

IDEA League

MASTER OF SCIENCE IN APPLIED GEOPHYSICS

RESEARCH THESIS

Application of seismic interferometry by Multidimensional Deconvolution to USArray data

Raied Al Sadan

August 7, 2015

Application of seismic interferometry by Multidimensional Deconvolution to USArray data

MASTER OF SCIENCE THESIS

for the degree of Master of Science in Applied Geophysics at
Delft University of Technology

ETH Zürich

RWTH Aachen University

by

Raied Al Sadan

August 7, 2015

Department of Geoscience & Engineering	·	Delft University of Technology
Department of Earth Sciences	·	ETH Zürich
Faculty of Georesources and Material Engineering	·	RWTH Aachen University



Delft University of Technology

Copyright © 2013 by IDEA League Joint Master's in Applied Geophysics:

Delft University of Technology, ETH Zürich, RWTH Aachen University

All rights reserved.

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying or by any information storage and retrieval system, without permission from this publisher.

Printed in The Netherlands

IDEA LEAGUE
JOINT MASTER'S IN APPLIED GEOPHYSICS

Delft University of Technology, The Netherlands
ETH Zürich, Switzerland
RWTH Aachen, Germany

Dated: *August 7, 2015*

Supervisor(s):

Dr. Cornelis Weemstra

Dr. Elmer Ruigrok

Committee Members:

Dr. Cornelis Weemstra

Dr. Elmer Ruigrok

Prof. Dr. Jan van der Kruk

Abstract

Seismic interferometry by crosscorrelation has found wide application over the last decade. With which, the Green's function is retrieved between two receiver positions. Because the bulk of the ambient seismic field consists of surface wave energy, most applications in seismology utilize the surface wave part of the Green's function. The correct reconstruction of interferometric (surface wave) responses, however, relies on a number of assumptions to be fulfilled. Violation of these assumptions reduces the accuracy of the retrieved responses. Seismic interferometry by multidimensional deconvolution (MDD) has been shown to improve the retrieved response on synthetic data. A point-spread function (PSF) computed from the same ambient noise is deconvolved from the response acquired by crosscorrelation and the result is a better focused response. In this thesis, it is demonstrated with numerical modeling that seismic interferometry by MDD yields better results than seismic interferometry by crosscorrelation. Moreover, a different formulation of the underlying theory is considered. The new formulation obviates the need to separate in-and outgoing wavefields, which is a drawback of the original formulation. The MDD method resulting from the new formulation is also numerically tested and subsequently applied to the data recorded by a passive deployment of broadband sensors (Transportable USArray) along the east of the continental USA. Limited response improvement is achieved mainly due to two reasons: receivers aparature (large nominal receiver spacing/limited width and height of receivers array) and illumination gaps.

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my supportive principle supervisor Dr. Cornelis Weemstra for the continuous support of my Masters thesis and related research, for his patience, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I would also like to extend my gratitude to my very helpful co-supervisor, Dr. Elmer Ruigrok. His deep insight, critical comments and coding skills helped me a great deal to tackle issues arisen. Last but not least, a big thank you for my very supportive parents and family.

Delft University of Technology
August 7, 2015

Raied Al Sadan

Table of Contents

Abstract	v
Acknowledgments	vii
Acronyms	xiii
1 Introduction	1
2 Theory	5
2-1 Green's function representation for seismic interferometry	5
2-2 Interferometry by Crosscorrelation	7
2-3 Interferometry by MDD	9
2-4 Reflecting boundary conditions	11
3 Numerical Modeling	13
3-1 1 D Model	13
3-1-1 Sources on one side	13
3-1-2 Sources on both sides	16
3-2 2 D Model	17
3-2-1 Effect of illumination on retrieved responses	20
3-2-2 Reflecting boundary condition	24
4 Application to USArray Data	33
4-1 Available Data	33
4-2 Preprocessing	35
4-3 Application of Interferometry	39
5 Conclusion	45
Bibliography	47

List of Figures

2-1	Green's function representation for seismic interferometry	6
2-2	Absorbing and reflecting boundary conditions	9
3-1	Source properties	14
3-2	Source from 1 direction crosscorrelation	15
3-3	Noise source from one direction	15
3-4	Simple 1D model survey setup (two directions)	16
3-5	Crosscorrelation of superpositions (receivers 1 and 11)	17
3-6	Crosscorrelation of individual responses and summing over sources (receivers 1 and 11)	17
3-7	Closer look into the cross-terms	18
3-8	Cross-terms do not contribute in uncorrelated noise sources	19
3-9	Effect of illumination in 1D setup	19
3-10	Perfect illumination	21
3-11	Irregular illumination	22
3-12	Regularly spaced sources with different power spectra	23
3-13	2D one sided illumination	25
3-14	2D one sided illumination (Noise)	26
3-15	Reflecting boundary conditions	27
3-16	Reflecting boundary conditions using actual USArray receivers locations	29
3-17	Reflecting boundary conditions using actual USArray receivers locations and real data average spectrum	31
4-1	USArray receivers locations	34
4-2	Available USArray Data	36
4-3	Frequency content of the USArray data set	37
4-4	Spectral homogenization of the USArray data set	38

4-5	Workflow flowchart	39
4-6	East west line crosscorrelations	40
4-7	North south line crosscorrelations	41
4-8	Comparison of crosscorrelations	42
4-9	Application of RBCMDD on the USArray data set	43
4-10	Beamforming analysis on the USArray data set	44

Acronyms

DUT Delft University of Technology

ETH Swiss Federal Institute of Technology

RWTH Aachen University

SI Seismic interferometry

MDD Multidimensional deconvolution

PSF Point spread function

CF Crosscorrelation function

ABC MDD Absorbing boundary condition multidimensional deconvolution

RBC MDD Reflecting boundary condition multidimensional deconvolution

RAMN Running-absolute-mean normalization

Chapter 1

Introduction

The term *Seismic Interferometry* (also referred to as Green's function retrieval) refers to the synthesis of seismic responses from virtual sources. This is typically implemented by crosscorrelating recordings of two receivers and the result is interpreted as the response recorded at one receiver due to a virtual source located at the other receiver. This holds for direct and reflected waves, given sufficient illumination by actual sources. The main application of seismic interferometry is the retrieval of seismic surface wave responses from ambient noise and the subsequent tomographic determination of the surface wave velocity distribution of the subsurface ([Shapiro et al., 2005](#)). Another application exploits body wave signal to create a virtual source. In this case the reflection response can be obtained by simple crosscorrelation ([Draganov et al., 2007](#)). This is done without necessarily knowing the position of the reflectors or the subsurface parameters.

The concept from which seismic interferometry stemmed, was first introduced by [Claerbout \(1968\)](#). He showed that for an acoustic layered medium, one side of the autocorrelation of the seismogram due to an impulsive source at depth is the seismogram due to an impulsive source, and a receiver, on the surface. This mathematical derivation was strictly one dimensional. The author later conjectured for the 3D case that "by crosscorrelating noise traces recorded at two locations on the surface, we can construct the wave field that would be recorded at one of the locations if there was a source at the other" ([Rickett and Claerbout, 1999](#)). Since then, several attempts have been made to make this conjecture work on real data ([Scherbaum, 1987a,b](#); [Cole, 1995](#); [Rickett and Claerbout, 1999](#)). The theory developed thereafter where it has been shown by various authors that the Greens function of a random medium or an irregular finite body can be obtained by cross correlating the recordings of a diffuse wave field at two receiver positions ([Weaver and Lobkis, 2001](#); [van Tiggelen, 2003](#); [Snieder, 2004](#)). These derivations holds only for diffusive fields. In 2002, Claerbouts conjecture was proven by [Wapenaar \(2003\)](#). He showed that the 3D generalization can be obtained along the same lines using a power reciprocity theorem. A year later, [Wapenaar \(2004\)](#) derived a relation, based on Rayleigh's reciprocity theorem, between the elastodynamic Greens function and the crosscorrelation of observed wave fields that holds at the free surface of random as well as

deterministic media. An alternative derivation based on physical arguments using the time-reversal invariance of the acoustic wave equation has been developed during the same time [Derode et al. \(2003\)](#).

Various authors have then shown that, in many cases, it is advantageous to replace crosscorrelation with deconvolution to relax some of the assumptions of interferometry by crosscorrelation, most importantly: the imprint of the source spectrum is removed and the deconvolution accounts for dissipation of energy. [Snieder et al. \(2006\)](#) employs a one dimensional deconvolution process to deconvolve passive wavefields observed at different depth levels and show that this leads to an estimate of the impulse response. [Schuster and Zhou \(2006\)](#); [Wapenaar et al. \(2008a\)](#) discuss a deconvolution both in space and time (MDD) in the context of controlled source seismic interferometry and later, [Wapenaar et al. \(2008b\)](#) proposed a method for seismic interferometry of passive data by MDD.

One can distinguish between controlled source and passive interferometry. Controlled source interferometry (as in exploration seismic), on top of crosscorrelation, also involves a summation of these correlations over all the sources. In passive seismic interferometry (as in noise and micro-earthquakes), a single crosscorrelation is sufficient as the recording is already the superposition of the responses due to individual sources. Given that the sources are mutually uncorrelated the same result is obtained as with controlled source seismic interferometry.

Interferometry by crosscorrelation assumes, among other assumptions, that the medium is lossless and the receivers are uniformly illuminated from all directions. This is rarely the case in reality. Also, the source power spectrum leaves an imprint on the retrieved Green's function. Interferometry by deconvolution relaxes those assumptions, removes the imprint of the source power spectrum and accounts for dissipative media. Interferometry by MultiDimensional Deconvolution (MDD), as a natural extension of 1D deconvolution, relaxes those assumptions even further and corrects the reconstructed responses for artifacts due to irregularities in the illumination pattern ([Wapenaar et al., 2011a](#)).

Conventional interferometry by MDD assumes an Absorbing Boundary Condition (ABC). This assumption heavily relies on the ability to separate inward and outward propagating wavefields which in turn, relies on assumptions that are not often fully satisfied. A revised formulation, assumes a Reflecting Boundary Condition (RBC), eliminates the need to separate the wavefields. However, it also introduces 'non-physical' reflections in the reconstructed response from the receivers boundary which can be muted if receivers aparature allows.

The main objective of this work is to apply seismic interferometry by RBCMDD to time-series of continuous seismic noise recorded by the Transportable USArray. At low frequencies (0.05-0.33 Hz), the ambient seismic wavefield is dominated by microseisms. These microseisms are the result of coupling between the oceans and the solid earth ([Longuet-Higgins, 1950](#); [Ardhuin et al., 2011](#)). The most energetic seismic waves that make up the microseismic field are Rayleigh waves. I therefore restrict myself to scalar fields because, at the free surface of

a layered medium, a single-mode dispersive surface wave behaves in the far field very similar to a scalar body wave in a 2-D homogeneous lossy medium ([Tsai, 2011](#)).

This work solely covers direct wave seismic interferometry in the context of surface waves recorded by a single (vertical) component receiver. A theoretical background is first introduced followed by numerical models that demonstrates seismic interferometry in 1 and 2D as well as highlights the advantages of each method. The theory is then applied to a real data set (subset of the Transportable USArray data) and the results are discussed.

Chapter 2

Theory

2-1 Green's function representation for seismic interferometry

The theory described in this section applies to acoustic media . The reason is that the ambient seismic noise data I will process, is dominated by fundamental-mode Rayleigh waves traveling along the Earth's surface. The behavior of such waves is very similar to waves propagating through a two-dimensional acoustic medium. The Green's function representations that form the basis for seismic interferometry can be derived from Rayleigh's reciprocity theorem (which is, in turn, can be obtained by substituting the equation of motion and a stress-strain relation for two states into the interaction quantity of [Hoop \(1988\)](#)). Integrating Rayleigh's reciprocity over an arbitrary spatial domain and applying the theorem of Gauss results in the convolution-type acoustic reciprocity theorem. Assuming a lossless medium, the time reversal invariance theorem can be applied on the convolution type reciprocity theorem which gives the correlation-type acoustic reciprocity theorem. Substituting Green's functions for the wavefield into both reciprocity theorems gives a correlation-type and convolution-type representation for the Green's functions ([Wapenaar and Fokkema, 2006](#)).

The configuration associated with the correlation-type Green's function representation is shown in [Figure 2-1a](#). A surface $\mathcal{S}$ bounds a lossless inhomogeneous volume $\mathcal{V}$ with possibly different medium parameters outside $\mathcal{S}$ and/or possibly different boundary conditions at $\mathcal{S}$. The Green's functions are denoted by $\hat{G}$ where the hat denotes frequency dependence (instead of time-domain dependence). The possibility of different medium or boundary conditions is reflected by the bar over $\hat{G}$ (still, inside $\mathcal{V}$ medium parameters are the same for $\hat{G}$ and $\bar{\hat{G}}$). Observation points $\mathbf{x}_A$ and $\mathbf{x}_B$ are contained within $\mathcal{V}$ and are uniformly illuminated by all sources on $\mathcal{S}$. The rays represent full responses, including primary and multiple scattering due to inhomogeneities inside as well as outside $\mathcal{S}$. The correlation-type representation for the causal and acausal acoustic Greens function between $\mathbf{x}_A$ and $\mathbf{x}_B$ for this is given by ([van Manen and Robertsson \(2005\)](#); [Wapenaar et al. \(2005\)](#))

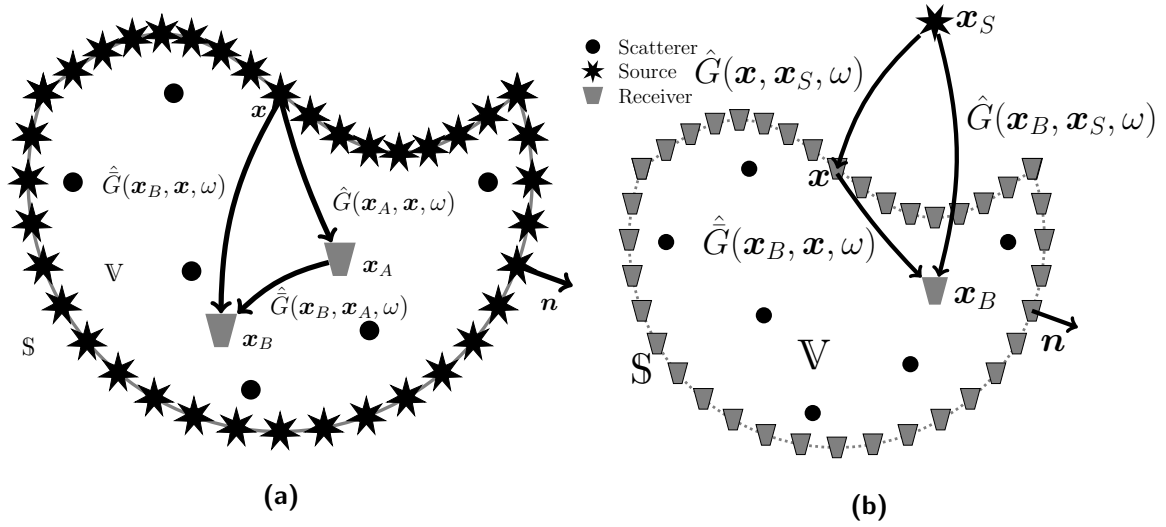


Figure 2-1: Green's function representation for seismic interferometry

(a) Configuration for the correlation-type Greens function representation (Eq. (2-1)). (b) Configuration for the convolution-type Greens function representation (Eq. (2-2)). Figures modified after [Weemstra et al. \(2015\)](#)

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_A, \omega) + \hat{G}^*(\mathbf{x}_B, \mathbf{x}_A, \omega) = - \oint_S \frac{1}{j\omega\rho(x)} (\partial_i \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega) - \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \partial_i \hat{G}^*(\mathbf{x}_A, \mathbf{x}, \omega)) n_i d\mathbf{x}, \quad (2-1)$$

where j is the unit imaginary number, $\rho(x)$ and $c(x)$ are the mass density and acoustic propagation velocity, respectively, on $\mathcal{S}$, ω is the angular frequency and $\mathbf{n}$ is an outward pointing normal vector to $\mathcal{S}$. The superscript asterisk denotes complex conjugation. If the parameters inside and outside $\mathcal{V}$ are the same, which is often the case, then the bar on G is omitted. The terms G and $\partial_i G n_i$ under the integral on the right hand side of the equation correspond to the responses of monopole and dipole sources at $\mathbf{x}$, respectively.

The convolution-type Green's function representation can be obtained by slightly modifying the theoretical configuration in Figure 2-1a. Here, a surface $\mathcal{S}$ bounds an inhomogenous volume $\mathcal{V}$ and $\mathbf{x}_A$ is taken outside $\mathcal{S}$ and renamed $\mathbf{x}_S$. For this configuration the convolution-type representation is given by ([Wapenaar et al. \(2011a\)](#))

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_S, \omega) = - \oint_S \frac{1}{j\omega\rho(x)} (\partial_i \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \hat{G}(\mathbf{x}, \mathbf{x}_S, \omega) - \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \partial_i \hat{G}(\mathbf{x}, \mathbf{x}_S, \omega)) n_i d\mathbf{x}. \quad (2-2)$$

Due to the absence of complex-conjugation signs, the products on the right-hand side correspond to crossconvolutions in the time domain. An important difference with the correlation-type representation is that this representation remains valid in media with losses (this comes from the fact that the time reversal invariance is not applied and hence, no lossless medium needs to be assumed).

2-2 Interferometry by Crosscorrelation

The correlation-type Green's function representation (2-1) forms the basis of seismic interferometry by crosscorrelation in open systems as it is an exact representation of the acoustic Greens function. However, three complications in its current form hinders its direct applicability in seismic interferometry. First, the right hand side of this equation contains a combination of two correlation products that has to be evaluated separately. Second, the monopole and dipole responses are assumed to be available for all sources at $\mathbf{x}$, and third, the sources are assumed to be impulsive sources which does not comply with reality. By approximating the normal derivatives of the Green's functions in the high frequency regime, the two crosscorrelation terms can be combined into a single term but a ghost term appears. This ghost term or "spurious arrivals" are caused by outward propagating waves (outward of $\mathbb{S}$) scattering by inhomogenities outside $\mathbb{S}$ and propagating back to observation receivers $\mathbf{x}_A$ and $\mathbf{x}_B$. By assuming an absorbing boundary condition, that is, the medium at and outside $\mathbb{S}$ is homogeneous with a constant velocity c and mass density ρ for both $\hat{G}$ and $\hat{\hat{G}}$ and hence the bar can be dropped, the ghost term cancels. To remove the dipole response, it is approximated (or written in terms of) the monopole response. This approximation is accurate when $\mathbb{S}$ is a sphere with a very large radius. In reality, this is not generally the case and this causes amplitude errors that can be significant as well as introduce spurious arrivals when the contributions from different stationary points along $\mathbb{S}$ cancel incompletely. This also means that the receivers whose observations are crosscorrelated are uniformly illuminated from all directions. With all the assumptions taken into account, (2-1) simplifies to (in time domain, Wapenaar et al. (2011a))

$$G(\mathbf{x}_B, \mathbf{x}_A, t) + G(\mathbf{x}_B, \mathbf{x}_A, -t) \approx \frac{2}{\rho c} \oint_{\mathbb{S}_{src}} G(\mathbf{x}_B, \mathbf{x}_A, t) * G(\mathbf{x}_B, \mathbf{x}_A, -t) d\mathbf{x}_{\mathbb{S}}. \quad (2-3)$$

The $*$ in the right part of the equation denotes temporal convolution, but the time reversal of the second Greens function turns the convolution into a correlation. This equation shows that the crosscorrelation of the two Green's functions observed by two receivers ($\mathbf{x}_A$ and $\mathbf{x}_B$), summed over sources, gives the Green's function between the two receivers plus a time reversed version.

