Institut für Erd- und Umweltwissenschaften
Allgemeine Geophysik

Ground-motion model selection and adjustment for seismic hazard analysis

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

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
Annabel Händel

Potsdam im Oktober 2018

Erstbetreuer: Prof. Dr. Frank Scherbaum
Zweitbetreuer: apl. Prof. Dr. Frank Krüger
Externer Gutachter: Prof. Dr. Donat Fäh

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-418123
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-418123

Eidesstattliche Erklärung

Ich versichere, dass ich die vorliegende Arbeit selbständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen oder anderen Quellen entnommen sind, sind als solche eindeutig kenntlich gemacht.

Die Arbeit ist in gleicher oder ähnlicher Form noch nicht veröffentlicht und noch keiner Prüfungsbehörde vorgelegt worden.

Potsdam, den 15. Oktober 2018 .....

Annabel Händel

Zusammenfassung

Erdbeben können starke Bodenbewegungen erzeugen und es ist wichtig, diese in einer seismischen Gefährdungsanalyse korrekt vorherzusagen. Üblicherweise werden dazu empirisch ermittelte Bodenbewegungsmodelle (GMPE) in einem logischen Baum zusammengefügt. Wenn jedoch die Bodenbewegung in einem Gebiet mit geringer Seismizität bestimmen werden soll, dann fehlen in der Regel die Daten, um regionsspezifische GMPEs zu entwickeln. In diesen Fällen ist es notwendig, auf Modelle aus anderen Gebieten mit guter Datengrundlage zurückzugreifen und diese an die Zielregion anzupassen. Zur korrekten Anpassung werden seismologische Informationen aus der Zielregion wie beispielsweise die standortspezifische Dämpfung κ_0 benötigt. Diese Parameter lassen sich jedoch ebenfalls nur unzuverlässig bestimmen, wenn die Datengrundlage schlecht ist.

In meiner Dissertation beschäftige ich mich daher mit der Auswahl von GMPEs für den logischen Baum beziehungsweise deren Anpassung an Regionen mit geringer Seismizität. Ich folge dabei zwei verschiedenen Strategien.

Im ersten Ansatz geht es um das Aufstellen eines logischen Baumes, falls kein regionsspezifisches Modell vorhanden ist. Ich stelle eine Methode vor, in der mehrere regionsfremde Modelle zu einem Mixmodell zusammengefügt werden. Die Modelle werden dabei je nach ihrer Eignung gewichtet und die Gewichte mittels der wenigen verfügbaren Daten aus der Zielregion ermittelt. Ein solches Mixmodell kann als sogenanntes 'Backbone'-Modell verwendet werden, welches in der Lage ist, mittlere Bodenbewegungen in der Zielregion korrekt vorherzusagen. Ich teste diesen Ansatz für Nordchile und acht GMPEs, die für verschiedene Subduktionszonen auf der Welt entwickelt wurden. Die Resultate zeigen, dass das Mixmodell bessere Ergebnisse liefert als die einzelnen GMPEs, die zu seiner Erzeugung genutzt wurden. Es ist außerdem ebenso gut in der Vorhersage von Bodenbewegungen wie ein Regressionsmodell, welches extra für Nordchile entwickelt wurde.

Im zweiten Ansatz beschäftige ich mich mit der Bestimmung der standortspezifischen Dämpfung κ_0. κ_0 ist einer der wichtigsten Parameter zur Anpassung eines GMPEs an eine andere Region. Mein Ziel ist es, κ_0 aus seismischer Bodenunruhe anstelle von Erdbeben zu ermitteln, da diese kontinuierlich aufgezeichnet wird. Mithilfe von Interferometrie kann die Geschwindigkeit und Dämpfung von seismischen Wellen im Untergrund bestimmt werden. Dazu werden lange Aufzeichnungsreihen seismischer Bodenunruhe entweder kreuzkorreliert oder entfaltet (Dekonvolution). Die Bestimmung der Dämpfung aus Bodenunruhe bei Frequenzen über 1 Hz und in geringen Tiefen ist jedoch nicht trivial. Ich zeige in meiner Dissertation die Ergebnisse von zwei Studien. In der ersten Studie wird die Dämpfung von Love-Wellen zwischen 1-4 Hz für ein kleines Testarray in Griechenland ermittelt. In der zweiten Studie verwende ich die Daten einer Bohrloch und einer Oberflächenstation aus dem Vogtland, um die Dämpfung von S-Wellen zwischen 5-15 Hz zu bestimmen. Diese beiden Studien stellen jedoch nur den Ausgangspunkt für zukünftige Untersuchungen dar, in denen κ_0 direkt aus der seismischer Bodenunruhe hergeleitet werden soll.

Abstract

The prediction of the ground shaking that can occur at a site of interest due to an earthquake is crucial in any seismic hazard analysis. Usually, empirically derived ground-motion prediction equations (GMPEs) are employed within a logic-tree framework to account for this step. This is, however, challenging if the area under consideration has only low seismicity and lacks enough recordings to develop a region-specific GMPE. It is then usual practice to adapt GMPEs from data-rich regions (host area) to the area with insufficient ground-motion recordings (target area). Host GMPEs must be adjusted in such a way that they will capture the specific ground-motion characteristics of the target area. In order to do so, seismological parameters of the target region have to be provided as, for example, the site-specific attenuation factor κ_0. This is again an intricate task if data amount is too sparse to derive these parameters.

In this thesis, I explore methods that can facilitate the selection of non-endemic GMPEs in a logic-tree analysis or their adjustment to a data-poor region. I follow two different strategies towards this goal.

The first approach addresses the setup of a ground-motion logic tree if no indigenous GMPE is available. In particular, I propose a method to derive an optimized backbone model that captures the median ground-motion characteristics in the region of interest. This is done by aggregating several foreign GMPEs as weighted components of a mixture model in which the weights are inferred from observed data. The approach is applied to Northern Chile, a region for which no indigenous GMPE existed at the time of the study. Mixture models are derived for interface and intraslab type events using eight subduction zone GMPEs originating from different parts of the world. The derived mixtures provide satisfying results in terms of average residuals and average sample log-likelihoods. They outperform all individual non-endemic GMPEs and are comparable to a regression model that was specifically derived for that area.

The second approach is concerned with the derivation of the site-specific attenuation factor κ_0. κ_0 is one of the key parameters in host-to-target adjustments of GMPEs but is hard to derive if data amount is sparse. I explore methods to estimate κ_0 from ambient seismic noise. Seismic noise is, in contrast to earthquake recordings, continuously available. The rapidly emerging field of seismic interferometry gives the possibility to infer velocity and attenuation information from the cross-correlation or deconvolution of long noise recordings. The extraction of attenuation parameters from diffuse wavefields is, however, not straightforward especially not for frequencies above 1 Hz and at shallow depth. In this thesis, I show the results of two studies. In the first one, data of a small-scale array experiment in Greece are used to derive Love wave quality factors in the frequency range 1-4 Hz. In a second study, frequency dependent quality factors of S-waves (5-15 Hz) are estimated by deconvolving noise recorded in a borehole and at a co-located surface station in West Bohemia/Vogtland. These two studies can be seen as preliminary steps towards the estimation of κ_0 from seismic noise.

Acknowledgements

In my PhD I had the luck to benefit from the supervision of not only one person but from several people. I am very grateful to Prof. Dr. Frank Scherbaum who gave me the opportunity to start my PhD at the University of Potsdam and who guided me through my first years. His enthusiasm for seismology was always inspiring and a source for motivation. I am especially thankful to Dr. Matthias Ohrnberger. Throughout my whole PhD he offered guidance, shared ideas and helped to solve problems. His comments and feedback led me to rethink and improve my results over and over again. Due to his continuous encouragement I was able to overcome also difficult times. I owe thanks to apl. Prof. Dr. Frank Krüger who joined to guide me through the second part of my PhD. Even though he was not an official supervisor he never hesitated to offer support or to answer questions. I am also grateful to Prof. Dr. Volker John from WIAS who gave advise with regard to simulation techniques.

My PhD was mainly funded by a scholarship of the graduate research school GeoSim. During several summer schools and workshops my fellow GeoSim students became friends. I am happy that I could be part of this inspiring group of people.

My colleagues at the University of Potsdam made work often fun. I enjoyed the office times, joined lunches, bicycle tours, Christmas dinners and the Mühlenberg choir. All members of the working group were always willing to discuss problems and to give feedback and support if needed. I found friends in my office and Mühlenberg mates - Conny Hammer, Kristin Vogel, Antonia Runge, Katrin Hannemann, Galina Kulikova, Christian Molkenthin, Sanjay Singh Bora, Agostiny M. Lontsi, Carsten Riggelsen, Nikos Gianniotis, Marius Kriegerowski, Martin Zeckra, Stefanie Donner, Nico Kühn and many others.

I am also glad about the joy that other friends bring into my life like Janin and Sven Danneberg, Alexandra and Stefan Mauerberger, Sabrina and Stefan Leibelt, the FC Schragen football members and Nadia Said. You always remind me that there are also other thinks in life than work.

I am grateful to my family who always accepted me as I am and supported me in all stages of my life with love and encouragement. They provide the foundation on which I can feel safe and build my life on.

Finally, I dedicate this work to Ben, Alena and Lio. You make life beautiful. Thanks for your love!

Contents

List of Figures

List of Tables

1 | Introduction

Earth's population is increasing vastly and more and more people are living in earthquake prone regions and cities (Bilham, 2009). Most fatalities are caused by collapsing buildings and structures exposed to strong ground shaking. It is therefore necessary to assess the ground motions that can be expected due to an earthquake at a certain site in order to develop building codes for earthquake resistant design.

Todays most widely used tool for the quantification of the earthquake shaking hazard is probabilistic seismic hazard analysis (PSHA) as introduced by Cornell (1968). PSHA takes into account the whole randomness and uncertainty associated with ground-motion generation. The outcome of a PSHA is a hazard curve which gives the rate at which a certain level of ground motion is exceeded at a particular site. The two main ingredients of a PSHA are (Bommer et al., 2010; Cornell, 1968):

1. A model for the occurrence of future earthquakes that incorporates all potentially relevant earthquake sources together with their locations in space, their sizes and activity rates.

2. A model for the propagation of generated seismic waves between the place where the earthquake occurs and the site where the hazard should be calculated. The model can either be a stochastic model (Boore, 1983, 2003), an empirical model (e.g Campbell, 2003b) or a physical model which is based on numerical simulations (e.g. Beresnev & Atkinson, 1997; Graves & Pitarka, 2010; Hartzell et al., 1999).

In the second step, empirical derived ground-motion prediction equations (GMPEs) are generally used to estimate the ground motion for a specific region. GMPEs are mathematical expressions that relate ground-shaking parameters (e.g. peak ground acceleration [PGA] or response spectral acceleration [SA]) to seismological parameters such as earthquake size and mechanism, distance from source to site, and geological structure directly below the site of interest (Campbell, 2003b). SA describes the response of a single-degree of freedom oscillator with a certain

natural frequency to an input ground-motion. GMPEs are derived by regression analysis of recordings obtained from past earthquakes. A multitude of different models exists that were either generated for specific regions or by combining data from several areas. Douglas (2011, 2017) gives an overview of models that have been published so far but the number of GMPEs is still growing. GMPEs are probabilistic models that provide estimates of the median and the logarithmic standard deviation (sigma) of ground-motion parameters. The sigma of a GMPE is related to aleatory variability which describes the intrinsic randomness of ground-motion generation. Not knowing which GMPE to consider as the true one, results in epistemic uncertainty (lack of knowledge). Hazard analysts usually combine a whole set of GMPEs within a logic tree framework (e.g. Bommer et al., 2005) to capture epistemic uncertainty. Thereby, alternative models occupy the branches of such a tree while the branch weights express the degree-of-belief of an expert into the corresponding models. The proper selection of GMPEs for a logic tree analysis has been recognized to be a crucial step in any seismic hazard study (e.g. Bommer et al., 2005; Scherbaum et al., 2005). The set of GMPEs has to reflect the whole range of possible ground motions that can occur in a region (e.g. Budnitz et al., 1997; Hanks et al., 2009) which, in reality, is not possible as GMPEs are derived from limited datasets. An additional problem arises for regions that are seismically active but where recorded data are sparse, either due a paucity in seismicity or due to a lack of instrumentation. Data amount is then insufficient to generate endemic empirical GMPEs. In order to perform a seismic hazard study in such a region the usual practice is to consider GMPEs from other areas of the world and to adjust them to the target area (Al Atik et al., 2014; Atkinson, 2008; Campbell, 2003a).

1.1 GMPE adjustment techniques

Several techniques have been proposed to adjust an empirical GMPE that was derived for a data-rich region (host region) to an area with insufficient ground-motion recordings (target region). In order to perform the adjustment, regional differences in seismological parameters between the host and the target area have to be taken into account. Seismological parameters with regard to the earthquake source (e.g. stress drop [$\Delta\sigma$]), the propagation path (e.g. intrinsic attenuation modeled by the quality factor Q) and the site (e.g. shear-wave velocity of the upper 30 m [v_{s30}] and site attenuation [κ_0]) have to be considered. Global sensitivity studies conducted by Molkenthin et al. (2017) revealed that the computed hazard is most sensitive to changes in $\Delta\sigma$ and κ_0. In the context of site-specific hazard evaluations and GMPE adaptions from soft soil to rock sites, v_{30}-κ adjustments turned out to be one of the key factors having a large impact on the hazard results (Biro & Renault, 2012; Laurendeau et al., 2013).

The parameter κ describes the spectral fall-off that is observed in the Fourier amplitude acceleration spectrum $A(f)$ of earthquakes at high frequencies (Anderson & Hough, 1984). The decay can be described as:

$$A(f) = A_0 \cdot \exp\left(-\pi\kappa f\right) \qquad f > f_e \tag{1.1}$$

where f_e is the corner frequency above which the decay of the spectrum is approximately linear in log-linear space. κ is associated with the intrinsic attenuation of S-waves on the travel path from source to site (Anderson & Hough, 1984). The path-independent component of κ (κ_0) is obtained at epicentral distance zero (Hough & Anderson, 1988). κ_0 is considered to represent the site-specific S-wave damping in the very shallow crust (several 100 m to a few kilometers; Anderson & Hough, 1984; Campbell, 2009) directly below the site. κ is usually not a predictive parameter in GMPEs. Nevertheless, it can have a significant impact when computing the hazard of critical facilities like nuclear power plants and bridges that are sensitive to the shaking at high frequencies (Ktenidou & Abrahamson, 2016). It is therefore important to consider κ when adjusting a GMPE from one region to another.

The hybrid empirical adjustment method (HEM) of Campbell (2003a) utilizes the point source stochastic approach (Boore, 1983, 2003) to address regional differences in seismological parameters. In the stochastic approach, ground-motion is modeled as Gaussian band-limited white noise with a random phase and an amplitude that is modified by multiplication factors to account for source, path and site characteristics of a region. Adjustment factors in HEM can be calculated from the ratio of stochastically simulated response spectra of the host and the target area. Ground-motion parameters from the target area are adapted to the host region by multiplication with the adjustment factor.

Atkinson (2008) introduced the referenced empirical approach (REA) as an alternative to the method proposed by Campbell (2003a). In HEM, adjustment factors are based on stochastic simulations and seismological models. In contrast to HEM, REA derives adjustment factors for GMPEs directly from empirical data recorded in the target area. Modification factors are estimated by fitting the observed residuals between host GMPEs and the target-region ground-motion database. The applied adjustments modify the overall level of the attenuation curves to account, for example, for differences in stress drop or event type, and the shape of the curves to accommodate differences in regional attenuation.

Another adjustment approach introduced by Al Atik et al. (2014) uses inverse random vibration theory (IRVT, Rathje et al.(2005)) for the computation of κ scaling factors for GMPEs that do not include κ as input parameter. (Note that these authors work with the whole-path κ and not only the site-specific component κ_0.) In the forward mode (Cartwright & Longuet-Higgins, 1956), random vibration theory (RVT) is utilized to convert a theoretical Fourier amplitude spectrum in combination with a model for the duration of ground motion into response spectral ordinates. The inverse mode (IRVT) is less straightforward. It employs extreme value statistics, properties of single-degree-of-freedom oscillator transfer functions, and spectral ratio correction to estimate Fourier amplitude spectra from response spectra. The κ adjustment in the IRVT approach of Al Atik et al. (2014) is achieved by

first converting the host response spectrum to a Fourier spectrum using IRVT. The host κ is derived from the Fourier spectral slope at high frequencies as described in Anderson & Hough (1984). The adjustment to the target area is obtained by multiplying the spectrum with $e^{-\pi f(\kappa_{target}-\kappa_{host})}$. Finally, the adjusted Fourier spectrum is converted back to response spectral space using RVT.

The adjustment methods presented above come with several limitations as the relationship between Fourier spectra and response spectra is not linear (Bora et al., 2016). Bora et al. (2014) therefore proposed to develop GMPEs from an empirical model of the Fourier amplitude spectrum. In combination with a model for the duration of ground-motion and by using RVT, response spectral GMPEs can be calculated. This opposes the traditional technique where GMPEs are directly derived by regression analysis from response spectral ordinates. The main advantage of the method proposed by Bora et al. (2014) is that the adjustment of a Fourier amplitude spectrum to different seismological conditions is physically consistent and can easily be performed.

One of the main problems of most of the existing GMPE adjustment techniques is that they require a sufficient number of seismological data in the target area to constrain local site conditions (v_{s30} and κ_0). This is generally not given in low seismicity regions which is especially disadvantageous if not a regional but a site-specific hazard study should be performed. In practice, when site-specific data is not available κ_0 is estimated from an ensemble of sites in an area as has been done by Ktenidou et al. (2017b) in southern Arizona. A more common approach is to infer κ_0 from existing v_{s30}-κ_0 relations (e.g. Ktenidou et al., 2014; Van Houtte et al., 2011). Yet, the scatter in v_{s30}-κ_0 relations is large owing to differences in the measurement techniques of v_{s30} and κ_0, to the heterogeneity of regions from which the relations are inferred, to the frequency band that is considered in the deviation of κ_0 or simply to the fact that κ_0 also relates to parameters other than v_{s30} (Edwards et al., 2015; Ktenidou et al., 2014). Often the adjustment of GMPEs has to be performed from active regions with soft rock to less active regions with hard rock (Biro & Renault, 2012; Ktenidou et al., 2017b). The situation is then even more complex. v_{s30} and κ_0 data pairs are too sparse for $v_{s30} > 1000$ m/s to estimate a clear relationship between the two parameters. Thus, there is a special need to derive local site parameters for target areas with few ground-motion recordings and sites with hard rock conditions.

1.2 Purpose of the thesis

In this thesis, I explore methods that facilitate the selection of GMPE or their adjustment from one region to another. The special focus is thereby on areas where seismological information and strong-motion recordings are sparse.

I follow two different strategies towards this goal which I will introduce in the following two subsections.

1.2.1 Mixture model approach

The first method is targeted to aid the setup of a ground-motion logic tree for areas without an indigenous GMPE and without a sufficient number of recordings to develop one. The most popular approach (Fig. 1.1a) is to select several models from other parts of the world and to adapt them to the new region. The choice for appropriate GMPEs and the assignment of tree weights for each of the models can be either achieved by the judgment of experts and/or by data-driven procedures (Arroyo et al., 2014; Azarbakht et al., 2014; Kale & Akkar, 2013; Mak et al., 2017; Roselli et al., 2016; Scherbaum et al., 2004a, 2009). This approach was, for example, also adopted in some large ground-motion projects like the Global Earthquake Model (GEM) Global GMPEs project (Stewart et al., 2015) and the Seismic Hazard Harmonization in Europe (SHARE) project (Delavaud et al., 2012).

In some recent hazard studies (e.g. Bommer et al., 2013; Rodriguez-Marek et al., 2014), few high-quality GMPEs were chosen to serve as backbone models in the logic tree. A backbone model is assumed to give good estimates of the median ground-motions in an area from which the upper and lower limits of the ground-motion distribution can be obtained by scaling the model up or down. Backbone GMPEs are, for example, modified by applying v_s-κ_0 adjustments and by accounting for possible differences in stress drop. This procedure was adopted by Bommer et al. (2013) for the Thyspunt site in South Africa leading to a set of site-specific GMPEs. The concept of the backbone model in a ground-motion logic tree is illustrated in Fig. 1.1b.

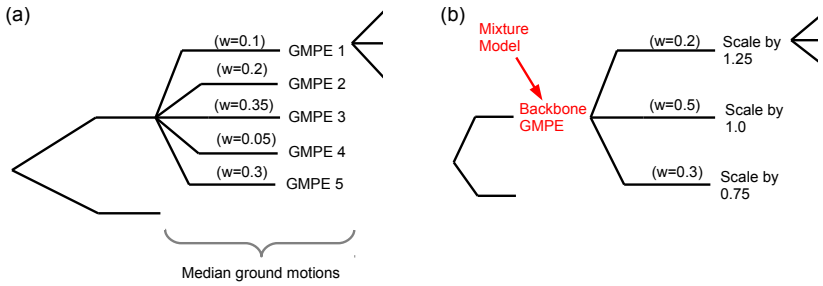

Figure 1.1: (a) Ground-motion logic tree using several GMPEs to predict the median ground motion in the area of interest. (b) Logic tree with a single high-quality backbone model that is scaled up and down to capture the body and range of median ground-motions. A mixture model can be used as an optimized backbone model.

In this thesis, I present a method in which an optimized backbone model is found for a region by combining several GMPEs as weighted components of a so called mixture model. As mentioned above, GMPEs are probabilistic models that provide estimates of the median and the logarithmic standard deviation of the log-normally distributed ground motion. The mixture is derived as a linear combination of these distributions where all weights are non-negative and sum up to one. The weights are inferred from observed data of the target area and the mixture model

is tuned to give an optimal performance for the region under study. The proposed method benefits from the large amount of observational data and expertise that has been used to generate the GMPEs that are combined in the mixture. Relevant information regarding, for example, the generation mechanisms of ground motion is thus transferred from the different host regions to the study area. At the same time, mixture weights for the GMPEs are derived in such a way that the new aggregated model captures the specific ground-motion characteristics of the target region.

The mixture model approach is especially helpful to generate backbone models in situations in which some ground-motion observations exist, but in which the number of data is still insufficient to generate a region-specific GMPE.

1.2.2 Estimation of κ_0 from seismic noise

The second approach is aimed at site-specific hazard evaluations for which it is especially important to know local site conditions to correctly adapt GMPEs from other regions. In this context, I put the focus on the estimation of the site-specific attenuation factor κ_0 as this parameter is one of the most important ones for GMPE adjustments having the largest effect on high frequency ground-motions (compare to Section 1.1).

In this thesis, I explore methods to derive κ_0 from ambient seismic noise. Seismic noise is quasi omnipresent and its use circumvents the problem of missing earthquake ground-motion recordings. By using noise it is in principle possible to determine κ_0 everywhere on Earth. After Bonnefoy-Claudet et al. (2006) noise below 0.5 Hz is mainly generated by natural sources as tides, interaction of the oceanic swell with the solid Earth and large-scale meteorological conditions. At frequencies around 1 Hz noise originates from local meteorological processes like wind effects. Above 1 Hz seismic noise generation is controlled by human activity as, for example, traffic or industrial production. Similar to earthquake generated waves, seismic noise wavefields travel through the ground and should thus contain information of the volume through which they propagate.

The use of seismic noise for the derivation of deterministic information about Earth's structure goes back to the pioneering works of Aki (1957) and Claerbout (1968). However, it was not until 2001 that ambient noise techniques became a widely used tool in seismology and other disciplines. Several authors have shown theoretically and experimentally that it is possible to estimate the time domain Green's function between a pair of receivers from the cross correlation of a diffuse wavefield (e.g. Lobkis & Weaver, 2001; Roux et al., 2005b; Sabra et al., 2005; Snieder, 2004; Wapenaar, 2004; Weaver & Lobkis, 2004). The Green's function is equivalent to the impulse response of the medium. It gives the signal at one receiver as if the other receiver was acting like a source. In recent years, many studies have shown that not only the elastic but also the anelastic Earth response can be extracted from seismic noise (e.g. Lawrence & Prieto, 2011; Lawrence et al., 2013; Lin et al., 2011; Liu et al., 2015; Prieto et al., 2011, 2009; Weaver, 2011, 2013; Weemstra et al., 2013; Zhang & Yang, 2013). The estimation of amplitude and,

hence, attenuation information from noise cross correlations is not straightforward. This is due to the fact that the amplitude of cross correlations is not only dependent on the anelastic structure of the Earth but also on the strength of the ambient field both in space and time and also on data pre-processing techniques (e.g Weaver, 2011).

Most of the studies listed above infer attenuation in the whole crust or upper mantle and within the ocean microseism band. The objective of this thesis is to estimate attenuation on a very local scale and at frequencies above 1 Hz. In particular, I am focusing on the extraction of the quality factor Q within the upper hundreds of meter below the surface as this is thought to be equivalent to the range where the site-dependent κ_0 originates from (Anderson & Hough, 1984; Campbell, 2009). Chapter 3 of this thesis presents estimates of Love-wave Q (Q_L) between several surface stations while Chapter 4 shows the derivation of S-waves quality factors (Q_s) between a borehole and a surface sensor. According to Hough & Anderson (1988), κ_0 can be regarded as a t* of vertically upward propagating S-waves through the local geological structure directly below the study site. Hence, κ_0 can be related to Q_s with a path integral over depth z or as a sum over N layers with thickness H (Ktenidou et al., 2015):

$$\kappa_0 = t^* = \int_{path} \frac{dz}{v_s(z)Q_s(z)} = \sum_{i=1}^{N} \frac{H_i}{v_{si}Q_{si}}. \tag{1.2}$$

Unfortunately, no earthquake-based κ_0 measurements are available for the studied site in order to compare our Q_s measurements via equation 1.2 with these values.

As will be shown in this thesis, the estimation of κ_0 values from ambient noise is very challenging. The presented applications may be seen as preliminary steps towards this goal that need further investigation in the future.

1.3 Outline of the thesis

The two research questions presented in the previous section guide my studies and are the basis for three publications. The following three chapters of this thesis are thus manuscripts from which the first and second one are already published while the third one is in review. The first publication introduces the mixture model concept and applies it to Northern Chile. The second and third manuscript are concerned with the estimation of attenuation parameters from seismic noise. Chapter 5 and a 6 of the thesis contain a discussion and conclusion of the proposed methods in the larger context of seismic hazard evaluation and ground-motion prediction.

The research documented in all three publications was carried out by the author of this thesis. Prof. Dr. Frank Scherbaum, Dr. Matthias Ohrnberger, apl. Prof. Dr. Frank Krüger and Dr. Nico Kühn provided ideas and suggestions and assisted in an advisory role. Sebastian Specht implemented the processing tool for ground-

motion recordings and helped to build the strong-motion database that is presented in the first publication.

Publication 1: *Mixtures of ground-motion prediction equations as backbone models for a logic tree: an application to the subduction zone in Northern Chile*

The paper introduces the aggregation of several GMPEs into a mixture model. The method is applied to Northern Chile, a region for which no indigenous GMPE existed by the time of the publication of this study. A new data set of strong-motion recordings obtained within the Integrated Plate Boundary Observatory Chile project serves as data basis. Eight subduction zone GMPEs are used to obtain mixture models either for interface- or intraslab-type events and for different oscillator periods. In addition to the mixture, a region-specific GMPE is derived by regression analysis. The performance of the GMPEs, the mixture and the regression model is evaluated using residual analysis and averaged sample log-likelihoods as proposed by Scherbaum et al. (2009). We show that the mixture model predicts the ground-motions in the target region better than any of the single GMPEs. It is also comparable in its performance to the regression model. We thus suggest that a mixture model might serve as a backbone model in a ground-motion logic tree analysis if data are too sparse to develop a region-specific GMPE.

Publication 2: *Extracting near-surface Q_L between 1-4 Hz from higher-order noise correlations in the Euroseistest area, Greece*

The second publication presents higher-order noise correlations (C^3 correlations) as introduced by Zhang & Yang (2013) for the derivation of near-surface attenuation parameters. To test the approach, we use data from a small-scale array experiment (station spacings $<$2 km) that was carried out in the Euroseistest area in Greece in 2011. The array was installed in the northern part of a sedimentary basin and stations were thus partly situated on weathered rock or on soft soil. As all receivers are surface stations only the surface wave response of the Green's function can be reconstructed. We estimate Love wave quality factors Q_L from the relative amplitude decay of Love waves in C^3 correlations with distance. Q_L is obtained between stations situated either purely on weathered rock or on soft soil and for frequencies between 1-4 Hz. As expected, attenuation is higher on soft soil than on weathered rock. The estimated quality factors are furthermore mainly in conformance with theoretical values derived from 1-D v_s and Q_s profiles from the Euroseistest area.

Publication 3: *Frequency dependent quality factors from the deconvolution of ambient noise recordings in a borehole in West Bohemia/Vogtland*

The last publication evaluates the retrieval of attenuation parameters from seismic noise that was recorded simultaneously in a borehole and at a surface station. We use a deconvolution procedure instead of cross correlations. Deconvolution interferometry effectively separates incoming and surface-reflected waves in the

wavefield. Attenuation can be inferred from the amplitude difference between these separated waves. The method is applied to a 87 m deep borehole in West Bohemia/Vogtland in Germany that is situated at a hard rock site. We employ the approach of Fukushima et al. (2016) to compute a transfer function of incoming and surface-reflected wave in the frequency domain. From the transfer function frequency dependent quality factors of S-waves (Q_s) can be obtained between 5-15 Hz. The retrieval of Q_s is compared for recordings of ambient noise and earthquakes. We can show that the deconvolution of ambient noise in a borehole provides a fast and valuable tool for the derivation of body wave quality factors in low-seismicity regions.

2 | Mixture model approach

Mixtures of ground-motion prediction equations as backbone models for a logic tree: an application to the subduction zone in Northern Chile

Annabel Haendel[1], Sebastian Specht[1], Nicolas M. Kuehn[2], Frank Scherbaum[1]

[1] *Institute of Earth- and Environmental Science, University of Potsdam, Karl-Liebknecht Str. 24-25, D-14476 Potsdam*
[2] *University of California, Berkeley, Berkeley, CA, USA*

Published in:
Bulletin of Earthquake Engineering
February 2015, Volume 13(2), pp. 483-501
DOI: 10.1007/s10518-014-9636-7
Link:
https://link.springer.com/article/10.1007/s10518-014-9636-7

Abstract In probabilistic seismic hazard analysis, different ground-motion prediction equations (GMPEs) are commonly combined within a logic tree framework. The selection of appropriate GMPEs, however, is a non-trivial task, especially for regions where strong motion data are sparse and where no indigenous GMPE exists because the set of models needs to capture the whole range of ground-motion uncertainty. In this study we investigate the aggregation of GMPEs into a mixture model with the aim to infer a backbone model that is able to represent the center of the ground-motion distribution in a logic tree analysis. This central model can be scaled up and down to obtain the full range of ground-motion uncertainty. The combination of models into a mixture is inferred from observed ground-motion data. We tested the new approach for Northern Chile, a region for which no indigenous GMPE exists. Mixture models were calculated for interface and intraslab type events individually. For each source type we aggregated eight subduction zone GMPEs using mainly new strong-motion data that were recorded within the Plate Boundary Observatory Chile project and that were processed within this study. We can show that the mixture performs better than any of its component GMPEs, and that it performs comparable to a regression model that was derived for the same dataset. The mixture model seems to represent the median ground motions in that region fairly well. It is thus able to serve as a backbone model for the logic tree.

Keywords Mixture model · Backbone model · Ground-motion prediction equation · Logic tree · Chile subduction zone

2.1 Introduction

The goal of probabilistic seismic hazard analysis (PSHA) is to quantify the probability that certain levels of ground shaking are exceeded at a particular site during a specified time period. The two main components of a PSHA are models for the space-time distribution, size and activity of relevant seismic sources, and secondly for the propagation of generated seismic waves to the point of interest at Earth's surface. The most common way to assess the latter one is the use of empirically derived ground-motion prediction equations (GMPEs), which give ground-shaking parameters [usually peak ground acceleration (PGA) or spectral acceleration (SA)] as a function of source-, path- and site-related predictor variables.

A major issue in PSHA is the correct quantification of uncertainty. Following the guidelines of the Senior Seismic Hazard Analysis Committee (SSHAC; Budnitz et al., 1997; Hanks et al., 2009; USNRC, 2012), GMPEs have to capture *the center, the body, and the range* (in SSHAC language) of the expected future ground-motion distribution in the area of interest. Because GMPEs are derived from limited datasets, often comprising data from more than one region, a single model is not assumed to capture the epistemic uncertainty of ground motion for a region of interest. This is especially the case for areas that lack an indigenous GMPE so that models from other regions have to be adopted. The hazard community therefore usually resorts to using sets of GMPEs, commonly combined within a logic tree framework as, for example, described in Bommer et al. (2005), where alternative models occupy the branches of the tree and where the branch weights

express the degree-of-belief of an expert into the corresponding models. The selection, ranking and weighting of appropriate GMPEs for a particular target area in a transparent and reproducible way becomes in this context the major challenge. The chosen set of models should on one hand be sufficiently small in order to keep the logic tree manageable, but needs to cover at the same time ground-motion uncertainty.

