2015-12-14

Innovations in Seismic Tomography, Their Applications and Induced Seismic Events in Carbon Sequestration

Peng Li
University of Miami, pli@rsmas.miami.edu

UNIVERSITY OF MIAMI

INNOVATIONS IN SEISMIC TOMOGRAPHY, THEIR APPLICATIONS AND INDUCED SEISMIC EVENTS IN CARBON SEQUESTRATION

By

Peng Li

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2015

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

INNOVATIONS IN SEISMIC TOMOGRAPHY, THEIR APPLICATIONS AND INDUCED SEISMIC EVENTS IN CARBON SEQUESTRATION

Peng Li

Approved:

Guoqing Lin Ph.D. Associate Professor of Marine Geology and Geophysics	Falk Amelung Ph.D. Professor of Marine Geology and Geophysics
Shimon Wdowinski Ph.D. Research Associate Professor of Marine Geology and Geophysics	Timothy Dixon Ph.D. Professor of School of Geosciences University of South Florida
Stephen McNutt Ph.D. Professor of School of Geosciences University of South Florida	Dean of the Graduate School

LI, PENG (Ph.D., Marine Geology and Geophysics)
Innovations in Seismic Tomography, Their Applications and Induced Seismic Events in Carbon Sequestration
(December 2015)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Guoqing Lin.
No. of pages in text. (116)

This dissertation presents two innovations in seismic tomography and a new discovery of induced seismic events associated with CO_2 injection at an Enhanced Oil Recovery (EOR) site. The following are brief introductions of these three works. The first innovated work is adaptive ambient seismic noise tomography (AANT). Traditional ambient noise tomography methods using regular grid nodes are often ill posed because the inversion grids do not always represent the distribution of ray paths. Large grid spacing is usually used to reduce the number of inversion parameters, which may not be able to solve for small-scale velocity structure. We present a new adaptive tomography method with irregular grids that provides a few advantages over the traditional methods. First, irregular grids with different sizes and shapes can fit the ray distribution better and the traditionally ill-posed problem can become more stable owing to the different parametrizations. Second, the data in the area with dense ray sampling will be sufficiently utilized so that the model resolution can be greatly improved. Both synthetic and real data are used to test the newly developed tomography algorithm. In synthetic data tests, we compare the resolution and stability of the traditional and adaptive methods. The results show that adaptive tomography is more stable and performs better in improving the resolution in the area with dense ray sampling. For real data, we extract the ambient noise signals of the seismic data near the Garlock Fault

region, obtained from the Southern California Earthquake Data Center. The resulting group velocity of Rayleigh waves is well correlated with the geological structures. High velocity anomalies are shown in the cold southern Sierra Nevada, the Tehachapi Mountains and the Western San Gabriel Mountains. The second innovated work is local earthquake tomography with full topography (LETFT). In this work, we develop a new three-dimensional local earthquake tomography algorithm with the inclusion of full topography that is integrated from the Digital Elevation Model data. We present both synthetic and real data tests based on the compressional (P) wave arrival time data for Kīlauea volcano in Hawai‘i. A total of 33,768 events with 515,711 P-picks recorded by 35 stations at the Hawaiian Volcano Observatory are used in these tests. The comparison between the new and traditional methods based on the synthetic test shows that our new algorithm significantly improves the accuracy of the velocity model, especially at shallow depths. In the real data test, the P-wave velocity model of Kīlauea shows some intriguing features. Velocity decrease from the surface to 2 km depth beneath Kīlauea caldera indicates a state change of the basalt. Low velocity zones beneath Pu'u'Ō'ō, Heiheiahulu and the Hilina fault system between 5 and 12 km are possible partial melting zones. High velocity anomalies are resolved below 6 km depth beneath the summit caldera, which may suggest the presence of consolidated gabbro-ultramafic cumulates. In the third work, we installed three broadband seismic stations (Test1, Test2 and Test3) in an Enhanced Oil Recovery field to monitor the potential seismic events associated with CO_2 injection. In the two years of continuous seismic data between October 2011 and October 2013, we observed a type of long duration (LD) events instead of typical micro earthquakes, with an average daily rate of 12. The LD events have the following characteristics: (1) their duration varies from $\sim$30 to $\sim$300 sec; (2) the amplitude changes smoothly from the beginning to the end of the LD event window; (3) they are

local seismic events and were not recorded by regional seismic stations (e.g., $\sim$200 km away); (4) the waveforms are very different from those of typical earthquakes, but similar to volcanic tremors; (5) the frequency content is mainly concentrated between 0.5 and 6 Hz, which is similar to the frequency band of volcanic tremors; and (6) the source of the LD event is not a single source and could migrate to complex fractures. We picked the LD events in the two-year time period, calculated their daily rate, and compared the results with the reservoir pressure data measured in the north block. The LD event daily rates of Test1 and Test2 have a similar variation pattern as the reservoir pressure. The peak of the LD event daily rate at Test2 is about two months delayed from the peak of the pressure, whereas the LD event daily rate at Test3 does not show similar pattern. We interpret that this is because Test3 is located in the south block and a sealing fault blocks the migration of the injected CO_2 from the north to the south block.

To my families

Acknowledgments

First I would like to express my most sincere gratitude to my advisor, Guoqing Lin, for her guidance and patience on my research. She always encouraged me to think independently and supported my immature idea, which made me understand how to conduct good geophysics research. I appreciate her insightful comments and suggestions. I am thankful for all her contributions of time, ideas and funding to make my Ph.D. experience productive.

Many thanks to my committee members: Falk Amelung, Shimon Wdowinski, Timothy Dixon and Stephen McNutt. They gave me a lot of valuable suggestions on research, presentations and papers. I also appreciate Timothy Dixon and Stephen McNutt's help for my life at Tampa and Miami.

I would also like to thank all the fellow people at Marine Geosciece Department, RSMAS and all my friend at UM and USF: Qiong Zhang, Shuangyu Ge, Bateer Wu, Wenliang Zhao, Heresh Fattahi, Emanuelle A. Feliciano, Fernando Greene, Talib Oliver, Daro Solano, Anieri Morales, Emre Havazli, Zhixuan Feng, Jian Zhao, Peng Wang, Chuntao Lu, Yang Liu, Xiao Du, Yan Wang, Honghai Zhang, Wenge Wei, Meng Wang, Jie Zhang, Chi Zhang, Rui Ran and many others. They made my life colorful for the last five years.

Lastly, I owe thanks to my family for their love and trust. My parents and my sister always supported me on all my pursuits. I treasure their unconditional love so much. And most of all, I would like to thank my loving, encouraging and patient wife Qian Yang. I appreciate her love, support on both my life and study during this Ph.D., which is so important to me.

My Ph.D. study and research have been supported by National Science Foundation and Department of Energy.

PENG LI

University of Miami

Fall 2015 graduation

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the last several decades, seismology has been developed rapidly, which benefits from the high growth of the performance of computers and the improvement of observation, for instance, the deployment of broadband seismometers. With abundance of seismic data and the high speed computers, new techniques are needed to improve the results in the seismic data processing and new information is able to be discovered in the high quality seismic data. In this dissertation, we develop two techniques in seismic tomography to obtain more reliable and stable results. We also observe a new type induced seismic event associated with CO_2 injection at an Enhanced Oil Recovery (EOR) site, which has much longer durations than the typical induced earthquakes and is called a long duration event. The following are brief introductions of the backgrounds of the three works.

1.1 An overview of ambient seismic noise tomography

Ambient seismic noise is a kind of continuous background noise in seismograms, whose sources are most likely from the ocean [Friedrich et al., 1998; Roult and Crawford, 2000; Rhie and Romanowicz, 2004; Webb, 2007]. It is thought to be caused by the shallow ocean waves interacting with the Earth's crust around the coast area and then

scatter when the waves propagate in the Earth's crust [Stehly et al., 2006; Yang and Ritzwoller, 2008]. Ambient seismic noise used to be treated as noise comparing to the seismic signals from the earthquakes. However, it is found that the cross-correlation of the ambient seismic noise record of two seismic stations could estimate the empirical Green's function between the two stations [e.g., Lobkis and Weaver, 2001; Shapiro and Campillo, 2004; Snieder, 2004; Larose et al., 2005]. So it can be used as a kind of signals for exploring shallow Earth structure in the past few years [e.g., Sabra et al., 2005; Shapiro et al., 2005; Brenguier et al., 2007; Liang and Langston, 2008; Lin et al., 2008], which is called ambient seismic noise tomography. This works because ambient seismic noise takes on the information of the Earth when it propagates in the Earth and is recorded by seismic stations. The advantages of ambient seismic noise include the continuous, globally-distributed sources and high energy in short-period surface waves, which are useful to constrain crustal structure [Yang et al., 2007]. Traditional ambient noise tomography uses regular blocks or grids, such as rectangles or squares, to invert two-dimensional (2-D) group or phase velocity maps. In most cases, the grid distribution is homogeneous in the entire or part of the study area. This kind of parameterization is straightforward and easy to implement in tomographic algorithms, but there are some limitations. First, since the station and source distributions are usually spatially inhomogeneous, the grids or blocks from traditional parameterizing methods are not correlated with the distribution of data. The tomography problem often becomes ill-conditioned and the results are not stable.

To deal with this problem, a large grid spacing is usually used to reduce the number of parameters. Damping and smoothing parameters are also applied to improve the stability of the inversion problem. In this case, small-scale velocity structures may not be well constrained or not recovered at all [Sambridge and Faletič, 2003]. Second, due

to the fixed shape and size of the regular grids or blocks, the traditional tomography method may perform poorly in recovering the shape of the velocity anomaly or the interface of the structure. To overcome the above problems of traditional tomography, irregular parameterization methods have been deployed into tomographic problems. One strategy of irregular parameterization is based on rectangles or squares. One idea that the grid configuration can be optimized by adjusting the positions of the grid nodes from an initial regular rectangle grid through a generalized curvilinear coordinate basis [e.g., Michelini, 1993, 1995]. Vesnaver [1996] adopted a method that eliminates uncrossed pixels or moves the pixel boundaries to make the problem overdetermined. Spakman and Bijwaard [2001] developed an automatic algorithm based on constructing larger cells in poorly sampled areas from original small cells. Another strategy of irregular parameterization is to use Delaunay triangles or a Voronoi diagram [Fortune, 1992], which are flexible in either shape or size. Sambridge and Gudmundsson [1998] proposed an algorithm of constructing the irregular parameters by Delaunay triangles and a Voronois diagram. Böhm et al. [2000] constructed irregular cells by eliminating the parameters with null space energy. Zhang and Thurber [2005] applied tetrahedral and Voronoi diagrams to body wave tomography. When rectangles or squares are applied to irregular parameterization, the homogeneous property will be kept locally due to the symmetry. On the contrary, triangles that are flexible in shape and size are able to fit the irregular ray distribution better. Moreover, regarding the property of maximizing the minimum angle over all triangulations, Delaunay triangles avoid the appearance of thinning triangles and keep the properties of triangles as well [De Berg et al., 2000].

We present a new adaptive tomography algorithm by constructing irregular grids based on Delaunay triangles and apply it to ambient noise data. The algorithm starts with an initial sparse regular grid, and then automatically optimizes the grid through

eliminating nodes in areas of low ray density and adding new nodes in area of high ray density. The coordinates of the new nodes are determined by weights of ray density of vertices in Delaunay triangles. The final new irregular grid is used as the inversion parameter for seismic tomography. Both synthetic and real data are used to evaluate our algorithm. In the synthetic data tests, we compare the performances of our new adaptive tomography algorithm with a traditional regular-grid based tomography algorithm when the data distribution is uneven and the velocity structure contains irregular anomalies in different scales. In real data application, we obtain a 2-D group velocity map of Rayleigh waves from ambient noise data recorded by broadband seismic stations in southern California. The new adaptive ambient noise tomography and its application to the Garlock fault area, southern California is presented in Chapter 2.

1.2 An overview of local earthquake tomography

Local earthquake tomography (LET) usually refers to seismic tomography using local earthquake data [e.g., Aki and Lee, 1976; Aki et al., 1977; Foulger and Toomey, 1989]. It has been widely used to solve velocity structure of the Earth's crust and uppermost mantle in local or regional scales (from tens to hundreds of kilometers) [e.g., Thurber, 1984; Eberhart-Phillips, 1990; Rowan and Clayton, 1993; Zhang and Thurber, 2003; Lin et al., 2007b; Park et al., 2009]. Since earthquake locations and origin times are coupled with velocity structure, these parameters are often solved simultaneously [e.g., Crosson, 1976; Bijwaard et al., 1998]. In LET, a study area is usually discretized by a series of cuboids without considering topography and the velocity structure is represented by the cuboids with constant velocity (Blocks method) [e.g., Aki et al., 1977; Humphreys and Clayton, 1988; Benz et al., 1992; Boschi and Dziewonski, 1999] or the vertices of cuboids (Nodes method) [e.g., Thurber, 1983; Eberhart-Phillips, 1986; Zhao

and Kanamori, 1992; Steck et al., 1998]. If the nodes method is used, interpolation of velocity at the nodes is needed to obtain the velocity structure of the whole study area. However, during the tomographic inversions in local scale, the neglect of topography may introduce errors in ray tracing and construction of tomographic equations, especially in areas with large topographic variations, such as volcanic or mountain areas. To reduce the effect of topography, traditional LET methods usually add extra layers above the surface and include station elevations in ray tracing [e.g., Thurber, 1992; Evans et al., 1994; Gentile et al., 2000; Karastathis et al., 2011]. However, this approximation can be problematic sometimes as rays may travel into the air and velocities above the surface may affect the construction of ray paths and inverse equations.

To overcome this problem, we develop a new LET algorithm including full topography, which we refer to as local earthquake tomography with full topography (LETFT). We test our algorithm in Kīlauea volcano, Hawai‘i using both P- and S-wave arrival times of local earthquakes between 1992 and 2009 recorded by the Hawaiian Volcano Observatory (HVO) seismic stations. We first conduct a synthetic data test to evaluate the new algorithm and compare it with the traditional method with only the inclusion of station elevations. We then apply the new algorithm to the real seismic data in Kīlauea volcano and invert for three-dimensional (3-D) P- and S-wave velocity models. The new velocity model reveals several notable features related with the geologic structures and magma activities in the study area. The new LETFT method and real data application in Kīlauea volcano, Hawai‘i are presented in Chapter 3.

1.3 Induced long duration (LD) seismic events

Injecting CO_2 into the underground and storage in oil and gas reservoirs is a practicable and economic way for carbon sequestration [e.g., Busch et al., 2008]. One of

the crucial points of carbon sequestration is the risk evaluation. There are two possible risks: One is leakage of CO_2 and the other is whether it could cause natural hazards, such as ground deformation or micro-earthquakes. The reason is that when CO_2 is injected into the reservoir, the pressure increase within the reservoir will cause ground deformation [e.g., Davis and Pennington, 1989; Vasco et al., 2010; Morris et al., 2011; Karegar et al., 2015; Yang et al., 2015]. Injected CO_2 is also possible to migrate around, leak from an existing fault [Dewhurst and Hennig, 2003], and even cause micro-earthquakes [e.g., Ake et al., 2005; Lei et al., 2008; Khatiwada et al., 2012; Mazzoldi et al., 2012; Martínez-Garzón et al., 2013; Keranen et al., 2013; Verdon, 2014; McGarr et al., 2015]. However, all these effects strongly depend on the local geological characteristics. Therefore, geophysical or other monitoring at each site is of great importance for carbon sequestration.

One of the most important geophysical monitoring techniques is seismic monitoring. It is likely to observe seismic events induced by the migration and leakage of the injected CO_2, which is very important for the safety evaluation of CO_2 storage. Seismic monitoring usually uses geophones or other narrow-band seismometers deployed on the surface of the Earth or in a well [e.g., Rutledge and Phillips, 2003; Bohnhoff and Zoback, 2010], which makes it difficult to record seismic events at lower frequencies. Broadband seismometers can record seismic information in a very wide band of frequency, which can provide much more information for the understanding of the characteristics of the induced seismic events. In previous research of seismic monitoring, most work are about the induced micro earthquakes [e.g., Talebi and Boone, 1998; Baisch and Harjes, 2003; Shapiro et al., 2007; Eisner et al., 2010; Zoback and Gorelick, 2012; Ellsworth, 2013; Sharma et al., 2013]. However, the geological characteristics of the oil and gas reservoir are very complex and include the involvement of fluids. So

it is possible to generate other types of seismic events with more complicated mechanisms. There are only a few works about long duration (LD) seismic events during hydraulic fracturing in shale-gas and tight-gas reservoirs [Das and Zoback, 2013; Eaton et al., 2013] and Enhanced Oil Recovery (EOR) fields [Martínez-Garzón et al., 2013] with less than a month of observations. With the long period of observations and taking advantage of broadband seismic stations, there would be more information to analysis the LD events.

In this study, we deployed three broadband seismic stations from October 2011 to October 2013 in an Enhanced Oil Recovery (EOR) field near Houston, TX. We observe a kind of long duration (LD) event in the seismic data with smooth head and tail. The LD events last from $\sim$30 sec to $\sim$300 sec with frequency of 0.5-6 Hz. In Chapter 4, we will analyze the characteristics of the LD events and compare the rate of LD events to the reservoir pressure in order to reveal more information about the LD events. The investigation of the LD events is presented in Chapter 4.

Chapter 2

Adaptive grid ambient noise tomography

2.1 Overview

Ambient noise signal is a continuous and global distributed seismic signal. It can be used to image the subsurface seismic velocity structures regardless of area with or without earthquakes. Traditional ambient noise tomography method discretized the study area by a series of square or rectangle nodes or cells. This configuration could cause the tomographic process instable. To solve this problem, we developed an adaptive grid method for ambient noise tomography, in which the grid nodes distribution fits the data distribution. As a result, the tomographic process becomes more stable and the resolution of the area with dense data is also improved. In the following, we will talk about the details of the adaptive grid ambient noise tomography. Then we conduct a synthetic test to make a comparison between traditional and adaptive grid ambient noise tomography. After the synthetic test, the new method is applied to the Garlock fault area, southern California to investigate the group velocity structure of Rayleigh wave.

2.2 Method

Ambient seismic noise likely origins from the ocean, then is diffused when the seismic waves encounter the numerous scatters in continent. Finally the seismic signals from these scatters are recorded by seismic stations as continuous daily background noise (see Figure 2.1). The basic idea of ambient noise tomography is extracting the Green's function between two stations by cross-correlation of the daily ambient noise signals recorded by the two stations. Then the Rayleigh wave part (or other phase) of the cross-correlation signals are used to invert the subsurface velocity structure. Ambient noise tomography includes three major parts, which are ambient seismic noise data processing, model parameterization and tomographic procedure to obtain the velocity model (Figure 2.2). The red dashed line in Figure 2.2 encloses the data processing part. The purpose of this part is extracting the ambient seismic noise signal. To be specific, we firstly obtain the ambient noise signal of station pairs by using cross-correlation of the temporal records of the station pairs. Then we calculate the group velocity of Rayleigh wave at different periods, which will be used as input data in the tomographic procedure. The reason we only use the Rayleigh wave part of the ambient noise signal is that the amplitude of the Rayleigh wave is larger than other phases and it has a good signal-to-noise ratio. The model parameterization discretizes the study area by a series of nodes [Kissling et al., 2001]. Because of the limit of data and the complexity of the subsurface structures, it is hard to represent the study area by unlimited or continuous parameters. In our method, the model parameters are automatically generated according to the distribution of data (or ray paths). The greatest advantage of this method is that it could improve the stability of the tomography and the resolution at the area with dense data. In the tomographic part, we use the group velocity of Rayleigh waves and model

parameters obtained by the previous two steps as input data to iteratively inverse for the two dimensional (2D) group velocity map of Rayleigh wave at different periods. As is typical for seismic tomography, it contains a forward part to calculate ray paths and a partial derivative matrix and inverse part to update the velocity model. In the forward part, we apply a finite difference method which is called Fast Marching method (FMM) to solve the Eikonal equation and then calculate the ray path. In the inverse part, Conjugate Gradient Least Square (CGLS) method is used to solve the linear equation for its high efficiency.

2.2.1 Ambient seismic noise data processing

There are three main steps in the data processing, (1) single station preparation, (2) ambient noise signal extraction, and (3) group velocity calculation. In step 1, we apply basic seismic data processing to each single station, including removals of instrument response, mean, linear trend and band-pass filter between 0.5 and 150 sec, which is wider than the period range we focus on. Time domain normalization is used to reduce the effect of earthquakes and non-stationary noise sources near the stations [Bensen et al., 2007]. Spectral normalization is also applied to all the data to broaden the band of ambient noise signals in cross-correlations since ambient noise is not flat in the frequency domain [Bensen et al., 2007]. In step 2, we compute cross-correlations for each pair of stations in one day length to extract the ambient noise signals. Cross-correlation results with the signal to noise ratio greater than 15 in the study time period are stacked to magnify the ambient noise signal. The signal level is the average of the absolute amplitudes of ± 10 sec from the maximum amplitude of the main Rayleigh wave envelope. Since the signals are in the range of 0-100 seconds, the data far from this range can represent the noise level. We use the average of the absolute amplitudes between 600

and 800 seconds away from the zero point of the cross-correlation result as the noise level. Figure 2.13 shows the six-month stack of the cross-correlation results observed for all the station pairs. There are clear signals between the typical low and high group velocity lines of 2 and 4 km/s. The signals in the positive lag time part are weaker than the negative part. The reason is that the source of ambient seismic noise in southern California has a strong directional dependence, with stronger signals in the west and southwest directions than others [Yang and Ritzwoller, 2008]. Figure 2.14 is an example of the stacked cross-correlation signals. The Rayleigh wave envelope is very clear in each band. Finally, we compute mean group velocities of station pairs at different periods from the ambient noise signals by using the frequency-time analysis (FTAN) method [Levshin et al., 1992]. The FTAN method uses a set of narrow-band Gaussian filters to obtain frequency-time amplitude maps, and then trace dispersion curves of surface wave group velocities. This method has been widely used in the group velocity detections of surface waves [e.g., Ritzwoller and Levshin, 1998; Moschetti et al., 2007; Zheng et al., 2008; Zhou et al., 2012]. In this step, we select the station pairs of each period only if the interstation distances are larger than one wavelength [Brenguier et al., 2007]. For example, if the period is 5 sec, the smallest interstation distances is 20 km, which is larger than the wavelength of 5 sec period Rayleigh wave.

2.2.2 Adaptive grid generation

The work flow of our adaptive tomography procedure is illustrated in Figure 2.4. It contains two major parts, the adaptive grid generation part and the tomography part. In the adaptive grid generation part, we start with sparse regular nodes and construct Delaunay triangles based on these initial nodes. Ray tracing is done through the Delaunay triangle grid and partial derivative matrix is calculated for the linearized travel time

equations. After that, we compute the ray density of each node according to the ray paths and partial derivative matrix, based on which nodes will be added or eliminated in the next step. When the ratio between the number of nodes and the number of data is above a certain threshold, we will perform ambient noise tomography. Otherwise, we will go back to the Delaunay triangle grid set-up and repeat the above steps until the threshold is reached or the total number of nodes does not change.

We use Delaunay triangles to represent the irregular grids because of the following advantages. The first one is the property of the maximum-minimum angle that the triangulations have few long and skinny triangles [Fortune, 1992]. Thus, the interpolation of the velocities in the forward part of the tomography will be able to avoid using the nodes far away from the interpolated points. Another advantage is that the size of Delaunay triangles is determined by the distribution of the nodes, called density-dependent size [Sambridge et al., 1995]. With this property, even extremely large variations of grid sizes can be readily represented in the model. In addition, the arbitrary shapes of the triangles enable the representation of complex models even when nodal distribution is highly irregular. The adaptive nodes, correlated with the ray distribution, are generated automatically by our newly developed algorithm. We start with sparse regular or irregular grid nodes and replace the grid by Delaunay triangles constructed by the Qhull method [Barber et al., 1996; Zhang and Thurber, 2005]. The vertices of the Delaunay triangles are the grid nodes. The regular grid nodes are optimized by adding and eliminating nodes based on the ray density of each initial node, similar to Zhang and Thurber [2005], although they used derivative weight sum (DWS) values, not the ray densities. In the adding nodes part, the nodes whose ray densities are greater than a given threshold are defined as the target nodes. To keep the stability of the inversion process and sufficiently utilize the data, the threshold for adding nodes was chosen to

be 10 after several tests. The result is not significantly sensitive to this threshold in the range of 8-14. If the vertices of a Delaunay triangle have only one target node, we will not insert any new node into this triangle. Otherwise the new node would be too close to the target node according to equation (2.1). When the vertices of a Delaunay triangle have two or more target nodes, a new node will be inserted into it. The coordinate of the new node is determined by the weight of ray density of each vertex in a Delaunay triangle by using the following expression,

$$\mathbf{R} = \frac{1}{W} \sum_{i=1}^{3} w_i \mathbf{r}_i \tag{2.1}$$

where $\mathbf{R}$ is the coordinate of the new node, i is the index of each vertex of the Delaunay triangle, $\mathbf{r}_i$ is the coordinate of each vertex of the Delaunay triangle, w_i is the ray density corresponding to each vertex, and W is the sum of all three w_i. The new node will be closer to the vertex with highest ray density according to equation (2.1) so that the area with dense rays will be covered by more nodes eventually. In the eliminating nodes part, a node will be deleted if its ray density is smaller than a given threshold. Generally large thresholds would increase the ray density of each node and the stability of the tomographic inversion. However, many nodes with relatively lower ray densities would be deleted at the same time, which may reduce model resolution. Considering this trade-off, we chose 4 as the final threshold for eliminating nodes. The thresholds of ray density in adding and eliminating nodes guarantee that the adaptive grid nodes have fairly even ray densities. Usually the optimal grid nodes cannot be constructed in one iteration. Therefore, after adding and eliminating nodes we use the new irregular nodes as inputs and repeat the above steps a few times. The iteration will stop when the total number of nodes does not change or a certain threshold is obtained. This threshold is represented by the ratio of the number of nodes and the number of data, which affects

the stability of inversion and the resolution of the tomography. In order to take the finite-frequency effect into account, we also require grid spacing and interstation distance greater than one wavelength of Rayleigh waves in the periods of our interest for robust model resolution [Brenguier et al., 2007].

2.2.3 Ray tracing method

Ray tracing is needed in both the adaptive grid construction and the forward part of tomography for ray path determination and partial derivative calculation of the linearized travel time equation. In this study, the ray tracing is divided into two steps, travel time prediction and ray path construction. The travel time (or wave front) is controlled by the eikonal equation [Lay and Wallace, 1995]:

$$|\nabla T(x,y,z)| = S(x,y,z), and\, S(x,y,z) = v^{-1}(x,y,z) \tag{2.2}$$

where $T(x,y,z)$ is the travel time at location (x, y, z). $S(x,y,z)$ is the slowness at location (x, y, z), which is the reciprocal of the velocity $v(x,y,z)$. In the adaptive ambient noise tomography, since we use the Rayleigh wave which is a kind of surface wave for seismic tomography, the eikonal equation (Equation (2.2)) is reduced to its two dimension form. Here we choose the fast marching method [Sethian and Popovici, 1999] as the travel time prediction method. The fast marching method (FMM) is a finite difference method that solves the eikonal equation directly and the two dimensional form is illustrated by the following equation,

$$\max\left(D_{ij}^{-x}T, D_{ij}^{+x}T, 0\right)^2 + \max\left(D_{ij}^{-y}T, D_{ij}^{+y}T, 0\right)^2 = S_{ij}^2 \tag{2.3}$$

where $D_{ij}T$ are the gradients of the travel time T at each grid node location (x, y) in Cartesian coordinates, i and j are the node indices in x and y directions, $+x$ and $-x$ mean the derivative in the positive and negative direction of the x coordinate, respectively, $+y$ and $-y$ are the same meaning in the y coordinate, and S_{ij} is the slowness.

