
PARTITIONING SIGNAL AND NOISE USING NON-LINEAR THRESHOLDING

Charles A. Langston

**Center for Earthquake Research and Information
University of Memphis
3876 Central Ave., Ste. 1
Memphis, TN 38152**

31 December 2020

Final Report

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED.



**AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
3550 Aberdeen Ave SE
AIR FORCE MATERIEL COMMAND
KIRTLAND AIR FORCE BASE, NM 87117-5776**

DTIC COPY

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by AFMC/PA and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RV-PS-TR-2021-0010 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNED//

Dr. Raymond J. Willemann
Program Manager/AFRL/RVB

//SIGNED//

Erin N. Pettyjohn, Chief
AFRL Geospace Technologies Division

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE (DD-MM-YYYY) 31-12-2020		2. REPORT TYPE Final Report		3. DATES COVERED (From - To) 25 Sep 2018 – 25 Sep 2020	
4. TITLE AND SUBTITLE Partitioning Signal and Noise Using Non-Linear Thresholding				5a. CONTRACT NUMBER FA9453-18-C-0064	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 62601F	
6. AUTHOR(S) Charles A. Langston				5d. PROJECT NUMBER 1010	
				5e. TASK NUMBER EF133213	
				5f. WORK UNIT NUMBER V1HA	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Center for Earthquake Research and University of Memphis 3876 Central Ave., Ste. 1 Memphis, TN 38152				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Space Vehicles Directorate 3550 Aberdeen Avenue SE Kirtland AFB, NM 87117-5776				10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RVBN	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RV-PS-TR-2021-0010	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. (AFRL-2021-0640 dtd 02 Mar 2021)					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT A comprehensive approach to phased array processing is developed around the use of continuous wavelet transforms that includes denoising, partitioning waveforms into component phases, scale-time gating of partitioned phases across the elements of a phased array, and beam forming in the wavelet domain. These methods were tested using an extensive waveform dataset from the 2016 IRIS Wavefields Community Experiment in northern Oklahoma that recorded natural seismicity and 12 explosions commissioned by AFTAC. Array analysis of explosion seismograms was used to develop a velocity model for near-surface structure under the array from surface wave dispersion inversion. High frequency Rayleigh wave data show multipathing and lateral refraction caused by lateral variations in the upper half km. Partitioning and beamforming earthquake waveforms allows the identification of near source surface reflections that constrain source depth for local seismicity. A comparison of standard frequency-wavenumber beam forming and CWT beam forming is made for explosions of the AFTAC experiment to show how CWT noise thresholding improves wave attribute determinations. The September 2017 North Korean nuclear test is analyzed at KSRS array using CWT techniques to show systematic changes in backazimuth of P and S regional phases. Empirical template correlation for detecting small events is theoretically extended to the CWT domain using empirical mother wavelets and is successfully tested using the AFTAC explosion dataset.					
15. SUBJECT TERMS signal analysis, continuous wavelet transform, denoising, seismogram decomposition, phased array processing					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Unlimited	18. NUMBER OF PAGES 86	19a. NAME OF RESPONSIBLE PERSON Dr. Raymond J. Willemann
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (include area code)

This page is intentionally left blank.

Table of Contents

1. Summary	1
2. Introduction.....	3
3. Technical Approach.....	5
3.1. Data.....	5
3.2. CWT Array Analysis	8
3.2.1 CWT Denoising	8
3.2.2 Signal Decomposition and Scale-Time Gating.....	10
3.2.3 Broad Band/Narrow Band Frequency-Wavenumber Analysis.....	14
3.2.4 Beam Forming in the Scale-Time Domain	17
3.3. Structure Modeling	20
3.3.1 Rayleigh Wave Phase and Group Velocity.....	20
3.3.2 Regional Velocity Structure.....	26
3.3.3 Verifying Dispersion Measurements	27
3.3.4 Inversion Method.....	30
3.4. An Earthquake Source Depth Estimate.....	33
3.5. Other Example Applications of CWT Array Processing.....	38
3.5.1 AFTAC Sources from Shotpoint 1	38
3.5.2 M6.3 North Korean Nuclear Test Recorded by KSRS	42
3.6 Using Empirical Source Templates as Mother Wavelets in the Continuous Wavelet Transform	48
3.6.1 The Empirical Wavelet	49
3.6.2 Data.....	51
3.6.3 Results.....	52
4. Discussion.....	57
5. Conclusions.....	63
References.....	65
List of Symbols, Abbreviations, and Acronyms.....	70

List of Figures

1. (a) Regional index map for Oklahoma and surrounding states. Location of the IRIS Wavefields experiment is shown by the star, cities by filled squares, and sources with filled circles. The focal mechanism of the M3.6 Langston, Oklahoma, earthquake shows largely strike-slip motion (Herrmann, 2016). (b) shows a smaller scale map of topography around the IRIS experiment, shot point 1 for the AFTAC explosion experiment, the location of the Kirk borehole which was logged for acoustic velocity, and stations of the IRIS YW array. Red dots denote locations of the 5Hz nodal instruments. The 6 tripartite clusters show the location of broadband instrumentation for the Golay 3x6 array and the dark cross denote the cross array stations used in this paper. The great circle path is drawn from the AFTAC shotpoint to the center of the cross array. The yellow line is the inferred position of the northern boundary of the Salt Fork River valley that may represent a horizontal velocity boundary. One possible ray path is drawn between the shotpoint and array that explains the back azimuth anomaly for 2Hz fundamental mode Rayleigh waves observed at the array. 6
2. Array geometry (left) and co-array diagram (right) of the cross array constructed from a subset of the IRIS experiment. Positions for stations 3016 and 3001 are shown on the array diagram. The co-array is a plot of all distances and azimuths between every pair of stations in the array and is an indicator of how well a seismic wavelength is sampled in space.....8
3. This is an example of scale-time gating the fundamental mode Rayleigh wave from denoised explosion vertical component seismograms of two stations of the cross array. Panel (a) shows the denoised seismogram (bottom) and its scalogram (top) after performing the Continuous Wavelet transform. A polygon is manually picked to isolate the fundamental mode energy on the scalogram at the reference station 3016. This portion of the scalogram is clipped out and then correlated with the scalograms of all other stations in the array. Panel (b) shows the shifted polygon encompassing the Rayleigh wave at station 3001. (c) shows the amplitude of the correlation vs time shift between the stations. The peak of the correlation occurs at an advanced time, relative to 3016, since station 3001 is closer to the source.....11
4. Panel (a) shows the scale-time gated Rayleigh waves obtained for stations of the cross array. Panel (b) is the broadband frequency-wavenumber result for the frequency band of 1 to 3 Hz encompassing the bandwidth of the Rayleigh waves. The small x's show the results of narrow band f-k for phase velocity as a function of frequency. Note that the bulk of the Rayleigh wave energy arrives approximately 15 degrees clockwise from the expected great circle path shown by the black line.....12
5. This figure illustrates two methods for scale-time gating the low scale/high frequency P wave train. The top panel (a) shows a choice of a block on the scalogram the encompasses the desired portion of the waveform. Panel (b) shows its complement where a block containing the high amplitude modes are chosen for correlation

instead. (c) displays the seismogram at the reference station 3016. P1 and P2 body wave phases are annotated.....13

6. Results of scale-time gating using the two different choices of CWT blocks shown in Figure 5. In (a), the P wave block in Figure 5a was used to correlate with all other station CWTs. However, results for some stations, shown by the delimited rectangular areas, were poor because the reference CWT block had higher correlation values with later arriving, high-amplitude modes. (b) This was alleviated by correlating with the modes block (Figure 5b) then removing this part of the CWT before inverse transforming.....14

7. Results of Rayleigh wave phase velocity dispersion analysis using narrow band f-k. (a) Slowness as a function of frequency. (b) Phase velocity as a function of frequency. (c) Back azimuth as a function of frequency. The heavy line at 311° is the theoretical great circle backazimuth. (d) Selected slowness spectra plots for the frequencies denoted by arrows. (e) Theoretical array responses for the corresponding plots shown in (d). Note that f-k spectra show the effects of Rayleigh wave multipathing for frequencies greater than 2.5Hz. f-k spectra display multiple peaks at different backazimuths and slownesses.....15

8. f-k spectra of P1 (left) and P2 (right) phases in the vertical components of the 2000lb explosion. Note the narrow beams showing that phases are highly correlated across the array.....16

9. CWT beamforming result for the P wave block of the 2000lb AFTAC explosion. The low scale wavetrain was determined from excluding the modes block (Figures 5b, 6b). Additionally, a scale bandpass of 0.1 to 0.5s was performed and is shown in (a). P1, P2, S, and Coda phases are annotated. (b) and (c) show the inferred phase velocity and backazimuth, respectively, at every 10^{th} time point in the trace. Red symbols denote well-formed single CWT beams with $R > 0.5$ and blue for $R < 0.5$. Representative slowness beams are indicated for the times shown by vertical lines...19

10. Array group velocity analysis example for 1.375Hz vertical component Rayleigh waves. Group arrival times are plotted against station distance from the source assuming the great circle path to the source in (a). Because the waves are arriving from an azimuth clockwise from the great circle backazimuth, times from the two linear arms of the cross array are inconsistent. In (b), an inversion for group slowness and backazimuth is performed to find the optimum backazimuth and group velocity for structure under the array. The data clearly indicate that Rayleigh waves arrive 22° clockwise from the great circle backazimuth, similar to the amount seen in the phase velocity determination.....23

11. (a) compares the group velocity determined by array analysis (filled circles) with the whole path group velocity determined by dividing the great circle distance by the group arrival time at station 3016 (filled squares). There seems to be no significant

difference in group velocities determine by two independent means. (b) shows the backazimuth determined from the array group velocity analysis as a function of frequency. Multipathing causes the larger errors at frequencies greater than 2.5 Hz.24

12. Waveforms and CWT scalograms for 2000 lb events at shot point 1 (a) and shot point 2 (b) recorded at station 3016 in the IRIS nodal array. Note the similarity in vertical waveform between the two shots although the time scale is 50% shorter for (b) compared to (a). Scalograms for both shots show distinct, well-developed regions of the CWT separating fundamental and higher mode Rayleigh waves.....25
13. Velocity model derived from Sarah Kirk Well sonic log contracted by the Mobil Oil Corporation, 1985.....27
14. Phase velocity dispersion for the fundamental mode and higher mode wave trains for Rayleigh waves generated by the 2000lb shot at shot point 1.....28
15. Profile of all vertical data from the 2000lb explosion at shot point 2. Major phase types are annotated. Note the high quality of waveform clearly showing moveout of all major phases. The data were instrument corrected but no filtering or noise thresholding were applied to remove the small amount of low frequency noise.....29
16. 1D f-k response of all vertical (a) and radial (b) components of explosion 1 at shot point 2. The white curves are theoretical dispersion curves derived from a velocity model determined from inversion of the fundamental mode data.....29
17. Inversion of the fundamental mode phase velocity data from the explosion at shot point 1. (a) A velocity structure is characterized by discrete nodes in shear wave velocity and depth. A linear gradient is assumed between nodes, except at layer boundaries. Error bars in (a) are estimated from the model covariance matrix, equation (35). (b) The theoretical dispersion curve fit is excellent.....32
18. The left panel displays the inversion result using the phase velocity data (middle panel) from the shot point 2 explosion. The right panel shows the change in the inversion least squares error with iteration.....32
19. Waveforms recorded at station 3016 for the M3.6 6/28/2016 Langston, Oklahoma, earthquake. Vertical, radial, and transverse component waveforms are shown. The data have been corrected for instrument response. P, PL, Rayleigh, and Love waves are annotated.....34
20. Scale-time gates chosen for the vertical component of the M3.6 earthquake. The top panel shows the scalogram for the seismogram shown below. Scale-time gate polygons were chosen for the low scale portion of the P wave and harmonic portion of the PL phase. S, Rg, and Coda phases are also annotated.....35

21. Results from scale-time gating the vertical component of station 3016 using the P and PL wave gates shown in Figure 20. The original data is shown in the top panel and the gated waveforms shown below.....36
22. (a) shows the gated P wave at the 81 stations of the cross array. The broadband f-k beam for 5 to 20Hz is shown in (b) and indicates that the two major phases, denoted by P and pP, have identical slownesses. (c) shows the instrument-corrected SH waves at the cross array with (c) the 1 to 10Hz broadband f-k reponse. S and sS phases are annotated on the seismograms.....37
23. a) Data recorded by the cross array for shotpoint 1-2 have been denoised using ECDF soft thresholding with a scale band reject for scales greater than 1s. (This is used for all denoising in this section). The data show artefacts from the denoising process within waveforms, such as the arrival near 15s at the farthest station. The Rayleigh wave after 30s arrival time is only partially recovered. The red line denotes a group velocity of 3km from the source. (b) shows the BBfk result for the 1 – 15Hz bandpass of the denoised data. Surface waves dominate the lower parts of the frequency band. (c) shows the BBfk result for th 5 – 15Hz bandpass. P and S wave trains are clearly imaged.....40
24. (a) Data from shotpoint 1-2 that have not been denoised. A 2-pole, phaseless butterworth filter for the bandpass of 1 – 15Hz has been applied to the data as well as a 5% Hanning taper to both ends of the time series. (b) shows the resulting BBfk response. The ambient noise field created by a power wind farm south of the array shows up as low velocity surface wave energy from southern directions. (c) shows the original data filtered in the 5 to 15Hz bandpass, with (d) being the BBfk response....41
25. (a) Denoised data at the cross array for the 250lb explosion at shotpoint1-3. The denoised seismograms show numerous high scale artefacts but are dominated by high velocity P and S phases from a small local earthquake. (b) shows the denoised data filtered between 15 and 30Hz. (c) BBfk of the filtered data. The P and S phase characteristics are highly focused.....43
26. (a) Data recorded at the cross array for the time of explosion 1-4. The data have been bandpassed filtered between 5 and 30Hz. There is very little evidence of the AFTAC explosion. Local noise source from north and south of the array dominate the seismograms, as seen by their moveout in time. (b) BBfk of the filtered data of (a). (c) Data after soft threshold denoising. The larger impulsive sources are much clearer although the smaller amplitude sources have been subdued by denoising. (d) BBfk of the denoised signals appear somewhat clearer than (b) and suggest sources from the south associated with the wind farm.....44
27. (a) Geometry of the CTBTO KSRS array located near Wonju, South Korea. The co-array diagram (b) shows good azimuthal performance for horizontal wavelengths from 2 to 20km. The co-array is a plot of all possible distances and azimuths between array elements showing how the array can sample a wavelength in space.....45

28. Theoretical broadband f-k responses for the KSRS array for frequency bands of 1-10Hz (a), 0.2-1Hz (b), and 0.05-0.2Hz (c).....46

29. CWT beam forming for the 3 September 2017 M6.3 nuclear test. The middle panels show the reference seismogram (d) at station KS01 near the center of KSRS. The time series below the seismogram show phase velocity (e) and backazimuth (f) as a function of time. CWT beamforms are indicated by the arrows to particular times. Backazimuth and phase velocity measurements are sifted into $R \geq 0.5$ (red) and $R < 0.5$ (blue) values. Low “R” values correspond to beamforms that have distributed energy across the slowness plane as in (h) and (i). Well focused beams generally have a single peak as in the other plots. Note that the P wave train (b) and (g) and the beginning of the S (c) have backazimuths close to the great circle path which is N12°E. Later S arrivals generally have backazimuths of roughly N10°W. The R metric breaks down where the time series is near zero as in (a).....47

30. Real (left) and negative Imaginary (right) parts of the empirical wavelet constructed from the 2000lb shot at shotpoint 1 for vertical component data recorded at station 3016 in the cross array. The imaginary part is derived from the real part through Hilbert transformation according to equation (40) and is, by construction, 90° out of phase from the real part. Waveform scaling for $a=0.5, 1.0, \text{ and } 1.5\text{s}$ are shown. Waveforms stretch with increasing scale.....51

31. Comparison of the instrument corrected vertical component data at station 3016 (left panel) and hard thresholding denoised data (right panel) for the explosions from shotpoint 2. SP21, SP22, SP23, and SP24 are 2000lbs, 500lbs, 250lbs, and 250lbs of explosives, respectively. The waveforms are effectively identical.....52

32. Results for shotpoint 1, explosion 1 (2000lb explosion). In this and next 3 figures, the data time series is shown as the top seismogram as in (a) and (c). The real part of the template mother wavelet for scale $a=1\text{s}$ is shown below the data. The right panels (b) and (d) show a 3D perspective view of the scale-time plane (scalogram). The black “x” shows the position of the peak value. (a) and (b) show the result of taking the CWT of the original template. (c) and (d) shows the result for the original noisy data, corrected for instrument response, for the 2000lb explosion.....53

33. CWT of explosion 2 (500lbs) using the explosion 1 mother wavelet. Same scheme as Figure 32.....54

34. Empirical wavelet CWT for explosions 3 (a) and (b), and 4 (c) and (d) at shotpoint 1. Same scheme as Figure 32.....55

35. Same scheme as Figure 3 showing CWT results for explosion 1 recorded at station 3001 ((a) and (b)) and station 1001 ((c) and (d)).....56

- 36. Shot point 1 fundamental and first higher mode phase velocity dispersion data with predictions from the fundamental mode inversion model. The higher mode (upper curve) is clearly too fast over most of the frequency band.....60

- 37. Frechet derivatives for the velocity model shown in the left panel for 1.5 Hz (middle) and 3.5 Hz (right). The fundamental mode derivatives are shown by black solid lines, the first higher mode by the red dashed line, and the second higher mode by the blue dashed line.....61

- 38. The body wave portion of the radon transform of the data displayed in Figure 15 for the 2000lb shot at shot point 2. The transform for the vertical components is dominated by waves in the P wave train. Various phases are annotated within the P Wave Group polygon shown in the left panel. Transform of the radial components (right panel) shows the S arrival (S1) clearly. Note that S1 is not very clear in the vertical component radon transform.....62

This page is intentionally left blank.

1. SUMMARY

Non-linear block thresholding of the continuous wavelet transform (CWT) of 2D phased array signals offers high time-resolution solutions for analyzing seismograms of local and regional seismic events. An initial denoising step on an array ensemble reveals those regions of the scale-time plane that contain high signal-to-noise arrivals. Individual seismic phase arrivals in ensemble, denoised seismograms can be partitioned using scale-time gating where CWT wave packets of an individual seismic phase on the scale-time plane for a reference array element are time-correlated with all other elements to find an optimum time shift for the phase across the array. The seismic phase is then clipped out of the CWT of each array element for further analysis using this optimal time shift.

The seismogram can be separated into component seismic waves for a detailed view of wave characteristics such as slowness and arrival back azimuth using conventional frequency-wavenumber methods. However, the process can be taken further by using the CWT of each phase to construct high time-resolution signal beams over CWT scale. Local explosion data from the 2016 IRIS Wavefields community experiment in northern Oklahoma are used to demonstrate these techniques in separating surface wave modes and body waves and examining the scattering of regional phases. High resolution CWT processing in the 1 to 3 Hz band for northern Oklahoma reveals horizontal Rayleigh wave refraction and multipathing. Rayleigh wave phase velocity dispersion measurements are used to construct a 1D velocity model under the IRIS experiment. Scale-time gating helps expose near-source surface reflections from a local M3.6 earthquake that are used to verify the 4km source depth obtained from a published regional moment tensor solution.

An empirical mother wavelet can be constructed from the time series of a template seismic event. Using the entire signal of a template event as a wavelet within the continuous wavelet transform produces the ultimate sparse representation for repeating events since, in principle, the detected event can be represented by a single wavelet coefficient in the scale-time domain. Use of empirical mother wavelets generalizes template correlation. The time scaling of the empirical wavelet mimics changes in relative arrival times between phases. A template derived from an event at one seismic

station might be appropriate for a close by station or for other events of similar mechanism close to the template event. This hypothesis is tested using local explosion data recorded by the 2016 IRIS Wavefields Community Experiment. Use of an empirical wavelet derived from the 2000lb shot approximately 35km from the IRIS experiment clearly detects the same explosion from noisy data at other stations in the array and for the 500lb explosion at the same shotpoint. The two 250lb explosions were not detected because of relatively higher levels of background noise. These results suggest the additional possibility of reconstructing waveforms from small events buried in the noise using the template event as the mother wavelet.

The application of CWT techniques for “de-signaling” in ambient noise correlation studies was investigated by Yang et al. (2020). Here, transients consisting of earthquakes and non-stationary noise sources were removed from daily seismograms by using thresholding before ambient noise correlation to obtain empirical Green’s functions. Resulting Green’s functions were shown to be higher quality than using previous time domain clipping techniques. An additional denoising step for the inferred Green’s functions greatly improved the final waveforms.

Curvelets are spatial wavelets that have directional properties. The use of curvelets in denoising of dense array data was investigated using data from the Long Beach seismic experiment. In the case of a dense array deployment where the wavefield is sampled finely in space, curvelets can be used to denoise over space for each time point in much the same way as thresholding in the scale-time domain using the CWT. In addition, curvelets allow for a decomposition of the wavefield based on images of the wavefield geometry. Scattering of teleseismic P into local high frequency Rayleigh waves was investigated for Long Beach by Zhang and Langston (2020). Results showed that structures associated with the Newport-Inglewood fault zone produced clear P-to-Rayleigh conversions that had both circular and planar wave fronts. Curvelets allowed the separation of the two kinds of waves.

This final report furnishes details of unpublished work performed in the second year of research on this 2-year contract. A rigorous denoising technique was developed in the first year and can be found in Langston and Mousavi (2019). This denoising method

forms the basis for initial denoising in the CWT array analyses described in the remainder of this report.

2. INTRODUCTION

The analysis of seismic signals using phased arrays of seismometers is the basis of much of nuclear verification seismology where detection, location, and discrimination of earthquakes from explosions are primary goals (Filson, 1975; Douglas, 2002; Rost and Thomas, 2002; Havskov and Alguacil, 2004). Numerous methods and array designs have been developed primarily for analyzing the principal wave attributes of frequency, horizontal phase velocity, and signal back azimuth for both source and structure studies (Capon et al., 1967; Haubrich, 1968; Capon, 1969; 1970; Claassen, 1985; Mykkeltveit et al., 1983; Mykkeltveit, 1985; Nawab et al., 1985; Abrahamson et al., 1987; Goldstein and Archuleta, 1987; Kvaerna, 1989; Gupta et al., 1990; Kvaerna and Ringdahl, 1992; Wagner and Owens, 1993; Kushnir, 1996; Lindfors, 1996; Stump et al., 2004; Harris and Kvaerna, 2010; Gibbons et al., 2011; Gibbons, 2014). An important attribute of phased array beam forming is the ability of the array to improve the signal-to-noise ratio (SNR) of incident signals by shifting, then stacking in the time or frequency domains at the appropriate azimuth and vector slowness of the incident wave. The improvement in SNR is often estimated by assuming a normally distributed background noise field so that noise is reduced by \sqrt{N} , where N is the number of sensors in the array (Capon et al., 1967; Aki and Richards, 1980). Improving the SNR improves the chances of event detection. However, wave coherence is a major issue in phased array signal processing and is a function of the wave scattering properties of Earth structure local to the array (e.g., Toksöz et al., 1991; Al-Shukri et al., 1995). Although there are many ideal array designs that can theoretically focus individual seismic phases, array performance is always degraded by noise and waveform decorrelation. Array design is almost always an issue as well, since field conditions may preclude optimum geometries or numbers of deployed sensors degrading the signal beam through creation of side lobes in wavenumber space making it difficult to separate interfering seismic phases.

In this report I extend the concepts introduced in Langston and Mousavi (2019), referred to as “L&M” in the remaining part of the report, for using the continuous

wavelet transform (CWT) to remove noise and partition waveforms into individual seismic phases to improve beam forming with a phased array. The initial step of denoising an array dataset using non-linear thresholding can improve the SNR by an order of magnitude or more before any array analysis is attempted. “Scale-time gating” will be introduced to separate target seismic phases over the array in the wavelet domain for individual frequency-wavenumber analysis. I then suggest an additional step of beam forming within the CWT domain that produces array beams at the sampling rate of the original time series. A relatively simple measure of beam coherence is also used to ascertain what parts of the seismic wavefield are coherent across the array and which aren’t to serve as an indication of the quality of the phase interpretation.

These methods will be demonstrated using local explosion data recorded by the 2016 IRIS Wavefields experiment in northern Oklahoma (Sweet et al., 2018). The array analysis will make it possible to investigate high frequency Rayleigh wave multipathing and the composition of body waves that propagate in the local waveguide of Paleozoic sediments. Although I apply the methods to local wave propagation and a high frequency seismic array, they can be used for any phased array.

Correlation techniques have found widespread use in many different problems in exploration and regional seismology (Anstey, 1966) including source detection and location (Gibbons and Ringdahl, 2004; Schaff and Richards, 2004; Schaff and Waldhauser, 2010; Withers et al., 1999, Zhang and Wen, 2015), focal mechanism studies (Shelly et al., 2016), and tomography (VanDecar and Crosson, 1990; Zhang and Thurber, 2003). Relying on the remarkable observation that many events recorded at the same station show the same waveform because of similar locations and wave propagation characteristics (Waldhauser and Ellsworth, 2000, among others), correlation of a template time series from a known, high signal-to-noise (SNR) event with low SNR seismograms at the same station can uncover additional small events well within the noise. I will incorporate the template correlation method into the CWT framework to propose a novel method that may be useful for recovering the waveforms of small events from the seismic noise field.

