12-2016

Investigation of Ambient Seismic Noise using seismic interferometry in the Midwestern United States

Joshua Dakota Boschelli
Purdue University

**PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By ___Joshua Dakota Boschelli___

Entitled
INVESTIGATION OF AMBIENT SEISMIC NOISE USING SEISMIC INTERFEROMETRY IN THE MIDWESTERN UNITED STATES

For the degree of ___Master of Science___

Is approved by the final examining committee:

Robert L. Nowack
_____
Chair
Hersh J. Gilbert
_____

Kenneth D. Ridgway
_____

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): ___Hersh J. Gilbert , Robert L. Nowack___

Approved by: ___Indrajeet Chaubey___   12-2-16

Head of the Departmental Graduate Program   Date

INVESTIGATION OF AMBIENT SEISMIC NOISE USING SEISMIC INTERFEROMETRY IN THE MIDWESTERN UNITED STATES

A Thesis

Submitted to the Faculty

of

Purdue University

by

Joshua Dakota Boschelli

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2016

Purdue University

West Lafayette, Indiana

I dedicate this thesis to my family, to my brother Lucas. I hope that this achievement will show what can happen when you apply yourself to something that captivates you.

iii

ACKNOWLEDGEMENTS

First I would like to thank my advisor Hersh Gilbert for not only his instruction and guidance in the field of geophysics, grueling field work, and with this project, but also for his patience and encouragement even when we failed to find a donut on its national namesake day. Additionally, I would like to thank my committee members, Ken Ridgway and Robert Nowack for their guidance over the years.

The completion of this project was also made possible by the support of my family who not only listened to my rants, but would humor me by throwing out ideas whenever I hit a roadblock. To my mother Mary, I thank her for unending support and boundless enthusiasm. To my stepfather James, for his wisdom, and drive which I strive to follow. To my father Mark, for his support from afar and my brother Lucas who continues to push me to achieve all that I can. Last but not least I want to thank my lab mates Chen, Dave, Haylee, Steeve, Roby, Leslie, Alex, and Alivia, who provided not only endless distractions but made 3261 a melting pot of ideas.

Finally, I would also like to acknowledge the input and support from all the individuals involved with the OIINK experiment. From the expertise of Gary Pavlis, Michael Hamburger and Terry Stigall to the numerous landowners that allowed me access to their land. Without your support none of this would have been possible.

ACKNOWLEDGEMENTS

First I would like to thank my advisor Hersh Gilbert for not only his instruction and guidance in the field of geophysics, grueling field work, and with this project, but also for his patience and encouragement even when we failed to find a donut on its national namesake day. Additionally, I would like to thank my committee members, Ken Ridgway and Robert Nowack for their guidance over the years.

The completion of this project was also made possible by the support of my family who not only listened to my rants, but would humor me by throwing out ideas whenever I hit a roadblock. To my mother Mary, I thank her for unending support and boundless enthusiasm. To my stepfather James, for his wisdom, and drive which I strive to follow. To my father Mark, for his support from afar and my brother Lucas who continues to push me to achieve all that I can. Last but not least I want to thank my lab mates Chen, Dave, Haylee, Steeve, Roby, Leslie, Alex, and Alivia, who provided not only endless distractions but made 3261 a melting pot of ideas.

Finally, I would also like to acknowledge the input and support from all the individuals involved with the OIINK experiment. From the expertise of Gary Pavlis, Michael Hamburger and Terry Stigall to the numerous landowners that allowed me access to their land. Without your support none of this would have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Boschelli, Joshua D., M.S., Purdue University, December 2016. Investigation of Ambient Seismic Interferometry in The Midwestern United States. Major Professor: Hersh Gilbert.

The portion of the North American craton occupied by the central United States is a cratonic platform, where a veneer of Phanerozoic sedimentary strata buries the Precambrian basement up to 7 km. Due to the sediment cover and low topographic relief, the at-depth structure of the region remains poorly understood. This study is focused on the crustal structure surrounding the Illinois Basin. This region is of interest because over the past half-billion years tectonic forces have resulted in the formation of epeirogenic provinces in a stable cratonic interior.

Using the OIINK flexible seismic array and the Earthscope Transportable Array, Ambient Seismic Noise Tomography was applied to investigate the crustal structure and produce high-resolution structural models of the region. For our analysis, we used the vertical component of seismograms recorded between January 2011 and December 2014, where spurious events were filtered out to establish the background seismic noise of the region.

Seismic observations based on the cross-correlations of seismic noise from 33,679 station pairs were used to obtain phase velocities at periods from 4 to 40 s. Phase velocity anomaly maps show that the Illinois Basin region is represented by a northwest striking low-velocity element. In contrast, the areas associated with the Ozark Dome and the westernmost portion of Tennessee exhibit high velocity trends. On the other hand, the Reelfoot Rift transitions from a low to high velocity element as period length increases. The contrast between these elements highlights the complex nature of the lithosphere.

The inversion of the phase velocity results produced a 1-D shear wave model, with a relatively fast lower crust overlain by a slow upper mantle. The observed high shear velocities highlight the possibility of mafic materials being emplaced into the lower crust from the nearby Reelfoot Rift. These materials are denser than those normally expected isostatic equilibrium, and could lead to the upper crust deforming. This deformation combined with the regional stress field would be able to deform the crust creating the observable geological features in this region. The most obvious differences between the results of this model and those observed in the Midwestern United States by previous continental scale experiments are due to the high degree of resolution of this study, which greatly improves the understanding of the cratonic geology of the Midwestern United States.

1. INTRODUCTION

Surface wave tomography in seismology has generally been restricted to regions where either earthquake occurs regularly or seismic activity can be easily recorded. A new technique to address these limitations has recently emerged, where the ever-present ambient seismic noise is used to investigate subsurface structures at shallow depths and at a high degree of resolution in both seismic and aseismic regions. Ambient Noise Tomography ("ANT") correlates the background seismic signal between two receivers to extract the response of the Earth's surface. In seismology, the ambient seismic signal was historically thought of as noise that obscured earthquakes and was thoroughly removed. However, it has since been shown that the background signal contains information about the geological structures through which it travels. By correlating the seismic record between two receivers the common characteristics of area are extracted, while the spurious signals are negated. These correlations can be treated as surfaces waves, thus expanding a field by using what was traditionally thought of as unwanted signals (Figure 1.1).
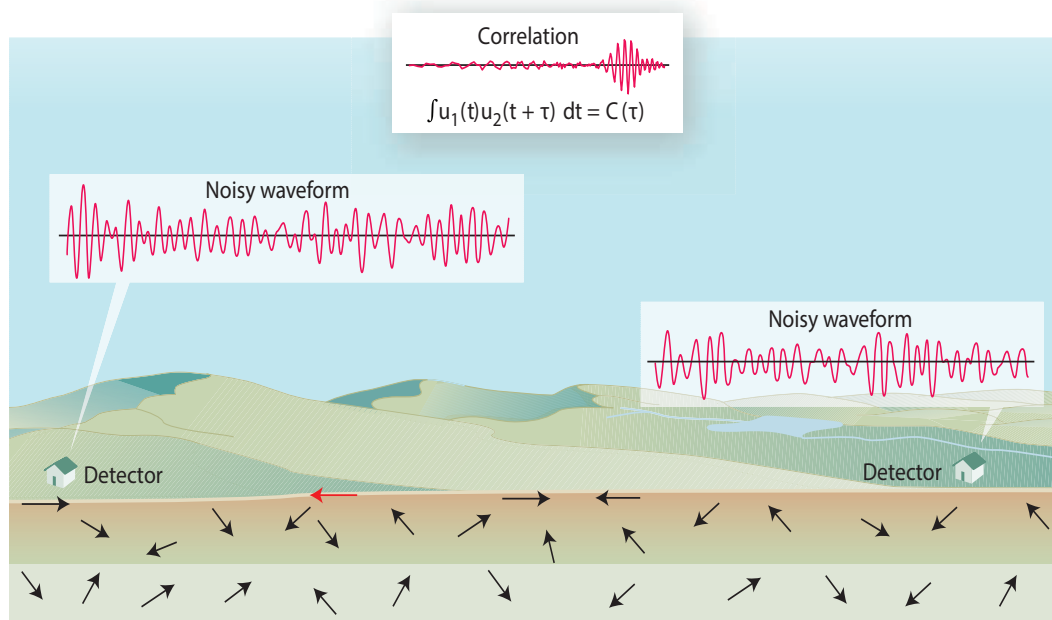
Figure 1.1

From *Weaver* 2005, a cartoon demonstrating how the signal within the Earth's crust can be correlated between two detectors to extract the response of the surface.

The portion of the North American craton occupied by the central United States is a cratonic platform, where a sequence of Phanerozoic sedimentary strata buries the Precambrian basement up to 7 km. Due to this sediment cover and low topographic relief, the at-depth structure of the region is poorly understood. Of the geological structures that are visible within the North American craton, the Illinois basin stands out as one of the few observable cratonic basins. The goal of this study is to understand the effect seismic waves have on the aseismic Midwestern United States, and to gain information about obscured geological structures. This region is of interest because tectonic forces in the past 500 Ma have resulted in the formation of epeirogenic provinces in a stable cratonic interior. Using the Ozarks Illinois Indiana Kentucky ("OIINK") flexible seismic array and the Earthscope Transportable Array, seismic ambient noise tomography was applied to investigate the crustal structure and produce high-resolution models of the region (Figure 1.2). This analysis used the vertical component of seismograms recorded between January 2011 and December 2014, where spurious events were filtered out to obtain the background seismic noise of the region.

The basic assumption of ambient noise tomography is that ambient seismic noise is composed of randomly distributed wave fields that surround a receiver. A perfectly random distribution of the sources of ambient noise would result in symmetric cross-correlations with equal energy arriving from both the positive and negative lags of the two-sided cross-correlation. However, these symmetric waveforms are rarely observed in practice, with a significant asymmetry more commonly observed. The asymmetry is created from stronger or closer ambient noise sources oriented toward or away from one station compared another. To systematically investigate the directions of the incoming ambient noise, the azimuthal distribution of the signal-to-noise ratio ("SNR") for each cross-correlation is examined and mapped. The orientation and magnitude of the SNR identifies the major sources the ambient noise field, with major peaks lining up with the contributors of the ambient noise field. Since the method involves extracting the signal from this ambient noise field it is important to understand its source as the region it occupies is examined.
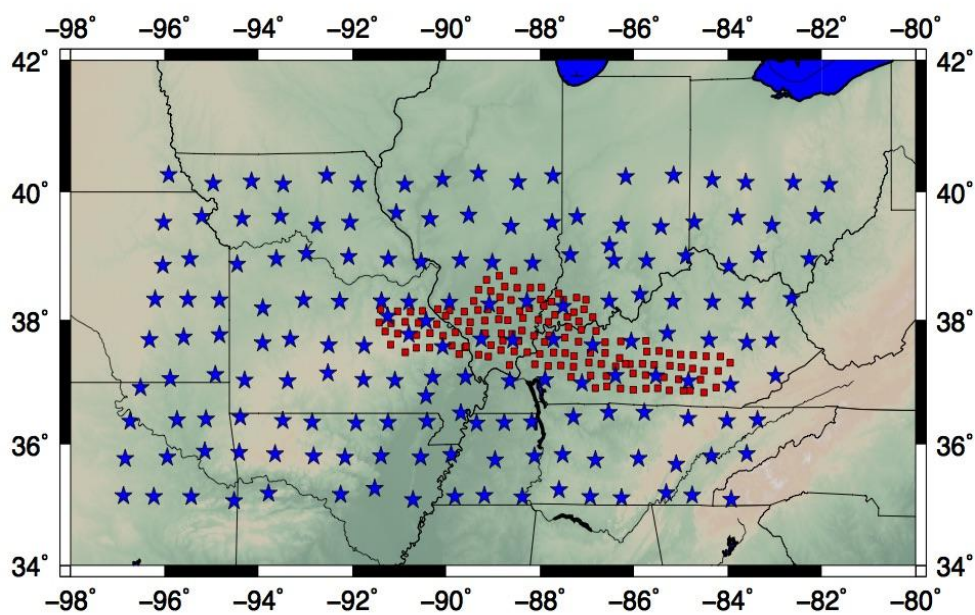
Figure 2.2

Seismic station locations across the Midwestern United States. Blue stars represent the stations of the USArray Transportable Array. Red squares represent the stations of the OIINK Flexible Array.

Sedimentary basins surrounded by denser bedrock are strongly affected by the ground motion of earthquakes because the energy is trapped and amplified within the basin (*Anderson et al.,* 1986). Seismologists have attempted to predict the ground shaking of future large earthquakes to gauge both their geological and ecological effects. Numerous uncertainties exist in such predictions; the major focus of this study concerns uncertainties due to basin amplification. Basin amplification refers to waveforms becoming trapped and amplified as they travel within a sedimentary basin. The amplification increases the vulnerability of the destructive energies of earthquakes. The ability to identify such an effect is critical to reliable seismic hazard analysis. Since the study region is aseismic, using seismic ambient noise allows the control of the disposition of virtual seismic sources, and hence the ability to measure the response in areas of concern.

Cratonic basins are generally poorly understood, since they represent the subsidence of some of the most stable parts of the lithosphere of the Earth. The formation mechanisms needed to alter crust that has remained unchanged since its formation in the Archean. The stability of the lithosphere also interferes with traditional seismic experiments, since only low magnitudes are recorded in their vicinities. Ambient noise tomography provides a means to work around this. With the Illinois Basin being in the Midwestern United States, an area that is heavily populated and provides significant agricultural contributions, the lack of knowledge of its formation is fundamental problem that must be addressed. Previous hypothesizes have varied from the subsidence being caused by the emplacement of dense mafic materials being emplaced in the lower crust to it occurring as a response to lithospheric extension that resulted in crustal thinning. Within this study we aim to provide observations that will aid in supporting an unified theory on what allowed the cratonic lithosphere to deform, creating the Illinois Basin and the adjacent geological features of the Midwestern United States.

2. BACKGROUND

2.1 The Illinois Basin

The focus of this study is the application of ANT to the lithosphere of the Midwestern United States, centering on the intracratonic Illinois Basin. Cratonic basins are locations of prolonged subsidence of the thick continental lithosphere, periodically filled with shallow water and deposition of terrestrial sedimentary rocks. They remain poorly understood because they form as oval shaped crustal depressions on stable and relatively thick continental lithosphere. During the Phanerozoic, subduction-initiated subsidence of the initial proto-Illinois Basin was likely caused by weakening of the lithosphere due to extension on its adjacent continental margins. The history of this subduction can be divided into two main stages: the subduction of Iapetus oceanic crust along the eastern margin during the Paleozoic, and the subduction of Pacific and Farallon oceanic crust along the western margin from the late Paleozoic to the Cenozoic (*Van der Pluijm et al.,* 1990; *McKerrow et al.,* 1991). With the lithosphere weakened, thermal cooling and isostatic adjustments in the crust ultimately caused failed rifts to form in adjacent regions, with the surrounding regions experiencing subsidence periodically through the Paleozoic.

The oval-shaped Illinois Basin is approximately 280,000 km^2 in area, encompassing parts of southern Illinois and Indiana, western Kentucky, Tennessee, and Missouri (Figure 2.1). The most significant seismic feature in the southern part of the basin is the New Madrid Seismic Zone, consisting of the Reelfoot Rift and Rough Creek Graben. Other structures include anticlines, synclines, and monoclines that resulted from regional tectonic   forces   that   uplifted   and   compressed   the   region.
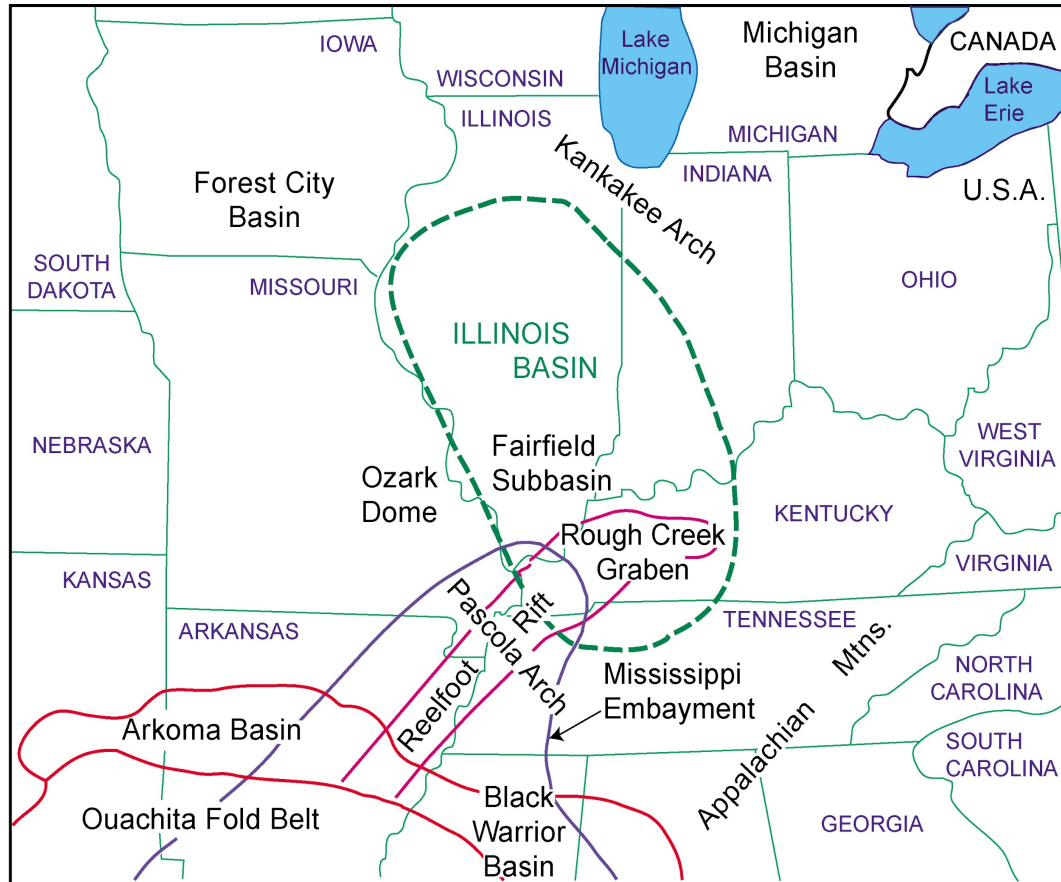
Figure 2.1

Outline of the major geological features of the Midwestern United States modified from *Buschbach and Kolata,* 1991. The green dashed line represents the generalized outline of the Illinois Bain. Other geological features that are important to this study include the Reelfoot Rift, the Rough Creek Graben, and the Ozark Dome.

2.2 The Saint Francois Mountains - Ozark Dome

The Saint Francois Mountains in eastern Missouri are the eroded remnants of a Proterozoic orogenic belt centered on the structural Ozark Dome. The Saint Francois Mountains began to form as an island in the Paleozoic seas surrounded by ancient reef complexes. After subsidence, burial, and induration, this terrane became concentration points for ore-bearing fluids, which later formed into rich lead-zinc veins that have been and continue to be mined today. Uplift related to the Grenville Orogeny and the subsequent withdrawal of the seas exposed the newly formed range during the late Cambrian and early Ordovician periods. The limestone and chert layers that underlay the area were eroded away, but are still observable in the valley walls and bluffs seen in the region today. During the Pennsylvanian, the region was further uplifted as a result of the Ouachita Orogeny as South America collided with North America, exposing the range to further erosion. The midcontinent has undergone extensive erosion, leaving the Ozark dome as an structural remnant.

2.3 The Reelfoot Rift

The Reelfoot Rift developed during the Late Proterozoic to early Paleozoic. As the supercontinent Rodinia began to break up, passive rifting and strike-slip faulting weakened the crust allowing the Midcontinent Rift System to form. The Reelfoot Rift represents one arm of this failed rift system and extends from east-central Arkansas into western Kentucky, abruptly terminating at the Arkansas Transform Fault. Although the Reelfoot Rift may once have extended further southwest, thrusting and sedimentation associated with the Ouachita Orogeny obliterated any surface expression. In the north, the Rift bends eastward merging with the Rough Creek Graben. This area of weakened crust is now known as the New Madrid Seismic Zone, which experiences low-level seismic events uncommon for the otherwise aseismic midcontinent. The bounding faults of the Reelfoot Rift have large normal displacements (*Howe* 1985; *Nelson and Zhang*, 1991; *Parrish and Van Arsdale*, 2004). Historically the Reelfoot reverse fault has the greatest

seismological hazard potential, due to how the surface deformation of the February 1812 earthquake occurred (*Csontos et al.,* 2008). The other faults in the area have displaced Paleozoic strata and thus have been active during the Phanerozoic (*Stark*, 1997).

2.4 The Rough Creek Graben

Located in southeastern Kentucky, the Rough Creek Graben is a northeast striking, 70 km wide feature with linear margins (*Kane et al.,* 1981: *Hildenbrand*, 1985a). The graben itself has an estimated sedimentation depth of roughly 3 km (subsequent igneous intrusions have complicated this estimate). The mechanism of formation of the graben has been a source of debate, with the geophysical basement that flanks it representing different tectonic histories. One hypothesis is that the southeast striking features northwest of the graben represent fracture systems in an old metamorphic terrane. *Hildenbrand,* 1985a used this structural system as a basis to propose that the graben developed along a shear zone separating contrasting basements. *Nelson and Zhang,* 1991 suggested an alternative mechanism, with the observed change in the rift being due to a westward offset in the south-trending Grenville front in central Kentucky. In either case, the presence of the graben indicates that tensional forces and crustal extension have resulted in prominent structural features in the region.

2.5 The New Madrid Seismic Zone

The New Madrid Seismic Zone (NMSZ) spans from southeastern Missouri to western Kentucky, covering a large portion of the Midwestern United States. The NMSZ has been the subject of increasing interest due to its standing as the largest seismic hazard zone east of the Rocky Mountains. This hazard stems from the recognition of a series of magnitude 7 to 8 earthquakes in 1811 to 1812, which caused significant damage to the area surrounding New Madrid, Missouri (*Nuttli*, 2009). While typically aseismic, the additional presence of nearly-constant low magnitude earthquakes has made the NMSZ a focal point for the study of intraplate seismicity.

The active faults in the NMSZ are poorly understood because they are not well expressed at the surface, with any traces being quickly eroded before they can be adequately studied. As previously discussed, the major faults of the area are associated with the Reelfoot Rift, which formed during the breakup of the supercontinent Rodinia during the Neoproterozoic. The resulting rift system failed to split the continent, but has remained as an aulacogen, and likely continues to mechanically weaken the crust. This relative weakness allowed the compressive forces affecting the North American Plate to reactivate once-extinct faults, making the area prone to earthquakes in spite of it being far from the nearest tectonic plate boundary. In addition, heating of the lithosphere may be causing basement rocks to behave more plastically, which concentrates the compressive stresses within the subsurface and further encourages earthquakes to occur (Figure 2.2).
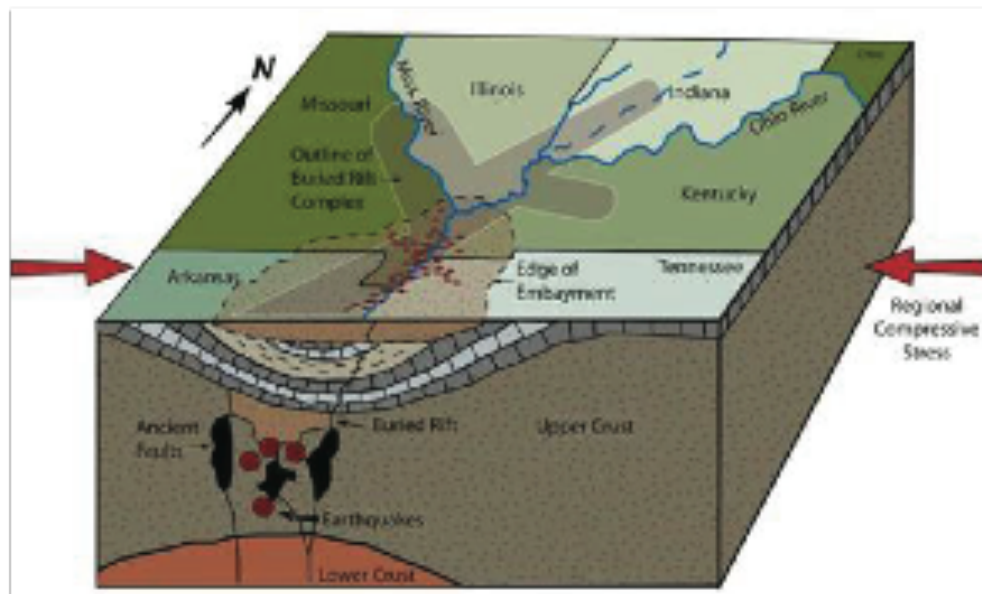
Figure 2.2

Conceptual geological and seismotectonic model of the New Madrid region, with dense mafic bodies being emplaced in the lower crust contributing to the deformation of the region. Modified from *Braile et al.,* 1984.

2.6 USArray Transportable Array

Since 2004 the USArray Transportable Array (TA), a set of 400 seismometers, has been moved gradually in two-year increments eastwards across the North American continent, from the Pacific coast across the Rocky Mountains to the eastern seaboard (Figure 2.3). The TA has yielded data that forms the best picture yet of the North American part of Earth's mantle, reaching hundreds of kilometers beneath the ground surface. The array has revealed many important geological features of North America, such as significant faults in the Pacific Northwest and the hot spot underneath Yellowstone National Park in Wyoming. The transportable array is also well-suited to detect small earthquakes and map the structure of Earth's interior beneath North America, revealing previously unknown geological features. While the OIINK array provides a focused coverage of the study area, the TA provides the spatial coverage needed for this study.

ON THE MARCH
Since 2004, 400 moveable seismometers have measured tremors in different parts of the United States.

UNITED STATES

Last station

First station

ALASKA

- 2004 • 2005 • 2006 • 2007 • 2008 • 2009 • 2010
- 2011 • 2012 • 2013 • Beginning 2014
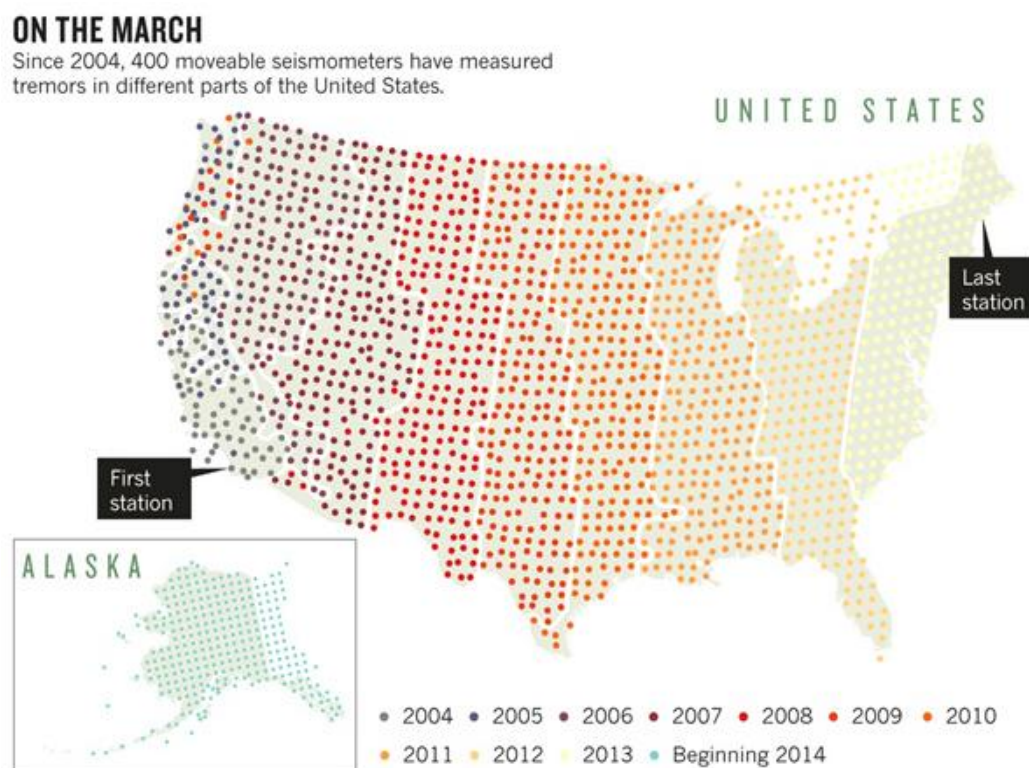
Figure 2.3

The distribution of the USArray Transportable Array as it moved across North America. Each dot color corresponds to the year of each stage of the deployment of the stations.

