BAYESIAN MONTE CARLO INVERSION OF MULTIPLE DATA SETS TO IMAGE THE CRUST AND UPPERMOST MANTLE BENEATH THE CONTINENTAL UNITED STATES

by

WEISEN SHEN

B.S., University of Science and Technology of China, 2008

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Physics

2014

This thesis entitled:

Baysian Monte Carlo inversion of multiple data sets to image the crust and uppermost mantle structure beneath the continental United States

written by Weisen Shen
has been approved for the Department of Physics

Michael H. Ritzwoller

Craig H. Jones

Peter Molnar

John Wahr

Shijie Zhong

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Shen, Weisen (Ph.D., Geophysics, Department of Physics)

Bayesian Monte Carlo Inversion of Multiple Data Sets to Image the Crust and Uppermost Mantle Beneath the Continental United States

Thesis directed by Professor Michael H. Ritzwoller

The recent deployment of the USArray/Transportable Array (USArray/TA) has stimulated new methods (e.g., ambient noise eikonal tomography; teleseismic Helmholtz tomography) to produce high resolution surface wave dispersion maps. These dispersion maps, combined with other geophysical data sets derived from the array such as receiver functions, present the opportunity to image the crust and uppermost mantle for the continental US at unprecedented resolution. However, new methods are needed to overcome the limitations of traditional methods that may generate unstable models and do not estimate model uncertainties.

In this thesis, I present a new approach that jointly interprets new surface wave observations with other geophysical observables using a Bayesian Monte Carlo framework. In this approach, prior constraints and assumptions are explicitly expressed as prior distributions, and data uncertainties are rigorously interpreted in the resulting models by the Monte Carlo sampling of the posterior distributions. Thus, a 3D model with attendant uncertainties at all depths and for all discontinuities is estimated.

I show that with this approach it is feasible to interpret both surface wave data and other geophysical data observed at most stations from the USArray/TA with simple models. In addition, the vertical resolution of the model is enhanced by improvements to estimates of Moho depth and upper crustal structures using both receiver functions and Rayleigh wave H/V ratio. By applying the new method to multiple data sets, a set of 3-D models is constructed for the crust and uppermost mantle beneath the contiguous US. These models reveal many geological features.

CONTENTS

TABLES

Table

FIGURES

Figure

CHAPTER I

INTRODUCTION

Synopsis of this Chapter

This chapter presents the background, purpose and roadmap of the thesis. In section 1.1, background information is presented. In section 1.2, we present the challenged faced by application of traditional linearized least-squares joint inversion of surface wave dispersion and receiver functions. In section 1.3, to these challenges, we list three guiding principles. In section 1.4, the new approach that jointly interprets surface wave and receiver functions using Bayesian Monte Carlo method with a simple model parameterization is presented. Finally in section 1.5, we present the arrangement of this thesis.

1.1 Background of the Study

Continents, the relatively stable and older parts of the earth, appear as complex assemblages whose detailed structure, composition and evolution have challenged generations of geoscientists. The North American continent, including the contiguous United States, for instance, retains a complicated continental lithosphere that has suffered from a variety of tectonic modifications (e.g., subduction, continental rifting, mantle plumes, etc) (Fig 1.1). To constrain the crustal and uppermost mantle structure beneath this region, scientists have developed a variety of comprehensive geophysical tools including active (e.g., the Consortium for Continental Reflection Profiling (COCORP) project) and passive seismic arrays. In recent years the latter has become the preferred means to image high-resolution earth structures. Traditionally, these seismic arrays were mainly deployed to focus exclusively on one tectonic feature (e.g., the Yellowstone Intermountain Seismic Array (YISA), the Rocky Mountain Front experiment (RMF), the Sierra Nevada EarthScope Experiment, and many others). But to perform a comprehensive study of continental tectonics, extensive seismic arrays at continental scales are needed.

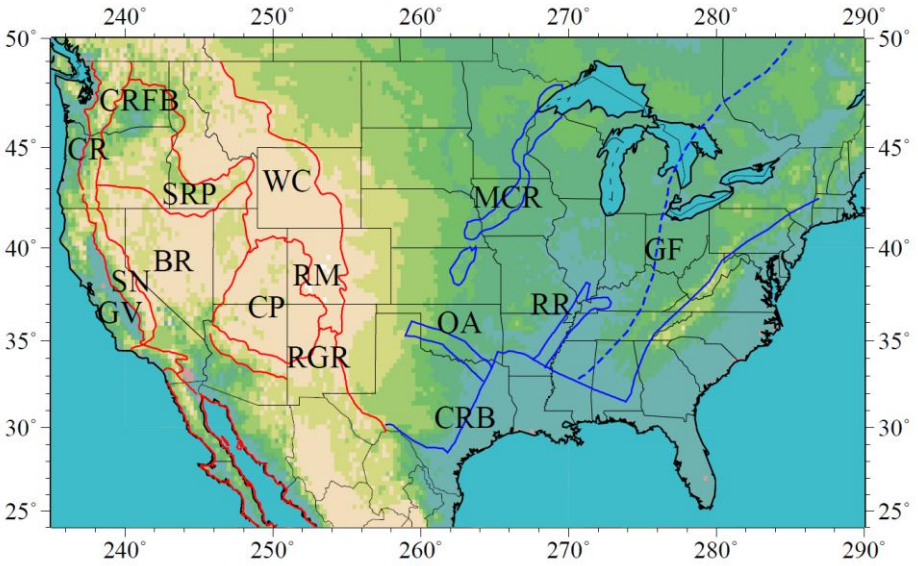

Figure 1.1 Geological/Tectonic features of the continental United States. In the western US, major geological/tectonic provinces are outlined with red lines and identified with abbreviations: the Columbia River Flood Basalt (CRFB), the Cascade Range (CR), the Snake River Plain (SRP), the Wyoming Craton (WC), the Basin and Range (BR), the Sierra Nevada (SN), the Great Valley (GV), the Rocky Mountains (RM), the Colorado Plateau (CP) and the Rio Grande Rift (RGR). In the eastern US, aulacogens, rifts, and other features are shown with blue lines, and named with abbreviations as well: the Midcontinental Rift (MCR), the Reelfoot Rift (RR), the Oklahoma Aulacogen (OA) and the the Continental Rift Boundary (CRB). The Greenville front (GF) is outlined with a dashed blue line. The seismic structures beneath these features will be presented and discussed in Chapters III, IV, and V.

Since the end of 20st century, large seismic observatories spanning the scale (thousands of km) have been proposed and deployed in several regions across around the world (e.g., the Virtual European Broadband Seismograph Network; the large seismic infrastructure in China sometimes referred to as ChinArray or CEArray). Among these, Earthscope, funded by the US National Science Foundation (NSF), stands out as the most comprehensive and has extraordinary data quality control. As a critical part of Earthscope, the USArray/Transportable Array (TA), comprising 400 portable, 3-component, broadband seismic stations with a uniform station spacing of about 70 km, has moved continuously over the past ten years to cover nearly the entire contiguous US from west to east in a rolling fashion (Fig. 1.2). By May 2013, stations at over 1,800 sites have been deployed. From the USArray/TA, high quality seismic data has been professionally collected, quality-controlled and archived by the IRIS DMC, and is freely distributed to scientists around the world. The deployment of the USArray/TA, combined with its professional data management and open data policy, has stimulated many innovations in the

seismology discipline, and ambient noise tomography is the most important one. Data from this array from its inception in 2004 until May 2013 form the basis for this thesis.

First applied to the USArray stations, the use of empirical Green's functions (EGFs) extracted from ambient noise to infer Rayleigh wave (Shapiro et al., 2005) group and phase velocity maps in continental areas has been well established. This technique (usually called ambient noise tomography) is typically performed between periods of 8 and 40 sec and used to infer crustal structure. Using this technique, numerous three-dimensional models of the crust and uppermost mantle have been constructed for the continental US (e.g., Bensen et al., 2009; Liang and Langston, 2008; Moschetti et al., 2010b; Stachnik et al., 2008;Yang et al., 2008b;Yang et al., 2011).

Another innovation based on USArray is the eikonal/Helmhotz tomography. The eikonal tomography method, designed to be applied to extensively and uniformly distributed seismic arrays, is developed to produce phase velocity maps from ambient noise directly using eikonal equation (Lin et al., 2009;). In this method, direct estimates of local phase speed are computed by applying spatial filters (e.g., gradient operator) to the travel time surfaces. By applying the filters on many travel time surfaces centered at different stations, the mean speed and 1 σ uncertainty estimate at a given location are thus quantified from a group of local phase speed measurements. The extension of eikonal tomography that incorporates the finite frequency correction term of the Helmholtz equation is called Helmholtz tomography (Lin and Ritzwoller, 2011). This tomography method is developed and applied to the surface waves generated from teleseismic earthquakes at longer periods (25-100 sec). Similar to eikonal tomography, meaningful uncertainties of the resulting phase velocity maps can also be estimated. By performing the ambient noise eikonal tomography and teleseismic Helmholtz tomography, broadband Rayleigh wave dispersion measurements (8-100 sec) with uncertainties can be obtained across a large array.

Overall, the deployment of the USArray/TA, the newly developed ambient noise tomography technique, together with recent innovations in array-based surface wave tomography, provide a comprehensive basis for imaging the earth's interior. Exploiting their capabilities to better

constrain the crustal and uppermost mantle structures beneath the US continent serves as the fundamental motivation of the studies presented in this thesis.

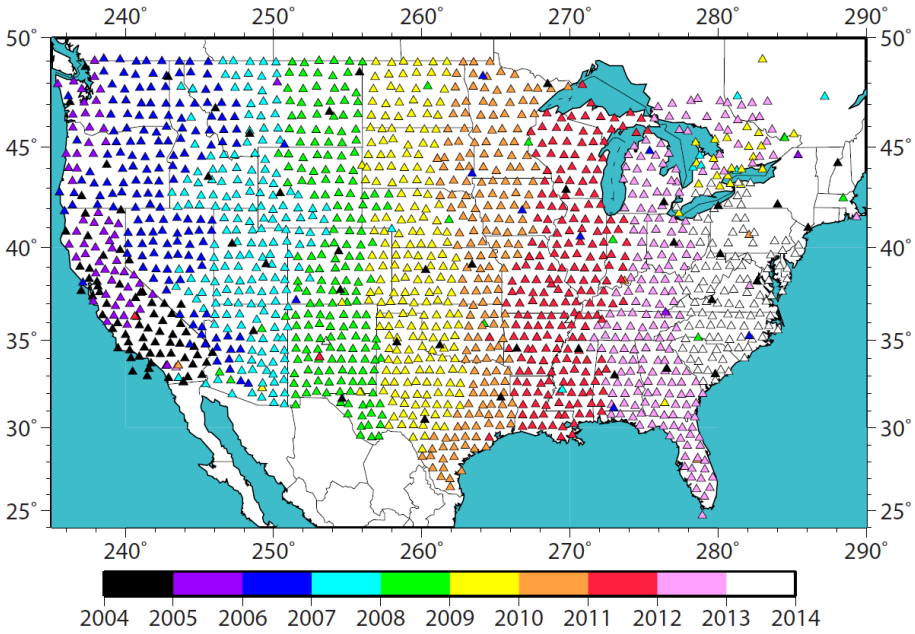

Figure 1.2 Earthscope/USArray seismic stations deployed before 2013 May. 1723 stations are presented as triangles with deployment dates marked by the colors. Transportable array (TA) stations roll across continental US from west coast since 2004 and cover 90% of the continental US in middle 2013. About 40 US backbone stations are scattered across the map with the deployment dates earlier than 2005. The average station spacing is about 70 km (40 miles).

1.2 Limitations of Traditional Joint Inversions of Surface Wave and Receiver Function Data

Large seismic arrays admit a wide variety of seismic techniques to address a diversity of geophysical problems, and seismic imaging is one of the most powerful choices. This is largely because 3-D information about the elastic properties of the deep earth, whose mathematical description is called a "model", directly constrains complex structures in the subsurface, providing information about density, composition, and temperature of earth's interior.

Seismic imaging techniques can be broadly categorized into two groups related to the types of seismic data used (M.G. Bostock, 1999). 1) Travel times of mechanical waves that propagates through the earth are analyzed to constrain earth's elastic properties. 2) Scattered waves generated from earth's velocity discontinuities are identified and analyzed to locate and characterize such discontinuities. Body-wave tomography and surface wave tomography are categorized into the first group. Receiver functions and deep refraction and reflection profiles belong to the second group. Combining both categories of data can be more powerful than applying either alone. For example, the joint inversion of surface wave dispersion and receiver function data has been well studied because of the complementarity of these data sets. Surface wave dispersion provides information about the average velocity within a given depth range while receiver functions provide information about jumps in velocity across interfaces.

The combination of surface wave dispersion information and receiver functions to infer earth's velocity structure was first proposed nearly two decades ago (e.g., Ozalaybey et al., 1997). Julia et al. (2000) presented a damped least-squares algorithm that jointly inverts the two data sets for local 1-D Vs structure, and this algorithm has been applied by many researchers (e.g., Julia et al., 2003; Endrun et al., 2004, Horspool et al., 2006, Gok et al., 2007, Tkalcic et al., 2006; Pasyanos et al., 2007; Yoo et al., 2007; Tokam et al., 2010, Salah et al., 2011, Bailey et al., 2012). This algorithm joinly inverts both data types using a thin-layer parameterization of the seismic velocities. A typical result using this algorithm is shown in Figure 1.3.

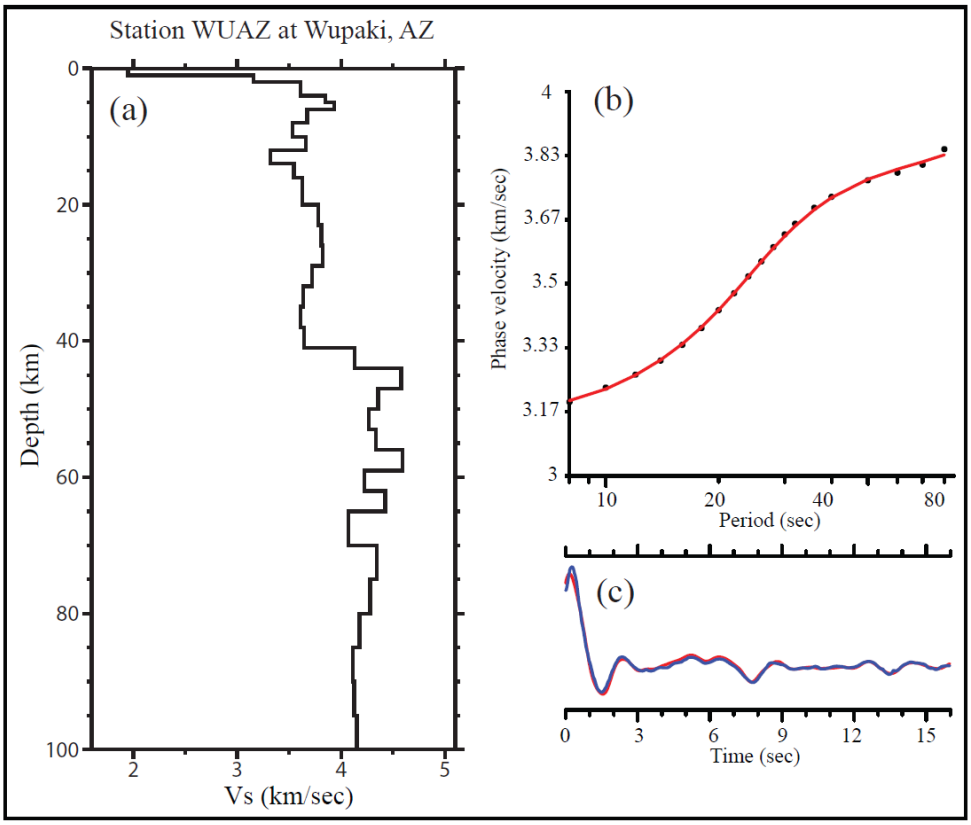

Figure 1.3. (a) 1-D Vs model from a traditional least-squares joint inversion of surface wave dispersion and a receiver function for station WUAZ located at Wupaki, AZ in the Colorado Plateau. (b) The observed Rayleigh wave phase velocity (black dots) and the predicted phase velocity curve (red line) from the model shown in (a) are presented. (c) The observed average receiver function waveform is shown with the blue line and the predicted receiver function from the model in (a) is shown as the red waveform.

Figure 1.3 presents an example joint inversion at station WUAZ in the Colorado Plateau in which both the surface wave dispersion and the receiver function waveform are fit perfectly (Fig. 1.3b-c). The resulting Vs model appears reasonable at first glance. However, the model possesses a number of questionable characteristics. At a depth of ~ 5km, Vs increases to about 4 km/sec in the upper crust, below which a low velocity layer (<3.3 km/sec) appears at ~ 15 km depth in the middle crust. Such velocity inversions are common in these types of inversions. In addition, to fit the receiver function, numerous discontinuities in Vs are generated in the uppermost mantle (45-80 km). Moreover, a "Moho" discontinuity appears at ~ 42 km depth, even though no clear Moho P-to-S (Ps) converted phase appears in the receiver function (Fig 1.3c). These phenomena (e.g., velocity inversions, many small scale discontinuities, Moho discontinuity without a Ps phase) are caused by the need to fit every small wiggle in the observed

receiver function. The introduction of alternating layers in the shallow crust generates reverberations in the receiver function that require discontinuities at greater depths (e.g, the "Moho" at ~ 40 km) in order to fit the relatively flat receiver function. The near perfect fit to the receiver function, therefore, results from the destructive interference of multiple phases generated at many discontinuities across a large depth range (in this case, from the surface to ~ 70 km depth), rather than from a simple, single structure (e.g., a gradient Moho structure, as discussed later in section 1.4).

This could admittedly be the way the earth is layered at this particular point, but such a fine interference condition would not be expected to be stable and to exist uniformly over large areas. Nearby receiver functions, therefore, would be expected to look different from one another because small perturbations to these variations in Vs will destroy the destructive interference and strongly change the receiver functions. In fact, similar receiver functions are observed to extend over large areas (e.g., the N.E. Colorado Plateau). Thus, the Vs model that appears from the joint inversion presented in Figure 1.3 is unstable, physically unrealistic, and uninterpretable.

Some seismologists would attribute the problems with the resulting model to vertical smoothing constraints being too weak during the inversion. In Figure 1.4, using data from stations R11A in the Basin and Range province, we present another example of a joint inversion using the traditional damped least-squares method, but with a stronger smoothing constraint. This example highlights another danger of the traditional least-squares inversion method.

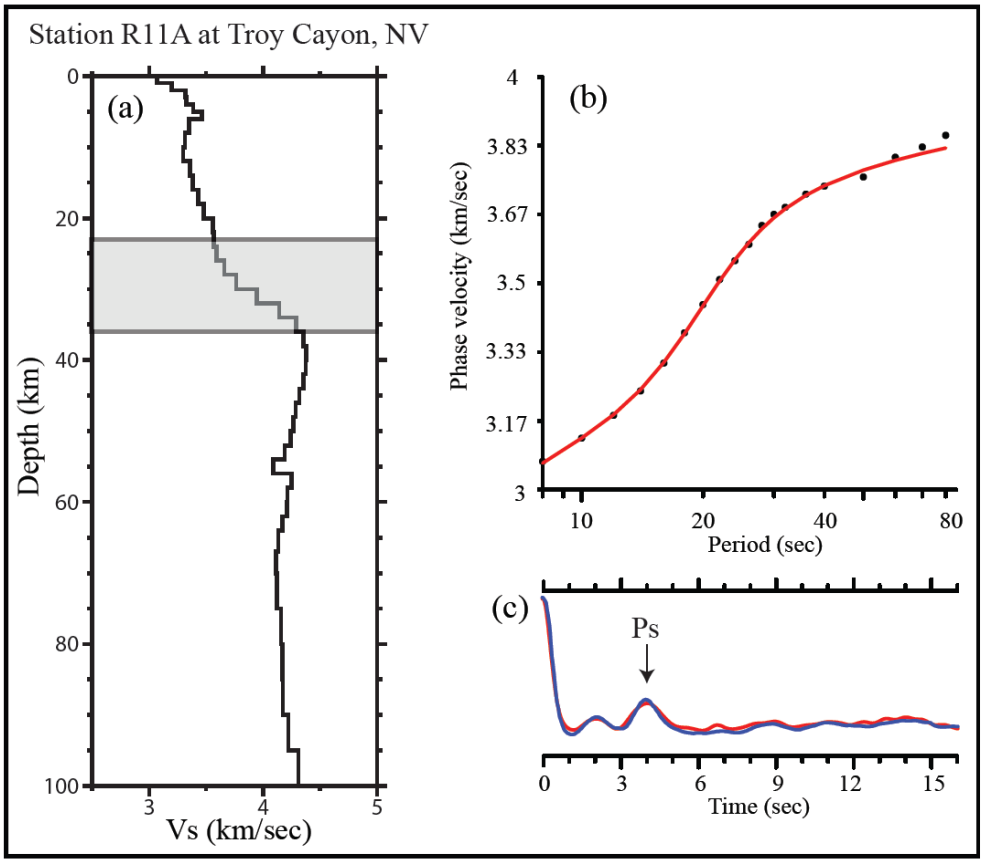

Figure 1.4 Another example of a joint inversion using the linearized least-squares method with stronger vertical smoothing, performed for station R11A in the Basin and Range. Similar to Figure 1.3, (a) presents the resulting Vs model and (b-c) present the observed and predicted data. Please note the Ps phase labeled in (c) and the gradient Moho structure shadowed in (a).

The model resulting from the joint inversion with a stronger smoothing constraint in Figure 1.4 again produces a dispersion curve and receiver function waveform that fit the observed data almost perfectly (Fig 1.4b-c). Compared with the model shown in Figure 1.3a, fewer small-scale Vs variations are observed with depth due to the stronger smoothing constraint. However, another problem is present: Although a strong Moho Ps phase appears in the observed receiver function and is fit by the model, the resulting model does not reveal the expected sharp and clear Moho. Instead, a gradient structure is seen between 25 and 36 km depth range (the shaded area marked in Fig. 1.4a). Again, the fit to the receiver function waveform is the outcome of interference between signals from many structures distributed over a range of depths. In this case, constructive interference occurs amongst the phases converted at the set of discontinuities in the 25-36 km depth range, not a single discontinuity.

Other studies using receiver function data alone (e.g., CCP stacking) present a different picture of a clear, singular Moho discontinuity beneath this region of the Basin and Range province (Gilbert et al., 2004, Gilbert, 2012). This discrepancy in interpretation reveals that the model presented in Figure 1.4a is biased by the ad-hoc selection of the smoothing constraint and is, therefore, unreliable.

The two examples presented here reveal a critical problem for the traditional least-squares inversion of surface wave dispersion and receiver functions: the inability to clearly and reliably recover information about the crust. This is caused by 1) the thin-layer model parameterization, 2) the ad-hoc smoothing constraint applied to the inversion, and 3) the attempt to fit every wiggle in the receiver functions regardless of data uncertainties. To solve this problem one has to carefully determine the appropriate smoothness constraint on a station-by-station basis, presumably by introducing local knowledge, which prohibits the automated application of the method systematically to a large number of stations. The application of this method to large seismic arrays such as USArray/TA has proven to be an insurmountable challenge.

Moreover, another critical disadvantage for the traditional least-squares method is that, although both models are unreliable, the inversion does not provide any information to determine the quality of the resulting models. Meaningful uncertainties are not provided with the resulting models from the linearized joint inversion scheme, and this further limits the usage/interpretation of the models produced with this method.

1.3 Basic Guidelines of this Study

To break through the limitations of the traditional linearized least-square inversion scheme summarized in the previous section, we look for a solution under a different philosophy. In this section, we describe the principles that guide our approach, which are summarized in three general Guidelines.

Guideline 1. The resulting models should be useful to scientists other than the model builders.

"Essentially, all models are wrong, but some are useful" (Box and Draper, 1987). We value the usefulness of the resulting models as the highest priority. We will show in this thesis how the results of this study, a set of 3-D models for the crust and uppermost mantle beneath the continental US, are generated under this guideline are designed to be used by other researchers (e.g, Levandowski et al., 2014). In this section we only discuss its implication and how it guides us to develop a new method.

To be useful to other scientists, meaningful uncertainties associated with the resulting earth model must be presented which improves usefulness in three ways. First, they help seismologists assess the resulting model. When two models (e.g., one from the joint inversion of surface wave dispersion and receiver functions, and another from surface wave inversion alone) are compared, uncertainties are important quality indicators. Second, predictions produced from the model (e.g., surface topography, body wave travel times, internal temperatures, surface or internal heat flow) can also include uncertainties, so scientists from another disciplines can better determine the solidity of conclusions that may be drawn from the predicted quantities. Third, the uncertainties themselves in some model parameters (e.g., crustal thickness) may have geological implications as well.

Finally, to ensure the model to be useful, we have to avoid models with physically unrealistic features from being generated during the inversion. This motivates the introduction of Guidelines 2 and 3.

Guideline 2. The assumptions that underlie the inversion should be well-understood, easily implemented, and explicitly presented.

Guideline 2 is often neglected in most geophysical inversions. All inversions contain preconditions and initial assumptions. The model parameterization is one of the preconditions but its effect is often poorly understood. For example, in traditional Vs inversion thin layers are often used as the model parameterization. In this case, model discontinuities occur only between these fixed depth interval layers. For instance, if a 5 km thickness layered model is used to perform the 1-D inversion, this parameterization would impose the very strong constraint that the Moho discontinuity and other crustal discontinuities can only exist at depths of 5 km, 10 km, 15 km, 20 km, etc. Other preconditions in traditional least-squares inversions include an ad-hoc damping/smoothing constraint that is usually ad-hoc and a starting model. Parameterization, ad-hoc smoothing constaint, and the starting point are the triad of preconditions that affect nearly every seismic inversion, and their impacts are typically poorly understood. In addition to these three, there are other constraints that are often applied in surfae wave inversions, such as the monotonic increasing of Vs in the crust, a positive Vs gradient in the uppermost mantle, etc.

These preconditions and assumptions listed above strongly affect the resulting model, but the ways in which they do so are typically not quantified or presented, and in some cases are hard to implement (e.g., the monotonic increasing of Vs cannot be easily applied in a linearized inversion). The new methods we develop will address these problems.

Guideline 3. The data should be fit with a simple model, in which the attempt is made to include all and only the structures that are needed to fit the data.

As discussed in last section, traditional thinly-layered models cause produce physically questionable and unreliable results. In the examples above, spurious model variations are introduced when the smoothness constraint is weakened, and a clear Moho is replaced by a gradient structure when the constraint is overly strong. Also, a thin-layered model requires the pre-selection of the layer thickness, which limits the location of the discontinuities required by the data. Third, the thin-layered model parameterization usually calls for numerous layers to represent the crust and uppermost mantle, which prohibits the usage of a non-linear parameter sampling algorithm. Guideline 3 forces us to abandon this thin-layer model parameterization to be replaced by another simpler parameterization.

1.4 Purpose of this Study

In this section, we present the purpose of this thesis, and discuss in detail how we choose a proper approach to achieve it.

To fully exploit the USArray and the recent innovations in surface wave tomography in order to better constrain the crustal and uppermost mantle structure of the continental US, a new joint inversion of surface wave dispersion and other seismic data including receiver functions is developed and applied exclusively for large array. As discussed in previous sections, the disadvantages of the traditional approach limit its application to the large array. To move over these limitations, we design this approach by following the three guiding principles listed in section 1.3: 1) the usefulness of the result; 2) the simplicity of the model and 3) the transparency of the assumptions.

To follow the Guideline 1, meaningful uncertainties of the input data must be estimated. This challenge has been partially solved recently for surface wave dispersion maps by the eikonal and Helmholtz tomography methods, which produce reliable error estimates for both ambient noise and earthquake phase velocity maps. In Chapter II, I will discuss a method to construct receiver functions, referred to as harmonic stripping, based on an azimuthal harmonic analysis (e.g., Girardin and Farra,1998; Bianchi et al., 2010) that determines azimuthally independent receiver functions and their uncertainties.

In addtion, these data uncertainties should be propagated into the resulting 3D models. This has traditionally challenged linearized inversion. The non-linear inversion scheme based on Bayesian theory provides an approach to solve this problem. Under Bayesian theory, the inversion can be stated as a problem in conditional probability: given the data observed and an explicit quantitative statement of the preconditions or assumptions, what is the earth likely to be? The answer to this question is presented as a probability, and this probability (usually called the posterior distribution) directly provides the uncertainties of the result (e.g, if the resulting Moho depth is provided as a Gaussian distribution with a given mean and variance, then the 1σ uncertainty of Moho depth can be defined as the standard deviation of the distribution). In practice, Bayesian inversion is often performed using distribution sampling techniques (e.g., Monte Carlo sampling using the Metropolis rule, Neighborhood sampling, and so forth). In this

thesis, we will show how Bayesian inversion using a Monte Carlo sampling scheme (we call it Bayesian Monte Carlo inversion below) is applied to produce model uncertainties both with surface wave data alone and jointly in conjunction with other kinds of geophysical data. In our view, this information greatly enhances the model's usefulness.

By applying a Bayesian Monte Carlo inversion, Guideline 2 is followed as well. Assumptions that underlie the inversion are expressed explicitly in terms of the prior distribution and models that satisfy these assumptions are obtained by sampling of the prior distribution. Furthermore, as the Bayesian Monte Carlo method does not involve an "inversion" but only performs the forward computation on the sampled models, some unachievable assumptions under a linear inversion are now possible.

To follow Guideline 3, we define an earth model that is totally different from a thinly-layered parameterization with ad-hoc constraints. In this study, both the crystalline crust and the uppermost mantle are parameterized by smooth cubic B-splines, and between them a Moho discontinuity is explicitly defined as a velocity jump. The location of this contrast and velocity across this contrast are clearly defined in the parameterization. This approach avoids the trying to discover the Moho in a model based on the thinly-layer model parameterization, which might not be easy in some instances (e.g., the model shown in Fig. 1.4a, which has an ill-defined Moho).

We also apply a monotonic increasing Vs constraint in the crust. This assumption, combined with the smooth model parameterization, ensures that rapidly varying models within the crust and uppermost mantle (such as swings of thin velocity anomalies) are prohibited, In consequence, the resulting models are physically more reaslist, more interpretable, and simply more scientifically useful. One could argue that this would prohibit possible low velocity zones or discontinuities within the crust, which might exist in some areas. This is true, but we see this as a hypothesis test: we start with simple models with strong prior constraints (e.g., monotonically increasing velocity in the crust and a radially smooth crust), and if the data are not fit we loosen the prior constraints (e.g., we can introduce extra layers in the crust or allow for a negative velocity gradient within the crust) and re-perform the Monte Carlo inversion. Essentially, such features will only be introduced when needed by the data. As we will show in

this thesis, the simple model hypothesis will be supported by the data observed at most stations in the continental US, but there are interesting exceptions.

1.5 New Methodology

To highlight the changes motivated by these guidelines, we present examples of the joint inversion results for the data sets shown from the two stations in Figures 1.3 and 1.4.

The result for station WUAZ (Fig. 1.3) in the Colorado Plateaue from the Bayesian Monte Carlo inversion is presented in Figure 1.5. In contrast with the linearized inversion result presented in Figure 1.3a, an ensemble of models rather than a single model is presented as the result of the inversion (gray corridor shown in Fig 1.5c). Each of the resulting models within the ensemble contains a smooth crust and mantle with a clearly identified Moho separating them and fit the Rayleigh wave phase speed data and receiver function (Fig 1.5a-b) within the estimated uncertainties.

For station WUAZ, as the observed receiver function does not display a clear Ps phase in the context of the estimated uncertainties (parallel lines shown in Fig 1.5b), each model in the ensemble possesses a weak Vs contrast or velocity jump across the Moho (Fig. 1.5c). As a result, the model ensemble represents a gradient Moho structure within the 40-52 km depth range (highlighted as the gray area in Fig 1.5c). This model ensemble can be used to quantify the uncertainties of the resulting average model. For example, the uncertainty of the Moho depth can be defined as the standard deviation of the Moho depths for all models in the ensemble, which is large for this station.

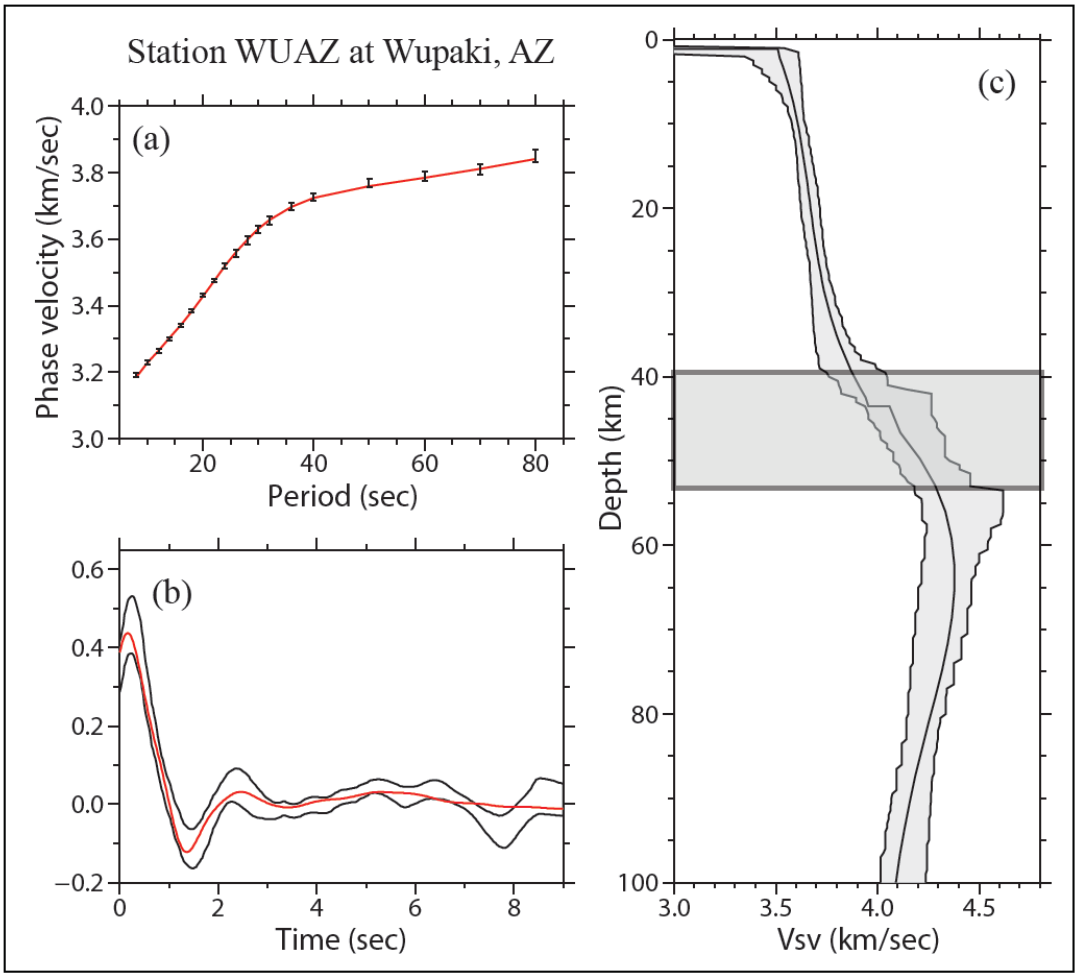

Figure 1.5. A typical result from the joint inversion of Rayleigh wave phase velocities and receiver function under the guidelines listed in section 1.3. (a) The observed dispersion curve with estimated uncertainties is presented as error bars. The red line is the predicted dispersion curve from the model shown in (c). (b) The observed receiver function with estimated uncertainties is shown as two black parallel lines. The red waveform is the fit to the receiver function from the model in (c). (c) The final model ensemble is shown as a gray corridor. The black line in this corridor represents one example from the ensemble of Vs models.

Another result of the Monte Carlo joint inversion for station R11A is shown in Figure 1.6. Compared with Figure 1.4, instead of fitting the Ps converted phase by a detailed constructive interference of many converted phases from discontinuities at a variety of depth ranges, we fit the observed receiver function by placing a sharp, clear Moho discontinuity at a depth of ~ 32 km. The average Moho depth (presented as the dashed line in Fig. 1.6c), that can be inferred from the distribution of Moho depths extracted directly from the accepted model ensemble, is

estimated to be ~ 32 km with a 1 σ uncertainty < 2 km. This is consistent with previous receiver function studies (Gilbert et al., 2004, Gilbert 2012).

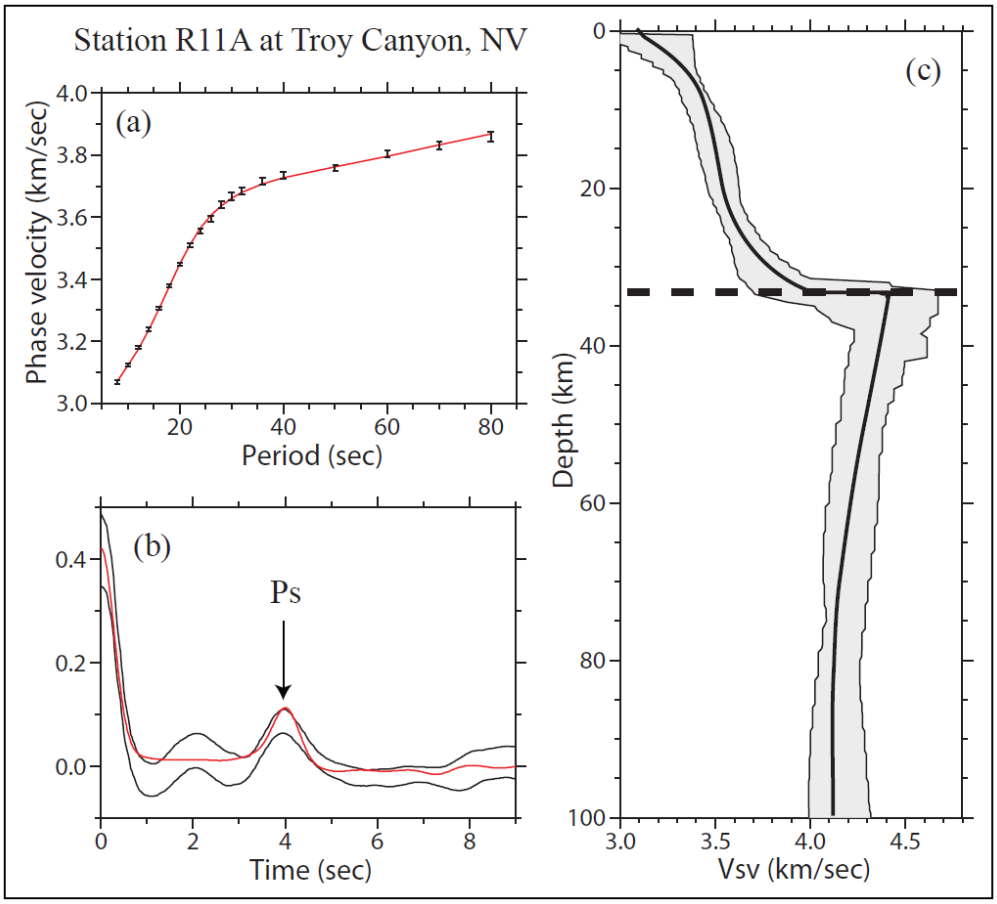

Figure 1.6 The same as Fig. 1.5, but for station R11A in the Basin and Range province.

The analysis of data uncertainty and its propagation into the resulting models, combined with a simple model parameterization, will make the resulting models more interpretable and we believe more useful to other scientists compared with the models presented in Figures 1.3 and 1.4.

1.6 Organization of the Thesis

The major parts of the thesis are divided into four chapters. First, Chapter II presents the details of the Bayesian Monte Carlo inversion and the test of this method using USArray/TA stations in the intermountain west area. Second, Chapter III presents a 3-D Vs model for the crust and uppermost mantle beneath the western US together with estimated uncertainties. Third, as an example of a more localized interpretation of the resulting model ensembles, Chapter IV presents

a 3-D model of the central US encompassing the Midcontinent Rift. In Chapter V, we extend the inversion algorithm to incorporate a new observable, Rayleigh wave H/V ratio, and apply the extended method to data from more than 1,700 stations from the USArray/TA. The thesis is then summarized in Chapter VI.

The study area of each chapter is identified in Figure 1.7, with the USArray/TA stations marked as red triangles and geological boundaries marked as red/blue lines in the background. For Chapter II, 185 stations in the intermountain west (blue box shown in Fig. 1.7) are used. For Chapter III, seismic data from more than 800 stations in the western US (black box shown in Fig. 1.7) are applied to image the crust and uppermost mantle structure. In chapter IV, ~ 122 stations covering the western arm of the Midcontinent Rift are used to produce a 3-D model for the area (red box shown in Fig. 1.7). In Chapter V, 1,723 stations acorss the US are used to construct the latest Rayleigh wave phase speed and H/V maps as well as receiver functions, and these data are analyzed simultaneously to construct a final crustal and uppermost mantle 3-D model covering ~80% of the continental US.

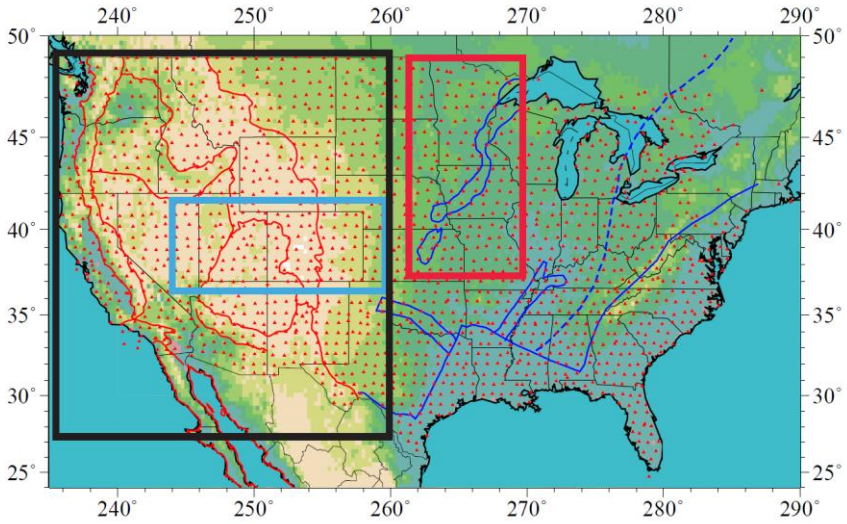

Figure 1.7 The study area for each chapter. The blue box outlines the testbed for the joint inversion of surface wave dispersion and receiver functions, which is the study area of Chapter II. The black box outlines the western US where a 3-D Vs model is presented in Chapter III. The red box shows the area encompassing the western arm of the Midcontinent Rift where chapter IV concentrates. All the stations are applied in Chapter V to perform the joint inversion of Rayleigh wave phase speed and H/V ratio together with receiver functions. Tectonic boundaries from Figure 1.1 and stations from Figure 1.2 are also shown.

CHAPTER II

BAYESIAN MONTE CARLO INVERSION OF SURFACE WAVE DISPERSION AND RECEIVER FUNCTIONS

Synopsis of this chapter

Following the three guiding principles presented in Chapter I, a new approach is laid out in this chapter that jointly interprets surface wave dispersion data and receiver functions. The techniques used by this approach, including Bayesian Monte Carlo inversion of multiple data sets and receiver function harmonic stripping, will be consistently used in later chapters. I start this chapter with an overview of the Bayesian Monte Carlo method in section 2.1, and then introduce the data processing of Rayleigh wave dispersion maps and receiver functions in section 2.2. Also in section 2.2, the harmonic stripping method, which is the basis for quality control and assessment and to produce azimuthally independent receiver functions, is described in detail. From section 2.3 to 2.6, the Monte Carlo inversion algorithm is laid out in 4 steps. First, the model parameterization, prior constraints, and Monte Carlo sampling of prior distributions are presented in section 2.3. Second, the Monte Carlo inversion of surface wave data alone is introduced in section 2.4. Third, the joint inversion of both surface wave and receiver function data sets is laid out and the effect of adding the receiver functions to the Monte Carlo inversion is discussed in section 2.5. Fourth, the capability of the Monte Carlo joint inversion to construct a 3-D model across large seismic arrays is applied as a test to 180 USArray/TA stations in the Intermountain West. Other aspects of the joint inversion (e.g., the constraint to the crustal Vp/Vs ratio; the use of denser arrays to produce better receiver functions) are discussed in the last section (2.7).

2.1. Overview of the Bayesian Monte Carlo Method

Monte Carlo and related parameter search algorithms (e.g., Sambridge, 1999a) are designed to map data misfit across a broad range of model space. Within a Bayesian framework, these models are interpreted by computing the a posteriori (or posterior) probability distribution, which is the probability distribution of the model parameters given the observed data. Bayes' theorem

allows the posterior distribution σ(m) for a model m to be computed from the prior information on model space (given by the prior probability density ρ(m) for model m) and the observed data as represented by the likelihood functional L(m) (the probability of observing the measured data given a particular model):

$$\sigma(m) \propto \rho(m)L(m) \tag{2.1}$$

Geophysical applications of Bayesian inference have been presented by Tarantola and Valette (1982), Mosegaard and Tarantola (1995), and Sambridge (1999b). The key steps are the expression of prior information as a probability density function and the computation of the likelihood functional, which is determined from data misfit. From a set of models distributed like the posterior distribution, one can then determine certain properties of the ensemble (e.g., best fitting model, mean model, median model), the covariance between model parameters, or the marginal distribution of particular model parameters or quantities derived from the model parameters.

The procedure we have developed to jointly invert surface wave dispersion information and receiver functions is similar in motivation and even is some of the details to the methods described by Bodin et al. (2011). We present a general overview of our methodology here as a guidepost to the principal results in the chapter. We will also point out some differences with the method of Bodin et al. (2011). The data used in the inversion are described in detail in section 2.3. The model parameterization and the construction of the prior distribution are presented in section 2.4. The prior information includes ranges in which individual model parameters are allowed to vary as well as rules that govern the relative values of different model parameters. The likelihood functional is simply related to the misfit function S(m) as follows:

$$L(m) = \exp\left(-\frac{1}{2}S(m)\right) \tag{2.2}$$

where

$$S(m) = \left(g(m) - D^{obs}\right)^T C_e^{-1} \left(g(m) - D^{obs}\right) \tag{2.3}$$

and D^{obs} is a vector of measured data, g(m) is the vector of data predicted from model m, C_e is the data covariance matrix, and T represents transpose of a vector. For surface wave dispersion here the vector D^{obs} consists of Rayleigh wave phase velocities observed on a discrete set period

grid and for receiver functions D^{obs} is the azimuthally independent receiver function over a fixed discrete time grid.

The Markov Chain Monte Carlo sampling of model space to generate the posterior distribution evolves as follows. A model m_j is selected from the prior distribution. A second model m_i is drawn and the likelihoods $L(m_i)$ and $L(m_j)$ are computed. The Metropolis law defines the probability of acceptance for model m_i:

$$P_{accept} = \begin{cases} 1 & \text{if } L(m_i) \geq L(m_j) \\ L(m_i)/L(m_j) & \text{if } L(m_i) < L(m_j) \end{cases} \quad (2.4)$$

That is, model m_i is accepted if its likelihood is greater than that of model m_j (i.e., its misfit is lower). Even if its likelihood is smaller than that of model m_j it may still be accepted, however. For example, if $P_{accept} = 0.4$, we use a uniformly distributed random deviate generated between 0 and 1. If the value is less than 0.4 we accept the model. If the new model m_i is accepted, we define a new perturbation based on this model and search on. If the new model is not accepted, we make a record of it and its associated misfit information and start a new perturbation based on the previously accepted model m_j to choose the next step. The evolution of the algorithm in terms of improving misfit for surface wave data alone is discussed in section 2.4.

An additional model acceptance criterion then is applied to define the posterior distribution. In this we diverge from strict Bayesian methodology. If the data covariance matrix were known accurately, the choice of the acceptance criterion would be straightforward. Misfits to surface wave and receiver function data would be commensurable and the acceptance criterion would involve only choosing a probability threshold in the posterior distribution. As described later, because of difficulty in estimating the off-diagonal elements of the data covariance matrices, we have assumed that the surface wave dispersion data and the receiver functions have error processes that are independently but not identically distributed so that each covariance matrix is diagonal with elements σ_i^2 and s_j^2, respectively, for phase velocity at period i and receiver function at time j. This assumption has the practical effect that the misfits between the two data types are incommensurable. We are, therefore, forced to introduce a misfit scaling parameter κ to normalize the misfits between the two data types, as described in section 2.6.1. The resulting joint misfit function, therefore, has the following form:

$$S_{\text{joint}}(m) = S_{SW} + \frac{1}{\kappa} S_{RF} = \sum_{i=1}^{N} \frac{\left(g_i(m) - D_i^{obs}\right)^2}{\sigma_i^2} + \frac{1}{\kappa} \sum_{j=1}^{M} \frac{\left(R_j(m) - A_0(t_j)\right)^2}{s_j^2} \quad (2.5)$$

where $g_i(m)$ is the phase velocity predicted for model m at period i on a discrete grid of N periods, D^{obs} is the observed phase velocity, the function $A_0(t_j)$ is the observed receiver function at time t_j on a discrete grid of M times and $R_j(m)$ is the predicted receiver function for model m at time t_j. Somewhat different model acceptance criteria are applied in the inversion with surface wave data alone in section 2.5 and in the joint inversion in section 2.6. Example posterior marginal distributions are discussed in section 2.4.

The methods that we present here differ from those of Bodin et al. (2011) in two principal ways. First, we choose smooth velocity profiles between specified boundaries at variable depths and fix the parameterization across the region of application. The nominal "transdimensionality" of the method of Bodin et al. refers to the variable parameterization of their method. We do not use an adaptive parameterization because, for the most part, we find that the parameterization we choose is sufficiently flexible to fit the data. Second, as discussed above, we do not estimate the off-diagonal elements of the covariance matrices. Inspired by the hierarchical Bayesian formalism of Bodin et al., we attempted to estimate full covariance matrices for receiver functions, but found first that the estimated covariance matrices were singular and did not yield to matrix regularization methods designed to approximate the inverse matrices. Second, the covariance matrices that we derived were appropriate for a raw receiver function not the estimated azimuthally independent receiver function $A_0(t)$ that we use in the inversion. This is a subtle point that requires further analysis. For both reasons, we moved forward with diagonal covariance matrices for both dispersion data and receiver functions. Although these differences with the more general formulation of Bodin et al. (2011), which attempted to estimate the full covariance matrices for receiver functions, were motivated in part by our inability to estimate satisfactory covariance matrices, they do accelerate the inversion, which facilitates the application of the method to a large array like the USArray. In addition, as we show, the method works well to improve 3D models relative to those produced from surface wave dispersion data alone.

2.2 Data Processing

2.2.1 Stations used in this study

As a testbed for the method, 185 stations from the TA in the Intermountain West (Figure 2.1) are chosen. After applying the joint inversion (and for comparison the inversion based on surface wave dispersion alone) at each station, the final 3D model of the crust and uppermost mantle is produced by interpolating between stations using simple-kriging at each model depth. Prior and posterior marginal distributions are presented to visualize changes in uncertainties for each data set. The study region is geologically diverse and structurally complex and provides an excellent proving ground for the methodology.

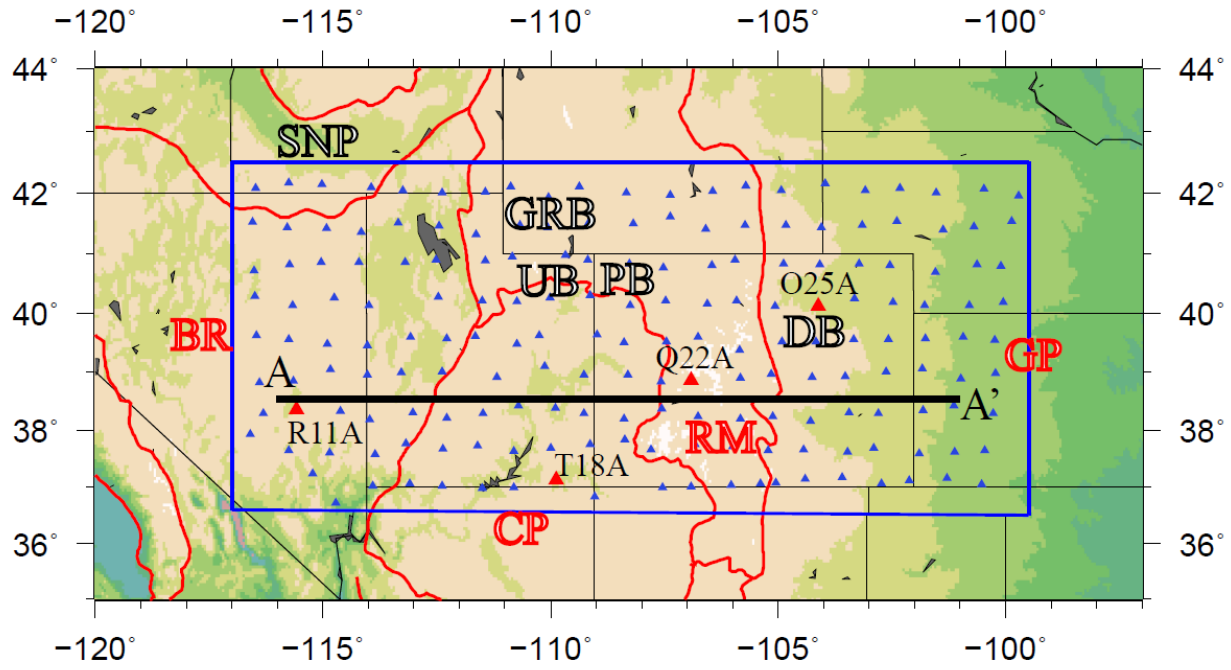

Figure 2.1 Stations of the Earthscope USArrayTransportable Array(TA) used in this study are shown with blue triangles. The main geological provinces are outlined with red contours and titled with red abbreviations (the Basin and Range (BR) province, the Colorado Plateau (CP), the Rocky Mountains (RM), and the Great Plains (GP)). Other regional geological features and basins are also identified with abbreviations (Snake River Plain (SNP), Green River Basin in Southern Wyoming (GRB), Uinta Basin in Northwestern Utah (UB), Piceance Basin in northwesternh Colorado (PB) and Denver Basin in Colorado (DB). Four stations that are used to demonstrate the methods are identified with red triangles. The black line indicates transect A-A' ranging from the BR to the GP.

2.2.2 Rayleigh wave data processing

2.2.2.1 Ambient noise tomography (ANT)

The method of ambient noise tomography is now well established, including cross-correlation of long pre-processed time series to generate empirical Green's functions, measuring both Rayleigh and Love wave phase and group velocity curves, and producing dispersion maps at particular periods (Shapiro et al., 2005, Yang et al., 2007, Moschetti et al., 2007, Yao et al., 2006). The data processing procedures that we adopt follow those of Bensen et al. (2007) and Lin et al. (2008), which we only briefly summarize here. First, raw vertical component seismograms recorded from 2005 to 2010 are downloaded from the IRIS DMC for the USArray TA stations in the western and central US and are cut into 1-day time series. The time series length of each station is usually about two years, but time series lengths are variable and we do not apply a minimum time series length restriction. Second, earthquake signals and other types of interference are removed by time domain normalization, in which the reciprocal of the mean of the absolute value of the waveform in a moving 80 sec time window is used to weight the data point at the center of the window. The weights are determined in the "earthquake band" between periods of 15 and 50 sec period, but then are applied to the unfiltered data. Third, the data are band-pass filtered between 5 and 100 sec period and cross-correlation is performed between all station pairs. Then, Rayleigh wave group and phase velocity dispersion measurements are obtained from the symmetric component of each inter-station cross-correlation by performing automatic frequency-time analysis (FTAN) (Bensen et al., 2007) . In this study, only Rayleigh wave phase velocity measurements are used. The automated FTAN dispersion measurements are winnowed by applying two criteria to select reliable measurements for surface wave tomography: (1) The inter-station distance must be greater than three wavelengths at each period to ensure the far-field approximation and sufficient separation from precursory noise and (2) signal-to-noise ratio (SNR) must be greater than 10 at each period for the measurement at that period to be accepted. Once all measurements are obtained, eikonal tomography (Lin et al., 2009) is then applied to produce phase velocity maps from 8 to 40 sec period. Eikonal tomography takes account of ray bending (off-great circle propagation) but does not model finite frequency effects (Lin and Ritzwoller, 2011) . One of the signature features of eikonal tomography is that it produces meaningful uncertainty estimates at each geographical location. This is done by accumulating phase velocity estimates at a given location from many central

stations, fitting and then removing a truncated Fourier Series over azimuth to estimate azimuthal anisotropy, and then measuring the standard deviation of the mean of the residual. Figure 2.2a-c presents phase velocity maps from ambient noise tomography at periods of 8, 20 and 36 sec. By 40 sec period, the SNR of ambient noise has decreased enough that high resolution phase velocity maps are no longer generated. Earthquake data are introduced to produce the longer period dispersion maps.

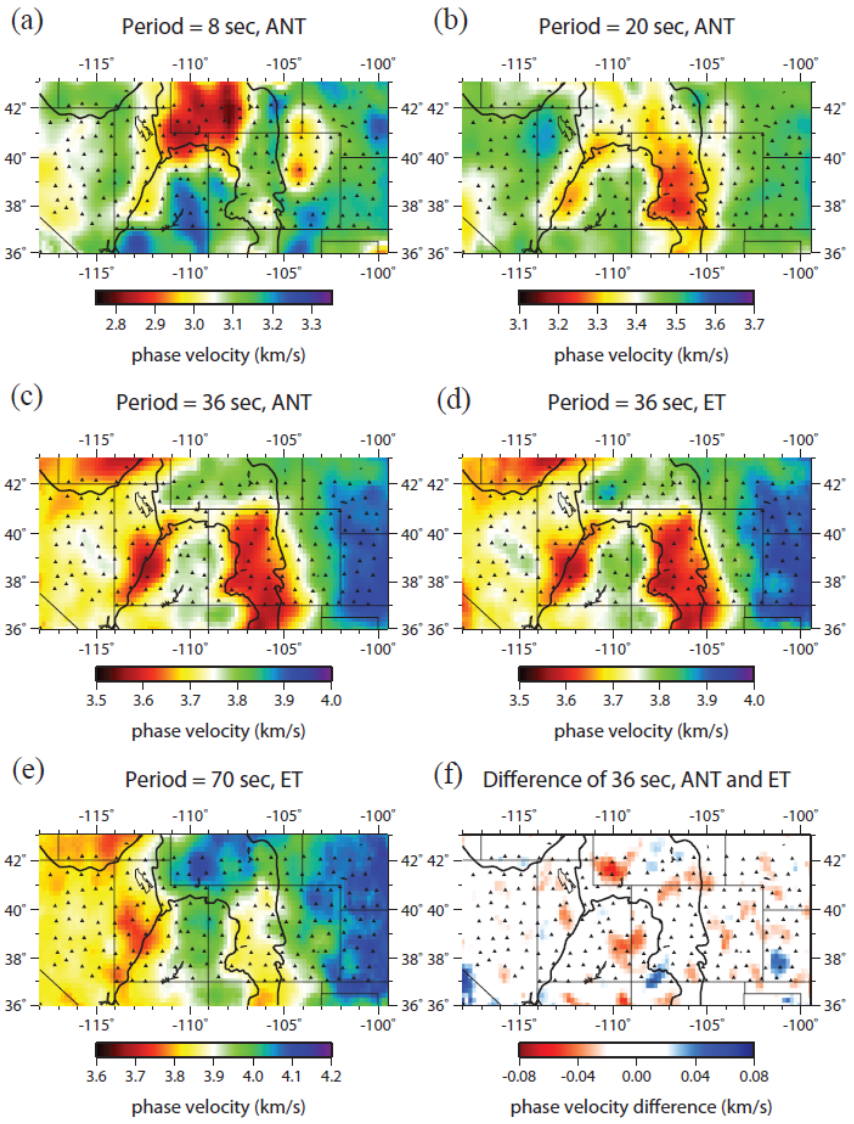

Figure 2.2 (a-c) Phase velocity maps at periods of 8, 20 and 36 sec from ambient noise eikonal tomography (ANT). (d-e) Phase velocity maps at 36 and 70 sec from earthquake Helmholtz tomography (ET). (f) Difference between the ANT and EQ results at 36 sec period.

2.2.2.2 Earthquake tomography (ET)

The data processing procedure for teleseismic earthquake data is discussed by Lin and Ritzwoller (2011) and is only briefly summarized here. First, teleseismic records at the USArray TA stations in the western US following earthquakes with surface wave magnitudes Ms ≥ 5.0 are downloaded from the IRIS DMC and cut according to the arrival window of surface waves before instrument responses are removed. Second, FTAN is applied to the cut seismograms and Rayleigh wavefront group and phase travel times and amplitudes are measured as a function of period. Phase ambiguity is resolved using a network-based approach in which travel time measurements are compared with measurements at nearby stations using the method of Lin and Ritzwoller (2011). Then Helmholtz tomography is performed to produce Rayleigh wave phase velocity maps from 32 to 80 sec period. Uncertainties are estimated similar to eikonal tomography applied to ambient noise data, but by accumulating phase velocities over many different earthquakes. In Helmholtz tomography, finite frequency effects are corrected through a term involving the local Laplacian of the amplitude field. Figure 2.2d,e shows the phase velocity maps from ET at periods of 36 and 70 sec, respectively. A comparison of the 36 sec phase velocity maps obtained from ANT and ET is shown in Figure 2.2f. The average difference is 5 m/s with slightly faster phase velocities emerging from ET.

2.2.2.3 Construction of phase velocity curves at station locations

After Rayleigh wave phase velocity maps are generated on a 0.2 °x 0.2 °grid from both ANT and ET, phase velocity curves are interpolated to each station location. At short periods (8-36 sec), phase velocity measurements and associated uncertainties are taken exclusively from ANT maps. At long periods (40-80 sec) results are taken solely from ET maps. Between periods of 36 sec and 40 sec, we weight the two velocity measurements by their local uncertainties and a smooth curve is produced. Figure 2.3 shows four dispersion curves and uncertainty estimates for TA stations R11A, T18A, Q22A and O25A whose locations are shown on Figure 2.1. In the overlapping period band, the phase velocity measurements from ANT and ET are overplotted, demonstrating the coherence between ANT measurements (red error-bars) and ET measurements (blue error-bars).

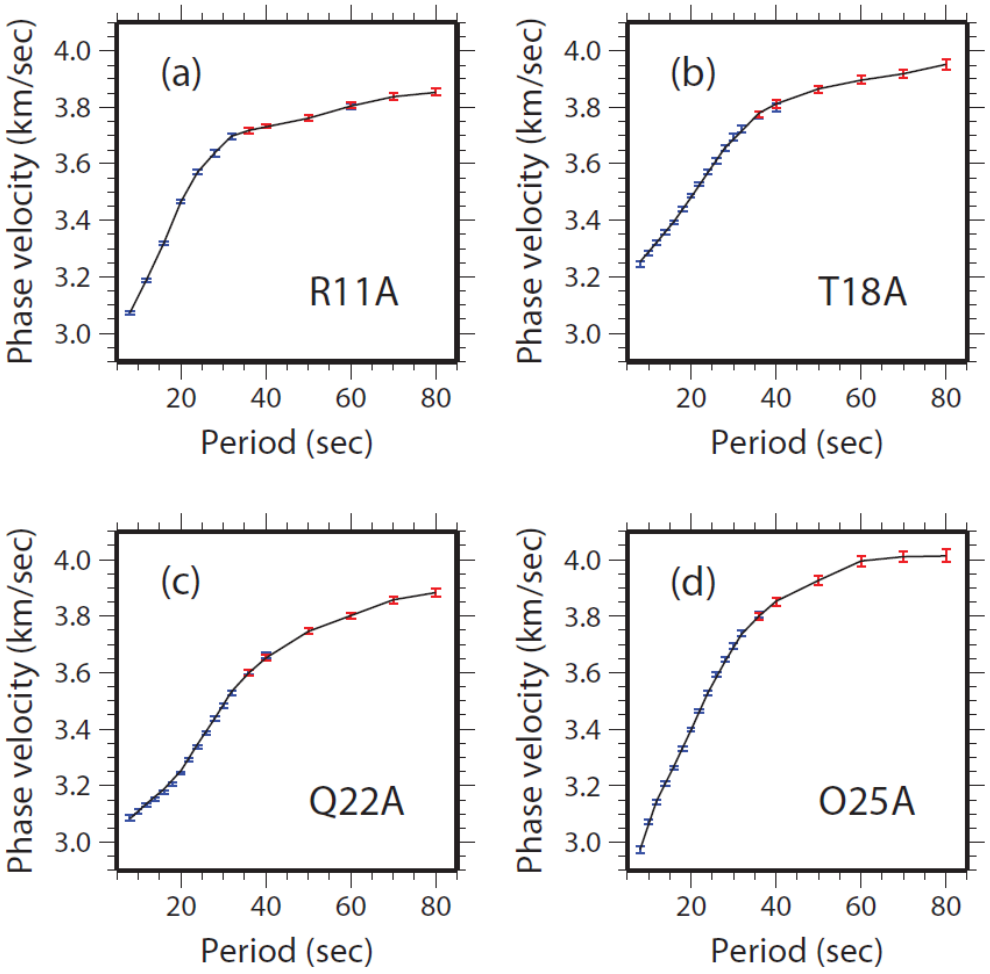

Figure 2.3 Phase velocity curves taken from the maps presented in Fig. 2.2 (and similar maps at the intervening periods) at the locations of stations R11A, T18A, Q22A and O25A, respectively. Blue symbols are uncertainties from the ambient noise tomography maps, red symbols are uncertainties for earthquake tomography maps, and the line is the joint dispersion curve from the two kinds of measurements.