3. TECHNICAL APPROACH

3.1. Data

The 2016 IRIS Wavefields seismic experiment was designed by an interested community group to collect densely sampled array data in an area of active induced seismicity in northern Oklahoma (Sweet et al., 2018) (Figure 1). 247 three-component, 5 Hz Fairfield Nodal Zland nodal seismometers were deployed in two NS 5 km lines and one EW 13 km line. 112 of these seismometers were deployed in an 800x800m, 7 level, nested gradiometer configuration. 18 broadband seismometers were deployed as a “Golay 3x6” regional array with an approximate aperture of 6km. 9 infrasound instruments were sited with 9 of the broadbands to help in discriminating acoustic signals that might be recorded seismically. Data from the experiment are stored at the IRIS Data Management Center under the network code YW for 2016.

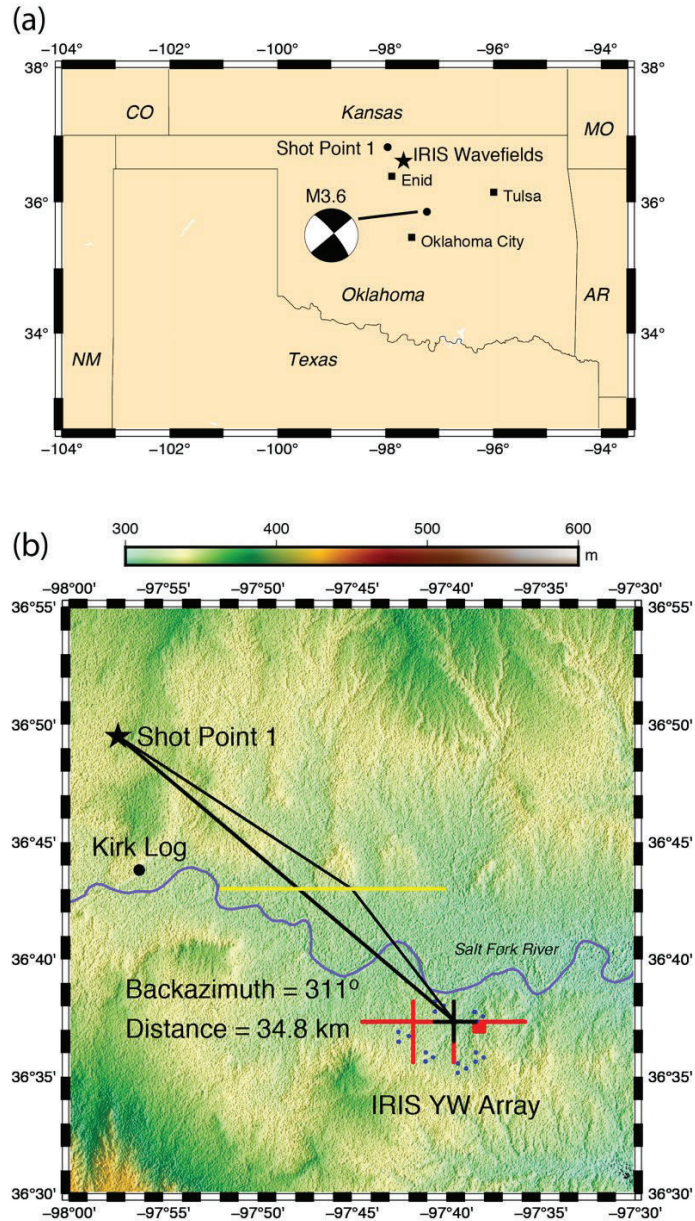


Figure 1– (a) Regional index map for Oklahoma and surrounding states. Location of the IRIS Wavefields experiment is shown by the star, cities by filled squares, and sources with filled circles. The focal mechanism of the M3.6 Langston, Oklahoma, earthquake shows largely strike-slip motion (Herrmann, 2016). (b) shows a smaller scale map of topography around the IRIS experiment, shot point 1 for the AFTAC explosion experiment, the location of the Kirk borehole which was logged for acoustic velocity, and stations of the IRIS YW array. Red dots denote locations of the 5Hz nodal instruments. The 6 tripartite clusters show the location of broadband instrumentation for the Goley 3x6 array and the dark cross denote the cross array stations used in this paper. The great circle path is drawn from the AFTAC shotpoint to the center of the cross array. The yellow line is the inferred position of the northern boundary of the Salt Fork River valley that may represent a horizontal velocity boundary. One possible ray path is drawn between the shotpoint and array that explains the back azimuth anomaly for 2Hz fundamental mode Rayleigh waves observed at the array.

The nodal seismometers were deployed over the time interval of 22 June to 27 July and passively recorded hundreds of local earthquakes, teleseisms, and other regional events. The broad band Golay array was kept in place until late November to catch aftershocks from the M5.8 Pawnee earthquake that occurred in September just before the originally planned decommissioning of the experiment.

On 14 July and 16 July 2016 the Air Force Technical Applications Command (AFTAC) commissioned an explosion seismic experiment performed by the IRIS Source Facility at the University of Texas, El Paso (Cleat Zeiler, personal communication 2017; Table 1). The experiment consisted of 4 borehole explosions of varying sizes set and detonated at 3 different azimuths and distances from the array. I use the data from the 2000 lb explosion detonated approximately 35 km to the northwest which is the same event as used in L&M (Figure 1).

Table 1 – Shot Information for the AFRL Experiment (courtesy of Cleat Zeiler, 2017)

Shotpoint #	Latitude	Longitude	Elev. (m)	Date	Time	Nominal Charge Size (lb.)	Comments
11	36.825026	-97.956602	343.8	14-Jul	14:26:00	2000	
12	36.824767	-97.956618	341.4	14-Jul	14:15:00	500	
13	36.824747	-97.956251	341.9	14-Jul	14:05:00	250	
14	36.825053	-97.956182	343.1	14-Jul	3:26:00	250	shotgunned
21	36.551591	-97.517971	304.6	16-Jul	0:37:00	2000	
22	36.551312	-97.517956	303.2	16-Jul	0:25:00	500	
23	36.551303	-97.517588	305.1	16-Jul	0:14:00	250	
24	36.551302	-97.517242	304.1	16-Jul	0:04:00	250	
31	36.113566	-97.248696	300.3	14-Jul	15:12:00	2000	
32	36.11365	-97.248307	300.1	14-Jul	15:02:00	500	
33	36.113896	-97.248479	299.8	14-Jul	14:50:00	250	
34	36.113816	-97.248777	301.3	14-Jul	3:40:00	250	geysered

A cross-shaped array was subset from the YW array from the stations located at the eastern crossing of the EW and NS arms of the linear deployments. This subarray has an aperture of 3 km and densely samples wave lengths of 3 to 3.5 km as demonstrated by its co-array diagram (Figure 2). This array was chosen to investigate the wave characteristics of the high frequency explosion waveforms in the band of 1 to 10 Hz.

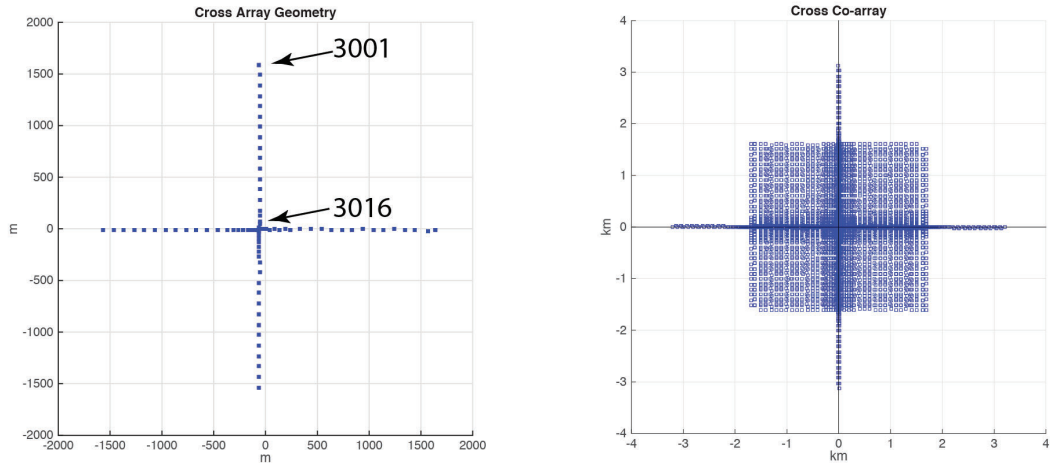


Figure 2 – Array geometry (left) and co-array diagram (right) of the cross array constructed from a subset of the IRIS experiment. Positions for stations 3016 and 3001 are shown on the array diagram. The co-array is a plot of all distances and azimuths between every pair of stations in the array and is an indicator of how well a seismic wavelength is sampled in space.

The raw waveform data were recorded at 250 sps. The only initial processing applied to the data was to remove the nominal instrument response of the 5 Hz nodal instruments and applying a trapezoidal zero-phase filter in the band 0.05 to 80 Hz with corner frequencies of 0.5 and 50 Hz.

3.2. CWT Array Analysis

3.2.1 CWT Denoising

L&M outline a method for using the CWT to efficiently remove noise from individual seismic traces and to partition a waveform into component seismic phases. This method will be quickly summarized before presenting its application to phased array processing.

The Morlet-Grossmann definition (Grossmann et al., 1989; Starck et al., 2010) for the CWT is

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t - \tau}{a} \right) dt \quad (1)$$

where the * represents the complex conjugate of the function. The CWT is simply a correlation of the signal, $f(t)$, with a scaled basis function $\psi(t)$. In general, the basis function is complex and is termed the “mother wavelet”. The wavelet coefficient, $W(a, \tau)$ is also complex and can be represented in the Fourier domain as

$$\hat{W}(a, \omega) = \sqrt{a} \hat{f}(\omega) \hat{\psi}^*(a\omega). \quad (2)$$

The CWT is a linear operation and has an exact inverse transform given by the double integral

$$f(t) = \frac{1}{C} \int_0^\infty \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} W(a, \tau) \psi\left(\frac{t-\tau}{a}\right) \frac{da d\tau}{a^2} \quad (3)$$

where C is found from a Parseval-like integral

$$C = \int_0^{+\infty} \frac{\hat{\psi}^*(\omega) \hat{\psi}(\omega)}{\omega} d\omega \quad (4)$$

that requires the basis function to have zero mean for this integral to be bounded at $\omega = 0$. I use Morlet's wavelet as the mother wavelet in most computations in this report unless otherwise indicated.

The first step in the phased array analysis is to remove the obvious noise contained in the seismograms of all array elements. This is done using a soft thresholding process where the wavelet coefficients less than a predetermined noise threshold are removed entirely and those above the threshold reduced by the noise estimate. This is given by

$$\tilde{W}(a, \tau) = \begin{cases} \text{sign}[W(a, \tau)] (|W(a, \tau)| - \beta(a)) & \text{if } |W(a, \tau)| \geq \beta(a) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

where

$$\text{sign}[W(a, \tau)] = \frac{W(a, \tau)}{|W(a, \tau)|}. \quad (6)$$

The threshold level, $\beta(a)$, is determined by examining the statistics of the wavelet coefficients for a noise time window before or after the signal and constructing its empirical cumulative distribution function (ECDF_a) where

$$\beta(a) = \text{ECDF}_a^{-1}(P = 0.99) \quad (7)$$

at the 99% confidence level.

This process was applied to the explosion data recorded by the cross array. In addition, as mentioned in L&M, a scale band reject filter was applied to the remaining coefficients on the scale-time plane for scales greater than 1s to remove long period, high scale, artifacts that remained from the soft threshold step. This improved the SNR from

less than 1 as seen visually in the instrument-corrected data to approximately 200 for the soft-thresholded waveforms, taking the maximum amplitudes of the pre-event noise as the noise measure and the maximum of the resulting signal as the signal measure. These pre-processing steps primarily find those regions of the scale-time plane that have high SNR. As we will see in the array processing, waveforms for the cross array are highly correlated creating close to theoretical beam patterns in the analysis showing that seismic phases are not significantly distorted in the denoising process. However, as pointed out in L&M, if the SNR is low even in the primary signal area of the scale-time plane, significant signal loss and distortion can occur.

3.2.2 Signal Decomposition and Scale-Time Gating

At this point the denoised waveforms can be processed using standard array beamforming techniques. However, the waveforms of the AFTAC explosion show many interesting and discrete seismic phases which appear to clearly separate on the scale-time plane (Figure 3). L&M show that it is straightforward to pick these phases on the scale-time plane by choosing an enclosing polygon and then using the inverse CWT to obtain their time domain waveform. This process can be tedious if manually applied to all 81 stations of the cross array.

The strategy for gating particular phases from individual elements of an array is to choose a reference station near the center of the array, delimit the phase of interest by encompassing it with a polygon on the CWT scalogram and then using this block of the CWT in a two-dimensional correlation with the scalograms of all other elements of the array. Figure 3 shows the result of using the vertical component of station 3016 to pick the fundamental mode surface wave and then finding the maximum of the correlation of the modulus of the CWT coefficient in this block with those for station 3001, the station located nearest the source.

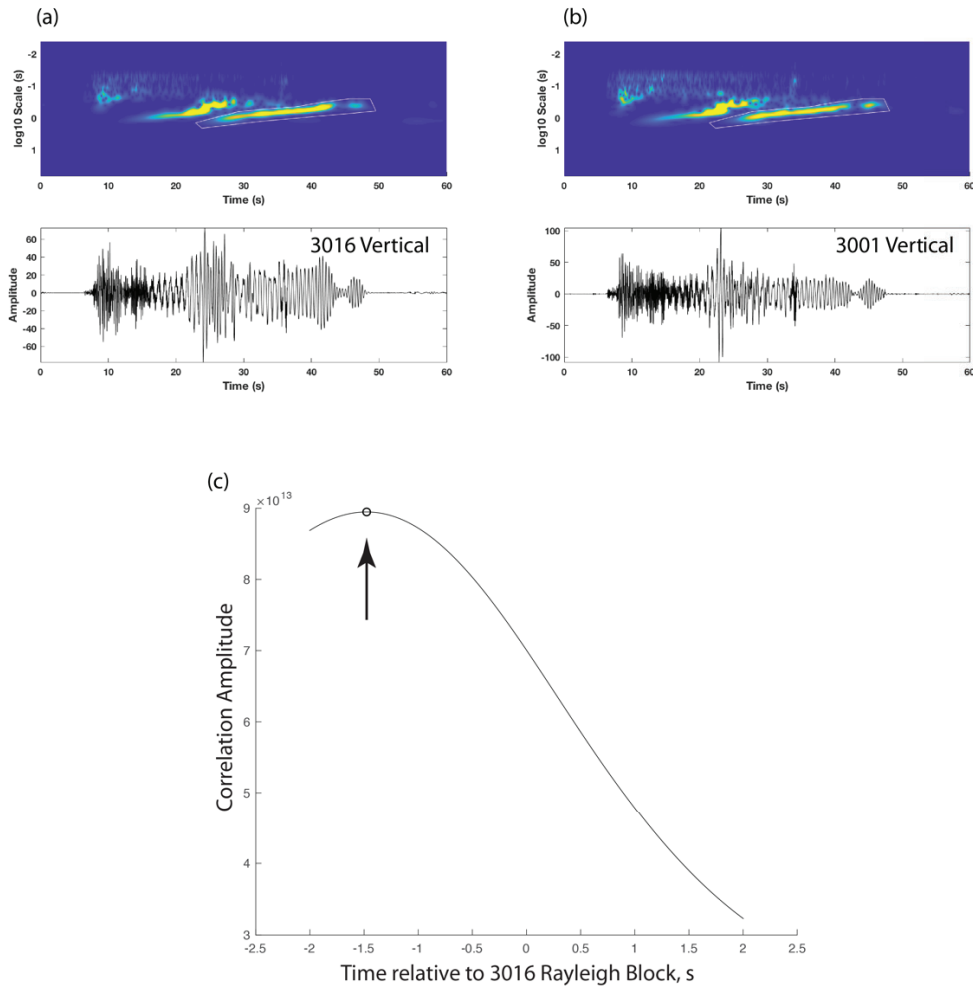


Figure 3 – This is an example of scale-time gating the fundamental mode Rayleigh wave from denoised explosion vertical component seismograms of two stations of the cross array. Panel (a) shows the denoised seismogram (bottom) and its scalogram (top) after performing the Continuous Wavelet transform. A polygon is manually picked to isolate the fundamental mode energy on the scalogram at the reference station 3016. This portion of the scalogram is clipped out and then correlated with the scalograms of all other stations in the array. Panel (b) shows the shifted polygon encompassing the Rayleigh wave at station 3001. Panel (c) shows the amplitude of the correlation vs time shift between the stations. The peak of the correlation occurs at an advanced time, relative to 3016, since station 3001 is closer to the source.

In the process of examining the time shifts obtained from the scale-time gating procedure, it was found that the time shifts were sometimes inconsistent with the expected move-out of the seismic phase, if the modulus of wavelet coefficients was used directly. Indeed, even performing the 2D correlation on the same signal that the block of interest was taken from did not always yield a zero time lag; adjacent wavelet coefficients

outside of the block could contaminate the correlation. Although the inferred time shifts were very small compared to the duration of the windowed signal, it appeared that non-unique correlations of the wavelet coefficient modulus could cause problems with the method. Through numerical experimentation, I found that using a power of $|W|^n$ before correlation was much more effective in giving appropriate time shifts. Using $n=3$ gave results in line with the observed moveout of the seismic phases. Using a power of the wavelet coefficients weights the correlation towards the highest amplitude coefficients in the block which, presumably, are less affected by noise.

Scale-time gating was effective in delimiting the fundamental mode Rayleigh wave (Figure 4) and first higher mode Rayleigh wave (not shown) from all array elements. Detailed results from the dispersion analysis will be discussed later.

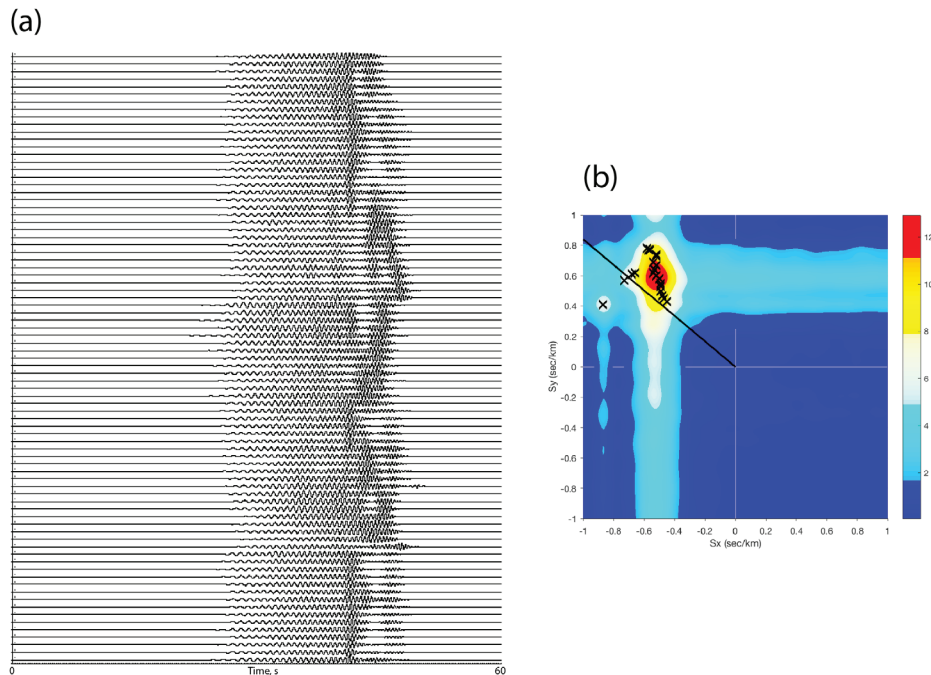


Figure 4 – Panel (a) shows the scale-time gated Rayleigh waves obtained for stations of the cross array. Panel (b) is the broadband frequency-wavenumber result for the frequency band of 1 to 3 Hz encompassing the bandwidth of the Rayleigh waves. The small x's show the results of narrow band f-k for phase velocity as a function of frequency. Note that the bulk of the Rayleigh wave energy arrives approximately 15 degrees clockwise from the expected great circle path shown by the black line.

Scale-time gating was found to work well with the surface wave modes but was problematical in an attempt at gating the P wave trains from the explosion. A straightforward choice of the P wave CWT block (Figure 5a) gave poor results (Figure

6a) since the correlation process often chose the first higher mode over the P wave train. This problem was solved by choosing a block defined by both the fundamental and first higher modes to get the optimum mode time shifts and then zeroing out the CWT coefficients within the mode block (Figure 5b). The remainder gives the desired P wave train (Figure 6b).

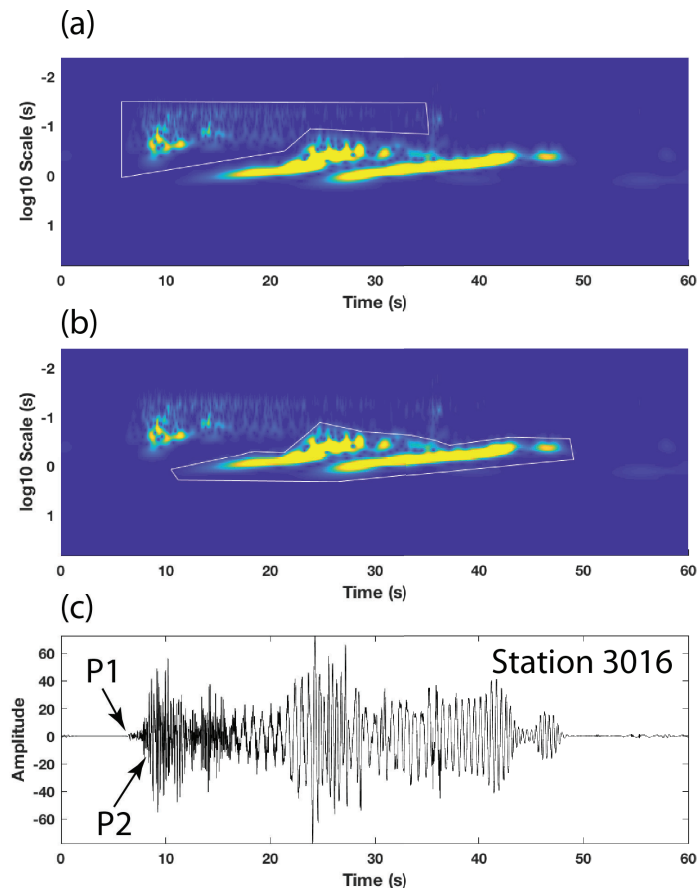


Figure 5 – This figure illustrates two methods for scale-time gating the low scale/high frequency P wave train. The top panel (a) shows a choice of a block on the scalogram the encompasses the desired portion of the waveform. Panel (b) shows its complement where a block containing the high amplitude modes are chosen for correlation instead. (c) displays the seismogram at the reference station 3016. P1 and P2 body wave phases are annotated.

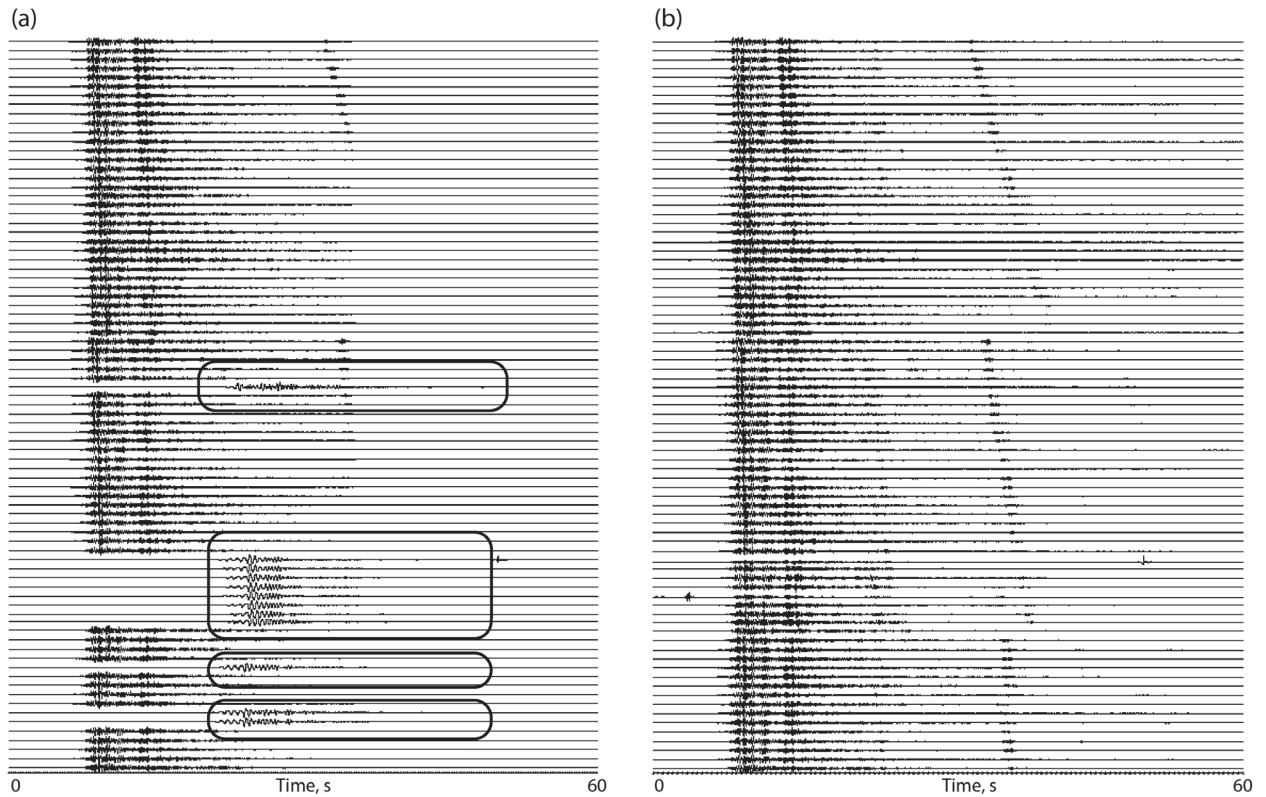


Figure 6 – Results of scale-time gating using the two different choices of CWT blocks shown in Figure 5. In (a), the P wave block in Figure 5a was used to correlate with all other station CWTs. However, results for some stations, shown by the delimited rectangular areas, were poor because the reference CWT block had higher correlation values with later arriving, high-amplitude modes. (b) This was alleviated by correlating with the modes block (Figure 5b) then removing this part of the CWT before inverse transforming.

