UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

COMPREHENSIVE ANALYSIS OF OKLAHOMA EARTHQUAKES: FROM EARTHQUAKE MONITORING TO 3D TOMOGRAPHY AND RELOCATION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

CHEN CHEN
Norman, Oklahoma
2016

COMPREHENSIVE ANALYSIS OF OKLAHOMA EARTHQUAKES: FROM

EARTHQUAKE MONITORING TO 3D TOMOGRAPHY AND RELOCATION

A DISSERTATION APPROVED FOR THE
CONOCOPHILLIPS SCHOOL OF GEOLOGY AND GEOPHYSICS

BY

_____
Dr. Kurt J. Marfurt, Chair

_____
Dr. G. Randy Keller

_____
Dr. Austin A. Holland

_____
Dr. Xiaowei Chen

_____
Dr. Xingru Wu

Acknowledgements

I would never have completed my Ph.D. degree without guidance and support of my committee members, help from my friends, and encouragement from my family.

I would like to express my sincere appreciation to Dr. G. Randy Keller, my research co-advisor, for all his time in guidance, mentoring and support throughout the past few years when he was in the School of Geology and Geophysics and retired. Thank you for bringing me in this research and supporting me when I pursued my Ph.D. degree. I will benefit from your knowledge and personality.

I also would like to thank Dr. Austin A. Holland for his guidance, support, and encouragement when he worked at the Oklahoma Geological Survey and moved to the U.S. Geological Survey. Thank you for your patience on me to learn geoscience skills from the scratch. Your great skills and scientific integrity will be wealth of great value for me.

I appreciate Dr. Kurt J. Marfurt to be my official committee chairman and give me valuable suggestions after Dr. G. Randy Keller retired. Your knowledge continuously provides me improvement direction on my research. I would like to thank Dr. Xiaowei Chen for specific technical suggestions and Dr. Xingru Wu for reviewing this manuscript.

I am deeply grateful to the Oklahoma Geological Survey for continuously providing me the financial support to my projects and for considering me to participate in this hot-topic research.

I would like to thank Dr. Kevin Crain, Stephen Holloway, and Jefferson Chang for providing and interpreting the data, as well as help on my search. I appreciate the

great time and helps from my friends Xiao Xu, Zonghu Liao, Ting Wang, Long Wu, Li Pan, Li Liu, Mu Liu, Tengfei Lin, Yifang Cheng and all my friends at University of Oklahoma and Colorado School of Mines.

Last but not the least, I want to express the deepest appreciation to my parents. Thank you for your support and understand. Without your help, I will never be able to complete my Ph.D. degree. My most special thanks are for my girl friend Yijia Zhang, for your love, sacrifice, patience, and understand, especially during the last year of my Ph.D. time. Without your encouragement and patience, I do not think I can achieve my goal.

Table of Contents

List of Tables

List of Figures

Chapter 1: PhasePApy: A Robust Pure Python Package for Automatic Identification of Seismic Phases

Chapter 2: Tomographic Study for Central Oklahoma by Integrating Gravity and Magnetic Data

Abstract

A better understanding of the basement, such as rock properties and faults orientation, is critical to study the earthquakes in central Oklahoma. The active seismicity in this region provides a great opportunity to study the velocity structures in the basement. The earthquake location is an important parameter for seismology such as hazard assessments, fault studies, induced seismicity, and others.

In order to advance our understanding of the increase in seismicity occurring within Oklahoma and the resultant seismic hazards, I improved our ability to locate the earthquakes in three-dimensional (3D) space and improved the geological and geophysical model for central and north-central Oklahoma. I developed a pure *Python* package for automatically identifying the seismic phase types for earthquake monitoring, which is used for further seismological analysis, such as seismic tomography, focal mechanism studies, and others. A major result is a new 3D velocity model for the central Oklahoma region. The 3D velocity model provides new insights on the deep geologic features to depth of ~10 km, and significantly improved the earthquake locations.

In order to process the massive amount of seismic data that the Oklahoma Geological Survey seismic network records, I developed a *Python* phase identification package: the *PhasePApy* for earthquake data processing and near real-time monitoring. All the data formats supported by *Obspy* (a *Python* open library for seismology) are supported within the *PhasePApy*. The *PhasePApy* has two sub-packages: the *PhasePicker* and the *Associator*, aiming to identify phase arrival onsets (picks) and associate them to phase types (P-wave and S-wave), respectively. The *PhasePicker* and

the *Associator* can work jointly or separately. Three autopickers are implemented in the *PhasePicker* sub-package: the frequency band picker, the Akaike Information Criteria function derivative picker, and the kurtosis picker. The *PhasePicker* triggers the pick with a dynamic threshold and can declare a pick with false pick filtering. Also, the *PhasePicker* identifies a pick polarity and uncertainty for further seismological analysis such as focal mechanism determination. There are two associators included in the *Associator* sub-package: the *one-dimensional* (*1D*) and *three-dimensional* (*3D*) *Associator*, which assign phase types to picks that best fit the predicted travel times of a potential earthquake by minimizing root mean square (RMS) residuals, respectively. Both associators use travel-time look up tables to determine the best estimation of the earthquake location and evaluate the phase type for picks. The *PhasePApy* package has been used extensively for local and regional earthquakes and can work for active source experiments as well.

Two 3D tomographic models for central Oklahoma are derived by employing different methods for central Oklahoma with P- and S-phase arrival times of local earthquakes. I created the first velocity model by using the FMTOMO package with 8194 M2+ earthquakes, the depths of which are greater than 3 km. I derived the second velocity model by using the SIMUL2000 package with 480 relative uniformly distributed composite events that include pick information from neighbor events within hypocentral distance less than 2 km. These two velocity models and Vp/Vs ratios show similar features on a large scale, but the model from SIMUL2000 has less velocity and Vp/Vs ratio variation. A strong lateral velocity heterogeneity within the Precambrian basement is present in both models, which indicates the presence of complex structures

in the upper crust of central Oklahoma. By comparing velocity models to gravity and magnetic data, most of the large-scale velocity anomalies generally correlate with the previous known features within crystalline basement in north central Oklahoma, such as mafic intrusion associate with the Osage County Microgranite and Spavinaw Granite, Central Oklahoma Granite Group, and southwestern extension of the Mid-Continent Rift (MCR) arm. A low Vp/Vs ratio zone is identified in the upper basement for depths of ~3 to 10 km, which may be due to the water-filled fractures.

I obtained high-accuracy earthquake locations for central Oklahoma from 01/01/2010 to 03/31/2016 employing the new 3D velocity model. The 3D velocity model algorithm improved the absolute locations of the earthquakes with the SIMUL2000. A 3D double-difference method (HYPODD 3D) followed to improve the relative locations for the clustered earthquakes. The relocated earthquakes show a hypocenter distribution centered at depth about 5 km, which demonstrate that most earthquakes occurred within the upper part of the crystalline basement. Numerous relocated clusters present clearer northeast-southwest (NE-SW) and northwest-southeast (NW-SE) orientations after the 3D model relocation, which are consistent with maximum horizontal stress state within the study region. Some of clusters suggest unmapped faults in the crystalline basement. I also relocated the cataloged earthquake locations with the 1D velocity double-difference method (HYPODD 1D) to compare to the results of the 3D velocity relocation. Both the 1D and the 3D velocity relocations improved the relative earthquake locations, and showed clear and narrow cluster orientations, but there were systematic shifts between the results. The southern

earthquakes mainly shifted westward, and the northern ones shifted southward. On average, the horizontal shift is about 0.7 km.

Introduction

Oklahoma has drawn a great deal of attention since 2009 due to the significant seismicity increase. Numerous earthquakes and seismometers in central Oklahoma can easily produce a large volume of data, which are almost impossible for analysts to identify all phase arrivals and earthquakes. I developed a robust and efficient software package *PhasePApy*, which is a very useful tool for automatically identifying a large number of earthquakes. Since 2009, the total number of Oklahoma earthquakes is much more than that of all events identified in the last few decades. Before 2009, the average number of earthquakes with magnitude greater than 3.0 was 1.6 per year. This number increased to 585 by 2014 and ~700 by 2015. The very active seismicity in central Oklahoma and the surrounding area provides a great opportunity to study the velocity structures in the crystalline basement of this region. A better understanding of the basement, such as optimal oriented faults and petrophysical properties of rock, is critical to study the increased seismicity in central Oklahoma. The earthquake location is an important parameter for seismology such as fault studies, induced seismicity, and others. The derived velocity model can improve the earthquake locations with high accuracy, which is useful to the following seismological research.

In this dissertation, there are three projects included. Three main chapters connect to each other. Chapter 1 focuses on the *Python* package for automatically identifying seismic phases. In Chapter 2, a tomographic study is discussed for central Oklahoma and the surrounding region by integrating gravity and magnetic data. A comparison between the modeling results, gravity, and magnetic data is conducted to make an interpretation in the upper basement. I correlate several significant previously

known geological features from gravity and/or magnetic data to the 3D velocity modeling results. In general, there are two uses for 3D velocity modeling via seismic tomographic approaches: interpret geological structures and improve source locations with the modeling results. Chapter 3 is the study about the earthquake relocation with the 3D velocity model that is derived from Chapter 2. The accurate earthquake location is critical for many subsequent applications such as fault studies, earthquake hazard assessment, seismic travel time tomography, stress field analysis, induced seismicity and others. The 3D velocity model improves the earthquake absolute and relative locations by considering the velocity bias; the improved accurate earthquake hypocenters help to identify the relationship between the earthquakes and nearby faults, injection wells, and other geological structures in the crystalline basement.

Chapter 1: PhasePApy: A Robust Pure Python Package for Automatic Identification of Seismic Phases

Introduction

In increasing numbers globally, modern seismic monitoring networks, with broadband and/or strong motion seismometers, can easily produce large enough volumes of waveform data such that manual picking of phase arrivals by analysts becomes nearly impossible. The phase identification for a large amount of data is tedious work for a human being. Moreover, to some extent, manual analysis of seismograms and phase picking is subjective, and may depend on different analysts who may introduce bias and inconsistencies in phase-arrival data (e.g., Leonard, 2000). Therefore, phase identification will always be a critical issue both for manual phase identification and automatic algorithmic phase identification. Inaccurate and inconsistent phase arrival times can bias all subsequent analyses such as earthquake location, travel time tomography, and focal mechanism analysis. In order to analyze manual picking errors, Zeiler and Velasco (2009) performed two experiments to define and isolate phase picking errors. They stated that the signal-to-noise ratio (SNR) is the main source of error for an individual analyst. Leonard (2000) conducted the comparison of 78 teleseismic phases manually picked by four analysts and three automatic picking algorithms to conclude that the average difference between the analysts is greater than the difference between analysts and an automatic picker.

There are many automatic phase arrival pickers based on different algorithms. Each kind of picker has its own advantages and limitations. Effectiveness and computational efficiency are two important measurements of an automatic algorithm.

Many factors can affect these two measures of an automatic picker, such as data quality and algorithm simplicity. The SNR can dramatically affect the behavior of automatic phase pickers (e.g., Zeiler and Velasco, 2009). The majority of picking algorithms can be classified into one of the following methods: energy transient methods in time-domain or frequency-domain (e.g., Withers *et al*., 1998; Vassallo *et al*., 2012), autoregressive (AR) methods (e.g., Leonard and Kennett, 1999), high-order statistical methods (e.g., Baillard *et al*., 2014), neural network methods (e.g., Gentili and Michelini, 2006), and wavelet transform methods (e.g., Bogiatzis and Ishii, 2015). The most popular picking algorithms compare the short-term average (STA) and long-term average (LTA) of the characteristic function (CF) of a signal to determine the phase arrivals; this is a classic energy transient method (e.g., Allen, 1978; Baer and Kradolfer, 1987; Lomax *et al*., 2012; Vassallo *et al*., 2012). These algorithms trigger and declare picks when the ratio of STA to LTA exceeds a predefined or dynamic threshold value. The STA/LTA method is an important automatic picking algorithm for early warning systems, due to its simplicity and effectiveness. Withers *et al*. (1998) discussed several STA/LTA algorithms including classic STA/LTA, delayed STA/LTA, recursive STA/LTA and Z-detector in time-domain and power spectrum density in frequency domain.

Another picking method is based on autoregressive Akaike Information Criteria (AR-AIC) function (e.g., Akaike, 1974; Kitagawa and Akaike, 1978; Leonard and Kennett, 1999; Sleeman and van Eck, 1999; Zhang *et al*., 2003). By autoregressive approach, the seismogram window is divided into two different local stationary segments, which contain the noise and signal, respectively. Two segments of different

statistical properties are modeled as an AR process. The AR-AIC picker can determine the optimal onset time by looking for the global minimum of the AIC function of a seismogram. For this reason, it is necessary to choose a window that only includes the segments of interest. Therefore, phase arrival identification is required to choose the appropriate window before using the AR-AIC method. The signal pre-detection processing will increase the computation time and reduce the efficiency.

Higher order statistics, such as kurtosis, is used to identify seismic phases due to the non-Gaussian distribution attribute of seismic waves (e.g., Panagiotakis *et al*., 2008; Baillard *et al*., 2014; Hibert *et al*., 2014). Artificial neural network methods are also used to search for phase onset by training the machine to recognize phase arrivals (e.g., Dai and MacBeth, 1995; Wang and Teng, 1997; Gentili and Michelini, 2006). In addition, wavelet transform is another approach for seismic signal analysis (e.g., Anant and Dowla, 1997; Akansu *et al*., 2010; Bogiatzis and Ishii, 2015). The two main types of wavelet analysis, discrete and continuous wavelet transform, can be used to detect both P- and S-phases.

In this paper, we introduce a software package to identify phase arrival times and associate the picks to the type of seismic wave. This software package is developed as a *Python* package: *PhasePApy*, in order to leverage the growing number of scientific libraries being written in *Python*, mostly notably *Obspy* (Beyreuther *et al*., 2010). The choice of making our algorithms compatible with *Obspy* means that all data formats and access methods supported by *Obspy* are naturally supported**;** these include SEED, MiniSEED, SAC, SEGY, and others. The other libraries on which *Obspy* and our algorithms rely are *Numpy* (Oliphant, 2007), *Scipy* (Jones *et al*., 2001), and *Matplotlib*

(Hunter, 2007). The *PhasePApy* is separated into two different tasks, which are implemented in two sub-packages: the *PhasePicker* and the *Associator*, aiming to pick phase arrivals and associate them to earthquakes, respectively. We implement three pickers based on different CFs: the frequency band picker (*FBpicker*), the AIC function derivative picker (*AICDpicker*), and the kurtosis picker (*KTpicker*). We also implement two associators: the *1D Associator* and *3D Associator*. Users can choose the appropriate picker and associator that best meet their requirements.

The task of the *PhasePicker* is simply making picks of phase arrivals from seismograms. Each picker only has a small number of different parameters to input. The processing method *FBpicker* adapts to sampling rate and instrument gain systematically. The *FBpicker* algorithm is an algorithm modified from that of Lomax *et al*. (2012), which does not require extensive tuning (Vassallo *et al*., 2012) as there is no need to specify triggering thresholds and other parameters for each analyzed channel. The *AICDpicker* and *KTpicker* determine the phase arrival times based on the derivative of the AIC function and kurtosis, respectively. Many picking algorithms pick and declare phase types by using specified thresholds, but the PhasePicker employs the dynamic threshold. The *PhasePicker* only detects phase arrivals and identifies the phase onsets, but not phase types. Phase polarities and picking uncertainties can be determined as well by the *PhasePicker*. The picker itself can be set to make picks from global phase arrivals to near-field phase arrivals and reflections.

Because the *PhasePicker* can pick any number of phase types that have sufficient energy contrast in the filtered waveforms, it is necessary to have a robust approach to automatically associate the picks to seismic phase types. The task of the

Associator is to associate the picks that best fit the different phase types associated with a particular earthquake. We developed two earthquake phase associators (fixed depth 1D velocity model associator and 3D velocity model associator) appropriate for local and regional earthquake monitoring, but these methods could be adapted to accommodate any distance. The *Associator* can provide estimation of event location and origin time for the determined earthquake events. The primary function is not to locate the earthquake but properly identify the phases associated with the earthquake.

Phasepicker Algorithms

Characteristic Function 1: FBpicker

We implement three pickers for pick identification; the first is the *FBpicker*, which is a modified transient energy method from Lomax *et al.* (2012). Before any processing, the *FBpicker* removes the mean and trend of the data with a least squares method. It is important to note that if the user is attempting to make picks for long-period phases, a sufficiently long window of data is required such that this initial processing does not affect the ability to identify phases. The *FBpicker* applies an octave-filter (doubling the central frequency) to the seismograms and generates several frequency bands for each seismogram, which automatically adapt to instrumentation sampling rate. The central frequency of each consecutive band keeps doubling from a user-defined minimum band, which includes low-frequency components of interest, until the high corner of the next band of the last filtering band exceeds the Nyquist frequency (see Fig. S1, available in the appendix). The user can reduce the number of bands by decimating the data, which decreases the Nyquist frequency. The *FBpicker* allows users to change the corner order of the filtering, which determines the

attenuation for frequencies outside of the bandwidth. In order to reduce the Gibbs effect from the band-pass filtering, the filtered data of each band are cosine tapered and also user configurable i.e., users can modify the tapering percentage of the data.

Following the method of Lomax *et al.* (2012), the *FBpicker* calculates the energy E_n of the filtered data for each band:

$$E_n[i] = BF_n[i]^2, \tag{1}$$

where $BF_n[i]$ is the band-pass filtered amplitude of *ith* sample of the *nth* band. We implemented RMS, and standard deviation (SD) mode to calculate the CF for each band in Equations (2) and (3), respectively:

$$CF_n^{RMS}[i] = \frac{E_n[i]}{RMS(E_n[i-1-l:i-1])}, \tag{2}$$

$$CF_n^{SD}[i] = \frac{E_n[i]-mean(E_n[i-1-l:i-1])}{SD(E_n[i-1-l:i-1])}, \tag{3}$$

where *l* is the window length in a sample. For RMS mode, $CF_n^{RMS}[i]$ is the ratio of the transient energy of sample *i* of frequency band *n* to the RMS value of energy in the previous moving window; for SD mode, $CF_n^{SD}[i]$ is the ratio of difference between the transient energy of sample *i* and the mean of its previous window for frequency band *n* to the standard deviation of the energy in the previous window. The CF quantifies the energy change relative to the energy level in the previous window. Therefore, the high $CF_n[i]$ indicates the high energy level of a sample compared to that in the previous window. Due to fewer numerical operations, the RMS mode can improve the computation time compared to the SD mode without significantly affecting the performance. The reason is that the trends of the data are removed in the data-conditioning phase. RMS mode reduces the computation time by about 30%, when computed on a 6-minute segment of data. The *FBpicker* allows the user to customize

the CF calculating mode (RMS or SD) depending on their requirements. Finally, the *FBpicker* summarizes the $CF_n[i]$ by taking the maximum value over all bands *n* for each sample *i* (Lomax *et al*., 2012) to obtain the CF.

Characteristic Function 2: AICDpicker

The AIC picker has been used in many P-wave picking algorithms for single or multiple component records by searching the global minimum of the AIC function calculated with the autoregressive technique (e.g., Leonard and Kennett, 1999; Sleeman and van Eck, 1999; Zhang *et al*., 2003). The AIC function of a seismogram of two signal segments can be represented as the function of division point *P* (Sleeman and van Eck, 1999):

$$AIC(P) = (P - M)\log\left(\sigma_{1,max}^2\right) + (N - M - P)\log\left(\sigma_{2,max}^2\right) + Const., \quad (4)$$

where P is the division point, M is the order of the autoregressive model, N is the total length of the data, and $\sigma_{1,max}^2$ and $\sigma_{2,max}^2$ indicate the variance of the seismogram in two segments. Another approach without using the autoregressive method determines the AIC function directly from the seismogram (*Maeda*, 1985):

$$AIC(P) = P\log\{var(x[1,P])\} + (N - P - 1)\log\{var(x[P+1,N])\}, \quad (5)$$

where division point P ranges over all the samples of the seismogram *x*. If the seismogram only contains the P-wave, the AIC picker can easily determine the phase onset by looking for the global minimum of AIC function. However, if the seismogram also contains the S-wave, it is difficult to identify the seismic phases by just searching for the global minimum of the AIC function. We take the absolute value of the first derivative AIC function as the CF for *AICDpicker*, because the first derivative of the AIC function is sensitive to the change of the AIC function.

Characteristic Function 3: KTpicker

The third picker in the *PhasePicker* is a kurtosis-based picker: the *KTpicker*. The kurtosis is a statistical measure to describe the degree of measurements concentration, which can be used as the CF of a seismogram. The kurtosis is high-order statistics of a variable, and has been widely used in identifying the seismic waves phase onset (e.g., Panagiotakis *et al*., 2008; Baillard *et al*., 2014; Hibert *et al*., 2014). We apply the kurtosis in a moving window through the time series to characterize the seismic signal. Kurtosis is defined as the standardized fourth moment about the mean of the measurements:

$$K = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2} = \frac{\mu_4}{\sigma^4}. \quad (6)$$

In equation (6), E is the expectation operator, X is the measurements, μ is the mean of the measurements, μ_4 is the fourth moment about the mean, and σ is the standard deviation. White noise is normally (Gaussian) distributed, the kurtosis of which is 3 (DeCarlo, 1997). Values of K greater than 3 result in a higher peak compared to the Gaussian distribution, while values less than 3 result in lower peak than the Gaussian distribution. Therefore, the kurtosis characterizes the signal based on the shape of the distribution instead of the SNR, and as such, the kurtosis can work well for the signals with low SNR.

Triggering, Declaration and False Pick Filtering

The *PhasePicker* identifies, or triggers, the pick by using a floating threshold level, which is determined by multiplying the RMS of the CF in a moving window with a user-defined coefficient, instead of a fixed threshold. The floating threshold level may make the pick triggering independent of the noise level of CF to some extent. This

dynamic threshold level is a method often applied in statistical process control called control charts (e.g., Alwan and Roberts, 1988) (e.g., 6-sigma). The threshold is determined by multiplying the RMS of the CF to a user-defined coefficient, which is roughly comparable to the number of standard deviations of the CF. The *Associator* allows the user to change threshold as a multiple of σ based on the requirements of each case. The *PhasePicker* triggers picks at the time when the CF exceeds the dynamic threshold. We design the *PhasePicker* with a rollback algorithm to ensure that picks are declared at the same time as much as possible for different threshold levels. The *PhasePicker* will roll backward the triggering time along the CF until it reaches the first local minimum, at which point a pick is declared. Without the rollback algorithm, there will be a few samples difference between higher and lower threshold levels. Therefore, pick declaration is roughly independent of the triggering threshold level. The SNR of CF (SNRCF) of each pick is determined by the ratio of the first local maximum CF after the pick to the RMS of CF in the previous window. This feature can assist users in choosing the appropriate threshold level for the pick-triggering process.

The SNR of a signal on different stations or time can vary over a wide range. Usually, the stations close to an earthquake's epicenter have higher SNR than the distant ones. As would be expected, it is easier to trigger picks for channels with high SNR rather than the ones with low SNR. A lower dynamic threshold level can identify a greater number of smaller amplitude signals (indicated by low CF spikes), but false picks are more likely to be identified. A high threshold-triggering level can trigger high SNR signals and introduce fewer false picks, but may miss weak signals. Therefore, there is a tradeoff between the threshold levels for triggering the picks. By tuning the

threshold level, one can remove a significant number of false picks. This is different than the method of Lomax *et al*. (2012), where static thresholds are applied and do not change without reconfiguration and as such cannot account for changing noise levels due to things such as wind or cultural noise. From our experience, it is often possible to find a single dynamic threshold coefficient that is appropriate for a broad range of channels within a region.

A high dynamic threshold level can identify picks that correspond to phase with high SNR. However, this approach often discards significant picks for phase arrivals with lower SNR. If users choose a relatively low dynamic threshold level, we implement two optional pick-filtering algorithms that attempts to preserve as many real picks as possible and remove questionable picks. The first utilizes a close-pick cleaning filter, which does not allow a following consecutive pick to occur within a user-defined amount of time. The second filter implements a short-period noise-cleaning filter, which determines whether a pick may be false or not by comparing the σ in the previous and following window of a pick. The energy levels before and after phase onset are different, which can be used to determine whether the detected signal is a false or correct earthquake phase. The short-period noise-cleaning filter removes the pick if the σ of the waveform amplitude in a short preceding window multiplied by a user-defined coefficient is greater than the σ of the waveform amplitude in the window after the pick. Otherwise, the pick is considered a correct phase pick. In the case that there is another pick falling in the previous or following window, the window size will automatically shrink to the interval between two picks. Although it is possible that some of the

remaining picks are still false, after both false pick-cleaning filters are applied, a significant number of false picks can be filtered out.

Polarity and Uncertainty

The *PhasePicker* includes a method to determine the phase polarity or direction of first motion as compression, dilatation, or uncertain. Human beings determine the phase polarity by visually inspecting the difference between the amplitude of the phase and background noise level. If an analyst considers the amplitude large enough, he or she can assign the direction of first motion. If the amplitude is approximately equivalent to the noise level, it is difficult to identify the first arrival and the polarity of the phase. The *PhasePicker* determines the polarity in a similar way by comparing the phase amplitude to the noise level. When the pick is declared, the *PhasePicker* searches for a local maximum or minimum value along the waveform after the triggering time. The amplitude of the pick is determined as the difference between the average in the pick epoch and the identified extreme value. If the absolute value of the pick amplitude is greater than σ of the seismic waveform in a previous window times a user-defined coefficient, the polarity will be declared.

The uncertainty of a pick is not easy to determine because it is not easy to distinguish how much of the true seismic phase is buried in the noise. The SNR is a significant factor that affects the accuracy of the uncertainty determination. We employ the RMS of the CF before the declared picks and a user-defined coefficient to evaluate the picking uncertainty. The uncertainty-triggered picking time $t_{triggered}^{uncert}$ is obtained by rolling forward the declared time $t_{declared}$ along the CF until the value of the CF exceeds the RMS times the user-defined coefficient. The *PhasePicker* determines the

uncertainty as the time difference between declared pick $t_{declared}$ and uncertainty-triggered pick $t_{triggered}^{uncert}$.

Results of PhasePicker

We take a 6-minute window of data as an example to show the implementation and performance. The example seismogram is from a magnitude 2.7 earthquake that occurred in central Oklahoma on June 16, 2013, which is recorded on the vertical component of the OKCFA station in the Oklahoma regional network (OK). In this example, we use the *FBpicker* to demonstrate seismic phase arrival picking performance. The *AICDpicker* and *KTpicker* have similar picking performance. The sampling rate of the seismogram is 100 Hz. By using 1 Hz as the central frequency of the starting band, the *FBpicker* determines six bands with central frequency at 1 Hz, 2 Hz, 4 Hz, 8 Hz, 16 Hz, and 32 Hz respectively (Fig. 1.1). In this example, we use the RMS and a 5-second moving window to calculate the CF for each band (Fig. 1.2). The strongest signal frequency component of this example is in the highest frequency band (32 Hz), which suggests that the seismogram corresponds to a local earthquake with higher frequency components.

Figure 1.3 demonstrates the behavior of the CF, which is normalized by the maximum absolute value, and the dynamic threshold with 20 seconds of the moving window and 6σ of dynamic threshold. We assign a value of zero for the very beginning windows (5 seconds and 20 seconds, respectively) of the normalized CF and dynamic threshold. In this example, the threshold of 6σ can identify not only the real phase arrivals, but also introduce some false triggering picks (Fig. 1.4a). By tuning up the dynamic threshold level, such as 8 or even higher, most false picks can be removed. In

order to demonstrate the removal of false picks from each component by applying the two cleaning filters, we use the window length of 0.78 seconds for the close-pick cleaning filter to remove close-picks, but there are still considerable false picks remaining (Fig. 1.4b). Then, we use the window length of 2 seconds and the coefficient of 2 as the short-period noise-cleaning filter setup. After applying the short-period noise-cleaning filter, most of the false picks are filtered out (Fig. 1.4c). The last picks on component N and Z in Figure 1.4c are false picks, which have not been removed. They could possibly be from some micro earthquakes or high-frequency cultural noise.

The *FBpicker* determines the polarity for declared picks (see Fig. S1.2, available in the appendix). We also show the normalized CF of the *AICDpicker* and *KTpicker* in Figures S1.3 and S1.4 (available in the appendix), but we only use picks from the *FBpicker* in the following processing demonstration. In these test data from vertical component of OKCFA, the picks from different automated picks are almost identical, with only a few samples difference. In our experience, the SNRCF of the declared picks is particularly useful information for choosing the appropriate coefficient for the dynamic threshold level, which can significantly affect the picking performance. By evaluating the SNRCF values of picks, the user will have a clue as to what threshold level can trigger the picks for actual phase arrivals, and avoid most false picks. The *Associator* can remove the remaining false picks by associating correct picks to the best-fitting earthquakes.

Associator Algorithm

The *Associator* uses identified picks and associates the picks with earthquakes using the following steps: pick aggregation, event candidate creation, origin time

analysis, and phase association. There are two associators in the sub-package: the *1D Associator* and the *3D Associator*. Both associators are using travel-time look up tables for P- and S-waves to determine earthquake location and to associate phases in this case for locally to regionally recorded earthquakes. The picks to associate can come from the *PhasePicker*, other picking algorithms, manual picks, or other sources.

Step 1: The *Associator* aggregates picks from different components on the same station if the seismometer has multiple components. Before aggregating picks, the *Associator* first stacks picks from different components on the same station (see Fig. S1.5, available in the appendix). Ideally, P- and S-waves are only identified on vertical and horizontal components, respectively. However, in some cases, P- and S-waves are picked on the other components, respectively. Pick stacking without differentiating components can make use of as many identified picks as possible for the further derivative analysis, such as tomography studies. The additional picks may allow for future advancements of phase association or further tuning of the picking algorithms. Automatic picks may occur on multiple channels at a seismic station and may have slightly different times. It is necessary to group the picks from different components of the same station. If the picks from different components have a time separation less than a user-defined pick link length, the picks are linked to create what we call modified picks. The *Associator* uses the minimum S-P time of the travel-time table multiplying a user-defined coefficient to determine the link length. The picks can be aggregated with either mean or median statistics mode, which takes the average value and median pick of the linked picks, respectively. Therefore, the aggregated pick is referred to as the modified pick. All the raw picks are associated with the modified picks for further

processing. Aggregating picks can reduce the number of picks necessary for the phase association and improves the computing efficiency.

Step 2: The *Associator* creates event candidates for every modified pick pair from the same station. The candidate pick pair separation (S-P interval in time) can determine S-P interval in distance (a measure of distance from the earthquake epicenter to a seismometer, S-P interval in distance refers to the epicentral distance for all the following cases in the *1D Associator*) by looking up the travel-time table. With the determined S-P interval in distance, the *Associator* calculates the origin time for the candidate by back projecting the P-wave travel time for the determined S-P distance. Each earthquake only has a single origin time; there must be some event candidates from different stations having similar origin times within some amount of uncertainty. Well-recorded or large earthquakes will naturally have a larger number of event candidates with origin times close to the actual earthquake origin time. The *Associator* is able to only process candidates with the S-P interval less than a certain user-configurable cut-off distance to improve processing efficiency and force the regional restriction we chose.

Step 3: The *Associator* analyzes the origin time clusters after all the event candidates are created. With a user-defined moving window through time, the *Associator* counts the number of unassociated event candidates with origin times falling within the moving window range. The origin time cluster analysis may roughly estimate when an earthquake occurred if the number of origin times in the cluster is greater than a predefined value. The minimum predefined value is 3, because triangulation for earthquake location requires at least 3 stations.

Step 4: The *Associator* determines the location for event candidate cluster by the count from high to low, and associates seismic phases to the unassociated picks. The *Associator* selects the first target cluster from all the candidate clusters that has the greatest number of event candidates. Once the event candidates in the first target cluster are used to associate to an earthquake, they will not be used in the cluster analysis for the second target cluster. Then the *Associator* re-runs the cluster analysis to determine the second target cluster among all the remaining clusters that has the greatest number of event candidates. An accurate earthquake location enhances pick association, although locating the earthquake is not the final objective of the *Associator*. The objective of the earthquake location algorithm in the *Associator* is to determine which pick pairs can fit the earthquake location best, and allow for picks not associated into pairs to possibly be identified as a phase arrival. The basis of the location algorithm in the *Associator* is to find a location with minimized RMS of misfits between S-P interval (in distance or epicentral distance for *1D Associator* and in time for *3D Associator*) and potential selected event candidates. Since any automatic phase picking algorithm may introduce false picks, a false pick close to a correct pick can introduce a false candidate with its origin time close to the one of the real event. In our package, one can reduce the false picks by tuning the *PhasePicker*, but meanwhile it is possible to miss the weak real seismic phase picks. If applying low dynamic threshold level, one may introduce false picks, which cause the false candidates. Due to the false candidates, two or even more event candidates from the same station may occur in an origin time cluster. The event candidates with close origin times can be viewed as concentric circles in a 2D perspective (Fig. 1.5a). The *Associator* divides the cluster with concentric circles into

several sub-clusters with event candidates from unique stations (no event candidates from the common station), calculates the RMS residuals for all the subsets, and determines the subset of candidates that minimizes RMS residuals as the best solution. In this way, the best fitting location can be determined, and false candidate(s) can be identified. The *Associator* includes an outlier evaluation function to determine whether an event candidate is an outlier by evaluating the residual. If the residual of an event candidate is greater than a user-defined parameter, the *Associator* will remove the candidate from the cluster as an outlier and repeat the location determination process until no outlier appears. Finally, in order to control the processing quality, we design a user-configurable RMS threshold and set the minimum number of observations to declare an earthquake. An earthquake is declared only if the RMS of misfits is less than a threshold, and the number of observations is greater than the minimum observation number.

Step 5: After all the earthquakes are identified, and all P- and S-phase arrival pairs are associated, it is still possible that there are individual modified picks that have not been associated with any earthquakes. Modified picks that do not have a P- and S-phase arrival pair may also represent phase arrivals associated with the associated earthquakes. These un-associated modified picks are then evaluated to determine whether the picks may be a single-phase arrival for an earthquake. The travel time of an individual single phase is calculated for phase arrivals in the travel time look up table based on the distance between the estimate of earthquake location and the seismometer. It is then possible to determine if the modified pick corresponds within uncertainty to the origin time of an identified earthquake. If an individual modified pick can be

associated to an earthquake then all its aggregated raw picks are associated to the earthquake with the appropriate phase.

Both the *1D Associator* and *3D Associator* take advantage of the fact that the *PhasePicker* does well at picking both P- and S-phase arrivals for local and regional earthquakes. For the *1D Associator*, the location algorithm determines the hypocenter of an earthquake by searching the best estimation that minimizes the RMS residuals between the S-P intervals in distance r_i and the distance from the trial epicenter to stations d_i:

$$R = \sqrt{\frac{\sum_{i=1}^{N}(r_i - d_i)^2}{N}}, \tag{7}$$

where N is the number of stations. A location determination example with four event candidates is shown in Figure 1.5b. With an initial trial epicenter, the *1D Associator* iterates the location search until the best estimation of hypocenter is determined. The final location will be close to the intersection of the circles (black star in Fig. 1.5b). The location uncertainty of the *1D Associator* is determined as the RMS residuals of distance:

$$U_{loc} = \sqrt{\frac{\sum_{i=1}^{N}(r_i - D_i)^2}{N}}, \tag{8}$$

where D_i is the distance from each station *i* to the final determined earthquake location.

For the *3D Associator*, the hypocenter is determined by using a grid-search algorithm, minimizing RMS residuals between the predicted travel-times t_i and the S-P intervals in time t_i^{s-p}:

$$R = \sqrt{\frac{\sum_{i=1}^{N}(t_i^{s-p} - t_i)^2}{N}}, \tag{9}$$

where N is the number of stations. The *3D Associator* uses a pyramid grid-search algorithm. The pyramid search is a search scheme breaking the searching grids into several sub-blocks (at most 8 sub-blocks in one search iteration) and finding the one with the least RMS of travel-time misfits. The identified sub-block that minimizes the RMS is then broken up and grid searched for the node that minimizes the RMS residuals until the sub-block cannot be split (less than or equal to three grids in each dimension). Finally, we employ a fine search to the node from the last iteration of pyramid search, in order to identify the node with the global minimum RMS residuals. The fine search compares the RMS residuals of all the surrounding nodes to the central node. If the node of minimum RMS residuals is in surrounding nodes, the fine search process is re-centered with the new node, and continues comparing the RMS residuals of the central node to the surrounding nodes until the node of minimum RMS residuals is in the center. The location uncertainty in the case of the *3D Associator* is determined as the RMS residuals of travel time:

$$U_{loc} = \sqrt{\frac{\sum_{i=1}^{N}(t_i^{s-p} - T_i)^2}{N}}, \tag{10}$$

where T_i is the predicted travel time from each station i to the final determined earthquake location. The grid size used for the 3D velocity model can affect the uncertainty since dense girds will improve the resolution of location search, but dramatically increase search times required to associate phases to an earthquake.

The location determination algorithms for the *1D* and the *3D Associator* are based on minimizing RMS misfits of S-P interval in distance (hypocentral distance) and S-P time, respectively. The hypocentral distance relates the S-P time with the known

velocity model. Therefore, the *1D* and the *3D Associator* determine the earthquake location with the same method, but by minimizing different RMS misfits.

Results of Associator

We choose minimum S-P time (0.78 seconds) to aggregate the picks for the 1day data test. The median mode is chosen in the picks aggregation for our example, so that outliers will not bias our results. There are 6023 picks identified, and 4832 modified picks remain after pick aggregation, which indicates that 1191 (~20%) picks are aggregated. Based on the dimensions of our study area we use 350 km as the cut-off distance to create an event candidate. A window length of 7 seconds is selected to process 1-day of data on June 16, 2013. A statistical result of the origin times falling in the moving window is shown in Figure 1.6a, including candidates from the common station (concentric circles in Fig. 1.5a). Figure 6b demonstrates the cluster analysis with concentric circles excluded, and clusters with counts greater than three probably indicate earthquakes. One of the events, which occurred at 15:38:50 on June 16, 2013, is located as an example shown in Figure 1.7. The star and its surrounding circle (very small: ~0.03°) indicate the epicenter and the location uncertainty of the earthquake. The example in Figure 1.7a indicates that the earthquake is a well-recorded event that occurred in central Oklahoma. The *1D Associator* determined S-P intervals could match the modeled travel-time curve well in the subplot to the left. The cross section plot (Fig. 1.7b) shows the waveforms and associated phases from all the stations in Figure 1.7a. The short bars are the associated picks, and the gray dots indicate where the bars cross the waveforms, which match the modeled P and S curves quite well. The *1D Associator* assigns picks matching the upper and lower curve as P- and S-phase respectively.

The same earthquake is processed with the *3D Associator* shown in Figure 1.8. We use 30, 40, and 4 grids in x, y, and z directions, respectively. Grid spacing is 0.1 degree for both latitude and longitude, and 3 km for depth. In the initial pyramid searching stage, grids in each direction are split into two segments, thereby having 8 sub-blocks for the first search. The *3D Associator* determines the deepest central node of the southeastern sub-block as the best solution for the first search (hexagon in Fig. 1.8a and the node depth in Fig. 1.8b). The pyramid search continues breaking up the identified sub-block and searching the central node with the minimum RMS residuals for each iteration (pentagon, square, and inverted triangle in Fig. 1.8). The depths of the determined node in the pyramid search are indicated in Figure 1.8b to the right with corresponding symbols. When the pyramid search finished, the fine search keeps comparing the RMS residuals of all surrounding nodes to the central node for several iterations to find the node with global minimum RMS residuals (circles in Fig. 1.8a and 1.8b).

Conclusions

We designed a *Python* automatic phase identification package: *PhasePApy*. The *PhasePApy* consists of two sub-packages: the *PhasePicker* and the *Associator*, which can identify picks and determine phase types, respectively. The *PhasePicker* is designed to automatically detect the seismic wave phase arrival times. Three different pickers: the *FBpicker*, *AICDpicker*, and *KTpicker*, are implemented in *PhasePicker*. The CFs for these three pickers are based on transient energy ratio, derivative of AIC function, and kurtosis, respectively. We employ the dynamic threshold method based on statistical process control (e.g., Alwan and Roberts, 1988) in the *PhasePicker* for pick triggering

to automatically adapt to instrumentation differences. The threshold level is determined based on both the noise level of the CF and a level control coefficient, which simplifies picker tuning over the *FilterPicker* (Lomax *et al*. 2012; Vassallo *et al*., 2012). In order to ensure the floating triggering threshold does not affect the declared picks, we implement a rollback algorithm to roll backward the triggering point along the CF to search for the local minimum to declare a pick. The *PhasePicker* has functions to determine the pick's polarity and uncertainty, which are designed for other processing purposes, such as focal mechanism analysis. The *Associator* identifies the phase types by searching for the picks that can best fit the determined earthquakes. Both *1D* and *3D Associator* aggregate the picks from different components on the stations to modify picks in order to improve computation efficiency and make use of as many picks as possible. By creating and back projecting the event candidates, the *Associator* conducts cluster analysis to search for the potential origin time cluster, in which event candidates have close origin times. The *Associator* evaluates RMS residuals to determine which picks can associate to the earthquakes.

There are only a few parameters to tune in the *PhasePApy* package. All parameters of the picker and associator are intuitive and easy to set. Many phase-picking algorithms can only identify P-wave phase arrival times; one advantage of our seismic phase identification system is that it has the capability to identify both P- and S-phase arrival times. Moreover, other phase types with sufficient SNR can be recognized by adapting this method to accommodate any distance. In addition, the *PhasePApy* can process and integrate the new incoming data with existing data, which can improve the earthquake location, origin time, magnitude, number of observed first motion of P-

waves, and as such can help further seismological analysis such as tomography and focal mechanism studies. Depending on users' requirements, the *PhasePicker* and the *Associator* can work separately or jointly.

Currently, *PhasePApy* is not designed to run in real time, but it can be utilized in a near real-time environment such as that employed for auto-locations and event notification at the Oklahoma Geological Survey (OGS). An older version of the *PhasePApy* package is being used to monitor the earthquakes in near-real time for OGS. The OGS, currently, automatically identifies and locates more than 1000 earthquakes each month in Oklahoma and surrounding regions. We compared the events identified by the *PhasePApy* to those manually identified, and the method presented here can automatically identify nearly all earthquakes of magnitude 2.0 or greater by using just a subset of the available stations. The *PhasePApy* can be applied beyond earthquake phase association and has been seen to perform well for active source experiments. The *PhasePApy* can clearly identify phase onsets and associate those phases to the active source. In addition, the inclusion of the *3D Associator* makes the use of the *PhasePApy* appropriate for micro-seismic monitoring of well stimulations, as well as other applications. The *PhasePApy* should work well for processing micro-seismic data independent of the surface or down-hole geometry.

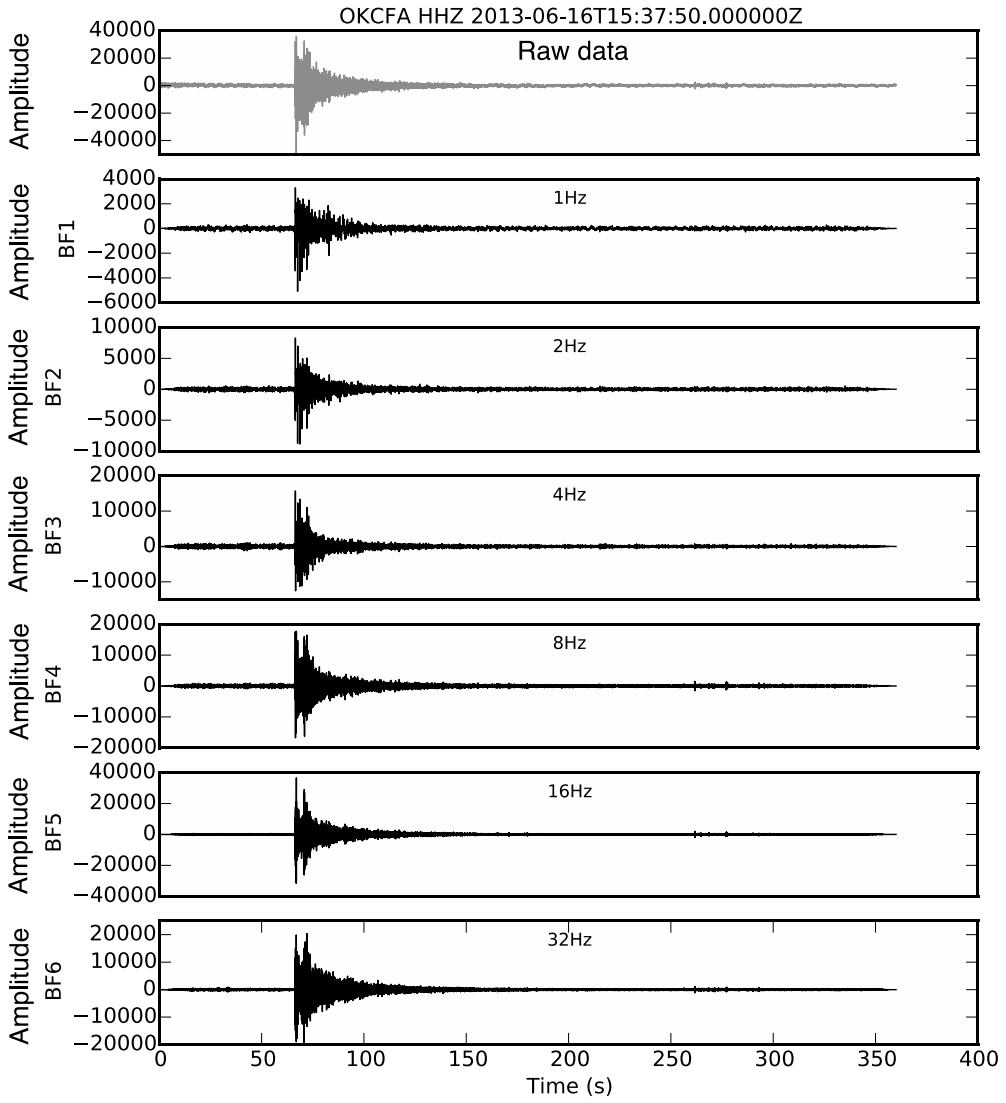


Figure 1.1 A de-trended example seismogram (top panel) from vertical component of station OKCFA in Oklahoma regional network (OK) is octave-filtered with 6 determined bands ($\boldsymbol{BF_1}$-$\boldsymbol{BF_6}$) from the *FBpicker*. The sampling rate of the data is 100 Hz. The central frequency of each band is labeled in the middle of each panel. A 10% tapering is applied to all bands. Amplitude is in counts.

