UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CHARACTERIZATION OF THE SUBSURFACE FROM EVALUATING LOW AMPLITUDE SEISMIC RECORDINGS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

RAYMOND NG
Norman, Oklahoma
2023

CHARACTERIZATION OF THE SUBSURFACE FROM

EVALUATING LOW AMPLITUDE SEISMIC RECORDINGS

A DISSERTATION APPROVED FOR THE

SCHOOL OF GEOSCIENCES

BY THE COMMITTEE CONSISTING OF

Dr. Jacob I. Walter, Chair

Dr. Xiaowei Chen, Co-Chair

Dr. Nori Nakata

Dr. Junle Jiang

Dr. Brett Carpenter

Dr. Xingru Wu

Table of Contents

List of Tables

List of Figures

Acknowledgment

"Never give up!" are words I have taped on my apartment wall and looked at every day since the start of my doctoral journey. I was transplanted away from family and familiarity in pursuit of the knowledge and growth I knew I would obtain at the University of Oklahoma. Where I had a void in my family, I managed to find and meet amazing individuals who encouraged and shared meals with me. My feelings of unfamiliarity were replaced with new experiences and perspectives of others, allowing me to further my personal growth. I am not sure I would have been able to complete my journey without those I have met and connected with along the way. I would like to express my deepest gratitude to those individuals.

First and foremost, I would like to thank my advisor, Dr. Nori Nakata, for being the first in the line to believe in my abilities and accept me as his Ph.D. student. His patience and encouragement in the early years of my program allowed me to initiate and develop my technical skills with coding. I'm not sure if I will ever be as great of a programmer as he is, but I will continue honing my skills for many years to come. Working with him, I was able to travel and explore many parts of Oklahoma where we battled boiling fierce heat, brutal cold temperatures, a barrage of insects, and seasonal allergies on many field deployments. His guidance helped me establish a baseline foundation where I am comfortable continuing my journey.

To Dr. Xiaowei Chen, who co-advised me and kept me focused on my research projects during the last 3 years. Multitasking projects are always challenging and she seemed to be able

to handle such tasks with grace and poise that it would appear mythical to me. Try as I might, I would struggle significantly with juggling multiple projects. I would like to thank Dr. Chen for having patience with me as I consistently fumbled my way through toward the end of my journey. I would especially like to express my thanks for bringing me with her for a sabbatical to the Seismological Laboratory at Caltech where we worked with distributed acoustic sensing. I am very grateful for the opportunity to have worked closely with such an emerging technology at one of the most prestigious research institutions in the country.

I would like to thank Dr. Jacob I. Walter, who accepted the impromptu role of the committee chair. Under his advisement, collaboration, and management, I can finalize my work and combine them into what I can only hope is a worthy dissertation. His patience and encouragement echoed “Never give up!” to me.

My sincere thanks to my committee members – Dr. Brett M. Carpenter, Dr. Xingru Wu, and Dr. Junle Jiang for their support in my endeavor and valuable feedback on improving this dissertation. I would like to extend my thanks to the school of geosciences staff – Dr. Gerilyn S. Soreghan, Dr. Shannon Dulin, Dr. Michael H. Engel, Leah Moser, Rebecca Fay, Gail Holloway, Ginger Leivas, Ashley Tullius, Robert L. Turner, and Stephen Holloway for their assistance throughout my doctoral program and enthusiastic lunchtime company, banter, and shenanigans.

I am grateful to the friends I have made in OU – Pranshu, Deepankar, Zhuobo, Yan, Collin, Alex, Peiyao, Jiewen, Jianhang, Marlena, Sophia, Julian, Roberto, Murphy, Mario, Laura, Diana, Pamela, Emily, and others for being wonderful confidants and for keeping the day to day life fun

and interesting. Finally, I want to give my deepest gratitude to my family for their unconditional support in all of my life's pursuits.

I knew the path toward success would be riddled with arduous twists and turns, oftentimes being lost or having to backtrack and start over. It certainly was not a simple feat arriving to where I am today without the care and support of others because they made "Never give up!" that much easier maintain.

Abstract of the dissertation

A geoscientist yearns to learn more about the subsurface and one way of characterizing the subsurface is through active seismic methods which utilize a local source to initiate a strong waveform (e.g., explosives, vibroseis, sledgehammer). However, active sources are generally costly to operate, and logistically complicated since, depending on location, they require approval from an authorizing body, and are very intrusive and disturbing to the environment. Alternatively, passive seismic methods are preferred due to their low cost, ease of instrument deployment, and low environmental impact. Passive seismic is not without its caveats. Passive seismic relies on the surrounding ambient seismic noise for measurements, which can be influenced by local noise sources (e.g., vehicles, strong wind), effectively lowering the signal-to-noise ratio. I will present an array of exciting results from low-amplitude seismic recordings and novel approaches to measuring them to understand the seismic noise environment and microseismic events, which have implications for the field of seismology. I will show an improved analysis of seismogenic potential from magnitude calibration, an interpreted resonant subsurface boundary determined from resonance frequencies, and a correlation of wind speed and seismic recordings. I use an induced seismicity catalog recorded by two nearby boreholes to apply a novel workflow of waveform correlation to 2074 earthquakes for magnitude calibration and subsequent evaluation of seismogenic potential. Waveform correlation reveals statistically significant differences in b-value and seismogenic potential after calibration and suggests a dependence of the earthquake statistics upon their optimal or suboptimal orientation to the local stress field. In another study, we utilize distributed acoustic sensing (DAS) recording, which

allows for a high density of receivers with minimal field deployment. The technique leverages novel sensors to measure backscatter from a telecommunication dark fiber, resulting in a time series of strain along the cable, effectively 100s to 1000s of “seismometers.” Comparison between this novel sensing device and traditional sensors reveals the higher noise level of DAS. However, the higher spatial resolution of DAS allows for stacking measurements along channels and identifying noise sources that were previously not easily recognizable in traditional arrays. Recorded measurements are often subject to increased background noise in an urban or noisy environment. From seismic recordings across Oklahoma from dense arrays, I examine the effects of wind and wind turbines on recorded seismic noise via single-station site response techniques to determine site resonance frequencies and noise source origin.

Chapter 1:

Introduction

1.1 Introduction

In our ever-changing landscape and growing society, it is not uncommon for the general population to go about their day woefully unaware of the destructive force beneath their feet. We live in a world that is constantly shifting and in a state of flux. However, every so often we are reminded of the energy that can be released in the form of an earthquake. Early civilization societies postulated the origin of earthquakes (Agnew et al., 2002) such as Aristotle (330BCE) attributed earthquakes to winds blowing in unground caverns (Missiakoulis, 2008), and early Japan gave credit to a large subterranean serpent-like catfish for earthquakes (Severn, 2012). Advancements in seismological thinking did not occur until the Lisbon earthquake of 1755 due to its destructiveness and evidence for causing motion at great distances (Frechet 2008). Public interest in destructive earthquakes shown in Tan and Maharjan 2018 showed an increased interest in earthquakes after large events and correlated public interest to the destructiveness of the earthquake.

It is commonly known that the destructiveness of earthquakes is directly linked to their magnitude (Gutenberg and Richter, 1942) and site-specific shaking intensity (Hartzell, 1992). Magnitude estimates have evolved with various techniques in calculation such as the least square technique for estimating b-value and maximum likelihood. These variations have led to prolonged debates about a preferred method for estimations. More importantly, the continued interest in geoscience research has aided in the development of recording for longer continuous durations for background noise and the detection of smaller earthquakes. Continuous ambient noise analysis has had a resurgence in recent decades, thanks mainly in part to advancements in commercial and personal computing where larger datasets may be processed at the same time.

One advantage of continuous data is that multiple datasets can be compared amongst each other. Large dataset recordings from regional networks have helped in monitoring the earthquake activity in Oklahoma where seismicity increased steadily in the earlier part of the millennium due to large-scale fluid injection as investigated in many studies (i.e. Rashidi and Ghassemieh, 2023). However, current regional networks still lack the spatial density to capture interstation lateral heterogeneities in the subsurface. High-density local deployments of seismic sensors can resolve the smaller-scale lateral variations in the subsurface that regional networks cannot capture.

The overall motivation for this dissertation is to explore new ways to investigate ground motion intensity. This is done broadly by first examining the earthquake magnitudes during a hydraulic fracturing survey and then with passive seismic techniques to investigate site response and the potential for ground motion amplification. I present the applications of passive seismic techniques at local deployments for high-spatial-resolution surveys of the shallow subsurface. Apart from the overall motivation, each section will have its motivation which will be introduced. The individual section motivation includes the – (i) improvement of magnitudes, (ii) resonance of a deep intracontinental basin, and (iii) local region noise sources' impact on acoustic seismic recordings.

1.2 Overview of the Dissertation

In Chapter 2, I investigate a classical dataset, the Cotton Valley hydraulic fracturing survey in Carthage, Texas. I will use waveform correlation to improve magnitudes calibration and apply newly calibrated magnitudes to quantify the seismogenic potential. Cross-correlation

is used to identify individual fracturing clusters for a b-value analysis and subsequently seismogenic potential. I show improvement in magnitude calibration and seismogenic potential results show a better correlation with the different fracture sets.

In Chapter 3, I use data collected from 3 component nodal sensors co-located with distributed acoustic sensing using a dark fiber array to compare and evaluate the local site conditions for resonance at a high spatial resolution (kilometer scale). 1-month of continuous passive acoustic seismic data was collected. I then compare the results with a velocity model to estimate geologic basin structures which show shallow and coherent resonance layers within the subsurface.

In Chapter 4, I collect 1-month of continuous passive acoustic seismic data from two locations in Northern Oklahoma and compare them with wind speed data. Ambient noise processing techniques are used to identify and analyze the dominant local noise source, wind turbines. I compare the results from the two arrays under similar wind speed conditions, but one lacking in wind turbines.

1.3 Structure of the Dissertation

This dissertation consists of 3 chapters, one of which is formatted in preparation to be submitted for publishing, and the other two are in preparation for submission. The chapters are as follows:

- **Chapter 2:** Waveform correlation improves magnitude calibration during hydraulic fracturing.
- In preparation for resubmission to the Bulletin of Seismological Society of America as:

Ng,R.,Chen,X.,Nakata,N.,Walter,J.I., (2023). Waveform correlation improves magnitude calibration during hydraulic fracturing. *Bulletin of the Seismological Society of America*

- **Chapter 3:** Site response of ambient seismic noise of Northern Oklahoma from the perspective of Nodal and DAS arrays.
- **Chapter 4:** Wind and wind turbine influence on measurements of seismic noise.

References:

Agnew, D. C., Lee, W. H. K., Kanamori, H., Jennings, P. C., & Kisslinger, C. (2002). History of seismology. *International handbook of earthquake and engineering seismology*, *81*(A), 3-11.

Chester, D. K. (2001). The 1755 Lisbon earthquake. *Progress in Physical Geography*, *25*(3), 363-383.

Fréchet, J. (2008). Past and future of historical seismicity studies in France. *Historical Seismology: Interdisciplinary Studies of Past and Recent Earthquakes*, 131-145.

Gutenberg, B., & Richter, C. F. (1942). Earthquake magnitude, intensity, energy, and acceleration. *Bulletin of the Seismological Society of America*, *32*(3), 163-191.

Hartzell, S. H. (1992). Site response estimation from earthquake data. *Bulletin of the Seismological Society of America*, *82*(6), 2308-2327.

Missiakoulis, S. (2008). Aristotle and earthquake data: A historical note. *International Statistical Review*, *76*(1), 130-133.

Rashidi, J. N., & Ghassemieh, M. (2023). Predicting the magnitude of injection-induced earthquakes using machine learning techniques. *Natural Hazards*, 1-26.

Severn, R. T. (2012). Understanding earthquakes: from myth to science. *Bulletin of Earthquake Engineering*, *10*(2), 351-366.

Tan, Y. J., & Maharjan, R. (2018). What Googling trends tell us about public interest in earthquakes. *Seismological Research Letters*, *89*(2A), 653-657.

Chapter 2:

Waveform correlation improves magnitude calibration during hydraulic fracturing

2.1 Abstract

Microseismic monitoring is an important technique to obtain detailed knowledge of in-situ fracture size and orientation during active well stimulation to maximize fluid flow throughout the rock volume and optimize production. Furthermore, considering that the frequency of earthquake magnitudes empirically follows a power law (i.e. Gutenberg-Richter), the accuracy of microseismic event magnitude distributions is potentially crucial for seismic risk management. In this study, we analyze microseismicity observed during four hydraulic fracture stages of the legacy Cotton Valley experiment in 1997 at the Carthage gas field of East Texas, where fractures with significant moment release were activated at the base of the sand-shale Upper Cotton Valley formation. We perform waveform cross-correlation to detect similar event clusters and calibrate event magnitudes based on relative amplitudes within those clusters from waveform cross-correlation. The calibrated magnitudes significantly reduce the deviations between magnitude differences and relative amplitudes of event pairs. This subsequently reduces the magnitude differences between clusters located at different depths. Reduction in magnitude differences between clusters suggests that attenuation-related biases could be effectively mitigated with waveform correlation. The maximum likelihood method is applied to understand the magnitude frequency distributions and quantify the seismogenic index of the geographic clusters. Statistical analyses after magnitude calibration suggest that fractures that are more favorably oriented for shear failure have lower b-value and higher seismogenic index, perhaps suggesting a higher potential for larger earthquakes, rather than fractures subparallel to maximum horizontal principal stress orientation.

2.2 Introduction

Microseismicity monitoring has been a useful technique in quality control of the operation during hydraulic fracturing in oil and gas exploration (Albright et al., 1982). Source location, fracture orientation, and fracture growth may be determined from microseismic monitoring to improve and calibrate treatment designs, as well as identify potential placement for additional hydraulic fracturing or wastewater injection wells within the desired unit. In some cases, it may be essential to maximize economical production by ensuring the rock volume is thoroughly stimulated and creates high permeability for long-term hydrocarbon recovery by examining such factors as proper fracture orientation, size, fractured volume, and hydraulic communication.

Hydraulic fracturing creates fractures and expands high-permeability fracture pathways for effective and efficient hydrocarbon extraction. This process is completed in stages because of the large amount of energy required to pump fluids and materials (e.g., sand proppant to hold fractures open) for generating fractures and is only practical to frack for a limited amount of time in each segment of the leased rock volume. The behavior of the seismicity that is induced during well stimulation is spatially and temporally controlled by stress relaxation and pore pressure fluctuations initiated at the injection location (Shapiro et al., 1997). Hydraulic fractures are created through a tensile opening in the direction of the least principal stress, and the expected hydraulic fracture usually aligns with the direction of maximum principal stress (Hubbert and Willis, 1957). While the tensile opening mode is expected for hydraulic fractures, shear slip has often been observed for microseismic events and larger events, which has been interpreted to represent an interaction between hydraulic fractures and natural fractures or

other sedimentary structures such as bedding planes (Maxwell and Cipolla, 2011; Rutledge et al., 2004; Hubbert and Willis, 1957).

Analyses of characteristics of event occurrence may provide a glimpse into the state of subsurface stress. The power-law distribution of earthquake occurrence in **equation 2.1**, commonly referred to as the Gutenberg-Richter (GR) relationship (Gutenberg and Richter, 1942), is described as

$$\log_{10} N = a - bM, \tag{2.1}$$

where N is the cumulative number of earthquakes above M magnitude, a is the intercept, and b is the slope of the power-law distribution. However, not all events are detected by a network of seismometers. The magnitude of completeness (M_c) describes the magnitude threshold, above which all events should be completely detected by the network. The b-value has been observed to vary across different tectonic zones, stress regimes, earthquake causal conditions, and even temporally for the same study region (e.g., Ghosh et al., 2008). It has been suggested that it may represent a relative indicator of stress distributions across fault zones (Bachmann et al., 2014; Schorlemmer et al., 2005). Bachmann et al. (2012) found that the b-value decreases with distance from the injection point, which could indicate reduced fault strength due to higher pore pressure near the injection source. The relationship between injected volume and magnitude evolution can help us forecast induced seismicity and evaluate a potential maximum magnitude for a fault zone (McGarr, 2014). Shapiro et al. (2010) developed the seismogenic index (SI) model to explain the relationship between injected

volume and event number for different magnitudes, where regions with higher seismogenic index have a higher probability of larger earthquakes with the same injected volume. The seismogenic index is a statistical approach to examine the occurrence probability of fluid injection-inducing earthquakes at a given injection site (Shapiro et al., 2010).

In this study, we analyze a legacy dataset for four hydraulic fracturing stages of the tight-gas reservoir within the sand-shale Cotton Valley formation in Carthage, Texas (May – July 1997) (Walker, 1997). High-resolution relocation and focal mechanism analyses by Rutledge et al. (2004) revealed different sets of fractures within the base of the Upper Cotton Valley formation. Hydraulic treatments with anomalously high moment releases and a high number of events occurred within several dense clusters, which delineate fault bends or jogs that are more favorably oriented for shear failure. Our aim in this study is to investigate whether statistical analyses can reveal different characteristics of fractures. We begin by introducing the dataset, then perform magnitude calibration using waveform cross-correlation analysis, followed by a detailed magnitude-frequency distribution and seismogenic index analyses. Finally, we discuss the improvements before and after magnitude calibrations and the implications of the improved results.

2.2.1 Experiment data

The hydraulic fracturing data used in our study was recorded in 1997 during injections at depths of 2757-2838 meters and 2615-2696 meters in the Cotton Valley Formation. The hydraulic fracturing experiment induced more than 4,000 microseismic events over six treatments, which were recorded by two borehole monitoring arrays (Walker, 1997; Walker et

al., 1998). Increasing the number of monitoring borehole arrays from one to two in the experiment for some treatments improved the ability to distinguish between coherent noise and microseismic events, in addition to overcoming low signal-to-noise ratio and low resolution, as described by Zinno et al. (1998).

For our analysis, we only use the available treatments A, B, C, and E as represented in **Figure 2.1,** and follow the same naming convention as Rutledge et al. (2004), which include 2547 events from the original Walker (1997) catalog (**Table 2.1**). Among these events, 2074 has available waveforms. The distribution of catalog events with complete waveforms is shown in **Figure 2.2**, which shows an evenly sampled subset of the original catalog. The relationship between the 2074 microseismic events and injection rate with time is shown in **Figure 2.3** for the four treatments analyzed here.

We observe a good correlation between the onset time of the injection rate and the origin time of the induced seismic events. In treatments A, B, and C the abrupt decrease in injection rate also correlated with the seismicity rate drop (**Figure 2.3**). However, in treatment E there is a relative anomalous increase in seismicity rate with injection. Rutledge et al., (2004) attribute anomalous seismic activity to pressurized fracture offsets or orientation changes that concentrated stress and fluid choke off. Typical right lateral movement places the left-stepping jog of treatment E in compression, but a right step is hypothesized by Rutledge et al., (2004) as a pressure sink which delayed seismicity. Therefore, seismicity increased even after shut-in, which is an industry term that denotes no further fluids were entering or exiting the wellhead.

In this analysis, we only use the 2074 cataloged events with complete waveforms that clear P and S-wave arrivals from four stages analyzed in Rutledge and Urbancic, (1999). Seismic moments are calculated by Rutledge et al. (2004) using methods described in Andrews (1986), which averages the values obtained from P and S phase recordings from several stations. The seismic moments from Rutledge et al. (2004) used in this study are mostly between the narrow range of 0.4 μNm – 8 μNm. The majority of the relocated microseismic events form clouds that delineate fractures aligned with the maximum horizontal stress (SH-max at N80E), especially during treatments A and C (**Figure 2.1**). During treatments B and E, microseismicity occurred in several clusters we identified using cross-correlation with some having anomalous orientations that are off-axis to fractures aligned with SH-max (Rutledge et al., 2004). As seen in **Figure 2.1**, most stations to event hypocenters distances are at similar equidistant ranges between 1km - 2km and at a similar depth.

2.3 Methods

In this study, we perform magnitude calibration based on the relative amplitude (Cleveland and Ammon, 2015) measured from waveform cross-correlation following Chen et al. (2018). Then, we perform magnitude-frequency distribution analysis to obtain the b-value for each treatment with the maximum-likelihood method (referred to as “MLE”) and compare results using both the original catalog magnitudes and the calibrated magnitudes. To better understand the improvement of calibrated magnitudes, we compare the frequency-magnitude distributions (FMD) of the different treatments and subclusters of treatment B. Finally, we link magnitudes with injected volume to obtain seismogenic indexes for each treatment following Shapiro et al. (2010) using corresponding b-values obtained from MLE. For both the b-value and

seismogenic index, we assess the uncertainty through bootstrap resampling. The relative variations of b-value and seismogenic index among treatments and subclusters are interpreted as related to the relative stress state of different fractures via Mohr circle analysis.

2.3.1 Magnitude calibration using cross-correlation

Precise magnitude estimates may be acquired from the cross-correlation of two earthquake waveforms with well-constrained reference event magnitudes (Gibbons and Ringdal, 2006; Schaff and Richards, 2014) and can be applied to earthquake clusters with similar faulting geometry and depth (Cleveland and Ammon, 2015). Magnitude calibration can reduce the scattering between the relative amplitude ratios and magnitude differences of event pairs, which is important to improve statistical analyses related to magnitude distributions (Chen et al., 2018; Shelly et al., 2015, 2016).

The dataset includes rotated three-component microseismic event waveforms at each geophone with sampling rates of 1000 Hz. For each treatment, only a subset of waveforms is available for the cataloged events from Rutledge et al. (2004) (**Table 2.1**). We first apply a short-term average/long-term average (STA/LTA) picker using the GISMO (Reyes and West, 2011) package to obtain P-wave arrival time on the vertical channel and S-wave arrival time on horizontal and transverse channels. Waveforms with no detectable picks are not used in the following analysis. Examination of waveforms suggests that the transverse component has the highest signal-to-noise ratio. Therefore, we use the S-wave from the transverse component for magnitude calibration.

We then measure relative amplitude using waveform cross-correlation (CC) with the following steps for each treatment:

1. A detailed inspection shows that many traces have a low signal-to-noise. To improve our data visualization, we apply a Butterworth bandpass between 20 Hz and 150 Hz where observed amplitudes for microseismic events are the largest.
2. Extract a 70 ms window around the initial arrival of the microseismic event (10 ms before and 60 ms after).
3. Cross-correlate each event pair at each station and measure the relative amplitude ratio using principal component analysis for waveform pairs with a CC coefficient greater than 0.6 (Wold et al., 1987). An example of relative amplitude measurement is shown in **Figure 2.4**. The waveform pairs are time-shifted based on the lag time, and the relative amplitude ratio is calculated by taking the maximum eigenvalue of the covariance of the shifted waveform pairs. Only event pairs with amplitude ratio measurements from more than 5 receivers are used for magnitude calibration.
4. Invert for new magnitudes based on amplitude ratios and cataloged seismic moment (Chen et al., 2018; Cleveland and Ammon, 2015; Schaff and Richards, 2014; Shelly et al., 2015, 2016) based on **equation 2.2**,

$$\begin{bmatrix} \log_{10} AR_{1,2} \\ \log_{10} AR_{1,3} \\ \log_{10} AR_{1,4} \\ \cdots \\ \log_{10} AR_{N-1,N} \\ \sum_{i=1}^{N} \log_{10} M_0^i \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & \cdots & \cdots & 0 \\ -1 & 0 & 1 & \cdots & \cdots & 0 \\ -1 & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & 0 & -1 & 1 \\ 1 & 1 & 1 & \cdots & \cdots & 1 \end{bmatrix} \times \begin{bmatrix} \log_{10} M_0^1 \\ \log_{10} M_0^2 \\ \log_{10} M_0^3 \\ \cdots \\ \log_{10} M_0^{N-1} \\ \log_{10} M_0^N \end{bmatrix} \quad (2.2)$$

where the logarithmic amplitude ratio ($AR_{j,k}$) between events j and k is the averaged amplitude ratio from all receivers with CC > 0.6. In the last row of **equation 2.2**, N is the total number of recorded earthquakes and $\sum_{i=1}^{N} \log_{10} M_0^i$ is the summation of the total seismic moment for events included in the calibration from the initial catalog, which constrains the summation of the calibrated seismic moment. The calibrated magnitudes are obtained via the moment magnitude relationship (**equation 2.3**) from Kanamori, (1977) where the moment (M_0) in dyne-cm is converted to N-m to obtain moment magnitude (M_w):

$$M_w = (\log_{10} M_0 - 9.1)/1.5 \tag{2.3}$$

2.3.2 b-value

Similar to Ibanez et al. (2012), we determine the Mc by maximum curvature method (MAXC) (Wyss et al., 2000) from an open-source z-map Matlab toolbox (Wiemer and Malone, 2001). MAXC determines Mc by the maximum value of the first derivative of the frequency-magnitude curve, the point of maximum curvature. However, a caveat in MAXC is the tendency to underestimate Mc in data as explained by Mignan and Woessner (2012). Using events with magnitudes above the estimated Mc, the b-values from the Gutenberg-Richter relationship (**equation 2.1**) are obtained using the maximum-likelihood method (MLE) (Wiemer and Malone, 2001). The estimated b-value is the logarithmic of $e/(M_{av} - M_{min})$ using the average magnitude (M_{av}) and minimum magnitude (M_{min}) (both before and after magnitude calibration) following **equation 2.4** (Aki, 1965):

$$b = \log_{10} e/(M_{av} - M_{min})$$

(2.4)

While MLE provides reliable estimates of b-values when the FMD follows a GR relationship (Milojevic, 2010), there is the possibility that MLE tends to bias lower magnitude events when Mc is too low and will result in a poor fit to the GR relationship (Ibanez et al., 2012).

To assess the uncertainty in b-values due to possible biases in Mc, we perform bootstrap resampling with 100 resampled datasets for each treatment. For each resampled dataset, a new Mc and b-value are estimated. Generally, b-values are expected to be ~1 in large sample sizes (Frohlich and Davis, 1993), which is typically observed for tectonic earthquakes (Kagan, 1999), but seismicity associated with fluid injection (Bachmann et al., 2011, 2012, 2014; Lei et al., 2008) or magma intrusion (Wiemer and McNutt, 1997) often have b-values greater than 1. Differences in styles of faulting also produce varying b-values (Schorlemmer et al., 2005). The spatiotemporal variations of the b-value often indicate stress distributions, with lower b-values indicating higher differential stress (Schorlemmer et al., 2005) or associated with fault activation (Shelly et al., 2016; Chen et al., 2018).

2.3.3 Mohr circle construction

A Mohr circle can be used to visualize the subcluster analysis of the treatments with clusters not orientated along to SHmax. Developed by Mohr (1900) to investigate fluid pressure and stress-controlling fracture opening, Mohr diagrams provide a useful tool for visualization of the relationship between the normal stress acting on a fracture surface and the shear stress

required to overcome the friction angle. We can describe the normal stress (σ_n) using the maximum (σ_1) and minimum (σ_3) principal stress, and (θ) the angle between the fracture plane and maximum principal stress. Mohr-Coulomb failure criterion (Heyman et al., 1971) included a failure envelope that is tangential to the Mohr circle where the apex between the failure envelope and the Mohr circle is (σ_n, τ_f). Normal stress is determined geometrically in **equation 2.5** using the horizontal component of the Mohr circle normal stress ($\frac{\sigma_1-\sigma_3}{2}\cos 2\theta$) and the principal average stress ($\frac{\sigma_1+\sigma_3}{2}$).

$$\sigma_n = \frac{\sigma_1+\sigma_3}{2} + \frac{\sigma_1-\sigma_3}{2}\cos 2\theta \tag{2.5}$$

Mohr-Coulomb failure criterion failure envelope describes the stress state at failure when the acting shear stress overcomes the internal friction angle (ϕ) and normal stress. The geometric relationship (**equation 2.6**) of the shear stress at failure (τ_f) to the normal stress and material cohesion (c) is given by

$$\tau_f = c + \sigma_n \tan\varphi \tag{2.6}$$

Since we are only interested in the relative variation of the fracture stress states on the Mohr circle, we only calculate a schematic Mohr diagram with normalized stress amplitude and assume a friction coefficient of 0.75 for the Coulomb failure envelope.

