DISS. ETH NO. 21328

STATISTICS OF AMBIENT SEISMIC NOISE: INVESTIGATIONS INTO SPECTRAL ATTRIBUTES AND MULTI-COMPONENT ARRAY PROCESSING

A dissertation submitted to
ETH ZURICH

for the degree of
DOCTOR OF SCIENCES

presented by

Nima Riahi
Dipl. Phys. ETH Zurich

born 21. September 1979

citizen of
Adliswil (ZH)

Accepted on the recommendation of

Prof. Dr. Jean–Pierre Burg, ETH Zurich (examiner)
PD Dr. Erik H. Saenger, ETH Zurich (co-examiner)
Dr. Gregor Hillers, Université Josef Fourier, Grenoble, F (co-examiner)
Dr. Alex V. Goertz, PGS Geophysical AS, Oslo, NO (co-examiner)

2013

We have to remember that what we observe is not nature herself, but nature exposed to our method of questioning.

Werner Karl Heisenberg, in *Physics and Philosophy*
[Heisenberg, 1958]

[. . .] given any rule, however "fundamental" or "necessary" for science, there are always circumstances when it is advisable not only to ignore the rule, but to adopt its opposite.

Paul Karl Feyerabend, in *Against Method*
[Feyerabend, 1975]

For Iran and Sadegh

Contents

Abstract

Seismic background vibrations, conventionally considered noise in seismology and exploration geophysics, can be used to study and monitor the Earth's subsurface. Suggested techniques loosely fall into three categories: wave field interferometry, array analysis, and spectral attributes of single receiver recordings. In this thesis we present methodological developments and results for the latter two categories.

Studies using spectral seismic attributes face two challenges: the composition of the wave field from which a passive seismic attribute is computed is not known (e.g. *surface-vs-body wave issue*) and correlations between spectral attributes and geologic targets can be confounded by unrelated variables such as noisy infrastructure, shallow resonances, or acquisition footprint *(correlation-vs-causation issue)*. We present a statistical strategy to address these two issues. First, joint statistics of spectral density in seismic power and polarization azimuth are used as an approximate means to know when surface waves clearly dominate. Second, a quantitative framework is used to describe seismic attribute-to-target correlations for several attribute parameterizations and different targets. Using data from a producing tight-gas field, we statistically test the hypothesis that hydrocarbon reservoirs can have a measurable effect on passive seismic attributes on the surface. The hypothesis cannot be falsified for the available data.

A three-component frequency-wavenumber processor is developed to jointly analyze back azimuth, phase velocity, and polarization of ambient surface waves. The processor is applied on an exploration scale to perform a time-lapse study of ambient surface wave anisotropy above an underground gas storage reservoir. Existing anisotropy, probably caused by the preferred orientation of local faults, is found to be accentuated between 0.4-0.6 Hz when the reservoir is close to its maximum fill. This implies that reservoir pore pressure could be monitored using seismic noise. The processor is also applied on a regional scale to study azimuthal surface wave anisotropy in Southern California. One year of ambient surface wave noise (2012) provides us with sufficient data points and illumination to constrain 2θ and 4θ Rayleigh wave anisotropy with unprecedented precision and sensitivity. The findings are a further testament to the inferential power afforded by array processing of seismic noise.

Zusammenfassung

Seismische Hintergrundvibrationen – in der Seismologie und Explorations-Geophysik häufig als Störsignal betrachtet – lassen sich zur Erkundung des Erduntergrundes nutzen. Hierzu vorgeschlagene Methoden können grob in drei Kategorien eingeteilt werden: Wellenfeld Interferometrie, Array Analyse und spektrale Attribute. Die vorliegende Arbeit stellt methodologische Entwicklungen und Resultate vor, die dem Bereich letzterer zwei Kategorien angehören.

Auf spektralen seismischen Attributen beruhende Untersuchungen haben zwei Schwierigkeiten zu bewältigen: Die Attribute basieren auf Wellenfeldern, deren Eigenschaften nur ungenügend bekannt sind (z.B. *Oberflächen-vs-Körperwellen Thematik*) und zweitens kann die Korrelation zu einem geologischen Zielobjekt tatsächlich verursacht sein durch oberflächennahe Eigenschaften oder der Signatur seismischer Lärmquellen (*Korrelation-vs-Kausalität*). Um diese Probleme anzugehen stellen wir eine zweistufige Strategie vor. Die Statistik der spektralen Dichte sowohl der seismischen Leistung wie auch des Polarisationsazimuths werden genutzt um zu erkennen in welchen Zeitfenstern Oberflächenwellen klar dominieren. Des Weiteren wird ein quantitativer Ansatz genutzt zur Beschreibung von Korrelationen zwischen seismischen Attributen unterschiedlicher Parametrisierung und mehreren Zielobjekten. Basierend auf Felddaten von einem aktiv betriebenen Gasfeld prüfen wir die Hypothese wonach Kohlenwasserstoff Lagerstätten einen messbaren Effekt auf an der Oberfläche gemessenen passiven Attributen haben können. Für den untersuchten Datensatz kann die Hypothese nicht falsifiziert werden.

Ein drei-komponenten Array Rechner wird entwickelt, der Einfallsrichtung, Phasengeschwindigkeit und Polarisation einfallender Umgebungsoberflächenwellen ermittelt. Wir wenden den Rechner auf der Explorationsskala an, um Oberflächenwellenanisotropie über einem unterirdischen Gasspeicher während zweier Zeitpunkte zu erfassen. Eine wahrscheinlich durch die Vorzugsrichtung der lokalen Bruchsysteme verursachte Anisotropie wird dabei zwischen 0.4-0.6 Hz verstärkt wenn der Speicher sich Nahe der operativen Maximalfüllung befindet. Die Beobachtung deutet auf eine mögliche Anwendung hin: Die Überwachung des Porendrucks in geologischen Lagerstätten mittels seismischen Hintergrundrauschens. Der Array Rechner wird auch auf regionaler Skala genutzt zur Messung von Oberflächenwellenanisotropie in Südkalifornien. Seismisches Oberflächenwellenrauschen aus einem Jahr (2012) liefert uns genügend

Datenpunkte und Ausleuchtung, um 2θ und 4θ Rayleigh Wellen Anisotropie mit bisher unerreichter Präzision und Empfindlichkeit zu messen. Die Resultate sind ein weiterer Ausdruck der Aussagekraft von Studien seismischen Rauschens mittels Arrays.

Chapter 1

Introduction

1.1 The case for noise

Dolphins locate and identify obstacles by producing broadband click sounds and listening to their reflections from those objects. This biological remote-sensing strategy depends on the fact that the time, location, and signature of the acoustic source are controlled by the animal. The principle of exploring and characterizing a wave field supporting medium using controlled or known sources is found not only in a host of other animal species, but also over a wide range of man-made remote-sensing technologies. Examples include air traffic radar, rain radar, submarine sonar, ultrasound imaging, georadar, non-destructive testing, upper atmosphere analysis with infrasound, seismic reflection and refraction surveys, and parts of earthquake seismology. The latter two applications are particularly relevant for this thesis because they use seismic waves as a probing signal. The earth is explored and monitored using known sources such as explosives, earthquakes, vibrotrucks, or sledge hammers. They represent a mature methodology whose scientific and economic value has been proven by the mining and hydrocarbon industries (e.g. Sheriff & Geldart [1995], Mooney [2007], Aki & Richards [2002]).

However, the fundamental notion that the source must be controlled or known for a remote-sensing operation to be successful has increasingly been questioned in the recent past. There has been an intensified focus on what the know-your-source paradigm considers seismic noise: the continuous signal observed on seismograms recorded anywhere on the planet when no transient or active sources are present. One part of this signal is indeed noise due to the recording equipment with

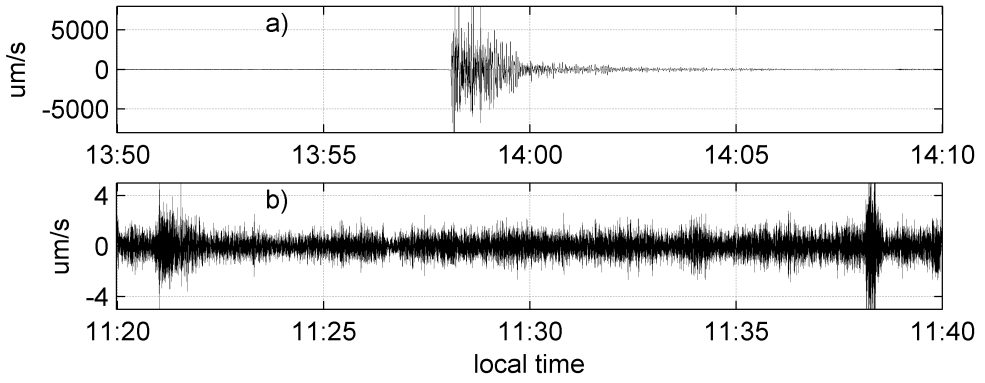

FIGURE 1.1: On August 26$^{\text{th}}$, 2012, a M=5.4 earthquake occurred about 8.2 km underneath the town of Brawley in Southern California. Panel (a) shows the resulting ground motion at a station about 35 km away from the event. Analogous to a flash of light, the event is sudden, clearly dominates the seismic motion, and we know fairly well when and where the energy of the event was released. In contrast, panel (b) shows the seismic motion at the same location 2.5 hours earlier (note the difference in scale). No single event can be made out, the ground motion is weak and continuous, and little is known about the sources that caused the vibrations. It is the latter type of signal that takes center stage in this thesis.

no relation to the Earth [Sleeman et al., 2006]. But the other part is due to actual background earth motion and is more aptly referred to as the *ambient seismic wave field*. In this thesis, we use these words to refer to any collection of seismic processes for which there is little to no deterministic information about the source signatures or even the sources themselves. Such indeterministic sources may be, among a long list of possibilities: volcanic activity [Julian, 1994], ocean gravity waves [Longuet-Higgins, 1950, Webb & Cox, 1986, Kedar et al., 2008, Hillers et al., 2012b], cultural activity [Hillers et al., 2012a], wind action [Withers et al., 1996], temperature variations [Hillers & Ben-Zion, 2011], and slow slip plate motion [Nadeau & Dolenc, 2005, Kao et al., 2006]. As a result of the wide range of possible sources, the ambient seismic wave field is a perpetual and ubiquitous phenomenon, albeit with characteristics that may vary with place and time. For illustration, Figure 1.1 compares the seismic trace of a transient earthquake arrival, which in this context may be considered the signature of a known source, to that of ambient seismic vibrations a few hours earlier.

The important aspect is that this wave field is seismic in nature. It *"consists of waves that reflect and refract around exactly the same subsurface heterogeneities as waves excited by active sources."* [Curtis et al., 2006]. This means it can, in

principle, offer the same scientific and economic benefits as active seismic and earthquakes studies. Ironically, exploiting seismic noise might in some cases be even more attractive than active source studies. For instance, areas with low seismicity or poor earthquake illumination can be studied with it. The continuous illumination with noise also lends itself well to monitoring for changes in the subsurface. Finally, the fact that passive studies require no active sources can significantly reduce costs and their environmental impact, potentially opening new possibilities for research and exploration in areas where operating active sources meets environmental or legal constraints.

1.2 Research context: interferometry, arrays, and spectral attributes

The challenge of studies based on ambient seismic noise is to turn the uncontrolled ambient seismic energy into a probing signal for the earth medium. Essentially, the domain where reasonable information is known about the sources shifts from the deterministic (amplitude, phase, location) to the statistic (spectral content, intensity, diffusivity). Various seismic techniques have been developed that are adapted to this new source paradigm.

Since about 2000, an increasing amount of research has been devoted to the study of cross-correlations between recordings of the ambient wave field from different locations. Both theory and laboratory experiments have shown that the impulse response of the studied medium (called the Green's function in seismology) can be retrieved if the ambient wave field satisfies some diffusivity assumptions (Larose et al. [2006], and references therein). Weaver & Lobkis [2006] state that, on theoretical grounds, the retrieval of the Green's function "is a robust consequence of any of a variety of definitions of diffuse fields." This is quite a promise: in theory the Green's function captures the totality of all seismic information between two locations. Techniques exploiting this phenomenon are often referred to as seismic interferometry. The applicability of this new paradigm has been shown in a wealth of studies. Among those were surface wave tomography studies [Shapiro et al., 2005, Sabra et al., 2005, Moschetti et al., 2007, Larose et al., 2008, Lin et al., 2008, Stehly et al., 2009, Mordret et al., 2013, Weemstra et al., 2013], monitoring applications [Brenguier et al., 2008, Durand et al., 2011, de Ridder & Biondi, 2013],

and body wave reflector imaging [Roux et al., 2005, Artman, 2006, Draganov et al., 2007, 2009, Ruigrok et al., 2011, Nakata et al., 2011, Hillers et al., 2012a, Poli et al., 2012a, Xu et al., 2012, Draganov et al., 2013]. The latter outline the potential for hydrocarbon exploration in environmentally sensitive areas and for deep targets that require strong sources.

An important limitation of interferometry, however, is that the diffusivity assumptions are not always valid in the real world, which can lead to bias [Weaver & Lobkis, 2006, Tsai, 2009, Harmon et al., 2010, Draganov et al., 2010]. Efforts to mitigate such errors are ongoing. Furthermore, the passive seismic paradigm has so far not been able to match the high resolution and quality of its active-source counterpart for seismic exploration.

Another group of techniques analyzes the ambient seismic wave field on a seismometer array by exploiting stable phase differences among them. The ambient wave field is considered a stochastic space-time variable which is then characterized by its frequency-wavenumber spectrum. Examples include the spatial autocorrelation method (SPAC) [Aki, 1957, Okada, 2003] and different flavors of frequency-wavenumber analysis or beamforming [Toksoz, 1964, Horike, 1985, Scherbaum et al., 2003, Parolai et al., 2005, Kind et al., 2005] which attempt to estimate the phase velocity dispersion of surface waves and then invert for subsurface parameters. Applications include geotechnical investigations, seismic hazard assessment, and the inference of velocity depth profiles. Array processing has also been used to study body waves. The arrival of continuous p-waves has been used to locate their generation areas in the ocean and monitor their intensity [Gerstoft et al., 2008b, Koper & de Foy, 2008, Zhang et al., 2009]. A similar application, often referred to as phased arrays, is implemented in acoustics to produce images of scattered ambient sounds, a bit like optical photography images scattered visible light [Buckingham et al., 1996, Epifanio et al., 1999, Chitre et al., 2012]. Array analysis is complementary to interferometry in that it requires dominant trains of coherent waves rather than a diffuse wave field.

There are also techniques that study ambient noise attributes in the frequency domain using single receivers. The H/V spectral ratio method (HVSR) studies the average ratio of horizontal to vertical energy to characterize the subsurface beneath the receiver. The approach is commonly used in geotechnical contexts and seismic hazard assessment [Fäh et al., 2001, Parolai et al., 2002, 2005]. Average spectral attributes of single-station recordings allow also detecting

voids [Nasseri Moghaddam et al., 2005] or heterogeneities such as hydrocarbons reservoirs [Saenger et al., 2009, Lambert et al., 2009b, 2012]. This application is different from the previous ones in that it focuses on body waves. Such reliance on body waves has led to some debate [Green & Greenhalgh, 2009a, Lambert et al., 2009a]. Spectral attributes are point measurements and more limited in their ability to characterize the wave field. Interpretation therefore plays a very important role. Spectral attribute studies are attractive in terms of logistics and cost: very few sensors (even one) would suffice and sensors do not necessarily have to record synchronously.

Techniques that study ambient seismic signals share some commonalities: array processing techniques and interferometric methods both rely on cross-correlations. The distinction hardly matters for two-sensor versions of SPAC [Chavez-Garcia et al., 2005, Hayashi et al., 2013] which are essentially a frequency-domain version of interferometry [Boschi et al., 2013], as shown by Yokoi & Margaryan [2008] and Tsai & Moschetti [2010]. A certain diffusivity of the ambient wave field is required explicitly by the HVSR technique [Fäh et al., 2001]. Still, there are substantial differences in how the ambient wave field is analyzed, which requirements it has to satisfy, and which instrumental setups are needed. Different techniques have different merits and drawbacks and which technique to use depends on the circumstances or available data.

1.3 Contributions of this thesis

This thesis contributes methodological developments and observations in the field of spectral seismic attributes and frequency-wavenumber array analysis. The guiding principle in the studies presented is to segment ambient seismic data into many short time windows and analyze them in the frequency domain. Using short time windows should increase the chance that the ambient vibrations approximate a stationary random wave field. Furthermore, ambient vibrations have been fruitfully modeled before as stochastic signals which, with reference to the spectral theorem (e.g. Yaglom & Silverman [1962]), have a natural representation in the frequency domain. Chapter 2 addresses two perennial challenges in the use and interpretation of ambient seismic spectral attributes. One issue is related to the lack of knowledge about the wave field that was used to compute the spectral attributes, in particular about the types of surface and body waves involved.

Fäh et al. [2001] and Lambert et al. [2011a] suggest processing measures in the time and frequency domains to alleviate such issues. In this chapter, we show how the statistics of the spectral density of both power and polarization of the ambient seismic particle motion can also be used. The other issue pertains to the interpretation of spectral attribute maps: spatial correlations to a feature in the deep subsurface, in this case a hydrocarbon reservoir, may be confounded with near-surface factors that can affect the seismic signal at a recording station (infrastructure noise, wind, near-surface resonance, and subsurface scatterers may all play a part). We introduce a quantitative framework that allows comparing correlations of the ambient wave field as a function of frequency and seismic "loudness" to a hydrocarbon reservoir at depth and several potential confounders near the surface. This chapter is published in *Geophysical Journal International* [Riahi et al., 2013b].

The other contribution of this thesis addresses the relative paucity of three-component frequency-wavenumber studies of ambient seismic noise (e.g. Poggi & Fäh [2010], Marano et al. [2012], Behr et al. [2013]). A frequency-wavenumber-polarization array processor was implemented according to existing ideas from earthquake seismology [Esmersoy et al., 1985, Wagner, 1996]. Frequency-wavenumber methods compute phase shifts among the receivers to infer propagation azimuth and phase velocity of wave trains. The processor we use includes phase shifts among the three-components of the array sensors to additionally infer the polarization information. In chapter 3 we apply this array processor on an exploration scale in a time-lapse surface wave anisotropy study above and underground gas storage (UGS). Our results suggest potential applications for UGS reservoir monitoring and are in review for the *Journal of Geophysical Research*. In chapter 4 the array processor was applied in Southern California to produce estimates of azimuthal Rayleigh wave anisotropy with unprecedented resolution. The results of this chapter illustrate how ambient noise studies can take advantage of substantially more data points in shorter analysis periods compared to earthquake studies. The study is planned to be submitted to *Geophysical Research Letters*. Issues of statistical significance are elucidated in both chapters using bootstrapping and model selection techniques.

Finally, appendix A gives an application where statistics of ambient seismic power density levels are used to model detection thresholds in hydraulic fracture

monitoring. This application is partially related to this thesis and was originally reported by Goertz et al. [2012a].

Chapter 2

A statistical strategy for ambient seismic wave field analysis: investigating correlations to a hydrocarbon reservoir

Abstract Theoretical work and modeling studies have led to the hypothesis that the ambient seismic wave field on the surface can be affected by hydrocarbon reservoirs (>800 m depth). Several field studies have linked spectral features on the vertical component between 1-10 Hz to reservoir locations. However, such evidence has been criticized due to concerns that surface recordings typically contain a large amount of surface wave noise and correlations to hydrocarbon targets could be caused by non-hydrocarbon variables such as topography or weathering layer thickness. In this paper we suggest a two-step analysis strategy to address such issues. First, spectral power is only averaged over time periods and frequencies where the distribution of polarization attributes show no obvious dominance of a few surface wave sources. An interferometric test reveals differences in the wave field composition between the filtered and unfiltered data. Second, the residual seismic power is correlated to hydrocarbon as well as non-hydrocarbon targets. The correlations are quantitatively compared using rank correlation and bootstrap confidence intervals. The method is illustrated on a passive seismic data set acquired with three-component, broad-band seismometers at the tight-gas Jonah field in Wyoming, USA. We find evidence that the wave field was dominated by a small number of surface sources in all of the data except for the quietest time

periods in the low-frequency range 1.5-3.0 Hz. Seismic power within this subset significantly correlates to a published reservoir map but not with a digital elevation model and less so with an infrastructure density map. The investigated hypothesis can thus not be rejected with this data.

This chapter is published in Geophysical Journal International *(doi: 10.1093/gji/ggs031). Authors: Nima Riahi, Alexander Goertz, Bradley Birkelo, Erik H. Saenger.*

2.1 Introduction

The continuous ambient seismic wave field is a function of its distributed, uncontrolled sources as well as the earth medium through which their vibrations travel. The source component can in theory be separated from the medium information if the source distribution and source types have certain statistical properties. For uniform source distributions ensemble averages of noise cross-correlations between two signals can be interpreted as the Green's function between those receivers [Wapenaar, 2004, Draganov et al., 2006]. This is used in applications such as crustal tomography [Shapiro et al., 2005] or reflector mapping [Draganov et al., 2007, Ruigrok et al., 2011]. Another approach to approximately separate medium effects is the division of spectra from different components. The horizontal-to-vertical spectral ratio (HVSR) technique is a single-station method and therefore not constrained by receiver geometry. It is routinely used in earthquake engineering to characterize the near surface weathering layer [Bard, 1999, D'Amico et al., 2008, Fäh et al., 2003, Bard, 2010].

Based on previous empirical observations and theoretical studies by different investigators, Lambert et al. [2011b] recently formulated an explicit hypothesis that the ambient seismic wave field can be modified by hydrocarbon reservoirs at depth and that these modifications can be detected at the surface. Plausibility for this hypothesis is provided by two observations. First, there is ample evidence for continuous body waves reaching the surface [Roux et al., 2005, Koper & de Foy, 2008, Gerstoft et al., 2008b, Zhang et al., 2009, Koper et al., 2009, 2010, Landes et al., 2010, Poli et al., 2012b] which would have passed the reservoir. Second, hydrocarbon reservoirs are particular inhomogeneities at low-frequencies: active seismic experiments show evidence of increased attenuation [Chapman et al., 2006] as well as increased reflectivity [Goloshubin et al., 2006] and these phenomena might in fact be linked [Korneev et al., 2004, Quintal et al., 2011]. In a modeling study that combined small-scale poroelastic theory with large-scale viscoelastic simulation, Lambert et al. [2012] propagate upward traveling incoherent seismic energy through a hydrocarbon-saturated inclusion and observe that it correlates with spectral attributes on the surface. Saenger et al. [2009] suggest an alternative explanation where non-linear pore effects produce a secondary wave field of p-waves. Both explanations predict a change in body wave energy at the surface in an a priori unknown frequency band. Whether these mechanisms would lead to an

increase or decrease in ambient seismic amplitudes in the vicinity of a hydrocarbon reservoir is not obvious a priori.

A body of empirical evidence collected over the last decade seems to support the hypothesis. Some case studies describe observations of foci at depth obtained by applying the time-reversal principle to array observations of seismic noise [Steiner et al., 2008, Witten & Artman, 2011, Goertz et al., 2012b]. Most other observations focus on simpler single-station analyses that correlate lateral spectroscopic features with reservoir locations [Dangel et al., 2003, van Mastrigt & Al-Dulaijan, 2008, Saenger et al., 2009, Birkelo et al., 2010, Goertz et al., 2012b]. However, such empirical studies have been subject to considerable controversy [Lambert et al., 2009b, Green & Greenhalgh, 2009a, Lambert et al., 2009a, Green & Greenhalgh, 2009b]. Two concerns are commonly brought forth: first, surface waves often dominate the wave field and it has been suggested that they might completely mask potential body waves [Ali et al., 2010]. Second, correlations to a reservoir do not confirm the hypothesis if they are in fact caused by non-hydrocarbon variables that are collocated with the reservoir. For instance, in a modeling study Hestholm et al. [2006] found that seismic energy can get trapped in topographic ridges and Hanssen & Bussat [2008] observed a correlation between vertical seismic amplitudes and dune height in a passive seismic dataset from the Libyan Desert. Shallow geology also affects seismic surface recordings, a phenomenon exploited by the HVSR technique [Fäh et al., 2003, Bonnefoy-Claudet et al., 2006]. Yet another factor is the potential effect that the data acquisition schedule can have: certain areas might have been measured during a period where natural or manmade sources were more energetic, leading to an imprint of the acquisition pattern on the seismic attribute map. Weather conditions, in particular wind speed, are also relevant variables to consider in the 1-10 Hz frequency range [Withers et al., 1996].

The principal contribution of this paper is to suggest a spectroscopic analysis strategy for passive seismic data sets that objectively addresses the surface wave problem and the risk of correlations to non-hydrocarbon variables. In the first part we reject time periods and frequencies that exhibit relatively high seismic power and laterally anisotropic polarization since these are considered strong indicators for a few surface sources dominating the signal. The impact of surface waves on the analysis is thus reduced. This single receiver analysis does not aim to identify all surface waves but it is easily implementable and poses few processing related constraints on the survey design. An interferometric test confirms that surface

waves from a dominating source were removed this way. In the second part the residual seismic power is spatially correlated to a hydrocarbon as well as to several non-hydrocarbon targets. We will test if the correlation to the hydrocarbon target is significant and stronger than to the non-hydrocarbon targets. Since a linear relation cannot be assumed, and to avoid strong sensitivity to outliers, we use the non-parametric rank correlation in combination with bootstrap confidence intervals.

A first data example is given for a passive seismic data set acquired with three-component broad-band seismometers at the tight-gas Jonah field in Wyoming, USA. Seismic power from appropriate sections of the ambient wave field recordings is first qualitatively compared to the acquisition pattern and then quantitatively correlated to reservoir parameter maps published by DuBois et al. [2004], elevation from the National Elevation Dataset [Gesch, 2007], and an estimated infrastructure-density map. We find that the hypothesis as formulated by Lambert et al. [2011b] can not be empirically rejected with the available data.

2.2 Data example

The tight-gas Jonah field in Wyoming, USA, was discovered in 1977 and has since been well studied, with a large body of knowledge available in Robinson & Shanley [2004]. The principal hydrocarbon reservoir is the Lance formation which is comprised of meandering fluvial sandstones intercalated with overbank siltstones and mudstones. The field is bounded by shear faults on the West and South sides and these faults form the updip and lateral trap for the field. The top of the formation lies at 2.4 km depth and dips down towards the Northeast to 2.8 km. The gross thickness increases downdip from 610 m to 915 m [DuBois et al., 2004]. Faults inside the field compartmentalize the reservoir, leading to an overall heterogeneous distribution of hydrocarbons and pressures. Porosity ranges from 8-12% and permeability from 0.01-0.9 mD. The Jonah field is being actively developed and produced. Operation and production noise in the area include truck traffic, drilling, well fracture stimulations, and compressors. The treeless area has a relatively flat topography and is partly covered with low-growing sagebrush. The surface geology in the survey area is uniform with shales and marlstones of Eocene age [Love & Christiansen, 1985]. Statics from active seismic surveys in the area typically show small lateral variation which are often correlated to elevation.

In a passive seismic survey in December 2009, three-component, broad-band particle velocity seismometers (Nanometrics Trillium T40) were deployed at 235 sites over the Jonah and the neighboring Pinedale gas fields. The instruments were chosen because they are able to resolve ground velocity below the low-noise model of Peterson [1993] in the frequency range between ocean microseisms and anthropogenic noise. Since the ocean microseism amplitudes vary smoothly within the survey area they can be used to detect sensors which deviate significantly with their neighbors, either due to coupling issues or bad instrument settings. Three-component geophones could be used as well, provided that their noise floor is sufficiently low after instrument correction.

The survey was acquired in a roll-along fashion with about 60 seismometers operating synchronously for about two days per location, set at a sampling rate of 100 Hz. A total of 88 measurement locations were collocated with data points from reservoir parameter maps published by DuBois et al. [2004]. Figure 2.1 shows an aerial map of the survey area with the measurement locations. Temperatures during the survey were below freezing point and did not allow for a thawing of the top soil. There was no notable precipitation and wind speeds remained below 3 m/s during more than 95% of the acquisition period. The same dataset has previously been investigated by Birkelo et al. [2011].

2.3 Analysis strategy

For the purpose of this paper we formulate the hypothesis to be tested as follows: *the ambient seismic wave field at the Jonah field surface exhibits variations in body wave power that are due to the subsurface hydrocarbon reservoir.* The objective of this paper is to present a statistical approach with which this hypothesis can be tested.

The hypothesis explicitly considers body waves because only they have the potential to carry information on the deeper subsurface (>800 m) above 1 Hz. The analysis of such potential body wave power can be strongly biased by surface wave energy which often dominates the wave field during a significant fraction or perhaps even all of the recording period. In general, the explicit and unambiguous removal of such surface wave energy requires a full description of the wave field which in turn necessitates synchronous and non-aliased spatial sampling. Such acquisition

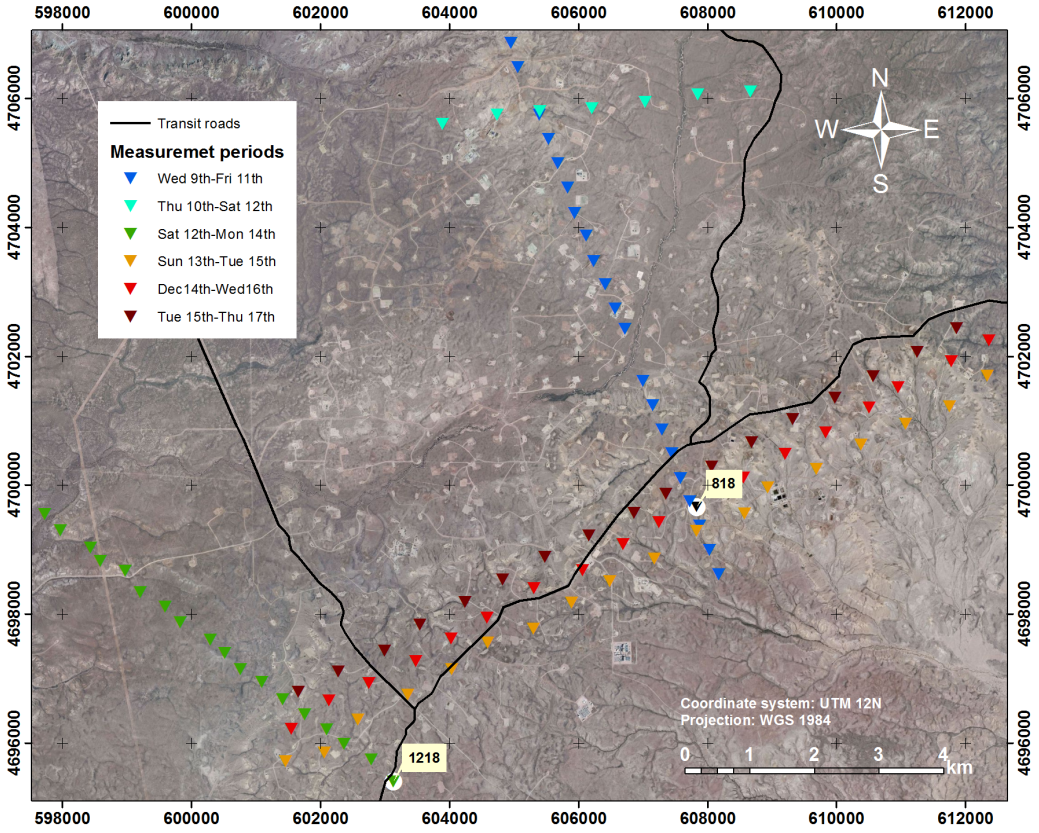

FIGURE 2.1: Aerial map of the survey area above the Jonah reservoir. The triangles show the sites where the broadband seismometers were deployed. The acquisition period is indicated by the color of the symbols. Recordings from the highlighted locations (1218,818) will be referred to later. The black lines show the transit roads in the survey area.

designs are hard to implement in a broad-band passive survey for practical reasons. Requirements on spatial sampling can be somewhat relaxed when analyzing move-out in cross-correlated receiver gathers. However, this happens at the cost of assuming an isotropic and/or stationary wave field, which can not always be guaranteed.

A less ambitious, but more readily implementable solution is to analyze single station polarization statistics to detect and reject time periods and frequencies where one or only a few surface sources dominate the signal. Under reasonable scattering conditions, the vibrations from a surface source seen by a surface receiver at a distance are bound to exhibit a degree of azimuthal directionality, independent of the mixture of wave types and modes involved. Even in the presence of a little more than one dominating source the distribution of polarization produced is likely to remain anisotropic. Because we deployed seismometers in a producing

and developing gas field, it can be expected that there are time periods at each recording site where a such a small number of surface sources will dominate, either by well pad activity or occasionally passing trucks. Clearly identifiable modes of the distribution of polarization angles over time are therefore a signature of time periods affected by one or a few dominant surface sources and should be rejected.

In the first step of the strategy, we therefore identify and select time periods and frequencies where the dominant polarization exhibits lateral isotropy, which is found to be the case for the quietest time periods of the recordings and at low-frequencies. This data selection strategy does not strictly achieve a complete removal of surface waves. However, by rejecting time periods and frequencies of high power and obvious azimuthal directivity, we preferentially reject surface waves from nearby surface noise sources and consequently increase the relative likelihood for the detection of body waves from other distant sources in the remaining data. Also, selecting the lowest power time periods of the wave field preferentially rejects transients (e.g. truck traffic) resulting in less time variability of the residual wave field. The seismic power density on the vertical component is then averaged over those time periods and frequencies only. Repeating the process for every measured location gives a map of seismic power density during the quiet time periods with no clearly distinguishable polarization. The concept of focusing on the quietest time periods was also applied by Hanssen & Bussat [2008] and has here been extended to additionally include the requirement for isotropic polarization.

In a second step, this seismic power density is now correlated to a hydrocarbon and as many non-hydrocarbon targets as are available for the measured locations. Correlations are quantified using a rank correlation that only measures monotonic trends without particular assumptions about the type of the relationship (linear or otherwise) and is also more robust against outliers compared to the linear correlation coefficient. Those correlations and their estimated uncertainties are then compared. The non-hydrocarbon variables considered are elevation, well density, and the acquisition pattern.

We test if the correlation to the hydrocarbon target is significant and also whether it is significantly stronger compared to correlations to non-hydrocarbon targets. If this is not the case, the hypothesis as stated above is rejected. Otherwise, the confidence in the hypothesis is increased. Note that empirical studies on observational data alone, which the ambient wave field inevitably is, can not be used to verify the causality as stated in the hypothesis.

