UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

COMPARING AMBIENT-NOISE-BASED SEISMIC VELOCITY VARIATIONS WITH DYNAMIC AND STATIC STRAIN CHANGES ASSOCIATED WITH MAJOR EARTHQUAKE RUPTURE AT PARKFIELD

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

HAOYU LI

Norman, Oklahoma

2023

COMPAIRING AMBIENT-NOISE-BASED SEISMIC VELOCITY VARIATIONS WITH DYNAMIC AND STATIC STRAIN CHANGES ASSOCIATED WITH MAJOR EARTHQUAKE RUPTURE AT PARKFIELD

A THESIS APPROVED FOR THE

SCHOOL OF GEOSCIENCES

BY THE COMMITTEE CONSISTING OF

Dr. Junle Jiang, Chair

Dr. Brett M. Carpenter

Dr. Jacob Walter

Acknowledgements

First of all, I would like to express my most heartfelt thanks to Dr. Junle Jiang for his meticulous and patient guidance as well as his serious and responsible attitude during my research and thesis writing process. In the past two years, Dr. Jiang always pushed me to do my best, and his guidance was an indispensable factor for my successful graduation. It was also under his supervision that I completed my first serious and complete research experience in my life. I learned from him not only a wealth of scientific knowledge, but also a meticulous attitude towards research. Dr. Jiang's words and teachings have enriched my experience in the past two years and filled me with interest and confidence in research. Throughout the process of writing this thesis, Dr. Jiang provided me with quick and detailed feedbacks, which made possible a huge improvement in the quality of my thesis. I feel myself so fortunate to have had the opportunity to work with such a dedicated and inspiring advisor.

I am grateful to Dr. Brett Carpenter and Dr. Jacob Walter for their support and time reading this thesis and providing valuable suggestions. Their insights and expertise are instrumental in the development of this work.

I also would like to express my greatest gratitude to my parents. If it was not them, I would never have the chance to pursue my dreams and freely explore my life. It is also their selfless and great support that made me what I am now.

Lastly, I would like to express my most sincere thanks to my girlfriend Xiangwang Gong for her companionship and encouragement during my hard times, and to my best

friend Haixiang Yin for his reliable and selfless assistance on programming. Their love and belief kept me motivated and hopeful all the time.

Thank you all again from the deepest part of my heart.

Abstract

Stress perturbations across a fault zone can change seismic wave velocity in the adjacent area, with implications for rock properties and fault mechanics. Ambient noise seismic interferometry is increasingly used to retrieve information about wave propagation between seismic stations, constrain seismic velocity changes, and potentially monitor the subsurface strain field. Comparing seismic velocity changes with strain changes in space and time may help us understand the damage and healing processes of fault rocks. Here we study the seismic velocity and strain changes in the shallow crust associated with the 2004 Mw 6.0 Parkfield earthquake in California. We process continuous seismic recordings from 13 stations of the High Resolution Seismic Network (HRSN) from 2001 to 2007. Then we use an open-source software NoisePy (Jiang and Denolle, 2020) to cut band-pass-filtered data into 1-hour chunks and perform cross-correlation, stack hourly cross-correlation functions into 30-day averages, and retrieve reference Green's functions based on a criteria of cross-correlation coefficients. To estimate the history of relative velocity changes, *dv/v*, we use a moving-window cross spectrum method with a moving window of 30 days. To resolve potential spatial variations in velocity changes, we first fit the *dv/v* history with a parametric function with seasonal, offset, and logarithmic terms, then select better-fitting station pairs based on signal-to-residual ratios and map the coseismic velocity changes at 5 frequency bands. We also assess the relationship between *dv/v* estimates, peak dynamic strains from strong-motion seismograms, and static strain changes inferred from coseismic and postseismic space-geodetic measurements. Our results show larger amplitudes of coseismic velocity drops in higher frequency bands for both network-averaged and station-pair *dv/v* values. We attribute the instantaneous drop of *dv/v* to ground shaking during the earthquake and the later gradual increase of *dv/v* to the postseismic recovery of shallow fault zones. While the coseismic *dv/v* changes suggest potential spatial variations across the fault, future efforts are needed to quantify and reduce uncertainties. Our study will help improve strategies to obtain robust *dv/v* time series and distinguish the causes of seismic velocity changes in the subsurface.

List of Figures

List of Tables

Table of Contents

1. Introduction

1.1. Ambient-Noise Interferometry

1.1.1. Background

In recent decades, advancements in computing power and innovative instrumentation have resulted in significant developments in the field of seismology. The exponential growth in the volume of recorded seismic data has led to the application of numerous new theories aimed at better understanding the Earth's structures and properties. Traditionally, seismologists have focused only on the portions of seismograms that display strong motion, such as the arrival of direct body waves, while disregarding other components like coda waves (the part of seismograph after direct arrivals) and background noise. Direct waves are straightforward to understand with classic ray theory. In contrast, coda waves, which are generated through multiple scattering and have complex diffusive characteristics, have not received much attention in research. While direct waves provide valuable information about the Earth's interior structures and properties, coda waves, as a complementary source, can offer additional insights, including information about the scattering properties of subsurface structures. However, their complex nature has made it challenging to understand and analyze them effectively.

Considering the conceptually easy and mathematically simple features, seismologists have long been using the direct body waves from earthquakes or controlled explosions to study the internal structure and properties of the Earth. These waves travel through the Earth's interior and their energy distribution in the phase space, which describes all parameters related to seismic waves such as amplitude and frequency, depends on the nature and position of the source and is influenced by the properties of the medium they pass through, such as the density and elastic properties of the rocks. By analyzing the point-to-point response of these waves, seismologists can obtain information about the structure and properties of the Earth's interior along the path of the waves. This information is reflected in seismic velocity variations, which can be measured by the arrival time of the body waves and dispersion curves of surface waves. One of the most significant applications of this method is the inversion of velocity variations, which provides “fundamental information about the Earth's interior… in terms of its spherically symmetric stratification… and anisotropic variations of seismic properties” (Shapiro et al., 2004). This technique has been used in various fields, such as seismic tomography, to gain insight into the Earth's structure and evolution. However, while these methods have proven to be reliable and have broad applications, they also have inherent deficiencies.

First, traditional seismic methods are limited by their dependence on earthquakes or artificial sources, which makes it impractical to make continuous or regular observations of changes in the Earth's interior. While thousands of earthquakes occur every day globally, most of these events are not useful for seismic analysis due to their low signal-to-noise ratio and potential contamination from noise. To overcome these challenges, seismologists typically rely on higher magnitude earthquakes, which occur less frequently and in more limited regions. Consequently, when high temporal resolution monitoring is required, seismologists must seek out other seismic sources and methods to meet their needs.

Second, one limitation of direct waves is that they only provide information on the Earth's interior structures and elastic properties along the paths of propagation. Variations in properties in spaces other than the direction of wave travel remain unobserved. Furthermore, the limited number and inhomogeneous distribution of large earthquakes results in a restricted spatial resolution. As a result, it is difficult to obtain a comprehensive and accurate understanding of the Earth's interior structure and properties merely using direct body wave measurements.

Third, to perform reliable inversions, seismologists need a good approximation of the location and properties of the earthquake source. However, obtaining accurate information about the source can be challenging, especially for remote earthquakes or those that occur in regions with poor seismic monitoring. In addition, uncertainties in the source location and properties can lead to significant errors in the inversion results. Therefore, an alternative approach that can better describe the Earth’s impulse response to a point source without depending on source information is favorable.

Fourth, while direct body waves are valuable for long-period measurements (> 50s) and can provide better constraints on the structures and properties of the Earth's deeper regions, they are not ideal for shallower regions due to the high degree of heterogeneity (Shapiro and Campillo, 2004). This is especially true for areas located within stable continents, where seismic activity is low, and it can be challenging to make measurements of shallow structures with sufficient resolution using teleseismic seismic (surface) waves.

1.1.2. Development of Ambient-Noise Interferometry

One approach that has gained attention in recent years as an alternative to traditional methods is ambient noise interferometry. This technique overcomes the limitations of traditional methods relying solely on earthquakes by utilizing continuous ambient noise wavefields to study the Earth's interior. The method involves cross-correlating the ambient noise wavefields at two seismic stations over a long period of time to extract a stable cross-correlation function, i.e., the empirical Green's function, for surface waves. This function describes the Earth's impulse response to a point source and is mathematically the solution to the partial differential equation of wave propagation in the elastic medium. Unlike traditional methods, no approximation of earthquake source location and properties is necessary, making it easier to extract information about the subsurface structures and properties with higher accuracy. Additionally, the method can also provide better resolution of shallow structures in areas of low seismicity, making it a valuable tool for seismologists.

The term "ambient noise" refers to the continuous and diffuse seismic signals that are present in the Earth’s background ‘humming’, originating from various sources. These sources can range from human activities to natural phenomena such as ocean and tidal forces, wind, and atmospheric variations (Stehly et al., 2006; Nicolson et al., 2012). The power spectral density of ambient noise exhibits distinct features that correspond to different sources, for example, the ocean source microseism has peak energy between 10 and 20s, allowing seismologists to identify and distinguish between them and choose the best frequency band for specific targets. Compared to direct waves from earthquakes or explosions, whose short-period signal can be obscured by attenuation, especially for long distances, ambient noise shows the ability to extract Rayleigh waves at short periods (5-40s) over hundreds of kilometers distance (Shapiro and Campillo, 2004). This feature enables the use of ambient noise for studying and monitoring the crustal and upper mantle structures of the Earth, where traditional seismic methods may not be effective. In addition, the diffusive nature of ambient noise makes it suitable for imaging the interior of the Earth in a way that is not limited by the sparsity and heterogeneity of earthquake sources. By cross-correlating ambient noise recorded at two or more seismic stations over a long period of time, the empirical Green's function for surface waves can be extracted, allowing for more precise and accurate measurements of the Earth's interior properties.

Various methods of retrieving Green’s function from a diffusive random noise wavefield, which are commonly referred as wavefield interferometry, have been developed and effectively applied in domains of oceanic acoustics (Roux et al., 2003), helioseismology (Duvall et al., 1993; Rickett et al., 2000), ultrasonic (Weaver et al., 2001) and geophysics. In the field of seismology, however, it was not until the early 2000s that seismic interferometry became widely known and underwent systematic mathematical research and analysis (Snieder et al., 2002; Campillo et al., 2003; Shapiro et al., 2004; Snieder et al., 2004, Stehly et al., 2006; Campillo, 2006; Wapenaar et al., 2010; Snieder et al., 2011; Weaver et al., 2013).

Ambient-noise interferometry, also known as seismic interferometry, involves the extraction of Green’s function from the ambient seismic waves using two fundamental techniques: cross-correlation of ambient noise wavefield and coda wave interferometry (Mao et al., 2022). The cross-correlation of ambient noise wavefield is utilized to reconstruct the Green's function of the surface wave between two points, which provides insights into the transmission of waveforms through the elastic medium. Coda wave interferometry exploits the sensitivity of the diffuse wavefield to perturbations by measuring the delay in the arrival time of coda waves. Coda waves are scattered waves that are generated by the interaction of seismic waves with the heterogeneous structure of the subsurface. Therefore, it contains much more information about crustal structures than direct body waves. This technique allows researchers to understand the characteristics of the Earth's crust with improved resolution.

The proposition of extracting surface waves from ambient noise was initially suggested by Aki (1957)using cross-coherence rather than cross-correlation. His study demonstrated that the Green's function of surface waves could be reconstructed if waves propagated homogeneously with uniform power. Subsequently, Clearbout (1968) mathematically proved that the autocorrelation, a specific type of cross-correlation, of seismograms recorded on a surface receiver from a buried source (transmission) is related to the cross-correlation of seismographs between two surface receivers (reflection). This was the first example of using cross-correlation of waves to retrieve reflection response despite their unknown phase spectrum.

It is generally agreed upon that seismic coda is the result of multiple backscattering of both body and surface waves due to heterogeneity in the medium, as noted by Aki et al. (1975). One important characteristic of seismic coda waves is that they accumulate changes in the medium along their paths. Thus, as the statistically averaged changes increase as a function of time lag, seismic coda becomes more sensitive to medium changes than direct body waves. This means that useful information about the temporal evolution of structural properties can be found in the diffusive wavefield, which can be extracted using seismic interferometry to measure the differences between wavelets. Lobkis and Weaver (2001) conducted a study that confirmed the possibility of retrieving Green's function by cross correlating a diffusive acoustic field, despite the existence of significant differences between the actual signal and the reconstructed Green's function. The authors attributed these discrepancies to the insufficient averaging of source positions, which highlights the importance of fully averaging the sources to obtain a more precise reconstruction of the Green's function. In light of their findings, Lobkis and Weaver (2001) also proposed the application of this method to seismic coda, which is considered a fully diffusive wavefield and carries information on the local stratigraphy.

In 2003, Campillo and Paul achieved the first successful extraction of the Green's function for surface waves from seismic coda. Their work utilized seismic signals recorded in Mexico to demonstrate that the cross-correlation of diffusive seismic coda contains coherent information about the elastic impulse response of the Earth between two seismic receivers. Furthermore, they found that the cross-correlation function between wavefields recorded at two stations possesses the same characteristics as surface Rayleigh waves. This study provided strong evidence for the feasibility and efficacy of using seismic interferometry techniques to analyze seismic coda and study the subsurface properties of the Earth. In early theories, it was believed that the approximation of the Green's function between two points required a homogeneous distribution of earthquakes to provide sufficient averaging of correlation functions. However, as this distribution pattern is unlikely to occur in reality, the authors proposed a solution to this problem by discussing the concept of modal equipartition of the diffuse wavefield. Equipartition refers to the uniform distribution of wave energy in phase space. Because diffusive wavefields, such as seismic coda and ambient noise, are composed of backscattering waves, their energy tends to homogenize in the diffusive regime. This property is independent of the heterogeneities of the medium and makes the distribution of earthquake locations less important.

Shapiro and Campillo (2004) and Snieder (2004) have shown that, by stacking a long periods of data, the incoherent signal from other paths can be cancelled, and the coherent signal between stations can be enhanced, allowing the inhomogeneous distribution of source locations to be ignored. However, it is crucial to ensure that the wavefield is fully diffused by using a long enough time stacking to retrieve a robust Green's function. Additionally, Snieder (2004) has emphasized that seismic sources located in the lobes on both ends of the inter-station path contribute most to the reconstruction of the Green's function. Stehly et al. (2006) have explored the asymmetry of the reconstructed Green's function and argued that the property discussed by Snieder (2004) could be used to decide the spatial density pattern of heterogeneous seismic sources and the main azimuth of energy flow. These studies have provided important insights into the retrieval of Green's function from seismic coda and its potential applications in seismic imaging and monitoring.

Cross-correlation of ambient noise produces Green’s functions between a pair of seismic stations. But a simple Green’s function between a station pair does not reflect information of temporal evolution. In order to obtain the changes in cross-correlation functions over time, seismic interferometry should be applied (Mao et al., 2022). The first implementation of seismic interferometry was done by Poupinet in 1984. Poupinet et al. (1984) utilized the doublet method to monitor velocity variations in the crust beneath the Calaveras Fault, California. In his work, Poupinet observed the progressive delay of coda wave arrivals and proposed that seismic interferometry can detect velocity changes as small as 0.01%. Poupinet's study was significant because it demonstrated that seismic interferometry could be used to monitor small changes in the subsurface, which was previously not possible with traditional seismic methods. But instead of using merely coda wave as most researchers do nowadays, Poupinet used whole seismograms for their calculations.

Coda-Wave Interferometry (CWI) was introduced as a method for studying the temporal and spatial variations of the subsurface medium by Snieder et al. in 2002. The theory assumes that the subsurface contains isotropic scattering points that randomly scatter seismic waves. The multiple scattering of waves leads to the formation of a diffusive field, or coda, which can be used to retrieve the Green's function of the medium. CWI involves cross-correlating coda waves recorded by a repeatable source at different receiver pairs and in different time windows. The mean and variance of wave arrival time changes, known as the differential travel times, can be calculated using the cross-correlation functions. Because point sources are excluded from this process, the differential travel times provide information on the temporal and spatial variations of the subsurface medium, including changes in seismic velocity, attenuation, and scattering strength. This method is also known as moving-window cross correlation method, and the linear relation between velocity change and time shift was introduced as shown in the equation $\frac{\delta v}{v} = -\frac{<\tau>_{(t,\tau)}}{t}$, where δ is the root mean square displacement of perturbed scatters, $< \tau >_{(t,\tau)}$ is the mean travel time change. However, Snieder et al. (2002) did not test the CWI method on real seismic data in their original paper. Instead, they validated the method using laboratory experiments on granite samples under different temperature conditions. The results of the experiments confirmed the feasibility of the CWI method for monitoring small changes in the medium with high precision.

Shapiro and Campillo (2004) demonstrated the effectiveness of cross-correlating continuous ambient noise recorded at different seismic stations for extracting coherent empirical Green’s functions of Rayleigh waves in a wide range of periods, by assuming that ambient noise fields are fully diffusive and the sources are isotropic. As a result, the retrieved Green’s functions contain statistically averaged information about any possible path and corresponding elastic properties between pairs of stations. This approach was used to retrieve the Rayleigh wave Green's functions and dispersion curve for a network of seismometers in southern California. The results matched well with observations of dispersion curves made using traditional methods and showed better measurements in short period ranges (5-40s). The success of this method has led to its widespread use in seismology and has opened up new avenues for studying the Earth's interior structure.

Since the seminal work by Shapiro and Campillo (2004), ambient noise interferometry (ANI) has rapidly developed and been applied in numerous studies across the globe. In the past decade, many comprehensive tutorials on data processing techniques have been published, providing general guides for the development of more advanced methods (Bensen et al., 2007). ANI has been extended to other fields such as civil engineering, and the noise sources have expanded from oceanic microseism to higher frequency sources like anthropogenic activities and freight trains. Nevertheless, ANI remains a powerful tool in monitoring and tomography studies, where its unique features of continuously repeating noise sources and high spectral power in short periods can capture temporal changes in crustal properties and complement the shortcomings of traditional methods.

ANI has been widely applied in monitoring and predicting volcanic eruptions (Brenguier et al., 2008a; Mordret et al., 2010), detecting seismic velocity changes caused by major earthquakes (Brenguier et al., 2008b; Hobiger et al., 2012; Liu et al., 2014; Taira et al., 2015; Wu et al., 2016; Hillers et al., 2019; Wang et al., 2019), aseismic slow slip events, non-volcanic tremors (River et al., 2014), and velocity changes caused by seasonal variations, ground water content, and tides (Wang et al., 2017; Mao et al., 2019, 2022). ANI has also been used for high-resolution tomography of the crust and mantle (Shapiro et al., 2005; Lin et al., 2008; Nicolson et al., 2012; Roux et al., 2016; Berg et al., 2018; Retaileau et al., 2020).

In addition, many open-source Python packages have been developed for ANI and related analysis, which allow easy access and manipulation of ANI data (Lecocq et al., 2014; Jiang et al., 2020). The development of these packages has greatly facilitated ANI applications and expanded the user community. Future work on ANI is expected to further improve the accuracy and resolution of tomography, as well as its applications in seismic hazard assessment, resource exploration, and environmental monitoring.

1.1.3. Basics of Ambient-Noise Interferometry

The mathematical derivation for Green's function reconstruction from diffuse wavefield is complicated. Here I show a simplified explanation from Snieder and Larcose (2004). We start with a one-dimensional scenario, where two stations A and B locate at x = R and x = 0, respectively. The incident wave with spectrum S_L comes from left at station B and the incident wave with spectrum S_R from right at station A. Both random incident waves are uncorrelated and have the same power spectrum. The superposed wave u_A and u_B recorded at two stations are

$$u_A = S_L e^{ikR} + S_R e^{-ikR} \quad , \; u_B = S_L + S_R \tag{1}$$

Where *k* is the wave number at angular frequency ω. Because the waveforms are represented in the frequency domain in this case, cross-correlation in frequency domain is simply multiplication of u_A and ${u^*}_B$, where the asterisk denotes the complex conjugate. Next we represent the expectation value of incident wave spectra with angled brackets<···>. Since both incident waves are uncorrelated and have the same power spectrum, we have

$$< |S_L|^2 > = < |S_R|^2 > = < |S|^2 > \ and \ < S_L {S^*}_R > = 0 \tag{2}$$

In this setting, the expectation value $< u_A {u^*}_B >$ of the cross-correlation simplifies to

$$< u_A {u^*}_B > = < |S|^2 > (e^{ikR} + e^{-ikR}) \tag{3}$$

We know that the Green's functions of a one-dimensional homogenous medium is $G_{1D}(R, \omega) = (-\frac{i}{2k})e^{ikR}$, so the formula for cross-correlation of two wavefields can be expressed as

$$< u_A {u^*}_B > = \ 2ik < |S|^2 > [G_{1D}(R, \omega) - G^*_{1D}(R, \omega)] \tag{4}$$

This equation shows that if the power spectrum of the noise is known, the difference of the Green's function and its complex conjugate can be obtained through cross correlation. The difference in frequency domain can be translated to the difference in time domain $G(R, t) - G(R, -t)$. Since the Green's function is causal, $G(R, t)$ is nonzero for $t > 0$ and $G(R, -t)$ is nonzero for $t < 0$. By breaking down the two solutions into their constituent parts and combining them, the complete Green's functions can be reconstructed. Shapiro and Campillo (2004) stated that the cross-correlation result differs only by an amplitude of spectral energy density from the actual Green's function and this result is usually referred to as the empirical Green's function for this reason. Snieder and Larcose (2004) also explored the mathematical expression of cross-correlation in 2-D and 3-D conditions. Their results showed that the expressions have the same form as the 1-D case but with different constants.

Coda wave interferometry (CWI) has enabled significant applications in monitoring the subsurface. A range of techniques has been proposed, including the moving-window cross-correlation method (Snieder et al., 2002), the moving-window cross-spectrum method, also known as the doublet method (Poupinet et al., 1984), the stretching method (Lobkis & Weaver, 2003), the dynamic time wrapping method (Mikesell et al., 2015), and wavelet cross-spectrum (Mao et al., 2019). Each of these methods targets the measurement of shifts that occur in either the time, frequency or wavelet domain. However, it should be noted that each approach exhibits different preferences and limitations.

The Moving-Window Cross-Correlation (MWCC) method is a time-domain technique that enables the measurement of seismic velocity variations. This approach operates by determining the value of time shift (*dt*) that maximizes the cross-correlation between waveforms. It should be noted, however, that an important assumption associated with this method is that the spectral content of the waveforms is similar. While this approach has shown utility in various applications (Gret et al., 2006; Snieder et al., 2006), it also exhibits limitations in its ability to accurately measure phase shifts in cases where significant spectral differences exist between waveforms.

The Moving-Window Cross Spectrum (MWCS) method is an extension of the MWCC method that enables the measurement of time shifts within specified frequency bands. Unlike MWCC, MWCS measures phase shifts in the frequency domain. This approach operates based on the assumption that *dt* increases linearly with lag time *t*, such that *dt/t* can be approximated through linear regression. One potential issue associated with MWCS is cycle-skipping, which can be mitigated through careful selection of window size; this ensures a certain level of accuracy and robustness in the measurement process.

The Stretching method is another approach that builds on the assumption that time shift, *dt*, linearly increases with lag time, *t*. By linearly stretching one waveform, the time shift can be determined as the value that maximizes the correlation coefficient between the two waveforms. However, this method has limitations. For example, because the stretching method measures time shifts over the entire trace, it fails to localize perturbations within the coda. As a result, it limits the ability to recognize changes associated with specific scatters. Additionally, this method is not well-suited for broad-band waveforms that have different spectral contents. Despite these limitations, the Stretching method represents a useful tool in certain applications where high accuracy in measuring time shifts is required (Zhan et al., 2013).

The Dynamic Time Warping (DTW) method is similar to the Stretching method in that it allows for adaptive stretching factors at each lag time. However, in DTW, the phase shifts, *dt*, are determined by the shortest warping path, which can help to mitigate the problem of cycle-skipping. Despite this advantage, DTW also has limitations. For example, it may lack accuracy in localizing scatters in coda waves.

The Wavelet Cross Spectrum (WCS) method estimates velocity changes in both the time and frequency domains, overcoming the limitations of poor time-frequency resolution associated with Fourier transform-based methods. WCS utilizes the wavelet transform, which generates a complex-valued time-frequency field. The time shifts retrieved in this field are functions of both time and frequency changes. While WCS shows promise in suppressing many of the shortcomings of other methods (e.g., cycle skipping), its measurement is only reliable when the time-frequency space has sufficient power. Furthermore, it also experiences the cycle-skipping problem.

As discussed above, each method in coda wave interferometry has its own advantages and disadvantages. Selecting the appropriate method requires careful consideration of the wavefield properties and the targeted questions. Among the various methods available, the most popular ones include MWCS, Stretching, and DTW. However, other methods are also being utilized for various purposes. Ultimately, the selection of the most suitable method (e.g., Mikesell et al., 2015; Yuan et al., 2021) is crucial for accurate and reliable measurements in coda wave interferometry.

1.2. Parkfield

1.2.1. Geologic Setting

In this thesis, I focus on studying the crustal structure in Parkfield, California. Parkfield is a small town located in the central coast region of California. It is situated in the southern part of Monterey County, about 200 miles northwest of Los Angeles and 175 miles southeast of San Francisco. The town is located at the boundary of the North American and Pacific tectonic plates along the San Andreas Fault, making it a region of great geological interest. The geology of Parkfield is characterized by a mix of sedimentary and volcanic rocks, which have been shaped by the complex tectonic activity in the area. This tectonic activity has resulted in the formation of several fault structures, making this area an ideal place for investigating the physical properties and processes that control deformation (Catchings et al., 2002).

