. (1991). Long-term seismicity of Istanbul and of the Marmara Sea region. *Terra Nova*, 3(5):527–539.
- Ambraseys, N. and Jackson, J. (2000). Seismicity of the Sea of Marmara (Turkey) since 1500. *Geophys. J. Int.*, 141:F1–F6.

- Ambraseys, N., Simpson, K., and Bommer, J. (1996). Prediction of horizontal response spectra in Europe. *Earthquake Engineering and Structural Dynamics*, 25:371–400.
- American Red Cross - Bogazici University (2002). *Earthquake risk assessment for Istanbul metropolitan area*. Kandilli Observatory and Earthquake Research Institute, Istanbul.
- Anderson, J. (2003). *Strong-motion Seismology*. in: International Handbook of Earthquake and Engineering Seismology (Part B), 937-965, Academic Press, Eds. W.H.K. Lee, H. Kanamori, P.C. Jennings, C. Kisslinger, ISBN: 0-12-440658-0.
- Anderson, J. and Hough, S. (1984). A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies. *Bull. Seism. Soc. Am.*, 74(5):1969–1993.
- Armijo, R., Meyer, B., Navarro, S., King, G., and Barka, A. (2002). Asymmetric slip partitioning in the Sea of Marmara pull-apart: a clue to propagation processes of the North Anatolian Fault? *Terra Nova*, 13(6):80–86.
- Atkinson, G. and Boore, D. (1995). Ground-motion relations for eastern North America. *Bull. Seism. Soc. Am.*, 85(1):17–30.
- Bakir, B., Sucuoglu, H., and Yilmaz, T. (2002). An overview of local site effects and the associated building damage in Adapazari during the 17 August 1999 Izmit earthquake. *Bull. Seism. Soc. Am.*, 92(1):509–526.
- Bard, P.-Y. and Tucker, B. (1985). Underground and ridge site effects: a comparison of observation and theory. *Bull. Seism. Soc. Am.*, 75(4):905–922.
- Benjamin, J. and Associates (1988). A criterion for determining exceedance of the operating basis earthquake. Technical report, Electric Power Research Institute, Palo Alto, California.
- Beresnev, I. and Atkinson, G. (1997). Modeling finite-fault radiation from the  $\omega^n$  spectrum. *Bull. Seism. Soc. Am.*, 87(1):67–84.
- Beresnev, I. and Atkinson, G. (1998). Finsim - a fortran program for simulating stochastic acceleration time histories from finite faults. *Seismol. Res. Lett.*, 69(1):27–32.
- Bishop, C. (1995). *Neural networks for pattern recognition*. Clarendon Press, Oxford.
- Boore, D. (1983). Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra. *Bull. Seism. Soc. Am.*, 73(6):1865–1894.
- Boore, D. (2003). Simulation of ground motion using the stochastic method. *Pure appl. geophys.*, 160:635–676.
- Boore, D. and Joyner, W. (1997). Site amplifications for generic rock sites. *Bull. Seism. Soc. Am.*, 87(2):327–341.