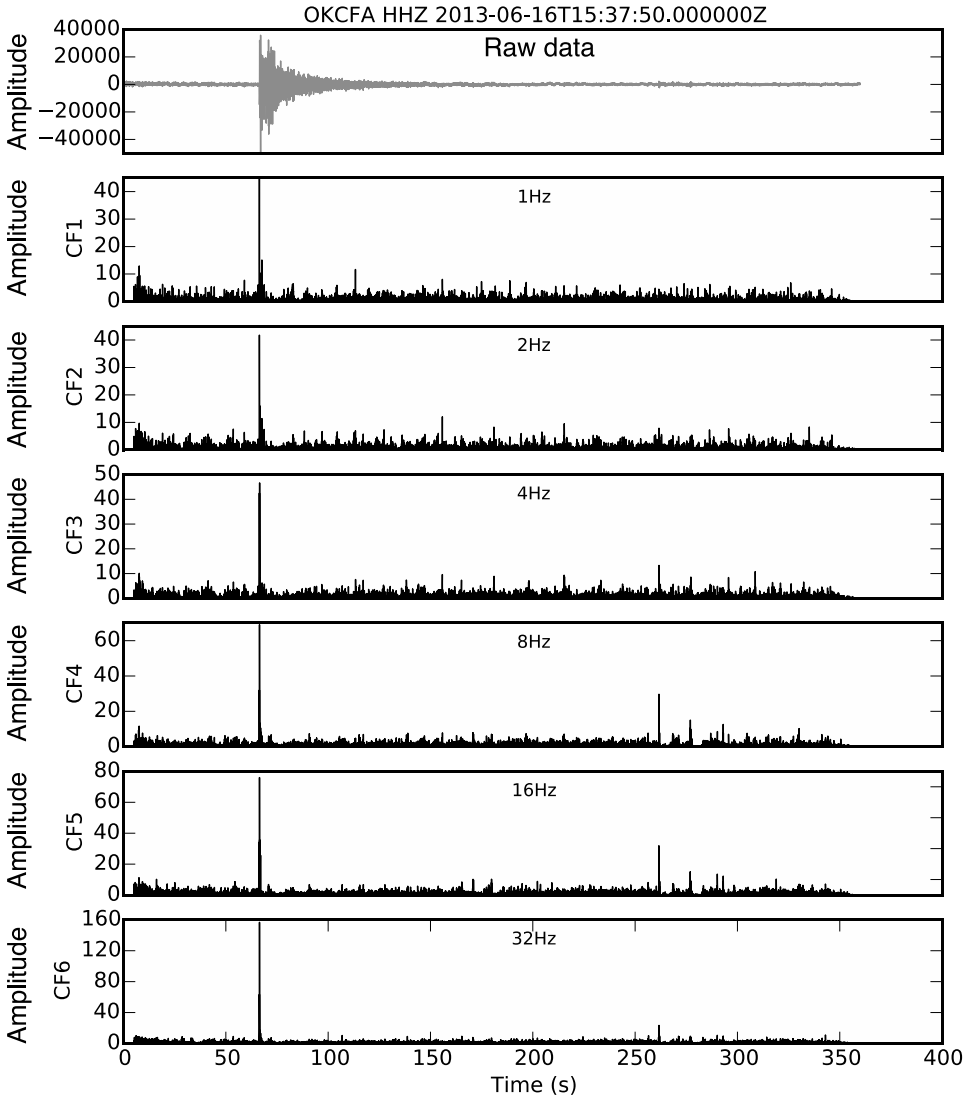


Figure 1.2. The de-trended seismogram on vertical channel from OKCFA (top panel) and the *FBpicker* determined the $\boldsymbol{CF_n}$ ($\boldsymbol{CF_1}$-$\boldsymbol{CF_6}$) with RMS mode. Amplitude is in counts.

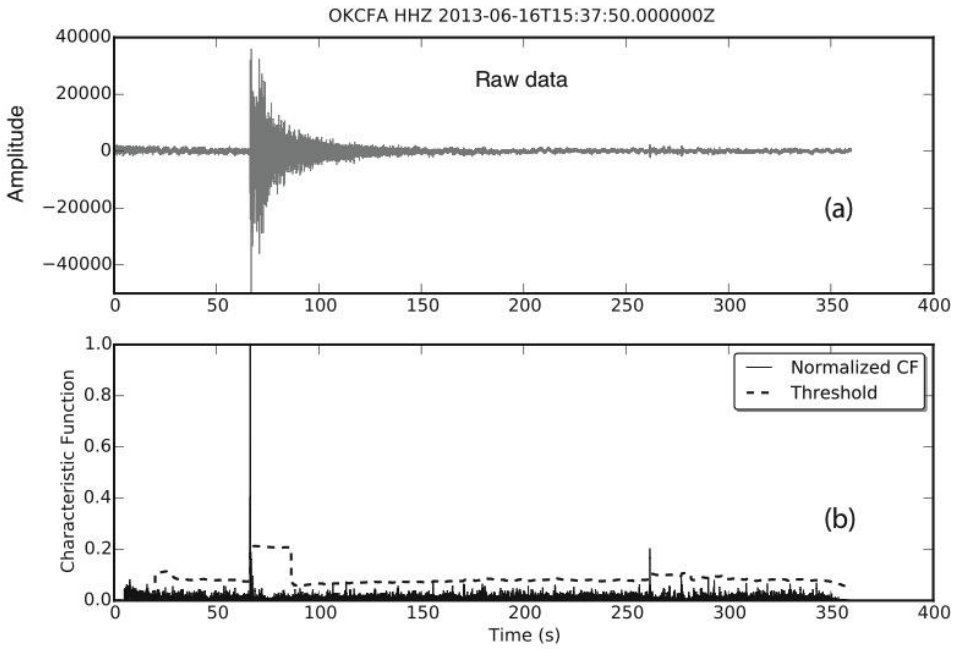


Figure 1.3. (a) The de-trended seismogram on vertical channel from OKCFA. (b) The *FBpicker* summarizes and normalizes the CF (solid line), and the dynamic threshold level of 6 shown as a dashed line. Amplitude is in counts.

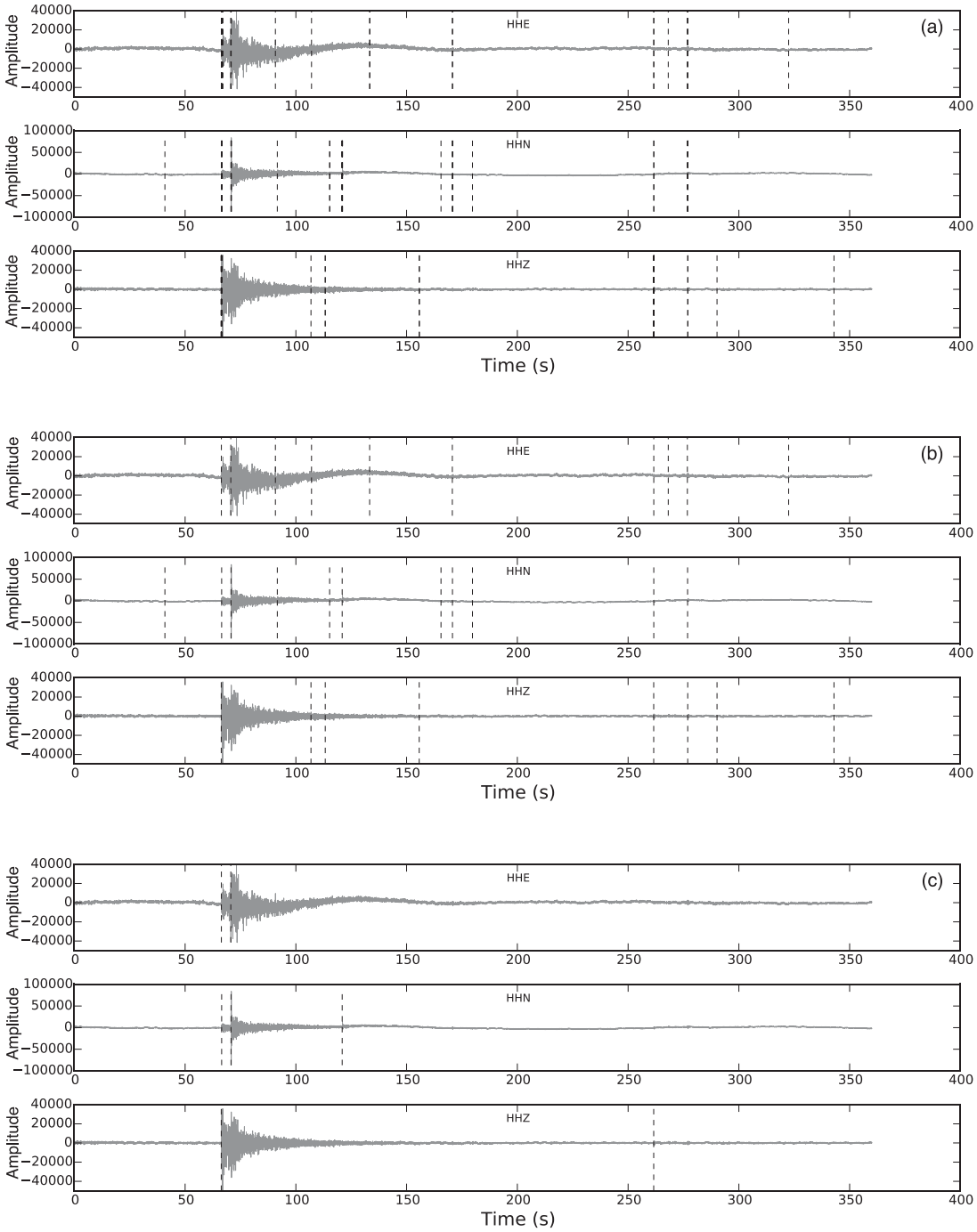
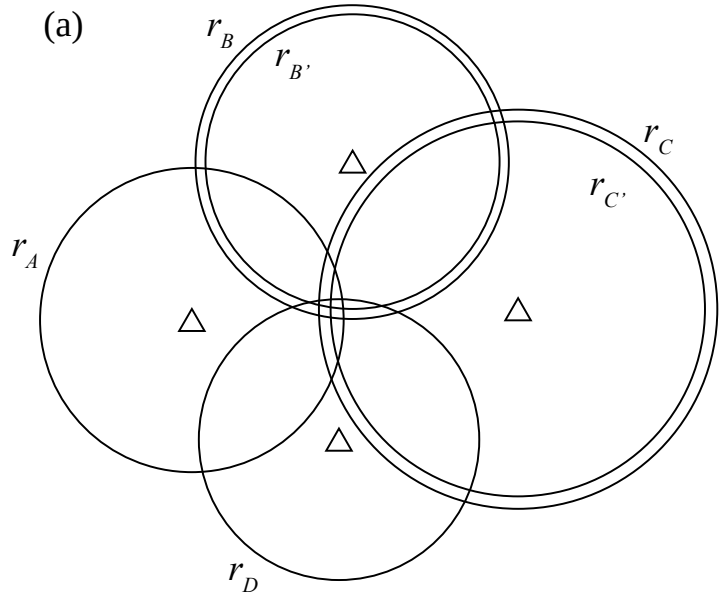

(a)
HHE
HHN
HHZ
Amplitude
Time (s)
(b)
HHE
HHN
HHZ
Amplitude
Time (s)
(c)
HHE
HHN
HHZ
Amplitude
Time (s)

Figure 1.4. The *FBpicker* false pick filtering on E, N and Z components. (a) All declared picks without applying false picks filters, (b) only applying the close-pick cleaning filter, (c) applying both close-pick cleaning filter and short-period noise-cleaning filter. Amplitude is in counts.


(a)
r_B
$r_{B'}$
r_C
$r_{C'}$
r_A
r_D

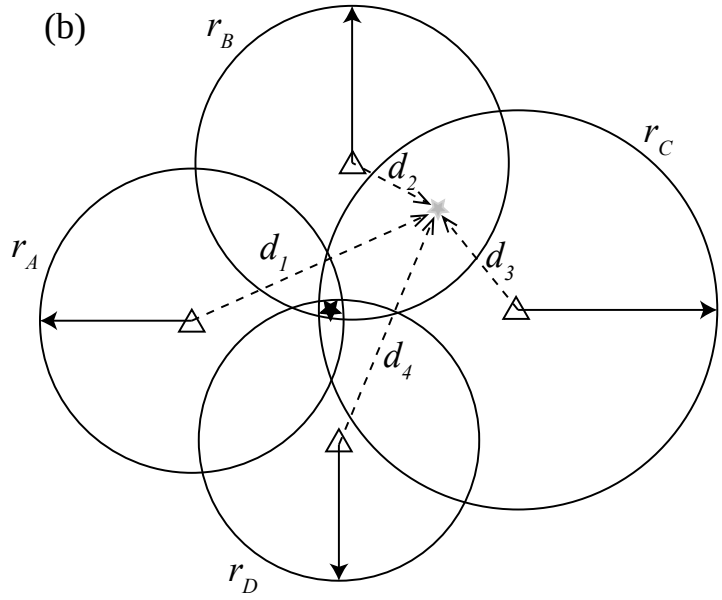

(b)
r_B
r_C
d_2
r_A
d_1
d_3
d_4
r_D

Figure 1.5. (a) Six event candidates in the origin time cluster, which are A, B, B', C, C', and D with the determined radius: $\boldsymbol{r_A}$, $\boldsymbol{r_B}$, $\boldsymbol{r'_B}$, $\boldsymbol{r_C}$, $\boldsymbol{r'_C}$, and $\boldsymbol{r_D}$. Event candidate B and B' are from the common station. Event candidate C and C' are from another common station. There are four combinations to create the sub-cluster with event candidates only from different station: S_1 = [A, B, C, D], S_2 = [A, B', C, D], S_3 = [A, B, C', D], and S_4 = [A, B', C', D]. Triangles represent the stations from where the event candidates come. (b) After excluding concentric circles, location determined by minimizing the RMS of the misfits between the S-P interval in distances $\boldsymbol{r_i}$ and the distances $\boldsymbol{d_i}$ from stations to trial epicenter (gray star). The final determined location is around the circles crossing area (black star).

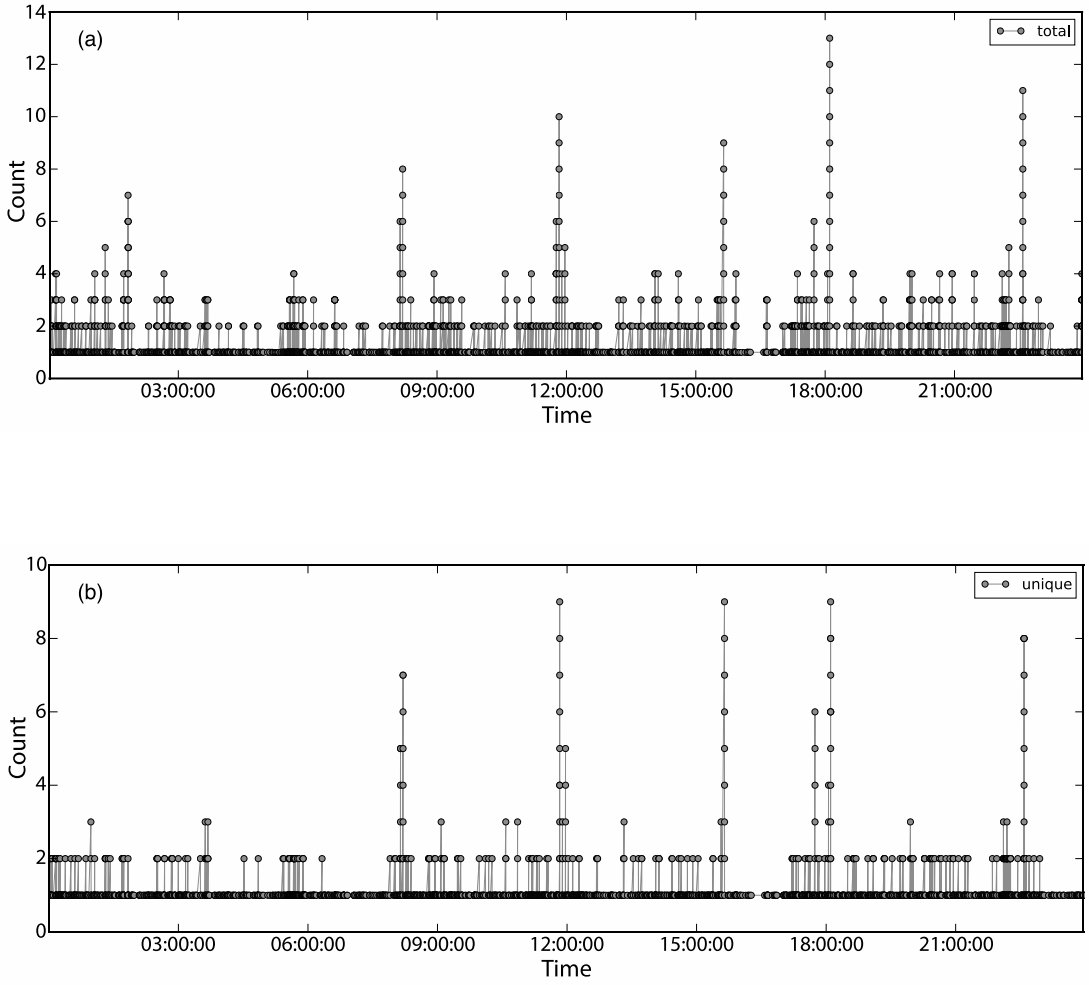
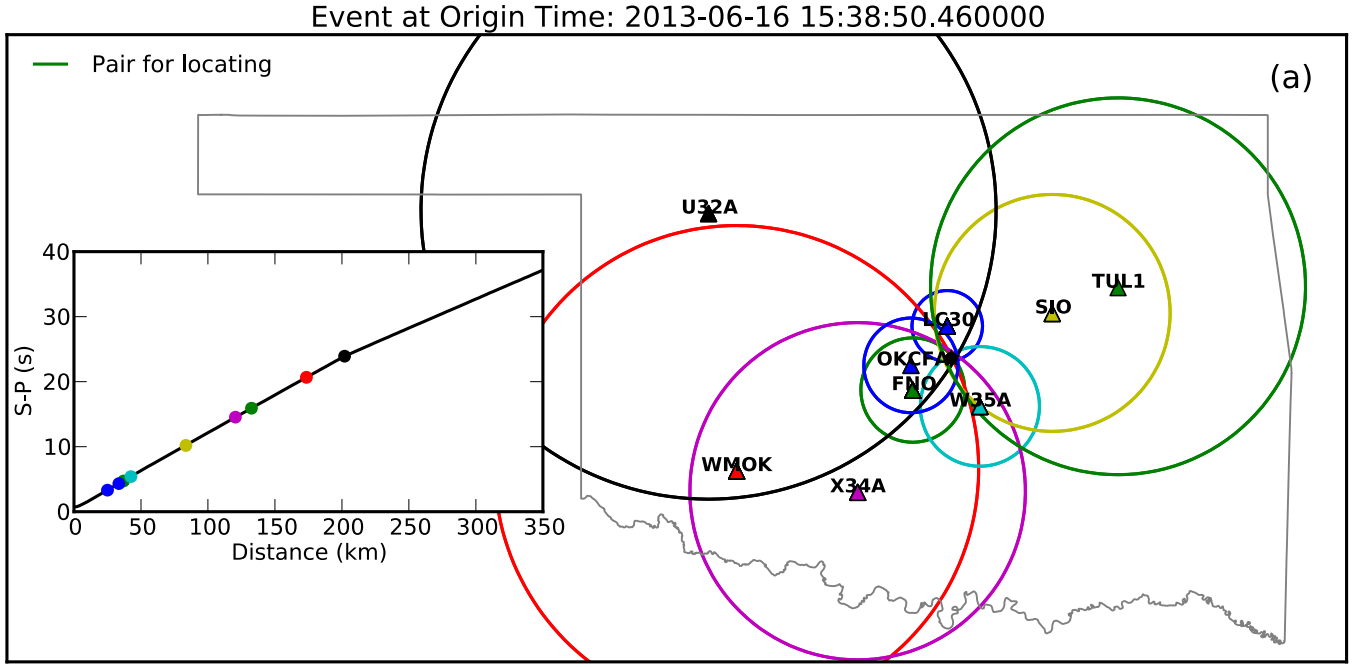


Figure 1.6. Cluster analysis of 1-day data on June 16, 2013. (a) Clusters including concentric circles. (b) Clusters excluding concentric circles.


Event at Origin Time: 2013-06-16 15:38:50.460000
Pair for locating
(a)
U32A
TUL1
SIO
LG30
OKCFAA
FNO
W35A
WMOK
X34A
S-P (s)
Distance (km)
0
10
20
30
40
0
50
100
150
200
250
300
350

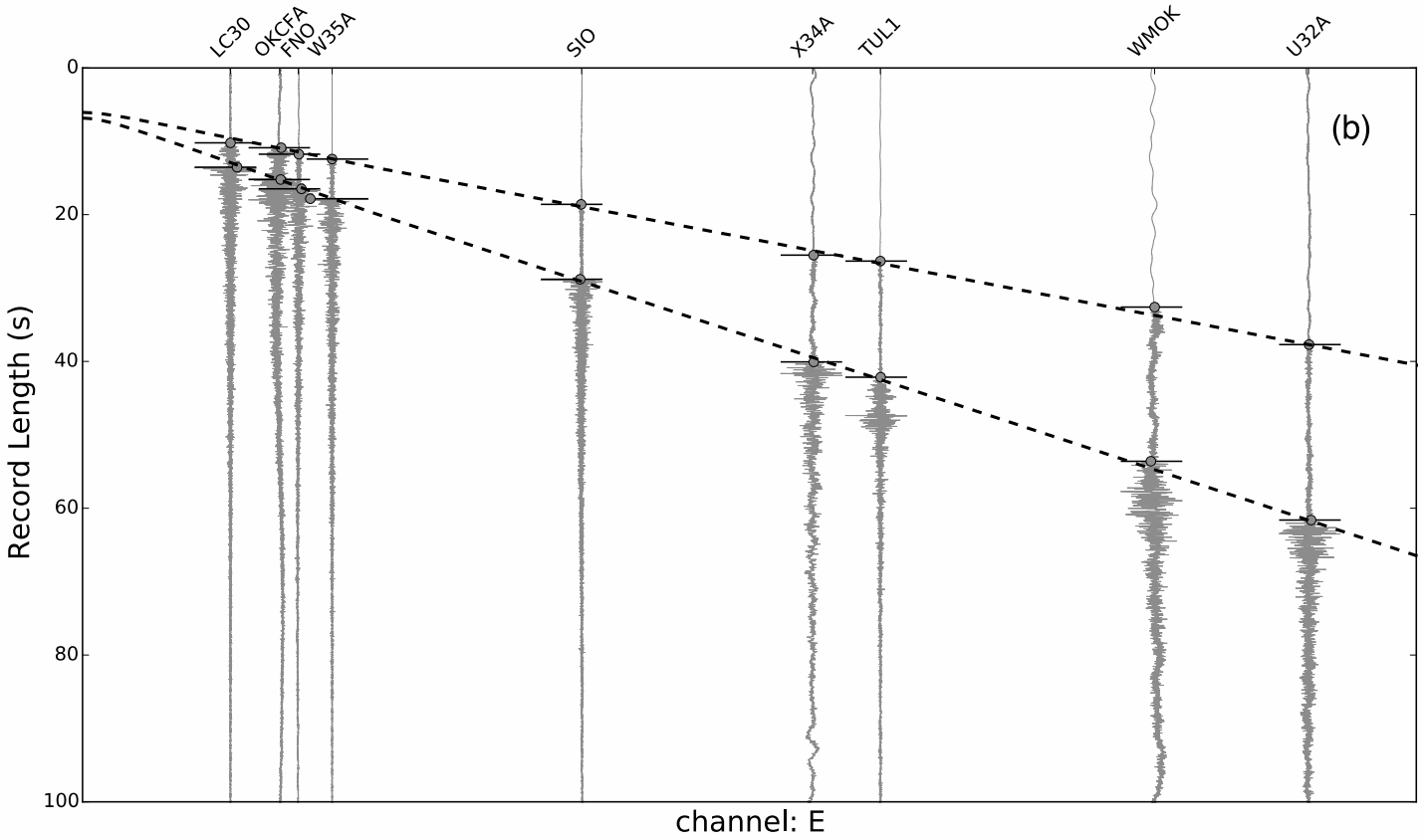

LC30
OKCFA
FNO
W35A
SIO
X34A
TUL1
WMOK
U32A
(b)
Record Length (s)
0
20
40
60
80
100
channel: E

Figure 1.7. Seismograms from an earthquake that occurred at 15:38:50 on June 16, 2013. (a) Map view of the event location determination. The triangles are the stations (green: TUL1 (big), FNO (small); blue: OKCFA (big), LC30 (small); red: WMOK; cyan: W35A; black: U32A; yellow: SIO; magenta: X34A). Circles are S-P interval in distance of event candidates. The star in the crossing area indicates the epicenter. The subplot shows the modeled S-P time curve and the *1D Associator* determined S-P times for event candidates. (b) Cross section plot of the *FBpicker* and *1D Associator* performance on channel E. Two dash lines are phases travel time curves (Upper: P; Lower: S). Short bars: the associated picks from the stations. Gray dots: where the bars cross waveforms.

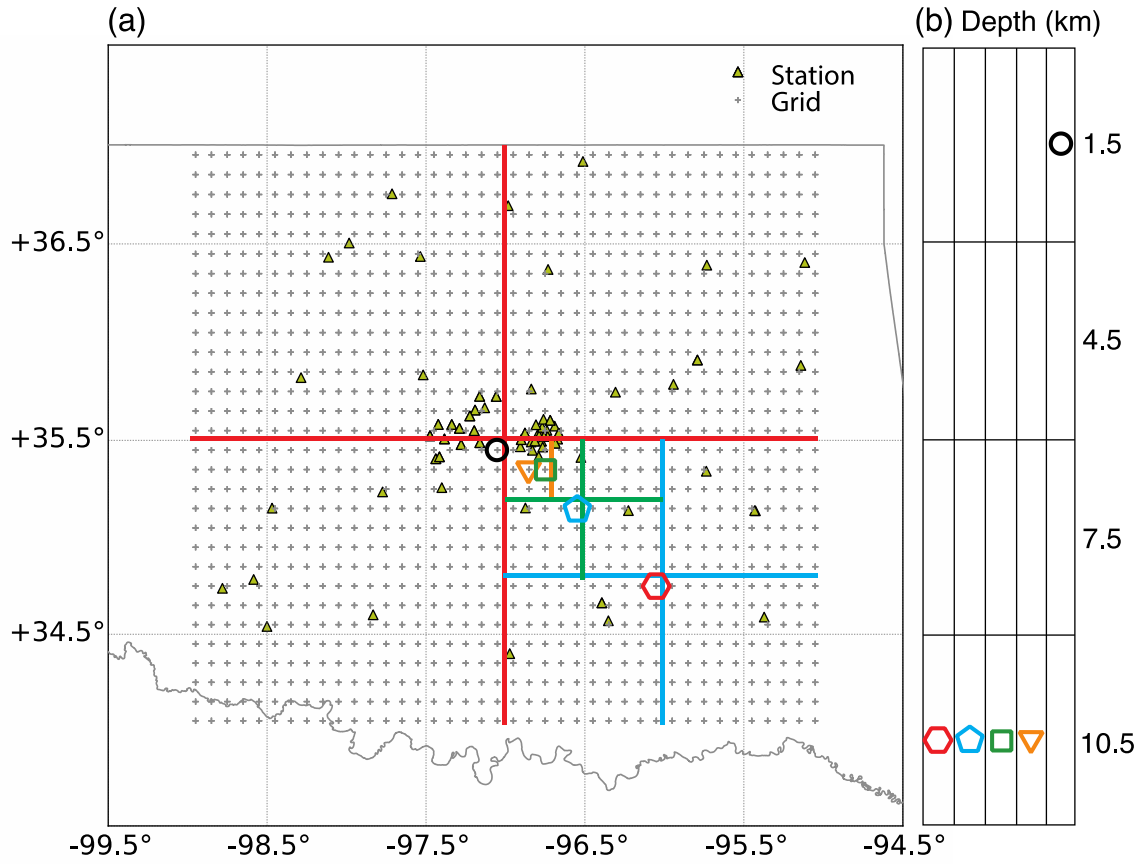


Figure 1.8. Grid search for the hypocenter of the same earthquake in Fig. 1.7(a). Map view of the searching process. Gray crosses: searching grids; yellow triangles: seismic stations; colored lines: sub-block boundaries; polygons and circles: the determined grid with least RMS residuals in each search. (b) Depths of the grids. The polygons and circles are consistent with the ones in (a).

Appendix: Supplementary material

Octave band-pass filtering of the *FBpicker* is shown in Figure S1.1. The consecutive band doubles the central frequency until the high corner of next band exceeds the Nyquist frequency. If the central frequency of a filtering band is N Hz, the low and high corner is $0.75\times N$ Hz and $1.5\times N$ Hz, respectively. For the next band, the central frequency is $2\times N$ Hz; the low and high corner is $1.5\times N$ Hz and $3\times N$ Hz, respectively, and as such can ensure that there is no gap and significant overlap in filtering band and all signal components below Nyquist frequency can be resolved. We take the window length of 10 in samples and the coefficient of 10 for pick polarity determination (Fig. S1.2). The polarity method determines the polarity of the first pick from the vertical component as dilatation (denoted as 'D'). The normalized characteristic functions of the *AICDpicker* and the *KTpicker* are in Figures S1.3 and S1.4, respectively.

The *Associator* stacks and aggregates picks with a user-defined link length. The picks from different components of the same station are stacked for aggregation. If the succeeding pick is close to the preceding one (less than the link length), the succeeding pick will be linked for the pick aggregation. However, if there is a pick from the same component that has already been linked, the succeeding pick will be treated as the first pick in the next pick aggregation process (Fig. S1.5). Although P1 and P2 are less than link length; they are treated as different aggregated picks.

The data from the Oklahoma Geological Survey OK network are archived and available at the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (DMC) through http://ds.iris.edu/ds/nodes/dmc/. The data from

station OKCFA can be downloaded using IRIS DMC utilities. The *PhasePApy* package is available on GitHub at https://github.com/austinholland/PhasePApy.

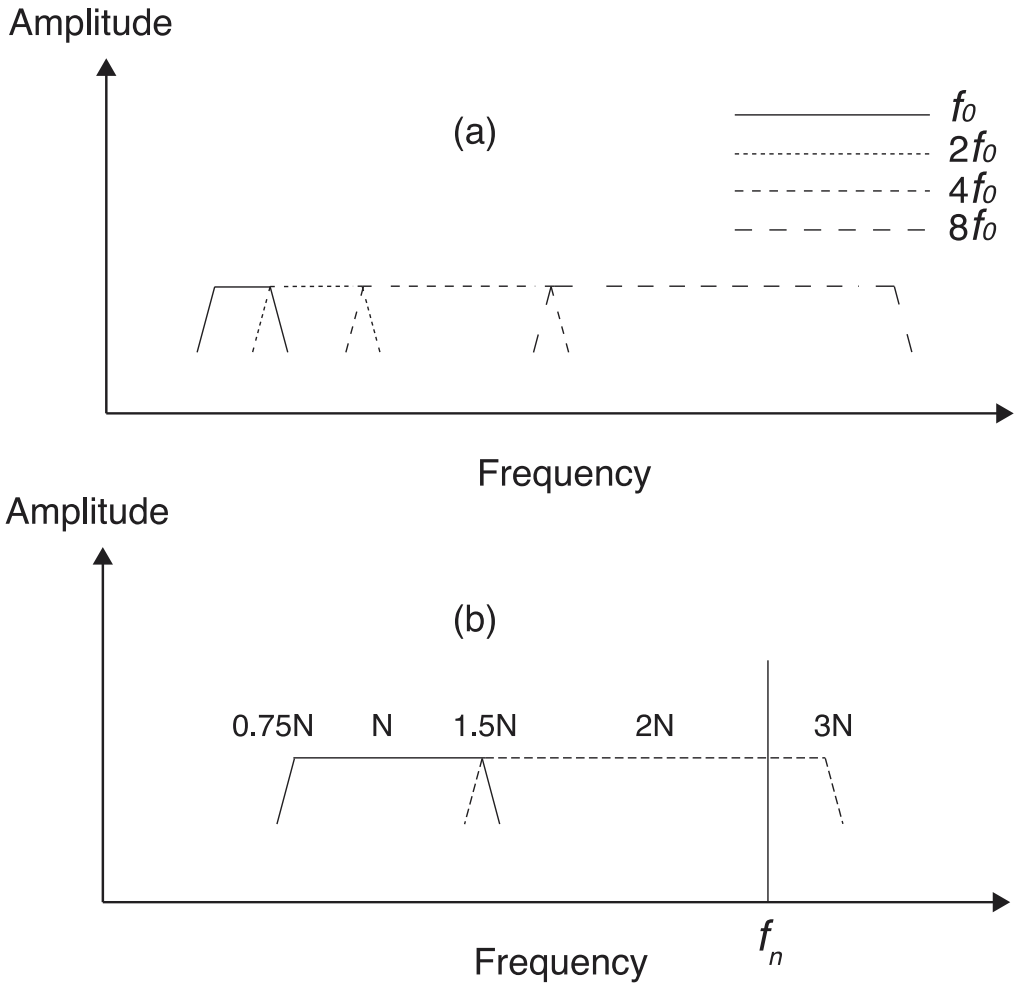


Figure S1.1. Determination of frequency bands for octave-filter from the *FBpicker*. (a) The central frequency of a band is doubled to that of its previous band. (b) The high-corner frequency of the last band (central frequency is *N* Hz) cannot exceed the Nyquist frequency f_N.

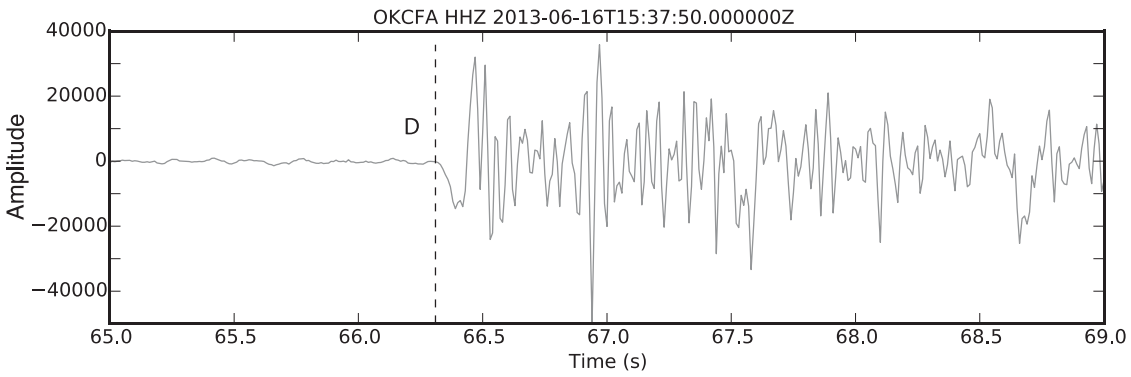


Figure S1.2. Example with the polarity determination (D stands for Down) of a pick on the vertical component signal, dashed line indicates the pick. Amplitude is in counts.

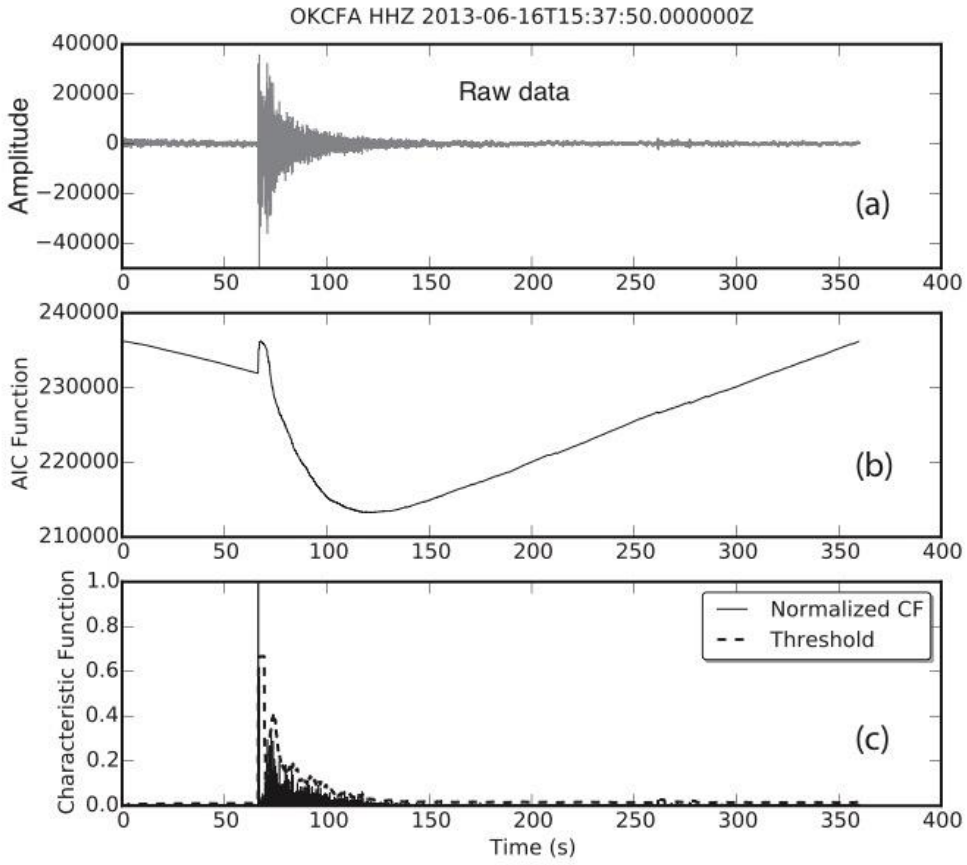


Figure S1.3. (a) De-trended raw seismogram on the vertical channel from station OKCFA. (b) AIC function of data. (c) The normalized derivative of the AIC function, and the dynamic threshold level of 6σ is in dashed line. Amplitude is in counts.

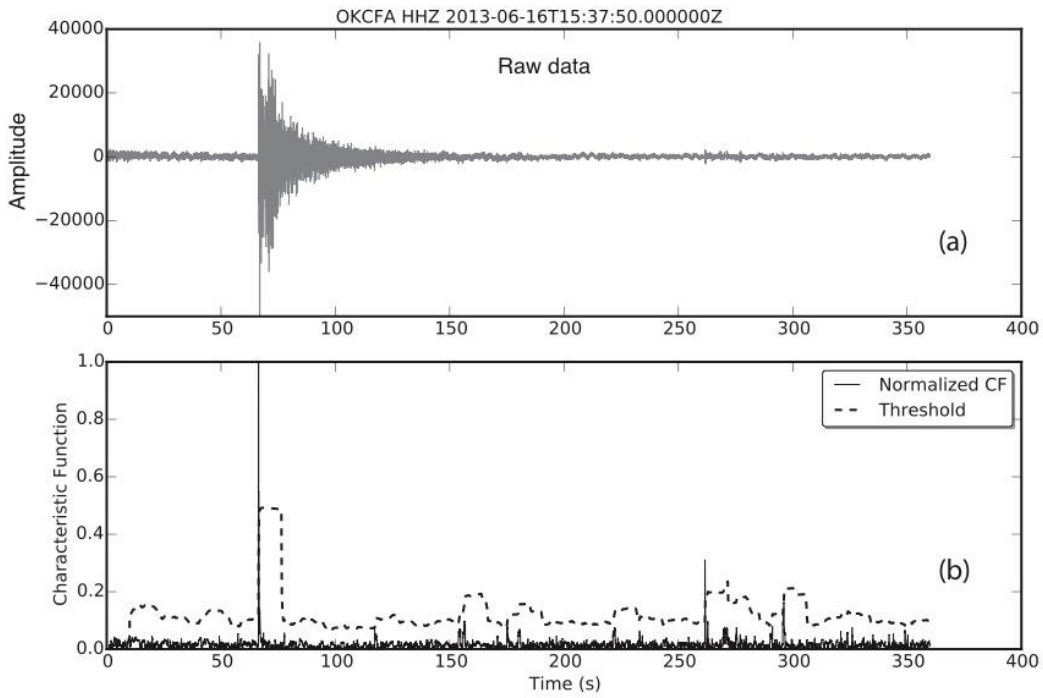


Figure S1.4. (a) The de-trended vertical channel seismogram from OKCFA. (b) The *KTpicker* normalized CF, and the dynamic threshold level of 6σ is in the dashed line. Amplitude is in counts.

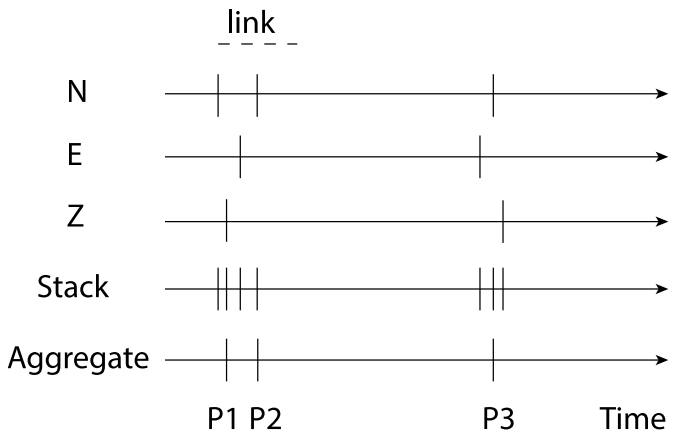


Figure S1.5. Stack and aggregate picks to obtain the modified picks from different components on the same station. N, E, and Z indicate two horizontal and a vertical components, respectively. The dashed line represents the link length to group the picks from different components for aggregation. P1, P2, and P3 are the aggregated picks from the stacking picks.

References

Akaike, H. (1974). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes, Annals of the Institute of Statistical Mathematics 26, no. 1, 363-387.

Akansu, A. N., W. A. Serdijn, and I. W. Selesnick (2010). Emerging applications of wavelets: A review, Physical. Communication **3**, no. 1, 1-18, doi: 10.1016/j.phycom.2009.07.001.

Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, Bulletin of the Seismological Society of America **68**, no. 5, 1521-1532.

Alwan, L. C. and H. V. Roberts (1988). Time-Series Modeling for Statistical Process Control, Journal of Business & Economic Statistics, **6**, no. 1, 87-95, doi: 10.2307/1391421.

Anant, K. S., and F. U. Dowla (1997). Wavelet transform methods for phase identification in three-component seismograms, Bulletin of the Seismological Society of America **87**, no. 6, 1598-1612.

Baer, M., and U. Kradolfer (1987). An automatic phase picker for local and teleseismic events, Bulletin of the Seismological Society of America **77**, no. 4, 1437-1445.

Baillard, C., W. C. Crawford, V. Ballu, C. Hibert, and A. Mangeney (2014). An automatic Kurtosis-based P- and S-phase picker designed for local seismic networks, Bulletin of the Seismological Society of America **104**, no. 1, 394–409, doi: 10.1785/0120120347.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010). ObsPy: A Python toolbox for seismology, Seismological Research Letters **81**, no. 3, 530-533, doi: 10.1785/gssrl.81.3.530.

Bogiatzis, P., and M. Ishii (2015). Continuous wavelet decomposition algorithm for automatic detection of compressional- and shear-wave arrival times, Bulletin of the Seismological Society of America **105**, no. 3, 1628-1641, doi: 10.1785/0120140267.

Dai, H., and C. MacBeth (1995). Automatic picking of seismic arrivals in local earthquake data using an artificial neural network, Geophysical Journal International **120**, no. 3, 758-774, doi: 10.1111/j.1365-246X.1995.tb01851.x.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis, Psychological Method 2, no. 3, 292-307, doi: 10.1037/1082-989X.2.3.292.

Gentili, S., and A. Michelini (2006). Automatic picking of P and S phases using a neural tree, Journal of Seismology **10**, no. 1, 39-63, doi: 10.1007/s10950-006-2296-6.

Hibert, C., A. Mangeney, G. Grandjean, C. Baillard, D. Rivet, N. M. Shapiro, C. Satriano, A. Maggi, P. Boisser, V. Ferrazzini, W. Crawford (2014). Automated identification, location, and volume estimation of rockfalls at Piton de la Fournaise volcano, Journal of Geophysical Research: Earth Surface **119**, no. 5, 1082-1105, doi: 10.1002/2013JF002970.

Hunter, J. D. (2007). Matplotlib: a 2D graphics evironment, Computing in Science & Engineering **9**, no. 3, 90-95, doi: 10.1109/MCSE.2007.55.

Jones, E., E. Oliphant, and P. Peterson, (2001). SciPy: Open source scientific tools for Python, see http://www.scipy.org.

Kitagawa, G., and H. Akaike (1978). A procedure for the modeling of non-stationary time series, Annals of the Institute of Statistical Mathematics **30**, no. 1, 351-363.

Leonard, M. (2000). Comparison of manual and automatic onset time picking, Bulletin of the Seismological Society of America **90**, no. 6, 1384-1390, doi: 10.1785/0120000026.

Leonard, M. and B. T. N. Kennett (1999). Multi-component autoregressive techniques for the analysis of seismograms, Physics of the Earth and Planetary Interiors **113**, no. 1, 247-263, doi:10.1016/S0031-9201(99)00054-0.

Lomax, A., C. Satriano, and M. Vassallo (2012). Automatic picker developments and optimization: FilterPicker—a robust, broadband picker for real-time seismic monitoring and earthquake early warning, Seismological Research Letters **83**, no. 3: 531-540, doi:10.1785/gssrl.83.3.531.

Maeda, N. (1985). A method for reading and checking phase times in autoprocessing system of seismic wave data, Journal of the Seismological Society of Japan **38**, no. 2: 365–379.

Oliphant, T. E. (2007). Python for scientific computing, IEEE Compuing in Science & Enginerring **9**, 10-20.

Panagiotakis, C., E. Kokinou, and F. Vallianatos (2008). Automatic P-phase picking based on local-maxima distribution, IEEE Transactions on Geoscience and Remote Sensing **46**, no. 8, 2280-2287, doi: 10.1109/TGRS.2008.917272.

Sleeman, R., and T. van Eck (1999). Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings, Physics of the Earth and Planetary Interiors **113**, no. 1, 265-275, doi:10.1016/S0031-9201(99)00007-2.

Vassallo, M., C. Santriano, and A. Lomax (2012). Automatic picker developments and optimization: a strategy for improving the performances of automatic phase pickers, Seismological Research Letters **83**, no. 3, 541-554, doi:10.1785/gssrl.83.3.541.

Wang, J. and T. Teng (1997). Identification and picking of S phase using an artificial neural network, Bulletin of the Seismological Society of America **87**, no. 5, 1140-1149.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection, Bulletin of the Seismological Society of America **88**, no. 1, 95-106.

Zeiler, C., and A. A. Velasco (2009). Seismogram picking error from analyst review (SPEAR): single-analyst and institution analysis, Bulletin of the Seismological Society of America **99**, no. 5, 2759-2770, doi: 10.1785/0120080131.

Zhang, H., C. Thurber, and C. Rowe (2003). Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, Bulletin of the Seismological Society of America **93**, no. 5, 1904-1912, doi: 10.1785/0120020241.

Chapter 2: Tomographic Study for Central Oklahoma by Integrating Gravity and Magnetic Data

Introduction

Oklahoma has drawn a great deal of attention because a large number of earthquakes have occurred in recent years (e.g., Ellsworth, 2013; McGarr et al., 2015; Petersen et al., 2015). Since 2009, the total number of Oklahoma earthquakes has exceeded that of all events identified since statehood. The average number of earthquakes with magnitude equal to or greater than 3.0 before 2009 was 1.6 per year, and by 2014 this number increased to 585 (Darold et al., 2015). Earthquakes are expected to occur within active tectonic regions. It is unusual that such large earthquake swarms occur in an intraplate region where the expected strain accumulation is low (Ellsworth, 2013). Most seismicity in Oklahoma occurred in the Precambrian basement, and exhibit general spatial correlation with areas where wastewater injection is common (Ellsworth, 2013; Holland, 2013; Keranen et al., 2013, 2014; Darold et al., 2014, 2015; Murray, 2014).