For each source, the FMM divides the grid nodes of the study area into three groups, alive points, narrow-band points and far-away points (Figure 2.5). Alive points are the points for which travel times have already been calculated. Narrow-band points are neighbors of alive points. Far-away points are other points in the study area. At first, only the source point is the alive point with travel time equal to 0. The travel times of narrow-band points around the sources will then be calculated according to equation (2.3). The narrow-band point that has the minimum travel time is chosen as the new alive point. After updating the set of alive points, the new narrow-band points are set to be the points that are neighbors of the new alive points, no matter whether they belong to the previous narrow-band points or far away points. The update of alive points and narrow-band points are iterated until all the points in study area are transformed to alive points, i.e., the travel times of all nodes of FMM are available. The travel times of all the source-station pairs will then be obtained through linear interpolation. The form of the travel time gradient $D_{ij}T$ is important because it affects the precision of the predicted travel times and the computational efficiency. We choose the second-order finite-difference algorithm in our calculations, which has an $O\left(h^2\right)$ error and feasible computation time (e.g., about 26 seconds with 40,000 points on a Linux system Ubuntu 10). The second-order algorithm is represented by,

$$
\begin{aligned}
D_{ij}^{-x}T &= \left(3T_{i,j} - 4T_{i-1,j} + T_{i-2,j}\right)/2h_x \\
D_{ij}^{+x}T &= \left(3T_{i+2,j} - 4T_{i+1,j} + T_{i,j}\right)/2h_x \\
D_{ij}^{-y}T &= \left(3T_{i,j} - 4T_{i,j-1} + T_{i,j-2}\right)/2h_y \\
D_{ij}^{+y}T &= \left(3T_{i,j+2} - 4T_{i,j+1} + T_{i,j}\right)/2h_y
\end{aligned}
\tag{2.4}
$$

where $D_{ij}^{-x}T$, $D_{ij}^{+x}T$, $D_{ij}^{-y}T$ and $D_{ij}^{+y}T$ are the components of the gradient in the negative and positive directions of the x and y coordinates, i and j are the x and y indices of

the node, and h_x and h_y are the grid spacings in the x and y coordinates, respectively. In the second step, for any given receiver point, ray paths can be constructed through back-propagation from the receiver point to each source point using the travel time information obtained from the first step. Since the ray path is always perpendicular to the wave front, the ray path construction can be represented by the following equation,

$$\begin{aligned} P_{x+1} &= -\alpha \left[max \left(D_{ij}^{-x} T, -D_{ij}^{+x} T, 0 \right) \right] \\ P_{y+1} &= -\alpha \left[max \left(D_{ij}^{-y} T, -D_{ij}^{+y} T, 0 \right) \right] \end{aligned} \tag{2.5}$$

The coordinate of the next point (P_{x+1}, P_{y+1}) along the ray path is determined by the gradient of the travel times of the previous point $D_{ij}T$ and the step length α. Indices i and j are the same as in equations (2.3) and (2.4). The form of the gradient $D_{ij}T$ is shown by equation (2.4). If the points along a ray path are not among the nodes of FMM, travel times and their gradients of these points will be obtained by linear interpolation from nearby points. After all the ray paths are determined, the partial derivative matrix of the linearized inversion problem is calculated by numerical integration along the ray paths.

2.2.4 Tomography part

The tomography part includes a forward part using the ray tracing method above and an inverse part. In the forward part, ray paths and partial derivatives of the linearized travel time equations are determined. In the inverse part, we adopt a damped least squares method to minimize the following misfit function [Tarantola, 2005]:

$$\Phi(\boldsymbol{m}) = ||(G\boldsymbol{m} - \boldsymbol{d}_{obs})||^2 + \lambda||(\boldsymbol{m} - \boldsymbol{m}_{prior})||^2 \tag{2.6}$$

where $\boldsymbol{d}_{obs}$ is data vector, $\boldsymbol{m}$ is model vector, $\boldsymbol{m}_{prior}$ is a priori model, G is the partial derivative matrix, and λ is the damping parameter. After one iteration of the forward

part and inverse part, the new model vector $\boldsymbol{m}$ will replace the $\boldsymbol{m}_{prior}$ in the next iteration, assuming the new model vector $\boldsymbol{m}$ is closer to the real model. We use the Conjugate Gradient Least Square method to solve equation (2.6) as it saves the memory of computer and converges fast [Aster et al., 2005].

2.3 Synthetic test

2.3.1 Model setup

We compare the new adaptive tomography method with a traditional algorithm using a two-dimensional synthetic model (Figure 2.6a). The velocity varies linearly from the upper left to the lower right corner. There are two velocity anomalies with the form of damped sinusoidal functions distributed in the high and low velocity areas, respectively. This model includes both large and small scale velocity structures, thus it is ideal to test the performances of the two algorithms on recovering the velocity structure in different scales. The numbers of the random synthetic sources and receivers are 10 and 40. The period of Rayleigh wave is set to be 8 seconds with the wavelength of $\sim$32 km (assuming a 4 km/s velocity), similar to the periods we use in the real data application. After removing the rays whose lengths are shorter than 32 km, we obtain 397 travel times used for the tomographic inversion. The minimum interstation distance is 37.4 km. In order to simulate the heterogeneity of the real data distribution in seismic tomography, we choose the non-uniform distributed sources and receivers in our tests. There are more sources and receivers in the lower right portion of the study area than other areas (Figure 2.6b).

In traditional tomography, damping parameters are usually applied to stabilize the inversion processes. The appropriate damping parameters are typically selected by using a data misfit versus model variance trade-off analysis. To find an optimal damping

parameter for the traditional tomography, we explored a wide range of damping values to make sure that we looked at the entire trade-off curve, instead of only a portion of it (Figure 2.7). We chose 100 and 700 as the damping values for our synthetic tests, which produced good compromises between data misfits and model variances. The velocity model with damping of 700 is expected to be smoother than that with damping of 100. However, a smoother model is not a necessity for tomography, since sometimes large damping parameters will artificially smooth the velocity model. In order to compare the two methods under the same conditions, we also use damping parameters of 100 and 700 for the adaptive tomography, which are in the optimal zones of their trade-off curves. The initial velocity model is the same for both methods, which is an uniform velocity model of 4 km/s.

In adaptive tomography, to better evaluate the variations of stability and resolution with different configurations of nodes, we chose two different thresholds of the upper bounds of the total number of adaptive nodes. One is 1/4 and the other is 1/3 of the total number of data. As a result, we obtain two adaptive grids with 111 nodes (referred to as adaptive model 1) and 161 nodes (referred to as adaptive model 2).

2.3.2 Tomographic results

Figure 2.8 shows the inverted velocity models from both the traditional and the new adaptive tomography algorithms with different damping parameters. Figure 2.8a-c show the traditional 9 by 9 evenly-distributed grid, adaptive grid model 1 with 111 nodes and adaptive grid model 2 with 161 nodes, respectively. The minimum node spacing for adaptive grid model 1 and 2 is about 47 km, which is larger than one wavelength of our synthetic Rayleigh wave ($\sim$32 km). As shown in Figure 2.8b-c, the grid nodes of the two adaptive models are dense in the lower right part of the study area and

sparse in the upper left part, consistent with the data distribution. In the homogeneous rectangle grids of the traditional method (Figure 2.8a), the redundant grid nodes in the area with relatively sparse data may decrease the stability of the inversion and the deficiency of grid nodes in the lower right area with dense data will result in low resolution. Figure 2.8d-f show the histograms of ray densities for each grid. The ray densities of the adaptive grids (Figure 2.8e-f) are more uniform with most densities concentrating in the range of 20 and 70. In contrast, the ray densities of the traditional regular grid spread in a wide range of 0 and 250. In addition, the regular grid has more nodes with ray densities close to zero.

Figure 2.8g-i show the tomography results for the traditional, the adaptive model 1 and model 2 with damping parameter of 100. Similar results are shown in Figure 2.8j-l for the same grid distribution, but with a different damping parameter of 700. From these results, we can see that both algorithms can recover the major structures of the synthetic velocity model, i.e., the linearly-varying velocity from the upper left to the lower right corner. However, the details of the synthetic velocity model are not recovered well in the result of the traditional algorithm. The resolution is low in the lower right area covered by dense ray paths. For instance, in Figure 2.8g, the velocity anomaly in the lower right part lacks the ring-shaped variations and is too coarse compared with the true structure of the damped sinusoidal function in this area. Another disadvantage of the traditional algorithm is that it may result in some artificial structures, most of which show in the left part of the study area, caused by the lack of data in this area. In this case, it is difficult to distinguish the real structures from the artificial effects for real data applications in the situation that we do not know the true velocity structure. For the results of the new adaptive tomography algorithm, there are much fewer artifacts in the left area than in the result of the traditional algorithm, although the resolution of

the velocity anomaly is not significantly higher than the traditional method. The reason is that there are not many sources and receivers and the grid nodes from automatic generation process are sparse in this area. If we add extra nodes here, artificial structures may appear, similar to the results of the traditional tomography algorithm. This type of low resolution in the model is caused by the lack of data and is difficult to be improved by varying tomographic methods. However, in the lower right area (Figure 2.8h-i) covered by dense ray paths, the resolutions of the adaptive models are much higher than that in the result of the traditional algorithm (Figure 2.8g) and the resolved structures are fairly close to the true velocity model. This is because the grid nodes of the adaptive method are dense in this area, which are able to sufficiently exploit information implied in the dense ray paths. When the damping parameter is increased to 700, the performances of the adaptive method are different from the traditional tomography. The result of the traditional tomography (Figure 2.8j) becomes much smoother compared to the results when using a damping parameter of 100. In contrast, the results of the adaptive tomography do not have obvious variations due to the increase of the damping parameter. This indicates that the adaptive grids may have fewer divergent nodes, which are strongly affected by damping parameters in the traditional tomography. To quantitatively evaluate the divergent nodes of traditional and adaptive grids, we calculated the singular values of the G matrix in equation (2.6) for the traditional grid, adaptive grid 1 and adaptive grid 2. The numbers of singular values smaller than 10 are 16, 1 and 3 for traditional grid, adaptive grid 1 and adaptive grid 2. The percentages of the nodes with smaller ($<$10) singular values among all the nodes are 19.8%, 0.9% and 1.8%, respectively. Thus the traditional grid has many more divergent nodes and is less stable than the adaptive grids in the inversion. When comparing the results of the two adaptive grids, we observe a better resolution in the lower right corner of the grid

model 2 (161 nodes) and the shape of the velocity anomaly is more similar to the true model. Moreover, the velocities in the lower left area (Figure 2.8i) are recovered better than grid model 1 (Figure 2.8h), where the velocities in the synthetic model are around 4 km/s.

Travel time residuals are often used to evaluate the quality of the tomographic models. We plot travel time residuals of the tomography results in Figure 2.9. Figure 2.9a-c are the residuals for the traditional tomography, adaptive grid model 1 and adaptive grid model 2 with damping parameter of 100, respectively. Figure 2.9d-f show similar results but with damping of 700. The black dashed line in each figure indicates the residuals for the initial uniform velocity model of 4 km/s, with a corresponding root mean square (RMS) of the residuals of 7.10 sec. Both the traditional and the adaptive tomography methods significantly reduce the RMS of residuals to below 1 sec. Compared with the traditional tomographic results, both adaptive tomographic model 1 and 2 result smaller RMS of residuals, which are 0.49 and 0.41 sec with damping parameter at 100, respectively. In addition, the taller and thinner shapes of the adaptive residuals in Figure 2.9b and c indicate the adaptive tomographic method fits the data better. When the damping parameter is increased to 700, the residuals are almost the same except that the RMS of the adaptive model 2 is increased by only 0.01 sec. This indicates that both the traditional and the two adaptive models fit the data equally well when the damping parameters are changed from 100 to 700.

The resolution and stability of traditional tomography may vary with grid size. To make the comparison between the traditional and the adaptive tomography more convincing, we also tested traditional tomography with different grid sizes. We chose 4 models with grid spacing of 2°, 1.25°, 0.83° and 0.43°, as shown in Figure 2.10a-d, respectively. The grid spacings of 2° and 0.43° are similar to the biggest and smallest

grid sizes in the adaptive grid, respectively. The damping parameter is chosen as 100 in this test. The corresponding residuals and inverted velocity models are shown in Figure 2.10e-h and Figure 2.10i-l. The RMS of residuals are 2.28 sec, 0.99 sec 0.74 sec and 0.64 sec in Figure 2.10e-h as the grid spacing decreases. The resolution in the lower right area with dense data is significantly improved with reduced grid spacing. For instance, the velocity anomaly in lower right area of Figure 2.10l has almost the same resolution as the velocity model of adaptive tomography in Figure 2.8h and i. However, the velocity model of small grid spacing is much more noisy than the results of large grid spacing. For instance, Figure 2.10l with grid spacing of 0.43° has more small scale artifacts than that of Figure 2.10i with the grid spacing of 2°. Our tests indicate that it is challenging for traditional tomography to obtain good model resolution and avoid inversion artifacts at the same time.

In addition, we calculated the correlation coefficients between the inverted models and the true model for both the traditional and adaptive methods in order to evaluate the model recovery. For a damping parameter of 100, the correlation coefficients are 0.59, 0.88 and 0.89 for the traditional, the adaptive model 1 and model 2, respectively. When the damping parameter is increased to 700, the correlation coefficients are changed to 0.74, 0.90 and 0.91, respectively. This indicates that the adaptive method recovers the model better than the traditional one.

2.3.3 Robustness test

In order to evaluate the robustness of the traditional and adaptive tomography methods, we add Gaussian distributed random noise to the input synthetic travel times to see how the results vary from those without noise. Gaussian distributed random noise is widely used for simulating the noise of data in seismic tomography [e.g., Ritter et al., 2001;

Li and Nowack, 2004; Matsubara et al., 2008]. In this study, the random noise and its standard deviation (STD) are represented by the form of a percentage. The noise percentage multiplied by each synthetic travel time is the real noise for each synthetic travel time. We add two different levels of Gaussian distributed random noise to the synthetic travel times. One has a STD of 1.5%, the other has a STD of 5%. A larger STD means that more synthetic travel times are added large noises. The purpose is to see the performances of the traditional and adaptive methods in different noise levels. The damping parameters chosen in both the traditional and adaptive methods are 100.

Figure 2.11 shows the inverted velocity models of the traditional method with grid distribution of Figure 2.8a and the adaptive method with grid distribution of Figure 2.8b. With the noise level of 1.5% STD, the velocity model from the traditional tomography method (Figure 2.11a) has more artifacts in the top and left parts covered by sparse data, but is better resolved in the lower right part with dense ray paths. This indicates that the stability of the traditional tomography is affected by the ray distribution in the study area due to the homogeneous grid distribution. In contrast, the velocity models from the adaptive method with grid model 1 (Figure 2.11b) has fewer artifacts and is approximately the same as the result without noise (Figure 2.8h). When the noise level is raised to 5% STD, the artifacts in both methods are increased in Figure 2.11c-d. The velocity model from the traditional tomography method (Figure 2.11c) has an artificially large and high velocity anomaly in the left central area and artificially low velocity anomalies in lower left and upper right corners. Similarly, artificially high velocity anomalies also appear in the lower left area of the velocity model from the adaptive method (Figure 2.11d), although not as intense as those in the traditional method result. In addition, there are some small artifacts in the lower right area of Figure 2.11d where the grid spacing is small.

Overall, the left and top parts of the velocity model from the traditional method is always unstable at different noise levels, whereas the lower right part is more stable due to the dense rays. The adaptive method performs well at lower noise level, consistent with the fact that the adaptive grid has much fewer divergent nodes, however, some artifacts also appear when noise level increases.

2.4 Application of adaptive grid ambient noise tomography at Garlock fault area, California

2.4.1 Geological background

We apply the adaptive ambient noise tomography method to the Garlock Fault (Figure 2.12), which is a northeast-southwest left-lateral strike-slip fault in southern California [Hamilton, 1969; McGill and Sieh, 1993; Meade and Hager, 2005]. It marks the boundary between the Sierra Nevada in the north and the Mojave block in the south [Davis and Burchfiel, 1973; Petersen and Wesnousky, 1994]. The Sierra Nevada is a rigid continental block [Wright, 1976; Dixon et al., 2000; Lutter et al., 2004; Yan and Clayton, 2007], whereas the Mojave block is crossed by many Quaternary northwest-striking right-slip faults [Garfunkel, 1974; Dokka and Travis, 1990; Petersen and Wesnousky, 1994]. The west end of the Garlock Fault is connected to the right-lateral San Andreas Fault [Dolan et al., 2007]. The San Joaquin Valley, filled with thick Cenozoic sediments [Kilkenny, 1951; Ingersoll, 1982; Goodman and Malin, 1992], is located to the west side of the Sierra Nevada. The area to the southwest side of the San Andreas Fault is dominated by the Transverse Ranges such as the west San Gabriel mountains. The diversity of the geological structure in the Garlock Fault zone makes it a perfect place to test our new algorithm. The grey polygon in Figure 2.12 is our study area after we blanked out the area in the corner, which is not covered by stations.

2.4.2 Data set

The seismic data used in this study are obtained from the Southern California Earthquake Data Center between January 2004 and July 2004 recorded by 14 broadband stations of the Southern California Seismic Network (blue triangles in Figure 2.12). The interstation distance ranges from 25 to 181 km. Ambient noise signals can be recorded at any time [Yang and Ritzwoller, 2008], although ambient noise signals must be extracted by cross-correlation between stations that share seismic records during the same time period. Therefore, any times of the seismic records are acceptable for ambient noise tomography as long as all the stations used in the tomography have continuous records in the same time span. In addition, six months of data are sufficient for the purpose of this study. For reference, Shapiro et al. [2005] used 2 months of data and Saygin and Kennett [2010] chose the durations of data between 15 days and several years. We only use the vertical component of the seismic records to focus on the 5-11 seconds of short period Rayleigh wave part of the ambient noise signal.

2.4.3 Data processing and results

There are three main steps in the data processing, (1) single station preparation, (2) ambient noise signal extraction, and (3) group velocity calculation. In step 1, we apply basic seismic data processing to each single station, including removals of instrument response, mean, linear trend and band-pass filter between 0.5 and 150 sec, which is wider than the period range we focus on. Time domain normalization is used to reduce the effects of earthquakes and non-stationary noise sources near the stations [Bensen et al., 2007]. Spectral normalization is also applied to all the data to broaden the band of ambient noise signals in cross-correlation since ambient noise is not flat in the frequency domain [Bensen et al., 2007]. In step 2, we compute cross-correlations for each

pair of stations in one day length to extract the ambient noise signals. Cross-correlation results with the signal-to-noise ratio greater than 15 in the study time period are stacked to magnify the ambient noise signal. The signal level is the average of the absolute amplitudes of ± 10 sec from the maximum amplitude of the main Rayleigh wave envelope. Since the signals are in the range of 0-100 seconds, the data far from this range can represent the noise level. We use the average of the absolute amplitudes between 600 and 800 seconds away from the zero point of the cross-correlation result as the noise level. Figure 2.13 shows the six-month stack of the cross-correlation results observed for all the station pairs. There are clear signals between the typical low and high group velocity lines of 2 and 4 km/s. The signals in the positive lag time part are weaker than the negative part. The reason is that the source of ambient seismic noise in southern California has a strong directional dependence, with stronger signals in the west and southwest directions than others [Yang and Ritzwoller, 2008]. Figure 2.14 is an example of the stacked cross-correlation signals. The Rayleigh wave envelope is very clear in each band. Finally, we compute mean group velocities of station pairs at different periods from the ambient noise signals by using the frequency-time analysis (FTAN) method [Levshin et al., 1992]. The FTAN method uses a set of narrow-band Gaussian filters to obtain frequency-time amplitude maps, and then traces dispersion curves of surface wave group velocities. This method has been widely used in the group velocity detections of surface waves [e.g., Ritzwoller and Levshin, 1998; Moschetti et al., 2007; Zheng et al., 2008; Zhou et al., 2012]. In this step, we select the station pairs of each period only if the interstation distances are larger than one wavelength [Brenguier et al., 2007].

We use mean group velocity measurements of Rayleigh waves from ambient noise data as input to invert a 2-D group velocity model in the Garlock Fault area. The

initial grid, adaptive grids and the corresponding ray distributions are shown in Figure 2.15 for 5, 7, 8 and 11 seconds of periods. We select a polygon that is covered by rays in our study area for robust resolution. The initial grid nodes are the 5 vertices of the polygon. The adaptive grids obtained after four iterations have more nodes in the central areas where the rays are more densely distributed. In Figure 2.15b-e, the minimum grid sizes are 21.3 km, 28.7 km, 32.9 km and 45.8 km for 5, 7, 8 and 11 seconds of periods, respectively, which are greater than one wavelength of the Rayleigh waves at the corresponding periods.

In the tomography part, we use the average value of the group velocity measurements at each period as the starting velocity model. The final group velocity results are obtained after five iterations when the group velocities change very little. Figure 2.16 shows the group velocities of Rayleigh waves at different periods. A rule-of-thumb is that the maximum sensitivity depth of a group velocity related to a shear wave is at its period expressed in kilometers [Lin et al., 2007a]. Therefore the group velocities in Figure 2.16 reflect the shear wave velocity at depths of about 4-6, 6-8, 7-9 and 10-12 km, respectively.

Our resulting group velocity model correlates well with the local geological structures. At 5 second period (Figure 2.16a), the high velocity anomalies in the cold southern Sierra Nevada crust in the north and the intersection of the Western San Gabriel Mountains and the Tehachapi Mountains in the south may be related with their crystalline basements [Dixon et al., 2000]. The shear wave velocity at 6 km depth implied by the V_p and V_p/V_s model for southern California by Lin et al. [2007b] agrees with these features in our model. In Figure 2.16b and c, similar structures are observed at 7 and 8 second periods, which are also consistent with the group velocity result by Sabra et al. [2005], except for the high velocity in the Western San Gabriel Mountains,

which may indicate the roots of the San Gabriel Mountains. In Figure 2.16d (11 second period), the velocity in the Western San Gabriel Mountains is not as high as previous periods. However, the high velocity anomaly in the southern Sierra Nevada is also very clear. In contrast to these high velocity anomalies, the Southern San Joaquin Valley on the west side is dominated by low velocities, especially at 7 and 8 second periods. These low velocity anomalies are consistent with the sedimentary deposits indicated by the seismic reflection and borehole studies, which showed that the thickness of the sediments is around 9 km [Kilkenny, 1951; Goodman and Malin, 1992]. The Western Mojave also shows low velocities, likely due to the disruption by the historical extension, strike-slip faulting [Glazner et al., 2002] and sedimentary deposits.

2.5 Summary

In this study, we developed a new adaptive grid ambient noise tomography. In the synthetic test, the results of the adaptive grid ambient noise tomography contain much fewer artifacts and have better resolution in the area with dense data. The evidence indicates that the adaptive grid method is more stable than the traditional method. The application of the adaptive grid ambient noise tomography at Garlock fault area reveals several notable group velocity structures of Rayleigh waves, such as high velocities at Sierra Nevada and low velocities at the Southern San Joaquin Valley.

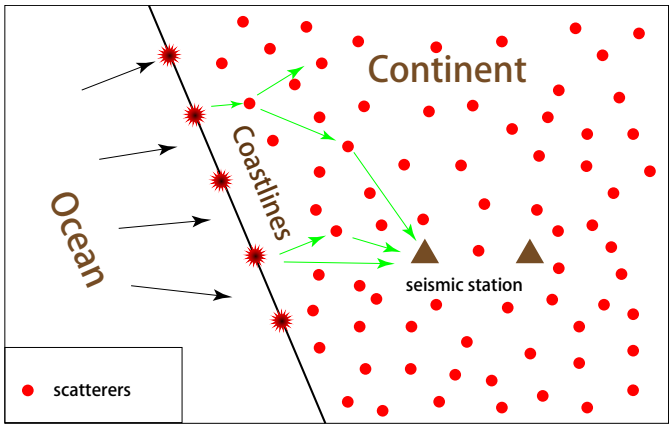

Figure 2.1: An illustration about the propagation of ambient noise signals.

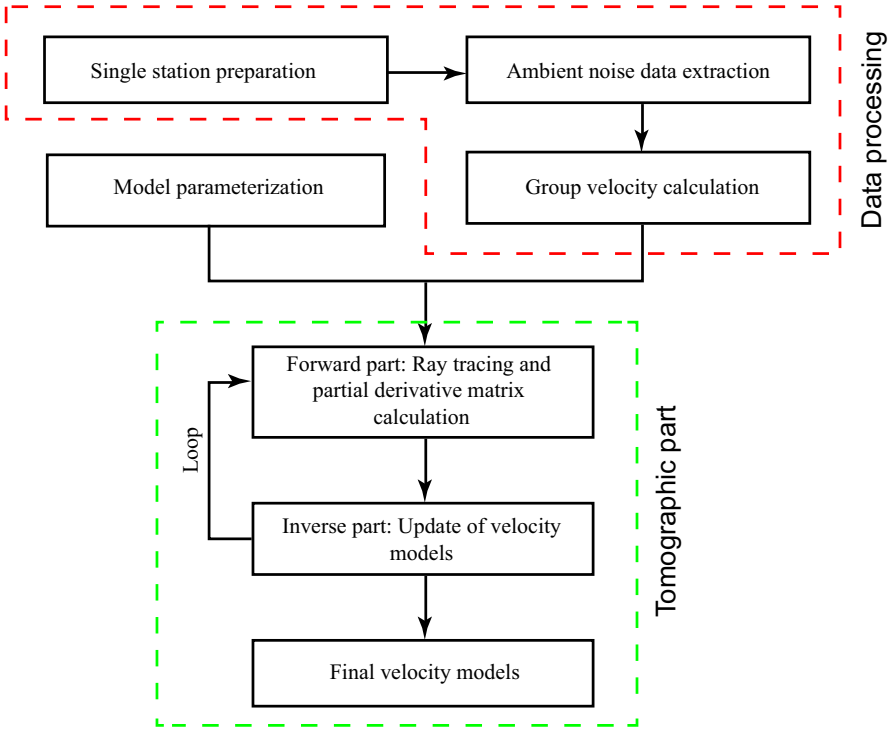

Figure 2.2: A work flowchart of our adaptive tomography in this study.

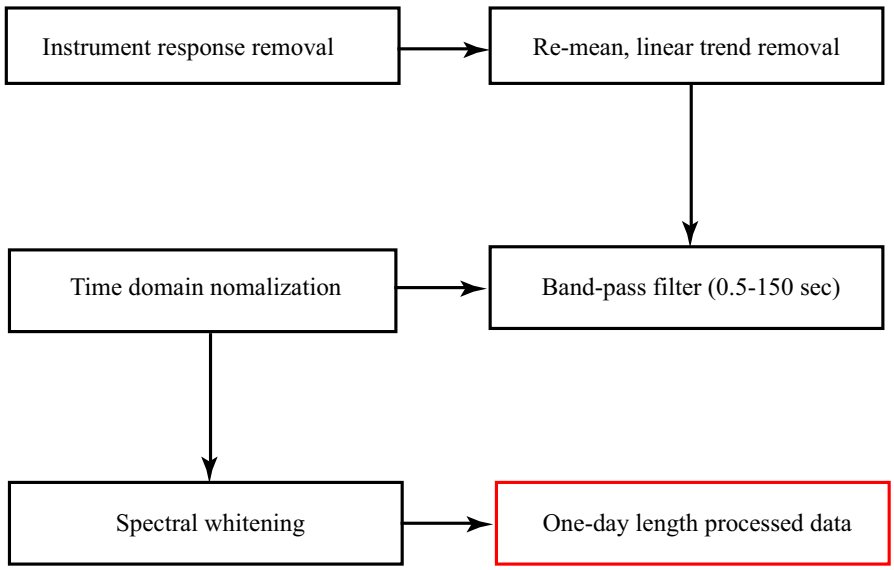

Figure 2.3: A work flowchart of single station preparation.