2.3.1 Rejection of dominant surface waves

The entire three-component recording is split into non-overlapping small time segments of fixed length, as is usual for the Short-Time Fourier Transform [STFT Gabor, 1946]. We found a segment length of T=20.48 seconds to be short enough to ensure that the stationarity assumption is met within most of those segments, yet long enough to provide reasonable estimates for frequencies down to 0.1 Hz. A 42.5 hour recording thus results in about 7500 time segments. Each of the three components of a segment starting at time t is made to have zero mean, has any linear trend removed (detrending), is tapered with a hanning taper of the same length as the segment, and finally transformed into the complex Fourier domain with the Fast Fourier Transform [Press, 2007]:

$$\mathbf{U}(f,t) = [U_E(f,t), U_N(f,t), U_V(f,t)]^T \,, \tag{2.1}$$

where $U_{E,N,V}(f)$ is the Fourier amplitude as a function of frequency, f, for the East, North, and Vertical seismometer component, respectively. The segment size of 20.48 seconds affords a frequency resolution of $1/T$=0.049 Hz from 0 to 50 Hz. From the Fourier vector $\mathbf{U}(f,t)$ we now estimate the cross-spectral density (CSD) matrix,

$$\hat{\mathbf{S}}(f,t) = k \cdot \mathbf{U}(f,t) \cdot \mathbf{U}^{\dagger}(f,t) \tag{2.2}$$

$$k = \frac{2}{f_s \cdot \sum_{\tau=0}^{n-1} |w(\tau)|^2} \,, \tag{2.3}$$

where f_s = 100 Hz is the sampling rate and $w(\tau)$ is the hanning taper applied before the Fourier transform. In this normalization the diagonal elements of $\hat{\mathbf{S}}$ are estimated power spectral densities (PSD) of the three components and the off-diagonal elements are their cross spectral densities. To reduce the statistical noise inherited from the Fourier amplitudes we compute a weighted time average of eleven adjacent spectral density matrices, with maximum weight for the center segment, t, and gradually decreasing weights to the flanks. This approach is similar to the commonly used block-averaging technique [Press, 2007].

The dominant particle motion at a given frequency and time segment is extracted by the eigenvector of $\hat{\mathbf{S}}(f,t)$ with the largest eigenvalue [Samson, 1983, Park et al., 1987]:

$$\hat{\mathbf{S}}(f,t) \cdot \mathbf{z}_{\max}(f,t) = \lambda_{\max}(f,t) \cdot \mathbf{z}_{\max}(f,t)\,. \tag{2.4}$$

For a normalized eigenvector the eigenvalue $\lambda_{\max}$ is an estimate of the seismic power density of the dominant polarization. Figure 2.2a shows $\lambda_{\max}$ as a function of frequency, f, and time, t, for site 1218 (location highlighted in Figure 2.1). From acquisition reports as well as from a more detailed analysis of the data (not shown) we know that the recording contains truck traffic, an M4.1 earthquake in the Gulf of California, and stationary and non-stationary noise from well operations at less than 120 m distance (hydraulic fracturing, perhaps also well drilling, fluid production or injection). Figure 2.2b gives the time series of $\lambda_{\max}$ at f=2.15 Hz for this site. The visible power density variations arise because the above mentioned sources may contribute different amounts of energy at different times into the 2.15 Hz bin. Figure 2.2c compares the power spectral density during hand-picked periods where different noise sources were active in the 1218 recording. The large dynamics (power variations up to 50 dB) and frequency dependence of the ambient wave field are illustrated. Nearby road traffic clearly dominates other noise sources. Figure 2.3 gives five minute examples of the three-component seismograms for some of the features highlighted in Figure 2.2a. The shown seismograms are typical for the measurements on the Jonah field. Visual inspection of the seismograms in Figure 2.3b,c,d confirms the stationarity assumption to the first order. The truck traffic shown in Figure 2.3a is clearly non-stationary on a minute-scale, but the transient nature of such sources makes them relatively easy to identify.

The three complex elements of $\mathbf{z}_{\max}(f,t)$ give the amplitudes and phases of three sinusoidal oscillations in the E, N, and V directions with the same frequency. A sketch of the resulting elliptical motion is shown in Figure 2.4. The semi-major axis, $\mathbf{a}$, and the semi-minor axis, $\mathbf{b}$, are found by maximizing the length of $\mathrm{Re}[\mathbf{z}\cdot e^{i\xi}]$ using a grid search over $\xi \in [0,\pi)$:

$$\xi_0 = \max_{\xi\in[0\,\pi)} \mathrm{Re}\,[\mathbf{z} \cdot \exp(i\xi)] \tag{2.5}$$

and then setting [Vidale, 1986]

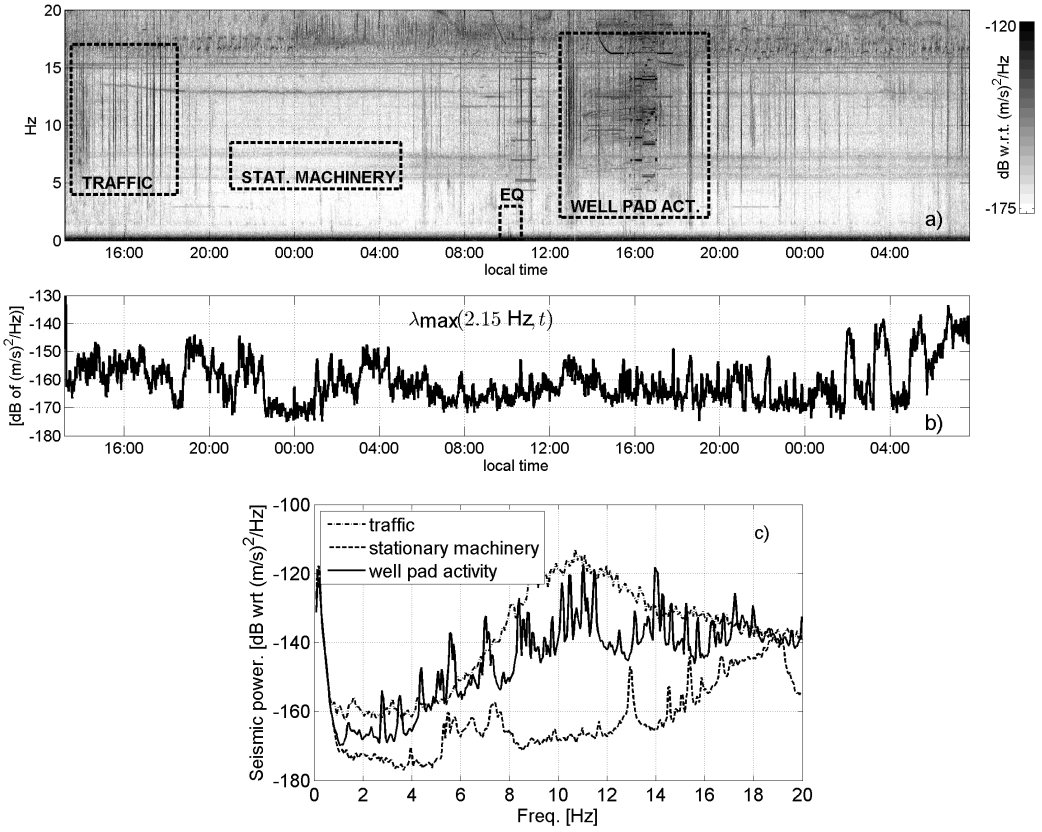

FIGURE 2.2: *a)* Seismic power density of the dominant polarization, $\lambda_{\max}$, as a function of frequency and time for the entire recording at site 1218, which was highlighted in Figure 2.1 (dark shades indicate higher power). Dashed boxes from left to right identify truck traffic, stationary industrial noise, a M4.1 earthquake from the Gulf of California, and industrial well pad activity. *b)* Time series of $\lambda_{\max}$ at the single frequency bin 2.15 Hz. *c)* Comparison of power spectral density during hand-picked time periods where different sources dominated.

$$\mathbf{a} = \mathrm{Re}[\mathbf{z} \cdot e^{i\xi_0}] \,, \tag{2.6}$$

$$\mathbf{b} = \mathrm{Re}[\mathbf{z} \cdot e^{i(\xi_0 + \pi/2)}] \,. \tag{2.7}$$

We now define three features of the polarization. The azimuth, $\phi \in [-180°, 180°]$, as the angle between North and the projection of the upward pointing semi-major axis, $\mathbf{a}$, to the horizontal plane, measured clockwise. The dip, $\theta \in [0°, 90°]$, is defined as the angle that the upward pointing semi-major axis spans with the horizontal plane (see Figure 2.4). Finally, the reciprocal ellipticity $\rho \in [0, 1]$ is defined as $|\mathbf{b}|/|\mathbf{a}|$. Linear and circular polarization are represented by ρ=0 and

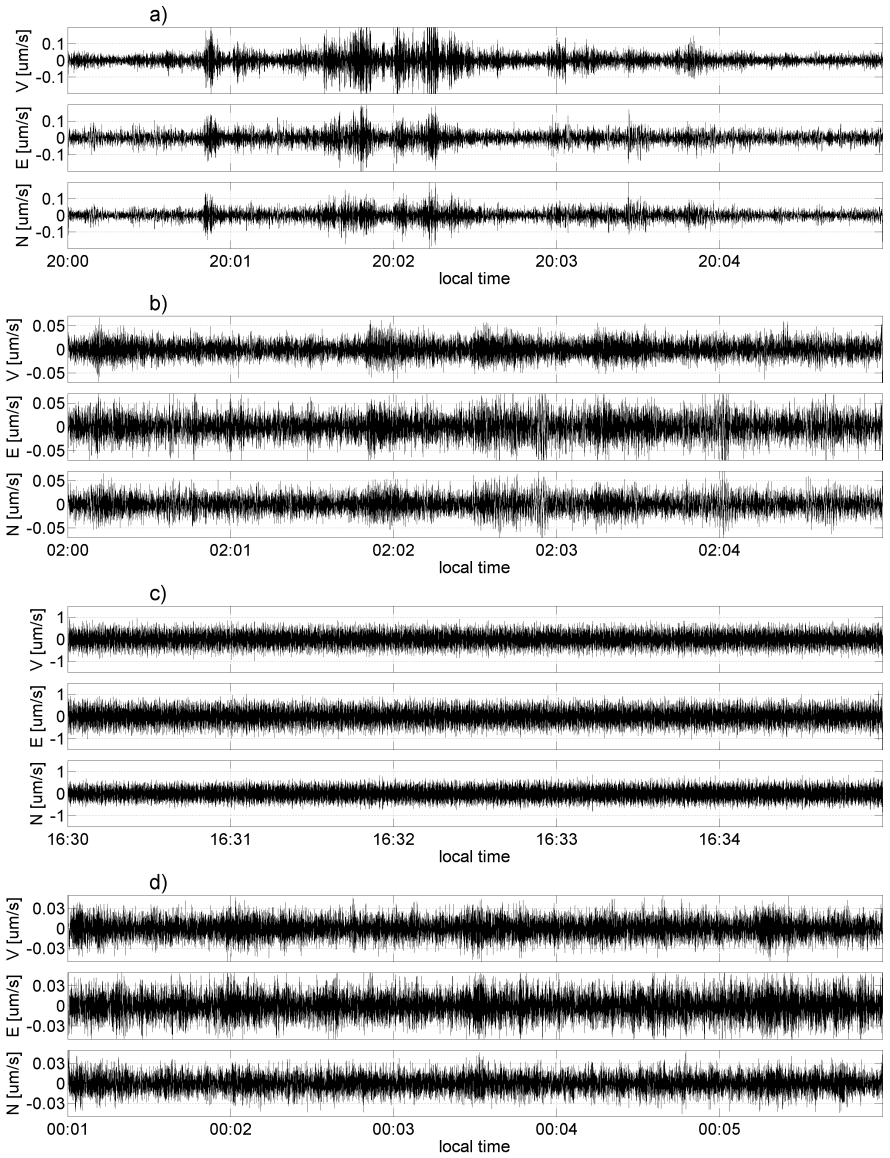

FIGURE 2.3: Five minute seismograms of the processes highlighted in Figure 2.2a. V, E, N refer to the vertical, East, and North components, respectively. A bandpass filter from 1-10 Hz was applied to all signals. *a)* Wave train generated by a truck passing outside the local surroundings (bandpass 3 to 10 Hz). *b)* Stationary background noise generated by machinery. *c)* Stationary background noise generated by well pad activities probably within a 300 m perimeter. *d)* Unidentified stationary background noise. Note the different scales of the vertical axes.

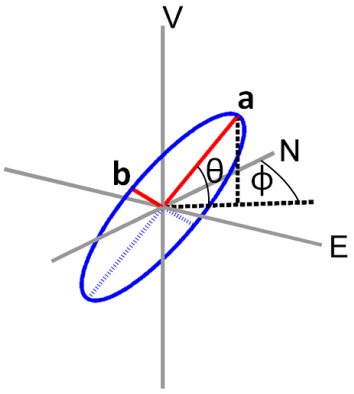

FIGURE 2.4: The dominant particle motion is estimated at each frequency by a polarization ellipse. This plot shows the semi-major axis, **a**, the semi-minor axis, **b**, the azimuth angle, ϕ, and the dip angle, θ.

ρ=1, respectively, while values in between imply elliptic motion. These three features are computed for all frequencies and all time segments.

For each frequency bin, f_0, we now take the 10% of time segments with the weakest seismic power density. This means that from the power density estimates at f_0 of the full recording we only consider those below the 10^{th} percentile, which for the Jonah recordings corresponds to an effective duration of 4.2 to 4.8 hours. Based on these segments we compute a histogram of the azimuth with equal-sized bins from $-180°$ to $180°$. The histogram values are normalized such that they integrate to one, whereby we get the empirical probability density function (PDF) of ϕ at frequency f_0 for the quietest time periods. Repeating this process for all frequencies gives a spectrum of such PDFs. Figure 2.5a shows the result for the frequency range 0 to 20 Hz. For frequencies above 3 Hz the plot reveals a mostly bimodal distribution of azimuth which changes slowly with frequency (note that the PDFs are periodic in azimuth). The modes are separated by about $180°$, as would be expected for surface waves from a distinct direction. This laterally anisotropic polarization property strongly suggests that most of the frequency band above 3 Hz is dominated by surface waves from one or only a few sources. Below 3 Hz these modes disappear and give way to a flat distribution, making this band very unlikely to be dominated by a few sources. Figure 2.5b gives the same visualization for polarization dip. Dark shades at low angles indicate dominant horizontal particle motion. Frequencies below 0.8 Hz show such dominant horizontal motion, probably due to oceanic microseisms. A concentration of low dips is also visible over most of the frequency band from 4 to 16 Hz, but less so for the frequency range 1 to 3 Hz. Figure 2.5c gives the

same visualization for reciprocal ellipticity. Except for some narrow-bands, the frequency range from 4 to 14 Hz exhibits polarizations far from linear. Between about 1 and 3 Hz, however, the ellipticites indicate a more isotropic particle motion. Finally, Figure 2.5d shows the visualization for power density. There is a low-power regime in the frequency range bounded by the oceanic microseisms around 1 Hz and noise processes above 5 Hz.

The above features are specific to the low-power time segments. Figure 2.6 gives PDF spectra of polarization parameters and seismic power density when considering time periods with $\lambda_{\max}$ values below the 50^{th} percentile level. Note how the azimuth PDFs for this segment group have stronger modes which now extend down to 0.5 Hz, indicating that also the low frequencies are now dominated by a small number of surface sources even though we still reject 50% of the data. Dominant horizontal motion is observed up to 15 Hz, including the frequency band 0.8-4 Hz, which showed considerably lower probability density values during the quietest 10% of the data, indicating that the signal now contains more horizontally polarized energy. Also, the reciprocal ellipticity in the band 1 to 3 Hz is less random than for the quiet subset of the data with a bias towards circular polarization.

Using a random number generator we produced three independent time series of 48 hours of Gaussian, white noise and used them jointly as an isotropic three-component test signal. Figure 2.7 gives the polarization PDFs of this synthetic noise. Comparing this plot with Figure 2.5a,b,c qualitatively confirms that the wave field is roughly isotropic between 0.5 and 3 Hz, and anisotropic almost everywhere else.

Plots as in Figures 2.5 and 2.6 were analyzed for all 85 measurement locations with valid three-component data within the survey area. We consistently observe such a pattern in the Jonah field: during the 10% quietest time periods the ambient wave field polarization below 3 Hz exhibits lateral isotropy. For less quiet time periods this observation did not hold reliably. Note that the value of 10% is data-driven and should be reevaluated for other data sets. For some receivers, the microseism energy visible at the lower end of the spectrum in Figure 2.5d occasionally extended up to 1.5 Hz. To avoid such interference we therefore compute the average seismic power density between 1.5-3.0 Hz, considering only the 10% quietest time periods at every frequency bin. The resulting power density is denoted by $\bar{P}$ and computed for every location.

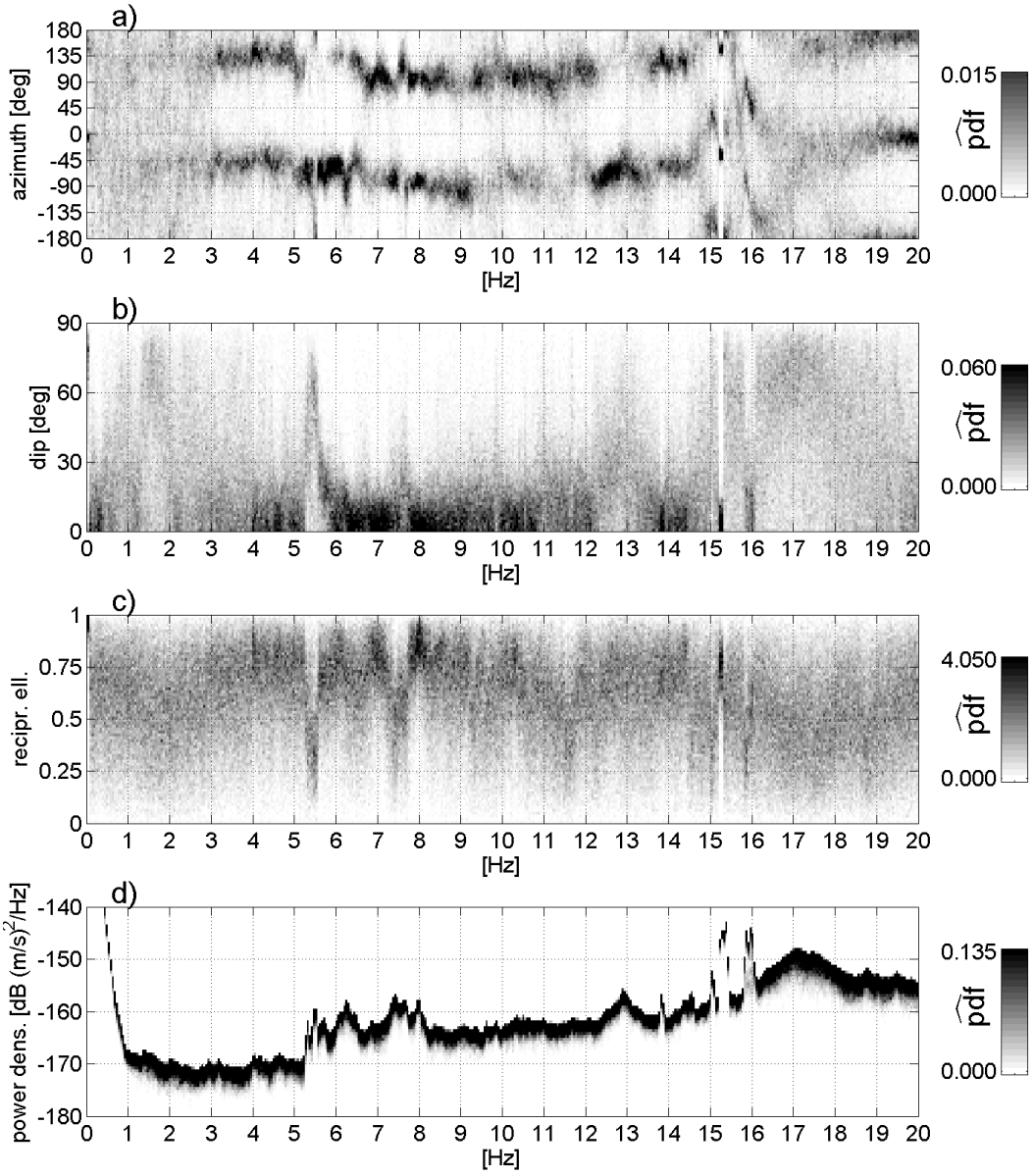

FIGURE 2.5: *(a)* Spectrum of azimuth PDFs from 0 to 20 Hz from site 1218, based on time segments with power density below the 10^{th} percentile of all estimates. Below 3 Hz the PDFs are almost flat. *(b)* Same visualization for polarization dip. Below 1 Hz there is an increase in horizontal motion (dip<30°), probably due to oceanic microseisms. *(c)* Same visualization for reciprocal ellipticity (0: linear polarization, 1: circular polarization). *(d)* Same visualization for the power density.

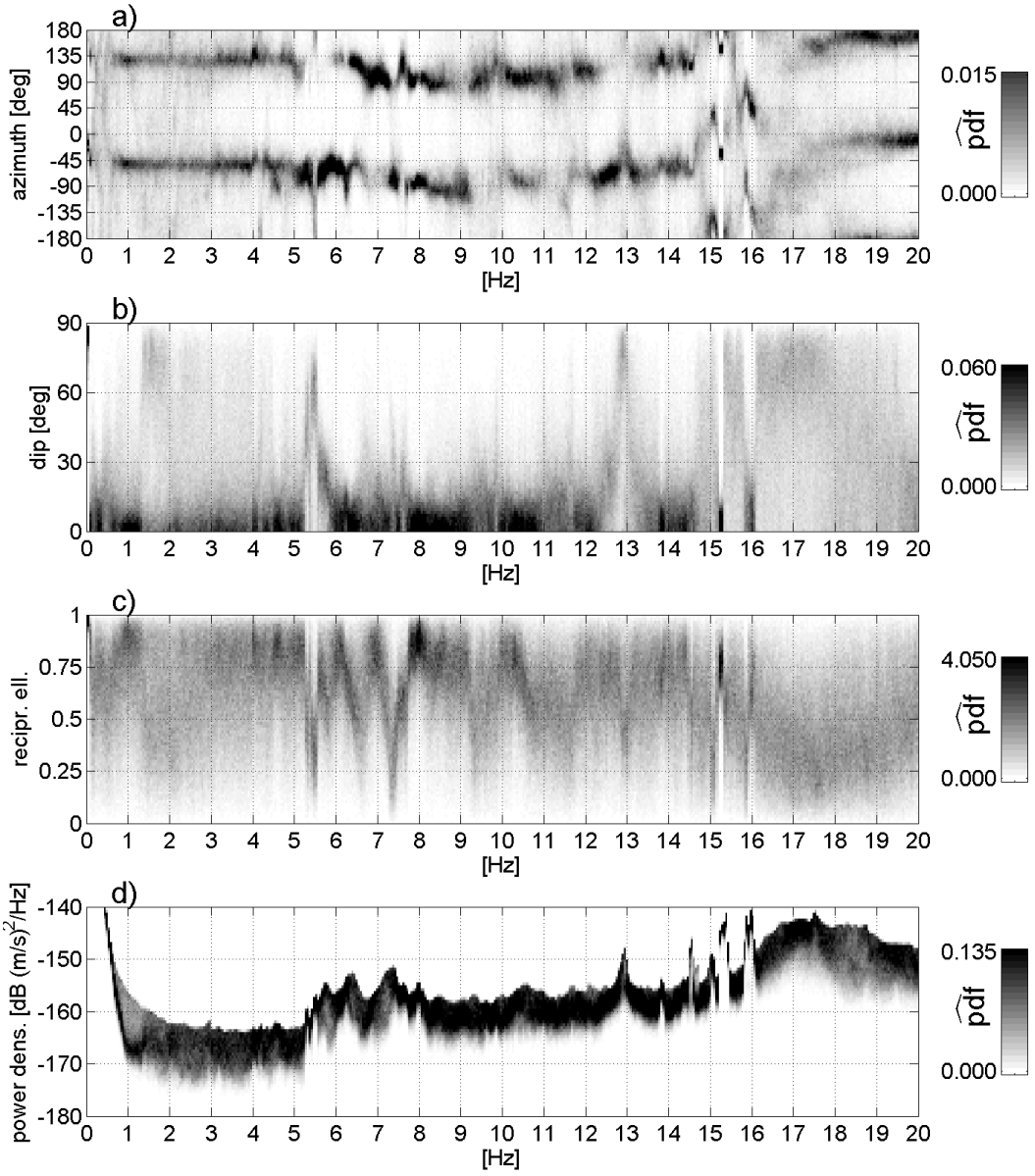

FIGURE 2.6: *(a)* Spectrum of azimuth PDFs from 0 to 20 Hz from site 1218, based on time segments with power density below the 50^{th} percentile of all estimates. The PDFs show distinct modes down to almost 0.5 Hz. *(b)* Same visualization for polarization dip. Horizontal motion dominates up to about 15 Hz. *(c)* Same visualization for reciprocal ellipticity (0: linear polarization, 1: circular polarization). The readings vary strongly but there is a trend towards more circular-like polarization. *(d)* Same visualization for the power density.

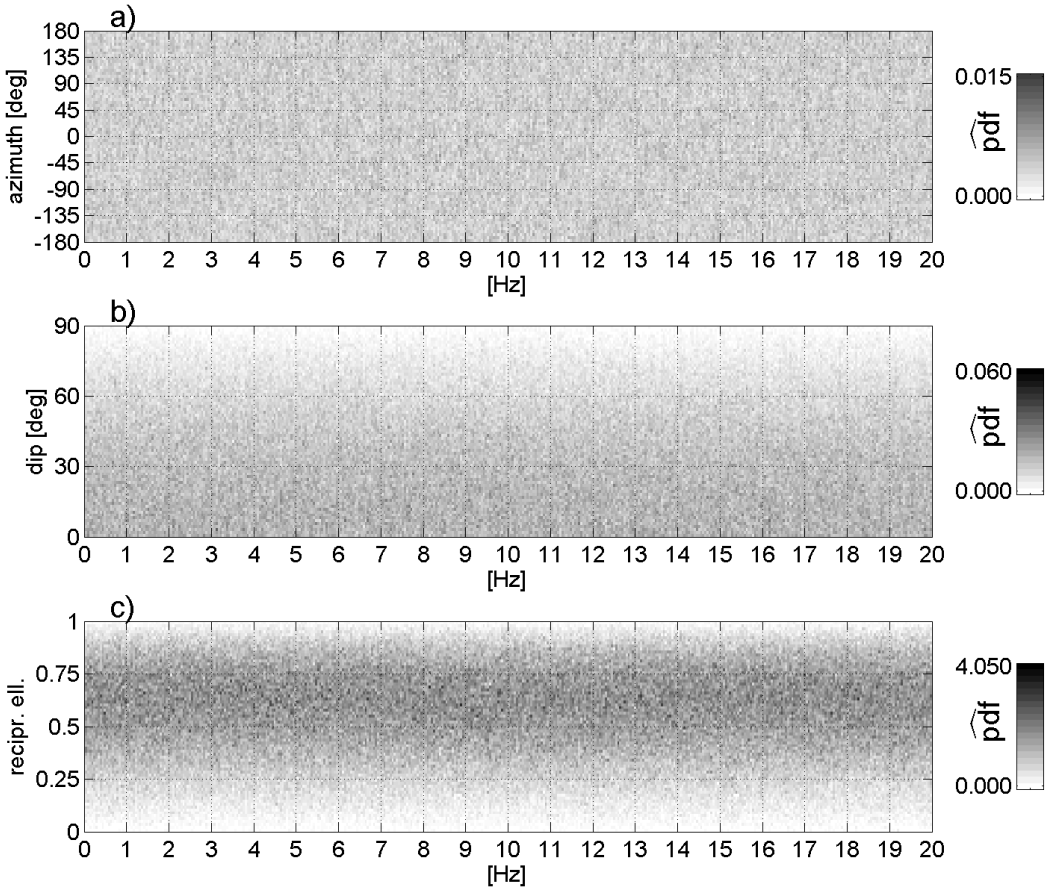

FIGURE 2.7: *(a)* Spectrum of azimuth PDFs from 0 to 20 Hz from Gaussian white, isotropic noise based on data of equal length as in Figure 2.5. *(b)* Same visualization for polarization dip. Note that the PDF scales as $\cos(\theta)$. *(c)* Same visualization for reciprocal ellipticity. The isotropic PDF is slightly skewed towards 1, with a mean value of ρ=0.57.

The focus on quiet time periods with lateral isotropy is a necessary precaution to avoid obvious surface waves but will not necessarily guarantee an increase of body wave energy relative to surface wave energy. Nevertheless, we tried to qualitatively assess the effect of the procedure on surface wave energy by analyzing a line of receivers stretching Northwestwards from station 1218 using the noise cross-correlation technique [Wapenaar, 2004, Snieder et al., 2010]. An active well pad is located along that line, close to station 1218. Therefore, the line can be considered to have a source in its stationary zone, ensuring an incident wave field coverage suitable for interferometry despite the anisotropy in the wave field.

We compute average cross-correlations of the vertical component of receiver 1218 with itself and all Northwestward receivers, up to a distance of 7 km. A two-pole, zero-phase bandpass filter between 1-3.5 Hz was applied to the raw data, which

was then split into segments of 13 sec duration for cross-correlation. Figure 2.8a shows the averaged trace-normalized cross-correlations during a day time period. A coherent wave train with linear move-out at apparent velocity of 1.0 km/s is detected. This is consistent with a surface wave that is traveling through station 1218 to all other stations. The fact that this wave train has no correspondence at negative times confirms that the well pad behind station 1218 is the main source of the energy and that there is no energy arriving from the opposite side of the line where also no infrastructure is found. Figure 2.8b shows an identical analysis on a quiet night time period where the 1-3.5 Hz range showed isotropic polarization. No coherent surface wave train is visible. It is therefore clear that the selection of the quietest 10% of the data can provide a means to reduce surface waves from the data.

The presented cross correlation test merely serves as an independent confirmation of the effectiveness of the data selection strategy with respect to reducing the influence of surface waves on our data. Owing to the observed anisotropy of the incident wave field over the used time span, it does not provide any additional discriminant between surface and body waves. Indeed, noise cross-correlations in Jonah often did not converge within the available recording period. A coherent event with near zero lag is visible up to about 4.5 km in Figure 2.5a. It might have been generated by broad-side arrivals of surface waves, although a clear polarization from that direction is missing in Figure 2.5a. An alternative explanation would be near-vertically traveling teleseismic body waves, e.g., of oceanic origin [Gerstoft et al., 2008b, Zhang et al., 2009, Landes et al., 2010, Poli et al., 2012b].

Furthermore, the data selection method by means of the power percentile does not ensure synchronous time windows over several locations and operates on a per-frequency basis. The cross-correlations of the quiet four hour nighttime window shown in Figure 2.8b are therefore only approximating the actual data that is used to compute $\bar{P}$.

2.3.2 Quantitative comparison of correlations

Figure 2.10 shows the average seismic power density, $\bar{P}$, as computed in the previous section (shaded circles). The measured $\bar{P}$ does not correlate qualitatively to the acquisition schedule indicated in Figure 2.1. The sites with high $\bar{P}$ values

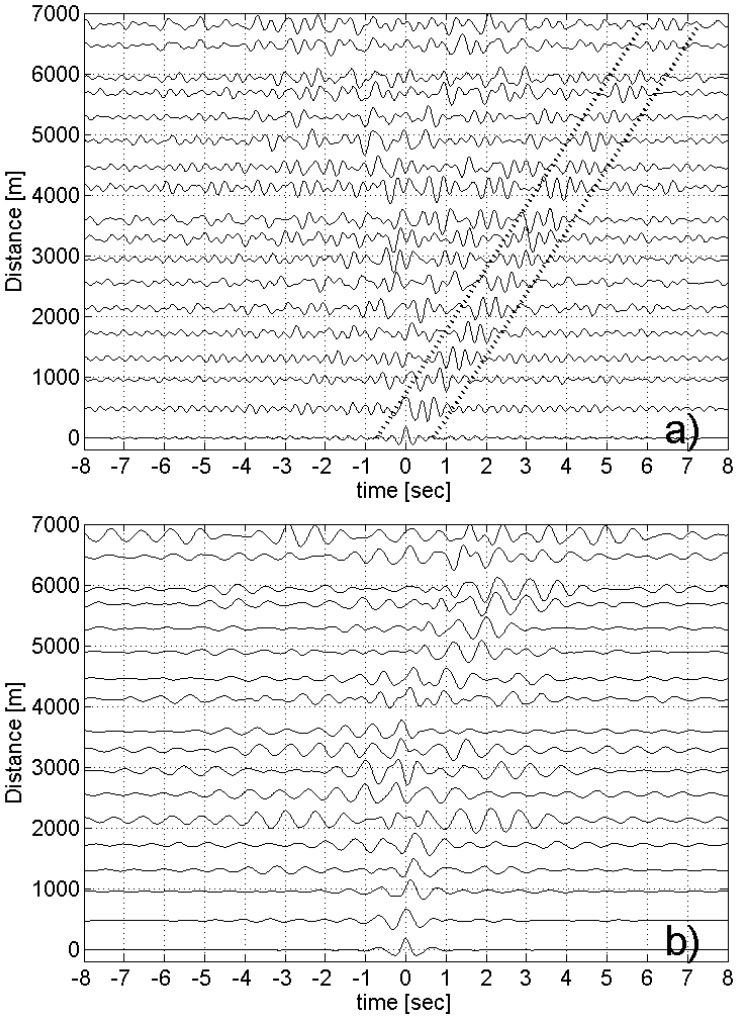

FIGURE 2.8: *(a)* Average noise cross-correlations of the vertical component of receiver 1218 with 18 receivers stretching out 7 km Northwestwards from station 1218. All traces were normalized to unity. This panel shows daytime recordings (1500h-1700h local time) which were bandpass filtered between 1-3.5 Hz. The traces are normalized and centered at the distance of the receiver to station 1218. The dotted lines highlight a coherent wave train with a linear move-out at apparent velocity 1.0 km/s, most probably caused by a surface wave. *(b)* Same as (a) but using night time recordings (0100h-0500h local time) where polarization was laterally isotropic. No coherent move-out can be detected.