To narrow down the large number of published GMPEs (listed e.g. in Douglas, 2011), quality criteria as proposed by Cotton et al. (2006) and updated by Bommer et al. (2010) can be applied to pre-select appropriate models. A model should, for example, be excluded if it belongs to an irrelevant tectonic regime or if the documentation of the model is incomplete. In the past, the final set of GMPEs and the corresponding tree-weights have been achieved purely by the judgment of experts. Recently data-driven procedures (Kale & Akkar, 2013; Scherbaum et al., 2004a, 2009) have been proposed to test the appropriateness of a model for a particular target area in a more consistent and reproducible way, aiding experts in the assignment of logic tree weights.

In practice the handling of logic trees can be rather challenging. Inconsistencies in the assignment of branch weights (for example treating branch weights as normalized degree-of-belief values rather than probabilities) can lead to the over- or under-estimation of epistemic uncertainty and hence to a too large or too narrow spread of the corresponding hazard curves (Scherbaum & Kuehn, 2011; Scherbaum et al., 2010). In addition, GMPEs are in general not equally appropriate for all magnitudes, distances and oscillator frequencies so that weighting factors may need to be defined separately for different magnitude-distance-frequency bins. When many models need to be considered, the logic tree very fast becomes highly complex, making it impossible to directly deduce how uncertainty related to the ground-motion part of the logic tree is transferred to the hazard curves.

In some recent hazard studies (Bommer et al., 2013; Rodriguez-Marek et al., 2014), single high-quality GMPEs have been used as so called backbone models to represent the center of the ground-motion (median values) distribution in the area of interest. From these models (for the medians of ground motion), the full ground-motion uncertainty for use in the logic tree was subsequently obtained by scaling the GMPE predictions up or down. In general, however, a single non-indigenous ground-motion model will rarely be able to perfectly represent the center of the ground-motion distribution (at least not for all magnitude and distance ranges), since this model may only capture some (but not all) aspects relevant for ground-motion generation in the region of interest.

In the present study, we propose a new method in which an optimized backbone model for median ground motion is generated by combining different GMPEs as weighted components of a mixture model. The most likely combination of models is inferred from observed ground-motion data from the region of interest. In this process, information is partially transferred from other regions to the region where the observations have been produced. It is thus possible to infer a model that can deliver predictions for a region for which no dedicated GMPE has been

developed. None of the component GMPEs is considered as *true* or *false* within the mixture model approach. Instead, each model is assumed to reflect the generation mechanism of a part of the possible ground motions in the target area. A mixture combines the prediction strength of several GMPEs in contrast to a traditional logic tree where only one GMPE can always be considered as the *true* model within a single hazard realization. Since the mixture is calibrated against observations of the target area, it can be regarded as a good estimate of the median ground motion in that region. It is thus better able to serve as a backbone reference model for the logic tree. Up and down scaling of the new backbone model should provide better estimates of the true ground-motion uncertainty compared to the up and down scaling of a single non-indigenous GMPE. As the uncertainty related to ground-motion prediction can now directly be read from the logic-tree, simple translation can be applied to directly deduce how this uncertainty is transferred into the final hazard curves.

The new method is used to generate a backbone reference model for Northern Chile, a region for which no indigenous GMPE, which covers a sufficiently wide magnitude range to capture all hazard-relevant earthquakes in the region, exists. Seismic data from a dense seismic network deployed as part of the Integrated Plate Boundary Observatory Chile (IPOC; Schurr et al., 2009) and additional strong-motion recordings collected by Arango et al. (2011) provide a unique new dataset to study the aggregation of different subduction zone GMPEs. To handle the large amount of IPOC data, we developed a semi-automatic tool in Matlab (and C) that cuts and processes recordings from the continuous station data streams from events with magnitude 5 or larger, and derives earthquake and record related information, which are needed as GMPE input parameters. Backbone models are then inferred for interface and intraslab type events from a set of 8 GMPEs each and for different oscillator frequencies. We assess the performance of the mixture compared to that of its component models in terms of average residuals and the average sample log-likelihood as proposed by Scherbaum et al. (2009). Additionally, we evaluate whether the mixture is able to predict new observations by applying the model to a testing dataset that was drawn from the complete dataset before the calculation of the mixture weights. The performance of the mixture is finally compared to that of a regression model that we specifically estimated for the Chilean dataset in order to understand if a mixture is equally appropriate to represent the center of the ground-motion distribution in that area.

2.2 Mixture models

GMPEs are probabilistic models that provide estimates of the median and the logarithmic standard deviation (sigma) of the ground-motion parameter of interest (or ground-motion intensity measure, GMIM) Z. Z is conditioned on various parameters, such as magnitude and distance, denoted here by $\mathbf{x}$. A common choice for Z is, for example, spectral acceleration. Z is generally assumed to be log-normally

distributed but for simplicity we will use $Y = ln(Z)$ instead, which is a normal distribution of the form $\mathcal{N}(Y|\mu(\mathbf{x}), \sigma^2(\mathbf{x}))$ with mean $\mu(\mathbf{x})$ and variance $\sigma^2(\mathbf{x})$.

Rather than using a single attenuation relation it is possible to employ an ensemble of models (Riggelsen et al., 2011) by combining several GMPEs as components of a standard mixture model. Thereby, we assume that observations from an unknown, true ground-motion distribution (nature) can be modeled by a convex combination of several probability distributions (GMPEs), where each observation is supposed to be drawn from only one of the models/mixture components. The term convex refers to a linear combination of models where all coefficients or weights are non-negative and sum up to one. The standard mixture distribution (Frühwirth-Schnatter, 2006; Titterington et al., 1985) of J normal distributions with weights w_j is given by

$$p(Y|x) = \sum_{j=1}^{J} w_j \mathcal{N}(Y|\mu_j(\mathbf{x}), \sigma_j^2(\mathbf{x})) \tag{2.1}$$

where $\sum_{j=1}^{J} w_j = 1$ and $0 \leq w_j \leq 1$. The weights contain information on how important a single model is for the mixture given this particular set of models. The mixture itself is a new probability distribution, which in general will not have the form of a normal distribution (which is not an issue in the present context since we focus on median predictions) but can have any shape, for example, that of a multimodal distribution.

Mixture weights can be inferred from observed ground-motion data by employing, for example, the Expectation Maximization (EM) algorithm (Dempster et al., 1977). The EM algorithm is an iterative method that attempts to compute the maximum likelihood estimate of parameters (here weights w_j) in the presence of missing data. The missing information in the context of mixture model estimation is the membership of an observation i to one of the models j. EM always increases the likelihood within each step; in order to guarantee that a global and not a local optimum is reached, we ran the EM algorithm 50 times using different random starting weights. We subsequently computed the mean of all runs to obtain the final mixture weights. It should be noted that in our experiments the weights obtained in 50 runs were always almost identical. An example mixture distribution that is obtained with EM for a given set of predictor variables is shown in Fig. 2.1 (left). The right hand site of Fig. 2.1 shows median PGA predictions of the mixture and of the component GMPEs against distance for magnitude Mw 6. The standard deviation of the inferred mixture is not shown and should only be handled with care since it is not purely aleatory but inherits epistemic uncertainty due to the combination of different GMPEs into one model.

In general, the weights of the individual models depend on their standard deviation, since this determines the value of the likelihood function. We have tested the approach using both the indigenous value of the standard deviation for each model as well as a common value of sigma for each model, and the results were

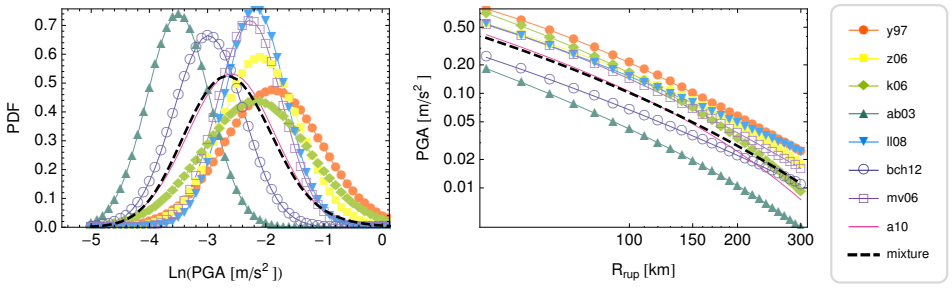

Figure 2.1: PGA predictions of the mixture (dashed) and the GMPEs (solid) for an Mw 6 event in 35 km depth and for NEHRP B. Left: Probability distributions in 120 km distance. Right: Median PGA versus distance. Compare to Fig. 2.3 for individual component weights and to Tab. 2.1 for GMPE abbreviations.

very similar. However, this can change for different applications, depending on the component GMPEs and the data set.

The mixture model is often able to predict ground motions that cannot be described by any of the component models alone (Riggelsen et al., 2011). However, the GMPEs are combined in a convex combination and a mixture model can only provide predictions within the range of its component distributions. If, for example, all models used for the derivation of the mixture over- or underpredict the observed ground motion in the area of interest the obtained mixture will also over- or underpredict the ground shaking. The same is true if the data used to derive the mixture weights are not representative samples of the ground-motion distribution in the target area. The mixture will then give incomplete predictions, covering only parts of the whole range of ground motion that is possible in that region. It is thus essential to use a diverse set of models and enough data in order to construct a reliable mixture model.

We remark that there are other possible methods to infer the most likely combination of models into a mixture besides the EM algorithm such as, for example, Bayesian inference. While the EM algorithm only returns point estimates (mean weights), the Bayesian approach provides the whole a posteriori distribution of weights.

2.3 Regression model

Parallel to the mixture, we also derive a new prediction equation that is obtained from the Chilean dataset by regression. The functional form is based on the GMPE of Zhao et al. (2006b), which we slightly modified. The coefficients of the model were estimated with the random effects algorithm of Abrahamson & Youngs (1992):

$$\begin{aligned}\ln Z =& aM_w + br_{rup} - (c + dM_w)\ln r_{rup} + e\begin{cases}h, & h \leq 125\,\text{km}\\ 125, & h > 125\,\text{km}\end{cases}\\ &+\begin{cases}q_i(M_w - 6.3)^2 + s_i, & \text{Interface}\\ q_s(M_w - 6.5)^2 + s_s + s_{sl}\cdot\ln r_{rup}, & \text{Intraslab}\end{cases}\\ &+\begin{cases}x, & \text{NEHRP} \neq \text{B}\\ 0, & \text{NEHRP} = \text{B}\end{cases} + \epsilon + \eta\end{aligned} \tag{2.2}$$

Here Z is PGA in cm/s^2, M_w is moment magnitude, r_{rup} is rupture distance and h is focal depth. a, b, c, d, e, q_i, q_s, s_i, s_s, s_{sl}, and x are the coefficients of the model that were derived through regression, where subscripts i and s equal interface and intraslab type events, respectively. q_i and q_s are the coefficients of the magnitude-squared term of the equation. The parameters s_i and s_s are used to describe ground-motion differences between interface and intraslab earthquakes, and s_{sl} is a magnitude-independent term to account for the more complex travel path of intraslab events. The coefficient x is a site class term. ϵ and η denote the within-event and between-event residuals, respectively, and are assumed to be normal distributed with zero mean and standard deviations ϕ and τ. The total standard deviation can be computed as $\sigma_{tot} = \sqrt{\phi^2 + \tau^2}$. The regression coefficients and the values for the standard deviation of within- and between-event residuals are listed in Online Resource 2.

We excluded the reverse-faulting parameter from the original functional form since it only applies to crustal events. As in Zhao et al. (2006b), we cap the depth term at 125 km, but we dropped the depth coefficient $h_c = 15$ km of their equation because only two events are shallower than 15 km in the Chilean database. The original term $\ln r_{rup} + ce^{dM_w}$ did not provide stable results using random effects regression; we therefore replaced the term by the equivalent form $(c + dM_w)\ln r_{rup}$. Instead of individual site terms for each class of the Zhao et al. (2006a) scheme, we used a very simple term which is zero for NEHRP B condition (majority of all the Chilean records) and takes the same value x for all other site conditions. Following equation 5 in Zhao et al. (2006b), we introduced a magnitude-correction term in our equation. However, this term was not derived from the between-event residuals of each source type but was obtained within the random effects regression itself. Due to trade-offs we had to exclude the linear and constant term in their correction function and only retained the magnitude-squared term. We did not change the coefficients $M_c = 6.3$ and 6.5 for interface and intraslab events, respectively, because the derived model provided sufficiently good results for the purpose of this study.

2.4 Processing of Chilean recordings

In this study we use accelerograms recorded within the Integrated Plate Boundary Observatory Chile (IPOC) project to infer the backbone GMPE for Northern Chile, a unique new dataset that forms the ideal basis to test the mixture model approach for a region for which no indigenous GMPE exists. Since 2006, the IPOC members (details e.g. in Schurr et al., 2009) operate a variety of geophysical stations to monitor the plate boundary segment in Northern Chile, which is assumed to be at the terminal stage of an interseismic seismic cycle (e.g. Comte & Pardo, 1991). In fact, on the 1st of April, during the review process of this manuscript, this seismic gap was partially ruptured by an Mw 8.1 offshore earthquake near the town of Pisagua. The key components of IPOC are 20 permanent, multi-parameter stations, which are installed throughout Northern Chile. Of special interest for this study are the continuously recording accelerometers that are operated at these sites. We also had access to eight triggered strong-motion instruments that are installed at different locations in the area. We refer to Online Resource 1 for details on the IPOC project and how the data were retrieved.

From 2006 until May 2012, more than 300 earthquakes with magnitude larger than five occurred along the subduction zone in Northern Chile and in Southern Peru. In order to be able to process the large number of recordings (more than 3,000) obtained from these events, we developed a semi-automatic tool in Matlab (currently translated into C) that queries earthquake information from seismological online platforms and subsequently cuts the corresponding recordings from the continuous station datastreams for processing. Earthquake metadata information was manually reviewed in order to deduce event (for example type of subduction zone earthquake) and record (for example rupture distance) related information. The next two subsections provide details on the processing tool and how metadata were obtained.

2.4.1 Earthquake and record related information

Earthquake metadata information was retrieved from different seismological data centers and publications. Only events with moment magnitudes (Mw) larger than 5.0 were considered as the majority of the GMPEs are calibrated against large magnitudes. Furthermore, we just took into account earthquakes with available focal mechanisms that are listed in the Harvard Centroid Moment Tensor (CMT) database (see Data and Resources section) in order to ensure that all relevant input information can be deduced for the models. Event locations coming from local agencies were preferred to those determined from teleseismic recordings. Date, time, epicentral location and focal depth of Chilean events were extracted automatically from the online catalog of the Geophysical Department of the University of Chile (GUC, see Online Resource 1). If no information could be obtained from the GUC database, the CMT solutions were used instead. For the Mw 7.7 Tocopilla earthquake and three of its aftershocks, improved event locations, focal mecha-

nisms and Mw measures were available from the study of Peyrat et al. (2010), which we substituted manually in our metadata list.

Subduction zone GMPEs generally differentiate between interface type events occurring at the coupled interface of the subducting and the overriding plate, and intraslab type events that take place within the downgoing slab. The classification into different source types was guided by the work of Delouis et al. (1996) for the Chilean earthquakes of this study. Delouis et al. (1996) identified different zones of geometry and stress regime in the Andean subduction zone between 22°S and 25°S that can be associated with interface and intraslab activity, respectively. We also used the definitions of Delouis et al. (1996) to discriminate the actual rupture fault plane from the auxiliary plane. The nodal plane of the focal solution dipping to the east was generally considered to be the true fault plane, but some cases remained ambiguous. Style-of-faulting was then assigned according to the rake angle of the rupture plane following the definitions of Wells & Coppersmith (1994). The final classification into interface and intraslab type events was performed manually and was based on the depth of an event, the distance to the trench, style-of-faulting and dip angle of the fault plane. We plotted vertical cross sections perpendicular to the trench using reported EHB events (Centennial Earthquake Catalog, see Online Resource 1) between 1960 and 2010 to visualize the downgoing Nazca plate, and to display the relative position of earthquakes in comparison to the slab. In doing so, crustal and other spurious events could be detected and removed from the final database. Earthquakes that could not clearly be classified as either interface or intraslab events were also excluded from the dataset.

Rupture and hypocentral distance were calculated automatically for each station-event pair. The rupture dimension of an earthquake was estimated following the work of Strasser et al. (2010) who developed empirical relations between rupture length and width and moment magnitude for subduction zone events. We preferred the relations of Strasser et al. (2010) to those of Blaser et al. (2010) because they provide relations for interface and intraslab type events rather than for different styles-of-faulting, an information which was missing for events for which the true fault plane could not be identified. The rupture plane of a Chilean event was first modeled in spherical coordinates using earthquake depth, rupture length, rupture width, strike and dip of the preferred focal plane, with the hypocenter lying in the center of the fault plane. We decided to determine the rupture plane first in spherical coordinates because it assures that the fault plane is slightly curved and that it does not come to lie outside of the seismogenic zone. Rupture plane and station coordinates were then converted to Cartesian coordinates and an iterative grid search was applied to find the shortest distance between the fault plane and the station. Two iterations were performed whereby the second one used a finer grid in the solution grid cell of the first iteration.

We used the relations of Scherbaum et al. (2004b) to determine rupture distance directly from hypocentral distance for recordings from those events for which the decision for the correct fault plane remained ambiguous. However, we modified the original distance conversion relations, which were derived for shallow crustal

activity, to make them applicable to subduction type events. We used, for example, the rupture dimension-Mw relations of Strasser et al. (2010) instead of the ones from Wells & Coppersmith (1994). Furthermore, the upper and lower depths limits of the simulated fault planes were set to 3 and 80 km for interface, and to 3 and 300 km for intraslab events. New dip angle distributions (truncated normal distributions) were chosen according to the dip angles observed for the Chilean events. For interface earthquakes, the angle was allowed to vary between 10° and 40° with a mean value of 25° and a standard deviation of 20°. For intraslab events, the mean and standard deviation of the dip distribution were set to 65° and 20°, respectively, and the distribution was confined to the range between 35° and 90°. 20.000 fault planes were constructed to determine the new distance conversion relations.

All continuously recording stations of the IPOC project (approximately 95 % of all records) are built on rock sites. Geological or geotechnical information for most of the eight triggered IPOC accelerographs are however missing. We only received site estimates for two of the sensors given in personal communication by Ruben Boroschek. Records from all other triggered stations were excluded from the database.

2.4.2 Record processing

The processing procedure implemented in our tool follows mainly the guidelines of Converse & Brady (1992) and the recommendations of the COSMOS strong-motion record processing workshop, which took place in 2004 (results described and summarized e.g. in Boore & Bommer, 2005; Shakal et al., 2005). First, we used the event times collected within the previous step to cut recordings from the continuous station datastreams. Thereby, we only extracted signals from stations lying within a circle with 500 km radius around an earthquake for processing. A record was cut 7,200 s before and 3,600 s after the reported event time and then stored. The relevant seismic signal was assumed to end 600 s after the reported earthquake origin time.

No instrument correction was applied to the recordings of this study, as the transfer functions of modern digital strong-motion instruments is usually flat in the frequency range that is of interest for engineers (up to 100 Hz or higher). The records were simply divided by the sensors absolute sensitivity at direct current to achieve acceleration units. An initial zeroth-order baseline correction was performed on the raw accelerograms. The mean of the pre-event part of a record was calculated and subtracted from the whole trace. We integrated some of our zeroth-order corrected records to velocity and displacement to check for long period drifts that would indicate changes of the reference baseline itself, which can occur due to strong shaking. The observed offsets were usually small or could be removed by high-pass filtering so that no further baseline adjustments schemes (as e.g. the one by Wang et al., 2011) were applied to the records of this study.

The accelerograms were then filtered using two passes of a 4th order Butterworth band- or high-pass filter, with one pass in forward and one pass in reverse

direction. The signals were padded with zeros prior to filtering in order to accommodate for transients that are introduced at both ends of the trace during the filtering process. The number of zeros T_{zpad} (in seconds) that needs to be added to a time series was calculated using the formula of Converse & Brady (1992), $T_{zpad} = 1.5n/f_c$, where n is the order and f_c is the corner frequency of the Butterworth filter. Half of the computed zeros was added at the beginning of the record, half of it at the end. In order to avoid sharp offsets between the padded section and the original trace, we replaced the parts of the signal lying before its first and after its last zero crossing with zeros.

We determined optimal filter corner frequencies for each record by comparing the Fourier amplitude spectrum (FAS) of the whole trace (containing signal and noise) with the FAS of the pre-event noise part. Discrete FAS's were computed for log-spaced frequencies and smoothed in the spectral domain by convolving them with a log-Gaussian function ($\sigma = 0.2$). Only signal parts with a signal-plus-noise to noise ratio (SNNR) of at least 3 (corresponding to a signal-to-noise ratio of 2) were accepted, and corner frequencies for a high- or band-pass filter were set accordingly. It was automatically checked how often the SNNR level of 3 was crossed. If no crossings were observed and the ratio was always larger than 3, a high-pass filter with fixed corner frequency of $f_c = 0.05$ Hz was applied. A high- or a band-pass filter was used if one or two crossings occurred above 0.05 Hz. The filter corners were then set to the corresponding frequencies of the SNNR level crossings. We visually inspected records where more than two crossings were detected. In general, those signals were of very low quality and the SNNR was barely larger than 3 over the whole frequency range. Such records were therefore rejected from the database. The final filtering results were randomly checked through visual inspection of the velocity and displacement curves, which are obtained by integrating the filtered accelerograms.

We decided to use the same filter parameters for all components of a triaxial recording to guarantee consistency in the frequency content between the channels. The choice of the filter corners was based on the horizontals since most GMPEs model horizontal ground motions. The SNNR was calculated for the geometrical mean of both horizontal FAS's to ensure that the final corner frequencies would not rely on the results of only one of the horizontal components.

Finally, PGA and peak ground velocity (PGV) were determined. Spectral accelerations for 5 % damping of critical were estimated using the method of Weber (2002). Following the suggestions of Akkar & Bommer (2006), Akkar et al. (2011), Boore (2005) and Douglas & Boore (2011), we did not use response ordinates below the high-pass corner frequency because the shape of the response spectrum strongly depends on the chosen filter parameters. In consistency with Abrahamson & Silva (1997), we set the lower bandwidth of the response spectrum to 1.25 times the high-pass corner frequency. The high frequency part of the response spectrum, on the other hand, is less affected by the choice of the low-pass filter corner. Douglas & Boore (2011) found that high frequency response ordinates are often dominated by signal frequencies much lower than the corner frequency of

the low-pass filter. We followed their conclusion that response spectral values can be used up to the Nyquist frequency, irrespective of the chosen filter parameters.

2.4.3 Final Chilean database

In order to extend the IPOC database to higher magnitudes, we included data from the work of Arango et al. (2011). The authors of this study collected and processed 98 accelerograms from 15 major subduction zone earthquakes that occurred in the Peru–Chile region between 1966 and 2007. Due to, for example, distance constrains we only used 47 records from 11 earthquakes recorded on 40 stations. We refer to their paper for details on the record processing, the stations or the collection of metadata.

Based on the applicability ranges of the used GMPEs, we only considered signals within a rupture distance range of 40–300 km and from earthquakes with magnitudes between Mw 5 and 8.1. All records where site conditions, source type information or other crucial predictor variables were missing were excluded from our database. The combined dataset used in this study consists of 1,094 triaxial recordings (interface: 374, intraslab: 720) from 138 events (interface: 48, intraslab: 90). The magnitude-distance distribution of the recordings is shown in Fig. 2.2.

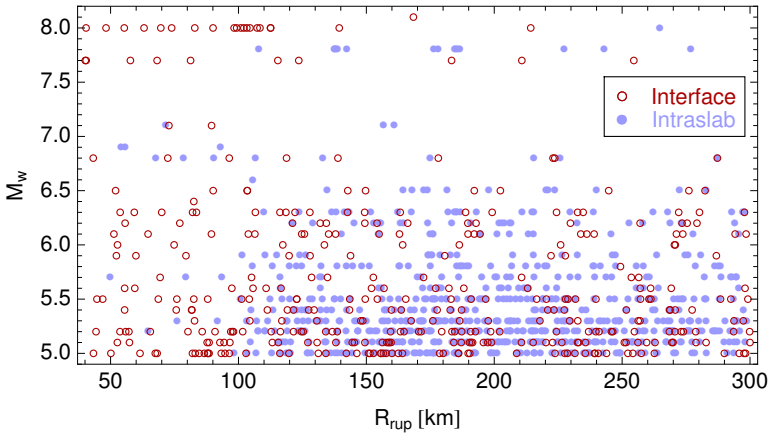

Figure 2.2: Magnitude-distance distribution of the Chilean recordings.

We extracted response spectral values for resonator frequencies between 0.3 and 30 Hz and for PGA, and computed the geometrical mean of two horizontal components for each record and for each frequency. The final list of earthquakes, a list of records and the computed spectral values for the different oscillator frequencies and for PGA can be found in the Online Resources 3–5.

Table 2.1: Characteristics of subduction zone GMPEs used in this study.

Model	Short	Main Region	Fosc (Hz)[a]	Mw[b]	Dist. type[c]	Dist. (km)[e]	ST[f]	H-Comp. type[g]	Site Class[h]
Arroyo et al. (2010)	a10	Mexico	0.20-25	5.0–8.0	rrup	20–350	IF	GMH	NEHRP B
Atkinson & Boore (2003)	ab03	Worldwide	0.33-25	5.0–8.5	rrup	20–300	IF/IS	RNDH	NEHRP (4)
BC Hydro (2012)	bch12	Worldwide	0.10-50	5.0–8.4	rrup, rhyp[d]	20–300	IF/IS	GMH	Vs30 (cont.)
García et al. (2005)	g05	Central Mexico	0.20-25	5.2–7.4	rrup	5–400	IS	GMH	NERHP B
Kanno et al. (2006)	k06	Japan	0.20-20	5.5–8.0	rrup	40–300	S/D	RVCH	Vs30 (cont.)
Lin & Lee (2008)	ll08	Taiwan	0.20-100	4.3–8.1	rhyp	40–300	IF/IS	GMH	NEHRP (2)
McVerry et al. (2006)	mv06	New Zealand	0.33-12.5	5.2–6.8	rrup	45–300	IF/IS	GMH	NZ (3)
Youngs et al. (1997)[i]	y87	Worldwide	0.33-12.5	5.0–8.2	rrup	20–500	IF/IS	RNDH	GMX (2)
Zhao et al. (2006b)	z06	Japan	0.20-20	5.0-8.3	rrup	10–300	IF/IS	GMH	HARD (5)

[a] Oscillator frequency range considered as good.
[b] Magnitude range considered as good. Mw is used in all GMPEs listed.
[c] Distance type: hypocentral distance (rhyp) or rupture distance (rrup).
[d] Rrup for interface events, rhyp for intraslab earthquakes.
[e] Distance range considered as good.
[f] Source type: interface (IF) or intraslab (IS). For Kanno et al. (2006): shallow (S) or deep (D) events (30 km) depth limit.
[g] Horizontal component definition: random component (RNDH), geometrical mean (GMH), vectorially resolved decomposition - square root of sum of squares of the two components in the time domain (RVCH).
[h] Site classification scheme. vs30: continuous (cont.) site term based on shear wave velocity of the top 30 m; NEHRP: Classes following the National Earthquake Hazards Reduction Program (NEHRP) (1997) scheme; GMX: Geomatrix C3 classification (Abrahamson & Silva, 1997); HARD: site classes according to Zhao et al. (2006a); NZ: New Zealand site definition as used by McVerry et al. (2006). The number of site classes is given in brackets. The equations of García et al. (2005) and Arroyo et al. (2010) are only valid for NEHRP B (rock).
[i] Youngs et al. (1997) considered their model as obsolete. However, the model is still in wide use (e.g. GEM; Douglas et al., 2012) so that we decided to keep it. Y97 receives mixture weights that are close to zero, showing that the mixture approach is stable even in the presence of a component GMPE that is not considered as good.

2.5 Subduction zone GMPEs

We use the same set of subduction zone models that has been pre-selected within the framework of the Global Earthquake Model (GEM) GMPEs Project (Douglas et al., 2012). Tab. 2.1 lists the models and their characteristics. The equations of García et al. (2005) and Arroyo et al. (2010) are only valid for intraslab and interface type events, respectively. All other candidate GMPEs provide either different sets of coefficients for interface and intraslab earthquakes (Atkinson & Boore, 2003), or account for the difference in source type by a switchable additive term in their equation (BC Hydro, 2012; Lin & Lee, 2008; McVerry et al., 2006; Youngs et al., 1997; Zhao et al., 2006b). The model of Kanno et al. (2006), however, discriminates between events shallower than 30 km (mainly interface and crustal earthquakes) and events deeper than 30 km (mainly intraslab), rather than between interface and intraslab type events.

The selected GMPEs combine the horizontal components in different ways. We adopt the geometrical mean of the two horizontals as the reference definition and adjust all other definitions. The correlations derived by (Beyer & Bommer, 2006) were used to convert random vector components to the geometrical mean. We note that these equations were derived for crustal motions and that it is unclear whether they are applicable to subduction zone regimes. The correction of the Kanno et al. (2006) model, which uses the time domain decomposition of the two horizontal components, was based on the work of Bragato & Slejko (2005), who reported that median ground motions computed with this definition are on average 27% higher than those using the geometrical mean.

Nearly all candidate models employ a different site classification scheme in their equation. We decided to transfer all schemes to the one of the National Earthquake Hazard Reduction Program (NEHRP). We did this by evaluating the models for rock conditions and then adding the NEHRP site term of the Atkinson & Boore (2003) model (which is zero for NEHRP B conditions) without the non-linearity term. The model of Lin & Lee (2008) already uses NEHRP classification, so no changes needed to be applied.

We are aware that Contreras & Boroschek (2012) established new equations for the prediction of spectral acceleration generated by large Chilean interface earthquakes. Their equations are, however, based on a very limited number of strong motion data (117 accelerograms from 13 interface type earthquakes), and are only valid for earthquakes with magnitudes larger than Mw 6.5. As can be seen in Fig. 2.2, the majority of the observations of this study is below Mw 6.5. We therefore think it is inappropriate to use the new model in the formation of the mixture.

2.6 Mixtures for Northern Chile

In this section we present the mixtures that are obtained for Northern Chile when aggregating the set of models introduced in the last paragraph. We computed mix-

ture weights individually for interface and intraslab type events and for different oscillator frequencies. The performance of each mixture is evaluated using

- average residuals, which show the difference between the mean predictions $\mu_j(\mathbf{x})$ of a model j and the observed data Y_{obs} for N observations: $AvRes_j := \sum_{i=1}^{N}(Y_{obs,i} - \mu_j(x_i))$
- the negative average sample log-likelihood (LLH) proposed by Scherbaum et al. (2009): $LLH_j := -\frac{1}{N}\sum_{i=1}^{N} log_2(p_j(x_i))$, where $p_j(x_i)$ is the likelihood that observation x_i was produced by the probability distribution of model j. Small LLH values indicate that a model is successful in describing the data, whereas the LLH value increases for models that mostly fail to predict the observations.

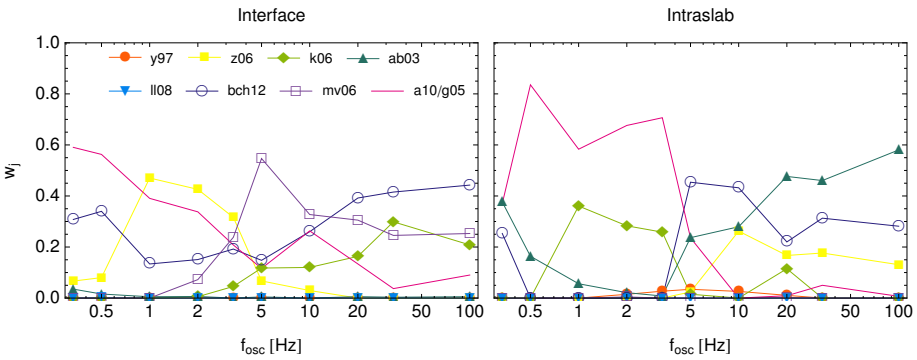

Figure 2.3: Mixture weights that are obtained for different oscillator frequencies when using the complete dataset of this study. The oscillator frequency of 100 Hz corresponds to PGA. We refer to Tab. 2.1 for GMPE abbreviations. Model a10 and g05 are only considered for interface and intraslab type events, respectively.

Fig. 2.3 shows the mixture weights that are obtained for different oscillator frequencies when applying EM on the complete interface and intraslab datasets of this study. As can be seen, the mixture is often only made up from a set of 3 or 4 models receiving high weights, whereas the other models are given weights close to zero. However, no single model contributes equally to the mixture over all spectral frequencies. Instead, some models receive, for example, high weights for low oscillator frequencies but not for high frequencies where instead another model is more important. Some GMPEs (y97, ll08 and ab03 for interface and y97, ll08 and mv06 for intraslab events) do not contribute at all to any of the mixtures and have zero weights (or very close to zero) over all oscillator frequencies.