In practice, what is observed by receivers is the Green's function convolved with the source spectrum $\hat{s}(\mathbf{x}, \omega)$ and not the Green's function. When the sources are transient: the observed wavefields $\hat{u}$ at $\mathbf{x}_A$, $\mathbf{x}_B$ and source power spectrum $\hat{s}$ are defined as follows

$$\hat{u}(\mathbf{x}_A, \mathbf{x}, \omega) = \hat{G}(\mathbf{x}_A, \mathbf{x}, \omega) \hat{s}(\mathbf{x}, \omega), \quad (2-4)$$

$$\hat{u}(\mathbf{x}_B, \mathbf{x}, \omega) = \hat{G}(\mathbf{x}_B, \mathbf{x}, \omega) \hat{s}(\mathbf{x}, \omega), \quad (2-5)$$

$$\hat{S}(\mathbf{x}, \omega) = \hat{s}^*(\mathbf{x}, \omega) \hat{s}(\mathbf{x}, \omega). \quad (2-6)$$

Using those definitions, the frequency domain version of equation (2-3) can be modified as follows (Wapenaar and Fokkema (2006))

$$2\mathbb{R}G(x_A, x_B, \omega)\hat{S}_0(\omega) \approx \frac{2}{\rho c} \oint_{\mathbb{S}} \hat{F}(\mathbf{x}, \omega) \hat{u}^*(\mathbf{x}_A, \mathbf{x}, \omega) \hat{u}(\mathbf{x}_B, \mathbf{x}, \omega) d\mathbf{x}_{\mathbb{S}}, \quad (2-7)$$

where $\hat{S}_0(\omega)$ is an arbitrarily chosen average power spectrum and $\hat{F}(\mathbf{x}, \omega)$ is a shaping filter defined as

$$\hat{F}(\mathbf{x}, \omega) = \frac{\hat{S}_0(\omega)}{\hat{S}(\mathbf{x}, \omega)}. \quad (2-8)$$

The shaping filter corrects for the differences in power spectra of different sources in case each source has a different wavelet. Equation (2-7) states that the sum of crosscorrelations of two receivers observations over all sources (and filtered in case different sources have different wavelets) gives the causal and acausal Green's function between those receivers convolved with the autocorrelation of the source wavelet or the average power spectrum in case different sources have different wavelets.

When the individual responses due to each source is unknown, like in the case of simultaneously acting ambient noise sources with a power spectrum $\hat{N}(\mathbf{x}, \omega)$, the receivers observations may be defined as

$$\hat{u}(\mathbf{x}_A, \omega) = \oint_{\mathbb{S}} \hat{G}(\mathbf{x}_A, \mathbf{x}, \omega) \hat{N}(\mathbf{x}, \omega) d\mathbf{x}_{\mathbb{S}}, \quad (2-9)$$

$$\hat{u}(\mathbf{x}_B, \omega) = \oint_{\mathbb{S}} \hat{G}(\mathbf{x}_B, \mathbf{x}', \omega) \hat{N}(\mathbf{x}', \omega) d\mathbf{x}_{\mathbb{S}}, \quad (2-10)$$

$$\langle \hat{N}^*(\mathbf{x}, \omega) \hat{N}(\mathbf{x}', \omega) \rangle = \delta(\mathbf{x} - \mathbf{x}') \hat{S}(\omega). \quad (2-11)$$

Equation (2-11) is accurate only when the two noise sources are mutually uncorrelated and their power spectrum $\hat{S}(\omega)$ is the same. The $\langle \cdot \rangle$ denotes ensemble average. In practice, this is replaced by averaging noise correlations over a long period of time. Using those definitions, the frequency domain version of equation (2-3) can be modified as follows ([Wapenaar and Fokkema \(2006\)](#))

$$2\mathbb{R}G(\mathbf{x}_A, \mathbf{x}_B, \omega)\hat{S}(\omega) \approx \frac{2}{\rho c} \langle \hat{u}(\mathbf{x}_A, \omega) \hat{u}(\mathbf{x}_B, \omega) \rangle. \quad (2-12)$$

Here, the individual responses due to each source does not need to be known. However, no correction can be made if different sources have different power spectra similar to the shaping filter $\hat{F}(\mathbf{x}, \omega)$ in (2-8). In other words, the power spectra of all noise sources has to be the same.

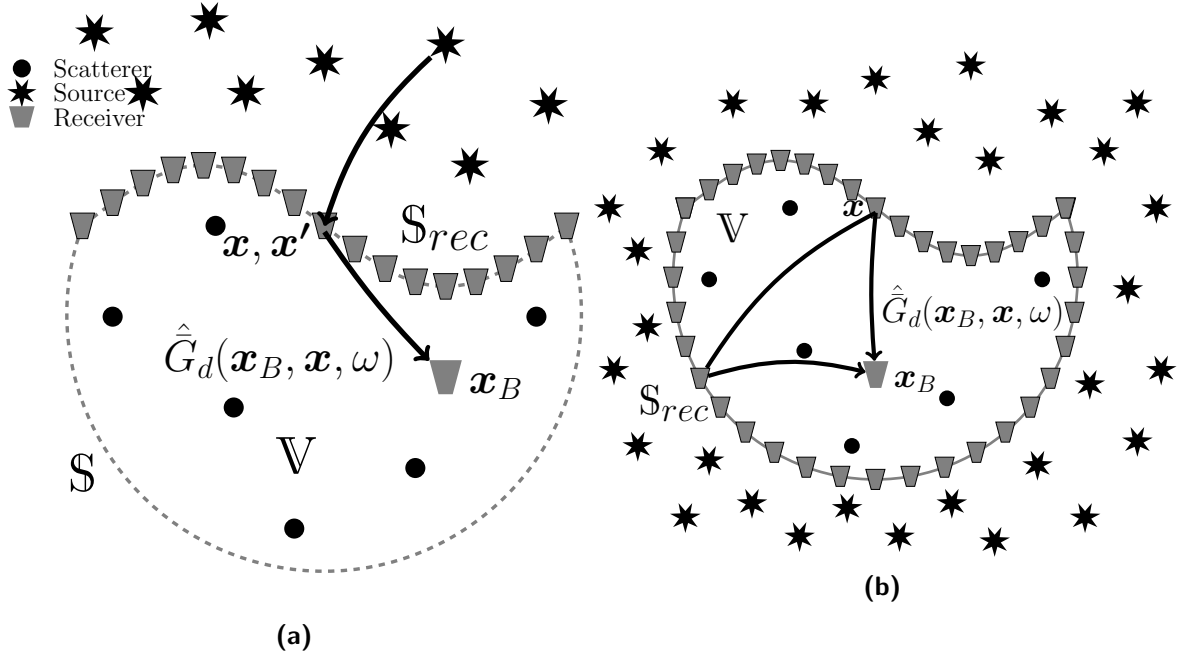


Figure 2-2: Absorbing and reflecting boundary conditions

(a) Configuration associated with the absorbing boundary condition and (b) Configuration associated with the reflecting boundary condition. Figures modified after [Weemstra et al. \(2015\)](#)

2-3 Interferometry by MDD

The convolution-type Green's function representation (equation (2-2)) is the basic expression for seismic interferometry by MDD. For the same reasons mentioned earlier in 2-2, equation (2-2) cannot be directly applied and needs simplification. The Green's function $\hat{G}(x_B, x, \omega)$ under the integral in the right hand side is the sought after and the one that needs to be resolved by MDD. The other two Green's functions are related to the observations and therefore defined in the actual medium inside as well as outside $\mathbb{S}$. Assuming the following: an absorbing boundary condition for $\hat{G}$ (homogeneous medium outside $\mathbb{S}$) which is assumed to be sufficiently smooth, ρ is constant on $\mathbb{S}$ and introducing the dipole Green's function $\hat{G}_d(x_B, x, \omega) = \frac{-2}{j\omega\rho}(n_i d_i \hat{G}(x_B, x, \omega))$, equation (2-2) simplifies to ([Wapenaar et al. \(2011b\)](#))

$$\hat{G}(x_B, x_S, \omega) = \oint_{\mathbb{S}_{rec}} \hat{G}_d(x_B, x, \omega) \hat{G}^{in}(x, x_S, \omega) dx, \quad (2-13)$$

where $\mathbb{S}_{rec}$ denotes that the integration surface contains the receivers of the Green's function $\hat{G}^{in}$. As absorbing boundary condition implies that only energy propagating into $\mathbb{V}$ is allowed, the integration of the open receiver boundary is taken as long as $\mathbb{S}_{rec}$ is sufficiently smooth and x_B and x_S lie on opposite sides of $\mathbb{S}_{rec}$ (one sided illumination, Figure 2-2a). Hence the closed contour integral in (2-13) is replaced with a normal integral. This is because of the radiation condition of the underlying convolution-type Green's function representation

(equation (2-2)) apply on the half sphere that closes the boundary. To replace the Green's functions with real observations, the observations due to transient sources are defined as

$$u^{in}(\mathbf{x}, \mathbf{x}_S^i, t) = G^{in}(\mathbf{x}, \mathbf{x}_S^i, t) s^i(\mathbf{x}, t), \quad (2-14)$$

$$u(\mathbf{x}_B, \mathbf{x}_S^i, t) = G(\mathbf{x}_B, \mathbf{x}_S^i, t) s^i(\mathbf{x}, t), \quad (2-15)$$

where the superscript i denotes a source index. Convolution both sides of equation (2-13) with $s^i(\mathbf{x}, \omega)$, taking the integration on the open receiver boundary and writing it in the time domain yields

$$u(\mathbf{x}_B, \mathbf{x}_S^i, t) = \int_{\mathbb{S}_{rec}} \bar{G}_d(\mathbf{x}_B, \mathbf{x}, t) u^{in}(\mathbf{x}, \mathbf{x}_S^i, t) d\mathbf{x}. \quad (2-16)$$

If the individual responses due to each source are not known, just like the case of simultaneously acting noise sources, equation (2-16) is summed over all sources to get an expression for the total observation as follows

$$\hat{u}(\mathbf{x}_B, \omega) = \int_{\mathbb{S}_{rec}} \hat{G}_d(\mathbf{x}_B, \mathbf{x}, \omega) \hat{u}^{in}(\mathbf{x}, \omega) d\mathbf{x}_{\mathbb{S}_{rec}}, \quad (2-17)$$

with the receivers observations $\hat{u}$ due to noise sources with spectrum $\hat{N}(\mathbf{x}, \omega)$ defined in a similar way as in equation (2-10). The sought after $\hat{G}_d(\mathbf{x}_B, \mathbf{x}, \omega)$ can be solved by first crosscorrelating both sides of (2-17) by the inward propagating wavefield observed at $\mathbb{S}_{rec}$ and taking the ensemble average. Also, introducing a coordinate $\mathbf{x}'$ on $\mathbb{S}_{rec}$ to distinguish between receivers whose observations are crosscorrelated (van Dalen et al. (2014))

$$\hat{C}(\mathbf{x}_B, \mathbf{x}', \omega) = \int_{\mathbb{S}_{rec}} \hat{G}_d(\mathbf{x}_B, \mathbf{x}, \omega) \hat{\Gamma}(\mathbf{x}, \mathbf{x}', \omega) d\mathbf{x}_{\mathbb{S}_{rec}}, \quad (2-18)$$

where $\hat{C}(\mathbf{x}_B, \mathbf{x}', \omega)$ is the Correlation Function (CF) and $\hat{\Gamma}(\mathbf{x}, \mathbf{x}', \omega)$ is the Point Spread Functions (PSF) defined as follows

$$\hat{C}(\mathbf{x}_B, \mathbf{x}', \omega) = \langle \hat{u}(\mathbf{x}_B, \omega) \hat{u}^{in*}(\mathbf{x}', \omega) \rangle, \quad (2-19)$$

$$\hat{\Gamma}(\mathbf{x}, \mathbf{x}', \omega) = \langle \hat{u}^{in}(\mathbf{x}, \omega) \hat{u}^{in*}(\mathbf{x}', \omega) \rangle. \quad (2-20)$$

Using the definition of uncorrelated sources (equation (2-11)), the ensemble averages in the CF and PSF expressions simplify to a sum over sources (in case of transient sources) or simply the superposition of responses due to all sources (total signal) in case of uncorrelated noise sources.

In words, the CF is the crosscorrelation of observations of the receivers $\mathbf{x}'$, that lies on $\mathcal{S}rec$ (also are the virtual sources) with the receivers inside $\mathbb{V}$, i.e, $\mathbf{x}_B$. The PSF is the crosscorrelation of individual receivers on $\mathcal{S}rec(\mathbf{x})$ with the other receivers on $\mathcal{S}rec(\mathbf{x}')$. This quantifies the smearing of the virtual source in space and time. The level of smearing depends on illumination, level of correlation of sources and the amount of losses in the medium. The more irregular the illumination is and the more the sources are correlated and the more losses in the medium the more smearing occurs. By deconvolving the CF by the PSF, this smearing is corrected for (but not fully). The virtual source then becomes better focused and, accordingly, the retrieved virtual source response more accurate. To achieve this, (2-18) is discretized along $\mathcal{S}rec$ and for each frequency, a system of equations is obtained that is finally inverted in a least squares sense to get $\hat{G}_d(\mathbf{x}_B, \mathbf{x}, \omega)$. The system of equations can be written in the matrix form

$$\hat{C} = \hat{G}_d \hat{\Gamma}, \quad (2-21)$$

with the rows and columns of these matrices related to different $\mathbf{x}_B$ and $\mathbf{x}'$ stations, respectively. Solving for the virtual source response in a least squares sense as follows

$$\hat{G}_d \cong \hat{C}(\hat{\Gamma} + \epsilon^2 I)^{-1}, \quad (2-22)$$

where ϵ^2 is a small stabilization number and I is the identity matrix. This inversion might not be possible when the noise sources are not mutually uncorrelated ($\langle \hat{N}^*(\mathbf{x}, \omega) \hat{N}(\mathbf{x}', \omega) \rangle = \hat{S}(\omega)$) as the resulting system of equations can be underdetermined (Wapenaar et al., 2012). This means that the more the sources are mutually uncorrelated (equation (2-11)), the more stable the inversion will be and the better the MDD results.

2-4 Reflecting boundary conditions

As seen in section 2-3, with the absorbing boundary condition, the ability of MDD to resolve a better (not smeared) virtual source response relies upon the ability to separate the inward and outward propagating wavefields (out and in $\mathbb{V}$). In modeling and controlled source experiments, this separation is achieved by only putting sources on one side of $\mathcal{S}rec$ and receivers ($\mathbf{x}_B$) on the other side (as in Figure 2-2a) so we are only left with the inward propagating wavefield (absorbing boundary condition assumes the medium outside $\mathbb{V}$ is homogeneous and therefore, there are no scattered back (into $\mathbb{V}$) waves). In ambient noise experiments, the sources are not controlled and are coming from all directions (Figure 2-2b). Separating wavefields can be done by choosing a preferable receiver configuration or by other methods that are often based on assumptions that are not fully satisfied. Therefore, considering the whole wavefield is beneficial, easier to work with and reduces the amount of assumptions by eliminating the assumptions in the wavefield separation process. The reflecting boundary conditions eliminate the necessity to separate the wavefield.

Assuming the pressure is zero on $\mathcal{S}rec$ inside $\mathbb{V}$, the convolution-type Green's function representation (equation (2-2)) simplifies to (Weemstra et al. (2015))

$$\hat{G}(\mathbf{x}_B, \mathbf{x}_S, \omega) = \oint_{\mathbb{S}rec} \hat{\hat{G}}_d(\mathbf{x}_B, \mathbf{x}, \omega) \hat{G}(\mathbf{x}, \mathbf{x}_S, \omega) d\mathbf{x}. \quad (2-23)$$

When compared to equation (2-3) the "in" superscript in $\hat{G}^{in}(\mathbf{x}, \mathbf{x}_S, \omega)$ disappears and the full wavefield is considered. Importantly, the sought after $\hat{\hat{G}}_d(\mathbf{x}_B, \mathbf{x}, \omega)$ contains reflections from the receiver boundary $\mathbb{S}rec$. The same procedure as for MDD with absorbing boundary conditions applies to obtain the dipole response. That is, convolving both sides with source spectrum, discretizing along $\mathbb{S}rec$ and inverting the resulting system of equations for each frequency in a least squares sense. The reflections contained within $\hat{\hat{G}}_d(\mathbf{x}_B, \mathbf{x}, \omega)$ can be simply muted, leaving the response of the direct arrival from the virtual source to the receiver. The ability of this method to work depends on the receivers configuration and propagation velocity/wavelength/frequency as well as the level of scattering within the medium. If the receiver ($\mathbf{x}_B$) is too close to the reflecting receivers boundary ($\mathbb{S}rec$) and the wavelength is large with respect to the time required for the reflection to come back, then the reflections would overlap (interfere, or get recorded in the same time as the direct arrival) hindering the ability to separate reflections from direct arrival and hence the ability to zero the reflections out. This also means that the waveform of the direct arrival is deformed by the reflections. Also, if scattering is strong within the medium, this will result in too many reflections getting recorded at the same time which also deforms the retrieved direct response.

Chapter 3

Numerical Modeling

Numerical models that demonstrates direct wave seismic interferometry are discussed in this chapter. Starting with a simple 1D model and building on it step by step to reach a representative model of the USArray data. A closely representative working model would show that seismic interferometry would work on the receivers configuration used to gather the USArray data, at least theoretically. Although the real (USArray) data is an elastodynamic data set, the modeling is mostly acoustic. All the modeling in this chapter is done in the frequency domain (frequency domain Green's functions), whereas (most of) the results are presented in time domain.

3-1 1 D Model

The source spectrum used in all of the 1D numerical examples is a Ricker wavelet with a central frequency of 10 Hz (Figure 3-1a) propagating in a lossless medium with a constant velocity of $c = 300m/s$. The data was created by convolving (multiplying in frequency domain) the wavelet with the 1D acoustic Green's function for a lossless medium

$$G = -\frac{i}{2k}e^{-ik|x-y|}, \quad (3-1)$$

where i is the imaginary unit number, $k = \frac{\omega}{c}$ is the wavenumber and $|x - y|$ is the distance between the source and the receiver.