2.2.3 Receiver function data processing

2.2.3.1. Receiver function data collection

We apply the time-domain iterative deconvolution method (Ligorria and Ammon, 1999) to a time window between 20 sec before and 30 sec after the predicted P-wave arrival to calculate radial component receiver functions and filtered to produce a pulse width of approximately 1 sec. P-waves at different distances have different incidence angles that affect both the timing and the amplitude of the P-to-S converted phase (Jones and Phinney, 1998; Levin and Park, 1997;

Bostock, 1998). We make corrections to the receiver functions in both time and amplitude by normalizing to a reference slowness of 0.06 deg/sec. The correction is based on a two-layer model with a crustal layer (Vs = 3.7km/sec) of 40 km thickness and a half-space mantle layer (Vs = 4.3 km/sec). Because this correction is designed to normalize the P-to-S conversion at the Moho discontinuity and reverberated phases have a different time and amplitude dependence on incidence angle, we discard the portion of the receiver function beyond 10 sec after the direct P-signal. Any shallow sedimentary multiples that are present in the receiver functions are still retained in this time series, but they are stacked down because of their different move-out behavior. After application of this correction, remaining variations in receiver functions over azimuth are mainly due to systematic effects such as dipping interfaces or anisotropy beneath the receiver, structural heterogeneity, scattering (Abers, 1998), and random noise in the seismograms. An example of the processed receiver functions observed at station R11A in the Basin and Range region is shown in Figures 2.4a, where the receiver functions are arrayed according to event back-azimuth. The receiver functions are then used in the 2nd step of quality control: harmonic stripping.

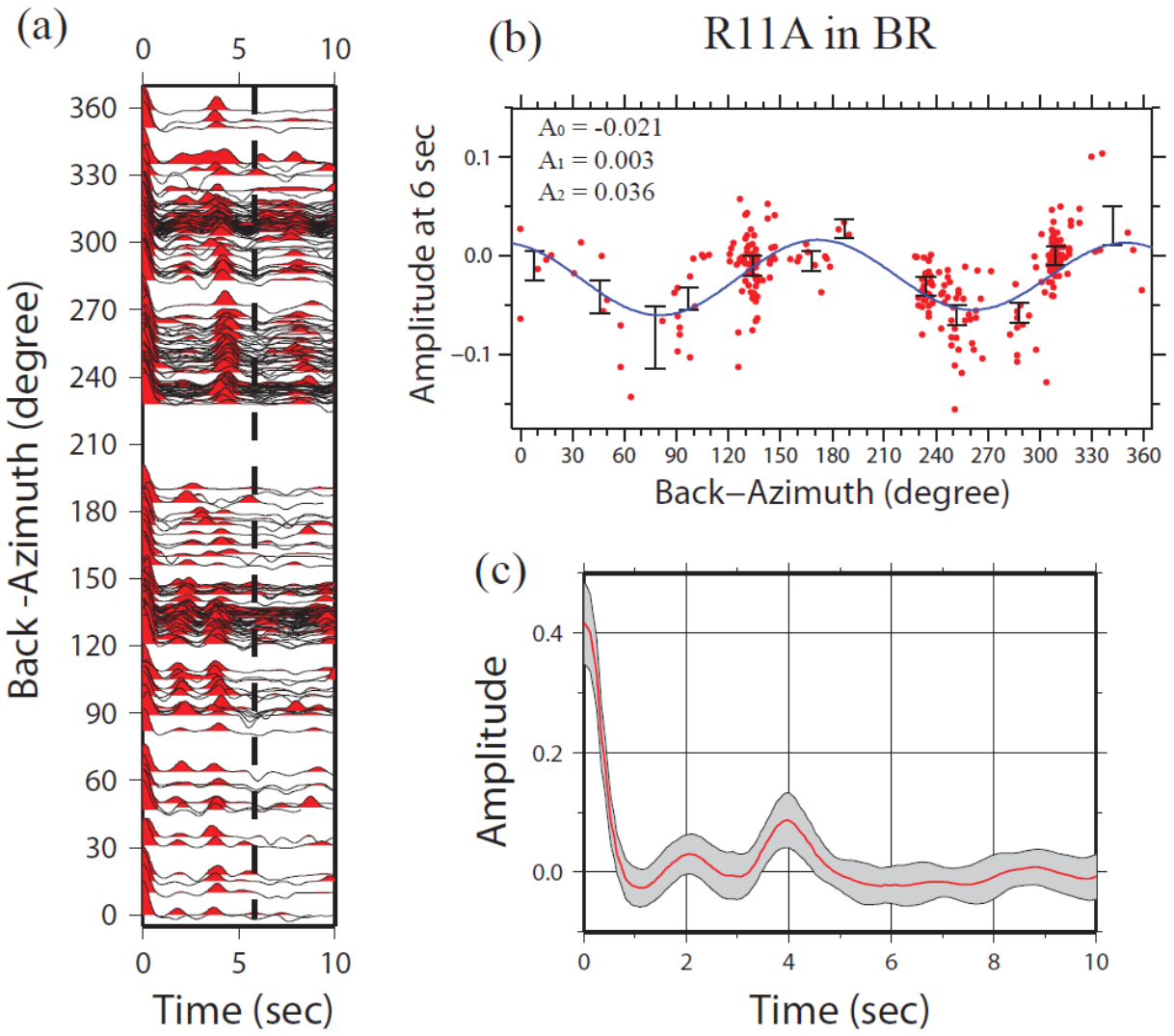

Figure 2.4. (a) Raw receiver function (RF) waveforms for station R11A in the Basin and Range are plotted as a function of back-azimuth, where 6 sec time is identified with the dashed line. (b) Red dots are RF amplitude measured at 6 sec time with azimuthally binned amplitudes plotted with black error-bars. The harmonic fitting result is shown with the blue curve and the amplitudes of the estimated coefficients A_0, A_1, and A_2 (eqn. (A1)) are shown. (c) The azimuthally independent $A_0(t)$ component of the RF that results from harmonic stripping is plotted (red curve) with the uncertainty range (grey corridor), which is the rms of the residual remaining after the harmonic fitting (eqn. (A2)) .

2.2.3.2. Harmonic stripping

Harmonic analysis of receiver functions is designed to analyze the azimuthal dependence of receiver function arrivals such that azimuthally smooth structural affects are detected and removed, notably potential dipping interfaces (e.g. Savage, 1998; Bianchi et al., 2010) and azimuthal anisotropy (e.g. Girardin and Farra, 1998). This is possible because both dipping interfaces and anisotropy generate sinusoidal variations as a function of back-azimuth with different azimuthal periodicity (Jones and Phinney, 1998). Decomposition into azimuthal components isolates contributions from horizontal symmetry axis anisotropy (two amplitude peaks and troughs over the complete backazimuthal range) and a much stronger plunging symmetry axis anisotropy signal or dipping isotropic interfaces (one peak and trough over backazimuth). The azimuthally-independent component of the resulting harmonic function represents an isotropic, horizontally-layered average of structure that is relatively unbiased by uneven and incomplete backazimuthal sampling, and is useful for joint inversion with surface waves:

$$H(\theta,t) = A_0(t) + A_1(t)\sin(\theta + \theta_1(t)) + A_2(t)\sin(2\theta + \theta_2(t)) \qquad (2.6)$$

where the A_i (i=0,1,2) are amplitudes of the three harmonic components, and the θ_i are initial phases for the azimuthally dependent components. An example of the amplitude vs. back-azimuth for the receiver functions observed at TA station R11A at 6 sec is shown in Figure 2.4b.

Under realistic circumstances where dipping interfaces and anisotropy beneath the receiver frequently exist, A_1 and A_2 will not be zero whereas A_0 will be the azimuthally-independent receiver function (Fig. 2.4c). Here, $A_0(t)$ is treated as the receiver function that is sensitive to the average horizontally-layered isotropic structure beneath the receiver and is, therefore, comparable to surface wave dispersion data. We refer to the method by which the azimuthally independent receiver function, $A_0(t)$, is estimated by fitting and removing the harmonic function $H(\theta, t)$ from the raw receiver functions as "harmonic stripping".

To estimate the uncertainty s(t) in $A_0(t)$ we use the RMS difference over azimuth between the observed receiver functions and the harmonic function defined by equation (2.7):

$$s(t) = \left[N^{-1} \sum_{i=1}^{N} \left(R_i(\theta_i,t) - H(\theta_i,t) \right)^2 \right]^{1/2} \qquad (2.7)$$

Here, $R_i(\theta_i,t)$ represents an observed quality-controlled (see below) receiver function at discrete azimuth θ_i for event i and N is the number of such receiver functions. This residual is a measure of the difference between the red dots and the fit line in Figure 2.4b, averaged over azimuth. Figure 2.5c shows the difference between the observed receiver functions and the harmonic function for station R11A. The uncertainty estimated in this way is presented in Figure 2.4c as a one standard deviation corridor about $A_0(t)$.

The RMS of the residuals over time for the receiver functions in Figure 2.4c is ~ 0.035, which is about twice the estimated noise in receiver functions (~ 0.015) found by Bodin et al. (2012), and is also about twice the RMS of the misfit between the observed receiver function and the receiver function from the best fitting model. We believe, therefore, that this definition of the uncertainty in receiver functions over-estimates the uncertainty of $A_0(t)$ that is produced from harmonic stripping. To compensate for this overestimation of the uncertainties, we make two corrections. First, we reduce uncertainties in the azimuthally independent receiver function between 3 sec and 8 sec by a factor of two in order to emphasize the fit to the P-to-S converted signal in the receiver functions. Subsequent plots of receiver function uncertainties possess this reduction (Figs. 5, 13-16). Second, we introduce a factor (κ) in the Monte Carlo inversion that acts to normalize the misfit to the surface wave and receiver function data as seen in equation (5). The determination of the appropriate value for κ is discussed in section 2.6.1.

2.2.3.3. Quality control (QC)

The observed azimuthally dependent receiver functions, $R_i(\theta_i,t)$, that go into the harmonic stripping procedure must first be quality controlled. We build on the quality control applied to the EARS (Earthscope Automated Receiver Study) database and seek only receiver functions that are consistent with the automatically QC'ed receiver functions from EARS used in their stacking procedure that are available through IRIS (Crotwell and Owens, 2005). The EARS data selection criterion is quite conservative (variance reduction in the iterative deconvolution of at least 80%) so that significant azimuthal gaps are left in the EARS receiver function station sets (e.g., Fig. 2.5b). We seek to expand the receiver function database relative to the EARS database, but retain the quality of the constitutive receiver functions in the EARS database.

To control the quality of the raw receiver functions (e.g., Fig. 2.5a for station R11A) we, first, perform harmonic stripping using the EARS data alone and get a preliminary estimate of the

harmonic function $\hat{H}(\theta,t)$. Then we consider each of the raw receiver functions in the database, filtered identically with those in the EARS database (the time windows used in the deconvolution differ somewhat between the EARS and our database, but this has little effect on the resulting receiver functions). Those that are similar to the preliminary estimate $\hat{H}(\theta,t)$ from EARS are retained, whereas those that are dissimilar are rejected. There are numerous ways to define similarity that would yield similar results. We define the difference functional between each raw receiver function at azimuth θ_i and the preliminary estimate of the harmonic fit from EARS as follows:

$$D(\theta_i) = \left\{ M^{-1} \sum_{j=1}^{M} \left[R_i(\theta_i, t_j) - \hat{H}(\theta_i, t_j) \right]^2 \right\}^{1/2} \qquad (2.8)$$

where M is the number of discrete time points in the receiver functions. We find that if we retain receiver functions where $D(\theta_i) < 0.05$, then we retain many more receiver functions than in the EARS database but at the same time do not greatly increase the noise level of the database. Figure 2.5c presents the results of this QC procedure for TA station R11A. The resulting database has been expanded by about a factor of two relative to EARS, but most of the noisy raw receiver functions have been discarded. We refer to this as the QC'ed receiver function database and it is used to estimate the final harmonic fit and azimuthally independent receiver function $A_0(t)$ with associated uncertainties s(t).

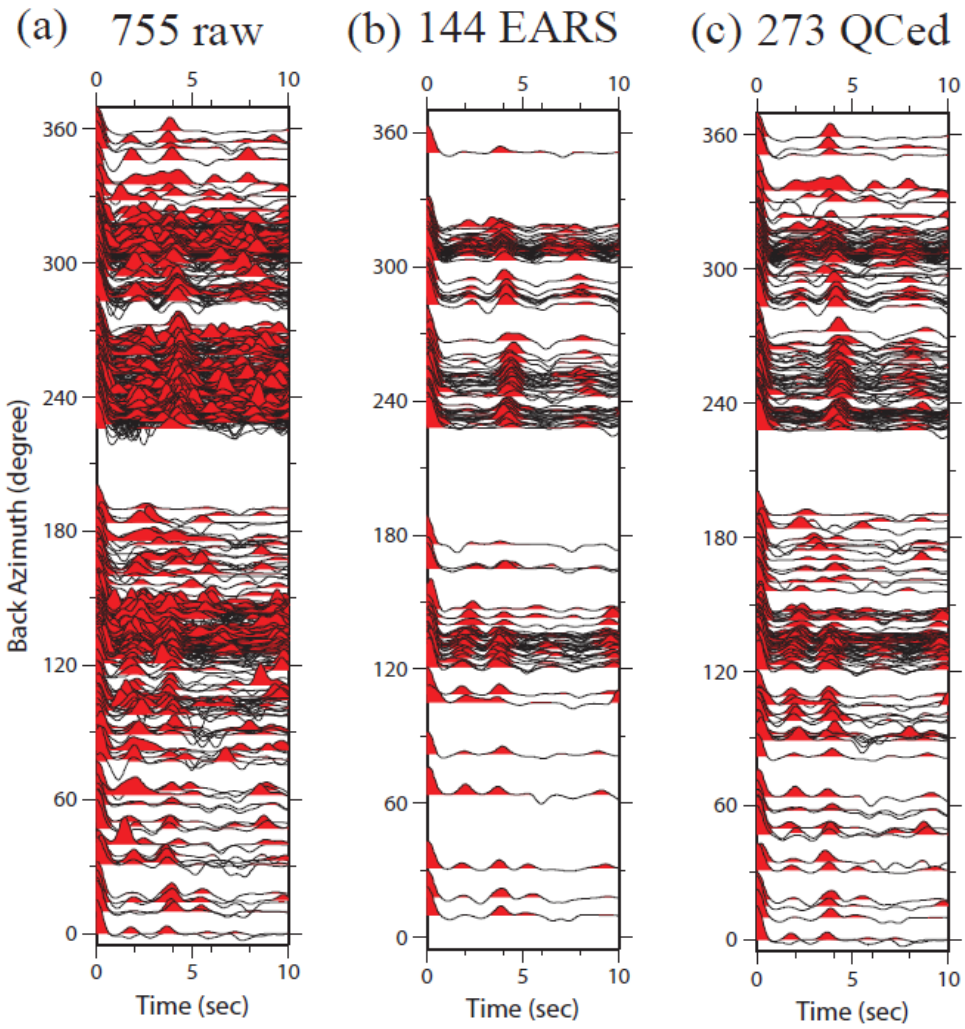

Figure 2.5. (a) The quality controlled receiver functions are plotted along back-azimuth for station R11A. (b) The estimated receiver functions, H(θ,t), from harmonic stripping. (c) The difference between (a) and (b), from which uncertainties in the azimuthally independent receiver function are determined. (d-f) The three estimated components from harmonic stripping.

2.2.3.4 Synthetic test of harmonic stripping

The goal of harmonic stripping is to remove the azimuthal periodicity of receiver function waveforms due to dipping interfaces, azimuthal anisotropy, or other structures beneath the receiver. Synthetic tests are presented here to illuminate the capabilities of harmonic stripping to remove structural effects and recovere the azimuthally independent receiver function for a horizontally-layered, isotropic effective medium.

First, we test the affects of structures that produce azimuthally smooth signals in the receiver functions; namely (1) a dipping Moho interface and (2) crustal azimuthal anisotropy, and show that the effects of these types of structures are dealt with effectively by harmonic stripping.

(1) We define a horizontally layered, isotropic model M_0 without a dipping Moho or anisotropy in Table 2.1.

Table 2.1 Horizontally-layered isotropic model M_0

Layer	Vsv (m/sec)	Vp (m/sec)	Density (kg/m^3)	Thickness (km)
1	3500	6125	2746	45
2	4300	7525	3350	~

The model, M_d, is the same as M_0 but with Moho tilted by 10° with the up-dip azimuth of 165° relative to North. Synthetic seismograms for every 10°in azimuth are computed (Frederiksen and Bostock, 2000), noise is added to the waveform before the time-domain deconvolution is performed to generate Ps receiver functions, $R(\theta,t)$, which are plotted in Figure 2.6a. The dipping Moho creates a sinusoidal variation to the P-s conversions at ~ 5.5 sec in both amplitude and arrival time. Harmonic stripping is applied to the synthetic receiver functions and the resulting estimates A_0, A_1, A_2 are shown in Figure 2.6d-f. As described by Jones and Phinney (1998), the dipping interface generates a 2π periodicity with the strongest amplitude at the updip azimuth and the arrival time will be delayed most at this azimuth. This phenomenon is observed in the harmonic stripping results. A_1 shows the strongest perturbation at 5-6 sec where the P-s conversion occurs and A_2 is subtle in this time range. Figure 2.6b,c show the harmonic fit, $H(\theta,t)$, and the difference between H and R, respectively. The random noise added to the seismograms is not mapped into the A_1 or A_2 estimates, but resides exclusively in the residuals which we take as the uncertainty of the final azimuthally independent receiver function.

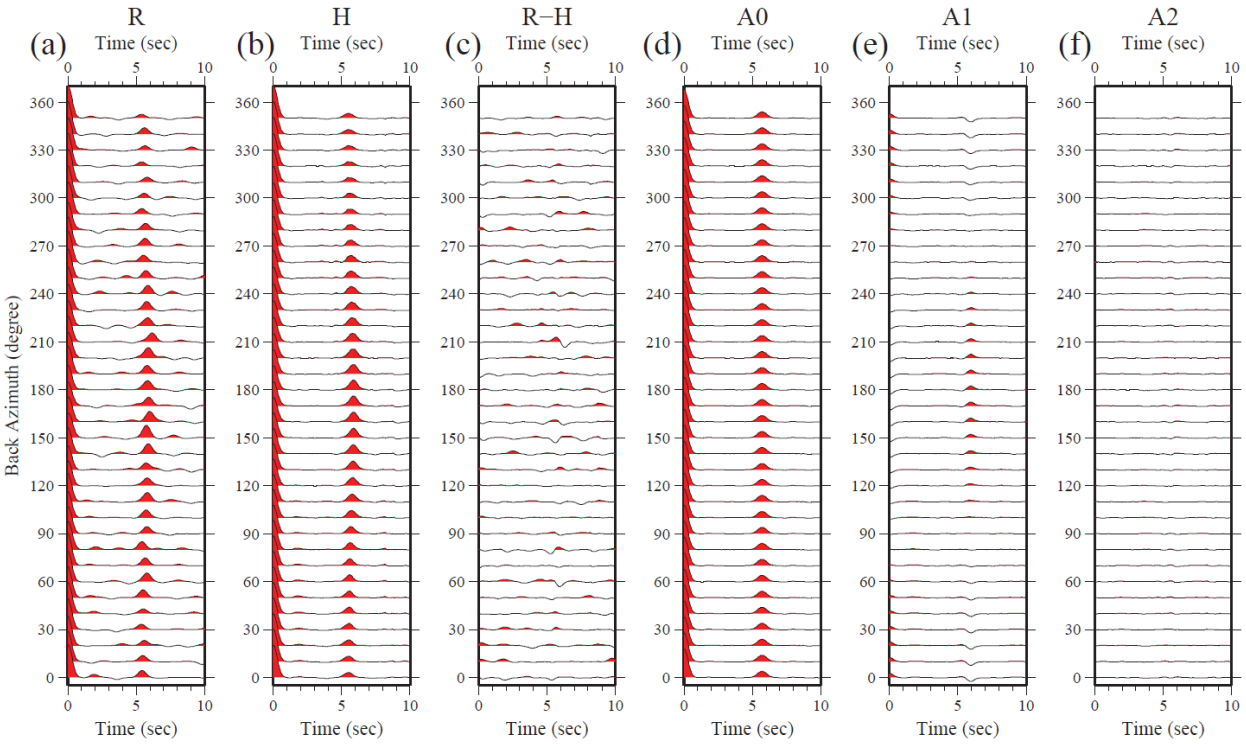

Figure 2.6. Synthetic experiment performed with Moho dipping 10 °with an up-dip azimuth of 165 °clockwise from North. (a) The synthetic receiver functions R(θ,t) including additive noise plotted at the appropriate back-azimuth. (b) The estimated receiver functions from harmonic stripping H(θ ,t). (c) The difference between (a) and (b. (d-e) The estimated A_0, A_1, A_2 components from harmonic stripping, respectively.

For comparison, the receiver functions computed from the horizontally layered model M_0 (blue waveform) and the estimate of A_0 (red dashed line) are shown in Figure 2.7a. The arrival times of the P-s conversion from both are similar, but the A_0 component recovered from the dipping model (red dashed line) displays a somewhat smaller amplitude because the harmonic stripping method does not fully reproduce the sinusoidal variation in arrival time of the input model (blue line). This inaccuracy, however, is encompassed by the enlargement of the uncertainties near the P-S conversion time as discussed further in section 2.8.2.

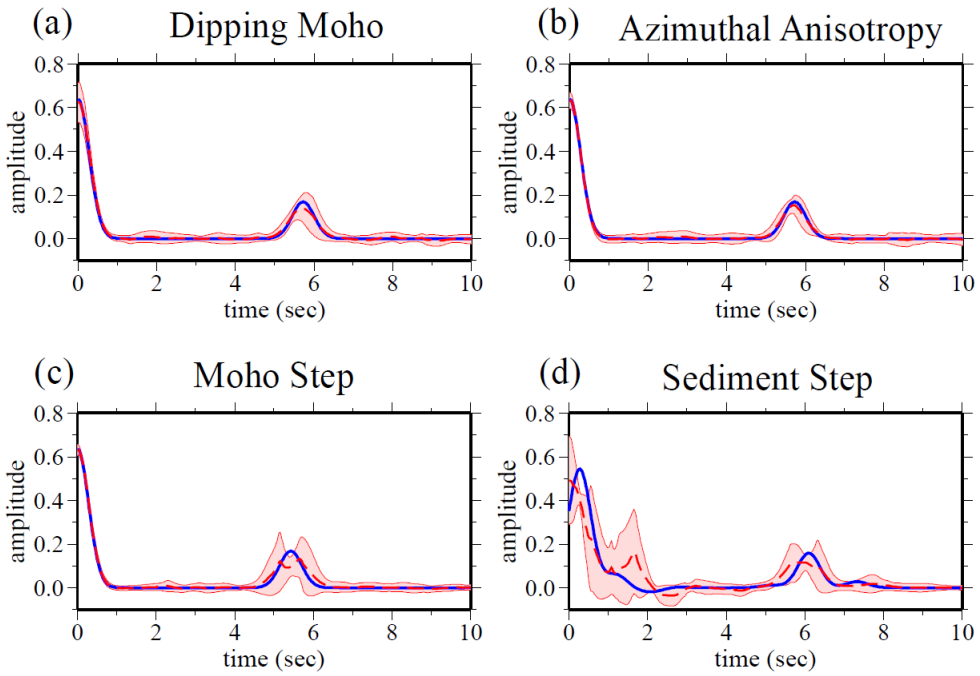

Figure 2.7. (a) The azimuthally independent A_0 receiver function estimated from harmonic stripping computed from a model with the dipping Moho is shown with the red dashed line. For comparison, the receiver function from the azimuthally averaged model (horizontal Moho) is shown with the blue line. Uncertainties s(t) are outlined by red lines with a light-red fill. (b) The same as (a), but for the A_0 receiver function estimated from synthetic receiver functions computed using a model with 4% anisotropy in the mid-to-lower crust. (c) The same as (a), but for the A_0 receiver function estimated from a model with a step in Moho depth of 5 km at the station. (d) The same as (a), but for a model with a 2 km step in sediment thickness at the station.

(2) We also test an azimuthally anisotropic model, M_{ani}, in which a 4% azimuthally anisotropic layer between 15-45 km depth range with a fast direction at 75 °to North is added to M_0. The resulting synthetic receiver functions contaminated by random noise and the results of harmonic stripping are shown in Figure 2.8. Azimuthal anisotropy introduces a π periodicity into the arrival times of the P-S conversions (Fig. 2.8a), which appears in the estimated A_2 component (Fig. 2.8f). The estimated A_0 component is shown in Figure 2.7b to agree with the receiver function computed from model M_0.

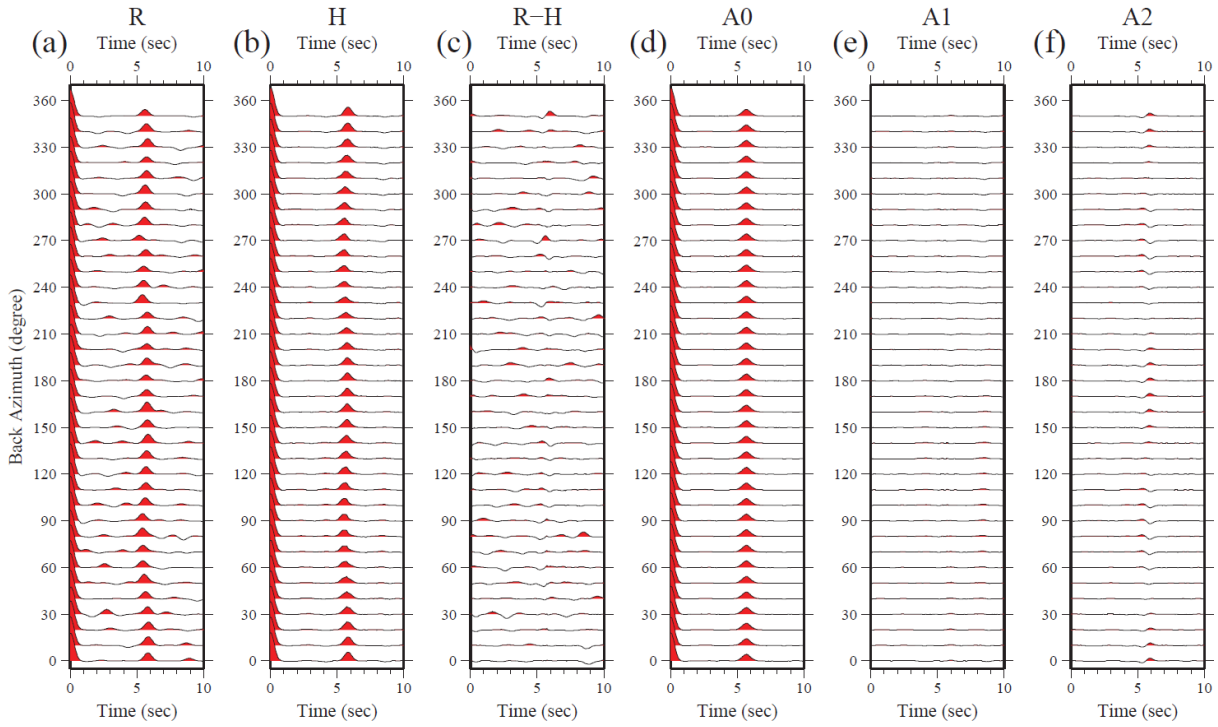

Figure 2.8 The same as Fig. 2.6, but for synthetic receiver functions computed from a model with 4% azimuthal anisotropy in the mid-to-lower crust with a fast-axis direction at 75 ° clockwise from North.

From these tests for structures that vary smoothly with azimuth, a dipping Moho or azimuthal anisotropy in the crust, we conclude that the harmonic stripping technique effectively retrieves the azimuthally-independent component which is an accurate estimate of the receiver function for an effective horizontally layered, isotropic medium.

Second, we test another type of model that possesses azimuthally discontinuous features: (3) a jump in Moho with azimuth and (4) a jump in sediment thickness with azimuth. We show that for azimuthally discontinuous structures, harmonic stripping does not reliably estimate the horizontally layered, isotropic effective medium. The method identifies this problem, however, by enlarging the uncertainties in the time band in which the estimated receiver function diverges from the receiver function of the effective medium.

(3) On the west side of the receiver (back-azimuthal angle between 180 °and 360 °) the model is the isotropic horizontally layered model M_0 with a Moho depth of 45 km, but on the east side Moho is 5 km shallower. Synthetic receiver functions from this model with additive noise are

plotted in Figure 2.9a, and the harmonic stripping results are shown in Figure 2.9d-f. The variations of the receiver functions with back-azimuth are not sinusoidal; thus harmonic stripping cannot fit them fully. As a result, the estimated A_0 component deviates from the receiver function of the average model (a model with a Moho depth of 42.5 km), which is shown as the blue waveform in Figure 2.7c. However, uncertainties in the A_0 component grow in the time window where A_0 differs from the receiver function of the average model. Thus, in this time band the receiver function will be down-weighted relative to surface wave data, as desired.

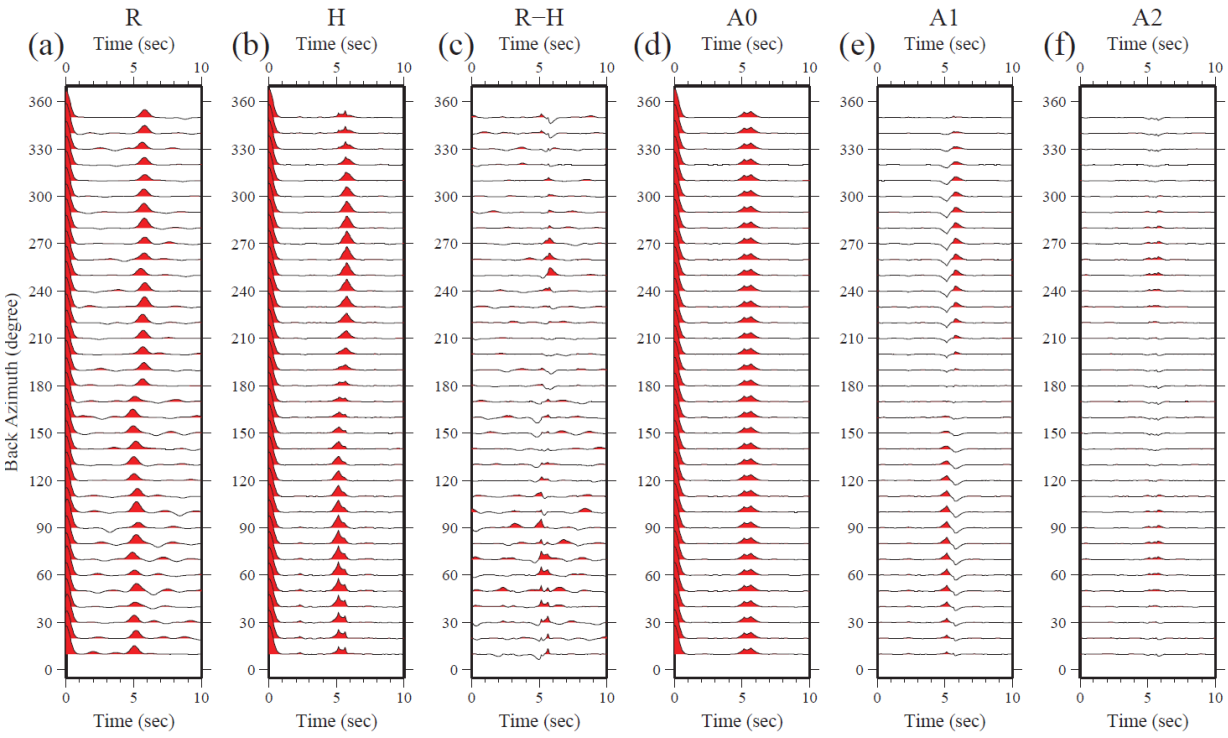

Figure 2.9 The same as Fig. 2.6, but for synthetic receiver functions from a model with a 5 km step of crustal thickness such that Moho is at 40 km depth east of the station and 45 km depth west of the station.

(4) Finally, we test of the model with a step in sediment thickness (Fig. 2.10). The model has a 2 km sedimentary layer west of the receiver and no sediments on its east side. Both models have the same Moho depth. The average model, therefore, has 1 km of sediments. The synthetic receiver functions in Figure 2.10a display a dichotomous pattern such that the Moho P-S conversion arrival times jumps near 180 °azimuth. The estimated A_0 component of the receiver function (Fig. 2.7d) shows an erroneous double Moho conversion peak as well as relatively large

uncertainties that reduce the weight of the receiver function relative to surface wave data in the inversion.

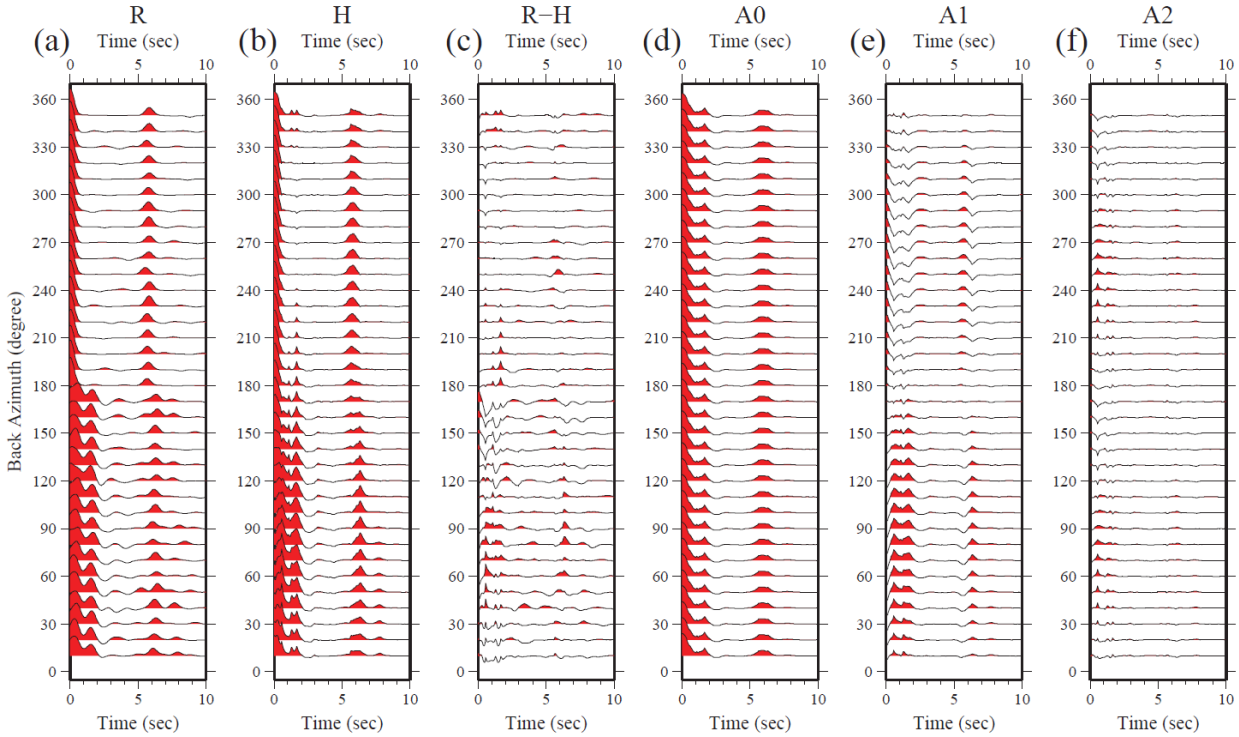

Figure 2.10 The same as Fig. 2.6, but for synthetic receiver functions from a model with a 2km jump of sedimentary layer thickness such that there are no sediments east of the station and 2 km of sediments west of the station.

2.2.3.5 Final Receiver Functions data set

The harmonic stripping method and quality control procedures have been applied to the 185 TA stations in the study region (Fig. 2.1). An example is shown in Figure 2.11 for the TA station R11A. The QC'ed receiver functions presented over azimuth are shown in Figure 2.11a and the harmonic function $H(\theta,t)$ fit to these receiver functions is shown in Figure 2.11b for comparison. As described in the Appendix, the residual between these functions is used to define the uncertainty, s(t), of the azimuthally independent receiver function and is presented in Figure 2.11c. The individual harmonic components are shown in Figure 2.11d-f. The functions $A_1(t)$ and $A_2(t)$ quantify the azimuthally dependent signals that may be produced by tilts on internal interfaces and anisotropy, respectively. Some of these arrivals are observed to undergo a move-

out at different azimuths. These functions could be applied to fine-tune the model presented here, but are not currently used for this purpose.

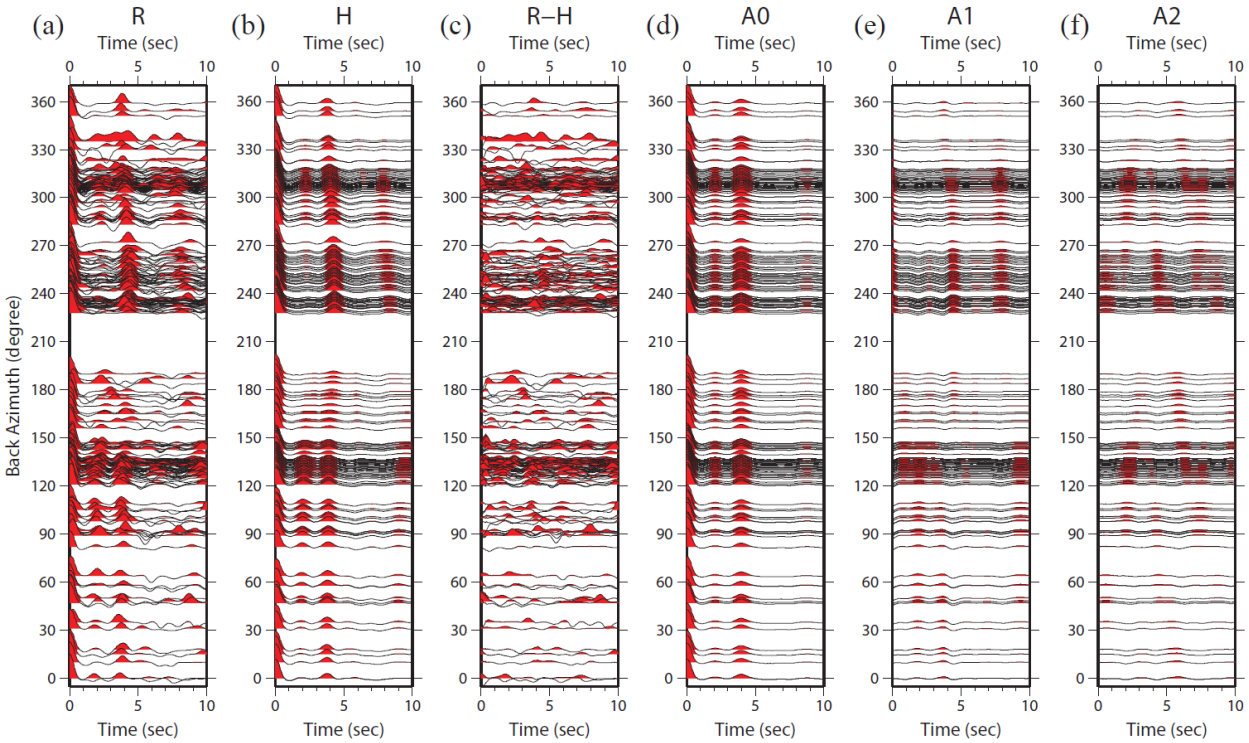

Figure 2.11 (a) The quality controlled receiver functions are plotted along back-azimuth for station R11A. (b) The estimated receiver functions, H(θ,t), from harmonic stripping. (c) The difference between (a) and (b), from which uncertainties in the azimuthally independent receiver function are determined. (d-f) The three estimated components from harmonic stripping.

Several examples of the final azimuthally independent receiver functions $A_0(t)$ and associated uncertainties are shown in Figure 2.12. The receiver functions display significant station-to-station variations. Those observed at station R11A in the Basin and Range province and station Q22A in the Rocky Mountains have strong P-to-S converted positive amplitude phases between 3 and 7 sec delay time, suggesting a sharp rather than gradient transition from crustal to mantle velocities. The receiver function for TA station T18A in the Colorado Plateau, in contrast, shows no Moho P-to-S conversion, which means that a gradient Moho is probable at this point. The receiver function for station O25A in the Great Plains has a strong negative arrival at ~ 4 sec caused by sedimentary layer reverberations at this location which probably interferes with observation of the Moho conversion.

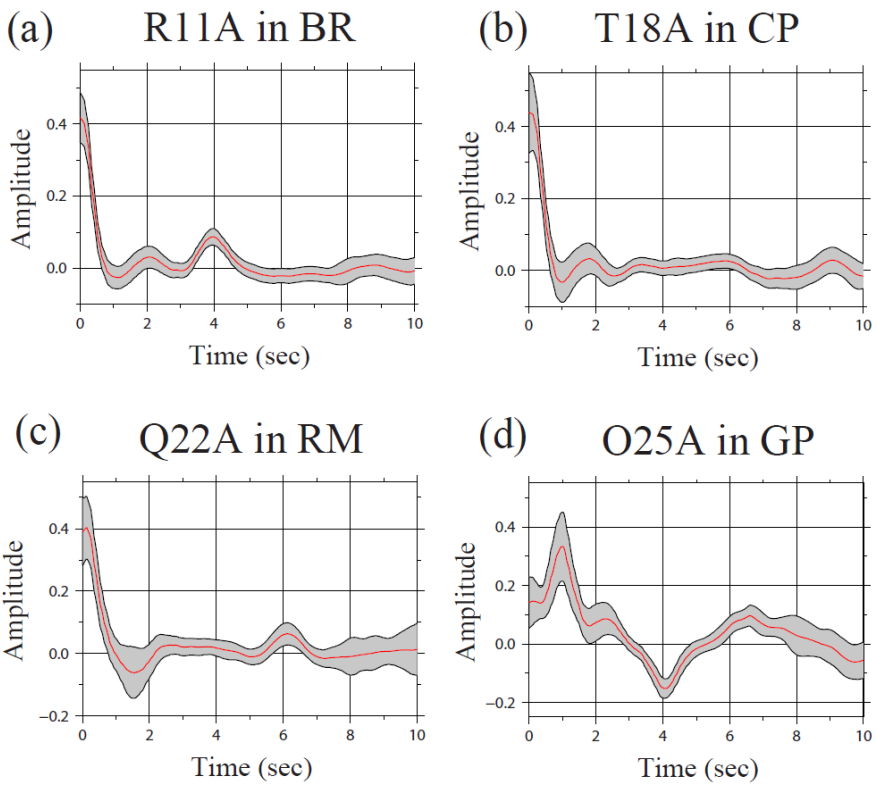

Figure 2.12: The estimated azimuthally independent receiver function ($A_0(t)$, red curve) and uncertainty (s(t), gray corridor) for four stations whose locations are identified in Fig. 2.1. The uncertainty is the rms of the residual remaining after the harmonic fitting (eqn. (A2)) to the azimuthally dependent receiver functions, reduced by a factor of two between 3 and 8 sec.

2.3 Model Parameterization and Determining the Prior Distribution

2.3.1 Model parameterization

Surface wave phase velocity curves and azimuthally averaged receiver functions are sensitive to the local average 1-D Vs structure beneath each location. For this reason, the inversion for a 3-D model reduces to a set of 1-D Vs inverse problems. We invert only for shear wave speed in the 200 km beneath the surface (reliable only to about 150 km) ecause the longest period of surface waves that we use here is 80 sec. Below 200 km, the model is assumed to be constant. Because we use only Rayleigh waves, which are predominantly sensitive to Vsv, we assume an isotropic Vsv model where Vs = Vsh = Vsv. We set the Vp/Vs ratio to 2.0 in the sedimentary layer and 1.75 in the crystalline crust and mantle. The Vp/Vs ratio is important in receiver function analysis and it is discussed further in section 8.1. For density we use the scaling relation (for 10

km) advocated by Christensen and Mooney (1995) in the crust and by Karato (1994) in the mantle. We apply a physical dispersion correction (Kanamori and Anderson, 1977) using the Q model from PREM (Dziewonski & Anderson, 1981), and the resulting model is reduced to 1 sec period.

In a traditional surface wave inversion a 1-D model may be parameterized with smooth functions in the crust and mantle (Megnin and Romanowicz, 2000), as a stack of layers (Yang and Forsyth, 2006), or a combination of both (e.g., Shapiro and Ritzwoller, 2002, Moschetti et al., 2010, Yang et al., 2008). Receiver function inversions are, however, typically parameterized with a set of fine layers (Cassidy, 1992, Julia et al., 2000). We parameterize our model similar to some surface wave inversions in which B-splines represent structure in both the crystalline crust and the uppermost mantle as follows. (1) There is one sedimentary layer with a linear gradient velocity. Three parameters are used to describe this layer: layer thickness and Vsv at the top and bottom of the layer. (2) There is one crystalline crustal layer. Five parameters are used to describe this layer: layer thickness (km) and four B-spline coefficients for Vsv. (3) There is one uppermost mantle layer to a depth of 200 km. Five parameters are used to describe this layer: five B-spline coefficients for Vsv. The thickness of this layer is controlled by the thicknesses of the top two layers.

This model-parameterization contains 13 free parameters (Fig. 2.13). We explicitly seek vertically smooth models that fit the data. Here, smooth means that the model has no more vertical structure than required to fit the data within a specified tolerance and also that the model is continuous within some depth ranges. In this model parameterization, two discontinuities are introduced because the predominant signals in receiver functions are from the Moho and, in the presence of sediments, from the sediment-basement contact. We demonstrate that a smooth parameterization can explain both the surface wave and receiver function data within data uncertainties across almost the entire study region. In other areas where a discrete higher-velocity lower crustal layer is observed (e.g., immediately north of our study area; Gorman et al., 2002), an additional crustal interface may be needed. The smoothness of the model is imposed by the parameterization so that ad hoc damping is not needed during the inversion.

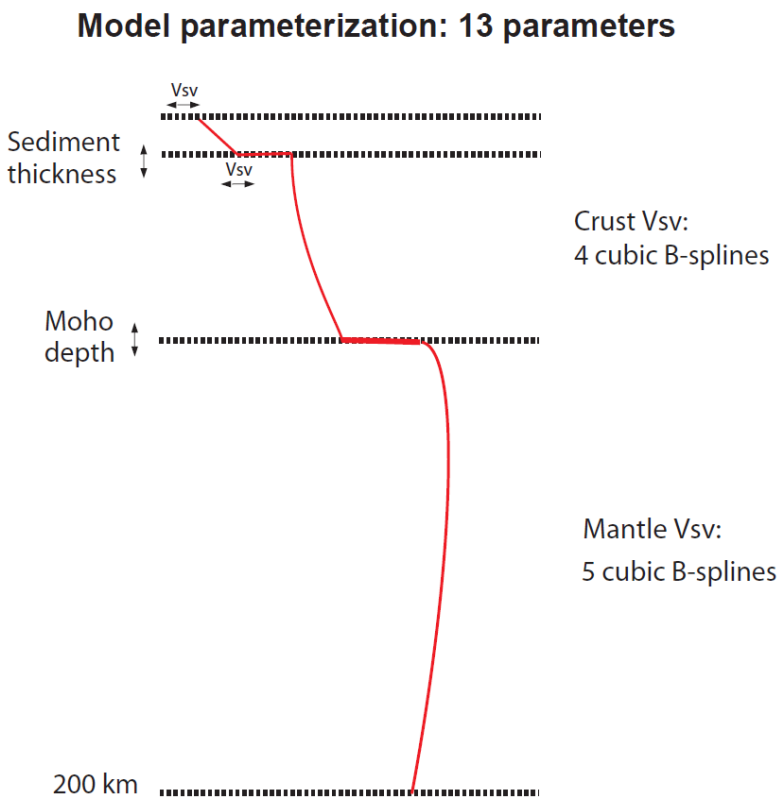

Figure 2.13. Model parameterization illustrating the 13 model parameters used in the Monte Carlo sampling of model space.

With this model parameterization, we construct the model space in which the prior and posterior distributions are determined. The construction of the model space is based on perturbations to a reference model, where sedimentary structure is taken from Mooney and Kaban (2010), crustal thickness from Bassin et al. (2000), and shear wave speeds in the crust and mantle from Shapiro and Ritzwoller (2002), as summarized in Table 2.2. At each location models are sampled around the reference in the range prescribed in the table. For example, the crystalline crustal thickness variation is ±25% of the input value. If the reference thickness were 40 km, values would be considered from 30 to 50 km.

Table 2.2 Model space and references.

Model Parameters	**Range**	**Reference**
Sediment thickness	0-2$\mathbf{m_0}$ (km)	Mooney and Kaban (2010)
Crystalline crustal thickness	$\mathbf{m_0}$ ±0.25$\mathbf{m_0}$ (km/s)	Bassin, C. et al. (2000)
Vsv, top of sedimentary layer	$\mathbf{m_0}$ ±0.2$\mathbf{m_0}$ (km/s)	Bassin, C. et al. (2000)
Vsv, bottome of sedimentary layer	$\mathbf{m_0}$ ±0.2$\mathbf{m_0}$ (km/s)	Bassin, C. et al. (2000)
B-spline coefficients, crust	$\mathbf{m_0}$ ±0.2$\mathbf{m_0}$ (km/s)	Shapiro and Ritzwoller (2002)
B-spline coefficients, mantle	$\mathbf{m_0}$ ±0.2$\mathbf{m_0}$ (km/s)	Shapiro and Ritzwoller (2002)

2.3.2 Determining the prior distribution

The prior distribution reflects the state of knowledge before data are introduced. Prior information that we apply in choosing models includes the following seven prior assumptions. (1) Models exist in the model space M, which is defined as the reference model subject to allowed perturbations (Table 2.3). (2) The model is continuous between Vsv discontinuities at the base of the sediments and Moho and is continuous in the mantle. (3) Velocity in the sedimentary layer increases with depth. (4) Velocity in the crystalline crust increases with depth (monotonicity constraint). (5) Velocity contrasts across the sedimentary basement and across the Moho discontinuity are positive. (6) Vs < 4.9 km/s throughout the model. (7) When surface wave data are used alone we apply the assumption of a positive velocity gradient in the uppermost mantle. Up to constraint (7), the prior distributions for inversions using only surface wave data or surface wave data jointly with receiver functions are identical. We discuss in section 7.3 how constraint (7) may introduce bias in the estimated model, and how it can be eliminated when receiver functions are introduced. Together, these assumptions reduce model complexity and also ameliorate some of the velocity-depth trade-offs that occur in the surface wave inversion.

To determine the prior distribution, we perform a random walk in model space with the following steps. First, we initiate the model search at a random point in model space M. If this random point does not obey the assumptions listed above, it is rejected and another random point is chosen. Second, we introduce a random jump in model space. We simultaneously perturb all 13 parameters by selecting a new value randomly governed by a Gaussian probability

distribution near the old value subject again to the previous assumptions. This Gaussian distribution is controlled by the width of the distribution (standard deviation), which is given in Table 2.3. These Gaussian widths have been chosen empirically to ensure the efficiency of the random walk in sampling the prior and posterior distributions. If the model produced is not contained in model space M (i.e., does not satisfy the seven constraints in the previous paragraph and does not fall within the ranges presented in Table 2.3), then the model is rejected and the process re-initiates.

Table 2.3 Width of the Gaussian distribution in the model perturbation.

Parameters	Gaussian Width
Thickness, sedimentary layer	0.1 km
Thickness, crystalline crust layer	1 km
B-spline coefficients, crust	0.05 km/sec
B-spline coefficients, mantle	0.05 km/sec
Velocity at top/bottom of sedimentary layer	0.05 km/sec

With a sufficient number of sample points (> 100,000), the prior distribution of each parameter can be viewed as histograms of marginal distributions of model characteristics at different depths such as those shown in Figure 2.14 at the location of the TA station R11A. For this figure we generated Vsv models from such parameter distributions by converting the B-spline coefficients to Vsv values as a function of depth. Due to the lack of prior information imposed on the thickness of the sedimentary layer or on the thickness of the crystalline crust, the marginal prior distributions of both appear as uniform distributions (Fig. 2.14a,c). However, the marginal distributions of Vs at different depths are not uniform because different parameters interact through the prior assumptions. The resulting marginal distributions for these variables are more similar to tilted Gaussian patterns because they are affected by prior restrictions to sampling such as the monotonicity constraint on crustal velocities.

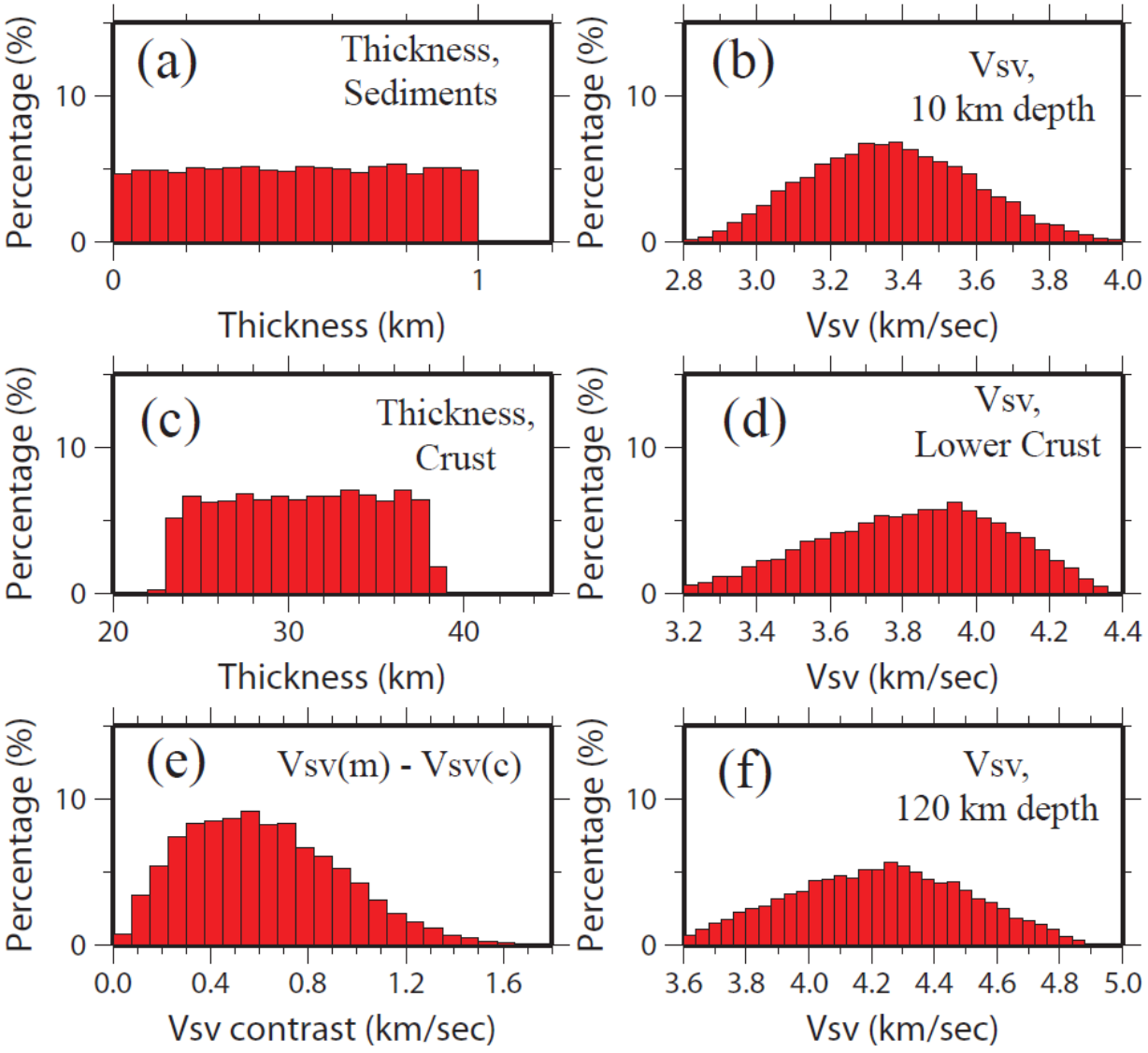

Figure 2.14. Prior distribution for several of the model parameters at the location of TA station R11A: (a) sediment thickness, (b) Vsv at 10 km depth, (c) crustal thickness, (d) Vsv in the lower crust, (e) velocity contrast from 4 km above to 4 km below Moho (mantle – crust), and (f) Vsv at 120 km depth.

2.4 Monte Carlo Inversion of Surface Wave Dispersion Data

In this section, we discuss the Bayesian Monte Carlo inversion based on surface wave data alone for later comparison with the joint inversion with receiver functions, which is presented in section 6. As above, we use the USArray TA station R11A in the Basin and Range province for examples and the input data for the inversion are shown in Figure 2.3a. At each location,

Rayleigh wave phase velocities exist on a fixed discrete grid from 8 to 80 sec period. Prior constraint (7) is not applied in this section.

2.4.1 The likelihood functional

For a problem in which the measurements are observed Rayleigh wave phase speeds at a particular location, which have independent but not identically distributed Gaussian observational uncertainties σ_i (period index i), the likelihood functional is computed from equation (2) where the misfit function comes from the first term on the right hand side of equation (5). The assumption of the frequency independence of error processes in surface wave dispersion is made here and in other studies (Liu et al., 2010, Bodin et al., 2012), because the covariance over frequency is not well understood. The correlation of the dispersion velocities over frequency deserves further investigation.

2.4.2 Determining the posterior distribution

The information about model space extracted by introducing the data is described by the posterior distribution. As defined by equation (1), the posterior distribution is the product of the prior distribution and the evaluation of fit to the observed data, which is summarized by the likelihood functional L. During model space sampling, when a new model m_i is generated by perturbing a given model m_j under the prior assumptions, the likelihood functional of this model is calculated through forward computation using the Thomson-Haskell method (computed using the code of Herrmann, http://www.eas.slu.edu/eqc/eqccps.html) with an earth-flattening transformation. Both of the likelihoods $L(m_i)$ and $L(m_j)$ are computed as discussed in section 5.1, where the Metropolis law, equation (4), defines the probability of acceptance for model m_i, as described in section 2.

Figure 2.15 shows an example of how the likelihood functional and misfit function of the models accepted during sampling of the posterior distribution evolve for TA station R11A. Accepted models are identified with circles, rejected models are not plotted. The trend shows that after about 100 samples, the likelihood functional will fluctuate near unity during the remaining iterations. We choose to stop iterating after 3000 sampling steps. To ensure that the starting point does not affect the sampling, we initiate the Monte Carlo sampling at ten different random models and the resulting statistics reflect these ten independent samplings. Figure 2.14 also

shows how the RMS misfit function, $\chi = \sqrt{N^{-1}S(m)}$, evolves. When $\chi < 1$, the fit lies within the estimated uncertainties, on average.

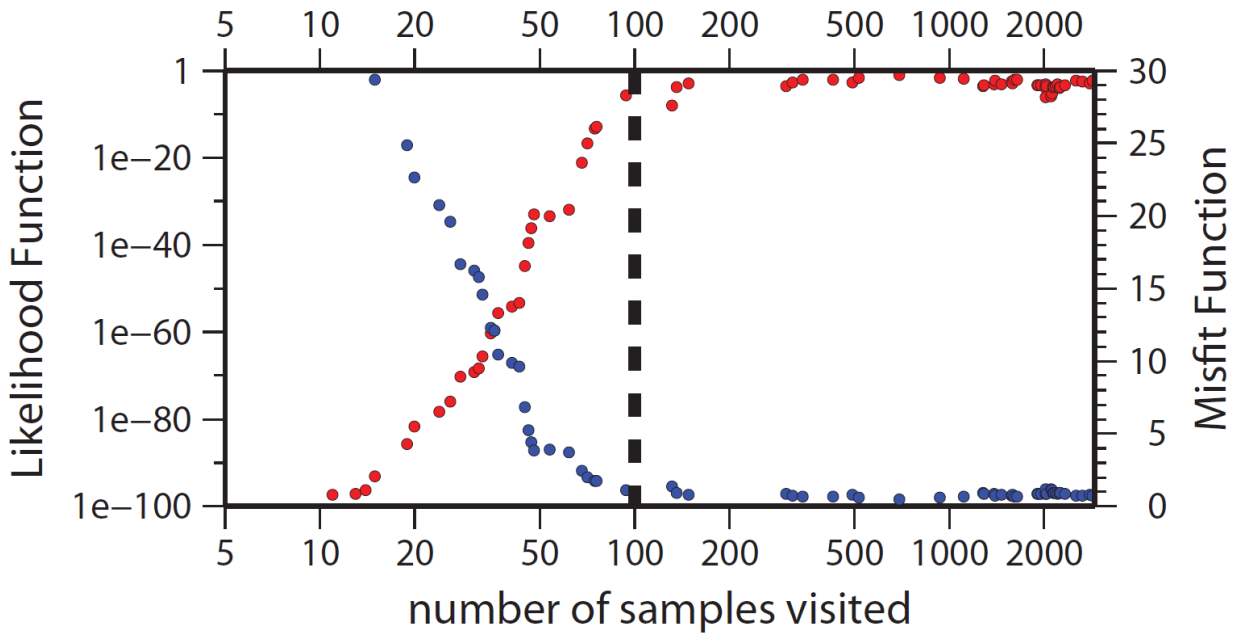

Figure 2.15 Convergence of the sampling of the posterior distribution of models when only surface wave dispersion data are used at the location of station R11A. Models are provisionally accepted according to the Metropolis Law in eqn. (4), and then replaced by other models as the procedure evolves. Red dots are the likelihood function (eqn. (2)) for each model that is accepted during the Metropolis sampling of model space, while blue dots are the rms-misfit (first term on the RHS of eqn. (15)). The convergence point is at ~100 models (dashed line).

As discussed in section 1, an additional model acceptance criterion is introduced as the basis for accepting models to form the posterior distribution. The minimum misfit χ_{min} is defined as the minimum value of χ found for all models visited during model space sampling. Once χ_{min} is found, we define the threshold of acceptance χ_{crit} for the surface wave inversion as follows:

$$\chi_{crit} = \begin{cases} 2\chi_{\min} & \text{if } \chi_{\min} \geq 0.5 \\ \chi_{\min} + 0.5 & \text{if } \chi_{\min} < 0.5 \end{cases} \qquad (2.9)$$

Thus, to define the ensemble of accepted models a posteriori, we accept any model m that is visited during the inversion as long as $\chi(m) \leq \chi_{crit}$. The choice of the critical threshold level below which models are accepted to form members of the posterior distribution is admittedly ad-

hoc and deviates from strict Bayesian practice. Equation (2.9) succeeds to produce posterior distributions that capture our degree of belief in the final models, however.

From the set of accepted models in the posterior distribution, we compute the distribution of each parameter and the marginal distribution of Vsv at each depth. Figure 2.16 presents example histograms of marginal distributions for several structural variables derived using surface wave data alone to compare with the prior marginal distributions shown in Figure 2.14.

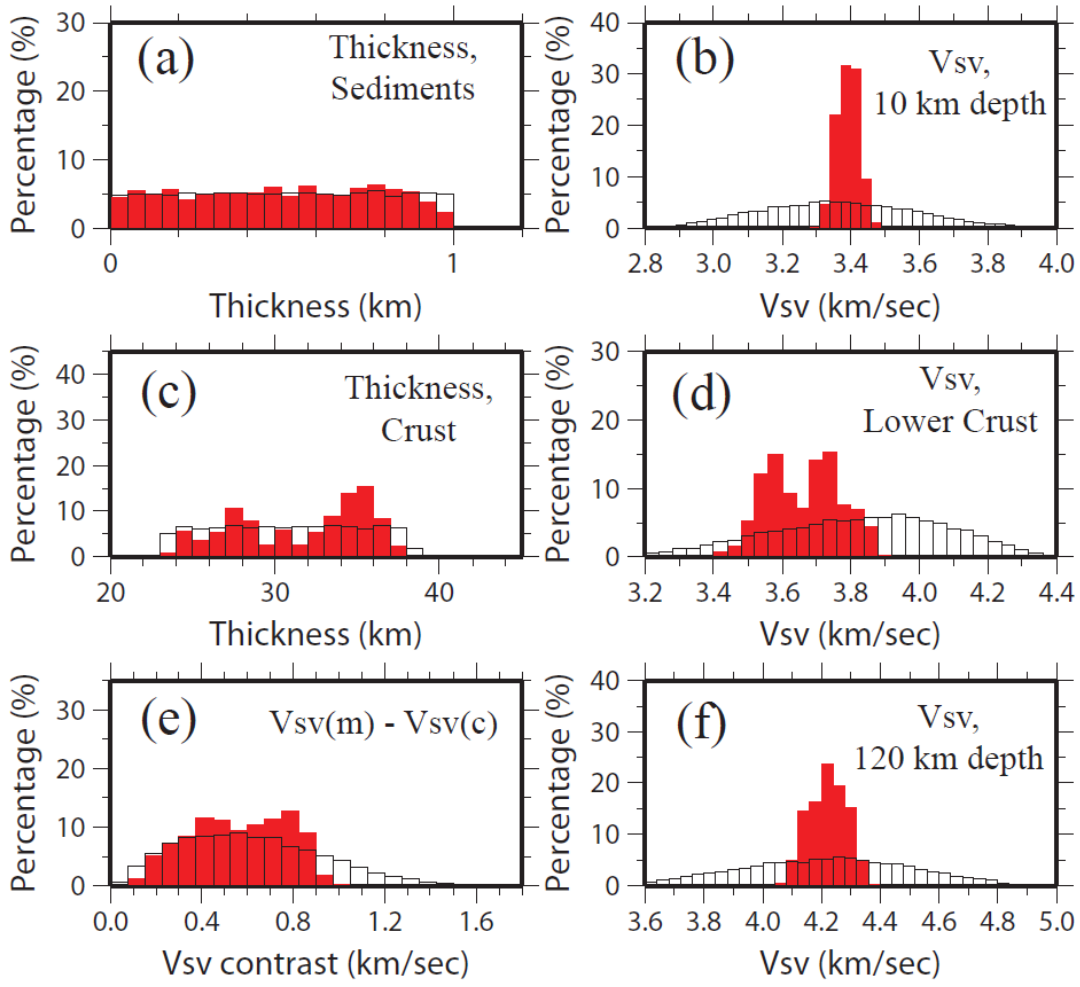

Figure 2.16. (a)-(f) The same as Fig. 2.14 for TA station R11A, but here the posterior distributions after Monte Carlo sampling using surface wave data alone (models fit the surface wave data) are plotted. The prior distributions from Fig. 2.14 are plotted as write histograms in the background. Surface wave data reduce the spread of structural velocities between boundaries, but have less effect on the depth to boundaries or velocity jumps across boundaries due to trade-offs between the parameters near the boundaries.