In order to assess the predictive performance of the mixture for new data, we randomly selected 100 recordings from interface type and 100 recordings from intraslab type events from the database. These recordings were removed from the complete database and stored as testing dataset. Mixture weights and the coefficients of the regression model were then derived for the remaining number of

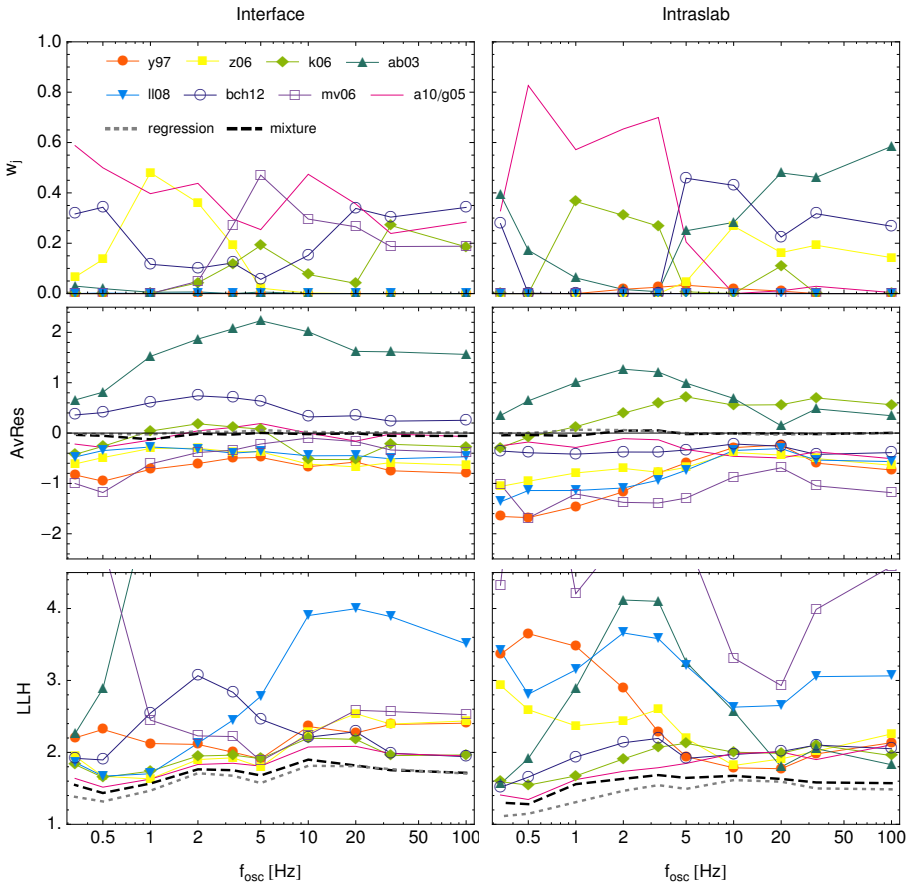

Figure 2.4: Mixture weights (top) that are obtained when inferring the mixture from the original dataset minus a test datasets (each 100 interface and intraslab records), and the regression model estimated for the same dataset. The performance of the mixture, the regression model and the individual GMPEs for the testing dataset is assessed in terms of average residuals (middle) and LLH values (bottom).

recordings (Fig. 2.4 left). We compared the model predictions of the resulting mixture (black, dashed line), the regression model (gray, dotted line) and the individual GMPEs with the extracted test observations in terms of residuals (Fig. 2.4 middle) and LLH values (Fig. 2.4 right). Both the mixture and the regression model always show average residuals very close to zero and LLH values that are smaller than the ones from the component GMPEs (which indicates a better fit). The inferred mixture and the derived regression model are comparable in terms of average residuals and LLH values for interface events. The regression model performs only slightly better for intraslab events when looking at LLH values, but still both the mixture and regression model have average residuals very close to zero. The results of the mixture and regression model, when estimated and applied for the complete dataset, look very similar to the test dataset case so that they are not shown here.

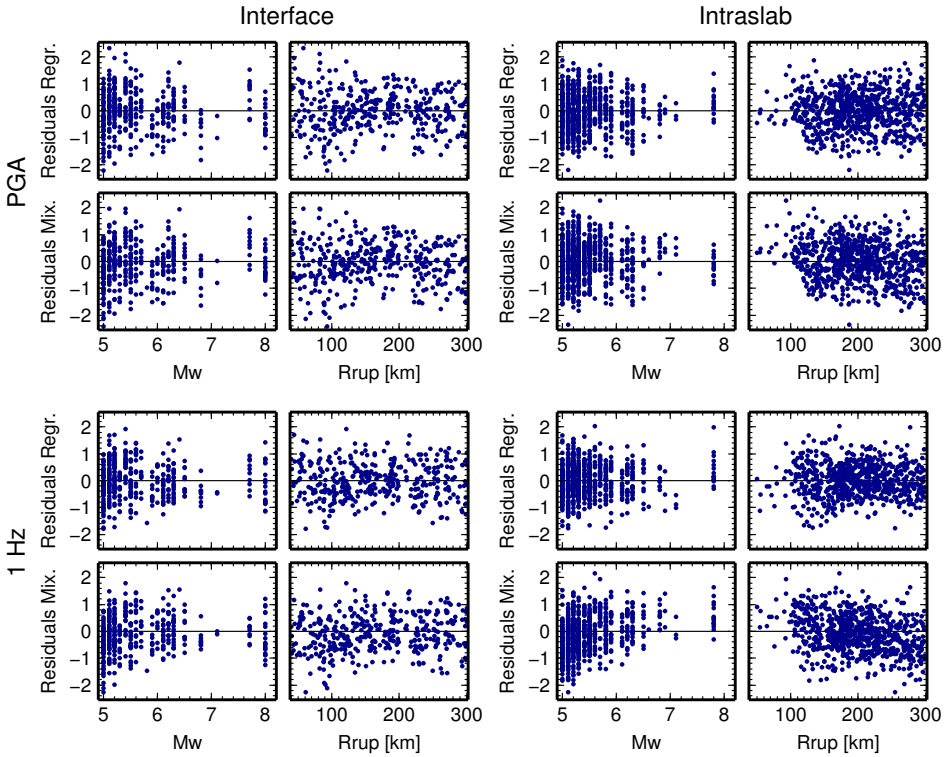

Figure 2.5: Residuals of the regression model and of the derived mixture as a function of magnitude and distance for both interface (left) and intraslab (right) type events estimated for the complete Chilean dataset. PGA: two upper panels, SA 1 Hz: two bottom panels.

Fig. 2.5 shows the residuals of the mixture and the regression model for PGA and 1 Hz as a function of magnitude and distance for both interface and intraslab type events. Neither the mixture nor the regression model shows a trend with magnitude or distance for interface events. However, for intraslab events especially at 1 Hz the mixture seems to underpredict ground motions at high magnitudes and

small distances, and to overpredict ground motions at small magnitudes and large distances.

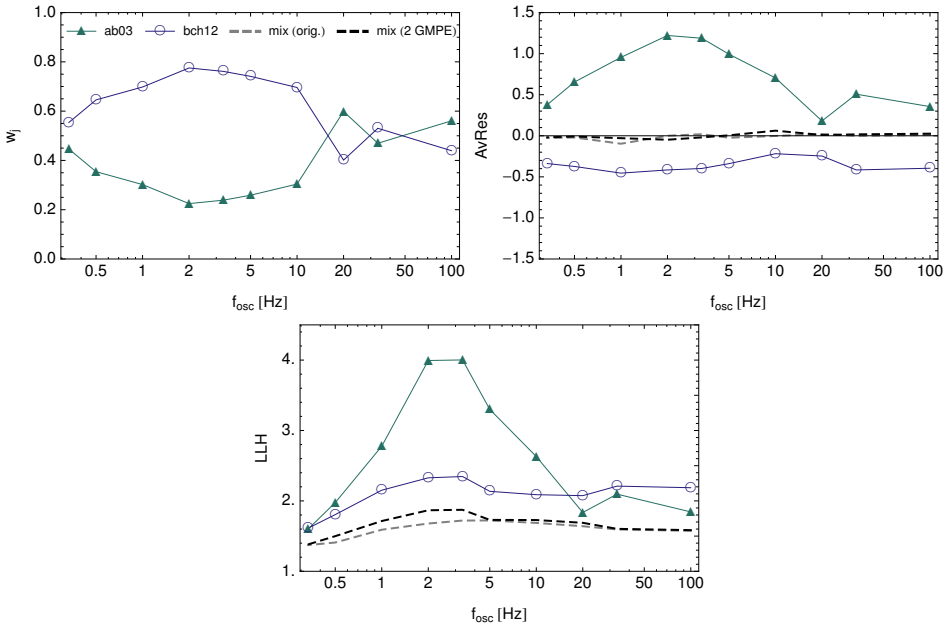

Figure 2.6: Mixture weights for intraslab type events (top left) that are obtained when aggregating only two GMPEs. The resulting residuals (top right) and LLH values (bottom) of the mixture and the two component GMPEs are shown in comparison to the performance of the mixture that was inferred for all eight GMPEs of this study.

We finally tested the performance of the mixture when aggregating only two GMPEs that are very different in terms of their average residual. We used the model ab03, which underpredicts the intraslab ground motions for all oscillator frequencies, and the model bch12, which generally overpredicts the ground shaking, and inferred a mixture made from both models. The weights, residuals and LLH values of the mixture and both component GMPEs are shown in Fig. 2.6. The performance of the two-component mixture is nearly as good as the one from the mixture obtained when aggregating all eight GMPEs (Fig. 2.6 grey dashed line).

2.7 Discussion

The results of the previous section show that a mixture performs better than any of the single GMPEs in terms of average residuals and LLH values. Only model a10 for interface and model g05 for intraslab event at low oscillator frequencies are close in their performance to the mixture. The mixture is even comparable to a model that was obtained by regression from the same dataset. The full potential of the mixture emerges for intraslab events at high oscillator frequencies, where

all GMPEs either under- or overpredict the Chilean observations and show quite large LLH values. The mixture performs then considerably better compared to all individual models.

As Fig. 2.6 shows, even two models might suffice to derive a good mixture model. However, both models need to cover different ranges of possible ground motions in order to derive a mixture that is able to represent the median ground motion in the area of interest. It has to be assured over the complete frequency range that not both models have either positive or negative average residuals. In practical applications and in situations where the hazard calculation should be automated, it is easier to employ a set of diverse models from which the essential ones will be picked within the mixture weight calculation step.

The lack of fit of the mixture for some magnitudes and distances (shown in Fig. 2.5) is not surprising, since we infer a single mixture for the whole dataset. It would be possible to allow the mixture weights to vary with magnitude and distance. However, as we propose to use the mixture approach in situations were strong-motion data are sparse, it would be hard to retain a sufficient number of recordings after splitting the dataset into different magnitude-distance bins, and this might lead to over-fitting. Furthermore, the idea is to keep the logic tree simple, using only a single model, which can easily be adjusted, for example, with respect to stress drop by applying scaling factors in the logic tree.

One could also derive a regression model if the available dataset of the target area is sufficiently large to do so. In general, however, more data are needed to derive a regression model compared to the combination of existing models into a mixture. Additionally, the hazard analyst needs to decide upon an appropriate functional form first, before running the regression, which is not always a trivial task. The mixture, on the other hand, is easier to handle and can give fast and reliable predictions. It is nevertheless dependent on the quality of the component GMPEs, but as long as the applied models span a wide range of possible ground-motion predictions the mixture will give adequate results.

Even though mixture weights are calibrated against observations of the target area and sum up to one, they should not be equated with logic tree weights. Logic tree weights give a notion on how good a single model is to describe observed and future ground motions in the area of interest. The combination of models into the mixture, on the contrary, is tuned to give an optimal overall performance of the mixture, sometimes irrespective of single model performances. In this context, it is possible that even *bad* performing models (in terms of average residuals and LLH values) are given high mixture weights because they provide an indispensable contribution to the mixture. When looking, for example, at Fig. 2.4 for intraslab events, model ab03 receives high weights for oscillator frequencies of 5 and 10 Hz. However, the model shows large positive average residuals and very large LLH values for the same frequencies and would most probably not receive high logic tree weights. But when combining different GMPEs into a mixture, the model is important to *drag* the aggregated model into the direction of positive residuals as nearly all other component GMPEs show negative residuals. The resulting mixture

then has residuals very close to zero. On the contrary, the model k06 performs very well for interface events at 1 Hz. Nevertheless, the model receives zero weight and does not contribute to the mixture. In general, however, performance in terms of average residuals and LLH values is often reflected in the mixture weights.

The median of the computed mixture can be used as backbone model for the logic tree of median ground motions. The standard deviation of the mixture, on the other hand, should not be regarded as a reliable (median) measure of the intrinsic ground-motion variability, because it inherits not only the randomness of ground motion but also epistemic uncertainty due to the combination of different GMPEs into one model. A median sigma model based, for example, on the single-station sigma concept (Atkinson, 2006) could be used instead, from which a full sigma logic-tree (to capture the epistemic uncertainty of sigma) could then be build as, for example, described in Rodriguez-Marek et al. (2014).

2.8 Conclusion and outlook

We have presented a new method that allows to aggregate existing GMPEs into a mixture model using observed data. The new approach was tested on a Chilean dataset, using eight subduction zone GMPEs for the prediction of interface and eight subduction zone GMPEs for the prediction of intraslab event ground motions. The derived model performs better than any of its component GMPEs, and performs comparable to a regression model that was also estimated on the basis of the Chilean dataset. We therefore conclude that the mixture model is a good estimate of the center of the ground-motion distribution in that area, and that the model could consequently be used as backbone model for the logic tree. The body and range of the ground-motion distribution can be obtained by scaling the mixture model up or down.

We believe that the mixture model approach might be helpful to generate backbone models in situations were some ground-motion observations exist, but in which the number of data is still insufficient to generate a region-specific GMPE. In areas where observations are extremely sparse or are known to be not representative, the generation of the mixture could be done within a Bayesian framework that allows the joining of experts' prior knowledge with data. Following this approach, mixture weights are distributed according to an a priori distribution (defined by experts believes), which is subsequently updated with observed data, leading to the a posteriori distribution of weights.

The mixture model is very flexible in that it allows the updating of existing mixture weights (estimated either purely data-driven or by the combination of expert knowledge and data) in a Bayesian framework as new earthquakes happen and new observations become available. The adjustment can be done fully automatic, making it possible to not only keep the ground- motion part of the logic tree *up to date*, but also the resulting hazard maps.

Data and Resources Accelerograms of this study were recorded within the Integrated Plate Boundary Observatory Chile (IPOC) project (Schurr et al., 2009, http://www.ipoc-network.org). IPOC data can be obtained from the Geofon Data Centre (http://geofon.gfz-potsdam.de/), but access to recordings of some of the instruments is restricted. We received the continuous data files from the multi-parameter sites upon request from the GFZ German Research Centre (contact person Bernd Schurr), and the accelerograms from the triggered sensors from the University of Chile (contact person Ruben Boroschek).

Recordings from the work of Arango et al. (2011) were received after contacting the authors of the study, but the accelerograms can now also be obtained from the COSMOS Strong-Motion Virtual Data Center (http://www.cosmos-eq.org/VDC/index.html).

The Global Centroid Moment Tensor Project (CMT) database was searched using www.globalcmt.org/CMTsearch.html (accessed 3 August 2012). Metadata information from the online catalog of the Geophysical Department of the University of Chile (GUC) were obtained from http://ssn.dgf.uchile.cl/ (accessed 10 August 2012), the website now moved to http://www.sismologia.cl/. Events from the Centennial Earthquake Catalog were downloaded from http://www.isc.ac.uk/ehbbulletin/ (accessed 21 February 2012).

Online resources of this manuscript can be dowloaded from https://link.springer.com/article/10.1007%2Fs10518-014-9636-7 and contain:

- Online Resource 2: List of regression coefficients
- Online Resource 3: List of earthquakes used in the study
- Online Resource 4: Final list of recordings used in this study
- Online Resource 5: Response spectral values computed for each record

Online resource 2 is also shown in Appendix A of this thesis.

Acknowledgements The two first authors of this study were partially supported by the Integrated Plate Boundary Observatory Chile (IPOC). Annabel Haendel would like to thank the Helmholtz graduate research school GeoSim (http://www.geo-x.net/geosim) for providing a scholarship. We are grateful to the IPOC members and Maria C. Arango for providing the accelerograms of this study, and we are especially grateful to Ruben Boroschek and Bernd Schurr for help and background information regarding the dataset. The authors would like to thank Helmut Staedtke for the implementation of a conversion tool, Elise Delavaud for guidance regarding the horizontal component conversion and Antonia Runge for many discussions about the mixture model approach. We are grateful to the two anonymous reviewers for valuable suggestions and comments, which helped to bring the manuscript to the present form.

3 | Near-surface Q_L from higher-order noise correlations

Extracting near-surface Q_L between 1-4 Hz from higher-order noise correlations in the Euroseistest area, Greece

Annabel Haendel[1], Matthias Ohrnberger[1], Frank Krüger[1]

[1] *Institute of Earth- and Environmental Science, University of Potsdam, Karl-Liebknecht Str. 24-25, D-14476 Potsdam*

Published in:
Geophysical Journal International
November 2016, Volume 207(2), pp. 655-666
DOI: 10.1093/gji/ggw295
Link: https://academic.oup.com/gji/article-abstract/207/2/655/2583627?rss=1

Abstract Knowledge of the quality factor of near-surface materials is of fundamental interest in various applications. Attenuation can be very strong close to the surface and thus needs to be properly assessed. In recent years, several researchers have studied the retrieval of attenuation coefficients from the cross correlation of ambient seismic noise. Yet, the determination of exact amplitude information from noise-correlation functions is, in contrast to the extraction of travel times, not trivial. Most of the studies estimated attenuation coefficients on the regional scale and within the microseism band. In this paper, we investigate the possibility to derive attenuation coefficients from seismic noise at much shallower depths and higher frequencies (>1 Hz). The Euroseistest area in northern Greece offers ideal conditions to study quality factor retrieval from ambient noise for different rock types. Correlations are computed between the stations of a small scale array experiment (station spacings <2 km) that was carried out in the Euroseistest area in 2011. We employ the correlation of the coda of the correlation (C^3) method instead of simple cross correlations to mitigate the effect of uneven noise source distributions on the correlation amplitude. Transient removal and temporal flattening are applied instead of 1-bit normalization in order to retain relative amplitudes. The C^3 method leads to improved correlation results (higher signal-to-noise ratio and improved time symmetry) compared to simple cross correlations. The C^3 functions are rotated from the ZNE to the ZRT system and we focus on Love wave arrivals on the transverse component and on Love wave quality factors Q_L. The analysis is performed for selected stations being either situated on soft soil or on weathered rock. Phase slowness is extracted using a slant-stack method. Attenuation parameters are inferred by inspecting the relative amplitude decay of Love waves with increasing interstation distance. We observe that the attenuation coefficient γ and Q_L can be reliably extracted for stations situated on soft soil whereas the derivation of attenuation parameters is more problematic for stations that are located on weathered rock. The results are in acceptable conformance with theoretical Love wave attenuation curves that were computed using 1-D shear wave velocity and quality factor profiles from the Euroseistest area.

Keywords Seismic attenuation · Interferometry · Coda waves

3.1 Introduction

Attenuation is an important ingredient of the propagation of seismic waves in Earth's interior for understanding and describing the spectral decay and the phase distortion of signals at high frequencies. Its knowledge is, for example, essential for ground-motion prediction in seismic hazard analysis (e.g. Campbell, 2009) or for the correction of path and site effects when investigating seismic sources (e.g. Hough, 1997; Müller, 1985).

Both, body and surface wave recordings from local or regional earthquakes (e.g. Sato et al., 2012) or from active source experiments (e.g. Langston et al., 2005; Xia et al., 2013) are used to study the attenuation structure in the crust. In seismic hazard analysis, near-surface attenuation (described by the parameter κ_0) is estimated from the spectral decay of S-waves traveling almost vertically through the uppermost layers of the ground to the surface (Anderson & Hough, 1984). Several studies, some of which employed downhole sensors to obtain at-

tenuation parameters as a function of depth, show that attenuation is very strong close to the surface but becomes less important at depth (e.g. Abercrombie, 1997; Anderson & Hough, 1984; Cranswick, 1988; Hanks, 1982, and references therein). Abercrombie (1997) and Aster & Shearer (1991), for example, reported that, for frequencies between ∼3-25 Hz and ∼20-80 Hz, respectively, up to 90 per cent of the total energy absorption of the crust occurs in the upper 3 km and 50 per cent in the upper 300 m. Attenuation studies from borehole experiments are rare due to cost of drilling and earthquake data is mainly available in regions of high seismic activity. Thus, we focus on quality factor Q estimation from ambient seismic noise as a potential low cost method for low seismicity regions.

In recent years, several researchers have studied the retrieval of attenuation properties from the cross correlation of seismic noise (e.g. Lawrence et al., 2013; Lin et al., 2011; Prieto et al., 2009; Weaver, 2011; Weemstra et al., 2013). The time derivative of the long term cross correlation of noise between two sensors converges to the Green's function between these sensors (Gouédard et al., 2008). It can be used to infer subsurface information using the same analysis techniques that are traditionally applied in earthquake studies or active source experiments. For example, the extraction of seismic velocities from seismic noise has been successful in many applications (e.g. Ekström et al., 2009; Hannemann et al., 2014; Pilz et al., 2012; Shapiro et al., 2005).

The retrieval of reliable amplitude information and based on that attenuation parameters is more difficult. Several factors contribute to the amplitude information of a seismic record like, for example, geometrical spreading, intrinsic attenuation, scattering, site amplification or focusing and defocusing effects. Noise CC amplitudes are additionally affected by variations of the ambient seismic field both in space and time and by the data pre-processing itself (e.g Weaver, 2011). Nevertheless, multiple studies demonstrated that seismic noise carries information about the anelastic structure of the Earth and that it can be reliably extracted. Prieto et al. (2009) fitted the real part of the azimuthally averaged coherency of the ambient seismic wavefield to a damped Bessel function to measure frequency dependent phase velocity and attenuation coefficients. This method was further investigated numerically by Lawrence et al. (2013) and successfully applied for imaging 3-D variations of the quality factor on the reservoir (∼1-20 km, Weemstra et al., 2013) and the regional scale (∼30-200 km, Lawrence & Prieto, 2011; Prieto et al., 2011). Lin et al. (2011) found that the spatially averaged amplitude decay of time-domain CCs is in good conformance with attenuation measurements from regional seismic events at the USArray. Weaver (2011, 2013) modelled noise source intensity by a radiative transfer equation and then simultaneously fitted the CC amplitudes from synthetic linear arrays for spatially varying attenuation coefficients and site amplification factors. Liu et al. (2015) developed an inversion algorithm to derive interstation attenuation from three-element linear arrays (station triplets).

Studies by Cupillard & Capdeville (2010) and Tsai (2011), on the other hand, showed that non-uniform distributions of noise sources can severely bias CC amplitudes and hence attenuation results. Zhang & Yang (2013) tried to overcome

these difficulties by computing higher order noise correlations (C^3) as was originally proposed by Stehly et al. (2008). Higher order correlations make use of the scattered coda portion of the first correlation step that is assumed to represent a more homogeneous source distribution than the original seismic noise field. Zhang & Yang (2013) successfully inferred the attenuation structure below the western United States for the seismic wave period of 18 s by applying C^3 to data of the US-Array. Their results were comparable with attenuation coefficients derived from earthquakes that were recorded at the same stations.

Most of the studies listed above focus on the retrieval of crustal and upper mantle attenuation parameters within the ocean microseism band. Our motivation lies in the estimation of Love wave phase slowness and quality factor Q in the very shallow subsurface (0-$\sim$500 m) and for frequencies closer to the frequency range of interest for seismic hazard and engineering applications ($>$1 Hz). We use in this study the notation Q_L for the apparent quality factor of Love waves. For a homogeneous half space, Q_L would be equivalent to the shear-wave quality factor Q_s. In stratified media Q_L is dependent on the eigenfunctions of Love waves and is therefore a function of frequency. The emphasis of this study is on the estimation of Q_L rather than the quality factor of Rayleigh waves because only then a direct relation with Q_s (and not also Q_p as for Rayleigh waves) is given. Q_s on the other hand is important for seismic hazard studies which generally utilize horizontal ground motions and shear wave attenuation factors in their analysis.

We adopt the method of Zhang & Yang (2013) and compute CCs between stations of a small scale array experiment (interstation distances $<$2 km) that was carried out in the Euroseistest area in northern Greece in 2011 (Hannemann et al., 2014). The correlograms are rotated to the ZRT system and we focus on the T-component and Love waves only. As the chosen test site has a complex surface geology, we estimate frequency dependent Q_L from the amplitude decay of Love waves for stations located either on soft soil or on weathered rock. One of our objectives is to test whether the chosen method is able to resolve attenuation differences between both rock types. The outcomes are compared with theoretical Q_L curves that are based on 1-D shear-wave velocity (v_s) and Q_s profiles that are reported for the Euroseistest valley below the stations. We furthermore test the applied methods by recovering the theoretical phase slowness and Q_L curves from synthetic seismograms that were computed for the same 1-D profiles.

3.2 Study Area and Data

We apply noise interferometry in the Mygdonia sedimentary basin in Northern Greece (Fig. 3.1) that is situated between lakes Lagada and Volvi and lies approximately 30 km to the northeast of the city of Thessaloniki. The Mygdonia basin is an European experimental test site (Euroseistest; Pitilakis et al., 2013) for integrated studies in earthquake engineering, engineering seismology, seismology and soil

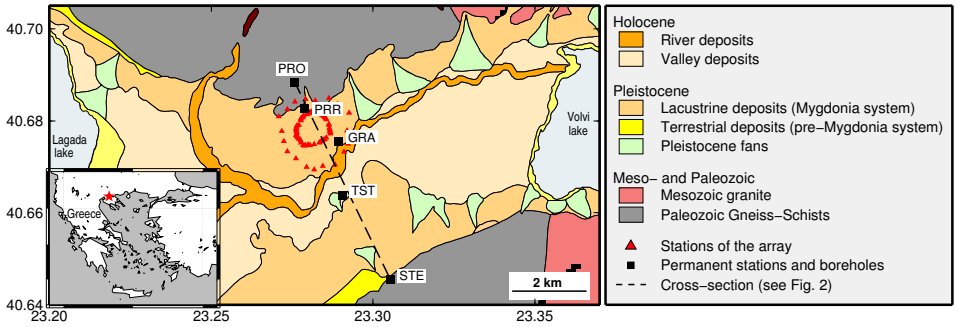

Figure 3.1: Euroseistest area in Northern Greece. Geological map of the study area and array configuration. The description of the geological units is taken from Hannemann et al. (2014) and Ktenidou et al. (2015). The positions of different permanent stations and boreholes are marked.

dynamics. Several geophysical and geotechnical experiments have been conducted in this area within the last 20 yr.

We use data from an array experiment that was carried out in August and September 2011 in the northern part of the Mygdonia basin. The array consisted in total of 27 instruments from which 19 stations were placed on fixed positions on an outer circle with a diameter of approximately 1.8 km, running for a period of two weeks. Eight mobile units were installed on an inner circle with a diameter of around 700 m. These stations were moved throughout investigation, running either for several hours during day or night time before being moved to a new position on the circle. Interstation distances range from a few tens of meters to approximately 2 km. All stations recorded with a sampling frequency of 100 Hz. We refer to Hannemann et al. (2014) for further details on the array and on the instrumentation.

Table 3.1: v_s and Q_s for the formations shown in Fig. 3.2. $Q_{s,lab}$ stems from laboratory testing as reported in Pitilakis et al. (1999). v_s and $Q_{s,geo}$ were determined in geophysical experiments by Jongmans et al. (1998).

Formation	A	B	C	D	E	F	G*	G
v_s (m/s)	130	200	300	450	650	800	1250	2600
$Q_{s,geo}$	15	20	30					
$Q_{s,lab}$	15	20	30	40	60	70	100	200

The Mygdonia valley is mainly a 2-D structure (compare to Fig. 3.1 and the NNW-SSE cross-section in Fig. 3.2). Site conditions range from weathered rock formations at the basin edge right below the northernmost stations of the array to around 200 m deep sediments at the basin center where the southeasternmost stations of the array are situated. We thus expect significant variations in velocity and attenuation estimates when performing noise CCs between stations located at different parts of the array.

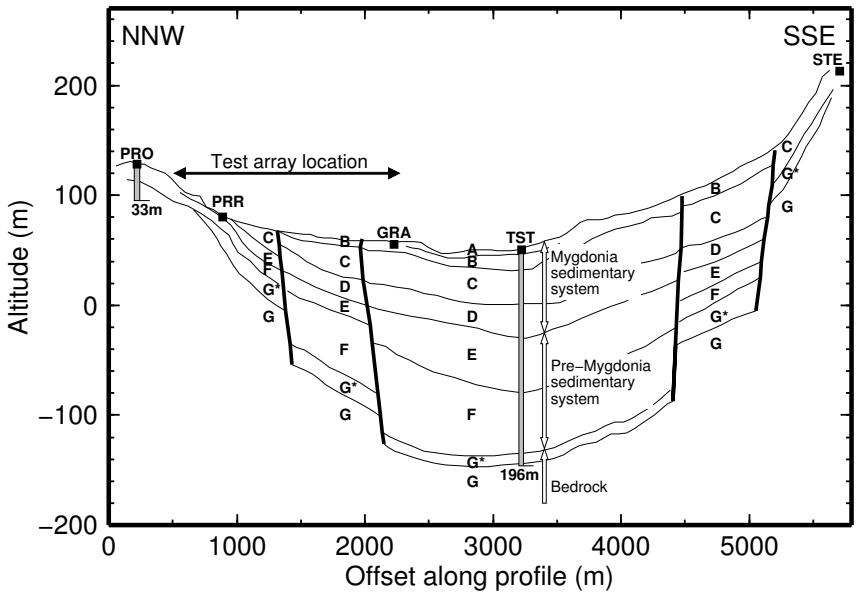

Figure 3.2: NNW-SSE profile (dashed line in Fig. 3.1) through the Mygdonia basin (after Pitilakis et al., 1999) showing the geological structure. The locations of the test array, the two boreholes at PRO and TST along with their corresponding depths and the positions of the two permanent stations PRR and GRA are indicated.

Q_s estimates of the Mygdonia valley are available from the studies of Jongmans et al. (1998) and Pitilakis et al. (1999) and are listed in Table 3.1. Jongmans et al. (1998) derived Q_s down to 40 m depth (formations A-C in Fig. 3.2 and Table 3.1) from attenuation measurements of Rayleigh waves. The authors assumed a frequency independent Q_s and analyzed their data between 1-4 Hz, the same frequency range as in our study. They performed their inversion for six profiles oriented parallel to the graben valley axis that were measured in different parts of the basin, and computed a mean Q_s profile from all curves. They observed that Q_s in the shallow layers is very similar for all profiles irrespective of the position within the basin ($Q_s \sim 15 - 30$ in the upper 15 m). Pitilakis et al. (1999) extracted damping parameters in the laboratory from samples of the TST borehole that is located in the center of the valley and penetrates down to the bedrock in 200 m depth (Figs 3.1 and 3.2 and Table 3.1). The laboratory values are in good agreement with the Q_s measurements of Jongmans et al. (1998) for formation A-C. Pitilakis et al. (1999) inferred a Q_s of 200 for formation G, the bedrock at the bottom of the borehole (see also Table 3.1).

The positions of the two permanent accelerometric stations GRA and PRR are shown in Figs 3.1 and 3.2. 1-D v_s and Q_s information below these stations are used in the following to compute theoretical phase slowness and Q_L curves for comparison with the noise based estimates of this study.

3.3 Noise correlation and C^3

It is well established that the time derivative of the CC of an equipartitioned random wavefield recorded at two positions converges to the Green's function. This in turn is equivalent to the impulse response obtained from an active source experiment (e.g. Lobkis & Weaver, 2001; Wapenaar, 2004). In theory, reconstruction of the full Green's function requires noise source energy to be emitted from all directions with equal strength (Snieder et al., 2007)—a condition that is generally not fulfilled in practice. In this context, scatterers play an important role, especially for sensor pairs oriented oblique or perpendicular to the predominant noise propagation direction, by acting as secondary sources.

Several authors used the homogenizing effect of scattering and computed correlations of earthquake coda recorded at two receivers (e.g. Campillo & Paul, 2003; Paul et al., 2005). Seismic coda at long lapse times is assumed to be at least partially diffuse and equipartitioned (Colombi et al., 2014; Paul et al., 2005) being thus better suited to converge to an accurate CC based Green's function. However, earthquake coda recordings are sparse while ambient noise is omnipresent.

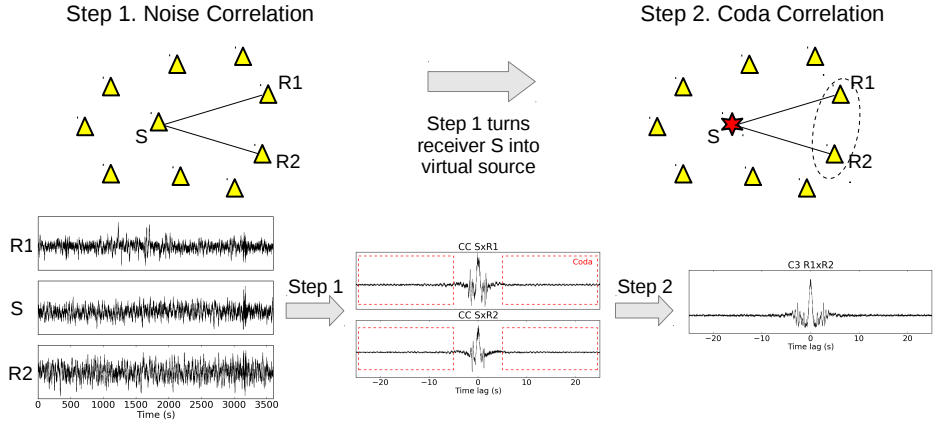

Figure 3.3: Illustration of C^3 procedure. Top: by correlating seismic noise recorded at two receivers R1 and R2 with noise recorded at a third station (coda station) S (step 1), a virtual source is placed at the position of station S. The virtual earthquake recording can be used to compute the coda correlation between R1 and R2 (step 2) that leads to $C^3_{R1\times R2}$. Bottom, from left to right: original noise recordings, the obtained cross-correlation results after step 1 and the final C^3 correlation trace after step 2.