3-1-1 Sources on one side

I use 11 receivers with a 30 meters spacing while one source is placed at a distance of 300 meters away from receiver 1 (Figure 3-2a). The source is located at $x = 0$ and emits a pulse at $t = 0$ that is rightward propagating. Crosscorrelating the recordings of any two receivers would result in the seismic response recorded at one receiver due to a source at the other receiver's

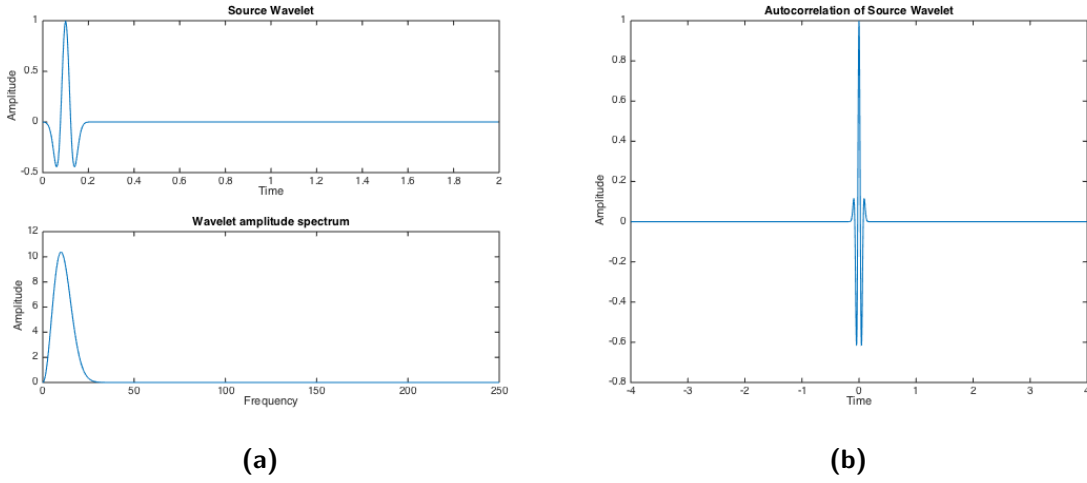


Figure 3-1: Source properties

(a) Source (Ricker) wavelet and amplitude spectrum, and (b) Autocorrelation of source wavelet

location. This response is the Green's function between those two receivers convolved with the autocorrelation of the source function (Figure 3-1b) as equation (2-7) states. Since the waves are coming only from one direction, only the causal part of the left hand side of equation (2-7) is retained. Also, only one source with one wavelet is used in this experiment so the shaping filter $\hat{F}(\mathbf{x}, \omega)$ needs not to be considered ($=1$). Figure 3-2b shows the crosscorrelation of the recordings of receivers 1 and 6 compared to the directly modeled response. The travel path associated with $G(\mathbf{x}_1, \mathbf{x}_S, t)$ and $G(\mathbf{x}_6, \mathbf{x}_S, t)$ have the path from $\mathbf{x}_S$ to $\mathbf{x}_1$ in common. This common travel path cancels in the process of crosscorrelation leaving the travel time along the ray path from $\mathbf{x}_1$ to $\mathbf{x}_6$, i.e. $t_6 - t_1 = (\mathbf{x}_6 - \mathbf{x}_1)/c = (450 - 300)/300 = 0.5s$ which means that both the location of the actual source and the propagation velocity needs not be known. Similarly, if the source activates at $t = t_s$ instead of $t = 0$, then the crosscorrelation result will be shifted by $t = t_s$. Here, the crosscorrelation completely overlaps the directly modeled response as all the assumptions of seismic interferometry by crosscorrelation is fulfilled in this model.

The principles of seismic interferometry holds true for any source function including noise. The next experiment uses the same exact setup as in Figure 3-2a but with a noise source $N(t)$ (Figure 3-3a). The noise is created by multiplying the spectrum of the Ricker wavelet used in the first experiment with random phases. Therefore, the amplitude spectrum as well as the autocorrelation of the noise are identical to those of the Ricker wavelet. A total of 40 seconds of noise was used and the crosscorrelation of observations of receivers 1 and 6 compared to the directly modeled response is shown in Figure 3-3b. Again, the common path is canceled in the crosscorrelation process and the response appears at $t_6 - t_1 = (\mathbf{x}_6 - \mathbf{x}_1)/c = (450 - 300)/300 = 0.5s$. The signal to noise ratio depends on the amount of noise used. Using an infinite noise series would achieve a 100% signal-to-noise and would result in identical image as in Figure 3-2b.

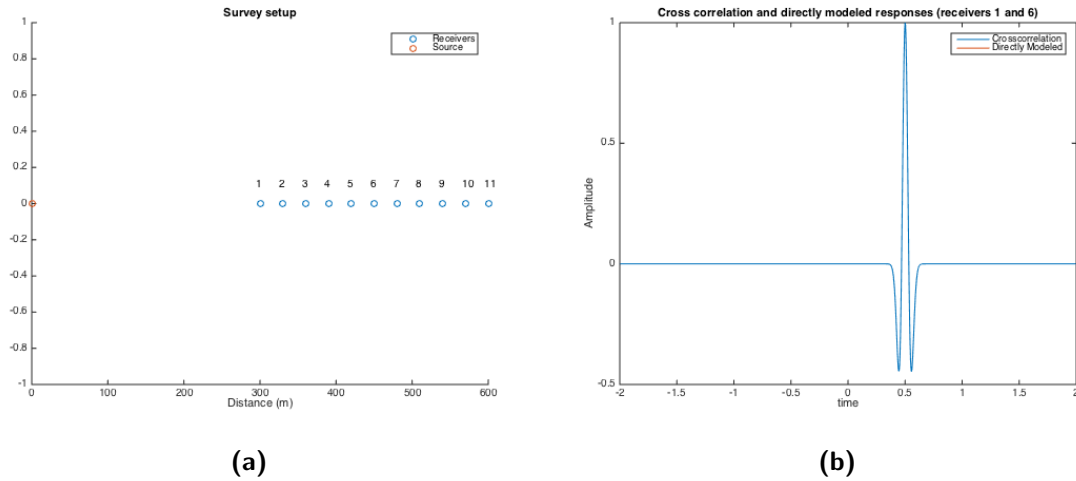


Figure 3-2: Source from 1 direction crosscorrelation
 (a) Simple 1D model survey setup (one direction), and (b) Crosscorrelation compared to directly modeled responses (receivers 1 and 6) using survey setup as in (a)

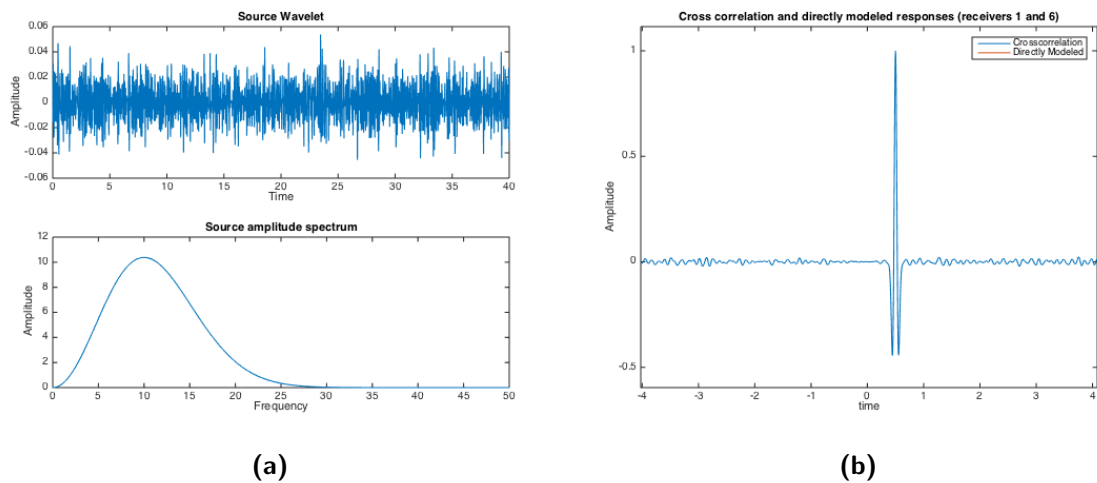


Figure 3-3: Noise source from one direction
 (a) Source noise and amplitude spectrum, and (b) Crosscorrelation compared to directly modeled responses (receivers 1 and 6) using survey setup as in Figure 3-2a

3-1-2 Sources on both sides

Now we consider two simultaneously acting transient sources that lie at both ends of the receivers array that have a spacing of 100 m (Figure 3-4). The west source is 400 m away from receiver 1 and the east source is 800 m away from receiver 11. The density of the lossless medium is $\rho = 1400 \text{ kg/m}^3$ with a water bulk modulus $K = 2.22 \times 10^9 \text{ Pascal}$ which makes the propagation velocity $c = \sqrt{\frac{K}{\rho}} = 1260 \text{ m/s}$. With more than one transient sources, the order of crosscorrelation and summation matters. The observation at $\mathbf{x}_1$ and $\mathbf{x}_{11}$, i.e., equations (2-4) and (2-5) becomes, in time domain: $u(\mathbf{x}_1, t) = \sum_{i=1}^2 G(\mathbf{x}_1, \mathbf{x}_s^i, t) * S(t)$ and $u(\mathbf{x}_{11}, t) = \sum_{i=1}^2 G(\mathbf{x}_{11}, \mathbf{x}_s^i, t) * S(t)$ respectively where $*$ donates temporal convolution. The crosscorrelation of these responses, shown in Figure 3-5, contain two cross-terms, on top of the causal and a causal parts, that has no physical meaning. Those cross-terms are arriving at the time required to travel the difference in distance between the west source and the first receiver used in crosscorrelation and east source and the second receiver (i.e., receivers 1 and 11 in this example). That is, $t_{\text{cross-terms}} = \pm \left| \frac{(\mathbf{x}_1 - \mathbf{x}_{s, \text{west}}) - (\mathbf{x}_{s, \text{east}} - \mathbf{x}_{11})}{c} \right| = \pm \frac{800 - 400}{1260} = 0.32 \text{ s}$ where $\pm$ here refers to plus and minus. Those terms disappear when the individual responses are crosscorrelated and then summed over sources as shown in Figure 3-6.

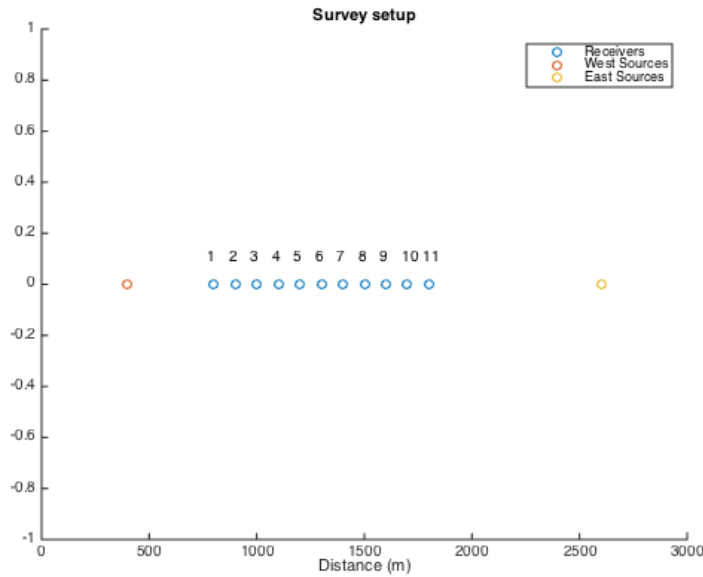


Figure 3-4: Simple 1D model survey setup (two directions)

To take a closer look into the nature of those cross terms, Figure 3-7 illustrates four different source configurations acting simultaneously on the same receivers array used in Figure 3-4. Here, the propagation velocity is reduced to $c = 300 \text{ m/s}$ to better separate the arrivals. Also, in any experiment with more than one source at a side, the sources are 80 m apart. Figure 3-7a shows the crosscorrelation of the superposition of responses for receivers 1 and 11 using two sources at equal distance from the receivers array. That is, the west source is 400 m west of receiver 1 and the east source is 400 m east of receiver 11. In this case, $t_{\text{cross-terms}} = \pm \left| \frac{400 - 400}{300} \right| = 0 \text{ s}$. That is, as shown in the Figure 3-7a, the two cross-terms superimpose at

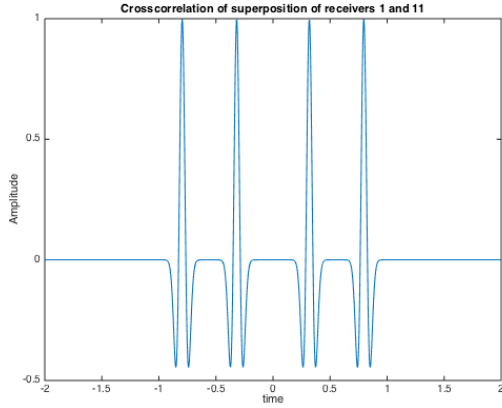


Figure 3-5: Crosscorrelation of superpositions (receivers 1 and 11)

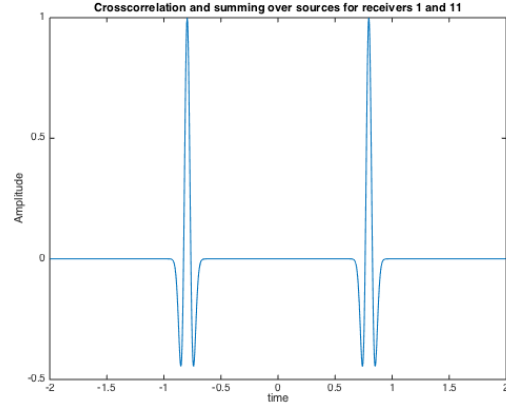


Figure 3-6: Crosscorrelation of individual responses and summing over sources (receivers 1 and 11)

$t = 0s$ resulting in double the amplitude of that of the causal and acausal parts. When sources are only acting from one side, only the causal response is retrieved. The cross terms, in this case, arrive at the time of the actual response $\pm$ the time required to travel the distance between the two sources. That is, $t_{cross-terms} = \frac{|\mathbf{x}_{11}-\mathbf{x}_1|}{c} \pm \frac{|\mathbf{x}_{s1}-\mathbf{x}_{s2}|}{c} = \frac{1000}{300} \pm \frac{80}{300} = 3.59s$ and $3.06s$ as shown in Figure 3-7b. Each additional source will add up two more cross-terms as demonstrated in Figure 3-7c where five west sources were used and a combination of effects is shown in Figure 3-7d where three sources were used on each side. To sum up, two sided illumination gives rise to “intermediary cross-terms” (between causal and acausal responses) and multiple sources on any side cause “associated cross-terms” to appear on the four main arrivals (causal and acausal responses and the two non physical “intermediary cross-terms”).

The situation is different for noise sources. Both the “intermediary” as well as the “associated” cross-terms do not contribute as long as the noise sources are mutually uncorrelated as in equation (2-11). This is demonstrated in Figure 3-8. Here, the same setup shown in Figure 3-4 is used with the reduced velocity ($c = 300m/s$). The noise sources are again the same Ricker wavelet as in Figure 3-1a multiplied with different random phases for each source.

3-2 2 D Model

The data in all the 2D numerical examples was synthesized by convolving a Ricker wavelet (multiplied by different random phases in case of noise sources) with a 2D Green’s function defined as follows

$$G = -\frac{i}{4}H_0^{(2)}(kr), \quad (3-2)$$

where i is the imaginary unit number, $H_0^{(2)}$ is the zeroth-order Hankel function of the second kind, $k = \frac{\omega}{c}$ is the wavenumber and r is the source receiver distance (propagation distance).

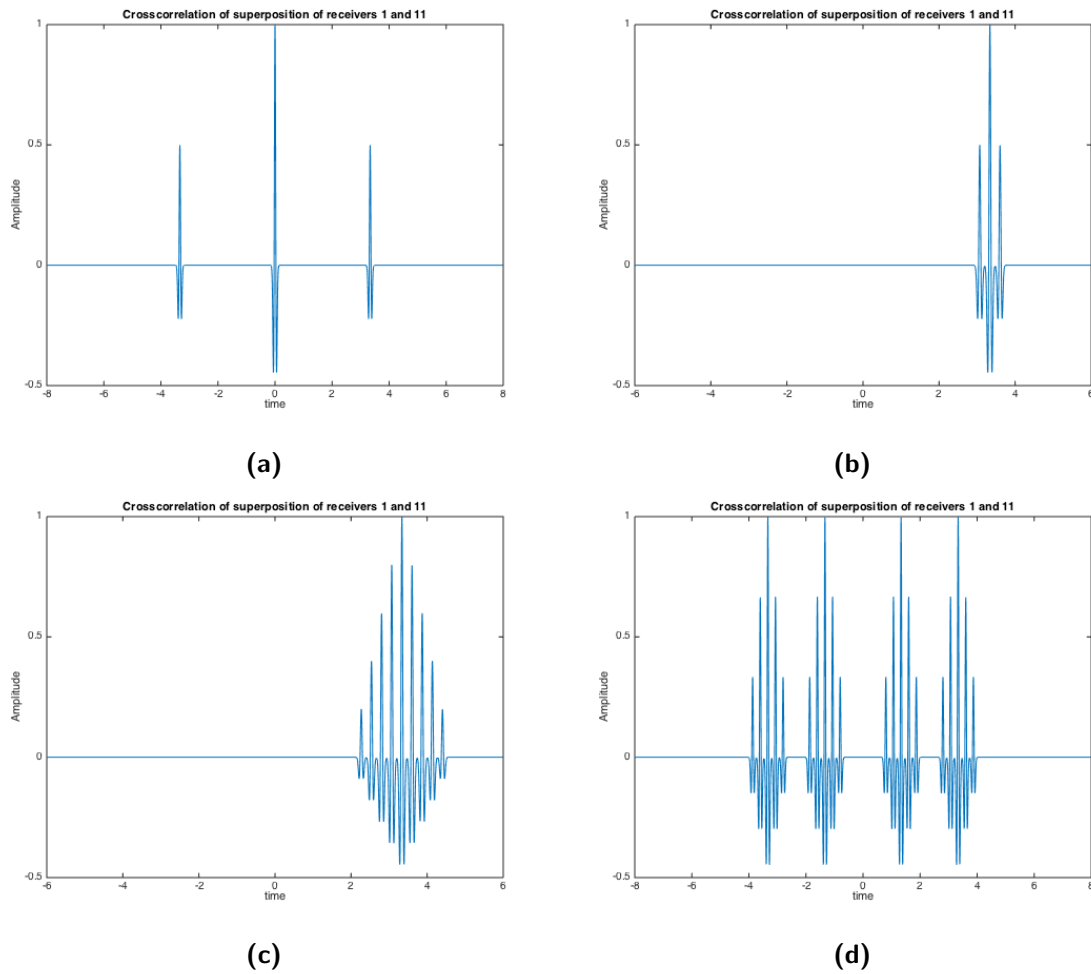


Figure 3-7: Closer look into the cross-terms

(a) Once source at each side at equal distance from the receivers array, and (b) Two west sources, (c) Five west sources and (d) Three sources on each side