Typically, the distributions of velocities between boundaries are narrowed. For example, Vsv at 10 km depth in the crust is narrowed; its standard deviation decreases by about a factor of six from ~ 0.3 km/sec to ~ 0.05 km/sec. However, posterior distributions of discontinuity depths (Moho, sediments) or velocity jumps across the discontinuities do not decrease as appreciably. For example, crustal thickness and Vsv contrast across Moho are not changed strongly. These histograms demonstrate the sensitivity of Rayleigh wave dispersion to Vs structure. (1) Sedimentary thickness does not change because of the lack of very short period (< 8 sec) surface wave dispersion measurements. (2) Vsv at 10 km depth shows a much narrower distribution compared with the prior distribution because ambient noise provides information about this depth. (3) The mean of the crustal thickness distribution is about 33 km with a 1σ width of ~ 3.6 km, showing that surface wave dispersion at this location possesses only weak sensitivity to the depth of the Moho discontinuity. The distribution is bimodal and not strongly peaked. (4) The marginal distributions of Vsv in the lower crust and the Vsv contrast across the Moho are narrowed somewhat but retain considerable uncertainties (0.15-0.2 km/sec) and the lower crustal distribution is also bimodal. (5) Vsv in the uppermost mantle (120 km) is narrowed appreciably because the earthquake data are sensitive to this depth, but it is not constrained as well as at 10 km in the crust. This demonstrates that the vertical resolution of surface waves to local structure degrades with depth. Therefore, we only report structure to 150 km. These findings make intuitive sense and are quantified with the posterior distributions.

An alternative view of the results is provided by computing the mean, the median, the 1σ uncertainty, and the full range of Vsv at each depth after all accepted models are identified. Figure 2.17 shows the resulting extent of accepted models for station R11A. The gray corridors in Figure 2.17a,b outline the extent of all accepted models that fit the data, while the red lines present the 1σ width around the mean model. Note that without assimilating receiver functions, the crustal thickness (Moho depth referenced to the surface in Fig. 2.17b) is not well resolved and the predicted receiver functions (red waveforms in Fig. 2.17c) do not fit the observed azimuthally independent receiver function well, on average. There is also a strong trade-off between the lower crustal velocity and the uppermost mantle velocity. We show in section 6 that these problems are ameliorated with the addition of receiver functions in the inversion.

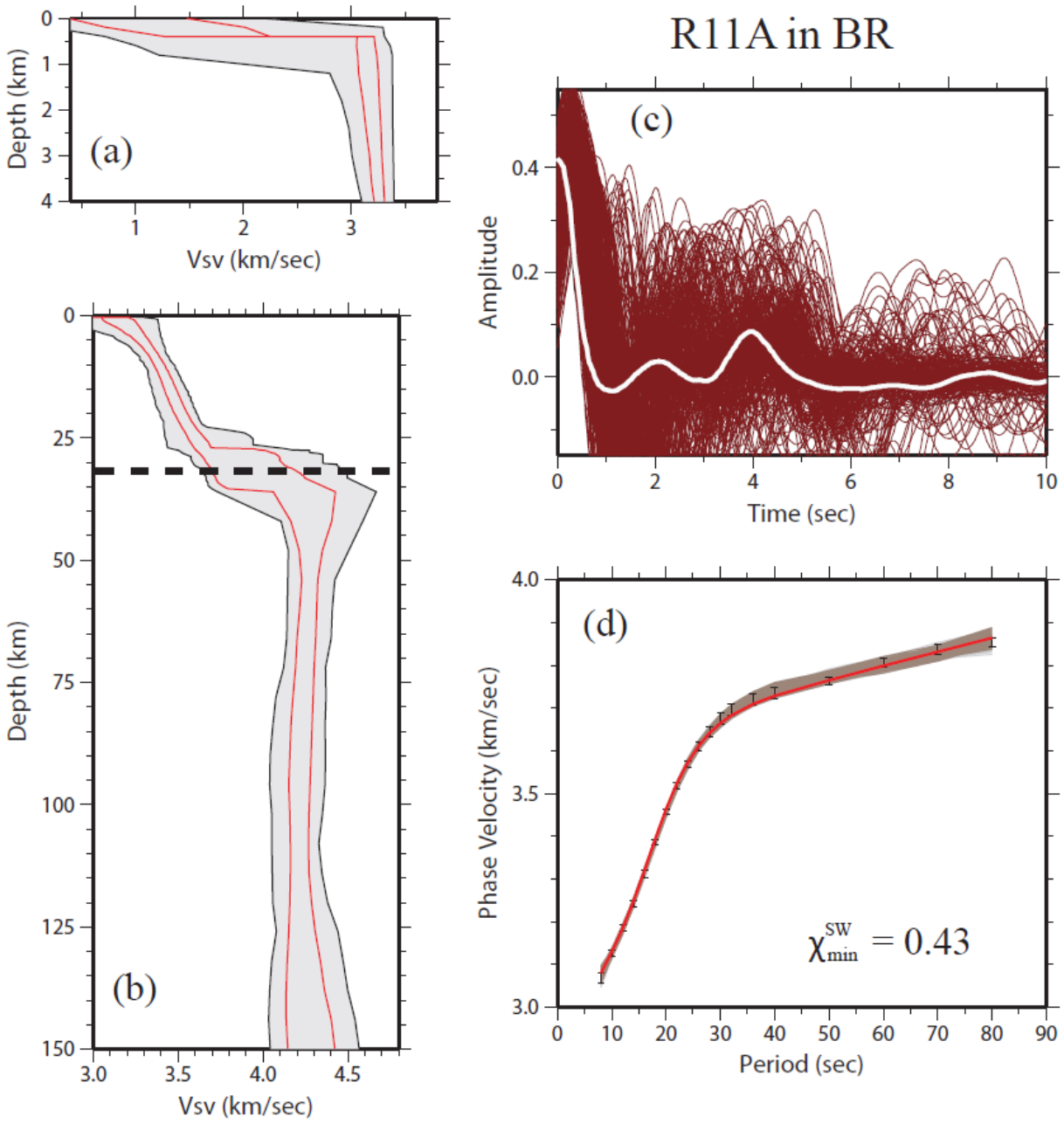

Figure 2.17 (a-b) The model ensemble at the location of TA station R11A in the Basin and Range province resulting from the Monte Carlo inversion of surface wave data alone is shown in full width (black lines with gray fill) and 1σ width (red corridor. Average Moho depth is identified as a dashed line at ~ 32 km. (c) The observed receiver function (white line) is plotted with predicted receiver functions (maroon lines) computed from all accepted models (eqn. (7)), showing that receiver functions are not well fit, on average, by models constrained by surface wave data alone. (d) The observed Rayleigh wave phase velocity dispersion data (black error bars) are plotted with predicted surface wave phase velocity curves computed from all accepted models (eqn. (7), gray lines). The red curve is the predicted phase velocity curve from the best fitting model.

2.5 Monte Carlo Inversion of Surface Wave Dispersion and Receiver Function Data

When receiver functions are assimilated into the Monte Carlo algorithm, the prior distribution of models largely remains the same, although as we discuss in section 7.3 we release the constraint on the positivity of the uppermost mantle velocity gradient. Here, we introduce the joint likelihood functional and then describe the posterior marginal distribution at TA station R11A and then for three other TA stations.

2.5.1 The joint likelihood functional

The joint likelihood functional is defined by equations (2) and (5), but because we do not know the diagonal elements of the data covariance matrices we need to estimate the relative scaling of the misfits for surface waves and receiver functions that is quantified through parameter κ. We consider values of κ ranging from 1 to 40, and choose 2.5 as a value that strikes the appropriate balance between the two data sets. By choosing this value both data sets are approximately weighted equally in the inversion and the resulting misfits of the two data sets are comparable. The joint misfit function of equation (2.8), therefore, becomes:

$$S_{\text{joint}}(m) = S_{SW} + \frac{1}{2.5} S_{RF} = N\chi^2_{SW} + \frac{M}{2.5}\chi^2_{RF} = \sum_{i=1}^{N} \frac{\left(g_i(m) - D_i^{obs}\right)^2}{\sigma_i^2} + \frac{1}{2.5}\sum_{j=1}^{M} \frac{\left(R_j(m) - A_0(t_j)\right)^2}{s_j^2} \quad (2.10)$$

where we have implicitly defined the χ^2 misfit for surface waves and receiver functions. RMS misfit is the square root of χ^2.

2.5.2 Determining the posterior distribution in the joint inversion

Figure 2.18 shows how the joint likelihood functional for station R11A evolves as model space is sampled. The convergence duration for this sampling is about 500 samples, which requires about five times more models than considered when surface wave data are used alone (Fig. 2.15).

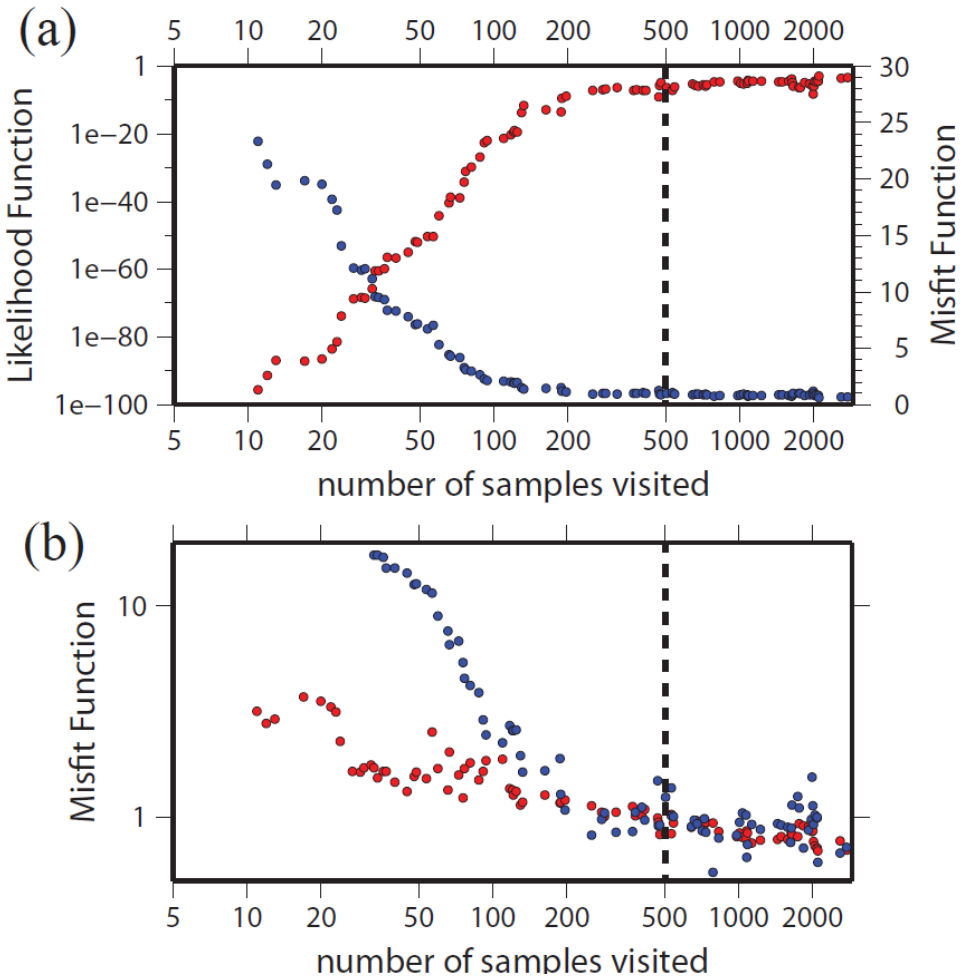

Figure 2.18 (a) The joint likelihood function (from eqns. (2) and (8), red dots) and the joint misfit function (eqn. (8), blue dots) for TA station R11A as a function of number of models sampled. (b) Misfit for each individual data set in the same search as (a). Blue dots are for surface wave phase velocity data and red dots are for the RF data. When the model converges to maximize the likelihood functional, both misfits typically converge to < 1. In both plots, the location of 500 sampled models is identified with a dashed line, where the fit approaches convergence.

As with the surface wave inversion, we repeat the procedure starting from ten different random points to find the range of models that fit the data acceptably. The forward calculation of receiver functions is based on the code developed by Shibutani et al. (1996), which has also been used by Sambridge (1999a). We introduce for the joint inversion a somewhat different criterion than for the surface wave inversion as the basis for accepting models to form the posterior distribution. For each location, minimum RMS misfits for surface wave dispersion $\chi^{SW}_{\min}$ and receiver functions $\chi^{RF}_{\min}$ are identified, although the models that minimize misfits for these two types of

data are generally not the same. The joint normalized relative RMS function χ^{joint} is defined as follows:

$$\chi^{\mathbf{joint}} = \frac{1}{2}\left(\frac{\chi^{SW}}{\chi^{SW}_{\min}} + \frac{\chi^{RF}}{\chi^{RF}_{\min}} \right) \tag{2.11}$$

This function is a combination of relative RMS misfits for the two data sets. The criterion for model acceptance is $\chi^{\text{joint}} < \chi_{crit}$, where χ_{crit} is defined as:

$$\chi_{crit} = \chi^{\text{joint}}_{\min} + 0.5 \tag{2.12}$$

For example, in the particular inversion for TA station R11A, the minimum misfit for the surface wave data is $\chi^{SW}_{\min}$ = 0.51, the minimum misfit for the receiver function data is $\chi^{RF}_{\min}$ = 0.53, and the minimum misfit to both data in the joint inversion is $\chi^{\text{joint}}_{\min}$ =1.15. Thus, the best fitting model that emerges from the joint inversion does not fit either data type optimally, but fits them both acceptably. In total, about 1000 models are found below the critical threshold $\chi_{crit} = 1.65$.

Example posterior marginal distributions are plotted and compared with prior (marginal) distributions in Figure 2.19. All the marginal distributions change shape relative to the prior distributions, especially for Vsv at different depths and for crustal thickness. Compared with the posterior distributions using surface wave data alone (Fig. 2.16), the distributions of velocities in the crust and upper mantle change only subtly (Fig. 2.19b,f). The other distributions change profoundly. For example, crustal thickness at station R11A sharpens (from σ = 3.6 km to σ = 1.3 km). Concerning the velocity difference across the Moho, the mean increases and the standard deviation approximately halves. Such significant changes are expected because receiver functions are most sensitive to velocity contrasts across layer boundaries beneath the station; therefore, the position of the Moho is better determined as are the values of model variables that trade-off with Moho depth in the surface wave inversion.

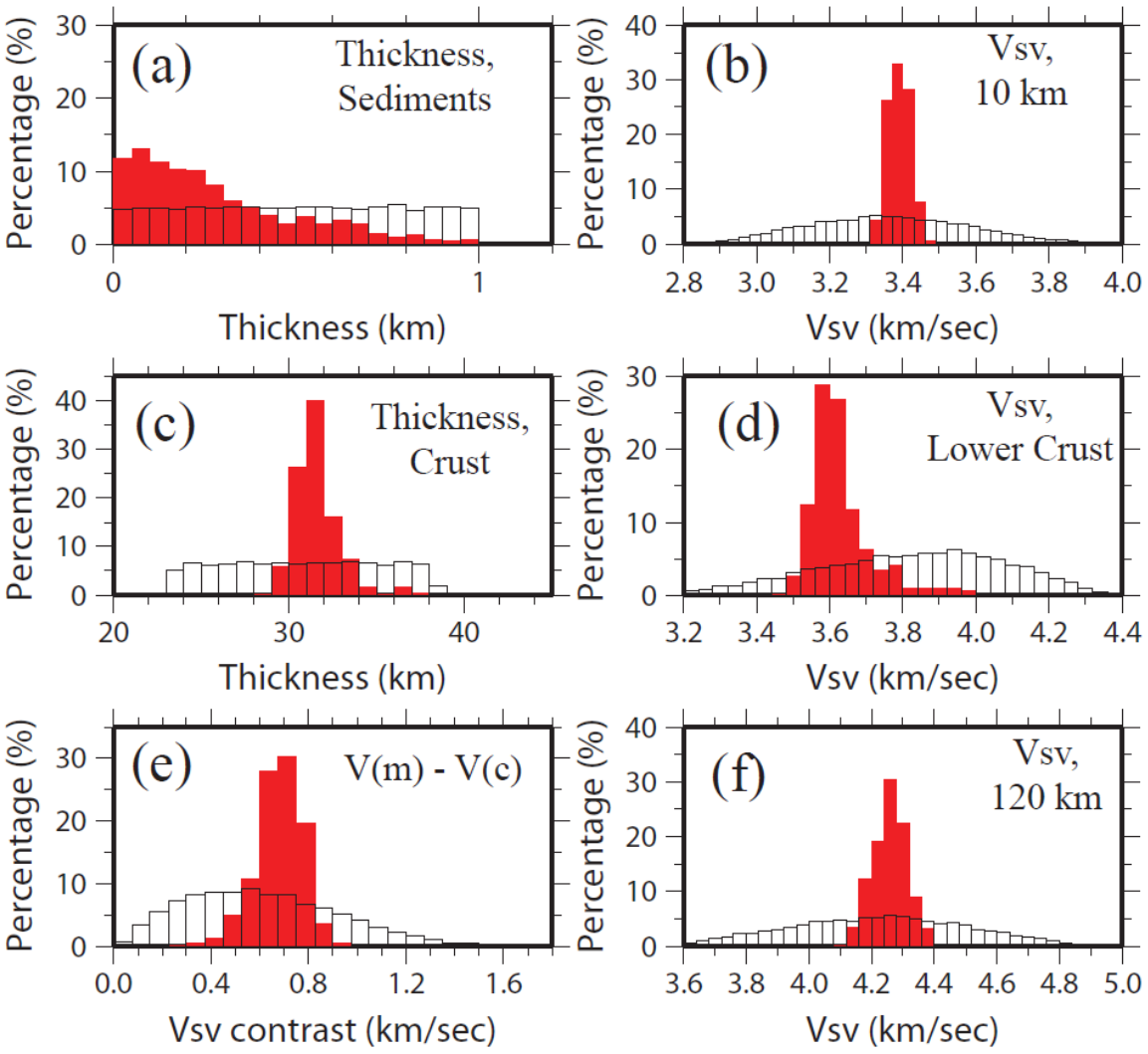

Figure 2.19 (a-f) The same as Fig. 2.16, but for the posterior distribution resulting from the joint inversion of surface wave phase velocities and receiver functions. Note the sharpening of distributions for parameters near the Moho in (c) – (e) compared to the distributions resulting from surface wave data alone (Fig. 2.16).

2.5.3 Examples of model ensembles

The model ensemble as well as the fit to the data at TA station R11A are shown in Figure 2.20. The 1σ width is less than half of the full width of the model ensemble, because the Vsv distribution at any depth for the joint inversion is approximately Gaussian except for the sedimentary layer (see Fig. 2.19). Figure 2.20c presents the azimuthally independent receiver function where the corridor outlined by black lines indicates the uncertainty range. The predicted receiver functions from the ensemble of accepted models are shown with gray lines and the

model that fits the receiver function best is shown with the red line. The most prominent signal is the peak at about 4 sec, which is fit quite well. However, the small trough near 2 sec is not fit because doing so would require introducing another crustal discontinuity, which would violate monotonicity. The fact that the algorithm does not automatically adapt the parameterization to accommodate other discontinuities is discussed further in section 8.2. Figure 2.20d identifies the model that fits the surface wave data best, but the model that fits the receiver function best misfits the surface wave data somewhat between about 40 and 60 sec period. The model that minimizes the joint misfit strikes a balance between these models.

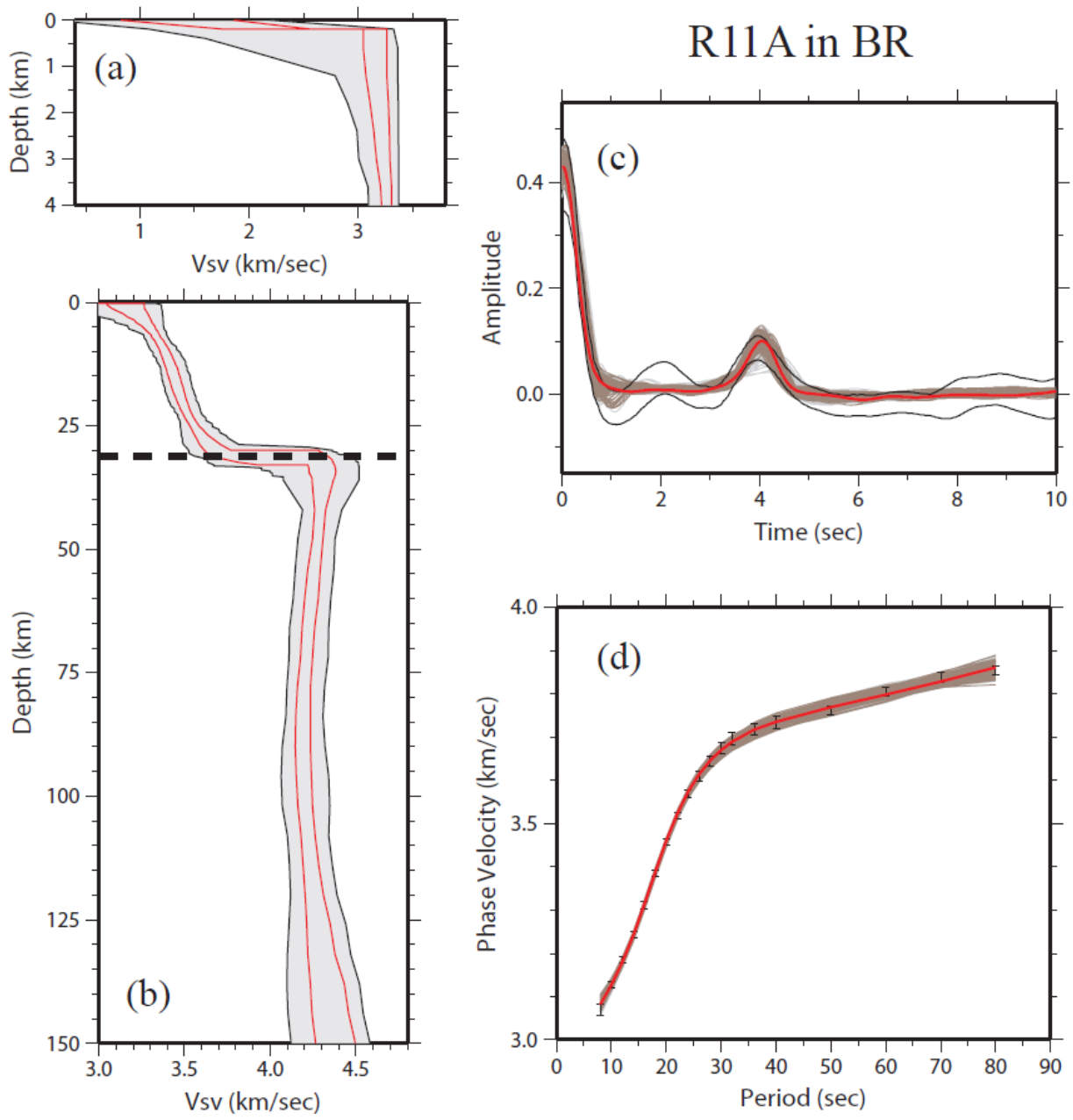

Figure 2.20 (a-b): The same as Fig. 2.17, but for results from the joint inversion of Rayleigh wave phase velocities and receiver functions at the same TA station (R11A). (c) The synthetic receiver functions from the accepted model ensemble are plotted with gray lines, with the best fitting receiver function shown as the red curve. The parallel black lines are the estimated uncertainty of the receiver function. There is a clear P-to-S conversion near 4 sec period, necessitating a large velocity jump at a shallow Moho. (d) The predicted surface wave dispersion curves from all accepted models are plotted with gray lines with the best fitting curve identified in red.

Figure 2.21 presents another example of the joint inversion for station T18A in the Colorado Plateau. In the observed receiver function, no clear peak is found from 3 sec to 8 sec where the Moho P-to-S converted phase would be seen in "normal" receiver functions. After inversion, the receiver function data is well fit without a P-to-S converted arrival based on the model ensemble shown in Figure 2.21b. The Moho discontinuity is not well defined in the ensemble of models and Vsv changes smoothly from the lower crust to the uppermost mantle.

Figure 2.22 summarizes the joint inversion result for station Q22A in the Colorado Rocky Mountains. In the observed receiver function, a peak at about 6 sec suggests that the Moho is deeper than at station R11A in the Basin and Range province. As a result, the crustal thickness distribution is centered at about 49 km depth with an uncertainty of 2.3 km. The relatively small uncertainty of crustal thickness generates a sharp transition between the lower crust and uppermost mantle, particularly compared with station T18A in the Colorado Plateau.

Figure 2.23 shows the joint inversion result for station O25A in a sedimentary basin, the Denver Basin region of the Great Plains. The peak at 1 sec on the receiver function and the trough at 4 sec indicate a thick sedimentary layer at shallow depths. After inversion, a thick sedimentary layer (~4 km) with a strong vertical velocity gradient is found and the crustal thickness is estimated to be 51.1 ±3.9 km. Signatures in receiver functions that are common for sedimentary basins are a broadened direct P-pulse or an apparent lack of a direct P-arrival at zero delay time followed by a large amplitude arrival in the first second or so – the latter feature being due to bending to vertical incidence and a strong conversion to shear energy at the sediment-basement contact. The high-amplitude apparent mid-crustal negative arrival is modeled here as a reverberation within the sedimentary basin with two shear and one compressional legs within the sediment layer. The constraint of monotonically increasing velocities within the crust aids suppression of sediment multiples in favor of imaging true deeper crustal structure. In a few other locations in which there are true mid-crustal low-velocity layers, such as the Rocky Mountain location shown in Figure 2.22, this may suppress a true feature, however. Extreme examples may be magma chambers (e.g., Wilson, 2003). Vsv in the sediments at O25A increases from less than 1 km/sec to more than 3 km/sec at depth of ~ 3km, representing the compaction of sediments in this layer. Vsv in the upper crust is ~ 3.4km/sec and reaches higher than 4 km/sec in

the lower crust, perhaps indicating a mafic lower underplated crust ("7.x layer") that has been proposed for parts of the High Plains (e.g. Gorman et al., 2002). A fast lithosphere (>4.5 km/sec) is observed in the upper mantle.

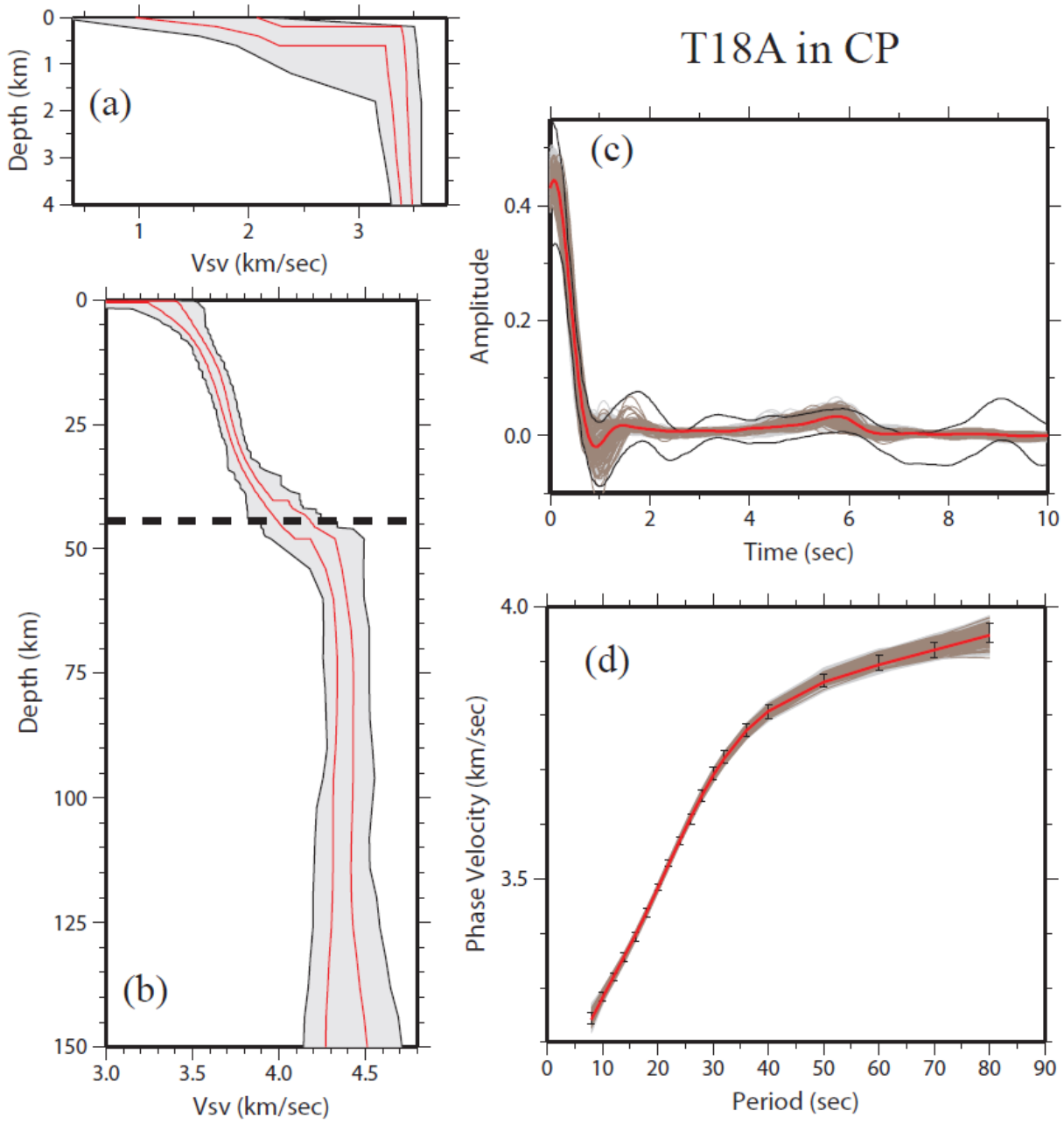

Figure 2.21 The same as Fig. 2.20, but for the result at TA station T18A located in the Colorado Plateau. No clear P-to-S Moho conversion is seen on the receiver function, probably implying a gradient Moho.

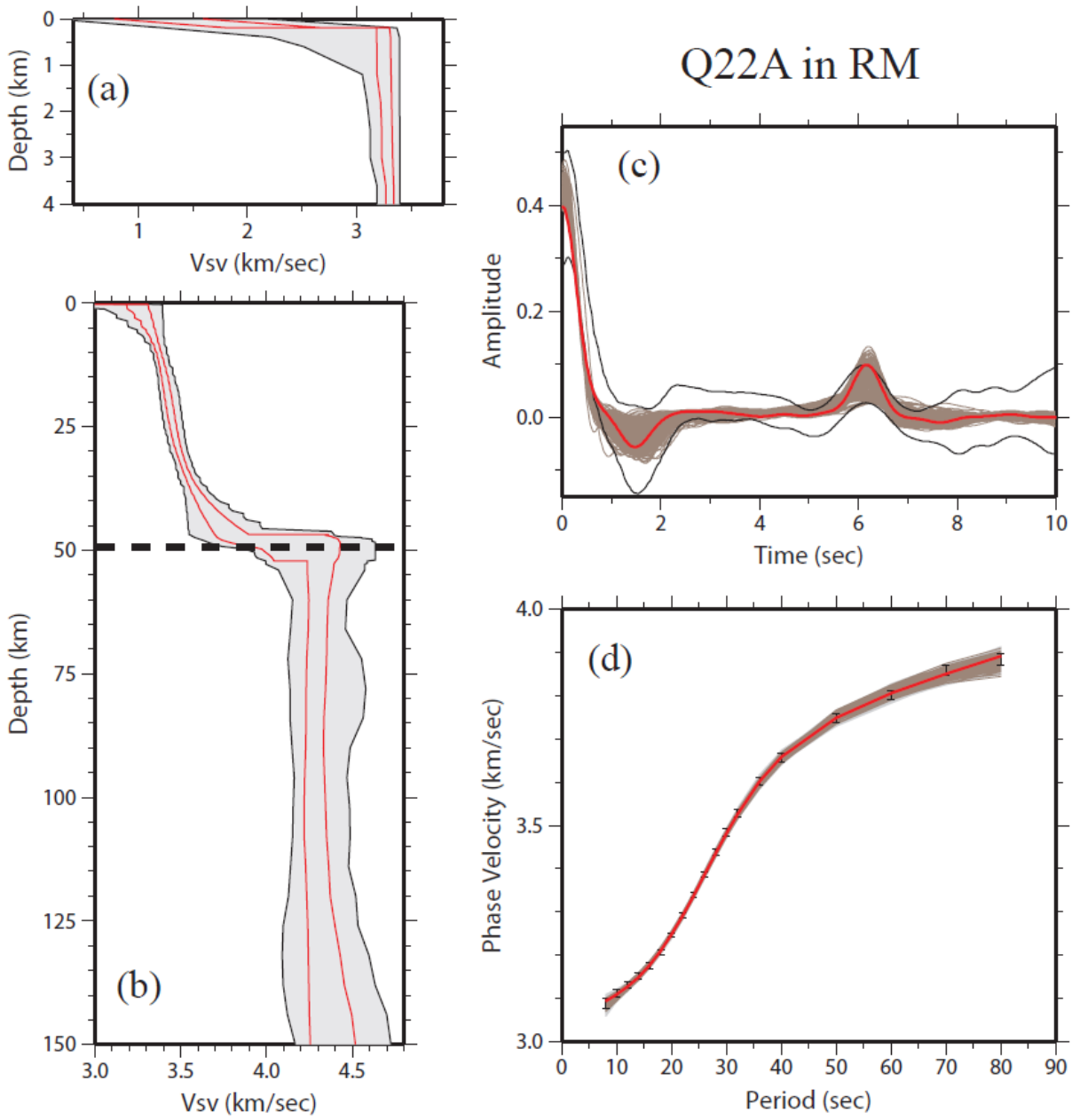

Figure 2.22 The same as Fig. 2.20, but for the result at TA station Q22A located in the Colorado Rocky Mountains. A clear P-to-S conversion is seen near 6 sec, requiring a large velocity jump at a deeper Moho than in Fig. 2.20.

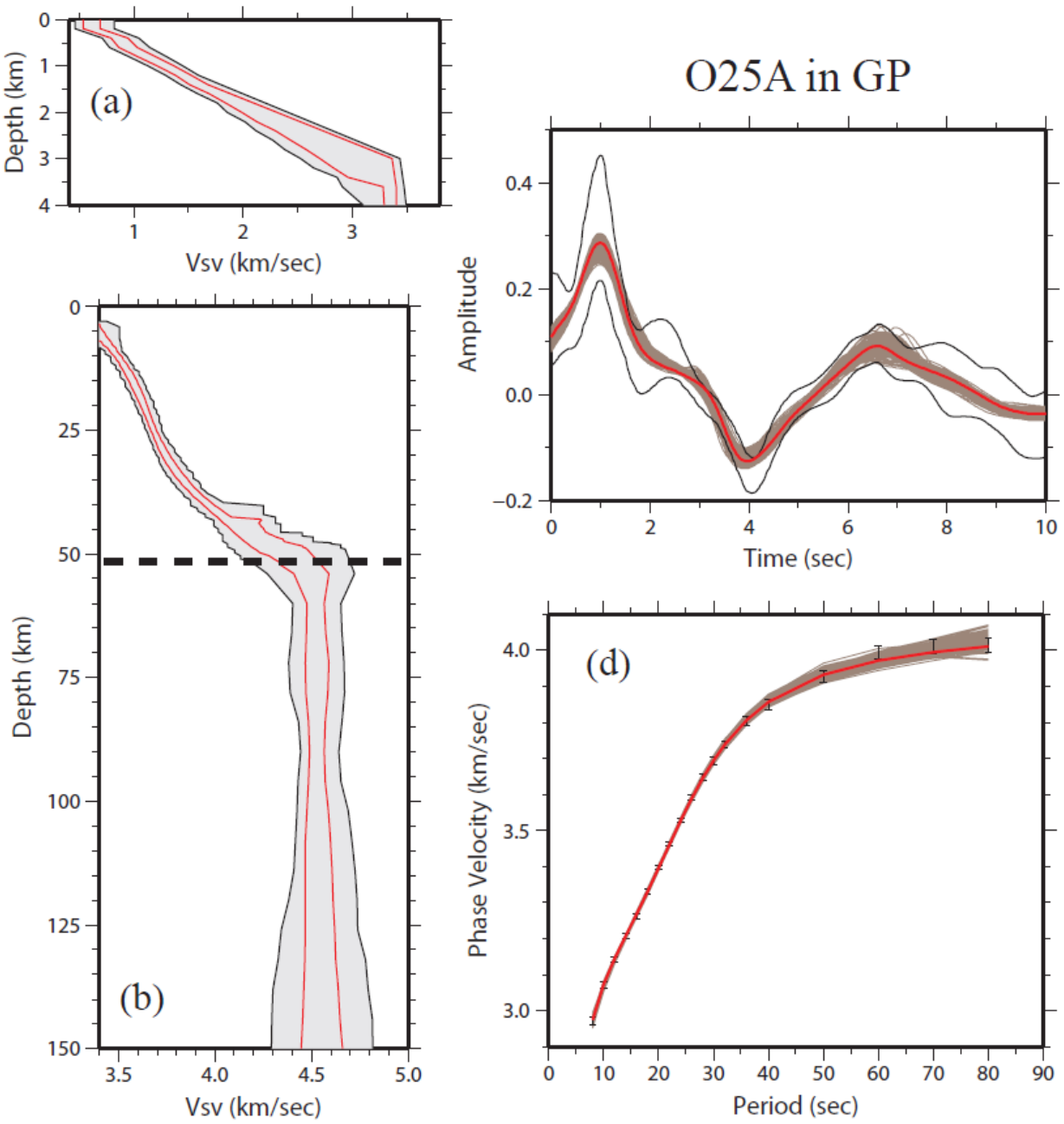

Figure 2.23 The same as Fig. 2.20, but for the result at TA station O25A located in the Great Plains. Strong sedimentary reverberations dominate the receiver function adding uncertainty in the location of and velocity jump at the Moho.

2.6 Systematic Application of the Joint Inversion Across the Intermountain West

In previous sections, the joint Bayesian Monte Carlo inversion of surface wave dispersion and receiver functions was shown to estimate reasonable 1-D models with attendant uncertainty information for stations situated in a wide variety of structural environments. We now report on the application of the joint inversion method to the 185 TA stations across the intermountain west (Fig. 2.1).

2.6.1 Construction of a 3-D Vsv model for the Intermountain West

We apply the Monte Carlo inversion to the 185 stations shown in Figure 2.1. For each joint inversion, we identify the model from the resulting model ensemble that fits the surface wave data (SW) or receiver functions (RF) best as the SW or RF best fitting model, respectively. The misfits to SW data by the SW best fitting models for the stations tested are plotted with the white histogram in Figure 2.24a, while the misfits to the RFs by the RF best fitting models are plotted in Figure 2.24b. For each station, the model with minimum joint misfit is identified, and is called the joint best fitting model. Misfits to SW and RF data from the joint best fitting models are shown in Figure 2.24 with red histograms. At most stations, the joint best fitting models fit the SW or RF data only slightly worse than the model that fits each single data type best. This indicates that there is some tension in fitting the two data types. Second, the joint best fitting models have a misfit < 1 for both data sets for almost all stations, which indicates that the simple model parameterization that we use in the joint inversion can reproduce the RFs with misfits below the RF uncertainties for most of the stations tested. Third, larger RF misfits appear for stations near the corner of Utah/Colorado/Wyoming and are sparsely distributed through the Basin and Range. The RFs at those stations have sharp back-azimuthal variations that vitiate the harmonic stripping method's attempt to estimate an azimuthally independent receiver function. Cases with sharp back-azimuthal signals in the RFs are discussed further in the Appendix, but in these cases the uncertainty in the receiver function increases substantially and the inversion reverts for the most part to fitting surface wave data. Fewer than three stations show a large misfit (>0.9) to surface wave data and these stations are sparsely distributed across the map. Finally, the comparable minimum and joint misfits for the two data sets also indicate that neither

data set is over-weighted during the inversion.

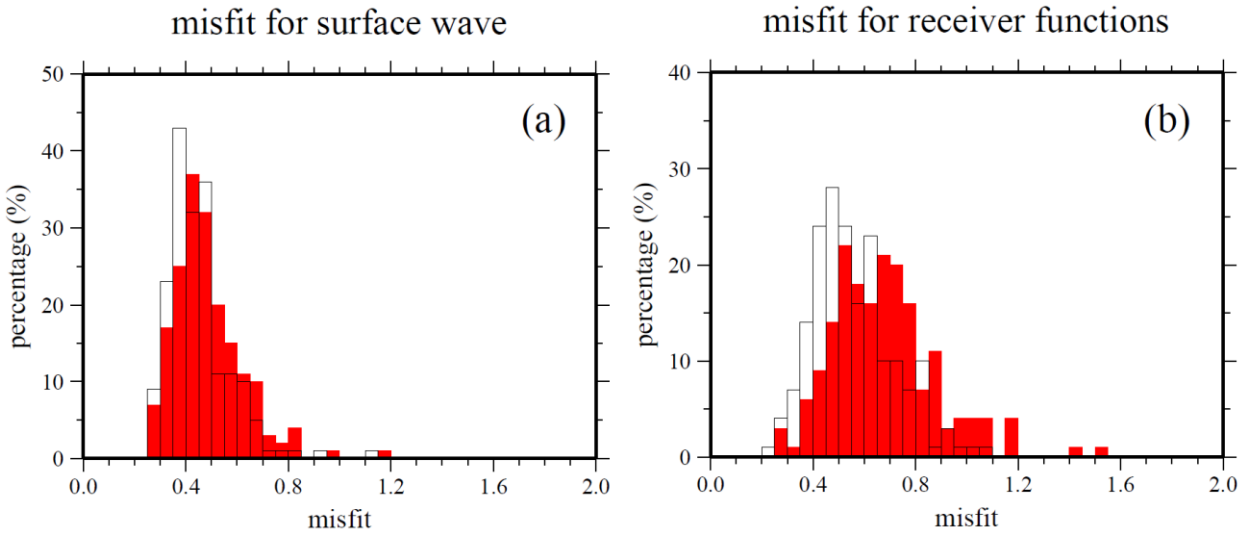

Figure 2.24 (a) Misfit histogram showing the misfits to surface wave data over the spatial grid covering the area of study. The white histogram is for the model that best fits the surface wave data at each point and the red histogram is for the model that jointly best fits both surface wave and receiver function data. At most grid points, the jointly best fitting models fit the surface wave data only slightly worse than the model that fits the surface wave data best. (b) Same as (a), but for receiver functions. The white histogram is for the model that best fits the receiver function data at each location and the red histogram is for the model that jointly best fits both data sets. At most grid points, the jointly best fitting models fit the receiver function data only slightly worse than the model that fits the receiver function data best.

After inversion is performed at all stations, we obtain Vsv model ensembles beneath the 185 stations with means and uncertainties at all depths. The distribution of stations forms an irregular grid. To produce a smooth Vsv model on a regular grid, simple kriging (Schultz et al., 1999) interpolation is applied to smooth the Vsv values at each depth based on the mean value at that depth and the estimated uncertainties. For a given depth, at each grid node we search for stations within a 1-degree radius. We weight the average Vsv of the model ensemble for at given station using a weighting function defined as follows:

$$W_i = \begin{cases} \dfrac{1}{(1+d_i)\sigma_i} & \text{if } d_i \le 1^{\circ} \\ 0 & \text{otherwise} \end{cases} \qquad (11)$$

where d_i and σ_i denote the distance to station i from the grid node and the uncertainty of the model at the specified depth for that station, respectively. By performing this simple-kriging interpolation, map views of the 3-D model on a 0.25 °x0.25 °spatial grid are produced.

2.6.2 Characteristics of the 3-D model

In the upper crust (10 km depth, Figure 2.25a), high velocities are observed beneath the Colorado Plateau and Great Plains, while the Rocky Mountains show low Vsv. Near the northern boundary of the Colorado Plateau low velocities are also observed, presumably caused by very thick sediments in the Green River Basin, Uinta Basin, etc. At these locations, the maximum allowed sediment thickness may be less than the true thickness, leading to smearing of low velocities into the upper portion of the crystalline crustal layer. In contrast, the Denver Basin is much thinner than 10 km and the structure there is well constrained by receiver function data. Uncertainties at 10 km depth are approximately homogeneous across the region, averaging about 27 m/s, which is about 0.8%.

In the lower crust (averaged from 4 km above the Moho to the Moho, Fig. 2.25c), the most prominent feature is the slow anomaly (<3.6 km/sec) encompassing the eastern Basin and Range province near the northwestern Colorado Plateau. Slow lower crust is found across the entire Basin and Range and also near the western edge of the Rocky Mountain province in southwestern Colorado. There is coincidence between the ~3.85 km/sec Vsv contour and the eastern Rocky Mountain topographic high in Colorado, implying a strong relationship between topography and crustal structure. A fast lower crustal anomaly extends through the Colorado Great Plains and southern Wyoming and also penetrates into the Colorado Plateau. Uncertainties are larger and more variable in the lower crust (Fig. 2.25d) than the upper crust, ranging from about 40 m/s in parts of the Colorado Plateau to more than 160 m/s in the Rocky Mountains of southern Colorado. Higher uncertainties have two causes. First, they appear where there is a large jump in velocities across the Moho, due to a trade-off between Moho depth and uppermost mantle structure. The trade-off has been ameliorated but not entirely eliminated through the addition of receiver functions in the inversion. Second, larger uncertainties also occur where receiver functions have a larger uncertainty in the amplitude of P-to-S conversions.

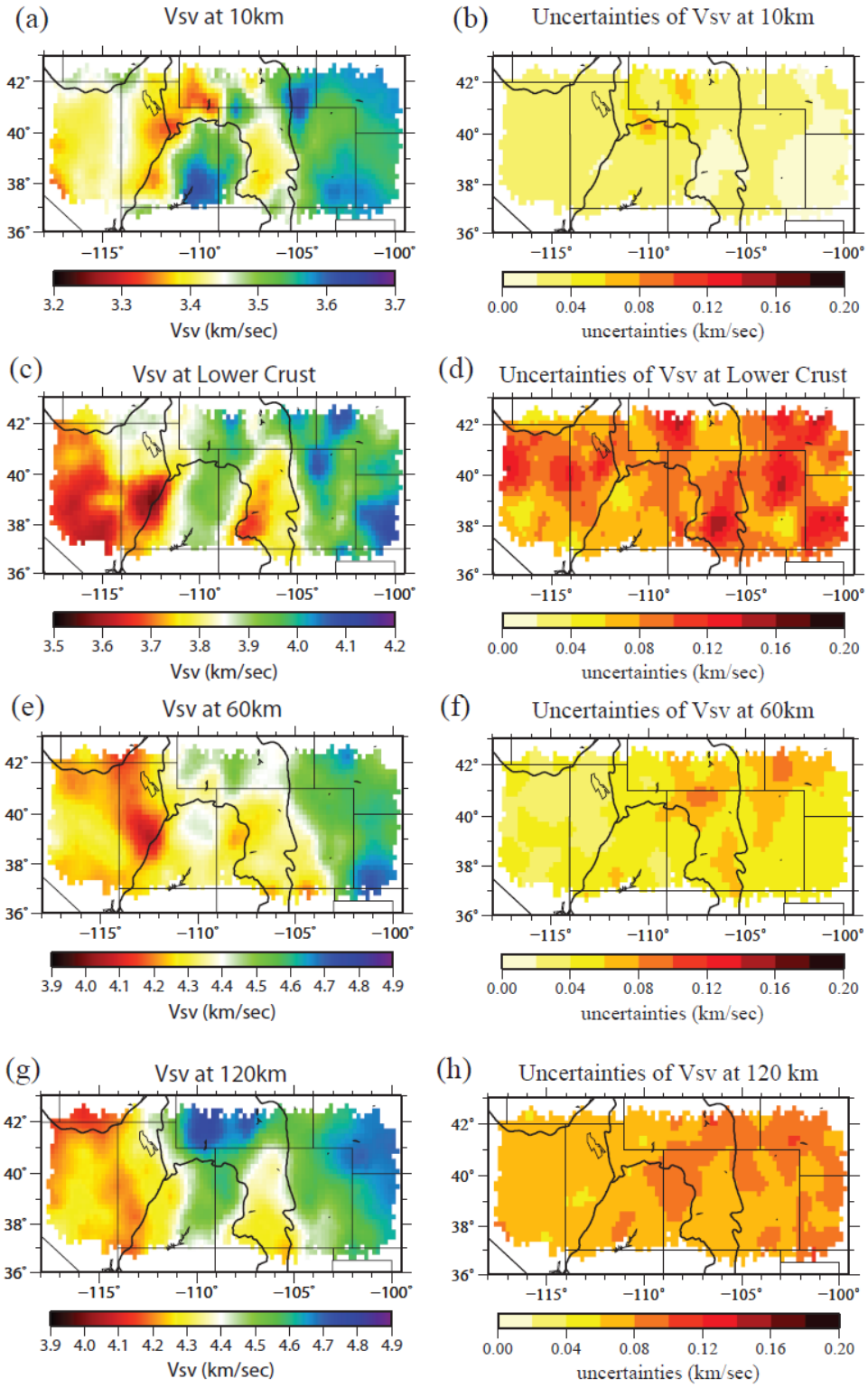

Figure 2.25 Map views of Vsv at different depths (left) with uncertainties (right). (a-b) Vsv at 10 km depth. (c-d) Vsv in the lower crust, 4 km above Moho. (e-f) Vsv at 60 km depth. (g-h) Vsv at 120 km depth.

At 60 km depth (Fig. 2.25e), which is in the uppermost mantle across the region, a strong contrast is observed between the eastern and western sides of the study region. A strong anomaly as low as 4.0-4.1 km/s is seen at the eastern edge of the Basin and Range, which may be related to the Cenozoic magmatism in this region (Roy et al., 2009). Much higher velocities are observed beneath the Great Plains, reaching up to 4.8 km/s. Uncertainties are largely homogeneous across the study region, averaging about 50 m/s (Fig. 2.25f). At a depth of 120 km (Fig. 2.25g), Vsv beneath the Great Plains is very high, but the strongest high velocity anomaly is detected beneath the Wyoming Craton in southern Wyoming. This high velocity feature continues beneath the Colorado Plateau, suggesting a strong, thick lithosphere beneath the northern Colorado Plateau except for its northwestern periphery. The Basin and Range and Rocky Mountains show relatively homogeneous low Vsv compared with other regions. Uncertainties (Fig. 2.25h) are fairly homogeneous and average about 65 m/s (< 1.5%) across the entire study region. Uncertainties are larger deeper in the mantle (120 km versus 60 km) because the surface wave dispersion information is less sensitive to deeper structure.

2.6.3 Changes in the 3D model compared with the surface wave inversion

There are several significant advantages to adding receiver functions to surface wave dispersion data in the Monte Carlo inversion. We focus on three topics: (1) determining crustal thickness and uncertainties, (2) determining the velocity jump across the Moho, and (3) constraining uppermost mantle structure.

By utilizing receiver functions in the inversion, there is a natural increase in the accuracy of estimates of Moho depth or crustal thickness. This is apparent on comparison between the crustal thickness distribution for station R11A from the inversion with surface wave data alone and that from the joint inversion of both data sets (Fig. 2.16c and Fig. 2.19c). This observation holds for all stations with clear P-to-S converted signals. Figure 2.26a,b presents crustal thickness and associated uncertainties from inverting surface wave dispersion data alone.

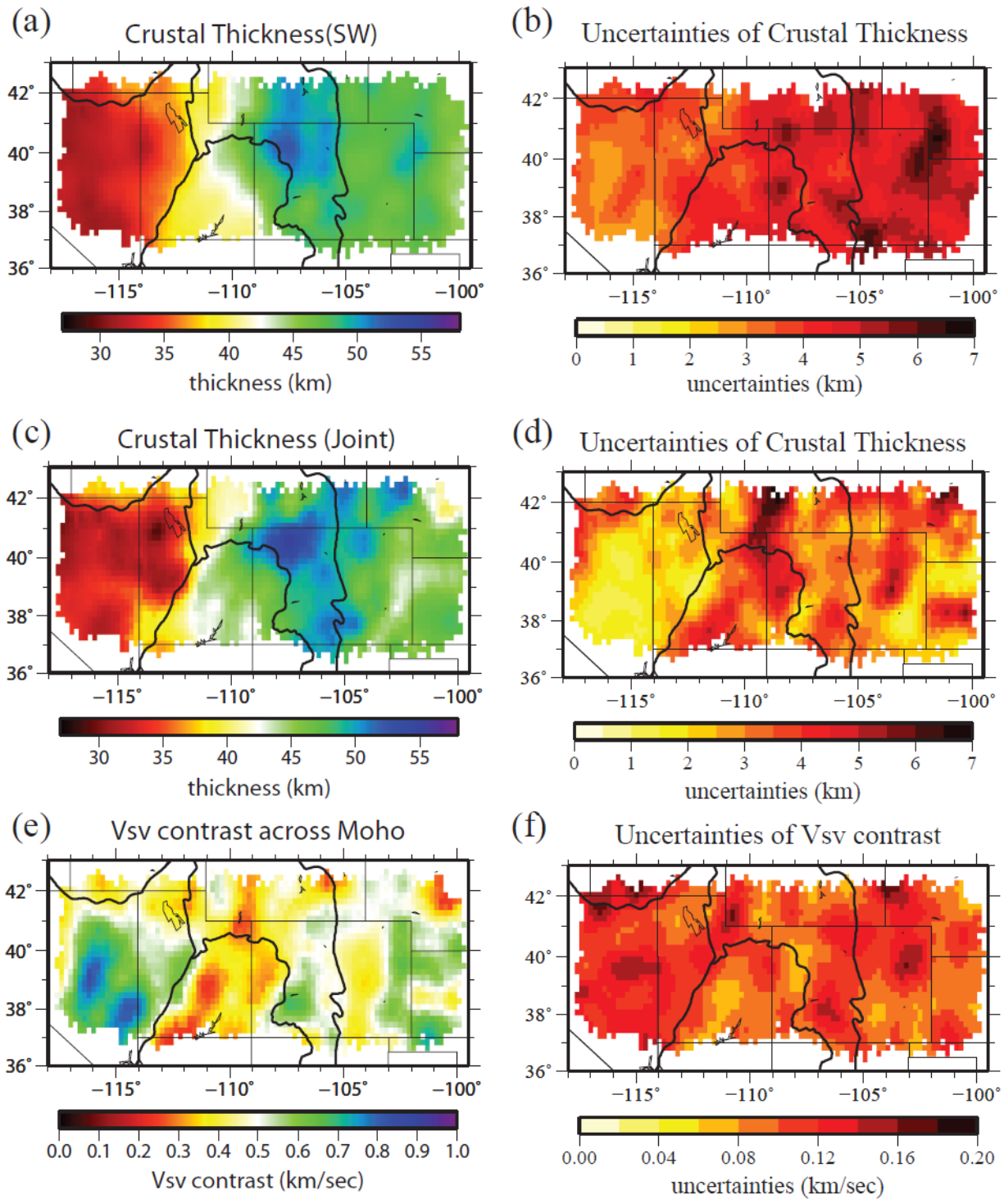

Figure 2.26 (a)-(b) Crustal thickness and its uncertainty resulting from inversion of surface wave data alone. (c)-(d) Crustal thickness and its uncertainty resulting from the joint inversion of surface wave and receiver function data. (e)-(f) Vsv contrast from 4 km below to 4 km above Moho and its uncertainty from the joint inversion of surface wave and receiver function data.

The uncertainty level scales with crustal thickness and is ~ 5 km on average with smaller values in the Basin and Range and larger values in the Great Plains. In contrast, crustal thickness and uncertainty determined from the joint inversion appears in Figure 2.26c,d. The uncertainty level from the joint inversion decreases in the Basin and Range, the central Rocky Mountains and parts of the Great Plains, where P-to-S converted phases are well observed. In the northern part of the Colorado Plateau and southern Wyoming craton where thick sediments exist, crustal thickness uncertainties are not reduced by adding receiver functions or may even increase where the P-to-S signals in the receiver functions are muted by sedimentary reverberations. In addition to improving the determination of Moho depth, receiver functions also help to determine whether a sharp Moho discontinuity exists or not. An example is shown in the inversion at station T18A in the Colorado Plateau (Fig. 2.21), where the raw receiver function has no dominant arrival from 3 sec to 7 sec where a P-to-S conversion should appear. The joint Monte Carlo inversion thus produces a model that has a gradient in Vsv at the depth where Moho is expected. The resulting crustal thickness distribution has larger uncertainty than when a clear P-to-S phase is observed, but none of the models in the distribution shows a sharp Moho.

Second, introducing receiver functions improves the determination of the Vsv contrast across the Moho, which is related to the amplitude of the P-to-S phase in the receiver functions. For instance, the receiver function at station R11A (Fig. 2.20) shows a strong Moho conversion while the converted phase in the receiver function at T18A is weak, resulting in a Vsv contrast at Moho that is stronger at R11A than at T18A (or perhaps a gradational crust-mantle transition at station T18A). A map of the Vsv contrast and uncertainty across the region is shown in Figure 2.26e,f. The features shown are coherent with geological province. A high Vsv contrast across Moho is observed beneath the Basin and Range as well as parts of the Rocky Mountains and the Great Plains east of the Denver Basin. In contrast, beneath the Colorado Plateau, especially under the northwestern Colorado Plateau, the Vsv contrast is very low. This feature is consistent with observations made by earlier studies (Sheehan et al., 1997, Gilbert and Sheehan, 2004, Levander et al., 2011, Bailey et al., 2012). Beneath the southern Wyoming Craton, we observe a small Moho Vsv contrast as well as high uncertainties in crustal thickness and thick sediments. This is because sedimentary reverberations dominate the receiver function where a Moho peak would be expected and the Vsv constrast across Moho is difficult to resolve. Uncertainties in the

velocity contrast across Moho range between about 80 m/s to 160 m/s, being smallest in the Colorado Plateau where the Vsv contrast is low in all accepted models.

The third advantage of the joint inversion method is a better determination of mantle structure below the Moho discontinuity. This is largely due to reduction of the trade-off between Moho depth and lower crustal velocity. Because of this trade-off, when surface wave dispersion is used alone to invert for a 3-D model, a prior constraint is often applied on the vertical velocity gradient in the uppermost mantle. For example, several studies have set the Vsv gradient in the uppermost mantle to be positive (Yang et al., 2008, Moschetti et al., 2010, Zheng et al., 2011). Figure 2.27 shows the result of various upper mantle constraints along the transect identified in Figure 2.1. In the model constructed by using surface wave data alone with an imposed positive Vsv gradient in the uppermost mantle (Fig. 2.27a), a slow anomaly belt is found immediately beneath the Moho. In contrast, if the constraint is changed in sign so that a negative Vsv gradient is imposed, a fast anomaly belt is observed (Fig. 2.27b). These two models show differences down to depths of more than 100 km and the placement of anomalies in depth is affected strongly. Lithospheric thickness and the depth to prominent asthenospheric anomalies are both changed by varying this constraint. However, both models fit the surface wave data equally well, which indicates that the surface wave data alone cannot distinguish between them. Figure 2.27c shows the model constructed from the joint inversion of surface wave dispersion and receiver function data with no constraint on the uppermost mantle velocity gradient. The result looks like a combination of Figures 2.27a and 2.27b. Beneath the Basin and Range where the Moho is prominent and the Vsv contrast across Moho is large, the model is more similar to Figure 2.27b constructed with a negative gradient constraint in the uppermost mantle. However, in places where a gradient Moho is expected, the model is more similar to models from the positive gradient constraint, Figure 2.27a. The assimilation of receiver functions in the inversion resolves the velocity-depth trade-off with data rather than with ad hoc prior constraints.

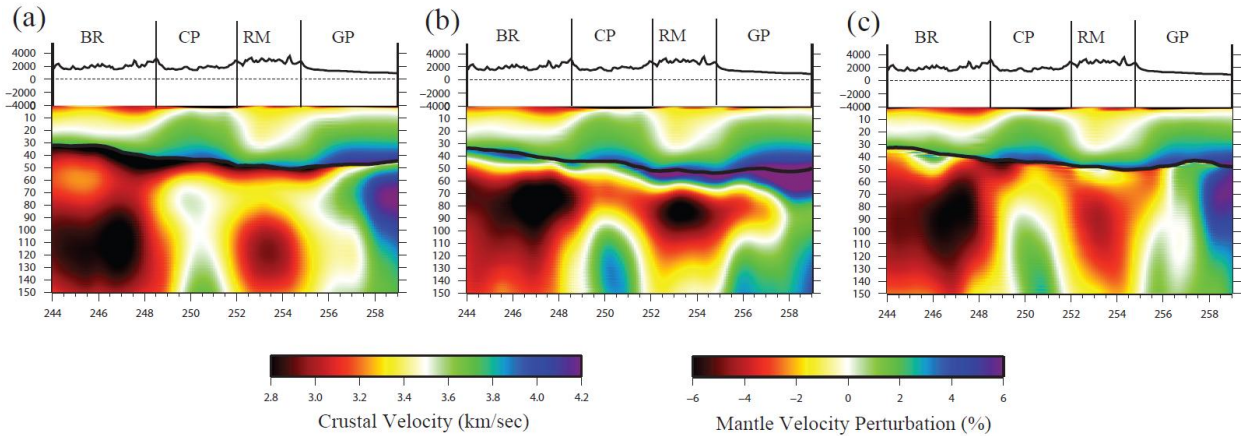

Figure 2.27 (a) Vsv model along transect AA' in Fig. 2.1 constructed using surface wave data alone with a positive gradient constraint in the uppermost mantle. (b) The same as (a), but with a negative constraint on the uppermost mantle velocity gradient. (c) Vsv model from the joint inversion of both surface wave and receiver function data with no constraint on the Vsv gradient in the uppermost mantle. Crustal structure is presented in absolute shear wave speed but mantle structure is presented as the percent perturbation relative to 4.4 km/s. Geological provinces are presented with abbreviations (Fig. 2.1) overlying surface topography.

2.7 Discussion

2.7.1 Sensitivity to the Vp/Vs ratio

Rayleigh wave phase velocity dispersion curves are only weakly senstive to Vp compared to Vs; hence, the Vp/Vs ratio is usually a fixed parameter in the inversion of surface wave data. However, the Vp/Vs ratio is important in receiver function analysis because it is needed to map the P-to-S conversion time delay to depth. The Vp/Vs ratio can be determined when Moho reverberations such as the PpPs or PsPs phases are used as discussed by Zhu and Kanamori (2000). However, these phases arrive later than the 10 sec time window used here in the Monte Carlo inversion. We omit these phases from consideration because at many locations in the western US Moho reverberations cannot be isolated cleanly due to lateral heterogeneity (e.g., the large scatter in the results of Lowry and Perez-Gussinye, 2011, Fig. 2a) and in some places a graditional crust-mantle transition. Rayleigh wave ellipticity and local amplification may also potentially be used to constrain the crustal Vp/Vs ratio (Lin et al. 2012a,b). However, this is beyond of the scope of this study.

Because we ignore crustal reverberation phases, we are unable to determine the three relevant parameters (Vs in the crust, crustal thickness, and Vp/Vs ratio) simultaneously. To demonstrate this result, we add an extra degree of freedom to the model space, the Vp/Vs ratio in the crystalline crust, which we allow to vary between 1.65 and 1.85 during the inversion. First, the posterior distribution of the Vp/Vs ratio as well as its prior distribution are plotted in Figure 2.28a. The posterior distribution of the Vp/Vs ratio does not center at any particular value. Second, once the Vp/Vs ratio is introduced as a variable in the inversion, the posterior distribution of the crustal thickness broadens (Fig. 2.28b) about 20% compared with the same posterior distribution from the inversion with fixed Vp/Vs ratio (Fig. 2.20c) from 1.4 km to ~ 1.7 km. Finally, we observe a strong trade-off between crustal thickness and the Vp/Vs ratio (Fig. 2.28c), which indicates that given a higher Vp/Vs ratio the estimated crustal thickness will be lower and vice versa.