Stehly et al. (2008) proposed to correlate the coda of noise CCs to obtain an improved CC function (CCF). The basic idea is that noise correlations also contain the coda part of the Green's function and that this coda is less sensitive to the source anisotropy of the original ambient field (Colombi et al., 2014). The correlation of the coda of the correlation procedure, in short C^3, is illustrated in Fig. 3.3. To extract the C^3 function between two receivers R1 and R2 one has to first calculate the correlation between a third station S (also termed coda station) and R1

and R2, which leads to $CC_{S\times R1}$ and $CC_{S\times R2}$, respectively. The obtained correlations represent the time series recorded at R1 and R2 that would be obtained for a virtual source placed at the position of sensor S. Correlation of the coda segments of $CC_{S\times R1}$ and $CC_{S\times R2}$ then gives $C^3_{R1\times R2}$. Different stations of the network can be chosen as coda station to obtain the same correlation $C^3_{R1\times R2}$. Averaging over time windows and coda stations leads to the final $C^3_{R1\times R2}$ function. The C^3 method combines the advantages of both using long noise time series for the computation of correlations and the diffuse character of the seismic coda. As a consequence, C^3 functions generally show improved time symmetry even if the original noise sources are not well distributed (Stehly et al., 2008).

Cupillard & Capdeville (2010) proved numerically that amplitudes of simple CCs are strongly dependent on the distribution of noise sources. The C^3 method, on the other hand, does not suffer from noise source anisotropy as has been shown empirically by Zhang & Yang (2013). Zhang & Yang (2013) were able to retrieve surface wave attenuation properties using the C^3 approach within the ocean microseism band. They demonstrated that C^3 based attenuation coefficients are much closer to attenuation coefficients retrieved from earthquakes than the ones obtained with conventional noise correlations. We therefore decided to adapt the C^3 method and apply it to our dataset. We focus in our work on the frequency band between 1-4 Hz and interstation distances much smaller than those by Zhang & Yang (2013) in order to obtain attenuation parameters within the uppermost layers of the subsurface.

3.4 Data Processing

The pre-processing of the array data was based on the works of Bensen et al. (2007) and Seats et al. (2012) and followed mainly the pre-processing steps of Hannemann et al. (2014) for the same array. We split the records into one hr data windows, removed the offset and applied a 5 per cent cosine taper at the edges of each window. The original response of each instrument was removed because different sensor types were used within the array. The data were then bandpass filtered between 0.5-30 Hz to minimize the impact of low-frequency signals on the frequency range of interest ($>$1 Hz). We did not apply 1-bit pre-processing to the recordings because it has been shown theoretically and numerically (Cupillard & Capdeville, 2010; Cupillard et al., 2011; Weaver, 2011) that relative amplitudes are not retained for non-uniform distributions of noise sources if such a strong non-linear operation is applied. Instead, we utilized the same pre-processing method as in the C^3 study of Zhang & Yang (2013) to reduce the influence of earthquake signals and temporal fluctuations of the noise field.

First, transient signals were removed by identifying signal windows with an amplitude larger than an hourly median amplitude level. To do so, we computed the envelope mean of each 5 min long data window and compared it to the median

of the envelope amplitudes of the whole hour data trace. The signal of a window was replaced with zeros if it exceeded two times the hourly amplitude level.

Second, we employed the temporal flattening technique that was originally proposed by Weaver (2011) to remove temporal fluctuations in the noise source intensity. For our data set, we expect fluctuations due to different day-night activity on the streets and cities surrounding the array or due to short-lived weather phenomena (e.g. wind gusts). Seasonal variations are not expected in the data given the short deployment time of two weeks. Temporal flattening was performed by normalizing each 1 hr data recording by the global noise amplitude level obtained for the whole array during this hour. The global noise amplitude level was defined as the quadratic mean (RMS value) of noise standard deviations obtained for all receiver and coda stations of the array (Weaver, 2011).

To finalize the processing steps of the first correlation, spectral whitening was applied and CCs were computed between all possible station pairs of the array for 1 hr long non-overlapping time windows resulting in noise CC's with a coda tail of 1800 s length. The CC results were not further stacked.

CC coda was then extracted between 20 and 1800 s for each of the original 1 hr noise windows. The beginning of the coda window was set to two times of the longest direct S-wave traveltime that can be observed in the array which in our case estimates to 20 s [minimum observed shear-wave velocity $v_{\mathrm{s,min}} \sim 200\,\mathrm{m\,s^{-1}}$, maximum interstation distance $r_{\mathrm{max}} \sim 2\,\mathrm{km}$; Hannemann et al. (2014)]. We tested different coda window lengths in the analysis. The signal-to-noise ratio (SNR) of the C^3 function improved with increasing window length. We compared increasing coda window lengths up to very large lapse times after the direct wave arrival (max. 1800 s time lag on the CCF). As interesting observation we found a continuously increasing SNR with lapse time. In order to validate these findings, we checked whether the C^3 function can be retrieved from coda windows on the CC traces between 1000 and 1800 s. The C^3 functions are almost identical (despite a lower SNR) to those obtained from CC coda windows between 20 and 1800 s. It follows that multiple scattered energy is present at very large lapse times in the CCF. We therefore use the whole coda tail starting from 20 s up to 1800 s for the computation of the C^3 function.

Following Zhang & Yang (2013), we added the coda from the positive lag CCF and the coda from the time-reversed negative lag CCF (mirror stacking) as this approach yielded the best C^3 results. A similar good result is obtained when computing C^3 individually for the positive and negative lag arm and summing them afterwards (Froment et al., 2011). Yet, this approach is computationally more expensive as the C^3 calculation step has to be performed twice.

The CC coda segments were split into 100 s long time windows with an overlap of 50 per cent (Welch's method, Seats et al., 2012; Welch, 1967). The offset was removed for each individual time window and spectral whitening was applied. For each station pair in the array, we then computed C^3 using several selected coda stations (selection criteria are described in section 3.5). C^3 was only obtained from the coda segments of receiver and coda station pairs that have been computed for

the same time instance of data recordings. For each hour, C^3 results of 25 s length were then written to disk.

Averaging over time windows and coda stations led to the final C^3 functions. Time averaging was performed after the C^3 step. Time stacking right after the CC computation leads to fewer coda time series and therefore lower C^3 SNR's as the SNR of a correlation function increases with the square root of data length (e.g. Bensen et al., 2007; Sabra et al., 2005). For example, for a receiver pair of the outer circle, that was recording for approximately two weeks, roughly 330 hr of data were used to compute the first correlation. The length of the selected coda window (1780 s) and the number of coda stations results in about 2900 hr of data for the computation of C^3. Only 9 hr of data remain if time averaging is performed directly after the CC step.

We rotate the C^3 functions derived for the ZNE components into the ZRT system and focus only on the transverse (T-) component. The rotation into a new coordinate system is usually performed on the raw noise or coda traces and correlations are then computed between Z-Z, R-R, T-T or between mixed components (e.g. Campillo & Paul, 2003; Paul et al., 2005; Poli et al., 2012). Yet, CC and rotation are linear operations and can be interchanged. We rotate the final C^3 instead of the raw traces because between-receiver and coda station-receiver azimuths are different for each combination of stations.

3.5 Quality of C^3

We evaluate the quality of C^3 depending on the chosen set of coda stations and C^3 performance in comparison to CC. For this reason, we compute the SNR of each trace by dividing the envelope maximum of a window around the theoretical signal arrival by the standard deviation derived for 6 s long coda taken from the end of the correlation trace. The choice of the signal window is based on the slowest (150 m s^{-1}) and fastest (2500 m s^{-1}) observed velocities within the array. The SNR is calculated on the T-component that is bandpass filtered between 1-4 Hz.

First, we test if coda stations that are far from the chosen receiver pair should be excluded from the final C^3 stack. Hannemann et al. (2014) excluded station pairs with an interstation distance larger than 1000 m from their analysis because they observed a group velocity change above this distance that they related to the propagation of higher mode surface waves. We therefore compute C^3 SNR for a receiver pair (T08 and T09; see Fig. 3.4a and Fig. 3.7a for the location of of the stations) and plot the results as a function of distance between each of the two receiver stations and the coda station. We expect that the C^3 SNR decreases with increasing receiver-coda station separation but both close and far distances lead to comparable SNRs. No distance penalty for far coda stations is thus employed.

The type of the chosen coda station (either a receiver of the inner circle or one of the outer circle), on the other hand, has a much larger influence on the SNR. Stations of the inner array circle (W stations) were only running for some hours

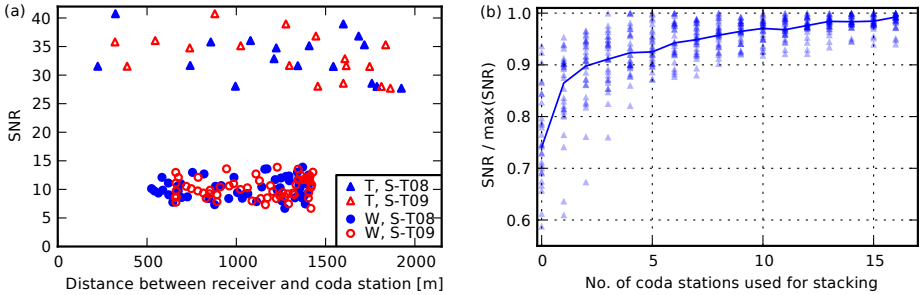

Figure 3.4: (a) C^3 SNR for receiver pair T08xT09 computed for single coda stations (see Fig. 3.7a for the respective station locations). The horizontal axis shows the distance between receiver T08 and the coda station (S-T08) or receiver T09 and the coda station (S-T09). T=coda station of outer circle running for two weeks. W=coda station of inner circle running only for some hours. (b) C^3 SNR (normalized to the maximum SNR) for several receiver combinations of the outer array circle when adding successively new coda stations to the C^3 stack. The solid line is the median of the data points.

while those of the outer circle (T stations) were in operation for several days. If two T stations are used as receiver pair, a T station as coda station will lead to a higher SNR as if a W station would be used that has a much shorter recording time. We exclude therefore all W stations as coda stations. Fig. 3.4(b) shows the SNR when adding successively new coda stations (only T stations) to the C^3 stack of several outer circle receiver pairs. The SNR increases rapidly for the first 5 stacked coda stations. Adding more coda stations to the stack improves the SNR to much lesser extent. A small number of coda stations is thus sufficient to obtain good C^3 SNRs.

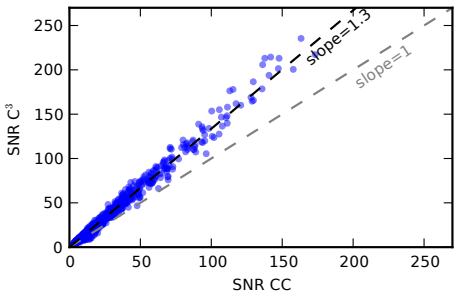

Figure 3.5: Comparison of CC and C^3 SNR for all possible receiver combinations of the array.

Fig. 3.5 compares CC and C^3 SNRs computed for the same receiver pairs. The C^3 SNR is on average 30 per cent higher than the CC SNR. We also investigate the time symmetry of CC and C^3 functions. We derive the correlation coefficient ρ between the positive and reversed negative lag arm of all computed CC and C^3's as a function of receiver azimuth. A value of $\rho = 1$ corresponds to high symmetry

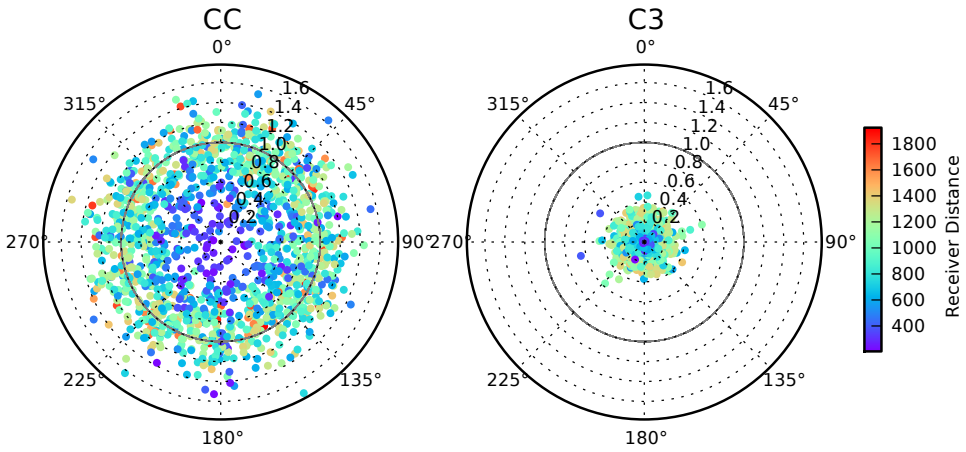

Figure 3.6: Correlation coefficient ρ between the positive and reversed negative lag arm of CC and C^3 as a function of receiver azimuth. The distance between the receivers is colour coded. The correlation coefficient is plotted as 1-ρ so that a value close to the center corresponds to high symmetry between positive and negative lag of a correlation function. Values close to two represent pure anticorrelation and values around one show asymmetry between the lag arms.

between positive and negative lag time. Fig. 3.6 shows for illustration purposes 1-ρ instead. Small values of 1-ρ that plot close to the center of the circle then correspond to high symmetry while higher values that scatter over larger parts of the circle indicate asymmetry. A value of two corresponds to pure anticorrelation. As can be expected, receivers with large separation have a stronger time asymmetry due to their low SNR and plot closer to the value of 1 in the circle.

In addition to the correlation coefficient, we directly check the symmetry of CC amplitudes. For each trace, we take the difference between maximum and minimum amplitude on the negative lag and divide this value by the same amplitude difference that is taken from the positive lag. CC amplitudes differ up to a factor of 2-3 between both lag arms while the deviation factor observed for C^3 functions is always less than 1.5. This confirms the results of Stehly et al. (2008) who noted that C^3 functions show a much stronger time symmetry than simple CCFs. The additional correlation step and the stacking over several coda stations, thus, greatly improves correlation quality for the Euroseistest array dataset.

3.6 Results

We estimate mean phase slowness and Q_L curves as a function of frequency for stations being situated either mainly on soft soil above the valley center in the southern part of the array or on weathered rock in the northern part of the array (Fig. 3.7a). Stations being located roughly parallel to the strike of the graben are chosen to ensure that the subsurface can be regarded as 1-D for stations being assigned to the same rock type. Throughout this paper, results that are associated with stations on soft soil are shown in gray whereas those on weathered rock are

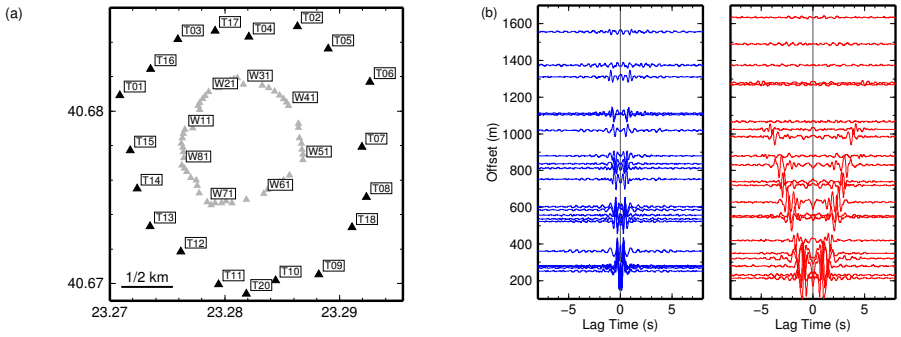

Figure 3.7: (a) Array configuration and chosen set of stations on soft soil (gray/red) and on weathered rock (black/blue). (b) Virtual shot gathers obtained from C^3 correlations for stations situated on soft soil and on weathered rock. Shown are the T-components filtered between 1-4 Hz.

colored in black (electronic version: soft soil in red and weathered rock in blue). We exclude the receivers T08 and T18 from the analysis of soft soil stations as they are situated on much softer ground than the adjacent receivers (s.a. Fig. 3.1). The waveforms recorded at T08 and T18 show much stronger dispersion and including both receivers in the analysis hampers the extraction of slowness and Q_L curves on soft soil. Furthermore, we exclude station pairs on weathered rock with distances smaller than 400 m because the Love wave arrivals on the positive and negative lag arm partly overlap at zero lag time. This is especially important for the analysis at lower frequencies.

We obtain virtual shot gathers for soft soil and for weathered rock stations (Fig. 3.7b) when plotting the T-component of the C^3 correlations as a function of distance. Love wave arrivals can be clearly identified in both cases propagating with much higher group velocity on weathered rock than on soft soil. As the C^3 results are symmetric (Stehly et al., 2008, and Fig. 3.6), we work in the following with the symmetric component of the CCFs (sum of positive and reversed negative lag arm) for deriving phase slowness and Q_L.

We cannot directly compare the v_s and Q_s measurements of Jongmans et al. (1998) and Pitilakis et al. (1999) with our phase slowness and Q_L estimates. Dependent on their wavelength, Love waves penetrate more or less deep into the ground. In order to make the observed results comparable with the reported structure of the Euroseistest valley we therefore compute theoretical phase slowness and Q_L dispersion curves that are based on selected 1-D profiles (Fig. 3.8).

Profiles GRA and PRO-1 are from Pitilakis et al. (1999) and correspond to the profiles S4 and S7 of their fig. 6, respectively. These profiles were taken along the NNW-SSE cross-section of the valley and formed the basis for the 2-D profile that is shown in Fig. 3.2. Profile S4 (here GRA) roughly coincides with the location of the permanent accelerometric station GRA and, taking the sensor locations of the test array into account, is considered to be representative for the structure below the soft soil stations. Profile S7 (in this study PRO-1) is located just south of the

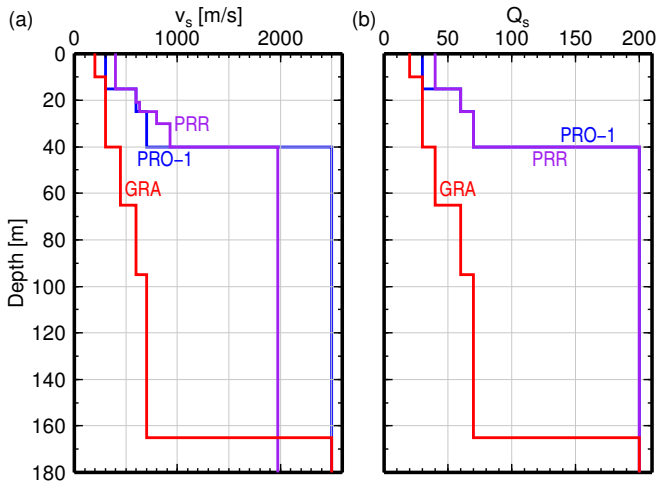

Figure 3.8: (a) Shear wave velocity and (b) shear wave quality factor as a function of depth for different 1-D profiles that represent the ground below the chosen soft soil (GRA) and weathered rock stations (PRR and PRO-1), respectively.

permanent station PRO and can be used to model the ground below the weathered rock type stations.

The v_s information of the profile below the permanent accelerometric station PRR shown in Fig. 3.8 is taken from the Euroseistest website and is based on the surface wave inversion results reported in Raptakis et al. (2000). Its location more or less coincides with the one of profile PRO-1. Nevertheless, reported velocities are slightly higher for layers above the bedrock and the bedrock velocity itself is lower than for profile PRO-1. No Q_s information is given for PRR, we therefore use the same Q_s profile as for PRO-1, changing only the Q_s value of the topmost layer from 30 to 40 as the shear-wave velocity of 400 m s^{-1} is more representative of layer D in Fig. 3.2 and Table 3.1. The position of both permanent stations GRA and PRR are shown in Figs 3.1 and 3.2.

3.6.1 Phase slowness estimation

We employ the slant stack method, a beamforming process, to derive phase slowness dispersion curves from the data (e.g. Yilmaz, 1987). The plane wavefield is transformed from the offset-time domain to the slowness-frequency domain. The transformation is achieved by correcting each correlation trace for a linear moveout time that is dependent on slowness and sensor offset and summing the amplitudes of all traces over the offset axis. Constructive interference occurs only if the correct phase slowness for a certain frequency is met so that the beampower maximizes. Gouédard et al. (2008) could show that the noise correlation slant stack method

provides very accurate phase velocity dispersion curves for Rayleigh and Love waves given that the medium is horizontally stratified.

We apply this method to the chosen set of stations on soft soil and weathered rock. A 10 and 5 s long window containing the Love wave arrival is cut for the analysis of soft soil and weathered rock stations, respectively, with a cosine taper of 2 per cent applied at the edges of each window. In a first step, amplitude offsets are removed from the records and each correlation trace is normalized to its maximum amplitude to account for amplitude differences due to different receiver separations. Moveout correction is then performed in the frequency domain for the specified slowness range. For each slowness and frequency, the semblance or normalized beampower (Neidell & Taner, 1971) of the phase shifted spectra is calculated using:

$$S(\omega, s) = \frac{|\frac{1}{N}\sum_{i=1}^{N} X_i(\omega, \tau_s)|^2}{\frac{1}{N}\sum_{i=1}^{N} |X_i(\omega, \tau_s)|^2} \tag{3.1}$$

with $\tau_s = s \cdot r$ the linear moveout time, s the slowness, r the sensor offset, ω the angular frequency and N the number of spectra over which is totalled. $X_i(\omega, \tau_s)$ corresponds to the ith phase shifted correlation spectrum.

Slant stacks for soft soil and weathered rock stations are shown in Fig. 3.9. Phase slowness dispersion curves are automatically extracted by fitting Gaussians to the slowness-semblance slices at every frequency. The estimated mean and standard deviations are exported and used as slowness mean and error estimate for the calculation of quality factor Q_{L}. Phase slowness on weathered rock could be extracted between 1.6-4.0 Hz. Phase velocity is around 2300 m s^{-1} at 1.6 Hz and 1550 m s^{-1} at 4.0 Hz. On soft soil, the usable frequency range lies between 1.1-2.6 Hz, where phase velocity is 550 m s^{-1} at 1.1 Hz and 330 m s^{-1} at 2.6 Hz.

The theoretical dispersion curves obtained for the fundamental and the first higher mode of the 1-D profiles PRO-1, PRR and GRA are shown as dashed-dotted and dotted lines in Fig. 3.9. The fundamental mode on soft soil coincides well with the dispersion curve that is obtained with the slant-stack method. Higher mode energy, on the contrary, cannot be identified in the frequency-slowness plot for soft soil indicating that the energy that is present in the correlograms is from the fundamental Love wave only. The dispersion curve on weathered rock fits better with the theoretical curve of profile PRR than PRO-1.

3.6.2 Attenuation retrieval

Love wave amplitude measurements inferred from C^3 correlations are used to retrieve Q_{L} information. The amplitude of a plane wave, single mode Love wave in the far-field travelling in a specific direction can be written as (e.g. Aki & Richards, 2002; Udias, 1999):

$$A_i(f, r_i) = A_0(f) \cdot \frac{1}{\sqrt{r_i}} \cdot e^{\frac{-\pi f r_i}{c(f) Q_L(f)}} \tag{3.2}$$

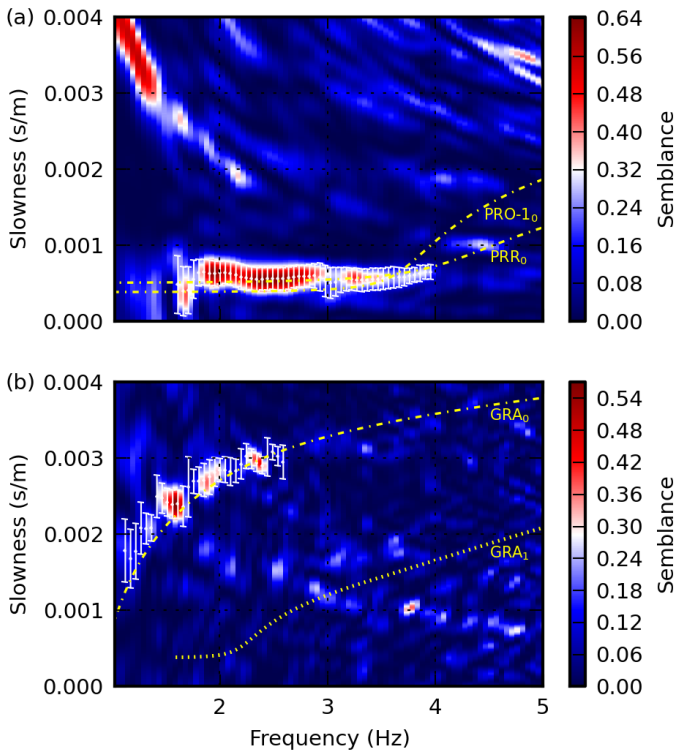

Figure 3.9: Frequency-slowness plots obtained for (a) weathered rock and (b) soft soil using the slant-stack method. The mean and standard deviation of the picked phase slowness curves are shown as white error bars. Dashed-dotted lines (PRO-1_0, PRR_0, GRA_0) are theoretical dispersion curves of the fundamental mode derived using the 1-D profiles shown in Fig. 3.8. The dotted line in the lower panel (GRA_1) is the theoretical dispersion curve for the first higher mode of profile GRA.

where A_0 is the unknown source term, c is phase velocity, Q_L is Love wave quality factor, r is distance and f is frequency. The term $1/\sqrt{r}$ accounts for geometrical spreading of surface waves, the exponential term describes material attenuation and scattering of the medium. We are working with amplitude ratios in order to get rid of the unknown amplitude term A_0 at the focus. Amplitudes are corrected for geometrical spreading before taking the ratio

$$R(f, \Delta r) = \frac{\sqrt{r_1} A_1(f, r_1)}{\sqrt{r_2} A_2(f, r_2)} = e^{\frac{-\pi f \Delta r}{c(f) Q_L(f)}} \tag{3.3}$$

where $\Delta r = r_1 - r_2$.

The theoretical arrival window for Love wave amplitude picking is computed for each trace individually. Window starting times are based on sensor offsets and the maximum derived phase velocity from the slowness analysis. Taking into account the slowest observed phase velocity, a window length of 6.0 s is chosen for stations on soft soil and a 3.0 s long window is selected for stations on weathered rock. We visually checked if the Love wave is fully contained in the chosen sig-

nal windows. Each trace is then Fourier transformed and amplitudes are directly picked in frequency domain. The amplitude spectra are smoothed prior to picking using a Gaussian narrow band-pass filter. The Gaussian filter has the form $\exp\left[-(2\pi(f-f_c))^2/(4b^2)\right]$ where f_c and b are the center frequency and bandwidth of the filter. After testing we decided to employ a bandwidth b of 1 Hz in the analysis because this value represents a good compromise between smooth Q_{L} curves while preserving the mean frequency dependence of Q_{L}.

The attenuation coefficient γ can be obtained by plotting the natural logarithm of the geometrical spreading corrected amplitude ratios against relative distance Δr:

$$\ln R(f,\Delta r) = \frac{-\pi f \Delta r}{c(f)Q_L(f)} = -\gamma(f)\cdot\Delta r \tag{3.4}$$

The slope is estimated by a least-square linear regression and corresponds to γ from which Q_{L} at a given frequency can be computed using the corresponding phase slowness value. From the available C^3 station combinations, 21 pairs on soft soil and 15 combinations on weathered rock are used for the estimation of the amplitude decay. As we are working with amplitude ratios, we always compute the amplitude of one C^3 station pair relative to another one. The station pairs are selected in such a way that they share the same 'virtual source' station and that part of their paths are overlapping. In doing so, 70 and 40 relative amplitude values can be formed between soft soil and weathered rock C^3 pairs, respectively, for use in the regression.

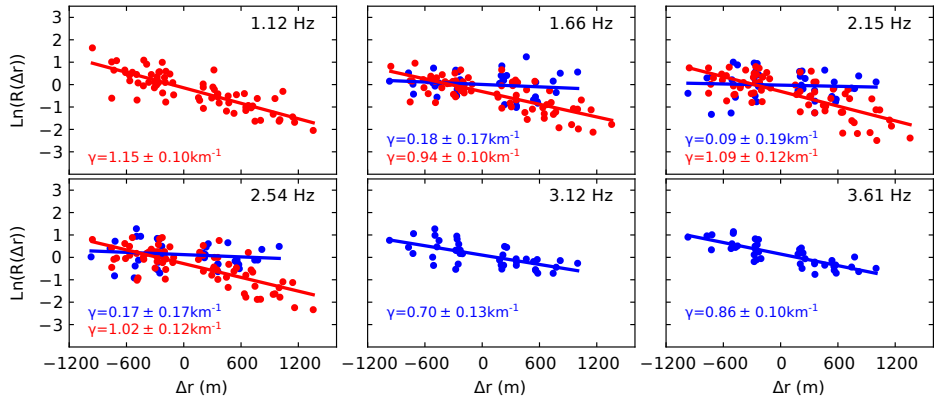

Figure 3.10: Amplitude decay with distance for different frequencies. Gray/red: stations on soft soil. Black/blue: stations on weathered rock. Fitted curves are shown as solid lines. The estimated slope corresponding to the attenuation coefficient γ along with its error is given at the bottom of each panel.

Fig. 3.10 shows the decay curves for different frequencies. The estimated attenuation coefficient γ and the corresponding standard deviation are given at the bottom of each panel. Although the scatter is large, it is clearly visible that the

amplitudes measured on soft soil decay faster with distance (almost constant γ of 1.0 km^{-1}) compared to those on weathered rock ($\gamma \sim 0.2$ km^{-1} at 1.7 Hz and 0.9 km^{-1} at 3.6 Hz).

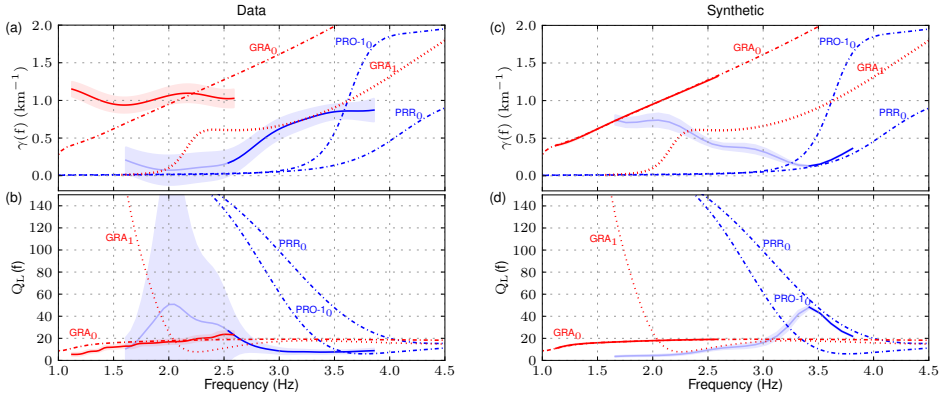

Figure 3.11: Top (a,c): γ and bottom (b,d): Q_L as a function of frequency. Left (a,b): solid lines show γ and Q_L computed from the C^3 correlations obtained in the Euroseistest area. Right(c,d): solid lines show the reconstruction of γ and Q_L for profiles GRA and PRR from synthetic seismograms. Standard deviation of the estimates ($\pm 1\sigma$) are indicated as shadowed areas and include, in the case of Q_L for real data (b), the slowness error. Curves for weathered rock below 3.4 Hz (synthetics) or 2.5 Hz (data) are plotted as thinner lines because values below this frequency are not reliable. Theoretical Q_L and γ curves obtained from 1-D profiles in the Euroseistest area (GRA, PRR, PRO-1, Fig. 3.8) are shown for comparison as dashed-dotted lines (fundamental mode) or dotted line (first higher mode). PRP-1$_0$, PRR$_0$, GRA$_0$ denote the theoretical fundamental and GRA$_1$ the first higher mode of the corresponding profiles.

Figs 3.11(a) and (b) show γ and Q_L as a function of frequency for weathered rock and soft soil (solid lines) as obtained from the data. We measure γ and Q_L only for frequencies were phase slowness values are available. Q_L for soft soil stations is in the range of 5.5-23 for frequencies between 1.2-2.6 Hz. For stations situated on weathered rock, Q_L lies between 10-50 for frequencies between 1.6-4 Hz. The shaded areas indicate the range of ± 1 standard deviation that takes into account the slope errors of the amplitude decay regression and, in the case of Q_L, also the estimated slowness errors. The theoretically derived γ and Q_L curves from the 1-D Euroseistest profiles are shown for comparison as dashed-dotted lines for the fundamental mode (PRO-1$_0$, PRR$_0$, GRA$_0$) and as dotted line for the first higher mode (GRA$_1$). The theoretical curves approach the Q_L or γ value of the uppermost layer at high frequencies and the value of the bedrock halfspace at low frequencies (this happens for soft soil below 1 Hz). For frequencies in between, Q_L and γ are influenced by the Love wave eigenfunction at the given frequency. The Love wave eigenfunction at a certain frequency is dependent on Q_s, v_s and the derivative of phase velocity with respect to v_s of not only one but several subsurface layers (Anderson et al., 1965).

