Application of State of the Art Seismic Techniques to California

By Bin Guo

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

At the
UNIVERSITY OF WISCONSIN-MADISON
2018

Date of the final oral examination: 12/21/2018

The dissertation is approved by the following members of the Final Oral Committee:

Clifford Thurber, Professor, Geoscience

Chuck DeMets, Professor, Geoscience

Lucas Zoet, Professor, Geoscience

Dante Fratta, Professor, Geological Engineering

Acknowledgements

I am sincerely grateful to my advisor Cliff Thurber for the mentorship he has provided me during my graduate school life at the University of Wisconsin-Madison. It was tough when I first landed in the United States as a foreigner who had never left China. As my M.S. and PhD advisor, Cliff has been enlightening me with his extraordinary expertise and broad view in seismology. Beyond the intellectual guidance for my research studies, Cliff supported me with patience and left me with space when I needed during the down times in my PhD life. Without Cliff, I cannot achieve what I have become, a better writer, researcher, and most importantly a more independent human being. Cliff is for sure my dream advisor.

I want to acknowledge all of my committee members, Cliff Thurber, Chuck Demets, Dante Fratta, Harold Tobin, and Lucas Zoet, for their kind support. I appreciate the help from Laurel Goodwin, Kurt Feigl, Michael Cardiff, Herb Wang, and Stephen Meyers during every class and seminar. One of the best memories I have is the field trip to Chile and Argentina with Brad Singer, Shanan Peters and, all other fellow students.

I would like to thank my coworkers, Steve Roecker, John Townend, Chris Rawles, Calum Chamberlain, Caroline Boese, Stephen Bannister, Jessica Feenstra, and Jennifer Eccles, for my Alpine Fault tomography study. The support and cooperation from David Shelly, Donna Eberhart-Phillips, and Haijiang Zhang are also acknowledged. It is enjoyable and beneficial to work with amazing researchers, Avinash Nayak, Ninfa Bennington, Federica Lanza, and Xiangfang Zeng.

I am thankful for all the help from the Department of Geoscience staff. Ben Abernathy and Patrick Kuhl have provided amazing IT support. Micheal Szabo, Jane Fox-Anderson, Shirley Baxa, Mary Schumann, and Judy Gosse made my time in grad school easier and smoother.

The funding support of a student assistantship from the Mark and Carol Ann Solien Graduate Assistantship and George P. Woollard-Sigmund I. Hammer Memorial Fund in Geology & Geophysics of the University of Wisconsin-Madison are also acknowledged.

Madison is the first city I visited and lived in the United States. During the time here, I have made friends, experienced different lifestyles, and been through the most critical part of my life. It is beyond words to express my feelings for the city and the school. This fantastic journey started when Cliff kindly offered me the life-changing opportunity five years ago. I have not and will never regret accepting the offer.

My parents and little sister are always there cheering for me. They provided me with the courage to face and concur the difficulties in life. I love you.

1. Introduction

Utilizing state of art geophysics techniques, I conduct studies focusing on the seismicity and seismic tomography in California. Four different topics are included in this thesis: 1) Non-volcanic tremor relocation error study in Parkfield, California; 2) Deep neural network seismic arrival time auto-picking algorithm; 3) Tomographic study in the San Francisco Bay Area, Northern California; 4) Double-difference attenuation tomography.

Debate about the tremor locations in central California, either scattered or clustered along the San Andreas Fault, has lasted for almost a decade. To investigate the possible reasons for this tremor location discrepancy, synthetic tests are conducted to test the stability of the station-pair location method, which was previously used to locate tremor events in central California. Based on synthetic relocation tests results, it is shown that the station-pair location method suffers from a local minimum problem. The station-pair relocation results show strong initial location dependency causing the estimate to become stuck in local minima near the starting location. Such initial location dependency causes relocation errors at a scale comparable to the scatter related to the long-lasting tremor location discrepancy.

In the second project, I develop a seismic arrival-time auto-picking algorithm based on a deep neural network technique adapted from computer vision studies. Application of traditional auto-pickers to different seismic datasets requires substantial effort in feature engineering (parameter tuning). With the deep neural network method, I develop an auto-picking algorithm that can be applied to general seismic data and extract useful features automatically. The trained neural network models can be easily maintained and updated. The training dataset is from a study along the Alpine fault, New Zealand, and its success of application to a dataset from Parkfield shows the potential portability of this neural network based auto-picking algorithm.

A three-dimensional P- and S-wave tomography study in the San Francisco Bay Area composes the third project in this PhD thesis. In this study, I am able to increase the number of P- and S-wave arrival times via seismic auto-picking algorithms. With the more comprehensive body-wave data and newly available surface-wave group velocities, P- and S-wave velocity models are jointly inverted. The P-wave model shows improved resolution for depths shallower than 4 km. It is the first time that a reliable S-wave model is obtained for this area. The newly inverted P- and S-wave velocity models provide additional insight for structural and seismic studies regarding the complex fault systems in this area.

Inspired by the double-difference seismic velocity tomography method, the last chapter focuses on developing and testing a double-difference attenuation tomography algorithm, which integrates absolute t^* and event-pair differential t^* data. The new double-difference attenuation tomography algorithm is expected to have improved source region resolution, similar to the double-difference seismic velocity tomography method. To derive more reliable event-pair δt^* values, an event-pair spectral ratio method is introduced. Via a series of synthetic tests, the event-pair spectral ratio method is shown to result in better performance for both corner frequencies and δt^* values compared to the traditional method that solves for t^* and corner frequency for each event separately. We conduct synthetic tests with The Geysers dataset as an initial test.

2. Non-volcanic tremor relocation error study in Parkfield, California

2.1. Abstract

The debate regarding the location pattern discrepancy between non-volcanic tremor (NVT) and low frequency earthquakes (LFEs) along the central San Andreas Fault (SAF) has lasted about a decade. The underlying question is if the very scattered locations in NVT studies along the SAF (Nadeau and Dolenc, 2005; Nadeau and Guilhem, 2009; Zhang et al., 2010; Guo et al., 2017) are real. Compared to these "cloudy" looking NVT clusters, the LFE locations in Shelly and Hardebeck (2010) are tightly distributed along the SAF. Hypotheses for this location discrepancy mainly fall into two categories: 1) there is a distinct difference in the sources of NVT versus LFEs; 2) the scattered NVT locations are caused by artifacts in the location method, which uses envelope correlation. If the second hypothesis is correct, the NVT and LFEs may be fundamentally the same phenomenon. In this study, I carried out a detailed study to assess the robustness of the NVT envelope cross-correlation procedures used by Nadeau and Guilhem (2009) and the station-pair double-difference (SPDD) location method of Zhang et al. (2010). The SPDD method relies only on station-pair differential times from NVT envelope cross-correlation. From our tests, we find that the SPDD result depends heavily on the initial locations and can become stuck in a local minimum. This may lead to the consistent scattering of NVT locations in multiple studies, in contrast to the tight clustering of LFEs.

2.2. Introduction

NVT, also called tectonic tremor, was first discovered in the Nankai Trough subduction zone in Japan (Obara, 2002). After the 1995 M 6.5 Kobe earthquake, the National Research Institute of Earth Science and Disaster Protection (NIED) built the nationwide high-sensitivity borehole seismic station network (Hi-net) (Obara, 2002). Hi-net enabled dense seismic station coverage in Japan for earthquake monitoring. In addition, the broadband instruments provided more low-frequency seismic signal information. Obara (2002) first found NVT as a noisy-looking waveform that appeared to be present at multiple stations at similar times with similar waveforms. After processing the non-impulsive seismic signals, Obara (2002) located their epicenters using relative arrival times of waveform envelopes. He confirmed that these noisy-looking signals were a new kind of seismic event different from regular earthquakes, and he termed it non-volcanic tremor. NVTs were located in the Southwest Japan subduction zone and distributed along the strike of Philippine Sea Plate at an inferred depth of about 30 to 40 km (Obara, 2002).

NVT differs from regular earthquakes because of some unique features. 1) NVT has no clear P- and S-wave arrivals; 2) the amplitude of the NVT waveform is usually very small and indistinct from background noise; 3) the frequency spectrum of NVT concentrates in the 1 – 10 Hz range, which is lower than regular earthquakes; 4) NVT can last days to weeks, rather than only seconds to minutes for regular earthquakes.

Low-frequency earthquakes (LFEs) were first reported in Japan, also benefiting from the construction of the Hi-net seismic array (Nishide et al., 2000; Katsumata and Kamaya, 2003). The Japan Meteorological Agency (JMA) discovered a type of seismic signal across many stations, which had no clear P-wave arrivals but detectable S-wave arrivals. JMA located these events with only the S-wave arrivals and found that most of them were located in the Tokai and Kii regions of the Nankai Trough at depths ranging from 25 to 35 km. The LFEs have some similar features to NVT. Signals with frequency lower than 10 Hz dominate in the LFE waveforms.

2.3. Previous studies of NVT locations beneath the SAF

Different from regular earthquakes, NVT signals usually last hours to days without clear P- and S-wave arrival times. This characteristic of NVT makes it impossible to locate them using traditional methods. Instead of using individual seismic phase arrivals, Nadeau and Guilhem (2009) used the root mean square (RMS) envelope of the NVT signal to obtain relative arrival times at pairs of stations from cross-correlation. Then they used a grid search location program to locate these NVT events with those relative arrival times (Nadeau and Guilhem, 2009). Based on their location results, there are two zones in which NVT was concentrated (Figure 2.1), one beneath the locked Cholame segment of the SAF with ~90% of the located NVTs, the other about 65 km to the northwest of Parkfield along the creeping segment of the SAF. All NVTs were distributed in depth from ~15 to 30 km. Furthermore, NVT scatters in a relative large range both horizontally and vertically. In other words, NVTs were not restricted to the presumably near-vertical SAF plane.

Zhang et al. (2010) showed slightly more clustered and deeper NVT locations in the central SAF region using the SPDD location method. The advantage of the SPDD location method is that by differencing travel time residuals for each NVT event at station pairs, the origin time of the NVT event is canceled. Therefore, the NVT locations can be estimated without solving for the origin time. Compared to the initial catalog NVT locations (Nadeau and Guilhem, 2009), the depths for station-pair located NVTs are deeper, ranging between 20 and 30 km (Figure 2.1), which is separated from the seismogenic zone that extends to 15 km depth in the crust (Thurber et al., 2006; Zhang et al., 2010). However, the NVT locations still appeared scattered in both the horizontal and vertical directions.

Shelly et al. (2006) hypothesizes that NVT is composed of LFE swarms based on similar location distributions and correlated occurrence time of LFEs and NVT in Shikoku, Japan. Shelly et al. (2010) identified LFEs within NVT signals and used the LFE locations to represent the NVT locations along the central SAF. LFEs from mid-2001 to early 2010 were cross-correlated to generate template waveforms of 88 LFE families on the Parkfield High Resolution Seismic Network (HRSN) borehole stations. After picking P- and S-wave arrivals from stacked LFEs, they used a 3D S-wave velocity model based on Thurber et al. (2006) and a grid search location method to locate all 88 LFE families. The LFEs locations were distributed tightly beneath the SAF with depths ranging from 18 to 28 km. The depths of the LFEs are also deeper than the indicated seismogenic zone depth of 15 km (Thurber et al., 2006). As in the case of NVT, the located LFEs fall into two groups, one adjacent to the locked Cholame segment of SAF and the other about 60 km northwest of Parkfield along the SAF (Figure 2.2). There is a clear discrepancy beneath the scattered locations of NVT and the clustered locations of LFEs.

Possible reasons for these relocation discrepancies are: 1) The observations used to determine the relative arrival times of NVT in the two methods are different; 2) NVT is more than just superimposed LFEs; 3) the SPDD location method is not reliable. According to the mathematical analysis of Wolfe (2002) and Zhang et al. (2010), due to too small differences in location partial derivations for station pairs that contain two very close-by stations, the location resolution can be impacted. The SPDD method can be insensitive to event depth when two close stations are directly above the event. The resulting variations in depth would be smeared into the horizontal directions due to the damping constraint in the SPDD. These cases could lead to a range of location scattering in all three directions. We explore these alternative explanations using both real and synthetic NVT and LFE data.

2.4. NVT signal simulation with LFEs

As hypothesized by Shelly et al. (2006) and Shelly and Hardebeck (2010), NVT is composed of a series of LFEs. If this hypothesis is correct, the NVT signal can be simulated approximately by superimposing the series of LFEs that occurred within the NVT episode. With simulated NVT waveforms, the envelope correlation procedure (Nadeau and Guilhem, 2009) and SPDD location method (Zhang et al., 2010) can be applied to investigate the possible causes of scattered NVT locations.

The LFE selection for each NVT event is a key step for NVT simulation. In Shelly and Hardebeck (2010), 88 LFE family templates were used to identify LFEs within continuous seismic data. Based on the catalog occurrence time of these repeating 88 LFE families (Shelly pers. comm.), the LFEs that occurred within each NVT event can be identified. As Nadeau and Guilhem (2009) used a 360-second window to detect NVT events, the LFEs that occurred within the 360-second period after each NVT's catalog reference time (Nadeau pers. comm.) are selected.

Five NVT events from the catalog of Nadeau (pers. comm.) are selected for simulation. These five NVT events and the LFEs within each NVT are plotted in Figure 2.3 NVT 669, 700 and 702 have clustered LFE events whereas NVT 660 and 688 have LFE events distributed over a large distance range along the SAF. These five selected NVTs are representative of the LFE and NVT location relationships in the catalog data.

Shelly and Hardebeck (2010) used linear stacking to construct the 88 LFE family template waveforms. In order to improve the signal to noise ratio (SNR), the Phase-Weighted Stacking (PWS) method (Thurber et al., 2014) was applied to generate impulsive waveforms for the selected LFEs. For each LFE family, the waveform signals were stacked separately for the east and vertical components on 14 High Resolution Seismic Network (HRSN) stations (Figure 2.4). Each PWS stacked LFE waveform contains 10 s before and 30 s after the catalog occurrence time.

For each station, the waveforms of LFEs are superimposed according to the NVT catalog reference time:

$$\mathrm{t}_{\mathrm{mn}} - t_{m0} = b_{mn} \tag{2.1}$$

where t_{mn} is the catalog arrival time of LFE n at station m, t_{m0} is the occurrence time of the selected NVT event for station m, b_{mn} is the begin time of LFE n on station m, with the time of 0 as the starting time for the simulated NVT event. All PWS stacked LFE waveforms are normalized so their maximum amplitudes are all equal to 1 initially. However, in nature, the LFEs maximum amplitude will change due to different energy release. When superimposing the PWS stacked LFE waveforms, random amplitude scaling between 0.5 and 0.95 is applied to each LFE so that the energy level will vary across different LFEs. Following the approach used by Nadeau and Guilhem (2009), noise signature for each station is generated based on the median amplitude over a 28-day signal. An SNR threshold of 1.65 was used in tremor detection for HRSN stations by Nadeau and Guilhem (2009). Therefore, the noise signature is added to each normalized simulated NVT waveform with 1.65 SNR (Figure 2.4).

After the simulated tremor is constructed, the RMS envelope is generated following the approach of Nadeau and Guilhem (2009). The sample rate for the simulated NVT signal is 20 Hz. A 101-sample boxcar with step size of 2 was used to calculate the RMS envelope, and the RMS envelope is decimated to a 10 Hz sample rate. Then a two-way 0.07 Hz low-pass filter was applied to produce the final RMS envelope for each simulated NVT (Figure 2.4).

2.5. Identification of closure error using simulated NVT signal

Zhang et al. (2010) used the SPDD method to relocate the NVT events. The input data for SPDD is the NVT station-pair differential times from RMS envelope cross-correlation (CC). There is no absolute time data used by SPDD. To generate the station-pair differential time for simulated

NVT events, CC is applied to the RMS envelopes of the simulated NVT events. Consider the three CC differential times, T_{ab}, T_{bc} and T_{ca} for stations a, b, and c. We define the closure error δ_{abc} as

$$\delta_{abc} = T_{ab} + T_{bc} + T_{ca} \tag{2.2}$$

If the station pair differential times are perfectly measured, then δ_{abc} should be zero. However, in most of our CC results, δ_{abc} is not zero. We also found the existence of closure errors in the station-pair CC differential time data from Zhang et al. (2010), which were originally from Nadeau and Guilhem (2009). Such closure errors may lead to relocation bias in the SPDD location method, as the SPDD is purely based on station pair differential time data.

Nadeau and Guilhem (2009) filtered out station-pair CC differential time outliers, defined as having closure errors larger than 0.5 s. Thus, the closure errors in the remaining dataset are all smaller than 0.5 s. Closure errors for all sets of three stations (a, b, c) are calculated from the station-pair CC differential time data (Nadeau and Guilhem, 2009; Zhang et al., 2010). The mean value and standard deviation have an average of ~ 0 s and ~ 0.22 s across all stations, respectively.

2.6. Effects on LFE relocation due to closure error

The CC station-pair differential times are calculated for HRSN stations with simulated NVT waveforms. We observe that many of the CC station-pair differential times are zero due to very similar simulated NVT RMS envelopes. In addition, the 14 HRSN stations have limited azimuthal coverage relative to the LFE and NVT locations (Figure 2.3). In Zhang et al. (2010), 67 stations including HRSN were used for NVT location in order to obtain comprehensive azimuthal coverage. Thus the results will likely be biased if we study closure error effects with the simulated NVTs and HRSN stations. Compared to the simulated NVTs, all selected LFEs are well located by Shelly and Hardebeck (2010) using P-and S-arrival times with a grid search method. Based on these considerations, we elected to conduct the tests with synthetic travel time data generated using selected LFEs and 67 stations in Zhang et al (2010) (Figure 2.3), rather than testing with simulated NVTs.

To test the effects of closure error on SPDD relocation results, we apply the SPDD method to relocate LFEs using synthetic data with and without closure error. Synthetic travel time data for the LFEs were generated for the 67 stations (Figure 2.3) with a one-dimensional (1D) S-wave velocity model. This 1D model is from the USGS for their NCSN earthquake catalog and was used to relocate tremors by Zhang et al (2010). For all station pairs in Zhang et al. (2010), we derive the synthetic station pair differential data by differencing the synthetic travel times for each LFE. Then we add in Gaussian noise with 0 s mean and 0.22 s standard deviation into the synthetic dataset to mimic the closure error in the real dataset from Zhang et al. (2010).

The SPDD method needs initial locations for all events to start the relocation process. To set up the initial location for each LFE, we examined the distribution of LFEs that occurred within the same NVT event. As shown in Figure 2.3, all LFEs in these simulated tremor events can be divided into two groups: 1) clustered LFEs within each NVT (669, 700, 702); and 2) scattered LFEs within each NVT (660, 688). For the LFEs that occurred in the same NVT event, the averaged LFE location is set to be their initial locations in our first set of tests: 1) LFE cluster center as the initial location, noise-free synthetic data (LNN); 2) LFE cluster center as the initial location, closure error noise added to the synthetic data (LWN). For LFEs in group 1, the average LFE location is closer to them than the corresponding NVT location. However, for LFEs in group 2, the corresponding NVT location can be closer to the LFEs than the averaged LFE location. To check whether setting initial locations closer to true locations can affect the relocation results, we add another set of tests, with each LFE initial location placed at the corresponding NVT location: 3) NVT catalog location as the initial location, noise free synthetic data (NNN); 4) NVT catalog location as the initial location, closure error noised added to the synthetic data (NWN). The initial locations should not affect the relocation results if the SPDD method always finds the global optimal location. If not, the relocation results can be different if SPDD becomes stuck in a local minimum.

Relocation results for the four tests are shown in Figure 2.5 and 2.6. To study the effect of closure error on relocation results we compared the relocation errors for two pairs of test datasets: (1) LNN and LWN; (2) NNN and NWN. The mean relocation error (MRE) differences are 0.92 and 0.23 km for (1) and (2), respectively. For the clustered LFEs in NVT 669, 700 and 702, the MRE differences are 0.65 and 0.20 km for (1) and (2), respectively. For the non-clustered LFEs in NVT 660 and 688, the MRE differences are 1.22 and 0.26 km for (1) and (2), respectively. For NNN and NWN, the MREs are 13.31 and 13.13 km, respectively, and the MRE difference is only about 1.7 – 1.8 %. For LNN and LWN, the MREs are 6.63 and 6.67 km, respectively, and the MRE difference is about 9.7 %-9.8 %. The comparison shows that closure error has a larger impact on the relocation results when the initial locations are closer to the true location.

The initial locations also affect the relocation results. In Figure 2.5 and Figure 2.6, relocations with average LFE locations as initial locations (LNN and LWN) are generally closer to the true location compared to those with NVT as initial locations (NNN and NWN). MREs for LNN and LNW are generally smaller than those for NNN and NWN. This observation is more apparent when LFEs are clustered and away from NVT locations. Average MREs for NVT 669, 700 and 702 are 9.59 km in NNN and 4.05 km in LNN, more than a 50% reduction.

2.7. Local minimum test for SPDD

In the synthetic relocation results shown previously, we found that the relocation error depends heavily on the initial guess of the location. If this initial location dependency is real, then we expect to see similar results when using real data. To further investigate this hypothesis, we used real data from Zhang et al. (2010) to test the impact of the initial location on relocation results. As shown in Figure 2.7, five NVT events from Zhang et al. (2010) are selected, which are close to the LFE events tested previously. For each selected event, we choose 10 points as the initial locations, equally spaced along the horizontal line connecting the catalog epicenter to the target location (red star in Figure 2.7). For all these initial locations, the depths are kept the same as in catalog. We used the same 1D S-wave velocity model from the USGS to relocate these events with the SPDD method. The final relocation results for these five events are shown in Figure 2.8 For each event, the final relocations are always closer to the initial locations than the catalog location. When the initial location is far away from the catalog location, the depths in the relocation results shift up and down around the catalog depth. The shifts in the Y axis direction (north-south) changes sign close to the middle point, whereas the shifts in the X axis direction (east-west) are almost all in the same direction. These relocation results using real data support our conclusion from the previous tests with synthetic data, that the relocation results are heavily dependent on the initial locations.

The SPDD method adopts the least squares solution for the non-linear inversion problem (Zhang et al., 2010). When solving a non-linear inversion problem with least squares, the algorithm can suffer from local minimum problems (Aster et al., 2013). During the inversion process of the SPDD algorithm, the initial locations can lead to different location perturbation directions, for which the residual decreases. It is therefore quite possible that the relocations from SPDD can become stuck in a local minimum during the inversion. To investigate the local minimum problem in SPDD, it is necessary to assess the spatial distribution of misfit. In order to assess the relocation residuals around the catalog locations, we use a simple grid search to evaluate misfit in a 3D volume around the catalog location of event 123, as an example. For each grid point in the volume, the mean residuals are calculated from the results for station pair differential time misfit.

For the local minimum test, we used the catalog location of NVT 123 and shifted the catalog location to create different initial locations (Figure 2.9). We shifted the catalog location in the X and Y directions from -5 km to 5 km from the catalog location with an interval of 0.5 km. In the Z direction, we shifted the depth from -2.5 km to 2.5 km from the catalog depth with an interval of 0.5 km. In total, the number of test locations in X, Y and Z directions are 21, 21 and 11 respectively. Then with each test point as initial location, we used the 1D S-wave model and the real data from Zhang et al. (2010) to get the relocation. Misfits on these relocations are calculated.

We find that the residuals range between 924 and 1072 ms, compared to the residual in the catalog location of 952.83 ms. The location with minimum residual values is at a depth of 30.5 km, whereas the catalog depth of event 123 is at 28.0 km. This indicates that inversions likely became stuck in a local minimum with different initial locations. The catalog location of NVT 123 in our cartesian coordinates is (-10.138, -56.217, -27.989) km. The residual distribution at depths 27.0, 27.8, and 28.6 km are shown in Figure 2.10. At depth 27.0 km, a clear multi-peaked residual distribution can be seen, which indicates that there are multiple local minima. From the valley between the these peaks, the location perturbation direction could be from either slope up to one of the “hills”. This can lead the result to become stuck in either local minima. At greater depths of 27.8 and 28.6 km, the prominent multi-peak distribution fades. However, the valley between the two large peaks is still visible. These peaks indicate local minima at each depth, one of which can potentially be the global minimum. These local minima can potentially trap the SPDD inversion and lead to the failure of finding a better location with a lower residual.

In Figure 2.11, we plot the isosurfaces of residual equal to 930 and 955 ms in 3-D space. Multiple separate isosurfaces for residual = 930 ms can be seen clearly and most of them are at greater depths than the catalog depth. These isosurfaces show that when the SPDD inversion searches the residual decrease direction, it can become trapped within one of these volumes, likely to the one closest to the initial location.

2.8. Discussion

In order to address the question of whether NVTs can be represented by a series LFEs, we simulated NVTs and carried out a relocation analysis for LFEs composing these simulated NVTs to check if the results are as scattered in Nadeau and Guilhem (2009) and Zhang et al. (2010). The simulated waveforms (Figure 2.4) are similar to real NVT waveforms. However, the amplitude of the real LFEs cannot be reproduced accurately, which will lead to the differences in the NVT waveform and RMS envelope between simulated and real NVTs. Following the NVT differential time estimation process in Nadeau and Guilhem (2009) and Zhang et al. (2010), we identified closure error in the station-pair differential data for both simulated NVTs and real NVTs. Even though a pre-filtering was done by Nadeau and Guilhem (2009) to remove data with large closure errors, closure errors with 0 s mean and 0.22 s standard deviation can result in a 9.8% increase in relocation error. However, a relocation change at a scale of 0.26 to 1.22 km cannot be counted as the major reason for the NVT-LFE location discrepancy, which is at a scale of 10 km.

While investigating the results from synthetic LFE relocation tests, we observed a more than 50% increase in relocation error when the initial location is far from the true location. This increase can be as large as ~ 29 km (event 41449 in Table 2.1). This is more comparable with the NVT-LFE location discrepancy. The results for relocating the same event with different initial locations (Figure 2.8) further supports the argument that the results from the SPDD method depends heavily on the initial location.

One major reason for the initial location dependency is that the inversion process can become stuck in a local minimum. The local minimum test results (Figure 2.10 and 2.11) show that the local minimum problem can be encountered by SPDD when searching for the minimum residual point.

2.9. Conclusion

In this study, a series of experiments are conducted to examine the reason why there is a NVT-LFE location discrepancy along the central SAF. From the NVT simulation tests, NVT-like waveforms can be produced by superimposing LFEs in synthetic NVT event with noise. By generating station pair differential times with simulated NVT waveforms, we identified closure error in both the simulated and real datasets. With synthetic LFE relocation tests, we find that the closure error can have a considerable impact on the relocation results, but the closure error should not be the main reason for the NVT-LFE location discrepancy. Results from synthetic LFE relocation tests indicate a strong initial location dependency for the SPDD method. The initial location dependency is caused by local minimum problems encountered by the SPDD method. Such dependency will lead to relocation changes at a scale similar to the NVT-LFE location discrepancy. Zhang et al. (2010, 2017) used NVT locations from Nadeau and Guilhem (2009) as initial locations. The NVT arrival times were inverted with CC station-pair differential times. This inverted arrival time dataset can lead to biased location results. Since the relocations in Nadeau and Guilhem (2009) are scattered from the SAF, NVTs in Zhang et al. (2010) are scattered from the SAF as well. Synthetic LFE relocation analysis from our study do not explain the reason for the scatter in the original Nadeau and Guilhem (2009) NVT locations.

2.10. Acknowledgements

This research is supported in part by the USGS, Department of the Interior, under USGS Award Number G14AP00056 to the University of Wisconsin-Madison. The support of a student assistantship from the Mark and Carol Ann Solien Graduate Assistantship and George P. Woollard-Sigmund I. Hammer Memorial Fund in Geology & Geophysics of the University of Wisconsin-Madison are also acknowledged.

2.11. References

Aster, R.C., Borchers, B. and Thurber, C.H., 2013, Parameter estimation and inverse problems. 2nd edn, Elsevier, 240 pp.

Guo, H., Zhang, H., Nadeau, R.M. and Peng, Z., 2017, High-resolution deep tectonic tremor locations beneath the San Andreas Fault near Cholame, California, using the double-pair double-difference location method. J. Geophys. Res.: Solid Earth, 122(4), pp.3,062-3,075.

Katsumata, A. and Kamaya, N., 2003, Low-frequency continuous tremor around the Moho discontinuity away from volcanoes in the southwest Japan. Geophys. Res. Lett., 30(1), 20-1.

McBride, J. H., and Brown, L. D., 1986, Reanalysis of the COCORP deep seismic reflection profile across the San Andreas fault, Parkfield, California. Bull. Seismol. Soc. Am., 76(6), 1,668-1,686.

Murray, J., and Langbein, J., 2006, Slip on the San Andreas fault at Parkfield, California, over two earthquake cycles, and the implications for seismic hazard. Bull. Seismol. Soc. Am., 96(4B), S283-S303.

Nadeau, R. M., and Guilhem, A., 2009, Nonvolcanic tremor evolution and the San Simeon and Parkfield, California, earthquakes. Science, 325(5937), 191-193.

Nadeau, R.M. and Dolenc, D., 2005, Nonvolcanic tremors deep beneath the San Andreas Fault. Science, 307(5708), 389.

Nishide, N., Hashimoto, T., Funasaki, J., Nakazawa, H., Oka, M., Ueno, H., . Yamada, N., Sasakawa, I., Maeda, K., Sugimoto, K. and Takashima, T., 2000, Nationwide activity of low-frequency earthquakes in the lower crust in Japan. In Abstr. Jpn. Earth and Planet. Sci. Joint Meeting, sk-p002.

Obara, K., 2002, Nonvolcanic deep tremor associated with subduction in southwest Japan. Science, 296(5573), 1,679-1,681.

Shelly, D. R., and Hardebeck, J. L., 2010, Precise tremor source locations and amplitude variations along the lower-crustal central San Andreas Fault. Geophys. Res. Lett., 37, L14301, doi:10.1029/2010GL043672

Shelly, D. R., Beroza, G. C., Ide, S., and Nakamula, S., 2006, Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip. Nature, 442(7099), 188-191.

Thurber, C. H., Zeng, X., Thomas, A. M., and Audet, P., 2014, Phase-weighted stacking applied to low-frequency earthquakes. Bull. Seismol. Soc. Am., 104(5), 2,567-2,572.

Thurber, C., Zhang, H., Waldhauser, F., Hardebeck, J., Michael, A., and Eberhart-Phillips, D., 2006, Three-dimensional compressional wavespeed model, earthquake relocations, and focal mechanisms for the Parkfield, California, region. Bull. Seismol. Soc. Am., 96(4B), S38-S49.

Waldhauser, F., and Schaff, D. P., 2008, Large-scale relocation of two decades of Northern California seismicity using cross-correlation and double-difference methods. J. Geophys. Res.: Solid Earth, 113, B08311, doi: 0.1029/2007JB005479.

Wolfe, C. J., 2002, On the mathematics of using difference operators to relocate earthquakes. Bull. Seismol. Soc. Am., 92(8), 2,879-2,892.

Zeng, X., Thurber, C.H., Shelly, D.R., Harrington, R.M., Cochran, E.S., Bennington, N.L., Peterson, D., Guo, B. and McClement, K., 2016, 3-D P-and S-wave velocity structure and low-frequency earthquake locations in the Parkfield, California region. Geophys. J. Int., 206(3), 1,574-1,585.

Zhang, H., Nadeau, R. M., and Toksoz, M. N., 2010, Locating nonvolcanic tremors beneath the San Andreas Fault using a station-pair double-difference location method. Geophys. Res. Lett., 37, L13304, doi:10.1029/2010GL043577

2.12. Table and Figures

Table 2.1 Synthetic LFE relocation results for four tests in closure error investigation.

NVT ID	LFE ID	Relocation distance from LFE catalog location (km)			
		LNN	LWN	NNN	NWN
669	40453	2.65	2.09	7.73	7.61
	33845	1.86	0.57	9.53	9.46
	8893	2.84	1.53	11.41	11.37
	58459	2.03	1.07	8.69	8.75
	61095	2.61	1.64	10.97	10.93
700	10824	4.47	4.26	6.83	6.95
	83647	2.06	1.29	12.87	12.77
	6196	1.47	0.89	11.01	10.79
	68539	4.30	3.66	16.18	15.97
702	37811	2.38	2.65	5.09	4.76
	15157	2.90	3.40	3.41	3.61
	16160	4.72	6.62	5.30	5.43
	45768	1.46	1.31	4.00	3.78
	30397	1.97	2.23	5.14	4.92
	34680	2.65	3.06	4.00	4.04
	77401	26.39	26.43	36.26	35.2
	37796	2.06	2.31	4.56	4.30
660	47706	4.76	4.18	12.83	12.91
	15157	2.88	2.33	3.97	3.99
	58755	2.59	1.89	4.13	4.07
	15025	3.26	2.60	4.20	4.00
	16160	1.79	1.36	6.18	5.39
	34680	2.10	1.51	3.97	4.04
	25753	4.63	4.14	11.61	11.2
688	58668	2.05	2.40	25.42	25.13
	75171	6.68	6.75	11.93	11.57
	58755	2.26	2.34	24.77	24.48
	70316	20.65	19.23	6.89	6.66
	40453	15.77	19.39	44.76	44.47
	41449	28.83	33.21	58.13	57.77
	36319	19.00	18.67	8.03	7.72
	58459	14.27	17.98	43.3	43.43
	9707	18.54	16.98	6.20	5.91

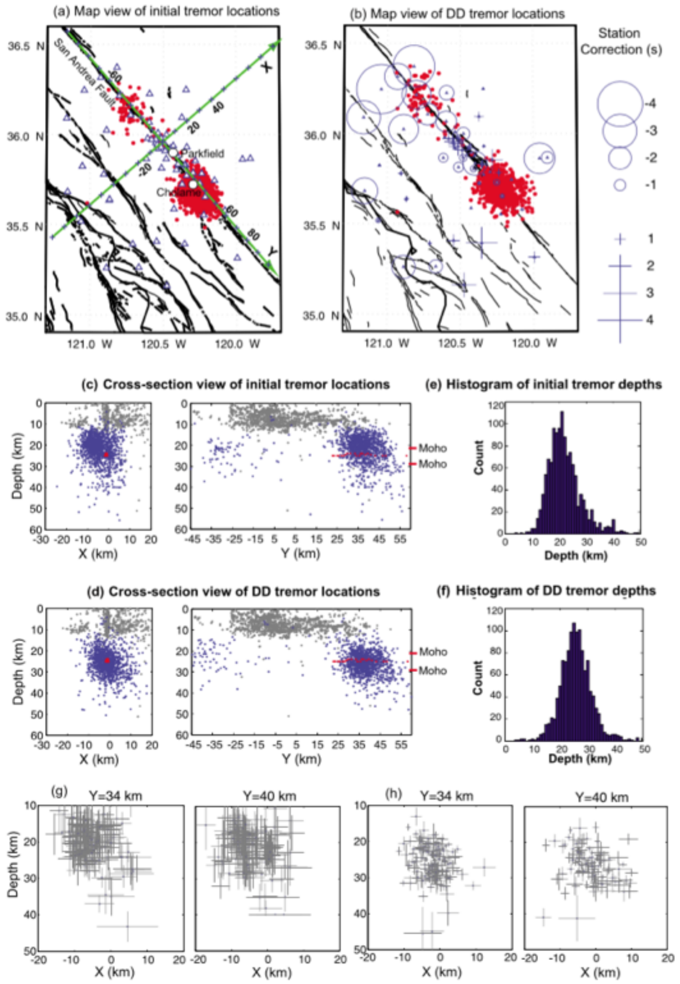


Figure 2.1 Catalog and DD tremor location comparison from Zhang et al., (2010). (a) Map view of catalog tremor locations (Nadeau and Guilhem, 2009). The coordinates are shown with green axes in X and Y direction, in km. Triangle: 64 seismic stations. Red dots: 1246 tremors. (b) Map view of DD tremor locations. Red dots: relocated 1246 tremors using station-pair DD location method. Station corrections are shown in seconds. (c) Cross-sections of catalog tremor locations in X-depth and Y-depth. (d) Cross-sections of DD tremor locations in X-depth and Y-depth. Red dots: LFE locations (Shelly et al., 2010). Blue dots: tremor catalog data. Gray dots: background seismicity (Thurber et al., 2006). Red lines are Moho depths of 22 and 29 km (Mcbride and Brown., 1986). (e) Catalog tremor location depth distribution. (f) DD tremor locations' depth distribution. (g) Across-fault cross-sections at Y = 34 and 40 km. 95% uncertainty bunds are shown for catalog tremor locations. (h) Across-fault cross-sections at Y = 34 and 40 km. 95% uncertainty bounds are shown for the DD tremor locations.

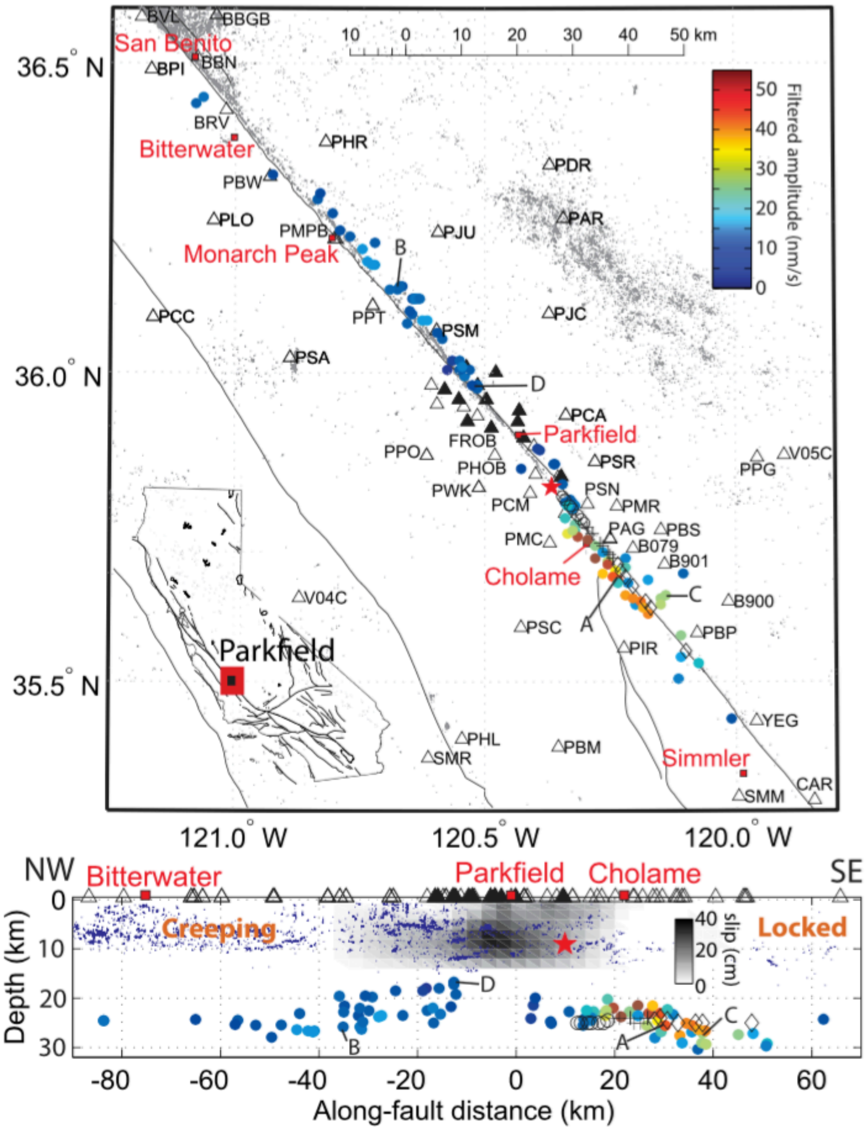


Figure 2.2 Template LFE family locations in central SAF region from Shelly et al., (2010). (top) Map view of LFE locations (colored circles) and seismic network stations. Different colors of LFEs show the amplitude potential. Black triangles are HRSN stations and white triangles are other stations in the CISN. Gray dots are microseismicity distributions from 1984 – 2003 (Waldhauser and Schaff, 2008). Red star is the hypocenter of 2004 M 6 earthquake. (Bottom) Cross-section along the SAF surface trace. Blue dots represent the microearthquakes located within 5 km to the fault trace. Gray shading shows the slip length during and the first 230 days after the 2004 earthquake (Murray and Langbein, 2006)

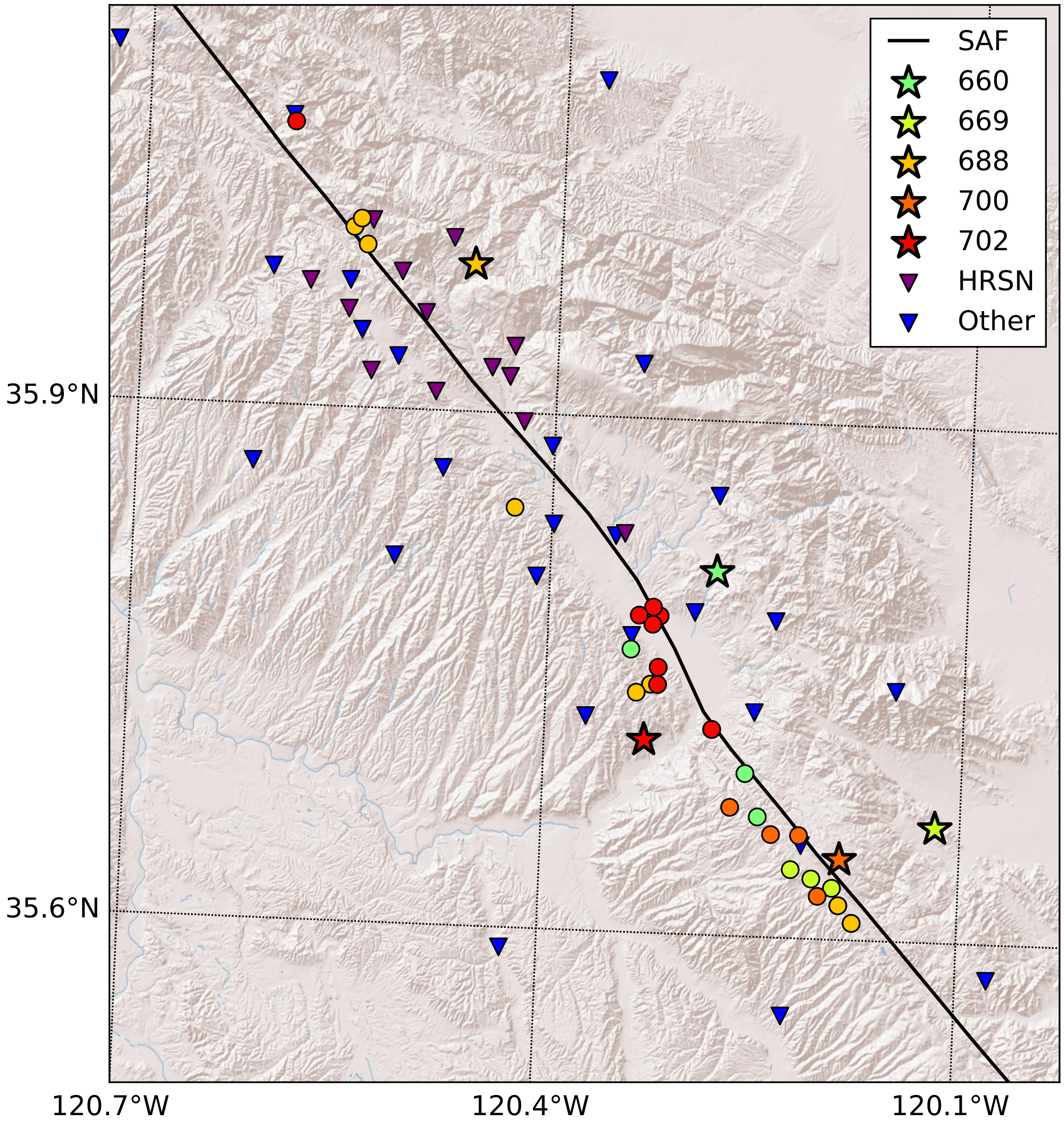


Figure 2.3 Map view of selected NVTs, LFEs and stations. Stars are five selected NVTs from Zhang et al. (2010). The circles are LFEs. For each NVT, the LFEs that occurred within its 360 s period are shown with the same color. Triangles are stations used in Zhang et al. (2010). The purple triangles are HRSN stations whereas the blue triangles are other stations. Solid black line is the SAF trace.

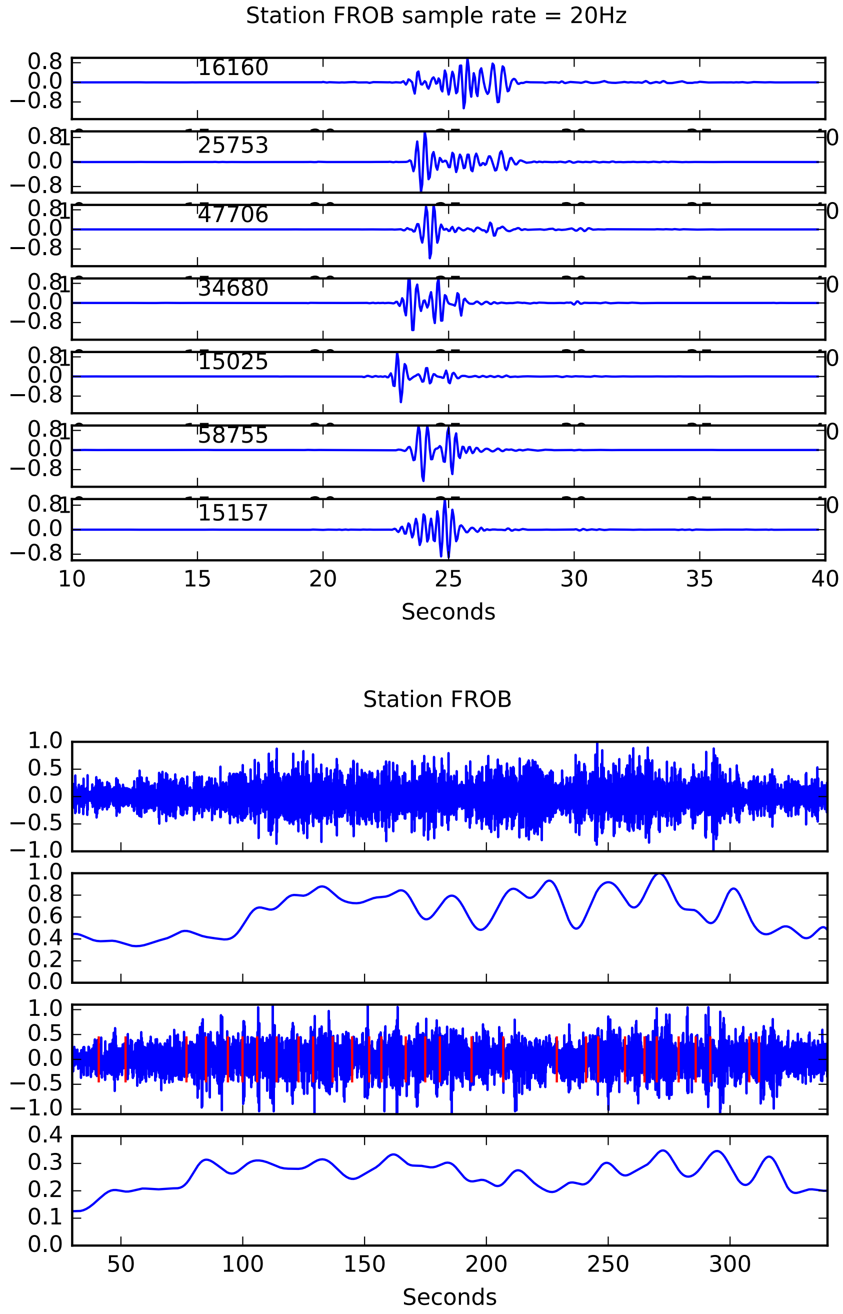
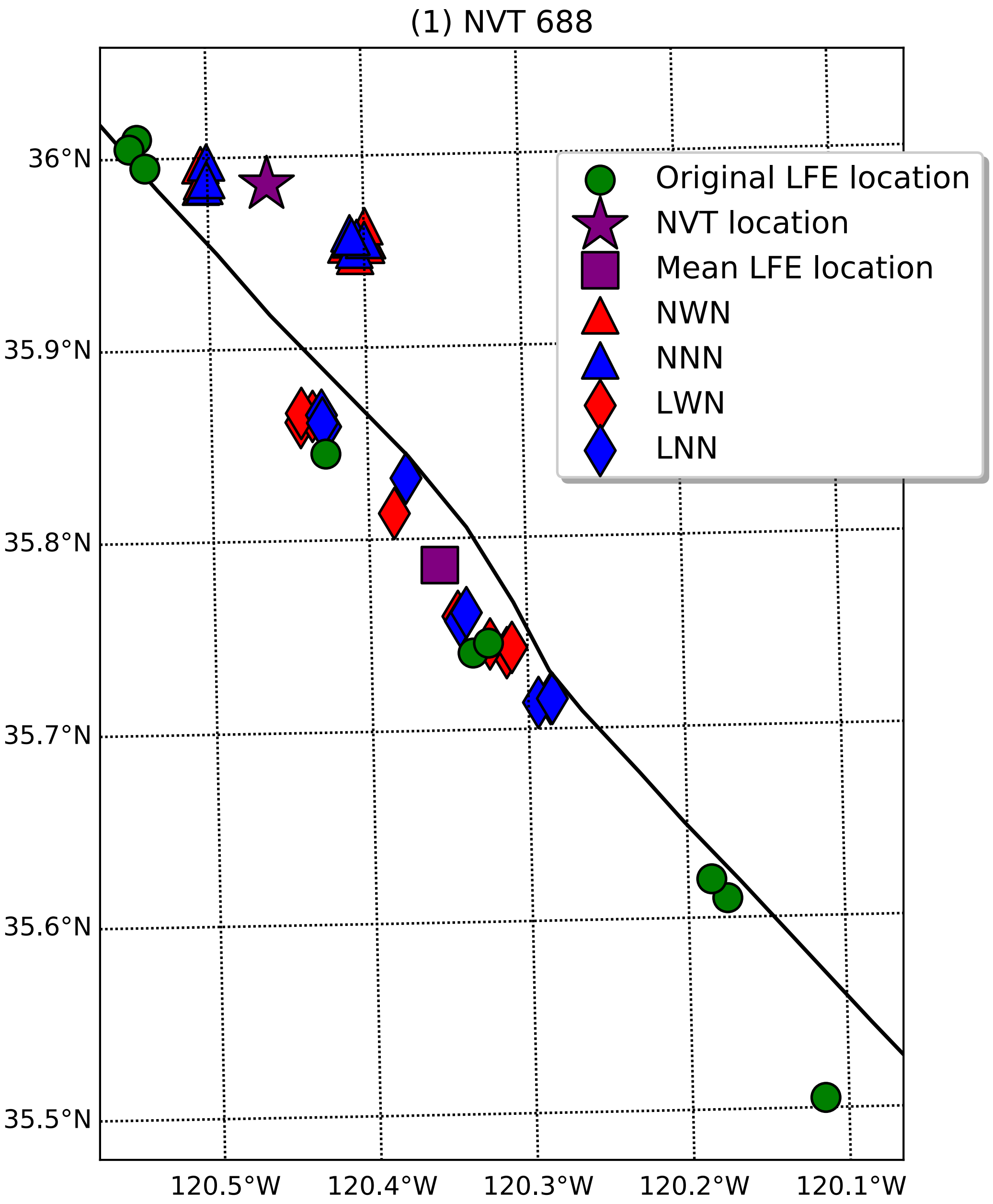
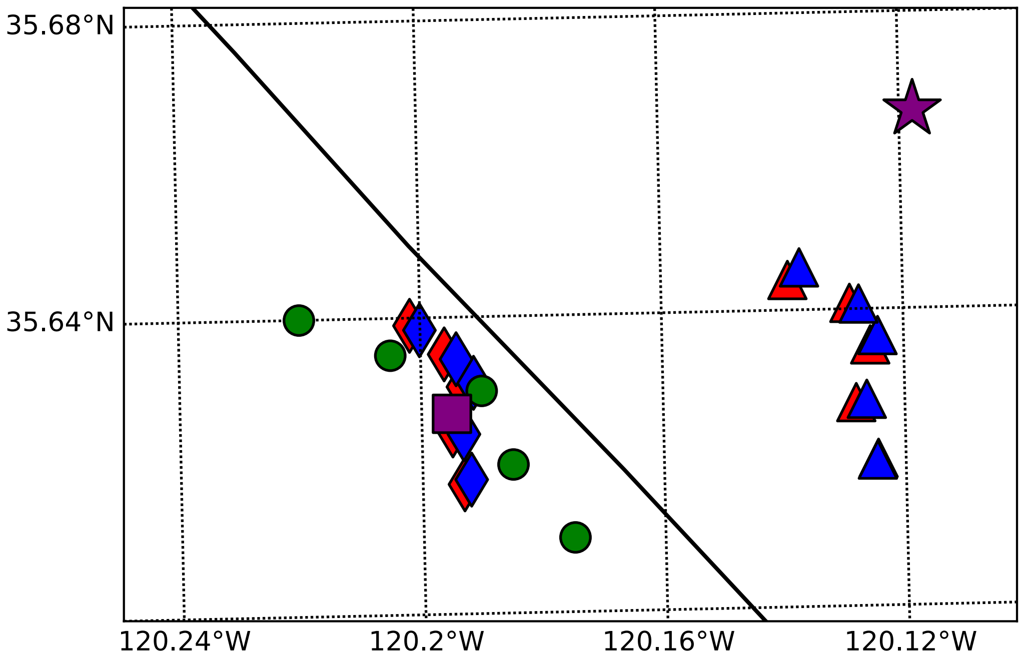


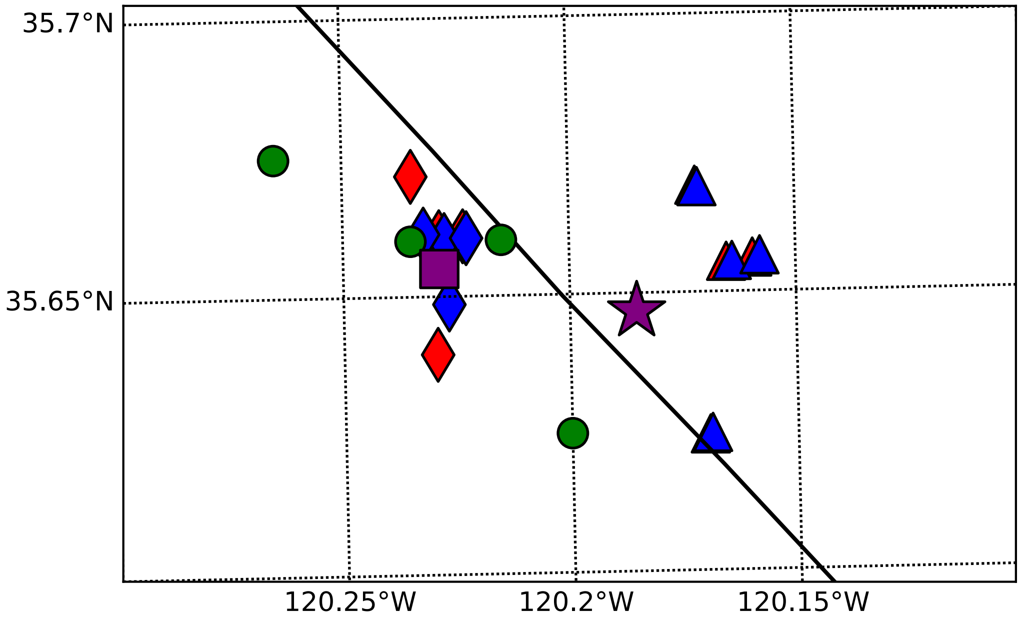
Figure 2.4 Upper panel: PWS stacked LFE waveforms for station FROB. Lower panel: comparison between real and simulated tremor waveforms for tremor event 660 station FROB. First trace: real tremor waveform. Second trace: RMS envelope of real tremor waveform. Third trace: simulated tremor waveform, red lines are marking the identified LFEs. Fourth trace: RMS envelope of simulated tremor.