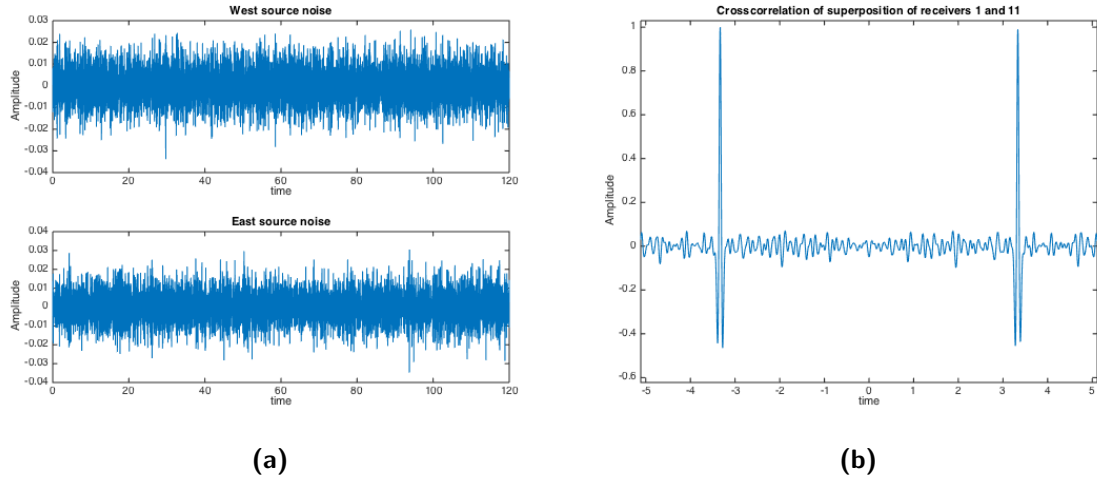


Figure 3-8: Cross-terms do not contribute in uncorrelated noise sources

(a) Two different noise series, and (b) Crosscorrelation of superposition of responses (receivers 1 and 11) using survey setup as in Figure 3-4

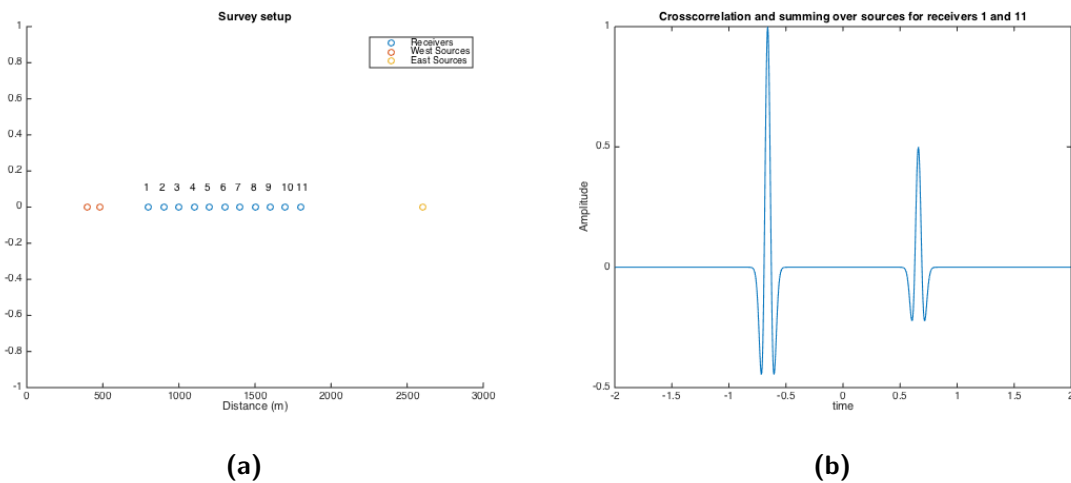


Figure 3-9: Effect of illumination in 1D setup

(a) Survey setup (two west and one east equal power sources), and (b) Crosscorrelation of responses (receivers 1 and 11) summed over sources using survey setup as in Figure 3-9a

3-2-1 Effect of illumination on retrieved responses

In the 1D models, illumination assumption is easily fulfilled as isotropic illumination in the 1D context, refers to equal illumination from the rightward and leftward propagating waves. That is, the total power coming from each direction is equal, i.e., equal number of sources on each side if their power spectra is identical. Violation of this assumption would not produce a time-symmetric response between the receivers (of which the causal part is the actual response) as seen in Figure 3-9. The situation is different in 2D and 3D as isotropic illumination means: sources are regularly distributed around the receivers, their power spectra are identical and in case of noise, the sources are mutually uncorrelated. The identical power spectra can also be replaced with the net power flux of the illuminating wavefield being (close to) zero (if the sources have different powers).

To look further into this, we first consider the numerical example shown in Figure 3-10. Here, the receivers are actual USArray receiver positions where real data is available and the sources are 10 Hz central frequency, equal power regularly spaced, every 1 degree along a circle surrounding the two receivers (Figure 3-10a). As the isotropic illumination assumption is fully fulfilled, the crosscorrelation of the observations of the two receivers completely coincide with the directly modeled response (Green's function convolved with autocorrelation of the source wavelet) as shown in Figure 3-10b. Also, the causal and acausal responses are time symmetric. Here, only the sources within the Fresnel zone contribute to the causal and acausal responses while other sources cancel each other. The Fresnel zone corresponds to the line connecting the two receivers plus a small area around this line on both sides as a consequence of the band limited sources. Those areas correspond to the peak and trough of the crosscorrelations of responses plotted as a function of source angle from the midpoint between the receivers (Figure 3-10c). If the equal power sources irregularly illuminate the receivers (Figure 3-11a), then two effects are combined. First, contributions from sources outside the Fresnel zone do not fully cancel resulting in contributions between the causal and acausal responses. Those contributions are also called "spurious arrivals" that, when appearing close to the actual response, can superimpose on it and deform the waveform Figure 3-11c. Second, the concentration of equal power sources on one side give the same effect as regularly spaces sources with higher power sources on one side than the other. Hence, the symmetry of the causal and acausal responses is violated (Figure 3-11). This effect is the same as we've seen in the 1D example (Figure 3-9). [Fan and Snieder \(2009\)](#) quantifies the spurious fluctuations that arrive between the causal and the acausal response by defining the fluctuation energy

$$E_m = \frac{1}{N_m} \sum_{i=1}^{N_m} A(i)^2, \quad (3-3)$$

in which N_m is the number of the discrete sample points in the middle part of the signal, $A(i)$ is the amplitude of the i th sample point in the middle part. This equation shows that the fluctuation energy (sum of squared amplitudes of the middle part over the number of sources) decay as a function of number of sources N . With 'enough' sources (depending on source configuration), the energy in-between the retrieved causal and acausal responses (spurious arrivals) dies out. [Weaver and Lobkis \(2005\)](#) showed that these fluctuations decay as N^{-1} if

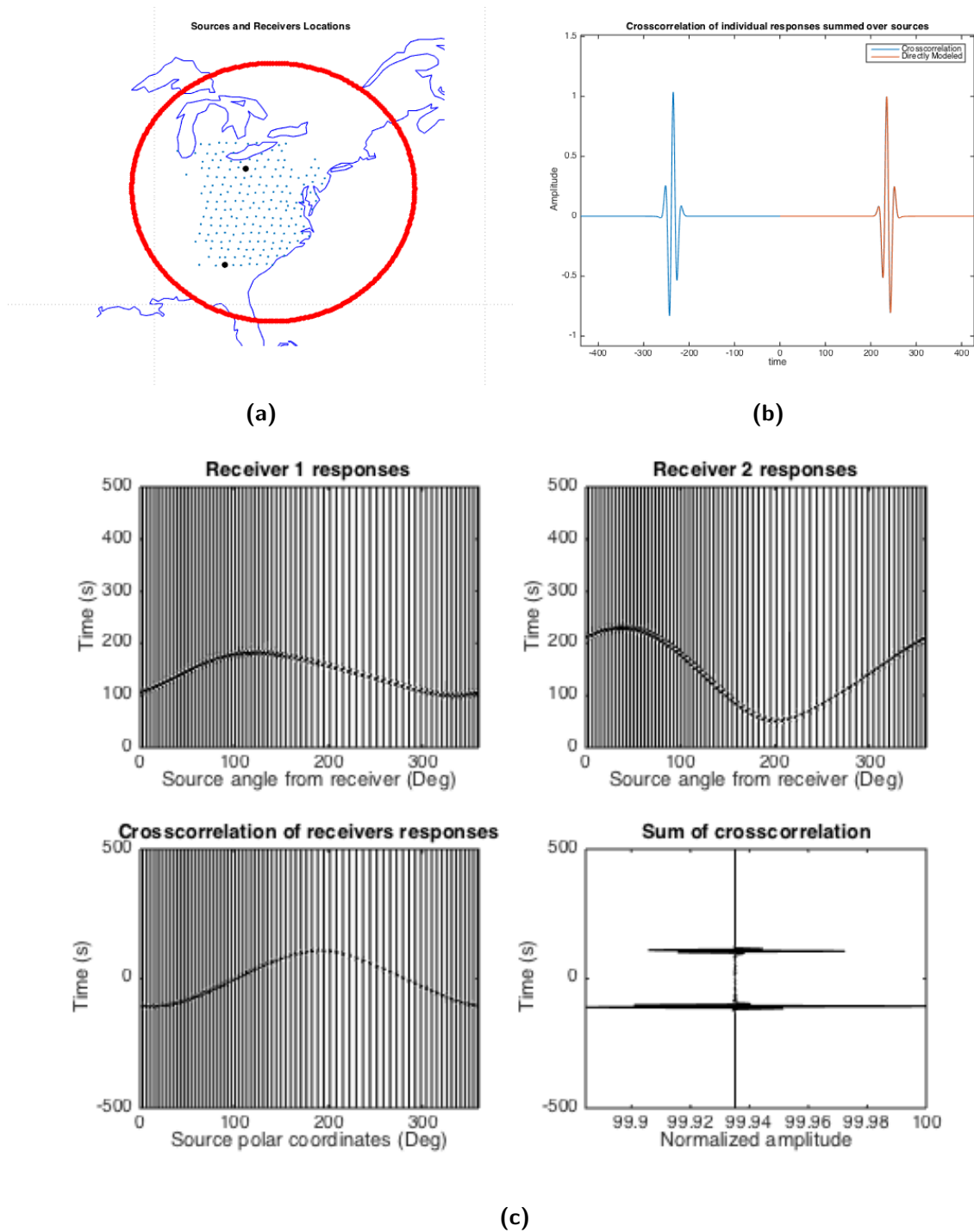


Figure 3-10: Perfect illumination

(a) Survey setup, source every 1 degree along a circle encompassing receivers, and (b) Crosscorrelation compared to directly modeled response, (c) receivers responses and crosscorrelations plotted against source angle from receivers and midpoints of receivers for the crosscorrelation, bottom right, sum of crosscorrelations shown in bottom left figure.

the sources are randomly distributed. Fan and Snieder (2009) showed that when the sources are uniformly distributed in angle, the decay rate is much faster than N^{-1} .

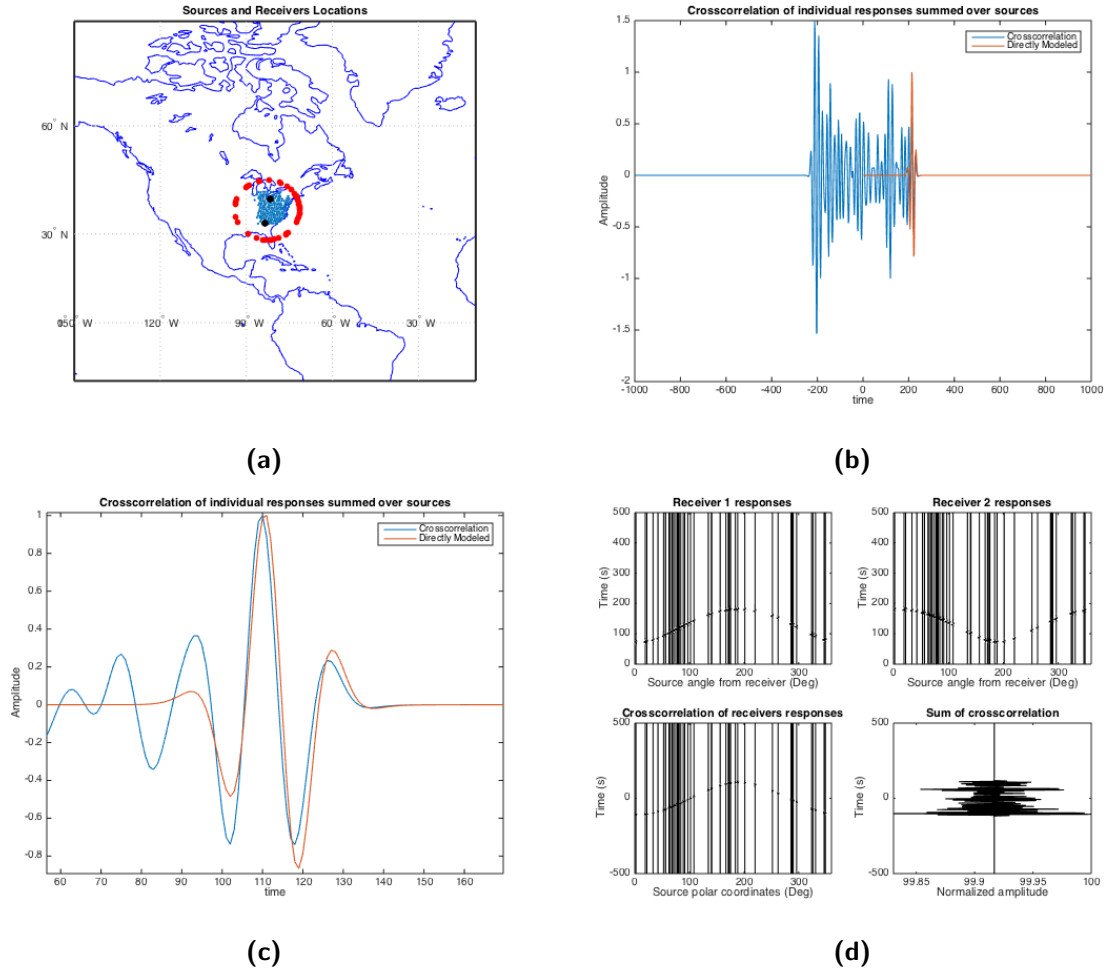


Figure 3-11: Irregular illumination

(a) Survey setup: 50 random sources along a circle encompassing receivers, and (b) Crosscorrelation compared to directly modeled response, (c) zoom into the causal part of (b) and (d) receivers responses and crosscorrelations plotted against source angle from receivers and midpoints of receivers for the crosscorrelation, bottom right, sum of crosscorrelations shown in bottom left figure.

The same spurious arrivals and time asymmetry effect is seen when the sources are regularly spaced but have different power spectra. This is demonstrated in Figure 3-12. Sources are regularly spaced, every 1 degree along a circle encompassing the receivers (Figure 3-12a). The sources are the same Ricker wavelets but with random amplitude spectra as shown in Figure 3-12c (only four randomly chosen source spectra are shown). Here, the modeling is a bit different as the underlying theory suggests. The data is synthesized by convolving the 2D acoustic Green's function (equation (3-2)) with the different source wavelets. Each crosscorrelation of responses due to a single source is multiplied (in frequency domain) by

the shaping filter $\hat{F}(\mathbf{x}, \omega)$ defined in equation (2-8) and crosscorrelations are then summed over sources. This gives, as equation (2-7) states, the Green's function (used for modeling (equation (3-2))), convolved with the average power spectrum $\hat{S}_0(\omega)$ which is shown as the directly modeled response in Figure 3-12b.

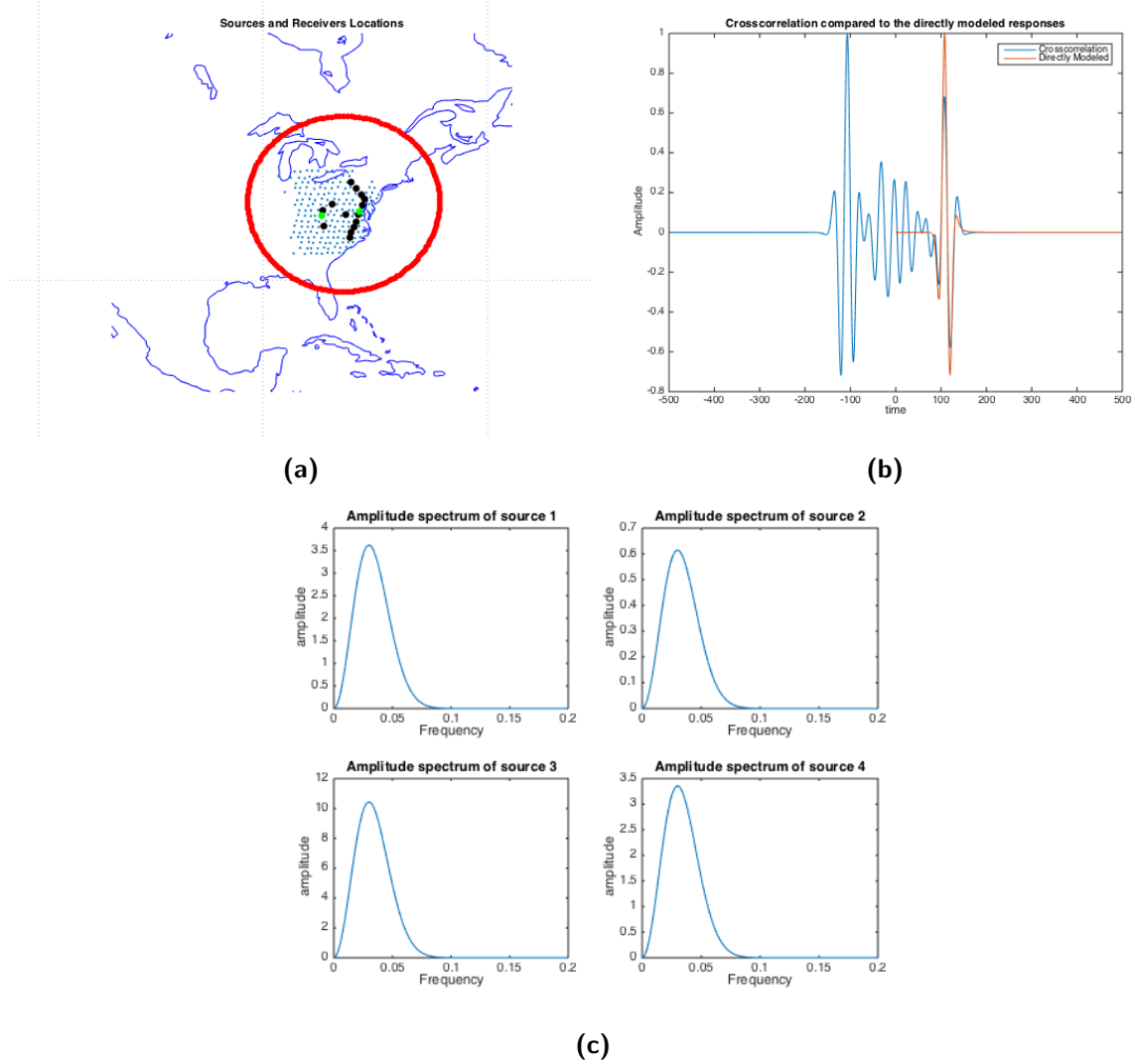


Figure 3-12: Regularly spaced sources with different power spectra

(a) Survey setup, source every 1 degree along a circle encompassing receivers, and (b) Crosscorrelation compared to directly modeled response for the two green receivers, (c) Power spectra of four randomly chosen sources.