3.2.3 Broad Band/Narrow Band Frequency-Wavenumber Analysis

The scale-time gated data for the various crustal phases can now be analyzed using standard techniques after the inverse CWT. As an example, I show a detailed dispersion analysis of the fundamental mode Rayleigh wave using a common broad band/narrow band method (Nawab et al., 1985). Figure 4b shows the Rayleigh wave slowness response over the entire frequency band of 1 to 3 Hz encompassing the fundamental mode data. The peak of the response is clearly about 15° north of the expected great circle wave path showing the effects of possible horizontal refraction and multipathing in these high frequency local surface waves.

Narrow band frequency-wavenumber spectra were taken for bandwidths of 0.125 Hz (Figures 4 and 7) and show smooth changes in phase velocity and back azimuth across

the array with frequency over the band of 1 to 2.5 Hz, probably due to horizontal refraction due to velocity heterogeneity in the upper kilometer of the crust. At frequencies higher than 2.5 Hz, the slowness spectra break up into multiple peaks showing several surface wave arrivals with different slowness and azimuths.

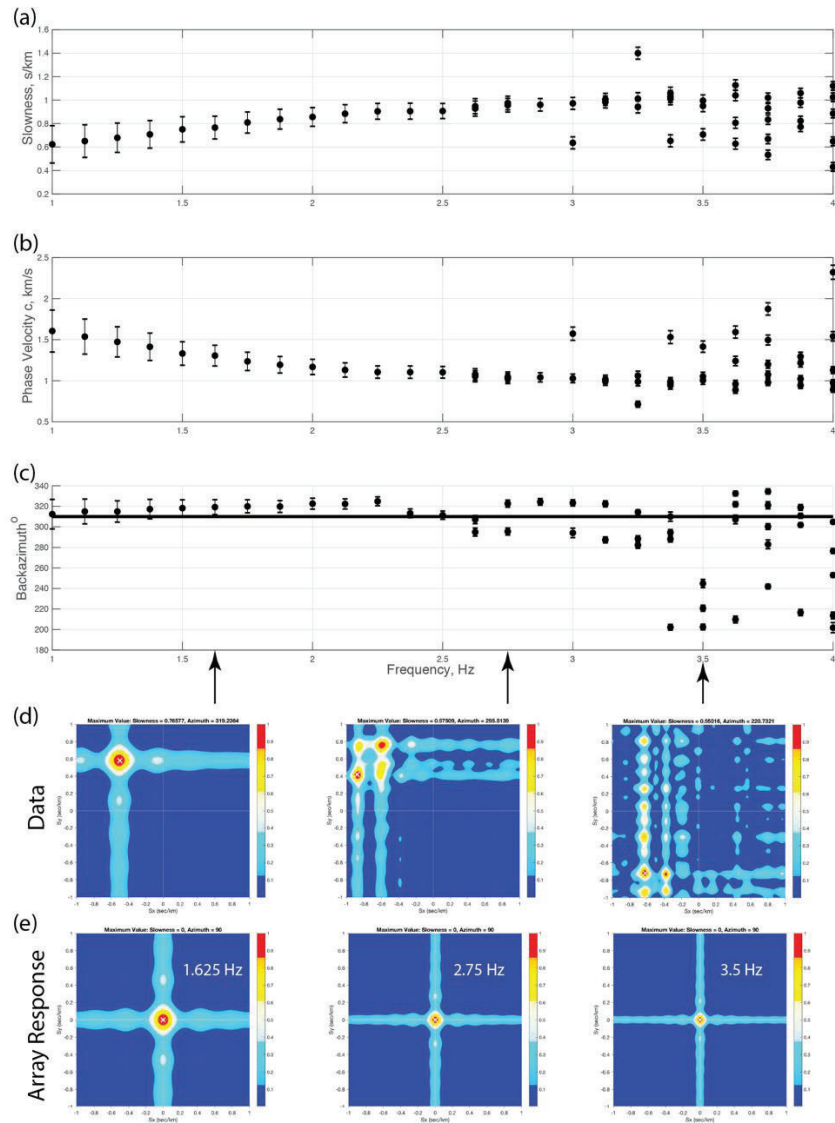


Figure 7 – Results of Rayleigh wave phase velocity dispersion analysis using narrow band f-k. (a) Slowness as a function of frequency. (b) Phase velocity as a function of frequency. (c) Back azimuth as a function of frequency. The heavy line at 311° is the theoretical great circle backazimuth. (d) Selected slowness spectra plots for the frequencies denoted by arrows. (e) Theoretical array responses for the corresponding plots shown in (d). Note that f-k spectra show the effects of Rayleigh wave multipathing for frequencies greater than 2.5 Hz. f-k spectra display multiple peaks at different backazimuths and slownesses.

Error bounds on slowness and back azimuth were estimated by an algorithm that uses the theoretical array response over the frequency band to estimate the variation within the 90% amplitude level. The theoretical response is shifted to the observed maximum peak, scaled by the observed peak amplitude and then subtracted from the data f-k spectrum. The peak of the remaining part of the spectrum is then taken as the signature of the next observed wave and the process repeated to estimate its errors. This process was done to find all peaks in the original slowness spectra down to a level of 0.3 of the maximum peak. As Figure 7 shows, the peaks of the theoretical narrow band slowness spectra have similar widths as the peaks seen in the data, so using the theoretical spectra for the error estimate is reasonable.

Likewise, broadband slowness spectra of P1 and P2 phases over the frequency band of 5 to 15 Hz show quite focused estimates for these P arrivals (Figure 8). Back azimuths are within a degree of the theoretical great circle path to the event showing little evidence for horizontal refraction due to velocity heterogeneity.

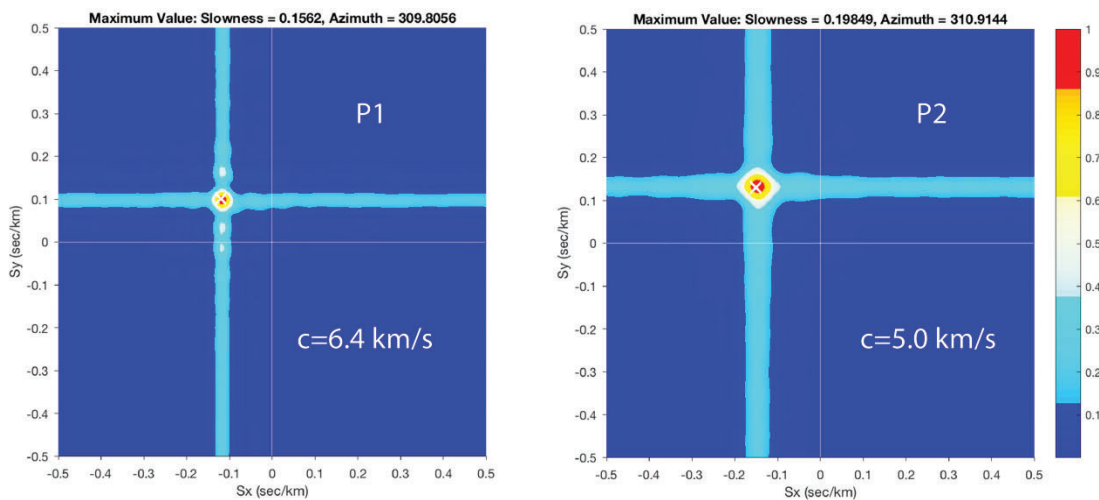


Figure 8 – f-k spectra of P1 (left) and P2 (right) phases in the vertical components of the 2000lb explosion. Note the narrow beams showing that phases are highly correlated across the array.

3.2.4 Beam Forming in the Scale-Time Domain

Up to this point, I have utilized the CWT to remove noise from the original seismograms, separate various phases in the seismogram, and then designed a scale-time gating method to analyze particular arrivals of interest using a phased array. The resulting seismograms were then analyzed using standard Fourier array beam forming such as narrow-band or broad-band methods. Narrow-band analysis is a natural method for determining surface wave dispersion since phase velocity spectral methods are tied directly to solutions for surface wave propagation in vertically inhomogeneous media. However, the question arises: can the wavelet coefficients themselves be used to estimate wave slowness and azimuth? If so, what are advantages to this kind of phased array analysis?

Using equation (1), the CWT of the k th array channel, $f_k(t)$, is given by

$$W_k(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_k(t) \psi^* \left(\frac{t - \tau}{a} \right) dt . \quad (8)$$

The wavelet coefficients, W_k , are the time series that make up the scalogram, or scale-time plot of the transformed time series data. Particular seismic phases have specific locations on the scale-time plane defined by time lags due to their wave propagation characteristics, frequency content of the medium's Green's function, and source excitation. Any time shifts undergone by waves in $f_k(t)$ will show up in the wavelet coefficients. This can be easily shown by assuming a shifted signal $f_k(t-b)$, where b is a constant:

$$I = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t-b) \psi^* \left(\frac{t - \tau}{a} \right) dt . \quad (9)$$

Let $T=t-b$ so that $t=T+b$, and $dT=dt$. This gives

$$\begin{aligned} I &= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_k(T) \psi^* \left(\frac{T - (\tau - b)}{a} \right) dT . \quad (10) \\ &= W(a, \tau - b) \end{aligned}$$

Therefore, beams of each wavelet scale time series can be made by shifting each scale by an assumed plane wave time shift across an array and stacking the shifted scalograms for each array element. A narrow-scale band beam can be defined by

$$W_{Beam}(a, \tau) = \sum_k W_k(a, \tau + p_x x_k + p_y y_k) \quad (11)$$

by summing the lagged wavelet series for a particular scale over all array elements after correcting for plane wave moveout across the array for an assumed slowness vector (p_x , p_y). In analogy to a broad-band stack, wavelet time series for a specified scale band can be summed:

$$\bar{W}_{Beam}(\tau) = \sum_j \sum_k W_k(a_j, \tau + p_x x_k + p_y y_k). \quad (12)$$

In practice, these sums are made by discretizing the slowness plane, making the plane wave time lag corrections to the scalogram and summing. Constructive interference between the array elements will give power maxima on the slowness plane where power is defined as the square of the beam moduli.

These array beams are continuous time series so they can give a continuous estimate of slowness as a function of time. An additional stacking process can be made by integrating the beam power over a time window to get the cumulative energy of discrete signals. An implication of these processes is that each sampled spot on the slowness plane will be a time series so that a very large three-dimensional array of slowness spectra data is generated.

Figure 9 displays an example of a broad scale band beam form according to equation (12). A scale bandpass of 0.1 to 0.5 s has been used to block filter the data. By “block filtering”, I mean that all wavelet coefficients between 0.1 and 0.5 s have been used in the beam form and that all other coefficients have been set to zero. Figure 9 shows the resulting broadband scale-slowness beam for 5 different time points. In addition, the quality of the beam form is assessed by using a simple statistic based on the power of the peak, P_{max} compared to the integrated power over the entire slowness plane, P_{Total} :

$$R = \frac{P_{max}}{P_{Total}} \quad (13)$$

In Figure 9, a phase velocity and back azimuth for a particular time is colored red when $R \geq 0.5$ or blue when $R < 0.5$. High values of R represent broadband scale slowness images that have single prominent peaks, such as for P1 and S in Figure 9, while low values describe diffuse beams or multiple peaks.

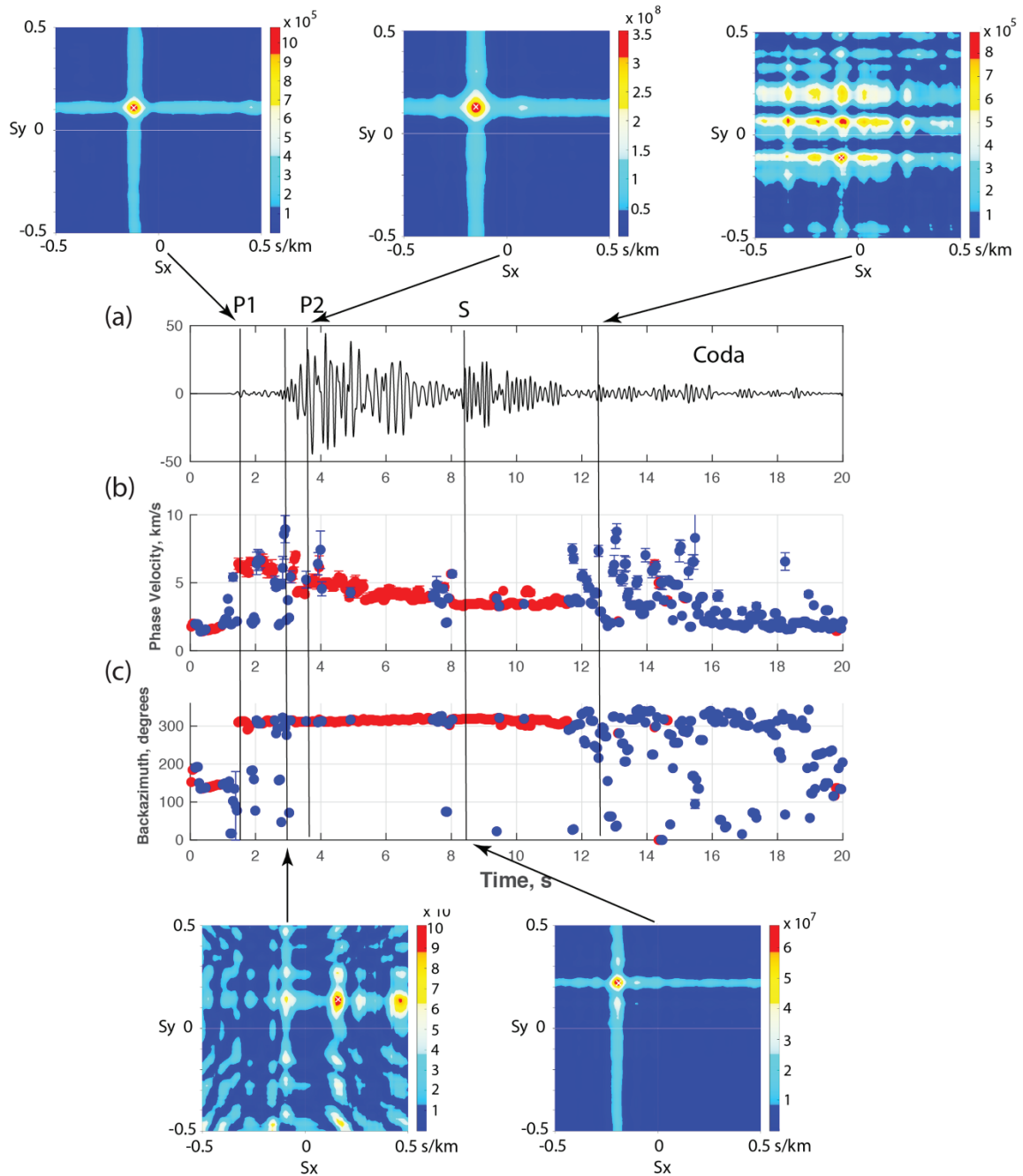


Figure 9 – CWT beamforming result for the P wave block of the 2000lb AFTAC explosion. The low scale wavetrain was determined from excluding the modes block (Figures 5b, 6b). Additionally, a scale bandpass of 0.1 to 0.5s was performed and is shown in (a). P1, P2, S, and Coda phases are annotated. (b) and (c) show the inferred phase velocity and backazimuth, respectively, at every 10th time point in the trace. Red symbols denote well-formed single CWT beams with $R \geq 0.5$ and blue for $R < 0.5$. Representative slowness beams are indicated for the times shown by vertical lines.

Note that this beamforming analysis can be used to identify arrivals in the seismogram that are relatively coherent across the array as well as arrivals that do not

produce good beams. Recall that this waveform was the result of removing the higher and fundamental Rayleigh modes from the denoised seismograms to expose the lower scale parts that seemed to be dominated by P wave arrivals. The CWT beam forming has uncovered what is likely the direct S wave from the explosion traveling at a horizontal phase velocity of about 3.5 km/s. The P1 phase appears to be a head wave with a velocity of about 6.4 km/s followed by a slower velocity P2 phase that may be the start of P reverberations in the near-surface Paleozoic sediments. The P2 wave train gradually gets slower with time attaining horizontal velocities of about 4 km/s until the S wave arrives. Coda waves show scatter in slowness and back azimuth indicating multiple arrivals with different speeds and directions. These are important attributes of the seismogram to know about before attempting to model waveform details for crustal structure.

3.3. Structure Modeling

3.3.1. Rayleigh Wave Phase and Group Velocity

The combination of a densely deployed seismic array, CWT denoising and array processing methods, and a well-controlled seismic source has allowed considerable detail to be discerned from waveforms that started out with a nominal SNR of less than one. It is unusual to observe such well-formed and coherent local seismic phases from events tens of kilometers from a station. The distinct Rayleigh wave dispersion determined for structure under the array (Figure 7) can be used to infer a velocity model that can find further use in locating nearby seismicity and understanding the propagation characteristics of local seismic wave trains.

Wave attributes inferred from an array reflect both the structure under the array and some of the history of wave propagation that the seismic wave experienced along the entire path. Rayleigh wave phase velocity inferred from the array analysis (Figure 7) is primarily related to the local structure under the array. However, the back-azimuth anomaly is an indication of wave propagation effects that occurred outside of the array. Group velocity, as it is usually measured, is a whole-path wave attribute since it is usually calculated by dividing the source-receiver distance by the group arrival time (e.g., Ewing et al., 1957; Capon, 1970). Lateral refraction due to horizontal velocity heterogeneity is governed by Snell's law and the phase velocity. Examining both phase

and group velocity may offer clues to the structure causing lateral refraction and the differences in array structure compared to the whole path.

Using station 3016 as the central reference array station, group velocity was determined for both the local array structure and along the entire path. The Rayleigh waveforms obtained after scale-time gating over the 81 array stations were narrow bandpass filtered with a bandwidth of 0.125 Hz around central frequencies starting at 1 Hz with a frequency increment of 0.125 Hz. The envelope functions of the bandpassed Rayleigh waves were computed using the instantaneous amplitude of the analytic signal and the arrival time of the peaks of the envelopes were automatically picked. Group velocity for the whole path was estimated by dividing the source – receiver distance of 34.7 km at station 3016 by the envelope peak time.

Group velocity under the array was determined by inverting the peak arrival times for group slowness in x (EW) and y (NS). Assume a Cartesian coordinate system with x being Easting and y being Northing. The group arrival time, t_i , at the i^{th} array station located at (x_i, y_i) is given by

$$t_i = \frac{-\vec{R}_i \cdot \vec{B}}{v} + t_0 \quad (14)$$

where v is the group velocity and t_0 a reference time. \vec{R}_i is the vector to the i^{th} station given by

$$\vec{R}_i = x_i \hat{e}_1 + y_i \hat{e}_2 \quad (15)$$

and \vec{B} is the negative of the wave direction vector given by

$$\vec{B} = \sin \theta \hat{e}_1 + \cos \theta \hat{e}_2 \quad (16)$$

with θ being the backazimuth. The x and y group slownesses are defined by

$$\begin{aligned} \gamma_x &= \frac{\sin \theta}{v} \\ \gamma_y &= \frac{\cos \theta}{v} \end{aligned} \quad (17)$$

Given n array stations with group arrival time observations, a linear system of equations can be set up to solve for the x and y group slownesses:

$$\begin{bmatrix} -x_1 & -y_1 \\ -x_2 & -y_2 \\ \vdots & \vdots \\ -x_n & -y_n \end{bmatrix} \begin{bmatrix} \gamma_x \\ \gamma_y \end{bmatrix} = \begin{bmatrix} t_1 - t_0 \\ t_2 - t_0 \\ \vdots \\ t_n - t_0 \end{bmatrix} \quad . \quad (18)$$

This is of the form of

$$\mathbf{Gm} = \mathbf{d} \quad (19)$$

with the least-squares solution

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d} \quad . \quad (20)$$

Define

$$\mathbf{G}_g^{-1} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \quad . \quad (21)$$

A very conservative estimate of the error is obtained by using the least squares error after inversion

$$lse = \|\mathbf{d} - \mathbf{Gm}\|_2 \quad (22)$$

to estimate the model parameter covariance matrix

$$\mathbf{C} = lse^2 \mathbf{G}_g^{-1} \mathbf{G}_g^{-1T} \quad (23)$$

where

$$\begin{aligned} \sigma_{\gamma_x}^2 &= C_{11} \\ \sigma_{\gamma_y}^2 &= C_{22} \end{aligned} \quad . \quad (24)$$

The total slowness is given by

$$\gamma = \left(\gamma_x^2 + \gamma_y^2 \right)^{1/2} \quad (25)$$

and to first order its error is

$$\sigma_\gamma^2 = \left(\gamma_x \sigma_{\gamma_x} + \gamma_y \sigma_{\gamma_y} \right). \quad (26)$$

Backazimuth is given by

$$\theta = \tan^{-1} \frac{\gamma_x}{\gamma_y} \quad (27)$$

and its error, also to first order, by

$$\sigma_\theta = \tan^{-1} \left(\frac{\gamma_y \sigma_{\gamma_x} + \gamma_x \sigma_{\gamma_y}}{\gamma_y^2} \right). \quad (28)$$

Surprisingly, the group velocity data for the array could clearly discriminate between the great circle back azimuth and the group arrival back azimuth for frequencies less than 2 Hz (Figures 10 and 11). The whole-path group velocity and group velocity under the array were not significantly different suggesting that any horizontal structure changes may be subtle (Figure 11).

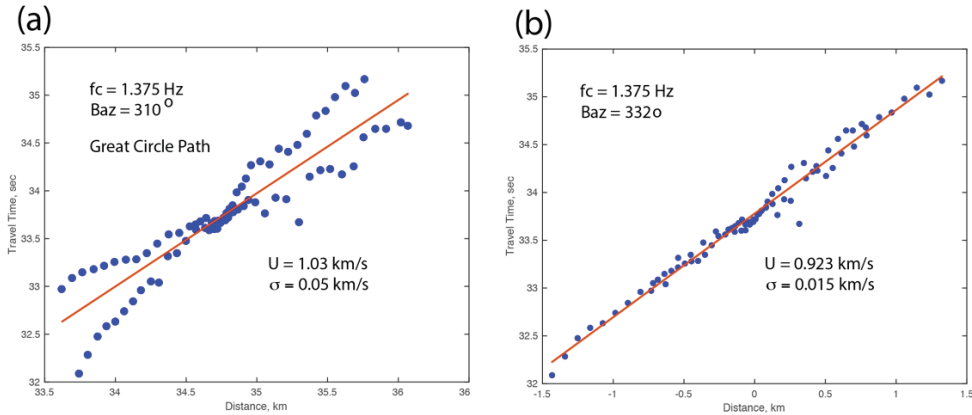


Figure 10 – Array group velocity analysis example for 1.375 Hz vertical component Rayleigh waves. Group arrival times are plotted against station distance from the source assuming the great circle path to the source in (a). Because the waves are arriving from an azimuth clockwise from the great circle backazimuth, times from the two linear arms of the cross array are inconsistent. In (b), an inversion for group slowness and backazimuth is performed to find the optimum backazimuth and group velocity for structure under the array. The data clearly indicate that Rayleigh waves arrive 22° clockwise from the great circle backazimuth, similar to the amount seen in the phase velocity determination.