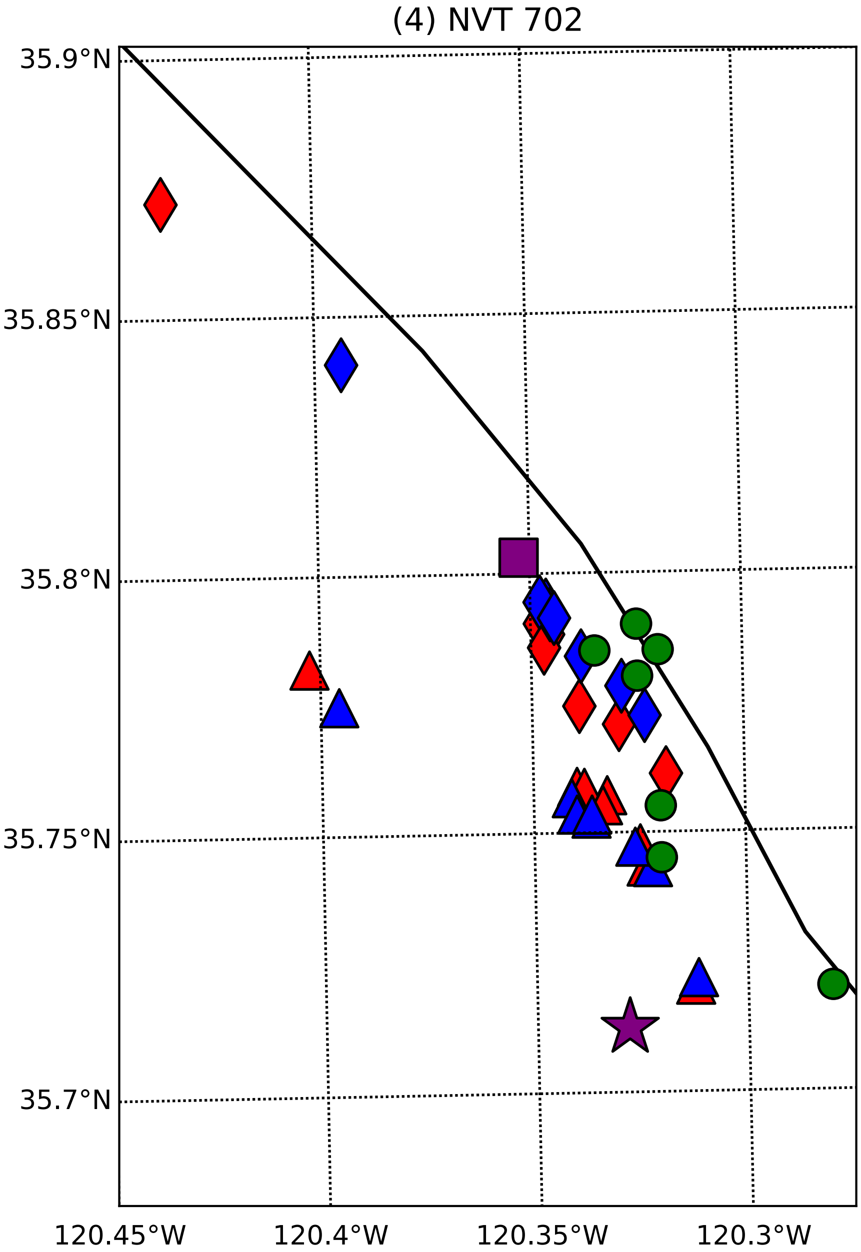

(1) NVT 688
36°N
35.9°N
35.8°N
35.7°N
35.6°N
35.5°N
120.5°W
120.4°W
120.3°W
120.2°W
120.1°W
Original LFE location
NVT location
Mean LFE location
NWN
NNN
LWN
LNN

(2) NVT 669



(3) NVT 700




(4) NVT 702
35.9°N
35.85°N
35.8°N
35.75°N
35.7°N
120.45°W
120.4°W
120.35°W
120.3°W

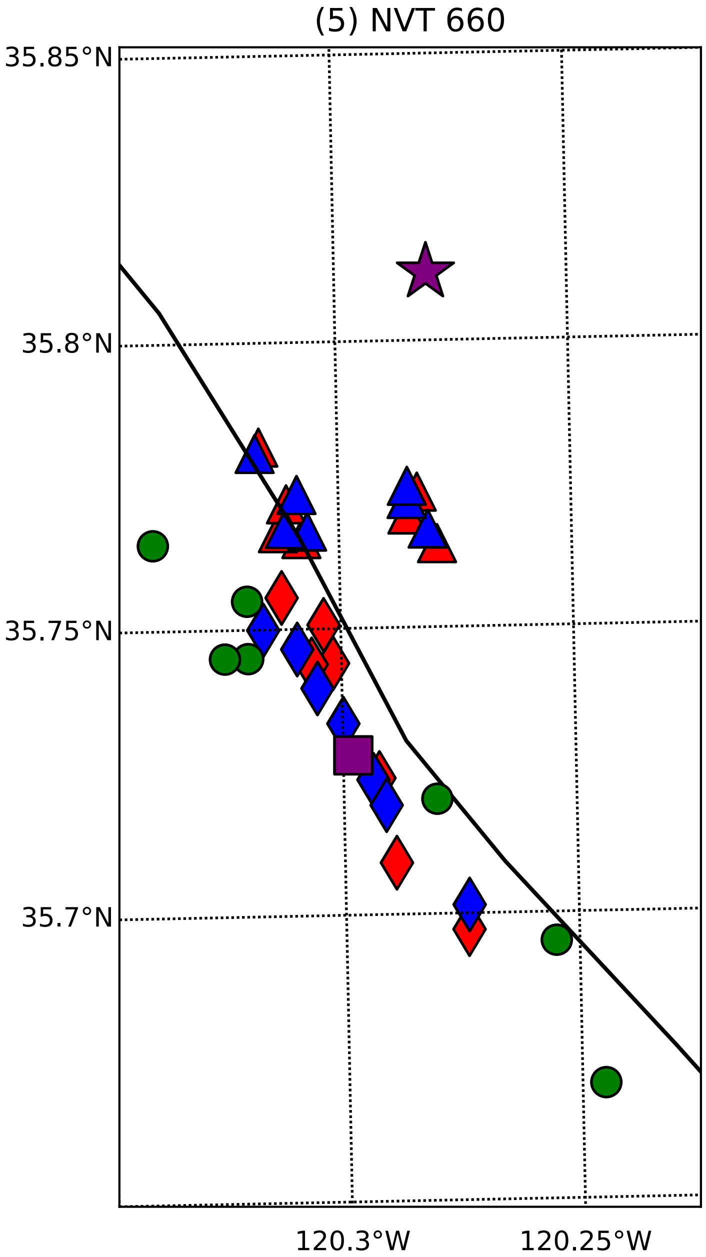
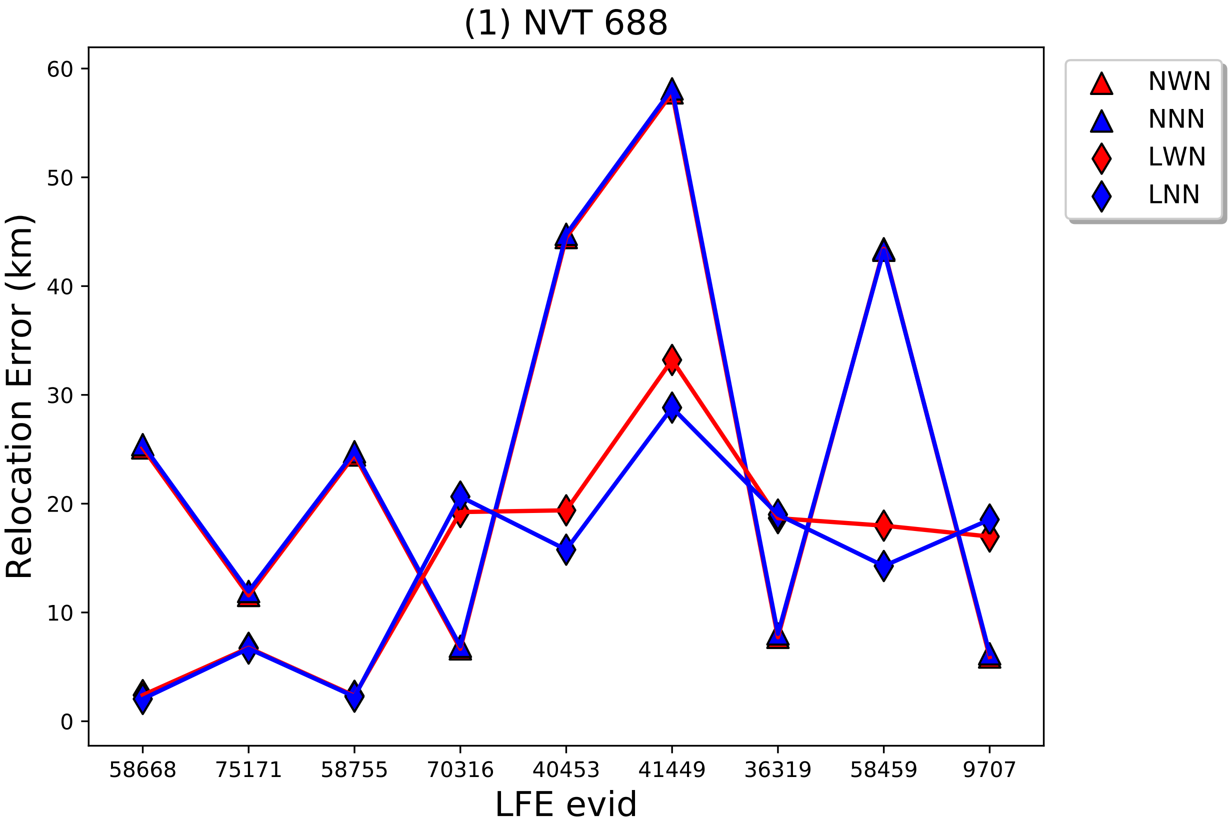
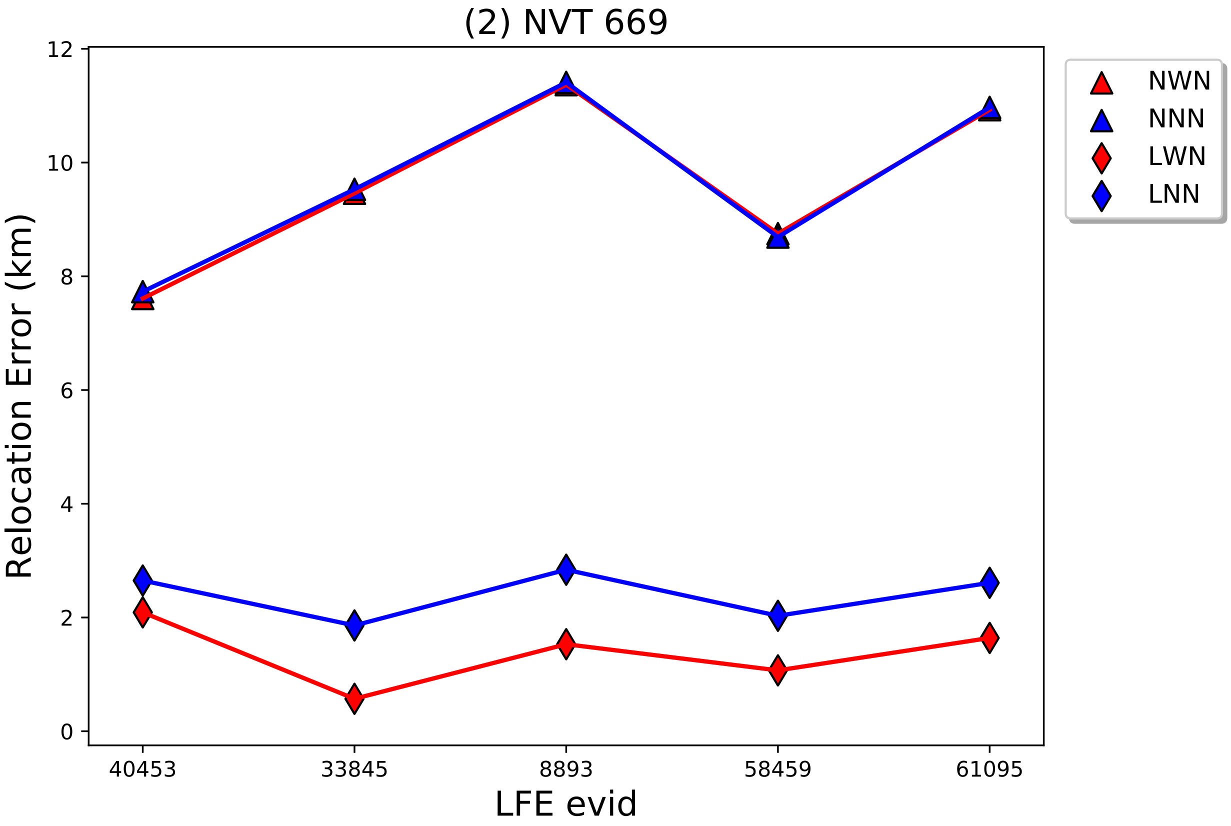
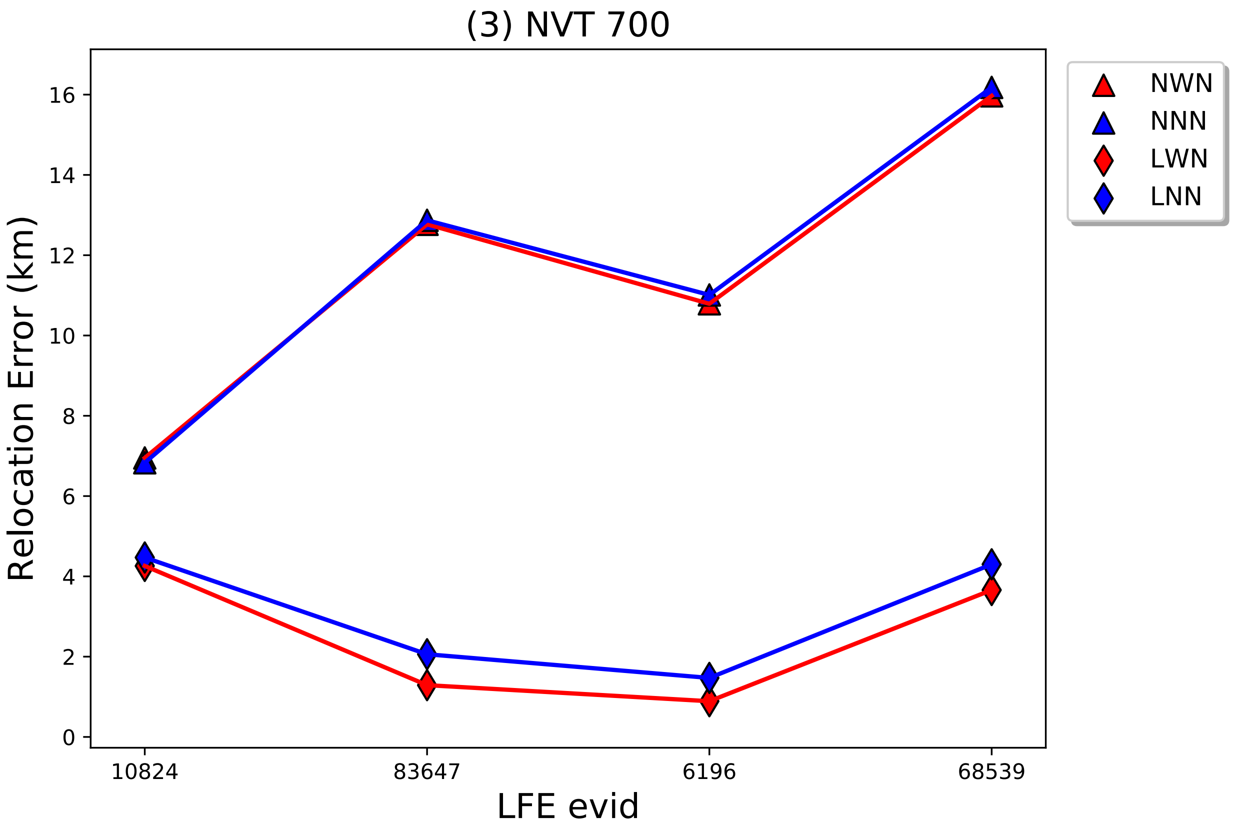
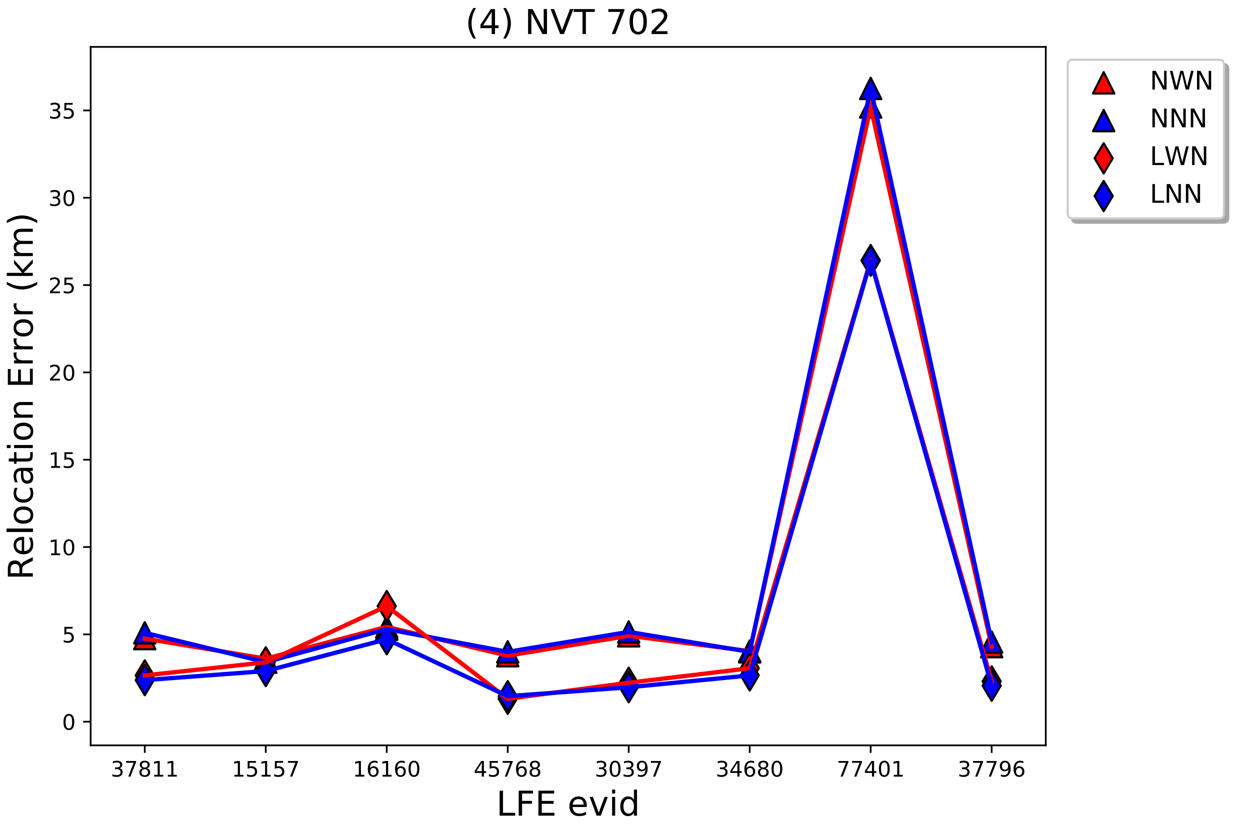


Figure 2.5 Relocation results for selected LFEs in the four tests. Circles are the catalog LFE locations. Purple square is the average LFE location. Purple star is the NVT catalog location. Triangles are LFE relocations for test1 and test2. Diamonds are LFE relocations for test3 and test4. Solid black line is the SAF trace.


(1) NVT 688
Relocation Error (km)
0
10
20
30
40
50
60
58668
75171
58755
70316
40453
41449
36319
58459
9707
LFE evid
NWN
NNN
LWN
LNN


(2) NVT 669
Relocation Error (km)
0
2
4
6
8
10
12
40453
33845
8893
58459
61095
LFE evid
NWN
NNN
LWN
LNN


(3) NVT 700
Relocation Error (km)
0
2
4
6
8
10
12
14
16
10824
83647
6196
68539
LFE evid
NWN
NNN
LWN
LNN


(4) NVT 702
Relocation Error (km)
0
5
10
15
20
25
30
35
37811
15157
16160
45768
30397
34680
77401
37796
LFE evid
NWN
NNN
LWN
LNN

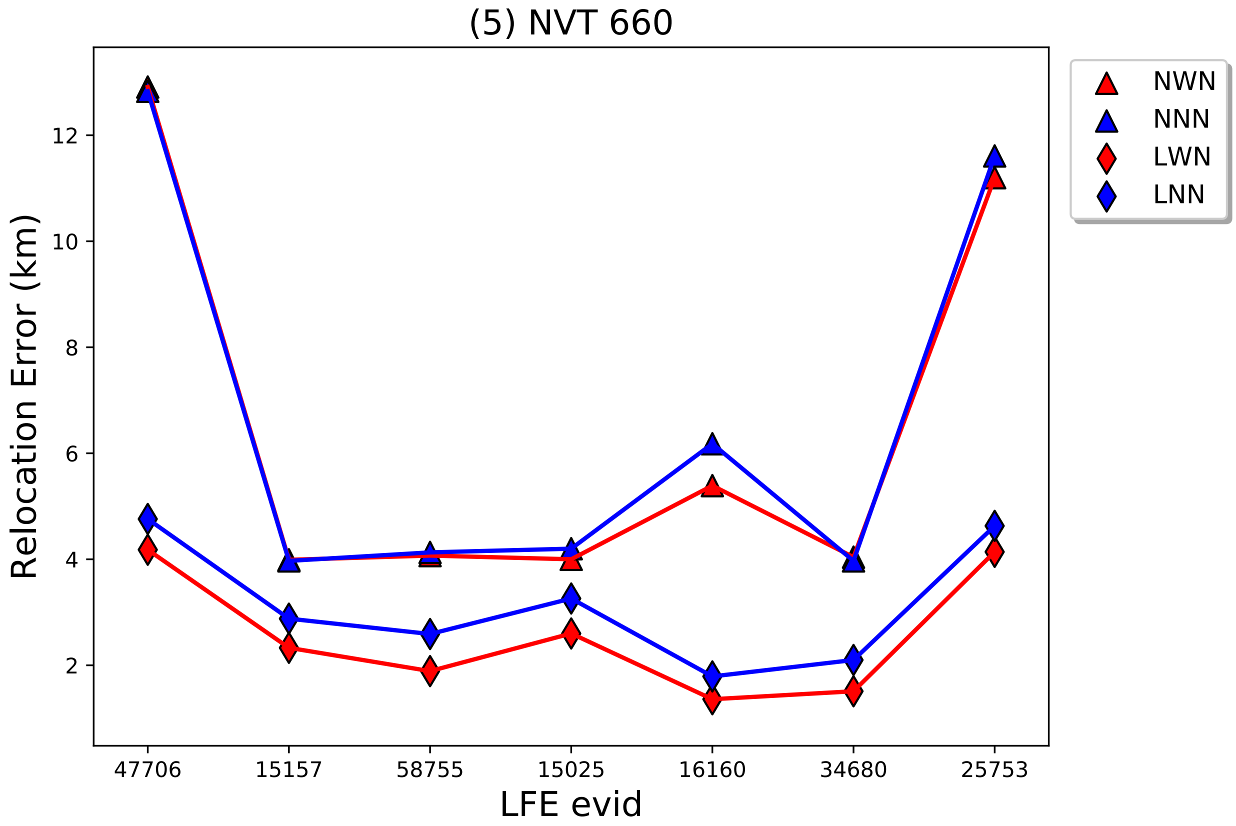


Figure 2.6 LFE relocation error for four tests. In each plot, the LFEs occurred within one NVT event. Triangles: relocation errors for test NNN (blue) and NWN (red). Diamonds: relocation errors for test LNN (blue) and LWN (red).

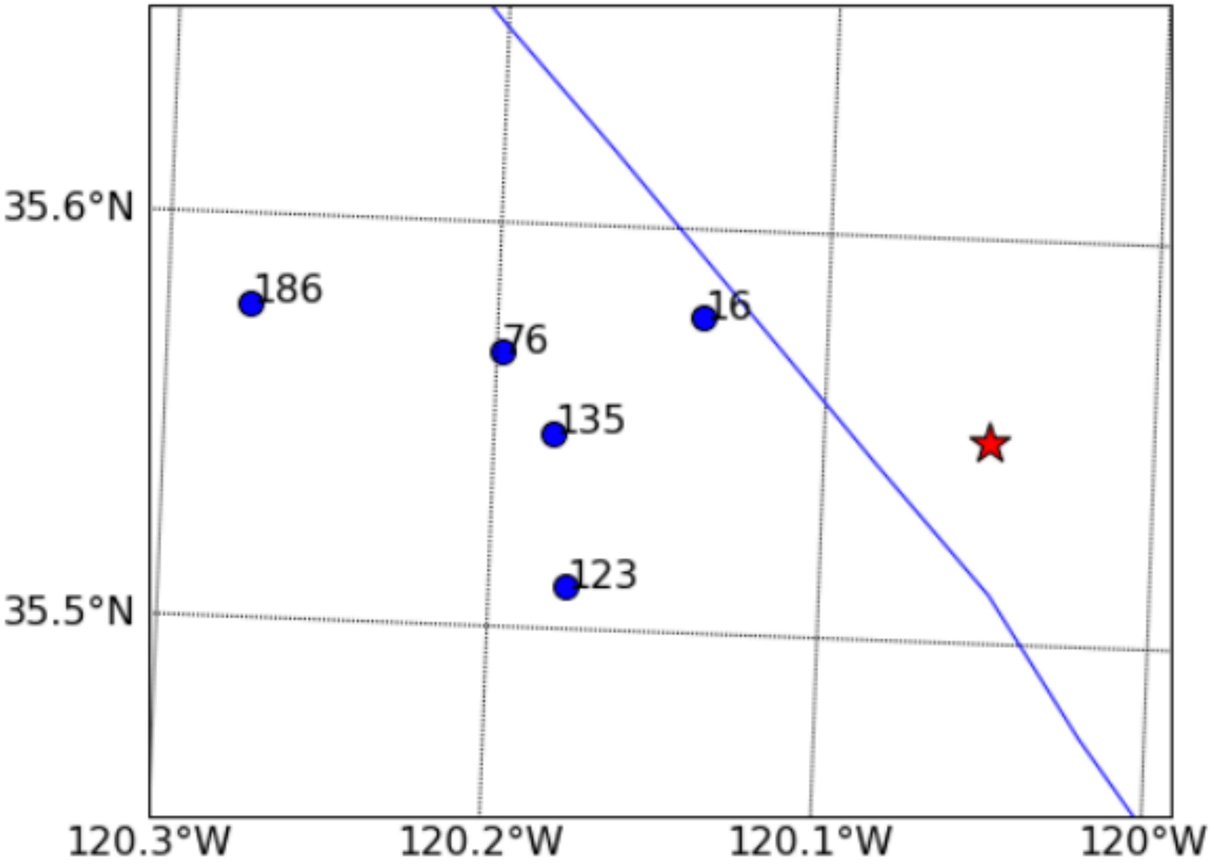
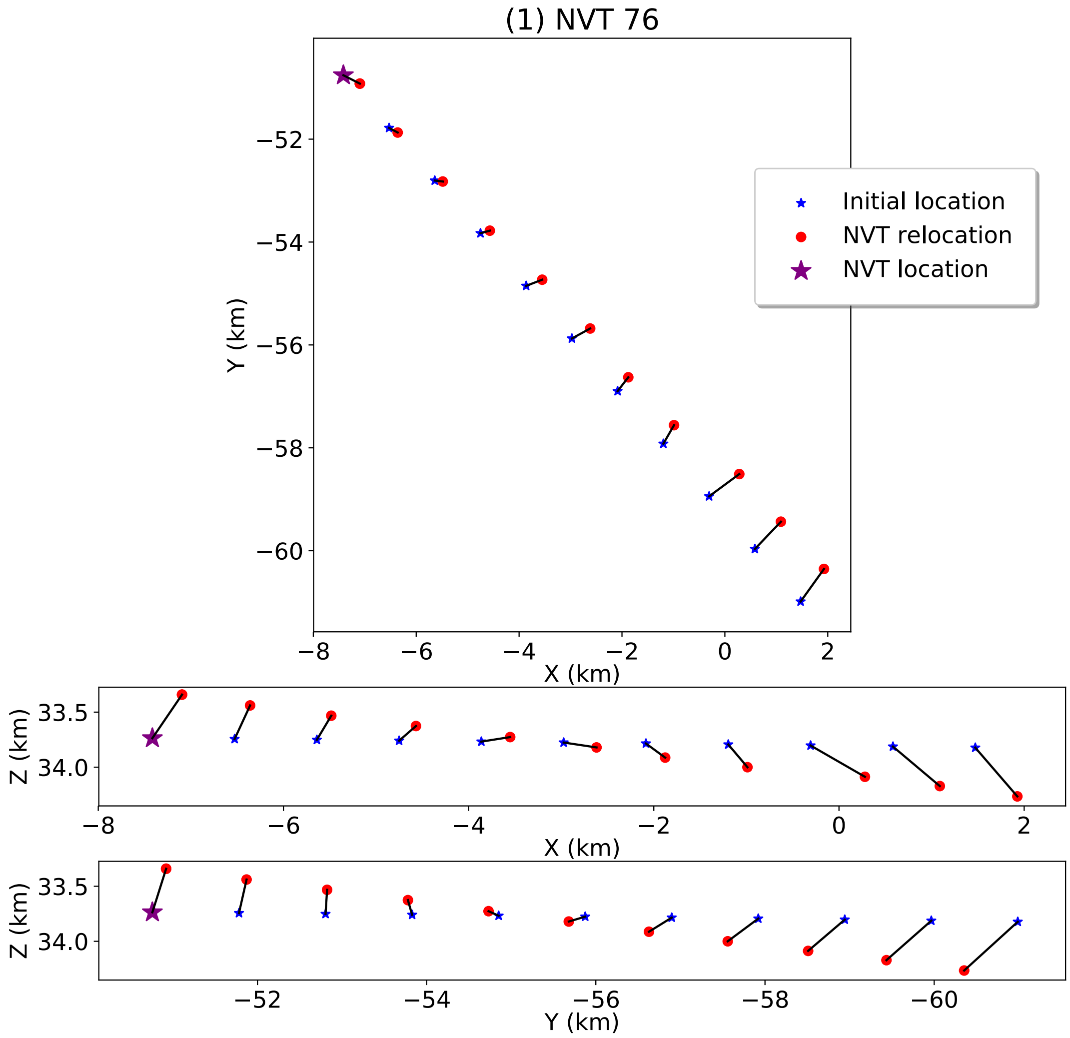
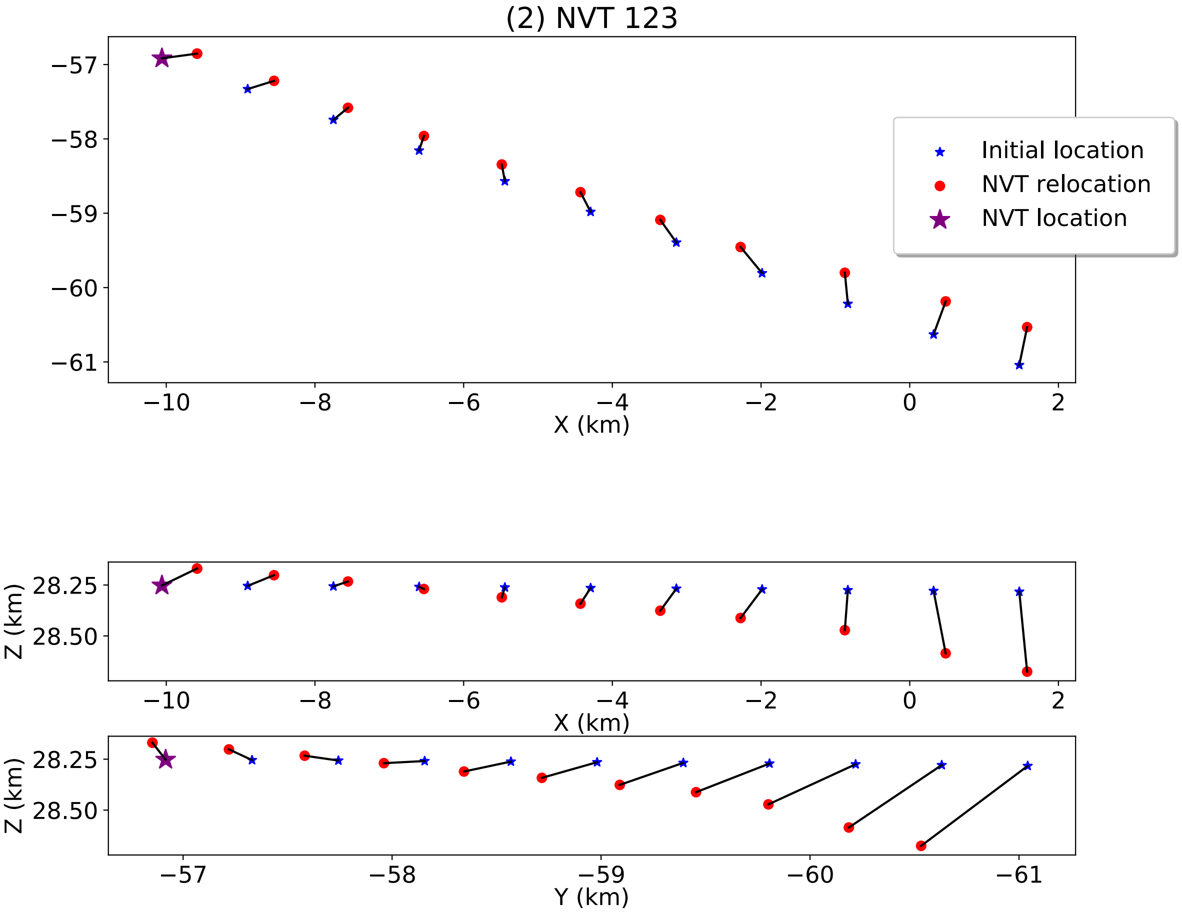
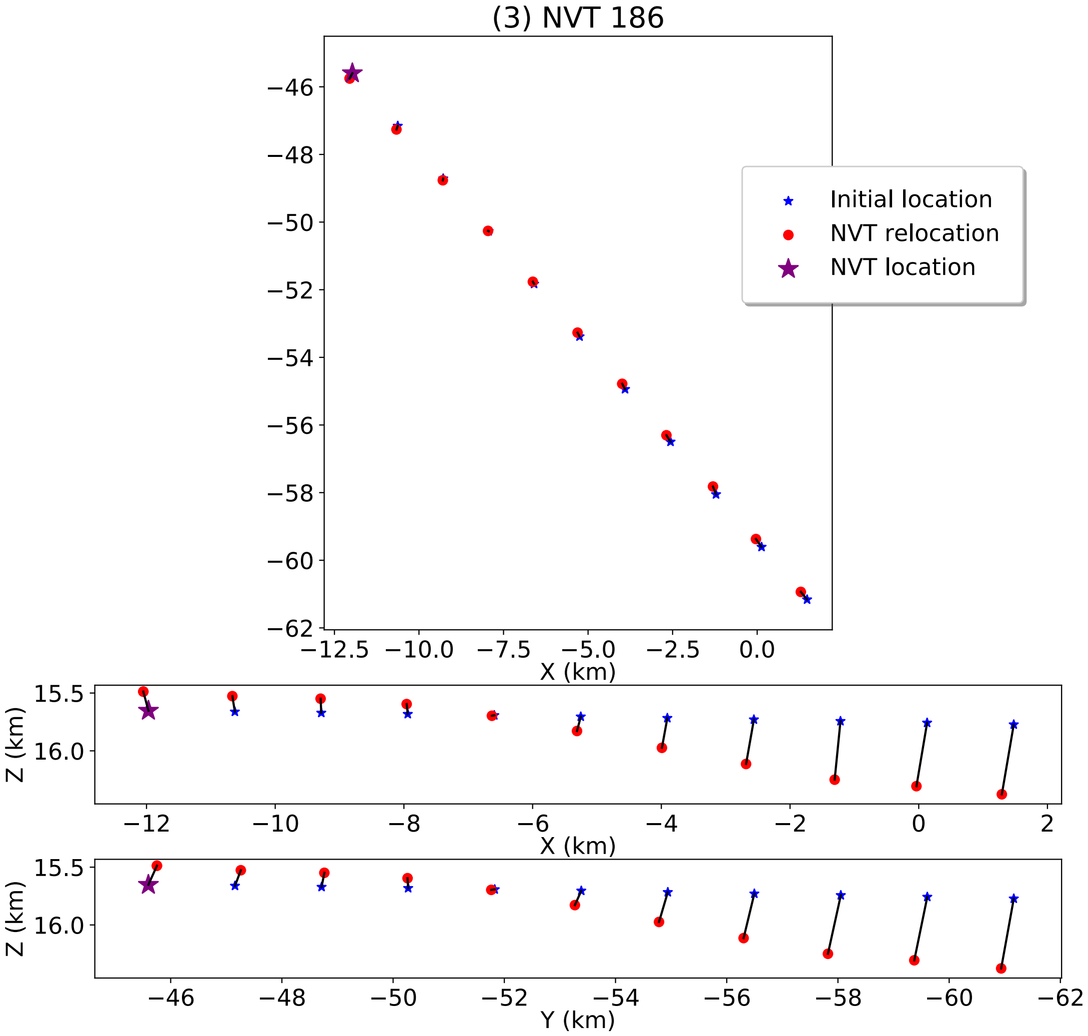
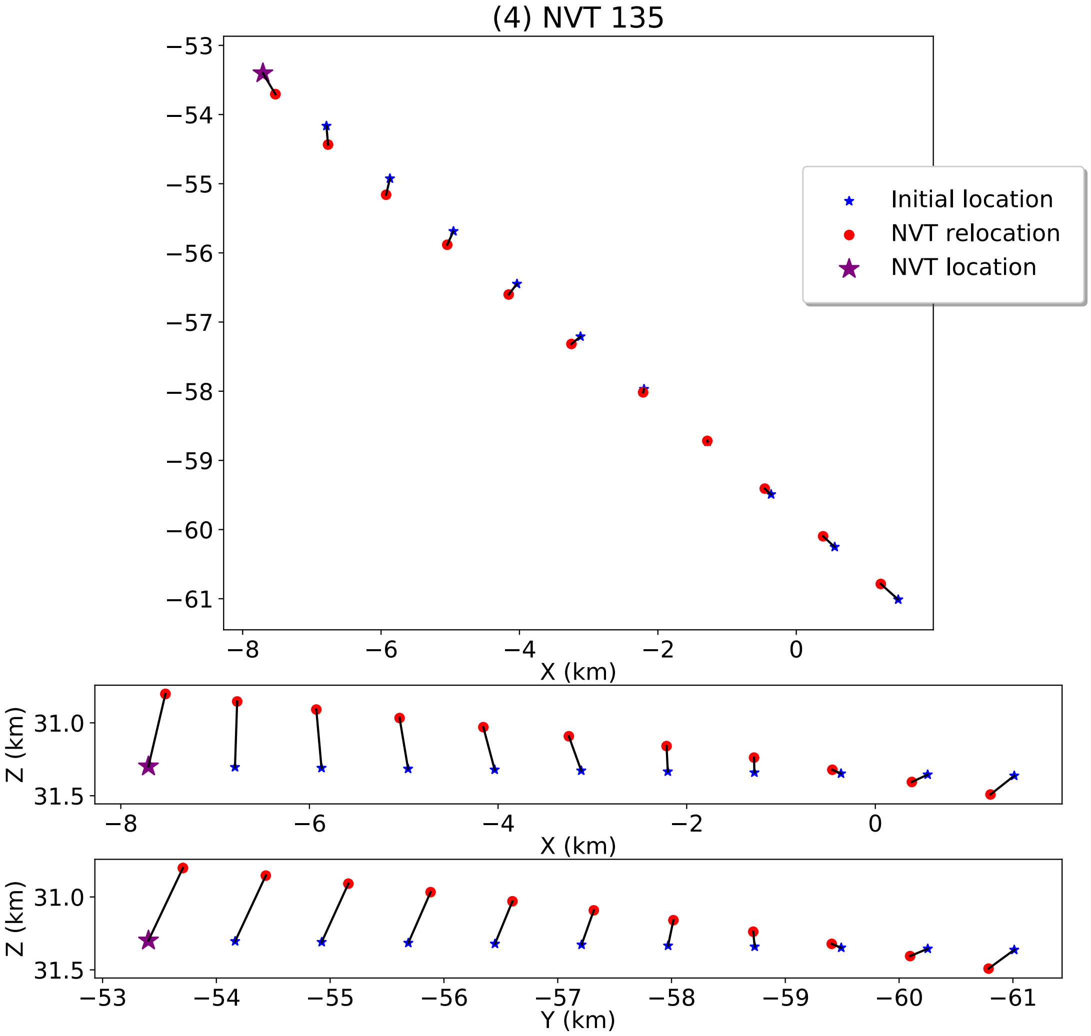


Figure 2.7 Five selected NVTs from Zhang et al. (2010). The red star is the target location. For each NVT, the 10 initial locations equally spaced along the horizontal line connecting the catalog epicenter to the target location. Blue circles are selected NVTs with the event ID next to it.


(1) NVT 76
Initial location
NVT relocation
NVT location
Y (km)
−52
−54
−56
−58
−60
X (km)
−8
−6
−4
−2
0
2
Z (km)
33.5
34.0
X (km)
Z (km)
33.5
34.0
−52
−54
−56
−58
−60
Y (km)


(2) NVT 123
−57
−58
−59
−60
−61
−10
−8
−6
−4
−2
0
2
X (km)
Initial location
NVT relocation
NVT location
Z (km)
28.25
28.50
−10
−8
−6
−4
−2
0
2
X (km)
Z (km)
28.25
28.50
−57
−58
−59
−60
−61
Y (km)


(3) NVT 186
−46
−48
−50
−52
−54
−56
−58
−60
−62
−12.5
−10.0
−7.5
−5.0
−2.5
0.0
X (km)
Initial location
NVT relocation
NVT location
Z (km)
15.5
16.0
−12
−10
−8
−6
−4
−2
0
2
X (km)
Z (km)
15.5
16.0
−46
−48
−50
−52
−54
−56
−58
−60
−62
Y (km)


(4) NVT 135
Initial location
NVT relocation
NVT location
X (km)
Y (km)
Z (km)
−53
−54
−55
−56
−57
−58
−59
−60
−61
−8
−6
−4
−2
0
31.0
31.5

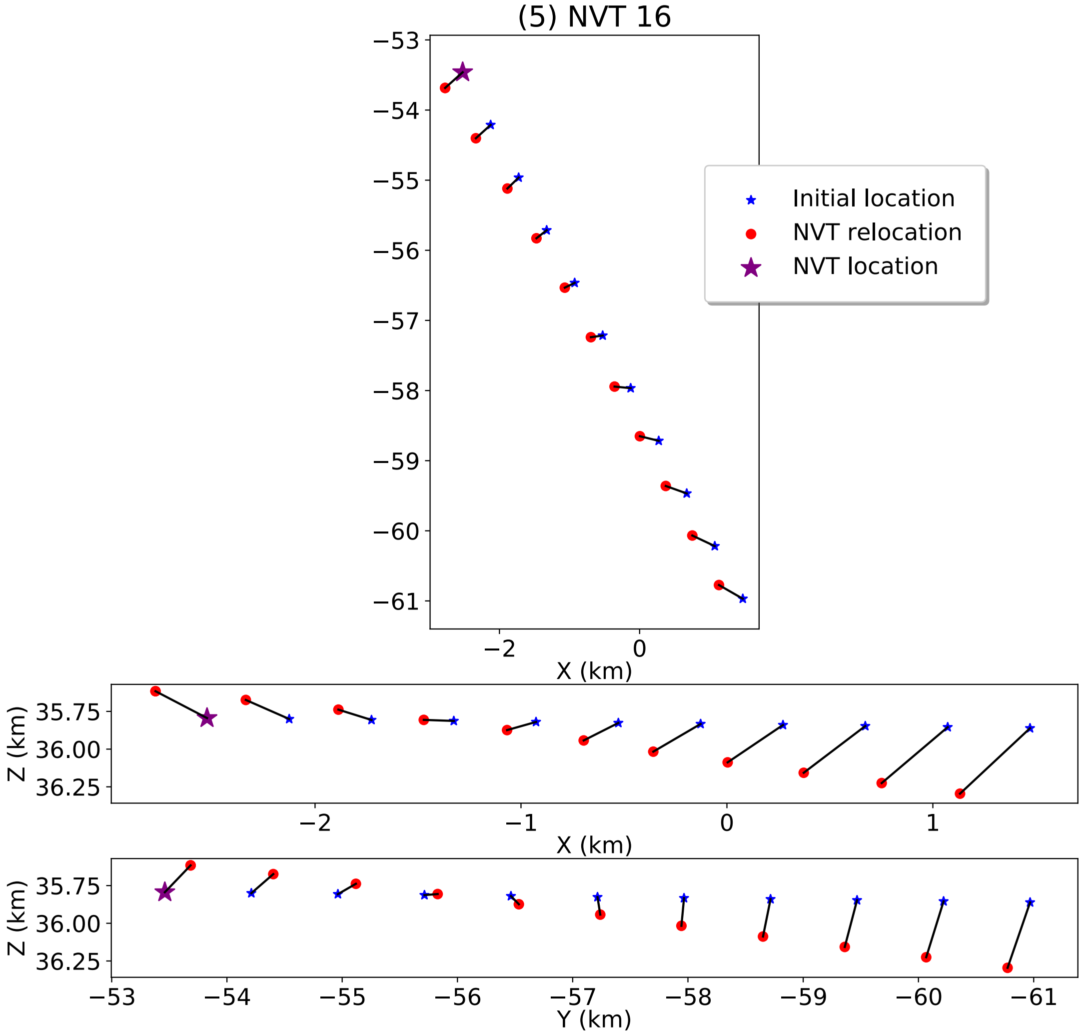


Figure 2.8 Relocation test results for five selected NVT from Zhang et al. (2010). Purple star is the catalog NVT location. The blue stars are different initial locations. Red circle is the relocation results. The relocation result is connected to its corresponding initial location with a black line. Note the differences in horizontal scale in the two cross-sections.

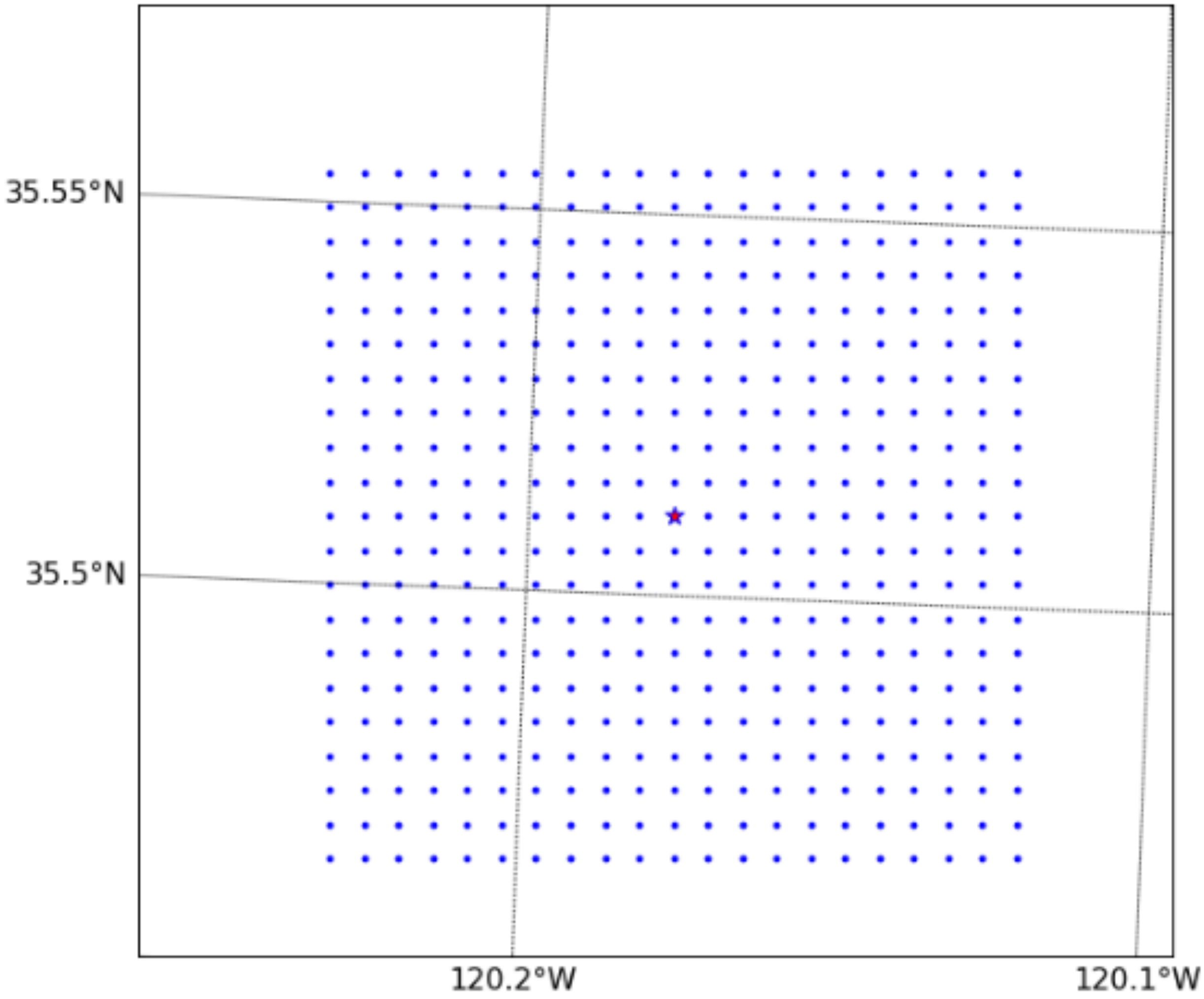
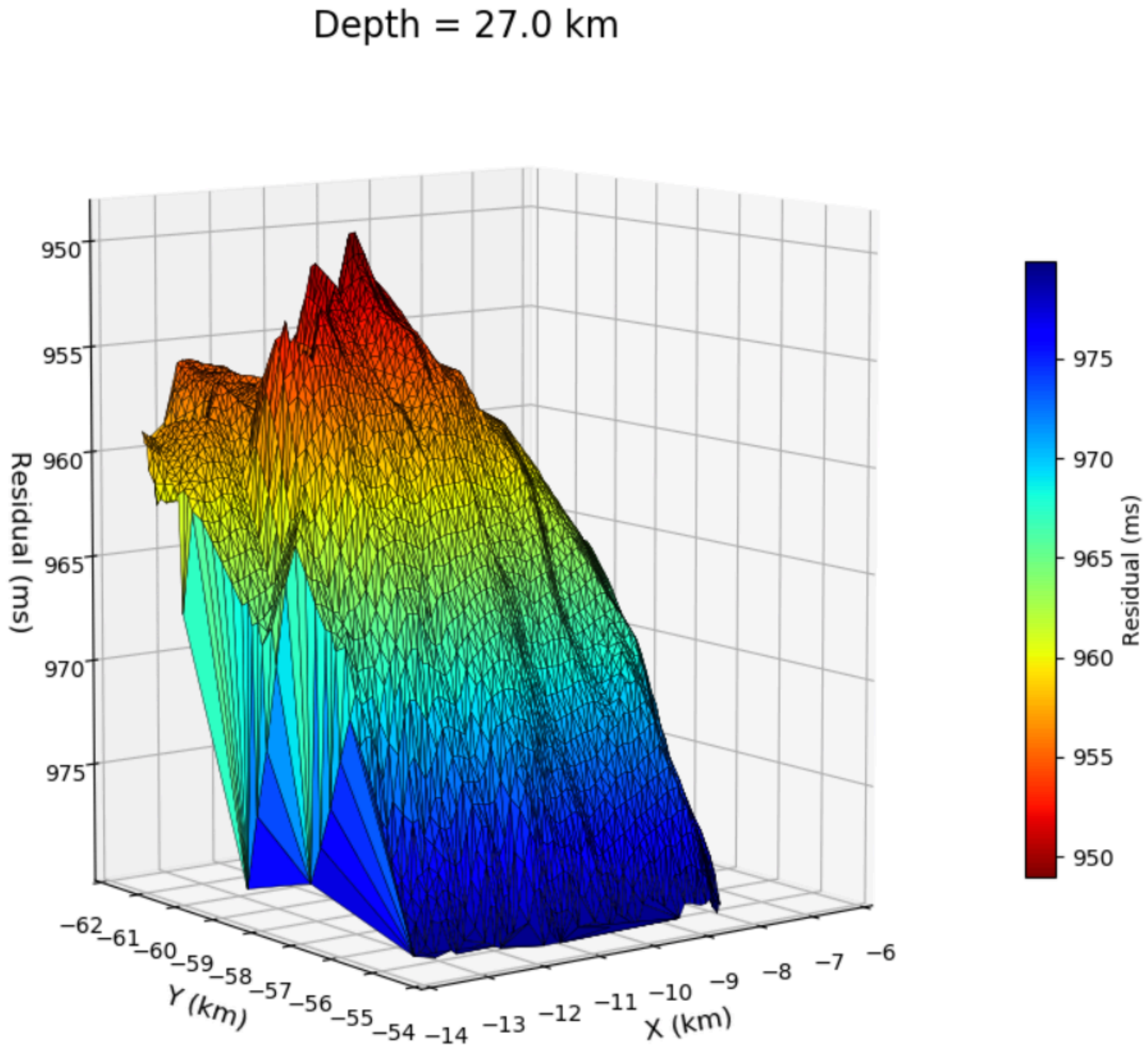
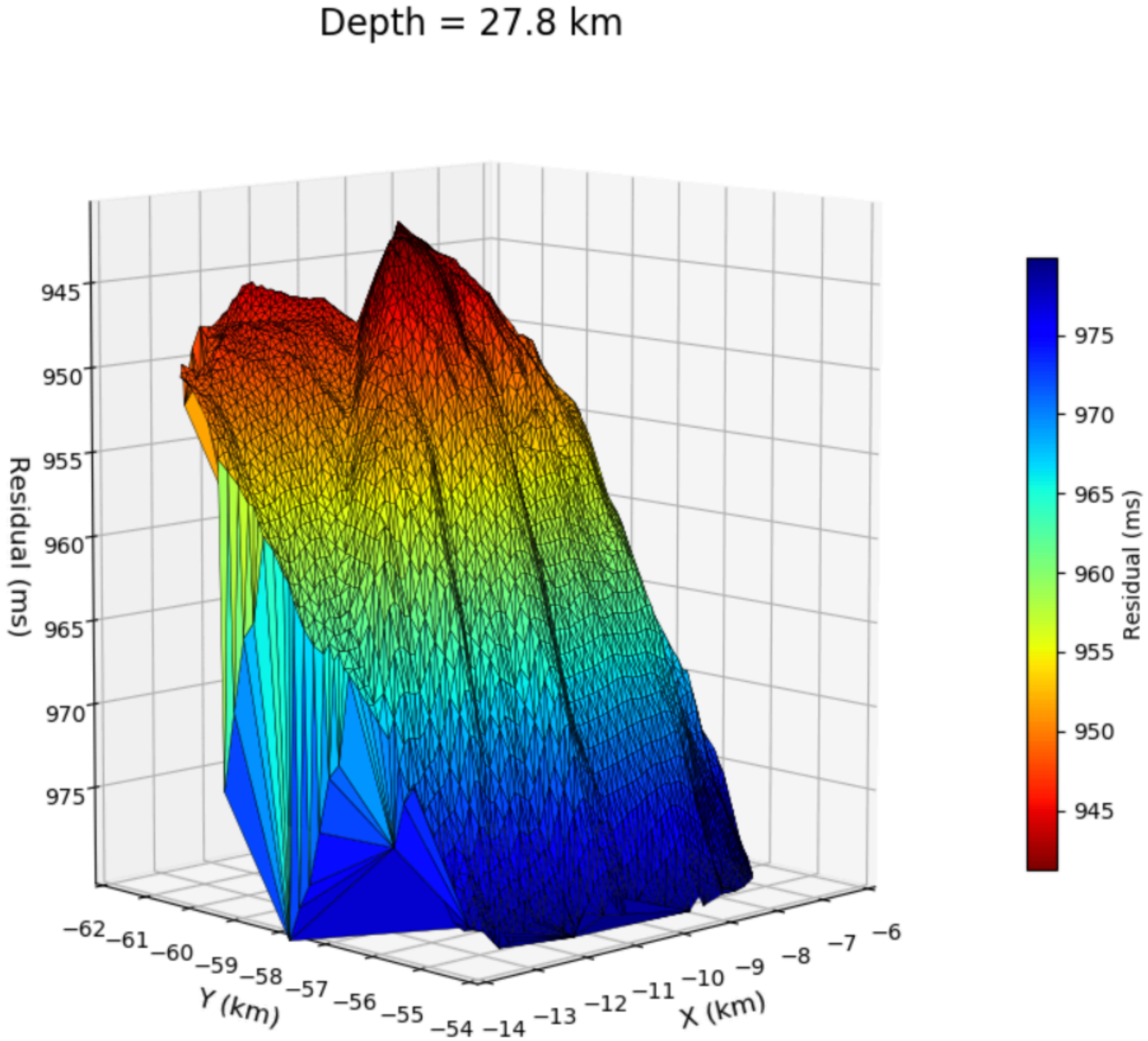


Figure 2.9 Initial locations in a horizontal plane. Red star is the catalog epicenter of event 123. Blue circles are initial locations shifted from the catalog epicenter.


Depth = 27.0 km
Residual (ms)
950
955
960
965
970
975
Y (km)
−62
−61
−60
−59
−58
−57
−56
−55
−54
X (km)
−14
−13
−12
−11
−10
−9
−8
−7
−6
Residual (ms)
975
970
965
960
955
950


Depth = 27.8 km
945
950
955
960
965
970
975
Residual (ms)
−62
−61
−60
−59
−58
−57
−56
−55
−54
Y (km)
−14
−13
−12
−11
−10
−9
−8
−7
−6
X (km)
975
970
965
960
955
950
945
Residual (ms)

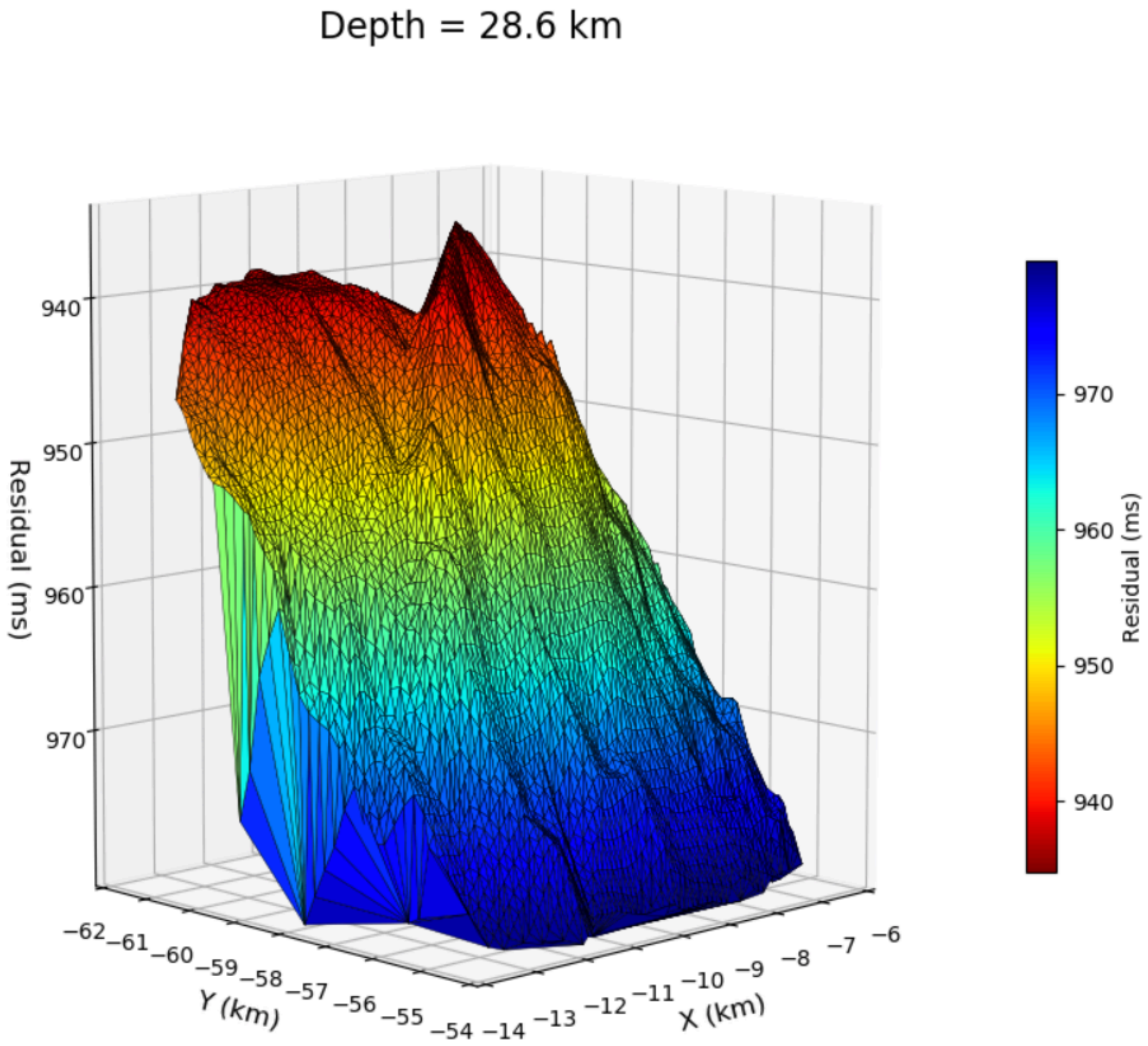


Figure 2.10 X-Y plane RMS distribution at 27, 27.8 and 28.6 km depths.

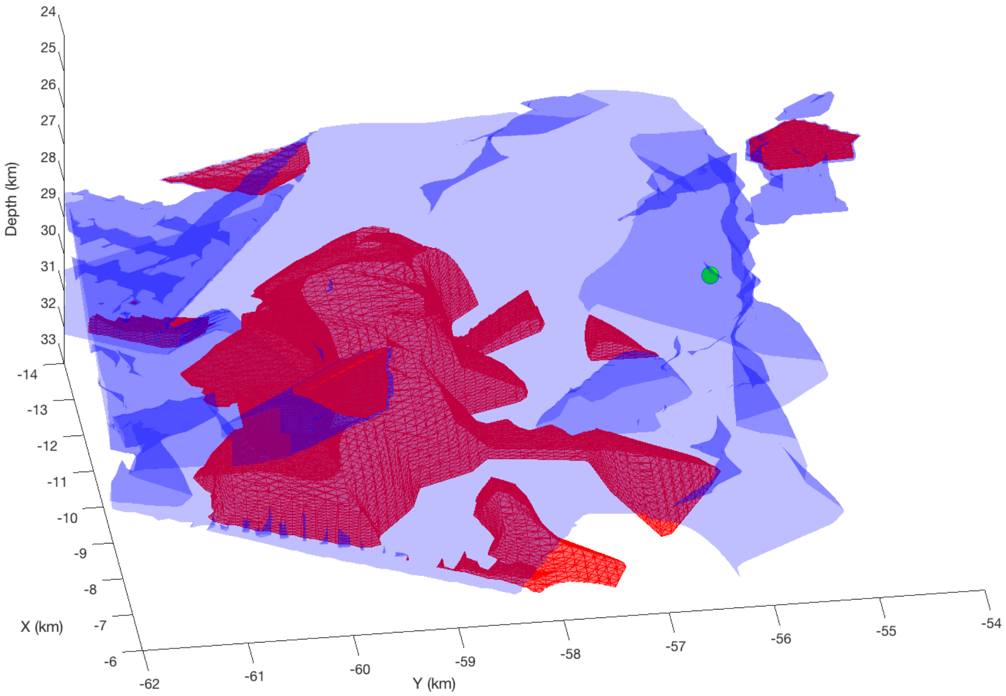


Figure 2.11 3D isosurface plot for residuals. Gray surface is the isosurface for RMS = 930 ms. Red isosurface is for RMS = 955 ms. Green circle is the catalog hypocenter for event 123.

3. Deep neural network seismic arrival time auto-picking algorithm

3.1. Abstract

Seismic phase arrival auto-picking plays an essential role in seismology studies. Application of traditional auto-pickers to different seismic datasets requires substantial effort in feature engineering (parameter tuning). In this study, I develop a seismic arrival-time auto-picking algorithm based on a deep neural network technique adopted from computer vision studies. With the deep neural network method, I develop an auto-picking algorithm that can be applied to general seismic data and extract useful features automatically. The trained neural network models can be easily maintained and updated. The training dataset is from a study along the Alpine fault, New Zealand, and its successful application to a dataset from Parkfield shows the potential portability of this neural network based auto-picking algorithm.

3.2. Introduction

In seismic studies, earthquake event arrival-time picking is essential for seismic data processing, such as hypocenter determination, seismic tomography, and earthquake early warning. To deal with the increasing volume of seismic data and public need for timely seismic hazard analysis, a variety of auto-picking algorithms have been introduced and applied to determine P- and S-wave arrival times. Traditional auto-picking algorithms identify seismic arrivals via substantial changes of single or multiple time series signal features. These signal features are calculated seismic wave characteristics in the time or frequency domain with single or multiple signal components. Some commonly used traditional algorithms are short- and long-time average ratios (STA/LTA) (Allen, 1978, Baer and Kradolfer, 1987; Earle and Shearer, 1994; Withers et al., 1998) and the Akaike information criterion (AIC) (Sleeman and Van Eck, 1999, Leonard, 2000, and St. Onge, 2010). Such auto-picking algorithms require parameter setups tailored for individual datasets, such as signal measure threshold and sliding window length. To achieve optimal auto-picking results in practice, extensive experiments are required to manually tune these parameters, which can be time consuming.

Another type of auto-picking algorithm uses artificial intelligence (AI) techniques such as nearest neighbor (Rawles and Thurber, 2015) and neural network (NN) (Dai and MacBeth, 1995; Wang and Teng, 1997; Gentili and Michelini, 2006). Such auto-picking algorithms aim to work as a self-adaptive tool, which can reduce the effort on parameter tuning. For example, in the training stage, the NN auto-picking algorithm can use different signal features as model input features (Wang and Teng, 1997) and learn the weights of each feature to properly combine them with respect to training examples. In the picking stage, the trained NN will apply these weights to incoming seismic signal features and pick the seismic arrivals once the measurement threshold is reached. However, current NN auto-picking algorithms (Dai and MacBeth, 1995; Wang and Teng, 1997; Gentili and Michelini, 2006) still requires much effort on feature engineering.

Deep learning methods can greatly reduce the effort of feature engineering for each individual dataset. By directly learning from massive training examples, a deep learning model is able to effectively learn the high-level features that can generalize well among different cases (Collobert et al., 2011; Hinton et al., 2012; Krizhevsky et al., 2012; He et al., 2016). I constructed a deep neural network (DNN) model for seismic arrival time auto-picking (DAP). The DAP model is designed to solve the seismic arrival time auto-picking problem with DNN in an end-to-end manner. Once the DNN model is trained, DAP can estimate high-quality earthquake arrival times by directly taking raw earthquake waveforms as input.

3.3. DNN background

Convolutional neural network (CNN) is a commonly used DNN model in processing two-dimensional (2D) images to perform classification and pattern recognition (Lawrence et al., 1997, Krizhevsky et al., 2012; Wei et al., 2016). In deep learning studies, 2D images are usually expressed by a 2D array of color values (coordinates are in the space domain). Such expression of 2D images is similar to the expression of seismic waveforms, which are one-dimensional (1D) arrays of amplitudes (coordinates are in the time domain). Potentially, CNN can also be applied to seismic signals with the change regarding input data dimension.

In computer vision, a specialized CNN, dilated CNN (Yu and Koltun, 2015) was designed for dense prediction, which generates pixel-wise labels for images. For each pixel in the image, the dense predicted label is the possibility of this pixel being a "positive" pixel. For instance, in portrait segmentation, which is to separate a human from the background, the positive pixels are these pixels that compose the human face. In a dilated CNN network, the convolution layers utilize dilated convolution instead of regular convolution to enlarge the sensing area of a certain image segment. In regular k by k 2D convolution, adjacent elements in a k by k matrix are used for convolution (Figure 3.1). In a k by k 2D dilated convolution with dilation rate n, k by k elements are used for convolution, with elements on the same axis that are n elements away from each other (Figure 3.1). With the enlarged receptive area of a certain convolution window, a dilated CNN network can produce dense predictions more efficiently by aggregating multi-scale contextual information (Yu and Koltun, 2015).

The seismic wave arrival auto-picking problem can also be viewed as a dense prediction problem in 1D time series data. A seismic signal can be treated as a 1D time domain amplitude "image." In this seismic "image," time-wise labels embody the possibility of a phase arrival for each time sample. For a certain type of seismic arrival, the label at the arrival time should have highest value. Once we convert the seismic arrival auto-picking problem into a dense prediction problem, the dilated CNN network in computer vision can provide us with an alternative approach for seismic arrival auto-picking.

3.4. DAP methodology

Developing a DAP picker is a challenging problem. To estimate high-quality arrival times, the picker needs to correctly extract long time dependencies from noisy input data. DAP solves this challenge in two steps: 1) DAP uses dilated convolutional layers to enlarge the receptive field of each convolution window. The enlarged reception field can help to reduce the sensitivity to high frequency noise in the input data; 2) DAP adopts the idea of multi-stage DNN structure that recursively refines its estimation. Such multi-stage structure can gradually refine the estimation by using the previous stage estimation as a kind of "attention." The "attention" leads the model to focus more on a certain segment of the input data with higher possibility of a phase arrival. Figure 3.2 from Wei et al. (2016) illustrates this idea on a T-stage pose recognition model, convolutional pose machines (CPM).

With the DAP model, seismic arrival time picking is performed in two steps: training step and picking step. In the training step, 1) the algorithm uses DNN model to generate features and produce phase-arrival predictions; 2) then the model minimizes the misfit between the predicted and true phase-arrivals for each input datum. In the picking step, the trained model is applied to incoming seismic signals to predict phase arrivals. For P- and S-wave arrivals, we trained DAP models for each type of arrival individually.

3.4.1. DAP model data processing

As shown in Figure 3.3, seismic wave arrivals can cause signal changes in both time and frequency domains. To utilize the useful information in both domains, we construct the DAP model input data to include seismic signal information in both time and frequency domains. In this way, we transform the 1D seismic waveform into a 2D seismic "image." With the 2D seismic "images" as input data, the DAP model can easily adopt DNN layers used in computer vision studies without the concern of the input dimension change.

To transform the 1D seismic signal (amplitudes in the time domain) into 2D seismic "image," we need to analyze the seismic waveform spectrum for each time sample. Following Guo et al. (2017), all raw seismic signals are first resampled to 100 Hz. To filter out the background noise in the frequency domain, we only included phase amplitude between 3 and 35 Hz, which was used by Guo et al. (2017) in their auto-picking process. In the spectral analysis, we used a 32-sample sliding window to measure the phase amplitude at different frequencies for each time spot. With a 100 Hz sample rate, a 32-sample sliding window length samples 10 phase amplitudes in the 3-35 Hz range. Then the time domain amplitudes, spectrum phase amplitudes and spectrum phases are concatenated along the time axis (X-axis) to compose the input 2D array for the DAP model (Figure 3.4). For P-arrival picking, only the vertical (Z) component is used. The constructed 2D input array for P-wave arrival picking is an n by 21 matrix, where n is the number of samples in the original waveform. For S-wave arrival picking, both north (N) and east (E) components of the seismic signals are used. As shown in Figure 3.4, we first concatenated the time domain amplitudes of the N and E components along the X-axis. Then the spectral information of the N and E components is concatenated along the X-axis as well. The constructed 2D input array for S-wave arrival picking is an n by 42 matrix, where n is the number of samples in the original waveform.

Since we define the DAP as a dense prediction problem for 1D time series data, it is necessary to label every sample in the input seismic signal from the training dataset:

$$\mathrm{Y}_{\mathrm{true}} = (\mathrm{y}_1, \dots, \mathrm{y}_\mathrm{n}) \tag{3.1}$$

where n is the number of samples in the input signal, and y_i is the label for the ith sample. At first, time samples with phase arrival picks are labeled as 1 whereas the rest are labeled as 0. However, there are uncertainties in the manually picked phase arrivals. In order to accommodate the human picking uncertainties, assumed as 0.1 s, the label values for samples around the phase arrival are adjusted using a Gaussian function with peak 1 and a standard deviation of 10 samples (based on the 100 Hz sample rate) (Figure 3.5). We term the labels $\mathrm{Y}_{\mathrm{true}}$ for the input seismic signal as a heat map. The heat map represents a probability mass function for the phase arrival in the time domain, with its peak indicating the manually picked phase arrival.

3.4.2. DAP model structure

DAP implements a three-stage DNN structure (Figure 3.6). Each stage performs as a predictor $\mathrm{g_s}(X_s);\ \mathrm{s} = 1, 2, 3$, which is trained to predict the heat map $\mathrm{Y_s} = (y_{s1}, \dots, y_{sn})$ of the phase arrival. In the predicted heat map $\mathrm{Y_s}$, $\mathrm{y_{si}}$ is the possibility of a phase arrival at the ith sample. In stage s, the predictor works in two steps:

$$\mathrm{g_s}(X_s) = p_s\big(f_s(X_s)\big) \tag{3.2}$$

where $\mathrm{f_s}$ is a feature extractor composed of five dilated convolution layers to extract the contextual features from input X_s, and $\mathrm{p_s}$ is the predictor with a single dilated convolution layer to produce the heat map $\mathrm{Y_s}$. In the DAP model, all convolution layers are 1D as the phase arrival auto-picking problem only considers the time domain. The number of kernels, small matrices used to convolute with the input data, for each convolution layer is 64. All these kernels have length of 4 (in the time axis) and dilation rate r. For traditional convolution, a 4-sample kernel can only extract information from a 0.04 s segment of data for signals with 100 Hz sample rate. The traditional convolution can be viewed as a dilated convolution with dilation rate 1. In dilated convolution with dilation rate r, the 4-sample kernel can extract information from a $\left(0.04 + \frac{3*(\mathrm{r}-1)}{100}\right) s$ segment for 100 Hz sampled signals. With dilated convolution, the sensing length of the input signal increases without increasing the kernel length. Therefore, the feature extractor f_i can extract contextual information more efficiently and be more resistant to high frequency noise in the input data (Yu and Koltun, 2015.) For the P- and S-arrival models, dilation rates are selected separately and the selections are discussed in later sections.