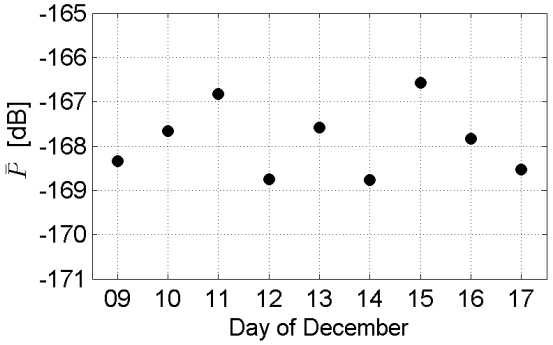

FIGURE 2.9: Average seismic power during the quietest 10% of time periods during consecutive 24 hour segments of the recording from reference station 818. The temporal variations are within a range of roughly 2 dB.

where the North-South line and the three East-West lines to the South cross were acquired within a time span of seven days. Sites in the Southwest corner of the area where the lowest values of $\bar{P}$ occur were acquired over a time span of five days including both a weekend and the beginning of a workweek. On the North-South line there are two station pairs that are separated by 60 m (Northern part) and 85 m (Southern part) which show seismic power variations of 2 dB and 1.7 dB, respectively. Figure 2.1 shows the location of reference station 818, which recorded during nine days. We computed the average seismic power during the 10% quietest periods for each day individually and plot the resulting time series in Figure 2.9. The power variations are contained within a range of about 2 dB, which is of the same order of differences observed at the station pairs near the line crossings. It is unclear to what extent the observed variations between the lines are a short-range phenomenon or a day-to-day effect. However, in relation with the total attribute dynamics of over 10 dB, this variation is modest. The spatial consistency of $\bar{P}$ is another indication that the natural background wave field during the acquisition period did not change significantly over time.

We now correlate $\bar{P}$ with one hydrocarbon and two non-hydrocarbon targets and compare those correlations. The first target is a reservoir hydrocarbon pore thickness ($S_g\phi H$) map in feet adapted from DuBois et al. [2004], also shown in Figure 2.10. This variable is a multiplication of estimated porosity, thickness, and hydrocarbon saturation of the reservoir layer and represents a relevant hydrocarbon target. The second target is an elevation map in meters retrieved from the National Elevation Dataset of the U. S. Geological Survey [Gesch, 2007], shown in Figure 2.11. This map will be used to test for near-surface geology

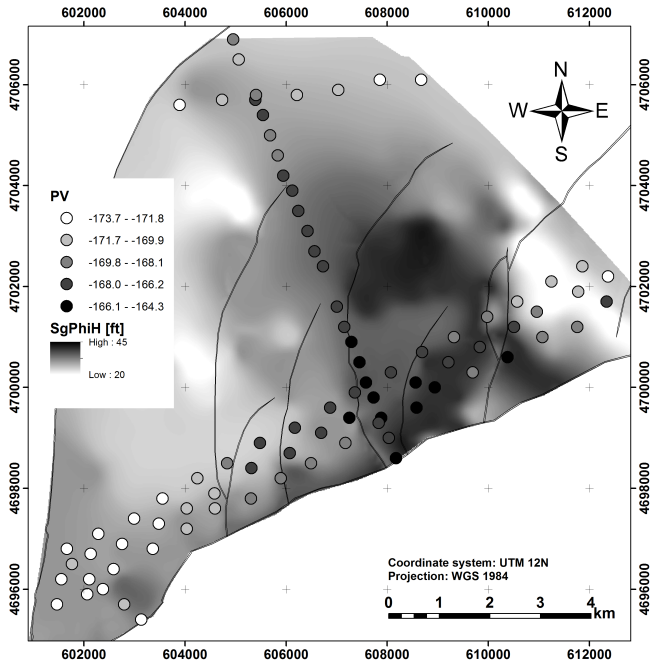

FIGURE 2.10: Contoured map of hydrocarbon pore thickness, $S_g\phi H$, of the Jonah reservoir adapted from DuBois et al. [2004]. The shaded circles show $\bar{P}$, the average seismic power density within 1.5-3.0 Hz during the quietest 10% of the recording [dB w.r.t. $(\text{m/s})^2/\text{Hz}$].

influence. The rationale behind this choice is given in the discussion. Third, a well pad density map shown in Figure 2.12 is used. The latter target is unitless and used as an infrastructure noise proxy where we considered well pads as potential surface wave sources that could radiate seismic energy within a range of about 500 m at about 2 Hz. The radius of 500 m is based on approximate observations of noise decay from machinery noise and the interferometry test. The map was modeled by placing two-dimensional Gaussian functions with standard deviation 250 m at all well locations and summing them up. Well pad locations as of 2009 were provided by field operators. Strictly speaking, the radius is a function of well pad type, noise environment, and near surface geology, among other factors. The map we use here is therefore primarily interpreted as representing a well density rather than expected seismic noise power due to the wells.

Figure 2.13a shows $\bar{P}$ plotted against total hydrocarbon pore thickness, $S_g\phi H$. The crossplot exhibits some degree of non-linearity, perhaps two separate trends (dashed lines). Quantifying the relationship with the conventional linear Pearson correlation coefficient might thus give misleading results. Instead, we use the non-parametric Spearman rank correlation coefficient ρ_{rank} [Spearman, 1904]. This coefficient quantifies to what degree a monotonic relationship exists between two

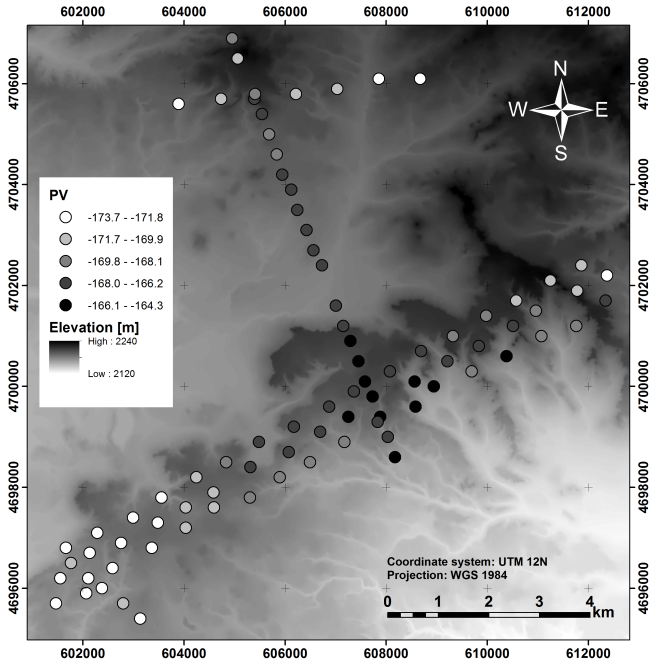

FIGURE 2.11: Elevation map at the Jonah field from the National Elevation Dataset [Gesch, 2007]. The shaded circles show $\bar{P}$, the average seismic power density within 1.5-3.0 Hz during the quietest 10% of the recording [dB w.r.t. $(\mathrm{m/s})^2/\mathrm{Hz}$].

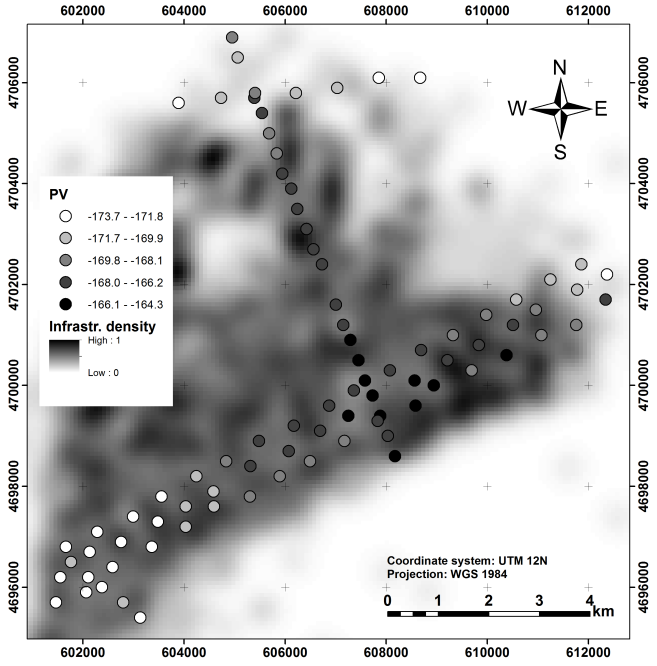

FIGURE 2.12: Well pad density map at the Jonah field. This map is used as a proxy for anthropogenic noise. The shaded circles show $\bar{P}$, the average seismic power density within 1.5-3.0 Hz during the quietest 10% of the recording [dB w.r.t. $(\mathrm{m/s})^2/\mathrm{Hz}$].

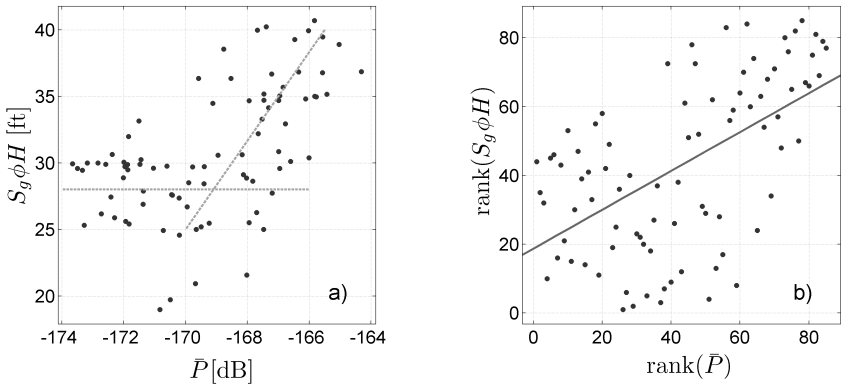

FIGURE 2.13: *(a)* The hydrocarbon target values $S_g\phi H$ are plotted against $\bar{P}$. The scatter exhibits a slight non-linearity, potentially two separate trends (dashed lines). The relation should not be described with the linear correlation coefficient. *(b)* The ranks of $S_g\phi H$ are plotted against the ranks of $\bar{P}$. The rank correlation in this case is the linear correlation in the ranks and quantifies the degree of monotonic relation between the two variables (gray line is best linear fit).

variables, independent of the type of that relationship. It is also more robust with respect to outliers. It ranges from -1 (monotonically decreasing) to $+1$ (monotonically increasing) with values close to zero indicating no monotonic relation between the variables. To compute this measure the values of a variable are replaced by their rank relative to the other values in that variable: the lowest value is replaced by 1, the second lowest by 2, and so on. In Figure 2.14a (left) the such transformed versions of $\bar{P}$ and $S_g\phi H$ are plotted against each other. The rank correlation is the conventional correlation coefficient computed on the ranks. The gray line in the panel shows the corresponding trend.

To include the uncertainty in the analysis, we do not estimate a single coefficient, but rather a confidence interval within which ρ_{rank} lies. This is attained by repeatedly estimating ρ_{rank} based on 1000 bootstrap resamples of the scatter points. From this resampling distribution a bias-corrected and accelerated 90% bootstrap confidence interval is computed [Efron & Tibshirani, 1993].

Figure 2.14a (right) shows the resulting distribution of ρ_{rank}. There exists a significant positive relation between $S_g\phi H$ and $\bar{P}$, as testified by the 90% confidence interval $\rho_{\text{rank}} \in [0.43, 0.67]$. Figure 2.14b gives the same plots for $\bar{P}$ and elevation, with $\rho_{\text{rank}} \in [-0.16, 0.24]$. Note that the confidence interval contains zero and therefore no significant monotonic relation can be claimed at the 90% level. Figure 2.14c shows the plots for $\bar{P}$ and well density, with $\rho_{\text{rank}} \in [0.06, 0.38]$.

There is a weak relationship, but one that is significantly weaker than the one to the hydrocarbon target.

Since the hydrocarbons in Jonah are mined we expect a certain correlation between the hydrocarbon target and infrastructure. In our data example the 90% confidence interval of the rank correlation between $S_g\phi H$ and well density $\rho_{\text{rank}} \in [0.31, 0.62]$. The fact that the attribute correlates better to the reservoir than it does to the confounder, while the confounder itself correlates well to the reservoir suggests that there is a different nature between these correlations. Specifically, the low $S_g\phi H$ values in the Southwest corner are captured better by the attribute, while contrasting with existing well pads there.

The above analysis was restricted to the quietest 10% of the recording durations. We also investigated how the rank correlations change as we used different amounts of data, from 2% to 100%. Figure 2.15a shows the resulting rank correlations color-coded as a function of frequency and percent of data used. The color-scale is white for insignificant correlation at the 90% significance level, which for 85 data points corresponds to $|\rho_{\text{rank}}| < 0.18$, and gradually saturates to blue or red for $\rho_{\text{rank}} \geq 0.55$ and $\rho_{\text{rank}} \leq -0.55$, respectively (strongest observed absolute correlations). Strongest correlations exist in the frequency range 1.5-3.0 Hz for up to about 50% of the data being used. Note that when using more than 10% of the data the necessary isotropy is not present. Weaker correlations also occur between 6-15 Hz. Correlations to elevation (Figure 2.15b) are mostly insignificant. Significant correlation to well density is observed in the frequency range 6-12 Hz (Figure 2.15c) and corresponds to the weaker correlations to $S_g\phi H$ in the same band, confirming a common concern that correlations to reservoirs are confounded by infrastructure noise. However, in the 1-4 Hz range, where best correlation to the reservoir is observed, there is no correlation to well density. All plots in Figure 2.15 use the same color-scale.

2.4 Discussion

We have computed the average seismic power density of the ambient wave field polarization at the Jonah tight-gas field in the frequency range 1.5-3.0 Hz, restricted to the 10% quietest time periods. This subset has low-power, is laterally isotropic and is likely to represent the isotropic, quasi-stationary background wave

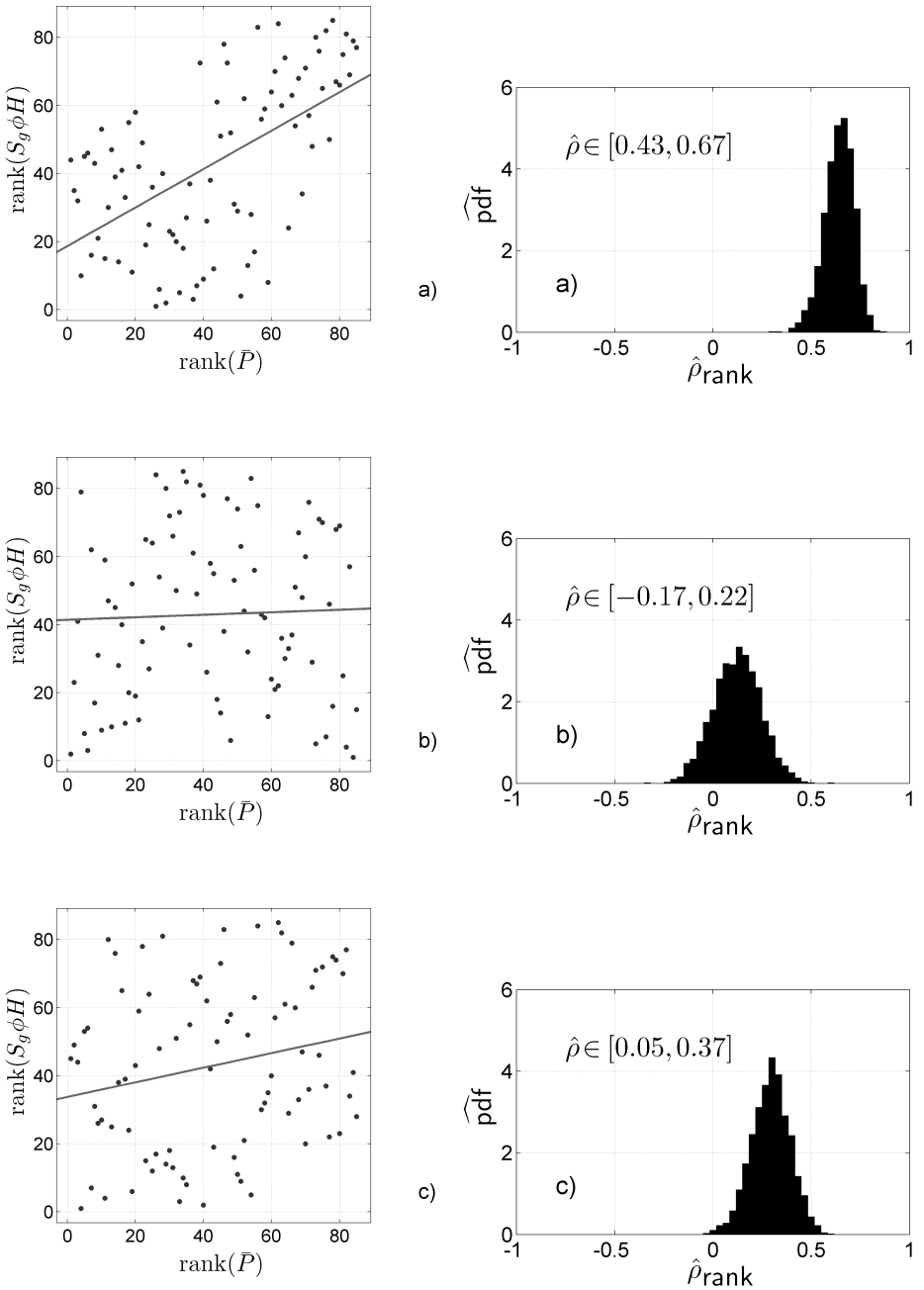

FIGURE 2.14: *(a)* Ranks of $S_g\phi H$ plotted against ranks of $\bar{P}$ (left) and bootstrap distribution of the estimated rank correlation coefficient ρ_{rank} (right). The 90% confidence interval for the rank correlation is given in the right panel. *(b)* Same plots for elevation against $\bar{P}$. *(c)* Same plots for well density against $\bar{P}$.

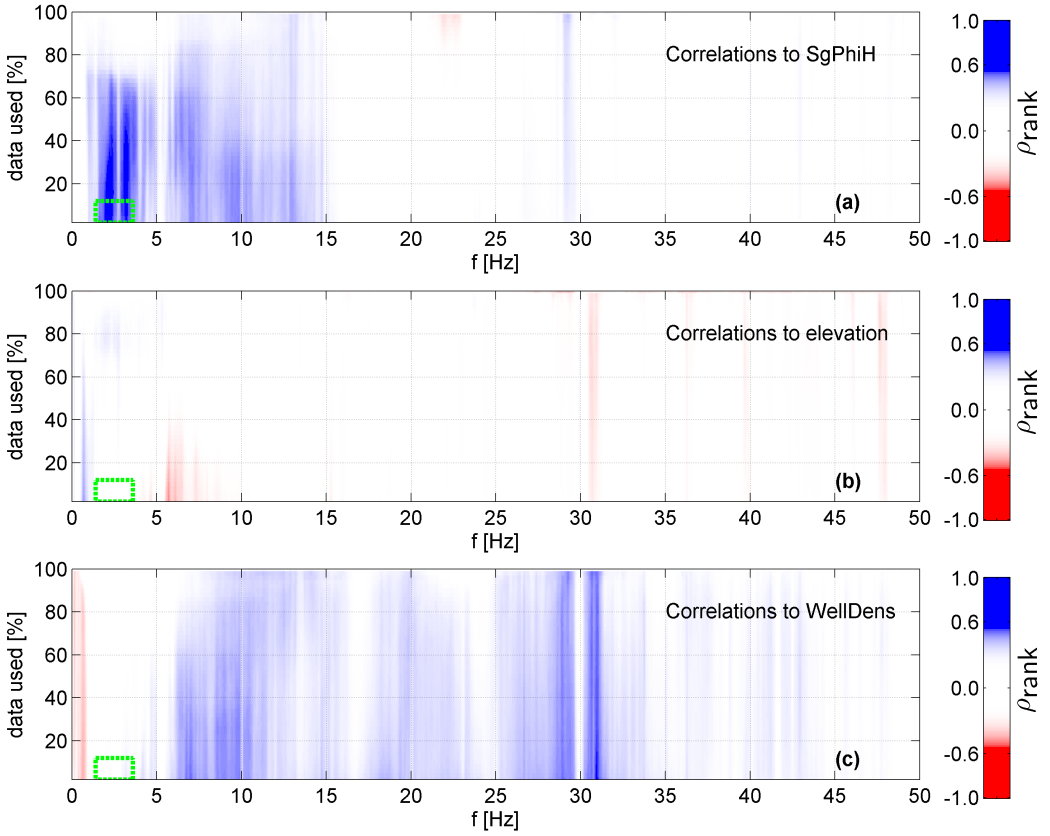

FIGURE 2.15: *(a)* The seismic power density at different frequencies is averaged over progressively larger percentages of the recording and then correlated to $S_g\phi H$. The graph shows the rank correlation coefficient color-coded as a function of frequency and percentage of data used. The colormap is white where the correlation is insignificant. The dashed green box highlights the data that was used in Figures 2.13 and 2.14. *(b)* Same visualization showing rank correlations to elevation. *(c)* Same visualization showing rank correlations to estimated well density.

field that consists of a mixture of body waves and surface waves originating from an apparently isotropic source distribution. Although the ratio of surface to body waves is unknown, filtering the wave field in this way increases the chance of uncovering weak components that may carry information about the deeper subsurface. An interferometric test independently confirms that the data selection strategy can indeed lead to a significant attenuation of surface waves from a specific direction.

The seismic power density correlates well with hydrocarbon pore thickness published by DuBois et al. [2004]. The same seismic attribute does not correlate to an elevation map and only weakly correlates to an estimated well density map. Also, no qualitative relation is visible to the data acquisition pattern that

spanned seven days. This is confirmed by the observation that day-to-day temporal variations of $\bar{P}$ were much smaller than the observed lateral variations (Figure 2.9).

The above three variables are thus unlikely to confound the correlation to the reservoir parameter map and we have a situation where the hypothesis can not be rejected at this stage. However, other non-hydrocarbon variables should still be included. Withers et al. [1996] investigated the effect of wind speed on the seismic background noise on a location in New Mexico with similar topography and vegetation. No precipitation occurred during the data acquisition period and wind speeds did not exceed the 3 m/s threshold above which Withers et al. [1996] observed an impact on seismic background noise on the surface. A significant atmospheric impact on the analysis is therefore unlikely.

Weathering layer thickness should still be included because of its well known influence on particle motion. Information on the near-surface geology in Jonah is very limited. The HVSR technique may still be used to estimate lateral variations of near-surface geology [Goertz et al., 2012b]. However, an analysis of H/V spectra in the measured locations showed that they were largely controlled by source characteristics, rendering them unusable for site characterization. No statics from active seismic surveys were available to us, but based on experience from the wider area, they often correlate with elevation. In this context, we used the elevation map as the closest available proxy for near-surface geology.

The well pad density map as a proxy for anthropogenic seismic noise should be considered with caution since it is based on the assumptions that each well isotropically radiates the same amount of surface wave energy during the analysis period and that the noise decay with distance does not depend on location. We use it in this work mainly to illustrate how infrastructure noise could be included in a comparative correlational analysis. Alternative proxies based on field logs of surface activity and dedicated seismic characterization of noise sources around the field would be preferable. Energy from near-by road traffic is unlikely to affect the quietest time periods, since this would require such transient events to occur during more than 90% of the recording period. The selected attribute also shows no visual relation to the transit roads in Figure 2.10.

The bootstrap method is likely to generate somewhat optimistic confidence intervals because the considered variables have a certain degree of spatial smoothness. A more appropriate method to use in such a situation is described

by Buhlmann [2002]. However, the simple technique used here is sufficient since the relatively high correlation to the subsurface reservoir is likely to hold even with larger confidence intervals and the statement of insignificant or weak correlation to non-hydrocarbon variables is even more supported by larger confidence intervals.

We note again that the lateral isotropy is a necessary but not sufficient requirement to identify sub-vertically traveling body waves. Noise cross-correlations are an interesting alternative to detect body waves [Draganov et al., 2006], but their applicability can be complicated by unsuited ambient source distributions and/or short recording periods. Another alternative to reduce ambiguities in the wave field analysis is offered by array analysis [Rost & Thomas, 2002]. Appropriately designed sensor arrays distributed around the area of interest can provide a means to detect sub-vertically incident body waves directly [Zhang et al., 2009, Gerstoft et al., 2008b, Koper et al., 2009, Birkelo et al., 2010].

2.5 Conclusions

Providing empirical evidence for the hypothesis that the ambient seismic wave field above 1 Hz carries information about the deeper subsurface (>800 m) is hampered by (i) uncontrolled surface wave sources and (ii) influences of the near-surface geology. Both effects can confound correlations of observed attributes to deeper targets.

In this work we suggest a strategy that addresses these issues in two main steps. First, we only average seismic power density over frequencies and time periods where the dominant polarization has low power and is laterally isotropic. This is considered a minimum criterion to reduce surface wave interferences from a few dominant sources. Second, we test whether the residual seismic power density significantly correlates to a hydrocarbon target and that this correlation is stronger than correlation to non-hydrocarbon targets. We recommend to do this quantitatively using the non-parametric rank correlation and bootstrap confidence intervals.

A first example for the proposed strategy is given for an ambient wave field dataset acquired over a producing tight-gas field. We find that the quietest 10% of time periods in the 1.5-3.0 Hz band meet the minimum requirements against few dominating surface sources. Seismic power density in this subset significantly

correlates to a hydrocarbon pore thickness map, while it does not significantly correlate to elevation, only weakly correlates to an estimated well density, and shows no apparent relation to the data acquisition days. We can therefore not falsify the hypothesis that the ambient wave field contains a measurable amount of body wave energy carrying information about the subsurface hydrocarbon target.

Acknowledgements The authors thank the Low Frequency Seismic Partnership (LFSP) for supporting this research and Spectraseis for providing the dataset. Hans-Rudolf Kunsch is thanked for helpful advice on the statistical analysis.

Chapter 3

Time-lapse analysis of ambient surface wave anisotropy: a three-component array study above an underground gas storage

Abstract We perform a time-lapse analysis of Rayleigh and Love wave anisotropy above an underground gas storage facility in the Paris Basin. The data were acquired with a three-component seismic array deployed during several days in April and November 2010. Phase velocity and back azimuth of Rayleigh and Love waves are measured in the frequency range 0.2-1.1 Hz using a three-component beamforming algorithm. In both snapshots, higher surface wave modes start dominating the signal above 0.4 Hz with a concurrent increase in back azimuth ranges. We fit anisotropy parameters to the array detections above 0.4 Hz using a bootstrap approach which also provides estimation uncertainty and enables significance testing. The isotropic phase velocity dispersion for Love and Rayleigh waves match for both snapshots. We also observe a stable fast direction of NNW-SSE for Love and Rayleigh waves which is aligned with the preferred orientation of known shallow (<300 m) and deeper (∼1000 m) fault systems in the area, as well as the maximum horizontal stress orientation. At lower frequencies corresponding to deeper parts of the basin, the anisotropic parameters exhibit higher magnitude in the November data. This may perhaps be caused by the higher pore-pressure changes in the gas reservoir in that depth range.

This chapter is published in Journal of Geophysical Research *(doi: 10.1002/jgrb.50375).*

Authors: Nima Riahi, Göetz Bokelmann, Paola Sala, Erik H. Saenger.

3.1 Introduction

Sensing seismic anisotropy in the crust and upper mantle provides important constraints on deformation and forces acting within the solid earth [Maupin & Park, 2007]. Both, body waves and surface waves have been frequently used for measuring such anisotropy over the last decades [Wüstefeld et al., 2009]. A still novel field in this respect is the application of array techniques for that purpose. This has been introduced for P-waves by Bokelmann [1995] who had used array analysis together with polarization of P-waves to constrain seismic anisotropy in the crust. Bear et al. [1999] had embedded such an approach in a three-component array processing procedure. Since surface waves provide good depth resolution, it is highly desirable to establish such an approach for surface waves, which is also natural considering the rich polarization properties of surface waves.

Numerous studies have addressed seismic anisotropy using surface waves [Maupin & Park, 2007], and more recently using ambient noise as a new data source to probe such anisotropy. The ambient seismic wave field is attractive because it carries significant seismic surface wave energy and is usually also available where natural seismicity is low. Many noise studies analyze seismic cross-correlations, often in combination with tomographic inversion, to produce maps of seismic anisotropy in the subsurface [Shapiro et al., 2004, Moschetti et al., 2010, Fry et al., 2010, Schaefer et al., 2011, Gallego et al., 2011, Pawlak et al., 2012, Adam & Lebedev, 2012]. Such techniques were also applied to study anisotropy changes after large earthquakes [Durand et al., 2011, Nakata & Snieder, 2012, Takagi & Okada, 2012, Tonegawa et al., 2013]. Alvizuri & Tanimoto [2011] used array analysis on teleseismic Rayleigh waves in Southern California to compute azimuthal anisotropy, but noise analyses with arrays are generally focused more on isotropic velocity structure inversion [Scherbaum et al., 2003, Kind et al., 2005, Parolai et al., 2005, Wathelet et al., 2005] and the study of oceanic microseism source mechanisms [Bromirski & Duennebier, 2002, Gerstoft et al., 2006, 2008b, Koper & de Foy, 2008, Koper et al., 2009, Zhang et al., 2009, 2010, Landes et al., 2010]. However, such studies generally do not make full use of the polarization properties, and we will show in this paper that including that information can be rather useful.

We analyze ambient seismic data acquired above an underground gas storage (UGS) facility about 170 km southwest of Paris during a few days in April and

November 2010. This represents two wave field snapshots of not only Spring and Autumn, but also of a low pressure and high pressure state of the UGS reservoir. The data were acquired in the frequency range 0.1-1.1 Hz where substantial energy from microseisms dominate the wave field and where non-negligible sensitivity to the storage reservoir can be expected. Our objective is to use frequency domain beamforming to test if the two snapshots differ in surface wave anisotropy at the surface, a phenomenon that might be linked to changes in the storage reservoir. The statistical variability in the beamformer estimates due to uncontrolled factors are taken into account by considering bootstrap-based uncertainty assessments on the anisotropy parameters.

Our results show that ambient surface wave anisotropy generally exhibits a fast axis roughly parallel to the preferred orientation of local fault systems. However, we observe that the anisotropy magnitude increases at lower frequencies in November. Pore pressure variations in the reservoir could plausibly be responsible for the observed change, although other causes cannot be ruled out. This study shows how array processing of the higher frequency flank of microseisms may open interesting possibilities for monitoring applications in the shallow crust and in sedimentary basins.

In section 3.2 we describe the geological setting of the study area and the data acquisition. Three-component array beamforming is explained in section 3.3. An overview of the array detections is given in section 3.4. Section 3.5 describes how the anisotropy parameters were fitted and shows the results for both snapshots. Discussion and conclusions follow in sections 3.6 and 3.7.

3.2 Geology of study area and data acquisition

The Paris Basin is a broad and circular intracratonic basin filled with sediments at low regional dips (on average less than 1°). The basin developed mostly during the Mesozoic and was affected by numerous deformation phases in the Tertiary. Deep fault systems are known to affect a large part of the stratigraphic column, from the deep Permo-Carboniferous levels up to Cenozoic levels [Debeglia & Debrand Passard, 1980, Beccaletto et al., 2011]. Some of these faults were reactivated during later tectonic phases and their orientation was inherited up to the shallow subsurface. Similar fault orientations can therefore be observed at different depths.

Our study area is located 170 km southwest of Paris near Chémery and lies above an underground gas storage (UGS) facility that is operating since the 1970s. The Chémery area is situated in one of the deepest parts of the Paris Basin with a sedimentary thickness above the basement of up to 2.6 km [Perrodon & Zabek, 1990]. The sedimentary sequence mainly consists of limestone, dolomite, shale, and fluvial sandstones. The gas reservoir levels are located below 1085 m depth and may have a thickness of up to 100 m [Hamon & Merzeraud, 2005]. They are known to be affected by NNW-SSE and E-W fault systems, as seen on a map of the Triassic sandstone shown in Fleury et al. [1997]. Some of the faults propagate through the entire sedimentary sequence up to the surface as suggested by a W-E cross section in Fleury et al. [1997]. Therefore, the same fault system is also recognizable in shallow formations and was recently reconstructed in a three-dimensional geological model [Sala et al., 2013]. Figure 3.1 shows a map of the study area with the shallow and deeper fault systems.

We analyze data acquired with a temporary seismic three-component array with a rectangular aperture of 3×6 km and consisting of 80-85 seismometers. The array was deployed during four days in April (effective operation 65 h) and eight days in November 2010 (effective operation 55 h). The Chémery UGS is France's largest gas storage facility and although precise pore pressure information was not available to us it is known that in April it was close to it lowest pore pressure due to Winter gas extraction while in November it was close to its maximum pore pressure after gas injection during the warmer months. The red triangles in Figure 3.1 show a typical array geometry from April 22$^{\text{nd}}$.

The acquisition geometry changed three times during each of the acquisition periods. This led to eight array geometries that shared the same aperture (3×6 km) and inter-station spacing (about 500 m) but had slightly different internal configurations due to permitting constraints and sensor failure. The variations in the array beam patterns, however, were negligible and also afforded a spatial sampling with little aliasing issues below 1 Hz. The seismic wave field was sensed using broad-band seismometers with a sensitivity of 1500 V/(m/s) and a relatively flat frequency response above 0.03 Hz. The cut-off frequency due to the sampling rate was 50 Hz. The sensors were placed in small holes of about 50 cm depth, oriented towards magnetic North using a hand compass, and covered with a wooden board to reduce wind noise and interactions with local fauna.