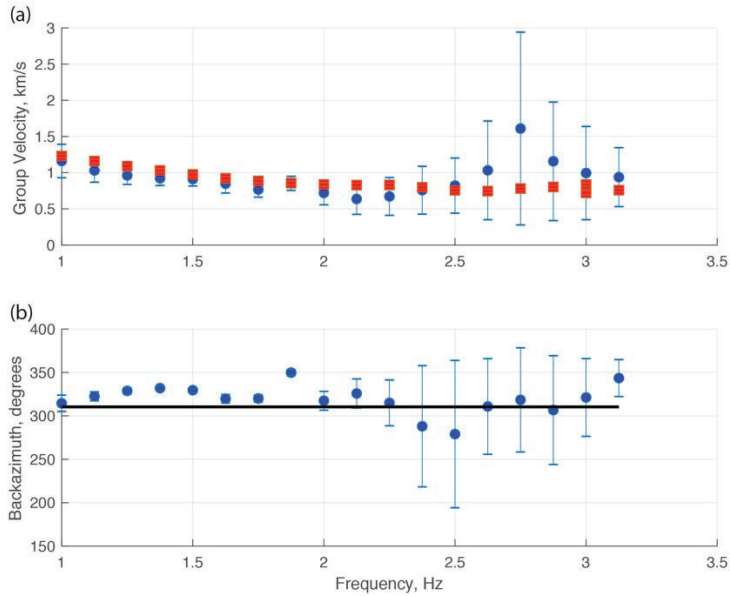


Figure 11 – (a) compares the group velocity determined by array analysis (filled circles) with the whole path group velocity determined by dividing the great circle distance by the group arrival time at station 3016 (filled squares). There seems to be no significant difference in group velocities determine by two independent means. (b) shows the backazimuth determined from the array group velocity analysis as a function of frequency. Multipathing causes the larger errors at frequencies greater than 2.5 Hz.

The explosion waveforms are dominated by two, well-dispersed wave trains of Rayleigh waves (Figure 12). Array waveforms show high frequency P and S wave trains along with two distinct bands of energy in the scalograms corresponding to the fundamental and higher modes of the crustal structure. Each mode packet can be isolated from the data using scale-time gating and then analyzed for phase velocity dispersion. Because the mode packets are so distinct and simple, an initial assumption going into this work was that a relatively simple layered earth model should be able to explain the dispersion. This did not turn out to be the case. A model could not be found that fit both the fundamental and the higher modes simultaneously. Evidence from the data suggests that scattering observed in the array analysis is due to lateral heterogeneities in the upper 0.5 km which severely perturbs the response of the first higher mode. Nevertheless, the data do put significant constraints on possible models and point to future interesting modeling studies on the effect of near-surface lateral heterogeneities on surface wave excitation and propagation. This can have significant implications for verification studies of small clandestine events if source locations are wanted from array observations of

wave slowness and backazimuth as well as estimating source depth from the surface wave amplitude.

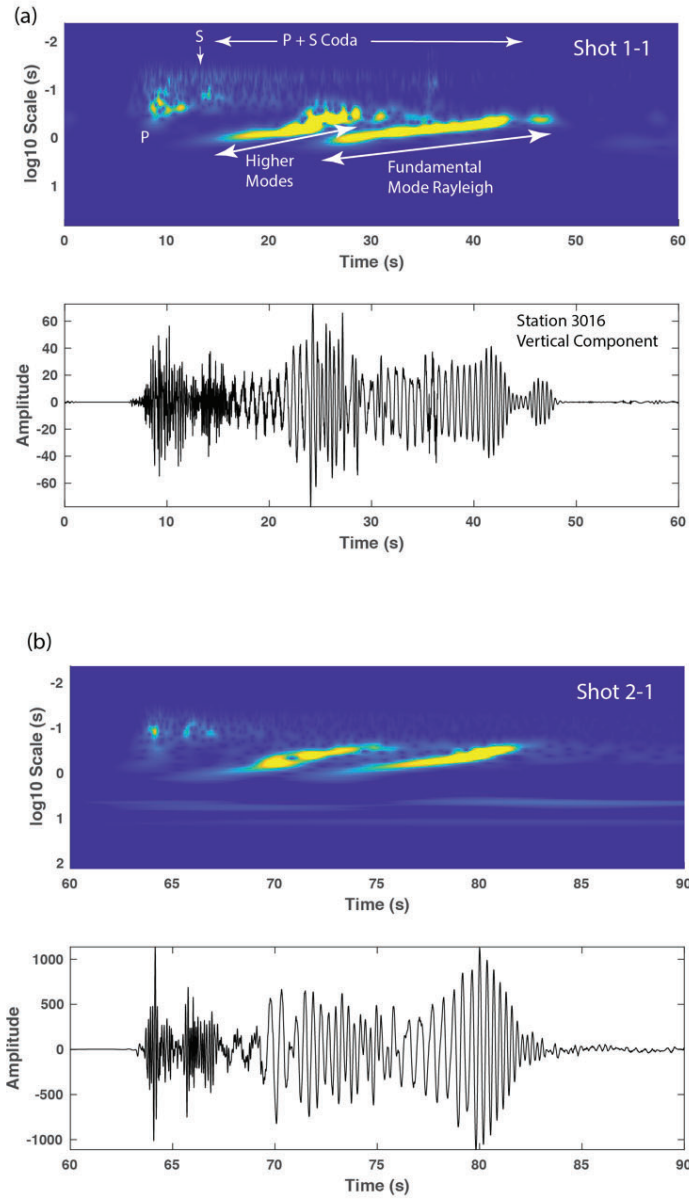


Figure 12 – Waveforms and CWT scalograms for 2000 lb events at shot point 1 (a) and shot point 2 (b) recorded at station 3016 in the IRIS nodal array. Note the similarity in vertical waveform between the two shots although the time scale is 50% shorter for (b) compared to (a). Scalograms for both shots show distinct, well-developed regions of the CWT separating fundamental and higher mode Rayleigh waves.

3.3.2 Regional Velocity Structure

The 2016 IRIS Wavefields Community experiment was installed northeast of the town of Enid, Oklahoma (Sweet et al., 2018). The experiment sits in the Anadarko shelf geologic province of Oklahoma just west of the Nemaha basement uplift (Johnson, 2008). The Anadarko shelf consists of relatively flat-lying rocks of Cambrian to Permian in age on top of preCambrian granite and gneiss. Permian clastic rocks dominate in the upper part of the ~2 km thick section under the experiment and grade downwards into interbedded limestones, dolomites, and shales.

A sonic log was obtained from the Oklahoma Geological Survey (Nori Nakata, personal communication, 2017) for an exploration borehole only 27 km away from the IRIS experiment and in the direction of shot point 1 of the AFTAC experiment (Figure 1). The integrated travel times of the log were digitized on a hundred foot depth interval and then converted to velocity. P wave velocity was converted to S wave velocity using Brocher's (2005) equation (6) and density using the Nafe-Drake relation, Brocher's (2005) equation (1). The sonic log starts at a depth of 600 ft so a layer 183 m in thickness with a P wave velocity of 2.59 km/s was added as the first layer; this velocity is consistent with near-surface structure for sedimentary sections but it is arbitrary. A better velocity will be inferred from the surface wave dispersion.

The resulting velocity model shows a low velocity, near-surface region, then alternating thin layers of high and low velocity down to about 1.6 km where velocities appropriate for crystalline rocks of 6 km/s and greater are attained (Figure 13). The average S wave velocity is predicted to be about 2.0-2.5 km/s in sedimentary rocks deeper than 400 m. An alternative explanation is that crystalline basement occurs at about 1900m and that high velocity limestones or dolomites are the source of the high velocity rocks near 1600m depth. The large negative excursion in P wave velocity at 1800m might suggest this, since a P wave velocity of 3.8 km/s is rather low for crystalline rocks. A geologic log was not included with the sonic log.

This kind of structure can be expected to create a pronounced low velocity, near-surface wave channel that will favor well-dispersed, slowly propagating surface waves.

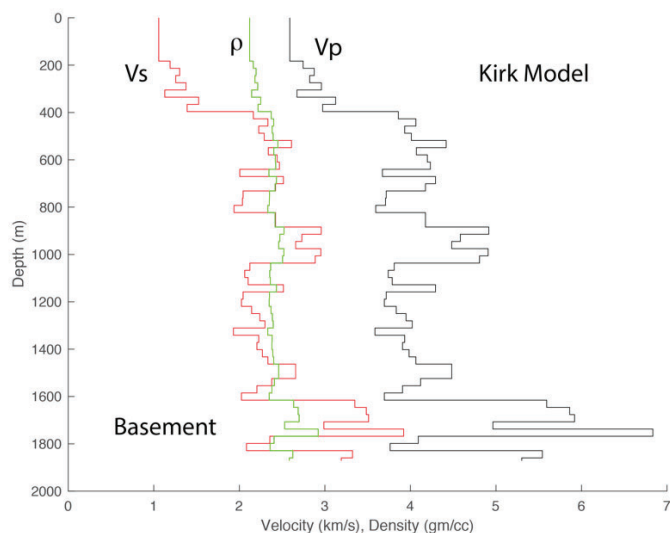


Figure 13 – Velocity model derived from Sarah Kirk Well sonic log contracted by the Mobil Oil Corporation, 1985.

3.3.3 Verifying Dispersion Measurements

The previous section outlined the CWT processing steps that went into denoising, scale-time gating, and then narrow band frequency-wavenumber filtering to obtain phase velocity dispersion for the cross array. Figure 14 shows the resulting measurements for the 2000 lb shot at shot point 1 35km northwest of the array. Measurement errors were estimated by using the 90% width of the power peak in the 2D f-k response. Fundamental mode dispersion for frequencies of 0.75 Hz and lower were estimated by using vertical component data from the entire nodal array, since resolution degraded at these frequencies for the 3km aperture cross array. Scale-time gating was performed on the data to separate the fundamental mode and the higher mode wave trains before dispersion measurements were taken.

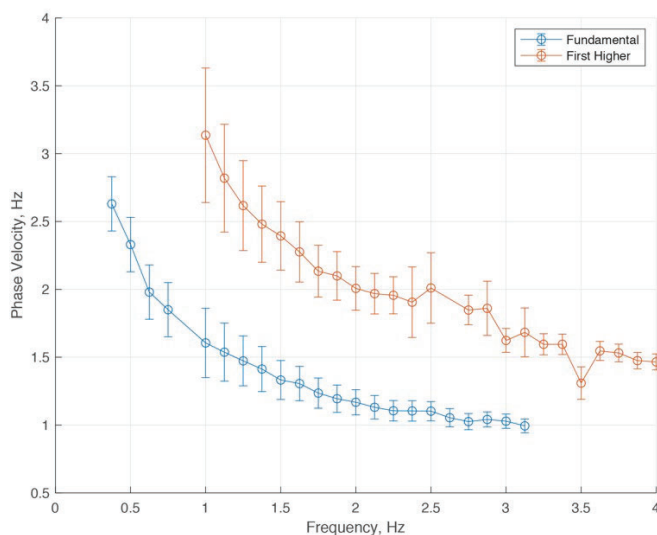


Figure 14 – Phase velocity dispersion for the fundamental mode and higher mode wave trains for Rayleigh waves generated by the 2000lb shot at shot point 1.

The inferred dispersion from Rayleigh waves generated from shot point 1 was checked using the waveforms from the 2000lb explosion at close range, shot point 2. The closest nodal station was about 10 km and the farthest at about 22km (Figure 15). The entire IRIS Wavefields array was utilized to determine Rayleigh wave dispersion from about 0.5Hz to 4Hz. The waveform data were arranged with respect to distance from the known shot point and then transformed using the 1D frequency-wavenumber transform for a plane wave. The resulting transform gives very precise estimates of fundamental mode phase velocity dispersion (Figure 16).

The phase velocity data were used in numerous calculations of forward models. Ultimately, the fundamental mode data from both sources were inverted to find a simple velocity model.

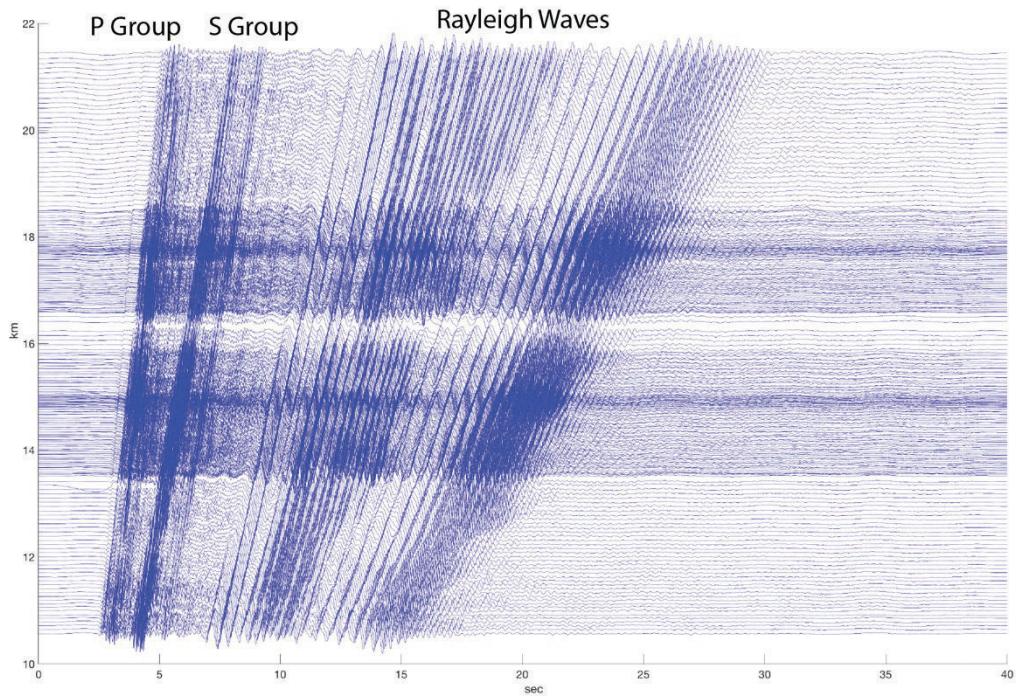


Figure 15 – Profile of all vertical data from the 2000lb explosion at shot point 2. Major phase types are annotated. Note the high quality of waveform clearly showing moveout of all major phases. The data were instrument corrected but no filtering or noise thresholding were applied to remove the small amount of low frequency noise.

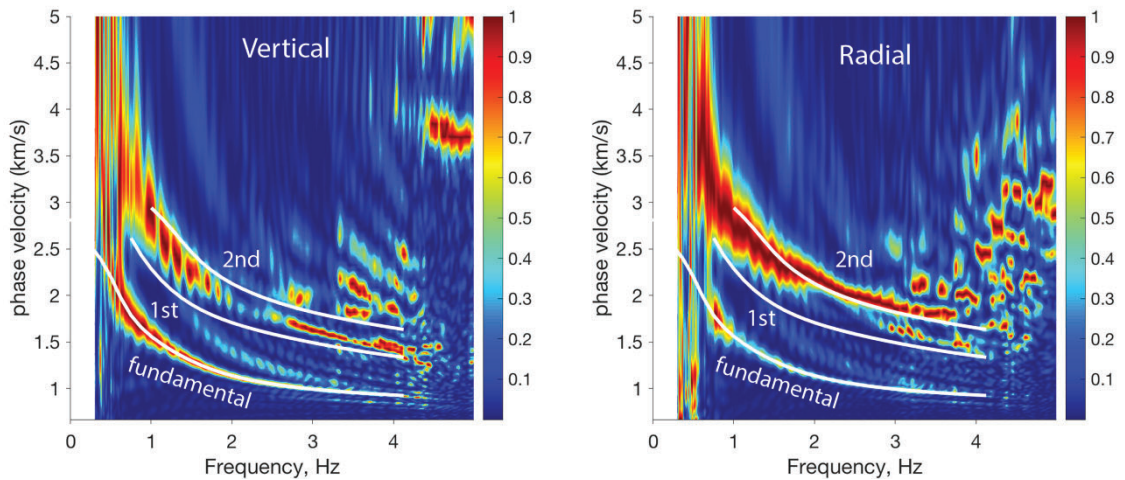


Figure 16 – 1D f-k response of all vertical (a) and radial (b) components of explosion 1 at shot point 2. The white curves are theoretical dispersion curves derived from a velocity model determined from inversion of the fundamental mode data.

3.3.4 Inversion Method

A shear wave velocity model is parameterized by specifying velocity nodes as a function of depth with a linear velocity gradient in between nodes, except at layer boundaries. An example is shown in Figure 17. Shear wave velocity and depth may be the object of the inversion but either parameter may be constrained depending on available data. For example, in Figure 17, the depths of the two nodes at 1.7km are constrained from the acoustic well log taken near the explosion shot point and array (Figure 1b). Because Rayleigh wave phase velocity is relatively insensitive to P wave velocity and density, these parameters are determined from the shear wave velocity. P wave velocity is derived from equation (9) of Brocher (2005) and density from the Nafe-Drake relation, equation (1) of Brocher (2005).

The inversion problem is set up using the “jumping” inversion method of Constable et al. (1987). Phase velocity, c_j , for frequency, f_j , may be a function of v_{s_i} and h_i , the depth of the node. The problem is linearized the usual way by expanding phase velocity in a Taylor’s series and truncating it at the first term:

$$c_j(\vec{m}) = c_j(\vec{m}_0) + \left. \frac{\partial c}{\partial v_{s_i}} \right|_{\vec{m}_0} (v_{s_i} - v_{s_i}^0) + \left. \frac{\partial c}{\partial h_i} \right|_{\vec{m}_0} (h_i - h_i^0) \quad . \quad (29)$$

where \vec{m}_0 is the starting model vector. Equation (29) is manipulated to

$$\left. \frac{\partial c}{\partial v_{s_i}} \right|_{\vec{m}_0} (v_{s_i} - v_{s_i}^0) + \left. \frac{\partial c}{\partial h_i} \right|_{\vec{m}_0} (h_i - h_i^0) = c_j(\vec{m}) - c_j(\vec{m}_0) \quad (30)$$

to give the matrix relation

$$\mathbf{G}\Delta\mathbf{m} = \Delta\mathbf{d}. \quad (31)$$

The partial derivatives in equation (30) were calculated using the central finite difference; two velocity models were calculated for each derivative.

The predicted data from the starting model are added to both sides of (31) to give a linear problem in the model parameters directly rather than the model parameter changes:

$$\begin{aligned} \mathbf{G}\Delta\mathbf{m} + \mathbf{G}\mathbf{m}_0 &= \Delta\mathbf{d} + \mathbf{G}\mathbf{m}_0 \\ \mathbf{G}(\Delta\mathbf{m} + \mathbf{m}_0) &= \mathbf{d} \quad . \quad (32) \\ \mathbf{G}\mathbf{m} &= \mathbf{d} \end{aligned}$$

This problem is iteratively solved using singular value decomposition (e.g., Aster et al., 2005) where

$$\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T \quad (33)$$

and

$$\mathbf{G}_g^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}^T \quad (34)$$

The problem is overdetermined so errors in the parameters are estimated by finding the covariance matrix of the parameters through

$$\mathbf{E}_m = \mathbf{G}_g^{-1}\mathbf{E}_d(\mathbf{G}_g^{-1})^T \quad (35)$$

where

$$\mathbf{E}_d = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix} \quad (36)$$

is the data covariance matrix containing the estimated variances of each measurement of phase velocity.

In application to finding the velocity model of Figure 17, convergence occurred within 4 iterations with the largest changes occurring in the first iteration. The error bars in Figure 17 are the standard deviations determined from the diagonal elements of the model covariance matrix and only give an approximate indication of the errors involved. The model covariance matrix has significant off-diagonal terms indicating large trade-off between some parameters such as node depth and velocity.

A MatLab code was written to compute phase velocity for plane layered Earth structure using the classical theory presented by Harkrider (1964). Layered models were constructed using the node velocity and positions with thin layers approximating the linear shear wave velocity gradients between nodes. P wave velocity and density were derived from each shear wave value using Brocher's equations mentioned above.

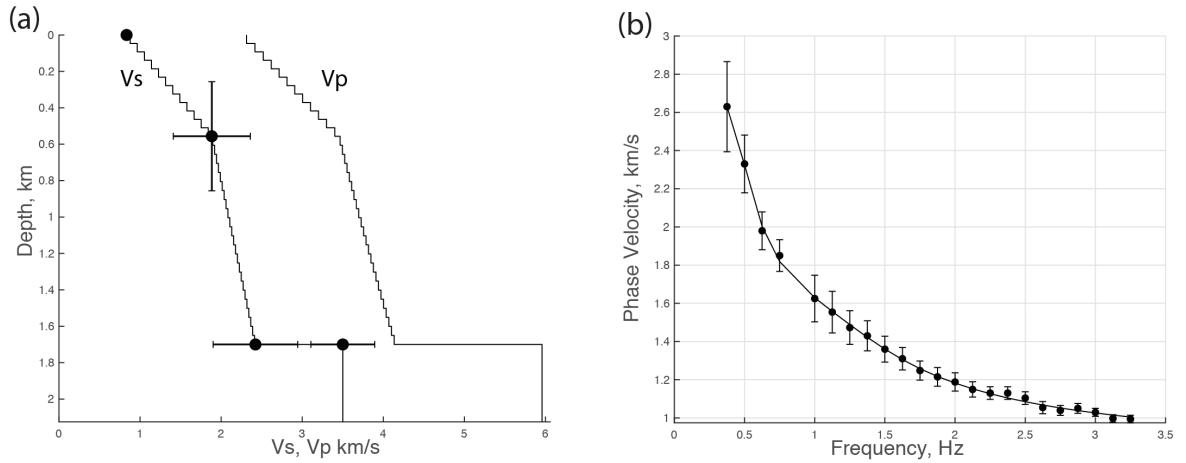


Figure 17 – Inversion of the fundamental mode phase velocity data from the explosion at shot point 1. (a) A velocity structure is characterized by discrete nodes in shear wave velocity and depth. A linear gradient is assumed between nodes, except at layer boundaries. Error bars in (a) are estimated from the model covariance matrix, equation (35). (b) The theoretical dispersion curve fit is excellent.

The same inversion process was used to invert the fundamental mode data from the shot point 2 explosion (Figure 18). Errors in the data are much smaller because of the large spatial bandwidth of the data yielding highly focused frequency-wavenumber images (e.g., Figure 16). The resulting model has the same characteristics of the model Figure 17) and appears to be much more precise in the upper half km.

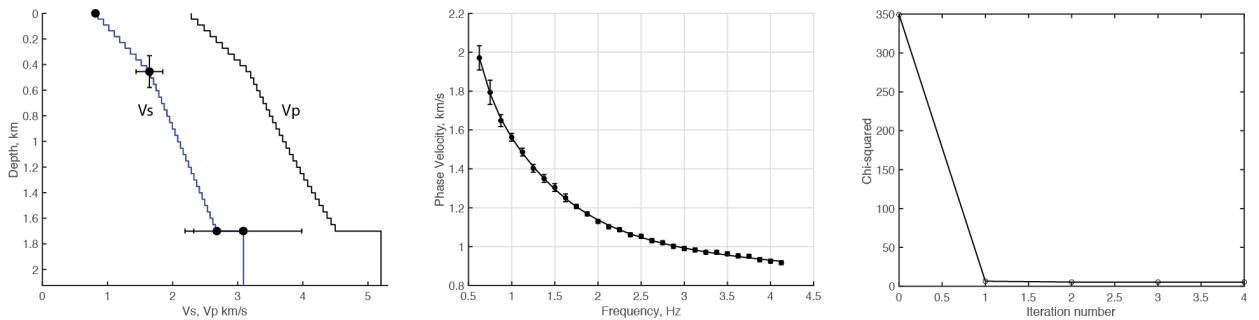


Figure 18 – The left panel displays the inversion result using the phase velocity data (middle panel) from the shot point 2 explosion. The right panel shows the change in the inversion least squares error with iteration.

The results show that a very simple velocity structure can fit the phase velocity data and that the data are most sensitive to the very near surface velocity gradient. This suggests that small changes in the near surface velocity could explain the back azimuth anomaly. A candidate structure is the roughly EW trending northern edge of the Salt Fork River plain (Figure 1b) where it could be expected that the river plain has somewhat lower shear wave velocities in the uppermost part of the structure compared to structure north of the river. A phase velocity contrast of 0.3 km/s could explain the back-azimuth anomaly for the 2Hz Rayleigh waves.

3.4. An Earthquake Source Depth Estimate

Waveforms from the M3.6 28 June 2016 Langston, Oklahoma, earthquake offer an example for using CWT processing techniques to infer source parameters, particularly observations of earthquake depth phases (Figure 19). There is only a small amount of background noise and the waveforms show a distinctive set of body and surface waves. Herrmann (2016) derived a moment tensor model for this event using available regional broadband data that were filtered in the 0.03 to 0.1 Hz frequency band. The mechanism shows right-lateral strike-slip faulting occurring on a NE-SW vertical plane or left-lateral strike-slip motion on an 80° dipping NW-SE striking plane (Figure 1). Source depth was inferred to be 4 km.

These waveforms were processed using the CWT thresholding and block manipulation techniques to remove noise and then separate the various phases in the seismograms. The denoised vertical component scalograms show several prominent phases including the fundamental mode Rayleigh wave, a harmonic and dispersed PL phase, along with high frequency P and associated coda (Figures 20 and 21).