2.7 The OIINK Flexible Array

The Ozark Illinois Indiana Kentucky (OIINK) Flexible Array was developed to improve our understanding of the regional-scale structures of the North American stable interior. The full array is centered on the depocenter of the Illinois Basin and can be viewed as a 3-fold oversampling of the USArray. Stage 1 of the deployment was from July 2011 to early June 2012 and centered on the Mississippi River Valley region between Missouri and Illinois. Phase 2 of the experiment began in early June 2012 and continued into the fall of 2013. Phase 3 was positioned in a swath from central Missouri to southwestern Indiana, with the removal of the all the stations in 2015. The final coverage of the array depended on the positioning of the stations eastward into southern Indiana and Kentucky over the Rough Creek Graben and across the Grenville Front. Over the four years that the OIINK array was operating, valuable data concerning the cratonic nature of the region were collected. The placement and operation of each station were made possible through the combination of the work of those involved and the hospitality of the landowners in region on whose property the OIINK stations were deployed

2.8 Ambient Noise Tomography

Recent developments in acoustics (e.g. *Weaver and Lobkis*, 2001; *Derode et al.,* 2003) and seismology (*Campillo and Paul*, 2003) have led to a method to measure the elastic response of the Earth by extracting the Green's Function from the diffuse wavefields propagating throughout the surface of the Earth. This low amplitude seismic signal is generated from events like waves striking a shoreline and local human activities such as farming and commuter traffic. Originally this "ambient noise" was filtered out from seismographs because it obscures the records of earthquakes. However, by using the coherent signal within this noise, aseismic areas can now be examined and imaged with the same methods that rely on earthquake signals. To retrieve the signal, long time series are correlated, extracting the response of the surface of the crust between two receivers. The resulting cross-correlation has two effects: the common signal is retrieved and amplified, while the incoherent energy is deconstructed and removed. By making measurements on these signals and using an array of stations for a multitude of ray paths,

a three-dimensional image of the subsurface can be constructed. This method is known as Ambient Noise Tomography ("ANT"), and creates new possibilities in imaging the subsurface at different scales and resolutions that traditional passive source tomography could not previously obtain.

ANT is built upon how a modal representation of a diffuse wave field inside an elastic body can be as expressed as:

$$\theta(x,t) = au(x)e^{i\mu t} \tag{1}$$

where x is position, t is time, u and μ are the eigenfunctions and eigenfrequencies of the Earth, and a is the modal excitation function (*Weaver and Lobkis*, 2004). A diffuse field is constructed from uncorrelated random variables:

$$< a_n a_m > = \gamma_{nm} F(\mu_n) \tag{2}$$

where $F(\mu)$ is the spectral energy density. A large set of these variables is needed to ensure that they are indeed random, thus requiring long time frames. The cross-correlation between the noise fields at locations x and y becomes:

$$C(x, y, \tau) = F(\mu)u(x)u(y)F(\mu)u(x)u(y)e^{-i\mu t} \tag{3}$$

Expression (3) differs only by an amplitude factor F from an actual Green's function between points x and y. Therefore, the Green's function between two locations can be extracted from the diffuse field with a cross-correlation taken over a sufficiently long time, thus forming the fundamental theory of ANT.

Since the beginning of its use in the early 2000's, ANT has been widely applied to both local and regional areas, such as California and the Pacific Northwest (*Moschetti et al.,* 2007) and Tibet (*Yao et al.,* 2006). Earlier studies focused on the microseism

frequency band below 20 seconds to obtain group velocities (*Shapiro et al.,* 2005). *Bensen et al.,* 2005 and 2007 extended the methods to explore longer periods and phase velocity calculations. The detailed procedures of data use for ANT have been summarized by *Bensen et al.,* 2007 and have become a standard for the field (see Figure 3.1). This study focuses on the processing of the ambient seismic noise to extract Rayleigh waves, the surface waves that include longitudinal and transverse motions, and that decrease exponentially in amplitude with increasing depth. Rayleigh wave phase velocity generally increases with period because it is sensitive to shear wave velocities at approximately one-third of its wavelength, while shear wave velocities largely increase with depth.

3. METHODS

3.1 Data Collection

The seismic data used in this study are from the OIINK flexible seismic array and the Transportable Array (TA) of the USArray, both components of the National Science Foundation ("NSF") EarthScope project. Data collection was done via a combination of fieldwork and data requests from the Incorporated Research Institutions for Seismology ("IRIS") consortium. The data consist of daylong records of the vertical component of each seismic station. Because the management and storage of the data was a concern, the daily record for each station was saved in one-hour segments that could be both easily managed and manipulated. A total of 306 stations covered the Midwestern region of the United States from January 2011 to December 2015 (Figure 1.2). Due to the timing and nature of each array, a number of stations were not operating during the same timeframe with the coverage of both arrays moving steadily eastward from 2011 to 2014.

In traditional seismic studies, the quality control of raw seismograms ensures that only reliable earthquake traces are used. For ANT, the data quality control procedure is more involved due to the dependency of background seismic noise on the signal propagating through it, the reliability of the seismograph, and the presence of spurious signals. Each of these factors needs to be considered, as hour-long raw seismograms are processed into time segments that are ready for the cross-correlation that ANT requires. These steps consist of: 1) band-pass filtering the seismograms, 2) removing instrument responses, 3) demeaning, detrending, and time-domain normalization, and 4) spectral whitening the frequency spectrum of the waveforms (Figure 3.1). The impact of these steps has a profound effect on the waveforms and will be examined in the following sections.

Phase 1:

| Raw data | → | Remove instrument response, remove mean, remove trend, band-pass filter, and cut to length of 1-day |

| Apply time domain normalization | → | Apply spectral whitening |

Phase 2:

| Compute cross-correlation | → | Stack day-correlations to desired number of days |

Phase 3:

| Measure group and/or phase velocity |

Phase 4:

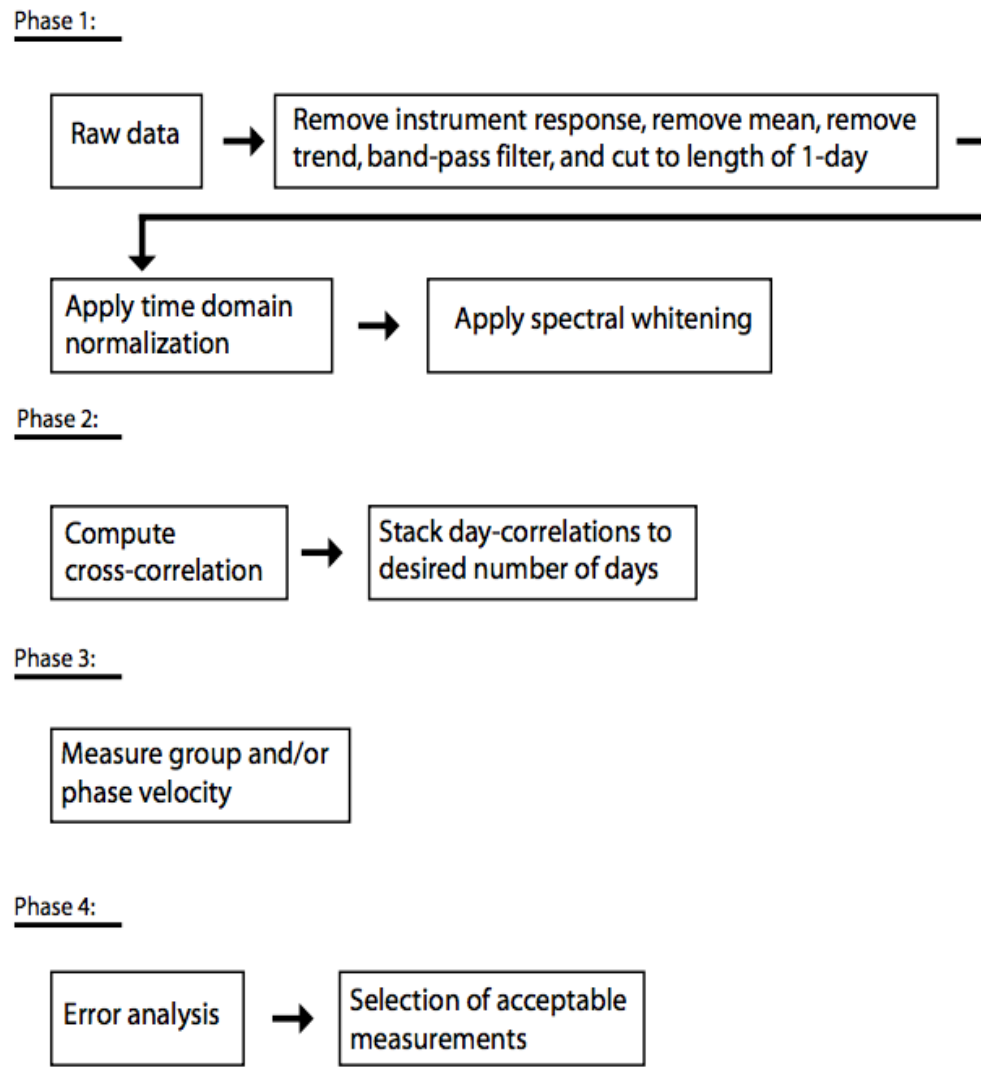| Error analysis | → | Selection of acceptable measurements |

Figure 3.1

Schematic representation of the data processing scheme for Ambient Noise Tomography, from *Bensen et. al,.* 2007. The overall process is broken down into four phases, with the goal of taking seismic noise and producing velocity measurements from it.

3.2 Temporal Normalization

Temporal normalization serves to reduce the effects of earthquakes, instrumental failure, and nonstationary noise sources near the station on the cross-correlations (*Bensen et al.,* 2007). These effects occur irregularly and contribute to spurious signals appearing in the cross-correlations, and can lead to inaccurate and unacceptable results. Of these signals, earthquakes are the most significant impediments to ANT data processing. While large earthquakes can be identified in earthquake catalogs, small earthquakes like those observed in the aseismic Midwestern United States are not recorded in the catalogs and are difficult to track. Other noise sources, such as pressure drops resulting from changing weather conditions, vehicle traffic, machinery, and other anthropogenic signals are similarly difficult to identify. These sources create spurious signals that otherwise obscure surface waves; removal of these signals is a crucial step in ANT.

While there are a number of methods to temporally normalize the data, this study employs the most aggressive method, known as "one-bit" normalization (Figure 3.2), which retains only the sign of the raw signal by replacing all positive and negative amplitudes with either a positive or negative one. This method has been shown to increase the SNR when employed in acoustic experiments in the laboratory *(Larose et al.,* 2004) and has been used in a number of early seismic studies of coda waves (*Yao et al.,* 2009) and ambient noise (*Campillo and Paul,* 2003*; Shapiro and Campillo,* 2004; *Shapiro et al.,* 2005; *Yao et al.,* 2006). With the presence of extensive anthropogenic noise, this aggressive approach neutralizes the spurious signals while maintaining the ambient signal.

3.3 Spectral Whitening

Spectral whitening acts to improve the frequency content of a time domain waveform to allow for both easier and accurate dispersion measurements to be obtained. Seismic ambient noise is not equally distributed across the frequency domain; rather it has peaks near 15 and 7.5 seconds (corresponding to the primary and secondary microseisms), and at 240 seconds to form a signal now referred to as "Earth hum" (*Rhie*

and Romanowiez, 2004). These irregularities impede smooth spectral measurement, leading to unstable results. Spectral whitening broadens the band of the ambient noise signal while also limiting degradation of the signal resulting from persistent monochromatic sources (Figure 3.2). This allows for the range of measurements to be expanded past the peaks, opening the spectrum. With an expanded spectrum, measurements have an enhanced stability as they resemble a Bessel function for distances larger than approximately one fourth a wavelength, which is traditionally observed in measured surface waves (*Weemstra,* 2013). Spectral whitening acts to improve the frequency content of a time domain waveform to allow for both easier and accurate dispersion measurements to be obtained.
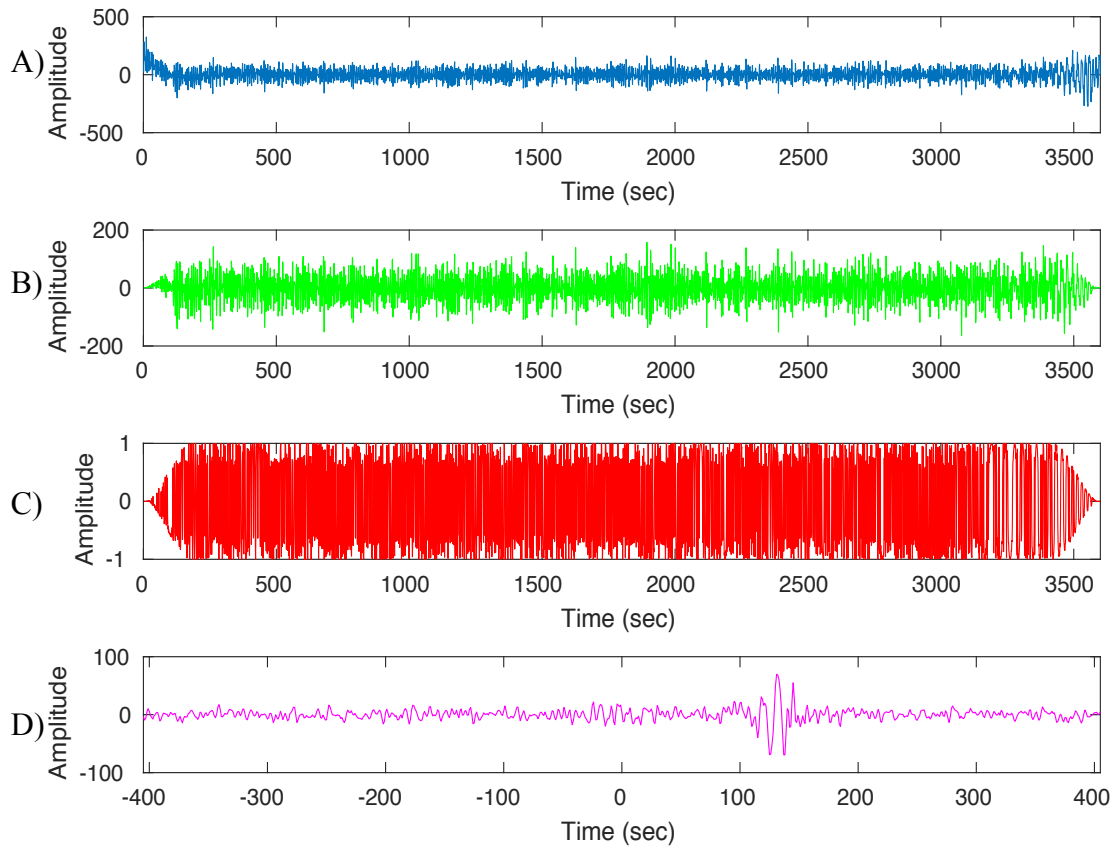
Figure 3.2

Example of the processing steps needed to go from a raw trace to a measurable cross-correlation. A) Raw one-hour seismogram from TA station Q51A. B) Same waveform from (A) with the waveform bandpass filtered and spectrally whitened. C) The waveform from (B) One-bit time domain normalized, the final step before cross-correlation. D) Two-sided time function that results from the correlation of a station pair, with both positive and negative lags present.

3.4 Cross-Correlation

After the preprocessing of the daily time series, the next step in the data processing scheme is to cross-correlate the waveforms. Because the interstation distances vary, cross-correlations were performed between all possible station pairs followed by a strict quality control step at a later time. This step produces a total number of station pairs following the relation:

$$N = \frac{n(n-1)}{2} \tag{4}$$

where N is the total number of station pairs and n is the number of stations. From this relation, we can expect to obtain tens of thousands of cross-correlations with 306 stations producing 33,679 station pairs that were operating at the same time. The data were then cross-correlated in incremented steps and the results stacked to obtain a two-sided time function (Figure 3.2). The length of the correlation window corresponds directly to the interstation distance; to accommodate those in this study 20-minute increments were used.

The two-sided time functions are characterized by having a positive and negative time lag, with the lags representing a waveform emanating and being recorded at each station in the pair. The positive lag portion of the cross-correlation is sometimes called the "causal" signal and the negative lag part of the "acausal" signal (Figure 3.3). If sources of the ambient wavefield were distributed homogeneously in azimuth, the causal and acausal signals would be symmetrical. However, in this study the waveforms exhibit a strong asymmetrical nature, the significance of which will be examined below.
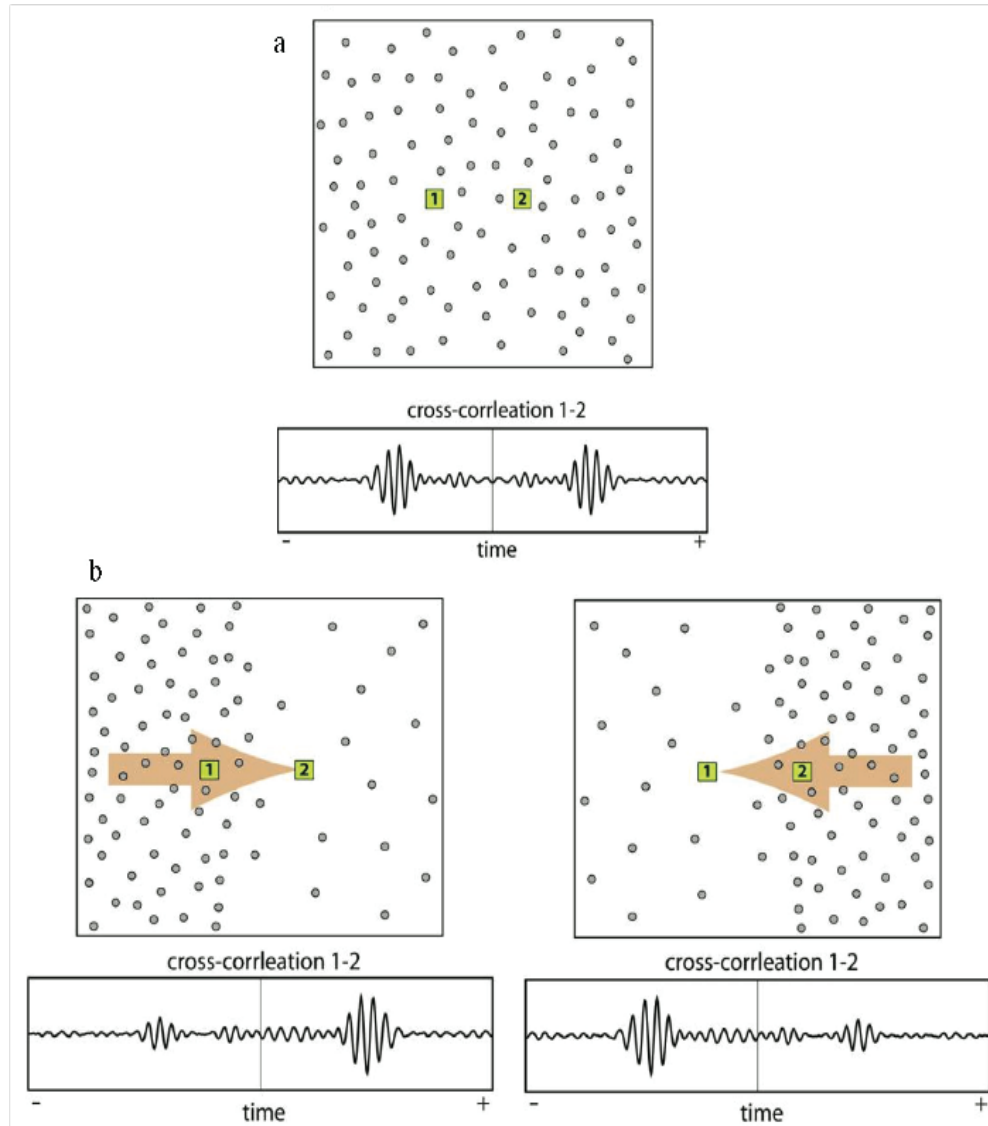
Figure 3.3

The effect of the distribution of the seismic noise has on the symmetrical nature of the two-sided time function that result from the cross-correlation of time series from two stations. From *Stehly et al., 2006.*

3.5 Waveform Energy Filtering and Stacking

Because the signal content of a waveform can change over time due to spurious factors, each daily waveform can be expected to vary. This variation must be accounted for in order to produce reliable waveforms for further measurement. Stacking the daily waveforms usually accomplishes this, where the common characteristics are enhanced and incoherent signals are diminished. The first step is to stack all the available data for each station pair to construct a baseline waveform. Following from this a secondary stacking procedure is conducted that involves computing the cross-correlation coefficient between each daily waveform and the results from the primary stacking procedure. This procedure is performed to separate the daily waveform that exhibits useful signal from that which is obscured by noise.

A range of coefficients to define days of useful data was initially difficult to establish due to the presence of spurious signals; values from 0.5 to 1 were eventually determined to be reasonable. Regardless, this eliminated noisy days being included in the stack and polluting the signal, as well as removed days that are identical to each other (Figure 3.4). Expanding the value range would allow more days to be used, but it would also add days with spurious signals, thus defeating the purpose of this two-step procedure. The impact of this procedure varies between each station pair, with some being affected more than others. This processing results in a waveform that exhibits the clearest response of the Earth's surface between two stations. Stacking the daily waveforms over increasingly long time-series improves the observed signal and thus the SNR ratio of the waveform itself.
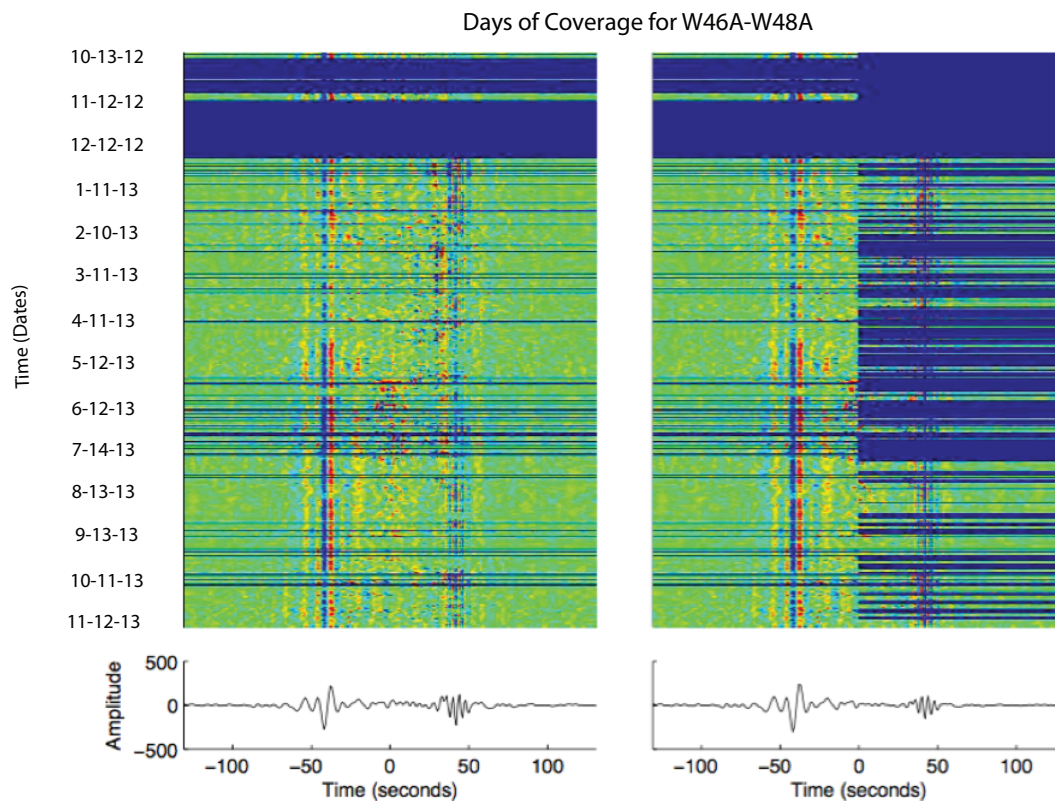
Figure 3.4

The daily cross-correlated waveforms of the station pair W46A and W48A over their shared operational period and the effect filtering the waveforms has on the them. The dark blue sections are times where their data is not available for the station pair. In the waveforms, the blue indicates relative lows, while Red indicates relative highs on each trace. The left figure represents the stack of all available data, and the right represents the stack where individual days without a correlation coefficient less than 0.5 have been removed. Below each stack are the waveforms that result from stacking each collection of data.

3.6 Signal-to-Noise Measurement

In the process of producing the clearest signal possible, the SNR was computed for each station pair. The first step was to separate the signal from the noise window for the cross-correlated waveforms. The signal window with bounds defined by the time it would take surface waves to travel at both 2.5 and 4.5 km/sec was identified, with the window depending on the distance between each station pair. The noise window follows the signal window and is equal in this study. While this served to measure a subset of the station pairs accurately, the methodology of *Bensen et al.,* 2007 was adopted instead, where the noise window follows the initial signal window. For the calculations of the SNR, the ratio between the root mean square of the signal and noise (Figure 3.5) is used. Because the correlated waveforms result in a two-sided time function, the SNR for both the causal and acausal lags was computed and the stronger of the two ratios was used to represent the SNR of the waveform.

The difference between the two SNR's further exhibits the asymmetric nature of the correlated waveforms in this region. To further constrain which waveforms should be used for measurement, a cutoff for the SNR was established. After trial and error, it was found that a waveform with an SNR less than 4 did not have adequate signal strength to provide a reliable phase velocity measurement. It is paramount to have a waveform with an adequate signal strength, since the entire waveform is utilized during the measurement process. With an established benchmark for the appropriate signal strength, the computed cross-correlated waveforms can be measured for their associated phase velocities while ensuring the reliability of the measurements. For this study, two different methods were used to measure for phase velocity, each providing a useful set of measurements to examine the Midwestern United States.
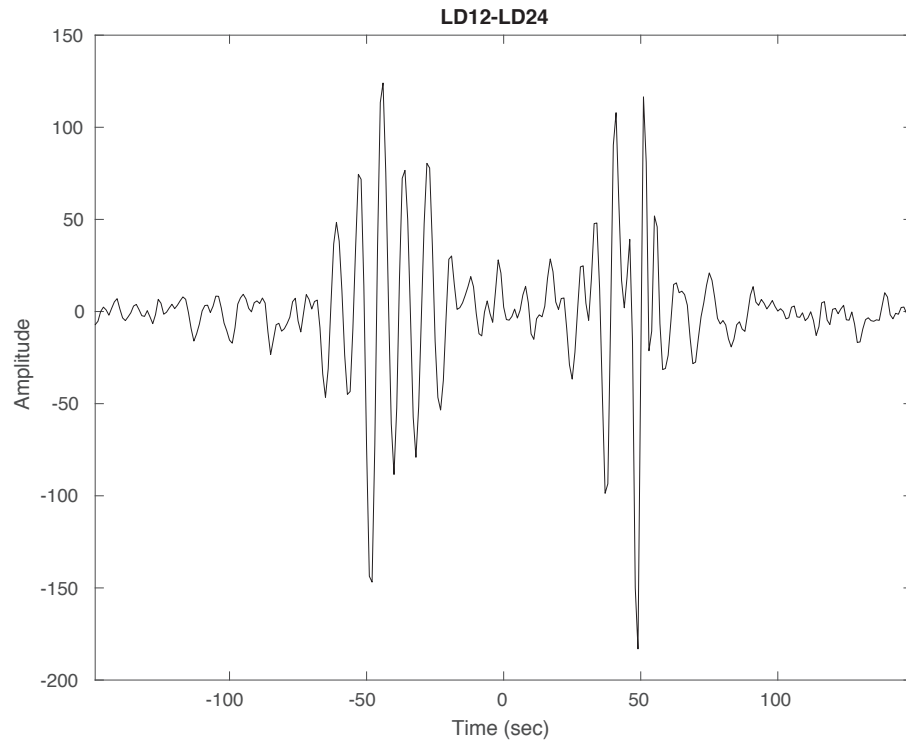
Figure 3.5

An example of a cross-correlated waveform from which the signal-to-noise ratio is measured from. To calculate the SNR, the root mean square of the signal window is divided by the root mean square of the corresponding noise window. The bounds of the signal windows are based on a wave traveling between 2.5 and 4.5 km/sec for each station pair's distance from one another. Because phase velocities are measured over a wide frequency band, the SNR at each corresponding period is also measured.
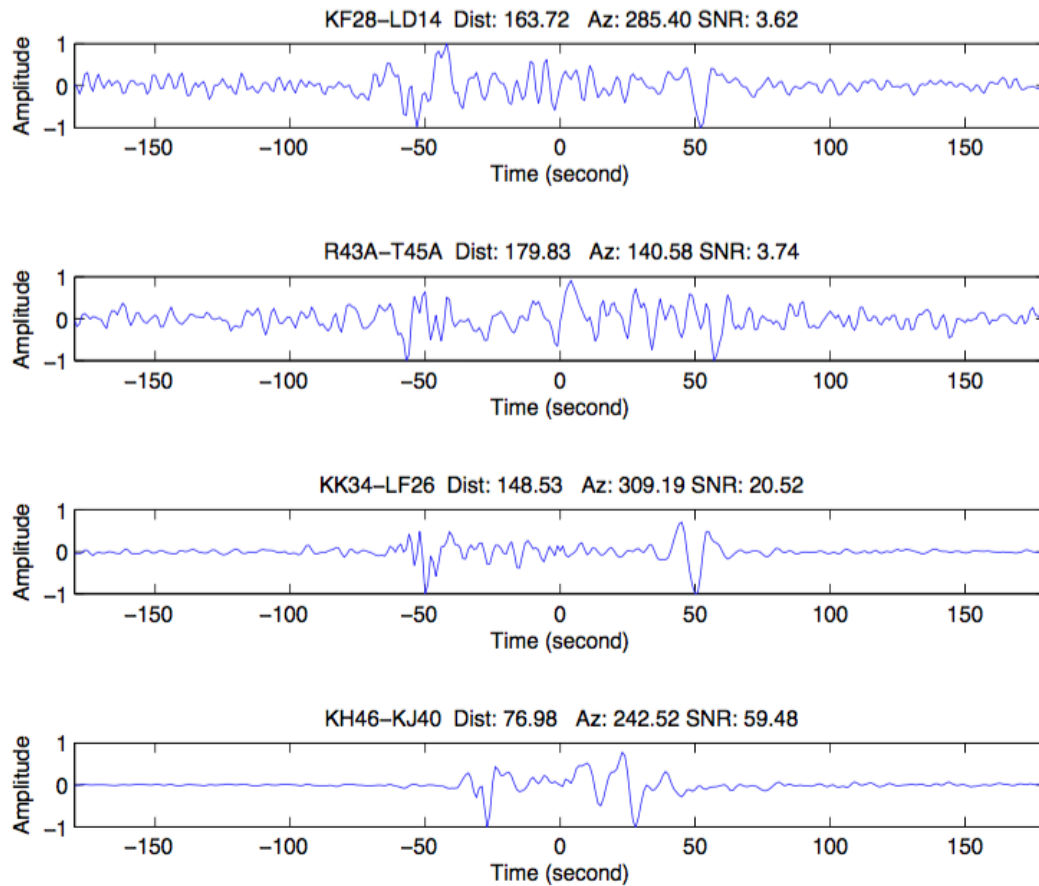
Figure 3.6

Four examples of waveforms to demonstrate the importance of having an SNR cutoff of 4. The top two waveforms have a low SNR, below 4, making it difficult to differentiate between the signal and noise windows, while the signal of the bottom two waveforms are clearly present with SNR's greater than 4.