- Boore, D., Joyner, W., and Fumal, T. (1997). Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: A summary of recent work. *Seismol. Res. Lett.*, 68(1):128–153.
- Böse, M. (2006). *Earthquake Early Warning for Istanbul using Artificial Neural Networks*. PhD thesis, Universität Karlsruhe (TH).
- Böse, M., Hauksson, E., Solanki, K., Kanamori, H., and Heaton, T. (2009). Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008  $m_w$ 5.4 Chino Hills earthquake. *Geophys. Res. Lett.*, 36:L00B03.
- Böse, M., Ionescu, C., and Wenzel, F. (2007). Earthquake early warning for Bucharest, Romania: Novel and revised scaling relations. *Geophys. Res. Lett.*, 34:L07302.
- Böse, M., Wenzel, F., and Erdik, M. (2008). Preseis: A neural network-based approach to earthquake early warning for finite faults. *Bull. Seism. Soc. Am.*, 98(1):366–382.
- Brown, H., Allen, R., and Grasso, V. (2009). Testing ElarmS in Japan. *Seismol. Res. Lett.*, 80(5):727–739.
- Brune, J. (1970). Tectonic stress and the spectra of seismic shear waves from earthquakes. *J. Geophys. Res.*, 75:4997–5009.
- Brune, J. (1971). Correction. *J. Geophys. Res.*, 76:5002.
- Building Seismic Safety Council (1997). *NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures, FEMA-302 and 303*. Prepared for the Federal Emergency Management Agency, Washington, D.C.
- Campbell, K. (2003). *Strong-Motion Attenuation Relations*. in: International Handbook of Earthquake and Engineering Seismology (Part B), 1003-1012, Academic Press, Eds. W.H.K. Lee, H. Kanamori, P.C. Jennings, C. Kisslinger, ISBN: 0-12-440658-0.
- Campbell, K. and Bozorgnia, Y. (2008). NGA ground motion model for the geometric mean horizontal component of PGA, PGV, PGD and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10 s. *Earthquake Spectra*, 24(1):139–171.
- Chávez-García, F., Sánchez, L., and Hatzfeld, D. (1996). Topographic site effects and HVSR. A comparison between observations and theory. *Bull. Seism. Soc. Am.*, 86(5):1559–1573.
- Chen, W.-F. and Scawthorn, C. (2003). *Earthquake Engineering Handbook*. CRC Press, ISBN: 0-8493-0068-1.
- Chernov, Y. and Sokolov, V. (1999). Correlation of seismic intensity with Fourier acceleration spectra. *Phys. Chem. Earth (A)*, 24(6):523–528.
- Cua, G. (2005). *Creating the Virtual Seismologist: developments in ground motion characterization and seismic early warning*. PhD thesis, California Institute of Technology.

- Cua, G., Fischer, M., Heaton, T., and Wiemer, S. (2009). Real-time Performance of the Virtual Seismologist Earthquake Early Warning Algorithm in Southern California. *Seismol. Res. Lett.*, 80(5):740–747.
- Cua, G. and Heaton, T. (2007). *The Virtual Seismologist (VS) method: a Bayesian approach to earthquake early warning*. in: Earthquake early warning systems, 97-132, Springer Verlag, Eds. P. Gasparini, G. Manfredi, J. Zschau, ISBN: 978-3-540-72240-3.
- Ellsworth, W. and Beroza, G. (1995). Seismic evidence for an earthquake nucleation phase. *Science*, 268:851–855.
- Erdik, M., Aydinoglu, N., Fahjan, Y., Sesetyan, K., Demircioglu, M., Siyahi, B., Durukal, E., Özbey, C., Biro, Y., Akman, H., and Yuzugullu, O. (2003a). Earthquake risk assessment for Istanbul metropolitan area. *Earthquake Engineering and Engineering Vibration*, 2(1):1–23.
- Erdik, M., Fahjan, Y., Ozel, O., Alcik, H., Mert, A., and Gul, M. (2003b). Istanbul Earthquake Rapid Response and the Early Warning System. *Bulletin of Earthquake Engineering*, 1:157–163.
- Espinosa-Aranda, J., Cuellar, A., Garcia, A., Ibarrola, G., Islas, R., Maldonado, S., and Rodriguez, F. (2009). Evolution of the Mexican Seismic Alert System (SASMEX). *Seismol. Res. Lett.*, 80(5):694–706.
- Espinosa-Aranda, J., Jimenez, A., Ibarolla, G., Alcantar, F., Aguilar, A., Inostroza, M., and Maldonado, S. (1995). Mexico City Seismic Alert System. *Seismol. Res. Lett.*, 66(6):42–53.
- Festa, G., Zollo, A., and Lancieri, M. (2008). Earthquake magnitude estimation from early radiated energy. *Geophys. Res. Lett.*, 35:L22307.
- Fleming, K., Picozzi, M., Milkereit, C., Kühnlenz, F., Lichtblau, B., Fischer, J., Zulfikar, C., Özel, O., and the SAFER / EDIM working groups (2009). The Self-organizing Seismic Early Warning Information Network (SOSEWIN). *Seismol. Res. Lett.*, 80(5):755–771.
- Goltz, J. (2002). Introducing Earthquake Early Warning in California: A Summary of Social Science and Public Policy Issues. A report to OES and the Operational Areas, Caltech Seismological Laboratory.
- Graizer, V., Shakal, A., Scrivner, C., Hauksson, E., Polet, J., and Jones, L. (2002). TriNet strong-motion data from the M7.1 Hector Mine, California, earthquake of 16 October 1999. *Bull. Seism. Soc. Am.*, 92(4):1525–1542.
- Gülkan, P. and Kalkan, E. (2002). Attenuation modeling of recent earthquakes in Turkey. *Journal of Seismology*, 6(3):397–409.
- Gülkan, P. and Kalkan, E. (2004). Attenuation characteristics of Turkey based on recent strong motion data. *ARI - The Bulletin of the Istanbul Technical University*, 54(2):45–51.