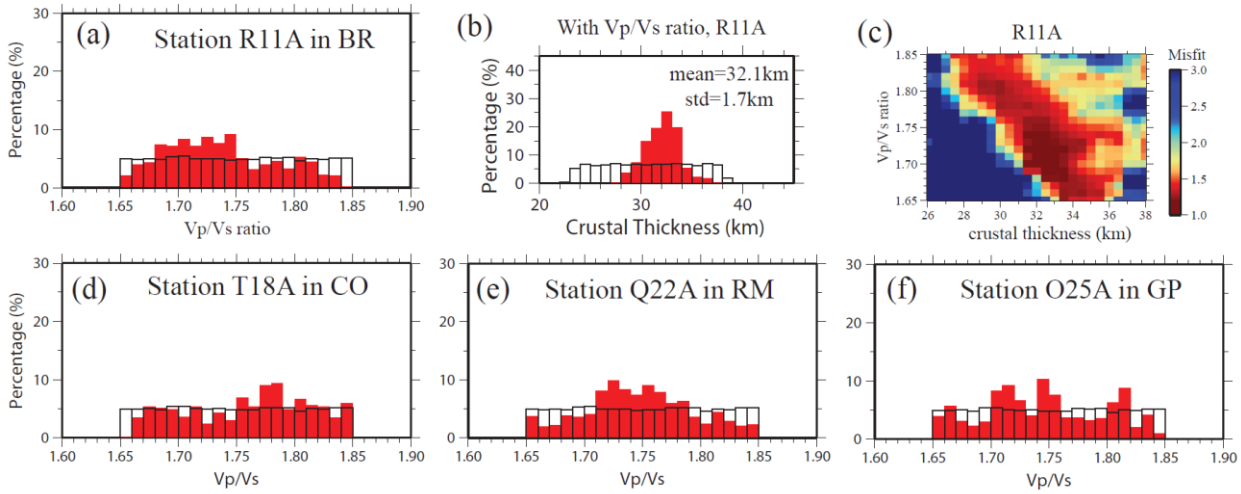

Figure 2.28 (a) Posterior distribution of Vp/Vs ratio when it is included as a parameter in the joint inversion. The prior distribution is shown with the white histogram outlined in black. (b) The posterior distribution of crustal thickness for station R11A from the inversion that includes Vp/Vs as a free parameter. (c) The joint misfit (χ^{joint}) is plotted as a function of crustal thickness and Vp/Vs ratio. Trade-off between crustal thickness and the Vp/Vs ratio is clear. (d)-(f) The same as (a), but for stations T18A, Q22A and O25A, respectively.

The fact that we are unable to determine a preferred Vp/Vs ratio from our data at station R11A also holds for other stations. Figure 2.28d-f shows the posterior distributions of Vp/Vs ratio at stations T18A in CP, Q22A in RM, and O25A in GP, respectively. Overall, we poorly constrain the Vp/Vs ratio from the data we are using. The direct effect is that when we present the 3-D

model in section 7.2, the uncertainty of crustal thickness is under-estimated by up to ~20% due to the fact that the Vp/Vs ratio is set to 1.75. To improve the determination of the Vp/Vs ratio, other information such as reverberation phases arriving after 10 sec in the receiver functions and Rayleigh wave ellipticity and local amplification (Lin et al., 2012a, b) would have to be included in the Monte Carlo sampling.

2.7.2 Limitations of the current method and potential refinements

The procedure that we have developed is intended to provide a practical method to invert surface wave dispersion and receiver functions jointly over large areas. As it currently exists, the method can be applied fruitfully across large arrays that have been developed around the globe. This includes all of the USArray in the US, as well as the Chinese Earthquake Array, the Virtual European Seismic Network, various PASSCAL experiments around the world, and F-net in Japan. However, the method also serves as a framework for future enhancements and improvements. In this regard we highlight four known limitations with the method as it is currently effected that may call for modification. A fifth limitation is discussed in section 8.1 regarding a variable Vp/Vs ratio.

First, the parameterization that we have defined is applied rigidly in the inversion. The algorithm has not been designed to sense misfit to the data and to adapt the parameterization accordingly, although other recent joint inversion algorithms include this feature (e.g., Bodin et al., 2011). For instance, in the inversion at station R11A (Fig. 2.20), a mid-crustal discontinuity was not introduced automatically to fit the negative arrival at ~1 sec on the receiver function. An adaptive parameterization would help the algorithm fit aspects of the data that are currently being ignored, but introducing more structure would increase uncertainties and may lead to over-interpretation of the data.

Second, the travel time variation of P-to-S conversions due to dipping interfaces (notably on the Moho) is not fit in our inversion and is removed in the harmonic stripping algorithm to estimate the azimuthally independent receiver function that we use in the joint inversion, $A_0(t)$. Not modeling the delay time variation explicitly means that the amplitude of P-to-S conversions will be underestimated if the variation is significant (i.e., ≥ 0.5 sec). The delay time variation does appear in the $A_1(t)$ and $A_2(t)$ components of the receiver function that are estimated in the

harmonic stripping algorithm but which are not used in the inversion. Optimally, the algorithm would employ information that exists in these two components to correct for the underestimation of the amplitude of the phase conversions from dipping interfaces. In principle, this information as well as transverse component receiver function amplitudes and delay time variations could also be used to estimate the dip on the interfaces, but this information would probably have to be interpreted independently from surface wave data.

Third, we do not use the receiver function (RF) produced from a multi-station imaging technique (e.g., the Common Conversion Point (CCP) stacked receiver function) but rather a single-station RF which is an average of the 3-D variation near each single station. In the future, this procedure could be replaced with the receiver function stacked at the common conversion point (CCP) rather than at the station. This procedure is inappropriate for the TA because the CCPs beneath nearby stations do not overlap, as Figure 2.29a illustrates. The procedure would be appropriate for a denser array such as PASSAL or EarthScope Flexible Array experiments. Figure 2.29b also presents a comparison between RFs computed using the TA and the CD-ROM PASSCAL experiment. The blue RFs are from three TA stations along the green transect in Figure 2.29a, and the red RFs are binned CD-ROM+TA RFs produced according to the Moho Conversion Points (MCP), where radial receiver functions that pierce the same bin at Moho depth are averaged laterally. The similarity between MCP RFs and the harmonic stripping $A_0(t)$ RFs demonstrate that the single station azimuthally indepenent RFs that we construct are consistent with the multi-station RFs. A smoothed image of the multi-station RFs is presented in Figure 2.29c, demonstrating the resolution that RFs from a tighter network geometry provide. Higher resolution features such as the mid-crustal signal at ~ 2.5 sec beneath the Cheyenne Belt (CB) are not captured by the sparse TA array. However, using higher resolution RFs would necessitate accommodating the lower resolution of surface wave dispersion maps so as not to alias sharp features from the RFs into larger scale features derived in the joint inversion. It remains unclear what maximum resolution may be possible for surface waves from a tighter array spacing than the TA, for example from the Flexible Array component of USArray. It is likely, however, that surface wave resolution will always lag resolution from receiver functions. Although the installation of the TA has ameliorated the differential resolution problem in the US, the issue may ultimately need to be resolved by low-pass filtering the RFs.

Fourth, the full error covariance matrices for surface wave dispersion data and receiver functions have not been utilized here. Rather, we have assumed that the matrices are diagonal (errors are independent) and have balanced the two data sets by introducing a scaling parameter (κ) that effectively normalizes the misfit found for the two data sets. Estimating the inverse covariance matrix for both data sets is not trivial, but it would improve the effectiveness of the algorithm.

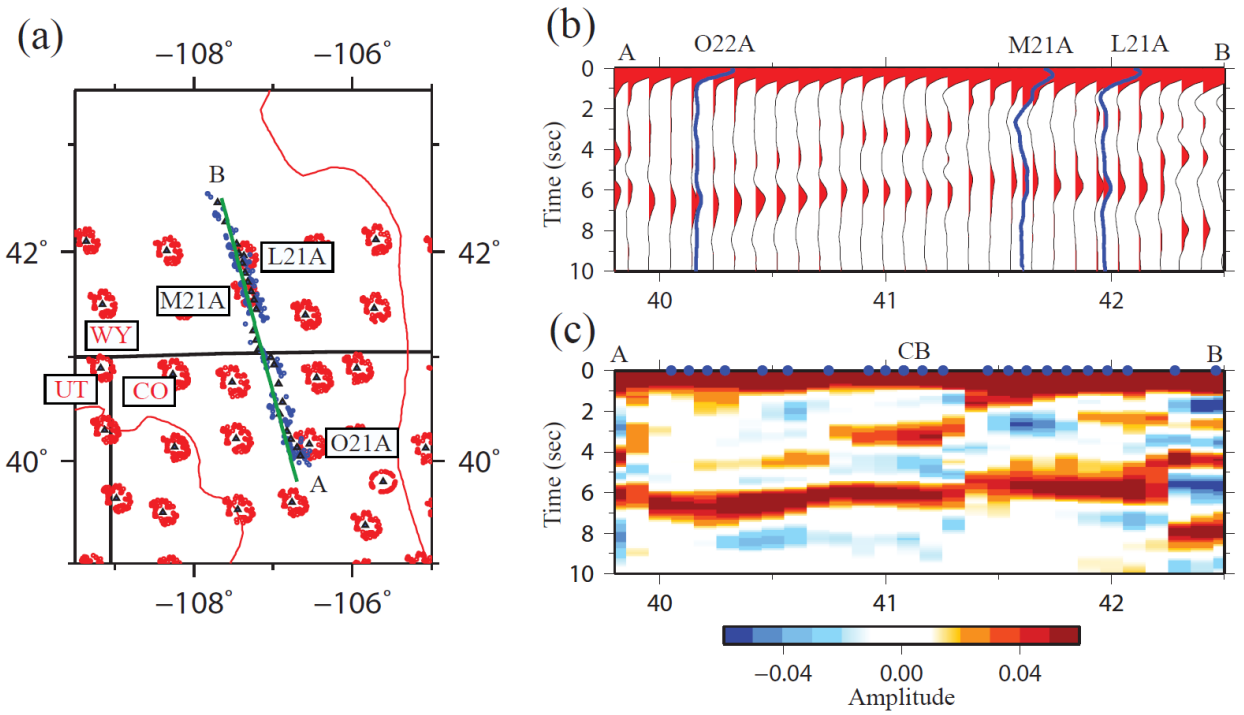

Figure 2.29 (a) Piercing points of the P-wave incident at the Moho for TA stations are shown with red dots. The piercing points for CD-ROM stations are shown with blue dots. The green line is the transect AB along the CD-ROM line from Colorado to Wyoming. Three TA stations near the transect are identified with names (L21A, M21A and O21A from north to south). State boundaries are outlined with black lines and red lines are geological provinces (Fig. 2.1). (b) Moho conversion point (MCP) stacked receiver functions are illustrated with red waveforms along transect AB in (a). For comparison, single-station processed receiver function s ($A_0(t)$) are shown with blue waveform for the three stations identified in (a). (c) The smoothed image of the red receiver functions in (b). Blue dots indicate the location of the CD-ROM stations. The location of the Cheyenne Belt is marked as CB.

2.8 Conclusions of this Chapter

We present a new method for joint inversion of surface wave dispersion data and receiver functions based on a Bayesian Monte Carlo scheme. When applied to receiver functions and surface wave data from ambient noise and earthquakes that are now emerging from extended

broadband seismic arrays, the method produces a 3-D model of the crust and uppermost mantle to a depth of about 150 km with associated uncertainties. The method is designed to be used in an automated fashion across a large number of stations and has been applied to data from 185 USArray Transportable Array (TA) stations in a geologically diverse part of the Intermountain West. The effect of the introduction of receiver functions to surface wave dispersion data is visualized through improvements in the posterior marginal distribution of model variables. By comparing the statistics of the posterior distributions, we find that adding receiver function data quantitatively improves the accuracy of estimates of Moho depth, improves the determination of the Vsv contrast across Moho, and improves uppermost mantle structure. Knowledge of uppermost mantle structure is improved because the assimilation of receiver function data makes it possible to relax ad-hoc structural constraints that are commonly invoked in inversions based on surface wave data alone.

Although the inversion method we describe can be applied robustly across large regions, there remain aspects of the method where refinements may prove beneficial. Three are particularly noteworthy: (1) development of an adaptive parameterization, particularly in the crust where low velocity layers may exist, (2) further investigation of the estimation of the full data covariance matrix and its inverse for both receiver functions (covariance over time) as was done by Bodin et al. (2011) and surface wave dispersion (covariance over frequency), (3) extension of the receiver functions past 10 sec to recover reverberations that may help to constrain the Vp/Vs ratio. These and other potential refinements to the method may reduce small biases that derive from current assumptions, but the current method produces results that are preferable to the use of surface wave dispersion or receiver function data alone.

Acknowledgments. Major part of this chapter is presented in a paper by Shen et al., (2013a). The co-authors include Michael H. Ritzwoller, Vera Schulte-Pelkum and Fan-chi Lin. the authors gratefully acknowledge insightful reviews from Thomas Bodin, Malcolm Sambridge, an anonymous reviewer, and the Associate Editor, Gabi Laske, that helped to improve this chapter. They are also grateful to Craig Jones for insights into receiver function analyses and to Anne Sheehan for discussions concerning the history of joint inversions with receiver functions and surface wave dispersion. The facilities of the IRIS Data Management System, and specifically the IRIS Data Management Center, were used to access the waveform and metadata required in

this study. The IRIS DMS is funded through the National Science Foundation and specifically the GEO Directorate through the Instrumentation and Facilities Program of the National Science Foundation under Cooperative Agreement EAR-0552316. This research was supported by NSF grants EAR-0711526, EAR-0844097, EAR-0750035, and EAR-1053291 at the University of Colorado at Boulder. F.-C. Lin is supported by the Director's Post-Doctoral Fellowship of the Seismological Laboratory at the California Institute of Technology.

CHAPTER III

A 3-D MODEL OF THE CRUST AND UPPERMOST MANTLE BENEATH THE CENTRAL AND WESTERN US BY JOINT INVERSION OF RECEIVER FUNCTIONS AND SURFACE WAVE DISPERSION

Synopsis of the chapter

By processing seismic data collected from over 800 USArray/TA stations deployed in the western US from the beginning of 2005 to the end of 2010, Rayleigh wave phase velocity measurements from 8 to 80 sec are acquired using ambient noise eikonal tomography and teleseismic earthquake Helmholtz tomography methods. Moreover, azimuthally independent receiver functions are constructed for these stations using the harmonic stripping technique described in Chapter II. These data sets are jointly interpreted to construct a 3-D Vs model for the crust and uppermost mantle beneath the western US using the joint Bayesian Monte Carlo inversion algorithm that is described in Chapter II. The resulting 3-D model is presented in the context of assessing the impact of introducing receiver functions to improve subsurface structures. I show that the assimilation of receiver functions improves the vertical resolution of the 3-D model by 1) reducing the range of estimated Moho depths, 2) improving the determination of the shear velocity jump across Moho, and 3) improving the resolution of the depth of anomalies in the uppermost mantle. In this chapter I also briefly describe a variety of geological and tectonic features revealed in the 3-D model. The use of this western US model by other geoscientists will be summarized in Chapter VI.

3.1 Introduction

Continental-scale arrays of seismometers with inter-station spacings between 50 and 100 km such as the EarthScope USArray Transportable Array (TA), the Chinese Earthquake Array, the Virtual European Broadband Seismic Network, or for that matter PASSCAL or USArray Flexible Array experiments that in some cases comprise more than 100 instruments, provide ideal data for surface wave tomography. The combination of ambient noise measurements, typically between about 8 and 40 sec period (e.g., in the US: Shapiro et al., 2005; Moschetti et al., 2007; Lin et al., 2008; Bensen et al., 2008), and earthquake-derived measurements, from

about 25 to 100 sec period (e.g., in the US: Pollitz, 2008; Lin et al., 2009, Lin and Ritzwoller, 2011a), produces broadband dispersion maps that constrain earth structure homogeneously through the crust to a depth of about 150 km in the uppermost mantle. Such broad-band measurements from ambient noise and/or earthquake data observed with USArray have been used by Yang et al. (2008), Pollitz and Snoke (2010), Moschetti et al. (2010a,b), Lin et al. (2010), Obrebski et al. (2011), and others to produce 3-D shear velocity models of the crust and uppermost mantle in the western US.

The use of surface wave dispersion data alone to produce models of the crust and uppermost mantle, however, presents significant non-uniqueness problems (e.g., Shapiro et al., 2002 and many others) because surface waves do not constrain the strength or location of jumps in shear velocity. Receiver functions, in contrast, provide the spatially discrete local response of seismic waves to discontinuities beneath receiver locations (Langston, 1979). As a consequence, combining surface wave data with receiver function data has been a natural direction for research and was introduced more than a decade ago (e.g., Last et al. 1997, Ozalaybey et al., 1997), with numerous realizations of the idea subsequently having been developed. In particular, joint inversions of receiver functions and surface wave dispersion within the context of non-linear, model-space sampling schemes have been developed in recent years (e.g., Chang et al., 2004; Lawrence and Wiens, 2004; Liu et al., 2010; Tokam et al., 2010; Bodin et al., 2011; Shen et al., 2013a).

Shen et al. (2013a) presents a non-linear Bayesian Monte-Carlo method to estimate a Vs model with uncertainties beneath stations by jointly interpreting surface wave dispersion and receiver functions and associated uncertainties. This method is designed for automated application to large arrays of broadband seismometers. Here, we apply this method to the joint inversion of surface wave dispersion maps and receiver functions with observations taken from 828 stations of the USArray Transportable Array (TA) as well as USArray reference network stations. The region of study extends eastward from the Pacific coast to 100 ºW longitude and covers the entire western US including parts of the Great Plains. This region extends about 1000 km eastward from earlier studies (Yang et al., 2008; Moschetti et al., 2010a,b), and includes data acquired through the year 2010, adding more than two years of TA data compared to these earlier studies. Significantly, as discussed here, the introduction of receiver functions into the inversion with

surface wave dispersion data from ambient noise and earthquake data significantly improves the vertical resolution of the model, revealing higher fidelity images of the crust and uppermost mantle across nearly half of the US.

3.2 Generation of the 3-D Model by Joint Inversion

3.2.1 Surface wave data

Rayleigh wave phase velocity measurements from 8 to 40 sec period were acquired from ambient noise using USArray TA stations deployed from the beginning of 2005 until the end of 2010. The data processing procedures described by Bensen et al. (2007) and Lin et al. (2008) were used to produce nearly 300000 dispersion curves between the 828 TA stations west of 100 ºW longitude and USArray backbone (or reference network) stations. Eikonal tomography (Lin et al., 2009) produced Rayleigh wave phase velocity maps for ambient noise from 8 to 40 sec period (e.g., Fig. 3.1a,b). Eikonal tomography is a geometrical ray theoretic technique that models off-great-circle propagation but not finite frequency effects (e.g., wavefront healing, back-scattering, etc.). Rayleigh wave phase velocity measurements from 25 to 80 sec period were obtained following earthquakes using the Helmholtz tomography method (Lin and Ritzwoller, 2011), also applied to TA data from 2005 through 2010. Example dispersion maps are shown in Figure 3.1c,e,f. A total of 1550 earthquakes were used with magnitude $M_s > 5.0$, of which on average about 270 earthquakes supplied measurements at each location. Helmholtz tomography is a finite frequency method that accounts for wavefield complexities that affect longer period surface waves, but Lin and Ritzwoller (2011) and Ritzwoller et al. (2011) argue that below about 40 sec period such corrections are not required.

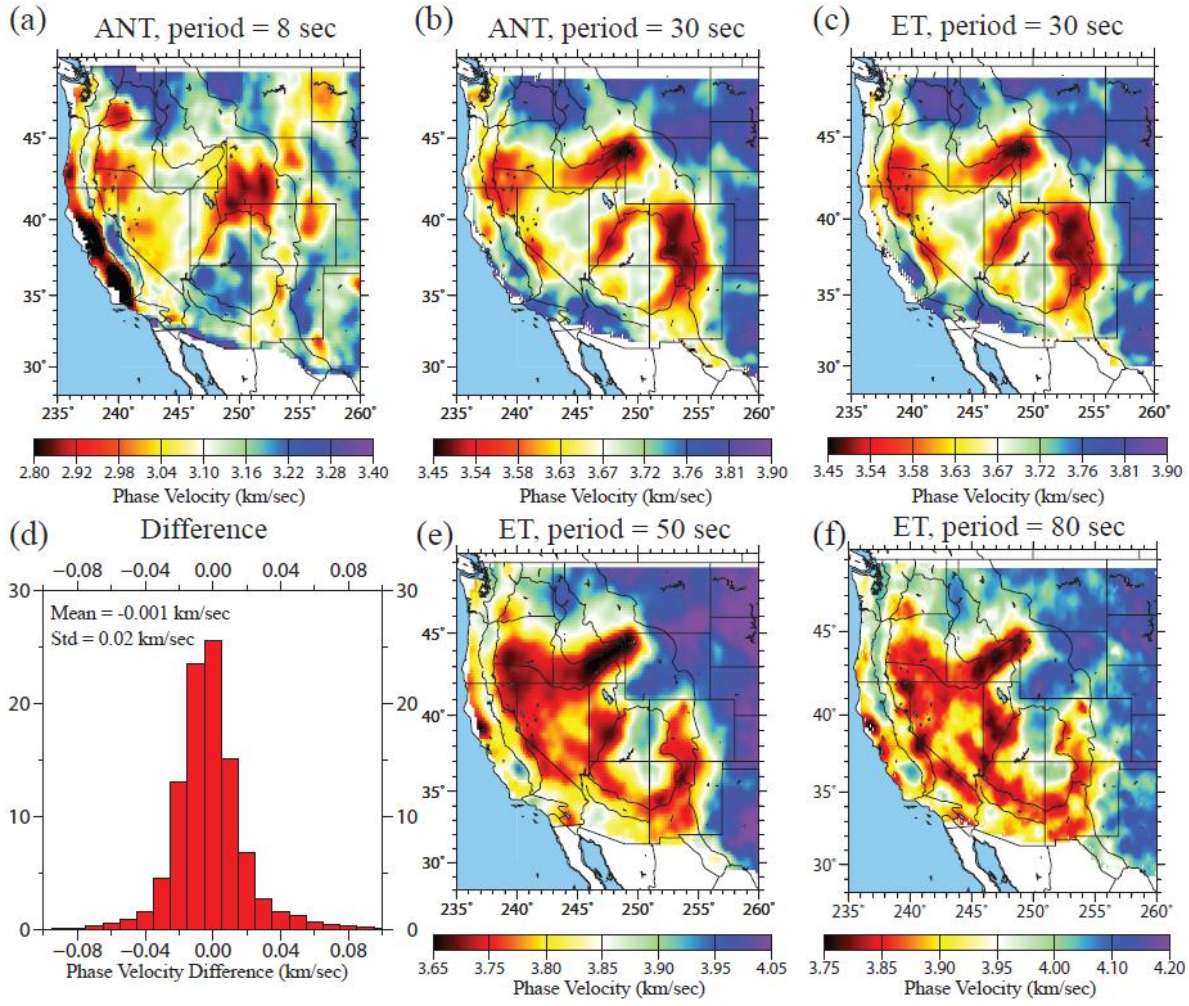

Figure 3.1 Example Rayleigh wave phase speed maps determined from (a,b) ambient noise data using eikonal tomography (ANT) and (c,e,f) earthquake data using Helmholtz tomography (ET) at the periods indicated. (d) Histogram of the differences in phase speeds at 30 sec period using ambient noise and earthquake data: mean difference is -1 m/s and the standard deviation of the difference is 20 m/s. Geological provinces are delineated by black lines in the maps.

Ambient noise and earthquake phase velocity maps overlap between 25 and 40 sec period. In this period band of overlap, the two measurements are averaged at each location based on their uncertainties. Uncertainties in Rayleigh wave dispersion maps derived from ambient noise average about 15 m/s between periods of 10 and 25 sec, and uncertainties of earthquake-derived maps also average about 15 m/s but between 30 and 60 sec period. At periods shorter and longer than these, uncertainties in each type of measurement grow. Therefore, the uncertainty of the

combined measurements is approximately flat, on average, at 15 m/s from 10 to 60 sec period, but grows at shorter and longer periods. An example of a dispersion curve with error bars at a point in the Basin and Range province is shown in Figure 3.2a. In the period band of overlap, the ambient noise and earthquake-derived maps agree approximately as well as expected given uncertainties in the maps, as Figure 3.1b,c,d illustrates at 30 sec period. The standard deviation of the difference between the maps is 20 m/s (Fig. 3.1d), consistent with the estimated uncertainty (~15 m /s) at 30 sec period.

3.2.2 Receiver functions

Single station receiver functions are constructed at each of the 828 TA stations west of 100 ℃W longitude using the method detailed described by Shen et al. (2012) and briefly summarized here. Earthquakes are used if they occur between 30 °-90 ° from the station with mb>5.0 during the lifetime of the station deployment. A time domain deconvolution method (Ligorria and Ammon, 1999) is applied to each seismogram windowed between 20 sec before and 30 sec after the P-wave arrival to calculate the radial component receiver functions with a low-pass Gaussian filter width of 2.5 s (pulse width ~ 1 sec). Corrections are made both to the time and amplitude of each receiver function, normalizing to a reference slowness of 0.06 sec/km (Jones and Phinney, 1998). The receiver function waveform is discarded after 10 sec because the slowness correction is appropriate for direct P-to-S conversions, but not for reverberations, and later arriving Moho reverberations are thus stacked down. An azimuthally independent receiver function, $R_0(t)$, for each station is computed by fitting a truncated Fourier Series at each time over azimuth and stripping the azimuthally variable terms using a method referred to as “harmonic stripping” (Shen et al., 2012), which exploits the azimuthal harmonic behavior in receiver functions (e.g., Girardin and Farra,1998; Bianchi et al., 2010). After removing the azimuthally variable terms at each time, the RMS residual over azimuth is taken as the 1σ uncertainty at that time. Shen et al. (2013a) describes procedures to assess and guarantee the quality of the receiver functions. On average, about 130 earthquakes satisfy the quality control provisions for each station across the region of study. An example of an azimuthally independent receiver function for a station in the

Basin and Range province is shown in Figure 3.2b as a pair of locally parallel black lines, which delineate the uncertainty at each time.

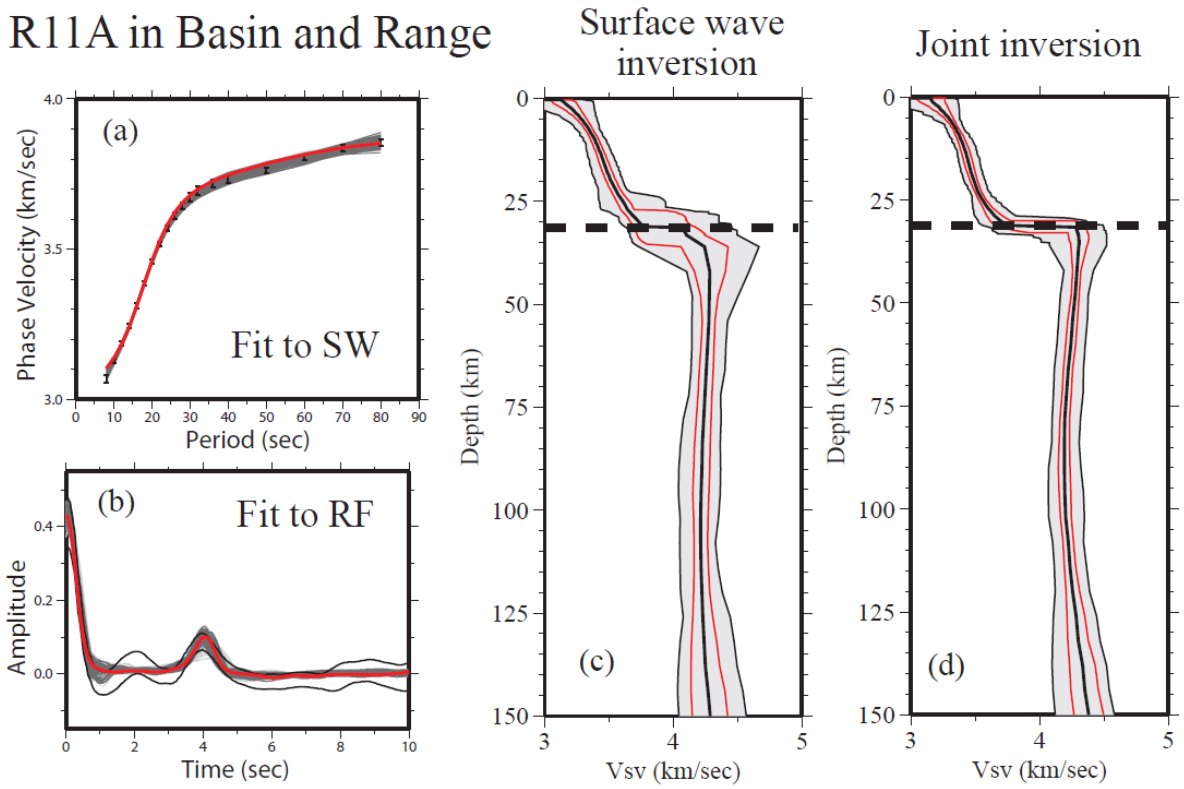

Figure 3.2 Example outcome of the joint inversion at USArray TA station R11A in the Basin and Range province in Currant, Nevada (38.35, -115.59). (a) Observed Rayleigh wave phase speed curve presented as 1σ error bars. Predictions from the ensemble of accepted models in (d) are shown (grey lines), as is the prediction from the best fitting model (redline). (b) The azimuthally independent receiver function $R_0(t)$ is shown with the black lines defining the estimated 1σ uncertainty. Predictions from the members of the ensemble in (b) are shown with grey lines, and the red line is the best fitting member of the ensemble. (c) Ensemble of accepted model using surface wave data alone. The full width of the ensemble is presented as black lines enclosing a grey-shaded region, the 1σ ensemble is shown with red lines, and the average model is the black curve near the middle of the ensemble. Moho is identified as a dashed line at ~ 32 km. (b) Ensemble of accepted models from the joint inversion.

3.2.3 Model parameterization

As a matter of practice, we seek models that have no more vertical structure than required to fit the data within a tolerance specified by the data uncertainties. At all points we attempt to fit the data with the same parameterization. There is one sedimentary layer with a linear velocity

gradient with depth. Three parameters are used to describe this layer: layer thickness and Vsv at the top and bottom of the layer. There is one crystalline crustal layer described by five parameters: layer thickness (km) and four B-splines for Vsv. Finally there is one upper-most mantle layer to a depth of 200 km described by five B-splines for Vsv. The smoothness of the model is imposed by the parameterization so that ad hoc damping is not needed during the inversion. In some places the model, which contains no discontinuities between the base of the sediments and the Moho and none in the mantle, may be too simple to fit the receiver function well or the dispersion data and receiver function jointly. We present examples of these types of cases below, and they provide the justification for an adaptive parameterization, which is used, for example, by Bodin et al. (2012).

Because only Rayleigh waves are used, there is predominant sensitivity to Vsv and we assume an isotropic Vsv model where Vs = Vsh = Vsv. We set the Vp/Vs ratio to 2.0 in the sedimentary layer and 1.75 in the crystalline crust and mantle. Reasonable variations around these values (e.g., Brocher, 2008) will produce little change in the isotropic shear wave speeds in the crust, but, as discussed by Shen et al. (2012), may change crustal thickness by up to 2 km in unusual circumstances. This value, however, lies within the estimated uncertainties for crustal thickness. For density we use a scaling relation in the crust that has been influenced by the studies of Christensen and Mooney (1995) and Brocher et al. (2005) and by Karato (1994) in the mantle where sensitivity to density structure is much weaker than in the crust. We also apply a physical dispersion correction (Kanamori and Anderson, 1977) using the Q model from PREM (Dziewonski & Anderson, 1981). The biggest deviations of Q from PREM probably occur in the uppermost mantle in regions of thin lithosphere, such as in the Basin and Range province. In such regions, low Q asthenosphere may extend nearly all the way upward to the crust. Replacing lithospheric Q=400 from PREM between Moho and 80 km depth with Q=100 would increase Vs estimated at 60 km depth in the Basin and Range province by less than 1%. This value also lies within stated 1σ uncertainties, but the inclusion of a more realistic Q model is a fruitful direction for future research. The model is reduced to 1 sec period.

3.2.4 Prior and posterior distributions of models

Prior models for Monte Carlo sampling are defined relative to a reference model (Shapiro and

Ritzwoller, 2002) subject to allowed perturbations (presented in Table 1 of Shen et al., 2013a) and model constraints. Constraints are that models are continuous between discontinuities at the base of the sediments and Moho, continuous in the mantle, and velocity increases linearly with depth in the sedimentary layer and monotonically with depth in the crystalline crust. The velocity contrasts across the sedimentary basement and across the Moho discontinuity are constrained to be positive and Vs < 4.9 km/s throughout the model. Examples of marginal prior distributions for several model variables are presented in Figure 3.3 as white histograms.

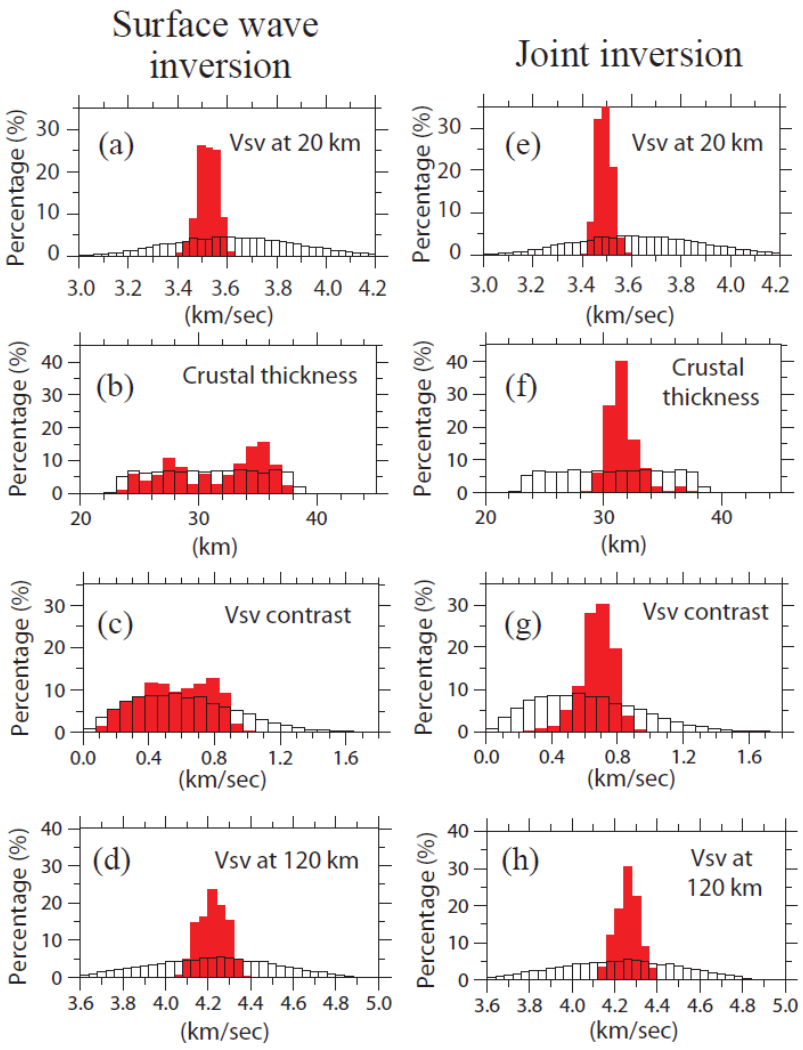

Figure 3.3 (a)-(d) Prior and posterior (surface waves only) marginal distributions of three model variables are presented with white and red histograms, respectively, for Vsv at 20 km depth , crustal thickness, the Vsv contrast across Moho, and Vsv at 120 km depth. (e)-(h) Same as (a)-(d), but the red histogram is for the posterior marginal distribution resulting from the joint inversion of receiver functions and surface wave phase velocities.

The Bayesian Monte Carlo joint inversion method described by Shen et al. (2013a) constructs a prior distribution of models at each location defined by allowed perturbations relative to the reference model as well as model constraints. The principal output is the posterior distribution of models that satisfy the receiver function and surface wave dispersion data within tolerances that depend on data uncertainties. The statistical properties of the posterior distribution quantify model errors. Examples of prior and posterior distributions for the inversion based on surface wave data alone are shown in Figure 3.3a-c, which illustrates that surface wave data alone do not constrain well crustal thickness or the jump in Vs across Moho but do determine Vs between discontinuities. In contrast, when receiver functions (e.g., Fig. 3.2b) and surface wave dispersion data (e.g., Fig. 3.2a) are applied jointly crustal thickness and the Vs jump across Moho are much more tightly constrained (Fig. 3.3d-f).

The details depend on the nature of the receiver function, but on average the vertical discontinuity structure of the crust is clarified and the vertical resolution of the model is improved by introducing receiver functions into the inversion with surface wave dispersion data. Figure 3.2c,d presents examples of model ensembles for a point in the Basin and Range province based on surface wave data alone compared with surface wave and receiver function data used jointly. Consistent with the observations of the marginal distributions shown in Figure 3.3, the introduction of receiver functions sharpens the image around the Moho, which reduces the trade-off between model variables in the lower crust and uppermost mantle, clarifying the thickness of the crust, the jump in Vs across the Moho, and reducing the spread of model velocities in the mantle.

Examples at other locations of data and resulting ensembles of models are presented by Shen et al. (2012) (in the Denver Basin, the Colorado Plateau, the Great Plains) and in Figures 3.4 and 3.5 here for complementary geological settings. The receiver functions in Figure 3.4a-e are typical and well-behaved in that the azimuthal variability is simple enough that the uncertainties are small and the azimuthally independent receiver function is well-defined. At these locations the surface waves and receiver functions can be fit well simultaneously and the introduction of receiver functions reduces the extent of the ensemble of accepted models, which are presented in Figure 3.5a-e.

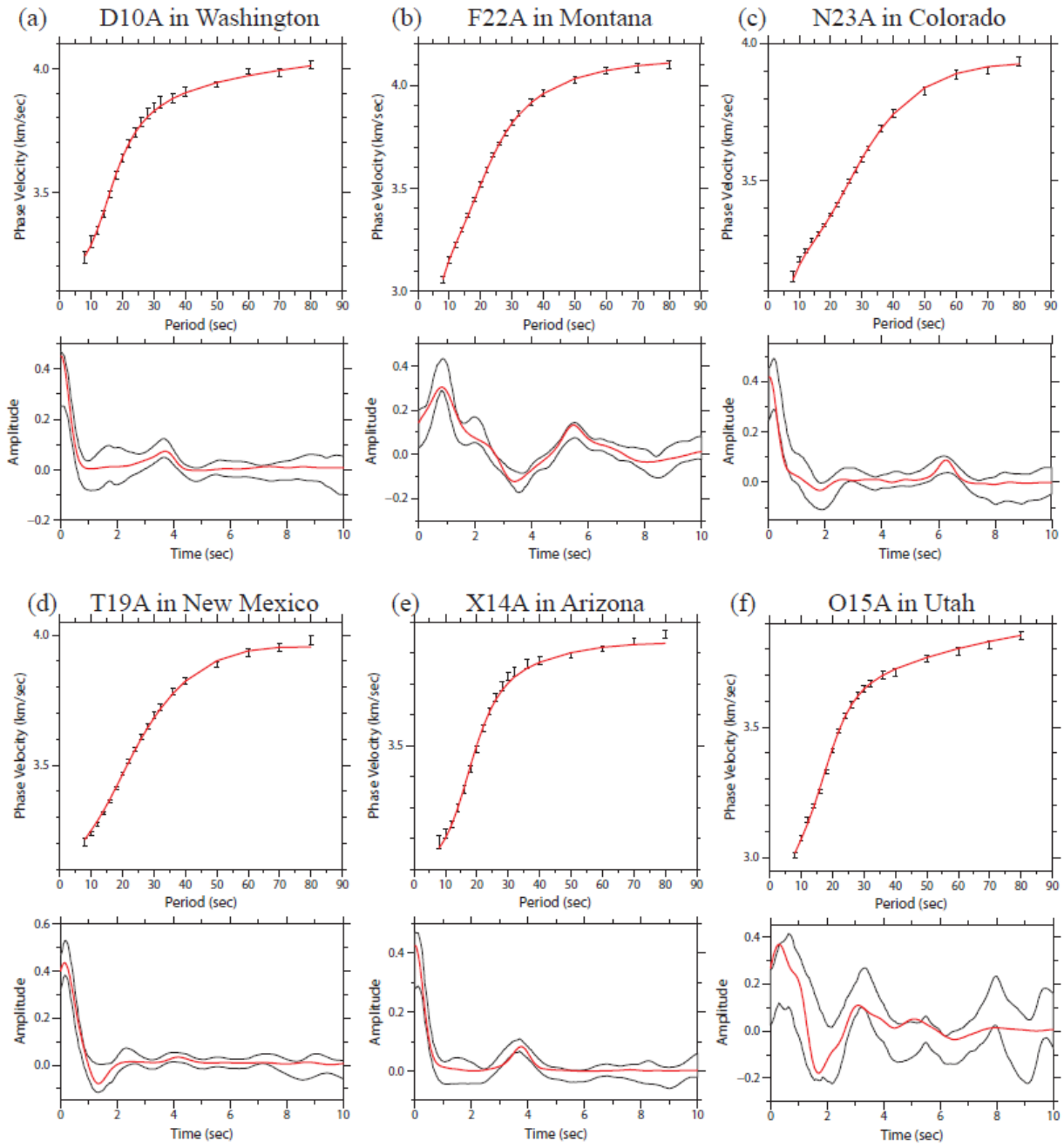

Figure 3.4 Six examples of azimuthally independent receiver functions (pair of black lines) and Rayleigh wave phase speed curves (error bars) compared with predictions from the best-fitting model from the joint inversion (red lines) found in Fig. 3.5. (a) TA station D10A, Oakesdale, WA (47.05, -117.28). (b) Station F22A, Rosebud, MT (45.78, -106.26). (c) Station N23A, Red Feather Lakes, CO (40.89, -105.94). (d) Station T19A, Beclabito, NM (36.83, -109.02). (e) Station X14A, Yava, AR (34.47, -112.89). (f) Station O15A, Rush Valley, UT (40.28, -112.47).

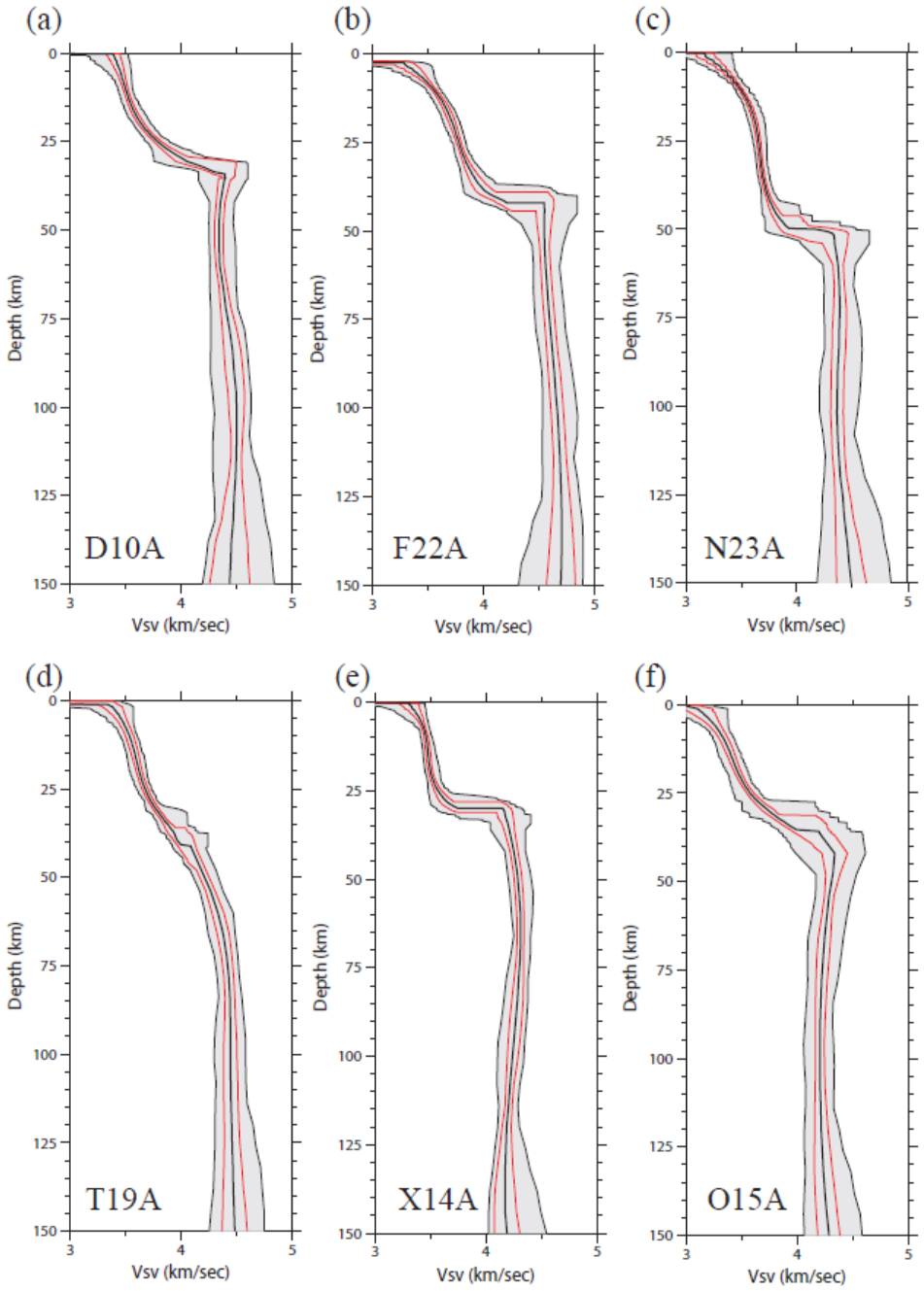

Figure 3.5 The ensembles of accepted models (defined as in Fig. 3.2c,d) from the joint inversion determined from the corresponding pairs of receiver functions and Rayleigh wave phase speed curves found in Fig. 3.4.

3.2.5 Places where receiver function uncertainties are large

In some locations, however, the receiver function is dominated by azimuthal variability so that uncertainties are very large. An example is seen in Figure 3.4f for a station in the Basin and Range province. For this station lateral heterogeneity between the local basin and adjacent mountain range is large enough to vitiate the azimuthally independent receiver function and the joint inversion reverts principally to fitting the surface wave data alone. At this point, the time-averaged uncertainty in the receiver function is about 0.11, which is the average half width of the corridor of the receiver function. As a consequence, the ensemble of models is somewhat broader in Figure 3.5f than at locations with a more accurate receiver function (e.g., Figs. 3.5a-e).

Such problems with receiver functions are relatively rare and mostly appear at discrete points rather than covering large regions. Figure 3.6a identifies the stations where the time-averaged uncertainty is larger than 0.07, which occurs for about 10% of the stations on the map. Stations where receiver function uncertainty is large because of sharp lateral variations appear in the Pacific Northwest, in parts of the Basin and Range province, and near the edges of the Rocky Mountain province. Large receiver function uncertainties also appear in the northern Great Plains but because of the reduction in the number of azimuthally dependent receiver functions caused by the greater distance from southwestern Pacific earthquakes. Even with relatively large uncertainties, however, the receiver functions provide important information in the inversion, but weaker constraints than at places where uncertainties are smaller.

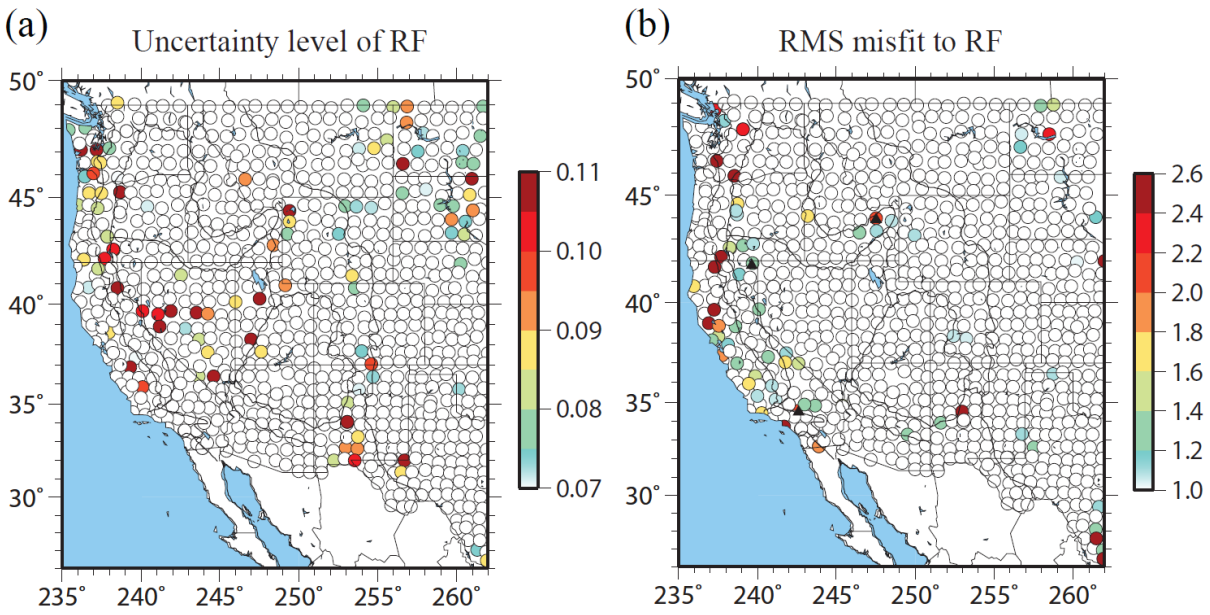

Figure 3.6 (a) The time-averaged uncertainty in the receiver function for each station is plotted across the region of study. Receiver function uncertainty is defined as the half-width of the receiver function corridor shown, for example, in Figs. 3.2b and 3.4. Values greater than 0.07 are considered large. (b) Misfit to the receiver function from the best-fitting model in the joint inversion is plotted. Misfit is defined as the square root of the reduced chi-squared value and referred to as χ_{min}. Black triangles are the locations of stations I15A, ADO, and MOD presented in Fig. 3.7.

3.2.6 Fit to the data

Surface wave data are fit acceptably in the joint inversion except near the far western periphery of the region where dispersion maps from ambient noise and earthquake tomography are most different in the period band of overlap. We believe this is caused mostly by degradation of the ambient noise maps due to azimuthal compression near the coast. In total, 817 out of the 829 stations have a surface wave reduced chi-squared value less than unity for the best fitting model in the accepted ensemble: $\chi^2_{min} < 1$.

Receiver functions are also usually well fit in the joint inversion such that 92% of the stations have a best-fitting reduced chi-squared value $\chi^2_{min} < 1$ for the receiver function. Receiver function misfits (reduced square root of chi-squared) from the jointly best fitting model at each location are presented in Figure 3.6b. The stations with larger receiver function misfits are

principally in California. This is largely due to differential resolution between the receiver function and surface wave data, particularly the sensitivity of the receiver function to relatively small-scale sedimentary basins, which are not reflected in the surface wave data. An example of this is shown in Figure 3.7a-c for a point in the Salton Trough in southern California. The receiver function (Fig. 3.7b) from 0-4 sec shows clear effects of deep sediments that are unfit by the receiver function predicted from the best-fitting model. In contrast, phase velocities at this location (Fig. 3.7c) do not display the low velocities at the short period end of the measurement band that are characteristic of the presence of sediments. At this point in the Salton Trough, the best fitting model fits the phase velocities exceptionally well and the receiver function much more poorly. Nevertheless, the observed receiver function at times greater than 4 sec is fit, meaning that it provides useful constraints on structures deeper in the crust.

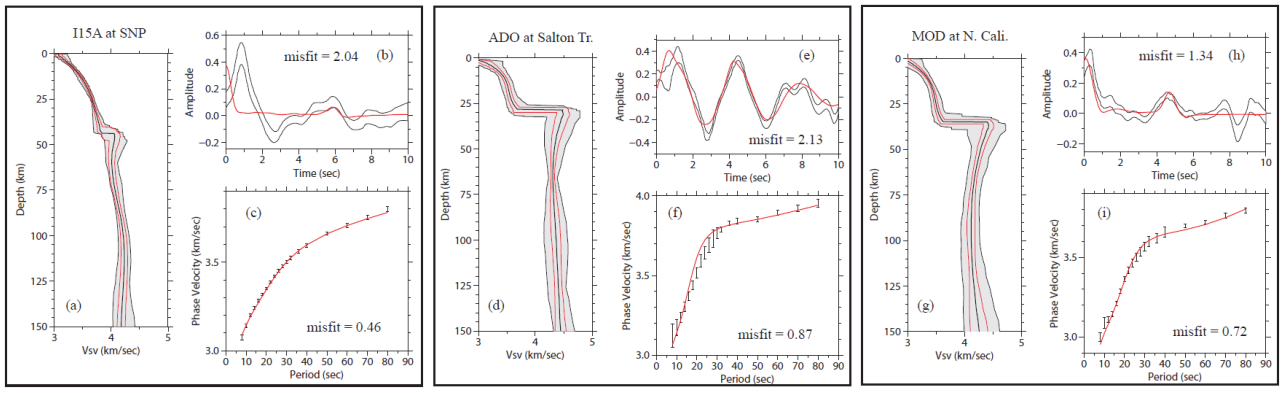

Figure 3.7 The three panels present examples of the joint inversion where the receiver functions ((b), (e), (h)) are poorly fit for stations in the Snake River Plain (I15A, SNP), the Salton Trough (ADO), and northern California (MOD). Station locations are presented in Fig. 3.6b. Plots (a), (d), and (g) are the ensemble of accepted models for each location. Fit to the local Rayleigh wave dispersion curves are presented in plots (c), (f), and (i). Formats of all plots are defined in Fig. 3.2.

Another example of a station where the receiver function is poorly fit is presented in Figure 3.7d-f for a point in the northeastern Snake River Plain. In this case, although the receiver function (Fig. 3.7e) is not fit well it is fit better than the receiver function in Figure 3.6b. However, the price is that the phase velocity data are not as well fit here. Again, this is a differential resolution problem in which the receiver function and surface wave data reflect different structures.

A final example of a receiver function in Northern California that is not well fit in the inversion

is presented in Figure 3.7g-i. This is an example of the effect of mis-parameterization, where the real earth possesses a structure that cannot be represented by the specified parameterization. The receiver function (Fig. 3.7h) shows an arrival near 8.5 sec presumably caused by a discontinuity or discontinuities in the uppermost mantle, which the smooth parameterization cannot reproduce. The phase velocity data also show the need for higher velocities in the uppermost mantle at periods between 50 and 70 sec. However, the receiver function is well fit at times before 6 sec, meaning that crustal structures are unaffected by the mantle mis-paramaterization. Issues like this can be resolved in a later round of model production by re-parameterizing the model.

As discussed by Shen et al. (2013a), more complicated Moho structure such as the imbricated Moho identified by Hansen and Dueker (2009) across the Cheyenne Belt in Wyoming would also not be modeled with our parameterization. The station spacing of the TA largely avoids the regions of the Cheyenne Belt with particularly distinct double Moho, however.

3.3. Discussion

The joint inversion is performed at the locations of the TA stations, producing an irregularly spaced mosaic of distributions of 1-D models separated, on average, by about 70 km. An example for the model at 120 km depth is shown in Figure 3.8a. At each depth, simple-kriging guided by model uncertainties (e.g., Schultz et al., 1998) is applied to interpolate the models onto a 0.25 °x0.25 ° grid as seen in Figure 3.8b. Views of averages of the distributions of accepted models are presented in Figures 3.8b, 9, 11, and 13.

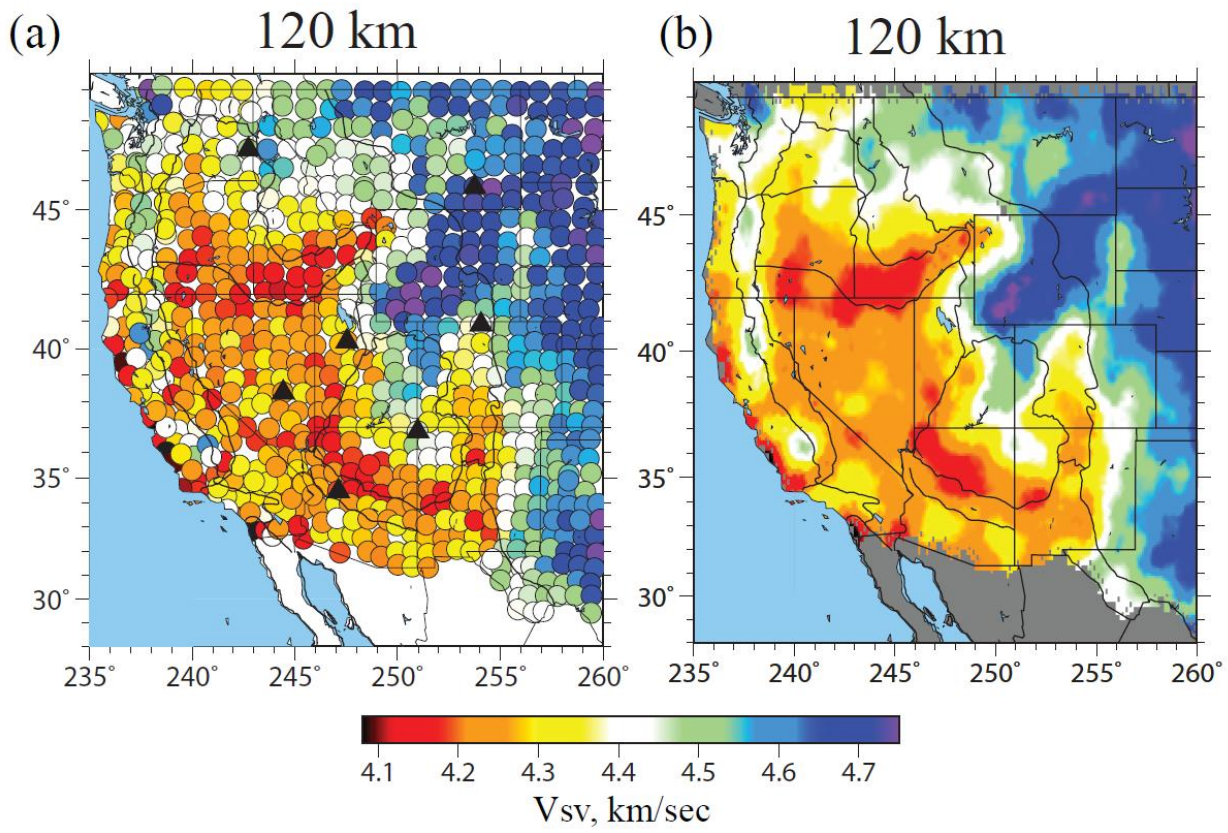

Figure 3.8 (a) Average of Vsv from the distribution of accepted models from the joint inversion in the depth range from 105 to 135 km, presented as a mosaic of results at each station location. Black triangles are station locations for results in Figs. 3-5. (b) The same as (a), but wave speeds have been interpolated onto a 0.25 °x0.25 ° grid by simple-kriging, in which model uncertainties guide the interpolation.

3.3.1 Features of the 3-D model

The upper crustal part of the model is dominated by the existence or absence of sediments. The low velocity features in Figure 3.9a are located at the major sedimentary basins of the central and western US, including the Central (or Great) Valley of California, the Pasco basin in Washington, most prominently the Wyoming-Utah-Idaho thrust belt covering southwestern Wyoming, the Powder River basin in northeast Wyoming, the Williston basin in North Dakota and eastern Montana, the Denver basin in northeast Colorado, and parts of the Permian and Anadarko basins near the edge of the map in Texas. The velocity anomaly of the Pasco basin is weaker than the others because of capping or interlayering of sediments by basalt flows. The sedimentary parameterization in the model is quite simple and inaccuracies in the resulting

sedimentary model may cause low velocities to leak into the upper crystalline crust, but no deeper. Although sediments are not prominent features of the vertical profiles presented in Figure 3.11, they do appear clearly beneath the Great (or Central) Valley of California in profiles B-B' and D-D'. High upper crustal velocities in Figure 3.9a, reflecting the lack of sediments, are correlated principally with extrusive volcanism (e.g. Columbia Plateau) or mountain ranges; e.g., the Rocky Mountain cordillera, the Colorado Plateau, the eastern and southern Basin and Range province, and the Sierra Nevada.

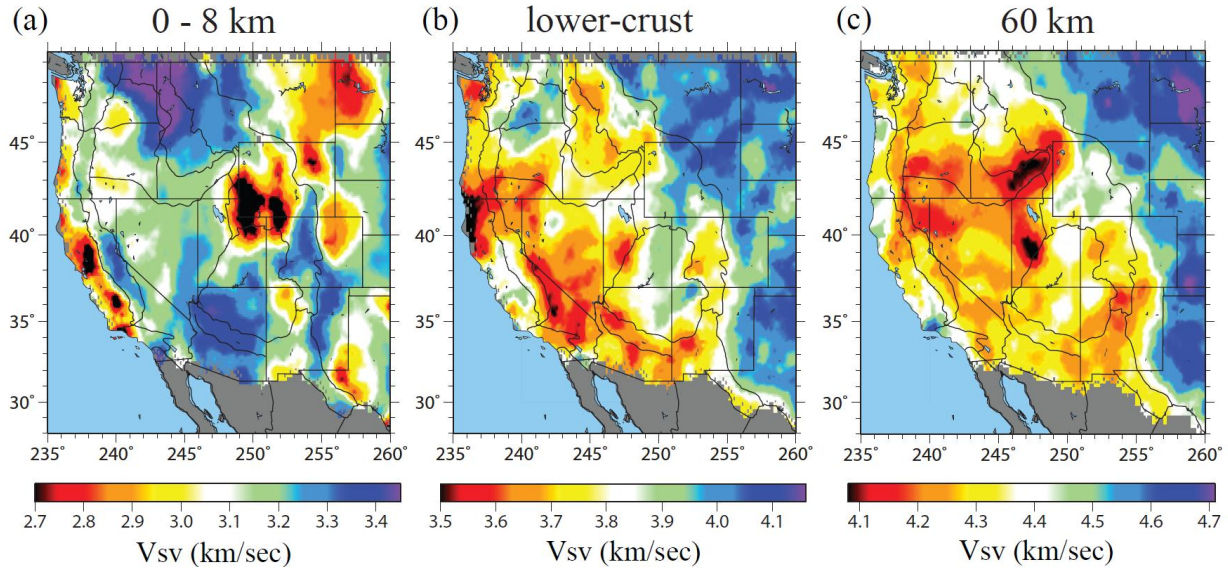

Figure 3.9 Average of Vsv from the distribution of accepted models from the joint inversion at three depths. (a) Average of the model between 0 and 8 km depth, (b) average of the model in the lower 4 km of the crust above Moho, and (c) the average of the model in the depth range between 50 and 70 km.

The predominant large-scale feature that appears in the lower crust (Fig. 3.9b) is the dichotomy between the high shear wave speeds in the Great Plains east of the Rocky Mountain front and the generally lower wave speeds in the tectonically deformed west. High lower crustal velocities in the west are associated with Pasco basin intrusives of the Columbia River flood basalt province, the Wyoming craton, and the interior (particular the northern interior) of the Colorado Plateau. Relative high velocities are found beneath the Snake River Plain. Mid-crustal velocities are higher relative to surroundings than lower crustal velocities. These high lower crustal velocities in the western US may be compositional in origin, reflecting either past volcanism or the presence of Archean cratonic blocks. A likely cause for fast lower crust under the Columbia

Plateau and Snake River Plain is a mafic underplate associated with basaltic magmatism. Low velocities in the lower crust are found prominently in the southern Basin and Range province, on the western and southern peripheries of the Colorado Plateau, and in coastal California north of Cape Mendocino near the southern end of the subducting Gorda plate. Fast lower crustal velocities appear conspicuously in the Great Plains on the vertical profiles of Figure 3.11, but also strikingly beneath the Colorado Plateau (A-A', D-D'). The persistence of high-velocity, presumably strong lower crust under the Laramide-affected Wyoming craton and the Colorado Plateau suggests that crustal strength may influence surface deformation. In contrast, the lower crustal velocity contrast tracking the Rocky Mountain deformation front and Rio Grande Rift truncates much older Proterozoic mafic lower crust (the "7.xx layer" seen by Gorman et al., 2002 in Wyoming and Montana and the granite-rhyolite province lower crustal restite proposed by Karlstrom et al., 2005) to the east, and cuts across the general NE-SW trend inherited from continental assembly (Whitmeyer and Karlstrom, 2007).

Crustal thickness and the jump in shear velocity across the Moho, two of the variables most improved by the introduction of receiver function waveforms in the inversion, are presented in Figures 3.10a,b. To first order, crustal thickness agrees with similar estimates made by Gilbert (2012) using receiver functions alone, except in regions where the Moho converted arrival, which provides information about crustal thickness, is of very low amplitude or absent, such as beneath parts of the Colorado Plateau (Figs. 3.4d, 3.5d). The thickest crust is in the northern Colorado Rockies (~60 km), consistent with the regional receiver function analysis of Karlstrom et al. (2012) using a denser array. Thinnest crust is less than 25 km in the Salton Trough. Even across the Great Plains, crustal thickness varies appreciably, from about 42 km depth in northern Nebraska to about 55 km near the Montana-North Dakota border. Another example of thinned crust is observed near the boundary of Wyoming and Montana, which was also observed by Gilbert (2012).