There are several previous studies to investigate the relationships between the seismicity and tectonics in Oklahoma (e.g., Luza and Lawson, 1981, 1982). Prior to 2009, the earthquakes in Oklahoma occurred over a broad region, without showing a clear relationship with previously known structures. However, the recent profound seismicity increase may be related to reactivation of faults and/or fractures within the basement, which could be triggered by wastewater injection that increases the pore-pressure to promote fault slip. The reactivation of subsurface faults may play an important role in the recent seismicity increase in central Oklahoma because some

earthquake clusters align with the mapped subsurface faults or extension of the mapped fault trends. However, the majority of the seismicity clusters do not well correlate with mapped faults, which probably suggest that many faults are deep in the basement, and have not been documented with shallow subsurface methods (McNamara et al., 2015). Most identified subsurface faults in central Oklahoma are northeast-southwest (NE-SW) or northwest-southeast (NW-SE) trending orientation. Holland (2013) quantified the probability (or risk) of triggered seismicity as function of the fault orientation and general stress field within Oklahoma. Darold and Holland (2015) updated the preliminary optimal Oklahoma fault orientations based on oil and gas industry data and published literature.

The basement structures that formed in multiple tectonic and metamorphic episodes control local structure evolution such as faulting and/or folding in the overlying sedimentary section (Gay, 1995). Therefore, a better understanding of the basement, such as optimally oriented faults and petrophysical properties of rock, is critical to study the increased seismicity in central Oklahoma. Although there were some seismic studies conducted in central Oklahoma, only a few of them image structures within crystalline basement (e.g., Tryggvason and Qualls, 1967; Toth, 2014). A three-dimensional (3D) tomographic study in central Oklahoma can be a good complementary to seismic reflection studies that were primarily for oil and gas exploration and only penetrate into subsurface sedimentary layers.

In order to investigate upper crustal structures of this region, which are difficult to obtain from wells data and reflection seismic data, we derive two 3D velocity models and Vp/Vs ratios with the FMTOMO (Rawlinson and Urvoy, 2006) and the

SIMUL2000 (Thurber, 1983) software packages, respectively. The 3D velocity and Vp/Vs ratio variation in a crystalline basement can provide new insights of geological structure. Furthermore, the 3D velocity model can be used to improve locations of the earthquakes within this area as well.

Geological Background

Oklahoma has several well-documented geological structures that formed in the Late Precambrian, Early Paleozoic, and Pennsylvanian (e.g., Keller et al., 1983; Denison et al., 1984; Perry, 1989). Several studies summarized basic evolution of the Precambrian and Cambrian basement of the southern Midcontinent region (e.g., Muehlberger et al., 1967; Denison et al., 1984; Adams and Keller, 1996; Keller and Stephenson, 2007), which suggest that the depth to basement is about 1 km in northern Oklahoma and >10 km in southern Oklahoma. The Paleozoic tectonic activity produced the major structures in Oklahoma (Northcutt and Campbell, 1995; Johnson and Luza, 2008). The asymmetric Anadarko Basin is the deepest basin in North America (>10 km) that formed in late Paleozoic, which is bounded by the inverted rift structures, Wichita Uplift, in southwestern Oklahoma. The Arkoma Basin is a large foreland basin in southeastern Oklahoma, which is overthrusted by the Ouachita Mountains from the south. The Arbuckle Uplift that formed during the Ouachita Orogeny in south central Oklahoma separates the Anadarko Basin and the Arkoma Basin. The Wichita and Arbuckle Uplift and southern Anadarko Basin are structural elements of a failed arm of a Cambrian triple junction that are evident in the gravity map shown in Figure 2.1. This triple junction formed a passive continental margin with one arm extending northeastward across southern Arkansas and another one extending southwestward

across eastern Texas. The NW-trending third arm formed the Southern Oklahoma Aulacogen (SOA) (e.g., Shatski, 1946; Hoffman et al., 1974; Keller et al., 1983; Brewer et al., 1983; Kruger and Keller, 1986; Perry, 1989; Keller and Stephenson, 2007; Bucky, 2012), which is a classic example of an intraplate failed rift arm that was reactivated and inverted in the Pennsylvanian to form a series of uplifts with structural relief greater than 10 km (e.g., Shatski, 1946; Hoffman et al., 1974; Gilbert, 1983). The Anadarko Shelf and Cherokee Platform are gently dipping areas in northern Oklahoma. The Nemaha Fault Zone (NFZ) in north central Oklahoma, the Wilzetta Fault Zone (WFZ) in central Oklahoma, and the thrusting faults produced by the late Paleozoic continental collision in southwestern, southern, and southeastern Oklahoma are the major fault system in Oklahoma.

In addition, the Mid-Continent Rift (MCR) is an important tectonic feature in Oklahoma that formed in Proterozoic (Fig. 2.1), which can be identified from gravity and magnetic data (e.g., Adams and Keller, 1996; Stein et al., 2014). The west arm of the MCR was interpreted to at least extend to southern Oklahoma from the outcrop area, seismic reflection profiles, and drilled samples from high gravity anomaly area (e.g., Denison, 1966; Adams and Keller, 1996; Merino et al., 2013; Stein et al., 2014). The interpreted southward extension of the MCR may associate with seismicity increase in central Oklahoma.

The crystalline basement in north central Oklahoma can be divided into four major Precambrian units (Fig. 2.2): the Washington County Volcanic Group (WCVG), the Spavinaw Granite Group (SGG), the Osage County Microgranite (OCM), and Central Oklahoma Granite Group (COGG) (Denison, 1981; Luza and Lawson, 1982).

The WCVG is rock unit primarily composed of rhyolite with volcanic extrusion over a board area in northeastern Oklahoma. The SGG refers to the magnetic low in eastern central Oklahoma, which is mainly composed of microgranite and thought to comprise of a sill complex underlies the cover of rhyolite. In north central Oklahoma, the OCM is the most petrographically uniform rock unit and believed to be intruded under the rhyolite cover. The COGG is a younger rock unit composed of medium-grained granite in a broad region of central Oklahoma.

Data and Methods

Earthquake data

Most of the earthquakes were recorded around Jones and Prague area in central Oklahoma from 2009 to 2013. Starting in late 2013, the earthquake activity expanded northward in a board region, even extending to southern Kansas. Due to a significant increase in seismicity in central Oklahoma, the Oklahoma Geological Survey (OGS), U.S. Geological Survey (USGS) and other groups have deployed a variety of instruments throughout the state to monitor the seismicity (e.g., Holland et al., 2011). The available stations in Oklahoma after 2009 include the Transportable Array (e.g., Astiz et al., 2014), permanent stations, and temporary stations deployed by the OGS, USGS, Incorporated Research Institutions for Seismology (IRIS), and others. The OGS located a total of >15000 earthquakes within the study area from 01/01/2010 to 07/31/2015. This large amount of seismic data provides a great opportunity to study the velocity model of upper crust with a tomographic method in central Oklahoma. There were 8194 M2+ earthquakes with depths >3.0 km, the number of station with observed arrivals >12, and root mean square (RMS) of traveltime residuals <0.6 s selected so as

to control the data quality and computation time (Fig. 2.3).

All the phase arrivals were manually picked by seismic analysts using SEISAN software package (e.g., Ottemöller et al., 2013) and identified by the automatic phase identification package *PhasePApy*, which is robust pure *Python* package for automatic identification of seismic phases (Chen and Holland, 2016). The analysts reviewed those automatic phase picks. Somehow there were some questionable picks in the selected data, which can significantly bias the tomography results, especially the outliers with unreasonable traveltimes (i.e. negative traveltimes). In order to better control the velocity model from the tomographic inversion, we filtered the picks with traveltime errors greater than 1.5 s comparing to predicted travel-time based on 1D velocity model (Table 2.1), and we also removed the picks of epicentral distance greater than 350 km, which is about the dimension of our study region (Fig. 2.4(a)). The 1D P-wave velocity model in Table 2.1 was derived from digitizing sonic logs near the WFZ (Toth, 2014) and surface wave dispersion curves (Robert Hermann pers. comm. 4/17/13). The Vp/Vs ratio is employed as the typical value of 1.73. After picks filtering, there were 153119 P- and 145949 S-picks left. The traveltime residuals, focal depth, and magnitude distribution are shown in Figure 2.4(b)-(d), respectively. The average focal depth of the earthquakes in central Oklahoma is more or less 5 km excluding those events with fixed depth (Darold et al., 2015). The filtered P-waves ray path can cover most area in central and north central Oklahoma (Fig. 2.5). The filtered S-waves ray path has the similar coverage to one of the P-waves due to they have the comparable number of picks.

Method 1

Many algorithms have been developed for local or regional traveltime

tomography studies (e.g., Thurber, 1983; Zhang and Thurber, 2003; Rawlinson and Urvoy, 2006). In this study, we created the first velocity model with the Fast Marching Tomography (FMTOMO) method (Rawlinson and Urvoy, 2006), which employs the multi-stage fast marching method (FMM) as the grid-based forward modeling eikonal solver to predict the traveltimes (Sethian and Popovici, 1999). The FMTOMO solves the non-linear inversion problem by repeatedly and iteratively adjusting model parameters to reconcile the objective function (velocity, interface depth, and/or source location) and traveltime perturbations. The FMTOMO can invert 3D velocity with multiple types of the seismic phase, such as local, region and/or teleseismic data. We used the filtered P- and S- picks from the 8194 M2+ local events to determine Vp, Vs, and Vp/Vs ratio, respectively.

Resolution of velocity model 1

We estimated the reliable region of both P- and S-wave tomography results with checkerboard tests. The alternatively perturbed velocity anomalies were imposed into the initial 1D velocity model to create a checkerboard pattern. The FMTOMO linearly interpolate the 1D velocity for the equal spacing depth nodes (Fig. 2.6). Positive and negative anomalies, 0.6 km/s, were added to P-wave velocity, and similarly ±0.3 km/s were added to S-wave velocity alternatively. The model was parameterized with an optimized node spacing about 0.15° (Δx: ~13 km) × 0.12° (Δy: ~13 km) × 2.8 km (Δz) after several tests with different node spacing (Fig. 2.7). The solution converged well, and the residuals were reduced to a reasonable level after 4 iterations. The checkerboard patterns of P-wave velocity model shallow than ~15 km were recovered for central and north central Oklahoma (depths of 3 km and 13 km in Fig. 2.8). The region of a good

resolution was indicated with a polygon outlined by the dashed lines in Figure 2.8. As would be expected, the resolution was diminished in the 13 km depth slice, since there are only a few earthquakes deeper than 10 km (Fig. 2.4(c)), which results in fewer ray paths penetrate to such depth.

The vertical cross-section slices in Figure 2.9 show areas of good resolution down to the depths about 15 km for both P- and S-wave velocity. The amplitude of the velocity perturbation is well recovered for depths <8 km. Although the resolution is reduced for depths from ~8 km to 15 km, the central portion of the checkerboard pattern can be recovered in this depth range. The patterns beneath 15 km were barely recovered due to the lack of deep events, although one can identify some poorly recovered patterns for depth deeper than 15 km. Therefore, we think that the most reliable area of the tomography with good resolution extends approximately from latitude 35° to 37.2° and longitude 96° to 98.5° (Fig. 2.8), and depth down to ~15 km (Fig. 2.9).

The checkerboard test of the S-wave velocity shows very similar resolution results, since the number of S-picks is quite comparable with the one of P-picks, only 4.7% less.

Method 2

We created the second velocity model by using the SIMUL2000 package (Thurber, 1983) with composite event method (Lin et al., 2007), which reduces the volume of the data so as to improve the computation time. Although the total number of picks is reduced, the composite event method preserves as much the original phase picks information as possible. In addition, this method makes the composite event distribution more uniform. We used 8194 M2+ events to start this tomography

processing. Due to the fact that close events have random picking error dominated residuals, we firstly used the HYPODD 1D (Waldhauser and Ellsworth, 2000; Waldhauser, 2001) method to relocate the earthquakes, which can provide better relative locations (Fig. 2.10). For those un-clustered and thrown away events due to the bad location, we added them back by using their initial locations in the catalog. Then a similar method as the one from Lin (2007) followed to construct the composite events (Fig. 2.11). We selected the first target earthquake that contains the most picks from the 1D velocity relocated data. A close event search algorithm identified the nearby earthquakes within a certain radius (2 km in our case) centered at the target event. For a certain station, we took the median value of all the picks from the events within the 2 km searching radius (including the target event) instead of taking the mean value as Lin (2007). The median value was assigned to the target event as the phase record from that certain station. A larger radius searching (5 km in our case) followed to exclude the events between two spheres (with radius 2 km and 5 km in our case, respectively) for candidate treatment of next target event in the future. There were 493 composite events with 13447 P-picks and 11293 S-picks constructed from 8194 M2+ events (Fig. 2.12). We used the first 485 composite events to conduct our second velocity modeling study, which have more than 99.5% picks from all the composite events.

We used a 30 km horizontal grid to invert for the 3D velocity model, which is about twice grid size of the FMTOMO model due to the computation capability restriction (Fig. 2.7). The vertical grid points are not equal spaced (Fig. 2.6), which can better represent our initial velocity model (Table 2.1). The Vp and Vp/Vs ratio damping parameters were optimized with a series of single-iteration inversion, respectively. For

the Vp damping determination, we used the damping parameters from 50 to 2000 to plot the data misfits versus model variation tradeoff curve (Fig. 2.13), while prohibiting the Vp/Vs ratio inversion. From the Vp tradeoff curve, we chose 300 as the optimized value since the data misfits can be reduced significantly without introducing too much model variance. By holding the Vp damping fixed at 300, we used a series of Vp/Vs ratio damping parameters from 10 to 1000 to determine the optimized value as 70. The tradeoff curve of Vp/Vs ratio is shown in Figure 2.14. We chose 300 and 70 as the values for Vp and Vp/Vs ratio, respectively, in our tomography inversions. Although the lower damping values can reduce the data misfits, they can introduce some sharper and strong velocity anomaly features that probably are due to artifacts.

Gravity, magnetic, and basement depth data

The free-air anomalies (FAA) gravity data and the pole reduced magnetic data used in this study were downloaded from gravity and magnetic database of Pan-American Center for Earth and Environmental Studies (PACES) at University of Texas at El Paso (UTEP) (Figs. 2.15 and 2.16). The gravity data present numerous short-wavelength artifacts, but should not affect our correlation with velocity structures at a large scale. The OGS developed a depth to basement top map based on well data and previous basement top contour maps from the USGS, the Colorado Geological Survey (CGS), the Kansas Geological Survey (KGS), the New Mexico Bureau of Geology, the Mineral Resources (NMBGMR), and others (Kevin Crain pers. comm., Oct./2015). The gravity, magnetic, and depth to basement top data were gridded to compare to two velocity models in this study region.

Modeling Results

In order to better understand the structures in the crystalline basement, we used the velocity models, gravity, and magnetic data to examine the geological correlation. We compared the velocity models from the FMTOMO and the SIMUL2000 methods to the gravity and magnetic data along 7 west-east (A1-A7) and 7 south-north (B1-B7) profiles, respectively. The locations of the profile are shown in Figures 2.15 and 2.16. The profiles go through most part of the region with good resolution in central Oklahoma. By comparing the 14 cross-sections of the Vp and Vp/Vs models from FMTOMO and SIMUL2000, we think that most of the large features in these two models are comparable (Figs. 2.17 and 2.18), although the FMTOMO model presents some small-scale features. We calculated the averaged velocity variations for the grid points at each depth in the cross-section region (Fig. 2.19) to show that the SIMUL2000 model has less velocity variation than the FMTOMO model. The sharper and small-scale velocity anomalies usually have large uncertainty. In addition, most of the profiles in the SIMUL2000 model show less Vp/Vs variation relative to the initial value 1.73 compared to the ones from the FMTOMO model (Fig. 2.20). The large FMTOMO model variation may be due to the small damping that we used in the inversion. Therefore, we chose the relatively smooth SIMUL2000 model to correlate gravity and magnetic anomalies for the following interpretations.

On a large scale, the velocity model correlates to gravity anomalies since the dense rocks generally have high velocities, and vice versa. The velocity models reveal strong lateral heterogeneities within the Precambrian crystalline basement, which indicates the complex structures in the upper portion of the crust in the study area.

Several evident anomalies can be identified in this study region from the magnetic map (Fig. 2.16), which were labeled from R1 to R4. The anomaly R1 between the NFZ and the WFZ is a strong magnetic feature to the south of our study region, whereas has not been mentioned by previous studies. The anomaly R2 (Spavinaw Granite), R3 (Lamont Ring Complex), and R4 (Osage Island Maximum) are previously known features in central Oklahoma (Luza and Lawson, 1982).

Cross-section A1: Since the western side of this Cross-section, beyond longitude about -98.2°, exceeds the polygon outlined region; we can only make interpretation from longitude about -98.2° to -96° at most. This cross-section goes through the magnetic high R1 (Fig. 2.16), and the magnetic data presents a relative high anomaly from longitude -97.4° to -97°. The gravity from longitude about -97.8° -to 96.6° also presents a broader high anomaly. There is a high Vp feature from longitude -97.6° to -97.0° for depth lower than ~8 km (Fig. 2.17a) in the SIMUL2000 model (use this model for the correlation in all the following cross-sections). The high Vp anomaly is bounded by the NFZ and the WFZ, which are vertical or near vertical in the sedimentary section. The high gravity, magnetic, and velocity anomaly suggest a presence of intrusion between the NFZ and the WFZ within the basement.

Cross-section A2: This cross-section crosses the south of the basement rock unit SGG (R2 in Fig. 2.16). Similarly, we do not interpret the velocity for the portion to the west of longitude -98.3°. There is only slight gravity high from longitude -97.8° to -96.5° in this profile, but a magnetic low centered at longitude about -97.1°. A strong high Vp anomaly is identified from longitude -97.5° to -96.4° for depth below ~8 km, which can be correlated to an intrusion related to the SGG.

Cross-section A3: This cross-section crosses the center of the unit SGG (R2 in Fig. 2.16). We use the range from longitude -98.5° to -96° for interpretation. A magnetic low presents from longitude about -97.2° to -96.4° and a strong high Vp anomaly body can be identified in this range for depth below ~8 km, which strongly suggest an intrusion below the SGG unit. However, a relatively low gravity anomaly shows in this range. We note that a low velocity zone at shallower depth to the east of NFZ overlies this strong Vp anomaly, which may reduce the total gravity observation. We restrainedly correlate a deep high Vp anomaly feature that centered at longitude 98° to the MCR.

Cross-section A4: This cross-section crosses the northern part of the unit SGG (R2 in Fig. 2.16) and the interpreted MCR. A magnetic low centered at longitude about -96.5° along this cross-section is underlain by a high Vp anomaly body below depth ~8 km, which suggest an intrusion. There is a slight high gravity anomaly centered at longitude about -96.3°, where is above the east portion of the intrusion. Another high velocity anomaly feature is centered at longitude about -98° that is interpreted as the MCR, where shows a corresponding high gravity anomaly. The velocity gradually decreases from the east portion of the intrusion to the MCR, which corresponds the gravity trend. The velocity and gravity present the local low anomalies at the common longitude -97.4°.

Cross-section A5: Two obvious features exist in this cross-section: the MCR to the west and southern part of the OCM (or WCVG) (R4 in Fig. 2.16) to the east, where underlain by two high Vp anomalies centered at longitude -98.2° and -96.2°, respectively. A relative magnetic high presents around the OCM unit, which indicates

that there is an intrusion in this area. The broad low gravity region from longitude -97.5° to -97° corresponds a broad low velocity zone in the same range. This broad low gravity and velocity region indicate a large OCGG rock unit in central Oklahoma, which is consistent with density determination from wells samples that OCGG has the lowest density of four rock units (Denison, 1981).

Cross-section A6: This cross-section crosses the MCR, southern part of the Lamont Ring Complex (R3 in Fig. 2.16), and the OCM (R4 in Fig. 2.16). Due to the east portion of this profile exceeds the reliable polygon; we only interpret the velocity as far as longitude about -96.5° to the east. Similar to the cross-section A5, a low gravity and velocity region is between the MCR and the OCM unit, which suggests a broad OCGG unit presence in this area. The gravity, magnetic, and velocity are generally consistent over this region, but the correlation between the velocity anomaly and Lamont Ring Complex is not obvious from this profile.

Cross-section A7: This cross-section crosses the MCR, northern part of the Lamont Ring Complex (R3 in Fig. 2.16), and north part of the OCM (R4 in Fig. 2.16). Similarly, we constrain our interpretation from west to the longitude -96.7°. The MCR and the intrusion around the OCM unit can be identified. This cross-section also shows a low gravity and velocity region in between, which suggest that the OCGG unit even extends northward to Kansas. The evidence of the source of the Lamont Ring Complex is still not clear from this profile.

Cross-section B1: This cross-section crosses eastern part of the SGG unit (R2 in Fig. 2.16). The resolvable range in this profile is from latitude about 35° to 36.2°. A magnetic low presents from latitude 35.5° to 36.2° and the gravity in this range is

slightly higher than that from latitude 35° to 35.5°. A high Vp anomaly exists in the magnetic low range correspondingly, which suggest an intrusion beneath the SGG unit.

Cross-section B2: The cross-section intersects the center of the SGG unit (R2 in Fig. 2.16) and western part of the OCM (R4 in Fig. 2.16). We only make the interpretation from latitude 35° to 36.7°. A strong magnetic low shows from latitude about 35.3° to 36.2°, where a high Vp anomaly body can be identified below depth ~8 km. We correlate the strong high Vp feature to the intrusion that related to the SGG unit. There is a high magnetic anomaly located north to latitude 36.5°, and the gravity is also slightly high in this range. The shallower high Vp feature at latitude 36.5° is correlated with the OCM unit. The reason that gravity above the SGG is lower than the OCM is probably due to the lower Vp rocks overlay the SGG unit and the OCM intrusion has shallower depth.

Cross-section B3: This cross-section crosses western part of the SGG unit (R2 in Fig. 2.16). A magnetic low is centered at latitude about 35.8°, where there is high Vp anomaly. The gravity decreases northward, where a low velocity zone presents. The high Vp feature can be interpreted as an intrusion that correlated with the SGG unit. The northern low Vp zone may relate to the younger COGG unit.

Cross-section B4: Southern part of this cross-section crosses the R1 magnetic high in Figure 2.16, and the northern part of this profile crosses the eastern part of the Lamont Ring Complex (R3 in Fig. 2.16). Both magnetic and gravity shows high anomalies at southern part over the high Vp anomaly area, where strongly suggests that there is an intrusion. A large low Vp zone presents from the latitude 36° to 36.6°, where a broad gravity low show. The high magnetic anomaly at latitude 36.6° represents the

Lamont Ring Complex. There is slightly high Vp anomaly from latitude 36.6° to 37°, where might be an intrusion related to the Lamont Ring Complex.

Cross-section B5: This cross-section crosses the western part of the Lamont Ring Complex (R3 in Fig. 2.16). The gravity shows slightly high anomaly at the southern part of the profile, where there is relatively high Vp anomaly. A magnetic high corresponding the Lamont Ring Complex from latitude 36.6° to 36.8° has a weak correlation with a slightly high Vp anomaly feature underneath. The weak high Vp anomaly might indicate an intrusion related to the Lamont Ring Complex.

Cross-section B6: Most part of this cross-section crosses the interpreted MCR. The magnetic anomaly does not show much variation in this profile. The gravity presents a relatively high anomaly in a broad region from latitude about 35.6° to 37°, where a broad high Vp anomaly high feature shows below depth ~8 km. There are two low gravity anomalies region on both ends, where two low Vp anomalies regions can be identified underneath.

Cross-section B7: This cross-section crosses the interpreted MCR. Due to less resolution at the southern part of this profile, we only make interpretation north to latitude 35.5°. There are two high local gravity anomalies centered at latitude 35.8° and 36.5°, respectively, where two corresponding separated high Vp anomalies can be identified underneath. The two high Vp features can be correlated to the MCR. A low Vp anomaly in between corresponds a local gravity low.

Figure 2.21 shows the SIMUL200 Vp perturbations relative to the 1D SIMUL2000 velocity model at depths 1.5 km, 5 km, 8 km, 12 km, and 15 km. The most area in the outlined polygon of the first 4 layers (1.5 km, 5 km, 8 km, and 12 km) has

the negative Vp anomaly. The low velocity area in central part of the study region may represent the board COGG rock unit. For the 8 km depth slice, a positive Vp anomaly starts to show between the NFZ and WFZ, which may relate to the intrusion that formed the SGG unit. The positive Vp anomaly becomes more obvious around the SGG unit at 12 km depth. Moreover, a high Vp anomaly presents at the south part of outline polygon, which suggests the intrusion related to the magnetic high R1 in Figure 2.16. The positive Vp anomaly is still prominent at depth 15 km. In addition, a north-south trending positive Vp anomaly presents in northwestern Oklahoma to the west of the NFZ in the 15 km slice, which strongly suggests that this feature is related to the MCR. The negative Vp anomalies around the NFZ suggest the OCGG unit.

We also show the Vp/Vs ratio anomalies of the SIMUL2000 model relative to 1.73 at depths 1.5 km, 5 km, 8 km, 12 km, and 15 km in Figure 2.22. The most area in the 1.5 km slice shows positive Vp/Vs ratio anomalies. The slice at depth 5 km starts to present some local negative Vp/Vs anomalies. The negative Vp/Vs variation at 8 km slice becomes prominent, but a relative narrow positive Vp/Vs ratio region presents at the SGG unit region. The positive Vp/Vs ratio is still prominent at 12 km and 15 km depth slices, respectively. Furthermore, by examining the variation of the Vp/Vs ratio in the 14 cross-sections, most of the profiles present a low Vp/Vs ratio region in upper crust (Figs. 2.17 and 2.18). We estimated the low Vp/Vs ratio zone roughly from depth ~3 – 10 km. The low Vp/Vs ratio zone also can be identified from the averaged Vp/Vs ratio variation in Figures 2.19 and 2.20 as well. The averaged basement top is ~3 km in Oklahoma. Therefore, the Vp/Vs ratio can be simplified to three layers: <3 km, ~3 to 10

km, and >10 km, although there is a strong variation of Vp/Vs ratio in some cross-sections.

The study from Lin et al. (2009) shows that low Vp/Vs ratio zone in the earthquake source regions relates to water-filled fractures or cracks. A number of studies support that the earthquakes in Oklahoma have a strong relationship with the deep injection of wastewater (e.g., Holland, 2013; Keranen et al., 2013, 2014; Ellsworth, 2013; Darold et al., 2014, 2015). A huge amount of fluids, including wastewater and brine, has been injected into the Arbuckle Group, which has little or no Cambrian sedimentary rock between it and the Precambrian basement in most of Oklahoma (e.g., Murray, 2014; Morgan and Murray, 2015; Weingarten et al., 2015; Walsh and Zoback, 2015). Due to the complex fault and fracture structures in the basement, it is highly possible that the wastewater can penetrate into the deep basement, even down to 10 km or more. We think the water-filled fractures can be used to explain the low Vp/Vs ratio zone in the upper basement of central Oklahoma.

Conclusions

We created two 3D velocity models for central Oklahoma with the FMTOMO and the SIMUL2000 methods, respectively. These two models present similar velocity features in a large scale. The FMOTMO model shows more small-scale velocity features since more earthquakes and dense grids were used in the inversion. The SIMUL2000 model shows the similar large features as the FMTOMO model, but has less model variation, which is due to fewer data and grids were used in the inversion. We use the SIMUL2000 model, gravity, and magnetic data to correlate the previously known geological features in central Oklahoma. The velocity model shows high Vp

regions where are generally consistent with gravity and magnetic profiles in the study area. The SGG, the OCM, and the COGG rock unit can be identified through our velocity model. The SGG region is underlain by an intrusion. The reason that this high Vp feature is not obvious from the gravity data is probably due to a low Vp rhyolite cover overlay the intrusion. The OCM or WCVG unit always causes the high gravity and magnetic anomalies in the Osage County, where is located at the northeast of our study region. All the profiles and depth slices including this unit will present high Vp anomalies, which are consistent with gravity and magnetic observation. From the velocity model, we think that the OCM intrusion has the shallower depth than the SGG intrusion. A broad region in central part of the study area shows low Vp anomalies near the NFZ. We interpret this feature as the younger OCGG unit. The drilled sample from the OCGG area has the smallest measured density among all the units in central Oklahoma (Denison, 1981), which can support our interpretation. We also identified an intrusion (R1 in Fig. 2.16) from the velocity model between the NFZ and WFZ at the southern part of the study region. No too much previous study was conducted on this high magnetic feature. There is only a weak correlation between the magnetic and velocity anomalies for the Lamont Ring Complex (R3 in Fig. 2.16). The previous studies suggest that this feature is probably a Precambrian meteorite impact, shield volcano, or from basement geology (Luza and Lawson, 1982). However, the source of the Lamont Ring Complex is still unclear and poorly studied. We do not find strong velocity evidence for this feature. Another important tectonic feature, the MCR, can be identified from velocity model. We think that the MCR might be relatively deep since this feature does not become continuous until 15 km.

Therefore, most of the velocity anomalies can correlate to the known and inferred features, such as igneous intrusion related to the OCM, the SGG, and magnetic high (R2 in Fig. 2.16). The OCGG unit is widely distributed in north central Oklahoma. The MCR is also obvious from the velocity model. In addition, the Vp/Vs ratio presents a high-low-high distribution along the depth. The simplified low Vp/Vs ratio zone is in the upper basement for depth from ~3 km to 10 km, which may relate to the water-filled fractures in the host rock.

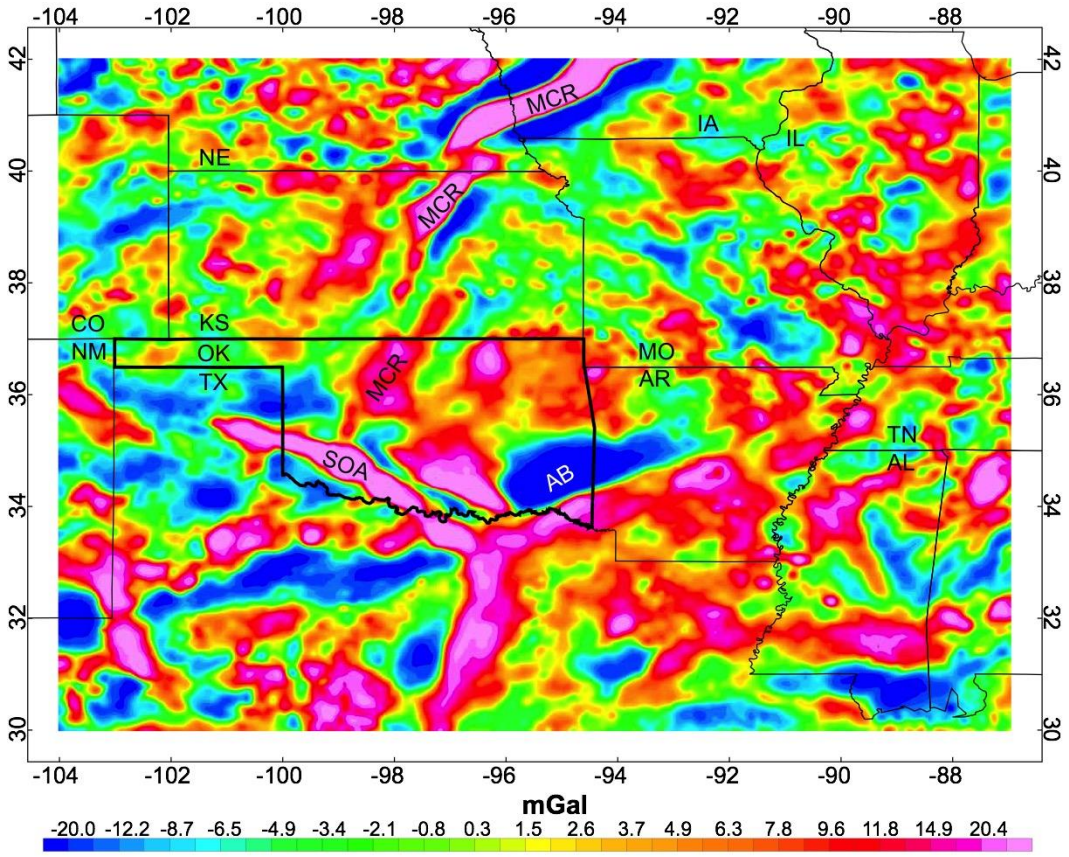


Figure 2.1. Gravity map showing the speculated west rift arm of Mid-Continent Rift (MCR) extending into Oklahoma. This image is computed by upward continuing the gridded complete Bouguer anomaly (CBA) data to 40 km and subtracting the resulting values from the gridded CBA values. Modified from Stein et al., (2014). SOA: Southern Oklahoma Aulacogen; AB: Arkoma Basin.

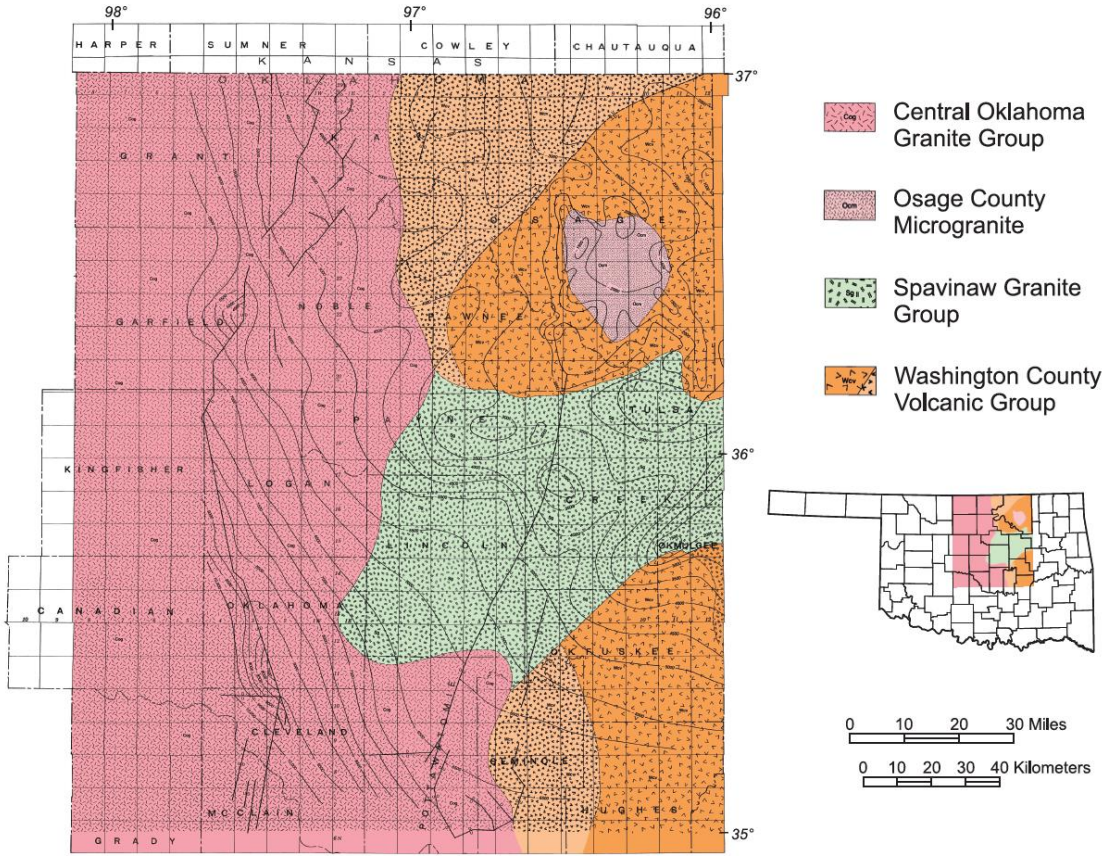


Figure 2.2. Basement rock unit map in north central Oklahoma (modified from Denison (1981) and Luza and Lawson (1982) by the Oklahoma Geological Survey).

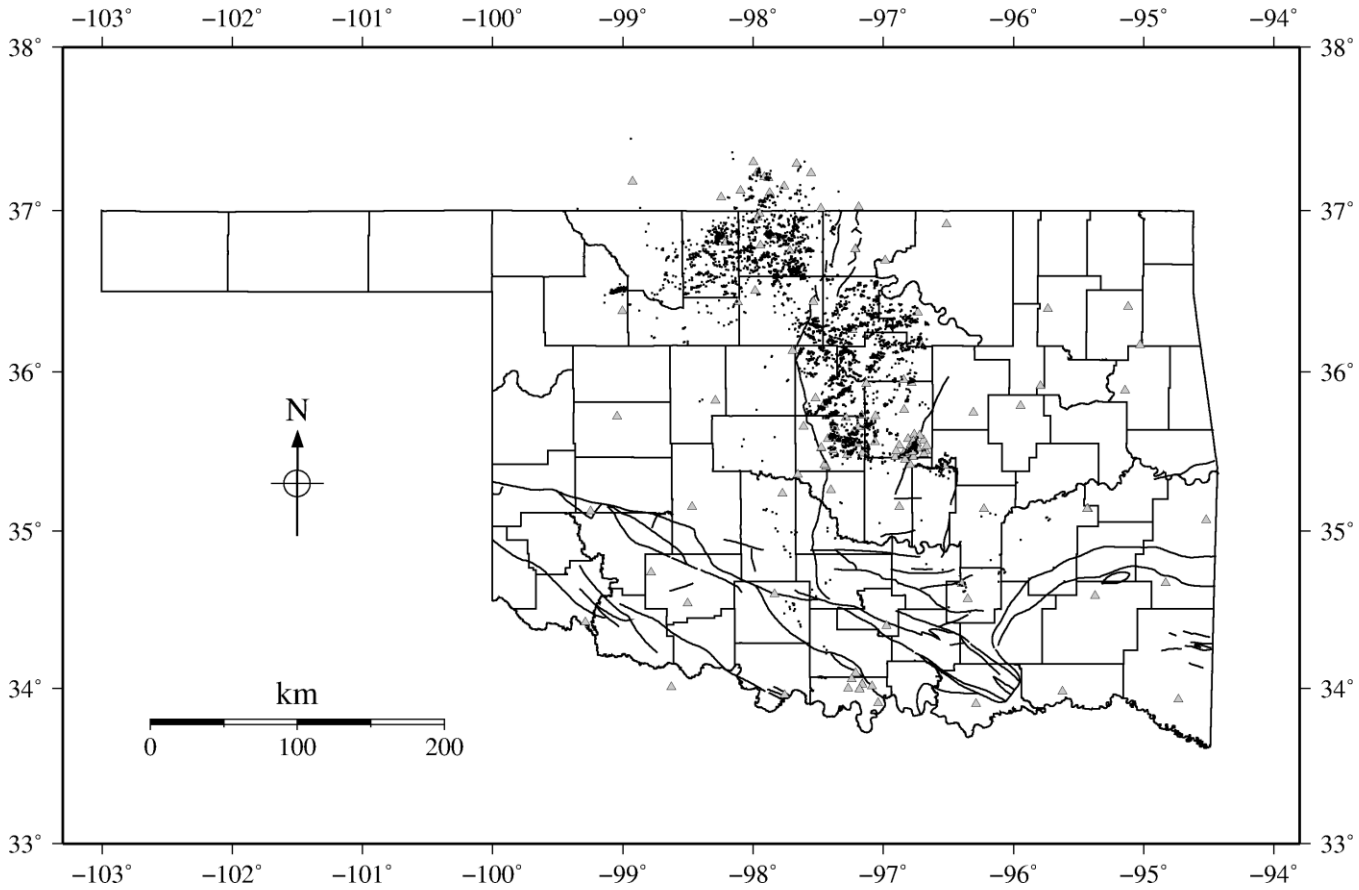


Figure 2.3. The 8194 M2+ earthquakes (01/01/2010 - 07/31/2015) were selected to determine the velocity model with FMTOMO. Black polygons are county boundaries, and black lines are major faults in Oklahoma. Gray triangles represent the stations.

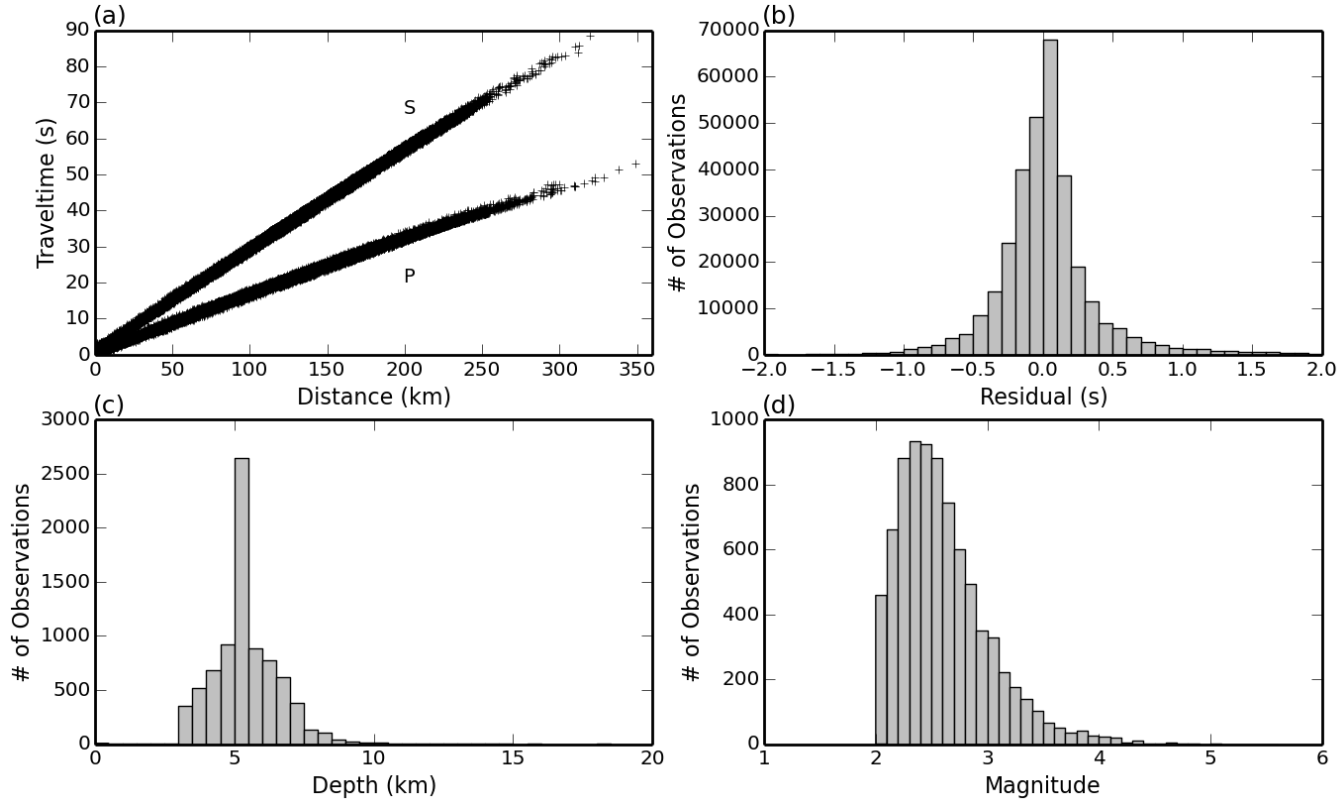


Figure 2.4. Statistics of selected data after filtering. (a) Filtered P- and S-picks travel-time vs distance. (b) Histogram of residuals. (c) Histogram of focal depth. (d) Histogram of magnitude.

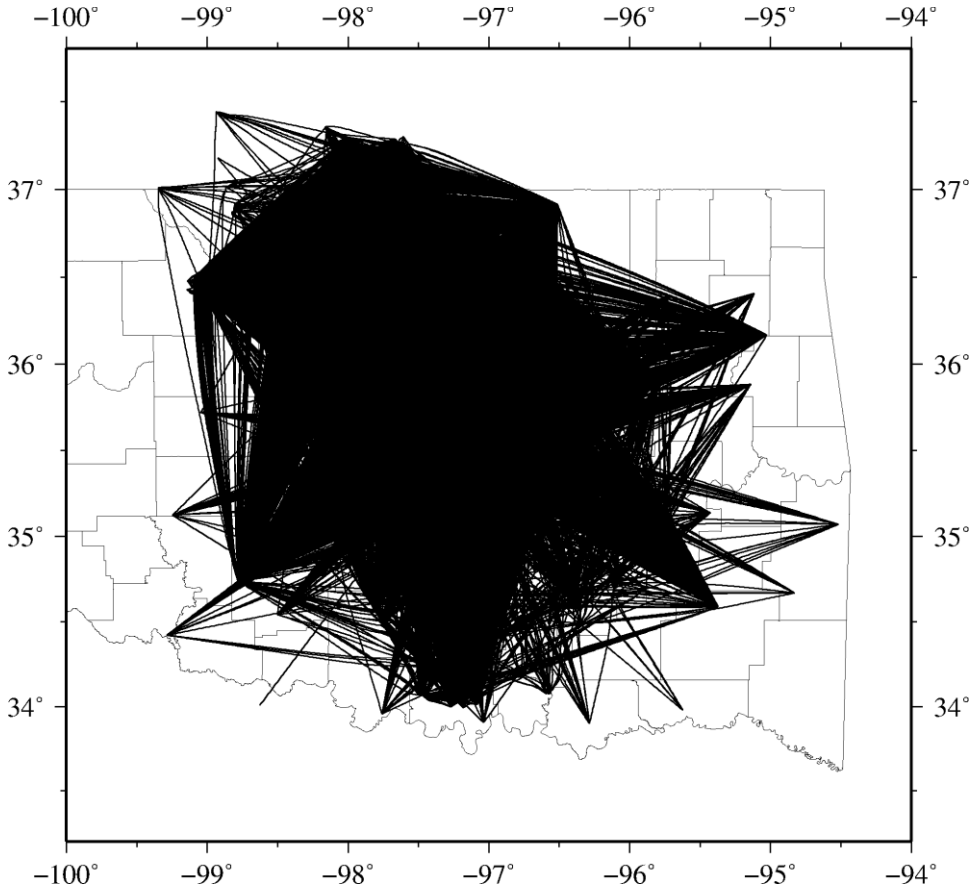


Figure 2.5. P-wave ray paths coverage in central Oklahoma. Gray polygons are county boundaries.