2.3.4 Seismogenic Index

Developed by Shapiro et al. (2010), the seismogenic index ∑ is a derived parameter based on the fluid-injection rate and seismicity rate of induced earthquakes that quantifies the

seismotectonic state at an injection site, originally empirically derived from geothermal case studies. The seismogenic index is theoretically independent of injection time and any injection characteristics (Shapiro et al., 2010), where the larger the index the larger the probability of a significant magnitude event. Shapiro et al. (2010) estimated the seismogenic index for a specific injection location using the following:

$$\sum(t) = \log_{10} N_{\geq M}(t) - \log_{10} Q_c(t) + bM \tag{2.7}$$

where N is the number of induced events with magnitude ≥ M larger than (M) as a function of injection time (t) and the cumulative injection fluid volume (Q_c). **Figure 2.3** depicts the relationship between the injection rate, time, magnitude, and the number of events which is used to determine the seismogenic index values in **equation 2.7**. We use the previously obtained b-values and Mc to calculate the seismogenic index at each time step of 0.5 hours and assume the average as the final index for each treatment. For each treatment, the seismogenic index is estimated using the catalog both before and after magnitude calibration.

The uncertainty of the seismogenic index is estimated with the same 100-resampled dataset for each treatment used for b-value uncertainty estimation. Both b-value and seismogenic index uncertainties are quantified with probability density functions based on results from these resampled datasets, where bootstrapped probability histogram results (h) are from the ratio of the cumulative (n) observations within a given bin size (k) at (i) bins shown in **equation 2.8**.

$$h = \frac{\sum_{i=1}^{k} n_i}{k}$$

(2.8)

2.4 Results

Out of the 2074 cataloged events with waveforms (see **Table 2.1**), we obtain calibrated magnitudes for 1963 events. We compare the relationship between magnitude difference and measured amplitude ratio to examine the improvement of calibrated magnitude. Then, we compare the distribution for the b-value and seismogenic index from the bootstrap resampling to discuss the geomechanical implications of those measurements.

2.4.1 Calibrated magnitude

For each stage, we use heat maps to examine the relationship between the logarithm of amplitude ratio measured from the waveform and the relative magnitude between available event pairs. The image resolution of the heat map is determined by the total number of events in each bin of amplitude ratio and relative magnitude. **Figure 2.5** clearly shows magnitude calibration significantly reduces the scatter for all treatments converging to a slope of 1.5 between amplitude ratio and relative magnitude as expected from **equation 2.3** (the moment magnitude relationship).

Before magnitude calibration, different treatments exhibit diverse behaviors. Both treatments A and B were conducted in treatment well 21-10 and monitored by wells 21-09 and 22-09. Both exhibited highly scattered distributions between amplitude ratio and relative magnitude before magnitude calibration. Several sub-parallel linear trends can be identified

with an approximate slope of 1.5 matching that of the inversion in **equation 2.3**. These sub-parallel groups and increased scatter in the uncalibrated data could be due to attenuation from different ray paths (see **Figure 2.1**). Treatments C and E were conducted in treatment well 21-09 and recorded by well 22-09. These two treatments exhibit tighter distributions compared to treatments A and B before magnitude calibration, which could be due to narrower depth distribution since ray paths to the monitoring well are similar (see **Figure 2.1**). To better understand the cause of the scattered magnitudes in the original catalog, we examine the FMD for similar event clusters during treatment B. We group events based on waveform similarity by defining similar event clusters with minimum CC of 0.6 from at least 8 stations. **Figure 2.6** shows events in each cluster confined within similar depth and spatial location, which is consistent with grouping based on location and polarity in Rutledge et al. (2004). These clusters have large magnitude differences before calibration, which decrease after calibration (**Figure 2.6b and 2.6d**). The measured amplitude ratios are based on filtered waveforms between 20 and 150 Hz, which is well below the expected corner frequency for the magnitude range (Urbancic et al., 1996). By measuring the amplitude ratio between event pairs at similar locations, the attenuation from the source region and path can be effectively canceled out as seen in **Figure 2.5**.

2.4.2 Frequency-Magnitude Distribution and the Seismogenic Index

Figure 2.7 shows the FMDs of the four treatments before and after magnitude calibration. **Table 2.2** provides the resultant seismogenic index before and after magnitude calibration as well as the b-value and Mc for each treatment. For all treatments, there exists some deviation from the GR relationship at higher magnitudes. Deviations from the GR

relationship have been observed for fluid-induced seismicity (Skoumal et al., 2015). This could explain the large difference in b-value between the MLE and least squares (LST) methods (see supplementary materials). The b-values with the MLE method for treatments A, B, and C are systematically lower than the b-value estimates from Dinske et al., (2013). However, those b-values from Dinske et al. (2013) agree well with our LST b-values (see supplementary materials), suggesting that the choice of b-value methods strongly influences the results when the FMD strongly deviates from the GR relationship (Skoumal et al., 2015). Despite the large differences in absolute b-values, the relative differences between b-values using the same method remain consistent. Both Dinske et al., (2013) and the MLE estimates here, found treatment B has a slightly lower b-value than treatment A and a much higher b-value for treatment C.

The probability density functions for the b-value and seismogenic index from the four treatments are compared in detail in **Figure 2.8**. The uncertainty for treatment C is significantly reduced after magnitude calibration, suggesting magnitude calibration improves the stability of subsequent calculations. The relative distribution of the seismogenic index is consistent before and after calibration: treatment E has the highest value, followed by treatment B, A, and C. The relative distribution of the b-value shows some differences. Before magnitude calibration, treatment B has the lowest b-value, followed by treatments E, A, and C, while after calibration, treatment E stands out with the lowest b-value, followed by treatments B, A, and C. Treatments B and A consistently have similar b-values, and treatment C always has the highest b-value. Both the seismogenic index and b-value can be associated with seismic hazard: a higher seismogenic index and lower b-value indicate a statistical population of earthquakes with a distribution composed of larger events within the range of magnitudes observed. In this respect, the relative

ranking of b-value and seismogenic index after magnitude calibration consistently indicates the highest seismogenic potential for treatment E, demonstrating that magnitude calibration perhaps provides a more coherent hazard assessment using different parameters.

Figure 2.6 shows that treatment B can be further separated into similar event clusters. A principal-component-analysis method is used to estimate the geometry of each cluster following Qin et al. (2022). Among all clusters, clusters 1 and 4 show a relatively larger angle, 20°, from SHmax orientation, while the other clusters have a similar orientation to SHmax. We then further assess the distribution of b-value and seismogenic index for events within clusters 1 and 4 (referred to as "C1-4") and the rest of the events (referred to as "others"). C1-4 consists of about 40% of all events in treatment B. **Figure 2.8** shows that C1-4 has a lower b-value and a higher seismogenic index than "others". The overall b-value of treatment B lies between C1-4 and "others". The events that are not within C1-4 have a similar seismogenic index compared to treatment A both before and after magnitude calibration. This suggests different seismogenic potentials for fractures of different orientations during the same treatment.

2.5 Discussion

Within the scope of this study, we determine seismogenic index values from calibrated magnitudes based on a limited range of data. The magnitude calibration using the full waveform provides significantly improved magnitudes, though the statistical analysis of the b-value and seismogenic index shows some differences. Here, we further discuss the influence of magnitude calibration on statistical analyses and the possible physical processes that cause the differences among different treatments and different fractures.

2.5.1 Factors influencing magnitude measurement

The significant improvement in agreement between amplitude ratio and magnitude difference in **Figure 2.5** suggests that analysis of the full waveform is needed for microseismic events. Catalog magnitude estimations were calculated based on seismic moments measured by Urbancic et al. (1996) and the moment magnitude relationship in **equation 2.3** (Rutledge et al., 2004). Influences such as attenuation (quality factor Q) may impact seismic moment measurements. Lower quality factors Q (high attenuation) were shown to result in a smaller seismic moment and magnitude estimation than similar earthquakes in higher Q regions (Chung and Bernreuter, 1981). Events with highly similar waveforms typically share similar ray paths and are closely located (Kane et al., 2013), so the amplitude ratios from full waveform are less influenced by attenuation. In other words, attenuation from the source region and path are effectively canceled out. Mitigating the influence of attenuation on the original magnitude estimations leads to significantly improved magnitude estimations. **Figure 2.6** clearly shows that magnitude calibration reduces the systematic magnitude differences among clusters located at different depths. **Figure 2.5** shows that the original magnitude difference exhibits a very scattered relationship with the amplitude ratios. The calibrated magnitudes mitigate the attenuation problem and exhibit better agreement with the amplitude ratios in **Figure 2.5**.

Cleveland and Ammon (2015) used a wide aperture of seismic networks for events used in amplitude ratio calculations. The borehole arrays used here have relatively limited azimuthal coverage and could lead to issues in radiation patterns. **Figure 2.6** shows that events of similar waveforms are grouped into compact clusters, which agree with clusters of similar focal mechanism solutions in Rutledge et al. (2004). This suggests that when measuring the ratio

between events of similar waveforms, the influence of radiation pattern would be relatively small, as long as the events have similar focal mechanisms or faulting styles.

2.5.2 Uncertainties in statistical measurement

The b-value tends to be approximately ~1 for tectonic earthquakes (Schorlemmer et al., 2005) and nearly ~1 for intercontinental regions such as Oklahoma (e.g., Walter et al., 2017). However, for induced seismicity that is triggered or associated with hydraulic fracturing, the b-value tends to be higher, which represents a greater quantity of lower magnitude events to larger magnitude events when compared to regions where the b-value tends to be approximately 1 (Eaton et al., 2014). Corrected magnitudes that are included in b-value estimation are important when characterizing induced seismicity since there is often a smaller dynamic range of magnitude units as compared to natural earthquakes because the catalogs are typically smaller. A relatively higher b-value implies lower distribution of larger events, which has been interpreted as being consistent with a conceptual model of relatively lower fault strength (Schorlemmer et al., 2005). In such a model, asperities remain small. On the other hand, event number and magnitude may depend on the injection pressure and, within the seismogenic index model, is proportional to injection volume (Dinske and Shapiro, 2016; Shapiro et al., 2007, 2010).

The seismogenic index estimation requires an estimation of the b-value (**equation 2.5**), so uncertainties in the b-value would influence seismogenic index estimations. The uncertainties are assessed via the bootstrap resampling method, and a new Mc is estimated for each resampled dataset. This could mitigate the potential bias due to possible underestimation

of Mc using the MAXC method (Mignan and Woessner, 2012). Sometimes, the FMD can deviate from the GR relationship (Skoumal et al., 2015), especially for earthquake catalogs that consist of small numbers of events. **Figure 2.7** shows the deviation from the GR relationship at larger magnitudes. This leads to some discrepancies between different methods, for example, LST tends to produce steep slopes that fit larger magnitude bins better. However, LST tends to have larger uncertainties using the bootstrap resampling method, and the probability density function features bi-modal distribution for treatment B (**Figure S2.1** in the supplemental material). Despite the larger uncertainty, LST shows some consistent relative distributions with MLE estimates: treatment E has a lower b-value and higher seismogenic index than "others" (non C1-4) from treatment B, followed by treatment A and C. The relative distribution between the whole treatment B (and C1-4) and treatment E shows some inconsistency due to the bi-modal distribution of b-values from the bootstrap resampling, suggesting a possible instability with the least-square approach, or some differences in the seismogenic potential of different sets of fractures during treatment B.

2.5.3 Difference in the seismogenic potential for different sets of fractures

The natural fracture system in the Cotton Valley formation is dominated by fractures with similar orientations to SHmax (Rutledge and Phillips, 2003). As suggested for other areas, the shear failure microseismic events represent an interaction between hydraulic fracture and natural fracture system (Maxwell and Cipolla, 2011). Treatments A and C involve fractures of similar orientation to SHmax at shallower depth, while treatments B and E involved fractures at deeper depth with some anomalous fractures that are more favorably oriented relative to the

stress field (Rutledge et al., 2004) (**Figures 2.2**). Rutledge et al. (2004) interpreted those anomalous clusters as fault jogs or bends that represent areas of stress concentration.

During treatment E, over 80% of events occurred along the fracture that is significantly deviating from SHmax, therefore, we do not separately analyze different event groups in treatment E. During treatment B, about 40% of events occurred within clusters 1 and 4 that strongly deviate from SHmax, so we separately analyzed different event groups (**Figure 2.8** and **Figure S2.1**). The b-value results from MLE show that C1-4 during treatment B behave more similarly to treatment E, while the other events occurring fractures aligned with Shmax behave more similarly to treatment A (**Figure 2.8**).

Based on the relative angle between the orientation of different fracture sets and the SHmax, **Figure 2.9** displays these fracture sets on a schematic Mohr diagram and a Coulomb failure envelope with a friction coefficient of 0.75. The largest cluster in treatment E and C1-4 in treatment B are more optimally oriented than the other fractures. The evolution of cumulative seismic moment shows that the seismic moment release during treatment B is mainly released during C1-4 as exhibited in **Figure 2.10**. The latter part of C1-4 and treatment E have much higher moment release per event than other treatments or event groups. Thus, the improved statistical measurements using MLE with calibrated magnitude are consistent with the geomechanical properties of different sets of fractures and agree with observed seismic moment release history. These observations are consistent with larger regional scale-induced seismicity observations in Oklahoma related to wastewater disposal. Qin et al. (2019) found that larger magnitude events in Oklahoma mostly occur along optimally orientated faults, and that non-optimally oriented faults with lower shear stress mostly occur within spatial areas of the

highest injection volume of wastewater. The consistency between large-scale observations and local fracture network suggests self-similarity, in that similar physical processes control the seismogenic processes at vastly different spatial scales.

In addition to fracture orientations, depth may play a role in affecting different statistics for different fracture sets. Clusters 1-4 during treatment B and treatment E are located at relatively deeper depths. This observation is generally in agreement with the decrease in the b-value with depth observed in Gerstenberger et al. (2001), likely related to increased stress with deeper depth. It may be possible the fracture orientation is depth related. However, Seeburger and Zoback (1982) examined 10 wells drilled in granitic rock from 3 different locations in North America and showed little depth dependence with fracture orientation in well fracture analysis.

2.6 Conclusion

We propose an addition to the statistical analysis workflow of microseismic events by incorporating a magnitude calibration method based on waveform cross-correlation. Using the available initial catalog magnitudes and magnitude calibration method, we obtain new magnitudes from microseismic events during four fracture stages. The observed improvement in magnitude difference and amplitude ratios after calibration suggests the new magnitudes are more accurate. With the new magnitudes, systematic differences in the b-value and seismogenic index between different fracture sets are revealed. Clusters 1 and 4 during treatment B and treatment E are more optimally oriented based on the schematic Mohr diagram, which is consistent with higher seismogenic potential based on the lower b-value and higher seismogenic index from the improved statistical analysis. Possible real-time

implementation of magnitude calibration can significantly improve microseismicity monitoring and assessment of seismic hazards and help identification of slip activation of optimally oriented fractures. However, additional comparisons will be required for other hydraulic fracturing datasets to support our findings.

2.7 Data and Resources

All data used in this paper came from Rutledge et al. (2004) listed in the references. We processed the SAC format waveform data using GISMO a MATLAB toolbox for seismic data analysis (Thompson, 2017). Supplementary materials contain an analysis of the dataset results affected by varying the number of input events.

2.8 Declaration of Competing Interest

The authors acknowledge that there are no conflicts of interest recorded.

2.9 Acknowledgements

We thank James Rutledge for providing seismic waveform data. We are grateful to the two anonymous reviewers for their insightful reviews that improved the manuscript. This research is partially supported by a graduate assistantship from the School of Geosciences and the McCoy fund from the University of Oklahoma.

2.10 Tables and Captions

Treatment	No. of raw catalog events	No. of common events	Final catalog events
P1S3 (A)	628	628	596
P1S2 (B)	888	644	581
P2S3 (C)	369	369	368
P2S1 (E)	662	433	418
Total events	2547	2074	1963

Table 2.1: Number of events in the full catalog, full catalog b-value, common events, and the final catalog.

Treatment	Depth range (m)	Catalog b-value	Percentage of events above Mc	Mc	Catalog seismogenic index	New b-value	Percentage of events above Mc	New Mc	New seismogenic index
A	2615 - 2696	1.98 ± 0.07	60.67%	-2.15	-4.5 ± 0.23	1.92 ± 0.07	61.25%	-2.10	-4.4 ± 0.18
B	2757 - 2838	1.79 ± 0.06	53.42%	-1.88	-3.6 ± 0.11	1.83 ± 0.08	61.10%	-1.86	-3.8 ± 0.20
C	2607 - 2643	2.62 ± 0.10	63.69%	-2.23	-6.4 ± 0.56	2.09 ± 0.08	74.72%	-2.28	-5.5 ± 0.24
E	2746 - 2763	1.88 ± 0.07	66.97%	-1.64	-3.1 ± 0.12	1.63 ± 0.06	72.82%	-1.69	-2.7 ± 0.11

Table 2.2 Caption: Maximum likelihood fitting for b-value and seismogenic index before and after magnitude calibration with depth ranges and percentage of event recovery above Mc.

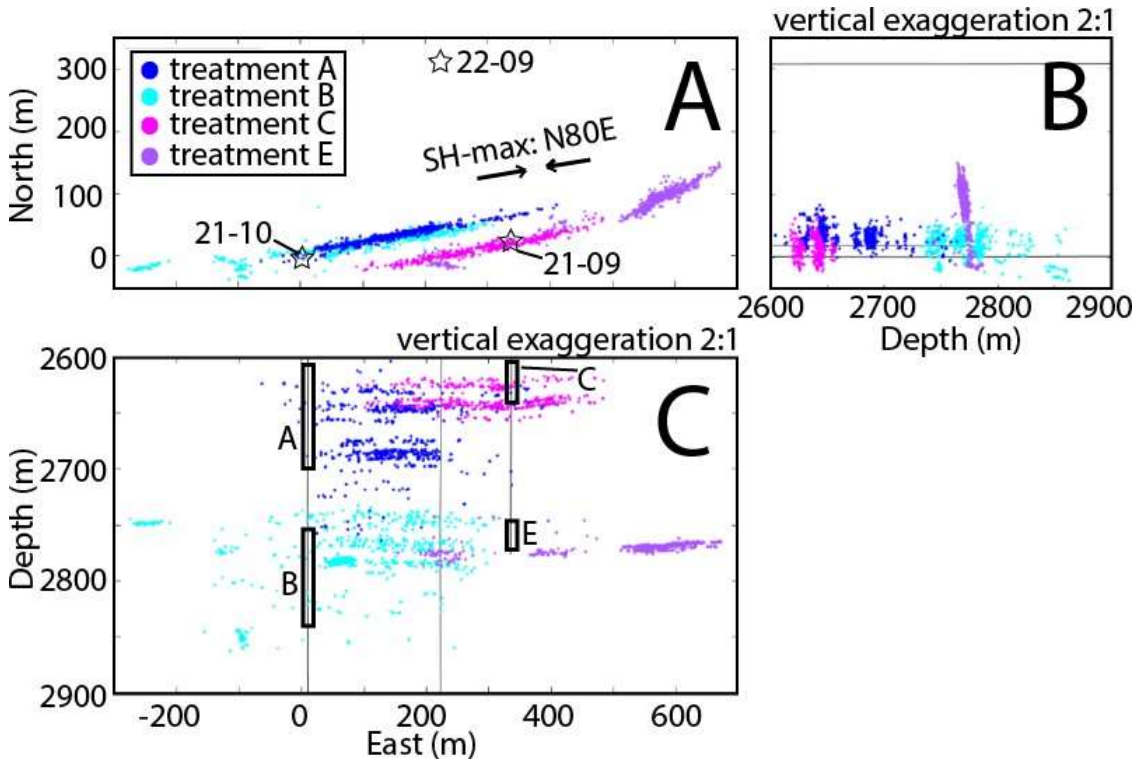

Figure 2.1: Array Geometry Caption: Overview of the experiment setting. The location of treatment wells 21-10 and 21-09 (stars) are shown by symbols and the thin black line. 21-09 (recorded treatments A and B) and 22-09 (recorded all four treatments) are monitoring wells. Note that well 21-09 is used for both monitoring and treatment. Microseismic events are shown by colored dots where cyan represents treatment B, blue represents treatment A, purple represents treatment E. and magenta represents treatment C. Maximum horizontal stress direction is illustrated by the black arrow. Injection interval depth is shown as black rectangles and labeled with the treatment interval name. Top left figure: North-East view. Top right figure: North-Depth view. Bottom left: Depth-East view. Vertical exaggeration is 2:1.

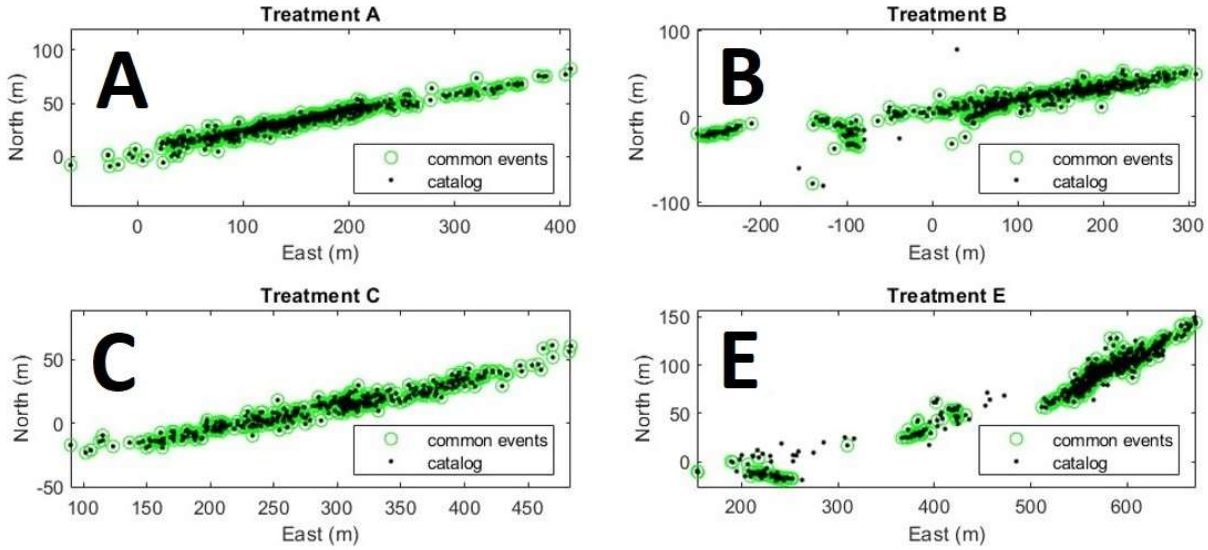

Figure 2.2: Distribution of available common events between catalog and available waveform data. Each quarter shows the map view of the spatial variability between catalog events in black and available waveform events in green.

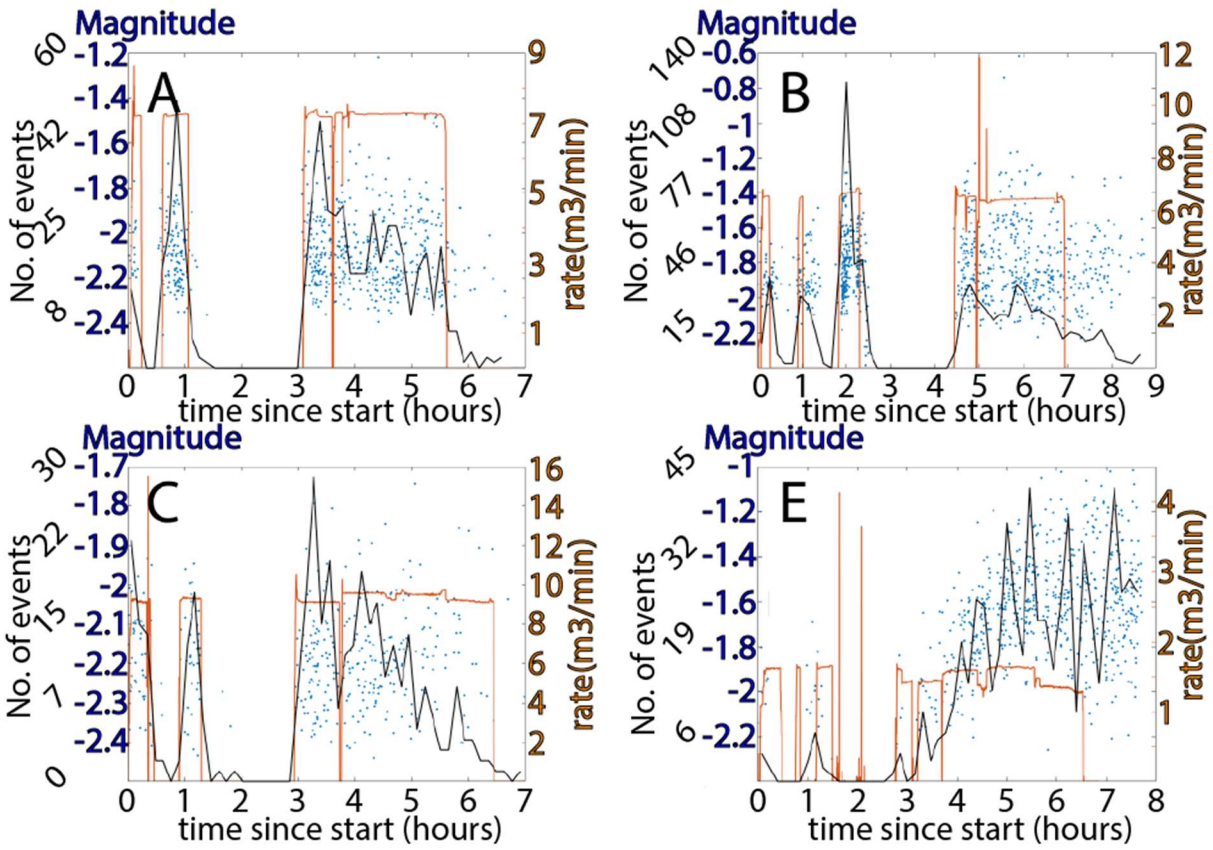

Figure 2.3: Injection volumes with common event-calibrated magnitudes are shown. Clockwise from the upper left is treatment A, B, C, and E. Injection rate is plotted in orange. Blue dots are the distribution of microseismic events where vertical placement is the magnitude and horizontal position is timing relative to the start of the hydraulic fracturing. The seismicity rate is shown in black by the number of events over time. The magnitude of completeness for each treatment is shown as the blue triangle

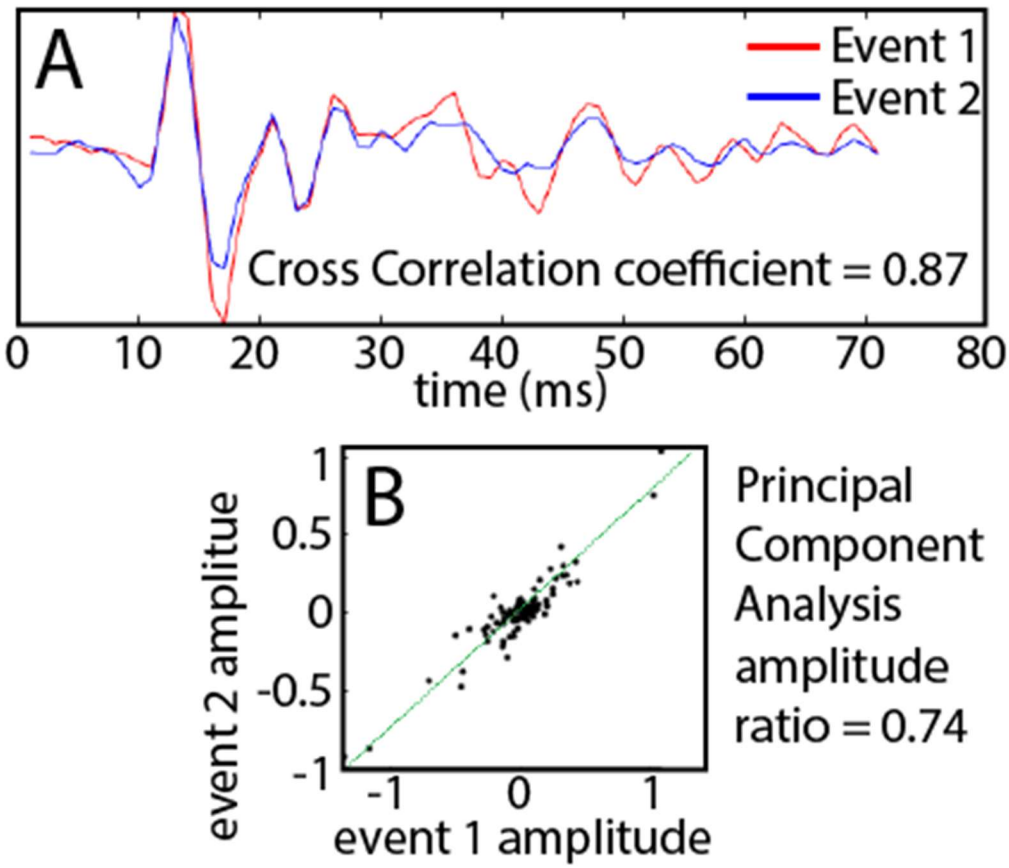

Figure 2.4: A: Comparison of two waveforms contrasted by color with a similarity coefficient of 0.87. B: the waveform amplitude value comparison between the two distinct events. Calculated amplitude ratios are denoted by the black dots where the principal component amplitude ratio is 0.74.