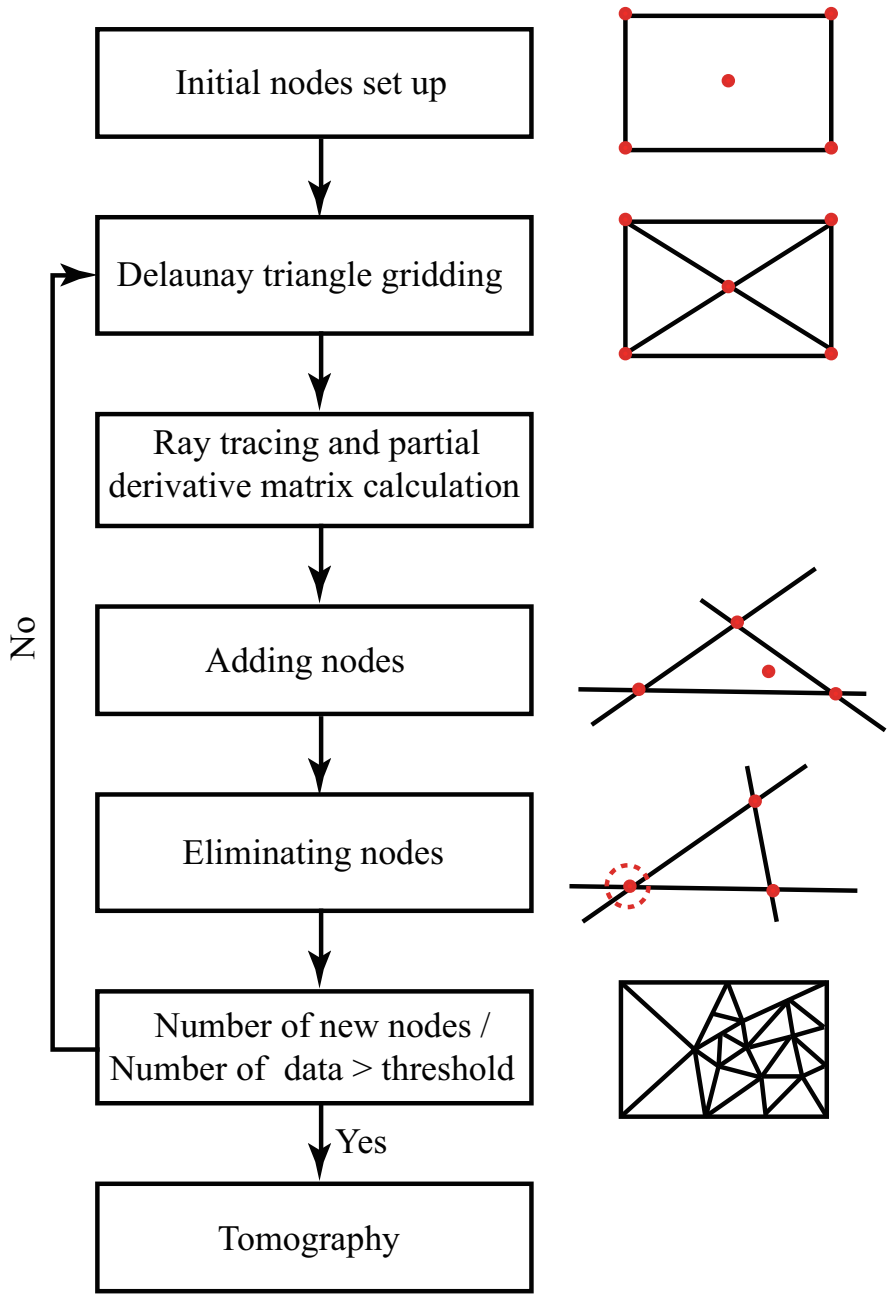

Figure 2.4: A work flowchart of model parameterization in this study.

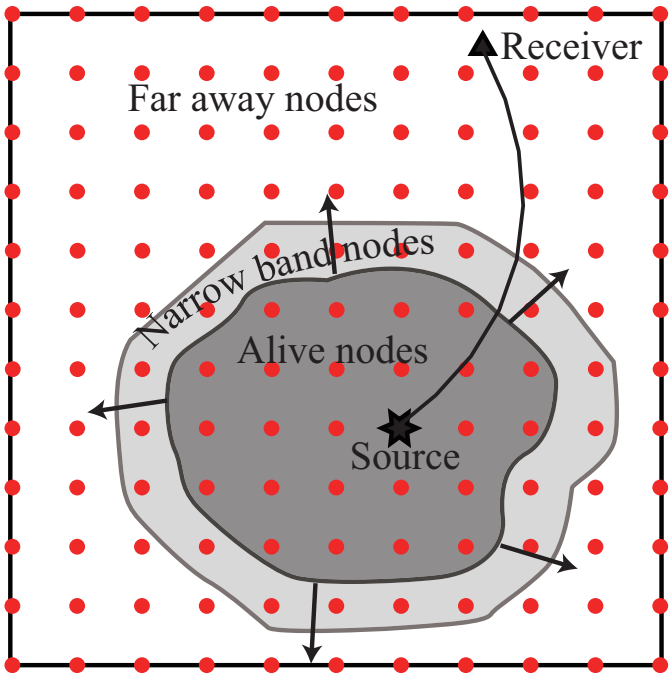

Figure 2.5: Illustration of ray tracing method. We firstly calculate the travel times of the wave in the study area by the Fast Marching method. Then a ray is constructed from the receiver to the source. The red points are the finite difference nodes used in the Fast Marching method. The edge of the dark grey area is the wave front. The arrows at the wave front show the propagating direction of the wave. The nodes in the dark grey area are alive nodes. The nodes in the light grey area are narrow band nodes which are neighbours of the alive nodes. The nodes in the white area are far away nodes.

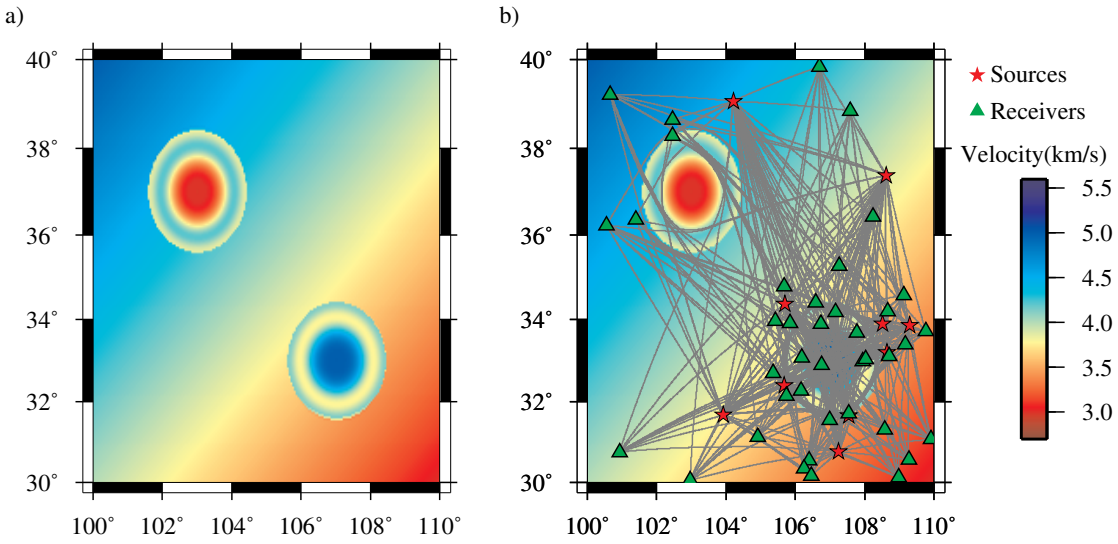

Figure 2.6: (a) True velocity model for the synthetic data tests. (b) Event (red stars), station (green triangles) and ray (gray lines) distributions in the synthetic data tests.

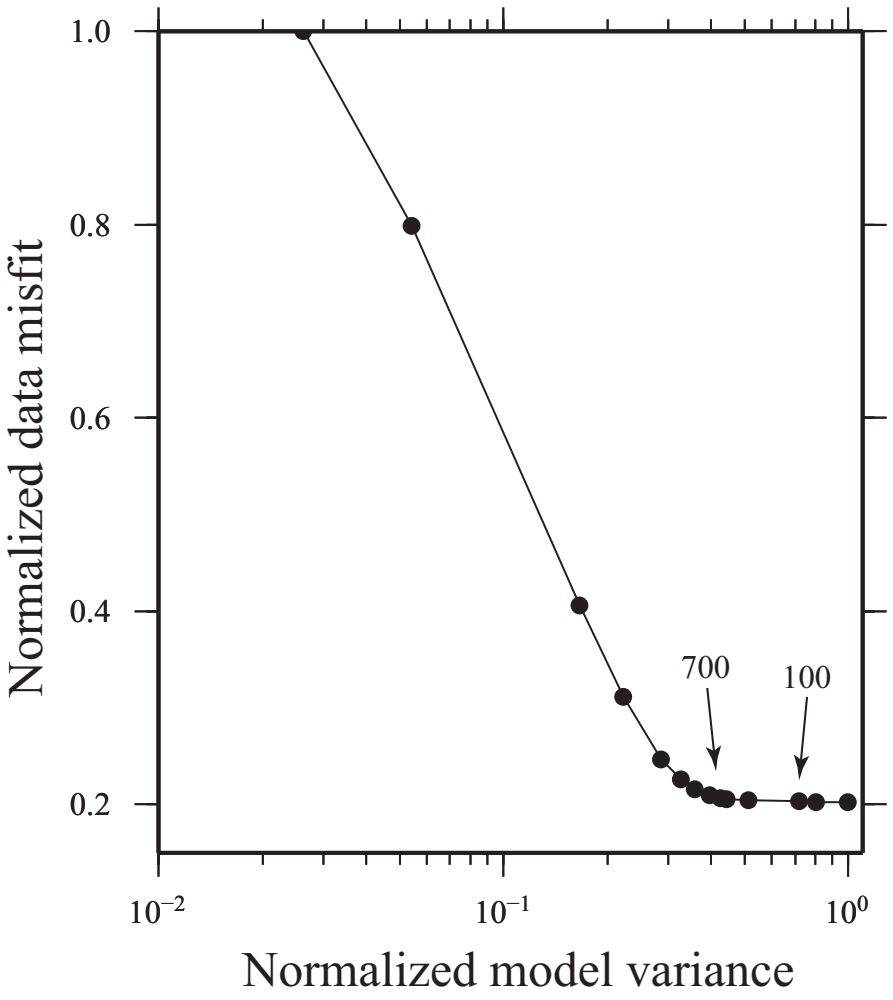

Figure 2.7: Trade-off curve between data misfit and model variance in the tomographic inversions. Damping parameters of 100 and 700 are chosen for the synthetic data tests.

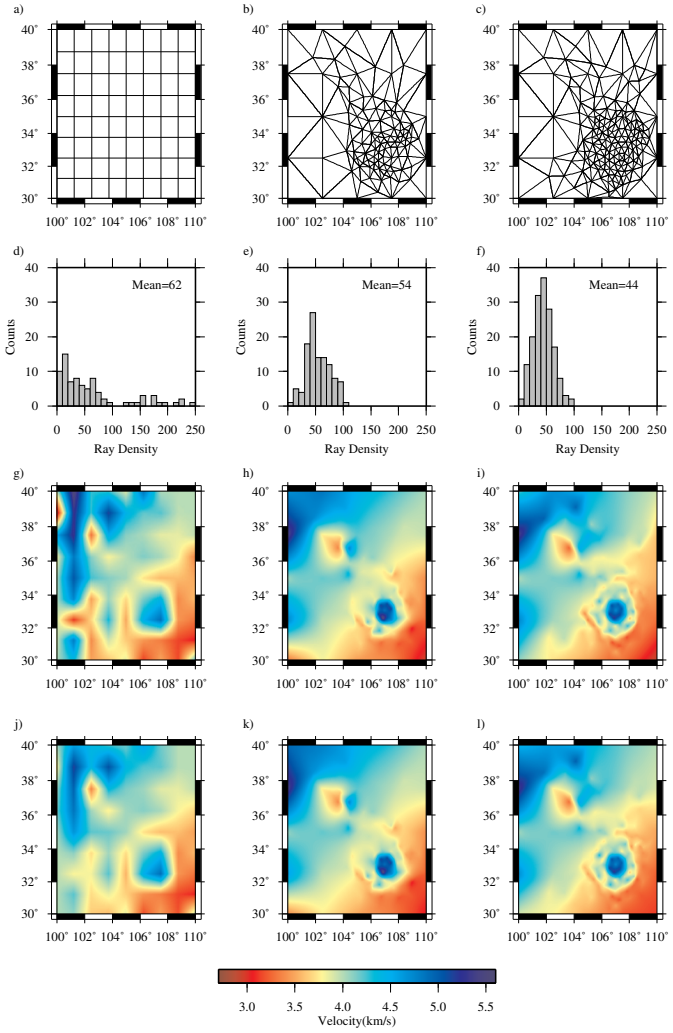

Figure 2.8: (a) Regular grid node for traditional tomography shown in (g) and (j) with grid size of $1.25° \times 1.25°$. (b) Adaptive grid with 111 nodes (model 1) for adaptive tomography shown in (h) and (k). The maximum and minimum grid spacings are about 2° and 0.4°, respectively. (c) Adaptive grid with 161 nodes (model 2) for adaptive tomography shown in (i) and (l). The maximum and minimum grid spacings are also about 2° and 0.4° but with denser grids than in (b). (d)-(f) Histograms of ray density for grid distributions in (a)-(c). The average number of rays through each grid is 62, 54 and 44 for three distributions, respectively. (g)-(i) Inverted velocity model for grid distributions in (a)-(c) with damping parameter at 100. (j)-(l) Inverted velocity model for grid distributions in (a)-(c) with damping parameter at 700.

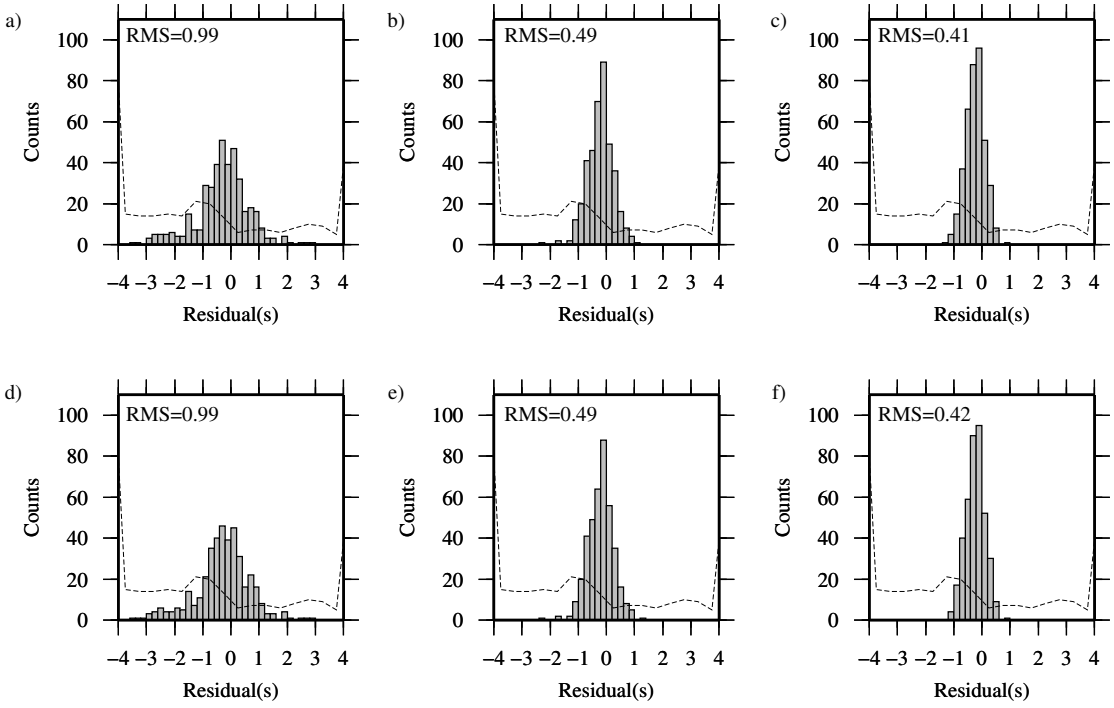

Figure 2.9: Histograms of travel time residuals and their corresponding root-mean-squares (RMS) from the inverted velocity models in Figure 2.8. (a)-(f) correspond to (g)-(l) in Figure 4. Black dash lines in all figures show the residual distribution from the uniform starting model and its RMS residual is 7.10 second.

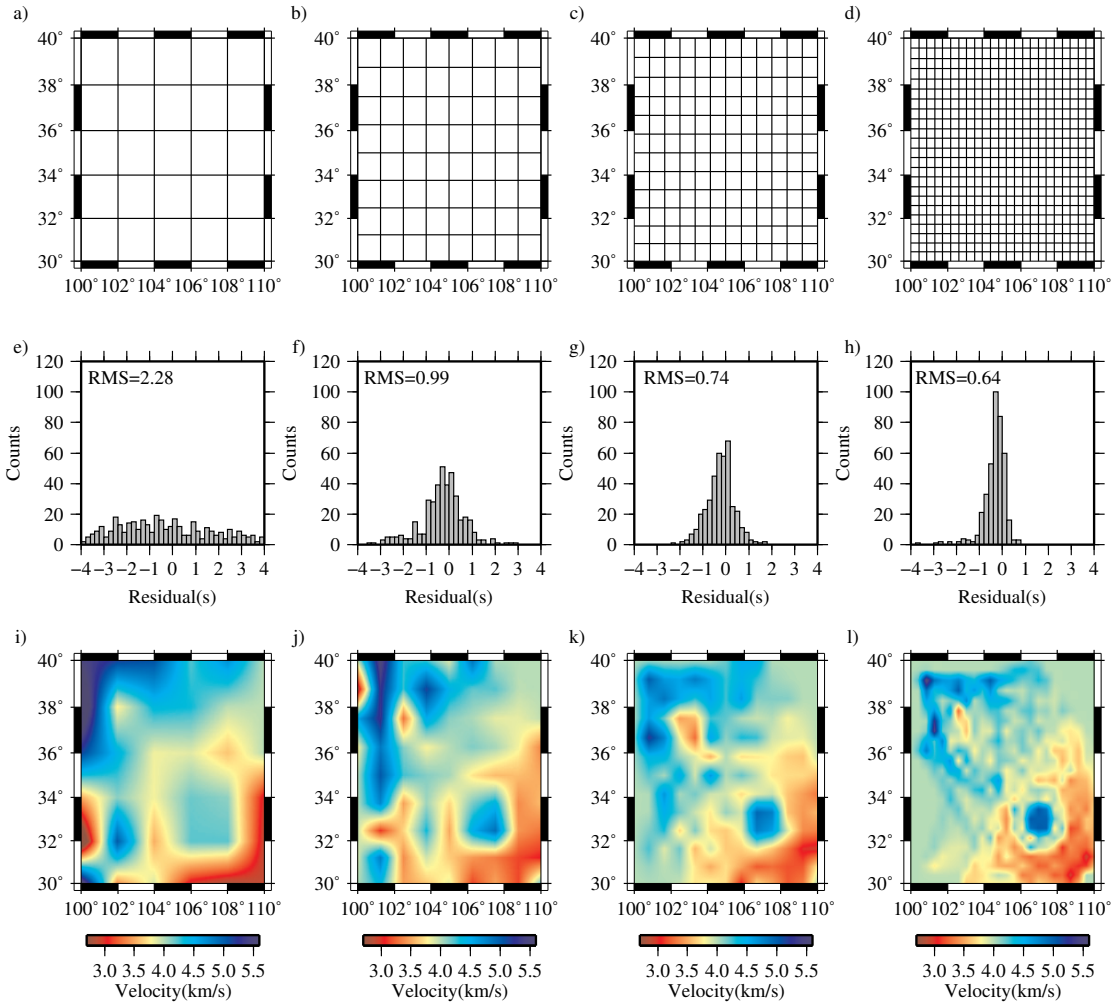

Figure 2.10: Inversion results of the traditional method with different grid spacing. (a)-(d) Grid node distributions; (e)-(h) residuals histograms; and (i)-(l) inverted models. The grid spacing is 2° in (a), 1.25° in (b), 0.83° in (c), and 0.43° in (d).

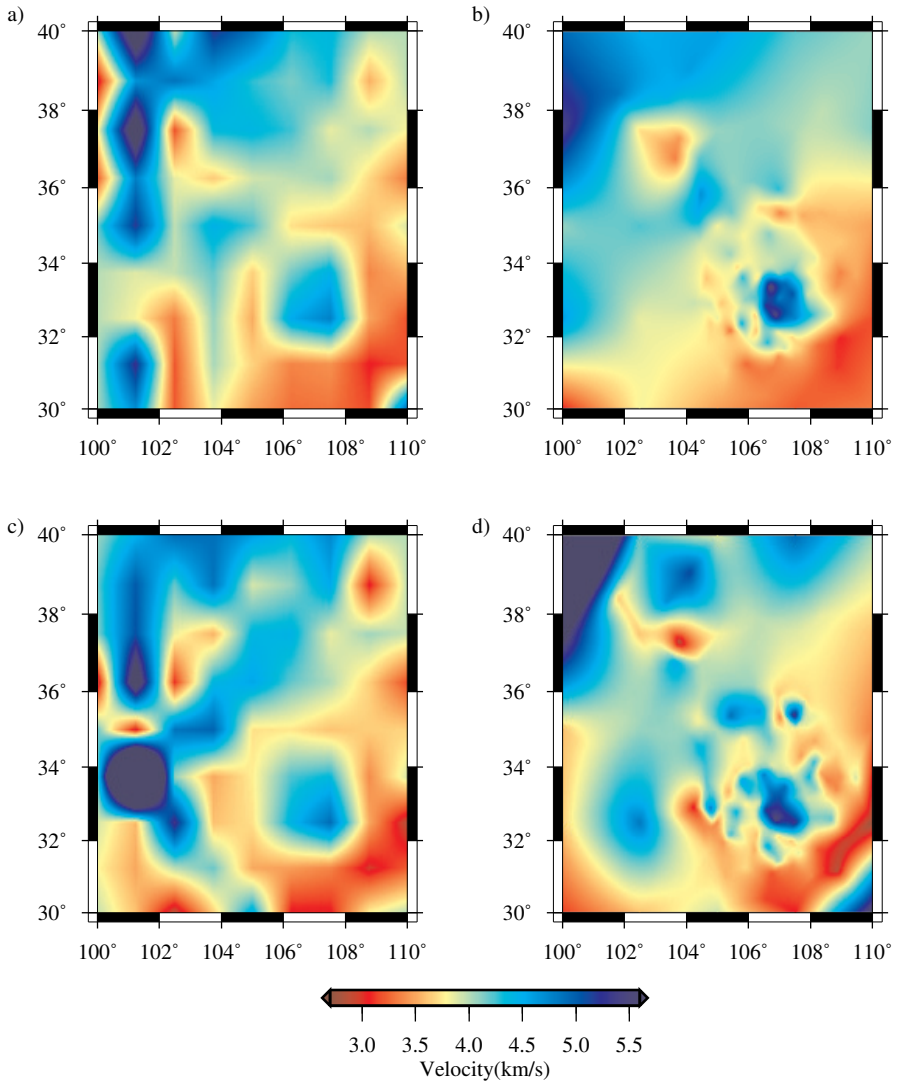

Figure 2.11: Inverted velocity models with Gaussian random noise added in the input data. (a) Result of the traditional method with the grid distribution in Figure 2.8a and a noise level of 1.5% STD. (b) Result of the adaptive method with the grid distribution in Figure 2.8b and noise level of 1.5% STD. (c) Same as in (a) but at a noise level of 5% STD. (d) Same as in (b) but at a noise level of 5% STD.

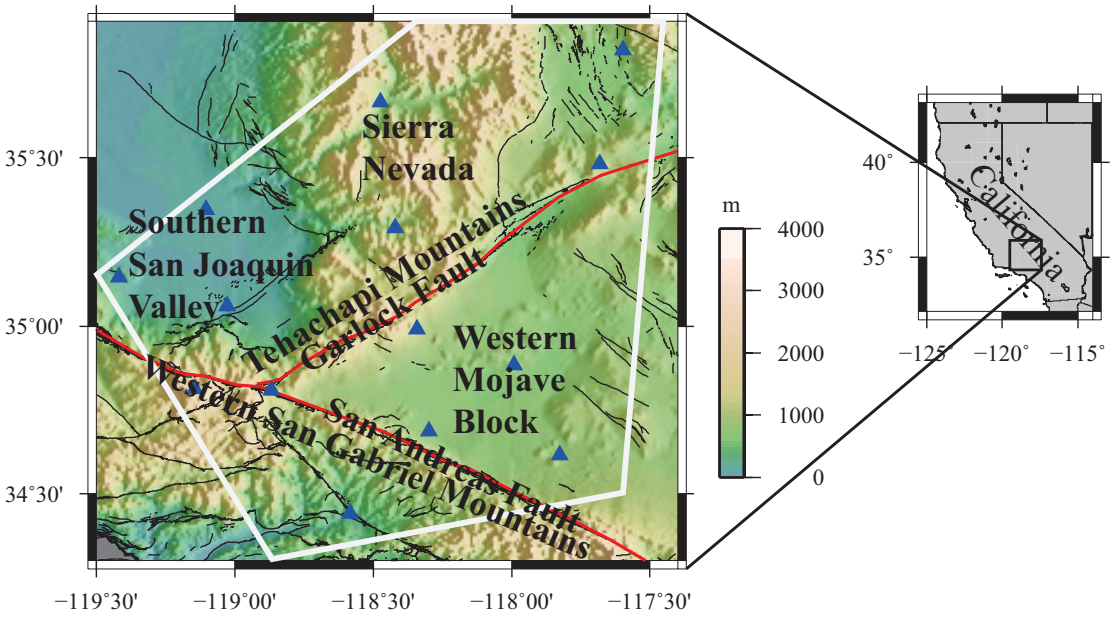

Figure 2.12: Major geological features near the Garlock Fault, California. Blue triangles represent the SCSN seismic stations used in the inversion. Grey polygon is our study area. Black lines are the faults and the red lines highlight the San Andreas Fault and Garlock Fault. The background is the topography from National Geophysical Data Center. The small panel on the right side shows the location of the Garlock fault area in California.

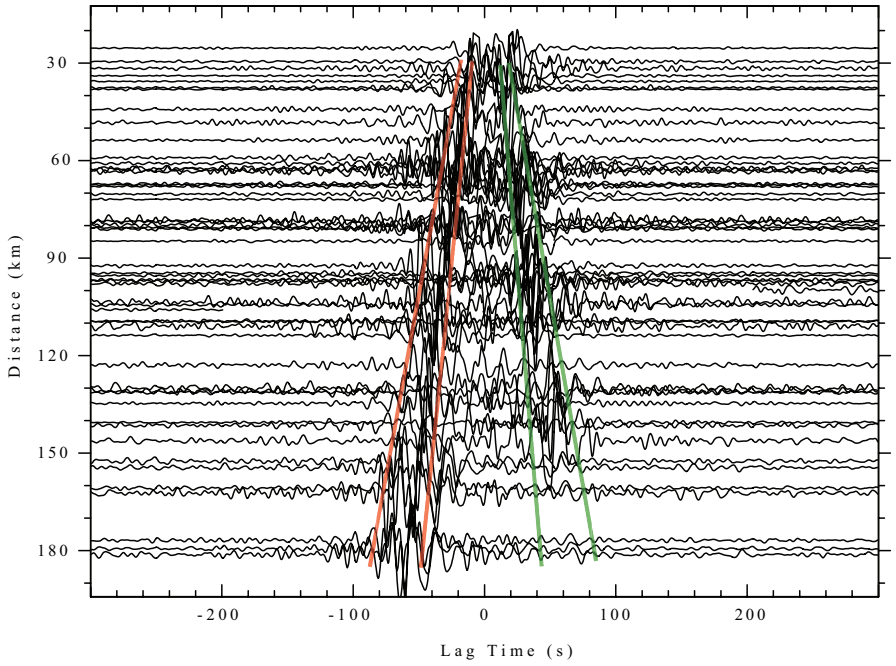

Figure 2.13: The ambient noise signals of the station pairs in the period of 0.5-150 second. The red and green lines in negative and positive lag times mark the velocity lines at 2 km/s and 4 km/s.

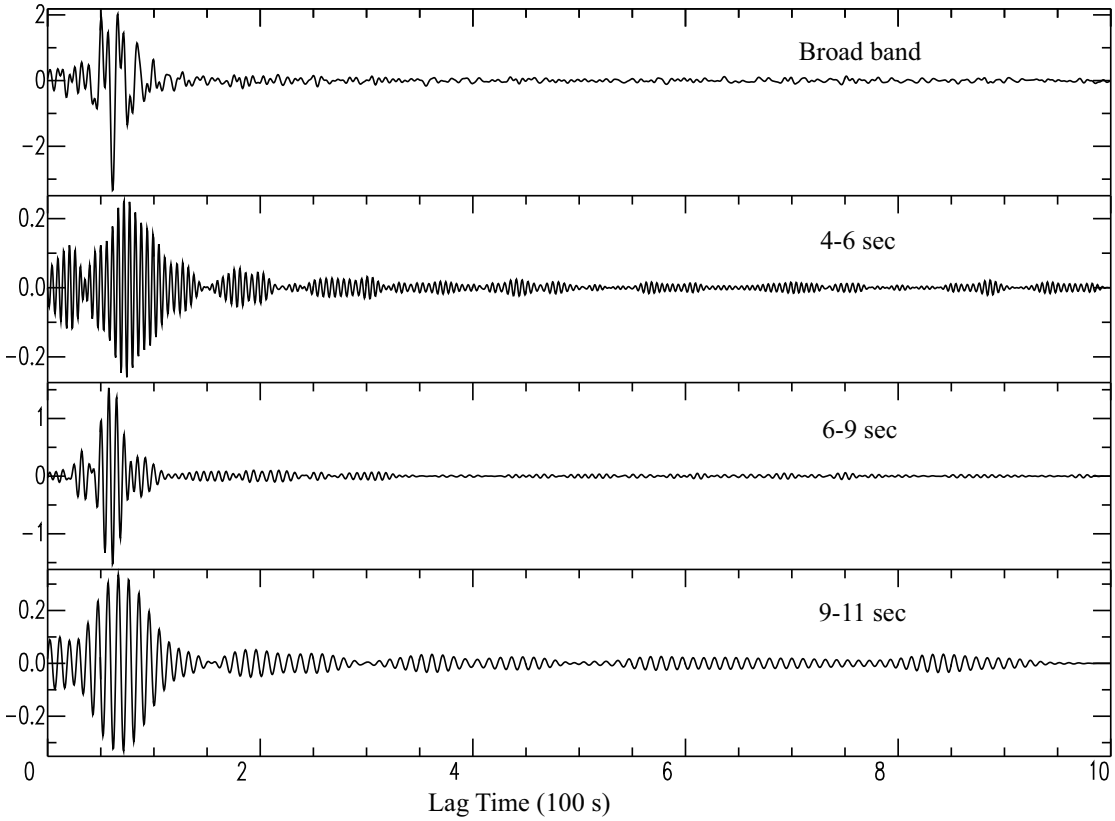

Figure 2.14: Stack of six months' cross-correlation signals observed for station pair of CLC-PDE filtered at different bands from the original broadband data.