3.7 Spectral Method

The first of the two methods used in this study for measuring phase velocities follows the work of *Aki* 1957 as illustrated by *Ekström et al.,* 2009:

$$\rho(r, \omega) = J(r \frac{\omega}{c(\omega)}) \tag{5}$$

which states that the frequency domain for a station pair separated by a distance *r* and frequency *ω* varies as *J*, the Bessel function of the first kind, where *c(ω)* is the phase velocity at frequency *ω*. The real parts of the frequency spectra resemble a Bessel function, which are the solutions to the Bessel differential equation but are nonsingular at the origin. This causes them to have peaks that do not decrease monotonically with frequency as expected. Because the amplitudes of the real part of the spectrum depend on both the background noise spectrum and nonlinear effects of the data processing, dispersion information cannot readily be deciphered from the spectrum. The locations of the zero crossings, where the spectrum crosses the frequency axis, are independent of the variations in the spectral power of the background noise, and therefore provide a reference point for making dispersion observations. Using this information, equation (5) can be modified into:

$$c(\omega) = \frac{\omega r}{z} \tag{6}$$

where *ω* denotes the frequency of the observed zero crossing and *z* denotes the zero crossings of the Bessel function. Figure 3.7 displays the locations of the zero crossings of the real spectra as points, as a function of frequency and how they can be used to construct a dispersion curve. Using these zero crossing points, a reliable dispersion curve can be generated from the full trace of the 33,697 station pair waveforms in the study over periods ranging between 4 and 40 seconds. Connecting the positive-to-negative zero crossings and the negative-to-positive zero crossings generates two separate dispersion curves over the frequency range of interest. The differences between each of these curves create bounds that can act as a quality criterion for selecting reliable measurements.
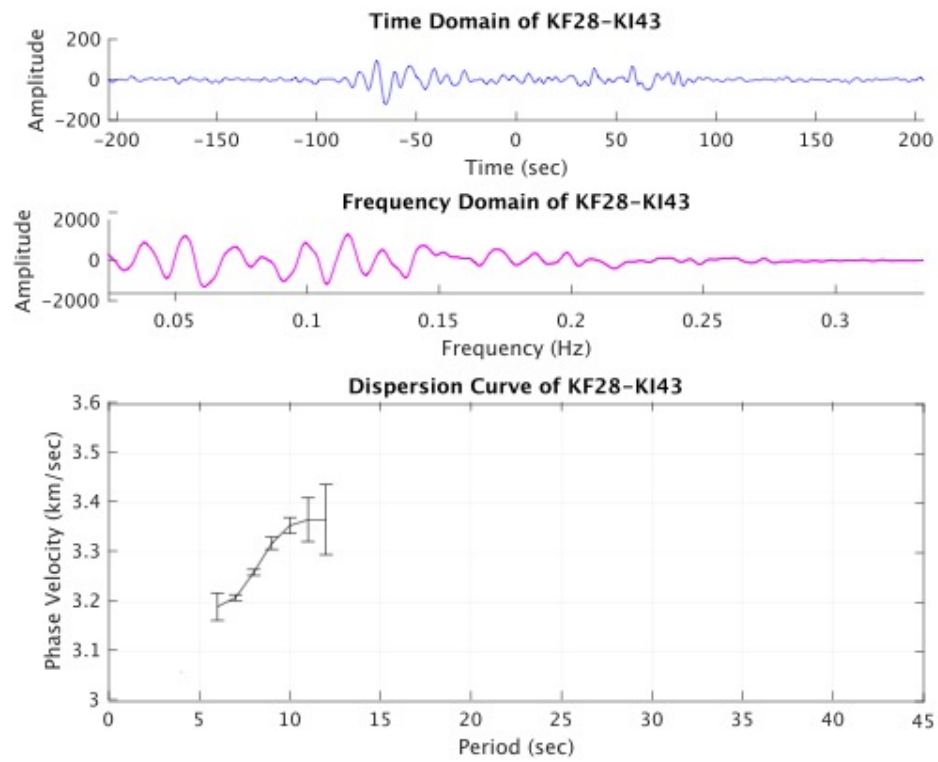
Figure 3.7

An example of the Spectral method measuring a station pair KF28-KI43's phase velocity. The top figure represents the waveform in the time domain. The middle figure represents the transformed waveform in the frequency domain. The bottom figure represents the measured phase velocity dispersion curve.

3.8 Frequency Time Analysis Method

The second measurement method focuses on using the traditional frequency time analysis (FTAN). The FTAN method used here is based on the work of *Levshin et al., 1999*, which obtains a measurement on a single waveform and involves significant analytical observation. By taking the Fourier transform of a waveform and the corresponding inverse Fourier transformation, it is possible to construct a smooth 2-D envelope frequency-time function. By analyzing the envelope function, a group velocity curve can be measured.

It is important to note that phase velocities cannot be derived directly from group velocities, but group velocities can be computed from phase velocities. To get the desired measurements let $U = \frac{\delta\omega}{\delta k}$ and $c = \frac{\omega}{k}$ be group and phase velocity, respectively, $s_u = U^{-1}$ and $s_c = c^{-1}$ are likewise the group and phase slowness, with representing k the wavenumber. Then using the relation:

$$s_u = \frac{\delta k}{\delta\omega} - \frac{\delta(\omega s_c)}{\delta\omega} \tag{7}$$

We obtain a first order differential equation relating the group and phase slowness at frequency ω:

$$\frac{\delta s_c}{\delta\omega} + \omega^{-1}s_c = \omega^{-1}s_u \tag{8}$$

Therefore if the phase velocity curve *c(ω)* is known, the group velocity curve *U(ω)* can be found directly from the above equation. Likewise if the group velocity curve is known, this differential equation (7) can be solved to find c(ω), which results in:

$$s_c = s_u + (\omega r)^{-1}(\varphi(t_u) + 2\pi N - \frac{\pi}{4}) \tag{9}$$

where $N = 0 = 0, \pm 1, \pm 2 \ldots$ and $\varphi(t_u)$ is the observed phase velocity at the observed group arrival time, $t_u = \frac{r}{U}$, and a phase shift is applied to account for possible dependences on the frequency and geographical location.
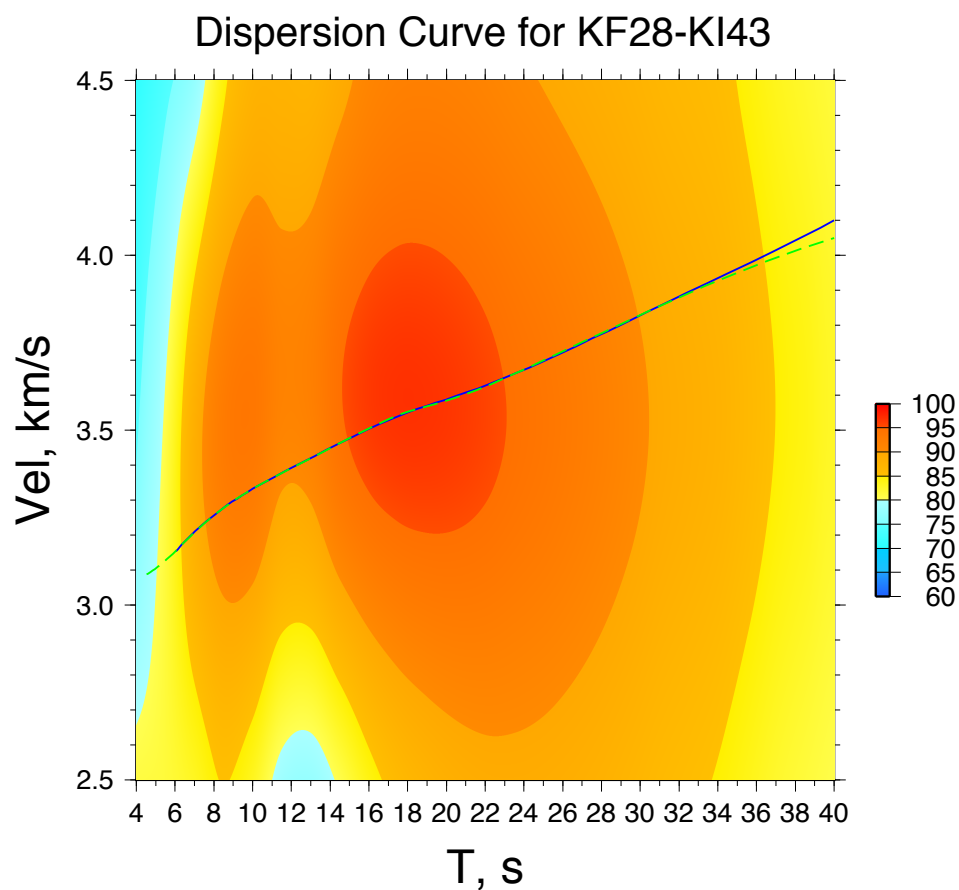
Figure 3.8

Phase velocity dispersion curve measure though the FTAN method for the station pair of KF28 and KI43. The color map corresponds to the confidence interval associated with the measurements. The measurements are made by filtering and measuring the cross-correlated waveforms. The dashed green line represents the raw measurements, and solid blue line the measured results smoothened over the period bands.
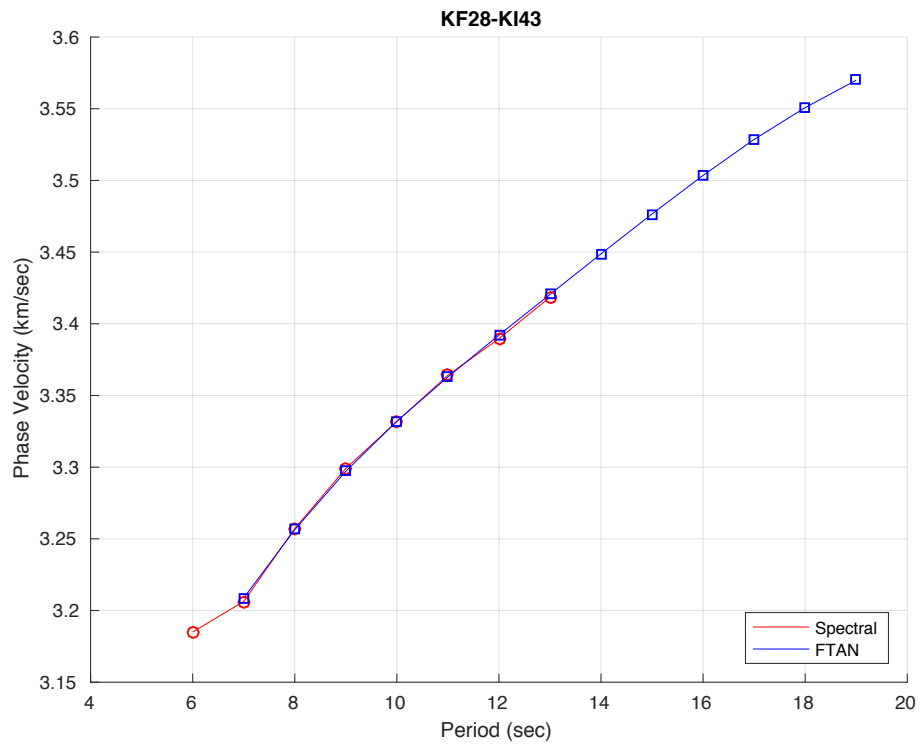
Figure 3.9

Phase velocity dispersion curves measured though the FTAN and Spectral method plotted on top of each other for the station pair of KF28 and KI43. Only the measurements that overlap are used for interpretation.

3.9 Ray Spatial and Method Comparison

Following the measurement process thousands of dispersion curves are obtained to ensure that only reliable measurements are used and whether additional quality control procedures are needed. Even though certain station pairs passed every error check in the measurement process as established by *Bensen et al.,* 2007, they still produce unrealistic dispersion curves that would introduce spurious features into the results. While the obvious step to address this would be to inspect every dispersion curve for error, the sheer number of dispersion curves (over 10,000) makes this step impracticable. Instead the phase velocities at each station were compared across all periods. The measurements that exceed the standard deviation at each station were removed, ensuring that similar velocities are being measured across the array. Figure 3.10 shows the effect of the grid comparison and how it eliminates the variability present at each station. Since each measurement effects the overall image of the region it covers, it is vitally important to only use those that will not introduce spurious results.

Because the phase velocities were measured using both the spectral and FTAN methods, two sets of measurements were obtained. Instead of favoring one set over another, a third set of measurements based on velocities agreed upon by both the spectral and FTAN sets was constructed. The FTAN measurements provide a range for each measurement, while the spectral method provides detailed dispersion of measurements. Figure 3.11 shows the benefit of using both sets of measurements. The notion of having a set of measurements built off two methods ensures the most accurate results possible. A range of values was tested allowing the measurements to differ, settling on a difference of 0.0125 km/sec as a cutoff. This value allows the measurements to be similar, while not forcing them to be identical. If the value was any lower, no benefit would be realized from this unnecessary working of the data. Following from these processing steps, the daily records of 306 stations were processed to obtain numerous and reliable phase velocities to examine geological structures of the Midwestern United States.

4 Seconds
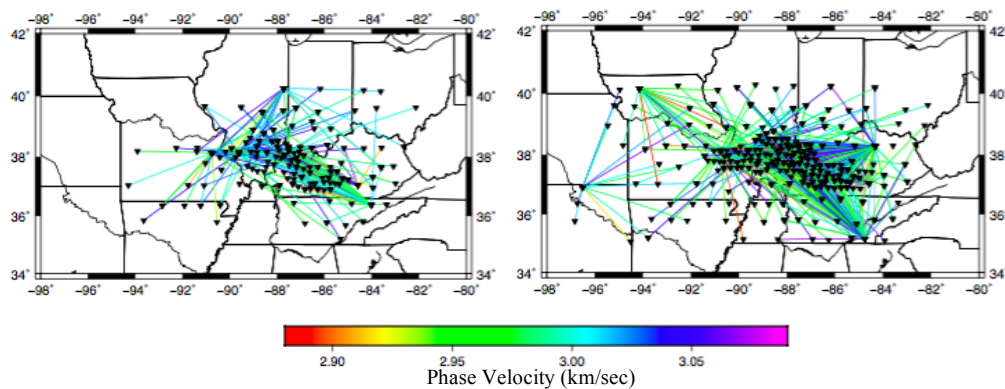
10 Seconds
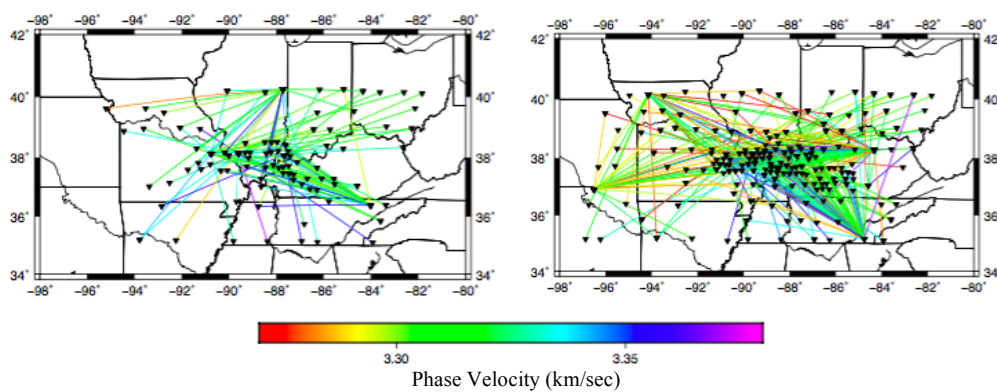
Figure 3.10

Example of the effect that the ray comparison has on the phase velocity measurements. Figures on the right represent all available phase velocity measurements, with figures on the left representing the treated measurements. Each ray is color coded based on the associated color bar.
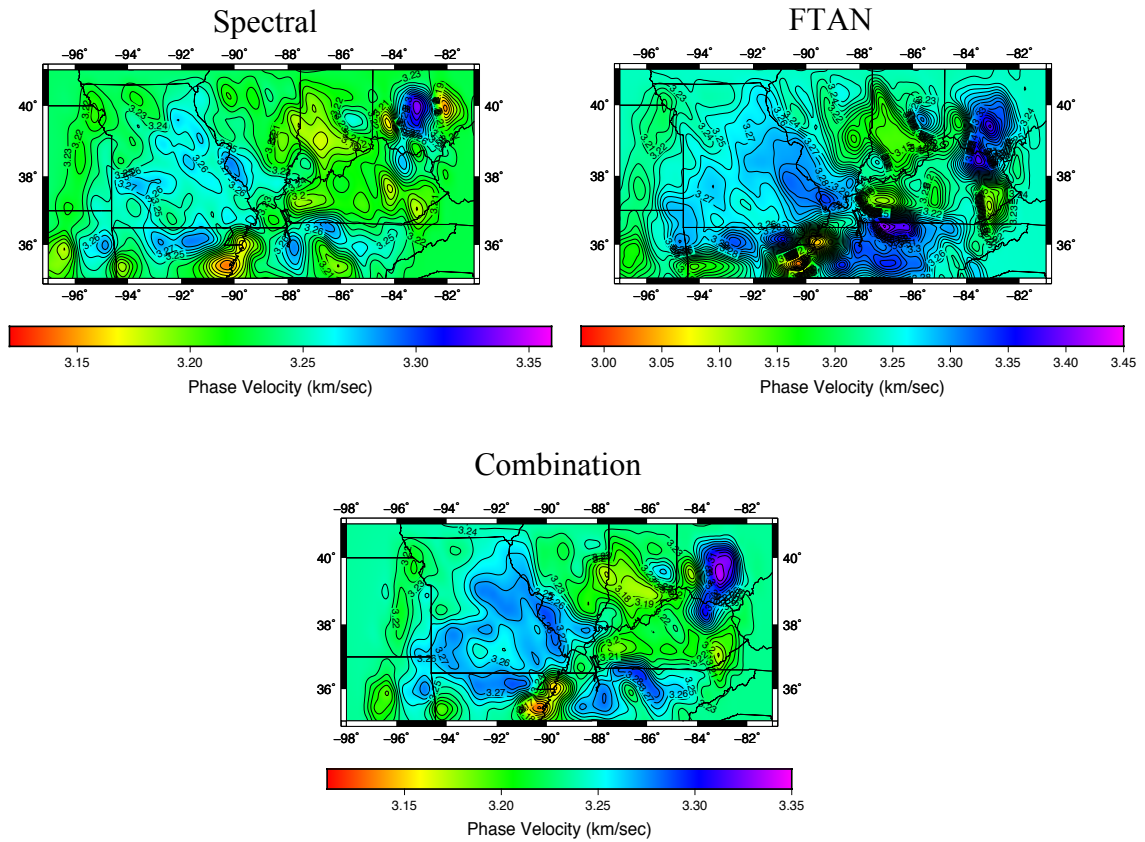
Figure 3.11

Example of the effect that the method comparison has on the phase velocity measurements. Each figure represents a 2-D phase velocity map at 8 seconds from each method used in this study. The maps are color coded based on the associated color bar.

4. RESULTS AND DISCUSSION

Despite variations within the ambient seismic noise field, relative trends in both the signal and its contents remain relatively consistent. These trends are dependent on the depths sampled, and so provide crucial information on the crustal structure of the region at specific depth intervals. General trends and specific features will be discussed in terms of both their relative depth and geographic distribution. Additional results, such as animations of surface wave propagating across the array can be found in the appendix section of this study. While there are few previous studies to which the results may be compared, the data themselves provide a large enough pool of measurements to conduct an extensive error analysis.

4.1 Phase Dispersion Curves

After completing the processing steps as described above, phase velocity dispersion curves from 21,506 station pairs were extracted from the possible 33,679 within the study area. While previous ANT studies focused on using frequencies less than 0.1 Hz, the coverage of the combination of the OIINK and TA arrays extended the range for this study to 0.25 Hz. Frequencies at this range are sensitive to the variations of earth materials from the upper crust to the upper mantle. The average measured phase velocities range from ~2.9 km/s at 4 s to ~4 km/s at 40 s. These measurements are faster than those observed by regional studies but agree with those measured using earthquake surface waves recorded by the OIINK array (*Chen et., 2016)*. The spread of the average phase velocities along with the distribution over the period bands are shown in Figure 4.1. With over 21,506 unique dispersion curves, this study provides adequate coverage to reveal the crustal structure in a large part of the Midwestern United States.
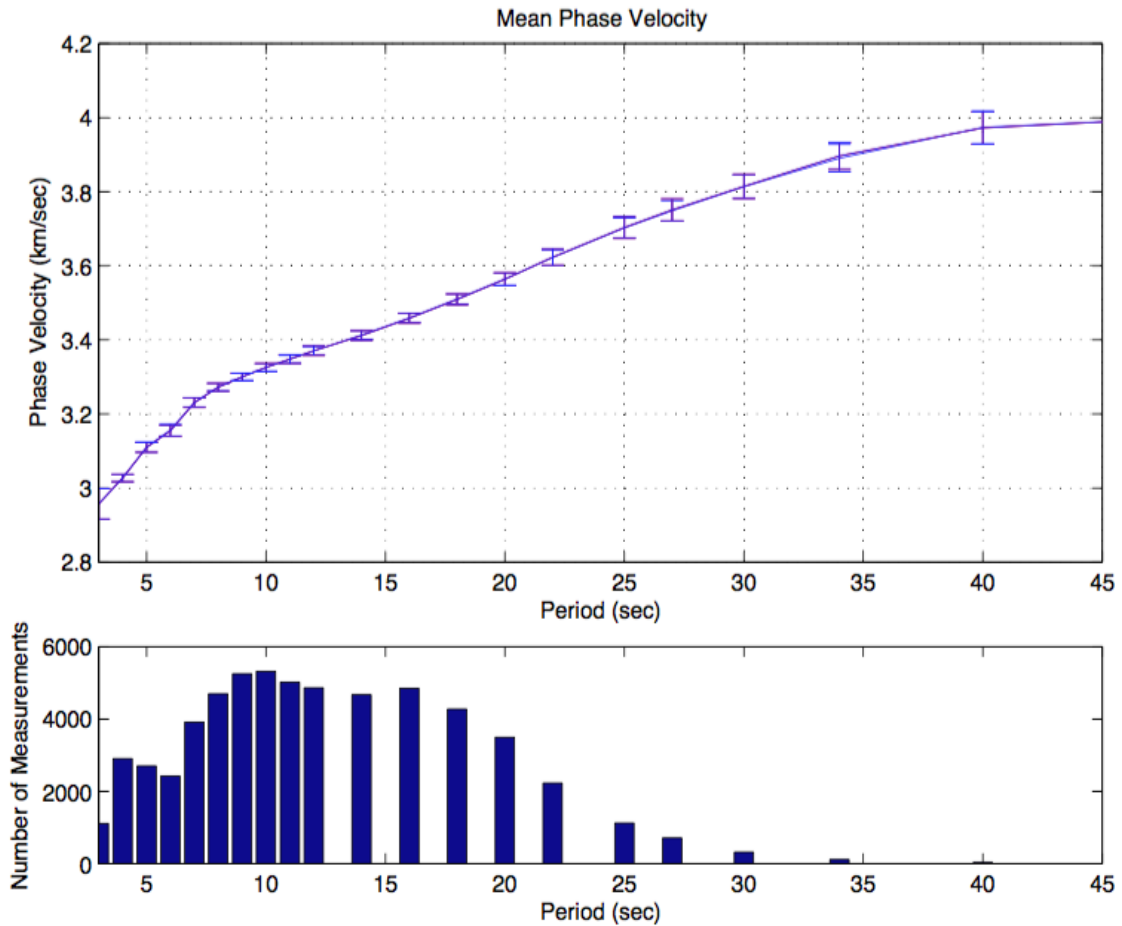
Figure 4.1

The mean phase velocity dispersion curve that results from the processing four years of OIINK and TA data. The measurements span from 2.9 km/sec at 4 s to 3.98 km/sec at 40 s. The lower figure demonstrates the spread of the measurements, with the greatest number of measurements ranging from 10 to 20 seconds.

4.2 Comparison Between Spectral and FTAN Measurements

Boschi et al., 2013 found consistency between phase velocity measurements made from cross-correlated waveforms using both the Spectral and FTAN methods. The velocities measured using the FTAN are reliable only for station pairs that are separated by at least three wavelengths. The Spectral method, on the other hand, produces measurements on station pairs separated by as little as one wavelength and remains consistent with pairs that are separated by greater distances. To compare the results of the methods, histograms were constructed at each observed period. Measurements were further divided into subsets based on their separations r relative to surface wave wavelength λ; (1) r > 3λ, (2) 3λ < r < 2λ and (3) 2λ < r < λ. These bounds push the respective limitations of either method and provide a base that is suitable for statistical comparison. The observed Gaussian distributions in this study are similar to that of *Boschi et al.* 2013, where the mean of differences are close to 0 m/s. This distribution extends into higher frequencies where it has been previously unobserved.