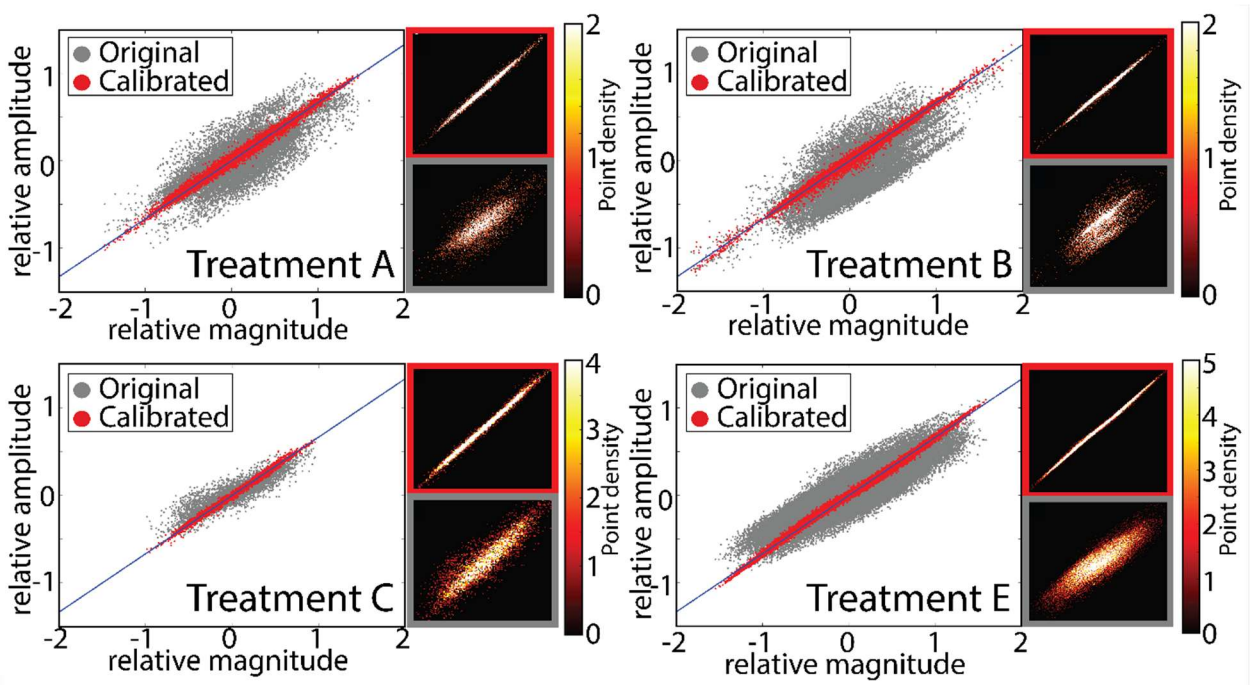

Figure 2.5: Principal component analysis for magnitude calibration of the four fracking stages clockwise starting from the top left panel: treatment A, B, C, and E. Each quadrant is sub-divided into: the main section which shows the relative magnitude and amplitude ratio before (gray) and after magnitude calibration (red) distribution; the upper right subsection that displays the density of the distribution in the main section after magnitude calibration as a heat map with color bar corresponding to density in log10 scale; lower right subsection displays the density of the distribution before magnitude calibration with the same color bar to the upper right subsection. Convergence in the principal component analysis is displayed by the calibrated points collapsing on the blue line that follows a slope of 1.5.

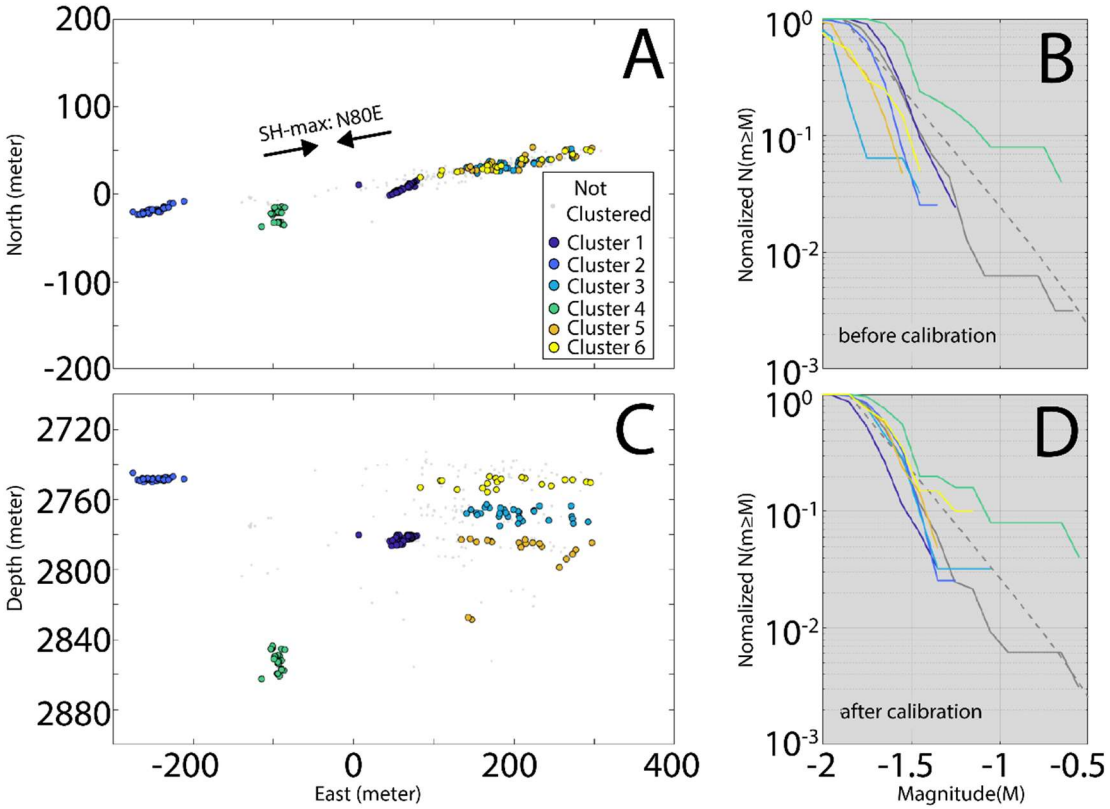

Figure 2.6: Comparison of the magnitude distribution for treatment B before and after magnitude calibration for different similar event clusters. A: map view of the seismicity location, SHmax is shown as a reference. C: cross-section view along EW direction. B and D show the Frequency Magnitude Distribution for the full sub-catalog of treatment B (grey solid line) and sub-clusters before (B) and after (D) calibration (solid lines with colors corresponding to legend in A). For comparison, the dashed lines show the b-values estimated for the full sub-catalog of treatment B. Vertical exaggeration is 2:1.

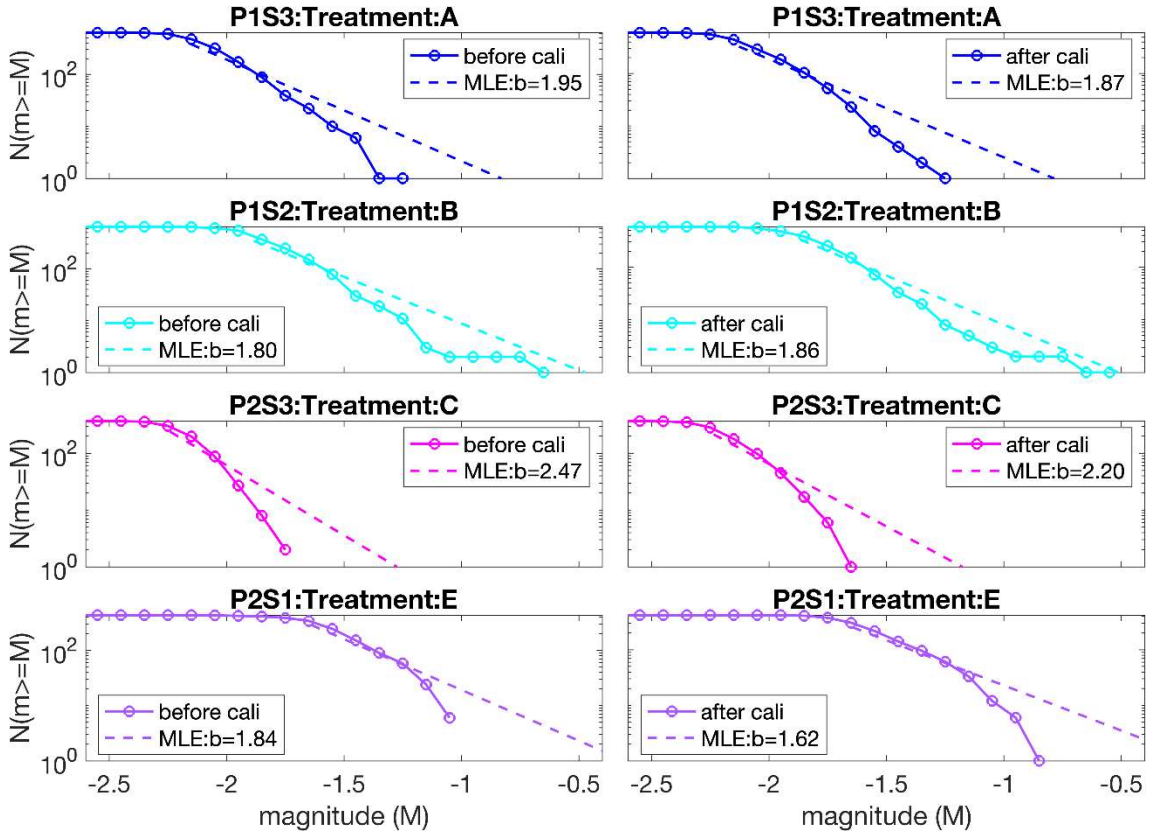

Figure 2.7: Comparison of Frequency-Magnitude-Distribution and corresponding b-values for treatments before(left) and after(right) magnitude calibration, where b-values are estimated by maximum likelihood (MLE). The vertical axis is the cumulative number of events and the horizontal axes is the event magnitude. Each treatment is color coded the same as in Figure 2.1.

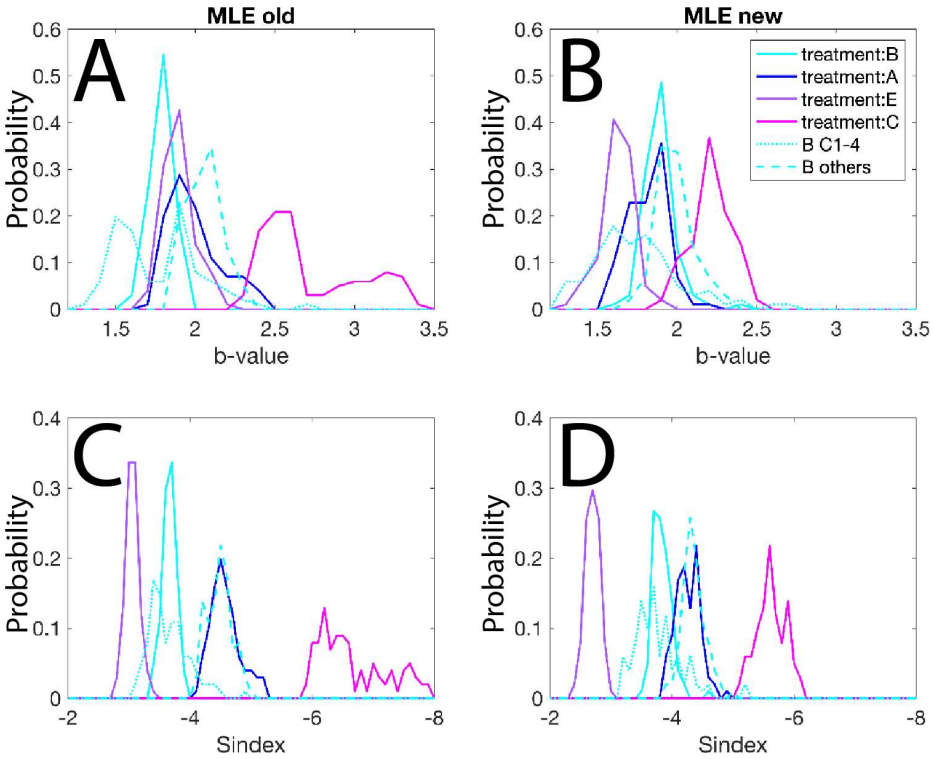

Figure 2.8: Comparison of the probability density function of b-value estimations from MLE (A and B), and the seismogenic index (C and D) before (left panel: A and C) and after (right panel: B and D) magnitude calibration. For all estimates, 100 bootstrap trials are calculated with MLE estimates for the b-value and seismogenic index.

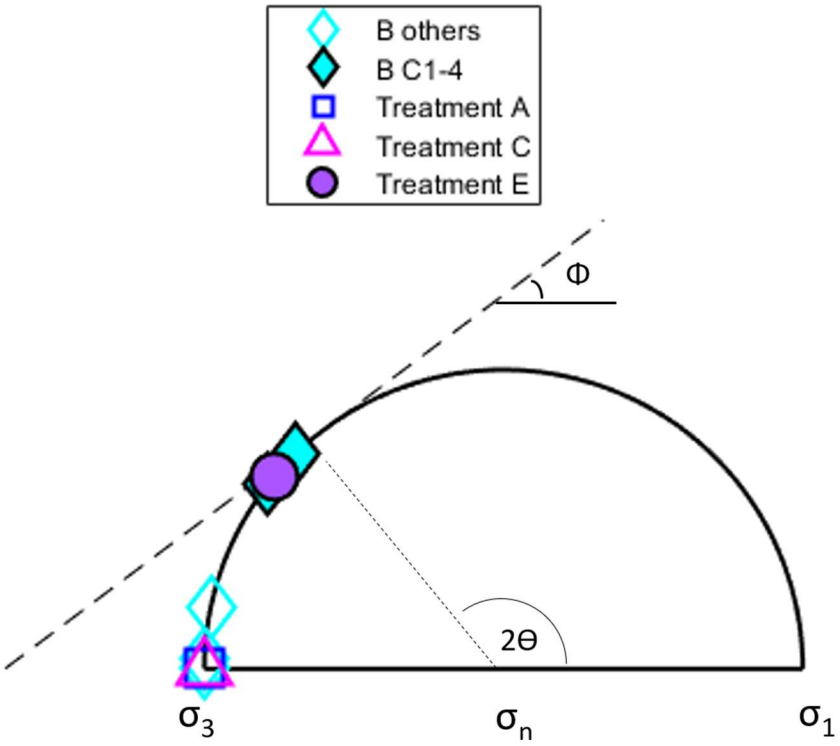

Figure 2.9: Comparison of different treatments and subclusters of treatment B as represented by a schematic Mohr circle. The dashed line shows the Coulomb Failure Envelope based on a friction coefficient of 0.75. Different symbols represent different clusters. Solid-filled symbols represent clusters and treatments with off-SH-max axis orientation. Non-filled symbols represent seismicity groups that are parallel to SH-max.

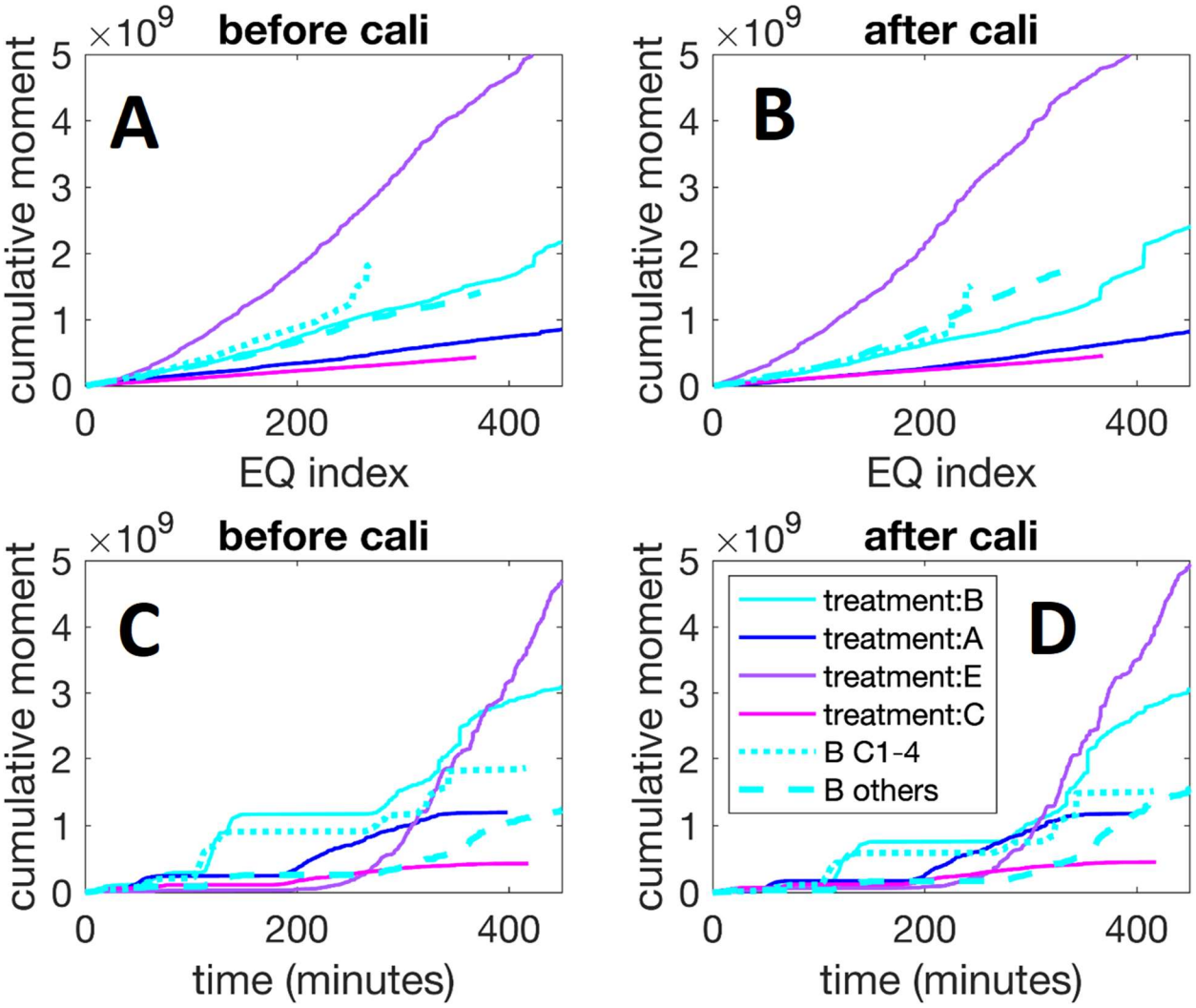

Figure 2.10: Cumulative moment growth with the increasing number of earthquakes (A and B) and time since the start of injection (C and D) for each treatment before (left panel: A and C) and after (right panel: B and D) magnitude calibration.

References

Abercrombie, R. E., Trugman, D. T., Shearer, P. M., Chen, X., Zhang, J., Pennington, C. N., Ruhl, C. J. (2021). Does Earthquake Stress Drop Increase With Depth in the Crust? Journal of Geophysical Research: Solid Earth, 126 (10). doi:10.1029/2021JB022314

Albright, J. N., Alamos, L., Laboratory, N., & Pearson, C. F. (1982). Acoustic Emissions as a Tool for Hydraulic Fracture Location: Experience at the Fenton Hill Hot Dry Rock Site. Society of Petroleum Engineers Journal , 22 (04),523–530.

Andrews, D. J. (1986). Objective Determination of Source Parameters and Similarity of Earthquakes of Different Size. Earthquake Course Mechanics, 37 , 259–267. doi:10.1029/GM037P0259

Bachmann, C. E., Foxall, W., & Daley, T. (2014). Comparing Induced Seismicity On Different Scales. PROCEEDINGS, Thirty-Ninth Workshop on Geothermal Reservoir Engineering.

Bachmann, C. E., Wiemer, S., Goertz-Allmann, B. P., & Woessner, J. (2012, May). Influence of pore-pressure on the event-size distribution of induced earthquakes. Geophysical Research Letters, 39 (9), 9302. doi:10.1029/2012GL051480

Bachmann, C. E., Wiemer, S., Woessner, J., & Hainzl, S. (2011). Statistical analysis of the induced Basel 2006 earthquake sequence: introducing a probability-based monitoring approach for Enhanced Geothermal Systems. Geophysical Journal International , 186 (2), 793–807. doi:10.1111/j.1365-246X.2011.05068.x

Boroumand, N., & Eaton, D. W. (2012). Comparing energy calculations - Hydraulic fracturing and microseismic monitoring. 74th European Association of Geoscientists and Engineers Conference and Exhibition 2012 Incorporating SPE EUROPEC 2012: Responsibly Securing Natural Resources , 4827–4831. doi:10.3997/2214-4609.20148187

Chen, X., Haffener, J., Goebel, T. H. W., Chang, J. C., & Al, C. E. T. (2018). Temporal Correlation Between Seismic Moment and Injection Volume for an Induced Earthquake Sequence in Central Oklahoma. Journal of Geophysical Research: Solid Earth, 123 (4), 1–18.

Chung, D., & Bernreuter, D. (1981). Regional Relationships Among Earthquake Magnitude Scales. Reviews of Geophysics, 19 (4), 649–663.

Dinske, C., & Shapiro, S. A. (2016). Performance test of the Seismogenic Index model for forecasting magnitude distributions of fluid injection induced seismicity. SEG Technical Program Expanded Abstracts 2016, 5022–5027.

Dinske, C., Shapiro, S. A., Dinske, C., & Shapiro, S. A. (2013). Seismotectonic state of reservoirs inferred from magnitude distributions of fluid-induced seismicity. Journal of Seismology, 17 (1), 13–25. doi:10.1007/s10950-012-9292-9

Eaton, D. W., & Boroumand, N. (2013). Estimating energy balance for hydraulic fracture stimulations: Lessons Learned from Basel. GeoConvention2013: Integration.

Eaton, D. W., Davidsen, J., Pedersen, P. K., & Boroumand, N. (2014). Breakdown of the Gutenberg-Richter relation for microearthquakes induced by hydraulic fracturing: influence of stratabound fractures. Geophysical Prospecting, 62 (4), 806–818. doi:10.1111/1365-2478.12128

Frohlich, C., & Davis, S. D. (1993). Teleseismic b values; or, much ado about 1.0. Journal of Geophysical Research, 98 (b1), 631–644. doi:10.1029/92JB01891

Gerstenberger, M., Wiemer, S., & Giardini, D. (2001). A systematic test of the hypothesis that the b value varies with depth in California. Geophysical Research Letters, 28 (1), 57–60. doi: 10.1029/2000GL012026

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. Geophysical Journal International, 165 (1), 149–166. doi: 10.1111/J.1365-246X.2006.02865.X

Gutenberg, B., & Richter, C. F. (1942). Earthquake Magnitude, Intensity, Energy, and Acceleration. Bulletin of the Seismological Society of America, 32 (3).

Hajati, T., Langenbruch, C., & Shapiro, S. A. (2015). A statistical model for seismic hazard assessment of hydraulic-fracturing-induced seismicity. Geophysical Research Letters, 42 (24), 10,601–10,606. doi:10.1002/2015GL066652

Heyman, J., de Coulomb, C. A., & Coulomb, C. A. (1972). Coulomb's memoir on statics: an essay in the history of civil engineering. CUP Archive.

Hubbert, M. K., & Willis, D. G. (1957). Mechanics of Hydraulic Fracturing. Society of Petroleum Engineers, 210 , 153–168.

Ibanez, J. M., De Angelis, S., Dıaz-Moreno, A., Hernandez, P., Alguacil, G., Posadas, A., & Perez, N. (2012). Insights into the 2011-2012 submarine eruption off the coast of El Hierro (Canary Islands, Spain) from statistical analyses of earthquake activity. Geophysical Journal International Geophys. J. Int, 191 , 659–670. doi:10.1111/j.1365-246X.2012.05629.x

Kagan, Y. Y. (1999). Universality of the Seismic Moment-frequency Relation. Pure appl. geophys, 155 , 537–573. doi:10.1007/s000240050277

Kanamori, H. (1977). The energy release in great earthquakes. Journal of geophysical research, 82(20), 2981-2987.

Kane, D. L., Kilb, D. L., & Vernon, F. L. (2013). Selecting empirical Green’s functions in regions of fault complexity: A study of data from the San Jacinto Fault Zone, Southern California.

Bulletin of the Seismological Society of America, 103 (2 A), 641–650. doi:10.1785/0120120189

Langenbruch, C., & Zoback, M. D. (2016). How will induced seismicity in Oklahoma respond to decreased saltwater injection rates? Science Advances, 2 (11), doi: 10.1126/SCIADV.1601542

Lei, X., Yu, G., Ma, S., Wen, X., & Wang, Q. (2008). Earthquakes induced by water injection at 3 km depth within the Rongchang gas field, Chongqing, China. Journal of Geophysical Research: Solid Earth, 113 (B10), 10310. doi:10.1029/2008JB005604

Maxwell, S. C., & Cipolla, C. (2011). What Does Microseismicity Tell Us About Hydraulic Fracturing? SEG Technical Program Expanded Abstracts 2011, 1565–1569. doi:10.2118/146932-MS

Michael Cleveland, K., & Ammon, C. J. (2015). Precise relative earthquake magnitudes from cross correlation. Bulletin of the Seismological Society of America, 105 (3), 1792–1796. doi:10.1785/0120140329

Mignan, A., & Woessner, J. (2012). Estimating the magnitude of completeness for earthquake catalogs. Community Online Resource for Statistical Seismicity Analysis, 1 . doi:10.5078/corssa-00180805

Milojevic, S. (2010). Power law distributions in information science: Making the case for logarithmic binning. Journal of the American Society for Information Science and Technology , 61 (12), 2417–2425. doi:10.1002/asi.21426

Mohr, O. (1900). Welche Umstände bedingen die Elastizitätsgrenze und den Bruch eines Materials. Zeitschrift des Vereins Deutscher Ingenieure, 46(1524-1530), 1572-1577.

Qin, Y., Chen, X., Walter, J. I., Haffener, J., Trugman, D. T., Carpenter, B. M., Kolawole, F. (2019). Deciphering the Stress State of Seismogenic Faults in Oklahoma and Southern Kansas Based on an Improved Stress Map. Journal of Geophysical Research: Solid Earth, 124 (12), 12920–12934. doi:10.1029/2019JB018377

Reyes, C. G., & West, M. E. (2011). The Waveform Suite: A Robust Platform for Manipulating Waveforms in MATLAB. Seismological Research Letters, 82 (1), doi:10.1785/gssrl

Rutledge, J. T., & Phillips, W. S. (2003). Hydraulic stimulation of natural fractures as revealed by induced microearthquakes, Carthage Cotton Valley gas field, east Texas. Geophysics, 68 (2), 441–452. doi:10.1190/1.1567214

Rutledge, J. T., Phillips, W. S., & Mayerhofer, M. J. (2004). Faulting induced by forced fluid injection and fluid flow forced by faulting: An interpretation of hydraulic-fracture microseismicity, Carthage Cotton Valley gas field, Texas. Bulletin of the Seismological Society of America, 94 (5), 1817–1830. doi:10.1785/012003257

Rutledge, J. T., & Urbancic, T. (1999). Cotton Valley Imaging Project Phase IV Completion of Microseismic Mapping of the Phase I and II Hydraulic Fractures Wells CGU21-10 and CGU21-09.

Saltiel, S., Boyle, K., Majer, E., Sciences, E., & Berkeley, L. (2011). Challenges in Determining b Value in the Northwest Geysers. Earth Sciences Division, Lawrence Berkeley National Lab, LBNL(4594E), 8.