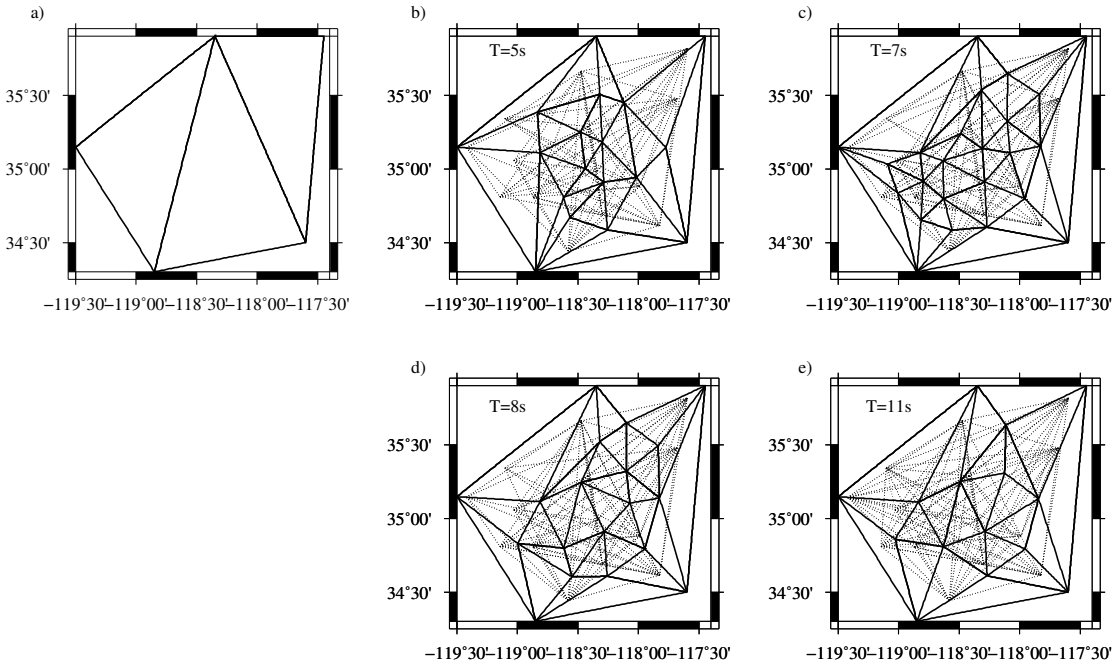

Figure 2.15: (a) Initial grid distribution of the adaptive tomography in our study area. (b)-(e) Adaptive grids for Rayleigh wave velocity inversions at 5, 7, 8 and 11 seconds, respectively. Dotted black lines represent rays used in the inversions for each period.

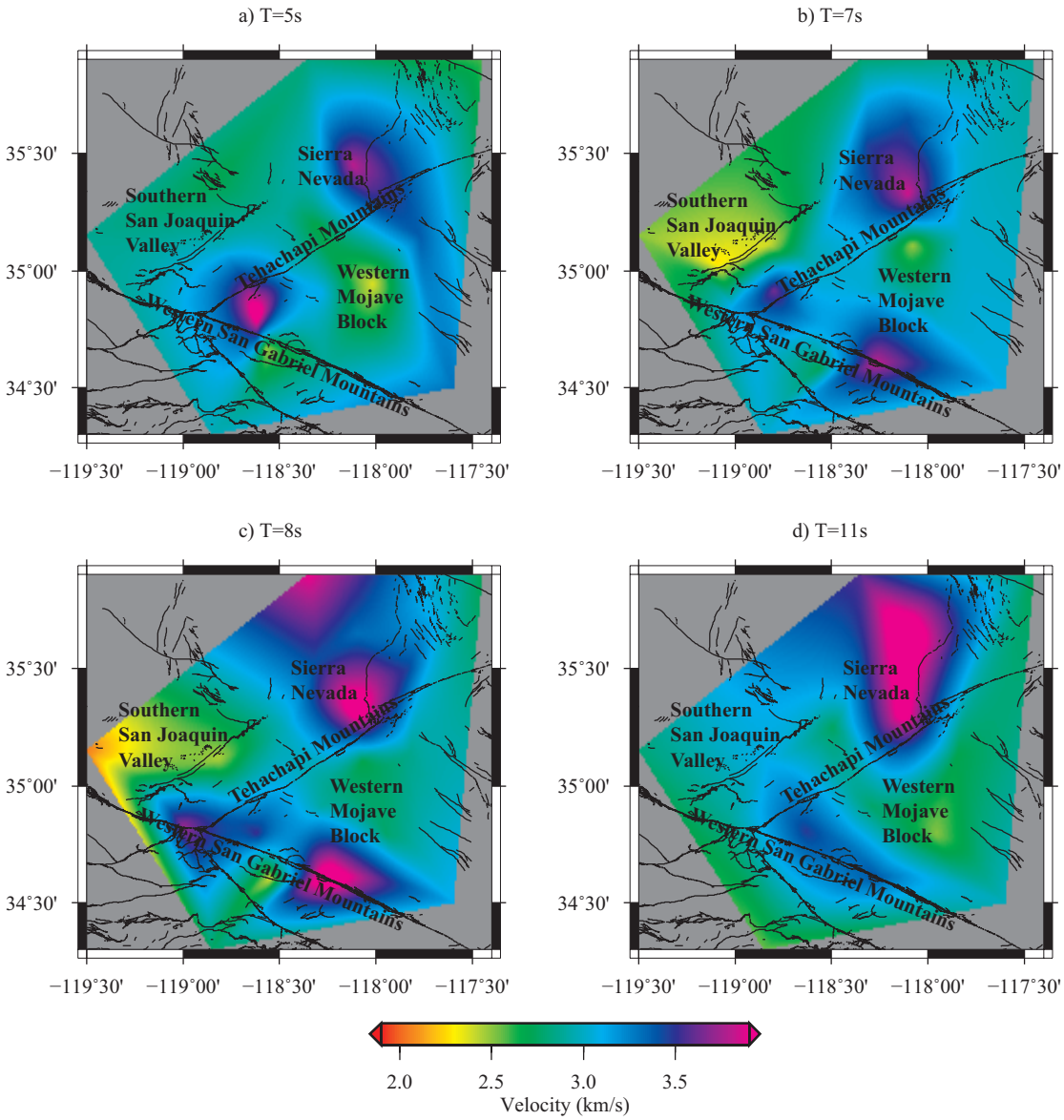

Figure 2.16: Group velocities of Rayleigh wave at different periods in our study area. Abbreviations are the same as those in Figure 2.12.

Chapter 3

Local earthquake tomography with topography

3.1 Overview

Local earthquake tomography uses P- or S-picks of the local earthquakes to invert the P- or S-wave velocity in the area with scales of 10s to 100s km. The local earthquakes are the earthquakes in the study area, and the epicentral distances are from 10s to 100s km. Considering the small scale of the local earthquake tomography, the topography could affect the tomographic result, especially in the area with rough topography, such as volcanoes or mountains. To reduce the effect of topography, we developed a new local earthquake tomography algorithm with including of full topography instead of using only station elevations as the traditional method does. We conduct a synthetic test at Kīlauea volcano to compare the results of traditional and new method. Then the new method is applied to the real data set at Kīlauea volcano to invert a new 3D P- and S-wave velocity structures. The following are the method, the synthetic test and the real data application of the local earthquake tomography with full topography.

3.2 Method

The main purpose of local earthquake tomography is investigating the P- or S-velocity structures of regional or local scale by using arrival times of the P- or S-phases. The equations of arrival times of P- and S-waves of an earthquake can be represented by,

$$\begin{aligned} T^p_{i,j} &= t_i + \int_{source_i}^{recever_j} \frac{dl}{v^p} \\ T^s_{i,j} &= t_i + \int_{source_i}^{recever_j} \frac{dl}{v^s} \end{aligned} \tag{3.1}$$

where i and j mean i_{th} earthquake and j_{th} seismic station, respectively. p and s indicate P- and S-wave, respectively. $T_{i,j}$ is the arrival time from i_{th} earthquake to j_{th} seismic station. dl is the small segment of the ray path. The arrival time is the sum of earthquake origin time t_i and the integral of the velocity (or travel time) along the ray path. Usually, the arrival time and location of the receiver (seismic station) are known. Location of source (earthquake), origin time, ray path and velocity model are unknown. The equation 3.1 can also be written in another form,

$$\begin{aligned} T^p_{i,j} &= t_i + \int_{source_i}^{recever_j} s^p dl, s^p = \frac{1}{v^p} \\ T^s_{i,j} &= t_i + \int_{source_i}^{recever_j} s^s dl, s^s = \frac{1}{v^s} \end{aligned} \tag{3.2}$$

where s^p and s^s are the slowness of P- and S-wave, respectively. The advantage of using slowness is that the arrival time would have linear relationship with the velocity structure.

Usually, the integral Equation 3.2 can not be solved directly by a computer. It can be linearized as the following:

$$\begin{aligned} dT^p_{i,j} &= \frac{\partial T^p_{i,j}}{\partial x_i} dx_i + \frac{\partial T^p_{i,j}}{\partial y_i} dy_i + \frac{\partial T^p_{i,j}}{\partial z_i} dz_i + \frac{\partial T^p_{i,j}}{\partial t_i} dt_i + \sum_{k=1}^{n} \frac{\partial T^p_{i,j}}{\partial s^p_k} ds^p_k \\ dT^s_{i,j} &= \frac{\partial T^s_{i,j}}{\partial x_i} dx_i + \frac{\partial T^s_{i,j}}{\partial y_i} dy_i + \frac{\partial T^s_{i,j}}{\partial z_i} dz_i + \frac{\partial T^s_{i,j}}{\partial t_i} dt_i + \sum_{k=1}^{n} \frac{\partial T^s_{i,j}}{\partial s^s_k} ds^s_k \end{aligned} \tag{3.3}$$

where i and j mean i_{th} earthquake and j_{th} seismic station, respectively. p and s indicate P- and S-wave, respectively. $dT^p_{i,j}$ and $dT^s_{i,j}$ are the arrival time perturbations of P- and S-waves from i_{th} earthquake to j_{th} seismic station. dx_i, dy_i, dz_i, and dt_i are the perturbations of i_{th} earthquake location in Cartesian coordinates and origin time. ds^p_k and ds^s_k are the discretized velocity parameters of P- and S-wave. Usually there are two ways to discretize the study area, which are the block method and the node method. The block method means that the study area is divided by a series of blocks and ds^p_k and ds^s_k mean the P- and S-velocity of k_{th} block. On the other hand, in the node method, the study area is represented by a series of nodes inside the study area and ds^p_k and ds^s_k are the P- and S-velocity of k_{th} node. Then the velocity of the whole study area is interpolated by the velocity at the nodes. $\frac{\partial T^p_{i,j}}{\partial x_i}$, $\frac{\partial T^p_{i,j}}{\partial y_i}$, $\frac{\partial T^p_{i,j}}{\partial y_i}$, $\frac{\partial T^p_{i,j}}{\partial t_i}$ and $\frac{\partial T^s_{i,j}}{\partial x_i}$, $\frac{\partial T^s_{i,j}}{\partial y_i}$, $\frac{\partial T^s_{i,j}}{\partial y_i}$, $\frac{\partial T^s_{i,j}}{\partial t_i}$ are partial derivatives of earthquake locations and origin times for P- and S-wave arrival times, respectively. $\frac{\partial T^p_{i,j}}{\partial s^p_k}$ and $\frac{\partial T^s_{i,j}}{\partial s^s_k}$ are partial derivatives of P- and S-wave slownesses, respectively.

Partial derivatives of earthquake locations are represented as [Thurber, 1983]:

$$\begin{aligned}
\frac{\partial T^p_{i,j}}{\partial x_i} dx_i &= -s^p_e (\frac{dz_e}{dl_e})^p_{i,j} \\
\frac{\partial T^p_{i,j}}{\partial y_i} dy_i &= -s^p_e (\frac{dy_e}{dl_e})^p_{i,j} \\
\frac{\partial T^p_{i,j}}{\partial z_i} dz_i &= -s^p_e (\frac{dz_e}{dl_e})^p_{i,j} \\
\frac{\partial T^s_{i,j}}{\partial x_i} dx_i &= -s^s_e (\frac{dx_e}{dl_e})^s_{i,j} \\
\frac{\partial T^s_{i,j}}{\partial y_i} dy_i &= -s^s_e (\frac{dy_e}{dl_e})^s_{i,j} \\
\frac{\partial T^s_{i,j}}{\partial z_i} dz_i &= -s^s_e (\frac{dz_e}{dl_e})^s_{i,j}
\end{aligned} \tag{3.4}$$

where s^p_e and s^s_e are the P- and S-velocity at the source of the earthquake, respectively. $\frac{dx_e}{dl_e}$, $\frac{dy_e}{dl_e}$ and $\frac{dz_e}{dl_e}$ are the x, y and z components of the ray paths at the source of

the earthquake, respectively. The partial derivatives of the origin time are much more simple. From the travel time equation (Equation 3.2) we know:

$$\begin{aligned} \frac{\partial T^p_{i,j}}{\partial t_i} dt_i &= 1 \\ \frac{\partial T^s_{i,j}}{\partial t_i} dt_i &= 1 \end{aligned} \tag{3.5}$$

Partial derivatives of the slowness depend on the method of parameterizing the velocity model of the study area. If the velocity model is represented by a series of blocks, the partial derivatives of the slowness of each block is just the sum of the ray segments length in it. If the velocity model is represented by a series of nodes, the partial derivatives of the slowness varies from the method of interpolation which is used to get the velocity model in the whole study area from the velocities at the nodes. Here we use the trilinear interpolation method to calculate the slownesses of the points in the study area. Figure 3.1 shows an example of the slowness calculation of a point in a single cuboid. The slowness at point $P(x, y, z)$ is determined by the eight vertices of the cuboid, which can be represented by the equation:

$$\begin{aligned} s(x,y,z) = {} & s_{111}\frac{(x_2-x)(y_2-y)(z_2-z)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{211}\frac{(x-x_1)(y_2-y)(z_2-z)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{121}\frac{(x_2-x)(y-y_1)(z_2-z)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{221}\frac{(x-x_1)(y-y_1)(z_2-z)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{112}\frac{(x_2-x)(y_2-y)(z-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{212}\frac{(x-x_1)(y_2-y)(z-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{122}\frac{(x_2-x)(y-y_1)(z-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \\ & + s_{222}\frac{(x-x_1)(y-y_1)(z-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)} \end{aligned} \tag{3.6}$$

where s_{111}, s_{211}, s_{121}, s_{221}, s_{112}, s_{212}, s_{122} and s_{222} represent the slownesses at vertices of the cuboid (See Figure 3.1). $x(x, y, z)$ is the slowness at the point $P(x, y, z)$ inside the cuboid. With the trilinear interpolation method, the partial derivatives of slowness are more complex than the sum of the ray segment lengths. It is more likely that the partial derivative of slowness at each node is the sum of the weighted ray segment length in the cuboid. Figure 3.2 shows how the partial derivative of slowness is calculated in a single cuboid. The study area is parameterized by a series of nodes and the nodes consist of the vertices of the cuboid in Figure 3.2. In general, the partial derivative of the slowness at the vertices in a single cuboid can be represented as:

$$
\begin{aligned}
(\frac{\partial T}{\partial s})_{111} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_2-x_{i,j})(y_2-y_{i,j})(z_2-z_{i,j})}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{211} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_{i,j}-x_1)(y_2-y_{i,j})(z_2-z_{i,j})}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{121} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_2-x_{i,j})(y_{i,j}-y_1)(z_2-z_{i,j})}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{221} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_{i,j}-x_1)(y_{i,j}-y_1)(z_2-z_{i,j})}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{112} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_2-x_{i,j})(y_2-y_{i,j})(z_{i,j}-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{212} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_{i,j}-x_1)(y_2-y_{i,j})(z_{i,j}-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{122} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_2-x_{i,j})(y_{i,j}-y_1)(z_{i,j}-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl \\
(\frac{\partial T}{\partial s})_{222} &= \sum_{j=1}^{a}\sum_{i=1}^{b_j}\frac{(x_{i,j}-x_1)(y_{i,j}-y_1)(z_{i,j}-z_1)}{(x_2-x_1)(y_2-y_1)(z_2-z1)}dl
\end{aligned}
\tag{3.7}
$$

where i is the index of the segments of a single ray path inside the cuboid. j is the index of the j^{th} rays in the cuboid. b_j is the total number of segments of j^{th} rays inside

the cuboid. a is the total number of the rays passed the cuboid. dl is the length of the segments of rays. $x_{i,j}$, $y_{i,j}$ and $z_{i,j}$ are the coordinates of the center point of the segment of the ray path. $(\frac{\partial T}{\partial s})_{111}$, $(\frac{\partial T}{\partial s})_{211}$, $(\frac{\partial T}{\partial s})_{121}$, $(\frac{\partial T}{\partial s})_{221}$, $(\frac{\partial T}{\partial s})_{112}$, $(\frac{\partial T}{\partial s})_{212}$, $(\frac{\partial T}{\partial s})_{122}$ and $(\frac{\partial T}{\partial s})_{222}$ are the partial derivative of slowness at each vertices of the cuboid. Considering that the different cuboids may have a common vertex, we need to calculate the partial derivative of slowness in each cuboid first, and then combine them together to obtain the partial derivative of slowness at each node.

If we use a computer to solve the equations (Usually it is impossible to handle such a large amount of data and equations by hand.), it is better to use the matrix form of Equation 3.3. In matrix form, the Equation 3.3 can be written as,

$$\begin{bmatrix} \boldsymbol{d}_p \\ \boldsymbol{d}_s \end{bmatrix} = \boldsymbol{G} \begin{bmatrix} \boldsymbol{m}_{evn} \\ \boldsymbol{m}_p \\ \boldsymbol{m}_s \end{bmatrix} \tag{3.8}$$

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{G}_{evn} & \boldsymbol{G}_p & \boldsymbol{G}_s \end{bmatrix} \tag{3.9}$$

where the left side is the data vector including the P wave arrival time as $\boldsymbol{d}_p$ and the S wave arrival time as $\boldsymbol{d}_s$. On the right side of the equation, matrix $\boldsymbol{G}$ is called the partial derivative matrix, which consists of all partial derivatives in Equation 3.3. The model vector on the right of matrix $\boldsymbol{G}$ includes the source parameters of the earthquake, P- and S-wave velocity parameters, which are $\boldsymbol{m}_{evn}$, $\boldsymbol{m}_p$ and $\boldsymbol{m}_s$, respectively.

3.2.1 Method for including full topography into tomography

The work-flow of the LETFT is shown in Figure 3.3. First, in order to reduce the amount of data and improve the data quality used in the tomography [Lin et al., 2007b], we pick the master events by dividing the study area into a series of small cubes and selecting the events with the greatest number of picks in each cube. We then add topography

data into the tomographic process together with the P- and S-wave phase data of the master events. The topography data used in this study is from the Hawaiian multibeam bathymetry synthesis (http://www.soest.hawaii.edu/hmrg/multibeam). The original resolution of the topography data is 50 m and we resample it to 200 m. Since the spacing of the velocity grid is 3 km in horizontal, 200 m resolution of topography data is much smaller than 3 km and is enough for our study. In the tomographic process, the velocity model and earthquake source parameters (i.e., location and origin time) are solved iteratively by three steps, including a forward part, a parameter separation part and a inverse part. In the forward part, ray paths, arrival time residuals and partial derivative matrices of the velocity model and earthquake locations are calculated based on the initial velocity model or the velocity model from a previous iteration. We then separate the source and velocity parameters by using the method presented by Pavlis and Booker [1980]. In the inverse part, the velocity model and source parameters are updated based on the partial derivative matrix and arrival time residuals obtained from the forward part. The updated velocity model and source parameters are used for the next iteration. The iteration stops when the change of the root-mean-square (RMS) of the arrival time residuals is smaller than a given threshold.

Figure 3.4 illustrates schematically the difference between the traditional LET method and the new LETFT method with the inclusion of full topography. In the traditional LET method (Figure 3.4a), the whole study area is represented in a Cartesian coordinate system with a flat top and extra layers added above sea level to include seismic stations into the study area, since they usually have certain elevations above sea level. The study area is then discretized into a series of blocks or nodes to represent the velocity model. Since velocities of the nodes or blocks above the Earth's surface are also updated in the tomographic processing, sometimes the rays could travel above the Earth's

surface, which is unrealistic for seismic P- and S-waves. Consequently, the ray paths are affected by the problematic velocity below and above the Earth's surface. In some cases, the rays could even travel above the surface. This problematic ray tracing could introduce errors in the velocity model and errors in earthquake source parameters in the topographic processing.

In the new LETFT method (Figure 3.4b), we avoid using the velocity nodes above the Earth's surface and the unrealistic ray paths by including the full topography. The method is implemented by the following settings. First, we extend the upper boundary of the study area to a level slightly above the highest elevation in the study area. Second, we insert the topography data into the 3-D space of the study area. We then set the velocity in the volume above the surface to a small constant number (e.g., 0.34 km/s or 1.5 km/s for P velocity in air or water), which does not vary during the tomography inversion. In this study, we choose the constant velocity above the Earth's surface as 1.5 km/s. The velocities below the Earth's surface are set to an initial one or three-dimensional velocity model in the first iteration or an updated velocity model from a previous iteration. The model is set this way to simulate the real circumstances in the study area and reduce the errors in predicting ray paths and partial derivative matrices in the forward part of tomography. Third, the study area is parameterized by a set of cuboidal vertices which are separated as two sets by the topography. Nodes above the Earth's surface are not used in either the forward or the inverse part, which means that they are not used in the interpolation of the velocities below the Earth's surface in either the forward or the inverse part. Thus, the velocity nodes above the surface have no influence on the tomographic process. The velocities of the nodes below the surface are updated in each iteration of the inverse part and used to interpolate the updated velocity model of the study area.

3.2.2 Forward and inverse parts of tomography

In the forward part of the new LETFT algorithm, we use the Fast Marching Method, which is a finite difference method solving the Eikonal equation [Sethian and Popovici, 1999], to predict the ray paths, arrival time residuals and partial derivative matrices of the velocity model and earthquake locations. This method is widely used in seismic tomography because of its high stability and accuracy [e.g., Rawlinson and Sambridge, 2004; Li and Lin, 2014]. With the above settings of the inclusion of real topography, ray paths avoid unrealistic penetration above the surface of the Earth and thus the accuracy of the partial derivative matrix and arrival time prediction are improved. After the forward processing, the parameters of the velocity model and source are separated by QR decomposition and updated subsequently. The velocity model is solved first by a damped least square method to minimize the following misfit function [Tarantola, 2005],

$$\Phi(\boldsymbol{m}) = ||(\boldsymbol{G}\boldsymbol{m} - \boldsymbol{d}_{obs})||^2 + ||\boldsymbol{\lambda}(\boldsymbol{m} - \boldsymbol{m}_{prior})||^2$$
$$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda_p} & 0 \\ 0 & \boldsymbol{\lambda_s} \end{bmatrix} \tag{3.10}$$

where G is the partial derivative matrix, $\boldsymbol{m}$ is the model vector, $\boldsymbol{d}_{obs}$ is the data vector, $\boldsymbol{m}_{prior}$ is a **priori** model, and $\boldsymbol{\lambda}$ is the damping parameter matrix determined by the trade-off curve between data misfit and model variance [e.g., Eberhart-Phillips, 1986]. Because the ray coverages of P- and S-wave are different, we use different damping parameters for P- and S-wave as $\boldsymbol{\lambda_p}$ and $\boldsymbol{\lambda_s}$, respectively. After getting the velocities at the inversion nodes, we use a trilinear interpolation method to obtain the velocity model in the study area beneath the surface, which only uses the nodes below the Earth's surface. Thus the nodes above the surface would not affect the subsurface velocity

model. The earthquake locations and origin times are then updated based on the new velocity model by using the singular value decomposition method [Golub and Reinsch, 1970].

3.3 Data set

We choose Kīlauea volcano in Hawai‘i (Figure 3.5) to test our method and algorithm because this area is one of the most seismically active areas in the world with a fair amount of topographic relief ($\sim$1 km above sea level) and there are numerous previous tomographic studies to compare with our results. The data used in this study are the 515,711 P-wave arrival times and 272,217 S-wave arrival times of 33,768 seismic events between 1992 and 2009 recorded by 35 seismic stations at the Hawaiian Volcano Observatory (HVO). We divide the study area into $2 \times 2 \times 2$ km^3 cubes and select the events with the greatest number of picks in each cube as the master events for the inversion. After this data processing, we obtain 1,800 master events with 30,172 P-picks and 12,673 S-picks for our study.

3.4 Synthetic data test

The synthetic P- and S-wave velocity models in our test are obtained by perturbing a one-dimensional (1-D) velocity model [Lin et al., 2014] with checkerboard-shape anomalies. The amplitudes of the anomalies are $\pm 10\%$ of the 1-D P- and S-wave velocity models, respectively (Figure 3.11b and 3.12b). The 1-D P- and S-wave velocity models (shown in Figure 3.11a and 3.12a) are also used as the initial models for both the synthetic and the following real data tests. The study area is parameterized by a series of cuboid shaped grid nodes. The horizontal spacing of the nodes is 3 km. In the vertical direction, the spacing of the nodes is 3 km above 11 km depth and 5 km below 11 km depth, respectively. The arrival times are calculated based on the synthetic

velocity model and the station and earthquake and ray distributions from the real data processing. We also add Gaussian-distributed noise with zero mean and 1 km standard deviation for the earthquake locations and 0.1 sec standard deviation for the origin times, and then use them as the initial values for the simultaneous P and S-wave velocity and source parameter inversions.

Before the final tomographic process, we tested a series of damping parameters ranging from 10 to 25,000 in order to find the optimal damping parameters (λ_p and λ_s in equation 3.10) for the inversion of P- and S-velocity models, separately. The strategy is similar to that described by Lin et al. [2014]. First, we fix λ_s at a number as large as 25,000 to find the optimal λ_p. The data misfits and model variances resulted from different damping parameters are calculated and plotted as the trade-off curves shown in Figure 3.7a. We select 200 as damping parameter λ_p, which provides an optimal compromise between data misfit and model variance. Second, we fix λ_p as 200 and find the damping parameter λ_s. From the trade-off curve of λ_s, we select 150 as the S-velocity damping parameter λ_s. We then fix λ_s as 150 with a series of λ_p and confirm that 200 is indeed the optimal λ_p. After this step, we use damping parameters of 200 and 150 for the inversions of P- and S-wave velocity models in our final tomography.

Whether the seismic tomography algorithm could converge is very important. Since seismic tomography algorithms usually use an iteration method, in which the result of the current iteration is depend on the previous iteration (i.e., the ray tracing and arrival time prediction in the forward part use the velocity model from the previous iteration), not converging means that the velocity model would be farther from the true model as the number of iterations is increasing. In order to test the convergency of the new LETFT method, we run a series of iterations with the damping parameters obtained above and monitor the variation of the residuals of the arrival time data. Figure

3.8 shows the convergency of the LETFT method. The horizontal axes is the number of iterations and we run 19 iteration in total. The vertical axes is the mean absolute deviation (MAD) of the residuals of the arrival time data. The MAD of the residual at the beginning is about 0.178 sec. At the first a few iterations, the MAD of the residual drops dramatically. After the 6th iteration, the MAD of the residual becomes stable, in another word, does not change too much. This evidence indicates that our new LETFT method is able to converge in the first several iterations. It also tells us that we do not need to run too many iterations because the improvement of the results is much smaller compared with the time consumption.