- Hanks, T. and Kanamori, H. (1979). A moment magnitude scale. *J. Geophys. Res.*, 84(B5):2348–2350.
- Hanks, T. and McGuire, R. (1981). The character of high-frequency strong ground motion. *Bull. Seism. Soc. Am.*, 71(6):2071–2095.
- Harben, P. (1981). Earthquake Alert System Feasibility Study. Technical report, Lawrence Livermore National Laboratory, UCRL-LR-109625.
- Hartzell, S. (1978). Earthquake aftershocks as Green's functions. *Geophys. Res. Lett.*, 5:1–4.
- Hauksson, E., Hutton, K., Jones, L., and Given, D. (2002). The M4.8 Yorba Linda, Orange County Earthquake of 3 September 2002. *California Integrated Seismic Network preliminary report*, 92:2 pps.
- Heaton, T. (1985). A Model for a Seismic Computerized Alert Network. *Science*, 228:987–990.
- Hergert, T. (2009). *Numerical modelling of the absolute stress state in the Marmara Sea region - a contribution to seismic hazard assessment*. PhD thesis, Universität Karlsruhe (TH).
- Horiuchi, S., Negishi, H., Abe, K., Kamimura, A., and Fujinawa, Y. (2005). An automatic processing system for broadcasting earthquake alarms. *Bull. Seism. Soc. Am.*, 95(2):708–718.
- Hoshiaba, M., Kamigaichi, O., Saito, M., Tsukada, S., and Hamada, N. (2008). Earthquake Early Warning Starts Nationwide in Japan. *EOS, Transactions, American Geophysical Union*, 89(8):73–74.
- Hsiao, N.-C., Wu, Y.-M., Shin, T.-C., Zhao, L., and Teng, T.-L. (2009). Development of earthquake early warning system in Taiwan. *Geophys. Res. Lett.*, 36:L00B02.
- Irikura, K. (1983). Semi-empirical estimation of strong ground motions during large earthquakes. *Bulletin of the Disaster Prevention Research Institute*, 33:63–104.
- Joyner, W. and Boore, D. (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bull. Seism. Soc. Am.*, 71(6):2011–2038.
- Joyner, W., Warrick, R., and Fumal, T. (1981). The effect of Quaternary alluvium on strong ground motion in the Coyote Lake, California, earthquake of 1979. *Bull. Seism. Soc. Am.*, 71(4):1333–1349.
- Kamigaichi, O. (2004). Jma Earthquake Early Warning. *Journal of Japan Association for Earthquake Engineering*, 4(3):134–137.
- Kamigaichi, O., Saito, M., Doi, K., Matsumori, T., Tsukada, S., Takeda, K., Shimoyama, T., Nakamura, K., Kiyomoto, M., and Watanabe, Y. (2009). Earthquake Early Warning in Japan: Warning the General Public and Future Prospects. *SRL*, 80(5):717–726.