The means of the differences are close to 0 m/s for the cross-correlated waveforms in all periods and at all three distance ranges, the largest differences at ~1.2 m/s for paths with r > 3λ. The standard deviations are about 8 m/s for paths with r > 3 λ and decrease to ~4 m/s for the other distance ranges. A systematic trend in standard deviation is not observed across any of the periods. These findings are consistent with those of *Luo et al.,* 2014; phase velocities measured by both methods indicate that both the FTAN and the Spectral methods provide dependable dispersion measurements from cross-correlated waveforms with a separation greater than one wavelength.
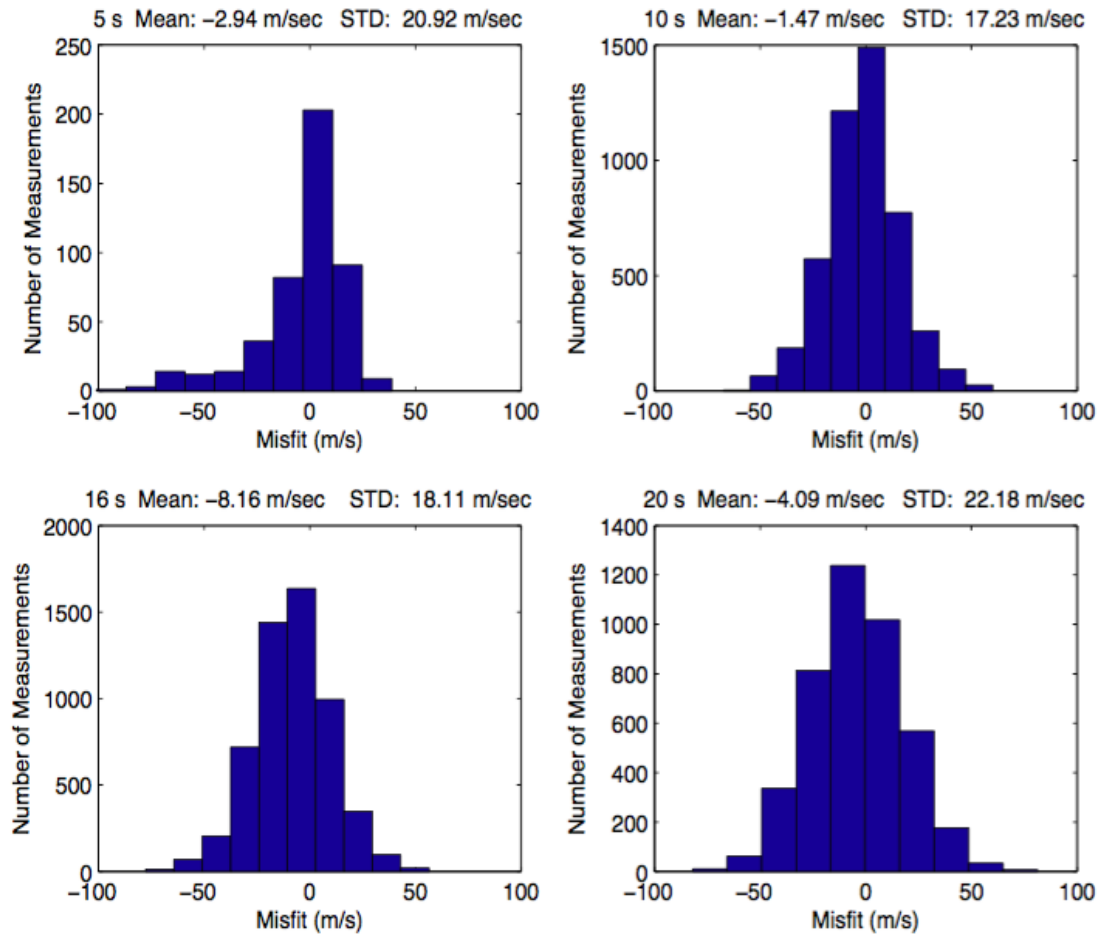
Figure 4.2

Histograms showing the difference between the Spectral and FTAN phase velocity measurements at four period lengths. The mean of differences is consistently close to 0 km/sec, with the standard deviation remaining consistent across all periods.

Having established the consistency of the two methods, only measurements that were generated by both methodologies were used. The largest differences span from 4 s to 10 s, which represents not only the sedimentary cover of the craton and upper crust, but also the outer edge of FTAN's effective range. From the mid to lower crust (10 s to 25 s) the difference becomes minimal. In the lower crust and upper mantle, from 25 s to 40 s, the differences increase due to the limitations of the Spectral method. Using a combination of both measurement methods, a stable set of measurements can be established upon which further observations can be made.

4.3 Temporal Variation

The study area experiences extreme seasonal variability in temperature, with additional variability being introduced from the impact of severe weather such as tornados and strong fronts. The daily cross-correlated waveforms of the station pairs were originally examined to identify days of noisy data (Figure 3.4). A subset of station pairs was identified that exhibited a unique behavior regarding the lack of surface waves within either lag of the correlated waveforms. Depending on the time of year, a cross-correlated waveform would have one or two observable surface waves in their respective lags (Figure 4.3).

Pawlak et al., 2011 observed that the stacks can be divided into two, representing the northern hemisphere summer and winter months. The northern summer is from May to September, while the winter is from November to March. The results of this study are consistent with the summer months being more azimuthally distributed with a higher SNR compared to the winter (*Pawlak et al., 2011)*. While two distinctive patterns to the behavior of the waveforms are observed, the extent exceeds what was observed by *Pawlak et al.,* 2011. The cause is likely related to the less dramatic seasonal transition in the Midwestern United States compared to that of northern Canada. Therefore the observed temporal variation spans two distinctive periods, with the summer months spanning February to September and the winter months spanning October to January.

Since only a subset of station pairs exhibit this variability in the signal content its presence is nominal at best.
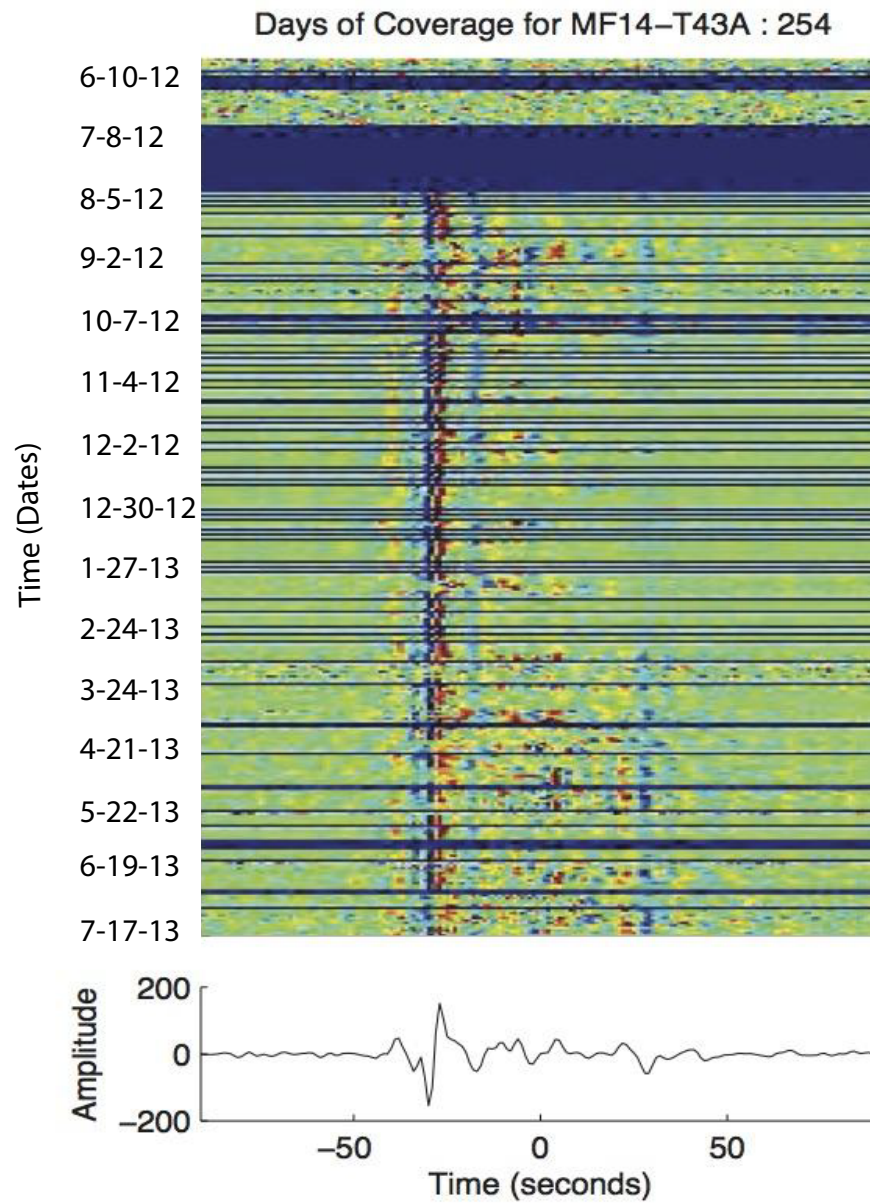
Figure 4.3

Example of the signal within the cross-correlations demonstrating a seasonal dependence for the station pair of MF14 and T43A based on strength of the amplitudes in either time lag. At certain points of the year the amplitudes in either lag vary, identifying a seasonal variation to the signal. The Seasonal variation is shown during the winter months, October to February, where a single wavefront is observed, while in the summer months, March to September both wavefronts are observed.

4.4 SNR Frequency and Directionality Dependence

As previously discussed, the emergence of Green's functions from the cross-correlation of a diffuse wavefield assumes that the sources of the ambient noise field are distributed homogeneously. This would produce two equal peaks in either lag of the cross-correlated waveform. Inspection reveals a strong asymmetry in which one lag has a significantly higher SNR than the other (Figure 3.6). This type of asymmetry is characteristic of stationary coastal noise sources (*Stehly et al.,* 2006). While this is to be expected with arrays that are located near large bodies of water it stands out in the Midwestern United States as an anomaly.

A similar trend is apparent when measuring the SNR of the waveforms. In order to remove bias from the results, the SNR is normalized by the square root of the station separation. The results, shown in Figure 4.4, demonstrate a frequency dependence to the SNR, with the signal strength decreasing with greater period length. The short periods (<20 s) are referred to as microseisms composed of the primary (10-20 s) and secondary (5-10s) microseism bands (*Stehly et al.,* 2006). The primary band is believed to be associated with oceanic swells interacting with oceanic shoals (*Capon,* 1973), while the secondary band represents the nonlinear interaction of P waves as they travel in opposing directions with similar frequencies (*Stehly et al.,* 2006). The frequency dependence of the SNR can also viewed through an animation, shown in Figure A.1, where the wave peaks are visible at the short periods and difficult to distinguish at greater period lengths. The asymmetrical nature and frequency content of the waveforms demonstrates how the energy generated from either coast easily travels into and affects the aseismic continental interior.

While the waveforms provided information about the content of their signal, directionality of the SNR can provide the source of the signal. Rose diagrams of the SNR were constructed for the overall array coverage using both the causal and acausal components of the cross-correlation (Figure 4.5). The majority of the signal can be traced to the coastal regions associated with Pacific and Atlantic oceans. The simplest

explanation for this is that the signal originating from either coast is enough to mask what little signal would travel from north or south of the array. The source of the ambient noise field further demonstrates how the pelagic sources that flank the Midwestern United States modify both the signal of the ambient noise field of the region.
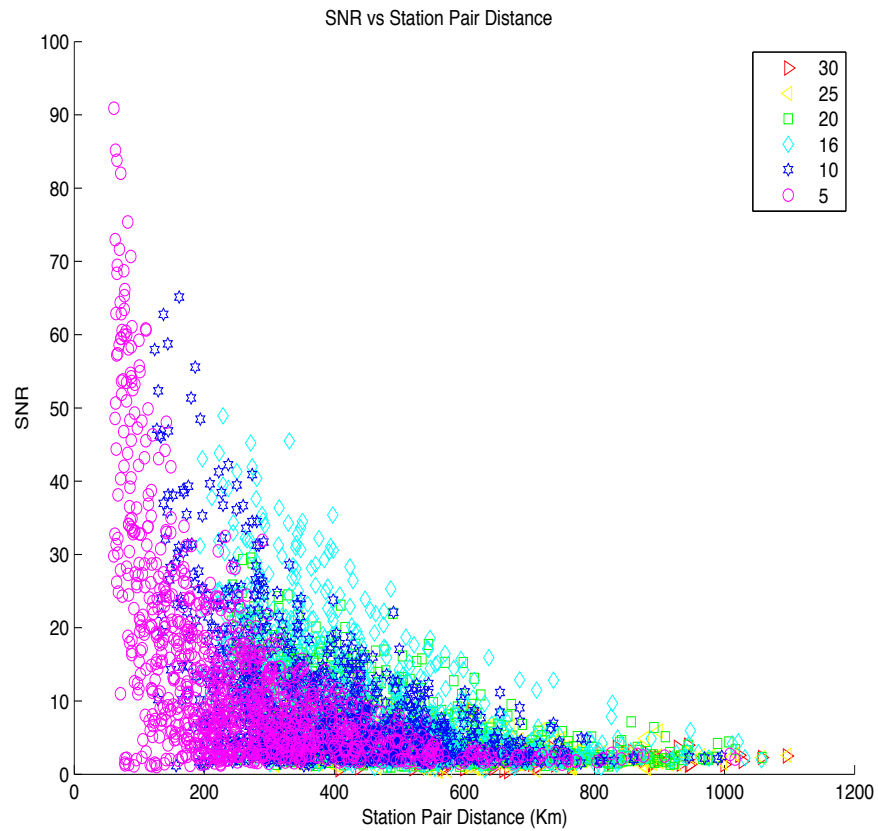
Figure 4.4

The SNR for each station pair plotted with their respective station pair distances. Each color represents a different period length. A correlation between the strength of the signal and the separation of the station pairs is clear. The signal is stronger for station pairs that are closer, since the energy of the waveforms has not dispersed. The SNR has been normalized by 1/r, where r represents the station pair distance.
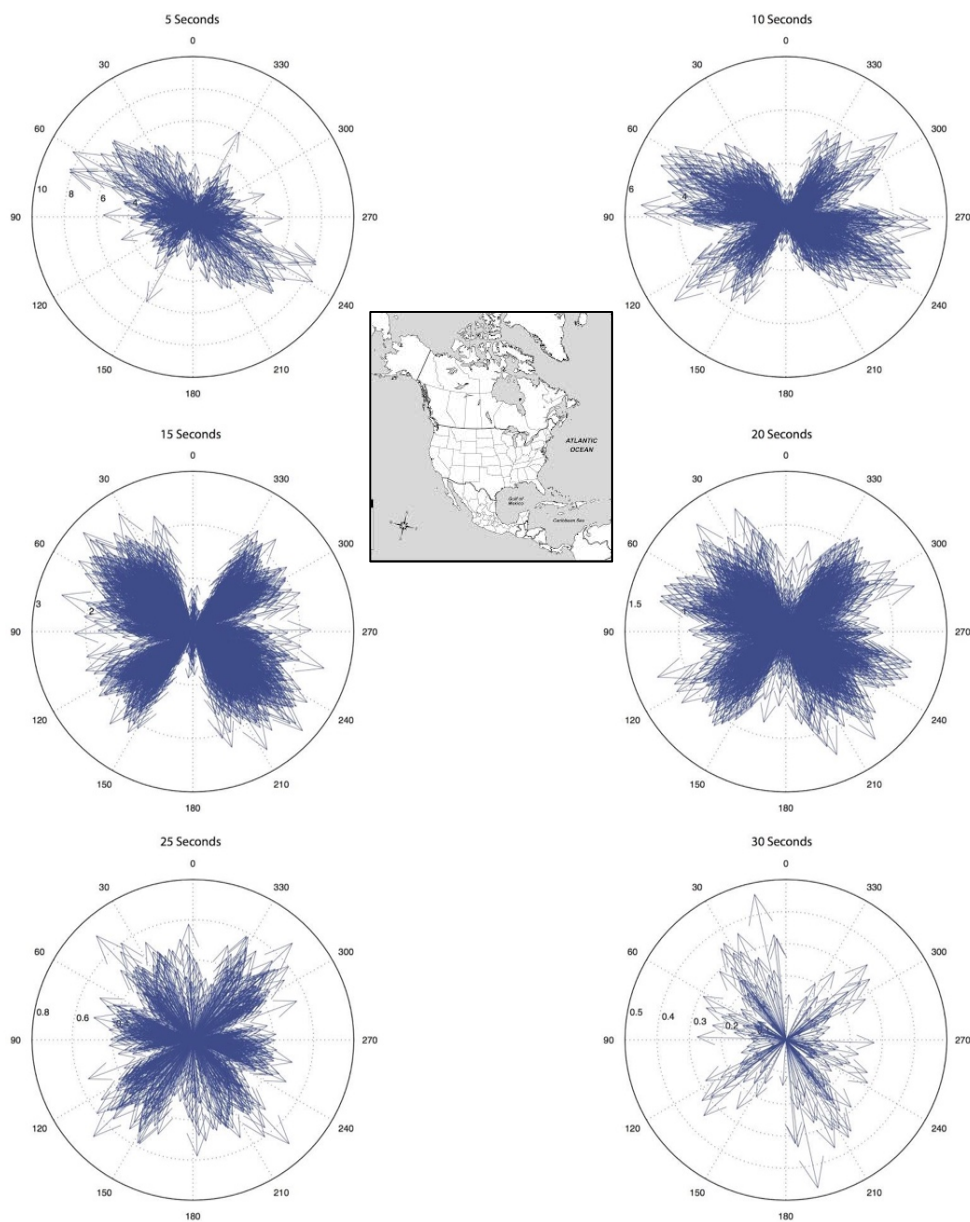
Figure 4.5

Rose plots of the SNR of the causal and acausal parts of the cross-correlated waveforms at 5 to 30 seconds, with a map of continental North America. The strong indicators in the eastern and western sections point to a pelagic source to the noise field.

4.5 Basin Amplification

By measuring the SNR at each period, geographical trends of the signal strength can be observed as a function of period length. Figure 4.6 shows examples of the ray paths between station pairs, with shading based on their SNR value. Since plotting every possible ray path would obscure any observable trend, only a subset of the ray paths was plotted. For the periods sensitive to the upper crust (< 15 s), the ray paths traveling through the Illinois Basin tend to be high compared to the mean SNR. Conversely the ray paths traveling along the margins of the basin exhibit a low trend. As period length increases these trends break down, with no observable trends present. This difference identifies a strong heterogeneity in the upper crust that transitions into more homogeneity at greater depth. Using the SNR as a proxy for signal strength and accounting for spurious ray paths, it is possible to observe how the Illinois Basin amplifies the seismic signal that travels through it.
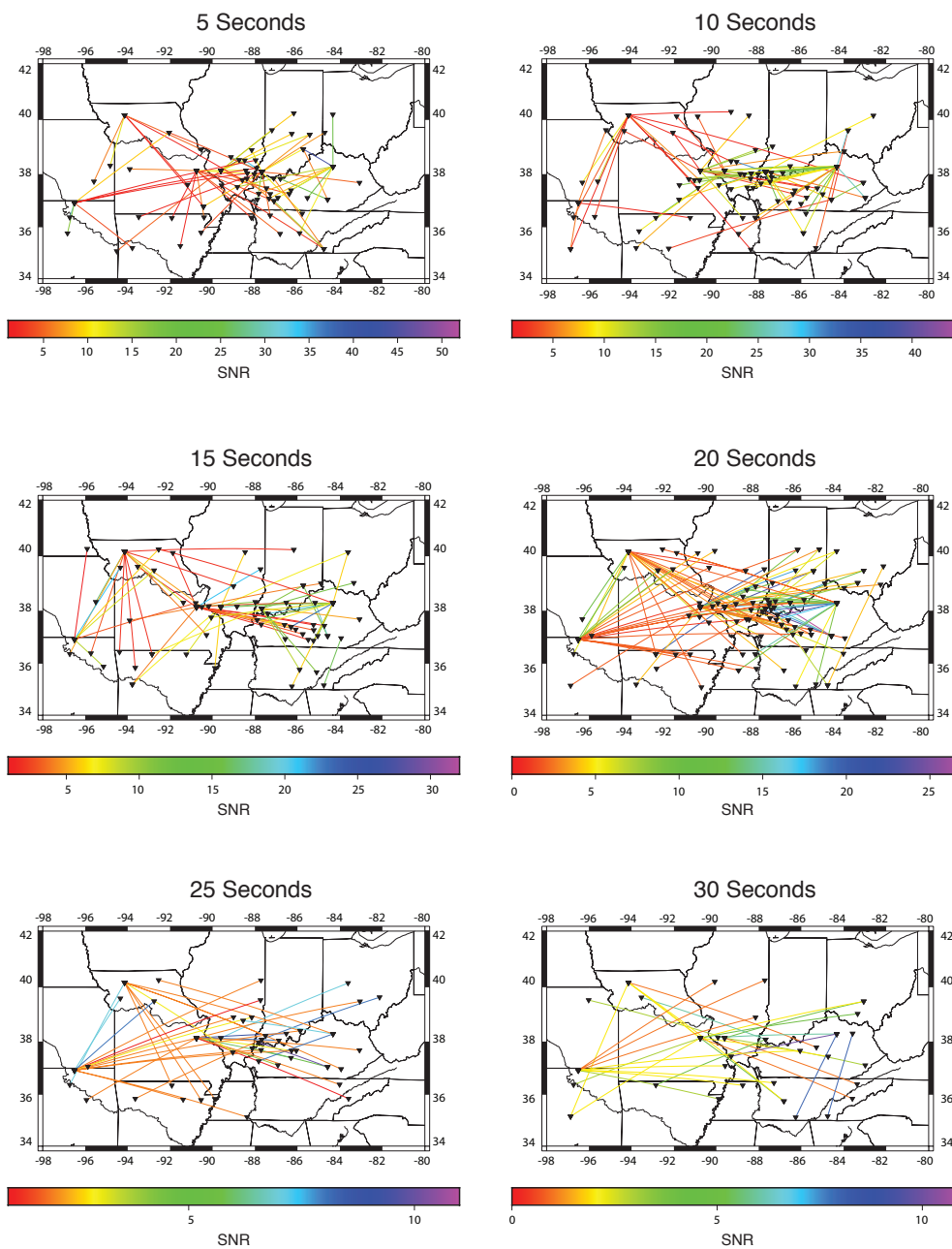
Figure 4.6

Ray paths for a subset of station pairs from 5 to 30 s. Each path is shaded based on their respective SNR. At shorter periods the area of the Illinois basin has a higher SNR compared to the regions surrounding it. At periods, longer than 20 s the signal of basin diminishes, suggesting a homogeneity to the strength of the signal.

4.6 Phase Velocity Maps

After obtaining phase velocity dispersions curves from all possible station pairs, the next step was to develop two-dimensional phase velocity maps. The numbers of ray paths for the different periods range from a high of 5,314 at 10 s, to a low of 320 at 30 s. The number of ray paths increases from 4s to 10s and decreases from 10 s to 30 s, establishing the best resolution centered at 10 s. Across all periods high resolution can be expected in the Illinois Basin, however at longer periods the structures are not well defined due to the paucity of measurements. Figure 4.7 demonstrates the difference in the coverage across the periods, with the densest coverage occurring from 9 to 20 s. The lack of measurements at either end of the measurement range results from the limitations of the measurement methods and the content of the signal within the region, both of which have been previously discussed. In any case, the overall coverage provides an adequate foundation upon which to examine the mid-crustal structural geology of the study region.

The Rayleigh wave phase velocity measurements at all station pairs are inverted for 2-D phase velocity maps on a grid of 0.5 by 0.5 using the method of *Barmin et al.,* 2001, Figure 4.10 and 4.11. The maps represent the observed behavior the material of the region at different period lengths. This method is based on minimizing a penalty function composed of a linear combination of data misfit, model smoothness and the perturbation of a reference model for isotropic wave speed (*Yang et al.,* 2007). The features of the maps vary with each period due to the frequency dependence of the depth sensitivity of Rayleigh waves, which is ⅓ of its wavelength. To quantify the best damping value to use for the study area, the work of *Ward et al.,* 2013 was followed to calculate the roughness of the model. The roughness is defined as:

$$\xi(\alpha,\lambda,\varphi) = \left\{ \sum_{i=1}^{18} \left[c(\alpha, T_{i+1}, \lambda, \varphi) - c(\alpha, T_i, \lambda, \varphi) \right]^2 \right\}^{1/2},$$

(10)

Where ξ is the roughness, α is the damping value, λ is the latitude in degrees, ϕ is the longitude in degrees, c is the phase velocity in km/s and T is the period in seconds (*Ward et al.,* 2013).

Ray Coverage at 5 Sec

Ray Coverage at 10 Sec

Ray Coverage at 15 Sec

Ray Coverage at 20 Sec

Ray Coverage at 25 Sec

Ray Coverage at 30 Sec
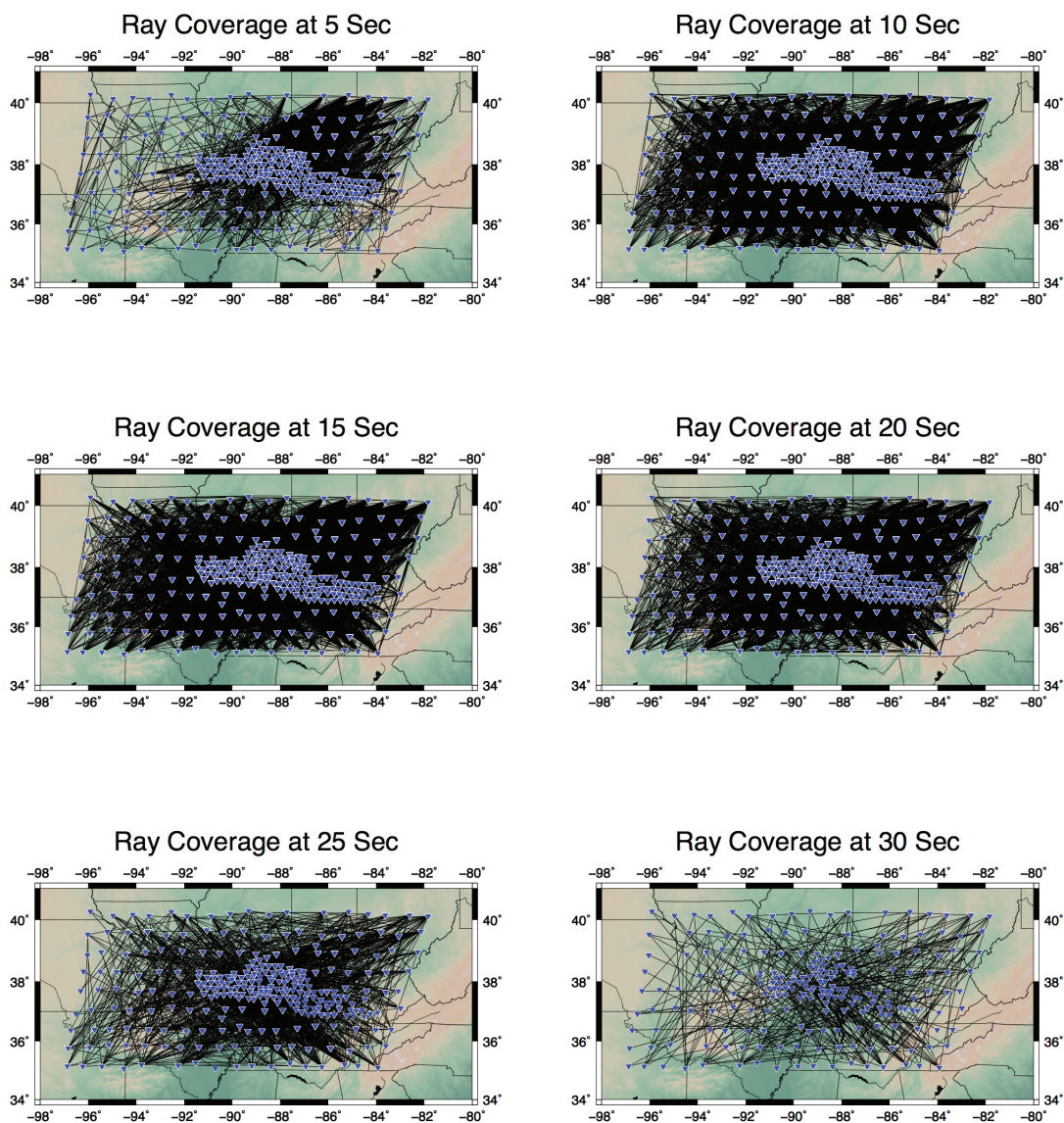
Figure 4.7

Ray coverage across periods 5, 10, 15, 20, 25, and 30 s. Black ray lines define the coverage and blue triangles represent the corresponding seismic stations at each period, with the densest coverage from 10 to 20 s.
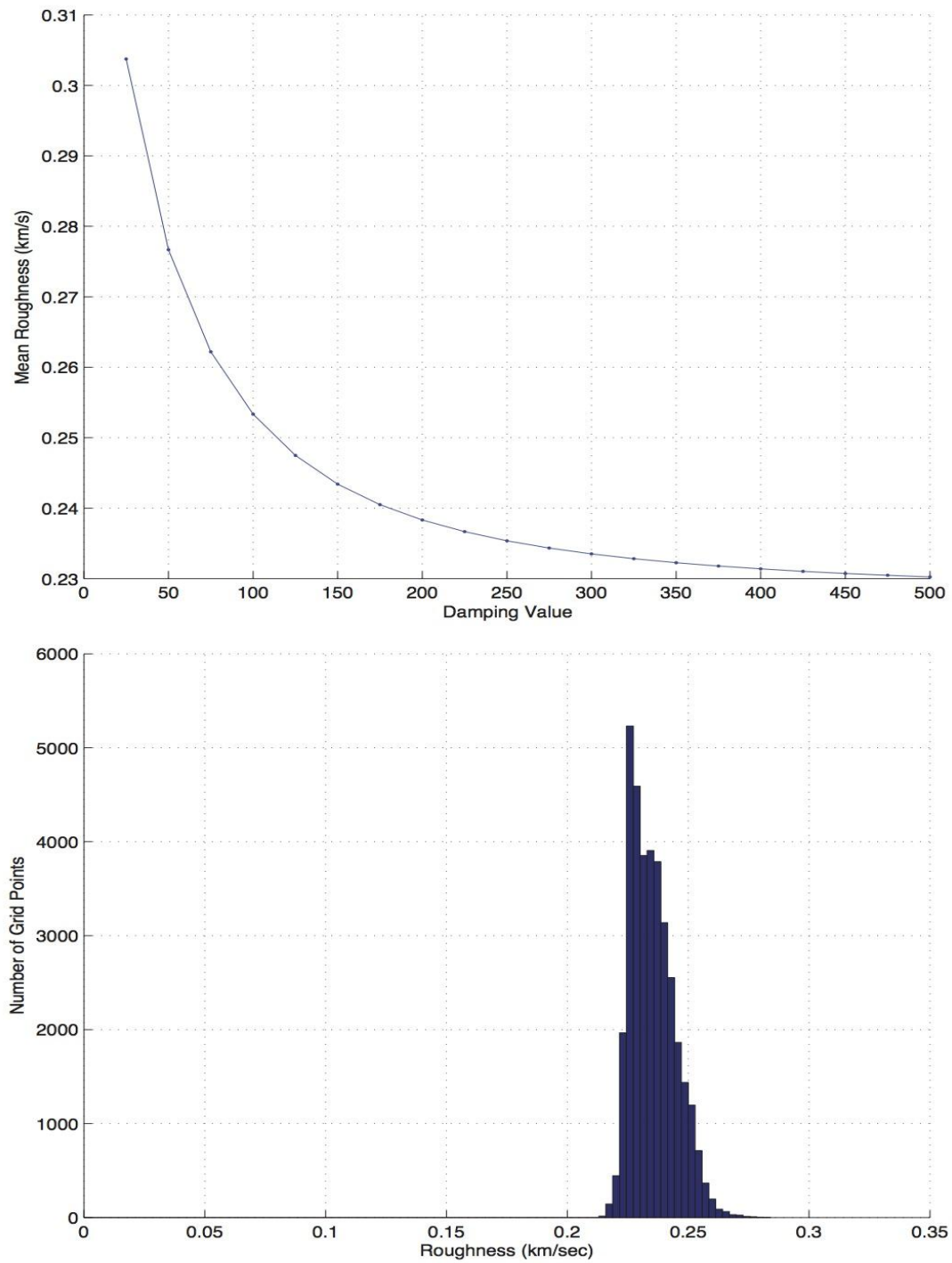
Figure 4.8