The first stage produces the contextual features $f_1(X_1)$ extracted from the input seismic 2D "image" X_1 and the heat map of phase arrival Y_1 (Figure 3.6). In later stage s $\in \{2,3\}$, the input X_s is composed of two parts: 1) extracted contextual features $f_1(X_1)$ from input seismic "image"; 2) heat map Y_{s-1} from the preceding stage. To keep the original input contextual features for later stages, $f_1(X_1)$ is used for all later stage inputs as they are computed purely based on the input 2D seismic "image" without any previous heat map. In stage s, g_s uses the previous stage heat map Y_{s-1} as attention attractor, which leads the phase arrival search to focus more on the time samples with higher possibility, to refine the possibilities of phase arrival for every time sample and produces a new heat map Y_s.

3.4.3. Learning in DAP

Within each stage, there are five dilated convolutional layers and one output layer. By using a three-stage DNN structure, DAP results in a deep architecture containing 18 layers. When training a network with multiple layers via the gradient descent method, the process is likely to encounter the vanishing gradients problem (Bengio et al., 1994; Bradley, 2010; Glorot and Bengio, 2010). In the gradient descent method, the loss function is used to measure the differences between prediction and truth. The gradients of the loss function will be calculated and back propagated from the output to the input layer. The parameters in each layer receive updates proportional to the gradients of the loss function. As observed in previous back-propagation studies, the magnitude of the layer parameter updates decreases as the number of layers increases. When there are too many intermediate layers between the input and output layers, the former layers will not learn anything as the loss function gradients become too small when back-propagated.

To address the vanishing gradients problem, we compose the final loss function with the losses from all three stages instead of only the loss from the last stage. By combing losses from all three stages, the gradients are enlarged, and the risk of the vanishing gradient problem is reduced. Besides, the first two stages provide direct contributions, instead of indirect contributions via the last stage, to the final loss function. Utilizing losses from all three stages enables a more efficient training process for the model (Wei et al., 2016). For model training, the loss function within each stage s is defined so as to minimize the mean squared distance between the predicted heat map Y_s and the true heat map Y_{true} for training data. Thus, the loss function to be minimized for each stage is given by:

$$l_s = \frac{1}{n}\sum_{i=1}^{n}(y_{is} - y_i) \tag{3.3}$$

Then, the overall cost function for all three stages is calculated by adding the losses at each stage:

$$L = l_1 + l_2 + l_3 \tag{3.4}$$

The optimizer RMSProp (Tieleman and Hinton, 2012) is used to train the model.

3.4.4. Phase-arrival picking

We slice the seismic signals into 512-sample segments, where multiple corresponding heat map segments are generated by the DAP model for each seismic signal. Then we reconstruct the final heat map for each seismic signal in two steps: 1) align these heat map segments along the time axis; 2) for overlapping time samples between two different segments, the larger heat map values are selected as the final predicted values. With the final heat map for each input signal, the phase arrival is chosen as the heat map peak position, which represents the most likely phase arrival time sample. For model performance evaluation, the accuracies are defined as PC1 and PC2 for the percentages of predicted phase-arrivals within 0.1 s and 0.2 s of the manually picked arrivals, respectively.

3.5. DAP for P-arrival picking

Manually picked seismic wave arrival data for regional earthquakes are available from a tomography study along the central Alpine Fault (AF), New Zealand (Guo et al., 2017). To train and evaluate our DAP model, we selected high quality manually picked waveforms as our training and validation datasets. In total, 4985 seismic traces containing P-arrivals are selected to develop the DAP model for P-wave arrival time picking.

3.5.1. Data processing

As discussed in section 3.4.1, the Z component signal for P-wave arrival time picking is transformed into a 2D array with dimension of n by 21, where n is the number of samples in the signal (Figure 3.4). The true heat maps for each 2D array are then generated (Figure 3.5). For each signal, we used a 512-sample sliding window with 256-sample step size to generate 2D arrays and heat map segments in window (b, e), with b 3 s before P-arrival and e one travel time after P arrival. The travel time is the time difference between event origin time and the picked P arrival. Within this window (b, e), background noise before the P arrival, P-wave coda, and S waves are all included to serve as negative samples (no P arrival present) in contrast to positive samples near the P arrival. These 2D arrays with dimension 512 by 21 are the input data for our DAP model. In each time segment, the positive samples are labeled with values greater than 0 whereas the negative samples are labeled 0. For each input signal, we only generate 2 segments containing positive samples. To increase the number of segments with positive samples, we shrink the step size to 64 samples in between 256 samples before and after the P-arrival. By this means, 8 segments containing positive samples are generated for each input signal. The positions of the P arrival in these 8 segments are different, which mimics the reality that we do not know where the P arrival is within each segment when picking new data. The (2D array, heat map) pairs are the instances used for later DAP model development. The number of instances generated for the P arrival DAP model is 45,945.

3.5.2. Epoch and dilation rate selection

The P-arrival dataset is split into a training set and validation set with a ratio of 4:1. For parameter selections, the model will be trained with the training set and the validation set will be used to evaluate the model performance.

One complete presentation of the training data to the DNN model is defined as one epoch. A proper selection of the total number of training epochs will ensure that the model has learned enough from the training data but has not been overly tailored to only perform well on the training data. To select the number of training epochs, we construct a DAP model with dilation rate of 3 and trained it with the training set for 300 epochs. As shown in Figure 3.7, the loss decreases rapidly in the first 50 epochs. Around epoch 250, the loss started to decrease slowly with many oscillations, so we selected 250 as the number of training epochs. In addition, the loss decreases from stage 1 to 3, which means the three-stage DNN structure does successfully refine the predicted heat maps.

For dilation rate selection, we tested dilation rates of 1, 2, 3, 4, 5, and 6. For each dilation rate, we trained the DAP model with the training set for 250 epochs. Then we examined the accuracy of the DAP model prediction for the validation set. As shown in Table 3.1, the DAP model with dilation rate of 5 performs the best in accuracy for both PC1 (92.5%) and PC2 (93.0%).

3.5.3. Model evaluation with cross-validation

In order to better evaluate the DAP model performance and avoid overfitting. We perform a five-fold cross-validation (Aster et al., 2003) with the entire dataset. The dataset is split into 5 subsets with each subset containing ~1000 seismic traces. In each cross-validation step, 4 subsets are used for training and the remaining 1 subset is used for evaluation. The average accuracies for predicted P arrivals PC1 and PC2 are 86.2% and 90.6%, respectively. Based on the cross-validation results, the DAP model for P-arrival picking shows stable performance with about 5000 station-event pair traces for training.

3.6. DAP for S-arrival picking

Both horizontal components (E and N) of seismic signals are used for the S-arrival DAP model. From the manually picked S-wave dataset in Guo et al. (2017), we selected 1,554 traces for DAP model development.

3.6.1. Data processing

Following the same data processing steps as shown in 3.4.1, we constructed 2D arrays for the E and N components of seismic signals (Figure 3.4). As both horizontal components are useful for S-arrival picking, we combined the 2D arrays of the E and N components. To combine these 2D arrays of E and N components for the same event-station pair, we put amplitudes in the time domain in the first two rows (E, N order), followed by the spectral information (E, N order). The final combined 2D array for E, N components has a dimension of (n, 42) where n is the total number of samples in the seismic signal (Figure 3.4).

A 512-sample sliding window is used to generate input segments for the DAP model. Note that we assume there are already picked P-arrivals when we proceed to pick the S arrivals. For each event-station pair, the signal within time window (b_s, e_s) is used to generate the input and heat map segments, where b_s is the midpoint between the P and S arrivals and e_s is 3 s after the S arrival. A 64-sample step size is used between the first and last window containing S arrivals. A step size of 256 samples is used for other window periods that do not contain S arrivals. In total, 14,720 (2D array, heat map) pairs are generated for S-arrival DAP model development.

3.6.2. Dilation rate selection

The S-arrival dataset is split into a training set and validation set with a ratio of 4:1. We trained the DAP models with dilation rates of 1, 2, 3, 4, 5, 6, 7, and 8 individually on the training set with epoch number of 250. Then the accuracies of PC1 and PC2 on S arrivals are evaluated with the validation set. As shown in Table 3.2, a dilation rate of 6 provides the best results for PC1 and PC2 for the validation set. For about 360 event-station pairs, PC1 is 72.1% and PC2 is 81.3%.

3.6.3. Model evaluation with cross-validation

Five-fold cross-validation is carried out to evaluate the DAP model performance for S waves with the same approach as in section 3.5.3. The average accuracies for predicted S arrivals PC1 and PC2 are 73.5% and 82.1%, respectively. Based on the cross-validation results, the DAP model for S-arrival picking shows stable performance with about 1600 station-event pair traces.

3.7. Test DAP on Parkfield dataset

With the trained DAP picker for P and S arrival time picking, a performance test is conducted on an independent set of seismic waveforms. Test set seismic waveforms are selected from events used in a Parkfield tomography study (Zeng et al., 2014). Similar to the training dataset from the AF tomography study, the events from the Parkfield dataset are also local to regional earthquakes that occurred near fault zones. 104 station-event pairs with high quality manually picked P and S arrivals are selected as the test set for the DAP model.

The data processing procedures for both P- and S-wave signals are the same as described previously. The test results for the P arrival are 96% and 99% for PC1 and PC2, respectively. For the S arrival, PC1 and PC2 are 82% and 93%, respectively. As shown in Figure 3.8, the predicted heat maps for both P and S arrivals preserve the envelope shape of true heat maps, even though there are oscillations near the true phase arrivals. As there is less background noise before the P arrival than the S arrival, the predicted P arrival heat maps usually have less oscillations than the S-arrival heat maps.

3.8. Discussion

By transforming the 1D seismic time series signals into 2D “images”, we are able to apply a deep learning model used in computer vision to provide an alternative solution for the phase-arrival auto-picking problem. With available event-station pair traces from manually picked data, DAP models for P and S arrivals are trained separately and evaluated with cross-validation. Due to limited available S-arrival data, the number of training data used for the S-arrival DAP model is only about 30% of that used for the P-arrival DAP model training. The average cross-validation results (PC1, PC2) for P- and S-arrival DAP models are (86.2%, 90.6%) and (72.1%, 81.3%), respectively. The lower PC1 and PC2 for S-arrival DAP model can be caused by 1) limited training data; 2) more complicated pre-S-arrival background noise than pre-P-arrival noise; 3) larger human label uncertainties than for the P-arrival label. The first reason can be solved by adding in more well-labeled S-arrival data into the training dataset. With an enlarged training dataset, the model can be generalized to handle the complicated background noise and approach human picking accuracy.

With the 104 test station-event pair traces, both of the DAP models result in higher accuracies (PC1 and PC2) compared to the cross-validation results. Cleaner waveforms in the test dataset than in the training dataset is the main reason for these higher accuracies. For the P arrival, the test PC1 and PC2 are 95% and 99%, respectively. For the S arrival, PC1 and PC2 are 82% and 93%, respectively. These results show the potential of our DAP model to handle both P- and S-wave signals. Moreover, with the training dataset from the AF study in New Zealand, the DAP models work well on earthquake signals in a SAF study. The character of the waveforms used in these two datasets are different, although the events are all local or regional earthquakes that occurred near fault zones. This is strong evidence regarding the capability of our DAP models to handle similar earthquake signals with no major location dependencies.

Compared to traditional and previous AI auto picking algorithms, DAP aims to perform high accuracy P- and S-wave arrival auto picking in an end-to-end manner with a simple user experience. Other than saving human effort on tuning auto-picking algorithm parameters, the trained models for each type of seismic arrival can be maintained, updated, and improved by training with more future training data.

3.9. Acknowledgements

This research is supported in part by the USGS, Department of the Interior, under USGS Award Number G14AP00056 to the University of Wisconsin-Madison. The support of a student assistantship from the Mark and Carol Ann Solien Graduate Assistantship and George P. Woollard-Sigmund I. Hammer Memorial Fund in Geology & Geophysics of the University of Wisconsin-Madison is also acknowledged. We thank Shuochao Yao and Yu Shi at the University of Illinois for their theoretical and technical support.

3.10. References

Allen, R., 1978, Automatic earthquake recognition and timing from single traces. Bull. Seismol. Soc. Am., 68, 1,521–1,532.

Baer, M., and Kradolfer, U., 1987, An automatic phase picker for local and teleseismic events. Bull. Seismol. Soc. Am., 77(4), 1,437-1,445.

Bengio, Y., Simard, P., and Frasconi. P., 1994, Learning long-term dependencies with gradient descent is difficult. IEEE Transactions on Neural Networks.

Bradley, D., 2010, Learning In Modular Systems. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA., 2010.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P., 2011, Natural language processing (almost) from scratch. Journal of Machine Learning Research 12: 2,493-2,537.

Dai, H., and MacBeth, C. 1995, Automatic picking of seismic arrivals in local earthquake data using an artificial neural network. Geophys. J. Int. 120(3), 758-774.

Earle, P. S., and Shearer, P. M., 1994, Characterization of global seismograms using an automatic-picking algorithm. Bull. Seismol. Soc. Am., 84(2), 366-376.

Gentili, S., and Michelini, A., 2006, Automatic picking of P and S phases using a neural tree. J. Seismol., 10, 39–63, doi:10.1007/ s10950-006-2296-6.

Glorot, X. and Bengio, Y., 2010, Understanding the difficulty of training deep feedforward neural networks. Proceedings of the thirteenth international conference on artificial intelligence and statistics, pp. 249-256.

Guo, B., Thurber, C.H., Roecker, S.W., Townend, J., Rawles, C., Chamberlain, C.J., Boese, C.M., Bannister, S., Feenstra, J. and Eccles, J.D., Guo, B., Thurber, C. H., Roecker, S. W., Townend, J., Rawles, C., Chamberlain, C. J., ... and Eccles, J. D. 2017, 3-D P-and S-wave velocity structure along the central Alpine Fault, South Island, New Zealand. Geophys. J. Int., 209(2), 935-947.

He, K., Zhang, X., Ren, S., and Sun, J., 2016, Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N. and Kingsbury, B., 2012, Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine 29.6, 82-97.

Krizhevsky, A., Sutskever, I., and Hinton, E., G., 2012, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, pp. 1,097-1,105.

Lawrence, S., Giles, C. L., Tsoi, A. C., and Back, A. D., 1997, Face recognition: A convolutional neural-network approach. IEEE transactions on neural networks, 8(1), 98-113.

Li, Y., 2018, Dilated Convolutions and Kronecker Factored Convolutions. https://medium.com/@lisulimowicz/dilated-convolutions-and-kronecker-factored-convolutions-b42ed58b2bc7

Rawles, C., and Thurber, C., 2015, A non-parametric method for automatic determination of P-wave and S-wave arrival times: application to local micro earthquakes. Geophys. J. Int., 202(2), 1,164–1,179.

Wang, J., and Teng, T. L., 1997, Identification and picking of S phase using an artificial neural network. Bull. Seismol. Soc. Am., 87(5), 1,140-1,149.

Withers, M., Aster, R., Young, C., Beiriger, J., Harris, M., Moore, S., and Trujillo, J., 1998, A comparison of select trigger algorithms for automated global seismic phase and event detection. Bull. Seismol. Soc. Am., 88(1), 95-106.

Yu, F., and Koltun, V., 2015, Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122.

3.11. Tables and Figures

Table 3.1 Validation set accuracies for different dilation rates in the P-arrival DAP model

Dilation rate	1	2	3	4	5	6
PC1	89.0%	90.8%	87.1%	92.1%	92.5%	91.9%
PC2	89.4%	91.1%	87.6%	92.2%	93.0%	92.2%

Table 3.2 Validation set accuracies for different dilation rates in the S-arrival DAP model

Dilation rate	1	2	3	4	5	6	7	8
PC1	45.7%	62.6%	64.9%	67.7%	70.2%	72.1%	69.4%	2.3%
PC2	64.3%	71.9%	72.4%	76.9%	77..2%	81.3%	80.7%	5.0%

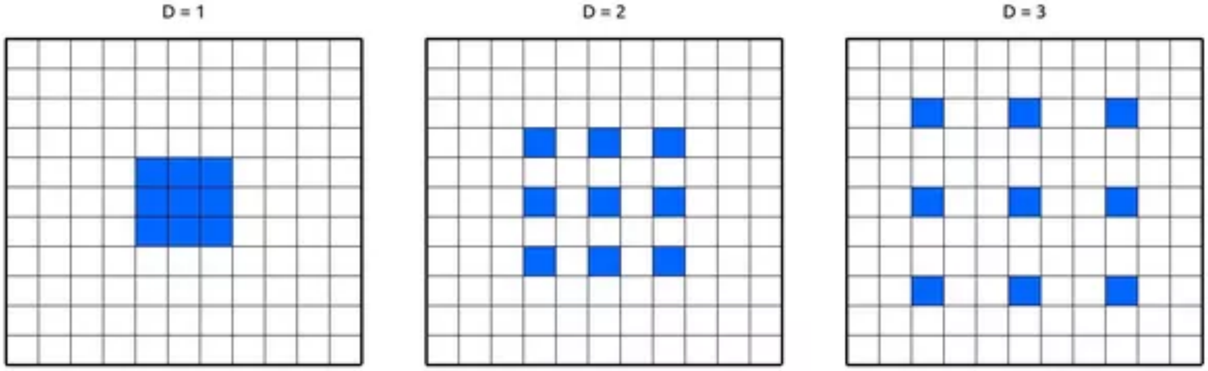


Figure 3.1 Illustration of dilated 3 by 3 convolution (Figure by Li, 2018). Elements marked in blue are elements used for convolution. Left, 1-dilated convolution. Middle, 2-dilated convolution. Right, 3-dilated convolution.

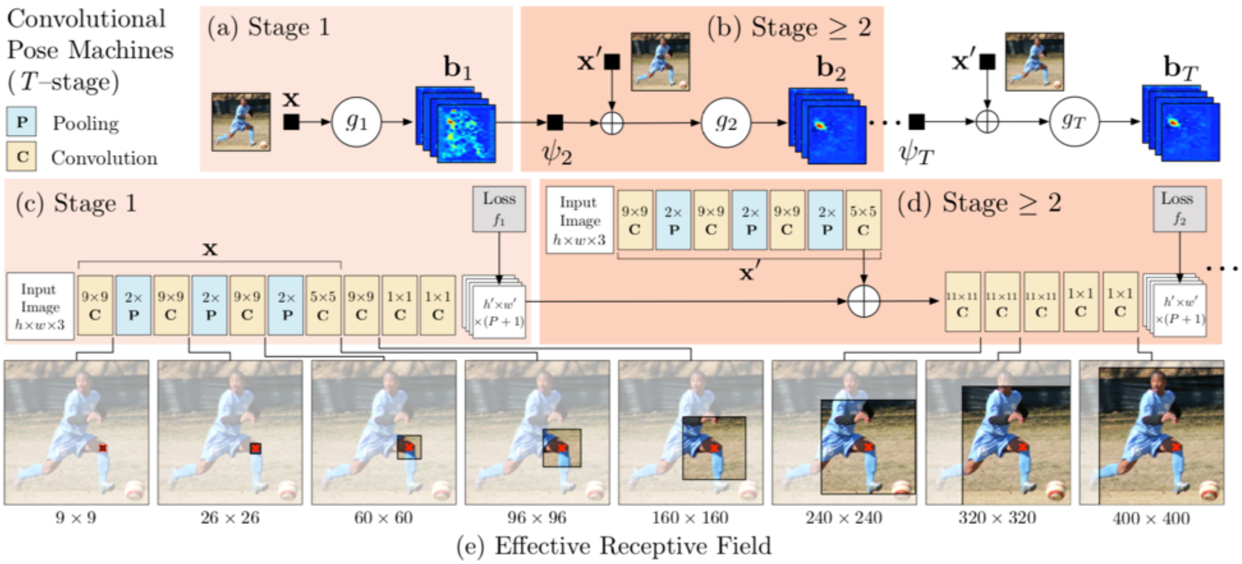


Figure 3.2 CPM illustration from Wei et al. (2016). These plots show the convolution network architecture and receptive fields across layers with the number of T stages. The pose machine and the attention layers are shown in (a) and (b), and the corresponding network structures are shown in (c) and (d). (a) and (c) show the first stage operations on the image. (b) and (d) show the operations on the following second to last stages. (e) show the corresponding receptive fields for the dilated convolution layers on the image

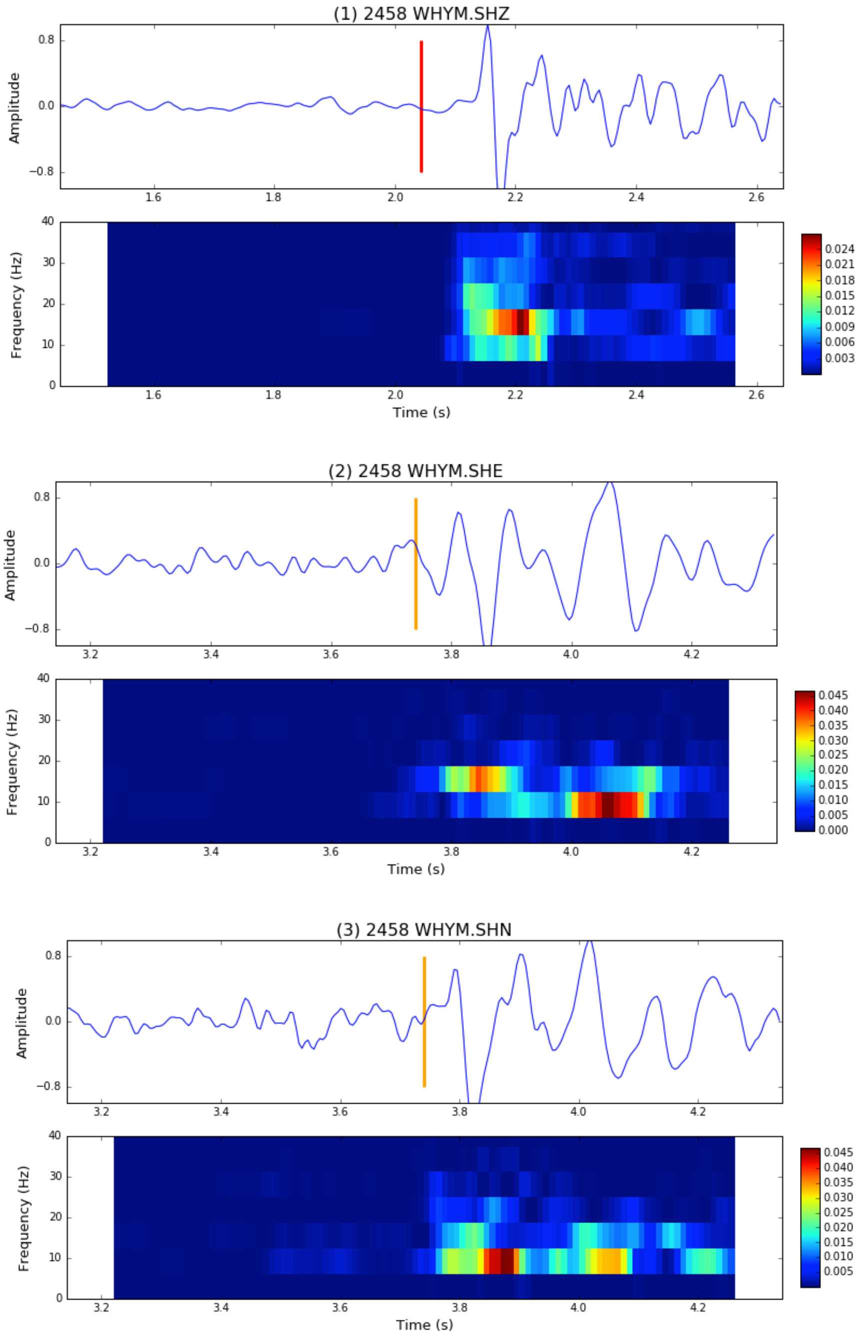


Figure 3.3 Seismic signal in the time and frequency domains for event 2458 on station WHYM. In each plot, the first row shows the seismic amplitude in the time domain, the second row shows the spectrum information (spectrogram) with the signal amplitude indicated by colors. (1) is the vertical component signal, the red line shows the P-wave arrival. (2), (3) are the east and north component signals, respectively, with an orange line indicating the S-wave arrival.

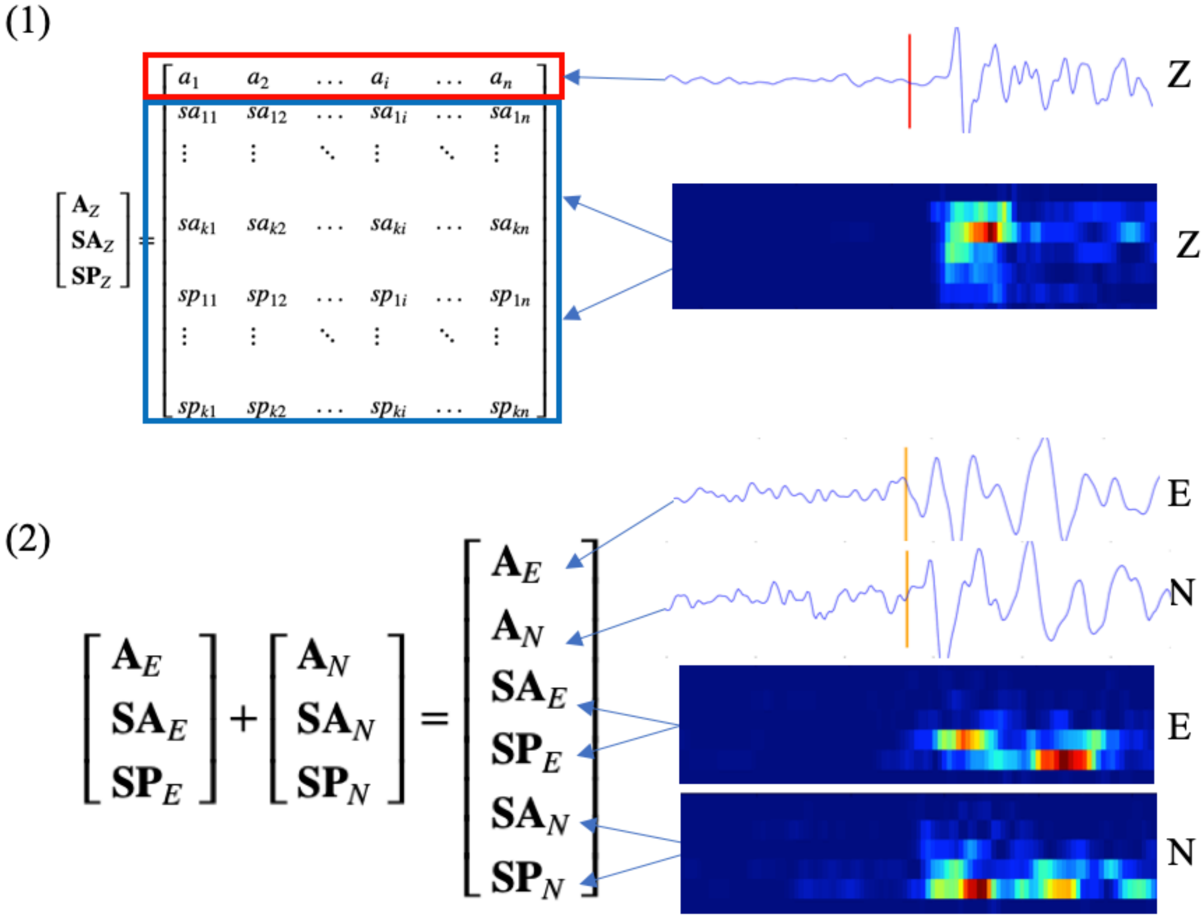


Figure 3.4 Transformation from 1D seismic signals into 2D seismic images. A_c, SA_c and SP_c are the matrices for time domain amplitude, spectral amplitude and spectral phase for $c \in \{Z, E, N\}$ component, respectively. The X-axis for these matrices is the time axis, $k \in \{1,2,3, \dots ,10\}$ is the index of each analyzed frequency. (1) For the P arrival, the Z component is used. The final 2D image dimension is n by 21, with n as the number of samples in the signal. (2) For the S arrival, the E and N components are used. The final 2D image dimension is n by 42.

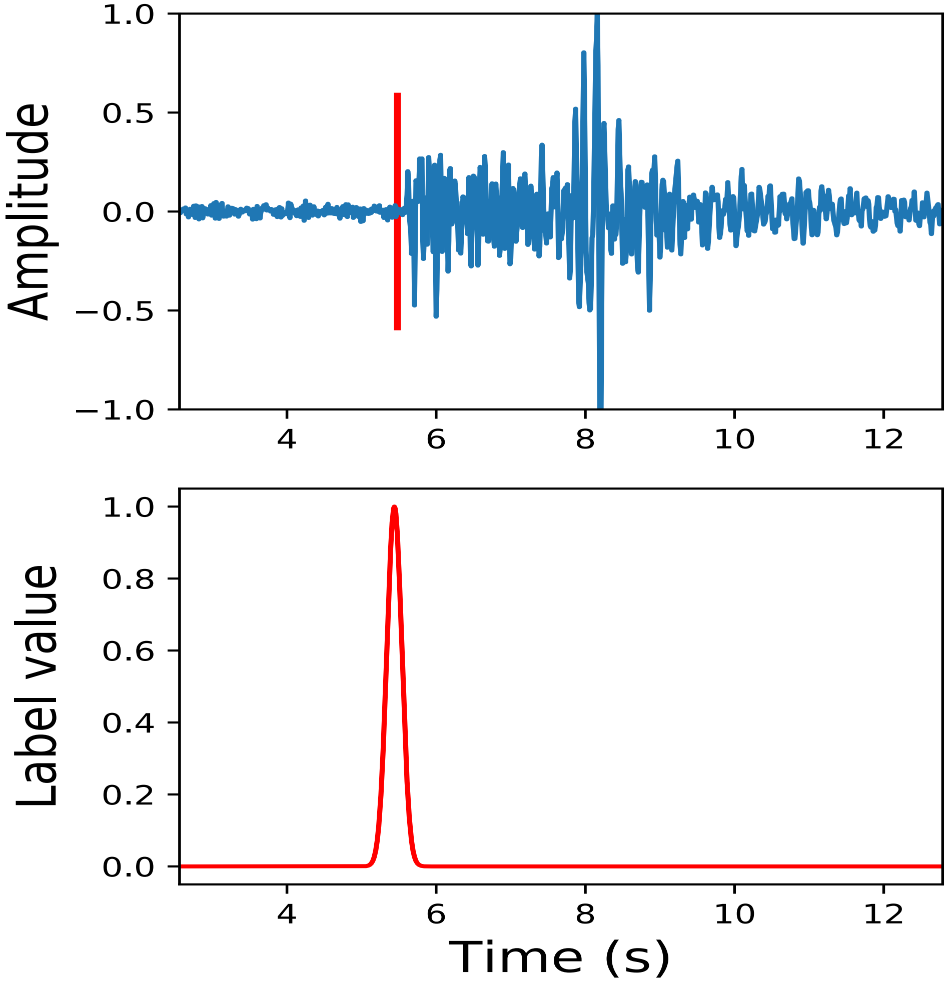


Figure 3.5 Heat map plot. In the first panel, the blue trace is original seismic waveform; the red line shows the manually picked phase-arrival. The second panel shows the true heat map generated for the seismic trace in the first panel.

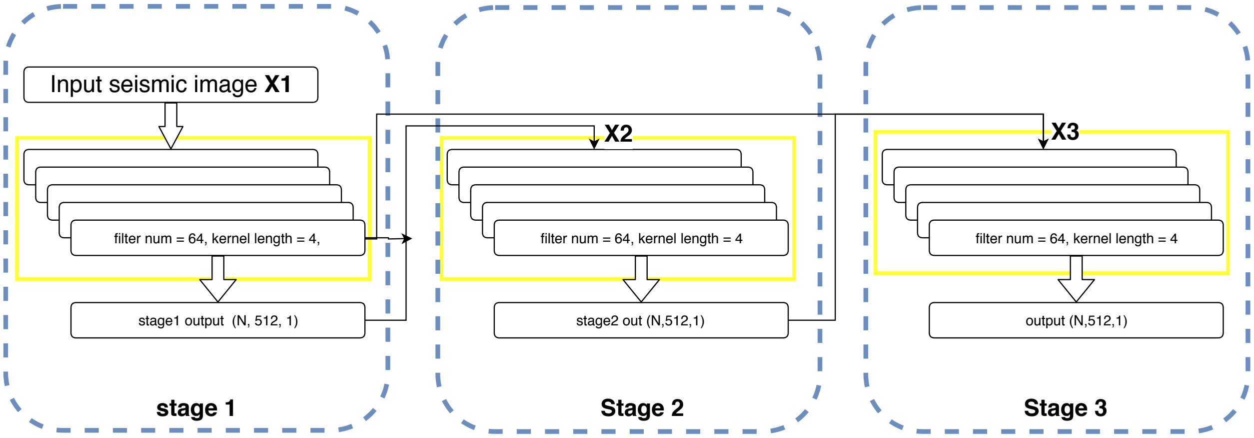


Figure 3.6 DAP architecture illustration. Within each stage, the layers within the yellow boxes are the feature extractors $f_s, s = 1,2,3$. The output layer in each stage is the predictor $p_s, s = 1,2,3$ to produce the heat map. The first stage takes the segments of 2D array input (X_1) with dimension of (N, 512, 21) for the P model and (N, 512, 42) for the S model. N is the number of instances, that is, processed 2D seismic "images," in the input data. For stage s (s = 2, 3), X_s is $f_1(X_1)$ concatenated with heat map Y_{s-1} along the X-axis.

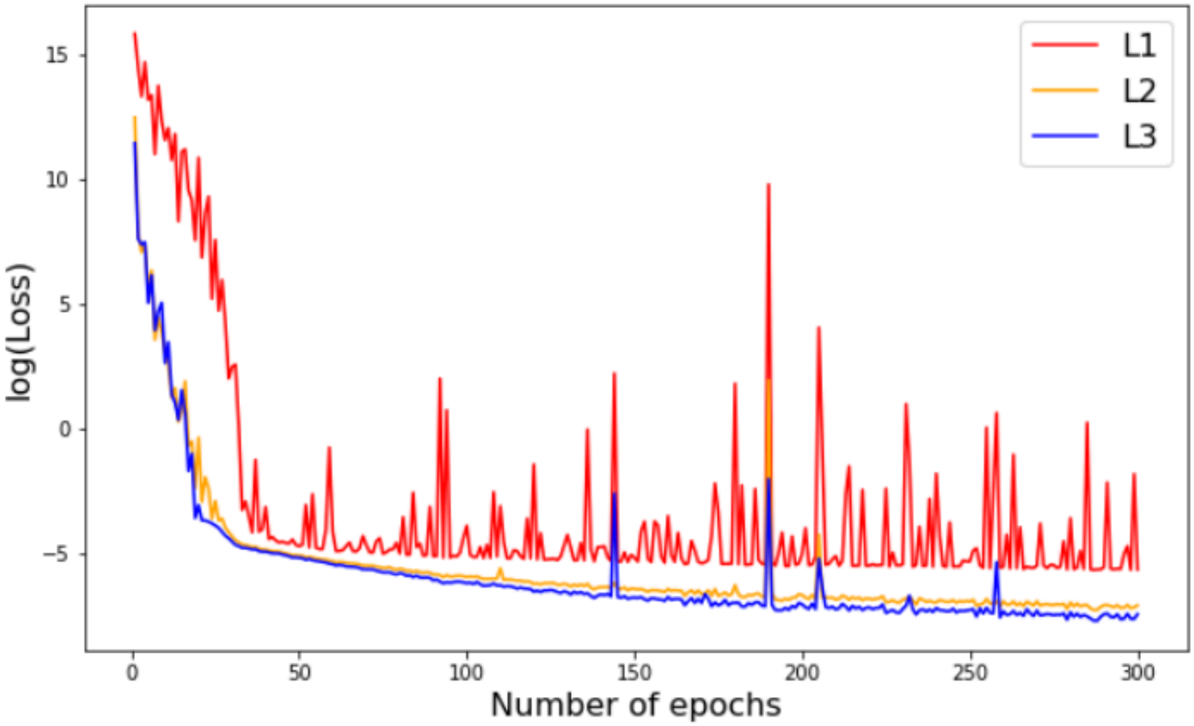
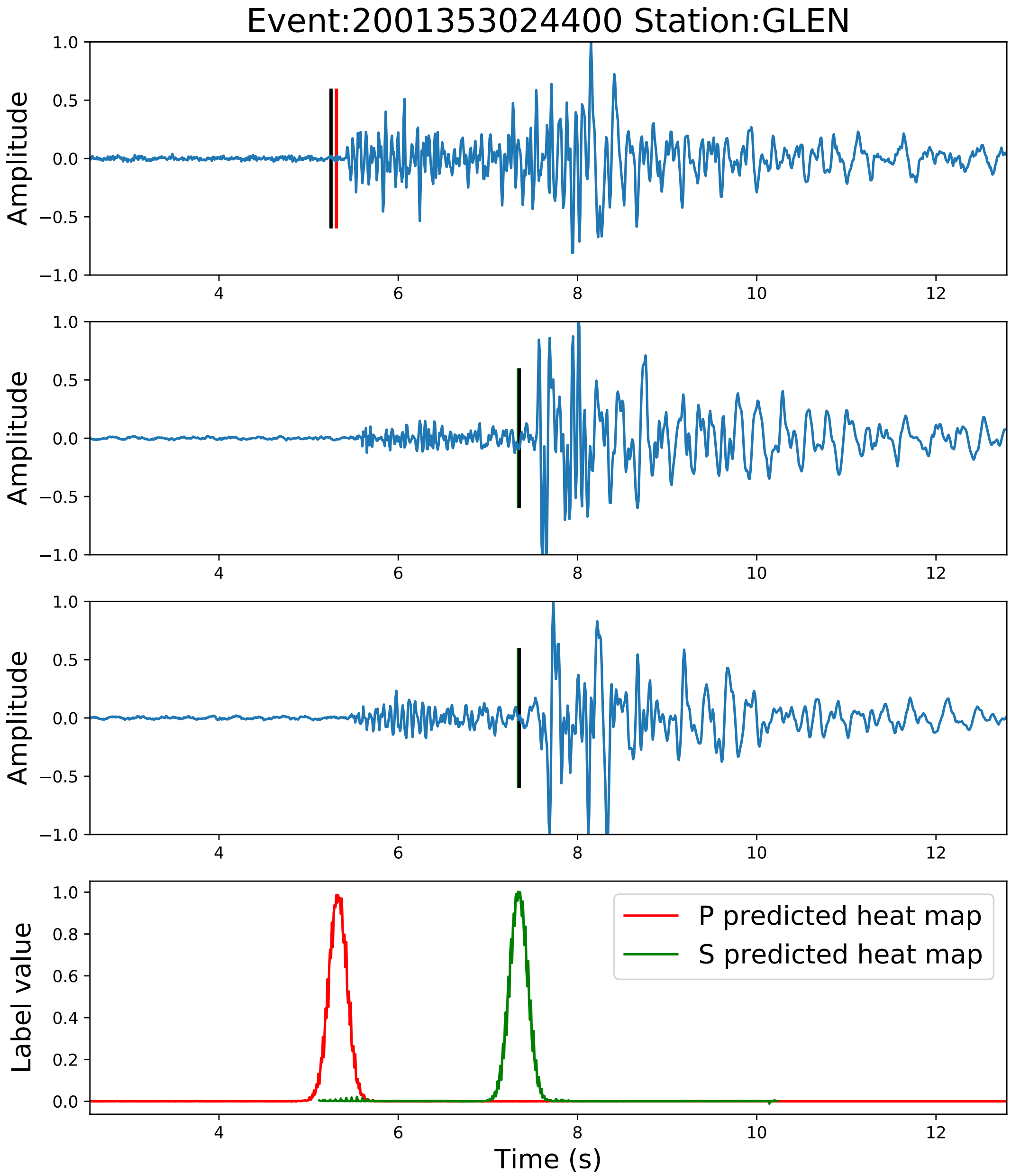
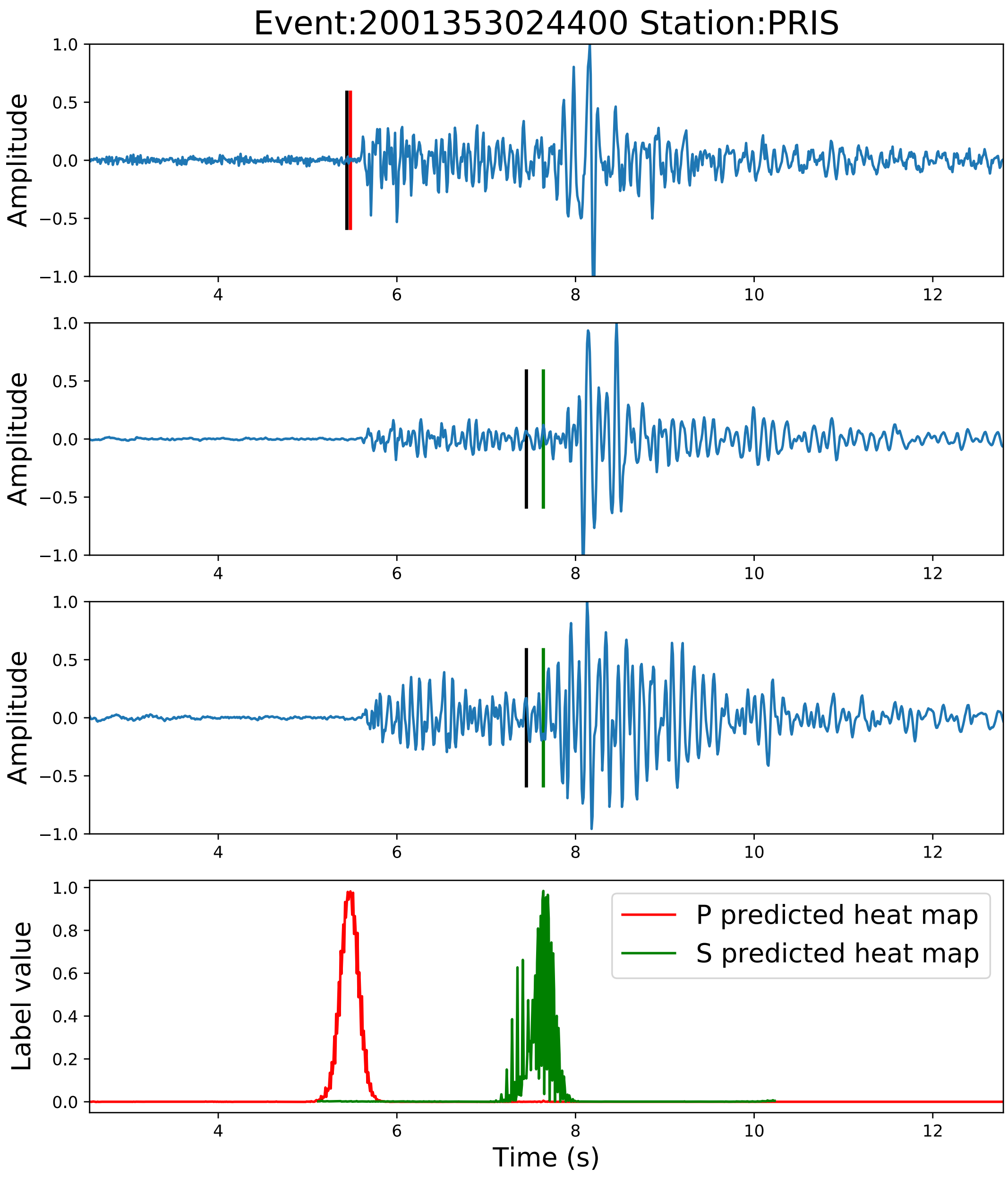


Figure 3.7 Loss for each stage versus the number of epochs. L1 is the loss for stage 1. L2 is the loss from stage 2. L3 is the loss from stage 3.


Event:2001353024400 Station:GLEN
Amplitude
1.0
0.5
0.0
−0.5
−1.0
4
6
8
10
12
Label value
0.0
0.2
0.4
0.6
0.8
1.0
P predicted heat map
S predicted heat map
Time (s)


Event:2001353024400 Station:PRIS
Amplitude
1.0
0.5
0.0
−0.5
−1.0
4
6
8
10
12
Label value
0.8
0.6
0.4
0.2
P predicted heat map
S predicted heat map
Time (s)

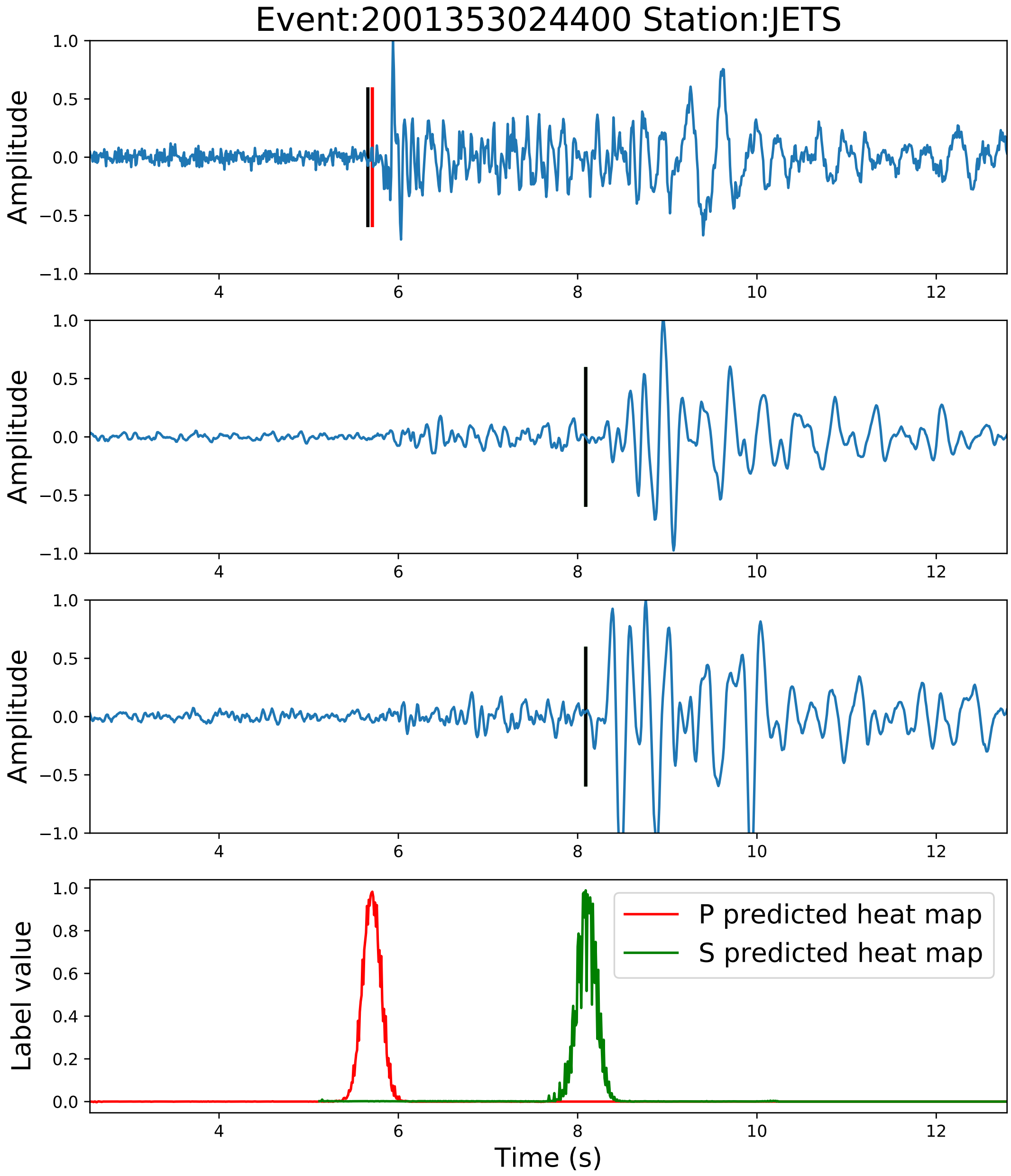


Figure 3.8 P- and S-arrival auto-picking results for 3 station-event pairs from the Parkfield study data. In each plot, the first three panels are the Z, E and N components, respectively. The red line in the first panel marks the predicted P arrival. The green line in the second and third panels show the predicted S arrival. The black lines in the first three panels show manually picked phase arrivals. The predicted and manually picked S arrivals overlap for the first and third examples. Note also the complex heat map for the second example, which is reflected in the different manual and automatic picks. The last panel shows the predicted P- and S-arrival heat maps.

4. Tomographic study in the San Francisco Bay Area, Northern California

4.1. Abstract

In previous tomography studies regarding the San Francisco Bay Area, the 3-D V_P model was obtained using body-wave arrival time data, and the 3-D V_S model was derived from surface-wave velocity model inverted by ambient noise tomography. Due to inadequate body-wave arrival time data, the available V_P model provides limited resolution for the South Napa earthquake region and the V_S model only resolves the near-surface features. In this study, we include passive and active source data between 1980 and 2016 to form an expanded body-wave dataset. The S-wave arrival time data are increased with an auto-picking algorithm. By combining the body-wave and surface-wave data, 3-D V_P and V_S models with improved resolution are jointly inverted. The newly inverted P- and S-wave velocity models provide additional insight for structural and seismic studies regarding the complex fault systems in this area.

4.2. Introduction

To solve geophysical problems, including seismic hazard assessment, detailed information on earthquake and fault properties needs to be derived. Improved knowledge of seismic velocity structure will help to shed light on the fault geometry, fault connectivity, rupture propagation, and ground motion prediction, etc. The goal of this study is to improve the 3-D P- and S-wave velocity models of the fault systems in the San Francisco Bay Area (SFBA) (Figure 4.1). This study focuses on the Hayward Fault (HF) and its junction with the Calaveras Fault (CF), the area of the 2014 South Napa earthquake, and the Sacramento-San Joaquin River Delta (SSJRD) area. Several key factors motivate the proposed study of these regions: 1) Field et al. (2013, 2015) pointed out the high estimated probability of a M ≥ 6.7 earthquake on the HF in the next 30 years; 2) Chaussard et al. (2015) reported a direct connection between the HF and CF allowing a through-going rupture; 3) the structural setting of the 2014 South Napa earthquake (Brocher et al., 2015); and 4) the vulnerability of the SSJRD levee system to potential large earthquakes (Salah-Mars et al., 2008). The improved 3-D seismic velocity models can help to address a number of key problems. Will the better resolved fault geometries support the connectivity between the HF and CF? Does the preferred rupture propagation direction correlate with the velocity contrast across various faults? In the deeper region of the 2014 South Napa earthquake, what structural features may be responsible for the aftershock distribution?

4.3. Previous work in the study region

The Hayward fault

Termed as a “tectonic time bomb” by the USGS, the HF has been characterized as “the most hazardous fault in the Bay Area given that a full average recurrence interval has passed since its last M ~7 earthquake in 1868 [Field et al., 2015]” (Chaussard et al., 2015). Moreover, the recent reported connectivity between the HF and CF increases the risk of seismic hazard due to potential through-going rupture for large earthquakes in the East Bay.

A number of seismic and geodetic studies have been focusing on the HF fault structure, including fault geometry, creeping and locked patches, and connectivity with adjacent faults (Andrews et al., 1993; Parsons et al., 2003; Ponce et al., 2004; Manaker et al., 2005; Chaussard et al., 2015). Three-dimensional seismic tomography studies have been carried out to image the 3-D structure of the HF (Zhang and Thurber, 2003; Hardebeck et al., 2007; Allam et al. 2014). For fault structure at shallow (a few kilometers) depth, the resolution from current P-wave velocity models is limited. Existing models may not be able to adequately test the hypothesis of the HF-CF connectivity and the rupture propagation direction. Because of the lack of S-wave arrival time data in this region, there is no reliable S-wave velocity model available based on body waves, although there is a recent S-wave model from ambient noise imaging (Li and Thurber, 2018).

The South Napa earthquake

The August 2014 South Napa earthquake was characterized by Brocher et al. (2015) as "the largest earthquake to strike the greater San Francisco Bay area since the Mw 6.9 1989 Loma Prieta earthquake." A number of temporary seismometers were deployed to record the aftershocks only a few days after the mainshock. The area between the Concord-Green Valley and Hayward Rogers Creek fault systems (Figure 4.1), where the South Napa earthquake and its aftershocks occurred, was previously devoid of instrumental seismicity. Thus, the recorded earthquake sequence since the 2014 South Napa earthquake will help to provide resolution of the 3-D seismic velocity structure in this region. The improved 3-D seismic velocity model and accurate earthquake locations can help to better understand the complex sub-parallel faults in this region.

Sacramento-San Joaquin River Delta

The levee system located within the SSJRD is responsible for about half of California's average annual streamflow, the majority of Californian's drinking water, and most of the farmland water usage (Lund et al., 2007). Several studies have been carried out to estimate the risk for the levee system failure caused by adjacent fault systems and major earthquakes in the SSJRD (Torres et al., 2000, Fletcher and Boatwright, 2007; Lund et al., 2007; Fletcher et al., 2008). Due to the existence of several significant faults and the increasing seismic hazard of the HF, all central SSJRD levees are threatened by medium to high risk of catastrophic failure (Lund et al., 2007).

Thurber et al. (2009) conducted a regional seismic tomography study of Northern California (NC). According to the P-wave velocity model, there are dramatic lateral variations in structure beneath the SSJRD. The V_P model shows a sharp lateral velocity contrast at the seismically active

Pittsburg/Kirby Hills fault. Due to limited seismic network coverage in the SSJRD in their dataset, the resolution in the horizontal directions is much worse than their horizontal node spacing (10 to 20 km). Again, there is no reliable S-wave velocity model based on body waves for the SSJRD either.

4.4. 3-D P- and S-wave tomography study in the SFBA

For previous SFBA tomography studies, there are 3-D P-wave velocity models inverted separately for studies of specific areas, such as the HF and CF region and SSJRD region. These available 3-D V_P models have limited resolution due to sparse seismicity and network coverage. For the region near the 2014 South Napa earthquake, which was previously devoid of seismicity, there is no 3-D V_P model available. Moreover, there is no reliable 3-D V_S model from previous studies, as the number of available S-wave arrivals is inadequate. In this study, we aim to provide a 3-D V_P model with improved resolution and the first reliable 3-D $\mathrm{V_s}$ model from a joint inversion, which utilizes the combination of an expanded P- and S-wave arrival dataset and Rayleigh wave group velocities.

With an enlarged body-wave arrival dataset, the body-wave tomography can improve resolution in regions with sufficient earthquake-station ray paths, for instance, the HF and CF region. However, one shortcoming of body-wave tomography using earthquake data is that the resolution of the results depends heavily on the station and earthquake distribution. Uneven distribution of earthquakes and stations will lead to insufficient ray coverage in certain areas, resulting in poor resolution of the seismic velocity structure. In regional body-wave tomography, lower velocities at shallow depth lead to nearly vertically traveling rays underneath the seismic stations, which smears the velocity features in the upper-most crust. Additionally, a lack of high quality S-wave arrival times creates many more uncertainties in the S-wave velocity model.

Joint inversion, which uses P- and S-wave arrival times and group and/or phase velocity data from Rayleigh and/or Love waves, can be used to improve the resolution of 3-D seismic velocity structure (Zhang et al., 2014; Fang et al., 2016). One way to do joint inversion is to directly relate the dispersion data to the 3-D V_P and V_S models. Fang et al. (2016) used this approach with body-wave data and Rayleigh wave group velocities to image the plate-boundary region around the San Jacinto Fault in southern California. These jointly inverted seismic velocity models provide a better fit to observed seismic waveforms than the Southern California Earthquake Center community model of Shaw et al. (2015). In our SFBA tomography study, the joint inversion algorithm from Fang et al. (2016) is used to invert for the 3-D P- and S-wave velocity models.

4.5. Dataset

4.5.1. Body-wave dataset

To create a comprehensive body-wave dataset, we examined the earthquakes in the Northern California Earthquake Data Center (NCEDC) recorded both by permanent and temporary stations between 1980 and 2016. The candidate earthquakes include events with magnitude greater than or equal to 3 that occurred between 1980 and 1999 and events with magnitude no less than 2.6 that occurred between 2000 and 2016. To filter out redundant events, we divide the study region into cells with dimension of 0.005 degrees in both latitude and longitude and 0.1 km in the vertical direction. Based on catalog information, within each cell, the two events with the most picked P-wave arrivals are selected, resulting in a set of 4,164 events. These selected NCEDC events include the 2014 South Napa earthquake and some of its aftershocks.

In addition to the selected events from the NCEDC, an additional 784 events are included from the dataset used for a previous SSJRD study (Teel, 2012). These events provide additional data coverage for the SSJRD area. From the previous Northern California (NC) tomography study (Thurber et al., 2009), 4,494 explosions and airgun shots are added to our body-wave dataset as well. The active source data can help to improve the P-wave velocity resolution at shallow depths. The final body-wave dataset includes 9,442 events (Figure 4.2) with 392,894 P-wave arrivals and 12,693 S-wave arrivals recorded on 1,093 temporary and permanent stations. Due to the difficulty of picking S-wave arrival times, more than half of the earthquakes have fewer than 4 S-wave arrivals in the catalog.

In order to increase the number of P and S arrivals in our dataset, we use an auto-picking algorithm, kpick, by Rawles and Thurber (2015) to pick additional P and S arrivals. As the S-wave arrival time auto-picking requires three-component waveforms, waveforms are downloaded from the NCEDC for events that were recorded by three-component stations within our study region, including 30 stations of the Berkeley Digital Seismic Network (BK) and 35 stations in the Northern California Seismic Network (NC) (Figure 4.2). Then, kpick is applied to these waveforms for P- and S-wave arrival auto-picking. After auto-picking, the number of P and S arrivals are 453,308 and 29,539, respectively. The number of P arrivals increased by 15.4% whereas the number of S arrivals increased by 132.8%. As shown in Figure 4.3, the number of events with more than 10 S-wave arrivals increased from 449 to 1,434.

The body-wave tomography part of the joint inversion algorithm by Fang et al. (2016) is based on the double difference tomography algorithm tomoDD, by Zhang and Thurber (2003). Thus, catalog event pair differential times need to be calculated. For our dataset, an event pair is defined as two events with less than 5 km catalog separation and more than 4 links. A link means two events are recorded by the same station. Based on the catalog information, 57,518 event pairs chosen. The number of differential times for P- and S-wave arrivals are 1,604,722 and 43,211, respectively.

4.5.2. Surface wave dataset

Li and Thurber (2018) determined 2-D group velocity maps for Rayleigh waves at 0.5 to 14 s period in the SFBA using the ambient noise tomography method. The resolution for the group velocity models is ~10 km at 3 s period and ~30 km at 10 s period. The 2-D Rayleigh wave group velocity data are used in our joint inversion study for the SFBA, providing 20,421 Rayleigh-wave travel times.

4.6. Joint inversion setup

The study region is meshed with 79 by 66 grid nodes with an interval of 0.05° in both latitude and longitude (Figure 4.2), centered on 37.75°N, 122.15°W. For the vertical direction, the nodes are positioned at -1, 0.0, 1.0, 2.0, 4.0, 6.0, 8.0, 10.0, 12.0, 15.0, 20.0, 25.0, and 35.0 km depths. The 3-D $\mathrm{V_P}$ model from Thurber et al. (2009) and the 3-D V_S model from Li and Thurber (2018) are used as the initial models for the joint inversion.

For the body-wave inversion, to integrate the absolute and catalog differential data types, we used a standard hierarchical weighting scheme (Zhang and Thurber, 2003). Initially, the absolute arrival data are weighted heavily relative to the differential time data. This allows the relocations and velocity structure to be established at large scale. Then the weight on catalog differential time data is increased for better constraint on the relative earthquake locations and higher resolution in source regions.

Fang et al. (2016) used a balancing parameter α to set the relative weighting between surface-wave data and body-wave data. The parameter α is chosen based on

$$\alpha = \sqrt{\frac{\mathrm{N_p}\sigma_{\mathrm{p}}^2 + \mathrm{N_p}\sigma_{\mathrm{s}}^2}{N_{sw}\sigma_{sw}^2}} \tag{4.1}$$

following Julia et al. (2000), where $\mathrm{N_p}, N_s,$ and $\mathrm{N_{sw}}$ are the number of P and S arrival times and surface wave travel times, respectively, and $\sigma_{\mathrm{p}}, \sigma_s,$ and σ_{sw} are the estimated uncertainties for the corresponding datasets. The uncertainties for P- and S-wave arrival time data based on body-wave only tomoDD inversion are ~0.09 s and ~0.14 s, respectively. The average Rayleigh wave data uncertainty from Li and Thurber (2018) is ~2.3 s. Therefore, the balancing parameter chosen for our dataset is 0.37.

4.7. Trade-off analysis

In the Fang et al. (2016) joint inversion algorithm, first-order Tikhonov smoothing is used to regularize the inversion (Aster et al., 2013). There are two separate groups of smoothing parameters for the body-wave and surface-wave inversion parts. A trade-off analysis, which helps to find the balance between minimizing data misfit versus minimizing model perturbations, is carried out to find the optimal smoothing parameters for the joint inversion. In the trade-off analysis, a trade-off curve, also known as an L-curve, is used to determine the optimal smoothing parameter quantitatively. As the smoothing parameter increases, the model perturbation decreases and the data misfit increases. The L-curve shows this relationship between model perturbation versus data misfit over different smoothing parameters. In general, the optimal smoothing parameter is found where the maximum curvature occurs in the L-curve (Aster et al., 2013).

4.7.1. Body-wave smoothing parameter

For the body-wave inversion part, the joint inversion algorithm (Fang et al., 2016) allows smoothing parameters to vary in different directions. As the scales for the two horizontal directions in our study region are similar, we set the smoothing parameters for the horizontal directions to be the same. The trade-off analysis is then carried out in two steps: 1) find the optimal horizontal direction smoothing parameters with a fixed ratio between the horizontal and vertical directions; 2) with the optimal horizontal smoothing parameters, search for the best smoothing parameter for the vertical direction.

To find the optimal smoothing parameters for the horizontal directions, we tested smoothing values of 20, 50, 100, 200, 300, 400, and 500. As the vertical direction scale is much smaller than in the horizontal directions, we initially fixed the ratio between horizontal and vertical smoothing parameters to be 2. As shown in Figure 4.4a, the decrease of model perturbation norm slows down substantially as the data root mean square (RMS) residual increases to about 400 ms at a smoothing parameter value of 100, which is chosen for the horizontal directions. With the horizontal smoothing parameters fixed at 100, we test smoothing values of 10, 20, 30, 40, 50, and 60 for the vertical directions. We note that the tested smoothing parameters are generally small, they can cause inversion instability for a nonlinear problem (Aster et al., 2003). In Figure 4.4b, the L-curve value at a smoothing value of 10 is a typical example of this inversion instability. From the L-curve shown in Figure 4.4b, we select 30 as the smoothing parameter for the vertical direction.

4.7.2. Surface-wave smoothing parameter

The surface-wave smoothing parameter is fixed to be uniform for all three directions in the joint inversion algorithm (Fang et al., 2016). As the body-wave smoothing parameters are 100, 100, and 30 for latitude, longitude and vertical directions, respectively, the surface-wave smoothing parameter should be at a similar scale. With the preselected body-wave smoothing parameters, we tested surface-wave smoothing parameter values of 20, 40, 60, 100, and 200. As shown in Figure 4.4c, the maximum curvature occurs for a smoothing parameter of 60, so 60 is selected as the surface-wave smoothing parameter.

4.8. Checkerboard test

Analysis of model resolution is essential as it helps to assess the adequacy of seismic ray coverage as well as the spatial resolution limits. In a checkerboard test, synthetic data are generated with a checkerboard anomaly model. Then the synthetic data and the initial model without anomaly are used to try to recover the model with anomalies. In our study, two sets of checkerboard tests are conducted to verify the reliability of the inverted seismic velocity models at different scales. For the first set of checkerboard tests, the dimension of alternative positive and negative anomalies is ~25 km in the horizontal directions and ~15 km in the vertical direction and the maximum anomaly amplitude is 5% (Figure 4.5a, 4.5b). In the second set of checkboard tests, the anomalies are ~40 km in the horizontal directions and ~15 km in the vertical direction, with a maximum anomaly amplitude of 10% (Figure 4.5c, 4.5d). Synthetic body-wave and surface-wave data are generated for each of the checkboard models individually. Uniformly distributed random noise within the range of 0.09 s and 0.14 s is added into the P- and S-wave synthetic arrival times, respectively. For the Rayleigh-wave travel times, random noise with uniform distribution within ± 2.3 s are added. These random noise levels are the same as the uncertainties analyzed from previous studies. The 1-D initial model that is the same as the initial model used in Thurber et al. (2009) is used for both sets of the checkboard tests, and the joint inversion results with body and surface waves are compared to those with body waves only.

As shown in Figure 4.5, the anomaly amplitude varies in all three directions, and the anomalies have a relatively large amplitude at depths Z = 0, 1, 10, and 12 km. At depths Z = 2, 6 and 8 km, the anomaly amplitudes are relatively small. For depths ranging between 3 to 5 km, the anomaly amplitudes are almost zero. We plot the results for depths with large anomaly amplitudes as it is easier to visualize the recovery difference (Figure 4.6 and 4.7). The checkerboard test results for other layers are shown in the Appendix Figure A4.1-4.3.