Now we consider a one sided illumination configuration (Figure 3-13b). Here, the sources are randomly distributed along east coast USA with more south sources than north ones. The sources are equal power with a reduced central frequency of 0.05 Hz to go along with the expected surface waves frequency of the real data. Also, the velocity linearly decreases from 3600 to 3200 m/s with linearly increasing frequency from 0 to 1 Hz. As the illumination is single sided, only the causal part is retrieved and as there are more south sources than

north ones, spurious arrival are expected and as a result, deforming the waveform of the retrieved response by crosscorrelation Figure 3-13d. Since the waves are coming in only from the east, the wavefield is already separated and we only have the $u^{in}(\mathbf{x}, \mathbf{x}_S^i, t)$. Therefore, Interferometry by MDD with an absorbing boundary condition (equation (2-16)) can be directly applied. As mentioned in theory part (section 2-3), the smearing of the virtual source in space and time (caused by the superposition of the spurious arrivals on the retrieved response) is quantified by the PSF. By deconvolving the crosscorrelation by the PSF, this smearing is corrected for and a much better and focused response is retrieved (Figure 3-13e).

With noise sources, another advantage of MDD is the increased signal to noise ratio as shown in Figure 3-14d. The same experiment as in Figure 3-13 is repeated but with two main differences. First, the sources are noise and second, a cluster of sources is added in the southern part to introduce strong spurious arrivals for MDD to correct for. With the low frequencies used (0.03 Hz central frequency), the amount of noise required to achieve a good signal to noise ratio is substantially higher than that used in the higher frequency 1D experiments (Figure 3-8). Here, a total of approximately 140 days of noise is used. The results are shown in Figure 3-14d (c) and (d). MDD not only provided a better response compared to the directly modeled one, but also resulted in a higher signal to noise ratio for the same amount of noise used for crosscorrelation and MDD.

3-2-2 Reflecting boundary condition

The application of seismic interferometry by MDD assuming an absorbing boundary condition relies on the ability to separate the incoming from the outgoing wavefields as modeled in Figure 3-13 and Figure 3-14. This separation is often hard with real data and assumptions are not fully fulfilled. When a closed boundary of receivers is available, it is easier to assume a reflecting boundary condition (S_{rec} is reflecting) for the application of MDD as shown in Figure 3-15.

Here, equal power sources uniformly illuminate a concentric closed circle boundary of receivers with an extra receiver located at the center. The RBC MDD results in Figure 3-15b shows a whole train of 'non-physical' reflections that are 'wrapping around' and deteriorating the quality of the retrieved response. Since the Hankel function used to model the Green's function (equation (3-2)) does not fully attenuate to zero amplitude at longer times (only geometrical spreading; no dissipation of energy), an extra attenuation term ($e^{-\alpha r}$) is introduced, with which the Hankel function is multiplied. Here α is a small number and r is the propagation distance. This term attenuates the wavefield at a faster pace which reduces the wrap-around effect and hence eases the modelling. Equation (3-2)) becomes

$$G = -\frac{i}{4}H_0^{(2)}(kr)e^{-\alpha r}. \quad (3-4)$$

With the extra attenuation added, the non-physical reflections can be limited to certain distances as can be seen in Figure 3-15c.

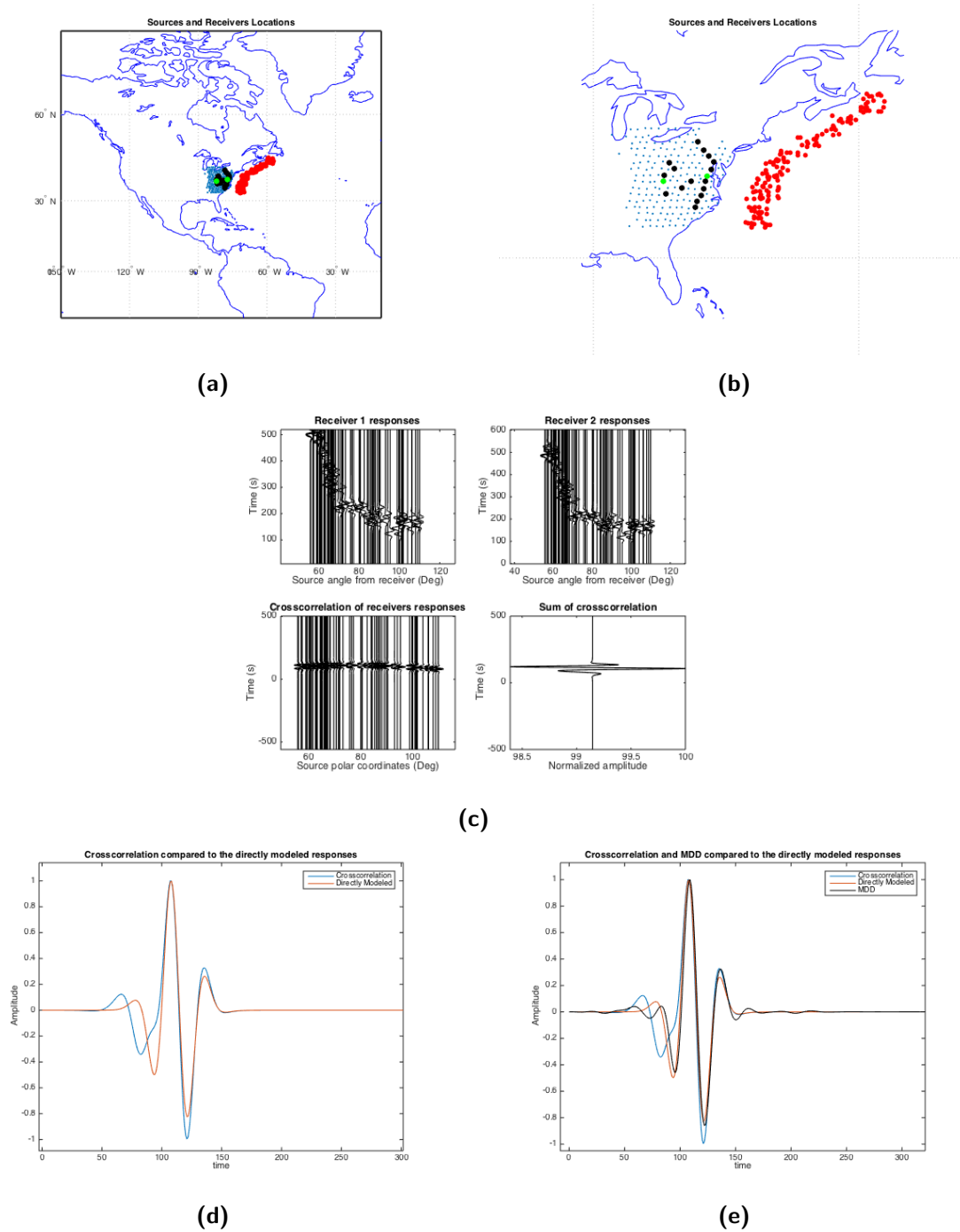


Figure 3-13: 2D one sided illumination

(a) Survey setup: random sources along east cost USA, and (b) zoom in version of (a), (c) receivers responses and crosscorrelations plotted against source angle from receivers and midpoints of receivers for the crosscorrelation, bottom right, sum of crosscorrelations shown in bottom left figure, (d) Crosscorrelation compared to directly modeled response and (e) Crosscorrelation and MDD compared to the directly modeled response

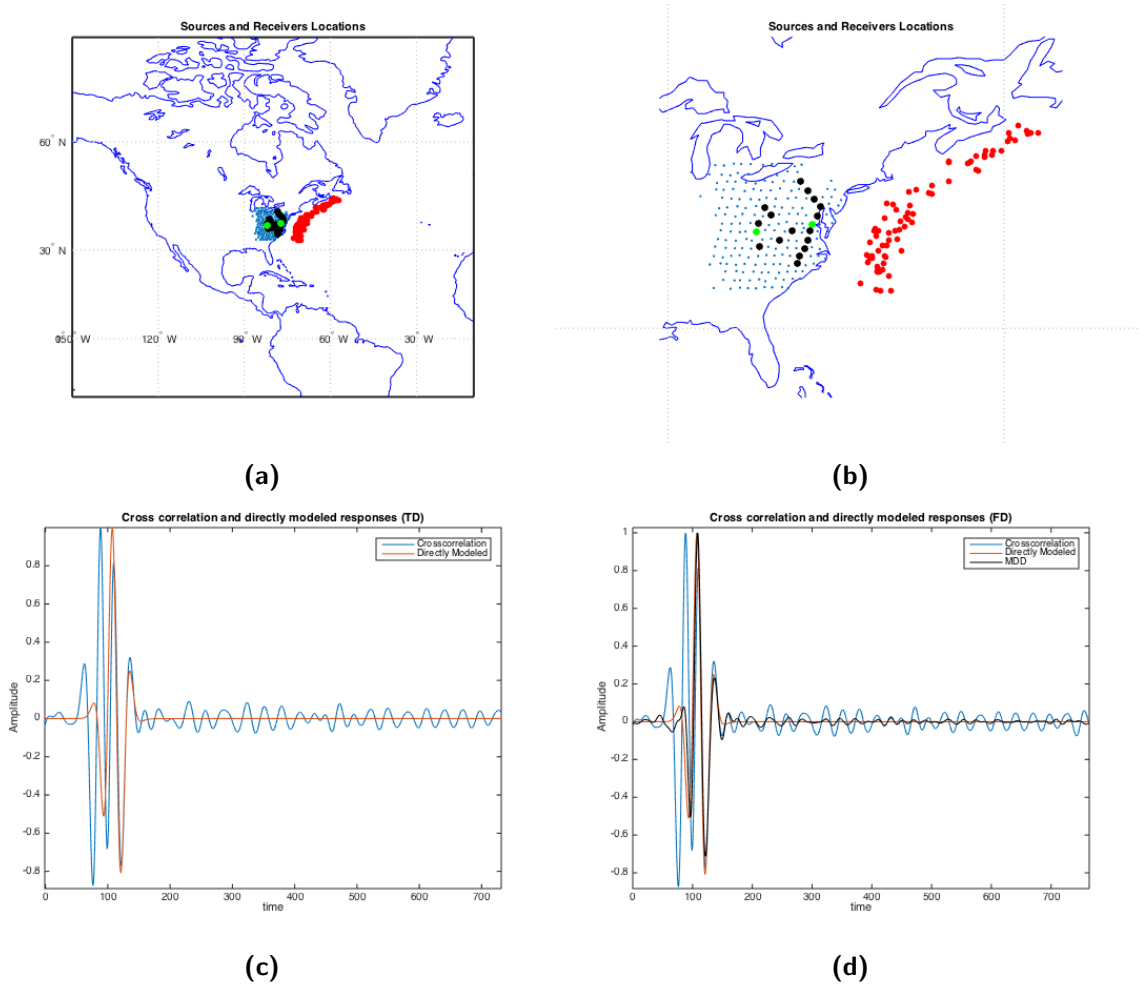


Figure 3-14: 2D one sided illumination (Noise)

(a) Survey setup: random sources along east coast USA, and (b) zoom in version of (a), (c) Crosscorrelation compared to directly modeled response and (d) Crosscorrelation and MDD compared to the directly modeled response

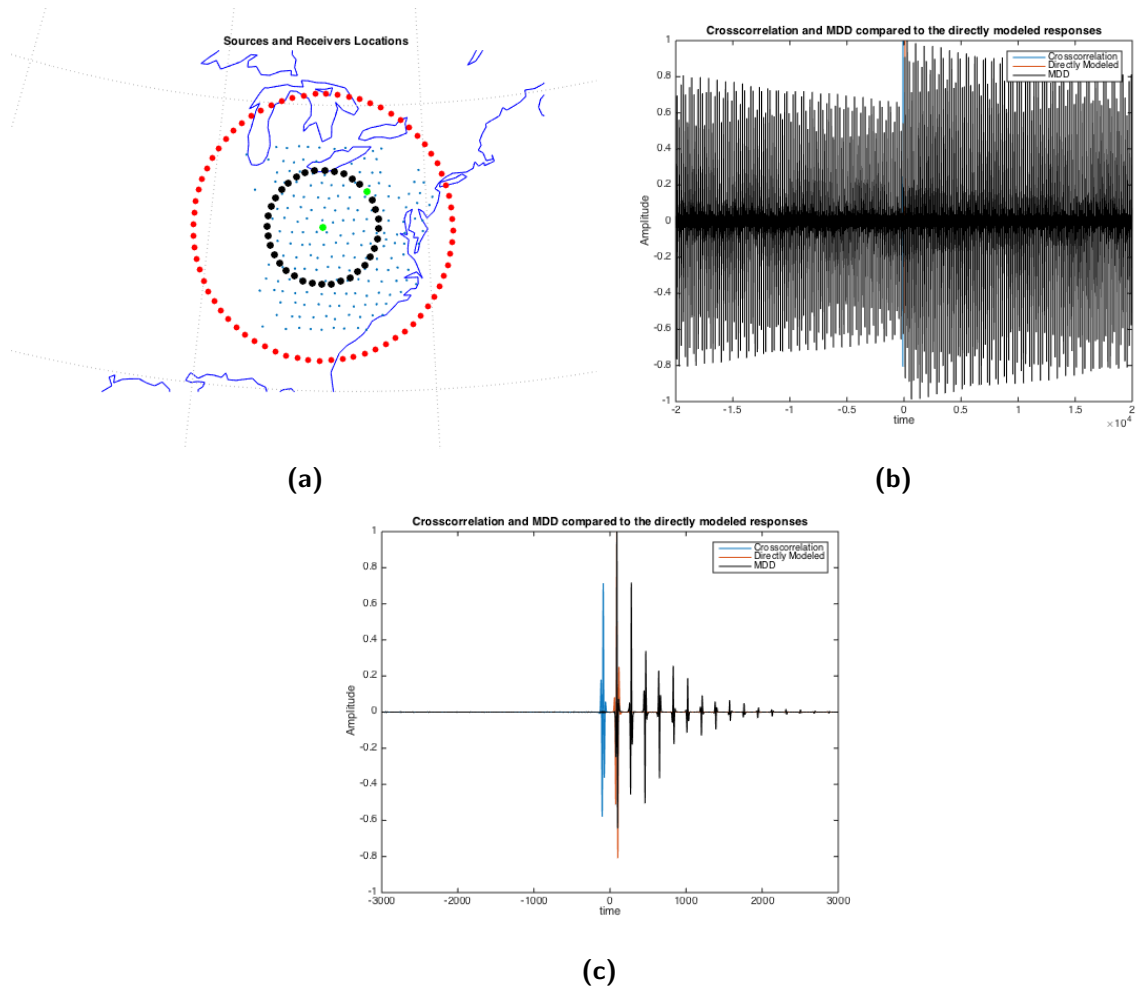


Figure 3-15: Reflecting boundary conditions

(a) Survey setup: equal power sources uniformly illuminating the receivers, (b) Crosscorrelation compared to directly modeled response and RBC MDD results, (c) Crosscorrelation and MDD with Increased attenuation compared to the directly modeled response

A RBC MDD model using actual USArray receiver locations is shown in Figure 3-16. Here, the sources are equal power, randomly distributed including clusters of sources around the closed boundary. Figure 3-16b shows the crosscorrelation compared to the directly modeled response. As discussed earlier, the source clustering, resulting in higher power from certain directions than other, introduces spurious arrivals as a consequence of the inadequate cancellation which results in deformed waveform of the retrieved response. This is corrected for in the retrieved MDD response (Figure 3-16c). However, it includes the ‘non-physical’ reflections from the receiver boundary as can be seen in Figure 3-16c. In order to prevent wrap-around effects, sufficiently long crosscorrelation windows are therefore required. In modelling, such wrap-around effect can be mitigated by increasing the attenuation rate of the retrieved MDD response (eq. (3-4)). In practice (i.e., using real data) this is of course not possible; the actual medium dissipation determines the attenuation of the retrieved MDD response. Again, the reflections can be simply zeroed out if the receivers apparatus and the frequency content of the data allows the separation of the direct response from the nonphysically reflected waves. That is, the end of the direct response is clearly distinguishable. In this configuration, it is hard to ‘model’ or anticipate the timing where reflections should show up. Reflections here are coming from all directions. In order to ‘anticipate’ when reflections should arrive, a separate model that computes the shortest distance between the virtual source and the receiver as well as second shortest distance (after reflection) and so on, for all directions, has to be computed. Even if computed and known, the short time between arriving reflections will superimpose (given the long wavelengths for the used frequencies) rendering individual reflections not clearly visible.

To improve upon the reflection scheme as well as to make the model as realistic as possible, the model in Figure 3-17 is created. Here, the receiver boundary is chosen to be as close to a perfect circle as possible using locations of actual USArray stations. Contrary to the ‘trapezoidal’ receiver boundary shown in Figure 3-16, the reflections here are solely due to the back and forth ray from the green receiver at the boundary, through the other green receiver in the center of the circle to the opposite end of the circle and back. Everything else cancels out due to the symmetry of the circle. The distance between the two green receivers is 283.4 km and the velocity is set to 3600 m/s. It can be seen from Figure 3-17c that the main arrival comes at exactly 78.8 seconds ($283.4/3.6$). The first reflection should come at 237 seconds ($(283.4*3)/3.6$) but it shows at 245 seconds and the second reflection should come at 394 seconds ($(283.4*5)/3.6$) but it shows at 435 s. This small difference is due to the fact that it is not a perfect circle and the receiver in the middle is not exactly in the center. Also, in this model, the slightly smoothed mean amplitude spectrum of the stacked real USArray data is used as source spectrum (Figure 3-17b) as can be seen in Figure 4-3c. Of course, this will not produce an identical situation as the real data because the recorded data is biased towards lower frequencies: the short periodic energy in the actual spectrum of the microseism noise has been attenuated along the path between the region of origin and the receivers. This becomes apparent from the spatial distribution of noise spectra (McNamara and Buland, 2004). However, it is more realistic than using the spectrum of a Ricker wavelet.