Schaff, D. P., & Richards, P. G. (2014). Improvements in magnitude precision, using the statistics of relative amplitudes measured by cross correlation. Geophysical Journal International, 197 (1), 335–350. doi:10.1093/GJI/GGT433

Schorlemmer, D., Wiemer, S., & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. Nature, 437 (7058), 539–542. doi:10.1038/nature04094

Seeburger, D. A., & Zoback, M. D. (1982). The distribution of natural fractures and joints at depth in crystalline rock. Journal of Geophysical Research: Solid Earth, 87 (B7), 5517–5534. doi:10.1029/JB087IB07P05517

Shapiro, S. A., Dinske, C., & Kummerow, J. (2007). Probability of a given magnitude earthquake induced by a fluid injection. Geophysical Research Letters, 34 (22), 22314. doi: 10.1029/2007GL031615

Shapiro, S. A., Dinske, C., Langenbruch, C., & Wenzel, F. (2010). Seismogenic index and magnitude probability of earthquakes induced during reservoir fluid stimulations. Leading Edge (Tulsa, OK), 29 (3), 304–309. doi:10.1190/1.3353727

Shapiro, S. A., Huenges, E., & Borm, G. (1997). Estimating the crust permeability from fluid-injection-induced seismic emission at the KTB site. Geophysical Journal International, 131 (2), F15–F18. doi:10.1111/J.1365246X.1997.TB01215.X

Shelly, D. R., Ellsworth, W. L., & Hill, D. P. (2016). Fluid-faulting evolution in high definition: Connecting fault structure and frequency-magnitude variations during the 2014 Long Valley Caldera, California, earthquake swarm. Journal of Geophysical Research, 121 (3), 1776–1795. doi:10.1002/2015JB012719

Shelly, D. R., Taira, T., Prejean, S. G., Hill, D. P., & Dreger, D. S. (2015). Fluid-faulting interactions: Fracture-mesh and fault-valve behavior in the February 2014 Mammoth Mountain, California, earthquake swarm. Geophysical Research Letters, 42 (14), 5803–5812. doi:10.1002/2015GL064325

Skoumal, R. J., Brudzinski, M. R., & Currie, B. S. (2015). Earthquakes Induced by Hydraulic Fracturing in Poland Township, Ohio. Bulletin of the Seismological Society of AmericaSociety of America, 105 (1), 189–197. doi:10.1785/0120140168

Thompson, G. (2017). GISMO: A MATLAB toolbox for seismic research, monitoring and education Volcano-seismic alarm systems View project Interactions between high intensity rainfall and volcanic activity View project. 2017 AGU Fall Meeting. doi:10.13140/RG.2.1.3085.7202

Urbancic, T. I., Trifu, C.-I., Mercer, R. A., Feustel, A. J., & Alexander, J. A. G. (1996). Automatic Time-Domain Calculation of Source Parameters for the Analysis of Induced Seismicity. Bulletin of the Seismological Society of America, 86 (5), 1627–1633.

Walker, R. N. J. (1997). Cotton Valley Hydraulic Fracture Imaging Project. SPE Annual Technical Conference and Exhibition.

Walker, R. N. J., Zinno, R. J., Gibson, J. B., Urbancic, T., & Rutledge, J. (1998). Carthage Cotton Valley Fracture Imaging Project - mag ng Methodology and Implications. Society of Petroleum Engineers (49194), 599–612.

Walter, J. I., Chang, J. C., & Dotray, P. J. (2017). Foreshock seismicity suggests gradual differential stress increase in the months prior to the 3 September 2016 M w 5.8 Pawnee earthquake. Seismological Research Letters, 88(4), 1032-1039.

Wiemer, S., & Malone, S. (2001). A SOFTWARE PACKAGE TO ANALYZE SEISMICITY: ZMAP. Seismological Research Letters, 72 (2).

Wiemer, S., & McNutt, S. R. (1997). Variations in the frequency-magnitude distribution with depth in two volcanic areas: Mount St. Helens, Washington, and Mt. Spurr, Alaska. Geophysical Research Letters, 24 (2), 189–192. doi:10.1029/96GL03779

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal Component Analysis. Chemometrics and intelligent laboratory systems, 2 (1-3), 37–52.

Wyss, M., Schorlemmer, D., & Wiemer, S. (2000). Mapping asperities by minima of local recurrence time: San Jacinto-Elsinore fault zones. Journal of Geophysical Research: Solid Earth, 105 (B4), 7829–7844. doi:10.1029/1999jb900347

Zinno, R., Gibson, J., Walker, R. N. J., & Withers, R. J. . (1998). Overview : Cotton Valley Hydraulic Fracture Imaging Project. SEG Expanded Abstracts, 1996–1999.

S2

Supplementary materials

The supplemental materials provided in this section aim to enhance the understanding of the research presented on hydraulic fracturing magnitude calibration improvements from waveform correlation. We describe in greater depth the b-value estimation using maximum likelihood and robust least square approximation and magnitude completeness protocols used in Matlab's Z-map for this study.

S2.1 Maximum likelihood B-value estimation

To determine the b-value, which describes the frequency-magnitude distribution of earthquakes, seismic studies frequently employ the statistical method known as maximum likelihood estimation (MLE). The relationship between the frequency and size of earthquakes in a given area is represented by the b-value estimated in **equation S2.1** based on Utsu's (1965) formula with the ratio of logarithmic Euler's constant, e, to the difference of the mean magnitude, $\bar{M}$, and half the magnitude binning, ΔM, subtracted from minimum magnitude, M_1. The fundamental idea behind MLE is to identify the parameter values that maximize the probability of seeing the provided data. MLE seeks to identify the most likely b-value that best fits the observed earthquake data in the context of b-value estimation, offering important insights into the seismicity of a region.

$$b = \frac{\log_{10} e}{\bar{M} - (M_1 - \frac{\Delta M}{2})} \quad \text{(S2.1)}$$

An earthquake magnitude dataset is necessary to apply MLE for b-value estimation. The magnitude range must first be divided into predetermined magnitude bins, ΔM, as the first step. The number of earthquakes within each bin is counted, and each bin represents a particular magnitude range. The likelihood function, which measures the likelihood of observing the given dataset for a particular b-value, is then computed using the observed earthquake frequencies in each magnitude bin. The Gutenberg-Richter formula, which denotes the relationship between earthquake magnitudes and frequencies, serves as the foundation for the likelihood function.

S2.2 B-VALUE ESTIMATION USING ROBUST LEAST SQUARE

The seismic analysis code is limited by the number of common events between the listed catalog and actual waveform data available. We apply the robust least square (**equation S2.2**) that reweights based on the residuals (**equation S2.3**), where (Han et al., 2015) show improvement in the robust least square over traditional least square solutions. Robust linear regression uses weights (**equation S2.4**) as part of the scaling factor which reduces the sensitivity to outliers and improves fit. Weights are automatically and iteratively determined where initially, weights are equal and then reweighted in each subsequent iteration, giving lower weights to points further from the previously iterated model predictions (**equation S2.5**). Iterations are terminated at the point of convergence between coefficients estimates within a given tolerance value. Dumouchel and O'Brien (1989) integrate the bisquare robust regression used in the ZMAP MATLAB package (Wiemer and Malone, 2001) as the function "robust-fit" (Dumouchel and O'Brien, 1989; Street et al., 1988; Holland and Welsch, 1977; Huber, 1981). This function determines the best fit by iteratively using the robust least squares with the bisquare weighting function by the following procedure:

1. Equal weight least square regression as expressed in the least square line (**equation S2.2**), where x_i is the magnitude, y_i , event count, ϵ_i, unobserved errors, and β, is the least squares estimate.

$$y_i = x_i\beta + \epsilon_i \quad i = 1, \dots, n \tag{S2.2}$$

2. Compute and adjust residuals, u_i, using Huber's function where h_i is the leverage that reduces weight for high leverage points, β_R, robust regression, S, scale estimate, and c, tuning constant of 4.685 for bisquare weighting.

$$u_i = \frac{y_i - x_i\widehat{\beta_R}}{cS\sqrt{1-h_i}} \tag{S2.3}$$

3. Compute bisquare weights, w_i, based on previous iteration residuals.

$$w = \begin{cases} (1-u_i^2)^2, & if\ |u_i| < 1 \\ 0, & if\ |u_i| \geq 1 \end{cases} \tag{S2.4}$$

4. Update the least square line by performing a weighted least squares regression with the weights

$$\sum_{i=1}^{n} w_i(y_i - x_i\hat{\beta}_R)x_{ij} \quad j = 1, \dots, p$$

(S2.5)

5. Iterate until convergence.

S2.3 LST results

We also observe the results from using LST in replacement of MLE in **Figure S2.1**. All steps are identical to the main text where MLE is used.

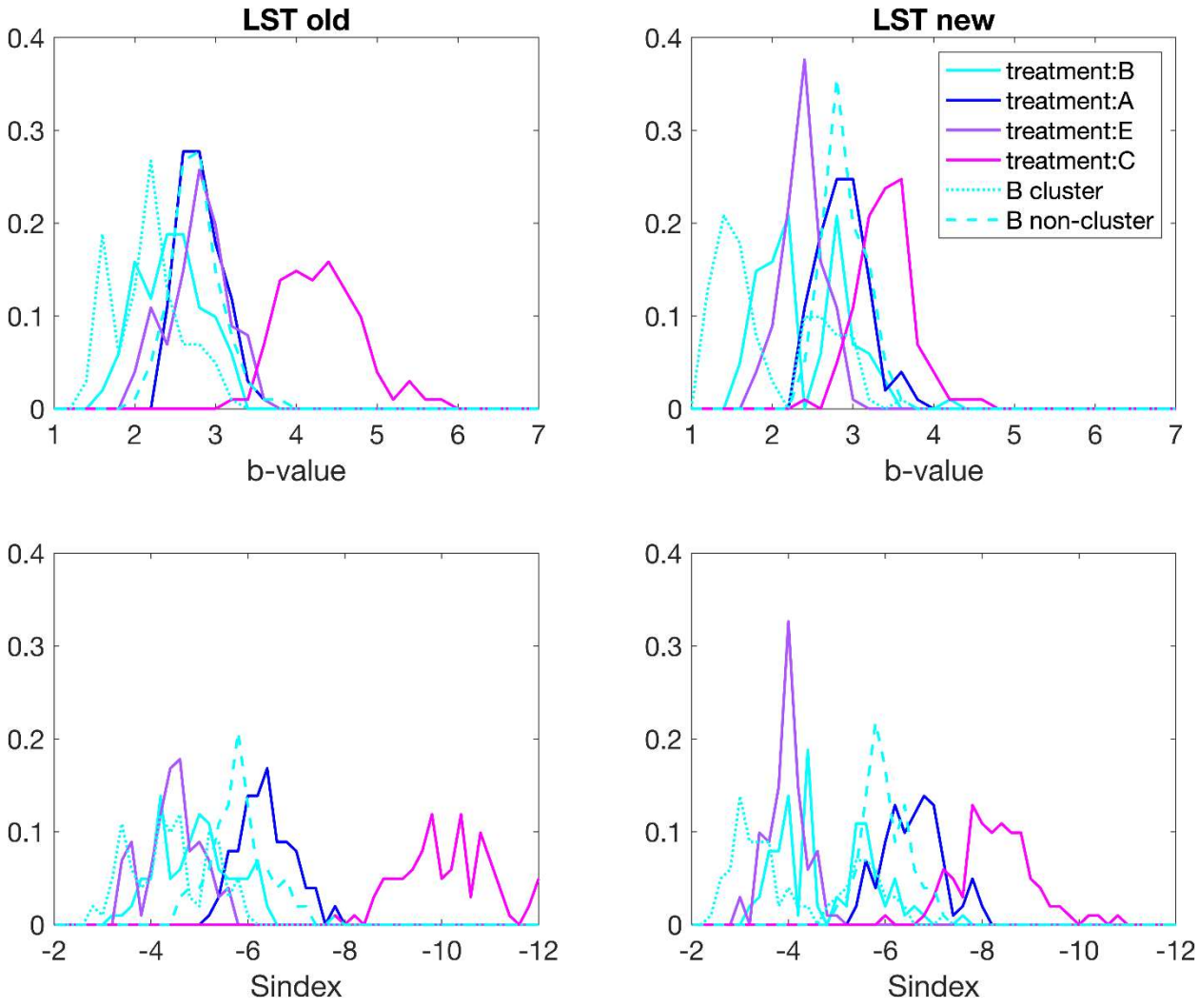

Figure S2.1: Comparison of probability density function of b-value estimations from LST (A and B), and the seismogenic index (C and D) before (left panel: A and C) and after (right panel: B and D) magnitude calibration. For all estimates, 100 bootstrap trials are calculated with LST estimates for the b-value and seismogenic index.

Chapter 3:

Site response of ambient seismic noise of Northern Oklahoma from the perspective of Nodal and DAS arrays

3.1 Abstract

The increased earthquake risk in a growing society is always prevalent and although Oklahoma is not known for large and damaging earthquakes, in the last decade it has experienced a more than 100-fold increase in smaller seismicity and the occasional moderate magnitude damaging earthquake. Though the frequent small events present themselves as mostly nuisance earthquakes, they are still influenced by site response that influences seismic damage and loss in urban areas. It is crucial to identify site characteristics like resonance frequency in Oklahoma at a high spatial resolution to account for site response, enhance our understanding of the destructive ground motion produced in earthquakes, and develop better seismic hazard assessment and mitigation in developed areas. We monitored an East-West transect near Enid, Oklahoma using two emerging instrument types, nodal sensors and distributed acoustic sensing for a higher spatial resolution understanding of site response. We compare the two instrument types to identify their strengths and weaknesses while in use. Finally, we measure site-specific fundamental frequencies where dominant peak fundamental frequencies are approximately 1Hz along the transect and estimate resonating layers using an existing velocity model.

3.2 Introduction: Environmental and earthquake monitoring

The high resource demand from a growing society has driven the development of infrastructure and residential buildings to accommodate the increased population density in many cities much like those in the state of Oklahoma. In Oklahoma, earthquake hazard is an increasing concern linked to the increase in wastewater injection (Ellsworth, 2013; Keranen et al., 2014) with three events larger than M5 occurring within the last decade. Few of these occurred on known faults as shown in **Figure 3.1.** However, most events occur on unmapped faults (Schoenball and Ellsworth, 2017a; Qin et al., 2019) which led to deployments of dense seismic arrays to monitor and manage the growing earthquake hazard in the region. In some cases, dense seismic arrays were deployed to monitor the increasing activity. A dense seismic array consists of multiple closely spaced seismic sensors that are strategically placed to monitor ground motion and seismic waves. These typically consist of nodal sensors, a self-contained data logger, a battery, and a sensor, for wavefield experiments. Examples of high-density arrays using nodal sensors for wavefield monitoring include over 1800 nodal sensor LArge-n Seismic Survey in Oklahoma (LASSO) array (Dougherty et al., 2019); the IRIS Community Wavefield Experiment in Oklahoma (Sweet et al., 2018) using 9 infrasound sensors, 18 broadband sensors, and 363 nodal sensors; and the Long Beach Dense array using 5200 nodal stations (Li et al., 2015). High-density arrays such as the ones previously mentioned have been increasing in occurrence due to their improved imaging resolution and desire to capture the unaliased spatial wavefield from earthquake sources. Amongst other insights, this is useful for earthquake monitoring as it allows for precise and accurate detection, location, and characterization of seismic events that would not otherwise be possible with the existing regional network. By

having a dense network of sensors, scientists can obtain high-resolution earthquake catalogs that, in part, provide a better understanding of the underlying geological structures and processes that may trigger earthquakes.

It is important to understand the local geology in present-day Oklahoma in any seismological survey, to make connections and bridge the knowledge gap between the convoluted geologic and seismo-tectonic history of the region. The Oklahoma crust, much like other areas of the North American craton, is the result of multiple tectonic activities that were especially active between the Proterozoic and Cambrian (Whitneyer and Karlstrom, 2007). The location of our study is east of the southern extent of the Nemaha uplift. The Nemaha uplift, a narrow paleo structural feature that stretches from southeastern Nebraska to Central Oklahoma, is 500 miles in length at 80 miles at its widest point. The structural history of the Nemaha uplift has repeated periods of regional warping, uplift, and erosion separated by periods of marine shelf sedimentation. The Nemaha uplift is a composite of several complex elements and features of separate fault blocks. At its surface, it is a gentle anticline plunging towards the south. The formation of the Nemaha uplift is attributed to left-lateral wrench fault movement as a crustal response to convergence at the continental margins (Berendsen 1986). Others have suggested similar strike-slip movements on the Central Oklahoma fault zone (Amsden, 1980) and uplift to vertical movement on a pre-existing Precambrian zone of weakness (Fath, 1920). Present-day interest in understanding the structure of the region is driven by hydrocarbon exploration of the Cambrian to Permian age rocks. The succession of paleozoic dolomites and limestones with subordinate clastics overlay the Precambrian crystalline basement. In the midcontinent, the ascending order of the stratigraphic sedimentary units is the Joins, Oil Creek, McLish, Tulip

Creek, and the Bromide which are a series of carbonates, shales, and sandstones. This unique depositional history and uplift create hydrocarbon plays which have been the economical target for oil and gas production due to the structural traps and reservoir rocks (Dolton and Finn, 1989).

While a dense seismic array will typically be utilized for passive detection of earthquake sources or during the acquisition of active-seismic data, the data can also be utilized for environmental monitoring through ambient seismic noise. Ambient noise is caused by natural sources such as ocean waves, wind, and atmospheric pressure, as well as cultural sources such as traffic and industrial activities. Ambient seismic noise studies have been used for over a century and can be traced back to Ernst von Rebeur-Paschwitz observations of the first recorded teleseismic earthquakes using a horizontal pendulum left to oscillate freely (Von Rebeur-Paschwitz, E., 1889).

In recent years, there is a resurgence in passive noise monitoring due to the greater availability of a wide variety of sensors such as the inexpensive Fairfield 3C Zland nodal sensor which leverages its compact self-contained size. Smaller inexpensive sensors and ease of deployment for passive noise experiments also reduce the requirement for local municipality permitting for active source experiments. By analyzing this ambient noise, valuable insights can be gained into various environmental factors such as soil properties, and groundwater levels (Larose et al., 2015). One property we are interested in our high-density observation of ambient seismic noise is the physical phenomenon of resonance. Resonance is a constructive interference event during wave propagation. In a ground motion, resonance can result in increased shaking intensity by having a larger wave amplitude. We compare different

instruments in this study and perform an analysis of the noise signals recorded by each instrument type in the time and frequency domain. Furthermore, the dense arrays are leveraged in an east-west site response study.

3.3 Array information

For this study, we deploy instruments along an East to West transect of the entire Garfield County in Oklahoma along highway US-412, a 4 lane two-way highway separated by a road verge, over three separate occasions in 2020 (EW1-40: June 09, 2020 - July 14, 2020) and 2021 (EW1-20A/B: April 21, 2021 - May 13, 2021, and EW2: August 1, 2021 – August 31, 2021) along the length of highway 412 and minor perpendicular arrays (NS1 and NS2: April 21, 2021 - May 13, 2021) in 2021 as shown in **Figure 3.2**. We perform a multi-deployment of temporary high-density nodal sensors that recorded continuous waveform for a one-month duration during each deployment time. We use the 2nd generation Fairfield Nodal Z-land sensors which are small, self-contained devices that are deployed in seismic exploration to detect and continuously measure continuous ground vibrations from anthropogenic sources created by a Vibroseis truck or natural sources such as local earthquake events for about a month or until the battery is depleted (**Figure 3.2)**. Nodes are easy to deploy and have a low environmental impact. However, nodal sensors have limited recording capacity and are prone to signal noise from improper isolation during deployment. Nodes are a 5Hz sensor and we set the recording sampling rate at 500Hz with a 12-decibel pre-amp gain. The recorded data are downsampled to 100Hz to conserve storage and improve computational resource management. In **Figure 3.3**, we show the day and night noise power distribution comparison between the two seismic instruments, nodal station 4013 and the corresponding fiber optic strain rate sensor. Elevated

daytime noise is observed in both recording types and is likely due to nearby anthropogenic sources. We deployed both in 2020 and twice in 2021 in a rolling deployment, details will be explained in subsequent subsections. For the 2021 deployments, we deploy the nodal sensors concurrently with a Distributed Acoustic Sensing (DAS) deployment using existing fiber optic cables from an Oklahoma telecommunication sub-terranean conduit-lain fiber optic bundle. The utilization of existing utility telecommunications cables for other purposes has been called dark fiber where Marra et al. (2018) used submarine telecommunication fiber lain on the seafloor for earthquake detection and location. As shown in **Figure 3.2**, we use the server room housing the fiber optic terminal ends at Northwestern Oklahoma State University as a midpoint for the arrays. The DAS unit we are using is the Silixa brand iDAS interrogator, a dynamic range sensing system that determines strain rate and distance from measuring phase change in the Rayleigh scattering signal. We record DAS measurements at the iDAS's maximum range of 1 kHz with a channel spacing of 2m and we later down-sample our recordings to 50Hz for noise analysis.

3.3.1 Deployment 1 - EW1-40 (nodal deployment)

We deploy 124 nodal sensors between June 09, 2020 - July 14, 2020, east of Northwestern Oklahoma State University (**Figure 3.2**). Initially, the plan for the deployment is about 400m spacing, but actual spacing varies from 300m to 400m spacing on account of GPS error and GPS updates when traversing the deployment line during deployment.

3.3.2 Deployment 2 - EW1-20A & B (Rolling nodal deployment and concurrent DAS)

As shown in **Figure 3.2**, EW1-20A uses 103 sensors on the initial deployment where 43 sensors are on the EW line (43km), 38 sensors NS1(23km) on Highway 74, and 22 sensors in NS2 (14km) on Highway 81. EW line spacing was 1-1.3km spacing, both NW lines were approximately 500-700m spacing. We use a rolling deployment, so half of the deployed sensors were collected shortly after the initial deployment of EW1-20A. All sensors east of NS1 were collected and half of both NS lines were collected within 2 weeks of deployment of 103 sensors. Existing sensors were replaced within 2 weeks of collection for an extended deployment to be co-located with the iDAS array deployment. On the redeployment of the EW1-20B, all sensors in NS1 and within the EW line (21.3km) were replaced with recharged sensors. The density in NS2 remained the same (14km), but NS1 was shortened (10km) and the density in the EW line was increased to have a spacing of 500-700m. The DAS system was installed shortly after the nodal deployment and was operational between April 21, 2021 - May 13, 2021.

3.3.3 Deployment 3 - EW2 (nodal deployment and concurrent DAS))

A total of 61 sensors are used for deployment 3, where 26 sensors are deployed west of station 4001 for 22.5km (**Figure 3.2**). Additionally, 35 sensors are redeployed on NS2 replacing deployed sensors from EW1-20B and extending the North–South deployment line to 23.5km between August 1, 2021 – August 31, 2021. The DAS system was installed before the nodal array this time on July 19, 2021, until August 11, 2021.

3.4. Comparison of DAS and Nodal array performance

Distributed Acoustic Sensing (DAS) and Seismic Nodal Sensors are two different types of sensors used in seismic surveys to acquire data on subsurface structures. While both technologies are used to record seismic waves, there are some significant differences in their performance.

DAS is a type of distributed sensing that uses fiber optic cables to capture seismic waves by measuring the backscatter from laser pulses as it travels along the fiber optic cable. Different types of backscatters can be measured by distributed sensing which provides insight into different physical characteristics affecting the fiber optic cable. Rayleigh, Raman, and Brillouin scattering are all phenomena related to the scattering of light. These light backscatters differ in terms of the physical mechanism responsible for the scattering, the wavelengths of light involved, and the information that can be obtained from the scattered light. As a laser pulse, from the interrogator, propagates through a length of fiber, the individual photons are affected by the fiber itself. Rayleigh scatter is when the backscatter frequency is the same as the initial frequency, Raman scatter is when the scatter has increased (anti-stokes) or decreased (stokes) its energy level from photon interaction with the fiber's natural molecular vibration, and Brillouin scatter is caused by the physical expansion and contraction of the fiber from photon energy, generating an acoustic response from the fiber. Phase changes in Rayleigh scattered signal can measure strain rate, backscatter of Raman stokes and anti-stokes measures temperature (Farahani and Gogolla, 1999), and Brillouin scattering measures both temperature and strain. Stokes and anti-stokes differ in terms of wavelength, whereas stokes have a longer wavelength than anti-stokes (Conway and Mondanos, 2015). Raman scattering occurs when

light interacts with molecules or crystals throughout the core of the fiber over the entire fiber length (Conway and Mondanos, 2015) (Raman, 1928). It is analyzed through Raman optical time-domain reflectometry where the backscatter from a short laser pulse contains information on loss and temperature along the length of the fiber (Farahani and Gogolla, 1999). Brillouin scattering occurs when light interacts with the acoustic waves (i.e., sound waves) that propagate through a material. The scattered light undergoes a frequency shift that corresponds to the frequency of the acoustic wave, and this shift can be used to measure the speed of sound in the material (Boyd, 2008). The iDAS interrogator from Silixia in our study uses Rayleigh backscatter for measuring strain rate along the fiber.

Traditional geophone sensors consist of three orthogonal (perpendicular) sensors that measure ground motion in three directions X, Y, and Z (Murphy, 1996) of displacement at their location. The more recent nodal sensors have been deployed in arrays or clusters across a seismic survey area (Li et al., 2018) and are then connected to a central recording system that collects and analyzes the data. The sensors are designed to be low-power and low-maintenance, and they can be left in place for extended periods to collect data. The operation of the Fairfield-Zland 3C nodal sensor is based on the principles of piezoelectricity. Piezoelectric materials generate an electric charge in response to the mechanical displacement of coils within the geophone (Fairfield Nodal, 2017). The generated voltage is proportional to the displacement of the moving coils within each geophone. The three aforementioned geophones are installed within each nodal sensor oriented orthogonal to each other, so they can measure ground motion in three directions. Voltages are then digitally converted to discrete velocities representing ground motion affecting the nodal sensor. The data from the geophones are

typically processed using digital signal processing techniques to remove noise and interference and to extract useful information about the seismic signals. This information can be used to generate images of the subsurface geology, locate oil and gas deposits, and monitor seismic activity.

Considering a typical regional network of sparsely spaced broadband seismometers; both DAS and nodal arrays are a major improvement on instrumentation density. A key difference between DAS and Seismic Nodal Sensors is their cost. DAS is generally more expensive than Seismic Nodal Sensors due to the cost of the fiber optic cables and the specialized equipment required to operate the system. Seismic Nodal Sensors, on the other hand, are relatively inexpensive and can be deployed in large numbers to cover a wide area. Both DAS and Seismic Nodal Sensors have advantages and disadvantages in seismic surveys. DAS offers high-resolution data over long distances but at a higher cost. Seismic Nodal Sensors are inexpensive and easy to deploy, but have limited recording capacity and may suffer from signal distortion. The choice of sensor technology depends on the specific needs of the survey and the available budget.

3.4.1 Noise level comparison of DAS and nodal

We compare the two recorded measurements from both DAS and nodal. For computational resource management, we resample both datasets to a reasonable 50Hz sampling for DAS and 100Hz sampling for nodal. To determine the initial data quality of the recordings, we investigate the influences on the acoustic seismic field from anthropogenic sources. Due to the proximity of a town and roadway parallel deployment, we assume that

human activity will influence the seismic measurements during traditional daytime business hours. We calculate power spectral density for both night and daytime hours, stack, and average the resultant spectral graphs. **Figure 3.4** shows a time-averaged spectrum for DAS channels located near a highway overpass and rail crossing, demonstrating showing how noise levels can vary depending on infrastructure. As expected, high amplitude noise levels are observed near anthropogenic sights prone to high activity and better ground coupling such as highway overpasses and rail crossings. It is observed that daytime noise levels are higher than nighttime noise even in areas with mostly farmland as shown in **Figure 3.2.**

3.4.2 Short-duration events

Throughout the recording duration of the experiment, both ambient noise and transient signals were detected. Notably, short-duration high-amplitude events are recorded, such as passing vehicles, and a few local earthquake events (Walter et al., 2020) which are shown in **Figure 3.2**. In **Figure 3.5,** we observe graphically linear and high-amplitude signals over different channels for a period. Observed high-amplitude signals are linear across multiple channels, coherent, and inversely graphically oriented. An approximate velocity of 58 mph may be estimated for the high-amplitude signal using the channel spacing of 2 m and arrival time approximation at each channel. Considering the 70 mph speed limit of US-412, we can interpret the traveling signal source to likely be a moving vehicle. In addition, we identify an earthquake arrival at 11:18:01 local time on May 2, 2021, as seen in **Figure 3.5** as a high-amplitude signal with similar arrival time amongst all channels. In **Figure 3.6**, both time series and spectra of different sources, background noise, moving vehicles, and earthquakes within a 10-minute DAS recording are compared. Even though the earthquake and vehicle time series have similar peak amplitude

values, their spectra are different in both time duration and frequency range, where the earthquake spectrum occupies lower frequencies, and the moving vehicle spectrum has higher frequencies.