3.7 Discussion

Before we discuss the observed attenuation curves we want to check if the applied methods (slant stack and amplitude decay) are suitable for the derivation of phase slowness, γ and Q_{L} within the frequency range that is considered in this study. We test this by performing forward calculations using the profiles GRA and PRR and computing synthetic seismograms for the same distance range as given by the data. A horizontal single force source emitting a dirac delta impulse is placed at the surface. We then use the same parameter settings and methods as applied to the C^3 functions to reproduce the theoretical phase slowness, γ and Q_{L} curves from the synthetic seismograms (Figs 3.11c and d).

The fundamental mode phase slownesses can be retrieved accurately both for soft soil and weathered rock profiles (not shown here). The same is true for γ and Q_{L} for soft soil profile GRA if only the fundamental mode Love wave is present in the synthetics. Slight deviations from the model curve occur if higher mode energy is allowed in the synthetic seismograms but these differences are very small and occur only above 2.3 Hz.

Major problems arise for the weathered rock profile PRR that shows deviations from the model curves within the complete frequency range that we investigate. These discrepancies are small for frequencies above 3.4 Hz but become larger for lower frequencies. The largest Q_{L} that is reached on weathered rock is 50 before it declines for decreasing frequencies. If we increase the distance range for the computation of weathered rock synthetics up to 30 km, γ and Q_{L} can be reproduced if at the same time a narrower Gaussian band pass filter is used (not shown). Only extremely small γ values and a Q_{L} above 150 remain troublesome. The very small γ values correspond to amplitude decay curves with a slope of almost zero. Tiny errors in the γ estimation then lead to large Q_{L} errors.

The γ and Q_{L} values that we derive from C^3 correlations for our dataset are given in Figs 3.11(a) and (b). The results are obtained using the assumption of a horizontally stratified medium without the effect of any station specific site amplification. The γ and Q_{L} values are mostly of the same order as the model values of GRA_0, PRR_0 and PRO-1_0 but some major differences in their frequency dependence can be observed.

For soft soil stations, the theoretical curve for the fundamental mode of profile GRA predicts an increase of γ with frequency. Using the C^3 dataset we observe the same frequency dependence between 1.7-2.2 Hz although the slope is different. At frequencies below 1.7 Hz or above 2.2 Hz γ values decrease with increasing frequency. Nevertheless, the resulting Q_{L} curve is roughly comparable with the Q_{L} curve of GRA_0 for all frequencies that we consider. Only the slope is slightly different from the model curve.

On weathered rock, γ is small at low frequencies and raises as the frequency increases. PRR_0 and PRO-1_0 share the same frequency dependence, yet, the frequency where the observed curve starts to bend upward is much lower. The resulting Q_{L} values are hence small above 3 Hz and only start to increase below this

frequency. As for the synthetics, the observed Q_L is always smaller than 50 and never reaches the bedrock Q_s of 200 at the lowest frequencies. We note that the derived γ curve for weathered rock shows errors that are of the same order or even larger than the γ value itself at frequencies below 2.5 Hz. As for the synthetic curves, γ at these frequencies is relatively small (below 0.2 km^{-1}) and even tiny errors in its value will lead to larger deviations in the Q_L estimate. That is why we plotted the γ and Q_L curves for weathered rock stations thinner for frequencies below 2.5 Hz. γ at these frequencies may be still close to the true value but the Q_L estimates are certainly not realistic. The shaded error ranges illustrate that much larger Q_L values are also possible.

The differences between the observations and the curves that are computed from the Euroseistest profiles can have several reasons. First, some of the assumptions that we make for the computation of attenuation parameters may not be met. We consider the ground below the soft soil and weathered rock stations as 1-D. Yet, the stations are not perfectly oriented parallel to the graben structure and the Euroseistest area has a rather complex geology. 2-D or even 3-D effects thus might affect the attenuation estimates. In addition, site effects have been observed in the valley and could be different for stations that we assigned to the same rock type. But this should mainly result in a larger scatter of the amplitude ratios. Mean amplitude decays with distance and, hence, the γ value itself should not be affected.

Second, we do not exclude higher mode Love waves from the C^3 correlograms and can therefore not rule out the possibility that they might influence the results. Yet, we do not observe any higher mode energy in the frequency-slowness plots constructed from the data. It is therefore unlikely that higher modes have a large influence on the results. Furthermore, problems with higher mode energy in the soft soil synthetics did only occur for frequencies above 2.3 Hz. The decrease of γ with frequency above 2.2 Hz that we observe in the data could thus maybe be attributed to this effect.

Third, amplitudes decrease only very slowly on weathered rock. We already noted in the synthetic tests that larger array apertures would be necessary to reliably extract γ and Q_L on weathered rock. Deviations from the model curves could thus also be effected by the limited distance range that we consider in this study. This could also be the case for γ values on soft soil at low frequencies where the frequency dependence is very different from the model curve GRA_0.

Finally, profiles GRA, PRR and PRO-1 might not be fully representative for the ground below the soft soil and weathered rock stations. While v_s in the Euroseistest area is rather well constrained, Q_s estimates are more uncertain, especially at depths below 30 m. Fig. 3.2 is a simplified model for the Euroseistest area but the real structure is more complex. Jongmans et al. (1998), for example, noted that, what is summarized as formation D in Fig. 3.2 could be composed of sediments in the center of the valley but could at the same time correspond to slope deposits or weathered bedrock at the edges of the basin. Different Q_s estimates must then be expected for the same formation but only a single value is given that was determined in the center of the valley at borehole TST. The Q_s values given in Table 3.1

for the deeper formations D-G were derived in the laboratory. These measurements are usually performed on small rock samples and at very high frequencies so that Q_s that one would observe at seismic frequencies and length scales might differ from the laboratory values. Attenuation estimates from the laboratory are generally a measure of intrinsic damping alone (e.g. Pujol & Smithson, 1991) and do not include scattering attenuation caused by material inhomogeneities. What we measure with C^3 correlations, on the other hand, is an effective Q that is made up from a combination of intrinsic and scattering attenuation. Scattering attenuation in the Euroseistest area could, for example, be caused by the deflection of surface waves from the valley borders or by small heterogeneities within single formations. Pitilakis et al. (1999) described the main geological features of each formation. He summarized that the upper formations contain several thin sublayers and mixtures of material fragments while some of the deeper formations are characterized by thick, alternating sublayers and fragments of stones and gravels. Such a geology will certainly lead to scattering attenuation. Ktenidou et al. (2015) compared their κ_0 estimates with the Q_s values given for the Euroseistest area. They noted that their κ_0 observations cannot be explained by the intrinsic Q_s values alone but that there is very likely a scattering contribution.

It would be possible to slightly modify the profiles GRA, PRO-1 and PRR to see if a better fit with the observed γ and Q_L curves is possible. Varying only the Q_s values and keeping v_s of each formation constant, an inverse Q_s profile (higher Q_s at the surface and lower values above bedrock) would be necessary to obtain the flat γ curve on soft soil. The bending of the γ curves on weathered rock, on the other hand, can only be shifted to lower frequencies if also v_s or the thickness of formations is varied. Changes in v_s or layer thicknesses, however, will also lead to differences in the predicted phase slowness curves shown in Fig. 3.9. A simultaneous fit to the observed phase slowness and γ curves would thus be necessary but is beyond the scope of this study.

3.8 Conclusion

In this work, we utilized higher order noise CCs to infer the attenuation structure of the shallow sub-surface in the Euroseistest area in Greece. We show that the C^3 method is successful in mitigating the effect of uneven noise source distributions and that it leads to correlation functions with a higher SNR than simple noise CCs. The employed procedure allows to extract mean phase slowness, the attenuation coefficient γ and quality factor Q_L of Love waves as a function of frequency on soft soil. The recovery of attenuation parameters from correlations obtained on weathered rock is more troublesome, especially for lower frequencies. This is mainly due to the fact that the amplitude decay with increasing interstation distance is very small at weathered rock conditions and that larger interstation distances would be necessary to reliably observe the amplitude decrease.

We are aware that the Q_L measurements could contain several sources of error like the negligence of site effects or focusing and defocusing that can be expected in a complex geological structure like the Euroseistest valley and that lead to a larger scatter in the amplitude decay curves. Yet, the attenuation coefficient estimates for Love waves are mainly in conformance with theoretical values derived from 1-D v_s and Q_s profiles from the Euroseistest area. The Q_s values that are used to model Q_L curves are purely intrinsic and do not include any scattering attenuation. Some differences between the observations and the theoretical γ and Q_L curves may thus be attributed to scattering in the Euroseistest valley or simply to the fact that the chosen profiles are not fully representative for the structure below the stations.

The measurement of near-surface attenuation is up to now bounded to regions of high seismicity or it requires the application of active source or VSP experiments. Our study moves a step further by extracting this information from ambient seismic noise. The only other study to our knowledge computing attenuation from seismic noise at similar distances and at frequencies above 1 Hz is the one of Liu et al. (2015). The authors infer Q from Rayleigh waves between triplets of stations in the San Jacinto fault zone. More research is needed to proof the reliability of these methods at shallow depths and for different subsurface conditions. The outcomes could then be utilized in seismic hazard analysis where near-surface attenuation (the parameter κ_0) is very important to successfully predict future ground motions.

Acknowledgements The array measurements were carried out with funds of the project "Efficient high-frequency surface wave tomography for seismological site characterization" of the IKY-DAAD2011 Greek-German scientific cooperation program. We are grateful to the team that conducted the array field work in 2011, namely K. Hannemann, A.M. Lontsi, M. Anthymidis and A. Savvaidis. The field measurements have been mainly realized with instruments of the WARAN instrument pool (Ohrnberger et al., 2006). Cross correlations and seismological data analysis were done using ObsPy, a Python framework for seismology (Beyreuther et al., 2010), and with the MIIC library (`https://github.com/miic-sw/miic`, last accessed 2015 October 1) that emerged from a project called Monitoring and Imaging based on Interferometric concepts. Some plots were made using the Generic Mapping Tools (GMT) version 4.5.6 (Wessel & Smith, 1998). Theoretical dispersion curves for phase velocity and γ and synthetic seismograms were computed with the code 'Computer Programs in Seismology', version 3.30 (Herrmann, 2013). The 1-D profile for station PRR was extracted from the Euroseistest website (`http://euroseisdb.civil.auth.gr/`, last accessed 2016 February 22). The first author of this publication was mainly funded by the Helmholtz graduate research school GeoSim. We are very grateful to the two anonymous reviewers who gave valuable suggestions and comments regarding the interpretation and discussion of our results which greatly improved the manuscript. The first author would also like to thank K. Hannemann and B. Bekeschus for helpful discussions and continuous support.

4 | Q_S from noise deconvolution in a borehole

Frequency dependent quality factors from the deconvolution of ambient noise recordings in a borehole in West Bohemia/Vogtland

Annabel Haendel[1], Matthias Ohrnberger[1], Frank Krüger[1]

[1] *Institute of Earth- and Environmental Science, University of Potsdam, Karl-Liebknecht Str. 24-25, D-14476 Potsdam*

Accepted for publication in:
Geophysical Journal International
Oktober 2018
DOI: 10.1093/gji/ggy422
Link: https://doi.org/10.1093/gji/ggy422

Abstract The correct estimation of site-specific attenuation is crucial for the assessment of seismic hazard. Downhole instruments provide in this context valuable information to constrain attenuation directly from data. In this study, we apply an interferometric approach to this problem by deconvolving seismic motions recorded at depth with those recorded at the surface. In doing so, incident and surface-reflected waves can be separated. We apply this technique not only to earthquake data but also to recordings of ambient vibrations. We compute the transfer function between incident and surface-reflected waves in order to infer frequency dependent quality factors for S-waves. The method is applied to a 87 m deep borehole sensor and a co-located surface instrument situated at a hard-rock site in West Bohemia/Vogtland, Germany. We show that the described method provides comparable attenuation estimates using either earthquake data or ambient noise for frequencies between 5-15 Hz. Moreover, a single hour of noise recordings seems to be sufficient to yield stable deconvolution traces and quality factors, thus, offering a fast and easy way to derive attenuation estimates from borehole recordings even in low to mid seismicity regions.

Keywords Seismic attenuation · Seismic interferometry · Seismic noise · Downhole methods

4.1 Introduction

Seismic waves undergo strong changes when propagating through the Earth and before reaching the surface. Of special importance are local site effects that are independent of the distance traveled from the source and can severely alter the appearance and frequency content of the seismic signal (Boore, 2003). The effects of local site geology are known to influence the signal significantly, for example, by basin effects, resonance effects or seismic wave attenuation. Attenuation is generally stronger close to the surface than in depth (e.g. Abercrombie, 1997) and acts as a low pass filter on the seismic signal. For site specific seismic hazard analysis there is often a lack of attenuation information at hard rock sites in particular for low seismicity regions (Ktenidou & Abrahamson, 2016). High risk facilities like nuclear power plants or dams are, however, usually constructed at hard-rock sites and it is thus especially important to assess the attenuation response under these conditions in order to compute the site-specific hazard.

Observations of seismic waves that are made both at the surface and within a co-located borehole can provide direct evidence for site specific seismic attenuation structure. Tonn (1991) gives an overview of several techniques to derive site attenuation from borehole records either in frequency or time domain. Taking the spectral ratio between borehole and surface is probably the most widely used approach (e.g. Abercrombie, 1997; Aster & Shearer, 1991; Bethmann et al., 2012). For shallow boreholes or at sites with fast wave velocities (hard rock sites) the spectral ratio method often fails due to the reflection of waves at the free surface. Incoming and surface-reflected waves overlay each other in the borehole recording leading to band-limited destructive interference in the amplitude spectrum which can complicate attenuation estimation (Shearer & Orcutt, 1987).

Several authors have therefore made use of an interferometric approach to separate incident and surface-reflected waves from earthquakes. The separation is achieved by deconvolving recordings obtained at depths with those obtained at the surface. Deconvolution interferometry (DCI) has been used to measure seismic velocities (Nakata & Snieder, 2012a; Parolai et al., 2009), to capture seismic velocity changes (Sawazaki et al., 2009), for imaging (Vasconcelos & Snieder, 2008), to study shear-wave splitting (Nakata & Snieder, 2012b), to constrain the input motion at the bottom of a borehole (Bindi et al., 2010) or to determine the response of a building (Bindi et al., 2015; Nakata & Snieder, 2014; Newton & Snieder, 2012; Snieder & Şafak, 2006).

There have been several efforts to derive the site-specific quality factor Q using DCI. Working in time domain, Trampert et al. (1993) derived attenuation factors for a 500 m deep borehole by inverting the SH propagator matrix. Following up on this work, Mehta et al. (2007) studied the same approach for the P-SV case. Raub et al. (2016) forward modeled the deconvolved wavefield in time domain to estimate seismic velocities and attenuation for P- and S-waves at the Tuzla vertical array in Turkey. In frequency domain, Parolai et al. (2010) fitted the Fourier transform of the deconvolved wavefield (the modulus and the acausal part) to a theoretical transfer function and estimated travel times and Q using a grid search procedure. Parolai et al. (2010) applied their method to the 140 m deep Ataköy vertical array with 4 downhole sensors. Following up on this work, Parolai et al. (2012) estimated attenuation by performing a full inversion of the spectrum. Fukushima et al. (2016) computed the transfer function of incident and surface-reflected wave in the deconvolved wavefield and derived frequency dependent quality factors for SH waves for several Kik-net stations in Japan. Finally, Snieder & Şafak (2006) and Newton & Snieder (2012) estimated Q for multiple-story buildings using earthquake records. Prieto et al. (2010), Nakata & Snieder (2014) and Bindi et al. (2015) applied DCI to ambient vibrations to study the Q retrieval in buildings.

Almost all studies listed above perform DCI using earthquake recordings or active sources to obtain attenuation information from a building or in a borehole. To our knowledge, there has been no study that adapts the deconvolution technique to borehole recordings of ambient vibrations for obtaining seismic attenuation information of the subsurface. Here, we deconvolve both ambient seismic noise and the signals of local earthquakes recorded at a pair of a 3-component borehole and a surface station to infer the inverse of the quality factor for S-waves (Q_s^{-1}). We use the method of Fukushima et al. (2016) to derive frequency dependent Q_s^{-1} from the transfer function of incident and surface-reflected waves in a deconvolved record. The data used were recorded at a 87 m deep borehole located at a hard-rock site in West Bohemia/Vogtland, Germany. We show that deconvolution of ambient vibrations provides equally good attenuation results as the deconvolution of earthquake recordings. Only a short duration of ambient vibration recordings is needed to obtain stable results. This is very promising because seismic noise is quasi-continuously available. The approach therefore may provide fast and reliable

site attenuation responses from borehole recordings even in mid to low seismicity regions.

4.2 Study Area

The study site is located in West Bohemia/Vogtland close to the village of Rohrbach at the Czech-German border. The borehole was originally planned for mineral water extraction by Bad Brambacher Mineralquellen GmbH & Co Betriebs KG (BBM) but was never used for production. With permission of BBM, a Lennartz borehole seismometer type LE-3D-BH (f_0=1 Hz, h=0.707) was placed at 87 m depth by the University of Potsdam and the GFZ German Research Centre for Geosciences in 2013. The borehole is drilled in a hard-rock site made up mainly from phyllite and mica schist that is weathered up to 40 m depth (personal communication with BBM). A Lennartz Electronic LE-3D-1s with the same instrument characteristics as the borehole sensor was installed at the surface next to the borehole. Data were digitized with Omnirecs data-cube^3 loggers. In this study we only analyze data that were recorded from June to August 2014. Data from both instruments were digitized with 400 Hz during this time period. The horizontal orientation of the borehole sensor has been derived from the data as described in Section 4.5.

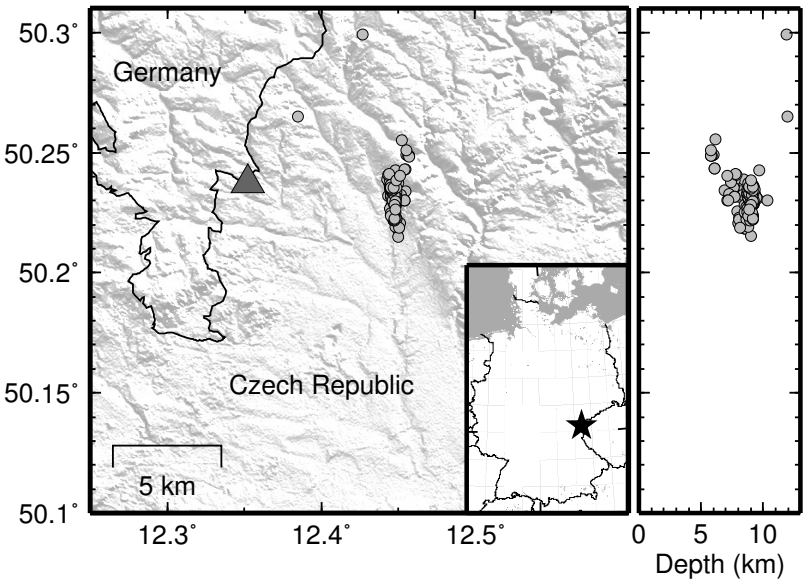

Figure 4.1: Map of the West Bohemia/Vogtland area at the Czech-German border. Shown is the position of the borehole (triangle). Locations of earthquakes from June to August 2014 with $0 \leq$ ML ≤ 1 are plotted as gray dots. A latitude-depth section of the earthquakes is shown on the right.

The West Bohemia/Vogtland region is well known for its repeating intra-continental earthquake swarm activity (Fischer et al., 2014). The seismic activity occurring in 2014 is unusual because it showed 3 typical mainshock-aftershock sequences triggered by M_L 3.5, 4.4 and 3.5 events on 24 May, 31 May and 3 August, respectively (Hainzl et al., 2016).

Fig. 4.1 shows the location of the borehole and of the earthquakes that were selected for deconvolution. We analyze both, ambient vibrations and earthquake recordings, in the following analysis.

4.3 Deconvolution interferometry

We use an interferometric approach that applies deconvolution analysis to either earthquake or seismic noise recordings to decompose the wavefield into up- and downgoing wavefields. DCI is preferred over the more commonly used cross-correlation interferometry (CCI). Newton & Snieder (2012) showed that CCI gives the correct phase but incorrect Fourier amplitudes which results in wrong attenuation estimates. DCI, on the other hand, allows for correct phase and amplitude estimation. Furthermore, as the source spectrum cancels in DCI it is generally preferable over CCI when measuring attenuation.

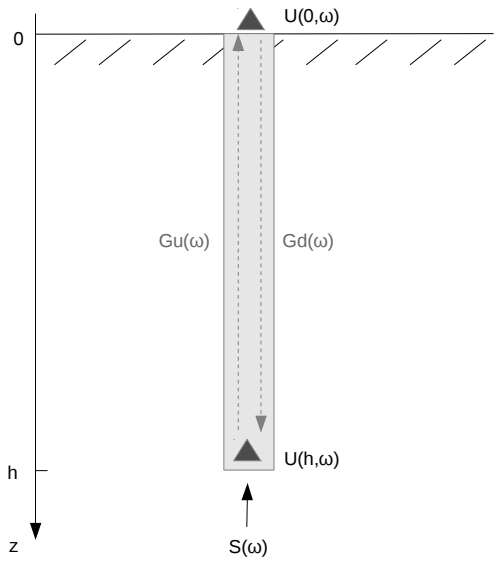

Figure 4.2: Motions recorded in a borehole due to a vertically incident plane wavefield $S(\omega)$ coming from below. $U(0,\omega)$ and $U(h,\omega)$ are the motions recorded at the surface and within the borehole at depth h. $G_u(\omega)$ and $G_d(\omega)$ denote Green's function response due to up- and downward propagating waves, respectively.

Figure 4.2 depicts the setup of the borehole configuration. A plane wavefield $S(\omega)$ is assumed to approach the borehole sensor with zero incidence angle. $S(\omega)$ includes the source function and all effects that occur between the source and the borehole instrument. The motion recorded by a sensor at the surface can be written in frequency domain as

$$U(0,\omega) = 2S(\omega)G_u(\omega). \tag{4.1}$$

$G_u(\omega)$ is the plane wave Green's function (propagation factor) between the sensor in the borehole and the one at the surface for a vertically upward propagating wave. The factor 2 accounts for the free surface effect.

The sensor at depth h records a superposition of the incoming wavefield $S(\omega)$ and the downward propagating wave that was reflected at the surface:

$$U(h,\omega) = S(\omega) + S(\omega)G_u(\omega)RG_d(\omega). \tag{4.2}$$

Here, $G_d(\omega)$ represents the Green's function between borehole and surface sensor for a downward propagating wave. R is the reflection coefficient observed at the free surface for a wavefield that is approaching the free surface from below. R=1 for vertical incidence.

Deconvolving the motion at depth with the motion at the surface yields

$$U(\omega) = \frac{U(h,\omega)}{U(0,\omega)} = \frac{S(\omega) + S(\omega)G_u(\omega)G_d(\omega)R}{2S(\omega)G_u(\omega)}. \tag{4.3}$$

$S(\omega)$ cancels and the resulting signal is the one that would be obtained for a source located at top of the borehole emitting a bandpass filtered delta impulse at $\tau = 0$:

$$\begin{aligned} U(\omega) &= \frac{1}{2G_u(\omega)} + \frac{G_d(\omega)}{2}R \\ &= U_1(\omega) + U_2(\omega). \end{aligned} \tag{4.4}$$

The first term $U_1(\omega)$ describes an incoming, upward propagating wavefield that can be observed at negative lag times (acausal arrival). The second term $U_2(\omega)$ is the surface-reflected phase arriving at positive lag times (causal arrival). The deconvolved wavefield is nonzero for negative times because the real incoming wavefield approaches the borehole from below. In order to generate a delta impulse at the top of the borehole at $\tau = 0$, a wave has to travel upward at negative times (Snieder, 2009).

$G_u(\omega)$ and $G_d(\omega)$ for an anelastic medium are given as (e.g. Aki & Richards, 2002)

$$G_{u,d}(\omega) = e^{-i\omega(h/c(\omega)-t)}e^{-\omega h/(2c(\omega)Q(\omega))} \tag{4.5}$$

where $\tau(\omega) = h/c(\omega)$ is the 1-way travel time between the bottom and the surface sensor. The first term in Eq. 4.5 describes the phase shift due to propagation with reference velocity $c(\omega)$. The second term gives the amplitude reduction due to the quality factor $Q(\omega)$. The two arrivals that are observed in the deconvolved wavefield according to Eq. 4.4 are thus separated in time by the 2-way travel time (2τ) between surface and borehole instruments. The amplitude difference of the pulses is related to the attenuation that the wave encounters on its path upwards from the borehole to the surface and back again. It is thus possible to derive the quality factor Q by comparing the amplitudes of the acausal and causal arrivals of the deconvolved trace.

The deconvolution process is due to spectral division potentially unstable. In order to prevent numerical instability a water-level stabilization (Clayton & Wiggins, 1976) is used:

$$U(\omega) = \frac{U(h,\omega)}{U(0,\omega)} \approx \frac{U(h,\omega)U(0,\omega)^*}{|U(0,\omega)|^2 + \epsilon < |U(0,\omega)|^2 >}. \tag{4.6}$$

The asterisk indicates complex conjugation. $< |U(0,\omega)|^2 >$ is the mean power spectrum of the surface record, and ϵ is a water-level constant that is chosen to keep the division stable. After testing, we set $\epsilon = 10^{-5}$ in our analysis.

4.4 Estimation of $Q^{-1}(f)$

We adopt the method of Fukushima et al. (2016) to estimate frequency dependent Q_s^{-1} from deconvolution. Fukushima et al. (2016) computed the transfer function between incident $U_1(\omega)$ (input) and surface-reflected wave $U_2(\omega)$ (output):

$$T(\omega) = \frac{U_2(\omega)}{U_1(\omega)} = \frac{U_2(\omega)U_1^*(\omega)}{|U_1(\omega)|^2} = \frac{U_{12}(\omega)}{U_{11}(\omega)}. \tag{4.7}$$

$U_1(\omega)$ and $U_2(\omega)$ correspond to the acausal and causal wave arrivals as given in Eq. 4.4. $U_{12}(\omega)$ is the cross-spectrum between the incident and the surface-reflected wave and $U_{11}(\omega)$ is the power spectrum of the incident wavefield.

Inserting Eqs 4.4 and 4.5 into Eq. 4.7 and taking only the amplitude part gives

$$|T(\omega)| = e^{-\omega\tau/Q(\omega)} \tag{4.8}$$

from which the frequency dependent Q^{-1} can be computed using

$$Q^{-1}(f) = -\frac{\ln|T(f)|}{2\pi f\tau}. \tag{4.9}$$

Q^{-1} is derived for different one hour long noise windows and for each earthquake record individually. In the following, we describe the processing steps for different noise windows but the procedure is similar for earthquakes. Following Fukushima et al. (2016), a mean Q^{-1} curve is calculated by taking the arithmetic

mean of the N number of logarithmized transfer functions that are obtained from different one hour long noise deconvolutions:

$$\langle \ln |T(f)| \rangle = \frac{1}{N} \sum_{i=1}^{N} \ln |T_i(f)|. \tag{4.10}$$

These curves will be denoted as log-mean curves within this publication.

The standard deviation of $\langle \ln |T(f)| \rangle$ is computed as

$$\sigma_{\ln T}(f) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\ln |T_i(f)| - \langle \ln |T(f)| \rangle)^2} \tag{4.11}$$

From Eqs 4.10 and 4.11 the mean and standard deviation of Q^{-1} can be estimated as

$$\langle Q^{-1}(f) \rangle \pm \sigma_{Q^{-1}}(f) = -\frac{(\langle \ln |T(f)| \rangle \mp \sigma_{\ln T}(f))}{2\pi f \langle \tau \rangle}. \tag{4.12}$$

with $\langle \tau \rangle = \frac{1}{N} \sum_{i=1}^{N} \tau_i$ being the mean travel time between the borehole and the surface sensor.

4.5 Data Analysis

In a first step, the correct horizontal orientation of the borehole instruments was estimated. The orientation was determined by analyzing 4 teleseismic earthquakes that were recorded at the borehole and the surface station. After the application of a 1 Hz lowpass filter the borehole traces were sequentially rotated until the best fit with the surface records was obtained. The borehole sensor turned out to be 54.0°±0.2° off the horizontal orientation off the surface sensor. In addition, the orientation of the surface instrument was verified with a Gyroscope. A horizontal error of 6.8° was detected. The horizontal components of the surface sensor were, thus, rotated by an azimuth of -6.8° while the borehole instrument was corrected by 47.2°.

No instrument correction was applied to the data because the borehole and the surface sensor have equivalent responses.

4.5.1 Intersensor travel-time differences

Travel times between the borehole and the surface sensor have to be known in order to compute the quality factor according to Eq. 4.9. For earthquake signals, intersensor travel times were estimated for each event by picking P- and S-wave arrivals within the borehole and at the surface. Intersensor travel times vary between 0.034 sec and 0.055 sec for P- and between 0.068 sec and 0.134 sec for S-waves. Computing the arithmetic mean of all values gives 0.043 sec and 0.085 sec for P- and S-waves, respectively. This leads to a v_p of around 2000 m/s (ranging between

1580 m/s to 2560 m/s) and a v_s of around 1000 m/s (ranging between 650 m/s to 1280 m/s) between borehole and surface sensor. The mean intersensor travel time $\langle\tau\rangle$ of S-waves is used for the computation of log-mean Q_s^{-1} curves (Eq. 4.12).

Seismic waves should reach the borehole instrument with vertical incidence in order for Eqs 4.1 to 4.4 to be fully valid. We randomly selected earthquakes from the dataset that took place at different distances and at different depths. The incidence angles that are deduced from the particle motion of the P-arrivals are 28-42° within the borehole and 14-25° at the surface.

For noise recordings, we estimated S-wave intersensor travel times directly from the mean noise deconvolution time traces in Fig. 4.3, left. The two arrivals that are observed in the deconvolved wavefields on the acausal and the causal part of the trace should be separated in time by 2τ. However, it is not immediately apparent where to set travel time picks in the deconvolutions. We simply picked the maxima of the acausal and the causal signal arrivals and find a S-wave travel time of 0.091 sec. A similar value is obtained when correcting the mean internensor travel time $\langle\tau\rangle$ obtained from the earthquake recordings to zero incidence angle. Dividing the $\langle\tau\rangle$ value by the cosine of a mean incidence angle between borehole and surface station of approximately 27° leads to 0.095 sec for S-waves. This indicates the presence of body waves in the incoming noise wavefield that impinge at the borehole sensor with almost zero incidence angle. Individual one hour long noise deconvolutions are very similar to the stacked noise deconvolution time series (see Fig. 4.5). We therefore use the same S-wave intersensor travel time of 0.095 sec for the computation of individual 1 hour and log-mean noise quality factors.

4.5.2 Ambient noise processing

We selected 9 full days of data (14-22 May 2014) for the computation of ambient noise deconvolution. These days lie before the start of the Bohemian earthquake activity on 24 May 2014. One hour long noise data were pre-processed by first removing the offset of each trace. Possible transient signals were eliminated by excluding signal windows from the data that have amplitudes larger than an hourly median amplitude level (Zhang & Yang, 2013). To do so, we computed for each 5 min long data window the envelope mean. The signal in this window was replaced with zeros if it exceeded two times the median envelope amplitude computed for the whole hour. The traces were then split into 100 sec long time windows with an overlap of 50 per cent (Welch's method, Seats et al., 2012; Welch, 1967). Windows were tapered at the edges with a 5 per cent cosine taper. The deconvolution was computed following Eq. 4.6. All deconvolved sequences originating from the same one hour long data window were stacked.

4.5.3 Earthquake data processing

We selected earthquakes from the 2014 relocated Webnet catalog that was used in Hainzl et al. (2016). Only events with $0 \leq$ ML ≤ 1 were chosen to avoid complex source time functions and non-linear site responses of the ground. We visually checked each event and removed records if, for example, earthquakes overlapped or if data quality was too bad. All earthquakes occurring in the time span from 24 May 2014 to 3 June 2014 had to be omitted due to datalogger malfunction of the surface station after a thunderstorm. In total, 194 earthquakes were selected. All earthquakes occurred approximately east of the borehole and close to each other (compare to Fig. 4.1) at distances between 5-12 km (average distance of 10.6 km) and at depths of 6-12 km (average depth of 8.8 km).

We took the whole 4 sec long earthquake signal (starting at the P-wave arrival time) for DCI calculation instead of selecting P- and S-wave windows. Mehta et al. (2007) and Parolai et al. (2009) showed that the deconvolutions are insensitive to the chosen signal window and only depend on the component of ground motion that is analyzed. P-wave energy is always observed on the deconvolved Z component and S-wave energy on the deconvolved horizontals irrespective of the chosen signal window. We tested this for all earthquakes and made the same observations. Deconvolution results and also Q^{-1} estimates are similar for selected P- and S-wave windows and for whole earthquake data segments.

The records were processed similar to noise recordings. First, signal offset was removed. The data windows were tapered at the edges with a 5 per cent cosine taper and the deconvolution was computed in frequency domain.