The results of the controlled source experiment crosscorrelation and RBC MDD compared to the directly modeled response between the green receivers is shown in Figure 3-17c and for all receivers along *Srec* as a function of azimuth in Figure 3-17e. Due to the large receiver

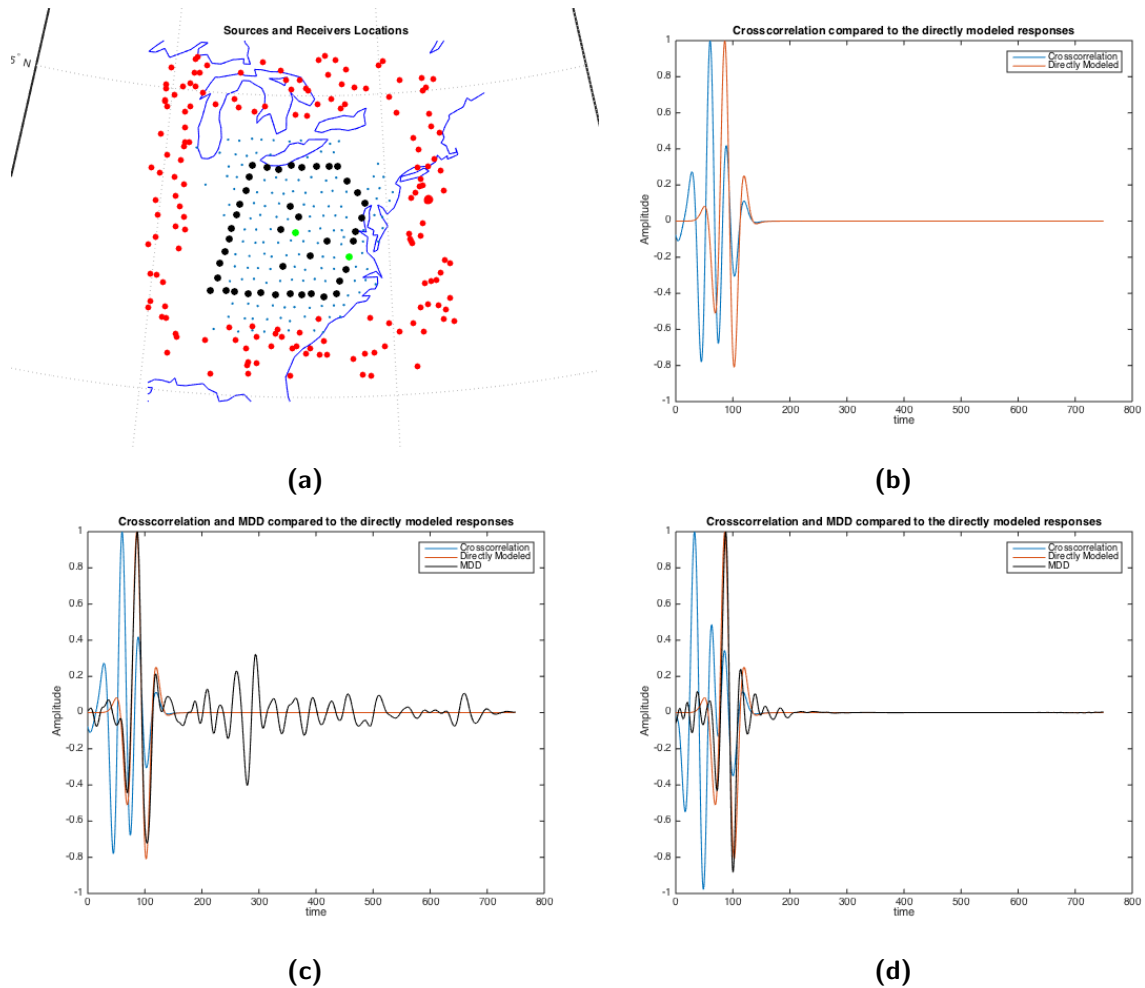


Figure 3-16: Reflecting boundary conditions using actual USArray receivers locations
 (a) Survey setup: random equal power sources around USArray receivers with source clusters, (b) Crosscorrelation compared to directly modeled response, (c) Crosscorrelation and MDD compared to the directly modeled response and (d) Increased attenuation

spacings (averaging to approximately 70 km with a maximum of approximately 95 km), the results had to be bandpass filtered between 0.02 and 0.04 Hz so it doesn't severely violate the Nyquist criteria. Still, those frequency are higher than what the configuration can support. This is evident in the MDD results where higher frequencies can still be observed at acausal times (the remaining part of it) than the causal part. Similar results achieved for all receivers along *Srec* as can be seen in Figure 3-17e. This effect is even more clearly visible in the noise results (1 month of stacked noise crosscorrelations) shown in Figure 3-17d and Figure 3-17f. For an (average) surface wave velocity of approximately 3500 m/s, the Nyquist wavelength associated with the nominal receiver spacing of 70 km is 140 km (two receivers per wavelength) supporting a maximum frequency of approximately 0.025 Hz ($2 \times 70 / 3500$). Using the maximum receiver spacing of 95 km, the maximum supported frequency computes to 0.0185 Hz ($3.5 / 190$). The bandpass filter used may therefore result in some aliasing. However, going below to even lower frequencies would substantially increase the wavelength and adding this to the limited distance available in the circle (diameter is approximately 560 km), the causal, acausal parts and reflections would be indistinguishable. Basically, it is a trade-off and clearly it would affect the the ability of MDD to improve the retrieved response in the real data.

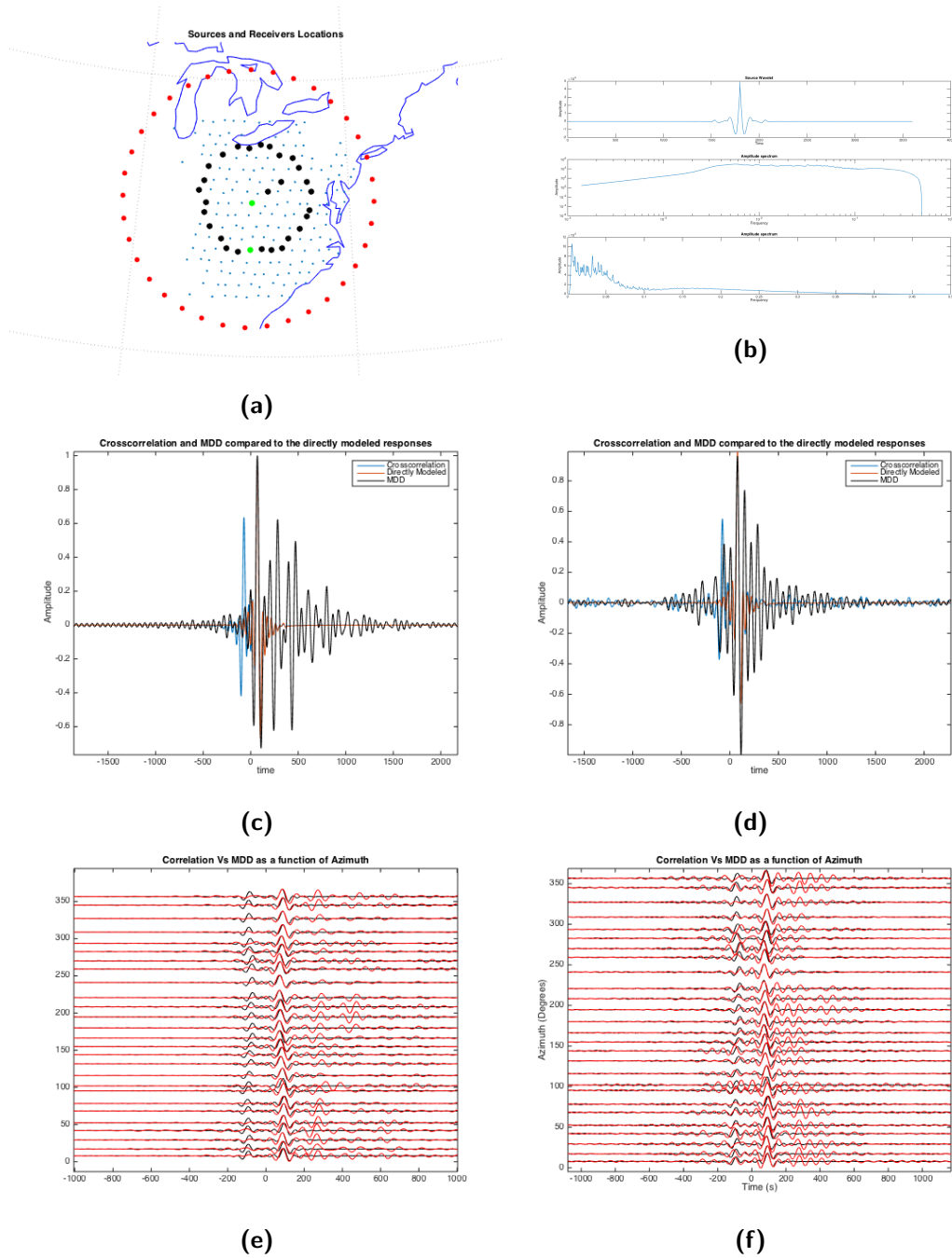


Figure 3-17: Reflecting boundary conditions using actual USArray receivers locations and real data average spectrum

(a) Survey setup: sources in red, receivers in black and receivers used for the responses in (c) and (d) in green, (b) Source power spectrum in time domain (top), amplitude spectrum on a log log scale (middle) and on linear scale (bottom), (c) Crosscorrelation and MDD compared to the directly modeled response (Controlled source) and for Noise in (d), (e) Crosscorrelation vs MDD as a function of azimuth for controlled source and for noise in (f), MDD is red, x axis is time and y axis is angle in degrees

Application to USArray Data

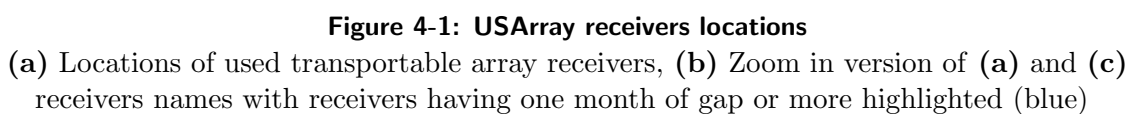
USArray is a 15-year program to place a dense network of permanent and portable seismographs across the continental United States. The seismographs record local, regional, and distant (teleseismic) earthquakes. It composes of four interrelated parts: Transportable Array, Flexible Array, Reference Network and Magnetotelluric Facility. The transportable array is composed of 400 very broadband seismographs that are placed first in the west USA with a mean spacing of 70 km, records for 18 to 24 months and then 'transported' eastwards in a rolling fashion. The flexible array is composed also of hundreds of very broadband seismographs that are dispatched to specific targets where a denser configuration is used for higher resolution mapping of structures in the upper 70 km of the lithosphere. The reference network is a uniformly spaced (300 km) permanent sensors distributed over all the USA to provide the baseline measurements for the Transportable Array and Flexible Array. In this work, only a subset of the Transportable Array data is considered.

The receivers used in the Transportable Array are a combination of: Streckeisen STS-2.0, Streckeisen STS-2.5, TrilliumV1, TrilliumV2 and Guralp CMG-3T seismometers. All receivers have three orthogonal components which are 24-bit digitized at 40 Hz and flat to velocity 0.01 to 50 Hz. Only the vertical component is considered in this work. The data loggers used are Quanterra Q330 and are the same for all 192 stations.

4-1 Available Data

A total of six months of recordings of 192 very broadband receivers, part of the Transportable Array, is considered. The data was recorded over the following time interval: 1372638600000 to 1388442600000 (Epoch timing) which is equivalent to the first of July till the 31st of December, 2013. The locations of the receivers are shown in Figure [4-1](#).

The data set, in general, has a very good continuity of recordings. However, some major gaps are still present. Figure [4-2](#) shows the availability of data per receiver station as a function of



time. As can be seen, 3 stations (L47A, O48A and X55A) suffer a major gap of one month or more of continually missing data. Those receivers locations are shown in Figure 4-1c.

4-2 Preprocessing

The original data set was digitized at 40 Hz, that is, a sample was taken every 0.025 seconds. This involves a lot of computations required in the application of crosscorrelation. Also, it is way oversampling the sought after hum and microseism frequencies (< 0.3 Hz) that will be used for the application of interferometry in this work. Therefore, it was first low-pass filtered (below 0.5 Hz (Nyquist of the new data set)), to avoid aliasing, and then decimated to 1 Hz. Since the data was recorded using a combination of different seismographs, the instrument response has to be corrected for to homogenize the data. The poles and zeros of each receiver type was acquired through the manufacturer's websites and the instrument response was then deconvolved from the data between 0.0025 and 0.5 Hz with a cosine taper using the following four poles: 0.0025, 0.005, 0.25 and 0.5 Hz. Since data loggers are the same for all receiver stations, no correction needed here. The data was then broken down to files of 1 hour with 50% overlap and a 3 minutes cosine taper in time domain is applied on both sides of each window (file). Finally, the data was converted from velocity to displacement which explains the relative high amplitudes at lower frequencies. These preprocessing steps were done in collaboration with supervisor Dr. Cornelis Weemstra.

The frequency content of the full data set is shown in Figure 4-3. It can be seen that the primary and secondary microseism peaks (≈ 0.06 and 0.22 Hz respectively) are actually present (more clearly in the low band). However, they are unusually overdominated by hum frequencies (< 0.03 Hz) which can be attributed to the presence of strong earthquake events and other large transient signals in the recorded dataset. In order to suppress those events, three methods separately tested. Running-absolute-mean normalization (RAMN), spectral normalization by mean amplitude of the frequency window of interest and spectral whitening.

The RAMN method is based on the following equation (Bensen et al., 2007)

$$d_n^{norm} = \frac{d_n}{W_n}, \text{ and } W_n = \frac{1}{N} \sum_{k=n-N}^{n+N} |d_k|, \quad (4-1)$$

where d_n^{norm} is the normalized data and W_n is the normalization coefficient. Bensen et al. (2007) determined an optimal window length of half the maximum period. This computes, for this data set to $\frac{1}{4f_{min}} = 25$ seconds and $N = 25 * 1 = 25$. The second method is basically normalizing all windows amplitude spectra by the mean of the amplitude spectrum in the frequency window of interest (0.02 - 0.04 Hz). The main motivation behind this is to boost the frequencies of interest. That is, make them dominate the spectrum. Ofcourse, this will only be true in the case that most 1 hour windows have dominating frequencies in that range. If most windows are dominated with the lower frequencies, then this method will only boost those lower frequencies. Last but not least is the spectral whitening. Basically, normalizing narrow amplitude spectrum windows by its mean to flatten the spectrum. Here, a window of 50 frequency samples is used (0.0138 Hz).

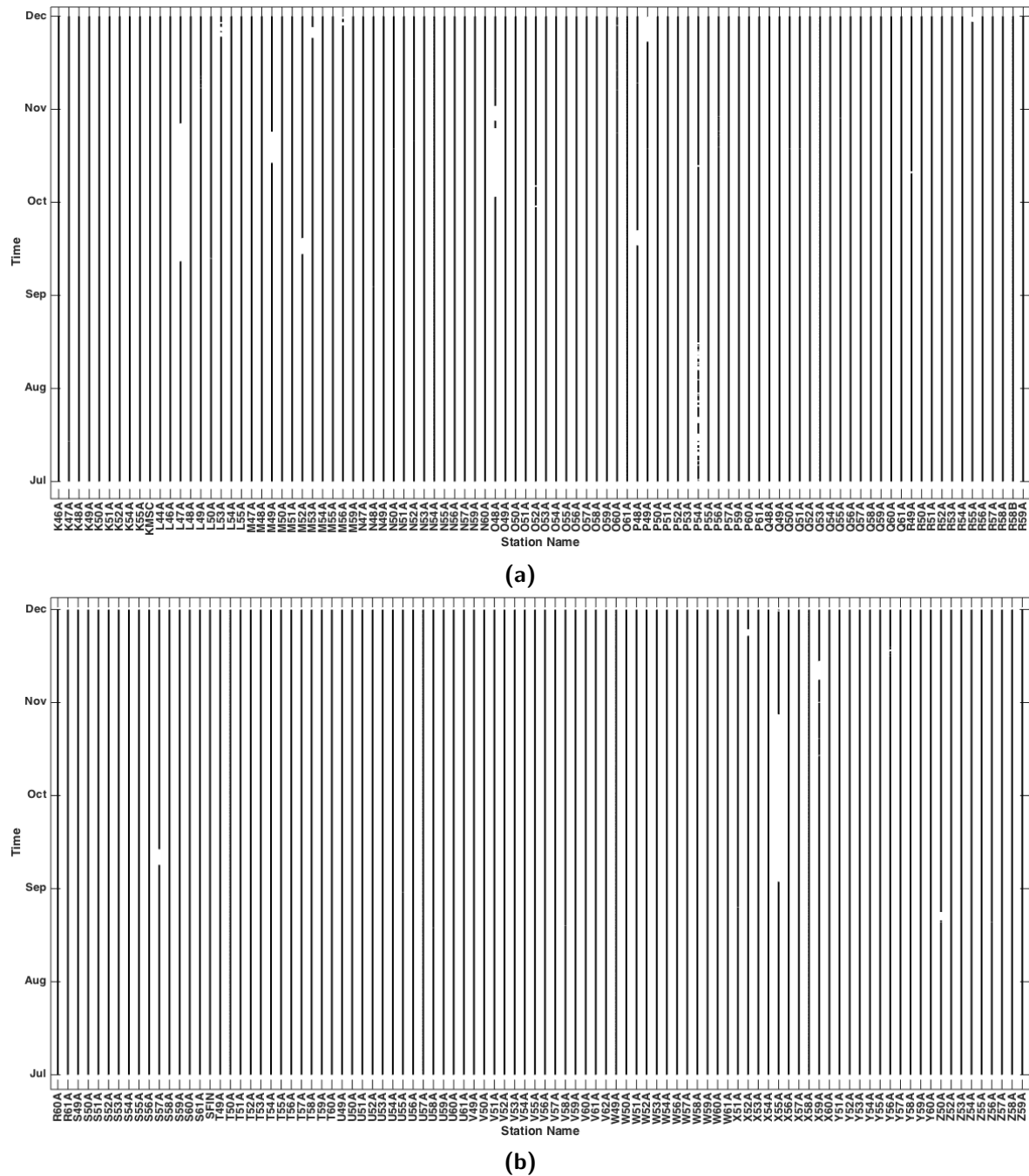


Figure 4-2: Available USArray Data

(a) The available data per receiver station as a function of time and (b) continuation of (a)