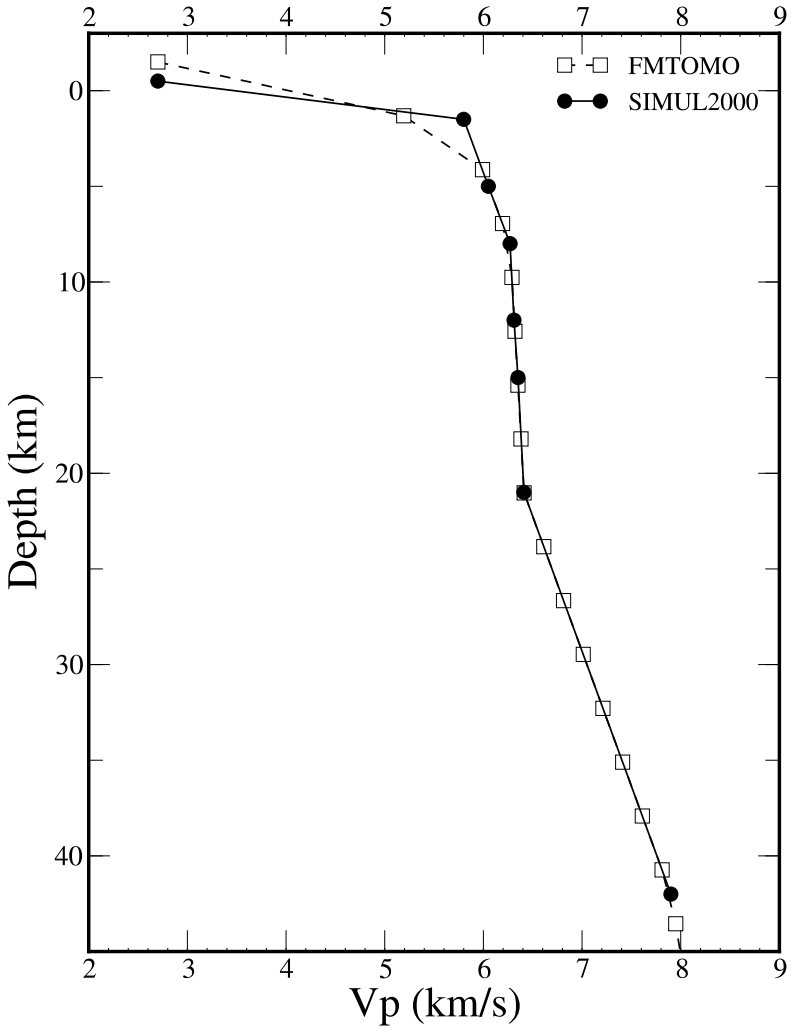


Figure 2.6. 1D velocity models for the FMTOMO (squares) and the SIMUL2000 (dots), respectively.

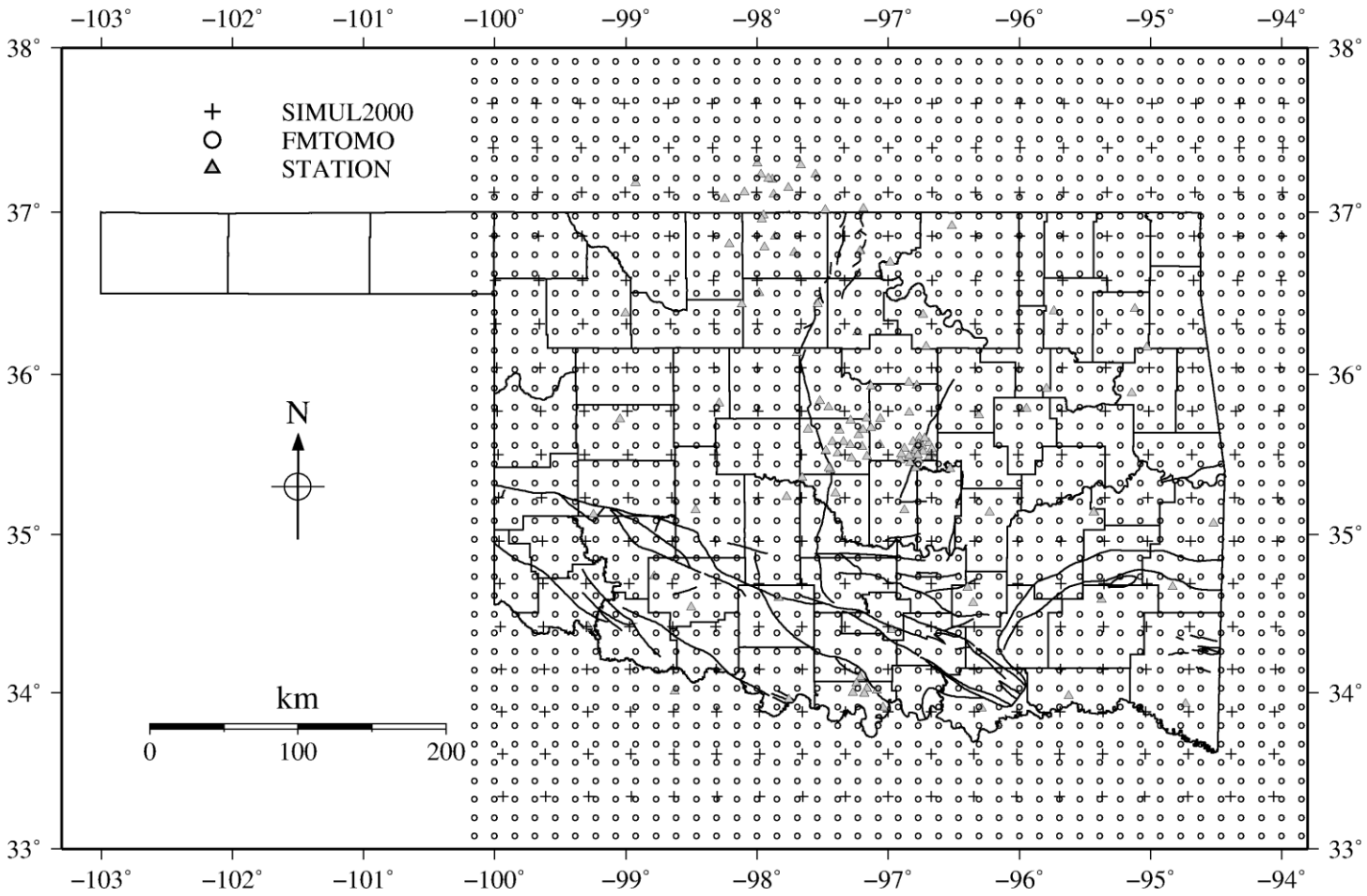
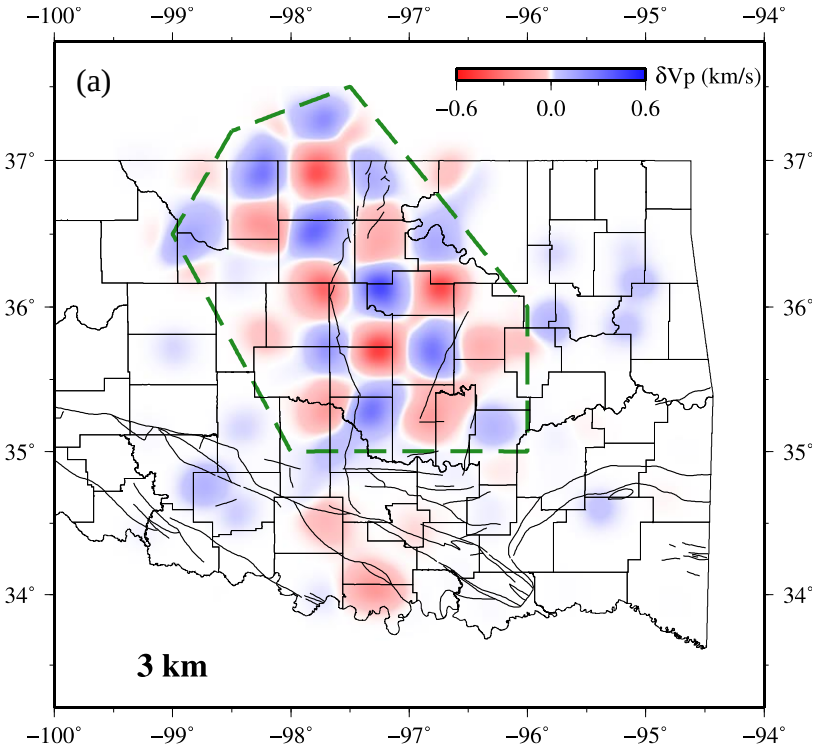


Figure 2.7. Grid configuration for the FMTOMO (open circles) and the SIMUL2000 (plus signs) methods. Gray triangles represent the stations. Black polygons are county boundaries, and black lines are major faults in Oklahoma.


(a)
δVp (km/s)
−0.6
0.0
0.6
−100°
−99°
−98°
−97°
−96°
−95°
−94°
37°
36°
35°
34°
3 km

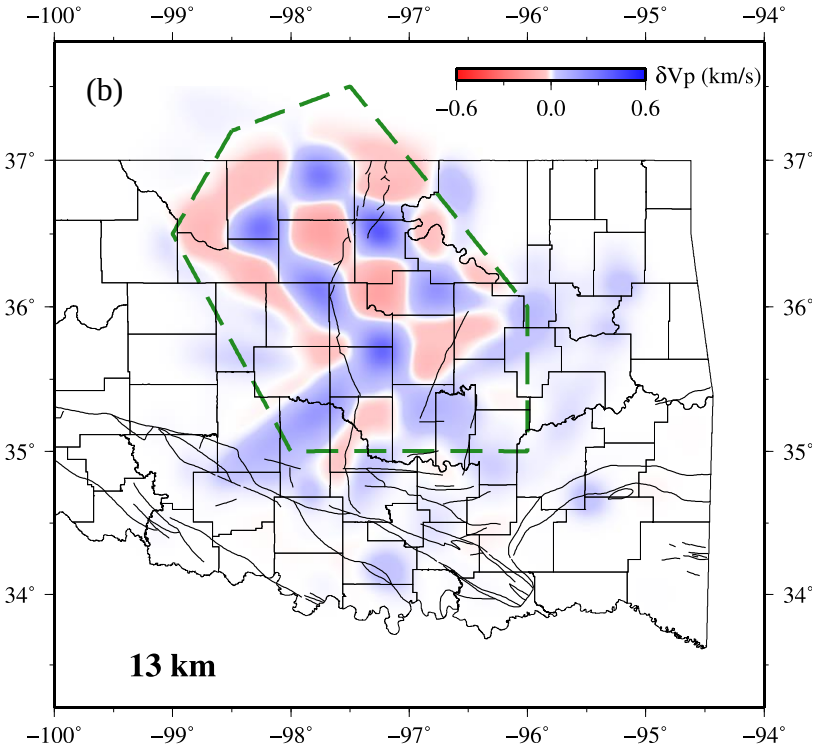
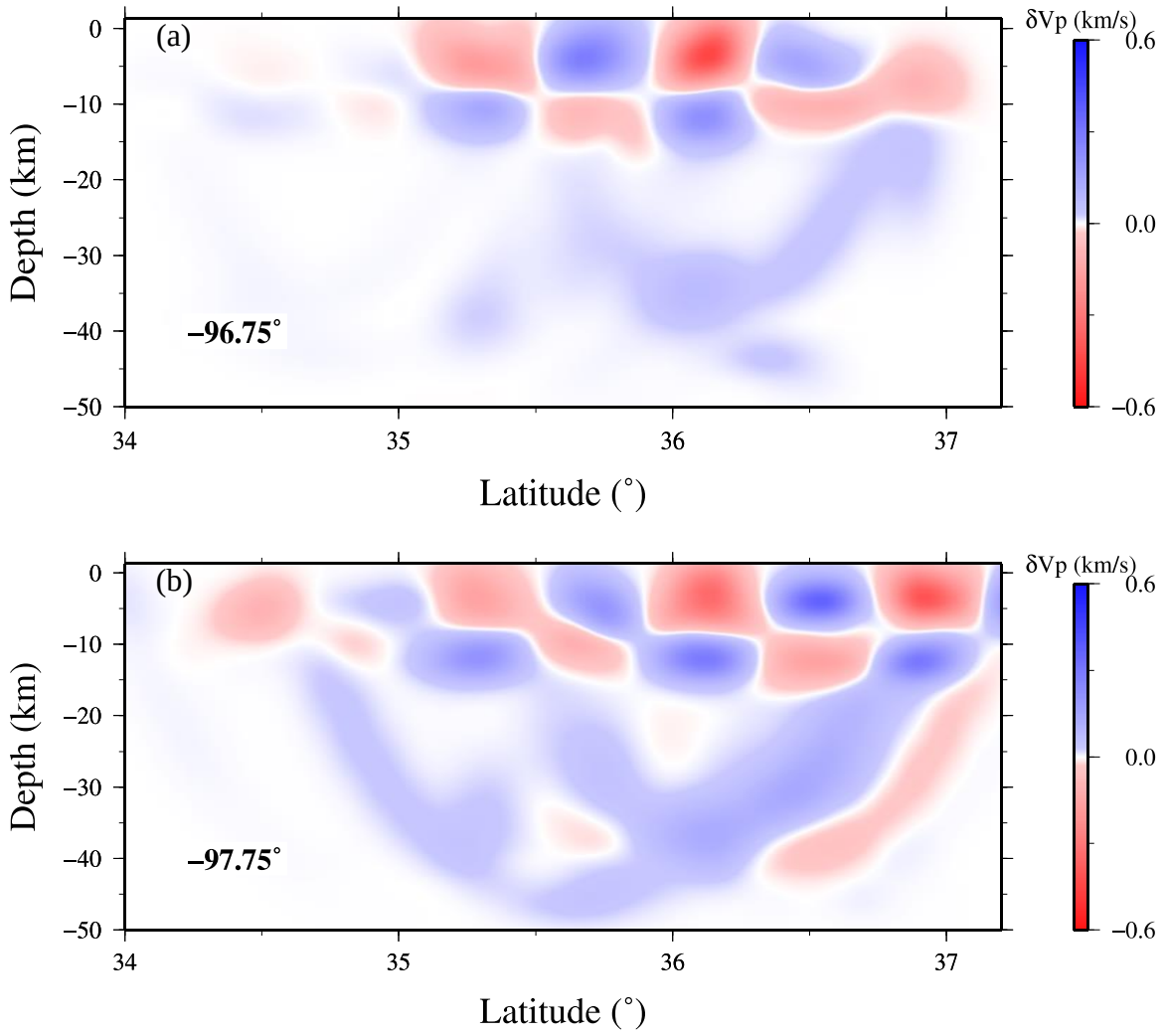

(b)
δVp (km/s)
−0.6
0.0
0.6
−100°
−99°
−98°
−97°
−96°
−95°
−94°
37°
36°
35°
34°
13 km

Figure 2.8. Checkerboard test of Vp depth slices at 3 km (a) and 13 km (b). Green polygon indicates the region with good resolution, gray polygons are county boundaries, and black lines are major faults in Oklahoma.


(a)
−96.75°
Depth (km)
0
−10
−20
−30
−40
−50
34
35
36
37
Latitude (°)
δVp (km/s)
0.6
0.0
−0.6
(b)
−97.75°
Depth (km)
0
−10
−20
−30
−40
−50
34
35
36
37
Latitude (°)
δVp (km/s)
0.6
0.0
−0.6

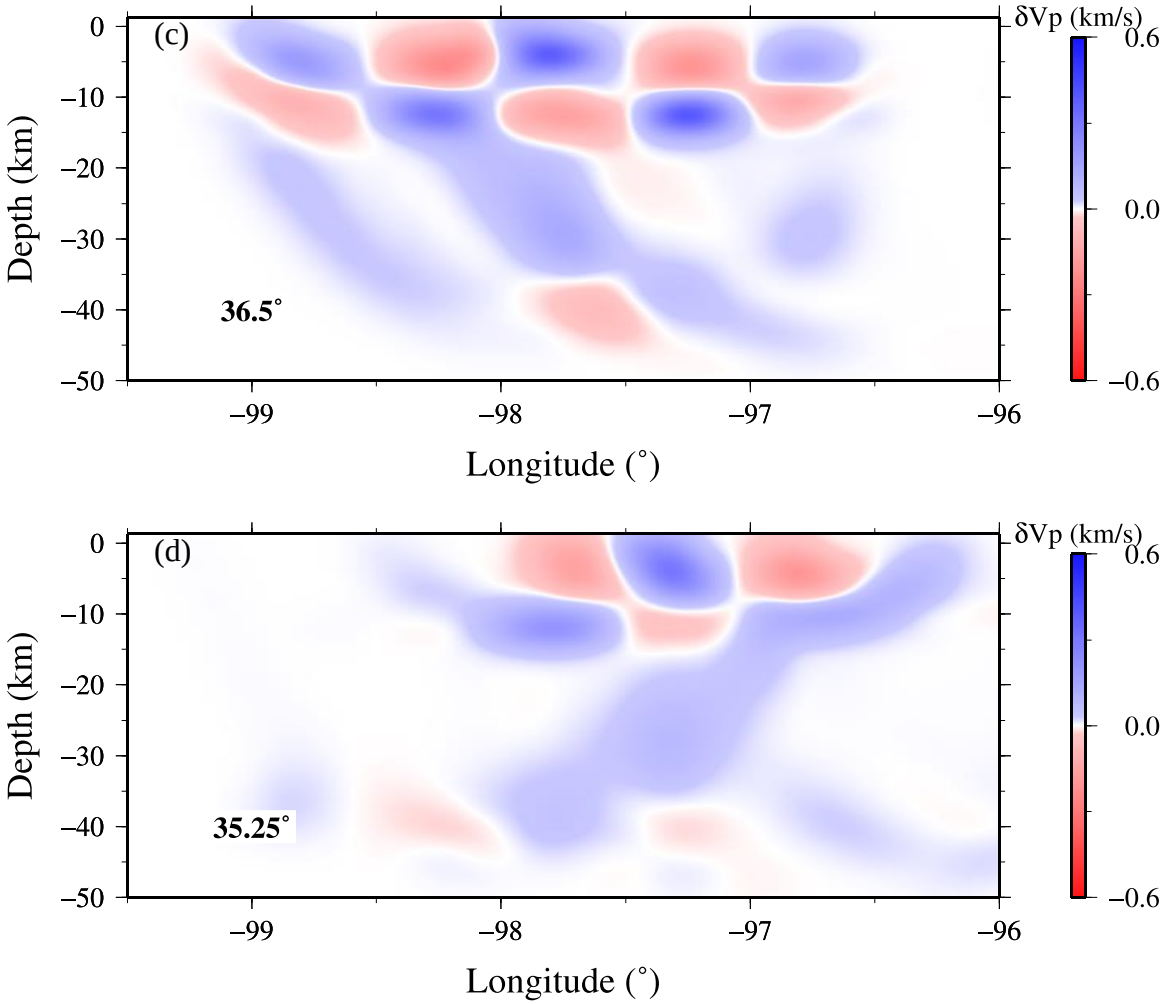

(c)
36.5°
Depth (km)
0
−10
−20
−30
−40
−50
−99
−98
−97
−96
Longitude (°)
δVp (km/s)
0.6
0.0
−0.6
(d)
35.25°
Depth (km)
0
−10
−20
−30
−40
−50
−99
−98
−97
−96
Longitude (°)
δVp (km/s)
0.6
0.0
−0.6

Figure 2.9. Checkerboard test of SN and WE cross-section for the Vp model. (a)-(d). The longitude or latitude of each slice is indicated at lower left corner.

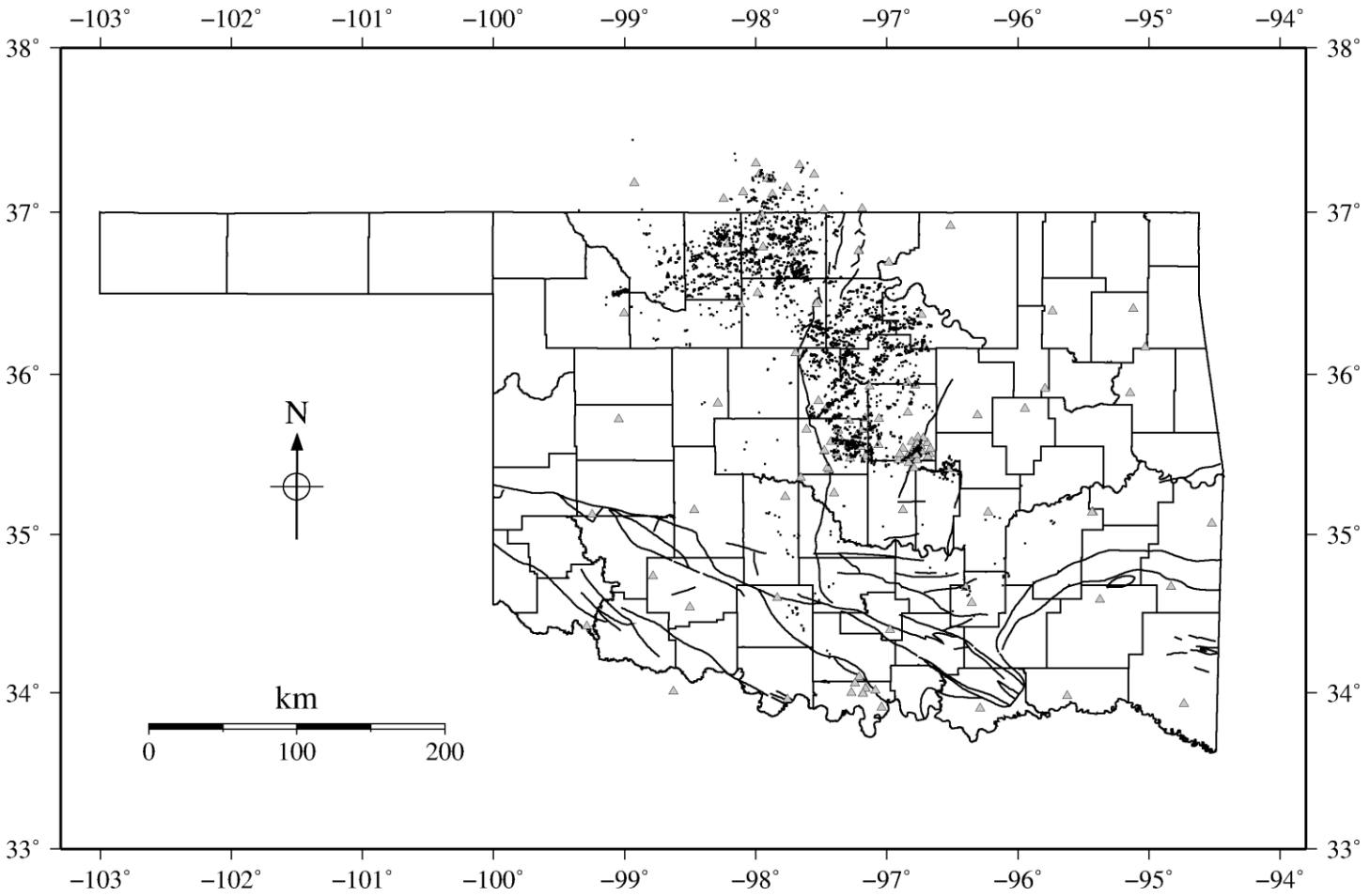


Figure 2.10. The relocated M2+ events (dots) with HYPODD 1D method. Black polygons are county boundaries, and black lines are major faults in Oklahoma. Gray triangles represent the stations.

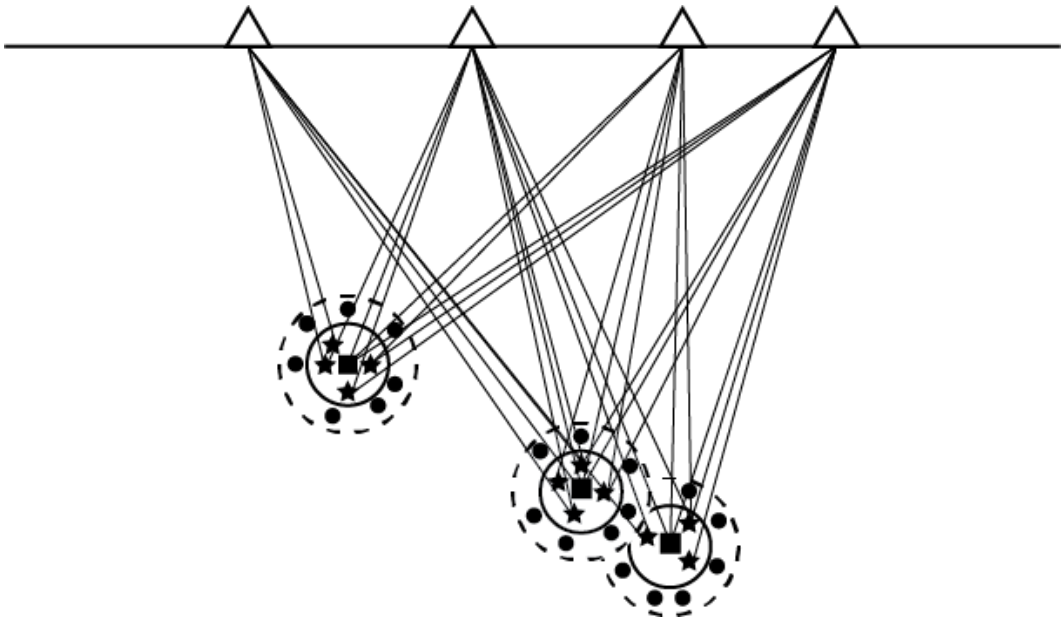


Figure 2.11. Illustration of the algorithm of constructing the composite events. Squares are the target master events. Stars in inner solid circle are the close events that contribute pick information to the composite events, and the dots between inner solid circles and outside dashed circles represent the events excluded for the future master event candidates. The triangles are the stations.

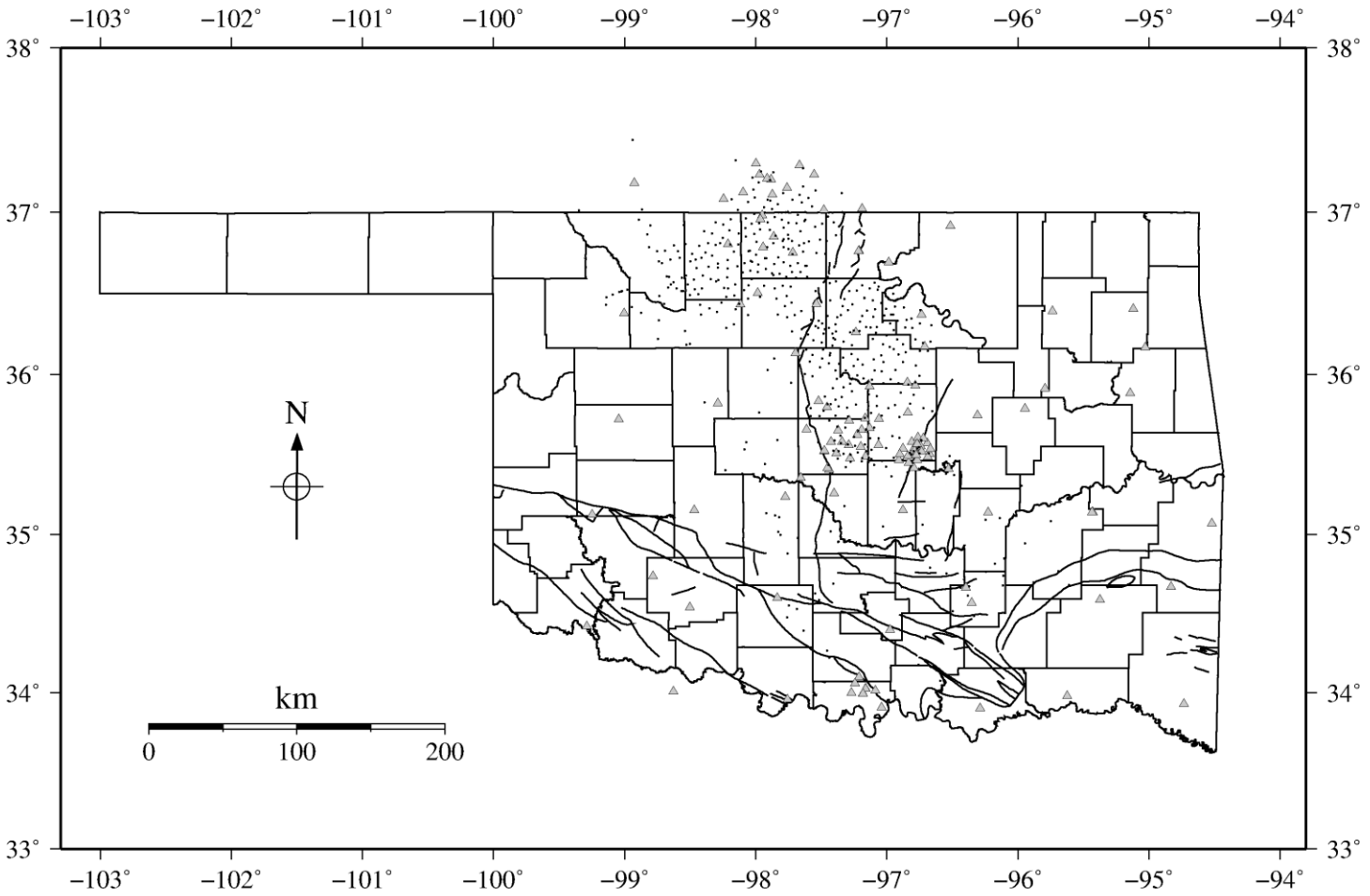


Figure 2.12. The composite events (dots) have relatively uniform distribution. Black polygons are county boundaries, and black lines are major faults in Oklahoma. Gray triangles represent the stations.

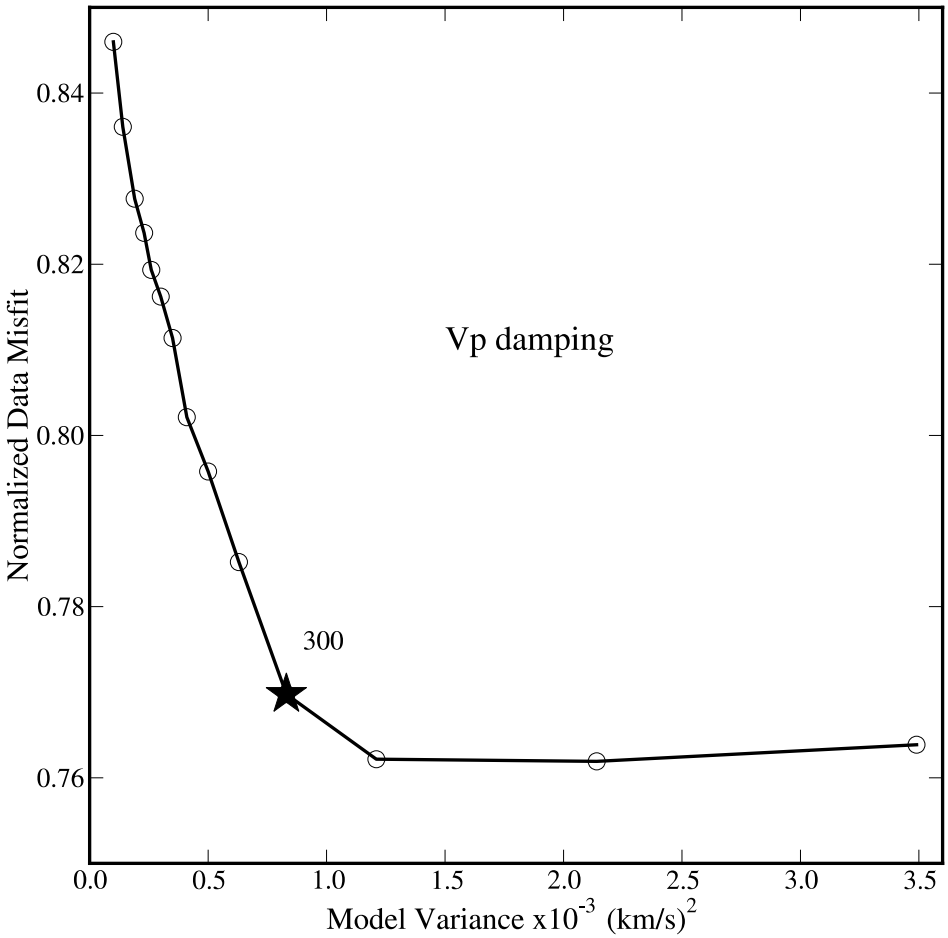


Figure 2.13. Tradeoff curve of Vp between model variance and data misfit.

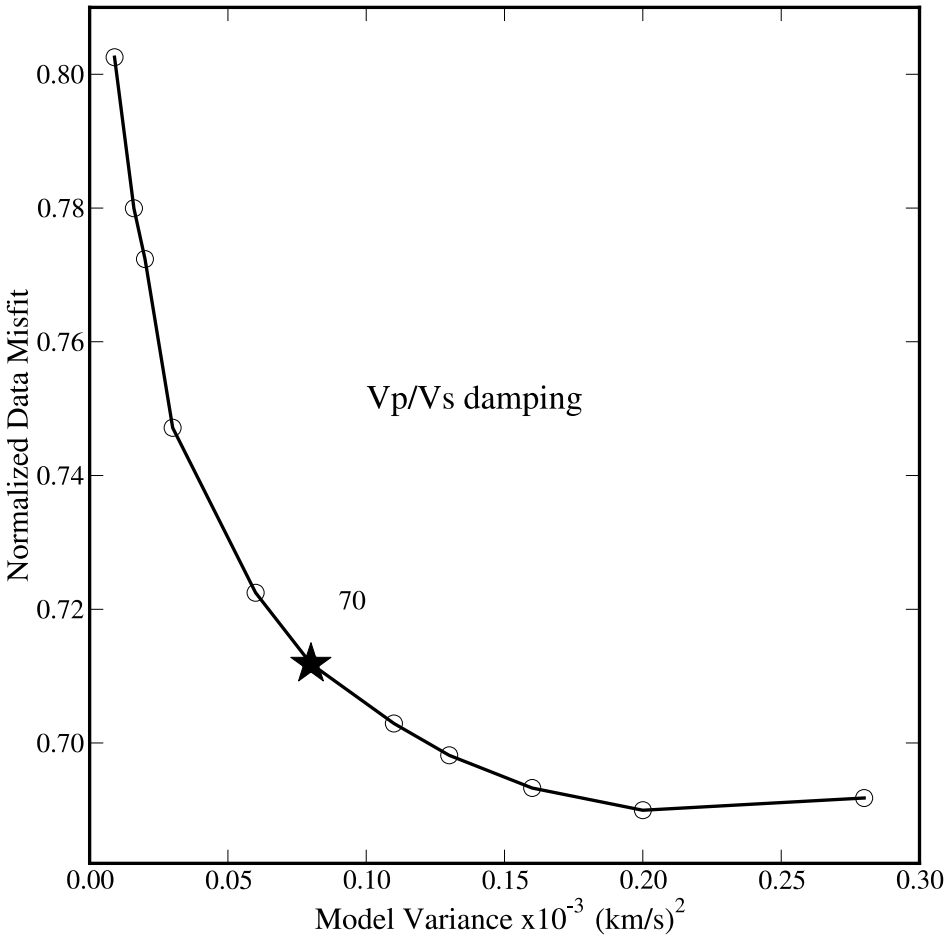


Figure 2.14. Tradeoff curve of Vp/Vs ratio between model variance and data misfit.

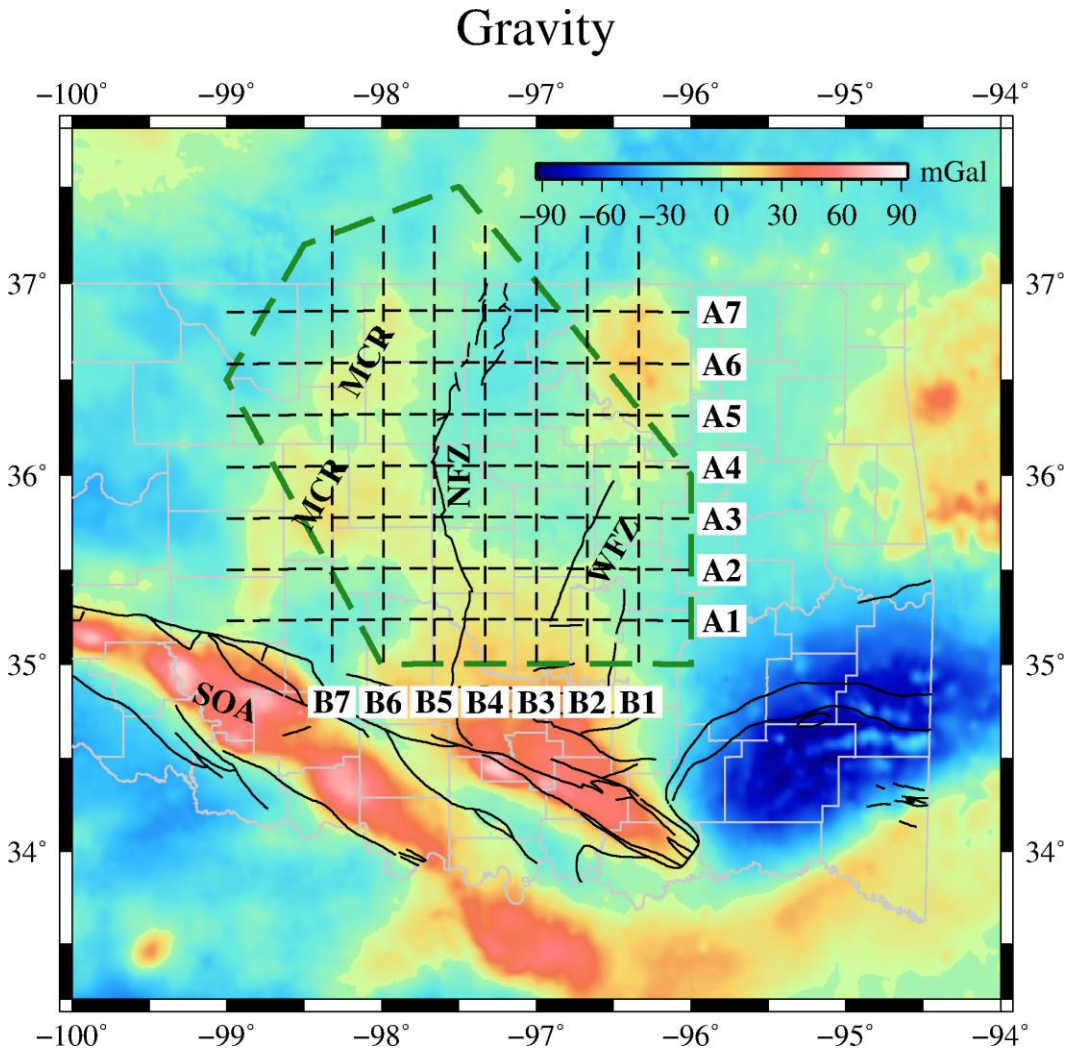


Figure 2.15. Free air gravity anomaly map of central Oklahoma. The data was downloaded from Pan-American Center for Earth and Environmental Studies (PACES). Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution. The dashed lines indicate the positions of the vertical cross-sections shown in Figures 16 and 17. MCR: Mid-Continent Rift; NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone; SOA: Southern Oklahoma Aulacogen.

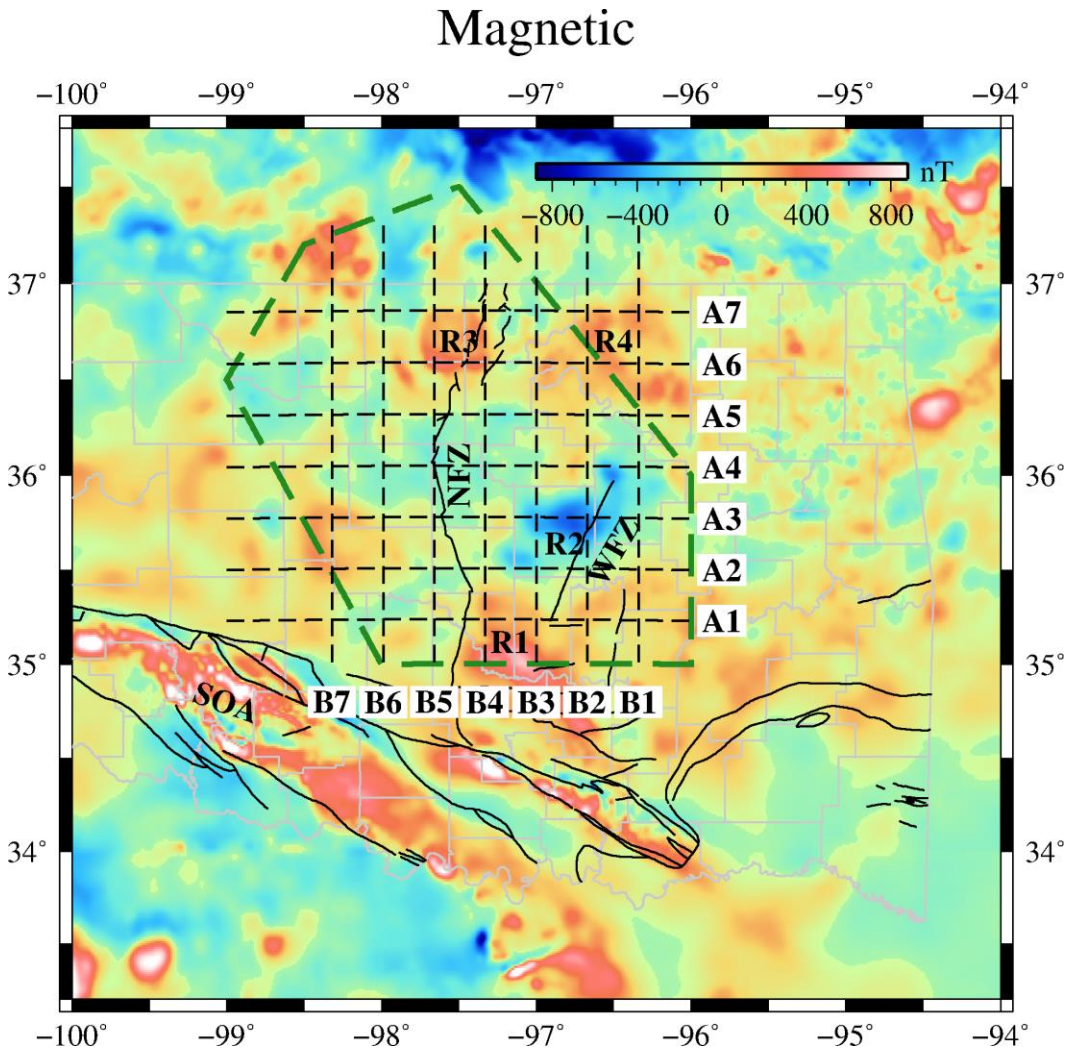


Figure 2.16. Reduced to the pole magnetic anomaly map. The data was downloaded from Pan-American Center for Earth and Environmental Studies (PACES). Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution. The dashed lines indicate the positions of the vertical cross-sections shown in Figures 16 and 17. NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone; SOA: Southern Oklahoma Aulacogen.