The San Andreas Fault is a major right-lateral transform fault that extends for more than 1,200 kilometers across California. The fault segment at Parkfield is known to host recurring M~6.0 earthquakes (Bakun & Lindh, 1985). This makes it an important site for earthquake research and monitoring, leading to the United States Geological Survey (USGS) Parkfield Earthquake Experiment in 1985 and 1993 (Bakun & Lindh, 1985). Many seismic arrays, such as the USArray and borehole sensors, are also installed in this region. The area has been the site of many seismic studies (Unsworth et al., 1997; Li et al., 2007; Brenguier et al., 2008; Zhao et al., 2010; Wu et al., 2016), focusing on properties of the San Andreas Fault, making Parkfield one of the most well-studied regions in the world, with a wealth of geological and geophysical data available for analysis.

Studying the crustal structure of Parkfield has the potential to provide insights into the complex geological processes that occur at plate boundaries. The presence of multiple fault structures and the absence of significant earthquake activity in the region make it a unique site for geological research, particularly in the study of fault behavior and earthquake prediction.

1.2.2. Previous Research

One of the most significant earthquakes at Parkfield was a magnitude 6 event in 2004. This earthquake has been extensively studied by scientists to better understand its properties and impacts. Researchers have used various methods to investigate the effects of the Parkfield earthquake, including changes in seismic velocity, direct body waves, and ambient noise interferometry.

Li et al. (2007) found a decrease in direct body waves after the Parkfield earthquake and concluded that the velocity variations occurred in a zone of approximately 200 meters wide and extended to depths of up to 6 kilometers. They suggested that the earthquake had a significant impact on the seismic velocity of the fault damage zone.

Brenguier et al. (2008) used ambient noise interferometry to detect small seismic velocity changes associated with the earthquake. Their continuous monitoring results showed an instant velocity drop after the earthquake, followed by a gradual recovery to pre-earthquake levels. However, they did not constrain the depth variation of the velocity variation.

Wu et al. (2016) expanded on the previous research by investigating the velocity changes at different depths by cutting broadband waveforms into subsequent narrow frequency bands. Their results showed that coseismic velocity change increased in amplitude with frequency band. So higher frequency bands, which correspond to shallower depths, demonstrate more apparent coseismic velocity changes. After careful inversion, they proposed that the S-wave velocity change extended to at least 1.2 km, providing new insights into the depth variation of the velocity change.

Other researchers, such as Rubinstein and Beroza (2005), Zhao et al. (2010), and Schaff (2012), have also studied the Parkfield earthquake using different methods. These studies have contributed to a better understanding of the properties and impacts of this significant earthquake.

1.3. Motivation and Objectives

Previous research by Brenguier et al. (2008) and Wu et al. (2016) has investigated the temporal evolution of seismic velocity variations and their depth changes. However, their findings were based on averaged measurements of dv/v across all station pairs, which only reflected overall changes in the study region and do not capture any spatial variations.

To address this issue, an alternative approach that avoids averaging between station pairs has been proposed. Hobiger et al. (2012) examined the coseismic velocity changes and postseismic recovery process of the 2008 Iwate earthquake using measurements made between individual station pairs. Their study demonstrated the feasibility of utilizing single station pair measurements to investigate spatial variations in seismic velocity changes. Similarly, Taira et al. (2015) also studied the spatial variation of seismic velocity changes associated with the 2014 Napa earthquake and found that the velocity changes were correlated with dynamic peak strain estimates, suggesting that they were caused by fractures during the mainshock.

In addition to exploring the spatial variation of seismic velocity changes, another objective of this project is to estimate how these changes are related to dynamic and static strain changes. Rivet et al. (2013) provides a good example of mapping strain rate with velocity variation and demonstrating their causal relationship. The goal of this study is to investigate similar patterns in Parkfield and gain a better understanding of the properties of the San Andreas Fault.

To find the spatial pattern of velocity variations and find the relation between strain changes, I generally follow methods described in Brenguier et al. (2008), Rivet et al. (2013), Taira et al. (2015) and Wu et al. (2016). The same datasets from the HRSN network will be used and cover a similar period. Crucial data processing methods and parameters will be made the same as well. More details about methods and workflow will be discussed in the following sections. Our results demonstrate similar seismic velocity change patterns as stated in previous research (Brenguier et al., 2008; Wu et al. 2016). Meanwhile, we investigate and map signal-station-pair coseismic velocity changes to compare with static and dynamic strain changes. We observe some relationships between seismic velocity and strain changes, but a clear quantitative relationship remains obscure.

2. Data

2.1. Seismic Data

This study involves the utilization of continuous seismic data obtained from 13 three-component borehole geophone seismometers from the High-Resolution Seismic Network (HRSN). This network has been specially designed to facilitate high-quality seismic data collection, essential for comprehensive geophysical analysis. The raw data in miniseed format was provided by the Northern California Earthquake Data Center's (NCEDC) FDSN web service. The data extraction employed the built-in NoisePy function, a powerful open-source Python package that assembles comprehensive ambient noise data inquiring and processing workflow (Jiang and Denolle, 2020).

The Berkeley Seismological Laboratory operates and maintains several geophysical instrumentation networks within northern and central California. These networks consist of broadband seismometers, borehole geophones, and GPS (Global Positioning System) stations, all contributing crucial data to better understand seismic activities in the region. The High-Resolution Seismic Network (HRSN) is a borehole geophone network deployed in the vicinity of Parkfield, California since 1986. Initially created as part of a controlled-source monitoring project, namely the Vibroseis project (HRSN, 2014), the network later underwent an upgrade to observe and monitor local microseismicity on the San Andreas Fault. The network code for HRSN is BP, which includes 13 stations distributed across the San Andreas Fault, as depicted in Figure 1. The geophone sensors, meticulously installed from 60 to 300 meters below the ground level, were designed to mitigate the effects of incoherent noise generated by temperature fluctuations, precipitation, or human activities.

For investigating seismic velocity alterations before and after earthquakes, we analyze data spanning from January 2001 to October 2007. This time span encapsulates both the 2003 M6.5 San Simeon earthquake and the 2004 M6.0 Parkfield earthquake. This extensive period of data collection ensures adequate scattering averaging, which is essential for reconstructing reliable Green's functions. Following the methodology employed by Brenguier et al., 2008, and Wu et al., 2016, we select the BP channels of the geophone sensors, which operate at a sampling rate of 20 Hz. To optimize our data processing, we restrict our calculations to the vertical component as it demonstrates higher sensitivity to surface waves compared to other directions. In Section 3.1.4, we will demonstrate that waveforms stacked using the vertical component yield a higher signal-to-noise ratio than those processed using other components. Table 1 provides an overview of all parameters required for downloading the raw data for reviewing purposes.

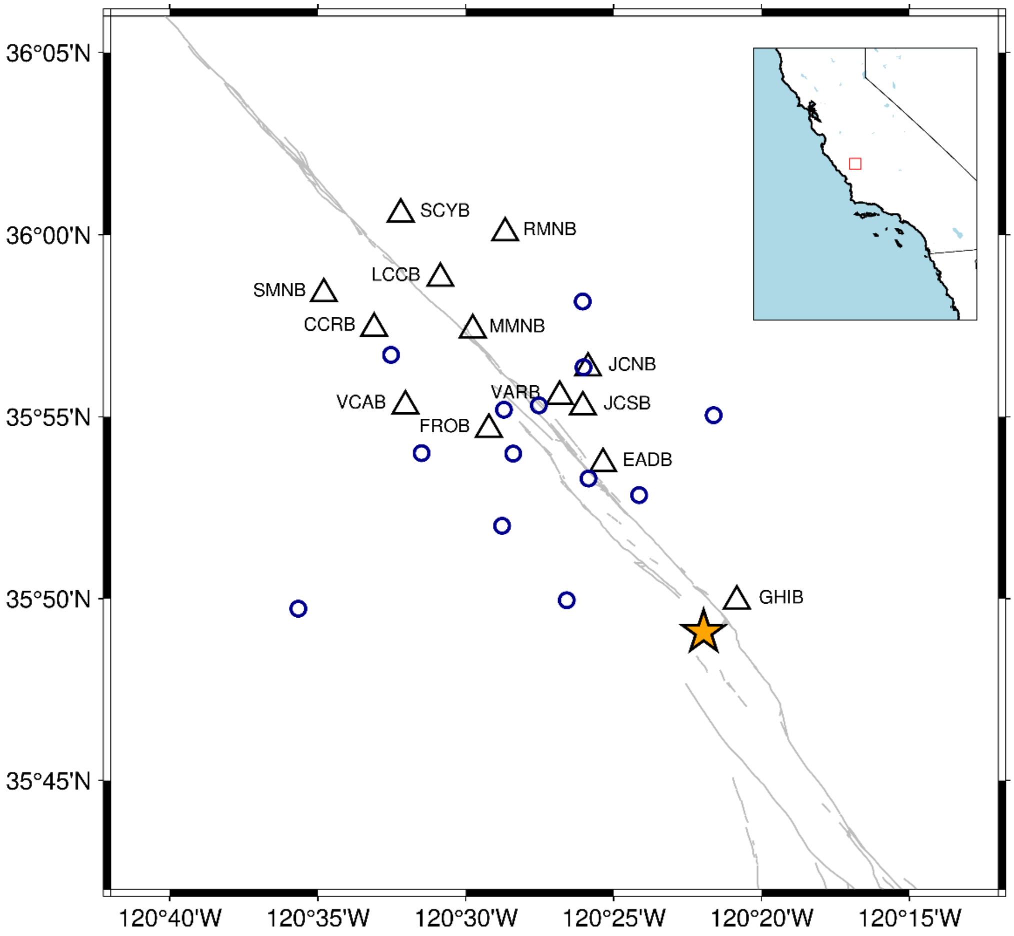


Figure 1. Station Map. Black triangular are seismic stations. The yellow star marks epicenter of 2004 Parkfield earthquake. The blue circles are GPS stations. The grey lines are traces of the San Andreas Fault. The inset map shows location of this area with blue marks the ocean.

Network Code	BP
Channels	BP1 (vertical)
Number of Stations	13
Period	01/01/2001 to 10/02/2007
Sampling Frequency	20 Hz
Frequency Band	0.08 to 2.0 Hz

Table 1. All parameters used for downloading raw seismic data

2.2. Raw GPS and Strong Motion Data

We use high-rate GPS displacement time series and strong motion seismograms (Shakal et al., 2006; Jiang et al., 2021) to estimate static and dynamic strains of the shallow crust. GPS data are downsampled from the combined 1 Hz and daily solutions recorded in a local real-time GPS network, which spans the entire coseismic and early postseismic phases of the 2004 Parkfield earthquake. Details of GPS data processing are described in Jiang et al. (2021). Network-wide common mode noise is reduced in the daily GPS solutions with a Principal Component Approach (Dong et al., 2002), whereas multipath effects are mitigated through a modified sidereal filtering approach (Choi et al., 2004). Furthermore, interseismic trends and offsets due to other events or artifacts have been corrected in the post-processing (Jiang et al., 2021). Velocity seismograms for the earthquake over durations of up to 30 s are obtained for 60 strong motion stations from the Center for Engineering Strong Motion Data (CESMD; https://strongmotioncenter. org/). Recording of these seismograms was triggered when the amplitude of acceleration reached 0.005 g at most stations.

3. Methods

In this section we introduce our workflow for data processing and seismic velocity variation computation. We also elaborate on the rationale behind the selection of various techniques and parameters. Lastly, this section also includes methods for processing geodetic and strong motion data to obtain both static and dynamic strain data.

3.1. Seismic Data Processing

3.1.1. Workflow

The methodology of ambient noise interferometry has been well-established and standardized in the past decade. Bensen et al. (2007) provides a comprehensive guide on processing ambient noise data to retrieve Green's function, which has become a standard workflow for ambient noise interferometry. However, it is important to note that many techniques outlined in this tutorial, such as normalization and stacking methods, along with other key parameters, should be adjusted accordingly based on the requirements and objectives of individual research projects. The data processing and analysis procedures implemented in our study generally follows the methodologies depicted in the works of Brenguier et al., (2008), Taira et al., (2015), and Wu et al., (2016). The overall workflow of data processing and analysis procedures is presented in Figure 2, which provides a step-by-step overview of the entire process form raw data acquisition to dv/v measurement.

The complete workflow of our study comprises four steps: preprocessing, cross-correlation, stacking, and measuring dv/v (relative velocity changes). The first step is filtering the data into user-defined frequency bands, checking for recording gaps, and segmenting the continuous data into chunks corresponding to specified time intervals. The preprocessed data are then stored in sequential h5 files, with each file containing complete information about the station and waveforms within that particular period.

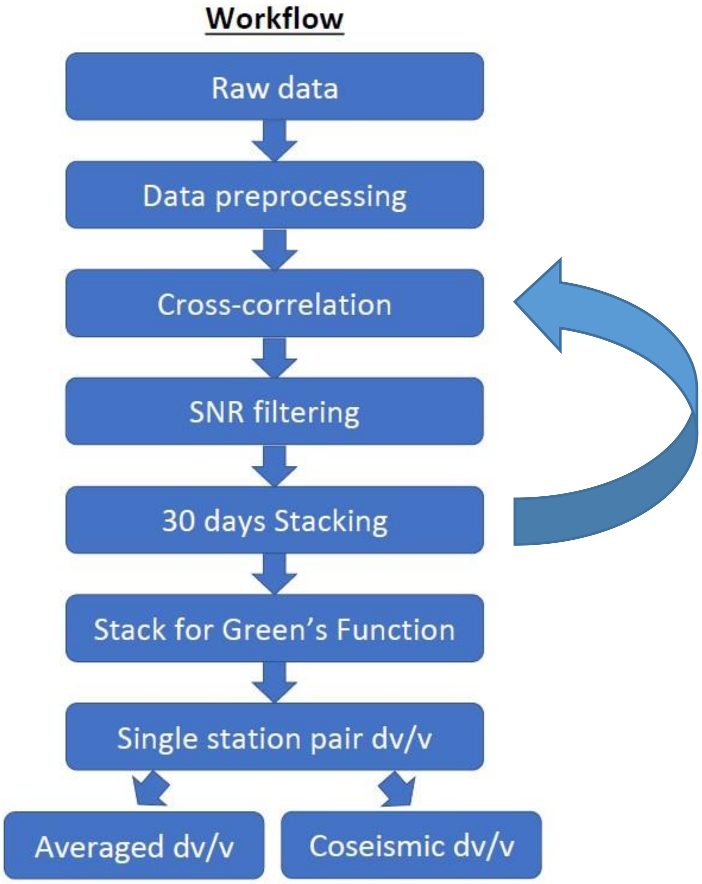


Figure 2. Complete workflow of data processing and dv/v calculation procedures. The arrow means return to SNR filtering step and proceed to the next steps again.

The second step involves data processing and cross-correlation. Here, preprocessed data are further processed to eliminate the effects of minor earthquakes or perturbations on the amplitude of seismic noise by clipping seismic recordings that are 10 times the overall standard deviation of the remaining seismograph, and to whiten the data in selected frequency domains. Following this data cleaning process, hourly cross-correlation functions are calculated using segments (traces) of continuous data then stacked to create substack waveforms for that time chunk. The outcomes of this step are 30-day-long substacks of cross-correlation functions between all station pairs, which are subsequently stored in the h5 files of successive time chunks.

The third step is the stacking of all substacks to generate the reference waveform, also known as the empirical Green's function of the Rayleigh wave. This reference waveform is computed by simply stacking all substacks using various methods. If necessary, the reference waveform can be employed to determine the center of the reconstructed direct wave, and we can revert to the previous cross-correlation step to execute filtering over traces based on the signal-to-noise ratio between direct and coda parts. Subsequently, all traces for substacks and the reference waveform are stacked. This approach enhances the signal-to-noise ratio of the reconstructed waveform, thereby improving the subsequent measurements. An appropriate stacking method is critical for retrieving robust reference waveforms because they act as standards for comparison with substack waveforms. Yang et al. (2022) provides a comprehensive evaluation of different stacking methods and outlines optimal selections for diverse applications that can used as a guideline for choosing the best stacking strategy.

The last step is calculating relative velocity changes (dv/v) between substacks (current waveforms) and reference waveform. Various methods can be employed to achieve this calculation, but the choice should be made on a case-by-case basis to ensure sufficient accuracy and computational efficiency. With measurements from a single station, they can be averaged to obtain an overall velocity variation in the area, or they can be used to reflect velocity changes between a station pair. Important parameters and methods utilized in the above steps are summarized in Table 2.

Frequency Bands	**0.2 to 0.6 Hz, 0.3 to 0.8 Hz, 0.5 to 1.0 Hz, 0.7 to 1.2 Hz, 0.9 to 1,4 Hz**
Cross-correlation Length	1 hour
Cross-correlation Step	0.5 hour
Cross-correlation Pair	All stations
Cross-correlation Channels	Vertical-Vertical
Time-domain Normalization	None
Freq-domain Normalization	Spectral Whitening
Earthquake Signal Removal Threshold	10*Std
Stacking Method	Linear
Reference Period	07/14/2005 to 10/02/2007
dv/v Measurement Method	Moving-Window Cross Spectrum

Table 2. Important parameters and methods used in the complete workflow

3.1.2. Software

With the increasing application of seismic interferometry, numerous open-source Python toolboxes have been developed to facilitate the processing of ambient noise data. We use NoisePy (Jiang et al., 2020), a versatile Python package capable of reconstructing Green's function from ambient noise and also incorporates application modules such as measuring dv/v and dispersion curve. The structure of NoisePy is quite straightforward and mirrors the procedure described in Bensen et al. (2007), ensuring user-friendly navigation and operation. In the main data processing segment, NoisePy features three scripts (denoted as S0, S1, and S2), each targeting one of the three steps involved in extracting the reference waveform (Figure 3). The output generated by each script is stored in separate, independent folders, which simplifies data management and retrieval. Moreover, each script implements the complete procedure for its respective step, minimizing dependencies on other scripts. This design facilitates modifications to the scripts, making them more adaptable to varying research needs. Importantly, it doesn’t necessitate advanced coding skills, making NoisePy an accessible tool for researchers with varying degrees of programming expertise.

The script S0A is utilized for downloading raw seismic data from the targeted network and webservice. The S1 script plays a critical role in data processing and calculating cross-correlations. In this step, data from each station are normalized first, then a Fourier transform is performed to convert the data to the frequency domain and calculate cross-correlation. Depending on the user-defined values, each time chunk may contain multiple or a single substack waveform. The script S2 is employed for stacking substacks and generating the reference waveform. This script also has the capability to rotate Cartesian coordinates to radial coordinates if needed. The outputs from this script are stored in h5 files corresponding to each station pair. After completing these steps, users can select application module scripts for further visualization or analysis. However, it is important to note that these scripts may not be universally applicable or comprehensive for most scenarios. As a result, they may require modifications to fit the specifics of a particular research project.

NoisePy distinguishes itself from other packages such as MSNoise by Lecocq et al. (2014), which uses seed or SAC format data and employs MySQL for parallel computing and data management. Unlike the predecessor, NoisePy initially converts miniseed data into ASDF (h5) format. ASDF format has significant advantages that can be easily read and transported in Python and other programming languages, and they can store large datasets with parallel input and output strategies. The implementation of parallel processing techniques enhances NoisePy's capacity to process data over extended periods, a capability that scales with the number of cores. This makes NoisePy particularly well-suited to the analysis of large datasets, such as those used in this study. However, there are still some improvements can be made. Functions can be added include, but not limited to, various types of filtering, selective stacking, and the capacity to process only one channel at a time. Another shortcoming of NoisePy is the lack of ability to generate various figures, which require additional work to plot desired figures for visualization and process monitoring. Overall, NoisePy provides diverse choices for each step of data processing and is a powerful tool for analyzing large datasets such as in this study (Table 3).

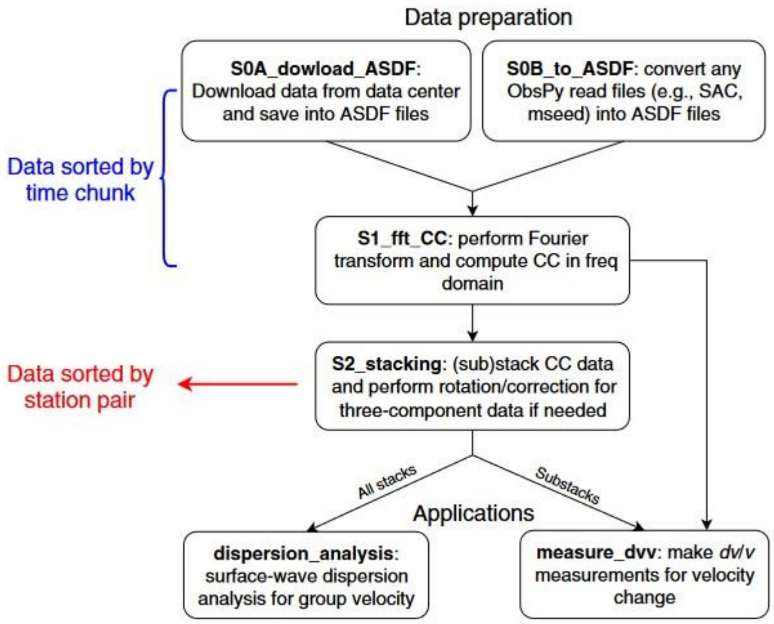


Figure 3. Workflow of NoisePy (Jiang et al., 2020).

Script	**Parameter**	**Available Methods**
	Time-domain normalization	running mean average, one-bit
	Frequency-domain normalization	running mean average, one-bit
S1	Cross-correlation methods	pure-correlation, coherency, deconvolution, phase-correlation
S2	Stacking methods	linear, phase weighted stacking, robust stacking, nroot, selective stacking, auto-covariance
Measuring dv/v	Methods	stretching, dynamic time wrapping, moving-window cross spectrum/correlation, wavelet stretching, wavelet cross spectrum

Table 3. List of some important parameters in each step and choices of methods.

3.1.3. Raw Data Preprocess

We downloaded a total of 70 months of continuous seismic data, spanning from 01/01/2001 to 10/02/2007, from the Northern California Earthquake Data Center (NCEDC). Along with the recorded waveforms from the 13 stations, additional information such as geographical coordinates was also downloaded and stored in h5 files as waveform tags. The raw data underwent an initial quality check for gaps and sampling rate. Data gaps were identified by calculating the ratio of data points with zero values to the total number of data points within each stream, an ObsPy object for data traces (Beyreuther et al., 2010). Streams exhibiting large data gaps (greater than 70%) were excluded, while the remaining data were merged after removing NaN/infinity values, mean values, and trends from each trace within the stream. Following this, the merged data was filtered to the frequency range of 0.08 to 2 Hz by using a Butterworth filter to minimize computational expense. We do not downsample data as the sampling rates for both the geophone and our requirements were 20 Hz. Instrument response removal is optional, and the last step is trimming continuous data to user-defined sequences. The output files from preprocessing procedure are 70 h5 files, with each file contains 30 days of data from all 13 stations.

3.1.4. Cross-correlation

Based on our choices of parameter values and data processing methods (Table 2), we performed cross-correlation function between seismic stations using script S1. In line with the methodology presented by Wu et al. (2016), we utilized only the vertical channel in this study. The vertical channel was chosen not only because it encapsulates more information about the surface wave, but also because it offers a high signal-to-noise ratio (SNR) for the reconstructed waveforms compared to other channels, as illustrated in Figure 4. Even though direct waves are not used for velocity change calculations, a high SNR indicates the high quality of the waveform and clearly demarcates the beginning of coda parts. Therefore, to ensure the reliability of the results, it's critical to maximize the SNR of the reconstructed waveforms as much as possible.

The preprocessed data from the previous step are segmented into five frequency bands to investigate the dependence of velocity variation on frequency and as a function of depth. These frequency bands are as follows: 0.2-0.6 Hz, 0.3-0.8 Hz, 0.5-1.0 Hz, 0.7-1.2 Hz, and 0.9-1.4 Hz. Employing narrower frequency bands can also reduce the spectral complexity of waveforms, potentially increasing their coherence. Parameters such as frequency- and time-domain normalization, data length for cross-correlation, step size, filtering methods, and substack length play a crucial role in producing coherent substack waveforms. In contrast to the approach taken by Brenguier et al. (2008) and Wu et al. (2016), which utilized both time- and frequency-domain normalization with a cross-correlation length of 24 hours, this study employs only spectral whitening and uses a cross-correlation length of 1 hour following the Welch's method (Seats et al., 2012).

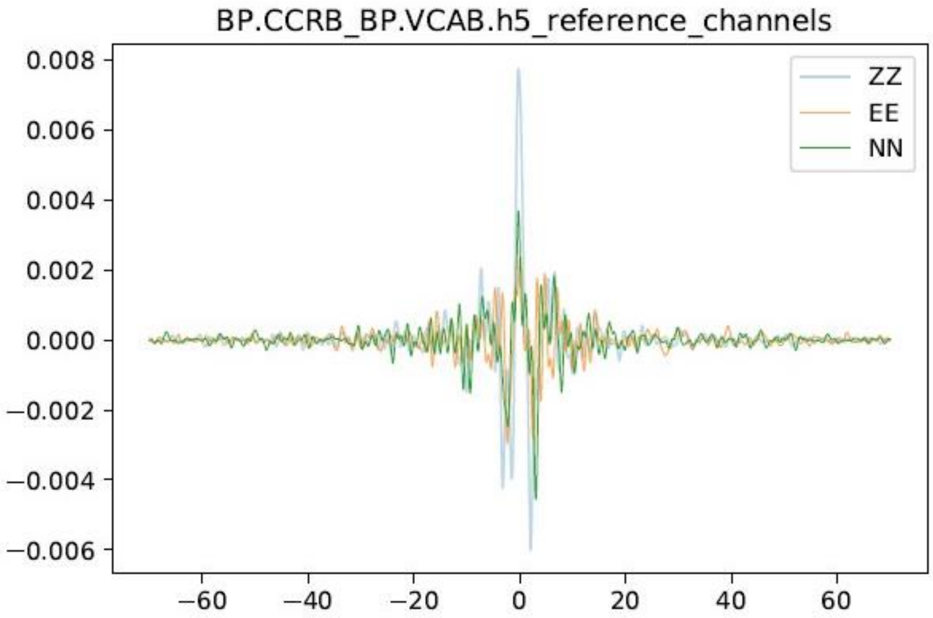


Figure 4. Reconstructed reference waveforms using different channels

Time-domain normalization can mitigate the effects of large amplitude difference by a few of orders of magnitude in ambient noise signal, which may overshadow the most energetic parts of the noise. Frequency-domain (spectral) whitening can broaden the bands of ambient noise, enabling the extraction of longer periods of data (Bensen et al., 2007). Without this, the spectral energy of ambient noise would concentrate mostly in the microseism band. Seats et al. (2012) suggests that with a cross-correlation length of 1 hour and 50% overlap, the stacked cross-correlation function swiftly converges to a robust noise correlation function. This occurs even faster than the normalized long correlation length time series. The authors have demonstrated that time-domain normalization, as widely assumed necessary before, is not required and does not accelerate the entire process.