The residuals of the arrival time data from the final P- and S-velocity model are shown in Figure 3.9. The residual histogram of the LETFT method is Figure 3.9a, while the residual histogram of the traditional LET method is Figure 3.9b. The residuals of the LETFT method are closer to zero and smaller than those of the traditional LET method. The RMS residual of LETFT and traditional LET method drop from 0.18 sec of the initial velocity model (Figure 3.10)to 0.013 sec and 0.016 sec, respectively. The small residual of arrival time data from the LETFT method indicates that the velocity results from it fit the data better. To better evaluate the performance of the two methods, we also need to investigate how they recover the velocity model. The final P- and S-velocity results of our synthetic data tests are shown in Figures 3.11 and 3.12. We compare our new LETFT method and the traditional LET method with the inclusion of station elevations along a profile (A-A' in Figure 3.6) across the study area from west to east. Note that zero depth represents mean sea level and the depths in the following sections are all relative to mean sea level, instead of the ground surface. Figures 3.11c-d and 3.12c-d are the inverted P- and S-wave velocity models along profile A-A', respectively. The black contours enclose the areas with the diagonal elements of the

resolution matrix greater than 0.4, which are considered well resolved. The new and the traditional methods have similar well-resolved areas since the resolution matrix is calculated from the same initial velocity model and damping parameter. For the P-wave velocity model, the patterns of the velocity anomalies from the LETFT (Figure 3.11c) and the LET method (Figure 3.11d) are similar below 5 km depth. However, the shallow portion (above 5 km depth) of the LET model is not as close to the synthetic model as the LETFT model. These features are more clear in Figures 3.11e and f, which show the velocity difference between the inverted and the synthetic models. The velocity difference shows that the new algorithm recovers velocity model better than the traditional method, especially at shallow depths, which indicates that the neglect of full topography may cause significant bias of the shallow velocity structure estimation. For the S-wave velocity model, the well resolved area (Figures 3.12c-d) is smaller than the P-wave velocity model due to the many fewer S-picks and sparser ray coverage. The S-wave velocity model from the LETFT method is more similar to the true model than that from the traditional LET method (Figures 3.12c-d). Although the well resolved area of the S-wave velocity model at shallow depths (above 5 km) is small, the S-wave velocity from the LETFT method in this depth range is still significantly improved compared to the model from the traditional LET method (Figure 3.12e-f).

In Table 3.1, we show a quantitative comparison of the inversion results between the new LETFT and the traditional LET methods. Earthquake location and origin time errors are the mean values of the differences between the inverted and the true values from all the input master events. The location errors from the LETFT method are 49.9 m, 56.2 m and 75.7 m in longitude, latitude and depth, respectively, whereas the errors from the traditional LET method are 103.7 m, 110.7 m and 157.3 m, which are about twice of those from the LETFT method. Similarly, the origin time error from the LETFT method

(0.018 s) is also much smaller than that from the traditional LET method (0.052 s). The velocity model error is the average value of the differences between the inverted and the true velocities on the well-resolved inversion nodes whose diagonal elements of the resolution matrix are greater than 0.4. The P-wave velocity errors of the LETFT and the traditional LET methods are 0.15 km/s and 0.20 km/s, respectively. The most significant difference between the two methods occurs at shallow depths. The velocity model errors for the well-resolved node above 5 km depth are 0.09 km/s and 0.20 km/s for the LETFT and the LET methods. The S-wave velocity errors of the LETFT and the traditional LET methods are 0.12 km/s and 0.20 km/s, respectively. Above 5 km depth, the S-wave velocity error of the LETFT method is 0.09 km/s, which is much small than the error of 0.21 km/s from the traditional LET method. We also calculate the root mean square (RMS) of the arrival time residuals from each method, which is 0.013 s for the LETFT method and 0.016 s for the traditional LET method. These quantitative comparisons between the two methods show that the LETFT method fits the data better and results in smaller errors in both the velocity model and source parameters.

3.5 Application to Kīlauea volcano

3.5.1 Model results

In the real data test, the velocity model is represented by a set of cuboidal nodes with 3 km horizontal interval. The vertical interval is 3 km above and 5 km below 11 km depth. The velocity of the nodes is updated in the inverse part of tomography and the velocity in the study area is determined by linear interpolation of the nodes below the Earth's surface. The velocity model is solved iteratively and we choose a threshold of terminating the tomographic iteration as 1%, which means that the RMS of the arrival time residuals varies $\sim$0.0014 s from the previous iteration, considering the final RMS

ot the arrival time residuals in our study is about 0.14 s. After we obtain the final P- and S-wave velocity models in Kīlauea, the RMS of arrival time residuals drops from 0.65 sec to 0.14 sec (Figure 3.13) in total. Among them, the RMS of P-wave arrival time residuals drops from 0.55 sec to 0.12 sec, while RMS of S-wave arrival time residuals drops from 0.83 sec to 0.16 sec.

Figures 3.16 and 3.17 show the well resolved areas of the final P- and S-wave velocity models along the profiles shown in Figure 3.6. The model is considered well resolved where the diagonal element of the resolution matrix is greater than 0.4, the same as in the synthetic data test. The white crosses are the projections of the inversion nodes within $\pm$1.5 km distance of the profile line. Since the velocities of the whole area are interpolated by the velocities of the inversion nodes, the area covered the white crosses indicates that the velocity is more accurate in this area. The black dots represent the relocated seismicity within $\pm$0.5 km distance of the profile lines from an updated location catalog based on Lin et al. [2014]. All the inversion nodes are below the ground surface. Zero depth again represents mean sea level.

Our new 3-D P- and S-wave velocity models reveal several interesting features. Figure 3.16 shows the P- and S-wave velocity models along four parallel profiles (1-1', 2-2', 3-3' and 4-4' in Figure 3.6) that are perpendicular to the East Rift Zone (ERZ) and passing through Kīlauea caldera and the Southwest Rift Zone (SWRZ). Profiles 1-1' is across the upper ERZ at Pu'u'Ō'ō crater. To the north of the upper ERZ along profile 1-1', there is a high P-wave velocity zone at $\sim$10 km depth (distance=0-13 km, Figure 3.16a). The Vp velocity beneath Pu'u'Ō'ō (distance=15-16 km) is $\sim$5.5 km/s (Figure 3.16a) at 6-10 km depth. The well resolved area along 1-1' is small and we only can see that the S-wave velocity is about 3.5 km/s at 5-8 km depth beneath Pu'u'Ō'ō (Figure 3.16e, distance=15-16 km). Profiles 2-2' is across the upper ERZ at Makaopuhi crater.

A high Vp zone is also observed to the north of ERZ (distance=0-13 km) with a velocity of ∼7.5 km/s (Figure 3.16b). It expands to 5-9 km beneath Makaopuhi (distance=14-15 km) in the upper ERZ. The P-wave velocity beneath the Hilina fault system (HFS, distance=21-22 km) is about 5.3 km/s. Along 2-2' profile, Vs varies from 4.0 km/s in the north (distance=0-11 km) to 3.5 km/s beneath Makaopuhi (distance=14-15 km, Figure 3.16f). Profile 3-3' passes through Kīlauea caldera and the Koa'e fault system (KFS, distance=13-14 km). In Figure 3.16c and g, high Vp (∼7.4 km/s) and Vs (∼4.1 km/s) dominate the depth between 5 km and 11 km beneath Kīlauea caldera. Beneath the KFS, Vp and Vs decrease to about 6.1 km/s and 3.6 km/s, respectively. The most significant feature along profile 4-4' that is across the SWRZ is a high Vp zone at 5-9 km depth with a velocity of ∼7.4 km/s (distance=7-8 km, Figure 3.16d). However, Vs is about 3.6 km/s in this area (Figure 3.16h), which is not significantly higher than the surrounding area.

Figure 3.17 shows the P- and S-wave velocity model along Kīlauea's summit caldera, the ERZ and the SWRZ (profiles B-B'-B" and C-C'-C" in Figure 3.6). At Kīlauea caldera of profile B-B'-B", the Vp and Vs are ∼3.9 km/s and ∼2.6 km/s from the Earth's surface to 1 km depth below MSL. which are decreased to ∼3.7 km/s and ∼2.3 km/s at 2 km depth. This variation of Vp and Vs is more clear along a depth profile beneath Kīlauea caldera (insets in Figure 3.17a and c). Another significant feature is that the high Vp zone (7.1-7.8 km/s) is not continuous in the ERZ (Figure 3.17a). It extends from 5-14 km depth at Kīlauea caldera to 5-9 km depth at Makaopuhi in the upper ERZ. The Vp drops to about 5.5 km/s beneath Pu'u'Ō'ō. The Vs along Kīlauea caldera and the ERZ has a similar trend as Vp (Figure 3.17c). However, the difference is that a low Vs zone is observed at 5-9 km depth at Makaopuhi. Along profile C-C'-C", there are several discontinuous high Vp zones divided by the surrounded relatively low

Vp bodies in the SWRZ (Figure 3.17b). High Vp and low Vs zone are still visible at 5-9 km depth close to Kīlauea caldera (Figure 3.17b and d), which are observed in Figure 3.16d and h.

3.5.2 Robustness test

The P- or S-wave arrival time data usually contain noise which could introduce errors in inverted velocity model. In order to make sure our P- and S- wave velocity model is reliable, we conduct a robustness test by adding Gaussian noise to P- and S-wave arrival time data. The mean and standard deviation (STD) are obtained from the arrival time residual (Figure 3.13) of the inverted P- and S-wave velocity model. Because the arrival time residual of the inverted velocity model consists of errors caused by model and noise of data, it could be larger or the same level as the noise of data. In this study, we add Gaussian noise with 0.0018 s mean and 0.14 s STD to the P- or S-wave arrival time data, then obtain inverted P- and S-wave velocity models with the noisy data. Figure 3.18 compares the results of well resolved areas from original and noisy data along profile A-A' in Figure 3.6. Figure 3.18a and b are P-wave velocity models. The high and low velocity features are basically same as each other, except that there is a little difference at the edge of the well resolved area (distance=45-55 km). S-wave velocity from the original data (Figure 3.18c) and noisy data (Figure 3.18d) are similar to each other as well, noticing only the absolute values of the velocity models have a little difference. For instance, the high velocity zone with distance at 20-25 km in Figure 3.18c is $\sim$0.1 km/s smaller than that in Figure 3.18d. The result of the robustness test indicates that our P- and S-wave velocity models are stable when data contains noise and are reliable.

3.6 Discussion

In traditional body wave LET, topography is usually neglected and sometimes only station elevations are used to approximate the effect of topography. This approximation could bring errors to the prediction of arrival time and partial derivative matrices of source and velocity parameters in the forward part of tomography. It could also bring errors to the velocity models in the inverse part of tomography. The amplitude of the errors depends on the variations of local topography. In our new LETFT algorithm, this kind of systematic error is avoided by including the full topography in both the forward and inverse parts of tomography and only updating the velocity below the Earth's surface. The synthetic data tests show that the new LETFT method is able to recover the true velocity better in both the shallow and deep parts of the crust compared to the traditional LET method. The calculated errors in the inverted velocity model, earthquake locations and arrival time residuals also provide quantitative proof of the improvement of the new method over the traditional one.

In the real data application to Kīlauea volcano, we observe several notable Vp and Vs structures. At Kīlauea caldera, the Vp ($\sim$3.9 km/s) and Vs ($\sim$2.6 km/s) are higher than the surrounding area from the surface to 1 km depth below MSL, which are decreased to 3.7 km/s and 2.3 km/s at 2 km depth. The abundant seismicity indicates that this area is very active. This low Vp zone was also observed by Thurber [1984] and Rowan and Clayton [1993] and was interpreted as the locus of a magma complex. However, low Vp can be caused by many factors, such as rock composition, fracture and fluid content [Lin, 2015]. The low Vs in the area gives more support to the hypothesis of the magma complex. Below 2 km depth, there is a high Vp zone ($>$7.0 km/s) at 5-14 km depth and high Vs zone ($>$3.9 km/s) at 5-11 km depth. This high Vp

zone was also observed in the previous studies [Dawson et al., 1999; Monteiller et al., 2005; Park et al., 2009; Lin et al., 2014] and was interpreted as consolidated intrusive gabbro-ultramafic cumulates [Okubo et al., 1997; Haslinger et al., 2001; Hansen et al., 2004]. The high Vs in this area also supports this interpretation.

To the north of the upper ERZ from Makaopuhi to Pu'u'Ō'ō, we observe high Vp (>7.0 km/s) and Vs (>3.9 km/s) zone at 8-12 km depth. In the upper ERZ, there is also a high Vp (>7.0 km/s) zone at 5-9 km depth beneath Makaopuhi. These high Vp zones are consistent with the high Bouguer gravity anomaly in this area [Kinoshita, 1965; Kauahikaua, 1993], which means that there are dense materials in the subsurface from the mantle plume [Montelli et al., 2004]. In the ERZ, the high Vp zone is not continuous and we observe a relatively low Vp (∼5.6 km/s) zone at 6-10 km depth beneath Pu'u'Ō'ō. This low Vp is similar to the velocity of olivine basalt or plagioclase basalt from the oceanic crust [Schreiber and Fox, 1977; Christensen, 1996]. Considering low Vs is not observed at 6-10 km depth beneath Pu'u'Ō'ō, we speculate that this area is dominated by the basaltic rocks of remnant oceanic crust. The similar discontinuous high Vp (>7.0 km/s) zone is also found below 4 km depth in the SWRZ. The high Vp zone of 4-10 km depth at the bend of the SWRZ (around C' on Figure 3.6) indicates a concentrated intrusion zone as a barrier to magma propagation [Klein et al., 1987]. Rowan and Clayton [1993] interpreted that this feature changed the direction of both the SWRZ and local stresses.

In the HFS, there is a thin high Vp layer (∼7.0 km/s) at 4-6 km depth and Vp decreases to as low as ∼5.7 km/s at 6-11 km depth, which has a sharp contrast to the high Vp in the ERZ. The low Vp zone was also observed by Rowan and Clayton [1993] and Haslinger et al. [2001]. The Bouguer gravity anomaly in the HFS is smaller than that in the ERZ [Kinoshita, 1965; Kauahikaua, 1993], which indicates that there are

lighter remnant oceanic crustal materials (compared to the dense materials from the mantle plume) at the subsurface of the HFS. With the evidence above, we suggest that the high Vp layer at 5-6 km is the intrusive ultramafic gabbro sills and the low Vp zone at 6-11 km beneath the HFS is the remnant oceanic crust. The horizontal linear distribution of seismicity at 6 km depth supports our suggestion since the seismicity likely occurs at the fractures caused by intrusive ultramafic gabbro sills. To the south KFS, Vp is $\sim$6.1 km/s and Vs is $\sim$3.5 km/s at 5-9 km depth. A similar Bouguer gravity anomaly [Kauahikaua, 1993] as in the HFS and Vp/Vs ratio of $\sim$1.74 indicate that this area is occupied by the remnant oceanic crust.

3.7 Summary

In this study, we developed a new local earthquake tomography algorithm with including of full topography. The new algorithm naturally avoids the errors introduced by topography in local earthquake tomography. In synthetic test, the results indicate that the new algorithm significantly improves the accuracy of the velocity models compared to the traditional method with only including the station elevations. In the real data application at Kīlauea volcano, we discovered a few interesting velocity features. For instance, the high Vp and Vs in the ERZ is discontinuous. And at Kīlauea caldera, there are a low velocity zone at about 2 km depth and a high velocity zone under 5 km depth.

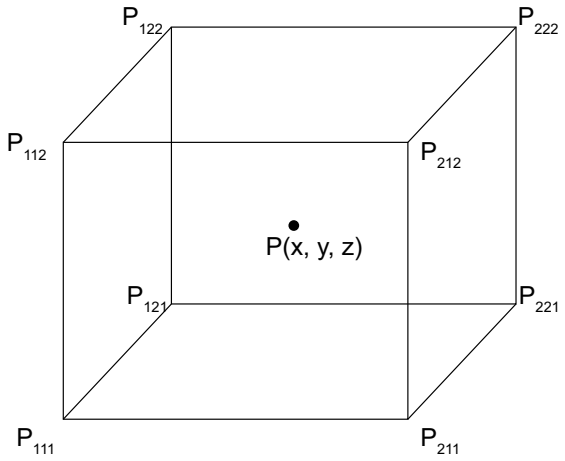

Figure 3.1: An illustration of trilinear interpolation.

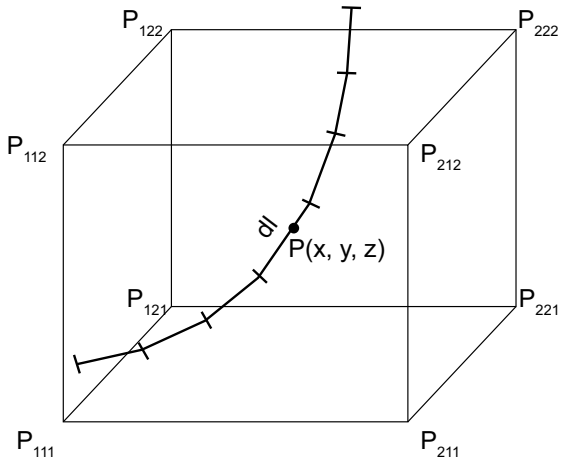

Figure 3.2: An illustration of partial derivatives of slowness calculation.

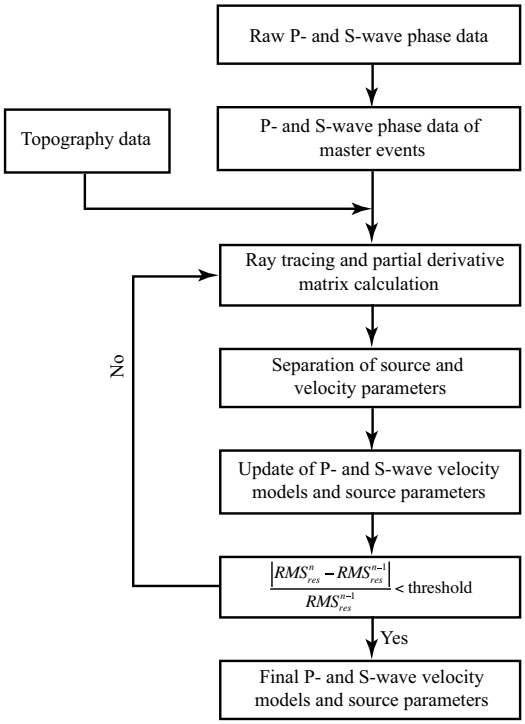

Figure 3.3: A work flowchart of our local earthquake tomography with full topography algorithm presented in this study.

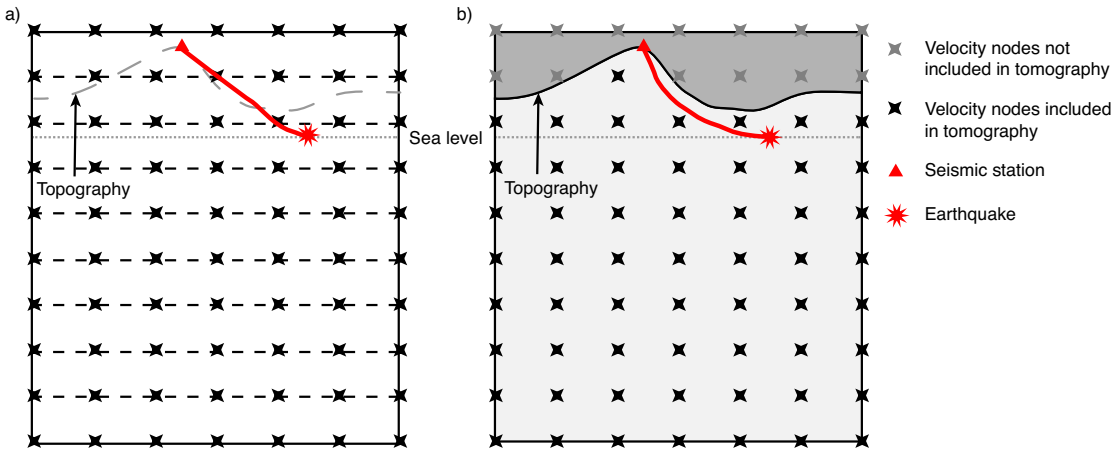

Figure 3.4: Schematic illustration of the difference between the traditional LET method (a) and the LETFT method (b).

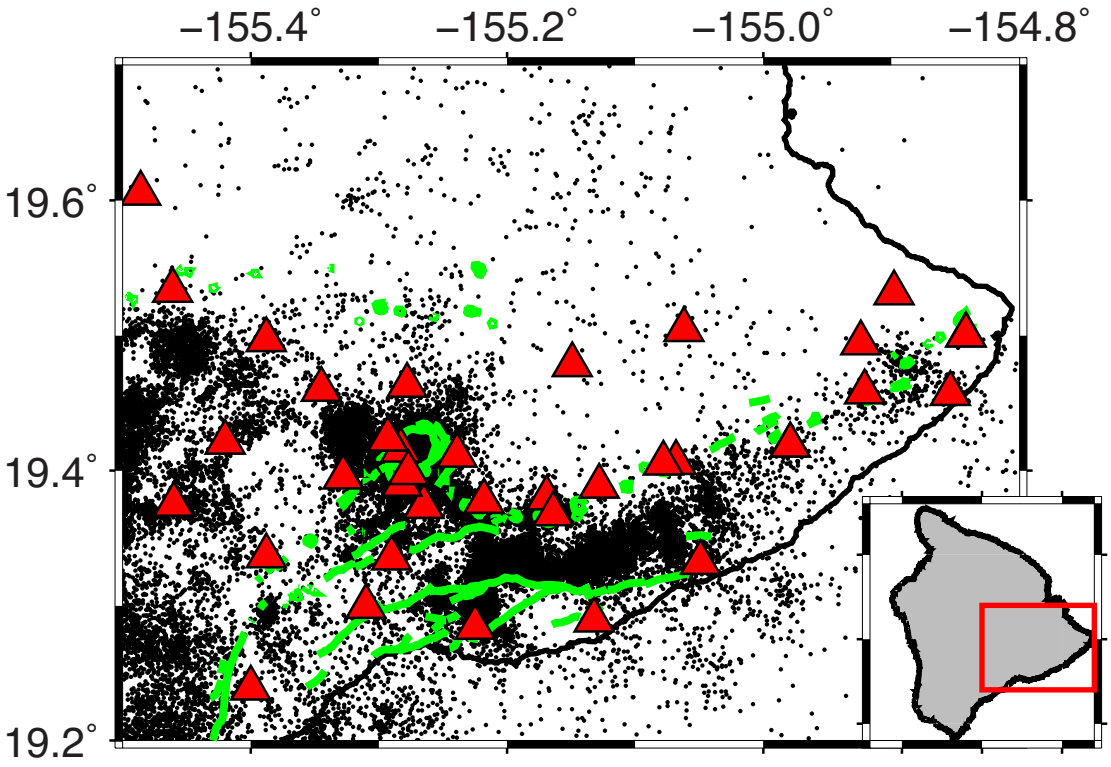

Figure 3.5: Seismic events and stations used in this study. Black dots represent earthquakes from 1992 to 2009 recorded by the seismic stations (red triangles) at the Hawaiian Volcano Observatory. Green lines are the surface traces of the major faults and craters. Red box in the inset map shows the location of our study area on the island.

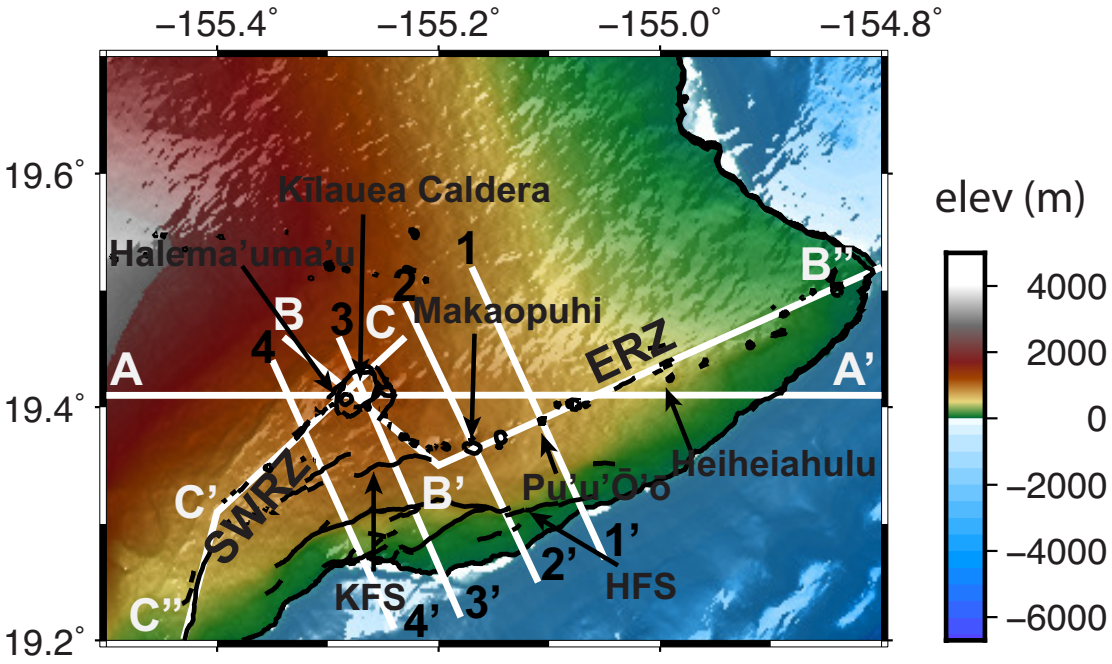

Figure 3.6: Geological features in our study area, including the Kīlauea summit caldera, the east rift zone (ERZ), the Southwest Rift Zone (SWRZ), the Hilina Fault System (HFS), and the Koa'e Fault System (KFS). The white lines are the profiles for the cross-sectional views in the following figures. The background is the topography basemap from the Hawaiian multibeam bathymetry synthesis (http://www.soest.hawaii.edu/hmrg/multibeam).

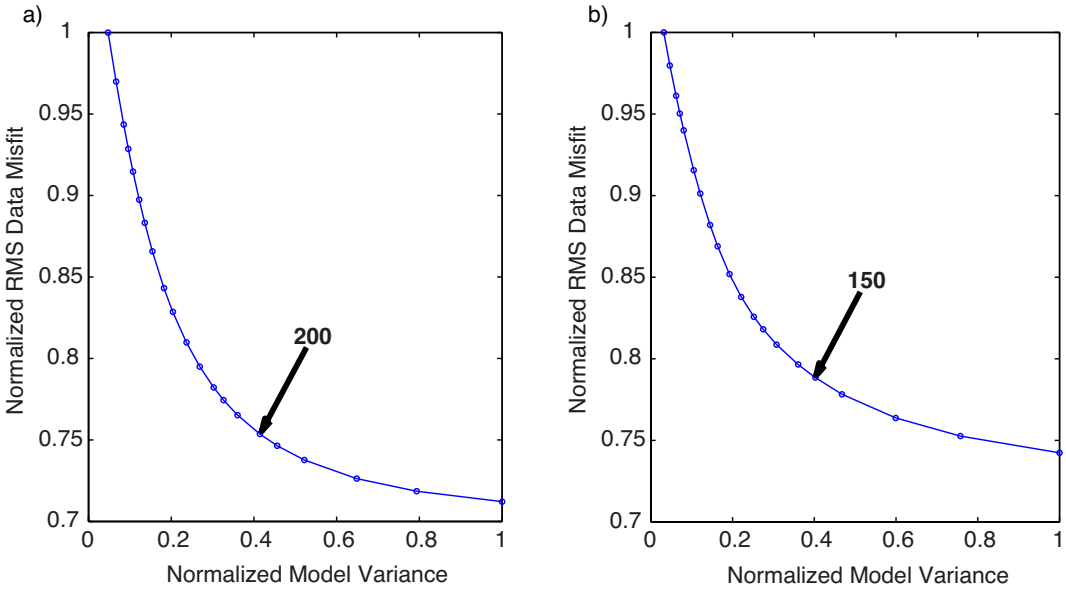

Figure 3.7: Trade-off curves between data misfit and model variance for P-wave (a) and S-wave (b) velocity models. Damping parameters of 200 and 150 are selected for P- and S-models, respectively.