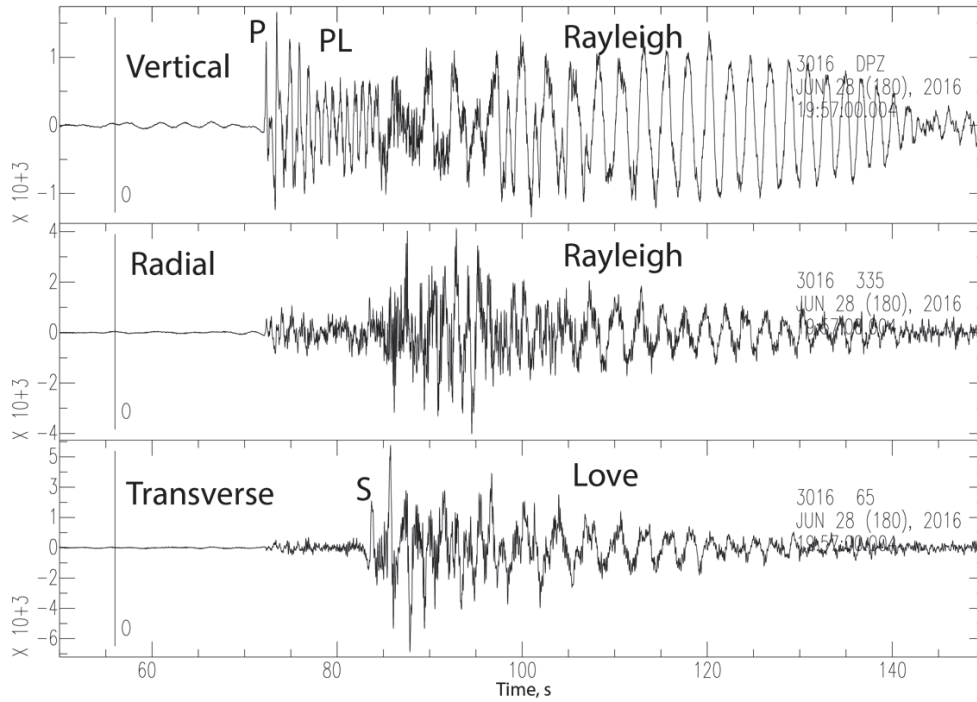


Figure 19 – Waveforms recorded at station 3016 for the M3.6 6/28/2016 Langston, Oklahoma, earthquake. Vertical, radial, and transverse component waveforms are shown. The data have been corrected for instrument response. P, PL, Rayleigh, and Love waves are annotated.

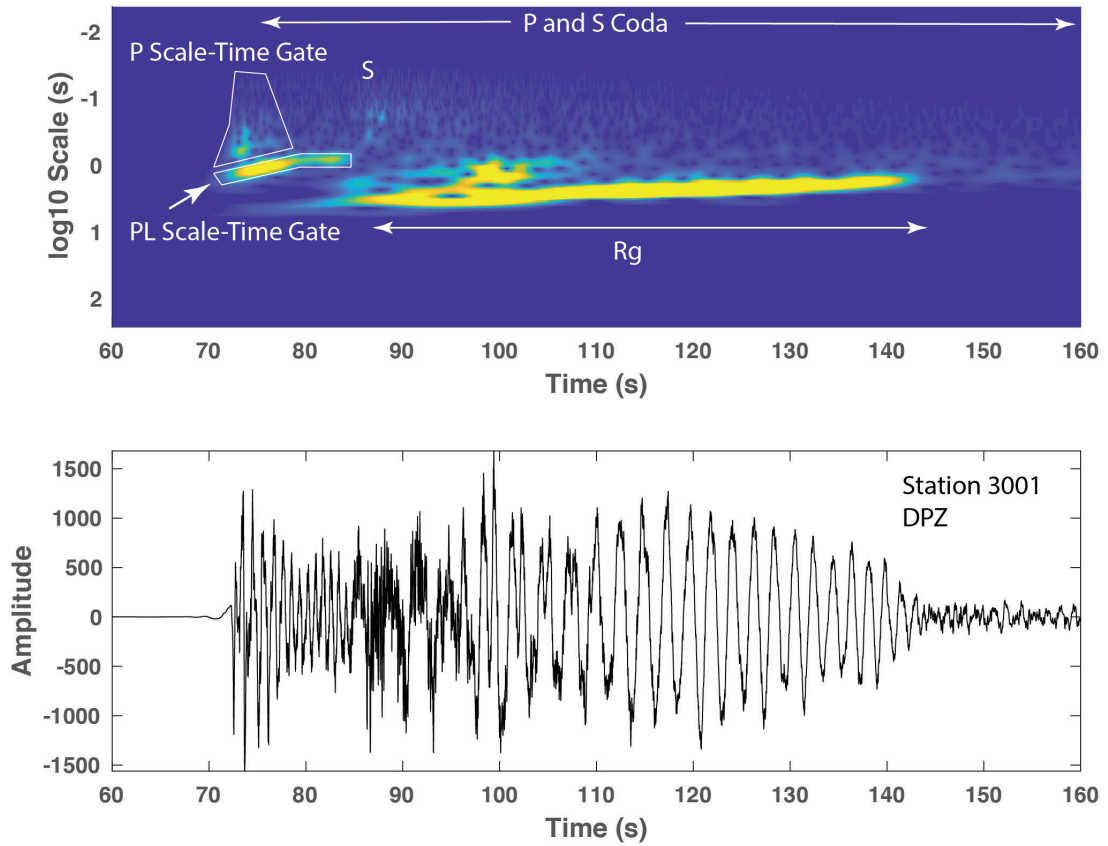


Figure 20 – Scale-time gates chosen for the vertical component of the M3.6 earthquake. The top panel shows the scalogram for the seismogram shown below. Scale-time gate polygons were chosen for the low scale portion of the P wave and harmonic portion of the PL phase. S, Rg, and Coda phases are also annotated.

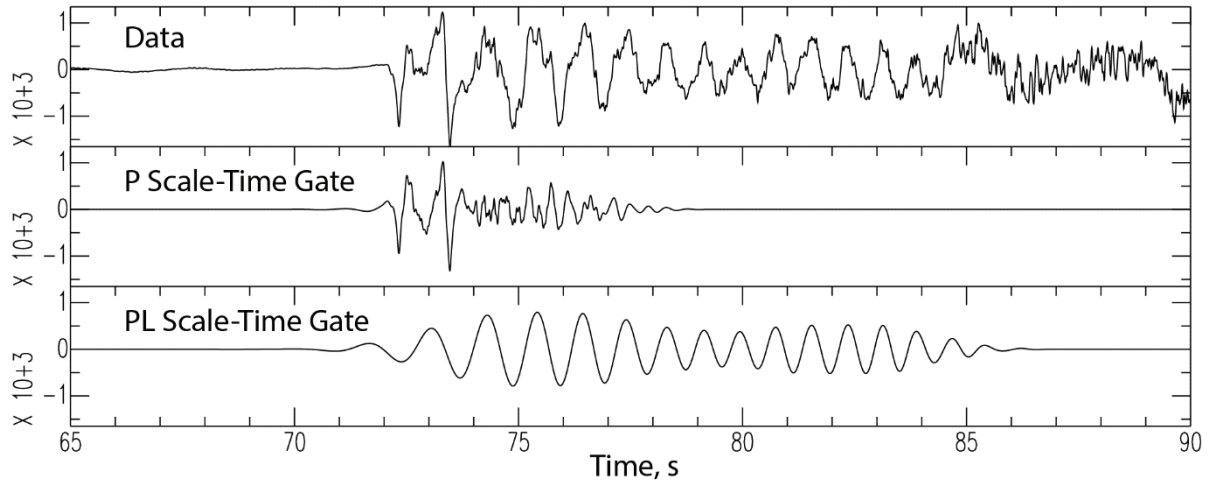


Figure 21 – Results from scale-time gating the vertical component of station 3016 using the P and PL wave gates shown in Figure 20. The original data is shown in the top panel and the gated waveforms shown below.

Choosing a scale-time gate around the low scale (high frequency) portion of the P waveform yields a distinctive pair of phases that are seen to propagate at a velocity of 6.4 km/s (Figure 22). Source depth can be estimated from the relative time of 1s between the two phases if the second phase is interpreted as pP where

$$t_{pP} - t_P = 2\eta_\alpha h \quad (37)$$

for the relative time between P and pP or if the second phase is sP so that

$$t_{sP} - t_P = (\eta_\alpha + \eta_\beta) h \quad (38)$$

where $\eta_\alpha = \left(\frac{1}{\alpha^2} - p^2 \right)^{\frac{1}{2}}$, $\eta_\beta = \left(\frac{1}{\beta^2} - p^2 \right)^{\frac{1}{2}}$, α and β are the average P and S wave

velocities above the source, respectively, p is the horizontal slowness, and h is the source depth (Langston, 1987). Assuming the second phase is pP yields a source depth of 4.0km for average P and S velocities of 5.0km/s and 2.9km/s. If the phase is sP, then the source depth is 2.3km/s.

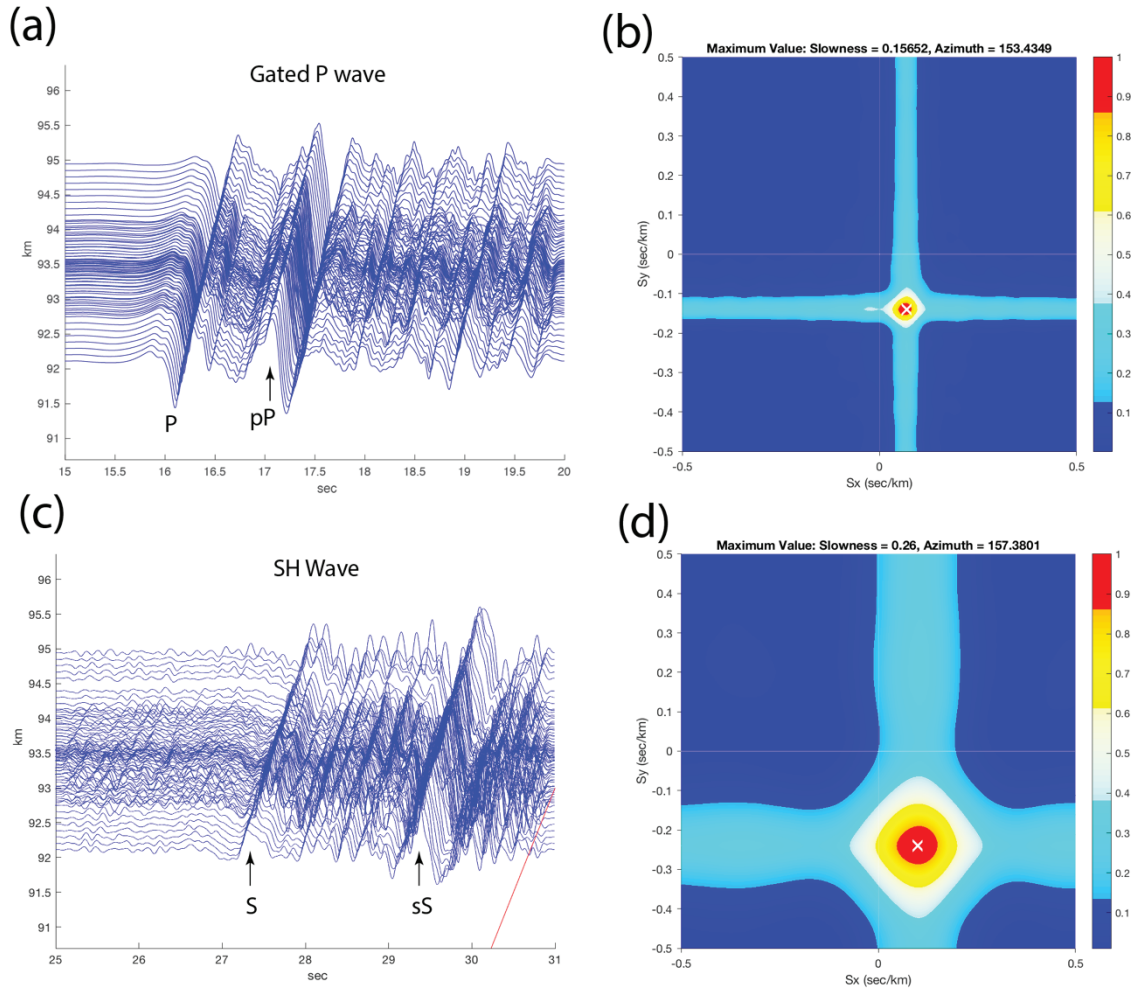


Figure 22 – (a) shows the gated P wave at the 81 stations of the cross array. The broadband f-k beam for 5 to 20Hz is shown in (b) and indicates that the two major phases, denoted by P and pP, have identical slownesses. (c) shows the instrument-corrected SH waves at the cross array with (c) the 1 to 10Hz broadband f-k response. S and sS phases are annotated on the seismograms.

Array examination of the transverse component waveforms gives an estimate of 0.26 s/km for horizontal slowness of the initial 3s of the SH waveforms (Figure 22) and observations of a distinct secondary arrival that is interpreted as sS with a relative time of 2s, implying a source depth of 4.4 km. The pP and sS times suggest that the source depth is between 4 and 4.4km which is consistent with the regional moment tensor depth of 4km (Herrmann, 2016). The discrepancy between pP and sS estimates is probably due to unknowns in the near-source velocity structure. This suggests that knowledge of the wave attributes obtained by precise decomposition of the seismogram for component

phases can significantly constrain the source depth. Source depth is very difficult to accurately constrain using standard earthquake location algorithms based on direct wave travel times for earthquakes within sparse networks or from relatively long-period moment tensor source inversion.

3.5. Other Example Applications of CWT Array Processing

The IRIS Wavefields Community experiment dataset may prove to be a canonical dataset for developing and testing many new array methods on a variety of natural and artificial sources. The AFTAC explosion experiment combined with a month-long deployment to capture thousands of natural earthquake sources offers a literal smorgasbord of seismic sources. These vary from large teleseismic earthquakes, regional moderate earthquakes, and local microearthquakes and explosions. Here I examine the AFTAC explosion sources at shotpoint 1 and show how CWT denoising operates in an environment of background seismicity and high ambient noise where the goal is to detect small known explosions.

Additionally, I will also analyze data from the M6.3, 3 September 2017 North Korean nuclear test recorded at the KSRS array near Wonju, South Korea. Because ambient noise for this large explosion event is not an issue, detailed CWT processing will be able to characterize the quality of wave propagation in the Korean peninsula which may inform future simulations of wave propagation in scattering media.

3.5.1 AFTAC Sources from Shotpoint 1

Shotpoint 1 was used to detonate a 500lb shot and two 250lb shots in addition to the 2000lb shot extensively used in this report (Table 1). The 500lb shot, shot 1-2, was large enough to be easily detected and processed using CWT soft thresholding (Figure 23). The somewhat smaller size results in smaller signal amplitudes, relative to the background noise. By definition, soft or hard thresholding allows about 1% of the noise to come through in the result using the empirical cumulative distribution function method of estimating the noise level. These artefacts become more noticeable when signal amplitude decreases. This is apparent in Figure 23a for the seismic trace farthest from shotpoint 1-2 that shows what looks like a sinusoidal waveform at about 15s arrival time.

This is clearly a high scale (low frequency), single wavelet coefficient artefact that could be interpreted as a surface wave. However, nearby stations do not show this false arrival so it does not contribute in a significant way to the beam. Many other stations, however, do show later arriving sinusoidal arrivals after 30s arrival time from the source that are not strictly artefacts but do represent only partial recovery of the fundamental mode Rayleigh wave.

Broadband f-k responses for the denoised data are consistent with the results obtained from the 2000lb shot. Overall, there is a back azimuth anomaly of about 12° for the surface wave arrivals, as seen in the BBfk for 1-15Hz (Figure 23b) and the high frequency bandpass shows distinct P and S arrivals (Figure 23c). The BBfk images are well-focused and do not show much in terms of the ambient noise which was taken out through soft thresholding.

It is interesting to compute the BBfk for data that have not been denoised but simply filtered in these respective passbands (Figure 24). The BBfk response for 1 – 15Hz data shows the same surface wave peak from the explosion as from the denoised data of Figure 23. However, the ambient noise field, consisting of low velocity surface waves, shows up as energy propagating from the south. This is quite consistent with ambient noise sources from the large power wind farm located just a few km south of the experiment. BBfk of the high frequency bandpass shows a strong P wave arrival but a somewhat weaker S wave (Figure 24d).

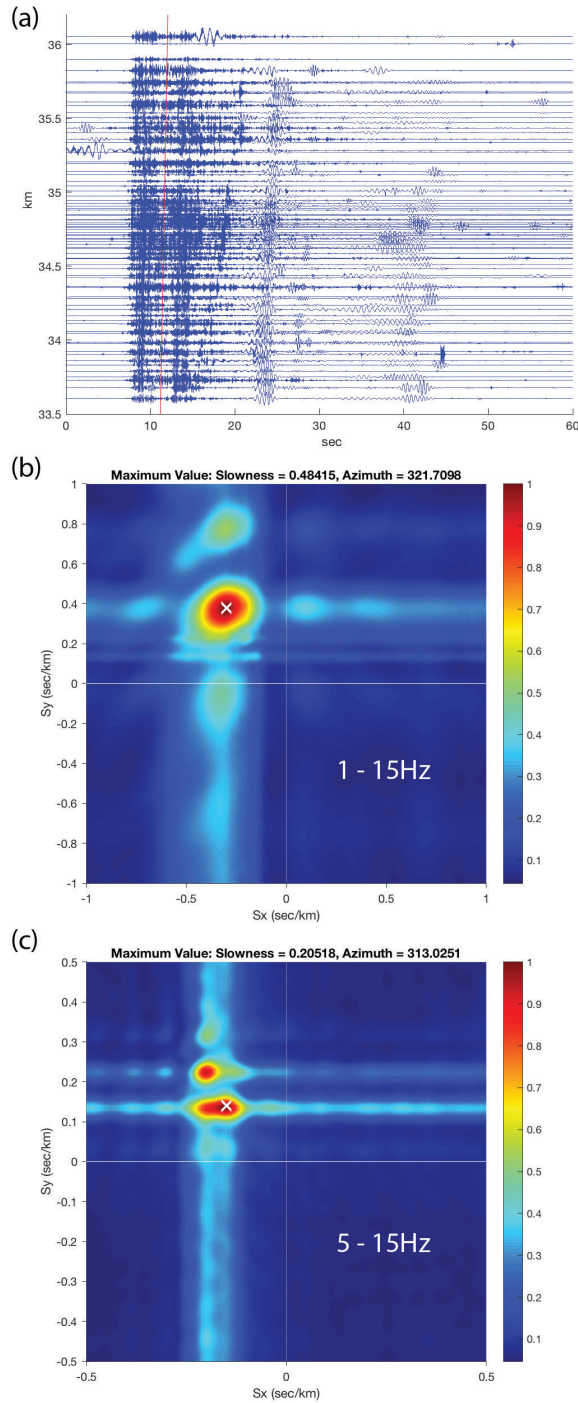


Figure 23 – (a) Data recorded by the cross array for shotpoint 1-2 have been denoised using ECDF soft thresholding with a scale band reject for scales greater than 1s. (This is used for all denoising in this section). The data show artefacts from the denoising process within waveforms, such as the arrival near 15s at the farthest station. The Rayleigh wave after 30s arrival time is only partially recovered. The red line denotes a group velocity of 3km from the source. (b) shows the BBfk result for the 1 – 15Hz bandpass of the denoised data. Surface waves dominate the lower parts of the frequency band. (c) shows the BBfk result for the 5 – 15Hz bandpass. P and S wave trains are clearly imaged.

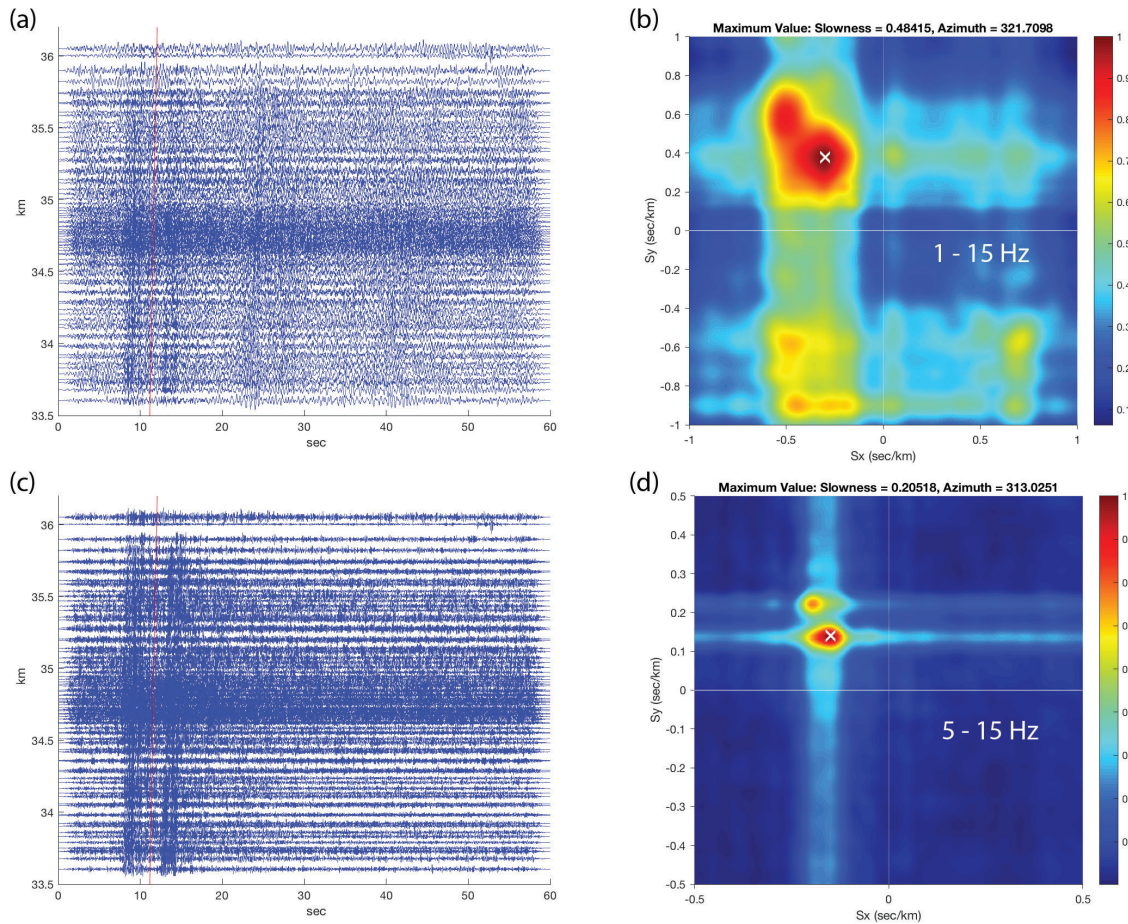


Figure 24 – (a) Data from shotpoint 1-2 that have not been denoised. A 2-pole, phaseless butterworth filter for the bandpass of 1 – 15Hz has been applied to the data as well as a 5% Hanning taper to both ends of the time series. (b) shows the resulting BBfk response. The ambient noise field created by a power wind farm south of the array shows up as low velocity surface wave energy from southern directions. (c) shows the original data filtered in the 5 to 15Hz bandpass, with (d) being the BBfk response.

Denoising clearly removes most of the ambient noise field revealing discrete arrivals in the explosion waveforms (Figure 23), compared to simply bandpassing the data (Figure 24). In fact, the denoised waveforms show that a small high frequency microearthquake occurs in the surface wavetrain of the explosion but this event is difficult to see even for the 5 -15Hz bandpassed data. Overall, the explosion is large enough to be easily seen in both sets of BBfk responses from the denoised and un-denoised waveforms. Using CWT denoising looks to be a useful tool in examining the nature of the ambient noise field as well as the signal.

Data from the 250lb shots are dominated by background seismicity and non-stationary ambient noise sources (Figures 25 and 26). After denoising, cross array data from shotpoint 1-3 shows significant high scale artefacts and high velocity P and S arrivals from a microearthquake under the array (Figure 25). The 15 – 30Hz BBfk clearly shows the microearthquake and no waves from the explosion.

Data from shotpoint 1-4 are completely dominated by southward and northward propagating, very slow, high frequency phases (Figure 26). These are probably local surface waves created from industrial activity at the wind farm and traffic on the roads around the array. There is no hint of the AFTAC explosion in the 5 – 30Hz BBfk responses in these data.

3.5.2 M6.3 North Korean Nuclear Test Recorded by KSRS

Data for the DPRK nuclear tests recorded at KSRS were obtained to perform a reconnaissance study of the use of CWT array techniques on a CTBTO array, more typical than the spatially oversampled IRIS Wavefields Experiment array (Greg Wagner, personal communication 2019). KSRS is a borehole array consisting of 19 vertical instruments and a reference three component broadband sensor (Figure 27). The theoretical BBfk beam patterns are quite symmetrical attesting to good array design (Figure 28). The data from the nuclear tests are of relatively high SNR so that CWT denoising only has marginal improvements over standard Fourier filtering.