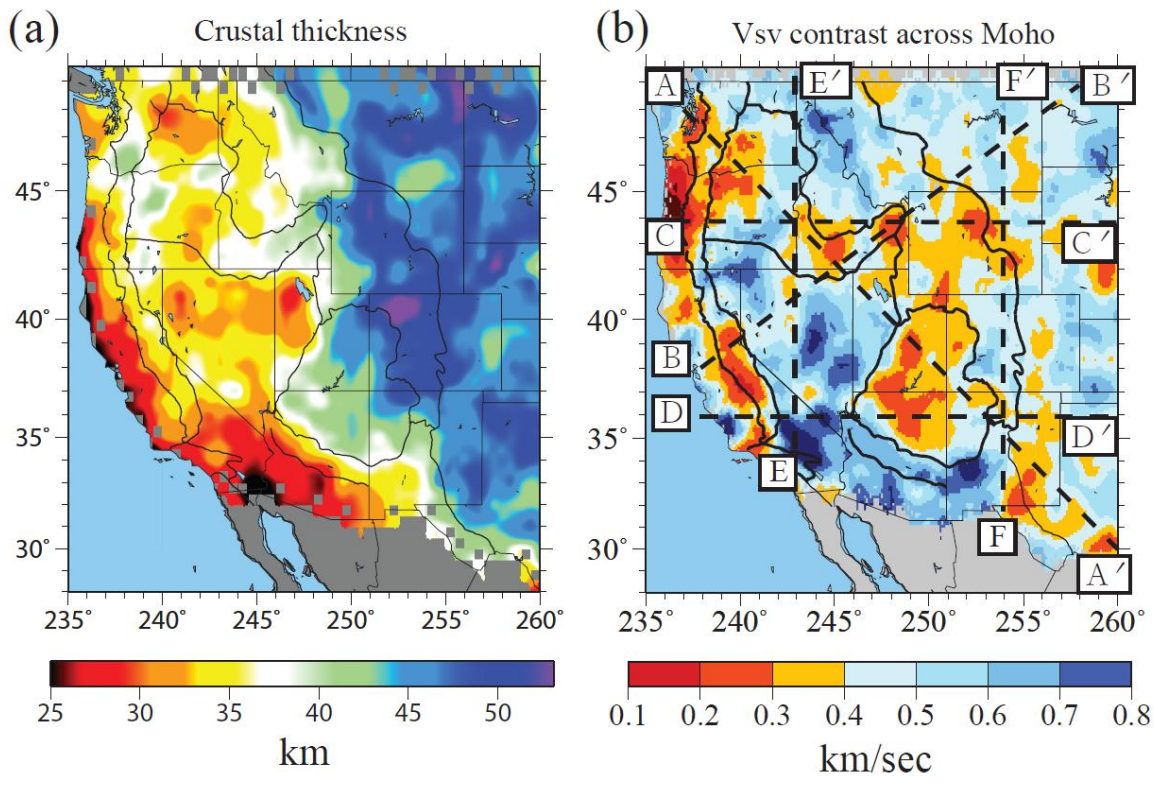

Figure 3.10 Average of (a) crustal thickness and (b) the jump in Vsv across the Moho from the ensemble of accepted models derived from the joint inversion. Vertical model profiles that are plotted in Fig. 3.11 are identified in (b).

Crustal thickness is not simply related to surface elevation; for example, there is little obvious difference between the crustal thicknesses near the Rocky Mountain front with the Great Plains. The 3D model presented here, however is the basis for the prediction of surface elevation from derived density variations within the crust and uppermost mantle (e.g., Levandowski et al., 2012). It is found that isostatic surface topography predicted from density variations derived from this 3D Vs model generally fits the observed topography across the western US and the exceptions are informative.

Regions that have a low amplitude for the Moho converted arrival in the receiver function appear in Figure 3.10b as small jumps in Vs across Moho, and are presented with warm colors. These regions include parts of the coastal Pacific Northwest of Oregon, the Sierra Nevada, the Snake River Plain, parts of Wyoming, and the Colorado Plateau. A relationship to lithospheric

delaminations or mantle drips is possible, as proposed in the Southern Sierra Nevada (Zandt et al., 2004, Boyd et al., 2004) and the Colorado Plateau (Levander et al., 2011, Bailey et al., 2012), which are underlain by relatively fast uppermost mantle. However, a reduced Moho velocity contrast is also consistent with high-velocity lower crust, such as an eclogitic root under the western Sierra Nevada foothills (Frassetto et al., 2011) or a mafic underplate. Most of the Basin and Range province has a large contrast in Vs across the Moho, as do parts of the Rocky Mountains and the Great Plains. The reduced contrast in Vs across the Moho west of the Cascade Range in the Pacific Northwest contains the region hypothesized to have serpentinized uppermost mantle (e.g., Bostock et al., 2002; Brocher et al., 2003).

Low mantle velocities (Fig. 3.8, 3.9c) are generally correlated with lower crustal velocities across the region probably because of the continuity of thermal anomalies across the Moho. Principal exceptions are the Pasco basin in southern Washington, the Snake River Plain, and the southwestern Basin and Range adjacent to the Sierra Nevada (Walker Lane), perhaps indicative of compositional heterogeneity in the crust caused by fast volcanic intrusives overlying slow hot mantle. The model in the mantle is resolved to scales of geological and tectonic relevance, and many features are observed on the vertical profiles of Figure 3.11 that are deserving of further detailed analysis and thermal modeling.

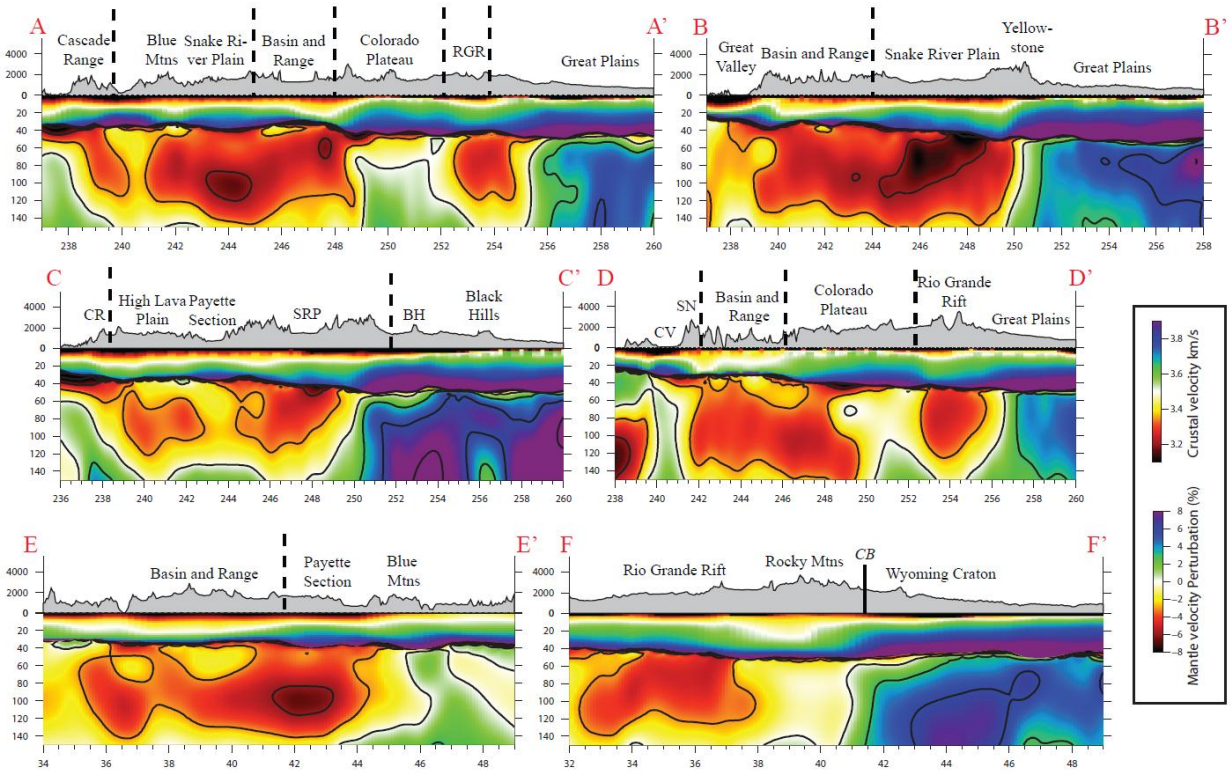

Figure 3.11 Vertical Vsv model profiles for the six transects identified in Fig. 3.10b with dashed lines. Speeds in the crust are in absolute units and those in the mantle are presented as percent perturbations relative 4.4 km/s, except along profile C-C' where it is 4.32 km/s. Mantle velocity contours appear at 3% intervals. (RGR: Rio Grande Rift, CR: Cascade Range, SRP: Snake River Plain, BH: Bighorn Range, CV: Central Valley; SN: Sierra Nevada; CB: Cheyenne Belt).

Examples on profile A-A' include the observation of low velocities in the supra-slab wedge beneath the Cascade Range, the Snake River Plain, the Basin and Range Province, and the Rio Grande Rift as well as high velocities beneath the Colorado Plateau. On profile B-B', the along hot spot track tilt of the low velocity anomaly underlying the Snake River Plain is seen as well as the rapid transition to high velocities beneath the Great Plains. The subducting Juan de Fuca slab appears on profile C-C' along with a resolved deep relative low velocity feature underlying the Black Hills of South Dakota. It is interesting and enigmatic that the Black Hills mantle anomaly is exclusively deep in the model. The mantle drip (Boyd et al., 2004; Zandt et al., 2004; Yang and Forsyth, 2006) underlying the Central Valley of California and the western Sierra Nevada is seen in profile D-D', along with low velocities beneath the Rio Grande Rift and the rapid

transition to the Great Plains province in eastern New Mexico, which is consistent with previous studies (Wilson et al., 2002, Gao et al., 2004, West et al., 2004). The eastward down-tilted low anomaly beneath the Basin and Range/western Colorado Plateau viewed in profile D-D' may be related to the complex embedment of the previous flat-subducted Farallon slab. The thin lithosphere observed beneath the Colorado Plateau in profiles A-A' and D-D' is consistent with earlier observations by Levander et al. (2011) and Bailey et al. (2012). Profile E-E' presents a south-north transect of the Basin and Range province illustrating relatively higher velocities directly underlying Moho that characterize parts of this province. Higher velocities underlie the Blue Mountains, consistent with body wave tomography (Hales et al., 2005) and dip northward with increasing depth. Finally, profile F-F' presents a south-north transect of the Rio Grande Rift and Rocky Mountains, illustrating the much lower shear wave speeds underlying the rift, and the rapid transition to high velocity mantle across the Cheyenne belt at the edge of the Archean Wyoming craton. The high velocities present in the lower crust of the Wyoming Craton in profile F-F' may diagnose the "double Moho" north of the Cheyenne Belt, which has been argued to result from underplating based on previous refraction studies (e.g. Gorman et al., 2002; Rumpfhuber et al., 2009a,b).

On average, the uppermost mantle structure is similar to previous models in the overlapping area (Yang and Ritzwoller, 2008; Moschetti et al., 2010), and is also similar to models constructed from surface wave/body wave tomography by Schmandt and Humphreys (2010) and Obrebski et al. (2011). However, the assimilation of receiver functions greatly improves the vertical resolution in the top 100 km of the model, whereas body wave tomography reveals much deeper structure.

3.3.2 Model uncertainties

One of the advantages of the Bayesian Monte Carlo method is that it provides uncertainties in the final model, determined from the standard deviation of the resulting marginal posterior distribution at each depth. Figures 3.8, 3.9, 3.10 and 3.11 present averages of the distributions at each depth, examples of which are shown in Figure 3.3d-f. Figures 3.12 and 3.13 summarize the uncertainties at the depths and for crustal thickness and Vs jump across Moho shown in Figures 3.8, 3.9, and 3.10. Uncertainties in the upper crust average about 100 m/sec, with larger values

near the Wyoming basins (150-200 m/sec) due to the complexity of sedimentary structure there. Uncertainties also grow near the periphery of the maps because surface wave dispersion uncertainties increase there. In the lower crust, uncertainties are slightly smaller than in the upper crust but are still substantial due to remaining trade-offs with Moho depth and uppermost mantle velocities. The pattern of uncertainties at this depth (Fig. 3.12b), therefore, is correlated with the uncertainties in crustal thickness (Fig. 3.11a). At 60 km depth, uncertainties are much smaller than in the upper and lower crust, about 50 m/s across much of the region, because this depth is separated from structural discontinuities across most of the region. In regions with thick crust, however, uncertainties increase to about 100 m/sec. At 120 km depth, uncertainties increase again because the structural sensitivity of surface waves decays with depth.

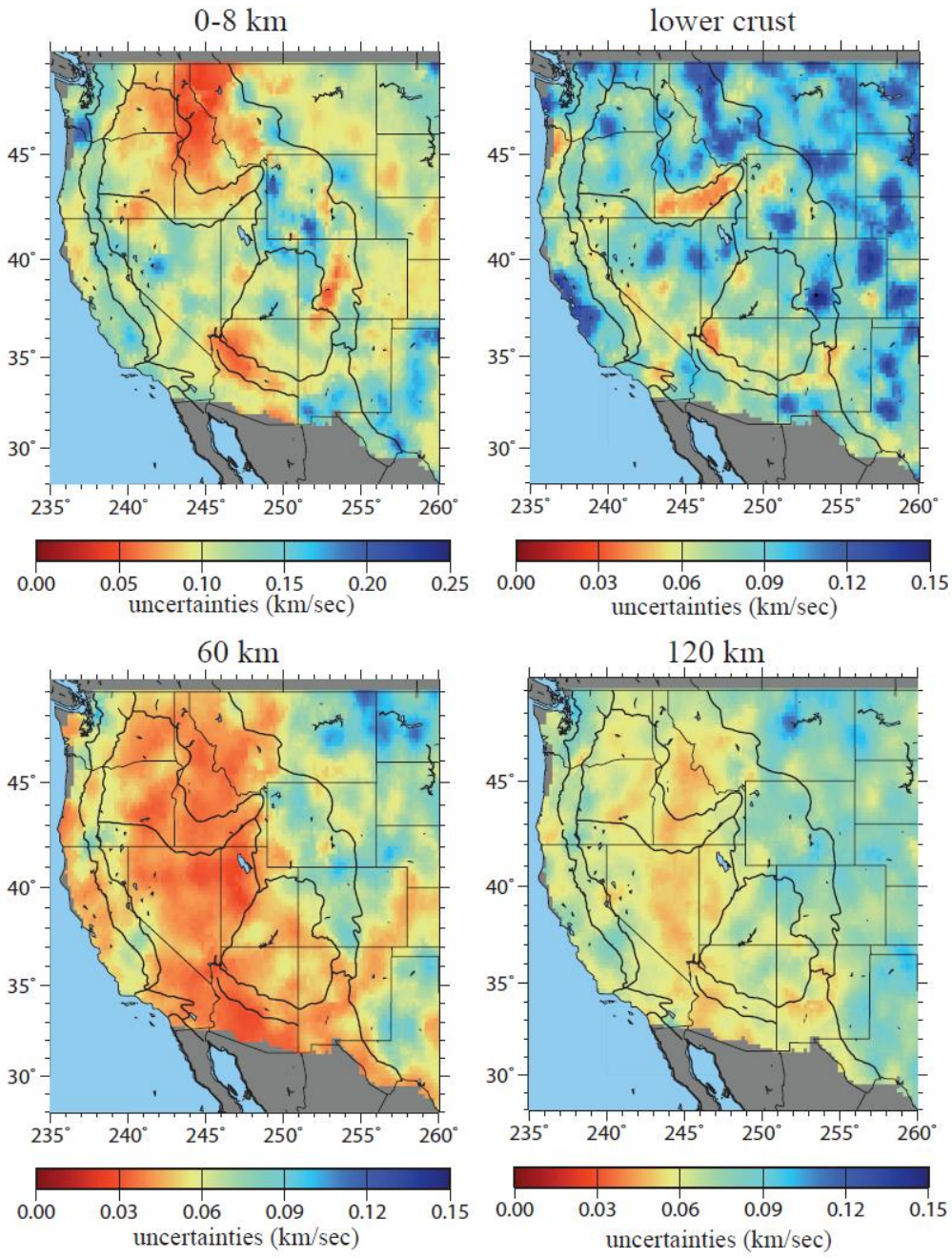

Figure 3.12 Estimates of uncertainties (1σ) in the 3-D model at three four depths: (a) 0-8 km average, (b) lower crustal average (4 km above Moho to Moho), (c) 60 km, and (d) 120 km.

Figure 3.12c presents the spatial average of the standard deviation of these distributions across the study region for both the joint inversion and the inversion with surface wave data alone. The introduction of receiver functions reduces model uncertainties predominantly between depths of 35 and 50 km, by introducing information about the depth to Moho and the velocity contrast

across it. This uncertainty, represented as the spread of the distribution, however, does not include bias. In fact, relative to the earlier studies of Yang et al. (2008) and Moschetti et al. (2010a,b) the improvements actually extend to greater depths because to reduce the spread of models near Moho those studies imposed a positivity constraint on the velocity derivative with depth, which is not needed in the joint inversion. In some locations that constraint obscured the location in depth of the mantle anomalies and its removal further improves the vertical resolution of the model.

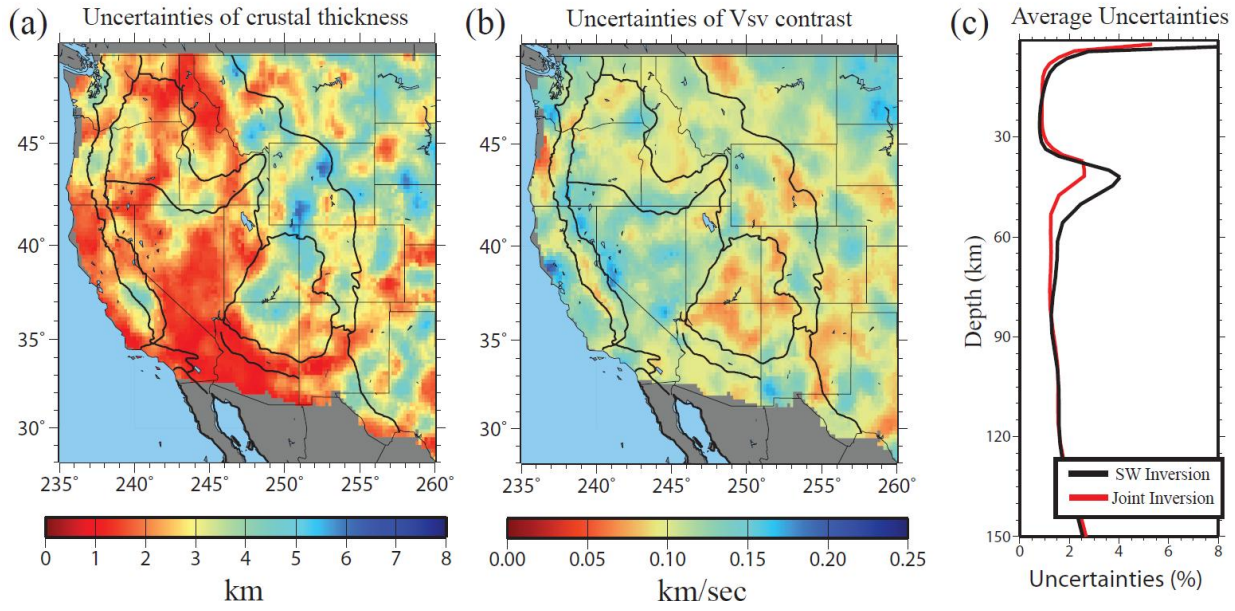

Figure 3.13 Estimates of uncertainty (1σ) for (a) estimated crustal thickness and (b) the jump in shear wave speed across the Moho. (c) Model uncertainties (1σ) averaged over the study region from the inversion with surface wave (SW) data alone (black line) and from the joint inversion (red line).

3.4. Conclusions

The joint analysis of surface waves from ambient noise and earthquake data produces high resolution (~70 km) Rayleigh wave dispersion maps from about 8 sec to 80 sec period. Finite frequency effects should be accounted for at periods above about 40 sec (Lin and Ritzwoller, 2011b) and are effected here through Helmholtz tomography (Lin and Ritzwoller, 2011a). Eikonal tomography (applied to ambient noise data) and Helmholtz tomography (applied to earthquake data) provide the uncertainty information needed to interpret the signals in ambient noise and earthquake maps quantitatively and reliably. In the frequency band of overlap between

ambient noise and earthquake data, the ambient noise and earthquake derived phase velocity maps agree remarkably well.

Bayesian Monte Carlo inversions of the surface wave dispersion data alone reveal that shear wave speeds at depths well separated from first-order discontinuities (base of sediments, Moho) are well recovered by surface wave data alone. Near the discontinuities, however, trade-offs between the interface depths and adjacent shear wave speeds spread the ensemble of accepted models. Ad-hoc constraints introduced to reduce such trade-offs have other deleterious effects, such as causing mantle anomalies to move to different depths. The assimilation of receiver functions into the joint inversion with Rayleigh wave dispersion curves overcomes these difficulties, improving the vertical resolution of the model by reducing the range of estimated Moho depths, improving the determination of the shear velocity jump across Moho, and improving the resolution of the depth of anomalies in the uppermost mantle. The joint inversion produces a distribution of models beneath each station, which is summarized by the mean and standard deviation of the distribution at each depth. Across the vast majority of the region studied, the surface wave dispersion and receiver functions can be fit jointly with a very simple model parameterization, producing a smooth model between the two internal discontinuities. There are, however, details in the receiver functions, signals arriving more than 10 seconds after the direct P-arrival, as well as azimuthal variations in the receiver functions that call for further refinement of the models with structures at greater depths or with local 2-D structures and anisotropy. But, these features are beyond the scope of the present study.

A great variety of geological and tectonic features are revealed in the 3-D model presented here that forms the basis for more detailed local to regional scale analysis and interpretation in the future.The 3-D model is useful as a basis for many other types of studies: studies of azimuthally/radially anisotropic structures, investigations of the density/thermal structure of the study region, and as a starting model for other types of seismic investigations such as Common-Conversion Point stacking and body wave tomography (e.g., Obrebski et al., 2011). In addition, the 3-D model can be used to improve regional seismic event location based on short period surface waves, examples of which have been discussed by Barmin et al. (2011) and Levshin et al. (2012).

Acknowledgments. The major component of the chapter is summarized in Shen et al. (2013b). Co-authors of the paper include Michael H. Ritzwoller and Vera Schulte-Pelkum. The authors thank two anonymous reviewers for helpful comments that improved this chapter. The facilities of the IRIS Data Management System, and specifically the IRIS Data Management Center, were used to access the waveform and metadata required in this study. The IRIS DMS is funded through the National Science Foundation and specifically the GEO Directorate through the Instrumentation and Facilities Program of the National Science Foundation under Cooperative Agreement EAR-0552316. This research was supported by NSF grants EAR- 0711526, EAR- 0844097, EAR-0750035, and EAR-1053291 at the University of Colorado at Boulder.

CHAPTER IV

CRUSTAL AND UPPERMOST MANTLE STRUCTURE IN THE CENTRAL US ENCOMPASSING THE MIDCONTINENT RIFT

Synopsis of the chapter

In this chapter, the joint Bayesian Monte-Carlo inversion described in Chapter II is applied to 122 USArray/TA stations in the central US encompassing the Midcontinent Rift (MCR), and a three dimensional Vs model for the crust and uppermost mantle is constructed, providing an overview of the seismic structure of the MCR and adjacent areas. In sections 4.1, 4.2 and 4.3, I present the geological background, the data processing of surface wave dispersion and receiver functions, and the application of the joint Bayesian Monte Carlo inversion to the study region, respectively. In section 4.4, I discuss the tectonic implications from the resulting 3-D model in detail leading to three major conclusions. 1) In the central US, a high correlation between the long wavelength gravity field and shallow Vs structure is observed and the MCR gravity high is obscured by clastic sediments in the shallow crust, which is consistent with an upper crustal origin to the MCR gravity anomaly as well as other anomalies in the region. 2) Thick crust (>47 km) underlies the MCR, which is evidence for post-rifting compression along the entire rift. 3) Crustal shear wave speeds vary across the Precambrian sutures (e.g., Great Lakes Tectonic Zone, Spirit Lakes Tectonic Zone), which reveals that the Precambrian sutures played an important role in the subsequent tectonic evolution of the central US. This chapter illustrates that detailed analysis and interpretation of the 3-D model at local to regional scales can be applied fruitfully to other regions using the models generated in Chapters III and V.

4.1. Introduction

The most prominent gravity anomaly in the central US (Woollard and Joesting, 1964) is a 2000-km-long gravity high with two arms merging at Lake Superior and extending southwest to Kansas and southeast into Michigan. This anomaly has been determined to mark a late Proterozoic tectonic zone (Whitmeyer and Karlstrom, 2007). Because of the existence of mantle-source magmatism and normal faults found along the anomaly, it has been generally accepted that this anomaly resulted from a continent rifting episode and as a consequence is known as the Midcontinent Rift system (MCR) (Schmus, 1992) or Keweenawan Rift system.

Geochronological evidence shows that the rift initiated at about 1.1 Ga and cut through several crustal provinces (Hinze et al., 1997). Although various geological and seismic studies have focused on the rift (Hinze et al., 1992; Mariano and Hinze, 1994; Woelk and Hinze, 1991; Cannon et al., 1989; Vervoort et al., 2007; Hollings et al., 2010; Hammer et al., 2010; Zartman et al., 2013), the mechanisms behind the opening and rapid deactivation and subsequent closure of the rift are still under debate. Fundamental questions, therefore, remain as to the nature and origin of the rift system (Stein el al., 2011).

Figure 4.1 outlines the location of the western arm of the MCR and its neighboring geological provinces. The MCR can be thought of as being composed of three large-scale components: the western arm through Minnesota, Iowa, and Kansas; the Lake Superior arm; and the eastern arm through Michigan. Our focus here is on the western arm. A free air gravity high defines its location (see 40 mgal anomaly contour in Fig. 1b), but it divides further into three segments: a northern one extending from near the south shore of Lake Superior along the Wisconsin-Minnesota boundary, a southern one extending from northeastern to southeastern Iowa, and a small segment in Nebraska and Kansas. Marginal gravity minima flank the MCR and have been interpreted as a signature of flanking sedimentary basins (Hinze et al., 1992). Gravity decreases broadly from 0 - 30 mgal in the north to -30 mgal in the south. Whether and how these gravity features relate to the structure of the crust and uppermost mantle is poorly understood.

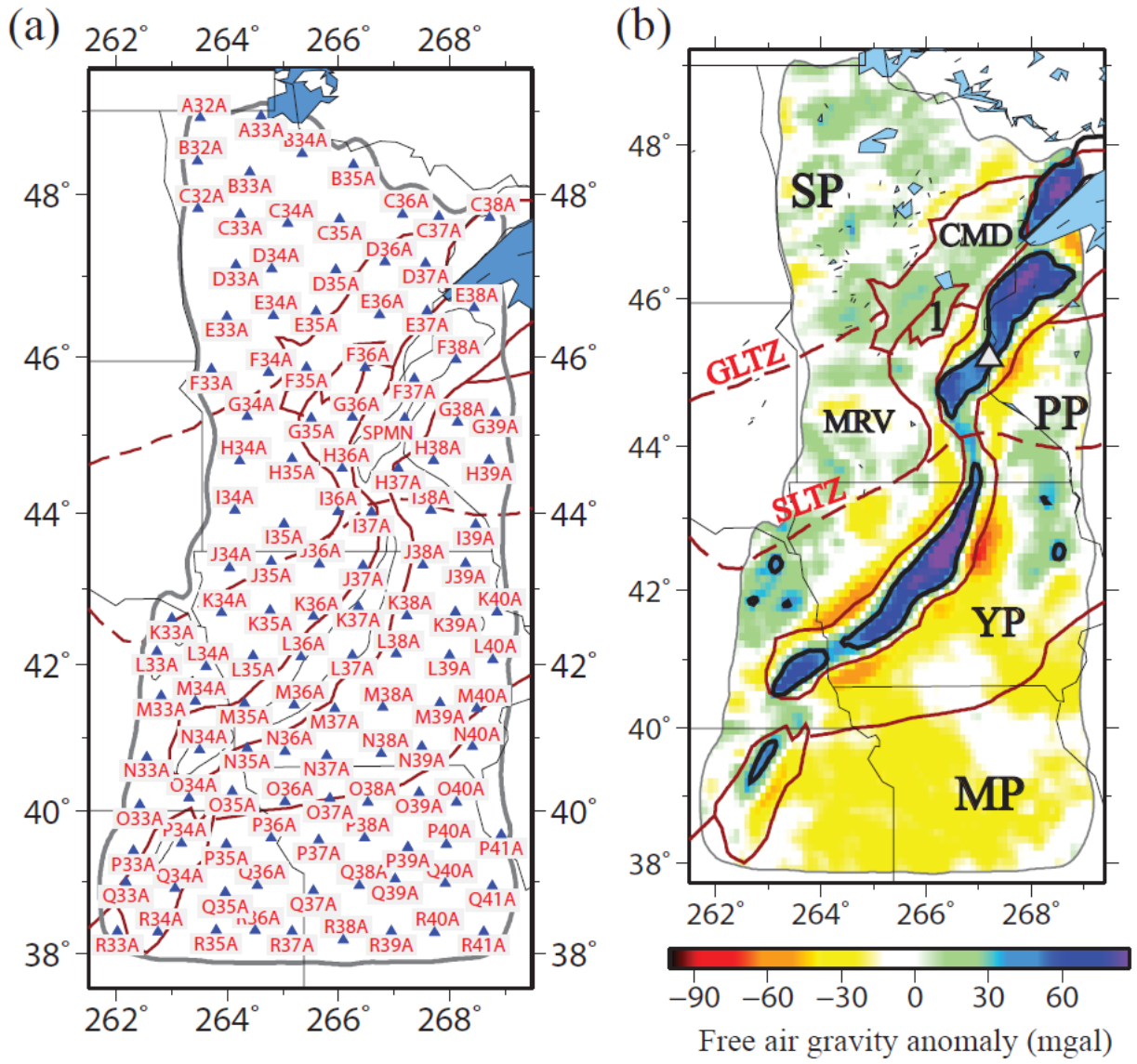

Figure 4.1. (a) The 122 seismic stations used in this study are named and shown with triangles. (b) The free-air gravity anomaly (Pavlis et al., 2012) is plotted in the background, with the 40 mgal level contoured with black lines to highlight the location of the major positive gravity anomaly along the Midcontinent Rift (MCR). Simplified tectonic boundaries are shown as solid dark red curves, which are identified with abbreviations: Superior Province (SP); Craton Margin Domain (CMD), Minnesota River Valley (MRV), Penokean Province (PP), Spirit Lake Tectonic Zone (SLTZ), Yavapai Province (YP), and Mazatzal Province (MP). A dashed line crossing the S. Dakota and Minnesota border is the Great Lakes Tectonic Zone (GLTZ), which separates the Superior Province into the Superior Greenstone Terrane to the north and MRV to the south. The number "1" indicates the location of the east-central Minnesota Batholith.

The western arm of the MCR remains somewhat more poorly characterized than the Lake Superior component of the MCR due to a veneer of Phanerozoic sediments (Hinze et al., 1997). About two decades ago active seismic studies were performed from northwestern Wisconsin to northeastern Kansas (Chandler et al., 1989; Woelk and Hinze, 1991) that revealed a structure similar to the Lake Superior component (Hinze et al. 1992). These similarities include crustal thickening to more than 48 km and high-angle thrust faults that appear to be reactivated from earlier normal faults. Van Schmus et al. (1993) and Cannon (1994) attributed these features to a post-rifting compressional episode during the Grenville orogeny.

The MCR cuts across a broad section of geological provinces of much greater age. The Superior segment of the MCR is embedded in the Achaean Superior Province (SP, 2.6-3.6 Ga), which continues into Canada. In Minnesota, this province is subdivided by the Great Lakes Tectonic Zone (GLTZ) west of the MCR into the 2.6-2.75 Ga Greenstone-Granite Terrane in the north, and the 3.4-3.6 Ga Gneiss Terrane or "Minnesota River Valley Sub-Province" (MRV) in the south (Sims and Petermar, 1986). During the Paleoproterozoic (1.8-1.9 Ga), the Penokean Province (PP) is believed to have been accreted to the southern edge of the Superior Province, adding vast foreland basin rocks and continental rocks along its margin. This is marked as a "craton margin domain (CMD)" (Holm et al., 2007) in Figure 1. From 1.7-1.8 Ga, the Yavapai province was added to the southern Minnesota River Valley and the Penokean provinces, which drove overprinting metamorphism and magmatism along the continental margin to the north. The East-Central Minnesota Batholith ("1" in Fig. 4.1) is believed to have been created during this time (Holm et al., 2007) and this accretion produced the continental suture known as the Spirit Lake Tectonic Zone (SLTZ). Later (1.65-1.69 Ga), the Mazatzal Province (MP) was accreted to the Yavapai Province, producing another metamorphic episode south of the Spirit Lake Tectonic Zone. Overall, the 1.1 Ga rift initiated and terminated in a context provided by geological provinces ranging in age from 1.6 to 3.6 Ga. During the Phanerozoic, this region suffered little tectonic alteration.

In this chapter, we aim to produce an improved, uniformly processed 3D image of the crust and uppermost mantle underlying the western arm of the MCR and surrounding Precambrian geological provinces and sutures. The purpose is to provide information on the state of the lithosphere beneath the region using a unified, well-understood set of observational methods. We are motivated by a long list of unanswered questions concerning the structure of the MCR,

including the following. (1) How are observed gravity anomalies related to the crustal and uppermost mantle structure of the region, particularly the gravity high associated with the MCR? (2) Is the crust thickened (or thinned) beneath the MCR, and how does it vary along the strike of the feature? (3) Is the MCR structurally a crustal feature alone or do remnants of its creation and evolution extend into the upper mantle? (4) Are the structures of the crust and uppermost mantle continuous across sutures between geological provinces or are they distinct and correlated with such provinces?

These are challenging questions for passive source seismology (ambient noise and earthquake tomography). The MCR is long and narrow and a difficult target tomographically. Thermal anomalies associated with the rifting and subsequent closure of the MCR have thermally equilibrated making the region relative homogeneous compared, for example, with the western US (e.g, Shen et al., 2013a). However, since 2010, the Earthscope/USArray Transportable Array (TA) left the tectonic western US and rolled over the region encompassing the western arm of the MCR, making it possible to obtain new information about the subsurface structure of this feature. The earlier deployment of USArray stimulated the development of new seismic data analysis and imaging methods. This includes ambient noise tomography (e.g., Shapiro et al., 2005; Lin et al. 2008; Ritzwoller et al., 2011) performed with new imaging methods such as eikonal tomography (Lin et al., 2009) as well as new methods of earthquake tomography such as Helmholtz tomography (Lin and Ritzwoller, 2011) and related methods (e.g., Pollitz, 2008; Yang et al., 2008b; Pollitz and Snoke, 2010). New methods of inference have also been developed based on Bayesian Monte Carlo joint inversion of surface wave dispersion and receiver function data (Shen et al., 2013a) that yield refined constraints on crustal structure with realistic estimates of uncertainties. The application of these methods together have produced a higher-resolution 3-D shear velocity (Vs) model of the western US (Shen et al., 2013b) with attendant uncertainties and have also been applied on other continents (e.g., Zhou et al., 2011; Zheng et al., 2011; Yang et al., 2012; Xie et al., 2013). In this chapter, we utilize more than 120 TA stations that cover the MCR region to produce high-resolution Rayleigh wave phase velocity maps from 8 to 80 sec period by using ambient noise eikonal and teleseismic earthquake Helmholtz tomography. We then jointly invert these phase velocity dispersion curves locally with receiver functions to produce a 3-D Vsv model for the crust and uppermost mantle beneath the western MCR and the surrounding region.

4.2 Data Processing

The 122 USArray stations used in this study are shown in Figure 1a as black triangles, which evenly cover the study area with an average inter-station distance of about 70 km. Based on this station set, we construct surface wave dispersion curves from ambient noise and earthquake data as well as receiver functions. Rayleigh wave phase velocity curves from 8 to 80 sec period are taken from surface wave dispersion maps generated by eikonal tomography based on ambient noise and Helmholtz tomography based on teleseismic earthquakes. We also construct a back-azimuth independent receiver function at each station by the harmonic stripping technique. Details of these methods have been documented in several papers (eikonal tomography: Lin et. al., 2009; Helmholtz tomography: Lin and Ritzwoller et al., 2011; harmonic stripping: Shen et al., 2013a) and are only briefly summarized here.

4.2.1 Rayleigh wave dispersion curves

We measured Rayleigh wave phase velocities from 8 to 40 sec period from ambient noise cross-correlations based on the USArray TA stations available from 2010 to May 2012. We combined the 122 stations in the study area with the TA stations to the west of the area (Shen et al. 2013b) in order to increase the path density. The ambient noise data processing procedures are those described by Bensen et al. (2007) and Lin et al. (2008), and produce more than 10,000 Rayleigh wave phase velocity curves in the region of study. As shown by Yang and Ritzwoller (2008) and many others, the azimuthal content of ambient noise in the US is sufficient to estimate phase and group velocities for Rayleigh and Love waves reliably. At these periods (8 to 40 sec), eikonal tomography (Lin et al., 2009) produces Rayleigh wave phase velocity maps with uncertainties based on ambient noise (e.g., Fig. 4.2a-c). For longer periods (25 to 80 sec), Rayleigh wave phase velocity measurements are obtained from earthquakes using the Helmholtz tomography method (Lin and Ritzwoller, 2011) in which finite frequency effects are modeled. A total of 875 earthquakes between 2010 and 2012 with Ms > 5.0 are used, and on average each station records acceptable measurements (based on a SNR criterion) from about 200 earthquakes for surface wave analysis. Example maps are presented in Figure 4.2d,e. In the period band of overlap between the ambient noise and earthquake measurements (25 to 40 sec), there is strong agreement between the resulting Rayleigh wave maps (Fig. 4.2f). The average difference is ~ 1

m/sec, and the standard deviation of the difference is ~ 12 m/sec, which is within the uncertainties estimated for this period (~15 m/sec).

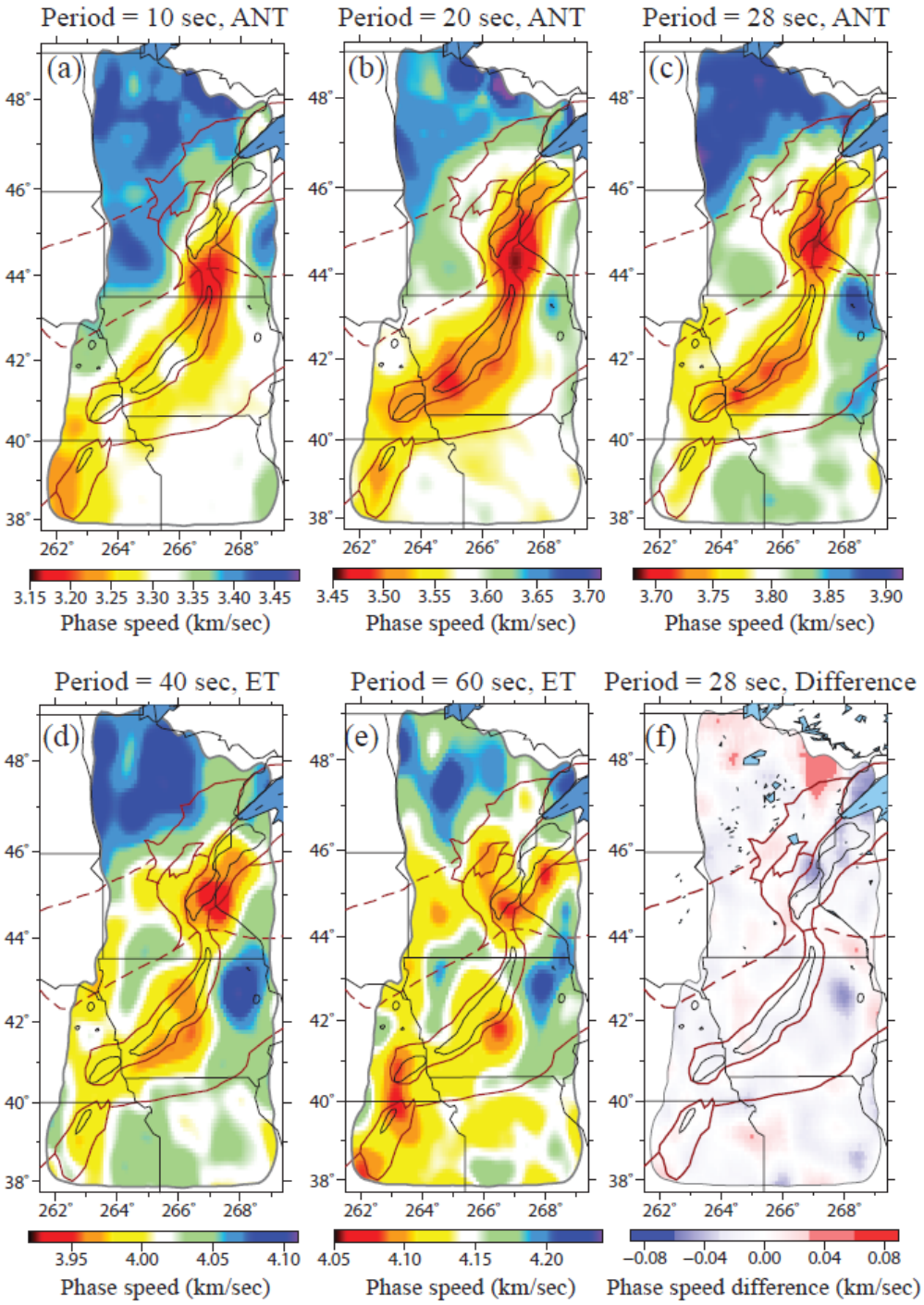

Figure 4.2 Rayleigh wave phase velocity maps from ambient noise tomography (ANT) and earthquake tomography (ET). (a-c) Maps from ambient noise eikonal tomography at periods of (a)10, (b) 20, and (c) 28 sec. (d-e) Maps at (d) 40 and (e) 60 sec from teleseismic earthquake

Helmholtz tomography. (f) Difference between the phase velocity map from ANT and ET at 28 sec period.

At 10 sec period, at which Rayleigh waves are primarily sensitive to sedimentary layer thickness and the uppermost crystalline crust, a slow anomaly is seen in the gap between the northern and southern MCR, and runs along the flanks of the MCR, particularly in the south. Wave speeds are high north of the Great Lakes Tectonic Zone (GLTZ) and average in the Mazatzal Province. Between 20 and 40 sec, the most prominent feature is the low speed anomaly that runs along the MCR, as was also seen by Pollitz and Mooney (2013). This indicates low shear wave speeds in the lower crust/uppermost mantle and/or a thickened crust beneath the MCR. Higher wave speeds at these periods appear mostly north of the Great Lakes Tectonic Zone. At longer periods, the anomaly underlying the MCR breaks into northern and southern segments with the lowest wave speeds shift off the rift axis near the southern MCR. Most of our study region is outside the area covered by the short period dispersion maps from the earlier data by Liang and Langston (2008), so comparison with that study is impossible. With these Rayleigh wave phase speed dispersion maps at periods between 8 and 80 sec, we produce a local dispersion curve at each station location. For example, the local Rayleigh wave phase velocity curve with uncertainties at station E33A in the southern Superior Province is shown in Figure 4.3a with black error bars. Other example Rayleigh wave curves are presented in Figure 4.3b-f.

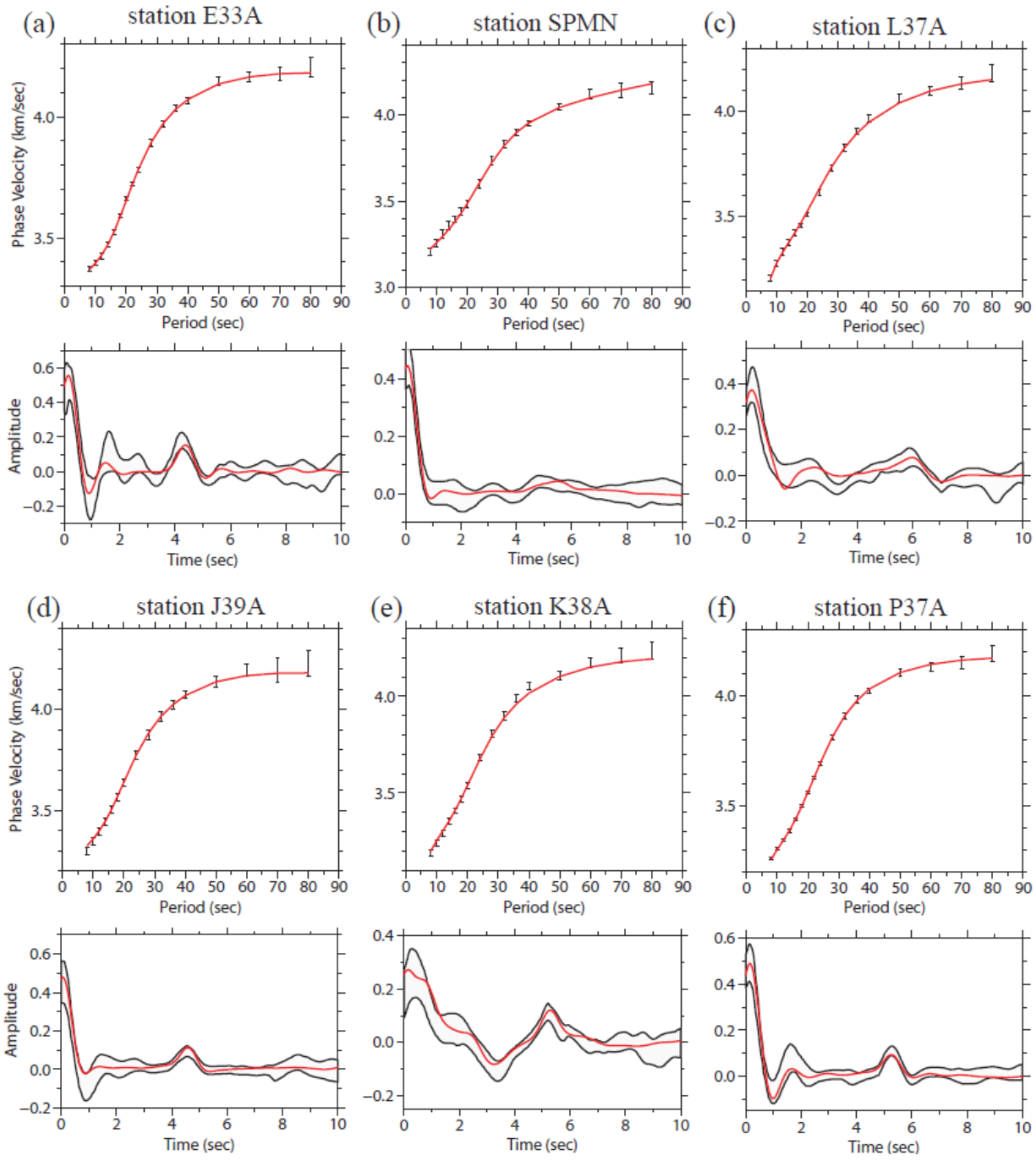

Figure 4.3 Examples of local Rayleigh wave phase velocity curves with uncertainty estimates (black error bars) and the azimuthally independent receiver functions (parallel black waveforms) are compared with predicted dispersion curves and receiver functions from the best fitting model at each location (red curves): (a) station E33A in the Southern Superior Province, (b) SPMN in the northern MCR, (c) L37A in the southern MCR, (d) J39A in northeastern Iowa east of the MCR, (e) K38A on the eastern flank of the southern MCR, and (f) P37A in the Mazatzal Province.

4.2.2 Receiver function processing

The method we use to process receiver functions for each station is described in detail in Shen et al. (2013b). For each station, we pick earthquakes from the years 2010, 2011 and 2012 in the distance range from 30° to 90° with mb > 5.0. We apply a time domain deconvolution method (Ligorria and Ammon, 1999) to each seismogram windowed between 20 sec before and 30 sec after the P-wave arrival to calculate radial component receiver functions with a low-pass Gaussian filter with a width of 2.5 s (pulse width ~ 1 sec), and high-quality receiver functions are selected via an automated procedure. Corrections are made both to the time and amplitude of each receiver function, normalizing to a reference slowness of 0.06 sec/km (Jones and Phinney, 1998). Finally, only the first 10 sec after the direct P arrival is retained for further analysis. We compute the azimuthally independent receiver function, $R_0(t)$, for each station by fitting a truncated Fourier Series at each time over azimuth and removing (or stripping) the azimuthally variable terms using a method referred to as "harmonic stripping" by Shen et al. (2013b). This method exploits the azimuthal harmonic behavior of receiver functions caused by sloping interfaces and anisotropy (e.g., Girardin and Farra, 1998; Bianchi et al., 2010). After removing the azimuthally variable terms at each time, the RMS residual over azimuth is taken as the 1σ uncertainty at that time.

On average, about 72 earthquakes satisfy the quality control provisions for each station across the region of study, which is about half of the average number of similarly high quality recordings at the stations in the western US (Shen et al., 2013b). This reduction in the number of accepted receiver functions results primarily from the distance range for teleseismic P (30^o to 90^o), which eliminates many events from the southwest Pacific (e.g., Tonga). The number of retained earthquakes varies across the region of study, being highest towards the southern and western parts of the study region and lowest towards the north and east. At some stations there are as few as 21 earthquake records retained and receiver functions at 15 stations display a large gap in back-azimuth, which prohibits applying the harmonic stripping method. For these stations, we use a simple, directly-stacked receiver function to represent the local average. Overall, the quality of the resulting azimuthally independent receiver functions is significantly lower than observed across the western US by Shen et al. (2013a,b) where more than 100 earthquakes are typically retained for receiver function analysis, but is high enough to provide reliable information at nearly every station in the study region.

Examples of receiver functions at six stations in the MCR region are shown in Figure 3 as parallel black lines that delineate the one standard deviation uncertainty at each time. At station E33A (Fig. 4.3a) in the southern Superior Province, a clear Moho conversion appears at ~ 4.3 sec after the direct P arrival, which indicates a distinct, shallow (~ 35 km) Moho discontinuity. In contrast, at station SPMN in the northern MCR (Fig. 4.3b) only a subtle Moho Ps conversion is apparent, which suggests a gradient Moho beneath the station. In the southern MCR, the receiver function at station L37A (Fig. 4.3c) has a Moho Ps signal at ~ 6 sec, implying the Moho discontinuity is at over 45 km depth. At station J39A to the east of the MCR (Fig. 4.3d), a Moho Ps conversion at ~4.5 sec is observed, indicating a much thinner crust. At K38A, which is located in the gravity low of the eastern flank of the southern MCR (Fig. 4.3e), a sediment reverberation appears after the P arrival. In the Mazatzal Province at station P37A (Fig. 4.3f) a relatively simple receiver function is observed with a Moho Ps conversion at ~ 5.3 sec, indicative of crust of intermediate thickness in this region.

4.3 Construction of the 3-D model from Bayesian Monte Carlo joint inversion

Here we briefly summarize the joint Bayesian Monte Carlo inversion of surface wave dispersion curves and receiver functions generated in the steps described in section 2. A 1-D joint inversion of the station receiver function and dispersion curve is performed on the unevenly distributed station grid and then the resulting models from all stations are interpolated into the 3-D model using a simple kriging method, as described by Shen et al. (2013b).

4.3.1 Model space and prior information

We currently only measure Rayleigh wave dispersion, which is primarily sensitive to Vsv, so we assume the model is isotropic: Vsv=Vsh=Vs. However, the possible existence of positive crustal radial anisotropy (Vsh > Vsv) would mean that our Vs estimate is low relative to a Voigt average. The Vs model beneath each station is divided into three principal layers. The top layer is the sedimentary layer defined by three unknowns: layer thickness and Vs at the top and bottom of the layer with Vs increasing linearly with depth. The second layer is the crystalline crust, parameterized with five unknowns: four cubic B-splines and crustal thickness. Finally, there is the uppermost mantle layer, which is given by five cubic B-splines, yielding a total of 13 free parameters at each location. The thickness of the uppermost mantle layer is set so that the total thickness of all three layers is 200 km. The model space is defined based on perturbations to a

reference model consisting of the 3D model of Shapiro and Ritzwoller (2002) for mantle Vsv, crustal thickness and crustal shear wave speeds from CRUST 2.0 (Bassin et al., 2000), and sedimentary thickness from Mooney and Kaban (2010). Because the reference sediment model is inaccurate in the region of study, we empirically reset the reference sedimentary thickness at stations that display strong sedimentary reverberations in the receiver functions. Similarly, in some regions where there is a gradient Moho we iteratively reset Moho depth in the reference model to the value from similar regions of the model without a gradient Moho (e.g,. northern MCR crustal thickness in the reference model has been iteratively reset to the estimated crustal thickness of the southern MCR). Following Shen et al. (2013b), the Vp/Vs ratio is set to be 2 for the sedimentary layer and 1.75 in the crystalline crust/upper mantle (consistent with a Poisson solid). Density is scaled from Vp by using results from Christensen and Mooney (1995) and Brocher (2005) in the crust and Karato (1993) in the mantle. The Q model from PREM (Dziewonski and Anderson, 1981) is used to apply the physical dispersion correction (Kanamori and Anderson, 1977) and our resulting model is reduced to 1 sec period. Increasing Q in the upper mantle from 180 to 280 will reduce the resulting Vs by less than 0.5% at 80 km depth.

In addition to the choices made in parameterizing the model, we impose three prior constraints in the Monte Carlo sampling of model space. (1) Vs increases with depth at the two model discontinuities (base of the sediments and Moho). (2) Vs increases monotonically with depth in the crystalline crust. (3) Vs < 4.9 km/sec at all depths. These prior constraints are imposed for two reasons. First, our philosophy is to seek simple models that fit the data and we allow complications in structure only where required by the data. Therefore, we see our inversions as testing the hypothesis that the Earth conforms to the simplifications represented by these constraints. If we are able to fit the data with very simple models, then we possess no evidence for complexity beyond our parameterization and constraints. That does not mean that such complexity is non-existent, however. If we are unable to fit aspects of the data, then we relax the constraints or generalize the parameterization accordingly. The strongest of these constraints is the second together with the choice of the B-spline parameterization in the crust; namely, the requirement that crustal shear wave speeds increase monotonically and continuously with depth through the crystalline crust. Discussion of whether this constraint may be too rigid in parts of the study region is presented in section 4.7. The second reason to apply the constraints is that

they reduce the extent of model space, which both reduces the computational burden and also yields error estimates that more closely reflect our belief in the estimated model.

As described by Shen et al. (2013b), the Bayesian Monte Carlo joint inversion method constructs a prior distribution of models at each location defined by allowed perturbations relative to the reference model as well as the model constraints described above. Examples of prior marginal distributions for crustal thickness at the six example stations are shown as white histograms in Figure 4. The nearly uniform distribution of the prior illustrates that we impose weak prior constraints on crustal thickness.

4.3.2 Joint Monte Carlo inversion and the posterior distribution

Once the data are prepared and the prior model space is determined, we follow Shen et al. (2013b) and perform a Markov Chain Monte Carlo process to compute the posterior distribution. At each location, we consider at least 100,000 trial models in which the search is guided by the Metropolis algorithm. Models are accepted into the posterior distribution or rejected according to the square root of the reduced χ^2 value. A model m is accepted if $\chi(m) > \chi_{min} + 0.5$, where χ_{min} is the χ value of the best fitting model. After the inversion, the misfit to the Rayleigh wave dispersion curve has a χ_{min} value less than 1 for all the stations. Discussion of the fit to receiver function data is presented later in section 4.7 in the context of determining whether the piecewise smooth parameterization that we use is sufficient to model receiver functions.

The principal output of the joint inversion at each station is the posterior distribution of models that satisfy the receiver function and surface wave dispersion data within tolerances that depend on the ability to fit the data and data uncertainties as discussed in the preceding paragraph. The statistical properties of the posterior distribution quantify model errors. In particular, the mean and standard deviation (interpreted as model uncertainty) of the accepted model ensemble are computed from the posterior distribution at each depth within the model. The posterior distribution both represents information from the data and prior constraints imposed on the inversion.

Figure 4.4 presents posterior distributions for crustal thickness for six example stations as red histograms. Compared with the prior distributions (white histograms), the posterior distributions narrow significantly at five of the six stations, meaning that at these stations crustal thickness is fairly tightly constrained ($\sigma < 2$ km) with a clear Moho Ps conversion in the receiver function

(Fig. 4.3). The exception is station SPMN (Fig. 3b) in the northern MCR where there is a weak Moho Ps conversion, but model uncertainty increases proportionally ($\sigma > 5$ km). In the six examples presented in Figure 4, crustal thickness ranges from about 35 to 48 km. Over the entire region of study, crustal thickness has a mean value of 44.8 km and an average 1σ uncertainty of about 3.3 km.

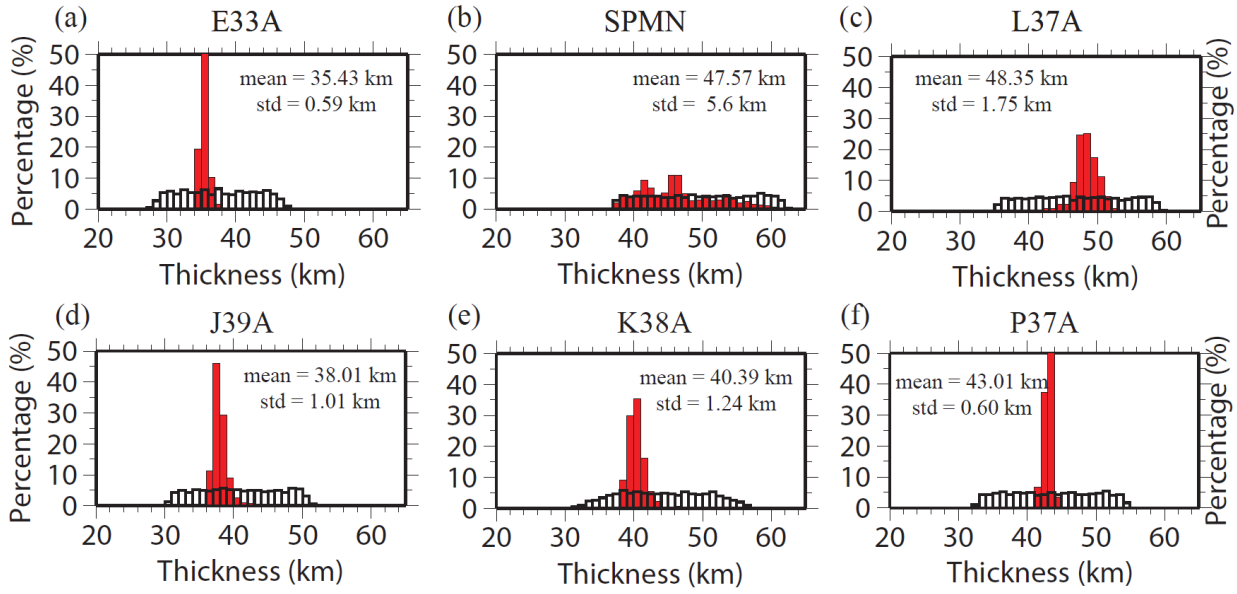

Figure 4.4 Prior and posterior distributions of crustal thickness for six example stations. (a): White histograms are the percentage distribution of the prior information for crustal thickness beneath station E33A. The red histogram centered at 35.43 km with 1 σ = 0.59 km represents the posterior distribution after the Bayesian Monte-Carlo inversion. (b-f): Same as (a), but for stations SPMN, L37A, J39A, K38A and P37A, respectively.

Inversion results for the six example stations are shown in Figure 4.5. The clear Moho with small depth uncertainty at station E33A reflects the strong Moho Ps signal in the back-azimuth averaged receiver function (Fig. 4.3a). Both the Rayleigh wave dispersion and the receiver function are fit well at this station (Fig. 4.3a). In contrast, at station SPMN (Fig. 4.5b) a gradient Moho appears in the model because the receiver function does not have a clear Moho Ps conversion (Fig. 4.3b).

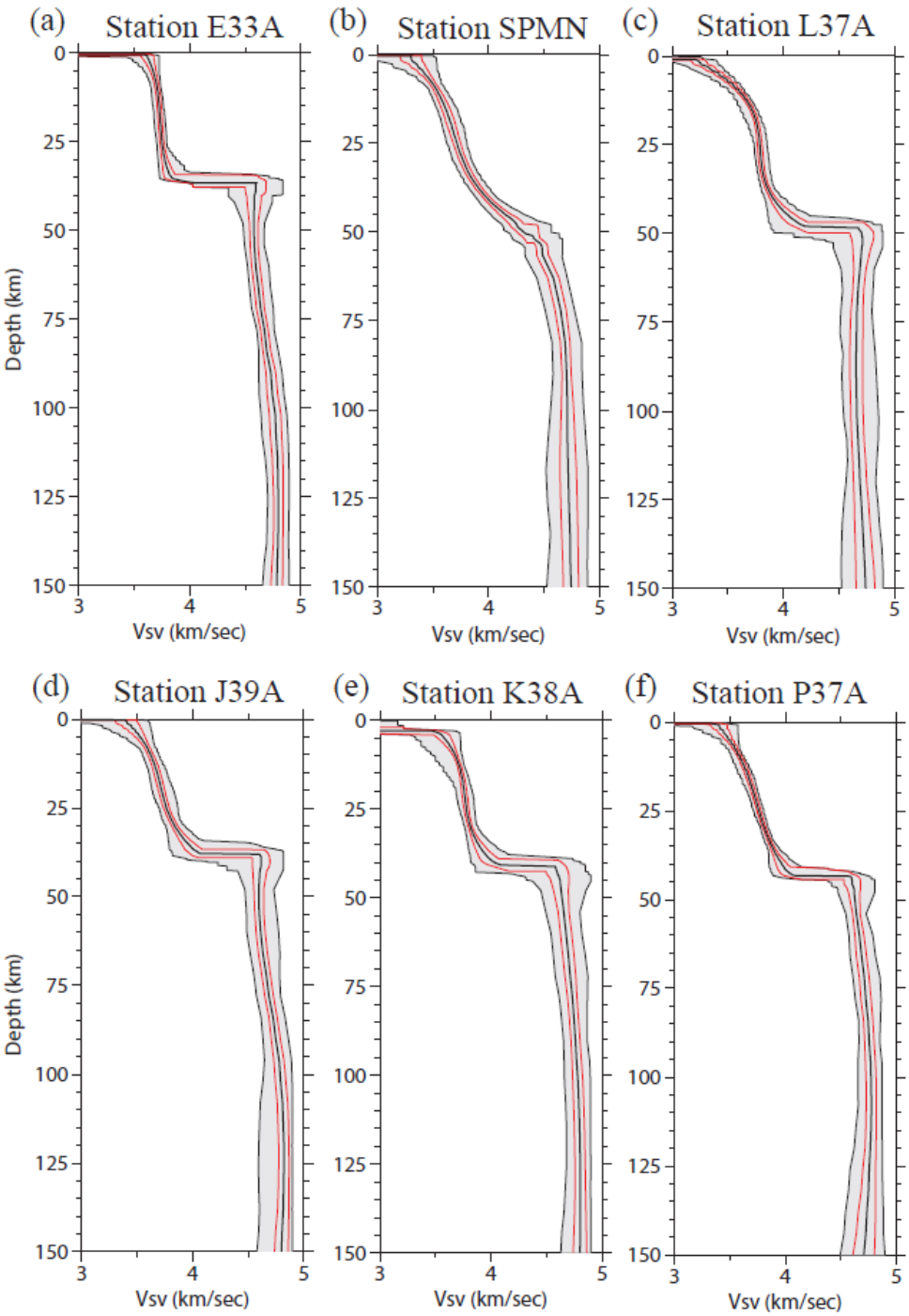

Figure 4.5 Resulting model ensembles that fit both Rayleigh wave and receiver function data for the six example stations of Figs. 4.3 and 4.4. (a) The resulting model ensemble for station E33A. The average of the posterior distribution is shown as the black line near the middle of the grey corridor, which defines the full width of the posterior distribution at each depth. The red lines represent the 1σ width of the distribution. (b-f) Same as (a), but for stations SPMN, L37A, J39A, K38A and P37A, respectively.

The resulting models for the other four stations (L37A, J39A, K38A, P37A) are shown in Figure 4.5c-f and the fit to the data is shown in Figure 4.3c-f. Station L37A is located near the center of the southern MCR. The receiver functions computed at this station show a relatively strong Moho Ps conversion at ~ 6 sec after the direct P arrival, indicating a sharp Moho discontinuity at ~ 50 km depth with uncertainty of about 1.75 km. For station J39A in northeastern Iowa, the clear Ps conversion at ~ 5 sec indicates a much shallower Moho discontinuity at ~ 40 km with an uncertainty of about 1 km. For station K38A near the eastern flank of the southern MCR, strong reverberations in the receiver function indicate the existence of thick sediments but there is also a clear Moho Ps arrival. Finally, a clear Moho with uncertainty less than 1 km is seen beneath station P37A in the Mazatzal province.