3.4.3 Long-duration signal

In addition, long-duration signals, such as trains, cross obliquely near the eastern half of the deployment close to nodal stations 4005 and 4006. The passing train signal does not present itself across the entire DAS and is only concentrated near the railroad crossing when there is a train present. Locating and identifying the train signal is performed best in the frequency domain since the identification of the train signal in the time series waveform is not obvious as depicted in **Figure 3.7,** where we isolate the train signal from other background signals by selecting an average spectral amplitude between 1-8hz. In the frequency domain, we average the frequency range 1-8hz together to exploit the broadband nature of the train signal against the average background noise. We identify high amplitudes in the average frequency stack between 1 Hz – 8 Hz by applying the kurtosis method, which is the ratio of a short-window average to a long-window average. However, to mitigate the false detection of trains we identify trains, with high spectral amplitude in consecutive periods over 120 seconds over 40 consecutive DAS channels. We transform the time series into the frequency domain for train detection to eliminate false detection created by consecutive noisy DAS channels or other noise sources that may generate signals within a long-time window.

3.4.4 Near–surface characteristics

Site response at any given site can vary significantly even within Northern Oklahoma as shown by the varying noise amplitudes over short distances in **Figure 3.3**. Determination of site characteristics, such as fundamental frequency and site amplification, is important to characterize potential earthquake hazards related to ground motion amplification in any microzonation study. Microzonation is the identification of potential for earthquake hazards in areas with an increased risk for ground motion amplification, liquefaction, and landslides. The purpose of microzonation studies is to assess these locations of potential risk and mitigate any potential economic loss such as damage to buildings or other structures (Ansal et al., 2009). The earliest microzonation site response study by Imamura (1913) utilized single-station microtremors following the 1854 Tokyo earthquake, where localized damage patterns were observed from amplified ground motions. Subsequent observations of large earthquakes have also shown localized damage patterns correlating to subsurface geologic structures such as basins (Frankel et al., 2002; Hall and Beck, 1986; Abbott, 2005; Kagami et al., 1986). Understanding this secondary hazard such as resonance has been an important feature in earthquake hazard risk reduction as shown in the well-documented 1985 Michoacan earthquake (Hall and Beck, 1986; Flores Estrella et al., 2006) where building structural damage was linked to earthquake shaking resonance, the 1987 Whittier earthquake (Hruby and Beresnev, 2003; Kawase and Aki, 1990) shown a basin edge effect in ground motion amplification, and the 1994 Northridge earthquake (Hruby and Beresnev, 2003; Beresnev et al., 1998; Hartzell et al., 1996) exhibiting basin resonance from excited ground motions. The geologic medium can influence the relative seismic acoustic wave amplitude which translates to

stronger or weaker shaking (Nakamura, 2000). Nakamura (1989) determined the cause of stronger localized shaking during a strong motion event is due to the harmonic wave constructively interfering through resonance.

There are several methods used to characterize the site response. Geotechnical investigations such as Cone Penetration Test (CPT), Standard Cone Penetration TEST (SPT), and soil borings are used in engineering applications for site characterization (Schmertmann, 1978). Information about the local surface geology is also important to consider as it can highly correlate with the observed seismic intensities (Wills et al., 2000). However, geotechnical investigations are often expensive, invasive, negatively affect the environment, and are not available for all locations. These methods are also focused on the determination of site amplification and do not consider the effects of resonance. In this study, we use the Horizontal-to-Vertical-Spectral-Ratio (HVSR) approach to ambient noise recorded on a high-density linear nodal array deployed in Enid, Oklahoma to estimate fundamental frequencies. It is known that deep sedimentary soil columns in basins can amplify earthquake ground motion and cause significant damage to buildings built on their surface. The extent of the damage is related to the building's specific height and soil site resonance frequency, fundamental frequency. We use the fundamental frequencies to estimate a geologic structure shape and determine subsurface sedimentary column resonance.

3.5 Method and Background for HVSR

The HVSR approach is used for site response investigations due to its ease of use, low cost, and flexibility in the use of data that is not dependent on strong ground motion or an

active source. Ambient noise HVSR studies have been conducted extensively and compared for reliability against other methods, as discussed by Bard et al. (2004) and Bard et al. (2005). Nakamura (1989) showed that the HVSR spectral peak frequency represents the fundamental frequency of the site soil column. This has been confirmed by later studies which tested the accuracy and reliability of the HVSR approach, as described in papers such as Cara et al. (2010), Guillier et al. (2007), and Parolai et al. (2004). Further investigation by Bard et al. (2004) in the Site Effects Assessment Using Ambient Excitation project (SESAME) compared the ambient noise HVSR results with those from earthquake-based HVSR analysis, and these results were further verified in Bard et al. (2004) and Bard et al. (2005). The primary objectives of the SESAME project were to better understand the physical basis of the HVSR approach, determine its purpose in site response, and propose guidelines for correct analysis. The SESAME project demonstrated a strong linear correlation between the spectral peak frequency determined through ambient noise HVSR and those determined at the same site through Standard Spectral Ratio (SSR) measurements from earthquake data (Bard et al., 2004). This result demonstrates that the peak frequency from the ambient noise HVSR may be interpreted as the expected peak frequency of earthquake-related ground motion. Based on this SESAME frequency comparison and investigations mentioned earlier, we will refer to the peak frequency measured from the HVSR graph as the site's fundamental frequency. The amplitude of the spectral ratio has been used by some authors as a representation of amplification relative to hard rock sites (e.g., Nakamura, 1989; Bard et al., 2004). The SESAME project also investigated the HVSR amplitude by comparison of peak amplitudes as determined by the HVSR and SSR methods. Although the HVSR and SSR peak amplitudes are not linearly related the HVSR peak amplitude can be

considered as a lower bound for site amplification (Bard et al., 2004). Thus, the HVSR method is a simple and effective technique for determining the first fundamental frequency of soil resonance. However, Bonnefoy-Claudet (2009) suggests the HVSR approach is inadequate for determining higher resonance modes, contrary to other assertions (Lermo et al., 1994).

In HVSR, the two horizontal components are combined using geometric mean in the numerator and the denominator is the vertical spectral component as exampled in **equation 3.1**.

$$H/V(f) = \sqrt{\frac{{H_T}^2 * {H_R}^2}{2V^2}} \tag{3.1}$$

where the root geometric mean square ratio of the two horizontal spectra, H, is taken over twice the vertical spectra, V.

One popular tool for HVSR is the open-source software, Geopsy (Wathelet et al., 2010). It is a user-friendly graphical user interface and has been well-tested by many. However, to improve on computational cost and speed we use the open-source Python HVSR software HVSRpy (Vantassel, 2020) which is comparable to the open-source Geopsy software (Wathelet et al., 2010) but accounts for azimuthal variability by using a frequency domain window rejection algorithm (Vantassel, 2020 and Cox et al., 2020). HVSRpy performs the rejection window by selecting unbiased spatial statistics, Voronoi tessellation, for fundamental site frequency. (Cheng et al., 2021). We observed the rejection of outlier spectral curves exampled in **Figure 3.8** and note the improved standard deviation of the estimated peak fundamental frequency.

3.5.1 Example data analysis

We determine the HVSR curve for each station and analyze the results per the SESAME guidelines in Chatelain et al. (2008) and Albarello et al. (2011) to implement the HVSR approach, which has three conditions for curve reliability, $H/V(f)$, and five criteria for identification of a peak frequency, f0, as a clear peak.

We use a 90-second time window, 30 Hz low-frequency passband, Konno and Ohmachi smoothing constant of 40, and a maximum of 50 iterations for spectral rejection in our calculations. The HVSR spectral curve in **Figure 3.8** shows a recording of the 11th hour on May 2nd calculated for HVSR using the aforementioned parameters.

3.5.2 HVSR variation along the array and interpretation

We use the HVSR spectral curve results in **Figure 3.9** and compare the coherent spectral peak frequencies with a 1D velocity profile to estimate depths of potential acoustic impedance. We use the 1D shallow velocity and forward model with spectral peak frequencies based on the quarter wavelength function of harmonics (**Figure 3.10**). The primary software identified a clear and reliable peak frequency of about 10 Hz, which is related to a depth layer at approximately 30 meters. Secondary HVSR peaks are interpreted to the geologic structure shown in **Figure 3.10**, which agrees with Caylor (1958) where nearby well correlations result in a basin structure that is shallower towards the East. On a local scale, Cary (1955) depicts a complex anticlinal structure along our deployment area which does support our interpreted structures at deeper depths.

A velocity model is created through passive processing of nodal seismic data as outlined in Behm et al. (2019) and Cheng et al. (2022), where noise cross-correlation is computed for all possible station pairs. Cross correlations are then applied to the multichannel analysis of surface waves (MASW) method (Park et al., 1998; Xia et al., 1999) for surface wave measurements from nodal data. Finally, frequency-time analysis (FTAN) (Bensen et al., 2007; Levshin et al., 1989; Hannermann et al., 2014) is applied to the recorded data. To obtain a velocity profile, we calculated dispersion curves from the interferometrically retrieved surface waves. Dispersion curves are then inverted to obtain a 1d velocity model, using a phase shift method (Xia et al., 1999). Dispersion curves are selected from energy trends and subsequently inverted for the 1D Vs model with Geopsy which follows a hybrid neighborhood algorithm (Wathelet et al., 2004).

Dense deployments are typically a drastic improvement over single-station observations. We increase the observation resolution significantly using nodal stations. However, even with nodal station density, details in recorded waveforms such as passing vehicles or trains would be easily overlooked when analyzing the recordings. Using very simple methods such as transforming data from time series into the frequency domain we can improve our observations and identification of anthropogenic noise sources such as moving vehicles and trains since they appear differently on the power spectrum (**Figure 3.6**). We can also compare our nodal results with those of DAS. Although DAS is recorded in strain rate, much of the processing and interpretation is identical to that of traditional acoustic seismic waveform measurements. With DAS, we can observe spatial temporal changes in the waveform stacks indicating features such as vehicles traveling in opposite directions as well as using kurtosis to identify a train passage.

However, with improvements in seismological sensing equipment, new logistical issues do appear that were not present prior in traditional single-station deployments. When using nodal stations, a crew of individuals is now required to effectively deploy each sensor. Adequate training for each individual, as well as systematic errors in each individual's deployment technique, may introduce unwanted deviation in recorded waveform. This issue and process are eliminated with DAS deployment where the primary setup is in the interrogator, but fiber cable installation and ground coupling may influence data quality.

In our DAS deployment, we are using dark fiber, which has potentially inconsistent coupling with the subsurface and inadequate fiber termination that introduced a significant amount of noise. Installing new fiber optic cables with good ground coupling may increase signal quality.

3.6 Conclusion

DAS application using dark fiber is an emerging technique in observational seismology. Observations of different sources such as trains, cars, and noise are recorded and compared between nodal stations and DAS channels. It is important to characterize the response and improve our understanding of dark fiber DAS and compare the measurements to traditional recording instrument types. In our observations, nodal stations do have stronger signal-to-noise ratios than DAS, likely due to ground coupling differences caused by the deployment of fiber optic cables and nodes. There is an inherent risk of data noise contamination from these surface effects due to the deployment methods. The high spatial density of stations allows for improved interstation resolution, which aids in providing information on slight variations in the wavefield between local sites.

We can see from the various HVSR spectral curves estimated from different recording times that the HVSR technique is rather stable and provides a good indication of the fundamental frequency of the subsurface. Comparing both the ambient noise HVSR and earthquake HVSR, we can see both are similar except for amplitude differences. This may be due to a lack of waveform recording length with earthquakes compared with the hundreds of hours of ambient noise. However, differences in HVSR amplitude between ambient noise and earthquake recordings are justified, also seen in Bard et al. (2004) where both earthquake HVSR and ambient noise HVSR are compared.

Using a velocity model, it is possible to determine a depth to acoustic impedance layer for the HVSR fundamental frequency measurements. Although primary HVSR studies focus on the prominent peak, it should be noted that coherent secondary peaks should not be ignored. In the forward model, we can interpret the assumed sedimentary layers on the western side of the Nemaha Ridge. We should not expect significant variations in the sedimentary structure of the adjacent Nemaha Ridge, but some variation in slope may be present.

Application of HVSR to telecommunication dark fiber DAS data is typically plagued with sensitivity issues due to poor ground coupling as well as a lack of instrument response that would allow for the conversion of the signal to scientific units of ground motion. This would allow for a more direct comparison with nodal measurements. However, the ability to successfully use an existing telecommunications fiber optic cable as a remote sensing network is significant to both research and commercial applications. To improve on and advance the method of dark fiber DAS, the instrument response of DAS must be resolved to easily compare

measurements to other instrument types. The popularity of using dark fiber for DAS is increasing and will become significant in the future as more studies include its deployment.

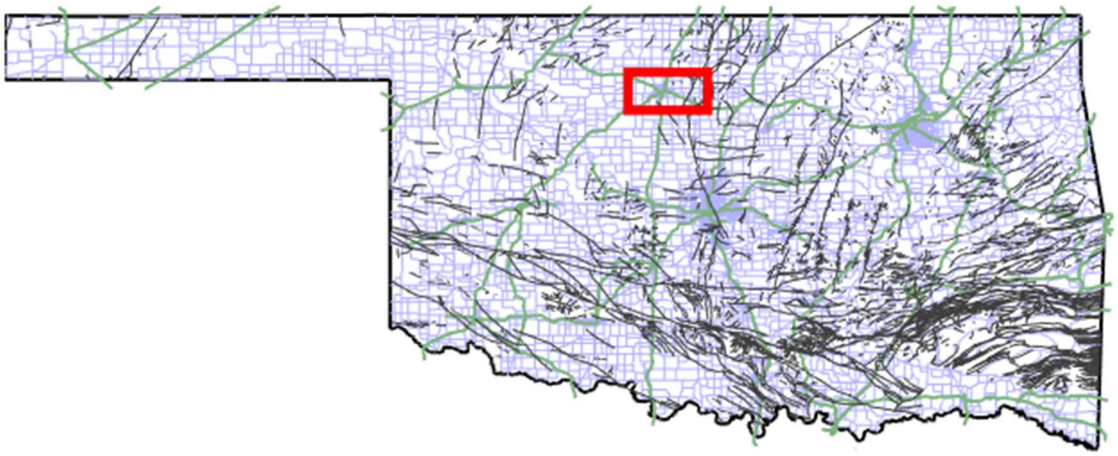

Figure 3.1: A map of Oklahoma with fault lines. The red box is the array location. Roads are shown as blue lines and railway tracks are green lines.

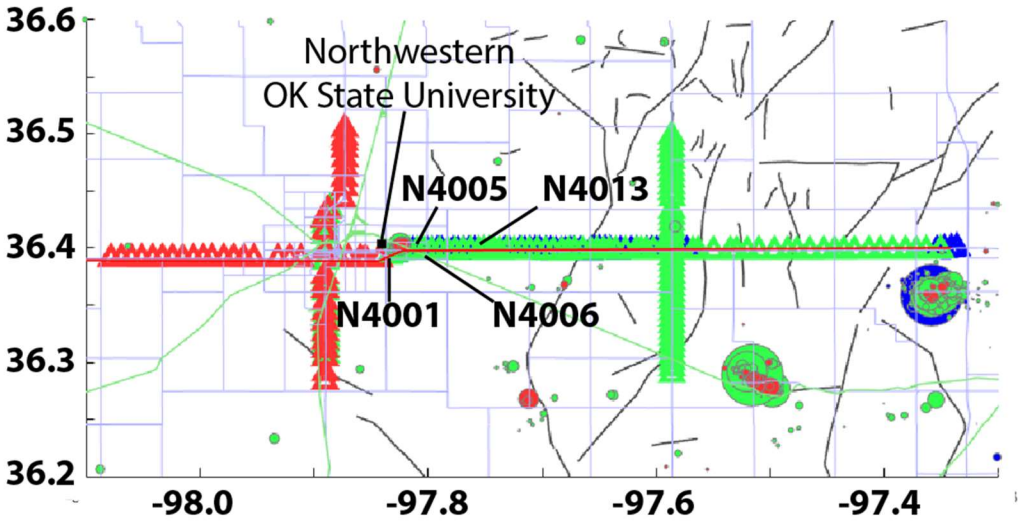

Figure 3.2 Experiment location from the red box in **Figure 1**. Local events are in dots where they are color coded for each array and magnitude would depict circle size. Nodal stations for each segment of the deployment are shown by the colored triangles where blue is EW1-40, green is EW1-20, and red is EW2.

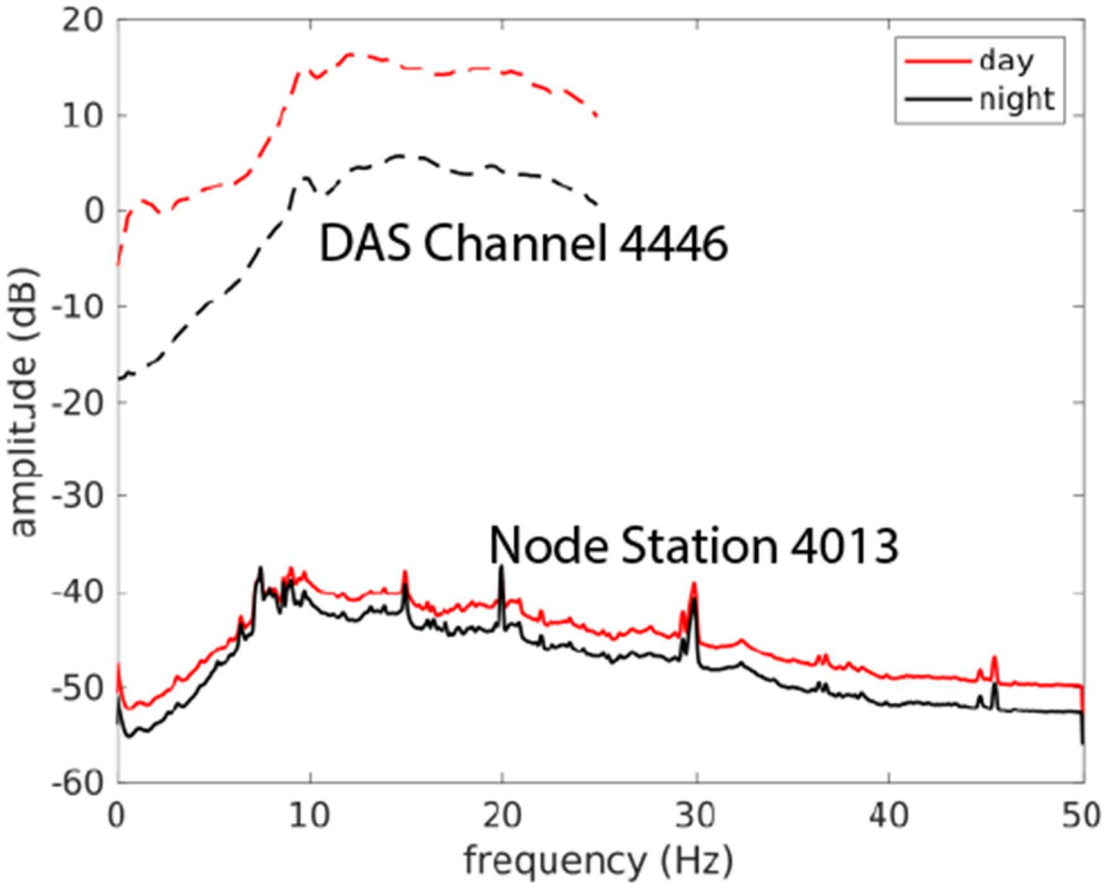

Figure 3.3 Day and nighttime noise for node station 4013 and corresponding DAS channel up to the Nyquist frequency of the down-sampled DAS and node time series.

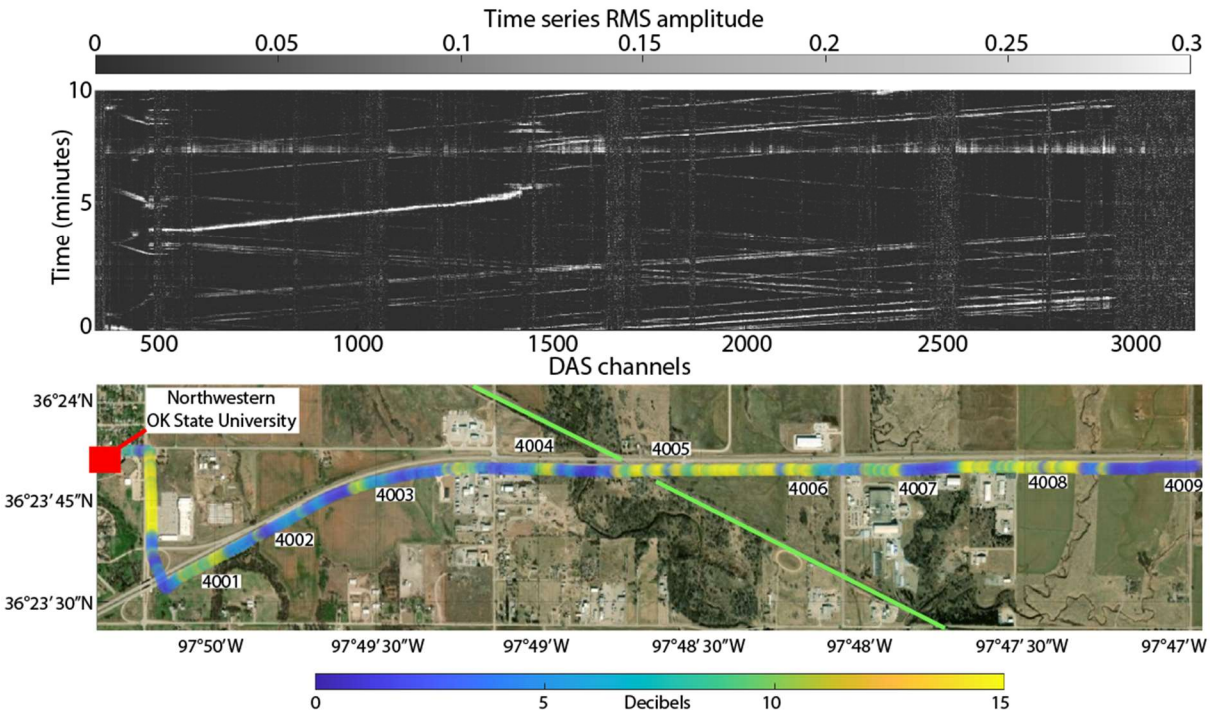

Figure 3.4 Deployment duration average noise spectrum of the DAS array for the first 3000 channels and an example 10-minute RMS time series capturing the 2.3 local magnitude earthquake on May 2, 2021.

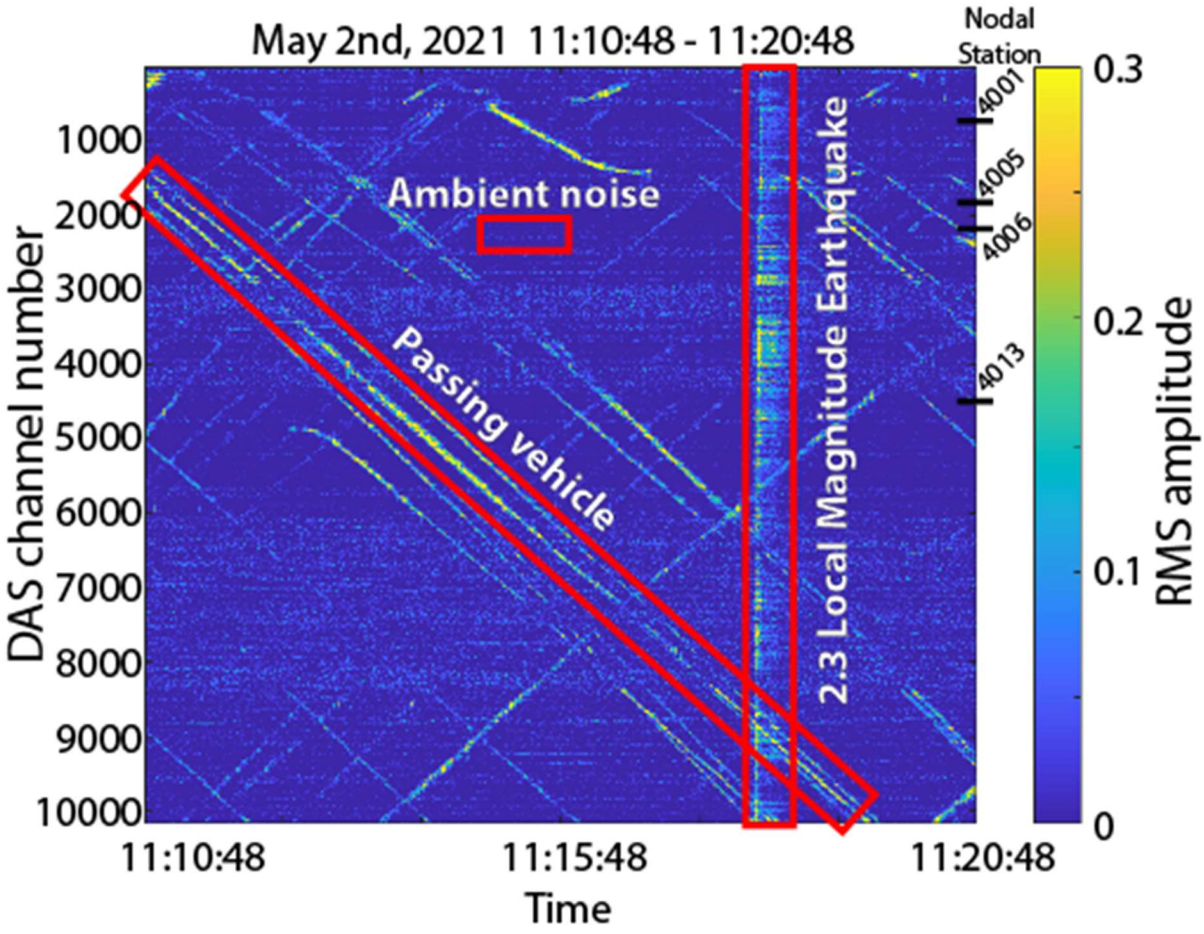

Figure 3.5 RMS of DAS data. Passing vehicles, ambient noise, and a local event are highlighted.

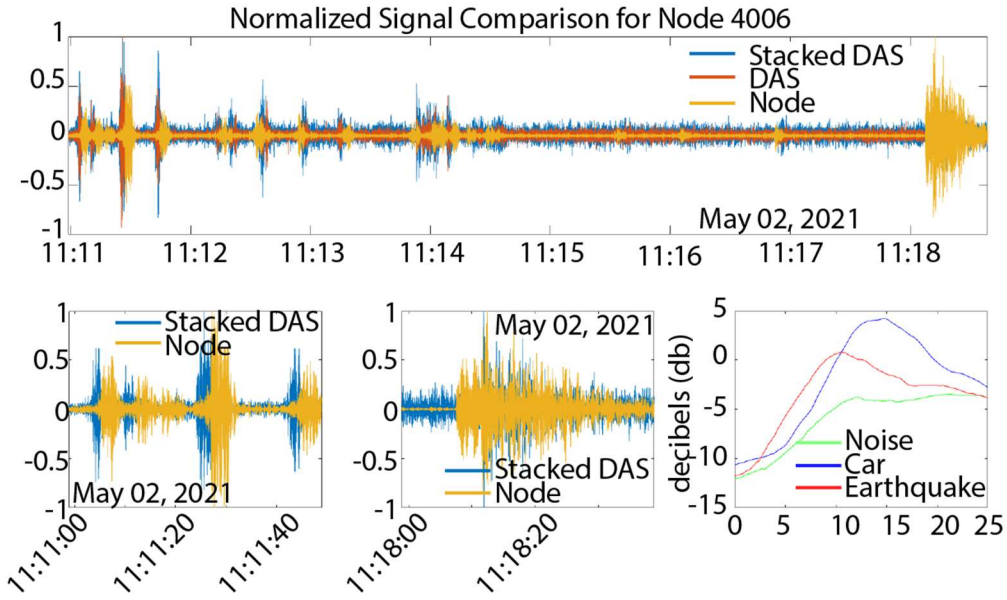

Figure 3.6 A waveform comparison of DAS and nodal station 4006. The top is stack DAS, single DAS, and Node. The bottom left is a car signal, the bottom middle is an earthquake, and the bottom right is averaged and smoothed signal spectrum.