4.5.4 Deconvolution time series processing

The time-window length for cutting incident and surface-reflected waves from deconvolutions has to be chosen carefully. Very short time windows cut off important parts of the signal and lead to lower frequency resolution while windows that are too long include not only signal but also too much noise. We decided to use the same time-window length for noise and earthquakes in order to make the attenuation results better comparable. After testing, a time window length of 0.3 sec (starting or ending at zero lag time for causal and acausal wave arrivals, respectively) was chosen. Cut traces were tapered with a 5 per cent cosine taper before taking the Fourier transform. Q_s^{-1} was computed from the N and E components and the mean curves were estimated using Eqs 4.9 and 4.12.

4.6 Results

Fig. 4.3 shows the deconvolution results obtained from seismic noise and from earthquake records after stacking over all available one hour long noise deconvolutions or single earthquake deconvolutions, respectively. The Z components display P-wave energy traveling with faster velocity than the S-waves that are visible on

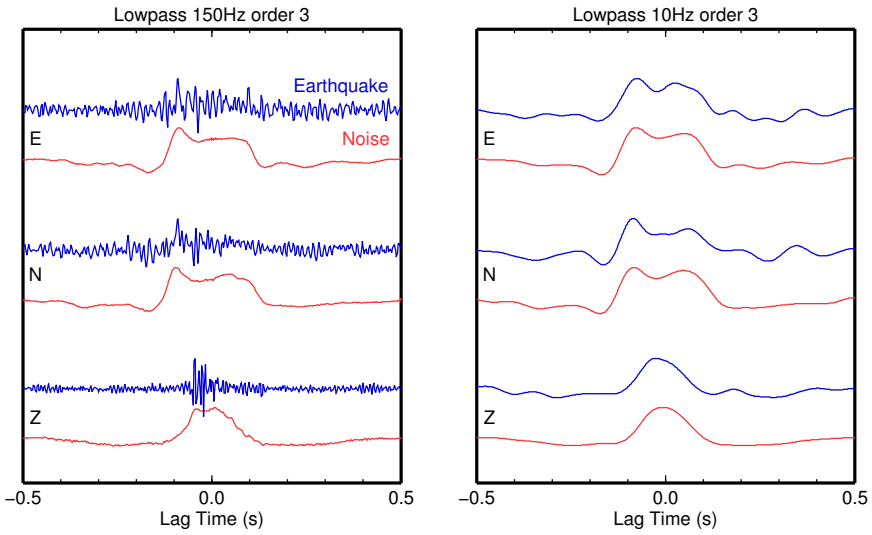

Figure 4.3: DCI results between borehole and surface sensor computed from ambient noise (red) and from earthquake records (blue) for all three components. Left: Deconvolved time traces filtered with a Butterworth lowpass filter of 150 Hz and order 3. Right: Traces filtered with a 10 Hz lowpass Butterworth filter of order 3. Note the similarity of noise and earthquake deconvolutions for low frequencies. The color version of this figure is available only in the electronic version.

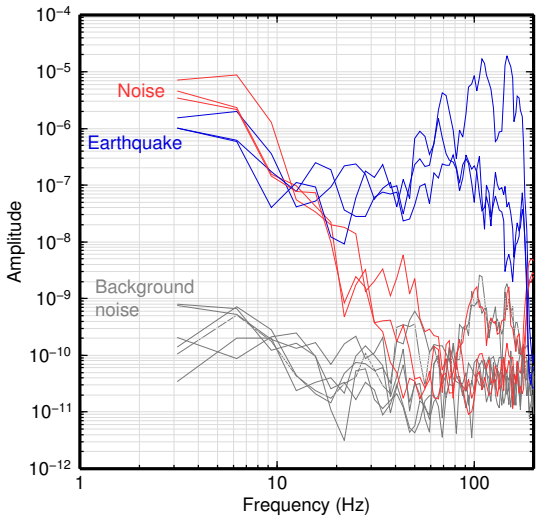

Figure 4.4: Fourier amplitude spectra of the deconvolution results shown in Fig. 4.3, left. Background noise spectra were taken from the first and last 1 sec of the in total ± 10 sec long noise deconvolution time series'. The color version of this figure is available only in the electronic version.

N and E. Deconvolution from earthquake records is richer in higher frequencies which is also visible in the deconvolution amplitude spectra of Fig. 4.4. Amplitudes of noise deconvolutions start to decrease above 6 Hz while the amplitude of earthquake deconvolutions remains high. Deconvolution results derived from noise and earthquakes are very similar for low frequencies as is shown in Fig. 4.3, right.

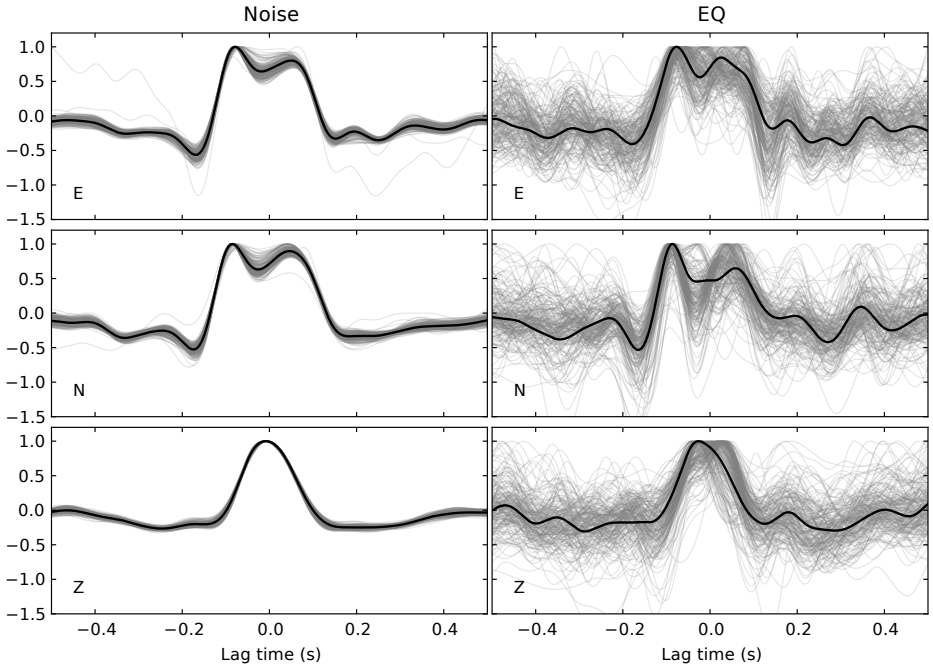

Figure 4.5: Deconvolution results for seismic noise (left) and earthquake recordings (right) for all three components of motion ZNE. Thin traces show the results of each of the one hour long noise segment or of each of the 194 events, respectively. Black lines are stacked results. Traces are filtered with an acausal Butterworth lowpass filter of 10 Hz and order 3. Each trace is normalized to its maximum.

Fig. 4.5 summarizes the lowpass filtered deconvolutions of each of the one hour long noise segments and each of the 194 earthquake records, respectively. The results are very similar for different noise windows. There are larger variations between individual event deconvolutions.

Up- and downgoing P-waves are badly separated in the deconvolution traces of the Z component. We therefore do not use the Z component in the further analysis and focus on the estimation of Q_s^{-1} from horizontal motions only. Fig. 4.6 shows the transfer functions computed with Eq. 4.7 from noise DCI (left) and earthquake DCI (right). Thin gray lines are transfer functions derived for different one hour long noise windows or different earthquakes, respectively. The thick solid and dashed blue lines represent the log-mean ± 1 standard deviation of all gray $T(f)$ curves computed according to Eqs 4.10 and 4.11. Individual one hour long noise $T(f)$ responses are fairly consistent below approximately 15 Hz. Above 15 Hz the

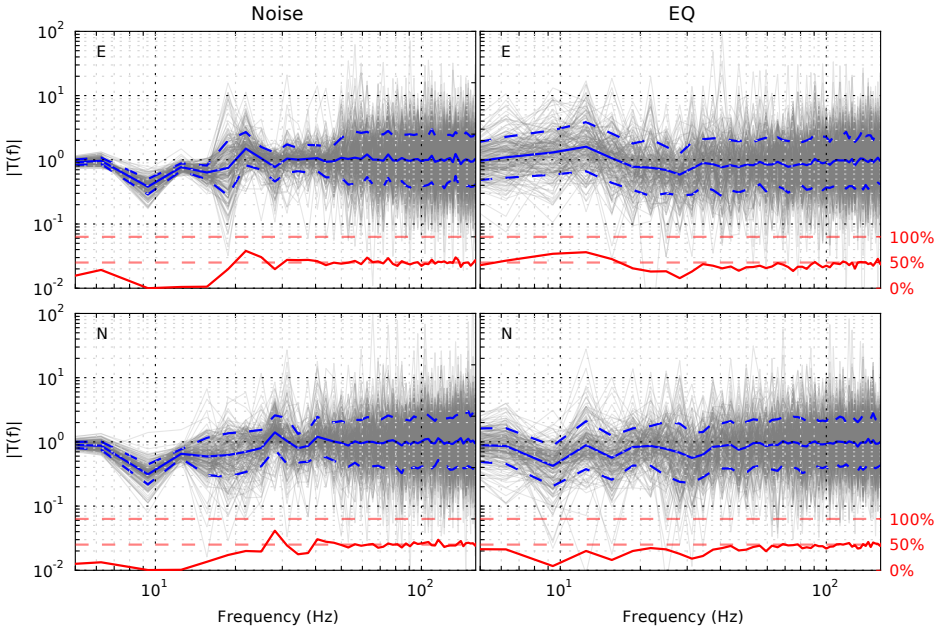

Figure 4.6: Transfer functions computed from noise DCI (left) and earthquake DCI (right). N components are displayed at the bottom and E at the top. Thin gray lines are transfer functions for individual one hour long noise windows or single earthquakes. The blue line is the logarithmic average of all transfer functions. Blue dashed lines show +/-1 standard deviation. Red lines of the bottom of each panel show the percentage of transfer functions with $T(f) > 1$. The color version of this figure is available only in the electronic version.

results show larger variations. This behavior is represented by small standard deviations of the log-mean curves below 15 Hz and larger standard deviations above this frequency. The differences between single earthquake $T(f)$ curves are generally high as is the standard deviation of the log-mean.

The red line at the bottom of Fig. 4.6 displays the percentage of data points with $T(f) > 1$. A value of $T(f) > 1$ will lead to a $Q_s^{-1} < 0$ and therefore to an amplitude increase instead of attenuation. Most of the noise transfer functions are smaller than one at frequencies below 15 Hz. The amount of curves with $T(f) > 1$ is around 50% at higher frequencies. On the contrary, many transfer functions derived from earthquakes have values larger than one irrespective of the chosen frequency range.

Fig. 4.7 shows $Q_s^{-1}(f)$. As in Fig. 4.6, thin gray lines are computed for individual one hour long noise windows or different earthquake recordings, respectively. The blue lines are computed from the log-mean $T(f)'s$. Because axes are plotted logarithmically curves are interrupted or not shown if $Q_s^{-1} < 0$.

The log-mean curves of Q_s^{-1} (all blue curves in Fig. 4.7) are summarized in Fig. 4.8 for the frequency range 5-20 Hz. Q_s^{-1} obtained from the E component of

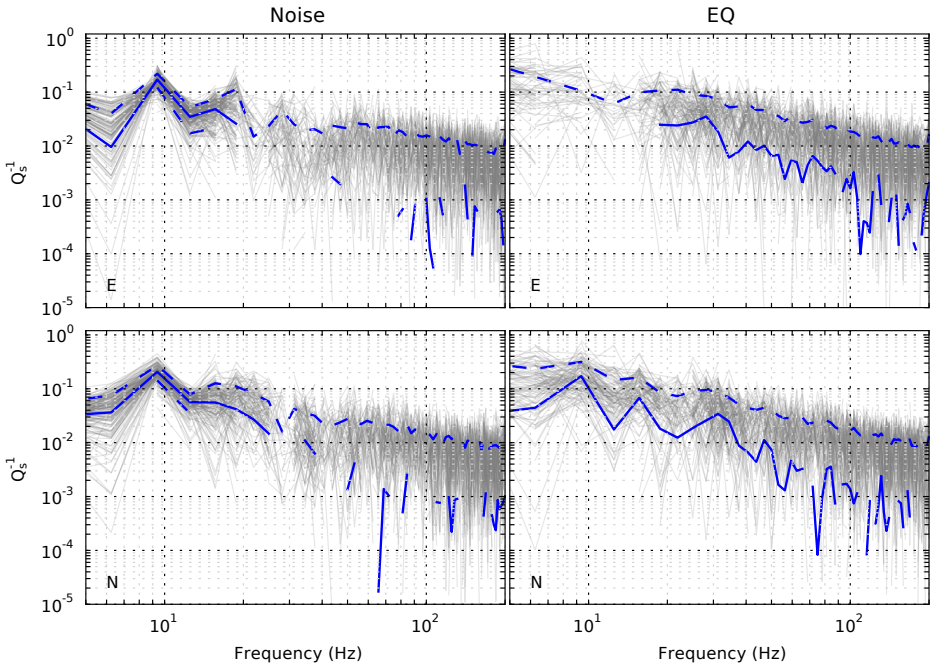

Figure 4.7: Frequency dependent Q_s^{-1} computed from noise DCI (left) and earthquake DCI (right). Q_s^{-1} derived from N and E components are shown at the bottom and at the top, respectively. Thin gray lines are Q_s^{-1} estimates calculated for individual one hour long noise windows or single earthquakes. The blue line is computed from the log-mean transfer function shown in Fig. 4.6. Blue dashed lines show +/-1 standard deviation. Curves are interrupted if $Q_s^{-1} < 0$. The color version of this figure is available only in the electronic version.

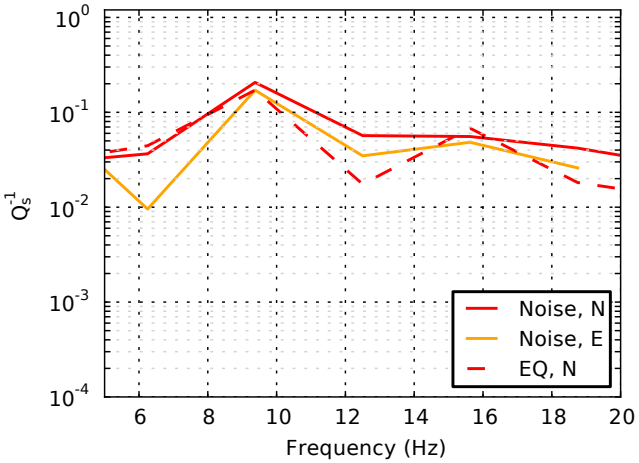

Figure 4.8: Comparison of Q_s^{-1} estimates (blue lines in Fig. 4.7) for frequencies between 5-20 Hz. Abbreviation: EQ=earthquake. The color version of this figure is available only in the electronic version.

earthquakes are negative within this frequency range and are therefore not shown. Noise and earthquake based Q_s^{-1} estimates are within the same order of magnitude.

4.7 Discussion

With DCI it is possible to observe the separation of the wavefield into incident and surface-reflected P- and S-waves in the West Bohemia/Vogtland area. This is not only possible with earthquake recordings but also with seismic noise. The stack in Fig. 4.3 was performed over 9 full days of noise data. However, deconvolution of only a single one hour long noise trace is sufficient to obtain a clear result (compare to Fig. 4.5). The situation is different for earthquake recordings. Earthquake deconvolutions have a higher frequency content than noise deconvolutions and the separation into incident and surface-reflected waves is not as clearly visible as for seismic noise without low pass filtering. Furthermore, up- and downgoing waves are only stable visible in the deconvolved earthquake traces when stacking over a sufficient number of events.

The transfer functions shown in Fig. 4.6 reveal a very similar pattern. The transfer function computed from a single one hour long noise deconvolution resembles the mean $T(f)$ of all noise segments very well at frequencies below 15 Hz. At higher frequencies, noise $T(f)$ curves are very different because the amplitude level of the noise deconvolution is almost at the same level as the amplitude of the background noise (Fig. 4.4). Transfer functions of earthquake recordings, on the contrast, vary strongly. This overall behavior of curves can also be observed for noise and earthquake derived Q_s^{-1} measures shown in Fig. 4.7.

The red percentage line in Fig. 4.6 can be regarded as a quality measure for the interpretation of log-mean transfer functions and Q_s^{-1} estimates. Mean Q_s^{-1} curves will be reliable for frequencies where only a small percentage of transfer functions is above one. This is the case for seismic noise below 15 Hz. Mean earthquake Q_s^{-1} measurements have to be taken with more care as many single event Q_s^{-1} estimates are below zero and therefore *drag* the log-mean curve to smaller Q_s^{-1} values. This is especially true for the E component below 15 Hz where the red line in Fig. 4.6 shows that more than 50 % of the curves have values of $T(f) > 1$. As a consequence, Q_s^{-1} cannot be recovered from the E component of earthquakes at low frequencies (compare to Figs 4.7 and 4.8).

One possible reason of the inferior performance of earthquake compared to noise recordings may be attributed to the non-zero incidence of earthquake waves. The assumption that the reflection coefficient at the free surface is one and that the free-surface factor is 2 is violated for waves that do not arrive steeply from below. This will affect the amplitudes of the causal and acausal wave arrivals in the deconvolved seismograms. In addition, SV-P conversions occur at non-zero incidence between the P- and the S-wave arrival and can further influence the amplitude of the downward reflected wave in the radial direction of motion. All earthquakes are located approximately east of the borehole so that the E component corresponds

roughly to the radial direction whereas the N component coincides with the transverse direction. SV-P conversions are thus expected to affect mainly the amplitudes of the reflection on the E component. This may be the reason why we are not able to recover Q_s^{-1} for the E component of earthquakes at frequencies below 15 Hz. Mean DCI results, $T(f)$ and Q_s^{-1} curves estimated from the N component of earthquakes are, however, very similar to noise derived curves.

Up- and downward propagating waves partly overlap in the deconvolution traces obtained from seismic noise or from earthquakes. We performed numerical simulations (see appendix) in order to evaluate the influence of the insufficient signal separation on the estimation of ${Q_s}^{-1}$. The results reveal that ${Q_s}^{-1}$ can also be recovered for shallow depths and for partially overlapping incident and surface-reflected waves in the deconvolution traces. Problems arise only at low frequencies were ${Q_s}^{-1}$ is overestimated. In addition, the scatter around the true value of ${Q_s}^{-1}$ is larger for shallower borehole recordings and improves as the sensor depth increases. The fluctuation of ${Q_s}^{-1}$ observed in Fig. 4.8 thus might be an effect of shallow borehole depth.

The deconvolution of noise recordings leads to very clear body wave signatures. Seismic noise is considered to be dominated by surface waves below 1 Hz while at higher frequencies the noise wavefield is suspected to be a mix of body waves and surface waves (Bonnefoy-Claudet et al., 2006). According to theory, the coherent noise sources contributing to the deconvolution have to be situated along the inter-station direction of the sensors, the so called end-fire lobes (Roux & Kuperman, 2004; Snieder, 2004). Sources outside the end-fire lobes interfere destructively if distributed homogeneously. Due to the strong impedance contrast at the free surface no significant amount of noise energy is assumed to enter the ground from above. This is confirmed by the observation that the deconvolved time series' are acausal and two-sided which is only the case if the incoming noise wavefield approaches the borehole from below (Snieder, 2009). Part of the seismic noise therefore needs to come from below the borehole station and cannot be explained by surface waves in the noise field. The amplitude spectra obtained from noise and earthquake deconvolutions as shown in Fig. 4.4 have a very different frequency content. Local events of very small or negative magnitude that may be present in the noise wavefield can therefore be ruled out as possible source of the body waves. One possible explanation for the presence of body waves in the noise wavefield might be scattering conversions at subsurface heterogeneities. Body waves could also be generated by local surface sources (e.g. close-by cities or roads) and arrive at the borehole sensor as diving waves from below. This was, for example, observed in a noise cross-correlation study of Hillers et al. (2012) at the TCDP borehole in Taiwan. Hillers et al. (2012) identified a cultural origin of body wave noise that follows the trajectory of a ballistic wave through the subsurface and enters at the borehole as coherent upward propagating body wave noise between 1-16 Hz. The origin of body wave noise in the West Bohemia/Vogtland may be a topic for future research.

The quality factors that are estimated within this study can be interpreted as effective ones including both, intrinsic and scattering attenuation. Eulenfeld & Wegler (2016) separated intrinsic and scattering attenuation of seismic shear waves by envelope inversion of local earthquakes in the Vogtland area. They found a scattering Q_s of 166-3000, an intrinsic Q_s of 100-2500 and a total Q_s of 67-1600 between 1-50 Hz. The Q_s values reported by Eulenfeld & Wegler (2016) were obtained for the whole travel path between source and site and are therefore higher than the near-surface values computed in this study. Nevertheless, their results reveal that scattering attenuation cannot be neglected in the Vogtland area. A superposition of intrinsic and scattering mechanism thus might explain the amplitude decay between incident and surface-reflected wave.

Q_s varies around a value of 20 between 5-20 Hz with a minimum value of 5 and a maximum value of 100. Typical Q values in the lithosphere are reported to be of the order of 10^2-10^3 in the frequency range 5-20 Hz (Sato et al., 2012, figure 5.1). Our results are, however, in conformance with attenuation measurements conducted at the KTB (Continental Deep Drilling Project) that is situated approximately 50 km to the southwest of the study site and that was drilled in a very similar crystalline environment down to a depth of about 9000 m. Li & Richwalski (1996) found Q_s to be between 8-25 at depth of 3-6 km and for frequencies between 11-22 Hz. Müller & Shapiro (2001) concluded that scattering attenuation plays an important role at the KTB site and might explain the low effective Q estimates. Several other borehole studies conducted at different rock sites of the world reported similarly high effective attenuations for depths shallower than 3 km (e.g. Abercrombie, 1997; Aster & Shearer, 1991, and references therein).

A comparison of our results with studies that employ the DCI method to earthquake recordings is only partly possible because these studies are usually conducted at sedimentary sites. Raub et al. (2016) obtain a Q_s of 20 between 0-288 m depth and at frequencies of 0.1-40 Hz in a limestone formation. Parolai et al. (2010) derive an average Q_s of about 30, 46 and 99 for the 0-50, 0-70, 0-140 m depth ranges, respectively, between 1-20 Hz for a site structure that is composed of limestone, clayey sand and sandstone layers.

The low Q_s estimates of our study are supported by slow P- and S-wave velocities that we observe between 0-87 m depth (compare to Section 4.5.1). The phyllites and mica schists are reported to be weathered up to at least 40 m depth (personal communication with BBM) which can certainly lead to low seismic velocities and high attenuation. As to our knowledge, there are no other studies in the area that investigate attenuation at similar depths and at similar frequencies as in our study.

4.8 Conclusion

We apply DCI to seismic noise recorded in a borehole configuration. We show that up- and downgoing waves are well observed in the deconvolved traces de-

spite the fact that noise recordings are usually assumed to be mainly composed of surface waves. Only short noise segments are necessary to obtain stable deconvolution results. DCI with ambient noise is thereby advantageous over DCI using earthquake recordings where several events are required to obtain a stable deconvolution. In addition, earthquake deconvolutions suffer from the non-zero incidence of earthquake waves so that Q_s^{-1} can only be recovered for the transverse component of motion due to P-SV conversion. On the contrary, Q_s^{-1} estimation from the deconvolution of seismic noise is limited to frequencies below 15 Hz in the West Bohemia/Vogtland area whereas earthquake deconvolutions are richer in higher frequencies. More research is needed to prove that DCI using ambient noise recorded in a borehole is also applicable to other areas or if the presented method is only valid if, for example, scattering attenuation is high. Furthermore, the origin of body waves in the noise wavefield needs to be investigated.

We are able to extract frequency dependent quality factors for S-waves between 5-15 Hz. The obtained quality factors are very small ($Q_s \approx 20$) but are in agreement with the results of other borehole studies conducted throughout the world. Available borehole logging information tells that the rock at the study location is weathered down to 40 m depth. This is probably the main reason for the low velocities and the high attenuation that we observe.

The presented method has the ability to provide an easy tool for the extraction of quality factors in the near surface if borehole recordings are available. This is especially the case for low- and mid-seismicity regions where earthquake recordings are sparse but where subsurface information is needed to assess, for example, the seismic hazard. The procedure may also be useful for continuous time-lapse monitoring of seismic velocity given the fact that a very short duration of ambient noise seems to suffice to obtain stable deconvolution results. Ambient noise interferometry thus provides a suitable alternative to earthquake-based borehole methods.

Acknowledgements The borehole is operated by Bad Brambacher Mineralquellen (BBM) GmbH & Co. Betriebs KG. We are very thankful to the BBM to make this borehole available for seismological studies. We owe further thanks to the GFZ German Research Centre for Geosciences Section 2.1 for logistical support and help during the field work. Deconvolution and seismological data analysis were done using ObsPy, a Python framework for seismology (Beyreuther et al., 2010), and with the MIIC library (`https://github.com/miic-sw/miic`, last accessed 1 October 2015) that emerged from a project called Monitoring and Imaging based on Interferometric concepts. Some plots were made using Generic Mapping Tools (GMT) version 4.5.6 (Wessel & Smith, 1998). We are grateful to Tomáš Fischer for providing the Webnet catalog with relocated events. QSEIS can be obtained from `https://www.gfz-potsdam.de`. The first author of this publication was partly funded by the section for equal opportunity of the Faculty of Science, University of Potsdam. She is furthermore a member of the Helmholtz graduate research school GeoSim. We are very grateful to the two anonymous reviewers who gave valuable suggestions and comments which led to an improvement of the manuscript.

5 | Discussion

In this thesis I presented two different approaches to tackle the question of GMPE selection and adjustment for regions of low to mid seismicity. In the following, I will summarize and discuss these two approaches separately and put them in the larger context of ground-motion prediction and seismic hazard assessment.

5.1 Mixture model approach

Chapter 2 introduces the mixture model concept where GMPEs coming from different regions are scored and merged in order to provide an optimized ensemble model. The aggregated model is able to deliver predictions in areas where no region-specific GMPE exists due to the scarcity of data. We tested the performance of the mixture for Northern Chile in terms of average residuals and average sample log-likelihoods on an independent dataset. We show that the mixture model outperforms all single GMPEs used for its construction. It is also comparable to a regression model specifically derived for that area.

Roselli et al. (2016) proposes an alternative method using a Bayesian information criterion (BIC) approximation to weight and combine GMPEs in an ensemble approach. They use BIC instead of the log-likelihood because part of the data that are taken to weight the models may have been also used in the construction of some of the GMPEs. Roselli et al. (2016) applied the approach to the Italian territory and to five selected GMPEs. They evaluated the performance of the uncombined and the merged GMPEs in BIC and they also found that the performance of the aggregated model is superior to those of single GMPEs. This is in conformance with our own findings. Yet, it is not clear whether the testing dataset that was used to evaluate the performance of the models in Roselli et al. (2016) is independent from the data that were employed to derive the weights in the ensemble approach. If the independence is not guaranteed it is not surprising that the aggregated model outperforms the other GMPEs that were derived from at least partially different datasets. Nevertheless, Roselli et al. (2016) argue that the choice of the scheme according to which GMPEs are merged (e.g. BIC or log-likelihood) may not be

of primary importance and that the ensemble approach will almost always give superior results compared to uncombined GMPEs.

A profuse number of new data has become available in Northern Chile since the publication of the mixture study in 2014. For example, on the 1st of April 2014 the Mw 8.2 Iquique earthquake ruptured parts of the Northern Chile seismic gap and triggered several large aftershocks with Mw>6. Two of the largest aftershocks occurred two days after the mainshock with Mw 6.5 and Mw 7.6. These events provide an important dataset to test the predictive power of the mixture model especially for large magnitudes for which data are less abundant. We extracted 45 records from the Iquique mainshock and the two largest aftershocks and evaluated the performance of the mixture, the regression model and the component GMPEs for this dataset (Fig. 5.1). The mixture model is not the best performing model but still among the best models to predict the new data. This is the case although the mixture model was mainly calibrated for events with Mw<6.5. All recordings with Mw>6.5 come from the catalog of Arango et al. (2011) and from earthquakes that occurred south of the study area. Incorporating these new data in the development of the mixture could improve its predictive power for future large events. Additional large magnitude events occurring in other parts of the Chile subduction zone like the Mw 8.3 Illapel in 2015 could possibly further augment the database in the high magnitude range.

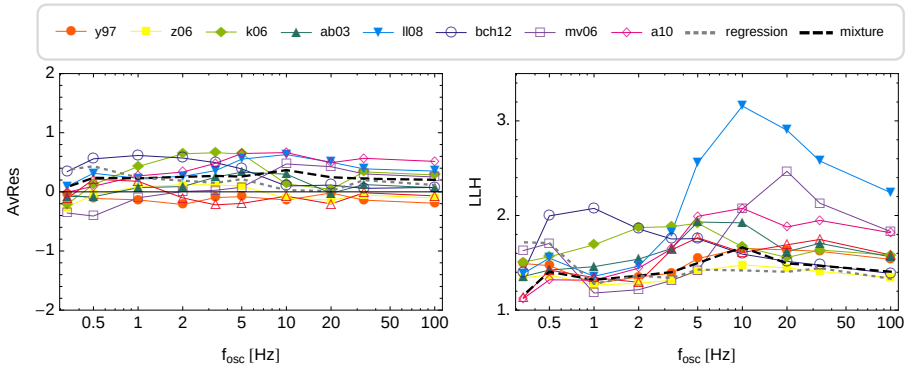

Figure 5.1: The performance of the mixture, the regression model and the individual GMPEs for the Iquique mainshock and the two largest aftershocks. The performance is assessed in terms of average residuals (left) and LLH values (right). Compare to Tab. 2.1 for GMPE abbreviations.

New GMPEs were developed for the Chilean subduction zone during the last years. Idini et al. (2017) and Montalva et al. (2017) derived regression models for interface and intraslab earthquakes for the distance range of around 30-400 km and for magnitudes up to Mw 7.8 (intraslab) and Mw 8.8 (interface). Both models predict spectral acceleration for oscillator periods between 0.01-10 Hz. It would be important to compare the mixture model to these local models in order to see if it is similarly able to predict ground-motions in Chile. However, it should always be kept in mind that the mixture is intended to give predictions in situations where

only little data are available. We used a rather large dataset for the calibration of the mixture (1094 records from 138 events) but data availability will certainly be different for other areas. Nevertheless, even if data amount is insufficient, a mixture model can still be derived by combining observed data with experts prior knowledge that can be incorporated into the mixture approach using a Bayesian framework.

Runge et al. (2013a) experimented with prior distributions that represent various expert beliefs regarding the appropriateness of different GMPEs in the formation of mixture models. Using Bayes theorem, these prior distributions are subsequently updated using observed data leading to the a posteriori distributions of the mixture weights. Runge et al. (2013a) applied their approach to Northern Chile and to the same set of data and GMPEs that was used in the study presented in Chapter 2 of this thesis. They found that expert beliefs and observed data can be combined but that it strongly matters how subjective expert knowledge is elicited and transferred into prior distributions (see also Runge et al., 2013b). The method proposed by Runge et al. (2013a) could be used to not only include expert beliefs in the construction of the mixture model but to subsequently update the mixture as new data become available with time.

The aim in setting up a ground-motion logic tree is to capture the epistemic uncertainty. If branch weights are interpreted as probabilities the Kolmogoroff axioms have to be followed. This implies that the set of models that occupies the branches of the tree must be mutually exclusive and collectively exhaustive (e.g. Bommer & Scherbaum, 2008). As mentioned in the introduction of this manuscript and as shown in Fig. 1.1 there are two main approaches to populate the branches of a ground-motion logic tree (Atkinson et al., 2014). In the first one, multiple GMPEs are used and weights are assigned to each model based on the judgment of experts in combination with data-driven selection procedures (Kale & Akkar, 2013; Scherbaum et al., 2009). Yet, using multiple GMPEs comes along with some issues. Many models that are used in a logic-tree analysis are derived from similar datasets (like the Next Generation (NGA)-West 2 GMPEs, Bozorgnia et al., 2014) or share a similar functional form. Models can then not be regarded as mutually exclusive. Secondly, the chosen set of models may not correctly represent the underlaying epistemic uncertainty of median ground motions (Atkinson et al., 2014). In fact, many models are too close in their predictions which leads to an underestimation of epistemic uncertainty (Douglas & Edwards, 2016). The set of GMPEs can then not be regarded as collectively exhaustive.

These difficulties can partially be overcome by using a single or a small set of GMPEs (backbone GMPEs) that represent the center of the ground-motion distribution in the area of interest (Atkinson et al., 2014). Adjustments applied to these GMPEs (e.g. using HEM) lead to alternative models that account for the upper and lower limits of the ground-motion distribution. Populating the logic tree with a representative suit of models and their scaled versions is assumed to better capture the underlaying epistemic uncertainty. In the case of a single backbone GMPE, the

branches of the tree become mutually exclusive and, if scaled properly, collectively exhaustive (Atkinson et al., 2014).

It is in the hands of experts to decide which GMPEs should be considered as backbone models and how they should be modified and scaled to capture the center, body and range of expected ground motions at the site of interest. The mixture model approach could provide in this context a valuable data-driven tool to find an optimal median backbone model. Mixture weights are inferred in such a way as to improve the ensemble prediction of the models. Thereby, GMPEs may be favored (in terms of high mixture weights) even if they perform bad for that area (in terms of average residuals and average sample log-likelihoods). On the contrary, it is possible that a good performing model receives zero mixture weights because information is not counted twice in the mixture model approach. If two models are very similar and would contribute the same informational content into the ensemble of models then only one GMPE obtains high mixture weights (or the weight is split between the two models). Thereby, mutually exclusiveness is guaranteed. A single central model like the mixture may thus be favorable over a set of backbone GMPEs that is not necessarily mutually exclusive. If data amount is sufficient, a mixture model captures the center of the ground-motions distribution in the target area sufficiently well. It may be easier for experts to assign scaling factors to this model (in order to define the upper and lower limits of the expected ground-motion) instead of scaling several GMPEs. Finally, a single central model which is scaled up and down results in a minimum number of branches and makes hazard calculations very efficient.