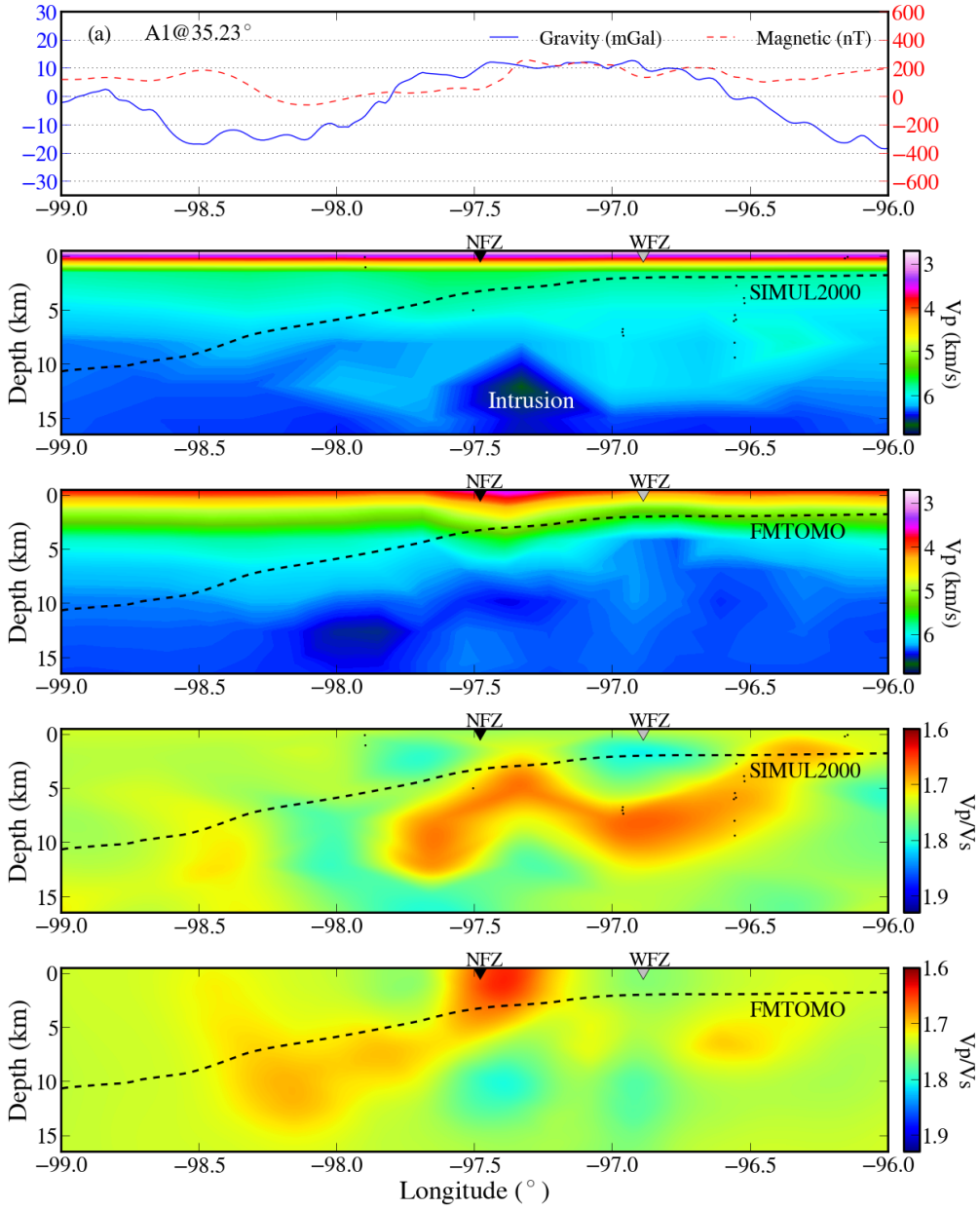


Figure 2.17a. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A1 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

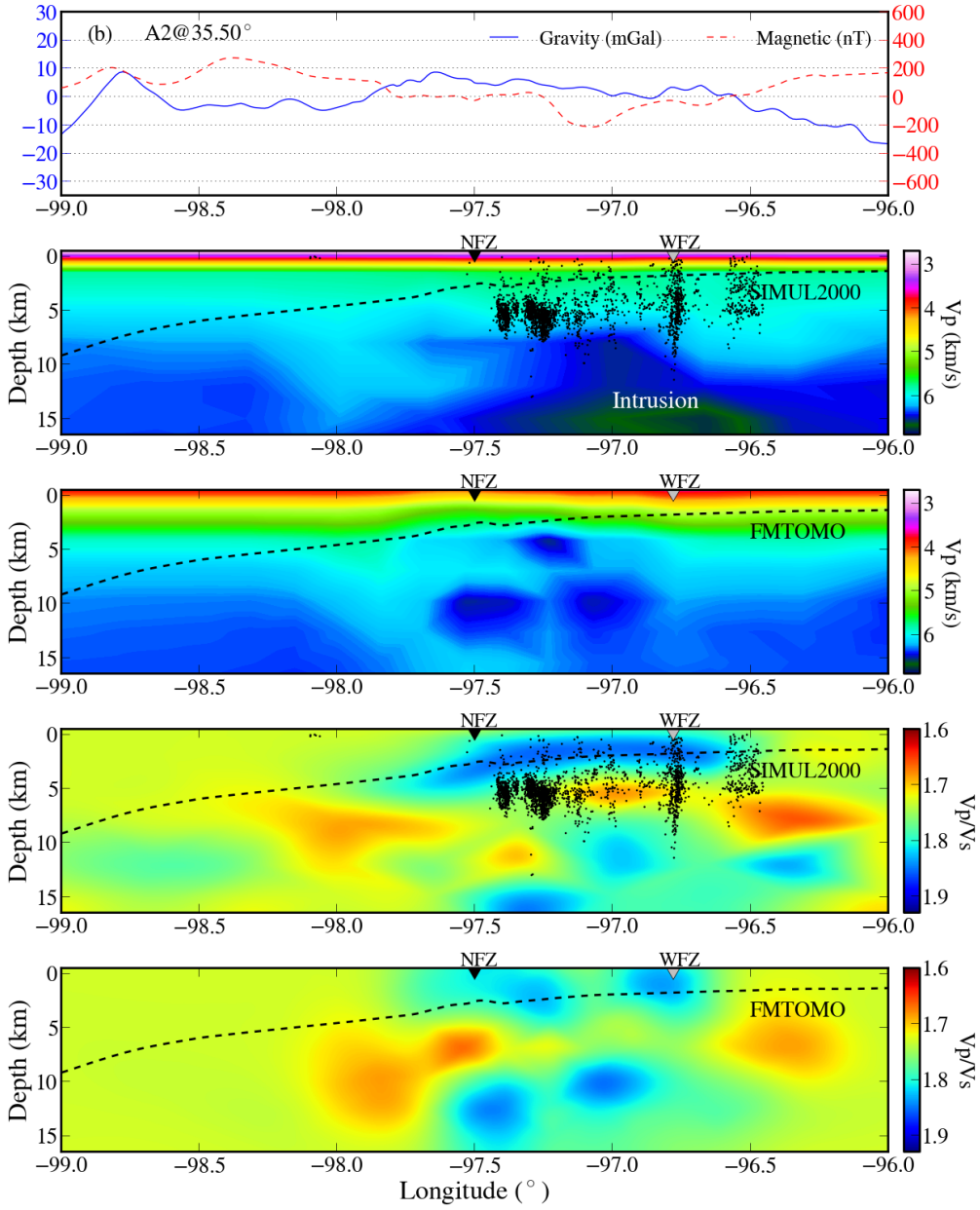


Figure 2.17b. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A2 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

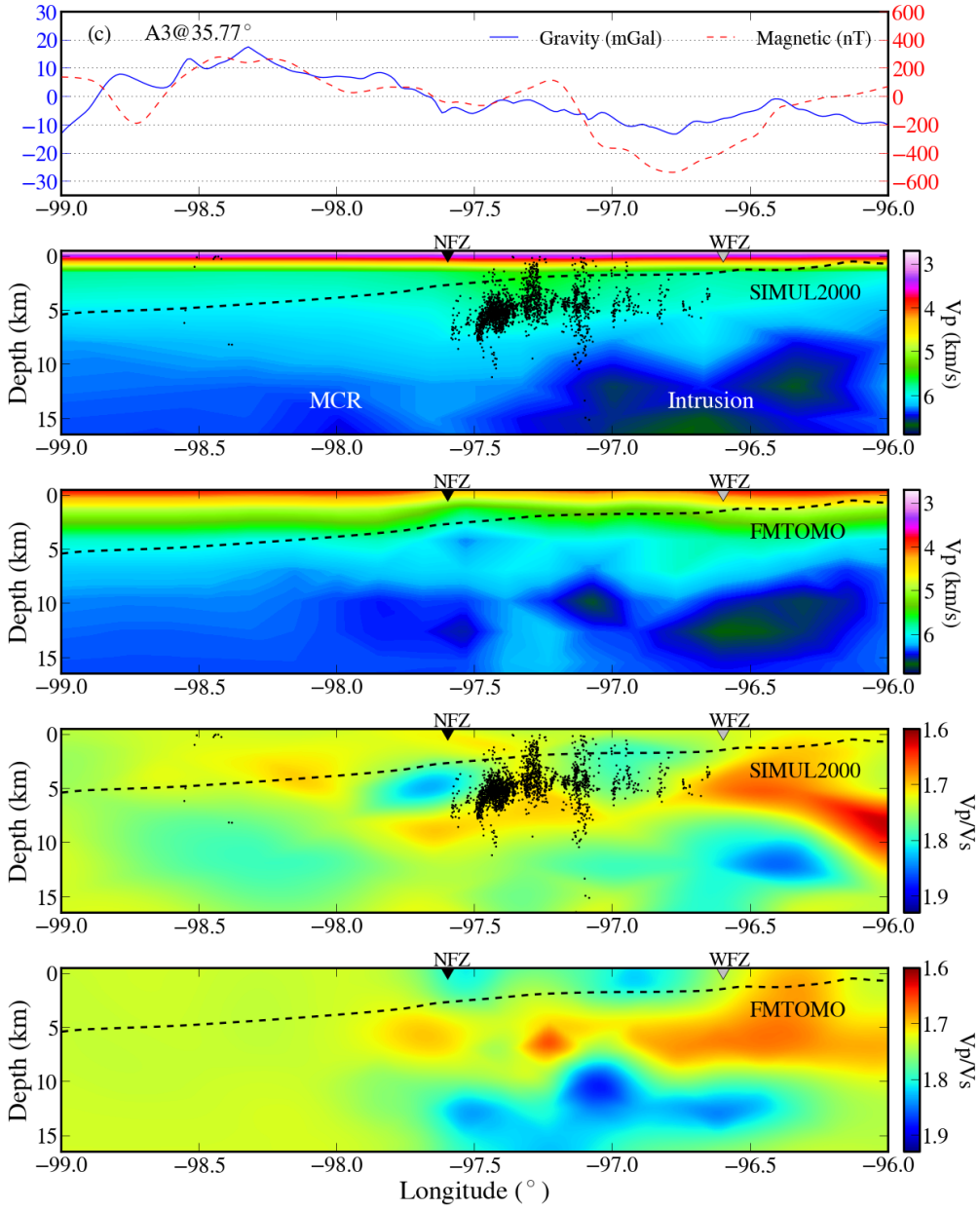


Figure 2.17c. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A3 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

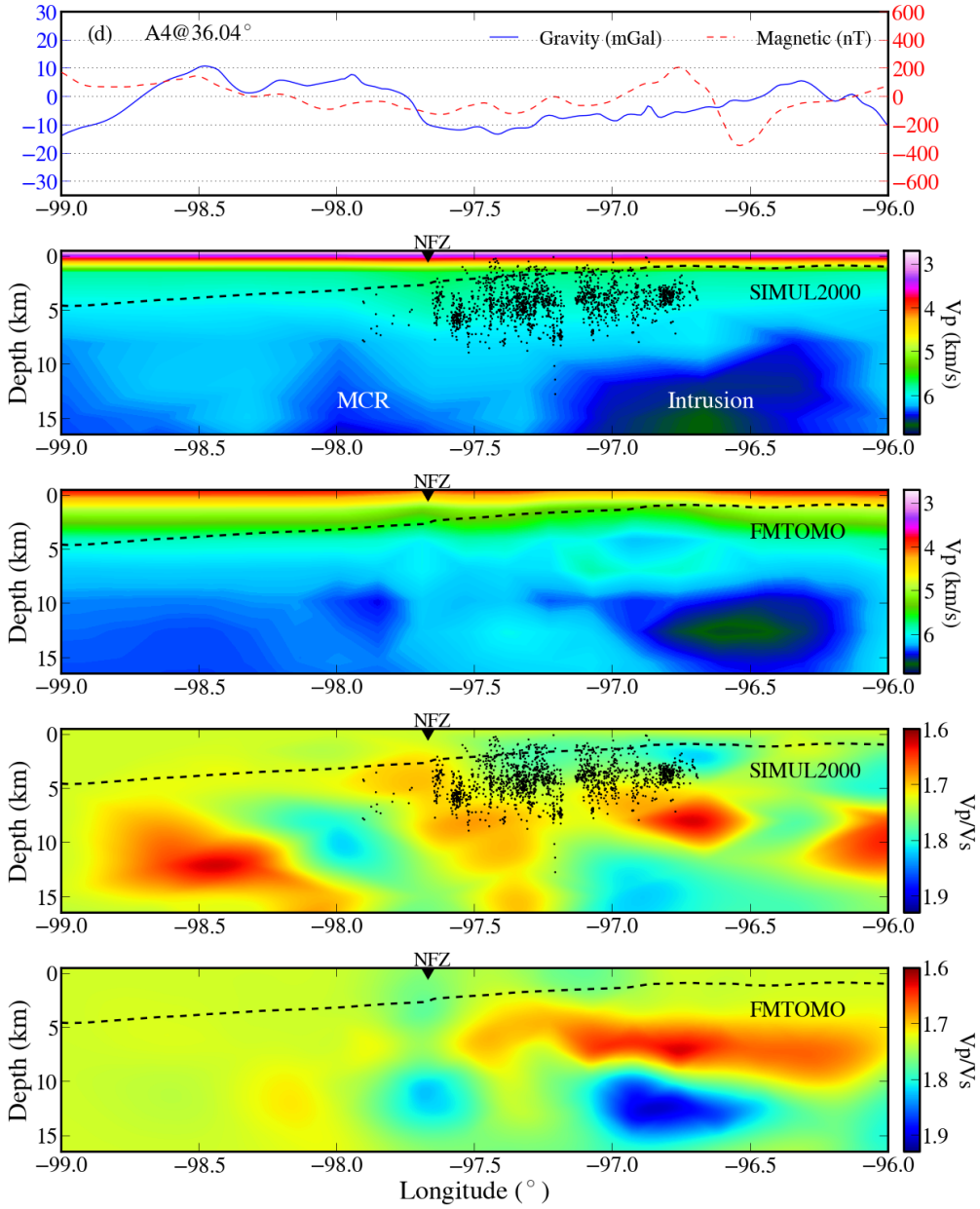


Figure 2.17d. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A4 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

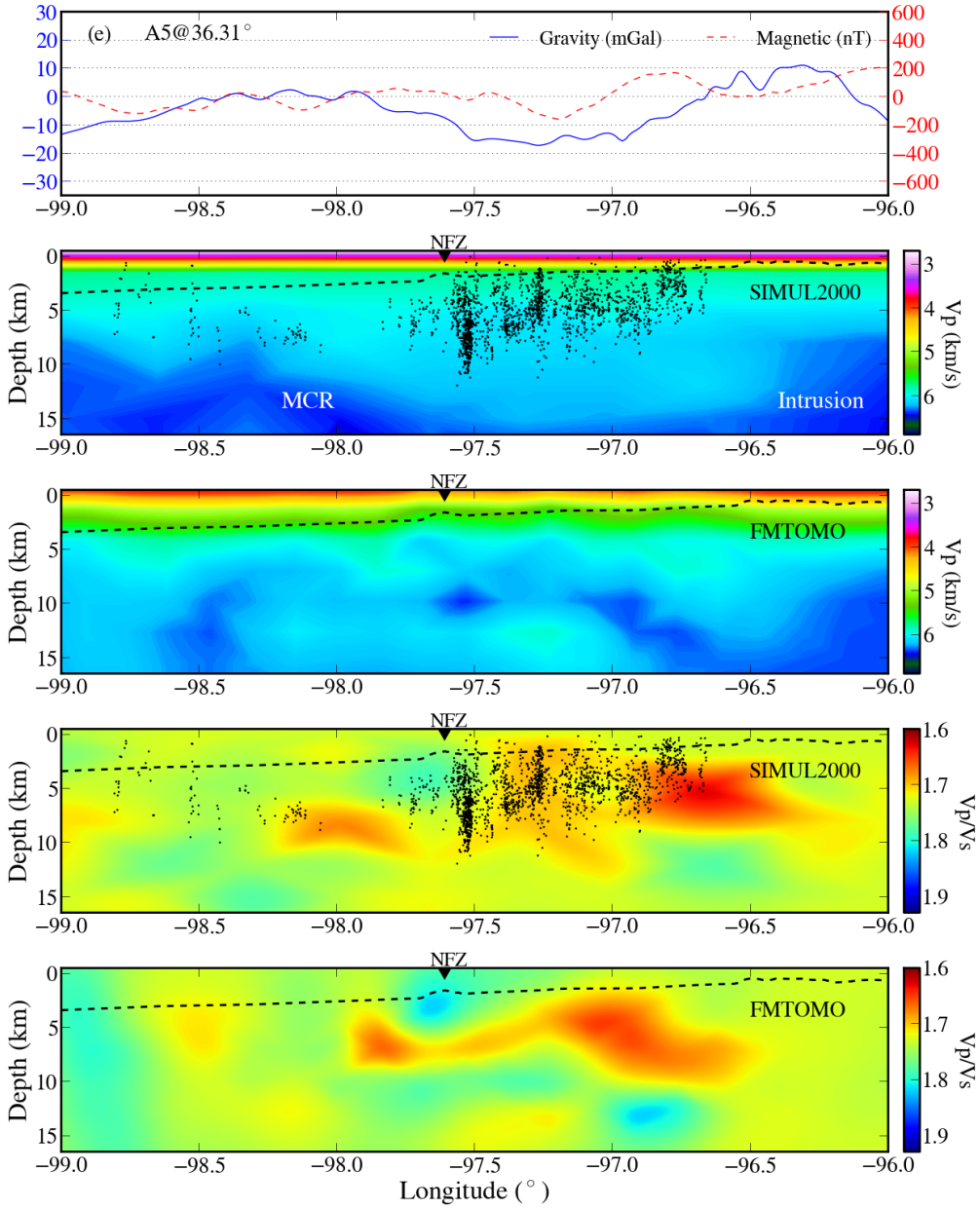


Figure 2.17e. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A5 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

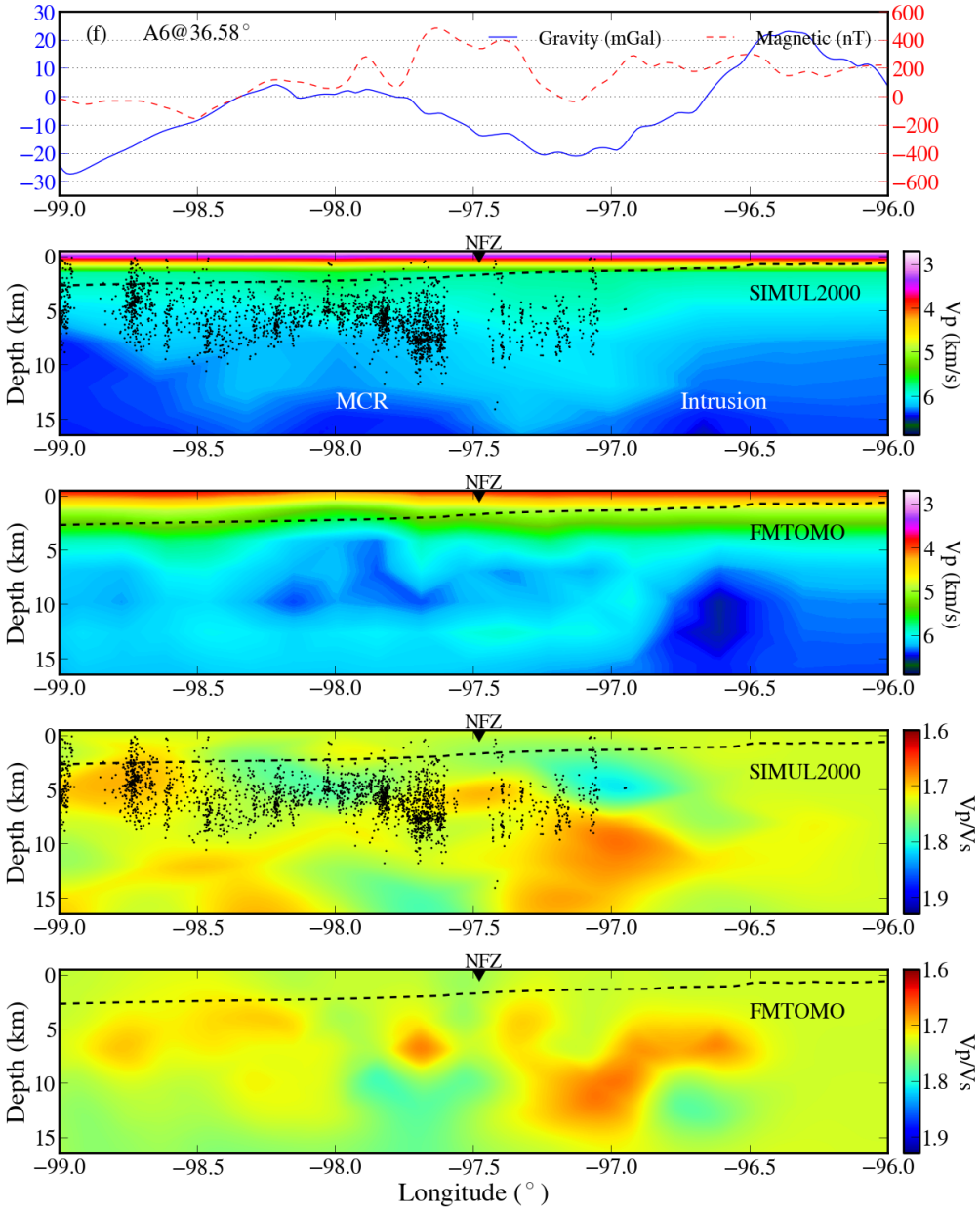


Figure 2.17f. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A6 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

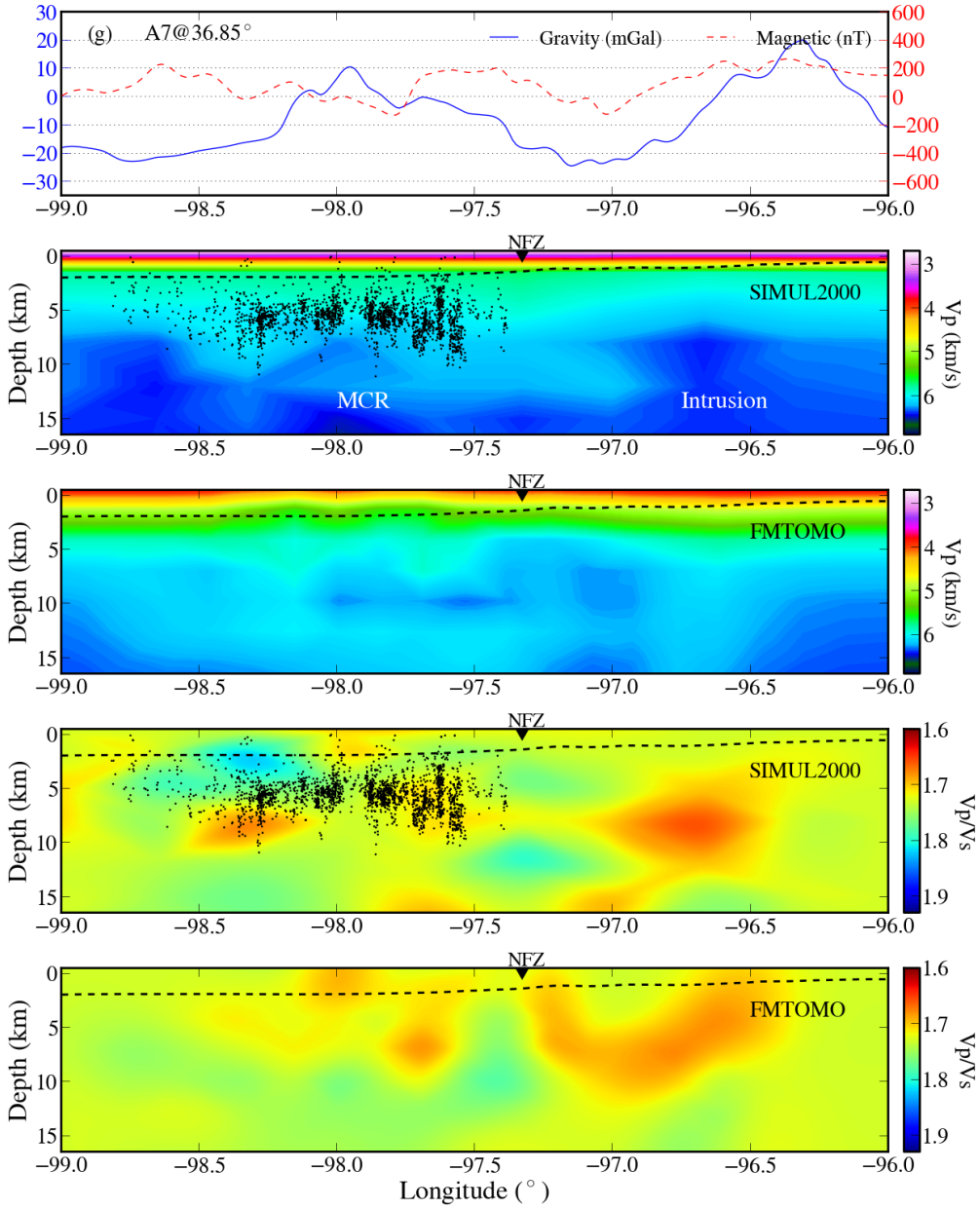


Figure 2.17g. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile A7 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

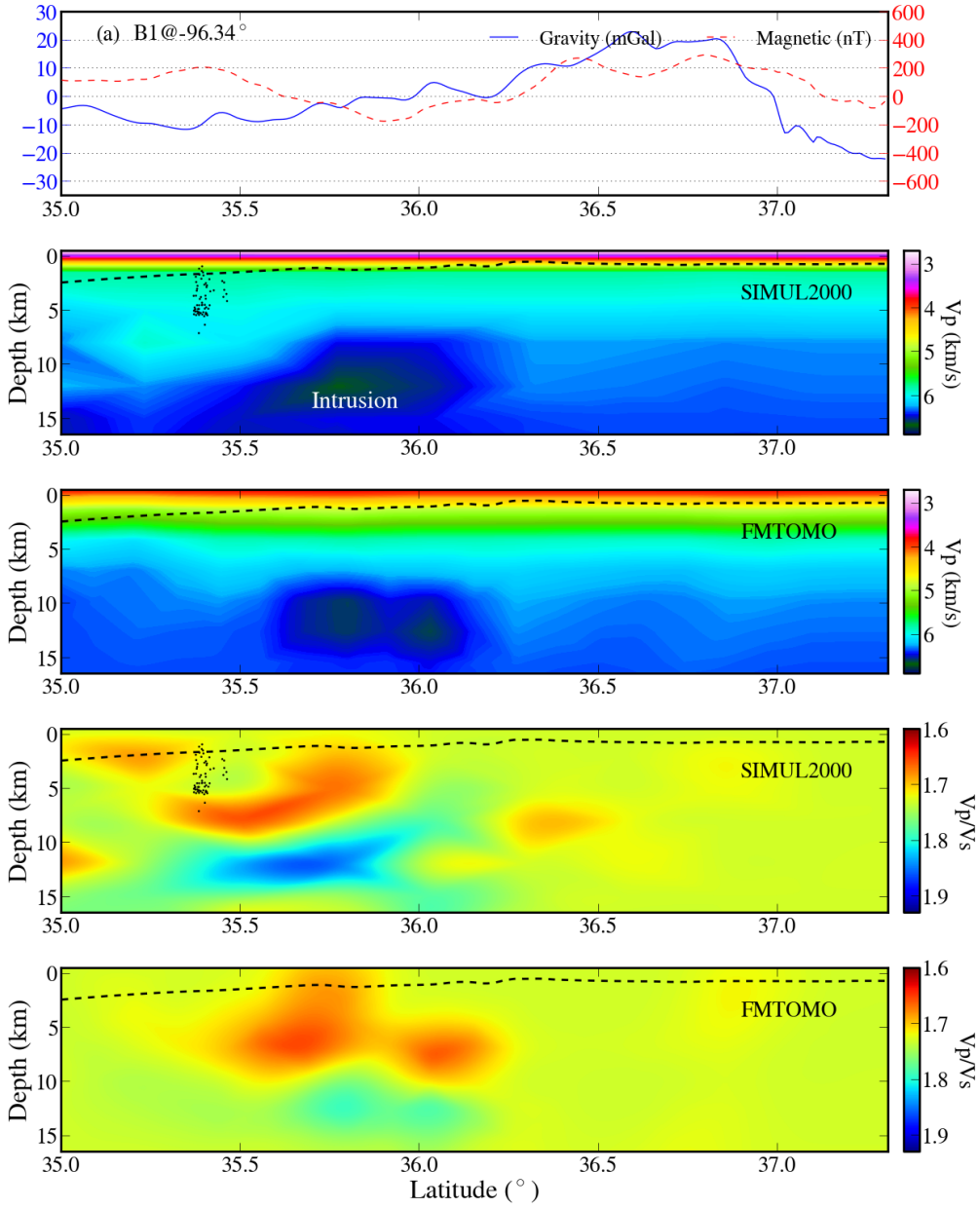


Figure 2.18a. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B1 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

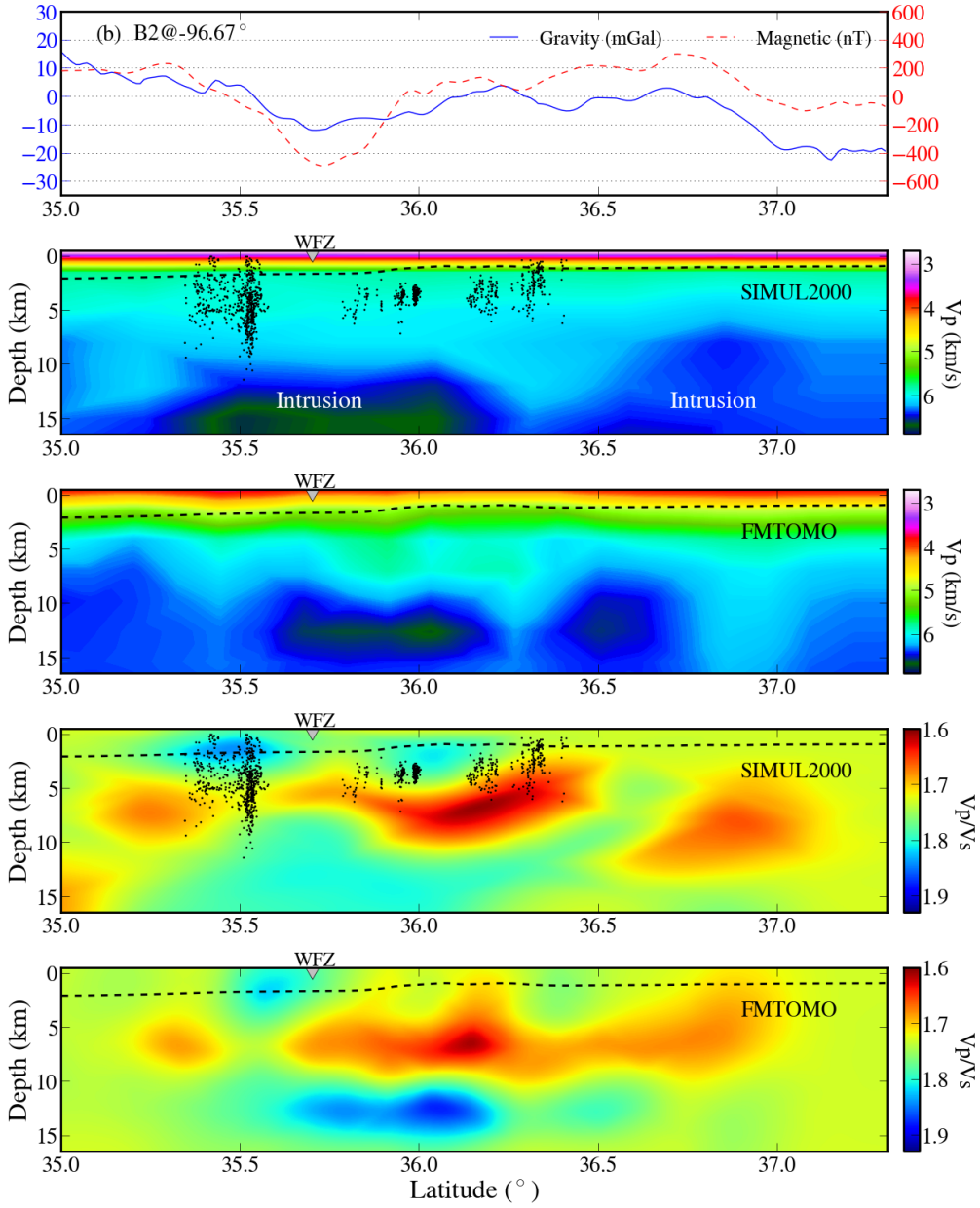


Figure 2.18b. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B2 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. WFZ: Wilzetta Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

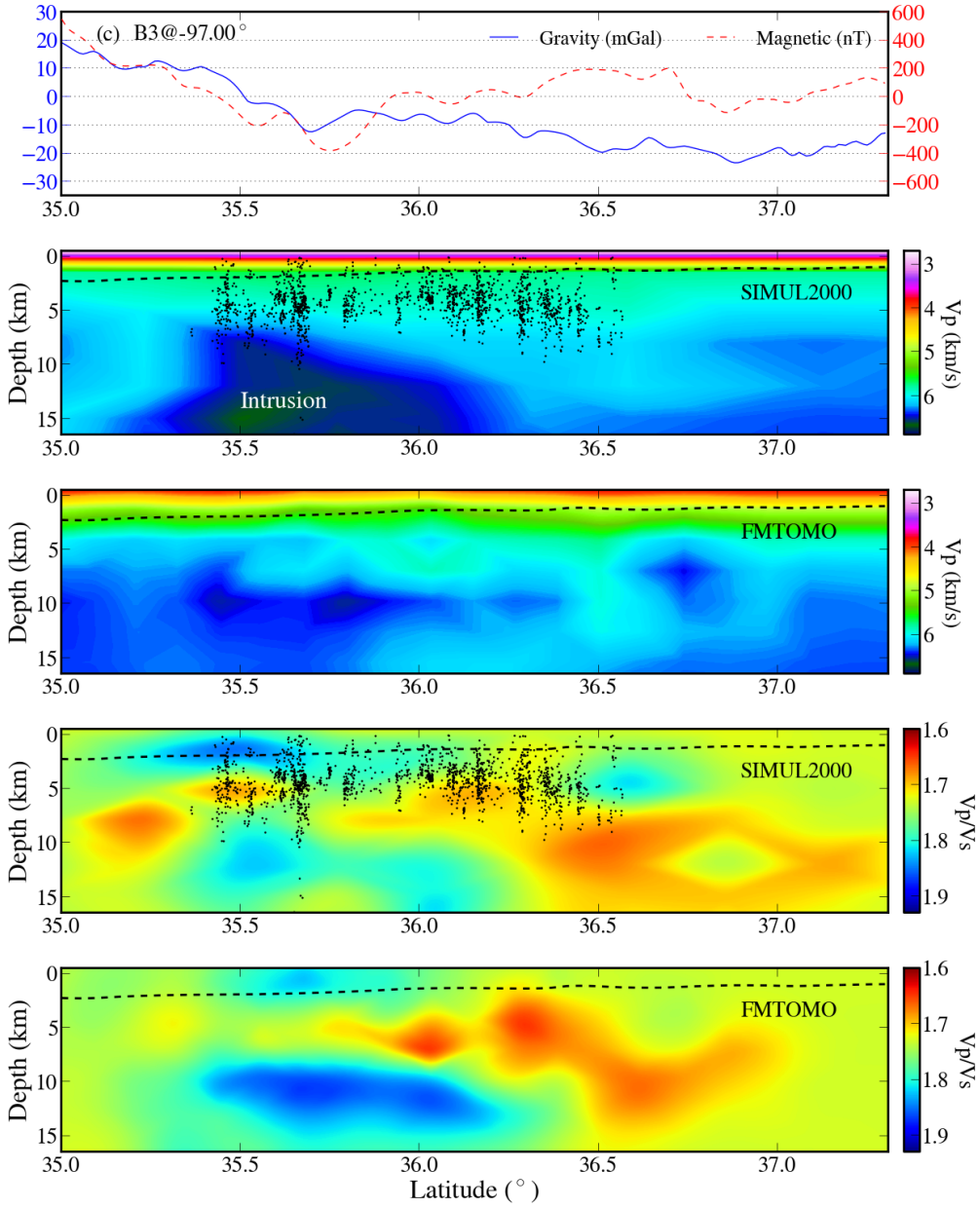


Figure 2.18c. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B3 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

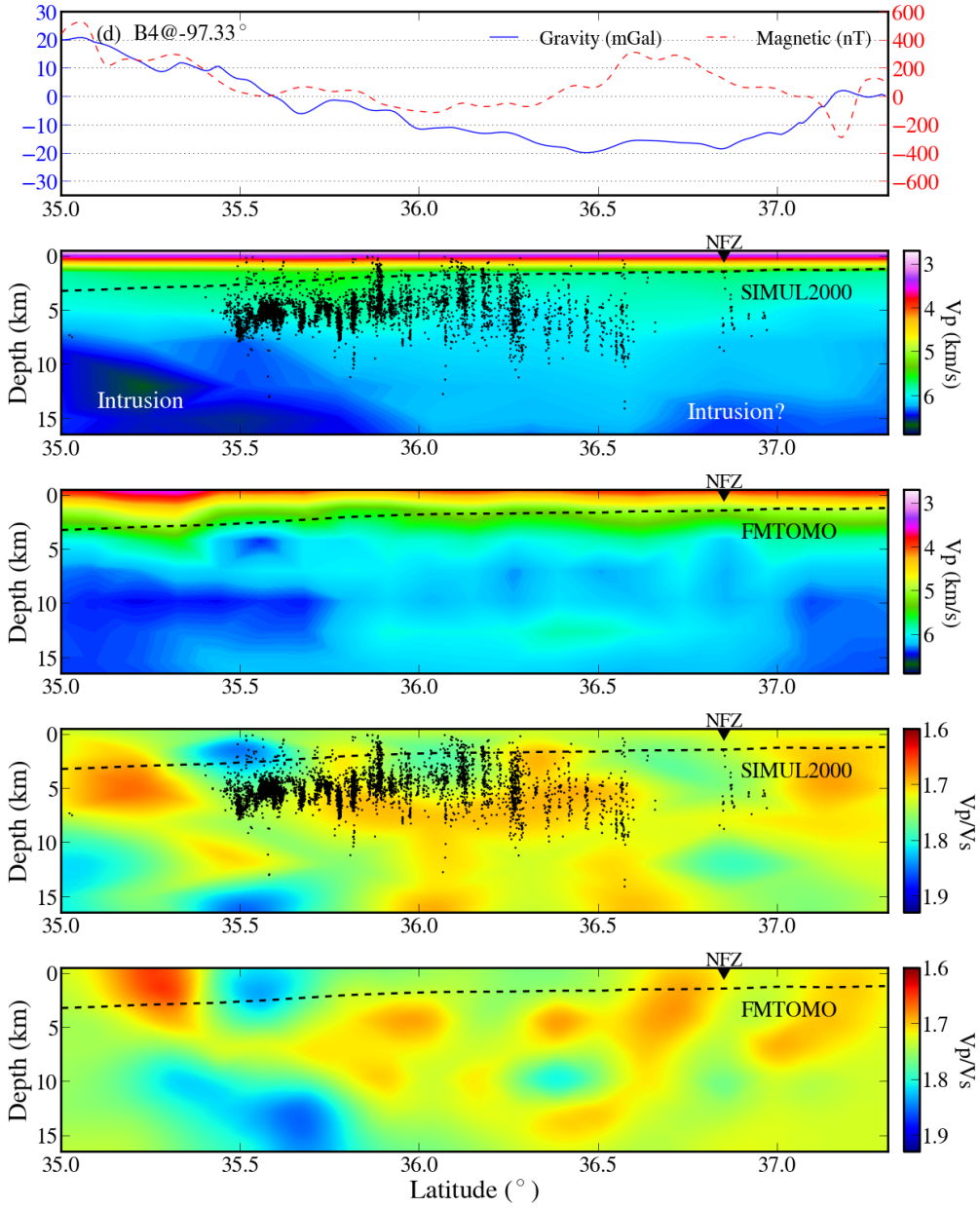


Figure 2.18d. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B4 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

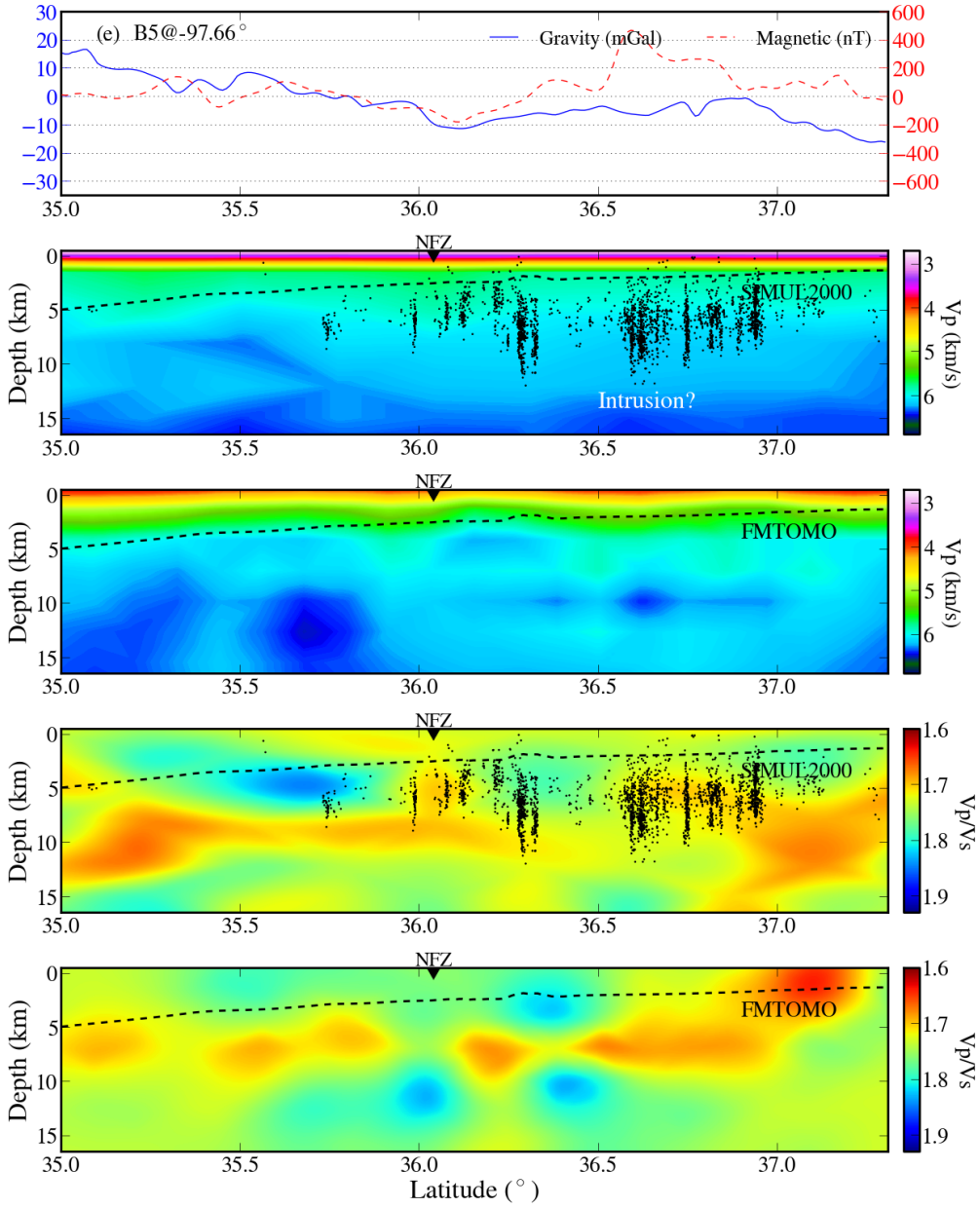


Figure 2.18e. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B5 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

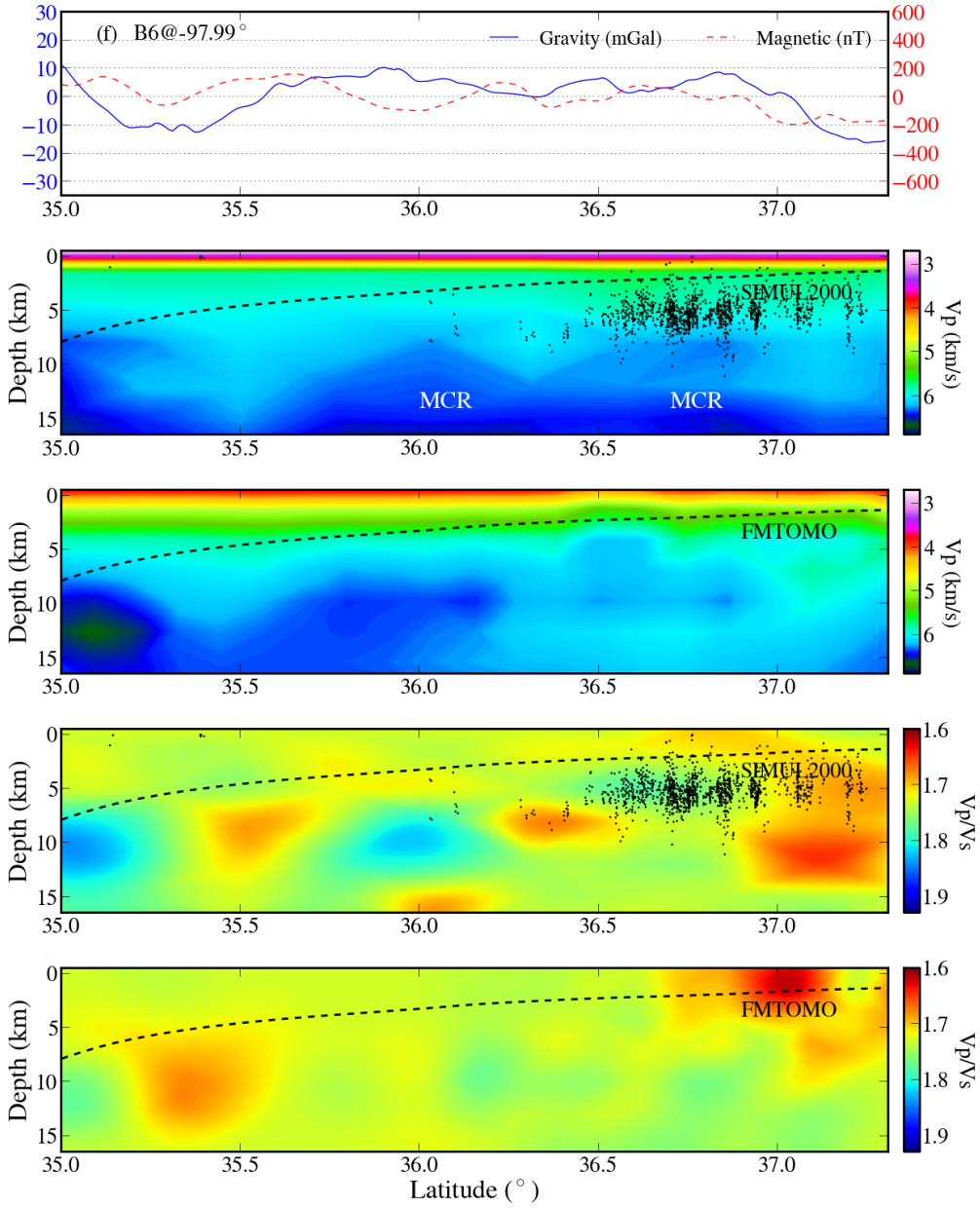


Figure 2.18f. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B6 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

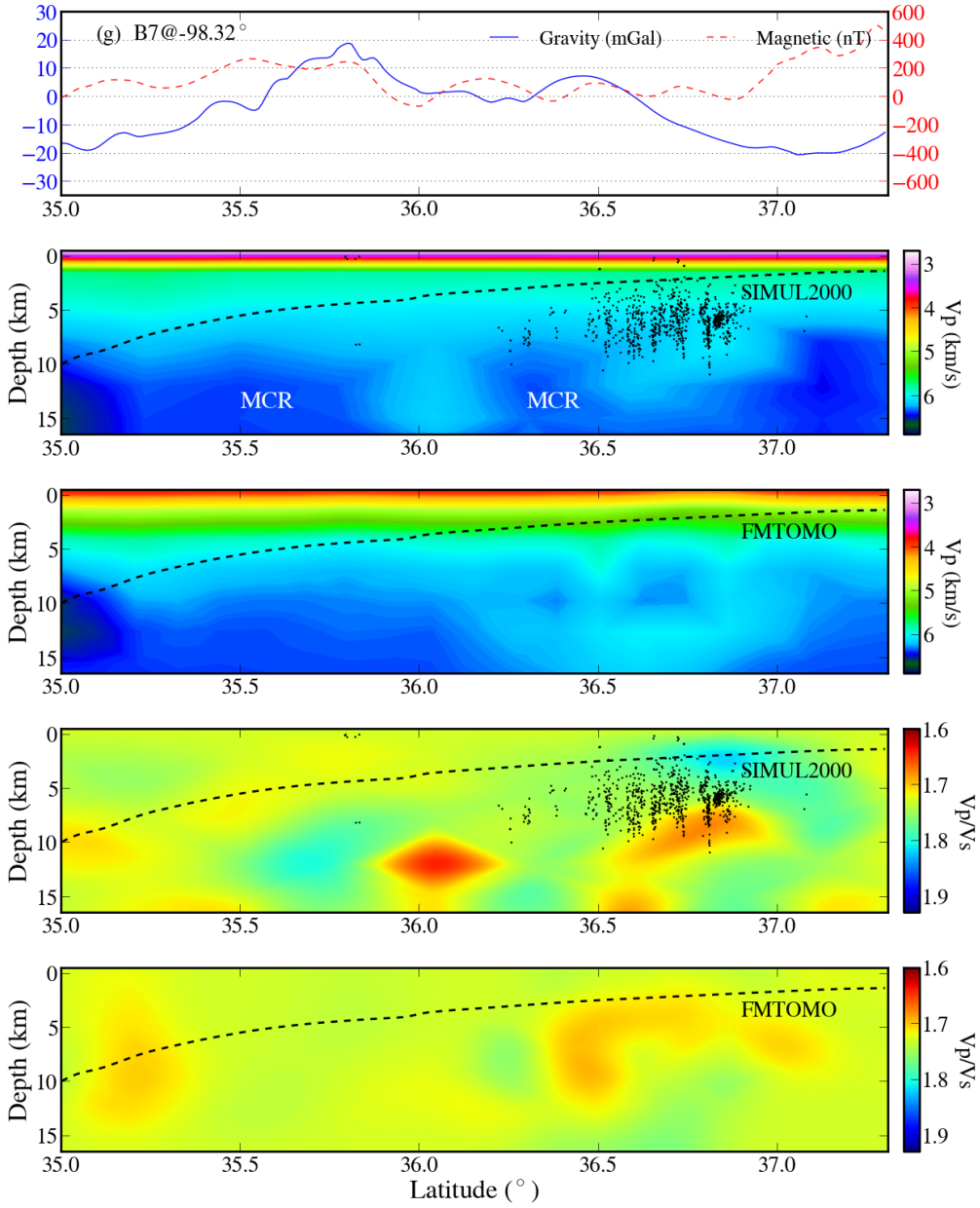


Figure 2.18g. Comparison of gravity, magnetic, Vp, Vp/Vs ratio from the SIMUL2000 and the FMTOMO model along profile B7 in Figs. 2.15 and 2.16. The dashed lines indicate the depth to the basement. The profile location is indicated at the upper left corner of the top panel. NFZ: Nemaha Fault Zone. Dots: HYPODD 3D relocated earthquakes from Chapter 3.