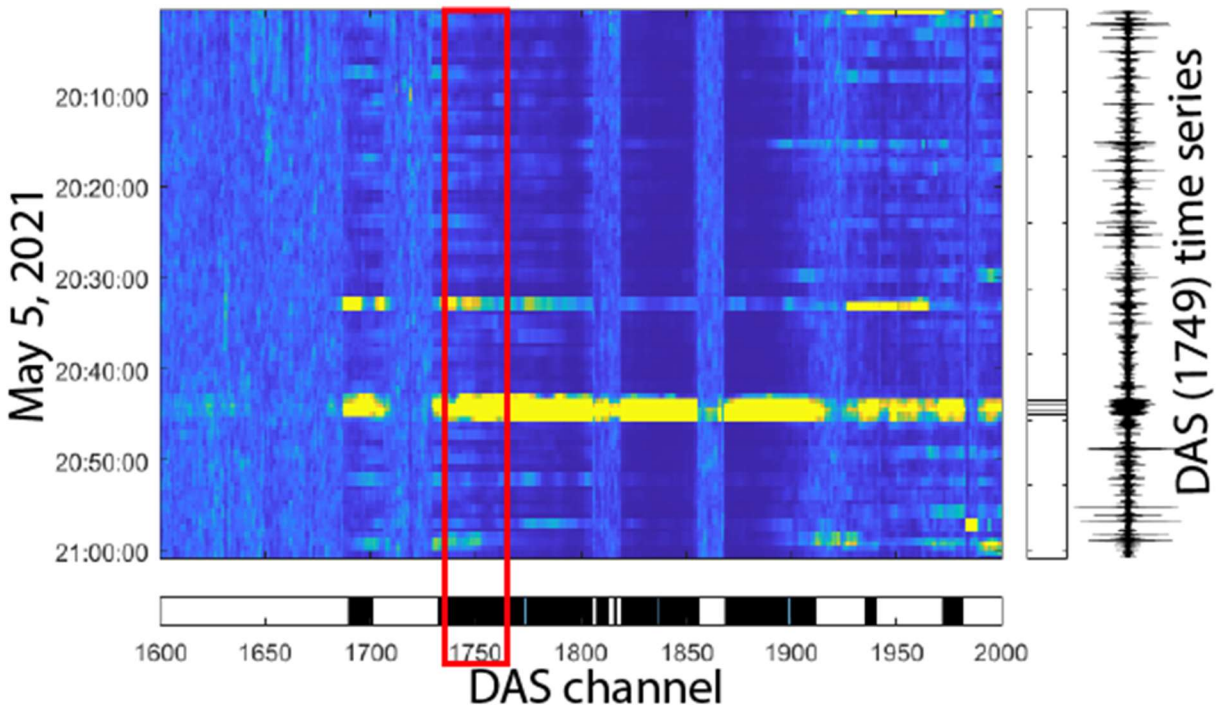

Figure 3.7 A train detection sparse array highlighting sections of both consecutive triggered DAS channels and consecutive time windows with large amplitude signals. An example time series is shown to illustrate a single-channel train signal.

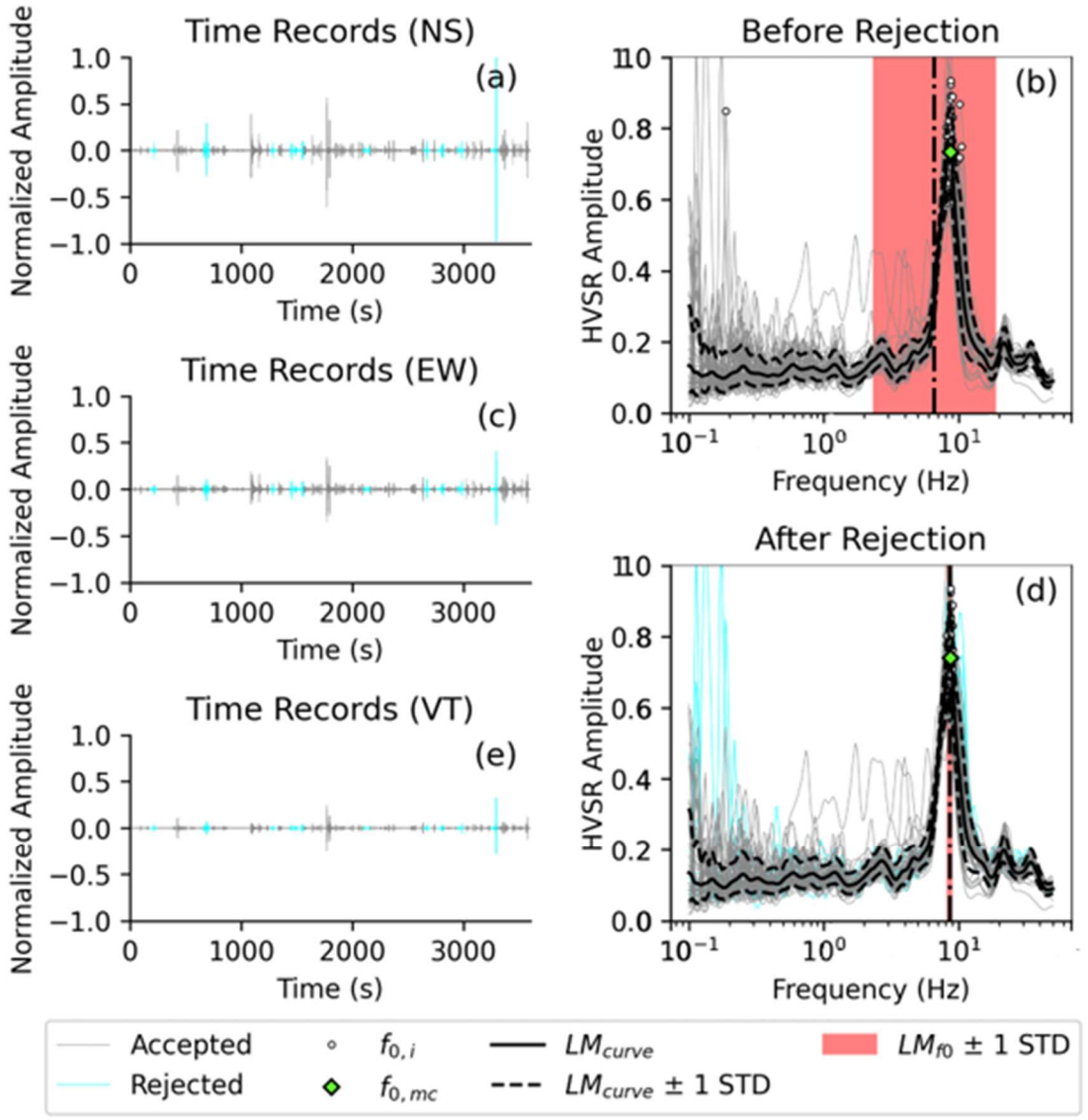

Figure 3.8 Example HVSRpy output for station 4006 where time series components and resultant HVSR curve are displayed. Spectral rejection (blue line) is shown here with improvement to standard deviation.

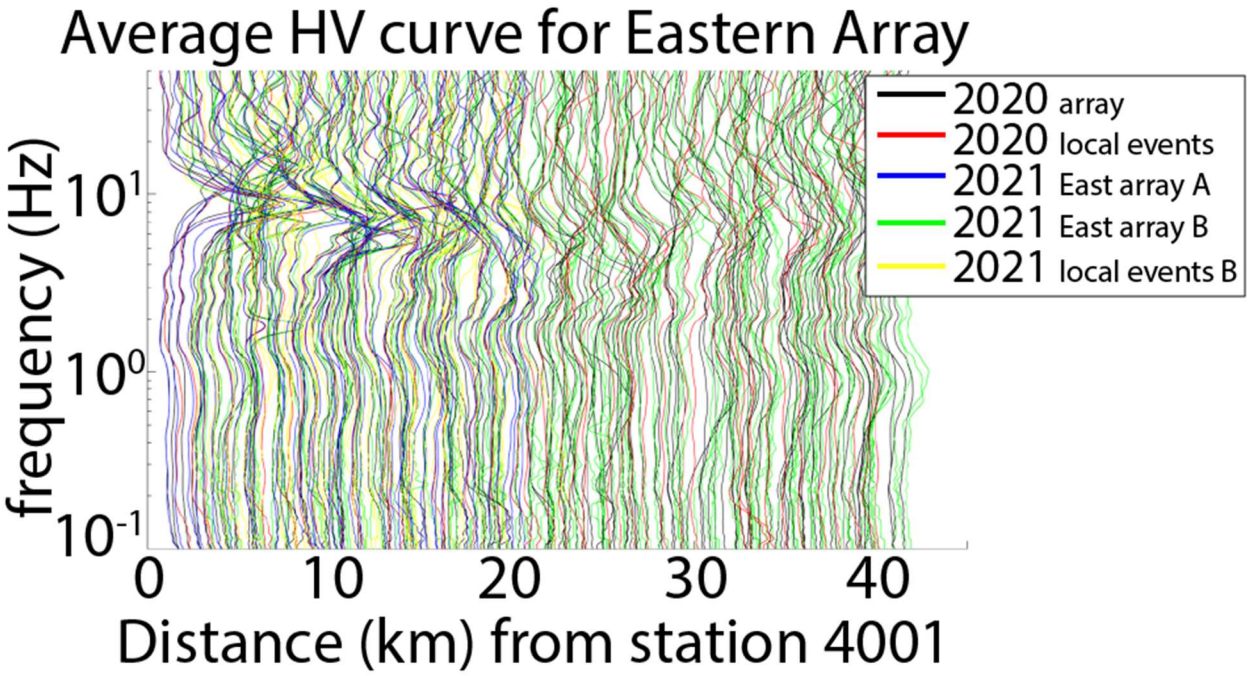

Figure 3.9 Compilation of HVSR curves for all deployments between 2020 and 2021 as defined by color.

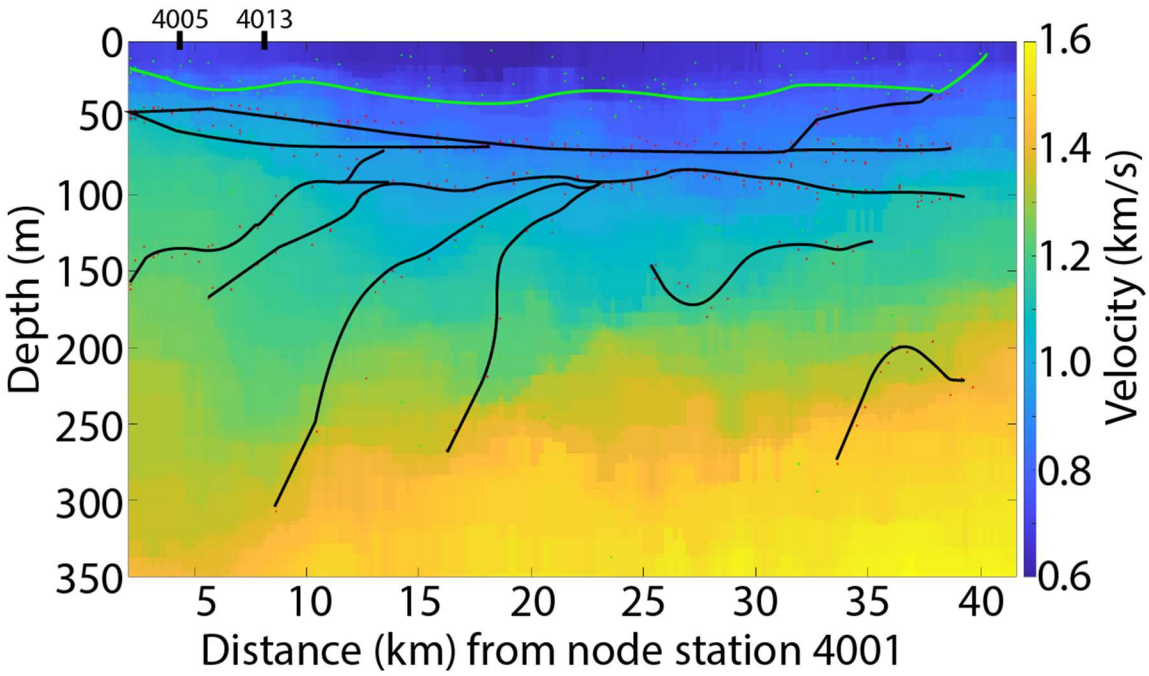

Figure 3.10 Interpreted geologic structure layers from forward modeling of the HVSR curves using an estimated velocity model to determine depth. The estimated fundamental frequency interpreted depth is shown as a green line and secondary frequency peaks are used to generate deeper interpreted structures (black lines).

References:

Abbott, J. (2005). Understanding and managing the unknown: The nature of uncertainty in planning. Journal of Planning Education and Research, 24(3), 237-251.

Albarello, D., Cesi, C., Eulilli, V., Guerrini, F., Lunedei, E., Paolucci, E., ... & Puzzilli, L. M. (2011). The contribution of the ambient vibration prospecting in seismic microzoning: an example from the area damaged by the April 6, 2009 L'Aquila (Italy) earthquake. Boll. Geof. Teor. Appl, 52, 513-538.

Amsden, T. W. (1980). Hunton Group (Late Ordovician, Silurian and Early Devonian) in the Arkoma Basin of Oklahoma.

Ansal, A., Tönük, G., & Kurtulucs, A. (2009). Microzonation for urban planning. Earthquakes and Tsunamis: Civil Engineering Disaster Mitigation Activities Implementing Millennium Development Goals, 133-152.

Bard, P.-Y., & SESAME Participants. (2004). The SESAME project: an overview and main results. proceedings of the 13th world conference on earthquake engineering, Vancouver. earthquake of September 19, 1985. Geophysical Research Letters, 13(6), 593-596.

Bard, P.-Y., Jongmans, D., Ohrnberger, M., & Wathelet, M. (2005). Site effects assessment using ambient excitations (SESAME). Europeam Commission - Research General Directorate, Final Repo(Project No. EVG1-CT-2000-00026 SESAME).

Beck, J. L., & Hall, J. F. (1986). Factors contributing to the catastrophe in Mexico City during the

Behm, M., Cheng, F., Patterson, A., & Soreghan, G. S. (2019). Passive processing of active nodal seismic data: estimation of VP/VS ratios to characterize structure and hydrology of an<? xmltex\break?> alpine valley infill. Solid Earth, 10(4), 1337-1354.

Benioff, H. (1949). Seismic evidence for the fault origin of oceanic deeps. Geological Society of America Bulletin, 60(12), 1837-1856.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., ... & Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. Geophysical journal international, 169(3), 1239-1260.

Berendsen, P., & Blair, K. P. (1986). Subsurface structural maps over the Central North American rift system (CNARS), central Kansas, with discussion (Vol. 8). Kansas Geological Survey.

Beresnev, I. A., Atkinson, G. M., Johnson, P. A., & Field, E. H. (1998). Stochastic finite-fault modeling of ground motions from the 1994 Northridge, California, earthquake. II. Widespread nonlinear response at soil sites. Bulletin of the Seismological Society of America, 88(6), 1402-1410.

Boyd, R. W. (2020). Nonlinear optics. Academic press.

Busanus, J. W. (1987). The Misener Strike-Valley Sandstone Reservoir, Grant and Garfield Counties, Oklahoma.

Cara, F., Di Giulio, G., Milana, G., Bordoni, P., Haines, J., & Rovelli, A. (2010). On the stability and reproducibility of the horizontal-to-vertical spectral ratios on ambient noise: the case study of Cavola, northern Italy. Bulletin of the Seismological Society of America, 100(3), 1263—-1275.

Cary, L. W. (1955). The Subsurface Geology of the Garber Area, Garfield County, Oklahoma.

Caylor, J. W. (1958). Subsurface geology of western Garfield County, Oklahoma.

Chatelain, J. L., Guillier, B., Cara, F., Duval, A. M., Atakan, K., Bard, P. Y., & Wp02 Sesame Team. (2008). Evaluation of the influence of experimental conditions on H/V results from ambient noise recordings. Bulletin of Earthquake Engineering, 6, 33-74.

Cheng, F., Xia, J., Behm, M., Hu, Y., & Pang, J. (2019). Automated data selection in the Tau–p domain: Application to passive surface wave imaging. Surveys in Geophysics, 40, 1211-1228.

Cheng, T., Hallal, M. M., Vantassel, J. P., & Cox, B. R. (2021). Estimating unbiased statistics for fundamental site frequency using spatially distributed HVSR measurements and Voronoi

tessellation. Journal of Geotechnical and Geoenvironmental Engineering, 147(8), 04021068.

Conway, C., & Mondanos, M. (2015, June). An introduction to fibre optic Intelligent Distributed Acoustic Sensing (iDAS) technology for power industry applications. In 9th International Conference on Insulated Power Cables, Jicable15—Versailles Jun (pp. 21-25).

Dolton, G. L., & Finn, T. M. (1989). Petroleum geology of the Nemaha uplift, central midcontinent. US Geological Survey Open-File Report, 88, 1-18.

Dougherty, S. L., Cochran, E. S., & Harrington, R. M. (2019). The Large-N seismic survey in Oklahoma (LASSO) experiment. Seismological Research Letters, 90(5), 2051-2057.

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. Physics of the earth and planetary interiors, 25(4), 297-356.

Ellsworth, W. L. (2013). Injection-induced earthquakes. Science, 341(6142), 1225942.

Fairfield Nodal (2017). ZLand 3C Node Specifications, available at http:// fairfieldnodal.com/equipment/zland/zland specs (last accessed November 2017).

Farahani, M. A., & Gogolla, T. (1999). Spontaneous Raman scattering in optical fibers with modulated probe light for distributed temperature Raman remote sensing. Journal of Lightwave Technology, 17(8), 1379.

Fath, A. E. (1920). The Origin of the Faults, Anticlines, and Buried" granite Ridge" of the Northern Part of the Mid-continent Oil and Gas Field.. (Vol. 128). University of Chicago.

Flores-Estrella, H., Yussim, S., & Lomnitz, C. (2007). Seismic response of the Mexico City Basin: A review of twenty years of research. Natural Hazards, 40(2), 357.

Frankel, A. D., Carver, D. L., & Williams, R. A. (2002). Nonlinear and linear site response and basin effects in Seattle for the M 6.8 Nisqually, Washington, earthquake. Bulletin of the Seismological Society of America, 92(6), 2090-2109.

Guillier, B., Chatelain, J.-L., Bonnefoy-Claudet, S., & Haghshenas, E. (2007). Use of Ambient Noise : From Spectral Amplitude Variability to H / V Stability. Journal of Earthquake Engineering, 11(6), 925–942. doi: 10.1080/13632460701457249

Hannemann, K., Papazachos, C., Ohrnberger, M., Savvaidis, A., Anthymidis, M., and Lontsi, A. M.: Three-dimensional shallow structure from high-frequency ambient noise tomography: New results for the Mygdonia basin-Euroseistest area, northern Greece, J. Geophys. Res, 119, 4979–4999, 2014.

Hartzell, S., Liu, P., & Mendoza, C. (1996). The 1994 Northridge, California, earthquake: Investigation of rupture velocity, risetime, and high-frequency radiation. Journal of Geophysical Research: Solid Earth, 101(B9), 20091-20108.

Hruby, C. E., & Beresnev, I. A. (2003). Empirical corrections for basin effects in stochastic ground-motion prediction, based on the Los Angeles basin analysis. Bulletin of the Seismological Society of America, 93(4), 1679-1690.

Imamura, A. (1913). Seismic intensity distribution in Tokyo and Osaka. Report of the Imperial Earthquake Investigation Committee, (77), 17-42.

Kagami, H., Okada, S., Shiono, K., Oner, M., Dravinski, M., & Mal, A. K. (1986). Observation of 1- to 5-second microtremors and their application to earthquake engineering. Part III. A two-dimensional study of site effects in the San Fernando Valley. Bulletin of the Seismological Society of America, 76(6), 1801-1812.

Kawase, H., & Aki, K. (1990). Topography effect at the critical SV-wave incidence: possible explanation of damage pattern by the Whittier Narrows, California, earthquake of 1 October 1987. Bulletin of the Seismological Society of America, 80(1), 1-22.

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., & Ge, S. (2014). Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. Science, 345(6195), 448-451.

Larose, E., Carrière, S., Voisin, C., Bottelin, P., Baillet, L., Guéguen, P., ... & Massey, C. (2015). Environmental seismology: What can we learn on earth surface processes with ambient noise?. Journal of Applied Geophysics, 116, 62-74.

Levshin, A.L., Yanovskaya, T.B., Lander, A.V., Bukchin, B.G., Barmin, M.P., Ratnikova, L.I. & Its, E.N., 1989. Seismic Surface Waves in a Laterally Inhomogeneous Earth, ed. Keilis-Borok, V.I., Kluwer, Norwell, Mass.

Li, Z., Peng, Z., Hollis, D., Zhu, L., & McClellan, J. (2018). High-resolution seismic event detection using local similarity for Large-N arrays. Scientific reports, 8(1), 1646.

Marra, G., Clivati, C., Luckett, R., Tampellini, A., Kronjäger, J., Wright, L., ... & Calonico, D. (2018). Ultrastable laser interferometry for earthquake detection with terrestrial and submarine cables. Science, 361(6401), 486-490.

Marsh, S., and A. Holland (2016), Comprehensive Fault Database and Interpretive Fault Map of Oklahoma, OGS Open File Report 2-2016, 16 pages, 2 plates.

Muanenda, Y., Oton, C. J., & Di Pasquale, F. (2019). Application of Raman and Brillouin scattering phenomena in distributed optical fiber sensing. Frontiers in Physics, 7, 155.

Murphy, P. (1996, April). Geophone design evolution related to non-geophysical applications. In AIP Conference Proceedings (Vol. 368, No. 1, pp. 49-56). American Institute of Physics.

Nakamura, Y. (1989). A method for dynamic characteristics estimation of subsurface using microtremor on the ground surface. Railway Technical Research Institute, Quarterly Reports, 30(1).

Park, C. B., Miller, R. D., Xia, J., Hunter, J. A., & Harris, J. B. (1999). Higher mode observation by the MASW method. In SEG Technical Program Expanded Abstracts 1999 (pp. 524-527). Society of Exploration Geophysicists.

Parolai, S., Richwalski, S. M., Milkereit, C., & Bormann, P. (2004). Assessment of the stability of H/V spectral ratios from ambient noise and comparison with earthquake data in the

Cologne area (Germany). Tectonophysics, 390(1), 57–73. doi: 10.1016/j.tecto.2004.03.024

Qin, Y., Chen, X., Walter, J. I., Haffener, J., Trugman, D. T., Carpenter, B. M., ... & Kolawole, F. (2019). Deciphering the stress state of seismogenic faults in Oklahoma and southern Kansas based on an improved stress map. Journal of Geophysical Research: Solid Earth, 124(12), 12920-12934.

Raman, C. V. (1928). A new radiation. Indian Journal of physics, 2, 387-398.

Schoenball, M., & Ellsworth, W. L. (2017). Waveform-relocated earthquake catalog for Oklahoma and southern Kansas illuminates the regional fault network. Seismological Research Letters, 88(5), 1252-1258.

Schmertmann, J. H. (1978). Guidelines for cone penetration test: performance and design (No. FHWA-TS-78-209). United States. Federal Highway Administration.

Sweet, J. R., Anderson, K. R., Frassetto, A., Beaudoin, B. C., Bilek, S. L., & Woodward, R. S. (2018, December). Evolution of the IRIS Portable Facility: New tools for Wavefield Imaging, Rapid Response, and Magnetotellurics. In AGU Fall Meeting Abstracts (Vol. 2018, pp. S43E-0677).

Vantassel, J. P., Cox, B. R., & Brannon, D. M. (2021). HVSRweb: An Open-Source, Web-Based Application for Horizontal-to-Vertical Spectral Ratio Processing. In IFCEE 2021 (pp. 42-52).

von Rebeur-Paschwitz, E. (1889). The earthquake of Tokio, April 18, 1889. Nature, 40(1030), 294-295.\

Wadati, K. (1928). Shallow and deep earthquakes. Geophys. Mag., 1, 162-202.

Walter, J. I., P. Ogwari, A. Thiel, F. Ferrer, I. Woelfel, J. C. Chang, A. P. Darold, and A. A. Holland (2020), The Oklahoma Geological Survey Statewide Seismic Network, Seismol. Res. Lett., 91 (2A): 611–621, doi:10.1785/0220190211.

Wathelet, M., Chatelain, J. L., Cornou, C., Giulio, G. D., Guillier, B., Ohrnberger, M., & Savvaidis, A. (2020). Geopsy: A user-friendly open-source tool set for ambient vibration processing. Seismological Research Letters, 91(3), 1878-1889.

Whitmeyer, S. J., & Karlstrom, K. E. (2007). Tectonic model for the Proterozoic growth of North America. Geosphere, 3(4), 220-259.

Wills, C., Petersen, M., Bryant, W., Reichle, M., Saucedo, G., Tan, S., . . . Treiman, J. (2000). A site-conditions map for California based on geology and shear-wave velocity. Bulletin of the Seismological Society of America, 90(6B), S187–S208.

Xia, J., Miller, R.D., Park, C.B., and Hunter, J., 1998, Comparison of shear wave velocities from MASW technique and borehole measurements in unconsolidated sediments of the Fraser River Delta: Kansas Geological Survey, Open-file Report No. 98-58.

Chapter 4:

Wind and wind turbine influence on measurements of seismic noise

4.1 Abstract

Renewable wind energy is a growing global industry that has led to the construction of wind farms across the United States, especially in midcontinent regions such as Oklahoma. Seismic noise generated adjacent to wind turbines includes both natural wind source coupling and resonance of the turbine blades and towers that are transmitted through the subsurface. We conducted a pilot study to characterize and classify such noise that could lead to the development of better techniques in ambient-noise seismology, noise suppression, identification of near-surface resonance, and improvement to signal-to-noise ratio. We investigate the wind and wind turbine-generated noise within the seismic field through the application of the power density function on 3-component waveforms collected from two temporary arrays of 5 Hz geophone sensors. The temporary array consists of 8 Fairfield nodes that were active for one month with varying distances (10 meters – 2000 meters) from wind turbine towers located in Grant County, Oklahoma. The second array is located approximately 50km northwest of those wind turbines. The spectral amplitudes and peak frequencies of the power spectrum show unique spatial-temporal variations in noise levels for the location of the wind turbine towers. Noise amplitude decreases exponentially with distance from the wind turbine, and wind speed correlates with power spectrum peak frequencies. We identify mechanical signal sources at 0.17 Hz, 0.29 Hz, 0.51 Hz, 0.87 Hz, 14.5 Hz, 24.83 Hz, and 27.82 Hz, and the resonant frequency of the wind turbine tower at 0.37 Hz. We compare the estimated fundamental frequencies of both arrays. The directivity of the seismic noise is characterized using a cross-correlation function. We observe the wind turbines to be a very clear source of

seismic noise with exponential power degrading at increasing distances. Closer examination shows wind, ground, and wind turbine coupling to be a complex field.

4.2 Introduction

The seismic wavefield has long been recorded and analyzed by researchers and professionals. Signals from both natural and anthropogenic sources such as earthquakes or hydrocarbon exploration propagate through the subsurface and are recorded by receivers. However, it is quite often that these signals are contaminated and distorted by acoustic seismic noise which can obscure important signals. Several noise types may exist, including, wind, ocean waves, and anthropogenic noise. To improve on the advancement of signal analysis it is important to delineate noise sources such as wind and wind turbines and characterize their influence on the acoustic seismic field.