The epistemic uncertainty can be captured, as described above, by applying appropriate scaling factors to the mixture model. The standard deviation (sigma) of GMPEs describes the aleatory variability that is associated with ground-motion generation. The standard deviation of a mixture is, in contrast to uncombined GMPEs, not a good representative of the aleatory variability because the mixture is not necessarily normally distributed. If models that are combined into the mixture and that obtain high mixture weights, have, for example, distributions that do not overlap then the mixture distribution will be multimodal. In addition, epistemic uncertainty may map into the standard deviation of the aggregated model due to the combination of GMPEs. Several studies (e.g. Anderson & Brune, 1999) have suggested that sigma of GMPEs should not be interpreted as aleatory variability anyway. Some components of sigma at a single site may be repeatable path and site effects and should not be considered as random (e.g. Al Atik et al., 2010). These components need to be removed from the aleatory description of uncertainty and transferred to the epistemic uncertainty captured in the logic tree. The remaining variability of strong motion at a specific site is referred to as single-station sigma (e.g. Atkinson, 2006; Rodriguez-Marek et al., 2014) which is usually smaller than the total sigma of GMPEs. The mixture model approach can thus be used in combination with a sigma logic tree as described in Rodriguez-Marek et al. (2014). Ktenidou et al. (2017a) proposed a method that estimates a single-station sigma from existing GMPEs rather than from a newly created local GMPE. This is of

special importance for areas with insufficient data amount and therefore also for studies where a mixture model approach may be considered due to the lack of data.

In recent years, numerical simulations of earthquake shaking (physic-based models) have become popular (e.g. Beresnev & Atkinson, 1997; Graves & Pitarka, 2010; Hartzell et al., 1999). Yet, numerical simulations require a full understanding of the physical system so that GMPEs will stay important for some time (Douglas & Edwards, 2016). However, when considering, for example, near-source ground motions of large earthquakes where data coverage is poor or in cases where the hazard is dominated by a certain fault, physical-based simulations can augment GMPE predictions (e.g. Atkinson & Goda, 2011). The mixture model concept could provide in this context a suitable framework to combine simulated and empirically derived model predictions. Another possibility would be to use physical-based predictions as 'observed' data in the construction of the mixture. Available data seldom cover all situations of practical interest which is especially the case in low-seismicity regions. Simulations could be conducted for all those scenarios where data are missing giving rise to a more complete picture of possible ground motions in the area of interest. Using such an 'augmented' dataset for the derivation of mixture weights could enhance the predictive power of the mixture model for possible future scenarios.

5.2 Estimation of κ_0 from seismic noise

The second part of the thesis is concerned with the derivation of the near-surface attenuation parameter κ_0 from ambient vibrations. I presented two preliminary studies towards this goal using interferometric approaches. In the study shown in Chapter 3, we estimated Love wave quality factors Q_L between 1-4 Hz close to the center and at the edge of a sedimentary basin from higher order noise correlations. Chapter 4 presents the derivation of Q_s between 5-15 Hz from the deconvolution of noise recorded simultaneously at a surface and at a borehole station in a hard rock site environment.

The initial goal of this thesis was to derive κ_0 directly from noise measurements. Yet, it proved to be very difficult to infer this parameter from diffuse wavefields. Seismic interferometry provides an excellent tool to derive virtual earthquake signals in the absence of real ones (e.g. Denolle et al., 2014). These can subsequently be used to investigate the subsurface structure of the Earth using traditional imaging techniques and, hence, also to estimate attenuation. With regard to κ_0 there are, however, some issues. κ_0 is associated with the attenuation of S-waves propagating vertically upwards through the geological structure directly below a site. The following factors hamper the successful extraction of κ_0 from seismic noise:

- Seismic noise is usually generated at Earth's surface by oceanic or meteorological processes or by human activity (Bonnefoy-Claudet et al., 2006). According to Forghani & Snieder (2010) noise sources can be distributed

anywhere along the receiver line for the retrieval of surface waves from cross correlations. Yet for the reconstruction of body waves, sources must be present also at depth below the receivers. In addition, for large receiver separations the coherent propagating noise energy is generally dominated by surface waves due to smaller geometrical loss even if body waves were excited (Roux et al., 2005a). As a result, Green's functions reconstructed from the cross correlation of seismic noise often only contain the surface wave component of the full impulse response (Shapiro et al., 2005).

- A large portion of the noise source energy is generated within the microseism band (5-20 sec, e.g. Friedrich et al., 1998) that travels coherently over long distances. Considerably less coherent noise energy is observed at large receiver separations and at higher frequencies where κ_0 is typically estimated (> 10 Hz). Interferometric experiments are therefore usually conducted at frequencies below 1 Hz.

- The estimation of subsurface parameters from ambient noise depends on the chosen pre-processing techniques and on the noise source distribution. If travel times and velocities should be extracted, only the phase information of the cross correlation is required and constraints on the pre-processing are less strict. This is not the case for the retrieval of attenuation parameters where the amplitude information of the cross correlation needs to be retained during pre-processing (e.g Cupillard & Capdeville, 2010; Tsai, 2011; Weaver, 2011).

I will discuss each of these points in the following and give possible solutions that might facilitate the extraction of κ_0 from ambient noise in the future.

With regard to the retrieval of body waves using seismic interferometry there has been some progress in recent years. Roux et al. (2005a) were the first to prove that P-waves can be extracted from noise cross correlations for closely-spaced receiver pairs and at frequencies between 0.1-1.3 Hz. Draganov et al. (2009) successfully obtained P-wave reflection images by correlating noise recorded along several receiver lines (distances smaller than 20 km) in Lybia between 6-24 Hz. Shirzad & Shomali (2015) recovered clear P- and S-wave arrivals at distances smaller than 35 km and at frequencies between 0.3-3 Hz using noise interferometry in combination with a rms-stacking method. Nakata et al. (2015) isolated P-waves between 3-15 Hz in a dense receiver network (8x4 km^2) at Long Beach, California by applying additional signal processing techniques to the cross-correlated signals. Mosher & Audet (2017) extracted P-waves at a network of borehole seismometers in Parkfield, California between 1-2 Hz. On a regional scale, Zhan et al. (2010) and Poli et al. (2012) revealed the retrieval of Moho reflections (SmS, PmP) from noise recordings. Using teleseismic interferometry, Boué et al. (2013), Lin et al. (2013) and Nishida (2013) could extract global body wave phases traveling through deep parts of the Earth like the inner and outer core.

Local interferometric body wave studies are especially important for the purpose of this thesis as they are conducted at frequencies similar to those that are usually applied to extract κ_0 from earthquake recordings. The studies listed above show that body waves can be extracted from seismic noise if, for example, additional signal processing techniques are applied (e.g. Nakata et al., 2015; Shirzad & Shomali, 2015). Takagi et al. (2014) proposed a new procedure to effectively separate body and Rayleigh waves using the cross terms of the cross-correlation tensor. They derived their concept for isotropic and homogeneous media but note that it might also be applicable to more realistic structures. Using this method, Takagi et al. (2014) successfully separated body and Rayleigh waves at the Tono array in Japan in the frequency range 0.5-2 Hz and at distances between 2.4-18 km. Procedures, as the ones of Takagi et al. (2014), may facilitate the retrieval of body waves from noise cross correlations in the future. However, the applicability of such methods with regard to amplitude extraction and, hence, attenuation parameter retrieval still needs to be appraised.

According to Bonnefoy-Claudet et al. (2006) seismic noise is dominated by surface waves below 1 Hz but has been observed to be a mix of body and surface waves at higher frequencies. Koper et al. (2010) analyzed seismic noise recorded at 18 arrays around the globe between 0.4-4 Hz. They found that the noise wavefield is dominated by Love waves within this frequency range but that, on average, 28% of the energy comes from P-waves. Yet, the origin of body waves in the ambient seismic field is still under debate, especially at frequencies above 1 Hz. Roux et al. (2005a) argue that a fraction of the dominant Rayleigh wave energy in the noise field might be locally converted into body waves by scattering at heterogeneities in the upper crust. Gerstoft et al. (2008) and Zhang et al. (2009) show that P-waves can be generated in the ocean by distant storms and can be observed up to frequencies of 2 Hz. Hillers et al. (2012) identified a cultural origin of body wave noise at the TCDP borehole in Taiwan that follows the trajectory of a ballistic wave through the subsurface. It is obvious that no consensus is reached upon the origin and strength of body waves in high-frequency seismic noise until now. It can also be assumed that body wave noise strongly varies from area to area depending on the available noise sources and the local geological structure.

If body waves are present in the ambient seismic field, it is easier to extract them by applying interferometry to vertically deployed stations instead of surface sensors. According to theory, a seismic wave can only be extracted from the cross correlation of noise between two stations if there is a physically wave in the noise field that propagates between these receivers (e.g. Snieder et al., 2009). In the context of body wave retrieval from surface sensors this implies that body wave noise must be incident at just the right angle in order to obtain body waves that are reflected at one receiver and then recorded at the other receiver (Snieder & Larose, 2013). All other incident waves will not contribute to the extraction of body waves. For downhole stations the constrains are less stringent. Diving body waves in the noise field coming from different directions and distances can contribute to the extraction of body waves between borehole receivers. This is probably the reason

why we do observe clear body waves in Chapter 4 of this thesis when applying deconvolution interferometry to a borehole and a surface sensor in the Vogtland. However, the origin of body waves in the noise field of the Vogtland is not yet clear. It is possible that the deconvolution approach applied to noise recorded in boreholes might not work for other areas if body wave noise is not sufficiently generated. The only other study that reports the extraction of body waves from the cross correlation of noise recorded between several downhole sensors is the one of Hillers et al. (2012). Hillers et al. (2012) observed coherent upward propagating body wave noise at the TCDP borehole in Taiwan at frequencies between 1-16 Hz. Beamforming and polarization analysis of noise correlations reveal that these waves are probably generated in the populated lowlands of western Taiwan that arrive at the vertical array as diving waves from below. Further studies need to be conducted in different areas in order to verify the finding that the deconvolution of noise recordings at vertically deployed sensors leads to the extraction of body waves.

We do not observe body waves in Chapter 3 of this thesis when applying higher-order noise correlations to surface stations. Yet, procedures as the one of Takagi et al. (2014) might enhance the extraction of body waves that are usually hidden in the surface wave signal of cross correlations given that body waves are present in the noise field and satisfy the origin constrains given by Snieder & Larose (2013).

As mentioned above, care has to be taken with regard to certain pre-processing techniques when extracting attenuation from seismic noise. For example, Cupillard & Capdeville (2010) and Tsai (2011) investigated numerically and theoretically how correlation amplitudes depend on the noise source distribution in combination with typical non-linear pre-processing operations like one-bit normalization and spectral whitening (Bensen et al., 2007). They observed that, if such pre-processing techniques are applied, attenuation parameters can only be retrieved accurately if the noise distribution is homogeneous. This is hardly the case in the real world. We overcame this difficulty by using higher-order noise correlations in Chapter 3 of this thesis. Higher-order correlations are less sensitive to the source anisotropy of the original ambient field so that unbiased attenuation estimates can be extracted (Zhang & Yang, 2013). In addition, we refrained from using pre-processing techniques like one-bit normalization in our analysis to ensure that the true amplitude information is retained. In Chapter 4, we used a deconvolution approach in combination with very basic pre-processing operations like offset removal and transient signal removal by muting. Changes in the amplitude content of the deconvolved signals due to nonlinear pre-processing procedures should therefore be insignificant in this case.

This thesis presents the estimation of quality factors for Love and S-waves within the shallow subsurface from ambient noise recordings. Using the relationship between κ_0 and Q_s that is presented in equation 1.2 it is easily possible to compare Q_s estimates obtained, for example, with deconvolution interferometry in a borehole with κ_0 values derived from earthquakes. Unfortunately, there are no

earthquake-based κ_0 values available for the Vogtland area for comparison. Even if earthquake-based κ_0 values were available, there is a multitude of measurement techniques for κ_0 (Ktenidou et al., 2014). To complicate the matter even more, there is until now no consensus upon the depth range and the physical mechanisms causing the high-frequency fall-off of the spectrum that leads to the observation of κ_0 (Ktenidou et al., 2014). κ_0 is mainly assumed to be an intrinsic (frequency independent) site attenuation parameter because of the linear decay of the acceleration spectrum in log-linear space. Recent studies (Ktenidou et al., 2015; Parolai et al., 2015; Pilz & Fäh, 2017) have provided evidence that κ_0 cannot be described by intrinsic attenuation alone but that scattering effects also contribute to the observed spectral fall-off. If attenuation estimation in low-seismicity regions is based on laboratory tests of small soil samples, which capture only the intrinsic material damping, then κ_0 is typically underestimated (Cabas et al., 2017). Noise-based κ_0 measurements may provide in this context a valuable source of information as they also include the scattering component of attenuation. In low-seismicity regions, noise measurements are then favorable over laboratory tests if no earthquake based estimation of κ_0 is possible.

Finally, recent site-specific hazard studies typically compute the ground-motion for a reference rock condition (e.g. Douglas & Edwards, 2016; Edwards et al., 2015). Site amplification and attenuation factors representative for the soil column above the rock are applied to these predictions to adapt the ground-motion to the specific site. It is necessary to give accurate κ_0 values for the rock reference site which is often critical due to the limited number of recordings on rock. This issue could be solved if κ_0 could be inferred from seismic noise.

6 | Conclusion

This thesis addresses the issue of GMPE selection and adjustment in low- to mid-seismicity regions. I presented two concepts to facilitate the adaption of foreign GMPEs if strong-motion recordings are too sparse to develop a region-specific model.

In the first approach, several non-indigenous GMPEs are combined in a mixture model using observed data from the target area. The mixture is aimed to serve as a backbone model for the region under study for use in a logic-tree analysis. Populating the logic tree with such an optimized median backbone model and its scaled versions can help to make the branches of the tree mutually exclusive and collectively exhaustive and, hence, to better capture the epistemic uncertainty of the GMPE selection process. The aleatory variability of ground-motion generation can be accounted for by using the mixture approach in combination with a single-station sigma logic tree as described in Rodriguez-Marek et al. (2014). The mixture model concept is very flexible in that it allows to incorporate new data as soon as they become available by using a Bayesian framework. In addition, it is possible to not only combine the predictions of empirically derived models but also from stochastic or numerical simulations. This is of special value if data coverage is poor in certain magnitude-distance ranges or if a site-specific hazard evaluation should be performed.

In the second approach, I explored techniques to extract the near-surface attenuation parameter κ_0 from ambient vibrations. I presented two studies where the quality factors of Love and S-waves could be derived from noise recordings. The direct estimation of κ_0 from seismic noise is very challenging. Borehole recordings will certainly provide the most easiest source of information to study the extraction of κ_0 from ambient noise as the constrains for body wave retrieval are less strict. Yet, the sources of body waves in seismic noise needs further investigation. Boreholes, as the 200 m deep borehole in the sedimentary basin of the Euroseistest area or the ICDP project in the hard rock environment of the Vogtland area, may be suitable candidates to compute κ_0 from noise and, if successful, to compare it with κ_0 values derived from earthquakes. It is a longer way to go to extract κ_0 from surface sensors only. Body waves at high frequencies are hardly observed in

cross-correlation traces. Advanced signal processing procedures as the one of Takagi et al. (2014) may facilitate the separation of body and surface waves in noise correlation in the future and therefore the derivation of κ_0.

The selection and adjustment of GMPEs in low- to mid-seismicity regions remains a challenge. Luckily, the number of strong-motion data is increasing fast so that it will be easier to develop region-specific GMPEs or site-specific physical models in the future. Yet, data of large magnitude events recorded at close distances will continue to be lacking in certain areas if the corresponding faults have not been active during the recording history of seismology. It will then still be necessary to resort to the information provided by foreign GMPEs or to physical simulations which need to be combined with local information.

The vastly emerging field of seismic interferometry has the potential to further aid the adjustment or development of GMPEs. An increasing number of studies deals with the derivation of the regional or local velocity and attenuation structure from ambient seismic noise. This additional source of information can be used to tune, for example, physically-based models or to derive parameters like κ_0. The studies shown in this thesis are only the first steps into this directions. Newly developed procedures will certainly enhance the extraction of GMPE adjustment parameters in the future.

Bibliography

Abercrombie, R., 1997. Near-surface attenuation and site effects from comparison of surface and deep borehole recordings, *Bulletin of the Seismological Society of America*, **87**(3), 731–744.

Abrahamson, N. A. & Silva, W. J., 1997. Empirical response spectral attenuation relations for shallow crustal earthquakes, *Seismological Research Letters*, **68**(1), 94–127.

Abrahamson, N. A. & Youngs, R. R., 1992. A stable algorithm for regression analyses using the random effects model, *Bulletin of the Seismological Society of America*, **82**(1), 505–510.

Aki, K., 1957. Space and time spectra of stationary stochastic waves, with special reference to microtremors, *Bulletin of the Earthquake Research Institute Tokyo University*, **25**, 415–457.

Aki, K. & Richards, P. G., 2002. *Quantitative Seismology*, University Science Books, Sausalito, California, US, 2nd edn.

Akkar, S. & Bommer, J. J., 2006. Influence of long-period filter cut-off on elastic spectral displacements, *Earthquake Engineering & Structural Dynamics*, **35**(9), 1145–1165.

Akkar, S., Kale, O., Yenier, E., & Bommer, J. J., 2011. The high-frequency limit of usable response spectral ordinates from filtered analogue and digital strong-motion accelerograms, *Earthquake Engineering & Structural Dynamics*, **40**, 1387–1401.

Al Atik, L., Abrahamson, N., Bommer, J. J., Scherbaum, F., Cotton, F., & Kuehn, N., 2010. The variability of ground-motion prediction models and its components, *Seismological Research Letters*, **81**(5), 794–801.

Al Atik, L., Kottke, A., Abrahamson, N., & Hollenback, J., 2014. Kappa (κ) scaling of ground-motion prediction equations using an inverse random vibration

theory approach, *Bulletin of the Seismological Society of America*, **104**(1), 336–346.

Anderson, D. L., Ben-Menahem, A., & Archambeau, C. B., 1965. Attenuation of seismic energy in the upper mantle, *Journal of Geophysical Research*, **70**(6), 1441–1448.

Anderson, J. & Hough, S., 1984. A model for the shape of the Fourier amplitude spectrum of acceleration at high frequencies, *Bulletin of the Seismological Society of America*, **74**(5), 1969–1993.

Anderson, J. G. & Brune, J. N., 1999. Probabilistic seismic hazard analysis without the ergodic assumption, *Seismological Research Letters*, **70**(1), 19–28.

Arango, M. C., Strasser, F. O., Bommer, J. J., Boroschek, R., Comte, D., & Tavera, H., 2011. A strong-motion database from the Peru-Chile subduction zone, *Journal of Seismology*, **15**(1), 19–41.

Arroyo, D., García, D., Ordaz, M., Mora, M. A., & Singh, S. K., 2010. Strong ground-motion relations for Mexican interplate earthquakes, *Journal of Seismology*, **14**(4), 769–785.

Arroyo, D., Ordaz, M., & Rueda, R., 2014. On the selection of ground-motion prediction equations for probabilistic seismic-hazard analysis, *Bulletin of the Seismological Society of America*, **104**(4), 1860–1875.

Aster, R. C. & Shearer, P. M., 1991. High-frequency borehole seismograms recorded in the San Jacinto fault zone, southern California. Part 2. Attenuation and site effects, *Bulletin of the Seismological Society of America*, **81**(4), 1081–1100.

Atkinson, G. M., 2006. Single-station sigma, *Bulletin of the Seismological Society of America*, **96**(2), 446–455.

Atkinson, G. M., 2008. Ground-motion prediction equations for eastern north america from a referenced empirical approach: implications for epistemic uncertainty, *Bulletin of the Seismological Society of America*, **98**(3), 1304–1318.

Atkinson, G. M. & Boore, D. M., 2003. Empirical ground-motion relations for subduction zone earthquakes and their application to Cascadia and other regions, *Bulletin of the Seismological Society of America*, **93**(4), 1703–1729.

Atkinson, G. M. & Goda, K., 2011. Effects of seismicity models and new ground-motion prediction equations on seismic hazard assessment for four canadian cities, *Bulletin of the Seismological Society of America*, **101**(1), 176–189.

Atkinson, G. M., Bommer, J. J., & Abrahamson, N. A., 2014. Alternative approaches to modeling epistemic uncertainty in ground motions in probabilistic seismic-hazard analysis, *Seismological Research Letters*, **85**(6), 1141–1144.

Azarbakht, A., Rahpeyma, S., & Mousavi, M., 2014. A new methodology for assessment of the stability of ground-motion prediction equations, *Bulletin of the Seismological Society of America*, **104**(3), 1447–1457.

BC Hydro, 2012. Probabilistic seismic hazard analysis (PSHA) model, *Engineering Report E658*, **vol. 1-4**, November 2012.

Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., Shapiro, N., & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophysical Journal International*, **169**, 1239–1260.

Beresnev, I. A. & Atkinson, G. M., 1997. Modeling finite-fault radiation from the ω^n spectrum, *Bulletin of the Seismological Society of America*, **87**(1), 67–84.

Bethmann, F., Deichmann, N., & Mai, P. M., 2012. Seismic wave attenuation from borehole and surface records in the top 2.5 km beneath the city of Basel, Switzerland, *Geophysical Journal International*, **190**(2), 1257–1270.

Beyer, K. & Bommer, J. J., 2006. Relationships between median values and between aleatory variabilities for different definitions of the horizontal component of motion, *Bulletin of the Seismological Society of America*, **96**(4A), 1512–1522.

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J., 2010. ObsPy: A Python Toolbox for Seismology, *Seismological Research Letters*, **81**(3), 530–533.

Bilham, R., 2009. The seismic future of cities, *Bulletin of Earthquake Engineering*, **7**(4), 839–887.

Bindi, D., Parolai, S., Picozzi, M., & Ansal, A., 2010. Seismic input motion determined from a surface-downhole pair of sensors: A constrained deconvolution approach, *Bulletin of the Seismological Society of America*, **100**(3), 1375–1380.

Bindi, D., Petrovic, B., Karapetrou, S., Manakou, M., Boxberger, T., Raptakis, D., Pitilakis, K. D., & Parolai, S., 2015. Seismic response of an 8-story RC-building from ambient vibration analysis, *Bulletin of Earthquake Engineering*, **13**(7), 2095–2120.

Biro, Y. & Renault, P., 2012. Importance and impact of host-to-target conversions for ground motion prediction equations in PSHA, in *Proc. of the 15th World Conference on Earthquake Engineering*, pp. 24–28, Lisbon, Portugal, September 24-28, 2012.

Blaser, L., Krüger, F., Ohrnberger, M., & Scherbaum, F., 2010. Scaling relations of earthquake source parameter estimates with special focus on subduction environment, *Bulletin of the Seismological Society of America*, **100**(6), 2914–2926.

Bommer, J. J. & Scherbaum, F., 2008. The Use and Misuse of Logic Trees in Probabilistic Seismic Hazard Analysis, *Earthquake Spectra*, **24**(4), 997–1009.

Bommer, J. J., Scherbaum, F., Bungum, H., Cotton, F., Sabetta, F., & Abrahamson, N. A., 2005. On the use of logic trees for ground-motion prediction equations in seismic-hazard analysis, *Bulletin of the Seismological Society of America*, **95**(2), 377–389.

Bommer, J. J., Douglas, J., Scherbaum, F., Cotton, F., Bungum, H., & Fah, D., 2010. On the selection of ground-motion prediction equations for seismic hazard analysis, *Seismological Research Letters*, **81**(5), 783–793.

Bommer, J. J., Coppersmith, K. J., Coppersmith, R. T., Hanson, K. L., Mangongolo, A., Neveling, J., Rathje, E., Rodriguez-Marek, A., Scherbaum, F., Shelembe, R., Stafford, P., & Strasser, F., 2013. A SSHAC Level 3 probabilistic seismic hazard analysis for a new-build nuclear site in South Africa, *Earthquake Spectra*, **31**(2), 661–698.

Bonnefoy-Claudet, S., Cotton, F., & Bard, P. Y., 2006. The nature of noise wavefield and its applications for site effects studies. A literature review, *Earth-Science Reviews*, **79**(3-4), 205–227.

Boore, D. M., 1983. Stochastic simulation of high-frequency ground motions based on seismological models of the radiated spectra, *Bulletin of the Seismological Society of America*, **73**(6A), 1865–1894.

Boore, D. M., 2003. Simulation of Ground Motion Using the Stochastic Method, *Pure and Applied Geophysics*, **160**, 635–676.

Boore, D. M., 2005. On pads and filters: Processing strong-motion data, *Bulletin of the Seismological Society of America*, **95**(2), 745–750.

Boore, D. M. & Bommer, J. J., 2005. Processing of strong-motion accelerograms: needs, options and consequences, *Soil Dynamics and Earthquake Engineering*, **25**(2), 93–115.

Bora, S. S., Scherbaum, F., Kuehn, N., & Stafford, P., 2014. Fourier spectral-and duration models for the generation of response spectra adjustable to different source-, propagation-, and site conditions, *Bulletin of Earthquake Engineering*, **12**(1), 467–493.

Bora, S. S., Scherbaum, F., Kuehn, N., & Stafford, P., 2016. On the relationship between Fourier and response spectra: Implications for the adjustment of empirical ground-motion prediction equations (GMPEs), *Bulletin of the Seismological Society of America*, **106**(3), 1235–1253.

Boué, P., Poli, P., Campillo, M., Pedersen, H., Briand, X., & Roux, P., 2013. Teleseismic correlations of ambient seismic noise for deep global imaging of the earth, *Geophysical Journal International*, **194**(2), 844–848.

Bozorgnia, Y., Abrahamson, N. A., Atik, L. A., Ancheta, T. D., Atkinson, G. M., Baker, J. W., Baltay, A., Boore, D. M., Campbell, K. W., Chiou, B. S.-J., et al., 2014. NGA-West2 research project, *Earthquake Spectra*, **30**(3), 973–987.

Bragato, P. & Slejko, D., 2005. Empirical ground-motion attenuation relations for the eastern Alps in the magnitude range 2.5–6.3, *Bulletin of the Seismological Society of America*, **95**(1), 252–276.

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

Cabas, A., Rodriguez-Marek, A., & Bonilla, L. F., 2017. Estimation of site-specific kappa (κ_0)-consistent damping values at KiK-net sites to assess the discrepancy between laboratory-based damping models and observed attenuation (of seismic waves) in the field, *Bulletin of the Seismological Society of America*, **107**(5), 2258–2271.

Campbell, K., 2009. Estimates of shear-wave Q and κ_0 for unconsolidated and semiconsolidated sediments in Eastern North America, *Bulletin of the Seismological Society of America*, **99**(4), 2365–2392.

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

Campbell, K. W., 2003b. Strong-motion attentuation relations, *International Geophysics Series*, **81**(B), 1003–1012.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**(5606), 547–549.

Cartwright, D. & Longuet-Higgins, M. S., 1956. The statistical distribution of the maxima of a random function, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 237, pp. 212–232, The Royal Society.

Claerbout, J. F., 1968. Synthesis of a layered medium from its acoustic transmission response, *Geophysics*, **33**(2), 264–269.

Clayton, R. & Wiggins, R., 1976. Source shape estimation and devoncolution of teleseismic body waves, *Geophysical Journal of the Royal Astronomical Society*, (47), 151–177.

Colombi, A., Chaput, J., Brenguier, F., Hillers, G., Roux, P., & Campillo, M., 2014. On the temporal stability of the coda of ambient noise correlations, *Comptes Rendus Geoscience*, **346**(11-12), 307–316.

Comte, D. & Pardo, M., 1991. Reappraisal of great historical earthquakes in the Northern Chile and Southern Peru seismic gaps, *Natural Hazards*, **4**, 23–44.

Contreras, V. & Boroschek, R., 2012. Strong ground motion attenuation relations for Chilean subduction zone interface earthquakes, in *Proceedings of the 15th World Conference on Earthquake Engineering (15WCEE)*, Lisbon, Portugal, September 24-28, 2012.

Converse, A. M. & Brady, A. G., 1992. BAP : Basic strong-motion accelerogram processing software; Version 1.0, *Open-File Report 92-296A*, U.S. Geological Survey.

Cornell, C. A., 1968. Engineering seismic risk analysis, *Bulletin of the seismological society of America*, **58**(5), 1583–1606.

Cotton, F., Scherbaum, F., Bommer, J. J., & Bungum, H., 2006. Criteria for selecting and adjusting ground-motion models for specific target regions: Application to Central Europe and rock sites, *Journal of Seismology*, **10**(2), 137–156.

Cranswick, E., 1988. The information content of high-frequency seismograms and the near-surface geologic structure of hard rock recording sites, in *Scattering and Attenuations of Seismic Waves, Part I*, vol. 128, pp. 333–363, eds Aki, K. & Wu, R.-S., Birkhäuser Basel.

Cupillard, P. & Capdeville, Y., 2010. On the amplitude of surface waves obtained by noise correlation and the capability to recover the attenuation: a numerical approach, *Geophysical Journal International*, **181**(3), 1687–1700.

Cupillard, P., Stehly, L., & Romanowicz, B., 2011. The one-bit noise correlation: a theory based on the concepts of coherent and incoherent noise, *Geophysical Journal International*, **184**(3), 1397–1414.

Delavaud, E., Cotton, F., Akkar, S., Scherbaum, F., Danciu, L., Beauval, C., Drouet, S., Douglas, J., Basili, R., Sandikkaya, M., Segou, M., Faccioli, E., & Theodoulidis, N., 2012. Toward a ground-motion logic tree for probabilistic seismic hazard assessment in Europe, *Journal of Seismology*, **16**(3), 451–473.

Delouis, B., Cisternas, A., Dorbath, L., Rivera, L., & Kausel, E., 1996. The Andean subduction zone between 22 and 25°S (northern Chile): precise geometry and state of stress, *Tectonophysics*, **259**, 81–100.

Dempster, A. P., Laird, N. M., & Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, **39**(1), 1–38.

Denolle, M., Dunham, E., Prieto, G., & Beroza, G., 2014. Strong ground motion prediction using virtual earthquakes, *Science*, **343**(6169), 399–403.

Douglas, J., 2011. Ground-motion prediction equations 1964-2010, *PEER Report No. 2011/102*, Pacific Earthquake Engineering Research Center, Berkeley, CA.

Douglas, J., 2017. *Ground motion prediction equations 1964-2017*, `http://www.gmpe.org.uk`, Last accessed January 31, 2018.

Douglas, J. & Boore, D. M., 2011. High-frequency filtering of strong-motion records, *Bulletin of Earthquake Engineering*, **9**(2), 395–409.

Douglas, J. & Edwards, B., 2016. Recent and future developments in earthquake ground motion estimation, *Earth-Science Reviews*, **160**, 203–219.

Douglas, J., Cotton, F., Di Alessandro, C., Boore, D. M., Abrahamson, N. A., & Akkar, S., 2012. Compilation and critical review of GMPEs for the GEM-PEER Global GMPEs Project, in *Proceedings of the 15th World Conference on Earthquake Engineering (15WCEE)*, Lisbon, Portugal, September 24-28, 2012.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A., & Wapenaar, K., 2009. Reflection images from ambient seismic noise, *Geophysics*, **74**(5), A63–A67.

Edwards, B., Ktenidou, O.-J., Cotton, F., Abrahamson, N., Van Houtte, C., & Fäh, D., 2015. Epistemic uncertainty and limitations of the κ_0 model for near-surface attenuation at hard rock sites, *Geophysical Journal International*, **202**(3), 1627–1645.

Ekström, G., Abers, G. A., & Webb, S. C., 2009. Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, *Geophysical Research Letters*, **36**(18), 5–9.

Eulenfeld, T. & Wegler, U., 2016. Measurement of intrinsic and scattering attenuation of shear waves in two sedimentary basins and comparison to crystalline sites in Germany, *Geophysical Journal International*, **205**(2), 744–757.

Fischer, T., Horálek, J., Hrubcová, P., Vavryčuk, V., Bräuer, K., & Kämpf, H., 2014. Intra-continental earthquake swarms in West-Bohemia and Vogtland: a review, *Tectonophysics*, **611**, 1–27.

Forghani, F. & Snieder, R., 2010. Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry, *The Leading Edge*, **29**(7), 790–794.

Friedrich, A., Krueger, F., & Klinge, K., 1998. Ocean-generated microseismic noise located with the Gräfenberg array, *Journal of Seismology*, **2**(1), 47–64.

Froment, B., Campillo, M., & Roux, P., 2011. Reconstructing the Green's function through iteration of correlations, *Comptes Rendus Geoscience*, **343**(8-9), 623–632.

Frühwirth-Schnatter, S., 2006. *Finite mixture and Markov switching models*, Springer Series in Statistics, Springer Science & Business Media, New York.

Fukushima, R., Nakahara, H., & Nishimura, T., 2016. Estimating S-wave attenuation in sediments by deconvolution analysis of KiK-net borehole seismograms, *Bulletin of the Seismological Society of America*, **106**(2), 552–559.