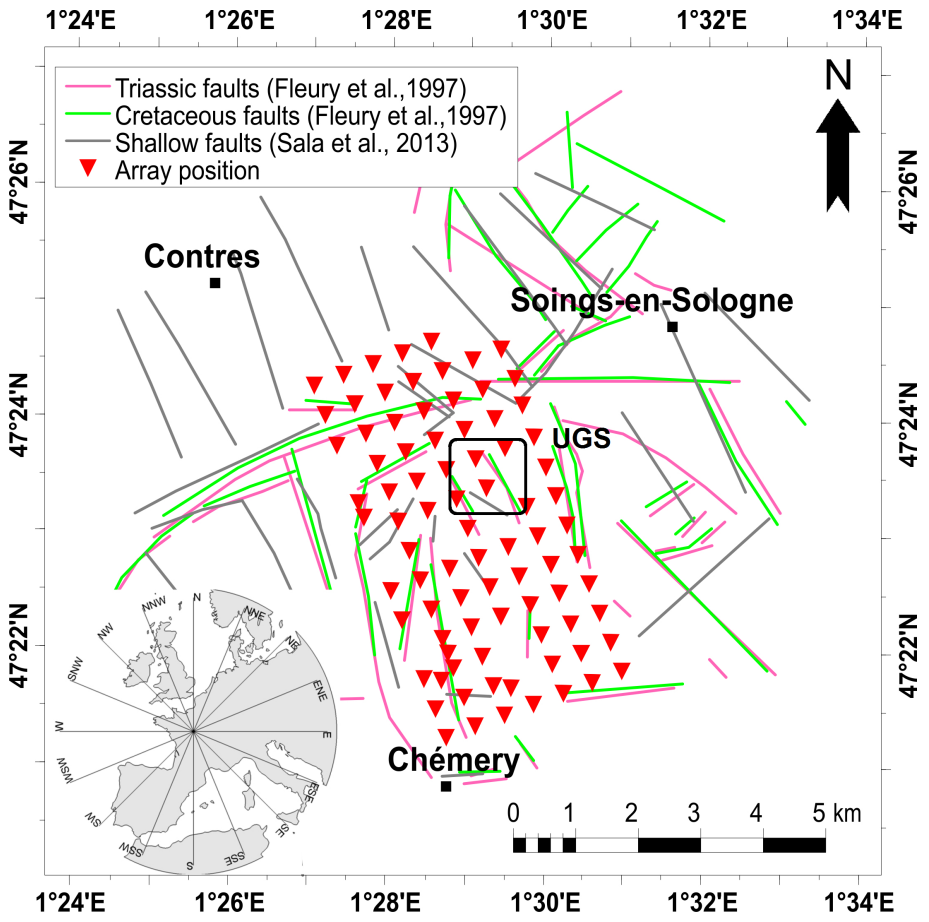

FIGURE 3.1: Map of the study area. The solid lines show fault systems characterized by Sala et al. [2013] (<300 m, gray) and Fleury et al. [1997] (1 km, red and green). The square shows the approximate location of the main UGS facilities and the triangles indicate the locations of the broad-band seismometers on April 22nd. The inlet shows an orthographic map of the region around the array..

For a small fraction of the receivers one or more components deviated strongly from the rest of the array members. These stations were removed from the analysis. Using teleseismic earthquake arrivals one recording was found to have a time shift of 2.6 sec and was corrected accordingly.

3.3 Three-component array processing

The frequency-wavenumber technique has been successfully applied in many applications of seismic noise ranging from microseism studies (e.g. Toksoz & Lacoss [1968], Gerstoft et al. [2008b]) to subsurface velocity inversion (e.g. Scherbaum et al. [2003], Kind et al. [2005]). The technique extracts back azimuths and phase

velocities of coherent wave trains as a function of frequency, thus providing a detailed characterization of the seismic wave field at the array location. We first give a brief formulation of this technique for single component arrays and then show how it can be extended to additionally decompose polarization for seismic three-component arrays.

3.3.1 Single component array

For a single component, say vertical, array the Fourier amplitudes of the signals on all M array sensors are summarized as a data vector $\mathbf{s} = [s_1, s_2, \ldots, s_M]^T$ (the frequency dependence will not be explicitly stated in the following equations). The relative phase shifts between these M signals are captured by the spectral density matrix $\mathbf{S} = \langle \mathbf{s} \cdot \mathbf{s}^\dagger \rangle$, where $\langle\ \rangle$ stands for the time ensemble operation. The diagonal elements s_{ii} are real and estimate the power spectral density of the signal on sensor i, while the off-diagonal elements s_{ij} are complex and estimate the cross spectral power density between sensors i and j. The phase and amplitude of s_{ij} corresponds to the relative phase difference between the sensors and the product of the expected amplitudes on the sensors. Note that $\mathbf{S}$ is therefore hermitian.

A typical model for a coherent signal across the array is the plane wave. The complex magnitude of the plane wave for a location $\mathbf{r}$ and time t is:

$$x(\mathbf{r}, t) = A \cdot \exp[2\pi i(\mathbf{k} \cdot \mathbf{r} - ft)]\,, \tag{3.1}$$

where $\mathbf{k}$ is the wave vector of the plane wave with a magnitude $1/\lambda$ and pointing in the direction of propagation [km^{-1}], f is the frequency of the plane wave oscillation in time [Hz], and A is a complex scalar describing the amplitude and phase of the plane wave. The theoretical phases and amplitudes observed on the M signals of the array by such a passing plane wave are captured in the so called mode vector or array response vector. This is a complex M-dimensional vector parameterized by the wave vector, $\mathbf{k}$:

$$\mathbf{a}(\mathbf{k}) = \frac{\mathbf{1}}{\sqrt{\mathbf{M}}} \begin{bmatrix} \exp(2\pi i\, \mathbf{k} \cdot \mathbf{r}_1) \\ \vdots \\ \exp(2\pi i\, \mathbf{k} \cdot \mathbf{r}_M) \end{bmatrix}, \tag{3.2}$$

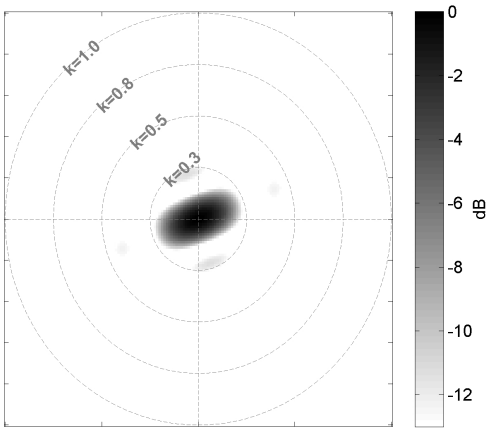

FIGURE 3.2: Array response for the array geometry shown in Figure 3.1 (units are [/km]). Besides the main lobe in the center four weak (<-12 dB) sidelobes are visible. Since the array geometry did not change substantially this response remained essentially the same for all acquisition days.

where $\mathbf{r}_n$ are the coordinates of sensor n in the array plane and the factor $1/\sqrt{M}$ normalizes $\mathbf{a}$ to unit length. For a single component array the vector elements vary only in complex phase, not amplitude. In conventional beamforming, given the observed phase variations in $\mathbf{S}$ the response of the array signals as a function of $\mathbf{k}$ is computed as:

$$R(\mathbf{k}) = \mathbf{a}(\mathbf{k})^{\dagger} \cdot \mathbf{S} \cdot \mathbf{a}(\mathbf{k})\,. \tag{3.3}$$

The response $R(\mathbf{k})$ is maximized when $\mathbf{k}$ matches with the wave vector of an actual plane wave impinging on the array [Lacoss et al., 1969]. In an asymptotic sense (very large arrays), this result also applies when there are several sources, each having an arbitrary stochastic time signal and mixed with random noise. However, sufficient time must be recorded to reduce the variance of the estimations. The effect of array geometry on beam-forming estimates is mainly controlled by the array response pattern which is the beam-response to a signal that has the same phase on all receivers (i.e. infinite velocity). Figure 3.2 shows the response of the array operating on April 22$^{\text{nd}}$. The central peak controls the beamformer resolution in wave number space and side lobes indicate potential leakage to other wave vectors. These aspects are discussed further below. As mentioned in section 3.2 the variations in the array responses of the eight array configurations present during the study were negligible.

3.3.2 Three-component array

Our formulation of the three-component array is closely related to that of Esmersoy et al. [1985]. For the case of a three-component array, let the frequency domain data vector of the $M \cdot 3$ array signals be

$$\mathbf{s}_{3C} = [s_{E,1}, \ldots, s_{E,M}, s_{N,1}, \ldots, s_{N,M}, s_{V,1}, \ldots, s_{V,M}]^T , \tag{3.4}$$

where $s_{\{E,N,V\},n}$ is the Fourier amplitude of the signal on the East, North, or Vertical component of receiver n. The spectral density matrix in this case is $\mathbf{S}_{3C} = \langle \mathbf{s}_{3C} \cdot \mathbf{s}_{3C}^{\dagger} \rangle$. The plane wave model of the single component array can now be extended to include a polarization. For a single receiver in the frequency domain such a polarization corresponds to three sinusoids on the three components with varying phases and amplitudes. These phases and amplitudes can be described by a complex three-dimensional vector $\mathbf{c} = [c_E, c_N, c_V]^T$ [Samson, 1983], where again the subscripts stand for the East, North, and Vertical component. For a laterally homogeneous and isotropic medium and a remote source only certain polarization states can propagate: retro/prograde Rayleigh-, SH-, SV-, and P-wave polarization. Rayleigh wave states are elliptical within the vertical plane of propagation, with the major axes aligned horizontally or vertically. They are parameterized by their degree of ellipticity which can be described by the ratio of the length of the horizontal major axis to the vertical major axis, sometimes called the H/V ratio. P- and SV-polarizations are parameterized by a dip angle ψ. Since there is only one SH-polarization state no parametrization is necessary there. Table 3.1 gives the parameter ranges of all 91 polarization states used in this study and Figure 3.3 illustrates two surface wave polarizations from this set. We describe the parameters of the polarization with ξ and the resulting three component phase shifts are $c(\xi)$. Since the length of $\mathbf{c}$ does not affect the described polarization it can be set to have unit length. The resulting signal model is written as:

$$\mathbf{x}(\mathbf{r}, t) = A \cdot \mathbf{c}(\xi) \cdot \exp[2\pi i(\mathbf{k} \cdot \mathbf{r} - ft)] , \tag{3.5}$$

where $\mathbf{x}$ is now a three-component time series describing particle motion velocity at location $\mathbf{r}$. Again, A is a complex scalar describing the amplitude and phase of the plane wave. Note that in eq. 3.5 the polarization ξ is modeled independently

from the wave vector $\mathbf{k}$. The relative phase variations of a fixed component, say East, among all M sensors therefore only depend on the wave vector and are still captured by the mode vector $\mathbf{a}(\mathbf{k})$. On the other hand, the relative phase variations between the three components due to polarization are the same for all receiver locations in this model. The $3 \cdot M$-dimensional, complex mode vector of the array can thus be written as:

$$\mathbf{w}(\mathbf{k}, \xi) = \mathbf{c}(\xi) \otimes \mathbf{a}(\mathbf{k}) \,, \tag{3.6}$$

with $\otimes$ being the Kronecker product. The first M elements of $\mathbf{w}$ describe the phase responses of all East components in the array, the next M elements those of the North components, and the last M elements those of the vertical components (the receiver order being the same for all components). Note that because both $\mathbf{a}$ and $\mathbf{c}$ are normalized also $\mathbf{w}$ has unit length. As in the single component case, the conventional beam-forming response as a function of wave vector $\mathbf{k}$ and polarization parameters ξ is:

$$R(\mathbf{k}, \xi) = \mathbf{w}(\mathbf{k}, \xi)^{\dagger} \cdot \mathbf{S}_{3\mathrm{C}} \cdot \mathbf{w}(\mathbf{k}, \xi) \,. \tag{3.7}$$

Again, the response $R(\mathbf{k}, \xi)$ is maximized when $\mathbf{k}$ and ξ match with the parameters of an actual plane wave impinging on the array. This also holds when there are several coherent waves because the mathematical representation of the extended model here is formally identical to the single component case [Schmidt, 1986].

Each of the three components of the M receivers is segmented into 40.96 sec windows, overlapping by 20.48 sec. We then compute $\mathbf{s}_{3\mathrm{C}}$ using the Fast Fourier Transform [Press, 2007]. For a given frequency bin, f_0, we estimate $\mathbf{S}_{3\mathrm{C}} = \langle \mathbf{s}_{3\mathrm{C}} \cdot \mathbf{s}_{3\mathrm{C}}^{\dagger} \rangle$ using block-averaging [Press, 2007] over 15 consecutive time windows, corresponding to a total window length of $\sim$5 min.

The array coherence $R(\mathbf{k}, \xi)$ (eq. 3.7) is computed over a discretized wave vector and polarization space. For the wave vector a polar grid is used with wave numbers ranging from 0.0056 to 0.45 km^{-1} in steps of 0.0056 km^{-1} and with propagation azimuth steps of 5°. We follow the convention of 0° for North and positive angles indicating clockwise rotation.

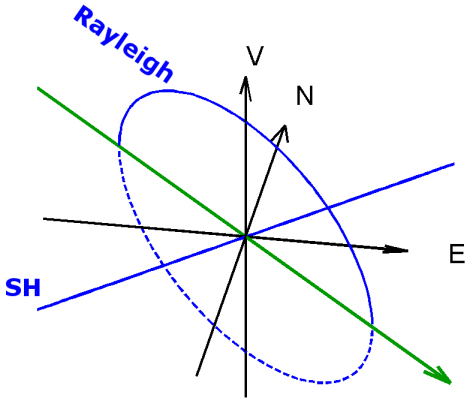

FIGURE 3.3: One Rayleigh and one SH polarization state are illustrated for a plane wave propagating from NW to SE. The elliptical Rayleigh states can have prograde (forward movement from top) and retrograde (backward movement from top) motion.

Polarization type	**Parameter**	**value range**
Rayleigh	H/V ratio	Inf,5,2.5,1.67,1.25,1,0.8,0.6,0.4,0.2,0
	orientation	prograde/retrograde
SH	no parameter	n/a
P	dip angle	0...90° in 2.5° steps
SV	dip angle	0...90° in 2.5° steps

TABLE 3.1: Parametrizations employed in the array processing for the various polarization states. A total of 91 polarization states were used at each wave vector.

The parameter space is thus discretized into 5760 wave vectors, with each wave vector connected to the 91 polarization states shown in table 3.1. The coherence R is now evaluated for each of the $5760 \cdot 91$ joint states. For each wave vector the maximum coherence over all polarizations at that wave vector is stored as well as the polarization that caused it. Figure 3.4 shows the resulting maximum coherence over the wave vector space for a sample time window at 0.73 Hz. To facilitate interpretation the wave numbers were scaled to represent slowness as $s = 1/v = k/f$ (v is the phase velocity) and the azimuth angles represent back azimuth (propagation azimuth +180°). The spectrum shown here is different from spectra computed using single-component beamforming because each peak in the slowness spectrum may correspond to another polarization state, depending on what polarization maximized the coherence. In order to capture not just the dominating wave train but also weaker ones, the strongest three peaks in the combined wave vector and polarization space are selected for further analysis.

The slowness spectrum shown in Figure 3.4 is typical for the observations made at the site in that a few clear peaks were visible in most time windows. For the time scale of the analysis window (5 min) it appears that a few distinct sources mostly dominated the wave field. When more than one coherent signal is present the coherence becomes a biased measure of the seismic power of the signal [Capon, 1969]. We follow the procedure described in Schmidt [1986] (equation 7 of that paper) to estimate the power for the coherent signals as well as an assumed incoherent noise. This also allows us to assess the signal-to-noise ratio of the detections. Appendix C of this thesis provides more detail on this topic.

The above process is applied to all Fourier frequencies from 0.1-1.1 Hz and then repeated for a window about 2.5 mins further ahead in time. The array response shown in Figure 3.2 contains small side lobes where seismic energy could leak in the analysis with an attenuation of slightly more than 12 dB. Despite that attenuation, these side lobes can lead to strong spectral leakage when the wave field is actually dominated by one single wave train, which was often the case below 0.3 Hz. To reduce such spurious peaks below 0.3 Hz, we discard detections that were more than 50% weaker than the strongest response for frequencies below 0.3 Hz. The resulting database of detections over the 65 hour and 55 hour recording periods in April and November 2010 will be visualized in the next section.

We also tested the high-resolution beamforming techniques proposed by Capon [Capon, 1969] and Schmidt [1986] and found very similar results, confirming the statement by Koper et al. [2010] that for a statistical characterization of wave field properties the choice of the array processing scheme is not critical.

3.3.3 Signal mixtures, polarization perturbations

At any given time there might be more than just one dominantly coherent wave train impinging on the array. The processing must be able to properly detect such mixtures with little bias. Also, anisotropy and lateral heterogeneity in the subsurface can cause polarizations to deviate from the isotropic polarization states defined above. One known phenomenon is a deviation of polarization angle and propagation direction [Maupin & Park, 2007, e.g.]. We assess the performance of the array processor in these two situations using synthetic stochastic signals.

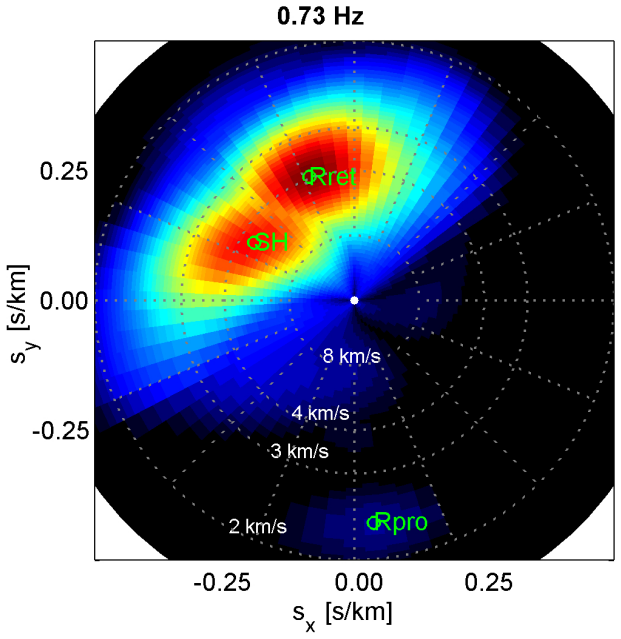

FIGURE 3.4: Slowness spectrum at 0.73 Hz for a sample time window centered around April 20th, 14:45h local time. For the three strongest peaks the polarization that was associated with the response is given (SH: transverse linear, Rret/Rpro: retrograde/prograde Rayleigh).

We considered the following scenario. A mixture of three polarized plane waves (without anisotropy effects) impinges on the array: (1) a retrograde Rayleigh wave with H/V amplitude ratio 2.5, traveling at 2.4 km·s^{-1} from direction -15° (∼NNW), (2) a prograde Rayleigh wave with H/V ratio 1, traveling at 3.5 km·s^{-1} from back azimuth -70° (∼WNW), and (3) a Love wave traveling at 2.8 km·s^{-1} from back azimuth -120° (∼WSW). The time signatures of the three plane waves are Gaussian stochastic processes which are independent of each other. Random, spatially white and unpolarized Gaussian noise was added to the entire setup. The simulation was done entirely in the frequency domain for a test frequency of 0.54 Hz. For each time window we simulated the frequency domain amplitudes of the three waves and the added noise by independently and uniformly drawing four amplitudes from a circularly symmetric complex normal distribution (these amplitudes correspond to the term A in equation 3.4). The frequency domain amplitudes were scaled such that, in the time domain, the expected signal amplitudes were equal and the ratio of the expected signal amplitudes to the expected noise amplitude was 0.25.

We simulated 100 realizations of the above synthetic signal model and processed them with the parameters given in the previous section. The resulting distribution

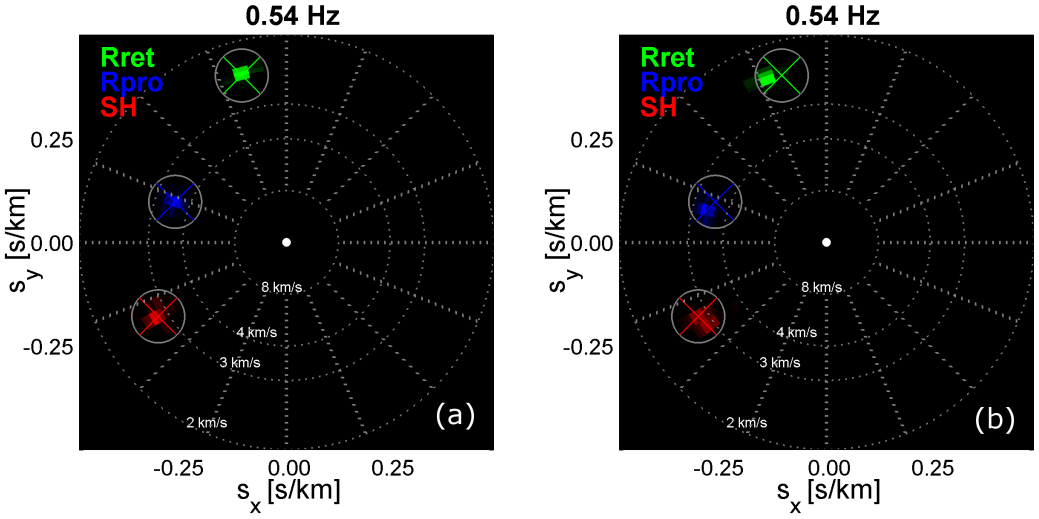

FIGURE 3.5: *(a)* Distribution of detections on 100 realizations of a synthetic mixture of three signals with signal-to-noise ratio 0.25 (details see text). The crosses indicate the true simulated slowness and backazimuths. The signals are correctly separated. *(b)* Same as in (a), but the particle motion of the plane waves was rotated along the vertical axis by 20° counterclockwise. Rayleigh and Love polarization thus substantially differs from the isotropic state assumed by the array processor. The effect on estimated slowness is small, but the back azimuths show a bias of 5° counterclockwise.

of detections is given in Figure 3.5a. The processor correctly identified all signals and their wave type. Most of the detections were made within the correct grid point in back azimuth and slowness.

Next, we rotated the particle motion by 20° counterclockwise around the vertical axis to construct out-of-plane polarization. This emulates a situation where lateral heterogeneity and/or anisotropy create polarization states that deviate substantially from the states that can propagate in a simplified model. Again, 100 realizations of spectral density matrices from these distorted signals were simulated and processed with the algorithm. Figure 3.5b shows the distribution of the resulting detections. The signals are still detected and properly identified, but there is a systematic bias in the back azimuth by about 5°.

The synthetic data used above represent perfect plane wave trains which are stochastically independent in time. In the real data such signals are unlikely to be encountered in this purity due to scattering, attenuation, multi-pathing, and other phenomena. A detailed discussion of such effects would go beyond the scope of this test but we mention them here to be clear about the limitations of the synthetic setup. However, the results qualitatively show that conventional

beam-forming can be used to characterize signal mixtures and that the isotropic polarization subspace is sufficient to classify Love and Rayleigh waves with moderate polarization perturbations. There is a small but significant estimation bias in back azimuth which should be kept in mind when analyzing surface wave anisotropy (see discussion). When two peaks in the slowness plane are too closely spaced they can merge into one peak. This limits the angular resolution of the beamformer but could also affect the estimated phase velocity. One controlling factor for angular resolution is the size of the main lobe in the array response shown in Figure 3.2. For the array geometry available in this study the angular resolution was about 15°.

3.4 Overview of detections

The 65 hour and 55 hour datasets in April and November 2010 provided 1681 and 1411 time windows, respectively, yielding a total of more than $280 \cdot 10^3$ detections over all 38 frequency bins between 0.2-1.1 Hz. Five earthquakes with moment magnitude M>5 occurred during data acquisition. They were teleseismic events and did not dominate the signal. Since we will analyze time distributions of the wave train parameters their influence would be negligible anyway, because they would affect less than 1% of the time windows. The UGS was operational during the data acquisition phases and can produce large amounts of seismic noise at high frequencies. This noise has been studied (not shown) but no clear signs of infrastructure noise were found below 1.1 Hz. In the following, we give a brief overview of the detections in three different frequency bands for both snapshots. Back azimuths are given with their geographical abbreviation, according to the inlet in Figure 3.1.

Figure 3.6a shows a 2D histogram over the slowness-back azimuth polar grid of all detections in April with Rayleigh- and SH-wave polarizations at 0.22 Hz, corresponding to a frequency almost at the spectral peak of the ocean microseism. The histograms were computed separately for each polarization type and a composite color spectrum was produced (red for SH/Love, green for retrograde Rayleigh, blue for prograde Rayleigh). Figure 3.6b shows the same visualization for the November data. The microseism peak consists of both Rayleigh and Love waves but the mixture is a function of back azimuth. Love waves dominate from WNW while retrograde Rayleigh waves dominate from NNW directions. Although

the slowness resolution is relatively poor at these low frequencies, the Rayleigh waves seem to propagate slightly faster compared to the Love waves. These patterns vary little in the two snapshots. Note that waves in this frequency band are generated by ocean gravity waves both near coasts as well as in the deep ocean [Ardhuin et al., 2011, Hillers et al., 2012b]. The source area cannot be unambiguously inferred with this analysis.

Figure 3.6c and 3.6d show the same visualizations for the frequency bin 0.54 Hz. Prograde Rayleigh waves have a wide range of back azimuths from South to NW (clockwise). In contrast to the April snapshot, the November snapshot shows fewer detections from North and South. The phase slowness estimations for Love waves are much more scattered than those for the Rayleigh waves which suggests problems with the beam-forming algorithm for this frequency and polarization. The brightness of the Love wave histogram has therefore been lowered to emphasize the more stable Rayleigh wave detections. There is evidence for phase velocity anisotropy: energy from the NNW shows lower slowness (i.e. higher velocity) compared to other directions. The solid line is a fit of anisotropy parameters to the slowness detections and is discussed in detail in section 3.5. Note that the best fit differs visibly between the snapshots.

Figure 3.6e and 3.6f show the same visualizations for the frequency bin 0.81 Hz. Phase slowness is much better constrained for both snapshots and both Love and Rayleigh waves have a wide azimuthal distribution, although their relative proportions still vary by back azimuth. Fewer detections are made from South to SSW (clockwise). The November snapshot has fewer detections from North compared to the April snapshot. Anisotropy in phase velocity is apparent for Rayleigh and Love detections with a fast axis roughly along the NNW-SSE axis. Again, the solid line is a fit of anisotropy parameters to the slowness detections.

Overall, detections above 0.4 Hz show a wider range of back azimuths compared to the microseism peak frequencies. The seasonal snapshots are relatively similar in their main features. No detections were made from the inland directions (NE-SE, clockwise) and only very few detections were made from the North Sea (N-NE) and the Mediterranean Sea (S-SE). Anisotropy in surface wave phase velocities is consistently observed with different magnitude over all frequency bands and for both seasonal snapshots. This surface wave anisotropy is quantified in the next section.

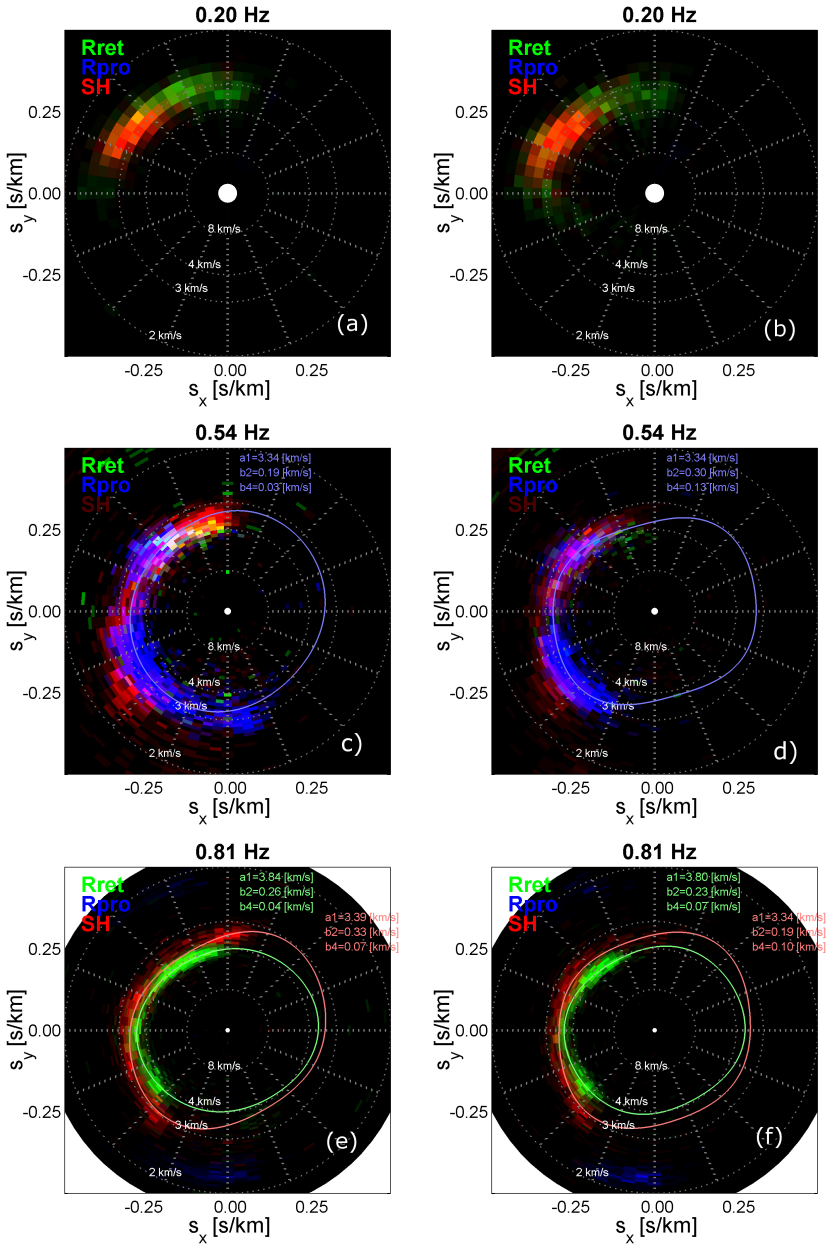

FIGURE 3.6: Distribution of detections in the phase slowness plane for various frequencies. *(a), (c), (e):* results for the April snapshot for 0.20 Hz, 0.54 Hz, and 0.81 Hz, respectively. Colors identify the dominant wave type of the detections in each slowness-back azimuth bin: green and blue for retrograde/prograde Rayleigh waves, red for SH/Love waves. The lines show the best fit of equation 3.8 to the data (see section 3.5). *(b), (d), (f):* results for the November snapshot for 0.20 Hz, 0.54 Hz, and 0.81 Hz, respectively.

3.5 Estimation of anisotropy parameters

A visual inspection of Figure 3.6 shows that the surface wave phase velocities vary with back azimuth over a wide range of frequencies. One likely explanation for such a phenomenon is anisotropy in the seismic parameters of the subsurface. Smith & Dahlen [1973] showed that for a stratified half-space such anisotropy would at first order cause a variation of surface wave phase velocities as follows:

$$v(\theta) = a_0 + a_1 \cos(2\theta) + a_2 \sin(2\theta) + a_3 \cos(4\theta) + a_4 \sin(4\theta)\,, \tag{3.8}$$

where v is the surface wave phase velocity [km·s^{-1}], θ is the direction of propagation measured clockwise from North, and a_i are parameters that depend on the subsurface. In this formulation, the magnitude of the 2θ and 4θ terms are

$$\begin{aligned} b_{2\theta} &= \sqrt{a_1^2 + a_2^2}\,, \\ b_{4\theta} &= \sqrt{a_3^2 + a_4^2}\,. \end{aligned} \tag{3.9}$$

Figures 3.7a,b,c show the observed distribution of phase velocity as a function of backazimuth over the population of all April detections of Rayleigh waves at 0.54 Hz and Rayleigh and Love waves at 0.81 Hz, respectively. A clear azimuthal pattern in the detections can be made out in these plots. The histogram bins stretch along the velocity axis with higher velocities which is due to the fact that the array processor response was computed on an even grid in slowness rather than phase velocity. To quantify the velocity anisotropy for a given frequency and polarization type we fit the Smith & Dahlen [1973] model on this data. Figure 3.7a at 0.54 Hz shows an example of suboptimal azimuthal illumination, while the other examples are better in this respect. Note that the distribution of velocities is non-symmetric with a heavy tail towards high velocities, in particular in Figures 3.7a,b. Such non-gaussian distributions can lead to substantial bias in a conventional least-squares fitting procedure. We therefore use the more robust least absolute deviations approach which minimizes the sum of absolute deviations rather than their squares [Bloomfield & Steiger, 1983]. The resulting best fits are shown in the panels of Figure 3.7 as dashed red lines with the best fit parameters given at the bottom right of the panels. Visual inspection confirms that the fitting routine

behaves as expected. Figure 3.8 shows the residuals from the fit and confirms both the asymmetric character and heavy upward tail of the distribution underlying the velocity estimates.

The 4θ parameters (a_3, a_4) are relatively small and the question arises as to their significance. We assess the uncertainty in the fit parameters as well as their statistical significance by recomputing the fitting process on bootstrap resamples of the data [Efron & Tibshirani, 1993]. This process in effect attempts to estimate the sampling distribution of the actual anisotropy parameters given the observed variability in the velocity estimates. Starting out from N azimuth and phase velocity pairs at a given frequency bin and polarization, we randomly sample (with replacement) an equally large set of N data points. The Smith & Dahlen [1973] model parameters are estimated using the above fitting routine and the entire process is repeated B times. This random resampling mimics the variability in the data and allows one to see how this variability is inherited by the anisotropy parameter estimates. The optimal number B of resamples is a function the true variability in the data and the structure of the hypothesis test. We defined it heuristically by using different numbers from B=10 to 1000 and taking the value above which the main features of the bootstrap distribution seize to change substantially. This criterion was attained at B=100. Figure 3.9a shows B=100 bootstrap estimates of the four anisotropy parameters (a_1, a_2) (2θ terms) and (a_3, a_4) (4θ terms) for 0.81 Hz and retrograde Rayleigh polarization, all given as a percentage relative to the isotropic phase velocity a_0. The elongated, diagonal shape of the distribution of the 2θ parameters is evidence that the parameter estimations can be correlated. If the true 4θ terms were zero, the (a_3, a_4) parameters would be distributed around the origin. Note that the distribution of the magnitude $b_{4\theta}$ is always positive, also in the case of $(a_3 = a_4 = 0)$, and no symmetric significance test can be used on it. To account for the latter two facts we test statistical significance in the two-dimensional (a_1, a_2) or (a_3, a_4) space, following a procedure described in Liu et al. [1999]: a convex hull is computed around the 90% of estimations that lie within the center of the (a_1, a_2) and (a_3, a_4) data clouds as defined by the Mahalanobis depth [Liu et al., 1999]. The Mahalanobis depth in this case accounts for the asymmetry in the parameter distributions that was observed in many cases. The convex hulls are shown for the 2θ and 4θ parameters in Figure 3.9a. If the origin (0,0) lies within the convex hull, the parameters are considered statistically insignificant at 90% level. Both the 2θ and 4θ terms are therefore statistically significant at the 90% level. Figure 3.9b

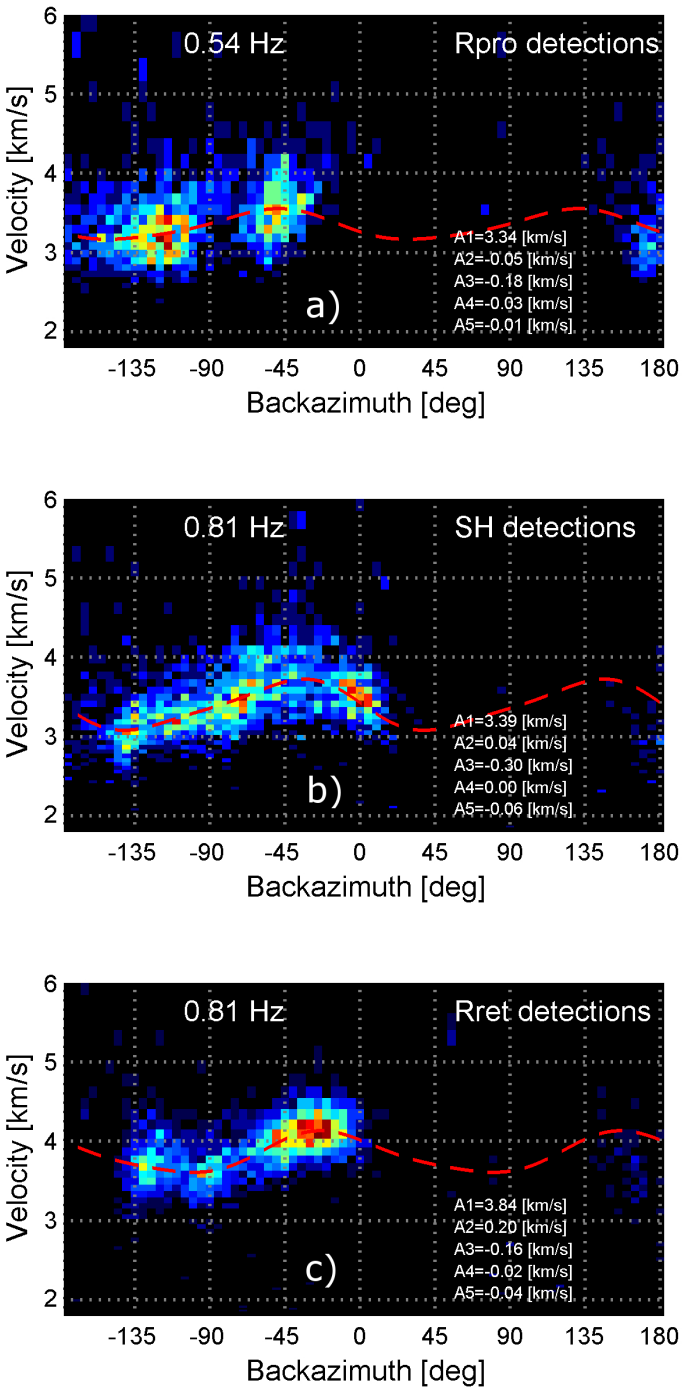

FIGURE 3.7: 2D histograms over the slowness-backazimuth space are computed for *(a)* 0.54 Hz and Rayleigh polarization, *(b)* 0.81 Hz and SH polarization, *(c)* 0.81 Hz and Rayleigh polarization. The dashed red lines indicate the best fit of eq. 3.8, with the coefficients shown on the bottom right.