Wind energy has been a global growing industry in recent years increasing its global power capacity exponentially from 2000-present. Growth in each wind turbine generation has also led to greater power outputs with an increase in rotor size (EWEA, 2012). Meo (2006) summarizes, in-depth, Oklahoma's expansion into wind energy due to partnerships, policies, infrastructure, economics, environmental, and community impact. These factors contribute to the overall growth of the wind energy industry in the state (Righter, 1996). Man-made structures, machinery, human activity, and natural surface forces such as wind and temperature may generate noise that is coupled and transmitted to the earth and subsequently recorded by seismographs, that detect ground motion small enough to be imperceptible to humans (Wilson, 1953 and Withers et al., 1996). As an example, structures such as wind turbines house complex vibration-inducing drive shafts and gears within the nacelle (Hemami, 2012) which may transmit vibrations into the subsurface. Less well known is the extent of the magnitude of source and

coupling and how it varies with distance as the energy propagates through the surface since source signals are unique to each source and are modified by the path and site effects.

Seismic noise generated by wind turbines poses a scenario where both natural wind source coupling and resonance frequencies of the turbine blades and towers are transmitted through the subsurface. Resonance from harmonic waves for much larger sources can amplify earthquake waves and cause structural damage, such as was observed during the 1987 Whittier earthquake (Vidale et al., 1991), where basin resonance caused significant localized damage to buildings. Resonance is primarily from earthquake shaking in the fundamental frequency but can occur from overtones (Rial et al., 1992). Previous studies in wind variations have shown a general increase in seismic noise with the wind with no apparent fundamental frequency at which the wind was observed (Muccaiarelli et al., 2005). However, other studies have suggested wind energy to be observed at higher frequency bands (Wilson, 1953; Withers et al., 1996; Young et al., 1996; and Teanby et al., 2016), though the measured source frequency bands vary between these studies. To improve our understanding of seismic signal sources such as earthquake or wind turbine fundamental frequencies, we must understand the characteristics of the noise as an essential first step before a desired signal can be isolated.

4.3 Data acquisition and processing

For this study, we select a high-noise environment of wind turbines to observe wind turbine-influenced signals. We compare our findings with measurements from a location of lesser noise and without wind turbines to determine the effects wind turbines have on the local acoustic seismic wavefield. Seismic data is acquired from a small temporary geophone (Fairfield

Zland 3T) array of 8 three-component sensors for a duration of 1 month from December 2017 – January 2018 (**Figure 4.1**) in Grant County, Oklahoma. A second temporary array of 65 additional Fairfield Zland 3T sensors were also deployed between November 2017 – December 2017 with 30 overlapping days in Alfalfa County, approximately 50 km northwest of the Grant County array. The nodes were recorded at the frequency of 500Hz and later down-sampled to 100Hz. An example of the continuous waveform recordings from both arrays is shown in **Figure 4.2** showing a transient low-magnitude earthquake and background noise. It shows variations in waveform amplitude between the two locations. At both array locations, the nodes are buried one inch below the ground surface in a relatively homogenous surface soil. Surface soil conditions at Grant County are varying degrees of silt loam whereas Alfalfa County's array is predominately on fine sandy loam north of Great Salt Plains Lake (Soil Survey Staff, 2018), with a shallow ground water table of approximately 3 m (USGS, 2016) for both locations. Ground-level weather conditions such as wind speed, direction, and temperature are measured by the Oklahoma Mesonet weather network and are publicly available (Brock et al., 1995; McPherson et al., 2007), with a sampling rate of 300 seconds. The nearest Mesonet station, Medford (MEDF), is 25 km to the northwest of Grant County. We compare the recordings to a concurrent array located 50 km to the northwest in Alfalfa County where wind turbines are absent. The wind turbines in Grant County are classified as monopole-horizontal-axis wind turbines used in electricity production by transforming slow turbine blade rotation to a high-speed rotation for a generator axle through gear ratios, similar to the transmission in an automobile. The monopole wind turbines in Grant County have a blade length of 41.25 m, a total height of 121.25 m, and an average rotation of 12 to 14 revolutions per minute. The time series measurements as

exampled in **Figure 4.2** from the nodal arrays, $f(t)$, are transformed using the fast Fourier transform, $F(\omega)$, in **equation 4.1** such as in **Figure 4.3,** and later compared with local weather conditions in **Figure 4.4**.

$$F(\omega) = \frac{1}{2\pi}\int_0^T f(t)e^{i\omega t}dt$$

(4.1)

Where T is the length of the time series, we use 300 seconds for each component, and ω is frequency. A noise-corrected power spectrum (Cooley and Turkey, 1965), $P(\omega)$, is computed by calculating the base-10 logarithm of the moving mean and mean spectrum ratio in **equation 4.2.**

$$P(\omega) = 10\log(\frac{(\frac{1}{k}\sum_{i=n-k+1}^{n} F(\omega)_i) + \epsilon}{\frac{F(\omega)}{n}})$$

(4.2)

The power spectrum is used to evaluate a broad range of noise at a given station and enhances the relative variations in seismic power, allowing the signal to be better-distinguished relative to background noise levels. We apply a moving average window, k=5, to a frequency spectrum array with respect to time, n, and divide by the average frequency spectrum for each

time series window, T. Adding epsilon, ϵ, to the numerator ensures that the division does not result in an undefined or infinite value, providing a small positive value that avoids numerical errors. We use spectra to observe differences between different conditions throughout the deployment duration such as day and nighttime hours and low and high wind conditions (**Figure 4.5**). Day hours, loosely defined as 8 am – 5 pm, are considered for their propensity for elevated noise levels when compared with night hours, 10 pm – 4 am. We define wind speeds up to 3 m/s as low-speed wind and above 5 m/s as high-speed wind since wind turbines are designed to turn on close to the low wind speed cut-off speed. Signal spectra are subsequently used to estimate site conditions and noise source location through methods such as HVSR for single-station site response and cross-coherence for tracking signal propagation.

4.3.1 Horizontal to Vertical Spectral Ratio

Horizontal to Vertical Spectral Ratio (HVSR) (Nakamura 1989) is often used in microzonation studies to quantify site effects in terms of resonance frequency, geologic surface conditions, and site amplification factor. Generally, it is recommended to use transient signal-free and calm environment recordings for ambient noise HVSR to analyze the background noise wavefield for interpretation of the subsurface structure through fundamental resonant frequencies. The HVSR technique is a commonly used method in geophysics and earthquake engineering to estimate the resonant frequency from the ratio of the spectral ratios of horizontal and vertical components, $H/V(f)$, of ground motion recorded at a site as expressed in **equation 4.3**. Where the root geometric mean square ratio of the two horizontal spectra, H, is taken over twice the vertical spectra. The resonant frequency is an important parameter for

seismic hazard assessment, as it determines the level of amplification of ground motion that occurs at a site during an earthquake.

The method can be applied to different types of data, such as seismic, acoustic, or ambient vibration data, and can be used to study a wide range of phenomena, including the effects of geologic and structural features on ground motion, the characterization of soil and rock properties, and the evaluation of the vulnerability of structures to earthquake damage.

We determine the HVSR curve for each station and analyze the results in accordance with the Site Effects Assessment Using Ambient Excitations (SESAME) guidelines Chatelain et al. (2008). These guidelines, which are empirically derived, provide an estimate of site amplification and meaning to the HVSR. Albarello et al. (2011) implemented the HVSR approach to have three conditions for curve reliability and five criteria for the identification of a peak fundamental frequency, f_0, as a clear peak. We use 30-second-long windows for our HVSR calculations with the open-source software HVSRpy (Vantassel, 2020).

$$H/V(f) = \sqrt{\frac{{H_T}^2 * {H_R}^2}{2V^2}} \tag{4.3}$$

4.3.2 Cross Coherence

Cross coherence is a method in signal processing and vibration analysis to quantify the relationship between two signals. The method is based on computing the coherence function between two signals, which measures the degree of similarity between them in the frequency domain. It is calculated by taking the cross-power spectral density of the two signals and

dividing it by the product of their individual power spectral densities. The cross-coherence function can provide valuable information on the frequency-dependent relationship between two signals, including the presence of common frequency components and the phase difference between them. One of the key advantages of the cross-coherence method is its ability to identify and quantify the coupling between two signals. The cross-coherence method can help to identify the frequency ranges and the degree of coupling between two systems, which is useful for understanding the underlying physical processes and designing appropriate control strategies.

We use cross coherence between station pairs A and B to determine the direction of travel for our noise source in the frequency domain (w) as described in (Nakata, 2013) and shown in **equation 4.4** below. We reject any transient signals and use rotated component 40-second overlapping time series windows from each station for the cross-coherence calculation.

$$C(B,A,w) = \sum_{t} \frac{u_t(B,w)u_t^*(A,w)}{|u_t(B,w)||u_t(A,w)| + e < |u_t(A,w)||u_t(B,w)| >} \tag{4.4}$$

With ambient noise recording at station A at time t, with complex conjugate denoted with *, e regularization, <…> representing the ensemble. Nakata (2011) shows the importance of power normalization for obtaining propagating waves between two stations where amplitude variations may be too great using other cross-correlation or deconvolution methods. In cross coherence with power normalization, the amplitude information is preserved and thus suitable for noisy data with varying amplitudes among traces.

4.4 Results

The time series in **Figure 4.2** exhibits an example of the Grant County and Alfalfa County array's recorded passive ground motion amplitudes under different conditions. Recordings from Grant County display greater amplitudes than Alfalfa County during and off business hours (11 pm - 4 am local time) with high and low wind conditions. This is also true in the frequency domain where we observe relatively greater noise in the averaged stations in both arrays in **Figure 4.5**. We show station 1001 for Grant County and 3003 for Alfalfa County where noise levels are separated by windy/nonwind conditions, day/night hours, and overall averaged power spectrums. Daytime noise is greater than nighttime noise, with higher noise amplitudes during high wind periods when compared with low wind periods. We observe little separation of day and nighttime intervals as shown in **Figure 4.5** which would suggest no significant anthropogenic noise caused by humans during working hours. Power spectrum observations show an increase in wind speed correlates to an increase in the overall ambient noise amplitude. As a result, we apply our analysis to our complete data set. We exclude the latter half of data from station 1008 due to an unexplained amplitude variation that is not observed on other stations. We suspect an undetermined software logic malfunction associated with the node itself, perhaps during signal digitization, to be the cause of the data anomaly.

At near-field stations 1001 and 1002, we observe the North component average power spectrums with narrow peak low frequencies approximately at 0.37 Hz, 0.5 Hz, and 0.8 Hz shown in **Figure 4.3**. Overtones of these signals are observed as well at far-field stations starting at 0.87 Hz which are multiples of the narrow peak frequencies. Significant power spectrum peak amplitudes are observed on all stations at 2.607 Hz, 7.823 Hz, and 14.5 Hz. Secondary smaller

peak amplitudes are observed in the range of 2 – 35 Hz as shown in the example power spectrum in **Figure 4.2**. The frequency band 0 Hz – 10 Hz is highlighted to show both lower frequency dependency and independence on wind energy.

Seismic power compared with atmospheric conditions suggests a positive correlation between frequency and wind speed using data from Oklahoma Mesonet station MEDF in **Figure 4.4**. As highlighted before, wind speed, azimuthal direction, and temperature are recorded at 300 seconds sampling rate. Wind speed, azimuthal direction, and temperature are compared with lower frequency bands in the power spectrum 0.32 – 0.64 Hz, 0.64 – 1.28 Hz, 1.28 – 2.56 Hz, and 2.56 – 5.12 Hz. No correlation is shown between temperature and noise power log amplitude. Wind azimuthal direction and strength shows varying positive correlation with the power spectrum (**Figure 4.3**). Stable power log amplitudes are observed between 0.32Hz and 2.56Hz and power log amplitudes increase at a higher frequency band of 2.56 Hz and 5.12 Hz. Contrary to other studies (Schofield, 2001) we do not observe a significant shift in frequency between windy and non-windy periods. We do observe stronger and weaker signals between the two conditions. Comparing Grant County's waveform spectral to atmospheric patterns we see a correlation between windspeed and spectral amplitude with respect to time. We only see a correlation with windspeed while there appears to be a lack of correlation with other weather conditions (**Figure 4.4**). We average power spectrums based on a range of wind speeds and show the gradual increase in average noise spectral amplitudes with an increase in wind speed. Median noise amplitude between stations at Grant County positioned at varying distances relative to wind turbine towers, **Figure 4.6**, shows the amplitude decay with distance from a wind turbine source fitting with a squared polynomial decay.

We observe the spectral curves and peak frequencies by applying the HVSR method in HVSRpy, as shown in **Figure 4.7**. Fourier transform of the three component waveform data is used for **equation 4.3** to generate HVSR curves which are then corrected through spectral curve rejection, based on unbiased spatial statistics, Voronoi tessellation (Cheng et al., 2021). Common HVSR peak frequencies are approximately 7 Hz and 10.5 Hz. However, HVSR curves from stations near wind turbines appear to have a significant 10.5 Hz peak frequency when compared to other stations. The measurements show the HVSR method is successful in examining the ambient seismic background without being influenced by atmospheric conditions and local noise sources. As shown in **Figure 4.8**, at broad frequency bands, we observe energy propagation through the array location. west to east. We determine our array location's wavefield source and sink directions by measuring station pair time lag from ambient noise cross coherence in **equation 4.4**.

4.5 Discussion

Wind-dependent frequencies have the highest amplitude at stations 1001 and 1002 and decrease in relative amplitude with stations further away from the wind turbine as shown with the average power spectrum in **Figure 4.3**. As in Saccorotti et al. (2011) revealed frequency ranges between 1-5 Hz noise were increased in wind farm operations and wind speed, which is similar to our findings. Significant peaks seen in **Figure 4.5** that show wind dependency are greatest in the range of frequencies from 1 Hz – 35 Hz. Stations 1001 and 1002 have an increase in relative noise level due to their proximity to the wind turbine tower compared with other far-field stations, **Figure 4.6**. Significant frequency peaks at 2.7 Hz, 7.8 Hz, 14.5 Hz, 24.83 Hz, and 27.82 Hz are observed strongest near-field, next to a wind turbine, and observed on all stations.

Additional frequencies between 25 Hz and 30 Hz show a decrease in amplitude with an increase in station distance to the wind turbine tower and is attributed to the energy propagation of the wind turbine and wind coupling, **Figure 4.6**. We identify a low-frequency peak of 0.29 Hz in the power spectrum on near-field stations only and a strong 0.87 Hz frequency with higher order modes on all other stations with decreasing amplitude with distance. We attribute 0.29 Hz to the low-speed shaft axial rotation within the nacelle and the third harmonic 0.87 Hz to the wind turbine's three-bladed sweep signal at 12-14 rpm which translates to an approximately 0.2Hz blade sweep frequency. Based on the assumption of narrow sharp peaks in the power spectrum are attributed to mechanical wind turbine noise; we also consider 14.5 Hz, 24.83 Hz, and 27.82 Hz to be attributed to the multi-stage shaft revolution of the high-speed generator axle.

The HVSR results show no clear distinction between day and nighttime periods with some power offset between windy and non-windy periods in **Figure 4.9**. This suggests that the HVSR method is well-suited for the analysis of the seismic background of both windy and non-windy periods. The HVSR curves are analyzed in accordance with the SESAME guidelines for ambient excitations where the spectral ratio curve is tested for reliability and clear peaks (Bard et al., 2005). All selected peaks pass the reliable curve criterion. However, some of the selected peaks fail the test for clear peaks due to insufficient amplitude difference in defining a peak. We attribute this to an excessively noisy environment of near-field stations next to a wind turbine tower. It is suggested in HVSR studies to avoid ambient vibrations in areas of heavy machinery and/or high anthropogenic noise sources to avoid including transient ground motion (Koller et al., 2004). Therefore, we compare the HVSR curve results based on the recordings near wind turbines to the second array in Alfalfa County. We observe a clear and distinct peak frequency

common to all stations located in Alfalfa County in the frequency band 3 Hz – 5 Hz. Lesser amplitude and clear peaks are also visible at 10-12 Hz for some stations. Grant County stations show consistent frequencies at approximately 0.37 Hz, 0.8 Hz, 2.7 Hz, 5.1 Hz, 7.8 Hz, and 10.5 Hz.

For further investigation, we compare the low wind environment power amplitude spectrum with HVSR and notice a stronger 0.51 Hz peak in the power spectrum. Higher modes of 0.51 Hz are also significant in non-windy conditions similar to higher modes of 0.8 Hz in windy conditions. We attribute 0.51 Hz to the corresponding speed of the wind turbine blade sweep.

In the HVSR curve, we observe a broad frequency peak at 0.8 Hz. A simple approximation for potential building resonance (Taranath, 2016) is provided by **equation 4.5**

$$F_b = \frac{10}{N_{stories}} \tag{4.5}$$

Where F_b is the resonance frequency in hertz and $N_{stories}$ is the height of the structure in units of floors. We apply this approximation to estimate the 80-meter-tall wind turbine monopole tower resonance frequency to be 0.37 Hz. We observe the natural frequency of 0.37 Hz at all stations located near the wind turbine towers, except station 1008, which is independent of wind forces that we can consider as the wind turbine monopole tower resonance frequency. Back of the envelope calculation with the observed frequency, we estimate the tower height to

be 84 meters, which is an adequate approximation given the significant figures that are overlooked in the conversion of stories to meters.

From ambient noise correlation, we determine the propagation direction of seismic wave energy at specific frequency bands. Intuitively, it would be assumed that noise frequencies would show propagation away from stations positioned near a wind turbine tower which suggests significant seismic noise caused by a direct effect of the wind turbine tower and wind interaction with the monopole. However, analysis of the noise coherence shows a general West to East trend for wave energy at all frequency bands. Wave energy propagating from West to East suggests signals are independent of any immediate local effects of atmospheric conditions and structures. This analysis is not sensitive to the narrow band frequencies previously identified as wind turbine mechanical noise. One hypothesis is that the low frequencies are energy from long-period waves generated by severe winter storms along the Western coastline. However, continental stations in Bromirski et al. (2005) did not observe any mid-ocean double-frequency microseisms when observing mid-ocean storm swells. Low, broad peak frequencies could be attributed to a deep structure causing harmonic resonance since observations are consistent with the HVSR results. The West to East propagation trend could be attributed to a near-surface broad frequency source such as the other wind turbines to the south-southwest of the array as shown in **Figure 4.1**. Finally, the noise in this frequency band could have another underdetermined source relatively closer to the Alfalfa array.

4.6 Conclusion

To improve our understanding of the ambient seismic field, the power density function is applied to the dataset recorded by the temporary array of 3 component sensors to highlight fundamental frequencies. Based on the results obtained from the power spectrum and Mesonet dataset comparison we can conclude the broad amplitude increase in power spectrum are likely due to the wind coupling with the ground based on power spectrum amplitude comparisons between stations and weather conditions, sharp narrow frequency peaks are due to the wind turbine mechanical noise coupling directly with the ground, and the wind turbine monopole natural frequency is 0.3 Hz. We assume that (1) the wind turbine tower radiates a constant amount of energy at each spin level, (2) All nearby wind turbines are always at the same spin level, (3) a simple laterally homogenous subsurface, and (4) wind turbine energy sums constructively. We conclude 0.3 Hz is related to the low-speed shaft axial rotation within the nacelle of the wind turbine at two different speeds. The third harmonic of these two signals at 0.51 Hz and 0.87 Hz are related to the wind turbine blade beats at the different corresponding wind turbine operating speeds. High frequency and larger amplitude peaks are corresponding to the frequency band of the wind constructively summing with the preexisting frequencies related to the wind turbine.

Alfalfa County station HVSR results show both common peaks between all stations at 3 Hz – 5 Hz and varying spectral curve peaks over a short spatial distance. We averaged all station HVSR curves between each array to account for any variations in instrument and lateral heterogeneity in site effects. The average HVSR curves are both stable in different time windows as well as between all stations. Therefore, differences between Alfalfa County and Grant

County's HVSR curves can most likely be attributed to differences in site location where the obvious difference is the existence of wind turbines. It is perplexing as to how Alfalfa County exhibits a greater HVSR amplitude and significant peak whereas, the Grant County HVSR curves are relatively flat in comparison. HVSR interpretation is commonly used to infer or otherwise deduce a subsurface geologic layer of high impedance. Therefore, it may be likely that the significant peak observed in Alfalfa County could be caused by a very shallow subsurface layer, whereas the Grant County array exhibits a subsurface layer of lesser acoustic impedance. Certainly, the two locations exhibit drastically different subsurface features as highlighted by the HVSR results. An alternative hypothesis is that other noise sources are not otherwise characterized by this study which is focused on wind turbine seismic noise. Nonetheless, these observational differences further support the importance of microzonation studies.

The directivity of seismic energy between stations is determined based on the cross-correlation approach where time lag indicates the energy transfer direction. There is directivity across all frequency bands. However, it is counter-intuitive the source of the noise is not coming from stations next to wind turbines. It may be likely that the primary noise for the region is the collective wind turbines for the wind farm.

4.7 Acknowledgement

We would like to acknowledge the kind Oklahoma landowner who permitted us to access his field to deploy our sensors and fellow graduate students at the University of Oklahoma who participated in deployment and sensor retrieval. Without their help, this project

would not have been able to be performed. To OSCER, for their continued support and assistance in troubleshooting regarding supercomputing.

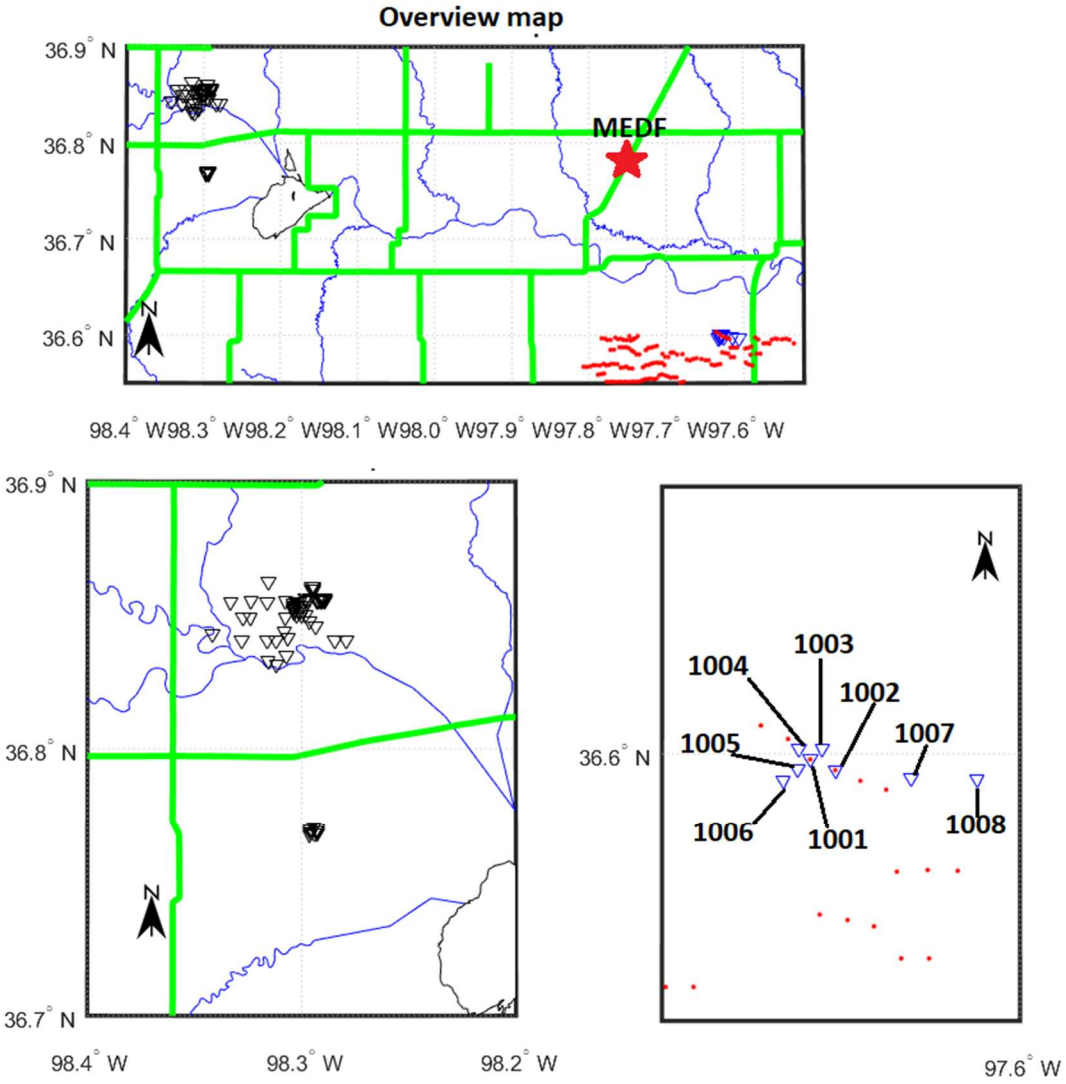

Figure 4.11: Local map in Grant County Oklahoma, Alfalfa County, and overview map showing the area of interest. geometry and individual stations are with triangles, neighboring wind

turbine locations are depicted with red dots, and roads are shown in green lines.

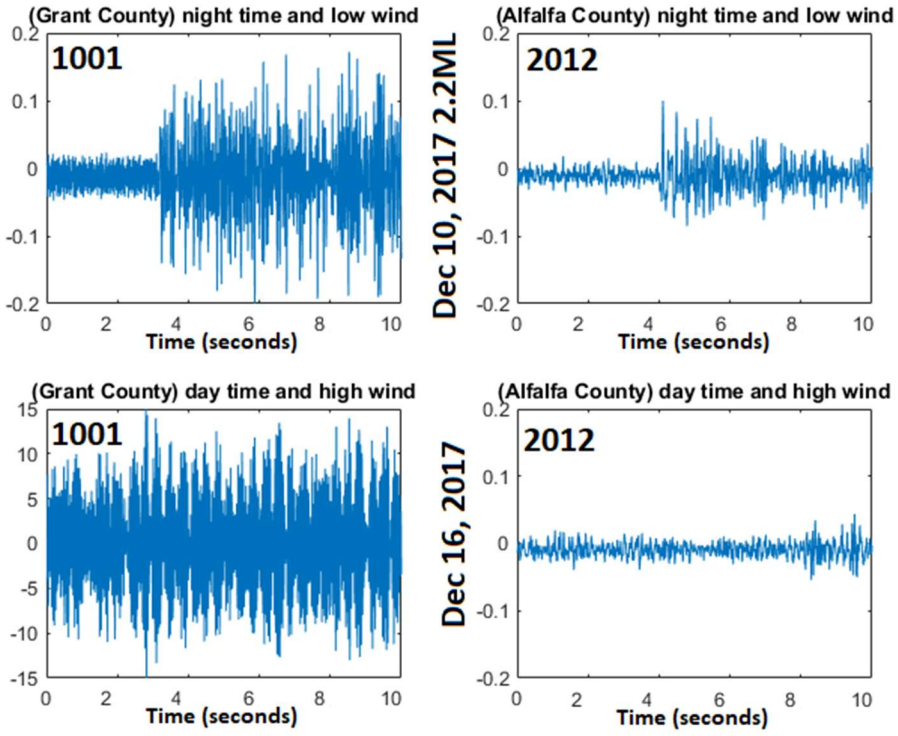

Figure 4.12: A 10-second time series window of the Z component for Stations 1001 from the Grant County array and Station 2012 from the Alfalfa County array. Two time periods are shown where Day time hours start at 7 pm and nighttime starts at 2 am CST. Note the different y-axis in the bottom left figure.