The checkboard tests results for V_P are shown in Figure 4.6 At depths Z = 10 and 12 km, the anomalies are recovered well from both body-wave inversion and joint inversion (Figure 4.6b). At depths Z = 0 and 1 km, the joint inversion yields better resolved anomalies compared to the body-wave inversion (Figure 4.6a). Based on the results shown in Figure 4.6 and 4.7, the V_P model from joint inversion has reasonable resolution overall to a depth of 20 km for anomaly size of ~25 km horizontally (Figure 4.6e). As the seismic ray density can be expressed by the derivative weight sum (DWS) (Thurber 1983) and used as a proxy for resolution, we find the regions with P-wave DWS values greater than 36 are reasonably well resolved.

For V_S, based on the results in Figure 4.6 and 4.7, joint inversion shows substantially improved resolution over the body-wave inversion at shallower depths, Z = 0 and 1 km, for both ~25 and ~40 km horizontal anomalies. The body-wave only inversion fails to recover anomalies at these depths. At a greater depth of Z = 10 and 12 km, even though body-wave inversion is able to recover some of the anomalies, the joint inversion recovers much clearer and more anomalies for both anomaly sizes. The maximum sensing depth for the jointly inverted V_S model is ~20 km (Figure 4.7e). The regions with S-wave DWS values greater than 16 are reasonably well resolved.

In order to quantify the differences between joint inversion and body-wave inversion results for all depths, root mean square (RMS) model recovery misfits are calculated by

$$r = \sqrt{\frac{\sum_i (\overline{m}_i - m_i)^2}{n}} \tag{4.2}$$

where r is the RMS misfit, $\overline{m}_i$ is the recovered model value at the ith node, m_i is the true model value at ith node, and nodes considered in this calculation are ones with DWS values greater than 36 (V_P model) and 16 (V_S model). The results are shown in Table 4.1.

For the V_P model, with anomaly size of ~25 km horizontally, the joint inversion shows better resolved anomalies at layers with large anomaly amplitudes. The average decrease of RMS model recovery misfit from body-wave inversion to joint inversion is 25.3% at Z = 0 and 1 km depths and 6.5% for deeper layers (Z = 10 and 12 km). The maximum decrease of RMS model recovery misfit of 40.9% is at Z = 0 km depth. When the anomaly amplitudes become too small, both joint inversion and body-wave only inversion result in worse resolved anomalies. The true anomaly amplitudes are almost zero at depths between 3 to 5 km (Figure A4.1), the very small anomalies at 4 km depth will affect the inversion result at 2 km depth when the vertical smoothing parameter smears the misfits into nearby layers. The body-wave inversion resolves the anomalies with over-exaggerated anomaly amplitudes whereas the joint inversion yields more smeared anomalies (Figure A4.2 and A4.3). The anomalies at greater depths (Z = 6, 8, 15, and 20 km) are constrained better than those at 2 and 4 km depths as the vertical anomaly scale (~14 km) is larger compared to that of the anomaly between 0 to 4 km depths. For layers at Z = 6, 8, 15, and 20 km depths, the joint inversion shows an average 5.1% RMS model recovery misfit decrease. For anomaly size of ~40 km, the average decrease of RMS model recovery misfit from body-wave inversion to joint inversion is 40.8% at Z = 0 and 1 depths and 5.0% for deeper layers (Z = 10 and 12 km). The maximum decrease of RMS model recovery misfit of 49.3% is at Z = 0 km depth. The average RMS misfit for depths at Z = 6, 8, 15, and 20 km is 13.7%. The decrease of RMS model recovery misfits for P-wave velocity demonstrates that the sensitivity for V_P from Rayleigh wave data.

For V_S, joint inversion shows improved resolution compared to the body-wave only inversion, especially at Z= 0 and 1 km depths. The average decrease of RMS model recovery misfit for layers Z = 0 and 1 km is 29.0% and 44.3% for the ~25 km and ~40 km anomalies, respectively. The maximum decrease of RMS model recovery misfit is at Z = 0 km depth for both the ~25 km anomaly (41.5%) and the ~40 km anomaly (51.4%). At 10 and 12 km depths, the corresponding average decrease of RMS misfit for anomaly sizes of ~25 and ~40 km is 5.2% and 3.4%. Though smearing for Z = 2 and 4 km depths also shows up in the V_S joint inversion results, the joint inversion shows much better recovered anomalies in region between latitude 37.75°N and 36.50°N (Figure A4.1-4.3). The average RMS misfit decrease for depths Z = 6, 8, 15, and 20 km is 5.8% and 12.5% for the ~25 km and ~40 km anomalies, respectively. For V_S, the greater decrease of RMS model recovery misfit shows that the Rayleigh wave data provide the most constraint on the near-surface S-wave velocity structure.

4.9. Joint inversion results and discussion

In previous Northern California 3-D V_P studies (Thurber et al., 2007; 2009), the trend of seismic velocity variations appeared to be predominantly perpendicular to the SAF. In order to make the inversion results easier to visualize and compare with previous models (Thurber et al., 2007; 2009), the region with reasonable resolution for V_P and V_S models (red box in Figure 4.2) is projected onto a Cartesian coordinate system with origin at (38° N, 122° W) and rotation angle 36° counter-clockwise from North, matching that of Thurber et al. (2009). The projected study region spans 0 to -160 km in the X direction and 50 to -300 km in the Y direction. In Figure 4.8 and 4.9, the V_P model (DWS $\geq$ 36) and V_S model (DWS $\geq$ 16) are plotted in map view at depths of 0, 1, 2, 4, 6, 8, 10, 12, 15, and 20 km. Cross-sections along the X axis are plotted for these models at Y = 20, 0, -20, -40, -60, -80, -100, -120, -140, -160, -180, -200, -220, -240, and -260 km (Figure 4.10 and 4.11).

4.9.1. Basement highs

For shallow layers at depths Z = 0 and 1 km, the joint inverted V_P and V_S model resolutions are improved substantially over the body-wave only inversion results based on the checkboard results. With the improved resolution, many expected basement highs (V_P > 5 km/s) are well identified at shallow depths. And some of the identified basement highs also show up in our V_S model. In these shallow layers, sedimentary rocks are expected to have V_P lower than 5 km/s (Brocher, 2005). Thurber et al. (2007) showed a number of basement highs for regions from Y = -260 to -50 km in our study box. Compared with the V_P model from Thurber et al. (2007), all expected basement highs are identified in our jointly inverted V_P model (Figure 4.12, features E-O). For regions from

Y = -50 to 50 km in our study box, an additional set of basement highs also is apparent (Figure 4.12 features A-D). These identified basement highs are: (A) Tertiary volcanic rocks northeast side of The Geysers and Hidden Valley Lake; ($B_{1,2}$) Franciscan rocks in the Coastal Ranges (B_1 Elk Range, B_2 Mayacamas Mountains) on the west and south sides of The Geysers; (C) Franciscan rocks northeast side of Pope Valley; (D) Franciscan complex in the Northern Coast Ranges (NCR) on the west side of the Rodgers Creek Fault (RCF); (E) Franciscan rocks of the Marin Hills and the San Francisco Bay block on the west side of the HF; (F) Salinian rocks of the Montara pluton; (G) Franciscan rocks of the Pilarcitos block; (H) Franciscan rocks of Mount Diablo (MD); (I) Franciscan rocks of Coyote Hills; (J) Salinian rocks of the Ben Lomond pluton; (K) Monte Vista fault hanging wall with Franciscan rocks; (L) Franciscan rocks of the Santa Cruz Mountains; (M) Franciscan rocks of the Diablo Range (DR); (N) Salinian rocks of the Gabilan Range (GR) pluton; and (O) Granites and metasediments of the Santa Lucia Range (SLR). In our V_P model, these basement highs continue from 0 km depth to 2 km depth with increasing velocities. At 2 km depth, the maximum V_P for most of these basement highs is greater than 5 km/s. As shown in Figure 4.12, V_S also shows high velocity anomalies for features B_1, E, G, H, I, K, M, N, and O in the regions with reasonable resolution. Identification of these expected geological features provides additional support for the reliability of our models.

The Franciscan complex contributes to most of the high velocity anomalies (V_P > 5 km/s) at depths shallower than 4 km. Features in the NCR ($B_{1,2}, A, D, E$), MD (H), and DR (M) are associated with large exposures of Franciscan complex. Furthermore, a mixture of Franciscan rocks and gabbro also results in a high V_P anomaly for feature K in the hanging wall of Monte Vista Fault and L in the Santa Cruz Mountains. Moreover, the granitic rocks lead to a high V_P anomaly in the Montara pluton (F), Ben Lomond pluton (J), GR (N) and SLR (O).

In contrast to the basement highs, places with sedimentary deposits can be recognized easily by the presence of low V_P anomalies. The prominent low V_P zone between features I and K lies within the SCV. The north side of feature H is the low V_P anomaly associated with the SSJRD. A low V_P anomaly lies in the region south of feature L and between the SAF and the CF, an area of Quaternary deposits.

4.9.2. HF and CF junction region

With the joint inverted $\mathrm{V_P}$ and $\mathrm{V_S}$ model, we are able to observe fault zone structures at shallow depths with greatly improved resolution. In previous tomographic studies regarding the HF structure, P-wave velocity across the HF changes from high to low at depth from west to the east (Zhang and Thurber, 2003; Hardebeck et al., 2007, Thurber et al., 2009). As shown in the cross-section plots in Figure 4.10f–i at Y = -80, -100, -120, and -140 km, the lateral velocity contrast across the HF is consistent from 5 to 20 km depth. However, at depths shallower than 5 km, the lateral velocity contrast fades away from Y = -80 to -140 km. This behavior corresponds to the geological setting associated with a basement high (feature E) on the northwest side of the HF. In the V_S model, a similar lateral velocity contrast across HF is observed (Fig. Figure 4.11f–i) between Z = 0 and 20 km depths. Due to limited resolution at greater depth, Li and Thurber (2018) only shows lateral $\mathrm{V_S}$ contrast across HF at depths shallower than 12 km.

For the CF, the lateral velocity contrast changes along the fault trace from north to south (Figure 4.10g–n). In the northern portion between Y = ~100 and -140 km, velocities on the southwest side of the CF are higher than those on the northeast side at depths shallower than 10 km. In the southern portion, between Y = ~-160 and -240 km, the lateral velocity contrast reverses. This lateral velocity contrast across the CF is consistent with our observation of the basement highs at depths shallower than 4 km.

Within the region from Y = -80 to -120 km, the seismicity beneath the HF trace trends almost vertically at depths between 2 and 10 km. For the CF, in the region Y = -160 to -180 km, earthquakes occured on a steeply dipping plane aligned under the CF surface trace at depths between 5 and 15 km. As they are both strike slip faults, such seismicity clearly indicates their fault planes at depth.

Complex seismicity and V_P changes in the junction region (red box in Figure 4.8) are observed at depths between 2 and 10 km. Using InSAR defined creeping fault traces and characteristically repeating earthquakes, Chaussard et al. (2015) hypothesized that there is potentially a connected fault plane between the HF and CF within the junction region. Earthquakes between the HF and CF traces appear to fall along a plane connecting the two fault traces at 4 – 10 km depths. These earthquakes also align along the velocity contrast between the low velocity anomaly under the SCV and the high velocity body extending from the end of the HF to the CF (red box in Figure 4.8). In the cross-section plot at Y = -140 km (Figure 4.10i), the earthquakes underneath each fault trace do not define a linear feature, instead, they are horizontally scattered at depths of 7 to 12 km beneath the CF trace. In Figure 4.13, events located in the CF and HF junction region are plotted in 3D. Within the region (Y ≤ -125 km) where the potential interception of the CF and HF will occur, the events fall onto a plane that extends from the south end of the HF surface trace and connects with the events located underneath the CF at depth of 5 to 12 km. In addition, there is a lateral V_P contrast at depth between 7 to 12 km on the west side of these clustered events. These observations at the junction region of the HF and the CF provide additional support for the potential connection of the two faults above 12 km depth (Chaussard et al., 2015).

4.9.3. South Napa earthquake region

In August, 2014, a M_W 6 earthquake occurred in southern Napa Valley. A series of aftershocks were recorded by temporary seismic stations. With the 2014 South Napa earthquake and aftershocks included, our jointly inverted models provide additional details for the seismic velocity structure in this region. As shown in the cross-section plot for Y = -60 km (Figure 4.10e), the mainshock and aftershocks (red dots) are distributed from 5 to 11 km depth. A small lateral velocity contrast appears in the location where these South Napa earthquake events are located (black box in Figure 4.10e). The V_P model from Thurber et al. (2009) also shows this lateral V_P variation, although there was no earthquake located within this region in their study.

The focal mechanism of the 2014 South Napa earthquake shows a right lateral strike-slip faulting with $157° – 165°$ strike and $85°$ dip (Brocher et al., 2015). In Fig.4.10e, the lateral velocity change at depths between 6 and 10 km shows a contrast at $V_P = 5.8$ km/s with a dip angle of $\sim 80°$. As the Y-axis of the projected Cartesian coordinates is slightly different from the proposed strike direction, the differences in these dip angles can be expected. Brocher et al. (2015) mentioned that the fault plane fit by the South Napa earthquakes is about 10 km in length starting from the mainshock epicenter to the northwest. In our model, this lateral V_P contrast associated with South Napa earthquakes fades away further north and south. At the cross-section to the north at Y = -40 km, the earthquakes located at X = -70 km at depth 5 to 10 km are on the west side of the high V_P anomaly. At the cross-section to the south at Y = -80 km (Figure 4.10f), the major lateral contrasts are from the HF and SAF faults. As observed in our V_P model (Figure 4.10e), it is clear that the South Napa earthquake and its aftershocks located along a ~80° plane that defines a lateral velocity contrast at depths of ~7 to ~12 km. This seismicity distribution and velocity contrast agree with the previous South Napa earthquake focal mechanism study by Brocher et al. (2015).

4.9.4. SSRJD region

Thurber et al. (2009) observed a sharp lateral V_P contrast underneath the Pittsburg/Kirby Hills fault (PKHF). This V_P contrast also shows up in our V_P model in the cross-section at Y = -100 km at depths from 2 to 20 km (Figure 4.10g). With improved V_P resolution at shallower depths, we find that the two low V_P zones underneath the Mt. Diablo Fault (MDF) and PKHF are interrupted by a high V_P zone beneath the Greenville Fault (GF) at depths between 2 and 5 km. In Figure 4.8 and 4.10, the low V_P zone on the east side of the PKHF around X = -50 km and Y = -100 km shows up clearly from Z = 0 to 20 km depths. Underneath the MDF, a low V_P zone is present at Z = 0 to 6 km depths (Figure 4.10h). For the V_S model, due to limited resolution to the east of the PKHF, a lateral velocity contrast is not observed, but a similar low V_S anomaly shows up underneath the MDF at depths from 0 to 6 km (Figure 4.11h).

4.9.5. Junction region of the SAF and CF

In our V_P and V_S models, the complex lateral velocity variations associated with the SAF and CF systems are well observed from 0 to 20 km depths. In the southern part of our study region, Y = -220 to -280 km, a sharp and dramatic lateral velocity contrast across the SAF at depths ≤ 12 km is observed in both the V_P and V_S models. At shallow depths ($Z \leq 4$), the high velocity body on the east side of the SAF matches the geological feature of the GR (feature N). The existence of the SCV leads to a low velocity body in between the SLR and GR. At greater depths, Z = 4 – 12 km, the high seismic velocity underneath the GR grows larger to the northwest along the SAF and west towards the SLR.

In the junction region between the SAF and CF (Y = -200 to 260 km), there is a dramatic low V_P zone east of the SAF and west of the DR at shallow depths ($Z \leq 4$ km). This low velocity zone also appears in our V_S model. Seismicity in this junction region appears only at depths below this low velocity zone. This low velocity zone corresponds to the presence of Quaternary deposits at shallow depths.

4.9.6. The Geysers

With densely distributed seismic events in region around The Geysers, the resolved V_P at shallow depths ($Z \leq 4$ km) matches well with surrounding geological features (feature A, $B_{1,2}$ and C). Earthquakes in The Geysers appear to locate within a high V_P body situated at depths $Z \leq 5$ km. At Y = 20 km, seismicity east of The Geysers is clustered on a velocity contrast with the low V_P zone on the east side of these earthquakes, corresponding to sediments from Quaternary deposits and the Great Valley Sequence. For the V_S model, resolution in the surrounding regions is limited, but a high V_S anomaly is observed where The Geysers earthquakes occur.

4.10. Conclusion

In this study, 3-D P- and S-wave velocity models for the SFBA are determined with joint inversion using a combination of body-wave and surface-wave data. With improved resolution for V_P at shallow depths, the high and low velocity zones clearly match up with the geological features. At greater depths (Z = 4 – 20 km), V_P structures near the SAF, CF, HF, MDF, SJJRD, and The Geysers are consistent with the previous NC V_P model from Thurber et al. (2009). As the spacing for our joint inversion is ~5 km in the horizontal directions, which is finer than that used in Thurber et al. (2009), the fault zone resolution is improved in our V_P model. A 3-D V_S model is also obtained for the SFBA with Rayleigh wave travel times and an increased number of S-wave arrivals. The jointly inverted V_S model shows overall much better resolution than the body-wave only V_S model (Figure 4.7, Table 4.1). The inverted V_S model shows generally consistent behavior to the V_P model in the SAF, CF, HF, MDF, and The Geysers regions.

4.11. Acknowledgements

This research is supported in part by the USGS, Department of the Interior, under USGS Award Number G14AP00056 to the University of Wisconsin-Madison. The support of a student assistantship from the Mark and Carol Ann Solien Graduate Assistantship and George P. Woollard-Sigmund I. Hammer Memorial Fund in Geology & Geophysics of the University of Wisconsin-Madison are also acknowledged. We thank Hongjian Fang for helping with the joint inversion code.

4.12. References

Allam, A. A., Ben-Zion, Y., and Peng, Z., 2014, Seismic imaging of a bimaterial interface along the Hayward fault, CA, with fault zone head waves and direct P arrivals. Pure Appl. Geophys., 171, 2993- 3011, doi: 10.1007/s00024-014-0784-0.

Andrews, D. J., Oppenheimer, D. H. and Lienkaemper, J. J., 1993, The Mission link between the Hayward and Calaveras Faults. J. Geophys. Res., 98, 12,083-12,095, doi:10.1029/93JB00712.

Aster, R.C., Borchers, B. and Thurber, C.H., 2013, Parameter Estimation and Inverse Problems. 2nd edn, Elsevier, 360 pp.

Brocher, T. M., Baltay, A. S., Hardebeck, J. L., Pollitz, F. F., Murray, J. R., Llenos, A. L., Schwartz, D. P., Blair, J. L., Ponti, D. J. and Lienkaemper, J. J., 2015, The M 6.0 24 August 2014 South Napa earthquake. Seismol. Res. Lett., 86, 307-308, doi:10.1785/0220150023.

Chaussard, E., Bürgmann, R., Fattahi, H., Nadeau, R. M., Taira, T., Johnson, C. W., and Johanson, I., 2015, Potential for larger earthquakes in the East San Francisco Bay Area due to the direct connection between the Hayward and Calaveras Faults. Geophys. Res. Lett., 42, 2,734-2,741. doi:10.1002/2015GL063575.

Fang, H., Zhang, H., Yao, H., Allam, A., Zigone, D., Ben-Zion, Y., Thurber C. and van der Hilst, R. D., 2016, A new three-dimensional joint inversion algorithm of body-wave and surface-wave data and its application to the Southern California Plate Boundary Region. J. Geophys. Res., 121, doi:10.1002/2015JB012702.

Field, E. H., and 2014 Working Group on California Earthquake Probabilities, 2015, UCERF3: A new earthquake forecast for California's complex fault system. U.S. Geological Survey FS 2015-3009, 6 pp., http://dx.doi.org/10.3133/fs20153009.

Field, E. H., Biasi, G. P., Bird, P., Dawson, T. E., Felzer, K. R., Jackson, D. D., Johnson, K. M., Jordan, T. H., Madden, C., Michael, A. J., Milner, K. R., Page, M. T., Parsons, T., Powers, P. M., Shaw, B. E., Thatcher, W. R., Weldon, II, R. J. and Zeng, Y., 2013, Uniform California Earthquake Rupture Forecast, version 3 (UCERF3)-The time-independent model, U.S. Geological Survey Open-File Report 2013- 1165, 97 pp.

Fletcher, J., and Boatwright, J., 2007, Site response in the San Joaquin/Sacramento River Delta. Eos Trans. AGU, 88, S53B-1270.

Fletcher, J., Boatwright, J., and Sell, R., 2008, Site response of the Sacramento/San Joaquin Delta. Seismol. Res. Lett., 79, 351.

Hardebeck, J. L., Michael, A. J., and Brocher, T. M., 2007, Seismic velocity structure and seismotectonics of the Eastern San Francisco Bay Region, California. Bull. Seismol. Soc. Am., 97, 826-842, doi:10.1785/0120060032.

Julia, J., Ammon, C.J., Herrmann, R.B. and Correig, A.M., 2000, Joint inversion of receiver function and surface wave dispersion observations. Geophys. J. Int., 143(1), pp.99-112.

Li, P. and Thurber, C., 2018, Rayleigh wave group velocity and shear wave velocity structure in the San Francisco Bay region from ambient noise tomography. Geophys. J. Int., 213(3), 1,599-1,607.

Lund, J., Hanak, E., Fleenor, W., Howitt, R., Mount, J., and Moyle, P., 2007, Envisioning futures for the Sacramento-San Joaquin Delta. Public Policy Institute of California.

Manaker, D. M., Michael, A. J., and Bürgmann, R., 2005, Subsurface structure and kinematics of the Calaveras-Hayward fault stepover from three-dimensional V_P and seismicity, San Francisco Bay region, California. Bull. Seismol. Soc. Am., 95, 446-470, doi:10.1785/0120020202.

Parsons, T., Sliter, R., Geist, E. L., Jachens, R. C., Jaffe, B. E., Foxgrover, A., Hart, P. E., and McCarthy, J., 2003, Structure and mechanics of the Hayward-Rodgers Creek fault step-over. Bull. Seismol. Soc. Am., 93, 2,187-2,200.

Ponce, D. A., Simpson, R. W., Graymer, R. W., and Jachens, R. C., 2004, Gravity, magnetic, and high- precision relocated seismicity profiles suggest a connection between the Hayward and Calaveras Faults, northern California. Geochem. Geophys. Geosyst., 5, Q07004, doi:10.1029/2003GC000684.

Rawles, C., Thurber, C., 2015, A non-parametric method for automatic determination of P-wave and S-wave arrival times: application to local micro earthquakes. Geophys. J. Int., 202(2), 1,164–1,179.

Salah-Mars, S., Rajendram, A., Kulkarni, R., McCann, Jr., M., Logeswaran, S., Thangalingam, K., Svetich, R., and Bagheban, S., 2008, Seismic vulnerability of the Sacramento-San Joaquin Delta levees, Geotech. Earthquake Eng. Soil Dyn. IV, 1-10, doi:10.1061/40975(318)182.

Shaw, J. H., Plesch, A., Tape, C., Suess, M. P, Jordan, T. H., Ely, G., Hauksson, E., Tromp, J., Tanimoto, T., Graves, R., Olsen, K., Nicholson, C., Maechling, P. J., Rivero, C., Lovely, P., Brankman, C. M, and Munster, J., 2015, Unified Structural Representation of the southern California crust and upper mantle. Earth Planet. Sci. Lett., 415, 1-15; http://scedc.caltech.edu/ research-tools/3d-velocity.html.

Teel, A. 2012, Seismic tomography of the Sacramento-San Joaquin River Delta : joint P-wave/gravity and ambient noise methods. PhD thesis, UW-Madison.

Thurber, C., Zhang, H., Brocher, T., and Langenheim, V., 2009, Regional three-dimensional seismic velocity model of the crust and uppermost mantle of northern California. J. Geophys. Res. Solid Earth, 114, B01304, doi:10.1029/2008JB005766.

Thurber, C.H., Brocher, T.M., Zhang, H. and Langenheim, V.E., 2007, Three-dimensional P wave velocity model for the San Francisco Bay region, California. J. Geophys. Res., Solid Earth, 112, B07313, doi:10.1029/2006JB004682.

Torres, R. A., Abrahamson, N., Brovold, F., Cosio, G., Driller, M., Harder, L., Marachi, D., Neudeck, C., O'Leary, L., Ramsbotham, M., and Seed, R., 2000, Seismic vulnerability of the Sacramento-San Joaquin Delta levees. CALFED Bay-Delta Program, Sacramento, California.

Zhang, H., Maceira, M., Roux, P. and Thurber, C., 2014, Joint inversion of body-wave arrival times and surface-wave dispersion for three-dimensional seismic structure around SAFOD. Pure Appl. Geophys., 171, doi:10.1007/s00024-014-0806-y.

Zhang, H., Thurber, C. H., 2003, Double-difference tomography: The method and its application to the Hayward fault, California. Bull. Seismol. Soc. Am., 93(5), 1,875–1,889.

4.13. Table and Figures

Table 4.1 Checkerboard RMS model recovery misfit for joint inversion and body-wave only inversion. The decrease percentage is the RMS for model recovery decrease from body-wave only inversion to joint inversion.

Checker Size (km)	Depth (km)	RMS [Vp, Joint] (km/s)	RMS [Vp, Body] (km/s)	Percentage change	RMS [Vs, Joint] (km/s)	RMS [Vs, Body] (km/s)	Percentage change
25	0	0.0214	0.0362	-40.88%	0.0226	0.0386	-41.45%
	1	0.0345	0.0382	-9.69%	0.0360	0.0431	-16.47%
	2	0.0463	0.0364	27.20%	0.0465	0.0385	20.78%
	4	0.0256	0.0229	11.79%	0.0261	0.0228	14.47%
	6	0.0256	0.0275	-6.91%	0.0266	0.0290	-8.28%
	8	0.0312	0.0326	-4.29%	0.0365	0.0380	-3.95%
	10	0.0323	0.0337	-4.15%	0.0391	0.0402	-2.74%
	12	0.0289	0.0317	-8.83%	0.0339	0.0367	-7.63%
	15	0.0249	0.0255	-2.35%	0.0287	0.0298	-3.69%
	20	0.0217	0.0233	-6.87%	0.0235	0.0253	-7.11%
40	0	0.0360	0.0710	-49.30%	0.0364	0.0749	-51.40%
	1	0.0497	0.0735	-32.38%	0.0529	0.0843	-37.25%
	2	0.0829	0.0637	30.14%	0.0857	0.0690	24.20%
	4	0.0426	0.0403	5.71%	0.0434	0.0402	7.96%
	6	0.0393	0.0505	-22.18%	0.0427	0.0541	-21.07%
	8	0.0528	0.0576	-8.33%	0.0600	0.0678	-11.50%
	10	0.0471	0.0485	-2.89%	0.0659	0.0692	-4.77%
	12	0.0375	0.0404	-7.18%	0.0601	0.0613	-1.94%
	15	0.0362	0.0398	-9.05%	0.0470	0.0490	-4.08%
	20	0.0458	0.0540	-15.19%	0.0502	0.0578	-13.15%

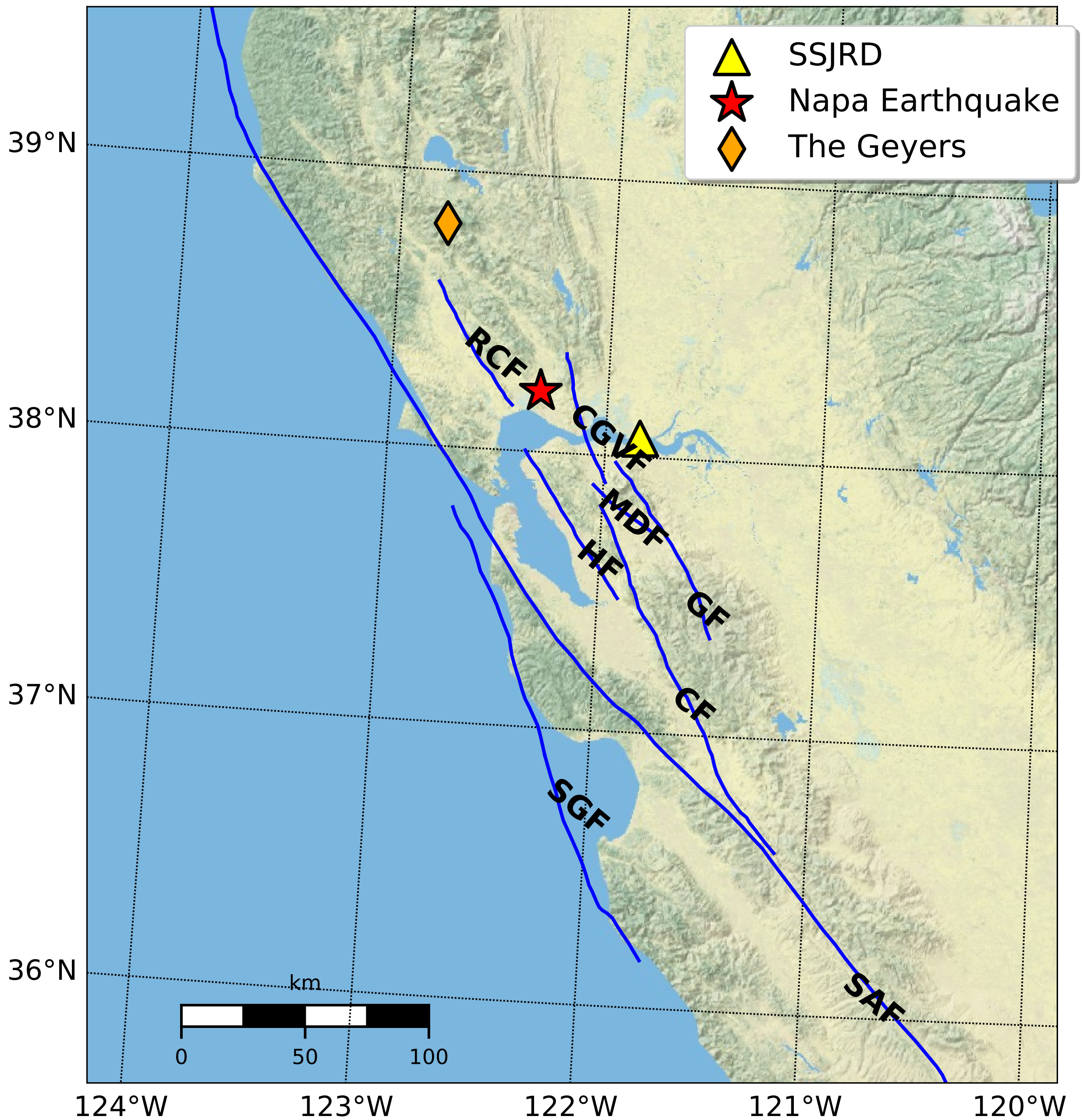


Figure 4.1 Map view of the San Francisco Bay Area. Major fault traces are marked with blue lines. SAF: San Andreas Fault. SGF, San Gregorio Fault. CF, Calaveras Fault. HF, Hayward Fault. MDF: Mr. Diablo Fault. GF: Greenville Fault. CGVF: Concord – Green Valley Fault. RCF: Rodgers Creek Fault. Yellow triangle: Sacramento-San Joaquin River Delta (SSJRD). Red star: epicenter of the 2014 South Napa earthquake. Orange diamond: The Geysers.

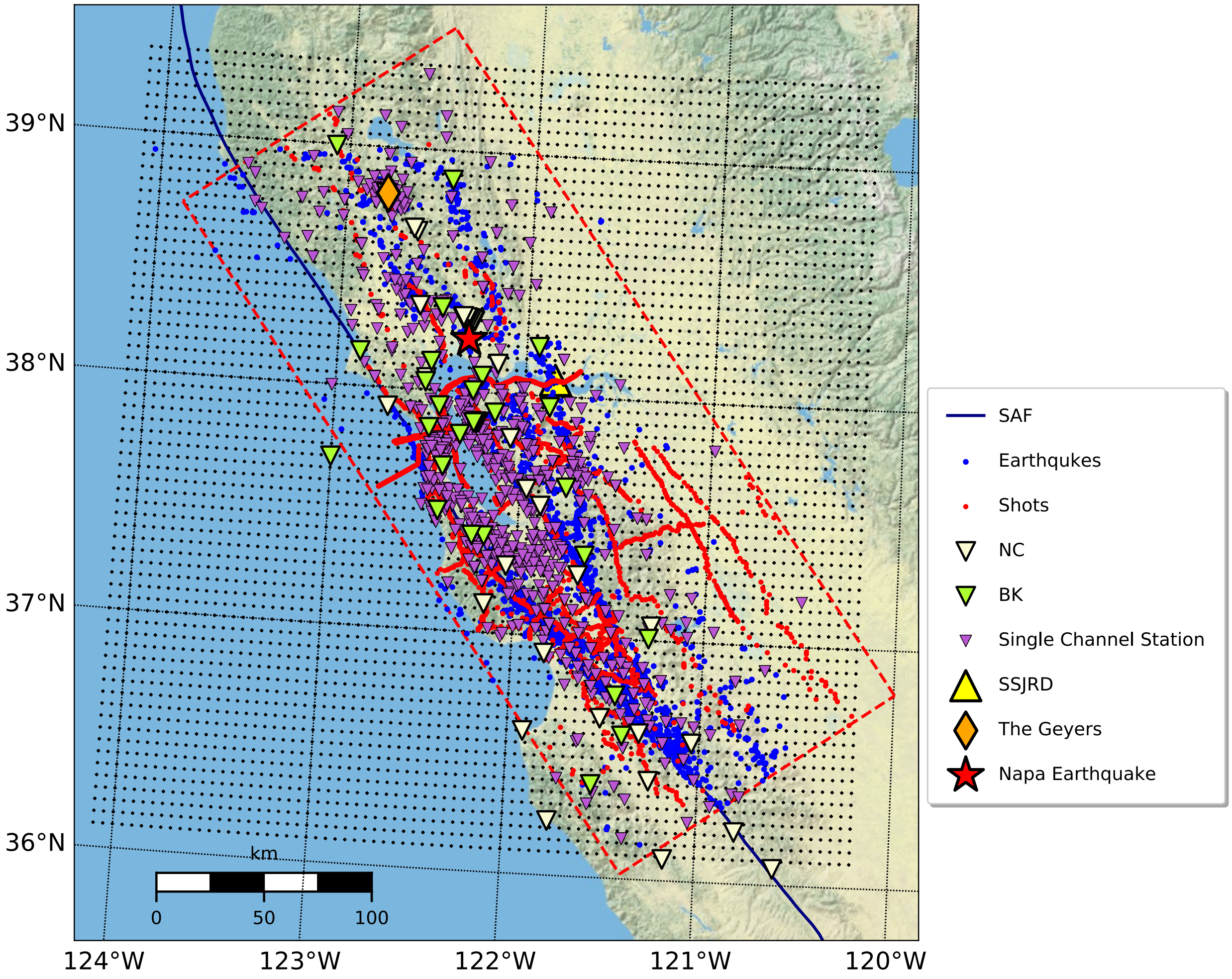


Figure 4.2 Map view of the events and stations. Red dots, active source events. Blue dots, earthquake events. Yellow triangle: Sacramento-San Joaquin River Delta (SSJRD). Red star: epicenter of the 2014 South Napa earthquake. Orange diamond: The Geysers. Purple flipped triangle: single channel stations. Green flipped triangle: three-component stations from the Northern California seismic network (NC). Light yellow flipped triangle: three-component stations from the Berkeley Digital Seismic Network (BK). The SAF is marked with a black line. Black dots: the nodes used in the joint inversion. Red dashed box: regions projected onto Cartesian coordinates used in Thurber et al. (2009).

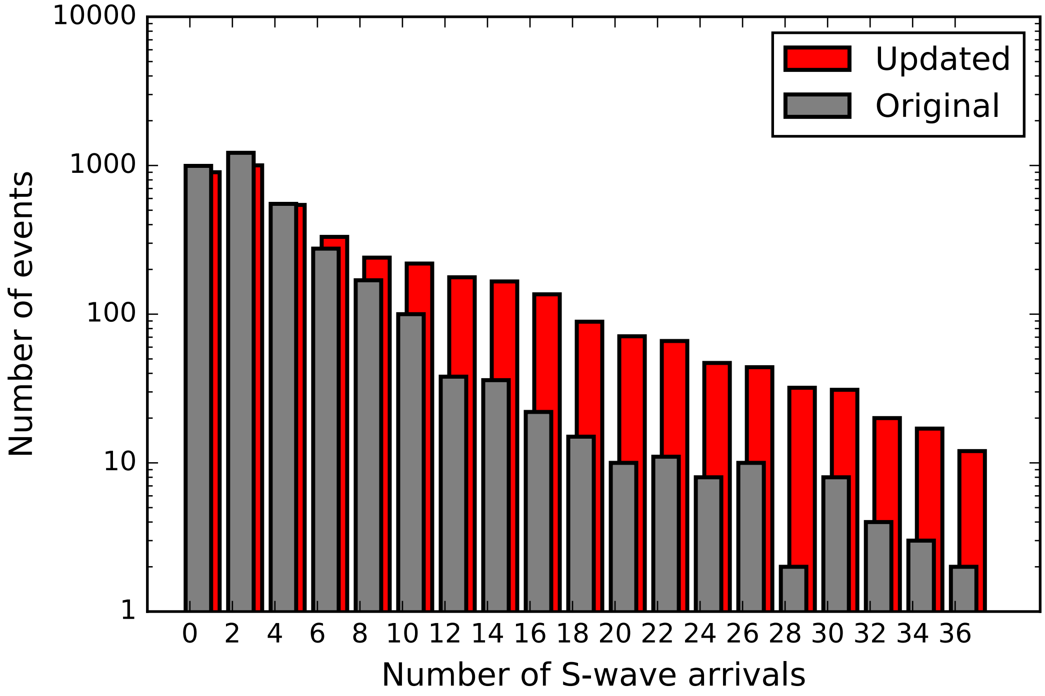
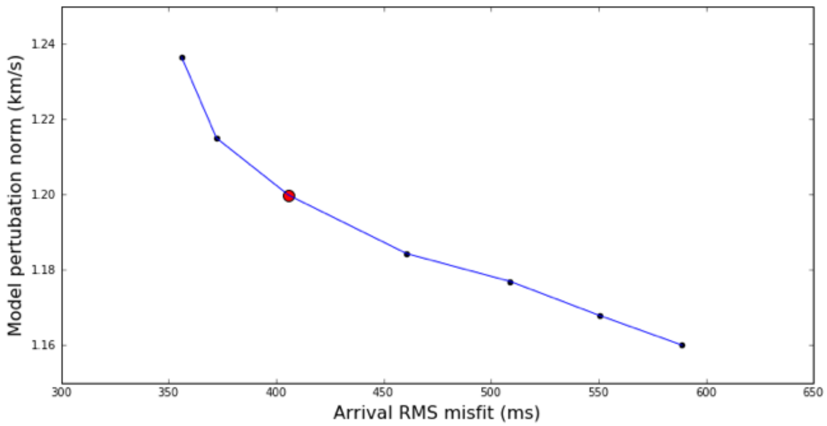
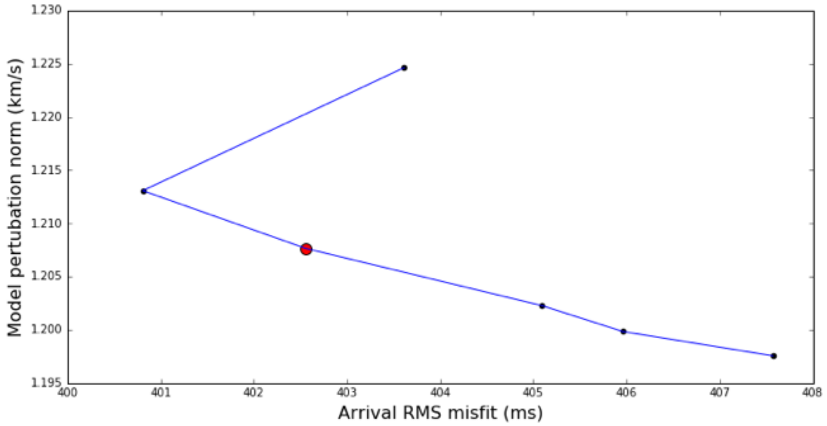


Figure 4.3 Histogram of number of S-wave arrivals for each event. Grey bins: original dataset. Red bins: updated dataset with auto-picked S-wave arrivals.

(a)


Model pertubation norm (km/s)
1.24
1.22
1.20
1.18
1.16
300
350
400
450
500
550
600
650
Arrival RMS misfit (ms)

(b)


Model pertubation norm (km/s)
1.230
1.225
1.220
1.215
1.210
1.205
1.200
1.195
400
401
402
403
404
405
406
407
408
Arrival RMS misfit (ms)

(c)

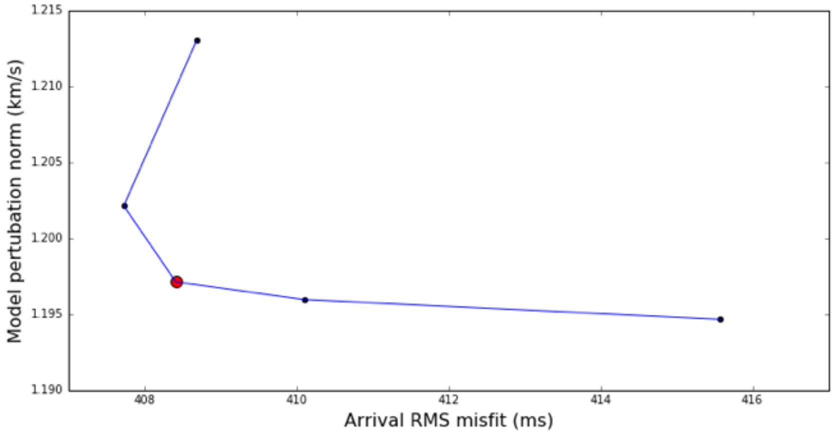
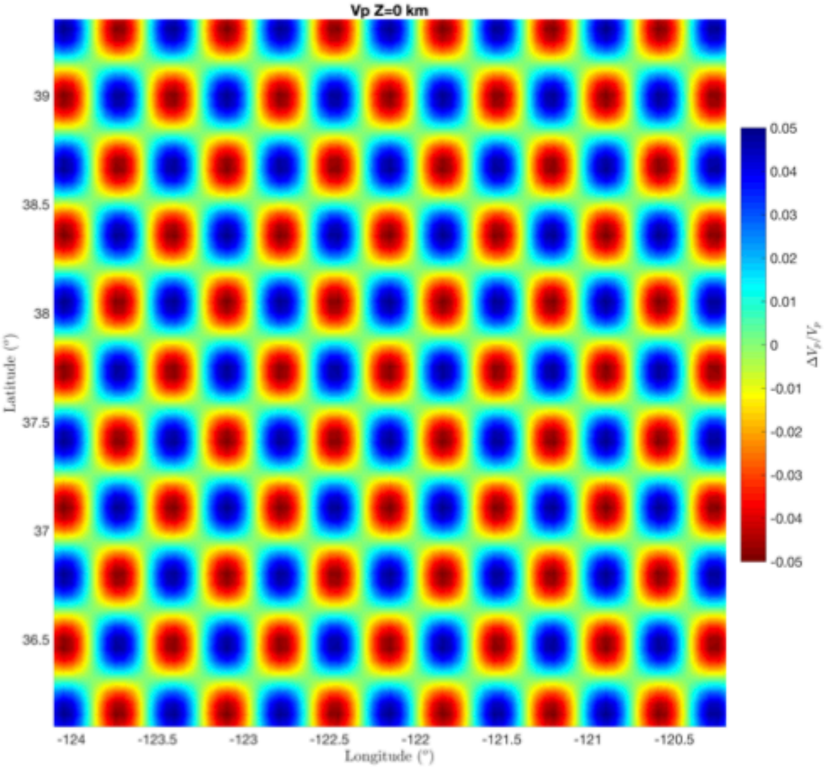
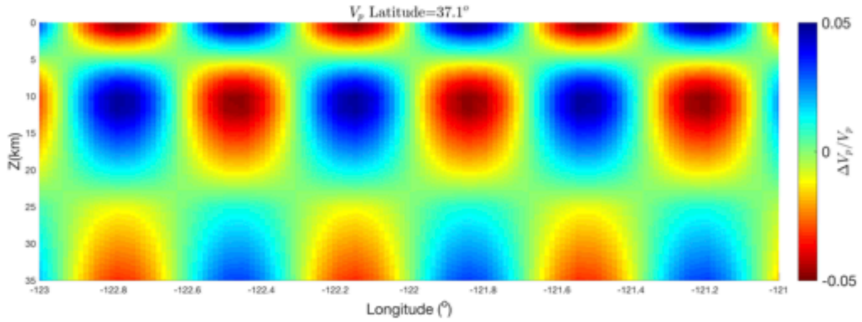


Figure 4.4 L-curves for trade-off analysis. (a) Fixed ratio (2:1) between horizontal smoothing parameters and the vertical smoothing parameters. Values tested for horizontal smoothing parameters are 20, 50, 100, 200, 300, 400, and 500. 100 is selected and marked with a red circle. (b) Fixed horizontal smoothing parameter value of 100 with varying vertical smoothing parameters (10, 20, 30, 40, 50, and 60). 30 is selected and marked with a red circle. (c) Fixed body wave smoothing parameters, testing smoothing parameters for surface wave. 60 is selected from 20, 40, 60, 100, and 200. A red circle shows the point associated with 60.

(a)


Vp Z=0 km
39
38.5
38
37.5
37
36.5
Latitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (°)
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
$\Delta V_p/V_p$

(b)


V_p Latitude=37.1°
0
5
10
15
20
25
30
35
Z(km)
-123
-122.8
-122.6
-122.4
-122.2
-122
-121.8
-121.6
-121.4
-121.2
-121
Longitude (°)
0.05
0
-0.05
$\Delta V_p/V_p$

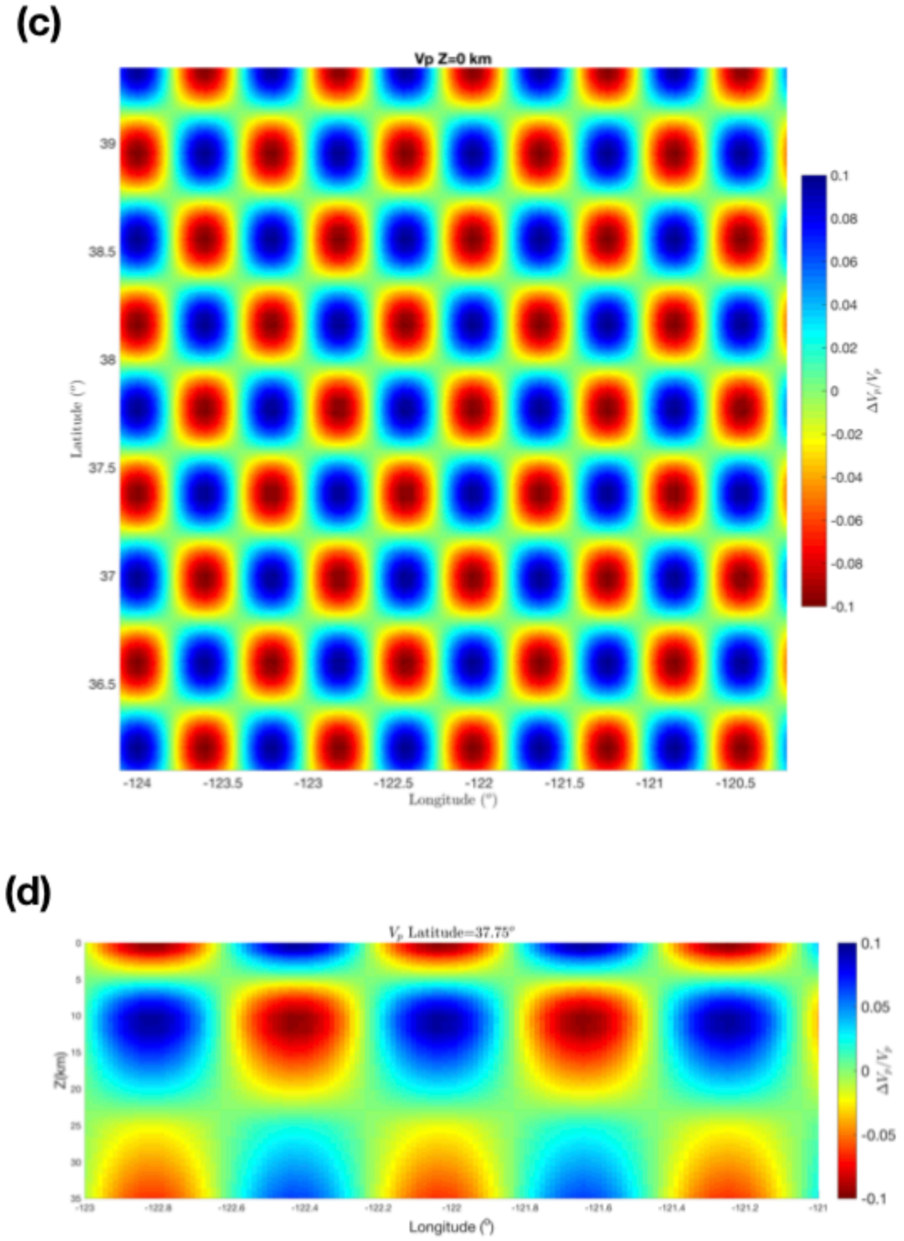
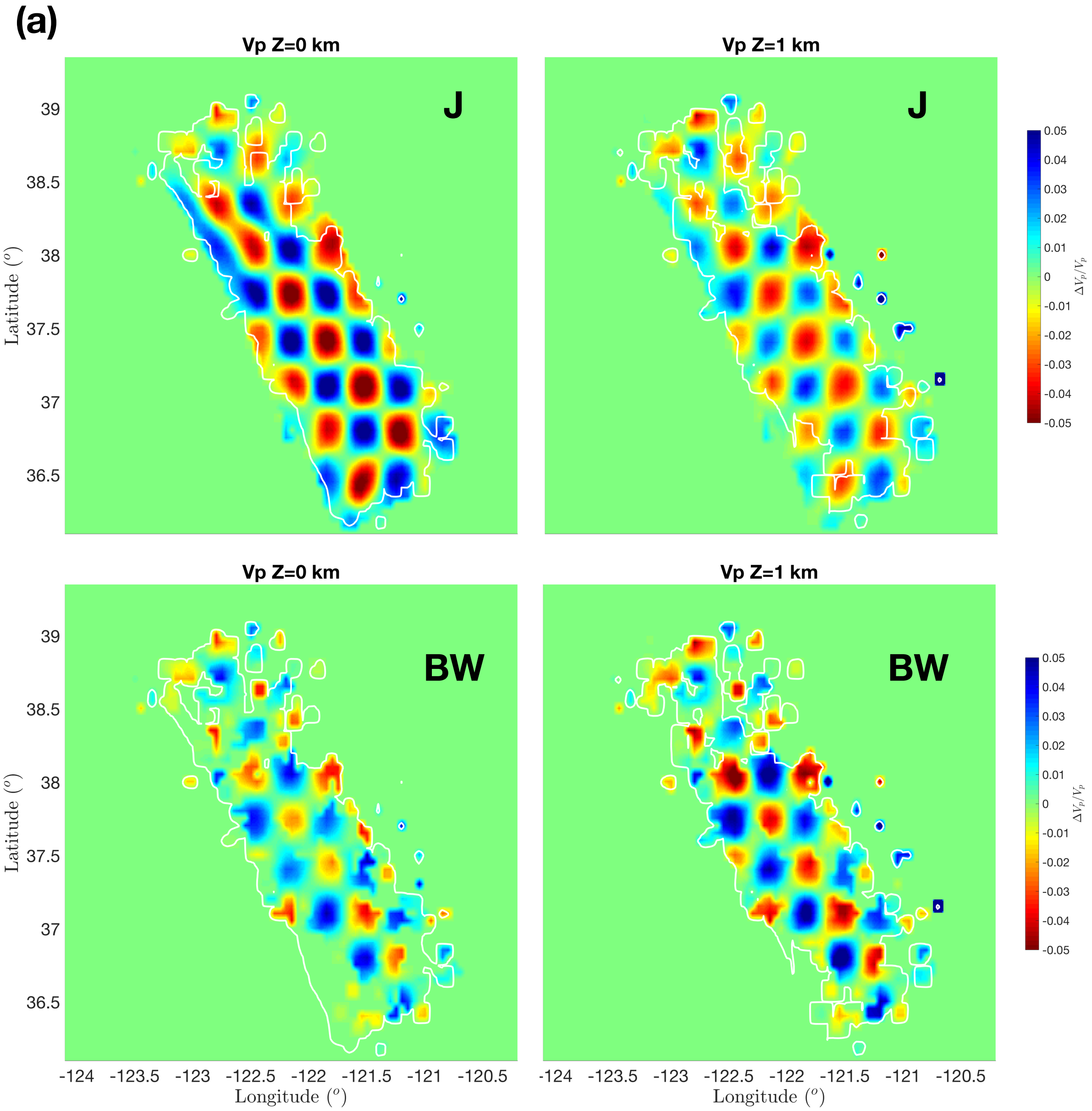
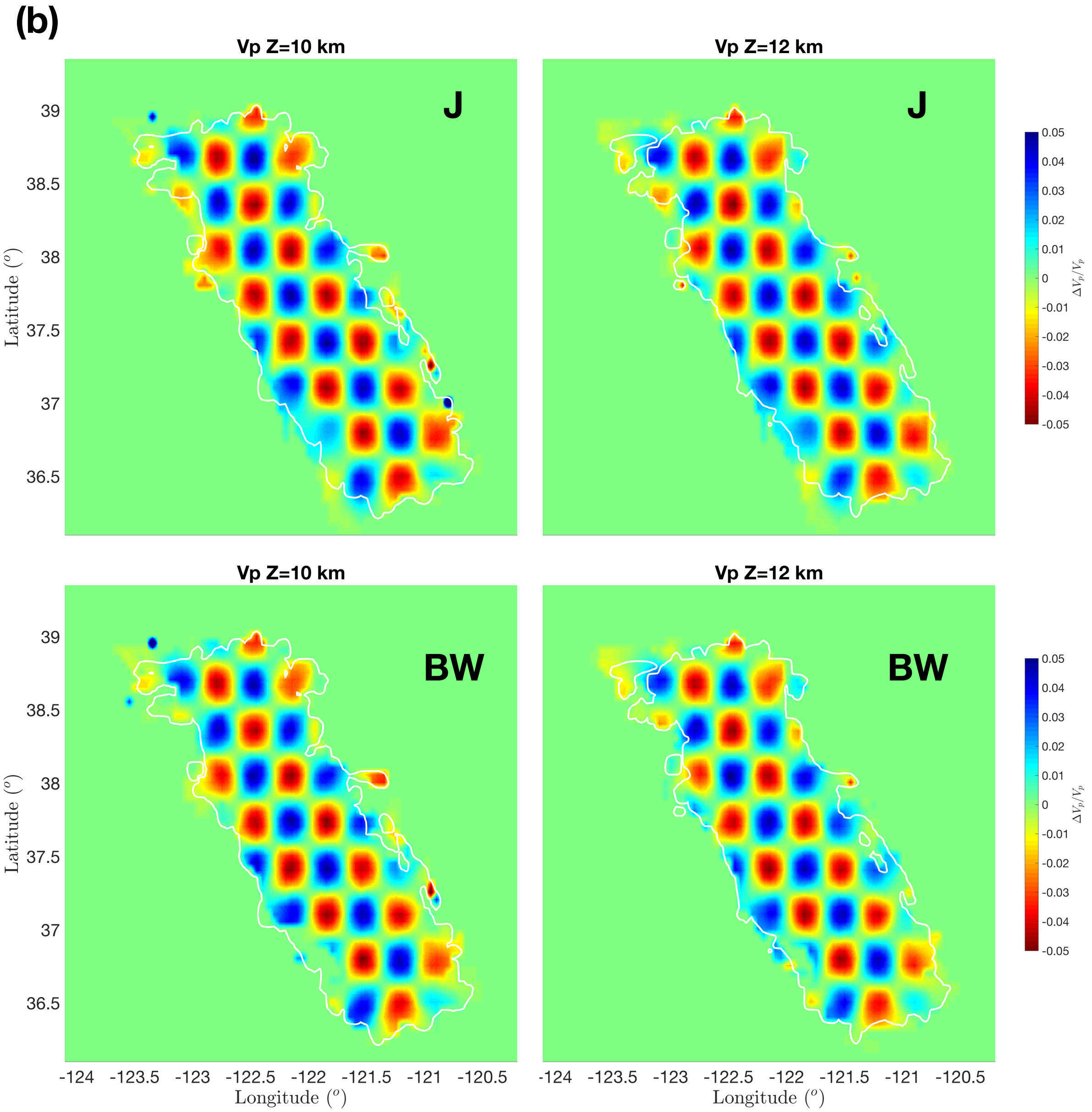


Figure 4.5 Checkerboard anomalies for the two sets of checkerboard tests. (a), (b) Map view at Z = 0 km and cross-section view at latitude = 37.1° for the V_P anomalies with horizontal dimension of ~25 km. The maximum anomaly amplitude is 5%. (c), (d) Map view at Z = 0 km and cross-section view at latitude = 37.75° for the V_P anomalies with horizontal dimension of ~40 km. The maximum anomaly amplitude is 10%.