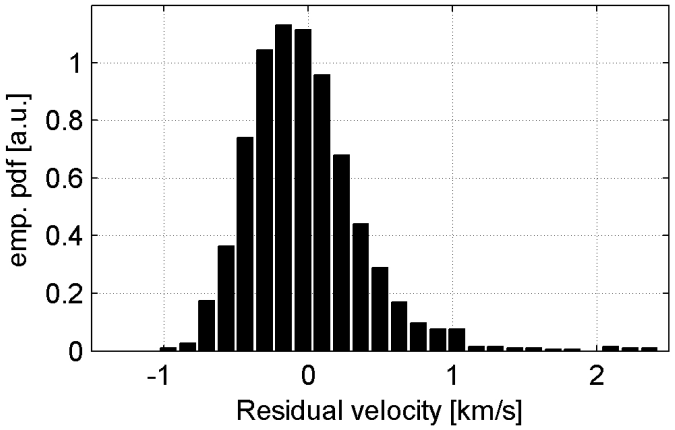

FIGURE 3.8: Empirical probability distribution of the residuals from the fit to the 0.54 Hz Rayleigh wave data. The distribution is asymmetric and has a heavy upward tail.

shows the same graph for the 0.54 Hz bin and prograde Rayleigh polarization, hence the less well constrained example shown in Figure 3.7a. Here the 2θ term is still statistically significant at 90% confidence, but this is no longer true for the 4θ term.

We repeat the above process for every Fourier bin in the analyzed frequency range and for all polarization types. The analysis was restricted to subsets that exhibit a back azimuth range wider than 100° and where there was no evidence of different surface wave modes of the same polarization being present. The latter is done by avoiding frequency bins and polarizations where we observe two different phase slownesses from the same back azimuth or sudden changes in phase slowness as a function of back azimuth. The best fitting models are superimposed as solid lines on the slowness spectra in Figure 3.6.

The above procedure produces spectra for (1) the direction of fastest phase velocity, (2) the isotropic phase velocity a_0, (3) the anisotropy term $b_{2\theta}$, and (4) the anisotropy term $b_{4\theta}$ as defined in eq. 3.9. Figures 3.10a-d show these four spectra for the April and November data, with different colors used to identify the two snapshots. Estimations that did not provide visually acceptable fits were not included in the Figure.

A striking feature in Figure 3.10a is the fact that above 0.7 Hz both Love and Rayleigh waves have a relatively constant fast direction up to more than 1 Hz, a feature that is visible in both snapshots. In this frequency band the fast direction

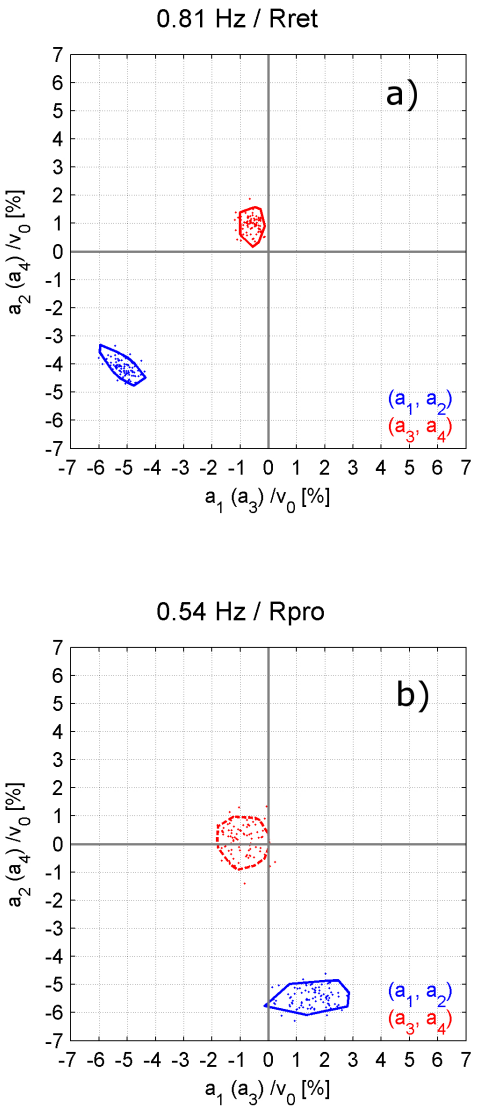

FIGURE 3.9: *(a)* 100 estimations of (a_1, a_2) and (a_3, a_4) were made at 0.81 Hz and SH polarization using bootstrap resamples. The point clouds visualize the estimations and show that the 2θ (blue) and 4θ (red) terms are significant at 90% confidence, with the latter being much smaller. The polygons delineate the convex hull containing the inner 90% of data points. *(b)* The same visualization as in (a) for 0.54 Hz and prograde Rayleigh motion. The convex hull of the 4θ point cloud contains (a_3,a_4)=(0,0) and is thus not significant at the 90% level.

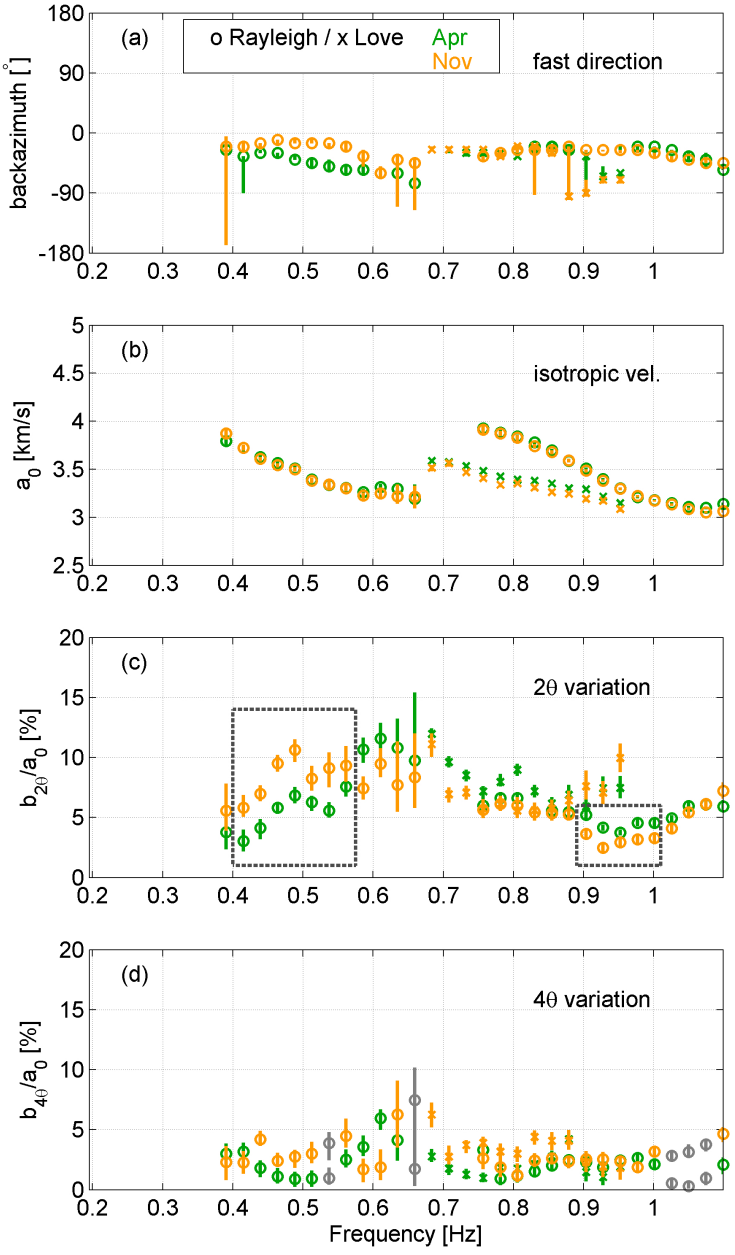

FIGURE 3.10: Spectra of the parameters from the fitting procedure for the April data (green) and November data (orange). The vertical bars in indicate the range within which 90% of the bootstrap estimates lie. Estimates for Rayleigh waves are marked by crosses, those for Love waves by circles. *(a)* Spectrum of the direction of fastest phase velocity. *(b)* Spectrum of the isotropic term a_0 from eq. 3.8. *(c)* Spectrum of $b_{2\theta}$ term. *(d)* Spectrum of $b_{4\theta}$ term. Gray color indicates estimations that are not statistically significant at the 90% level. Note how in those cases the vertical bars cannot necessarily be used for significance testing. The dashed boxes in (b) are discussed in the text.

lies within a back azimuth range of -35 to -20° for both the Rayleigh and Love mode. For the prograde Rayleigh mode below 0.7 Hz this fast direction is less constant over frequency and varies between -60 to -10°. These fast directions (NNW-SSE) are roughly aligned with the preferred fault orientation of shallow (<300 m) and deeper (∼1 km) fault systems described by Sala et al. [2013] and shown in Figure 3.1. Below 0.6 Hz the fast direction of the November data has a more northwards trend compared to the April data.

Figure 3.10b shows the best fitting isotropic dispersion curves for the observed surface waves. Three distinct branches are visible, clearly separated by polarization and phase velocity. These branches probably correspond to three surface wave modes. While the Rayleigh wave velocities match between the snapshots, the Love wave velocities are consistently slower in November by between 10-100 m/s.

Figure 3.10c and 3.10d are a quantitative estimate of the 2θ and 4θ terms, respectively. Markers in gray indicate estimates that were not statistically significant at the 90% level. It is clear that the 2θ term is stronger over most of the frequency band. But the 4θ estimations in most frequency bins are still statistically significant at 90% confidence. Both terms seem to depend primarily on frequency and less so on the polarization of the surface waves. The inappropriateness of using the variability of anisotropy magnitudes for significance testing is made obvious here: the 90% variation range (vertical bars) do not contain zero even where there is no statistical significance.

Comparing the 2θ anisotropy in Figure 3.10c for the two snapshots we note certain features: between 0.6-0.85 Hz the Rayleigh wave anisotropy roughly matches within the 90% variability ranges. Above 0.9 Hz there is a small decrease of 1-2% for November (dashed box to the right). On the other hand, below 0.6 Hz we observe an increase of anisotropy by 3-5% (dashed box to the left). This coincides with the mismatch observed for the fast direction in Figures 3.10a. Love wave anisotropy does not change between 0.85-0.95 Hz but is lower in November by about 1-3% between 0.7-0.85 Hz. The 4θ terms are stronger in November for almost all frequency bins. The observed patterns are discussed in the following section.

3.6 Discussion

We have presented a three-component seismic array processing scheme with which phase velocity and back azimuth of ambient Rayleigh and Love waves can be captured. Ambient surface waves in the frequency range between the secondary ocean microseism peak at 0.2 Hz up to 1.1 Hz were characterized for two snapshots in April and November 2010. The microseism peak energy at 0.2 Hz consists of similar proportions of Love and Rayleigh wave detections coming from a narrow range of back azimuths. However, around 0.4 Hz we observe that higher surface wave modes start to dominate the wave field (Figure 3.10b) with a concurrent widening of the range of back azimuths (Figures 3.6c-f). The increased back azimuth range might be due to higher frequency surface wave scattering on heterogeneities. On the other hand, it is also conceivable that the particular conditions under which ocean gravity waves and the solid earth interact to produce surface waves are satisfied on more locations at higher frequencies.

The array processor assumes isotropic polarization states while the observations show obvious anisotropy in the wave field. However, as tested in section 3.3.3 the sensitivity of the estimator to polarization perturbations seems rather limited and the bias in back azimuth should be small enough to not affect the anisotropy results shown in this study. The clear identification of three surface wave branches in the dispersion curves of Figure 3.10b are further evidence that perturbations in Rayleigh and Love wave polarization are probably not too strong. But such perturbations can still be included in the array processor and might provide additional constraints on the anisotropy in the subsurface [Tanimoto, 2004] although this must be done with caution [Maupin, 2004]. As illustrated in appendix B of this thesis, errors due to random inaccuracies in sensor orientation or location are unlikely.

Anisotropy was quantified by fitting the Smith & Dahlen [1973] model to the data at every frequency bin and for Love and Rayleigh waves separately. Using least absolute deviations rather than least squares accounts for the non-gaussian, heavy-tailed distribution of the phase slowness estimations from the array processing. We approximated the distribution of the anisotropy parameter estimations by a bootstrap distribution and also showed evidence of correlation among the (a_1, a_2) (2θ terms) and (a_3, a_4) (4θ terms) parameter estimates. The resulting uncertainties and significance tests on the anisotropy parameters capture

uncertainties due to the velocity estimation variability, though they do not account for systematic bias in the beamformer detections.

From Figures 3.10a and 3.10b we see that the fast direction and the isotropic part of the phase velocities are relatively stable between the two snapshots with two exceptions: the fast direction has a northwards trends in November between 0.45-0.6 Hz and Love wave phase velocities were found to be consistently lower in November compared to April. For an area in Japan, Nakata & Snieder [2012] observed that shear-wave velocity within less than 500 m depth could be lowered by an increase in precipitation. Although they studied higher frequencies (1-13 Hz) than in this analysis, precipitation may still be a plausible cause of the lower Love wave velocities.

The 2θ variation (Figure 3.10c) between 0.4-0.6 Hz are higher in November compared to April by about 3-5% while at 0.9-1.0 Hz they are lower by about 1-2%. The weaker 4θ terms (Figure 3.10d) were mostly estimated at higher levels in November.

There are several potential explanations for the differences between the snapshots. Diffraction and mode interference will in all likelihood introduce unknown estimation bias within and between the snapshots and the bootstrap uncertainties and significance tests do not address this bias. The orientation of the sensors was done manually using hand compasses which is another potential source of bias. The latter error source would presumably have a random character and it is therefore unlikely that the observed difference between spring and autumn would be produced by it.

Perhaps the most interesting difference, however, is that of the subsurface itself. As mentioned above, differences in precipitation between the seasons might affect the shallow subsurface which could explain the 2θ variations at higher frequencies (right box in Figure 3.10c). Also, the state of the UGS, above which the measurements took place, was distinctly different between the snapshots. In April it was mostly depleted due to Winter demand with minimum pore pressure. In November, on the other hand, it was close to its maximum fill with accordingly high pore pressure. The increased pore pressure might lead to an extension of fractures and/or cracks in the reservoir interval, thereby pronouncing the existing effect of the fault system on surface wave anisotropy. For instance, Tonegawa et al. [2013] used an oblate spheroidal crack model where changes in the crack aspect

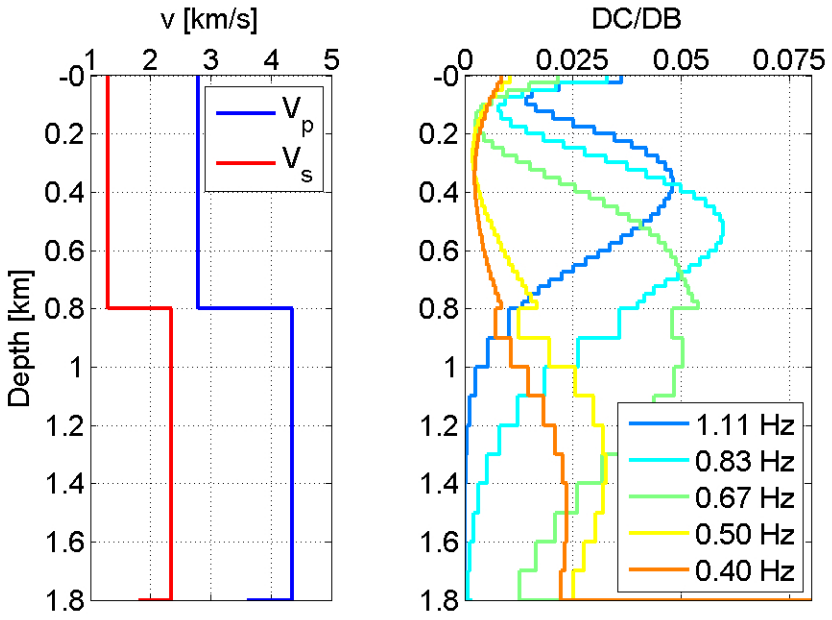

FIGURE 3.11: *(a)* Simplified velocity profile for the Paris Basin based on well data from Bush & Crampin [1991]. *(b)* Depth sensitivity kernels for phase velocity of Rayleigh wave fundamental modes at five different frequencies within the analyzed band.

ratios explained variations in s-wave anisotropy in shallow marine sediments. Considering the depth of the reservoir, this explanation might be more pertinent to the Rayleigh wave anisotropy variations below 0.6 Hz (left box in Figure 3.10c). Another potential source of variation is the overburden. Teatini et al. [2011] described cyclic subsidence and uplift on the vertical and horizontal component above a gas storage facility in the Po plain in Italy. Geomechanical effects on such a large scale might change effective anisotropy measurably. Both explanations are in general consistent with the near-constant fast direction between 0.4-1.1 Hz which matches the preferred direction of both shallow (<300 m) and deeper (1 km) fault systems as shown in Figure 3.1. That orientation also corresponds to the orientation of maximum compressive horizontal stress (e.g., World Stress Map, www.world-stress-map.org).

We want to qualitatively assess the plausibility whether pore pressure changes in the reservoir could be responsible for the observed anisotropy change at 0.4-0.6 Hz. To this end, we compute depth sensitivity kernels [Herrmann, 1996] for surface waves at various frequencies of our observed anisotropy range. Bush & Crampin [1991] published a profile of seismic velocity and density from a well in the Paris Basin that reached down to 2.8 km. We use a simplified version of

that profile in Figure 3.11a which gives the V_p and V_s depth profile for the top 1.8 km, to calculate partial derivatives of phase velocity with respect to changes of in-situ shear-wave velocity. We do not consider the dependence on compressional velocity or density, since they are much weaker. The resulting depth sensitivity kernels for Rayleigh waves are shown in Figure 3.11b. The depth range to which phase velocities are sensitive generally become shallower with increasing frequency. The higher frequencies in this study for which we do not detect any differences are sensitive to the top-most kilometer. The frequencies 0.67 Hz, 0.50 Hz, and 0.4 Hz, where we notice different anisotropic parameters between spring and autumn, have considerable sensitivity to the depth range 1000-1200 m where the reservoir is located. This suggests that the temporal variation of seismic anisotropy might be caused by the different pore pressures in the reservoir between spring and autumn. However, our test only computes sensitivity to shear-wave velocity and not to azimuthal anisotropy directly. That the reservoir interval is responsible for the observed changes therefore remains speculative. Future research could attempt to invert the spectra of anisotropy parameters underlying Figure 3.10 for anisotropy in the subsurface [Montagner & Nataf, 1986]. Since both Love and Rayleigh wave are detected such a scheme could also take advantage of Love-Rayleigh wave discrepancies to address vertical transverse isotropy. Although the beamformer in this study was not optimized for it, the H/V ratio could be used as another wave field parameter to interpret for the cause of the variation.

Finally, the type of seismic anisotropy that we observe here is generally consistent with the orientation of fractures in the area, which also corresponds to the orientation of the maximum horizontal stress in the area. Opening of fractures and cracks can have a significant effect on seismic anisotropy, but the precise mechanism is beyond the scope of this paper.

3.7 Conclusions

We used data from a temporary three-component seismic array in the Paris Basin to characterize ambient Love and Rayleigh waves in terms of their phase velocity distribution as a function of back azimuth. The data was acquired above an underground gas storage during a few days in April and November 2010. The analyzed frequency range of 0.2-1.1 Hz covered the secondary microseism peak and its higher frequency flank.

For both snapshots the microseism peak was around 0.2 Hz and consisted of both retrograde Rayleigh and Love wave modes with back azimuths distributed between North and West. Above 0.4 Hz, however, higher surface wave modes started to dominate with a much wider range of back azimuths ($\sim$160°). The Smith & Dahlen [1973] equation for surface wave anisotropy fits well to most of the data and we estimate spectra of its isotropic and anisotropic parameters.

According to a bootstrap test the Rayleigh waves exhibit substantially higher 2θ anisotropy in November below 0.6 Hz, increasing from 3-7% to 6-11%. The isotropic part of the phase velocity dispersion for Love and Rayleigh waves, meanwhile, match for both snapshots. We also observe a stable fast direction of NNW-SSE for Love and Rayleigh waves which is aligned with the preferred orientation of shallow (<300 m) and deeper (1000 m) fault systems in the area. We speculate that these observations might be due to geomechanical effects in the reservoir interval caused by increased pore pressure.

Acknowledgments The authors thank H. .R. Kuensch from the seminar for statistics at the ETH Zurich for helpful advice. The Low Frequency Seismic Partnership is thanked for providing the seismic data and supporting this research. Simon Lloyd is thanked for help with the calculation of sensitivity kernels. Our gratitude also extends to Victor Tsai and two anonymous reviewers for greatly improving the quality of this paper through their valuable comments and constructive criticism.

Chapter 4

Rayleigh and Love wave anisotropy in southern California using seismic noise

Abstract We use three-component frequency-wavenumber analysis to study one year (2012) of ambient seismic noise from the Southern California Seismic Network. The rich statistics from the continuous data stream allow estimating azimuthal anisotropy with unprecedented precision. Statistically significant 2θ and 4θ anisotropy for Rayleigh waves is observed over most of the frequency range 15 to 100 mHz. The estimates are consistent with previous array analyses and shear wave splitting studies in the region. We also show preliminary results for Love waves but ambient Love wave illumination in Southern California may not be sufficient to reliably constrain anisotropy.

This chapter is published in Geophysical Research Letters *(doi: 10.1002/2013GL 058518). Authors: Nima Riahi, Erik H. Saenger.*

4.1 Introduction

Seismic anisotropy in the crust and mantle can elucidate geodynamic processes in the Earth (e.g. Long [2013], and references therein). One observational constraint on such anisotropy is the azimuthal variation of surface wave phase velocities. Surface waves offer the advantages of depth resolution [Wüstefeld et al., 2009] and illumination of areas with low seismicity. Smith & Dahlen [1973] showed that for laterally homogeneous media a weak but otherwise arbitrary anisotropy at depth results in relatively simple azimuthal perturbations of phase velocity: a sum of two sinusoids with 180° and 90° periodicity, typically dubbed 2θ and 4θ anisotropy. These parameters may be inverted for a laterally homogeneous and anisotropic subsurface model [Montagner & Nataf, 1986].

Azimuthal anisotropy has been estimated at various scales using surface wave tomography with earthquakes [Trampert & Woodhouse, 2003, Zhang et al., 2007, Deschamps et al., 2008] or noise [Fry et al., 2010, Gallego et al., 2011, Pawlak et al., 2012, Moschetti et al., 2010]. Tomographic techniques effectively use only a limited amount of data to constrain anisotropy at any given point on the surface. Uncertainty in those estimates is accordingly high. Alvizuri & Tanimoto [2011] analyzed teleseismic Rayleigh waves on the Southern California Seismic Network (SCSN) with array beamforming. Using 190 events no 2θ anisotropy could be detected below 30 mHz and 4θ anisotropy was not detected at all despite evidence for it in the region [Montagner & Tanimoto, 1990, Trampert & Woodhouse, 2003, Deschamps et al., 2008].

In this work we demonstrate how using three-component frequency-wavenumber (FK) analysis on ambient seismic noise [Riahi et al., 2013a] (chapter 3 of this thesis) allows for a much more sensitive detection, and precise estimation, of surface wave anisotropy. Statistical significance is assessed with F-tests and the estimation precision due to the uncertainty of the array processor is captured by bootstrapping. The study highlights the potential offered by array processing of ambient noise to provide high-precision estimates of array-averaged seismic parameters.

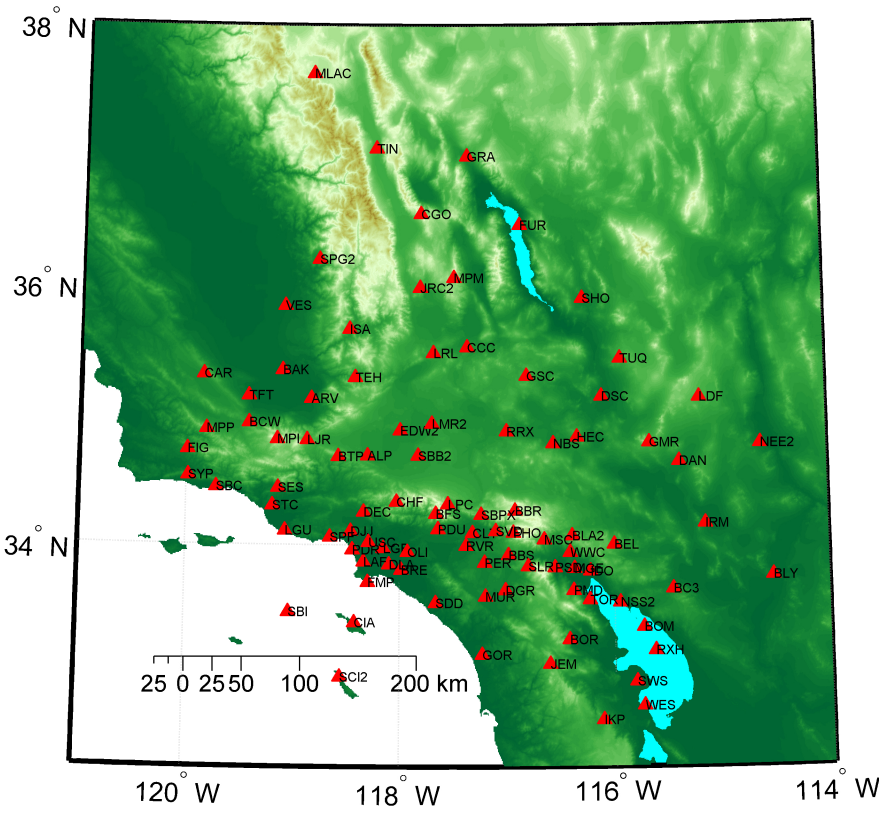

FIGURE 4.1: The locations of the SCSN stations used to analyze data from December 2012 are shown with red triangles (the list of stations used varied based on monthly quality checks).

4.2 Data and method

We study ambient seismic data recorded by the Southern California Seismic Network (SCSN) shown in Figure 4.1. Data from vertical channels of this network has been used for array studies in the past, both to analyze teleseismic events [Tanimoto & Prindle, 2007, Alvizuri & Tanimoto, 2011] as well as ambient surface and body waves [Gerstoft et al., 2006, Gerstoft & Tanimoto, 2007, Gerstoft et al., 2008b]. We refer to Alvizuri & Tanimoto [2011] for a more detailed analysis of the SCSN network as a seismic array.

In this study, we use all three-components from the high-gain, long period channels (LHZ, LHE, LHN) in the frequency range 10 to 100 mHz. The raw data of each station was conservatively quality checked for each month of the year 2012. If a significant portion of the waveform contained gaps, instrument failures, or narrow band spectral peaks, the station was excluded for that month. The three-component beamforming method is a relatively straight-forward generalization of

conventional single component beamforming and is described in detail by Riahi et al. [2013a] (see chapter 3). We only provide a very brief summary of the processing steps here. When a polarized plane wave impinges on a three-component array, it will impose relative phase and amplitude variations between the $3N$ array channels. This variation is described in the Fourier domain by a $3N$ dimensional response vector $\mathbf{k}$, the Kronecker product of two complex-valued vectors:

$$\mathbf{w}(\mathbf{k}, \xi) = \mathbf{c}(\xi) \otimes \mathbf{a}(\mathbf{k}) \,. \tag{4.1}$$

$\mathbf{c}(\xi)$ is a complex-valued three element vector of relative amplitudes and phase shifts between the East, North, and Vertical component, thus defining a polarization ellipse represented by ξ. The complex-valued N element vector $\mathbf{a}(\mathbf{k})$ describes the relative phase shifts between the N station locations due to the wave vector $\mathbf{k} = \frac{f}{v} \cdot \mathbf{n}$, where f is frequency, v is phase velocity, and $\mathbf{n}$ is the wave propagation direction. The factor $\mathbf{c}$ can be tuned to represent transverse motion (Love waves) or elliptical motion within the propagation plane (Rayleigh waves). The amplitudes and relative phases between all $3N$ channels of the array are captured by the spectral density matrix:

$$\mathbf{S}_{3\mathrm{C}} = \langle \mathbf{s}_{3\mathrm{C}} \cdot \mathbf{s}_{3\mathrm{C}}^{\dagger} \rangle \,, \tag{4.2}$$

where $\mathbf{s}_{3\mathrm{C}}$ is a column vector containing the $3N$ Fourier amplitudes of all channels, $\cdot^{\dagger}$ means conjugate transpose, and $\langle \cdot \rangle$ indicates time averaging. The beamformer output is:

$$R(\mathbf{k}, \xi) = \mathbf{w}(\mathbf{k}, \xi)^{\dagger} \mathbf{S}_{3\mathrm{C}} \mathbf{w}(\mathbf{k}, \xi) \,. \tag{4.3}$$

Back azimuth, phase velocity and polarization of dominant wave trains can now be estimated by searching for maxima in $R(\mathbf{k}, \xi)$. We compute Fourier amplitudes on synchronous 512 s windows of all $3N$ channels between 10-100 mHz. The analyzed windows started on January 1$^{\mathrm{st}}$ and were advanced by 256 sec, scanning the entire year 2012. The spectral density matrix $\mathbf{S}_{3\mathrm{C}}$ is estimated by averaging over 11 consecutive windows, with one estimate made every 6 windows. The data is thus characterized every 25 min, using 51 min worth of data each time.

A radial wavenumber grid is used with wavenumber intervals of 8.1310^{-5} km^{-1} and angular intervals of 5°. The search includes one SH polarization state and 11 Rayleigh polarization states with varying ellipticity and orientation of motion (prograde/retrograde). At each window the back azimuth, phase velocity, and polarization of sufficiently strong maxima in R are stored (the number of retained detections varies by signal quality).

4.3 Ambient surface wave summary

For the analyzed year 2012 the array analysis yielded between about $11{\cdot}10^3$ to $33{\cdot}10^3$ detections per Fourier bin and polarization type. To investigate anisotropy, we study the distribution of phase velocity as a function of azimuth for each frequency bin. Figure 4.2 shows histograms of phase velocity against back azimuth of detected Rayleigh wave trains for 22, 43, 61, and 96 mHz. Each graph shows a 2D-histogram with the color indicating the logarithm of the counts made in any bin. The statistical variation in estimated velocities is similar for different back azimuths and a pattern is apparent, in particular for frequency bins 43, 61, 96 mHz. Graphs as those in Figure 4.2 were computed for all 46 frequency bins from 11-100 mHz but the four bins shown here are representative in terms of illumination and variability. The dashed red line is discussed in the next section.

Figure 4.3 shows the same histograms for the Love wave detections, which numbered between $5{\cdot}10^3$ to $8{\cdot}10^3$ depending on frequency. Below 40 mHz Rayleigh waves dominated over the Love waves making their detection less reliable, which is why these frequencies are not shown.