We perform the joint inversion for all 122 TA stations in the region of study and construct a mean 1-D model with uncertainties below each station. We then interpolate those 1-D models onto a regular 0.25° × 0.25° grid by using a simple kriging method in order to construct a 3-D model for the study region (Shen et al. 2013a). Simple kriging is only one possible method to interpolate the model between stations and may be optimal only in simplicity and computational efficiency. In simple kriging, model interpolation between stations is guided exclusively by the posterior distribution at all neighboring stations. Another possibility would be to use the surface wave phase speed maps to guide the interpolation because they exist on a grid finer than the station spacing. In this case, one would re-perform the inversion at the interstation locations using only the phase speed information but with a prior distribution determined from the posterior distributions at the neighboring stations. We effected a simple version of this algorithm and found that differences relative to the kriging method are mostly subtle and largely confined to the shallow crust where the amplitude but not the distribution of features are increased somewhat. Therefore, this alternative method of interpolation does not present first-order changes in the model relative to the kriging method or affect the conclusions of this chapter. We believe further development efforts of methods like this are advisable because of the general importance of assimilating data in inversions that may differ in type, grid location, and grid spacing, but such a method may be most useful in regions where inter-station spacing is larger than the 70 km that characterizes the USArray TA.

Maps of the 3-D model for various model characteristics are shown in Figures 4.6-4.8. Figure 4.6 presents map views of the 3-D model within the crust: average thickness and Vs of the sedimentary layer (Fig. 4.6a,b, respectively), Vs at 10 km depth (Fig. 4.6c), middle crust defined as the average in the middle 1/3 of the crystalline crust (Fig. 4.6d), and lower crust defined as the average from 80% to 100% of the depth to the Moho (Fig. 4.6e). Moho depth, uncertainty in Moho depth, and the Vs contrast across the Moho (the average difference between Vs in the uppermost 1 km of the mantle and lowermost 1 km of the crust) are shown in Figure 7a-c. Deeper structures in the mantle are presented in Figure 8 with Vs maps at 80 km depth (Fig. 4.8a) and 120 km (Fig. 4.8b). Three vertical slices that cross the MCR are shown in Figure 4.9 along profiles identified as A-A', B-B' and C-C' in Figure 8b. The model is discussed in more detail in section 4.4. Although the 3-D model extends to 200 km below the surface of the earth, the Vs uncertainties increase sharply with depth below 150 km due to the lack of vertical resolution. Therefore, we only discuss the top 150 km of the 3-D model.

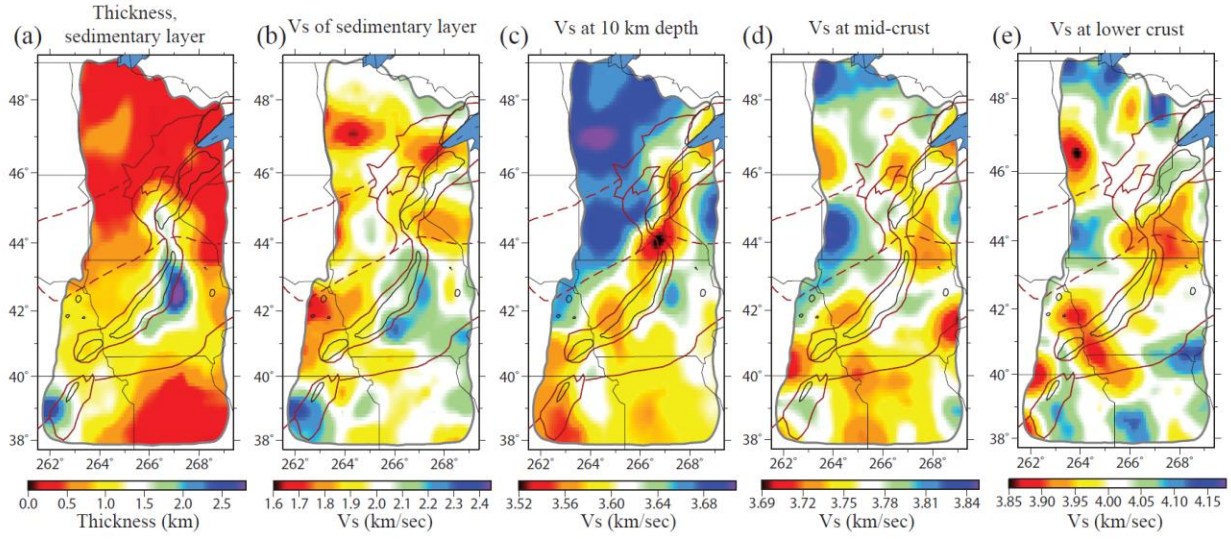

Figure 4.6 Maps of the mean of the posterior distribution for crustal structure of the study area. (a) Sedimentary thickness and (b) average shear wave speed. (c-e): Maps of Vsv at 10 km depth, in the middle crust (averaged in the middle 1/3 of the crystalline crust), and in the lower crust (averaged for 80% - 100% of crustal depth), respectively.

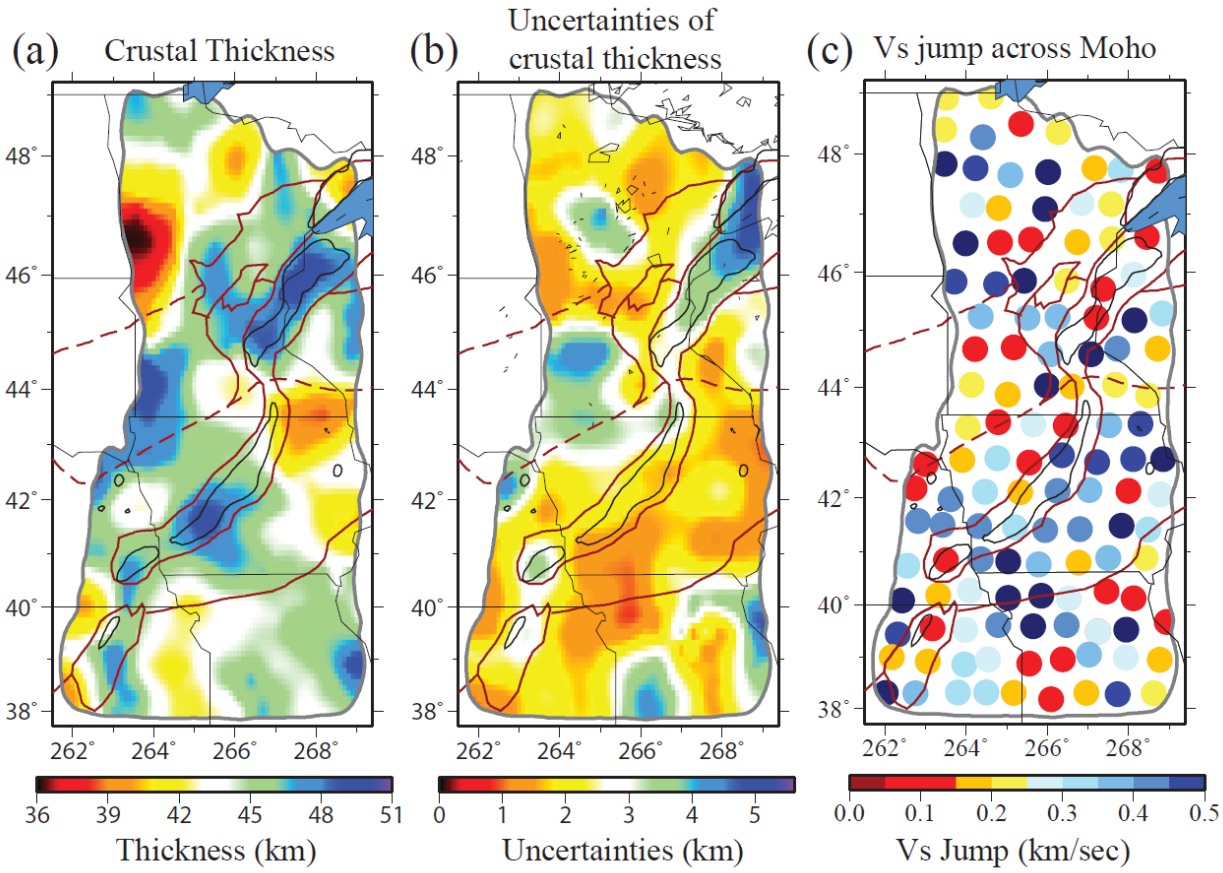

Figure 4.7 (a) Map of the mean of the posterior distribution for (a) crustal thickness. (b) Map of the standard deviation of the posterior distribution for crustal thickness, interpreted as its uncertainty. (c) Mean of the posterior distribution for the shear velocity contrast across Moho (Vsv difference between the uppermost 1 km of the mantle and the lowermost 1 km of the crust) plotted at the station locations.

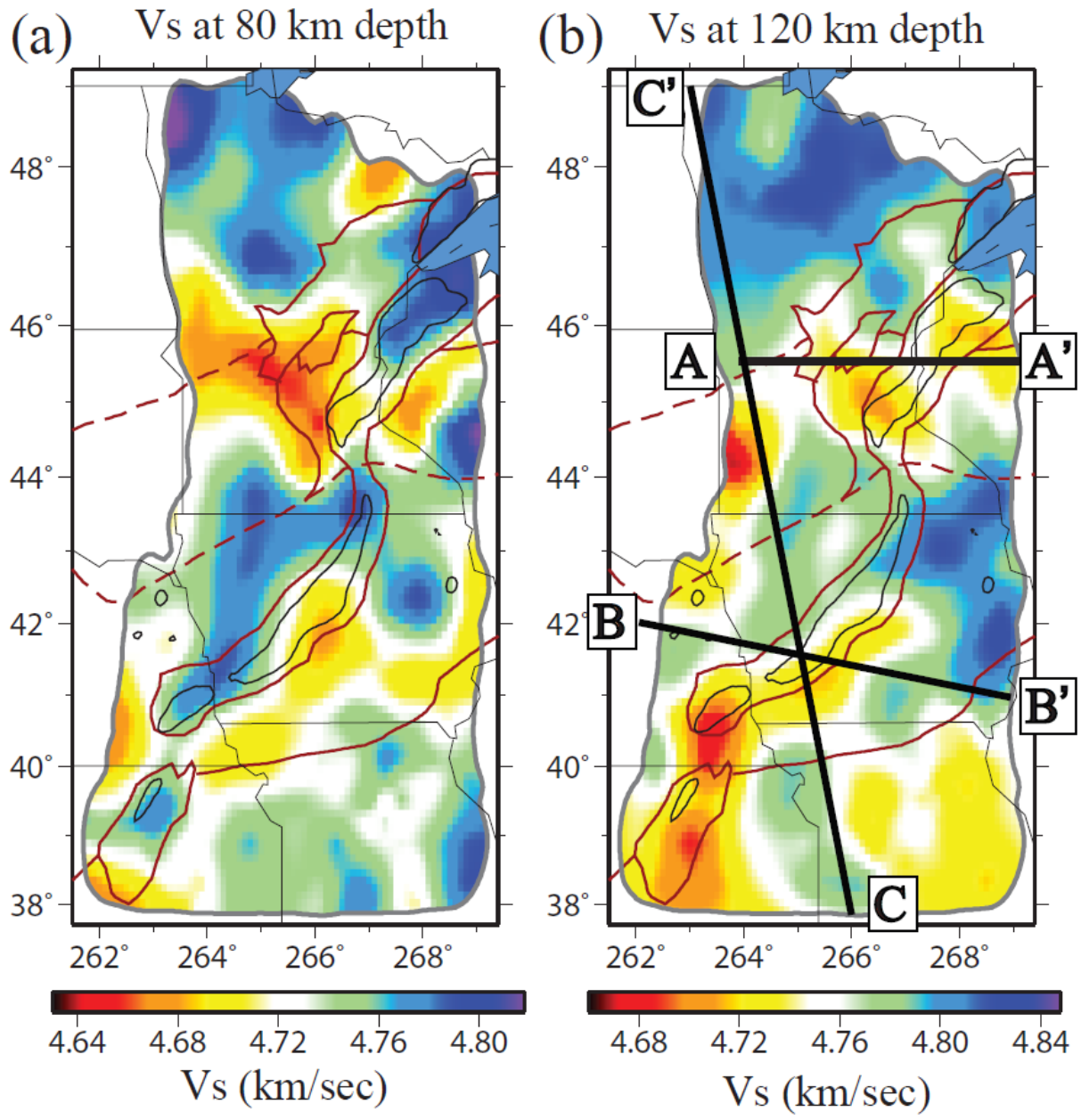

Figure 4.8 Maps of the mean of the posterior distribution for uppermost mantle Vsv at (a) 80 km depth and (b) 120 km depth. The black lines labeled A-A', B-B', and C-C' indicate the locations of the three vertical transects presented in Fig. 4.9.

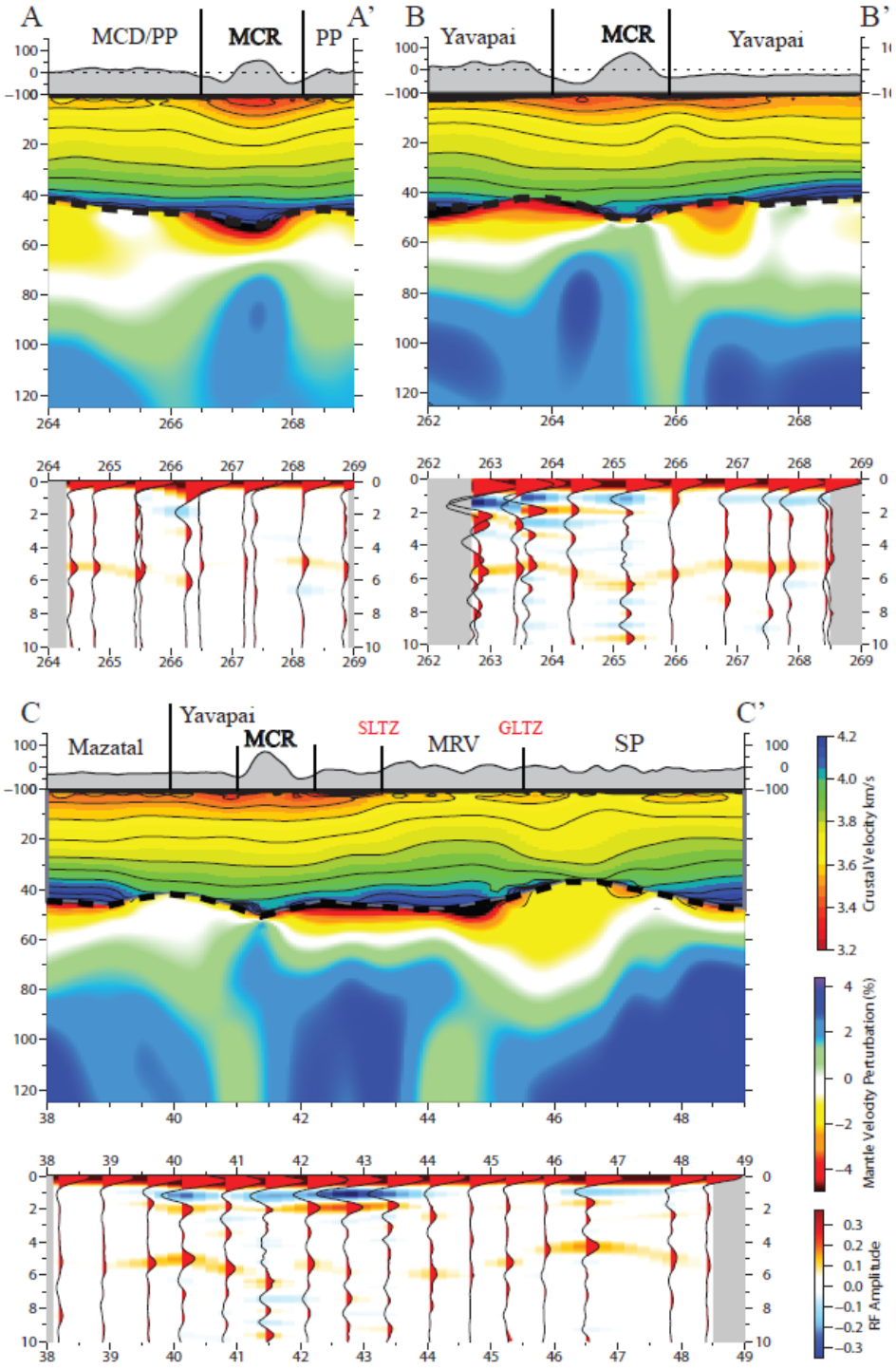

Figure 4.9 Vertical transects of the means of the posterior distribution of Vsv along profiles A-A', B-B', and C-C' whose locations are identified in Fig. 4.8b. In the upper panel of each pair, absolute shear velocity (km/s) is shown in the crust, the Moho is identified by the thick dashed line, and percent perturbation relative to 4.65 km/s is presented in the mantle. In the lower panel of each pair, receiver functions at stations close to the transects are shown as black waveforms and filled with warm color for positive amplitudes and cool colors for negative amplitudes.

4.4 Results and Discussion

4.4.1 Sedimentary layer

The sedimentary layer structure is shown in Figure 4.6a,b. Thick sediments (> 2 km) are observed near the eastern flank of the southern MCR, thinning southward. Another thick sedimentary layer appears near the southern edge of the MCR in Kansas. In the rest of the area, the sediments are relatively thin (<1 km). Notably, sediments are thin in the region between the northern and southern MCR even though short period Rayleigh wave phase speeds are slow (Fig. 2a). This inference is guided by the receiver functions at nearby stations, which are inconsistent with strong sediments in this area. However, the resulting sedimentary distribution may be spatially aliased due to the high lateral resolution of the receiver functions (< 5 km) with a low spatial sampling rate from the station locations (~70 km). For this reason interpretation of data from the Superior Province Rifting Earthscope Experiment (SPREE) experiment (Stein, 2011) is needed to more fully illuminate the shallow structure in this region.

The receiver functions do indicate the existence of sediments with particularly low shear wave speeds in some areas. For example, strong reverberations observed in the receiver function for station E33A in the first 2 sec may be fit by a Vs model with a thin (< 0.5 km) but slow Vs layer (< 1.8 km/sec) near the surface (Fig. 4.3a). Figure 4.6b shows the pattern of the inferred Vs in the sedimentary layer, which differs from pattern of sedimentary thickness. Very slow sedimentary shear wave speeds are found in northern Minnesota, which may be due to the moraine associated with the Wadena glacial lobe (Wright, 1962). Some of the slow sediments generate strong reverberations in the receiver functions that coincide in time with the Moho signal, resulting in large uncertainties in the crustal thickness map (Fig. 4.7b). At some other stations, sedimentary reverberations do not obscure the Moho Ps arrival; e.g., K38A (Fig. 4.3e). Sedimentary reverberations in the receiver functions can also be seen in Figure 9 beneath the Yavapai Province in transects B-B' and C-C', beneath the southern MCR in transect C-C', and north of the southern MCR in transect C-C'.

4.4.2 Shallow crystalline crust and correlation with the observed gravity field

The MCR gravity high (40 mgal anomaly outlined in the free air gravity map of Fig. 4.1) is poorly correlated with the shear velocity anomalies presented in Figures 4.6-4.8. Because positive density anomalies should correlate to positive velocity anomalies (e.g., Woollard, 1959; Brocher, 2005), the expectation is that high velocity anomalies underlie the MCR or the crust is thin along the rift. In fact, the opposite is the case. At 10 km depth, low velocity anomalies run beneath the rift and, on average, the crust is thickened under the rift. Our 3D model, therefore, does not explain the gravity high that runs along the MCR. There are two possible explanations for this. First, the high-density bodies that cause the gravity high may be too small to be resolved with surface wave data determined from the station spacing presented by the USArray. Second, small high shear wave speed bodies that cause the gravity high may be obscured by sediments in and adjacent to the rift. We believe the latter is the more likely cause of the anti-correlation between observed gravity anomalies and uppermost crustal shear velocity structure beneath the rift. If this is true, however, the high-density bodies that cause the gravity high probably would be in the shallow crust, or else they would imprint longer period maps that are less affected by sediments. This is consistent with the study of Woelk and Hinze (1991) who argue that the uppermost crust beneath the MCR contains both fast igneous rocks and slow clastic rocks. Under this interpretation, shallow igneous rocks must dominate the gravity field while the clastic rocks dominate the shear wave speeds. A shallow source for the gravity anomaly is also supported by the observation that the eastern arm of the MCR, which is buried under the Michigan Basin, has a much weaker gravity signature than the western arm imaged in this study (Stein et al., 2011).

The 3D shear velocity model is better correlated with the longer wavelengths in the gravity map (Fig. 4.1), which displays a broad gradient across the region (von Frese et al., 1982). The free-air gravity southeast of the MCR is lower (-30 mgal) than in the northwestern part of the map (10-20 mgal). It has been argued that this gradient is not due to variations in Precambrian structure across the sutures (Hinze et al., 1992), but may be explained by a density difference in an upper crustal layer. Our results support an upper crustal origin because the correlation of high shear velocities with positive long wavelength gravity anomalies exists primarily at shallow depths. At 10 km depth, which is in the uppermost crystalline crust (Fig. 4.6c), the most prominent shear velocity feature is a velocity boundary that runs along the western flank of the MCR. This follows the Minnesota River Valley Province-Yavapai Province boundary in the west and the northeastern edge of the Craton Margin Domain in the east. North of this boundary, Vs is

between 3.65 and 3.7 km/sec in the southern Superior Province, while to the south it decreases to between 3.5 and 3.6 km/sec in the Minnesota River Valley, Yavapai Province and Mazatzal Province. This boundary lies near the contrast in free air gravity. Similar features do not appear deeper in the model (Figs. 4.6d,e, 7).

Because receiver functions are sensitive to the discontinuity between the sediments and the crystalline crustal basement, the commonly unresolved trade-off between crustal structure and deeper structure in traditional surface wave inversions (e.g., Zheng et al., 2010; Zhou et al., 2011) is ameliorated in the model we present here and smearing of sedimentary velocities into the crystalline crust (e.g., Fig 4.6c) should not be strong. The observed high correlation of Vs with the long-wavelength gravity field supports this conclusion. Therefore, we believe that the very slow anomaly (< 3.5 km/sec) that is observed in the gap between the northern and southern MCR at 10 km depth (Fig. 4.6c) does in fact reside in the crystalline crust and is not a smearing effect of sediments to greater depth. It is not clear to us, however, why such low shear wave speeds appear in the upper crust at this location.

Finally, there are also correlations between shallow Vs structure and short wavelength gravity anomalies. In the gravity map, the lowest amplitudes appear near station K38A on the eastern flank of the southern MCR where thick sediments are present in the model (Fig. 4.5e). Thus, local gravity minima may be due to the presence of local sediments at this point.

4.4.3 Relationship between Precambrian sutures and observed crustal structures

4.4.3.1 Great Lakes Tectonic Zone

In the northern part of the study region, the Great Lakes Tectonic Zone (GLTZ) suture lies between the 2.7 - 2.75 Ga greenstone terrane to the north and the 3.6 Ga granulite-facies granitic and mafic gneisses Minnesota River Valley sub-province to the south cutting the southern end of Superior Province into two sub-provinces (Morey and Sims, 1976). The eastern part of Great Lakes Tectonic Zone in our study region is covered by the Craton Margin Domain (CMD of Fig. 1), which contains several structural discontinuities (Holm et al., 2007).

Beneath this northernmost suture, a Vs contrast is observed in the 3-D model through the entire crust, with the contrast becoming stronger with depth. In the upper crust (Fig. 4.6c), Vs is ~ 3.7 km/sec beneath the Superior Province (SP) greenstone terrane and ~ 3.68 km/sec beneath the

Minnesota River Valley (MRV) with a relatively slow Vs belt beneath the eastern part of the suture. In the middle crust (Fig. 4.6d), the Vs contrast is stronger. A fast anomaly (> 3.8 km/sec) is observed beneath the MRV itself, perhaps indicating a more mafic middle crust, while in the north the SP is about 0.08 km/sec slower than the MRV. This difference across the Great Lakes Tectonic Zone strengthens with depth to about 0.15 km/sec in the lowermost crust (Fig. 4.6e).

These variations in crustal structure are also reflected in Moho depth, which is discussed further in section 4.4. North of the GLTZ, a clear, large-amplitude Moho signal is seen as early as 4.3 sec (Fig. 4.3a), although the receiver functions at some stations display large reverberations from the thin slow sediments. Combined with relatively fast phase velocities observed at 28 sec period in this area, the inversion yields a relatively shallow Moho at about 36 km depth at station E33A and neighboring points. To the south of the GLTZ, thicker crust is found in the MRV with an average crustal thickness of about 46 km with a maximum thickness of about 48 km. The average uncertainties of crustal thickness in the MRV are greater than 3 km suggesting that the Moho is more of a gradient than a sharp boundary (Fig. 4.7c). A seismic reflection study in this area (Boyd and Smithson, 1994) reveals localized Moho layering probably due to mafic intrusions related to post-Achaean crustal thickening events in this area. Our large Moho depth and fast middle to lower crust (Fig. 4.6d,e) are consistent with this interpretation.

4.4.3.2 Spirit Lakes Tectonic Zone

The boundary between the Superior (SP) and the Yavapai (YP) provinces is the Spirit Lakes Tectonic Zone (SLTZ), which extends east through the middle of the MCR into Wisconsin. East of the MCR, the SLTZ separates the Penokean Province to the north from the Yavapai Province to the south. As described in section 4.2, west of the MCR this suture forms a boundary within the upper crystalline crust that correlates with the transition from mild gravity highs to the north to gravity lows to the south. The Yavapai province is the region with thickest sediments across the region. Structural differences between the two provinces across the suture are particularly striking at 10 km depth but continue into the lower crust, with faster Vs in the Minnesota River Valley subprovince and slower Vs in the Yavapai. In terms of Moho topography, the Yavapai Province has relatively thinner crust (~ 44-45 km) than the Minnesota River Valley, particularly east of the southern MCR (~ 39 km). For example, the receiver function at station J39A (Fig. 4.3d) displays a clear Moho Ps conversion at about 4.5 sec after the direct P arrival. The resulting

model for station J39A is shown in Figure 4.4d with a crustal thickness of about 38±1.5km. This is the thinnest crust in the vicinity of the rift, but is still deeper than in the Greenstone terrane in the western part of the Superior Province.

4.4.3.3 Boundary between the Yavapai and Mazatzal provinces

The third and southernmost suture in the study region is the boundary between the Yavapai (YP) and Mazatzal provinces (MP) near the Iowa-Missouri border, extending in the NE-SW direction. Compared with the structural variations across the more northerly sutures, the variations across this suture are more subtle both in crustal velocities and crustal thickness. However, lower crustal Vs (Fig. 4.6e) is slower (< 4 km/sec) in the YP than it in the MP (> 4 km/sec) and the velocity jump across Moho (Fig. 4.7c) is larger in the YP, on average.

In summary, the three major Precambrian sutures in the region are associated with crustal seismic structural variations, especially across the northern (GLTZ) and middle (SLTZ) sutures in the MCR region. Later cumulative metamorphism of early Proterozoic accretionary tectonics (Holm et al., 2007) may have obscured structural variations across the Yavapai – Mazatzal boundary at least in the shallow crust.

4.4.4 Variations in crustal thickness

A clear and profound difference between average crustal and mantle shear wave speeds is observed across the entire region of study, as Figures 6d and 8a show. Average mid-crustal shear wave speeds are about 3.76 km/s whereas at 80 km depth they average about 4.72 km/s, nearly 1 km/s higher. However, we seek to determine crustal thickness where and if a jump in velocity occurs between the crust and uppermost mantle. The inference of this jump is guided principally by the receiver functions. An advantage of the joint inversion of surface wave dispersion and receiver functions is the amelioration of trade-offs that occur near structural discontinuities such as the base of the sediments and the Moho, which hamper inversion of surface wave data alone. As argued by Shen et al. (2013a,b) and many others (e.g., Bodin et al., 2012; Lebedev et al., 2013), estimates of depth to Moho as well as the velocity contrast across it are greatly improved. But, a clear Moho Ps converted phase is not observed at every station. Where it is observed, there is evidence for a jump of seismic velocities between the crust and uppermost mantle and we can estimate crustal thickness with considerable accuracy. Where this phase is not observed (e.g., Fig. 4.3b), we have evidence that the transition between crustal and mantle shear wave

speeds is probably gradual, which we refer to as a gradient Moho (e.g., Fig. 4.5b). In these places, crustal thickness is not well defined which may result from underplating or interleaving of crustal and mantle rocks in a finite Moho transition zone. In both instances whether we observe or do not observe a clear Moho Ps conversion on the receiver function, our estimate of the uncertainty in crustal thickness reflects our knowledge (e.g., Fig. 4.7b).

Therefore, our discussion of crustal thickness begins with an assessment of where we have definite information from receiver functions for a discrete jump in seismic wave speeds from the crust to the mantle and, hence, can estimate crustal thickness accurately. We present this information in Figure 7c as an estimate of the jump in shear wave speeds from the crust to the uppermost mantle. Stations at which this jump is below about 0.15 km/s (colored red in Fig. 4.7c) are located where no Moho Ps conversion is seen and, hence, were there is a gradient Moho. Stations where the jump is larger than about 0.25 km/s (colored blue in Fig. 4.7c) represent a clear sharp Moho. Stations underlain by a gradient Moho are scattered throughout the study region and number 22 of the 122 stations we used (or just less than 20% of the region). The locations of these stations are highly correlated with estimates of large uncertainty in crustal thickness (Fig. 4.7b).

We make two observations about the location of a gradient or a sharp Moho across the region. First, a sharp Moho is observed beneath the southern MCR but not beneath the northern MCR. Uncertainties in crustal thickness for the northern segment of the MCR are larger (> 4km) than for the southern segment (<2 km), as Figure 7b shows. Between the northern and southern segments, there is a shallow Moho (< 42 km) that extends eastward to the eastern Penokean Orogen and perhaps further east. Second, a sharp Moho is observed across much of the Yavapai province whereas a gradient Moho is seen beneath much of the Minnesota River Valley province, which is consistent with a previous reflection seismic survey in the area (Boyd and Smithson, 1994).

Figure 4.7a presents the resulting map of Moho topography and shows that the MCR is estimated to have a deep Moho (>47 km, peaking at ~50 km) in all three segments (Wisconsin/Minnesota, Iowa, Nebraska/Kansas), although crustal thickness in the northern MCR is not well determined because it is a gradient Moho. The crust beneath the MCR is about 5 km thicker, on average, than crustal thickness averaged across the study region. For the northern MCR, the crustal

thickening mostly occurs within the gravity anomaly and extends to the northeastern edge of the Craton Margin Domain. In the southern MCR, crustal thickening is not uniform along the rift but is most pronounced in the southern half of this segment. For the Nebraska/Kansas segment, thickened crust (> 47 km) is also present, which is consistent with a previous reflection study for this area (Woelk and Hinze, 1991), although the northernmost part of this small region appears to have a gradient Moho. Analyzing older data, Moidake et al. (2013) also measured crustal thickness using Ps information and inferred locally thick crust (~53 km thick) beneath a narrow section of the southern MCR near 41.5N, 94.5W, which is near the area with the thickest crust we find beneath the southern MCR. (Fig. 4.7a).

Notable crustal thickness variations are observed throughout the rest of the study area as well: a significantly thinned crust is seen near the western border of Minnesota within the Superior Province, which changes to a thick crust with a gradient Moho at about 50 km depth in the Gneiss Terrane of the Minnesota River Valley to the south. Another gradient Moho is observed north of the Great Lakes Tectonic Zone in the Superior Province, Further south, crustal thickness lies between 42 km and 46 km in the Mazatzal Province.

Three transects (identified as A-A', B-B', C-C' in Fig. 4.8b) across the MCR are presented as pairs of panels in Figure 4.9. In the top half of each pair, absolute Vs in the crust beneath the three transects is shown with 0.1 km/sec contours outlined by black lines and the Moho identified by a thick dashed line. In the mantle, Vs is shown as the percent perturbation relative to 4.65 km/sec. Transects A-A' and B-B' cut the northern and southern segments of the MCR, respectively, and transect C-C' cuts across the study region in the N-S direction and intersects with transect B in the southern MCR. In the lower half of each pair of panels of Figure 4.9, receiver function waveforms are shown for stations within a distance to each transect of 0.4°. We observe in the receiver functions two major features beneath the MCR. (1) As noted previously, Moho Ps conversion across the northern MCR (A-A') is obscure. (2) There is a clearer Moho Ps conversion at ~6 sec for the southern MCR (B-B' and C-C'). As a result of the gradient Moho beneath the northern MCR, crustal thickness is poorly determined (1 σ uncertainty > 4 km), lowermost crustal wave speeds are fast (> 4 km/sec), and uppermost mantle wave speeds are slow (< 4.4 km/sec). It is possible that this layer results from magmatic intrusion or underplating (Furlong and Fountain, 1986). However, the underplating cannot be continuous along the entire MCR, because beneath the southern MCR this intermediate-velocity Vs layer is not present. In

the adjacent area, another gradient Moho feature is seen beneath the Minnesota River Valley, with a Moho Ps conversion in the receiver functions that is weaker than those in the Superior or Yavapai Provinces (transect C-C'). This is consistent with a seismic reflection study in this sub-province (Boyd and Smithson, 1994) where Moho layering has been inferred due to mafic intrusion in the lower crust. The other features seen in these transects include the relatively thin crust (~ 40 km) near the flanks of the MCR (e.g., SMCR-Yavapai boundary) and in the southern Superior Province north of the Great Lakes Tectonic Zone. The latter region is the thinnest crust across the area of study (<38 km) and the cause of this thinning deserves further investigation.

4.4.5 Evidence that the MCR is a compressional feature

Currently active rifts such as the East African Rift (e.g., Braile et al., 1994; Nyblade and Brazier, 2002), the Rio Grande Rift (e.g., West et al., 2004; Wilson et al., 2005; Shen et al., 2013a)), the West Antarctic Rift (Ritzwoller et al., 2001), and the Baikal rift (Thybo and Nielsen, 2009) as well as hot spots (e.g., Snake River Plain; e.g, Shen et al., 2013a) show crustal thinning. At some locations the thinned crust has been rethickened by mafic crustal underplating; for example, the Baikal rift (Nielsen and Thybo, 2009) and perhaps also the Lake Superior portion of the MCR (Cannon et al., 1989). Although thermal anomalies dominantly produce low Vs in the mantle underlying active rifts (e.g., Bastow et al., 1998), compositional heterogeneity in the crust due to mafic underplating and intrusions can overcome the thermal anomaly to produce high crustal wave speeds even in currently active regions. After the thermal anomaly has equilibrated, as it has had time to do beneath the MCR, high crustal wave speeds would be expected. In actuality, we observe a thickened and somewhat slow crust within the MCR. We discuss here evidence that the observed crustal characteristics reflect the compressional episode that followed rifting (Cannon, 1994).

The presence of low velocities in the upper and middle crust and crustal thickening beneath the MCR has been discussed above (e.g., Figs. 4.6-8). Figure 4.9 presents vertical profiles along with receiver functions profiles shown for reference. Transect AA', extending from the Superior Province to the Penokean Province, illustrates that the upper crust beneath the rift is slightly slower than beneath surrounding areas and the crust thickens to about 50 km. In the upper and middle crust, lines of constant shear wave speed bow downward beneath the northern MCR, but this is not quite as clear in the southern MCR as transects B-B' and C-C' illustrate. The gradient

Moho beneath transect A-A' appears as lower Vs in the uppermost mantle in Transect A-A'. The sharper Moho beneath transects B-B' and C-C' appears as higher Vs in the uppermost mantle.

These observations of a vertically thickened crust with the downward bowing of upper crustal velocity contours contradict expectations for a continent rift. They are, in fact, more consistent with vertical downward advection of material in the crust, perhaps caused by horizontal compression and pure shear thickening. Geological observations and seismic reflection studies in the region also indicate a compressional episode occurring after rifting along the MCR. (1) Thrust faults form a horst-like uplift of the MCR, showing crustal shortening of about 20 to 35 km after rifting (Anderson, 1992; Cannon and Hinze, 1992, Chandler et al., 1989; Woelk and Hinze, 1991). (2) Uplift evidence from anticlines and drag folds along reverse faults are also observed (Fox, 1988; Mariano and Hinze, 1994a). The horst-like uplift combined with reverse faults have been dated to ca. 1060 Ma (Bornhorst et al. 1988; White, 1968; Canno and Hinze,1992), which is about 40 Ma after the final basalt intrusion (Cannon, 1994). (3) Seismic reflection studies show a thickened crust beneath certain transects (Lake Superior: Cannon et al., 1989; Kansas: Woelk and Hinze, 1991).

In summary, we argue from our 3-D model combined with these other lines of evidence that the present-day MCR is a compressional feature in the crust. The compressive event thickened the crust beneath the MCR and may have advected material downward in the crust. More speculatively, rifting (ca 1.1 Ga) followed by compression may have weakened the crust, which allowed for the extensive volcanism in the neighboring Craton Margin Domain that appears to have occurred in response to continental accretion to the south (Holm et al., 2007).

A potential alternative to tectonic compression as a means to produce crustal thickening beneath the MCR may be magmatic underplating that occurred during the extensional event that created the rift (e.g., Henk et al., 1997). Although the gradient Moho that is observed beneath parts of the northern MCR may be consistent with magmatic underplating, the clear Moho with the large jump in velocity across it in the southern MCR is at variance with underplating. The general absence of high velocity, presumably mafic, lower crust also does not favor magmatic intrusions into the lower crust. Thus, although magmatic underplating cannot be ruled out to exist beneath parts of the MCR, particularly in the north, it is an unlikely candidate for the unique cause of

crustal thickening along the entire MCR. In addition, it cannot explain the downward bowing of shear velocity isolines in the upper and middle crust.

4.4.6 Uppermost mantle beneath the region

Not surprisingly for a region that has not undergone tectonic deformation for more than 1 Ga, the upper mantle beneath the study region is seismically fast. The average shear wave speed at 100 km depth beneath the study region is 4.76 km/s. By comparison, at the same depth the upper mantle beneath the US west of 100°W is 4.39 km/s. The slowest Vs is about 4.62 km/sec at 80 km depth near the border of the east-central Minnesota Batholith. This is still faster than the Yangtze Craton (4.3 km/sec at 140 km depth, Zhou et al., 2012) or the recently activated North China Craton (~ 4.3 km/sec at 100 km, Zheng et al., 2011), but is similar to the Kaapvaal craton in South Africa (Yang et al., 2008a). The rms variation across the region of study is about 0.05 km/s, which is much less than the variation across the western US (rms of 0.18 km/s). Thus, mantle variability across the central US is small in comparison to more recently deformed regions.

Although upper mantle structural variation is relatively small across the study region, Figures 4.8 and 4.9 show that prominent shear velocity anomalies are still apparent. In general, the Vs structure of the uppermost mantle is less related to the location of the Precambrian provinces and sutures than is crustal structure. One exception is deep in the model (120 km, Fig. 4.8) where there is a prominent velocity jump across the Great Lakes Tectonic Zone. The principal mantle anomalies appear as two low velocity belts. One is roughly contained between the Great Lakes Tectonic Zone and the Spirit Lakes Tectonic Zone, and then spreads into the Penokean Province east of the northern MCR. The other extends along the southern edge of the Southern and the Nebraska/Kansas segments of the MCR, particularly at depths greater than 100 km. Beneath the MCR itself, shear wave speeds in the uppermost mantle are variable. The main high velocity anomaly exists beneath the Superior Province with the shape varying slightly with depth. This anomaly terminates at the Great Lakes Tectonic Zone, being particularly sharp at 120 km depth. The jump in velocity at the Great Lakes Tectonic Zone is seen clearly in transect C-C' (Fig. 4.9c).

There are three major factors that contribute to variations in isotropic shear wave speeds in the uppermost mantle: temperature, the existence of partial melt or fluids, and composition (e.g.,

Saltzer and Humphreys, 1997). The fast average Vs in the upper mantle compared with tectonic regions and recently rejuvenated lithosphere suggests the lack of partial melt. Similarly, velocity anomalies in the region probably do not have a tectonothermal origin because they have had time to equilibrate in the past 1.1 Ga. However, low velocity anomalies at greater depth may still reflect thinner lithosphere, which we speculate may be the case on the southern edge of the southern MCR. Nevertheless, the most likely cause of much of the variability in velocity structure in the uppermost mantle is compositional heterogeneity.

An alternative interpretation of the relatively low Vs is a lower depletion in magnesium in the mantle. Jordan (1979) argued that magnesium depletion will lower density but increase seismic velocities in the upper mantle. Thus, the lower wave speeds observed between the Great Lakes and Spirit Lakes Tectonic Zones may be due to less depleted material from the mantle rejuvenation that occurred during the rifting. Beneath the MCR near Lake Superior, basalts have been observed that were generated from a relatively juvenile mantle source (Paces and Bell, 1989; Nicholson et al., 1997), indicating the possible emplacement of less depleted material at shallower depth from the upwelling during the the rifting. This possible rejuvenation process may leave an enriched mantle remnant at depths greater than 100 km beneath the MCR and its surroundings (e.g., the craton margin domain), causing slower Vs compared to the rest of more depleted sub-cratonic lithosphere. Schutt and Lesher (2006), however, argued that mantle depletion would cause relatively little change in Vs in the upper mantle. Thus, the cause of the observed velocity variability in the uppermost mantle remains largely an open question that deserves further concerted investigation.

4.4.7 Effect of smoothness and monotonicity constraints on crustal structure

The results presented to this point and the interpretation that followed from them are based on the application of both velocity smoothness and monotonicity constraints in the crystalline crust. The earth may not be smooth or vary monotonically with depth so we consider here if there is evidence for either mid-crustal discontinuities or vertical velocity minima in the crust in the study region.

Evidence for mid-crustal discontinuities would come from the inability to fit the receiver functions with a continuous, smooth model between the base of the sediments and Moho. We find that we have significant trouble fitting the receiver function at only 4 of the 122 stations

used in this study: TA stations H37A, N36A, L37A, and K39A (identified in Fig. 4.1). In all four cases, we believe the receiver functions are good and reflect reliable information beneath each point. Figure 4.10 presents the two worst fit receiver functions at TA stations H37A and N36A. In both cases, a Moho Ps converted phase near 6 seconds (Fig. 10b,e) is well fit by the continuous, monotonic crustal model but there are oscillations in the receiver function at earlier times at both stations that are not fit (red lines Fig. 4.10b,e) with this parameterization (red lines in Fig. 4.10a,d). The introduction of a mid-crustal discontinuity (blue lines Fig. 10a,d) does allow the receiver functions to be fit acceptably (blue lines Fig. 4.10b,e). Three conclusions are worth drawing from these observations. First, because there are only four isolated stations at which receiver functions are not well fit by the continuous, monotonic crustal model there is no evidence that mid-crustal discontinuities are general features of any of the tectonic zones in the region. Second, the models that result from the simplified parameterization agree with the models parameterized with a mid-crustal discontinuity in crustal thickness, shallow crustal structure, and mantle structure. Therefore, the introduction of mid-crustal discontinuities does not change the principal conclusions or arguments of this chapter. Finally, discontinuous structures such as those shown by the blue line in Figure 10a are physically questionable. Further work is needed, presumably using data derived with a tighter station spacing (e.g., SPREE, Stein, 2011), which allows different receiver function stacking methods to be applied, to produce more realistic velocity profiles that fit the receiver functions at stations like H37A.

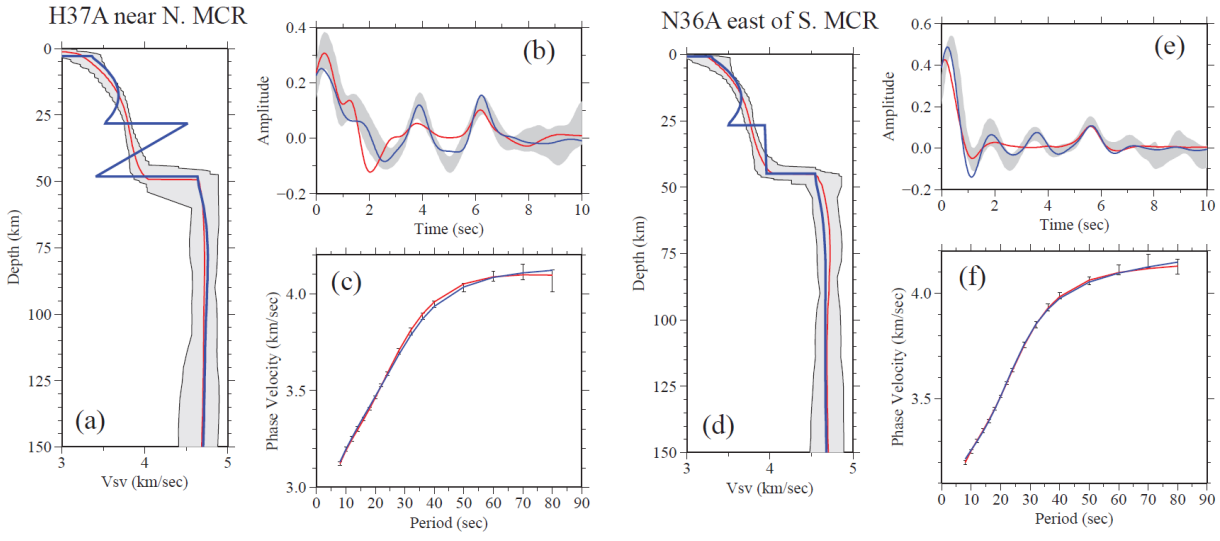

Figure 4.10 Comparison of models and data fit with and without a mid-crustal discontinuity at two (H37A, N36A) of the four stations needing the discontinuity. (a) & (d) Like Fig. 5 but where the model ensemble (grey shaded region) and best fitting model (red line) are compared with the best-fitting model containing the mid-crustal discontinuity. (b) & (e) Fit to the observed receiver function (grey shaded region) from the model with (blue line) and without (red line) the mid-crustal discontinuity. (c) & (f) Fit to the observed Rayleigh wave phase velocity curve (one standard deviation error bars) from the model with (blue line) and without (red line) the mid-crustal discontinuity.

We have also tested the effect of the monotonicity constraint by re-performing the inversion across the region without the constraint applied. Figure 4.11 presents a comparison between the models in the uppermost and lowermost crust with and without the imposition of the monotonicity constraint. In general, the results are very similar, although there is a tendency for the unconstrained model to have larger amplitudes particularly in the lower crust. The reason the monotonicity constraint has a relatively weak effect is because of our use of B-splines as basis functions in the crust. These splines automatically disallow local model excursions and the monotonicity constraint merely makes this explicit. With the introduction of more freedom to the parameterization, perhaps through layerization or adding more B-splines, both local velocity minima and maxima would appear, but they are not required to fit the data.

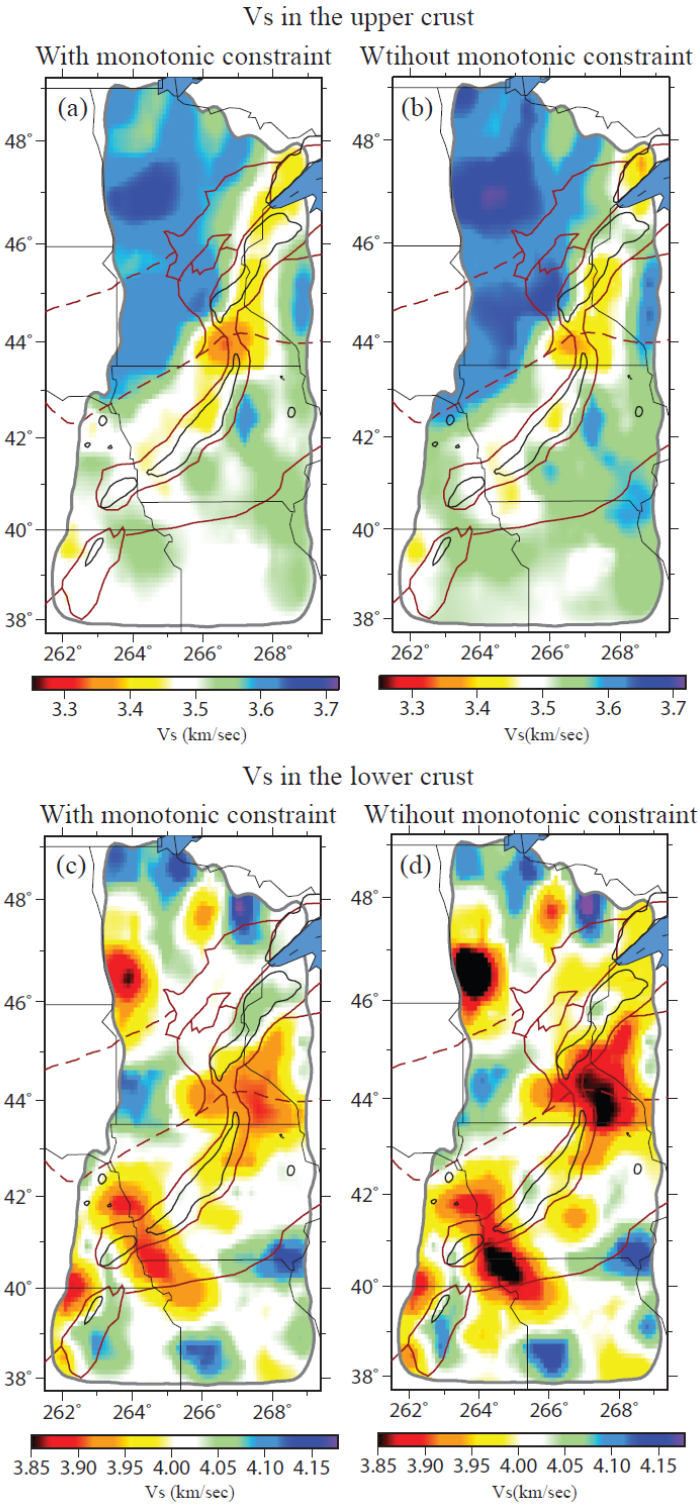

Figure 4.11 Comparison yof upper and lower crustal velocities from inversions in which the monotonic constraint either has or has not been imposed. Velocities are averaged across the upper and lower 20% of crustal thickness in each case, respectively.

4.5 Conclusions

Based on two years of seismic data recorded by the USArray/Transportable Array stations that cover the western arm of the Mid-Continent Rift (MCR) and its neighboring area, we applied ambient noise tomography using the eikonal tomography method and teleseismic earthquake tomography using the Helmholtz tomography method to construct Rayleigh wave phase velocity maps from 8 to 80 sec across the region. By performing a joint Bayesian Monte Carlo inversion of the phase velocity measurements with receiver functions, we construct posterior distributions of shear wave speeds in the crust and uppermost mantle from which we infer a 3D model of the region with attendant uncertainties to a depth of about 150 km. This model reveals three major features of the crust and uppermost mantle in this area.

First, the observed free air gravity field correlates with sediments and upper crustal structures in three ways. (1) A thick sedimentary layer contributes to the negative gravity anomalies that flank the MCR. (2) The slow upper crust at the gap between the northern and southern MCR masks the high gravity anomaly that runs along the rift. (3) Shear velocities in the uppermost crystalline crust are associated with a long wavelength gravity anomaly that is observed across the study area. However, our 3D model does not explain the existence of the gravity high along the rift because the crust beneath the MCR is seismically slow or neutral, on average. High-density anomalies must either be smaller than resolvable with our data or obscured by sediments. We believe the latter is the primary reason as the uppermost crust beneath the MCR probably contains both fast igneous rocks and slow clastic rocks such that shallow igneous rocks dominate the gravity field while the clastic rocks dominate the shear wave speeds.

Second, crustal thickening is found along the entire MCR, although along-axis variations exist and the gradient Moho in the northern MCR makes uncertainties large there. Analysis of local faults and seismic reflection studies in this area provide additional evidence for a compressional inversion of the rift and crustal thickening during the Grenville orogeny (French et al., 2009). Thicker crust and a deeper Moho cause a decrease in mid-crustal shear wave speeds and in Rayleigh wave phase velocities at intermediate periods (15-40 sec). Although the uppermost mantle beneath the MCR is faster than the average of the study region, velocity anomalies associated with the MCR are dominantly crustal in origin.

Third, the seismic structure of the crust, particularly the shallow crust, displays discrete jumps across the three major Precambrian sutures across the study region. This implies that although the Superior Greenstone Terrane in the north collided with the Minnesota River Valley more than 2 Ga ago, preexisting structural differences beneath these two subprovinces are preserved. Other sutures (e.g., Spirit Lakes Tectonic Zone, Yavapai/Mazatzal boundary) also represent seismic boundaries in the crust. The mantle beneath the entire region is faster than for cratonic areas that have undergone significant tectonothermal modification and lithospheric thinning (e.g., North China Craton), with the Superior Greenstone Terrane being the least affected by events of tectonism across the region.

In summary, the 3-D model presented here combined with other lines of evidence establishes that the MCR is a compressional feature of the crust. Presumably, the closing of the rift produced compressive stresses that thickened the crust beneath the MCR, advecting material downward in the crust under pure shear. The position of the slow thickened crust directly under the MCR suggests that crustal weakening during extension and subsequent thickening under compression occurred as pure shear (McKenzie, 1978), rather than under simple shear conditions. Simple shear would have resulted in a lateral offset between surface versus deep crustal features (Wernicke, 1985). Finally, since the MCR has been inactive for long enough that thermal signals associated with tectonic activity should have long decayed, our results provide a useful context for distinguishing between compositional and thermal influences on seismic velocities in active continent rifts (Ziegler and Cloetingh, 2004).

In closing, we note several topics for further research. (1) There is evidence at some stations for a mid-crustal discontinuity, which deserves further focused investigation. (2) The USArray TA data do not provide ideal inter-station spacing for receiver function analyses of the rift, and spatial aliasing of structures is possible. Finer sampling at select areas along the rift may appreciably improve the model. (3) Our model does not reveal structures deeper than about 150 km, which makes the determination of variations in lithospheric thickness difficult. (4) The physical cause of the low shear wave speeds in the uppermost crystalline crust (10 km depth) near the gap between the northern and southern MCR is unknown to us. (5) We also find mysterious the cause of very thin crust in the western Superior Province. (6) The spatial distribution of sharp versus gradient Moho across the region is fairly random and requires further

investigation into the cause of this variability. (7) In particular, the tectonic cause of the gradient Moho beneath the northern MCR compared to the sharp Moho beneath the southern is unclear to us. These issues call for further work with a denser seismic array, such as the SPREE array which has already been installed in this area (Stein et al., 2011), as well as the input of other types of geophysical data. Nevertheless, the 3-D model provides a synoptic view of the crust and uppermost mantle across the region that presents an improved basis for further seismic/geodynamic investigation of the MCR.

CHAPTER V

CRUSTAL AND UPPERMOST MANTLE STRUCTURE BENEATH THE UNITED STATES BY JOINT INVERSION OF SURFACE WAVE DISPERION, RECEIVER FUNCTIONS, AND RAYLEIGH WAVE ELLIPTICITY

Synopsis of the Chapter

In this chapter, I introduce a new seismic observable, the Rayleigh wave Horizontal-to-Vertical (H/V) ratio measured from teleseismic earthquakes, to the joint Monte Carlo inversion in three steps. First, this new observable is measured using over 1,700 USArray/TA stations, and a set of Rayleigh wave H/V ratio maps across the US continent are constructed. Second, I show that the measured H/V ratio can be interpreted along with surface wave dispersion and receiver functions using the joint Monte Carlo inversion algorithm described in Chapter II. Third, I apply the joint Monte Carlo inversion of surface wave dispersion, Rayleigh wave H/V ratio and receiver functions to the USArray/TA stations deployed between 2005 and May 2013, and construct a 3-D Vs model for the crust and uppermost mantle beneath nearly the entire continental US. I show in this chapter that by introducing the Rayleigh wave H/V ratio shallow structures (<5 km), including sedimentary layers and the uppermost crust, can be better constrained much better. New structural features are revealed in the resulting 3-D model, such as a low velocity anomaly in the uppermost mantle beneath the Reelfoot Rift. The model presented in this chapter, which extends nearly across the entire US, is the culmination of my thesis.

5.1 Introduction

USArray/Transportable Array (TA), one of the principal components of EarthScope, has been deployed and been evolving for about 10 years. Since 2005, the USArray/TA has deployed 400 3-component broadband portable seismic stations with a station spacing ~ 70 km. This array has crept continuously across the US until it has now nearly spanned the continent. This seismic observatory has stimulated many innovations in seismology that help better constrain the crustal and uppermost mantle structure of the earth. For example, the ambient noise cross-correlation technique was first applied to the USArray/TA(Shapiro et al., 2005) and has been combined with new tomographic methods (e.g., eikonal tomography by Lin et al., 2009) to produce high

resolution dispersion maps between periods of 8 and 40 sec. More recently, surface wave dispersion measurements at longer periods (25-100 sec) have been derived from newly developed array-based tomography (Lin and Ritzwoller, 2011a; Ritzwoller et al., 2011). These dispersion maps constrain the earth's structure through the crust to a depth of about 150 km in the uppermost mantle.

In practice, broad-band surface wave measurements from ambient noise and/or earthquake data are often jointly used with other geophysical observations to reduce the depth-velocity trade-off that characterized surface wave inversions, and receiver functions are commonly chosen (e.g., Liu et al., 2010; Tokam et al., 2010; Bodin et al., 2011; Bailey et al., 2012, Shen et al., 2013a) because they provide constraints on velocity discontinuities. In particular, Shen et al. (2013a) developed a new approach to jointly invert surface wave dispersion and azimuthally independent receiver functions for local Vs structures (as described in Chapter II). In this approach, data uncertainties are rigorously acquired and propagated into the resulting model uncertainties. Moreover, simple model assumptions are made so that all and only the features needed by the data are introduced during the inversion, which enhances the simplicity of the results. This approach has been applied on both a sub-continental scale for the western US (Shen et al., 2013b) and a regional scale for the region surrounding the Midcontinent Rift (Shen et al., 2013c), but it has not been applied systematically to the USArray/TA stations in the central/eastern US at a continental scale. Furthermore, Lin et al. (2012a) presented a new observable: Rayleigh wave H/V ratio (ellipticity) estimated using the USArray/TA across the western US and demonstrated that this new observable can be combined with Rayleigh wave phase velocity dispersion curves to improve the determination of Vsv structures for the western US. This new observable is sensitive, in particular, to shallow Vsv structures and is complementary to surface wave dispersion and receiver functions. However, the joint inversion of the H/V ratio together with receiver functions and surface wave dispersion for local structures has not been achieved previously.

Here we make observations for this new data set using over 1,700 USArray/TA stations deployed before 2013 May (Figure 1) and incorporate this data set with high resolution Rayleigh wave phase/group velocity dispersion data and receiver functions under a Bayesian Monte Carlo framework to construct a new 3-D crustal and uppermost mantle model. The region of study

extends from the Pacific coast to 80 ℃W longitude and covers over 80% of the contiguous US, about 2,000 km eastward from earlier continental-scale studies (Shen et al. 2013b), and more than two years of TA data are added compared to these earlier studies. Significantly, as discussed in earlier chapters, the introduction of receiver functions into the inversion significantly improves determination of Moho depth and structures near the crust-mantle transition, and the introduction of the H/V ratios significantly improves estimates of structures in the top few km of the crust. Overall, combining all three data sets improves the vertical resolution for the crust and uppermost mantle, and the resulting 3-D Vs model reveals higher fidelity structures for the crust and uppermost mantle across nearly the entire US.

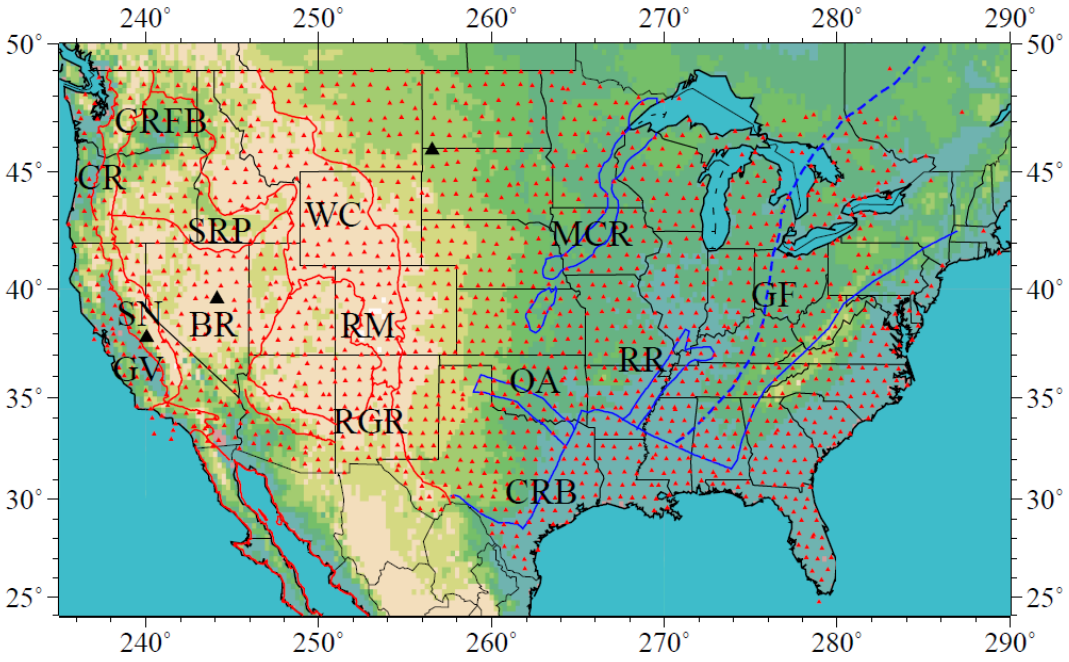

Figure 5.1. The 1,723 USArray/TA Stations used in this study are marked as red triangles. In the background, surface topography is shown, and some of the geological/tectonic features of the continental United States are outlined. Black triangles shown at California, Nevada and North Dakota-South Dakota boundary present the locations of three example stations used to exemplify the method. In the western US, major geological/tectonic provinces are outlined with red lines and identified with abbreviations: the Columbia River Flood Basalt (CRFB), the Cascade Range (CR), the Snake River Plain (SRP), the Wyoming Craton (WC), the Basin and Range (BR), the Sierra Nevada (SN), the Great Valley (GV), the Rocky Mountains (RM), the Colorado Plateau (CP) and the Rio Grande Rift (RGR). In the eastern US, aulacogens, rifts, and other features are shown with blue lines and named with abbreviations as well: the Midcontinent Rift (MCR), the

Reelfoot Rift (RR), the Oklahoma Aulacogen (OA) and the Continental Rift Boundary (CRB). The Greenville front (GF) is outlined with a dashed blue line.

5.2 Data Processing

The 1,723 stations used in this study are shown in Figure 1 as red triangles. These stations are evenly distributed across nearly the entire US continent with an average inter-station spacing of about 70 km. Utilizing these stations, we construct 1) Rayleigh wave dispersion curves from ambient noise and earthquake data, 2) receiver functions, and 3) Rayleigh wave ellipticity measurements (H/V ratio). Rayleigh wave phase velocity curves from 8 to 80 sec period are taken from surface wave dispersion maps generated by eikonal tomography (for ambient noise data) and Helmholtz tomography (for teleseismic earthquake data). Rayleigh wave group velocity curves are generated by straight-ray tomography from 8 to 40 sec. We also construct a back-azimuth independent receiver function at each station by the harmonic stripping technique. Details of these methods have been documented in several papers (eikonal tomography: Lin et. al., 2009; Helmholtz tomography: Lin and Ritzwoller et al., 2011; straight ray tomography: Barmin et al., 2001; harmonic stripping: Shen et al., 2013a; Rayleigh wave H/V ratio: Lin et al., 2012a), in the earlier chapters of this thesis, and are only briefly summarized here.

5.2.1. Rayleigh wave phase velocity curves

We measured Rayleigh wave phase velocities from 8 to 40 sec period from the ambient noise cross-correlations based on the USArray TA stations available from Jan 2005 to May 2013. The ambient noise data processing procedure follows Bensen et al. (2007) and Lin et al. (2008). More than 650,000 cross-correlations in the study region are produced. At short periods (8 to 40 sec), eikonal tomography (Lin et al., 2009) produces Rayleigh wave phase velocity maps with uncertainties based on ambient noise (e.g., Fig. 2a-c). For longer periods (28 to 80 sec), Rayleigh wave phase velocity measurements are obtained from earthquake data using the Helmholtz tomography method (Lin and Ritzwoller, 2011). Data from a total of 8,786 earthquakes between 2005 and 2013 with Ms > 5.0 are acquired, Rayleigh wave dispersion is measured and Helmholtz tomography is applied. Sample maps are presented in Figure 2e,f. In the periods that ambient noise and earthquake measurements overlap (28 to 40 sec), there is strong agreement between

the maps (Fig. 2c,e). The average difference is ~ 0.001 km/sec, and the standard deviation of the difference is ~ 0.012 km/sec, which is within the estimated average uncertainty (~0.015 km/sec).

With these Rayleigh wave phase velocity dispersion maps (8 to 80 sec), we produce local phase velocity dispersion curves at each station location. For example, the local Rayleigh wave phase velocity curve with uncertainties at station P12A in the Basin and Range is shown in Figure 5b with black error bars near the red curve.