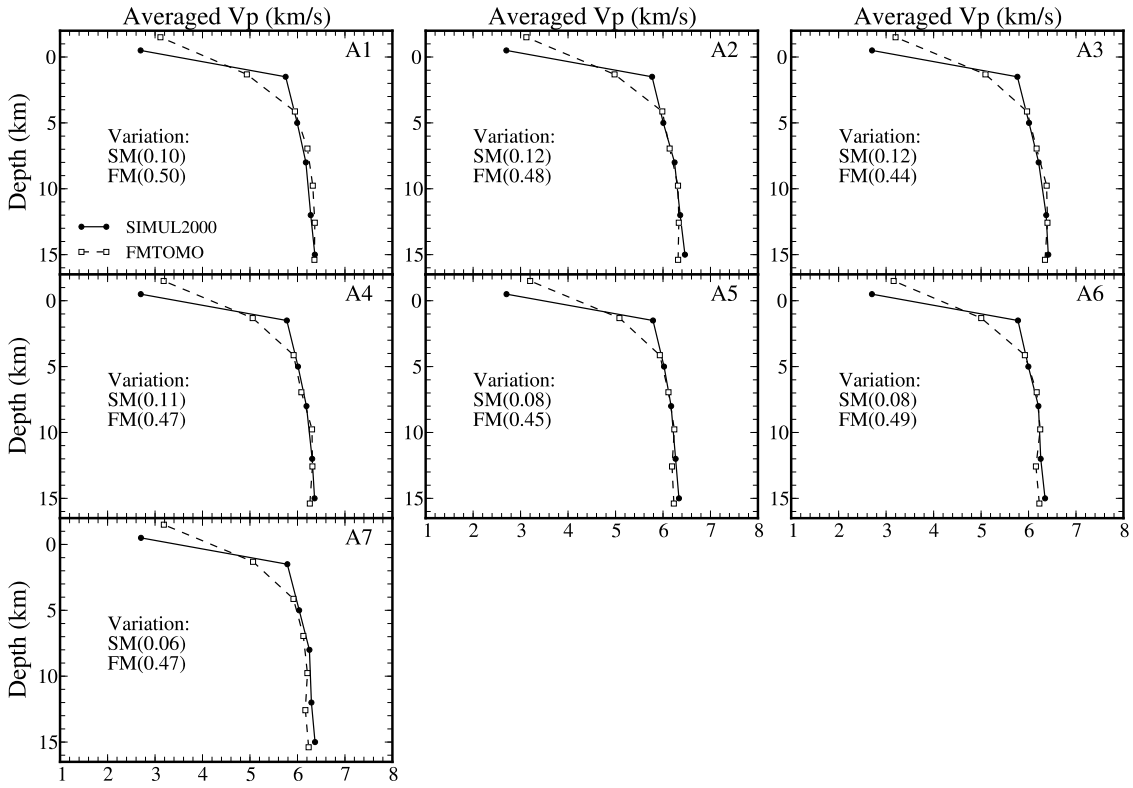


Figure 2.19a. Averaged Vp of the SIMUL2000 and the FMTOMO models for WE cross-sections (A1-A7). The variations of the models are shown in the parenthesis. SM: SIMUL2000; FM: FMTOMO.

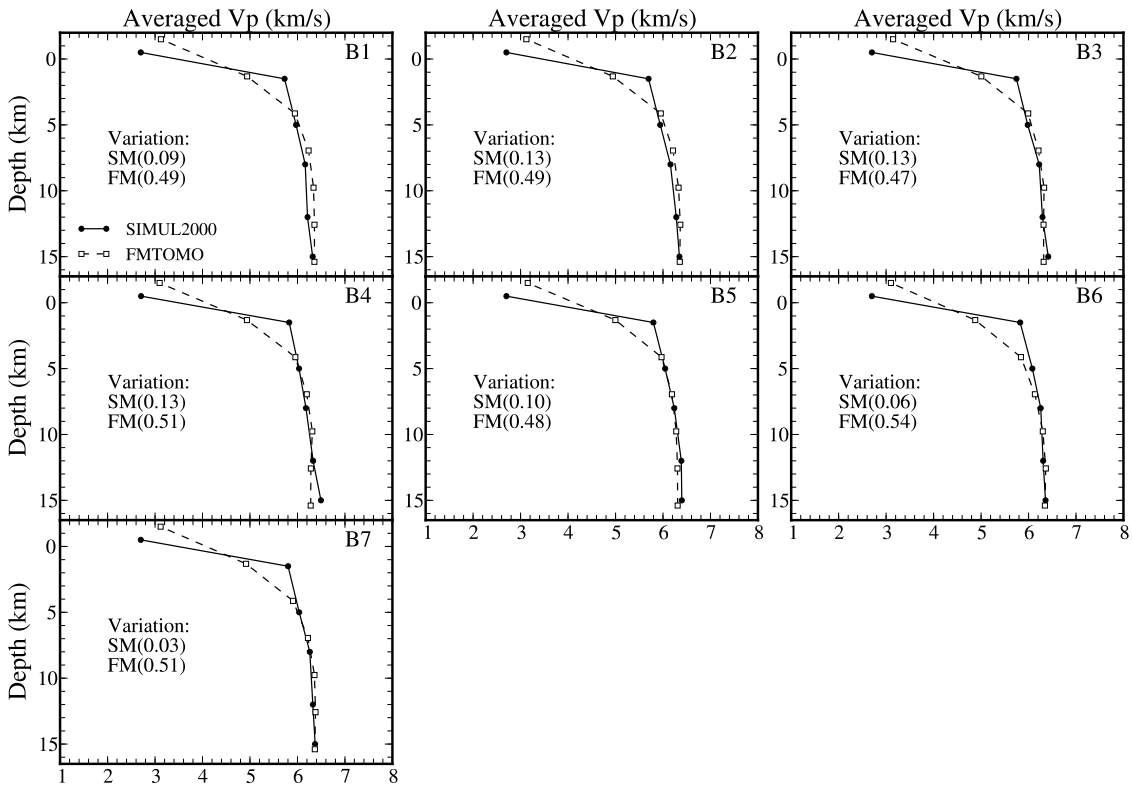


Figure 2.19b. Averaged Vp of the SIMUL2000 and the FMTOMO models for SN cross-sections (B1-B7). The variations of the models are shown in the parenthesis. SM: SIMUL2000; FM: FMTOMO.

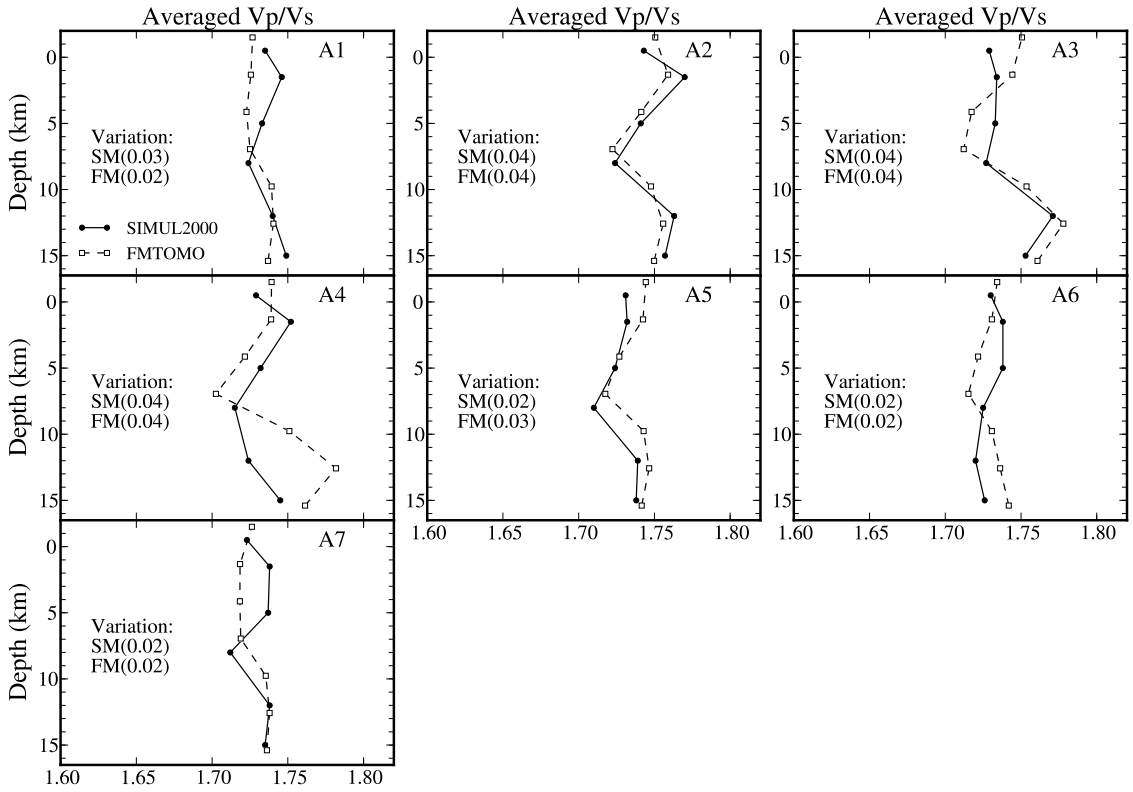


Figure 2.20a. Averaged Vp/Vs ratio of the SIMUL2000 and the FMTOMO models for WE cross-sections (A1-A7). The variations of the models are shown in the parenthesis. SM: SIMUL2000; FM: FMTOMO.

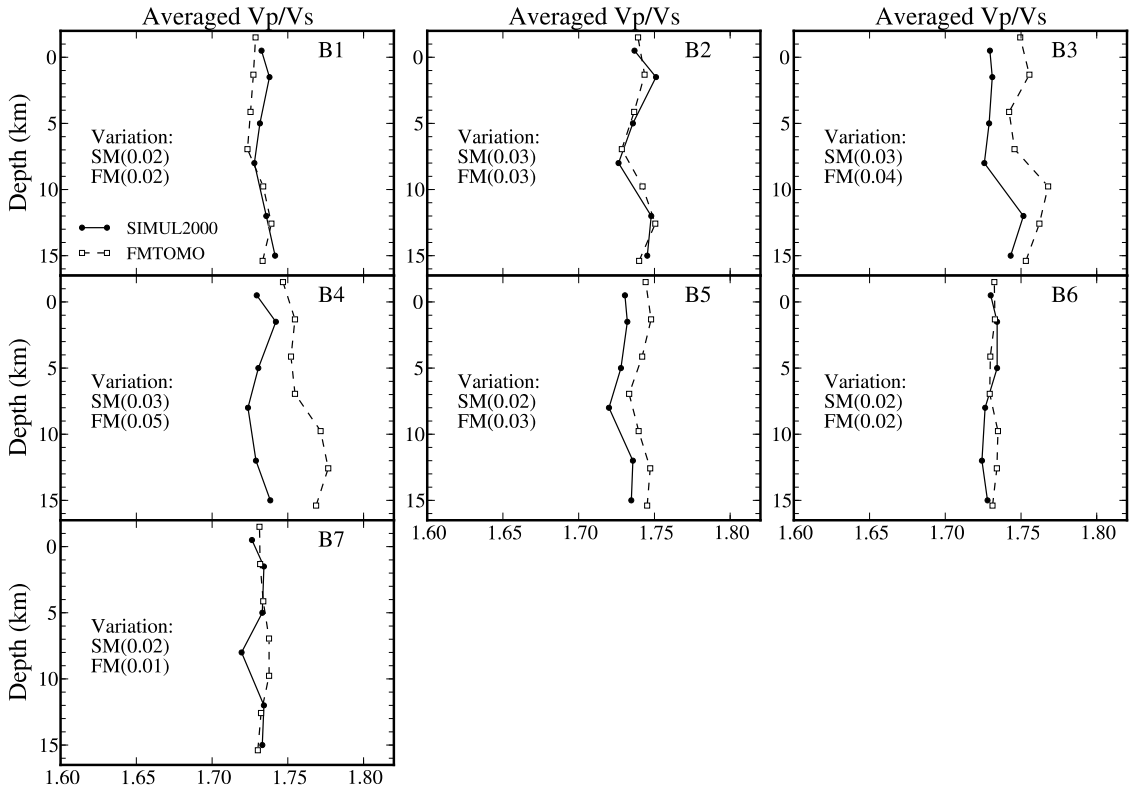


Figure 2.20b. Averaged Vp/Vs ratio of the SIMUL2000 and the FMTOMO models for SN cross-sections (B1-B7). The variations of the models are shown in the parenthesis. SM: SIMUL2000; FM: FMTOMO.

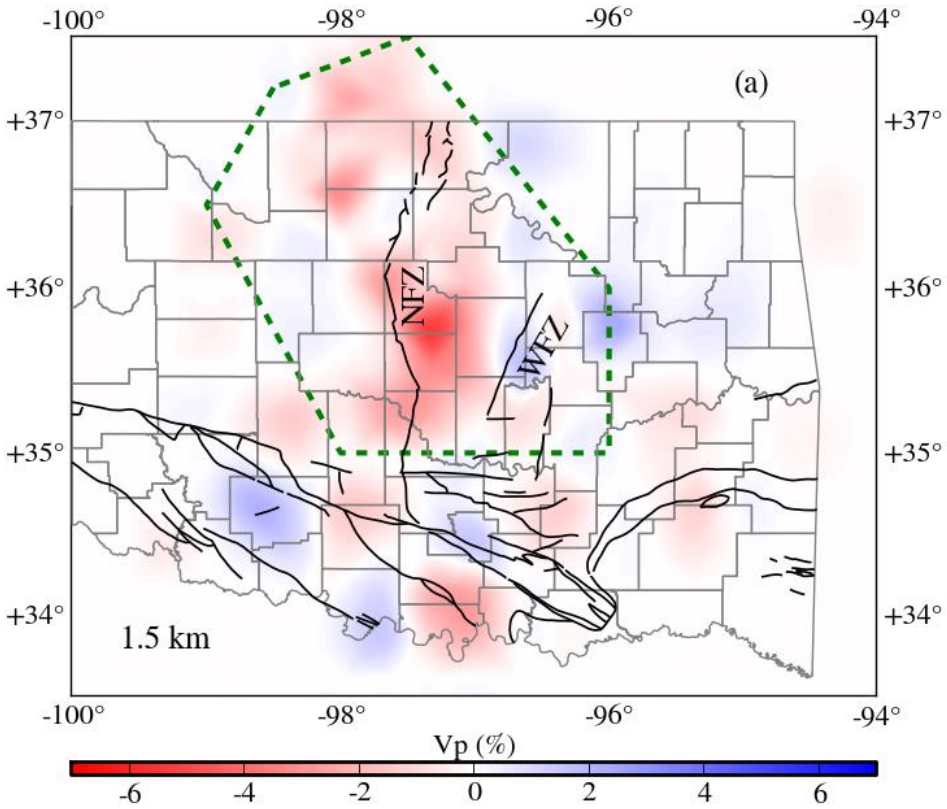


Figure 2.21a. Depth slices of Vp determined by SIMUL2000 at 1.5 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

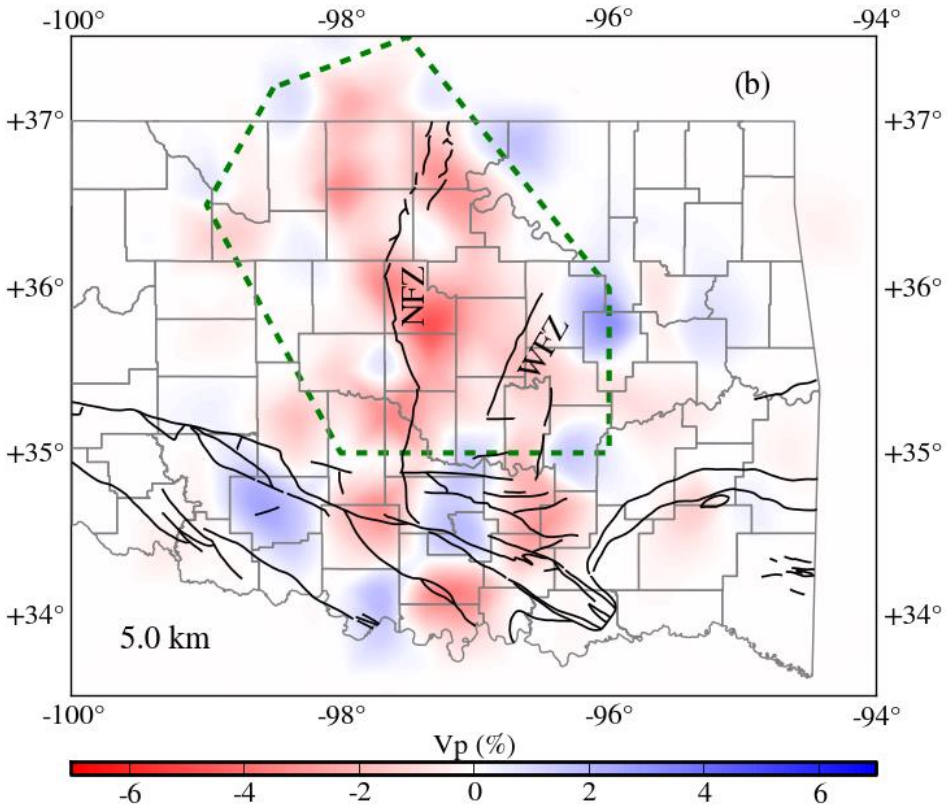


Figure 2.21b. Depth slices of Vp determined by SIMUL2000 at 5 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

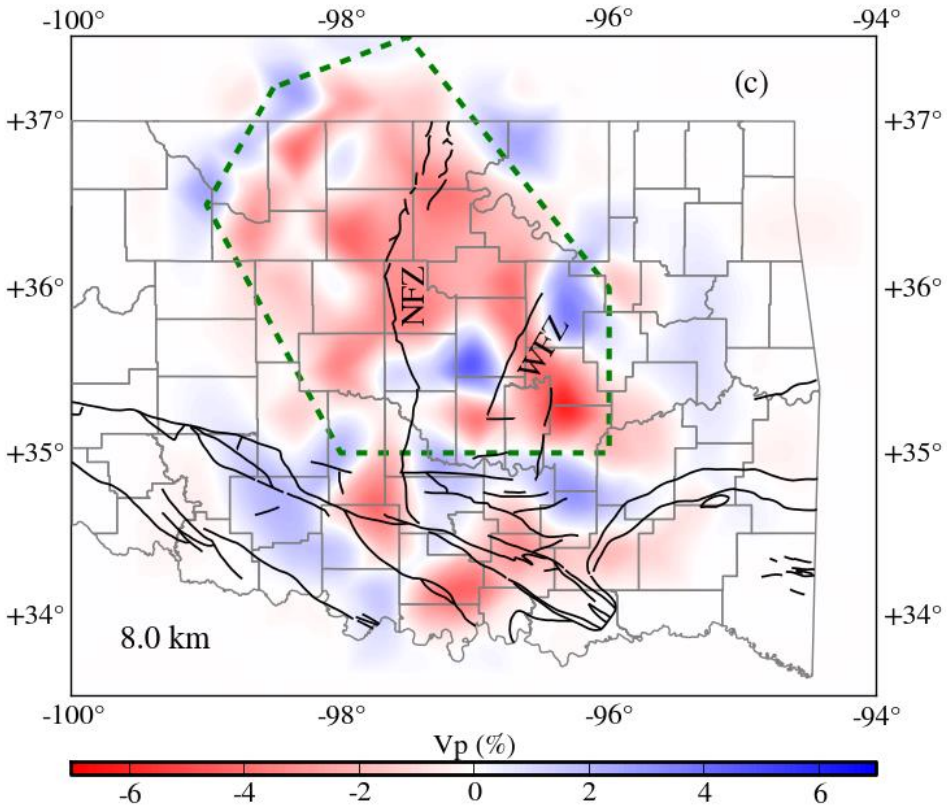


Figure 2.21c. Depth slices of Vp determined by SIMUL2000 at 8 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

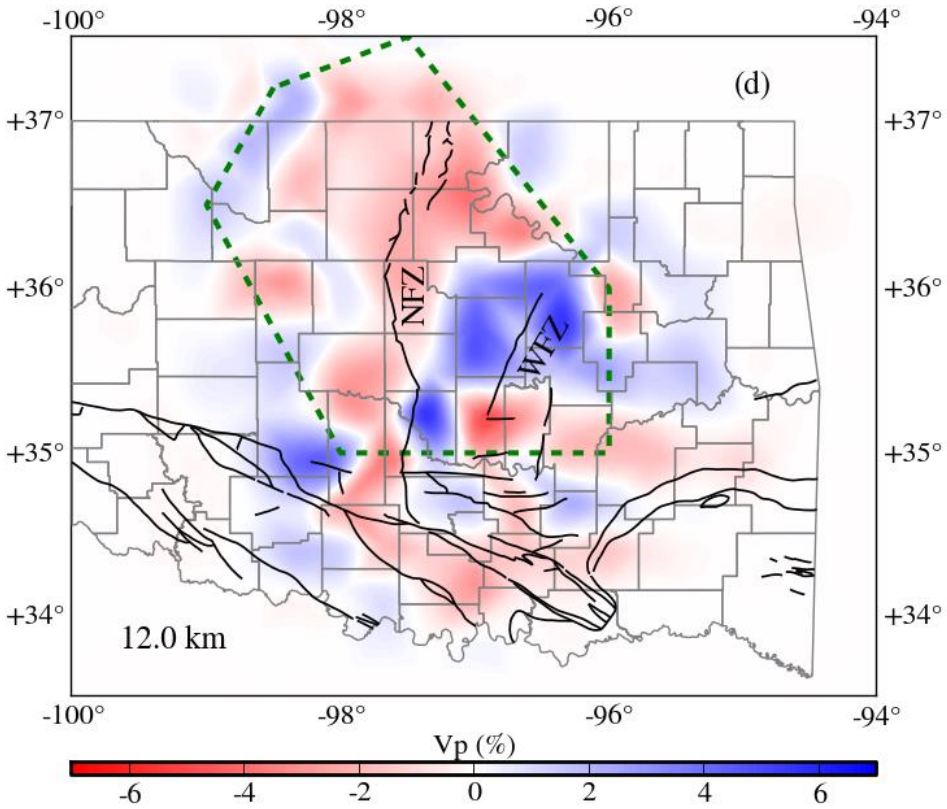


Figure 2.21d. Depth slices of Vp determined by SIMUL2000 at 12 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

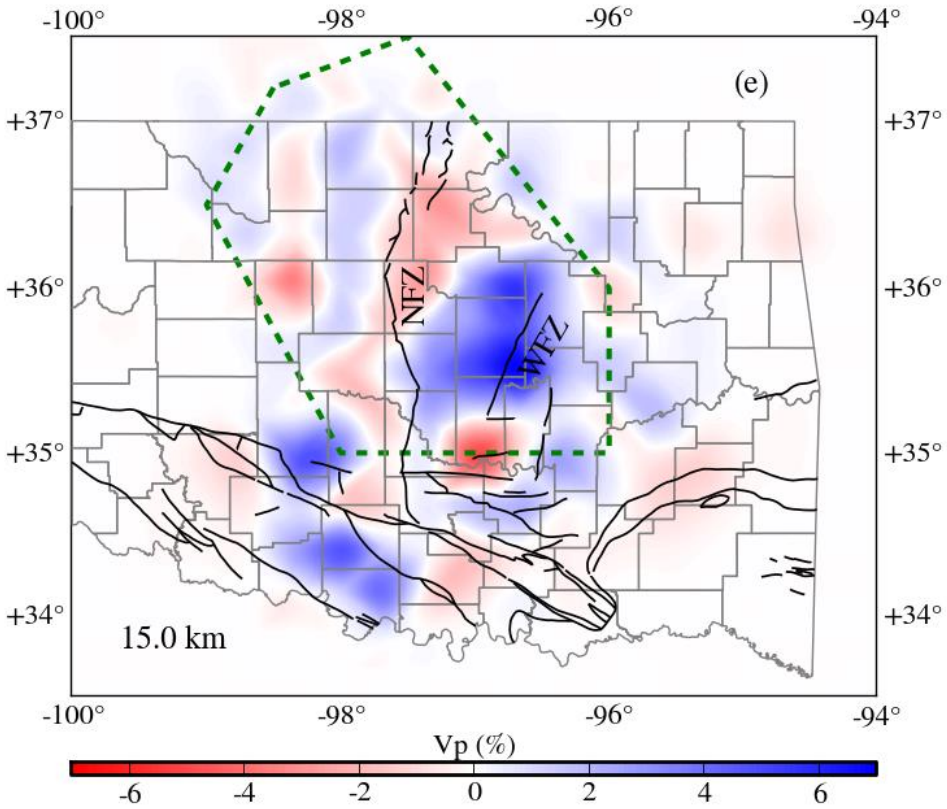


Figure 2.21e. Depth slices of Vp determined by SIMUL2000 at 15 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

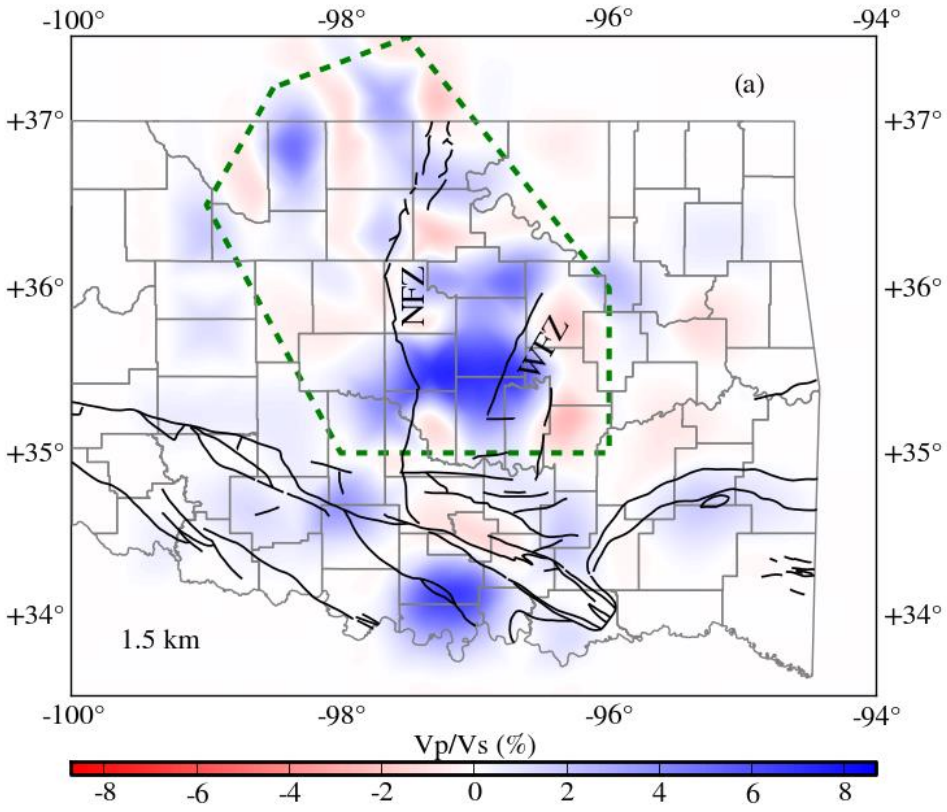


Figure 2.22a. Depth slices of Vp/Vs ratio determined by SIMUL2000 at 1.5 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

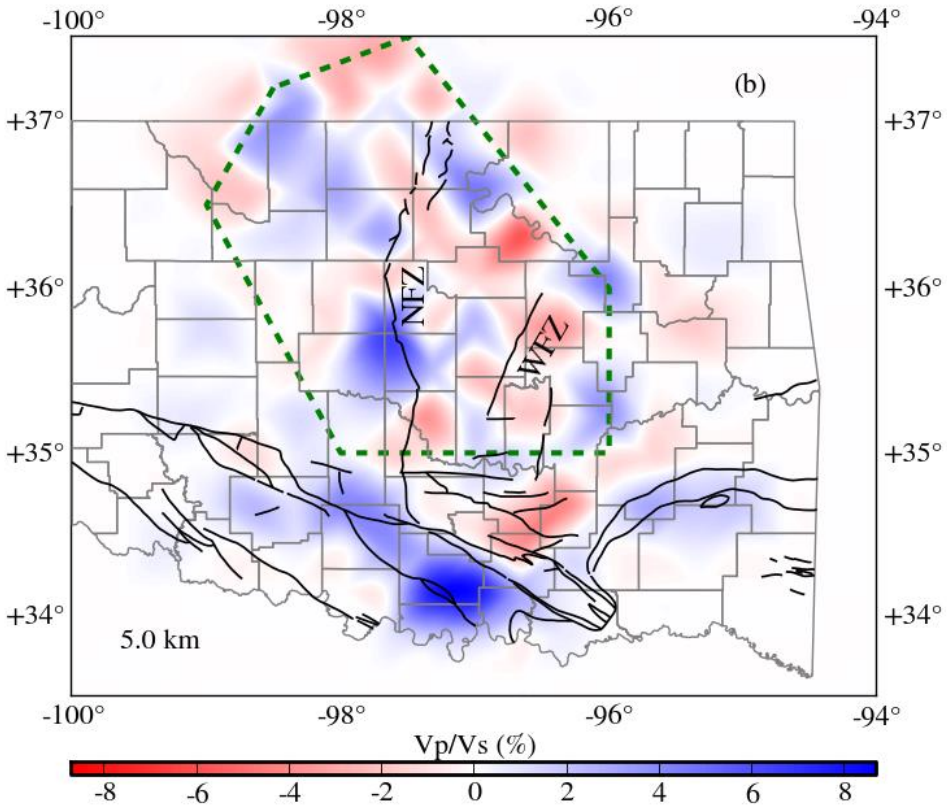


Figure 2.22b. Depth slices of Vp/Vs ratio determined by SIMUL2000 at 5 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

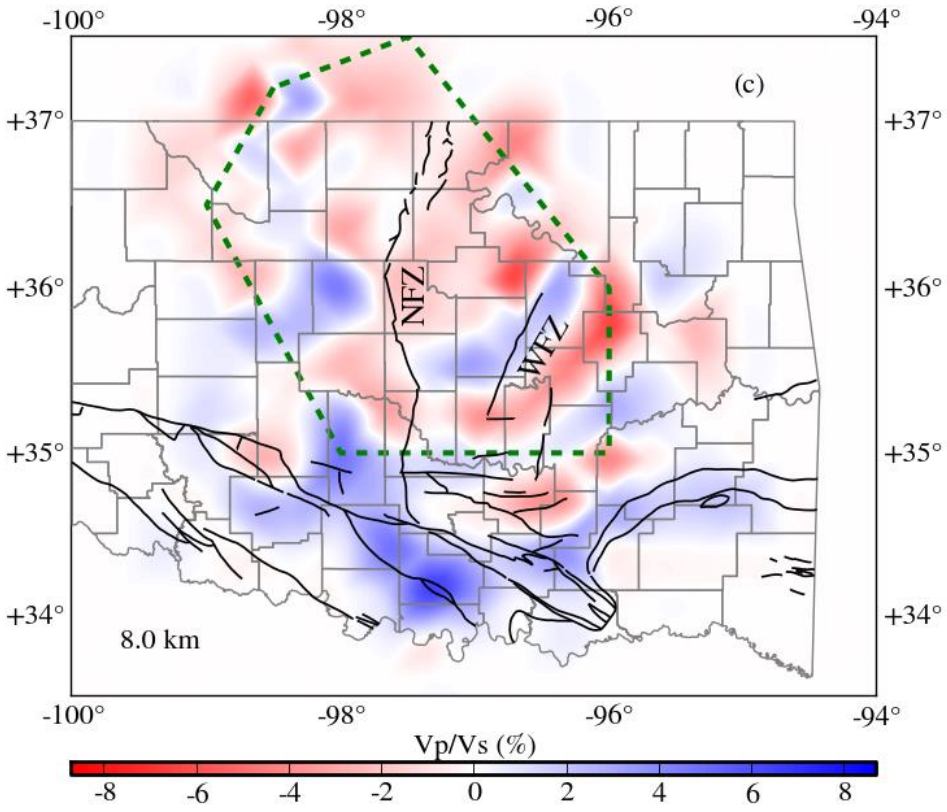


Figure 2.22c. Depth slices of Vp/Vs ratio determined by SIMUL2000 at 8 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

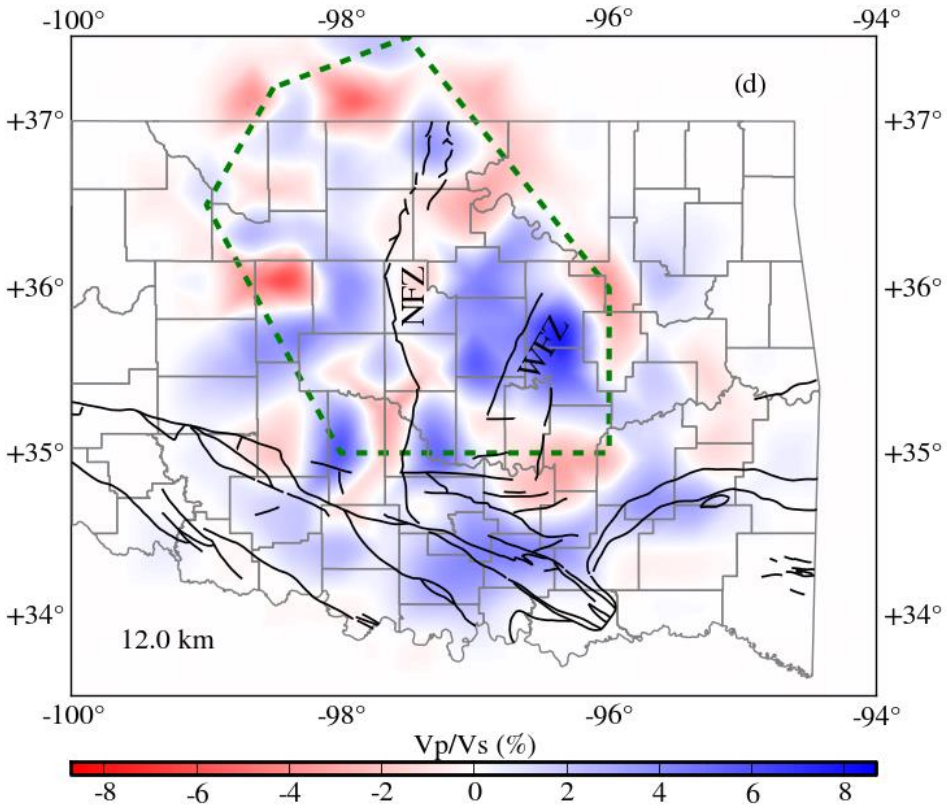


Figure 2.22d. Depth slices of Vp/Vs ratio determined by SIMUL2000 at 12 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

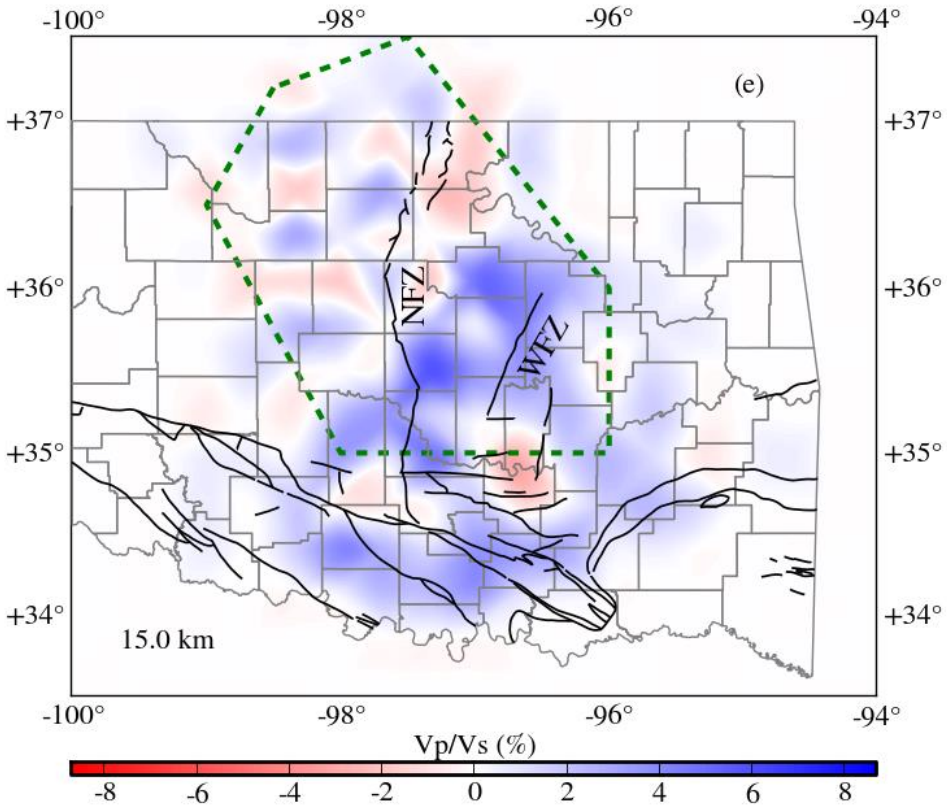


Figure 2.22e. Depth slices of Vp/Vs ratio determined by SIMUL2000 at 15 km. Green polygon indicates region with good resolution. Slice depth is shown at lower left of each plot. Gray polygons are county boundaries, and black lines are major faults in Oklahoma. Green polygon indicates the region with good resolution.

Depth (km)	0	0.3	1.0	1.5	8.0	21	42	50	80
Vp (km/s)	2.70	2.95	4.15	5.80	6.27	6.41	7.90	8.15	8.50

Table 2.1. Initial 1D P-wave velocity model.

References

Adams, D. C., and G. R. Keller (1996), Precambrian basement geology of the Permian Basin region of West Texas and eastern New Mexico: A geophysical perspective, AAPG Bulletin, **80**, 410–431.

Astiz, L., J. A. Eakins, V. G. Martynov, T. A. Cox, J. Tytell, J. C. Reyes, R. L. Newman, G. H. Karasu, T. Mulder, M. White, G. A. Davis, R. W. Busby, K. Hafner, J. C. Meyer, and F. L. Vernon (2014), The array network facility seismic bulletin: Products and an unbiased view of United States seismicity, Seismological Research Letter, **85**(3), 576–593, doi:10.1785/0220130141.

Brewer, J. A., R. Good, J. E. Oliver, L. D. Brown, and S. Kaufman (1983), COCORP profiling across the Southern Oklahoma aulacogen: Overthrusting of the Wichita Mountains and compression with the Anadarko Basin, Geology, **11**, 109–114.

Buckey, A. (2012), An Integrated geophysical analysis of crustal structure in the Wichita uplift region of southern Oklahoma, The Shale Shaker, 62(6), 432-452.

Chen, C., and A. A. Holland (2016), PhasePApy: A robust pure Python package for automatic identification of seismic phases, Seismological Research Letter. **87**, no. 6, 2016.

Darold, A. P., A. A. Holland, C. Chen, and A. Youngblood (2014), Preliminary Analysis of Seismicity Near Eagleton 1-29, Carter County, July 2014, Oklahoma Geological Survey Open-File Report, OF2-2014.

Darold, A. P., A. A. Holland, J. K. Morris, and A. R. Gibson (2015), Oklahoma earthquake summary report 2014, Oklahoma Geological Survey Open-File Report, OF1-2015.

Darold, A. P., and A. A. Holland (2015), Preliminary optimal Oklahoma fault orientations, Oklahoma Geological Survey Open-File Report, OF4-2015.

Denison, R. E., 1966, Basement rocks in adjoining parts of Oklahoma, Kansas, Missouri and Arkansas, Ph.D Dissertation, University of Texas, 291.

Denison, R. E., 1981, Basement rocks in northeastern Oklahoma, Oklahoma Geological Survey Circular 84.

Denison, R. E., E. G. Lidiak, M. E. Bickford, and E. B. Kisvarsanyi (1984), Geology and geochronology of the Precambrian rocks in the central interior region of the United States, U.S. Geological Survey Professional Paper, 1214 C, 1-10.

Ellsworth, W. L. (2013), Injection-induced earthquakes, Science, 341, 1225942, doi:10.1126/science.1225942.

Gay, S. P. (1995), The basement fault block pattern: its importance in petroleum exploration, and its delineation with residual aeromagnetic techniques, in R.W. Ojakangas, ed., Proceedings of the 10th International Basement Tectonics Conference, 159-207, doi: 10.1007/978-94-017-0831-9_23.

Havskov, J., and L. Ottemoller (1999), SeisAn earthquake analysis software, Seismological Research Letters, **70**, 532-534.

Hoffman, P., J. F. Dewey, and K. Burke (1974), Aulacogens and their genetic relations to geosynclines, with a Proterozoic examples from Great Slave Lake, Canada, Modern and ancient geosynclinal sedimentation, Society of Economic Paleontologists and Mineralogists Special Publication, **19**, 38-55.

Holland, A. A. (2011), A. A. Gibson, C. Chen, G. R. Keller, and K. V. Luza, Incorporating near real-time transportable array data into the regional Oklahoma

seismic network, Seismological Research Letter Annual Meeting, **82**(2), 358.

Holland, A. A. (2013), Earthquakes triggered by hydraulic fracturing in south-central Oklahoma, Bulletin Seismological Society of America, 103(3), 1784-1792, doi:10.1785/0120120109.

Holland, A. A. (2013), Optimal fault orientations within Oklahoma, Seismol. Res. Lett., **84**, 876-890, doi: 10.1785/022012015.

Johnson, K. S., and K. V. Luza (2008), Earth sciences and mineral resources of Oklahoma, Oklahoma Geological Survey Educational Publication-9.

Keranen, K. M., H. M. Savage, G. A. Abers, and E. S. Cochran (2013), Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 MW 5.7 earthquake sequence, Geology, **41**, 699-702, doi:10.1130/G34045.1.

Keranen, K. M., M. Weingarten, G. A. Abers, B. A. Bekins, and S. Ge (2014), Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection, Science, **345**, 448-451, doi: 10.1126/science.1255802.

Keller, G. R., E. G. Lidiak, W. J. Hinze, and L. W. Braile (1983), The role of rifting in the tectonic development of the Midcontinent, U.S.A., Tectonophysics, **94**, 391-412.

Keller, G. R., and R. A. Stephenson (2007), The Southern Oklahoma and Dniepr-Donets Aulacogens: a comparative analysis, GSA Memoir, **200**, 127-143, doi:10.1016/0040-1951(83)90026-4.

Kruger, J. R., and G. R. Keller (1986), Interpretation of crustal structure from regional gravity anomalies, Ouachita Mountains area and adjacent Gulf coastal plain,

AAPG Bulletin, **70**, 667–689.

Lin, G., P. M. Shearer, E. Hauksson, and C. H. Thurber (2007), A three-dimensional crustal seismic velocity model for southern California from a composite event method, Journal of Geophysical Research, **112**, B11306, doi:10.1029/2007JB004977.

Lin, G., and P. M. Shearer (2009), Evidence for water-filled cracks in earthquake source regions, Geophysical Research Letters, **36**, L17315, doi:10.1029/2009GL039098.

Luza, K. V., and J. E. Lawson (1981), Seismicity and tectonic relationships of the Nemaha Uplift in Oklahoma, part III, Oklahoma Geological Survey Special Publication, 81-3.

Luza, K. V., and J. E. Lawson (1982), Seismicity and tectonic relationships of the Nemaha Uplift in Oklahoma, part IV, Oklahoma Geological Survey Special Publication, 82-1.

McGarr, A., B. Bekins, N. Burkhardt, J. Dewey, P. Earle, W. Ellsworth, S. Ge, S. Hickman, A. Holland, E. Majer, J. Rubinstein and A. Sheehan (2015), Coping with earthquakes induced by fluid injection, Science, **347,** 830-831, doi:10.1126/science.aaa0494.

Merino, M, G. R. Keller, S. Stein, and C. Stein (2014), Variations in Mid-Continent Rift magma volumes consistent with microplate evolution, Geophysical Research Letter, **40**, 1513-1516, doi**:** 10.1002/grl.50295.

Morgan, B. C., and K. E. Murray (2015), Characterizing small-scale permeability of the Arbuckle Group, Oklahoma. Oklahoma Geological Survey Open-File Report OF2-2015.

Muehlberger, W. R., R. E. Denison, and E. G. Lidiak (1967), Basement rocks in continental interior of United States: AAPG Bulletin, **51**, 2351-2380.

Murray, K. E. (2014), Class II Underground Injection Control Well Data for 2010–2013 by Geologic Zones of Completion, Oklahoma, Oklahoma Geological Survey Open-File Report, OF1-2014.

Nishi, K. (2001), A three-dimensional robust seismic ray tracing for volcanic regions, Earth Planets Space, **53**, 101-109, doi:10.1186/BF03352367.

Northcutt, R. A., and J. A. Campbell (1995), Geologic provinces of Oklahoma, Oklahoma Geological Survey Open-File Report 5-95.

Ottemöller, L., P. Voss, and J. Havskov (2013), SEISAN earthquake analysis software for Windows, Solaris, Linux and MacOSX. Department Earth Science University of Bergen: Norway, 1-416.

Petersen, M. D., C. S. Mueller, M. P. Moschetti, S. M. Hoover, J. L. Rubinstein, A. L. Llenos, A. J. Michael, W. L. Ellsworth, A. F. McGarr, A. A. Holland and J. G. Anderson (2015), Incorporating Induced Seismicity in the 2014 United States National Seismic Hazard Models: Results of 2014 Workshop and Sensitivity Studies: U.S. Geological Survey Open-File Report, 2015-1070, 69 p., http://dx.doi.org/10.3133/ofr20151070.

Perry, W. J., Jr. (1989), Tectonic evolution of the Anadarko Basin region, Oklahoma, U.S. Geological Survey Professional Paper, 1866-A.

Rawlinson, N., M. Urvoy (2006), Simultaneous inversion of active and passive source datasets for 3-D seismic structure with application to Tasmania, Geophysical Research Letter, **33**, L24313, doi:10.1029/2006GL028105.

Shatski, N.S. (1946), The Great Donets Basin and the Wichita System, in Comparative tectonics of ancient platforms, Akademii Nauk SSSR, Izvestia, Seriya Geologicheskaya, **1**, 15-62 (in Russian).

Sethian, J. A., and A. M. Popovici (1999), 3-D traveltime computation using the fast marching method, Geophysics, **64**, 516-523, doi:10.1190/1.1444558.

Stein, C. A., S. Stein, M. Merino, G. R. Keller, L. M. Flesch, and D. M. Jurdy (2014), Was the Midcontinent Rift part of a successful seafloor-spreading episode?, Geophysical Research Letters, **41**, 10.1002/2013GL059176, doi:10.1002/2013GL059176.

Thurber, C. H. (1983), Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, Journal of Geophysical Research, **88**, 8227-8326, doi:10.1029/JB088iB10p08226.

Toth, C. R. (2014), Separation of the Earthquake Tomography Inverse Problem to Refine Hypocenter Locations and Tomographic Models: A Case Study from Central Oklahoma, University of Oklahoma, Norman, OK, USA.

Tryggvason, E., and B. R., Qualls (1967), Seismic refraction measurements of crustal structure in Oklahoma: Journal of Geophysical Research, **72**, 3738-3740.