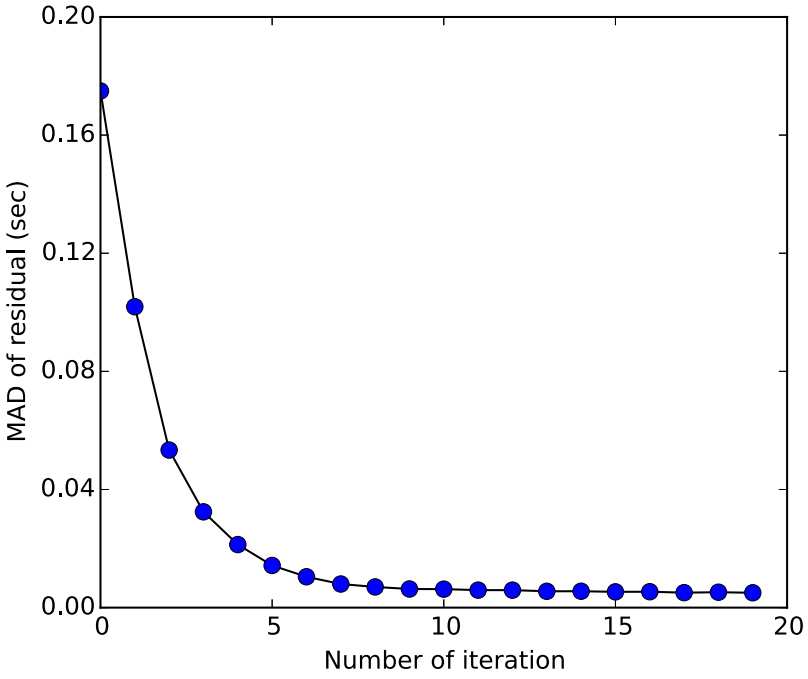

Figure 3.8: The convergency of LETFT method. The horizontal axes is the number of iteration. The vertical axes is the mean absolution deviation (MAD) of the residuals.

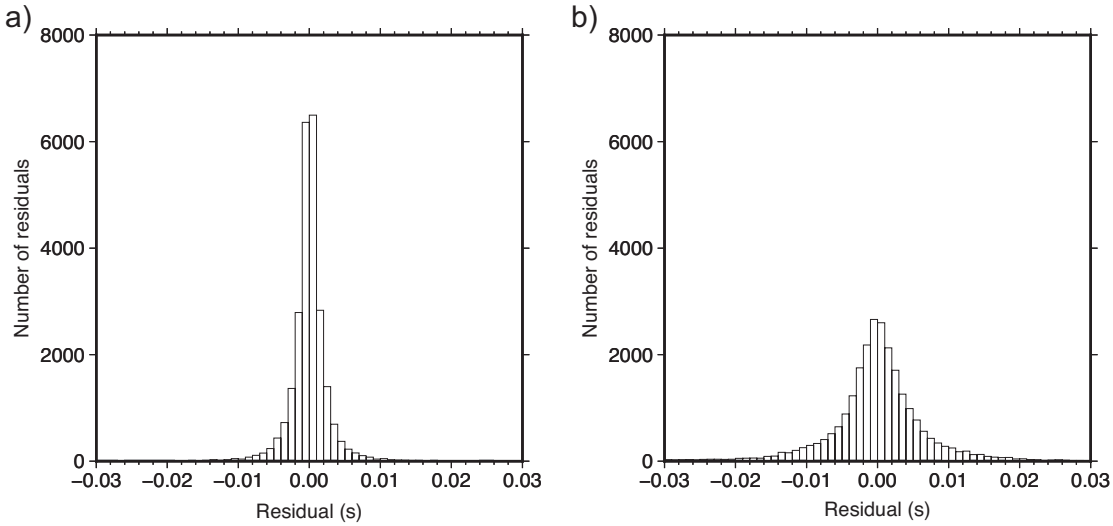

Figure 3.9: Histograms of the travel time residuals from the inverted 3-D velocity model of (a) LETFT and (b) traditional LET method. The RMS residual of LETFT and traditional LET method are 0.013 sec and 0.016 sec, respectively.

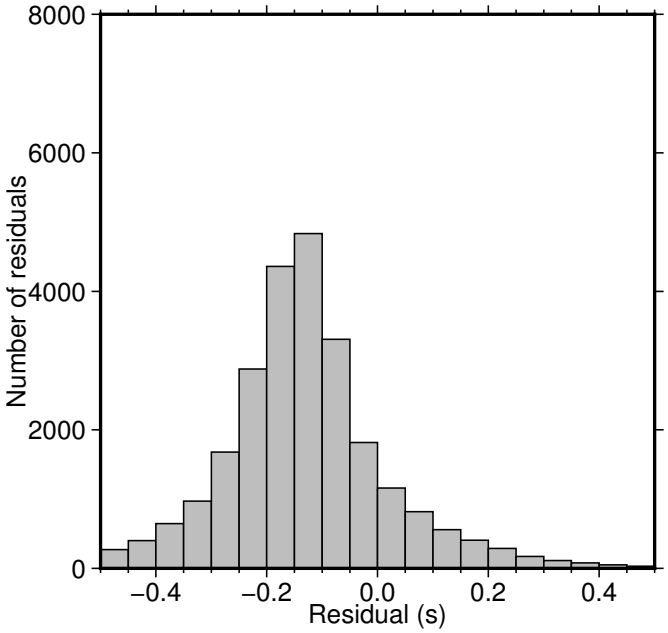

Figure 3.10: Histograms of the travel time residuals from the initial velocity model. The RMS residual is 0.18 sec.

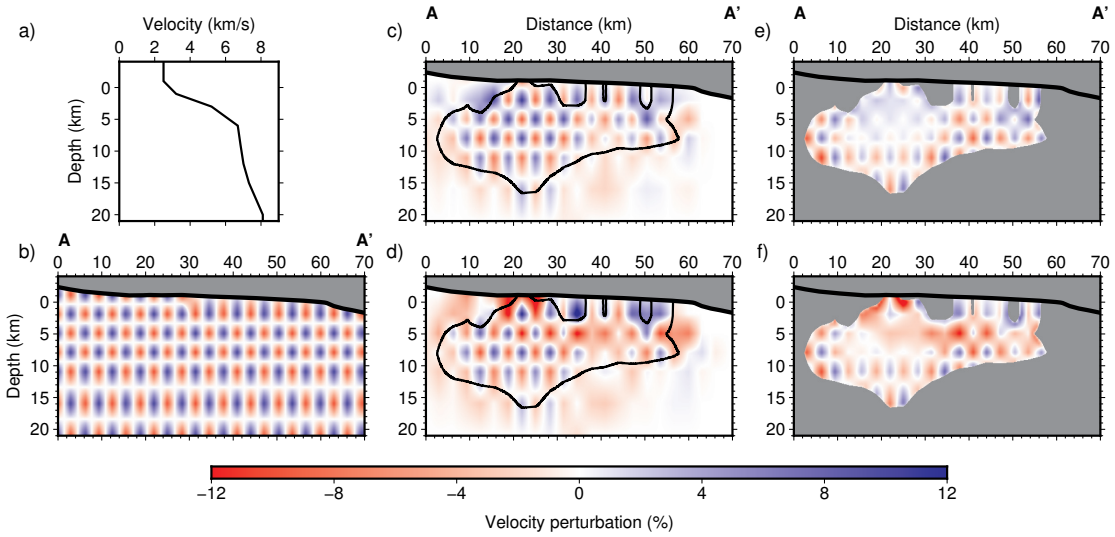

Figure 3.11: Synthetic data test comparison of P-velocity model between the traditional local earthquake tomography (LET) and local earthquake tomography with full topography (LETFT). (a) 1-D initial P-wave velocity model used in this study. (b) True P-velocity model used in the synthetic data tests along profile A-A' in Figure 3.6. Thick black curve at top illustrates the local topography. (c, d) Inverted P-velocity model from the LETFT and the LET, respectively. Black contours enclose the area with the diagonal element of the resolution matrix greater than 0.4, which are considered well-resolved. (e, f) P-velocity model difference for the well-resolved area of the LETFE and the LET model relative to the true model, respectively.

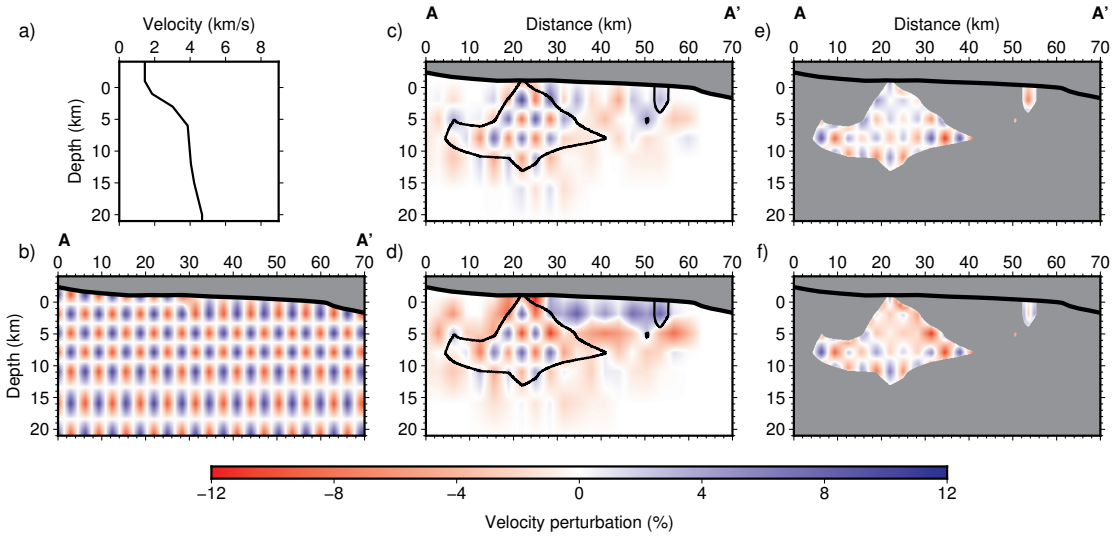

Figure 3.12: Synthetic data test comparison of S-velocity model between the traditional local earthquake tomography (LET) and local earthquake tomography with full topography (LETFT). (a) 1-D initial S-wave velocity model used in this study. (b) True S-velocity model used in the synthetic data tests along profile A-A' in Figure 3.6. Thick black curve at top illustrates the local topography. (c, d) Inverted S-velocity model from the LETFT and the LET, respectively. Black contours enclose the area with the diagonal element of the resolution matrix greater than 0.4, which are considered well-resolved. (e, f) S-velocity model difference for the well-resolved area of the LETFE and the LET model relative to the true model, respectively.

Table 3.1: Tomographic inversion result comparison between the LETFT and traditional LET methods. Definitions of the errors are explained in the main text. RMS is short for Root Mean Square.

	LETFT	LET
Location error in longitude (m)	49.9	103.1
Location error in latitude (m)	56.2	110.7
Location error in depth (m)	75.7	157.3
Origin time error (sec)	0.018	0.052
P velocity model error (km/s)	0.15	0.20
Shallow (z$<$5 km) P velocity error (km/s)	0.09	0.20
S velocity model error (km/s)	0.12	0.21
Shallow (z$<$5 km) S velocity error (km/s)	0.09	0.21
RMS of residual (sec)	0.013	0.016

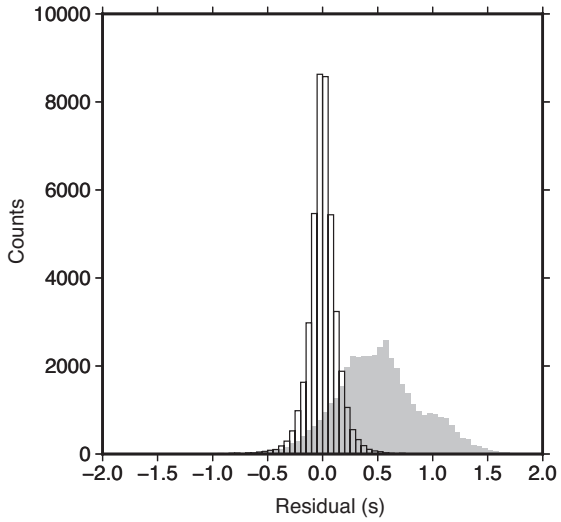

Figure 3.13: Histograms of the travel time residuals from the initial 1-D velocity model (gray) and the inverted 3-D velocity model (black and white) for Kīlauea area. The RMS residual is reduced from 0.65 to 0.14 s.

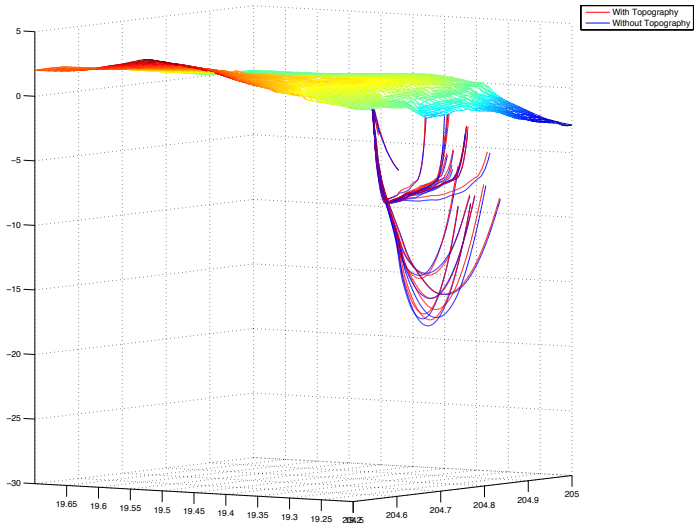

Figure 3.14: Comparison of 20 rays calculated from the final models of the LETFT method (Red lines) and the traditional LET method (Blue lines) with a common seismic station.

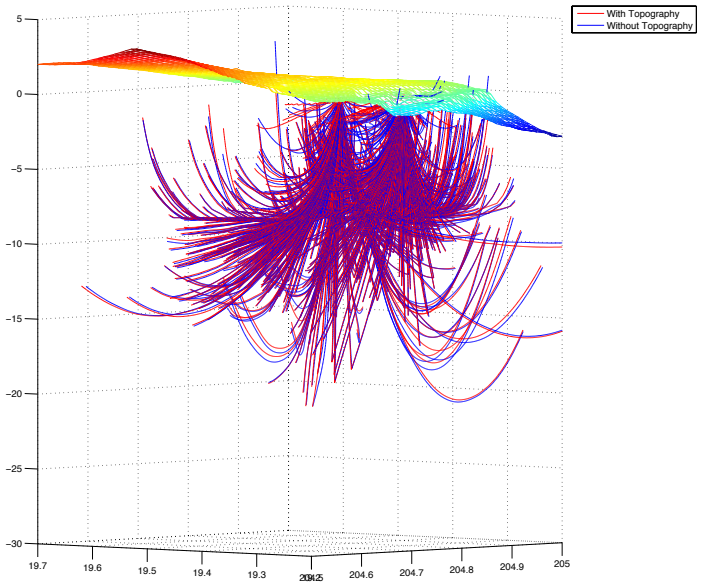

Figure 3.15: Comparison of an example of 1000 rays calculated from the final models of the LETFT method (Red lines) and the traditional LET method (Blue lines).

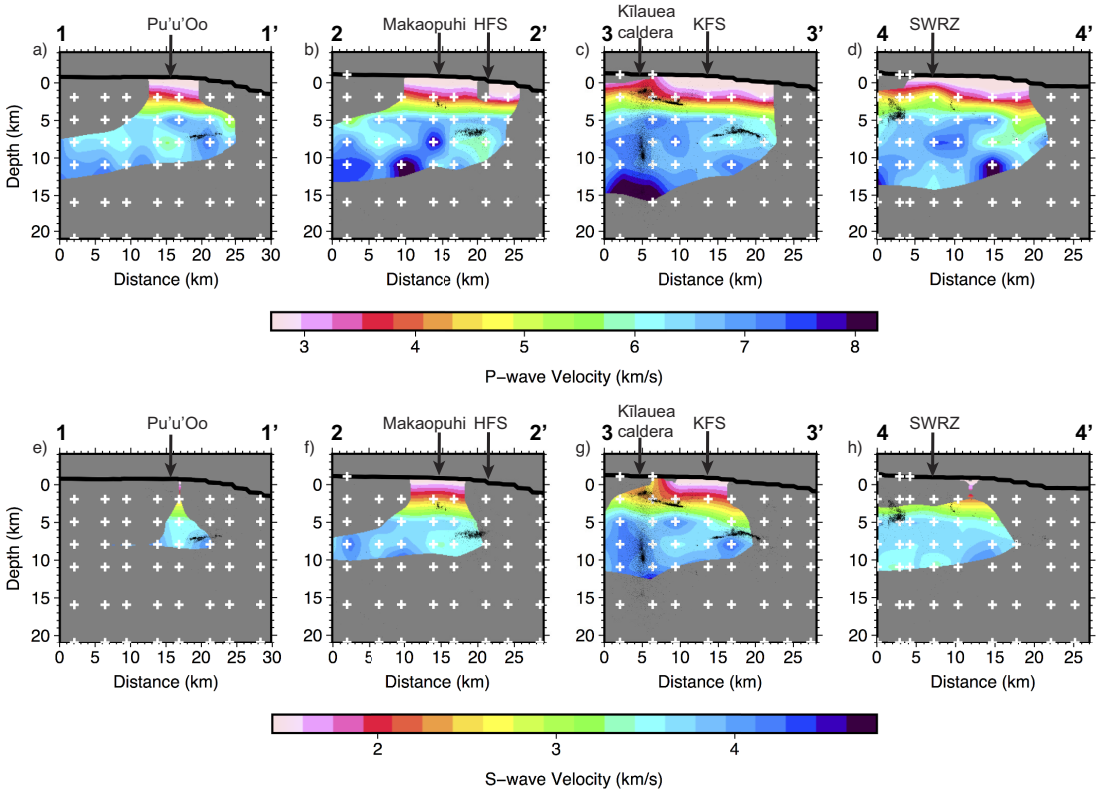

Figure 3.16: Cross-sectional views of the inverted P-wave velocity model (a-d) and S-wave velocity model (e-h) along profiles 1-1', 2-2', 3-3' and 4-4' in Figure 3.6. White crosses are the inversion nodes projected onto the vertical profiles. Black dots are relocated seismicity from [Lin et al., 2014]

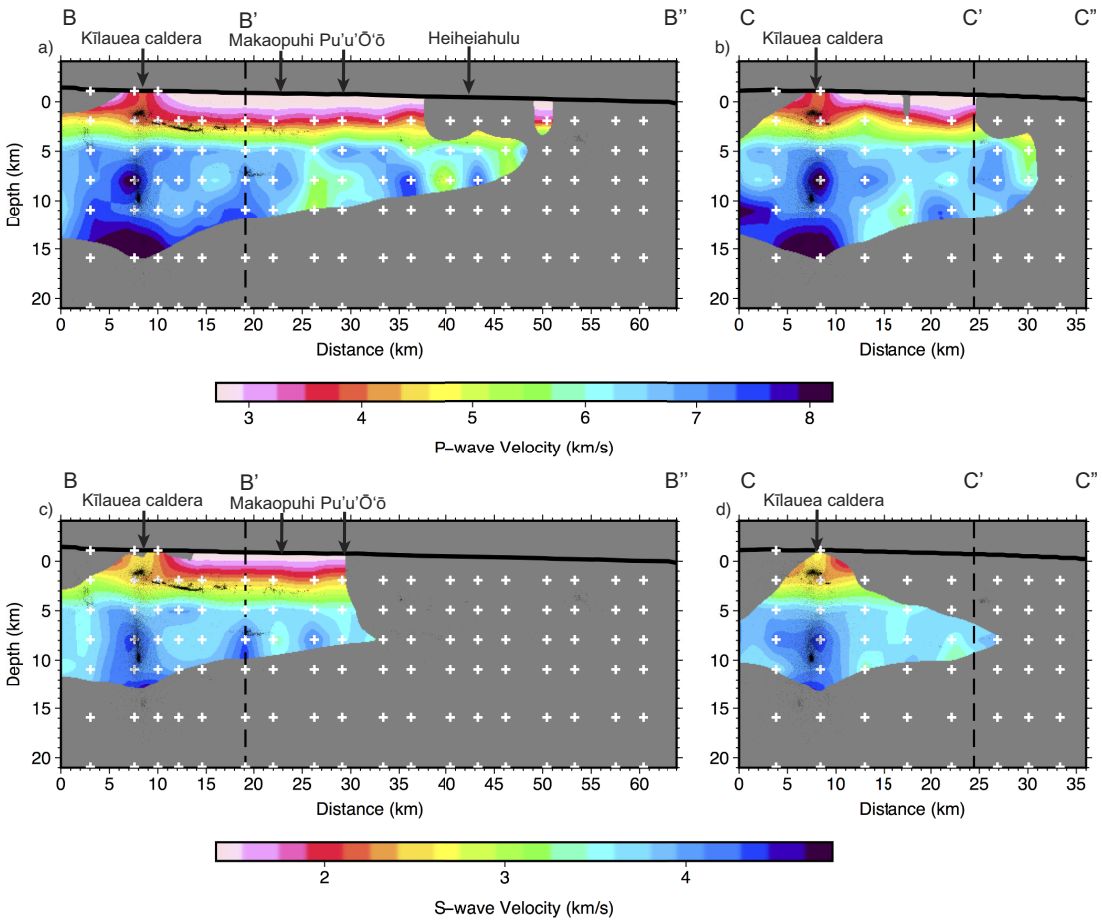

Figure 3.17: Cross-sectional views of the inverted P-wave velocity model (a-b) and S-wave velocity model (c-d) along profiles A-A'-A" and B-B'-B" in Figure 3.6. White crosses are the inversion nodes projected onto the vertical profiles. Black dots are relocated seismicity from [Lin et al., 2014]

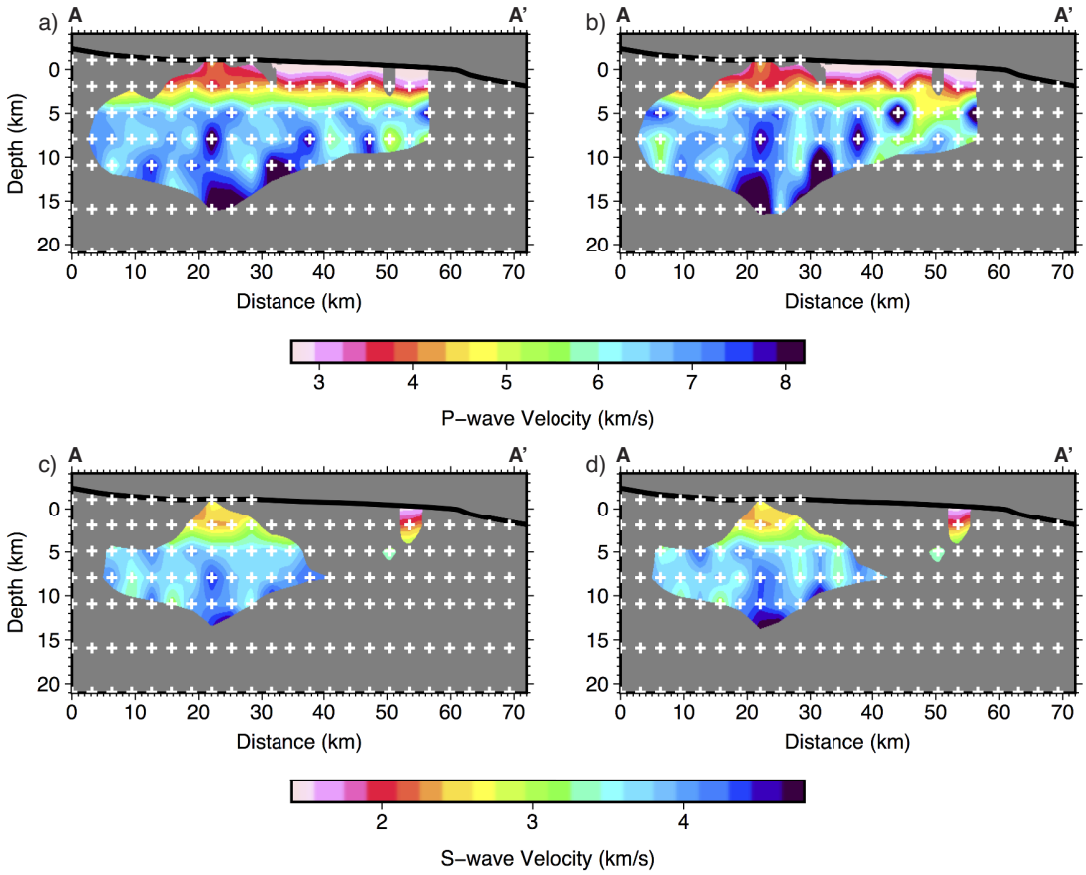

Figure 3.18: Robustness test. (a) and (c) are the inverted P- and S-velocity models from the original catalog data. (b) and (d) are the inverted P- and S-wave velocity models from data with 0.0018 s mean and 0.14 s standard deviation Gaussian distributed noise.

Chapter 4

Induced long duration seismic events

4.1 Overview

Injecting CO_2 in enhanced oil recovery (EOR) field is a practical and economic way for carbon sequestration. Since the injection of CO_2 can increase the pressure in the oil reservoir, seismic monitoring is important for evaluation of the risk. The previous studies mainly focused on the induced micro-earthquakes. However, the fluid involved source in the oil reservoir is also able to generate other types of seismic events. In this study, we focus on a new type of induced seismic event, which is called long duration (LD) event. The LD events lasts from 10s to 100s sec, which is much longer than the typical micro-earthquakes (several seconds or shorter). In the following, we use different method to find characteristics of the LD events and calculate the rate-per-night of the LD events, which provides the useful information to understand the nature of the LD events.

4.2 Data

We installed three broadband seismic stations (Test1, 2 and 3) in an EOR field near Houston in October 2011. The frequency response band of the seismic stations is be-

tween 120 seconds and 150 Hz. The study area is composed of two blocks divided by a southwest-northeast normal fault (Figure 4.1). The fault is considered as a seal preventing the oil or gas transporting from one block to the other. The depth of the reservoir is about 1600 m, which consists of two sand stone units with one called Upper Frio of $\sim$60 m thick and the other called Lower Frio of $\sim$200 m thick. Above the reservoir, there is a $\sim$200 m thick Anahuac shale acts as impermeable caprock [Karegar et al., 2015]. The continuous injection of CO_2 and saline water starts in late 2010, and lasted through our two-year monitoring. The production of oil began in February, 2012. Figure 4.1 shows the locations of the three seismic stations. The distance is 0.8 km between Test1 and Test2, and 1.5 km between Test2 and Test3. Test1 and Test2 are located in the north block and Test3 in the south block. We collected 2 years of continuous seismic data between October 2011 and October 2013. During this period, there is a 2-month data gap at Test1 in early 2012 due to power outage. Test3 was flooded between December 2011 and January 2012, which affected the data quality during this time.

In the continuous seismic data, we observe a type of LD events instead of the typical micro earthquakes. The duration (the time from the start to the end of an event) of the LD events varies from $\sim$30 to $\sim$300 seconds, which is much longer than the micro earthquakes. Figure 4.2 shows an example of a one-day seismogram and a LD event. The amplitude of the LD event gradually increases at the beginning and decreases at the end of the event window. The LD events happen in both daytime and nighttime. However, because of the low signal-to-noise ratio during the daytime due to human activities, we only use the data in the nighttime in the following analyses.

4.3 Characteristics of long duration events

4.3.1 Spectrum analysis

We first calculate the spectrogram of the one-day waveform as shown in Figure 4.3. The purpose is to visually inspect the frequency band of the LD events. We could also separate the signal of the LD event from the noise by taking advantage of the frequency information. The spectrogram of the one-day waveform is shown in Figure 4.3a. The left and right halves of the spectrogram are the night and day, respectively. The frequency band of the LD event is from 0.5 Hz to 6 Hz, which is much lower than the LD events (10-80 Hz) in hydraulic fracturing area of shale-gas reservoir [Das and Zoback, 2013]. The band of the low-frequency continuous noise from the ocean and earth oscillations is below 0.5 Hz. There are also some narrow band high frequency noise at 3 Hz, 4 Hz, 4.7 Hz and 10 Hz, which last through the day and night. In the daytime, because the high frequency noises between 0.5 Hz and 8 Hz arise, it is hard to distinguish the frequency information of the LD events.

4.3.2 Comparison with regional earthquake data

In order to reveal more characteristics of the LD events, we compare the LD events with other types of seismic events, such as natural earthquakes and volcanic tremors. We search earthquakes near (less than 500 km distance) our study area between 2011 and 2013 from the National Earthquake Information Center catalog. The nearest earthquake with a significant magnitude is a M4.8 earthquake, at 08:12, May 17th, 2012 (Coordinated Universal Time). This reference earthquake (white star in Figure 4.4) is about 285 km from the northeast side of our stations. We also chose the nearest reference seismic station US.NATX from the USArray to compare with our stations. The one-day seismograms for the reference earthquake are shown for both our Test3 and the

reference station in Figure 4.5a. Not only does Test3 record the reference earthquake, but also records many LD events. In contrast, the reference station only records the reference earthquake, which indicates that the LD events are local effects and attenuate fast. On the zoomed seismograms in Figure 4.5b, the waveform of the reference earthquake is very different from the LD event. The earthquake has an impulsive arrival compared to the smooth variation of the LD event's amplitude. It is much easier to pick the arrival time for the earthquake than for the LD event. Typical earthquakes are caused by sudden slip on faults by stress release, which is mainly an elastic problem. The environment for the LD events is more complex with the effect of fluid in the reservoir.