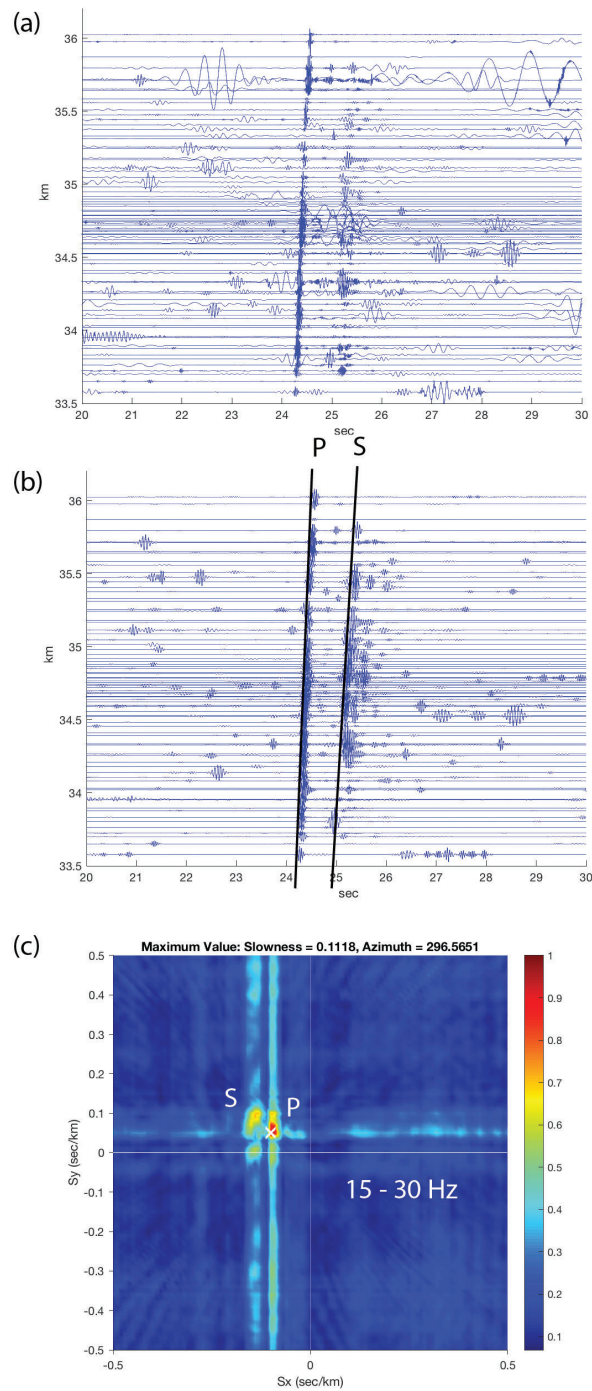


Figure 25 – (a) Denoised data at the cross array for the 250lb explosion at shotpoint 1-3. The denoised seismograms show numerous high scale artefacts but are dominated by high velocity P and S phases from a small local earthquake. (b) shows the denoised data filtered between 15 and 30Hz. (c) BBfk of the filtered data. The P and S phase characteristics are highly focused.

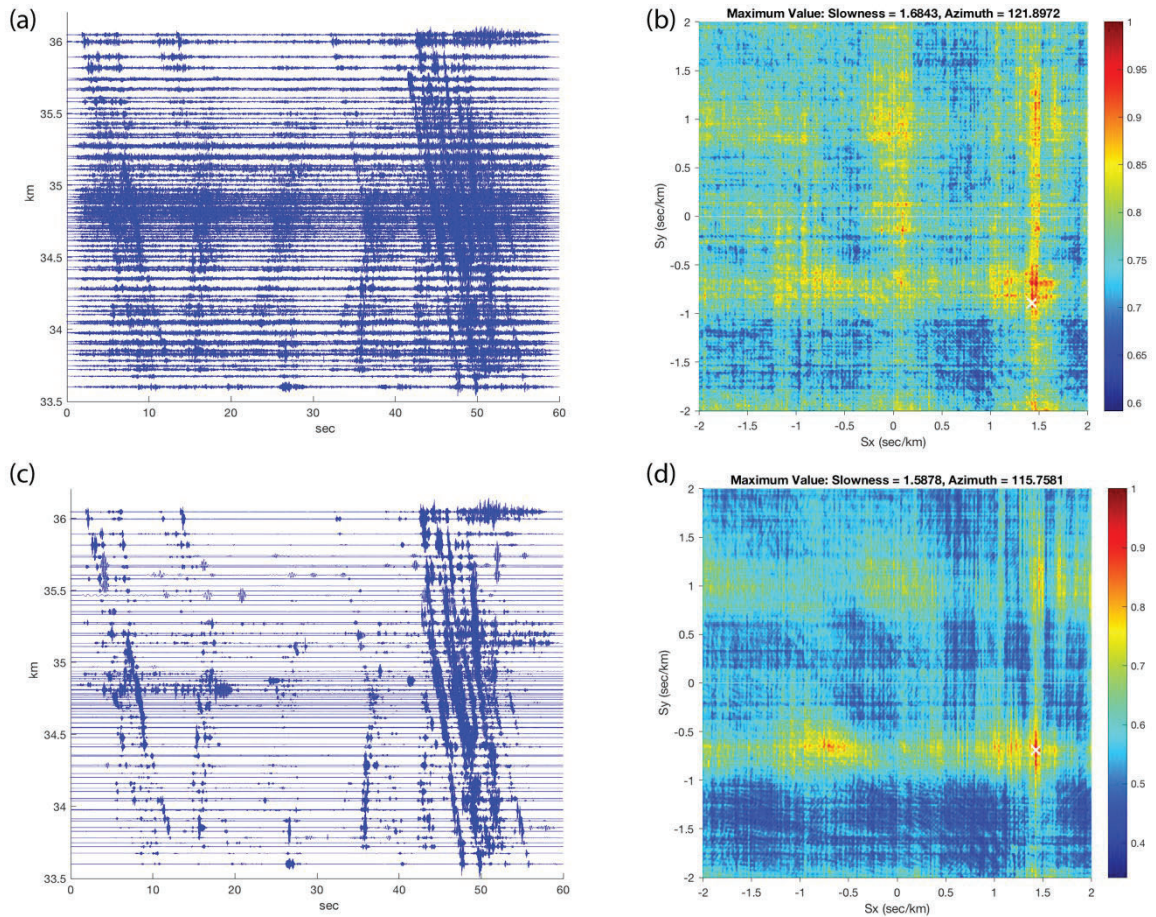


Figure 26 – (a) Data recorded at the cross array for the time of explosion 1-4. The data have been bandpassed filtered between 5 and 30Hz. There is very little evidence of the AFTAC explosion. Local noise source from north and south of the array dominate the seismograms, as seen by their moveout in time. (b) BBfk of the filtered data of (a). (c) Data after soft threshold denoising. The larger impulsive sources are much clearer although the smaller amplitude sources have been subdued by denoising. (d) BBfk of the denoised signals appear somewhat clearer than (b) and suggest sources from the south associated with the wind farm.

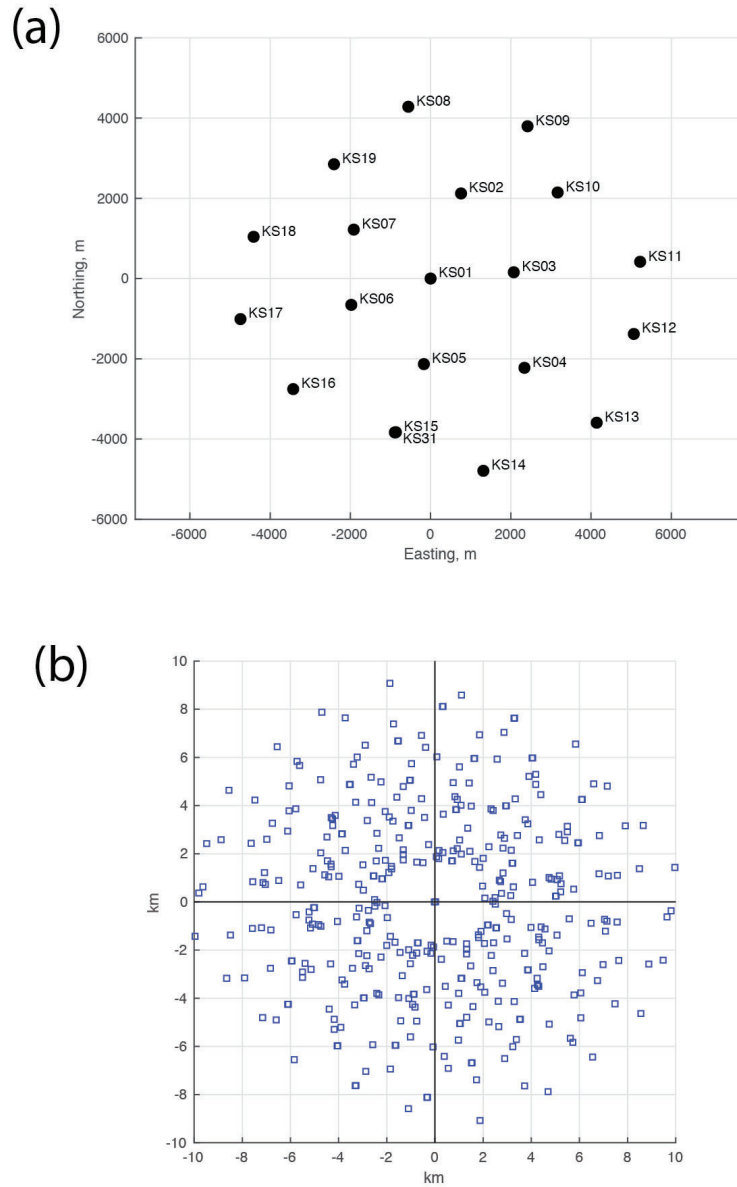


Figure 27 – (a) Geometry of the CTBTO KSRS array located near Wonju, South Korea. The co-array diagram (b) shows good azimuthal performance for horizontal wavelengths from 2 to 20km. The co-array is a plot of all possible distances and azimuths between array elements showing how the array can sample a wavelength in space.

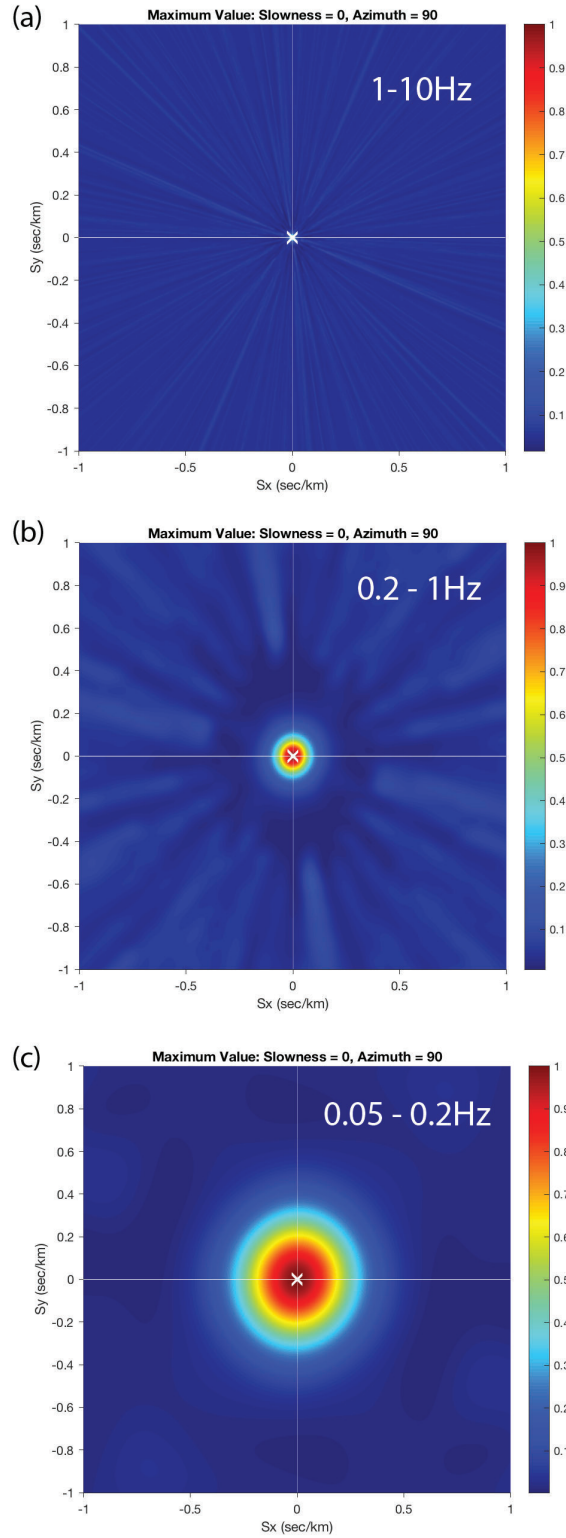


Figure 28 – Theoretical broadband f-k responses for the KSRS array for frequency bands of 1-10Hz (a), 0.2-1Hz (b), and 0.05-0.2Hz (c).

The KSRS data from the largest explosion of 3 September 2017 were analyzed using CWT beamforming to examine the stability of regional phases within the waveforms (Figure 29). Ideally, assuming a vertically inhomogeneous Earth structure, all phases should arrive from the direction of the great circle backazimuth and any scattering should be confined to this backazimuth.

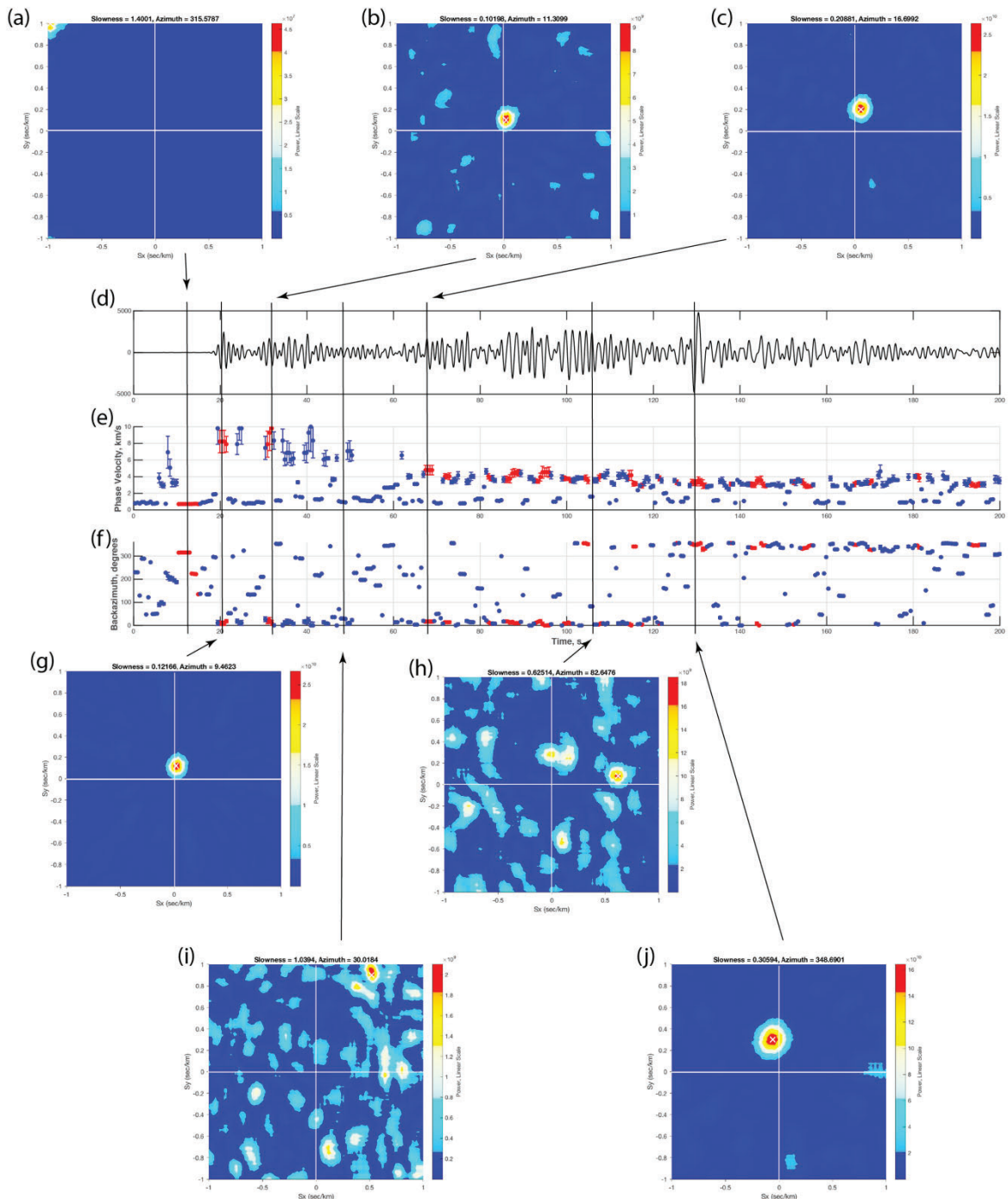


Figure 29 (previous page) – CWT beam forming for the 3 September 2017 M6.3 nuclear test. The middle panels show the reference seismogram (d) at station KS01 near the center of KSRS. The time series below the seismogram show phase velocity (e) and backazimuth (f) as a function of time. CWT beamforms are indicated by the arrows to particular times. Backazimuth and phase velocity measurements are sifted into $R \geq 0.5$ (red) and $R < 0.5$ (blue) values. Low “R” values correspond to beamforms that have distributed energy across the slowness plane as in (h) and (i). Well focused beams generally have a single peak as in the other plots. Note that the P wave train (b) and (g) and the beginning of the S (c) have backazimuths close to the great circle path which is N12°E. Later S arrivals generally have backazimuths of roughly N10°W. The R metric breaks down where the time series is near zero as in (a).

The data show particular arrivals in the wave train that are well focused and arrive with the correct backazimuth. However, most of the seismogram, judging by the number of blue points vs red, consists of scattered arrivals where the beam forms show distributed energy in azimuth and phase velocity. Much of the scattered energy appears with relatively large slowness (low velocity) suggesting that scattering may be more local to the array compared to the distance from the explosion source. Furthermore, later well-focused arrivals in the S wave train tend to have backazimuths to the west of north compared to the great circle path which is east of north. This change in backazimuth within the S wave train suggests that the causative structure may be smoothly changing between source and receivers and not just under the array, such as with a local dipping crustal interface. Information like this could be useful in future studies of lateral heterogeneities of the crust and mantle.

3.6 Using Empirical Source Templates as Mother Wavelets in the Continuous Wavelet Transform

The continuous wavelet transform (CWT) represents an interesting theoretical as well as phenomenological connection with the straightforward process of template correlation. This can be seen from the definition of the CWT (Grossman et al., 1989; Starck et al., 2010) where

$$W(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t - \tau}{a} \right) dt \quad . \quad (1)$$

In equation (1), $W(a, \tau)$ is the wavelet time series for scale, a , $f(t)$ is the function to be transformed, and $\psi^*(t)$ the complex conjugate of the mother wavelet. For any particular scale, a , equation (1) is simply the correlation of $f(t)$ with the scaled wavelet. From this

point of view, correlation of $f(t)$, a test seismogram, with a real-valued mother wavelet, the event template, yields a wavelet time series for scale $a=1$. Provided the mother wavelet has zero mean, equation (1) can form a valid transformation of $f(t)$ into the scale-time plane for a spectrum of scales, a . In practice, an event detection occurs when the correlation produces one peak over the correlogram. In CWT parlance, this peak represents the amplitude of one wavelet coefficient representing the entire wavelet used to derive it. It also represents, perhaps, the sparsest representation of the hidden event by being the amplitude of the small event waveform relative to the large template event. The entire seismogram has been reduced to one number, an excellent compression of information.

In this section, I extend this idea by creating an empirical mother wavelet based on a known, denoised event time series and computing the CWT of a test time series using this mother wavelet. Data come from the AFTAC explosion experiment recorded by the 2016 IRIS Wavefields Community experiment in northern Oklahoma described elsewhere in this report. In particular, I will concentrate on events from shotpoint 1 approximately 35km to the northwest of the IRIS arrays. The distribution of explosion sizes from 2000lbs to 250lbs with known origin times allows for unambiguous interpretations to be made from the results.

3.6.1 The Empirical Wavelet

I choose to construct the empirical wavelet from the denoised vertical component of station 3016 for the 2000lb explosion detonated at shotpoint 1. There is nothing special about choosing this station other than it was a reference station used in the array analyses of previous sections of the report. Mother wavelets may be entirely real functions but there are aspects of actual wave propagation that are affected by changes in phase of the waves. For example, Rayleigh waves are not rectilinear but have elliptical particle motions. Small changes in distance from the source will cause Rayleigh waves to change phase differently compared to a rectilinear wave, such as a P wave. Having a mother wavelet with the ability to model the change in phase with scale is desirable.

In analogy to the exponential in Euler's relation

$$e^{i\phi(\omega)} = \cos\phi(\omega) + i\sin\phi(\omega) \quad (39)$$

I use the analytic signal (e.g., Farnbach, 1975) to define an empirical mother wavelet by

$$\begin{aligned}\psi(t) &= w(t) - iH[w(t)] \\ &= |\psi(t)| e^{i\phi(t)}\end{aligned}\quad (40)$$

where, $w(t)$ is the empirical template time series taken for an event using the vertical component of motion, $H[\]$ is the Hilbert transform, the $\psi(t)$ is the mother wavelet. The phase, ϕ , in Euler's relation is given by

$$\phi(\omega) = \tan^{-1} \frac{\text{Im } e^{i\phi(\omega)}}{\text{Re } e^{i\phi(\omega)}} \quad (41)$$

and the instantaneous phase in the analytic signal is given by

$$\begin{aligned}\phi(t) &= \tan^{-1} \frac{\text{Im } \psi(t)}{\text{Re } \psi(t)} \\ &= \tan^{-1} \frac{-H[w(t)]}{w(t)}.\end{aligned}\quad (42)$$

These relations are not used directly in computing the CWT but they do demonstrate that signal phase can be represented by a complex mother wavelet.

The empirical mother wavelet used here is constructed from the vertical component denoised data from station 3016 in the cross array (Figure 30). As a reminder, the data were denoised using the ECDF method with soft thresholding proposed by L&M with an additional scale band rejection for scales greater than 1s. A scale parameter a of 1s gives the original denoised waveform for the real part of the wavelet. Greater or lesser scales stretch or contract the waveform according to the definition of the CWT, equation (1) (Figure 30).

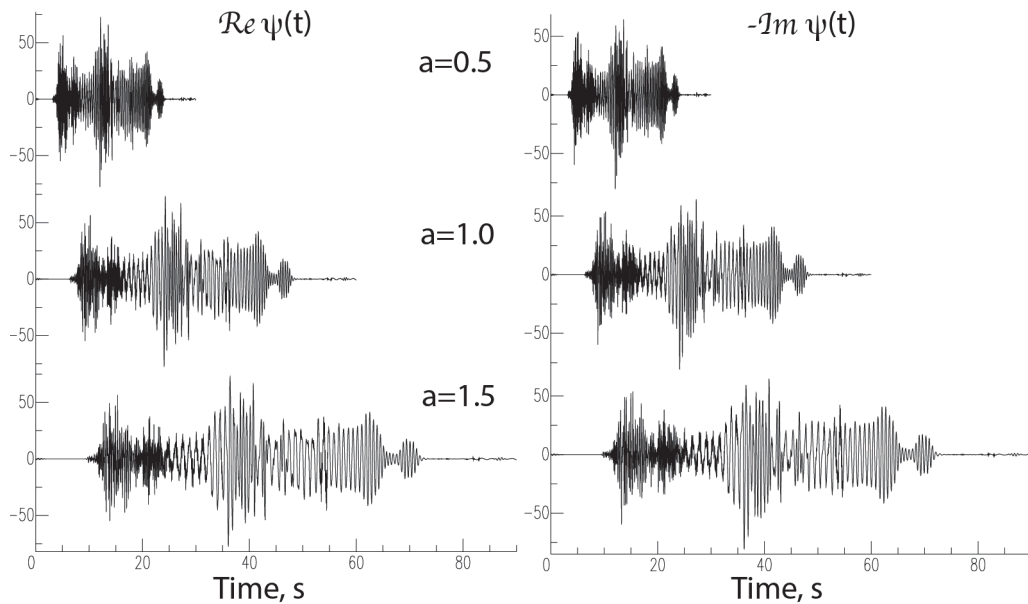


Figure 30 – Real (left) and negative Imaginary (right) parts of the empirical wavelet constructed from the 2000lb shot at shotpoint 1 for vertical component data recorded at station 3016 in the cross array. The imaginary part is derived from the real part through Hilbert transformation according to equation (40) and is, by construction, 90° out of phase from the real part. Waveform scaling for $a=0.5$, 1.0 , and 1.5 s are shown. Waveforms stretch with increasing scale.