4.4 Estimation of anisotropy parameters

For a laterally homogeneous half-space with anisotropic seismic properties, surface wave phase velocities would at first order be subject to azimuthal variation as follows [Smith & Dahlen, 1973]:

$$v(\theta) = a_0 + a_1 \cos(2\theta) + a_2 \sin(2\theta) + a_3 \cos(4\theta) + a_4 \sin(4\theta)\,, \qquad (4.4)$$

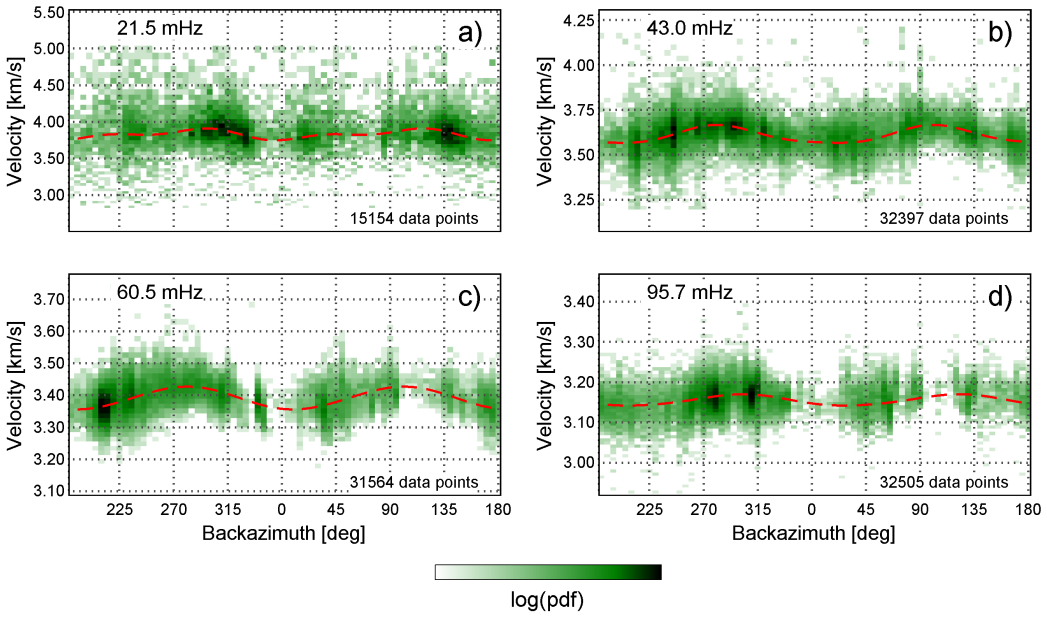

FIGURE 4.2: 2D histogram of phase velocity vs. back azimuth over all Rayleigh wave detections at *(a)* 21.5 mHz, *(b)* 43.0 mHz, *(c)* 60.5 mHz, *(d)* 95.7 mHz. Darker shades indicate higher counts (logarithmic colorscale). The red dashed line is the best robust fit to the anisotropy model. The variability of the velocity estimates ranges from 0.7 km/s (21.5 mHz) to less than 0.2 km/s (above 60 mHz).

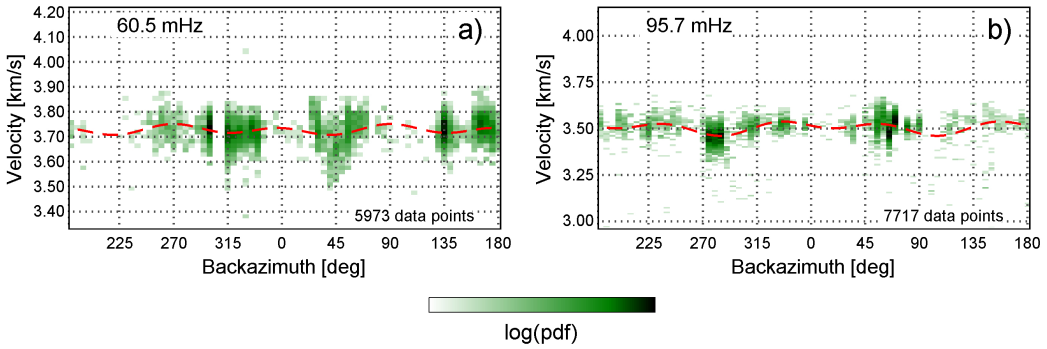

FIGURE 4.3: The same visualization as in Figure 4.2 but for Love wave detections at 60.5 mHz *(a)* and 95.7 mHz *(b)*. Note how illumination is drastically reduced compared to Rayleigh waves. The dashed red line is the best robust fit to the anisotropy model.

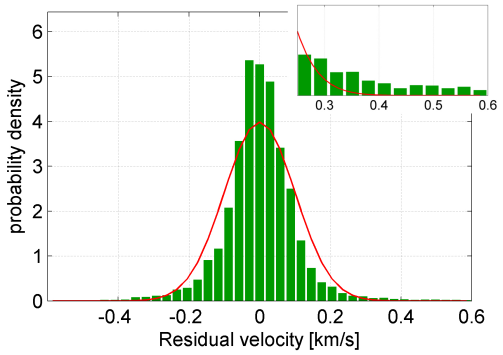

FIGURE 4.4: The distribution of residuals from best fit of eq. 4.4 to the Rayleigh detections at 43 mHz (green bars). The red line is the theoretical normal distribution with same variance. The residuals are more centered and have heavier tails.

where v is surface wave phase velocity [km·s^{-1}], θ is the azimuth of propagation clockwise from North, and a_i are five parameters that depend on the subsurface. We fit the anisotropy model eq. (4.4) using the estimates of phase velocity and back azimuth at every frequency bin and surface wave type (dashed lines in Figures 4.2 and 4.3). As shown later, the distribution of phase velocities at any given back azimuth is generally more heavy-tailed than for a normal distribution. The heavy-tails can bias conventional least-squares fitting procedures and we therefore use the robust ℓ_1 norm minimization approach that minimizes the sum of absolute deviations [Bloomfield & Steiger, 1983]. Figure 4.4 gives the distribution of residuals for the Rayleigh wave detections at 43 mHz where a large number of reliable detections was made. The distribution confirms the non-normality of the errors and similar distributions were observed for frequencies below 53 mHz and between 66-82 mHz.

Using the model parameters $a_0, \ldots, a_4$ we estimate the isotropic phase velocity (a_0), the magnitude of 2θ and 4θ anisotropy ($b_{2\theta} = \sqrt{a_1^2 + a_2^2}$ and $b_{4\theta} = \sqrt{a_3^2 + a_4^2}$, respectively) and the azimuth of fastest surface wave velocity. The statistical variability of the measured phase velocities will be inherited by these anisotropy measures. This variability should be low since the number of model parameters (five) is much smaller than the number of data points used to constrain them (thousands). We assess the variability using a bootstrap approach [Efron & Tibshirani, 1993]: the N data points are randomly sampled with replacement to produce a resample of the same size N. The anisotropy measures are estimated for this set and the entire procedure is repeated $B = 100$ times, which

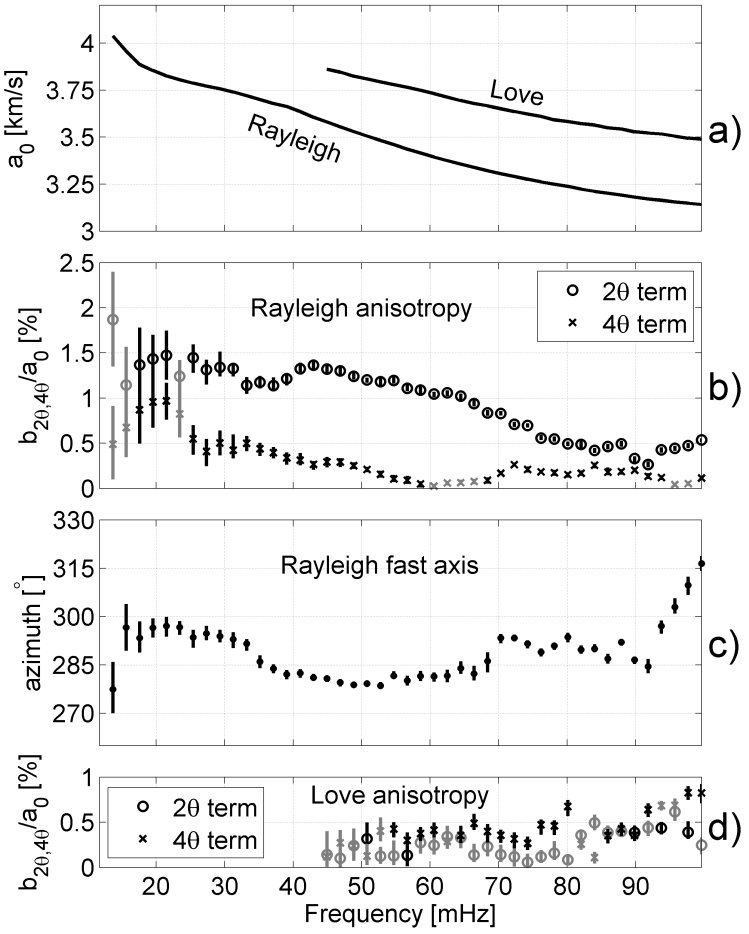

FIGURE 4.5: *(a)* Fitted isotropic phase velocity term (a_0). *(b)* Azimuthal anisotropy parameters for Rayleigh waves. The vertical bars represent the bootstrap variability range. *(c)* Azimuth of fast axis of Rayleigh waves. *(d)* Azimuthal anisotropy parameters for Love waves. Gray markers indicate insignificant terms.

approximates the sampling distribution of the estimated anisotropy parameters given the variability in the velocity estimates of the year 2012. The value for B is defined heuristically as the smallest value above which the distribution stabilizes.

The procedure is repeated for every frequency bin and surface wave type and the final result is summarized in the spectra in Figures 4.5a-d. The vertical lines show the range of bootstrap estimates. No vertical lines are visible for the isotropic phase velocity spectrum (Figure 4.5a) because the variability is smaller than the line width. The Rayleigh wave mode there is consistent with previous studies in Southern California [Alvizuri & Tanimoto, 2011].

The range of bootstrap estimates in Figures 4.5b and 4.5d merely provides an idea on the precision of anisotropy estimates if the full model in eq. (4.4) is used but

does not test whether simpler models would explain the data better. To test if 2θ or 4θ terms are significantly constrained by the data we use a sequential F test (see also appendix D). We compare fits between the full model (2θ+4θ), one with parameters a_0, a_1, a_2 (2θ model), one with parameters a_0, a_3, a_4 (4θ model), and an isotropic model with only a_0 (0θ model). Each model's misfit to the data is captured by the sum of its squared residuals: $SSR = \sum_{i=1}^{N}(v_i(\theta_i) - \hat{v}_i(\theta_i))^2$, where $\hat{v}$ is the phase velocity estimate of the best fit of that model. Increasing the number of parameters of a model typically reduces residuals of the best fit and a direct comparison of SSR measures is therefore misleading. Assuming that a simple model i (e.g. 2θ) is true and nested in a more complicated model j (e.g. $2\theta + 4\theta$) the F statistic offers a means to see this:

$$\bar{F}_{ij} = \frac{(SSR_i - SSE_j)/(k_j - k_i)}{SSR_j/(N - k_j - 1)}\,, \tag{4.5}$$

where k_i and k_j are the number of parameters of models i and j. The F statistic measures the improvement in the sum of squared residuals per additional parameter in the bigger model, normalized by an estimate of the variability of the data. It follows a Fisher distribution with degrees of freedom $\nu_1 = k_j - k_i$ and $\nu_2 = N - k_j - 1$ [Miller, 1990] (for large numbers of N the distributional character of the data points becomes less important. If under the simpler model the probability of observing a value at or above the observed F statistic is less likely than a certain threshold probability, say p=0.01, then the simple model is rejected in favor of the richer model. If the 2θ model or 4θ model better fits the data than the 0θ model, we respectively conclude that 2θ or 4θ anisotropy is significantly constrained. To conclude that both terms are present in the data we additionally require that the 2+4θ model fits better than the 2θ or 4θ model. This procedure is applied to the data in every frequency bin using a threshold level of p=0.01. The gray markers in Figures 4.5b and 4.5d indicate statistically insignificant terms.

As seen in Figure 4.5b), the variability of the Rayleigh wave anisotropy estimates above ∼35 mHz is almost negligible, as indicated by the small vertical bars at the measurement points (<0.1%). Second, both 2θ and 4θ anisotropy are statistically significant over much of the frequency band above 16 mHz. Only the frequency bands 60-70 mHz and 95-98 mHz show insignificant 4θ anisotropy. Note that anisotropy as small as 0.1% was significantly detected in the data. Furthermore,

the fast direction for Rayleigh waves (Figure 4.5c) is mostly contained between 275° and 300°, with a more Westwards orientation (278..285°) around 40-70 mHz.

Figure 4.5d shows the results for Love wave anisotropy. Most frequencies have no significant 2θ anisotropy and also the 4θ term is often insignificant. The reliability of the Love wave results are discussed below.

4.5 Discussion

As summarized in Figure 4.5b, one year of ambient seismic noise provides evidence of significant Rayleigh wave 2θ and 4θ anisotropy over most of the frequency band 15-100 mHz. Anisotropy magnitudes of less than 0.2% are significantly constrained. This extends a previous study of Alvizuri & Tanimoto [2011] based on teleseismic arrivals over a nine-year period where Rayleigh wave 2θ anisotropy from 30-60 mHz was significantly constrained but 4θ anisotropy could not be detected between 9-60 mHz. Compared to hand-picked teleseismic surface wave arrivals, ambient noise is expected to have a lower signal-to-noise ratio and therefore provide less reliable estimates. We attribute the increased precision to the larger amount of data points (up to tens of thousands as opposed to hundreds) and broad illumination provided by the use of ambient seismic data. Consequently, for Love waves where the number of detections is only in the thousands and illumination is poorer, our anisotropy estimates are less conclusive.

The fact that 4θ anisotropy is found for Rayleigh waves is consistent with global tomography for Southern California in the analyzed frequency range [Montagner & Tanimoto, 1990, Trampert & Woodhouse, 2003, Deschamps et al., 2008]. The fast axis revealed in this study (278-284° in the range 40-60 mHz) is close to the direction given by Alvizuri & Tanimoto [2011] for a similar frequency band, albeit trending slightly more westwards. A shear-wave splitting study in the region by Polet & Kanamori [2002] found similar fast directions but we note that the connection between shear-wave splitting and surface wave anisotropy is not well understood [Kosarian et al., 2011]. For Rayleigh waves the results confirm that 2θ anisotropy has higher magnitude than the 4θ type, as argued by Montagner & Nataf [1986].

The bootstrap procedure only captures variability of parameters assuming a fixed 2θ+4θ model. The approach does not account for deviations of the true velocity-vs-azimuth model from that of eq. (4.4) or systematic bias in the phase velocity and back azimuth estimations. Both errors may occur due to lateral heterogeneities or multipathing [Rost & Thomas, 2002] and are hard to quantify. In particular, Alvizuri & Tanimoto [2011] considered multipathing possible at the SCSN above 60 mHz. For the same reasons, also the F test must be interpreted as being approximate and our choice of a 99% confidence level should not be confused as confidence in the test's strict admissibility.

4.6 Conclusions

We find significant 2θ and 4θ Rayleigh wave anisotropy in Southern California Seismic for most frequencies between 15-100 mHz. Our analysis uses three-component frequency-wavenumber analysis of ambient seismic noise from the year 2012. The broad illumination and large data volume make the approach both precise and sensitive: above 35 mHz the anisotropy estimation variability as well as the magnitude of the weakest significant detections is about 0.1%. Results for Love waves are less conclusive, probably because of the substantially lower illumination seen by the array processing.

Our observations confirm regional and global studies that found 4θ Rayleigh wave anisotropy in Southern California in this frequency range. We conclude that given sufficient ambient illumination, array processing of ambient noise on dense sensor networks can be a powerful tool to constrain surface wave anisotropy.

Acknowledgments The authors are grateful to Peter Gerstoft for helpful comments. The seismic data used in this study was made available by the Southern California Earthquake Center (SCEC), which is funded by NSF Cooperative Agreement EAR-0529922 and USGS Cooperative Agreement 07HQAG0008.

Chapter 5

Discussion and conclusion

In the introduction, we described the value that ambient seismic wave field studies offer to geophysical investigations. This thesis contributes methodological developments to the problem of calculating and interpreting passive spectral attributes as well as to applied three-component array analysis of seismic noise. We review key aspects from chapters 2, 3, and 4 and discuss a promising direction for future development and applications.

Passive seismic attributes In chapter 2 we developed a strategy that jointly analyzes the spectral density statistics of seismic power and polarization on a processing level. For data from a producing tight-gas field in Wyoming this revealed evidence of azimuthal diffusivity only during quiet time periods and in a specific frequency band (1-3.5 Hz). We argue that such a polarization diffusivity is a necessary requirement (in particular for an industrial environment) to avoid surface waves which above 1 Hz will hardly carry any information on a reservoir target at 2.5 km depth. This argument is not sufficient to avoid surface waves but we make the case that it increases the likelihood of picking up seismic signatures from depth. The strategy is complemented on the interpretation level by a quantitative framework to describe attribute-to-target correlations for many different attribute parameterizations and different targets.

We applied our approach to test the hypothesis that hydrocarbon reservoirs affect passive seismic attributes on the surface. We find that only an attribute that satisfies the diffusivity constraint also has, firstly, positive correlation to the hydrocarbon reservoir and, secondly, low to insignificant correlation to the

tested confounders. The hypothesis could thus not be rejected, however, the exact mechanism to produce the hypothesized ambient wave field modifications is not well known and there are still many confounders to be tested. For further empirical studies of the hypothesis, we suggest to apply our strategy to more oil or gas fields, in particular for cases where the near surface and production infrastructure are well known.

Noise analysis with three-component arrays As part of this thesis a frequency-wavenumber-polarization processor was developed to study ambient noise on three-component arrays. In chapter 3 we used this processor in conjunction with statistical techniques to perform a time-lapse array analysis of ambient Rayleigh and Love wave anisotropy. Results from data acquired above an underground gas storage (UGS) reservoir during its low and high pore pressure states reveal substantial anisotropy variations in the range 0.4-0.6 Hz. This gives rise to the hypothesis that pore pressure variations in the reservoir may be able to accentuate existing fault-system related azimuthal anisotropy. Ambient seismic surface waves might therefore be used to monitor UGS pore pressure, a task that conventionally requires monitoring wells. Future research could attempt to verify the hypothesis through geomechanical modeling or to improve the sensitivity to the reservoir using higher resolution techniques such as the double-beamforming technique [Boué et al., 2013] or interferometry.

In chapter 4 we presented a three-component array analysis of ambient noise waveforms from the Southern California Seismic Network. Using only one year of seismic noise yields sufficient data to estimate regional 2θ and 4θ Rayleigh wave anisotropy between 15-100 mHz with unprecedented precision. Our results are consistent with other anisotropy studies in the area. Regional sensor networks where this method can be applied are increasing, e.g. the Transportable Array component of USArray (US), RESIF (France), Hi-net (Japan). Array beamforming of ambient seismic noise on those datasets could therefore improve the existing constraints on crustal and upper mantle dynamics for different locations.

Using all three seismometer components for array processing is still uncommon in seismology and the studies in chapters 3 and 4 are therefore a contribution in a direction for which there is a recognized need in seismology [Koper & Ammon, 2013]. Our array processing implementation uses simple beamforming

(i.e. no high-resolution technique) and does not attempt to correct for many real world complications, as discussed in appendix B. While unrealistic or implausible detections were made occasionally, the processor gave overall stable and consistent estimates of average surface wave phase velocities, in particular for the Southern California study. The tolerance of the array studies against invalid assumptions are interpreted as the beneficial effect of the law of large numbers. First, between 76 and 98 sensors were used (depending on data quality) for the array processing. This relatively large number of array elements helps enhancing coherent signals and mitigate site effects. In the case of Southern California, the array sampled over several wavelengths which reduces effects of lateral heterogeneity through spatial averaging. Second, the phase velocity statistics were based on a massive amount of time windows: 1000-2000 per snapshot in the Paris Basin time-lapse study and tens of thousands in Southern California. The sheer number of resulting data points dilutes occasional erroneous detections.

The large number of data points also plays an important part in enabling the improved precision of the anisotropy estimates in chapter 4 over previous studies. In that case, however, the precision was best where the ambient seismic illumination was widest.

Outlook: joint array processing and interferometry Array processing requires coherent signals across a group of sensors to work well. In this sense it is complementary to interferometry that ideally requires diffusivity. The reality of the ambient wave field will lie somewhere between these situations depending on time scales and study method, among other things. We conclude this thesis by emphasizing the improvements that will be gained by developing techniques combining array processing ideas with interferometry. Promising results have been attained in this regard in the field of ocean acoustics through a technique called passive fathometer [Gerstoft et al., 2008a, Harrison & Siderius, 2008, Traer & Gerstoft, 2011, Leroy et al., 2012, Menon et al., 2012]. This combination could boost the often elusive body wave part of the Green's function or it could reduce the required cross-correlation averaging time by avoiding (or exploiting) interferences to accelerate the convergence to the Green's function. When applied to higher frequencies (>1 Hz) the former effect may allow a better use of the ambient seismic wave field for high-resolution imaging and monitoring of subsurface reservoirs.

Shorter convergence times would improve the temporal resolution in applications where seismic noise is used to monitor an earth (or generally an elastic) system.

Appendix A

Modeling detection thresholds of microseismic monitoring networks

A.1 Summary

We present a method to estimate the detection threshold of seismic monitoring arrays, based on the estimated spectral amplitude of microseismic events, and the expected noise level at the recording station. The aim is to develop objective criteria for network optimization, such as the best combination of surface and downhole networks, or the optimal depth of shallow borehole monitoring arrays. We present some simplified generic examples for illustration of the general principles, and discuss the most critical parameters that should be known beforehand for an informed decision on network design.

This appendix was presented at the Society of Exploration Geophysicists (SEG) annual meeting and conference in Las Vegas, 2013, and is published in the proceedings of the conference.
Authors: Alexander Goertz, Nima Riahi, Toni Kraft, Marc Lambert.

A.2 Introduction

Different array geometries for monitoring microseismicity of hydrofrac operations have been the subject of a sometimes heated debate over the past few years (e.g.,

Warpinski [2010], Duncan & Eisner [2010]). Surface monitoring arrays oftentimes have a lower detection threshold owing to greater source-receiver distances and higher levels of cultural noise. While much closer to the action in a typically quieter environment with correspondingly lower detection thresholds, borehole arrays are subject to very constrained geometries due to the prohibitive cost of drilling monitoring wells. In this paper we present a simple methodology to estimate the detection threshold of a microseismic monitoring array. We compare estimates of the anticipated spectral amplitude of an earthquake of given moment magnitude with expected ambient noise levels either at the surface or in a borehole, based on representative broadband recordings of surface ambient noise. The statistical description of the noise level allows calculating a probability of detecting an event of given magnitude. By calculating the surface wave eigenfunctions for a given velocity model, we can estimate the frequency-dependent decay of the noise amplitude with depth. From the latter, we can derive an estimate of either a threshold magnitude for a given deployment depth, or a minimum deployment depth for a given maximum magnitude threshold. We also consider the fact that stacking a multitude of densely spaced surface receivers can effectively lower the detection threshold of such arrays. While we concentrate on the seismic noise for the detection threshold, the sensitivity of the used recording equipment has to be considered as well if it is to be expected that the seismic noise could be lower than the instrument noise. Aside from the detection threshold, we can also estimate the anticipated dominant frequency of microseismic events and frequency ranges that might be masked by typical features of cultural noise. These estimates can be used to select the proper bandwidth of recording equipment with respect to instrument response and sampling rate. The method aims at providing simple yet objective criteria for the design of frac monitoring networks, and defining the most relevant medium parameters that should be known a priori in the network design process. For illustration purposes, we present a generic example, loosely based on a shale play.

A.3 Spectral amplitude of microearthquakes

We compute the source spectrum using the model of Brune [1970, 1971]:

$$A(f) = \frac{\Omega_0}{(1 + \frac{f}{f_C})^2} \, . \tag{A.1}$$

The low-frequency displacement spectral amplitude Ω_0 relates to the seismic moment [Boatwright, 1984] as

$$M_0 = \frac{4\pi}{F^S} \sqrt{\rho_s \rho_r \beta_s^5 \beta_r} \, R \, \Omega_0 \, , \tag{A.2}$$

where R denotes the distance of the travel path between source s and receiver r. Letters ρ and β denote the density and shear wave velocity, respectively. The factor F^s depends on the source radiation pattern and the free-surface amplification. For the sake of simplicity, we assume a constant factor of 1.1, based on values given by Deichmann [2006]. We translate the seismic moment to moment magnitude using the definition of Hanks & Kanamori [1979],

$$M_W = \frac{2}{3} \log_{10} M_0 = 6.07 \ \text{[Nm]} \, . \tag{A.3}$$

The corner frequency f_C scales with magnitude and depends on the stress drop $\delta\sigma$. Assuming the source spectra to represent S arrivals, we use [Edwards et al., 2010]

$$\Delta\sigma = M_0 \left(\frac{f_C}{0.49\beta} \right) \, . \tag{A.4}$$

To consider attenuation, we multiply the source spectrum with a factor of $e^{-\frac{\pi f t}{Q}}$, where t denotes the traveltime between source and receiver and Q is the average quality factor of the medium. Finally, we convert the displacement spectrum to velocity by multiplying with $2\pi f$, and square it to obtain a power spectral density (PSD) that can be compared to ambient noise spectra. We do not consider the directional cosines between the receiver orientation and the emergence angle of the arrival, which is equivalent to assuming that the full vectorial character of the wave field is captured with multicomponent receivers.

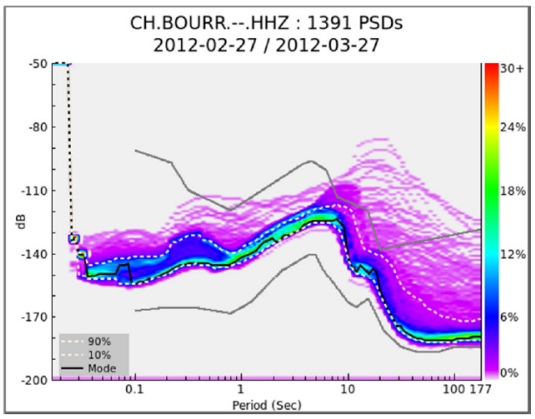

FIGURE A.1: Monthly average noise level calculated using PQLX for a quiet broadband station in a Swiss vault.

A.4 Estimating the ambient seismic noise level

The ambient seismic wave field has a unique spectral character with some common features that can be observed worldwide. This is illustrated in Figure A.1 that shows the monthly average noise level at a quiet broadband station in a vault in Switzerland. The spectral average has been calculated using the method of McNamara & Buland [2004]. The most prominent feature is the ocean microseism peak at about 5 seconds, caused by ocean gravity wave energy coupling into the solid earth. Cultural noise is emerging at periods below 1 s, where we can observe a splitting of the distribution into two modes, reflecting the diurnal variation of human activity. The instrument response and sampling rate of the sensor prevents the analysis of frequencies above 50 Hz. White dashed lines denote the 10th and 90th percentile of the distribution. We can use these percentiles to estimate the probability of detecting a microseismic event.

To account for the possibility of enhancing the detection capability of densely spaced surface arrays with N stations through stacking, we subtract a factor of $\log_{10} N$ from the power spectra in some of the examples $\sqrt{N}$ in amplitude corresponds to N in power). This approach is somewhat simplified in the sense that we consider the stacking to be ideal, i.e., well-behaved, or well-known velocity models and statics.

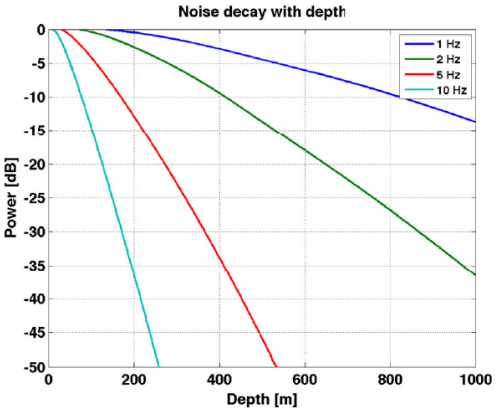

FIGURE A.2: Depth decay of the noise power of fundamental-mode Rayleigh waves for a layered velocity model.

A.5 Noise decay with depth

Ambient noise consists to a large part of surface waves, and decays with depth, depending on the frequency. We estimate this depth decay by forward-calculating the eigenfunctions of the fundamental-mode Rayleigh wave for a layered velocity model using a propagator matrix method [Aki & Richards, 2002]. We notice that over the frequency range of interest for microseismic monitoring (mainly above 5 Hz), the depth decay of surface waves is relevant only in the upper 500 m, and plays a role in selecting the optimal depth of shallow-hole monitoring arrays. At deeper levels, other constituents of the ambient wave field such as body waves or tube waves become the dominant part and the eigenfunction decay is not applicable anymore.

A.6 Shale gas example

To illustrate the above considerations, we use numbers for a typical North American shale play at a target depth of 8000 ft (2440 m). We assume an average bulk quality factor of 100 (optimistically), and a shear wave velocity of 2454 m/s at the source. These numbers are loosely based on values encountered in the Barnett. For the purpose of this calculation, we assume a stress drop of 1 MPa. Densities are estimated using Gardners rule [Gardner et al., 1974]. Our noise estimates are based on 24-hr averages from broadband measurements near a frac operation. The noise level is typically much stronger near the wellhead of the treatment well and decays away from it. Our on-pad example was recorded about 100 m from

a treatment well, and the offpad example is about 2 km away in a quiet open-range area. To illustrate the influence of different parameters, we assume three network configurations: (i) a densely spaced surface array of point receivers, (ii) a sparse network of borehole stations in 300 m deep dedicated monitoring wells, and, (iii) a deep borehole array at target depth within a few hundred feet of the perforation zone. Figure A.3 shows a comparison of the measured ambient noise distribution with source spectra for different magnitudes at the indicated source-receiver distance. Near the well pad (Figure A.3 top), the median noise spectrum is mostly below the M_w -0.5 curve, translating to a >50% probability of detecting an M_w -0.5 event. A M_w 0.0 event would be always detected, and a M_w -1.0 event would be rarely detected ($\geq$10% probability). Around the perimeter of the array, the noise is lower, with less variance, and the listening distance is greater. This greatly increases the probability of detecting a M_w -0.5 event, and a M_w -1.0 event would now be detected with a probability of >50 %. A M_w -1.5 event would never be detected. In order to consider the possibility of stacking a dense surface array we assume an array of 200 stations and subtract a bulk 23 dB from the noise power spectra. This translates into a decrease of the overall detection threshold by almost 1 magnitude unit (Figure A.4). In reality, it is to be expected that the stacking is frequency dependent. While it may work well at the low frequencies, it will be increasingly difficult to achieve a SNR improvement at higher frequencies.

Cultural noise consists predominantly of surface waves and decreases dramatically with depth. We simulate the detection threshold for a sparse shallow borehole network in 300 m deep monitoring wells by multiplying the noise spectra with the forward-modeled eigenfunction for fundamental-model Rayleigh waves. It is possible, if not likely, that in this case the seismic noise drops below the instrument noise level. We assume an instrument noise level of -190 dB, corresponding approximately to state-of-the-art borehole sensors. In reality, the noise level at depth may be higher due to body- or tube waves [Goertz et al., 2011].

Figure A.5 compares the expected noise levels at 300 m depth to theoretic source spectra. We observe that the strong 10-30 Hz noise band that is visible in Figure A.3 and A.4 is now fully attenuated. The overall detection threshold did not decrease compared to the 200-fold stack due to the instrument noise limit which underscores the importance of high instrument sensitivity. For the deep downhole array, we assume receiver levels at distances of 100-500 m to the sources in a similar formation. In this environment, we expect the instrument noise floor to be the

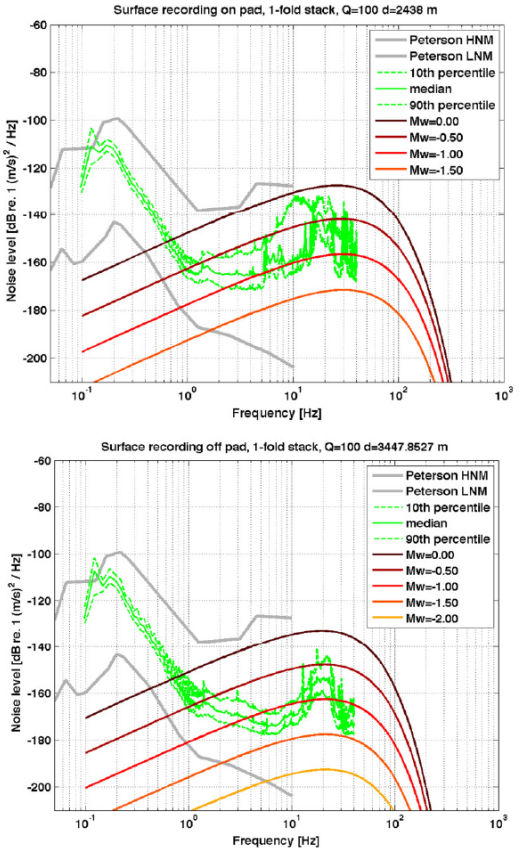

FIGURE A.3: Surface detection threshold close (top) and 2 km away (bottom) from a treatment well. Source spectra in 0.5 magnitude steps (black to orange) are compared to median and 10th/90th percentile of seismic noise observed during a frac (green). Peterson [1993] low noise model (gray) shown for reference.

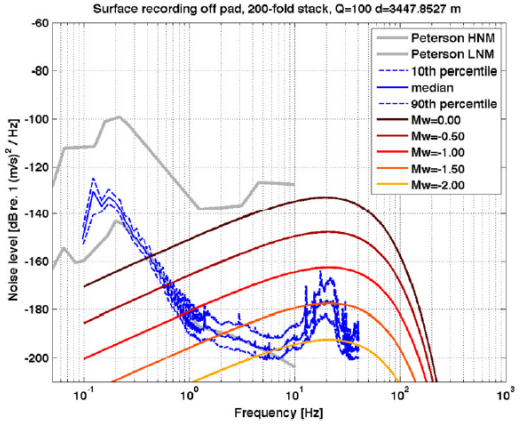

FIGURE A.4: Surface recording away from the pad, assuming a 23 dB improvement in SNR by a 200-fold stack (blue curves).