Figure 5 compares waveforms obtained using Welch's method (labeled 1h) and traditional method (labeled 24h). Two waveforms are almost identical and the one generated using Welch's method has slightly higher SNR. Therefore, it is hard to decide which cross-correlation length is better because no major changes are observed between waveforms. But if we consider consistency of the reconstructed waveforms over time, the Welch's method performs much better than the traditional method as shown in Figures 6 and 7. These figures show waveform interferometry of monthly substacks generated by using Welch's method and traditional method, respectively. Figure 6 shows higher coherence between traces than Figure 7, meaning more stable reconstruction of Green's function. Since many of the dv/v measuring methods rely on similarities between waveforms, higher coherency over time is ideal for obtaining stable and accurate dv/v results. Furthermore, higher coherency between waveforms also increases the number of good measurements that can be used for further analysis and increase the overall quality of our study. Additionally, Seats et al. (2002) showed that better time averaging leads to better spatial averaging, which is the primary challenge in retrieving Green's functions. Because the Welch's method uses hourly traces, it increases the number of averaging by an order of magnitude compared to the conventional method using 24-hour traces. Hence the Welch's method is more favorable than the traditional method due to its better performance in reconstructing stable substack waveforms. The differences between the two methods can be explained by the number of correlation functions averaged in the process. More correlation functions used result in better construction of a fully diffusive wavefield, and thus closer to the wavefield settings required by the math calculation. Figure 8 compares waveforms using one-bit normalization or not. We can see two waveforms are very similar, but the one without time-domain normalization has high SNR. So, the effect of time-domain normalization can be achieved in better quality by averaging enough number of correlation functions.

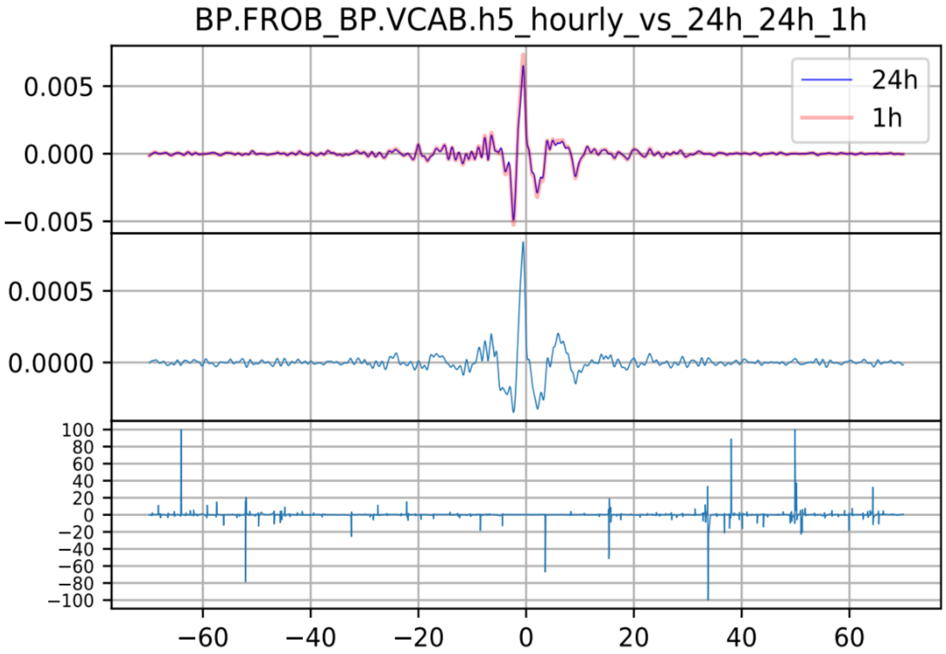

BP.FROB_BP.VCAB.h5_hourly_vs_24h_24h_1h

Figure 5. Upper panel: comparison between reference waveforms stacked by using daily and hourly traces. Middle panel: amplitude difference between the two waveforms. Lower panel: percentage difference between the two waveforms.

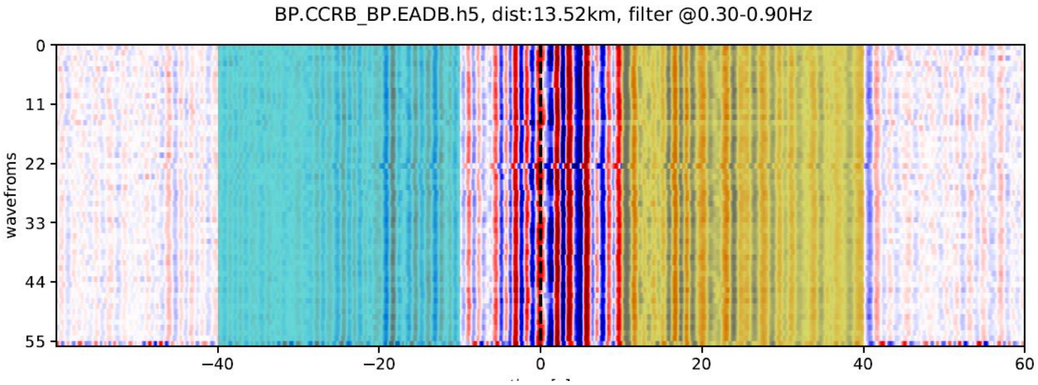

BP.CCRB_BP.EADB.h5, dist:13.52km, filter @0.30-0.90Hz

Figure 6. Substack waveform interferometry stacked from hourly traces. The vertical axis is waveform numbers, the horizontal axis is length in seconds.

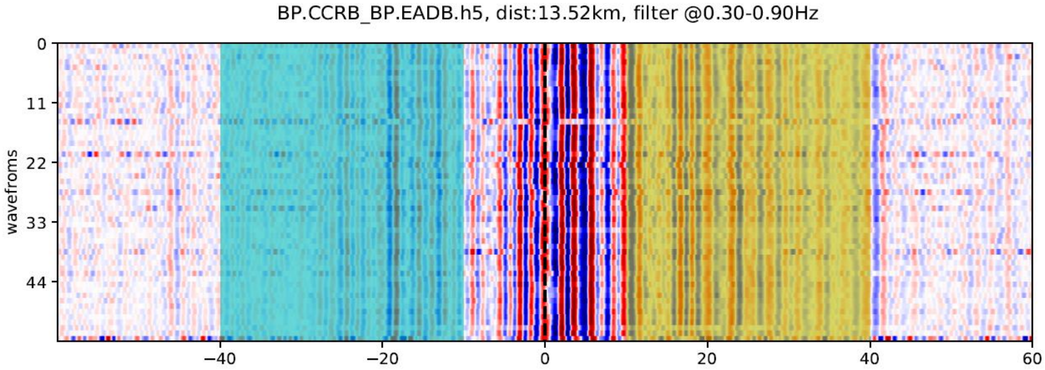


Figure 7. Substack waveform interferometry stacked from daily traces. The vertical axis is waveform numbers, the horizontal axis is length in seconds.

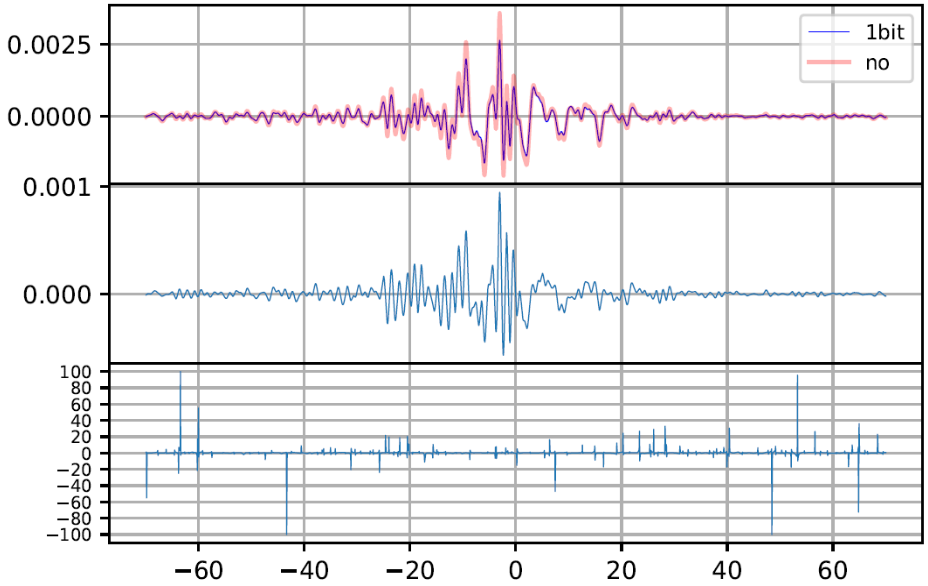


Figure 8. The upper panel shows comparison between reference waveforms of using 1-bit normalization and not. The middle panel shows the amplitude differences between the waveforms. The lower panel shows relative changes in percentage.

The choice of stacking strategy is critical in preserving the completeness of information contained in the seismic data, particularly when focusing on the coda parts of the waveforms. Similar to stacking over substacks for reference waveform using different methods such as linear stacking or phase-weighted stacking, the stacking of hourly traces also has various choices. Figure 9 plots waveforms stacked by using linear and phase-weighted methods. Phase-weighted stacked waveform has higher SNR than the other one, but it loses information on the coda parts. Therefore, it is not ideal for analysis done using coda waves. Linearly stacked waveform preserves information on the coda parts even though it has a lower SNR. Other stacking strategies have the same issue for over simplifying coda part if too many traces are stacked. Since the monthly waveforms will be stacked again in subsequent step, more original information should be preserved at this point. Otherwise, the measurement of dv/v will not be robust.

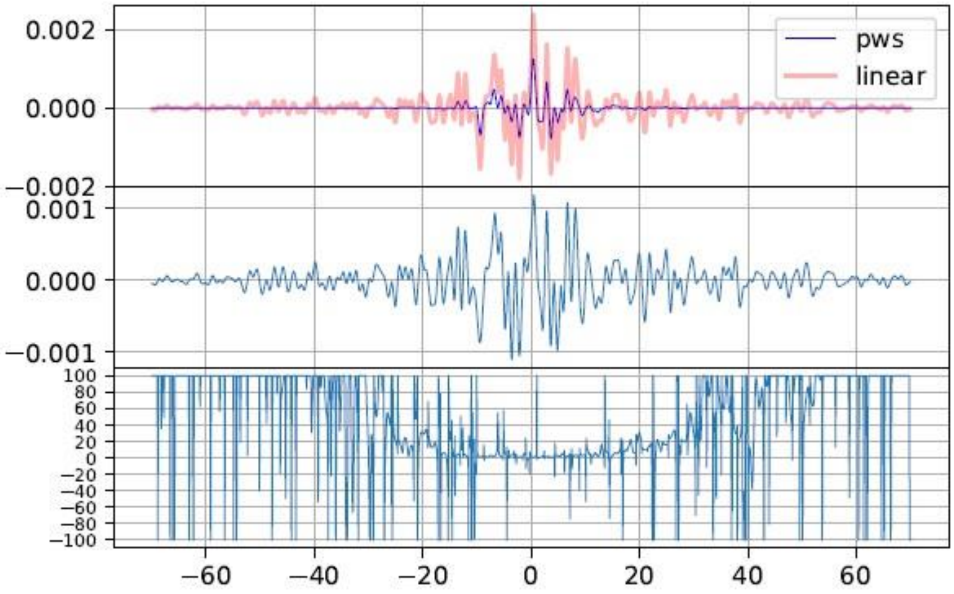


Figure 9. The upper panel shows difference between waveforms stacked using pws and linear methods over hourly traces. The middle panel shows amplitude difference between waveforms. The lower panel shows relative changes in percentage.

An additional capability of this script is implementing filtering using the single-to-noise ratio (SNR). In the process of seismic data analysis, the accurate determination of the central point, which is the highest amplitude of the reference waveform, for all traces is imperative. Once the center has been established, the script is run again to compute the SNR for individual traces between user-defined signal parts and noise parts. For the purposes of this study, the signal part is defined as a window spanning 20 seconds, centered at the maximum point. On the other hand, the noise parts consist of two windows that are 40 seconds wide, located on both sides of the signal part. The SNR is quantified by calculating the ratio of the root-mean-square of the absolute amplitudes in the signal and noise parts.

Choosing the appropriate filtering threshold based on SNR is arbitrary. If the threshold is set too high, there's a risk of discarding too many traces, leaving insufficient data for the stacking process. Conversely, if it's set too low, the filtering process loses its effectiveness and fails to sufficiently differentiate between signal and noise. The objective of the SNR filtering process is to remove traces that are dominated by low SNR and incoherent noise. This step enhances the coherence of the reconstructed substacks, ultimately leading to more reliable measurements. Figures 10A and 10B demonstrate the traces before and after filtering, respectively. After applying the filter, the signal parts in Figure 10B are qualitatively more coherent and the code parts have less variations in amplitude than the other plot.

To calculate cross-correlation between station pairs in the frequency domain, we employ a wavelet transform to convert the results into a time-series format. These time-series are essentially correlation functions which are subsequently stacked and preserved for the next stage of analysis. In this study, we use a duration of 30 days as the length for substacks. This means that each h5 file comprises one monthly substack for each station pair.

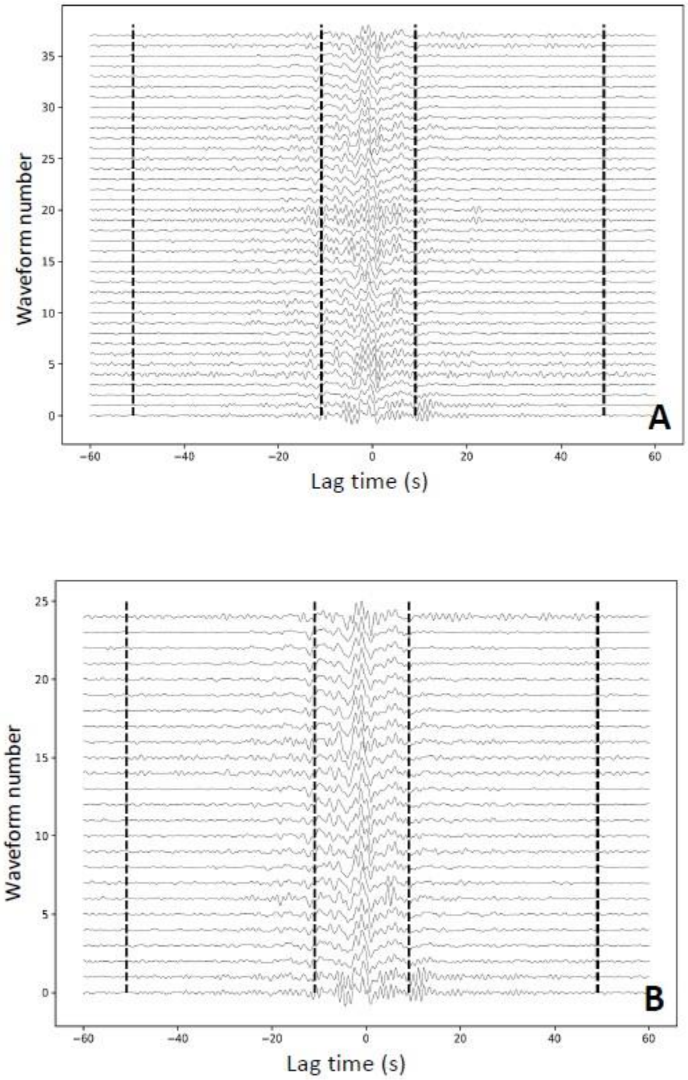


Figure 10. (A) Traces before SNR filtering. The vertical dash lines mark signal and noise parts. (B) Traces after applying SNR filter with threshold of 3.0. The vertical dash lines mark signal and noise parts

3.1.5. Stacking

The script S2 is responsible for executing the waveform stacking procedure, offering a range of stacking methods and their associated parameters to choose from. In this section, our primary focus will be on discussing the linear stacking method, which was selected for use in this study, along with other parameter options.

In this study, we took a different approach to reconstruct the reference waveform compared to most previous studies. Rather than stacking all monthly substacks from the cross-correlation stage, we selectively stacked only parts of the monthly substacks to produce the reference waveform. The reference waveform in this study is expected to provide a representation of wave propagation through the medium under normal, undisturbed conditions between two seismic stations. Aiming at accurately monitoring velocity variations over time, we need to establish a benchmark reference waveform that accurately resolve the properties of undisturbed rocks. Therefore, we determined that it is crucial to exclude from the stacking process any substacks from periods in proximity to major earthquakes. During these periods, the rock properties undergo significant changes, which lead to the contamination of random signals into the waveforms. These so-called “contaminated” substacks can compromise the coherence of the reference waveform compared to other substacks, consequently leading to less reliable and fluctuating dv/v measurements. In consideration of this, we carefully selected the time frame of July 14, 2005, to October 2, 2007, as our stacking period. This choice was informed by estimations of fault recovery speed (i.e., when dv/v returns to pre-earthquake levels) derived from previous research (Brenguier et al., 2008; Wu et al., 2014; Okubo & Denolle, 2020).

To illustrate the difference our approach makes, Figure 11 provides a comparison between reference waveforms generated using our selective method in the time domain, versus those produced using the non-selective methods commonly employed in other studies.

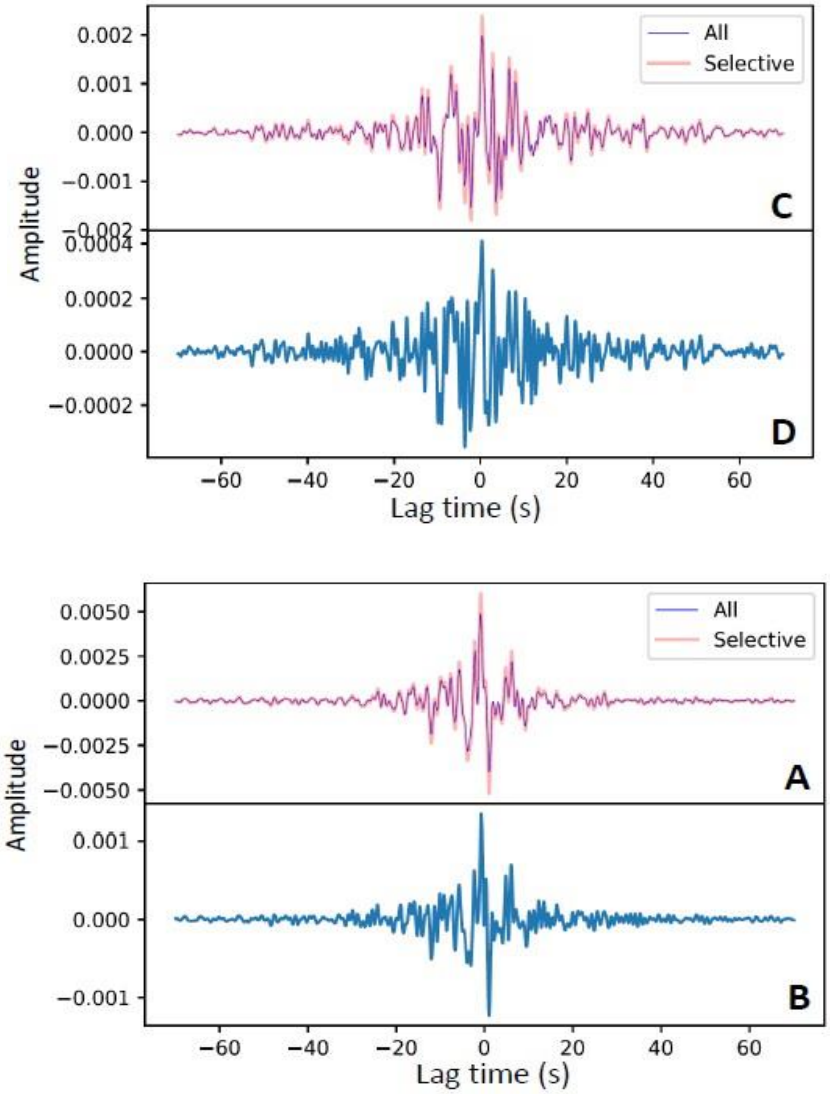


Figure 11 (A) & (C): Reference waveform between different station pairs stacked by using all substacks (blue) and only selective periods (red). (B) & (D): Amplitude difference between the two reference waveforms in each panel A and C.

There is clear increase in amplitude in both reference waveforms, especially in the direct wave parts. And the changes in amplitudes are concentrated in high frequency bands as shown in panels B and D, which is reasonable because strong motion involves changes in higher frequency bands. Therefore, by using period-selective stacking, some contamination caused by earthquakes can be removed and the resulting reference waveform becomes more coherent and robust.

In our pursuit of monitoring temporal phase changes across waveforms, maintaining coherence between the reference and current correlation functions is essential. This is particularly true for the coda parts of the waveforms. As such, selecting an appropriate stacking method that can rapidly converge towards the empirical Green's function with minimal distortion in the power spectrum, becomes a critical step in the entire process. In this study we used several stacking methods to stack the same set of short-time (1 hour) correlation functions and compared the differences between resulting waveforms. Figure 12 illustrates a selection of the short-time correlation functions we used to construct the reference waveform. While the signal parts appear noisy, there is still a discernible coherent peak around zero time. The noise window also show incoherency across waveforms but the amplitude variance is less clear. Figure 13 displays the reference waveforms generated using linear stacking, phase-weighted stacking, nroot stacking, and robust stacking methods. All these methods exhibit similar features between -25s and 25s, including a peak amplitude around zero. However, some discrepancies become apparent in the post-arrival parts of the waveforms. Both the linear and robust methods manage to preserve some amplitude variations and possibly some phase information in the coda parts, as depicted in the traces in Figure 12. The other two methods, on the other hand, do not exhibit the same level of detail. These differences can be attributed to the divergent focuses of these methods. Some methods accentuate the signal part, while others aim to average the traces. As demonstrated in Figure 13, from the standpoint of monitoring transient velocity changes which requires preserving phase information and maintaining coherence, the linear and robust stacking methods are good choices.

Yang et al. (2022) systematically examined performance of multiple stacking strategies from convergence, dispersion stability, signal-to-noise ratio, similarity index and other aspects. The authors proposed that the optimal stacking methods for both velocity tomography and seismic monitoring are linear and robust stacking for processed datasets. Yang et al. (2022) also pointed out that phase-weighted-stacking method failed the test on obtaining velocity variation because of the very low

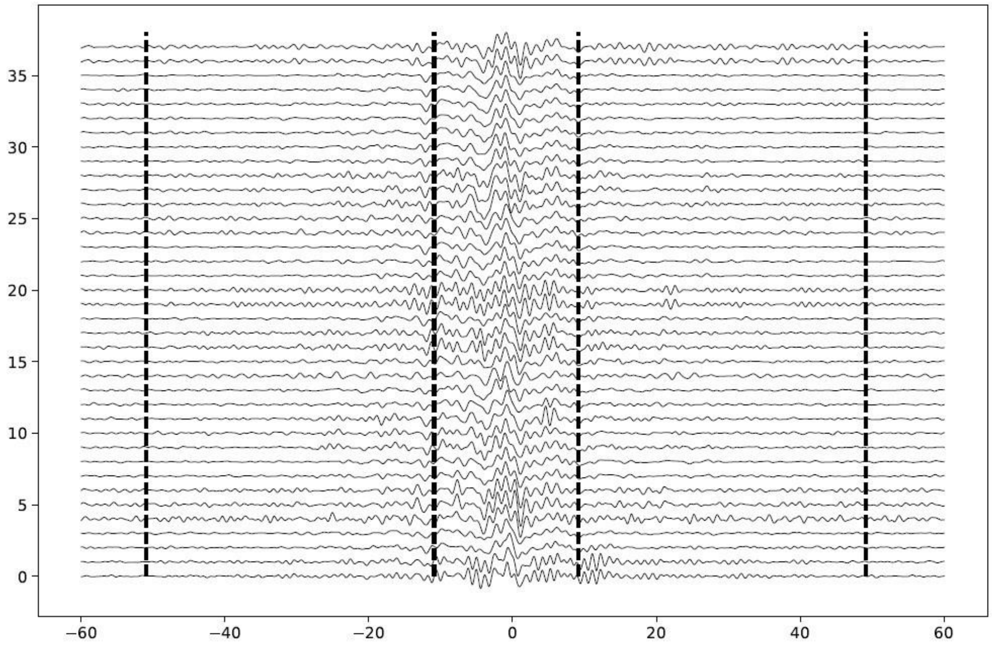


Figure 12. Some traces of correlation functions of station pair EADB-FROB. The traces are normalized in each row. Dashed lines mark windows for signal and noise parts.