Plot of mean roughness versus damping values used in the 2-D phase velocity inversion. Top

figures show the trend of the roughness as function of the damping value. The bottom figure

demonstrates the roughness as a function of grid points.
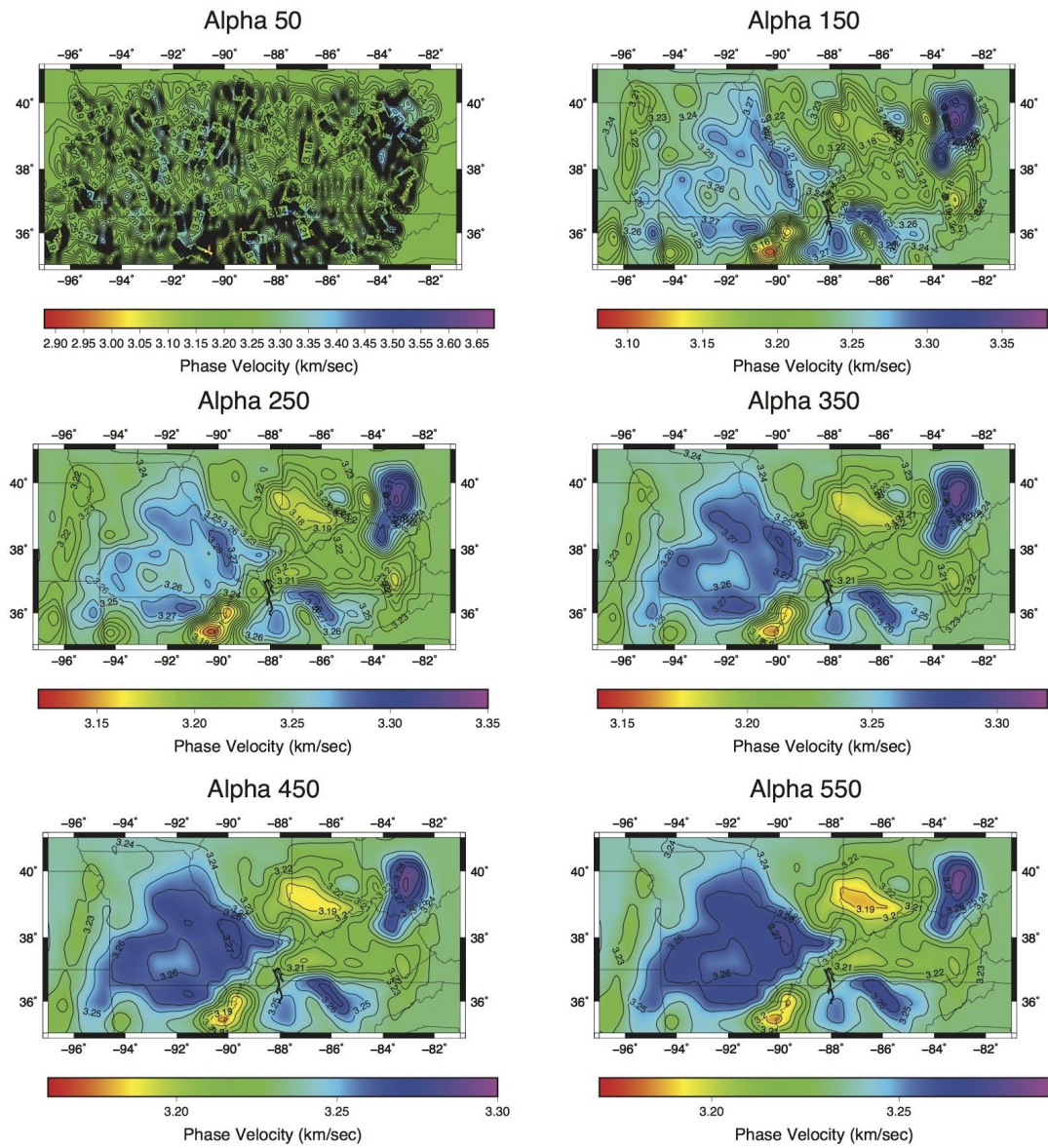
Figure 4.9

Examples of the effect of the damping factor alpha has on the 2-D inversion of the phase velocity measurements at 7 s. As the amount of damping increases the overall map becomes clearer, however too much will over simplify the results.

The purpose of the roughness is to gauge the average rate of change in velocity between periods for a phase velocity profile as a function of the damping value used in the inversion. The mean roughness value has no true physical meaning, however a plot of the mean roughness value as a function of the corresponding damping value provides a quantitative measure of the effect the damping value has on the 2-D inversion maps (Figure 4.8). Each point of the resulting curve demonstrates the overall transition across each measured period at the corresponding damping value. As the damping values increase the overall roughness drops. While this does eliminate anomalies in the maps, making them easier to interpret, it can lead to results that are over simplified and misrepresent the data. The slope of the curve representing the transition from underdamped, with measurements changing rapidly, to overdamped, where measurements are forced to be similar. The goal is to find a roughness value that results in stability between the periods while not overly damping the data. Through this method a damping value of 225 was found to be preferred due to its measured roughness equaling the overserved mean.
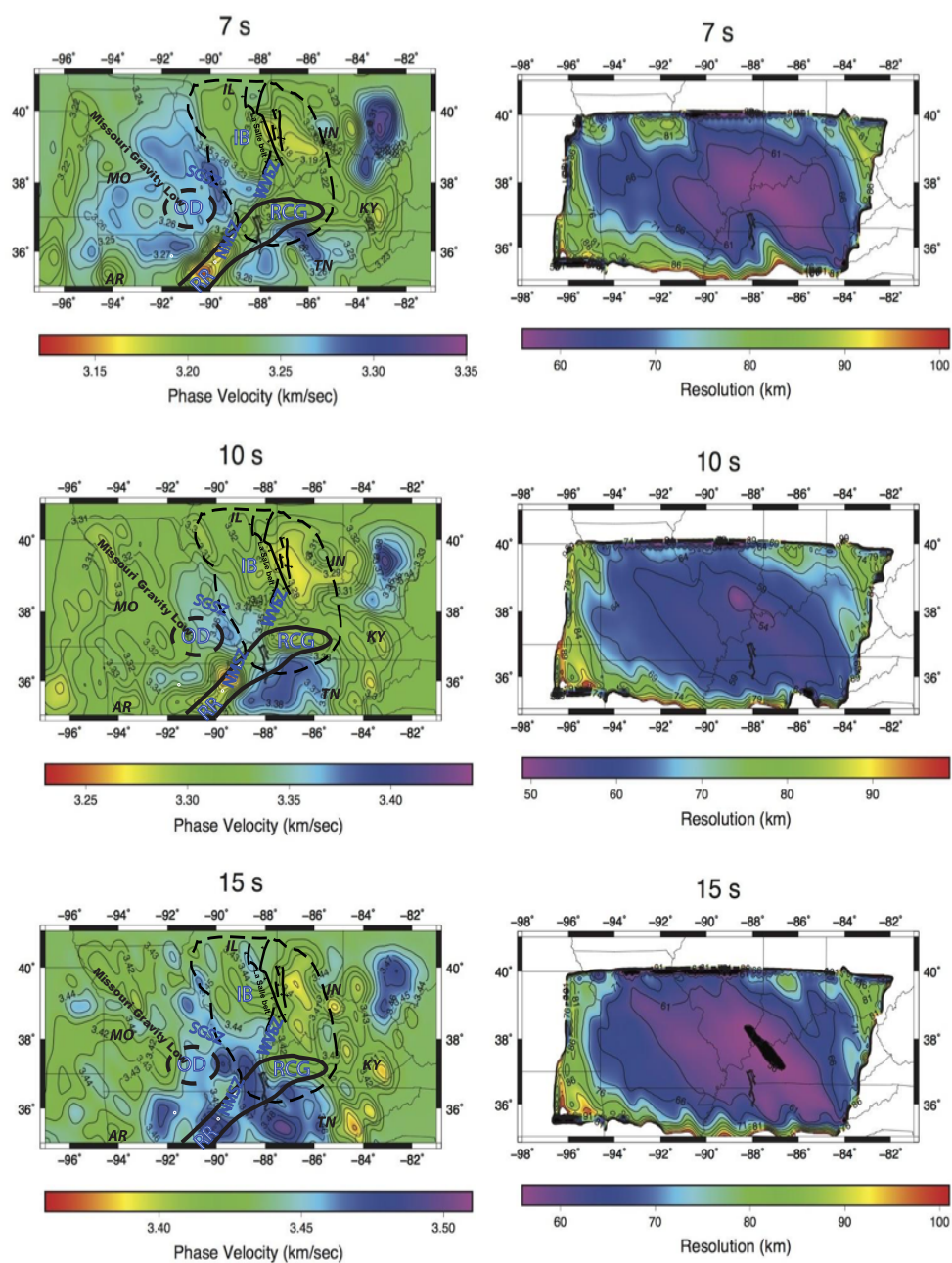
Figure 4.10

Phase velocity anomaly tomography maps from ambient seismic noise using data from the OIINK and USArray in the Midwestern United States from 7 to 15 s, which are sensitive to shear wave velocities in the upper to mid-crust. Figures on the left represent the phase velocity maps, with the figures on the right representing horizontal resolution. The geological structures of the region are outlined on the figures to the left.
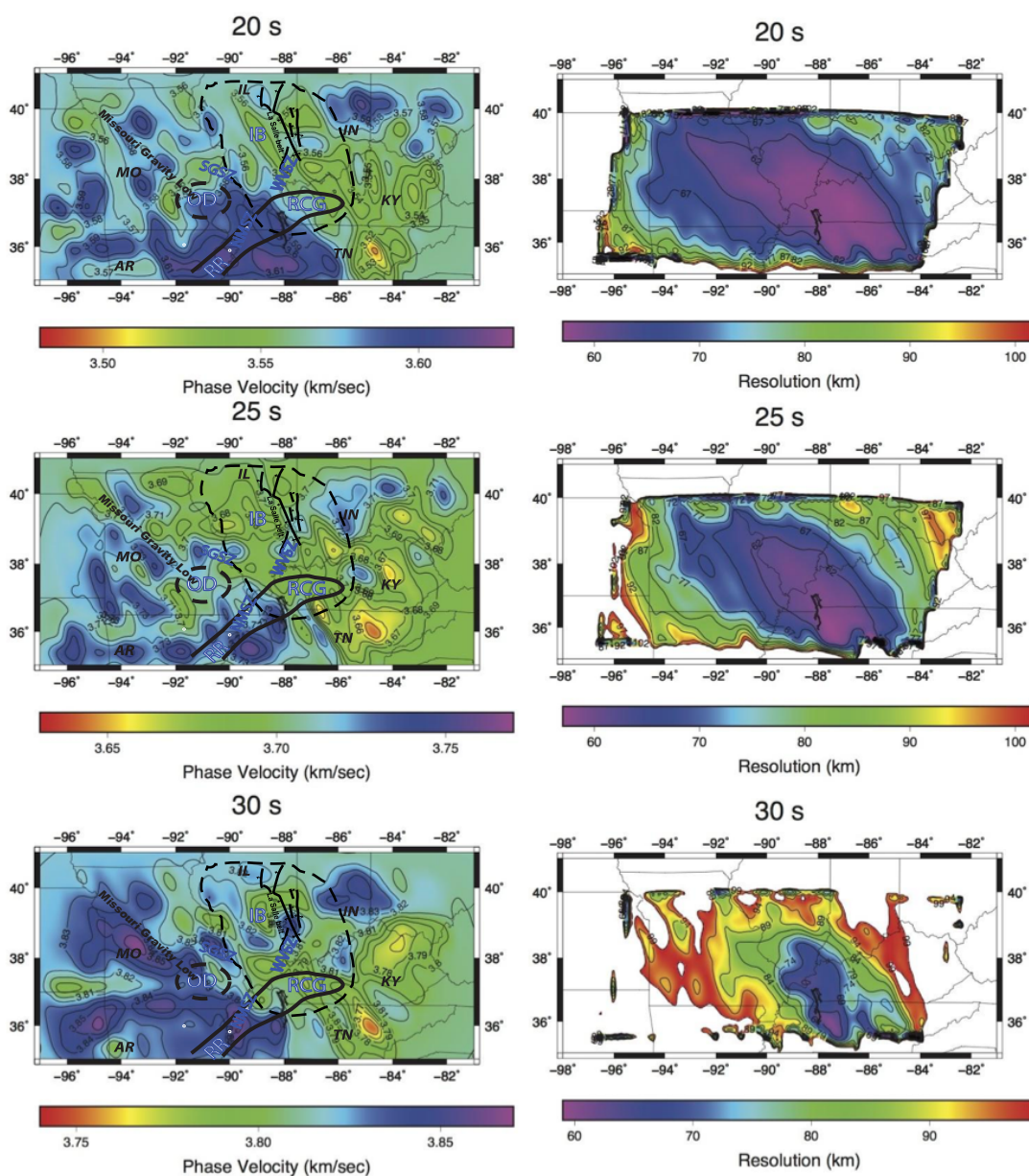
Figure 4.11

Phase velocity anomaly tomography maps from ambient seismic noise using data from the OIINK and USArray in the Midwestern United States from 20 to 30 seconds, which are sensitive to shear wave velocities in the mid to lower crust. Figures on the left represent the phase velocity maps, with the figures on the right representing horizontal resolution. The geologicl structures of the region are outlined on the figures to the left.
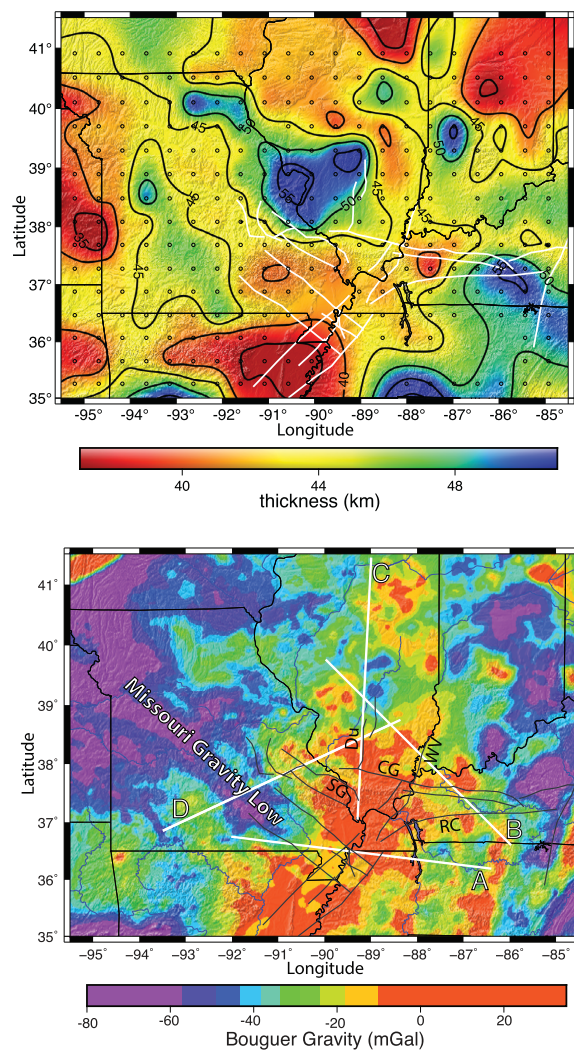
Figure 4.12

From *McGlannan et al.,* 2015, The top figure represents the measured crustal thickness of the Midwestern United States and the bottom figure shows the measured Bouguer gravity.

For periods that are most sensitive to the upper crust (< 15 seconds), the region within the Illinois Basin appears as a conspicuous, low-velocity element that strikes northwest. In contrast, the areas associated with the Ozark Dome and the western edge of Tennessee exhibit high-velocity trends. The Reelfoot Rift is represented by a low-velocity that increases in velocity as period length increases. Other features observable with this methodology include a high-velocity element located in central Ohio, which lines up with the Kankankee Arch, and a consistent low-velocity trend the closely correlates with the Rough Creek Graben. For increasing periods above 15 s, both the low- and high-velocity anomalies increase with size and demonstrate a strong northwest strike. The lack of ray paths at the longer periods (> 30 seconds) reduces the reliability of these anomalies compared to those observed at the shorter periods.

Strong contrast between anomalies points to key crustal differences in this portion of the North American craton. The majority of the slow velocities are observed in the eastern edge of the study area, likely corresponding to areas of thicker crust. In areas to the west, higher velocities could be caused by a closer proximity to mantle materials. One mechanism for the proximity of these anomalies could be melt from the mantle causing altering the crust, causing it sink and thereby expanding its presence to the longer period observed. When compared to crustal thicknesses calculated from gravity measurements the observed velocity anomalies demonstrate a strong correlation. The thickest section of crust observed within Illinois Basin corresponds well with the observed low velocity anomaly observed at the shorter periods (< 15 s). The overlapping of these observations and the Illinois Basin's velocity variation suggests that the basin cannot be characterized by a single crustal structure. The New Madrid Seismic Zone lies on relatively thin 40 km thick crust, while the Ozark Plateau represents some of the thickest crust in the Midwestern United States (*McGlannan et al.,* 2016). With the Reelfoot rift transitioning from a slow to fast anomaly as period length increases and the Ozark Plateau remains relatively fast. It is possible that magmatic underplating from the rift has emplaced mafic materials in the crust flanking the Illinois Basin. This would explain the thick crust and relatively fast velocities on the western edge of the basin, and

the rift's transition from a slow to fast anomaly as the warm mantle interacted with the cooler crust. This alteration could be the cause of deformation ultimately leading to subsidence of the cratonic lithosphere that formed the Illinois Basin. To examine this further the measured depth-sensitive phase velocities must be inverted to solve for shear velocity.
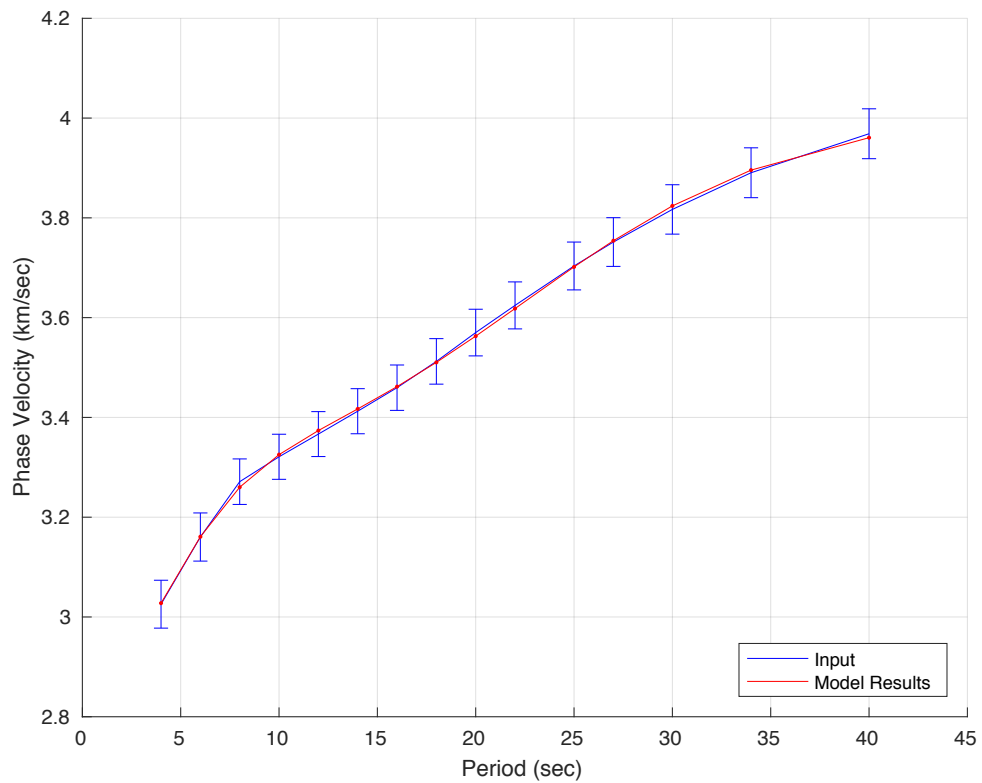
Figure 4.13

Regional average dispersion of the phase velocities spanning from 4 to 40 s. The input is the average dispersion curve resulting from inversion of the unique 21,506 dispersion curves with the error bars corresponding to the observed standard deviations of the phase velocities.

4.7 1-D Shear Velocity Profile

As an outgrowth of the 2-D inversion of the 21,506 dispersion curves, the average phase velocities at each period were used to construct a 1-D shear wave model of the Midwestern United States (Figure 4.13). The methods of *Hermann,* 2013 were used to invert for the 1-D shear velocity structure at each grid point by adjusting an initial velocity model to minimize the misfit between its predicted dispersion curve and the curve observed at each grid point (*Saito,* 1988). We used a reference velocity profile based on the ak135 model (*Kennett et al.,* 1995) for the shear wave inversion to have an established reference point. Two experiments were conducted on the starting velocity model to test the sensitivity of the inversion and the appropriate level of damping on the resulting shear velocity structure. The first set focused on the appropriate amount of damping to apply during the inversion (Figure 4.14). For the second set, the starting velocity of the input model was varied in increments of 0.2 km/sec to test the stability of the resulting models (Figure 4.15).

By varying the amount of damping applied we are able to establish a 1-D shear profile that accurately represents the data, while not being overly biased towards the input model. The level of damping applied varied from .01 to 1, representing a range from very low to very high damping. To establish the appropriate level, the dispersion of the input phase velocities was compared to those resulting from the inversion; a damping factor of 0.06 was established as the appropriate level for the inversion process (Figure 4.14). Varying the starting velocity of the 1-D model the stability of resulting shear wave model was then tested (Figure 4.15). It was observed that the starting velocity does not influence the resulting shear wave model, as all the results overlap. The results of these tests indicate the level of sensitivity of the inversion, and the appropriate level of damping needed to be applied as the measured phase velocities are inverted to shear velocities.
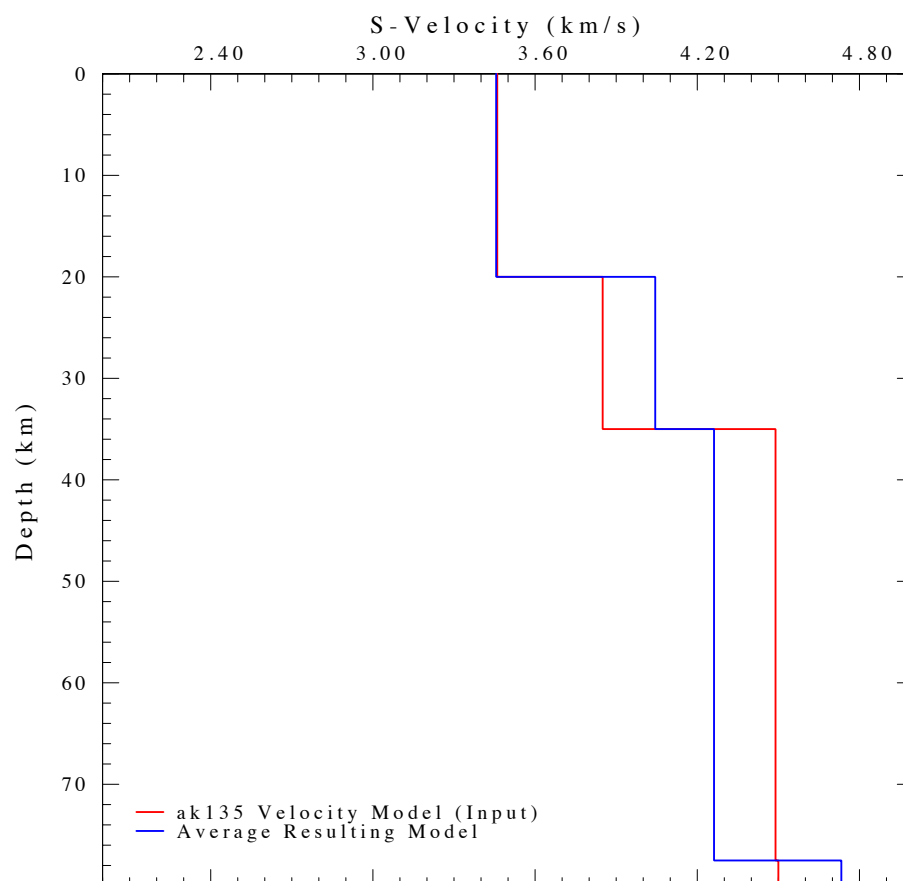
Figure 4.14

The initial and resulting average 1-D shear model based on averaging together all the points

across the 0.1° x 0.1° grid of the 2-D phase velocity inversions.
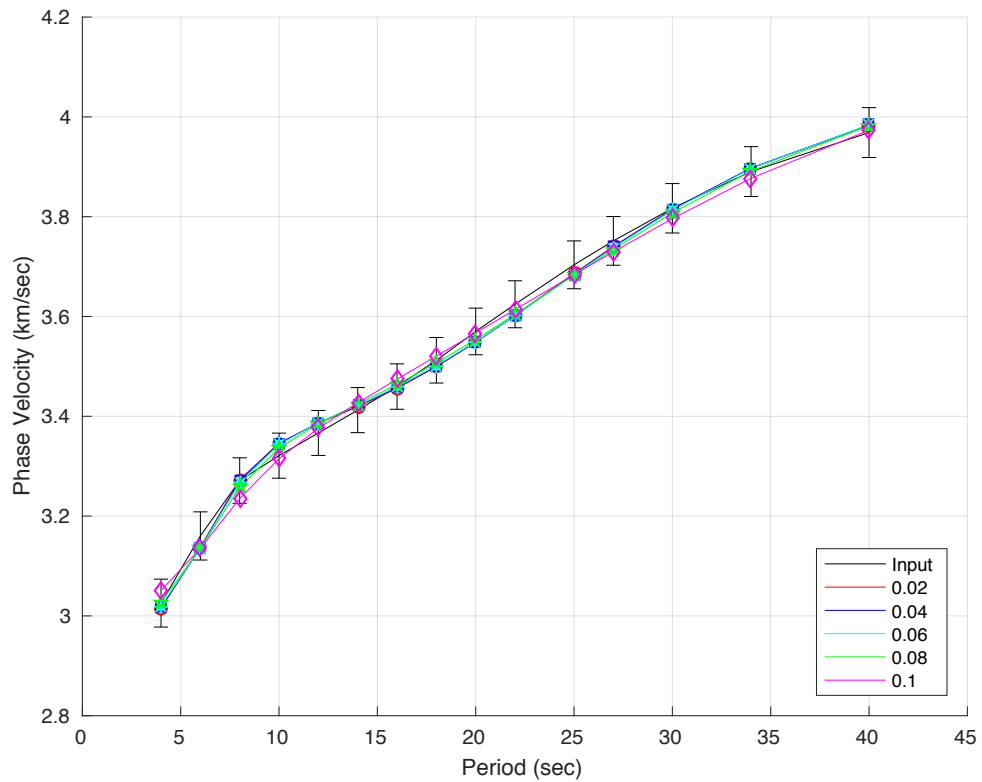
Figure 4.15

Dispersion of phase velocities from 4 to 40 s measured after the 1-D shear wave inversion. Average phase velocities from the inversion of 21,506 dispersion curves are shown in black. Phase velocities with different amounts of damping applied during the inversion are color coded.
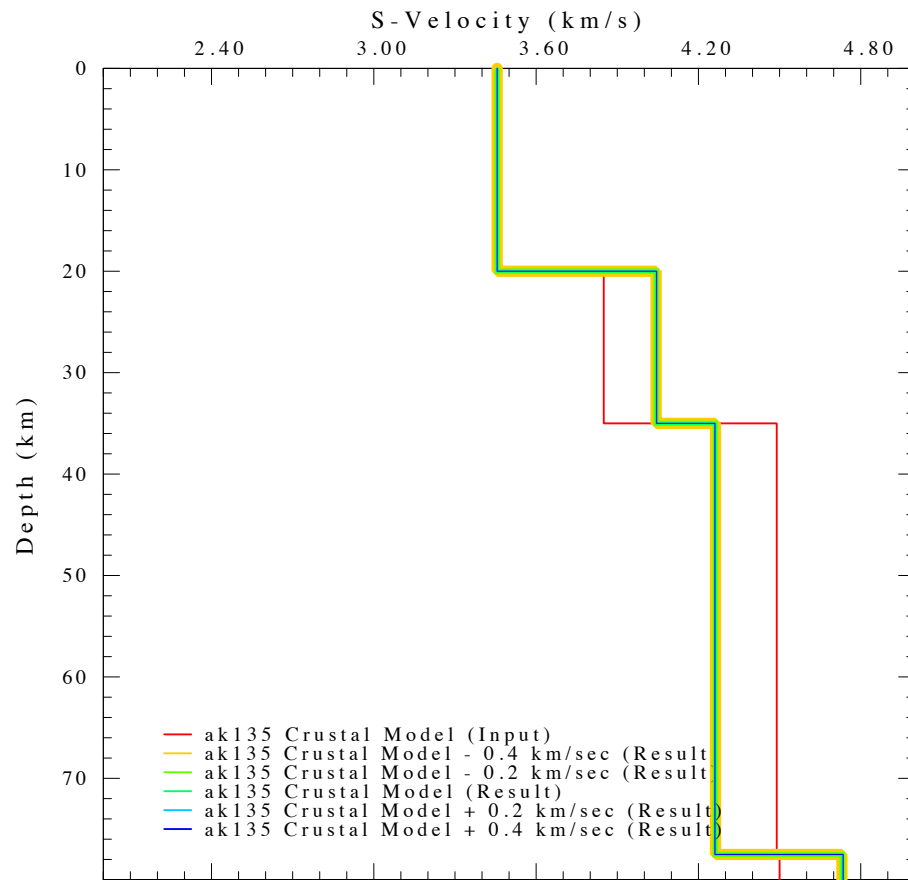
Figure 4.16

1-D shear models testing the input velocity during the 2-D inversion of the phase velocities from the ANT study at point [39.0, -90.0]. The results of each crustal model overlap one another demonstrating the overall stability of the results.