Waldhauser, F., and W. L. Ellsworth (2000), A double-difference earthquake location algorithm: method and application to the northern Hayward Fault, California, Bulletin Seismological Society of America, **90**, 1353-1368, doi:10.1785/0120000006.

Waldhauser, F. (2001), HypoDD: a computer program to compute double-difference hypocenter locations, U.S. Geological Survey Open File Report 01-113.

Walsh F. R., and M. Zoback (2015), Oklahoma's recent earthquakes and saltwater disposal, Science Advances, **1**(5), e1500195, doi:10.1126/sciadv.1500195.

Weingarten, M., S. Ge, J. W. Godt, B. A. Bekins, and J. L. Rubinstein (2015), High-rate injection is associated with the increase in US mid-continent seismicity, Science **348**(6241), 1336–1340, doi:10.1126/science.aab1345.

Zhang, H. J., and C. H. Thurber (2003), Double-difference tomography: The method and its application to the Hayward Fault, California, Bulletin Seismological Society of America, **93**, 1875–1889, doi:10.1785/0120020190.

Chapter 3: Earthquake Relocations Using A Three-Dimensional Velocity Model and Double Differencing: Relocations of Earthquakes in Oklahoma

Introduction

Earthquake location estimation is one of the most important inverse problems in seismology study. Most earthquake location methods are based on exploiting phase arrivals from the seismic waveforms and the velocity model. The accurate earthquake location is important for many applications, such as fault studies, hazard assessment, seismic tomography, stress field, induced seismicity, and others. The early earthquake location determination employed simple grid search (e.g., Reid, 1910) or Geiger's method (1912). With the power of computer increasing, numerous methods have been developed to improve the accuracy of absolute locations with the single event method for local and global earthquakes in the past few decades (e.g., Thurber, 1983; Moser et al., 1992; Lomax et al., 2000). There are several other algorithms developed to locate multiple events at the same time that can improve the relative locations, such as joint hypocenter determination (JHD) (Douglas, 1967; Frohlich, 1979), hypocentroidal decomposition (HDC) (Jordan and Sverdrup, 1981), and source-specific station terms (SSST) method (Richards-Dinger and Shearer, 2000). Double-difference (DD) method is another algorithm for improving relative locations between the nearby events in a network (Waldhauser and Ellsworth, 2000; Waldhauser, 2001). The DD methods significantly improve the relative earthquake location accuracy either with a 1D or 3D velocity model. The improved relative earthquake locations can produce the sharper

clustered seismicity patterns, which can help to identify the fault location and orientation. In order to relocate the increased seismicity in central Oklahoma since 2009, we derived the three-dimensional (3D) velocity model with the SIMUL2000 algorithm (Thurber, 1983) in Chapter 2.

Most of the earthquakes in central Oklahoma are clustered and presented northeast-southwest (NE-SW) or northwest-southeast (NW-SE) trending orientation that are consistent with maximum horizontal stress state within the region (e.g., McNamara, 2015). With the 3D velocity model, we relocated the cataloged earthquakes. The high-accuracy earthquake locations improved the resolution of fault locations by producing sharper patterns of the seismicity. Furthermore, the improved the earthquake locations can better understand the relationship between the injection wells and the induced seismicity. In addition, the improvement of the earthquake locations can help with identifying primary fault planes from focal mechanism, crustal deformation, and others.

Data and Methods

Our 3D velocity relocation strategy is to first relocate the cataloged earthquakes with the single event method, which is used to improve the absolute earthquake locations. Then, a DD method follows to improve the relative locations for the clustered nearby events with the 3D velocity model. The DD technique takes the advantage of the fact that if the hypocentral separation between the nearby events is much smaller than the event-station distance and velocity variation scale, the ray paths from the event to the common station are similar (Waldhauser and Ellsworth, 2000; Waldhauser, 2001). In this case, the relative location between the nearby events can be improved by the

double-differential traveltimes and the hypocentral separation, even though the traveltimes are biased by the 3D velocity structures.

The Oklahoma Geological Survey (OGS) has used the earthquake analysis package SEISAN (Havskov and Ottemoller, 1999) since 2010. From the OGS catalog, there were a great number of earthquakes identified in central Oklahoma and the surrounding area from early 2010 to July 2016. We selected 17385 earthquakes in this study region from the OGS catalog from 01/01/2010 to 03/31/2016 by requiring events having records at least 6 stations with a root mean square (RMS) residuals 2.0 s cutoff (Fig. 3.1). Due to the phase arrival outliers existed in the database, the picks with traveltime errors greater than 1.5 s comparing to modeled travel-time curve from the 1D velocity model (Table 5 in Darold et al., 2015) were removed, and we also removed the picks with epicentral distance greater than 350 km, which is about the dimension of our study region (Fig. 3.2(a)). There were 282531 P-picks and 274698 S-picks left in the earthquake relocation process after the bad pick filtering. The histograms of the data are shown in Figure 3.2(b)-(d).

Step 1: Initially, we used the SIMUL2000 to relocate 17385 selected cataloged earthquakes with the 3D velocity model. The 3D velocity model was derived with the composite event method (Lin et al., 2007) from 8194 earthquakes with magnitude greater than 2.0 in central Oklahoma and the surrounding area (please refer to Chapter 2 for more details). Simultaneously inverting for the velocity and location is usually used for the local earthquakes (e.g., Thurber, 1983; Zhang and Thurber, 2003). Ideally, we would like to invert for the velocity model using all the available earthquakes and relocate all the earthquakes simultaneously. However, large data volume makes the

computation time intensive. The computation capability limitation forced us to choose the composite event method to derive the velocity model and relocate earthquakes separately. The SIMUL2000 allows one only to relocate earthquakes by keeping the 3D velocity fixed. We applied 6 iterations in the relocation procedure.

Step 2: During the single event relocation in Step 1, some poorly located earthquakes of high residuals were thrown out. We added them back with their initial cataloged locations, so as to make the catalog completed. Then, we clustered the earthquakes with 2 km hypocentral separation between the nearby events with the 3D version of HYPODD package and used the double-differential traveltimes to improve the relative locations. Only the double-differential catalog data were used in the relocation since the waveforms cross correlation data were not available yet at the moment of documenting this work. We applied 3 sets of iteration in the relocation. Each set had 3 iterations with decreasing residual threshold weight (9, 8, and 7, respectively). The catalog P-picks weights were 1.0 for the 3 sets of iteration, and we used 0.6, 0.65, and 0.7 for the weights of S-picks for each iteration set, respectively. The max distance threshold between the linked pairs was set as 2 km. The damping values were 100, 80, and 60 for the 3 sets of iteration, respectively.

Step 3: We also used the 1D version of HYPODD package with the 1D velocity model to relocate all the selected cataloged earthquakes. We investigated how the velocity model biased the earthquake location by examining the location shift between the 1D and the 3D velocity models relocation for some large or important clusters in central Oklahoma. In order to compare the relocation results from the 1D and 3D velocity relocations, all the parameters of the iteration sets were identical except each

set has 4 iterations. (i.e. In this 1-D velocity relocation, there were also 3 sets of iteration. We determined to use 4 iterations for each set by running several tests to check the condition number during the relocation. See the user manual for more details about the condition number.)

Results

Single event method relocation results with the 3D velocity model

The 17385 selected earthquakes were relocated with the SIMUL2000. There are 16361 relocated earthquakes relocated with the single event relocation method in Figure 3.3, which shows sharper cluster patterns than the initial catalog. Less than 5.9% events were discarded due to the large residuals. Since the DD method works well for the clustered earthquakes, the clustered events with the sharper pattern are the better starting locations for the following DD method to improve the relative locations. The traveltime residuals met a reasonable convergence minimum value after 6 iterations. The RMS residual of each iteration is shown in Figure 3.4. The RMS residual reduces much for the first two iterations. From iteration 3 to 6, the RMS residuals do not change significantly. The normalized histograms of the relocation uncertainty along x, y, and z directions are shown in Figure 3.5. The prominent uncertainties are less than 0.2 km along x, y, and z directions. From the cumulative percentage, one can see that more than 80% events have the uncertainties of x and y directions less than 0.4 km. For the z (depth) direction, the uncertainty range is wider than the ones of the x and y directions. There are more than 80% events having the uncertainties in 1.2 km for z direction.

DD method relocation results with the 3D velocity model

The 3D DD relocation follows the 3D single event relocation to improve the relative locations. There were 14550 earthquake events that were clustered with 2 km hypocentral separation before the DD method relocation, and 13818 earthquakes were relocated in this step (Fig. 3.6). The seismicity swarms show a sharpening of the clustered earthquakes compared to the single event method relocation results. About 5% events with high residuals were discarded during the relocation process. The RMS traveltime residuals reduce to about 0.26 s after 9 iterations (Fig. 3.7(a)). Comparing to the original catalog data, most of the earthquakes show the reduced RMS traveltime residuals within 0.2 s after relocation (Fig. 3.7(b)), and focal depths show a symmetric distribution centered at about 5 km (Fig. 3.7(c)). We estimate the uncertainty of the earthquake locations by employing 50 bootstrap tests with 80% resampled data. The location uncertainty is ~40 m horizontally and ~100 m vertically (Fig. 3.8).

DD method relocation results with the 1D velocity model

Although the 3D velocity model can improve the absolute earthquake locations and relative locations, the DD relocation with the 1D velocity model still has important meaning for this study. There were 15756 cataloged earthquakes that were clustered, and 14866 clustered earthquakes were relocated with the 1D DD relocation. We noted that the number of the clustered earthquakes (15756) of 1D velocity relocation is more than that (14550) of 3D velocity relocation. The reason maybe that the single event method relocation with the 3D velocity model not only changes the epicenters, but also changes the focal depth for most of the earthquakes, which probably make the 2 km hypocentral separation can link fewer events to a cluster. About 5.6% earthquakes with

high residuals were discarded after this relocation. The relocated earthquakes are shown in Figure 3.9, which also present clear and narrow clusters. The shapes and orientations of the clusters are almost consistent with the 3D velocity relocated ones.

After 12 relocation iterations, the RMS traveltime residuals reduce to about 0.29 s (Fig. 3.10(a)), which is slightly higher than that of the 3D velocity relocation. The slightly higher RMS residuals with more iteration numbers may indicate that the 3D velocity model has the better relocation performance. Figure 3.10(b) shows the similar histogram of the residuals to Figure 3.7(b). Figure 3.10(c) shows more earthquakes around 5 km depth (Fig. 3.10(c)) than that in Figure 3.7(c), since the 3D velocity model already improved the earthquake hypocenters before the DD relocation.

Comparison between the 3D and 1D DD relocations

Although the 1D and the 3D velocity relocations show similar shapes and orientations of clusters, there are systematic shifts between the results. For comparison, we plotted differences between the 1D and the 3D velocity DD relocation results in Figure 3.11. The vectors point to the 3D velocity DD relocation from the 1D velocity DD relocation and are color coded for directions. Two large earthquake groups have different shifting directions (Figure 3.11). Most of the clusters in the southern group, which is between the NFZ and the WFZ, basically shift westward and northwestward (in greenish color). The clusters in the northern group to the west of the NFZ mainly shift southward (in reddish color). Figure 3.12 shows the histogram of the epicentral shift distances between the 1D and 3D velocity relocation results. The dominant shift distance is about 0.5 km, and the average shift amount is 0.7 km. More than 80% of the earthquakes shift less than 1 km. There are only a few events shifting greater than 2 km.

Figure 3.13 shows 8 sub-regions (black dashed boxes in Fig. 3.11) to examine cluster shifting in small scales.

Sub-region A in Figure 3.13a shows the relocated clusters near the town of Jones, OK. The 3D velocity relocated earthquakes (dots) generally present a systematic shifting orientation compared to the 1D velocity relocation (vector tails). (The symbols are the same for all the sub-regions in Figure 3.13a-h. The vectors point to the 3D DD relocation from the 1D DD relocation.) The prominent shifting directions for the earthquakes in this area are northwest, although there are a few earthquakes moving in the other directions. The clusters in this are do not seem to present clear orientations.

Sub-region B in Figure 3.13b shows the Prague sequence, which is the largest seismic sequence that had drawn a great deal of attention (e.g., Keranen, 2013; McNamara, 2015). Three moderate damaging earthquakes (Mw 4.8, 5.6, and 4.8) occurred in early November 2011 near Prague, OK. The aftershocks present a clear NE-SW trending orientation that very likely suggests a rupture, and an interpreted fault was inferred from this sequence. The earthquakes in the sequence tail show more consistent southwest shifting (in yellowish color). The northwestern earthquakes in the sequence head shift southward (in reddish color) to the fault, and the southeastern earthquakes in the sequence head primarily shift northwestward (in greenish color) to make the seismicity pattern sharpening.

Sub-region C in Figure 3.13c shows several clusters near Guthrie, OK, where an NE-SW trending fault is to the southeast of the clusters. Most of the earthquakes slightly shift to west (in greenish color). The largest cluster in this area with very consistent shifting presents a clear conjugate-planes pattern (NE-SW and NW-SE). The

cluster in the southwest corner is likely to show two parallel separated small clusters with the primary northwest shifting and a few other shifting directions. Several earthquakes at northeastern corner shift to the northeast (in bluish color).

There are two NE-SW trending clusters around Cushing, OK in Figure 3.13d. Two earthquake clusters in this sub-region D present fairly consistent northwest shifting orientations. The earthquakes around this region have tremendous meaning to the energy security of Unite States, where one of the world largest crude oil storage facilities is located in (McNamara et al., 2015). In October 2014, two M4+ earthquakes shook the Cushing oil hub that is a strategic infrastructure of Unite States. McNamara et al., (2015) interred a WNW-ESE fault based on their earthquake relocation study for the southern cluster and made a series of analysis. However, our relocation result shows NE-SW trending orientations for it, which is different than McNamara et al., (2015). We think that more studies may be necessary for further details to demonstrate the fault structures in this area.

Sub-region E in Figure 3.13e is close to the gap between two large earthquake groups. Most earthquakes of two obvious clusters in this sub-region shift southwestward (in orange and yellowish colors). These two clusters also present NE-SW and NW-SE orientations, which are consistent with the maximum horizontal stress state in Oklahoma. Only a few earthquakes shift to other directions.

Several parallel NE-SW trending clusters in sub-region F are shown in Figure 3.13f. This area is located in the large northern earthquake group. Most of the earthquakes in this area shift southward (in reddish color). The consistent cluster orientations may indicate similar fault strike orientations in this sub-region.

Similar to sub-region F, clusters in sub-region G also present systematic southward shifting (in reddish color) in Figure 3.13g. Some clusters show NE-SW trending orientations, and some of them show NW-SE orientations, which are close to nearby parallel fault segments.

Sub-region H in Figure 3.13h shows an NE-SW trending cluster in northwestern Oklahoma. The earthquakes in this cluster primarily shift eastward (in purplish color). The cluster aligns well with the mapped fault to the southwest. This fault may be blamed for the most moderate earthquakes in this cluster occurred in early 2016.

Discussion

Demonstrating the relationship between the earthquakes and geological structures in Oklahoma is important, because the seismicity is very active in this region after 2009. The 3D velocity model can improve the earthquake locations, which helps to correlate the seismicity with the mapped faults, interpret unmapped faults, and determine the connection between the earthquakes and the wastewater injection wells. Some of the relocated clusters present narrower and linear trends, which may indicate the potential faults that related to the clusters. Therefore, the correlation between the earthquake locations and faults may be used to decide where to build large-scale structures with low ricks, such as large-rise facilities, oil storage infrastructures, pipelines, and others. In addition, the correlation may provide information to evaluate insurance rates of residential buildings in Oklahoma as well (Luza and Lawson, 1982).

Most of the clusters present narrower linear trends with NE-SW or NW-SE orientation after the 3D velocity relocation, which are consistent with the favorable fault orientations in central Oklahoma (Holland, 2013). Several studies suggest that recent

earthquakes in Oklahoma were caused by reactivation of the ancient faults, which cut through Arbuckle Group and extend into the upper basement (e.g., Holland, 2013; Alt and Zoback, 2014; McNamara et al., 2015). The Prague sequence was studied with a variety of focal mechanism data to show that the most of the earthquakes in Prague sequence have strike-slip motion, which is consistent with the knowledge that Wilzetta Fault is a vertical or near vertical fault, at least in sedimentary section and upper basement (Holland, 2013; Sumy et al., 2014; McNamara et al., 2015). However, not all of clusters have been well studied in Oklahoma. For many relocated clusters, there is no identified fault nearby, but some of the clusters may be associated with large geological structures in central Oklahoma, since they align well with mapped faults. Due to the strong heterogeneity in the basement (Chapter 2), basement faults may be complicated that are difficult to associate with the seismicity increase.

Conclusions

We relocated the earthquakes in central Oklahoma and the surrounding regions from the 01/01/2010 to 03/31/2015 with 3D velocity model. Firstly, we applied the single event method to improve the absolute locations with the 3D velocity model. Next, the DD method was used to improve the relative locations with the 3D velocity model. We also relocated the earthquakes with the DD method using the 1D velocity model for comparison to determine to the location shifts. The majority of the earthquakes shift less than 1 km, and the average shift amount is 0.7 km over the whole region. Most of the clusters present NE-SW or NW-SE orientations, which are consistent with maximum horizontal stress in the mid-continent (Zoback, 1992). Some clusters show a clear correlation with fault splays, or align with the mapped faults. For

some clusters, there are no close faults to correlate, which may suggest new inferred faults to support the fault database. Some of the clusters show shallow depth distributions in the sedimentary section.

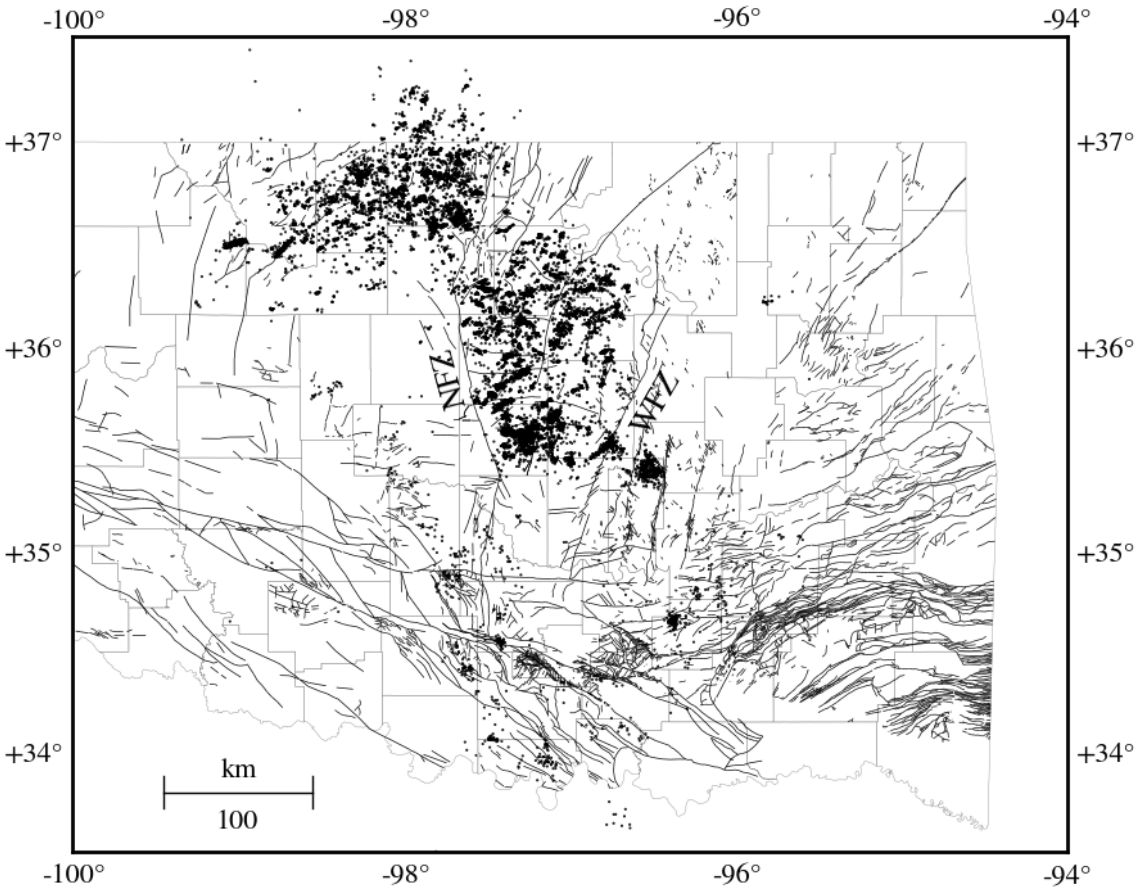


Figure 3.1. Map view of the initial cataloged earthquakes (dots) from 01/01/2010 to 03/31/2016 in central Oklahoma and the surrounding region. The cataloged earthquakes were located with the SEISAN package (Havskov and Ottemoller, 1999) by the OGS. Gray polygons are the county boundaries. The black lines are the preliminary Oklahoma faults (Holland, 2015). NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone.

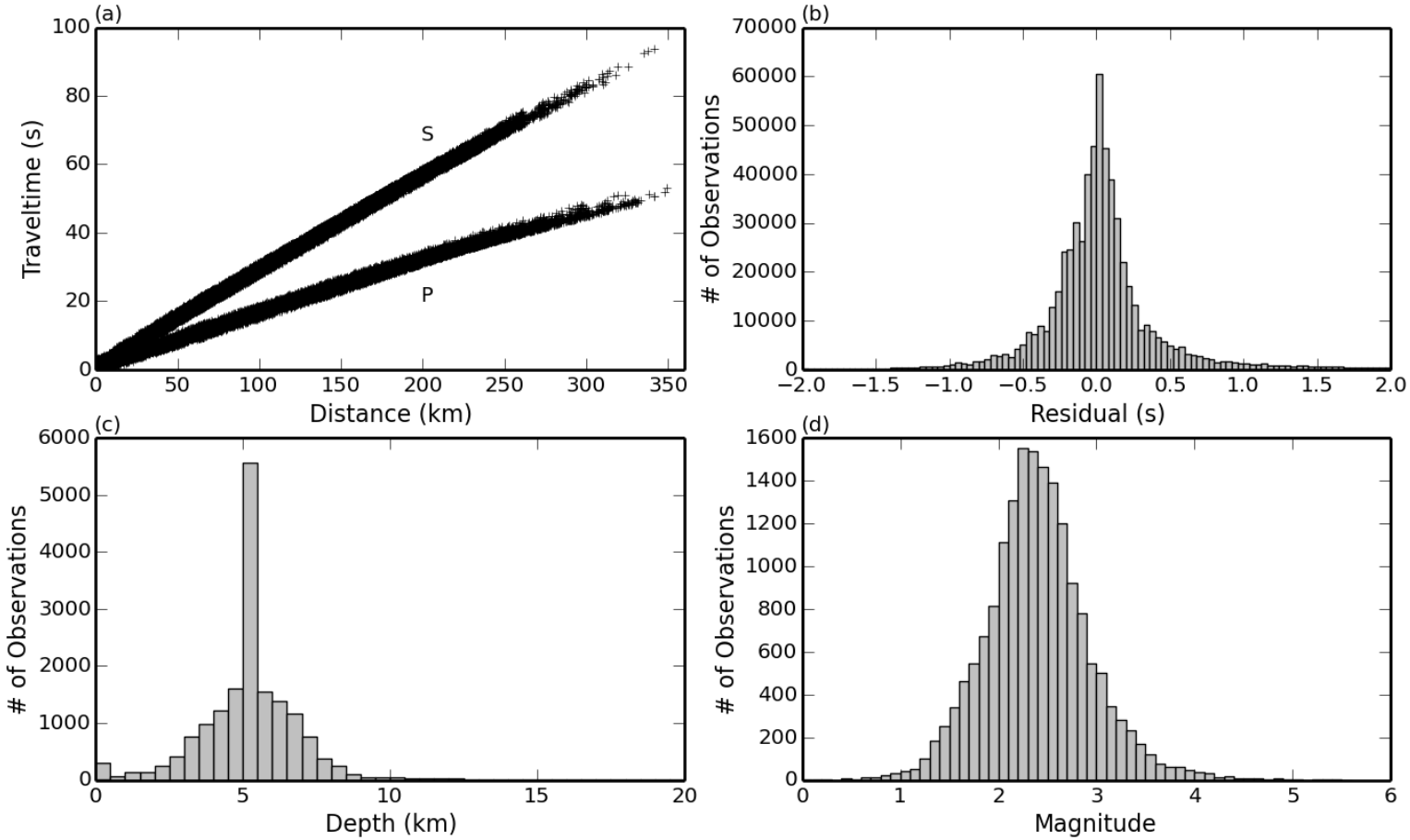


Figure 3.2. Statistics of selected 17385 earthquakes after bad pick filtering. (a) Filtered P- and S-picks travel-time vs distance. (b) Histogram of residuals. (c) Histogram of focal depth. (d) Histogram of magnitude.

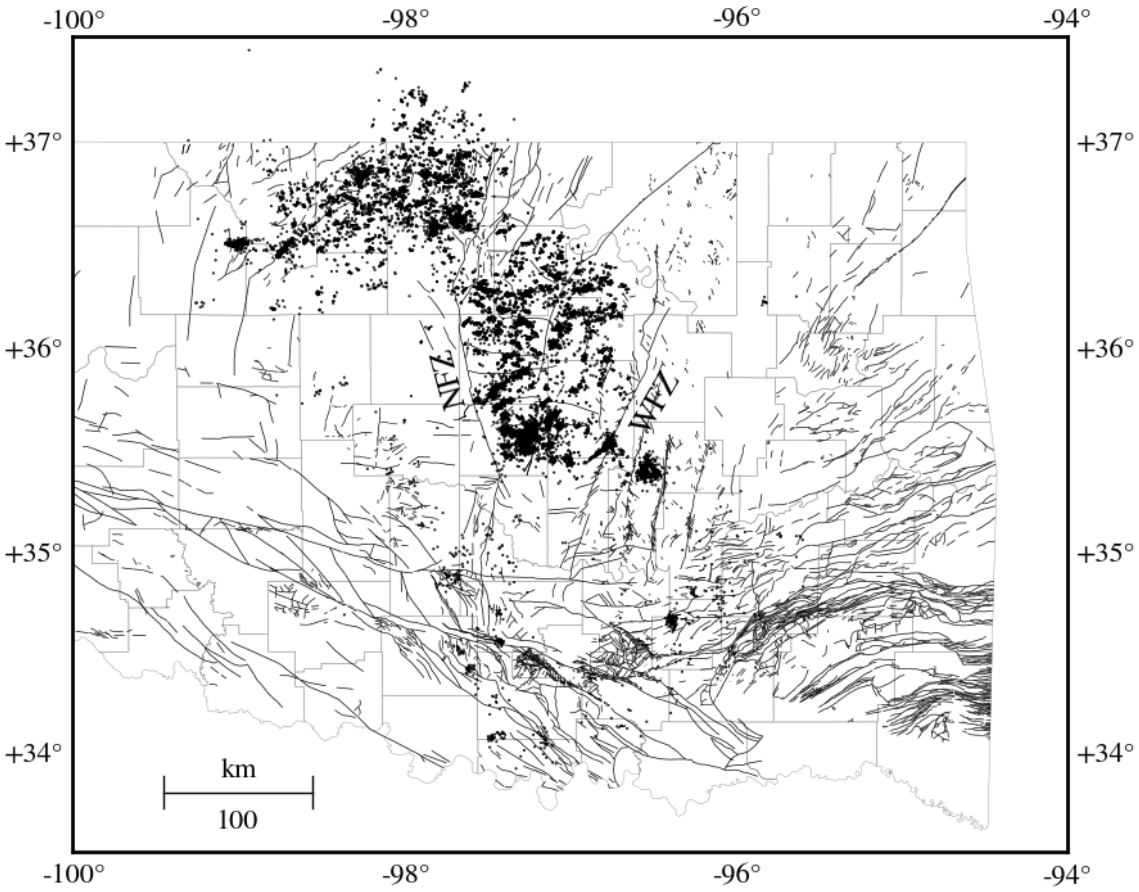


Figure 3.3. Map view of the relocated earthquakes (dots) with the single event method from the SIMUL2000. Gray polygons are the county boundaries. The black lines are the preliminary Oklahoma faults (Holland, 2015). NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone.

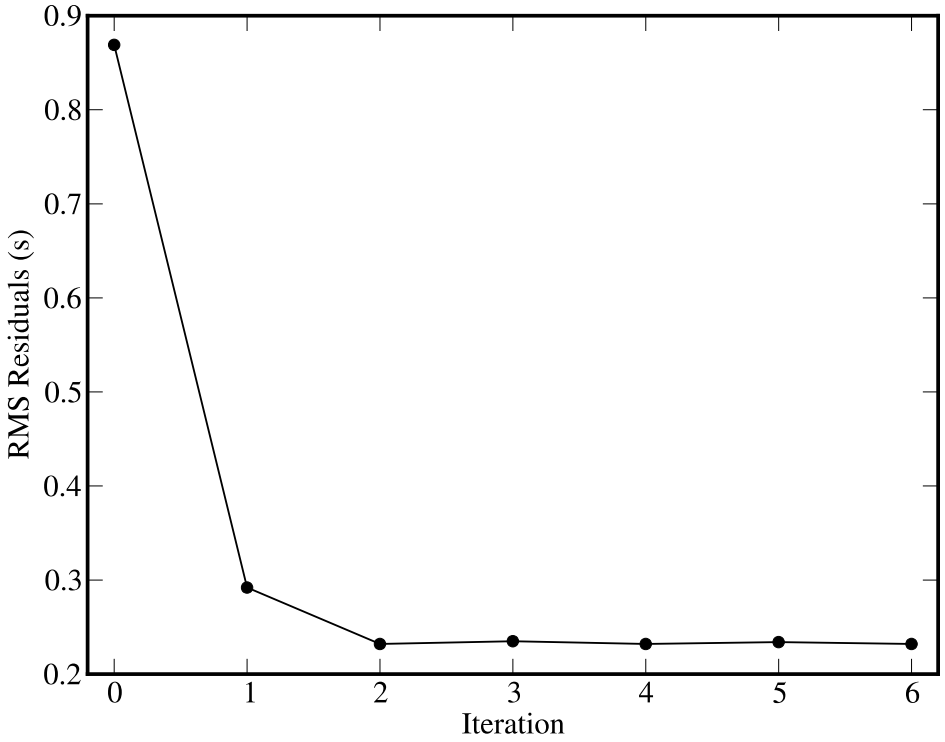


Figure 3.4. Reduction of the RMS traveltime residual with iterations of the single event method relocation (SIMUL2000).

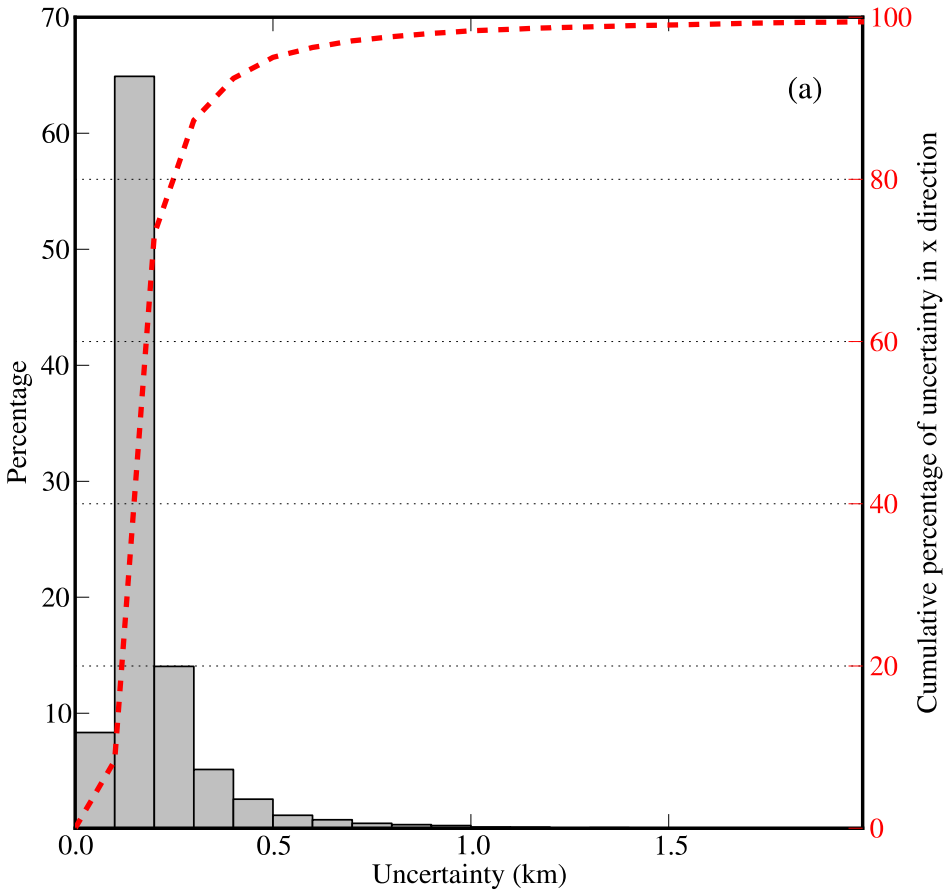


Figure 3.5a. The normalized uncertainty from the single event method relocation along x-axis. Red dashed line: cumulative percentage on the x-axis (uncertainty).

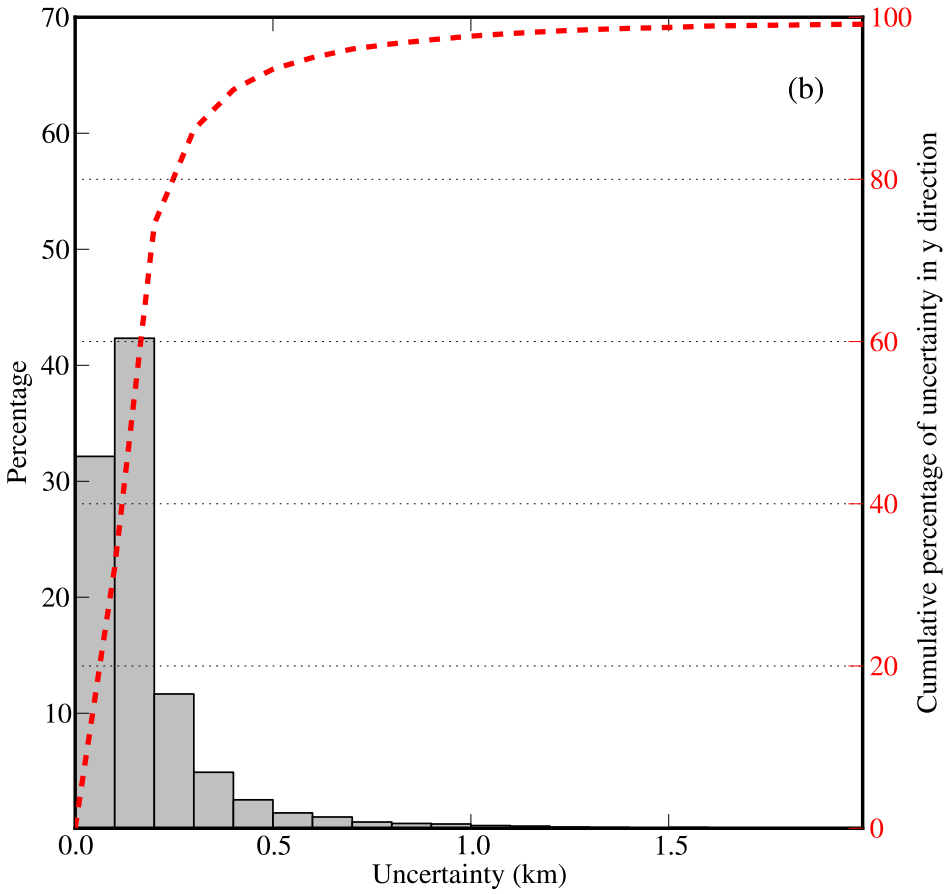


Figure 3.5b. The normalized uncertainty from the single event method relocation along y-axis. Red dashed line: cumulative percentage on the y-axis (uncertainty).

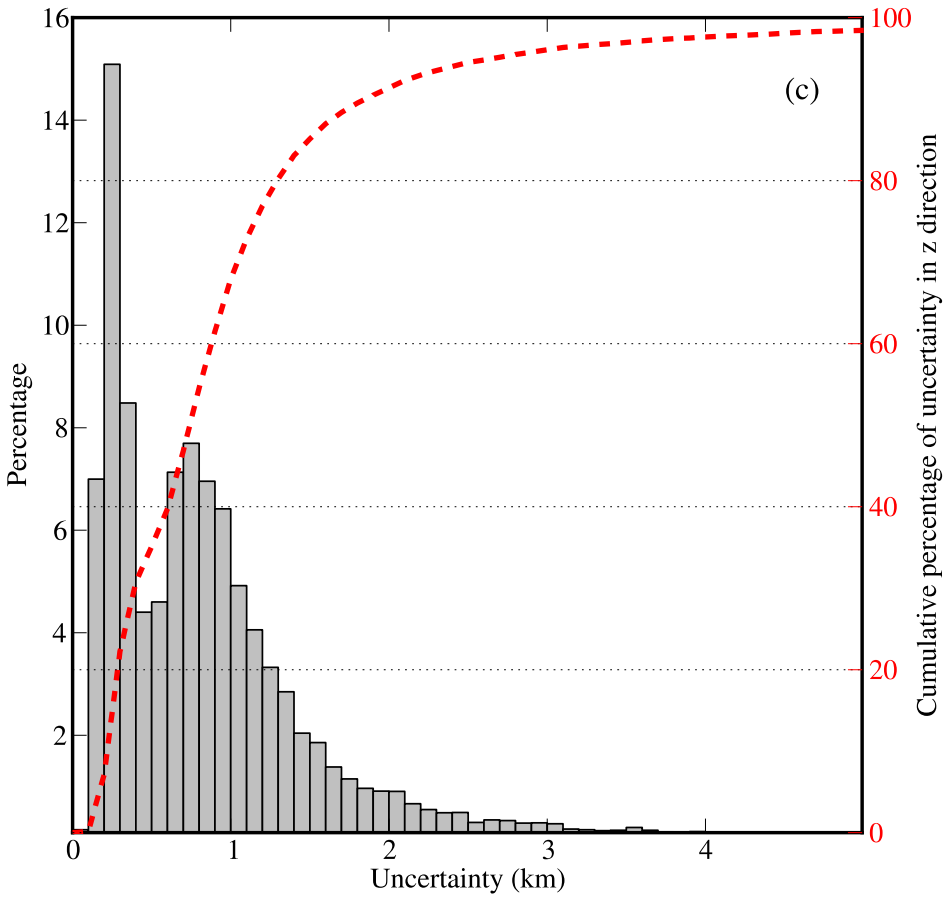


Figure 3.5c. The normalized uncertainty from the single event method relocation along z-axis. Red dashed line: cumulative percentage on the z-axis (uncertainty).

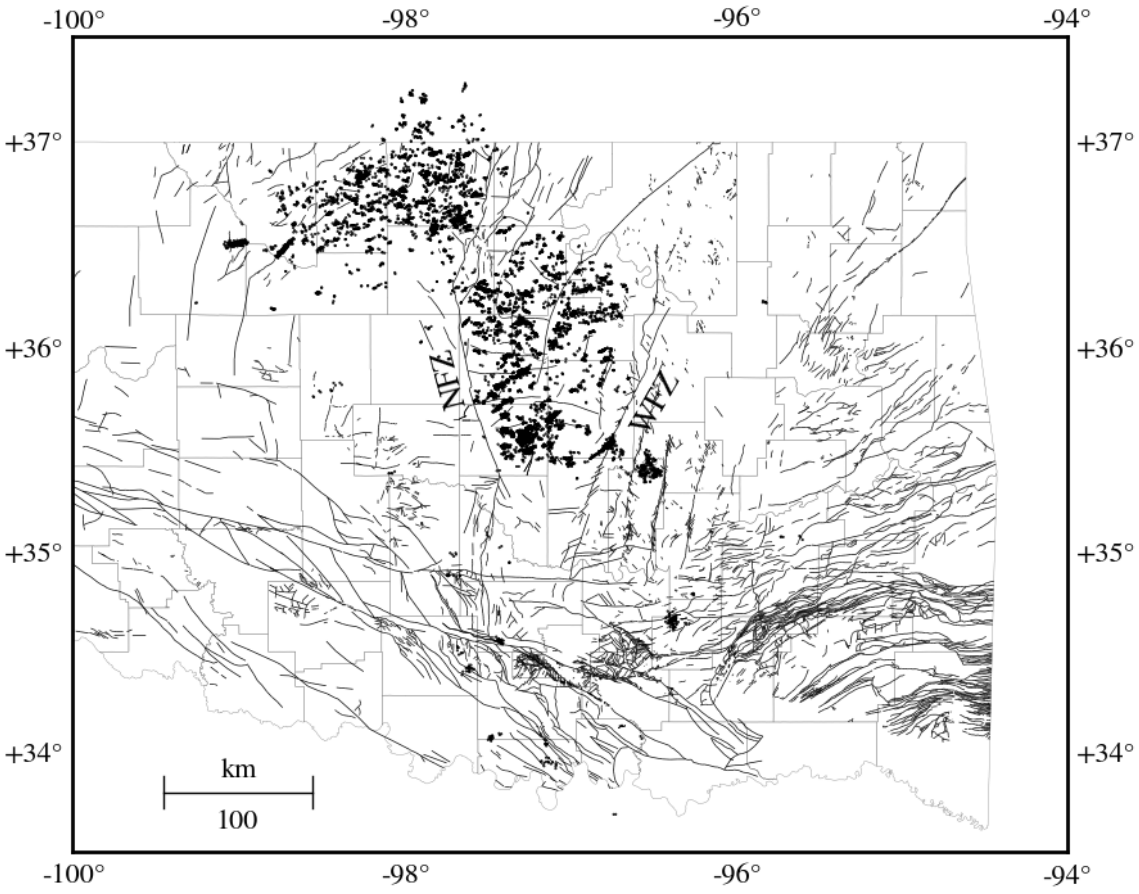


Figure 3.6. Map view of the relocated earthquakes (dots) with the 3D DD method. Gray polygons are the county boundaries. The black lines are the preliminary Oklahoma faults (Holland, 2015). NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone.

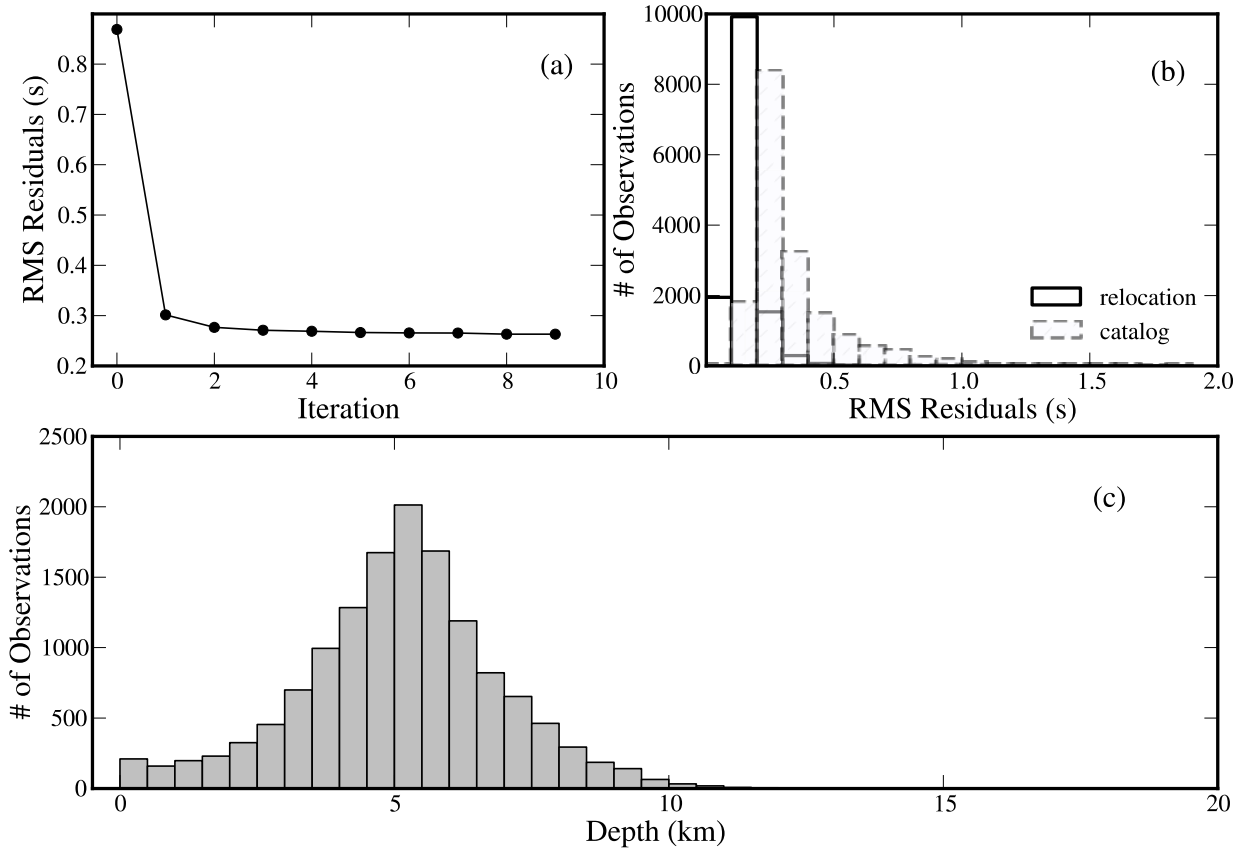


Figure 3.7. (a) Reduction of the RMS traveltime residual with iterations for the 3D DD method. (b) Histograms of residual distribution for initial cataloged data (shaded) and relocation data (white). (c) Histogram of the depth after the relocation.

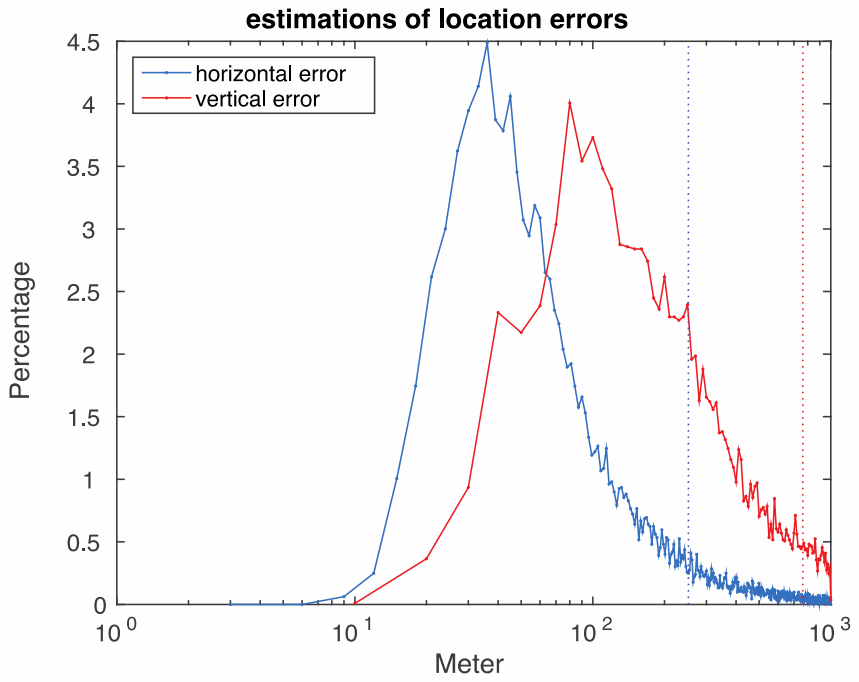


Figure 3.8. The normalized uncertainty of the 3D DD method relocation with 50 bootstrap tests.