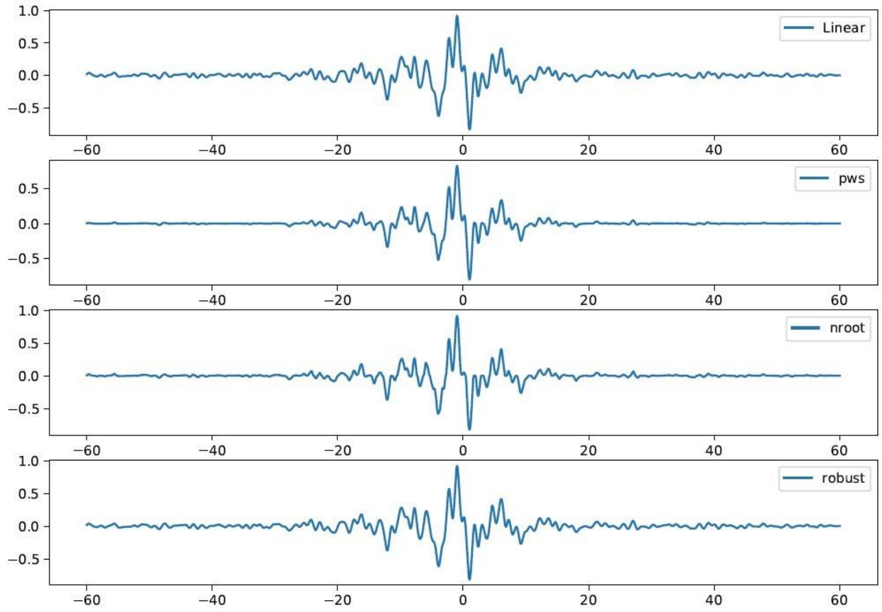


Figure 13 Reference waveform between station EADB-FROB stacked by different methods

correlation coefficient between current and reference waveform as a result of over distortion and simplifying of coda waves in the long-time waveform. Following previous research, we also use linear stacking method in this case from the consideration of both computation expense and reliability.

The last crucial parameter in this step is the length of the substacking window, which effectively is the temporal resolution of the dv/v measurements. Previous seismic monitoring studies have used window lengths ranging from as short as 5 days to as long as 50 days (Taira et al., 2015). A shorter window length results in a higher temporal resolution, but it also has its drawbacks. Namely, the use of a narrow window can result in insufficient averaging of correlation traces, hindering the convergence to a robust Green’s function and subsequently leading to a low correlation coefficient with the reference waveform. This can cause the velocity variations measured using these substacks to fluctuate significantly, sometimes to the point where it becomes too noisy to extract any useful information. A longer substacking window, on the contrary, can provide stable substack waveforms and smoother dv/v measurements. However, despite the smooth velocity variation curve it produces, this approach can potentially overshadow or obscure some weaker and transient signals. Hence, selecting an appropriate substacking window length should be selected based on the specific goals of the study.

In this study, we adopted a 30-day substacking window, following the precedent set by Brenguier et al. (2006) and Wu et al. (2016), which worked with the same datasets and also focused on seismic monitoring. In addition to window length, several pioneering studies have employed moving windows for substacking. For instance, Brenguier et al. (2006) and Wu et al. (2016) shifted their window by one day for each substacking calculation, while Rivet et al., (2014) used a 10-day step for each substacking window. In our study, we do not overlap the substacking windows as the temporal resolution remains unaffected by the size of the overlap. Theoretically, overlapping tends to smooth the dv/v measurement curve but does not alter the overall trend. Therefore, any strong signals should still be discernible even without overlapping.

3.1.6. The dv/v Measurement

The final step in this workflow is the *dv/v* measurement, which is calculated using the script 'measure_dvv.py'. Each monthly substack is compared with the reference waveform to determine time shifts as a function of time lags (*dt/t*), and then converted into velocity variations (*dv/v*). A variety of methods have been developed to measure seismic travel-time changes across time, frequency, and wavelet domains. Each method has its strengths and limitations (Mao et al., 2020; Yuan et al., 2021). However, the results across these methods are generally comparable. While more advanced and complex methods have been introduced, such as the continuous wavelet transform (CWT) and wavelet transform dynamic time warping (WTDTW), we favor the MWCS method, primarily due to its proven effectiveness with the same dataset that we are using (Brenguier et al. 2008; Wu et al., 2016).

In the MWCS method, time shifts are calculated in the frequency domain based on the assumption that phase shifts increase linearly with frequency. Consequently, the slope of the linear regression of phase changes against frequencies between the current and reference waveforms within a moving window corresponds to the time shift (dt) centered at a specific time lag (t). Another important assumption is homogenous medium along the path. This assumption results in the expectation that the time shift will increase proportionally with distance or time. By obtaining time shifts at different time lags and performing another linear regression, the resulting slope is the negative of velocity change: $-dt/t = dv/v$. The accuracy and reliability of MWCS measurements is therefore susceptible to some parameters such as the similarity between waveforms and the length of the window utilized in the analysis. We discuss the effects of these two parameters in the following context.

One common challenge encountered in MWCS measurements is cycle skipping, which refers to the mismatching of phases over multiple wave cycles between two waveforms, leading to inaccurate time shift approximations. Therefore, selecting an appropriate window size for comparison is vital. If the size is too large or too small (greater than 2.5 times the longest wavelength or smaller than 1 times the longest wavelength), the risk of cycle skipping increases. Meanwhile, if the size is too small, there may be insufficient information to perform a reliable linear regression.

Figure 14 illustrates the effect of window length on the final *dv/v* measurements. As the window length increases from 0.6 to 2.0 times the longest wavelength within the frequency range (which spans 0.3 to 0.8 Hz in this instance), the *dv/v* measurements become more dependable, indicated by the narrowing error bars. Furthermore, the velocity drop associated with the occurrence of the Parkfield earthquake becomes more evident relative to the size of the error bars. Based on these results, we decided to use a window size equal to twice the longest wavelength in the specific frequency range for this study. Figure 14 also demonstrates that the number of windows used to perform linear regression significantly influences the results. The more data points that are used, the more reliable the regression becomes. To increase the number of windows, one should avoid overly large windows and enhance the overlap between windows. In this study, we adopted a 50 percent overlap to ensure robust linear regression and to accommodate lower quality windows that are filtered out.

The MWCS method calculates the cross-spectrum between the reference and current Green's functions to detect minor phase shifts, necessitating a certain level of similarity between the waveforms. We utilize the Pearson correlation coefficient to measure the linear relationship between two sets of variables, which in our case can be considered as a measure of similarity. The Pearson correlation coefficient ranges from -1 to 1, with 0 indicating no correlation, 1 for perfect positive correlation and -1 for perfect negative correlation. However, in real-world scenarios, we regard a cc value higher than 0.6 as indicative of good correlation due to the inherent complexity of seismic waveforms. This threshold has also been employed in other studies as a lower limit to discard unreliable measurements (Hobiger et al., 2012; Rivet et al., 2014).

We also conducted tests to ascertain the effects of cc values on dv/v measurements, as displayed in Figure 15. The correlation coefficient is calculated between the coda parts of the current and reference waveforms, rather than between each moving window. As shown in Figure 15, a lower cc threshold provides a better resolution of the velocity change during an earthquake, and the dv/v measurements feature smaller error bars. Additionally, the number of averaged dv/v measurements increases with a lower cc threshold, offering a more robust approximation of the velocity variations. Given the complex nature of seismic waveforms, even minor stretching can induce a significant change in the correlation coefficient. Therefore, a high threshold could potentially filter out some velocity change signals. While Brenguier et al. (2008) and Wu et al. (2016) did not include this step in their analyses, we incorporate it into our study to enhance the reliability of our results.

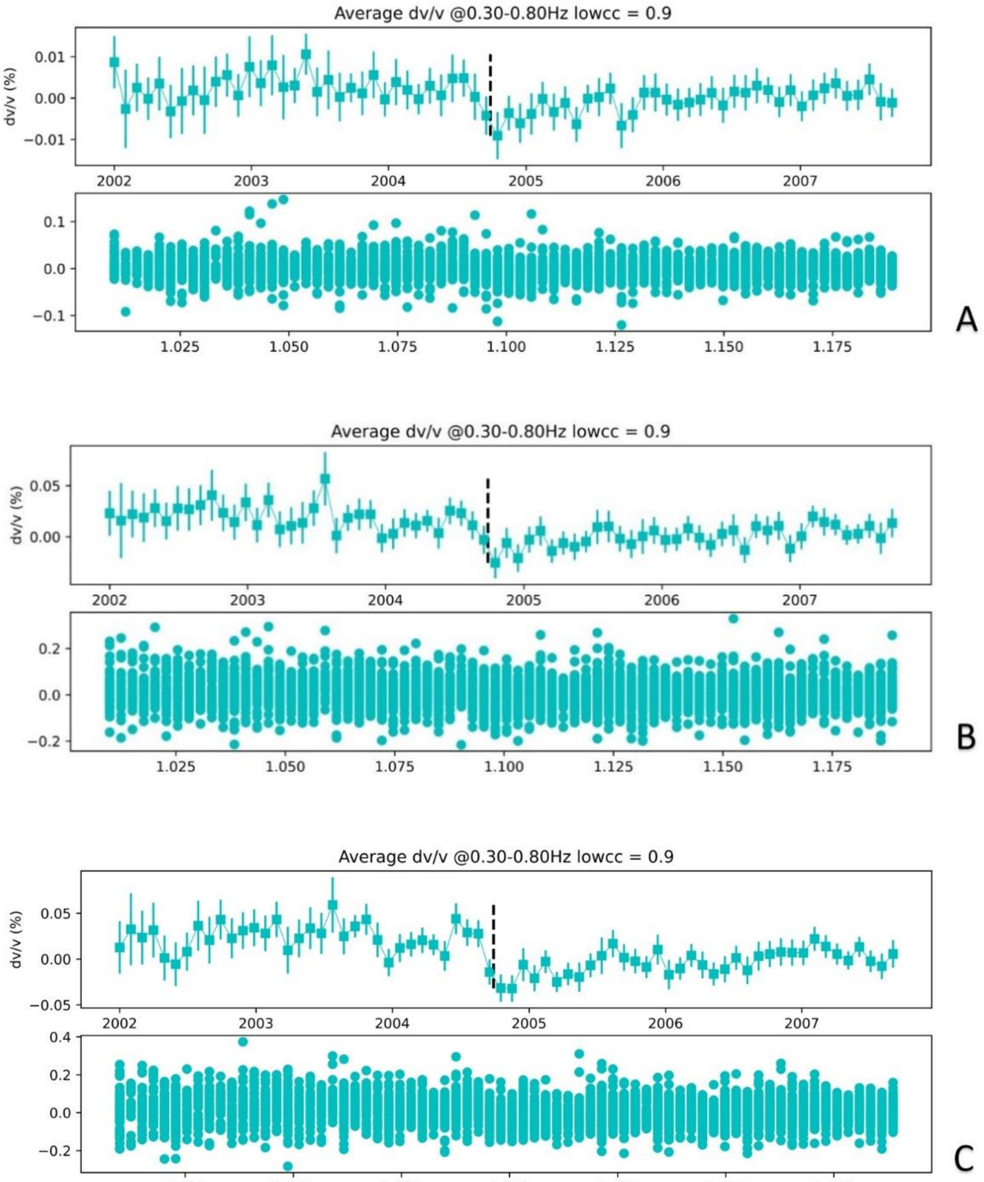


Figure 14. From (A) to (C) the upper panels show averaged dv/v measurements with the window lengths increase from 0.6x to 1.0x to 2.0x the longest wavelength in the frequency range (0.3 to 0.8 Hz), respectively. The lower panels show values used to average at that time.

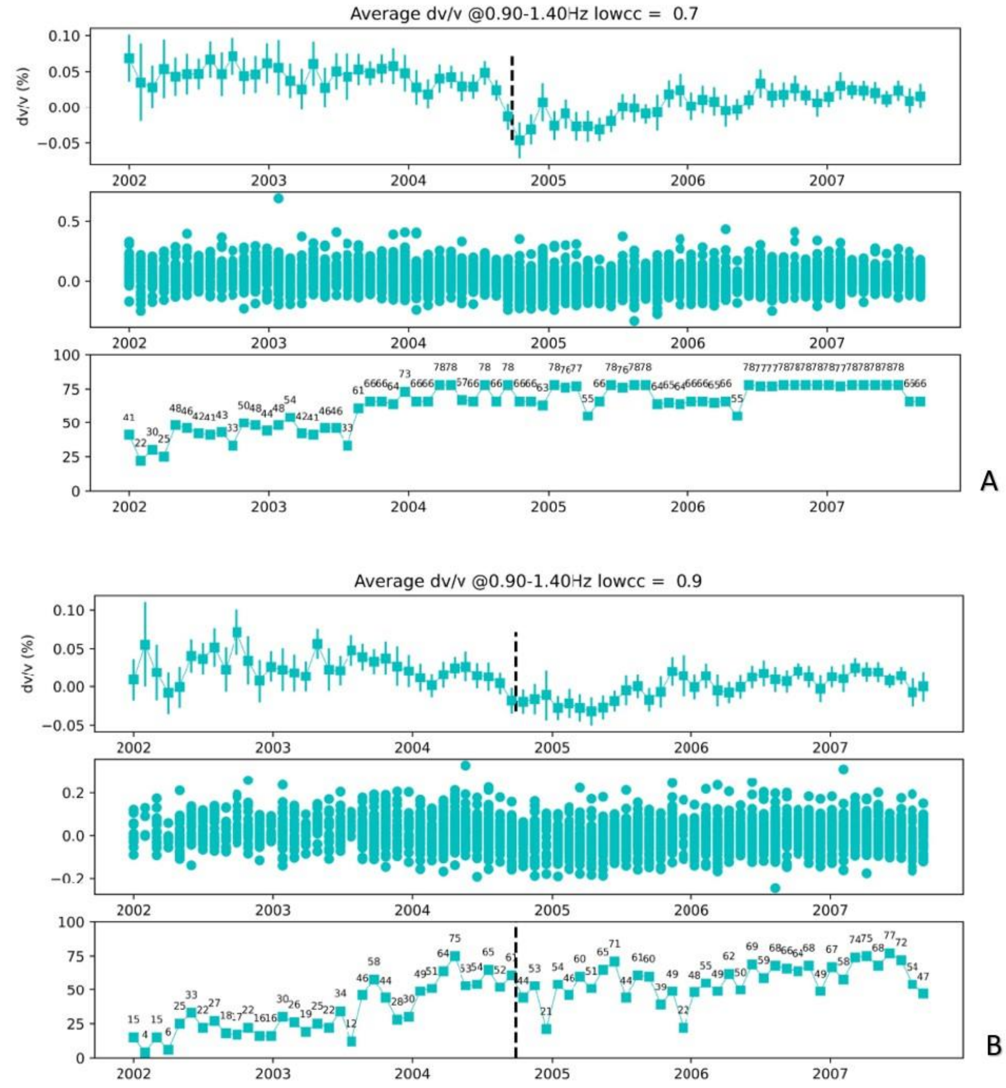


Figure 15(A) upper panel shows averaged dv/v measurements using cc threshold of 0.7; middle panel shows distribution of dv/v values used to average on that timestamp; lower panel shows the number of dv/v values used to average on that timestamp. Figure 15(B) shows the same plots using a cc threshold of 0.9

3.2. GPS and Strong Motion Data Processing

We calculate the strain components ϵ_{xx}, ϵ_{yy} and ϵ_{xy} using relative displacements between each pair of GPS stations with distances smaller than 9 km and then derive the second strain invariant using $I_2 = \frac{1}{2}(\epsilon_{xx}\epsilon_{yy} - \epsilon_{xy}\epsilon_{xy})$. The metric serves as a simple proxy for the amplitude of strain changes for the coseismic and specified postseismic periods. We focus on the decimal logarithm of the square root of the second invariant, $\log_{10}(\sqrt{I_2})$ to assess the spatial variations of the strain changes.

The dynamic strains during the earthquake are estimated from strong-motion records, with the peak ground velocities (PGV) divided by the assumed S-wave phase speed (2.7 km/s) (e.g., Hill et al., 1993; Taira et al., 2015). We compare the spatial variations of maximum logarithmic values of dynamic strains in Section 4.

4. Results

4.1. The dv/v Results

After implementing the ambient noise interferometry workflow for each station pair, we proceeded to generate dv/v measurements for each of the five distinct frequency bands. To resolve the subtle velocity variations, we examine the averaged dv/v over all station pairs, (Figure 16), and individual station pair dv/v values (Figures 17–21). Each substack is averaged over a period of 30 days with no overlap, providing us with a total of 70 data points for each frequency band. To accurately evaluate the coseismic velocity changes (Δ_0), we model the dv/v time series using a five-parameter function, following Hobiger et al. (2012):

$$f(t) = A + B * \ln\left(\frac{t-t_0}{\tau}\right) * H(t-t_0) + E * \cos(\omega t + \varphi) \tag{5}$$

where t is the date of measurement, t_0 is the date of the Parkfield earthquake, and $H(t-t_0)$ is the Heaviside function defined by $H(t-t_0) = 0$ for $t < t_0$ and $H(t-t_0) = 1$ for $t > t_0$. The five parameters in this equation are A, B, τ, E and φ, where A is a constant offset; B is a coefficient that describes the coseismic velocity change; τ is the postseismic recovery time in years; E and φ are the amplitude and phase shift for the seasonal variation model with a fixed period, ω, of 2π/year. We note that the model behaves like a simple cosine function with some offset before the Parkfield earthquake and incorporates a logarithmic function after the event to capture the postseismic recovery process. Since the logarithmic function $B * \ln\left(\frac{t-t_0}{\tau}\right)$ approaches negative infinity for very small $(t - t_0)$ values, it describes an infinite coseismic velocity drop, which is unrealistic in real scenarios. Therefore, we need to use a different definition to estimate the coseismic velocity jump (Δ_0). As we use an averaging window, d, with a length of 30 days for current waveforms, the coseismic velocity change (Δ_0) can be estimated using the difference between the beginning and end of the averaging window centered on the date of earthquake. Following Hobiger et al. (2012), we use the following equation to calculate Δ_0:

$$\Delta_0 = B * \ln(\frac{d}{2\tau}) \tag{6}$$

where d is the substack smoothing window length in days, τ is the number of postcoseismic recovery time in days and B is the best fitting parameter for each station in Eq. 5. Because of the usage of averaging window, a factor of 2 is added to the denominator to calculate the Δ_0 value at the center point of the smoothing window.

For each station pair, we execute the grid-search algorithm on the five parameters to minimize the root-mean-square error (RMSE) of the fitting curve. Table 4 lists the ranges used in grid-search for the five parameters in Eq. 5. The ranges are estimated from reasonable extreme values and based on previous studies (We et al., 2016; Hobiger et al., 2012; Taira et al., 2015; Okubo & Denolle, 2020).

The optimal combination of these parameters, which yields the most accurate fitting curve, is subsequently displayed in each dv/v curve figure. Even though the final best-fitting curves have the lowest RMSE, it is still insufficient to reflect the realistic and correct coseismic velocity variations because the calculated $\Delta_0$value may be only a random number within the error range that happen to show the expect value, and thus cannot be considered as clear and robust unless it is statistically valid. So, in order to make sure our coseismic velocity change measurements are solid and meaningful, i.e., not within the uncertainty range of a predicted value, we need to introduce an additional quality metric to select the Δ_0 values. We also ascertain the robustness of the fitting curves by calculating the ratio between the RMSE and Δ_0 values. If a fitting curve has a ratio that is less than 30%, it is deemed to be within the acceptable range. In this way, we can focus on calculated Δ_0 values that are at least 3.3 times higher than the model fitting error. Thus, only fitting curves that meet this criterion are selected and presented for further discussion and demonstration.

Constant offset (A)	0 – 0.2%
Coseismic Velocity Change (B)	0 – 0.3%
Recovery Time (τ)	0 – 100 years
Seasonal Variations Amplitude (E)	0 – 0.1%
Seasonal Variations Phase (φ)	0 – 150 days

Table 4. List of ranges for the five parameters in performing grid-search.

As indicated in Figure 16, we observe an increase in the amplitude of coseismic dv/v drop, from 0.06% at 0.2-0.6 Hz to 0.13% at 0.9-1.4 Hz. Concurrently, we witness a decrease in the RMSE error of the fitting curve with the increment of the frequency band. This trend suggests a better characterization of the distribution of data points with higher frequencies. Examining all averaged dv/v measurements, we observe a fluctuation in the values prior to the earthquake. However, these values exhibit relatively less variability in the period following the seismic event. We hypothesize that the observed characteristic is attributable to a relatively small number of values utilized for averaging before Sept. 2004 due to data gap and relatively low quality of recorded data. However, in general the averaged dv/v measurements present consistent patterns of the coseismic velocity drop and the subsequent postseismic recovery process. Moreover, the trend of an increasing dv/v at higher frequency bands, and the magnitude of velocity change, are consistent with the findings of previous research (Wu et al., 2016).

The single station dv/v measurements across different frequency bands, as illustrated in Figures 17-21 and Appendix Figures 1-4, reveal more features compared to the averaged dv/v plots shown in Figure 16. A notable feature, increasingly prominent with higher frequency bands, is the larger quantity of good fitting curves. Based on a previous study (Campillo and Paul, 2003), this change in better averaging is indicative of improving quality dv/v measurements at higher frequencies. The cause behind this phenomenon is not entirely clear, but we suspect it may be related to the comparatively smaller reactions to perturbations at lower frequencies, resulting in low amplitudes that are harder to detect in both time and frequency domains.

Figures 17-21 also demonstrate a pattern that errors are lower at high frequencies and higher at low frequencies. Large error represents more randomly distributed data points, but we expect them to arrange in certain patterns. Therefore, a small error is also critical for accurately estimating coseismic velocity changes using fitting curves. In contrast with averaged dv/v measurements, single-station-pair measurements exhibit greater fluctuation across all frequency bands. This is mainly caused by insufficient averaging in the single-station-pair measurements. In the averaging process, incoherent values will be cancelled out and the coherent part will be emphasized. But because single station pair dv/v measurements do not go through such procedure, the larger fluctuation is expected.

While all single station pairs exhibit patterns of velocity change to a certain degree, the quality of these patterns is somewhat hindered by the fluctuation of data points. Another issue contributing to this problem is the incoherency between the current and reference waveforms, especially in the coda portions. Some monthly waveforms possess a higher overall amplitude, thus having more weight when combined with other waveforms. This results in a more significant contribution to the reference Green's functions, which in turn reduces the coherency with other current waveforms. We have also observed time shifts in some waveforms, which could potentially result in abnormal dv/v measurements. However, we currently lack a method to validate this assumption. Despite these challenges, all Figures from 17 to 21 exhibit similar patterns as seen in Figure 16, namely, that coseismic velocity changes increase and errors decrease with increasing frequencies. Notably, some station pairs even display a coseismic velocity change as high as 0.27%.

To illustrate the spatial pattern at different frequency bands, we map out coseismic velocity change values for all station-pairs (Figure 22-26). The selected station pairs are connected with lines that are color-coded based on coseismic dv/v values. To facilitate comparisons across frequencies, figures for all frequency bands use the same color bar and range. These map views reiterate the overall increase in coseismic dv/v at higher frequency bands and highlight some intriguing spatial variations. For instance, stations situated on the eastern side of the San Andreas Fault register higher velocity changes compared to other stations operating at the same frequency band. Conversely, some stations located in the southwest part of the array demonstrate higher velocity variations only at higher frequency bands.

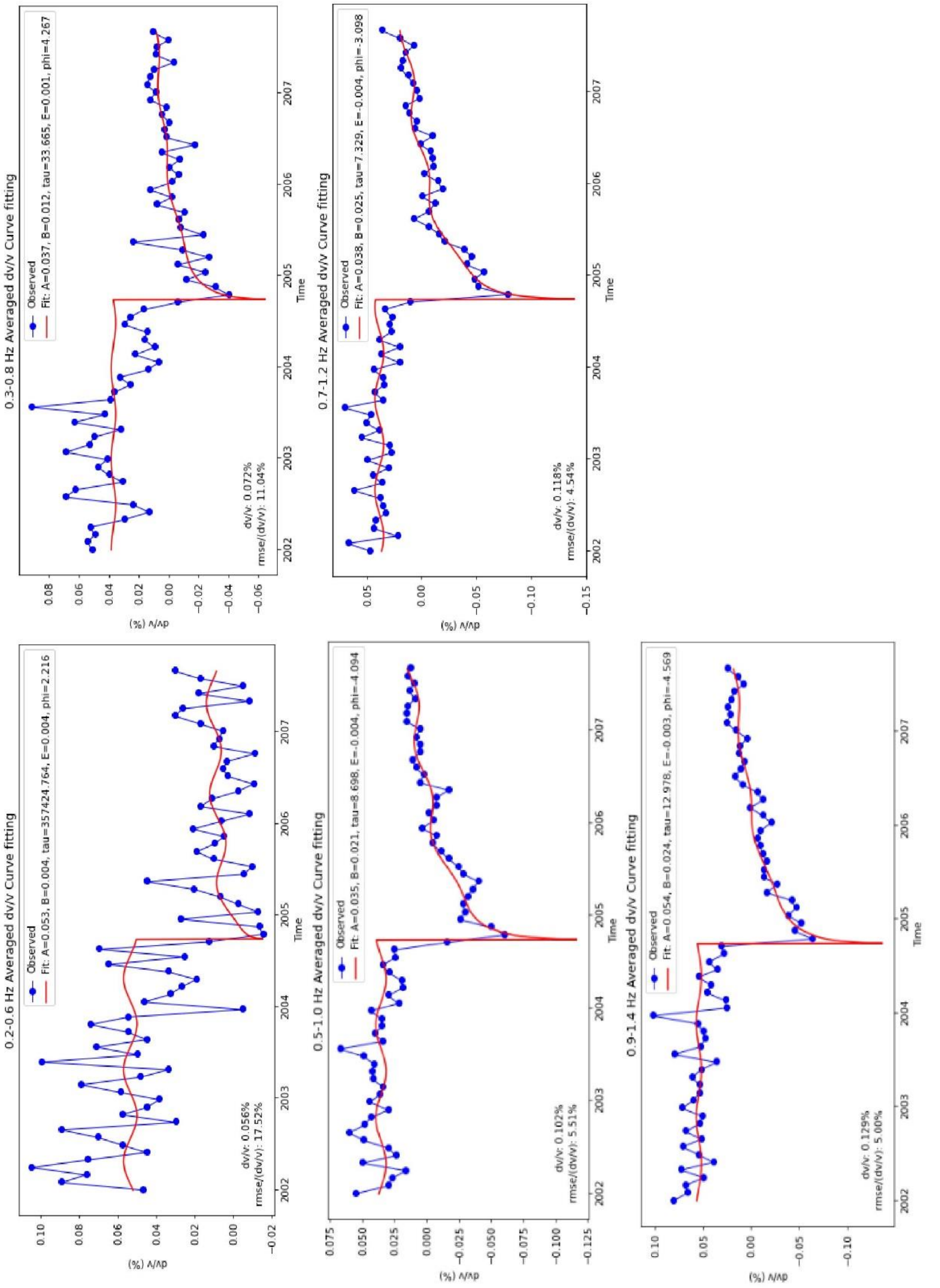


Figure 16. Averaged dv/v measurements at 0.2-0.6, 0.3-0.8, 0.5-1.0, 0.7-1.2, 0.9-1.4 Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.