The average shear wave velocity in the upper crust was found to be approximately ~3.5 km/sec, which is similar to those observed in the ak135 model (*Kennett et al.,* 1995) (Figure 4.13). Within the lower crust, from 20 to 35 km, the observed velocities are ~0.2 km/sec faster those seen in the ak135 model. In the upper mantle the velocities increase to ~4.3 km/sec, which is slower than those previously observed for upper mantle of this area, but is likely due drop off of measurements and sensitivity of our longer period phase velocity measurements (*Chen et al.,* 2016). It is also worth noting that at ~4.25 km/sec, these velocities are slower than those predicted by the ak135 model. At greater depths, > 78 km, the velocities increase to be above 4.6 km/sec. At depths greater than these, the sensitivity of the short period phase velocities used here are too low for further consideration.

The increased velocity of the lower crust, shows that it has been altered. Possible causes for the velocity anomalies observed are heterogeneities in temperature, composition, water content, or a combination of the three (*Al-Shukri et al.,* 1987, *Bear et al.,* 1997, *Braile et al.,* 1986, and *Chen et al.,* 2016). With the observed high shear velocities, it is possible that mafic materials have been emplaced into the lower crust. With these materials being denser than those normally expected isostatic equilibrium would lead to the upper crust deforming. This deformation along with the regional stress field would be able to deform the crust creating the observable geological features in this region. The relatively slow upper mantle is indicative of rifting, where the mantle is warmer than expected leading to a reduction in the observed shear velocity (*West et al., 2003).* However, the Reelfoot Rift's thermal anomaly has had ample time to disparate, and its effects cannot completely explain the velocity reduction. Instead a combination of heterogeneities in temperature, composition and water content have been proposed to cause the shear velocities to be reduced in the mantle, which can all be introduced by rifting (*Chen et al., 2016).* Connecting the slow upper mantle with a fast-lower crust argues rifting and subsequent migration of melt into lower crust played a significant role in geological development of the region. The differences between this model and that observed in the Midwestern United States by *Bensen et al.,* 2009 likely results from our

focus on a particularly anomalous region. Further investigation of the structure at depth will improve our understanding of the development of the regional geological structures.

5. SUMMARY

This high-resolution study of surface wave dispersion using ambient seismic noise tomography on a combination of the OIINK flexible seismic array and the Transportable Array (TA) of the USArray produced insight into the subsurface geologic structures in the Midwestern United States. A total of 306 broadband seismic stations were used, with data coverage spanning from January 2011 to December 2014. By correlating four years of the vertical component of the available seismograms in 20 minute increments, Rayleigh wave phase velocity dispersion curves spanning from 4 to 40 s were obtained. Further analysis of the correlated data revealed geologic and seismic characteristics of the region, including potential sources of the ambient noise field and factors influencing its signal content.

During the measurement process, the components of the physical characteristics of signal within the ambient noise were identified. The majority of the noise signal can be traced to the coastal regions associated with Pacific and Atlantic oceans. The data suggest that the signal originating from either coast is enough to mask what little signal would travel from the north or the south of the array. Additionally, the variability of the weather in the Midwestern United States does not affect its ambient noise field. The source and behavior of the ambient noise field is strongly modified by the geological processes that flank the Midwestern United States.

By inverting the dispersion curves 2-D phase velocity maps at 4 s to 40 s were developed. While these results are not sensitive to the structure at depth, they reveal the basic structural characteristics of the region. A concentration of low phase velocities corresponding   to   areas   of   thicker   crust   in   Illinois

and Indiana is also observed, coincident with the Illinois Basin. With the Reelfoot Rift and Rough Creek Graben being located in zones of high phase velocities, a close proximity to mantle materials can be inferred. When compared to measured crustal thickness, the notion of magmatic underplating becomes a possibility. The thicker crust and relatively fast velocities on the Ozark Plateau, and the Reelfoot Rift's transition from a slow to fast anomaly as the warm mantle material interacts with a thinned crust suggests the two are closely related. This alteration could be the cause of the deformation that would lead to the subsidence of the cratonic lithosphere as it equilibrates with the mantle material beneath it. While this would thicken some regions of the crust, it would also allow the Illinois Basin to form as the surrounding crust is pulled down around it.

With the phase velocity data, 1-D shear-wave profiles were constructed that further illustrate the complex geological structures of the region. The 1-D shear velocity profile demonstrates how the lithosphere of the Midwestern United States differs from the average crustal ak135 model, where shear velocity increases with depth. The relative high and low velocities of the lower crust and upper mantle suggest that warm material, possibly from rifting, has been transmitted into and fundamentally altered the crust. With the crust altered, the strong cratonic lithosphere would be susceptible to deformation, possibly explaining the subsidence of the Illinois Basin within cratonic crust. These observations broadly agree with those proposed by *Braile et al.,* 1986, where dense mafic bodies have been emplaced and subsequently lead to the deformation of cratonic crust. While the 1-D model suggests only a general trend, it corroborates previous work seeking to understand the structural framework of the Midwestern United States. It is likely that the failed Reelfoot Rift was instrumental in the development of the Illinois Basin.

LIST OF REFERENCES

LIST OF REFERENCES

Aki, K. (1957.). Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors. *Bull. Earthq. Res. Inst*, (35), 415-457.

Barmin, M. P., Ritzwoller, M. H., & Levshin, A. L. (2001). A Fast and Reliable Method for Surface Wave Tomography. *Monitoring the Comprehensive Nuclear-Test-Ban Treaty: Surface Waves*, 1351-1375. doi:10.1007/978-3-0348-8264-4_3

Bassin, C., Laske, G., & Masters, G. (2000). The current limits of resolution for surface wave tomography in North America. *EOS Trans AGU, 81*.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International, 169*(3), 1239-1260. doi:10.1111/j.1365-246x.2007.03374.x

Bensen, G. D., Ritzwoller, M. H., & Shapiro, N. M. (2008). Broadband ambient noise surface wave tomography across the United States. *Journal of Geophysical Research, 113*(B5). doi:10.1029/2007jb005248

Bensen, G. D., Ritzwoller, M. H., & Yang, Y. (2009). A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise. *Geophysical Journal International, 177*(3), 1177-1196. doi:10.1111/j.1365-246x.2009.04125.x

Boschi, L., Weemstra, C., Verbeke, J., Ekstrom, G., Zunino, A., & Giardini, D. (2012). On measuring surface wave phase velocity from station-station cross-correlation of ambient signal. *Geophysical Journal International*, *192*(1), 346-358. doi:10.1093/gji/ggs023

Braile, L. W., Hinze, W. J., Keller, G., Lidiak, E. G., & Sexton, J. L. (1986). Tectonic development of the New Madrid rift complex, Mississippi embayment, North America. *Tectonophysics*, *131*(1-2), 1-21. doi:10.1016/0040-1951(86)90265-9

Buschbach, T., & Kolata, D. (1991). Regional setting of the Illinois Basin, in Leighton. *American Association of Petroleum Geologists Memoir*, *51*, 29-58.

Campillo, M. (2003). Long-Range Correlations in the Diffuse Seismic Coda. *Science*, *299*(5606), 547-549. doi:10.1126/science.1078551

Campillo, M., & Paul, A. (2002). Correlations in the seismic coda and elastic Green's functions. *The Journal of the Acoustical Society of America*, *112*(5), 2319. doi:10.1121/1.4779349

Capon, J. (2009). Analysis of Microseismic Noise at LASA, NORSAR and ALPA†. *Geophysical Journal of the Royal Astronomical Society*, *35*(1-3), 39-54. doi:10.1111/j.1365-246x.1973.tb02413.x

Chen, C., Gilbert, H., Andronicos, C., Hamburger, M. W., Larson, T., Marshak, S., Yang, X. (2016). Shear velocity structure beneath the central United States: implications for the origin of the Illinois Basin and intraplate seismicity. *Geochemistry, Geophysics, Geosystems*, *17*(3), 1020-1041. doi:10.1002/2015gc006206

De Hoop, M. V., & Solna, K. (2009). Estimating a Green's Function from "Field-Field" Correlations in a Random Medium. *SIAM Journal on Applied Mathematics*, *69*(4), 909-932. doi:10.1137/070701790

Derode, A., Larose, E., Tanter, M., De Rosny, J., Tourin, A., Campillo, M., & Fink, M. (2003). Recovering the Green's function from field-field correlations in an open scattering medium (L). *The Journal of the Acoustical Society of America*, *113*(6), 2973. doi:10.1121/1.1570436

Pawlak, E. A., & Eaton, W. D. (2010). Crustal Structure Beneath Hudson Bay from Ambient-noise Tomography - Implications for basin formation. *72nd EAGE Conference and Exhibition incorporating SPE EUROPEC 2010*. doi:10.3997/2214-4609.201401367

Ekström, G., Abers, G. A., & Webb, S. C. (2009). Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation. *Geophysical Research Letters*, *36*(18). doi:10.1029/2009gl039131

Hildenbrand, T. G. (1985). Rift Structure of the Northern Mississippi Embayment from the analysis of gravity and magnetic data. *J. Geophys. Res*, *90*(B14), 12607. doi:10.1029/jb090ib14p12607

Howe, J. R. (1989). Tectonics, Sedimentation, and Hydrocarbon Potential of Reelfoot Aulacogen: ABSTRACT. *Bulletin*, *73*. doi:10.1306/44b4a3a2-170a-11d7-8645000102c1865d

Kane, M. F., Hildenbrand, T. G., & Hendricks, J. D. (1981). Model for the tectonic evolution of the Mississippi embayment and its contemporary seismicity. *Geol*, *9*(12), 563. doi:10.1130/0091-7613(1981)92.0.co;2

Kennett, B. L., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108-124. doi:10.1111/j.1365-246x.1995.tb03540.x

Levshin, A. L., Ritzwoller, M. H., & Resovsky, J. S. (1999). Source effects on surface wave group travel times and group velocity maps. *Physics of the Earth and Planetary Interiors*, *115*(3-4), 293-312. doi:10.1016/s0031-9201(99)00113-2

Lin, F., 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. doi:10.1111/j.1365-246x.2008.03720.x

Lin, F., Ritzwoller, M. H., Townend, J., Bannister, S., & Savage, M. K. (2007). Ambient noise Rayleigh wave tomography of New Zealand. *Geophysical Journal International*, *170*(2), 649-666. doi:10.1111/j.1365-246x.2007.03414.x

Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the Green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, *110*(6), 3011. doi:10.1121/1.1417528

Luo, Y., Yang, Y., Xu, Y., Xu, H., Zhao, K., & Wang, K. (2015). On the limitations of interstation distances in ambient noise tomography. *Geophysical Journal International*, *201*(2), 652-661. doi:10.1093/gji/ggv043

Lewan, M. D., Henry, M. E. (2002). Material-balance assessment of the New Albany-Chesterian petroleum system of the Illinois basin. *AAPG Bulletin*, *86*. doi:10.1306/61eedb8e-173e-11d7-8645000102c1865d

Mac Niocaill, C., Van der Pluijm, B. A., & Van der Voo, R. (1997). Ordovician paleogeography and the evolution of the Iapetus ocean. *Geology*, *25*(2), 159. doi:10.1130/0091-7613(1997)0252.3.co;2

McBride, J. H., & Kolata, D. R. (1999). Upper crust beneath the central Illinois basin, United States. *Geological Society of America Bulletin*, *111*(3), 375-394. doi:10.1130/0016-7606(1999)1112.3.co;2

McGlannan, A. J., & Gilbert, H. (2016). Crustal signatures of the tectonic development of the North American midcontinent. *Earth and Planetary Science Letters*, *433*, 339-349. doi:10.1016/j.epsl.2015.10.048

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*(8). doi:10.1029/2007gc001655

Nelson, K., & Zhang, J. (1991). A COCORP deep reflection profile across the buried reelfoot rift, south-central United States. *Tectonophysics*, *197*(2-4), 271-293. doi:10.1016/0040-1951(91)90046-u

Nuttli, O. W. (2009). CHAPTER 9: The central Mississippi River Valley and its great earthquakes of 1811–1812. *Geological Society of America Special Papers*, 83-102. doi:10.1130/2009.3455(9)

Parrish, S., & Van Arsdale, R. (2004). Faulting along the Southeastern Margin of the Reelfoot Rift in Northwestern Tennessee Revealed in Deep Seismic-reflection Profiles. *Seismological Research Letters*, *75*(6), 784-793. doi:10.1785/gssrl.75.6.784

Pawlak, A., Eaton, D. W., Bastow, I. D., Kendall, J., Helffrich, G., Wookey, J., & Snyder, D. (2010). Crustal structure beneath Hudson Bay from ambient-noise tomography: implications for basin formation. *Geophysical Journal International*, *184*(1), 65-82. doi:10.1111/j.1365-246x.2010.04828.x

Pollitz, F. F., & Snoke, J. A. (2010). Rayleigh-wave phase-velocity maps and three-dimensional shear velocity structure of the western US from local non-plane surface wave tomography. *Geophysical Journal International*, *180*(3), 1153-1169. doi:10.1111/j.1365-246x.2009.04441.x

Rapine, R., Tilmann, F., West, M., Ni, J., & Rodgers, A. (2003). Crustal structure of northern and southern Tibet from surface wave dispersion analysis. *J. Geophys. Res*, *108*(B2).

Rhie, J., & Romanowicz, B. (2006). A study of the relation between ocean storms and the Earth's hum. *Geochemistry, Geophysics, Geosystems*, *7*(10), n/a-n/a. doi:10.1029/2006gc001274

Ritzwoller, M. H., & Levshin, A. L. (1998). Eurasian surface wave tomography: Group velocities. *Journal of Geophysical Research: Solid Earth*, *103*(B3), 4839-4878. doi:10.1029/97jb0262

Sabra, K. G. (2005). Extracting time-domain Green's function estimates from ambient seismic noise. *Geophysical Research Letters*, *32*(3). doi:10.1029/2004gl021862

Sabra, K. G., Roux, P., & Kuperman, W. A. (2005). Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function. *The Journal of the Acoustical Society of America*, *118*(6), 3524. doi:10.1121/1.2109059

Scotese, C. R., & Mckerrow, W. S. (1991). Ordovician Plate Tectonic Reconstructions. doi:10.4095/132195

Shapiro, N. M. (2005). High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. *Science*, *307*(5715), 1615-1618. doi:10.1126/science.1108339

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, *31*(7), n/a-n/a. doi:10.1029/2004gl019491

Shapiro, N. M., & Ritzwoller, M. H. (2002). Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle. *Geophysical Journal International*, *151*(1), 88-105. doi:10.1046/j.1365-246x.2002.01742.x

Stark, T. J. (1997). The East Continent rift complex: Evidence and conclusions. *Special Paper 312: Middle Proterozoic to Cambrian rifting, central North America*, 253-266. doi:10.1130/0-8137-2312-4.253

Stehly, L., Campillo, M., & Shapiro, N. M. (2006). A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research*, *111*(B10). doi:10.1029/2005jb004237

Van der Pluijm, B. A., Johnson, R. J., & Van der Voo, R. (1990). Early Paleozoic paleogeography and accretionary history of the Newfoundland Appalachians. *Geology*, *18*(9), 898. doi:10.1130/0091-7613(1990)0182.3.co;2

Ward, K. M. (2015). Ambient noise tomography across the southern Alaskan Cordillera. *Geophysical Research Letters*, *42*(9), 3218-3227. doi:10.1002/2015gl063613

Ward, K. M., Porter, R. C., Zandt, G., Beck, S. L., Wagner, L. S., Minaya, E., & Tavera, H. (2013). Erratum: Ambient noise tomography across the Central Andes. *Geophysical Journal International*, *196*(2), 1264-1265. doi:10.1093/gji/ggt429

Weaver, R. L. (2005). GEOPHYSICS: Information from Seismic Noise. *Science*, *307*(5715), 1568-1569. doi:10.1126/science.1109834

Weaver, R. L. (2008). Ward identities and the retrieval of Green's functions in the correlations of a diffuse field. *Wave Motion*, *45*(5), 596-604. doi:10.1016/j.wavemoti.2007.07.007

Weaver, R. L., & Lobkis, O. I. (2004). Diffuse fields in open systems and the emergence of the Green's function (L). *The Journal of the Acoustical Society of America*, *116*(5), 2731. doi:10.1121/1.1810232

Weemstra, C., Boschi, L., Goertz, A., & Artman, B. (2013). Seismic attenuation from recordings of ambient noise. *GEOPHYSICS*, *78*(1), Q1-Q14. doi:10.1190/geo2012-0132.1

Wheeler, R. L. (1997). Boundary Separating the Seismically Active Reelfoot Rift from the Sparsely Seismic Rough Creek Graben, Kentucky and Illinois. *Seismological Research Letters*, *68*(4), 586-598. doi:10.1785/gssrl.68.4.586

Yang, Y., Ritzwoller, M., Moschetti, M., & Forsyth, D. (2007). Ambient seismic noise and teleseismic tomography in the western USA: High-resolution 3-D model of the crust and upper mantle from Earthscope/USArray. *The Journal of the Acoustical Society of America*, *121*(5), 3101. doi:10.1121/1.4782010

Yao, H., Van der Hilst, R. D., & De Hoop, M. V. (2006). Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps. *Geophysical Journal International, 166*(2), 732-744. doi:10.1111/j.1365

APPENDIX

APPENDIX

The following MATLAB scripts were used to process seismic data stored in SEED file formatting for the above ambient noise study. Of the included scripts, only those used during the Spectral method are unique to this study. Those used during the FTAN method and Shear wave inversion are available from the following sources.

- FTAN analysis of ambient noise cross-correlation: available from Colorado University at http://ciei.colorado.edu/Products/

- Computer Programs in Seismology. A description of the programs can be followed from:

  o Herrmann, R. B. (2013) Computer programs in seismology: An evolving tool for instruction and research, Seism. Res. Lettr. 84, 1081-1088, doi:10.1785/0220110096

  o Software package can be found at http://www.eas.slu.edu/eqc/eqccps.html

1. run_step_n.bash

```bash
#!/bin/bash
#Just a wrapper for step_1.m and step_2.m to avoid memory problems with extended
matlab runs.  Adjust parameters and maybe numdays (number of days to process before
restarting matlab).
#########################
#PARAMETERS to be set  #
#########################
component=LHZ
# Format is from number of days the will be processed, again day 0 is the first day of data
startmonth=0
endmonth=100
year=2011
#####################
#END PARAMETERS #
#####################
if [ $# -ne 2 ]
then
    echo "Usage: $0 {full_input_db_path} {output-dir}"
    echo "    on taupo, a good DB for CAFE can be found at:"
    echo '   /Users/jcalkins/CAFE/DB/cafplusLH_07only'
    echo "    on josh's laptop, a good DB for CAFE can be found at:"
    echo '   /Users/jcalkins/Research/Cafe/DB/cafplusLH'
    exit 1
fi
yrepochsec=`TZ=UTC date --date="$year-01-01 05:00:00" +%s`
echo "yrepochsec is " $yrepochsec
db_path=$1
outdir=$2
numdays=0
iterdays=$(( $numdays + 1 ))
mfile=step_1
#mfile=step_2
echo starting at `date`
[ ! -d $outdir ] && mkdir $outdir
cd $outdir
for (( i=$startmonth; i<=$endmonth; i++ ))
 do
 startday=$(( $i * $iterdays ))
 export MNBSJDAY=$startday
 export MNBNUMDAYS=$numdays
 export MNBCOMPONENT=$component
 export YREPOCHSEC=$yrepochsec
 export MNDB=$db_path
```

```
    #DEBUG
    echo -e "\n\nRunning $mfile with following env vars:"
    echo "    MNBSJDAY --- $startday"
    echo "    MNBNUMDAYS --- $numdays"
    echo "    MNBCOMPONENT --- $component"
    echo "    YREPOCHSEC --- $yrepochsec"
    echo -e "    MNDB --- $db_path\n\n"
mkdir $i
cd $i
    matlab -r $mfile
cd ..
done
cd ../
echo end
```

2. step_1.m

```
%Step 1 of the ambient noise processing code, as used by Josh Boschelli. This is run by
the bash script run_step_n.bash. Saves antelope data into .mat format
clear all; %removes any variables that might exist%
setup
%Connects antelope with MATLAB, depending on what the version of antelope that is
being used a specific version of MATLAB might be needed.
stations={'BLO' };
preproc_total=zeros(length(stations),24,3600);
%creates a data base where all the data is saved. It is in the form of stations x n hour x
mins in n hour
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%           Data Input from Wrapper          %%
%%                                            %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
start_julday=getenv('MNBSJDAY');start_julday=str2num(start_julday);
%The start of data sample range.
numdays=getenv('MNBNUMDAYS');numdays=str2num(numdays);
%The number of days that are being processed.
component=getenv('MNBCOMPONENT');
%The component of the seismic data.
yrepochsec=getenv('YREPOCHSEC');yrepochsec=str2num(yrepochsec);
%The length of the data window.
full_db_path=getenv('MNDB');%The path to the antelope dbe.
 if exist('start_julday')~=1;
    disp('FATAL ERROR: start_julday not defined.  Exiting');
    return
 end
 if exist('component')~=1;
    disp('FATAL ERROR: component not defined.  Exiting');
    return
 end
 if exist('numdays')~=1;
    disp('FATAL ERROR: numdays not defined.  Exiting');
    return
 end
 if exist('yrepochsec')~=1;
    disp('FATAL ERROR: yrepochsec not defined.  Exiting');
    return
 end
 if exist('full_db_path')~=1;
    disp('FATAL ERROR: full_db_path not defined.  Exiting');
    return
 end
```

```matlab
  ts= yrepochsec+(86400*(start_julday-1));%the time step, currently one day in terms of
seconds
  db=dbopen(full_db_path,'r'); %opens the database path
  kdayall=1;
dt=1; %data sample interval
npoles=5; % strength of filters
hi_corner=0.3; %filter corners in Hz
lo_corner=0.003;
fullwin_len=86400; %length of window to preprocess
procwin_len=3600; %for outer time loop length
xwin_len=1200;
filt_flag=1;      % make sure to have this on!
rminst_flag=0;     % remove instrument -- not necessary if whitening
whiten_flag=0;
savepref='raw_Day_data';   % prefix for outputs
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%       End parameters to be set    %%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
ts0=ts;
samprate=1/dt;
nyq=0.5/dt;
wn=[lo_corner/nyq, hi_corner/nyq];
[B,A]=butter(npoles,wn);
%Window & sample specifics
nproc_wins=floor(fullwin_len/procwin_len);
proc_win_samps=floor(procwin_len*samprate);
fflen=2^(nextpow2(procwin_len)+1);
%Strings for plot labels & filenames
procs='Processes: ';
if (rminst_flag==1), procs=[procs,'RMINST ']; end;
if (filt_flag==1), procs=[procs,'FILT ']; end;
procs=sprintf('%s PW %d XW %d',procs,procwin_len, xwin_len);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Antelope
Commands%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
db=dbopen(full_db_path,'r');
dbwf=dblookup_table(db,'wfdisc');
subsetcomp=sprintf('chan=~/%s/',component);
dbwf=dbsubset(dbwf,subsetcomp);
dbsi=dblookup_table(db,'site');
 %%%db queries that only need to be done once per stn
 slats=zeros(1,length(stations)); slons=zeros(1,length(stations)); %gather stn info
 ind=1;
 for stnind1=1:length(stations)-1
  for stnind2=stnind1+1:length(stations)
    stnpair_labels(ind)=strcat(stations(stnind1),'-',stations(stnind2));
```

```matlab
    [stnpair_dist(ind),
stnpair_az(ind)]=distance(slats(stnind1),slons(stnind1),slats(stnind2),slons(stnind2));
    ind=ind+1;
  end
 end
 %Calculate instrument response corrections once and assume the instruments don't
change
 if (rminst_flag==1)
   dbsen=dblookup_table(db,'sensor');
   dbin=dblookup_table(db,'instrument');
   dbsnin=dbjoin(dbsen,dbin);

resp=zeros(length(stations),fflen);
   dtr=zeros(length(stations),1);
   for ind=1:length(stations)
     thisstn=char(stations(ind));
     [resp(ind,:) dtr(ind,:)]=calcinstresp(dbsnin,thisstn,'LHZ',-1, fflen, lo_corner);
   end
 end
 %%%%%%%%%%%%%%
 %     Preprocess     %
 %%%%%%%%%%%%%%%%
 daycntr=0;
 i=1;
 %Create an index with dimensions (number of stations x number of proc_wins x number
of days )
 % into kdayall. i.e., good_data_ind(:,:,1) corresponds to the day in kdayall(1),
 good_data_ind=zeros(length(stations),nproc_wins,length(kdayall));
 juldayindex=zeros(length(kdayall),1);
 iii = 0;
 for kday=kdayall% LOOP Over 1-day segments
   savepref=sprintf('Raw_TA_Day_data_%i',kday);   % prefix for outputs
   daycntr=daycntr+1;
   %Create a data array for all procwins, all stations, this kday
   preproc_data=zeros(length(stations),nproc_wins,procwin_len);
   ts=ts0+86400*kday;    % start
   julday=yearday(ts);
   juldayindex(i,1)=julday;
   i=i+1;
   te=ts+fullwin_len;
   for stncntr=1:length(stations)  %%Begin loop through stations
     thisstn=char(stations(stncntr));
     disp(sprintf('Preprocessing station %s for day %d',thisstn,julday));
      substr1=sprintf('(wfdisc.time <= %d && wfdisc.endtime > %d && sta =~/%s/)
||(wfdisc.time > %d && wfdisc.time <%d && sta =~/%s/)',ts,ts,thisstn,ts,te,thisstn);
```

```
        dbtr1=dbsubset(dbwf,substr1);
        if (dbquery(dbtr1,'dbRECORD_COUNT'))==0
         disp(sprintf('NO DATA for stn: %s day: %s',thisstn,strdate(ts)));
          preproc_data(stncntr,:,:)=NaN;
continue;
        end
        %%%%Read full window from DB
        trptr1=trload_css(dbtr1,ts,te);
        trsplice(trptr1,50);
        if (dbnrecs(trptr1)~=1)
           disp(sprintf('day: %d %s stn: %s MULTIPLE TRACES',julday,strdate(ts),thisstn));
           good_data_ind(stncntr,:,daycntr)=0;
           preproc_data(stncntr,:,:)=NaN;
           trdestroy(trptr1);
           continue;
        end
        [tim1 nsmp1]=dbgetv(trptr1,'time','nsamp');
        if (nsmp1 < (fullwin_len) || abs(tim1-ts)>dt)
           disp(sprintf('day: %d %s stn %s SHORT TRACE or START TIME MISMATCH:
length= %d, ts=%d, tim1=%d',...
              julday,strdate(tim1),thisstn,nsmp1,ts,tim1));
           good_data_ind(stncntr,:,daycntr)=0;
           preproc_data(stncntr,:,:)=NaN;
           trdestroy(trptr1);
           continue;
        end
        rawtr1=trextract_data(trptr1);
        trdestroy(trptr1);
        fullwin1=rawtr1;
        for nouter=0:nproc_wins-1
         proc_win1=fullwin1(1+(nouter*procwin_len):(nouter*procwin_len)+procwin_len);
         proc_win1=detrend(proc_win1);
         if rminst_flag==1
           f1=fft(proc_win1,fflen);
           if rminst_flag==1
              f1=f1.*resp(stncntr,:)';
              if (abs(dt-dtr(stncntr))/dt > 1.e-4)
                 disp(sprintf('Instrument response dt mismatch %s',thisstn));
                 return;
              end
            end
          proc_win1=ifft(f1);
          proc_win1=real(proc_win1(1:proc_win_samps));
          proc_win1=taper10(real(proc_win1(1:proc_win_samps)));
          end
```

```
      if filt_flag==1
         proc_win1=filtfilt(B,A,proc_win1);
       end
     preproc_data(stncntr,nouter+1,:)=proc_win1;
     preproc_total(stncntr,nouter+1,:)=proc_win1;
     end
   end
    clear preproc_data rawtr1 ta1 proc_win1 henv fullwin1 f1 stncntr;
 end
eval(['save ', savepref, '.mat preproc_total stations juldayindex  stnpair_labels wn dt ts']);
 clear preproc_data rawtr1 ta1 proc_win1 henv fullwin1 f1 stncntr;
dbclose(db);
 quit
```