(a)
Vp Z=0 km
Vp Z=1 km
J
J
Vp Z=0 km
Vp Z=1 km
BW
BW
Latitude (^{o})
39
38.5
38
37.5
37
36.5
Longitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
$\Delta V_p/V_p$
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05



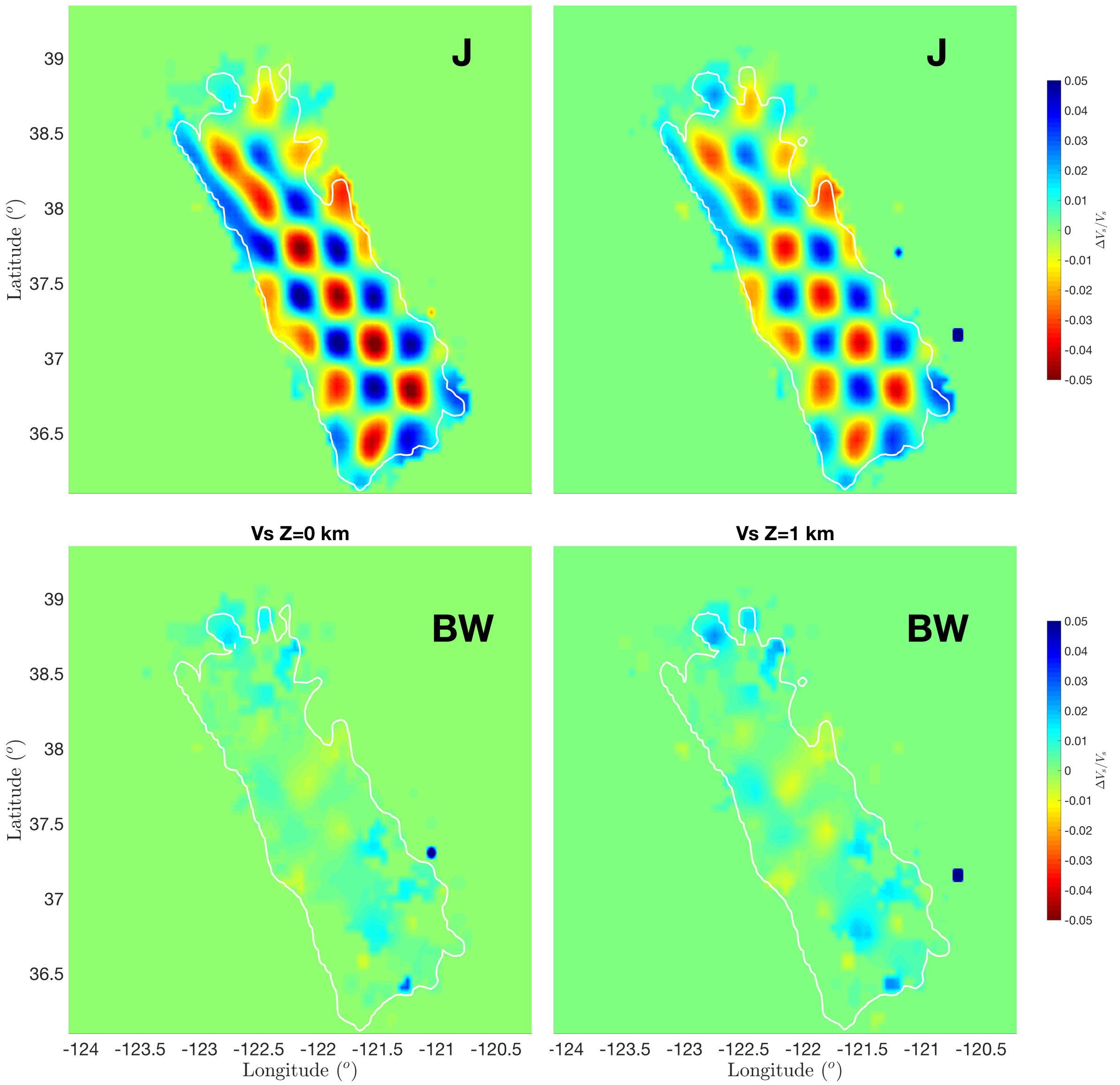
Vp Z=10 km

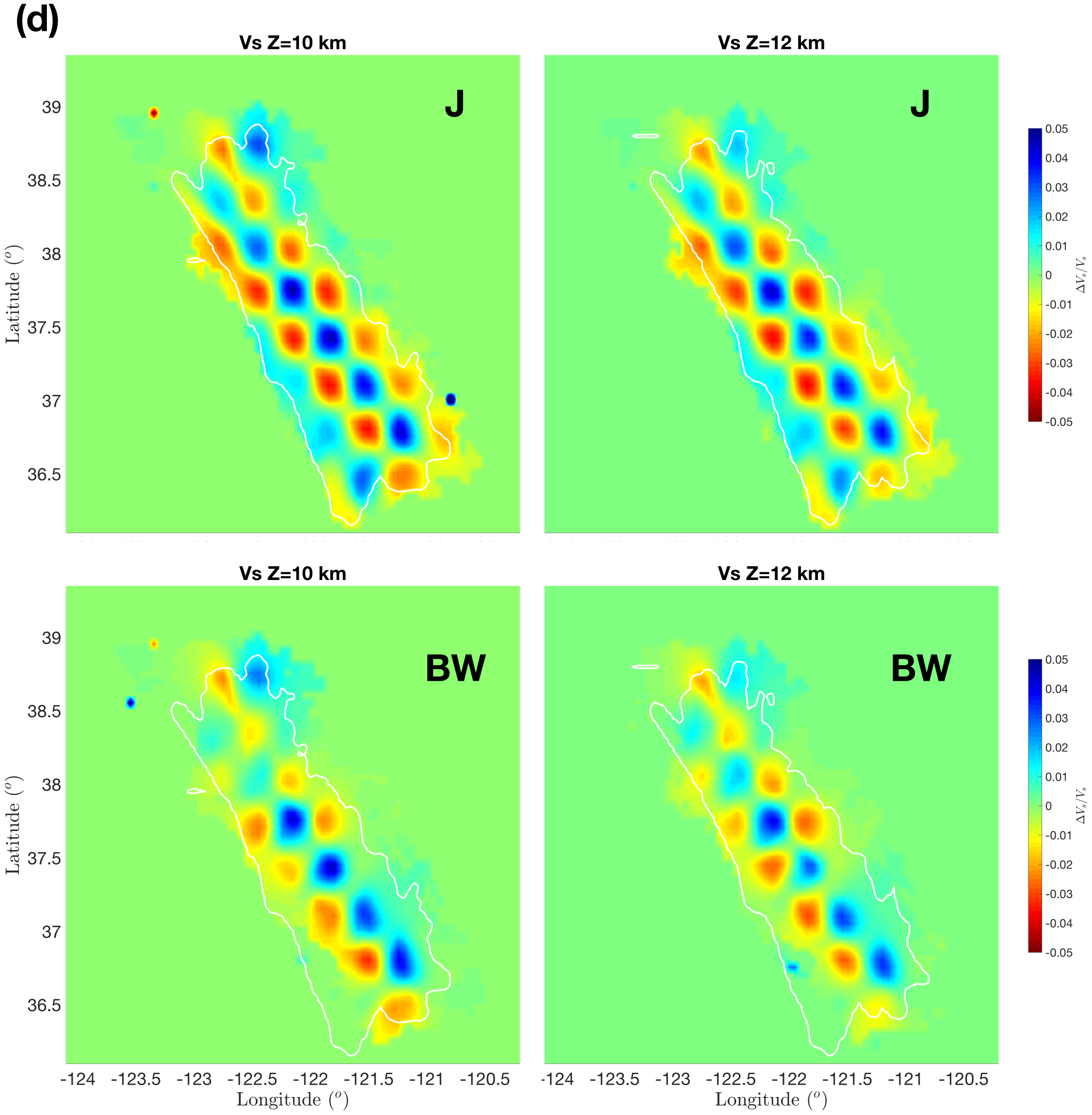
Vp Z=12 km

Vp Z=10 km

Vp Z=12 km

(c)


Vs Z=0 km
Vs Z=1 km
J
J
Vs Z=0 km
Vs Z=1 km
BW
BW
Latitude (^{o})
39
38.5
38
37.5
37
36.5
Longitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
$\Delta V_s/V_s$
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05


(d)
Vs Z=10 km
Vs Z=12 km
J
J
Vs Z=10 km
Vs Z=12 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
$\Delta V_s/V_s$

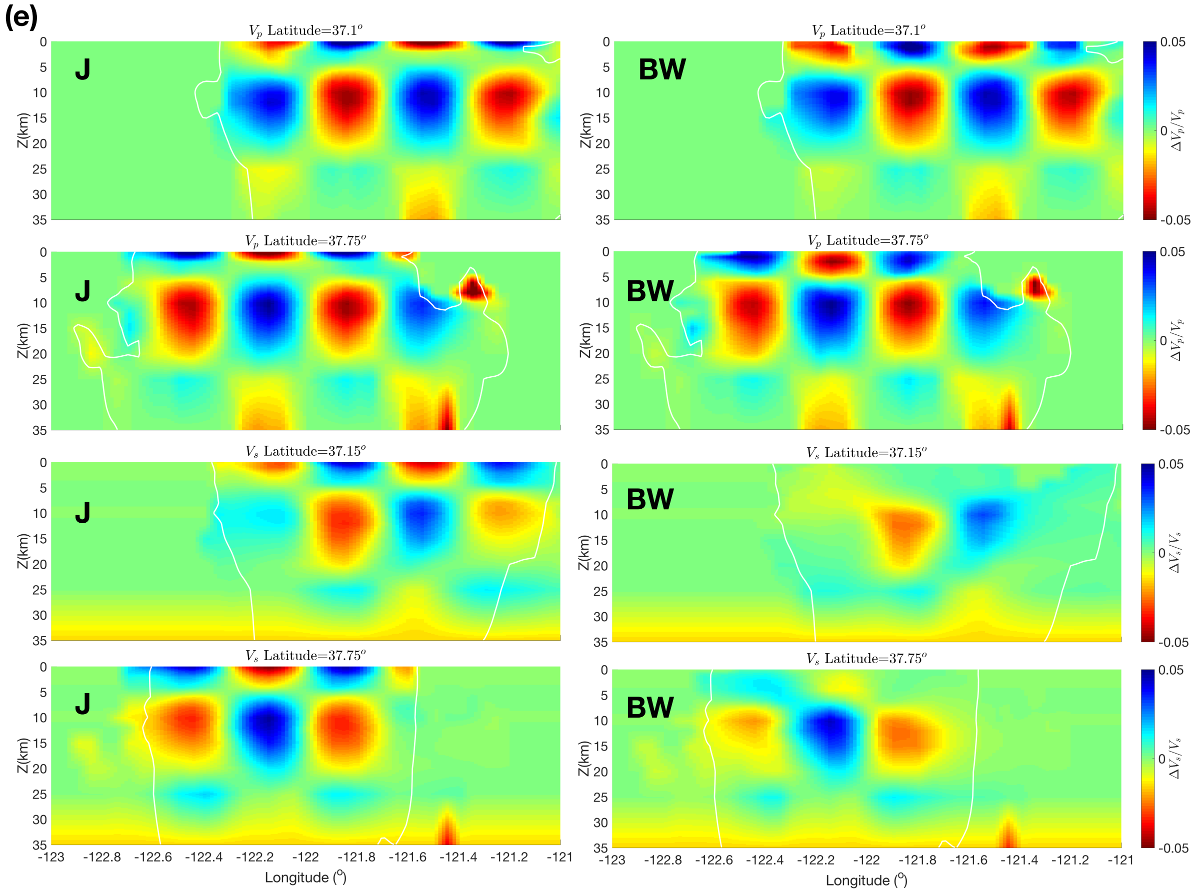
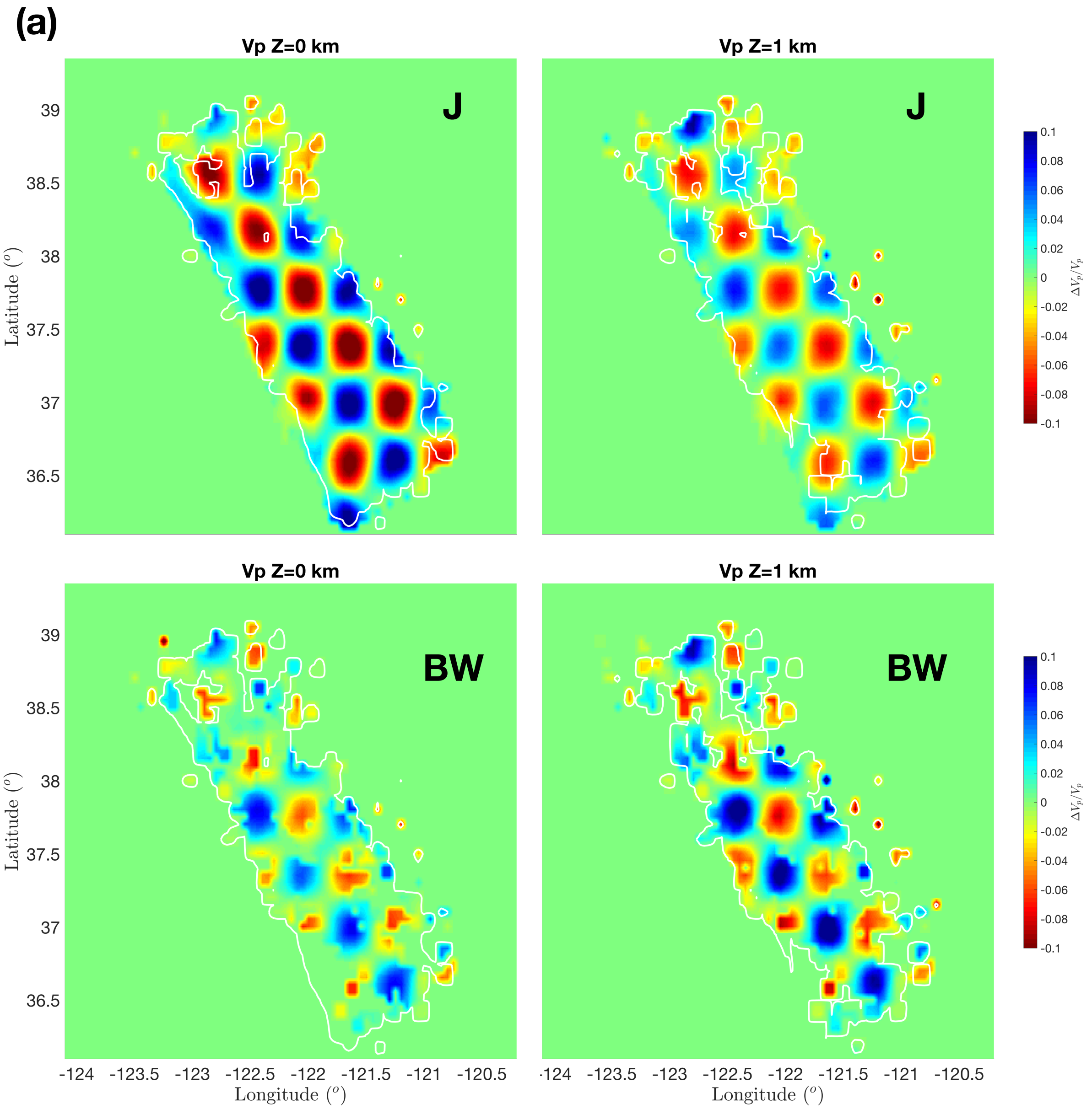
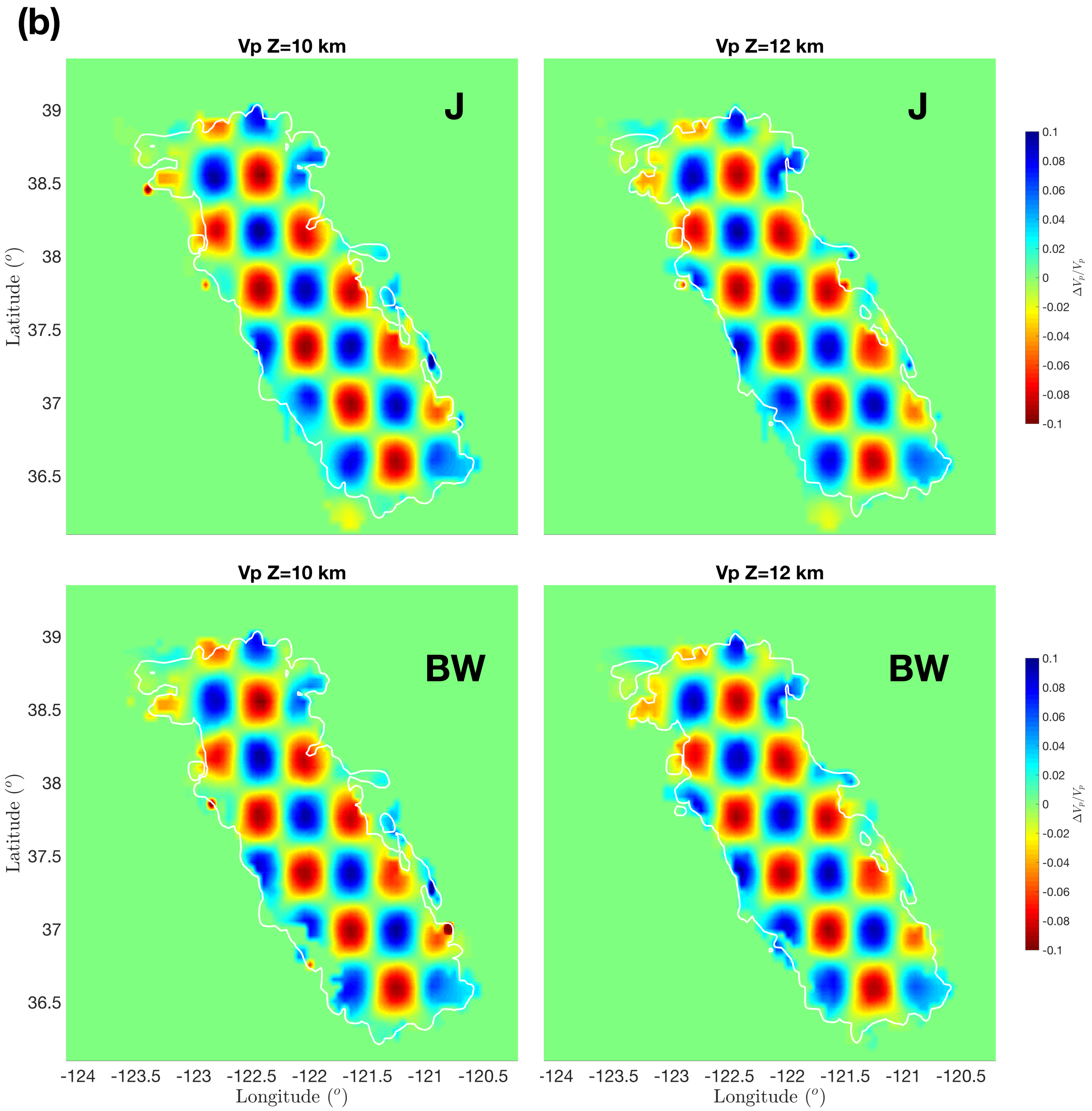
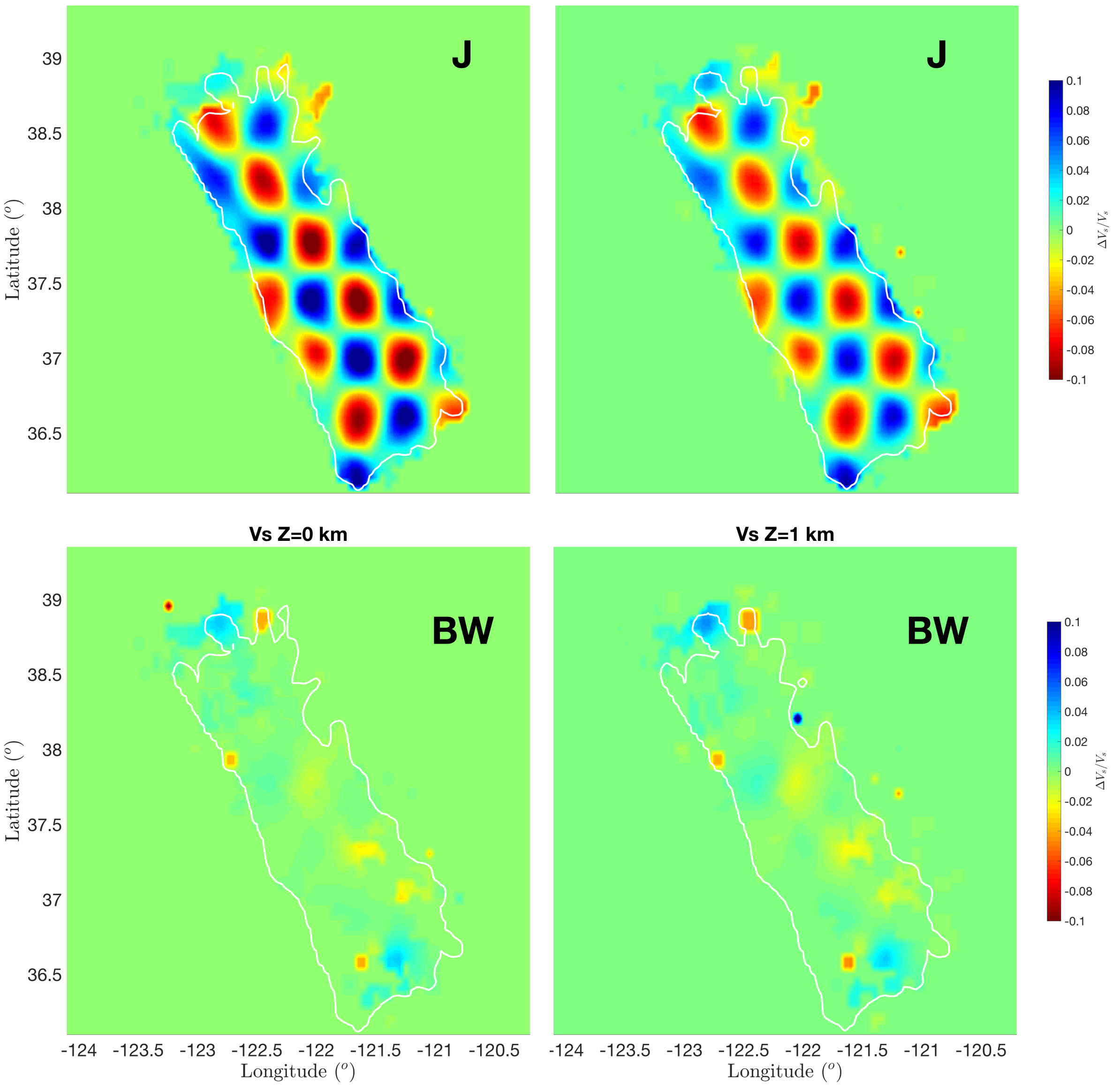


Figure 4.6 Checkerboard test results for anomalies with ~25 km horizontal dimension. For each four-plot group (a), (b), (c), (d), the first row shows the joint inversion results (J) and the second row shows the body-wave only inversion results (BW). (a) V_P results at Z = 0 and 1 km depths. (b) V_P results at Z = 10 and 12 km depths. (c) V_S results at Z = 0 and 1 km depths. (d) V_S results at Z = 10 and 12 km depths. (e) cross-sections for latitude = 37.15° and 37.75°, left column shows the joint inversion results and right column shows the body-wave only results; top two panels are V_P, bottom two panels are V_S. White contours are shown for P-wave model with DWS = 36 and S-wave model for DWS = 16.


(a)
Vp Z=0 km
Vp Z=1 km
J
J
Vp Z=0 km
Vp Z=1 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
$\Delta V_p/V_p$


(b)
Vp Z=10 km
Vp Z=12 km
J
J
Vp Z=10 km
Vp Z=12 km
BW
BW
Latitude (°)
39
38.5
38
37.5
37
36.5
Longitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
$\Delta V_p/V_p$
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1

(c)

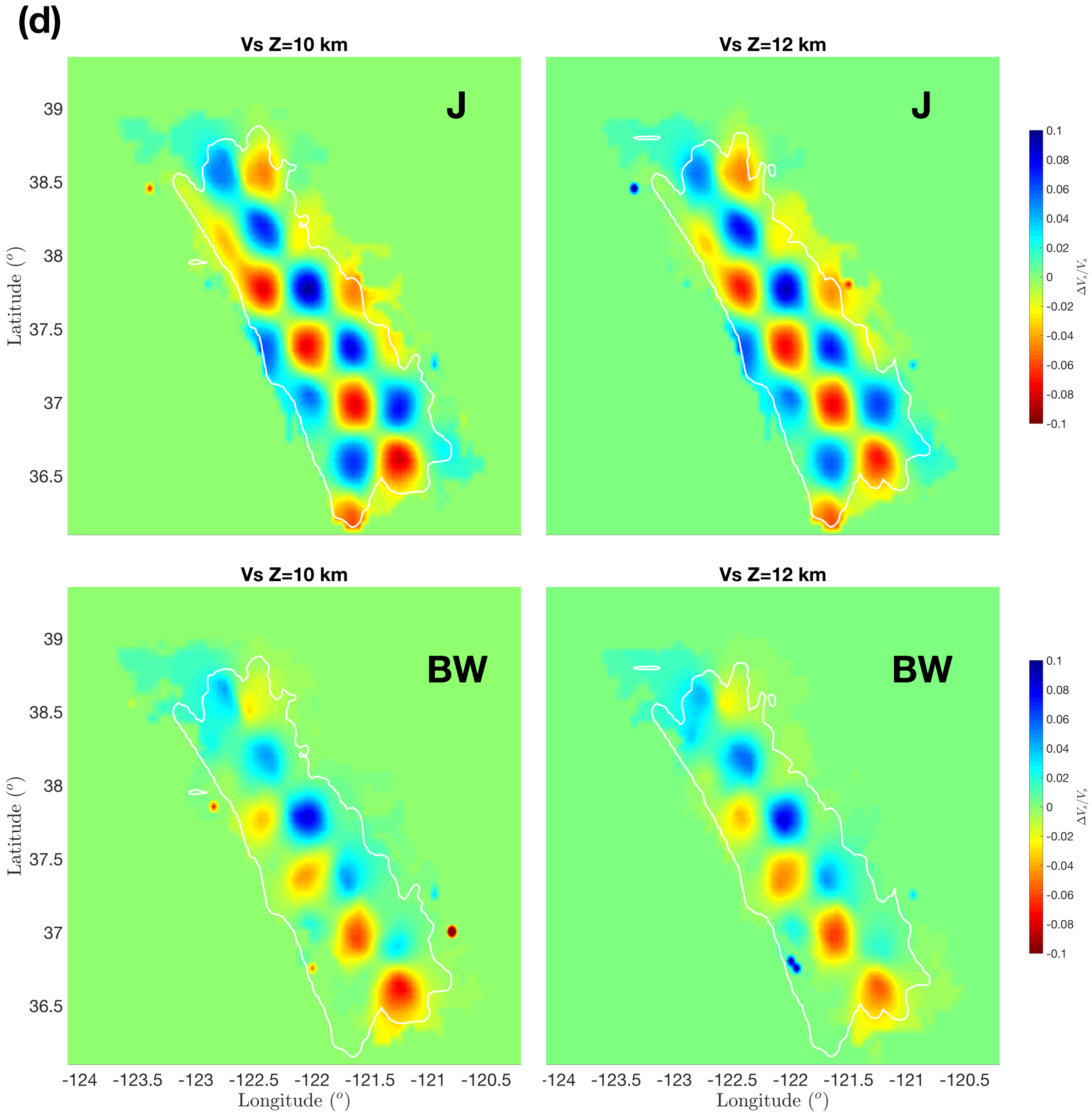


Vs Z=0 km

Vs Z=1 km

Vs Z=0 km

Vs Z=1 km


(d)
Vs Z=10 km
Vs Z=12 km
J
J
Vs Z=10 km
Vs Z=12 km
BW
BW
Latitude (°)
39
38.5
38
37.5
37
36.5
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
$\Delta V_s/V_s$
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (°)

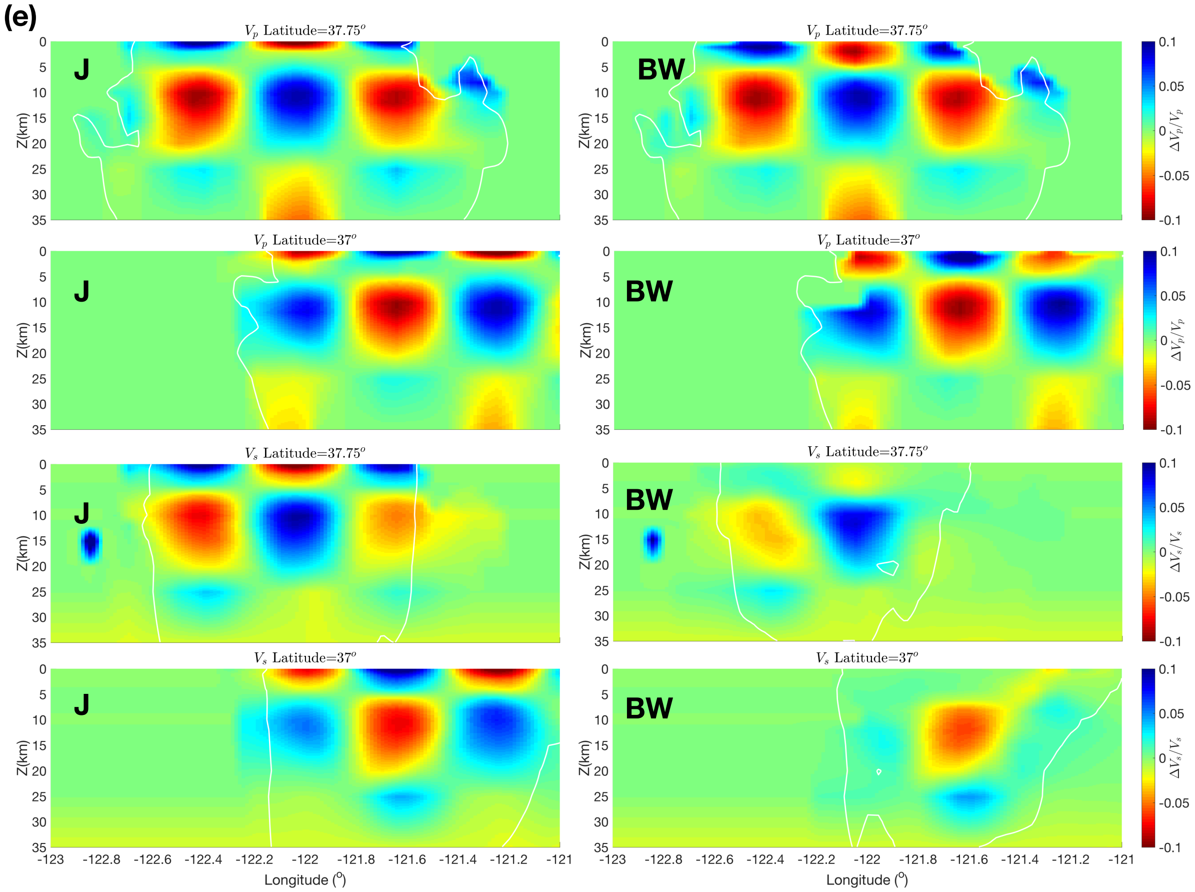
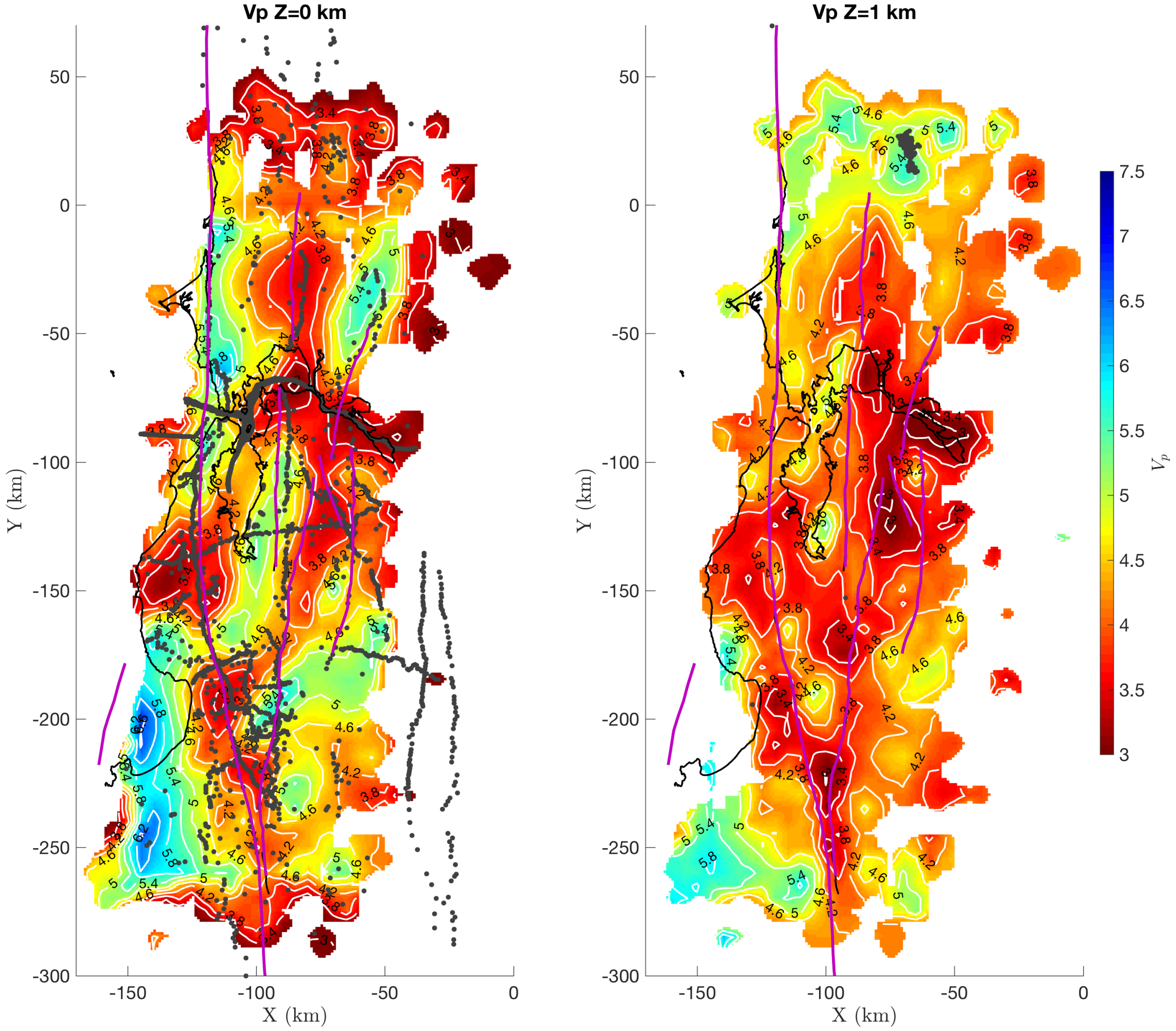
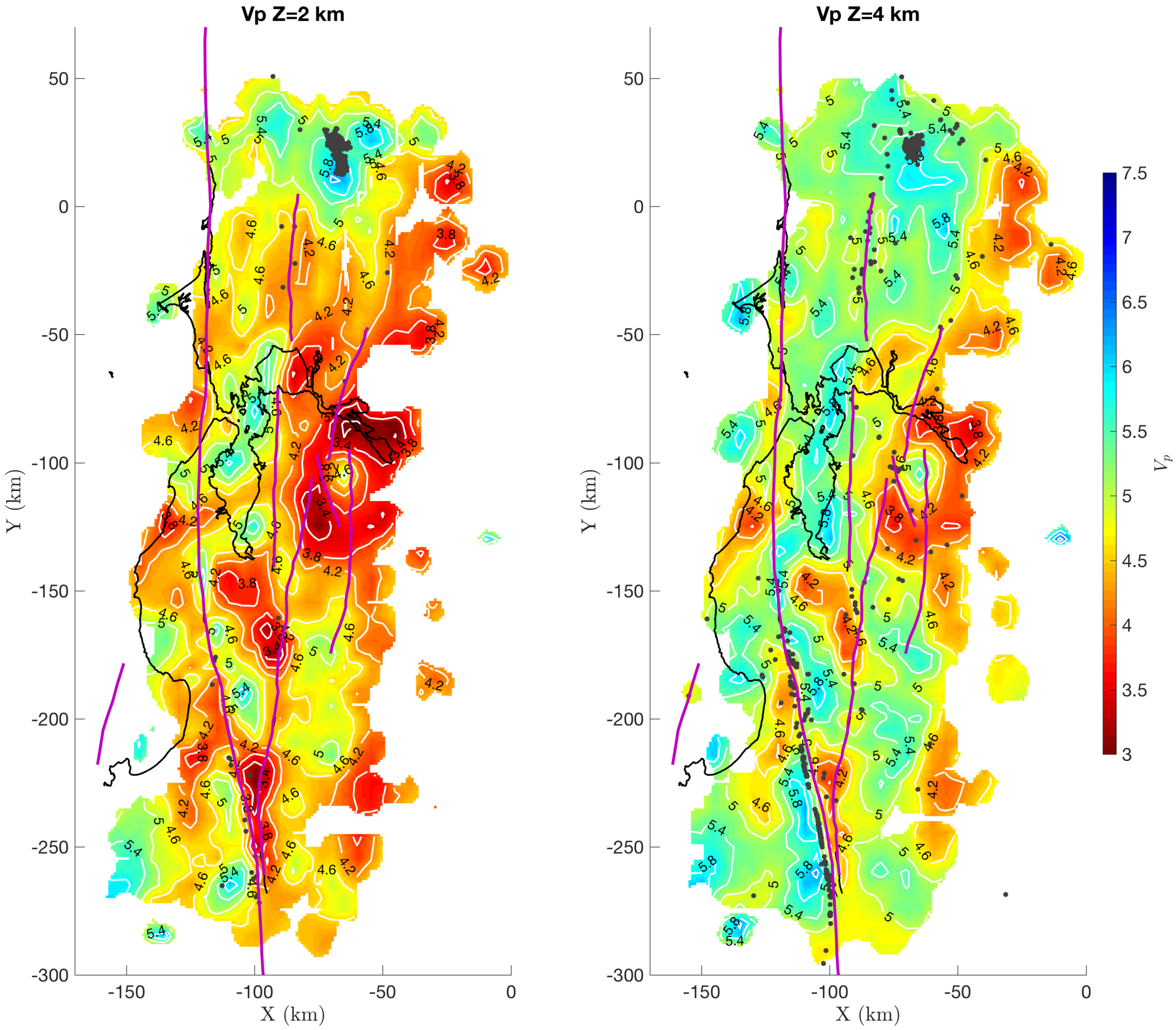
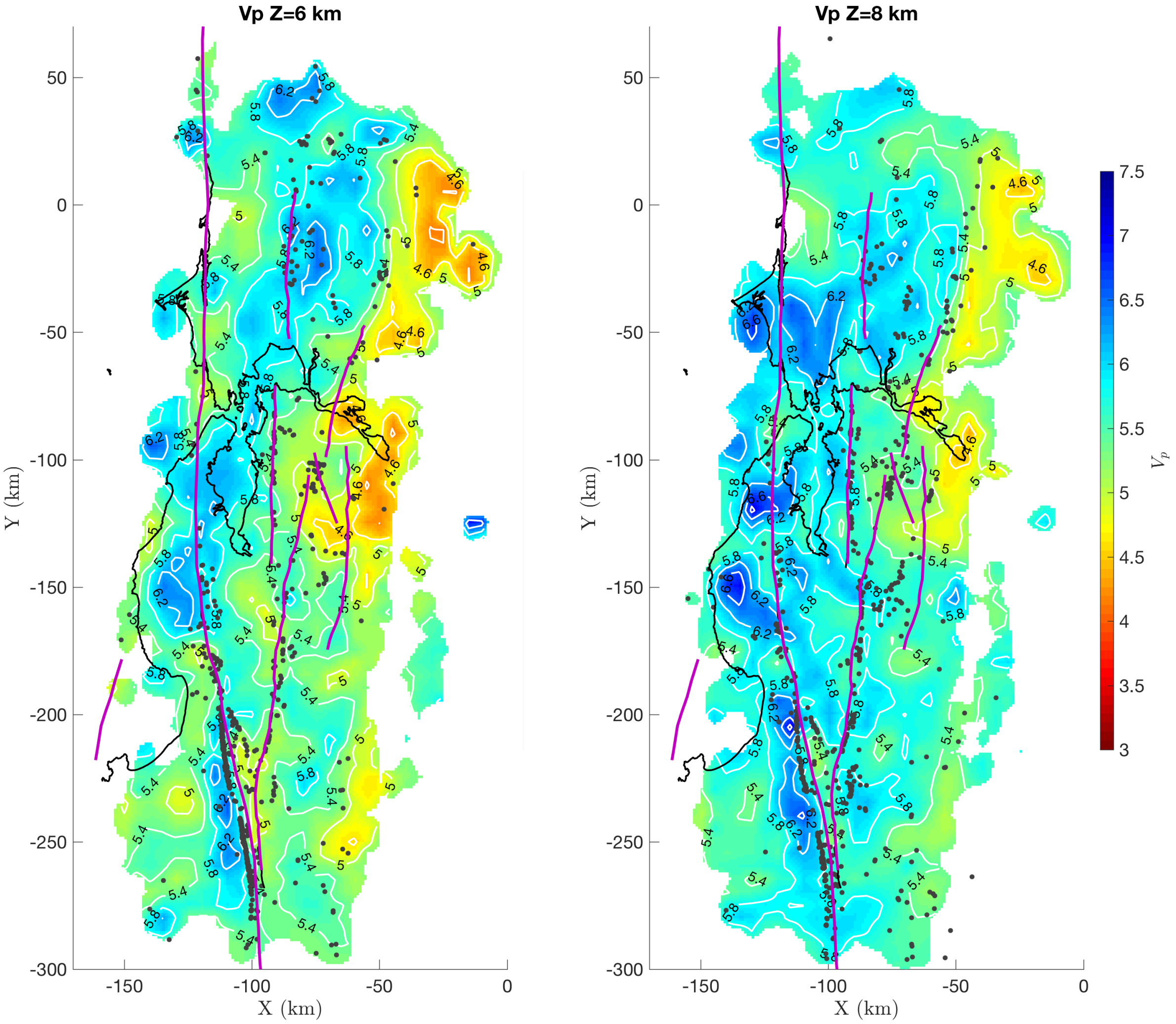
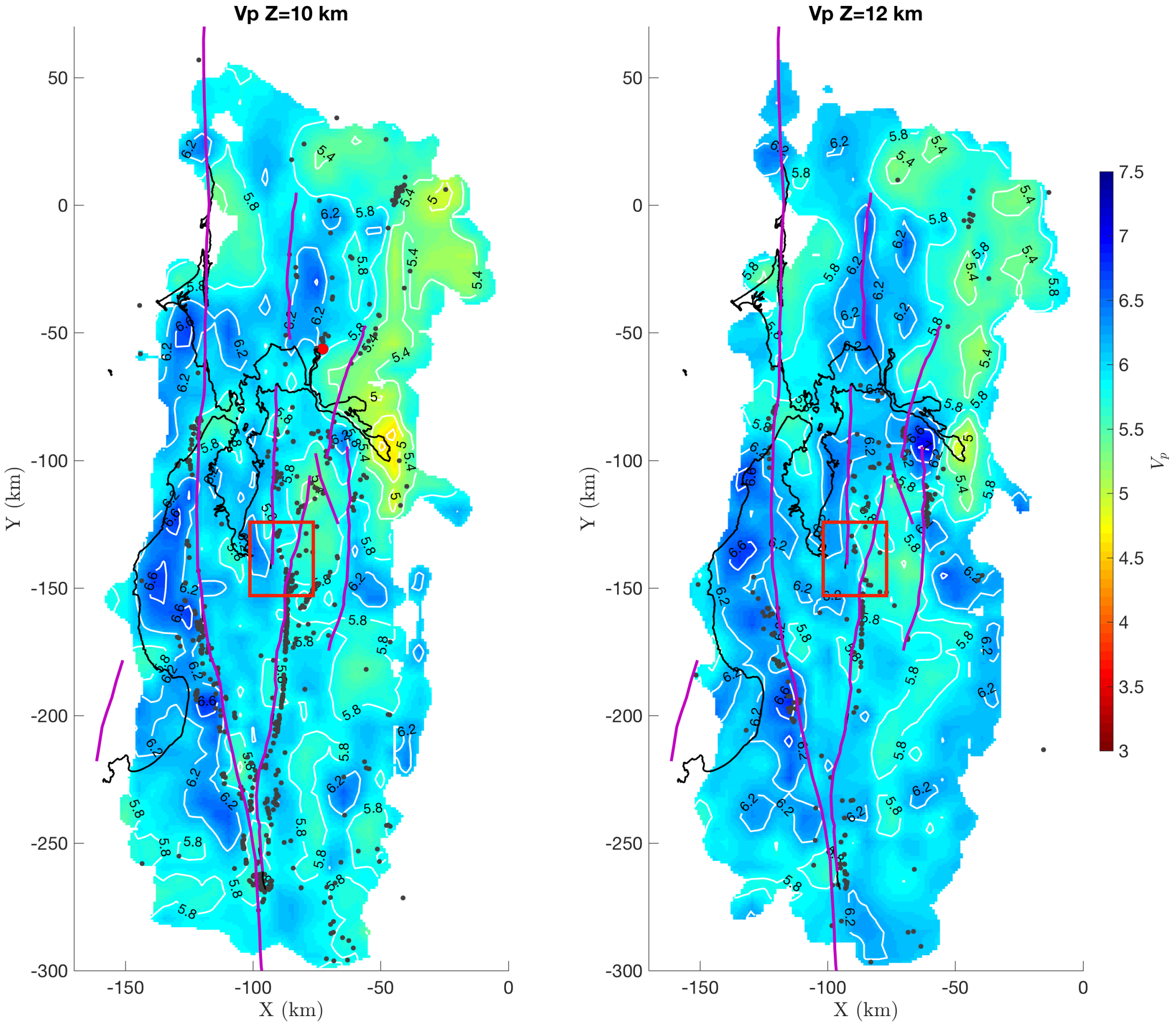


Figure 4.7 Checkerboard test results for anomalies with ~40 km horizontal dimension. For each four-plot group (a), (b), (c), (d), the first row shows the joint inversion results (J) and the second row shows the body-wave only inversion results (BW). (a) V_P results at Z = 0 and 1 km depths. (b) V_P results at Z = 10 and 12 km depths. (c) V_S results at Z = 0 and 1 km depths. (d) V_S results at Z = 10 and 12 km depths. (e) cross-sections for latitude = 37° and 37.75°, left column shows the joint inversion results and right column shows the body-wave only results; top two panels are V_P, bottom two panels are V_S. White countours are shown for P-wave model with DWS = 36 and S-wave model with DWS = 16.


Vp Z=0 km
Vp Z=1 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
V_p
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3


Vp Z=2 km
Vp Z=4 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
V_p
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3


Vp Z=6 km
Vp Z=8 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
V_p
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3


Vp Z=10 km
Vp Z=12 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
V_p
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3

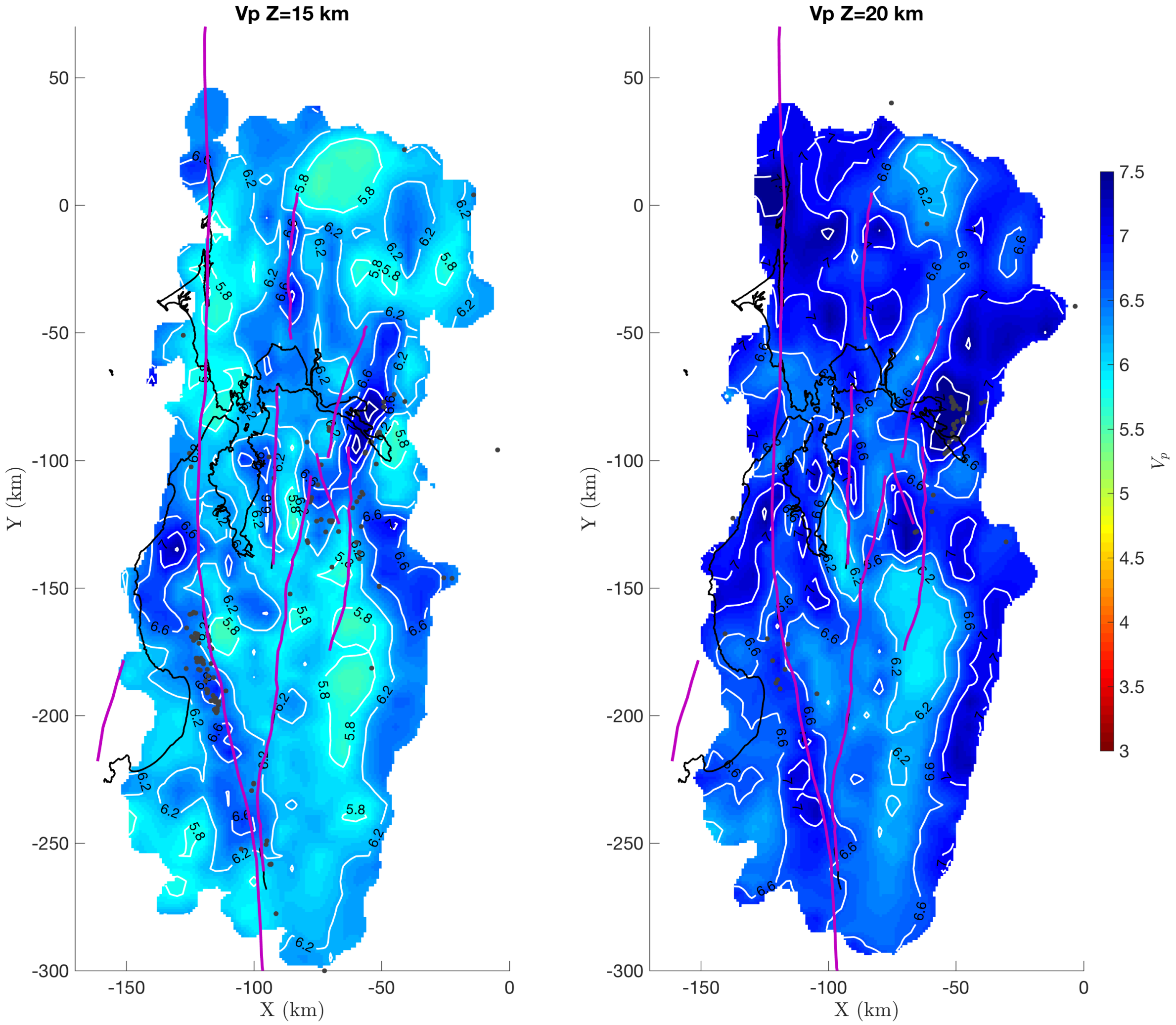
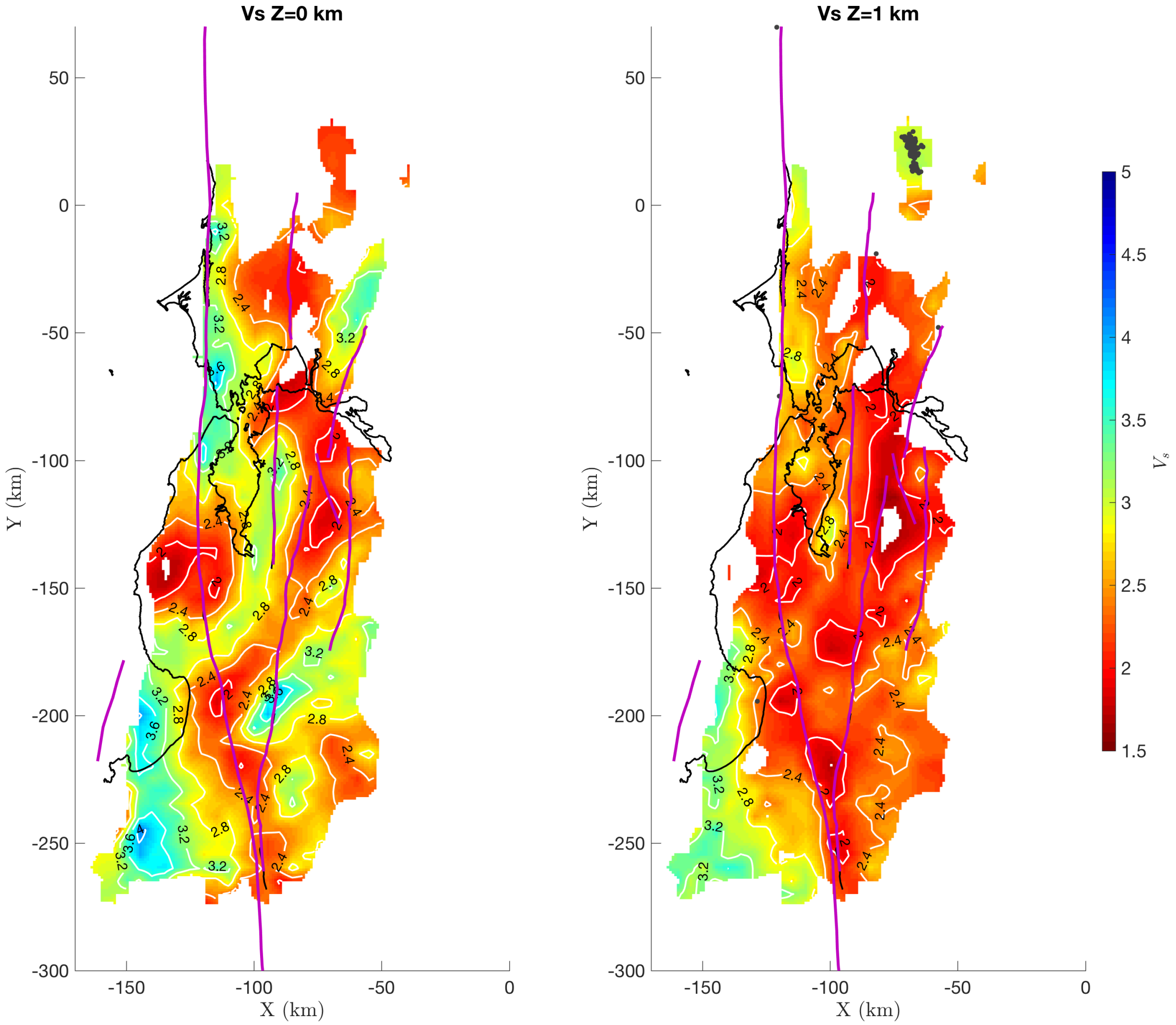
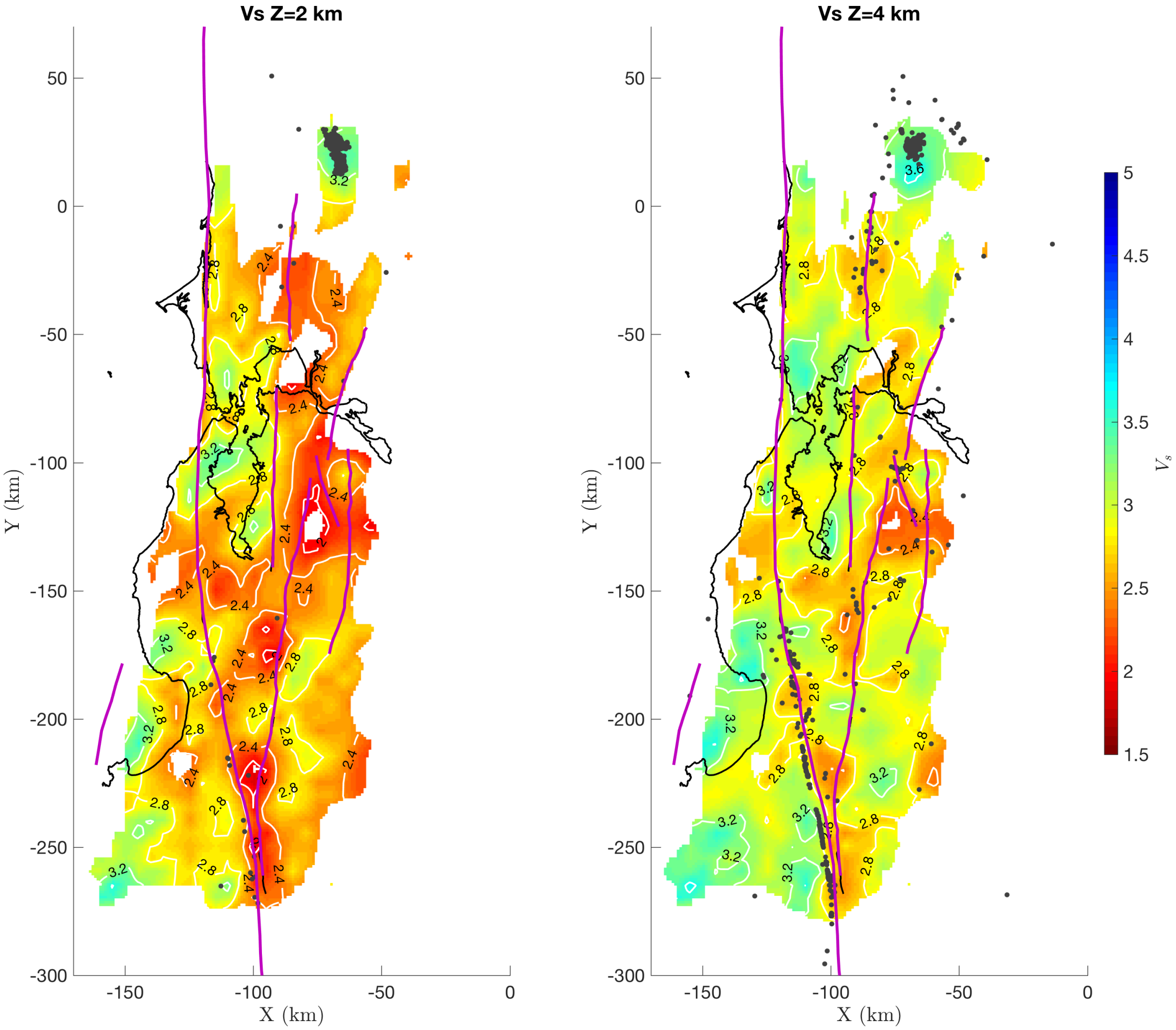
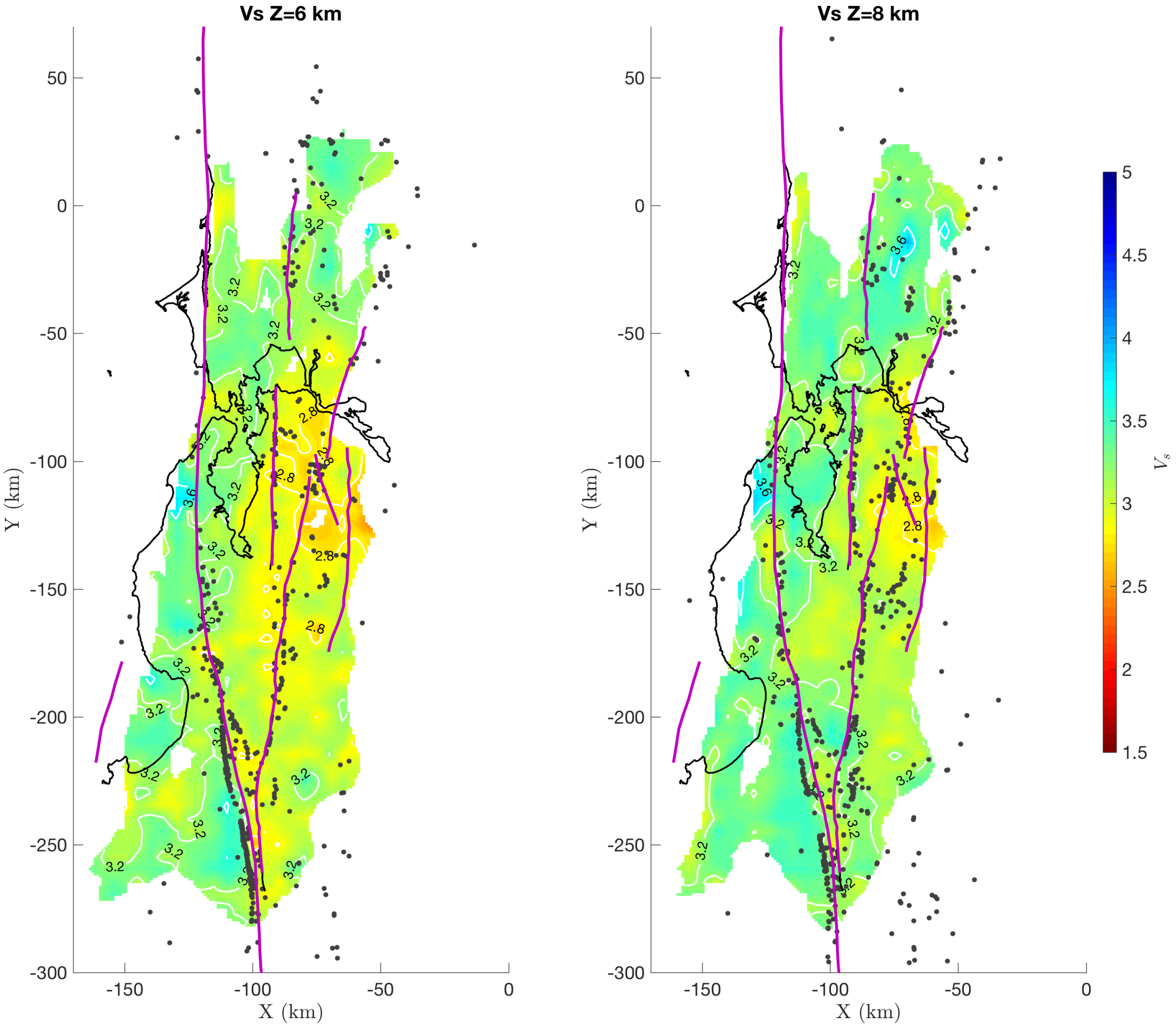
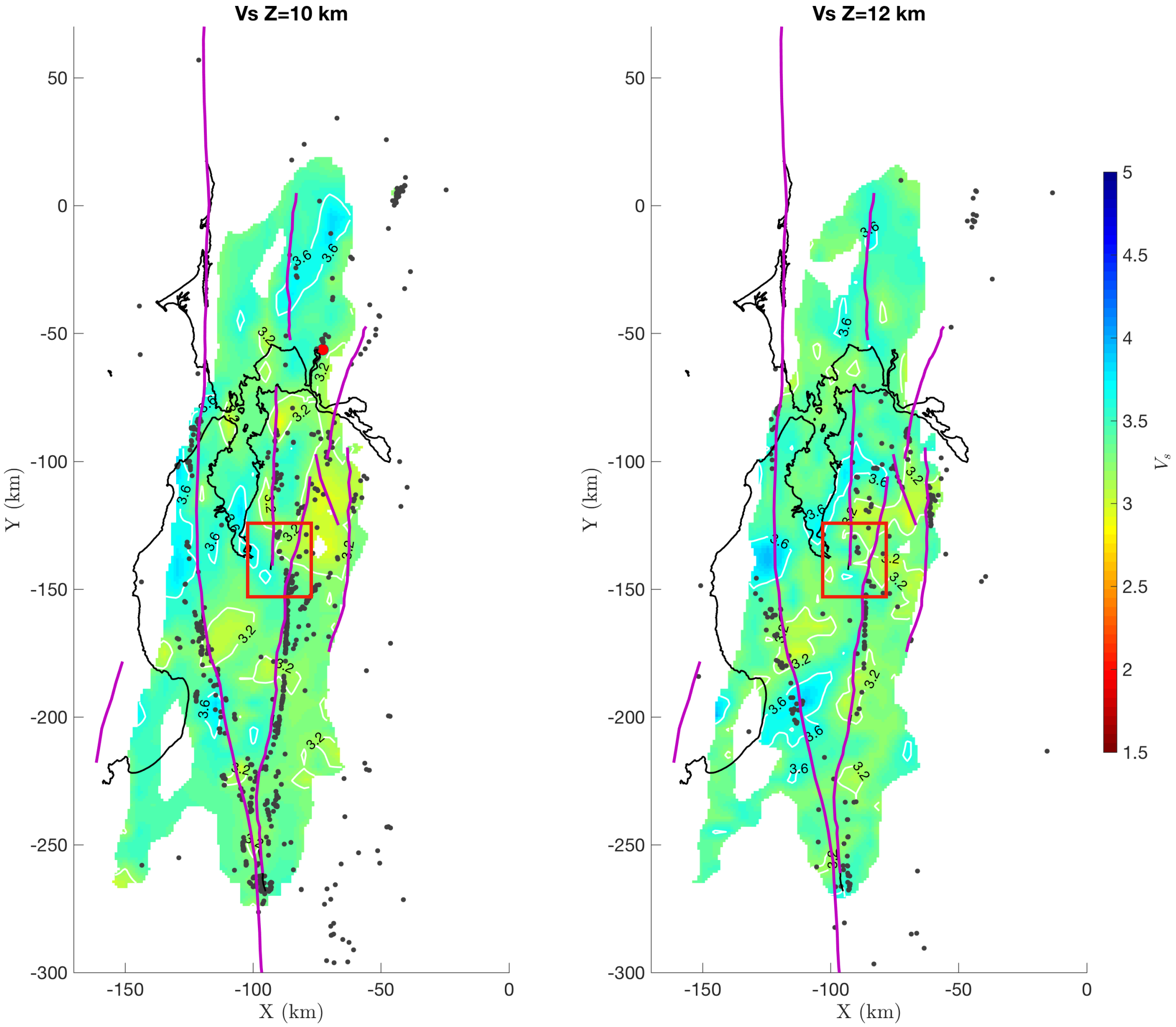


Figure 4.8 Map views of the inverted V_P model at different depths. Purple lines: major fault traces. Black dots: earthquakes. Red dot: 2014 South Napa earthquake mainshock. Areas with P-wave DWS ≤ 36 are masked.