García, D., Singh, S. K., Herráiz, M., Ordaz, M., & Pacheco, J. F., 2005. Inslab earthquakes of Central Mexico: Peak ground-motion parameters and response spectra, *Bulletin of the Seismological Society of America*, **95**(6), 2272–2282.

Gerstoft, P., Shearer, P. M., Harmon, N., & Zhang, J., 2008. Global P, PP, and PKP wave microseisms observed from distant storms, *Geophysical Research Letters*, **35**(23).

Gouédard, P., Cornou, C., & Roux, P., 2008. Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique, *Geophysical Journal International*, **172**(3), 971–981.

Graves, R. W. & Pitarka, A., 2010. Broadband ground-motion simulation using a hybrid approach, *Bulletin of the Seismological Society of America*, **100**(5A), 2095–2123.

Hainzl, S., Fischer, T., Čermáková, H., Bachura, M., & Vlček, J., 2016. Aftershocks triggered by fluid intrusion: Evidence for the aftershock sequence occurred 2014 in West Bohemia/Vogtland, *Journal of Geophysical Research: Solid Earth*, **121**(4), 2575–2590.

Hanks, T., 1982. fmax, *Bulletin of the Seismological Society of America*, **72**(6), 1867–1879.

Hanks, T. C., Abrahamson, N. A., Boore, D. M., Coppersmith, K. J., & Knepprath, N. E., 2009. Implementation of the SSHAC guidelines for Level 3 and 4 PSHAs—experience gained from actual applications, *US Geological Survey Open-File Report*, **1093**, 66 p.

Hannemann, K., Papazachos, C., Ohrnberger, M., Savvaidis, A., Anthymidis, M., & Lontsi, A. M., 2014. Three-dimensional shallow structure from high-frequency ambient noise tomography: New results for the Mygdonia basin-Euroseistest area , northern Greece, *Journal of Geophysical Research*, **119**(6), 4979–4999.

Hartzell, S., Harmsen, S., Frankel, A., & Larsen, S., 1999. Calculation of broadband time histories of ground motion: Comparison of methods and validation using strong-ground motion from the 1994 Northridge earthquake, *Bulletin of the Seismological Society of America*, **89**(6), 1484–1504.

Herrmann, R. B., 2013. Computer programs in seismology: an evolving tool for instruction and research, *Seismological Research Letters*, **84**(6), 1081–1088.

Hillers, G., Campillo, M., Lin, Y.-Y., Ma, K.-F., & Roux, P., 2012. Anatomy of the high-frequency ambient seismic wave field at the TCDP borehole, *Journal of Geophysical Research: Solid Earth*, **117**(B6).

Hough, S., 1997. Empirical Green's function analysis: Taking the next step, *Journal of Geophysical Research*, **102**(B3), 5369–5384.

Hough, S. & Anderson, J., 1988. High-frequency spectra observed at Anza, California: Implications for Q structure, *Bulletin of the Seismological Society of America*, **78**(2), 692–707.

Idini, B., Rojas, F., Ruiz, S., & Pastén, C., 2017. Ground motion prediction equations for the Chilean subduction zone, *Bulletin of Earthquake Engineering*, **15**(5), 1853–1880.

Jongmans, D., Pitilakis, K., Demanet, D., Raptakis, D., Riepl, J., Horrent, C., Tsokas, G., Lontzetidis, K., & Bard, P.-Y., 1998. EURO-SEISTEST: determination of the geological structure of the Volvi basin and validation of the basin response, *Bulletin of the Seismological Society of America*, **88**(2), 473–487.

Kale, O. & Akkar, S., 2013. A new procedure for selecting and ranking ground-motion prediction equations (GMPEs): The Euclidean distance-based ranking (EDR) method, *Bulletin of the Seismological Society of America*, **103**(2A), 1069–1084.

Kanno, T., Narita, A., Morikawa, N., Fujiwara, H., & Fukushima, Y., 2006. A new attenuation relation for strong ground motion in Japan based on recorded data, *Bulletin of the Seismological Society of America*, **96**(3), 879–897.

Koper, K. D., Seats, K., & Benz, H., 2010. On the composition of Earth's short-period seismic noise field, *Bulletin of the Seismological Society of America*, **100**(2), 606–617.

Ktenidou, O. & Abrahamson, N. A., 2016. Empirical estimation of high-frequency ground motion on hard rock, *Seismological Research Letters*, **87**(6), 1465–1478.

Ktenidou, O.-J., Cotton, F., Abrahamson, N. A., & Anderson, J. G., 2014. Taxonomy of κ: A review of definitions and estimation approaches targeted to applications, *Seismological Research Letters*, **85**(1), 135–146.

Ktenidou, O.-J., Abrahamson, N. A., Drouet, S., & Cotton, F., 2015. Understanding the physics of kappa (κ): Insights from a downhole array, *Geophysical Journal International*, **203**(1), 678–691.

Ktenidou, O.-J., Roumelioti, Z., Abrahamson, N., Cotton, F., Pitilakis, K., & Hollender, F., 2017a. Understanding single-station ground motion variability and uncertainty (sigma): lessons learnt from euroseistest, *Bulletin of Earthquake Engineering*, pp. 1–26.

Ktenidou, O.-J., Silva, W. J., Darragh, R. B., Abrahamson, N. A., & Kishida, T., 2017b. Squeezing kappa (κ) out of the transportable array: A strategy for using bandlimited data in regions of sparse seismicity, *Bulletin of the Seismological Society of America*, **107**(1), 256–275.

Langston, C. A., Bodin, P., Powell, C., Withers, M., Horton, S., & Mooney, W., 2005. Bulk sediment Qp and Qs in the Mississippi embayment, central United States, *Bulletin of the Seismological Society of America*, **95**(6), 2162–2179.

Laurendeau, A., Cotton, F., Ktenidou, O.-J., Bonilla, L.-F., & Hollender, F., 2013. Rock and stiff-soil site amplification: Dependency on v_{s30} and kappa (κ_0), *Bulletin of the Seismological Society of America*, **103**(6), 3131–3148.

Lawrence, J. F. & Prieto, G. A., 2011. Attenuation tomography of the western United States from ambient seismic noise, *Journal of Geophysical Research*, **116**(B6), 1–11.

Lawrence, J. F., Denolle, M., Seats, K. J., & Prieto, G. A., 2013. A numeric evaluation of attenuation from ambient noise correlation functions, *Journal of Geophysical Research*, **118**(12), 6134–6145.

Li, X.-P. & Richwalski, S., 1996. Seismic attenuation and velocities of P- and S-waves in the German KTB area, *Journal of Applied Geophysics*, **36**(2-3), 67–76.

Lin, F., Ritzwoller, M., & Shen, W., 2011. On the reliability of attenuation measurements from ambient noise cross-correlations, *Geophysical Research Letters*, **38**, L11303.

Lin, F.-C., Tsai, V. C., Schmandt, B., Duputel, Z., & Zhan, Z., 2013. Extracting seismic core phases with array interferometry, *Geophysical Research Letters*, **40**(6), 1049–1053.

Lin, P.-S. & Lee, C.-T., 2008. Ground-motion attenuation relationships for subduction-zone earthquakes in Northeastern Taiwan, *Bulletin of the Seismological Society of America*, **98**(1), 220–240.

Liu, X., Ben-zion, Y., & Zigone, D., 2015. Extracting seismic attenuation coefficients from cross-correlations of ambient noise at linear triplets of stations, *Geophysical Journal International*, **203**(2), 1149–1163.

Lobkis, O. & Weaver, R., 2001. On the emergence of the Green's function in the correlations of a diffuse field, *J. acoust. Soc. Am.*, **110**(6), 3011–3017.

Mak, S., Clements, R. A., & Schorlemmer, D., 2017. Empirical evaluation of hierarchical ground-motion models: Score uncertainty and model weighting, *Bulletin of the Seismological Society of America*, **107**(2), 949–965.

McVerry, G. H., Zhao, J. X., Abrahamson, N. A., & Somerville, P. G., 2006. New Zealand acceleration response spectrum attenuation relations for crustal and subduction zone earthquakes, *Bulletin of the New Zealand Society for Earthquake Engineering*, **39**(1), 1–58.

Mehta, K., Snieder, R., & Graizer, V., 2007. Extraction of near-surface properties for a lossy layered medium using the propagator matrix, *Geophysical Journal International*, **169**(1), 271–280.

Molkenthin, C., Scherbaum, F., Griewank, A., Leovey, H., Kucherenko, S., & Cotton, F., 2017. Derivative-based global sensitivity analysis: Upper bounding of sensitivities in seismic-hazard assessment using automatic differentiation, *Bulletin of the Seismological Society of America*, **107**(2), 984–1004.

Montalva, G. A., Bastías, N., & Rodriguez-Marek, A., 2017. Ground-motion prediction equation for the chilean subduction zone, *Bulletin of the Seismological Society of America*, **107**(2), 901–911.

Mosher, S. G. & Audet, P., 2017. Recovery of P Waves from ambient-noise interferometry of borehole seismic data around the San Andreas fault in Central California, *Bulletin of the Seismological Society of America*, **108**(1), 51–65.

Müller, T. & Shapiro, S., 2001. Seismic scattering attenuation estimates for the German KTB Area derived from well-log statistics, *Geophysical Research Letters*, **28**(19), 3761–3764.

Müller, G., 1985. Source pulse enhancement by deconvolution of an empirical Green's function, *Geophysical Research Letters*, **12**(1), 33–36.

Nakata, N. & Snieder, R., 2012a. Estimating near-surface shear wave velocities in Japan by applying seismic interferometry to KiK-net data, *Journal of Geophysical Research*, **117**, B01308.

Nakata, N. & Snieder, R., 2012b. Time-lapse change in anisotropy in Japan's near surface after the 2011 Tohoku-Oki earthquake, *Geophysical Research Letters*, **39**(11), 1–6.

Nakata, N. & Snieder, R., 2014. Monitoring a building using deconvolution interferometry. II: Ambient-vibration analysis, *Bulletin of the Seismological Society of America*, **104**(1), 204–213.

Nakata, N., Chang, J. P., Lawrence, J. F., & Boué, P., 2015. Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry, *Journal of Geophysical Research: Solid Earth*, **120**(2), 1159–1173.

National Earthquake Hazards Reduction Program (NEHRP), 1997. Recommended provisions for seismic regulations for new buildings and other structures, *Technical Report Fema 302*, U.S. Federal Emergency Management Agency, Washington, D.C.

Neidell, N. S. & Taner, M. T., 1971. Semblance and other coherency measures for multichannel data, *Geophysics*, **36**(3), 482–497.

Newton, C. & Snieder, R., 2012. Estimating intrinsic attenuation of a building using deconvolution interferometry and time reversal, *Bulletin of the Seismological Society of America*, **102**(5), 2200–2208.

Nishida, K., 2013. Global propagation of body waves revealed by cross-correlation analysis of seismic hum, *Geophysical Research Letters*, **40**(9), 1691–1696.

Ohrnberger, M., Vollmer, D., & Scherbaum, F., 2006. WARAN – A mobile wireless array analysis system for in-field ambient vibration dispersion curve estimation, in *First European Conference on Earthquake Engineering and Seismology (ECEES), Abstract-ID 2017*, Geneva, Switzerland, 3-8 September 2006.

Parolai, S., Ansal, A., Kurtulus, A., Strollo, A., Wang, R., & Zschau, J., 2009. The Ataköy vertical array (Turkey): Insights into seismic wave propagation in the shallow-most crustal layers by waveform deconvolution, *Geophysical Journal International*, **178**(3), 1649–1662.

Parolai, S., Bindi, D., Ansal, A., Kurtulus, A., Strollo, A., & Zschau, J., 2010. Determination of shallow S-wave attenuation by down-hole waveform deconvolution: a case study in Istanbul (Turkey), *Geophysical Journal International*, **181**(2), 1147–1158.

Parolai, S., Wang, R., & Bindi, D., 2012. Inversion of borehole weak motion records observed in Istanbul (Turkey), *Geophysical Journal International*, **188**(2), 535–548.

Parolai, S., Bindi, D., & Pilz, M., 2015. k0: The role of intrinsic and scattering attenuation, *Bulletin of the Seismological Society of America*, **105**(2A), 1049–1052.

Paul, A., Campillo, M., Margerin, L., & Larose, E., 2005. Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves, *Journal of Geophysical Research*, **110**(B8), B08302.

Peyrat, S., Madariaga, R., Buforn, E., Campos, J., Asch, G., & Vilotte, J. P., 2010. Kinematic rupture process of the 2007 Tocopilla earthquake and its main aftershocks from teleseismic and strong-motion data, *Geophysical Journal International*, **182**(3), 1411–1430.

Pilz, M. & Fäh, D., 2017. The contribution of scattering to near-surface attenuation, *Journal of Seismology*, **21**(4), 837–855.

Pilz, M., Parolai, S., Picozzi, M., & Bindi, D., 2012. Three-dimensional shear wave velocity imaging by ambient seismic noise tomography, *Geophysical Journal International*, **189**(1), 501–512.

Pitilakis, K., Raptakis, D., Lontzetidis, K., Tika-Vassilikou, T., & Jongmans, D., 1999. Geotechnical and geophysical description of Euro-Seistest, using field, laboratory tests and moderate strong motion recordings, *J. Earthq. Eng.*, **3**(3), 381–409.

Pitilakis, K., Roumelioti, Z., Raptakis, D., Manakou, M., Liakakis, K., Anastasiadis, a., & Pitilakis, D., 2013. The EUROSEISTEST strong-motion database and web portal, *Seismol. Res. Lett.*, **84**(5), 796–804.

Poli, P., Pedersen, H. A., Campillo, M., & Group, P. W., 2012. Emergence of body waves from cross-correlation of short period seismic noise, *Geophysical Journal International*, **188**(2), 549–558.

Prieto, G., Denolle, M., Lawrence, J., & Beroza, G., 2011. On amplitude information carried by the ambient seismic field, *Comptes Rendus Geoscience*, **343**(8), 600–614.

Prieto, G. A., Lawrence, J. F., & Beroza, G. C., 2009. Anelastic Earth structure from the coherency of the ambient seismic field, *Journal of Geophysical Research*, **114**(B7), B07303.

Prieto, G. A., Lawrence, J. F., Chung, A. I., & Kohler, M. D., 2010. Impulse response of civil structures from ambient noise analysis, *Bulletin of the Seismological Society of America*, **100**(5A), 2322–2328.

Pujol, J. & Smithson, S., 1991. Seismic wave attenuation in volcanic rocks from VSP experiments, *Geophysics*, **56**(9), 1441–1455.

Raptakis, D., Chávez-Garcıa, F., Makra, K., & Pitilakis, K., 2000. Site effects at Euroseistest—I. Determination of the valley structure and confrontation of observations with 1D analysis, *Soil Dyn. Earthq. Eng.*, **19**, 1–22.

Raub, C., Bohnhoff, M., Petrovic, B., Parolai, S., Malin, P., Yanik, K., Kartal, R. F., & Kilic, T., 2016. Seismic-wave propagation in shallow layers at the GONAF-Tuzla site, Istanbul, Turkey, *Bulletin of the Seismological of America*, **106**(3), 912–927.

Riggelsen, C., Gianniotis, N., & Scherbaum, F., 2011. Learning aggregations of ground-motion models using data, in *Proceedings of the 11th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP'11, Zürich, Switzerland, 1-4 August 2011*, p. 721–728, Taylor and Francis Group/CRC Press.

Rodriguez-Marek, A., Rathje, E. M., Bommer, J. J., Scherbaum, F., & Stafford, P. J., 2014. Application of single-station sigma and site-response characterization in a probabilistic seismic-hazard analysis for a new nuclear site, *Bulletin of the Seismological Society of America*, **104**(4), 1601–1619.

Roselli, P., Marzocchi, W., & Faenza, L., 2016. Toward a new probabilistic framework to score and merge ground-motion prediction equations: The case of the italian region, *Bulletin of the Seismological Society of America*, **106**(2), 720–733.

Roux, P. & Kuperman, W. A., 2004. Extracting coherent wave fronts from acoustic ambient noise in the ocean, *The Journal of the Acoustical Society of America*, **116**(4), 1995–2003.

Roux, P., Sabra, K. G., Gerstoft, P., Kuperman, W., & Fehler, M. C., 2005a. P-waves from cross-correlation of seismic noise, *Geophysical Research Letters*, **32**(19).

Roux, P., Sabra, K. G., Kuperman, W. A., & Roux, A., 2005b. Ambient noise cross correlation in free space: Theoretical approach, *The Journal of the Acoustical Society of America*, **117**(1), 79–84.

Runge, A. K., Händel, A., Riggelsen, C., Scherbaum, F., & Kühn, N. M., 2013a. A smart elicitation technique for informative priors in ground-motion mixture modelling, in *EGU General Assembly Conference Abstracts*, vol. 15, Vienna, Austria, April 7-12, 2013.

Runge, A. K., Scherbaum, F., Curtis, A., & Riggelsen, C., 2013b. An interactive tool for the elicitation of subjective probabilities in probabilistic seismic-hazard analysis, *Bulletin of the Seismological Society of America*, **103**(5), 2862–2874.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C., 2005. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophysical Research Letters*, **32**(3), 1–5.

Sato, H., Fehler, M. C., & Maeda, T., 2012. *Seismic wave propagation and scattering in the heterogeneous earth*, vol. 484, Springer.

Sawazaki, K., Sato, H., Nakahara, H., & Nishimura, T., 2009. Time-lapse changes of seismic velocity in the shallow ground caused by strong ground motion shock of the 2000 Western-Tottori earthquake, Japan, as revealed from coda deconvolution analysis, *Bulletin of the Seismological Society of America*, **99**(1), 352–366.

Scherbaum, F. & Kuehn, N. M., 2011. Logic tree branch weights and probabilities: Summing up to one is not enough, *Earthquake Spectra*, **27**(4), 1237–1251.

Scherbaum, F., Cotton, F., & Smit, P., 2004a. On the use of response spectral-reference data for the selection and ranking of ground-motion models for

seismic-hazard analysis in regions of moderate seismicity: The case of rock motion, *Bulletin of the Seismological Society of America*, **94**(6), 2164–2185.

Scherbaum, F., Schmedes, J., & Cotton, F., 2004b. On the conversion of source-to-site distance measures for extended earthquake source models, *Bulletin of the Seismological Society of America*, **94**(3), 1053–1069.

Scherbaum, F., Bommer, J. J., Bungum, H., Cotton, F., & Abrahamson, N. A., 2005. Composite ground-motion models and logic trees: methodology, sensitivities, and uncertainties, *Bulletin of the Seismological Society of America*, **95**(5), 1575–1593.

Scherbaum, F., Delavaud, E., & Riggelsen, C., 2009. Model selection in seismic hazard analysis: An information-theoretic perspective, *Bulletin of the Seismological Society of America*, **99**(6), 3234–3247.

Scherbaum, F., Kuehn, N. M., Ohrnberger, M., & Koehler, A., 2010. Exploring the proximity of ground-motion models using high-dimensional visualization techniques, *Earthquake Spectra*, **26**(4), 1117–1138.

Schurr, B., Asch, A., Sodoudi, F., Manzanares, A., Ritter, O., Klotz, J., Chong-Diaz, G., Barrientos, S., Villotte, J.-P., & Oncken, O., 2009. The International Plate Boundary Observatory Chile (IPOC) in the northern Chile seismic gap, in *EGU General Assembly Conference Abstracts*, vol. 11, p. 11040.

Seats, K., Lawrence, J., & Prieto, G., 2012. Improved ambient noise correlation functions using Welch's method, *Geophysical Journal International*, **188**(2), 513–523.

Shakal, A. F., Boore, D., Chiou, B. S.-J., Iwan, W. D., & O'Connell, D. R., 2005. Guidelines and recommendations for strong-motion record processing and commentary, *COSMOS Publication No. CP-2005/01*.

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface-wave tomography from ambient seismic noise., *Science*, **307**, 1615–1618.

Shearer, P. & Orcutt, J., 1987. Surface and near-surface effects on seismic waves—theory and borehole seismometer results, *Bulletin of the Seismological Society of America*, **77**(4), 1168–1196.

Shirzad, T. & Shomali, Z.-H., 2015. Extracting seismic body and Rayleigh waves from the ambient seismic noise using the rms-stacking method, *Seismological Research Letters*, **86**(1), 173–180.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Physical Review E*, **69**(4), 046610.

Snieder, R., 2009. Extracting the time-domain building response from random vibrations, in *Coupled Site and Soil-Structure Interaction Effects with Application to Seismic Risk Mitigation*, pp. 283–292, eds Schanz, T. & Iankov, R., Springer.

Snieder, R. & Larose, E., 2013. Extracting earth's elastic wave response from noise measurements, *Annual Review of Earth and Planetary Sciences*, **41**, 183—-206.

Snieder, R., Wapenaar, K., & Wegler, U., 2007. Unified Green's function retrieval by cross-correlation; connection with energy principles, *Phys. Rev. E*, **75**(3), 036103.

Snieder, R., Miyazawa, M., Slob, E., Vasconcelos, I., & Wapenaar, K., 2009. A comparison of strategies for seismic interferometry, *Surveys in Geophysics*, **30**(4-5), 503–523.

Snieder, R. & Şafak, E., 2006. Extracting the building response using seismic interferometry: Theory and application to the Millikan Library in Pasadena, California, *Bulletin of the Seismological Society of America*, **96**(2), 586–598.

Stehly, L., Campillo, M., Froment, B., & Weaver, R., 2008. Reconstructing Green's function by correlation of the coda of the correlation (C^3) of ambient seismic noise, *Journal of Geophysical Research*, **113**(B11306), 1–10.

Stewart, J. P., Douglas, J., Javanbarg, M., Bozorgnia, Y., Abrahamson, N. A., Boore, D. M., Campbell, K. W., Delavaud, E., Erdik, M., & Stafford, P. J., 2015. Selection of ground motion prediction equations for the global earthquake model, *Earthquake Spectra*, **31**(1), 19–45.

Strasser, F. O., Arango, M. C., & Bommer, J. J., 2010. Scaling of the source dimensions of interface and intraslab subduction-zone earthquakes with moment magnitude, *Seismological Research Letters*, **81**(6), 941–950.

Takagi, R., Nakahara, H., Kono, T., & Okada, T., 2014. Separating body and Rayleigh waves with cross terms of the cross-correlation tensor of ambient noise, *Journal of Geophysical Research: Solid Earth*, **119**(3), 2005–2018.

Titterington, D. M., Smith, A. F. M., & Makov, U. E., 1985. *Statistical analysis of finite mixture distributions*, Wiley series in probability and mathematical statistics, Wiley, New York.

Tonn, R., 1991. The determination of the seismic quality factor Q from VSP data: a comparison of different computational methods, *Geophys. Prospect.*, **39**, 1–27.

Trampert, J., Cara, M., & Frogneux, M., 1993. SH Propagator Matrix and Qs Estimates From Borehole- and Surface-Recorded Earthquake Data, *Geophysical Journal International*, **112**(2), 290–299.

Tsai, V. C., 2011. Understanding the amplitudes of noise correlation measurements, *Journal of Geophysical Research*, **116**(9), 1–16.

Udias, A., 1999. *Principles of seismology*, Cambridge University Press.

USNRC, 2012. Practical implementation guidelines for SSHAC Level 3 and 4 hazard studies, *NUREG-2117*, US Nuclear Regulatory Commission, Washington, D.C.

Van Houtte, C., Drouet, S., & Cotton, F., 2011. Analysis of the origins of κ (kappa) to compute hard rock to rock adjustment factors for GMPEs, *Bulletin of the Seismological Society of America*, **101**(6), 2926–2941.

Vasconcelos, I. & Snieder, R., 2008. Interferometry by deconvolution: Part 2 — Theory for elastic waves and application to drill-bit seismic imaging, *Geophysics*, **73**(3), S129–S141.

Wang, R., 1999. A simple orthonormalization method for stable and efficient computation of green's functions, *Bulletin of the Seismological Society of America*, **89**(3), 733–741.

Wang, R., Schurr, B., Milkereit, C., Shao, Z., & Jin, M., 2011. An improved automatic scheme for empirical baseline correction of digital strong-motion records, *Bulletin of the Seismological Society of America*, **101**(5), 2029–2044.

Wapenaar, K., 2004. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.*, **93**(25), 1–4.

Weaver, R. L., 2011. On the amplitudes of correlations and the inference of attenuations, specific intensities and site factors from ambient noise, *Comptes Rendus Geoscience*, **343**(8-9), 615–622.

Weaver, R. L., 2013. On the retrieval of attenuation and site amplifications from ambient noise on linear arrays: further numerical simulations, *Geophysical Journal International*, **193**(3), 1644–1657.

Weaver, R. L. & Lobkis, O. I., 2004. Diffuse fields in open systems and the emergence of the Green's function, *The Journal of the Acoustical Society of America*, **116**(5), 2731–2734.

Weber, B., 2002. Tragwerksdynamik, *Lecture notes*, pp. 1–178, ETH Zürich, Abteilung für Bauingenieurwesen, Institut für Baustatik und Konstruktion (in German).

Weemstra, C., Boschi, L., Goertz, A., & Artman, B., 2013. Seismic attenuation from recordings of ambient noise, *Geophysics*, **78**(1), Q1–Q14.

Welch, P., 1967. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Transactions on Audio and Electroacoustics*, **15**(2), 17–20.

Wells, D. L. & Coppersmith, K. J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bulletin of the Seismological Society of America*, **84**(4), 974–1002.

Wessel, P. & Smith, W. H. F., 1998. New, improved version of generic mapping tools released, *EOS, Trans. Am. Geophys. Un.*, **79**(47), 579.

Xia, J., Yin, X., & Xu, Y., 2013. Feasibility of determining Q of near-surface materials from Love waves, *J. Appl. Geophys.*, **95**, 47–52.

Yilmaz, O., 1987. *Seismic Data Processing*, Society of Exploration Geophysicists.

Youngs, R. R., Chiou, B. S.-J., Silva, W. J., & Humphrey, J. R., 1997. Strong ground motion attenuation relationships for subduction zone earthquakes, *Seismological Research Letters*, **68**(1), 58–73.

Zhan, Z., Ni, S., Helmberger, D. V., & Clayton, R. W., 2010. Retrieval of moho-reflected shear wave arrivals from ambient seismic noise, *Geophysical Journal International*, **182**(1), 408–420.

Zhang, J. & Yang, X., 2013. Extracting surface wave attenuation from seismic noise using correlation of the coda of correlation, *Journal of Geophysical Research*, **118**(5), 2191–2205.

Zhang, J., Gerstoft, P., & Shearer, P. M., 2009. High-frequency P-wave seismic noise driven by ocean winds, *Geophysical Research Letters*, **36**(9).

Zhao, J. X., Irikura, K., Zhang, J., Fukushima, Y., Somerville, P. G., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., & Ogawa, H., 2006a. An empirical site-classification method for strong-motion stations in Japan using H/V response spectral ratio, *Bulletin of the Seismological Society of America*, **96**(3), 914–925.

Zhao, J. X., Zhang, J., Asano, A., Ohno, Y., Oouchi, T., Takahashi, T., Ogawa, H., Irikura, K., Thio, H. K., Somerville, P. G., Fukushima, Y., & Fukushima, Y., 2006b. Attenuation relations of strong ground motion in Japan using site classification based on predominant period, *Bulletin of the Seismological Society of America*, **96**(3), 898–913.

A | Appendix: Supplementary material mixture model approach

A.1 List of regression coefficients

Table A.1: List of regression coefficients

Period[s]	a	b	c	d	e	s_i	s_s
0.00	0.7401	-0.0098	1.0273	-0.1099	0.0071	0.2841	4.1648
0.03	0.6705	-0.0091	1.2395	-0.1161	0.0082	1.4474	5.1601
0.05	0.7702	-0.0090	1.3255	-0.1055	0.0083	1.7109	5.0176
0.10	1.1981	-0.0128	0.3231	-0.0171	0.0066	-2.0628	0.4953
0.20	1.1514	-0.0115	0.2009	-0.0197	0.0063	-2.3766	0.8320
0.30	1.1851	-0.0108	0.1132	-0.0253	0.0033	-3.2995	0.7593
0.50	0.9312	-0.0084	0.6730	-0.0937	0.0026	-1.6475	2.7501
1.00	1.0526	-0.0047	1.0174	-0.1027	0.0027	-2.3731	3.0927
2.00	1.2670	-0.0040	1.0513	-0.0900	0.0031	-4.4886	0.5335
3.00	1.1306	-0.0043	1.3145	-0.1383	0.0040	-4.3997	0.2381

Period[s]	s_{sl}	x	q_i	q_s	Φ	τ	σ_{tot}
0.00	-0.6199	0.4998	-0.0662	-0.1807	0.3745	0.1568	0.7289
0.03	-0.5782	0.6268	-0.0688	-0.1934	0.3708	0.1720	0.7367
0.05	-0.4598	0.3452	-0.0059	-0.1294	0.4131	0.1870	0.7746
0.10	-0.2977	0.3238	-0.0243	-0.1352	0.4413	0.1854	0.7916
0.20	-0.5038	0.6891	-0.1150	-0.2745	0.4019	0.1424	0.7378
0.30	-0.6676	0.7220	-0.1294	-0.3320	0.4073	0.1392	0.7392
0.50	-0.7935	0.7786	-0.2236	-0.3403	0.3807	0.1418	0.7229
1.00	-1.0475	0.6917	-0.2363	-0.3231	0.2995	0.1105	0.6403
2.00	-0.9444	0.7777	-0.2641	-0.3362	0.2660	0.0938	0.5999
3.00	-0.9130	0.8795	-0.3722	-0.3407	0.2436	0.0884	0.5762

B | Appendix: Supplementary material deconvolution approach

B.1 Numerical tests

The causal and acausal signal components partially overlap in the deconvolution traces. We conducted numerical tests in order to study the influence that this overlap can have on the estimation of Q_s^{-1}.

We used QSEIS (Wang, 1999), a code that employs the orthonormal propagator algorithm, to compute synthetic seismograms for a viscoelastic 1-D subsurface model. According to the results obtained for the West Bohemia/Vogtland area within the present study (Subsection 4.5.1), v_p and v_s were set to 2000 m/s and 1000 m/s, respectively, and Q_s was chosen to be 20. A source was placed at 4.2 km depth on an arbitrarily chosen fault plane with a strike, dip and rake of 170°, 75° and -30°, respectively. Synthetic seismograms were calculated for receivers that are placed at the surface and at different depths (50-500 m) directly above the source. White noise with 10% of the maximum amplitude was added to the synthetics. We applied the same procedure to the synthetics as to the real data by first computing the deconvolution with respect to the surface sensor according to Eq. 4.6. Epsilon had to be set to 0.1 in order to obtain stable deconvolutions. Q_s^{-1} was then derived using Eq. 4.9. The results are shown for the transverse component of motion.

Fig. B.1 shows the deconvolutions obtained for borehole receivers at depths of 50 m, 87 m, 100 m, 200 m and 500 m. At shallow depths, the causal and signal components partially overlap.

Fig. B.2 summarizes the Q_s^{-1} results that are obtained for different sensor depths. The input Q_s^{-1} of 0.05 can be recovered for frequencies between 1-12 Hz. The result is most clear for the deepest sensor at 500 m. For shallow receivers the results tend to fluctuate around the true Q_s^{-1} value. Q_s^{-1} is overestimated at frequencies below 6 Hz for the receiver at 50 m depth. This might be caused by the insufficient causal and acausal signal separation at shallow depth which will be most prominent at low frequencies.

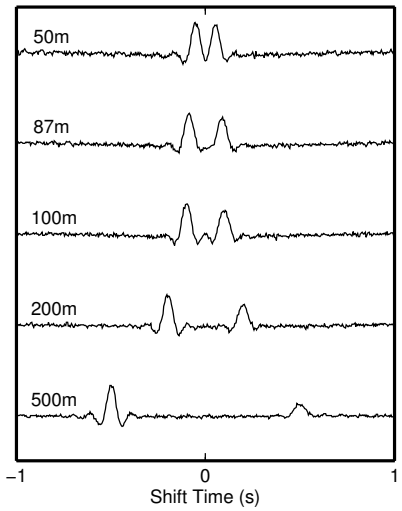

Figure B.1: Deconvolution traces obtained from synthetic seismograms that were computed using QSEIS. Shown is the transverse component and the results for different sensor depths after deconvolution with the motion recorded at a surface sensor.

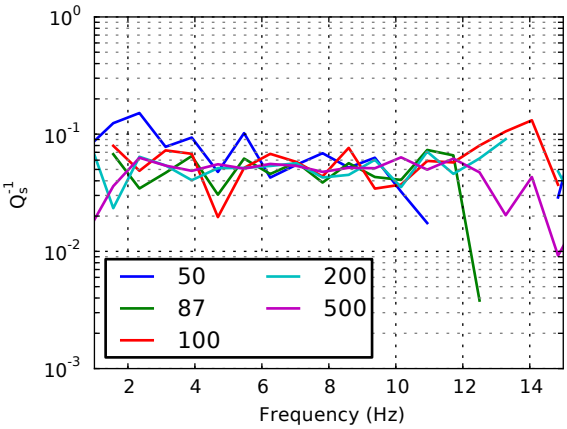

Figure B.2: Q_s^{-1} that is obtained from the deconvolution traces shown in Fig. B.1. The sensor depth is color coded.