4.3.3 Comparison with volcanic tremor

A type of commonly recognized fluid-involved seismic signal is volcanic tremor, often associated with magma transportation. Thus, we compare the LD events in this study with the volcanic tremors. Figure 4.6 shows the comparison of an LD event recorded by Test3 and a volcanic tremor recorded at Arenal volcano, Costa Rica, 2011. The waveforms of the two types of seismic events look very similar. In addition, the typical frequency of volcanic tremors is 1-5 Hz [McNutt, 2011]. The majority of our LD event energy is also concentrated in 0.5-6 Hz band, which indicates that the mechanism of the LD events may have some similarities with volcanic tremors.

4.3.4 Comparison with traffic data

Given the distance of $\sim$500 m between our stations and the main road, it is natural to raise the question of whether the LD events are caused by the traffic or heavy trucks on the road. In order to examine this, we conducted an experiment on the traffic effect near a similar road. We deployed a test station $\sim$30 m away from the road and record

the signals due to the traffic for an hour. The times that took heavy trucks to pass by are recorded by hand so that we can easily extract the corresponding seismic signals. Figure 4.7 compares the LD event (Figure 4.7a) and the signals of heavy trucks (Figure 4.7b and c). We only plot the vertical component seismograms here. Although the shape of the signals of the heavy trucks looks similar the LD event, their durations are quite different. The duration of the LD event in Figure 4.7a is $\sim$200 seconds long, whereas the signals of heavy trucks last only $\sim$5 seconds. This comparison excludes the possibility that the LD events were caused by the traffic.

4.3.5 Particle motion of LD events

Particle motion is the ground motion recorded by seismometers, which can help understand the mechanism of the seismic events. Figure 4.8 shows the particle motions of a single LD event at the beginning part, the middle part and the peak. The time length of the plotted particle motion is 4 seconds. Figures 4.8b, c and d are the particle motions of the beginning part at three orthogonal planes with one horizontal and two vertical planes, respectively. Similarly, Figures 4.8e, f and g are the particle motions of the middle part and Figures 4.8h, i and j are the particle motions at the peak. In the beginning part of the LD event, there seems no significant polarization of the particle motions, which is likely generated by isotropic inflation or explosive source. In the middle part, there is a slight polarity of the particle motions in the vertical plane, which indicates the source migrates to a small fault or tunnel with a preferred direction. Finally, at the peak of the LD event, the polarity of the particle motions is more distinct and the principal direction is west-east direction. The change in the polarity direction indicates that the source of the LD event may not be a single source and could migrate from an isotropic inflation source to complex fractures with different directions.

4.3.6 Event rate

To extract the LD event signals, we need to pre-process the data to increase the signal to noise ratio. Since the LD events are mainly concentrated in the frequency band of 0.5-6 Hz as shown in Figure 4.3, we apply 0.5-6 Hz bandpass filter to all our data in order to enhance the LD event signals. Because the noise level in the daytime is much greater than that at night, we only use seismic data at nighttime to avoid any potential artificial picks during daytime in the following analyses. Given the continuous injection day and night throughout our study period, the rate of the LD events should not have significant difference between daytime and nighttime. The characteristics of the LD events at night should be representative for the entire day.

Our LD event picking method is similar to the short-term average over long-term average ratio (STA/LTA) method, which has been widely used in earthquake auto-detection [e.g., Allen, 1982; Boschetti et al., 1996]. In this study, we use the average over the entire 12-hour nighttime instead of a "long-term" (typically tens of seconds) average. This method can be expressed by:

$$\phi(t) = \overline{A^2_{stw}} / \overline{A^2_{whole}} \tag{4.1}$$

where $\overline{A^2_{stw}}$ is the short-term average of the squared amplitudes, $\overline{A^2_{whole}}$ is the average of the squared amplitudes in the entire 12-hour window per night, and $\phi(t)$ is the objective function. Since the duration of the LD events ranges from $\sim$30 to $\sim$300 seconds, the short-term time window should not be too long or too short in order to capture the entire event sequence. After testing different lengths of short-term time windows, the results are stable when the short-term time window varies between 20 and 80 seconds. For the final results of this study, we use 50 seconds as the short-term time window. The short-term time window moves from the start to the end of each half day seismogram

and $\phi(t)$ is calculated for short-term time window of each segment. When $\phi(t)$ is larger than a pre-defined threshold, we set the event to be a potential LD event. We try different thresholds and found that the results did not change significantly for thresholds between 2 and 5. Therefore, we chose 3 as the final threshold. Given the shape and duration of the LD event, $\phi(t)$ increases when the short-term time window meets an LD event and then declines when the window leaves the LD event. If this trend of $\phi(t)$ is confirmed, then the final LD event is detected. This process prevents duplicate count of the LD events since $\phi(t)$ is always larger than the threshold before the short-term time window leaves the LD event.

Each component of the three stations is used for automatic detection of the LD events separately. If the method is reliable, the results from the three components of the same station should be similar. After automatically picking the LD events of the two years of seismic data, we calculate the rate of the LD events per night. Figure 4.9 shows the rate per night (RPN) of the LD events for Test1, Test2 and Test3. The average RPN of the LD events is 12, and we observe about 8700 LD events in two years.

4.3.7 Comparison with reservoir pressure data

We compare the results from all the three stations with the oil reservoir pressure data recorded in a well close to Test1 and Test2 (Figure 4.9). The pressure at the reservoir kept increasing from 2011 and arrived at the peak of $\sim$2850 psi before the production in February 2012, and then dropped to $\sim$2700 psi when the production began. After that, the pressure at the reservoir became stable. The trend of the LD event rate at Test2 is similar to the variations of the reservoir pressure. The highest part of the LD event rate is at the period of rapid change of reservoir pressure between February to June 2012, which indicates that the LD event rate is correlated with the variation of pressure at the

reservoir. Although the LD event rate of Test1 was high in the first half of 2012, it is difficult to determine the peak of the LD event rate since Test1 lacks about two months of data due to power outage. Unlike at Test1 and Test2, the LD event monthly rate of Test3 in the south block has a different pattern with poorer correlation with the reservoir pressure than that in the north block. We attribute this to the fault between the north and south blocks, which may seal the injected CO_2 in the north block. As a result, some events were not recorded by Test3 for their low energy and fast attenuation.

4.4 Summary

In above, we talked about a new type of induced seismic event related to carbon sequestration in an EOR field, which is called long duration (LD) event for it is last much longer than the typical micro-earthquakes. The LD event has an envelop-like shape of waveform, which has a smooth head and tail. The waveform of LD event is not like typical earthquake, but similar to volcanic tremor. This evidence indicates that the mechanism of LD event is not likely a pure elastic process and could be associated with fluid. The average rate-per-night of LD event is 12. And the LD event rate is correlated with the variation of pressure at the reservoir.

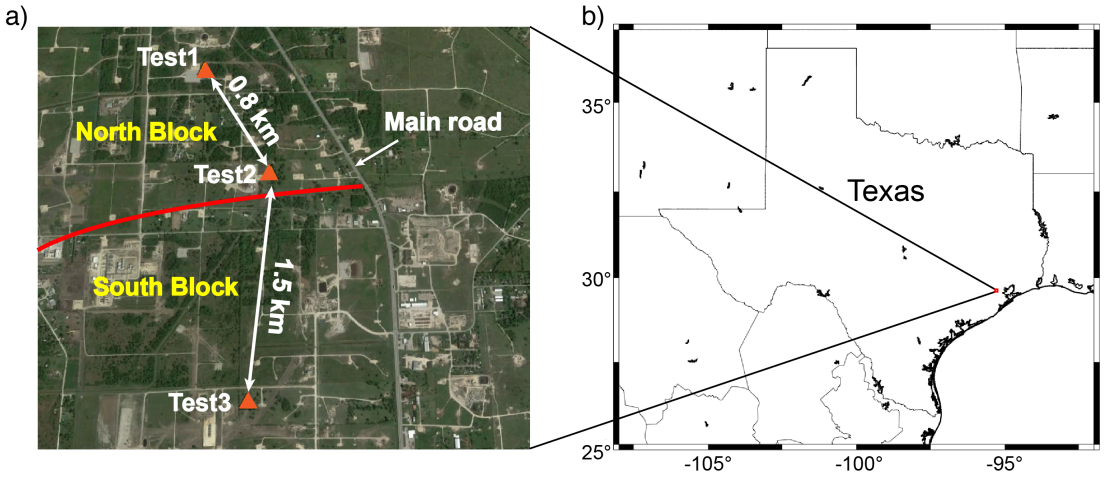

Figure 4.1: A map of our site (a) and its location in Texas (b). Orange triangles are our 3 seismic stations. The red curve is the main fault that divides the study area to the north and south blocks.

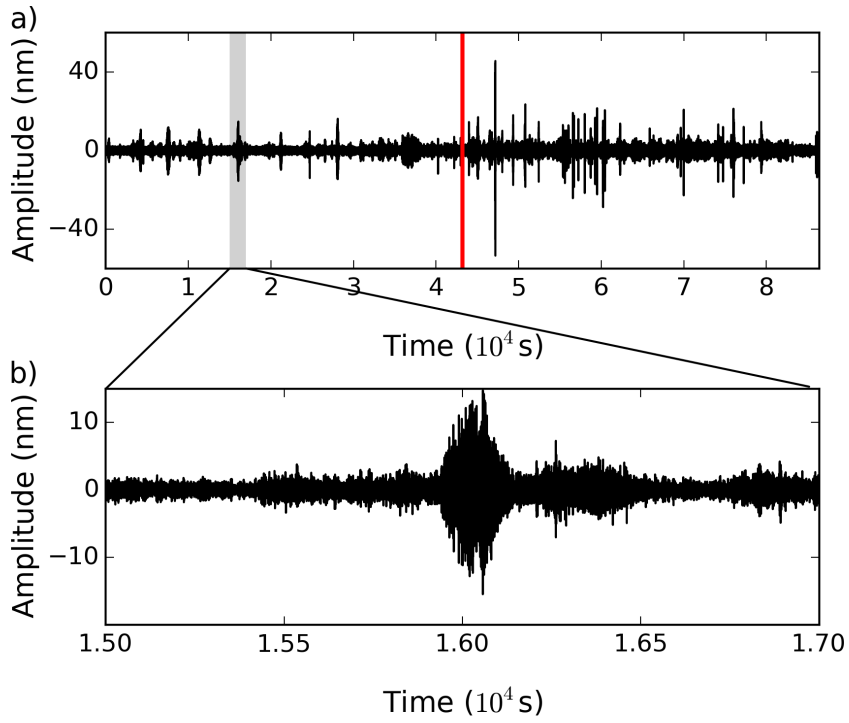

Figure 4.2: (a) One-day seismogram recorded by the vertical component of station Test3. (b) Zoom-in of a single long duration (LD) event. The left and right halves separated by the vertical red line in (a) are records during the nighttime and daytime, respectively.

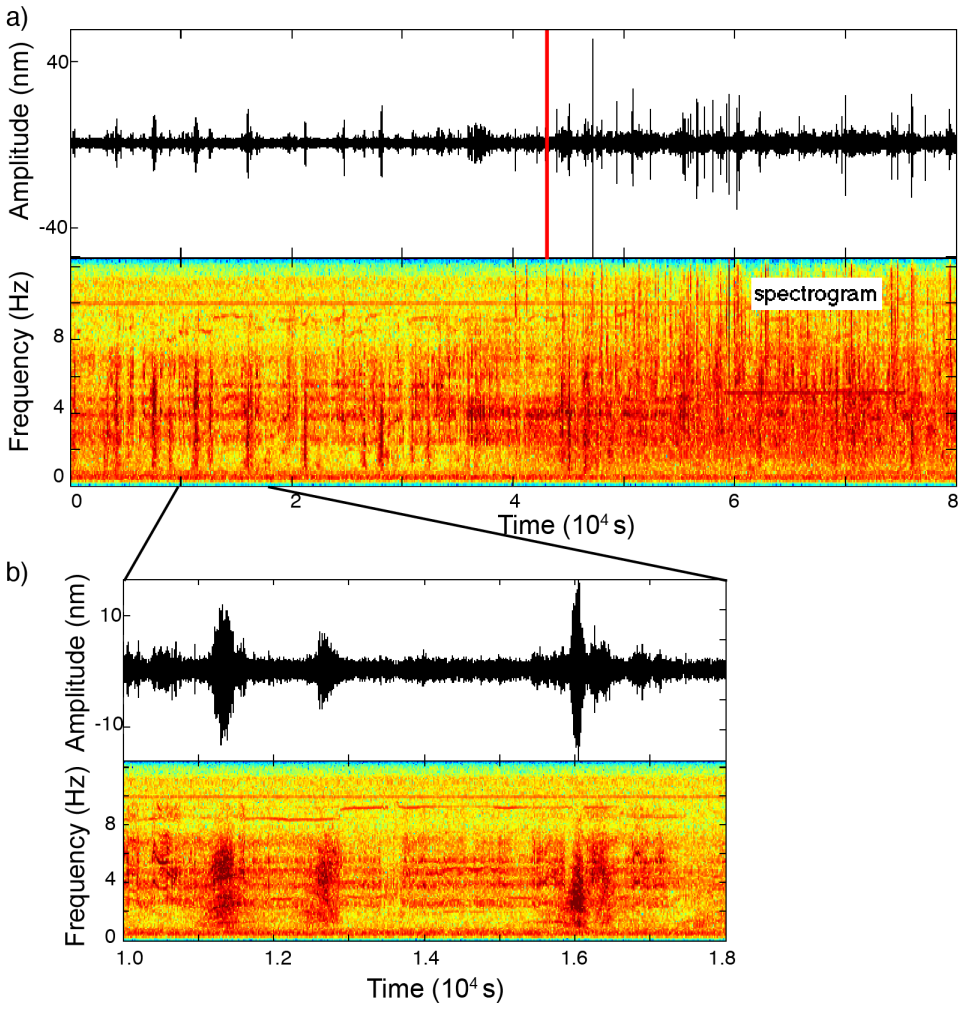

Figure 4.3: (a) A one-day seismogram and its spectrogram recorded by the vertical component of a seismic station. The left and right halves separated by the vertical red line are records during the nighttime and daytime, respectively. (b) LD events and their spectrogram within 18000 seconds.

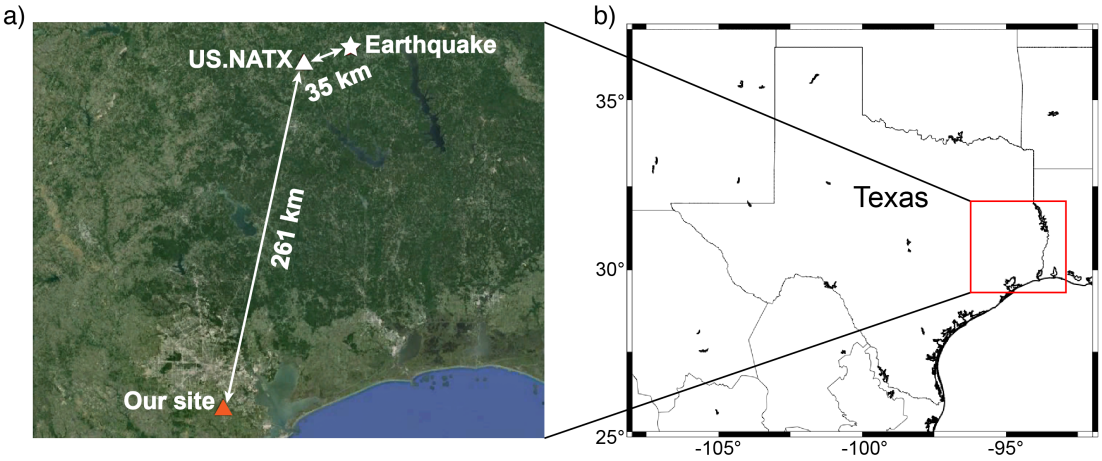

Figure 4.4: (a) Locations of the USArray reference station (white triangle) and reference earthquake (white star). The orange triangle represents our site. The reference station US.NATX is 261 km away from our site and the reference earthquake is about 35 km to the northeast of US.NATX. (b) The location of the area in Texas.

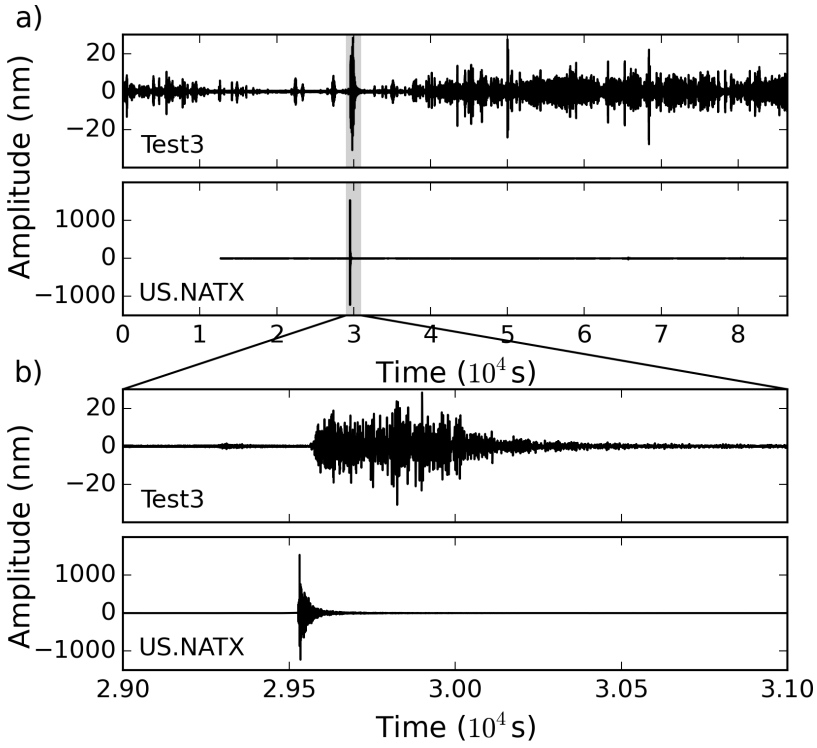

Figure 4.5: (a) Comparison of one-day seismograms recorded by the vertical component of Test3 and the reference station US.NATX. (b) Zoomed seismograms near the arrival time of the reference earthquake.

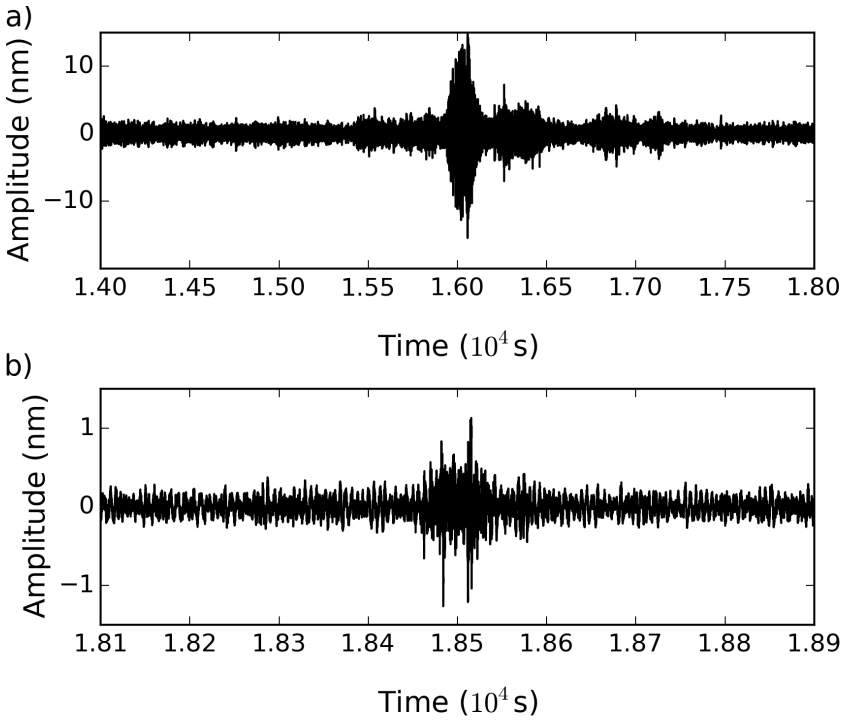

Figure 4.6: Comparison of an LD event (a) and a volcanic tremor at Arenal volcano, Costa Rica (b) recorded by the vertical component of the seismic stations.

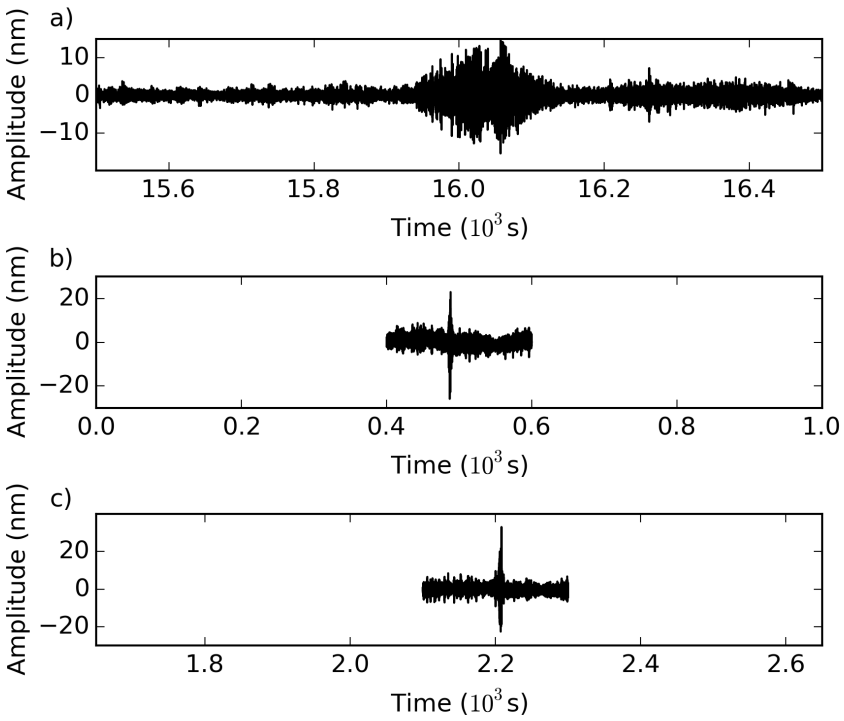

Figure 4.7: Comparison of an LD event (a) and two seismic signals of heavy trucks (b and c) in the same time scale.

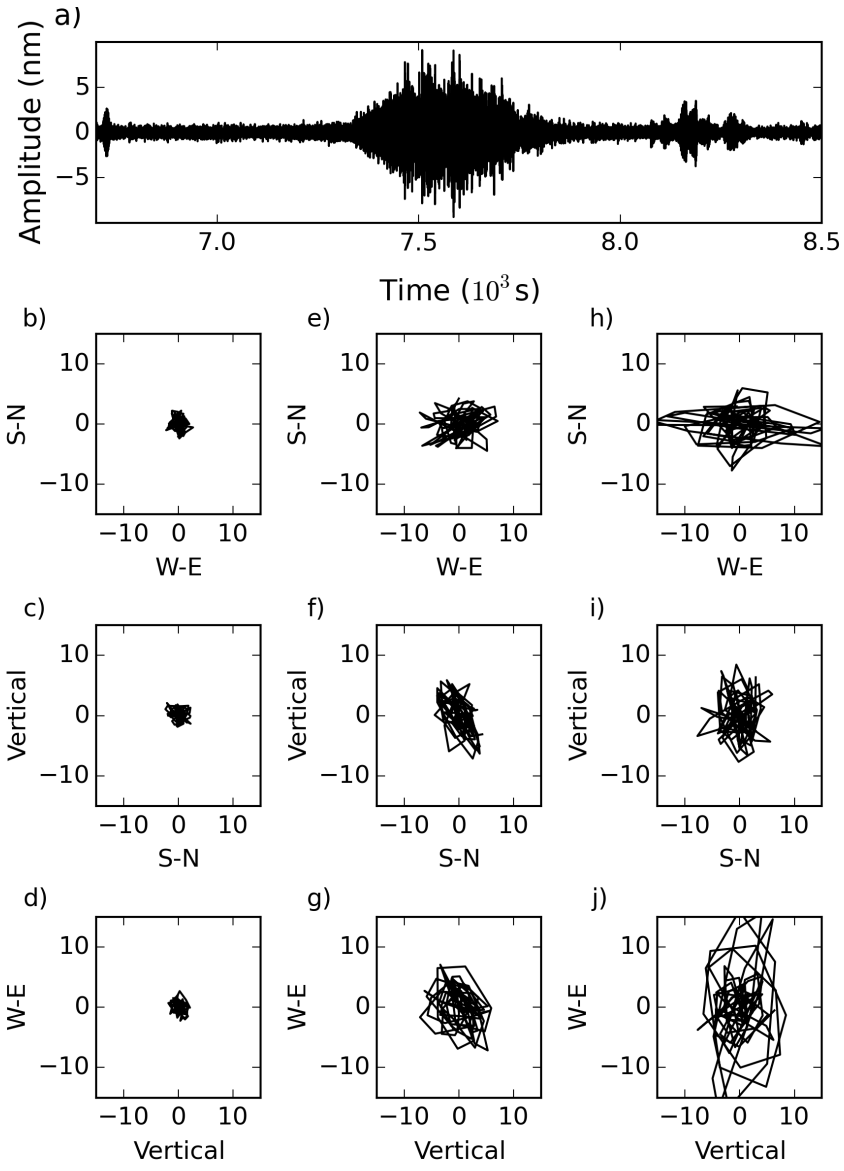

Figure 4.8: Particle motions of an LD event (a) at the beginning part (b, c and d), middling part (e, f and g) and peak (h, i and j). S-N is in the south-north direction and W-E is in the west-east direction.

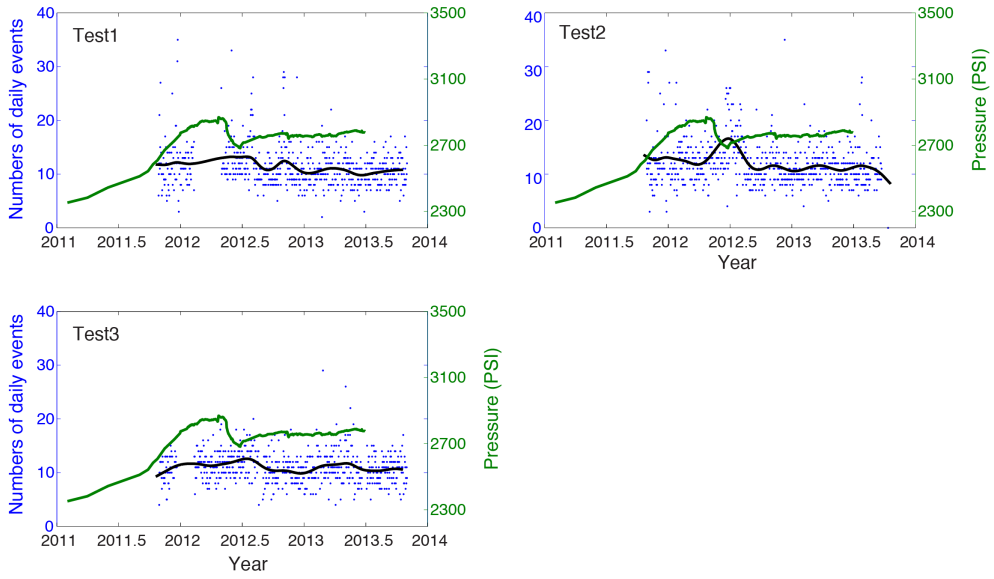

Figure 4.9: Comparison of the reservoir pressure (Green) and the LD event rates at the vertical components of the three stations. The blue dots are the rate of the LD events per night. The solid black lines are the smoothed fitting of the rate from cubic smoothing spline. The location of the pressure well is closer to stations Test1 and Test2.

Chapter 5

Conclusions

5.1 Overview

In this dissertation, we developed new techniques for seismic tomography, which includes (1) adopting adaptive parametrization of velocity model to improve the stability and resolution of ambient noise tomography, and (2) integrating full topography into local earthquake tomography to improve the accuracy of the inverted velocity model. On the other hand, we find a new kind of induced seismic event associated with CO_2 injection at an Enhanced Oil Recovery (EOR) site, which is called Long duration (LD) event for its very long duration of $\sim$30 sec to $\sim$300 sec. The overview of the three works are shown as following.