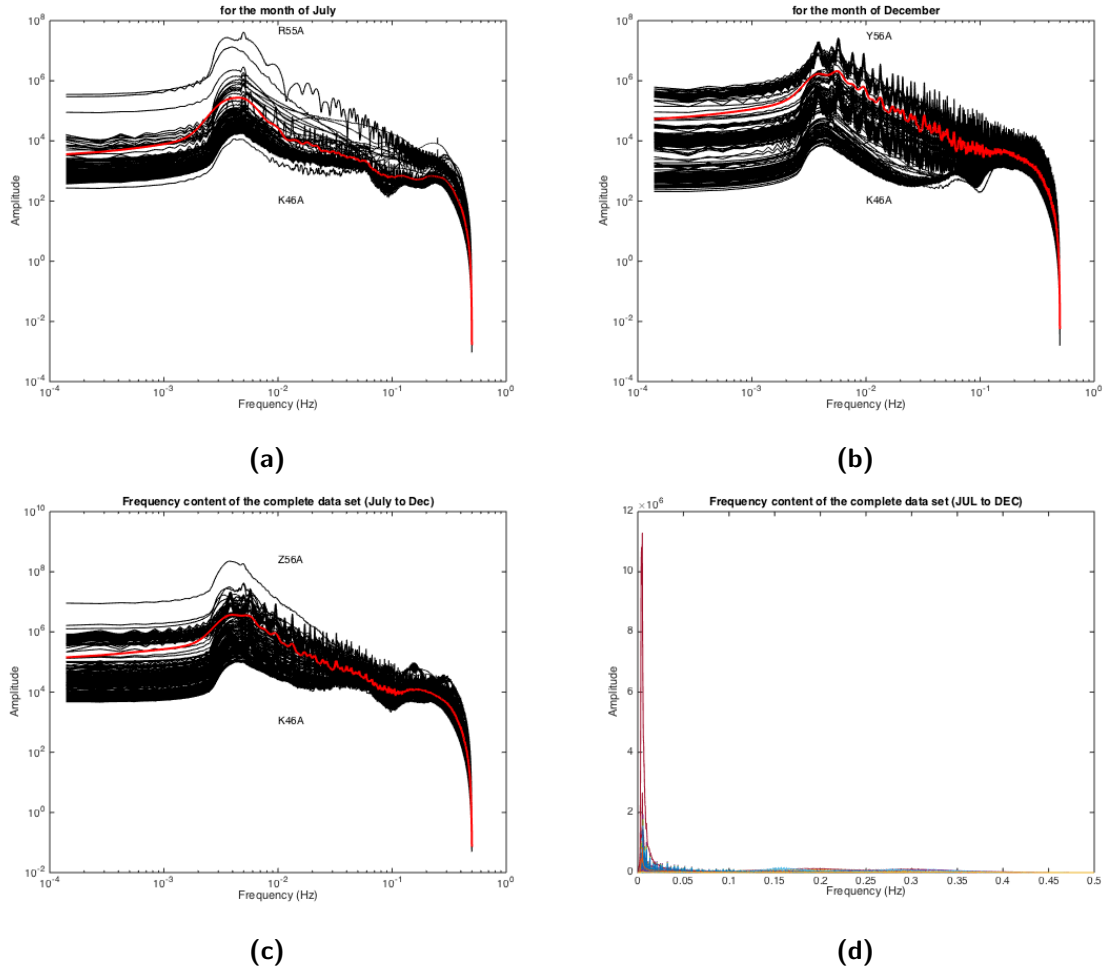


Figure 4-3: Frequency content of the USArray data set

(a) Amplitude spectrum for all 192 receivers in summer (July) and for winter (December) in (b), (c) Amplitude spectra for all 192 receivers for the complete dataset (July to December) on a log log scale and on a linear scale in (d). Slightly smoothed mean (red) is shown in each plot as well as the names of stations with lowest and highest spectra

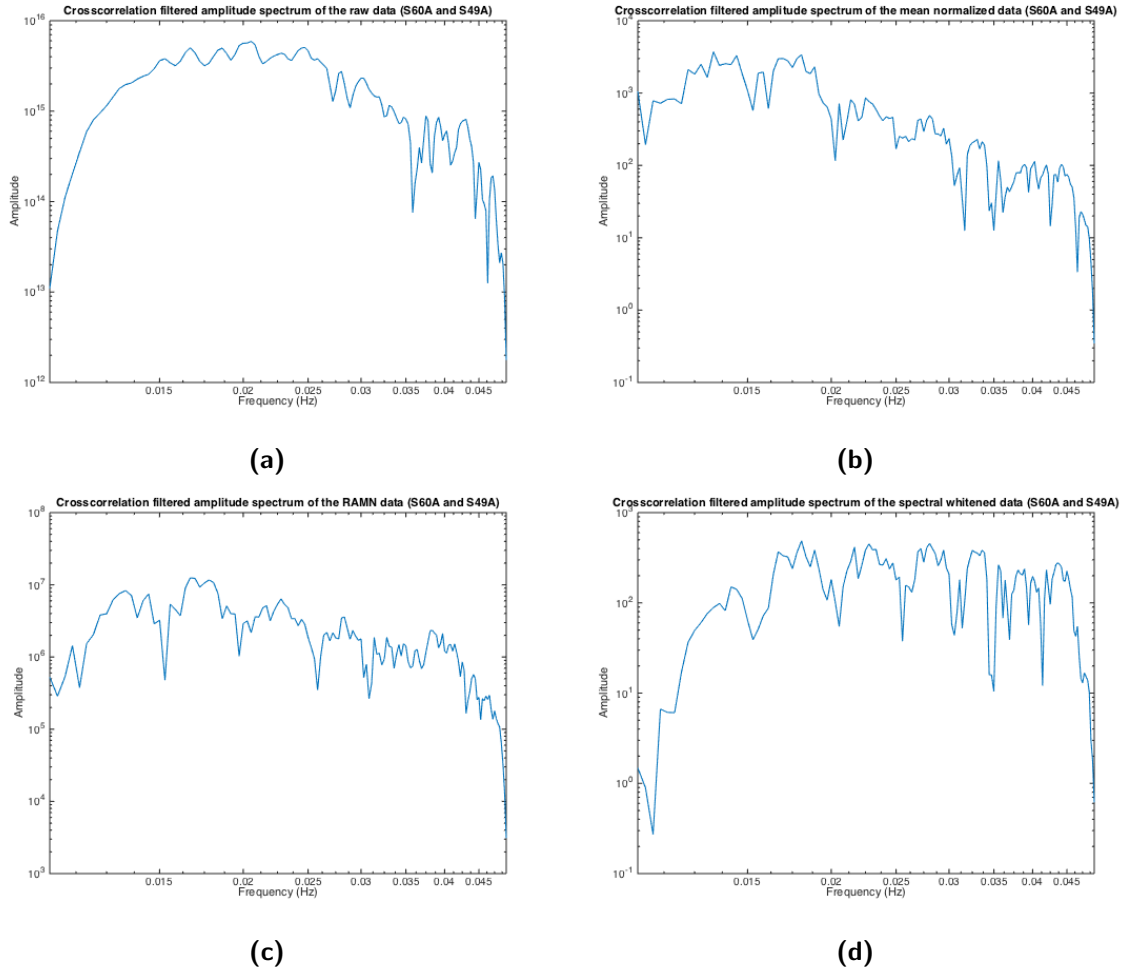


Figure 4-4: Spectral homogenization of the USArray data set

(a) Sample crosscorrelation filtered amplitude spectrum for the raw data, (b) for the mean normalized data, (c) for the RAMNed data and (d) for the spectral whitened data. All crosscorrelations are between stations S60A and S49A

A sample 6 months east west stacked crosscorrelation spectrum (between stations S60A and S49A) is shown in Figure 4-4. All the three methods were applied on the raw 1 hour window data prior to crosscorrelation, stacking and filtering. The stacked crosscorrelation spectrum is then filtered using a four poles filter with a cosine taper (0.01 0.02 0.04 0.05 Hz). Those frequencies, as discussed in the last model of the numerical modeling chapter are the ones that will be used for the application of RBCMDD. It can be seen from the raw cross-spectrum that it is dominated by frequencies less than 0.01 Hz (350 km wavelength based on 3500 m/s velocity). Given that the radius of the MDD circle receivers along *Srec* is approximately 280 km, the retrieved RBCMDD response would be degraded by the superposition with the reflections and therefore, the spectrum has to be more or less dominated by higher frequencies. This is best done by the spectral whitening method as can be seen in Figure 4-4d and therefore, it will be used for the application of RBCMDD.

4-3 Application of Interferometry

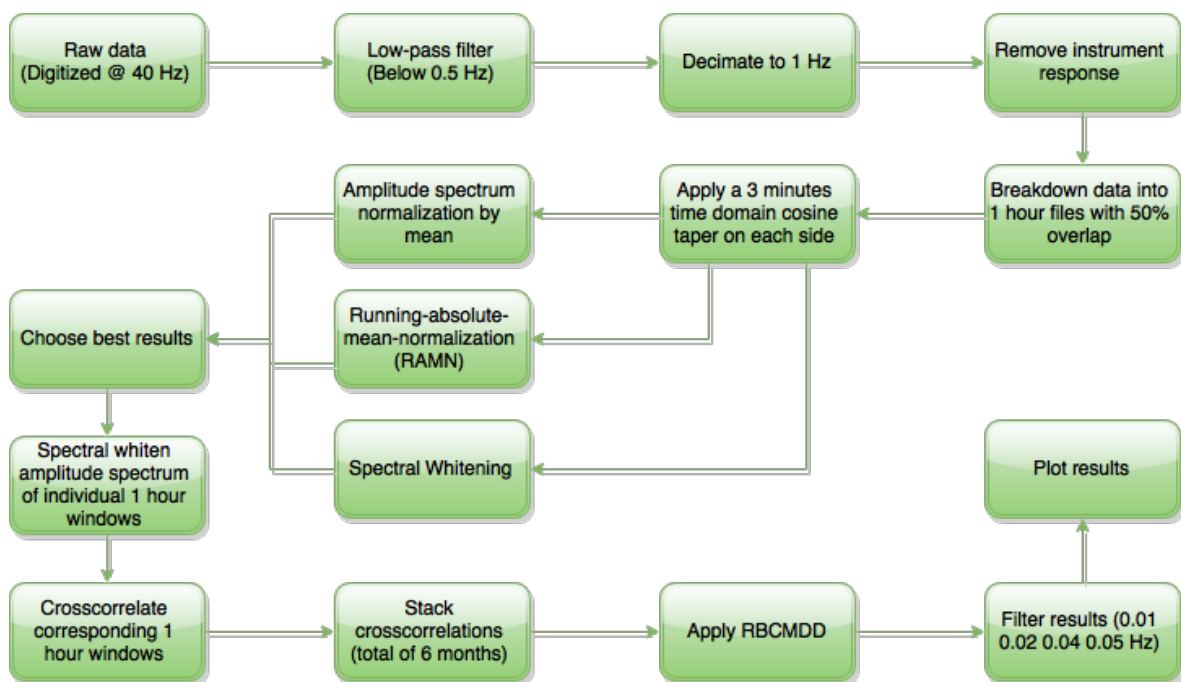


Figure 4-5: Workflow flowchart

Flowchart summarizing processing steps from raw data to results

A flow chart summarizing the workflow used to achieve results is shown in Figure 4-5. Crosscorrelations along an east west line using the three spectral manipulation methods are shown in Figure 4-6, along a north south line in Figure 4-7 and a comparison of all of them in Figure 4-8. Comparing the crosscorrelations of each method with the others, it can be seen that all methods produce similar results with the spectral whitened data crosscorrelations looks in general cleaner, more focused and better. Along the east west line, it is notable that the causal and acausal parts don't fully separate until distances of 550 km and larger. At

this distance, they are a bit less than two wavelengths apart. On the north south line, it can be seen that the signal-to-noise ratio is way lower than that of the east west results. Also, most of the traces are lacking the causal part of the response. This can be interpreted as not having enough north and south sources (or intensity of sources). As microseism events/noise are generated from water waves along oceans and lakes, at this location, stronger oceans are more abundant in the east and west than the little lakes in the north.

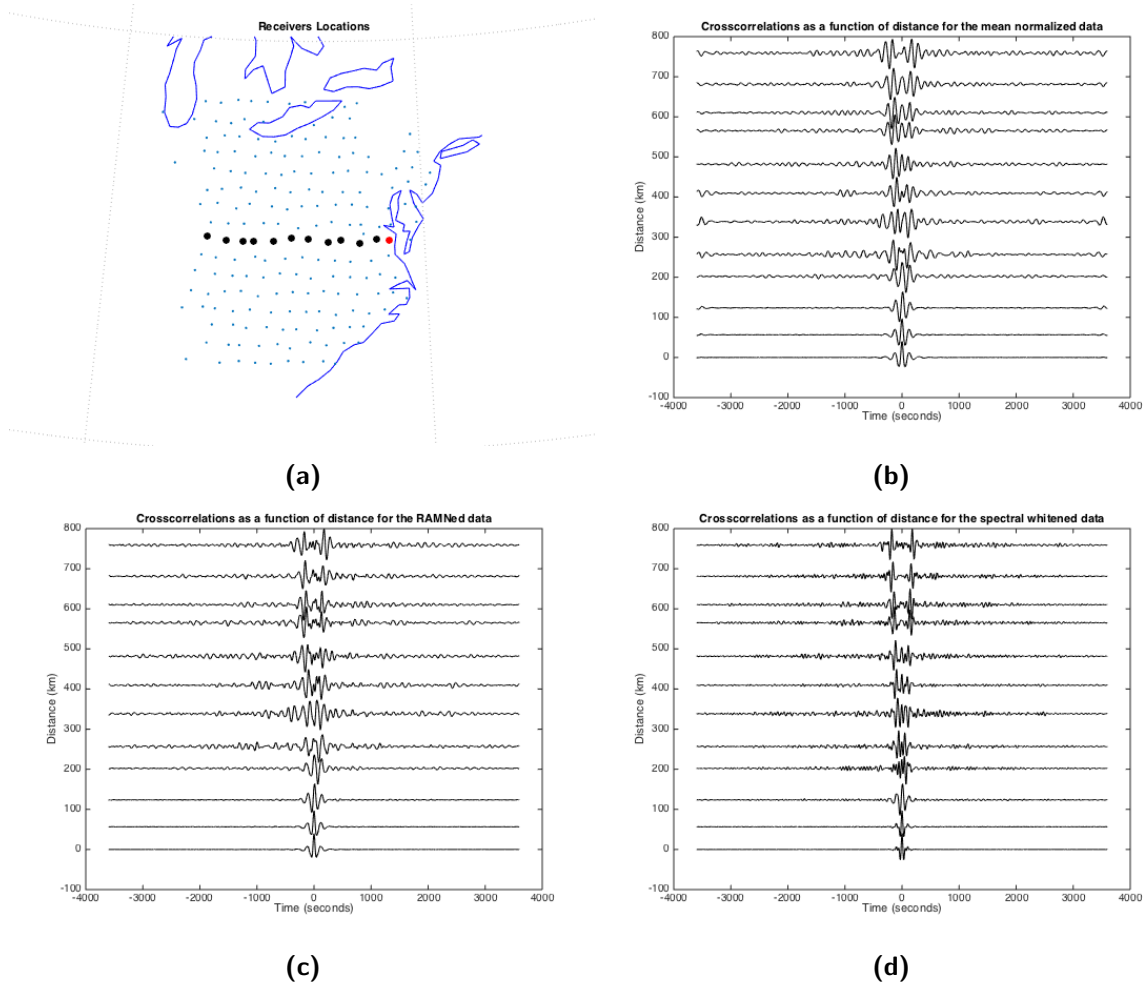


Figure 4-6: East west line crosscorrelations

(a) East west receivers locations with red being virtual source, (b) Crosscorrelations as a function of distance for the mean normalized data and for the RAMNed data in (c) and for the spectral whitened data in (d)

The results of RBCMDD on the complete data set (stacked 6 months of data) is shown in Figure 4-9. Those results are acquired by using the same procedure as the crosscorrelations in Figure 4-6 and Figure 4-7. That is, load raw one hour windows, apply a time domain three minutes taper on both ends, whiten the spectrum (14 frequency samples window * 0.00028 Hz per sample), crosscorrelate, stack, apply MDD and bandpass filter (0.02 0.04 Hz). It can be seen that there isn't much improvement on the retrieved response. This can be attributed,