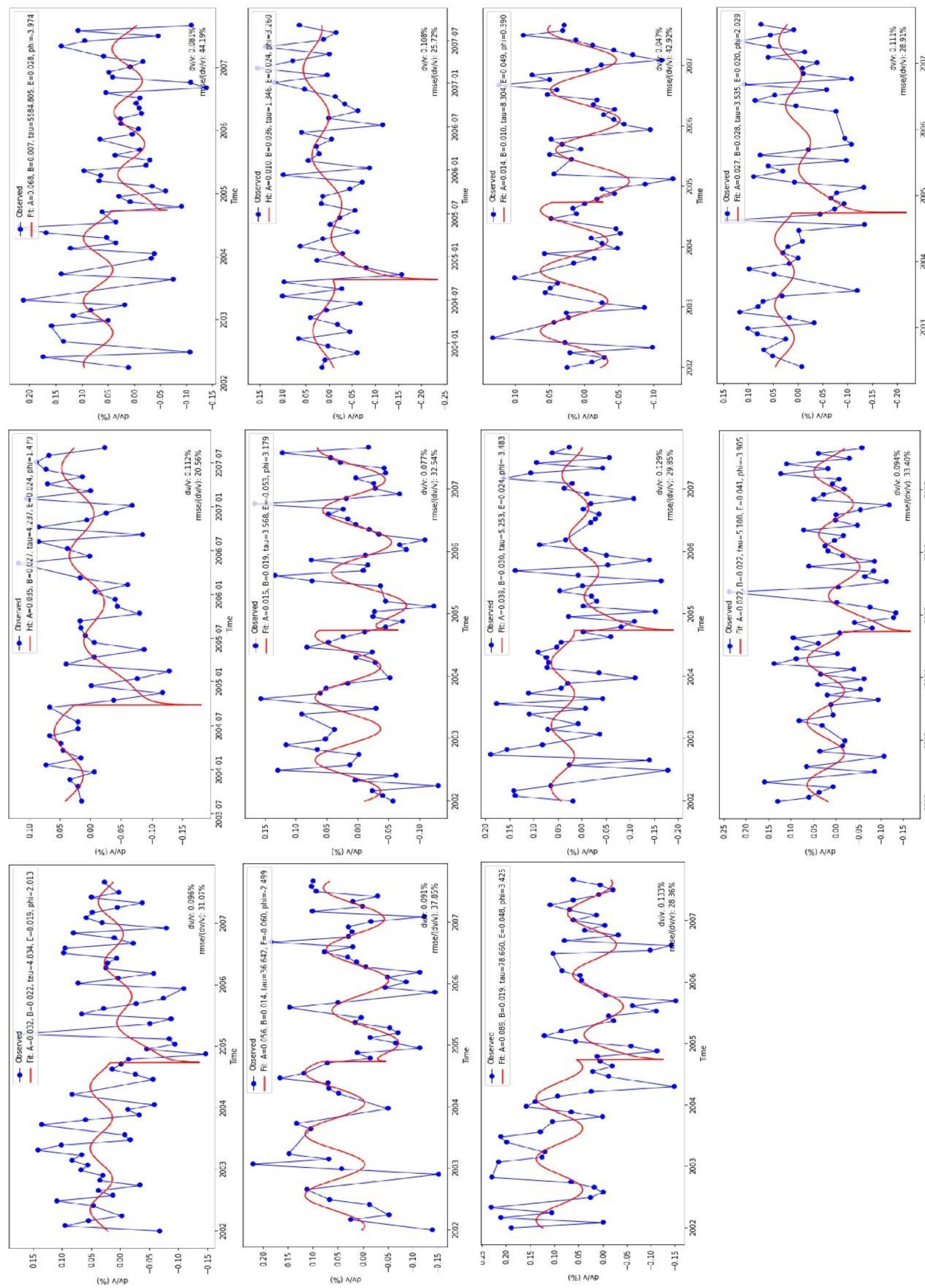


Figure 17. Selected single station pair dv/v measurements at frequency band ***0.2-0.6*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

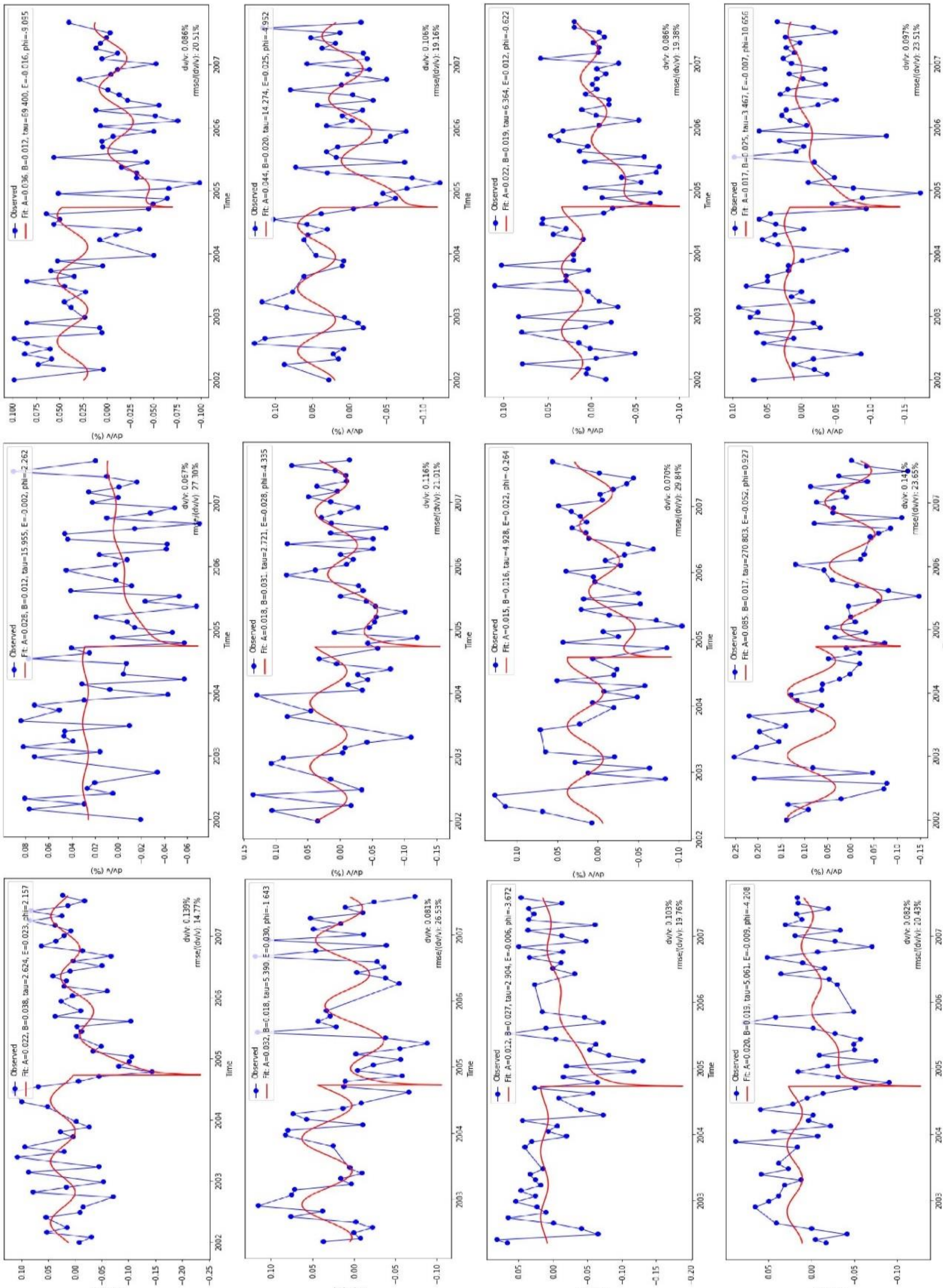


*Figure 18. Selected single station pair dv/v measurements at frequency band **0.3-0.8** Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

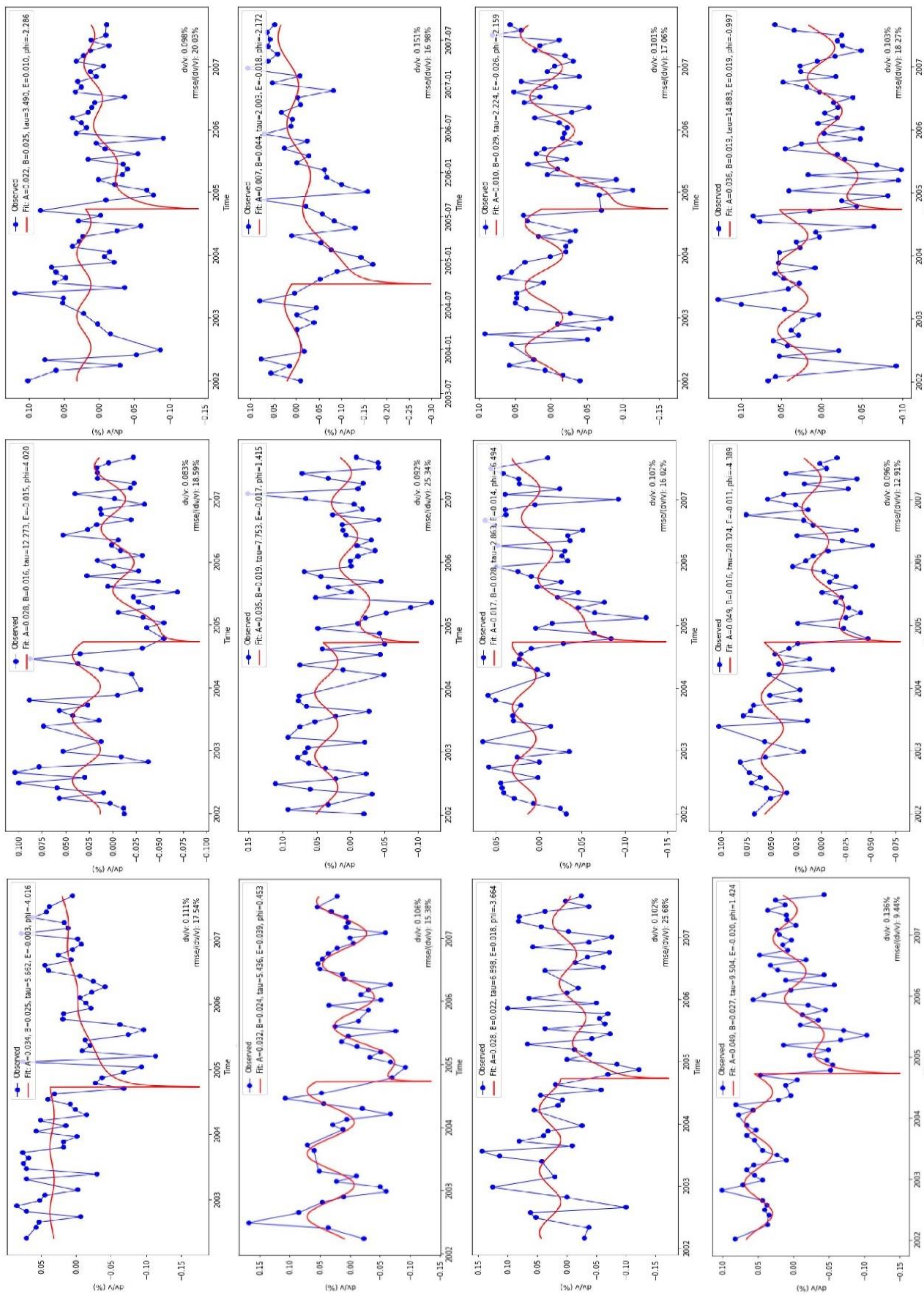


Figure 19. Selected single station pair dv/v measurements at frequency band ***0.5-1.0*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

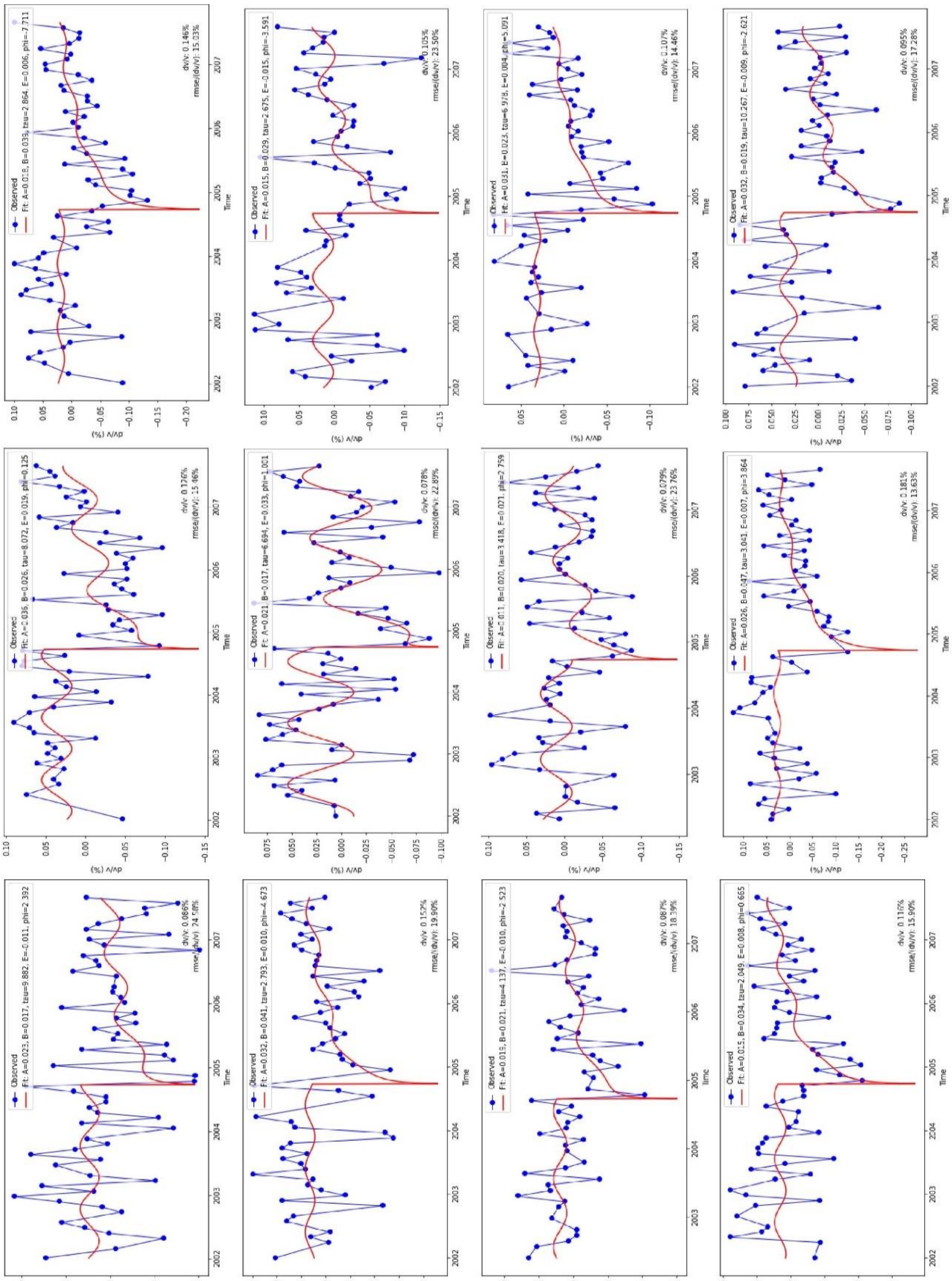


Figure 20. Selected single station pair dv/v measurements at frequency band ***0.7-1.2*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

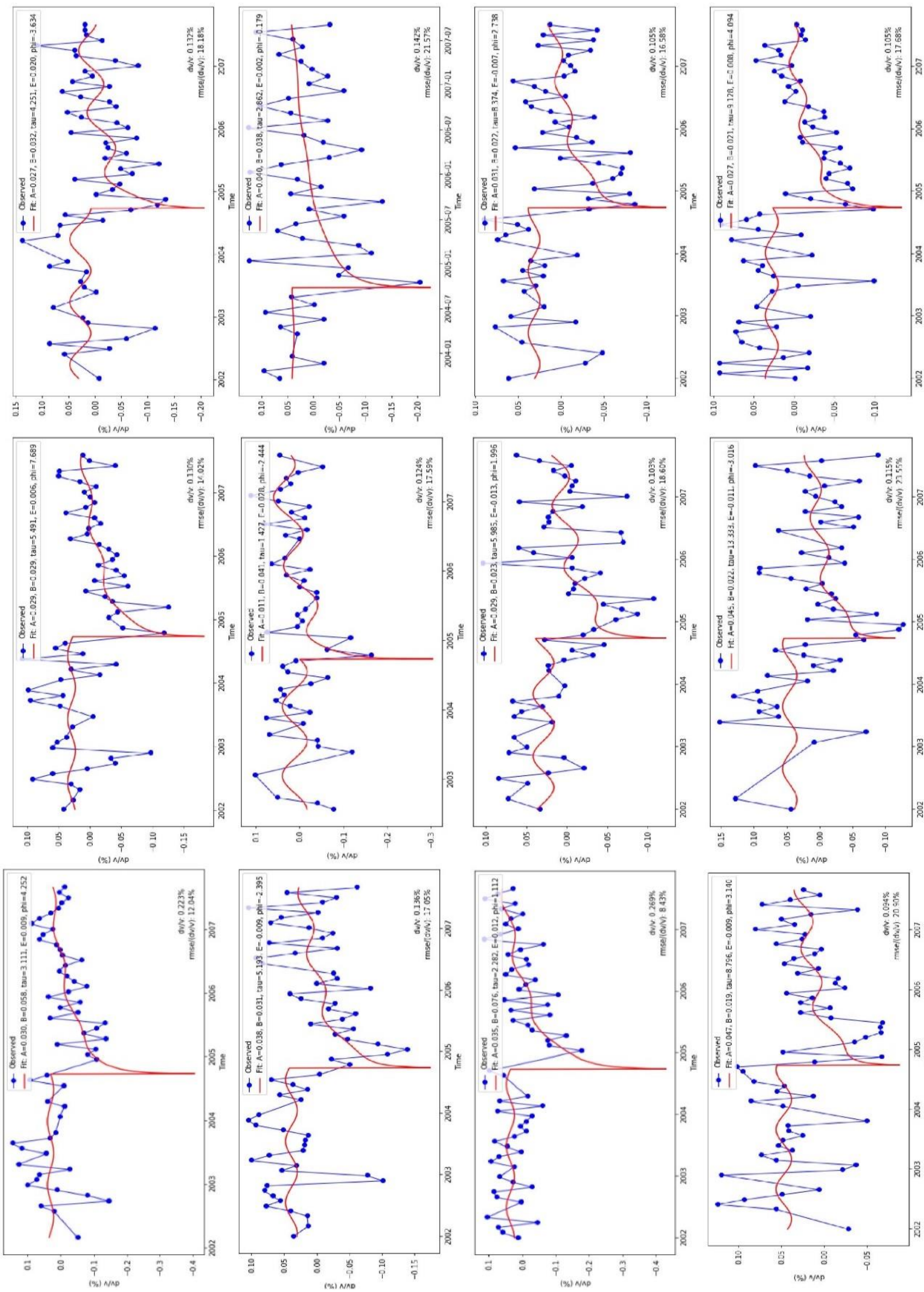


Figure 21. Selected single station pair dv/v measurements at frequency band ***0.9-1.4*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

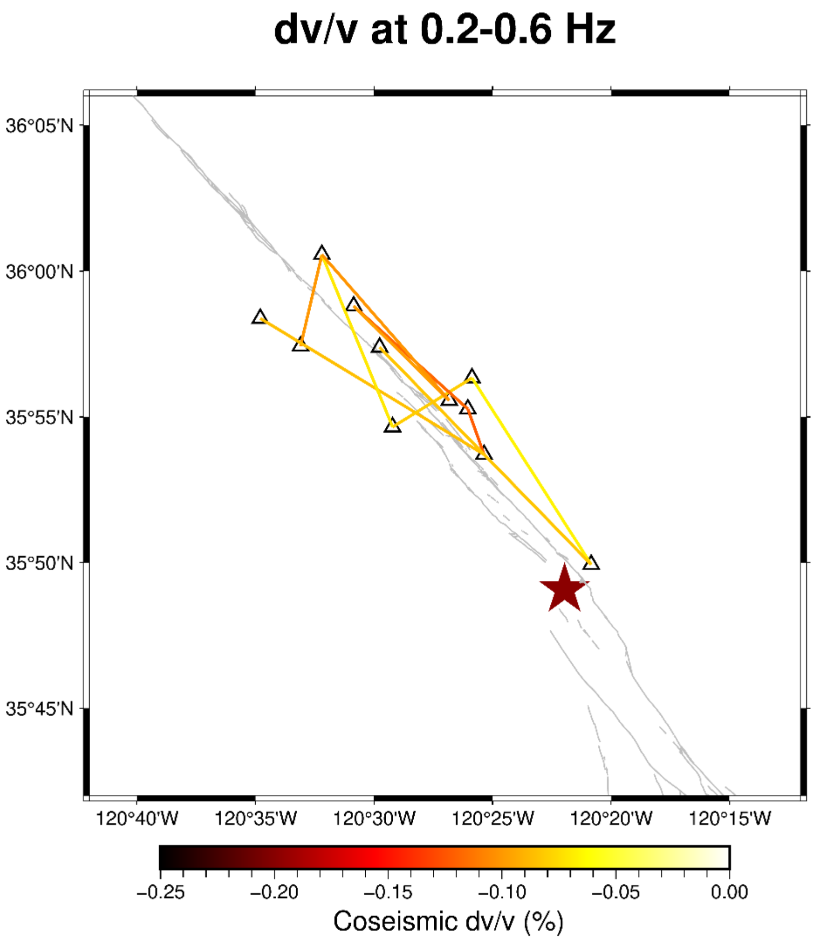


Figure 22. Map view of selected single station pair coseismic dv/v measurements at ***0.2-0.6 Hz*** *frequency band. Black triangles are seismic stations and the red star marks epicenter of the Parkfield earthquake. The grey lines are fault traces.*

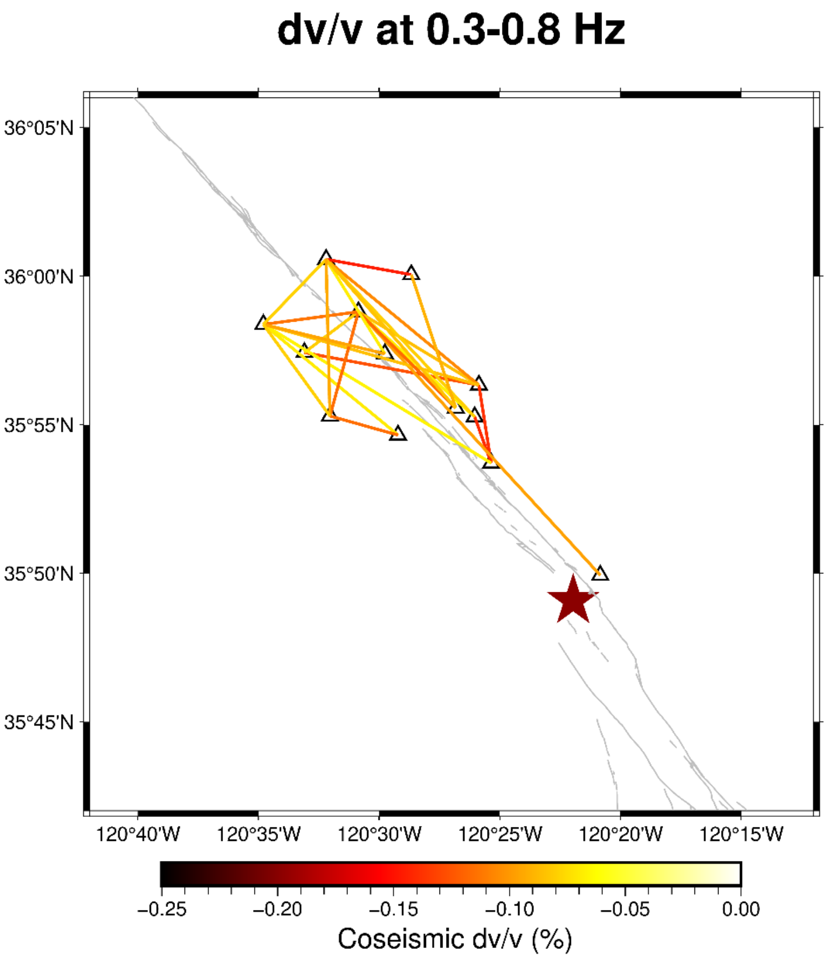


Figure 23. Map view of selected single station pair coseismic dv/v measurements at ***0.3-0.8 Hz*** *frequency band. Black triangles are seismic stations and the red star marks epicenter of the Parkfield earthquake. The grey lines are fault traces.*