- Kanamori, H. (2005). Real-time seismology and earthquake damage mitigation. *Annu. Rev. Earth Planet. Sci.*, 33:195–214.
- Kawase, H. (2003). *Site Effects on Strong Ground Motions*. in: International Handbook of Earthquake and Engineering Seismology (Part B), 1013-1030, Academic Press, Eds. W.H.K. Lee, H. Kanamori, P.C. Jennings, C. Kisslinger, ISBN: 0-12-440658-0.
- Kinoshita, S. (2003). *Kyoshin Net (K-NET), Japan*. in: International Handbook of Earthquake and Engineering Seismology (Part B), 1049-1056, Academic Press, Eds. W.H.K. Lee, H. Kanamori, P.C. Jennings, C. Kisslinger, ISBN: 0-12-440658-0.
- Köhler, N., Cua, G., Wenzel, F., and Böse, M. (2009). Rapid source parameter estimations of southern California earthquakes using PreSEIS. *Seismol. Res. Lett.*, 80(5):748–754.
- Lancieri, M. and Zollo, A. (2008). A Bayesian approach to the real-time estimation of magnitude from the early P and S wave displacement peaks. *J. Geophys. Res.*, 113:B12302.
- Lay, T. and Wallace, T. (1995). *Modern Global Seismology*. International Geophysics Series, Vol. 58, Academic Press, ISBN: 0-12-732870-X.
- Le Pichon, X., Seng-Åúr, A., Demirbag, E., Rangin, C., Imren, C., Armijo, R., Görür, N., Cagatay, N., Mercier de Lepinay, B., Meyer, B., Saatçilar, R., and Tok, B. (2001). The active Main Marmara Fault. *Earth Planet. Sci. Lett.*, 192:595–616.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly Journal of Applied Mathematics*, 2:164–168.
- Lockman, A. and Allen, R. (2007). Magnitude-period scaling relations for Japan and the Pacific Northwest: Implications for earthquake early warning. *Bull. Seism. Soc. Am.*, 97(1B):140–150.
- Munich Re Group (2009). *Topics Geo. Natural catastrophes 2008. Analyses, assessments, positions*. Münchener Rückversicherungs-Gesellschaft, <http://www.munichre.com/>.
- Nakamura, Y. and Saita, J. (2007). *UrEDAS, the earthquake warning system: Today and tomorrow*. in: Earthquake early warning systems, 249-281, Springer Verlag, Eds. P. Gasparini, G. Manfredi, J. Zschau, ISBN: 978-3-540-72240-3.
- Odaka, T., Ashiya, T., Tsukada, S., Sato, S., Ohtake, K., and Nozaka, D. (2003). A new method of quickly estimating epicentral distance and magnitude from a single seismic record. *Bull. Seism. Soc. Am.*, 93(1):526–532.
- Olson, E. and Allen, R. (2005). The deterministic nature of earthquake rupture. *Nature*, 438:212–215.
- Özbey, C., Sari, A., Manuel, L., Erdik, M., and Fahjan, Y. (2004). An empirical attenuation relationship for Northwestern Turkey ground motion using a random effects approach. *Soil Dyn. Earthq. Eng.*, 24:115–125.