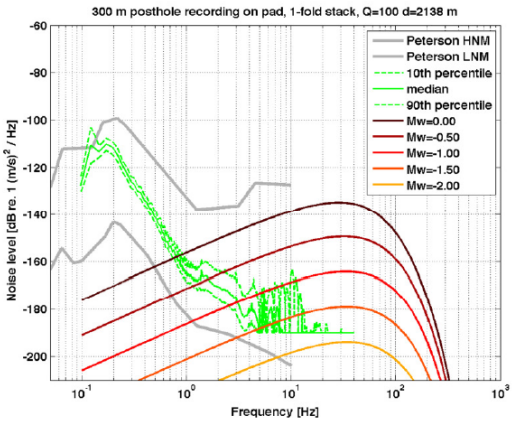

FIGURE A.5: Detection threshold at 300 m depth. Noise spectra are multiplied with the fundamental-mode Rayleigh eigenfunctions to account for the decay of surface waves with depth. Imposing a lower limit of -190 dB accounts for instrument noise.

main factor limiting the detection threshold. We calculate the instrument noise floor for a 15 Hz omnidirectional high-temperature phone after Rodgers [1992]. We observe in Figure A.6 that the detection threshold is now greatly reduced to M_w -3.3 or greater at close distances (100 m), respectively M_w -2.3 at 500 m distance. Note the increased frequency content of the spectra, warranting a faster sampling rate. The advantage of more sensitive equipment (if available) is evident, particularly at the lower frequencies which are important for accurate magnitude determination. The increase of the detection threshold with distance depends on geometrical spreading and attenuation such that, simplified for constant velocity,

$$M_{\min} \propto \log R + \frac{cR}{Q} \,. \tag{A.5}$$

A.7 Comparison to expected seismicity level

We model a magnitude distribution by drawing a fixed number of events from a Gutenberg-Richter statistic with a b-value of 1.5. We equally distribute 10,000 events between 0 2000 ft from the receiver array, and then apply the distance-dependent detection threshold criterion as stated in the above equation. This procedure results in about 1000 recorded events with magnitudes from M_w < -4 to M_w=-1.2 (Figure A.7). This synthetic distribution corresponds roughly to actually recorded event distributions in the Barnett [Maxwell et al., 2006]. To

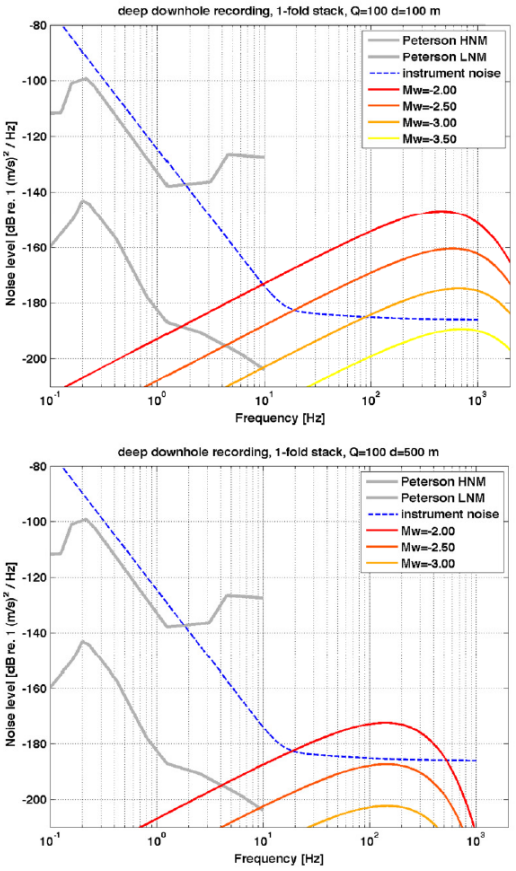

FIGURE A.6: Deep downhole recording at 100 m (top) and 500 m (bottom) distance compared to instrument noise floor (blue).

summarize our results, we plot the obtained best detection thresholds for the surface and shallow borehole array in Figure A.7. We note that neither the 200-station surface array nor the shallow borehole array would have been successful in this particular case. Note however, that this result should not be generalized as it strongly depends on the underlying parameters which vary significantly from play to play and from project to project. We have used generic values that appear reasonable, for demonstration purposes only. For example, surface arrays have been used successfully in the Barnett (e.g. Lakings et al. [2006]). Differences can be due to different target depths, velocities, or Q values. In addition, surface or near-surface borehole arrays may offer better location accuracy which may offset pure detection threshold considerations.

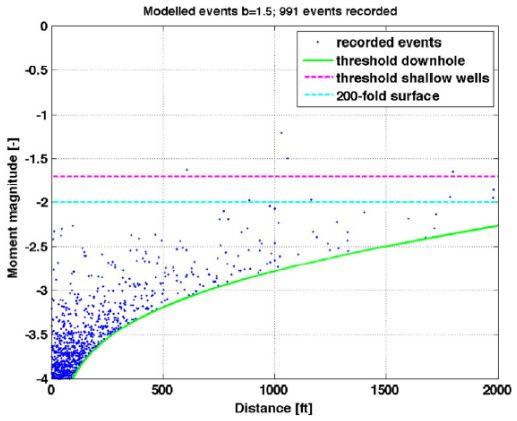

FIGURE A.7: Forward-modeled seismicity cloud (blue) after applying detection threshold for deep downhole recording (green) compared to surface and shallow well array thresholds (cyan and magenta).

A.8 Conclusions

We have developed a fast and simple forward-modeling method to assess the capability of microseismic network configurations to detect a certain magnitude event. Comparison with the distribution of ambient seismic noise levels allows an estimation of the detection probability. We compare modeled detection thresholds with the expected seismicity rate based on Gutenberg-Richter statistics.

Acknowledgments Part of this work was co-funded by Spectraseis and the Low Frequency Seismic Partnership (LFSP). Mid-continent frac monitoring data examples are courtesy of Spectraseis.

Appendix B

Polarization-wavenumber analysis: error sensitivity

The polarization-wavenumber analysis used in chapters 3 and 4 makes several assumptions. This introduces the possibility for systematic errors whenever they are not met. We explicitly state some important assumptions here:

- location and orientation of array sensors are correct,
- only polarizations within or perpendicular to the propagation plane exist,
- coherent arrivals are delta distributed in wavenumber space.

A geophysical interpretation of the latter two assumptions is that of an isotropic, laterally homogeneous subsurface without attenuation. Also no multiphathing or reflections and refractions of a source signal should occur within or in the vicinity of the array. The second assumption was studied in section 3.3.3 and resulting errors were assessed to have a relatively small effect on the estimation of anisotropy parameters. Since the polarization-wavenumber processor of this thesis is a linear operator, one can assume that moderate violations of the above assumptions can be expected to have moderate effects on the results. In this chapter we investigate the effect of random placement and orientation errors on the three-component beamforming performance. A final example will illustrate that correlation between sources is another factor that can affect performance.

We apply the processor to simulations of wave signals under varying degrees of random errors in sensor placement and orientation. Random errors in this context mean that errors are independent between the sensors and invariant over time. This would capture sensor placement problems in the field due to local site conditions or GPS related position retrieval errors. Since this is equivalent to imposing time-invariant random phase errors in the observed Fourier amplitudes the errors could also be considered a result of random short-wavelength heterogeneity or anisotropy in the earth. The reliability of the array processor with respect to such deviations from the assumptions is also tested.

As a reference for the test we use the original array geometry of the study in chapter 3 (Figure 3.1). For a reference frequency at 0.54 Hz we computed synthetic Fourier amplitudes for two impinging plane waves with parameters shown in Table B.1. The time series of the plane waves are modeled as independent Gaussian random processes. We define the signal-to-noise ratio (SNR) at any give frequency as the ratio between the power densities of the signals (-150 dB) and the noise process at that frequency (see appendix C). In the simulation we set SNR=2.

For the scenario we simulate 11 realizations of Fourier amplitudes for all array channels. The power spectral density matrix is computed by averaging over these 11 windows. The wavenumber-polarization processor then estimates the plane wave parameters as described in section 3.3. The entire process is repeated 300 times to assess the performance statistics. Figure B.1 shows the resulting distributions of detections in slowness space with red, green, and blue identifying Love, retrograde Rayleigh, and prograde Rayleigh waves. The sensor locations and orientations are unperturbed in this scenario and the visualization thus represents the array processor performance under ideal conditions (all above assumptions are satisfied). One can verify the estimator's unbiasedness: the detections are well centered around the true plane wave slownesses and the polarizations were correctly identified. The graph also allows us to assess the statistical variability of the estimator, which in this case was about ±200 m/s around the true value, which

f	v [m/s]	ϕ	Type	p [dB w.r.t. $(\mathrm{m/s})^2/\mathrm{Hz}$]
0.54	2400	-15°	Rayleigh, retrograde, H/V=1.7	-150 dB
0.54	3000	-100°	Love	-150 dB

TABLE B.1: Parameters of the two plane waves used for the synthetic study: frequency, phase velocity, back azimuth, polarization, power density.

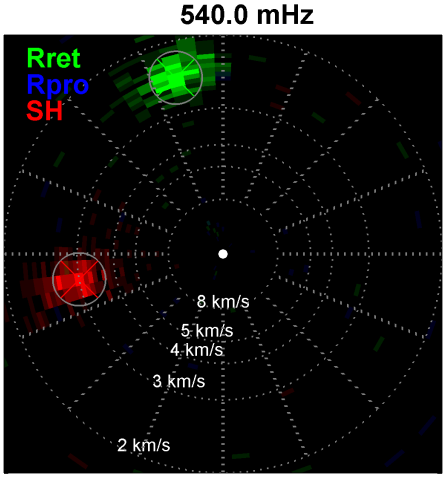

FIGURE B.1: Two plane surface waves with deterministic spatial parameters but stochastic time behavior were analyzed. The wavenumber-polarization processor was applied to 300 realization of these waves. The resulting distribution of detections is shown. Red, green, and blue color indicate a detected Love, retrograde Rayleigh, and prograde Rayleigh wave, respectively. The simulations contained no location and orientation errors.

represents the array processor standard error in phase velocity. The standard error in orientation is about $\pm 8°$.

Next we randomly perturb the location of the array sensors using a 2D normal distribution with standard deviation 15 m, which should capture relatively large localization errors due to insufficient GPS signal. Additionally, the North orientation of the sensors is perturbed with a normally distributed error with standard deviation 10°. Orientation errors larger than 10° are highly unlikely and would require significant impediments during the field operations such as failure of the compass and low visibility to landmark cues from the area surrounding the deployment site. The estimator is then run again on this for 300 repetitions, with the random number generator initialized with the same seed as for the non-perturbed case. This is done to exclude variations due to the generator. Figure B.2a shows the resulting distribution of detections in the slowness space. The results are remarkably similar to the unperturbed case. This is less surprising for the location errors, since 15 m correspond to less than 0.3% of the wavelengths involved. However, the orientation errors are significant and the fact that the results appear essentially unchanged is remarkable.

Figure B.2b shows results where the standard deviation of location errors was 400 m (∼9% of wavelength). A bias starts to become visible towards lower back

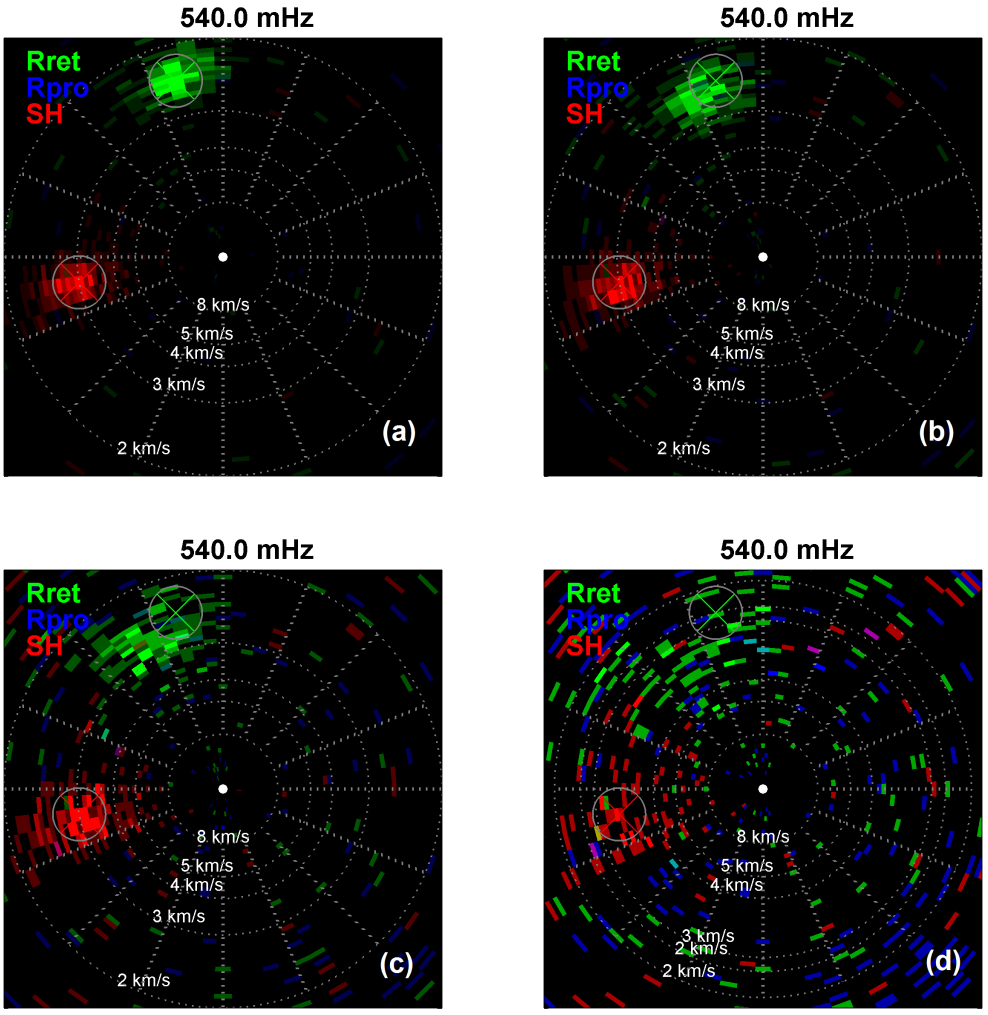

FIGURE B.2: The same visualization as in Figure B.1 is shown for different degrees of location and orientation errors. *(a)* Location errors follow a 2D normal distribution with standard deviation 15 m, orientation errors follow normal distribution with standard deviation 10° (errors fixed for all 300 repetitions). *(b)* Location perturbations were increased to 400 m, orientation error remains 10°. *(c)* Location perturbations were increased to 700 m, orientation error remains 10°. *(d)* Location perturbations were increased to 1000 m, orientation error remains 10°.

azimuths (counterclockwise error). The precision of the estimates does not seem strongly affected. In Figure B.2c results for 700 m error (∼16% of wavelength) are shown. The Rayleigh wave is still visible but shows substantial deviations from the true parameters. The number of spurious detections all over the slowness space has also increased. Finally, Figure B.2 shows results for 1000 m error (∼22% of wavelength). The array processor fails to retrieve the correct wave parameters.

f	v [m/s]	ϕ	Type	p [dB w.r.t. $(m/s)^2/Hz$]
0.54	2400	-15°	Rayleigh, retrograde, H/V=1.7	-150 dB
0.54	2400	-80°	Rayleigh, retrograde, H/V=2.5	-150 dB

TABLE B.2: Parameters of the two plane waves used to test processor performance in the presence of two correlated sources. In one scenario the sources are uncorrelated, in the other they have a correlation coefficient (in time) of 0.9.

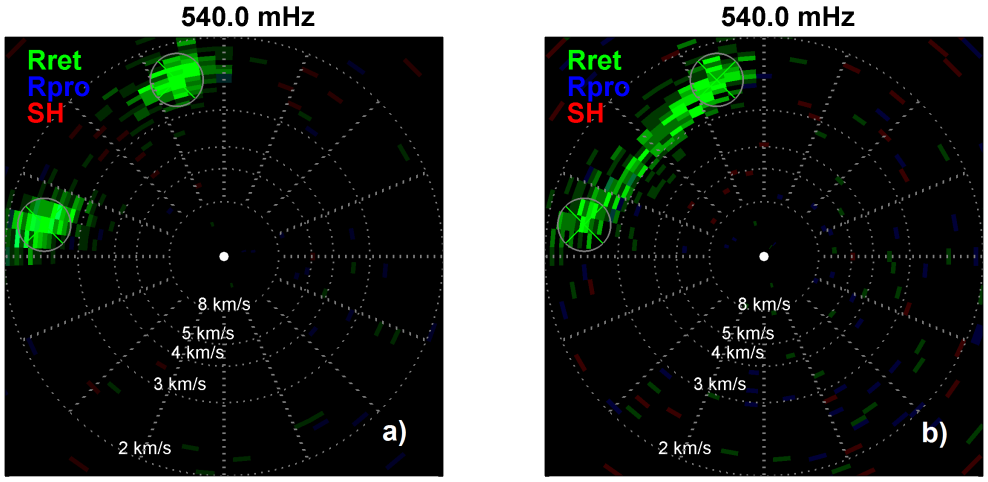

FIGURE B.3: *(a)* Detection statistics for the simulation with two uncorrelated Rayleigh waves (green crosses). *(b)* Same as (a) but the Rayleigh wave signals have a correlation coefficient (in time) of 0.9.

Table B.2 shows a scenario with two Rayleigh wave signals, where again SNR=2. We use this setup to illustrate the effect of source correlation on array performance. The two waves, incident from the Northwest quadrant and separated in azimuth by 65 °, were simulated once as uncorrelated sources and once with a correlation coefficient (in time) of 0.9. Figures B.3a,b compare the detection statistics for the uncorrelated and correlated cases, respectively. The processor is successful for the uncorrelated case but fails for the case of strongly correlated sources: detections are made over the full azimuthal range connecting the two sources with the intermediate detections having lower phase velocity. This example serves to illustrate that source correlation is relevant for array performance. This is a consequence of the limited array aperture and generally performance can improve or deteriorate with increasing correlation. Finally, we note that there are array processing techniques such as MUSIC [Schmidt, 1986] where performance is much less affected by source correlation.

Appendix C

Polarization-wavenumber analysis: SNR

In the polarization-wavenumber analysis used in chapters 3 and 4, for each detected wave train a signal-to-noise ratio (SNR) was estimated. This measure was only briefly touched upon in the chapters. Since it is not commonly described in the literature we derive the SNR estimate in this section.

In a situation where several coherent signals exist and the assumptions of appendix B are met, Schmidt [1986] introduced an elegant formalism to represent the spectral density matrix of all channels of the array. We use his formula to compute a signal-to-noise ratio measure based on those power densities (in the following the notation of section 3.3 is used). Let M be the number of successfully (and presumably correctly) detected coherent plane wave signals impinging on the array. Each of these plane waves has an associated mode vector $\mathbf{w}_i(\mathbf{k_i}, \xi_i)$, $i = 1, \ldots, K$, as shown in eq. (3.6). If we assume that the M plane waves are random processes in time with Fourier amplitudes f_i, then we can write the Fourier amplitude vector observed at the $3K$ array channels as (frequency dependence is implicit):

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{3K} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{w}(\mathbf{k_1}, \xi_1) & \mathbf{w}(\mathbf{k_2}, \xi_2) & \ldots & \mathbf{w}(\mathbf{k_M}, \xi_M) \\ | & | & & | \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_M \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{3K} \end{bmatrix}, \qquad \text{(C.1)}$$

or in shorthand matrix notation

$$\mathbf{s} = \mathbf{W}\mathbf{f} + \mathbf{n}\,, \tag{C.2}$$

where n_i are $3K$ noise processes on the three-components of all K receivers. All signal and noise sources are now assumed to be independent stochastic signals with their power density given by $p_i = \langle f_i^2 \rangle$ and $\sigma^2 = \langle n_i^2 \rangle$ for all i, where $\langle\rangle$ is the expectation operator. There exist methods to handle correlated source and noise signals [Bienvenu & Kopp, 1983, Nadakuditi & Silverstein, 2010] which were not used in chapters 3 and 4. The spectral density matrix of the measured $\mathbf{s}$ now becomes:

$$\mathbf{S} = \langle \mathbf{s}\mathbf{s}^\dagger \rangle = \mathbf{W}\langle \mathbf{f}\mathbf{f}^\dagger \rangle \mathbf{W}^\dagger + \langle \mathbf{n}\mathbf{n}^\dagger \rangle\,. \tag{C.3}$$

Due to the independence assumptions about the signal and noise processes this simplifies to:

$$\mathbf{S} = \mathbf{W}\mathbf{P}\mathbf{W}^\dagger + \sigma^2\mathbf{I}\,, \tag{C.4}$$

with $\mathbf{P} = \langle \mathbf{f}\mathbf{f}^\dagger \rangle$ a diagonal M-dimensional square matrix with p_i as diagonal entries and $\mathbf{I}$ the identity matrix. Since $\mathbf{S}$ is a hermitian matrix its eigenvectors are orthogonal. The MUSIC technique [Schmidt, 1986] is based on the realization that the the $3K - M$ smallest eigenvalues of $\mathbf{S}$ are identical to the noise power σ^2. The corresponding eigenvectors are said to span the noise space and the orthogonal complement space, called the signal space, must then contain the signal mode vectors $\mathbf{w}(\mathbf{k_i}, \xi_\mathbf{i})$. Eq. (C.4) offers the possibility to directly compute the power densities p_i of the impinging plane waves from the spectral density matrix $\mathbf{S}$ (eq. 7 in Schmidt [1986]):

$$\mathbf{P} = (\mathbf{W}\mathbf{W}^\dagger)^{-1}\mathbf{W}^\dagger(\mathbf{S} - \sigma^2\mathbf{I})\mathbf{W}(\mathbf{W}\mathbf{W}^\dagger)^{-1}\,. \tag{C.5}$$

Since the detections of the plane waves are unbiased also the diagonal matrix elements of $\mathbf{P}$, p_i, will be unbiased estimates of the power spectral density of plane wave i. The ratio of these power densities to the incoherent and isotropic noise power density is the measure of signal-to-noise ratio that we want to estimate.

To estimate $\hat{\mathbf{P}}$ and hence $\hat{p}_i$, we use $\mathbf{W} = \hat{\mathbf{W}}$, the matrix constructed from the mode vectors of the detected waves, $\mathbf{S} = \hat{\mathbf{S}}$ the spectral density estimated by block averaging and $\sigma = \hat{\sigma}$, the average over the $3K - M$ smallest eigenvalues of $\hat{\mathbf{S}}$. The signal-to-noise ratio for source i is then estimated as:

$$\mathrm{SNR}_i = \hat{p}_i / \hat{\sigma} \, . \tag{C.6}$$

Appendix D

Model selection using the F test

In chapter 4 the following problem arose: given a set of observations of phase velocity and back azimuth, find a model that best describes the data. Among the models on offer was the simple model that assumes no azimuthal dependency and subsets of the Smith & Dahlen [1973] model that allow for azimuthal variations by a sinusoidal function of 2θ and 4θ. Note that this will produce nested models, i.e. the models with more parameters contain the simpler models. For instance, the Smith & Dahlen [1973] model contains a constant, which corresponds to the no-azimuthal-dependency model.

To choose which of several nested models best describes a collection of observation there is a statistical argument that is based on the so called *F statistic* which we will describe here.

D.1 Formalization of problem

Let y be a random variable defined in the following linear model:

$$y(x_1, \ldots, x_p) = a_1 \cdot x_1 + \ldots + a_p \cdot x_p + e(\sigma^2)\,, \tag{D.1}$$

where a_i are model parameters, x_i are (non-random) predictor variables and e is a random independent and identically distributed variable following a normal distribution with zero mean and variance σ^2. Consider a second model that contains the first model but adds m new parameters for m new predictor variables:

$$y(x_1, \ldots, x_{p+m}) = a_1 \cdot x_1 + \ldots + a_p \cdot x_p + \\ a_{p+1} \cdot x_{p+1} + \ldots + a_{p+m} \cdot x_{p+m} \quad + e(\sigma^2)\,. \tag{D.2}$$

In both cases the model parameters a_i may be found using the least-squares algorithm. However, because the richer model contains the other as a special case, its last squares fit will on average be better than the fit to the simpler model, independent of the merit of the additional parameters. In some sense part of the variability in the data is absorbed into the model parameters, a phenomenon called overfitting. For this problem scenario the F statistic offers a quantitative method to decide if the simpler model should be used. Assume that model M_p (having p variables) and model M_{p+m} (having $p+m$ variables) were both fitted to the data and the sum of squared residuals (difference between observed and modeled values, $y(x) - \hat{y}(x)$) was computed as SSR_p and SSR_{p+m}. The F statistic is defined as

$$\bar{F}_{p,m} = \frac{(SSR_p - SSR_{p+m})/(m)}{SSR_{p+m}/(N-p-m-1)}\,, \tag{D.3}$$

where N is the number of data points used in the fitting. It turns out that if the simpler model M_p is correct, then the F statistic follows a Fisher distribution with degrees of freedom $\nu_1 = (m)$ and $\nu_2 = N - p - m - 1$ (e.g. Miller [1990]). This allows to test for the significance of the simpler model being true. If in a study the F statistic has a value of $\bar{F}$, the probability that this value or a larger one would be observed under a true model M_p is $P = 1 - F(f, \nu_1, \nu_2)$, where F is the Fisher cumulative probability distribution with degrees of freedom ν_1 and ν_2. By setting a threshold for P, say 0.05, one can discard the simple model with 95% confidence if $P < 0.05$.

D.2 Example

We use the data observed in chapters 3 and 4 to simulate random phase velocities as a function of back azimuth under these model assumptions:

$$v(\theta) = a_0 + a_1 \sin(2\theta) + a_2 \cos(2\theta) + a_3 \sin(4\theta) + a_4 \cos(4\theta) + e(\sigma^2) \,, \qquad \text{(D.4)}$$

where we define $e(\sigma^2)$ as a normally distributed random error. In a first example we simulate the following model

$$a_0 = 3, a_1 = 0.04, a_2 = -0.01, a_3 = a_4 = 0 \,, \qquad \text{(D.5)}$$

which corresponds to 1.4% 2θ anisotropy and 0% 4θ anisotropy. Figure D.1a shows the distribution of N=2000 simulated phase velocities, where the imperfect illumination seen in chapter 4 was emulated by reducing the number of detections for certain back azimuths.

We produce 8000 realizations of data clouds as shown in Figure D.1a and fit the 2θ and more complicated 2θ+4θ model each time. For each of the 8000 realizations we compute the resulting F statistic. Figure D.1b shows the simulated distribution of the F statistic (black bars) with the theoretic Fisher distribution for the case where the simpler model was the correct one (red line). Since the data were produced based on the simple 2θ model the histogram matches the theoretic distribution. In Figure D.1c the same distribution was computed for the same model except that the normal error distribution was replaced by a double-exponential or Laplace distribution with similar variance. The Laplace distribution has wider tails and better models occasional outliers. The simulated distribution still matches well with the theoretical distribution, indicating that this property is not very sensitive to the type of error function.

Next we simulate 8000 realization of 2000 phase velocities using the following model:

$$a_0 = 3, a_1 = 0.04, a_2 = -0.01, a_3 = -0.004, a_4 = 0.004 \,, \qquad \text{(D.6)}$$

which corresponds to 1.4% 2θ anisotropy and 0.2% 4θ anisotropy. Figure D.2a shows one realization of 2000 phase velocities. As can be expected by the low magnitude of the 4θ anisotropy the resulting changes compared to Figure D.1a are very subtle. However, the resulting distribution of the F statistic (Figure D.2b)

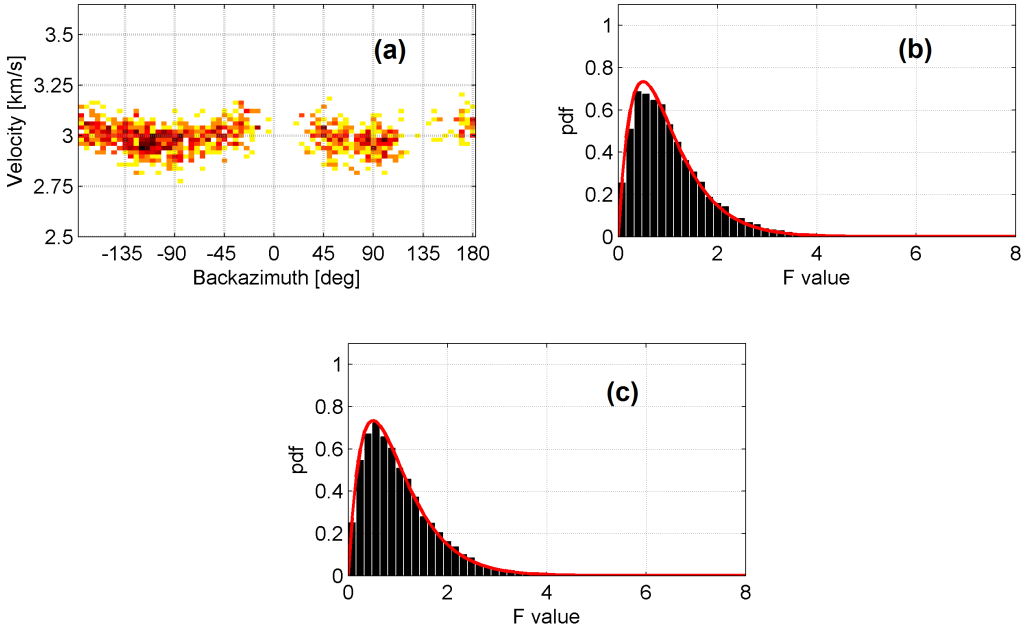

FIGURE D.1: *(a)* Distribution of 2000 simulated phase velocities using a 2θ dependency on back azimuth. The azimuthal coverage was restricted to emulate imperfect illumination in the real world. *(b)* Histogram over 8000 simulated F values (black bars). The red line shows the theoretically expected distribution. *(c)* Same as (b) but the normal error distribution was replaced with a double-exponential error with similar variance (black bars). The F distribution is not very sensitive to the nature of the error distribution.

shows that even for a small complication of the simpler model the simulated F distribution starts to deviate from the theoretical one (red line).

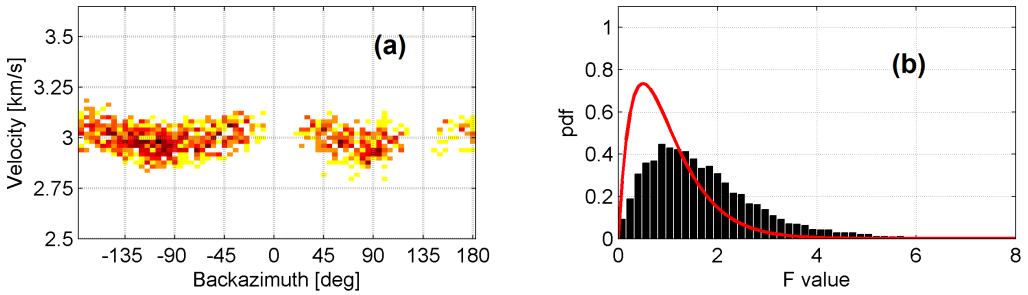

FIGURE D.2: *(a)* As in Figure D.1a, 2000 phase velocities where simulated but in addition to the 2θ a 4θ term was added with about one eighth of the magnitude of the 2θ term. The change is hardly visible. *(b)* Histogram over 8000 simulated F values for a phase velocity distribution as shown in (a) (black bars). Despite the weak 4θ term, the F distribution changes substantially.

Acknowledgements

At the origin of this thesis stood not only a personal desire to engage in academic research but also the encouragement and support of Jean-Pierre Burg and Erik Saenger. They have created and maintained through the three years of the project a stimulating environment of academic freedom. Erik has tirelessly accompanied my research efforts with a keen sense for effective prioritization and I deeply cherish the friendship that has developed over the course of our collaboration. Jean-Pierre, a true enabler, has provided me with a reliable and free atmosphere and and showed a remarkable dedication to help me with issues both small and large. The existence of this thesis is as much due to the commitment of these two men as it is to mine.

Countless exciting and stimulating discussions have affected my work and my thinking. Of direct consequence for this thesis have been conversations with Hans-Roduolf Künsch, Alex Goertz, Peter Gerstoft, and several anonymous reviewers. More generally, but no less important, I have benefited from engaging exchanges with Kees Weemstra, Beatriz Quintal, Paola Sala, Marcel Frehner, Evangelos Moulas, Valerio Poggi, Daniela Hunziker, Jessica Kind, Mara Beinerte, Regula Schälchli, and many more. I thank Gregor Hillers for agreeing to review my thesis as an external referee and for the insightful discussions we had on the field of ambient seismic studies.

My former employer, Spectraseis AG, has supported my decision to enter academic research and has straightforwardly provided me with copious amounts of data which, as the practitioner knows, are a crucial ingredient to all scientific endeavors. I thank Rob Habiger for inviting me to visit their offices in Denver/CO in November 2012 and learning more about current technical challenges facing passive seismic services in the industry. I am also thankful to my former colleagues Marc Lambert, Brad Artman, Paul West, Konrad Cieslik, Brad Birkelo, and many more for engaging discussions.

As a geology-outsider working in the structural geology group at ETH, I have come to appreciate the many outstanding, hard-working researchers in this group. They have all helped to make my stay enjoyable and memorable. I am happy to have had the chance to befriend these good people and be part of their world. Those with whom I shared the office are specifically thanked for their patience in enduring my frequent phone calls and over-the-desk discussions (with the exception, perhaps, of the occasional projectile launched by Marine Collignon).

My gratitude also extends to Arash Riahi for his relentless efforts in *reality-anchoring* me and to Joëlle Turrian for being the great woman she is and for being at my side during the more introvert moments of this work's creation. I thank Marco Venturini for continuously challenging my thoughts on the value of academic research and Dana Frei for our fascinating conversations and for letting me enjoy some insights from her razor-sharp analytical mind. Finally, I am grateful to my parents Iran and Sadegh Riahi for the greatest gift imaginable: unconditional love and unconditional support.

Bibliography

Adam, J. M. C. & Lebedev, S., 2012. Azimuthal anisotropy beneath southern Africa from very broad-band surface-wave dispersion measurements, *Geophysical Journal International*, **191**(1), 155–174.

Aki, K., 1957. Space and time spectra of stationary stochastic wave, with special reference to microtremors, *Bulletin of the Seismological Society of America*, **35**, 415–456.

Aki, K. & Richards, P. G., 2002. *Quantitative seismology*, University Science Books, Sausalito, Calif., 2nd edn.

Ali, M. Y., Berteussen, K. A., Small, J., & Barkat, B., 2010. Low-frequency passive seismic experiments in Abu Dhabi, United Arab Emirates: implications for hydrocarbon detection, *Geophysical Prospecting*, **58**(5), 875–899.

Alvizuri, C. & Tanimoto, T., 2011. Azimuthal anisotropy from array analysis of Rayleigh waves in Southern California, *Geophysical Journal International*, **186**(3), 1135–1151.

Ardhuin, F., Stutzmann, E., Schimmel, M., & Mangeney, A., 2011. Ocean wave sources of seismic noise, *Journal of Geophysical Research-Oceans*, **116**.

Artman, B., 2006. Imaging passive seismic data, *Geophysics*, **71**(4), SI177–SI187.

Bard, P. Y., 1999. *Microtremor measurements: A Tool for site effect estimation?*, pp. 1251–1279, Balkema, Rotterdam.

Bard, P. Y., 2010. *From Non-invasive Site Characterization to Site Amplification: Recent Advances in the Use of Ambient Vibration Measurements*, chap. 5, pp. 105–123, Geotechnical, Geological, and Earthquake Engineering 17.

Bear, L. K., Pavlis, G. L., & Bokelmann, G. H. R., 1999. Multi-wavelet analysis of three-component seismic arrays: Application to measure effective anisotropy at Pinon Flats, California, *Bulletin of the Seismological Society of America*, **89**(3), 693–705.