3.6.2 Data

The AFTAC explosion experiment provides for an excellent data set to investigate extending the template correlation method with the CWT. The locations and origin times are known with great accuracy (Table 1). Also, the experiment itself gives direct evidence that the waveforms for all explosions at a particular shotpoint should nearly be identical (Figure 31). Even though the source waveforms from shotpoint 1 are buried in the noise, my assumption is that they are identical, similar to what is seen with the events from shotpoint 2 (Figure 31).

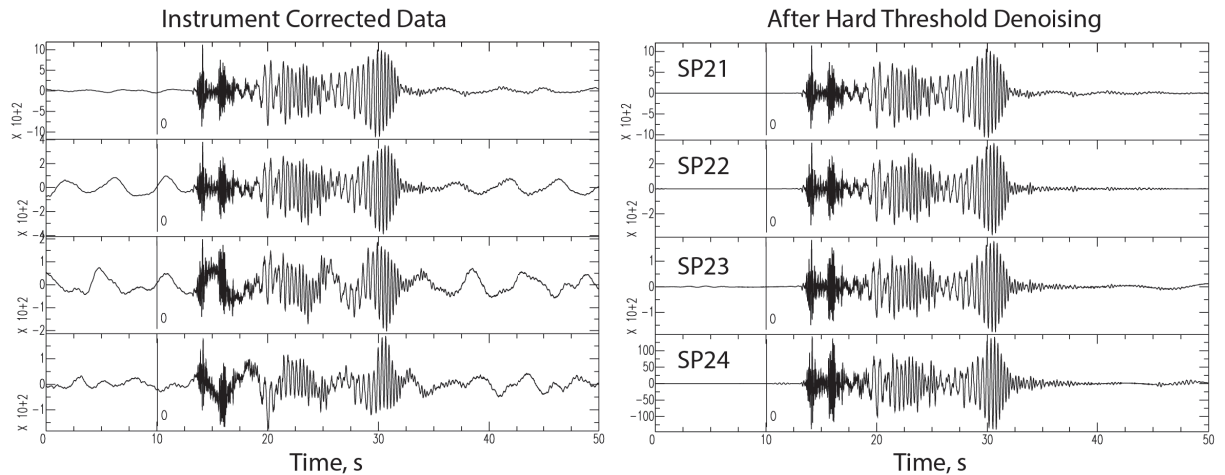


Figure 31 – Comparison of the instrument corrected vertical component data at station 3016 (left panel) and hard thresholding denoised data (right panel) for the explosions from shotpoint 2. SP21, SP22, SP23, and SP24 are 2000lbs, 500lbs, 250lbs, and 250lbs of explosives, respectively. The waveforms are effectively identical.

The data from shotpoint 1 were corrected for instrument response and time windowed such that the expected correlation peak from each event would occur at the same time of 60s in the correlogram. This consistency in time guarantees that results can be judged to be successful or not.

3.6.3 Results

The first experiment is to investigate the characteristics of the CWT on the scale-time plane for the original denoised seismogram used to construct the empirical wavelet (Figure 32). The scalogram shows a very narrow peak at a scale of 1s and zero time shift indicating that the empirical wavelet has clearly found itself on the scale-time plane. In computing these transforms, scale was limited to values between 0.5 and 1.5s representing a 50% contraction and 50% expansion, respectively, of the original template waveform. From a physical point of view, this might mean corresponding decreases and increases in the relative times of arrival of the major phases in the seismogram. The scale interval was also linearly sampled to give 64 values, also called “voices”.

Applying the empirical wavelet to the original noisy data gave a transform with the exact time and scale coordinates expected for the SP11 event (Figure 32) showing that the noise was not much of an impediment to the process. Previous sections of this report have shown that the denoising procedure essentially finds those portions of the scale-time

plane that have high SNR. Thus, using the high SNR data as the empirical wavelet guarantees that it will correlate quite well with the original data.

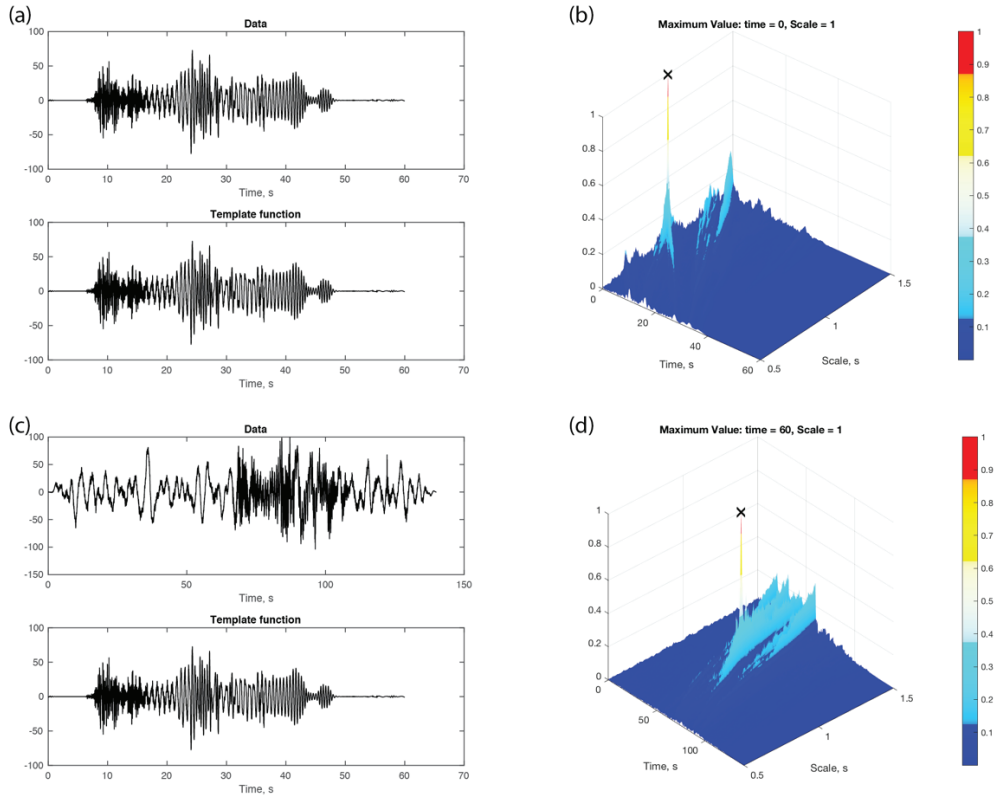


Figure 32 - Results for shotpoint 1, explosion 1 (2000lb explosion). In this and next 3 figures, the data time series is shown as the top seismogram as in (a) and (c). The real part of the template mother wavelet for scale $a=1$ s is shown below the data. The right panels (b) and (d) show a 3D perspective view of the scale-time plane (scalogram). The black “x” shows the position of the peak value. (a) and (b) show the result of taking the CWT of the original template. (c) and (d) shows the result for the original noisy data, corrected for instrument response, for the 2000lb explosion.

Explosion 2 at shotpoint 1 was only 500lbs in size and is not discernable to the naked eye in the instrument-corrected data (Figure 33). However, the scalogram displays a highly focused peak at the correct time and scale showing that the event has been detected using the empirical wavelet. Comparing Figures 32d and 33b, it can be seen that the floor of the background noise is more diffuse and higher amplitude for explosion 2 than for explosion 1.

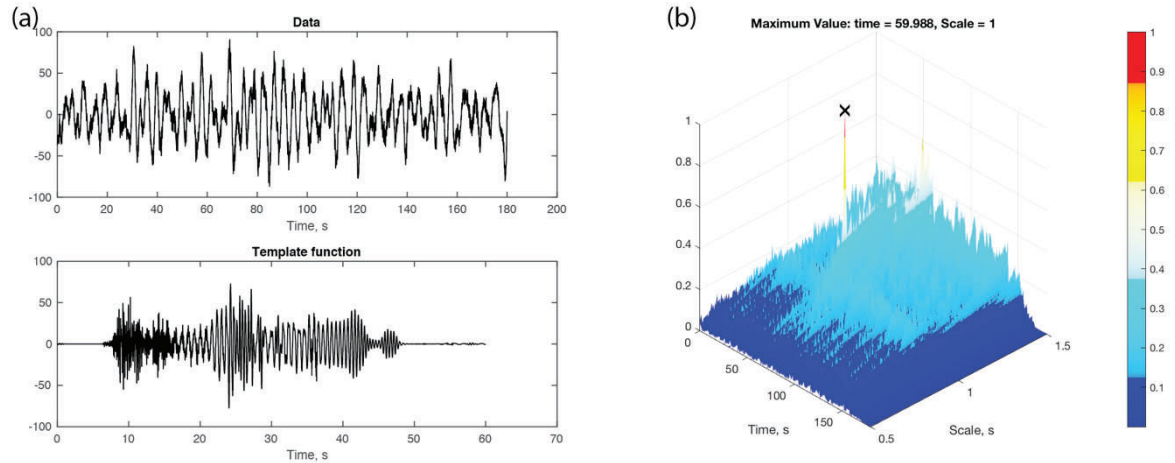


Figure 33 – CWT of explosion 2 (500lbs) using the explosion 1 mother wavelet. Same scheme as Figure 32.

Explosions 3 and 4 were only 250lbs in size. Application of the empirical mother wavelet to time series containing these explosions shows no indication of detection at the correct lag time of 60s and scale of 1s (Figure 34). Instead, peaks on the scalogram occur at higher scale which suggests that lower frequency components of the empirical wavelet are correlating with the background noise which is lower frequency than the signal of interest. Based on the relative amplitude scaling of explosions at shotpoint 2 (Figure 31), I expect that waves from the 250lb explosions should be 15% to 20% the size of waves from the 2000lb explosion at shotpoint 1. Detection fails for the small explosions at a distance of 35km.

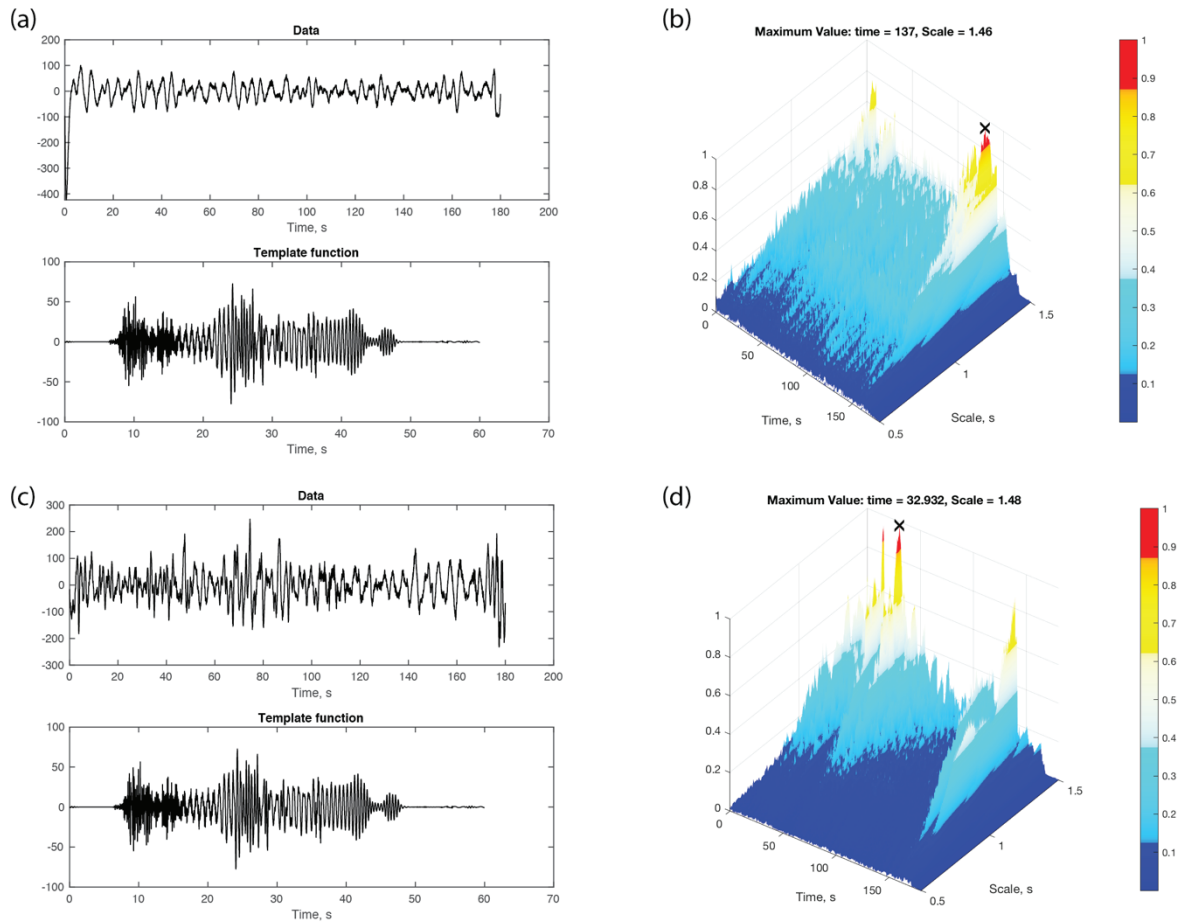


Figure 34 – Empirical wavelet CWT for explosions 3 (a) and (b), and 4 (c) and (d) at shotpoint 1. Same scheme as Figure 32.

Application of the empirical wavelet to other stations of the IRIS experiment illustrates the performance of the CWT method in trying to capture wavefield variations due to wave propagation phase changes (Figure 35). Stations within 100m of station 3016 show highly peaked scalograms similar to station 3016 itself indicating that the underlying waveform has not changed much over short distances (this is not shown here). However, as distance from station 3016 increases, the scalograms show that a single peak becomes a series of peaks along discrete directions on the scale-time plane (Figure 35). These linear streaks are a result of the stretching of a template seismogram that has 3 arrivals – P wave train, first higher mode, and fundamental mode Rayleigh – and correlating with the observed seismogram that has similar but differently phased arrivals. Evidence of these linear streaks can be seen in the self-test data of Figure 32. Station 3001, approximately 1.5km from station 3016 displays a peak close to the expected time

(59.6s) and scale (0.98s) but also shows significant spreading of wavelet coefficients along one streak direction. Station 1001 is the westernmost station of the IRIS experiment approximately 7km from station 3016. The scalogram shows a wall of high amplitude peaks spread over scale, some that are close to the nominal scale of 1.

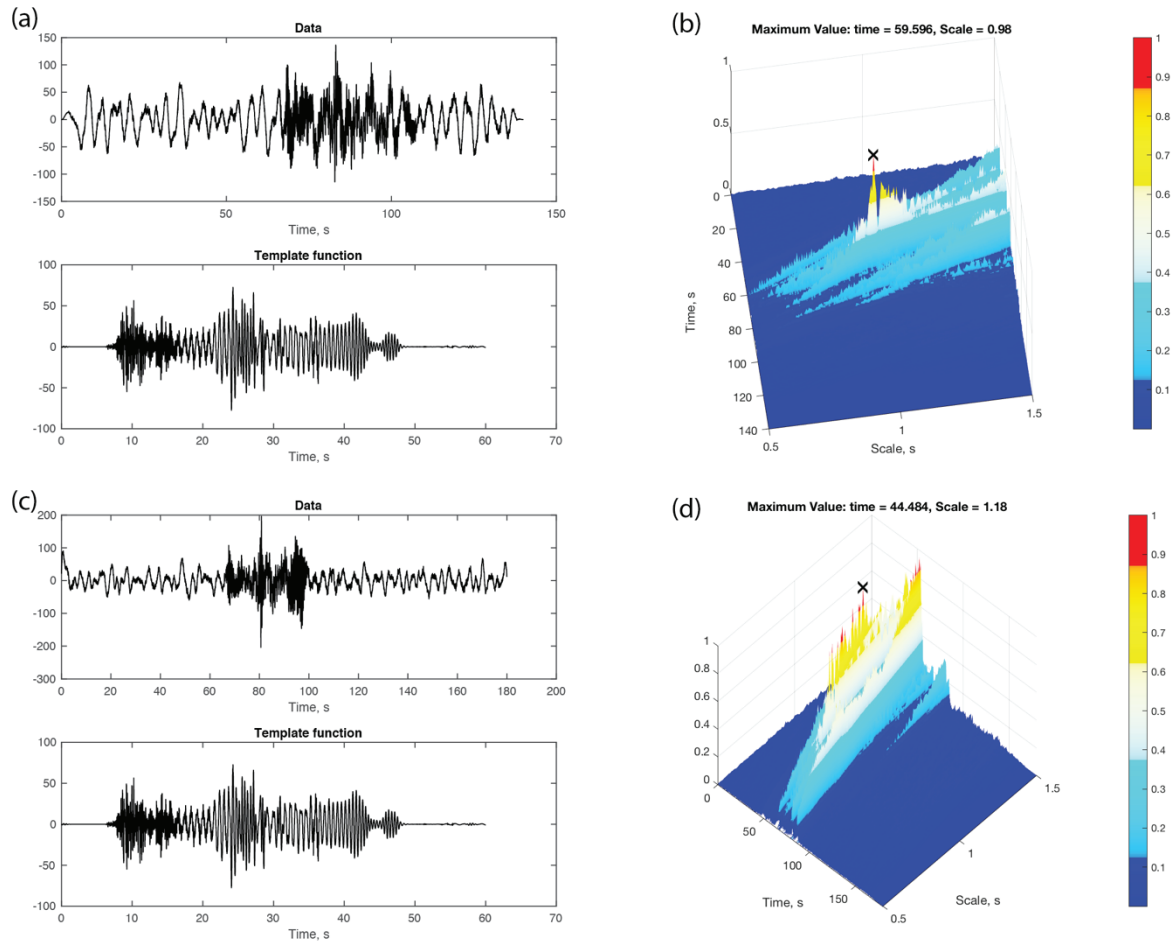


Figure 35 – Same scheme as Figure 3 showing CWT results for explosion 1 recorded at station 3001 ((a) and (b)) and station 1001 ((c) and (d)).

4. DISCUSSION

The use of time-frequency representations of seismic signals literally offers another dimension to view and manipulate time series of interest. The techniques outlined here using the continuous wavelet transform can improve the ability to recognize those regions of the scale-time plane in the wavelet domain that contain high SNR portions of wanted seismic signals. Finding these high SNR data spaces is an important step towards optimizing a processing flow for detection and analysis of small seismic events in geographic regions of interest using seismic arrays. The additional capabilities of investigating the wave composition of seismograms by separating seismic phases through block manipulation is a useful tool for studying details of wave propagation and source excitation in phased array processing. It should also be helpful in understanding the structure effect when modeling data using synthetic seismograms.

Because using the CWT in seismological signal analysis is not commonplace, a typical question concerns the amount of distortion that noise thresholding and block manipulation might introduce in the filtered seismograms that could affect correlation across a phased array. The answer to this question is that “it depends”. Any signal processing method has limits. My experience with CWT techniques is that noise thresholding for high SNR portions of the scale-time plane does little to degrade time domain waveforms across a phased array. Denoised waveforms from the 2000lb explosion displayed in this paper showed very high correlation across the dense nodal array. Array beams using standard frequency-wavenumber methods and CWT beam forming often approached the theoretical beam response of the array showing that seismic phase waveforms were highly similar (e.g., Figure 9). However, I tried the same analysis for the 250lb shot near the shotpoint of the 2000lb explosion with negative results. Signal for the smaller explosion was comparable to the noise, even in that part of the scale-time plane where the 2000lb shot had high SNR, so that soft thresholding removed both signal and noise. It can be expected that signal distortion will occur from thresholding when the SNR is relatively low and not much greater than one (1).

Likewise, care should be taken when performing the signal decomposition step to separate seismic phases. A polygon can be arbitrarily drawn around any area on the scale-time plane of a signal. Clipping out an arbitrary portion of the CWT will yield a

time domain waveform that could look interesting but be completely meaningless in terms of reflecting some aspect of real wave propagation. Signal decomposition should be based on robust features of the CWT that have high energy and distinct attributes that isolate the feature from others. This is in keeping with the philosophy of “sparsity” (Starck et al., 2010) where an attempt is made to describe a signal with the minimum amount of information. A classic Fourier example is the description of a long duration time domain sinusoid by its single frequency in the frequency domain. In the case of the explosion Rayleigh wave phases observed here, the fundamental mode portion of the seismogram was generally separate from other portions of the CWT on the scale-time plane and showed dispersion through decreasing scale with time. Perhaps a better way to think about the phase decomposition analysis is that it is a hypothesis test where the hypothesis, the separated signal, is analyzed by other means to test its usefulness. However, there is no doubt that scalograms of interesting signals have similar attributes as seen across a phased array. This alone suggests that these new signal attributes could be incorporated into pattern recognition or machine learning algorithms.

Signal decomposition using CWT methods will also help in a common problem encountered in wave gradiometry. A geodetic array (Spudich et al., 1995) is a variant of a phased array where the array aperture is a small fraction of the target wavelength. Provided that there are a sufficient number of elements of the array that have matching instrument responses, wave attributes are inferred by estimating wave spatial gradients by representing the observed local wavefield in terms of a Taylor’s series in displacement. The inferred first order spatial gradients are, of course, the strains and rotations induced by wave motion under the array. The spatial gradients are related to the time gradient (velocity) through the wave equation so that wave horizontal slowness, azimuth of propagation, and other attributes can be estimated directly from amplitudes (Langston, 2007a; b; c; Langston and Liang, 2008). A central assumption in the method is that a single wave is the object of analysis. Obviously, this may not be true for arbitrary wavefields. Being able to separate waves in a seismogram using the CWT and block manipulation has the potential to improve gradiometry wave attributes since the disrupting effects of wave interference can be minimized. This is the subject of ongoing work.

It was relatively easy to find a simple, yet robust model that fit the fundamental mode dispersion data from both explosions (Figures 17 and 18). It is also gratifying that the models are not significantly different, considering the errors in the data, since both sets of data apply to structure directly under the IRIS experiment. The models show that structure in the upper half km is best resolved by the linear gradient node representation.

Although the fundamental mode is well fit, it proved quite difficult to find a self-consistent model that also fit the presumed first higher mode data. Several hundred forward model calculations were made to find a model to do so. Models included multi-node as well as multi-layer models based on the Kirk well log. Figure 36 shows the predicted fit of the fundamental mode model of Figure 17 to the first higher mode data for the shot point 1 explosion. Every model attempted underpredicted the phase velocity of the first higher mode as a function of frequency.

Figure 16 compares the 1D fk image with the predicted phase velocity curves for the fundamental, first, and second higher modes of the 2000lb explosion at shotpoint 2. The vertical component data shows that the observed higher mode energy generally falls between the first and second higher mode prediction. Perhaps this higher mode energy is almost pure second higher mode?

To check this hypothesis, 1D fk dispersion was determined using the radial components of motion for the shot point 2 explosion (Figure 16 – right). It seems that the observed mode energy may be from the second higher mode after all. There still seems to be a remnant of the first higher mode between 3 and 4 Hz in the radial 1Dfk, however, that falls between the two curves.

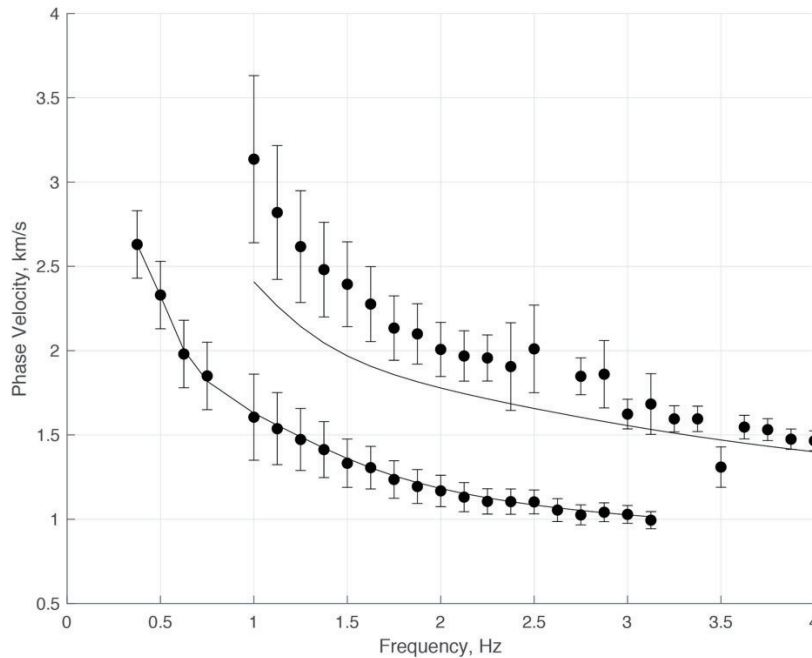


Figure 36 – Shot point 1 fundamental and first higher mode phase velocity dispersion data with predictions from the fundamental mode inversion model. The higher mode (upper curve) is clearly too fast over most of the frequency band.