Vs Z=0 km
Vs Z=1 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
V_s
5
4.5
4
3.5
3
2.5
2
1.5
3.6
3.2
2.8
2.4
2


Vs Z=2 km
Vs Z=4 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
2.4
2.8
3.2
3.6
5
4.5
4
3.5
3
2.5
2
1.5
V_s


Vs Z=6 km
Vs Z=8 km
Y (km)
X (km)
50
0
-50
-100
-150
-200
-250
-300
-150
-100
-50
0
3.2
3.6
2.8
V_s
5
4.5
4
3.5
3
2.5
2
1.5


Vs Z=10 km
Vs Z=12 km
50
0
-50
-100
-150
-200
-250
-300
Y (km)
-150
-100
-50
0
X (km)
3.6
3.2
5
4.5
4
3.5
3
2.5
2
1.5
V_s

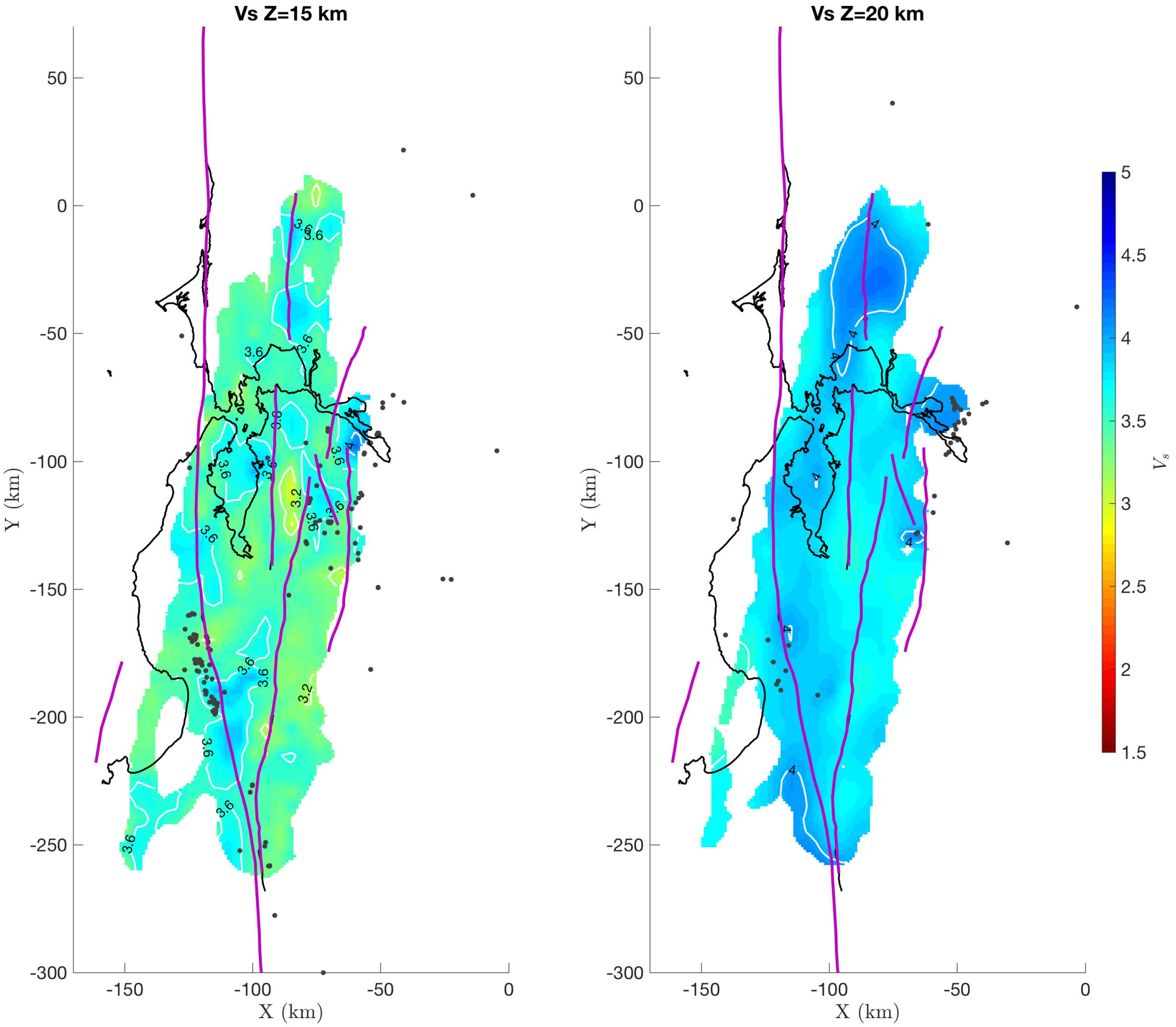
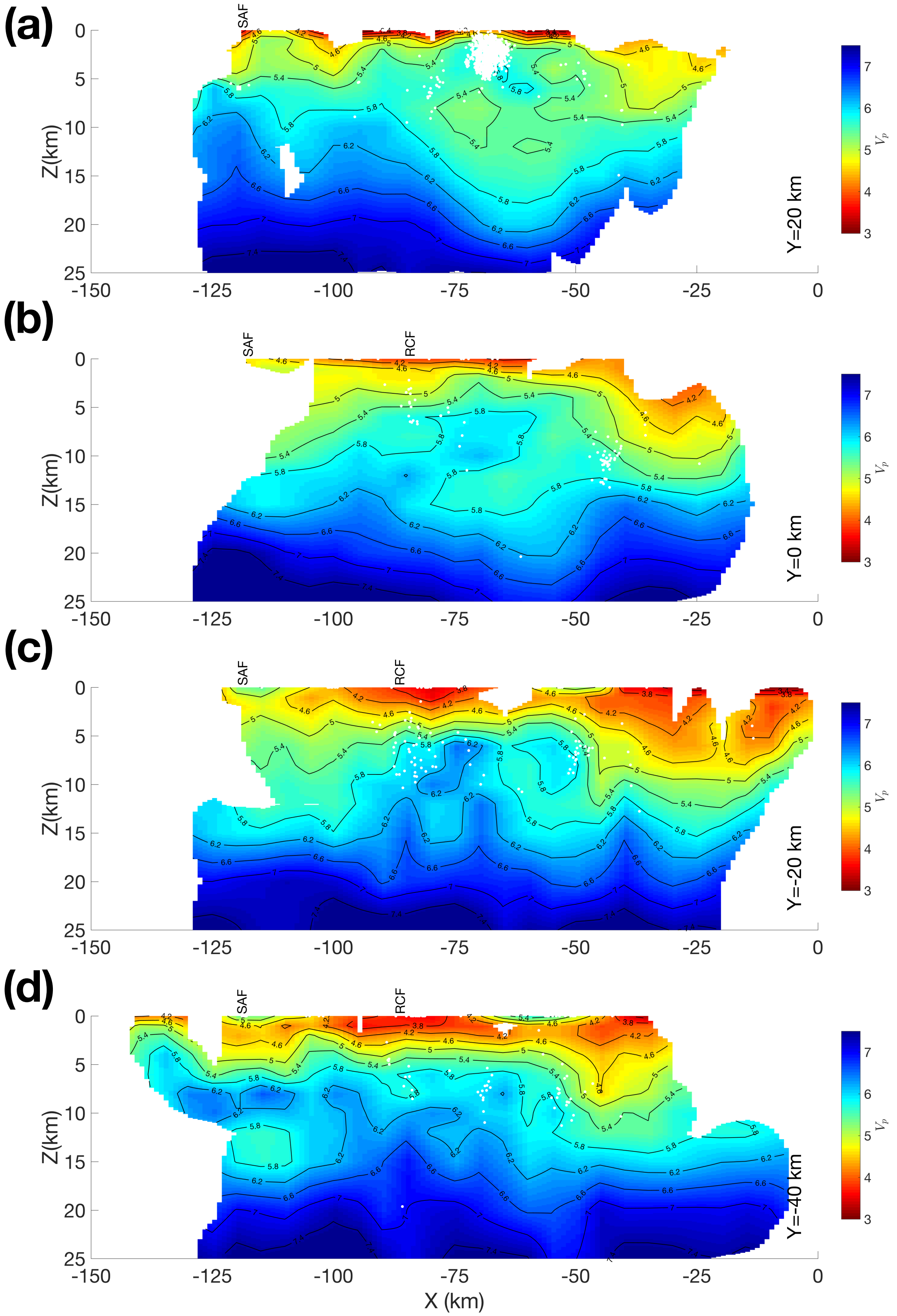
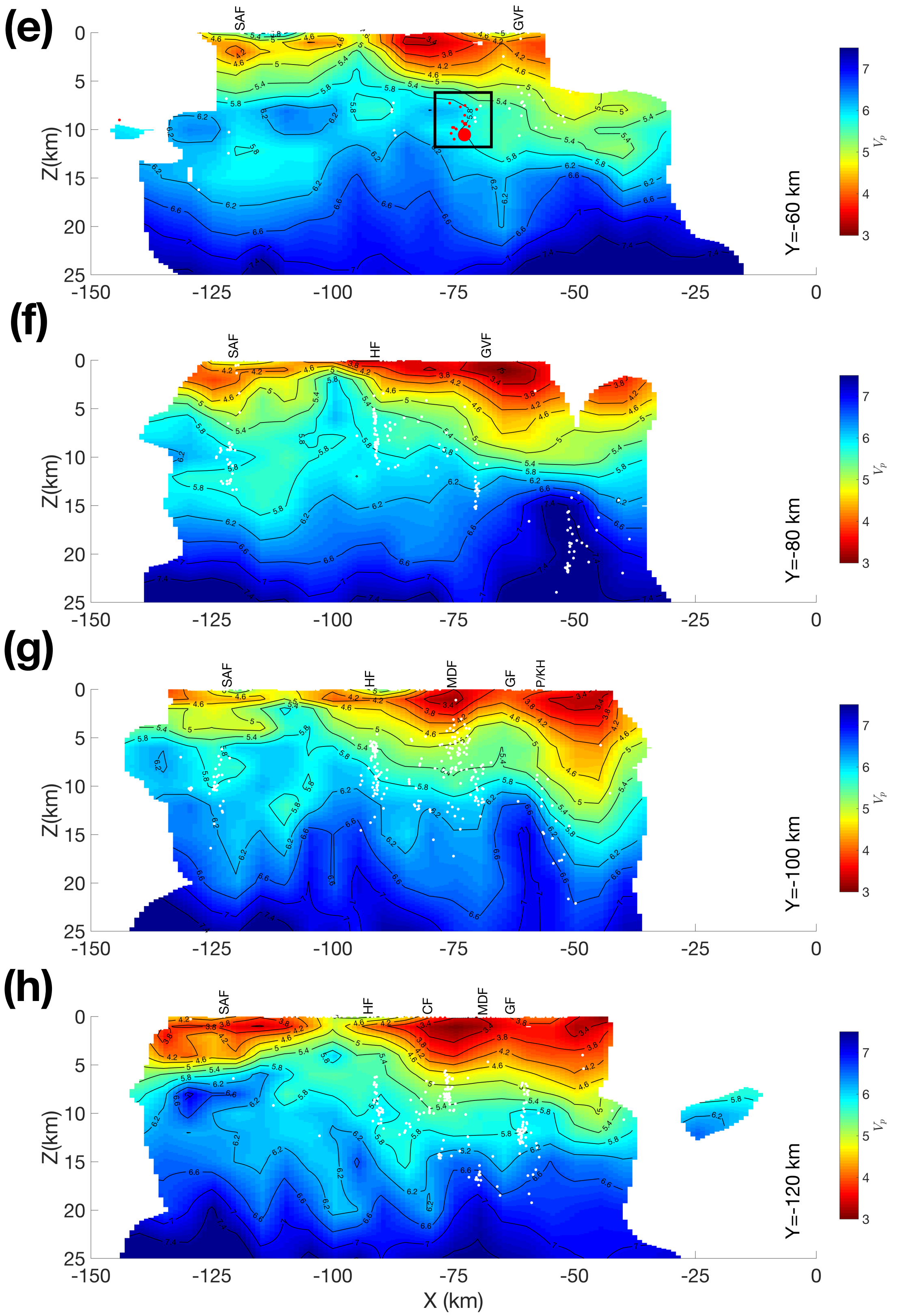
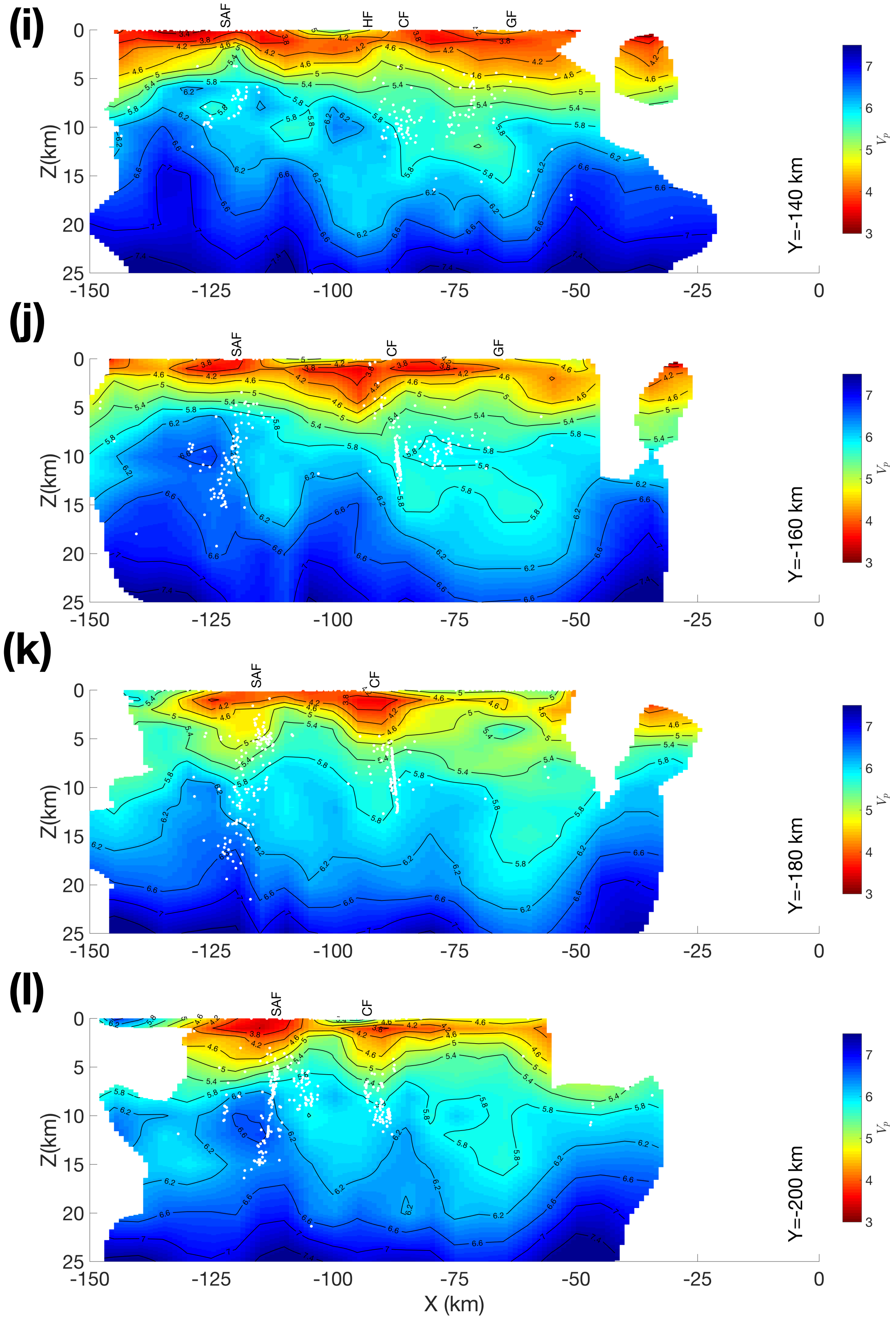


Figure 4.9 Map views of the inverted V_S model at different depths. Purple lines: major fault traces. Black dots: earthquakes. Red dot: 2014 South Napa earthquake mainshock. Areas with S-wave DWS ≤ 16 are masked.


(a)
SAF
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
Y=20 km
V_p
7
6
5
4
3
(b)
SAF
RCF
Z(km)
Y=0 km
(c)
SAF
RCF
Z(km)
Y=-20 km
(d)
SAF
RCF
Z(km)
Y=-40 km
X (km)


(e)
SAF
GVF
Y=-60 km
(f)
SAF
HF
GVF
Y=-80 km
(g)
SAF
HF
MDF
GF
P/KH
Y=-100 km
(h)
SAF
HF
CF
MDF
GF
Y=-120 km
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
X (km)
V_p
3
4
5
6
7


(i)
SAF
HF
CF
GF
Z(km)
Y=-140 km
(j)
SAF
CF
GF
Z(km)
Y=-160 km
(k)
SAF
CF
Z(km)
Y=-180 km
(l)
SAF
CF
Z(km)
Y=-200 km
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
X (km)
V_p
3
4
5
6
7

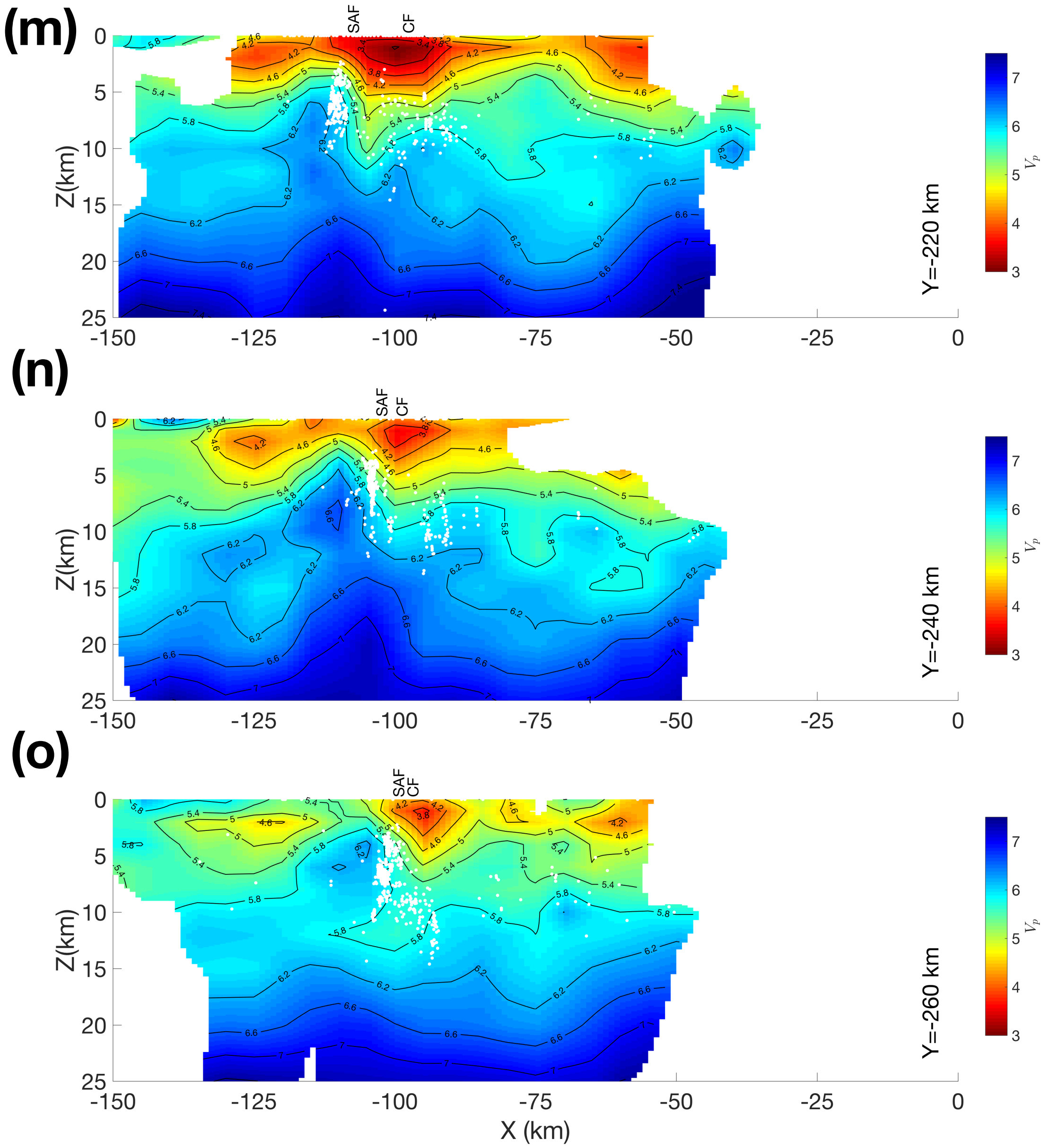
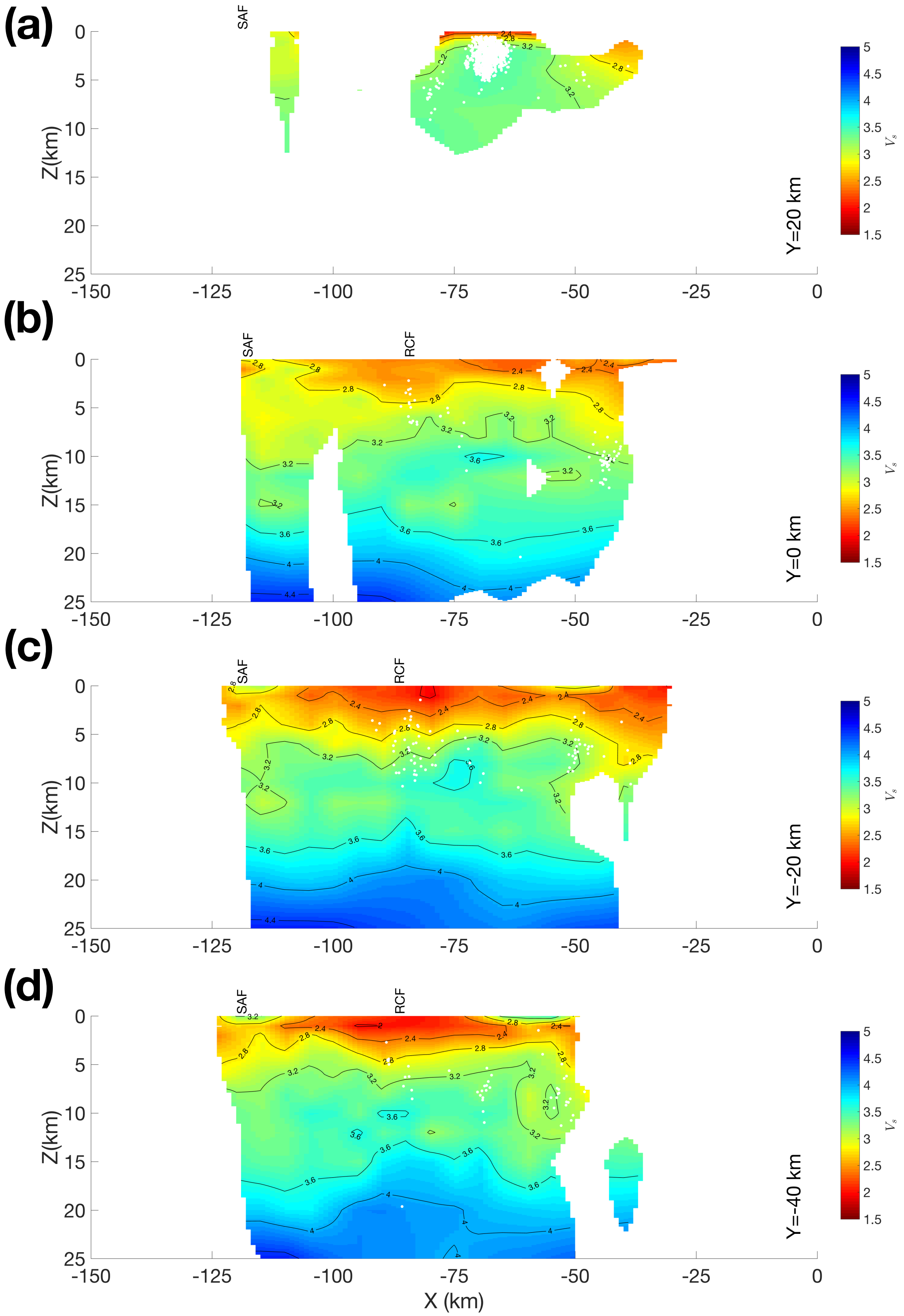
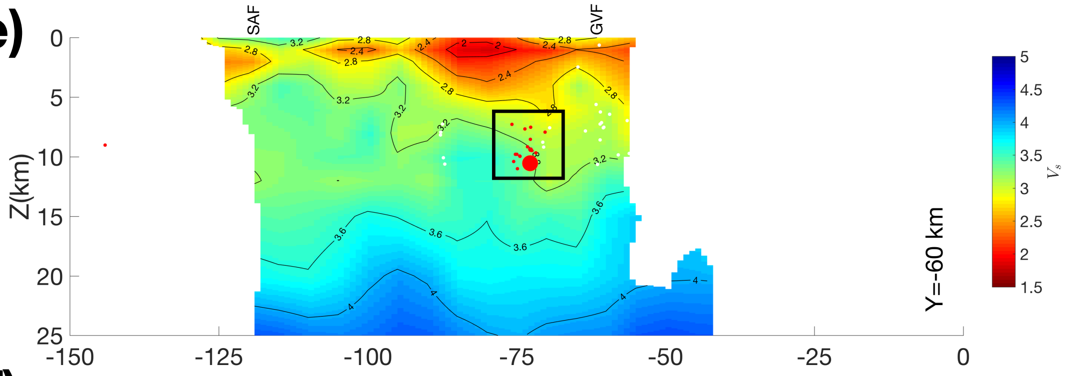


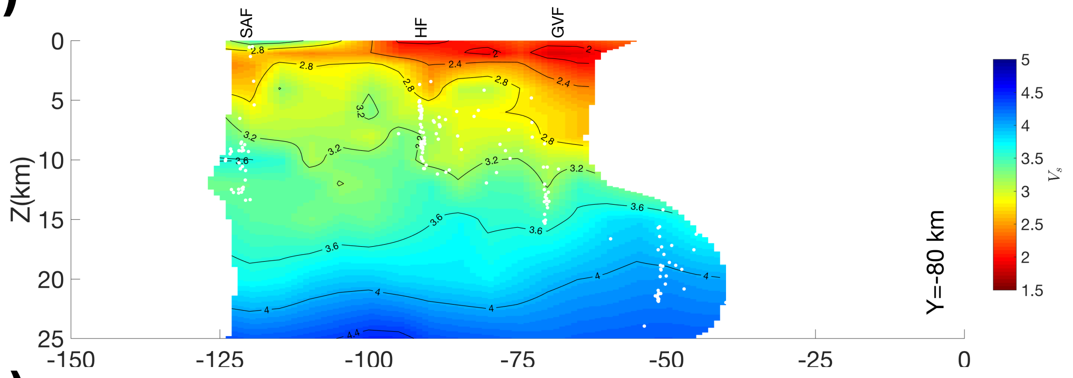
Figure 4.10 Cross-sections of the inverted V_P model. White dots: earthquakes and shots. Red dots: 2014 South Napa earthquake mainshock (larger dot) and aftershocks (smaller dots). SAF: San Andreas Fault. SGF: San Gregorio Fault. CF: Calaveras Fault. HF: Hayward Fault. MDF: Mt. Diablo Fault. GF: Greenville Fault. GVF: Green Valley Fault. P/KH: Pittsburg/Kirby Hills fault. Areas with P-wave DWS ≤ 36 are masked.


(a)
SAF
Y=20 km
(b)
SAF
RCF
Y=0 km
(c)
SAF
RCF
Y=-20 km
(d)
SAF
RCF
Y=-40 km
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
X (km)
V_s
5
4.5
4
3.5
3
2.5
2
1.5

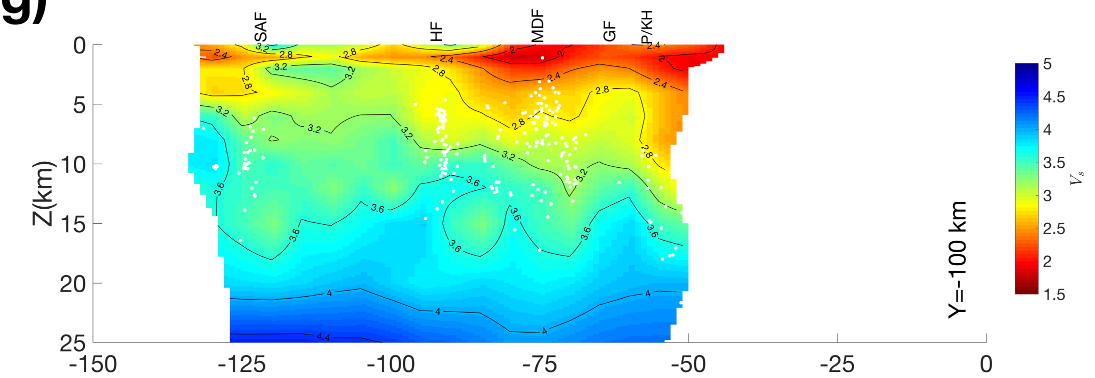
(e)


SAF
GVF
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
Y=-60 km
V_s
5
4.5
4
3.5
3
2.5
2
1.5

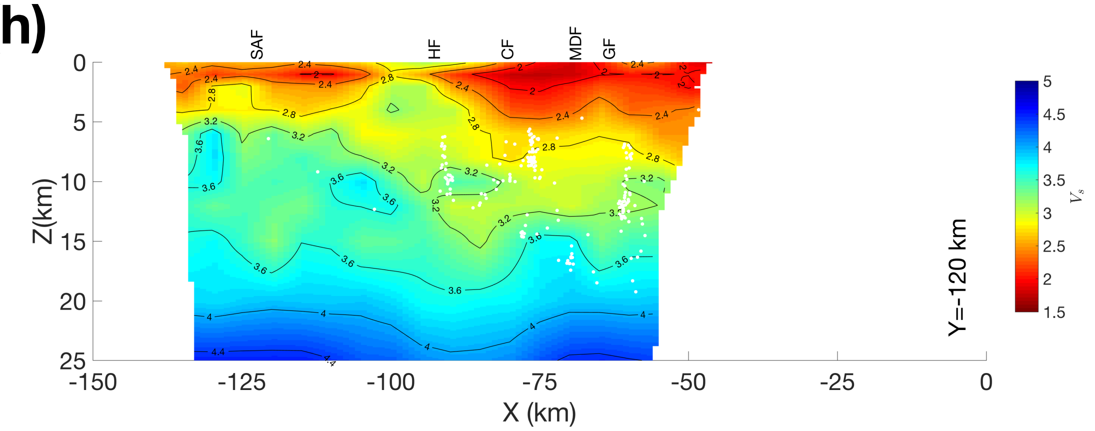
(f)

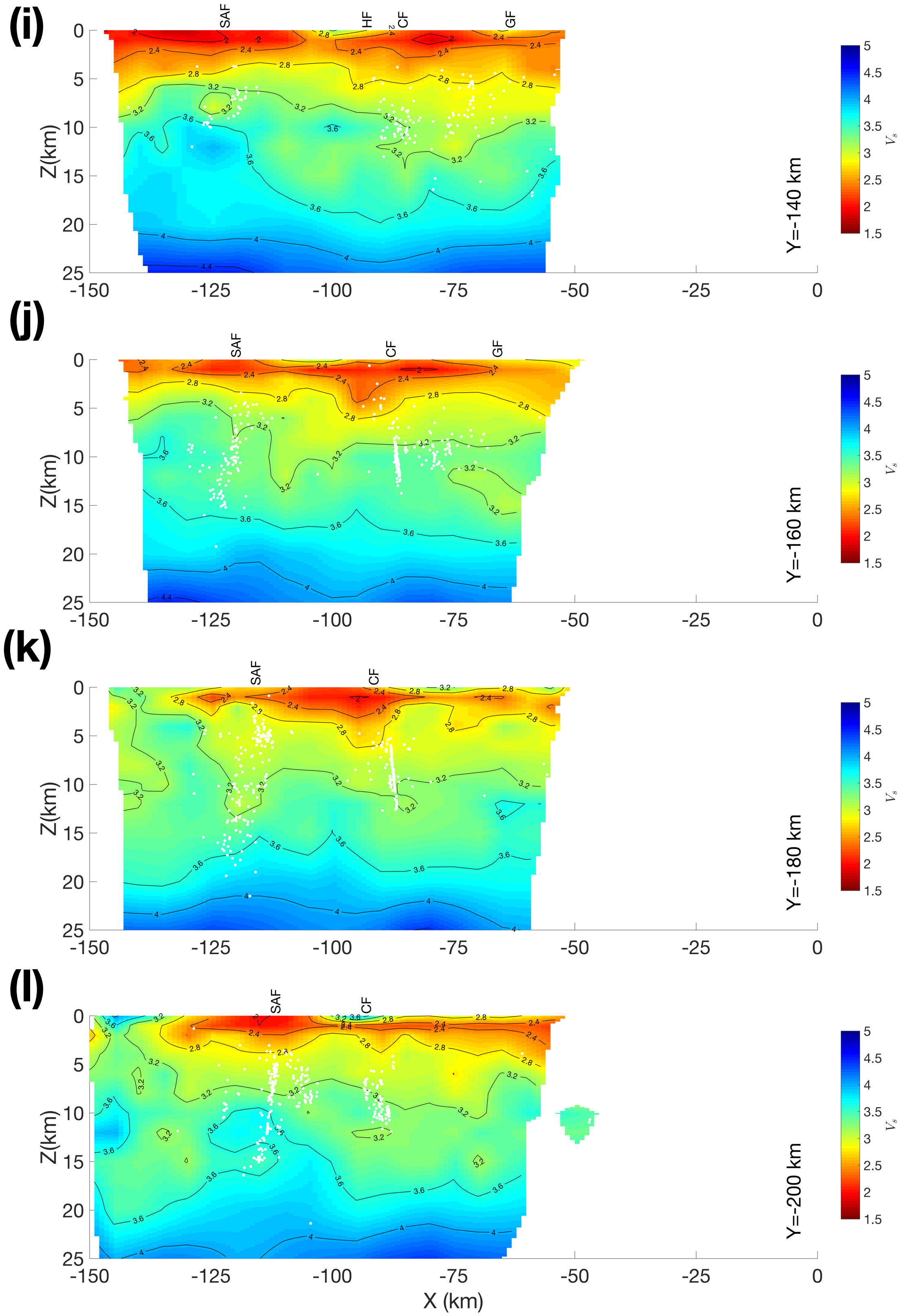

SAF
HF
GVF
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
Y=-80 km
V_s
5
4.5
4
3.5
3
2.5
2
1.5

(g)


SAF
HF
MDF
GF
P/KH
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
Y=-100 km
V_s
5
4.5
4
3.5
3
2.5
2
1.5

(h)


SAF
HF
CF
MDF
GF
Z(km)
0
5
10
15
20
25
-150
-125
-100
-75
-50
-25
0
X (km)
Y=-120 km
V_s
5
4.5
4
3.5
3
2.5
2
1.5


(i)
SAF
HF
CF
GF
Z(km)
Y=-140 km
(j)
SAF
CF
GF
Z(km)
Y=-160 km
(k)
SAF
CF
Z(km)
Y=-180 km
(l)
SAF
CF
Z(km)
Y=-200 km
X (km)
-150
-125
-100
-75
-50
-25
0
0
5
10
15
20
25
V_s
1.5
2
2.5
3
3.5
4
4.5
5

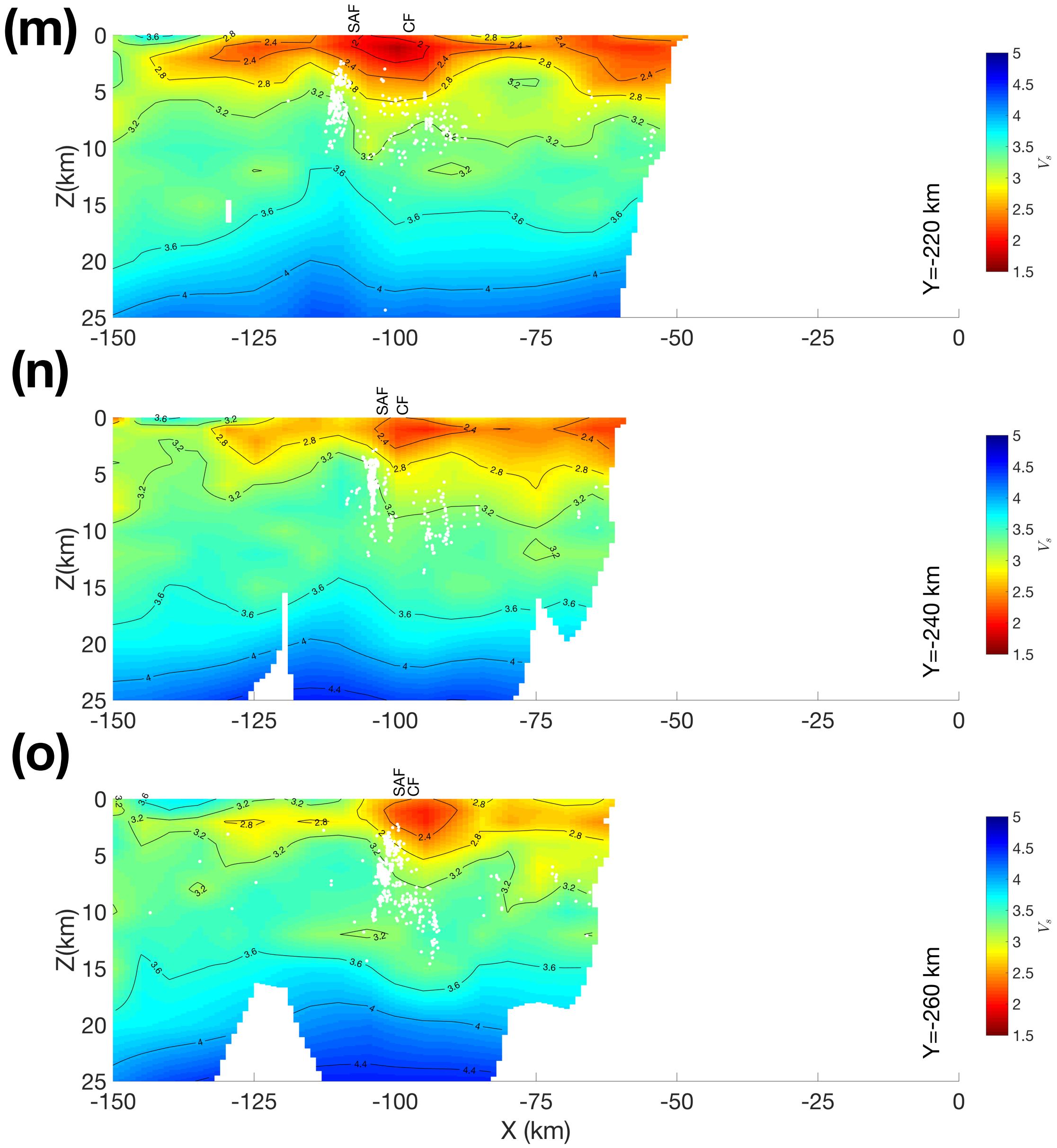


Figure 4.11 Cross-sections of the inverted V_S model. White dots: earthquakes (shots are not plotted for the V_s model). Red dot: 2014 South Napa earthquake mainshock (larger dot) and aftershocks (smaller dots). SAF: San Andreas Fault. SGF: San Gregorio Fault. CF: Calaveras Fault. HF: Hayward Fault. MDF: Mt. Diablo Fault. GF: Greenville Fault. GVF: Green Valley Fault. P/KH: Pittsburg/Kirby Hills fault. Areas with S-wave DWS $\leq$ 16 are masked.

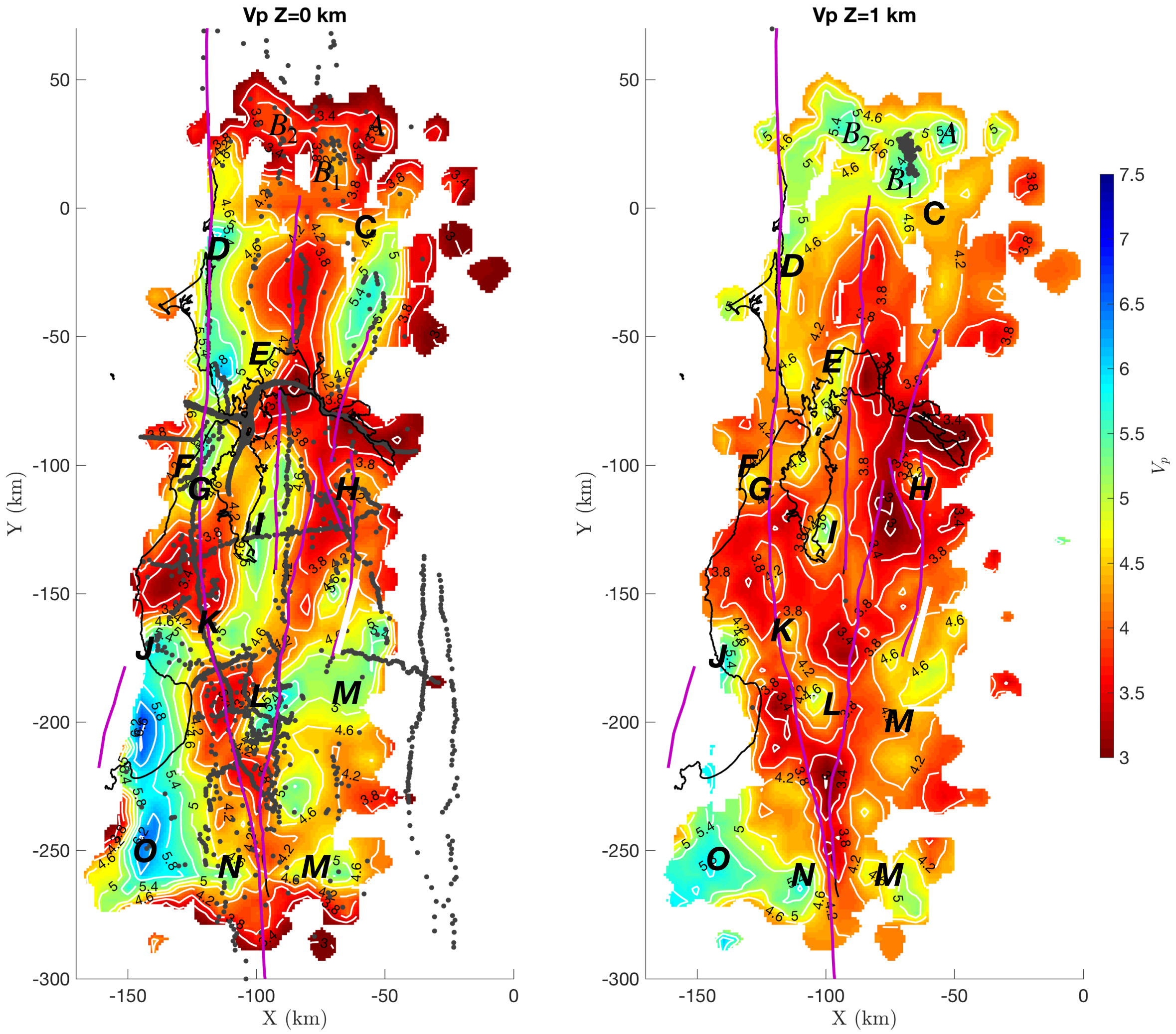


Figure 4.12 Map view of identified basement highs at depths Z = 0 and 1 km. Purple lines are the major fault traces. (A) Tertiary volcanic rocks on the northeast side of The Geysers and Hidden Valley Lake; ($B_{1,2}$) Franciscan rocks in the coastal ranges (B_1 Elk Range, B_2 Mayacamas Mountains) west and south sides of The Geysers; (C) Franciscan rocks northeast side of the Pope Valley; (D) Franciscan complex in Coast Ranges west side of the Rodgers Creek Fault; (E) Franciscan rocks of the Marin hills and the San Francisco Bay block west side of the HF; (F) Salinian rocks of the Montara pluton; (G) Franciscan rocks of the Pilarcitos block; (H) Franciscan rocks of Mount Diablo; (I) Franciscan rocks of the Coyote Hills; (J) Salinian rocks of the Ben Lomond pluton; (K) Monte Vista fault hanging wall with Franciscan rocks; (L) Franciscan rocks of the Santa Cruz Mountains; (M) Franciscan rocks of the Diablo Range; (N) Salinian rocks of the Gabilan Range pluton; and (O) Granites and metasediments of the Santa Lucia Range.

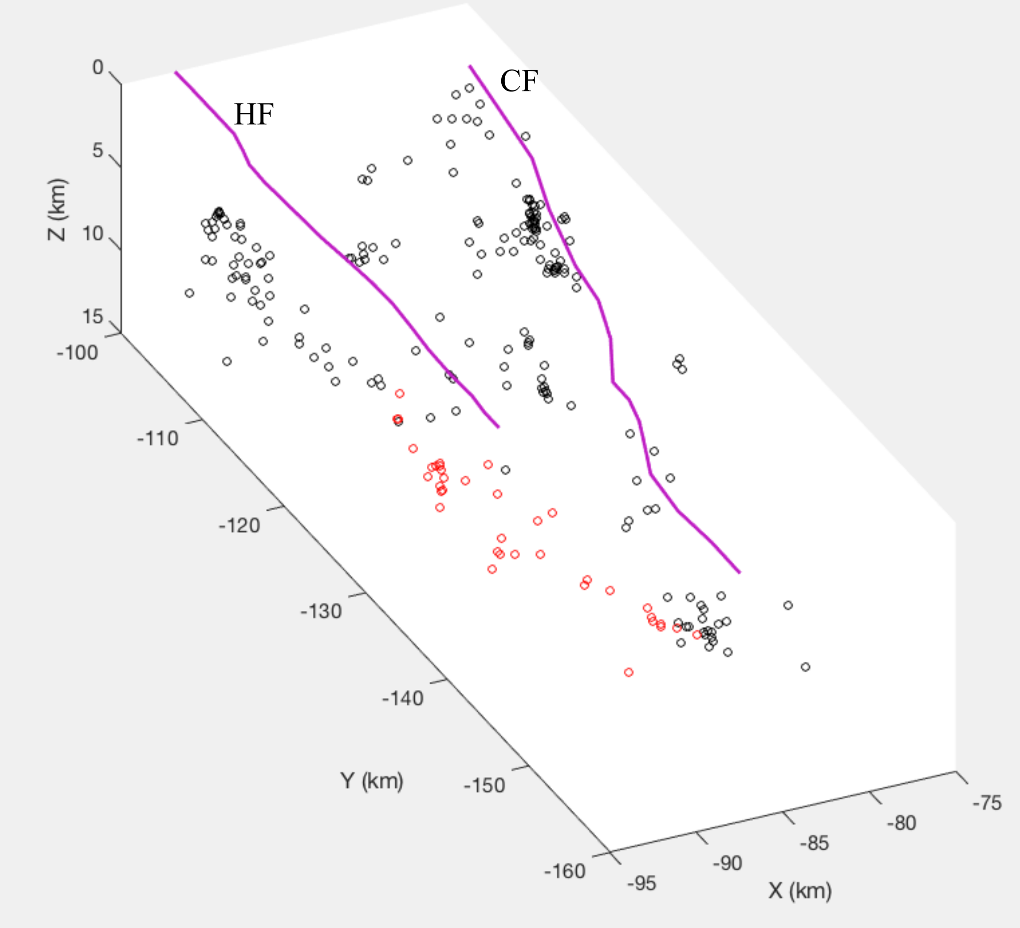
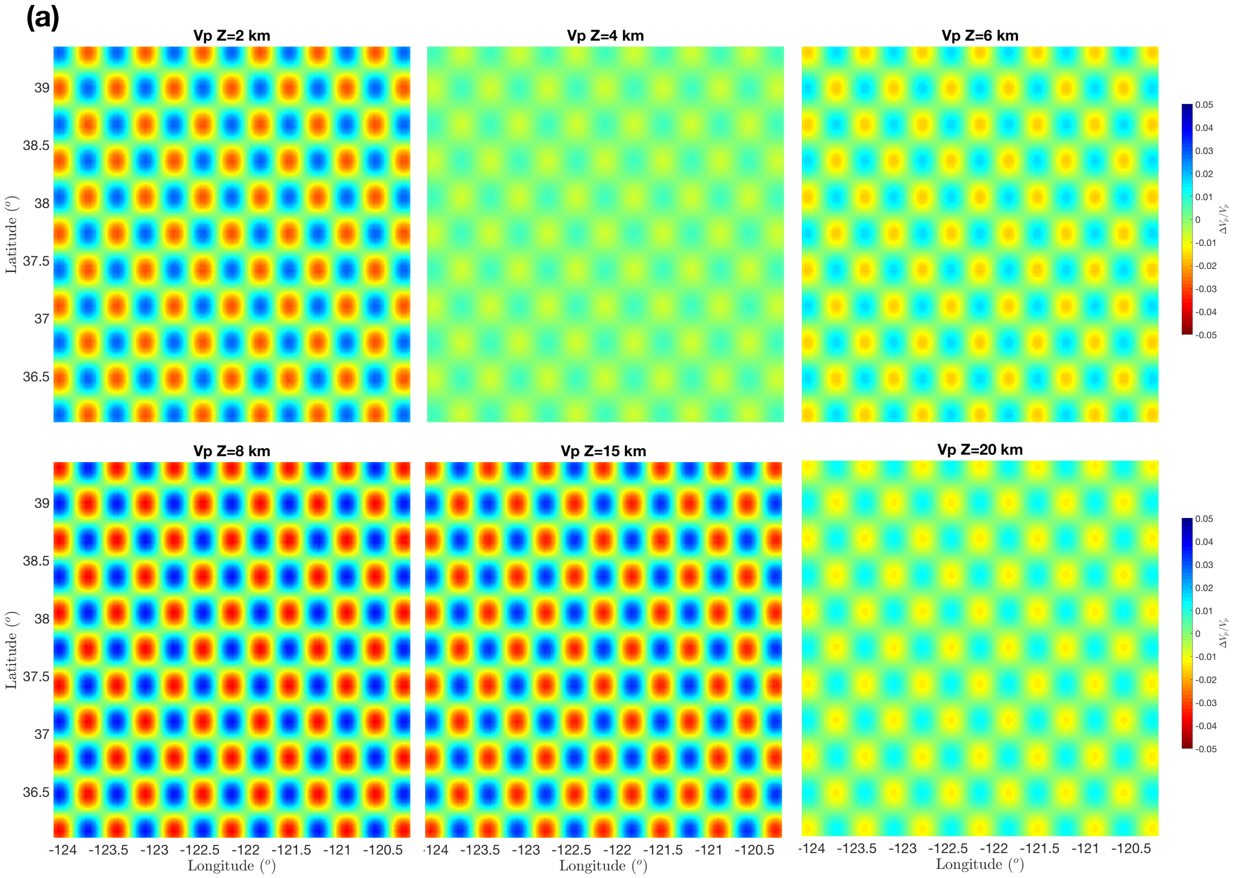


Figure 4.13 Earthquakes distributed within the CF and HF junction region. The HF and CF surface traces are plotted with purple lines at 0 km depth. Events distributed within region (Y ≤ -125km and X ≤ -87.5 km) are marked with red circles. Other events located underneath the HF and CF fault surface traces are plotted with black circles.

4.14. Appendix



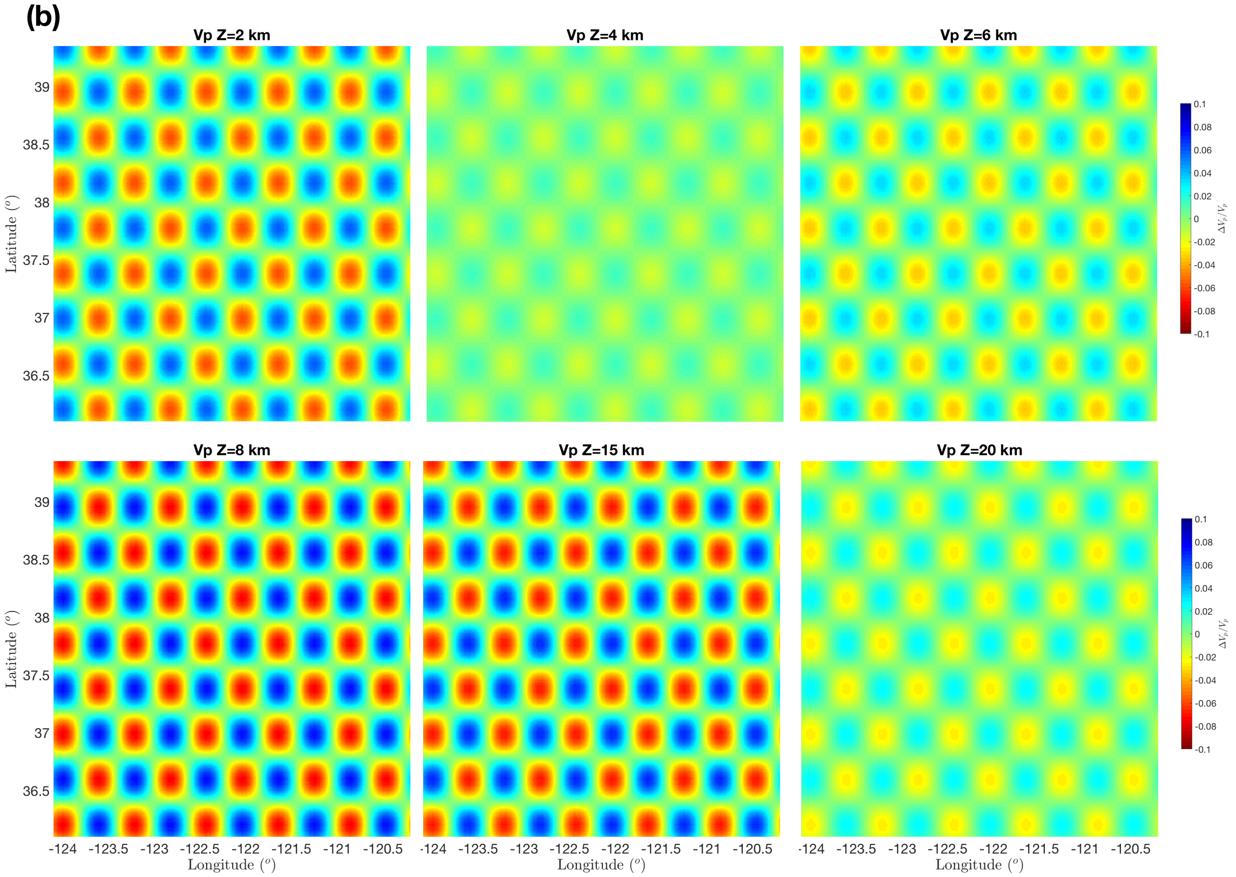
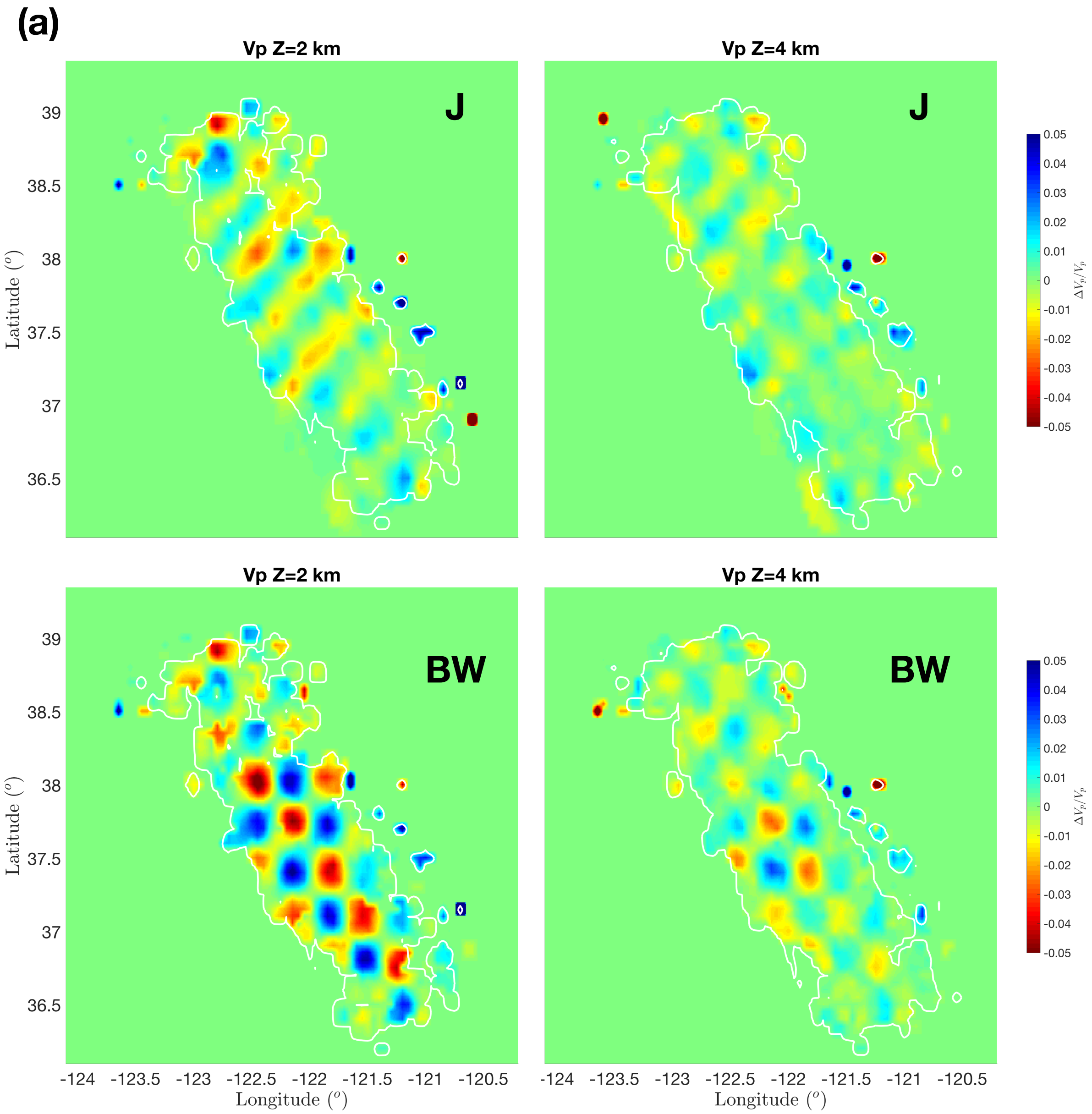
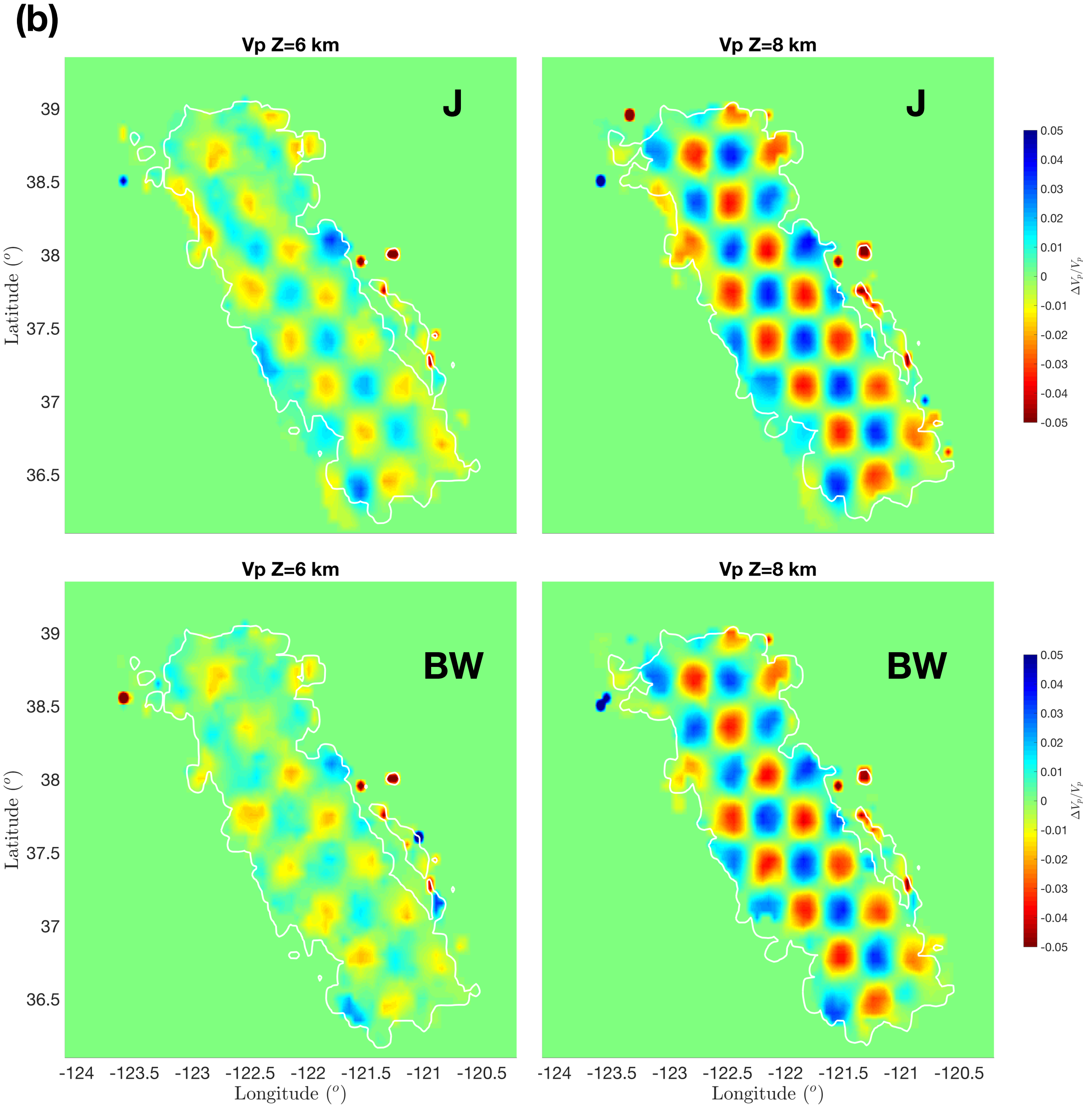
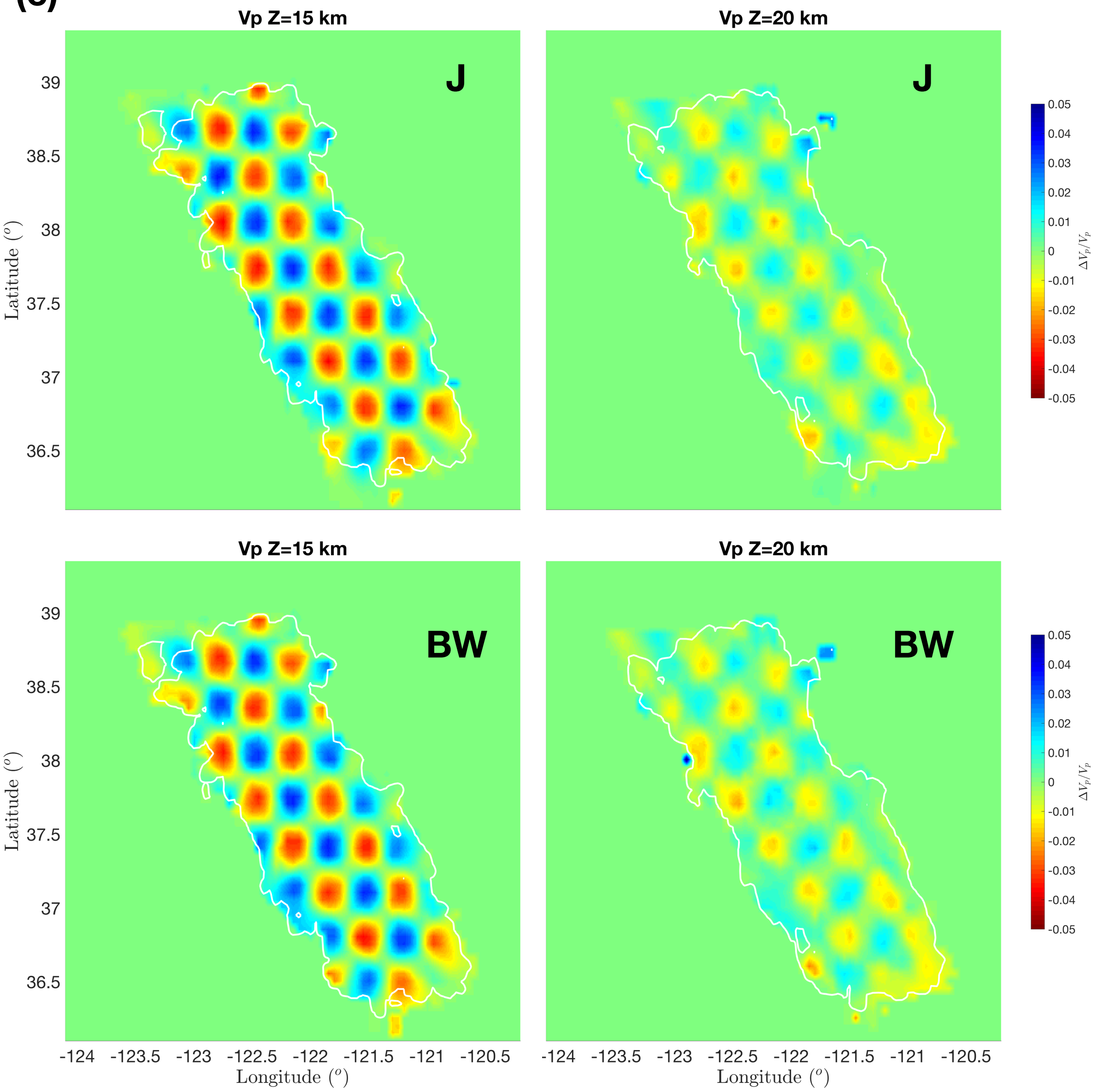


Figure A 4.1 Checkerboard anomalies for the two sets of checkerboard tests. (a) Map views at Z = 2, 4, 6, 8, 15, and 20 km for the V_P anomalies with horizontal dimension of ~25 km. The maximum anomaly amplitude is 5%. (b) Map views at Z = 2, 4, 6, 8, 15, and 20 km for the V_P anomalies with horizontal dimension of ~40 km. The maximum anomaly amplitude is 10%.