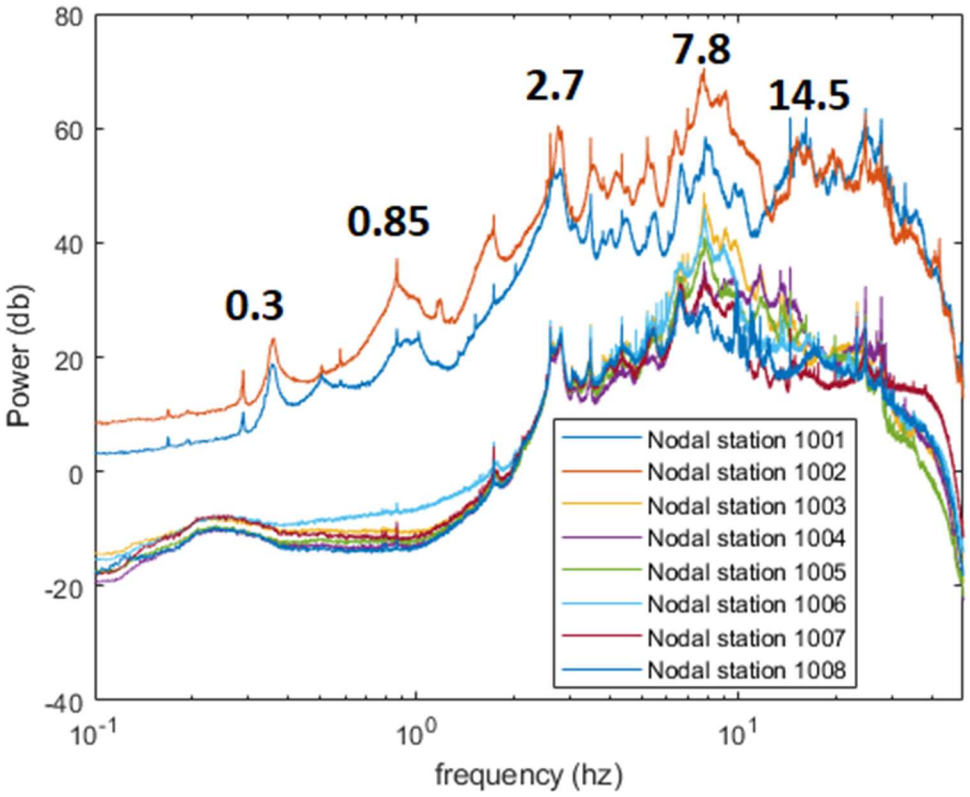

Figure 4.13: North component power spectrum of each nodal station in Grant County. Significant peak frequencies of interest are identified. Nodal station 1008 is furthest from the cluster of turbines.

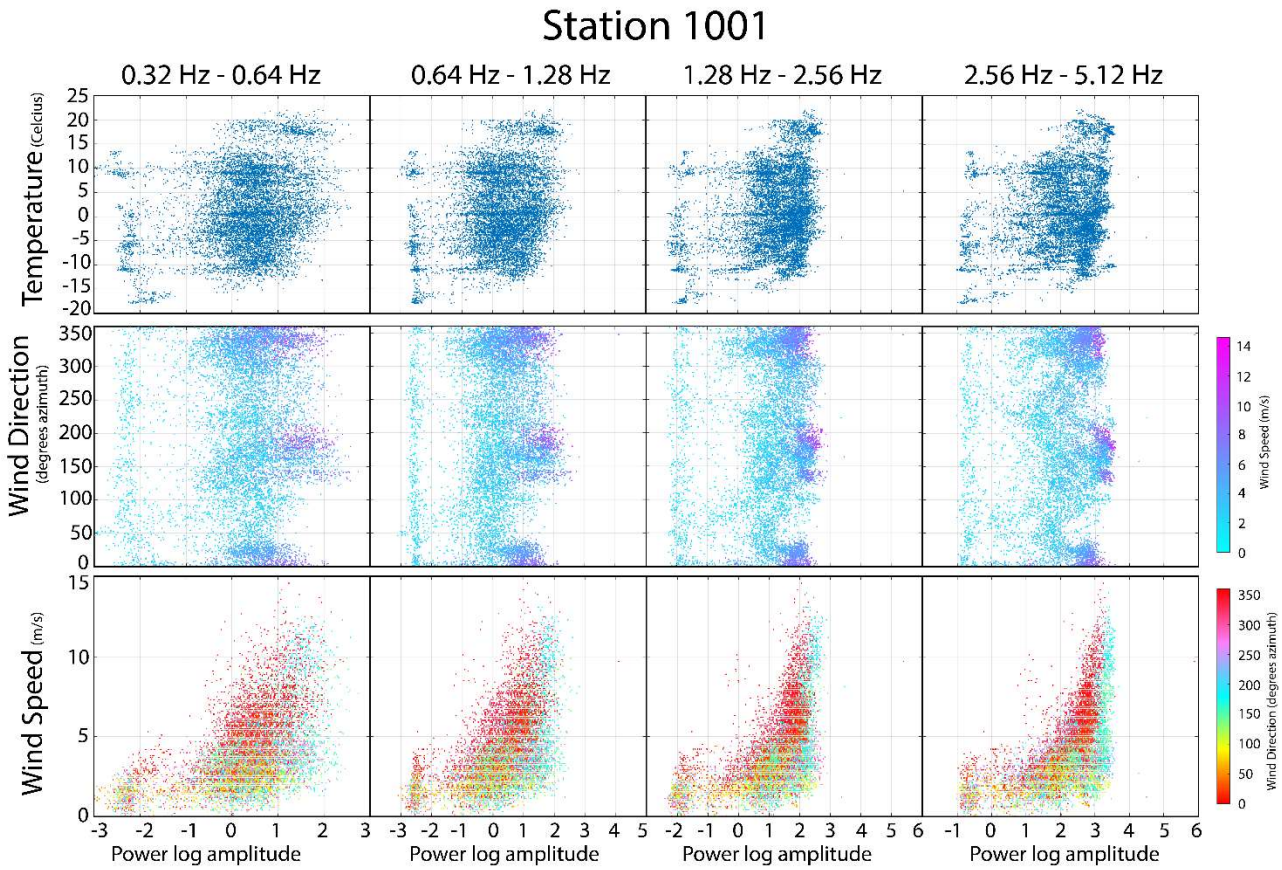

Figure 4.4: Correlation plots of Station 1001 according to pass band frequencies expressed in power log amplitudes. Mesonet station MEDF temperature, wind direction, and wind speed data are used for this comparison. The top row is temperature, the center row is Wind direction with color showing wind speed, and the bottom row is wind speed with color showing wind direction. Color bars on the right show wind speed and wind direction.

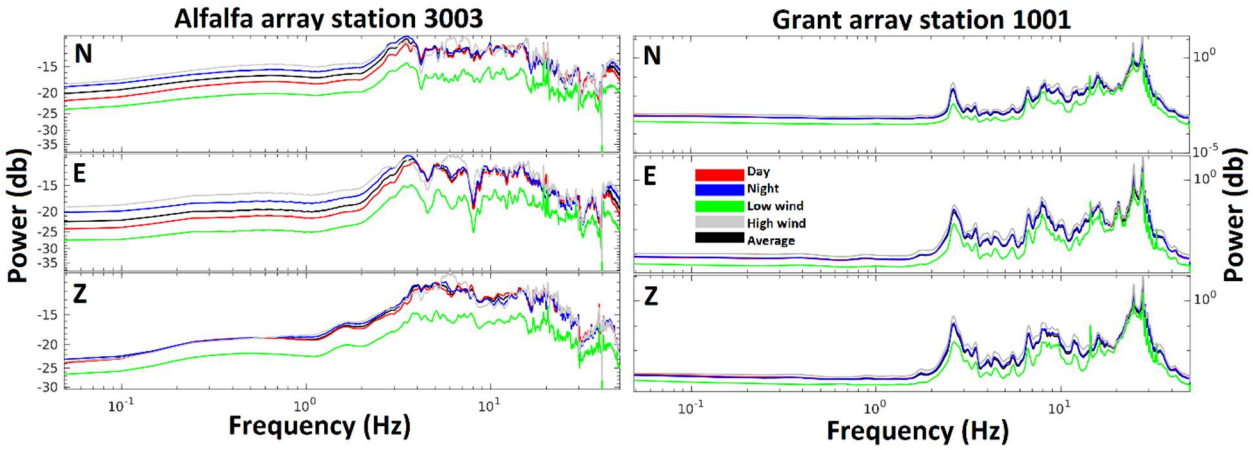

Figure 4.5: Average power spectrum of each component for all stations in both arrays. Spectral curves show differences in day, night, windy, non-windy, and total average.

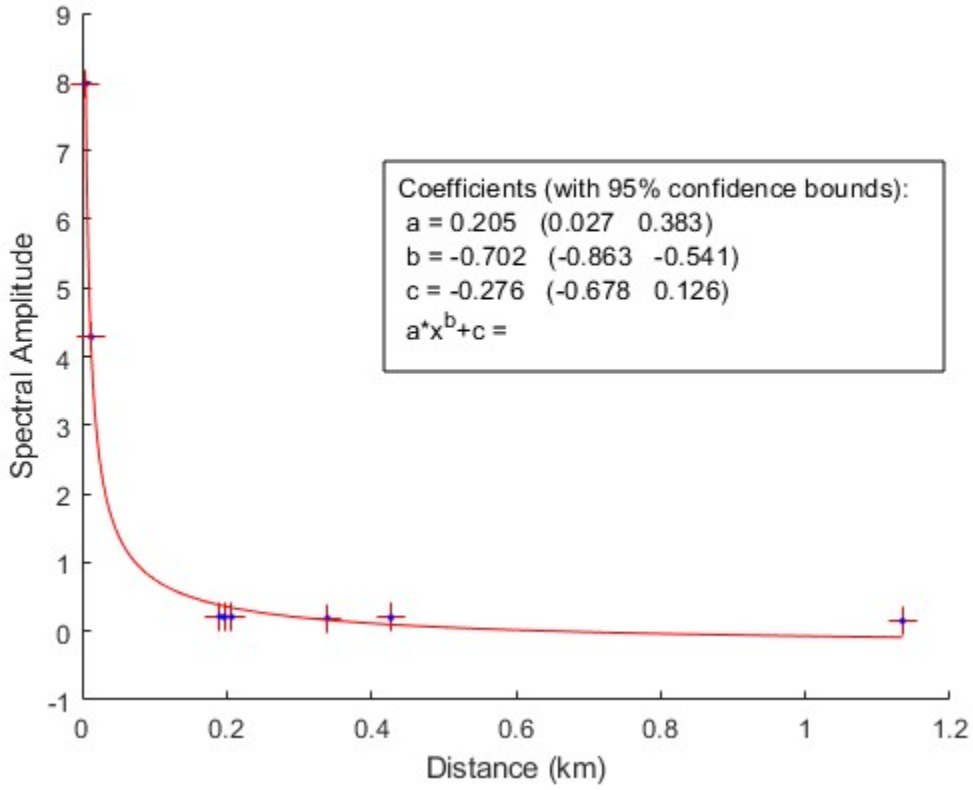

Figure 4.6: Amplitude spectrum median average noise per station listed by distance from the nearest wind turbine. A squared polynomial best-fit line is used to approximate a spatial signal strength decay.

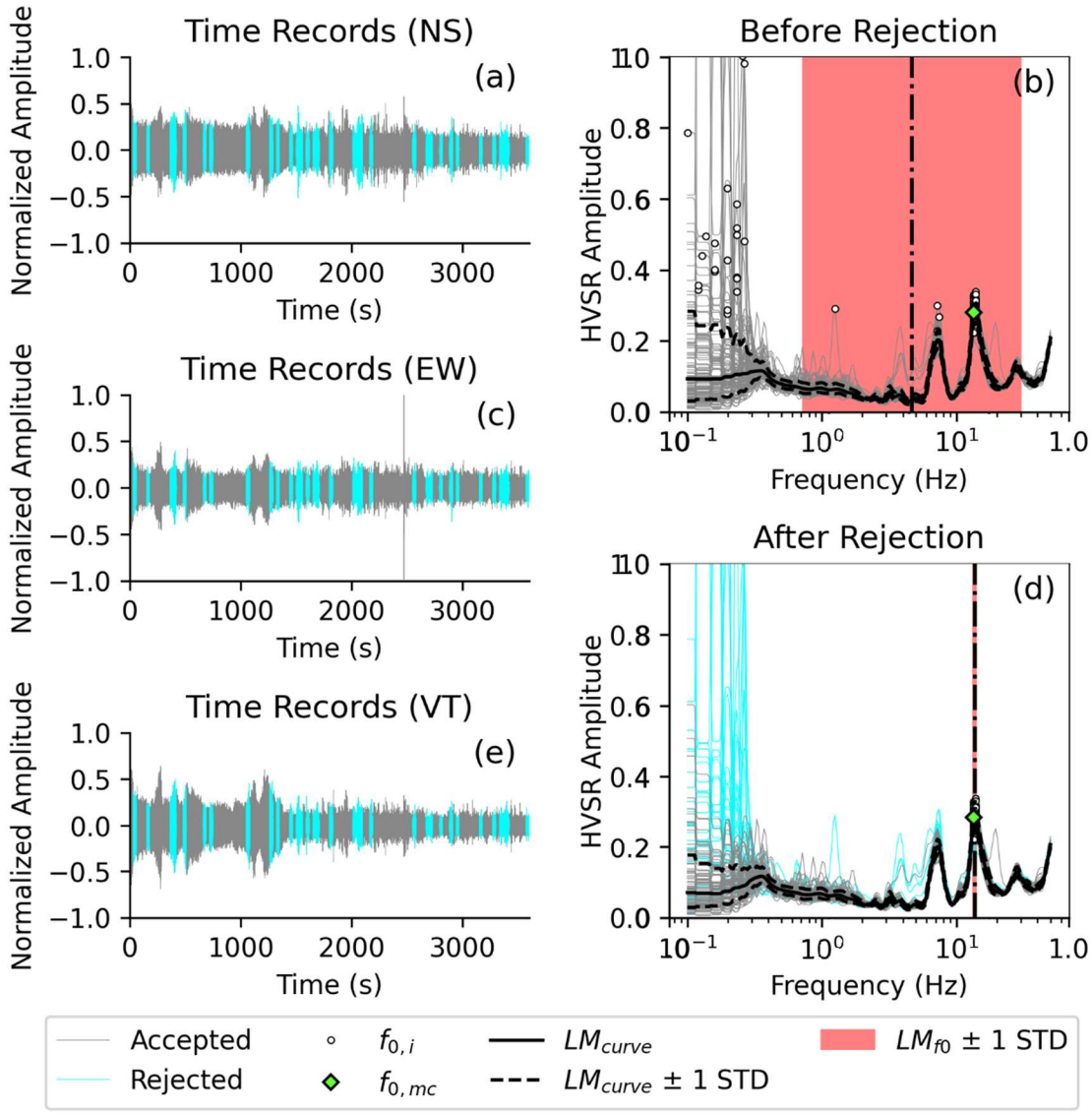

Figure 4.14: HVSR results for day 353 at hour 12 from using HVSRpy. Spectral rejection is shown to improve HVSR peak frequency selection and standard deviation.

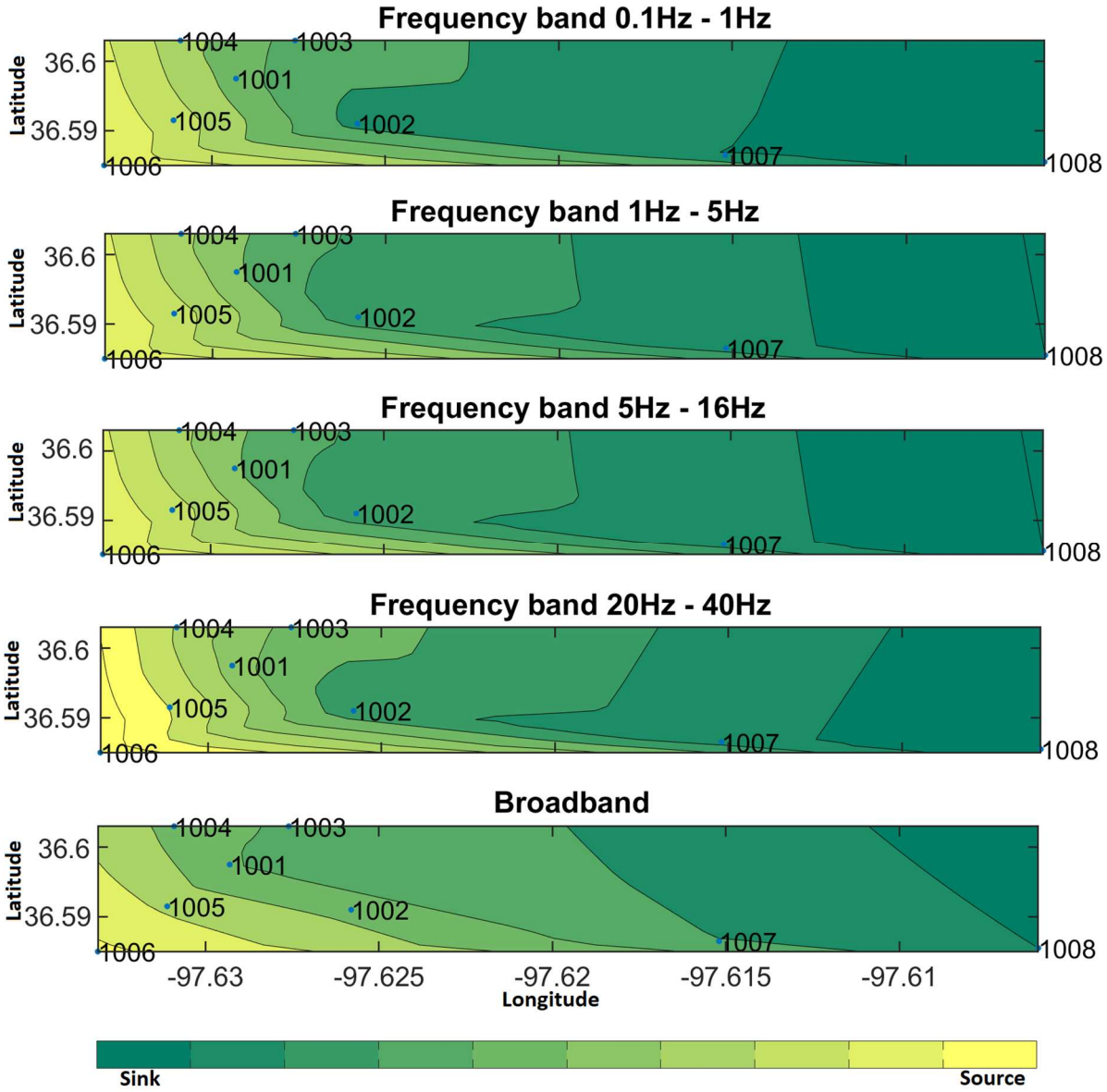

Figure 4.8: Ambient seismic wave propagation directivity from frequency bands within 0 - 50 Hz and broadband cross coherence for station pairs. The seismic wave source is shown in yellow compared to the direction of travel increasing in green tint.

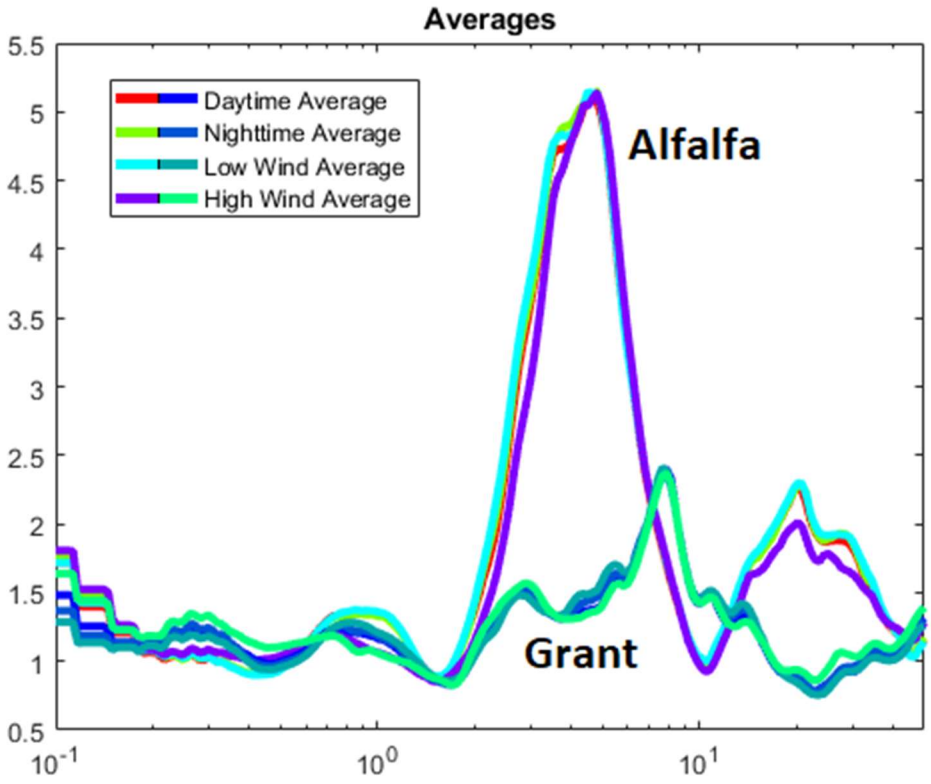

Figure 4.9: Average HVSR spectral curve between both Grant County and Alfalfa County. Differences in day, night, low wind, and high wind periods are insignificant for the HVSR method.

References:

Bard, P.Y., Jongmans, D., Ohrnberger, M., & Wathelet, M. (2005). Site effects assessment using ambient excitations (SESAME). Europeam Commission Research General Directorate, Final Repo(Project No. EVG1-CT-2000-00026 SESAME).

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, (1995) The Oklahoma Mesonet: A technical overview. J. Atmos. Oceanic Technol., 12, 5-19.

Bromirski, P. D., Duennebier, F. K., & Stephen, R. A. (2005). Mid-ocean microseisms. Geochemistry, Geophysics, Geosystems, 6(4).

Chatelain, J. L., Guillier, B., Cara, F., Duval, A. M., Atakan, K., Bard, P. Y., & Wp02 Sesame Team. (2008). Evaluation of the influence of experimental conditions on H/V results from ambient noise recordings. Bulletin of Earthquake Engineering, 6, 33-74.

Cheng, T., Hallal, M. M., Vantassel, J. P., & Cox, B. R. (2021). Estimating unbiased statistics for fundamental site frequency using spatially distributed HVSR measurements and Voronoi tessellation. Journal of Geotechnical and Geoenvironmental Engineering, 147(8), 04021068.

Cooley, James W., and John W. Tukey (1965). "An algorithm for the machine calculation of complex Fourier series." Mathematics of computation 19.90, 297-301.

Estrella, H. F., Korn, M., & Alberts, K. (2017). Analysis of the Influence of Wind Turbine Noise on Seismic Recordings at Two Wind Parks in Germany. Journal of Geoscience and Environment Protection, 5(05), 76.

European Wind Energy Association. (2012). Wind energy-the facts: a guide to the technology, economics and future of wind power. Routledge.

FEMA. 2018. “FEMA's Hazus program.” Federal Emergency Management Agency. Accessed August 18, 2018. https://www.fema.gov/hazus

Giocoli, A., Stabile, T. A., Adurno, I., Perrone, A., Gallipoli, M. R., Gueguen, E., ... & Piscitelli, S. (2015). Geological and geophysical characterization of the southeastern side of the High Agri Valley (southern Apennines, Italy). Natural Hazards and Earth System Sciences, 15(2), 315-323.

Hemami, A. (2012). Wind turbine technology. Cengage Learning.

Kloesel, S. J. Stadler, A. D. Melvin, A.J. Sutherland, and H. Shrivastava, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. J. Atmos. Oceanic Technol. 24, 301–321.

Li, C., Li, Z., Peng, Z., Zhang, C., Nakata, N., & Sickbert, T. (2018). Long-period long-duration events detected by the IRIS Community Wavefield Demonstration Experiment in Oklahoma: Tremor or train signals?. Seismological Research Letters, 89(5), 1652-1659.

McNamara, D. E., & Buland, R. P. (2004). Ambient noise levels in the continental United States. Bulletin of the seismological society of America, 94(4), 1517-1527.

Meng, H., & Ben-Zion, Y. (2018). Characteristics of airplanes and helicopters recorded by a dense seismic array near Anza California. Journal of Geophysical Research: Solid Earth.

Meo, M. (2006, January). The Oklahoma wind rush: A collaborative approach to sustainable energy development. In workshop on energy, environmental impacts, and sustainability, Manhattan, Kansas. http://www. engg. ksu. edu/CHSR/events/che650/200601/docs/09Meopaper. pdf (accessed October 7, 2006).McPherson, R. A., C. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A.

O'Neal, R., Hellweg Jr, R. D., & Lampeter, R. M. (2009). A Study of Low Frequency Noise and Infrasound from Wind Turbines. Epsilon Associates Inc, Maynard.

OWRB. 2018. "Groundwater Wells, Standards, and Protection in Oklahoma." Oklahoma Water Resources Board. Accessed August 18,2018. Https://www.owrb.maps.arcgis.com

Rial, J. A., Saltzman, N. G., & Ling, H. (1992). Earthquake-induced resonance in sedimentary basins. American Scientist, 80(6), 566-578.

Righter R (1996) Wind energy in America: a history. University of Oklahoma, Norman

Sigurosson, G. (2015). Seismic response of wind turbine structures in the near-fault region. University of Iceland.

Schofield, R. (2001). Seismic measurements at the stateline wind project. Rept No LIGO T020104-00-/, Laser Interferometer Gravitational Wave Observatory available at http://www. ligo. caltech. edu/docs.

Saccorotti, G., Piccinini, D., Cauchie, L., & Fiori, I. (2011). Seismic noise by wind farms: a case study from the Virgo Gravitational Wave Observatory, Italy. Bulletin of the Seismological Society of America, 101(2), 568-578.

Taranath, B. S. (2016). Tall Building Design: Steel, Concrete, and Composite Systems. CRC press.

Vidale, J. E., Bonamassa, O., & Houston, H. (1991). Directional site resonances observed from the 1 October 1987 Whittier Narrows, California, earthquake and the 4 October aftershock. Earthquake Spectra, 7(1), 107-125.

Chapter 5:

Conclusion

In the pursuit of expanding our understanding in the field of geosciences through investigating recorded waveforms, I can definitively state that there is much more to learn and uncover. Although it is quite popular to analyze large events when they happen due to the attention large-magnitude earthquakes can attract, these events are infrequent. A substitution of small events and passive seismic ideal since they can provide a proxy for large events. Passive seismic methods techniques do not require an active source where timing and positioning are known. Economically, passive seismic analysis is low-cost and simpler to logistically carry out.

I have attempted to raise issues and address the shortcomings of traditional studies of large earthquakes. In one of our studies, I analyzed a limited dataset from a classical hydraulic fracturing experiment to improve magnitude estimations and quantify the seismogenic state. I found an improvement in magnitude estimation through waveform cross-correlation and subsequent seismogenic index. In addition, the improved magnitudes and seismogenic index revealed subclusters that are orientated in a way where they are more prone to failure.

For our passive seismic experiments, I compared two newly developed seismic sensing instrument types and performed a local site analysis. DAS using dark fiber's recent introduction to seismic sensing has generated significant interest in its capabilities especially when paired with nodal sensors. I examine the noise level differences between the two sensor types and follow up with a site characterization of the array length to estimate fundamental frequencies for shallow subsurface interpretation of geologic structures.

Finally, I compare what is considered to be a high noise level environment with wind turbines to a low noise level environment without wind turbines and observe how seismic

recordings are affected in nodal data. Intuitively, I observe noise levels drop exponentially further away from a wind turbine. However, the cross-coherence of station pairs shows a general West to East wave propagation direction.

5.1 Summary of Results

Some key results are as follows:

(i) I use waveform cross-correlation to identify different fracture sets for b-value estimates and subsequent seismogenic index calculation.

(ii) In Enid (Oklahoma, US) I use both DAS and nodal array to investigate a long (20km) east-west transect. Recordings for both instruments are compared, and site response is estimated using the nodal data. Using a velocity model, I estimate the resonant subsurface layer.

(iii) In my wind turbine analysis, I compare nodal data with wind speed data to analyze the influence a wind turbine generator may have on the acoustic seismic field.

5.2 Future work

The feasibility of microseismic and passive seismic analysis is described here for the small case studies in Oklahoma. However, the application and possibilities reach beyond the scope of this dissertation. The magnitude calibration and subsequent seismogenic index approximation may have applications in other cases outside of hydraulic fracturing. Current research in the space of carbon sequestration and geothermal energy may be dependent on accurately tracking fracture networks to facilitate proper operation without the risk of an

anomalous fracture network that may have a negative impact. Extrapolation of our novel approach in microseismicity analysis can improve fracture monitoring.

DAS is developing into an interesting monitoring system capable of being successful in many environments. It has evolved significantly since its early applications in boreholes and ocean bottoms. However, noisy environments such as within cities or roadside still pose a challenge to DAS. One of the biggest challenges that became apparent with this study is the coupling of the fiber optic cable when using a dark fiber array. One way I attempted to overcome this is by normalizing the noise level for each channel.