3. step_2.m

```matlab
%Step 2 of the ambient noise processing code, as used by Josh Boschelli. This can be run
by the bash script run_step_2.bash.
%Data is cross-correlated and saved in day by day step, doesn't require a connection to
antelope
%Modified from matnoise_v3_batch.m by Josh Calkins, and Geoff Abers
clear all
start_julday=getenv('MNBSJDAY');start_julday=str2num(start_julday);day=start_julday;
numdays=getenv('MNBNUMDAYS');numdays=str2num(numdays);
component=getenv('MNBCOMPONENT');
yrepochsec=getenv('YREPOCHSEC');yrepochsec=str2num(yrepochsec);
full_db_path=getenv('MNDB');
 if exist('start_julday')~=1;
    disp('FATAL ERROR: start_julday not defined.  Exiting');
    return
 end
 if exist('component')~=1;
    disp('FATAL ERROR: component not defined.  Exiting');
    return
 end
 if exist('numdays')~=1;
    disp('FATAL ERROR: numdays not defined.  Exiting');
    return
 end
 if exist('yrepochsec')~=1;
    disp('FATAL ERROR: yrepochsec not defined.  Exiting');
    return
 end
 if exist('full_db_path')~=1;
    disp('FATAL ERROR: full_db_path not defined.  Exiting');
    return
 end
ts= yrepochsec+(86400*(start_julday-1));
numdays=1;
kdayall=1; %index of days to loop over, relative to ts
dt=1; %data sample interval
hi_corner=0.30; %filter corners in Hz
lo_corner=0.003;
fullwin_len=86400; %length of window to preprocess
procwin_len=3600; %for outer time loop length
xwin_len=1200; %length of window to xcorr in seconds - inner time loop -lengthen for
long lags
npoles=5;   % strength of filters
filt_flag=1;
```

```
rminst_flag=1;     % remove instrument -- not necessary if whitening
onebit_flag=1;     % change signal to sign(signal) - non-preferred
envnorm_flag=1;
whiten_flag=1;
taper_flag=1;
ampscl_flag=1;
plot_flag=0;       %1 to plot results for every station pair – dangerous
wlev=1;
savepref='batchXC_TAstn';   % prefix for outputs
%%%%%%%%%%%%%%%%%
%End parameters to be set%
%%%%%%%%%%%%%%%%%%%
ts0=ts;
samprate=1/dt;
nyq=0.5/dt;
wn=[lo_corner/nyq, hi_corner/nyq];
[B,A]=butter(npoles,wn);
%Window & sample specifics
nproc_wins=floor(fullwin_len/procwin_len);
nxc_wins=floor(nproc_wins*procwin_len/xwin_len);
xc_data_samps=floor(samprate*xwin_len);
xc_length=2*xc_data_samps-1;
proc_win_samps=floor(procwin_len*samprate);
fflen=2^(nextpow2(procwin_len)+1);
%Strings for plot labels & filenames
procs='Processes: ';
if (rminst_flag==1), procs=[procs,'RMINST ']; end;
if (whiten_flag==1), procs=[procs,'WHITEN ']; end;
if (filt_flag==1), procs=[procs,'FILT ']; end;
if (onebit_flag==1), procs=[procs,'1BIT ']; end;
if (envnorm_flag==1), procs=[procs,'ENVNORM ']; end;
if (ampscl_flag==1), procs=[procs,'AMPSCL ']; end;
procs=sprintf('%s PW %d XW %d',procs,procwin_len, xwin_len);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%   Information Being Loaded In From step_1.m %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
load raw_Day_data
%%%%%%%%%%%%%%%%%%%%%%%
%%            XCROSS              %%
%%%%%%%%%%%%%%%%%%%%%%%
ind=1;
 for stnind1=1:length(stations)-1
   for stnind2=stnind1+1:length(stations)
    stnpair_labels(ind)=strcat(stations(stnind1),'-',stations(stnind2));
    [stnpair_dist(ind),
```

```matlab
stnpair_az(ind)]=distance(slats(stnind1),slons(stnind1),slats(stnind2),slons(stnind2));
    ind=ind+1;
  end
 end
 daycntr=0;
 xc_all_stack=zeros(length(stations)*(length(stations)-1)/2,xc_length);%follow n(n-1)/2
 for total station pairs
 xc_stack_cnt=zeros(1,length(stations)*(length(stations)-1)/2); %a count of the stacks for
 each stn pair
 counter= juldayindex;
 nouter = 24;
for kday=kdayall
  daycntr=1;
  ts=ts0+86400.*kday;
  julday=counter(daycntr);
  te=ts+fullwin_len;
%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Checks to see if each stations has data%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  if max(good_data_ind(:,:,daycntr))<1
   disp(sprintf('No good data found for any station on julday %d... SKIPPING', julday));
   continue;
  end
  stnpair_index=0;
  for stnind1=1:length(stations)-1
   for stnind2=stnind1+1:length(stations)
     stnpair_index=stnpair_index+1;
     if length(find(good_data_ind(stnind2,:)==1))==0 ||
length(find(good_data_ind(stnind1,:)==1))==0
          disp(sprintf('No good data found for julday %d stn1=%s and stn2=%s...
SKIPPING',...
          julday,char(stations(stnind1)),char(stations(stnind2))));
       continue;
     end
     disp(sprintf('working on DAY %d for stn1=%s and
stn2=%s',julday,char(stations(stnind1)),char(stations(stnind2))));
     for nouter1=1:nproc_wins %goes throught the hours of the day, 1:24
      if good_data_ind(stnind2,nouter1)==0 || good_data_ind(stnind1,nouter1)==0
         disp(sprintf('bad procwin for julday %d stn1=%s and stn2=%s procwin#%d...
SKIPPING',...
               julday,char(stations(stnind1)),char(stations(stnind2)),nouter));
continue;
      end
      ninner=0;
      for jj=1:nxc_wins/nproc_wins
```

```
xc_win1=preproc_total(stnind1,nouter1,1+(ninner*xwin_len):(ninner*xwin_len)+xwin_l
en);

xc_win2=preproc_total(stnind2,nouter1,1+(ninner*xwin_len):(ninner*xwin_len)+xwin_l
en);
        xc_win1=taper10(squeeze(xc_win1)); xc_win2=taper10(squeeze(xc_win2));
        this_xcorr=xcorr(xc_win1,xc_win2);
          if ampscl_flag==1
          this_xcorr=this_xcorr/max(abs(this_xcorr));
          end
          if sum(isnan(this_xcorr))==0
           xc_all_stack(stnpair_index,:)=xc_all_stack(stnpair_index,:)+this_xcorr';
           xc_stack_cnt(stnpair_index)=xc_stack_cnt(stnpair_index)+1;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Once the full hour of data has been added together over the 24 hours%
%           prints out the data for each station pair                  %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
          if jj==nxc_wins/nproc_wins&&nouter1==24
            filename =
sprintf('%s_%s_day_%d',char(stations(stnind1)),char(stations(stnind2)),julday);
            end
          else
        this_xcorr(isnan(this_xcorr))=0;
        end
         ninner=ninner+1;
       end %for jj=1:nxc_wins/nproc_wins
     end %for nouter1=1:nproc_wins
   end %for stnind2
 end %for stnind1
 clear preproc_data; %Avoid accidentally operating on the wrong data
 end %for kday=kdayall
 savepref='batchXC_TAstn'
 eval(['save ' savepref,'jd',num2str(day),'_',component,'.mat xc_all_stack xc_stack_cnt
stnpair_labels stnpair_dist stnpair_az wn dt ts']);
 quit
```

4. step_3.m

```
%step_3 of the noise processing. Takes daily results from runs of step_2.m and does
some post processing (sorting, SNR check), then saves out stacked data for the whole
range covered, as well as (optional) data subsets
tic
clear;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%      Loads in batch files from step_2.m      %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
days=0:1460;
for i=1:length(days)
    filename=sprintf('batchXC_TAstnjd%d_LHZ.mat',days(1,i));
    files{i}=filename;
end
%%%%%%%%%%%%%%%%%%%%%%%%%
%%      Parameters to be set      %%
%%%%%%%%%%%%%%%%%%%%%%%%%
min_vel=2.5; %Group velocities to determine windows for SNR processing, depending
on the station distances min_vel needs to be modified
max_vel=4.5;
outfile_prefix='Station_stack';
subset_flag=0;
single_station_flag=0; %if you wish to have day sets of each station pair be saved to used
later turn on
Acasual_flag=0; %if you wish to run the data with just the negative lag of the stacked
data
Casual_flag=0;  %if you wish to run the data with just the positive lag of the stacked data
Compare_flag=0;
flip_flag=0;
%%%%%%%%%%%%%%%%%%%%%
%end Params to be set      %
%%%%%%%%%%%%%%%%%%%%%
for ii=1:length(days)
eval(['load ',char(files(ii))]);
xc_all_stack(~isnan(xc_all_stack))=0;
save xc_all_stack xc_all_stack
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%Stacks the daily records and removes those %%
%%that are noisy.                              %%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for ii=1:length(days)
days(ii)
eval(['load ',char(files(ii))]);
```

```
samprate=1/dt;
xc_all_stack(isnan(xc_all_stack))=0;
midpt = (size(xc_all_stack,2)+1)/2;
            if days(ii)==days(1)
                full=load('xc_all_stack');


for k=1:size(full.xc_all_stack,1)
                    if sum(full.xc_all_stack(k,:))~=0


full.xc_all_stack(k,:)=full.xc_all_stack(k,:)./max(abs(full.xc_all_stack(k,:)));
                    end
                end
        xc_yr_stack=xc_all_stack;
        xc_yr_stack_cnt=xc_stack_cnt;
        xc_full_cnt = xc_stack_cnt;
        pos_cor = zeros(size(xc_all_stack,1),length(days));
        pos_cor(:,1)=0;
        neg_cor = pos_cor;
        continue
        end
    for k=1:size(xc_all_stack,1)
        if round(deg2km(stnpair_dist(k)))>1199
            stnpair_dist(k)=km2deg(1198);
        end
        if sum(xc_all_stack(k,:))~=0
        xc_all_stack(k,round(midpt):size(xc_all_stack,2)) =
xc_all_stack(k,round(midpt):size(xc_all_stack,2))./max(abs(xc_all_stack(k,round(midpt):
size(xc_all_stack,2))));
        xc_all_stack(k,1:round(midpt)) =
xc_all_stack(k,1:round(midpt))./max(abs(xc_all_stack(k,1:round(midpt))));
        end

nep=((corrcoef(xc_all_stack(k,midpt:round(midpt+deg2km(stnpair_dist(k)))),full.xc_all_
stack(k,midpt:round(midpt+deg2km(stnpair_dist(k)))))));
    pos_cor(k,ii)=nep(1,length(nep));
    nep=((corrcoef(xc_all_stack(k,round(midpt-
deg2km(stnpair_dist(k))):midpt),full.xc_all_stack(k,round(midpt-
deg2km(stnpair_dist(k))):midpt))));
    neg_cor(k,ii)=nep(1,length(nep));
        if pos_cor(k,ii)<.5
            xc_all_stack(k,1:round(midpt))=0;
        end
        if neg_cor(k,ii)<.5
            xc_all_stack(k,round(midpt):size(xc_all_stack,2))=0;
        end
```

```matlab
            if pos_cor(k,ii)<.5&neg_cor(k,ii)<.5
                xc_stack_cnt(k)=0;
            end
            if k==round(size(xc_all_stack,2)./2)
                disp('Half Done');
            end
        end

xc_yr_stack(:,round(midpt):size(xc_all_stack,2))=xc_yr_stack(:,round(midpt):size(xc_all_stack,2))+xc_all_stack(:,round(midpt):size(xc_all_stack,2));
xc_yr_stack(:,1:round(midpt))=xc_yr_stack(:,1:round(midpt))+xc_all_stack(:,1:round(midpt));
    xc_yr_stack_cnt=xc_yr_stack_cnt+xc_stack_cnt;
end
xc_all_stack=xc_yr_stack;
xc_stack_cnt=xc_yr_stack_cnt;
eval(['save ',outfile_prefix,'.mat stnpair_dist stnpair_az stnpair_labels xc_all_stack xc_stack_cnt']);
toc
```

5. mn_fitxcspec.m

```matlab
%%%%%% Spectral Measurement Method  %%%%%%%%%%%%
%%%%%% modified from mn_fitxcspec.m by Josh Calkins
clear all; close all;
setup
inpref='Station_stack';
outpref=strcat('fit19_',inpref);
adat=load([inpref '.mat']);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% FLAGS
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
plot_flag=0;
cull_flag=1;
comp_flag=0;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% REFERENCE CURVES
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
fref=[.01;.0115;.013;.015;.017;.02;.022;.025;.029;.033;.037;.04;.045;.05;.07;.1;.12;.14;.1
6;.2;.25;.33];
phvref=[4.25;4.155;4.15;4.1;4.08;4.06;4.05;3.98;3.91;3.83;3.76;3.72;3.64;3.57;3.43;3.34;
3.28;3.22;3.15;3.1;3.025;2.95];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% PARAMETERS TO BE SET
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A = exist('Station pair info');
if A == 0
mkdir('Station pair info')
end
fr_max=0.25;
order_ampwin=5;
SNRmin=20;
deltkm=deg2km(adat.stnpair_dist);
numprs=length(deltkm);
samprate = 1;
dt =1/samprate;
fnyq=0.5/dt;
df =1;
nyq=fnyq;
hi_corner=0.3;
lo_corner=0.003;
wn=[lo_corner/nyq, hi_corner/nyq];
hp_filt=wn(1)*fnyq;
lp_filt=wn(2)*fnyq;
farray=adat.farray;
```

```matlab
nf=length(farray);
df=0.02;
fkmin=0.025;
fkmax=fr_max;
        fknot=[1/40,1/34,1/30,1/27,1/25,1/22,1/20,1/19,1/18,1/17,1/16,1/15,1/14,1/13,1/2,
        1/11,1/10,1/9,1/8,1/7,1/6,1/5,1/4,1/3]';
fknot'
lp_corner=max(fknot);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% MATRIX STORAGE
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
m_fdatr=zeros(numprs,nf); %Real part of the spectrum (input)
m_ffit=zeros(numprs,nf); %Freqs at which data are fit
m_fzc=zeros(numprs,length(fknot)); %Zero crossings of the real spectrum
m_phv_zc=zeros(numprs,length(fknot)); %phase vels at observed fzc
m_pknot=zeros(numprs,length(fknot)); %interpolated slownesses 1/(phvel at fknots)
m_errpvel=zeros(numprs,length(fknot)); %error in pknot (only calc if do_linfit=1)
m_ppred=zeros(numprs,nf); %spline fit to pknots @ ffit
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Begin Loop over all stn pairs
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
good_fits=0;
for thispair=1:numprs
 fwrap=adat.fwrap(thispair,:);
 plabl=char(adat.stnpair_labels(thispair));
 if (plot_flag==1)
     disp(sprintf('-- Working On %d %s --',thispair,plabl));
 end
 krange=find(farray>hp_filt & farray<lp_filt)';
 nrg=length(krange);
 ffrg=reshape(farray(krange),nrg,1);
 fdatr=real(fwrap(1:nf));
 datrg=reshape(fdatr(krange),nrg,1);
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% FIND ZERO CROSSINGS & PRELIMINARY PHASE ESTIMATES
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 vtol=.025;
 vtol2=1;
 a=datrg(1:nrg-1).*datrg(2:nrg);
 kzc=find(a<0);
 nzc=length(kzc);
 if adat.xc_stack_cnt(thispair)<1
     disp(sprintf('WARNING - ZERO or ONE zero crossing(s) found for stnpair % s -
SKIPPING',char(adat.stnpair_labels(thispair))));
     adat.xc_all_stack(thispair,:)=0;
```

```matlab
        continue
  end
  if nzc<2
      disp(sprintf('WARNING - ZERO or ONE zero crossing(s) found for stnpair %s -
SKIPPING',char(adat.stnpair_labels(thispair))));
adat.xc_all_stack(thispair,:)=0;
             continue

end
  if nzc<2
      disp(sprintf('WARNING - ZERO or ONE zero crossing(s) found for stnpair %s -
SKIPPING',char(adat.stnpair_labels(thispair))));
adat.xc_all_stack(thispair,:)=0;
      continue
  end
  dsdf=(datrg(kzc+1)-datrg(kzc))./(ffrg(kzc+1)-ffrg(kzc));   % slope/sign of zero crossings
  fzc0=ffrg(kzc)-datrg(kzc)./dsdf;
  dsdf0=dsdf;
  if (dsdf(1)>0)
      disp(sprintf('WARNING-1 %d %s 1st zero has wrong sign',thispair,plabl));
      kzc=kzc(2:nzc);
      dsdf=dsdf(2:nzc);
      nzc=nzc-1;
  end
  fzc=ffrg(kzc)-datrg(kzc)./dsdf;
  noff=0;
  j0zeros=j0zeroseries((1:nzc)',noff);
  phv_zc=2.*pi.*fzc.*deltkm(thispair)./j0zeros;
  ph1pred=interp1(fref,phvref,fzc,'pchip');
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% CHECK PREDICTED PHASE VEL TO THOSE MEASURED   %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  if abs(ph1pred(1)-phv_zc(1))/ph1pred(1)>vtol
      pvc=2.*pi.*fzc.*deltkm(thispair)./j0zeros(1);
      kkk=find((ph1pred-pvc)./ph1pred<vtol & dsdf<0);
      if (ph1pred(1)-phv_zc(1)/ph1pred(1))>0
          disp(sprintf('WARNING %d %s SLOW phvel at 1st Zero Crossing; toss 1st
crossings',thispair, plabl));
      end
      if (ph1pred(1)-phv_zc(1)/ph1pred(1))<0
          disp(sprintf('WARNING %d %s FAST phvel at 1st Zero Crossing; toss 1st
crossings',thispair, plabl));
end
      if isempty(kkk)
          disp(sprintf('WARNING %d %s all observed zero crossings give
```

```
INACCURATE PhVel  -SKIPPING',thispair,plabl));
adat.xc_all_stack(thispair,:)=0;
continue
      end
      nzc1=kkk(1);   % hope this one works: throw out earlier ones
      dsdf=dsdf(nzc1:nzc);
      fzc=fzc(nzc1:nzc);
      nzc=nzc-nzc1+1;
      j0zeros=j0zeros(1:nzc);
      phv_zc=2.*pi.*fzc.*deltkm(thispair)./j0zeros;
  end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%% FIX K+1 MEASUREMENTS--FZC,PHV_ZC fzc %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 k=2;
 while (k<=nzc)
     php=interp1(fref,phvref,fzc(k),'pchip');
     fzcdiff=0;      % Change in spacing
     if (k>2)
          fzcdiff=(fzc(k)-fzc(k-1))./(fzc(k-1)-fzc(k-2))-1;
     end
     if ((php-phv_zc(k))/php>vtol && fzc(k)<max(fref))||(phv_zc(k-1)-
phv_zc(k))/phv_zc(k)>vtol2   % velocity too slow, skip this one
         if (plot_flag==1)
             disp(sprintf('Skipping %.2f at %.3f because << %.2f
or %.2f',phv_zc(k),fzc(k),php,phv_zc(k-1)));
         end
         if (k<nzc)
             fzc=[fzc(1:(k-1));fzc((k+2):nzc)];
         else
             fzc=fzc(1:(k-1));
         end
         nzc=nzc-2;
         if (nzc<2)
             disp(sprintf('WARNING  %d %s SKIPPING, could not fix slow
velocities',thispair,plabl));
     adat.xc_all_stack(thispair,:)=0;
             continue
         end
j0zeros=j0zeros(1:nzc);
         phv_zc=[phv_zc(1:(k-1)); 2.*pi.*fzc(k:nzc).*deltkm(thispair)./j0zeros(k:nzc)];
      elseif ((phv_zc(k)-php)/php>vtol && fzc(k)<max(fref) ) || abs(fzcdiff)>vtol
     pvc=2.*pi.*fzc(k).*deltkm(thispair)./j0zeroseries(k,(noff-4):2:(noff+4));
      dv=1./phv_zc(k-1)-1./pvc;
      kkk=find(abs(dv)==min(abs(dv)));
```

```
         if (kkk==3)   % same noff as before
            k=k+1;
         else
            noff=noff+(kkk-3)*2;
            j0zeros(k:nzc)=j0zeroseries(k:nzc,noff);
            phv_zc(k:nzc)=2.*pi.*fzc(k:nzc).*deltkm(thispair)./j0zeros(k:nzc);
            k=k+1;
         end
      else
         k=k+1;
      end

  end
  if nzc<2
     adat.xc_all_stack(thispair,:)=0;
    continue
  end
  %  Solve for piecewise Hermite cubic / Cardinal Spline (like pchip)
  %  interpolate on slowness not velocity
%Tension parameter can be -1 to 1, but doesn't seem to do much
% relatively high damping (in the range 0.1 - 1)  helps with the closer station pairs
  tens=0; damp=0.1; ireg=1;

  mn=find(fzc>=min(fknot));
   fzc=fzc(mn:length(fzc));
 if length(fzc) < 4
  adat.xc_all_stack(thispair,:)=0;
  continue
  end
nzc=length(fzc);
  phv_zc=phv_zc(mn:length(phv_zc));
  [pknot0, scov, chifit]=csplinefit2(fzc,1./phv_zc, fknot,tens,damp,ireg);
  ppknot=interp1(fzc,1./phv_zc,fknot,'pchip');
  phv0=1./pknot0;
  pknot_err=sqrt(diag(scov));
  kk=find(fknot>=fzc(nzc));   % deal with high-end instability
  if ~isempty(kk)
  phv0(kk)=phv_zc(nzc);
end   % keep const-vel past last ZC
  pknot0=1./phv0;
  nknot=length(fknot);
  kfit=find(farray>=fknot(1) & farray<=fknot(nknot));
  ffit=farray(kfit);
  pknot=pknot0;
  ppred=interp1(fknot,pknot,ffit,'pchip');
```

```matlab
errpvel=pknot_err.*phv0.*phv0;
x = interp1(1./fzc(2:2:length(fzc)),phv_zc(2:2:length(fzc)),1./fknot,'pchip');
xx = interp1(1./fzc(1:2:length(fzc)),phv_zc(1:2:length(fzc)),1./fknot,'pchip');
xxx = interp1(1./fref,phvref,1./fknot,'pchip');
errpvel=abs(x-xx)./2;
lambdaref=(interp1(fref,phvref,fknot,'linear'))./fknot; %reference wavelength
for i=1:length(fknot)
  if deltkm(thispair)<3*lambdaref(i)
    pknot(i)=NaN;
    errpvel(i)=NaN;
  end
end %for i=1:lenght(fknot)
%%%%%%%%%%%%%%%%%%%
%% Plot results -DEBUG ONLY%
%%%%%%%%%%%%%%%%%%%%%%%
if plot_flag==1
    mth=0;
    figure(2)
    clf
    zeroscan(fzc0,dsdf0,deltkm(thispair),mth);   % Probably move, eventually, more logic into this
    figure(1)
   hamp00=abs(hilbert(real(fwrap(1:nf))));
   p=polyfit(farray,hamp00,order_ampwin);
   hamp0=polyval(p,farray);
   dpred=besselj(0,2.*pi.*ffit.*deltkm(thispair).*ppred);
   penv=abs(hilbert(dpred));
   p2=polyfit(ffit,penv,order_ampwin);
   dpred=dpred.*spline(farray,hamp0,ffit)./polyval(p2,ffit);

   figure(1);
   clf
   subplot(211)
   plot(farray,fdatr,'g',ffit,dpred,'r');
   legend('Alldat','FitPred')

   xlim([0 fr_max]);
 xlabel('Hz')

   hold on;
   plot(fzc,zeros(size(fzc)),'k+')
   title(sprintf('%s(pr %.0f) Delta=%.0f km SNR=%.2f damp=%.2f tens=%.1f ireg=%.0f',...

char(adat.stnpair_labels(thispair)),thispair,deltkm(thispair),snr(thispair),damp,tens,ireg))
```

```
subplot(212)
plot(ffit,1./ppred,'-',fknot,1./pknot,'o',fknot,1./pknot0,'k+');
hold on;
plot(fzc, phv_zc,'k^','MarkerFaceColor','g','MarkerSize',8);
legend('fit','fit-knot','starting','from Zcross');
errorbar(fknot,1./pknot,errpvel,'.');
xlim([0 fr_max]);
ylim([1. 6]);
xlabel('Frequency, Hz')
ylabel('phase velocity, km/s');
title('Triangles: PhVel based on zero-crossings; error bars: from linear fit');
plot(fref,phvref,'r:')

figure(2);
hold on;
plot(ffit,ppred,'k');
figure(3)
clf
title(sprintf('Comparison of Post zero-crossings to Neg zero-crossings
for %s',adat.stnpair_labels{thispair}))
hold on
plot(1./fzc(2:2:length(fzc)),phv_zc(2:2:length(fzc)),'c^')
plot(1./fzc(1:2:length(fzc)),phv_zc(1:2:length(fzc)),'mo')
errorbar(1./fknot,1./pknot,errpvel,'r')
legend('- Zero','+ Zero','interp1','Location','SouthEast')
plot(1./fknot,x,'c')
plot(1./fknot,xx,'m')
xlim([0,40])
ylim([2.5,4.5])
xlabel('Period, sec')
ylabel('Phase Velocity, km/s')
filename=sprintf('Zero_cross_comp_%s',adat.stnpair_labels{thispair});
print(figure(3),filename,'-dpdf')
figure(4)

subplot(2,1,1)
plot([-1199:1199],adat.xc_all_stack(thispair,:))
title(sprintf('%s Time
Domain %f',adat.stnpair_labels{thispair},adat.xc_stack_cnt(thispair)./72))
xlabel('Time (second)')
ylabel('Amplitude')
errpvel
keyboard;

end %if plot_flag==1
```

```
if plot_flag==1
movefile('Zero_cross*','Station pair info')
end
  m_fdatr(thispair,1:length(fdatr))=fdatr;
  m_ffit(thispair,1:length(ffit))=ffit;
  m_fzc(thispair,1:length(fzc))=fzc;
  m_ppred(thispair,1:length(ppred))=ppred;
  m_pknot(thispair,1:length(pknot))=pknot;
   m_pknot(find(m_pknot==0))=NaN;
  m_phv_zc(thispair,1:length(phv_zc))=phv_zc;
  m_errpvel(thispair,1:length(errpvel))=errpvel;
   m_errpvel(find(m_errpvel==0))=NaN;
  good_fits=good_fits+1;
disp('Calculating SNR...')
  snr_cal
  snr_11(thispair,:)=snr_1;
  snr_22(thispair,:)=snr_2;
  snr_3(thispair,:)=snr;
end %for thispair=1:length numprs
for vv=1:size(adat.xc_all_stack,1)
me(vv)=sum(abs(adat.xc_all_stack(vv,:)));
end
k = find(me>1);
nips = size(adat.xc_all_stack,1);
disp(sprintf('Successfully fit %i spectra of %i attempts',good_fits, nips));
m_pknot_cp=m_pknot;
child=0;
m_pknot_cp=1./m_pknot_cp;
m_pknot_cp(isnan(m_pknot_cp))=0;
for men=1:size(m_pknot,1)
for women=1:size(m_pknot,2)
if m_pknot_cp(men,women)>0
child=child+1;
end
end
end
total_of=size(m_pknot,1)*size(m_pknot,2);
        disp(sprintf('Total number of measurments %i out of a

        possible %i',child,total_of));

%Clean up a few more filenames for output
stnpair_labels=adat.stnpair_labels;
stnpair_dist=adat.stnpair_dist;
stnpair_az=adat.stnpair_az;
```

```
wn=adat.wn;
samprate=adat.samprate;

%Get rid of outliers
if cull_flag==1
  sigmatol=2.5;
  for i=1:size(m_pknot,2)
    [junk,idx]=find(~isnan(m_pknot(:,i)'));
    mean_slow(i)=mean(m_pknot(idx,i));
    std_slow(i)=std(m_pknot(idx,i));
    m_pknot(find(m_pknot(:,i)>(mean_slow(i)+sigmatol*std_slow(i))),i)=NaN;
    m_pknot(find(m_pknot(:,i)<(mean_slow(i)-sigmatol*std_slow(i))),i)=NaN;
  end
end %if cull_flag=1
bn = size(m_errpvel,2);
xc_stack_cnt = adat.xc_stack_cnt;
load location
me = load('stnpair_labels');
for mm=1:length(stnpair_labels)
k = find(strcmp(me.stnpair_labels,stnpair_labels(mm)));
lat1(mm,1)=locations(k,1);
lat2(mm,1) = locations(k,3);
lon1(mm,1) = locations(k,2);
lon2(mm,1) = locations(k,4);
end
cd Station' pair info'/
dell=deg2km(stnpair_dist);
snr_2=snr_22;
snr_1=snr_11;
for vv=1:46665
dell(vv)=(dell(vv)./max(dell));
end
%%%%%%%%%%%%%%%%%%%%
%%%%%   Process one    %%%
%%%%%   SNR CUT OFF   %%
%%%%%%%%%%%%%%%%%%%%%%
count=0;
for vv=1:size(xc_all_stack,1)
        for yy=1:length(fknot)

if snr_3(vv,yy)<4
count=count+1;
m_pknot(vv,yy)=NaN;
m_errpvel(vv,yy)=NaN;
end
```

```
end
end
disp(sprintf(' %i measurements were removed in process one',count));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%    Process two    %%%%%%%%
%%%%%    ERROR CUT OFF  %%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%
count=0;
for mn=1:size(m_pknot,2)
hen=find(m_errpvel(:,mn)>.1);
m_pknot(hen,mn)=NaN;
m_errpvel(hen,mn)=NaN;
count=length(hen)+count;
end
disp(sprintf(' %i measurements were removed in process two',count));
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%    Process three     %%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
count = 0;
mk_invert_files=1;
for f_cntr=1:length(fknot)
been=find(isnan(1./m_pknot(:,f_cntr))~=1);
wen=find((1./m_pknot(been,f_cntr))>=(xxx(f_cntr)+xxx(f_cntr)*.025));
m_pknot(been(wen),f_cntr)=NaN;
m_errpvel(been(wen),f_cntr)=NaN;
ben=find((1./m_pknot(been,f_cntr))<=(xxx(f_cntr)-xxx(f_cntr)*.025));
m_pknot(been(ben),f_cntr)=NaN;
m_errpvel(been(ben),f_cntr)=NaN;
count=length(ben)+length(wen)+count;
end
disp(sprintf(' %i measurements were removed in process three',count));
disp('COMPARING RAY PATHS');
ray_compare
child=find(isnan(m_pknot)~=1);
total_of=size(m_pknot,1)*size(m_pknot,2);
disp(sprintf('Final total number of measurments %i out of a
possible %i',length(child),total_of));
if mk_invert_files==1
  for pair_cntr=1:numprs
   c_pair=char(adat.stnpair_labels(pair_cntr));
   st1(pair_cntr)={c_pair(1:(findstr(c_pair,'-'))-1)};
   st2(pair_cntr)={c_pair(findstr(c_pair,'-')+1:end)};
  end %for pair=1:length(numprs)
for f_cntr=1:length(fknot)
   outtextfile=sprintf('phvel_%4.3f',1./fknot(f_cntr));
```

```matlab
    fid = fopen(outtextfile,'w+');
    for pr_cntr=1:numprs
     if ~isnan(m_pknot(pr_cntr,f_cntr)) && ~isnan(m_errpvel(pr_cntr,f_cntr)) &&
m_errpvel(pr_cntr,f_cntr)<.1

fprintf(fid,' %s  %5.0f %5.2f %5.0f  %7.4f %7.4f  %7.4f %7.4f   %5.4f   %5.4f\n',stnpai
r_labels{pr_cntr},snr_3(pr_cntr,f_cntr),deltkm(pr_cntr),adat.xc_stack_cnt(pr_cntr)./72,lat
1(pr_cntr,1),lon1(pr_cntr,1),lat2(pr_cntr,1),lon2(pr_cntr,1),1/m_pknot(pr_cntr,f_cntr),m_
errpvel(pr_cntr,f_cntr));
      end %quality and distance checks
    end %for pr_cntr=1:numprs
    fclose(fid);
  end %for f_cntr=1:size(fknot)
end %if mk_invert_file=1
if comp_flag==1
disp('Time to compare the dispersion curves from the different methods')
Disp_com(plot_flag)
disp('Time to compare the domains for the different methods')
freq_sing(plot_flag)
end
cd ..
eval(['save ',outpref,'.mat samprate dt stnpair_az stnpair_dist stnpair_labels farray df
m_fdatr m_ffit m_fzc m_ppred fknot m_pknot m_phv_zc m_errpvel tens damp ireg snr
snr_1 snr_2 snr_3']);
```