(a)
Vp Z=2 km
Vp Z=4 km
J
J
Vp Z=2 km
Vp Z=4 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (°)
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
$\Delta V_p/V_p$


(b)
Vp Z=6 km
Vp Z=8 km
J
J
Vp Z=6 km
Vp Z=8 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
$\Delta V_p/V_p$

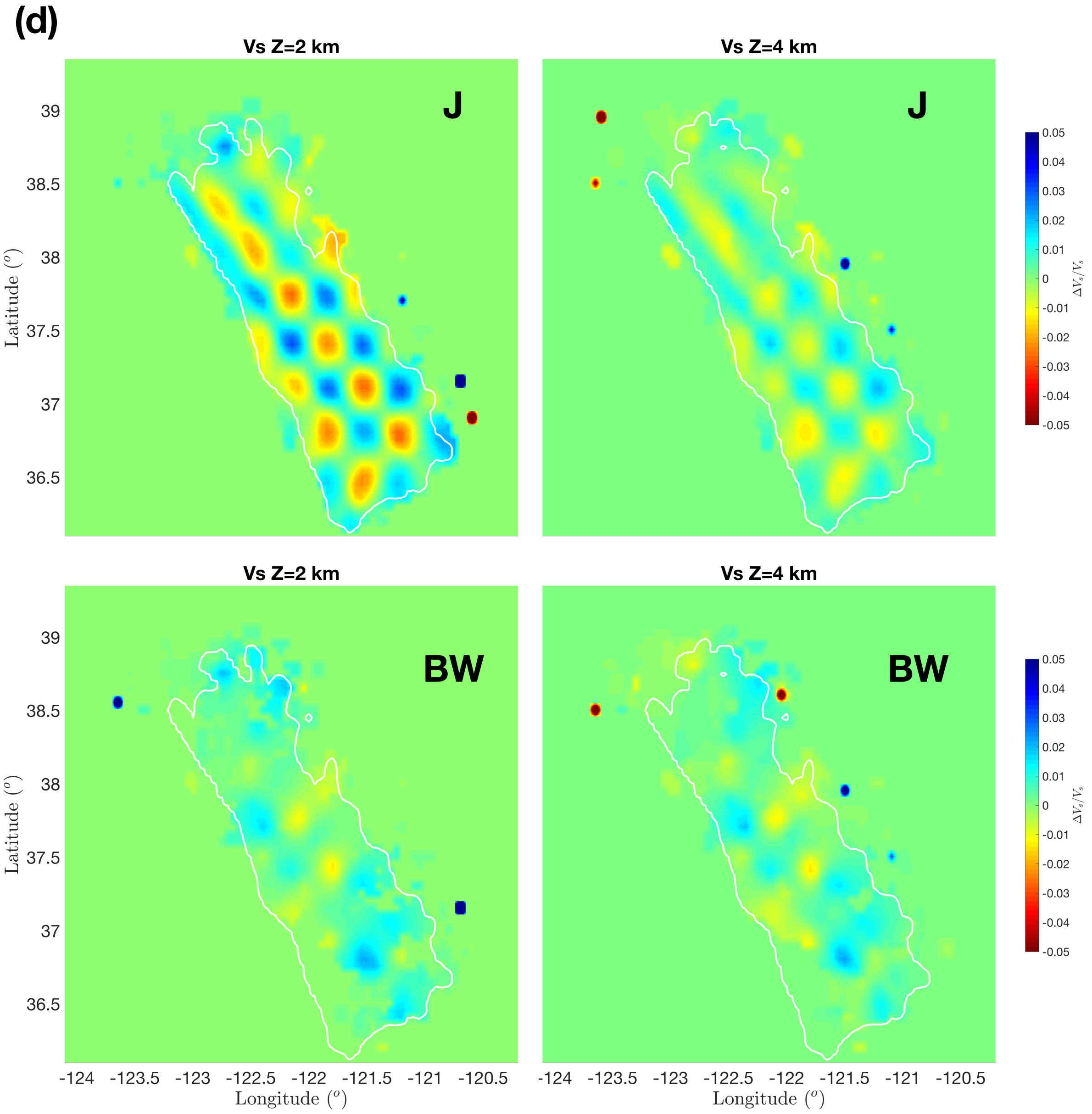
(c)


Vp Z=15 km

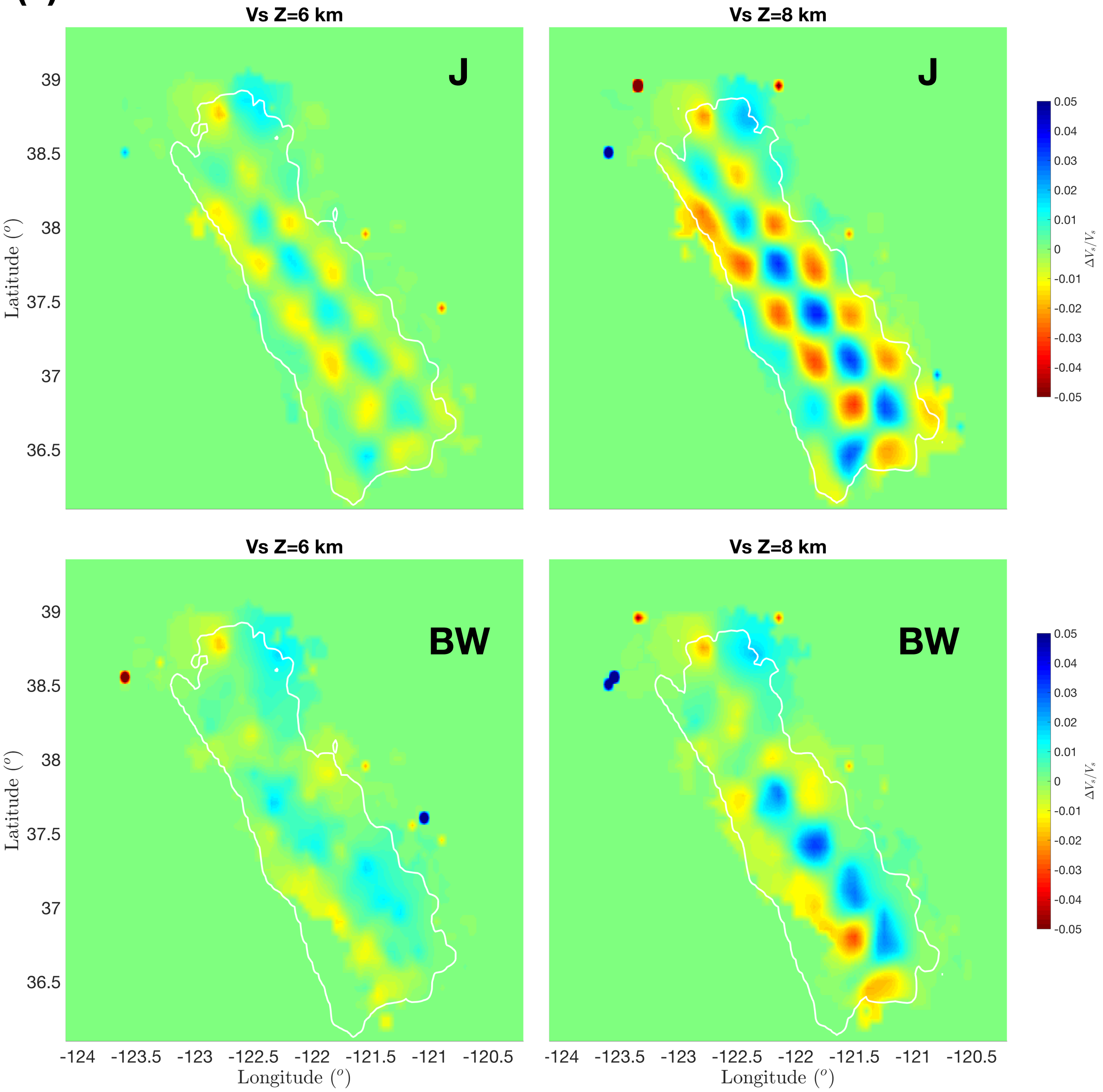
Vp Z=20 km

Vp Z=15 km

Vp Z=20 km


(d)
Vs Z=2 km
Vs Z=4 km
J
J
Vs Z=2 km
Vs Z=4 km
BW
BW
Latitude (°)
39
38.5
38
37.5
37
36.5
Longitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
ΔVs/Vs

(e)


Vs Z=6 km
Vs Z=8 km
J
J
Vs Z=6 km
Vs Z=8 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.05
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
-0.05
$\Delta V_s/V_s$

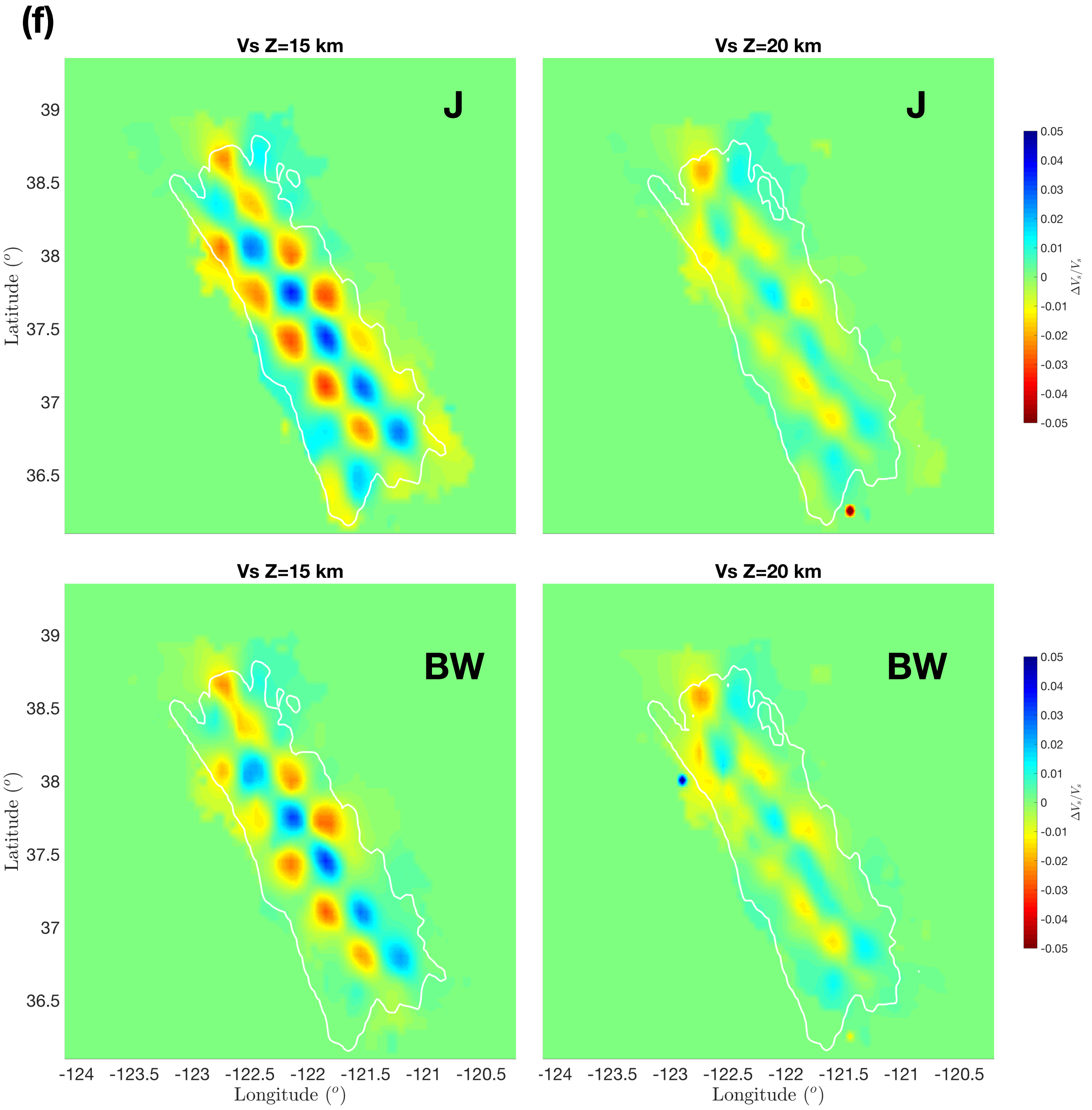
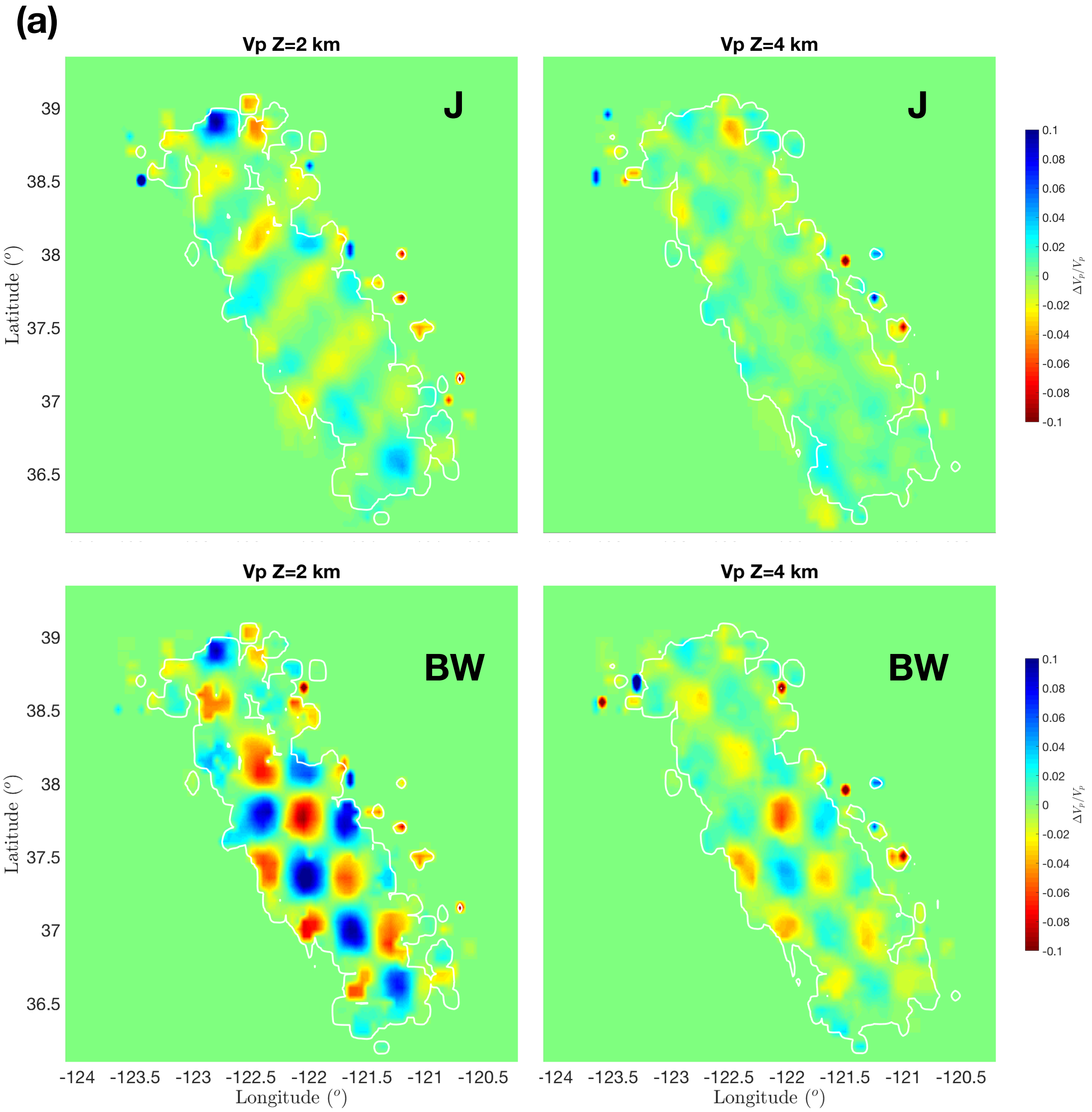
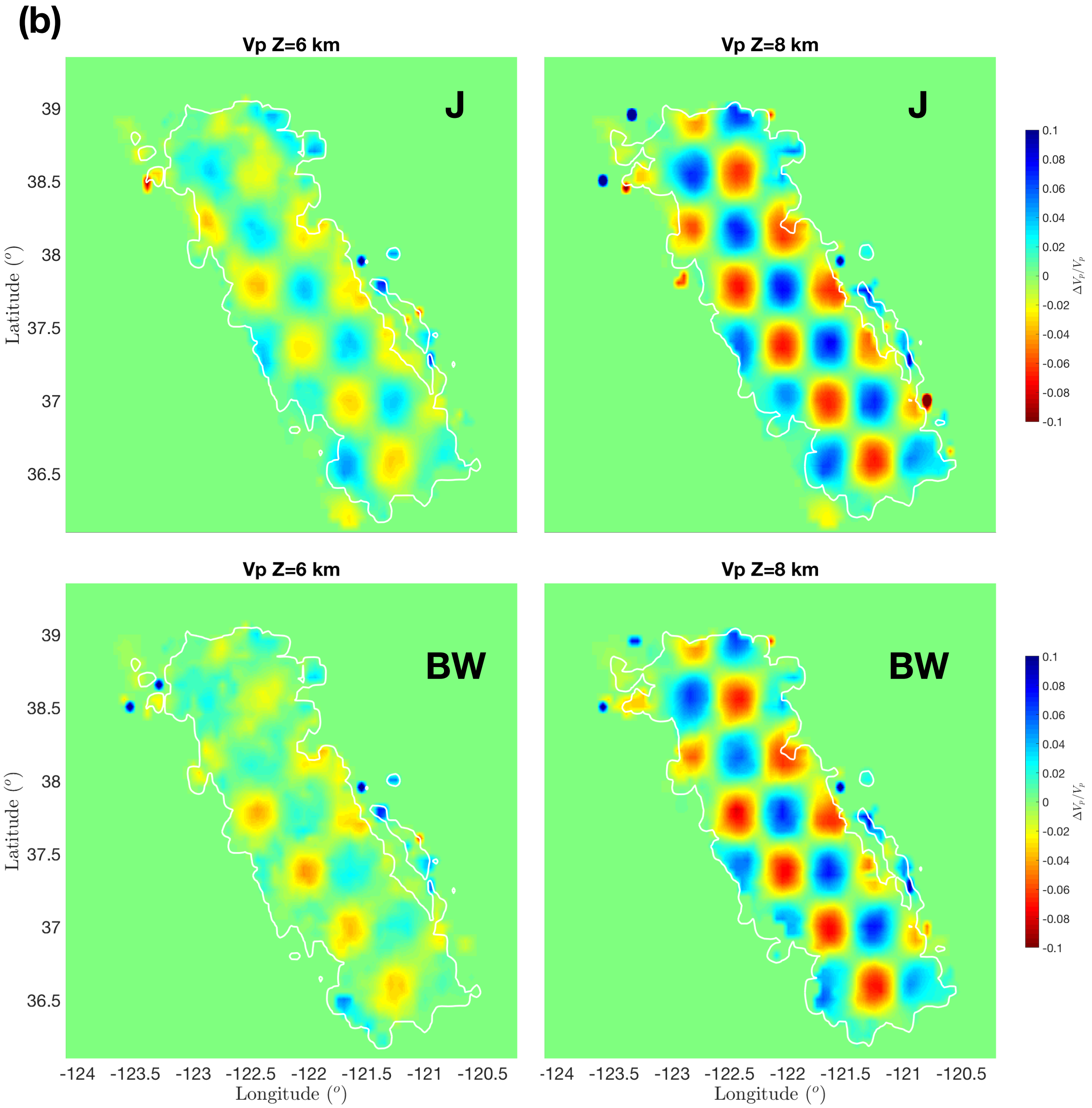
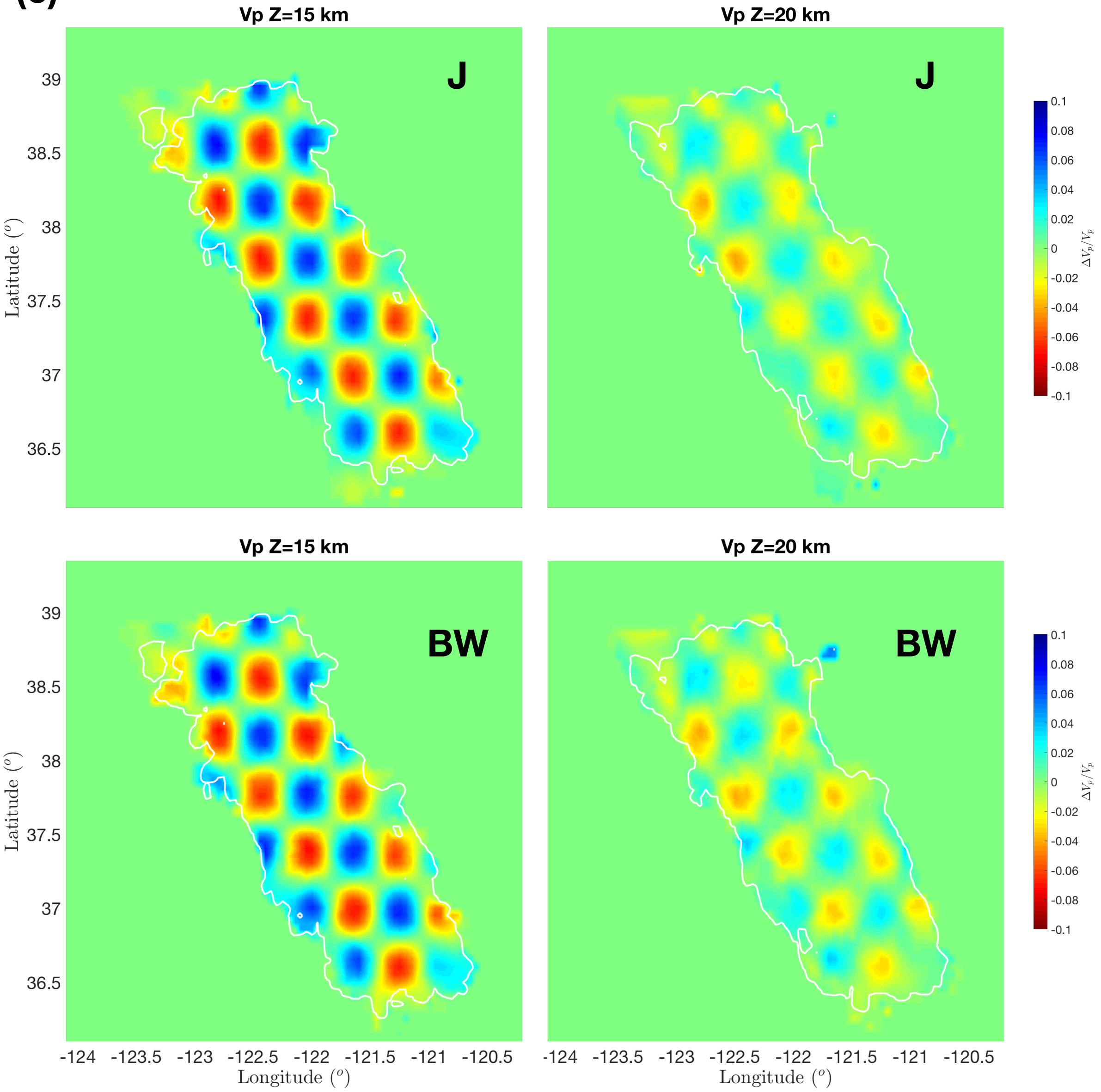


Figure A 4.2 Checkerboard test results for anomalies with ~25 km horizontal dimension. For each four-plot group (a)-(f), the first row shows the joint inversion results (J) and the second row shows the body-wave only inversion results (BW). (a) V_P results at Z = 2 and 4 km depths. (b) V_P results at Z = 6 and 8 km depths. (c) V_P results at Z = 15 and 20 km depths. (d) V_S results at Z = 2 and 4 km depths. (e) V_S results at Z = 6 and 8 km depths. (f) V_S results at Z = 15 and 20 km depths. White contours are shown for P-wave model with DWS = 36 and S-wave model for DWS = 16.

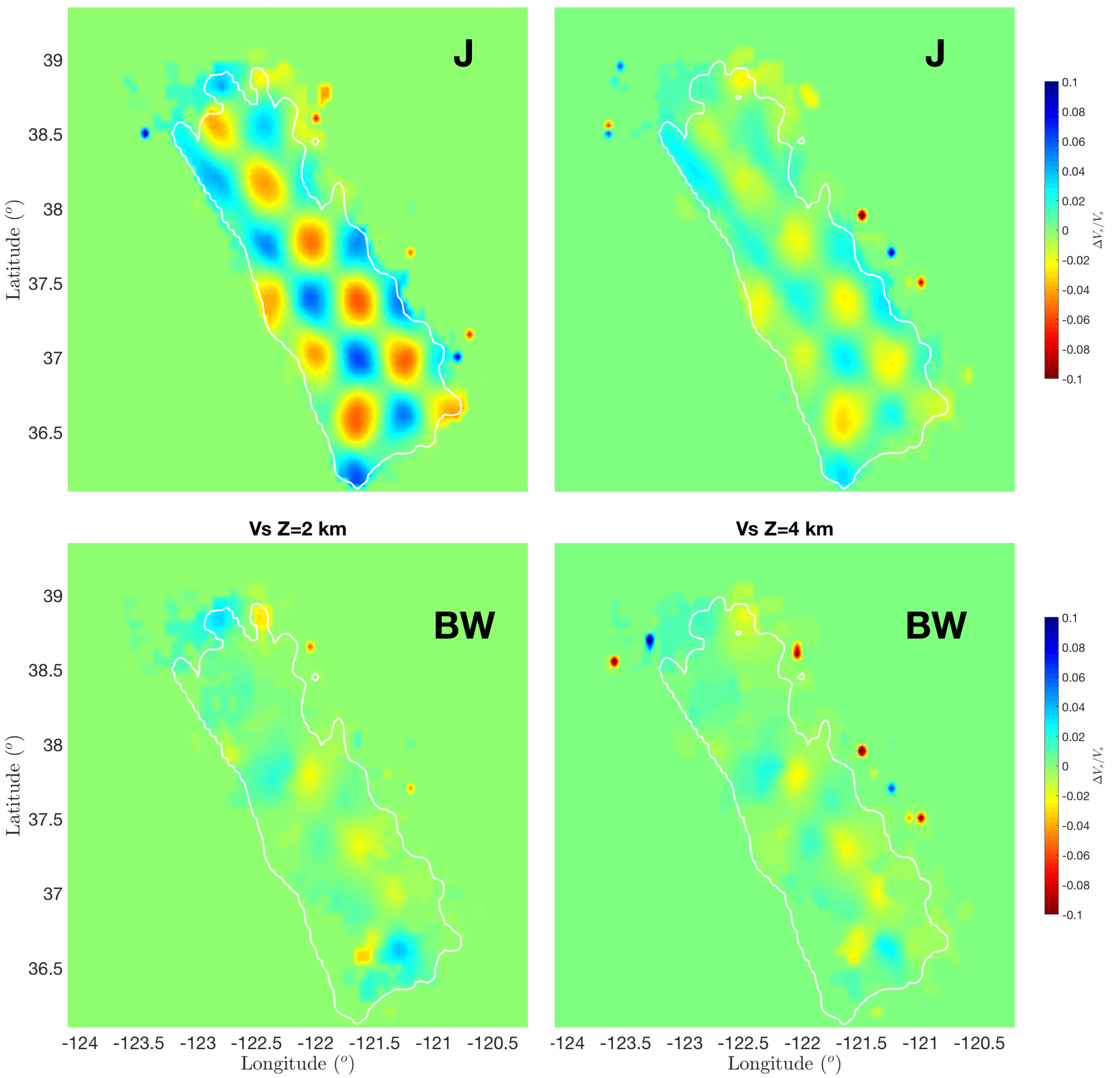

(a)
Vp Z=2 km
Vp Z=4 km
J
J
Vp Z=2 km
Vp Z=4 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
$\Delta V_p/V_p$


(b)
Vp Z=6 km
Vp Z=8 km
J
J
Vp Z=6 km
Vp Z=8 km
BW
BW
Latitude (°)
39
38.5
38
37.5
37
36.5
Longitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
$\Delta V_p/V_p$
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1

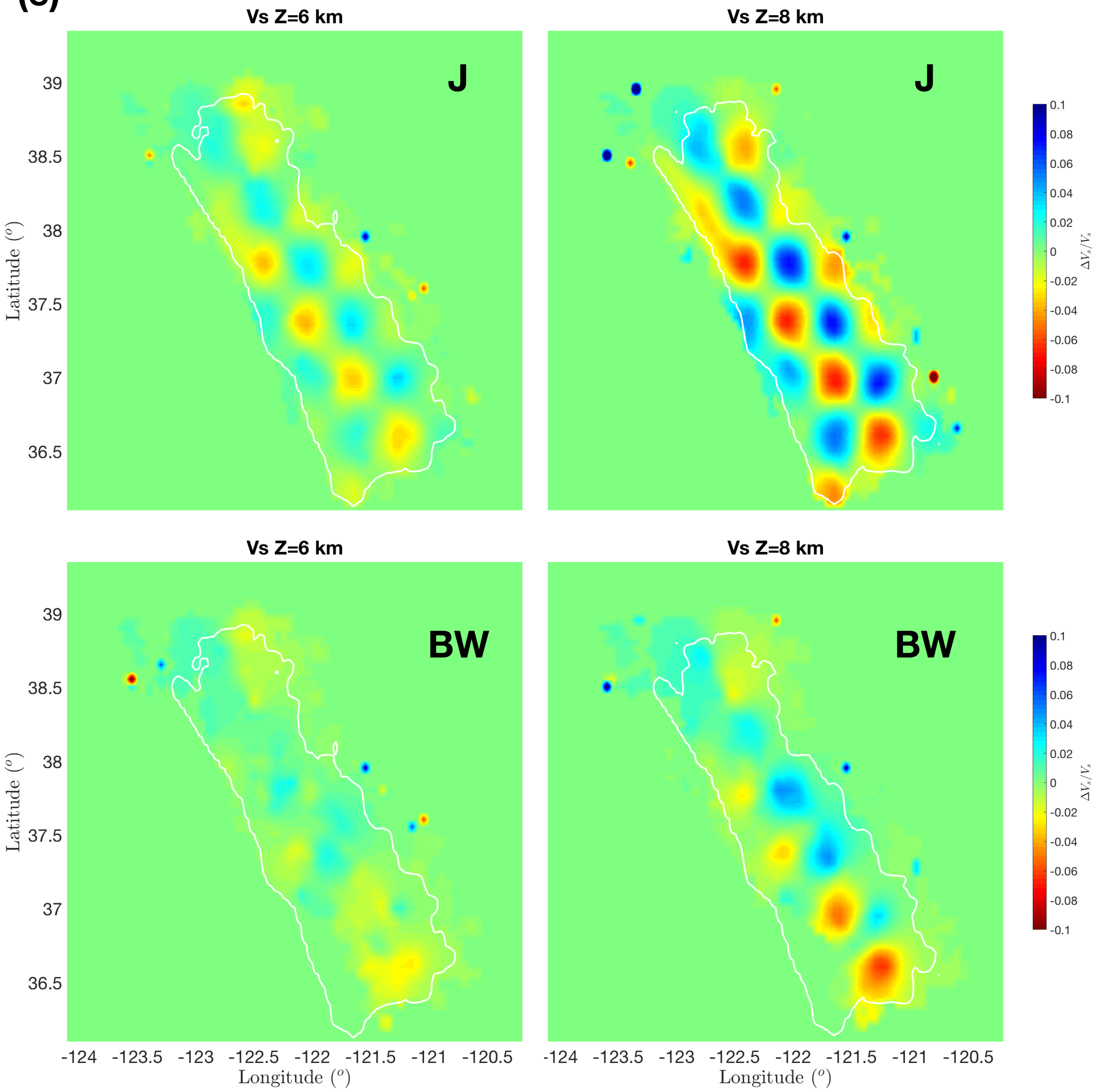
(c)


Vp Z=15 km
Vp Z=20 km
J
J
Vp Z=15 km
Vp Z=20 km
BW
BW
39
38.5
38
37.5
37
36.5
Latitude (^{o})
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
$\Delta V_p/V_p$

(d)


Vs Z=2 km
Vs Z=4 km
J
J
Vs Z=2 km
Vs Z=4 km
BW
BW
Latitude (^{o})
39
38.5
38
37.5
37
36.5
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
Longitude (^{o})
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
$\Delta V_s/V_s$

(e)


Vs Z=6 km
Vs Z=8 km
J
J
Vs Z=6 km
Vs Z=8 km
BW
BW
Latitude (°)
39
38.5
38
37.5
37
36.5
Longitude (°)
-124
-123.5
-123
-122.5
-122
-121.5
-121
-120.5
$\Delta V_s/V_s$
0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1

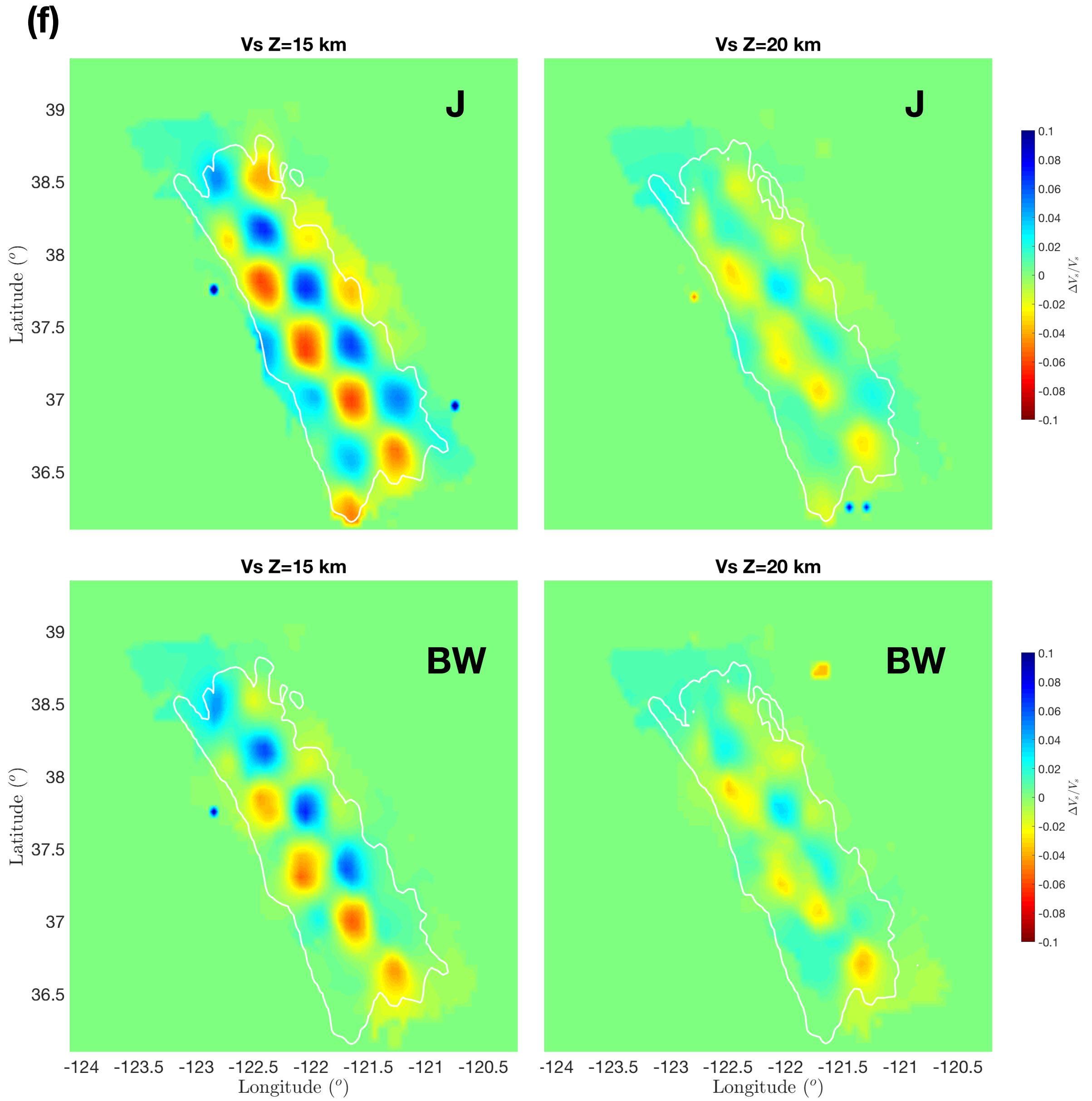


Figure A 4.3 Checkerboard test results for anomalies with ~40 km horizontal dimension. For each four-plot group (a)-(f), the first row shows the joint inversion results (J) and the second row shows the body-wave only inversion results (BW). (a) V_P results at Z = 2 and 4 km depths. (b) V_P results at Z = 6 and 8 km depths. (c) V_P results at Z = 15 and 20 km depths. (d) V_S results at Z = 2 and 4 km depths. (e) V_S results at Z = 6 and 8 km depths. (f) V_S results at Z = 15 and 20 km depths. White contours are shown for P-wave model with DWS = 36 and S-wave model for DWS = 16.

5. Double-difference attenuation tomography

5.1. Abstract

Double-difference seismic velocity tomography utilizes the body-wave catalog and cross-correlation differential arrival time data for nearby event pairs to improve the seismic velocity structure resolution in the earthquake source region. Inspired by the double-difference velocity tomography method, we propose to integrate the differential t^* data for adjacent event pairs in the traditional Q tomography, which only uses catalog t^* data, to improve the Q structure resolution in the source region. In this study, we investigated the advantage of solving event pair differential t^* with spectral ratio method over deriving differential t^* by subtracting individually solved t^* values. Based on the results from the checkerboard test and temporal Q structure change recovery tests, including event pair differential t^* in the Q tomography improves the resolution in the source region.

5.2. Introduction

In high-precision earthquake relocation studies, event-pair body-wave differential arrival times derived from waveform cross-correlations have been used to provide a sharper image of local earthquake distributions in various tectonic settings (Deichmann and Garcia-Fernandez, 1992; Got et al., 1994; Slunga et al., 1995; Gillard et al., 1998; Waldhauser et al., 1999; Rubin et al., 1999; Waldhauser and Ellsworth, 2000). We will refer to this technique as the double-difference (DD) location method. Later on, the DD velocity tomography method was introduced by Zhang and Thurber (2003) to do earthquake relocation and 3-D seismic velocity structure inversion at the same time. DD tomography jointly uses body-wave absolute arrival times, body-wave absolute differential times, and body-wave cross-correlation differential times. The resolution of seismic velocities in the region of earthquake sources can be enhanced via DD tomography (Zhang and Thurber, 2003, 2006).

Body-wave attenuation tomography uses absolute t* values (Eberhart-Phillips and Chadwick, 2002), where t^* is the whole path attenuation parameter (Shearer, 2002). Using a similar idea as DD tomography, we propose to use both absolute t* and event-pair differential t* (δt*) for 3-D attenuation structure tomography. The differential t* values include both absolute t* differences and δt* values derived from spectral ratios of adjacent event-pairs. We term this method DD attenuation (Q) tomography. DD Q tomography is expected to provide sharper structural resolution in regions of earthquake sources, analogous to DD velocity tomography.

In this study, we will develop the algorithm to derive δt* values from spectral ratios. Then we will conduct a series of test with synthetic data from The Geysers region to investigate the advantages of the new DD Q attenuation tomography method over the traditional Q attenuation tomography method.

5.3. Algorithm to derive event-pair δt*

Commonly used approaches to derive δt* involve calculating the spectral ratio between either event-pair records for same phase from the same source at different stations or phase pair records from different phases (P and S, S and SS, etc.) at the same station (Roth et al., 1999). These two types of δt* measurements are termed "station pair δt*" and "phase pair δt*." As discussed by Roth et al. (1999), the P- and S-wave spectra from the same earthquake can be written as:

$$P(\omega) = C_P\, F(\omega)\, I(\omega)\, R(\omega)\, A_P(\omega) \tag{5.1a}$$

$$S(\omega) = C_S\, F(\omega)\, I(\omega)\, R(\omega)\, A_S(\omega) \tag{5.1b}$$

where C_P and C_S are constants determined by the radiation pattern and geometric spreading for the P- and S-wave, respectively, F(ω) is the source spectrum, I(ω) is the instrument response, R(ω) is the crustal response, and $A_P(\omega)$ and $A_S(\omega)$ are the P and S attenuation functions, respectively. For the same earthquake source, the phase pair (P and S) δt* can be obtained by taking the spectral ratio under some reasonable assumptions. Also, Roth et al. (1999) showed that the station pair δt* can be generated from the slope of the log of the spectral ratio for a single phase from the same event.

Similar to measuring event-pair differential time in DD velocity tomography, we develop the algorithm to calculate the spectral ratios between adjacent event pairs and derive event-pair δt*. For two nearby earthquakes 1 and 2 recorded at station j, their spectra A_{1j} and A_{2j} can be written as

$$A_{1j}(f) = \Omega_1 S_j(f) \frac{1}{1+\left(\frac{f}{f_{c1}}\right)^{\alpha}} \exp(-\pi f t^*_{1j}) \tag{5.2a}$$

$$A_{2j}(f) = \Omega_2 S_j(f) \frac{1}{1+\left(\frac{f}{f_{c2}}\right)^{\alpha}} \exp(-\pi f t^*_{2j}) \tag{5.2b}$$

where the Ω_i's are the zero-frequency spectral amplitudes, α is the source spectrum fall-off, f is frequency, $S_j(f)$ is the site response at station j for frequency f, f_{ci} is corner frequency for event i, and t_{ij}* is the t* value for event i at station j. In standard attenuation studies (Scherbaum, 1990; Lees and Lindley, 1994; Rietbrock, 2001; Eberhart-Phillips and Chadwick, 2002; Bennington et al., 2008), (5.2) is solved for a single event or group of events to estimate t*, Ω, and f_c.

For the event-pair δt* value estimation, we don't need to solve (5.2) separately for each event. With the assumption that site responses for close-by events at a single station will be the same, we relate the event-pair spectral ratio $A_{1j}(f)/A_{2j}(f)$ to δt_j* by

$$\frac{A_{1j}(f)}{A_{2j}(f)} = \frac{\Omega_1}{\Omega_2}\frac{1+\left(\frac{f}{f_{c2}}\right)^{\alpha}}{1+\left(\frac{f}{f_{c1}}\right)^{\alpha}}\exp\left(-\pi f\delta t_j^*\right) \tag{5.3}$$

In (5.3), the unknown values for the event pair at station j are δt*, two corner frequencies (f_{c1}, f_{c2}) and the ratio between Ω_1 and Ω_2. Solving (5.3) for δt* can avoid solving (5.2) individually for each t* and then deriving δt* by subtraction, which more likely would be affected by errors introduced by site response and the source spectrum model.

When solving (5.3), we use the least-squares inversion with zeroth order Tikhonov regularization (Aster et al., 2003). In the δt^* inversion, we regard the Ωs as known constants inverted from a previous t^* inversion using (5.2). Following the Brune formula, $\alpha = 2$ is chosen in the inversion for δt^*. The unknown parameters are the two corner frequencies and the δt^* for each station. The damped least square solution for this inversion can be expressed as

$$\Delta m = (G^T * G + \eta I)^{-1} * G^T * r \tag{5.4}$$

where Δm is the vector of unknown parameter perturbations, G is the sensitivity matrix, G^T is the transpose of G, η is the damping value, I is an identity matrix, and r is the residual between the calculated and observed spectral amplitude ratio. In order to construct the sensitivity matrix, we first take the natural log of (5.3):

$$\ln(rat_i) = \ln(C_i) + \ln\left(1+\left(\frac{f}{f_{c2}}\right)^{\alpha}\right) - \ln\left(1+\left(\frac{f}{f_{c1}}\right)^{\alpha}\right) - \pi f\delta t_i^* \tag{5.5}$$

where rat_i is the spectral ratio between event 1 and event 2 at station i, C_i is the zero-frequency spectral amplitude ratio between events 1 and 2 at station i, δt_i^* is the differential t^* value at station i, f is frequency, f_{c1} and f_{c2} are corner frequencies for events 1 and 2, respectively, and α is the source spectrum fall-off. The partial derivatives for each unknown parameter in (5.5) are given by the following equations

$$\frac{\partial \ln\left(rat_i\right)}{\partial f_{c2}} = -\alpha f^{\alpha} f_{c2}^{-(\alpha+1)}\left(1+\left(\frac{\mathrm{f}}{f_{c2}}\right)^{\alpha}\right)^{-1} \tag{5.6a}$$

$$\frac{\partial \ln\left(rat_i\right)}{\partial f_{c1}} = \alpha f^{\alpha} f_{c1}^{-(\alpha+1)}\left(1+\left(\frac{\mathrm{f}}{f_{c1}}\right)^{\alpha}\right)^{-1} \tag{5.6b}$$

$$\frac{\partial \ln\left(rat_i\right)}{\partial \delta t_i^*} = -\pi f \tag{5.6c}$$

Then the matrices and vectors in (5.4) can be expressed as

$$G = \begin{bmatrix} \frac{\partial \ln\left(rat_1\right)}{\partial f_{c1}} & \frac{\partial \ln\left(rat_1\right)}{\partial f_{c2}} & \frac{\partial \ln\left(rat_1\right)}{\partial \delta t_1^*} & 0 & \cdots & 0 \\ \frac{\partial \ln\left(rat_2\right)}{\partial f_{c1}} & \frac{\partial \ln\left(rat_2\right)}{\partial f_{c2}} & 0 & \frac{\partial \ln\left(rat_2\right)}{\partial \delta t_2^*} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \ln\left(rat_n\right)}{\partial f_{c1}} & \cdots \frac{\partial \ln\left(rat_n\right)}{\partial f_{c1}} & 0 & 0 & \cdots & \frac{\partial \ln\left(rat_n\right)}{\partial \delta t_n^*} \end{bmatrix} \tag{5.7a}$$

$$\Delta m = \begin{bmatrix} \Delta f_{c1} \\ \Delta f_{c2} \\ \Delta \delta t_1^* \\ \vdots \\ \Delta \delta t_n^* \end{bmatrix} \tag{5.7b}$$

$$r = \begin{bmatrix} rat_1 - rat_{obs_1} \\ \vdots \\ rat_n - rat_{obs_n} \end{bmatrix} \tag{5.7c}$$

where Δf_{c1} and Δf_{c2} are perturbations to f_{c1} and f_{c2}, respectively, $\Delta \delta t_i^*$ is the perturbation to δt_i^*, and rat_{obs_i} is the observed spectral amplitude ratio between event 1 and event 2 at station i.

5.4. Synthetic tests for solving for event-pair δt*

A simplified simulation is carried out to illustrate the effectiveness of solving for δt* with (5.3) compared to (5.2). For each event, the site response is set to be 1 for all frequencies at all stations and the Ω values are also set to 1. Following the Brune model, we choose $\alpha = 2$ in (5.2) and (5.3). We assume a pair of nearby earthquakes are recorded by 8 stations. The correct corner frequencies for events 1 and 2 are 4 Hz and 6 Hz, respectively. t* values for event 1 at the 8 stations are chosen to range from 0.02 to 0.09, and for event 2 they range from 0.025 to 0.13. The true spectra for the two events at these 8 stations are shown in Figure 5.1. Then we use δt* values shifted by -0.005 s from the correct values and shifted corner frequencies of 4.5 and 5.5 Hz for events 1 and 2, respectively as initial values for the inversion.

5.4.1. Site responses effects on event-pair δt^* inversion

When solving for event-pair δt^* with (5.3), site responses cancel out and will not affect the inversion process. After 15 iterations of damped least-squares inversions with $\eta = 0.4$. Equation (5.3) provides corner frequency solutions within 1.9 % or better (3.97 Hz and 5.94 for event 1 and event 2, respectively). As shown in Fig 5.2, the δt* solutions are within 2.4% or better compared to the correct values.

When using (5.2) to solve for t^* and f_c for each event individually, we need to consider the effect of site responses. For simplicity, we treat the site response as constants. For the inversion, we set all site responses to be off by 5% and other initial values the same as in the previous inversion for (5.3). Corner frequency solutions by (5.2) for event 1 and event 2 are 4.16 and 6.29 Hz, respectively. The errors for corner frequency solutions are 3.9 – 4.8%. The inverted δt^* for all stations are shown in Figure 5.3a. The errors for all stations are greater than those solved by (5.3). When we set the site responses to be 10% off, the errors for corner frequencies are 8.2 – 10.1% and the errors for δt^* can be as great as 6.4% (Figure 5.3b). According to this experiment, we can conclude that solving δt* with (5.3) is more effective than taking differences from individually solved t* from (5.2), as deriving δt^* from individually solved t^* is very sensitive to site response errors whereas solutions from the event-pair spectral ratio method are not.

5.4.2. Source spectrum model sensitivity

In the previous synthetic tests, we assumed the source spectrum model follows the Brune model ($\alpha = 2$). It is important to investigate the sensitivity to different source spectrum models when solving for event-pair δt^* with $\alpha = 2$ in (5.3). We generated two additional sets of synthetic spectral data for these two events for 8 stations with source spectrum models $\alpha = 1.5$ and 2.5, but assumed $\alpha = 2$ in the inversion. The inverted event-pair δt^* values for all stations are shown in Figure 5.4. The errors are the largest for station 1, 5.8% ($\alpha = 1.5$) and 9.5% ($\alpha = 2.5$). As the δt^* value for station 1 is the smallest (0.005), the absolute errors for station 1 are actually quite small at 0.00029 and 0.00047, respectively. For the larger δt^* value at station 8 (0.04), the errors are 0.7% and 1.2% for $\alpha = 1.5$ and 2.5, respectively. The corner frequency solutions are within 8.4% and 7.1% for $\alpha = 1.5$ and 2.5, respectively.

For comparison, we also test the sensitivity of solutions from (5.2) to the source spectrum model. With the same synthetic data sets, we use (5.2) with $\alpha = 2$ to invert for t^* for each event at all stations individually. True site responses are used in the inversion. For $\alpha = 1.5$, errors for inverted t^* are up to 54.9% for event 1 at station 1 and 43.4% for event 2 at station 1. For $\alpha = 2.5$, errors for inverted t^* at station 1 are 54.8% for event 1 and 43.1% for event 2. As for the δt^* values, the errors are 2.2% and 3.9%, respectively, for $\alpha = 1.5$ and 2.5. We observe that even though the errors for individually solved t^* values are substantial they are always in the same direction with our experiment setup. As a result, the errors of individually solved t^*s are largely canceled out when deriving the final δt^* by subtraction. The corner frequency solutions are within 10.3% and 9.3% for $\alpha = 1.5$ and 2.5, respectively.

According to these test results, solving δt^* with (5.3) and the Brune model does show some sensitivity to different source spectrum models ($\alpha = 1.5$ and 2.5). The errors for the δt^* solutions from (5.3) are reasonably small, although the percent error can be large when the absolute value of δt^* is very small. The errors in individually solved t^*s can be as large as ~55% when the true source spectrum is not the Brune model. However, in our experiment, the errors for the individually solved t^*s are in the same direction, which results in smaller errors in the derived δt^* values. For corner frequencies, solutions from (5.3) show less error compared with solutions from (5.2).

5.4.3. Ω value sensitivity

In the previous synthetic tests, the Ωs are all set to be 1 for simplicity. In reality, we need to include Ω in the inversion process. If we regard Ω_{ij} as an unknown parameter for event j at station i, then the number of unknown parameters for each event-pair increases from n+2 to 2n+2, where n is the number of stations. The number of constraints, the spectral ratio observations, is only n. Thus, it would be impossible to solve for δt^* properly for such an ill-conditioned problem. As in previous standard attenuation tomography studies (Rietbrock, 2001; Eberhart-Phillips and Chadwick, 2002; Bennington et al., 2008), Ω is initially inverted together with t^* and f_c for each event-station pair. We assume the inverted Ω values are close enough to the true values and check that our inversion with (5.3) is not that sensitive to the Ω value uncertainties from the previous inversions.

To investigate the sensitivity of our method to Ω values with uncertainties, we conduct synthetic tests to solve event-pair δt^* with variations in the Ω ratios. When generating synthetic spectral data, we use Ω values of 200, 220, 240, 260, 280, 300, 320, and 340 for event 1 and Ω values of 300, 320, 340, 360, 380, 400, 420, and 440 for event 2 at the 8 stations. The true and initial corner frequencies and true and initial δt^* values are the same as in the previous synthetic tests. When solving (5.3), we used Ω ratios between event 1 and 2 to be 80%, 90%, 100%, 110%, and 120% of the true ratios. The inversion results with damping value $\eta = 0.4$ for all 5 tests are shown in Table 5.1. When Ω ratios in the inversion are smaller than the true values, the errors are larger compared to those when the Ω ratios are larger than the true values. When Ω ratios are 80% of the true values, the error in δt^* for station 1 is 13.8%. As the δt^* value at station 1 is very small (0.005), this error percentage corresponds to a δt^* error of 0.0007. For the test with 90% of the true Ω ratios, the δt^* errors are all within 5.2%. When the Ω ratios are greater than the true values, the errors are within 1.8% or better. According to these results, the δt^* inversion with event-pair spectral ratio is stable when there are variations in the Ω ratios.

5.5. Synthetic DD Q tomography tests with The Geysers dataset

5.5.1. Checkerboard test

Ample fluid related seismicity in The Geysers makes it a suitable region to conduct the DD Q tomography study. We conduct a synthetic test with earthquakes that occurred in The Geysers region to investigate the advantages of the DD Q tomography over the traditional Q tomography. As shown in Figure 5.5, 865 events that occurred in 2010 are selected and 37 3-component stations of the Lawrence Berkeley National Laboratory seismic network are used to generate the synthetic t^* values. For the DD Q tomography, we need to define earthquake clusters in order to calculate the δt^* of event pairs within each cluster. For the 865 selected earthquakes, we used the k-nearest neighbor algorithm (Mitchell 1997) to find cluster centers and earthquakes within each cluster. With k = 10 and a cluster radius of 2 km, 27 clusters including 602 events are obtained.

With a coordinate origin at (38.835°N, 122.784°W) origin and 0° rotation, inversion nodes are projected into Cartesian coordinates. Nodes in the Y direction from South to North are: -9.0, -5.4, -1.8, 0.0, and 3.6 km. Nodes in the X direction from West to East are: -9.0, -5.4, -3.6, -2.9, -2.3, -1.6, -1.0, -0.3, 0.3, 1.0, 1.6, 2.3, 2.9, 3.6, 5.4, and 9.0 km. In the vertical direction, the layers are at 0.0, 0.7, 1.4, 2.2, 2.9, 3.6, and 5.4 km depths. To generate synthetic data, we first created synthetic V_P model and Q_P models for the study region. For simplicity, we used a homogeneous V_P model with a value of 4 km/s and a Q_P model with a value of 75. Then we created a checkerboard anomaly model by assigning positive and negative Q_P perturbations of 30% to alternating groups of eight nodes (Figure 5.6). Then synthetic t^* data are generated with the homogeneous V_P model and checkerboard Q_P model for all 37 stations. For event pairs within the same clusters, δt^* values are derived from the synthetic t^*.

Checkerboard tests are carried out for both the traditional Q and the DD Q tomography methods with the homogeneous Q_P model as the initial model. Checkerboard results for cross-sections with reasonable earthquake distributions (Y = -5.4, -1.8, and 0.0 km) are shown in Figure 5.7. To compare the differences between the recovered anomalies between the DD Q and the traditional Q methods, we calculated the anomaly misfit difference by

$$misfit_{diff} = \frac{abs\left(anomaly_{ddq} - anomaly_{true}\right) - abs\left(anomaly_{q} - anomaly_{true}\right)}{abs\left(anomaly_{true}\right)} \tag{5.8}$$

where $anomaly_q$ and $anomal_{ddq}$ are the recovered anomaly percentage for the traditional Q tomography and the DD Q tomography, respectively, $anomaly_{true}$ is the true anomaly percentage, and $misfit_{diff}$ is the anomaly recovery difference between the DD Q tomography and the traditional Q tomography. The anomaly misfit differences are calculated for regions where the anomaly is recovered by at least one method. When the event-pair δt^* data are used in the inversion, the mean anomaly recovery misfits decrease by 6.8%, 8.9%, and 6.6% for the Y = -5.4, -1.8, and 0.0 km cross-sections, respectively. The checkerboard test demonstrates that DD Q tomography can improve the source region resolution over traditional Q tomography.

5.5.2. Temporal change recovery test

Gritto and Jarpe (2014) showed temporal V_P/V_S ratio changes during geothermal powerplant operations in The Geysers. It is possible that there are temporal variations in the Q values as well.

Thus, detecting changes in Q values is essential for monitoring temporal Q structure changes in The Geysers region. We conduct a series of synthetic tests to investigate the ability of each Q tomography method to recover anomalies with different amplitudes. In this synthetic test, we created two Q_P anomalies located at 0 to 2.9 km depth (Figure 5.8). The low Q_P anomaly locates between X = -2.9 and -0.3 km and the high Q_P anomaly locates between X = 1.0 and 3.6 km. For each anomaly recovery test, synthetic t^* data are generated with a 1D V_P model and the Q_P model with the anomalies for all stations. In the 1D V_P model, V_P is 4 km/s at 0 km depth and increases 0.2 km/s for every 1 km depth. By assuming the 20% anomaly amplitude is the normal state, we generated synthetic t^* data for Q_P anomaly amplitudes of 10% and 30%, and then use the 20% anomaly amplitude Q_P model as initial model to test the recovery of these two perturbed Q_P models with both the traditional Q and the DD Q methods.

To quantify the anomaly recovery improvements of DD Q method over traditional Q method, the recovery differences between the DD Q method and traditional Q method are calculated by

$$r_{diff} = \frac{abs(r_q - anomaly_{true}) - abs(r_{ddq} - anomaly_{true})}{abs(\Delta anomaly)} \tag{5.9}$$

where r_q and r_{ddq} are the recovered anomaly for traditional Q tomography and DD Q tomography, respectively, $anomaly_{true}$ is the true anomaly, $\Delta anomaly$ is the perturbation of the Q_P anomaly, and r_{diff} is the anomaly recovery improvement percentage of the DD Q tomography over the traditional Q tomography. Overall, the DD Q tomography improves the recovery percentage compared to the traditional Q tomography in regions with clustered events (Figure 5.9, 5.10). For nodes within each Q_P anomaly, the mean recovery for each Q_P anomaly in cross-sections Y = -5.4, -1.8, and 0.0 km are shown in Table 5.2. For the low Q_P anomaly, the average recovery improvement is 9.9% and 8.0%, respectively, for the 10% anomaly amplitude decrease and increase. For the high Q_P anomaly, the average recovery improvement is 8.7% and 8.9%, respectively, for the 10% anomaly amplitude decrease and increase. These synthetic recovery tests show the improvement of DD Q tomography over traditional Q tomography when the Q_P anomaly amplitude changes.

Based on the synthetic tests with The Geysers dataset, the DD Q tomography improves the Q_P resolution for regions with adequate earthquakes compared to the traditional Q tomography. The next step will be using real data from The Geysers region and applying the DD Q tomography to investigate the Q_P structures and its possible temporal change.

5.6. Discussion

Compared to deriving δt^* with individually inverted t^* values, solving δt^* using event-pair spectral ratios has one main advantage: because the events are close to each other, the site responses for the event pair at the same station should completely cancel out. When the source spectrum model changes, the errors in δt^* solved by (5.3) are reasonably small.

With a well-established method to solve t^*, f_c, and Ω for each event individually (Eberhart-Phillips and Chadwick, 2002; Bennington et al., 2008; Bisrat et al., 2014), we can use inverted Ω ratios as constants, derived δt^* values from inverted t^* as initial δt^* values, and inverted f_c values as initial f_c values in the inversion. In this way, the initial values should be reasonable estimates of the true δt^* and f_c.

This study shows the reliability of solving event-pair δt^* with spectral ratios and the potential source region resolution improvement when δt^* data are included in the Q tomography. As a next step, we need to modify the traditional attenuation tomography algorithm in order to integrate the derived δt^* and inverted event-pair δt^* in the attenuation inversion process. Derived δt^* is calculated using individually inverted t^* values and the inverted δt^* is solved with our event-pair spectral ratio method. Similar to ideas in tomoDD, including derived and inverted event-pair δt^* should improve the resolution in source regions of the event-pairs.

5.7. Acknowledgements

This research is supported in part by the USGS, Department of the Interior, under USGS Award Number G14AP00056 to the University of Wisconsin-Madison. Thanks for the support from NSF, Division of Earth Sciences, under NSF Award Number 1724685. The support of a student assistantship from the Mark and Carol Ann Solien Graduate Assistantship and George P. Woollard-Sigmund I. Hammer Memorial Fund in Geology & Geophysics of the University of Wisconsin-Madison are also acknowledged.

5.8. References

Aster, R. C., Borchers, B., Thurber, C. H., 2013, Parameter Estimation and Inverse Problems. 2nd edition, Elsevier, 93 pp.

Bennington, N., Thurber, C. and Roecker, S., 2008, Three-dimensional seismic attenuation structure around the SAFOD site, Parkfield, California. Bull. Seismol. Soc. Am., 98(6), 2,934-2,947.

Deichmann, N., and Garcia-Fernandez, M., 1992, Rupture geometry from high-precision relative hypocentre locations of microearthquake clusters. Geophys. J. Int., 110(3), 501-517.

Eberhart-Phillips, D., and Chadwick, M. 2002, Three-dimensional attenuation model of the shallow Hikurangi subduction zone in the Raukumara Peninsula, New Zealand. J. Geophys. Res.: Solid Earth, 107(B2), ESE 3-1 – ESE 3-15.

Gillard, D., Rubin, A. M., and Okubo, P., 1998, Highly concentrated seismicity caused by deformation of Kilauea's deep magma system. Nature 384, 343-346, doi:10.1038/384343a0.

Got, J.-L., Fréchet, J., and Klein, F. W., 1994, Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea. J. Geophys. Res. 99, 15,375-15,386.

Gritto, R. and Jarpe, S.P., 2014, Temporal variations of Vp/Vs-ratio at The Geysers geothermal field, USA. Geothermics, 52, pp.112-119.

Lees, J. M., and Lindley, G. T., 1994. Three-dimensional attenuation tomography at Loma Prieta: Inversion of t* for Q. J. Geophys. Res.: Solid Earth, 99(B4), 6,843-6,863.

Mitchell, T.M., 1997, Machine learning. Burr Ridge, IL: McGraw Hill, 45(37), 231 pp.

Rietbrock, A., 2001, P wave attenuation structure in the fault area of the 1995 Kobe earthquake. J. Geophys. Res.: Solid Earth, 106(B3), 4,141-4,154.

Roth, E. G., Wiens, D. A., Dorman, L. M., Hildebrand, J., and Webb, S. C., 1999, Seismic attenuation tomography of the Tonga-Fiji region using phase pair methods. J. Geophys. Res.: Solid Earth, 104(B3), 4,795-4,809.

Rubin, A. M., Gillard, D., and Got, J.-L., 1999, Streaks of microearthquakes along creeping faults. Nature 400, 635-641, doi:10.1038/23196.

Scherbaum, F., 1990, Combined inversion for the three-dimensional Q structure and source parameters using microearthquake spectra. J. Geophys. Res., 95, 12,423-12,438, doi:10.1029/JB095iB08p12423.

Shearer, P.M., 2009, Introduction to seismology. Cambridge University Press, 163 pp.

Slunga, R., Rögnvaldsson, S. T, and Bödvarsson, R., 1995, Absolute and relative locations of similar events with application to microearthquakes in southern Iceland. Geophys. J. Int. 123, 409-419, doi: 10.1111/j.1365-246X.1995.tb06862.

Waldhauser, F. and Ellsworth, W.L., 2000, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. Bull. Seismol. Soc. Am., 90(6), 1,353-1,368.

Waldhauser, F., Ellsworth, W. L., and Cole, A., 1999, Slip-parallel seismic lineations on the northern Hayward Fault, California. Geophys. Res. Lett. 26, 3,525-3,528.

Zhang, H., and Thurber, C., 2006, Development and applications of double-difference seismic tomography. Pure Appl. Geophys, 163(2), 373-403.

Zhang, H., Thurber, C. H., 2003, Double-difference tomography: The method and its application to the Hayward fault, California. Bull. Seismol. Soc. Am., 93(5), 1,875–1,889.

5.9. Tables and figures

Table 5.1 δt^* error percentages for solutions with different percentage of true Ω ratios

Ω ratios	80%	90%	100%	110%	120%
Station 1	13.8	5.2	1.3	0.5	1.8
Station 2	6.9	2.6	0.7	0.3	0.9
Station 3	4.6	1.7	0.4	0.2	0.6
Station 4	3.4	1.3	0.3	0.1	0.5
Station 5	2.8	1.0	0.3	0.1	0.4
Station 6	2.3	0.9	0.2	0.1	0.3
Station 7	2.0	0.7	0.2	0.1	0.3
Station 8	1.7	0.7	0.2	0.1	0.2

Table 5.2. Q_P anomaly recovery for different anomaly perturbation amplitudes.

Q_P anomaly amplitude change	Y (km)	Low Q_P recovery		Recovery improvement	High Q_P recovery		Recovery improvement
		Q	DD Q		Q	DD Q	
-10%	-5.4	68.0%	74.8%	10.0%	38.9%	44.6%	14.7%
	-1.8	66.6%	71.1%	6.8%	86.1%	89.6%	4.1%
	0.0	41.4%	46.8%	13.0%	70.9%	76.0%	7.2%
10%	-5.4	76.0%	80.4%	5.8%	34.6%	39.7%	14.7%
	-1.8	72.3%	76.0%	5.1%	84.3%	88.3%	4.7%
	0.0	47.8%	54.1%	13.2%	67.9%	72.9%	7.4%

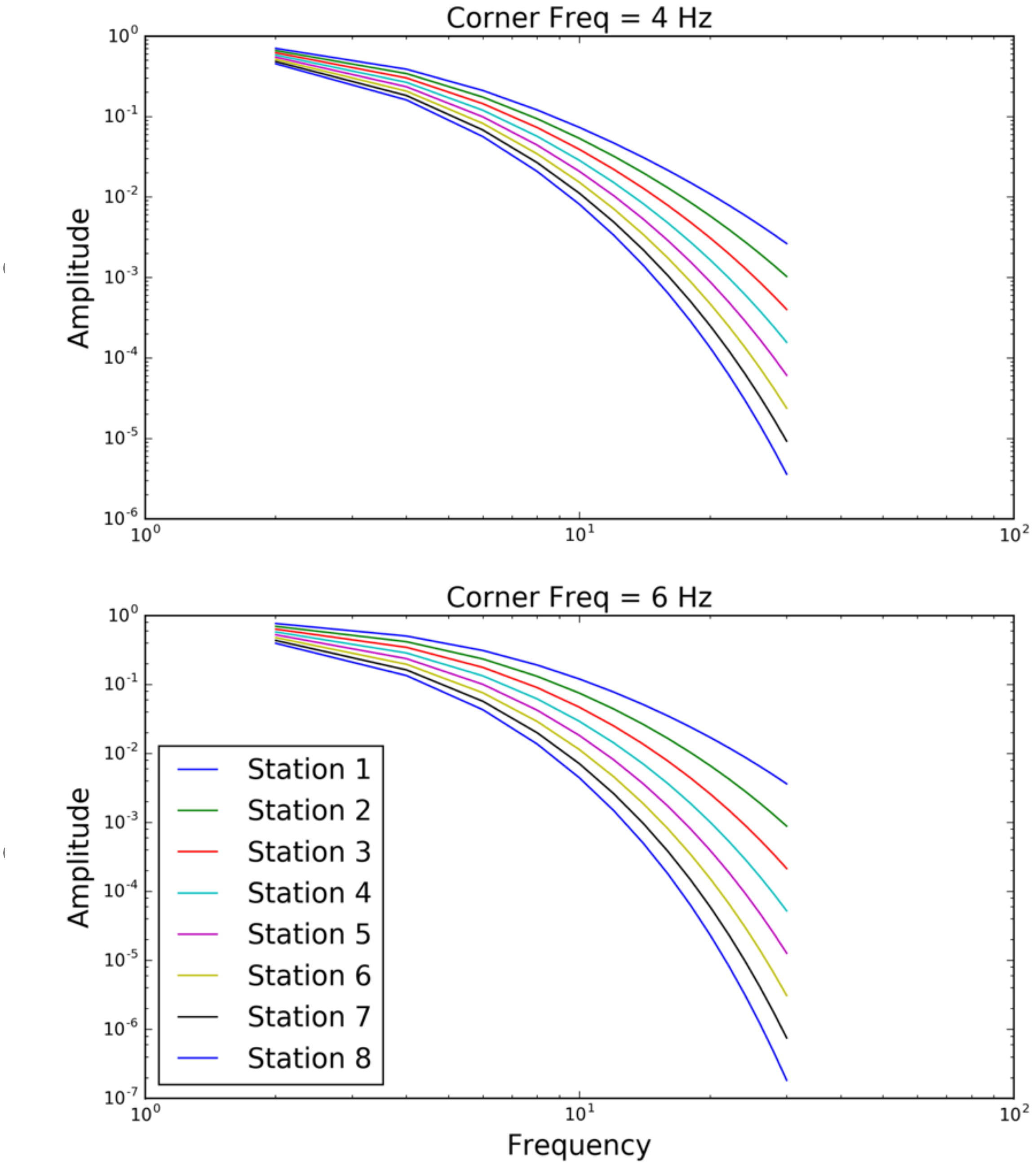


Figure 5.1 Synthetic spectra for (a) event 1 and (b) event 2 for the 8 stations.

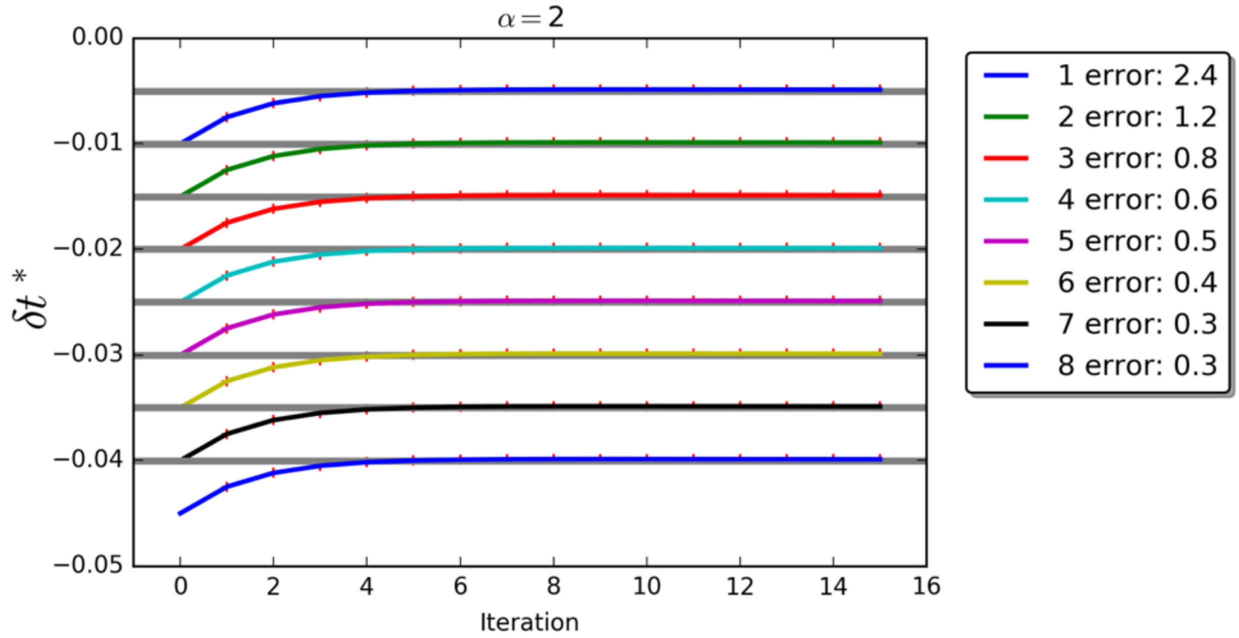


Figure 5.2 Convergence of the δt^* solutions from the event-pair spectral ratio method. Source spectral model is the Brune model, $\alpha = 2$. The errors are percentages from the inverted δt^* to the true δt^*.