Frechet derivatives of phase velocity as a function of depth were calculated at frequencies of 1.5 and 3.5Hz (Figure 37). They clearly show that the first higher mode is most sensitive to structure above 0.2 km at 1.5Hz but all modes are equally sensitive at higher frequency. The detailed array analysis made in previous sections showed that multipathing and lateral refraction becomes important for Rayleigh waves above 2.5 Hz. This suggests that the gradient zone near the surface is irregular in space causing modes to scatter and for waves to refract horizontally. Changes in the top gradient zone may affect the first higher mode more strongly compared to the fundamental and second higher mode based on how the plane layer derivatives behave.

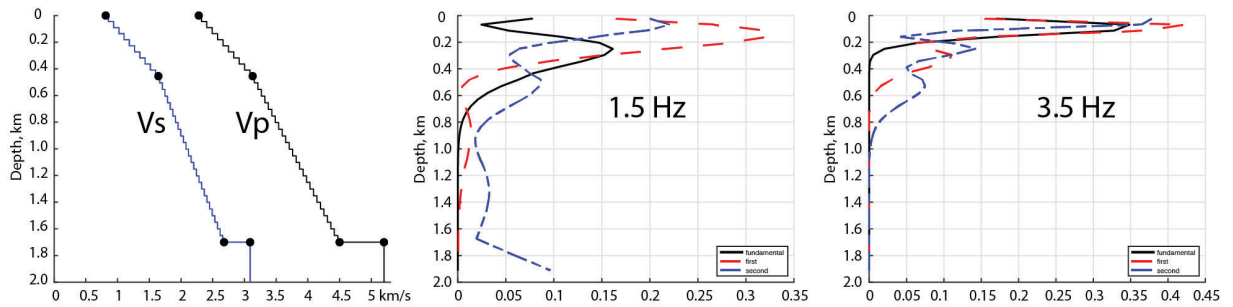


Figure 37– Frechet derivatives for the velocity model shown in the left panel for 1.5 Hz (middle) and 3.5 Hz (right). The fundamental mode derivatives are shown by black solid lines, the first higher mode by the red dashed line, and the second higher mode by the blue dashed line.

While this seems like an esoteric problem in wave propagation for this particular structure for this particular part of the world, I think there is some importance to investigating it further. From an Earth structure point view, it may be possible to map out lateral heterogeneities related to changes in geology and near-surface velocity structure using array-derived mode behavior. In addition, details of the source excitation of the modes and how they scatter in the waveguide could be important in explaining the creation of Lg and Lg coda, as well as the conversion of P waves into shear waves from explosions. An approach to this problem would be to employ numerical methods for wave propagation in 3D media. At a minimum, it is important to know the limitations of using standard techniques, such as dispersion measurements or synthetic seismogram computations for plane layered media to model local earthquake and explosion waveforms.

Much of the work spent on this project went into computation and inversion of mode dispersion to obtain velocity models simply because the data were dominated by modes. However, the body wave data are an equally rich source of information, as Figure 15 demonstrates. The P wave group shows a clear head wave moving out in front of higher amplitude critical reflections from interfaces in the shallow waveguide. S waves are also prominent begging the question of how they were generated by the explosion source. The continuity and correlation of these high frequency waveforms over 12 km is a testament to the high quality of the nodal instruments and their installation.

Figure 38 shows the result of τ - p “inversion” or the application of the radon transform. The data of Figure 15 have been slant stacked to derive the τ function, or zero distance intercept, as a function of ray parameter, p . This function has physical meaning for turning rays and discrete ray arrivals and can be used to derive a velocity model for the structure. The P wave group is quite strong and shows the high velocity head wave (P1) with secondary arrivals P2, P3, P4 showing progressively slower phase velocities or higher slowness. These waves probably are trapped in the sediment waveguide. The P5 phase may be a shallow reflection or lower order multiple compared to P3 or P2.

The radial component slant stack shows S1 as dominating the response. Notably, even though there are S phases in the vertical component, S1 does not show up prominently at all. These observations suggest that there is considerable richness in the body wave data set that can be used to develop a more refined model than the dispersion model that may approach the detail of structure seen in the sonic log.

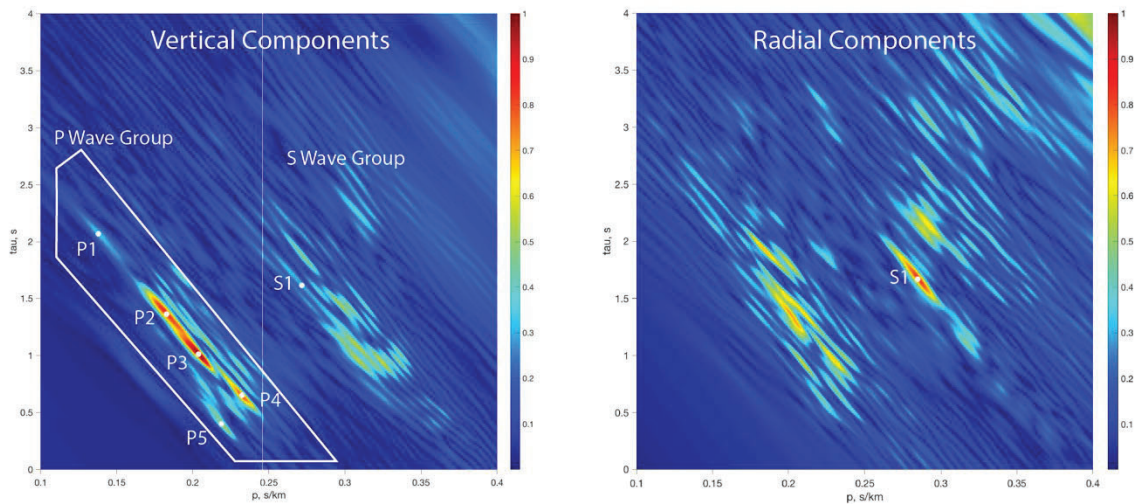


Figure 38 – The body wave portion of the radon transform of the data displayed in Figure 15 for the 2000lb shot at shot point 2. The transform for the vertical components is dominated by waves in the P wave train. Various phases are annotated within the P Wave Group polygon shown in the left panel. Transform of the radial components (right panel) shows the S arrival (S1) clearly. Note that S1 is not very clear in the vertical component radon transform.

Using an empirical wavelet as the mother wavelet at a scale of 1s is equivalent to results obtained with template correlation. From a theoretical point of view, it is interesting to note that an unknown seismogram can be reconstructed from a basis

function derived from a reference seismogram. In a sense, the CWTs shown here are “explosion 1 CWT transforms” that can be used to reconstruct the unknown signal. For example, the CWT transform for station 3001 shown in Figure 6b could be thresholded by choosing only the highest peaks in the scale-time plane. Inverse transforming should recover an estimate of the explosion signal at the station without noise. Essentially, correlation is used to detect an event and then to reconstruct its waveform from the behavior of the CWT while simultaneously removing the noise from the seismogram. It is conceivable that the reconstructed waveform will contain important wave propagation information, such as the timing between P and S phases as well as relative amplitude information that can be used for relative locations and focal mechanism studies. Use of the CWT extends the technique of template correlation to more than just detection. This is a very interesting topic that will be investigated in future work.

5. CONCLUSIONS

Phased array processing can benefit from application of CWT denoising and block manipulation techniques. Noise thresholding on the scale-time plane can be used to find high SNR parts of the CWT spectrum of a signal that can be further studied using standard frequency-wavenumber beamforming. Array seismograms can be separated into component phases using scale-time gating where specific seismic phase blocks in the CWT domain are correlated from a reference station to all other stations in the array. The separated seismic phases can then be analyzed using standard techniques or by constructing array beams by directly using the CWT.

These techniques were used on explosion data recorded by the 2016 IRIS Wavefields experiment to determine Rayleigh wave phase velocity and group velocity dispersion under a cross-shaped subarray and the wider full array. The 1-3Hz Rayleigh waves display horizontal refraction along the wave path from source to array and were used to develop a shallow velocity model for the upper 1.7km. In addition, near-source pP and sS reflections were detected in seismograms from a M3.6 local earthquake to constrain source depth to 4 to 4.4 km, consistent with a previously determining regional moment tensor model of the earthquake.

The explosion array data are a rich source of information on the propagation and excitation of the fundamental and higher mode Rayleigh waves affected by structure under the array. Inversion of the fundamental mode data yields an Earth model with a steep positive velocity gradient near the surface but is relatively insensitive to deeper structure in the sediment waveguide and crystalline basement. The array data suggest lateral heterogeneity in the shallow structure causing lateral refraction of Rayleigh waves and distortion of the first higher mode in phase velocity and amplitude.

Application of CWT methods to the low SNR data of shot point 1 is seen to yield robust results comparable to directly using profiles of high SNR data across the entire IRIS experiment from the shot point 2 explosion. This is a good reality check on the veracity of CWT noise thresholding, block manipulation through scale-time gating, and beam forming in the wavelet domain.

Application of CWT denoising methods to data from the AFTAC explosion experiment gives the sensible result that once signal approaches the background noise on the scale-time plane, it will not survive the thresholding process. Other sources, such as microearthquakes or impulsive, non-stationary artificial sources will be exposed at the expense of the desired signal. Thresholding was effective in removing the pervasive ambient noise field from the wind farm south of the array.

CWT beam forming of the M6.3 North Korean explosion suggests that the crust and mantle of the Korean peninsula contains both significant near-array scatterers as well as smoothly varying lateral structure along the path from source to receivers.

The CWT extends the well-known template correlation method for detecting small events buried in the noise. Template correlation is a special case of the CWT using the template as the real part of the mother wavelet at a scale of 1s. Creating a complex mother wavelet using the template plus its Hilbert transform was effective in detecting the 500lb explosion from shotpoint 1 using the 2000lb waveform as the template. Further work is needed to explore the possibility of recovering the waveform of small events by thresholding the energy around the correlation peak at a scale of one.

REFERENCES

- Abrahamson, N. A., B. A. Bolt, R. B. Darragh, J. Penzien, and Y.B. Tsai (1987), The SMART-I accelerograph array (1980-1987): A review. *Earthquake Spectra*, 3, pp. 263-287.
- Aki, K. and P. G. Richards (1980), *Quantitative Seismology - Theory and Method*,. San Francisco, CA, W. H. Freeman and Co.
- Al-Shukri, H., G. L. Pavlis, and F. L. Vernon (1995), Site effect observations from broadband arrays, *Bull. Seism. Soc. Am.*, 85, pp. 1785-1769.
- Anstey, N. A. (1966), Correlation Techniques - A Review, *Can. J. Expl. Geophys.*, 2, pp. 55-82.
- Aster, R. C., B. Borchers, and C. H. Thurber (2005), *Parameter estimation and inverse problems*, Elsevier, Inc.
- Brocher, T. M. (2005), Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. Seism. Soc. Am.*, 95, pp. 2081-2092, doi 10.1785/0120050077.
- Capon, J. (1969),. High resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, 57, pp. 1408-1418.
- Capon, J. (1970), Analysis of Rayleigh-wave multipath propagation at LASA, *Bull. Seism. Soc. Am.*, 60, pp. 1701-1731.
- Capon, J., R. J. Greenfield, and R. J. Kolker (1967), Multidimensional maximum likelihood processing of a large aperture seismic array, *Proc. IEEE*, 55, pp. 192-211.
- Claassen, J. P. (1985), Design criteria for sizing regional arrays, In A. U. Kerr (Ed.), *The VELA Program: A Twenty-Five Year Review of Basic Research*, pp. 506-515, United States, Dianne L. Carlson.
- Constable, S. C., R. L. Parker, and C. G. Constable (1987), Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, 52, pp. 289-300.
- Douglas, A. (2002), Seismometer arrays - their use in earthquake and test ban seismology, In W. H. K. Lee, H. Kanamori, P. C. Jennings, & C. Kisslinger (Eds.), *International Handbook of Earthquake and Engineering Seismology*, pp. 357-367, San Diego, CA, Academic Press.

- Ewing, W.M., W.S. Jardetsky, and F. Press, *Elastic Waves in Layered Media*, McGraw-Hill, New York, NY, 1957.
- Farnbach, J. S. (1975), The complex envelope in seismic signal analysis, *Bull. Seism. Soc. Am.*, *65*, pp. 951-962.
- Filson, J. (1975), Array seismology. *Annual Review of Earth and Planetary Sciences*, *3*, pp. 157-181, doi:10.1146/annurev.ea.03.050175.001105.
- Gibbons, S. J. (2014), The applicability of incoherent array processing to IMS seismic arrays, *Pure Appl. Geophys.*, *171*, pp. 377-394, doi:10.1007/s00024-012-0613-2.
- Gibbons, S. J. and F. Ringdahl (2004), The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.*, *165*, pp.149-166.
- Gibbons, S. J., J. Schweitzer, R. Ringdal, T. Kværna, S. Mykkeltveit, and B. Paulsen (2011), Improvements to seismic monitoring of the European Arctic using three-component array processing at SPITS, *Bull. Seism. Soc. Am.*, *101*, pp. 2737-2754, doi:10.1785/01201100109
- Goldstein, P. and R. J. Archuleta (1987), Array analysis of seismic signals, *Gephys. Res. Lett.*, *14*, pp. 13-16.
- Grossmann, A., R. Kronland-Martinet, and J. Morlet (1989), Reading and understanding the continuous wavelet transform, In J. Combes, A. Grossmann, & P. Tchamitchian (Eds.), *Wavelets: Time-Frequency Methods and Phase-Space*, pp. 2-10, New York, NY, Springer.
- Gupta, I. N., C. S. Lynnes, T. W. McElfresh, and R. A. Wagner, R. A. (1990), F-k analysis of NORESS array and single-station data to identify sources of near-receiver and near-source scattering, *Bull. Seism. Soc. Am.*, *80*, pp. 2227-2241.
- Harkrider, D. G. (1964), Surface waves in multilayered elastic media: I. Rayleigh and Love waves from buried sources in a multilayered elastic half-space, *Bull. Seism. Soc. Am.*, *54*, pp. 627-679.
- Harris, D. B. and T. Kværna (2010), Superresolution with seismic arrays using empirical matched field processing. *Geophys. J. Int.*, *182*, pp. 1455-1477, doi:10.1111/j.1365-246x.2010.04684.x.
- Haubrich, R. A. (1968), Array Design. *Bull. Seis. Soc. Am.*, *58*(3), pp. 977-991.

- Havskov, J. and G. Alguacil (2004), *Instrumentation in Earthquake Seismology*, Springer.
- Herrmann, R. (2016), http://www.eas.slu.edu/eqc/eqc_mt/MECH.NA/20160628195756/index.html, last accessed December 7, 2020.
- Kushnir, A. F. (1996), Algorithms for adaptive statistical processing of seismic array data. In E. S. Husebye & A. M. Dainty (Eds.), *Monitoring a Comprehensive Test Ban Treaty*, pp. 565-586, Netherlands, Kluwer Academic Publishers.
- Kværna, T. (1989), On exploitation of small-aperture NORESS type arrays for enhanced P-wave detectability, *Bull. Seis. Soc. Am.*, 79, pp. 888-900.
- Kværna, T. and F. Ringdal (1992), Integrated array and three-component processing using a seismic microarray, *Bull. Seis. Soc. Am.*, 82, pp. 870-882.
- Johnson, K.S. (2008), Geologic history of Oklahoma, Earth sciences and mineral resources of Oklahoma: Oklahoma Geological Survey Educational Publication, 9, pp.3-5.
- Langston, C. A. (1987), Depth of faulting during the 1968 Meckering, Australia, earthquake sequence determined from waveform analysis of local seismograms, *J. Geophys. Res.*, 92, pp. 11561-11574.
- Langston, C. A. (2007a), Spatial gradient analysis for linear seismic arrays, *Bull. Seism. Soc. Am.*, 97, pp. 265-280, doi:10.1785/0120060100.
- Langston, C. A. (2007b), Wave Gradiometry in the time domain, *Bull. Seism. Soc. Am.*, 97, pp. 926-933, doi:10.1785/0120060152.
- Langston, C. A. (2007c), Wave Gradiometry in two dimensions, *Bull. Seism. Soc. Am.*, 97, pp. 401-416, doi:10.1785/0120060138.
- Langston, C. A. and C. Liang (2008), Gradiometry for polarized seismic waves, *J. Geophys. Res.*, 113, doi:10.1029/2007JB005486.
- Langston, C. A. and S. M. Mousavi (2019), Separating signal from noise and from other signal using nonlinear thresholding and scale-time windowing of continuous wavelet transforms, *Bull Seism. Soc. Am.*, 109, pp. 1691-1700, doi:10.1785/0120190073.

- Lindfors, A. (1996), Array processing by phase regression, In E. S. Husebye & A. M. Dainty (Eds.), *Monitoring a Comprehensive Test Ban Treaty*, pp. 629-644, Netherlands, Kluwer Academic Publishers.
- Mykkeltveit, S. (1985), A new regional array in Norway: design work and results from analysis of data from a provisional installation, In A. U. Kerr (Ed.), *The VELA Program, A Twenty Five Year Review of Basic Research*, pp. 546-553.
- Mykkeltveit, S., K. Astebol, D. J. Doornbos, and E. S. Husebye (1983), Seismic array configuration optimization, *Bull. Seism. Soc. Am.*, 73, pp. 173-186.
- Nawab, S., F. Dowla, and R. Lacoss (1985), Direction determination of wideband signals, *IEEE Trans. Acous. Speech Sig. Proc.*, 33, pp. 1114-1122.
- Rost, S. and C. Thomas (2002), Array seismology: methods and applications, *Rev. Geophys.*, 40, doi:10.1029/2000RG000100.
- Schaff, D. P. and P. G. Richards (2004), Repeating seismic events in China, *Science*, 303, pp. 1176-1178, doi:10.1126/science.1093422.
- Schaff, D. P. and F. Waldhauser (2010), One magnitude unit reduction in detection threshold by cross correlation applied to Parkfield (California) and China seismicity, *Bull Seism. Soc. Am.*, 100, pp. 3224-3238. doi:10.1785/0120100042.
- Shelly, D. R., J. L. Hardebeck, W. L. Ellsworth, and D. P. Hill (2016), A new strategy for earthquake focal mechanisms using waveform-correlation-derived relative polarities and cluster analysis: Application to the 2014 Long Valley Caldera earthquake swarm, *J. Geophys. Res.: Solid Earth*, 121, pp. 8622-8641, doi:10.1002/2016JB013437.
- Spudich, P., L. K. Steck, M. Hellweg, J. B. Fletcher, and L. M. Baker (1995), Transient stresses at Parkfield, California, produced by the M7.4 Landers earthquake of June 28, 1992: Observations from the UPSAR dense seismograph array, *J. Geophys. Res.*, 100, pp. 675-690.
- Starck, J.-L., F. Murtagh, and J. M. Fadili (2010), *Sparse Image and Signal Processing*. New York, NY, Cambridge University Press.
- Stump, B., M. Jubn, C. Hayward, J. Jeon, I.-Y. Che, K. Thomason, S. M. House, and J. McKenna (2004), Small Aperture Seismo-Acoustic Arrays: Design, Implementation and Utilization, *Bull. Seism. Soc. Am*, 94, pp. 220-236.

- Sweet, J. R., K. R. Anderson, S. Bilek, M. Brudzinski, X. Chen, H. DeShon, C. Hayward, M. Karplus, K. Keranen, C. Langston, F.-C. Lin, M. B. Magnani, and R. L. Woodward (2018), A community experiment to record the full seismic wavefield in Oklahoma, *Seism. Res. Lett.*, *89*, pp. 1923-1930, doi:10.1785/0220180079.
- Toksoz, M. N., A. M. Dainty, and E. E. Charette (1991), Coherency of ground motions at regional distances and scattering, *Phys. Earth and Planet. Int.*, *67*, pp. 162-179.
- VanDecar, J. C., and R. S. Crosson (1990), Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seism. Soc. Am.*, *80*, pp. 150-169.
- Wagner, G. S. and T. J. Owens (1993), Broadband bearing-time records of three-component seismic array data and their application to the study of local earthquake coda, *Geophys. Res. Lett.*, *20*, pp. 1823-1826.
- Waldhauser, F. and W. L. Ellsworth (2000), A double-difference earthquake location algorithm: method and application to the Northern Hayward fault, California, *Bull Seism. Soc. Am.*, *90*, pp. 1353-1368.
- Withers, M., R. Aster, and C. Young (1999), An automated local and regional seismic event detection and location system using waveform correlation, *Bull. Seism. Soc. Am.*, *89*(3), pp. 657-669.
- Yang, Y., C. Liu, and C. A. Langston (2020), Processing seismic ambient noise data with the continuous wavelet transform to obtain reliable empirical Green's functions, *Geophy. J. Int.*, *222*, pp. 1224-1235, doi:10.1093/gji/ggaa243.
- Zhang, J. and C. A. Langston (2020), Separating the scattered wavefield from teleseismic P using curvelets on the long beach array data set, *Geophy. J. Int.*, *220*, pp. 1112-1127, doi:10.1093/gji/ggz487.
- Zhang, H. and C. H. Thurber (2003), Double-difference tomography: the method and its application to the Hayward fault, California, *Bull Seism. Soc. Am.*, *93*, pp. 1875-1889.
- Zhang, M. and L. Wen (2015), An effective method for small event detection: match and locate (M&L), *Geophy. J. Int.*, *200*, pp. 1523-1537. doi:10.1093/gji/ggu466.

List of Symbols, Abbreviations, and Acronyms

a	wavelet scale (s)
α	P wave velocity
AFRL	Air Force Research Laboratory
AFSPC	Air Force Space Command
AFTAC	Air Force Technical Applications Command
β	threshold criterion, also used as S wave velocity in context
\vec{B}	negative of the wave direction vector
BBfk	Broad Band frequency-wavenumber
c	phase velocity, km/s
C	model covariance matrix
CWT	Continuous Wavelet Transform
CTBTO	Comprehensive Test Ban Treaty Organization
DPRK	Democratic People's Republic of Korea
DPZ	vertical component of motion for nodal seismometers
E_d	data covariance matrix
E_m	model covariance matrix
\hat{e}_1, \hat{e}_2	x and y unit vectors
$ECDF_a$	Empirical Cumulative Distribution Function for scale a
$ECDF_a^{-1}$	Quantile Function for scale a
η_α, η_β	vertical wave slownesses for P and S waves
fk or f-k	frequency-wavenumber
f_c	center frequency, Hz
γ	total group slowness
γ_x, γ_y	group slowness for x and y
G	matrix
G_g^{-1}	least-squares or generalized inverse
$H[]$	Hilbert transform
h	node depth
Im	imaginary part of a complex number
m	model vector
d	data vector
Hz	unit of frequency, Hertz
IRIS	Incorporated Research Institutions in Seismology
km	unit of kilometer
KSRS	Korean Seismic Research Station
L&M	Langston and Mousavi (2019) reference
lse	least squares error
Λ	diagonal matrix of singular values associated with singular value decomposition
p	ray parameter, s/km

P	Probability or compressional wave depending on context
pP	near source P wave underside reflection from the surface
P1	a first P wave arrival
P2	a second P wave arrival
P3, P4, P5	other P wave arrivals
PL	a P wave wave-guide phase
P_{\max}	maximum of power on the slowness plane
P_{Total}	integrated power over the slowness plane
φ	instantaneous phase
ϕ	Fourier phase
R	ratio of power
Re	real part of a complex number
R_g	high frequency Rayleigh wave
\vec{R}_i	distance vector to the i^{th} station of an array from array center
ρ	density, gm/cc
s	unit of seconds
sS	near source S wave underside reflection from the surface
S	S wave arrival
S1	principal S wave arrival
SV	vertically polarized S wave
SVD	singular value decomposition
SH	horizontally polarized S wave
SNR	Signal-to-Noise Ratio
SP11	Shot point 1, explosion 1 in AFTAC experiment
SP12	Shot point 1, explosion 2 in AFTAC experiment
SP13	Shot point 1, explosion 3 in AFTAC experiment
SP14	Shot point 1, explosion 4 in AFTAC experiment
t	time, s
t_P	arrival time for P
t_{pP}	arrival time for pP
t_S	arrival time for S
t_{sS}	arrival time for sS
$\sigma_{\gamma_x}, \sigma_{\gamma_y}$	group slowness standard deviation in x and y
σ_{θ}	backazimuth standard deviation
σ_i	data standard deviation
TFR	Time-Frequency Representation
θ	backazimuth
t_i	i^{th} group arrival time
t_0	event origin time
τ	correlation time lag (s)
U	eigenvector matrix associated with singular value decomposition
V	eigenvector matrix associated with singular value decomposition
v	group velocity
v_s	S wave velocity, km/s
V_p	P wave velocity, km/s

V_s	S wave velocity, km/s
W	wavelet coefficient
w	vertical displacement
W_{Beam}	narrow scale band beam of wavelet coefficients
\bar{W}_{Beam}	broad scale band beam of wavelet coefficients
YW	Array code for the IRIS Wavefields experiment
ψ	Mother wavelet
ω	circular frequency
1Dfk or 1Df-k	one dimensional frequency-wavenumber analysis
□	

DISTRIBUTION LIST

DTIC/OCP 8725 John J. Kingman Rd, Suite 0944 Ft Belvoir, VA 22060-6218	1 cy
AFRL/RVIL Kirtland AFB, NM 87117-5776	1 cy
Official Record Copy AFRL/RVB/Dr. Raymond J. Willemann	1 cy

This page is intentionally left blank.