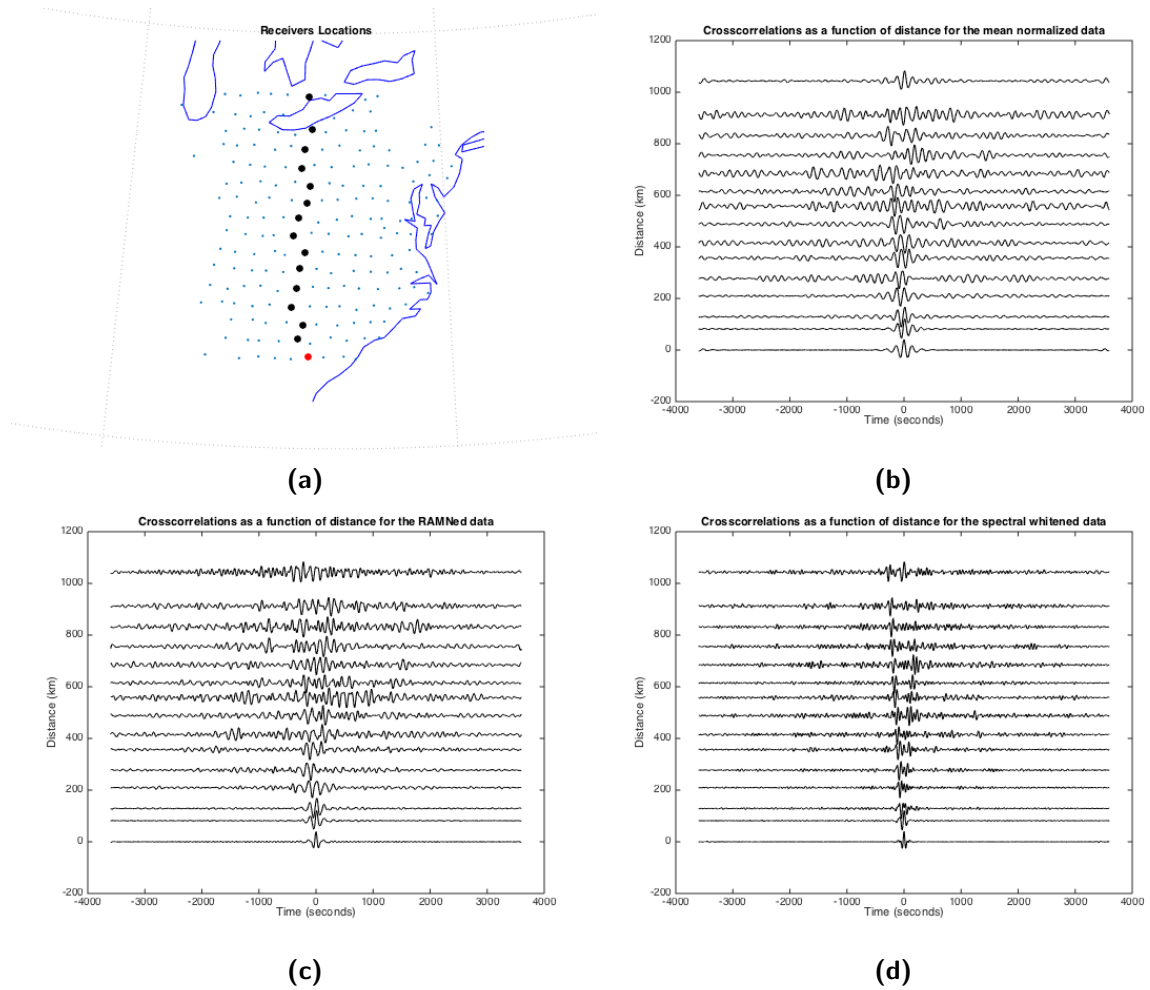


Figure 4-7: North south line crosscorrelations

(a) North south receivers locations with red being virtual source, (b) Crosscorrelations as a function of distance for the mean normalized data and for the RAMNed data in (c) and for the spectral whitened data in (d)

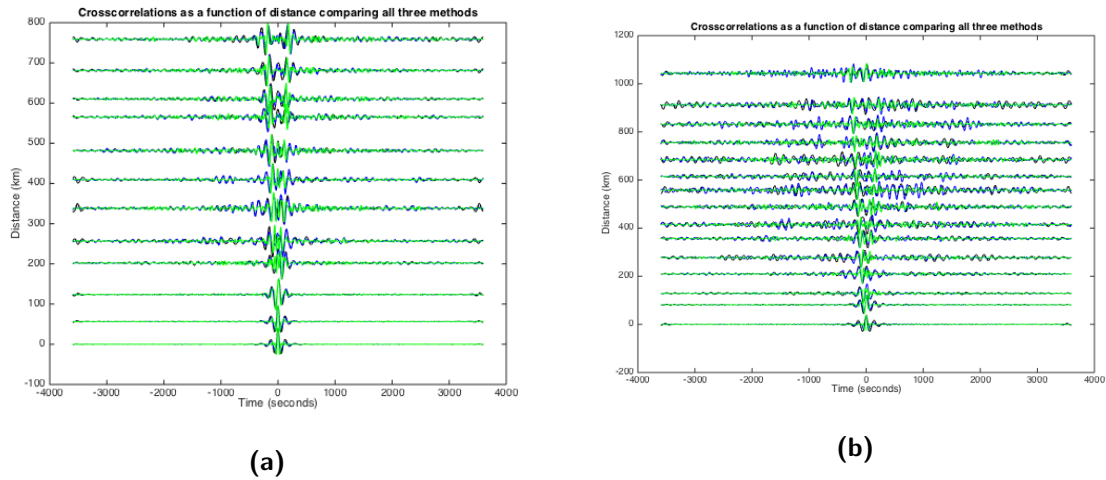


Figure 4-8: Comparison of crosscorrelations

(a) Comparison of crosscorrelations of recordings along the east west line (Figure 4-6a) using three spectral homogenization methods and for the north south line (Figure 4-7a) in (b). Black: mean normalization, blue: RAMN and green: spectral whitening

as mentioned earlier, to the aliasing of the frequencies used rendering the ability of MDD to deconvolve the temporal and spatial smearing. Also, from the south-north line (Figure 4-8), it can be seen that the causal part is mostly lacking. This is due to the lack of sources (or source intensity) in the north as the model in Figure 3-9 demonstrates. The model shows that more source intensity from one side produces stronger causal/acausal peak than the counterpart and if sources are totally lacking from one side, only causal/acausal peak will be retained (Figures 3-2 and 3-3). The two sources just to the west of receiver 1 in Figure 3-9a can be compared to the energy coming from the Gulf of Mexico while the source to the east of receiver 11 is comparable to the energy coming from the north lakes. The same effect can be seen in the 2D model shown in Figure 3-11. While MDD corrects for irregular illumination, it doesn't improve responses with 'illumination gaps'.

To take a closer look into the distribution of energy coming to the array, a beamforming analysis is performed (in collaboration with supervisor Dr. Elmer Ruigrok). In order for results not to be biased, a uniformly wide range of azimuths and distances needs to be sampled. The PSF computed to get results in Figure 4-9 exactly fits this purpose and therefore, it was used as an input for the beamforming algorithm. The results are shown in Figure 4-10. It can be seen that for the lower frequencies (< 0.015 Hz), there is actually a good non-uniform illumination. However, with increasing frequencies, major illumination gaps are present which are, again, hindering the ability of MDD to improve the retrieved response. If only the lower frequencies are used, the receivers apparatus won't allow for the separation of the causal, acausal and non-physical reflections (all will be superimposed into one event). If the receivers boundary had at least double the current diameter (with the same current receivers spacing), RBCMDD would improve the retrieved response by crosscorrelation on this data set for those lower frequencies.

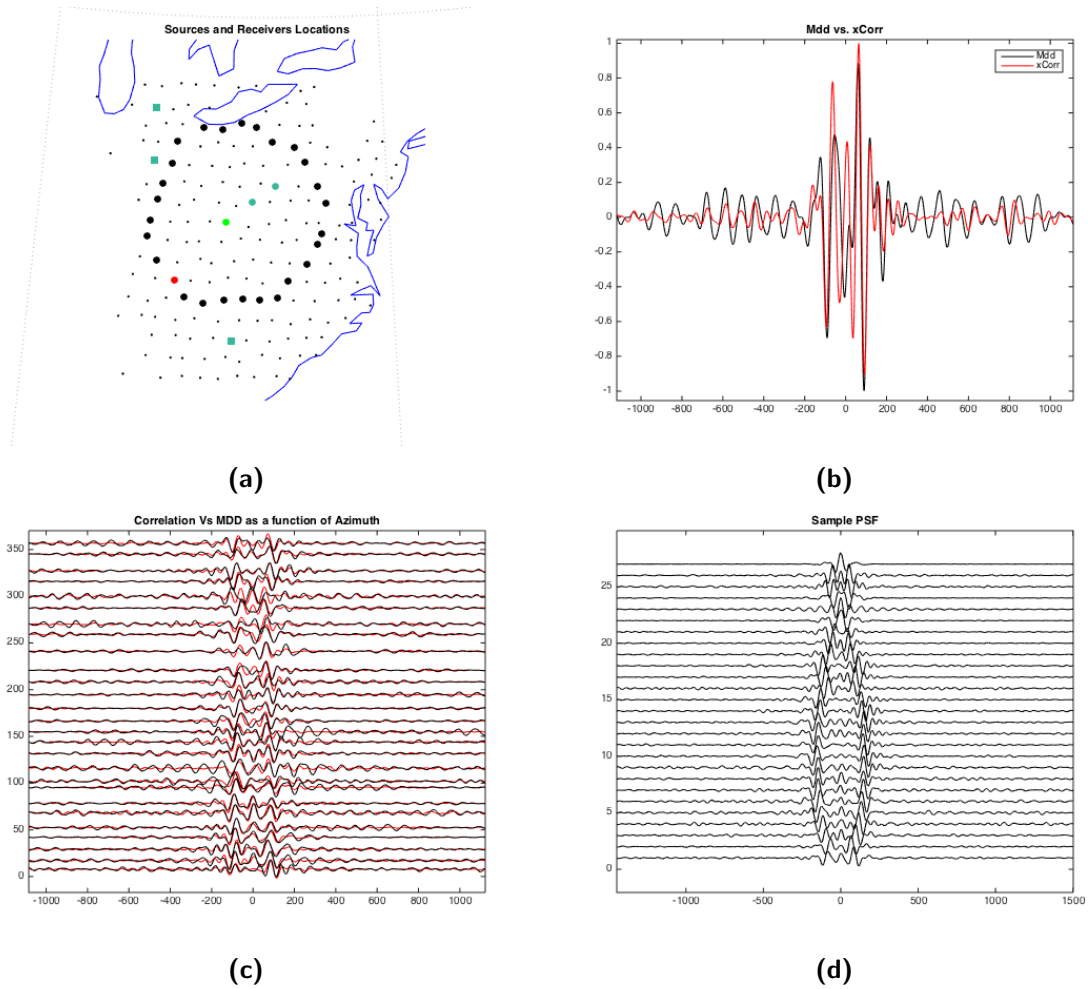


Figure 4-9: Application of RBCMDD on the USArray data set

(a) Receivers set up, MDD receivers along S_{rec} in black, receivers with major data gaps in green squares, red is virtual source location and bright green circle is receiver used for the response shown in (b). (c) Crosscorrelations compared to RBCMDD retrieved responses as a function of azimuth, (d) Sample PSF (corresponding to the virtual source)

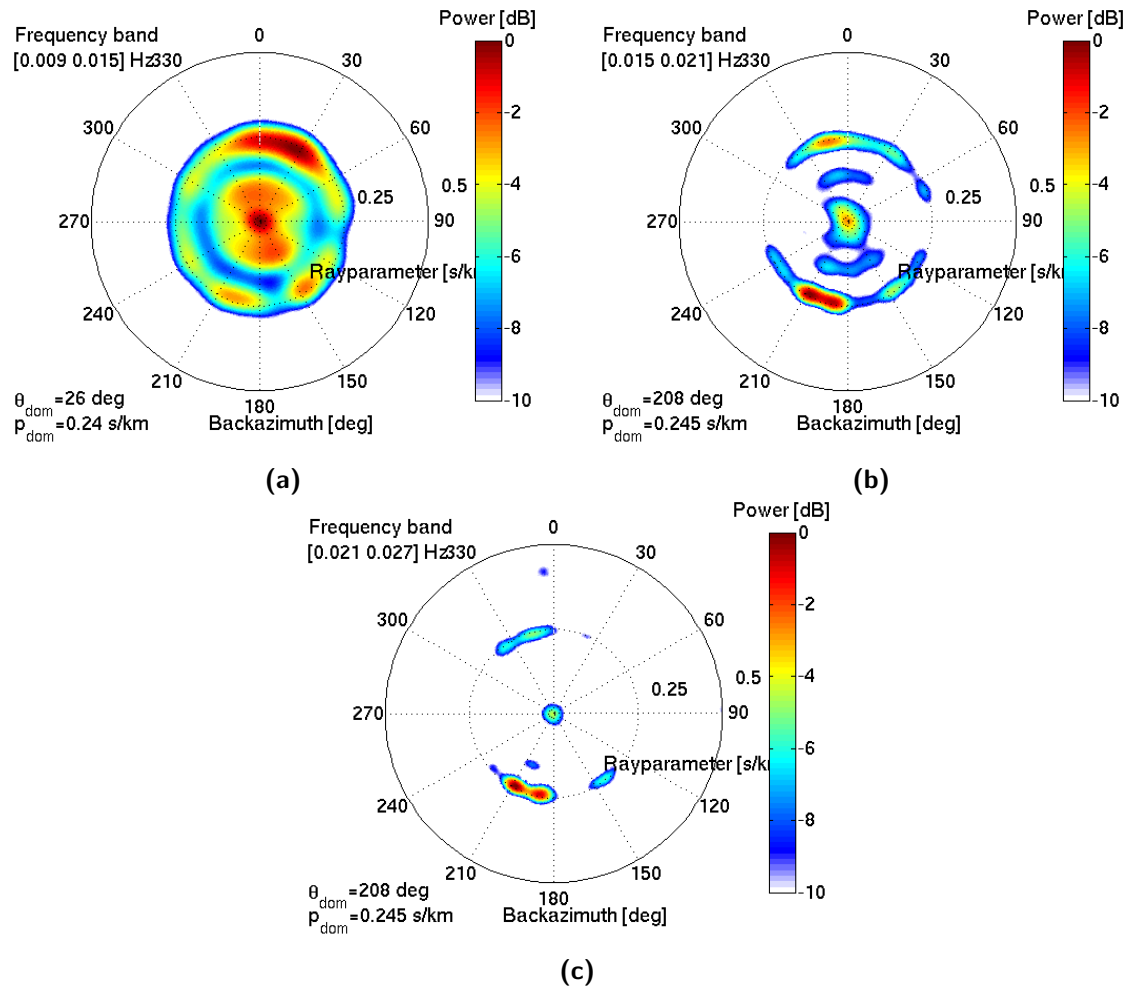


Figure 4-10: Beamforming analysis on the USArray data set

(a) Beamforming analysis for frequency band 0.009 - 0.015 Hz, (b) for 0.015 - 0.021 Hz and (c) for 0.021 - 0.026 Hz

Chapter 5

Conclusion

Seismic interferometry by RBC MDD has been applied on a subset of the Transportable US-Array data. The revised formulation eliminates the need to separate incoming from outgoing wavefields by assuming a reflecting boundary condition. The ability of response improvement is the same of that for an absorbing boundary condition with the distinction of the presence of non-physical reflections from the closed receivers boundary in the retrieved response. On this data set, not much response improvement is gained by applying RBC MDD over conventional seismic interferometry by crosscorrelation. This is attributed to the presence of illumination gaps as well as large nominal receivers spacing limiting the maximum frequencies that can be used. Also, small receivers aparature limiting the minimum frequencies that can be used.

Bibliography

- Ardhuin, F., Stutzmann, E., Schimmel, M., and Mangeney, A. (2011). Ocean wave sources of seismic noise. *Journal of Geophysical Research*.
- Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Gophysical Journal International*.
- Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, 33:264–269.
- Cole, S. (1995). *Passive seismic and drill bit experiments using 2-D arrays*. PhD thesis, Stanford University.
- Derode, A., Larose, E., Campillo, M., and Fink, M. (2003). How to estimate the greens function of a heterogeneous medium between two passive sensors? application to acoustic waves. *Applied Physics Letters*.
- Draganov, D., Wapenaar, K., Mulder, W., Singer, J., and Verdel, A. (2007). Retrieval of reflections from seismic background-noise measurements. *Geophysical Research Letters*.
- Fan, Y. and Snieder, R. (2009). Required source distribution for interferometry of waves and diffusive fields. *Gophysical Journal International*.
- Hoop, A. T. D. (1988). Time-domain reciprocity theorems for acoustic wave fields in fluids with relaxation. *Journal of the Acoustical Society of America*.
- Longuet-Higgins, M. S. (1950). A theory of the origin of microseisms. *Philosophical Transactions of the Royal Society*.
- McNamara, D. E. and Buland, R. P. (2004). Ambient noise levels in the continental united states. *Bulletin of the Seismological Society of America*.
- Rickett, J. and Claerbout, J. (1999). Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring. *The Leading Edge*, 18:957960.

- Scherbaum, F. (1987a). Seismic imaging of the site response using microearthquake recordings: Part i. method. *Bulletin of the Seismological Society of America*.
- Scherbaum, F. (1987b). Seismic imaging of the site response using microearthquake recordings: Part ii: Application to the swabian jura, southwest germany, seismic network. *Bulletin of the Seismological Society of America*.
- Schuster, G. T. and Zhou, M. (2006). A theoretical overview of model-based and correlation-based redatuming methods. *Geophysics*.
- Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*.
- Snieder, R. (2004). Extracting the greens function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*.
- Snieder, R., Sheiman, J., and Calvert, R. (2006). Equivalence of the virtual-source method and wave-field deconvolution in seismic interferometry. *Physical Review E*.
- Tsai, V. C. (2011). Understanding the amplitudes of noise correlation measurements. *Journal of Geophysical Research*.
- van Dalen, K. N., Mikesell, T. D., Ruigrok, E. N., and Wapenaar, K. (2014). Retrieving surface waves from ambient seismic noise using seismic interferometry by multidimensional deconvolution. *Journal of Geophysical Research: Solid Earth*.
- van Manen, D.-J. and Robertsson, J. O. A. (2005). Modeling of wave propagation in inhomogeneous media. *Applied Physics Letters*.
- van Tiggelen, B. A. (2003). Green function retrieval and time reversal in a disordered world. *Physical Review Letters*.
- Wapenaar, K. (2003). Synthesis of an inhomogeneous medium from its acoustic transmission response. *Geophysics*.
- Wapenaar, K. (2004). Retrieving the elastodynamic greens function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*.
- Wapenaar, K. and Fokkema, J. (2006). Greens function representations for seismic interferometry. *Geophysics*.
- Wapenaar, K., Fokkema, J., and Snieder, R. (2005). Retrieving the greens function in an open system by cross correlation: A comparison of approaches (I). *Acoustical Society of America*.
- Wapenaar, K., Slob, E., and Snieder, R. (2008a). Seismic and electromagnetic controlled-source interferometry in dissipative media. *Geophysical Prospecting*.
- Wapenaar, K., van der Neut, J., and Ruigrok, E. (2008b). Passive seismic interferometry by multidimensional deconvolution. *Geophysics*.

- Wapenaar, K., van der Neut, J., Ruigrok, E., Deyan Draganov, a. J. H., Slob, E., Thorbecke, J., and Snieder, R. (2011a). Seismic interferometry by crosscorrelation and by multidimensional deconvolution: a systematic comparison. *Geophysical Journal International*, 185:1335-1364.
- Wapenaar, K., van der Neut, J., Ruigrok, E., Draganov, D., Hunziker, J., Slob, E., Thorbecke, J., and Snieder, R. (2011b). Seismic interferometry by crosscorrelation and by multidimensional deconvolution: a systematic comparison. *Geophysical Journal International*.
- Wapenaar, K., van der Neut, J., and Thorbecke, J. (2012). On the relation between seismic interferometry and the simultaneous source method. *Geophysical Prospecting*.
- Weaver, R. L. and Lobkis, O. I. (2001). Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies. *Physical Review Letters*.
- Weaver, R. L. and Lobkis, O. I. (2005). Fluctuations in diffuse field correlations and the emergence of the greens function in open systems. *Journal of the Acoustical Society of America*.
- Weemstra, C., Wapenaar, K., and van Dalen, K. N. (2015). Reflecting boundary conditions for interferometry by multidimensional deconvolution. *SEG Technical Program Extended Abstract*, 34:.