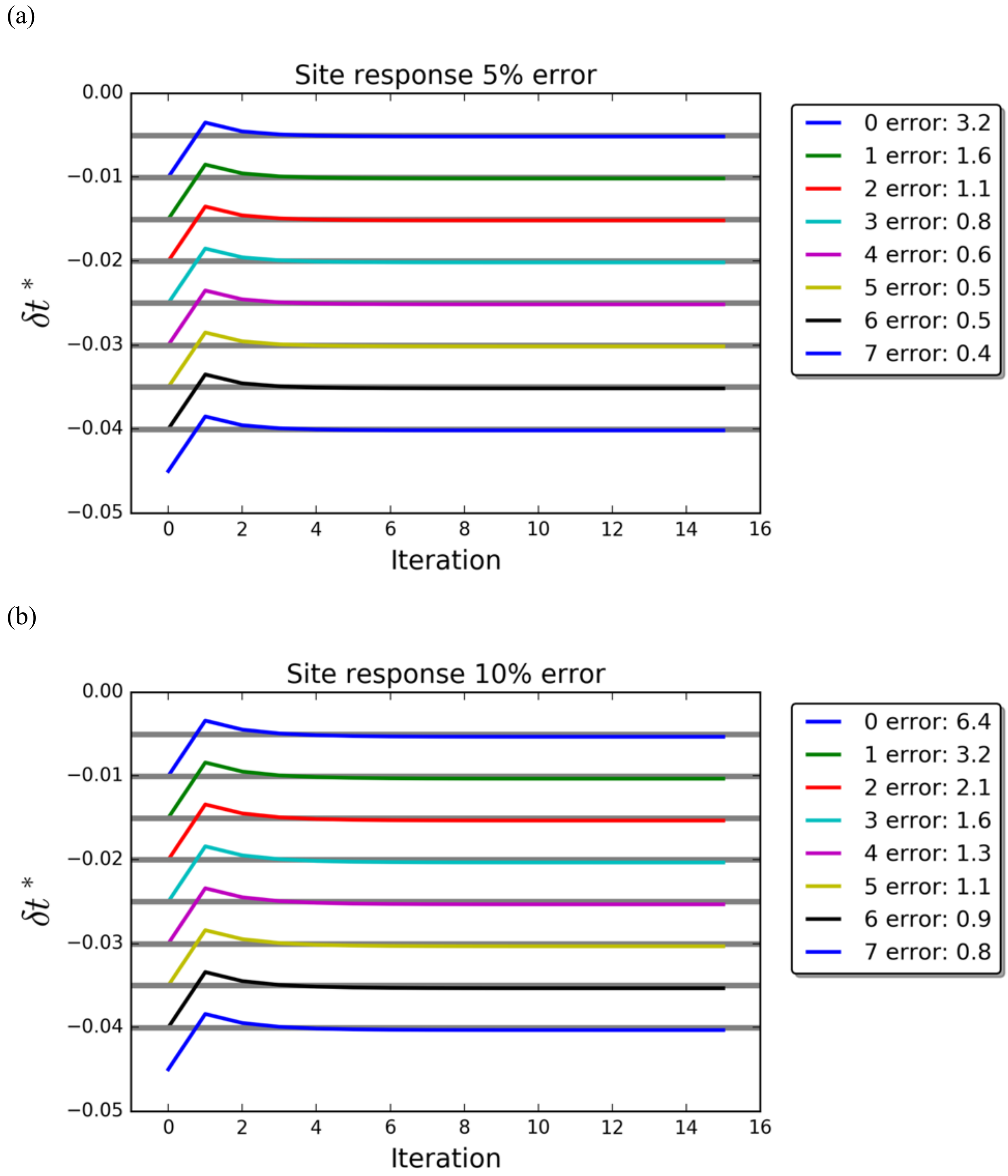


Figure 5.3 δt^* solved using individually derived t^*. The source spectral model is the Brune model, $\alpha = 2$. The errors are percentages from the inverted δt^* to the true δt^*. (a) Inversion results with 5% error in site responses. (b) Inversion results with 10% error in site responses.

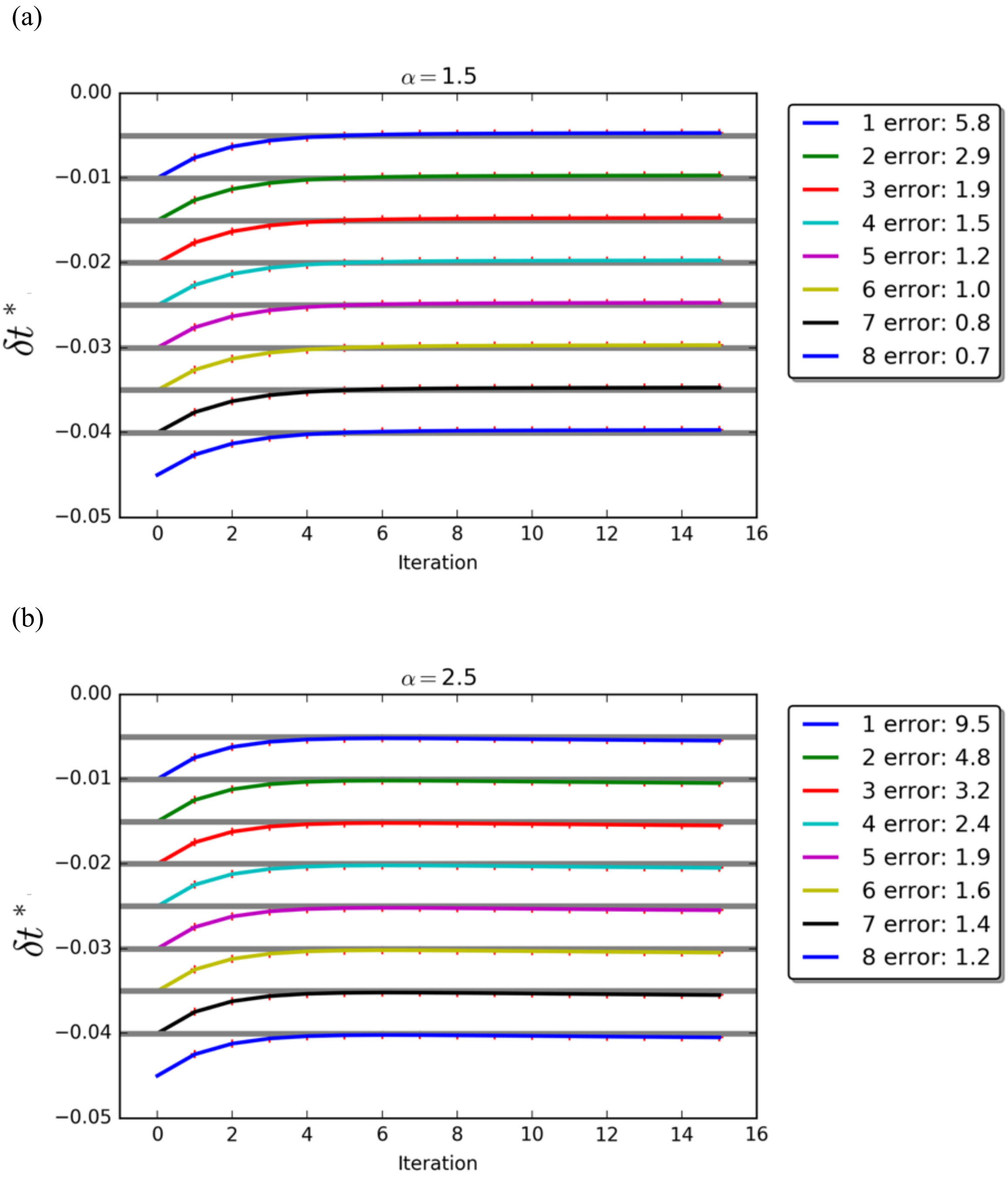


Figure 5.4 δt^* solutions from the event-pair spectral ratio method for an incorrect α. In the inversion, the source spectrum model is assumed as to be the Brune model, $\alpha = 2$. The errors are percentages from the inverted δt^* to the true δt^*. (a) Solutions for synthetic spectral data generated with $\alpha = 1.5$. (b) Solutions for synthetic spectral data generated with $\alpha = 2.5$.

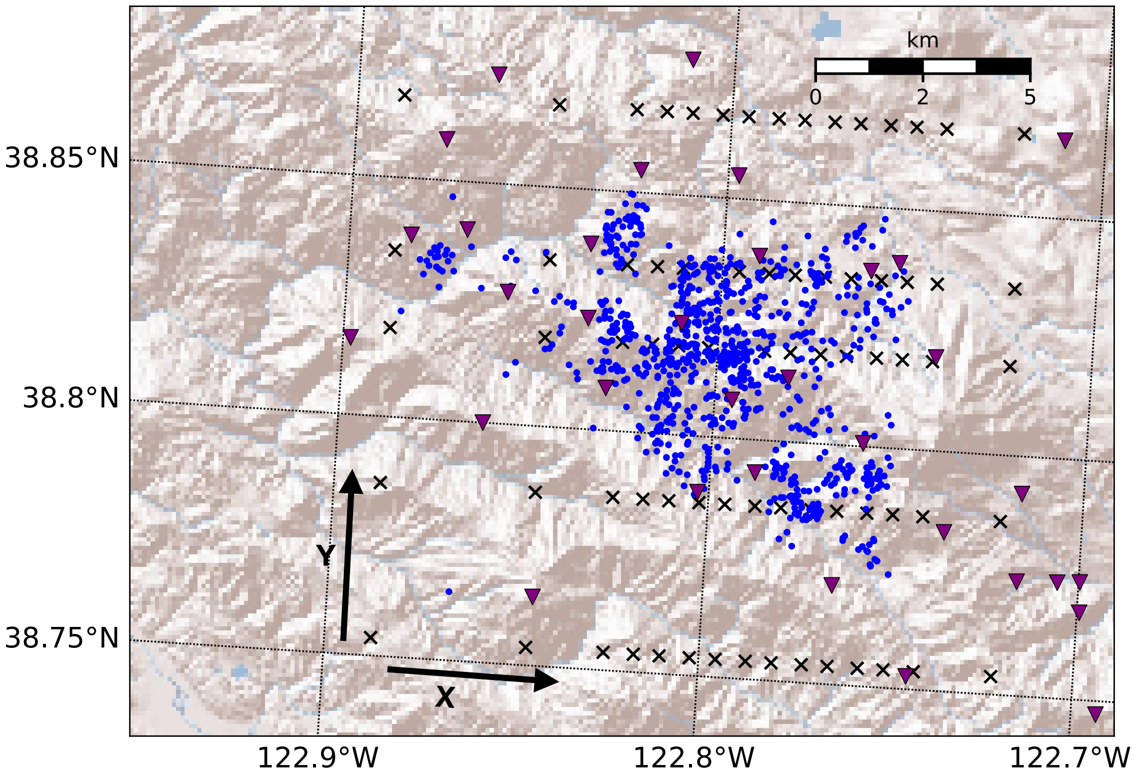
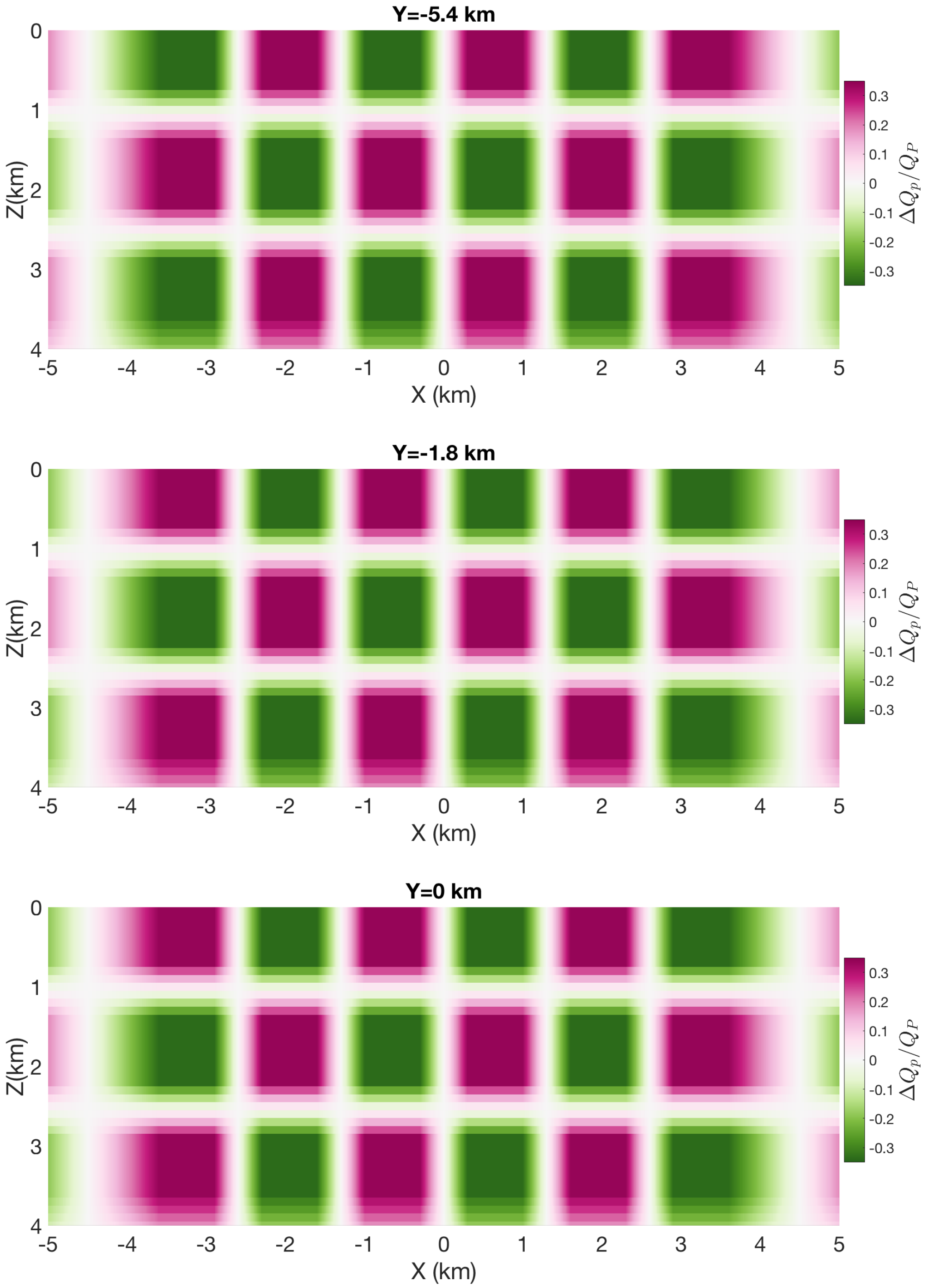


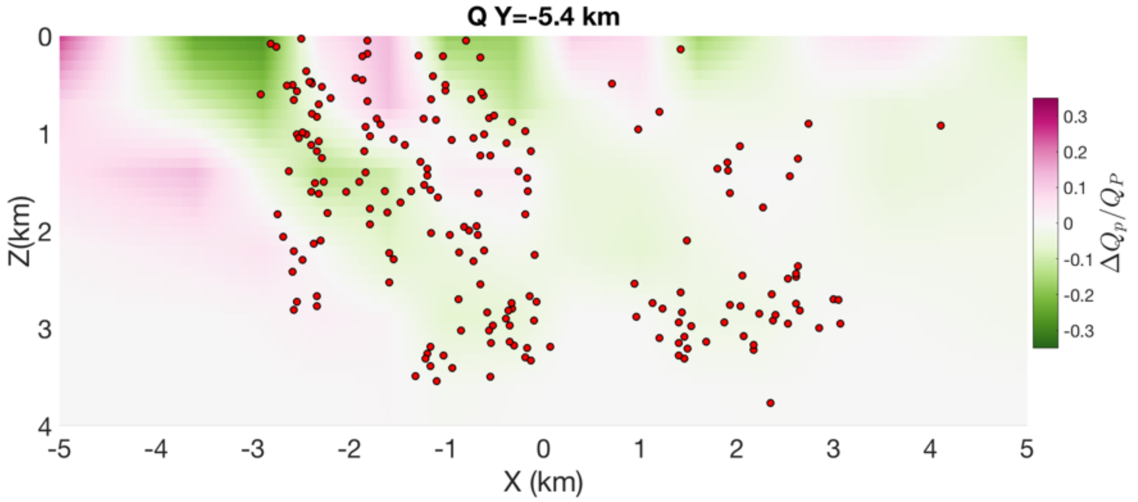
Figure 5.5 Map view of all events and stations in The Geysers study region. Purple triangles: 37 stations. Blue dots: 865 selected earthquakes. Black crosses: inversion nodes used in the Q tomography. Nodes in the Y direction from South to North are: -9.0, -5.4, -1.8, 0.0, and 3.6 km. Nodes in the X direction from West to East are: -9.0, -5.4, -3.6, -2.9, -2.3, -1.6, -1.0, -0.3, 0.3, 1.0, 1.6, 2.3, 2.9, 3.6, 5.4, and 9.0 km.


Y=-5.4 km

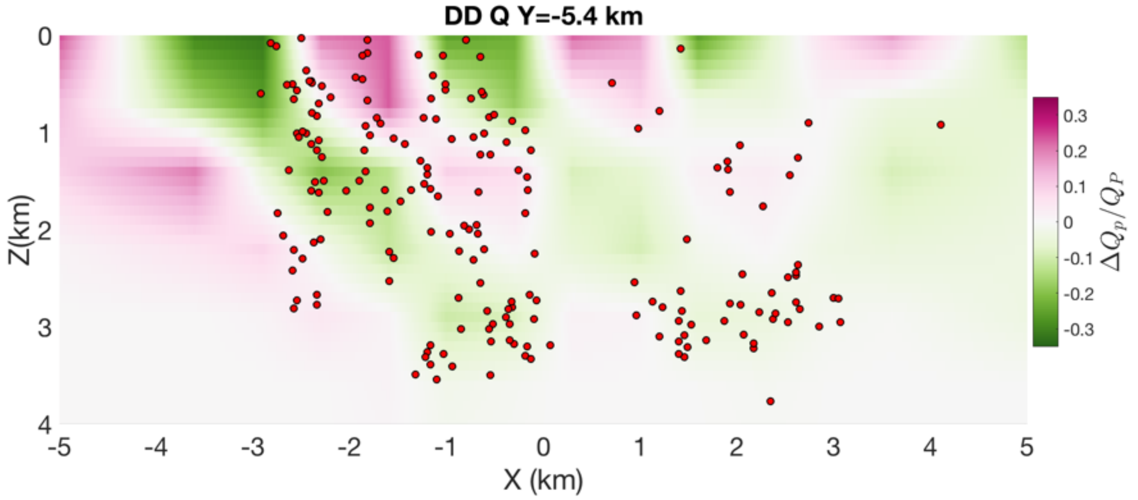
Y=-1.8 km

Y=0 km

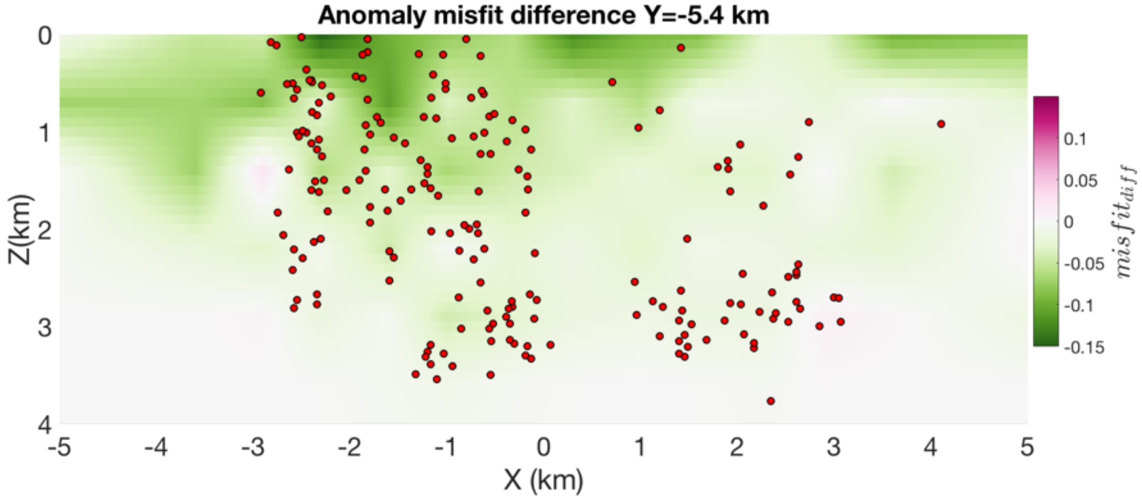
Figure 5.6 Cross-section plots for true checkerboard anomaly for the Q_P model.



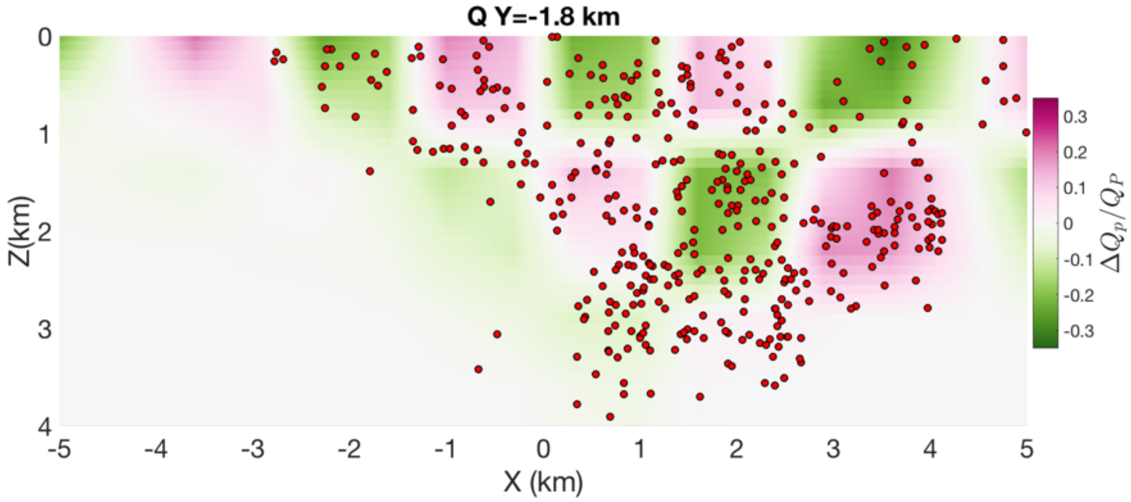
Q Y=-5.4 km

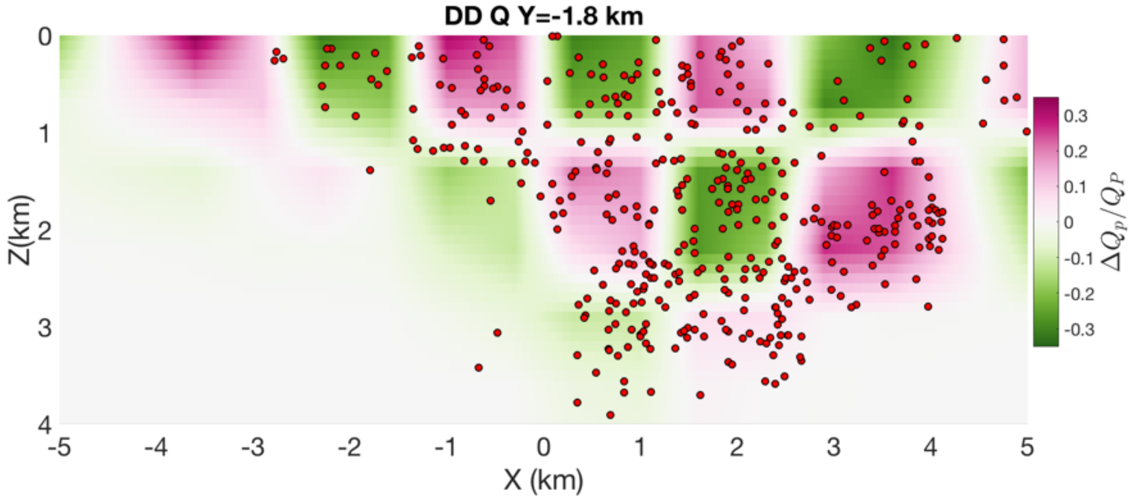


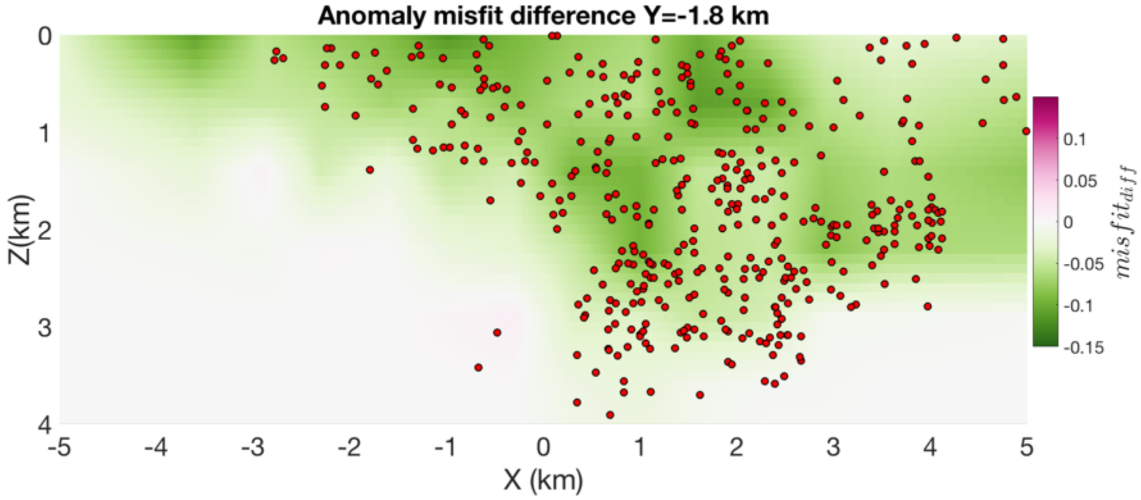
DD Q Y=-5.4 km



Anomaly misfit difference Y=-5.4 km


Q Y=-1.8 km
Z(km)
X (km)
0
1
2
3
4
-5
-4
-3
-2
-1
0
1
2
3
4
5
$\Delta Q_p/Q_P$
0.3
0.2
0.1
0
-0.1
-0.2
-0.3


DD Q Y=-1.8 km
Z(km)
X (km)
0
1
2
3
4
-5
-4
-3
-2
-1
0
1
2
3
4
5
$\Delta Q_p/Q_P$
0.3
0.2
0.1
0
-0.1
-0.2
-0.3


Anomaly misfit difference Y=-1.8 km
Z(km)
X (km)
0
1
2
3
4
-5
-4
-3
-2
-1
0
1
2
3
4
5
$misfit_{diff}$
0.1
0.05
0
-0.05
-0.1
-0.15

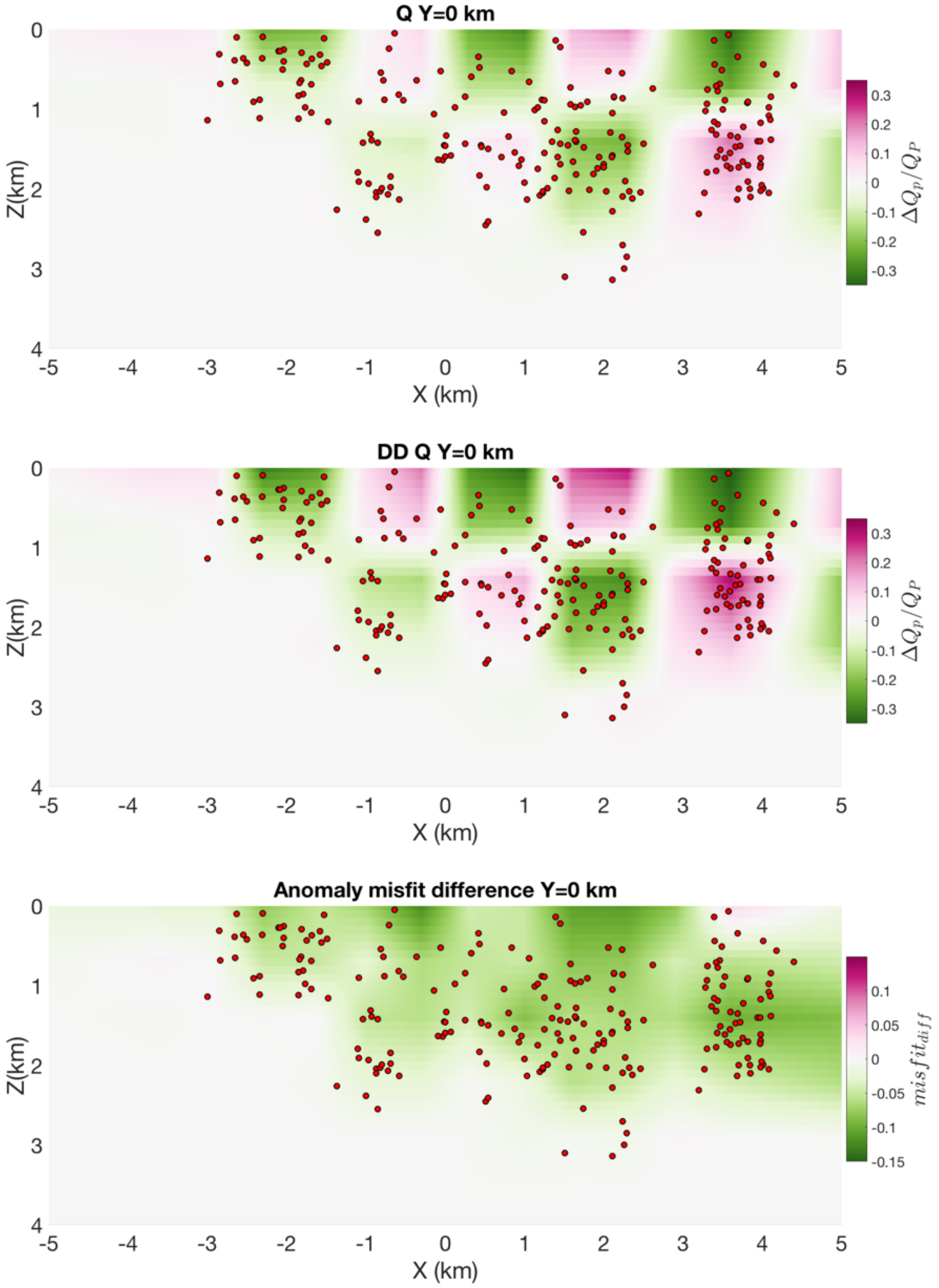


Figure 5.7 Cross-section plots for the checkerboard recovery results. Within each three-plot group, the first row shows the checkerboard results from the traditional Q tomography, the second row

shows the checkerboard results from the DD Q tomography, and the last row shows the anomaly misfit difference between the DD Q tomography and the traditional tomography (negative means decrease in anomaly misfit). Red dots: earthquakes.

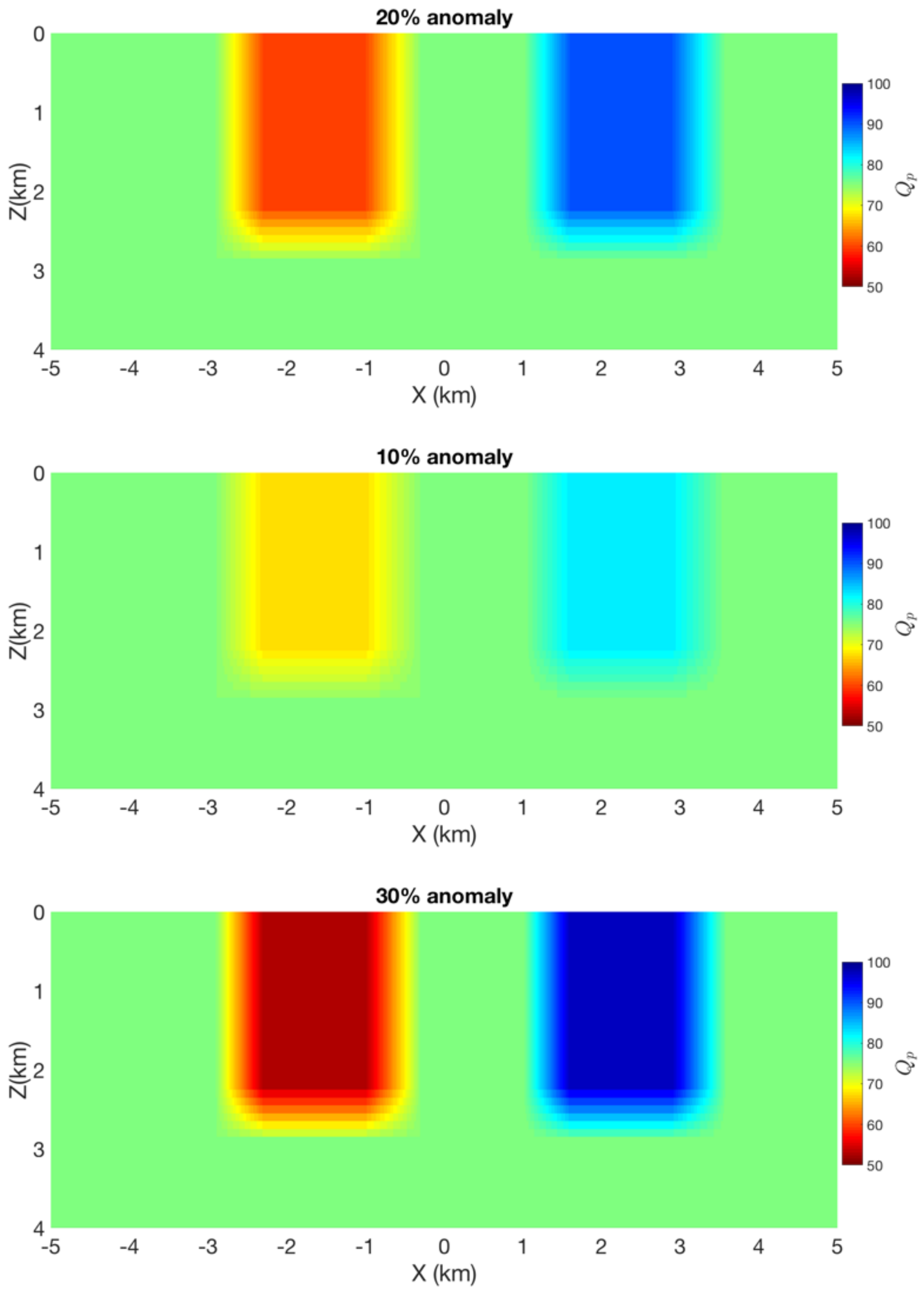
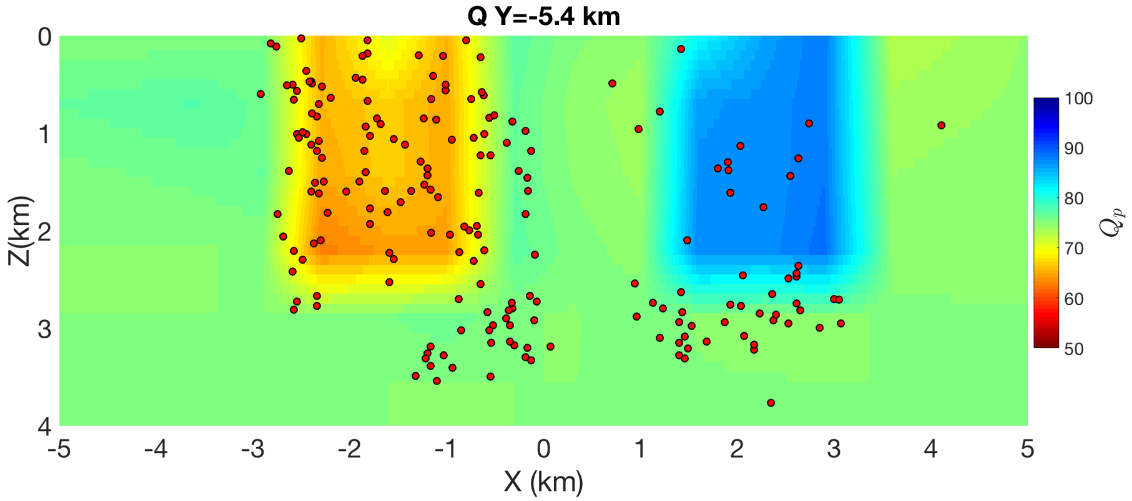
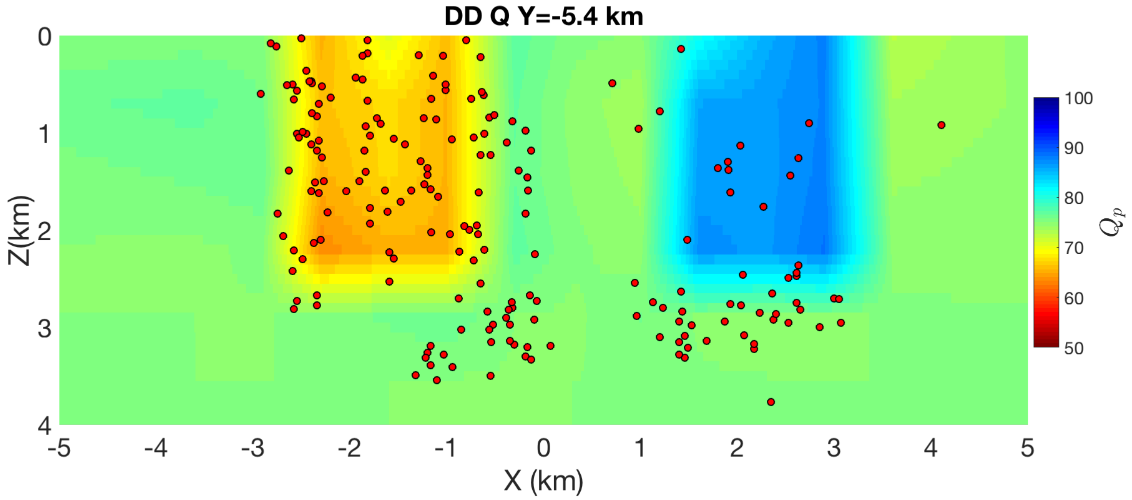
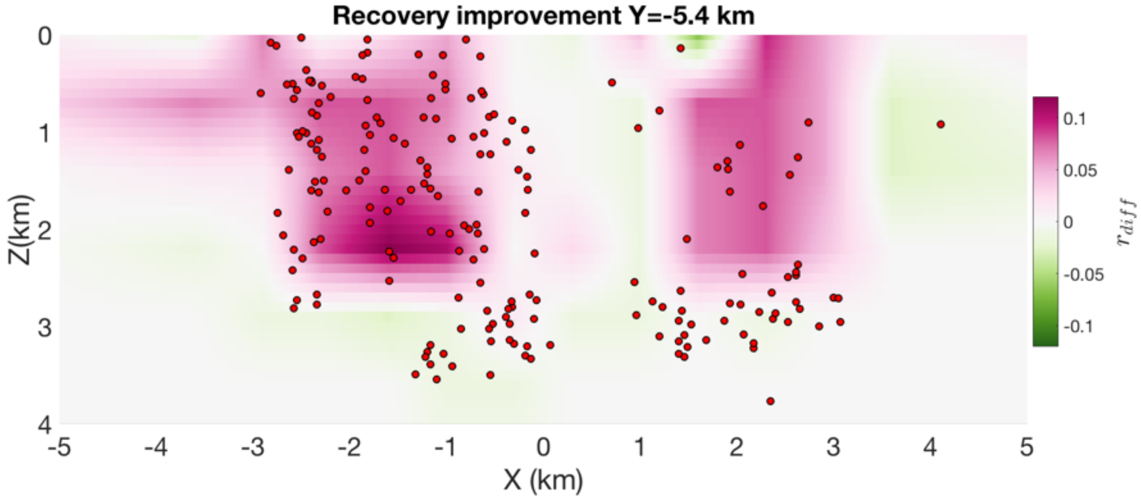
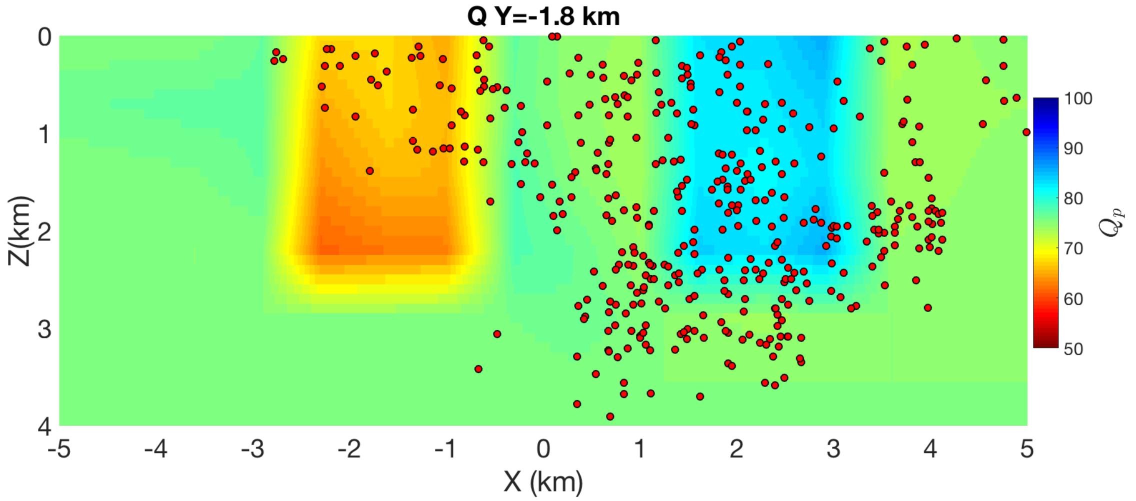


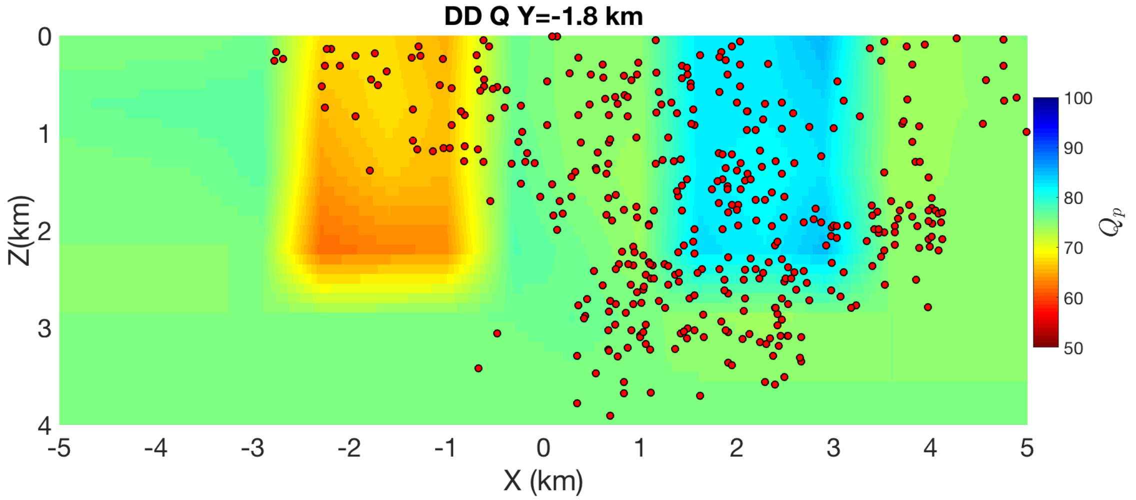
Figure 5.8 Q_P anomaly model cross-section plots. The first row is the Q_P anomaly with 20% amplitude. The second row is the Q_P anomaly with 10% amplitude. The third row is the Q_P anomaly with 30% amplitude. Black stars: nodes used in inversion.

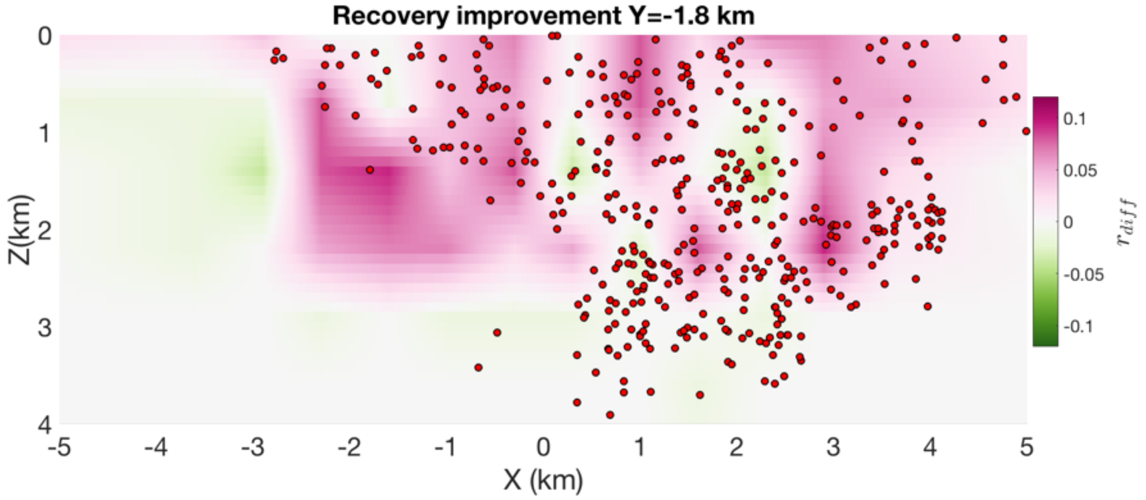

Q Y=-5.4 km
Z(km)
0
1
2
3
4
X (km)
-5
-4
-3
-2
-1
0
1
2
3
4
5
Q_p
100
90
80
70
60
50


DD Q Y=-5.4 km
Z(km)
0
1
2
3
4
X (km)
-5
-4
-3
-2
-1
0
1
2
3
4
5
Q_p
100
90
80
70
60
50


Recovery improvement Y=-5.4 km
Z(km)
0
1
2
3
4
X (km)
-5
-4
-3
-2
-1
0
1
2
3
4
5
r_{diff}
0.1
0.05
0
-0.05
-0.1


Q Y=-1.8 km


DD Q Y=-1.8 km


Recovery improvement Y=-1.8 km

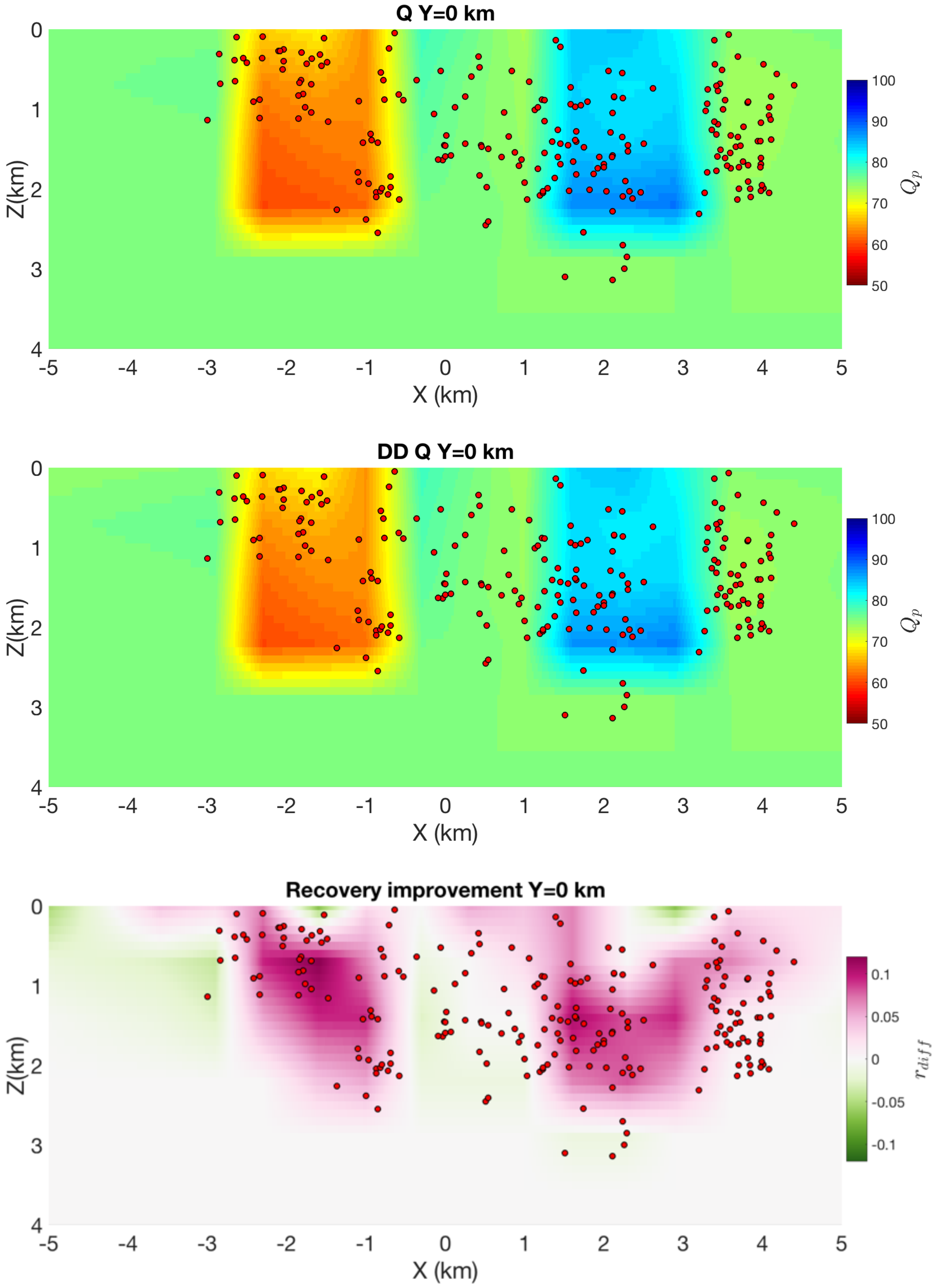
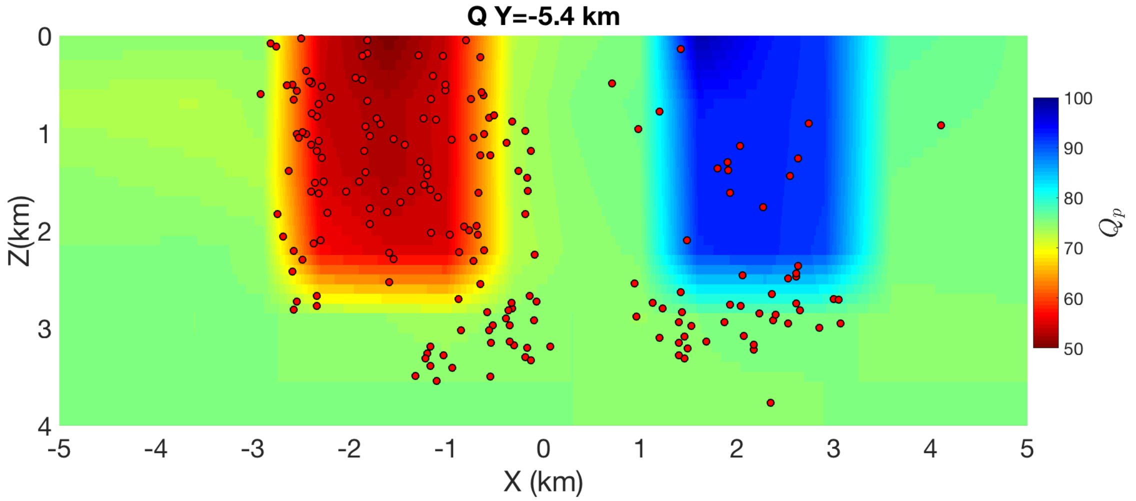
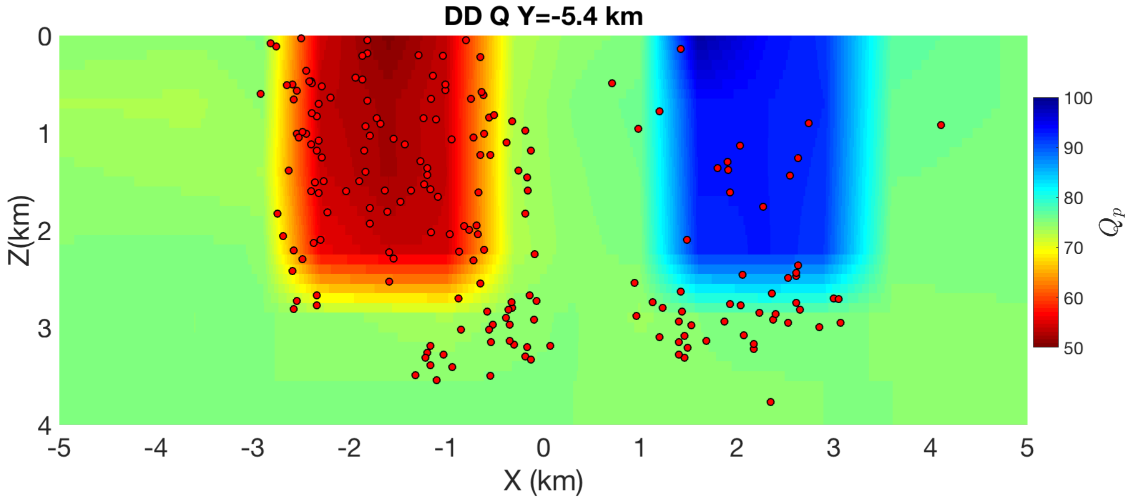
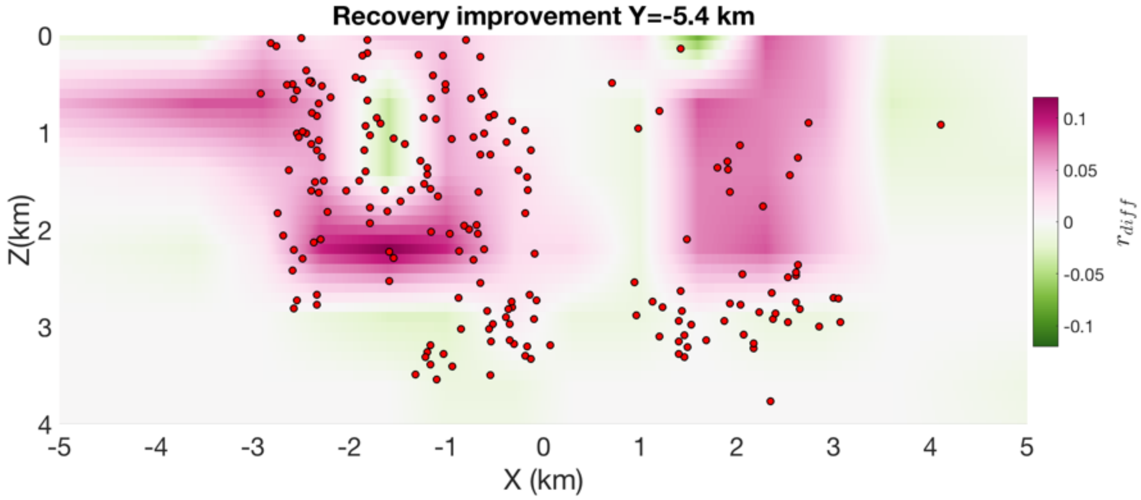


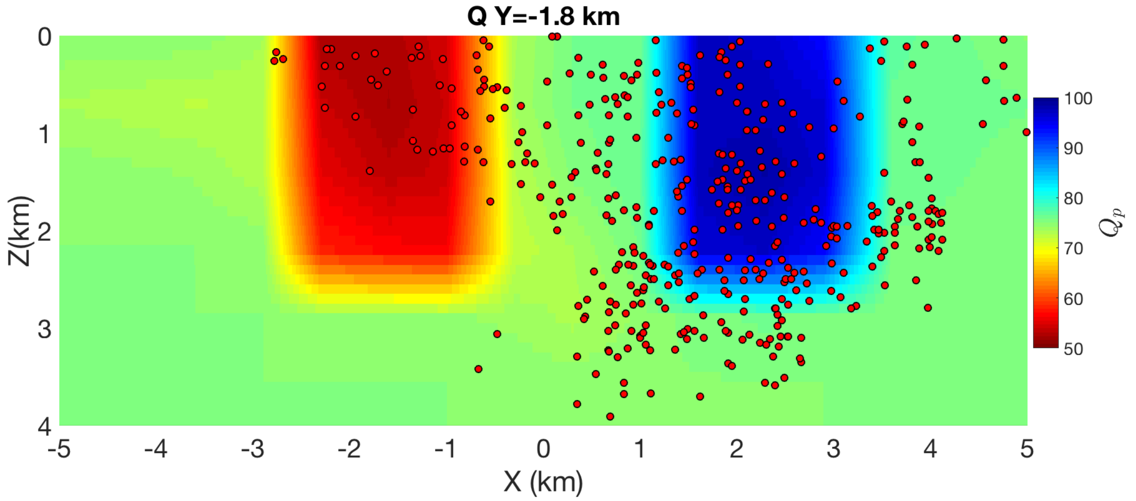
Figure 5.9 Cross-sections of results from 10% Q_P anomaly recovery test. For each three-plot group, the first row is the results from traditional Q tomography, the second row is the results from DD

Q tomography, and the last row shows the recovery improvement percentage of DD Q tomography over traditional Q tomography. Red dots: earthquakes.

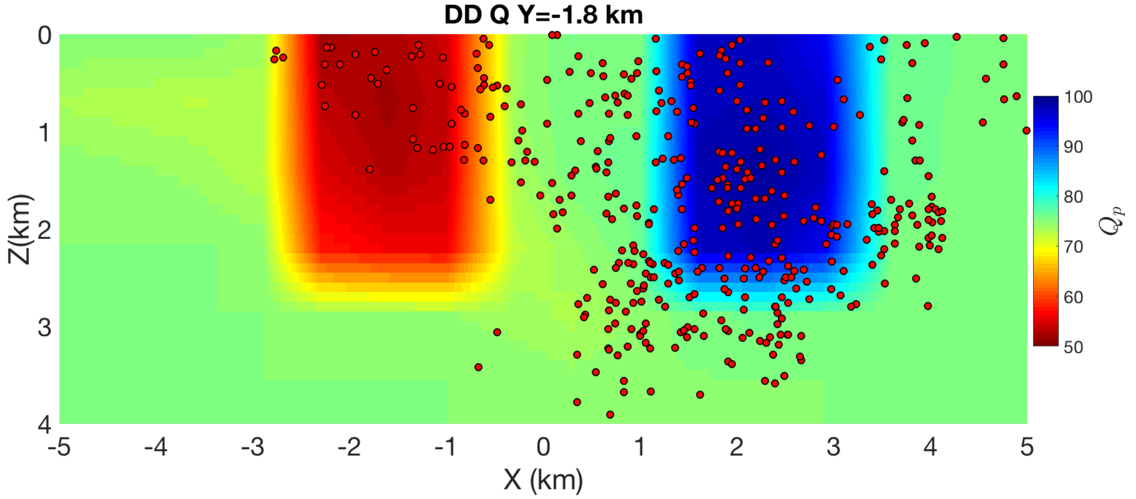

Q Y=-5.4 km


DD Q Y=-5.4 km

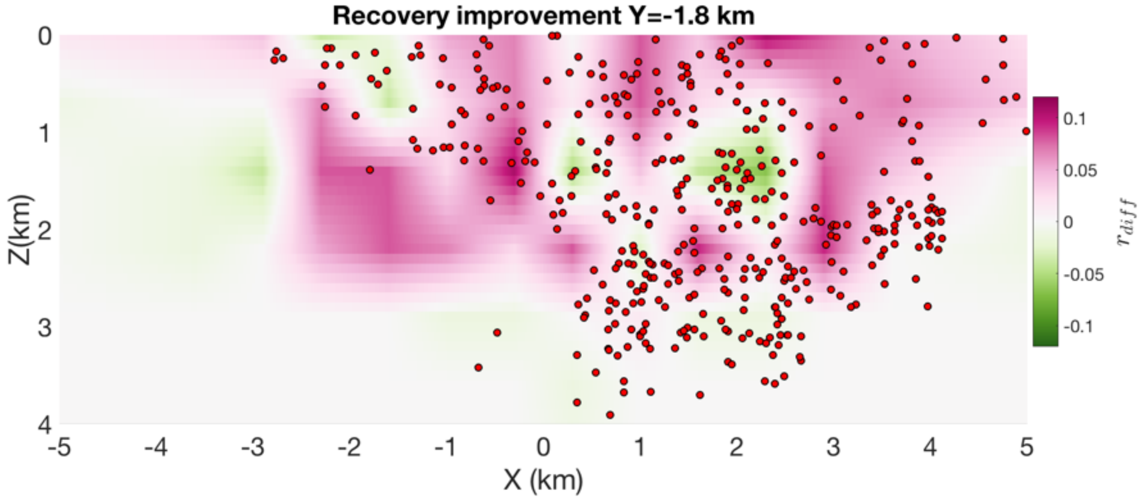

Recovery improvement Y=-5.4 km



Q Y=-1.8 km



DD Q Y=-1.8 km



Recovery improvement Y=-1.8 km

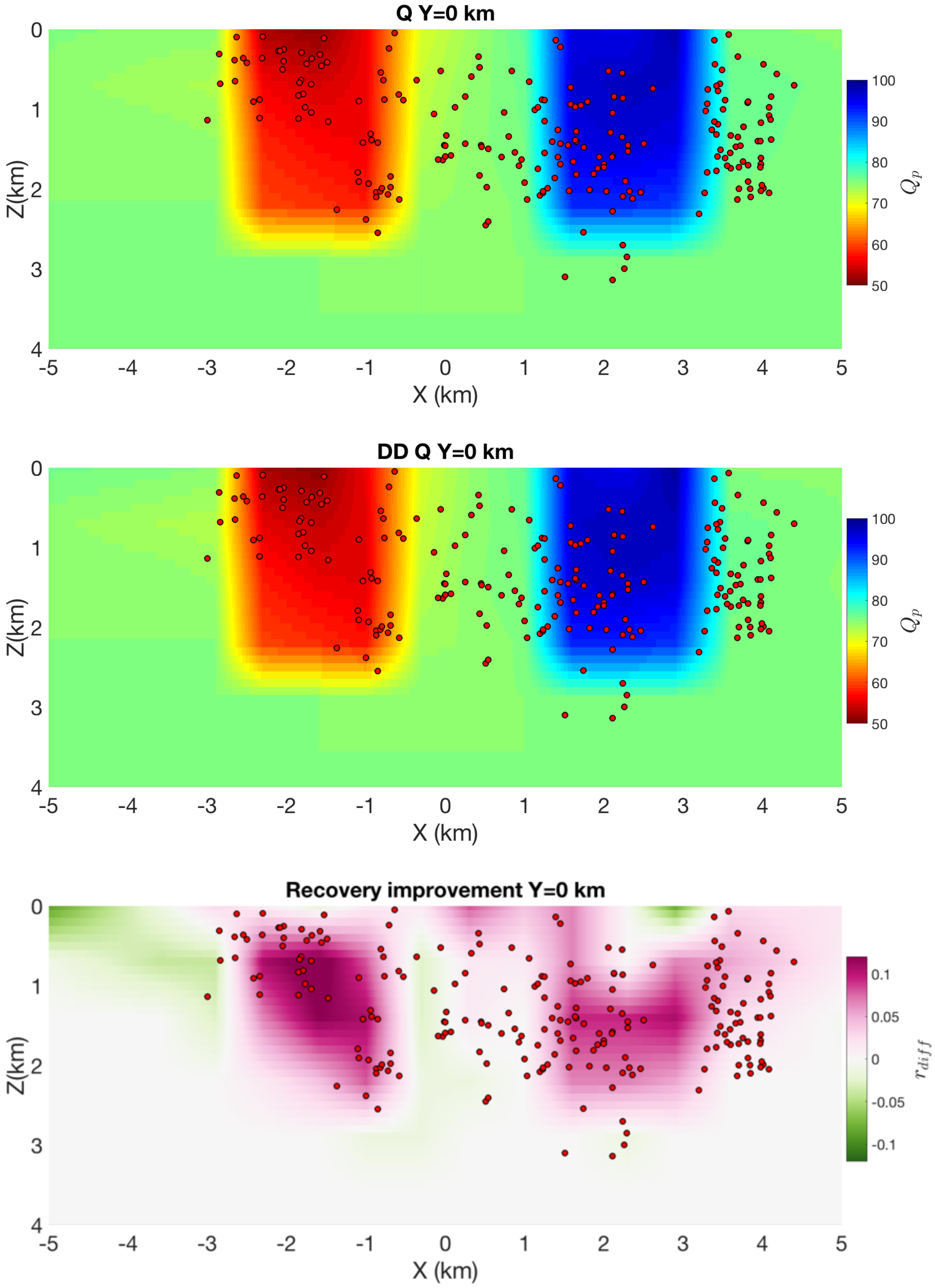


Figure 5.10 Cross-sections of results from the 30% Q_P anomaly recovery test. For each three-plot group, the first row is the results from traditional Q tomography, the second row is the results from

DD Q tomography, and the last row shows the recovery improvement percentage of DD Q tomography over traditional Q tomography. Red dots: earthquakes.