6. snr_cal.m

```matlab
%%%%%%%%%%%%%%%%%%%%%%
%       SNR measurement          %
%%%%%%%%%%%%%%%%%%%%%%
%%      Parameters to be set        %%
%%%%%%%%%%%%%%%%%%%%%%
min_vel=2.5; %Group velocities to determine windows for SNR processing, depending
on the station distances min_vel needs to be modified
max_vel=4.5;
plot_flag =0;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%      Calculate SNR                                                        %
%Note that the tapering from matnoise means the first and last 60 seconds  %
% of the xcorr aren't indicative of the real noise, so I've moved the noise    %
% window in 100 seconds from the end of the trace                          %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
xc_all_stack = adat.xc_all_stack;
stnpair_az = adat.stnpair_az;
stnpair_dist = adat.stnpair_dist;
stnpair_azs = stnpair_az+180;
xc_length = size(xc_all_stack,2);
midpt =(size(xc_all_stack,2)+1)/2;
np1=floor((xc_length+1)/2);
dt = 1;
inner_win = floor(deg2km(stnpair_dist)/max_vel/dt);
out_win=floor(deg2km(stnpair_dist)/min_vel/dt);
periods = 1./fknot';
for vv =1:size(periods,2)
upper_corner=periods(vv)-2;
lower_corner=periods(vv)+2;

if upper_corner <= 2
upper_corner = 2.1;
end;
nyq=0.5/dt;
clear wn B A
wn=[1/lower_corner/nyq, 1/upper_corner/nyq];
[B,A]=butter(5,wn);
for bbb=1:size(xc_all_stack,1)
 xc_all_stack_3(bbb,:)=(xc_all_stack(bbb,:)+fliplr(xc_all_stack(bbb,:)))./2;
 xc_1=xc_all_stack(bbb,:);
 xc_2=xc_all_stack_3(bbb,:);
 if sum(xc_all_stack(bbb,:))>1
 xc_1=filtfilt(B,A,xc_all_stack(bbb,:));
```

```matlab
xc_2=filtfilt(B,A,xc_all_stack_3(bbb,:));
end
swin1=xc_1(1,(midpt+inner_win(bbb)):(midpt+out_win(bbb))); %signal window at pos
lag
swin3=xc_2(1,(midpt+inner_win(bbb)):(midpt+out_win(bbb))); %signal window at pos
lag
swin2=xc_1(1,(midpt-out_win(bbb)):(midpt-inner_win(bbb))); %signal window at neg

lag

swin4=xc_2(1,(midpt-out_win(bbb)):(midpt-inner_win(bbb))); %signal window at neg
lag
nwin1=xc_1(1,(midpt+out_win(bbb)):size(xc_all_stack,2));
nwin3=xc_2(1,(midpt+out_win(bbb)):size(xc_all_stack,2));
nwin2=xc_1(1,(1:midpt-out_win(bbb)));
nwin4=xc_2(1,(1:midpt-out_win(bbb)));
snr_1(bbb,vv)=sqrt((mean((abs(swin1))).^2))./sqrt((mean((abs(nwin1))).^2));
snr_2(bbb,vv)=sqrt((mean((abs(swin2))).^2))./sqrt((mean((abs(nwin2))).^2));
snr_3(bbb,vv)=sqrt((mean((abs(swin3))).^2))./sqrt((mean((abs(nwin3))).^2));
snr_4(bbb,vv)=sqrt((mean((abs(swin4))).^2))./sqrt((mean((abs(nwin4))).^2));
if plot_flag ==1
g = figure
meet = ones(1,length(swin1));
meet1 = (snr_1(bbb,vv)*meet)*max(abs(xc_1(bbb,:)));
meet2 = (snr_2(bbb,vv)*meet)*max(abs(xc_1(bbb,:)));
leet1 = [abs((midpt+inner_win(bbb))-midpt):abs((midpt+out_win(bbb))-midpt)];
leet2 = [(midpt-out_win(bbb)-midpt):(midpt-inner_win(bbb)-midpt)];
clf
subplot(2,1,1)
plot([-1199:1199],xc_all_stack(bbb,:))
xlim([-stnpair_dist(bbb)*111.2,stnpair_dist(bbb)*111.2])
title('Unfilter')
ylabel('Amplitude')
xlabel('Time (sec)')
subplot(2,1,2)
hold on
plot([-1199:1199],xc_1(1,:))
xlim([-stnpair_dist(bbb)*111.2,stnpair_dist(bbb)*111.2])
plot(leet1,meet1,'g','LineWidth',1)
plot(leet2,meet2,'g','LineWidth',1)
title(sprintf('%s neg SNR: %.2f pos
SNR: %.2f',stnpair_labels{bbb},snr_2(bbb,vv),snr_1(bbb,vv)))
ylabel('Amplitude')
xlabel('Time (sec)')
keyboard
```

```
end % plot flag
end % bbb
end % vv
```

7. spider.m

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Compares the phase velocity at every station at all periods%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
if exist('spec') ~= 0
m_pknot = spec.m_pknot;
m_errpvel = spec.m_errpvel;
end
if exist('FTAN') ~= 0
m_pknot = FTAN.m_pknot;
m_errpvel = FTAN.m_errpvel;
end
% Checks the phase velocity for each station at each period to find a station pair to that
station that is inconsistent with all the rest and sets it to NAN
a=[0.125:.125:2];
count=0;
m_pknot(find(m_pknot==0))=NaN;
m_pknot(find(1./m_pknot==0))=NaN;
cp = m_pknot;
m_errpvel_cp=m_errpvel;
m_pknot_cp=1./m_pknot; % s/km to km/s
m_pknot=1./m_pknot;
m_pknot_cp(isnan(m_pknot_cp))=0; % gets rid of nans for summing purposes
m_pknot(isnan(m_pknot))=0; % gets rid of nans for summing purposes
load bee
for stn = 1:306
clear vel_stack
for eb = 1:305
for j=1:size(m_pknot,2)
    vel_stack(eb,j)=m_pknot(bee(stn,eb),j);
end
end
for j=1:size(m_pknot,2)
    clear vel_stat
    vel_stat=unique(vel_stack(:,j),'rows'); % isolates the velocities
    if length(vel_stat)>2
    vel_stat=vel_stat(2:length(vel_stat),1); % gets rid of the inf due to 1/0
    end
    vel(j)=mean(vel_stat); % mean for n at period j
    vel_std(j)=std(vel_stat); % std for n at period
end
for j = 1:size(m_pknot,2)
for eb = 1:305
 if abs(vel_stack(eb,j)-vel(j))>=min(vel_std(j),0.05)
```

```matlab
if isnan(cp(bee(stn,eb),j))~=1
            count=count+1;
        end
            m_pknot_cp(bee(stn,eb),j)=NaN; % sets outliers to 0
                m_errpvel_cp(bee(stn,eb),j)=NaN;

end
end
end
end
m_errpvel=m_errpvel_cp;
m_pknot=1./m_pknot_cp;
m_pknot(find(m_pknot==0))=NaN;
m_pknot(find(1./m_pknot==0))=NaN;
tt=size(m_pknot,2)*size(m_pknot,1);
unt=(count/tt)*100;
disp(sprintf('%.2f percent were removed',unt));
if exist('spec') ~= 0
spec.m_pknot = m_pknot;
spec.m_errpvel= m_errpvel;
end
if exist('FTAN') ~= 0
FTAN.m_pknot = m_pknot;
FTAN.m_errpvel= m_errpvel;
end
```
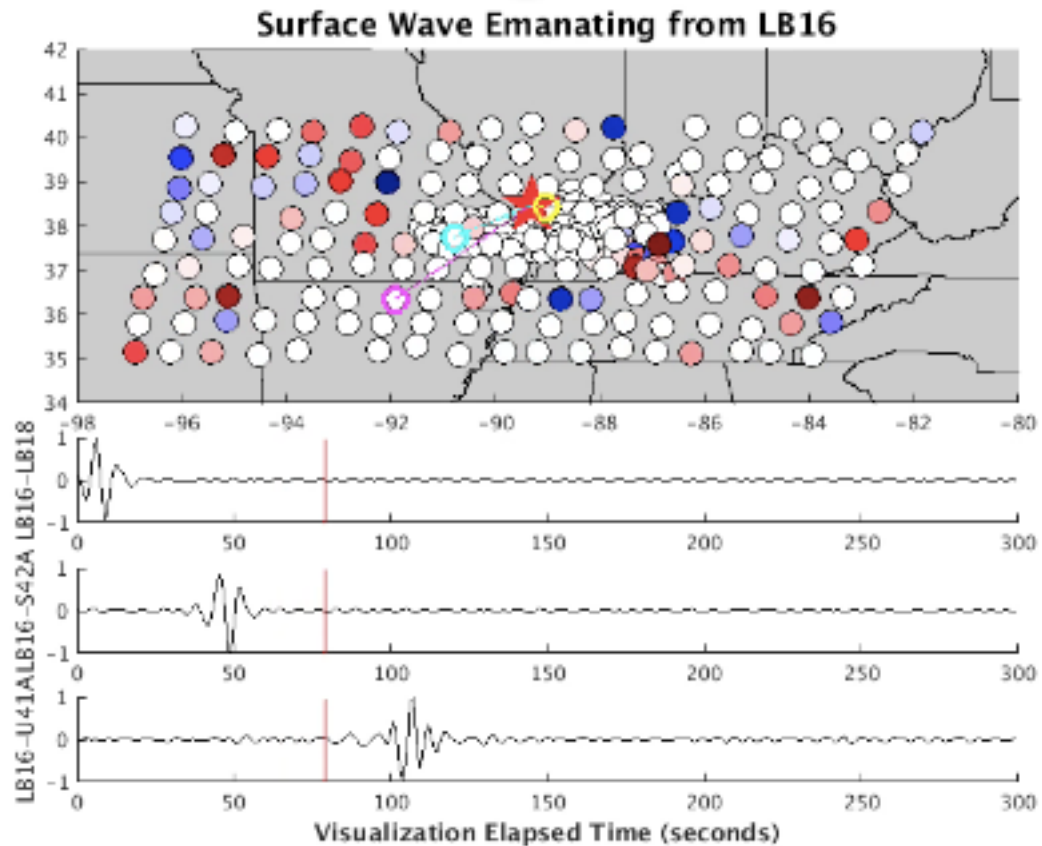
Figure A.1

Snap shot of a waveform animation that can be produced with movie_ses.m, where the correlated waveforms are treated as surface waves emanating from an adjustable epicenter at different period bands. The red star indicates the epicenter, or central station, with the three sub traces represented by the yellow, teal, and magenta stations. The color of the dots vary over time depending on the peaks and troughs of the waveforms. The waveforms in the amination demonstrate the frequency dependence of the SNR, with the wave peaks being visible at the shorter period lengths (< 20 s) and harder to distinguish at greater period lengths.

8. movie_ses.m

```matlab
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Makes animations showing the move out of the cross-correlations %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear all
close all
load movie_pkg  % Has station labels, locations and waveforms
wn=[1/40,1/5];
[B,A]=butter(2,wn./.5);
dots=zeros(1,length(stations_for_mov));
for i=1:141
dots(i)=50;
end
for i=142:length(stations_for_mov)
dots(i)=175;
end
check_station=0;
while check_station==0
prompt = 'Which Station Will Act As the Epicenter?\n';
EPICENTER=input(prompt,'s');
index=find(strcmp(stations_for_mov,EPICENTER));
if isempty(index)==0
check_station=1;
else
disp('Station Is Out Of Study Area, Please Choose Another Station.');
end
end
ind = 1;
for stnid=1:length(stations_for_mov)-1
for stnid2=stnid:length(stations_for_mov)
dot_size(ind,1)=dots(stnid);
dot_size(ind,2)=dots(stnid2);
ind=ind+1;
end
end
clear dots
ind = 1;
for stnind1=1:index
    stn_labels_1(ind)=strcat(stations_for_mov(stnind1),'-',EPICENTER);
    back_1(ind)=stn_labels_1(ind);
    ind=ind+1;
    stn_labels_1(ind)=strcat(EPICENTER,'-',stations_for_mov(stnind1));
    ind=ind+1;
end
```

```matlab
ind=1;
for stnind1=index+1:length(stations_for_mov)
   stn_labels_2(ind)=strcat(EPICENTER,'-',stations_for_mov(stnind1));
   back_2(ind)=stn_labels_2(ind);
   ind=ind+1;
   stn_labels_2(ind)=strcat(stations_for_mov(stnind1),'-',EPICENTER);
   ind=ind+1;
end
stn_labels_3=[stn_labels_1,stn_labels_2];
stn_label=unique(stn_labels_3);
ind=1;
for i=1:length(stn_label)
   k=find(strcmp(stnpair_labels,stn_label{i}));
   if isempty(k)~=1
   stn_label_4{ind}=stn_label{i};
   ind=ind+1;
   end
end
clear stn_label
stn_label=stn_label_4;
for t=1:length(stn_label)
k=find(strcmp(stnpair_labels,stn_label{t}));
   if abs(sum(xc_all_stack(k,:)))==0
   stn_label{t}='';
   else
   xc_all_stack(k,:)=filtfilt(B,A,xc_all_stack(k,:));
   end
end
stn_label_1=unique(stn_label);
clear stn_label
for t=2:length(stn_label_1)
stn_label{t-1}=stn_label_1{t};
end
rand_wave = randi([1,length(stn_label)],1,1);
for i=1:length(stn_label)
   pr_num(i)=find(strcmp(stnpair_labels,stn_label(i)));
end
kip =
find((deg2km(stnpair_dist(pr_num))<deg2km(stnpair_dist(pr_num(rand_wave)))/2)&(de
g2km(stnpair_dist(pr_num))>deg2km(stnpair_dist(pr_num(rand_wave)))/4));
temp=stn_label(kip);
rand_wave_1 = randi([1,length(temp)],1,1);
kip=find(strcmp(stn_label,temp(rand_wave_1)));
rand_wave_1=kip;
kip = find(deg2km(stnpair_dist(pr_num))<deg2km(stnpair_dist(pr_num(rand_wave)))/4);
```

```matlab
temp=stn_label(kip);
rand_wave_2 = randi([1,length(temp)],1,1);
kip=find(strcmp(stn_label,temp(rand_wave_2)));
rand_wave_2=kip;
for i=1:46665
xc_all_stack(i,:)=xc_all_stack(i,:)+fliplr(xc_all_stack(i,:));
end
shore    = shoreline_US;
political = political_US;
shore(:,2) = asinh(tan(shore(:,2)*pi/180.))* 180.0 / pi;
political(:,2) = asinh(tan(political(:,2)*pi/180.)) * 180.0 / pi;
political(:,2) = political(:,2);
shore(:,2) = shore(:,2)-2.5 ;
shore(:,1) = shore(:,1);
political(:,2) = political(:,2)-2.5;
for i=1:length(stn_label)
    pr_num(i)=find(strcmp(stnpair_labels,stn_label(i)));
end
for i=1:length(stn_labels_1)
        k=find(strcmp(stn_label,stn_labels_1{i}));

if isempty(k)~=1
    kk =find(strcmp(stnpair_labels,stn_label{k}));
    xc_all_stack(kk,:)=fliplr(xc_all_stack(kk,:));
    end
end
order = zeros(length(pr_num),2);
for i=1:length(pr_num)
    order(i,2) = stnpair_dist(pr_num(i))*111.2;
    order(i,1) = i;
end
sett = order;
for i=1:length(stn_label)
    kk =find(strcmp(stnpair_labels,stn_label{i}));
    xc_all_stack(kk,:)=xc_all_stack(kk,:)./max(abs((xc_all_stack(kk,:))));
end
for i=1:length(pr_num)
latt(i) = locations(pr_num(i),3);
lonn(i) = locations(pr_num(i),4);
end
lats = zeros(size(pr_num,2),1);
dots = zeros(size(pr_num,2),1);
lons = lats;

ind = 1;
```

```matlab
for i=1:length(stn_labels_1)
    k=find(strcmp(stn_label,stn_labels_1{i}));
    if isempty(k)~=1
    lats(ind) = locations(pr_num(k),1);
    lons(ind) = locations(pr_num(k),2);
    lo = find(strcmp(OIINK_OIINK_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
    dots(ind) = 50;
    end
    lo = find(strcmp(TA_TA_stnpair_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
    dots(ind) = 175;
    end
    lo = find(strcmp(OIINK_TA_stnpair_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
dots(ind) = 175;
    end
    ind = ind + 1;
    end
end
for i=1:length(stn_labels_2)
    k=find(strcmp(stn_label,stn_labels_2{i}));
    if isempty(k)~=1
    lats(ind) = locations(pr_num(k),3);
    lons(ind) = locations(pr_num(k),4);
    lo = find(strcmp(OIINK_OIINK_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
    dots(ind) = 50;
    end
    lo = find(strcmp(TA_TA_stnpair_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
    dots(ind) = 175;
    end
    lo = find(strcmp(OIINK_TA_stnpair_labels,stnpair_labels(pr_num(k))));
    if isempty(lo)~=1
    dots(ind) = 175;
    end
    ind = ind + 1;
    end
end
k=find(strcmp(back_1,stn_label{rand_wave}));
if isempty(k)~=1
EPIC_lat = locations(pr_num(rand_wave),3);
EPIC_lon = locations(pr_num(rand_wave),4);
RAND_lat = locations(pr_num(rand_wave),1);
```

```
RAND_lon = locations(pr_num(rand_wave),2);
RAND_dots = dots(rand_wave,1);
end
k=find(strcmp(back_1,stn_label{rand_wave_1}));
if isempty(k)~=1
RAND_lat_1 = locations(pr_num(rand_wave_1),1);
RAND_lon_1 = locations(pr_num(rand_wave_1),2);
RAND_dots_1 = dots(rand_wave,1);
end
k=find(strcmp(back_1,stn_label{rand_wave_2}));
if isempty(k)~=1
RAND_lat_2 = locations(pr_num(rand_wave_2),1);
RAND_lon_2 = locations(pr_num(rand_wave_2),2);
RAND_dots_2 = dots(rand_wave,1);
end
k=find(strcmp(back_2,stn_label{rand_wave}));
if isempty(k)~=1
EPIC_lat = locations(pr_num(rand_wave),1);
EPIC_lon = locations(pr_num(rand_wave),2);
RAND_lat = locations(pr_num(rand_wave),3);
RAND_lon = locations(pr_num(rand_wave),4);
RAND_lon = locations(pr_num(rand_wave),4);
RAND_dots = dots(rand_wave,1);
end
k=find(strcmp(back_2,stn_label{rand_wave_1}));
if isempty(k)~=1
RAND_lat_1 = locations(pr_num(rand_wave_1),3);
RAND_lon_1 = locations(pr_num(rand_wave_1),4);
RAND_dots_1 = dots(rand_wave,1);
end
k=find(strcmp(back_2,stn_label{rand_wave_2}));
if isempty(k)~=1
RAND_lat_2 = locations(pr_num(rand_wave_2),3);
RAND_lon_2 = locations(pr_num(rand_wave_2),4);
RAND_dots_2 = dots(rand_wave,1);
end
pr_cnt=pr_num;
tstart=(size(xc_all_stack,2)+1)/2;
tend=300+1200;
tstep=1;
 vidHeight = 600;
 vidWidth  = 800;
 figure('Color',[0.2 0.2 0.2],'Position',[1 200 vidWidth vidHeight],'Color','white')
 load dots
 colormap(cmap)
```

```
subplot(6,1,1:3)
set(gca,'Color',[0.8 0.8 0.8])
hold on
plot(shore(:,1),shore(:,2),'k')
plot(political(:,1),political(:,2),'k')
title(sprintf('Surface Wave Emanating from %s',EPICENTER),'FontSize',16)
xlim([-98,-80])
ylim([34,42])
subplot(6,1,4:4)
hold on
plot([0:tend-1200],xc_all_stack(pr_num(rand_wave_2),tstart:tend),'k')
xlim([0 tend-1200])
ylabel(sprintf('%s',stnpair_labels{pr_num(rand_wave_2)}))
subplot(6,1,5:5)
hold on
plot([0:tend-1200],xc_all_stack(pr_num(rand_wave_1),tstart:tend),'k')
xlim([0 tend-1200])
ylabel(sprintf('%s',stnpair_labels{pr_num(rand_wave_1)}))
subplot(6,1,6:6)
hold on
plot([0:tend-1200],xc_all_stack(pr_num(rand_wave),tstart:tend),'k')
xlim([0 tend-1200])
ylabel(sprintf('%s',stnpair_labels{pr_num(rand_wave)}))
xlabel(' Visualization Elapsed Time (seconds) ','FontSize',12,'FontWeight','bold')
set(gca,'CLim',[-1.0 1.0]);
%set(gca,'Box','on','XTick',[],'YTick',[],'DataAspectRatio',[1 1 1]);
 k = 1;
  for i=1:length(pr_num)
     amp(i,:) = xc_all_stack(pr_num(i),:);
  end
 while tstart < tend
       subplot(6,1,1:3)
      h1= scatter(lons,lats,135,amp(:,tstart),'filled','MarkerEdgeColor','k');
      h2= plot(EPIC_lon,EPIC_lat,'rp','MarkerSize',30,'MarkerFaceColor','r');
      h3= scatter(RAND_lon,RAND_lat,135,'m','LineWidth',3);
      h4= scatter(RAND_lon_1,RAND_lat_1,135,'c','LineWidth',3);
      h5= scatter(RAND_lon_2,RAND_lat_2,135,'y','LineWidth',3);
      h6= plot([EPIC_lon,RAND_lon],[EPIC_lat,RAND_lat],'m');
      h7= plot([EPIC_lon,RAND_lon_1],[EPIC_lat,RAND_lat_1],'c');
      h8= plot([EPIC_lon,RAND_lon_2],[EPIC_lat,RAND_lat_2],'y');
       subplot(6,1,4:4)
      h9= plot([tstart-1199,tstart-1199],[-1,1],'r');
       subplot(6,1,5:5)
      h10= plot([tstart-1199,tstart-1199],[-1,1],'r');
       subplot(6,1,6:6)
```

```matlab
        h11= plot([tstart-1199,tstart-1199],[-1,1],'r');
         pause(0.1);
        tstart = tstart + 1;
        F(k) = getframe(gcf);
        delete(h1,h2,h3,h4,h5,h6,h7,h8,h9,h10,h11)
        k = k + 1;

if tstart == 2400
    break
end
  end
movie2avi(F,sprintf('EPICENTER-%s-%.0f-%.0f.avi',EPICENTER,1./wn(2),1./wn(1)));
```