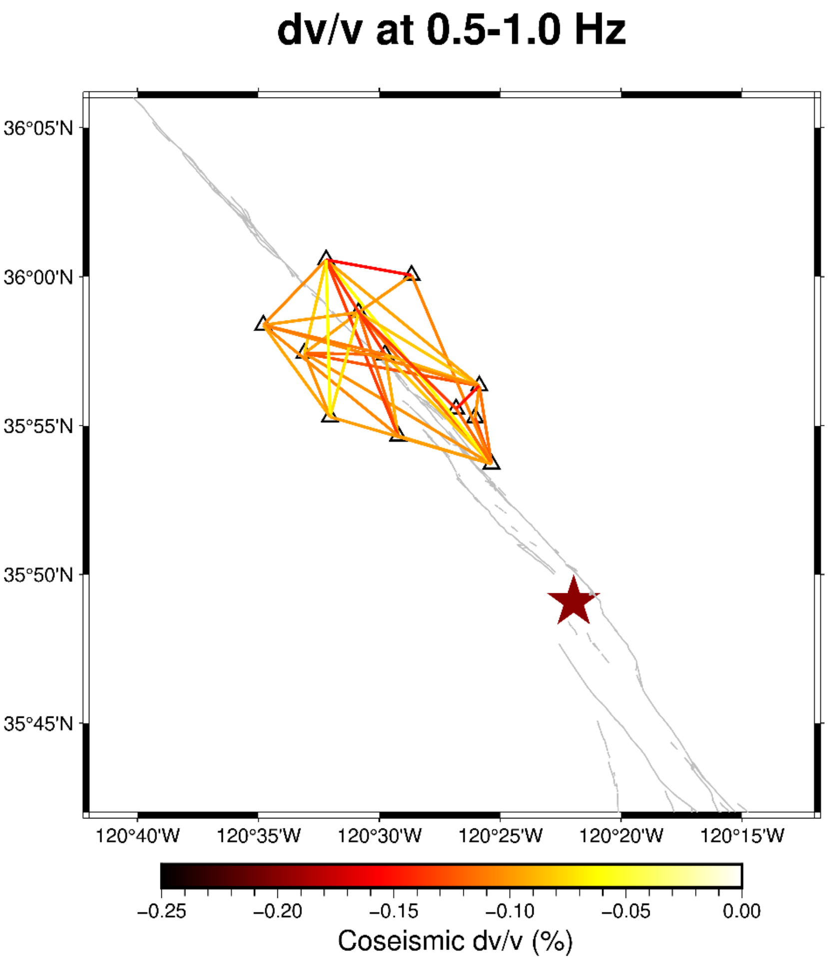


Figure 24. Map view of selected single station pair coseismic dv/v measurements at ***0.5-1.0 Hz*** *frequency band. Black triangles are seismic stations and the red star marks epicenter of the Parkfield earthquake. The grey lines are fault traces.*

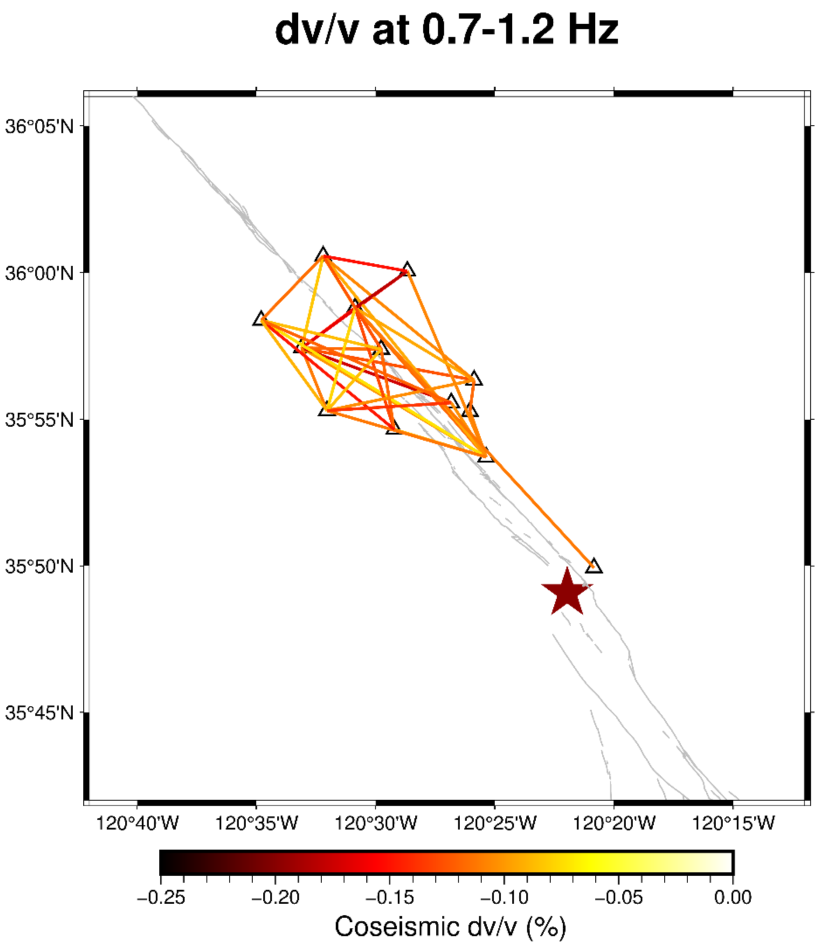


*Figure 25. Map view of selected single station pair coseismic dv/v measurements at **0.7-1.2 Hz** frequency band. Black triangles are seismic stations and the red star marks epicenter of the Parkfield earthquake. The grey lines are fault traces.*

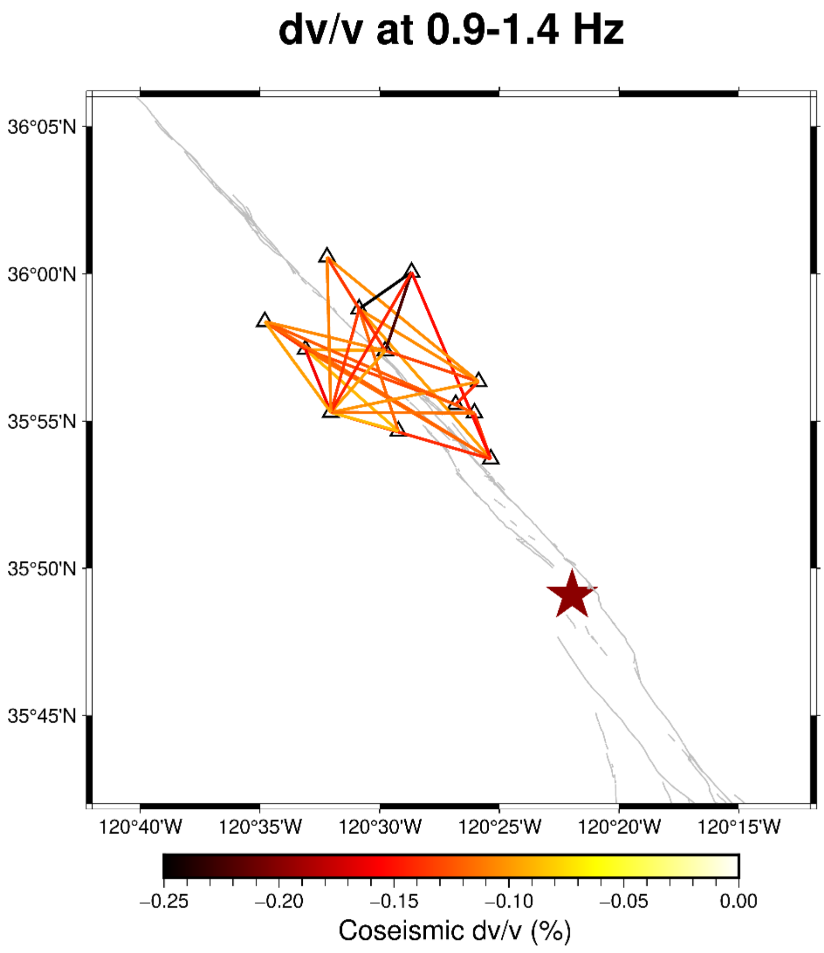


Figure 26. Map view of selected single station pair coseismic dv/v measurements at ***0.9-1.4 Hz*** *frequency band. Black triangles are seismic stations and the red star marks epicenter of the Parkfield earthquake. The grey lines are fault traces.*

4.2. Dynamic and Static Strain Results

The results of dynamic and static strain are shown in the map view (Figures 27-29). These figures elucidate the coseismic and postseismic displacements and strain changes that have been observed at GPS stations and calculated using records from several seismometers. Figures 27 and 28 reveal that static strain is more pronounced between stations located on either side of the fault. The associated displacements are approximately in the order of 10 mm. As for the peak dynamic strain changes (Figure 29), these strain alterations are largely concentrated on the fault and its immediate surroundings. Areas situated at a further distance from the fault display smaller peak dynamic strain changes. For both static and dynamic strain changes, the amplitude are very similar.

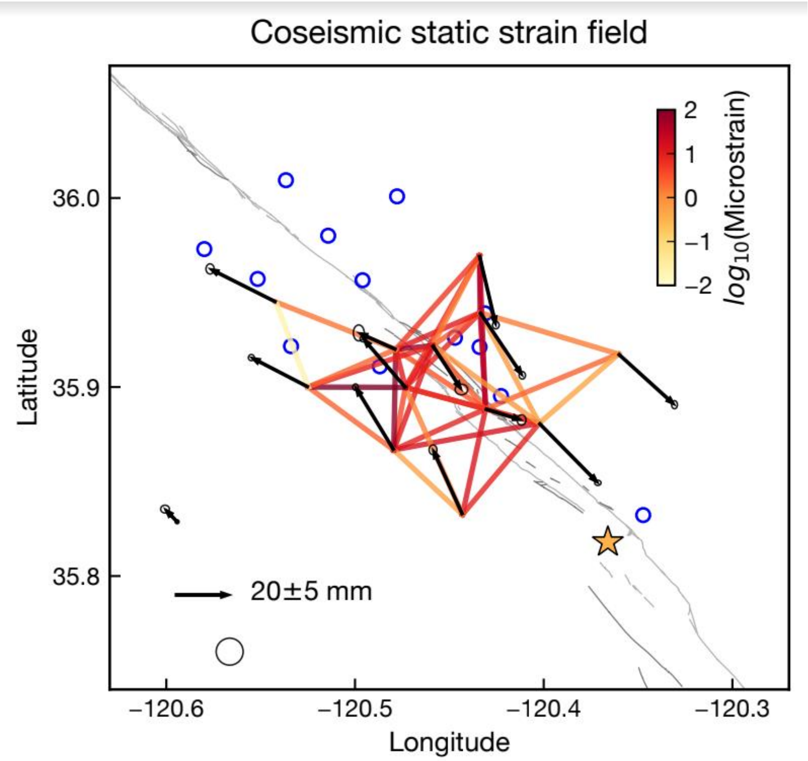


Figure 27. Map view of coseismic static strain changes. The black circles are GPS stations. The blue circles mark seismic stations. The yellow star marks the epicenter. The arrow marks displacement and direction. Red color means higher strain changes and yellow means smaller strain changes.

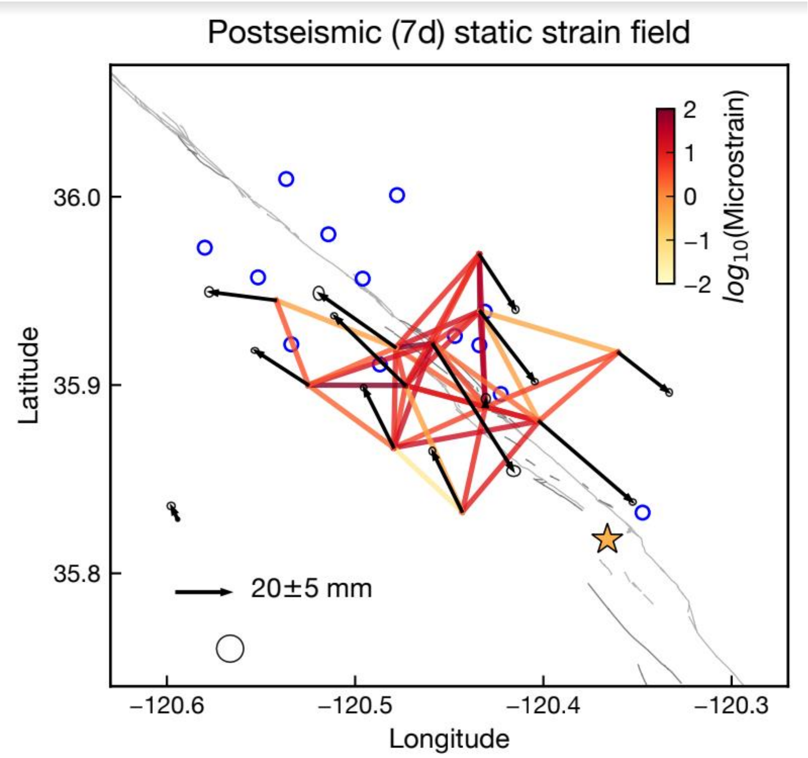


Figure 28. Map view of postseismic static strain changes. The black circles are GPS stations. The blue circles mark seismic stations. The yellow star marks the epicenter. The arrow marks displacement and direction. Red color means higher strain changes and yellow means smaller strain changes.

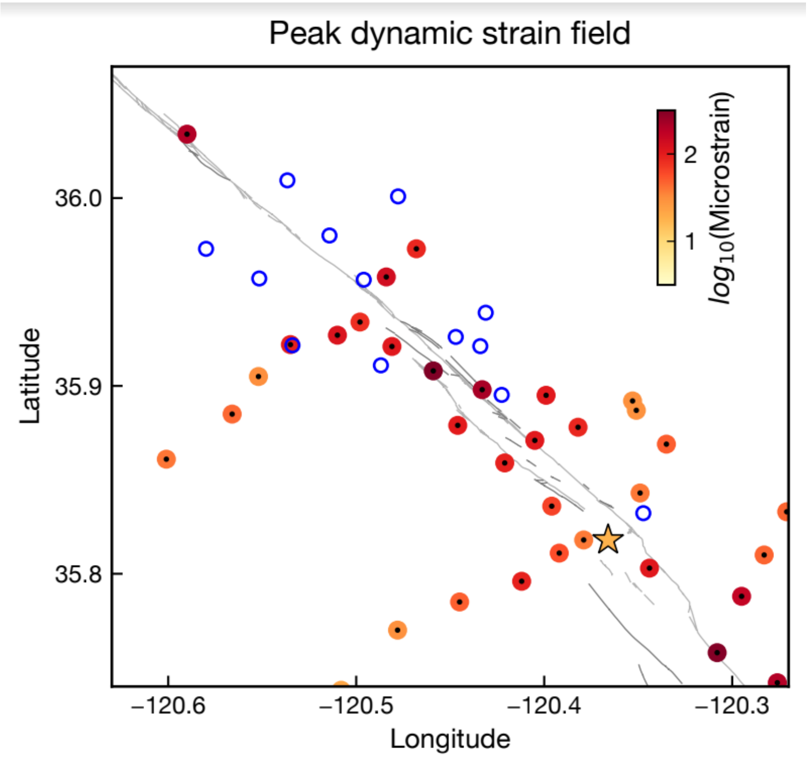


Figure 29. Map view of peak dynamic strain changes. The black dots are seismometer stations. The blue circles mark BP stations. The yellow star marks the epicenter. Red color means higher strain changes and yellow means smaller strain changes.

5. Discussion

5.1. Relation between dv/v and Strain Changes

Brenguier et al. (2008) and Wu et al. (2016) have previously employed the ambient noise interferometry method on the same data set, utilizing all three components and the vertical component exclusively, respectively. Wu et al. (2016) conducted an examination of average velocity variations across different frequency bands and established constraints for the depth of velocity variation. In this study, we have expanded upon the analysis to single station pairs, offering insights into the patterns of spatial velocity variations. Our observations are consistent with the previously noted increase in the amplitude of the velocity drop at higher frequencies and the gradual trend of postseismic recovery. We conclude that a higher coseismic velocity change at higher frequency bands indicates more pronounced changes in rock elastic properties and potential inelastic deformation at shallower depths, with these effects diminishing at greater depths. Our results also imply that the coseismic velocity change may be generally larger on the eastern side of the San Andreas Fault. However, at higher frequencies, the southwest part of the BP station array also experiences significant velocity changes. From a theoretical standpoint, we would expect to see the highest dv/v changes spatially concentrated along the fault line, decreasing with distance from the fault. This expectation assumes that rocks in closer proximity to the fault would undergo more damage during earthquake rupture than those located further away. However, such patterns are not evident in our results, prompting further investigation and analysis to better understand these spatial velocity change patterns. The spatial distribution of seismic velocity changes presents a more nuanced picture of the seismic phenomena associated with different frequency bands and may offer insights into the geodynamic processes at play in different areas surrounding the San Andreas Fault.

Previous studies have explained the sudden velocity drop concurrent with large earthquakes with a variety of factors. These include the opening of cracks, reduced packing within weak sedimentary layers and fault zones, or an increase in water flow into cracks (Li et al., 2007). However, the entire process remains a subject of debate and is not fully understood. Since Figures 27-29 only map amplitudes of strain changes and neglect the directions of strain changes, we cannot yet investigate the relationship between velocity changes and tensional and compression strains. However, we can still observe that coseismic dynamic strains are generally several times larger than coseismic static strain changes, especially in areas further away from the fault. Therefore, coseismic dynamic strains likely contributed more to the seismic velocity drops, as suggested in Wu et al. (2009). The postseismic seismic velocity recovery may be related to postseismic static strain changes because dynamic strain changes only exist coseismically. Nonetheless, the relation between the spatial change patterns of dv/v and strain remains inconclusive. This ambiguity could be due to the distribution of seismic and GPS stations and uncertainties and potential biases in estimates of strain and dv/v. The connection between strain and velocity changes is complex and achieving a comprehensive understanding of it would require a more detailed study with a denser network of stations.

To check the reliability of our curve fitting results, we compare values of best-fitting model parameters such as the recovery time (*tau*) with results of previous studies (Brenguier et al., 2008; Wu et al., 2016; Okubo & Denolle, 2020). For averaged *dv/v* measurements, higher frequency bands agree better with results from Brenguier et al. (2008) and Wu et al. (2016). At frequency bands higher than 0.5-1.0 Hz, our models have a recovery time of around 10 years, which is comparable with previous studies (Okubo & Denolle, 2020). However, model parameters are relatively random at lower frequency bands, which can be a result of large data noise.

5.2. Limitations and Potential Improvements

While we have successfully captured some key features of velocity variations between different station pairs, there are several areas where improvements could be made. Firstly, higher waveform coherence would yield more accurate measurements, so the coherence between the current and reference waveform could be further increased by selectively stacking correlation functions. However, care must be taken not to excessively remove correlation functions. Secondly, to better illustrate the dv/v change over time, we could introduce a moving window for stacking substack waveforms. This would allow us to visualize and analyze temporal changes more effectively and in detail. When it comes to calculating *dv/v* values, we could also consider more comprehensive methods like the Weighted Time-Domain Cross-Correlation (WTDTW) method. This approach accounts for both time- and frequency-domain features of waveforms, offering a more holistic view of the data. Our single station *dv/v* results still have room for improvement. The data points have large fluctuations, making the velocity variation pattern not very clear, particularly for low-frequency bands. In addition, to better understand the possible relation between strain and seismic velocity changes, we should examine local rock properties maps, the relation between seismic velocity changes and distance from the fault, the spatial distribution of modeling parameters, and better estimation of errors in both modeling and calculating *dv/v* values.

Finally, our method for examining spatial variation currently relies on first-order connections between stations (Mao et al., 2020). Future studies could incorporate inversion techniques and more advanced methodologies to better constrain and understand the spatial patterns of velocity variations.

6. Conclusion

In this study, we implemented the ambient noise interferometry method to process continuous ambient noise data from 13 BP borehole stations. By exploiting and expanding the functions of an open-source Python package, NoisePy, we were able to execute the entire workflow of ambient noise interferometry with careful selection of important parameters and methods. Our results illustrate clear changes in dv/v measurements both over time and during earthquakes. The observed coseismic velocity variation aligns well with previous findings, and we further extended our analysis to single station pairs to resolve spatial variations. Our results suggest that the eastern part of the San Andreas Fault may exhibit larger velocity changes. We utilized static and dynamic strain change data for comparison with our results. Our analysis suggests that the velocity drop is primarily due to dynamic strain change, while the postseismic recovery process is predominantly related to static strain change. Several aspects of our study can be improved in the future. The coherence between substack and reference waveforms could be further improved to yield more accurate results. Additionally, other recently introduced methods could also be employed for calculating velocity variations. Lastly, seismic inversions could be used to better model the velocity changes at different depths and directions. More comprehensive information can be revealed in the future through careful data processing and the selection of appropriate techniques, thus deepening our understanding of seismic activity and fault behavior.

7. References

Abercrombie, R. E. (2013). Comparison of direct and coda wave stress drop measurements for the Wells, Nevada, earthquake sequence. *Journal of Geophysical Research: Solid Earth*, *118*(4), 1458–1470. https://doi.org/10.1029/2012JB009638

Aki, K., & Chouet, B. (1975). Origin of coda waves: Source, attenuation, and scattering effects. *Journal of Geophysical Research*, *80*(23), 3322–3342. https://doi.org/10.1029/JB080i023p03322

Aki K. (1957). Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bull. Earthq. Res. Inst*. 35:415–56

Baig, A. M., Campillo, M., & Brenguier, F. (2009). Denoising seismic noise cross correlations. *Journal of Geophysical Research*, *114*(B8), B08310. https://doi.org/10.1029/2008JB006085

Bakun, W. H., & Lindh, A. G. (1985). The Parkfield, California, Earthquake Prediction Experiment. *Science*, *229*(4714), 619–624. https://doi.org/10.1126/science.229.4714.619

Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., & Waldhauser, F. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. Nature, 437(7061), 969-974.

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

Berg, E. M., Lin, F. -C., Allam, A., Qiu, H., Shen, W., & Ben-Zion, Y. (2018). Tomography of Southern California Via Bayesian Joint Inversion of Rayleigh Wave Ellipticity and Phase Velocity From Ambient Noise Cross-Correlations. *Journal of Geophysical Research: Solid Earth*, *123*(11), 9933–9949. https://doi.org/10.1029/2018JB016269

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python Toolbox for Seismology. *Seismological Research Letters*, *81*(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E. (2008). Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations. *Science*, *321*(5895), 1478–1481. https://doi.org/10.1126/science.1160943

Brenguier, F., Shapiro, N. M., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., & Nercessian, A. (2008). Towards forecasting volcanic eruptions using seismic noise. *Nature Geoscience, 1*(2), 126–130. https://doi.org/10.1038/ngeo104

Campillo, M. (2006). Phase and Correlation in `Random' Seismic Fields and the Reconstruction of the Green Function. *Pure and Applied Geophysics*, *163*(2–3), 475–502. https://doi.org/10.1007/s00024-005-0032-8

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

Catchings, R. D. (2002). High-Resolution Seismic Velocities and Shallow Structure of the San Andreas Fault Zone at Middle Mountain, Parkfield, California. *Bulletin of the Seismological Society of America*, *92*(6), 2493–2503. https://doi.org/10.1785/0120010263

Choi, K., Bilich, A., Larson, K. M., & Axelrad, P. (2004). Modified sidereal filtering: Implications for high-rate GPS positioning. Geophysical research letters, 31(22).

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269. https://doi.org/10.1190/1.1439927

Clarke, D., Zaccarelli, L., Shapiro, N. M., & Brenguier, F. (2011). Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise: MWCS: assessment of resolution and accuracy. *Geophysical Journal International*, *186*(2), 867–882. https://doi.org/10.1111/j.1365-246X.2011.05074.x

Dong, D., Fang, P., Bock, Y., Cheng, M. K., & Miyazaki, S. I. (2002). Anatomy of apparent seasonal variations from GPS‐derived site position time series. Journal of Geophysical Research: Solid Earth, 107(B4), ETG-9.

Grêt, A., Snieder, R., & Özbay, U. (2006). Monitoring *in situ* stress changes in a mining environment with coda wave interferometry. *Geophysical Journal International*, *167*(2), 504–508. https://doi.org/10.1111/j.1365-246X.2006.03097.x

Gupta, H. K. (Ed.). (2021). *Encyclopedia of Solid Earth Geophysics*. Springer International Publishing. https://doi.org/10.1007/978-3-030-58631-7

Hill, D. P., et al. (1993), Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, Science, 260, 1617–1623.