- Özel, O., Cranswick, E., Meremonte, M., Erdik, M., and Safak, E. (2002). Site effects in Av-cilar, West of Istanbul, Turkey, from strong- and weak-motion data. *Bull. Seism. Soc. Am.*, 92(1):499–508.
- Parsons, T. (2004). Recalculated probability of  $M \geq 7$  earthquakes beneath the Sea of Marmara, Turkey. *J. Geophys. Res.*, 109:B05304.
- Parsons, T., Toda, S., Stein, R., Barka, A., and Dieterich, J. (2000). Heightened odds of large earthquakes near Istanbul: an interaction-based probability calculation. *Science*, 288:661–665.
- Paz, M. (1994). *International Handbook of Earthquake Engineering*. Chapman & Hall, New York.
- Rydelek, P. and Horiuchi, S. (2006). Is earthquake rupture deterministic? *Nature*, 441:E5–E6.
- Sambridge, M. (1999a). Geophysical inversion with a Neighbourhood Algorithm - I. Searching a parameter space. *Geophys. J. Int.*, 138:479–494.
- Sambridge, M. (1999b). Geophysical inversion with a Neighbourhood Algorithm - II. Appraising the ensemble. *Geophys. J. Int.*, 138:727–746.
- Saragoni, G. and Hart, G. (1974). Simulation of artificial earthquakes. *Earthq. Eng. Struct. Dyn.*, 2:249–267.
- Satriano, C., Lomax, A., and Zollo, A. (2008). Real-time evolutionary earthquake location for seismic early warning. *Bull. Seism. Soc. Am.*, 98(3):1482–1494.
- Singh, S., Mena, E., and Castro, R. (1988). Some aspects of source characteristics of the 19 September 1985 Michoacan earthquake and ground motion amplification in and near Mexico City from strong motion data. *Bull. Seism. Soc. Am.*, 78(2):451–477.
- Sokolov, V. (2002). Seismic intensity and Fourier acceleration spectra: Revised relationship. *Earthquake Spectra*, 18(1):161–187.
- Stein, S. and Wysession, M. (2003). *An Introduction to Seismology, Earthquakes, and Earth Structure*. Blackwell Publishing, ISBN: 0-86542-078-5.
- Suárez, G., Novelo, D., and Mansilla, E. (2009). Performance Evaluation of the Seismic Alert System (sas) in Mexico City: A Seismological and a Social Perspective. *Seismol. Res. Lett.*, 80(5):707–716.
- Weber, E., Convertito, V., Iannaccone, G., Zollo, A., Bobbio, A., Cantore, L., Corciulo, M., Di Crosta, M., Elia, L., Martino, C., Romeo, A., and Satriano, C. (2007). An advanced seismic network in the southern Apennines (Italy) for seismicity investigations and experimentation with earthquake early warning. *Seismol. Res. Lett.*, 78(6):622–634.

- Wenzel, F., Bonjer, K., Fiedrich, F., Lungu, D., Marmureanu, G., and Wirth, W. (2004). Real-time earthquake information systems. In *2nd Indo-German Workshop on Seismic Safety of Structures, Risk Assessment and Disaster Mitigation, February 16-17, 2004, IIT Madras, Chennai, India*.
- Wenzel, F., Erdik, M., Köhler, N., Zschau, J., Milkereit, C., Picozzi, M., Fischer, J., Redlich, J., Kühnlenz, F., Lichtblau, B., Eveslage, I., Christ, I., Lessing, R., and Kiehle, C. (2009). EDIM - Earthquake Disaster Information System for the Marmara Region, Turkey. In Stroink, E. L., editor, *GEOTECHNOLOGIEN Science Report - Early Warning Systems in Earth Management. Abstracts of the Status Seminar, Munich, 12-13 October*, pages 84–95. Koordinierungsbüro GEOTECHNOLOGIEN, Potsdam.
- Wenzel, F., Oncescu, M., Baur, M., Fiedrich, F., and Ionescu, C. (1999). An Early Warning System for Bucharest. *Seismol. Res. Lett.*, 70(2):161–169.
- Wolfe, C. (2006). On the properties of predominant-period estimators for earthquake early warning. *Bull. Seism. Soc. Am.*, 96(5):1961–1965.
- Wu, Y.-M. and Teng, T. (2002). A Virtual Subnetwork Approach to Earthquake Early Warning. *Bull. Seism. Soc. Am.*, 92(5):2008–2018.
- Wurman, G., Allen, R., and Lombard, P. (2007). Toward earthquake early warning in northern California. *J. Geophys. Res.*, 112:B08311.
- Yamada, T. and Ide, S. (2008). Limitation of the predominant-period estimator for earthquake early warning and the initial rupture of earthquakes. *Bull. Seism. Soc. Am.*, 98(6):2739–2745.
- Zollo, A., Iannaccone, G., Lancieri, M., Cantore, L., Convertito, V., Emolo, A., Festa, G., Gallovic, F., Vassallo, M., Martino, C., Satriano, C., and Gasparini, P. (2009). Earthquake early warning system in southern Italy: Methodologies and performance evaluation. *Geophys. Res. Lett.*, 36:L00B07.
- Zollo, A., Lancieri, M., and Nielsen, S. (2006). Earthquake magnitude estimation from peak amplitudes of very early seismic signals on strong motion records. *Geophys. Res. Lett.*, 33:L23312.