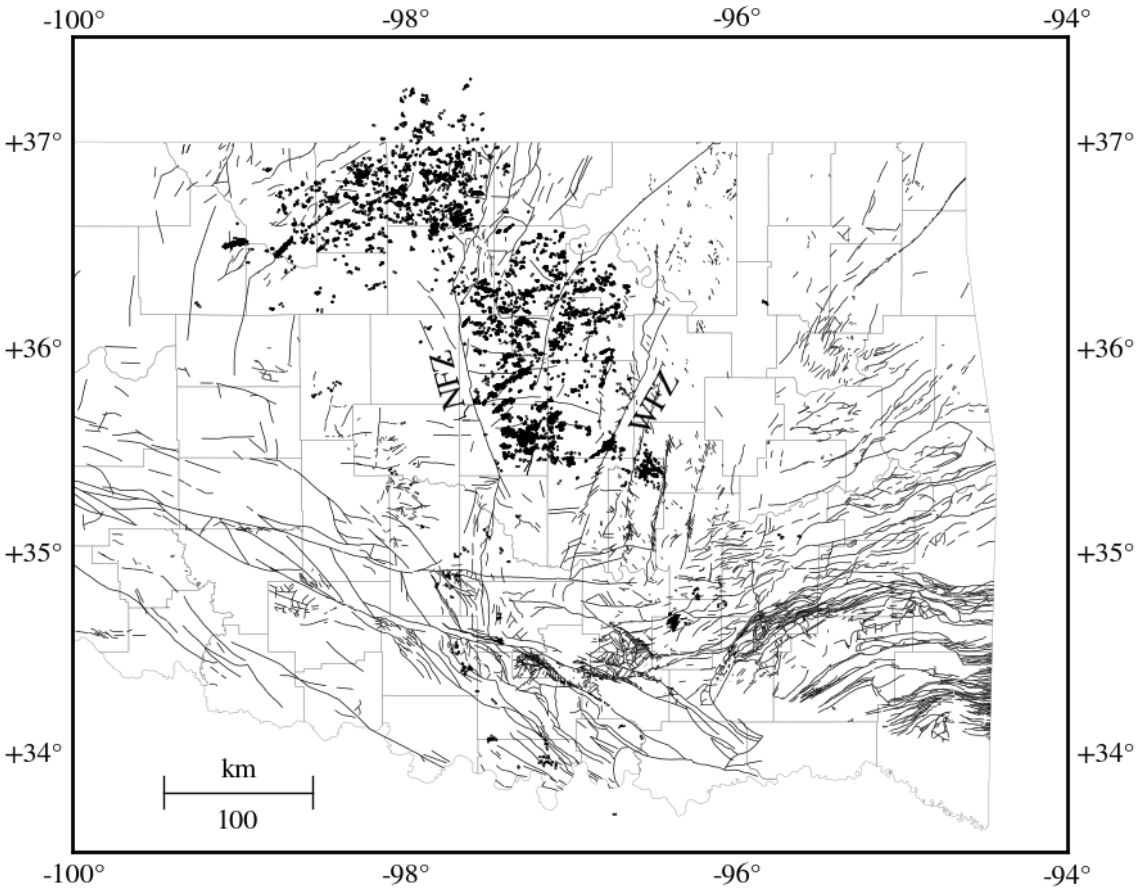


Figure 3.9. Map view of the relocated earthquakes (dots) with 1D DD method. Gray polygons are the county boundaries. The black lines are the preliminary Oklahoma faults (Holland, 2015). NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone.

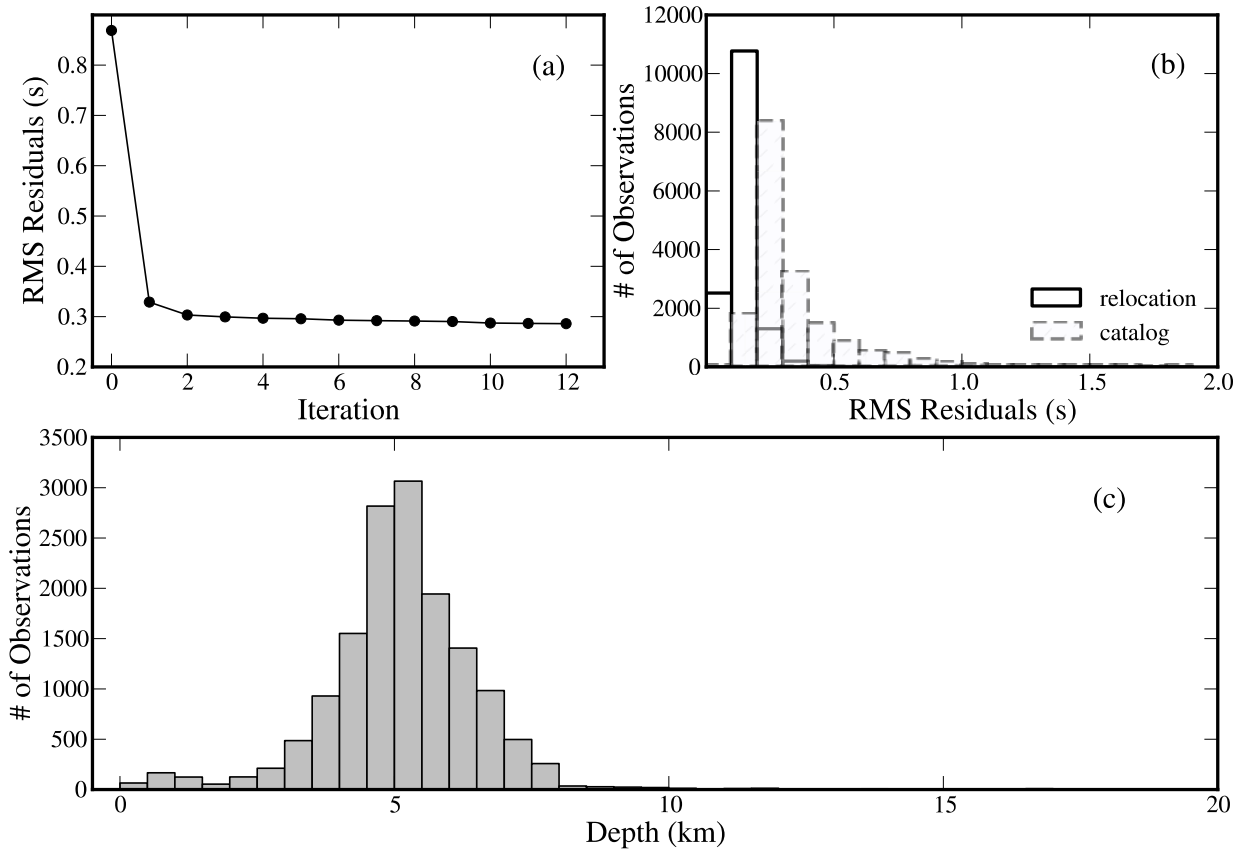


Figure 3.10. (a) Reduction of the RMS traveltime residual with iterations for the 1D DD method. (b) Histograms of residual distribution for initial cataloged data (shaded) and relocation data (white). (c) Histogram of the depth after the relocation.

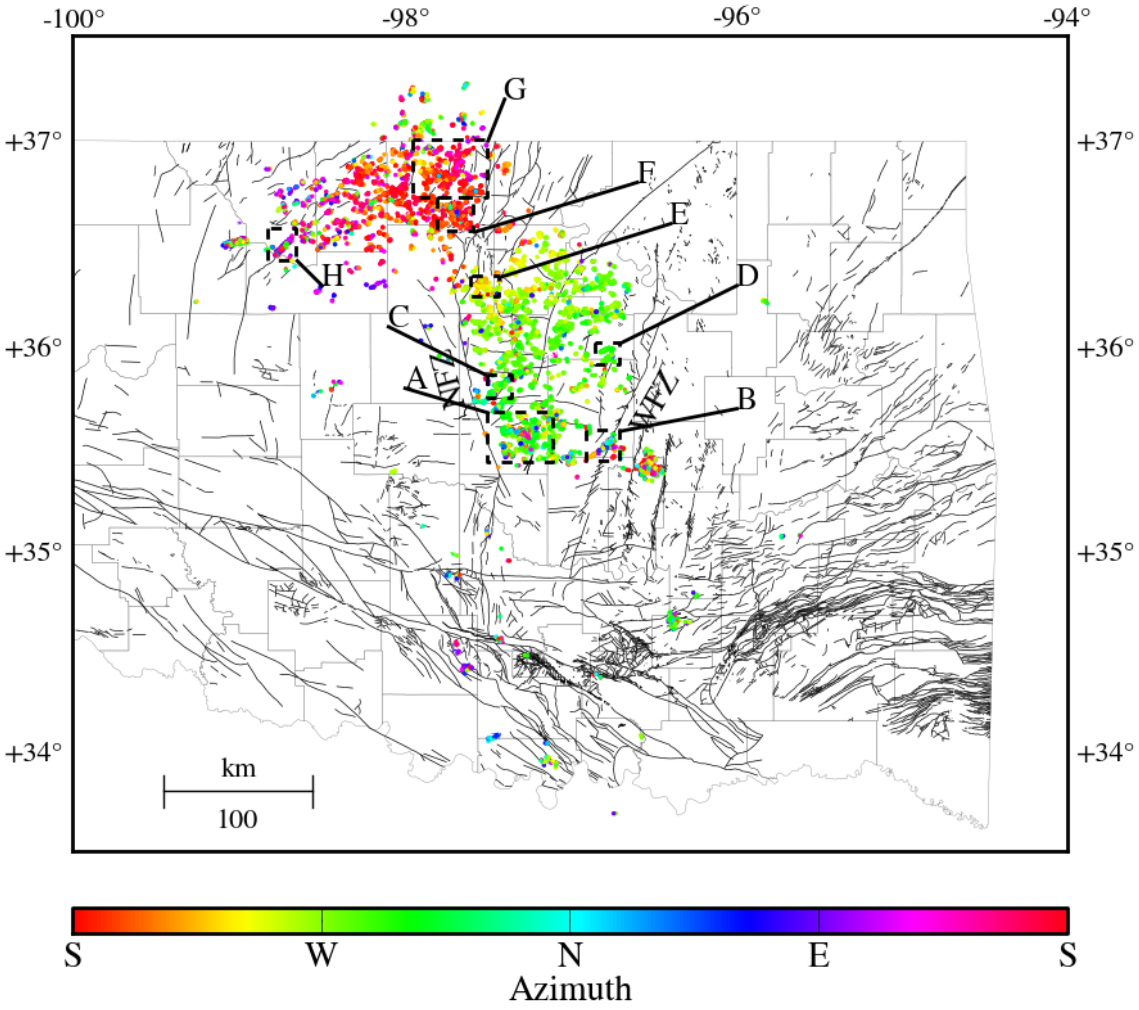


Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results. The colored vectors (very small in this scale) point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. Eight sub-regions (A-H) are shown in Figure 3.13 to examine cluster shift in the small scale. Gray polygons are the county boundaries. The black lines are the preliminary Oklahoma faults (Holland, 2015). NFZ: Nemaha Fault Zone; WFZ: Wilzetta Fault Zone.

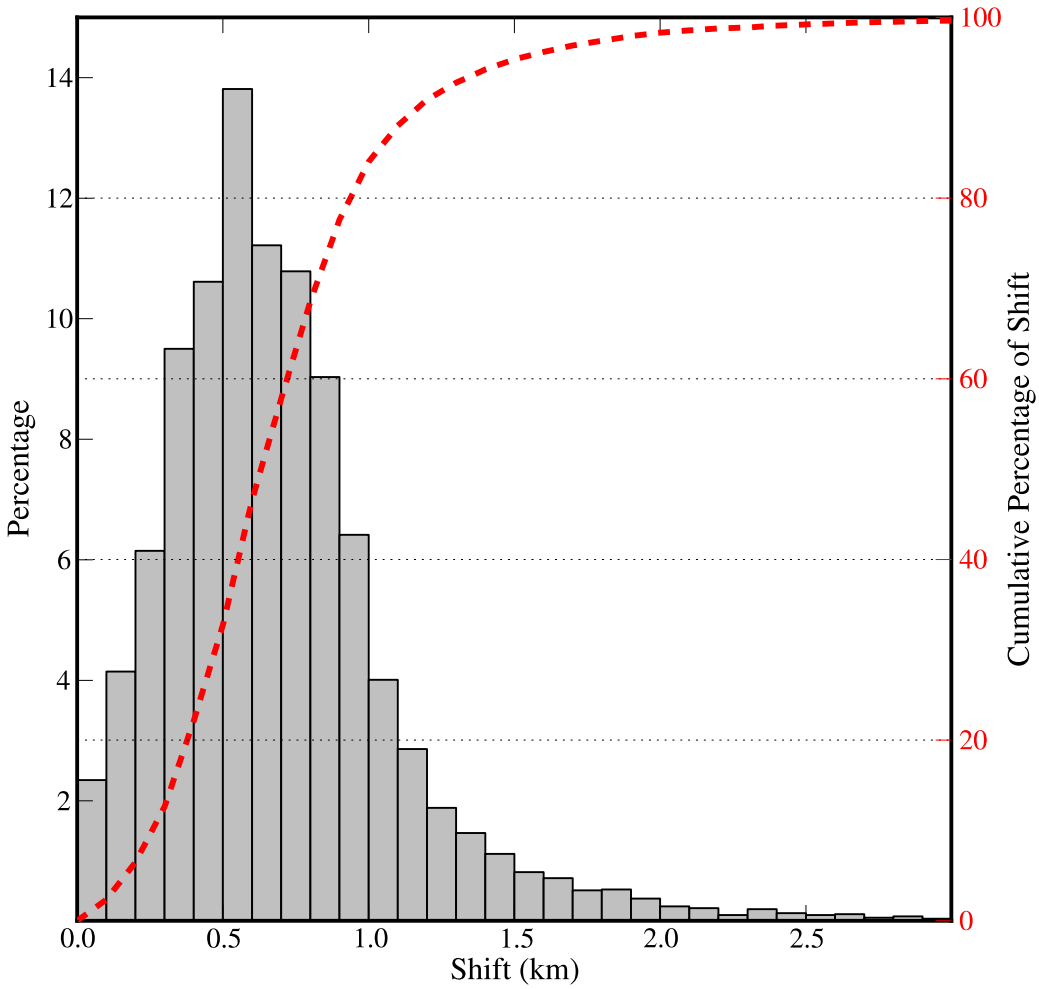


Figure 3.12. Histogram of the epicentral shift distances between the 1D and the 3D velocity relocations.

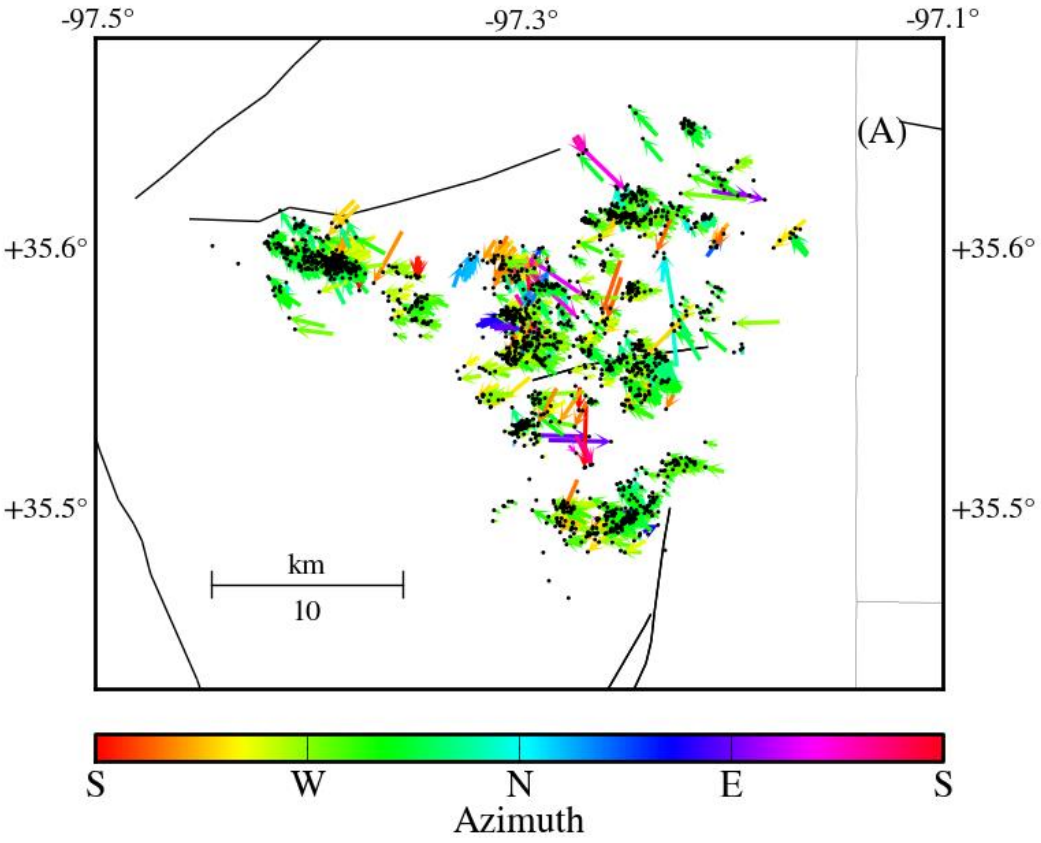


Figure 3.13a. Sub-region A in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

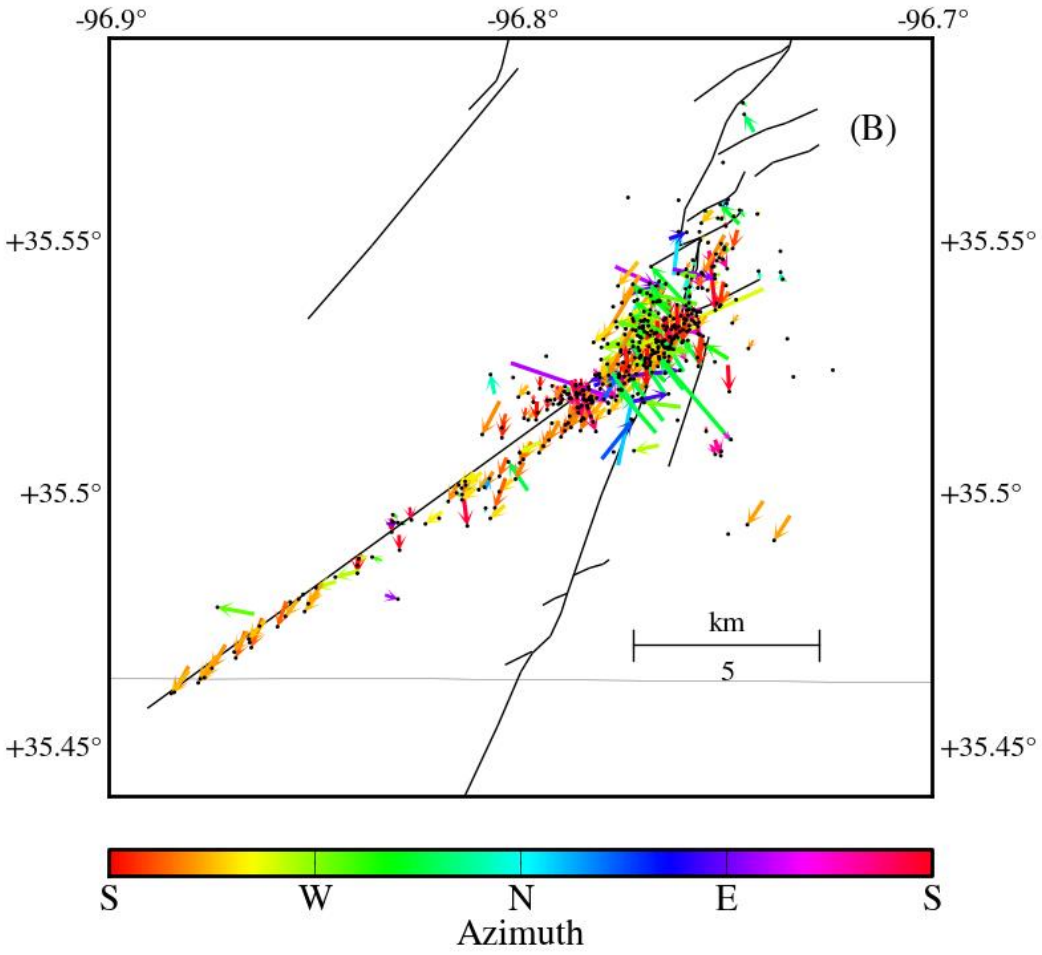


Figure 3.13b. Sub-region B in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

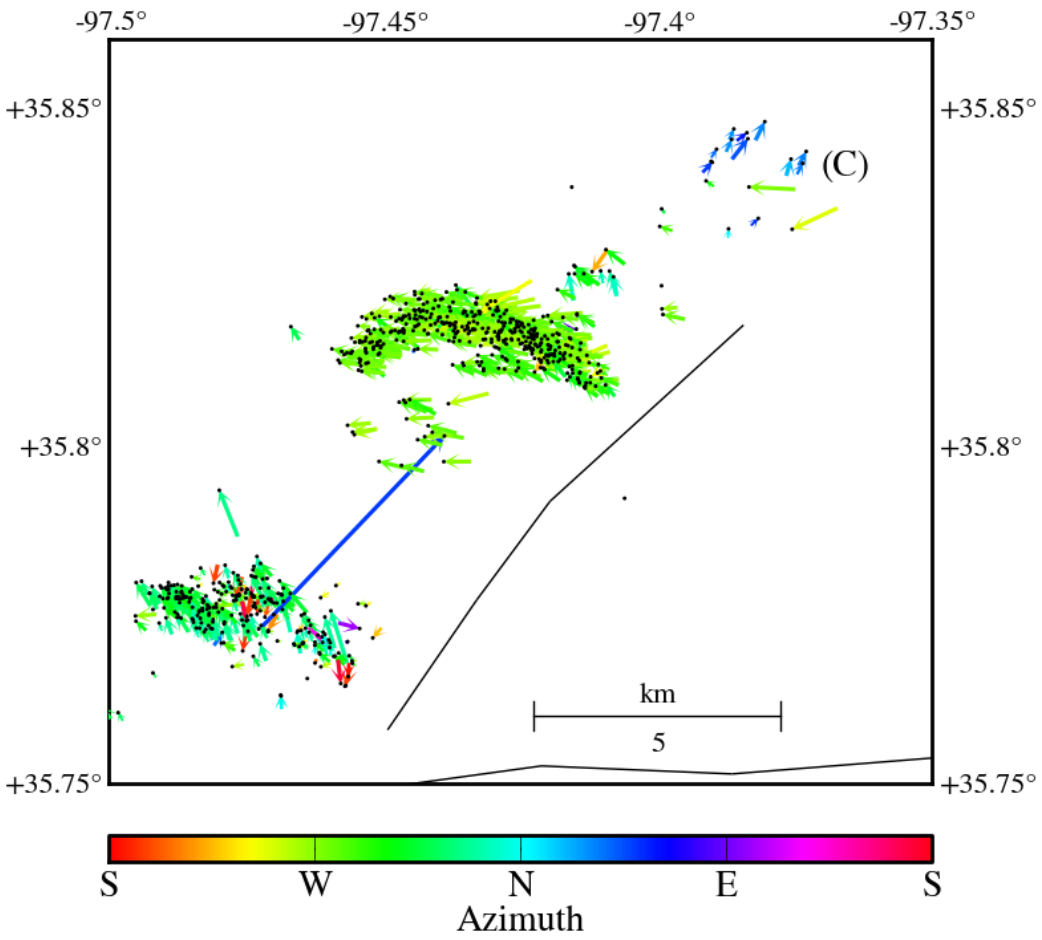


Figure 3.13c. Sub-region C in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015).

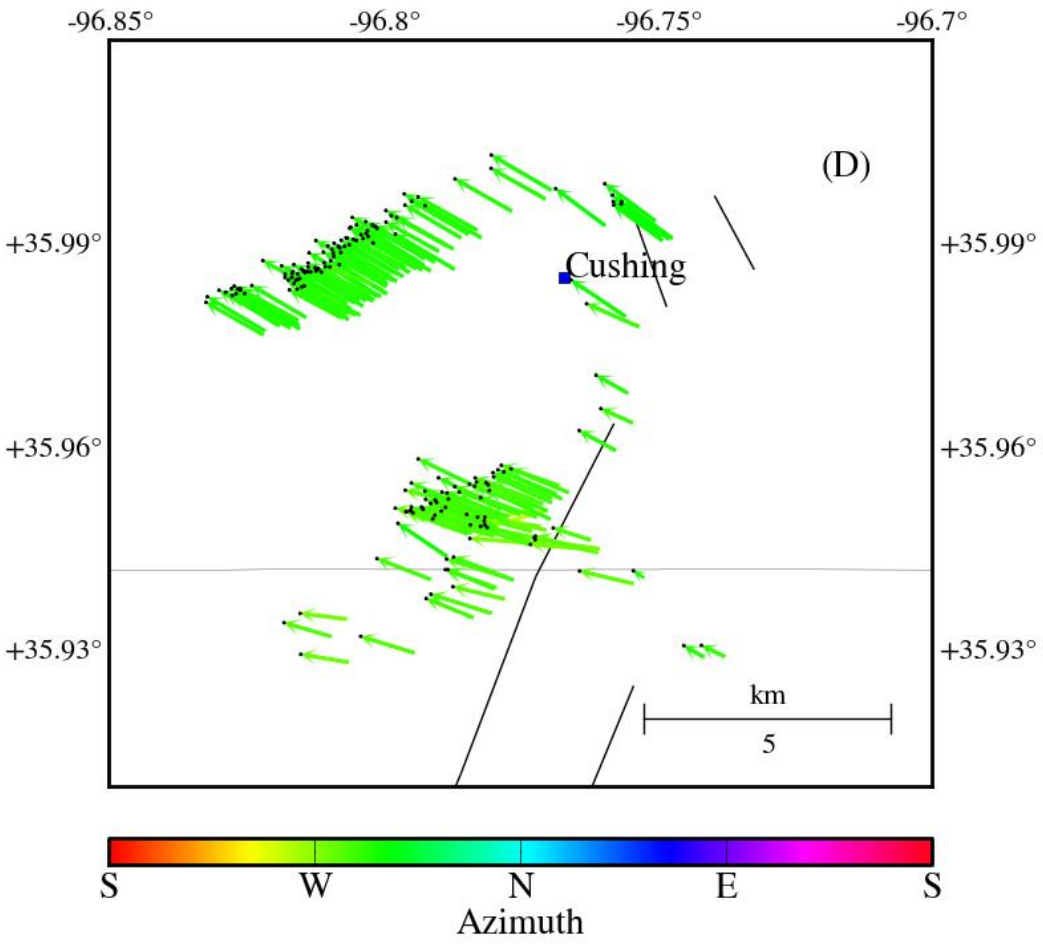


Figure 3.13d. Sub-region D in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

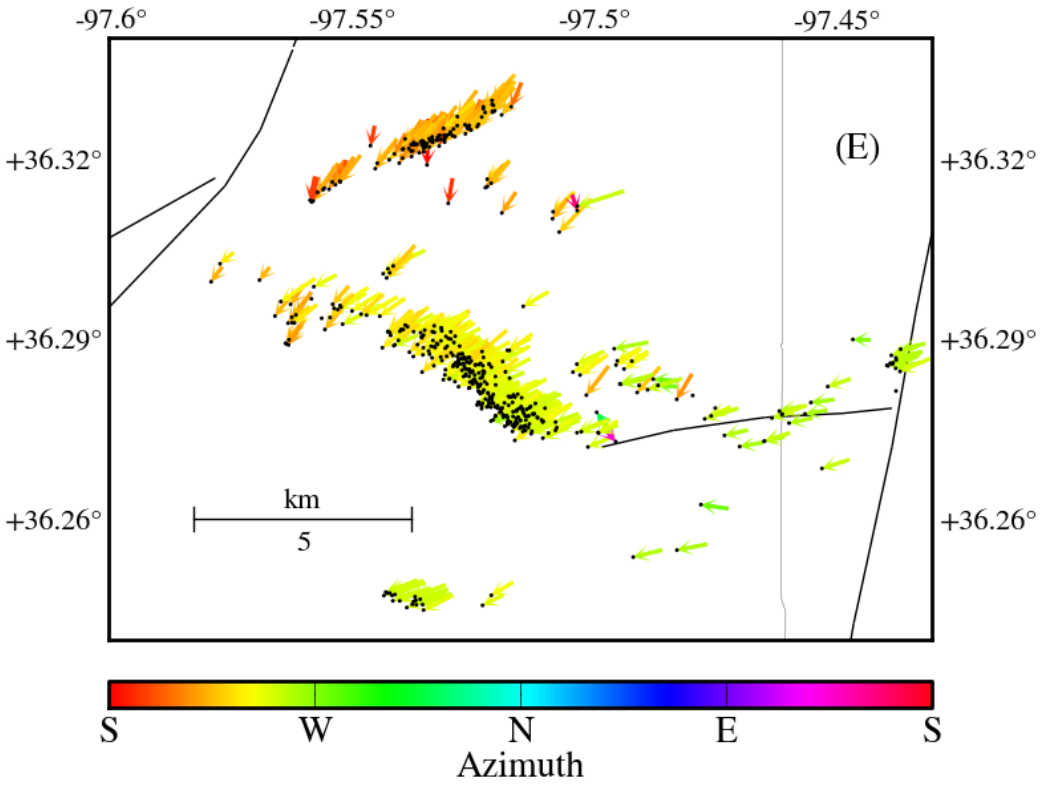


Figure 3.13e. Sub-region E in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

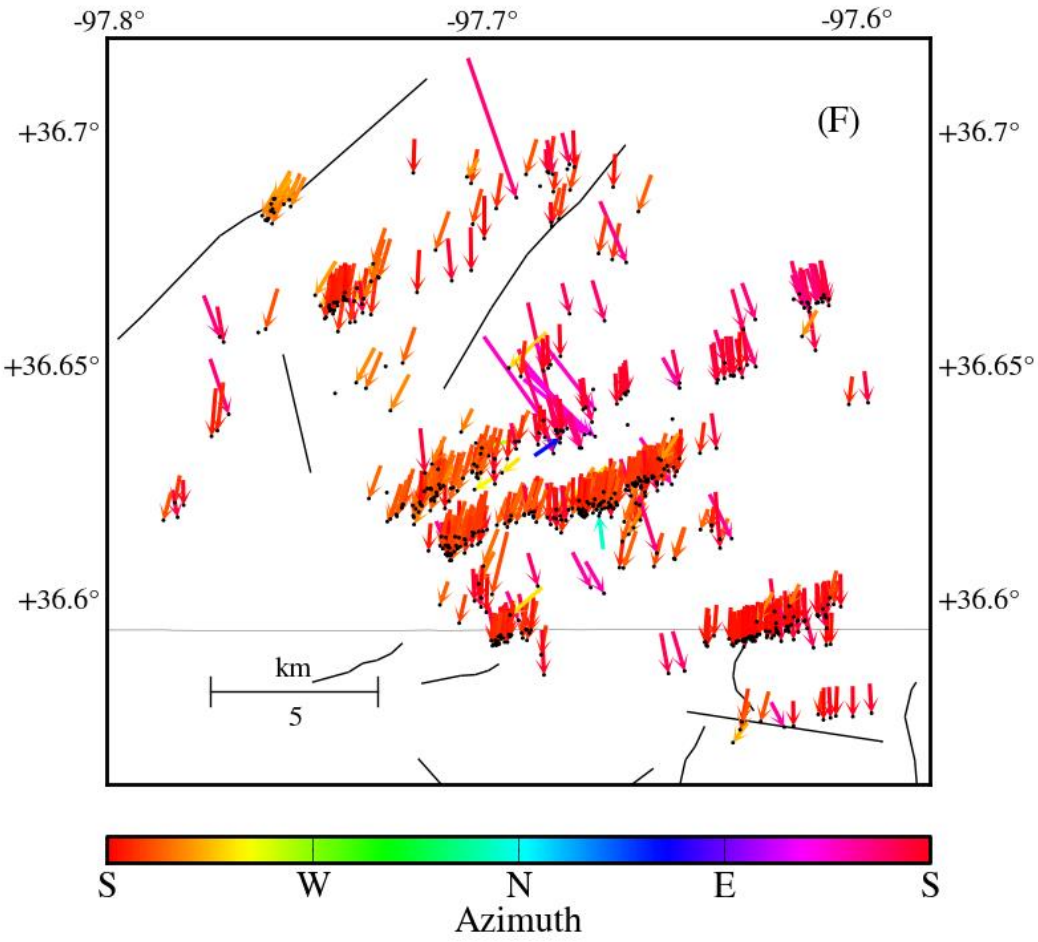


Figure 3.13f. Sub-region F in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

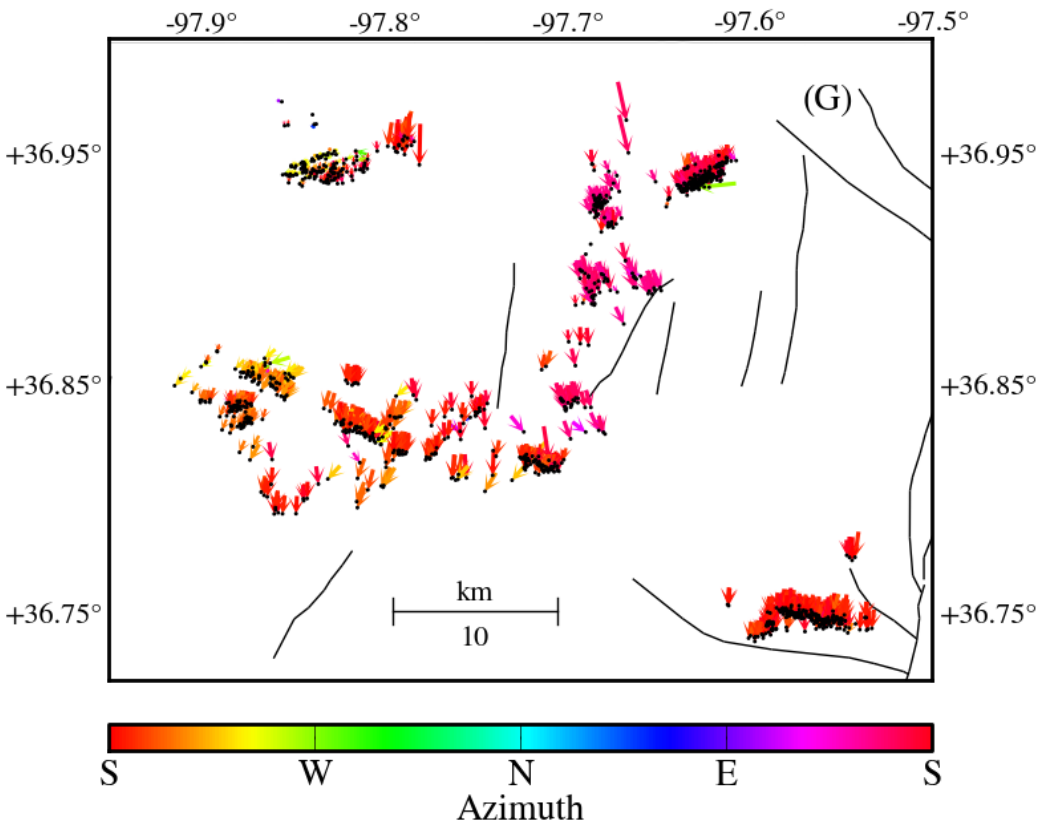


Figure 3.13g. Sub-region G in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015).

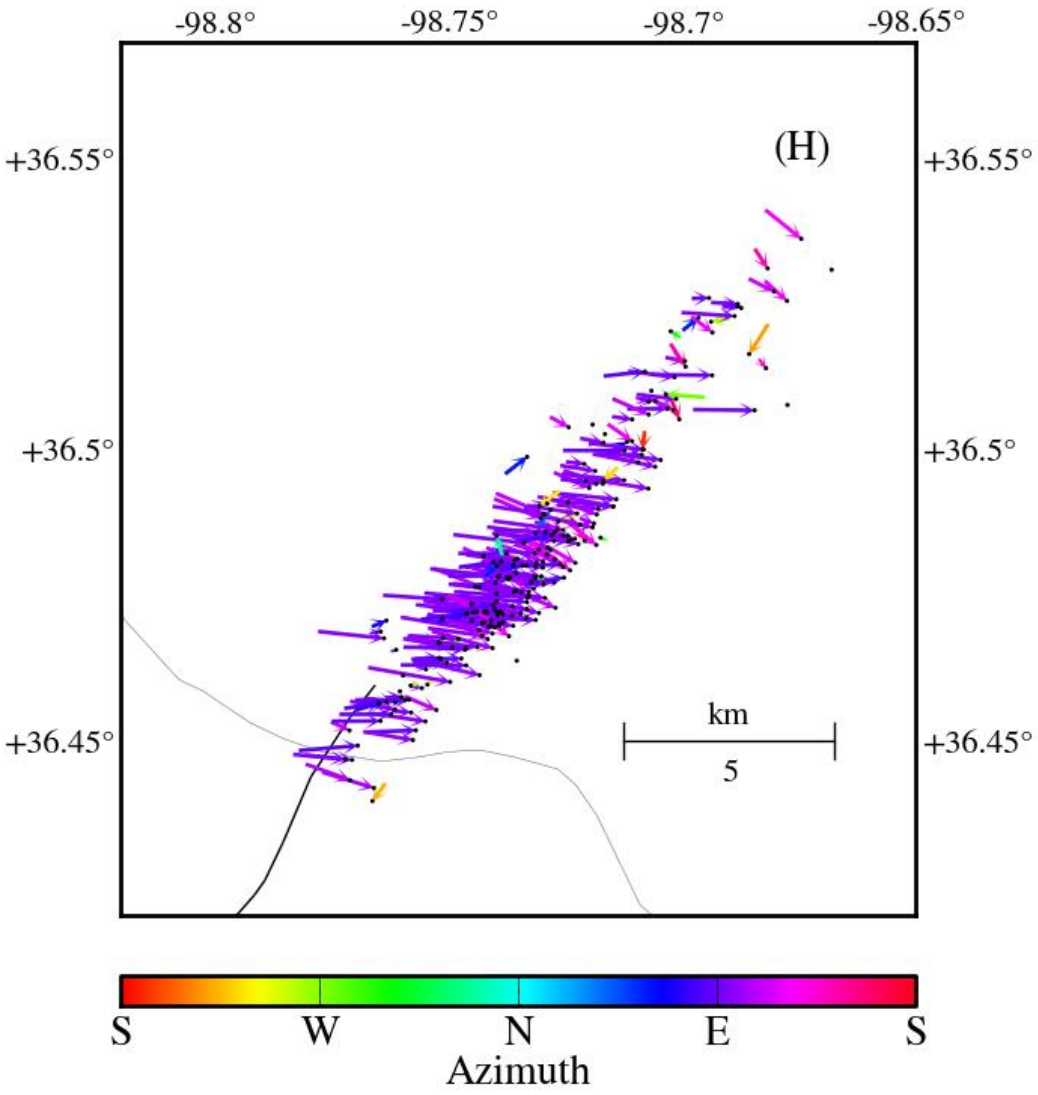


Figure 3.13h. Sub-region H in Figure 3.11. Epicentral shifts between the 1D and 3D velocity DD relocation results (black dots). The colored vectors point to the 3D velocity relocated earthquakes from the same events of 1D velocity relocation. Color indicates the shift azimuth. The thick black lines are the preliminary Oklahoma faults (Holland, 2015). The gray lines are the county boundaries.

References

Alt, R., and M. Zoback (2014), Development of a detailed stress map of Oklahoma for avoidance of potentially active faults when siting wastewater injection wells, AGU Fall Meeting 2014, abstract.

Darold, A. P., A. A. Holland, J. K. Morris, and A. R. Gibson (2015), Oklahoma earthquake summary report 2015, Oklahoma Geological Survey Open-File Report, OF1-2015.

Darold, A. P., and A. A. Holland (2015), Preliminary optimal Oklahoma fault orientations, Oklahoma Geological Survey Open-File Report, OF4-2015.

Douglas, A. (1967), Joint epicentre determination, Nature, **215**, 47–48.

Geiger, L. (1912), Probability method for the determination of earthquake epicenters from the arrival time only (translated from Geiger's 1910 German article), Bull. St. Louis University, **8**, 56-57.

Havskov, J., and L. Ottemoller (1999), SeisAn earthquake analysis software, Seismological Research Letters, **70**, 532-534.

Holland, A. A. (2013), Optimal fault orientations within Oklahoma, Seismol. Res. Lett., **84**, 876-890, doi: 10.1785/022012015.

Holland, A. A. (2015), Preliminary fault map of Oklahoma, 1 plate with supplement, Oklahoma Geological Survey Open-File Report, OF3-2015.

Jordan, T. H., and Sverdrup, K. A. (1981), Teleseismic location techniques and their application to earthquake clusters in the South-Central Pacific, Bulletin Seismological Society of America, **71**, 1105–1130.

Keranen, K. M., H. M. Savage, G. A. Abers, and E. S. Cochran (2013), Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 MW 5.7 earthquake sequence, Geology, 41, 699-702, doi:10.1130/G34045.1.

Lin, G., P. M. Shearer, E. Hauksson, and C. H. Thurber (2007), A three-dimensional crustal seismic velocity model for southern California from a composite event method, Journal of Geophysical Research, **112**, B11306, doi:10.1029/2007JB004977.

Lomax, A., J. Virieux, P. Volant, and C. Berge (2000), Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations, in Advances in Seismic Event Location, edited by C. H. Thurber and N. Rabinowitz, 101–134, Kluwer, Amsterdam.

Luza, K. V., and J. E. Lawson (1982), Seismicity and tectonic relationships of the Nemaha Uplift in Oklahoma, part IV, Oklahoma Geological Survey Special Publication, 82-1.

McNamara, D. E., H. M. Benz, R. B. Herrmann, E. A. Bergman, P. Earle, A. Holland, R. Baldwin, and A. Gassner (2015), Earthquake hypocenters and focal mechanisms in central Oklahoma reveal a complex system of reactivated subsurface strike-slip faulting, Geophys. Res. Lett., **42**, 2742–2749, doi:10.1002/2014GL062730.

McNamara, D. E., G.P. Hayes, H. M. Benz, R. A. Williams, N. D. McMahon, R. C. Aster, A. Holland, T. Sickbert, R. Herrmann, R. Briggs, G. Smoczyk, E. Bergman, and P. Earle (2015), Reactivated faulting near Cushing, Oklahoma: Increased potential for a triggered earthquake in an area of United States strategic infrastructure, Geophys.

Res. Lett., **42**, 8328–8332, doi:10.1002/2015GL064669.

Moser, T. J., T. van Eck, and G. Nolet (1992), Hypocenter determination in strongly heterogeneous Earth models using the shortest path method, Journal of Geophysical Research, **97**, 6563-6572, doi:10.1029/91JB03176.

Reid, H. F (1910), The mechanics of the Earthquake, Vol. II of the California earthquake of April 18, 1906, report of the State Earthquake Investigation Commission, A. C. Lawson, Chairman, Carnegie Institution of Washington Publication 87.

Richards-Dinger, K. B., and Shearer, P. M. (2000), Earthquake locations in Southern California obtained using source-specific station terms, J. Geophys. Res. Solid Earth, **105**, 10939–10960.

Sumy, D. F., E. S. Cochran, K. M. Keranen, M. Wei, and G. A. Abers (2014), Observations of static Coulomb stress triggering of the November 2011 M5.7 Oklahoma earthquake sequence, *J. Geophys. Res. Solid Earth*, **119**, 1904–1923, doi:10.1002/2013JB010612.

Thurber, C. H. (1983), Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, Journal of Geophysical Research, **88**, 8227-8326, doi:10.1029/JB088iB10p08226.

Waldhauser, F., and W. L. Ellsworth (2000), A double-difference earthquake location algorithm: method and application to the northern Hayward Fault, California, Bulletin Seismological Society of America, **90**, 1353-1368, doi:10.1785/0120000006.

Waldhauser, F. (2001), HypoDD: a computer program to compute double-difference hypocenter locations, U.S. Geological Survey Open File Report 01-113.

Zhang, H. J., and C. H. Thurber (2003), Double-difference tomography: The method and its application to the Hayward Fault, California, Bulletin Seismological Society of America, **93**, 1875–1889, doi:10.1785/0120020190.

Zoback, M. L. (1992), First- and second-order patterns of stress in the lithosphere: The world stress map project, J. Geophys. Res., **97**, 11703–11728, doi:10.1029/92JB00132.

Conclusions

This dissertation firstly introduced a method to automatically identify seismic phase types. I developed the package: *PhasePApy*, which consists of two sub-packages: the *PhasePicker* and the *Associator*. The *PhasePicker* was designed to automatically detect seismic wave phase arrival times. The *FBpicker*, the *AICDpicker*, and the *KTpicker*, are implemented in the *PhasePicker*. I use the dynamic threshold method in the *PhasePicker* for pick triggering. The threshold level is determined from the noise level of the CF and a control coefficient. The *PhasePicker* has functions to determine the pick's polarity and uncertainty. The *1D* and *3D Associator* identify the phase types by searching for the picks that best fit an earthquake. The *Associator* evaluates RMS residuals to determine which picks can associate to the earthquakes. There are only a few parameters to tune in the *PhasePApy* package. All parameters of the picker and associator are intuitive and easy to set. Depending on users' requirements, the *PhasePicker* and the *Associator* can work separately or jointly. The *PhasePApy* should work well for processing micro-seismic data independent of the surface or down-hole geometry.

I created two 3D velocity models for central Oklahoma, which present similar velocity features at a large scale. The SIMUL2000 model shows the similar large features as the FMTOMO model, but has less model variation, which is due to less data and grids being used in the inversion. I decided to use the SIMUL2000 model, gravity, and magnetic data to correlate with the previously known geological features. Most of the velocity anomalies correlate with known and inferred features, such as igneous intrusions related to the OCM (Osage County Microgranite), the SGG (Spavinaw

Granite Group), and the magnetic high feature in this region. The OCGG Central Oklahoma Granite Group) unit is widely distributed in north central Oklahoma. At mid-crustal levels, a southward extension of the MCR (Mid-Continent Rift) from Kansas into central Oklahoma is suggested from the velocity model. The Vp/Vs ratio generally displays a high-low-high distribution vertically. The low Vp/Vs ratio zone may relate to the water-filled fractures in the host rock.

I relocated the earthquakes from the 01/01/2010 to 03/31/2016 with 3D velocity model. The single event method is firstly used to improve the absolute locations with the 3D velocity model. Then, the DD (double-difference) method was employed to improve the relative locations with the 3D velocity model. The 1D velocity model DD method relocation results were used to determine location shifts. Most earthquake locations shifted less than 1 km, and the average shift amount was about 0.7 km. The relocated clusters provided narrower NE-SW or NW-SE orientations, which are consistent with maximum horizontal stress in this region. Some clusters correlate with fault splays, or align with mapped faults. Some clusters suggest new inferred faults due to no nearby publically known faults correlating with them.