5.1.1 Adaptive ambient noise tomography

In this work, we present a new adaptive algorithm for ambient noise tomography. The irregular grids constructed by our algorithm match the ray distribution well. The synthetic data tests show that the adaptive tomography algorithm provides better resolution in the area with high ray density owing to more efficient use of the data information than the regular tomography. The fact that the results of the adaptive tomography show no

obvious variations with damping parameters indicates that the adaptive grids have fewer divergent nodes, which are strongly affected by damping parameters in the traditional evenly-spaced grid inversion. When Gaussian distributed noise is added into the input data, areas with sparse rays in the velocity model are always unstable in the traditional method at different noise levels, although the area with dense rays is more stable. The adaptive method stays fairly stable at lower noise levels, although some artifacts also appear when the noise level increases. In addition, the tests of the traditional tomography with different grid sizes show a trade-off between obtaining a good resolution and avoiding inversion artifacts at the same time. When applied to real ambient noise data near the Garlock fault in southern California, the new adaptive tomography algorithm is able to recover the major structures of the group velocity structure. We observe high velocity anomalies in the cold southern Sierra Nevada crust, crystalline basement of Tehachapi Mountains, and brittle roots of the Western San Gabriel Mountains. Low velocity anomalies are seen near the southern San Joaquin Valley and western Mojave.

5.1.2 Local earthquake tomography with full topography

We develop and present a new local earthquake tomography algorithm (referred to as LETFT) with the inclusion of full topography integrated from the digital elevation model (DEM) data. The algorithm naturally avoids possible errors that may be caused by the neglect of local topography and improves the accuracy of the resulting velocity model, especially at shallow depths. We compare our new algorithm with the traditional LET method using both synthetic and real data. Our synthetic data tests show that the new algorithm is able to recover the true velocity well in both the shallow and deep parts of the crust. In contrast, the traditional method does not perform as well as the new one to solve for the shallow velocity structure. The calculated errors in the inverted velocity

model, earthquake locations and arrival time residuals also provide quantitative proof of the improvement of the new method over the traditional one.

The application of the new algorithm to Kīlauea volcano, Hawai‘i, reveals several notable Vp structures. The velocity decreased from 4.3 km/s near the surface to 3.8 km/s at $\sim$2 km depth beneath the summit caldera and could be associated with a state change of olivine basalt. We also observe several relatively low P-wave velocity zones in the ERZ and HFS that may indicate some potential partial melting zones. The low velocity zone beneath the Pu'u'$\bar{O}$'$\bar{o}$ crater at 5-10 km depth is similar to the velocity of olivine basalt or plagioclase basalt from the oceanic crust.

The new LETFT algorithm shows better performance than the traditional method at shallow depths in the synthetic data tests and resolves several interesting P-wave velocity features in Kīlauea area, Hawai'i, where the highest elevation is about 1 km above sea level. We expect it will uncover more robust velocity structures in areas with larger topography variation, such as other volcanic areas or mountain areas.

5.1.3 Induced long duration seismic events

In the two years of seismic monitoring at an EOR field near Houston, TX, we observe a type of LD micro-seismic events with a duration of 30-300 seconds, instead of the typical micro earthquakes which have much shorter durations for comparable amplitudes. The amplitudes of the LD events vary smoothly. The waveforms of the LD events are very different from the typical earthquake and are effect of local features, which were not recorded by the regional seismic stations (e.g., $\sim$200 km away). The waveforms and frequency information of the LD events are similar to volcanic tremors. This indicates that the fluid-involved seismic events such as the LD events and volcanic tremors may share similar mechanisms. After a traffic effect test, we exclude the possibility that

the LD events are caused by the local traffic. With the particle motion information at different parts of the LD event, we find that the source of the LD events may not be an single source and could migrate to a complex fracture area.

We also picked the LD events of two years data and calculate their rate per night. We then compare the results with the reservoir pressure data recorded in the north block of the study area. The LD event rates at Test1 and 2 show relations with the variations of the reservoir pressure. The highest part of LD event rate at Test2 happens in the period of rapid change of the reservoir pressure. The LD event rate of Test3 in the south block does not show a clear pattern, which could be due to the sealing fault between the two blocks preventing the migration of the injected CO_2 from the north block to the south block.

5.2 Future work

The work above is not the end, it is just a small step on the journey of seismology research. I also found some problems that need to be resolved when I worked on the old problems. (1) For adaptive ambient noise tomography, although the results are more reliable and stable, I found it is a little difficult to explain the resolution of the velocity model. Because the node distribution is not homogeneous, the area with dense nodes could resolve smaller scale structures than the area with sparse nodes. It is easier to mislead the readers that the area with sparse nodes has only larger velocity features instead of small velocity features from the velocity map. A question for the future work is how to evaluate the resolution of the results in the future, in order to help people understand the information from the results. (2) In the work of local earthquake tomography with full topography, the remaining problems are how to quantitatively evaluate the effects of topography and finding a better way to parameterize the velocity

model. The difficulty of the former one is that the topography is complex and depends on the area of research. To solve the latter one, we need to find the optimal node structure near the Earth's surface. (3) For induced LD events at EOR field, we still don't quite know the mechanism of the LD events. More information, for instance, the locations of the LDEs, local geological structures, and detailed injection information are needed to better understand these events. Thus data from more stations are needed. Simulation about the fluid induced seismicity also could provide evidence to help solve this problem.

Bibliography

Ake, J., Mahrer, K., O'Connell, D., and Block, L. (2005). Deep-injection and closely monitored induced seismicity at Paradox Valley, Colorado. *Bull. Seismol. Soc. Am.*, 95(2):664–683.

Aki, K., Christoffersson, A., and Husebye, E. S. (1977). Determination of the three-dimensional seismic structure of the lithosphere. *J. Geophys. Res.*, 82(2):277–296.

Aki, K. and Lee, W. (1976). Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model. *J. Geophys. Res.*, 81(23):4381–4399.

Allen, R. (1982). Automatic phase pickers: Their present use and future prospects. *Bull. Seismol. Soc. Am.*, 72(6B):S225–S242.

Aster, R., Borchers, B., and Thurber, C. (2005). *Parameter estimation and inverse problems*. Academic Press, Burlington, MA.

Baisch, S. and Harjes, H.-P. (2003). A model for fluid-injection-induced seismicity at the KTB, Germany. *Geophys. J. Int.*, 152(1):160–170.

Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The Quickhull algorithm for convex hulls. *Trans. Math. Software*, 22(4):469–483.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophys. J. Int.*, 169(3):1239–1260.

Benz, H., Zandt, G., and Oppenheimer, D. (1992). Lithospheric structure of northern California from teleseismic images of the upper mantle. *J. Geophys. Res.*, 97(B4):4791–4807.

Bijwaard, H., Spakman, W., and Engdahl, E. R. (1998). Closing the gap between regional and global travel time tomography. *J. Geophys. Res.*, 103(B12):30055–30078.

Böhm, G., Galuppo, P., and Vesnaver, A. (2000). 3D adaptive tomography using Delaunay triangles and Voronoi polygons. *Geophys. Prospect.*, 48(4):723–744.

Bohnhoff, M. and Zoback, M. D. (2010). Oscillation of fluid-filled cracks triggered by degassing of CO2 due to leakage along wellbores. *J. Geophys. Res.*, 115(B11).

Boschetti, F., Dentith, M. D., and List, R. D. (1996). A fractal-based algorithm for detecting first arrivals on seismic traces. *Geophysics*, 61(4):1095–1102.

Boschi, L. and Dziewonski, A. M. (1999). High-and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling. *J. Geophys. Res.*, 104(B11):25567–25594.

Brenguier, F., Shapiro, N. M., Campillo, M., Nercessian, A., and Ferrazzini, V. (2007). 3-D surface wave tomography of the Piton de la Fournaise volcano using seismic noise correlations. *Geophys. Res. Lett.*, 34.

Busch, A., Alles, S., Gensterblum, Y., Prinz, D., Dewhurst, D. N., Raven, M. D., Stanjek, H., and Krooss, B. M. (2008). Carbon dioxide storage potential of shales. *Int. J. Greenh. Gas Control*, 2(3):297–308.

Christensen, N. I. (1996). Poisson's ratio and crustal seismology. *J. Geophys. Res.*, 101(B2):3139–3156.

Crosson, R. S. (1976). Crustal structure modeling of earthquake data: 1. Simultaneous least squares estimation of hypocenter and velocity parameters. *J. Geophys. Res.*, 81(17):3036–3046.

Das, I. and Zoback, M. D. (2013). Long-period, long-duration seismic events during hydraulic stimulation of shale and tight-gas reservoirs -Part 1: Waveform characteristics. *Geophysics*, 78(6):KS97–KS108.

Davis, G. A. and Burchfiel, B. C. (1973). Garlock fault: An intracontinental transform structure, southern California. *Geol. Soc. Am. Bull.*, 84(4):1407–1422.

Davis, S. D. and Pennington, W. D. (1989). Induced seismic deformation in the Cogdell oil field of west Texas. *Bull. Seismol. Soc. Am.*, 79(5):1477–1495.

Dawson, P., Chouet, B., Okubo, P., Villaseñor, A., and Benz, H. (1999). Three-dimensional velocity structure of the Kilauea caldera, Hawaii. *Geophys. Res. Lett.*, 26(18):2805–2808.

De Berg, M., Van Kreveld, M., Overmars, M., and Schwarzkopf, O. (2000). *Computational geometry: algorithms and applications*. Springer, New York, 2nd edition.

Dewhurst, D. N. and Hennig, A. L. (2003). Geomechanical properties related to top seal leakage in the Carnarvon Basin, Northwest Shelf, Australia. *Petrol. Geosci.*, 9(3):255–263.

Dixon, T. H., Miller, M., Farina, F., Wang, H., and Johnson, D. (2000). Present-day motion of the Sierra Nevada block and some tectonic implications for the Basin and Range province, North American Cordillera. *Tectonics*, 19(1):1–24.

Dokka, R. K. and Travis, C. J. (1990). Late Cenozoic strike-slip faulting in the Mojave Desert, California. *Tectonics*, 9(2):311–340.

Dolan, J. F., Bowman, D. D., and Sammis, C. G. (2007). Long-range and long-term fault interactions in Southern California. *Geology*, 35(9):855–858.

Eaton, D., van der Baan, M., Tary, J.-B., Birkelo, B., Spriggs, N., Cutten, S., and Pike, K. (2013). Broadband microseismic observations from a Montney hydraulic fracture treatment, northeastern BC, Canada. *Focus*.

Eberhart-Phillips, D. (1986). Three-dimensional velocity structure in northern California Coast Ranges from inversion of local earthquake arrival times. *Bull. Seismol. Soc. Am.*, 76(4):1025–1052.

Eberhart-Phillips, D. (1990). Three-dimensional P and S velocity structure in the Coalinga region, California. *J. Geophys. Res.*, 95(B10):15343–15363.

Eisner, L., Hulsey, B., Duncan, P., Jurick, D., Werner, H., and Keller, W. (2010). Comparison of surface and borehole locations of induced seismicity. *Geophys. Prospect.*, 58(5):809–820.

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

Evans, J. R., Eberhart-Phillips, D., and Thurber, C. (1994). User's manual for SIMULPS12 for imaging Vp and Vp/Vs: A derivative of the " Thurber" tomographic inversion SIMUL3 for local earthquakes and explosions. Open-File Report 94-431, US Geological Survey.

Fortune, S. (1992). *Voronoi diagrams and Delaunay triangulations*. World Scientific, Singapore.

Foulger, G. and Toomey, D. (1989). Tomographic inversion of local earthquake data from the Hengill-Grensdalur central volcano complex, Iceland. *J. Geophys. Res.*, 94(B12):17–497.

Friedrich, A., Krüger, F., and Klinge, K. (1998). Ocean-generated microseismic noise located with the Gräfenberg array. *J. Seismol.*, 2(1):47–64.

Garfunkel, Z. (1974). Model for the late Cenozoic tectonic history of the Mojave Desert, California, and for its relation to adjacent regions. *Geol. Soc. Am. Bull.*, 85(12):1931–1944.

Gentile, G., Bressan, G., Burlini, L., and De Franco, R. (2000). Three-dimensional Vp and Vp/Vs models of the upper crust in the Friuli area (northeastern Italy). *Geophys. J. Int.*, 141(2):457–478.

Glazner, A. F., Walker, J. D., Bartley, J. M., and Fletcher, J. M. (2002). Cenozoic evolution of the Mojave block of southern California. In Glazner, A. F., Walker, J. D., and Bartley, J. M., editors, *Geologic evolution of the Mojave Desert and southwestern Basin and Range*, volume 195, pages 19–41. Geological Society of America.

Golub, G. H. and Reinsch, C. (1970). Singular value decomposition and least squares solutions. *Numer. Math.*, 14(5):403–420.

Goodman, E. D. and Malin, P. E. (1992). Evolution of the Southern San Joaquin Basin and mid-Tertiary "transitional" tectonics, central California. *Tectonics*, 11(3):478–498.

Hamilton, W. (1969). Mesozoic California and the underflow of Pacific mantle. *Geol. Soc. Am. Bull.*, 80(12):2409–2430.

Hansen, S., Thurber, C., Mandernach, M., Haslinger, F., and Doran, C. (2004). Seismic velocity and attenuation structure of the east rift zone and south flank of Kilauea Volcano, Hawaii. *Bull. Seismol. Soc. Am.*, 94(4):1430–1440.

Haslinger, F., Thurber, C., Mandernach, M., and Okubo, P. (2001). Tomographic image of P-velocity structure beneath Kilauea's east rift zone and south flank: Seismic evidence for a deep magma body. *Geophys. Res. Lett.*, 28(2):375–378.

Humphreys, E. and Clayton, R. W. (1988). Adaptation of back projection tomography to seismic travel time problems. *J. Geophys. Res.*, 93(B2):1073–1085.

Ingersoll, R. (1982). Initiation and evolution of the Great Valley forearc basin of northern and central California, USA. In *Geological Society, London, Special Publications*, volume 10, pages 459–467. Geological Society of London.

Karastathis, V., Papoulia, J., Di Fiore, B., Makris, J., Tsambas, A., Stampolidis, A., and Papadopoulos, G. (2011). Deep structure investigations of the geothermal field of the North Euboean Gulf, Greece, using 3-D local earthquake tomography and Curie Point Depth analysis. *J. Volcanol. Geotherm. Res.*, 206(3):106–120.

Karegar, M. A., Dixon, T. H., Malservisi, R., Yang, Q., Hossaini, S. A., and Hovorka, S. D. (2015). GPS-based monitoring of surface deformation associated with CO2 injection at an enhanced oil recovery site. *Int. J. Greenh. Gas Control*, 41:116–126.

Kauahikaua, J. (1993). Geophysical characteristics of the hydrothermal systems of Kilauea volcano, Hawaii. *Geothermics*, 22(4):271–299.

Keranen, K. M., Savage, H. M., Abers, G. A., and Cochran, E. S. (2013). Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology*, 41(6):699–702.

Khatiwada, M., Adam, L., Morrison, M., and van Wijk, K. (2012). A feasibility study of time-lapse seismic monitoring of CO2 sequestration in a layered basalt reservoir. *J. Appl. Geophys.*, 82:145–152.

Kilkenny, J. E. (1951). San Joaquin Valley. *AAPG Bull.*, 35(2):215–218.

Kinoshita, W. (1965). A gravity survey of the island of Hawai'i. *Pac. Sci.*, 19(3):349.

Kissling, E., Husen, S., and Haslinger, F. (2001). Model parametrization in seismic tomography: a choice of consequence for the solution quality. *Phys. Earth Planet. In.*, 123(2):89–101.

Klein, F., Koyanagi, R., Nakata, J., and Tanigawa, W. (1987). The seismicity of Kilauea's magma system. *US Geol. Surv. Prof. Pap*, 1350:1019–1185.

Larose, E., Derode, A., Clorennec, D., Margerin, L., and Campillo, M. (2005). Passive retrieval of Rayleigh waves in disordered elastic media. *Physical Review E*, 72(4):046607.

Lay, T. and Wallace, T. C. (1995). *Modern global seismology*, volume 58. Academic press.

Lei, X., Yu, G., Ma, S., Wen, X., and Wang, Q. (2008). Earthquakes induced by water injection at 3 km depth within the Rongchang gas field, Chongqing, China. *J. Geophys. Res.*, 113(B10).

Levshin, A., Ratnikova, L., and Berger, J. (1992). Peculiarities of surface-wave propagation across central Eurasia. *Bull. Seismol. Soc. Am.*, 82(6):2464–2493.

Li, C. and Nowack, R. L. (2004). Application of autoregressive extrapolation to seismic tomography. *Bull. Seismol. Soc. Am.*, 94(4):1456–1466.

Li, P. and Lin, G. (2014). Adaptive ambient noise tomography and its application to the Garlock fault, southern California. *Geophys. J. Int.*, page ggu073.

Liang, C. and Langston, C. A. (2008). Ambient seismic noise tomography and structure of eastern North America. *J. Geophys. Res.*, 113.

Lin, F. C., Moschetti, M. P., and Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophys. J. Int.*, 173(1):281–298.

Lin, F. C., Ritzwoller, M. H., Townend, J., Bannister, S., and Savage, M. K. (2007a). Ambient noise Rayleigh wave tomography of New Zealand. *Geophys. J. Int.*, 170(2):649–666.

Lin, G. (2015). Seismic velocity structure and earthquake relocation for the magmatic system beneath Long Valley Caldera, eastern California. *J. Volcanol. Geotherm. Res.*, 296:19–30.

Lin, G., Shearer, P. M., Hauksson, E., and Thurber, C. H. (2007b). A three-dimensional crustal seismic velocity model for southern California from a composite event method. *J. Geophys. Res.*, 112(B11).

Lin, G., Shearer, P. M., Matoza, R. S., Okubo, P. G., and Amelung, F. (2014). Three-dimensional seismic velocity structure of Mauna Loa and Kilauea volcanoes in Hawaii from local seismic tomography. *J. Geophys. Res.*, 119(5):4377–4392.

Lobkis, O. I. and Weaver, R. L. (2001). On the emergence of the Greens function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6):3011–3017.

Lutter, W. J., Fuis, G. S., Ryberg, T., Okaya, D. A., Clayton, R. W., Davis, P. M., Prodehl, C., Murphy, J. M., Langenheim, V. E., Benthien, M. L., Godfrey, N. J., Christensen, N. I., Thygesen, K., Thurber, C. H., Simila, G., and Keller, G. R. (2004). Upper crustal structure from the Santa Monica mountains to the Sierra Nevada, southern California: tomographic results from the Los Angeles Regional Seismic Experiment, Phase II (LARSE II). *Bull. Seismol. Soc. Am.*, 94(2):619–632.

Martínez-Garzón, P., Bohnhoff, M., Kwiatek, G., Zambrano-Narváez, G., and Chalaturnyk, R. (2013). Microseismic monitoring of CO2 Injection at the Penn West Enhanced Oil Recovery pilot project, Canada: Implications for detection of wellbore leakage. *Sensors*, 13(9):11522–11538.

Matsubara, M., Obara, K., and Kasahara, K. (2008). Three-dimensional P-and S-wave velocity structures beneath the Japan Islands obtained by high-density seismic stations by seismic tomography. *Tectonophysics*, 454(1):86–103.

Mazzoldi, A., Rinaldi, A. P., Borgia, A., and Rutqvist, J. (2012). Induced seismicity within geological carbon sequestration projects: maximum earthquake magnitude and leakage potential from undetected faults. *Int. J. Greenh. Gas Control*, 10:434–442.

McGarr, A., Bekins, B., Burkardt, N., Dewey, J., Earle, P., Ellsworth, W., Ge, S., Hickman, S., Holland, A., Majer, E., et al. (2015). Coping with earthquakes induced by fluid injection. *Science*, 347(6224):830–831.

McGill, S. and Sieh, K. (1993). Holocene slip rate of the central Garlock fault in southeastern Searles Valley, California. *J. Geophys. Res.: Solid Earth*, 98(B8):14217–14231.

McNutt, S. R. (2011). Earth science: Volcanic tremor wags on. *Nature*, 470(7335):471–472.

Meade, B. J. and Hager, B. H. (2005). Block models of crustal motion in southern California constrained by GPS measurements. *J. Geophys. Res.: Solid Earth*, 110(B3).

Michelini, A. (1993). Velocity model inversion using parametric curves. *Geophys. J. Int.*, 115(2):337–343.

Michelini, A. (1995). An adaptive-grid formalism for traveltime tomography. *Geophys. J. Int.*, 121(2):489–510.

Monteiller, V., Got, J.-L., Virieux, J., and Okubo, P. (2005). An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to the Kilauea volcano. *J. Geophys. Res.*, 110(B12).

Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. R., and Hung, S.-H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303(5656):338–343.

Morris, J. P., Hao, Y., Foxall, W., and McNab, W. (2011). A study of injection-induced mechanical deformation at the In Salah CO2 storage project. *Int. J. Greenh. Gas Control*, 5(2):270–280.

Moschetti, M. P., Ritzwoller, M. H., and Shapiro, N. M. (2007). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps. *Geochem. Geophys. Geosyst.*, 8(8).

Okubo, P. G., Benz, H. M., and Chouet, B. A. (1997). Imaging the crustal magma sources beneath Mauna Loa and Kilauea volcanoes, Hawaii. *Geology*, 25(10):867–870.

Park, J., Morgan, J. K., Zelt, C. A., and Okubo, P. G. (2009). Volcano-tectonic implications of 3-D velocity structures derived from joint active and passive source tomography of the island of Hawaii. *J. Geophys. Res.*, 114(B9).

Pavlis, G. L. and Booker, J. R. (1980). The mixed discrete-continuous inverse problem: Application to the simultaneous determination of earthquake hypocenters and velocity structure. *J. Geophys. Res.*, 85(B9):4801–4810.

Petersen, M. D. and Wesnousky, S. G. (1994). Fault slip rates and earthquake histories for active faults in southern California. *Bull. Seismol. Soc. Am.*, 84(5):1608–1649.

Rawlinson, N. and Sambridge, M. (2004). Multiple reflection and transmission phases in complex layered media using a multistage fast marching method. *Geophysics*, 69(5):1338–1350.

Rhie, J. and Romanowicz, B. (2004). Excitation of Earth's continuous free oscillations by atmosphere–ocean–seafloor coupling. *Nature*, 431(7008):552–556.

Ritter, J. R., Jordan, M., Christensen, U. R., and Achauer, U. (2001). A mantle plume below the Eifel volcanic fields, Germany. *Earth Planet. Sci. Lett.*, 186(1):7–14.

Ritzwoller, M. H. and Levshin, A. L. (1998). Eurasian surface wave tomography: Group velocities. *J. Geophys. Res.: Solid Earth*, 103(B3):4839–4878.

Roult, G. and Crawford, W. (2000). Analysis of backgroundfree oscillations and how to improve resolution by subtracting the atmospheric pressure signal. *Phys. Earth Planet. In.*, 121(3):325–338.

Rowan, L. R. and Clayton, R. W. (1993). The three-dimensional structure of Kilauea Volcano, Hawaii, from travel time tomography. *J. Geophys. Res.*, 98(B3):4355–4375.

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

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., and Fehler, M. (2005). Surface wave tomography from microseisms in Southern California. *Geophys. Res. Lett.*, 32.

Sambridge, M., Braun, J., and McQueen, H. (1995). Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophys. J. Int.*, 122(3):837–857.

Sambridge, M. and Faletič, R. (2003). Adaptive whole Earth tomography. *Geochem. Geophys. Geosyst.*, 4(3).

Sambridge, M. and Gudmundsson, Ó. (1998). Tomographic systems of equations with irregular cells. *J. Geophys. Res.*, 103(B1):773–781.

Saygin, E. and Kennett, B. L. N. (2010). Ambient seismic noise tomography of Australian continent. *Tectonophysics*, 481(1):116–125.

Schreiber, E. and Fox, P. J. (1977). Density and P-wave velocity of rocks from the FAMOUS region and their implication to the structure of the oceanic crust. *Geol. Soc. Am. Bull.*, 88(4):600–608.

Sethian, J. A. and Popovici, A. M. (1999). 3-D traveltime computation using the fast marching method. *Geophysics*, 64(2):516–523.

Shapiro, N. M. and Campillo, M. (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise. *Geophysical Research Letters*, 31(7).

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715):1615–1618.

Shapiro, S., Dinske, C., and Kummerow, J. (2007). Probability of a given-magnitude earthquake induced by a fluid injection. *Geophys. Res. Lett.*, 34(22).

Sharma, N., Convertito, V., Maercklin, N., and Zollo, A. (2013). Ground-motion prediction equations for the Geysers geothermal area based on induced seismicity records. *Bull. Seismol. Soc. Am.*, 103(1):117–130.

Snieder, R. (2004). Extracting the Greens function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4):046610.

Spakman, W. and Bijwaard, H. (2001). Optimization of cell parameterizations for tomographic inverse problems. *Pure Appl. Geophys.*, 158(8):1401–1423.

Steck, L. K., Thurber, C. H., Fehler, M. C., Lutter, W. J., Roberts, P. M., Baldridge, W. S., Stafford, D. G., and Sessions, R. (1998). Crust and upper mantle P wave velocity structure beneath Valles caldera, New Mexico: Results from the Jemez teleseismic tomography experiment. *J. Geophys. Res.*, 103(B10):24301–24320.

Stehly, L., Campillo, M., and Shapiro, N. (2006). A study of the seismic noise from its long-range correlation properties. *J. Geophys. Res.*, 111(B10).

Talebi, S. and Boone, T. J. (1998). Source parameters of injection-induced microseismicity. *Pure Appl. Geophys.*, 153(1):113–130.

Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadephia, PA.

Thurber, C. H. (1983). Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California. *J. Geophys. Res.*, 88(B10):8226–8236.

Thurber, C. H. (1984). Seismic detection of the summit magma complex of Kilauea volcano, Hawaii. *Science*, 223(4632):165–167.

Thurber, C. H. (1992). Hypocenter-velocity structure coupling in local earthquake tomography. *Phys. Earth. Planet. In.*, 75(1):55–62.

Vasco, D., Rucci, A., Ferretti, A., Novali, F., Bissell, R., Ringrose, P., Mathieson, A., and Wright, I. (2010). Satellite-based measurements of surface deformation reveal fluid flow associated with the geological storage of carbon dioxide. *Geophys. Res. Lett.*, 37(3).

Verdon, J. P. (2014). Significance for secure CO2 storage of earthquakes induced by fluid injection. *Environ. Res. Lett.*, 9(6):064022.

Vesnaver, A. L. (1996). Irregular grids in seismic tomography and minimum-time ray tracing. *Geophys. J. Int.*, 126(1):147–165.

Webb, S. C. (2007). The Earths humis driven by ocean waves over the continental shelves. *Nature*, 445(7129):754–756.

Wright, L. (1976). Late Cenozoic fault patterns and stress fields in the Great Basin and westward displacement of the Sierra Nevada block. *Geology*, 4(8):489–494.

Yan, Z. and Clayton, R. W. (2007). Regional mapping of the crustal structure in southern California from receiver functions. *J. Geophys. Res.: Solid Earth*, 112(B5).

Yang, Q., Zhao, W., Dixon, T. H., Amelung, F., Han, W. S., and Li, P. (2015). InSAR monitoring of ground deformation due to CO2 injection at an enhanced oil recovery site, West Texas. *Int. J. Greenh. Gas Control*, 41:20–28.

Yang, Y., Ritzwoller, M., Levshin, A., and Shapiro, N. (2007). Ambient noise Rayleigh wave tomography across Europe. *Geophys. J. Int.*, 168(1):259–274.

Yang, Y. and Ritzwoller, M. H. (2008). Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochem. Geophys. Geosyst.*, 9(2).

Zhang, H. and Thurber, C. H. (2003). Double-difference tomography: The method and its application to the Hayward fault, California. *Bull. Seismol. Soc. Am.*, 93(5):1875–1889.

Zhang, H. and Thurber, C. H. (2005). Adaptive mesh seismic tomography based on tetrahedral and Voronoi diagrams: application to Parkfield, California. *J. Geophys. Res.*, 110:B04303.

Zhao, D. and Kanamori, H. (1992). P-wave image of the crust and uppermost mantle in southern California. *Geophys. Res. Lett.*, 19(23):2329–2332.

Zheng, S., Sun, X., Song, X., Yang, Y., and Ritzwoller, M. H. (2008). Surface wave tomography of China from ambient seismic noise correlation. *Geochem. Geophys. Geosyst.*, 9(5).

Zhou, L., Xie, J., Shen, W., Zheng, Y., Yang, Y., Shi, H., and Ritzwoller, M. H. (2012). The structure of the crust and uppermost mantle beneath South China from ambient noise and earthquake tomography. *Geophys. J. Int.*, 189(3):1565–1583.

Zoback, M. D. and Gorelick, S. M. (2012). Earthquake triggering and large-scale geologic storage of carbon dioxide. *Proc. Natl. Acad. Sci.*, 109(26):10164–10168.