5.2.2 Rayleigh wave group velocity dispersion curves

When automated frequency-time analysis (AFTAN) is applied to ambient noise cross-correlations, both phase and group velocity dispersion curves are measured. As discussed in section 5.2.1, eikonal tomography is performed on the phase time measurements from ambient noise. However, the eikonal equation governs propagation of phase but not group times; thus, for group velocities we use the straight ray tomography method of Barmin et al. (2001) with confidence, given the similarity of these two methods evidenced in Lin et al., 2009 and Zhou et al., 2012. We ignore finite frequency effects here, because they are weak in the period band considered (8~40s; e.g. Lin & Ritzwoller 2011; Ritzwoller et al. 2011)

The 8 and 28 sec period group velocity maps are shown in Figure 2b and 2d. At 8 sec period, group velocity is most sensitive to structure in the top 10 km of the crust. Similar to the 8 sec phase velocity map, major basins exhibit slow group velocities (< 2.4 km/sec). At 28 sec period, group velocity shows a slightly different pattern from the phase velocity map because of the relatively shallower sensitivity. Because a damped least-squares inversion is used to generate the group velocity maps, meaningful uncertainties are not obtained through the inversion. The uncertainty of group velocity is scaled from uncertainties of phase velocity using the relationship described by Moschetti et al. (2012). On average, group velocity uncertainty is ~ 3 times larger than phase velocity uncertainty. In this study, we allow group velocity measurements to be used only when the horizontal resolution is better than 100 km, so some group velocity measurements at in the eastern edge of the map are discarded. Finally, local dispersion curves are obtained for 1,702 out of 1,723 USArray/TA stations.

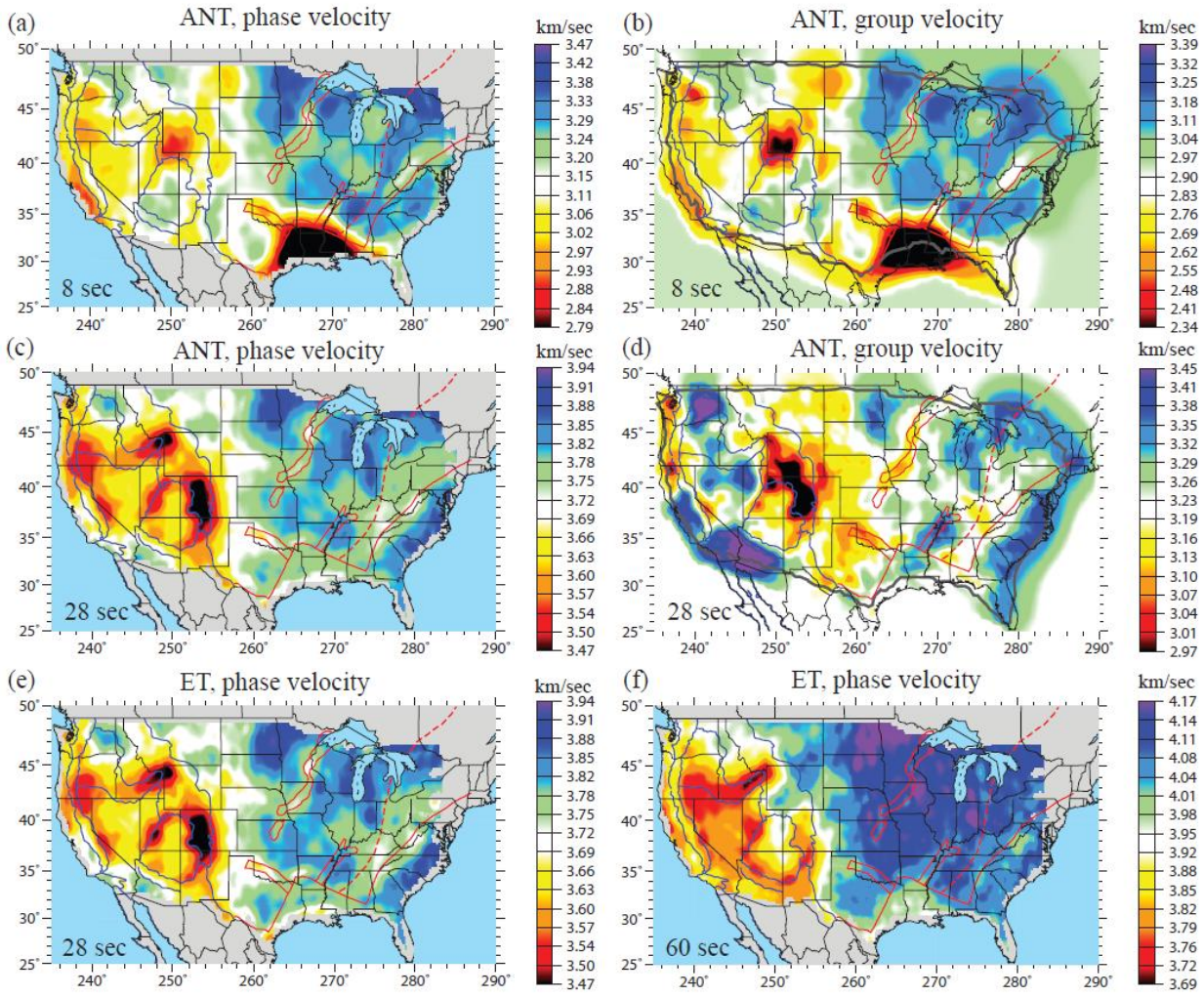

Figure 5.2. Phase and group velocity maps. (a-b) Phase and group velocity maps at 8 sec are shown. The phase velocity map is constructed using ambient noise eikonal tomography method, and the group velocity map is constructed using ambient noise straight-ray tomography method. (c-d): similar to the (a-b), but at 28 sec period. (e-f) Phase velocity maps at 28 and 60 sec generated by teleseismic earthquakes are shown.

5.2.3 Receiver function data processing

The method we use to process receiver functions for each station is described in detail by Shen et al. (2013b) and discussed in Chapter II of this thesis. For each station, we pick earthquakes from Jan 2005 to May 2013 with epicentral distance from 30 °to 90 °, and with magnitude mb > 5.0. We apply a time domain deconvolution method (Ligorria and Ammon, 1999) to each seismogram windowed between 20 sec before and 30 sec after the direct P-wave arrival to calculate the radial component receiver function using a low-pass Gaussian filter with a width of

2.5 s (pulse width ~ 1 sec), and high-quality receiver functions are selected through an automated procedure. Corrections are made both to the time and amplitude of each receiver function, normalizing to a reference slowness of 0.06 sec/km (Jones and Phinney, 1998). Finally, only the first 12 sec after the direct P arrival are retained for further analysis. We compute the azimuthally independent receiver function, $R_0(t)$, for each station by fitting a truncated Fourier Series at each time over azimuth and stripping the azimuthally variable terms using a method called "harmonic stripping" by Shen et al. (2013b). After removing the azimuthally variable terms at each time, the RMS residual over azimuth is taken as the 1σ uncertainty at that time. On average, 84 receiver functions from different earthquakes are collected for each station. For stations with less than 10 receiver functions that passed quality control, joint inversion is not be applied to them. Finally, azimuthally independent receiver functions are obtained for 1,493 USArray/TA stations covering over 80% of the continental US. Three typical receiver functions are presented in Figure 5c, 5f, and 5i.

5.2.4 Rayleigh wave H/V ratio measurements

As described above, data following more than 8,000 teleseismic earthquakes are collected to perform teleseismic Helmholtz tomography for the continental US. The same data set is also used to measure the Rayleigh wave H/V ratio. To measure the H/V ratio with these data, we follow the processing procedure presented by Lin et al. (2012a), which we summarize briefly here.

For each earthquake recorded by each available station, 3-component seismographs are cut according to the Rayleigh wave travel time predicted by a global model (Shapiro and Ritzwoller, 2002), and the mean, linear trend and the station response are removed. The horizontal components (E and N) are then rotated to radial (R) and transverse (T) according to the great-circle path. AFTAN (Bensen et al., 2007) is applied to determine Rayleigh wave travel times, and amplitudes of both the V and R components are measured between 18 and 80 sec. The amplitude ratio between the two components (R/V) is used to evaluate the H/V ratio as a function of period at the station location.

To further insure the quality of the H/V measurements, we impose a 4-step quality control procedure: First, signal to noise ratio should be smaller than 15 for Rayleigh waves in boththe

radial and horizontal directions. Second, we apply a phase difference criterion such that measurements with $|T(R)-T(V)-\pi/4| > 2$ sec are removed, where T represents the observed phase of the radial (R) or vertical (V) component. Third, high H/V measurements (>10) are removed. Fourth, for each station, we discard the measurements outside the 2σ range to further stabilize the estimate of the mean H/V ratio.

After these quality control steps, an ensemble of H/V ratio measurements are obtained for a given period at each station. For the ensemble with number of measurements ≥ 20, the mean and standard deviation of the mean are computed to represent the H/V ratio measurement and its uncertainty for this period. We discard the ensemble with measurement number smaller than 20 to enhance the reliability of the estimate of the mean. Moreover, at each station, the H/V measurements estimated from different events are similar, although small variations (<2%) dependent on back-azimuth can be observed. However, we also find that for about 15 stations the H/V ratio measurements possess large variations (>20%) with time mostly due to the differential signal output error of the seismometer sensors (personal communications with IRIS/DMC). The H/V ratio measurements that are estimated during the malfunction period are discarded when the differential signal output happens on the E/N components; otherwise they are corrected to be consistent with normal H/V measurements estimated outside the malfunction epoch.

The resulting H/V maps at 30 and 60 sec are presented in Figure 3a,b. They are very similar to the maps shown by Lin et al. (2012a), although the H/V ratio presented in this study is smaller on average and they extend over a larger area. At 30 sec, high H/V ratio is correlated with the sedimentary basin distribution across the continental US (e.g., the Central Valley in California; the Williston Basin near Eastern Montana and Western North Dakota; the coastal basins near the Mississippi embayment). Low H/V ratio is observed in major mountain ranges (e.g., the Sierra Nevada, Rocky Mountains) in the western US, and in the Superior upland province and the Appalachian highlands in the eastern US. Beneath the Yellowstone hotspot and southern Sierra Nevada, the lowest H/V ratio measurements (<0.65) are observed. At longer periods, the H/V ratio is still partially sensitive to shallow Vs structure, so the effect from major basins is still observed: the Green River Basin in southwestern Wyoming and the Mississippi embayment areas possess the highest H/V at this period. But sensitivity of H/V ratio penetrates into the

mantle and is positively correlated with Vsv, so low H/V ratio is also observed beneath the Rocky Mountains and the Snake River Plains where the mantle Vsv is low (Fig. 2f).

The H/V ratio dispersion curves for 1,619 stations are constructed from 18 to 80 sec. Typical H/V ratio curves with uncertainties for three example stations are presented in Figure 5a, 5d and 5g.

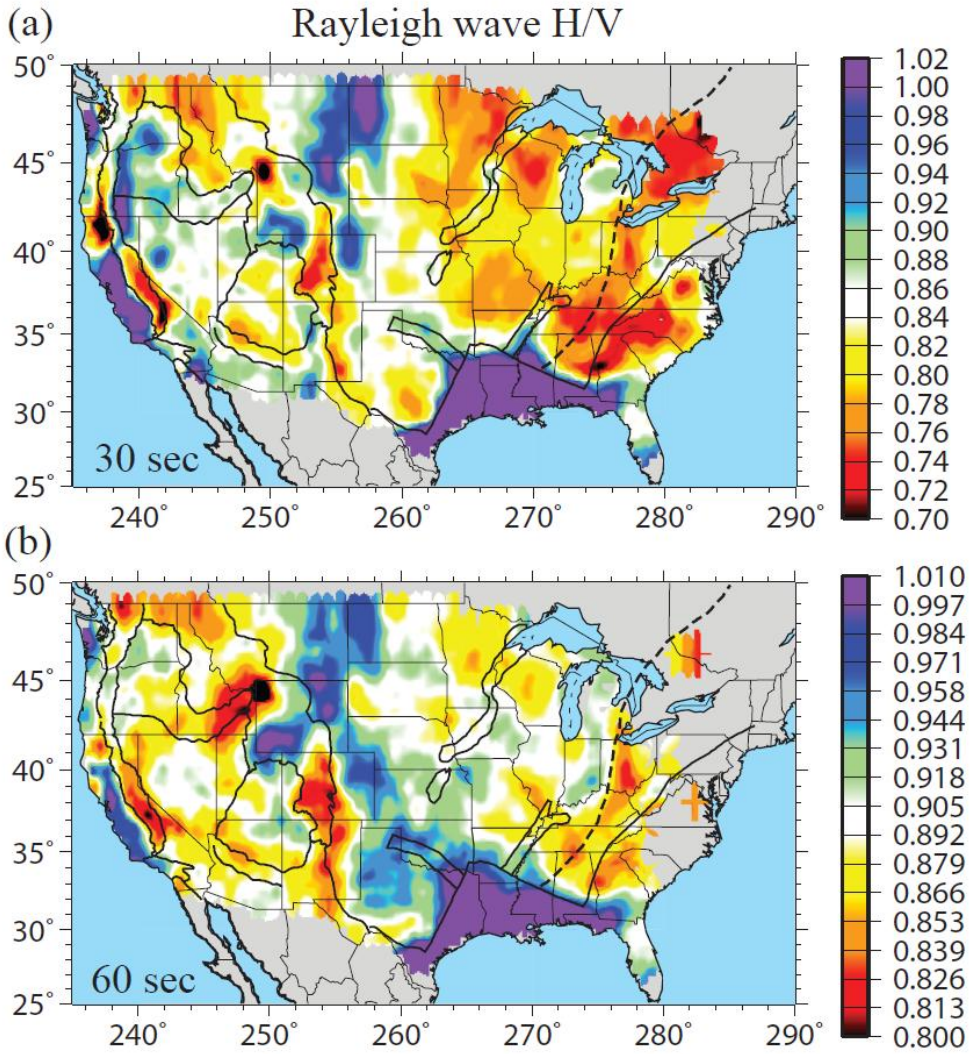

Figure 5.3. (a) The map views of the Rayleigh wave H/V ratio at 30 sec. (b) Same as (a), but for the H/V ratios at 60 sec.

5.3 Joint Monte Carlo Inversion of Surface Wave Dispersion, Receiver Functions and Rayleigh Wave H/V Ratio.

In this section we describe the joint Bayesian Monte Carlo inversion that is extended to incorporate measurements of H/V ratios in three steps. In the first step, we perform two initial Monte Carlo inversions without the involvement of H/V ratio meaurements. Such inversions produce two models of the crust and uppermost mantle Vs for the continental US: 1) a model constrained only by surface wave dispersion data, which is called the "SW alone" model and 2) a model constrained by surface wave dispersion and receiver functions jointly, which is called the "SW+RF" model. Based on such models, the Rayleigh wave H/V ratio maps are predicted and compared with the observed H/V ratio measurements to demonstrate the need to impose the H/V data. In the second step, H/V ratios are incorporated in the Monte Carlo inversion, and three example stations are chosen to demonstrate the compatibility of the three data sets to be inverted under the Bayesian Monte Carlo framework. In the final step, the new 3-D model constrained by all three types of data (the "SW+RF+H/V" model) is generated, and the fit to the data are discussed. But, before the introduction of the three-step inversion, we first briefly describe the parameterization and model spaces used by these Monte Carlo inversions in section 5.3.1.

5.3.1 Model space and prior constraints

We currently only use Rayleigh wave data (and not Love wave data), which is primarily sensitive to Vsv, so we assume the model is isotropic: Vsv=Vsh=Vs. The Vs model beneath each station is divided into three principal layers. The top layer is the sedimentary layer defined by three unknowns: layer thickness and Vs at the top and bottom of the layer with Vs increasing linearly with depth. The second layer is the crystalline crust, parameterized with five unknowns: four cubic B-splines and crustal thickness. Finally, there is the uppermost mantle layer, which is given by five cubic B-splines, yielding a total of 13 free parameters at each location. The thickness of the uppermost mantle layer is set so that the total thickness of all three layers is 200 km. For the initial surface wave inversion, the model space is defined based on perturbations to a reference model consisting of the 3D model of Shapiro and Ritzwoller (2002) for mantle Vs, crustal thickness and crustal shear wave speeds from CRUST 2.0 (Bassin et al., 2000), and sedimentary thickness from Mooney and Kaban (2010). For the joint inversion of surface wave

data and receiver functions, the model space is defined based on perturbations to the resulting average model from the initial surface wave inversion.

In addition, the following three prior constraints are introduced in the Monte Carlo sampling of model space. (1) Vs increases with depth at the two model discontinuities (base of the sediments and Moho). (2) Vs increases monotonically with depth in the crystalline crust. (3) Vs < 4.9 km/sec at all depths. These prior constraints reduce the model space effectively. Vp and density are scaled from Vs according to Brocher (2005):

$$Vp = 0.9409 + 2.0947Vs - 0.8206Vs^2 + 0.2683Vs^3 - 0.0251Vs^4; \quad (1)$$

$$\rho = 1.22679 + 1.53201Vs - 0.83668\ Vs^2 + 0.20673Vs^3 - 0.01656Vs^4; \quad (2)$$

The scaling relationship of Vp compared to Vs makes the Vp/Vs ratio higher than 2 in the sedimentary layer where Vs < 3 km/sec, and Vp/Vs ~ 1.72 in the crystalline crust. In the mantle, Vp/Vs is set to be 1.75 and density is scaled from Vs perturbations with 10 kg/m^3 per 1% velocity change (Hacker and Abers, 2003). The Q model from PREM (Dziewonski and Anderson, 1981) is used to apply the physical dispersion correction (Kanamori and Anderson, 1977) and the resulting model is reduced to 1 sec period.

5.3.2 Initial Bayesian Monte Carlo inversions without H/V ratios

After the model space is determined, the Monte Carlo inversion of surface wave dispersion alone and the joint inversion of surface wave dispersion and receiver functions are applied. The technique has been well discussed by Shen et al. (2013a) and in Chapter II and its systematic application to USArray have been documented by Shen et al.(2013b,c) and in Chapters III and IV, so we skip the details of the inversions here.

The first “SW alone” inversion is applied to all 1,703 stations where local dispersion curves are obtained. At each station, an ensemble of models is obtained from the Monte Carlo inversion. From this ensemble, the mean and 1 standard deviation are calculated at each depth to represent the average model and associated uncertainties. We then compute the H/V ratio based on the average model at each station location.

Figure 4a presents the 30 sec H/V ratio maps predicted by the “SW alone” model. To first order,

it is very similar to the observed H/V ratio map (Fig 3a) in pattern: high H/V ratio is observed near major sedimentary basins and low H/V ratio is seen along mountain ranges and the continental core. Quantitative comparisons are presented in Figure 4c as differences between the predicted H/V ratios and the observed H/V ratios normalized by the estimated uncertainties of observations. Red stripes (e.g., the Sierra Nevada) in Figure 4c show that the observed H/V ratio is much lower (< -4σ), while the blue stripes (e.g., the Williston Basin) imply that observed H/V ratio is higher (>4σ). In many areas (e.g., the Basin and Range) the differences between the observed and predicted H/V ratios are relatively small (e.g., the midcontinent areas), implying that the 3-D model constrained by surface waves alone possesses a shallow structure that is not totally incorrect. In contrast, Lin et al. (2012a) found a large misfit between the observed and predicted H/V ratios from surface wave phase velocity inversion alone. This is because the group velocity dispersion curves introduced in this study improve constraints to the structures at shallow depths.

For the initial joint inversion of surface wave and receiver functions, local 1-D Vsv models at 1,493 stations are generated, and the H/V ratios are predicted based on the models. Figure 4b,d presents the predicted H/V ratio at 30 sec from the "SW+RF" model and the normalized misfit to the observed data. Compared with the H/V ratios predicted by "SW alone" model, the high H/V ratios observed at the Williston Basin and the Denver Basin are more prominently predicted by the "SW+RF" model, and the pattern of the high H/V ratio in southwestern Wyoming is also different. To quantify the fit to observed data, we define the mean misfit to H/V ratio as:

$$M = \left(\frac{1}{N}\sum_{i=1}^{N}\left(H_i^{obs} - H_i^{m}\right)^2/\sigma^2\right)^{1/2} \quad (3)$$

where N represents the number of stations; H_i^{obs} and σ represents the observed H/V ratio and uncertainty at the ith station; H_i^{m} represents the calculated H_i^{m} for the same station based on the resulting model. H/V predicted by the "SW+RF" model is more similar to the observed H/V ratio map, with a mean misfit ~5.20, while the mean misfit for the H/V ratio predicted by the "SW alone" model is ~ 6.24. The improvement of the H/V ratio prediction implies that the introduction of receiver functions helps to constrain the shallow Vsv structures. However, the relatively large misfit (>5) to observed H/V measurements for both models suggests that the H/V ratio is needed to be incorporated in the joint inversion.

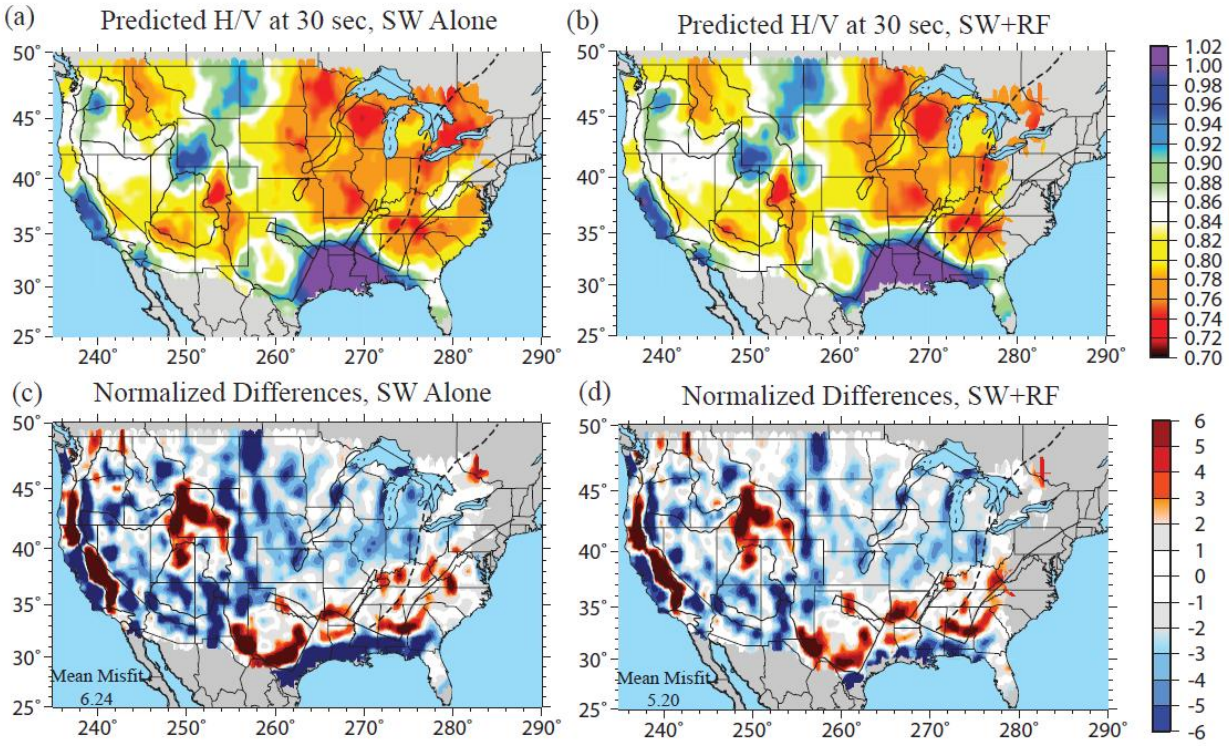

Figure 5.4. (a) 30 sec Rayleigh wave H/V ratio map predicted from model constructed by surface wave inversion alone. (c) Same as (a), but for 30 sec H/V ratio predicted by model constructed by joint inversion of surface wave dispersion and receiver functions. (c) Differences between H/V ratio shown in (a) and the observed H/V ratio shown in Fig. 3(a) normalized by uncertainties. (d). Similar to (c), but shows the difference between (b) and Fig. 3(a).

5.3.3 Joint inversion of surface wave dispersion, receiver functions and H/V ratio

As discussed previously, although the Vs model constrained by surface wave dispersion and receiver functions can predict the general spatial pattern of the H/V ratio, the observed H/V ratio is needed to be incorporated in the inversion. Lin et al. (2012a) demonstrated that the Rayleigh wave phase velocity dispersion and Rayleigh wave H/V ratio can be combined to invert for local Vs structure, and Shen et al. (2013a) (and Chapters II and III in this thesis) demonstrated that Rayleigh wave phase velocity dispersion and receiver functions can be combined as well. In this section we provide evidence that all three data sets can be interpreted together to invert for simple a Vs structure under a Monte Carlo framework. Three USArray/TA stations marked as black triangles in Figure 1 at various geological provinces are selected: P12A at McGill, Nevada in the Basin and Range; F25A at Bowman, North Dakota in the Williston Basin; and S06C at Mather, California in the Sierra Nevada mountain range. Such stations are chosen because they are located at various tectonic provinces, and the observed data at such stations show great variety.

Figure 5 presents the observed Rayleigh wave H/V ratio, phase and group velocity dispersion curves as error bars, and azimuthally independent receiver functions as two parallel lines for the example stations. At station P12A in the Basin and Range province, the observed Rayleigh wave H/V ratio is low at short periods and peaks at ~ 50 sec (Fig. 5a). The phase velocity dispersion curve increases between 8 and 35 sec dramatically, and the group velocity curve possesses an Airy phase at ~ 18 sec. The receiver function at this station possesses a clear Ps conversion at the Moho discontinuity at ~ 4 sec. In contrast, the Rayleigh wave H/V ratio observed at station F25A located in Williston Basin is high at short periods (< 30 sec), and the group velocity is much lower (2.6 km/sec) at 8 sec. The receiver function observed at this location is dominated by the reverberations generated by shallow discontinuities. At station S06C in the Sierra Nevada, the H/V ratio is very low at short periods, and the receiver function is quite flat without major Ps conversions generated by any discontinuities. Based on these observed data sets, a joint inversion of surface wave dispersion, receiver functions and H/V ratio is performed for the three stations.

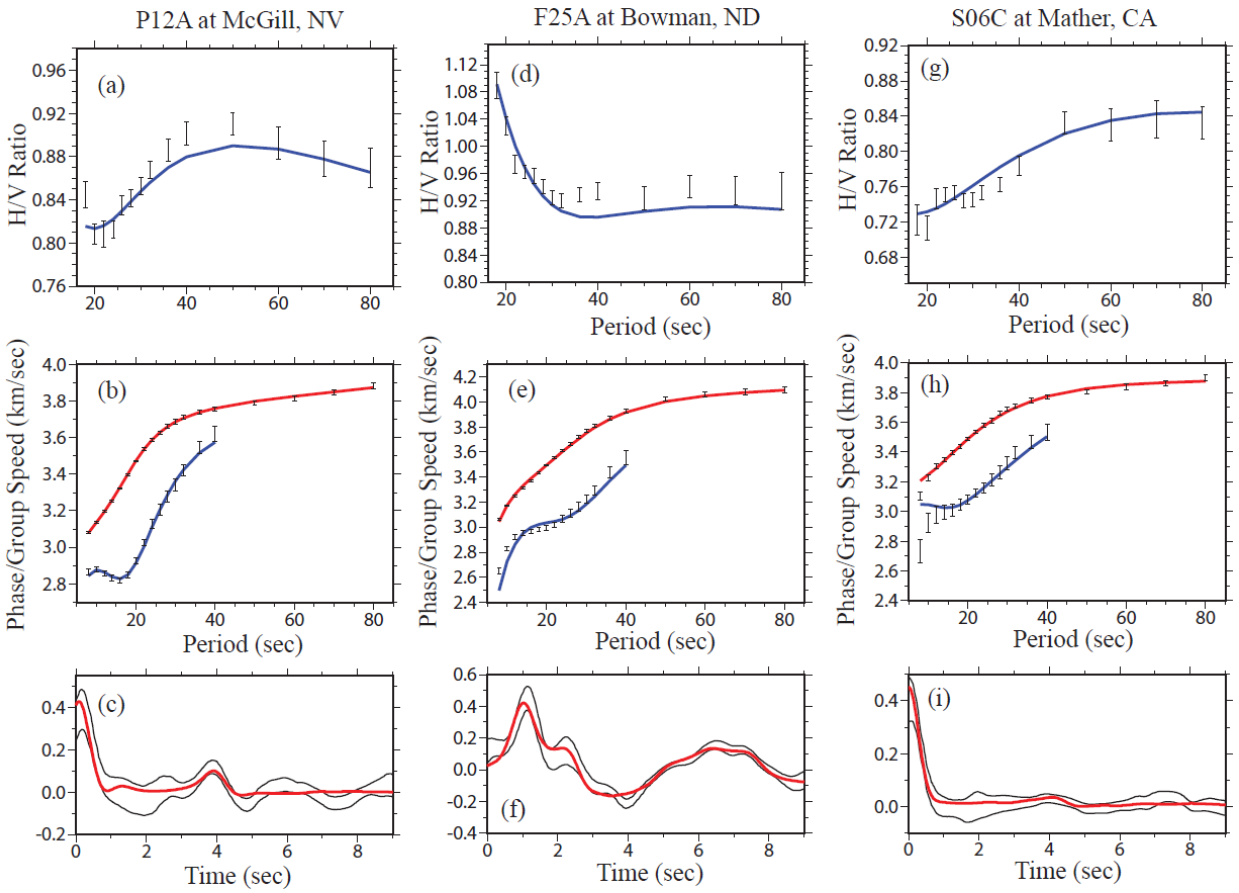

Figure 5.5. (a-c) Observed Rayleigh wave H/V ratio (error bars in upper panel), Rayleigh wave phase velocity curves (error bars in the middle panel), and azimuthally independent receiver functions (pair of black lines in lower panel) for station P12A at McGill, Nevada. These data sets are presented in comparison with predictions based on the Vs models shown in Fig. 6. (d-f) Similar to (a-c), but for station F25A at Bowman, ND. (g-i) Similar to (a-c), but for station S06C at Mather, CA.

The joint inversion of multiple data sets have been established by Shen et al. (2013a) and this algorithm is utilized here. After the inversion, the result of the joint Monte Carlo inversion is not a single model, but an ensemble of models. Figure 6 presents the ensemble of models for the three example stations. For station P12A, the most prominent feature in the resulting model is the sharp boundary between the crust and uppermost mantle at ~ 32 km depth (Fig 6b). At shallow depths, a relatively thin sedimentary layer (~ 1km) sits on top of the crystalline crust (Fig. 6a). For station F25A at the Williston Basin, a thick sedimentary layer is seen in the resulting model (Fig 6c), with a slow Vs (< 1km/sec) at the top of sedimentary layer. This is consistent with local geological investigations which indicate weakly consolidated sediments in the upper few hundred meters underlain by multiple sedimentary strata of Ordovician to Cretaceous ages

(Blackwell et al., 2006). The crustal thickness is not well determined for this station, mainly because the Moho Ps conversion is contaminated by reverberations at shallow discontinuities (e.g, the bottom of the sedimentary layer). For station S06C, the sedimentary layer is less than 0.5 km thick, and the absence of the Ps phase leads to a resulting model with a gradient Moho structure.

For each station, from the ensemble of models, the mean and the standard deviations of Vsv at each depth can be calculated. The surface wave dispersion, receiver functions as well as the Rayleigh wave H/V ratio can be calculated from the mean model. Based on the model ensembles presented in Figure 6, we calculate all three types of data and draw them with the observed data in Figure 5. Overall, the resulting models are able to successfully recover the observed data. For example, at station P12A, the χ misfit for the average model to phase velocity dispersion, group velocity dispersion, receiver functions, and H/V ratios are 1.00, 1.10, 0.82, and 1.24, respectively. The χ misfits to the observed data are below 2 for the other two stations presented here as well, which demonstrates the compatibility of all three types of data to invert for local 3-D model.

However, we also observe a small discrepancy between the predicted and observed data in some cases. For station S06C in the Sierra Nevada, the observed short periods phase and group velocity dispersion cannot be perfectly fit by the resulting model. This is because both receiver functions and H/V ratios are sensitive to localized structures but dispersion curves estimated from surface wave travel time measurements are affected by neighboring structures (e.g., Central Valley for this station). As a result, the slow phase/group velocities observed at short periods are not completely compatible with the receiver functions and H/V ratios. Another small discrepancy is for the H/V ratio. For example, the observed H/V ratio is systematically higher than the predicted values between 30 and 60 sec for station P12A, and between 40 and 80 sec for station F25A. A similar misfit pattern is seen by Lin et al. (2012) as well. These small misfits of the H/V ratios will be discussed in the following section in detail.

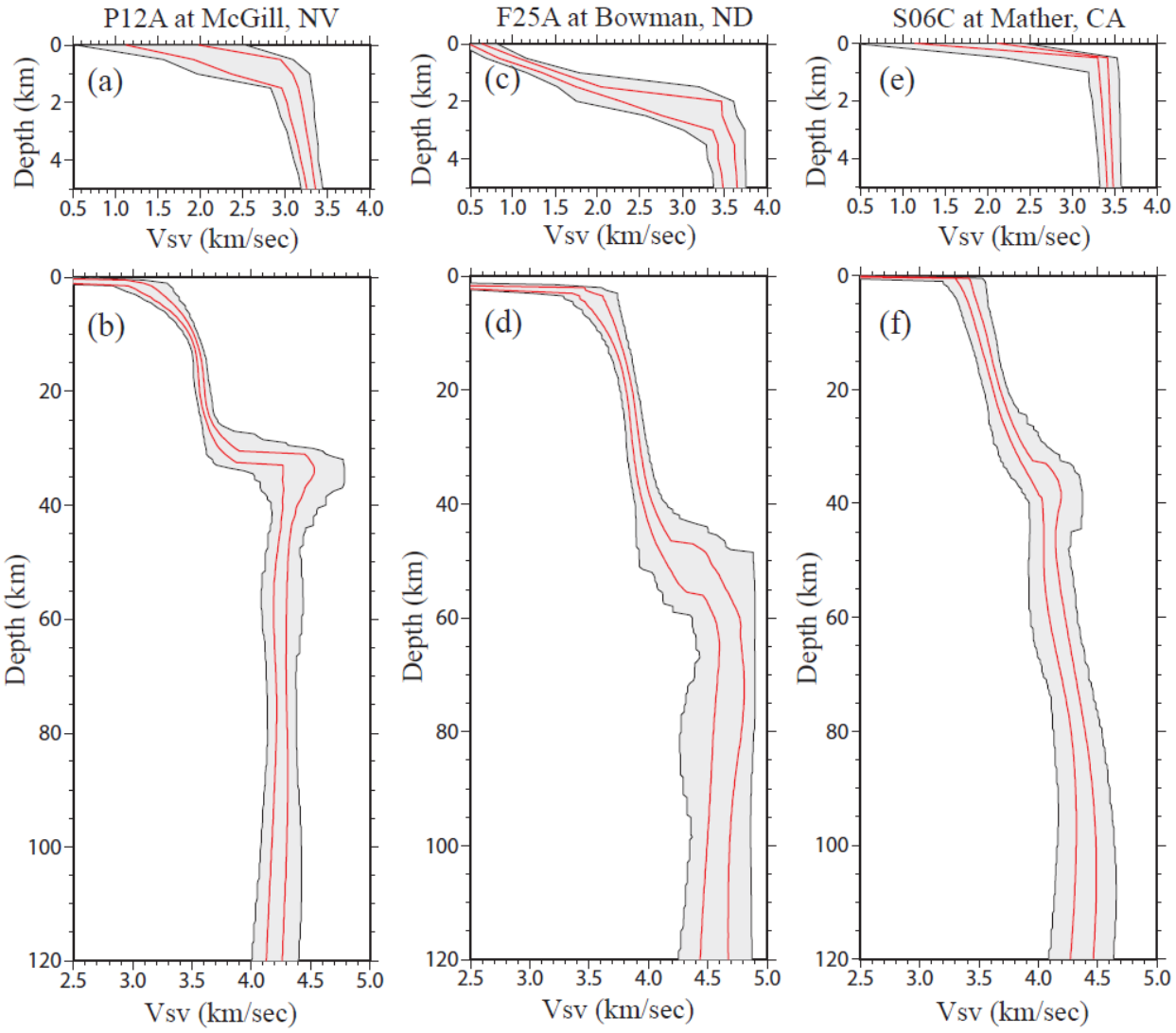

Figure 5.6. (a-b) Example outcome of the joint inversion at USArray TA station P12A in the Basin and Range province in McGill, Nevada. The full width of the resulting model ensemble is presented as black lines enclosing a grey-shaded region; the 1σ ensemble is shown with red lines.

5.3.4 Fit to H/V Ratio data

The joint Monte Carlo inversion is then applied to all 1,392 stations, at which all three types of data (surface wave dispersion, receiver function and H/V ratio) can be obtained. After the inversion, we again calculate the H/V ratio based on the resulting models, and the improvement of prediction of H/V data at 30 sec is summarized in Figure 7. Predicted from the new model, H/V ratio map at 30 sec (Fig. 7a) shows more consistent with the observed H/V at the same period (Fig. 3a), and the average misfit decreases to 2.02 from 5.20. As shown in Figure 7c, the normalized difference to H/V ratio at 30 sec is within (-2,2), which can be viewed as good.

However, the misfit is systematically rather than randomly distributed: The predicted H/V ratios are mostly lower than observed data in the Great Plains (e.g., station F25D shown in Fig. 5b) and some areas in the Basin and Range province (e.g., station P12A shown in Fig. 5a). At longer periods, the systematical misfit is still observed (Figure 7b,d) in the Great Plains and Colorado Plateau areas. In fact, the differences between the predicted and observed H/V ratios are larger at 60 sec, but because the uncertainties are also usually greater at longer periods, thus the average normalized misfit is smaller at 60 sec (~1.70).

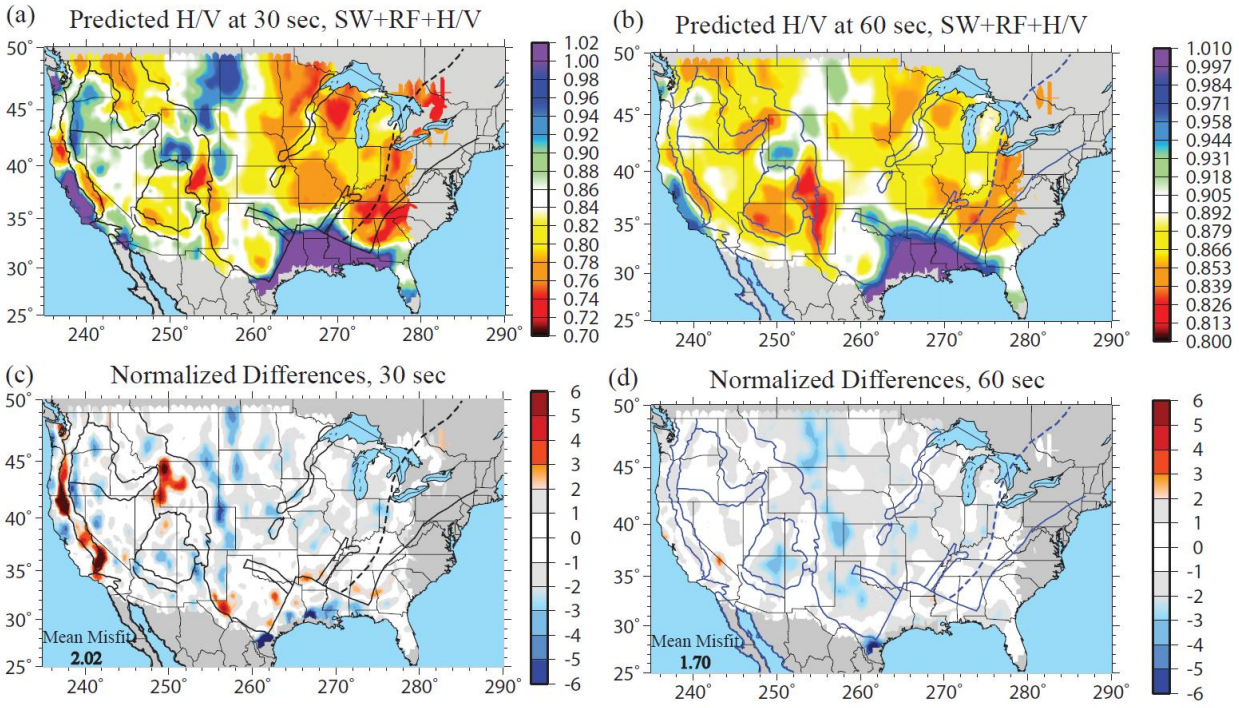

Figure 5.7. (a) 30 sec Rayleigh wave H/V ratio map predicted from model constructed by joint Monte Carlo inversion of surface wave dispersion, receiver function, and Rayleigh wave H/V ratio. (b) Same as (a), but for 60 sec H/V ratios. (c) Differences between H/V ratio shown in (a) and the observed H/V ratio shown in Fig. 3(a) normalized by uncertainties. (d) Same as (c) but for 60 sec H/V ratios.

The systematical misfit between the predicted and observed H/V ratio occurs may be caused by two reasons. For the large misfit located around major basins (mainly the Denver Basin and Williston Basin), probably the data requires much higher Vp/Vs ratio suggested by equation (1) or a much lower density than the relationship in equation (2) in the sedimentary layer. If this is the case, the misfit can be solved by including Vp/Vs ratio and density in the sedimentary layer as free parameters. But for the areas off the major sedimentary basins (e.g., the Colorado Plateau,

the Colorado-Kansas boundary), perturbing density or Vp/Vs ratio in the sedimentary layer may not solve the systematical misfit. In this case, it is possible that the observed data is biased by the 3-D wave effect (e.g., back-scattering and interference). To further investigate these factors that affect the inversion, the azimuthal variation of H/V ratio measurements should be investigated and the waveform modeling should be carried out. But these further investigations are beyond the scope of this study.

5.4 Results and discussion

As discussed in previous section, 1,392 average 1-D Vsv models and the adjunct uncertainties are obtained for the USArray/TA after the joint Monte Carlo inversion of surface wave dispersion, receiver functions and H/V ratio. Following Shen et al (2013ab), a simple kriging is applied to these models to produce a uniform 3-D Vs model ("SW+RF+H/V" model) for crusta and uppermost mantle beneath the continental US at 0.25° by 0.25° grid. In this section, the features in this 3-D model is presented in comparison with models constructed without the evolvement of H/V ratio ("SW alone" and "SW+RF" models).

5.4.1 Features of the 3-D model

As demonstrated by Lin et al. (2012a), the introduction of H/V ratio in the inversion will greatly improve the shallow structure. In Figure 8, we summarize the improvement of shallow Vsv structure (top 5km) as the receiver functions and H/V ratio are incorporated in the inversion stepwise. Figure 8a and 8b shows the map views of Vsv of top 5 km of the "SW alone" model and "SW+RF" model, respectively. With the introduction of receiver functions, the change to the shallow structure is minor and focused on the Williston Basin and Denver Basin. Furthermore, the Green River Basin near southwestern Wyoming shows higher velocity. Figure 8c presents the top 5 km Vsv map of the "SW+RF+H/V" model, in which the major basins in the continental US are more pronounced with Vsv < 2.5 km/sec (e.g., the Central Valley in California, the Williston basin at Wyomint-North Dakota boundary, the Powder River basin in northeastern Wyoming, the Denver basin at eastern Colorado/Wyoming, the Washita basin in northwestern Colorado and the Michigan basin in central Michigan). Moreover, the H/V ratio also improves the determination of non-basin structures. For example, a fast anomaly (>2.9 km/sec) shows up at the Wyoming-South Dakota boundary, which is correlated with the volcanic rocks in the Black Hill. This

feature is not prominent in the "SW alone" model because the eikonal/Helmholtz tomography is more sensitive to travel time difference between nearby stations but not localized anomalies. However, the H/V ratio and receiver functions are both observed locally and thus the localized structures are better imaged when both data are incorporated.

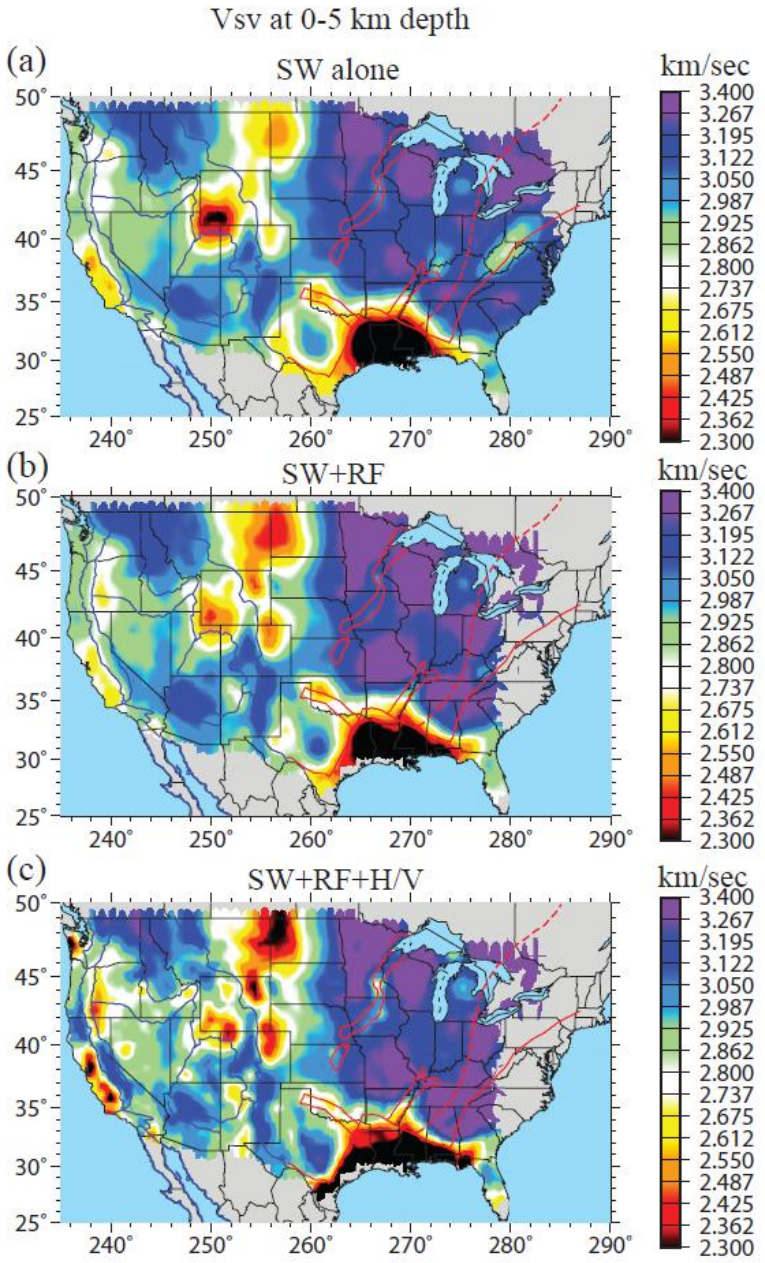

Figure 5.8. (a) Map view of the average Vsv in the uppermost crust (0-5km) for the 3-D model constructed by using surface wave dispersion alone. (b) Same as (a), but for the 3-D model from joint inversion of surface wave dispersion and receiver functions. (c) Same as (a), but for 3-D

model from joint inversion of surface wave dispersion, receiver functions and Rayleigh wave H/V ratios.

In the middle crust, the Vsv also varies when receiver functions and H/V ratios are introduced as well. Figure 9a-9c illustrates the effect of receiver functions and H/V ratio on the Vsv in the middle 1/3 of the crystalline crust. When receiver functions are introduced into the Monte Carlo inversion (Fig 9b), they are not constraining the middle crustal Vsv directly because the little sensitivity to the absolute velocity. Indeed, the effects to the Vsv given by receiver functions are imposed through the modification to the crustal thickness. In contrast, the H/V ratio will directly constrain the middle crustal Vsv because of its broader sensitivity. As a result, the middle crustal Vsv from "SW+RF+HV" model contains more localized features introduced by H/V ratio data (Fig 9c). In the western US, the fast middle crust exists beneath the Yakima Fold Belt near the northern Oregon, the Snake River Plains, and the Colorado Plateau, and the fastest Vsv is found beneath the northern Great Plains spanning from western North Dakota to eastern Montana extending to northern Wyoming. The slow Vsv in the crust is mainly located beneath the Basin and Range, the Rocky Mountains and the Idaho Batholith. In the eastern US, velocity is uniformly higher than the western US. The highest Vsv (>3.8 km/sec) in this region is found beneath the Oklahoma Aulacogen and the Reelfoot Rift areas. Figure 9 also presents the lower crustal velocity (2/3 of the crust to Moho) from "SW+RF+H/V" model. Overall the lower crust in the western US is less than 3.9 km/sec and the central and eastern US possess a faster lower crust (>4km/sec). Fast Vsv in the lower crust spreads the eastern Wyoming, Montana and the northeastern part of North Dakota. Fast speeds are also seen beneath large areas extending from northern Texas to Tennessee, which can be seen in the vertical transect shown in Figure 12b. The distributions of these fast lower crust region is consistent with the "7.X" layer revealed by seismic refractions across the continent (Braile, 1989; Gorman et al., 2002;)

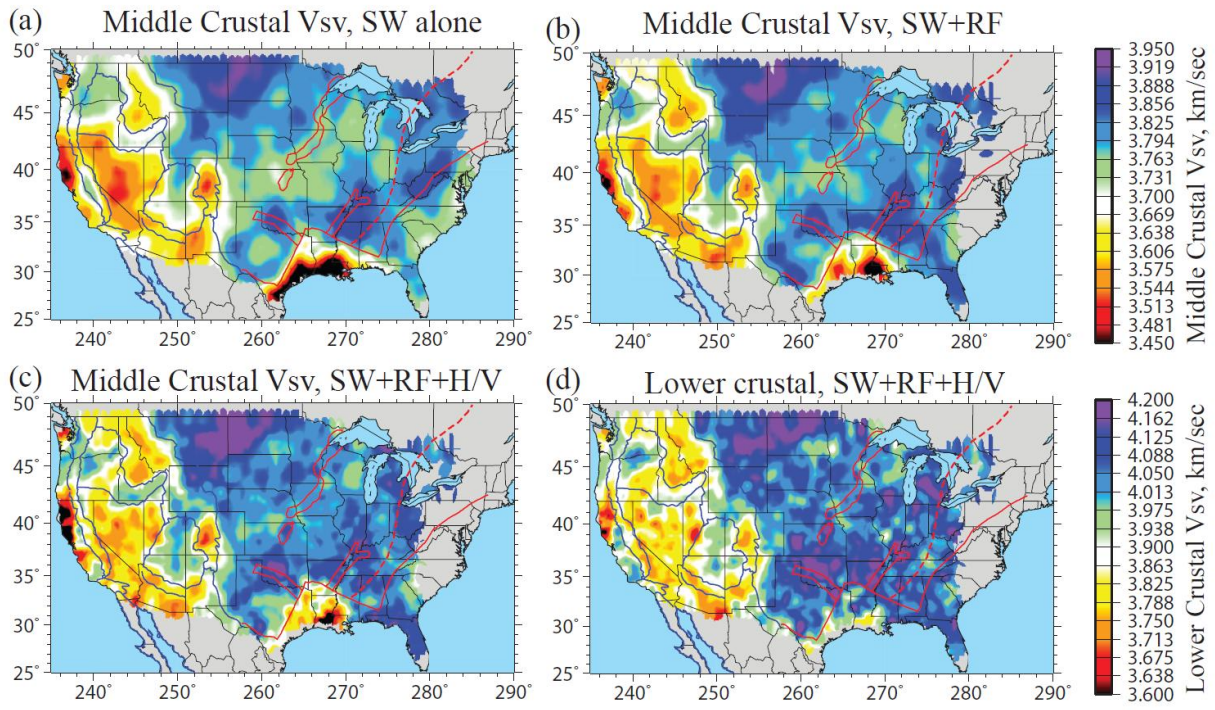

Figure 5.9. (a) Map view of the Vsv in the middle crust of the 3-D model constructed by using surface wave dispersion alone. (b) Same (a), but for the 3-D model from joint inversion of surface wave dispersion and receiver functions. (c) Same as (a), but for 3-D model from joint inversion of surface wave dispersion, receiver functions and Rayleigh wave H/V ratios. (d) Map view of the Vsv in the lower crust for the 3-D model from joint inversion of all three data sets.

For crustal thickness, models "SW+RF" and "SW+RF+H/V" possess similar values (Fig 10b,c), which differ from the results for the "SW alone" model (Fig. 10a). This is mainly because this variable is mostly improved by the introduction of the receiver functions rather than the H/V ratios. On average, the crustal thickness maps derived by the incorporation of receiver functions are more variant, and display larger contrast for small scale features, while the crustal thickness from surface wave inversion alone is smoother. For example, by using the surface wave dispersion alone, crustal thickness beneath the Rocky Mountain areas is ~ 50-55 km uniformly. But with receiver functions, thicker crust (>55 km) is found in north Colorado and relatively thinner crust is revealed in central Colorado (~50 km). Such small scale variations are consistent with recent study of the crustal structure using denser seismic arrays in this area (Hansen et al., 2013).

In the western US, crustal thickness variations agree with similar estimates from Shen et al. (2013b). In the central-eastern US, two crustal thinning features can be identified, one is located

at southeastern North Dakota extending to Minnesota and South Dakota with crustal thickness ~ 37 km, which has been partially revealed by previous study (Shen et al. 2013c). Thin crust is also observed beneath the Mississippi embayment areas where the Moho is ~ 30-35 km deep, which extends north along the Reelfoot Rift to western Tennessee. Thick crust is shown beneath the Midcontinent Rift areas, eastern of the Grenville Front, and at the ends of the Reelfoot Rift and Oklahoma aulakogen. In particular, the Grenville Front serves as the major boundary of the Moho topography. The tectonic implications behind this feature should be further investigated.

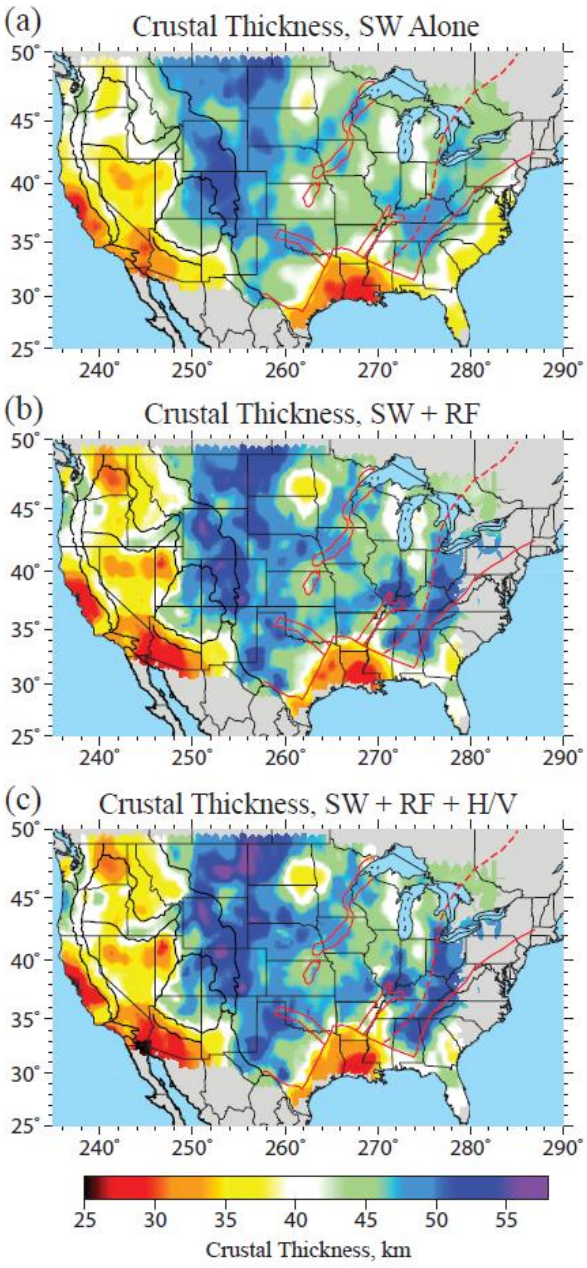

Figure 5.10. Crustal thickness maps from the models generated by: (a) Monte Carlo inversion of surface wave dispersion alone; (b) joint Monte Carlo inversion of surface wave dispersion and receiver functions; (c) joint Monte Carlo inversion of surface wave dispersion, receiver functions and Rayleigh wave H/V ratio.

Deep in the mantle, the introduction of receiver functions and H/V ratios improve the Vsv structures. Figure 11a,b present the Vsv at 80 km depth from "SW alone" and "SW+RF+H/V" models, respectively. Similarly to the crustal features, mantle heterogeneities at 80 km depth are sharpened by the introduction of receiver functions and H/V ratios. For example, the low Vs anomaly (<4.2 km/sec) in the "SW+RF+H/V" model beneath the Rio Grande Rift is largely narrowed at its northern end located at Colorado-New Mexico boundary. Another improved feature is the low velocity anomaly (~ 4.5 km/sec) beneath the Reelfoot Rift. For the "SW alone" model, a broad slow velocity zone extends from eastern Texas to northern Illinois with a width ~ 500 km. In contrast, a more localized slow anomaly is revealed beneath the Reelfoot Rift area and ends at western Kentucky at 80 km depth. This anomaly shifts toward east at greater depths (Fig. 11c) and the pattern breaks at depth ~ 150 km (Fig. 11d). To better illustrate many features revealed from these maps, three vertical transects along the profiles outlined in Figure 11d are presented in Figure 12.

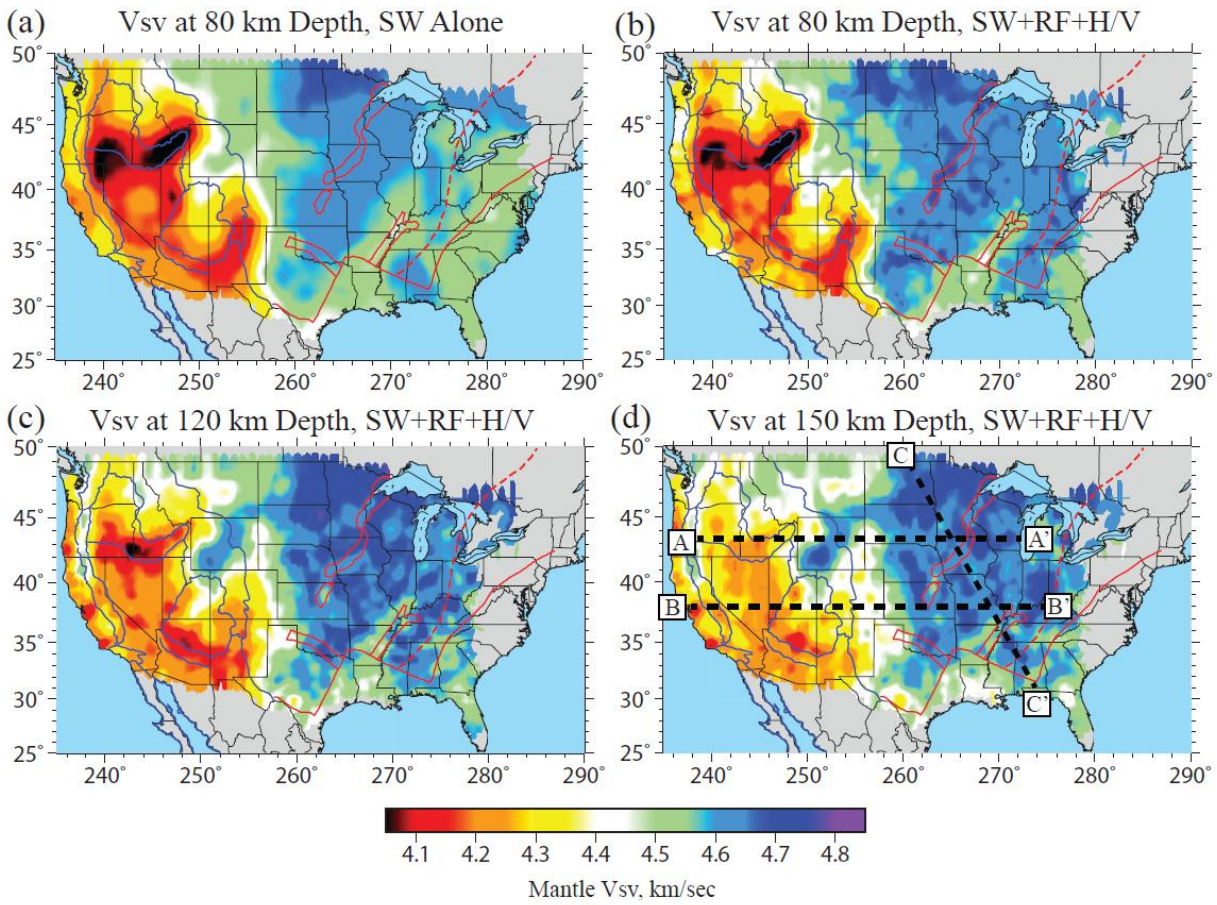

Figure 5.11. Mantle structures from the inversions. (a) average of Vsv between 70 and 90 km depths from the distribution of accepted models from the Monte Carlo inversion using surface wave dispersion alone. (b) Same as (a), but from the joint Monte Carlo inversion of surface wave dispersion, receiver functions, and H/V ratio. (c) Average of the model between 110 and 130 km depths. (d) Average of the model between 140 and 160 km depths.

On profile A-A', two major low velocities (<4.1km/sec) beneath the northern Basin and Range and the Snake River Plain are revealed between 60 - 120 km depth, and another resolved slow anomaly (~4.35 km/sec) beneath the Black Hills of South Dakota is shown at greater depths (>100 km). In this transect, A thick and fast lithosphere is observed beneath the Wyoming Craton, which is separated by the Black Hills low velocity from the lithosphere root beneath the central US. Beneath the Midcontinental Rift, the low velocities in the upper crust between the northern and southern segments of the rift are seen, with little variation in the uppermost mantle. On profile B-B', a eastward down-tilted slow velocity in the mantle beneath the Basin and Range/western Colorado Plateau is observed as before (Shen et al., 2013b). The fast velocity lid in the uppermost mantle produces the large velocity contrast across the Moho required by

receiver functions. Beneath the western Colorado Plateau, a fast lower crust and slow uppermost mantle is observed, consistent with the earlier observations (Levander et al. (2011) and Bailey et al. (2012)). Slow velocities are preserved beneath the Rocky Mountains from surface to over 150 km depth, indicating that the high topography is supported by both the crust and uppermost mantle. In the central/eastern part of the profile, the most pronounced anomaly is the relatively low velocities (~4.5km/sec) beneath the Reelfoot Rift embedded in the fast background mantle (~ 4.65km/sec). This anomaly, decaying at depth > 120km in this profile, is seen clearly in transect CC' (Fig. 12c), which cut through the Midcontinent Rift and the Reelfoot Rift directly. In this profile, a pronounced slow anomaly is presented beneath the Reelfoot Rift extending westward with depth. The fastest velocities are found beneath the Superior Province, which possesses the thinnest crust in the central US.

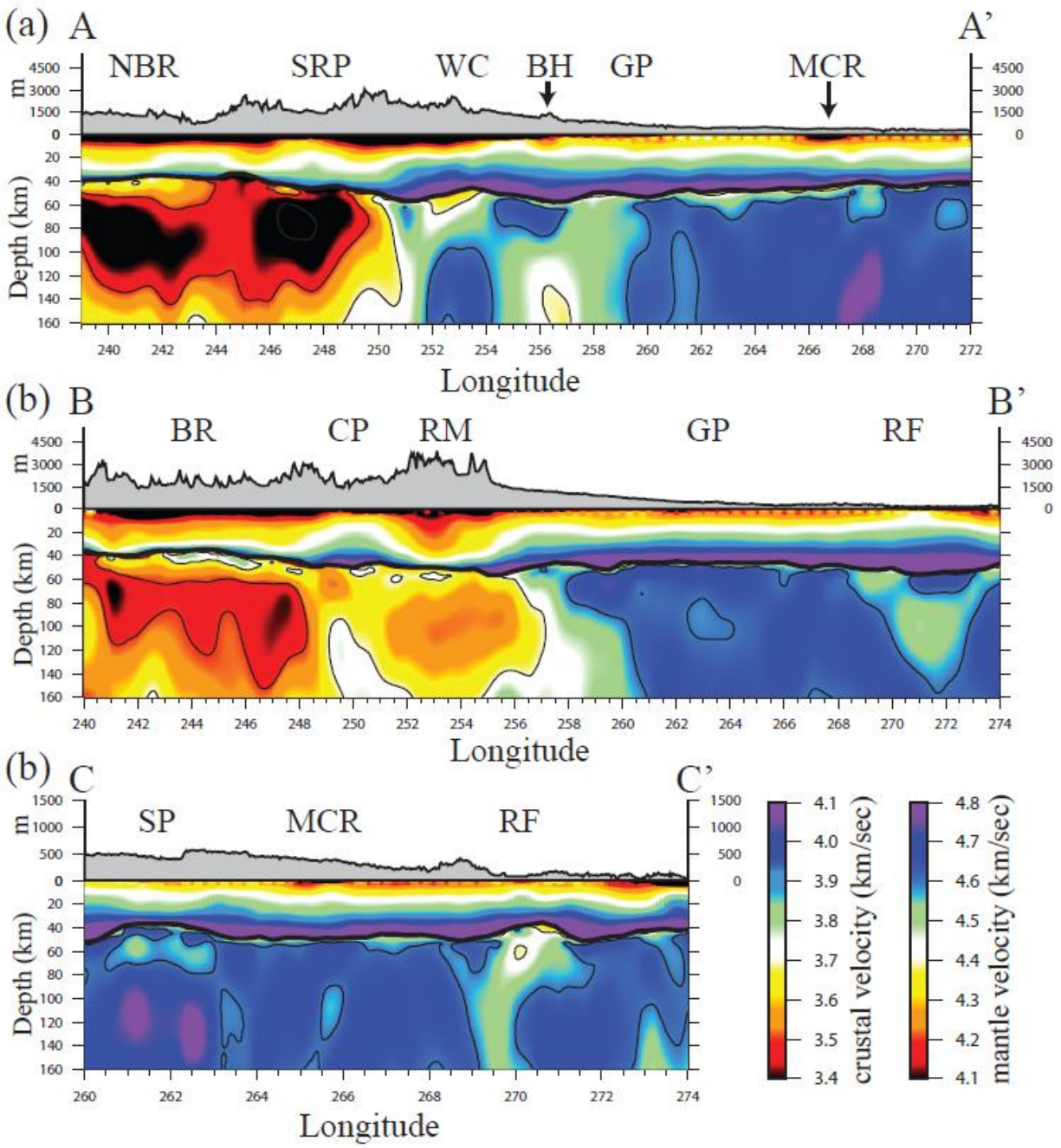

Figure 5.12 Vertical Vsv model profiles for the three transects identified in Fig. 3.10b with dashed lines. Speeds in the crust and uppermost mantle are plotted with absolute velocities using different color scales. Mantle velocity contours appear at 0.2 km/sec intervals. Surface topography is outlined above each transect. Tectonic/Geological features are identified with abbreviations: (N)BR: the (northern) Basin and Range; SRP: the Snake River Plains; WC: the Wyoming Craton; BH: the Black Hills; GP: the Great Plains; RF: the Reelfoot Rift; SP: the Superior Province; MCR: the Midcontinent Rift.