Hillers, G., Campillo, M., Brenguier, F., Moreau, L., Agnew, D. C., & Ben‐Zion, Y. (2019). Seismic Velocity Change Patterns Along the San Jacinto Fault Zone Following the 2010 *M* 7.2 El Mayor‐Cucapah and *M* 5.4 Collins Valley Earthquakes. *Journal of Geophysical Research: Solid Earth*, *124*(7), 7171–7192. https://doi.org/10.1029/2018JB017143

Hobiger, M., Wegler, U., Shiomi, K., & Nakahara, H. (2012). Coseismic and postseismic elastic wave velocity variations caused by the 2008 Iwate-Miyagi Nairiku earthquake, Japan: 2008 IWATE EARTHQUAKE VELOCITY CHANGES. *Journal of Geophysical Research: Solid Earth*, *117*(B9). https://doi.org/10.1029/2012JB009402

HRSN (2014), High Resolution Seismic Network. UC Berkeley Seismological Laboratory. Dataset. doi:10.7932/HRSN

Jiang, C., & Denolle, M. A. (2020). NoisePy: A New High‐Performance Python Tool for Ambient‐Noise Seismology. *Seismological Research Letters*, *91*(3), 1853–1866. https://doi.org/10.1785/0220190364

Jiang, J., Bock, Y., & Klein, E. (2021). Coevolving early afterslip and aftershock signatures of a San Andreas fault rupture. Science Advances, 7(15), eabc1606.

Kumar, U., Legendre, Cédric. P., Zhao, L., & Chao, B. F. (2022). Dynamic Time Warping as an Alternative to Windowed Cross Correlation in Seismological

Applications. *Seismological Research Letters*, *93*(3), 1909–1921. https://doi.org/10.1785/0220210288

Lecocq, T., Caudron, C., & Brenguier, F. (2014). MSNoise, a Python Package for Monitoring Seismic Velocity Changes Using Ambient Seismic Noise. *Seismological Research Letters, 85*(3), 715–726. https://doi.org/10.1785/0220130073

Li, Y.-G., Chen, P., Cochran, E. S., & Vidale, J. E. (2007). Seismic velocity variations on the San Andreas fault caused by the 2004 M6 Parkfield Earthquake and their implications. *Earth, Planets and Space*, *59*(1), 21–31. https://doi.org/10.1186/BF03352018

Lin, F.-C., Moschetti, M. P., & Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International*, *173*(1), 281–298. https://doi.org/10.1111/j.1365-246X.2008.03720.x

Liu, Z., Huang, J., Peng, Z., & Su, J. (2014). Seismic velocity changes in the epicentral region of the 2008 Wenchuan earthquake measured from three-component ambient noise correlation techniques: TEMPORAL CHANGES IN WENCHUAN EPICENTER. *Geophysical Research Letters, 41*(1), 37–42. https://doi.org/10.1002/2013GL058682

Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the Green’s function in the correlations of a diffuse field. *J. Acoust. Soc. Am.*, *110*(6).

Mao, S., Campillo, M., Hilst, R. D., Brenguier, F., Stehly, L., & Hillers, G. (2019). High Temporal Resolution Monitoring of Small Variations in Crustal Strain by Dense Seismic Arrays. *Geophysical Research Letters*, *46*(1), 128–137. https://doi.org/10.1029/2018GL079944

Mao, S., Lecointre, A., van der Hilst, R. D., & Campillo, M. (2022). Space-time monitoring of groundwater fluctuations with passive seismic interferometry. *Nature Communications*, *13*(1), 4643. https://doi.org/10.1038/s41467-022-32194-3

Mao, S., Mordret, A., Campillo, M., Fang, H., & van der Hilst, R. D. (2020). On the measurement of seismic traveltime changes in the time–frequency domain with wavelet cross-spectrum analysis. *Geophysical Journal International*, *221*(1), 550–568. https://doi.org/10.1093/gji/ggz495

Mikesell, T. D., Malcolm, A. E., Yang, D., & Haney, M. M. (2015). A comparison of methods to estimate seismic phase delays: Numerical examples for coda wave interferometry. *Geophysical Journal International*, *202*(1), 347–360. https://doi.org/10.1093/gji/ggv138

Mordret, A., Jolly, A. D., Duputel, Z., & Fournier, N. (2010). Monitoring of phreatic eruptions using Interferometry on Retrieved Cross-Correlation Function from Ambient Seismic Noise: Results from Mt. Ruapehu, New Zealand. *Journal of Volcanology and Geothermal Research*, *191*(1–2), 46–59. https://doi.org/10.1016/j.jvolgeores.2010.01.010

Mordret, A., Shapiro, N. M., & Singh, S. (2014). Seismic noise-based time-lapse monitoring of the Valhall overburden. *Geophysical Research Letters*, *41*(14), 4945–4952. https://doi.org/10.1002/2014GL060602

Nicolson, H., Curtis, A., Baptie, B., & Galetti, E. (2012). Seismic interferometry and ambient noise tomography in the British Isles. *Proceedings of the Geologists' Association*, *123*(1), 74–86. https://doi.org/10.1016/j.pgeola.2011.04.002

Obermann, A., Planès, T., Larose, E., Sens-Schönfelder, C., & Campillo, M. (2013). Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium. *Geophysical Journal International, 194*(1), 372–382. https://doi.org/10.1093/gji/ggt043

Poupinet, G., Ellsworth, W. L., & Frechet, J. (1984). Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5719–5731. https://doi.org/10.1029/JB089iB07p05719

Radiguet, M., Perfettini, H., Cotte, N., Gualandi, A., Valette, B., Kostoglodov, V., Lhomme, T., Walpersdorf, A., Cabral Cano, E., & Campillo, M. (2016). Triggering of the 2014 Mw7.3 Papanoa earthquake by a slow slip event in

Guerrero, Mexico. *Nature Geoscience*, *9*(11), 829–833. https://doi.org/10.1038/ngeo2817

Ratdomopurbo, A., & Poupinet, G. (1995). Monitoring a temporal change of seismic velocity in a volcano: Application to the 1992 eruption of Mt. Merapi (Indonesia). *Geophysical Research Letters*, *22*(7), 775–778. https://doi.org/10.1029/95GL00302

Retailleau, L., Boué, P., Li, L., & Campillo, M. (2020). Ambient seismic noise imaging of the lowermost mantle beneath the North Atlantic Ocean. *Geophysical Journal International*, *222*(2), 1339–1351. https://doi.org/10.1093/gji/ggaa210

Rivet, D., Campillo, M., Radiguet, M., Zigone, D., Cruz-Atienza, V., Shapiro, N. M., Kostoglodov, V., Cotte, N., Cougoulat, G., Walpersdorf, A., & Daub, E. (2014). Seismic velocity changes, strain rate and non-volcanic tremors during the 2009–2010 slow slip event in Guerrero, Mexico. *Geophysical Journal International*, *196*(1), 447–460. https://doi.org/10.1093/gji/ggt374

Roux, P., Moreau, L., Lecointre, A., Hillers, G., Campillo, M., Ben-Zion, Y., Zigone, D., & Vernon, F. (2016). A methodological approach towards high-resolution surface wave imaging of the San Jacinto Fault Zone using ambient-noise recordings at a spatially dense array. *Geophysical Journal International*, *206*(2), 980–992. https://doi.org/10.1093/gji/ggw193

Seats, K. J., Lawrence, J. F., & Prieto, G. A. (2012). Improved ambient noise correlation functions using Welch's method: Improved ambient NCFs using Welch's method. *Geophysical Journal International*, *188*(2), 513–523. https://doi.org/10.1111/j.1365-246X.2011.05263.x

Shakal, A., Haddadi, H., Graizer, V., Lin, K., & Huang, M. (2006). Some key features of the strong-motion data from the M 6.0 Parkfield, California, earthquake of 28 September 2004. Bulletin of the Seismological Society of America, 96(4B), S90-S118.

Shapiro, N. M., & Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise: CORRELATIONS OF THE

SEISMIC NOISE. *Geophysical Research Letters*, *31*(7), n/a-n/a. https://doi.org/10.1029/2004GL019491

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H. (2005). High-Resolution Surface-Wave Tomography from Ambient Seismic Noise. *Science*, *307*(5715), 1615–1618. https://doi.org/10.1126/science.1108339

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, *69*(4), 046610. https://doi.org/10.1103/PhysRevE.69.046610

Snieder, R., Grêt, A., Douma, H., & Scales, J. (2002). Coda Wave Interferometry for Estimating Nonlinear Behavior in Seismic Velocity. *Science*, *295*(5563), 2253–2255. https://doi.org/10.1126/science.1070015

Snieder, R., & Larose, E. (2013). Extracting Earth's Elastic Wave Response from Noise Measurements. *Annual Review of Earth and Planetary Sciences*, *41*(1), 183–206. https://doi.org/10.1146/annurev-earth-050212-123936

Stehly, L., Campillo, M., & Shapiro, N. M. (2006). A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research*, *111*(B10), B10306. https://doi.org/10.1029/2005JB004237

Taira, T., Brenguier, F., & Kong, Q. (2015). Ambient noise-based monitoring of seismic velocity changes associated with the 2014 M_w 6.0 South Napa earthquake. *Geophysical Research Letters*, *42*(17), 6997–7004. https://doi.org/10.1002/2015GL065308

Unsworth, M. J., Malin, P. E., Egbert, G. D., & Booker, J. R. (1997). Internal structure of the San Andreas fault at Parkfield, California. *Geology*, *25*(4), 359. https://doi.org/10.1130/0091-7613(1997)025<0359:ISOTSA>2.3.CO;2

Wang, Q., Campillo, M., Brenguier, F., Lecointre, A., Takeda, T., & Hashima, A. (2019). Evidence of Changes of Seismic Properties in the Entire Crust Beneath Japan After the M_w 9.0, 2011 Tohoku-oki Earthquake. *Journal of Geophysical Research: Solid Earth*, *124*(8), 8924–8941. https://doi.org/10.1029/2019JB017803

Wang, Q.-Y., Brenguier, F., Campillo, M., Lecointre, A., Takeda, T., & Aoki, Y. (2017). Seasonal Crustal Seismic Velocity Changes Throughout Japan: SEASONAL SEISMIC VELOCITY CHANGES. *Journal of Geophysical Research: Solid Earth*, *122*(10), 7987–8002. https://doi.org/10.1002/2017JB014307

Wapenaar, K., Slob, E., Snieder, R., & Curtis, A. (2010). Tutorial on seismic interferometry: Part 2 — Underlying theory and new advances. *GEOPHYSICS*, *75*(5), 75A211-75A227. https://doi.org/10.1190/1.3463440

Weaver, R. L., Hadziioannou, C., Larose, E., & Campillo, M. (2011). On the precision of noise correlation interferometry: Precision of noise correlation interferometry. *Geophysical Journal International*, *185*(3), 1384–1392. https://doi.org/10.1111/j.1365-246X.2011.05015.x

Wu, C., Delorey, A., Brenguier, F., Hadziioannou, C., Daub, E. G., & Johnson, P. (2016). Constraining depth range of *S* wave velocity decrease after large earthquakes near Parkfield, California: DEPTH RANGE OF *S* WAVE VELOCITY DECREASE. *Geophysical Research Letters*, *43*(12), 6129–6136. https://doi.org/10.1002/2016GL069145

Yang, X., Bryan, J., Okubo, K., Jiang, C., Clements, T., & Denolle, M. (2022). *Optimal Stacking of Noise Cross-Correlation Functions* [Preprint]. Geophysics. https://doi.org/10.1002/essoar.10511292.2

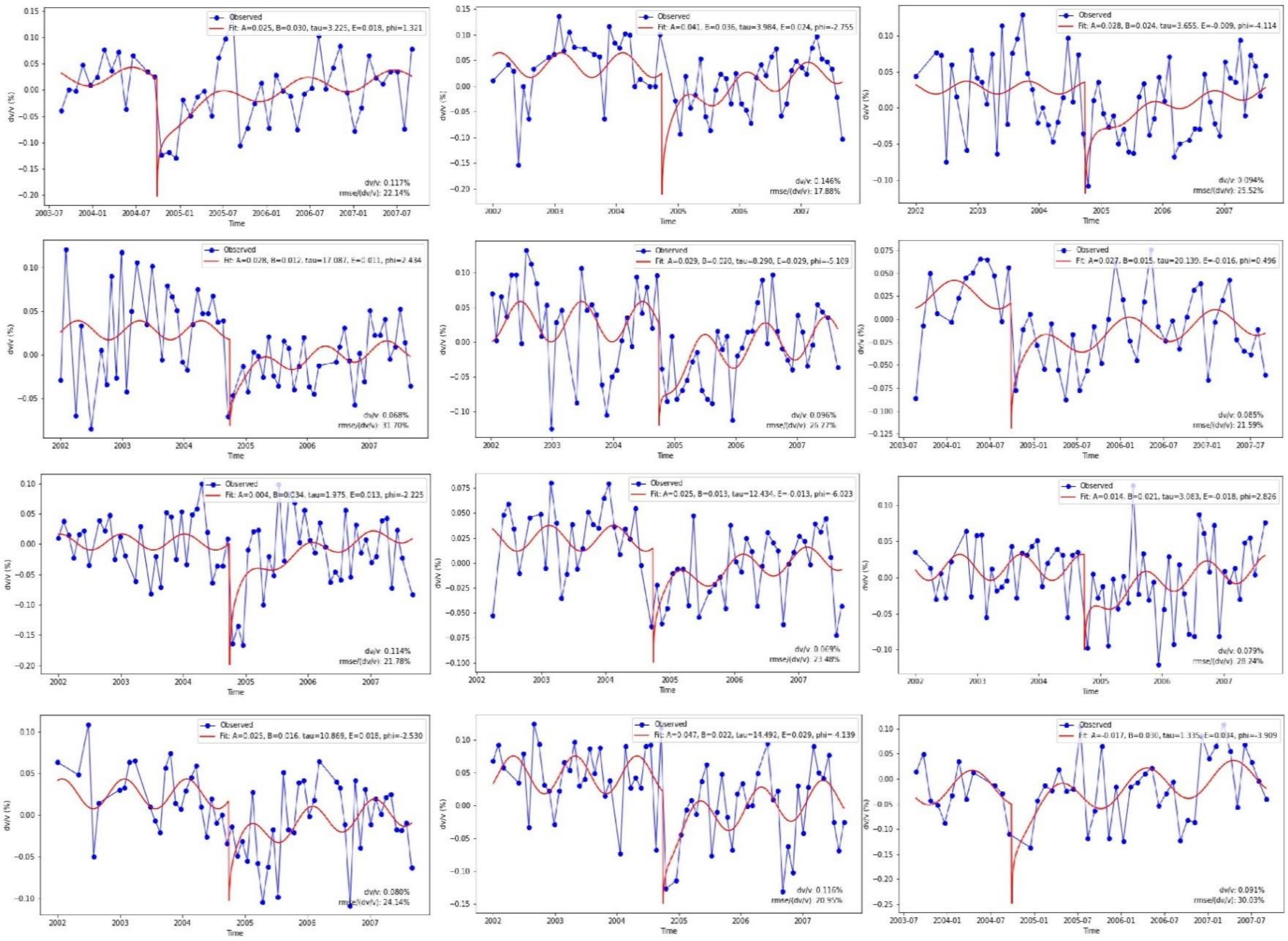
Yuan, C., Bryan, J., & Denolle, M. (2021). Numerical comparison of time-, frequency- and wavelet-domain methods for coda wave interferometry. *Geophysical Journal International*, *226*(2), 828–846. https://doi.org/10.1093/gji/ggab140

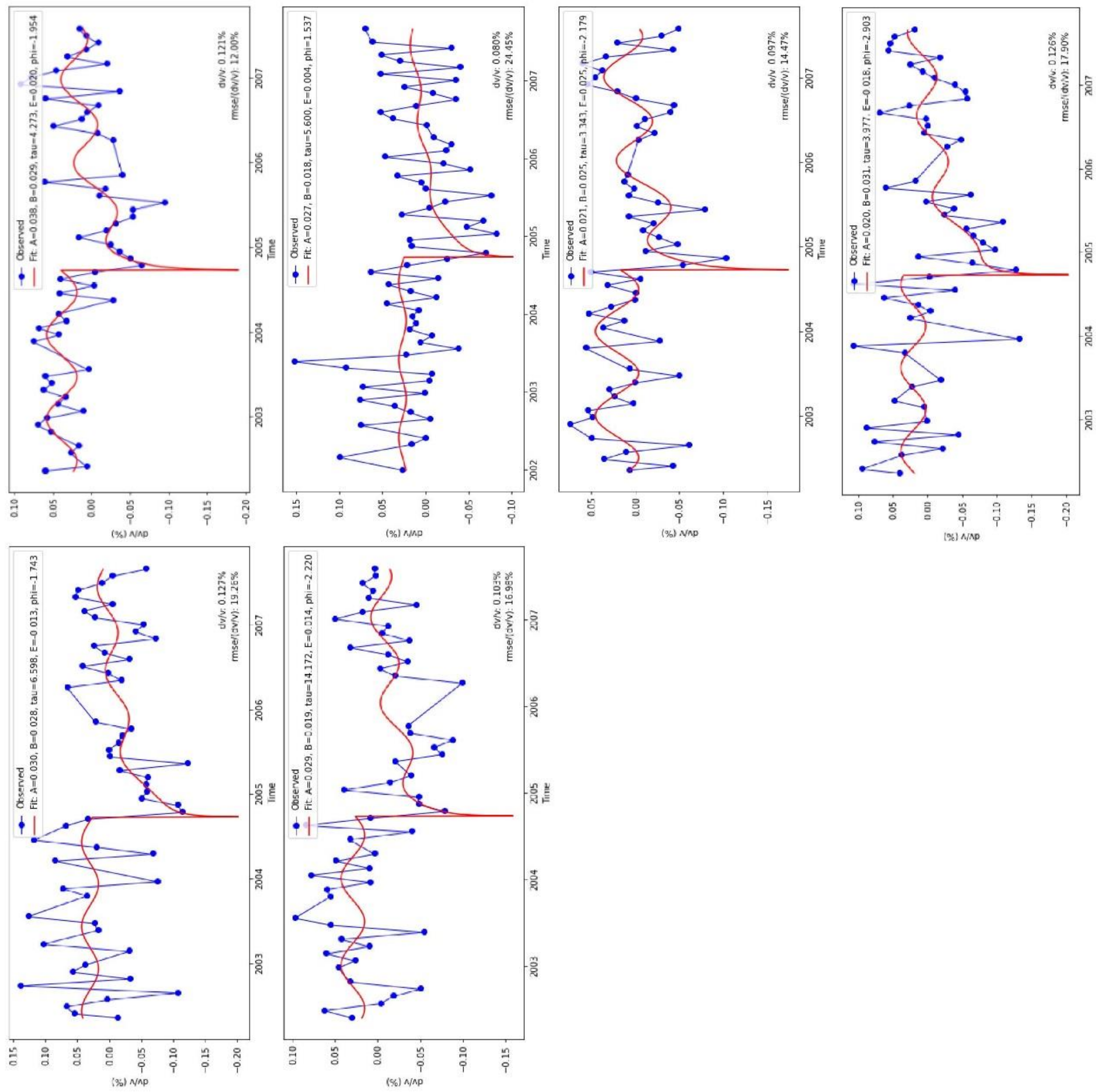
Zhan, Z., Tsai, V. C., & Clayton, R. W. (2013). Spurious velocity changes caused by temporal variations in ambient noise frequency content. *Geophysical Journal International*, *194*(3), 1574–1581. https://doi.org/10.1093/gji/ggt170

Zhao, P., Peng, Z., Shi, Z., Lewis, M. A., & Ben-Zion, Y. (2010). Variations of the velocity contrast and rupture properties of M6 earthquakes along the Parkfield section of the San Andreas fault. *Geophysical Journal International*, *180*(2), 765–780. https://doi.org/10.1111/j.1365-246X.2009.04436.x

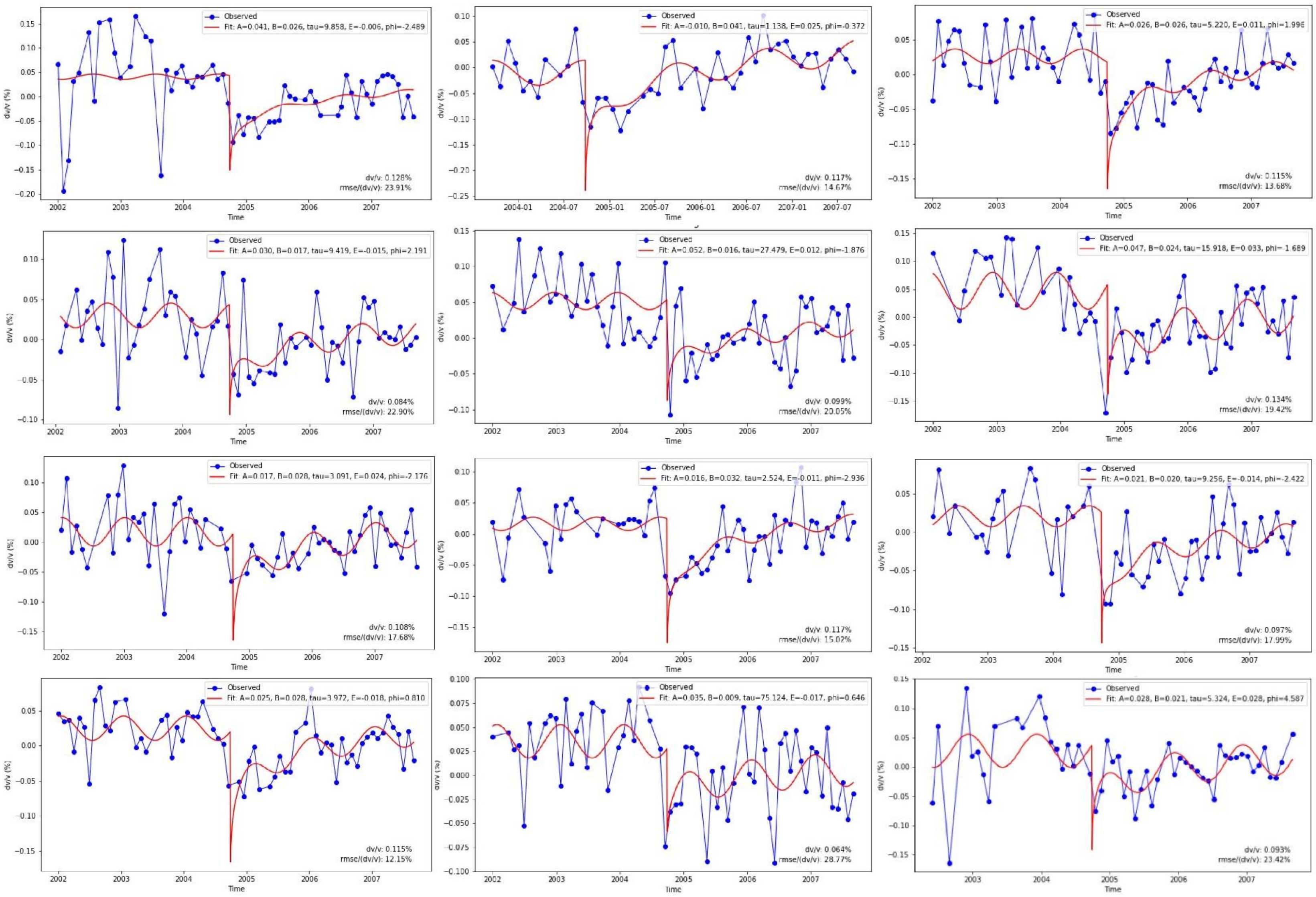
8. Appendix

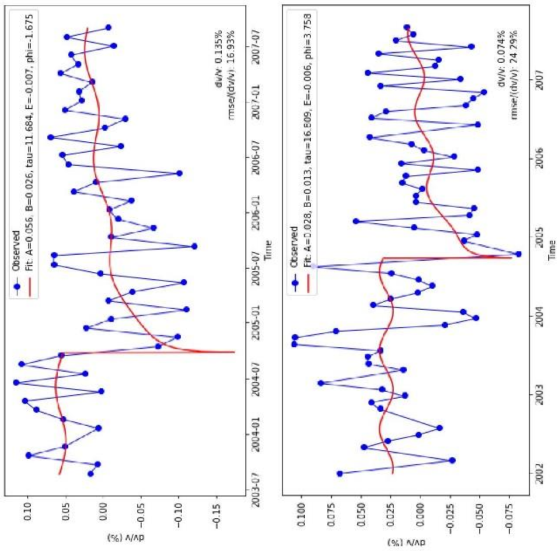
Appendix Figures 1-4 are complementary to Figures 21-24, showing all single station dv/v plots that cannot fit in the main text. By providing the complete sets of dv/v plots at 5 selected frequency bands, we can better observe and understand the features of velocity variations and check quality of curve fitting.