Beccaletto, L., Hanot, F., Serrano, O., & Marc, S., 2011. Overview of the subsurface structural pattern of the Paris Basin (France): Insights from the reprocessing and interpretation of regional seismic lines, *Marine and Petroleum Geology*, **28**(4), 861–879.

Behr, Y., Townend, J., Bowen, M., Carter, L., Gorman, R., Brooks, L., & Bannister, S., 2013. Source directionality of ambient seismic noise inferred from three-component beamforming, *Journal of Geophysical Research: Solid Earth*, pp. n/a–n/a.

Bienvenu, G. & Kopp, L., 1983. Optimality of High Resolution Array Processing Using the Eigensystem Approach, *Ieee Transactions on Acoustics Speech and Signal Processing*, **31**(5), 1235–1248.

Birkelo, B., Duclos, M., Artman, B., Schechinger, B., Witten, B., Goertz, A., Weemstra, K., & Hadidi, M., 2010. A passive low-frequency seismic survey in Abu Dhabi - Shaheen project, *SEG Technical Program Expanded Abstracts*, **29**(1), 2207–2211.

Birkelo, B., Goertz, A., Cieslik, K., & LaBarre, E., 2011. Locating high-productivity areas of tight gas-sand reservoirs using LF seismic surveys, *EAGE 77th Annual International Meeting, Expanded Abstracts*, p. C016.

Bloomfield, P. & Steiger, W. L., 1983. *Least absolute deviations : theory, applications, and algorithms*, Progress in probability and statistics, Birkhauser, Boston.

Boatwright, J., 1984. Seismic Estimates of Stress Release, *Journal of Geophysical Research*, **89**(Nb8), 6961–6968.

Bokelmann, G. H. R., 1995. P-wave array polarization analysis and effective anisotropy of the brittle crust, *Geophysical Journal International*, **120**(1), 145–162.

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**, 205–227.

Boschi, L., Weemstra, C., Verbeke, J., Ekstrom, G., Zunino, A., & Giardini, D., 2013. On measuring surface wave phase velocity from station-station cross-correlation of ambient signal, *Geophysical Journal International*, **192**(1), 346–358.

Boué, P., Roux, P., Campillo, M., & Cacqueray, B. d., 2013. Double beamforming processing in a seismic prospecting context, *Geophysics*, **78**(3), V101–V108.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E., 2008. Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations, *Science*, **321**(5895), 1478–1481.

Bromirski, P. D. & Duennebier, F. K., 2002. The near-coastal microseism spectrum: spatial and temporal wave climate relationships, *Journal of Geophysical Research*, **107**(B8), ESE5–1–ESE5–ESE5–20ESE5–ESE5–20.

Brune, J. N., 1970. Tectonic Stress and Spectra of Seismic Shear Waves from Earthquakes, *Journal of Geophysical Research*, **75**(26), 4997.

Brune, J. N., 1971. Correction, *Journal of Geophysical Research*, **76**(20), 5002.

Buckingham, M. J., Potter, J. R., & Epifanio, C. L., 1996. Seeing underwater with background noise, *Scientific American*, **274**(2), 86–90.

Buhlmann, P., 2002. Bootstraps for time series, *Statistical Science*, **17**(1), 52–72.

Bush, I. & Crampin, S., 1991. Paris Basin VSPs - case-history establishing combinations of fine-layer (or lithologic) anisotropy and crack anisotropy from modeling shear wave-fields near point singularities, *Geophysical Journal International*, **107**(3), 433–447.

Capon, J., 1969. High-resolution frequency-wavenumber spectrum analysis, *Proceedings of the IEEE*, **57**(8), 1408–1418.

Chapman, M., Liu, E. R., & Li, X. Y., 2006. The influence of fluid-sensitive dispersion and attenuation on AVO analysis, *Geophysical Journal International*, **167**(1), 89–105.

Chavez-Garcia, F. J., Rodriguez, M., & Stephenson, W. R., 2005. An alternative approach to the SPAC analysis of microtremors: Exploiting stationarity of noise, *Bulletin of the Seismological Society of America*, **95**(1), 277–293.

Chitre, M., Kuselan, S., & Pallayil, V., 2012. Ambient noise imaging in warm shallow waters; robust statistical algorithms and range estimation, *Journal of the Acoustical Society of America*, **132**(2), 838–847.

Curtis, A., Gerstoft, P., Sato, H., Snieder, R., & Wapenaar, K., 2006. Seismic interferometry - turning noise into signal, *The Leading Edge*, pp. 1082–1092.

D'Amico, V., Picozzi, M., Baliva, F., & Albarello, D., 2008. Ambient noise measurements for preliminary site-effects characterization in the urban area of Florence, Italy, *Bulletin of the Seismological Society of America*, **98**(3), 1373–1388.

Dangel, S., Schaepman, M. E., Stoll, E. P., Carniel, R., Barzandji, O., Rode, E. D., & Singer, J. M., 2003. Phenomenology of tremor-like signals observed over hydrocarbon reservoirs, *Journal of Volcanology and Geothermal Research*, **128**(1-3), 135–158.

de Ridder, S. A. L. & Biondi, B. L., 2013. Daily Reservoir-Scale Subsurface Monitoring Using Ambient Seismic Noise, *Geophysical Research Letters*.

Debeglia, N. & Debrand Passard, S., 1980. Principaux accidents tectoniques issus des corrlations entre les donnes gophysiques et les donnes de terrain dans le Sud-Ouest du Bassin de Paris, *Bulletin de la Socit gologique de France*, **22**, 639–664.

Deichmann, N., 2006. Local magnitude, a moment revisited, *Bulletin of the Seismological Society of America*, **96**(4), 1267–1277.

Deschamps, F., Lebedev, S., Meier, T., & Trampert, J., 2008. Azimuthal anisotropy of Rayleigh-wave phase velocities in the east-central United States, *Geophysical Journal International*, **173**(3), 827–843.

Draganov, D., Wapenaar, K., & Thorbecke, J., 2006. Seismic interferometry: Reconstructing the earth's reflection response, *Geophysics*, **71**(4), Si61–Si70.

Draganov, D., Wapenaar, K., Mulder, W., Singer, J., & Verdel, A., 2007. Retrieval of reflections from seismic background-noise measurements, *Geophysical Research Letters*, **34**(4), –.

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

Draganov, D., Ghose, R., Ruigrok, E., Thorbecke, J., & Wapenaar, K., 2010. Seismic interferometry, intrinsic losses and Q-estimation, *Geophysical Prospecting*, **58**(3), 361–373.

Draganov, D., Campman, X., Thorbecke, J., Verdel, A., & Wapenaar, K., 2013. Seismic exploration-scale velocities and structure from ambient seismic noise (¿ 1Hz), *Journal of Geophysical Research: Solid Earth*, pp. n/a–n/a.

DuBois, D. P., Wynne, P. J., Smagala, T. M., Johnson, J. L., Engler, K. D., & McBride, B. C., 2004. *Geology of the Jonah Field, Sublette County, Wyoming*, AAPG Studies in Geology.

Duncan, P. M. & Eisner, L., 2010. Reservoir characterization using surface microseismic monitoring, *Geophysics*, **75**(5), 75A139–75A146.

Durand, S., Montagner, J. P., Roux, P., Brenguier, F., Nadeau, R. M., & Ricard, Y., 2011. Passive monitoring of anisotropy change associated with the Parkfield 2004 earthquake, *Geophysical Research Letters*, **38**(13), L13303.

Edwards, B., Allmann, B., Fah, D., & Clinton, J., 2010. Automatic computation of moment magnitudes for small earthquakes and the scaling of local to moment magnitude, *Geophysical Journal International*, **183**(1), 407–420.

Efron, B. & Tibshirani, R., 1993. *An introduction to the bootstrap*, Monographs on statistics and applied probability, Chapman and Hall, New York.

Epifanio, C. L., Potter, J. R., Deane, G. B., Readhead, M. L., & Buckingham, M. J., 1999. Imaging in the ocean with ambient noise: the ORB experiments, *Journal of the Acoustical Society of America*, **106**(6), 3211–3225.

Esmersoy, C., Cormier, V. F., & Toksoz, M. N., 1985. *Three-component array processing*, pp. xviii, 964 p., Executive Graphic Services, United States.

Fäh, D., Kind, F., & Giardini, D., 2001. A theoretical investigation of average H/V ratios, *Geophysical Journal International*, **145**(2), 535–549.

Fäh, D., Kind, F., & Giardini, D., 2003. Inversion of local S-wave velocity structures from average H/V ratios, and their use for the estimation of site-effects, *Journal of Seismology*, **7**(4), 449–467.

Feyerabend, P., 1975. *Against method : outline of an anarchistic theory of knowledge*, Humanities Press, Atlantic Highlands, N.J.

Fleury, R., Charnet, F., Corpel, J., Debrandpassard, S., Gros, Y., & Maget, P., 1997. *Carte gol. France (1/50 000), feuille Romorantin (460).*

Fry, B., Deschamps, F., Kissling, E., Stehly, L., & Giardini, D., 2010. Layered azimuthal anisotropy of Rayleigh wave phase velocities in the European Alpine lithosphere inferred from ambient noise, *Earth and Planetary Science Letters*, **297**(1-2), 95–102.

Gabor, D., 1946. Theory of communication, *Journal of the Institution of Electrical Engineers*, **93**, 429–457.

Gallego, A., Panning, M. P., Russo, R. M., Comte, D., Mocanu, V. I., Murdie, R. E., & Vandecar, J. C., 2011. Azimuthal anisotropy in the Chile Ridge subduction region retrieved from ambient noise, *Lithosphere*, **3**(6), 393–400.

Gardner, G. H. F., Gardner, L. W., & Gregory, A. R., 1974. Formation Velocity and Density - Diagnostic Basics for Stratigraphic Traps, *Geophysics*, **39**(6), 770–780.

Gerstoft, P. & Tanimoto, T., 2007. A year of microseisms in southern California, *Geophysical Research Letters*, **34**(20).

Gerstoft, P., Fehler, M. C., & Sabra, K. G., 2006. When Katrina hit California, *Geophysical Research Letters*, **33**(17).

Gerstoft, P., Hodgkiss, W. S., Siderius, M., Huang, C. F., & Harrison, C. H., 2008a. Passive fathometer processing, *Journal of the Acoustical Society of America*, **123**(3), 1297–1305.

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

Gesch, D. B., 2007. *Chapter 4 - The National Elevation Dataset*, pp. 61–82, Amercian Society for Photogrammetry and Remote Sensing, Bethesda, Maryland.

Goertz, A., Cieslik, K., Hauser, E., Watts, G., McCrossin, S., & Zbasnik, P., 2011. A combined borehole/surface broadband passive seismic survey over a gas storage field.

Goertz, A., Riahi, N., Kraft, T., & Lambert, M., 2012a. Modeling detection thresholds of microseismic monitoring networks, in *SEG conference and meeting*, SEG.

Goertz, A., Schechinger, B., Witten, B., Koerbe, M., & Krajewski, P., 2012b. Extracting subsurface information from ambient seismic noise A case study from Germany, *Geophysics*, **77**(4), 1–19.

Goloshubin, G., VanSchuyver, C., Korneev, V., Silin, D., & Vingalov, V., 2006. Reservoir imaging using low frequencies of seismic reflections, *The Leading Edge*, **25**(5), 527–531.

Green, A. G. & Greenhalgh, S., 2009a. Comment on Low-frequency microtremor anomalies at an oil and gas field in Voitsdorf, Austria by Marc-Andre Lambert, Stefan Schmalholz, Erik H. Saenger and Brian Steiner, *Geophysical Prospecting*, **58**, 335–339.

Green, A. G. & Greenhalgh, S., 2009b. Microtremor spectra: a proven means for estimating resonant frequencies and S-wave velocities of shallow soils/sediments, but a questionable tool for locating hydrocarbon reservoirs, *First Break*, **27**(7).

Hamon, Y. & Merzeraud, G., 2005. *Nouvelles donnes sur le Trias de Sologne (Chemery, Sud-Ouest du Bassin de Paris): Stratigraphie et environnements de depots*, pp. 3–22.

Hanks, T. C. & Kanamori, H., 1979. Moment Magnitude Scale, *Journal of Geophysical Research*, **84**(Nb5), 2348–2350.

Hanssen, P. & Bussat, S., 2008. Pitfalls in the analysis of low-frequency passive seismic data, *First Break*, **26**, 111–119.

Harmon, N., Rychert, C., & Gerstoft, P., 2010. Distribution of noise sources for seismic interferometry, *Geophysical Journal International*, **183**(3), 1470–1484.

Harrison, C. H. & Siderius, M., 2008. Bottom profiling by correlating beam-steered noise sequences, *Journal of the Acoustical Society of America*, **123**(3), 1282–1296.

Hayashi, K., Martin, A., Hatayama, K., & Kobayashi, T., 2013. Estimating deep S-wave velocity structure in the Los Angeles Basin using a passive surface-wave method, *The Leading Edge*, **32**(6), 620–626.

Heisenberg, W., 1958. *Physics and philosophy; the revolution in modern science*, World perspectives, v 19, Harper, New York,, 1st edn.

Herrmann, R., 1996. Computer Programs in Seismology: 3.0, *Saint Louis University, Saint Louis, Missouri*.

Hestholm, S., Moran, M., Ketcham, S., Anderson, T., Dillen, M., & McMechan, G., 2006. Effects of free-surface topography on moving-seismic-source modeling, *Geophysics*, **71**(6), T159–T166.

Hillers, G. & Ben-Zion, Y., 2011. Seasonal variations of observed noise amplitudes at 2-18 Hz in southern California, *Geophysical Journal International*, **184**(2), 860–868.

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

Hillers, G., Graham, N., Campillo, M., Kedar, S., Landes, M., & Shapiro, N., 2012b. Global oceanic microseism sources as seen by seismic arrays and predicted by wave action models, *Geochemistry Geophysics Geosystems*, **13**.

Horike, M., 1985. Inversion of Phase-Velocity of Long-Period Microtremors to the S-Wave-Velocity Structure down to the Basement in Urbanized Areas, *Journal of Physics of the Earth*, **33**(2), 59–96.

Julian, B. R., 1994. Volcanic Tremor - Nonlinear Excitation by Fluid-Flow, *Journal of Geophysical Research-Solid Earth*, **99**(B6), 11859–11877.

Kao, H., Shan, S. J., Dragert, H., Rogers, G., Cassidy, J. F., Wang, K. L., James, T. S., & Ramachandran, K., 2006. Spatial-temporal patterns of seismic tremors in northern Cascadia, *Journal of Geophysical Research-Solid Earth*, **111**(B3).

Kedar, S., Longuet Higgins, M., Webb, F., Graham, N., Clayton, R., & Jones, C., 2008. The origin of deep ocean microseisms in the North Atlantic Ocean, *Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences*, **464**(2091), 777–793.

Kind, F., Fah, D., & Giardini, D., 2005. Array measurements of S-wave velocities from ambient vibrations, *Geophysical Journal International*, **160**(1), 114–126.

Koper, K. D. & Ammon, C. J., 2013. Planning a Global Array of Broadband Seismic Arrays, *Eos, Transactions American Geophysical Union*, **94**(34), 300–300.

Koper, K. D. & de Foy, B., 2008. Seasonal Anisotropy in Short-Period Seismic Noise Recorded in South Asia, *Bulletin of the Seismological Society of America*, **98**(6), 3033–3045.

Koper, K. D., de Foy, B., & Benz, H., 2009. Composition and variation of noise recorded at the Yellowknife Seismic Array, 1991-2007, *Journal of Geophysical Research- Solid Earth*, **114**.

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.

Korneev, V. A., Goloshubin, G. M., Daley, T. M., & Silin, D. B., 2004. Seismic low-frequency effects in monitoring fluid-saturated reservoirs, *Geophysics*, **69**(2), 522–532.

Kosarian, M., Davis, P. M., Tanimoto, T., & Clayton, R. W., 2011. The relationship between upper mantle anisotropic structures beneath California, transpression, and absolute plate motions, *Journal of Geophysical Research-Solid Earth*, **116**.

Lacoss, R. T., Kelly, E. J., & Toksoz, M. N., 1969. Estimation of Seismic Noise Structure using Arrays, *Geophysics*, **34**(1), 21–38.

Lakings, J., Duncan, P., Neale, C., & Theiner, T., 2006. Surface-based Microseismic Monitoring of a Hydraulic Fracture Well Stimulation in the Barnett Shale.

Lambert, M., Schmalholz, S. M., Saenger, E. H., & Steiner, B., 2009a. Reply to comment on Low-frequency microtremor anomalies at an oil and gas field in Voitsdorf, Austria by Marc-Andre Lambert, Stefan M. Schmalholz, Erik H. Saenger and Brian Steiner, *Geophysical Prospecting*, **58**, 341–346.

Lambert, M. A., Schmalholz, S. M., Saenger, E. H., & Steiner, B., 2009b. Low-frequency microtremor anomalies at an oil and gas field in Voitsdorf, Austria, *Geophysical Prospecting*, **57**(3), 393–411.

Lambert, M.-A., Nguyen, T., Saenger, E. H., & Schmalholz, S. M., 2011a. Spectral analysis of ambient ground-motion–Noise reduction techniques and a methodology for mapping horizontal inhomogeneity, *Journal of Applied Geophysics*, **74**(2-3), 100–113.

Lambert, M.-A., Nguyen, T., Saenger, E. H., & Schmalholz, S. M., 2011b. Spectral analysis of ambient ground-motion–Noise reduction techniques and a methodology for mapping horizontal inhomogeneity, *Journal of Applied Geophysics*, **74**(2-3), 100–113.

Lambert, M. A., Saenger, E. H., Quintal, B., & Schmalholz, S. M., 2012. Numerical simulation of ambient seismic wavefield modification caused by pore-fluid effects in an oil reservoir, *Geophysics*, **78**.

Landes, M., Hubans, F., Shapiro, N. M., Paul, A., & Campillo, M., 2010. Origin of deep ocean microseisms by using teleseismic body waves, *Journal of Geophysical Research-Solid Earth*, **115**.

Larose, E., Margerin, L., Derode, A., van Tiggelen, B., Campillo, M., Shapiro, N., Paul, A., Stehly, L., & Tanter, M., 2006. Correlation of random wavefields: An interdisciplinary review, *Geophysics*, **71**(4), SI11–SI21.

Larose, E., Stehly, L., & Campillo, M., 2008. *Imaging the solid Earth with seismic noise*, vol. 118 of **Journal of Physics Conference Series**, pp. 12003–12003.

Leroy, C., Lani, S., Sabra, K. G., Hodgkiss, W. S., Kuperman, W. A., & Roux, P., 2012. Enhancing the emergence rate of coherent wavefronts from ocean ambient noise correlations using spatio-temporal filters, *Journal of the Acoustical Society of America*, **132**(2), 883–893.

Lin, F. C., Moschetti, M. P., & Ritzwoller, M. H., 2008. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophysical Journal International*, **173**(1), 281–298.

Liu, R. Y., Parelius, J. M., & Singh, K., 1999. Multivariate analysis by data depth: Descriptive statistics, graphics and inference, *Annals of Statistics*, **27**(3), 783–840.

Long, M. D., 2013. Constraints on subduction geodynamics from seismic anisotropy, *Reviews of Geophysics*, pp. n/a–n/a.

Longuet-Higgins, M. S., 1950. A theory of the origin of microseisms, *Philosophical Transactions of the Royal Society of London. Series A (Mathematical and Physical Sciences)*, (857), 1–3535.

Love, J. D. & Christiansen, A. C., 1985. Geologic map of Wyoming.

Marano, S., Reller, C., Loeliger, H. A., & Fah, D., 2012. Seismic waves estimation and wavefield decomposition: application to ambient vibrations, *Geophysical Journal International*, **191**(1), 175–188.

Maupin, V., 2004. Comment on 'The azimuthal dependence of surface wave polarization in a slightly anisotropic medium' by T. Tanimoto, *Geophysical Journal International*, **159**(1), 365–368.

Maupin, V. & Park, J., 2007. *Theory and Observations Wave Propagation in Anisotropic Media*, vol. 1, Elsevier, Amsterdam ; Boston.

Maxwell, S., Waltman, C., Warpinski, N., Mayerhofer, M., & Boroumand, N., 2006. Imaging seismic deformation induced by hydraulic fracture complexity.

McNamara, D. E. & Buland, R. P., 2004. Ambient noise levels in the continental United States, *Bulletin of the Seismological Society of America*, **94**(4), 1517–1527.

Menon, R., Gerstoft, P., & Hodgkiss, W. S., 2012. Cross-correlations of diffuse noise in an ocean environment using eigenvalue based statistical inference, *Journal of the Acoustical Society of America*, **132**(5), 3213–3224.

Miller, A. J., 1990. *Subset selection in regression*, Monographs on statistics and applied probability, Chapman and Hall, London England ; New York.

Montagner, J. P. & Nataf, H. C., 1986. A Simple Method for Inverting the Azimuthal Anisotropy of Surface-Waves, *Journal of Geophysical Research-Solid Earth and Planets*, **91**(B1), 511–520.

Montagner, J. P. & Tanimoto, T., 1990. Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *Journal of Geophysical Research-Solid Earth and Planets*, **95**(B4), 4797–4819.

Mooney, W. D., 2007. *Crust and Lithospheric Structure Global Crustal Structure*, vol. 1, chap. 11, pp. 361–417, Elsevier.

Mordret, A., Shapiro, N., Singh, S., Roux, P., & Barkved, O., 2013. Helmholtz tomography of ambient noise surface wave data to estimate Scholte wave phase velocity at Valhall Life of the Field, *Geophysics*, **78**(2).

Moschetti, M. P., Ritzwoller, M. H., & Shapiro, N. M., 2007. Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps, *Geochemistry Geophysics Geosystems*, **8**.

Moschetti, M. P., Ritzwoller, M. H., Lin, F., & Yang, Y., 2010. Seismic evidence for widespread western-US deep-crustal deformation caused by extension, *Nature*, **464**(7290), 885–U94.

Nadakuditi, R. R. & Silverstein, J. W., 2010. Fundamental Limit of Sample Generalized Eigenvalue Based Detection of Signals in Noise Using Relatively Few Signal-Bearing and Noise-Only Samples, *Ieee Journal of Selected Topics in Signal Processing*, **4**(3), 468–480.

Nadeau, R. M. & Dolenc, D., 2005. Nonvolcanic tremors deep beneath the San Andreas Fault, *Science*, **307**(5708), 389–389.

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

Nakata, N., Snieder, R., Tsuji, T., Larner, K., & Matsuoka, T., 2011. Shear wave imaging from traffic noise using seismic interferometry by cross-coherence, *Geophysics*, **76**(6), SA97–SA106.

Nasseri Moghaddam, A., Cascante, G., & Hutchinson, J., 2005. A new quantitative procedure to determine the location and embedment depth of a void using surface waves, *Journal of Environmental and Engineering Geophysics*, **10**(1), 51–64.

Okada, H., 2003. *The microtremor survey method*, Geophysical monograph series, Society of Exploration Geophysicists, Tulsa, OK.

Park, J., Vernon, F. L., & Lindberg, C. R., 1987. Frequency-dependent polarization analysis of high-frequency seismograms, *Journal of Geophysical Research-Solid Earth and Planets*, **92**(B12), 12664–12674.

Parolai, S., Bormann, P., & Milkereit, C., 2002. New relationships between V-s, thickness of sediments, and resonance frequency calculated by the H/V ratio of seismic noise for the Cologne area (Germany), *Bulletin of the Seismological Society of America*, **92**(6), 2521–2527.

Parolai, S., Picozzi, M., Richwalski, S. M., & Milkereit, C., 2005. Joint inversion of phase velocity dispersion and H/V ratio curves from seismic noise recordings using a genetic algorithm, considering higher modes, *Geophysical Research Letters*, **32**(1).

Pawlak, A., Eaton, D. W., Darbyshire, F., Lebedev, S., & Bastow, I. D., 2012. Crustal anisotropy beneath Hudson Bay from ambient noise tomography: Evidence for post-orogenic lower-crustal flow?, *Journal of Geophysical Research-Solid Earth*, **117**.

Perrodon, A. & Zabek, J., 1990. Paris Basin, in Interior Cratonic Basins, *Am. Assoc. Petr. Geol. Memoir*, **51**, 653–679.

Peterson, J., 1993. Observations and modeling of seismic background noise.

Poggi, V. & Fäh, D., 2010. Estimating Rayleigh wave particle motion from three-component array analysis of ambient vibrations, *Geophysical Journal International*, **180**(1), 251–267.

Polet, J. & Kanamori, H., 2002. Anisotropy beneath California: shear wave splitting measurements using a dense broadband array, *Geophysical Journal International*, **149**(2), 313–327.

Poli, P., Campillo, M., Pedersen, H., & Grp, L. W., 2012a. Body-Wave Imaging of Earth's Mantle Discontinuities from Ambient Seismic Noise, *Science*, **338**(6110), 1063–1065.

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

Press, W. H., 2007. *Numerical recipes : the art of scientific computing*, Cambridge University Press, Cambridge, UK ; New York, 3rd edn.

Quintal, B., Schmalholz, S. M., & Podladchikov, Y. Y., 2011. Impact of fluid saturation on the reflection coefficient of a poroelastic layer, *Geophysics*, **76**(2), N1–N12.

Riahi, N., Bokelmann, G., Sala, P., & Saenger, E. H., 2013a. Time-lapse analysis of ambient surface wave anisotropy: A three-component array study above an underground gas storage, *Journal of Geophysical Research: Solid Earth*, pp. n/a–n/a.

Riahi, N., Goertz, A., Birkelo, B., & Saenger, E. H., 2013b. A statistical strategy for ambient seismic wavefield analysis: investigating correlations to a hydrocarbon reservoir, *Geophysical Journal International*, **192**(1), 148–162.

Robinson, J. W. & Shanley, K. W., 2004. *Jonah field : case study of a tight-gas fluvial reservoir*, AAPG studies in geology, American Association of Petroleum Geologists, Tulsa, Oklahoma.

Rodgers, P. W., 1992. Frequency limits for seismometers as determined from signal-to-noise ratios: 1. the electromagnetic seismometer, *Bulletin of the Seismological Society of America*, **82**(2), 1071–1098.

Rost, S. & Thomas, C., 2002. Array seismology: Methods and applications, *Reviews of Geophysics*, **40**(3).

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

Ruigrok, E., Campman, X., & Wapenaar, K., 2011. Extraction of P-wave reflections from microseisms, *Comptes Rendus Geoscience*, **343**(8-9), 512–525.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C., 2005. Surface wave tomography from microseisms in Southern California, *Geophysical Research Letters*, **32**(14).

Saenger, E. H., Schmalholz, S. M., Lambert, M. A., Nguyen, T. T., Torres, A., Metzger, S., Habiger, R. M., Muller, T., Rentsch, S., & Mendez Hernandez, E., 2009. A passive seismic survey over a gas field: Analysis of low-frequency anomalies, *Geophysics*, **74**(2), O29–O40.

Sala, P., Frehner, M., Tisato, N., & Pfiffner, O. A., 2013. Building a 3D near-surface geological and petrophysical model based on borehole data A case-study from Chmery, Paris Basin, France, *submitted to AAPG Bulletin*.

Samson, J. C., 1983. Pure states, polarized waves, and principal components in the spectra of multiple, geophysical time-series, *Geophysical Journal of the Royal Astronomical Society*, **72**(3), 647–664.

Schaefer, J. F., Boschi, L., Becker, T. W., & Kissling, E., 2011. Radial anisotropy in the European mantle: Tomographic studies explored in terms of mantle flow, *Geophysical Research Letters*, **38**.

Scherbaum, F., Hinzen, K. G., Ohrnberger, M., & Herrmann, R. B., 2003. Determination of shallow shear wave velocity profiles in the Cologne, Germany area using ambient vibrations, *Geophysical Journal International*, **152**(3), 597–612.

Schmidt, R. O., 1986. Multiple Emitter Location and Signal Parameter Estimation, *Ieee Transactions on Antennas and Propagation*, **34**(3), 276–280.

Shapiro, N. M., Ritzwoller, M. H., Molnar, P., & Levin, V., 2004. Thinning and flow of Tibetan crust constrained by seismic anisotropy, *Science*, **305**(5681), 233–236.

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, **307**(5715), 1615–1618.

Sheriff, R. E. & Geldart, L. P., 1995. *Exploration seismology*, Cambridge University Press, Cambridge ; New York, 2nd edn.

Sleeman, R., van Wettum, A., & Trampert, J., 2006. Three-channel correlation analysis: A new technique to measure instrumental noise of digitizers and seismic sensors, *Bulletin of the Seismological Society of America*, **96**(1), 258–271.

Smith, M. L. & Dahlen, F. A., 1973. The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, *Journal of Geophysical Research*, **78**(17), 3321–3333.

Snieder, R., Miyazawa, M., Slob, E., Vasconcelos, I., & Wapenaar, K., 2010. *A Comparison of Strategies for Seismic Interferometry*, Arrays and Array Methods in Global Seismology, Springer, New York.

Spearman, C., 1904. The proof and measurement of association between two things, *American Journal of Psychology*, **15**, 72–101.

Stehly, L., Fry, B., Campillo, M., Shapiro, N. M., Guilbert, J., Boschi, L., & Giardini, D., 2009. Tomography of the Alpine region from observations of seismic ambient noise, *Geophysical Journal International*, **178**(1), 338–350.

Steiner, B., Saenger, E. H., & Schmalholz, S. M., 2008. Time reverse modeling of low-frequency microtremors: Application to hydrocarbon reservoir localization, *Geophysical Research Letters*, **35**(3).

Takagi, R. & Okada, T., 2012. Temporal change in shear velocity and polarization anisotropy related to the 2011 M9.0 Tohoku-Oki earthquake examined using KiK-net vertical array data, *Geophysical Research Letters*, **39**.

Tanimoto, T., 2004. The azimuthal dependence of surface wave polarization in a slightly anisotropic medium, *Geophysical Journal International*, **156**(1), 73–78.

Tanimoto, T. & Prindle, K., 2007. Surface wave analysis with beamforming, *Earth Planets and Space*, **59**(5), 453–458.

Teatini, P., Castelletto, N., Ferronato, M., Gambolati, G., Janna, C., Cairo, E., Marzorati, D., Colombo, D., Ferretti, A., Bagliani, A., & Bottazzi, F., 2011. Geomechanical response to seasonal gas storage in depleted reservoirs: A case study in the Po River basin, Italy, *Journal of Geophysical Research-Earth Surface*, **116**.

Toksoz, M. N., 1964. Microseisms and an attempted application to exploration, *Geophysics*, **29**(2), 154–.

Toksoz, M. N. & Lacoss, R. T., 1968. MICROSEISMS - MODE STRUCTURE AND SOURCES, *Science*, **159**(3817), 872–.

Tonegawa, T., Fukao, Y., Nishida, K., Sugioka, H., & Ito, A., 2013. A temporal change of shear wave anisotropy within the marine sedimentary layer associated with the 2011 Tohoku-Oki earthquake, *Journal of Geophysical Research-Solid Earth*, **118**(2), 607–615.

Traer, J. & Gerstoft, P., 2011. Coherent averaging of the passive fathometer response using short correlation time, *Journal of the Acoustical Society of America*, **130**(6), 3633–3641.

Trampert, J. & Woodhouse, J. H., 2003. Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s, *Geophysical Journal International*, **154**(1), 154–165.

Tsai, V. C., 2009. On establishing the accuracy of noise tomography travel-time measurements in a realistic medium, *Geophysical Journal International*, **178**(3), 1555–1564.

Tsai, V. C. & Moschetti, M. P., 2010. An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, *Geophysical Journal International*, **182**(1), 454–460.

van Mastrigt, P. & Al-Dulaijan, A., 2008. Seismic Spectroscopy Using Amplified 3C Geophones, in *EAGE 70th Annual International Meeting, Expanded Abstracts*, p. B047.

Vidale, J. E., 1986. Complex polarization analysis of particle motion, *Bulletin of the Seismological Society of America*, **76**(5), 1393–1405.

Wagner, G. S., 1996. Resolving diversely polarized, superimposed signals in three-component seismic array data, *Geophysical Research Letters*, **23**(14), 1837–1840.

Wapenaar, K., 2004. Retrieving the Elastodynamic Green's Function of an Arbitrary Inhomogeneous Medium by Cross Correlation, *Physical Review Letters*, **93**(25), 254301.

Warpinski, N., 2010. The Physics of Surface Microseismic Monitoring, Tech. rep.

Wathelet, M., Jongmans, D., & Ohrnberger, M., 2005. Direct inversion of spatial autocorrelation curves with the neighborhood algorithm, *Bulletin of the Seismological Society of America*, **95**(5), 1787–1800.

Weaver, R. L. & Lobkis, O. I., 2006. Diffuse fields in ultrasonics and seismology, *Geophysics*, **71**(4), SI5–SI9.

Webb, S. C. & Cox, C. S., 1986. Observations and modeling of sea-floor microseisms, *Journal of Geophysical Research-Solid Earth and Planets*, **91**(B7), 7343–7358.

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

Withers, M. M., Aster, R. C., Young, C. J., & Chael, E. P., 1996. High-frequency analysis of seismic background noise as a function of wind speed and shallow depth, *Bulletin of the Seismological Society of America*, **86**(5), 1507–1515.

Witten, B. & Artman, B., 2011. Signal-to-noise estimates of time-reverse images, *Geophysics*, **76**(2), MA1–MA10.

Wüstefeld, A., Bokelmann, G., Barruol, G., & Montagner, J. P., 2009. Identifying global seismic anisotropy patterns by correlating shear-wave splitting and surface-wave data, *Physics of the Earth and Planetary Interiors*, **176**(3-4), 198–212.

Xu, Z., Juhlin, C., Gudmundsson, O., Zhang, F. J., Yang, C., Kashubin, A., & Luth, S., 2012. Reconstruction of subsurface structure from ambient seismic noise: an example from Ketzin, Germany, *Geophysical Journal International*, **189**(2), 1085–1102.

Yaglom, A. M. & Silverman, R. A., 1962. *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall.

Yokoi, T. & Margaryan, S., 2008. Consistency of the spatial autocorrelation method with seismic interferometry and its consequence, *Geophysical Prospecting*, **56**(3), 435–451.

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

Zhang, J. A., Gerstoft, P., & Bromirski, P. D., 2010. Pelagic and coastal sources of P-wave microseisms: Generation under tropical cyclones, *Geophysical Research Letters*, **37**.

Zhang, X., Paulssen, H., Lebedev, S., & Meier, T., 2007. Surface wave tomography of the Gulf of California, *Geophysical Research Letters*, **34**(15).