On average, the uppermost mantle structure is similar to previous models in the overlapping area (Yang and Ritzwoller, 2008; Moschetti et al., 2010, Shen et al., 2013a,b,c), and is also similar to models constructed from surface wave/body wave tomography by Schmandt and Humphreys (2010) and Obrebski et al. (2011). In this model, the assimilation of receiver functions greatly improves the vertical resolution in the top 100 km of the model, whereas body wave tomography reveals much deeper structure.

5.4.2 Model uncertainties

One of the advantages of the Bayesian Monte Carlo method is that it provides uncertainties in the final model, determined from the standard deviation of the resulting posterior distributions.. Figures 3.12 summarize the uncertainties for crustal thickness shown in Figures 3.10.

Figure 3.12c presents the spatial average of the standard deviation of these distributions across the study region for both the joint inversion and the inversion with surface wave data alone. The introduction of receiver functions reduces model uncertainties predominantly between depths of 35 and 50 km, by introducing information about the depth to Moho and the velocity contrast across it. This uncertainty, represented as the spread of the distribution, however, does not include bias. In fact, relative to the earlier studies of Yang et al. (2008) and Moschetti et al. (2010a,b) the improvements actually extend to greater depths because to reduce the spread of models near Moho those studies imposed a positivity constraint on the velocity derivative with depth, which is not needed in the joint inversion. In some locations that constraint obscured the location in depth of the mantle anomalies and its removal further improves the vertical resolution of the model. As discussed by Shen et al. (2013a,b), uncertainties of the crustal thickness is highly correlated with the Moho structure.

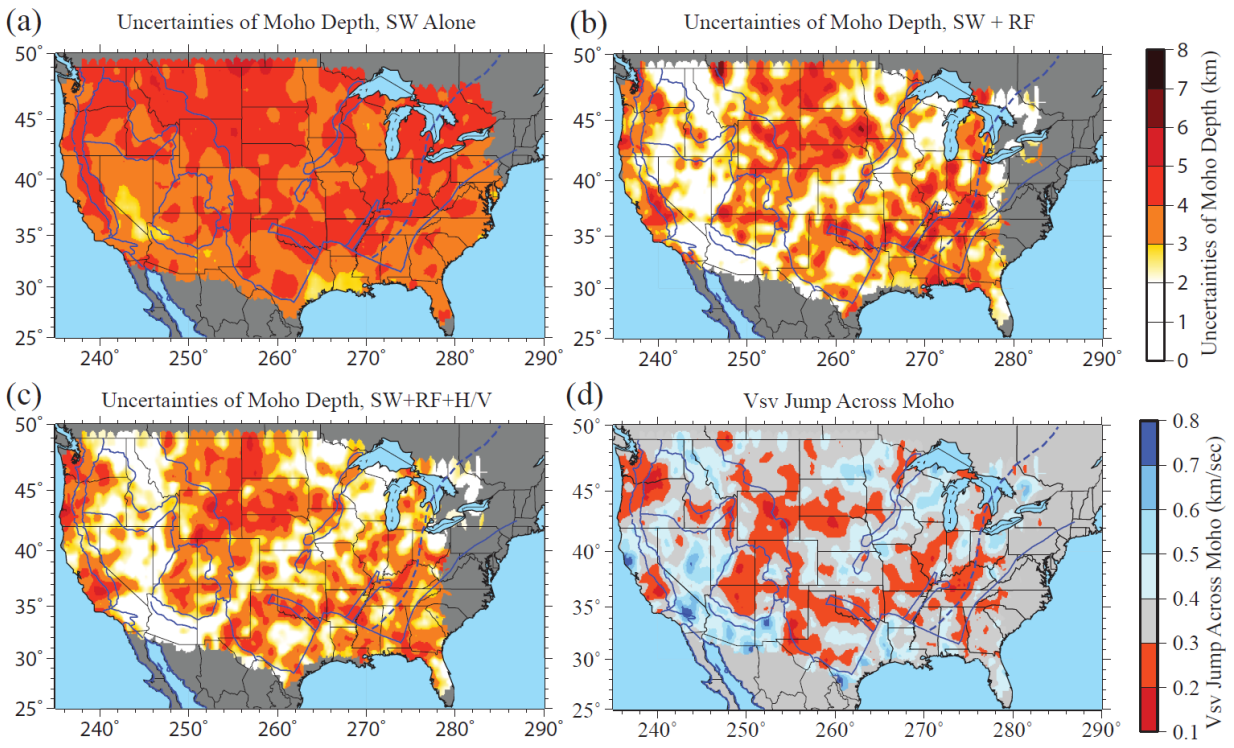

Figure 5.13. Uncertainties of Moho depth. (a) Uncertainties of Moho depth in the model constructed by surface wave dispersion alone. (b) Same as (a), but for the model constructed by joint inversion of surface wave dispersion and receiver functions. (c) Same as (a), but for the model constructed by joint inversion of surface wave dispersion, receiver functions and Rayleigh wave H/V ratio. (d) The map view of Vsv contrast across the Moho for the final Vs model.

Uncertainties in the top 5 km of the crust are about 100 -150 m/sec. These large values are due to the remaining trade-off between the sedimentary velocities and the sediments thickness. In the middle crust, uncertainties decrease to ~ 20 m/sec, with larger values distributed near the periphery of the map due to the larger uncertainties in the dispersion data. In the lower 1/3 of the crust, uncertainties grow up to ~60-80 m/sec due to remaining trade-offs with Moho depth and uppermost mantle velocities. At 80 km depth, uncertainties are about 50 m/s across much of the region, because this depth is separated from structural discontinuities across most of the region. In regions with thick crust, however, uncertainties increase to about 100 m/sec. At 120 and 150 km depth, uncertainties increase again because the structural sensitivity of surface waves decays with depth.

5.5 Summary and Conclusions

The Rayleigh wave H/V ratio, a new seismic measurement that is sensitive to shallow Vs structures, is estimated for the continental US using USArray/TA stations deployed from 2005 to May 2013. Surface wave data recorded by the same array are collected to construct phase velocity maps from ambient noise and teleseismic earthquake data using eikonal/Helmholtz tomography methods. Moreover, group velocity dispersion maps are also obtained using a traditional ambient noise tomography method to provide extra constraints on subsurface structures. Azimuthally independent receiver functions are also collected for the same period to add constraints on the subsurface discontinuities (e.g., Moho). To demonstrate the need for the H/V ratio measurements in the inversion, models that are only constrained by surface wave alone or jointly constrained by surface wave and receiver function datga are produced first. Both models predict the H/V ratios that fit the observations qualitatively but details are not matched quantitatively, particularly in major basin areas.

In this study, the Bayesian Monte Carlo inversion is extended to incorporate surface wave dispersion, receiver functions, and Rayleigh wave H/V ratios to invert for local Vs structures. In this approach, surface wave dispersion helps to constrain the shear wave speeds at depths separated from first-order discontinuities (e.g., the middle crustal velocity); receiver functions help to determine the structures near discontinuities (base of sediments, Moho); Rayleigh wave H/V ratios help to determine the local and shallow velocity structures. The three data sets are proven compatible and complementary with each other. Moreover, the joint Monte Carlo inversion produces a distribution of models beneath each station, which is summarized by the mean and standard deviation of the distribution at each depth. Across the vast majority of the region studied, the surface wave dispersion, receiver functions and H/V ratios can be fit jointly with a very simple model parameterization, producing a smooth model between the base of the sediments and Moho.

However, there are misfits in the Rayleigh wave H/V ratios that call for further investigation. For example, the accuracy of the Rayleigh wave H/V ratio measurements should be improved to further reduce the bias introduced by noise and 3-D wave effects. Further refinement of the 3-D model with perturbations of density should also be explored, but these improvements are beyond the scope of the present study. Moreover, to improve the model at shallow depths (<1km), the

recently obtained short period Rayleigh wave H/V ratios observed from ambient noise cross-correlations (Lin et al., 2014) should also be incorporated in the joint inversion. To improve estimates of the density structure, Rayleigh and perhaps Love wave local amplification should also be incorporated (Lin et al., 2012b).

A great variety of geological and tectonic features in the central/eastern US are better resolved in the 3-D model with constraints from H/V ratios and receiver functions. A pronounced low-velocity anomaly is observed in the uppermost mantle beneath the Reelfoot Rift, indicating the retained compositional signature of the Proterozoic rifting process. Other interesting features revealed by this model include the thickened crust to the east of the Grenville Front line and the relatively slow/neutral uppermost mantle beneath the Great Plains near the Rocky Mountain Front Range. These features call for more detailed local-to regional-scale analysis and interpretation in the future. This new 3-D model is useful as a basis for many other types of studies: studies of azimuthally/radially anisotropic structures, investigations of the density/thermal structure of the study region, and as a starting model for other types of seismic investigations such as Common-Conversion Point stacking and body wave tomography (e.g., Obrebski et al., 2011). In addition, the 3-D model can be used to improve regional seismic event location based on short-period surface waves, examples of which have been discussed by Barmin et al. (2011) and Levshin et al. (2012).

CHAPTER VI

CONCLUSION AND FUTURE PLAN

Synopsis of the Chapter

This chapter presents a summary of the thesis and is subdivided into three parts. In section 6.1, the major parts of the thesis are summarized and the conclusions of the thesis are presented. In section 6.2, other applications of the Bayesian Monte Carlo inversion of surface wave data that I developed are presented. In section 6.3, some future fruitful directions for further research are listed to conclude the thesis.

6.1 Summary and Conclusions

To better image the crust and uppermost mantle beneath the continents using large seismic arrays, I have developed a new approach that combines surface wave dispersion data and other geophysical observables including receiver functions. The development of the approach and its applications to the USArray/TA form the major component of the thesis.

Following three principle guidelines (the transparency of the assumptions; the simplicity of the model; and the usefulness of the results), the Monte Carlo inversion of multiple data sets is chosen as the defining characteristic of the new approach. Compared with traditional practices, this new approach possesses three significant advantages. First, by explicitly defining the prior distribution that underlies the Monte Carlo inversion, we ensure that the assumptions and prior constraints applied in the inversion are fully understood and clearly presented. Second, the method utilizes a simple model parameterization for the Monte Carlo inversion. This simple model parameterization 1) represents all and only the structures needed to fit the data, 2) ensures that the physically realistic models are generated, and 3) circumvents the need to apply the ad-hoc damping and smoothing constraints that result from the use of traditional thinly-layered parameterizations. Thir, the new approach rigorously 1) utilizes uncertainty information for surface wave dispersions estimated from eikonal/Helmholtz tomography, 2) constructs back-azimuth independent receiver functions with uncertainties using a harmonic stripping technique, and 3) infers propagates the data errors into model errors using a Bayesian Monte Carlo algorithm. The presentation of this uncertainty information substantially enhances the usefulness of the resulting 3-D models.

Once the development of the Bayesian Monte Carlo joint inversion was accomplished, we applied it to the USArray/TA to construct a set of 3-D crustal and uppermost mantle models for the continental US. The application is presented in four steps. In step 1, the Monte Carlo joint inversion of surface wave dispersion data and receiver functions is tested and proven successful using ~ 185 USArray/TA stations in the intermountain west. A 3-D crust and uppermost mantle is then constructed for this region. This step is summarized in Chapter II. In the second step, the method is applied to over 800 USArray/TA stations in the western US, and a 3-D model to a depth of 150 km is constructed. In this model a variety of geological/tectonic features are revealed in the crust and uppermost mantle beneath the western US. This step is described in Chapter III. In the third step, the Monte Carlo joint inversion of surface wave dispersion and receiver function data is applied to the central US encompassing the western arm of the Midcontinent Rift. Using ~ 122 USArray/TA stations, a 3-D model for the Midcontinent Rift and its neighboring areas is generated. The 3-D model and its tectonic implications are presented in Chapter IV. In the fourth and last step, the Bayesian Monte Carlo inversion is extended to incorporate the Rayleigh wave H/V data, which is sensitive to the Vs structure at shallow depths. The method is applied to over 1,700 USArray/TA stations, and a new 3-D Vs model is constructed for the crust and uppermost mantle beneath nearly the entire continental US. This step is presented in Chapter V and represents the culmination of my thesis.

The major conclusions drawn from the thesis are summarized here. We first present the conclusions related to the development of inversion techniques and then briefly summarize some conclusions that derive from the resulting 3-D models.

Technically, the thesis confirms that:

- The Bayesian Monte Carlo inversion is shown to be feasible and successful in inverting the surface wave dispersion data along with other geophysical data sets, including receiver functions and Rayleigh wave H/V ratio for local Vs structure. This method can be systematically applied to large arrays to constrain the crustal/uppermost mantle structures with unprecedented precision.

- Meaningful uncertainties of receiver functions can be inferred using the newly developed harmonic stripping technique. These uncertainties, together with the surface wave dispersion uncertainties estimated by the eikonal/Helmhotz tomography methods and the

uncertainties estimated for Rayleigh wave H/V ratio, can be meaningfully propagated into the uncertainties of the resulting models using the Bayesian Monte Carlo method.

- The model uncertainties inferred from the Monte Carlo inversion can help to better determine the stability and reliability of the 3-D model and may also help other scientists meaningfully utilize the 3-D model.

- The simple model parameterization has proven successful in the context of Bayesian Monte Carlo inversion with surface wave and other data. The smooth parameterization prohibits physically unrealistic features from being generated. For most stations in USArray/TA, the simple parameteriztion contains all and only the features needed to fit the data, but there are interesting exceptions. In some areas, a more complicated model parameterization is needed (e.g., a mid-crustal discontinuity is needed for some stations close to the MCR (Chapter IV); upper mantle Vs discontinuities may be needed by some stations in northern California (Chapter III)).

- By introducing receiver functions into the surface wave inversion, the sediment and Moho structures are better determined for many areas in the continental US. The Moho structure is improved under two situations. First, where the observed receiver functions contain clear Moho Ps phases, clear and sharp boundaries between the crust and uppermost mantle are obtained. Second, where the Moho Ps phase is absent, the Moho discontinuity is vague and a gradient transition between the crust and mantle will be obtained. For some areas where the Moho Ps phases are contaminated by reverberations generated from shallow structures, the Moho will be obscured, resulting in large uncertainties for the Vs model. For the areas where the Moho can be better determined, the Vs structures in the lower crust and uppermost mantle are also better imaged as the depth trade-off is largely reduced.

- By introducing the Rayleigh wave H/V ratio from teleseismic earthquakes into the joint inversion, shallow structures (<5 km) including sedimentary layers and uppermost crust are imaged better.

From the 3-D models constructed using the joint Monte Carlo inversion,several geological/tectonic implications can be drawn as well. In this thesis, we only discuss the detailed

tectonic implications for the central US encompassing the Midcontinent Rift area (MCR) (Chapter IV). As a result, the conclusions based on the resulting models concentrate principally in the central US, although the western US model also reveals many interesting features.

- There is a high correlation between the long wavelength gravity field and shallow Vs structure, especially in the central US. However, the MCR gravity high is obscured by clastic sediments in the shallow crust. This is consistent with an upper crustal origin to the MCR gravity anomaly as well as other anomalies in the region.

- Thick crust (>47 km) underlies the MCR, but there is a gradient Moho in the northern part of the rift and a sharp Moho in the south. Thickened crust beneath the MCR is evidence for post-rifting compression with pure-shear deformation along the entire rift; along-rift differences in lower crustal structure may signify magmatic underplating in the northern rift.

- Crustal shear wave speeds vary across the Precambrian sutures (e.g., Great Lakes Tectonic Zone, Spirit Lakes Tectonic Zone). This reveals that the Precambrian sutures played an important role in the subsequent tectonic evolution of the central US.

- In the uppermost mantle, the Reelfoot Rift preserves a slow anomaly at ~ 80 km, and the low velocity anomaly decays and spreads out at greater depths. This anomaly is probably a compositional signature resulting from the Proterozoic rifting process.

6.2 Other Applications of the Monte Carlo Inversion

As discussed in section 6.1, two major results are presented in this thesis: 1) the Bayesian Monte Carlo inversion algorithm for surface wave related problems and 2) a set of 3-D models for the continental US. The first, the Bayesian Monte Carlo inversion of surface wave and other geophysical data sets, has been applied not only to the USArray/TA, but also to other seismic networks around the world, for both on-land and ocean- bottom seismograph (OBS) arrays. These applications include the following.

Since 2010, a set of 3-D isotropic Vs models have been constructed for eastern Asia using large seismic arrays (e.g., CEArray) across that region. First, an isotropic 3-D model was constructed for northeastern China and the Sea of Japan. In this work, surface wave dispersion data using

ambient noise tomography are inverted by the Monte Carlo inversion algorithm to image the crust and uppermost mantle structure. This work was performed in collaboration with Dr. Yong Zheng from Chinese Academy of Sciences and is summarized in Zheng et al., (2011). Second, an isotropic 3-D model for he crust and uppermost mantle beneath south China region was produced using the Monte Carlo inversion. In this work, surface wave dispersion from ambient noise and teleseismic earthquakes were used. This work was performed in collaboration with Dr. Longquan Zhou from Chinese Earthquake Administration (CEA) and is summarized in Zhou et al. (2012). Third, Yang et al. (2012) applied the Monte Carlo inversion method to surface wave dispersion data from ambient noise to image the low velocity zone in the crust of the Tibetan Plateau. These models reveal many tectonic features in the lithosphere and set up a basis for the future application of Monte Carlo joint inversion of surface wave and receiver functions for this region.

Off the continents, the Bayesian Monte Carlo inversion also has been applied to an OBS array. Tian et al. (2013) applied the method to the Cascadia Initiative OBS array deployed on the Juan de Fuca plate. In this study, a 2-D crustal and uppermost mantle model near the Juan de Fuca ridge was constructed and a low velocity zone under the young (<1.5 Ma) Juan de Fuca plate is found between the 20 and 50 km depths, indicating the existence of a partial melt zone near the ridge.

Overall, the Monte Carlo inversion technique has been used by other studies for surface wave inversion exclusively. However, we expect many more applications of the joint inversion algorithm in the near future (e.g., a regional 3-D crust and uppermost mantle model is under construction for the Qinling-Dabie area in central China, Zheng et al., 2014) as more seismologists begin to combine the surface wave and azimuthally independent receiver functions using the Bayesian Monte Carlo method.

Another result presented in the thesis is a set of models for the crustal and uppermost mantle beneath the continental US. These models, together with the models constructed for Eastern Asia discussed above, are open to public access through the website “3D models from CU-Boulder” by Professor Michael H Ritzwoller at the University of Colorado at Boulder (http://ciei.colorado.edu/Models/).

Other scientists have used the resulting models to conduct various studies. Here, we briefly introduce some of the studies that have utilized the 3-D model for the western US (Shen et al., 2013b, Chapter III). For example, the 3-D western US model has been utilized by Boue et al. (2013) to validate the usability of the P delay times derived from cross-correlation techniques. In this study, the relative delay times of direct P phases derived from ambient noise cross-correlations were measured for the USArray/TA stations and then compared with our 3-D Vs model for the western US. The high correlation between the two quantities was used to validate the potential usefulness of the body wave phases to image earth structure. Another example is the work of Levandowski et al. (2014), which used the 3-D western US model to explore the major causes of surface topography in the western US. By converting the Vs model to density, they successfully interpreted the western US surface topography by the temperature, compositional and melt content variations. It should be noted, in addition, that the model ensembles from the Bayesian Monte Carlo inversion improved these authors' ability to estimate the uncertainties of other geophysical information (e.g., densities, surface topography), and thus helped them improve their geological/tectonic interpretations.

6.3 Future Work

This section describes some future plans that address the unfinished or ongoing work.

The first incomplete task is to produce a complete 3-D model for the continental US. As discussed in Chapters II through V, the joint Bayesian Monte Carlo inversion method has been applied to USArray/TA stations to construct a 3-D model for ~ 80% of the continental US. To obtain a complete and more accurate crustal and uppermost mantle model for the continental US, at least two projects should be carried out. First, the method should continue to be applied to the remainder of USArray/TA stations on the eastern edge of the United States as well as in Alaska, as these stations are installed. Second, to enhance the horizontal resolution, Earthscope/flexible arrays, such as the Superior Rifting Earthscope Experiment (SPREE), the Bighorns Arch Seismic Experiment (BASE) and the Seismic Investigation of Edge Driven Convection Associated with the Rio Grande Rift (CIEDCAR), should be integrated with the USArray/TA. These denser arrays will contribute at least two benefits to the resulting models. First, these arrays, with a station spacing ~ 30 km, allow the application of array-based receiver functions (e.g., CCP

stacking method), which will help to investigate finer structures within the crust (e.g. the mid-crustal discontinuities discussed in Chapter IV, Figure 4.10). Second, they may also help to improve the horizontal resolution of the surface wave maps, resulting in better local surface wave dispersion curves.

The second planned work is the development of the Bayesian Monte Carlo algorithm to perform thermodynamically constrained inversion for the continental US. Shapiro and Ritzwoller (2004) established such an inversion scheme that directly inverts seismic data (surface wave dispersion) to infer the thermal conditions of the lithosphere. A similar approach can be taken with the higher resolution surface wave dispersion data from USArray/TA and the harmonically stripped receiver functions.

The third planned work is the construction of a more accurate 3-D crustal and uppermost mantle model for Eastern Asia. As discussed in section 6.2, using the dense seismic arrays in China (CEArray) and Japan (F-net), we have produced a set of 3-D regional models with surface wave data alone. Similar to USArray/TA, seismic data from these arrays can be collected to produce high quality receiver functions as well, but only after careful data quality control. These receiver functions can be combined with surface wave dispersion to invert jointly for a more accurate 3-D model using the Bayesian Monte Carlo inversion algorithm.

The fourth work on the list is the expansion of the data sets used in the joint inversion. For instance, the receiver function waveform contains Moho reverberations that are discarded in the current joint inversion practice. These reverberations have proven helpful in determining the Vp/Vs ratio in the crust (Zhu and Kanamori, 2000). In the future, these reverberations should be incorporated in the joint inversion. Another seismic oberservable that will help image the crustal structure is the surface wave local amplification. As discussed by Lin et al. (2012), Rayleigh wave local amplification will help constrain the elastic velocity, density, and anelastic attenuation, and it certainly serves as a potential data set used by the Monte Carlo joint inversion. Another observable that can be integrated is the short period H/V (< 20 sec) estimated using ambient noise. This piece of data will help constrain the very shallow structure (<2 km) to an unprecedented level.

For the fifth planned work, we should further extend the Monte Carlo inversion to constrain the anisotropic structure beneath the continental US. To help constrain the radial anisotropy, Love wave dispersion information should be integrated into the joint inversion, while to help constrain the azimuthal anisotropy, Rayleigh wave azimuthal anisotropy dispersion measured from eikonal/Helmholtz tomography should be incorporated. Moreover, a new approach that utilizes ambient noise to better estimate the surface wave aizmuthal anisotropy at long periods (Yang et al., 2014) for both Rayleigh and Love waves should also be introduced.

BIBLIGRAPHY

Abers, G. A. (1998), Array measurements of phases used in receiver-function calculations: Importance of scattering, *Bulletin of the Seismological Society of America*., 88, 313-318

Ammon, C. J., G. E. Randall, and G. Zandt (1990), On the nonuniqueness of receiver function inversions, *J. Geophys. Res.*, 95(B10), 15,303–15,318, doi:10.1029/JB095iB10p15303.

Ammon, C.J. & Zandt, G. (1993), Receiver structure beneath the southern mojave block, California, *Bulletin of the Seismological Society of America*, 83, 737-755.

An, M., and M. S. Assumpção (2004), Multi-objective inversion of surface waves and receiver functions by competent genetic algorithm applied to the crustal structure of the Paraná Basin, SE Brazil, *Geophys. Res. Lett.*, 31, L05615, doi:10.1029/2003GL019179.

Bailey, I. W., M. S. Miller, K. Liu, and A. Levander (2012), V_S and density structure beneath the Colorado Plateau constrained by gravity anomalies and joint inversions of receiver function and phase velocity data, *J. Geophys. Res.*, 117, B02313, doi:10.1029/2011JB008522.

Bannister, S., Yu, J., Leitner, B. & Kennett, B.L.N., (2003). Variations in crustal structure across the transition from West to East Antarctica, Southern Victoria Land, *Geophysical Journal International*, 155, 870-884.

Bassin, C., Laske, G. and Masters, G., The Current Limits of Resolution for Surface Wave Tomography in North America, (2000). *EOS Trans AGU*, 81, F897.

Bastow, I. D., Stuart, G. W., Kendall, J., & Ebinger, C. J. (2005). Upper-mantle seismic structure in a region of incipient continental breakup: northern Ethiopian rift. *Geophysical Journal International*, *162*(2), 479-493.

Basuyau, C. & Tiberi, C. (2011), Imaging lithospheric interfaces and 3D structures using receiver functions, gravity, and tomography in a common inversion scheme, *Computers & Geosciences*, 37, 1381-1390.

Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., (2007). Processing seismic ambient noise data to obtain reliable broad-

band surface wave dispersion measurements, *Geophysical Journal International*, 169, 1239-1260.

Bensen, G. D., M. H. Ritzwoller, and N. M. Shapiro (2008), Broadband ambient noise surface wave tomography across the United States, *J. Geophys. Res.*, *113*, B05306, doi:10.1029/2007JB005248.

Bensen, G.D., M.H. Ritzwoller, and Y. Yang (2009), A 3D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise, *Geophysical Journal International*, 177(3), 1177-1196.

Bianchi, I., J. Park, N. Piana Agostinetti, and V. Levin (2010), Mapping seismic anisotropy using harmonic decomposition of receiver functions: An application to Northern Apennines, Italy, *J. Geophys. Res*., 115, B12317, doi:10.1029/2009JB007061.

Bodin, T., M. Sambridge, H. Tkalčić, P. Arroucau, K. Gallagher, and N. Rawlinson (2012), Transdimensional inversion of receiver functions and surface wave dispersion, *J. Geophys. Res.*, 117, B02301, doi:10.1029/2011JB008560.

Bostock, M. G. (1998), Mantle stratigraphy and evolution of the Slave province, J. Geophys. Res., 103(B9), 21,183–21,200, doi:10.1029/98JB01069.

Boyd, N. K., and S. B. Smithson (1994), Seismic profiling of Archean crust: Crustal structure in the Morton block, Minnesota River Valley subprovince, *Tectonophysics*, 232(1–4), 211–224, doi:10.1016/0040-1951(94)90085-X.

Braile, L.W., B. Wang, C.R. Daudt, G.R. Keller, and J.Pl. Patel (1994), Modelling the 2-D seismic velocity structure across the Kenya rift, *Tectonophysics*, 236, 217-249.

Brocher, T.M. (2005), Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. Seism. Soc. Am*., 95, 2081-2092, doi:10.1785.0120050077.

Cannon, W. F., A. G. Green, D. R. Hutchinson, M. Lee, B. Milkereit, J. C. Behrendt, H. C. Halls, J. C. Green, A. B. Dickas, G. B. Morey, R. Sutcliffe, and C. Spencer (1989), The North American Midcontinent Rift beneath Lake Superior from GLIMPCE seismic reflection profiling, *Tectonics*, 8(2), 305-332, doi: 10.1029/TC008i002p00305.

Cannon, W. F., and W. J. Hinze (1992), Speculations on the origin of the North American Midcontinent rift, *Tectonophysics*, 213(1–2), 49–55.

Cannon, W. F. (1994), Closing of the Midcontinent rift-A far—field effect of Grenvillian compression, *Geology*, *22*(2), 155–158.

Cassidy, J. F. (1992), Numerical experiments in broad-band receiver function-analysis, *Bulletin of the Seismological Society of America*, 82(3), 1453-1474.

Chandler, V. W., P. L. McSwiggen, G. B. Morey, W. J. Hinze, and R. R. Anderson (1989), Interpretation of Seismic Reflection, Gravity, and Magnetic Data Across Middle Proterozoic Mid-Continent Rift System, Northwestern Wisconsin, Eastern Minnesota, and Central Iowa, *AAPG Bulletin*, *73*(3), 261–275.

Chang, S.J., Baag, C.E. & Langston, C.A. (2004), Joint analysis of teleseismic receiver functions and surface wave dispersion using the genetic algorithm, *Bulletin of the Seismological Society of America*, 94, 691-704.

Christensen, N.I. & Mooney, W.D. (1995), Seismic velocity structure and composition of the continental crust: A global view, *J. Geophys. Res.*,100(B6): 9761–9788.

Clitheroe, G., O. Gudmundsson, and B. L. N. Kennett (2000), The crustal thickness of Australia, *J. Geophys. Res.*, 105(B6), 13,697–13,713, doi:10.1029/1999JB900317.

Crotwell, H. P., and T. J. Owens (2005), Automated receiver function processing, *Seismological Research Letters*, 76(6), 702-709.

Du, Z.J. & Foulger, G.R., (1999). The crustal structure beneath the northwest fjords, Iceland, from receiver functions and surface waves, *Geophysical Journal International*, 139, 419-432.

Dziewonski, A. and D. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Int.*, 25(4): 297–356.

Endrun, B., Meier, T., Bischoff, M. & Harjes, H.P. (2004), Lithospheric structure in the area of Crete constrained by receiver functions and dispersion analysis of Rayleigh phase velocities, *Geophysical Journal International*, 158, 592-608.

Frassetto, A. M., G. Zandt, H. Gilbert, T. J. Owens, and C. H. Jones (2011), Structure of the

Sierra Nevada from receiver functions and implications for lithospheric foundering, *Geosphere*, 7, 898-921.

Frederiksen, A.W., Folsom, H. & Zandt, G., (2003). Neighbourhood inversion of teleseismic Ps conversions for anisotropy and layer dip, *Geophysical Journal International*, 155, 200-212.

French, S.W., K. M. Fischer, E. M. Syracuse, and M. E. Wysession (2009), Crustal structure beneath the Florida-to-Edmonton broadband seismometer array, *Geophys. Res. Lett.*, 36, L08309, doi: 10.1029/2008GL036331.

Furlong, K. P., & Fountain, D. M. (1986). Continental crustal underplating: Thermal considerations and seismic-petrologic consequences. *Journal of Geophysical Research: Solid Earth* (1978–2012), 91(B8), 8285-8294.

Gao, W., Grand, S.P., Baldridge, W.S., Wilson, D., West, M., Ni, J.F. & Aster, R., 2004. Upper mantle convection beneath the central Rio Grande rift imaged by P and S wave tomography, *Journal of Geophysical Research-Solid Earth*, 109, 16.

Gilbert, H. (2012), Crustal structure and signatures of recent tectonism as influenced by ancient terranes in the western United States, *Geosphere*, 8, 141-157, doi:10.1130/GES00720.1.

Gilbert, H. J., and A. F. Sheehan (2004), Images of crustal variations in the intermountain west, *J. Geophys. Res*., 109, B03306, doi:10.1029/2003JB002730.

Girardin, N. and V. Farra (1998), Azimuthal anisotropy in the upper mantle from observations of P-to-S converted phases: application to southeast Australia, *Geophysical Journal International*, 133, 615-629.

Gok, R., Pasyanos, M.E. & Zor, E. (2007), Lithospheric structure of the continent-continent collision zone: eastern Turkey, *Geophysical Journal International*, 169, 1079-1088.

Gorman, A. R., Clowes, R. M., Ellis, R. M. et al., (2002), Deep Probe: Imaging the roots of western North America, *Can. Journ. Earth Sci*., 39(3), 375-398.

Hales, T.C., Abt, D.L., Humphreys, E.D. & Roering, J.J., 2005. A lithospheric instability origin for Columbia River flood basalts and Wallowa Mountains uplift in northeast Oregon, *Nature*, 438, 842-845.

Hammer, P., R. M. Clowes, F. A. Cook, A. J. van der Velden, and K. Vasudevan (2010), The Lithoprobe trans-continental lithospheric cross sections: imaging the internal structure of the North American continent, *Can. J. Earth Sciences*, 47(5), 821-857, doi: 10.1139/E10-036.

Hetenyi, G. & Bus, Z., (2007). Shear wave velocity and crustal thickness in the Pannonian Basin from receiver function inversions at four permanent stations in Hungary, *Journal of Seismology*, 11, 405-414.

Hinze, W. J., D. J. Allen, A. J. Fox, D. Sunwood, T. Woelk, and A. G. Green (1992), Geophysical investigations and crustal structure of the North American Midcontinent Rift system, *Tectonophysics*, 213(1–2), 17–32, doi:10.1016/0040-1951(92)90248-5.

Hinze, W. J., D. J. Allen, L. W. Braile, and J. Mariano (1997), The Midcontinent Rift System: A major Proterozoic continental rift, *Geological Society of America Special Papers*, *312*, 7–35, doi:10.1130/0-8137-2312-4.7.

Hollings, P., M. Smyk, L. H. Heaman, and H. Halls (2010), The geochemistry, geochronology and paleomagnetism of dikes and sills associated with the Mesoproterozoic Midcontinent Rift near Thunder Bay, Ontario, Canada, *Precambrian Research*, 183(3), 553-571, doi: 10.1016/j.precamres.2010.01.012.

Holm, D. K., R. Anderson, T. J. Boerboom, W. F. Cannon, V. Chandler, M. Jirsa, J. Miller, D. A. Schneider, K. J. Schulz, and W. R. Van Schmus (2007), Reinterpretation of Paleoproterozoic accretionary boundaries of the north-central United States based on a new aeromagnetic-geologic compilation, *Precambrian Research*, 157(1–4), 71–79, doi:10.1016/j.precamres.2007.02.023.

Horspool, N.A., Savage, M.K. & Bannister, S. (2006), Implications for intraplate volcanism and back-arc deformation in northwestern New Zealand, from joint inversion of receiver functions and surface waves, *Geophysical Journal International*, 166, 1466-1483.

Jones, C. H., and R. A. Phinney (1998), Seismic structure of the lithosphere from teleseismic converted arrivals observed at small arrays in the southern Sierra Nevada and vicinity, California, *J. Geophys. Res*., 103(B5), 10,065–10,090, doi:10.1029/97JB03540.

Jordan, T. H. (1979). Mineralogies, densities and seismic velocities of garnet lherzolites and their geophysical implications, in The Mantle Sample: Inclusions in Kimberlites and Other Volcanics, *Proceedings of the Second Internaitonal Kimberlite Conference*, vol. 2, edited by F.R. Boyd and H.O. Meyer, .Special Publications, 16, pp. 1-14, AGU, Washington, D.C..

Julia, J., Ammon, C.J. & Herrmann, R.B. (2003), Lithospheric structure of the Arabian Shield from the joint inversion of receiver functions and surface-wave group velocities, *Tectonophysics*, 371, 1-21.

Julia, J., Ammon, C.J., Herrmann, R.B. & Correig, A.M. (2000), Joint inversion of receiver function and surface wave dispersion observations, Geophysical Journal International, 143, 99-112.

Kaban, M.K., Schwintzer, P., Artemieva, I.M. & Mooney, W.D., 2003. Density of the continental roots: compositional and thermal contributions, *Earth and Planetary Science Letters*, 209, 53-69.

Kanamori, H. and D. Anderson (1977), Importance of physical dispersion in surface wave and free oscillation problems : Review, Revs. Geophys. Space Phys., 15(1):105-112.

Karato, S. (1993), Importance of anelasticity in the interpretation of seismic tomography, Geophys. Res. Lett., 20(15), 1623–1626, doi:10.1029/93GL01767.

Karlstrom, K.E., Whitmeyer, S.J., Dueker, K., Williams, M. L., Bowring, S., Levander, A., Humphreys, E. D., Keller, G. R., CD-Rom Working Group, 2005, Synthesis of results from the CD-ROM experiment: 4-D image of the lithosphere beneath the Rocky Mountains and implications for understanding the evolution of continental lithosphere, in: Karlstrom, K. E., and Keller, G. R., eds.,The Rocky Mountain region - an evolving lithosphere. *AGU geophysical monograph*, 154, DOI: 10:1029/154GM31.

Karlstrom, K.E., Coblentz, D., Dueker, K., Ouimet, W., Kirby, E., Van Wijk, J., Schmandt, B., Kelley, S., Lazear, G., Crossey, L.J., Crow, R., Aslan, A., Darling, A., Aster, R., MacCarthy, J., Hansen, S.M., Stachnik, J., Stockli, D.F., Garcia, R.V., Hoffman, M., McKeon, R., Feldman, J., Heizler, M., Donahue, M.S. & Grp, C.W., 2012. Mantle-driven dynamic uplift of the Rocky Mountains and Colorado Plateau and its surface response: Toward a unified hypothesis, *Lithosphere*, 4, 3-22.

Khan, A., Zunino, A. and Deschamps, F. (2011) The thermo-chemical and physical structure beneath the North American continent from Bayesian inversion of surface-wave phase velocities , *J. Geophys. Res*., 116, B09304, doi:10.1029/2011JB008380.

Langston, C. A. (1979), Structure Under Mount Rainier, Washington, Inferred From Teleseismic Body Waves, *J. Geophys. Res.*, 84(B9), 4749–4762, doi:10.1029/JB084iB09p04749.

Last, R. J., A. A. Nyblade, C. A. Langston, and T. J. Owens (1997), Crustal structure of the East African Plateau from receiver functions and Rayleigh wave phase velocities, *J. Geophys. Res.*, 102(B11), 24,469–24,483, doi:10.1029/97JB02156.

Lawrence, J.F. & Wiens, D.A., (2004). Combined receiver-function and surface wave phase-velocity inversion using a niching genetic algorithm: Application to Patagonia, Bulletin of the Seismological Society of America, 94, 977-987.

Lebedev, S., et al., (2013), Mapping the Moho with seismic surface waves: A review, resolution analysis, and recommended inversion strategies, *Tectonophysics*, http://dx.doi.org/10.1016/j.tecto.2012.12.030.

Levander, A., B. Schmandt, M. S. Miller, K. Liu, K. E. Karlstrom, R. S. Crow, C. T. A. Lee, and E. D. Humphreys (2011), Continuing Colorado plateau uplift by delamination-style convective lithospheric downwelling, *Nature*, 472(7344), 461-U540.

Levandowski, W., C. Jones, W. Shen, M.H. Ritzwoller and V. Schulte-Pelkum, Origins of topography in the Western US: Mapping crustal and upper mantle density variations using a uniform seismic velocity model, *J. Geophys. Res*., doi: 10.1002/2013JB010607, in press, 2014.

Levin, V. & Park, J., (1997). P-SH conversions in a flat-layered medium with anisotropy of arbitrary orientation, *Geophysical Journal International*, 131, 253-266.

Levshin, A.L. and M. H. Ritzwoller (1995), Characteristics of surface waves generated by events on and near the Chinese nuclear test site, *Geophys. J. Int*., 123, 131-149.

Levshin, A.L., M.P. Barmin, M.P. Moschetti, C. Mendoza, and M.H. Ritzwoller (2012), Refinements to the method of epicentral location based on surface waves from ambient seismic noise: Introducing Love waves, *Geophys. J. Int.*, submitted.

Levin, V., N.M. Shapiro, J. Park, and M.H. Ritzwoller (2002), Seismic evidence for catastrophic slab loss beneath Kamchatka, *Nature*, 418, 763-767.

Levshin, A.L., M.H. Ritzwoller, and N.M. Shapiro (2005), The use of crustal higher modes to constrain crustal structure across Central Asia, *Geophys. J. Int.*, 160, 961-972.

Liang, C., Langston, C. A. (2008), Ambient seismic noise tomography and structure of eastern North America. *J. Geophys. Res.* 113, B03309, doi:10.1029/2007JB005350.

Ligorria, J. P., and C. J. Ammon (1999), Iterative deconvolution and receiver-function estimation, *Bulletin of the Seismological Society of America*, 89(5), 1395-1400.

Lin, F.C., Moschetti, M.P. & Ritzwoller, M.H., (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophysical Journal International*, 173, 281-298.

Lin, F.C., M.H. Ritzwoller, Y. Yang, M.P. Moschetti, and M.J. Fouch (2011), Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States, *Nature Geoscience*, Vol 4, Issue 1, 55-61.

Lin, F.C., M.H. Ritzwoller, and W. Shen, On the reliability of attenuation measurements from ambient noise cross-correlations, *Geophys. Res. Letts*., 38, L11303, doi:10.1029/2011GL047366, 2011.

Lin, F.C., Ritzwoller, M.H. & Snieder, R., (2009). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array, *Geophysical Journal International*, 177, 1091-1110.

Lin, F.C. & Ritzwoller, M.H., (2011). Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, Geophysical Journal International, 186, 1104-1120.

Lin, F.C.,V. Tsai, and M.H. Ritzwoller (2012a), The local amplification of surface waves: A new observable to constrain elastic velocities, density, and anelastic attenuation, *J. Geophys. Res*., 117, B06302, doi:10.1029/2012JB009208.

Lin, F.-C.,B. Schmandt, and V.C. Tsai (2012b), Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: constraining velocity and density structure in the upper crust, *Geophys. Res. Letts.*, 39, L12303, doi:10.1029/2012GL052196, 2012

Liu, Q.Y., Li, Y., Chen, J.H., van der Hilst, R.D., Guo, B.A., Wang, J., Qi, S.H. & Li, S.C., (2010). Joint inversion of receiver function and ambient noise based on Bayesian theory, *Chinese Journal of Geophysics-Chinese Edition*, 53, 2603-2612.

Lowry, A. R., & Pérez-Gussinyé, M. (2011). The role of crustal quartz in controlling Cordilleran deformation. *Nature*, 471(7338), 353-357.

Lucente, F. P., N. Piana Agostinetti, M. Moro, G. Selvaggi, and M. Di Bona (2005), Possible fault plane in a seismic gap area of the southern Apennines (Italy) revealed by receiver function analysis, *J. Geophys. Res.*, 110, B04307, doi:10.1029/2004JB003187.

Maraschini, M. & Foti, S., 2010. A Monte Carlo multimodal inversion of surface waves, *Geophysical Journal International*, 182, 1557-1566.

Mariano, J., and W. J. Hinze (1994), Structural interpretation of the Midcontinent Rift in eastern Lake Superior from seismic reflection and potential-field studies, *Canadian Journal of Earth Sciences*, *31*(4), 619–628, doi:10.1139/e94-055.

McKenzie, D. (1978), Some remarks on the development of sedimentary basins, *Earth Planet. Sci. Lett.*, 22, 108-125.

Megnin, C., and B. Romanowicz (2000), The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophysical Journal International*, 143(3), 709-728.

Moidaki, M., Gao, S. S., Liu, K. H., Atekwana, E. (2013), Crustal thickness and Moho sharpness beneath the Midcontinent rift from receiver functions, *Research in Geophysics*, doi:10.4081/rg.2013.e1

Molnar, S., S.E. Dosso and J.F. Cassidy, (2010). Bayesian inversion of microtremor array dispersion data in southwestern British Columbia, *Geophys. J. Int.*, 183, 923-940.

Mooney, W. D., and M. K. Kaban (2010), The North American upper mantle: Density, composition, and evolution, *J. Geophys. Res*., 115, B12424, doi:10.1029/2010JB000866.

Moorkamp, M., A. G. Jones, and S. Fishwick (2010), Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data, *J. Geophys. Res*., 115, B04318, doi:10.1029/2009JB006369.

Moschetti, M.P., M.H. Ritzwoller, and F.C. Lin (2010), Seismic evidence for widespread crustal deformation caused by extension in the western USA, *Nature*, 464, Number 7290, 885-889.

Moschetti, M. P., M. H. Ritzwoller, and N. M. Shapiro (2007), Surface wave tomography of the western United States from ambient seismic noise: Rayleigh wave group velocity maps, *Geochem. Geophys. Geosyst*., 8, Q08010, doi:10.1029/2007GC001655.

Moschetti, M. P., M. H. Ritzwoller, F. C. Lin, and Y. Yang (2010a), Crustal shear wave velocity structure of the western United States inferred from ambient seismic noise and earthquake data, *Journal of Geophysical Research-Solid Earth*, 115.

Moschetti, M.P., M.H. Ritzwoller, F.C. Lin, and Y. Yang, Crustal shear velocity structure of the western US inferred from amient noise and earthquake data (2010b), *J. Geophys. Res*., 115, B10306, doi:10.1029/2010JB007448.

Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res*., 100(B7), 12,431–12,447, doi:10.1029/94JB03097.

Mosegaard K. and Sambridge M. 2002. Monte Carlo analysis of inverse problems. *Inverse Problems* 18, 29–54.

Nielsen, C. A. and H. Thybo (2009), No Moho uplift below the Baikal Rift Zone: Evidence from a seismic refraction profile across southern Lake Baikal, *Journal of Geophysical Research*, 114, B08306, doi:10.1029/2008JB005828.

Nicholson, S. W., K. J. Schulz, S. B. Shirey, and J. C. Green (1997), Rift-wide correlation of 1.1 Ga Midcontinent rift system basalts: implications for multiple mantle sources during rift development, *Canadian Journal of Earth Sciences*, 34(4), 504–520, doi:10.1139/e17-041.

Nicholson, T., Bostock, M. & Cassidy, J.F., (2005). New constraints on subduction zone structure in northern Cascadia, *Geophysical Journal International*, 161, 849-859.

Nyblade, A.A. and R.A. Brazier (2002), Precambrian lithospheric controls on the development on the development of the East African Rift system, *Geology* 30(8), 755-758.

Obrebski, M., Allen, R.M., Pollitz, F. & Hung, S.H., 2011. Lithosphere-asthenosphere interaction beneath the western United States from the joint inversion of body-wave traveltimes and surface-wave phase velocities, *Geophysical Journal International*, 185, 1003-1021.

Ozalaybey, S., Savage, M.K., Sheehan, A.F., Louie, J.N. & Brune, J.N., (1997). Shear-wave velocity structure in the northern Basin and Range province from the combined analysis of receiver functions and surface waves, *Bulletin of the Seismological Society of America*, 87, 183-199.

Paces, J. B., & Bell, K. (1989). Non-depleted sub-continental mantle beneath the Superior Province of the Canadian Shield: Nd-Sr isotopic and trace element evidence from Midcontinent Rift basalts. *Geochimica et Cosmochimica Acta*,53(8), 2023-2035.

Pasyanos, M.E., Tkalcic, H., Gok, R., Al-Enezi, A. & Rodgers, A.J., (2007). Seismic structure of Kuwait, *Geophysical Journal International*, 170, 299-312.

Pavlis, N. K., Holmes, S. A., Kenyon, S. C., & Factor, J. K. (2012). The development and evaluation of the Earth Gravitational Model 2008 (EGM2008).*Journal of Geophysical Research: Solid Earth* (1978–2012), 117(B4).

Piana Agostinetti, N., F. P. Lucente, G. Selvaggi, and M. Di Bona (2002), Crustal Structure and Moho Geometry beneath the Northern Apennines (Italy), *Geophys. Res. Lett.*, 29(20), 1999, doi:10.1029/2002GL015109.

Piana Agostinetti, N. & Chiarabba, C., 2008. Seismic structure beneath Mt Vesuvius from receiver function analysis and local earthquakes tomography: evidences for location and geometry of the magma chamber, *Geophysical Journal International*, 175, 1298-1308.

Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophysical Journal International*, 181, 858-872.

Pollitz, F.F. (2008), Observations and interpretation of fundamental mode Rayleigh wavefields recorded by the Transportable Array (USArray), *Geophys. J. Int.*, 173,189-204.

Pollitz, F. F. & Snoke, J. A. (2010), Rayleigh-wave phase-velocity maps and three dimensional shear velocity structure of the western US from local non-plane surface wave tomography. *Geophys. J. Int.*, 180, 1153–1169.

Pollitz, F.F. and W.D. Mooney (2013), Mantle origin for stress concentration in the New Madrid seismic zone, *Earth Planet. Sci. Letts*., submitted.

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

Ritzwoller, M.H., A.L. Levshin, L.I. Ratnikova, and A.A. Egorkin (1998), Intermediate period group velocity maps across Central Asia, Western China, and parts of the Middle East, *Geophys. J. Int.*, 134, 315-328.

Ritzwoller, M. H., N. M. Shapiro, A. L. Levshin, and G. M. Leahy (2001), Crustal and upper mantle structure beneath Antarctica and surrounding oceans, *J. Geophys. Res.*, 106(B12), 30,645–30,670, doi:10.1029/2001JB000179.

Ritzwoller, M.H., N.M. Shapiro, and G.M. Leahy, A resolved mantle anomaly as the cause of the Australian-Antarctic Discordance (2003), *J. Geophys. Res.*, 108, no. B12, 2559, doi:10.1029/2003JB002522.

Ritzwoller, M.H., N.M. Shapiro, S. Zhong (2004), Cooling history of the Pacific lithosphere, *Earth Planet. Sci. Letts.*, 226, 69-84.

Ritzwoller, M.H., F.C. Lin, and W. Shen, Ambient noise tomography with a large seismic array, *Compte Rendus Geoscience*, (2011), 13 pages, doi:10.1016/j.crte.2011.03.007.

Roy, M., T. H. Jordan, and J. Pederson (2009), Colorado Plateau magmatism and uplift by warming of heterogeneous lithosphere, *Nature*, 459(7249), 978-U102.

Salah, M.K., Chang, S.J. & Fonseca, J., 2011. Crustal structure beneath the Lower Tagus Valley, southwestern Iberia using joint analysis of teleseismic receiver functions and surface-wave dispersion, *Geophysical Journal International*, 184, 919-933.

Sambridge, M., 1999a. Geophysical inversion with a neighbourhood algorithm - I. Searching a parameter space, *Geophysical Journal International*, 138, 479-494.

Sambridge, M., 1999b. Geophysical inversion with a neighbourhood algorithm - II. Appraising the ensemble, *Geophysical Journal International*, 138, 727-746.

Sambridge M. 2001. Finding acceptable models in nonlinear inverse problems using a neighbourhood algorithm. *Inverse Problems* 17, 387–403.

Sambridge M. and Mosegaard K. 2002. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics* 40, 1–29

Saltzer, R. L., & Humphreys, E. D. (1997). Upper mantle P wave velocity structure of the eastern Snake River Plain and its relationship to geodynamic models of the region. *Journal of Geophysical Research: Solid Earth* (1978–2012), 102(B6), 11829-11841.

Savage, M. K. (1998), Lower crustal anisotropy or dipping boundaries? Effects on receiver functions and a case study in New Zealand, *J. Geophys. Res.*,103, 15069-15087.

Schultz, C.A., Myers, S.C., Hipp, J. & Young, C.J., 1999. Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location and identification, *Physics of the Earth and Planetary Interiors*, 113, 321-338.

Schmandt, B. & Humphreys, E., 2010. Complex subduction and small-scale convection revealed by body-wave tomography of the western United States upper mantle, *Earth and Planetary Science Letters*, 297, 435-445.

Schmus, W. R. (1992), Tectonic setting of the Midcontinent Rift system, *Tectonophysics*, 213(1–2), 1–15, doi:10.1016/0040-1951(92)90247-4.

Schutt, D. L., & Lesher, C. E. (2006). Effects of melt depletion on the density and seismic velocity of garnet and spinel lherzolite. *Journal of Geophysical Research: Solid Earth* (1978–2012), 111(B5).

Shapiro, N.M., Campillo, M., Stehly, L. & Ritzwoller, M.H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, 307, 1615-1618.

Shapiro, N.M. & M. H. Ritzwoller, 2002. Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophysical Journal International*, 151, 88-105.

Shapiro, N.M. & M.H. Ritzwoller, 2004. Thermodynamic constraints on seismic inversions, *Geophys. J. Int*., 157, 1175-1188, doi:10.1111/j.1365-246X.2004.02254.x.

Shapiro, N. M., M. H. Ritzwoller, and E. R. Engdahl (2008), Structural context of the great Sumatra-Andaman Islands earthquake, *Geophys. Res. Lett*., 35, L05301, doi:10.1029/2008GL033381.

Sheehan, A. F., C. H. Jones, M. K. Savage, S. Ozalaybey, and J. M. Schneider (1997), Contrasting lithospheric structure between the Colorado Plateau and Great Basin: Initial results from Colorado Plateau - Great Basin PASSCAL Experiment, *Geophys. Res. Lett.*, 24(21), 2609–2612, doi:10.1029/97GL02782.

Shen, W., M.H. Ritzwoller, V. Schulte-Pelkum, F.-C. Lin (2013a), Joint inversion of surface wave dispersion and receiver functions: A Bayesian Monte-Carlo approach, *Geophys. J. Int.*, 192, 807-836, doi:10.1093/gji/ggs050.

Shen, W., M.H. Ritzwoller, and V. Schulte-Pelkum (2013b), A 3-D model of the crust and uppermost mantle beneath the central and western US by joint inversion of receiver functions and surface w QASZXDXCFave dispersion, *J. Geophys. Res.*, 118, 1-15, doi:10.1029/2012JB009602,

Shen, W., M.H. Ritzwoller, and V. Schulte-Pelkum (2013c), Crustal and uppermost mantle structure in the central US encompassing the Midcontinent Rift, *J. Geophys. Res*., 118, 4325-4344, doi:10.1002/jgrb.50321, 2013.

Shibutani, T., M. Sambridge, and B. Kennett (1996), Genetic algorithm inversion for receiver functions with application to crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.*, 23(14), 1829–1832, doi:10.1029/96GL01671.

Sims, P. K., and Z. E. Petermar (1986), Early Proterozoic Central Plains orogen: A major buried structure in the north-central United States, *Geology*, *14*(6), 488–491.

Socco, L.V. and D. Boiero, Improved Monte Carlo inversion of surface wave data (2008), *Geophys. Prospecting*, 56, 357-371.

Stein, S. (2011), Learning from failure: The SPREE Mid-Continent Rift Experiment, *GSA Today*, 21(9), 5–7, doi:10.1130/G120A.1.

Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys*., 20(2), 219–232, doi:10.1029/RG020i002p00219.

Thybo, H. and C. A. Nielsen (2009), Magma-compensated crustal thinning in continental rift zones, *Nature*, 457, 873-876, doi:10.1038/nature07688.

Tkalčić, H., M. E. Pasyanos, A. J. Rodgers, R. Gök, W. R. Walter, and A. Al-Amri (2006), A multistep approach for joint modeling of surface wave dispersion and teleseismic receiver functions: Implications for lithospheric structure of the Arabian Peninsula, *J. Geophys. Res.*,111, B11311, doi:10.1029/2005JB004130.

Tokam, A.P.K., Tabod, C.T., Nyblade, A.A., Julia, J., Wiens, D.A. & Pasyanos, M.E. (2010), Structure of the crust beneath Cameroon, West Africa, from the joint inversion of Rayleigh wave group velocities and receiver functions, *Geophysical Journal International*, 183, 1061-1076.

Van Schmus, W.R., Bickford, M.E., Anderson, J.L., Bender, E.E., Anderson, R.R., Bauer, P.W., Robertson, J.M., Bowring, S.A., Condie, K.C., Denison , R.E., Gilbert, M.C., Grambling, J.A., Mawer, C.K., Shearer, C.K., Hinze, W.J., Karlstrom, K.E., Kisvarsanyi, E.B., Lidiak, E.G., Reed, J.C., Sims, P.K., Jr., Tweto, O., Silver, L.T., Treves, S.B., Williams, M.L., and Wooden, J.L. (1993), Transcontinental Proterozoic provinces, in Reed, yellwstJ.C., Jr., Bickford, M.E., Houston, R.S., Link, P.K., Rankin, D.W., Sims, P.K., and Van Schmus, W.R., eds., Precambrian: Conterminous U.S.: Boulder, Colorado, *Geological Society of America*, *Geology of North America*, v. C-2, p. 171-334.

Vervoort, J. D., K. Wirth, B. Kennedy, T. Sandland, and K. S. Hatpp (2007), The magmatic evolution of the Midcontinent rift: New geochronologic and geochemical evidence from felsic magmatism, *Precambrian Research*, 157(1-4), 235-268, doi: 10.1016/j.precamres.2007.02.019.

Villasenor, A., M.H. Ritzwoller, A.L. Levshin, M.P. Barmin, E.R. Engdahl, W. Spakman, and J. Trampert (2001), Shear velocity structure of Central Eurasia from inversion of surface wave velocities, *Phys. Earth Planet. Int.*, 123(2-4), 169 - 184.

Vinnik, L.P., 1977. Detection of waves converted from P to SV in the mantle, *Phys. Earth Planet. Inter.*, 15, 39-45.

Vinnik, L.P., Aleshin, I.M., Kaban, M.K., Kiselev, S.G., Kosarev, G.L., Oreshin, S.I. & Reigber, C., 2006. Crust and mantle of the Tien Shan from data of the receiver function tomography, *Izvestiya-Physics of the Solid Earth*, 42, 639-651.

Vinnik, L.P., Reigber, C., Aleshin, I.M., Kosarev, G.L., Kaban, M.K., Oreshin, S.I. & Roecker, S.W., 2004. Receiver function tomography of the central Tien Shan, *Earth and Planetary Science Letters*, 225, 131-146.

Wernicke, B. P. (1985), Uniform-sense normal simple shear of the continental lithosphere, *Can. J. Earth Sci*., 22, 108-125.

West, M., Gao, W., S. Grand (2004), A simple approach to the joint inversion of seismic body and surface waves applied to the southwest U.S. *Geophys. Res. Lett*., 31, L15615, doi:10.1029/2004GL020373.

West, M., Ni, J., Baldridge, W.S., Wilson, D., Aster, R., Gao, W. & Grand, S., 2004. Crust and upper mantle shear wave structure of the southwest United States: Implications for rifting and support for high elevation, *Journal of Geophysical Research-Solid Earth*, 109, 16.

Whitmeyer, S. J., and K. E. Karlstrom (2007), Tectonic model for the Proterozoic growth of North America, *Geosphere*, 3(4), 220–259, doi:10.1130/GES00055.1.

Wilson, C. K., Jones, C. H., and Gilbert, H. J., 2003. Single-chamber silicic magma system inferred from shear wave discontinuities of the crust and uppermost mantle, Coso geothermal area, California, *J. Geophys. Res.*, 108(B5), 2226.

Wilson, D., R. Aster, M. West, J. Ni, S. Grand, W. Gao, W. S. Baldridge, S. Semken, and P. Patel (2005) Lithospheric structure of the Rio Grande rift, *Nature, 433*, 851-854.

Wright, H. E. (1962), Role of the Wadena Lobe in the Wisconsin Glaciation of Minnesota,

Geological Society of America Bulletin, 73(1), 73–100, doi:10.1130/0016-7606.

Woelk, T. S., and W. J. Hinze (1991), Model of the midcontinent rift system in northeastern Kansas, *Geology*, 19(3), 277–280, doi:10.1130/0091-7613.

Woollard, G. P. (1959). Crustal structure from gravity and seismic measurements. *Journal of Geophysical Research*, 64(10), 1521-1544.

Woollard, G. P., and H. R. Joesting (1964), Bouguer gravity anomaly map of the United States, map, 1:2,500,000, U.S. Geol. Surv., Washington, D.C..

Xie, J., M.H. Ritzwoller, W. Shen, Y. Yang, Y. Zheng, and L. Zhou, Crustal radial anisotropy across eastern Tibet and the western Yangtze craton, *J. Geophys. Res*., 118, 4226-4252, doi:10.1002/jgrb.50296, 2013.

Yang, Y. & Forsyth (2006), D. Regional tomographic inversion of the amplitude and phase of Rayleigh waves with 2-D sensitivity kernels, *Geophys. J. Int*., 166, 1148–1160.

Yang, Y.J., Ritzwoller, M.H., Levshin, A.L. & Shapiro, N.M., 2007. Ambient noise rayleigh wave tomography across Europe, *Geophysical Journal International*, 168, 259-274.

Yang, Y., and M. H. Ritzwoller (2008a), Teleseismic surface wave tomography in the western U.S. using the Transportable Array component of USArray, *Geophys. Res. Lett*., 35, L04308, doi:10.1029/2007GL032278.

Yang, Y. and M.H. Ritzwoller (2008b), The characteristics of ambient seismic noise as a source for surface wave tomography, *Geochem., Geophys., Geosys.*, 9(2), Q02008, 18 pages, doi:10.1029/2007GC001814.

Yang, Y., A. Li, and M.H. Ritzwoller (2008a), Crustal and uppermost mantle structure in southern Africa revealed from ambient noise and teleseismic tomography, *Geophys. J. Int*., doi:10.1111/j.1365-246X.2008.03779.x.

Yang, Y., M. H. Ritzwoller, F.-C. Lin, M. P. Moschetti, and N. M. Shapiro (2008b), Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography, *J. Geophys. Res*., 113, B12310, doi:10.1029/2008JB005833.

Yang, Y., M. H. Ritzwoller, and C. H. Jones (2011), Crustal structure determined from ambient noise tomography near the magmatic centers of the Coso region, southeastern California, *Geochem. Geophys. Geosyst*., 12, Q02009, doi:10.1029/2010GC003362.

Yang, Y., M. H. Ritzwoller, Y. Zheng, W. Shen, A. L. Levshin, and Z. Xie (2012), A synoptic view of the distribution and connectivity of the mid-crustal low velocity zone beneath Tibet, *J. Geophys. Res.*, 117, B04303, doi:10.1029/2011JB008810.

Yang,Y., W. Shen and M.H. Ritzwoller, Surface wave tomography in a large-scale seismic array combining ambient noise and teleseismic earthquake data, *Earthquake Science*, 24, 55-64, 2011.

Yao, H.J., van der Hilst, R.D. & de Hoop, M.V., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps, *Geophysical Journal International*, 166, 732-744.

Ye, T., W. Shen, and M.H. Ritzwoller, Crustal and uppermost mantle shear velocity structure adjacent to the Juan de Fuca Ridge from ambient seismic noise, *Geochem. Geophys. Geosyst*., 14(8), 3221-3233, doi:10.1002/ggge.20206, 2013.

Yoo, H.J., Herrmann, R.B., Cho, K.H. & Lee, K., 2007. Imaging the three-dimensional crust of the Korean Peninsula by joint inversion of surface-wave dispersion and teleseismic receiver functions, *Bulletin of the Seismological Society of America*, 97, 1002-1011.

Yoshizawa, K. and B.L.N. Kennett (2002), Non-linear waveform inversion for surface waves with a neighbourhood algorithm – application to multimode dispersion measurements, *Geophys. J. Int.* 149, 118-133.

Zandt, G., Gilbert, H., Owens, T.J., Ducea, M., Saleeby, J. & Jones, C.H., 2004. Active foundering of a continental arc root beneath the southern Sierra Nevada in California, *Nature*, 431, 41-46.

Zartman, R. E., P. D. Kempton, J. B. Paces, H. Downes, I. S. Williams, G. Dobosi, and K. Futa (2013), Lower-Crustal Xenoliths from Jurassic Kimberlite Diatremes, Upper Michigan (USA): Evidence for Proterozoic Orogenesis and Plume Magmatism in the Lower Crust of the Southern Superior Province, *J. Petrology*, 54(3), 575-608, doi: 10.1093/petrology/egs079.

Zhao, L.S., Sen, M.K., Stoffa, P. & Frohlich, C., 1996. Application of very fast simulated annealing to the determination of the crustal structure beneath Tibet, *Geophysical Journal International*, 125, 355-370.

Zheng, X. F., Z. X. Yao, J. H. Liang, and J. Zheng (2010), The role played and opportunities provided by IGP DMC of China National Seismic Network in Wenchuan earthquake disaster relief and researches, *Bull. Seismol. Soc. Am.*, 100(5B), 2866-2872.

Zheng, Y., W. Shen, L. Zhou, Y. Yang, Z. Xie, and M. H. Ritzwoller (2011), Crust and uppermost mantle beneath the North China Craton, northeastern China, and the Sea of Japan from ambient noise tomography, *J. Geophys. Res.*, 116, B12312, doi:10.1029/2011JB008637.

Zhou, L., J. Xie, W. Shen, Y. Zheng, Y. Yang. H. Shi, and M.H. Ritzwoller (2012), The structure of the crust and uppermost mantle beneath South China from ambient noise and earthquake tomography, *Geophys. J. Int.*, doi: 10.1111/j.1365-246X.2012.05423.x.

Zhu, L., and H. Kanamori (2000), Moho depth variation in southern California from teleseismic receiver functions, *J. Geophys. Res*., 105(B2), 2969–2980, doi:10.1029/1999JB900322.

Ziegler, P., and Cloetingh, S., 2004, Dynamic processes controlling evolution of rift basins, *Earth Science Reviews*, v. 64, p. 1–50, 2004.