Observed
Fit: A=0.025, B=0.030, tau=3.225, E=0.018, phi=1.321
dv/v: 0.117%
rmse/(dv/v): 22.14%
Observed
Fit: A=0.041, B=0.036, tau=3.984, E=0.024, phi=-2.755
dv/v: 0.146%
rmse/(dv/v): 17.88%
Observed
Fit: A=0.028, B=0.024, tau=3.655, E=-0.009, phi=-4.114
dv/v: 0.094%
rmse/(dv/v): 25.52%
Observed
Fit: A=0.028, B=0.012, tau=17.087, E=0.011, phi=2.434
dv/v: 0.068%
rmse/(dv/v): 31.70%
Observed
Fit: A=0.029, B=0.020, tau=8.290, E=0.029, phi=-5.109
dv/v: 0.096%
rmse/(dv/v): 26.27%
Observed
Fit: A=0.027, B=0.015, tau=20.139, E=-0.016, phi=0.496
dv/v: 0.085%
rmse/(dv/v): 21.59%
Observed
Fit: A=0.004, B=0.034, tau=1.975, E=0.013, phi=-2.225
dv/v: 0.114%
rmse/(dv/v): 21.78%
Observed
Fit: A=0.025, B=0.013, tau=12.434, E=-0.013, phi=-6.023
dv/v: 0.069%
rmse/(dv/v): 23.48%
Observed
Fit: A=0.014, B=0.021, tau=3.083, E=-0.018, phi=2.826
dv/v: 0.079%
rmse/(dv/v): 28.24%
Observed
Fit: A=0.025, B=0.016, tau=10.869, E=0.018, phi=-2.530
dv/v: 0.080%
rmse/(dv/v): 24.14%
Observed
Fit: A=0.047, B=0.022, tau=14.492, E=0.029, phi=-4.139
dv/v: 0.116%
rmse/(dv/v): 20.95%
Observed
Fit: A=-0.017, B=0.030, tau=1.335, E=0.034, phi=-3.909
dv/v: 0.091%
rmse/(dv/v): 30.03%
dv/v (%)
Time

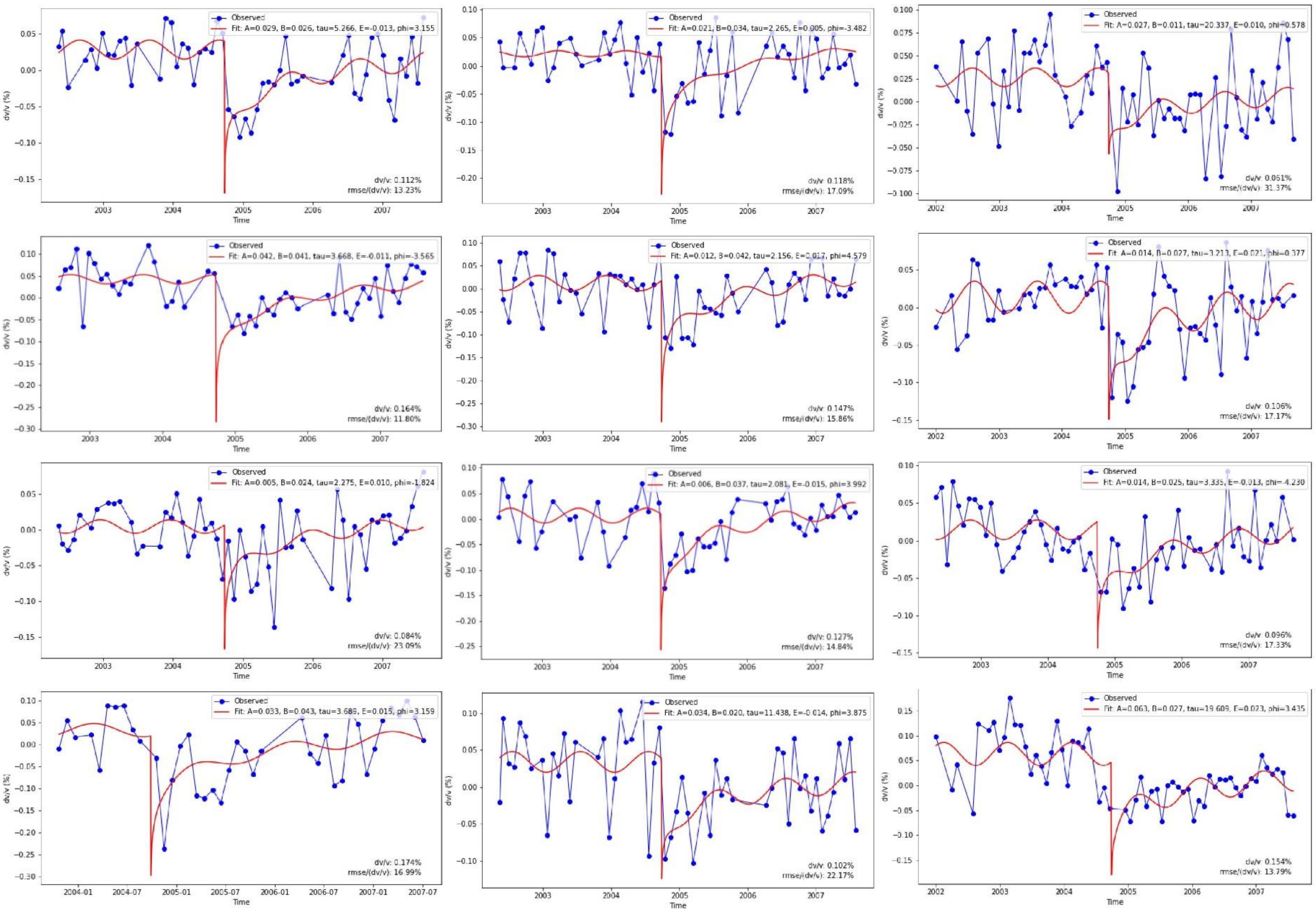


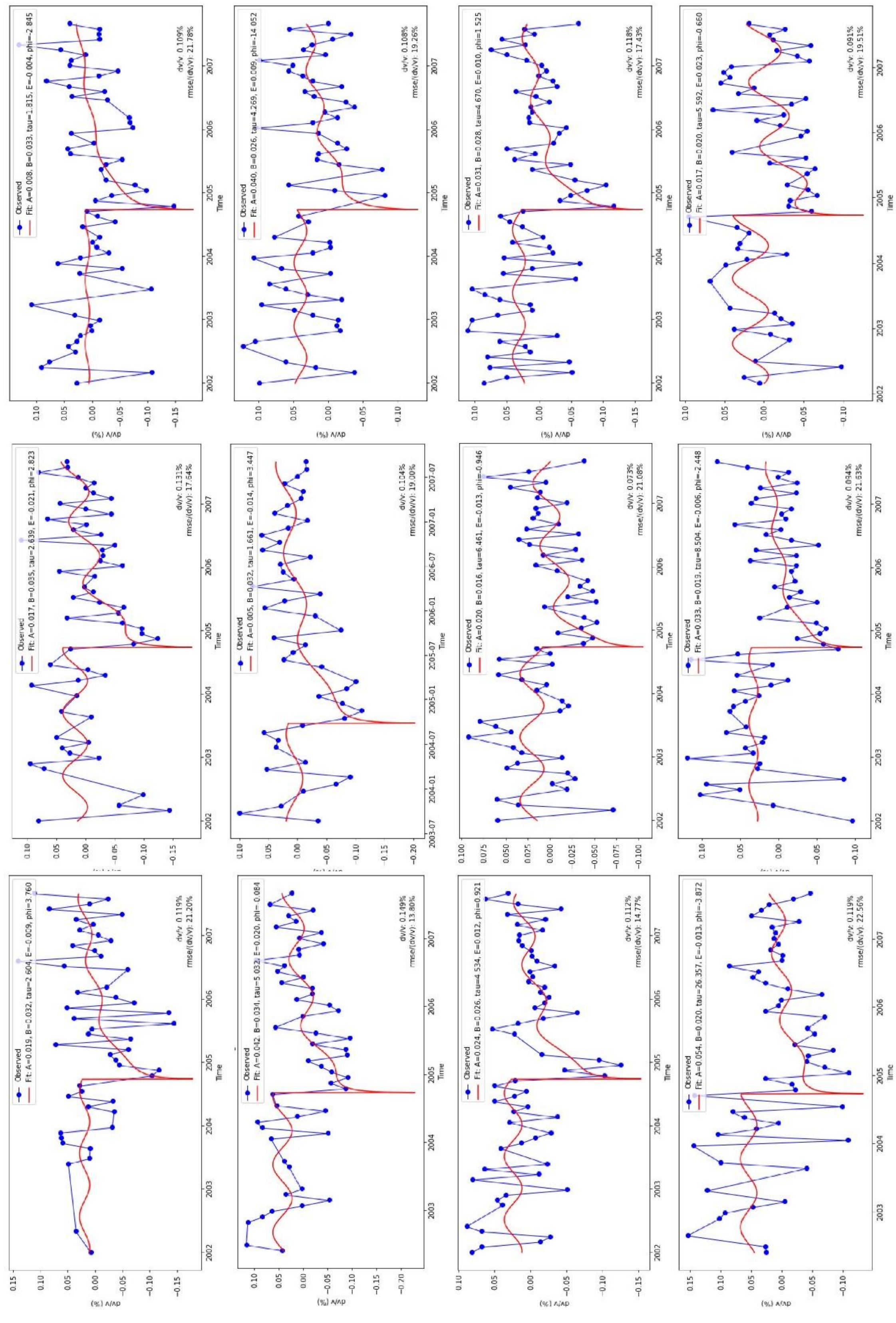
*Appendix Figure 1. Selected single station pair dv/v measurements at frequency band **0.3-0.8** Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*


Observed
Fit: A=0.041, B=0.026, tau=9.858, E=-0.006, phi=-2.489
dv/v: 0.128%
rmse/(dv/v): 23.91%
Observed
Fit: A=-0.010, B=0.041, tau=1.138, E=0.025, phi=-0.372
dv/v: 0.117%
rmse/(dv/v): 14.67%
Observed
Fit: A=0.026, B=0.026, tau=5.220, E=0.011, phi=1.996
dv/v: 0.115%
rmse/(dv/v): 13.68%
Observed
Fit: A=0.030, B=0.017, tau=9.419, E=-0.015, phi=2.191
dv/v: 0.084%
rmse/(dv/v): 22.90%
Observed
Fit: A=0.052, B=0.016, tau=27.479, E=0.012, phi=-1.876
dv/v: 0.099%
rmse/(dv/v): 20.05%
Observed
Fit: A=0.047, B=0.024, tau=15.918, E=0.033, phi=-1.689
dv/v: 0.134%
rmse/(dv/v): 19.42%
Observed
Fit: A=0.017, B=0.028, tau=3.091, E=0.024, phi=-2.176
dv/v: 0.108%
rmse/(dv/v): 17.68%
Observed
Fit: A=0.016, B=0.032, tau=2.524, E=-0.011, phi=-2.936
dv/v: 0.117%
rmse/(dv/v): 15.02%
Observed
Fit: A=0.021, B=0.020, tau=9.256, E=-0.014, phi=-2.422
dv/v: 0.097%
rmse/(dv/v): 17.99%
Observed
Fit: A=0.025, B=0.028, tau=3.972, E=-0.018, phi=0.810
dv/v: 0.115%
rmse/(dv/v): 12.15%
Observed
Fit: A=0.035, B=0.009, tau=75.124, E=-0.017, phi=0.646
dv/v: 0.064%
rmse/(dv/v): 28.77%
Observed
Fit: A=0.028, B=0.021, tau=5.324, E=0.028, phi=4.587
dv/v: 0.093%
rmse/(dv/v): 23.42%
dv/v (%)
Time
2002
2003
2004
2005
2006
2007
2004-01
2004-07
2005-01
2005-07
2006-01
2006-07
2007-01
2007-07

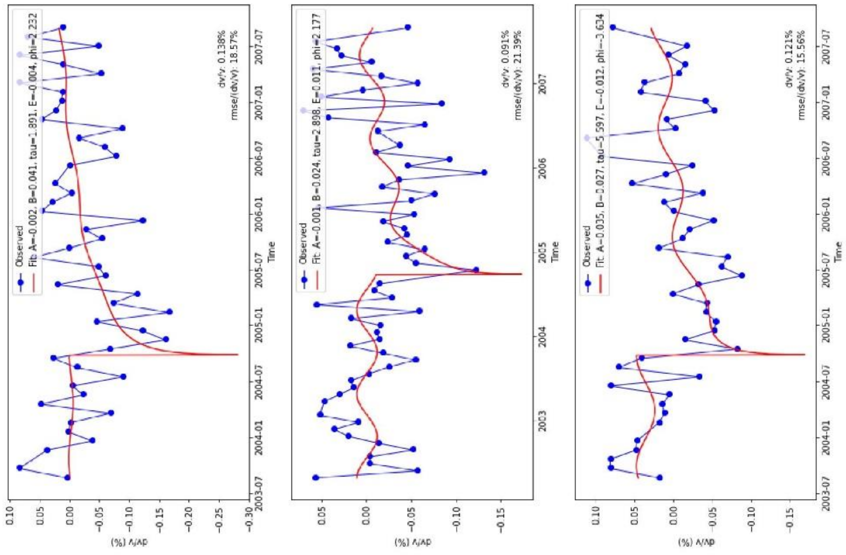


Appendix Figure 2. Selected single station pair dv/v measurements at frequency band ***0.5-1.0*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*

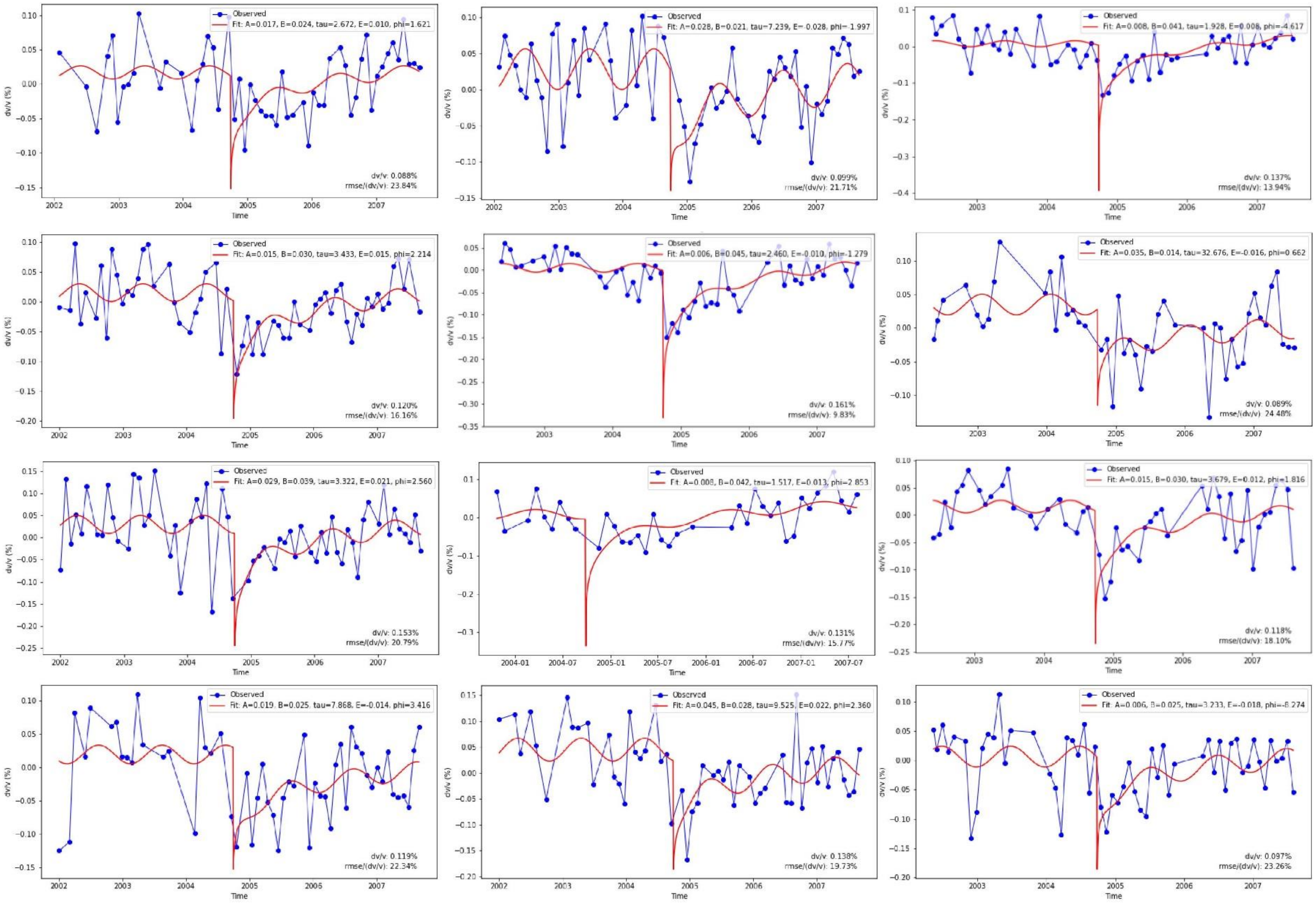

Observed
Fit: A=0.029, B=0.026, tau=5.266, E=-0.013, phi=3.155
dv/v: 0.112%
rmse/(dv/v): 13.23%
Observed
Fit: A=0.021, B=0.034, tau=2.265, E=0.005, phi=-3.482
dv/v: 0.118%
rmse/(dv/v): 17.09%
Observed
Fit: A=0.027, B=0.011, tau=20.337, E=0.010, phi=0.578
dv/v: 0.061%
rmse/(dv/v): 31.37%
Observed
Fit: A=0.042, B=0.041, tau=3.668, E=-0.011, phi=-3.565
dv/v: 0.164%
rmse/(dv/v): 11.80%
Observed
Fit: A=0.012, B=0.042, tau=2.156, E=0.017, phi=4.579
dv/v: 0.147%
rmse/(dv/v): 15.86%
Observed
Fit: A=0.014, B=0.027, tau=3.213, E=0.021, phi=0.377
dv/v: 0.106%
rmse/(dv/v): 17.17%
Observed
Fit: A=0.005, B=0.024, tau=2.275, E=0.010, phi=-1.824
dv/v: 0.084%
rmse/(dv/v): 23.09%
Observed
Fit: A=0.006, B=0.037, tau=2.081, E=-0.015, phi=3.992
dv/v: 0.127%
rmse/(dv/v): 14.84%
Observed
Fit: A=0.014, B=0.025, tau=3.335, E=-0.013, phi=-4.230
dv/v: 0.096%
rmse/(dv/v): 17.33%
Observed
Fit: A=0.033, B=0.043, tau=3.668, E=0.015, phi=3.159
dv/v: 0.174%
rmse/(dv/v): 16.99%
Observed
Fit: A=0.034, B=0.020, tau=11.438, E=-0.014, phi=3.875
dv/v: 0.102%
rmse/(dv/v): 22.17%
Observed
Fit: A=0.063, B=0.027, tau=19.609, E=0.023, phi=3.435
dv/v: 0.154%
rmse/(dv/v): 13.79%
dv/v (%)
Time
2002
2003
2004
2005
2006
2007
2004-01
2004-07
2005-01
2005-07
2006-01
2006-07
2007-01
2007-07

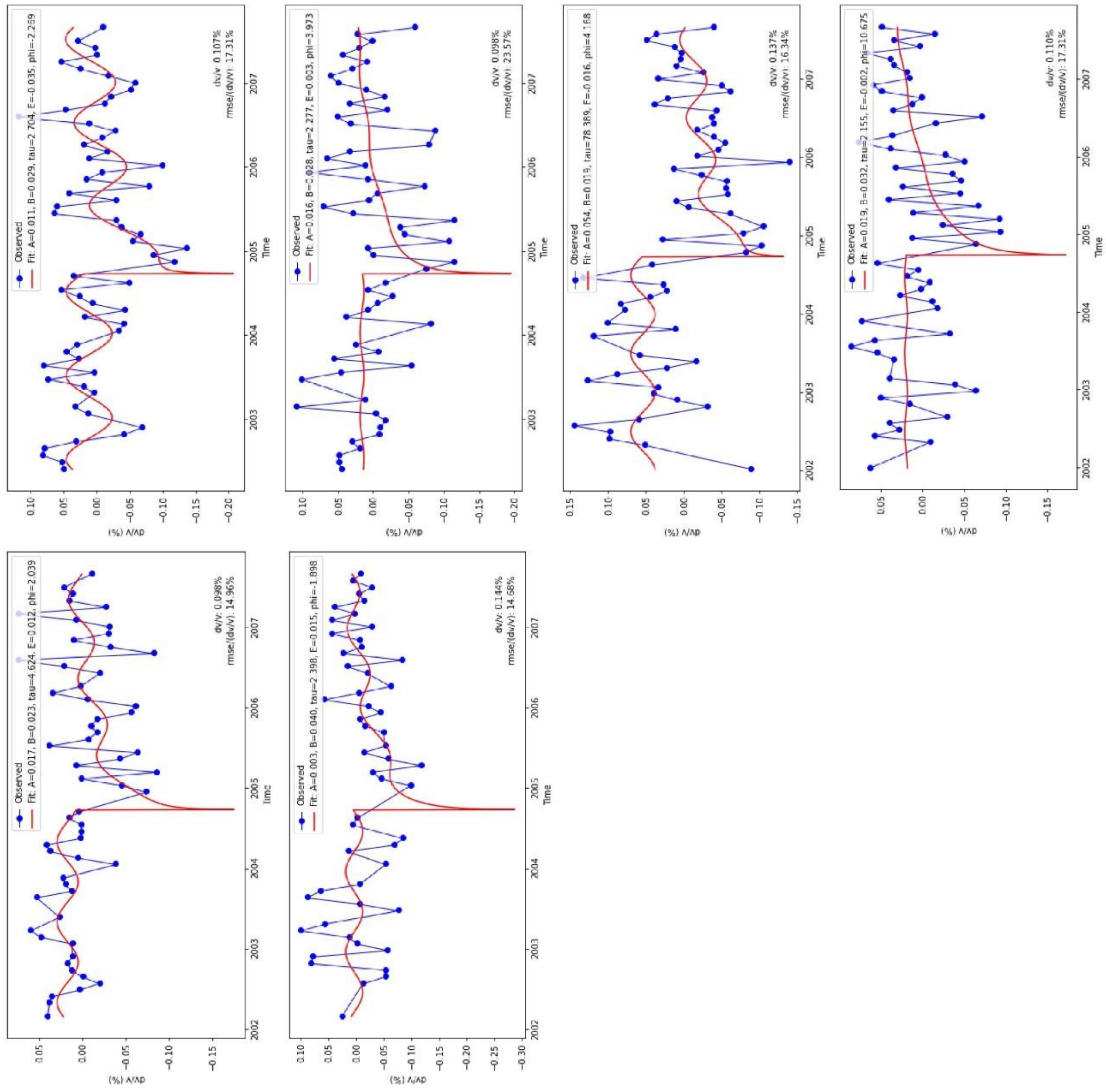


Observed
Fit: A=0.008, B=0.033, tau=1.815, E=-0.004, phi=-2.845
dv/v: 0.109%
rmse/(dv/v): 21.78%
Observed
Fit: A=0.040, B=0.026, tau=4.269, E=0.009, phi=-14.052
dv/v: 0.108%
rmse/(dv/v): 19.26%
Observed
Fit: A=0.031, B=0.028, tau=4.670, E=0.010, phi=1.525
dv/v: 0.118%
rmse/(dv/v): 17.43%
Observed
Fit: A=0.017, B=0.020, tau=5.592, E=0.023, phi=-0.660
dv/v: 0.091%
rmse/(dv/v): 19.51%
Observed
Fit: A=0.017, B=0.035, tau=2.639, E=-0.021, phi=2.823
dv/v: 0.131%
rmse/(dv/v): 17.54%
Observed
Fit: A=0.005, B=0.032, tau=1.661, E=-0.014, phi=3.447
dv/v: 0.104%
rmse/(dv/v): 19.00%
Observed
Fit: A=0.020, B=0.016, tau=6.461, E=-0.013, phi=-0.946
dv/v: 0.073%
rmse/(dv/v): 21.08%
Observed
Fit: A=0.033, B=0.013, tau=8.504, E=-0.006, phi=-2.448
dv/v: 0.094%
rmse/(dv/v): 21.53%
Observed
Fit: A=0.019, B=0.032, tau=2.604, E=-0.009, phi=3.760
dv/v: 0.119%
rmse/(dv/v): 21.20%
Observed
Fit: A=0.042, B=0.034, tau=5.032, E=-0.020, phi=-0.084
dv/v: 0.149%
rmse/(dv/v): 13.80%
Observed
Fit: A=0.024, B=0.026, tau=4.534, E=0.012, phi=0.921
dv/v: 0.112%
rmse/(dv/v): 14.77%
Observed
Fit: A=0.054, B=0.020, tau=26.357, E=-0.013, phi=-3.872
dv/v: 0.119%
rmse/(dv/v): 22.56%
dv/v (%)
Time



Appendix Figure 3. Selected single station pair dv/v measurements at frequency band ***0.7-1.2*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*


Observed
Fit: A=0.017, B=0.024, tau=2.672, E=0.010, phi=1.621
dv/v: 0.088%
rmse/(dv/v): 23.84%
Observed
Fit: A=0.028, B=0.021, tau=7.239, E=-0.028, phi=-1.997
dv/v: 0.099%
rmse/(dv/v): 21.71%
Observed
Fit: A=0.008, B=0.041, tau=1.928, E=0.008, phi=-4.617
dv/v: 0.137%
rmse/(dv/v): 13.94%
Observed
Fit: A=0.015, B=0.030, tau=3.433, E=0.015, phi=2.214
dv/v: 0.120%
rmse/(dv/v): 16.16%
Observed
Fit: A=0.006, B=0.045, tau=2.460, E=-0.010, phi=-1.279
dv/v: 0.161%
rmse/(dv/v): 9.83%
Observed
Fit: A=0.035, B=0.014, tau=32.676, E=-0.016, phi=0.662
dv/v: 0.089%
rmse/(dv/v): 24.48%
Observed
Fit: A=0.029, B=0.039, tau=3.322, E=0.021, phi=2.560
dv/v: 0.153%
rmse/(dv/v): 20.79%
Observed
Fit: A=0.008, B=0.042, tau=1.517, E=0.013, phi=2.853
dv/v: 0.131%
rmse/(dv/v): 15.77%
Observed
Fit: A=0.015, B=0.030, tau=3.679, E=0.012, phi=1.816
dv/v: 0.118%
rmse/(dv/v): 18.10%
Observed
Fit: A=0.019, B=0.025, tau=7.868, E=-0.014, phi=3.416
dv/v: 0.119%
rmse/(dv/v): 22.34%
Observed
Fit: A=0.045, B=0.028, tau=9.525, E=0.022, phi=2.360
dv/v: 0.138%
rmse/(dv/v): 19.73%
Observed
Fit: A=0.006, B=0.025, tau=3.233, E=-0.018, phi=-8.274
dv/v: 0.097%
rmse/(dv/v): 23.26%
dv/v (%)
Time



Appendix Figure 4. Selected single station pair dv/v measurements at frequency band ***0.9-1.4*** *Hz. The blue circles are observed values. The red curve represents the best fitting curve. The parameters of the best fitting curve are shown in the label.*