## Used Software

The simulations of seismic ground motions in this work were performed using a modified version of *FINSIM* (Beresnev and Atkinson, 1997, 1998) written in Fortran 90. The neural network computations as well as any other data processing and computations were performed with MATLAB (The MathWorks, Inc.).

All maps were produced using The Generic Mapping Tools (GMT). The further plots were made using MATLAB and Xfig.

This thesis was written using the document preparation system L<sup>A</sup>T<sub>E</sub>X.



## Danksagung / Acknowledgments

Ich danke Prof. Dr. Friedemann Wenzel für seine Betreuung und dafür, mir die Möglichkeit zu dieser interessanten Arbeit gegeben zu haben. Die fortwährenden Möglichkeiten, an internationalen Projekttreffen, Konferenzen und Workshops teilzunehmen, haben nicht nur zu meiner fachlichen Qualifikation beigetragen, sondern haben es mir auch erlaubt, viele interessante und wertvolle internationale Kontakte zu knüpfen.

I thank Prof. Dr. Mustafa Erdik for being the co-reviewer of my thesis, which I appreciate very much. The close cooperation with him and the kind supply of data and valuable information contributed considerably to the success of this work. I would also like to thank his working group for answering many questions, most of all Cem Yenidoğan, Hakan Alçık, and Aydın Mert.

Großer Dank gebührt Maren Böse. Sie hat mich nicht nur zu Beginn meiner Doktorarbeit geduldig eingearbeitet, sondern war während der drei Jahre auch aus großer Ferne immer eine verlässliche Hilfe für mich. Ihr offenes Ohr und ihre Anregungen haben sehr zum Gelingen dieser Arbeit beigetragen. Nicht zu vergessen möchte ich ihr auch sehr für das Korrekturlesen dieser Arbeit danken.

Ich danke Tanja und Eike für ihre fortwährende Unterstützung, sowohl bei meiner Arbeit als auch moralisch, für die vielen hilfreichen Diskussionen und für das Korrekturlesen dieser Arbeit. Ich hatte immer sehr viel Spaß mit Euch und werde mich immer gerne an unsere nette Gemeinschaft in unserem Büro erinnern. Meinem Kollegen Joachim möchte ich für seine stete Bereitschaft zur Hilfe danken, und für die vielen unterhaltsamen und verlässlichen Kaffeepausen, auf die ich mich jeden Tag gefreut habe. Adrienne danke ich für seine Beiträge und die wertvollen Diskussionen - die Zusammenarbeit mit ihm hat mir sehr viel Spaß gemacht. Danken möchte ich ihm auch für seine viele unterhaltsame Gesellschaft auf zahlreichen Projekttreffen und Tagungen.

I would like to thank Dr. Georgia Cua for the valuable cooperation and for providing me with her dataset from Southern California which was a great contribution to this work. The time at ETH Zurich was a pleasure for me.

Ich danke Tobias Hergert für die Bereitstellung seiner Modelldaten bzgl. meiner Simulation des 1509-Bebens in Istanbul.

Außerdem möchte ich all meinen Kollegen aus dem EDIM-Projekt für die interessante und wertvolle Zusammenarbeit danken, die mir viel Freude bereitet hat.

Zu guter Letzt möchte ich mich bei allen Mitarbeiterinnen und Mitarbeitern des Geophysikalischen Instituts für die schöne Zeit und die netten Plaudereien in den Kaffeepausen bedanken, und

außerdem für die zusätzliche fachliche Unterstützung, die der eine oder andere mir hat zukommen lassen. Mein besonderer Dank gilt dem Sekretariat und Petra Knopf, die immer bereitwillig und schnell mit Hilfe zur Stelle waren.